

This article was downloaded by:[EBSCOHost EJS Content Distribution]
On: 21 May 2008
Access Details: [subscription number 768320842]
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Control

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t713393989>

An optimal expansion of Volterra models using independent Kautz bases for each kernel dimension

Alex da Rosa ^a; Ricardo J. G. B. Campello ^b; Wagner C. Amaral ^a

^a School of Electrical and Computer Engineering, State University of Campinas (UNICAMP), Campinas-SP, Brazil

^b Department of Computer Sciences, University of São Paulo (USP), São Carlos-SP, Brazil

First Published: June 2008

To cite this Article: da Rosa, Alex, Campello, Ricardo J. G. B. and Amaral, Wagner C. (2008) 'An optimal expansion of Volterra models using independent Kautz bases for each kernel dimension', *International Journal of Control*, 81:6, 962 — 975

To link to this article: DOI: 10.1080/00207170701599070
URL: <http://dx.doi.org/10.1080/00207170701599070>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

An optimal expansion of Volterra models using independent Kautz bases for each kernel dimension

ALEX DA ROSA*†, RICARDO J. G. B. CAMPELLO‡ and WAGNER C. AMARAL†

†School of Electrical and Computer Engineering, State University of Campinas (UNICAMP),
13083-852, Campinas-SP, Brazil

‡Department of Computer Sciences, University of São Paulo (USP), São Carlos-SP,
13560-970, Brazil

(Received 3 April 2007; in final form 29 July 2007)

A new solution for the problem of selecting poles of the two-parameter Kautz functions in Volterra models is proposed. In general, a large number of parameters are required to represent the Volterra kernels, although this difficulty can be overcome by describing each kernel using a basis of orthonormal functions, such as the Kautz basis. This representation has a Wiener structure consisting of a linear dynamic generated by the orthonormal basis followed by a non-linear static mapping represented by the Volterra series. The resulting Wiener/Volterra model can be truncated into fewer terms if the Kautz functions are properly designed. The underlying problem is how to select the arbitrary complex poles that fully parametrize these functions. This problem has been approached in previous research by minimizing an upper bound for the error resulting from the truncation of the kernel expansion. The present paper goes even further in that each multidimensional kernel is decomposed into a set of independent Kautz bases, each of which is parametrized by an individual pair of conjugate Kautz poles intended to represent the dominant dynamic of the kernel along a particular dimension. An analytical solution for one of the Kautz parameters, valid for Volterra models of any order, is derived. A simulated example is presented to illustrate these theoretical results. The same approach is then used to model a real non-linear magnetic levitation system with oscillatory behaviour.

1. Introduction

In recent years, there has been growing interest in the use of orthonormal basis functions (OBFs) in studies involving the identification and control of dynamic processes (Schetzen 1989, Dumont and Fu 1993, Bokor and Schipp 1998, Oliveira *et al.* 2000, Doyle III *et al.* 2002, Heuberger *et al.* 1995, 2005). This approach consists of representing a given system in terms of an orthonormal basis for the space of interest. Many problems of control and signal processing can be formulated as optimizing a certain cost-function over the class of stable systems using orthonormal exponentials (Wahlberg and Mäkilä 1996). Such optimization provides good parameters for modelling this

class of systems, because the property of orthonormality of those functions facilitates the solution of the problem and hence leads to a simpler model. Furthermore, OBF-based models have an output-error structure, which takes advantage that the deterministic component of the model can be estimated consistently whenever the system noise is uncorrelated with the system input (Nelles 2001). One important issue regarding the use of orthonormal bases in the model structure is the incorporation of approximate knowledge about the dynamics of the system into the identification process (den Hof *et al.* 1995, Ninness and Gustafsson 1997, Nelles 2001), reducing in this way, the number of parameters to be estimated and consequently increasing the accuracy of the model.

Models using OBFs can generally be constructed with a reduced number of terms to describe a given system.

*Corresponding author. Email: alex@dca.fee.unicamp.br

When properly parametrized, such functions lead to an increase in the speed of convergence in problems of identification (den Hof *et al.* 1995, Heuberger *et al.* 1995), since the parsimony of OBF-based models depends strongly on the choice of the bases poles. Moreover, the orthonormal functions correspond to all-pass filters, which are robust in the implementation of numerical computations (Wahlberg 1994). The most commonly used orthonormal bases of functions are the Laguerre and Kautz bases (Broome 1965, Wahlberg and Mäkilä 1996; Bokor and Schipp 1998), which are more suitable for modelling systems having kernels with dominant dynamics of first and second order, respectively. To model more complex dynamics, the generalized orthonormal basis functions (GOBFs) (den Hof *et al.* 1995, Ninness and Gustafsson 1997, Heuberger *et al.* 2005) are more appropriate, although they involve a more complex parametrization.

The poles of the orthonormal bases are arbitrary within a region of stability. Despite this, their selection can be optimized. Such an optimization allows for reducing the number of functions associated with a given series truncation error, thus reducing the resulting finite-dimensional representation. Many papers have been written about how this can be done, especially by means of analytical approaches, as will be discussed further in this paper. Analytical approaches used for deriving optimal Laguerre poles are possible since the Laguerre functions satisfy a suitable difference equation, which provides further reduction of the mathematical complexity associated with that problem. To the best of the authors' knowledge, however, since an analogous difference equation relative to the Kautz functions has not yet been established, no fully analytical solution for optimal Kautz poles has been obtained so far.

Laguerre functions involve rational transforms with a simple recursive form. Since such functions are completely parametrized by a real-valued pole, they are more appropriate for representing well-damped dynamic systems. The choice of the best Laguerre poles has been extensively addressed and is well established in the literature. The first study concerning this subject which optimized the performance of Laguerre functions based models in the case of linear continuous-time systems seems to be (Clowes 1965). Later, the problem of selecting Laguerre poles in discrete-time linear systems was investigated in Masnadi-Shirazi and Ahmed (1991) by minimizing the error between the impulse response of the system and the corresponding Laguerre model. The main drawback of this strategy is that it requires finding roots of high-order polynomials. In Fu and Dumont (1993) and Tanguy *et al.* (1995), respectively, a cost-function and an energy error are minimized for

obtaining analytical formulae for optimal Laguerre poles. In the context of non-linear systems, an analytical optimization of Laguerre bases for the orthonormal series expansion of second-order Volterra models has first been derived in Campello *et al.* (2001) and further extended to any-order Volterra models in Campello *et al.* (2003, 2004). Such a research represents a generalization of the study found in Fu and Dumont (1993), whose solution is only valid for first-order Volterra models (linear models). In Campello *et al.* (2006), a strict global optimal solution was derived by decomposing each multidimensional kernel of the model using a set of independent orthonormal bases, each of which is parametrized by an individual Laguerre pole associated with the kernel dynamics along a particular dimension. The list of works dealing with the Laguerre pole location also include Silva (1994), which derived optimality conditions for linear truncated Laguerre networks. These conditions of great theoretical interest can however result in complicated computations in practical cases, as already observed in Tanguy *et al.* (2002).

Poorly damped dynamics are not accurately approximated with a small number of Laguerre functions. Indeed, these functions are not well suited to approximate signals with strong oscillatory behaviour (Silva 1995, Tanguy *et al.* 2000, Nelles 2001). This drawback has led to an increasing interest in the two-parameter Kautz functions, first introduced in Kautz (1954). These functions can better approximate systems with oscillatory behaviour because they are parametrized by resonant poles. The problem of approximating linear time-variant stable systems by a finite expansion of the Kautz functions can be found in Wahlberg (1994), which provided a frequency-domain approach in system identification. Optimality conditions for truncated Kautz series in linear discrete-time models have been derived in den Brinker *et al.* (1996) by minimizing the error between the impulse response of a given system and its corresponding Kautz model. In the context of pole location of the Kautz bases, a sub-optimal choice of Kautz poles in the representation of discrete-time linear systems was proposed in Tanguy *et al.* (2002), whereas the corresponding non-linear counterpart was later addressed in da Rosa *et al.* (2005, 2007). Such approaches are said sub-optimal because they consider the optimal selections of only one of the Kautz parameters, and involve the minimization of an upper bound for the kernel truncation error.

Much research concerning the problem of pole location of OBFs require prior information about the system kernel(s) (impulse response in the linear case). When such information is unavailable, the kernels need to be estimated from input-output data measured from the system. An iterative strategy for selecting the

Kautz poles has been presented in Sarroukh *et al.* (2001). A recent method with high computational cost based on an exhaustive search of GOBFs poles has been proposed in Kibangou *et al.* (2005b). This drawback has been ameliorated by the analytical determination of the poles, as presented in Kibangou *et al.* (2005a). Other papers present experimental results regarding the use of orthonormal bases for the identification of real-world systems. In this context, one can cite Nalbantoğlu *et al.* (2003), which studied the pole location via frequency-domain techniques, and Ziaei and Wang (2006), where the system identification based on GOBFs with both real and complex poles is presented. In Patwardhan and Shah (2005), it is proposed a decomposed strategy to estimate only the GOBFs poles by nonlinear iterative search and the orthonormal expansion coefficients analytically. One consequence of this method is the reduction in dimensionality of the optimization problem. Surveys of the studies involving GOBFs also are found in Ninness *et al.* (1999) and Silva (1997), with the latter presenting stationary conditions for optimal GOBFs-based linear models.

In the present paper, an analytical solution for one of the parameters related to the Kautz pole is derived when any-order Volterra kernels are decomposed into a set of independent orthonormal bases, each of which parametrized by an individual pair of conjugate Kautz poles associated with the dominant dynamic of the kernel along a particular dimension. Such a solution is based on the minimization of an upper bound of the error resulting from the truncated approximation of Volterra kernels using Kautz functions. This is an extension of the results presented in da Rosa *et al.* (2005, 2007), where the solution involves a single Kautz basis for expanding a given kernel along all its dimensions. Using independent bases for each kernel dimension is expected to reduce the truncation error when the dominant dynamics along the multiple dimensions are different from one another. The results to be presented here can also be seen as a generalization of the results found in Campello *et al.* (2006), from Laguerre to Kautz bases.

This paper is outlined as follows. In the next section, orthonormal basis functions are presented in the context of Volterra and Wiener/Volterra models. In §3, an upper bound for the kernel truncation error in expansion of Volterra models using Kautz functions is derived. By minimizing this bound, an optimization problem for the selection of the Kautz poles is formulated and solved. An example illustrating the theoretical results is provided in §4, and a comparison to previous work is presented as well. Section 5 presents an application of the proposed method for the modeling of a real magnetic levitation system with strong

non-linear behaviour. Finally, §6 addresses the conclusions and the perspectives for future research.

2. Volterra and Wiener/Volterra models

A Volterra model is essentially an input–output functional expansion of a non-linear system whose structure is given by a straightforward generalization of the unit-impulse response model (Eykhoff 1974, Nelles 2001). The direct link of these models to the non-linear realization theory and their ability to represent a wide class of non-linear systems make them very attractive. In the discrete-time domain, the mathematical description of a Volterra model relates the output $y(k)$ of a physical process to its input $u(k)$ as (Rugh 1981, Schetzen 1989)

$$y(k) = \sum_{\eta=1}^{\infty} \sum_{\tau_1=0}^{\infty} \dots \sum_{\tau_{\eta}=0}^{\infty} h_{\eta}(\tau_1, \dots, \tau_{\eta}) \prod_{l=1}^{\eta} u(k - \tau_l), \quad (1)$$

where the multidimensional functions $h_{\eta}(\tau_1, \dots, \tau_{\eta})$ are the η -th-order Volterra kernels. Although these models can describe a wide class of non-linear systems, their practical use is limited due to the usually large number of coefficients that are needed to be estimated, even for simple problems. The reason is that Volterra models, like impulse response models, do not depend on past output signals. Such a drawback can be avoided by expanding the Volterra kernels using orthonormal bases of functions. The number of parameters necessary to represent the models can thus be drastically reduced if properly designed bases of functions are adopted.

Approximating a given system by means of a truncated orthonormal series was first suggested by Wiener (1966). A representation of Wiener type (Rugh 1981, Nelles 2001) consists of a linear dynamic, here composed by a set of orthonormal filters, followed by a nonlinear static mapping, here represented by the Volterra series. The basic idea of such Wiener/Volterra models is to describe the kernels h_{η} by means of an expansion using OBFs in such a way that one needs to determine the coefficients of this expansion in lieu of the coefficients of the kernel. In the more general case of using independent bases of functions $\{\psi_{l,n}\}$ for expanding the kernel along its multiple dimensions ($l=1, \dots, \eta$), the expansion is as follows (Schetzen 1989):

$$h_{\eta}(k_1, \dots, k_{\eta}) = \sum_{i_1=1}^{\infty} \dots \sum_{i_{\eta}=1}^{\infty} \alpha_{i_1, \dots, i_{\eta}} \prod_{l=1}^{\eta} \psi_{l, i_l}(k_l) \quad (2)$$

which assumes that the kernels are absolutely summable on $[0, \infty)$. In practice, this condition can be assured by forcing the long memory terms of the kernels to be null,

which is possible provided that the system to be modeled is stable. In other words, $h_\eta(k_1, \dots, k_\eta)$ is assumed to be zero for $k_l > \epsilon, \forall l \in \{1, \dots, \eta\}$. An appropriate value for $\epsilon < \infty$ can be set based on the saddle or rise time of the system.

The kernel expansion coefficients $\alpha_{(\cdot)}$ in (2) can be derived using the orthonormality property of the sets $\{\psi_{l,n}\}$, i.e., $\sum_{k=0}^{\infty} \psi_{l,q}(k)\psi_{l,r}(k) = \delta_{qr}$, where δ_{qr} is the Kronecker delta, as

$$\alpha_{i_1, \dots, i_\eta} = \sum_{k_1=0}^{\infty} \dots \sum_{k_\eta=0}^{\infty} h_\eta(k_1, \dots, k_\eta) \prod_{l=1}^{\eta} \psi_{l,i_l}(k_l). \quad (3)$$

If the input signal $u(k)$ in (1) is bounded so that $|u(k)| < 1 \quad \forall k$, then the higher-order kernels can be ignored in such a way that the resulting Volterra model is truncated to a finite-order N (Eykhoff 1974). Furthermore, for computational reasons, equation (2) is, in practice, approximated with a finite number M of functions, as follows:

$$\tilde{h}_\eta(k_1, \dots, k_\eta) = \sum_{i_1=1}^M \dots \sum_{i_\eta=1}^M \alpha_{i_1, \dots, i_\eta} \prod_{l=1}^{\eta} \psi_{l,i_l}(k_l). \quad (4)$$

Hence, if it is assumed that $u(k) = 0$ for $k < 0$, then equation (1) can be approximated as

$$\tilde{y}(k) = \sum_{\eta=1}^N \left[\sum_{i_1=1}^M \dots \sum_{i_\eta=1}^M \alpha_{i_1, \dots, i_\eta} \prod_{l=1}^{\eta} \left(\sum_{\tau_l=0}^k \psi_{l,i_l}(\tau_l) u(k - \tau_l) \right) \right]. \quad (5)$$

In Boyd and Chua (1985), it is shown that truncated Volterra models, such as those in (5), can approximate to desired accuracy any time-invariant analytical non-linear system with fading memory and bounded input. Indeed, any prescribed approximation accuracy can be obtained by setting the number of kernels, N , and functions, M , to appropriate finite values.

Since model (5) is linear-in-the-parameters $\alpha_{(\cdot)}$, these parameters can be straightforwardly computed using a least-squares algorithm. Clearly, the overall number of parameters depends upon the number of functions, M , used in the kernels expansions. This number represents a trade-off between accuracy and parsimony of the model and can be significantly reduced by optimally designing the set of orthonormal filters $\{\psi_{l,i_l}\}$. However, the actual value for M needed to provide an accurate representation depends on the complexity of the specific problem in hand. Dynamic systems with multiple dominant modes, for example, typically require models with a larger number of functions.

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

2.1 Orthonormal basis functions

The use of rational orthonormal filters for representing signals and systems has a long history, since the pioneering proposal of Takenaka (1925), which investigated orthonormal networks in system modelling. The problem of building a set of continuous-time orthonormal functions was introduced in Kautz (1954), whereas the corresponding discrete case was solved in Broome (1965). Discrete-time generalized orthonormal basis functions are defined in the complex z -domain as (den Hof *et al.* 1995, Ninness and Gustafsson 1997, Heuberger *et al.* 2005)

$$F_{l,n}(z) = \frac{z \sqrt{1 - |\beta_{l,n}|^2}}{z - \beta_{l,n}} \prod_{j=1}^{n-1} \left(\frac{1 - \bar{\beta}_{l,j} z}{z - \beta_{l,j}} \right) \quad n = 1, 2, \dots, \quad (6)$$

where $\beta_{l,n}, \bar{\beta}_{l,n} \in \mathbb{C}$ are the poles of the GOBFs. The corresponding realizations in the time-domain, namely $f_{l,n}(k)$, are given by the inverse Z -transform of equation (6) and satisfy the property of orthonormality. The set $\{f_{l,n}\}$ is complete on $\ell^2[0, \infty)$ if and only if $\sum_{n=1}^{\infty} (1 - |\beta_{l,n}|) < \infty$ (Heuberger *et al.* 1995, 2005), so any finite energy signal (including absolutely summable kernels) can be approximated with any prescribed accuracy by linearly combining a certain finite number of such functions. In general, the functions $f_{l,n}(k)$ will be complex, although this is physically unrealistic in 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, consisting of a linear combination of the complex functions generated by (6).

When all the poles of (6) are real-valued and equal to each other for any value of n , i.e., $\beta_{l,n} = \bar{\beta}_{l,n} = c_l$, one gets the Laguerre basis (Fu and Dumont 1993, Silva 1994, Tanguy *et al.* 1995)

$$\Phi_{l,n}(z) = \frac{z \sqrt{1 - c_l^2}}{z - c_l} \left(\frac{1 - c_l z}{z - c_l} \right)^{n-1} \quad n = 1, 2, \dots \quad (7)$$

with c_l denoting the Laguerre pole of the l th basis. By setting $c_l = 0$, the Laguerre functions are simplified to the pulse basis $\Phi_{l,n}(z) = z^{-(n-1)}$, and model (5) is reduced to an ordinary non-linear finite impulse response (NFIR) Volterra model, i.e., a truncated version of (1).

The particular case of GOBFs in which the set of poles $\{\beta_{l,n}\}$ in (6) is $\{\beta_l, \bar{\beta}_l, \beta_l, \bar{\beta}_l, \dots\}$, with $\beta, \bar{\beta} \in \mathbb{C}$, results in the so-called two-parameter Kautz functions. These functions constitute a second-order generalization

of (7) and are defined as follows (Wahlberg 1994, Ninness and Gustafsson 1997):

$$\Psi_{l,2n}(z) = \frac{z\sqrt{(1-c_l^2)(1-b_l^2)}}{z^2 + b_l(c_l-1)z - c_l} \left[\frac{-c_l z^2 + b_l(c_l-1)z + 1}{z^2 + b_l(c_l-1)z - c_l} \right]^{n-1}$$

$$\Psi_{l,2n-1}(z) = \frac{z(z-b_l)\sqrt{1-c_l^2}}{z^2 + b_l(c_l-1)z - c_l} \left[\frac{-c_l z^2 + b_l(c_l-1)z + 1}{z^2 + b_l(c_l-1)z - c_l} \right]^{n-1}$$

$n = 1, 2, \dots,$ (8)

where b_l and c_l are real-valued constants related to the pair of Kautz poles $(\beta_l, \bar{\beta}_l)$ as

$$b_l = \frac{(\beta_l + \bar{\beta}_l)}{(1 + \beta_l \bar{\beta}_l)} \tag{9}$$

$$c_l = -\beta_l \bar{\beta}_l. \tag{10}$$

Expressions analogous to (8) can be found elsewhere, e.g., in Silva (1995) and den Brinker *et al.* (1996).

3. Selection of Kautz parameters in expansions of Volterra kernels

In this section, the problem of sub-optimal choice for the Kautz poles based on the minimization of an upper bound for the kernel approximation error is presented. The approach consists of the adaptation of the original (Kautz) problem into a transformed (Laguerre) problem with known solution. The underlying problem considered here is how to select the Kautz parameters b_l and c_l in (8) so as to minimize the upper bound resulting from the truncated series expansion in (4). The simultaneous optimal selection of both b_l and c_l is still under investigation, but it is possible to set one of them as constant and obtain the best choice for the other. In the proposed method, each multidimensional kernel is decomposed into a set of independent orthonormal bases, each of which is parametrized by an individual Kautz pole. Details are given below.

By defining the norm $\|h_\eta\|$ as

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

and using equations (2) through (4), as well as the orthonormality property of the sets $\{\psi_{l,n}\}$, the normalized quadratic error (NQE) of the approximation of kernel h_η , defined as $NQE \triangleq (\|h_\eta - \bar{h}_\eta\|^2) / \|h_\eta\|^2$, can be written as follows:

$$NQE = \frac{\sum_{i_1=M+1}^{\infty} \dots \sum_{i_\eta=M+1}^{\infty} \alpha_{i_1, \dots, i_\eta}^2}{\sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \alpha_{i_1, \dots, i_\eta}^2}, \tag{11}$$

where $\alpha_{i_1, \dots, i_\eta}$ are the coefficients of the expansion of $h_\eta(k_1, \dots, k_\eta)$ according to equation (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 $\{\phi_{l,n}\}$ be time-domain Laguerre bases, i.e. the inverse Z-transform of (7), parametrized by parameters c_l ($l=1, \dots, \eta$). Also, let $\alpha_{i_1, \dots, i_\eta}$ be the coefficients of the expansion of the kernel $h_\eta(k_1, \dots, k_\eta)$ as in equation (3) using Kautz bases $\{\psi_{l,n}\}$ with the same parameters c_l as $\{\phi_{l,n}\}$. Now define the following functions for $l=1, \dots, \eta$:

$$g_{\text{even},l}(k_1, \dots, k_\eta) \triangleq \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1, \dots, 2i_\eta} \prod_{l=1}^{\eta} \phi_{l,i_l}(k_l) \tag{12}$$

$$g_{\text{odd},l}(k_1, \dots, k_\eta) \triangleq \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1-1, \dots, 2i_\eta-1} \prod_{l=1}^{\eta} \phi_{l,i_l}(k_l). \tag{13}$$

The truncated approximation error of the Volterra kernel h_η decomposed into η M -term Kautz bases (one independent basis along each kernel dimension) is bounded by

$$NQE \leq \frac{2}{\eta(M+1)\|h_\eta\|^2} \sum_{l=1}^{\eta} \left[\frac{m_{2,l}c_l^2 - 2m_{1,l}c_l + m_{3,l}}{1 - c_l^2} \right] \tag{14}$$

where the terms $m_{p,l}$ ($p=1, 2, 3$) are computed as

$$m_{1,l} = \mu_{1,l}(g_{\text{even},l}) + \mu_{1,l}(g_{\text{odd},l}) \tag{15}$$

$$m_{2,l} = \mu_{2,l}(g_{\text{even},l}) + \mu_{2,l}(g_{\text{odd},l}) \tag{16}$$

$$m_{3,l} = \mu_{2,l}(g_{\text{even},l}) + \mu_{2,l}(g_{\text{odd},l}) + \eta\mu_{3,l}(g_{\text{even},l}) + \eta\mu_{3,l}(g_{\text{odd},l}) \tag{17}$$

and the moments $\mu_{1,l}(x)$, $\mu_{2,l}(x)$, $\mu_{3,l}(x)$ are given by

$$\mu_{1,l}(x) = \sum_{k_1=0}^{\infty} \dots \sum_{k_l=0}^{\infty} \dots \sum_{k_\eta=0}^{\infty} k_l x(k_1, \dots, k_l, \dots, k_\eta) \times x(k_1, \dots, k_l - 1, \dots, k_\eta) \tag{18}$$

$$\mu_{2,l}(x) = \sum_{k_1=0}^{\infty} \dots \sum_{k_l=0}^{\infty} \dots \sum_{k_\eta=0}^{\infty} k_l x^2(k_1, \dots, k_l, \dots, k_\eta) \tag{19}$$

$$\mu_{3,l}(x) = \left(\frac{1}{\eta}\right) \sum_{k_1=0}^{\infty} \dots \sum_{k_l=0}^{\infty} \dots \sum_{k_\eta=0}^{\infty} x^2(k_1, \dots, k_l, \dots, k_\eta). \tag{20}$$

Proof: See Appendix A. □

The results of Theorem 1 can be used to obtain an optimal choice for the Kautz parameter c_l by solving the optimization problem

$$\min_{|c_l| < 1, l=1, \dots, \eta} L(c_1, \dots, c_\eta) = \frac{2}{\eta(M+1)\|h_\eta\|^2} \times \sum_{l=1}^{\eta} \left[\frac{m_{2,l}c_l^2 - 2m_{1,l}c_l + m_{3,l}}{1 - c_l^2} \right]. \tag{21}$$

Since $\|h_\eta\|$ is a (non-null) constant for a given system, the necessary conditions for solving (21) are $\partial L/\partial c_l = 0$, $\forall l = \{1, \dots, \eta\}$. From equations (19) and (16), it is straightforward to verify that $m_{2,l} > 0$. Consequently, function $\varpi(c_l) \triangleq m_{2,l}c_l^2 - 2m_{1,l}c_l + m_{3,l}$ is convex. It is also differentiable. Moreover, $\varpi(c_l)$ is non-negative for all $c_l \in]-1, 1[$; otherwise NQE would be negative according to equation (14), which is not possible by definition. The function $\nu(c_l) \triangleq 1 - c_l^2$, in turn, is differentiable, concave and positive for $c_l \in]-1, 1[$. Hence, $L(c_1, \dots, c_\eta)$ is a pseudo-convex function for $|c_l| < 1$, which implies that the conditions $\partial L/\partial c_l = 0$, $\forall l = \{1, \dots, \eta\}$, are necessary and sufficient for solving the problem (21) (Bazaraa *et al.* 1993).

The optimality conditions $\partial L/\partial c_l = 0$ are satisfied if and only if

$$m_{1,l}c_l^2 - (m_{2,l} + m_{3,l})c_l + m_{1,l} = 0 \quad l = 1, \dots, \eta. \tag{22}$$

Then, defining ξ_l as $(m_{2,l} + m_{3,l})/(2m_{1,l})$, the solution of (22) is given by

$$c_{opt,l} = \begin{cases} \xi_l - \sqrt{\xi_l^2 - 1} & \text{if } \xi_l > 1 \\ \xi_l + \sqrt{\xi_l^2 - 1} & \text{if } \xi_l < -1 \end{cases} \quad l = 1, \dots, \eta. \tag{23}$$

Equation (23) is thus an analytical solution for the selection of the parameter c_l of the Kautz functions according to the criterion (21). After the setting of the value of parameter b_l , it can be used to minimize the upper bound $L(c_1, \dots, c_\eta)$ for the squared norm of the error resulting from the truncated expansion of the Volterra kernels.

It is possible to show that the absolute value of variable ξ_l is greater than unity. To see this, notice

first that the following inequality holds for $\forall l = \{1, \dots, \eta\}$:

$$0 < \sum_{k_1=0}^{\infty} \dots \sum_{k_l=0}^{\infty} \dots \sum_{k_\eta=0}^{\infty} k_l [x(k_1, \dots, k_l, \dots, k_\eta) \pm x(k_1, \dots, k_l - 1, \dots, k_\eta)]^2 = 2\mu_{2,l}(x) \pm 2\mu_{1,l}(x) + \eta\mu_{3,l}(x). \tag{24}$$

By making $x = g_{even,l}$ and $x = g_{odd,l}$ in (24), one gets (25) and (26) respectively

$$2\mu_{2,l}(g_{even,l}) \pm 2\mu_{1,l}(g_{even,l}) + \eta\mu_{3,l}(g_{even,l}) > 0 \tag{25}$$

$$2\mu_{2,l}(g_{odd,l}) \pm 2\mu_{1,l}(g_{odd,l}) + \eta\mu_{3,l}(g_{odd,l}) > 0. \tag{26}$$

The sum of (25) and (26) yields

$$2\mu_{2,l}(g_{even,l}) + 2\mu_{1,l}(g_{odd,l}) + \eta\mu_{3,l}(g_{even,l}) + \eta\mu_{3,l}(g_{odd,l}) > \pm [2\mu_{1,l}(g_{even,l}) + 2\mu_{1,l}(g_{odd,l})]$$

which allows the conclusion that:

$$\left| \frac{2\mu_{2,l}(g_{even,l}) + 2\mu_{2,l}(g_{odd,l}) + \eta\mu_{3,l}(g_{even,l}) + \eta\mu_{3,l}(g_{odd,l})}{2\mu_{1,l}(g_{even,l}) + 2\mu_{1,l}(g_{odd,l})} \right| > 1. \tag{27}$$

Finally, the use of equations (15)–(17) into (27) leads to the result that $|(m_{2,l} + m_{3,l})/(2m_{1,l})| = |\xi_l| > 1$.

Theorem 2: *If h_η is symmetric (Schetzen 1989), which means that it takes the same value for any permutation of its arguments (e.g. $h_2(k_1, k_2) = h_2(k_2, k_1)$ for the second-order kernel), and if all parameters b_l are set equally for all $l = 1, \dots, \eta$, then the solution (23) is reduced to the special case in which the expansion of the kernel using a single Kautz basis is adopted, i.e. (da Rosa *et al.* 2005, 2007):*

$$c_{opt} = \begin{cases} \xi - \sqrt{\xi^2 - 1} & \text{if } \xi > 1 \\ \xi + \sqrt{\xi^2 - 1} & \text{if } \xi < -1, \end{cases} \tag{28}$$

where $\xi = \sum_{l=1}^{\eta} \xi_l$.

Proof: The proof is straightforward since the terms $m_{1,l}$, $m_{2,l}$, $m_{3,l}$, $\mu_{1,l}$, $\mu_{2,l}$, $\mu_{3,l}$ in (15)–(20) are constant for $l = 1, \dots, \eta$ when kernel h_η is symmetric and when $b = b_1 = \dots = b_\eta$. In this case, the single solution in (28) becomes clearly equivalent to each individual solution in (23), i.e., $c_{opt} = c_{opt,1} = \dots = c_{opt,\eta}$. □

Note that, from the perspective of the output of the Volterra model in (1), it is possible to replace any non-symmetric kernel with a symmetric equivalent by means of an ordinary symmetrization procedure (Schetzen 1989). However, the symmetric kernels are equivalent to their non-symmetric counterparts only in terms of the model output, i.e., they are not equivalent to each other as multidimensional functions to be

described by means of a truncated series expansion. This means that the symmetrization of a given kernel does not ensure that its conventional expansion using a single Kautz basis is equivalent to the expansion of the original kernel using a particular basis for each dimension. The advantage of using an independent basis for each dimension $l=1, \dots, \eta$, as solved in equation (23), is that it enhances flexibility and is expected to reduce the truncation error when the kernel has different dominant dynamics along its multiple dimensions. For instance, independent bases can provide a better approximation of a second-order kernel with well-damped dynamics along one of the axial directions and poorly damped dynamics along the other. In summary, Theorem 2 suggests that if one begins by identifying a symmetric Volterra model, as usual in practice, then there is no reason to use multiple Kautz bases to approximate each kernel. On the other hand, if one does decide to adopt multiple bases, then the use of asymmetric kernels may provide better approximation results.

When using equations (12) and (13) for computing the functions $g_{\text{even},l}$ and $g_{\text{odd},l}$, one must explicitly compute the coefficients $\alpha_{(\cdot)}$ given by (3). However, such computations can be performed independently of $\alpha_{(\cdot)}$ using the following theorem.

Theorem 3: Functions $g_{\text{even},l}(k_1, \dots, k_\eta)$ and $g_{\text{odd},l}(k_1, \dots, k_\eta)$, defined in equations (12) and (13), respectively, can be written in terms of both the η th-order kernel h_η and the Kautz functions as

$$g_{\text{even},l}(k_1, \dots, k_\eta) = \sum_{\tau_1=0}^{\infty} \dots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \dots, \tau_\eta) \prod_{l=1}^{\eta} \hat{\psi}_{l,2(k_l+1)}(\tau_l) \tag{29}$$

$$g_{\text{odd},l}(k_1, \dots, k_\eta) = \sum_{\tau_1=0}^{\infty} \dots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \dots, \tau_\eta) \prod_{l=1}^{\eta} \hat{\psi}_{l,2(k_l+1)-1}(\tau_l) \tag{30}$$

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

Proof: See Appendix B.

Theorem 3 extends previous results (Tanguy *et al.* 2002, da Rosa *et al.* 2007) in the sense that different Kautz poles can now be set for each dimension $l=1, 2, \dots, \eta$ of the η th kernel $h_\eta(\tau_1, \dots, \tau_\eta)$. It states that $g_{\text{even},l}$ and $g_{\text{odd},l}$ depend solely on the η th-order kernel h_η and on parameter b_l of the l th Kautz basis. The same holds true for $m_{1,l}, m_{2,l}, m_{3,l}, \mu_{1,l}, \mu_{2,l}, \mu_{3,l}$ in equations (15)–(20). Thus, the analytical solution to an optimal selection of parameter c_l , given by (23), also depends solely on b_l and h_η .

The method proposed here can be summarized by the following steps. For each dimension $l=1, \dots, \eta$ of kernel $h_\eta(k_1, \dots, k_\eta)$

- (a) choose an arbitrary value for Kautz parameter $b_l \in]-1, 1[$;
- (b) once kernel h_η is known, compute the functions $g_{\text{even},l}$ and $g_{\text{odd},l}$ using equations (29) and (30), respectively;
- (c) compute the terms $\mu_{1,l}, \mu_{2,l}, \mu_{3,l}$ from equations (18)–(20) and the terms $m_{1,l}, m_{2,l}, m_{3,l}$ using (15)–(17);
- (d) calculate $c_{\text{opt},l}$ using equation (23).

The pair obtained $(b_l, c_{\text{opt},l})$ thus represents the Kautz parameters that minimize the upper bound $L(c_1, \dots, c_\eta)$ in (21) for the squared norm of the error resulting from the truncated expansion of each Volterra kernel with this specific value of b_l .

4. Illustrative example

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

$$h_2(k_1, k_2) = (k_1 - 2k_2) \exp(-\rho_1 k_1 - \rho_2 k_2) \cos(\omega_1 k_1 + \omega_2 k_2) \tag{31}$$

for $k_1, k_2 \geq 0$. For negative values of k_1 or k_2 , the kernel $h_2(k_1, k_2)$ is assumed to be null (causal system). Memory terms of h_2 which are longer than 30 lags are considered null, i.e., the multiple summations in equation (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 $\epsilon < \infty$ such that $h_\eta(k_1, \dots, k_\eta)$ is assumed to be null for $k_l > \epsilon, \forall l \in \{1, \dots, \eta\}$. The real-valued constant ρ_l ($l=1, 2$) can be seen as the decay rate of the kernel in (31) along the l th axis, whereas ω_l is the frequency with which the kernel oscillates in that direction. For this example, the values $\rho_1 = 0.45, \rho_2 = 0.7, \omega_1 = 100$ and $\omega_2 = 1$ were used.

The method proposed in the previous section will be carried out to compute the Kautz poles for each dimension of the kernel in (31). For instance, the values $b_1 = 0.6$ for the first axial direction and $b_2 = 0.5$ for the second have been arbitrarily chosen. Computational simulations provided the following values, computed using equation (23): $c_{\text{opt},1} = -0.2321$ and $c_{\text{opt},2} = -0.3058$. With the values of Kautz parameters $(b_1, c_{\text{opt},1}) = (0.6, -0.2321)$ and $(b_2, c_{\text{opt},2}) = (0.5, -0.3058)$, the approximation of kernel h_2 can be computed using equation (4). The error associated with this approximation, calculated via (11), is shown in table 1 for different numbers of Kautz functions.

The kernel in (31) is depicted in figure 1(a), whereas figure 1(b) illustrates the corresponding approximation of this kernel with $M=8$. The Kautz poles are calculated from (9) and (10) as $\beta_1=0.3696 \pm i0.3090$ for the first direction ($l=1$) and $\beta_2=0.3264 \pm i0.4463$ for the second direction ($l=2$).

The method proposed in da Rosa *et al.* (2005, 2007), in which equal parameters (poles) are considered for each kernel dimension, has also been applied to the kernel in equation (31). For the sake of comparison, the same values of the parameters b_l just selected (0.6 and 0.5) have been adopted for this new computational simulation; the numerical results are summarized in table 2.

As can be seen by comparing the results in tables 1 and 2, the use of an independent basis for each dimension generally reduces the truncation error when the dominant dynamics of the kernel along its multiple dimensions are different from one another. By varying parameters b_1 and b_2 within the region of feasibility $(b_1, b_2) \in]-1, 1[\times]-1, 1[$, each set of parameters

$(b_1, c_{opt,1})$ and $(b_2, c_{opt,2})$ provides an approximation error. The values of c_1 and c_2 providing the best approximation are those for which NQE is the lowest. It is obtained from equation (23) by choosing $b_1=0.770$ and $b_2=0.345$, which results in $c_{opt,1}=-0.4583$ and $c_{opt,2}=-0.2473$, respectively. For $(b_1, c_{opt,1})=(0.770, -0.4583)$ and $(b_2, c_{opt,2})=(0.345, -0.2473)$, the Kautz poles are $\beta_1=0.5615 \pm i0.3783$ and $\beta_2=0.2152 \pm i0.4483$. Equation (4) with $M=8$ gives the corresponding approximation of the kernel in (31). The error associated with this approximation is shown in figure 2, resulting in $NQE = 3.6082 \times 10^{-7}$.

5. Modelling of a magnetic levitation system

The magnetic levitation system considered here, schematically shown in figure 3, consists of upper and lower drive coils that produce a magnetic field in response to a DC current. One or two magnets travel along a

Table 1. Approximation errors of the orthogonal expansion of the kernel in (31) with different numbers of Kautz functions.

Number of functions (M)	$NQE = \frac{\ h_2 - \tilde{h}_2\ ^2}{\ h_2\ ^2}$
2	0.8110
4	0.1036
6	0.0077
8	8.1136×10^{-4}

Table 2. Approximation errors of the orthogonal expansion of the kernel in (31) using the method presented in (da Rosa *et al.* 2005, 2007).

Number of functions (M)	NQE for $b_1=b_2=0.6$	NQE for $b_1=b_2=0.5$
2	0.8422	0.7878
4	0.1081	0.2029
6	0.0086	0.0133
8	8.2405×10^{-4}	0.0046

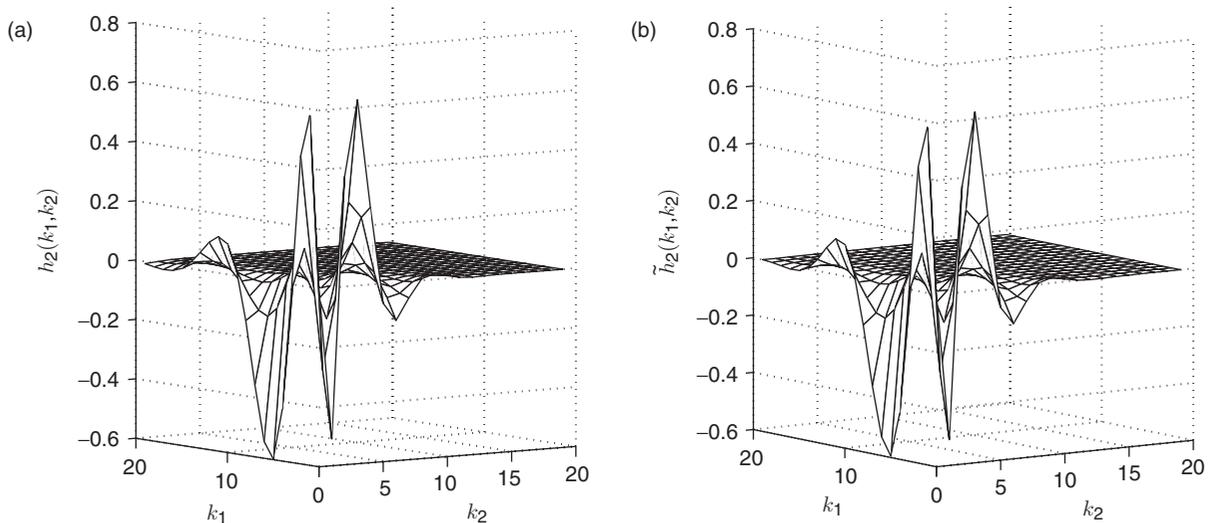


Figure 1. Second-order kernel $h_2(k_1, k_2)$ and its approximation $\tilde{h}_2(k_1, k_2)$ with $(b_1, c_{opt,1})=(0.6, -0.2321)$, $(b_2, c_{opt,2})=(0.5, -0.3058)$ and $M=8$.

precision ground glass guide rod. By energizing the lower coil, a single magnet is levitated by a repulsive magnetic force. As current in the coil increases, the field strength also increases and the height of the levitated magnet is increased too. For the upper coil, the levitating force is attractive. Two laser-based sensors measure the magnet positions. The magnets are of an ultra-high strength rare earth (NeBFe) type and are designed to provide large levitated displacements to clearly demonstrate principles of levitation and motion control (ECP 1999).

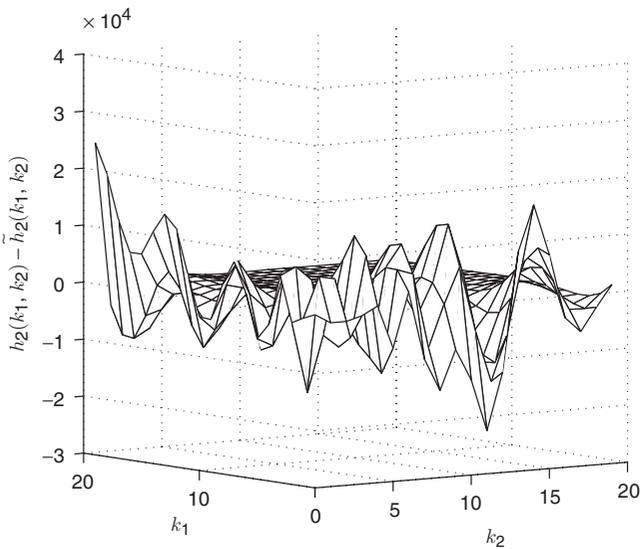


Figure 2. Error surface $h_2(k_1, k_2) - \tilde{h}_2(k_1, k_2)$ for the optimal Kautz poles $\beta_1 = 0.5615 \pm i0.3783$ and $\beta_2 = 0.2152 \pm i0.4483$.

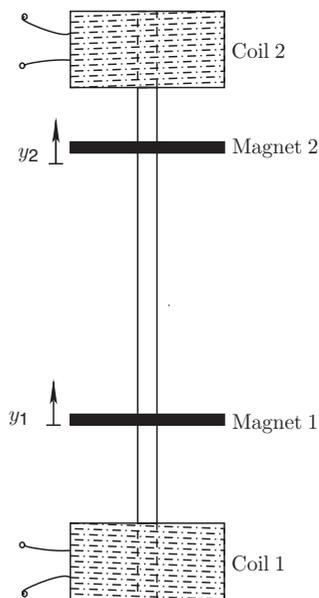


Figure 3. Schematic sketch of the magnetic levitation system.

Let m_1 be the mass of the lower magnet, v_1 the viscous friction coefficient between this magnet and air, and g the acceleration of gravity. The equation that describes the movement of the lower magnet (Magnet 1) is given by

$$m_1 \ddot{y}_1 + v_1 \dot{y}_1 + F_{m12} = F_{u11} - F_{u21} - m_1 g,$$

where y_1 is the position of Magnet 1, F_{u11} is the magnetic force from Coil 1 interacting with Magnet 1, F_{u21} is the magnetic force from Coil 2 interacting with Magnet 1, and F_{m12} is the mutual magnetic force between the two magnets. These forces are described by the following non-linear equations:

$$F_{m12} = \frac{c}{(y_c + y_2 - y_1 + d)^4}$$

$$F_{u11} = \frac{i_1}{a(k_s y_1 + b)^4}$$

$$F_{u21} = \frac{i_2}{a(y_c + k_s y_1 + b)^4},$$

where

- y_2 : position of Magnet 2;
- y_c : distance between Coils 1 and 2;
- i_1 : current through Coil 1;
- i_2 : current through Coil 2;
- a, b, c, d : real constants.

In §5.1, some experimental results for modeling a laboratory-scale plant of the magnetic levitation system just described are presented.

5.1 Simulation results

Experimental data have been acquired by keeping constant the DC current applied to Coil 1 of the plant while varying current through Coil 2. The input signal u (current through Coil 2) has been designed as a sequence of steps with different amplitudes so as to excite different modes of the system. The measured output signal y has been taken as the position of Magnet 1 (y_1). Figure 4 shows the input and output data available for estimation of the model. Before estimation, these data were sampled using a sampling period of 0.017 seconds, and normalized within the interval $[-1, 1]$ in order to avoid numerical problems. Another similar yet independent set of data has also been acquired and reserved for further model validation.

The model to be used relates the input u and output y in figure 4 by means of a second-order ($N=2$) Volterra

representation, as usual in the literature (Billings 1980, Dumont and Fu 1993)

$$\hat{y}(k) = \sum_{\tau_1=0}^{\epsilon} h_1(\tau_1)u(k - \tau_1) + \sum_{\tau_1=0}^{\epsilon} \sum_{\tau_2=0}^{\epsilon} h_2(\tau_1, \tau_2)u(k - \tau_1)u(k - \tau_2) \quad (32)$$

in which kernels h_1 and h_2 are to be expanded using a truncated Kautz basis $\{\psi_{l,n}\}$ according to equation (4).

In model (32), only the input and output signals are known. In this case, no prior information on the model kernels is available; therefore it is necessary to estimate them from input–output data measured from the system in order to obtain the optimal Kautz poles. The underlying idea adopted here consists of recovering the Volterra kernels from a numerical least-squares estimation of the Kautz coefficients followed by the computation of the corresponding optimal poles. This procedure continues until either the difference between kernels in successive steps is less than a given threshold or a maximum number of steps is exceeded.

The procedure thus begins with an initial Kautz pole for each direction $l=1, \dots, \eta$ of each kernel $h_{\eta}(k_1, \dots, k_{\eta})$. Then, the coefficients $\alpha_{(.)}$ in (5) are estimated by using available input–output data and a least-squares algorithm. Next the Volterra kernels are computed using equation (4). Finally, for each kernel direction: find the Kautz parameters b_l and c_l that provide the lowest kernel approximation error (NQE) by varying b_l over the interval $]-1, 1[$ and computing the respective optimal values for c_l by means of equation (23).

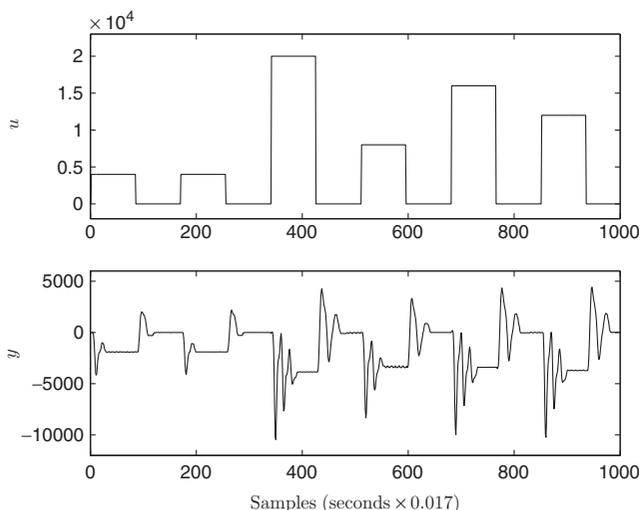


Figure 4. Above: current through Coil 2 (input signal u). Below: position of Magnet 1 (output signal y).

Based on the time response of the magnetic levitation system, long memory terms of the kernels – longer than 50 memory lags – are considered to be null. In other words, $\epsilon=50$ is adopted in equation (32). Moreover, $M=8$ is used for the number of Kautz functions. Choosing, for instance, the initial Kautz poles as $0.5 \pm i0.5$ for the first-order kernel as well as for the two directions of the second-order kernel, the corresponding initial values for the Kautz parameters, according to equations (9) and (10), are $b_1=b_2=0.6666$ and $c_1=c_2=-0.5$. Table 3 presents the values of the Kautz poles after the convergence of the above procedure, which took 17 steps. The adopted stop criteria was that the difference between the quadratic norms of the kernels in successive steps is less than 10^{-3} . Simulations with different initial poles have provided similar results.

By using the values of the Kautz poles shown in table 3, the kernels $h_1(k_1)$ and $h_2(k_1, k_2)$ can be calculated according to equation (4). They are illustrated in figures 5 and 6, respectively.

Figure 7 illustrates the model prediction of output of the system, $\hat{y}(k)$, with respect to another set of data (validation data), jointly with the actual (measured)

Table 3. Final values for the Kautz parameters after 17 steps: first and second order kernels.

Kernel order (η)	Axial direction of h_{η}	b_l	c_l	$(\beta_l, \bar{\beta}_l)$
1	$l=1$	0.9047	-0.6832	$0.7614 \pm i0.3216$
2	$l=1$	0.9323	-0.7604	$0.8206 \pm i0.2949$
	$l=2$	0.9434	-0.6521	$0.7793 \pm i0.2116$

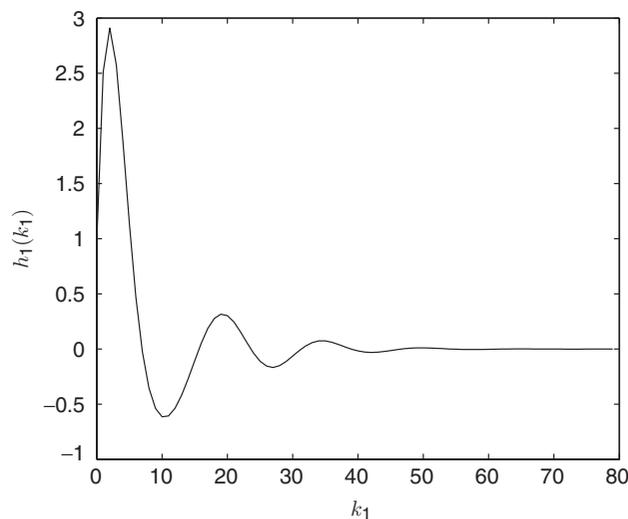


Figure 5. Estimated first-order kernel for model (32).

output value $y(k)$. It can be seen that the model matches the output of the system almost perfectly.

6. Conclusions

An analytical solution for the optimal selection of one of the Kautz parameters in Wiener/Volterra models has been derived. It is based on the decomposition of each multidimensional Volterra kernel using a set of independent orthonormal bases, each of which is parametrized by an individual pair of conjugate Kautz poles associated with the kernel dynamics along a particular dimension. The strategy adopted consists of adapting the original (Kautz) problem into a transformed (Laguerre) problem with known solution. Such a transformation makes it possible to set one of the

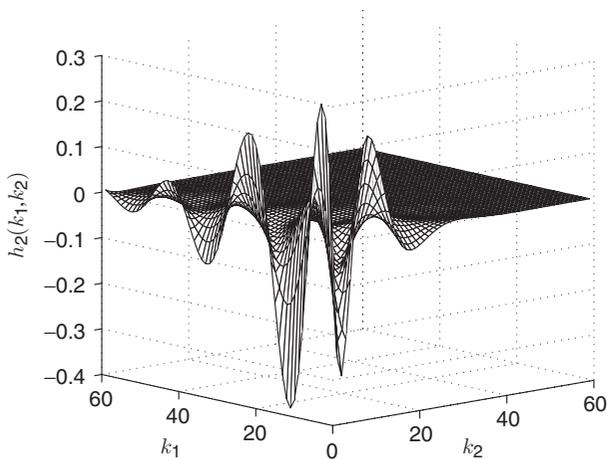


Figure 6. Estimated second-order kernel for model (32).

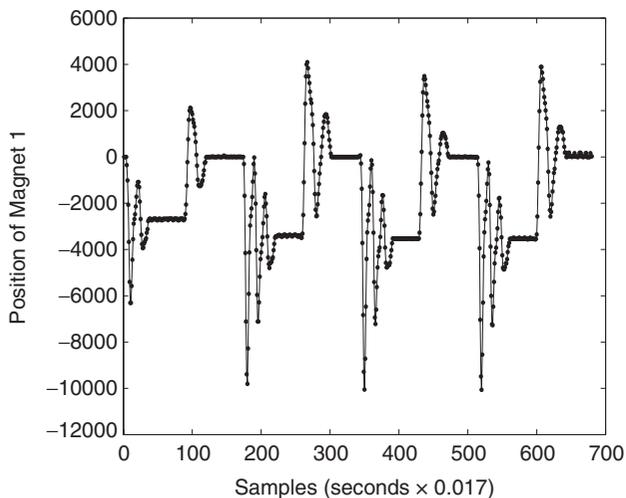


Figure 7. System (solid line) and model (dotted line) outputs: $y(k)$ and $\hat{y}(k)$.

Kautz parameters as constant in order to obtain the best choice for the other. The resulting solution, valid for models of any order, is based on the minimization of an upper bound for the error resulting from the truncated approximation of Volterra kernels using the two-parameter Kautz functions. It indirectly minimizes the number of functions associated with a given series truncation error.

In the proposed method, each multidimensional kernel is decomposed into independent orthonormal bases along its multiple dimensions. This approach provides a more general way for selecting the Kautz poles and represents an extension of a former work found in da Rosa *et al.* (2005, 2007), where an analytic solution was obtained for the particular case in which a single Kautz basis is used for expanding a given kernel along all its dimensions. It has been proven here that the particular and extended solutions are equivalent to each other when the Volterra kernels are symmetric. These results can be seen as a generalization, to the Kautz domain, of the work in Campello *et al.* (2006), where an analogous problem was investigated within the scope of Laguerre functions.

One illustrative example has shown that system identification with orthonormal basis functions is a very suitable framework for modeling non-linear systems when prior information about the system kernels is available. When compared to results found in previous work (da Rosa *et al.* 2007), it has been observed that the use of an independent basis for each kernel dimension reduces the truncation error when the kernel has different dominant dynamics along its multiple dimensions. Moreover, a practical application of the theoretical results has been made to the computation of the Kautz poles from input–output data measured from a non-linear magnetic levitation system with strong oscillatory behaviour. A very accurate second-order Volterra model, which represents almost perfectly the dynamics of the levitation system, has been obtained.

As stated in this paper, Laguerre and Kautz bases are preferable for modeling systems with dominant dynamics of first and second order, respectively. Systems having more complex dynamics are better represented using models based on generalized orthonormal bases, because the mathematical description of such bases involves multiple poles. For this reason, an important subject for future research concerns the extension of the results presented here to generalized orthonormal bases. Within the narrower scope of Kautz functions, an open problem that still deserves investigation is how to obtain an analytical solution to the simultaneous optimization of the two Kautz parameters.

Appendix A: Proof of Theorem 3.1

Consider a non-null function $x(k_1, \dots, k_\eta)$, which is null for $k_l < 0$ ($l=1, 2, \dots, \eta$). Suppose that x is absolutely summable on $[0, \infty)$, i.e.,

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

Now let $\alpha_{i_1, \dots, i_\eta}$ be the expansion coefficients of the η th kernel h_η decomposed into a set of independent Kautz bases, according to equation (4), and let each of the real constants c_l represent one of the Kautz parameters. The following relationship is valid for all $l=1, \dots, \eta$:

$$\begin{aligned} & -2c_l \sum_{k_1=1}^{\infty} \dots \sum_{k_l=0}^{\infty} \dots \sum_{k_\eta=0}^{\infty} k_l x(k_1, \dots, k_l - 1, \dots, k_\eta) \\ & \times x(k_1, \dots, k_l, \dots, k_\eta) \\ & + (1 + c_l^2) \sum_{k_1=0}^{\infty} \dots \sum_{k_l=0}^{\infty} \dots \sum_{k_\eta=0}^{\infty} k_l x^2(k_1, \dots, k_l, \dots, k_\eta) \\ & + \sum_{k_1=0}^{\infty} \dots \sum_{k_l=0}^{\infty} \dots \sum_{k_\eta=0}^{\infty} x^2(k_1, \dots, k_l, \dots, k_\eta) \\ & = (1 - c_l^2) \sum_{i_1=1}^{\infty} \dots \sum_{i_l=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} i_l \alpha_{i_1, \dots, i_\eta}^2 \end{aligned} \tag{A1}$$

The use of equations (18)–(20) allows rewriting (A1) for $l=1, \dots, \eta$ as

$$\begin{aligned} & \sum_{i_1=1}^{\infty} \dots \sum_{i_l=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} i_l \alpha_{i_1, \dots, i_\eta}^2 \\ & = \frac{-2c_l \mu_{1,l}(x) + (1 + c_l^2) \mu_{2,l}(x) + \eta \mu_{3,l}(x)}{1 - c_l^2} \end{aligned} \tag{A2}$$

and summing up equation (A2) with respect to l yields:

$$\begin{aligned} & \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} (i_1 + \dots + i_\eta) \alpha_{i_1, \dots, i_\eta}^2 \\ & = \sum_{l=1}^{\eta} \left[\frac{-2c_l \mu_{1,l}(x) + (1 + c_l^2) \mu_{2,l}(x) + \eta \mu_{3,l}(x)}{1 - c_l^2} \right]. \end{aligned} \tag{A3}$$

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

$$g_{\text{even},l}(k_1, \dots, k_\eta) = \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \gamma_{i_1, \dots, i_\eta} \prod_{l=1}^{\eta} \phi_{l,i_l}(k_l) \tag{A4}$$

$$g_{\text{odd},l}(k_1, \dots, k_\eta) = \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \sigma_{i_1, \dots, i_\eta} \prod_{l=1}^{\eta} \phi_{l,i_l}(k_l) \tag{A5}$$

By comparing equations (12) and (13) with (A4) and (A5), it can be verified that the coefficients $\alpha_{(\cdot)}$ of the expansion of kernel $h_\eta(k_1, \dots, k_\eta)$ using Kautz functions are related to the coefficients of the expansions above as $\gamma_{i_1, \dots, i_\eta} = \alpha_{2i_1, \dots, 2i_\eta}$ and $\sigma_{i_1, \dots, i_\eta} = \alpha_{2i_1-1, \dots, 2i_\eta-1}, \forall i_1, \dots, i_\eta$.

The upper bound given by (14) is thus deduced from the following inequality (Campello *et al.* 2004):

$$\begin{aligned} & \eta(M+1) \sum_{i_1=M+1}^{\infty} \dots \sum_{i_\eta=M+1}^{\infty} \alpha_{i_1, \dots, i_\eta}^2 \\ & \leq \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} (i_1 + \dots + i_\eta) \alpha_{i_1, \dots, i_\eta}^2 \\ & = \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} (2i_1 + \dots + 2i_\eta) \alpha_{2i_1, \dots, 2i_\eta}^2 \\ & \quad + \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} [(2i_1 - 1) + \dots + (2i_\eta - 1)] \alpha_{2i_1-1, \dots, 2i_\eta-1}^2 \\ & \leq 2 \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} (i_1 + \dots + i_\eta) \gamma_{i_1, \dots, i_\eta}^2 \\ & \quad + 2 \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} (i_1 + \dots + i_\eta) \sigma_{i_1, \dots, i_\eta}^2. \end{aligned} \tag{A6}$$

The (quadratic) approximation error of kernel $h_\eta(k_1, \dots, k_\eta)$, given by (11), appears on the left-hand side of equation (A6) multiplied by $\eta(M+1)\|h_\eta\|^2$. Then, dividing (A6) by $\eta(M+1)\|h_\eta\|^2$ and using (A3) results in:

$$\begin{aligned} \text{NQE} & \leq 2 \sum_{l=1}^{\eta} \left[\frac{-2c_l \mu_{1,l}(g_{\text{even},l}) + (1 + c_l^2) \mu_{2,l}(g_{\text{even},l}) + \mu_{3,l}(g_{\text{even},l})}{\eta(M+1)\|h_\eta\|^2(1 - c_l^2)} \right] \\ & \quad + 2 \sum_{l=1}^{\eta} \left[\frac{-2c_l \mu_{1,l}(g_{\text{odd},l}) + (1 + c_l^2) \mu_{2,l}(g_{\text{odd},l}) + \mu_{3,l}(g_{\text{odd},l})}{\eta(M+1)\|h_\eta\|^2(1 - c_l^2)} \right]. \end{aligned} \tag{A7}$$

Finally, equations (15), (16), and (17) as well as inequality (A7) yield

$$\text{NQE} \leq \frac{2}{\eta(M+1)\|h_\eta\|^2} \sum_{l=1}^{\eta} \left[\frac{m_{2,l} c_l^2 - 2m_{1,l} c_l + m_{3,l}}{1 - c_l^2} \right]$$

which completes the proof.

Appendix B: Proof of Theorem 3

Using (3) for rewriting equation (12) yields

$$\begin{aligned}
 &g_{\text{even},l}(k_1, \dots, k_\eta) \\
 &= \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1, \dots, 2i_\eta} \prod_{l=1}^{\eta} \phi_{l,i_l}(k_l) \\
 &= \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \left(\sum_{\tau_1=0}^{\infty} \dots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \dots, \tau_\eta) \prod_{l=1}^{\eta} \psi_{l,2i_l}(\tau_l) \right) \\
 &\quad \times \prod_{l=1}^{\eta} \phi_{l,i_l}(k_l) \\
 &= \sum_{\eta=0}^{\infty} \dots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \dots, \tau_\eta) \\
 &\quad \times \left(\sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \prod_{l=1}^{\eta} \psi_{l,2i_l}(\tau_l) \phi_{l,i_l}(k_l) \right). \tag{B1}
 \end{aligned}$$

Then, taking the Z-transform with respect to τ_l (for $l = 1, 2, \dots, \eta$) of the term between parentheses above results in:

$$\begin{aligned}
 &\mathcal{Z} \left[\sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \prod_{l=1}^{\eta} \psi_{l,2i_l}(\tau_l) \phi_{l,i_l}(k_l) \right] \\
 &= \sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \left\{ \prod_{l=1}^{\eta} \left[\frac{\sqrt{(1-c_l^2)(1-b_l^2)z_l}}{z_l^2 + b_l(c_l-1)z_l - c_l} \right. \right. \\
 &\quad \left. \left. \times \left(\frac{-c_l z_l^2 + b_l(c_l-1)z_l + 1}{z_l^2 + b_l(c_l-1)z_l - c_l} \right)^{i_l-1} \phi_{l,i_l}(k_l) \right] \right\} \\
 &= \left[\prod_{l=1}^{\eta} \frac{\sqrt{(1-c_l^2)(1-b_l^2)z_l}}{z_l^2 + b_l(c_l-1)z_l - c_l} \right] \\
 &\quad \times \left[\sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \prod_{l=1}^{\eta} \phi_{l,i_l}(k_l) w_l^{1-i_l} \right], \tag{B2}
 \end{aligned}$$

where the following simplification has been made

$$w_l \triangleq \frac{z_l^2 + b_l(c_l-1)z_l - c_l}{-c_l z_l^2 + b_l(c_l-1)z_l + 1} \quad l = 1, \dots, \eta. \tag{B3}$$

The last term in (B2) can be rewritten as

$$\begin{aligned}
 &\sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \prod_{l=1}^{\eta} \phi_{l,i_l}(k_l) w_l^{1-i_l} \\
 &= \prod_{l=1}^{\eta} \left(w_l \sum_{i_l=1}^{\infty} \phi_{l,i_l}(k_l) w_l^{-i_l} \right) \\
 &= \prod_{l=1}^{\eta} \left[\frac{w_l \sqrt{1-c_l^2}}{w_l + c_l} \left(\frac{1+c_l w_l}{w_l + c_l} \right)^{k_l} \right] \tag{B4}
 \end{aligned}$$

after having used the following valid relationship for the Laguerre functions $\phi_{l,n}(k)$ (Tanguy *et al.* 2002):

$$\sum_{n=1}^{\infty} \phi_{l,n}(k) w_l^{-n} = \frac{\sqrt{1-c_l^2}}{w_l + c_l} \left(\frac{1+c_l w_l}{w_l + c_l} \right)^k.$$

Substituting (B3) into (B4) and using the resulting equation to rewrite (B2) gives the following (after certain algebraic manipulations):

$$\begin{aligned}
 &\mathcal{Z} \left[\sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \prod_{l=1}^{\eta} \psi_{l,2i_l}(\tau_l) \phi_{l,i_l}(k_l) \right] \\
 &= \prod_{l=1}^{\eta} \left[\frac{\sqrt{1-b_l^2}}{z_l - b_l} \left(\frac{1-b_l z_l}{z_l - b_l} \right)^{k_l} z_l^{-k_l} \right] \\
 &= \prod_{l=1}^{\eta} \mathcal{Z} \{ [\psi_{l,2(k_l+1)}(\tau)]_{c_l=0} \} \\
 &= \mathcal{Z} \left\{ \prod_{l=1}^{\eta} \{ [\psi_{l,2(k_l+1)}(\tau)]_{c_l=0} \} \right\}. \tag{B5}
 \end{aligned}$$

Equation (B5) is thus rewritten as:

$$\sum_{i_1=1}^{\infty} \dots \sum_{i_\eta=1}^{\infty} \prod_{l=1}^{\eta} \psi_{l,2i_l}(\tau_l) \phi_{l,i_l}(k_l) = \prod_{l=1}^{\eta} \hat{\psi}_{l,2(k_l+1)}(\tau_l), \tag{B6}$$

where $\hat{\psi}_{l,n}(\eta) \triangleq [\psi_{l,n}(\tau)]_{c_l=0}$. Substituting (B6) into (B1) completes the proof.

The proof for $g_{\text{odd},l}(k_1, \dots, k_\eta)$ is analogous.

Acknowledgments

The authors would like to acknowledge the financial support from the Brazilian National Council for Scientific and Technological Development (CNPq) and from the Research Foundation of the State of São Paulo (Fapesp).

References

M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 2nd ed., John Wiley & Sons, 1993.
 S.A. Billings, "Identification of nonlinear systems – a survey", *IEE Proc. Pt D*, 127, pp. 272–285, 1980.
 J. Bokor and F. Schipp, "Approximate identification in Laguerre and Kautz bases", *Automatica*, 34, pp. 463–468, 1998.
 S. Boyd and L.O. Chua, "Fading memory and the problem of approximation nonlinear operators with Volterra series", *IEEE Trans. on Circuits and Systems*, 32, pp. 1150–1161, 1985.
 P.W. Broome, "Discrete orthonormal sequences", *Journal of the Association for Comp. Machinery*, 12, pp. 151–168, 1965.
 R.J.G.B. Campello, W.C. Amaral and G. Favier, "Optimal Laguerre series expansion of discrete Volterra models", in *Proc. European Control Conference*, Porto, Portugal, pp. 372–377, 2001.
 R.J.G.B. Campello, G. Favier and W.C. Amaral, "Optimal expansions of discrete-time Volterra models using Laguerre functions",

- in *Proc. of the 13th IFAC Symposium on System Identification*, Rotterdam, The Netherlands, pp. 18454–1849, 2003.
- R.J.G.B. Campello, G. Favier and W.C. Amaral, “Optimal expansions of discrete-time Volterra models using Laguerre functions”, *Automatica*, 40, pp. 815–822, 2004.
- R.J.G.B. Campello, W.C. Amaral and G. Favier, “A note on the optimal expansion of Volterra models using Laguerre functions”, *Automatica*, 42, pp. 689–693, 2006.
- G.J. Clowes, “Choice of time-scaling factor for linear system approximation using orthonormal Laguerre functions”, *IEEE Trans. on Auto. Cont.*, 10, pp. 487–489, 1965.
- A. da Rosa, W.C. Amaral and R.J.G.B. Campello, “Choice of free parameters in expansions of discrete-time Volterra models using Kautz functions”, in *Proc. 16th IFAC World Congress*, Prague, Czech Republic, 2005.
- A. da Rosa, R.J.G.B. Campello and W.C. Amaral, “Choice of free parameters in expansions of discrete-time Volterra models using Kautz functions”, *Automatica*, 43, pp. 1084–1091, 2007.
- A.C. den Brinker, F.P.A. Benders and T.O. Silva, “Optimality conditions for truncated Kautz series”, *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 43, pp. 117–122, 1996.
- P.M.J.V. den Hof, P.S.C. Heuberger and J. Bokor, “System identification with generalized orthonormal basis functions”, *Automatica*, 31, pp. 1821–1834, 1995.
- F.J. Doyle III, R.K. Pearson and B.A. Ogunnaike, *Identification and Control Using Volterra Models*, Berlin: Springer-Verlag, 2002.
- G.A. Dumont and Y. Fu, “Non-linear adaptive control via Laguerre expansion of Volterra Kernels”, *Int. J. Adaptive Control and Signal Processing*, 7, pp. 367–382, 1993.
- ECP, *Manual for Model 730 – Magnetic Levitation System*, California, USA: Educational Control Products, 1999.
- P. Eykhoff, *System Identification: Parameter and State Estimation*, John Wiley & Sons, 1974.
- Y. Fu and G.A. Dumont, “An optimum time scale for discrete Laguerre network”, *IEEE Trans. on Autom. Cont.*, 38, pp. 934–938, 1993.
- P.S.C. Heuberger, P.M.J.V. den Hof and O.H. Bosgra, “A generalized orthonormal basis for linear dynamical systems”, *IEEE Transactions on Automatic Control*, 40, pp. 451–465, 1995.
- P.S.C. Heuberger, P.M.J.V. den Hof and B. Wahlberg, *Modelling and Identification with Rational Orthogonal Basis Functions*, London: Springer, 2005.
- W.H. Kautz, “Transient synthesis in time domains”, *IRE Transactions on Circuit Theory*, 1, pp. 29–39, 1954.
- A.Y. Kibangou, G. Favier and M. Hassani, “Iterative optimization method of GOB-Volterra filters”, in *Proc. 16th IFAC World Congress*, Prague, Czech Republic, 2005a.
- A.Y. Kibangou, G. Favier and M. Hassani, “Selection of generalized orthonormal bases for second-order Volterra filters”, *Signal Processing*, 85, pp. 2371–2385, 2005b.
- M.A. Masnadi-Shirazi and N. Ahmed, “Optimum Laguerre networks for a class of discrete-time systems”, *IEEE Trans. on Signal Proc.*, 39, pp. 2104–2108, 1991.
- V. Nalbantoğlu, J. Bokor, G. Balas and P. Gaspar, “System identification with generalized orthonormal basis functions: an application to flexible structures”, *Control Engineering Practice*, 11, pp. 245–259, 2003.
- O. Nelles, *Nonlinear System Identification*, Berlin: Springer-Verlag, 2001.
- B. Ninness and F. Gustafsson, “A unifying construction of orthonormal bases for system identification”, *IEEE Transactions on Autom. Cont.*, 42, pp. 515–521, 1997.
- B. Ninness, H. Hjalmarsson and F. Gustafsson, “The fundamental role of general orthonormal bases in system identification”, *IEEE Trans. on Autom. Cont.*, 44, pp. 1384–1406, 1999.
- G.H.C. Oliveira, W.C. Amaral, G. Favier and G.A. Dumont, “Constrained robust predictive controller for uncertain processes modeled by orthonormal series functions”, *Automatica*, 36, pp. 563–571, 2000.
- S.C. Patwardhan and S.L. Shah, “From data to diagnosis and control using generalized orthonormal basis filters. Part I: Development of state observers”, *Journal of Process Control*, 15, pp. 819–835, 2005.
- W.J. Rugh, *Nonlinear System Theory – The Volterra/Wiener Approach*, The Johns Hopkins University Press, 1981.
- B.E. Sarroukh, S.J.L. van Eijndhoven and A.C. den Brinker, “An iterative solution for the optimal poles in a Kautz series”, in *Proc. IEEE Intl. Conf. Acoustics, Speech and Signal Processing*, Liscataway, New Jersey, USA, 2001, pp. 3949–3952.
- M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, Robert Krieger Publishing Company, 1989.
- T.O. Silva, “Optimality conditions for truncated Laguerre networks”, *IEEE Transactions on Signal Processing*, 42, pp. 2528–2530, 1994.
- T.O. Silva, “Optimality conditions for truncated Kautz networks with two periodically repeating complex conjugate poles”, *IEEE Trans. on Autom. Cont.*, 40, pp. 342–346, 1995.
- T.O. Silva, “Stationary conditions for the L^2 error surface of the generalized orthonormal basis functions lattice filter”, *Signal Processing*, 56, pp. 233–253, 1997.
- S. Takenaka, “On the orthogonal functions and a new formula of interpolation”, *Japanese Journal of Mathematics*, II, 1925.
- N. Tanguy, R. Morvan, P. Vilbé and L.C. Calvez, “Optimum choice of free parameter in orthonormal approximations”, *IEEE Transactions on Automatic Control*, 40, pp. 1811–1813, 1995.
- N. Tanguy, R. Morvan, P. Vilbé and L.C. Calvez, “Pertinent parameters for Kautz approximation”, *Electronic Letters*, 36, pp. 769–771, 2000.
- N. Tanguy, R. Morvan, P. Vilbé and L.C. Calvez, “Pertinent choice of parameters for discrete Kautz approximation”, *IEEE Transactions on Automatic Control*, 47, pp. 783–787, 2002.
- B. Wahlberg, “System identification using Kautz models”, *IEEE Transactions on Automatic Control*, 39, pp. 1276–1282, 1994.
- B. Wahlberg and P.M. Mäkilä, “Approximation of stable linear dynamical systems using Laguerre and Kautz functions”, *Automatica*, 32, pp. 693–708, 1996.
- N. Wiener, *Nonlinear Problems in Random Theory*, New York: The MIT Press, 1966.
- K. Ziaei and D.W.L. Wang, “Application of orthonormal basis functions for identification of flexible-link manipulators”, *Control Engineering Practice*, 14, pp. 99–106, 2006.