TOM-BASED BLIND IDENTIFICATION OF CUBIC 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 – In this paper, we extend our previous studies on blind cubic nonlinear system identification from the second-order moment (SOM) domain into the third-order moment (TOM) domain. It will be shown that under the given sufficient conditions, more subsets of truncated sparse Volterra systems can be blindly identified using TOM instead of SOM. This is consistent with the fact that more statistical knowledge can be obtained in the third-order statistics domain for blind system identification. Simulation results confirm the validity and usefulness of our proposed algorithm.

1. INTRODUCTION

Nonlinear system identification is a fundamental signal processing technique aimed at choosing an appropriate mathematical nonlinear model to match an unknown dynamic system in terms of its input and output. In the case where the input signal to the system to be handled is not directly available, the system identification has to be carried out based on the system output signal along with some statistical knowledge of the input signal, thus the name "*blind*". There exist various nonlinear systems: classes of Volterra system, Wiener system, Hammerstein system and neural network-based system, etc. Among them the Volterra system model is the most popular representation in nonlinear signal processing, communication, fault diagnosis, and biomedical applications due to the linearity of the model output as a function of the system parameters, i.e., Volterra kernels [1]. *Blind* identification of Volterra systems is unquestionably a common but difficult problem since the number of kernels in a *full-sized* Volterra system (i.e., all kernels are nonzero) is huge. This challenge motivated many researchers to search for closed-form solutions for blind identification of some specific *truncated* Volterra systems such as quadratic, cubic, or bilinear systems. In [2], blind identifiability of quadratic models in the second-order cumulant (SOC) and third-order cumulant (TOC) domain was established in a closed form. The authors in [3] provided an explicit description between the TOC of the quadratic system output and the system kernels. In [4], particular bilinear systems were identified blindly in the fourth order cumulant domain. These results imply that for a full-sized quadratic or bilinear system, unique matrix inversion solution of blind identification does not exist because of the coupling of unknown kernels and the shortage of available statistical information. For some applications [5], there is a need for higher order *sparse* Volterra filters (i.e., a Volterra system that has zero coefficients for specified kernels). Matrix inversion solution for some subclass of sparse Volterra systems may be possible. In [6], the authors proposed a blind identification algorithm in SOM domain for sparse Volterra systems such as $y(t) = h(k_1,k_1)u^2(t-k_1) + h(k_2,k_2)u^2(t-k_2)$, where $k_1, k_2 = 0,1,2,\cdots,n$. In this paper, we propose a TOM based approach to blind identification of more complex but still sparse Volterra systems. The organization of the paper is as follows. In Section 2, we briefly present the blind identification problem for cubic systems in the TOM domain. The sufficient conditions for blind identifiability are derived and discussed 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. PROBLEM FORMULATION

Consider the following cubic systems [7]

$$\begin{cases} y(t) = \sum_{l=1}^{3} \sum_{k_1,k_2,\cdots,k_l=0}^{n} h(k_1,k_2,\cdots,k_l)x(t-k_1)x(t-k_2)\cdots x(t-k_l) \\ y_o(t) = y(t) + n_w(t) \end{cases} \tag{1}$$

where $\{y(t)\}$ is the output of the system, $\{y_o(t)\}$ is the output measurement contaminated by an additive white Gaussian noise $\{n_w(t)\}$ which is independent of $\{y(t)\}$. $\{x(t)\}$ is an unobservable zero-mean i.i.d. signal with any distribution whose nonzero statistics $\gamma_{\beta\cdot x} = E[x^\beta(t)] \neq 0$, are known up to β=9. Note that n is the memory length. The cubic kernels are then defined as $h(k_1)$, $h(k_1,k_2)$ and $h(k_1,k_2,k_3)$, where $k_1,k_2,k_3 \in [0,n]$ and $\sum_{k_1,k_2,\ldots,k_l} \left|h(k_1,k_2,\cdots,k_l)\right| < \infty$, l=1, 2, 3, indicating bounded input-output, causality and stability. After some straightforward derivation, the TOM of $\{y_o(t)\}$ is given by

$$\begin{aligned} m_3^{y0}(\tau_1,\tau_2) &= E\{y_0(t)y_0(t+\tau_1)y_0(t+\tau_2)\} \\ &= m_3^{y}(\tau_1,\tau_2) \text{ if } \tau_1 \neq \tau_2 \neq 0 \end{aligned} \tag{2}$$

which suggests that under some conditions, the TOM of noisy output observation $\{y_o(t)\}$ is insensitive to the Gaussian noise. Here $E\{.\}$ denotes an expectation operation and τ_1, τ_2 refers to time lags of the sequence. In this paper, we will consider sparse third-order Volterra systems where only some of the coefficients $h(k_1)$, $h(k_1,k_2)$ and $h(k_1,k_2,k_3)$ of system in Equ. (1) are nonzero. Our goal is to determine these coefficients based only on knowledge of $m_3^{y0}(\tau_1,\tau_2)$ in Equ. (2).

3. BLIND IDENTIFIABILITY OF CUBIC SYSTEMS IN TOM DOMAIN

After removing all redundant entries in Equ. (1) such as $h(0,1)x(t)x(t\text{-}1)$ and $h(1,0)x(t\text{-}1)x(t)$, or $h(0,1,2)x(t)x(t\text{-}1)x(t\text{-}2)$, $h(0,2,1)x(t)x(t\text{-}2)x(t\text{-}1)$, $h(1,0,2)x(t\text{-}1)x(t)x(t\text{-}2)$, $h(1,2,0)x(t\text{-}1)x(t\text{-}2)x(t)$, $h(2,0,1)x(t\text{-}2)x(t)x(t\text{-}1)$, and $h(2,1,0)x(t\text{-}2)x(t\text{-}1)x(t)$, …, the system in Equ. (1) is described as

$$y_o(t) = \sum_{i=1}^{3} H_{ri}^T X_{ri}(t) + n_w(t) \tag{3}$$

The vectors $X_{r1}(t)$ and H_{r1}, $X_{r2}(t)$ and H_{r2} and $X_{r3}(t)$ and H_{r3} have the dimension of $q_1 \times 1$, $q_2 \times 1$ and $q_3 \times 1$ respectively, given by

$$X_{r1}(t) = [x(t)\ x(t-1)\ x(t-2) \cdots x(t-n)]^T \tag{4a}$$

$$X_{r2}(t) = [x^2(t)\ \ x(t)x(t-1)\ \cdots\ x(t)x(t-n)\ \ x^2(t-1)\ \cdots\ x(t-1)x(t-n) \cdots x^2(t-n)]^T \tag{4b}$$

$$X_{r3}(t) = [x^3(t)\ \ x^2(t)x(t-1)\ \cdots\ x(t)x^2(t-n)\ \ x^3(t-1)\ \ x^2(t-1)x(t-2) \cdots x^3(t-n)]^T \tag{4c}$$

and

$$H_{r1} = [h(0)\ h(1)\ h(2) \cdots h(n)]^T \tag{5a}$$

$$H_{r2} = [h(0,0)\ \ h(0,1)\ \cdots\ h(0,n)\ \ h(1,1)\ \cdots\ h(1,n) \cdots h(n,n)]^T \tag{5b}$$

$$H_{r3} = [h(0,0,0)\ \ h(0,0,1)\ \cdots\ h(0,n,n)\ \ h(1,1,1)\ \cdots\ h(1,1,n)\ \cdots\ h(n,n,n)]^T \tag{5c}$$

where

$$\begin{cases} q_1 = n+1 \\ q_2 = (n+1)(n+2)/2 \\ q_3 = q_2 + \{q_2 - q_1\} + \{q_2 - [(n+1)+n]\} \\ \qquad + \cdots + \{q_2 - [(n+1)+n+\cdots+2]\} \end{cases} \tag{6}$$

and each term in the RHS is set to zero if it is evaluated as negative. For instance, if n=1, p=3, then $q_1 = 2, q_2 = 3$, and $q_3 = 4$. Substituting Equ. (3) into Equ. (2) and expanding $m_3^{y0}(\tau_1,\tau_2)$ over the region of $\tau_1 \in [l_1,l_2]$ and $\tau_2 \in [l_3,l_4]$, $l_1,l_2,l_3,l_4 \geq 0$ leads to

$$M_{3x}^T \cdot H = M_{3y_o} - \Delta \tag{7}$$

where

$$M_{3x} = [M_3^x(l_1,l_3)^T, M_3^x(l_1,l_3+1)^T, \cdots, M_3^x(l_1,l_4)^T, \cdots, M_3^x(l_2,l_4)^T]^T \tag{8a}$$

$$\begin{aligned} M_3^x(\tau_1,\tau_2) = [&E\{X_{r1}(t)\otimes X_{r1}(t-\tau_1)\otimes X_{r1}(t-\tau_2)\}, \cdots, \\ &E\{X_{r1}(t)\otimes X_{r1}(t-\tau_1)\otimes X_{r3}(t-\tau_2)\}, \cdots, \\ &E\{X_{r3}(t)\otimes X_{r3}(t-\tau_1)\otimes X_{r3}(t-\tau_2)\}]^T \end{aligned} \tag{8b}$$

$$H = [H_{r1}\otimes H_{r1}\otimes H_{r1}, \cdots, H_{r1}\otimes H_{r1}\otimes H_{r3}, \cdots, H_{r3}\otimes H_{r3}\otimes H_{r3}]^T \tag{8c}$$

$$M_{3y_o} = [m_3^{y_o}(l_1,l_3), m_3^{y_o}(l_1,l_3+1), \cdots, m_3^{y_o}(l_1,l_4), \cdots, m_3^{y_o}(l_2,l_4)]^T \tag{8d}$$

$$\Delta = m_1^y \cdot [m_2^{n_w}(l_1) + m_2^{n_w}(l_3) + m_2^{n_w}(l_3 - l_1), \cdots, m_2^{n_w}(l_2) + m_2^{n_w}(l_4) + m_2^{n_w}(l_4 - l_2)]^T \tag{8e}$$

where $\otimes$ is the Kronecker product. Clearly if the terms with $\tau_1 = \tau_2 = 0$ and $\tau_1 = \tau_2$ are removed from M_{3y_o} and M_{3x}, $M_{3y_o} = M_{3y}$ and Δ=0. In other words, the effect of Gaussian noise has been eliminated.

On the other hand, the kronecker product in Equ. (8c) has also introduced redundant entries that must be eliminated before solving Equ. (7) for H. Indeed, one can see that $H_{ri}\otimes H_{ri}\otimes H_{rj} = H_{ri}\otimes H_{rj}\otimes H_{ri} = H_{rj}\otimes H_{ri}\otimes H_{ri}$ for i, j=1, 2, 3 and $i \neq j$;
$H_{ri}\otimes H_{rj}\otimes H_{rk} = H_{ri}\otimes H_{rk}\otimes H_{rj} = H_{rj}\otimes H_{ri}\otimes H_{rk} = H_{rj}\otimes H_{rk}\otimes H_{ri} = H_{rk}\otimes H_{ri}\otimes H_{rj} = H_{rk}\otimes H_{rj}\otimes H_{ri}$ for i, j, k=1, 2, 3 and $i \neq j \neq k$. Also, in $H_{ri}\otimes H_{ri}\otimes H_{ri}$,

$$\begin{aligned} &h(\alpha_1,\alpha_2,\cdots,\alpha_i)h(\beta_1,\beta_2,\cdots,\beta_i)h(\chi_1,\chi_2,\cdots,\chi_i) = \\ &h(\alpha_1,\alpha_2,\cdots,\alpha_i)h(\chi_1,\chi_2,\cdots,\chi_i)h(\beta_1,\beta_2,\cdots,\beta_i) = \\ &h(\beta_1,\beta_2,\cdots,\beta_i)h(\alpha_1,\alpha_2,\cdots,\alpha_i)h(\chi_1,\chi_2,\cdots,\chi_i) = \\ &h(\beta_1,\beta_2,\cdots,\beta_i)h(\chi_1,\chi_2,\cdots,\chi_i)h(\alpha_1,\alpha_2,\cdots,\alpha_i) = \\ &h(\chi_1,\chi_2,\cdots,\chi_i)h(\alpha_1,\alpha_2,\cdots,\alpha_i)h(\beta_1,\beta_2,\cdots,\beta_i) = \\ &h(\chi_1,\chi_2,\cdots,\chi_i)h(\beta_1,\beta_2,\cdots,\beta_i)h(\alpha_1,\alpha_2,\cdots,\alpha_i), \end{aligned}$$

for i=1, 2, 3 and $\alpha_i, \beta_i, \chi_i = 0,1,\cdots,n$.

Define H_r as the kernel vector obtained from H by removing the above redundancies and $\overline{M}_{3x}$ as the corresponding M_{3x}. An equivalent form for Equ. (7) is

$$\overline{M}_{3x}^T \cdot H_r = M_{3y_o} - \Delta \tag{9}$$

where the dimension of $\overline{M}_{3x}$, H_r and M_{3y} are respectively $s_r \times (l_2 - l_1 + 1)(l_4 - l_3 + 1)$, $s_r \times 1$ and $(l_2 - l_1 + 1)(l_4 - l_3 + 1) \times 1$. Here s_r is given by

$$s_r = \sum_{\substack{i,j,k=1 \\ i \leq j \leq k}}^{3} q_{rijk} \tag{10}$$

where

$$q_{rijk} = \begin{cases} \frac{q_i(q_i+1)(q_i+2)}{6} & \forall i = j = k \\ q_i q_j q_k & \forall i \neq j \neq k \\ \frac{q_k q_i (q_k+1)}{2} & \forall i \neq j = k \\ \frac{q_i q_k (q_i+1)}{2} & \forall i = j \neq k \end{cases} \quad (11)$$

To blindly identify the system of Equ. (1), we need to determine all unknown Volterra kernels in H_r uniquely based on Equ. (9). The following theorem provides sufficient conditions of blind identifiability for cubic systems with finite memory.

Theorem 1 The *full-sized* cubic system with memory length *n* in Equ. (1) is not blindly identifiable. However, systems with *sparse kernels* can be identified blindly if $\overline{M}_{3x}$ is of full row rank and

(i) $s_r \leq \frac{n(3n-1)}{2}$ in a noise-free situation or

(ii) $s_r \leq \frac{3n^2+5n+2}{2}$ in a noisy situation.

Proof. If the matrix $(\overline{M}_{3x} \cdot \overline{M}_{3x}^T)$ is nonsingular, the non-trivial solution of H_r in Equ. (9) is obtained by

$$H_r = (\overline{M}_{3x} \cdot \overline{M}_{3x}^T)^{-1} \cdot \overline{M}_{3x} \cdot (M_{3y_o} - \Delta) \quad (12)$$

As we will see, however, for full-sized cubic system in Equ. (1), the solution given by Equ. (12) does not exist.

Substituting *y*(*t*) of Equ. (1) into Equ. (2) leads to

$$\begin{aligned} m_3^y(\tau_1,\tau_2) = \sum_{v1,v2,v3=1}^{3} \sum_{\substack{(i_1,i_2,\cdots,i_{l_1})=0 \\ (j_1,j_2,\cdots,j_{l_2}) \\ (k_1,k_2,\cdots,k_{l_3})}}^{n} \{h(i_1,i_2,\cdots,i_{v1}) \cdot h(j_1,j_2,\cdots,j_{v2}) \\ \cdot h(k_1,k_2,\cdots,k_{v3})\} \cdot E\{[x(t-i_1)x(t-i_2)\cdots x(t-i_{v1})] \\ \cdot [x(t-j_1-\tau_1)x(t-j_2-\tau_1)\cdots x(t-j_{v2}-\tau_1)] \\ \cdot [x(t-k_1-\tau_2)x(t-k_2-\tau_2)\cdots x(t-k_{v3}-\tau_2)]\} \end{aligned} \quad (13)$$

Recall the assumption that $\tau_1 \in [l_1, l_2]$ and $\tau_2 \in [l_3, l_4]$, $l_1, l_2, l_3, l_4 \geq 0$. For $l_1 = 0$ and $l_2 = n$ in Equ. (13), it follows that $m_3^y(n+\delta,\tau_2) = m_3^y(n,\tau_2)$ for $\delta = 1, 2, \cdots$. With the symmetry of TOM $m_3^y(\tau_1,\tau_2) = m_3^y(\tau_2,\tau_1)$, we have $l_3 = \tau_1$ and $l_4 = 2n$. Likewise, $m_3^y(\tau_1, 2n+\delta) = m_3^y(\tau_1, 2n)$. Thus the dimension of the matrix $\overline{M}_{3x}$ becomes $s_r \times l_r$ where $l_r = \frac{3n^2+5n+2}{2}$. It is quite clear that s_r as given by Equ. (10) is $s_r >> l_r$ and matrix $\overline{M}_{3x}$ is not of full row rank. It follows immediately from Equ. (2) that it is impossible to identify a *full-sized* cubic system given by Equ. (1) blindly. However, this result also implies that blind identifiability is possible for a subclass of *sparse* Volterra models that meets the relationship of $s_r \leq l_r$. Since the time lag pairs of TOM include terms with $\tau_1 = \tau_2 = 0$ and $\tau_1 = \tau_2$, the obtained kernel estimates are biased due to effect of the Gaussian noise in this case, namely, $\hat{H}_r = (\overline{M}_{3x} \cdot \overline{M}_{3x}^T)^{-1} \cdot \overline{M}_{3x} \cdot (M_{3y_o} - \Delta)$ where $\Delta \neq 0$.

For example, consider the sparse cubic model $y(t) = h(k_0)x(t-k_0) + h(k_1,k_1)x^2(t-k_1) + h(k_2,k_2,k_2) \cdot x^3(t-k_2)$, where $k_2 \geq k_1 \geq k_0$ [5], it is straightforward to show that $s_r = 10$, $l_r = \frac{3k_2^2+5k_2+2}{2}$. So when $k_2 \geq 2$ and rank($\overline{M}_{3x}$)$= s_r$, the model can be identified blindly in a noisy environment. It should be noted that in SOM domain, the relevant matrix is always singular and a unique solution of $\hat{H}_r$ cannot be achieved.

On the other hand, after removing terms with $\tau_1 = \tau_2 = 0$ and $\tau_1 = \tau_2$ from Equ. (13), l_r reduces to $l_r' = \frac{n(3n-1)}{2}$. In this case, $\hat{H}_r = (\overline{M}_{3x} \cdot \overline{M}_{3x}^T)^{-1} \cdot \overline{M}_{3x} \cdot M_{3y_o}$; i.e., the effect of Gaussian noise has been completely suppressed. Thus, if $k_2 \geq 3$, i.e., $s_r < l_r' = \frac{k_2(3k_2-1)}{2}$ and rank($\overline{M}_{3x}$)$= s_r$, the above model can be blindly identified without any influence of Gaussian noise. Theorem 1 has thus been proven.□

4. SIMULATIONS

Consider the following sparse cubic systems

$$y_o(t) = x(t-1) + 0.8x^2(t-2) - 0.4x^3(t-3) + n_w(t) \quad (14)$$

where *h*(1)=1, *h*(2,2)=0.8, and *h*(3,3,3)=-0.4. According to Theorem 1, $s_r = 10 < l_r = 22$ and rank($\overline{M}_{3x}$)$= s_r$ suggesting that the system can be identified blindly in TOM domain. However, as can be seen in [6], this is not the case in the SOM domain. Generally an i.i.d. exponentially distributed random sequence {*x*(*t*)} with zero mean is generated as the input signal. We assume the model output observations are corrupted by white Gaussian noise $\{n_w(t)\}$, where signal-to-noise-ratio $SNR = 10\log_{10}\frac{E(y^2(t))}{E(n_w^2(t))}$=10 dB. Expectation operations in Equ. (9) are approximated by their estimation. The asymptotically unbiased estimates of $m_3^{y_o}(\tau_1,\tau_2)$ in (8d) are obtained as [8]

$$\hat{m}_{3(k)}^{y_o}(\tau_1,\tau_2) = \frac{1}{M(\tau_1,\tau_2)} \cdot$$

$$\cdot \sum_{i=1+\max(0,-\tau_1,-\tau_2)}^{M-\max(0,\tau_1,\tau_2)} y_{o(k)}(i)\, y_{o(k)}(i+\tau_1)\, y_{o(k)}(i+\tau_2)$$

$$i=1,\ 2,\ \ldots,\ M \qquad (15a)$$

$$M(\tau_1,\tau_2)=[M-\max(0,\tau_1,\tau_2)]\text{-}\max(0,-\tau_1,-\tau_2) \qquad (15b)$$

$$\hat{m}_3^{y_o}(\tau_1,\tau_2)=\frac{1}{K}\sum_{k=1}^{K}\hat{m}_{3(k)}^{y_o}(\tau_1,\tau_2) \qquad (15c)$$

where the output data with N_Ω length are segmented into K records of M samples each. These N_Ω samples are divided into K (4 or 16) records, each containing M (256 or 1024) samples. To verify the theoretical analysis, 50 Monte Carlo runs are performed. The final results, kernel estimates $\hat{h}(1)$, $\hat{h}(2,2)$, $\hat{h}(3,3,3)$ as well as the corresponding standard deviations are summarized in Table 1. It is clear that the estimated sparse cubic kernels are very close to their true values even under a lower SNR environment.

5. CONCLUSION

In this paper, TOM based blind identification of cubic systems is considered. Although *full-sized* cubic kernels cannot be estimated based on the TOM of the system output, some classes of *sparse* systems are blindly identifiable if the number of nonzero kernels meets some conditions involving the memory length *n*. The contribution of the proposed approach is that a larger subset of sparse cubic systems can be identified blindly as compared to the SOM-based approach due to the fact that additional statistical knowledge is available. Simulation results verify the effectiveness of the proposed method for blind identification of sparse nonlinear cubic systems.

REFERENCES

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

[2] P. Bondon, “Blind identifiability of a quadratic stochastic system”, *IEEE Trans. on Information Theory*, Vol. 41, No. 1, 1995, pp. 245-254.

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

[4] N. Kalouptsidis, P. Koukoulas, V. J. Mathews, “Blind identification of bilinear systems”, *IEEE Trans. on Signal Procsssing*, Vol. 51, No. 2, 2003, pp. 484-499.

[5] L. Tan, J. Jiang, “Adaptive Volterra Filters for Active Control of Nonlinear Noise Processes”, *IEEE Trans. on Signal Processing*, vol. 49, pp. 1667-1675, Aug. 2001.

[6] H.-Z. Tan, T. Aboulnasr, “Blind identifiability of the third-order Volterra systems”, *Proc. ICASSP 2003*, Hong Kong, 2003, pp. 665-668.

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

[8] C. L. Nikias, J. M. Mendel, “Signal processing with higher-order spectra”, *IEEE Signal Processing Magazine*, Vol. 10, No. 3, 1993, pp. 10-37.

Table 1 True and estimates of cubic model
with their means (standard deviations) under 50 Monte Carlo runs

	$SNR = 10$ dB	
	$N_\Omega = 1024$	$N_\Omega = 16384$
True $h(1)$	1	1
TOM-based estimate $\hat{h}(1)$	1.0105 (0.1969)	1.0014 (0.0932)
SOM-based estimate $\hat{h}(1)$ [5]	N/A	N/A
True $h(2,2)$	0.8	0.8
TOM-based estimate $\hat{h}(2,2)$	0.8171 (0.1884)	0.7969 (0.0854)
SOM-based estimate $\hat{h}(2,2)$ [5]	N/A	N/A
True $h(3,3,3)$	-0.4	-0.4
TOM-based estimate $\hat{h}(3,3,3)$	-0.4022 (0.0411)	-0.3999 (0.0246)
SOM-based estimate $\hat{h}(3,3,3)$ [5]	N/A	N/A