On the Use of Equality Constraints in the Identification of Volterra-Laguerre Models

Chérif El Valid Diouf, Mihai Telescu, Pascale Cloastre, Noël Tanguy

On the Use of Equality Constraints in the Identification of Volterra–Laguerre Models

C. Diouf, M. Telescu, *Member, IEEE*, P. Cloastre, *Member, IEEE*, and N. Tanguy, *Member, IEEE*

***Abstract*—This letter focuses on the a posteriori correction of Volterra–Laguerre models in order to meet specific static or dynamic requirements. The authors set a general theoretical framework and provide an illustrative example.**

***Index Terms*—Lagrange multiplier, linear constraint, nonlinear system, polynomial filter, Volterra–Laguerre model.**

I. Introduction

VOLTERRA series are best described as a generalization of the impulse response function and often serve as an elegant input-output representation for a large class of nonlinear systems [1]. Applications of Volterra series include physiological modeling [2], automatic control [3], nonlinear device modeling [4], [5]. Volterra-based structures of polynomial filters have also proven very efficient in signal processing applications [6]–[9]. This letter explores the possibility of identifying nonlinear systems under equality constraints using discrete-time Volterra–Laguerre expansions. The authors show that Lagrange multipliers may be conveniently used to compute an a posteriori correction of Volterra–Laguerre expansions strictly enforcing specific static or dynamic properties on the models. In practice this may prove useful in various applications: an efficient input-output model of a digital device, for example, besides accounting for the nonlinearities of the original system also requires to precisely match its static behavior; in non-linear filter design, forcing a specific static gain function may be desirable; when simulating radar systems one may need to guarantee a specific behavior at precise frequencies, etc.

The letter is organized as follows: Section II provides a brief overview of the Volterra series and of the projection of Volterra kernels on Laguerre basis functions; Section III presents two strategies for applying equality constraints by a posteriori correction of Volterra–Laguerre coefficients and, finally, Section IV provides an illustrative example.

Manuscript received July 20, 2012; revised September 06, 2012; accepted September 21, 2012. Date of publication October 09, 2012; date of current version October 17, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Hsiao-Chun Wu.

The authors are with the Université Européenne de Bretagne, Université de Brest ; CNRS, UMR 6285, Lab-STICC, CS93837-29238 Brest Cedex 3, France.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2012.2223463

II. Volterra–Laguerre Models

The relation that binds input $x(k)$ and output $y(k)$ of a nonlinear M-order ($M \in N^*$) Volterra system is given by

$$y(k) = \sum_{m=1}^{M} \sum_{k_1=0}^{k} \cdots \sum_{k_m=0}^{k} h_m\left(k_1, \ldots, k_m\right) \prod_{l=1}^{m} x\left(k - k_l\right), \tag{1}$$

$h_m\left(k_1, \ldots, k_m\right)$ being the so-called Volterra kernels. Due to the multidimensional nature of the kernels, Volterra series present a high parametric complexity. It is therefore convenient to project each kernel $h_m\left(k_1, \ldots, k_m\right)$ on an orthonormal basis. A classical choice is that of the Laguerre functions [10], [11] $\phi_{m,i}(k)$ which form a complete orthonormal set in $l^2\left(\mathrm{N}\right)$ and may be conveniently defined by their z-transform as follows

$$\Phi_{m,i}(z) = \sqrt{1 - a_m^2} \frac{z}{z - a_m} \left(\frac{1 - a_m z}{z - a_m}\right)^i \tag{2}$$

where $a_m \in \left]-1, 1\right]$. Each kernel may therefore be rewritten as

$$h_m\left(k_1, \ldots, k_m\right) = \sum_{i_1=0}^{\infty} \cdots \sum_{i_m=0}^{\infty} C_{m,i_1,\ldots,i_m} \phi_{m,i_1}\left(k_1\right) \ldots \phi_{m,i_m}\left(k_m\right).$$

In practice a finite number of Laguerre functions, $I_m \geq 1$, is used which leads to an approximation of the Volterra model.

In order to simplify subsequent notations, the Volterra–Laguerre coefficients $C_{m,i_1,\ldots,i_m}$ are reshaped as a vector i.e., $\vec{D}_m = \left[D_{m,j}\right]_{j=0,\ldots,I_m^m-1}$ with $j = \sum_{n=0}^{I_m-1} i_n I_m^n$ and $D_{m,j} = C_{m,i_1,\ldots,i_m}$. Moreover, the products of the $\phi_{m,i}(k)$ functions are conveniently denoted:

$$\psi_{m,j}\left(k_1, \ldots, k_m\right) = \phi_{m,i_1}\left(k_1\right) \times \cdots \times \phi_{m,i_m}\left(k_m\right).$$

Finally, let

$$\overline{\psi}_{m,j}\left(k\right) = \sum_{k_1=0}^{k} \cdots \sum_{k_m=0}^{k} \psi_{m,j}\left(k_1, \ldots, k_m\right) \prod_{l=1}^{m} x\left(k - k_l\right),$$

the expression of the approximated system output, $\widetilde{y}(k)$, may thus be written as

$$\widetilde{y}\left(k\right) = \sum_{m=1}^{M} \sum_{j=0}^{I_m^m-1} D_{m,j} \overline{\psi}_{m,j}\left(k\right). \tag{3}$$

The observation of the system response to a well-chosen identification signal allows the estimation of the coefficients' spectrum $\vec{D} = \left[\vec{D}_1^T \ldots \vec{D}_M^T\right]^T$ using various techniques. In practice this operation is not always trivial, an in-depth discussion of the identification algorithms is however beyond the scope of this letter (see [10] and its references for details). The key issue we address here is not how to improve the overall quality of a Volterra–Laguerre model with respect to a general error criterion but rather how to operate minimal modifications of an already available model and strictly enforce specific properties.

III. Applying Equality Constraints

At this stage an M-order Volterra–Laguerre model of a given system is considered to be available, i.e., one disposes of the truncated kernels $\widetilde{h}_m(k_1,\ldots,k_m) \cong h_m(k_1,\ldots,k_m)$

$$\widetilde{h}_m(k_1,\ldots,k_m) = \sum_{j=0}^{I_m^m-1} D_{m,j}\psi_{m,j}(k_1,\ldots,k_m).$$

The question this section tackles is how to operate a minimal, a posteriori correction of the $\widetilde{h}_m(k_1,\ldots,k_m)$ kernels ($m \in [1;M]$) in order to have a new model (defined by new $\widetilde{h}'_m(k_1,\ldots,k_m)$ kernels) satisfying equality constraints.

Let all m-dimension real functions $f(k_1,\ldots,k_m)$ and $g(k_1,\ldots,k_m)$ belong to $l^2(\mathrm{N}^m)$. The inner product is given by $\langle f,g\rangle_m = \sum_{k_1=0}^{\infty}\cdots\sum_{k_m=0}^{\infty} f(k_1,\ldots,k_m)\, g(k_1,\ldots,k_m)$ and $\|f\|_m = \sqrt{\langle f,f\rangle_m}$ is the norm of f.

Let $Q_m\left(\vec{D}'_m\right)$ define the quadratic error introduced by conditioning the kernel $\widetilde{h}_m(k_1,\ldots,k_m)$, $Q_m\left(\vec{D}'_m\right) = \left\|\widetilde{h}_m - \widetilde{h}'_m\right\|_m$, the overall model quadratic error being $Q\left(\vec{D}'\right) = \sum_{m=1}^{M} Q_m\left(\vec{D}'_m\right)$.

While it is obvious that constraint enforcement is achieved by a perturbation of the Volterra–Laguerre spectrum, it is interesting to note that two distinct strategies exist. One may either operate individual kernel-wise corrections minimizing $Q_m\left(\vec{D}'_m\right)$ or a global correction minimizing $Q\left(\vec{D}'\right)$, the practical difference between the two will be illustrated in the example. Both are problems of optimization under constraints and can be solved by means of Lagrange multipliers.

In the first approach for each $\widetilde{h}_m$ kernel, we consider l_m constraint equations $\vec{K}_m\left(\vec{D}'_m\right) = 0$ and $[\lambda_i]_{i=1,\ldots,l_m}$, the associated Lagrange multipliers.

For the second strategy a number l of constraints equations $\vec{K}\left(\vec{D}'\right) = 0$ and the respective Lagrange multipliers $[\lambda_i]_{i=1,\ldots,l_m}$ are considered. The mathematical derivations are essentially the same for both approaches and only the first will be explicitly given below.

The Lagrange function to minimize is $Q'_m\left(\vec{D}'_m\right) = Q_m\left(\vec{D}'_m\right) + \sum_{i=1}^{l_m}\lambda_{m,i}K_{m,i}\left(\vec{D}'_m\right)$. Differentiating with respect to $D'_{m,j}$ and equating to 0 yields

$$\frac{\partial Q'_m\left(\vec{D}'_m\right)}{\partial D'_{m,j}} + \sum_{i=1}^{l_m}\lambda_{m,i}\frac{\partial K_{m,i}\left(\vec{D}'_m\right)}{\partial D'_{m,j}} = 0, \tag{4}$$

for $j = 0,\ldots,I_m^m - 1$.

Owing to $\partial Q'_m\left(\vec{D}'_m\right)/\partial D'_{m,j} = -2\left\langle \widetilde{h}_m - \widetilde{h}'_m, \psi_{m,j}\right\rangle$, (4) is equivalent to

$$\left\langle \widetilde{h}'_m, \psi_{m,j}\right\rangle_m + \frac{1}{2}\sum_{i=1}^{l_m}\lambda_{m,i}\frac{\partial K_{m,i}\left(\vec{D}'_m\right)}{\partial D'_{m,j}} = \left\langle \widetilde{h}_m, \psi_{m,j}\right\rangle_m. \tag{5}$$

Merging all equations given in relation (5) for $j = 0,\ldots,I_m^m - 1$, the problem to solve becomes

$$\Psi_m\vec{D}'_m + \frac{1}{2}K_{D'_m}\vec{\Lambda}_m = \vec{\widetilde{h}}_{\psi_m}, \tag{6}$$

where

$$\Psi_m = \left[\langle\psi_{m,i},\psi_{m,j}\rangle_m\right]_{i=0,\ldots,I_m^m-1,\ j=0,\ldots,I_m^m-1},$$
$$K_{D'_m} = \left[\frac{\partial K_{m,i}\left(\vec{D}'_m\right)}{\partial D'_{m,j}}\right]_{i=1,\ldots,l_m,j=0,\ldots,I_m^m-1},$$
$$\vec{\Lambda}_m = [\lambda_{m,i}]_{i=1,\ldots,l_m}$$

and $\vec{\widetilde{h}}_{\psi_m} = [\langle h_m, \psi_{m,i}\rangle]_{i=0,\ldots,I_m^m-1}$.

Owing to the Laguerre functions being orthogonal, (6) becomes

$$\vec{D}'_m + \frac{1}{2}K_{D'_m}\vec{\Lambda}_m = \vec{\widetilde{h}}_{\psi_m}. \tag{7}$$

Consider the particular case of constraints $\vec{K}_m\left(\vec{D}'_m\right) = 0$ given by a linear system of equations $L_m\vec{D}'_m - \vec{\sigma}_m = 0$. It has been shown in [12] and [13] that the coefficients $\vec{D}'_m = 0$ defining the constrained system are given by the following equation with minimal quadratic error

$$\vec{D}'_m = \vec{D}_m - L_m^+\left(L_m\vec{D}_m - \vec{\sigma}_m\right), \text{ for } m = 1,\ldots,M \tag{8}$$

where L_m^+ is the pseudo-inverse of L_m.

In the case of global linear constraints with $L\vec{D}' - \vec{\sigma} = 0$ the contribution of each kernel is corrected according to

$$\vec{D}'_m + \frac{1}{2}K_{D'_m}\vec{\Lambda} = \vec{\widetilde{h}}_{\psi_m} \tag{9}$$

where $K_{D'_m} = \left[\partial K_{m,i}\left(\vec{D}'_m\right)/\partial D'_{m,j}\right]_{i=1,\ldots,l_m,j=0,\ldots,I_m^m-1}$. Thus, a counterpart of (8) can be written as

$$\vec{D}' = \vec{D} - L^+\left(L\vec{D} - \vec{\sigma}\right) \tag{10}$$

where $\vec{D} = \left[\vec{D}_1^T \ldots D_M^T\right]^T$.

The theoretical difference between the two approaches is best seen when comparing (7) and (9). It is obvious that in the first case one attempts to solve a local problem related to a specific kernel. In the second case the problem is global, but one may split it along the different kernels and subsequently correct the contribution of each kernel. The practical difference between

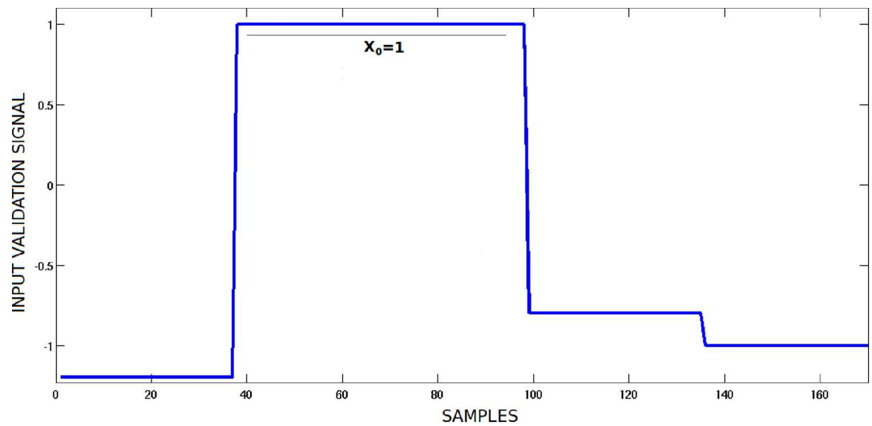

Fig. 1. Input validation signal.

the two strategies will be discussed in Section IV and illustrated by an example.

IV. Application

The following application is a classical case of black-box system identification and it is used to illustrate how the techniques in Section III can be used to improve the static quality of the model. Consider a two-kernel Volterra class nonlinear system proposed in [10]. The first kernel $h_1(k)$ defined is by its z-transform

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

The second kernel is defined in time domain by $h_2(k_1,k_2) = 0.25 \times h_{20}(k_1) \times h_{20}(k_2)$ where $h_{20}(k)$ is given by its z-transform

$$H_{20}(z) = \frac{z\,(z+1)}{(z-0.3)\,(z-0.2)}.$$

A Volterra–Laguerre model of the system is subsequently computed using a white noise as an identification signal. The well chosen Laguerre parameters (a1 $= 0.509$; a2 $= 0.801$) of [10] are used with $I_1 = 4$ Laguerre functions for the first order and $I_1 = 2$ for the second order. An a posteriori correction of this model is performed in order to force a better agreement between its static characteristic and that of the original system.

A. Kernel-Wise Constraints

A Volterra–Laguerre system has a polynomial static characteristic. In the particular case of the selected example for a constant input X and output Y the following relation holds

$$Y = A_1 X + A_2 X^2,$$

where each coefficient A_m depends on kernel h_m (in our example $m = 1{,}2$). It follows that tuning h_m via kernel-wise constraints allows us an optimal static behavior. In practice, this optimal polynomial (defined by its coefficients A'_m) is obtained by fitting the original static characteristic, an exploration phase is thus necessary.

The constraints equations are given by equation $L\vec{D}' - \vec{\sigma} = 0$ amounting to

$$\left(\frac{1+a_m}{1-a_m}\right)^{m/2} \sum_{j=0}^{I_m^m-1} D'_{m,j} - \vec{\sigma}_m = 0$$

where $\vec{\sigma}_m = A'_m$ and $L_m = [L_{m,i}]_{i=0,\ldots,I_m^m-1}$ with $L_{m,i} = ((1+a_m)/(1-a_m))^{m/2}$. The correction is then performed by applying (8).

B. Global Constraints

Point-wise global constraints can also be set. For a given static input X_0 the model is constrained to exactly match the system's original static output $Y'_0 = \sigma_0$. In the given example we have chosen $X_0 = 1$. The global constraints equation $L\vec{D}' - \vec{\sigma} = 0$ amounts to

$$X_0\sqrt{\frac{1+a_m}{1-a_m}} \sum_{j=0}^{I_1-1} D'_{1,j} + X_0^2\sqrt{\frac{1+a_m}{1-a_m}} \sum_{j=0}^{I_2^2-1} D'_{2,j} - \sigma_0 = 0$$

where $L = [L_i]_{i=0,\ldots,I_1+I_2^2-1}$ with $L_i = \sqrt{(1+a_m)/(1-a_m)}\,X_0$ for $i < I_1$ and $L_i = ((1+a_m)/(1-a_m))\,X_0^2$ for $i \geq I_1$. Equation (10) is used to compute the constrained model.

C. Results and Discussion

A stair-case like validation signal (Fig. 1) is used to contrast the response of the unconstrained model and those of the two constrained models described above. The original system response is provided as a reference (Fig. 2) and deviation plots are provided to emphasize the effect of the correction (Fig. 3).

It should be noted that the improvement in terms of static quality is visible while the dynamic behavior of the system suffers no significant damage. Obviously, when using point-wise global constraints more points could be added to the constraints equation to better anchor the static characteristic of the model.

Generally however, fitting the static characteristic of the original system with a polynomial and subsequently constraining the individual kernels of the Volterra–Laguerre model accordingly seems a better approach since a small number of con-

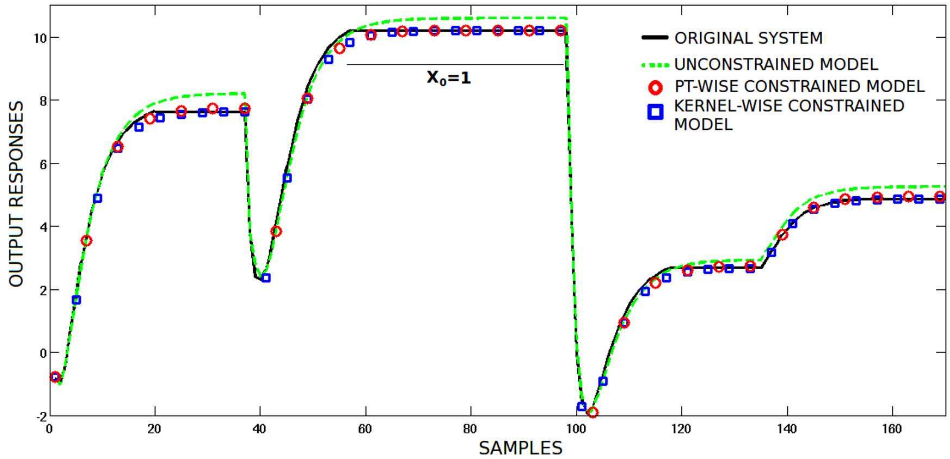

Fig. 2. Responses of the original system (solid line), identified model (dashed lines) and constrained modes (circles and squares).

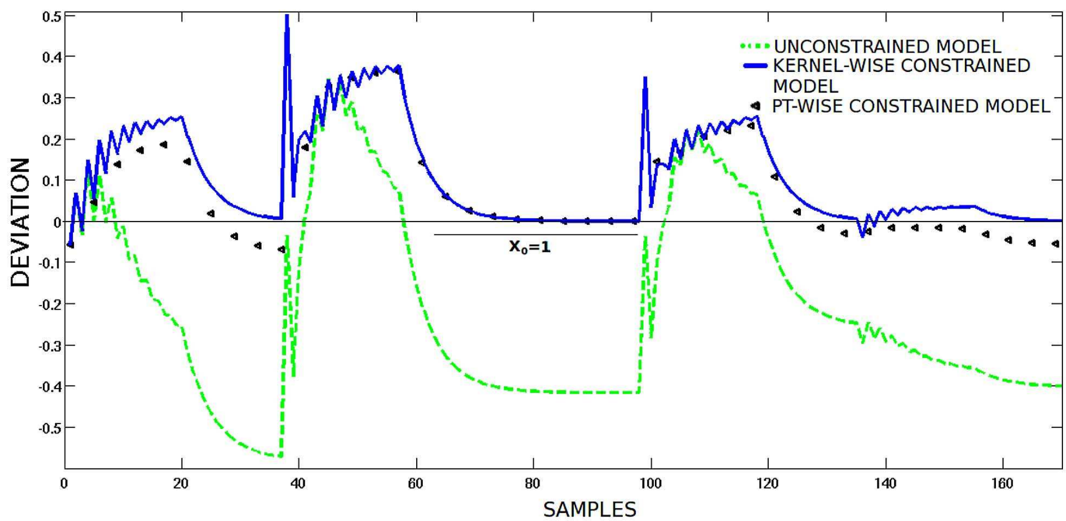

Fig. 3. Deviation with respect to the original system response.

straints equations lead to a significant overall improvement of the model. It should be noted that in practice the whole static characteristic may not be of interest, in which case the fit and subsequent correction should only concentrate on the specific working zone.

V. Conclusion

In this letter the issue of constraints in Volterra–Laguerre models has been addressed. Two approaches have been presented: kernel-wise constraints and global model constraints, both successfully illustrated by an example of static correction. This kind of approach may prove useful in signal processing applications using nonlinear filters, in the behavioral modeling of electronic devices, and, generally, in any identification application where specific system properties need to be enforced.

References

[1] J. W. Rugh, *Nonlinear System Theory The Volterra/Wiener Approach*. Baltimore, MD: Johns Hopkins Univ. Press, 1981.

[2] V. Z. Marmarelis, "Identification of nonlinear biological systems using Laguerre expansions of kernels," *Ann. Biomed. Eng.*, vol. 21, no. 6, pp. 573–589, 1993.

[3] G. Dumont and Y. Fu, "Non-linear adaptive control via Laguerre expansion of Volterra kernels," *Int. J. Adapt. Contr. Signal Process.*, vol. 7, pp. 367–382, 1993.

[4] A. Zhu and T. Brazil, "Behavioral modeling of rf power amplifiers based on pruned volterra series," *IEEE Microwave Wireless Compon.*, vol. 14, no. 12, pp. 563–565, 2004.

[5] M. G. Telescu, I. S. Stievano, F. G. Canavero, and N. Tanguy, "An application of Volterra series to IC buffer models," in *14th IEEE Workshop on Signal Propagation on Interconnects*, Hildesheim, Germany, May 9–12, 2010, pp. 93–96.

[6] J. Stapleton and S. Bass, "Adaptive noise cancellation for a class of nonlinear, dynamic reference channels," *IEEE Trans. Signal Process.*, vol. SP-32, pp. 143–150, Feb. 1985.

[7] V. J. Mathews, "Adaptive polynomial filters," *IEEE Signal Process. Mag.*, vol. 8, no. 3, pp. 10–26, Jul. 1991.

[8] L. Tan and J. Jiang, "Adaptive Volterra filters for active control of nonlinear noise processes," *IEEE Trans. Signal Process.*, vol. 49, pp. 1667–1676, Aug. 2001.

[9] L. Azpicueta-Ruiz, M. Zeller, A. Figueiras-Vidal, J. Arenas-Garcia, and W. Kellermann, "Adaptive combination of volterra kernels and its application to nonlinear acoustic echo cancellation," *IEEE Trans. Audio, Speech Lang. Process.*, vol. 19, pp. 97–110, Mar. 2010.

[10] A. Y. Kibangou, G. Favier, and M. M. Hassani, "Laguerre-volterra filters optimization based on Laguerre spectra," *J. Appl. Signal Process.*, vol. 2005, no. 17, pp. 2874–2885, 2005.

[11] R. J. G. B. Campello, G. Favier, and W. C. Amaral, "Optimal expansions of discrete-time Volterra models using Laguerre functions," *Automatica*, vol. 40, pp. 815–822, 2004.

[12] J. C. Boulbry, "Approximation de Signaux et Modélisation de Systèmes Linéaires," Thèse de doctorat, Univ. Bretagne Occidentale, Brest, France, 1989.

[13] L. C. Calvez, P. Vilbé, and P. Glouannec, "Orthonormal and nonorthonormal least squares approximation of a function subject to linear equality constraints," *IEEE Trans. Circuits Syst.*, vol. CAS-32, pp. 851–853, Aug. 1985.