Tensor-based Identification of the Structure of Block-Oriented Nonlinear Systems

Alain Kibangou, Gérard Favier

Tensor-based Identification of the Structure of Block-Oriented Nonlinear Systems

Alain Y. Kibangou * **Gérard Favier** **

* *GIPSA-Lab, System Control Department, CNRS, 961 rue de la Houille Blanche, B.P. 46- 38402 Saint Martin d'Hères, France (e-mail: Alain.Kibangou@gipsa-lab.inpg.fr)*
** *Laboratoire I3S, University of Nice Sophia Antipolis, CNRS, Les Algorithmes- Bât. Euclide B, 2000 Route des lucioles, B.P. 121 - 06903 Sophia Antipolis Cedex, France, (Phone:+33 4 92 94 27 36, Fax: +33 4 92 94 28 96, e-mail: favier@i3s.unice.fr)*

Abstract: In this paper, we propose a new tensor-based approach to identify the structure of a block-oriented nonlinear system (Hammerstein, Wiener, and Wiener-Hammerstein systems). The proposed method makes use of one time-domain Volterra kernel of an arbitrary order higher than two, which can be viewed as a tensor. We develop a tensor analysis for carrying out the identification of the structure of block-oriented nonlinear systems. The performance of the proposed identification scheme is illustrated by means of some simulation results.

Keywords: Nonlinear system identification; Hammerstein systems; Wiener systems; Wiener-Hammerstein systems; Volterra kernels; PARAFAC decomposition.

1. INTRODUCTION

The class of nonlinear dynamic systems which can be represented by a cascade of linear dynamic and nonlinear static subsystems, also called block-oriented models, plays an important role in many fields of application. Three kinds of block-oriented models are generally encountered: the Wiener-Hammerstein (also called Sandwich or LNL model, formed by a nonlinear static subsystem in sandwich between two linear dynamic subsystems), Wiener (a linear dynamic subsystem followed by a nonlinear static one), and Hammerstein (a nonlinear static subsystem followed by a linear dynamic one) models. They have been used to model fluid-structure dynamic interactions, Bendat (1990), human ankle stiffness dynamics, Kearney et al. (1997), biological systems, Korenberg and Hunter (1986), digital communication systems, Feher (1993), and chemical systems, Kalafatis et al. (1997), among other applications.

Nonlinear system identification using block-oriented models has been studied for many years (see Giannakis and Serpedin (2001) for an extensive bibliography). Existing methods are in general structure dependent. Prior knowledge about the structure of the nonlinear system to be identified is required before selecting the adequate parameter estimation method.

Structure-independent identification procedures valid for quite different models can be obtained by considering the equivalent representations in Volterra series, Haber (1989). It is indeed well known that any mild nonlinear system with fading memory can be approximated with a given degree of accuracy by a Volterra model, Boyd and Chua (1985). Unlike the block-oriented models, the Volterra model is linear in its parameters.

As pointed out in Haber (1989) and Chen (1995), the analysis of the estimated Volterra kernels can reveal the underlying structure of the unknown plant. For example, if the second-order Volterra kernel is diagonal, then the model is of Hammerstein type, Marmarelis and Naka (1974); if it gives rise to a rank one matrix then the model is of Wiener type, Haber (1989), else a Wiener-Hammerstein structure can be adopted. Once the structure is determined, one can make use of one of the structure-dependent estimation algorithms developed in the literature. For a Wiener-Hammerstein structure, many of the proposed algorithms begin by estimating the frequency response of the linear subsystems from the first and second order frequency-domain Volterra kernels (see Korenberg and Hunter (1986); Weiss et al. (1998); Vandersteen and Schoukens (1999); Hui Tan and Godfrey (2002)). In addition, they generally make use of Gaussian or multisine input signals. As it is well known, the input signal design is a very important step for system identification, since a badly designed input can introduce unwanted distortions and/or an excitation default in the considered frequency range.

In this paper, by considering Volterra kernels of order higher than two as tensors, or multi-way arrays, we show that the analysis of such tensors can reveal the structure of the system to be identified. The proposed scheme is more general than that requiring first or second order Volterra kernels in time or frequency domain.

The rest of this paper is organized as follows. Section 2 briefly recalls a very useful tensor decomposition called PARAFAC, Harshman (1970); Caroll and Chang (1970). In Section 3, we derive a tensor analysis based on a pth-order time-domain Volterra kernel that allows determining the structure of the nonlinear plant to be identified. The

proposed identification method is illustrated by means of some simulation results in Section 4, before concluding the paper in Section 5.

Notations: Vectors are written as boldface lower-case letters ($\mathbf{a}$,$\mathbf{b}$,$\cdots$), matrices as boldface capitals ($\mathbf{A}$,$\mathbf{B}$,$\cdots$) and tensors as blackboard letters ($\mathbb{A}$, $\mathbb{B}$,$\cdots$). $\mathbf{A}^T$ stands for the transpose of the matrix $\mathbf{A}$. $diag(.)$ is the operator that forms a diagonal matrix from its vector argument. The symbol $\circ$ denotes the vector outer product defined as follows:

$$\mathbf{u} \in \Re^I,\ \mathbf{v} \in \Re^J \rightarrow \mathbf{u} \circ \mathbf{v} \in \Re^{I \times J} \leftrightarrow (\mathbf{u} \circ \mathbf{v})_{ij} = u_i v_j.$$

The Kronecker product is denoted by $\otimes$, whereas $\odot$ denotes the Khatri-Rao product defined for two arbitrary matrices $\mathbf{A}$ and $\mathbf{B}$, having the same number of columns N, as $\mathbf{A} \odot \mathbf{B} = (\ \mathbf{A}_{.1} \otimes \mathbf{B}_{.1} \ \cdots \ \mathbf{A}_{.N} \otimes \mathbf{B}_{.N}\ )$. $\odot^p \mathbf{A}$ stands for the Khatri-Rao product of $\mathbf{A}$ p times by itself. We have the following property:

$$(\mathbf{A} diag\,(\mathbf{c})) \odot \mathbf{B} = \mathbf{A} \odot (\mathbf{B} diag\,(\mathbf{c})) = (\mathbf{A} \odot \mathbf{B})\, diag\,(\mathbf{c}) \tag{1}$$

We also make use of the following property: for matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, with respective dimensions $M \times N$, $N \times P$, and $Q \times M$, if $\mathbf{C}$ is full column rank and $\mathbf{B}$ is full row rank then:

$$rank(\mathbf{AB}) = rank(\mathbf{CA}) = rank(\mathbf{A}) \tag{2}$$

For a vector $\mathbf{a} = (a_1 \ \cdots \ a_M)^T$, with $a_1 \neq 0$ and $a_M \neq 0$, we denote by $\mathcal{T}_{M+N-1,N}(\mathbf{a})$ the $(M+N-1) \times N$ Toeplitz matrix of rank N, given by:

$$\mathcal{T}_{M+N-1,N}(\mathbf{a}) = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ a_2 & a_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & & a_1 \\ \vdots & \vdots & & \vdots \\ a_M & a_{M-1} & \cdots & a_{M-N+1} \\ 0 & a_M & \ddots & \vdots \\ \vdots & & \ddots & a_{M-1} \\ 0 & \cdots & \cdots & a_M \end{pmatrix}.$$

Assuming $x(j) = 0\ \forall j < 0$ and $\forall j > N-1$, $x(0) \neq 0$ and $x(N-1) \neq 0$, the set of equations $y(j) = \sum_{i=1}^{M} a_i x(j-i)$, $j = 1, \cdots, M+N-1$, can be written as:

$$\mathbf{y} = \mathcal{T}_{M+N-1,N}(\mathbf{a})\mathbf{x} = \mathcal{T}_{M+N-1,M}(\mathbf{x})\mathbf{a}, \tag{3}$$

where

$$\mathbf{y} = (y(1) \ \cdots \ y(M+N-1))^T$$

and

$$\mathbf{x} = (x(0) \ \cdots \ x(N-1))^T$$

with $rank(\mathcal{T}_{M+N-1,N}(\mathbf{a})) = N$, $rank(\mathcal{T}_{M+N-1,M}(\mathbf{x})) = M$.

2. THE PARAFAC DECOMPOSITION

Let $\mathbb{X}$ be a tensor of order N, also called a N-way array, with dimensions $I_1 \times I_2 \times \cdots \times I_N$ and entries $x_{i_1,\cdots,i_N}$. This tensor is diagonal if $x_{i_1,\cdots,i_N} = 0$ unless $i_1 = i_2 = \cdots = i_N$. We shall say that a N-way array is cubical if its N dimensions are identical ($I_n = I$, $n = 1, \cdots, N$) Kolda and Bader (2009). A cubical array is called supersymmetric if its entries do not change under any permutation of its N indices, i.e. $x_{i_1,\cdots,i_N} = x_{\pi(i_1),\cdots,\pi(i_N)}$, where $\pi(.)$ is any permutation of the indices $i_1, \cdots, i_N$ (See Comon (2000)).

Any real valued tensor $\mathbb{X}$, of order N, can always be decomposed as:

$$\mathbb{X} = \sum_{r=1}^{R} \mathbf{A}_{.r}^{(1)} \circ \cdots \circ \mathbf{A}_{.r}^{(N)}, \tag{4}$$

where $\mathbf{A}_{.r}^{(n)}$ is the r-th column of the matrix factor $\mathbf{A}^{(n)} \in \Re^{I_n \times R}$, $n = 1, \cdots, N$. The tensor rank is defined as the smallest integer R such that this decomposition holds exactly. This decomposition, called PARAFAC (PARAllel FACtor analysis), Harshman (1970), or CANDECOMP (CANonical DECOMPosition), Caroll and Chang (1970), admits the following scalar formulation:

$$x_{i_1,\cdots,i_N} = \sum_{r=1}^{R} \prod_{n=1}^{N} a_{i_n,r}^{(n)}, \tag{5}$$

$a_{i_n,r}^{(n)}$ being the entries of the matrix factor $\mathbf{A}^{(n)}$.

It is convenient to represent tensors by means of matrices. Writing a tensor into a matrix form is called matricization or unfolding. For a tensor $\mathbb{X}$ of an arbitrary order N, we have N unfolded matrix representations $\mathbf{X}_n \in \Re^{I_{n+1}\cdots I_N I_1 \cdots I_{n-1} \times I_n}$, corresponding to an unfolding of $\mathbb{X}$ along each mode $n = 1, \cdots, N$. There are different ways to define the unfolded matrices Kolda and Bader (2009). In this paper, the tensor element $x_{i_1,\cdots,i_N}$ is located in the unfolded matrix $\mathbf{X}_n$ with the column number i_n and the row number j given by:

$$j = 1 + \sum_{k=1,\ k \neq n}^{N} (i_k - 1) J_k, \tag{6}$$

where

$$J_k = \begin{cases} \prod_{m=k+1}^{n-1} I_m, & \text{for } k = 1, \cdots, n-2 \\ 1 & \text{for } k = n-1 \\ J_N \prod_{m=k+1}^{N} I_m, & \text{for } k = n+1, \cdots, N-1 \\ \prod_{m=1}^{n-1} I_m, & \text{for } k = N \end{cases} \tag{7}$$

For a third-order tensor $\mathbb{X}$, there are three unfolded matrices: $\mathbf{X}_1 \in \Re^{I_2 I_3 \times I_1}$, $\mathbf{X}_2 \in \Re^{I_3 I_1 \times I_2}$, and $\mathbf{X}_3 \in \Re^{I_1 I_2 \times I_3}$ defined as:

$$\mathbf{X}_1 = \begin{pmatrix} \mathbf{X}_{.1.} \\ \vdots \\ \mathbf{X}_{.I_2.} \end{pmatrix}, \quad \mathbf{X}_2 = \begin{pmatrix} \mathbf{X}_{..1} \\ \vdots \\ \mathbf{X}_{..I_3} \end{pmatrix}, \quad \mathbf{X}_3 = \begin{pmatrix} \mathbf{X}_{1..} \\ \vdots \\ \mathbf{X}_{I_1..} \end{pmatrix} \tag{8}$$

where

$$\mathbf{X}_{i_1..} = \begin{pmatrix} x_{i_1,1,1} & \cdots & x_{i_1,1,I_3} \\ \vdots & \ddots & \vdots \\ x_{i_1,I_2,1} & \cdots & x_{i_1,I_2,I_3} \end{pmatrix} \in \Re^{I_2 \times I_3}$$

$$\mathbf{X}_{.i_2.} = \begin{pmatrix} x_{1,i_2,1} & \cdots & x_{I_1,i_2,1} \\ \vdots & \ddots & \vdots \\ x_{1,i_2,I_3} & \cdots & x_{I_1,i_2,I_3} \end{pmatrix} \in \Re^{I_3 \times I_1}$$

$$\mathbf{X}_{..i_3} = \begin{pmatrix} x_{1,1,i_3} & \cdots & x_{1,I_2,i_3} \\ \vdots & \ddots & \vdots \\ x_{I_1,1,i_3} & \cdots & x_{I_1,I_2,i_3} \end{pmatrix} \in \Re^{I_1 \times I_2}.$$

In general, slice-wise expressions do not easily extend beyond three dimensions.

For an arbitrary order N, we have N unfolded matrices $\mathbf{X}_n$, $n = 1, \cdots, N$, linked to the PARAFAC matrix factors as follows:

$$\mathbf{X}_n = \left(\mathbf{A}^{(n+1)} \odot \mathbf{A}^{(n+2)} \odot \cdots \odot \mathbf{A}^{(n+N-1)}\right) \mathbf{A}^{(n)T}, \quad (9)$$

with the convention

$$\mathbf{A}^{(k+N)} = \mathbf{A}^{(k)}, \qquad k < N.$$

For a third-order tensor, we get:

$$\mathbf{X}_1 = \left(\mathbf{A}^{(2)} \odot \mathbf{A}^{(3)}\right) \mathbf{A}^{(1)T},$$
$$\mathbf{X}_2 = \left(\mathbf{A}^{(3)} \odot \mathbf{A}^{(1)}\right) \mathbf{A}^{(2)T},$$
$$\mathbf{X}_3 = \left(\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right) \mathbf{A}^{(3)T}.$$

while for $N = 5$ and $n = 3$ we get:

$$\mathbf{X}_3 = \left(\mathbf{A}^{(4)} \odot \mathbf{A}^{(5)} \odot \mathbf{A}^{(1)} \odot \mathbf{A}^{(2)}\right) \mathbf{A}^{(3)T}.$$

In some particular cases, the rank of the unfolded matrices can reveal that of the tensor, as shown in the following Lemma.

Lemma 1. Let $\mathbb{X}$ be a $I_1 \times I_2 \times \cdots \times I_N$ tensor of rank R and $\mathbf{A}^{(n)} \in \Re^{I_n \times R}$, $n = 1, \cdots, N$, the matrix factors associated with its PARAFAC decomposition. If these matrices are full column rank, i.e. $rank(\mathbf{A}^{(n)}) = R$, $n = 1, \cdots, N$, then the unfolded matrices $\mathbf{X}_n$, $n = 1, \cdots, N$, are also full column rank:

$$rank(\mathbf{X}_n) = R = rank(\mathbb{X}), \quad n = 1, \cdots, N. \quad (10)$$

Proof: Since $\mathbf{A}^{(n)} \in \Re^{I_n \times R}$, $n = 1, \cdots, N$, are full column rank matrices, then from property (2) we have:

$$rank(\mathbf{X}_n) = rank(\mathbf{A}^{(n+1)} \odot \mathbf{A}^{(n+2)} \odot \cdots \odot \mathbf{A}^{(n+N-1)}).$$

Applying the Lemma 1 in Sidiropoulos et al. (2000), we can deduce that a matrix equal to Khatri-Rao products of full column rank matrices is also full column rank, which implies (10).■

3. TENSOR-BASED METHOD FOR IDENTIFYING THE STRUCTURE OF A BLOCK-ORIENTED NONLINEAR SYSTEM

Let us consider the block-oriented nonlinear model depicted in Fig. 1. We denote by $u(n)$, $y(n)$, $v_1(n)$, and $v_2(n)$, the input, the output, and intermediate signals respectively. Assuming that the nonlinearity is continuous within the considered dynamic range, then, from the Weierstrass theorem, it can be approximated to an arbitrary degree of accuracy by a polynomial $C(.)$ of finite degree P, with coefficients c_p. So, the model is constituted with a polynomial $C(.)$, sandwiched between two linear filters with impulse responses $l(.)$ and $g(.)$, and memories M_l and M_g respectively, i.e.:

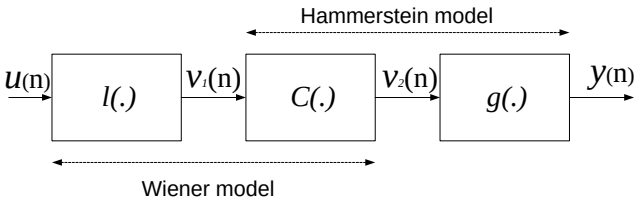

Fig. 1. Block-oriented nonlinear model

$$v_1(n) = \sum_{i=0}^{M_l-1} l(i)u(n-i), \;\; v_2(n) = \sum_{p=1}^{P} c_p v_1^p(n),$$
$$y(n) = \sum_{i=0}^{M_g-1} g(i)v_2(n-i),$$

with $l(0) = g(0) = 1$, a standard constraint for guaranteing the model uniqueness, and $l(M_l-1) \neq 0$ and $g(M_g-1) \neq 0$, where $M_l \geq 1$ and $M_g \geq 1$, $M_l = 1$ corresponding to a Hammerstein model whereas $M_g = 1$ corresponds to a Wiener model.

The estimation of the parameters that characterize each block of this model is not simple since this model is not linear in its parameters. So, the equivalent Volterra model, linear in its parameters, is easier to identify. The parameters of both linear and nonlinear subsystems can then be deduced from estimated Volterra kernels. In the literature, several methods based on this approach were proposed by considering first and second-order Volterra kernels in time, Haber (1989), or frequency ,Weiss et al. (1998); Vandersteen and Schoukens (1999); Hui Tan and Godfrey (2002), domain. In the sequel, we do not study the parameter estimation problem but only that of determining the model structure.

The output of the Volterra model associated with the above block-oriented nonlinear system can be written as

$$y(n) = \sum_{p=1}^{P} \sum_{i_1,\cdots,i_p=1}^{M_v} h_p(i_1-1, \cdots, i_p-1) \prod_{j=1}^{p} u(n-i_j+1) \quad (11)$$

where $h_p(.)$ denotes the pth-order Volterra kernel given by (see Kibangou and Favier (2006)):

$$h_p(i_1-1, \cdots, i_p-1) = c_p \sum_{i=1}^{M_g} g(i-1) \prod_{k=1}^{p} l(i_k-i), \quad (12)$$

$i_k = 1, \cdots, M_v$, with $M_v = M_l + M_g - 1$ the memory of the nonlinear plant that is assumed to be known. Obviously, this kernel is nonzero if $c_p \neq 0$.

The block oriented nonlinear model depicted in Fig. 1 is of Wiener-Hammerstein type. Since Hammerstein and Wiener models are subclasses of the Wiener-Hammerstein one, from (12) we can deduce their corresponding Volterra

kernels. Indeed, we get a Wiener model by setting $g(i-1)=\delta_{i,1}$, with δ the Kronecker symbol, whereas a Hammerstein model is obtained by setting $l(i-1)=\delta_{i,1}$. So we get respectively:

$$h_p(i_1-1,\cdots,i_p-1)=c_p\prod_{k=1}^{p}l(i_k-1), \tag{13}$$

$$h_p(i_1-1,\cdots,i_p-1)=\begin{cases} c_p g(i-1) & \text{if } i_1=\cdots=i_p=i \\ 0 & \text{elsewhere}\end{cases} \tag{14}$$

The first and second order Volterra kernels can respectively be viewed as a vector and a matrix whereas higher order kernels are supersymmetric tensors since their entries remain unchanged under any permutation of their indices.

In the sequel, we assume that the pth-order Volterra kernel is known. Note that the algorithms using second order frequency-domain Volterra kernels such as those in Weiss et al. (1998) and Hui Tan and Godfrey (2002) assume the existence of a quadratic nonlinearity. Herein, the considered nonlinearity is of arbitrary degree p. In other words, our approach is able to deal with more general systems including those for which there is no quadratic term, as it is the case for nonlinear systems containing only odd-order polynomial terms.

3.1 Volterra kernel based tensor construction

Now, we build a tensor whose analysis allows identifying the structure of a block-oriented non-linear system.

Let $f(i)$, $i=1,\cdots,M_f$, with $M_f\geq M_v$, be a set of nonzero coefficients randomly chosen. We denote by $\mathbb{F}_p$ an $M\times M\times\cdots\times M$ pth-order tensor with entries $f_{i_1,\cdots,i_p}$, $i_k=1,\cdots,M$, $k=1,\cdots,p$, obtained as follows:

$$f_{i_1,\cdots,i_p}=\sum_{i=1}^{M_f}f(i)h_p(i_1-i,\cdots,i_p-i). \tag{15}$$

Taking the memories of the kernel $h_p(.)$ and the filter $f(.)$ into account, we have :

$$\begin{aligned}M&=M_f+M_v-1=M_f+M_l+M_g-2\\&=M_l+\mu-1,\end{aligned} \tag{16}$$

where

$$\mu=M_f+M_g-1. \tag{17}$$

By using (12), equation (15) can be rewritten as:

$$\begin{aligned}f_{i_1,\cdots,i_p}&=c_p\sum_{i=1}^{M_f}\sum_{j=1}^{M_g}f(i)g(j-1)\prod_{k=1}^{p}l(i_k-i+1-j)\\&=c_p\sum_{i=1}^{M_f}\sum_{j=1}^{i+M_g-1}f(i)g(j-i)\prod_{k=1}^{p}l(i_k-j)\\&=\sum_{i=1}^{\mu}\bar{f}_p(i)\prod_{k=1}^{p}l(i_k-i),\end{aligned} \tag{18}$$

with

$$\bar{f}_p(i)=c_p\sum_{j=1}^{M_f}f(j)g(i-j). \tag{19}$$

One can note that a simple analysis of the nonzero tensor $\mathbb{F}_p$ reveals the structure of the block-oriented nonlinear model. Indeed, according to (14) and (15), if the tensor $\mathbb{F}_p$ is diagonal then the structure is of Hammerstein type. Wiener and WH structures correspond to non-diagonal tensors $\mathbb{F}_p$. In the sequel we show that the evaluation of the rank of the tensor $\mathbb{F}_p$ allows distinguishing the three structures.

Defining $a^{(k)}_{i_k,i}=l(i_k-i)$, $k=1,\cdots,p-1$, and $a^{(p)}_{i_p,i}=\bar{f}_p(i)l(i_p-i)$, we can rewrite (18) as:

$$f_{i_1,\cdots,i_p}=\sum_{i=1}^{\mu}\prod_{k=1}^{p}a^{(k)}_{i_k,i}. \tag{20}$$

Comparing (5) with (20), we conclude that (20) represents the PARAFAC decomposition of the tensor $\mathbb{F}_p$. The number of factors involved in this decomposition is equal to μ. It is obvious that $rank(\mathbb{F}_p)\leq\mu$ and equality occurs when all the p matrix factors are full column rank. These p matrix factors, of dimensions $M\times\mu$, are given by:

$$\mathbf{A}^{(n)}=\begin{cases}\mathbf{L}=\mathcal{T}_{M,\mu}(\mathbf{l}) & \text{for } n=1,\cdots,p-1,\\ \mathbf{L}diag(\bar{\mathbf{f}}_p) & \text{for } n=p\end{cases} \tag{21}$$

with

$$\mathbf{l}=\left(l(0)\ \cdots\ l(M_l-1)\right)^T,$$
$$\bar{\mathbf{f}}_p=\left(\bar{f}_p(1)\ \cdots\ \bar{f}_p(\mu)\right)^T$$

Using equation (3), the vector $\bar{\mathbf{f}}_p$, defined in (19), can be written as :

$$\bar{\mathbf{f}}_p=\mathcal{T}_{\mu,M_g}(\mathbf{f})\bar{\mathbf{g}}_p,$$

with

$$\mathbf{f}=\left(f(1)\ \cdots\ f(M_f)\right)^T,$$
$$\bar{\mathbf{g}}_p=c_p\mathbf{g}=c_p\left(g(0)\ \cdots\ g(M_g-1)\right)^T.$$

As $l(0)\neq0$ and $l(M_l-1)\neq0$, $\mathbf{L}$ is full column rank, implying that $\mathbf{A}^{(n)}$, $n=1,\cdots,p-1$, are full column rank. We note that the matrix factor $\mathbf{A}^{(p)}$ is also full column rank if the entries of $\bar{\mathbf{f}}_p$ are all nonzero, i.e. $\bar{f}_p(i)\neq0$, $i=1,\cdots,\mu$.

Taking equation (19) into account, with nonzero randomly chosen coefficients $f(j)$, $j=1,\cdots,M_f$, we conjecture that the coefficients $\bar{f}_p(i)$ are all nonzero, which implies that all the matrix factors $\mathbf{A}^{(n)}$, $n=1,\cdots,p$, of $\mathbb{F}_p$ are full column rank. Applying Lemma 1, we deduce that: $rank(\mathbb{F}_p)=\mu$. Therefore, from (16) and (17), we can conclude that

$$rank(\mathbb{F}_p)=M-M_l+1, \tag{22}$$
$$=M_g+M_f-1. \tag{23}$$

As we know that:

$$M_l=1\Rightarrow\text{Hammerstein structure}$$
$$M_g=1\Rightarrow\text{Wiener structure}$$
$$M_l>1\ \text{and}\ M_g>1\Rightarrow\text{WH structure}$$

the rank of $\mathbb{F}_p$ can be used to determine the structure of the block-oriented nonlinear system by means of the following rules deduced from (22) and (23).

$$rank(\mathbb{F}_p) = M \Rightarrow \text{Hammerstein structure}$$
$$rank(\mathbb{F}_p) = M_f \Rightarrow \text{Wiener structure}$$
$$M_f < rank(\mathbb{F}_p) < M \Rightarrow \text{WH structure}$$

3.2 Structure identification algorithm

Unlike matrices, computing the rank of a tensor is not a trivial task. However, as shown in Lemma 1, when all the matrix factors are full column rank, the tensor rank can be determined as the rank of one of its unfolded matrix representations.

Let us consider the unfolded matrix of $\mathbb{F}_p$ along the pth dimension:

$$\begin{aligned}\mathbf{F}_p &= (\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)} \odot \cdots \odot \mathbf{A}^{(p-1)})\mathbf{A}^{(p)T} \\ &= \left(\odot^{(p-2)}\mathbf{L}\right) diag(\bar{\mathbf{f}}_p)\mathbf{L}^T.\end{aligned}$$

We have:

$$rank(\mathbf{F}_p) = \mu = rank(\mathbb{F}_p). \quad (24)$$

Using this result, we derive the following tensor-based algorithm for identifying the structure of block-oriented nonlinear systems.

Algorithm: TENsor-based Structure IDentification algorithm (TENSID):

Given the estimate of the pth-order kernel $h_p(.)$, with $p \geq 3$, of the Volterra model associated with the nonlinear system to be identified, and its memory M_v.

(1) Generate the random coefficients $f(i)$, $i = 1, \cdots, M_f$, of a FIR filter, with $M_f \geq M_v$.

(2) Generate the elements $f_{i_1,\ldots,i_p}$ of the pth-order tensor $\mathbb{F}_p \in \Re^{M \times M \times \cdots \times M}$ by convoluting the kernel $h_p(., \cdots, .)$ with the impulse response of the FIR filter according to (15).

(3) Construct the unfolded matrix $\mathbf{F}_p$ of $\mathbb{F}_p$, with dimensions $M^{p-1} \times M$ and entries $f_{j,i_p} = f_{i_1,\cdots,i_{p-1},i_p}$, with j given by (6) and (7).

(4) Determine the rank of the tensor $\mathbb{F}_p$ as $rank(\mathbb{F}_p) = rank(\mathbf{F}_p)$.

(5) Deduce the structure of the system:
- $rank(\mathbb{F}_p) = M \Rightarrow$ Hammerstein structure
- $rank(\mathbb{F}_p) = M_f \Rightarrow$ Wiener structure
- $M_f < rank(\mathbb{F}_p) < M \Rightarrow$ WH structure

and then the values of M_g and M_l using (23) and (22) respectively.

In practice, the estimated Volterra kernel is corrupted by estimation errors. Such errors can impact the rank evaluation. So, we propose the following rule for determining the numerical rank of the tensor, that consists in discarding the less significant singular values of $\mathbf{F}_p$. Denoting by σ_i, $i = 1, \cdots, M$, the computed singular values of $\mathbf{F}_p$, ranged in a non-increasing order, the rank r is chosen as the smallest integer k such that:

$$\sum_{i=1}^{k-1} \sigma_i < \epsilon \sum_{i=1}^{M} \sigma_i \leq \sum_{i=1}^{k} \sigma_i, \quad (25)$$

where ϵ is a constant close to 1.

4. SIMULATION RESULTS

In this section, we evaluate the TENSID algorithm by considering third-order block-oriented nonlinear systems, with memory $M_v = 5$, drawn from a Gaussian distribution. For systems with a Wiener-Hammerstein structure we have $M_l = M_g = 3$. In order to carry out such an evaluation independently from the Volterra kernel estimation method, we build the exact associated third-order Volterra kernel and then we add a Gaussian noise to simulate the estimation errors. We consider different levels of estimation error evaluated by means of the NMSE (Normalized mean square error):

$$NMSE = 10 \log_{10} \frac{\|h_3(.,.,.)\|_F^2}{\|e_3(.,.,.)\|_F^2},$$

where $\|.\|_F$ denotes the Frobenius norm, $e_3(.)$ being the additive tensor of estimation errors.

The M_f coefficients of the FIR filter were generated with a Gaussian distribution, and $M_f = M_v$.

Tables 1, 2 and 3 present the probability of good structure identification obtained in averaging the results over 100 independent Monte Carlo runs, and using the criterion (25), for each class of block-oriented systems and for different values of NMSE and ϵ.

First of all, without using the numerical rank given by the criterion (25), the TENSID algorithm always gives the actual structure in the noiseless case. When an estimation error is added to the exact Volterra kernel, the structure provided by the TENSID algorithm is almost always the Hammerstein structure, which corresponds to a full rank $\mathbf{F}_p$.

By using the criterion (25), even in very noisy cases, the TENSID algorithm can provide the actual structure (see Tables 1, 2, and 3).

Obviously, the probability of good structure identification increases with the NMSE for the three considered values of ϵ. For NMSE values lower than 40 dB, decreasing ϵ allows increasing the total probability of good identification. On the contrary, for NMSE values higher than 40 dB, the total probability of good identification slightly decreases with ϵ. In addition, identification of the Hammerstein structure is more sensitive to the value of ϵ. The best performance of Hammerstein structure identification is obtained for higher values of ϵ, while identification of Wiener structures is degraded when epsilon increases.

Table 1. Structure identification performance for $\epsilon = 0.999$

Structure	NMSE (dB)					
	10	20	30	40	50	60
Hammerstein	98%	97%	97%	95%	87%	88%
Wiener	0%	3%	21%	42%	76%	92%
WH	8%	40%	70%	85%	95%	100%
Total	35%	47%	62%	74%	86%	93%

Table 2. Structure identification performance for $\epsilon = 0.997$

Structure	NMSE (dB)					
	10	20	30	40	50	60
Hammerstein	89%	86%	86%	81%	73%	70%
Wiener	2%	19%	60%	79%	88%	95%
WH	44%	78%	92%	95%	98%	98%
Total	45%	61%	79%	85%	86%	87%

Table 3. Structure identification performance for $\epsilon = 0.995$

Structure	NMSE (dB)					
	10	20	30	40	50	60
Hammerstein	81%	84%	70%	68%	60%	64%
Wiener	9%	31%	60%	78%	96%	95%
WH	67%	77%	88%	92%	98%	96%
Total	52%	64%	72%	79%	84%	85%

5. CONCLUSION

In this paper, we have presented a new scheme for identifying block-oriented nonlinear systems with a nonlinearity degree higher than two. It is based on the PARAFAC decomposition of a pth-order tensor obtained by convoluting a random FIR filter with the pth-order Volterra kernel associated with the block-oriented nonlinear system. The analysis of this tensor reveals the structure of the nonlinear system to be identified. The structure identification is carried out through the evaluation of the tensor rank determined as the numerical rank of a given unfolded matrix representation of the considered tensor. The sensitivity of the proposed criterion is to be studied and the derivation of a more robust criterion is to be investigated. In addition, the estimation of the linear subsystems can be achieved by estimating the PARAFAC decomposition of the involved tensor whose matrix factors are in Toeplitz form. The authors have shown in Kibangou and Favier (2008) that such a decomposition can be estimated in using a closed-form solution. Recalling that the proposed identification method relies on the estimation of only one Volterra kernel of a given order higher than or equal to 3, the possibility for estimating separately the Volterra kernels by means of robust closed form expressions makes the proposed identification method very attractive.

REFERENCES

J.S. Bendat. *Nonlinear system analysis and identification from random data*. John Wiley & Sons, 1990.

S. Boyd and L.O. Chua. Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Trans. on Circuits and Systems*, 32(11): 1150–1161, 1985.

J.D. Caroll and J.J. Chang. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 35:283–319, 1970.

H.-W. Chen. Modeling and identification of parallel nonlinear systems: Structural classification and parameter estimation methods. *Proc. of the IEEE*, 83(1):39–66, January 1995.

P. Comon. Tensor decompositions: state of the art and applications. In *IMA Conf. Mathematics in Signal Processing*, Warwick, U.K, December 2000.

K. Feher. *Digital communications- Satellite/Earth station engineering*. Englewood Cliffs NJ: Prentice Hall, 1993.

G.B. Giannakis and E. Serpedin. A bibliography on nonlinear system identification. *Signal Processing*, 81 (3):533–580, March 2001.

R. Haber. Structural identification of quadratic block-oriented models. *Int. J. Systems Science*, 20(8):1355–1380, 1989.

R.A. Harshman. Foundation of the PARAFAC procedure: models and conditions for an "explanatory" multimodal factor analysis. *UCLA working papers in phonetics*, 16: 1–84, 1970.

A. Hui Tan and K. Godfrey. Identification of Wiener-Hammerstein models using linear interpolation in the frequency domain (LIFRED). *IEEE Trans. on Instr. and Measurement*, 51(3):509–521, June 2002.

A.D. Kalafatis, L. Wang, and W.R. Cluett. Identification of Wiener type nonlinear systems in a noisy environment. *International Journal of Control*, 66(6):923–941, 1997.

R.E. Kearney, R.B. Stein, and L. Parameswaran. Identification of intrinsinc and reflex contributions to human ankle stiffness dynamics. *IEEE Trans. on Biomedical Engineering*, 44(6):493–504, 1997.

A.Y. Kibangou and G. Favier. Wiener-Hammerstein systems modeling using diagonal Volterra kernels coefficients. *IEEE Signal Proc. Letters*, 13(6):381–384, June 2006.

A.Y. Kibangou and G. Favier. A non-iterative solution for PARAFAC decomposition with Toeplitz constraints. (Submitted to Signal Processing), 2008.

T.G. Kolda and B.W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3), September 2009.

M.J. Korenberg and I.W. Hunter. The identification of nonlinear biological systems: LNL cascade models. *Biological Cybernetics*, 55:125–134, 1986.

P.Z. Marmarelis and K.-I. Naka. Identification of multi-input biological systems. *IEEE Trans. on Biomedical Engineering*, 21(2):88–101, March 1974.

N.D. Sidiropoulos, R. Bro, and G.B. Giannakis. Parallel factor analysis in sensor array processing. *IEEE Trans. on Signal Processing*, 48(8):2377–2388, 2000.

G. Vandersteen and J. Schoukens. Measurement and identification of nonlinear systems consisting of linear dynamic blocks and one static nonlinearity. *IEEE Trans. on Automatic Control*, 44(6):1266–1271, 1999.

M. Weiss, C. Evans, and D. Rees. Identification of nonlinear cascade systems using paired multisine signals. *IEEE Trans. on Instrum. and Measurement*, 47(1):332–336, February 1998.