CHOICE OF FREE PARAMETERS IN
EXPANSIONS OF DISCRETE-TIME
VOLterra MODELS USING KAUTZ
FUNCTIONS

Alex da Rosa† Wagner C. Amaral†
Ricardo J. G. B. Campello*†

†DCA/FEEC/UNICAMP, 13083-852, Campinas-SP, Brazil
*COPOP/UNISANTOS, 11070-906, Santos-SP, Brazil

Abstract: The present paper involves the approximation of nonlinear systems using Wiener/Volterra models with Kautz orthonormal functions. It focuses on the problem of selecting the free complex pole which characterizes these functions. The problem is solved by minimizing an upper bound of the error arising from the truncated approximation of Volterra kernels using Kautz functions. An analytical solution for the optimal choice of one of the parameters related to the Kautz pole is thus obtained, with the results valid for any-order Wiener/Volterra models. An example illustrates the application of the mathematical results derived.

Copyright © 2005 IFAC

Keywords: Nonlinear Systems, Volterra Series, Kautz Functions, Optimization.

1. INTRODUCTION

Discrete-time Volterra models relate the output $y(k)$ of a physical process to its input $u(k)$ as (Schetzen, 1989; Rugh, 1991):

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

(1)

where the functions $h_{\eta}(\tau_1, \ldots, \tau_{\eta})$ are the $\eta$-order Volterra kernels. Equation (1) is a generalization of the impulse response model (Eykhoff, 1974), traditionally used for the analysis of linear systems.

Wiener/Volterra models mathematically express the kernels $h_{\eta}$ as an expansion using an orthonormal function basis $\{ \psi_n \}$:

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

(2)

which assumes that the kernels are quadratically summable in $[0, \infty)$, i.e.

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

Expansion (2) can equivalently be defined in the $z$-domain as:

$$H_{\eta}(z_1, \ldots, z_{\eta}) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_{\eta}=1}^{\infty} \alpha_{i_1, \ldots, i_{\eta}} \prod_{j=1}^{\eta} \Psi_{i_j}(z_j)$$

(3)
where $\Psi_{k_j}(z_j)$ is the one-sided Z-transform of
the sequence $\psi_{k_j}(k_j)$, i.e. $\Psi_{k_j}(z_j) = \mathcal{Z}[\psi_{k_j}(k_j)]$.
Given the orthonormality property of the set
$\{\psi_n\}$, the coefficients $\alpha_{i_1,\ldots,i_n}$ can be computed from the following:

$$
\alpha_{i_1,\ldots,i_n} = \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} h_\eta(k_1,\ldots,k_n) \prod_{j=1}^n \psi_{i_j}(k_j)
$$

(4)

Two of these sets are widely used in the approxi-
mation, modelling and identification of systems:
Laguerre and Kautz functions (Broome, 1965),
which are characterized by real and complex poles,
respectively.

The use of Wiener/Volterra models with Laguerre
and Kautz functions has been widespread in stud-
ies involving the identification and control of dy-
namic processes (Tanguy et al., 2002; Wahlberg
et al., 1996; Wahlberg, 1994). These models require
a reduced number of terms in order to represent a
given system by means of a truncated orthonormal
approximation of equation (1). Since the param-
eters of the orthonormal functions are free within a
stability region, their selection has been the focus
of many studies.

One of the first papers aimed at solving the
above-mentioned parameter selection problem
(Masnadi-Shirazi et al., 1991) investigated the op-
timal approximation of a particular class of linear
systems using Laguerre functions. Later, an ana-
sylitical solution for the optimal choice of the La-
guerre pole in the case of stable linear systems was
achieved by minimizing a quadratic cost function
(Fu et al., 1993). More recently, Campello et al.
extended this solution to any-order Volterra mod-
els (Campello et al., 2003; Campello et al., 2004).
In all of these papers an analytical solution de-
ending only on the system kernels was derived.

However, poorly damped dynamics are difficult
to be approximated with a small number of La-
guerre functions. This has led to an interest in
Kautz functions, which can better approximate
systems with oscillatory behavior. A sub-optimal
analytical solution for the choice of the Kautz pole
in the representation of continuous and discrete
linear systems has been presented in (Tanguy et al.,
2000) and (Tanguy et al., 2002), respectively.
In the present paper the results of this last cited
work have been extended to any-order Volterra
models in such a way that an analytical solution
for the optimal selection of one of the parameters
related to the Kautz pole is obtained.

This paper is organized as follows. In the next
section, Kautz functions are presented in the con-
text of Wiener/Volterra models. In Section 3
the optimization problem for the selection of the
Kautz pole is discussed, and a theorem facilitating
computational implementation of the models is
formulated. In Section 4 an illustrative example is
presented, and Section 5 adresses the conclusions.

2. EXPANSION OF VOLterra MODELS
USING Kautz FUNCTIONS

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

$$
\tilde{h}_\eta(k_1,\ldots,k_n) = \sum_{i_1=1}^{M} \cdots \sum_{i_n=1}^{M} \alpha_{i_1,\ldots,i_n} \prod_{j=1}^n \psi_{i_j}(k_j)
$$

(5)

Defining the norm $\|h_\eta\|$ as:

$$
\|h_\eta\|^2 = \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} h^2_\eta(k_1,\ldots,k_n)
$$

and using equations (2) and (5) makes it possible
to deduce that the Normalized Quadratic Error
(NQE) of the approximation of the kernel $h_\eta$,
defined as $NQE = \frac{\|h_\eta - \tilde{h}_\eta\|^2}{\|h_\eta\|^2}$, can be written as follows (Da Rosa, 2004):

$$
NQE = \frac{\sum_{i_1=M+1}^{\infty} \cdots \sum_{i_n=M+1}^{\infty} \alpha_{i_1,\ldots,i_n}^2}{\sum_{i_1=1}^{\infty} \cdots \sum_{i_n=1}^{\infty} \alpha_{i_1,\ldots,i_n}^2}
$$

(6)

where $\alpha_{i_1,\ldots,i_n}$ are the coefficients of the expansion
$h_\eta(\{k_1,\ldots,k_n\})$ according to equation (4).

When the set of orthonormal functions is the
Laguerre basis, such functions are defined in the
$z$-domain as follows (Wahlberg, 1994):

$$
\Phi_n(z) = \sqrt{1-c^2} \frac{z^{n-1}}{(z-c)^n}
$$

(7)

where $n = 1, 2, \ldots$ and $c$ $(|c| < 1)$ is the Laguerre
(real) pole.

Kautz functions, which constitute a second-order
generalization of (7), are defined as follows
(Wahlberg, 1994):

$$
\Psi_{2n}(z) = \sqrt{1-c^2} \frac{z^{n-1}}{2^2 + b(c-1)z - c^2 + b(c-1)z - c}
$$

$$
\Psi_{2n-1}(z) = \sqrt{1-c^2} \frac{z^{n-1}}{2^2 + b(c-1)z - c^2 + b(c-1)z - c}
$$

(8)

where $b$ and $c$ are real constants that characterize
functions $\Psi_{\{,\}(z)}$. They are related to the Kautz
pole $\beta$ through $b = (\beta + \bar{\beta})/(1 + \beta \bar{\beta})$ and $c = -\beta \bar{\beta}$. Parameters $b$ and $c$ are restricted to $|b| < 1$ and $|c| < 1$. The Kautz basis is complete in the Lebesgue space $\ell^2(0, \infty)$ for $|\beta| < 1$, so any finite energy signal (including quadratically summable kernels) can be approximated with any prescribed accuracy by truncating the infinite expansion.

In the next section, the solution of the problem of the sub-optimal choice for the Kautz pole based on the minimization of an upper bound for the approximation error in equation (6) is presented.

3. SUB-OPTIMAL SELECTION OF THE KAUTZ POLE

Kautz functions, defined in equation (8), depend upon two real parameters ($b$ and $c$); the selection of these parameters has a direct influence on the computation of the coefficients $\alpha_{i_1, \ldots, i_q}$ in (4). The optimal selection of both at the same time is still under investigation. It is possible, however, to set one of them as constant in order to obtain the best choice for the other. The solution to be presented here consists of the adaptation of the original (Kautz) problem into a transformed (Laguerre) problem with known solution. It also considers to set $b$ as constant in order to optimize $c$. Details are given below.

3.1 Sub-Optimal Expansion of the $\eta$-Order Volterra Kernel

Let $\Phi_n$ be the Laguerre functions and $\alpha_{i_1, \ldots, i_q}$ be the coefficients of the expansion of the kernel $h_\eta(k_1, \ldots, k_q)$ using Kautz functions. Now define the following functions:

$$G_{\text{even}}(z_1, \ldots, z_\eta) \triangleq \sum_{i_1=1}^{\infty} \cdots \sum_{i_q=1}^{\infty} \alpha_{2i_1, \ldots, 2i_q} \cdot \prod_{j=1}^{\eta} \Phi_{i_j}(z_j)$$

(9)

$$G_{\text{odd}}(z_1, \ldots, z_\eta) \triangleq \sum_{i_1=1}^{\infty} \cdots \sum_{i_q=1}^{\infty} \alpha_{2i_1-1, \ldots, 2i_q-1} \cdot \prod_{j=1}^{\eta} \Phi_{i_j}(z_j)$$

(10)

From equation (6) and the inequality

$$\sum_{i_1=1}^{\infty} \cdots \sum_{i_q=1}^{\infty} (i_1 + \cdots + i_q) \alpha_{i_1, \ldots, i_q} \geq \eta(M + 1) \sum_{i_1=M+1}^{\infty} \cdots \sum_{i_q=M+1}^{\infty} \alpha_{i_1, \ldots, i_q}$$

it is possible to show\, that:

$$\text{NQE} \leq L(c) = \frac{2(m_2c^2 - 2m_1c + m_3)}{\eta(M+1){\|h_\eta\|^2}(1-c^2)}$$

(11)

where the terms $m_p$ ($p = 1, 2, 3$) are defined as:

$$m_1 = \mu_1(g_{\text{even}}) + \mu_1(g_{\text{odd}})$$

(12)

$$m_2 = \mu_2(g_{\text{even}}) + \mu_2(g_{\text{odd}})$$

(13)

$$m_3 = \mu_2(g_{\text{even}}) + \mu_2(g_{\text{odd}}) + \eta \mu_3(g_{\text{even}}) + \eta \mu_3(g_{\text{odd}})$$

(14)

with $g_{\text{even}}(k_1, \ldots, k_q)$ and $g_{\text{odd}}(k_1, \ldots, k_q)$ denoting the inverse Z-transforms of $G_{\text{even}}(z_1, \ldots, z_\eta)$ and $G_{\text{odd}}(z_1, \ldots, z_\eta)$ in (9) and (10), respectively. The moments $\mu_1(x)$, $\mu_2(x)$ e $\mu_3(x)$ are given by the following equations:

$$\mu_1(x) = \sum_{l=1}^{\eta} \sum_{k_1=0}^{\infty} \cdots \sum_{k_1=0}^{\infty} \sum_{k_q=0}^{\infty} k_l \cdot x(k_1, \ldots, k_1, k_q)x(k_1, \ldots, k_1, k_q - 1, \ldots, k_q)$$

(15)

$$\mu_2(x) = \sum_{l=1}^{\eta} \sum_{k_1=0}^{\infty} \cdots \sum_{k_1=0}^{\infty} \sum_{k_q=0}^{\infty} k_l \cdot x^2(k_1, \ldots, k_1, k_q)$$

(16)

$$\mu_3(x) = \sum_{k_1=0}^{k_q=0} \cdots \sum_{k_q=0}^{k_q=0} x^2(k_1, \ldots, k_q)$$

(17)

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

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

(18)

Since $\|h_\eta\|$ is a (non-null) constant for a given system, and $L(c)$ is a pseudo-convex function for $|c| < 1$, as shown in (Da Rosa, 2004), the necessary and sufficient condition for solving (18) is:

$$\frac{\partial L(c)}{\partial c} = \frac{2[(2m_2c - 2m_1)(1-c^2) + 2c(m_2c^2 - 2m_1c + m_3)]}{\eta(M+1){\|h_\eta\|^2}(1-c^2)^2} = 0$$

The previous equation is satisfied if and only if

1 The proof follows the lines of (Tanguy et al., 2002) generalized to $\eta$-dimensional moments. Further details can be found in (Da Rosa, 2004).
\[ m_1c^2 - (m_2 + m_3)c + m_1 = 0 \]  

Then, defining \( \xi = (m_2 + m_3)/(2m_1) \), the solution of (19) is given by:

\[
c_{\text{opt}} = \begin{cases} 
\xi - \sqrt{\xi^2 - 1} & \text{if } \xi > 1 \\
\xi + \sqrt{\xi^2 - 1} & \text{if } \xi < -1 
\end{cases}
\]

which is also the strict global optimal solution for (18). It can be shown that the condition \( \xi = 1 \) is unfeasible since kernel \( h_\eta(k_1, \ldots, k_\eta) \) is assumed to be summable.

Equation (20) is an analytical solution for the selection of parameter \( c \) of the Kautz functions according to criterion (18). It can be used, after setting the value of parameter \( b \), in order to minimize the upper bound \( L(c) \) for the squared norm of the error resulting from the truncated expansion of the Volterra kernels.

### 3.2 Theorem

The computation of functions \( G_{\text{even}} \) and \( G_{\text{odd}} \) can more easily be performed in time-domain. The following theorem is useful for such computations when only experimental data is available or if the system’s kernels are not known analytically.

**Theorem 1.** Functions \( G_{\text{even}}(z_1, \ldots, z_\eta) \) and \( G_{\text{odd}}(z_1, \ldots, z_\eta) \), defined in equations (9) and (10), respectively, are written in time-domain as:

\[
g_{\text{even}}(k_1, \ldots, k_\eta) = Z^{-1}[G_{\text{even}}(z_1, \ldots, z_\eta)] \\
= \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \ldots, \tau_\eta) \cdot \prod_{j=1}^{\eta} \hat{\psi}_{2(j+1)}(\tau_j) \tag{21}
\]

\[
g_{\text{odd}}(k_1, \ldots, k_\eta) = Z^{-1}[G_{\text{odd}}(z_1, \ldots, z_\eta)] \\
= \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} h_\eta(\tau_1, \ldots, \tau_\eta) \cdot \prod_{j=1}^{\eta} \hat{\psi}_{2(j+1)-1}(\tau_j) \tag{22}
\]

where \( \hat{\psi}_{(j)}(\tau) \) denotes the Kautz functions in time-domain with \( c = 0 \) and \( h_\eta(\tau_1, \ldots, \tau_\eta) \) is the \( \eta \)-order Volterra kernel.

**Proof:** The proof is in Appendix A.

### 4. Illustrative Example

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

\[
h_2(k_1, k_2) = (k_1 - 2k_2) \exp(-\rho_1 k_1 - \rho_2 k_2) \cdot \cos(\omega_1 k_1 + \omega_2 k_2) \tag{23}
\]

for \( k_1 \geq 0 \) and \( k_2 \geq 0 \). For negative values of \( k_1 \) or \( k_2 \), \( h_2(k_1, k_2) \) is assumed to be zero (causal system). The real constant \( \rho_i (i = 1, 2) \) can be seen as the decay rate of the kernel (23) along the \( i \)-th axis, whereas \( \omega_i \) is the frequency with which the kernel oscillates in that direction.

Figure 1 illustrates kernel (23) for \( \rho_1 = 0.45, \rho_2 = 0.7, \omega_1 = 100 \) and \( \omega_2 = 1 \).

![Fig. 1. Second-order kernel \( h_2(k_1, k_2) \)](image)

The choice of \( b = 0.4 \) results in \( c_{\text{opt}} = -0.20833 \), computed via (20). For \( (b, c) = (0.4, -0.20833) \) the normalized quadratic error associated with the approximation of \( h_2 \), computed using (6), is shown in Table 1 for \( M = 2, 4, 6 \) Kautz functions.

**Table 1.** Approximation error of the kernel in (23) based on the number of Kautz functions used in its expansion

<table>
<thead>
<tr>
<th>Number of functions (M)</th>
<th>NQE</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.73525</td>
</tr>
<tr>
<td>4</td>
<td>0.28299</td>
</tr>
<tr>
<td>6</td>
<td>0.05877</td>
</tr>
</tbody>
</table>

By varying parameter \( b \) in the interval \((-1, 1)\), each pair \((b, c_{\text{opt}})\) gives an approximation error. The value of \( c \) providing the best approximation is that for which NQE is the lowest. It is obtained choosing \( b = 0.593 \), which results in \( c_{\text{opt}} = -0.2594 \). For \( (b, c_{\text{opt}}) = (0.593, -0.2594) \), the Kautz pole is \( \beta = 0.37341 \pm 0.34636 \). Equation (5) with \( M = 6 \) gives the corresponding approximation of kernel (23), as illustrated in Figure 2. The error associated with this approximation is given by the difference between the surfaces...
in Figures 1 and 2, shown in Figure 3. With \((b, c_{\text{opt}}) = (0.593, -0.2594), \) NQE = \(6.621 \cdot 10^{-3}\).

Fig. 2. Approximation of kernel \(h_2(k_1, k_2)\) with \(b = 0.593\) and \(c = -0.2594\)

5. CONCLUSIONS

An analytical solution for the optimal selection of one of the Kautz parameters has been obtained, which is based on the minimization of an upper bound of the error in the approximation of Volterra kernels using Kautz functions. The use of this solution requires setting one of the parameters as constant, since a formula for the optimization of both has yet to be devised. Simulation results have shown that the method proposed provides a good approximation of non-linear systems with oscillatory behavior. More illustrative examples of these results can be found in (Da Rosa, 2004). In future works, the authors intend to extend the results found in this paper with respect to generalized orthonormal functions (Van den Hof et al., 1995).

Appendix A. PROOF OF THEOREM 1

Writing equation (9) in time-domain yields:

\[
g_{\text{even}}(k_1, \ldots, k_\eta) = \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \alpha_{2i_1, \ldots, 2i_\eta} \cdot \prod_{j=1}^{\eta} \Phi_{ij}(k_j)
\]

\[
= \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} \left( \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} h_{\eta}(\tau_1, \ldots, \tau_\eta) \cdot \prod_{j=1}^{\eta} \Phi_{ij}(k_j) \right)
\]

\[
= \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_\eta=0}^{\infty} \left( \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \prod_{j=1}^{\eta} \psi_{2i_j}(\tau_j) \phi_{ij}(k_j) \right)
\]  
(A.1)

Then, taking the Z-transform (with respect to \(\tau_j\)) of the term between parentheses above results in:

\[
Z \left[ \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \prod_{j=1}^{\eta} \psi_{2i_j}(\tau_j) \phi_{ij}(k_j) \right] =
\]

\[
= \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \left[ \prod_{j=1}^{\eta} \left[ \frac{\sqrt{(1-c^2)(1-c^2)}}{z_j^2 + b(c-1)z_j - c} \right] \cdot \frac{i_j-1}{z_j^2 + b(c-1)z_j - c} \phi_{ij}(k_j) \right]
\]

\[
= \prod_{j=1}^{\eta} \frac{\sqrt{(1-c^2)(1-c^2)}}{z_j^2 + b(c-1)z_j - c} \cdot \sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \prod_{j=1}^{\eta} \phi_{ij}(k_j) w_j^{1-i_j}
\]  
(A.2)

where the following simplification has been made (for \(j = 1, 2, \ldots, \eta\)):

\[
w_j \triangleq \frac{z_j^2 + b(c-1)z_j - c}{-cz_j^2 + b(c-1)z_j + 1}
\]  
(A.3)

The last term in (A.2) can be rewritten as:

\[
\sum_{i_1=1}^{\infty} \cdots \sum_{i_\eta=1}^{\infty} \prod_{j=1}^{\eta} \phi_{ij}(k_j) w_j^{1-i_j} =
\]
by using the following valid relationship for the Laguerre functions $\phi_l(k)$ (Tanguy et al., 2002):

$$\sum_{l=1}^{\infty} \phi_l(k) w^{-l} = \frac{1 - c^2}{w + c} \left( \frac{1 + cw}{w + c} \right)^k$$

Substituting (A.3) into (A.4) and using the resulting equation to rewrite (A.2) gives the following (after certain algebraic manipulations):

$$Z \left[ \sum_{i_1=1}^{\infty} \cdots \sum_{i_n=1}^{\infty} \prod_{j=1}^{n} \psi_{2i_j}(\tau_j) \phi_{i_j}(k_j) \right] = \prod_{j=1}^{n} \frac{\sqrt{1 - b^2 z_j - b}}{z_j - b} \left( \frac{1 - b z_j}{z_j - b} \right)^{k_j} z_j^{-k_j} = \prod_{j=1}^{n} Z \left\{ \tilde{\psi}_{2(k_j+1)}(\tau_j) \right\}_{c=0} = Z \left\{ \prod_{j=1}^{n} \tilde{\psi}_{2(k_j+1)}(\tau_j) \right\}_{c=0}$$

Equation (A.5) is thus rewritten as:

$$\sum_{i_1=1}^{\infty} \cdots \sum_{i_n=1}^{\infty} \prod_{j=1}^{n} \psi_{2i_j}(\tau_j) \phi_{i_j}(k_j) = \prod_{j=1}^{n} \tilde{\psi}_{2(k_j+1)}(\tau_j)$$

(A.6)

where $\tilde{\psi}_n(\tau) = \left[ \psi_n(\tau) \right]_{c=0}$. Substituting (A.6) into (A.1), one has:

$$g_{\text{even}}(k_1, \ldots, k_n) = \sum_{\tau_1=0}^{\infty} \cdots \sum_{\tau_n=0}^{\infty} h_{\eta}(\tau_1, \ldots, \tau_n) \cdot \prod_{j=1}^{n} \tilde{\psi}_{2(k_j+1)}(\tau_j)$$

The proof for $g_{\text{odd}}(k_1, \ldots, k_n)$ is analogous.

REFERENCES


