Available online at www.sciencedirect.com

ScienceDirect

automatica

Automatica 43 (2007) 1084–1091

www.elsevier.com/locate/automatica

Brief paper

Choice of free parameters in expansions of discrete-time Volterra models using Kautz functions[☆]

Alex da Rosa[a,*], Ricardo J.G.B. Campello[a,b], Wagner C. Amaral[a]

[a]*DCA/FEEC/UNICAMP, 13083-852, Campinas-SP, Brazil*

[b]*SCC/ICMC/USP, 13560–970, São Carlos-SP, Brazil*

Received 30 March 2006; received in revised form 28 November 2006; accepted 1 December 2006

Abstract

This work tackles the problem of modeling nonlinear systems using Volterra models based on Kautz functions. The drawback of requiring a large number of parameters in the representation of these models can be circumvented by describing every kernel using an orthonormal basis of functions, such as the Kautz basis. The resulting model, so-called Wiener/Volterra model, can be truncated into a few terms if the Kautz functions are properly designed. The underlying problem is how to select the free-design complex poles that fully parameterize these functions. A solution to this problem has been provided in the literature for linear systems, which can be represented by first-order Volterra models. A generalization of such strategy focusing on Volterra models of any order is presented in this paper. This problem is solved by minimizing an upper bound for the error resulting from the truncation of the kernel expansion into a finite number of functions. The aim is to minimize the number of two-parameter Kautz functions associated with a given series truncation error, thus reducing the complexity of the resulting finite-dimensional representation. The main result is the derivation of an analytical solution for a sub-optimal expansion of the Volterra kernels using a set of Kautz functions.

Keywords: Nonlinear systems; Volterra series; Two-parameter Kautz functions; Optimization; System identification

1. Introduction

Models based on orthonormal basis functions have received increasing interest in recent years. This subject has become an important topic of research in studies involving the identification and control of dynamic processes (Bokor & Schipp, 1998; Doyle, Pearson, & Ogunnaike, 2002; Dumont & Fu, 1993; Heuberger, Van den Hof, & Bosgra, 1995; Heuberger, Van den Hof, & Wahlberg, 2005; Oliveira, Amaral, Favier, & Dumont, 2000; Schetzen, 1989). The use of orthonormal series for representing signals and systems was first addressed by Wiener (1958). This approach consists of representing a given signal or system in terms of an orthonormal basis for the space of interest. Many control problems can be formulated as optimizing a certain cost-function over the class of stable systems, and orthonormal exponentials provide good parameterizations for this class of systems (Wahlberg & Mäkilä, 1996).

The main reason for using orthonormal basis functions in modeling for control and signal processing is that the corresponding approximate problem usually has a simplified solution (Heuberger et al., 2005). The orthonormality property of those functions often yield a simpler general model. An important issue regarding the use of a generalized orthonormal basis model structure is the incorporation of approximate knowledge about the system dynamics into the identification process (Heuberger et al., 1995). This way, the number of model parameters to be estimated and, accordingly, the variance of their estimation, can be reduced. The consequence is an improvement to the accuracy of the models.

Models using orthonormal functions may require a reduced number of terms to represent a given system by means of

[☆] This paper was presented at the 16th IFAC World Congress, Prague, Czech Republic, July 2005. This paper was recommended for publication in revised form by Associate Editor Johan Schoukens under the direction of Editor Torsten Söderström.

[*] Corresponding author. Tel.: +55 19 3251 3823.

E-mail addresses: alex@dca.fee.unicamp.br (A. da Rosa), campello@icmc.usp.br (R.J.G.B. Campello), wagner@dca.fee.unicamp.br (W.C. Amaral).

0005-1098/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2006.12.007

a truncated orthonormal series. When properly selected, the orthonormal series can increase the speed of convergence in problems of identification (Van den Hof, Heuberger, & Bokor, 1995; Heuberger et al., 1995). Another advantage is that such functions correspond to all-pass filters, which are robust to implement in numerical computations. Laguerre and Kautz bases (Bokor & Schipp, 1998; Broome, 1965; Wahlberg, 1994; Wahlberg & Mäkilä, 1996) are probably the most used orthonormal basis functions in the approximation, modeling and identification of systems. They are adequate to model systems having kernels with dominant dynamic of first or second order, respectively. To model more complex dynamics, the generalized orthonormal basis functions (GOBFs) (Van den Hof et al., 1995; Ninness & Gustafsson, 1997) are more suitable. The price to pay, however, is a much more complex parameterization of that basis.

A notorious advantage of Laguerre functions is that they have transforms that are rational functions with simple recursive form. Since these functions are completely parameterized by a real-valued pole, they are more suitable for representing well-damped dynamic systems. The Laguerre functions have certain properties that facilitate the problem of optimizing their poles in an analytical strategy, as such functions satisfy a specific second-order difference equation. The optimization of the Laguerre pole was addressed for the linear case in Clowes (1965), Masnadi-Shirazi and Ahmed (1991), Fu and Dumont (1993), Tanguy, Morvan, Vilbé, and Calvez (1995), Silva (1994). More recently, analytical formulae for any-order Volterra models have been presented in Campello, Favier, and Amaral (2004) and Campello, Amaral, and Favier (2006). Poorly damped dynamics, however, are difficult to approximate with a small number of Laguerre functions. Indeed, these functions are not well suited to approximate signals with strong oscillatory behavior (Silva, 1995; Tanguy, Morvan, Vilbé, & Calvez, 2000). This drawback has led to an increasing interest in the two-parameter Kautz functions, introduced in the 1950s by Kautz (1954). Such functions can better approximate systems with oscillatory behavior because they are parameterized by resonant poles. A sub-optimal analytical solution for the choice of the Kautz poles in the representation of discrete linear systems was proposed in Tanguy, Morvan, Vilbé, and Calvez (2002), and the corresponding nonlinear counterpart was addressed in da Rosa (2005), and da Rosa, Amaral, and Campello (2005). The list of works on studies of pole location also includes (Silva, 1997), which presented stationary conditions for optimal linear models based on GOBFs.

In this paper, the results presented in Tanguy et al. (2002) are extended to any-order Volterra models in such a way that an analytical solution for the selection of one of the parameters related to the Kautz pole is obtained. An optimization of the Kautz bases for the orthonormal series expansion of discrete-time Volterra models is given, representing a sub-optimal solution for the choice of the Kautz poles. This solution is based on the minimization of an upper bound of the error resulting from the truncated approximation of Volterra kernels using Kautz functions. The proposed approach requires the system kernels to be known in advance.

The outline of this paper is as follows. In the next section, orthonormal basis functions are presented in the context of Wiener/Volterra models. In Section 3, the optimization problem for the selection of the Kautz pole is discussed. In Section 4, an illustrative example is presented. Finally, Section 5 addresses the conclusions.

2. Volterra and Wiener/Volterra models

Discrete-time Volterra models relate the output $y(k)$ of a physical process to its input $u(k)$ as (Rugh, 1981; Schetzen, 1989)

$$y(k)=\sum_{\eta=1}^{\infty}\sum_{\tau_1=0}^{\infty}\cdots\sum_{\tau_\eta=0}^{\infty}h_\eta(\tau_1,\ldots,\tau_\eta)\prod_{j=1}^{\eta}u(k-\tau_j), \tag{1}$$

where the functions $h_\eta(\tau_1,\ldots,\tau_\eta)$ are the ηth-order Volterra kernels. Eq. (1) is a generalization of the impulse response model (Eykhoff, 1974), traditionally used for the analysis of linear systems. The main drawback of these models is over-parameterization. Such a drawback can be avoided by expanding the Volterra kernels using an orthonormal basis of functions. This leads to the so-called Wiener/Volterra models.

Wiener/Volterra models have a linear dynamic, composed of filters of orthonormal functions, followed by a static nonlinear mapping (Rugh, 1981). Such models mathematically describe the kernels h_η using orthonormal bases of functions $\{\psi_n\}$ as (Schetzen, 1989):

$$h_\eta(k_1,\ldots,k_\eta)=\sum_{i_1=1}^{\infty}\cdots\sum_{i_\eta=1}^{\infty}\alpha_{i_1,\ldots,i_\eta}\prod_{j=1}^{\eta}\psi_{i_j}(k_j) \tag{2}$$

which assumes that the kernels are absolutely summable on $[0,\infty)$. Given the orthonormality property of the set $\{\psi_n\}$, i.e. $\sum_{k=0}^{\infty}\psi_q(k)\psi_r(k)=\delta_{qr}$, where δ_{qr} is the Kronecker delta, the coefficients $\alpha_{(\cdot)}$ can be computed from (2) as

$$\alpha_{i_1,\ldots,i_\eta}=\sum_{k_1=0}^{\infty}\cdots\sum_{k_\eta=0}^{\infty}h_\eta(k_1,\ldots,k_\eta)\prod_{j=1}^{\eta}\psi_{i_j}(k_j). \tag{3}$$

For computational reasons, Eq. (2) is, in practice, approximated with a finite number M of functions, as follows:

$$\tilde{h}_\eta(k_1,\ldots,k_\eta)=\sum_{i_1=1}^{M}\cdots\sum_{i_\eta=1}^{M}\alpha_{i_1,\ldots,i_\eta}\prod_{j=1}^{\eta}\psi_{i_j}(k_j). \tag{4}$$

If the input signal $u(k)$ in (1) is bounded and normalized so that $|u(k)|<1\ \forall k$, then the higher-order kernels can be ignored in such a way that the resulting Volterra model is truncated into a finite-order N (Eykhoff, 1974). Furthermore, if it is assumed that $u(k)=0$ for $k<0$, then Eq. (1) becomes

$$\tilde{y}(k)=\sum_{\eta=1}^{N}\left[\sum_{i_1=1}^{M}\cdots\sum_{i_\eta=1}^{M}\alpha_{i_1,\ldots,i_\eta}\cdot\prod_{j=1}^{\eta}\left(\sum_{\tau_j=0}^{k}\psi_{i_j}(\tau_j)u(k-\tau_j)\right)\right]. \tag{5}$$

Truncated Volterra models such as those in Eqs. (1) and (5) can approximate to desired accuracy a broad class of stable nonlinear systems (Boyd & Chua, 1985; Schetzen, 1989).

The orthonormal basis functions that are most commonly used in signal and system representations are presented in the sequel.

2.1. Orthonormal basis functions

The first mention of rational orthonormal bases seems to have occurred in Takenaka (1925). Three decades later, the problem of orthonormalizing a set of continuous functions was presented in Kautz (1954), whereas the corresponding discrete case was solved in Broome (1965). The (discrete-time) GOBFs are defined in the complex z-domain for $n=1, 2, \ldots$ as (Van den Hof et al., 1995; Heuberger et al., 2005; Ninness & Gustafsson, 1997)

$$F_n(z) = \frac{z\sqrt{1-|\beta_n|^2}}{z-\beta_n}\prod_{l=1}^{n-1}\left(\frac{1-\bar{\beta}_l z}{z-\beta_l}\right), \tag{6}$$

where $\beta_l, \bar{\beta}_l \in \mathbb{C}$ are the poles of the GOBFs. These functions are the so-called Takenaka–Malmquist functions. The corresponding realizations in the time-domain, $f_n(k)$, are given by the inverse z-transform of (6) and satisfy the orthonormality property. The set $\{f_n\}$ is complete on $\ell^2[0,\infty)$ if and only if $\sum_{n=1}^{\infty}(1-|\beta_n|)<\infty$ (Heuberger et al., 1995, 2005), so any finite energy signal (including absolutely summable kernels) can be approximated with any prescribed accuracy by truncating the infinite expansion.

In general, the functions f_n will be complex, which is physically unreasonable in real system identification problems. In Ninness and Gustafsson (1997), it is shown that this drawback can be overcome by constructing a new orthonormal basis of functions with real impulse responses as a linear combination of the complex functions generated by (6). Using this approach, when all the poles of (6) are real-valued and equal to each other for any value of n, i.e. $\beta_n = \bar{\beta}_n = c$, one gets the particular case of Laguerre functions (Wahlberg, 1991; Silva, 1994)

$$\Phi_n(z) = \sqrt{1-c^2}\,\frac{z}{z-c}\left(\frac{1-cz}{z-c}\right)^{n-1} \tag{7}$$

with c denoting the Laguerre pole. By setting $c=0$, the Laguerre functions simplify to an ordinary pulse basis $\Phi_n(z) = z^{-(n-1)}$ and model (5) reduces to the ordinary nonlinear finite impulse response (NFIR) Volterra model in (1).

The particular case of GOBFs in which the set of poles $\{\beta_n\}$ in (6) is $\{\beta, \bar{\beta}, \beta, \bar{\beta}, \ldots\}$, with $\beta, \bar{\beta} \in \mathbb{C}$, results in the so-called two-parameter Kautz functions. They constitute a second-order generalization of (7) and are defined as follows (Wahlberg, 1994)

$$\Psi_{2n}(z) = \frac{\sqrt{(1-c^2)(1-b^2)}z}{z^2+b(c-1)z-c}\left[\frac{-cz^2+b(c-1)z+1}{z^2+b(c-1)z-c}\right]^{n-1},$$

$$\Psi_{2n-1}(z) = \frac{\sqrt{1-c^2}z(z-b)}{z^2+b(c-1)z-c}\left[\frac{-cz^2+b(c-1)z+1}{z^2+b(c-1)z-c}\right]^{n-1}, \tag{8}$$

where b and c are real-valued constants which are related to the pair of Kautz poles $(\beta, \bar{\beta})$ as

$$b = (\beta+\bar{\beta})/(1+\beta\bar{\beta}), \tag{9}$$

$$c = -\beta\bar{\beta}. \tag{10}$$

Expressions analogous to (8) can be found e.g. in den Brinker, Benders, and Silva (1996), and Heuberger et al. (2005).

3. Selection of the Kautz poles

In this section, a solution to the problem of determining the Kautz pole based on the minimization of an upper bound for the kernel approximation error is presented. The approach to be described here consists of an adaptation of the original (Kautz) problem into a transformed (Laguerre) problem with known solution.

Kautz functions, defined in Eq. (8), depend upon two real-valued parameters (b and c). The selection of these parameters has a direct influence on the computation of the coefficients $\alpha_{i_1,\ldots,i_\eta}$ in (3). The simultaneous optimal selection of both b and c is still under investigation.[1] It is possible, however, to set one of these parameters as constant in order to obtain the best choice for the other according to a certain criterion. Details are given below.

First, let the norm $\|h_\eta\|$ be defined as

$$\|h_\eta\|^2 = \sum_{k_1=0}^{\infty}\cdots\sum_{k_\eta=0}^{\infty} h_\eta^2(k_1,\ldots,k_\eta).$$

By using Eqs. (2)–(4), as well as the orthonormality property of set $\{\psi_n\}$, it is not difficult to deduce that the normalized quadratic error (NQE) of the approximation of kernel h_η, defined as $\mathrm{NQE} \triangleq (\|h_\eta - \tilde{h}_\eta\|^2)/\|h_\eta\|^2$, can be written as follows:

$$\mathrm{NQE} = \frac{\sum_{i_1=M+1}^{\infty}\cdots\sum_{i_\eta=M+1}^{\infty}\alpha^2_{i_1,\ldots,i_\eta}}{\sum_{i_1=1}^{\infty}\cdots\sum_{i_\eta=1}^{\infty}\alpha^2_{i_1,\ldots,i_\eta}}, \tag{11}$$

where $\alpha_{i_1,\ldots,i_\eta}$ are the coefficients of the expansion of $h_\eta(k_1,\ldots,k_\eta)$ according to Eq. (3). An upper bound for (11) when the Kautz functions in (8) are considered can be obtained by means of the following theorem.

Theorem 1. *Let ϕ_n $(n=1,2,\ldots)$ be the Laguerre functions in the time-domain, i.e. the inverse z-transform of* (7), *parameterized by the Kautz parameter c. Also, let $\alpha_{i_1,\ldots,i_\eta}$ be the coefficients of the expansion of kernel $h_\eta(k_1,\ldots,k_\eta)$ using the same Kautz basis and consider the following functions*:

$$g_{\mathrm{even}}(k_1,\ldots,k_\eta) \triangleq \sum_{i_1=1}^{\infty}\cdots\sum_{i_\eta=1}^{\infty}\alpha_{2i_1,\ldots,2i_\eta}\prod_{j=1}^{\eta}\phi_{i_j}(k_j), \tag{12}$$

[1] The required difference equations relative to the discrete-time Kautz functions have not yet been established (Tanguy et al., 2002), at least to the best of the authors' knowledge.

$$g_{\mathrm{odd}}(k_1,\ldots,k_\eta) \triangleq \sum_{i_1=1}^{\infty}\cdots\sum_{i_\eta=1}^{\infty}\alpha_{2i_1-1,\ldots,2i_\eta-1}\prod_{j=1}^{\eta}\phi_{i_j}(k_j). \quad (13)$$

Then, the truncated approximation error of the Volterra kernel h_η *decomposed into an M-term Kautz basis, NQE in* (11), *is bounded by*

$$\mathrm{NQE} \leqslant L(c) \triangleq \frac{2(m_2c^2-2m_1c+m_3)}{\eta(M+1)\|h_\eta\|^2(1-c^2)}, \quad (14)$$

where the terms m_p $(p=1,2,3)$ *are computed as*

$$m_1=\mu_1(g_{\mathrm{even}})+\mu_1(g_{\mathrm{odd}}), \quad (15)$$

$$m_2=\mu_2(g_{\mathrm{even}})+\mu_2(g_{\mathrm{odd}}), \quad (16)$$

$$m_3=\mu_2(g_{\mathrm{even}})+\mu_2(g_{\mathrm{odd}})+\eta\mu_3(g_{\mathrm{even}})+\eta\mu_3(g_{\mathrm{odd}}) \quad (17)$$

with $g_{\mathrm{even}}(k_1,\ldots,k_\eta)$ *and* $g_{\mathrm{odd}}(k_1,\ldots,k_\eta)$ *defined in Eqs.* (12) *and* (13), *respectively. The moments* $\mu_1(x),\mu_2(x),\mu_3(x)$ *are given by*

$$\mu_1(x)=\sum_{j=1}^{\eta}\left[\sum_{k_1=0}^{\infty}\cdots\sum_{k_j=0}^{\infty}\cdots\sum_{k_\eta=0}^{\infty}k_j \cdot x(k_1,\ldots,k_j,\ldots,k_\eta) \cdot x(k_1,\ldots,k_j-1,\ldots,k_\eta)\right], \quad (18)$$

$$\mu_2(x)=\sum_{j=1}^{\eta}\left[\sum_{k_1=0}^{\infty}\cdots\sum_{k_j=0}^{\infty}\cdots\sum_{k_\eta=0}^{\infty}k_j \cdot x^2(k_1,\ldots,k_j,\ldots,k_\eta)\right], \quad (19)$$

$$\mu_3(x)=\sum_{k_1=0}^{\infty}\cdots\sum_{k_\eta=0}^{\infty}x^2(k_1,\ldots,k_\eta). \quad (20)$$

Proof. See Appendix A. □

Computing the functions g_{even} and g_{odd} via Eqs. (12) and (13), respectively, requires that the coefficients $\alpha_{(\cdot)}$ given by (3) be known. However, such computation can be performed independently of $\alpha_{(\cdot)}$ using the following auxiliary theorem.

Theorem 2. *The functions* g_{even} *and* g_{odd}, *defined in Eqs.* (12) *and* (13), *respectively, can be written in terms of both the ηth-order kernel* h_η *and the Kautz functions as*

$$g_{\mathrm{even}}(k_1,\ldots,k_\eta)=\sum_{\tau_1=0}^{\infty}\cdots\sum_{\tau_\eta=0}^{\infty}h_\eta(\tau_1,\ldots,\tau_\eta) \cdot \prod_{j=1}^{\eta}\hat{\psi}_{2(k_j+1)}(\tau_j), \quad (21)$$

$$g_{\mathrm{odd}}(k_1,\ldots,k_\eta)=\sum_{\tau_1=0}^{\infty}\cdots\sum_{\tau_\eta=0}^{\infty}h_\eta(\tau_1,\ldots,\tau_\eta) \cdot \prod_{j=1}^{\eta}\hat{\psi}_{2(k_j+1)-1}(\tau_j), \quad (22)$$

where $\hat{\psi}_{(\cdot)}$ *denotes the Kautz functions in time-domain with* $c=0$.

Proof. The proof follows the lines of reasoning of an analogous theorem provided in Tanguy et al. (2002), whose results are valid only for the first-order Volterra kernel. This reference provides the essential arguments and mathematical foundations also used for proving the generalized result in Theorem 2. The main steps of the proof can be found in da Rosa et al. (2005). □

An optimal choice for parameter c of the Kautz functions can thus be derived from the solution of the following optimization problem:

$$\min_{|c|<1} L(c)=\frac{2(m_2c^2-2m_1c+m_3)}{\eta(M+1)\|h_\eta\|^2(1-c^2)}. \quad (23)$$

Since $\|h_\eta\|$ is a (nonnull) constant for a given system, a necessary condition for solving (23) is $\partial L(c)/\partial c=0$. From Eqs. (16) and (19), it is straightforward to verify that $m_2>0$. Consequently, function $\varpi(c)\triangleq 2(m_2c^2-2m_1c+m_3)$ is convex. It is also differentiable. Moreover, $\varpi(c)$ is nonnegative for all $c\in]-1,1[$, otherwise NQE would be negative according to Eq. (14), which is not possible by definition. Function $\upsilon(c)\triangleq 1-c^2$, in turn, is differentiable, concave and positive for all $c\in]-1,1[$. Hence, $L(c)$ is a pseudo-convex function for $|c|<1$, which implies that $\partial L(c)/\partial c=0$ is a necessary and sufficient condition for solving problem (23) (Bazaraa, Sherali, & Shetty, 1993).

The optimality condition $\partial L(c)/\partial c=0$ is satisfied if and only if:

$$m_1c^2-(m_2+m_3)c+m_1=0. \quad (24)$$

Then, defining $\xi\triangleq(m_2+m_3)/(2m_1)$, the solution of (24) is given by

$$c_{\mathrm{opt}}=\begin{cases}\xi-\sqrt{\xi^2-1} & \text{if } \xi>1,\\ \xi+\sqrt{\xi^2-1} & \text{if } \xi<-1.\end{cases} \quad (25)$$

It is possible to show that the condition $-1\leqslant\xi\leqslant 1$ is unfeasible. See details in Appendix B.

Remark 3. Theorem 2 states that the functions g_{even} and g_{odd} depend solely on the ηth-order kernel h_η and on parameter b of the Kautz basis. Accordingly, the same holds for the terms m_1,m_2,m_3 (see Eqs. (15)–(17)) and for the moments μ_1,μ_2,μ_3 (see Eqs. (18)–(20)). Thus, the analytical solution to the selection of parameter c, given by (25), also depends solely on b and h_η.

The method proposed here can be summarized by the following steps. For every kernel $h_\eta(k_1, \ldots, k_\eta)$ do:

1. Choose an arbitrary value for the Kautz parameter $b \in]-1, 1[$.
2. Once kernel h_η is known, compute the functions g_{even} and g_{odd} using Eqs. (21) and (22), respectively.
3. Compute the terms μ_1, μ_2, μ_3 from Eqs. (18) to (20) and the moments m_1, m_2, m_3 using (15)–(17).
4. Compute c_{opt} using Eq. (25).

The resulting pair (b, c_{opt}) represents the Kautz parameters that minimize the upper bound $L(c)$ for the squared norm of the error resulting from the truncated expansion of the Volterra kernels with this specific value of b.

4. Illustrative example

Suppose that a specific system has the following second-order Volterra kernel:

$$h_2(k_1, k_2) = (k_1 - 2k_2)\exp(-\rho_1 k_1 - \rho_2 k_2) \cdot \cos(\omega_1 k_1 + \omega_2 k_2) \tag{26}$$

for $k_1, k_2 \geqslant 0$. For negative values of k_1 or k_2, $h_2(k_1, k_2)$ is assumed to be null (causal system). Long memory terms of the kernels—longer than 30 memory lags—are considered to be null. In other words, the multiple summations in Eq. (3) go until $k_1, k_2 = 30$. The selection of this factor represents a practical truncation for the Volterra kernels, i.e., a constant $\varepsilon < \infty$ such that $h_\eta(k_1, \ldots, k_\eta)$ is assumed to be null for $k_j > \varepsilon$, $\forall j \in \{1, \ldots, \eta\}$. This value can be set based on the saddle or rise time of the system. The real-valued constant ρ_j $(j = 1, 2)$ can be seen as the decay rate of kernel (26) along the jth axis, whereas ω_j is the frequency with which the kernel oscillates in that direction. For this example, let $\rho_1 = 0.45$, $\rho_2 = 0.7$, $\omega_1 = 100$ and $\omega_2 = 1$.

The choice of $b = 0.4$, for instance, results in $c_{\mathrm{opt}} = -0.2083$, computed via Eq. (25). For $(b, c) = (0.4, -0.2083)$ the NQE associated with the approximation of h_2, computed using (11), is shown in Table 1 for different numbers of Kautz functions. The values of the upper bound (14) are presented as well.

For each value of parameter b in the interval $]-1, 1[$, the corresponding c_{opt} is given by (25). The lowest upper bound $L(c)$ results from the optimal pair $(b, c_{\mathrm{opt}}) = (0.593, -0.2594)$. For these values, the Kautz poles are computed from (9) and (10) as $(\beta, \bar{\beta}) = 0.3734 \pm \mathrm{i}0.3464$. The associated upper bound is $L(c) = 0.0239$ when using $M = 8$ Kautz functions. Fig. 1 illustrates the kernel (26) and Fig. 2 shows the corresponding approximation of this kernel computed using Eq. (4). The error associated with this approximation is obtained from (11) as $\mathrm{NQE} = 1.242 \times 10^{-3}$.

Table 1
Approximation errors and upper bounds for the orthonormal series expansion of kernel (26) with different numbers of Kautz functions

Number of functions (M)	NQE	$L(c)$
2	0.7352	0.8305
4	0.2830	0.3183
6	0.0588	0.1274
8	6.936×10^{-3}	0.0853

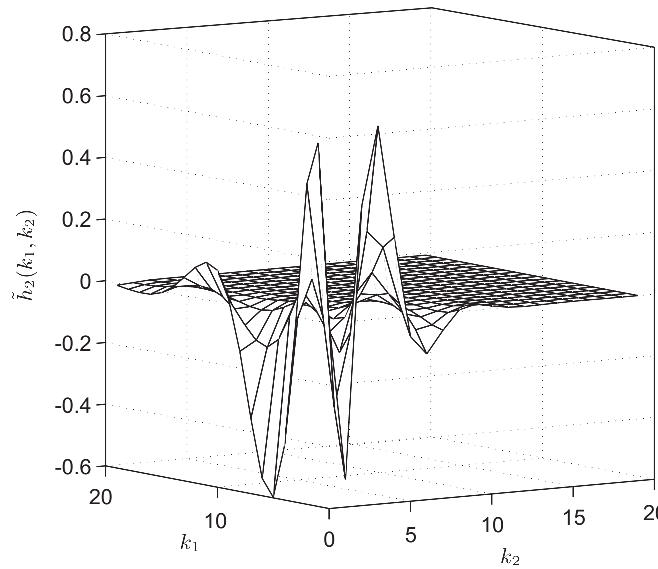

Fig. 1. Second-order kernel (26).

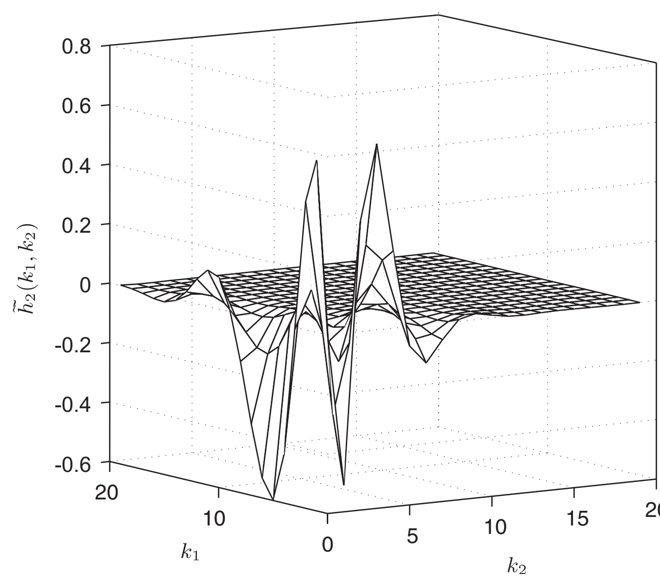

Fig. 2. Approximation of kernel (26) for $b = 0.593$, $c = -0.2594$, and $M = 8$ Kautz functions.

5. Conclusions

In this paper, an optimization of a two-parameter Kautz basis for the orthonormal series expansion of discrete-time Volterra models of any order has been addressed. An analytical solution for the choice of one of the Kautz parameters in Wiener/Volterra models has been derived. This solution is based on the minimization of an upper bound of the error resulting from the truncated approximation of Volterra kernels using Kautz functions and presumes that the kernels are *a priori* known. It indirectly minimizes the number of functions associated with a given series truncation error.

The results reported here represent an extension of the results found in Tanguy et al. (2002), where a solution has been obtained for the particular case in which the Kautz basis is used for expanding the first-order Volterra kernel (linear model). By means of an illustrative example, it has been seen that the use of orthonormal basis functions is a suitable framework for modeling nonlinear systems when prior information about the system kernels is available. Simulation results have shown that the proposed method can provide satisfactory approximations of nonlinear systems with oscillatory behavior.

Acknowledgments

The authors acknowledge financial support from the Brazilian National Council for Scientific and Technological Development (CNPq).

Appendix A. Proof of Theorem 1

Let $x(k_1, \ldots, k_\eta)$ be a nonnull function, which is null for $k_j < 0$ $(j=1, 2, \ldots, \eta)$. Suppose that x is absolutely summable on $[0, \infty)$, i.e.:

$$\sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} |x(k_1, \ldots, k_\eta)| < \infty.$$

The coefficients $\vartheta_{i_1,\ldots,i_\eta}$ of the Laguerre expansion of function x are given by

$$\vartheta_{i_1,\ldots,i_\eta} = \sum_{k_1=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} x(k_1, \ldots, k_\eta) \prod_{j=1}^{\eta} \phi_{i_j}(k_j), \tag{A.1}$$

where ϕ_n is the nth Laguerre function. In (da Rosa, 2005), it is shown that these coefficients satisfy

$$\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta)\vartheta^2_{i_1,\ldots,i_\eta} = \frac{-2c\mu_1(x) + (1+c^2)\mu_2(x) + \eta\mu_3(x)}{1-c^2} \tag{A.2}$$

with $\mu_1(x)$, $\mu_2(x)$ and $\mu_3(x)$ given by Eqs. (18)–(20).

Consider now the Laguerre expansions of the functions g_{even} and g_{odd} in (12) and (13) with coefficients $\gamma_{i_1,\ldots,i_\eta}$ and $\sigma_{i_1,\ldots,i_\eta}$, respectively, such that

$$g_{\text{even}}(k_1, \ldots, k_\eta) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \gamma_{i_1,\ldots,i_\eta} \prod_{j=1}^{\eta} \phi_{i_j}(k_j), \tag{A.3}$$

$$g_{\text{odd}}(k_1, \ldots, k_\eta) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \sigma_{i_1,\ldots,i_\eta} \prod_{j=1}^{\eta} \phi_{i_j}(k_j). \tag{A.4}$$

By comparing Eqs. (12) and (13) to (A.3) and (A.4), it is straightforward to verify that the coefficients $\alpha_{(\cdot)}$ of the expansion of kernel $h_\eta(k_1, \ldots, k_\eta)$ using Kautz functions are related to the coefficients of the expansions above as $\gamma_{i_1,\ldots,i_\eta} = \alpha_{2i_1,\ldots,2i_\eta}$ and $\sigma_{i_1,\ldots,i_\eta} = \alpha_{2i_1-1,\ldots,2i_\eta-1}$, $\forall i_1, \ldots, i_\eta$. Then, by using the inequality (Campello et al., 2004)

$$\eta(M+1) \sum_{i_1=M+1}^{\infty} \cdots \sum_{i_\eta=M+1}^{\infty} \alpha^2_{i_1,\ldots,i_\eta} \leqslant \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta)\alpha^2_{i_1,\ldots,i_\eta} \tag{A.5}$$

it follows that

$$\begin{aligned}
&\sum_{i_1=M+1}^{\infty} \cdots \sum_{i_\eta=M+1}^{\infty} \alpha^2_{i_1,\ldots,i_\eta} \\
&\leqslant \frac{1}{\eta(M+1)} \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta)\alpha^2_{i_1,\ldots,i_\eta} \\
&= \frac{1}{\eta(M+1)} \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (2i_1 + \cdots + 2i_\eta)\alpha^2_{2i_1,\ldots,2i_\eta} \\
&\quad + \frac{1}{\eta(M+1)} \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} [(2i_1-1) + \cdots + (2i_\eta-1)] \\
&\quad \cdot \alpha^2_{2i_1-1,\ldots,2i_\eta-1} \\
&\leqslant \frac{2}{\eta(M+1)} \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta)\gamma^2_{i_1,\ldots,i_\eta} \\
&\quad + \frac{2}{\eta(M+1)} \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} (i_1 + \cdots + i_\eta)\sigma^2_{i_1,\ldots,i_\eta}.
\end{aligned} \tag{A.6}$$

The left-hand side of inequality (A.6) is the (quadratic) approximation error of kernel $h_\eta(k_1, \ldots, k_\eta)$ (Campello et al., 2004, 2006). Then, dividing (A.6) by $\|h_\eta\|^2$ and using (A.2) results in:

$$\begin{aligned}
&\text{NQE} \\
&\leqslant 2 \cdot \left[\frac{-2c\mu_1(g_{\text{even}}) + (1+c^2)\mu_2(g_{\text{even}}) + \eta\mu_3(g_{\text{even}})}{\eta(M+1)\|h_\eta\|^2(1-c^2)}\right] \\
&\quad + 2 \cdot \left[\frac{-2c\mu_1(g_{\text{odd}}) + (1+c^2)\mu_2(g_{\text{odd}}) + \eta\mu_3(g_{\text{odd}})}{\eta(M+1)\|h_\eta\|^2(1-c^2)}\right].
\end{aligned} \tag{A.7}$$

Finally, the substitution of Eqs. (15)–(17) into inequality (A.7) completes the proof. □

Appendix B. Proof of the feasibility of ξ

First and foremost, note that

$$\begin{aligned}
0 &< \sum_{j=1}^{\eta} \left(\sum_{k_1=0}^{\infty} \cdots \sum_{k_j=0}^{\infty} \cdots \sum_{k_\eta=0}^{\infty} k_j [x(k_1, \ldots, k_j, \ldots, k_\eta) \right. \\
&\quad \left. \pm x(k_1, \ldots, k_j - 1, \ldots, k_\eta)]^2 \right) \\
&= 2\mu_2(x) \pm 2\mu_1(x) + \eta\mu_3(x).
\end{aligned} \tag{B.1}$$

By making $x = g_{\text{even}}$ and $x = g_{\text{odd}}$ in (B.1), one gets the following two inequalities:

$$2\mu_2(g_{\text{even}}) \pm 2\mu_1(g_{\text{even}}) + \eta\mu_3(g_{\text{even}}) > 0, \tag{B.2}$$

$$2\mu_2(g_{\text{odd}}) \pm 2\mu_1(g_{\text{odd}}) + \eta\mu_3(g_{\text{odd}}) > 0. \tag{B.3}$$

The sum of (B.2) and (B.3) yields

$$2\mu_2(g_{\text{even}}) + 2\mu_2(g_{\text{odd}}) + \eta\mu_3(g_{\text{even}}) + \eta\mu_3(g_{\text{odd}}) > \pm\left[2\mu_1(g_{\text{even}}) + 2\mu_1(g_{\text{odd}})\right]$$

which allows to conclude that

$$\left|\frac{2\mu_2(g_{\text{even}}) + 2\mu_2(g_{\text{odd}}) + \eta\mu_3(g_{\text{even}}) + \eta\mu_3(g_{\text{odd}})}{2\mu_1(g_{\text{even}}) + 2\mu_1(g_{\text{odd}})}\right| > 1. \tag{B.4}$$

Finally, the use of Eqs. (15)–(17) into (B.4) leads to $|(m_2+m_3)/(2m_1)| = |\xi| > 1$. $\square$

References

Bazaraa, M. S., Sherali, H. D., & Shetty, C. M. (1993). *Nonlinear programming: theory and algorithms*. New York: Wiley.

Bokor, J., & Schipp, F. (1998). Approximate identification in Laguerre and Kautz bases. *Automatica*, *34*(4), 463–468.

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

Broome, P. W. (1965). Discrete orthonormal sequences. *Journal of the Association for Computing Machinery*, *12*(2), 151–168.

Campello, R. J. G. B., Amaral, W. C., & Favier, G. (2006). A note on the optimal expansion of Volterra models using Laguerre functions. *Automatica*, *42*(4), 689–693.

Campello, R. J. G. B., Favier, G., & Amaral, W. C. (2004). Optimal expansions of discrete-time Volterra models using Laguerre functions. *Automatica*, *40*(5), 815–822.

Clowes, G. J. (1965). Choice of time-scaling factor for linear system approximation using orthonormal Laguerre functions. *IEEE Transactions on Automatic Control*, *10*(4), 487–489.

da Rosa, A. (2005). *Expansion of discrete-time Volterra models using Kautz functions*. M.Sc. thesis, School of Electrical and Computer Engineering of the State University of Campinas (FEEC/UNICAMP), Campinas-SP, Brazil (in Portuguese).

da Rosa, A., Amaral, W. C., & Campello, R. J. G. B. (2005). Choice of free parameters in expansions of discrete-time Volterra models using Kautz functions. In *Proceedings of the 16th IFAC world congress*. Prague, Czech Republic.

den Brinker, A. C., Benders, F. P. A., & Silva, T. O. (1996). Optimality conditions for truncated Kautz series. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, *43*(2), 117–122.

Van den Hof, P. M. J., Heuberger, P. S. C., & Bokor, J. (1995). System identification with generalized orthonormal basis functions. *Automatica*, *31*(12), 1821–1834.

Doyle, F. J., III, Pearson, R. K., & Ogunnaike, B. A. (2002). *Identification and control using Volterra models*. Berlin: Springer.

Dumont, G. A., & Fu, Y. (1993). Non-linear adaptive control via Laguerre expansion of Volterra kernels. *International Journal of Adaptive Control and Signal Processing*, *7*(5), 367–382.

Eykhoff, P. (1974). *System identification: parameter and state estimation*. New York: Wiley.

Fu, Y., & Dumont, G. A. (1993). An optimum time scale for discrete Laguerre network. *IEEE Transactions on Automatic Control*, *38*(6), 934–938.

Heuberger, P. S. C., Van den Hof, P. M. J., & Bosgra, O. H. (1995). A generalized orthonormal basis for linear dynamical systems. *IEEE Transactions on Automatic Control*, *40*(3), 451–465.

Heuberger, P. S. C., Van den Hof, P. M. J., & Wahlberg, B. (2005). *Modelling and identification with rational orthogonal basis functions*. Berlin: Springer.

Kautz, W. H. (1954). Transient synthesis in the time domain. *IRE Transactions on Circuit Theory*, *1*(3), 29–39.

Masnadi-Shirazi, M. A., & Ahmed, N. (1991). Optimum Laguerre networks for a class of discrete-time systems. *IEEE Transactions on Signal Processing*, *39*(9), 2104–2108.

Ninness, B., & Gustafsson, F. (1997). A unifying construction of orthonormal bases for system identification. *IEEE Transactions on Automatic Control*, *42*(4), 515–521.

Oliveira, G. H. C., Amaral, W. C., Favier, G., & Dumont, G. A. (2000). Constrained robust predictive controller for uncertain processes modeled by orthonormal series functions. *Automatica*, *36*(4), 563–571.

Rugh, W. J. (1981). *Nonlinear system theory—the Volterra/Wiener approach*. Baltimore, MD: The Johns Hopkins University Press.

Schetzen, M. (1989). *The Volterra and Wiener theories of nonlinear systems*. USA: Robert Krieger Publishing Company.

Silva, T. O. (1994). Optimality conditions for truncated Laguerre networks. *IEEE Transactions on Signal Processing*, *42*(9), 2528–2530.

Silva, T. O. (1995). Optimality conditions for truncated Kautz networks with two periodically repeating complex conjugate poles. *IEEE Transactions on Automatic Control*, *40*(2), 342–346.

Silva, T. O. (1997). Stationary conditions for the L^2 error surface of the generalized orthonormal basis functions lattice filter. *Signal Processing*, *56*(3), 233–253.

Takenaka, S. (1925). On the orthogonal functions and a new formula of interpolation. *Japanese Journal of Mathematics*, *II*, 129–145.

Tanguy, N., Morvan, R., Vilbé, P., & Calvez, L. C. (1995). Optimum choice of free parameter in orthonormal approximations. *IEEE Transactions on Automatic Control*, *40*(10), 1811–1813.

Tanguy, N., Morvan, R., Vilbé, P., & Calvez, L. C. (2000). Pertinent parameters for Kautz approximation. *Electronic Letters*, *36*(8), 769–771.

Tanguy, N., Morvan, R., Vilbé, P., & Calvez, L. C. (2002). Pertinent choice of parameters for discrete Kautz approximation. *IEEE Transactions on Automatic Control*, *47*(5), 783–787.

Wahlberg, B. (1991). System identification using Laguerre models. *IEEE Transactions on Automatic Control*, *36*(5), 551–562.

Wahlberg, B. (1994). System identification using Kautz models. *IEEE Transactions on Automatic Control*, *39*(6), 1276–1282.

Wahlberg, B., & Mäkilä, P. M. (1996). Approximation of stable linear dynamical systems using Laguerre and Kautz functions. *Automatica*, *32*(5), 693–708.

Wiener, N. (1958). *Nonlinear problems in random theory*. New York: Wiley.

Alex da Rosa was born in Goiânia-GO, Brazil, on December 19, 1980. He received the BSc degree in Electrical Engineering from the Federal University of Goiás (UFG) in 2002, and the MSc degree in Electrical Engineering from the State University of Campinas (UNICAMP) in 2005. Currently, he is a doctoral student at the School of Electrical and Computer Engineering of UNICAMP. His research interests include the areas of modeling and identification of nonlinear dynamic systems.

Ricardo J.G.B. Campello was born in Recife, PE, Brazil. He received the BSc degree in Electronics Engineering from UNESP (*Universidade Estadual Paulista*), Ilha Solteira, SP, in 1994 and the MSc and PhD degrees in Electrical Engineering from the School of Electrical and Computer Engineering of UNICAMP (*Universidade Estadual de Campinas*), Campinas, SP, in 1997 and 2002, respectively. In 2002 he was a visiting scholar at the Laboratoire D'Informatique, Signaux et Systèmes de Sophia Antipolis (I3S),

Université de Nice, Sophia Antipolis, France. He is currently with the Department of Computer Sciences of USP (Universidade de São Paulo), São Carlos, SP, Brazil. He is also a researcher of the Brazilian National Research Council (CNPq) and of the Research Foundation of the State of São Paulo (Fapesp). His research interests include computational intelligence, data mining, and dynamic systems identification and control.

Wagner Caradori do Amaral was born in Campinas, SP, Brazil on March 25,1952. He received the BS, MS and doctorate degrees in Electrical Engineering from State University of Campinas, UNICAMP, in 1974, 1976 and 1981, respectively. Since 1991 he has been a Full Professor at the Department of Computer Engineering and Industrial Automation at UNICAMP. From 1995 to 1999 he was the Director of the School of Electrical and Computer Engineering at UNICAMP. His research interests are in the areas of modeling, identification and predictive control.