XXth National Conference
CIRCUIT THEORY AND ELECTRONIC NETWORKS
Kołobrzeg, Poland, October 21-24, 1997

Włodzimierz Greblicki
Zygmunt Hasiewicz
Institute of Engineering Cybernetics
Technical University of Wrocław
Wybrzeże S. Wyspiańskiego 27, 50-370 Wrocław

Identification of Non-linear Systems by Orthogonal Series

Key words: system identification, non-parametric approach, orthogonal series

ABSTRACT

The paper deals with recovering non-linearities in memoryless, Hammerstein, and Wiener systems. The *a priori* information about the systems is non-parametric. To estimate non-linear characteristics of examined systems, a class of adequate non-parametric algorithms is proposed. This class is based on orthogonal expansions and covers the algorithms using both classical orthogonal series and orthogonal wavelets.

1. INTRODUCTION

In this paper, we examine the non-parametric approach to system identification, suited to the case when our *a priori* information about the system to be identified is much smaller than in parametric inference ([1]). We propose a class of algorithms to estimate non-linearities in memoryless, Hammerstein, and Wiener systems. The identification algorithms are based on the idea of orthogonal expansions and are presented in a unified framework which covers two useful subclasses: *a*) classical orthogonal systems, and *b*) orthogonal wavelets with compact support. The paper is a companion to [6], where kernel-type non-parametric identification algorithms were presented.

2. ORTHOGONAL SYSTEMS

In the case of classical orthogonal systems, we shall use various complete families of orthonormal functions $\{\varphi_k; k = 0, 1, 2, ...\}$ in the respective sets D (being specified below), and we assume that all φ_k's vanish outside D and possess the properties:

$$|\varphi_k(u)| \le c(u)(k+1)^{\alpha}, \tag{2.1a}$$

some $c(u)$ independent of k,

$$\sup_{u \in D} |\varphi_k(u)| \le d_1(k+1)^{\beta}, \tag{2.2a}$$

some d_1 independent of k, and

$$\sup_{u \in D} |\varphi_k'(u)| \le d_2(k+1)^{\gamma}, \tag{2.3a}$$

some d_2 independent of k. Examples of orthogonal systems satisfying (2.1a)-(2.3a) are: *1.* the trigonometric series - orthonormal in the interval $D = [-\pi, \pi]$, for which the conditions hold for $\alpha = \beta = 0$ and $\gamma = 1$, *2.* the Legendre series - orthonormal in the interval D=[-1,1] for which we have $\alpha = \beta = 1/2$ and $\gamma = 5/2$, and *3.* the Hermite series - orthonormal in the whole real line, $D = R$, for which $\alpha = -1/4$, $\beta = -1/12$ and $\gamma = 5/12$ - see [9] and [10] for the details.

In the case of wavelets, the orthogonal system $\{\psi_{kl};\ k,l = ...,-1,0,1,...\}$ is of the form

$$\psi_{kl}(u) = 2^{k/2}\,\psi(2^k u - l)$$

and the particular basis functions are constructed in an automatic way from a single initial function $\psi(u)$ (called a "mother" wavelet) by scaling (index k) and shifting (index l). They constitute a series of functions orthonormal in the set $D = R$. We assume that $\psi(u)$ has compact support (i.e. equals zero outside some compact set, $[s_1, s_2]$ say), and moreover that

$$|\psi_{kl}(u)| \le c(u)\,2^{\alpha k}, \qquad \text{all } l \tag{2.1b}$$

some $c(u)$ independent of k,

$$\sup_{u\in R} |\psi_{kl}(u)| \le d_1\,2^{\beta k}, \qquad \text{all } l \tag{2.2b}$$

some d_1 independent of k, and

$$\sup_{u\in R} |\psi'_{kl}(u)| \le d_2\,2^{\gamma k}, \qquad \text{all } l \tag{2.3b}$$

some d_2 independent of k. Notice that (2.1b)-(2.3b) correspond to the conditions (2.1a)-(2.3a) with 2^k in the role of k+1 and ψ_{kl} in place of φ_k. Examples of wavelets satisfying the above requirements are *1.* the Haar wavelets - for which $\alpha = \beta = 1/2$ in (2.1b)-(2.2b) ((2.3b) does not hold since they are not differentiable), and *2.* the Daubechies wavelets - fulfiling (2.1b)-(2.3b) with $\alpha = \beta = 1/2$ and $\gamma = 3/2$, for smooth versions - see [2], [11].

3. MEMORYLESS SYSTEM IDENTIFICATION

Consider first a memoryless system shown in Fig. 1. Its input is a stationary white random noise $\{U_n; n = ..., -1, 0, 1, 2, ...\}$ with finite variance. The probability density of U_n exists and is denoted by f. The system is disturbed by stationary random noise $\{Z_n; n = ..., -1, 0, 1, 2, ...\}$. The noise is an output of an asymptotically stable linear system (filter) driven by a stationary random noise $\{\xi_n; n = ..., -1, 0, 1, 2, ...\}$ with zero mean and finite variance. Processes $\{U_n\}$ and $\{\xi_n\}$ are mutually independent. The non-linear characteristic m of the system is completely unknown and has to be estimated from measurements $(U_0,Y_0), (U_1,Y_1), ..., (U_n,Y_n)$.

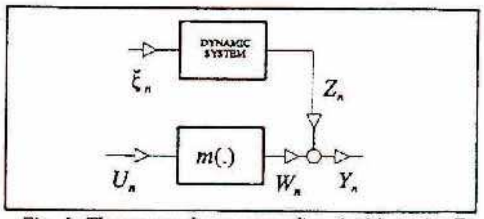

Fig. 1. The memoryless system disturbed by noise Z_n

Let us notice that

$$E\{Y_n | U_n = u\} = m(u), \tag{3.1}$$

i.e., that m is a regression function. Observe moreover that

$$m(u) = g(u)/f(u), \tag{3.2}$$

where $g(u) = E\{Y_n|U_n = u\}f(u)$, and f is the probability density of U_n. We assume that

$$|m(u)| \le a_1 |u| + a_2$$

some a_1 and a_2, and that $\int_D f^2(u)du < \infty$. Since then $\int_D g^2(u)du < \infty$, the nominator g and denominator f in (3.2) may be expanded in a series $\{\varphi_k ; k = 0,1,2,...\}$ of functions orthonormal in a set D

$$g(u) \sim \sum_{k=0}^{\infty} a_k \varphi_k(u) \quad \text{and} \quad f(u) \sim \sum_{k=0}^{\infty} b_k \varphi_k(u),$$

where

$$a_k = E\{Y_0 \varphi_k(U_0)\} \quad \text{and} \quad b_k = E\varphi_k(U_0),$$

which leads to the following natural estimate $\hat{m}(u)$ of $m(u)$:

$$\hat{m}(u) = \sum_{k=0}^{N(n)} \hat{a}_k \varphi_k(u) \Big/ \sum_{k=0}^{N(n)} \hat{b}_k \varphi_k(u), \tag{3.3a}$$

where $\hat{a}_k$ and $\hat{b}_k$ (estimates of a_k's and b_k's) are computed from the (random) observations $\{(U_i,Y_i);\ i = 0,1,...,n\}$ of the system input and output as follows:

$$\hat{a}_k = n^{-1} \sum_{i=0}^{n-1} Y_i \varphi_k(U_i) \quad \text{and} \quad \hat{b}_k = n^{-1} \sum_{i=0}^{n-1} \varphi_k(U_i) \tag{3.4a}$$

and $N(n)$ is a sequence of integers depending on the number of data n. Respectively, for compactly supported orthogonal wavelets $\{\psi_{kl};\ k,l = ...,-1,0,1,...\}$ we can write

$$g(u) \sim \sum_{|k|=0}^{\infty} \sum_{|l|=0}^{\infty} a_{kl}\, \psi_{kl}(u) \quad \text{and} \quad f(u) \sim \sum_{|k|=0}^{\infty} \sum_{|l|=0}^{\infty} b_{kl}\, \psi_{kl}(u),$$

where

$$a_{kl} = E\,\{Y_0\, \psi_{kl}(U_0)\} \quad \text{and} \quad b_{kl} = E\, \psi_{kl}(U_0)$$

and as a wavelet estimate $\hat{m}(u)$, we take

$$\hat{m}(u) = \sum_{|k|=0}^{N(n)} \sum_{l=L_{\min}}^{L_{\max}} \hat{a}_{kl}\, \psi_{kl}(u) \Big/ \sum_{|k|=0}^{N(n)} \sum_{l=L_{\min}}^{L_{\max}} \hat{b}_{kl}\, \psi_{kl}(u) \tag{3.3b}$$

where

$$\hat{a}_{kl} = n^{-1} \sum_{i=0}^{n-1} Y_i\, \psi_{kl}(U_i) \quad \text{and} \quad \hat{b}_{kl} = n^{-1} \sum_{i=0}^{n-1} \psi_{kl}(U_i) \tag{3.4b}$$

and

$$L_{\min} = [2^k u - s_2] + 1 \quad \text{and} \quad L_{\max} = [2^k u - s_1]$$

with $[v]$ the integer part of v.

We assume that the number sequence $N(n)$ satisfies the following restrictions:

$$N(n) \to \infty \text{ as } n \to \infty, \tag{3.5}$$

and

$$n^{-1} \sum_{k=0}^{N(n)} k^{2\alpha} \sum_{k=0}^{N(n)} k^{2\beta} \to 0 \quad \text{as} \quad n \to \infty \tag{3.6a}$$

in the case of standard orthogonal estimate (3.3a), or

$$n^{-1}\, 2^{2(\alpha+\beta)N(n)} \to 0 \quad \text{as} \quad n \to \infty \tag{3.6b}$$

in the case of wavelet estimate (3.3b).

Theorem 1. Let the orthogonal system satisfy (2.1a)-(2.2a) in the standard case, or (2.1b)-(2.2b) in the case of wavelets, and let the number sequence $N(n)$ fulfil (3.5) and (3.6a), or respectively (3.5) and (3.6b). Then

$$\hat{m}(u) \to m(u) \quad \text{as} \quad n \to \infty \quad \text{in probability}$$

at every point $u \in D$ at which $f(u) > 0$, and

$$\sum_{k=0}^{n} a_k \varphi_k(u) \to m(u)f(u) \quad \text{and} \quad \sum_{k=0}^{n} b_k \varphi_k(u) \to f(u) \quad \text{as} \quad n \to \infty \tag{3.7a}$$

for standard orthogonal system, or respectively for wavelets

$$\sum_{|k|=0}^{n} \sum_{l=L_{min}}^{L_{max}} a_{kl} \psi_{kl}(u) \to m(u)f(u) \text{ and } \sum_{|k|=0}^{n} \sum_{l=L_{min}}^{L_{max}} b_{kl} \psi_{kl}(u) \to f(u) \text{ as } n \to \infty \tag{3.7b}$$

Applying different orthogonal series specified in Section 2, we can obtain different particular versions of our algorithm and the above theorem. Examples are given in Section 6.

4. HAMMERSTEIN SYSTEM IDENTIFICATION

Consider a Hammerstein system shown in Fig. 2. The linear dynamic part is by assumption asymptotically stable. Properties of the exciting signal $\{U_n; n = ...-1, 0, 1, 2, ...\}$ and the noise $\{Z_n; n = ...-1, 0, 1, 2, ...\}$ are the same as in Section 3. So is the *a priori* knowledge about the nonlinear characteristic m. The goal is to identify the non-linearity m from input-output observations of the whole system (U_0,Y_0), (U_1,Y_1), ..., (U_n, Y_n).

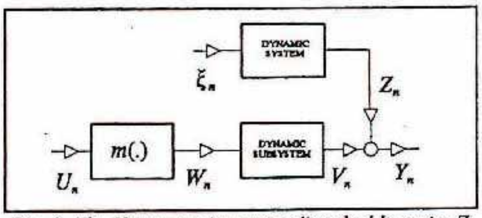

Fig. 2. The Hammerstein system disturbed by noise Z_n

The basic observation is that

$$E\{Y_n | U_{n-1} = u\} = c_1 m(u), \tag{4.1}$$

where c_1 is a constant (we have assumed for simplicity that EW_n=0; this is the case for ,e.g., m odd and f even). Thus, to recover $c_1 m(u)$ (scaled non-linearity), we can use the algorithm (3.3a) or (3.3b) with the modification that (see (3.4a))

$$\hat{a}_k = n^{-1} \sum_{i=1}^{n} Y_i \varphi_k(U_{i-1}) \quad \text{and} \quad \hat{b}_k = n^{-1} \sum_{i=1}^{n} \varphi_k(U_i) \tag{4.2a}$$

and the same for (3.4b):

$$\hat{a}_{kl} = n^{-1} \sum_{i=1}^{n} Y_i \psi_{kl}(U_{i-1}) \quad \text{and} \quad \hat{b}_{kl} = n^{-1} \sum_{i=1}^{n} \psi_{kl}(U_i) \tag{4.2b}$$

Theorem 2. Let m be odd and f even. Let all the assumptions of theorem 1 hold. Then, including (4.2a)-(4.2b),

$$\hat{m}(u) \to c_1 m(u) \quad \text{as} \quad n \to \infty \quad \text{in probability}$$

at every point $u \in D$ at which $f(u) > 0$, and (3.7a), respectively (3.7b), holds - with $c_1 m(u)$ instead of $m(u)$.

Observe that the convergence conditions are the same as in Section 3.

5. WIENER SYSTEM IDENTIFICATION

In the Wiener system, Fig. 3, we assume that additionally U_n's have a normal distribution with zero mean, the characteristic m is differentiable and invertible, i.e., m^{-1} exists, and that the

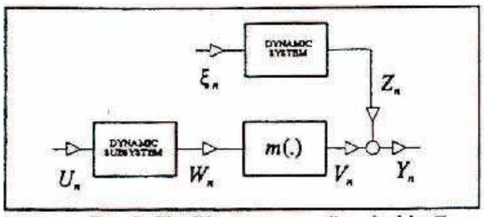

Fig. 3. The Wiener system disturbed by Z_n

density of the system output, $f_y(y)$, is square integrable. One can then observe that

$$E\{U_n | Y_{n+1} = y\} = c_2 m^{-1}(y), \tag{5.1}$$

where c_2 is a constant. Therefore, for estimating the scaled inverse $c_2 m^{-1}$, we can apply the same algorithm (3.3a) or (3.3b) - however at present with the system output y instead of the system input u as an argument in the formulae, and the following modifications of the weighting coefficients:

$$\hat{a}_k = n^{-1} \sum_{i=0}^{n-1} U_i \varphi_k(Y_{i+1}) \quad \text{and} \quad \hat{b}_k = n^{-1} \sum_{i=0}^{n-1} \varphi_k(Y_i) \tag{5.2a}$$

in the case of (3.4a), and

$$\hat{a}_{kl} = n^{-1} \sum_{i=0}^{n-1} U_i \psi_{kl}(Y_{i+1}) \quad \text{and} \quad \hat{b}_{kl} = n^{-1} \sum_{i=0}^{n-1} \psi_{kl}(Y_i) \tag{5.2b}$$

in the case of (3.4b).

Theorem 3. Let m be invertible and differentiable. Let all the assumptions of theorem 1 hold and let, moreover, the orthogonal system satisfy (2.3a), and (2.3b) for wavelets. Let the number sequence $N(n)$ fulfil in addition

$$n^{-1} \sum_{k=0}^{N(n)} k^{2\alpha} \sum_{k=0}^{N(n)} k^{\beta+\gamma} \to 0 \quad \text{as} \quad n \to \infty \tag{5.3a}$$

and for wavelets

$$n^{-1} 2^{(2\alpha+\beta+\gamma)N(n)} \to 0 \quad \text{as} \quad n \to \infty \tag{5.3b}$$

Then, with the above modifications,

$$\hat{m}(y) \to c_2 m^{-1}(y) \quad \text{as} \quad n \to \infty \quad \text{in probability}$$

at every point $y \in D$ at which $f_y(y) > 0$, and (3.7a), respectively (3.7b), holds - at present with y in the role of u, $c_2 m^{-1}(y)$ instead of $m(u)$, and $f_y(y)$ in place of $f(u)$.

Observe that convergence conditions are now more restrictive than that in Sections 3 and 4.

6. EXAMPLES

Let us consider memoryless system, and apply the trigonometric series. From Theorem 1 in Section 3, the facts referred in Section 2 and standard results concerning convergence of trigonometric expansions [9], we get

Corollary 1. If (3.5) holds and $N^2(n)/n \to 0$ as $n \to \infty$, then

$$\hat{m}(u) \to m(u) \quad \text{as} \quad n \to \infty \quad \text{in probability}$$

at every point $u \in [-\pi, \pi]$ at which $f(u) > 0$, and both m and f are differentiable.

Consider, in turn, Hammerstein system and let the Haar wavelets be implemented. Employing Theorem 2 in Section 4, the properties quoted in Section 2 and the known facts concerning convergence of the Haar expansions [8], we obtain

Corollary 2. If (3.5) holds and $2^{2N(n)}/n \to 0$ as $n \to \infty$, then

$$\hat{m}(u) \to c_1 m(u) \quad \text{as} \quad n \to \infty \quad \text{in probability}$$

at every point $u \in (-\infty, \infty)$ at which $f(u) > 0$, and both m and f are continuous.

We can similarly proceed for Wiener system and other orthogonal series.

7. FINAL REMARKS

Algorithms presented here are non-parametric, i.e., may be implemented when the *a priori* knowledge of the identified system is very small. They successfully recover non-linearities and need no modification for colored noise (for comparison see [3,4] where only white noise was admitted). Compared with the kernel algorithms ([6]), they require a much smaller amount of computer memory (now, a finite number ($\sim N(n)$) of coefficients in (3.3a) or (3.3b) is sufficient to be memorized instead of the whole set of measurement data (i.e. $n \gg N(n)$ input-output pairs), as it was the case for kernel estimates in [6]). Moreover, our algorithms - consisting generally in estimating regression - need only elementary computations. Let us emphasize that the algorithms converge at different sets of points in D, depending on the particular orthogonal system being implemented - see e.g. examples in Section 6. This is just the reason for which different orthogonal systems should be taken into account (Section 2). Proofs of the presented convergence theorems were omitted for shortness. They can be found in [5].

REFERENCES

[1] J. S. Bendat, *"Nonlinear System Analysis and Identification"*, Wiley, New York, 1990.

[2] I. Daubechies, *"Ten Lectures on Wavelets"*, SIAM Edition, Philadelphia, 1992.

[3] W. Greblicki, *"Nonparametric Orthogonal Series Identification of Hammerstein Systems"*, Int. J. Systems Sci., vol. **20**, pp. 2355-2367, 1989.

[4] W. Greblicki, *"Nonparametric Identification of Wiener Systems by Orthogonal Series"*, IEEE Trans. Automat. Control, vol. **AC-39**, pp. 2077-2086, 1994.

[5] W. Greblicki, Z. Hasiewicz, *"Non-linear System Identification in the Presence of Correlated Noise"*, Technical Report, series **PRE-37/95**, Wrocław, 1995.

[6] W. Greblicki, Z. Hasiewicz, *"Identification of Non-linear Systems Corrupted by Colored Noise"*, Proceedings XIX National Conference on Circuit Theory and Electronics Circuits, vol. **1**, pp. II/33-II/38, Kraków-Krynica, 1996.

[7] G. Härdle, *"Applied Nonparametric Regression"*, Cambridge University Press, Cambridge, UK, 1990.

[8] S. Kelly, M. Kon, and L. Raphael, *"Pointwise Convergence of Wavelet Expansions"*, Bull. Amer. Math. Soc., vol. **30**, pp. 87-94, 1994.

[9] G. Sansone, *"Orthogonal Functions"*, Interscience Publishers Inc., 1959.

[10] G. Szegö, *"Orthogonal Polynomials"*, Amer. Math. Soc. Coll. Publ., 1959.

[11] G. Walter, *"Wavelets and Other Orthogonal Systems with Applications"*, CRC Press Inc., Boca Raton, 1994.