A NON PERSISTENT EXCITATION INPUT SEQUENCE FOR VOLTERRA AND NARMAX FILTER REPRESENTATION

J.-P. Costa, E. Thierry and T. Pitarque

Laboratoire I3S, CNRS-UNSA 2000 route des Lucioles, 06410 Biot France costa@i3s.unice.fr, pitarque@i3s.unice.fr

ABSTRACT

This paper presents an original approach for nonlinear Volterra and NARMAX filter identification. We propose here to use a non persistent excitation of input sequence to excite few terms of the Volterra and NARMAX filter, which must satisfy a set of physical constraints. Expressions have been developped to evaluate the accurate number of excited terms of the Volterra and NARMAX filter.

1. INTRODUCTION

The necessary characteristics required for linear filter identification are well understood [1]; however, the requirements for nonlinear filter excitation have not been yet studied in great details.

This paper deals with input signals for the identification of nonlinear discrete time systems via a Volterra and a NAR-MAX (Nonlinear Autoregressive Moving Average model with eXogenous inputs) filter. The Volterra filter is attractive because it is a straightforward generalization of the linear system description and the behaviour of many physical systems can be described with a Volterra filter [2]. In a similar manner, the NARMAX filter can, by including information from both lagged inputs and outputs, provide a very concise representation for nonlinear physical systems [3].

It has been shown in [2] that the q-level RMS (Random Multilevel Sequence) input sequence guarantees the persistence of excitation (PE) condition of the Volterra filter with nonlinearities of order D with q > D + 1. Unfortunately for a high order Volterra and NARMAX filter, the increasing number of terms leads to a ill-conditioned data matrix (presented in section 2.).

A recent approach developped in [4] and [5] uses a non PE input sequence for Volterra filter identification. The goal of this approach is to use a input sequence which excite as few as possible terms of the Volterra filter (section 3. and 4.), and so the number of parameters of the Volterra filter to be estimated is reduced. Moreover, this approach has been successfully validated on real underwater sonar signals for a nonlinear Jammer application [6].

The contribution of this paper is to extend this approach for the NARMAX filter and to define expressions which evaluate the accurate number of excited terms of these compact Volterra and NARMAX filters for a particular BPSK sequence (section 5.). Illustrative examples are also proposed in this section.

2. PERSISTENCE OF EXCITATION CONDITION FOR THE VOLTERRA FILTER

2.1. Volterra filter

The discrete-time invariant Volterra filter with memory length m_x and order of nonlinearity D, with L samples, is defined by

$$y_k \stackrel{\triangle}{=} \sum_{i=1}^D \sum_{j_1, \cdots, j_n=0}^{m_x - 1} h_i(j_1, \cdots, j_n) x_{k-j_1} \cdots x_{k-j_n} \quad (1)$$

or with a matrix representation $\boldsymbol{y}_k = \boldsymbol{X}_k \theta + \eta_k$, where $\boldsymbol{y}_k = [y_k, \cdots, y_{k-L+1}]^{t}$ is the observed output sequence associated with the input sequence \boldsymbol{x}_k, θ is the vector of the Volterra parameters and η_k is a noise sequence and \boldsymbol{X}_k is the $(L \times \mathbf{p})$ data matrix defined by $\boldsymbol{X}_k \stackrel{\Delta}{=} [\mathbf{x}_k, \cdots, \mathbf{x}_{k-L+1}]^{t}$

with

$$\mathbf{x}_{k} \stackrel{\Delta}{=} [x_{k-j_{1}} \ x_{k-j_{1}} x_{k-j_{2}} \ \dots \ x_{k-j_{1}} x_{k-j_{2}} \dots x_{k-j_{D}}]^{\mathsf{t}}$$

for $j_{1}, \dots, j_{D} = 0 \dots m_{x} - 1$.

If $X_k^t X_k$ is full rank, then the least squares estimate of the filter parameters is given by

$$\hat{\theta} = (\boldsymbol{X}_{k}^{\mathsf{t}} \boldsymbol{X}_{k})^{-1} \boldsymbol{X}_{k}^{\mathsf{t}} \boldsymbol{y}_{k}$$
(2)

The number of unique products in Eq. (1) is given by

$$\mathbf{p}_{V} = \frac{(D+m_{x})!}{D!m_{x}!} - 1$$
(3)

For D = 3, $m_x = 14 \mathbf{p}_V = 679$.

2.2. Persistence of excitation condition

Once, the Volterra parameters of the system described in Eq. (1) are identified, the PE condition can be defined as in [2] by :

Definition 1 Assume that the sequence x_k is a stationary random process. If the correlation matrix R_x of the Volterra filter

$$R_{\mathbf{X}} \stackrel{\triangle}{=} E[\mathbf{x}_k \mathbf{x}_k^{\mathsf{t}}] \tag{4}$$

exists and is nonsingular, then the input sequence x_k is said to be PE of degree m_x and order D.

The RMS input signal is usually used as a PE sequence because the RMS sequence defined below satisfy the PE condition if $q \ge D + 1$ [2]

Definition 2 Let x_k be an i.i.d. sequence which takes on a finite number of distinct values l_1, l_2, \dots, l_q with corresponding probabilities p_1, p_2, \dots, p_q $\sum_{i=1}^{q} p_i = 1$. Then x_k is called a q-level RMS,

However for high order Volterra filters the increasing number \mathbf{p}_V leads inevitably to a ill-conditioned matrix $R_{\mathbf{X}}$. Direct inversion of the data matrix $\mathbf{X}_k^{t} \mathbf{X}_k$ becomes numerically ill-posed. To reduce this number \mathbf{p}_V , an alternative solution is proposed with a non persistence of excitation input sequence.

The originality of this approach is to use a cyclostationnary Binary Phase Shift Keying (BPSK) input sequence which takes a finite number of distinct values. This sequence doesn't respect the PE condition because it's not an i.i.d. sequence. Moreover this BPSK sequence excite few terms of the Volterra filter.

3. NON PERSISTENCE OF EXCITATION INPUT FOR THE VOLTERRA FILTER

This approach proposes a BPSK input sequence which allows to reject Volterra filter terms by the evaluation of R_x ; a minimal filter is constructed with only the excited terms.

3.1. The BPSK input sequence

The BPSK signal is a cyclostationary signal (see [7] for its properties). Taking the sampling frequency as an integer multiple of the baud rate of a continuous BPSK signal, the discrete-time BPSK signal x_k is expressed as

$$x_k = \sum_{m=0}^{\infty} a_m p(k - mR) \cos(2\pi f_0 k) \quad k = 0, 1, 2 \dots$$
(5)

where R corresponds to the number of points per keying interval, f_0 is the reduced carrier frequency, $\{a_m\}$ is an i.i.d. sequence equal to ± 1 and p(k) is the rectangular pulse such that p(k) = 1 for $k = 0, \ldots, R - 1$ and 0 otherwise.

3.2. Extension of the definition of the correlation matrix

In order to obtain the minimal Volterra filter, we must extend the definition of the correlation matrix for an asymptotically stationary sequence [4],

$$R_{\mathbf{x}} \stackrel{\triangle}{=} \lim_{T \to \infty} \frac{1}{T} \sum_{l=0}^{T-1} E[\mathbf{x}_{l+k} \mathbf{x}_{l+k}^{\mathsf{t}}]$$
(6)

The excited terms of the Volterra filter $\mathbf{x}'_k (\dim(\mathbf{x}'_k) \ll \dim(\mathbf{x}_k))$ correspond to a full rank sub-matrix $R_{\mathbf{x}'}$. The theoretical calculation of this correlation matrix has been studied in the white case in [4] and for all BPSK sequences in [5].

3.3. Construction of the minimal Volterra filter

In order to evaluate the minimal Volterra model the following procedure is used :

- 1. the correlation matrix R_X for a cyclostationary BPSK sequence is computed,
- the rank of R_x allows to determine a parsimonious vector x'_k which only takes terms of x_k associated to non zero eigenvalues of R_x,
- 3. the reduced correlation matrix $R_{\mathbf{x}'}$ is now full rank and the minimal Volterra model is given by the \mathbf{x}'_k terms. The evaluation of \mathbf{x}'_k is computed iteratively.

4. INFLUENCES OF THE CARRIER FREQUENCY

In the sequel the input sequence is assumed to have a large bandwidth.

4.1. Frequency representation

Signal Bandwidth B of the BPSK input sequence is inversely proportional to R (for R = 1, BPSK signal is white).

To generate a BPSK sequence with $B = 2 \Delta_f$ ($B \in [f_1 - \Delta_f, f_1 + \Delta_f]$) centered on the frequency f_1 , we must take $R = 1/\Delta_f$ and $f_0 = f_1$. Figure 1 represents the power spectral density of the BPSK sequence for $f_0 = f_1 = 0.20$ and for R=1, 5 and 10.

We can now show the infuence of f_0 on the finite number of distinct values of the sequence.



Figure 1: Frequency representation of a BPSK sequence : $f_0 = 0.2$.

4.2. Number of distinct values of the BPSK input sequence

The finite number of distinct values of the BPSK input sequence is directly dependent of the carrier frequency. Figure 2 represents the evolution of the number of levels of the sequence when f_0 varies and takes the following values $f_0 = 0.01, \ldots, 0.49$.



Figure 2: Number of distinct values of the BPSK input sequence versus the reduced carrier frequency.

The number of distinct values is comprised between 3 levels for $f_0 = 0.25$ and 50 levels for $f_0 = 0.09$, $f_0 = 0.13$ and $f_0 = 0.19$. The number of distinct values also plays a role in the number of excited terms of the Volterra filter.

4.3. Relationship between the number of distinct values and the number of excited terms of the Volterra filter

Figure 3 shows the evolution of the number of excited terms when the number of points per keying interval (*R*) increases and for three differents carried frequencies, using the procedure described in §3.3. We can see that best performances are obtained with $f_0 = 0.25$ (3 levels) and with $f_0 = 0.20$ (6 levels).



Figure 3: Number of excited terms of a Volterra filter versus the keying interval. Volterra filter D = 3 and $m_x = 14$: 679 initial terms

For R = 1 the BPSK input sequences are white. The white BPSK input sequence with $f_0 = 0.25$ (3 levels) corresponds to the more efficient non PE sequence (see figure 3). The number of terms of the minimal Volterra filter, in this particular case, is evaluated in the following section.

5. EVALUATION OF THE NUMBER OF EXCITED TERMS

The BPSK sequence with R = 1 and $f_0 = 0.25$ can be expressed as $x_k = \pm 1 \cos(k\pi/2)$. This sequence contains three distincts values ± 1 and 0.

5.1. Volterra filter

Using this remark we can evaluate the number of excited terms of the Volterra filter. Let $\mathbf{p}'_V(D=2)$ and $\mathbf{p}'_V(D=3)$ respectively be the number of excited terms for a second and third polynomial order of the Volterra filter. These expressions can be defined as the following form :

$$\begin{array}{l} \mathbf{0} \quad D = 2 \\ \mathbf{\bullet} \ m_x \text{ even :} \end{array}$$

$$\mathbf{p}_V'(D=2) \stackrel{\triangle}{=} \left(\frac{m_x}{2}\right)^2 + \frac{m_x}{2} + 2$$



Figure 4: Number of parameters \mathbf{p}' to be estimated for D = 2.

• m_x odd :

$$\mathbf{p}_V'(D=2) \stackrel{\triangle}{=} (\frac{m_x}{2})^2 + \frac{m_x}{2} + \frac{9}{4}$$

2 D = 3

• m_x even :

$$\mathbf{p}_V'(D=3) \stackrel{\triangle}{=} \frac{m_x^3}{24} + \frac{20}{24}m_x + 2$$

• m_x odd :

$$\mathbf{p}_V'(D=3) \stackrel{\triangle}{=} \frac{m_x^3}{24} + \frac{23}{24}m_x + 2$$

Figures 4 and 5, respectively D = 2 and D = 3, show the evolution of the number of excited terms (i.e. number of parameters to be estimated) when the memory length m_x varies ($m_x = 1, \ldots, 30$) for the BPSK input sequence ($f_0 = 0.25$ and R = 1) and for a PE input sequence using Eq. (3). The BPSK input sequence allows to reject a lot of terms of the Volterra filter because only few terms are excited (due to some aliasing in the model structure). For $m_x = 30$ and for D = 2 (D = 3) only 242 terms (1152 terms) are excited for a initial base composed of 495 terms (5455 terms).

5.2. The NARMAX filter

5.2.1. Input/Output relationship

The NARMAX filter is a general parametric filter (including Volterra and bilinear models) [8]. This filter is defined by

$$y_k = f(y_{k-1}, \cdots, y_{k-m_y}, x_k, \cdots, x_{k-m_x}, \\ \epsilon_{k-1}, \cdots, \epsilon_{k-m_{\epsilon}}) + \epsilon_k$$



Figure 5: Number of parameters $\mathbf{p}'(m_x)$ to be estimated for D = 3.

where ϵ_k is the prediction error sequence and m_{ϵ} is the memory length prediction sequence.

A second order NARMAX model without prediction error sequence is described by

$$y_{k} = \sum_{i=0}^{m_{x}-1} a_{i}x_{k-i} + \sum_{j=1}^{m_{y}} b_{j}y_{k-j} + \sum_{i=0}^{m_{x}-1} \sum_{j=1}^{m_{y}} c_{ij}x_{k-i}y_{k-j}$$
$$+ \sum_{i=0}^{m_{x}-1} \sum_{j=0}^{m_{x-1}} d_{ij}x_{k-i}x_{k-j} + \sum_{i=1}^{m_{y}} \sum_{j=1}^{m_{y}} e_{ij}y_{k-i}y_{k-j}$$

The number of parameters of this model is

$$\mathbf{p}_N = m_x + m_y + m_x m_y + \frac{1}{2}(m_x + m_y + m_x^2 + m_y^2)$$
(7)

For high order NARMAX filter, the number of terms increases dramatically and this induces numerical problems.

5.2.2. Definition

We can evaluated the expressions of the accurate number of excited terms of the NARMAX filter using a similar procedure. Let $\mathbf{p}'_V(D = 2)$ and $\mathbf{p}'_N(D = 3)$ respectively be the number of excited terms for a second and third polynomial order of the NARMAX filter. These expressions take the following form :

0
$$D = 2$$

•
$$m_x$$
 even :

$$\mathbf{p}_N'(D=2) \stackrel{\triangle}{=} \left(\frac{m_x}{2}\right)^2 + \frac{m_x}{2} + 2 + m_y + m_y m_x + \frac{(m_y+1)m_y}{2}$$

• m_x odd :

$$\mathbf{p}_N'(D=2) \stackrel{\triangle}{=} \left(\frac{m_x}{2}\right)^2 + \frac{m_x}{2} + \frac{9}{4} + m_y + m_y m_x + \frac{(m_y+1)m_y}{2}$$

 $\begin{array}{l} \mathbf{2} \quad D = 3 \\ \bullet \ m_x \text{ even} : \end{array}$

$$\mathbf{p}_{N}'(D=3) \triangleq \frac{m_{x}^{3}}{24} + \frac{20}{24}m_{x} + 2 + \frac{(3+m_{y})!}{3!m_{y}!} + m_{x}m_{y} + m_{x}\frac{(m_{y}+1)m_{y}}{2} + m_{y}((\frac{m_{x}}{2})^{2} - \frac{m_{x}}{2} + 2)$$

• m_x odd :

$$\mathbf{p}_{N}'(D=3) \stackrel{\triangle}{=} \frac{m_{x}^{3}}{24} + \frac{23}{24}m_{x} + 2 + \frac{(3+m_{y})!}{3!m_{y}!} + m_{x}m_{y} + m_{x}\frac{(m_{y}+1)m_{y}}{2} + m_{y}((\frac{m_{x}}{2})^{2} - \frac{m_{x}}{2} + \frac{9}{4})$$

Note that the generalization for high polynomial order of the Volterra and NARMAX filter is not straightforward. It's worthwhile noticing that when the memory length m_x of the Volterra and NARMAX filter increases, the reduction of the number of initial terms is far more important.

5.3. Illustrative examples

EXAMPLE $n^{\circ}1$: let $y_k = x_k + 0.2x_{k-1}^2 - 0.3x_{k-1}x_k - 0.3x_{k-1}^3$ be the non linear Input-Ouput relationship of an unknown structure. In a first step, we identify this system by a 3-level RMS PE input sequence and in the same time with a BPSK non PE input sequence both using a D = 2 and $m_x = 2$ Volterra filter. This Volterra filter contains five terms : $[x_k, x_{k-1}, x_k^2, x_{k-1}^2, x_k x_{k-1}]$. Note that the cubic term x_{k-1}^3 of Input-Ouput relationship is not represented in this Volterra filter. Only 4 terms are excited by the BPSK input sequence : $[x_k, x_{k-1}, x_k^2, x_{k-1}^2]$.

In a second step two new input sequences are generated (RMS and BPSK) in order to compare the PE and non PE Volterra filter identification. A commonly measure of identification error is the normalized mean square error Erdefined by

$$Er = \frac{\sum_{i=1}^{L} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{L} y_i^2}.$$

One sequence is used for the identification of this system and 50 realizations are used to compute the normalized mean square error.

We obtain $Er = +7 \, dB$ for the PE input sequence and $Er = -33 \, dB$ for the non PE input sequence. This big difference of performance is due to the fact that the x_{k-1}^3

term is equivalent to the x_{k-1} term for the BPSK input sequence, so this term is contained in the minimal Volterra filter (4 terms).

EXAMPLE $n^{\circ}2$: let $y_k = x_k + 0.2x_{k-2} - 0.1x_{k-1}y_{k-1}^2$ be the non linear Input-Ouput relationship of an unknown structure. The cross-term $x_{k-1}y_{k-1}^2$ is not represented by the Volterra filter ($m_x = 2$ and D = 2). We obtain Er = $-8.9 \ dB$ for the RMS sequence and $Er = -14.1 \ dB$ for the BPSK sequence. Best identification results are obtained by the non PE sequence ($\Delta Er = 5 \ dB$).

6. CONCLUSION

An alternative method of the classical polynomial filter identification is proposed for some specific applications where the input sequence has the same properties than the reference sequence used for the identification of the unknown system : it is the case for the jammer application.

New formulae have been evaluated to compute the accurate number of the excited terms for a particular BPSK sequence for the Volterra and NARMAX filters.

7. REFERENCES

- T. Soderstrom and P. Stoica. System Identification. Cambridge, UK, Prentice Hall, 1989.
- [2] R. D. Nowak and B. D. Van Veen. "Random and Pseudorandom Inputs for Volterra Filter Identification". *IEEE trans. on Signal Processing*, 42(8), august 1994.
- [3] S.A. Billings and W. S. Voon. "A prediction-error and stepwise-regression estimation algorithm for nonlinear systems". *Int. J. Control*, 44(3):803–822, 1986.
- [4] J-P. Costa, P. Rostaing, and T. Pitarque. "Filtrage adaptatif non linéaire utilisant une entrée cyclostationnaire". Proc. Sixteenth GRETSI symposium on signal and image processing, Gretsi97, Grenoble, France, 2:1387–1390, 1997.
- [5] J-P. Costa and P. Rostaing. "Nonlinear Adaptive Filtering For a Jammer Application". Proc. 13th International conference on Digital Signal Processing, DSP97, Santorini, Greece, 1997.
- [6] J-P. Costa, P. Rostaing, and T. Pitarque. "Nonlinear filtering for a new nimble Jammer". Oceans'98, IEEE/OES Conference, 3:1491–1495, 1998.
- [7] W. A. Gardner. Introduction to Random Processes with Application to Signal & Systems. MC Graw-Hill, 1986.
- [8] S. Chen and S.A. Billings. Representation of nonlinear systems : the NARMAX model. *Int. J. Control*, 49(3):1013– 1032, 1989.