BLIND IDENTIFIABILITY OF THIRD-ORDER VOLTERRA NONLINEAR SYSYTEMS

Hong-Zhou Tan[†], Tyseer Aboulnasr[‡]

School of Information Technology and Engineering, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5
[†]hongzhou@site.uottawa.ca, [‡]aboulnasr@genie.uottawa.ca

***Abstract* –**A novel approach to blindly estimate kernels of a Volterra nonlinear system up to the third order is proposed in this paper. The system is excited by an unobservable i.i.d. random sequence. Blind identifiability is achieved using second order statistics (SOS) rather than using higher order statistical (HOS) information to ensure lower complexity. Since the output of the Volterra system is linearly dependent upon its kernel parameters, conventional LMS or RLS algorithms can be used and consistent estimation of Volterra kernels can be achieved provided some conditions of persistent excitation (PE) are satisfied. The simulation demonstrated the ability of the proposed method to achieve a good estimation performance.

1. INTRODUCTION

While linear models no doubt form most common way of describing dynamic systems, for most real-world practical applications, however, there are advantages to using nonlinear models to characterize the inherent nonlinear relationships. Discrete Volterra series is used to describe the input-output relation in a nonlinear domain, so that characterization, analysis and synthesis are easily amenable. Applications of Volterra nonlinear system theory have played an ever-increasing role in nonlinear signal processing, communication and control during recent decades [1]-[3]. One common difficulty encountered when one wants to apply the Volterra functional representation to nonlinear problem involves the determination of the Volterra kernels [3]. As a result, kernel identification has always been a subject of many studies, especially in the case of *blind* identification, which has to rely on the output signals and the statistical properties of the input signals. Much work on blind identification for nonlinear models has been done recently [4]-[6]. Although identification of the specific nonlinear Hammerstein systems was studied in [4], only the sampled linear sub-systems can be identified blindly. The algorithm in [5] considered the neural-network based quadratic kernel estimation, since there exists a complicated nonlinear relationship between kernels and system outputs in HOS domain. In [6], blind identifiability of quadratic models using both the second- and the third-order statistics was discussed, but no simulations were provided to verify the theoretical results.

In this paper, the SOS-based blind kernel identification is considered up to the third-order Volterra systems. It is known that for blind identification, the SOS introduces an unknown bias into the final estimates if the output measurements are contaminated by additive Gaussian noises of unknown covariances. As will be shown, however, due to the linear relationship between the system output and the kernels in the SOS domain, unique kernel estimates exist if the system is persistently excited. Conventional adaptive signal processing algorithms such as LMS or RLS [7][8][9] then provide unbiased kernel estimates which means the influence of the noise at the output site is removed in the statistical sense for white noises. The organization of the paper is as follows. The problem as well as assumptions on the system model is delineated in Section 2. We present the procedure for blind identification of the third-order Volterra systems in SOS domain in Section 3. Simulations are provided in Section 4 to verify the performance of the proposed identification algorithms and finally, a conclusion is drawn in Section 5.

2. DESCRIPTION OF THE PROBLEM

Assume that a system output $\{y(t)\}$ is generated through a third-order Volterra model driven by a stationary input sequence $\{x(t)\}$. The observation $\{y_o(t)\}$ is corrupted by an additive white Gaussian or non Gaussian noise $\{n_w(t)\}$ which is independent of $\{y(t)\}$. The system is then described as follows:

$$\begin{aligned} y(t) &= \sum_{i=0}^{n} h(i)x(t-i) \\ &+ \sum_{i=0}^{n}\sum_{j=0}^{n} h(i,j)x(t-i)x(t-j) \\ &+ \sum_{i=0}^{n}\sum_{j=0}^{n}\sum_{k=0}^{n} h(i,j,k)x(t-i)x(t-j)x(t-k) \end{aligned} \tag{1a}$$

$$y_o(t) = y(t) + n_w(t) \tag{1b}$$

where the excited sequence of the model $\{x(t)\}$ is an unobservable zero-mean i.i.d. signal with known $g_{ix} = E[x^i(t)] \neq 0$, $\forall\, i$=2, 3, 4, 5, 6. Unknown kernels include $\{h(i), h(i,j), h(i,j,k)\,; \forall\, i, j, k$=0, 1, ... , $n\}$. Without any loss of generality, the kernels satisfy the following properties [3] (a) bounded: $h(i) = h(i,j) = h(i,j,k) = 0$, $\forall\, i, j, k>n$; (b) causal: $h(i) = h(i,j) = h(i,j,k) = 0$, $\forall\, i, j, k<0$; and (c) stable: $\sum_i |h(i)| < \infty$, $\sum_{i,j} |h(i,j)| < \infty$ and $\sum_{i,j,k} |h(i,j,k)| < \infty$.

For any observed stationary random sequence $\{y_o(t)\}$, its second-order moment is given by

$$m_2^{y_o}(\boldsymbol{t}) = E\{[y_o(t)][y_o(t-\boldsymbol{t})]\}$$
$$= m_2^y(\boldsymbol{t}) + m_2^{n_w}(\boldsymbol{t}) \qquad (2)$$

where E denotes an expectation operation and $\boldsymbol{t}$ refers to time lags of the sequence.

The objective of this study is to determine the third-order Volterra kernels $\{h(i), h(i,j), h(i,j,k); \forall\, i, j, k=0, 1, \dots, n\}$ of model (1) based on the SOS properties of the measured output sequence $\{y_o(t)\}$ in Eq. (2).

3. BLIND IDENTIFIABILITY OF THIRD-ORDER VOLTERRA SYSTEMS

The following analysis bridges the gap between the SOS of the system output observations $\{y_o(t)\}$ and the system itself $\{h(i), h(i,j), h(i,j,k); \forall\, i,j,k=0, 1, \dots, n\}$.

Let $\{x(t)\}$ be an arbitrary zero-mean, i.i.d. random sequence that excites a third order Volterra system as in Eq.(1). Using Theorem 1 in [5] and the definition in Eq.(2), the second-order moment of output measurements of this system are obtained as follows

$$\begin{aligned}
m_2^{y_o}(\boldsymbol{t}) = & \sum_{i,j=0}^{n} h(i)h(j) f_1(\boldsymbol{g}_{2x}, \boldsymbol{t}, i, j) \\
& + \sum_{i,j,k=0}^{n} h(i)h(j,k) f_2(\boldsymbol{g}_{3x}, \boldsymbol{t}, i, j, k) \\
& + \sum_{i,j,k,l=0}^{n} h(i)h(j,k,l) f_3(\boldsymbol{g}_{2x}^2, \boldsymbol{g}_{4x}, \boldsymbol{t}, i, j, k, l) \\
& + \sum_{i,j,k=0}^{n} h(i,j)h(k) f_4(\boldsymbol{g}_{3x}, \boldsymbol{t}, i, j, k) \\
& + \sum_{i,j,k,l=0}^{n} h(i,j)h(k,l) f_5(\boldsymbol{g}_{2x}^2, \boldsymbol{g}_{4x}, \boldsymbol{t}, i, j, k, l) \\
& + \sum_{i,j,k,l,s=0}^{n} h(i,j)h(k,l,s) f_6(\boldsymbol{g}_{2x}\boldsymbol{g}_{3x}, \boldsymbol{g}_{5x}, \boldsymbol{t}, i, j, k, l, s) \qquad (3) \\
& + \sum_{i,j,k,l=0}^{n} h(i,j,k)h(l) f_7(\boldsymbol{g}_{2x}^2, \boldsymbol{g}_{4x}^2, \boldsymbol{t}, i, j, k, l) \\
& + \sum_{i,j,k,l,s=0}^{n} h(i,j,k)h(l,s) f_8(\boldsymbol{g}_{2x}\boldsymbol{g}_{3x}, \boldsymbol{g}_{5x}, \boldsymbol{t}, i, j, k, l, s) \\
& + \sum_{i,j,k,l,s,v=0}^{n} h(i,j,k)h(l,s,v) \\
& \qquad f_9(\boldsymbol{g}_{2x}\boldsymbol{g}_{4x}, \boldsymbol{g}_{3x}^2, \boldsymbol{g}_{6x}, \boldsymbol{t}, i, j, k, l, s, v) \\
& + m_2^{n_w}(\boldsymbol{t})
\end{aligned}$$

where $f_1(\boldsymbol{g}_{2x}, \boldsymbol{t}, i, j)$ is a function of $\boldsymbol{g}_{2x}, \boldsymbol{t}, i, j$, $f_2(\boldsymbol{g}_{3x}, \boldsymbol{t}, i, j, k)$ is a function of $\boldsymbol{g}_{3x}, \boldsymbol{t}, i, j, k$, …, and so on and are given by:

$$\left\{\begin{aligned}
& f_1(\boldsymbol{g}_{2x}, \boldsymbol{t}, i, j) \\
& = m_2^x(\boldsymbol{t} + j - i) \\
& f_2(\boldsymbol{g}_{3x}, \boldsymbol{t}, i, j, k) \\
& = m_3^x(\boldsymbol{t} + j - i, \boldsymbol{t} + k - i) \\
& f_3(\boldsymbol{g}_{2x}^2, \boldsymbol{g}_{4x}, \boldsymbol{t}, i, j, k, l) \\
& = m_4^x(\boldsymbol{t} + j - i, \boldsymbol{t} + k - i, \boldsymbol{t} + l - i) \\
& f_4(\boldsymbol{g}_{3x}, \boldsymbol{t}, i, j, k) \\
& = m_3^x(j - i, \boldsymbol{t} + k - i) \\
& f_5(\boldsymbol{g}_{2x}^2, \boldsymbol{g}_{4x}, \boldsymbol{t}, i, j, k, l) \\
& = m_4^x(j - i, \boldsymbol{t} + k - i, \boldsymbol{t} + l - i) \\
& f_6(\boldsymbol{g}_{2x}\boldsymbol{g}_{3x}, \boldsymbol{g}_{5x}, \boldsymbol{t}, i, j, k, l, s) \\
& = m_5^x(j - i, \boldsymbol{t} + k - i, \boldsymbol{t} + l - i, \boldsymbol{t} + s - i) \\
& f_7(\boldsymbol{g}_{2x}^2, \boldsymbol{g}_{4x}, \boldsymbol{t}, i, j, k, l) \\
& = m_4^x(j - i, k - i, \boldsymbol{t} + l - i) \\
& f_8(\boldsymbol{g}_{2x}\boldsymbol{g}_{3x}, \boldsymbol{g}_{5x}, \boldsymbol{t}, i, j, k, l, s) \\
& = m_5^x(j - i, k - i, \boldsymbol{t} + l - i, \boldsymbol{t} + s - i) \\
& f_9(\boldsymbol{g}_{2x}\boldsymbol{g}_{4x}, \boldsymbol{g}_{3x}^2, \boldsymbol{g}_{6x}, \boldsymbol{t}, i, j, k, l, s, v) \\
& = m_6^x(j - i, k - i, \boldsymbol{t} + l - i, \boldsymbol{t} + s - i, \boldsymbol{t} + v - i)
\end{aligned}\right. \qquad (4)$$

The terms in Eq.(4) are given as follows for an i.i.d. random sequence $\{x(t)\}$:

$$m_2^x(\boldsymbol{a}) = \begin{cases} \boldsymbol{g}_{2x}, \forall \boldsymbol{a} = 0 \\ 0, \quad \text{otherwise} \end{cases} \qquad (5a)$$

$$m_3^x(\boldsymbol{a}_1, \boldsymbol{a}_2) = \begin{cases} \boldsymbol{g}_{3x}, \forall \boldsymbol{a}_1 = \boldsymbol{a}_2 = 0 \\ 0, \quad \text{otherwise} \end{cases} \qquad (5b)$$

$$m_4^x(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3) = \begin{cases} \boldsymbol{g}_{2x}^2, \forall \boldsymbol{a}_i = 0, i = 1,\dots,3 \\ \quad \text{and } \boldsymbol{a}_j = \boldsymbol{a}_k, \ j,k = 1,\dots,3, \ j \neq k \neq i \\ \boldsymbol{g}_{4x}, \forall \boldsymbol{a}_1 = \boldsymbol{a}_2 = \boldsymbol{a}_3 = 0 \\ 0, \quad \text{otherwise} \end{cases} \qquad (5c)$$

$$m_5^x(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4) = \begin{cases} \boldsymbol{g}_{2x}\boldsymbol{g}_{3x}, \forall \boldsymbol{a}_i = 0, i = 1,\dots,4, \text{and} \\ \quad \boldsymbol{a}_j = \boldsymbol{a}_k = \boldsymbol{a}_l, \ j,k,l = 1,\dots,4, \ j \neq k \neq l \neq i \\ \quad \text{or } \forall \boldsymbol{a}_i = \boldsymbol{a}_j = 0, i, j = 1,\dots,4, \ i \neq j, \text{and} \\ \quad \boldsymbol{a}_k = \boldsymbol{a}_l, k, l = 1,\dots,4, \ k \neq l \neq i \neq j \\ \boldsymbol{g}_{5x}, \quad \forall \boldsymbol{a}_1 = \boldsymbol{a}_2 = \boldsymbol{a}_3 = \boldsymbol{a}_4 = 0 \\ 0, \quad \text{otherwise} \end{cases}$$
(5d)

$$m_6^x(\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5)$$

$$=\begin{cases}\gamma_{2x}\gamma_{4x}, \forall \alpha_i = 0, i=1,\ldots,5, \text{ and} \\ \quad \alpha_j=\alpha_k=\alpha_l=\alpha_s,\ j,k,l,s=1,\ldots,5,\ j\neq k\neq l\neq s\neq i \\ \quad \text{or } \forall \alpha_i=\alpha_j=\alpha_k=0,\ i,j,k=1,\ldots,5,\ i\neq j\neq k, \text{ and} \\ \quad \alpha_l=\alpha_s, l,s=1,\ldots,5, l\neq s\neq i\neq j\neq k \\ \gamma_{3x}^2, \forall \alpha_i=\alpha_j=0, i,j=1,\ldots,5, i\neq j, \text{ and} \\ \quad \alpha_k=\alpha_l=\alpha_s=0, k,l,s=1,\ldots,5, k\neq l\neq s\neq i\neq j \\ \gamma_{6x}, \quad \forall \alpha_1=\alpha_2=\alpha_3=\alpha_4=\alpha_5=0 \\ 0, \quad \text{otherwise}\end{cases} \tag{5e}$$

By substituting $\alpha_1,\alpha_2,\ldots,\alpha_5$ in Eq. (5a)~(5e) for the arguments of $m_i^x(.)$, i=2,…, 6 in Eq. (4), we determine $f_1(\gamma_{2x},\tau,i,j)$ as a function of γ_{2x},τ,i,j, $f_2(\gamma_{3x},\tau,i,j,k)$ as a function of γ_{3x},τ,i,j,k, …, and so on. For example,

$$f_1(\gamma_{2x},\tau,i,j)=m_2^x(\tau+j-i)=\begin{cases}\gamma_{2x}, \forall \tau+j-i=0 \\ 0, \quad \text{otherwise}\end{cases}$$

Therefore, Eq. (3), the result, follows immediately.

Indeed, Eq. (3) can be rewritten as a matrix representation. Assume the time lag $\tau\in[1,p]$, we construct a $p\times p$ matrix M_x as

$$M_x=\begin{bmatrix} f_1(1) & f_2(1) & \cdots & f_9(1) \\ f_1(2) & f_2(2) & \cdots & f_9(2) \\ \vdots & \vdots & \vdots & \vdots \\ f_1(p) & f_2(p) & \cdots & f_9(p)\end{bmatrix} \tag{6}$$

where p is defined as $p=(n+1)^2+2(n+1)^3+3(n+1)^4+2(n+1)^5+(n+1)^6$. In Eq. (6) $f_1(.)$ is a $1\times(n+1)^2$ vector $f_1(\tau)=[f_1(\gamma_{2x},\tau,0,0)\ldots f_1(\gamma_{2x},\tau,n,n)]$, τ =1, 2, …, p. Similarly, $f_2(.)$ is a $1\times(n+1)^3$ vector $f_2(.)=[f_2(\gamma_{3x},\tau,0,0,0)\ldots f_2(\gamma_{3x},\tau,n,n,n)]$, τ =1, 2, …, p, …, and finally $f_9(.)$ is a $1\times(n+1)^6$ vector $f_9(.)=[f_9(\gamma_{2x}\gamma_{4x},\gamma_{3x}^2,\gamma_{6x},\tau,0,0,0,0,0,0)\ldots$ $f_9(\gamma_{2x}\gamma_{4x},\gamma_{3x}^2,\gamma_{6x},\tau,n,n,n,n,n,n)]$, τ =1, 2, …, p.

The corresponding $p\times 1$ output moment vector M_y, $p\times 1$ noise moment vector M_{n_w} and $p\times 1$ system kernel vector H are respectively given by

$$M_y=[m_2^{y_o}(1),m_2^{y_o}(2),\ldots,m_2^{y_o}(p),]^T \tag{7}$$

$$M_{n_w}=[m_2^{n_w}(1),m_2^{n_w}(2),\ldots,m_2^{n_w}(p),]^T \tag{8}$$

$$H=[H(i,j),H(i,j,k),H(i,j,k,l),\ldots,H(i,j,k,l,s,v)]^T \tag{9}$$

$$H(i,j)=[h^2(0),h^2(1),\ldots,h^2(n),h(0)h(1),\ldots,h(n-1)h(n)]^T \tag{10}$$

$$H(i,j,k)=[h(0)h(0,0),h(0)h(1,1),\ldots,h(n)h(n,n), h(0)h(0,1),\ldots,h(n)h(n-1,n)]^T \tag{11}$$

$$H(i,j,k,l)=[h(0)h(0,0,0),h(0)h(1,1,1),\ldots,h(n)h(n,n,n), h(0)h(0,0,1),\ldots,h(n)h(n-1,n-1,n)]^T \tag{12}$$

$$H(i,j,k,l,s,v)=[h^2(0,0,0),h^2(0,0,1),\ldots,h^2(n,n,n), h(0,0,0)h(0,0,1),\ldots,h(n,n,n)h(n-1,n-,n-1)]^T \tag{13}$$

Consequently, it is straightforward to formulate the following matrix expression of Eq. (3)

$$M_y=M_xH+M_{n_w} \tag{14}$$

Our goal is to determine the elements of H (for a third order Volterra system) given M_x, M_y and M_{n_w}. This blind identification of the third-order Volterra systems is then provided in the following theorem.

[Theorem 1] The third-order Volterra system (1) is blindly identifiable if it is persistently excited (PE); i.e. if the $p\times p$ matrix M_x is nonsingular.

Proof. Since $m_2^{n_w}(\tau)=0$, $\forall\tau\neq 0$, M_{n_w}=0. If M_x is nonsingular, it is easy to find that

$$H=M_x^{-1}M_y \tag{15}$$

which yields the desired result.

Bearing in mind that a consistent estimate of $m_2^{y_o}(\tau)$ is obtained by [9]

$$\hat{m}_2^{y_o}(\tau)=\lim_{N_\Omega\to\infty}\frac{1}{N_\Omega}\sum_{t\in\Omega}y_o(t)y_o(t+\tau) \tag{16}$$

where N_Ω is the number of output samples in region Ω. Substituting $m_2^{y_o}(.)$ of (7) by the estimated $\hat{m}_2^{y_o}(.)$ in (16) implies the consistent kernel estimate

$$\hat{H}=\lim_{N_\Omega\to\infty}M_x^{-1}M_y(N_\Omega) \tag{17}$$

It is clear that an adaptive algorithm such as LMS or RLS [7][8][9] realizes the matrix equation in (15) in an iterative way with an expected performance in a statistical sense.

Particularly noteworthy is the fact that if the matrix M_x is singular, the kernel estimates can also be achieved through (i) reducing the $p\times p$ matrix M_x to a $q\times p$ matrix $\overline{M}_x$ where q<p and rank ($\overline{M}_x$)=q in terms of singular value decomposition (SVD) [9] or (ii) using steepest descent strategy [9] or (iii) presetting some of Volterra parameters as zeros in the sense that sparse Volterra systems are considered, which is reasonable in some practical situations.

4. SIMULATIONS

In this section, we provide two simulated numerical examples. Let the unknown truncated Volterra systems be

Model 1: $y_o(t) = h(0,0)x^2(t) + n_w(t)$ (18)

Model 2: $y_o(t) = h(0,0)x^2(t) + h(1,1,1)x^3(t-1) + n_w(t)$ (19)

where $h(0,0)$=1 for model 1 and $h(0,0)$=2, $h(1,1,1)$=1 for model 2. Generally an i.i.d. exponential distributed random sequence $\{x(t)\}$ with zero mean is generated as the input signals. We assume the model output observations are corrupted by white Gaussian noise $\{n_w(t)\}$, where signal-to-noise-ratio $SNR = 20\log_{10}\frac{E(y^2(t))}{E(n_w^2(t))}$ =18 dB. It is concluded from Eq. (16) that a larger N_Ω is capable of providing a more accurate estimate of $m_2^{y_o}(\boldsymbol{t})$. In the study, 4096 ($N_\Omega$) output samples which are computed by convolving the random inputs with the true model (18) and (19) are utilized. To verify the theoretical analysis, 50 Monte Carlo runs are performed. It is not difficult to find that these two models are persistently excited. The final results, kernel estimates as well as the corresponding standard deviations are summarized in Table 1 and 2, respectively for model 1 and 2. It is indicated that the estimated Volterra kernels are very close to their true values even under a lower SNR environment.

5. CONCLUSION

In this paper, we considered blind identification of third-order Volterra systems using only Second Order Statistics in the time domain. The unobservability assumption of input i.i.d. random sequence requires that the kernel estimation be performed based on the output observations only. It is shown that under a persistent excitation condition, the determination of the Volterra kernels is reduced to a linear modeling problem and can be resolved by typical adaptive techniques. While the HOS-based approach with significantly more computational cost is capable of removing the effect of *Gaussian* noise corrupting the output measurements, the SOS-based kernel estimation scheme can remove *white* noise with any distribution and offers significant reduction in the computational burden. Simulation results indicate the effectiveness of the proposed method for blind identification of any truncated third-order Volterra nonlinear systems. The compromise between HOS and SOS-based use is an interesting one for further study.

REFERENCES

[1] V. J. Mathews, G. L. Sicuranza, *Polynomial Signal Processing*, John Wiley & Sons, 2000.

[2] C. Jaszczur, T. Aboulnasr, "Low complexity, real exponent high order polynomial filters", *IEEE Int. Sym. Sig. Process. and Inf. Tech.*, 2002.

[3] J. Rugh, *Nonlinear Model Theory: The Volterra/Wiener Approach*, Johns Hopkins Press, 1981.

[4] E.-W. Bai, M. Fu, "A blind approach to Hammerstein model identification", *IEEE Trans. Signal Processing*, Vol. 50, No. 7, 2002, pp. 1610-11619.

[5] H.-Z. Tan, T. W. S. Chow, "Blind identification of quadratic nonlinear models using neural networks with higher order cumulants", *IEEE Industrial Electronics*, Vol. 47, No. 3, 2000, pp. 687-696.

[6] P. Bondon, "On the identifiability of a quadratic stochastic model", *IEEE Signal Processing Workshop on Higher-Order Statistics*, 1993, pp. 46-50.

[7] T. Aboulnasr, K. Mayyas, "Complexity reduction of the NLMS algorithm via selective coefficient update", *IEEE Trans. Signal Processing*, Vol. 47, No. 5, 1999, pp. 1421-14224.

[8] T. Aboulnasr, K. Mayyas, "A robust variable step-size LMS-type algorithm: analysis and simulations", *IEEE Trans. Signal Processing*, Vol. 45, No. 3, 1997, pp. 631-639.

[9] J. Mendel, *Lessons In Estimation Theory ForSignal Processing, Communications, and Control*, Prentice Hall, 1995.

Table 1 True and estimated Volterra kernels of the model 1, with the corresponding means (standard deviations) under 50 Monte Carlo runs

$h(0,0)$	$SNR=\infty$	$SNR=18$ dB
true	1	1
estimated	0.9976(0.1030)	1.0997(0.0948)

Table 2 True and estimated Volterra kernels of the model 2, with the corresponding means (standard deviations) under 50 Monte Carlo runs

	$SNR=\infty$	$SNR=18$ dB
true $h(0,0)$	2	2
est. $h(0,0)$	2.0948(0.1777)	2.1477(0.2606)
true $h(1,1,1)$	1	1
est. $h(1,1,1)$	0.9605(0.0689)	0.9420(0.0903)