MMAR'97 26-29 August 1997, Międzyzdroje, Poland
Fourth International Symposium on **Methods and Models in Automation and Robotics**

WAVELET IDENTIFICATION OF WIENER SYSTEMS

WŁODZIMIERZ GREBLICKI*, ZYGMUNT HASIEWICZ*

*Institute of Engineering Cybernetics, Technical University of Wrocław, ul. Janiszewskiego 11/17, 50-372 Wrocław, Poland, zhas@ict.pwr.wroc.pl

Abstract. The paper deals with the non-parametric identification of Wiener systems, based on the idea of orthogonal wavelet expansions. A general orthogonal wavelet algorithm for Wiener system identification is proposed, and the conditions for its in-probability convergence to the true non-linear characteristic under correlated noise are given. Then a particular version of the algorithm, using Daubechies wavelets, is discussed and a comparison with the algorithm employing classical Hermite polynomials is made. The advantages of applying wavelets to system identification are characterized.

Key Words. Wiener system, wavelets, system identification, non-parametric estimation.

1. INTRODUCTION

The paper deals with identification of Wiener systems. These are non-linear dynamical cascades, consisting of the linear dynamic block followed by a static non-linearity - Fig.1. Such tandem connections are often met in applied cybernetics as well as communication theory, image processing, or biocybernetics. Examples of such applications can, e.g., be found in [10].

For this class of systems, Greblicki proposed earlier the so-called kernel identification algorithms [6] and the orthogonal series algorithms, employing the trigonometric, Legendre, and Hermite orthogonal series [7]. In this paper we propose a class of algorithms applying compactly supported orthogonal wavelets. In the paper, the wavelet-based identification algorithm will be first investigated in a general form. We will show that under mild conditions the algorithm converges (in probability) to the (scaled) inverse of the unknown system nonlinearity (assumed to exist). Next, employing Daubechies wavelets, the particular realization of the algorithm will be discussed and compared with the classical Hermite series algorithm (both applicable on the whole real line). The advantages of applying wavelets to Wiener system identification will be pointed out.

The paper can be considered as a companion to [9] where Hammerstein systems, i.e., the reverse connection of blocks in Fig. 1, were investigated.

2. THE WIENER SYSTEM

The Wiener system, Fig. 1, consists of a linear dynamic and nonlinear memoryless subsystems connected in a cascade. We assume that the system is driven by a sequence $\{U_n; n = ..., -1, 0, 1, ...\}$ of independent zero-mean *Gaussian* random variables.

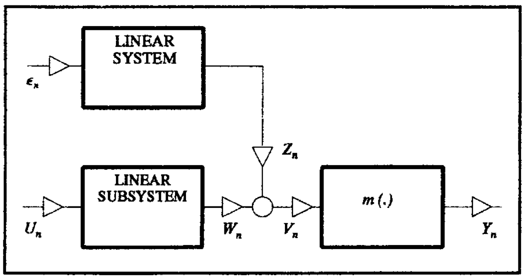

Fig. 1. The Wiener system

The linear subsystem is described by the following state-space equation:

$$\left.\begin{aligned} X_{n+1} &= AX_n + bU_n \\ W_n &= c^T X_n \end{aligned}\right\}$$

and is asymptotically stable. The output W_n of the subsystem is disturbed by stationary random noise $\{Z_n; n=...,-1,0,1,...\}$, yielding $V_n = W_n + Z_n$ as an input to the static part of the system. The noise Z_n is itself the output of an asymptotically stable linear system (filter), excited by $\{\epsilon_n; n=...,-1,0,1,...\}$ - a sequence of independent *Gaussian* random variables with zero mean. In result, disturbance $\{Z_n; n=...,-1,0,1,...\}$ is a sequence of correlated Gaussian random variables. The external inputs $\{\epsilon_n; n=...,-1,0,1,...\}$ and $\{U_n; n=...,-1,0,1,...\}$ are by assumption mutually independent. The characteristic of the nonlinear subsystem is denoted by m. Only qualitative preliminary information about unknown m is preassumed, which makes the class of admissible m broad and needs in the following non-parametric approach. Namely, we assume that m is a Borel measurable function defined in the entire real line, satisfying the following conditions:

Assumption 1.

m is differentiable and strictly monotonous

and

$$\sup_{v\in(-\infty,\infty)} |dm(v)/dv| \le \rho, \quad \text{some } \rho.$$

We also assume that V_n, the internal signal interconnecting both parts of the system, is not accessible for measurements and only the outer input-output signals of the whole system can be measured.

The goal is to identify the system non-linearity, i.e., to estimate the characteristic m, from input-output observations $(U_0,Y_0),(U_1,Y_1),...,(U_n,Y_n)$ of the overall system.

3. THE WAVELET-BASED IDENTIFICATION ALGORITHM

3.1. Background

Let us observe that under the quoted assumptions the pair (U_0,V_1) has a zero-mean Gaussian distribution with marginal variance σ_U^2 and σ_V^2, and the correlation coefficient $c^T b\sigma_U/\sigma_V$. Therefore, $E\{U_{n-1}|V_n=v\} = \lambda v$ where $\lambda = c^T b\sigma_U^2/\sigma_V^2$. Hence

$$E\{U_{n-1}|Y_n=y\} = \lambda m^{-1}(y) \tag{1}$$

with λ - an unknown constant, and thus estimating the regression in (1) (of input on output) we can recover in fact the inverse of the unknown characteristic of the nonlinear part of the system (which does exist) up to a scaling factor (provided that $\lambda \neq 0$). This is the basic observation.

3.2. Identification Algorithm

We shall give the algorithm estimating the scaled inverse $\lambda m^{-1}(y)$ (regression in (1)), using compactly supported orthogonal wavelets $\{\psi_{kl}; k,l = ...,-1,0,1,...\}$. They form an orthonormal basis for $L^2(R)$, and are easily generated from one single initial function $\psi(y)$ ("mother" wavelet) by scaling and shifting [2,16]:

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

with $\psi(y)$ vanishing outside some compact set $[s_1, s_2]$. To this end, let us notice that (1) can be rewritten as follows

$$\lambda m^{-1}(y) = g(y)/f(y),$$

where

$$g(y) = E\{U_{n-1}|Y_n=y\}f(y),$$

and f is the probability density of Y_n. On the density f, we impose the following

Assumption 2.

In the system,

$$\int_R f^2(y)dy < \infty.$$

Observing that moreover $\int_R g^2(y)dy < \infty$, the nominator g and denominator f of the above expression can be decomposed in a wavelet series, yielding

$$g(y) \sim \sum_{|k|=0}^{\infty}\sum_{l=L_{min}}^{L_{max}} a_{kl}\,\psi_{kl}(y)$$

and

$$f(y) \sim \sum_{|k|=0}^{\infty}\sum_{l=L_{min}}^{L_{max}} b_{kl}\,\psi_{kl}(y),$$

where

$$a_{kl} = E\{U_0\psi_{kl}(Y_1)\} \quad\text{and}\quad b_{kl} = E\,\psi_{kl}(Y_0)$$

and

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

with $[v]$ the integer part of v. This immediately results in the following natural, and convenient, wavelet-based estimate of the (scaled) inverse $\lambda m^{-1}(y) = g(y)/f(y)$:

$$\hat{v}(y) = \frac{\sum_{|k|=0}^{N(n)} \sum_{l=L_{min}}^{L_{max}} \hat{a}_{kl} \psi_{kl}(y)}{\sum_{|k|=0}^{N(n)} \sum_{l=L_{min}}^{L_{max}} \hat{b}_{kl} \psi_{kl}(y)} \qquad (2)$$

where

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

are the usual estimates of a_{kl} and b_{kl} above and are computed from (random) observations of the overall system input and output, and $N(n)$ is a sequence of integers depending on the number of data n.

3.3. Convergence

We assume the following

Assumption 3.

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

some $c(y)$ independent of k,

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

some d_1 independent of k, and

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

some d_2 independent of k. Using the above, for m satisfying Assumption 1 one can establish that ([8])

$$E\,\hat{a}_{kl} = a_{kl} \qquad (3a)$$

and

$$\operatorname{var}\hat{a}_{kl} = O(2^{2\beta k}/n) + O(2^{(\beta+\gamma)k}/n) \qquad (3b)$$

The same can be shown for $\hat{b}_{kl}$. Using these facts, the following general convergence theorem can be proved for the estimate (2) ([8]).

Theorem

Under the Assumptions 1-3, if

$$N(n) \to \infty \text{ as } n \to \infty, \qquad (4)$$

and

$$n^{-1}2^{2(\alpha+\beta)N(n)} \to 0 \quad \text{as} \quad n \to \infty \qquad (5a)$$

and

$$n^{-1}2^{(2\alpha+\beta+\gamma)N(n)} \to 0 \quad \text{as} \quad n \to \infty \qquad (5b)$$

then

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

at every point $y \in R$ at which $f(y) > 0$, and

$$\sum_{|k|=0}^{n} \sum_{l=L_{min}}^{L_{max}} b_{kl}\psi_{kl}(y) \to f(y) \text{ as } n \to \infty \qquad (6a)$$

and

$$\sum_{|k|=0}^{n} \sum_{l=L_{min}}^{L_{max}} a_{kl}\psi_{kl}(y) \to \lambda m^{-1}(y) f(y) \text{ as } n \to \infty \qquad (6b)$$

Sketch of the proof: By virtue of (3a), $\lim_{n\to\infty} E\hat{g}(y) = g(y)$, at every point $y \in R$ at which (6b) holds, where $\hat{g}(y)$ is the nominator in (2). Examining variance and using (3b), we get

$$\operatorname{var}\hat{g}(y) = O(2^{2(\alpha+\beta)N(n)}/n) + O(2^{(2\alpha+\beta+\gamma)N(n)}/n)$$

Therefore, $\hat{g}(y) \to g(y)$ as $n \to \infty$ in probability at every point $y \in R$ at which (6b) is in force. Similar convergence can be shown for the denominator of the estimate. □

4. THE DAUBECHIES WAVELET ALGORITHM

The Daubechies wavelets are families of compactly supported functions $\mathcal{D}^s = \{D^s_{kl};\, |k|, |l| = 0,1,2,\ldots\}$, $s = 1,2,3,\ldots$, where, as usual

$$D^s_{kl}(y) = 2^{k/2} D^s(2^k y - l) \qquad (7)$$

however for $s > 1$ the mother wavelet $D^s(y)$ is not given by an explicit formula but is computed from a two-stage iterative procedure as follows ([2, Ch. 6], [16, Ch. 3]):

Stage 1: Determine the 'scaling function' $\phi^s(y)$ such that

$$\phi^s(y) = \sum_l c^s_l\, \phi^s(2y - l) \qquad (8a)$$

by solving, for each y (typically, at dyadic points $y = l/2^k$), the recursion

$$\phi^s_{j+1}(y) = \sum_l c^s_l\, \phi^s_j(2y - l), \qquad j = 0,1,\ldots$$

where $\phi^s_0(y) = \chi_{[0,1)}(y)$ (indicator function of [0,1)) and $\{c^s_l\}$ are sequences of $2s$ non-zero coefficients, specific for each s. They can be found in [2, p. 195, Table 6.1].

Stage 2: Compute $D^s(y)$, for each y (practically, on the grid as above), as

$$D^s(y) = \sum_l (-1)^l c_{1-l}^s \phi^s(2y-l) \tag{8b}$$

where $\{c_l^s\}$ are the same as before.

The smooth differentiable wavelets are obtained for $s \geq 3$. They constitute orthonormal bases for $L^2(R)$ and are regular, which means that

$$|D^s(y)|, |D^{s\prime}(y)| \leq \rho^s/(1+|y|)^2, \quad \text{all } y \in R$$

for the proper choice of constants $\rho^s > 0$. Therefore, the following obvious relations hold

$$|D_{kl}^s(y)| \leq c^s 2^{k/2}, \qquad \text{all } l$$

some c^s independent of k,

$$\sup_{y\in R} |D_{kl}^s(y)| \leq d_1^s 2^{k/2}, \qquad \text{all } l$$

some d_1^s independent of k, and

$$\sup_{y\in R} |D_{kl}^{s\prime}(y)| \leq d_2^s 2^{k/2+k}, \qquad \text{all } l$$

some d_2^s independent of k, i.e. the requirements in Assumption 3 are satisfied with $\alpha = \beta = 1/2$ and $\gamma = 3/2$.

The respective Daubechies expansion of a square integrable function g has the form:

$$\sum_{|k|=0}^{\infty} \sum_{l=L_{\min}}^{L_{\max}} a_{kl} D_{kl}^s(y)$$

where for each fixed y

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

and the support $[s_1,s_2]$ of $D^s(y)$ is defined by the number of defining coefficients $\{c_l^s\}$ in (8a) as follows: $s_1=1-N_s$, $s_2=N_s$, where $N_s=(l_{\max} - l_{\min} + 1)/2$ is called a wavelet number, and where $l_{\max}$ is the greatest odd integer and $l_{\min}$ - the least even integer, such that $c_l^s \neq 0$ for $l_{\min} \leq l \leq l_{\max}$ ([1]). In this expansion $a_{kl} = \langle g, D_{kl}^s \rangle$ are respective Fourier coefficients. The following convergence takes place:

$$\lim_{n\to\infty} \sum_{|k|=0}^{n} \sum_{l=L_{\min}}^{L_{\max}} a_{kl} D_{kl}^s(y) = g(y)$$

at every point $y \in (-\infty,\infty)$ at which g is continuous, [11, Theorem 2.1].

Applying smooth Daubechies wavelets (for $s \geq 3$), we get the algorithm (see (2)):

$$\hat{v}_D^s(y) = \frac{\sum_{|k|=0}^{N(n)} \sum_{l=L_{\min}}^{L_{\max}} \hat{a}_{kl,D}^s D_{kl}^s(y)}{\sum_{|k|=0}^{N(n)} \sum_{l=L_{\min}}^{L_{\max}} \hat{b}_{kl,D}^s D_{kl}^s(y)} \tag{9}$$

where for given y and given s,

$$L_{\min} = [2^k y - N_s] + 1 \quad \text{and} \quad L_{\max} = [2^k y + N_s - 1]$$

with N_s specified above, and

$$\hat{a}_{kl,D}^s = \frac{1}{n}\sum_{i=1}^{n} U_{i-1} D_{kl}^s(Y_i) = \frac{2^{k/2}}{n} \sum_{\{i:R_i\in[1-N_s,N_s)\}} U_{i-1} D^s(R_i) \tag{9a}$$

and

$$\hat{b}_{kl,D}^s = \frac{1}{n}\sum_{i=1}^{n} D_{kl}^s(Y_i) = \frac{2^{k/2}}{n} \sum_{\{i:R_i\in[1-N_s,N_s)\}} D^s(R_i) \tag{9b}$$

where $R_i = 2^k Y_i - l$ and the values $D^s(R_i)$ are computed according to (8a)-(8b). Notice that the estimates in (9a)-(9b) can be calculated recursively.

Invoking the general Theorem in Section 3 and using the above-mentioned facts, concerning the basic properties of the Daubechies families of wavelets and convergence of the Daubechies wavelet expansions, we come to the following theorem.

Theorem D: Let Assumptions 1 and 2 hold. For the Wiener system and the Daubechies wavelet-based identification algorithm if, in addition to (4), also

$$2^{3N(n)}/n \to 0 \quad \text{as} \quad n \to \infty, \tag{10}$$

then

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

at every point $y \in (-\infty,\infty)$ at which $f(y) > 0$, and f is continuous.

Observe that here only continuity, at a given point, of the system output probability density $f(y)$ is additionally required.

5. THE HERMITE SERIES ALGORITHM

For comparison, we present here the correspondent identification algorithm based on the classical Hermite series which is also an orthonormal system on the whole real line.

Hermite polynomials are defined in the following way

$$H_k(y) = e^{y^2}(d^k/dy^k)e^{-y^2}, \tag{11a}$$

$k = 0,1,2,\ldots$, [14, Ch.4]. For example, $H_0(y) = 1$, $H_1(y) = -2y$, $H_2(y) = 4y^2 - 2$, $H_3(y) = -8y^3 + 12y$, and so on - which is however not so automatic and easy to compute as in the case of wavelet basis.

It is well known that the series

$$h_k(y) = e^{-y^2/2}H_k(y)/\sqrt{2^k k!\sqrt{\pi}} \tag{11b}$$

$k = 0,1,2,\ldots$, is orthonormal in the entire real line. Moreover, it is known that

$$|h_k(y)| \le c(y)(k+1)^{-1/4},$$

some $c(y)$ independent of k and

$$\sup_{y\in R}|h_k(y)| \le d_1(k+1)^{-1/12},$$

some d_1 independent of k, [15, p. 242], and

$$\sup_{y\in R}|h_k'(y)| \le d_2(k+1)^{5/12},$$

some d_2 independent of k, which fulfils Assumption 3 with $h_k(y)$ instead of $\psi_{kl}(y)$, $k+1$ in place of 2^k, and $\alpha = -1/4$, $\beta = -1/12$ and $\gamma = 5/12$.

The Hermite expansion of a square integrable function g is of the following form:

$$\sum_{k=0}^{\infty} a_k h_k(y),$$

where $a_k = \langle g,h_k\rangle$ are Fourier coefficients, and it holds the convergence:

$$\lim_{n\to\infty}\sum_{k=0}^{n} a_k h_k(y) = g(y),$$

at every point $y\in(-\infty,\infty)$ at which g is differentiable, [15, p. 247].

Applying the Hermite series, by analogy to (9-9a-9b) we obtain the following identification algorithm of $\lambda m^{-1}(y)$ (with single sums in the nominator and denominator and $h_k(y)$ in place of $D_{kl}(y)$):

$$\check{v}_h(y) = \frac{\sum_{k=0}^{N(n)} \check{a}_{k,h}\, h_k(y)}{\sum_{k=0}^{N(n)} \check{b}_{k,h} h_k(y)}, \tag{12}$$

where

$$\check{a}_{k,h} = \frac{1}{n}\sum_{i=1}^{n} U_{i-1}h_k(Y_i), \text{ and } \check{b}_{k,h} = \frac{1}{n}\sum_{i=1}^{n} h_k(Y_i)$$

and computation can be made recursively.

Observe that the estimate can be reduced to a more convenient form ([7])

$$\check{v}_h(y) = \frac{\sum_{k=0}^{N(n)} \check{a}_{k,H}\, H_k(y)}{\sum_{k=0}^{N(n)} \check{b}_{k,H} H_k(y)}, \tag{12a}$$

with

$$\check{a}_{k,H} = \frac{1}{n}\sum_{i=1}^{n} U_{i-1}H_k(Y_i), \text{ and } \check{b}_{k,H} = \frac{1}{n}\sum_{i=1}^{n} H_k(Y_i)$$

where one does not calculate the time consuming function exp(.) but deals only with polynomials $H_k(y)$.

Now, an application of the respective counterpart of the Theorem in Section 3 - see Remark *H* below and [8] for the details, and the usage of the Hermite series properties reported above lead to the theorem (compare Theorem 3 in [7] for the white noise case):

Theorem H: Let Assumptions 1 and 2 be in force. If (4) holds and

$$N^{7/3}(n)/n \to 0 \quad \text{as} \quad n\to\infty, \tag{13}$$

then

$$\check{v}_h(y) \to \lambda m^{-1}(y) \quad \text{as} \quad n\to\infty \quad \text{in probability}$$

at every point $y\in(-\infty,\infty)$ at which $f(y) > 0$, and f is differentiable.

Notice that actually differentiability, at a point, of the density $f(y)$ of the output signal is needed.

Remark H: Theorem *H* can be easily concluded from the general Theorem in Section 3 by replacing (5a) and (5b) with

$$n^{-1}\sum_{k=0}^{N(n)}(k+1)^{2\alpha}\sum_{k=0}^{N(n)}(k+1)^{2\beta}\to 0 \text{ as } n\to\infty \tag{14a}$$

and

$$n^{-1}\sum_{k=0}^{N(n)}(k+1)^{2\alpha}\sum_{k=0}^{N(n)}(k+1)^{\beta+\gamma}\to 0 \text{ as } n\to\infty \quad (14b)$$

respectively, using appropriate versions of (6a) and (6b) (with single sums and $h_k(y)$ in place of $\psi_{kl}(y)$) and lying $\alpha = -1/4$, $\beta = -1/12$ and $\gamma = 5/12$. Observe that (14a)-(14b) certainly correspond to the conditions (5a)-(5b), with $k+1$ in the role of 2^k, as

$$\sum_{k=0}^{N(n)} q^k = O(q^{N(n)})$$

for $q > 1$ and $N(n)$ large. More explanation of the unified framework, covering both classical orthogonal systems and orthogonal wavelets, can be found in [8,9].

6. FINAL REMARKS

1. The wavelet-based identification algorithm is computationally much simpler than the algorithm based on the classical Hermite series. It needs only elementary operations - scaling and translating of a constant function (mother wavelet) and of the data (see (7) and (9)), on the contrary to the Hermite series algorithm which requires more involved and less automatic computation - even in the simplified version (compare (11) and (12)-(12a)). Moreover, computational cost of implementing the wavelet estimator can be further diminished by the use of recursive versions of (9a)-(9b) and of fast wavelet algorithms ([12],[13],[5]).
2. The wavelet-based algorithm is attractive because of its satisfactory convergence properties, as well. The condition (10) should be not discouraging as compared with (13) in Theorem *H* because it is very well known that for smooth functions (Assumption 1) and smooth wavelets, generally moderate N is sufficient for good function recovering ([3],[4]). Moreover, in addition to Assumptions 1 and 2, only continuity - at a point - and not differentiability (as in Theorem *H*) of the density $f(y)$ of the output signal is required to assure (pointwise) convergence of the algorithm to the unknown non-linearity.
3. From the estimate (9) (or (12)), one can derive the estimate of the (dilated) non-linearity characteristic $m(v/\lambda)$ by computing a pseudo-inversion - see [7].
4. The true scaling factor λ (dilation constant $1/\lambda$) cannot be estimated by the approach. This is a consequence of the cascade structure of the system and the 'assembling' character of the measurement data used.
5. The algorithm works well despite correlation of the noise (Section 2). Its form does not depend on the particular correlation structure of the noise.
6. To memorize the estimate, only a finite number of coefficients in (9) (or (12)) is sufficient to be stored in a computer - much less than the whole number of data n.
7. The algorithm can be used under poor *a priori* knowledge of the system, when no parametrization of the non-linear characteristic m is known.

7. REFERENCES

1. Antoniadis A., Gregoire G., McKeague I. W.: Wavelet methods for curve estimation, J. Amer. Statist. Assoc., vol. 89, 1994, pp. 1340-1353
2. Daubechies I.: Ten Lectures on Wavelets, SIAM Edition, Philadelphia, 1992
3. Donoho D., Johnstone I. M.: Adapting to unknown smoothness via wavelet shrinkage, J. Amer. Statist. Assoc., vol. 90, 1995, pp. 1200-1224
4. Donoho D., Johnstone I. M., Kerkyacharian G., Picard D.: Wavelet shrinkage - asymptopia?, J. Royal Statist. Soc., vol. 57, 1995, pp. 301-369
5. Fabre E.: New fast smoothers for multiscale systems, IEEE Trans. Signal Proc., vol. 44, 1996, pp. 1893-1911
6. Greblicki W.: Nonparametric identification of Wiener systems, IEEE Trans. Inform. Theory, vol. 38, 1992, pp. 1487-1493
7. Greblicki W.: Nonparametric identification of Wiener systems by orthogonal series, IEEE Trans. Autom. Control, vol. 39, 1994, pp. 2077-2086
8. Greblicki W., Hasiewicz Z.: Non-linear system identification in the presence of correlated noise, Reports of the Inst. Engng. Cybern., Tech. Univ. of Wrocław, no. 37/95, 1995
9. Greblicki W., Hasiewicz, Z.: System identification in the presence of correlated noise; non-parametric approach, Proc. 3rd Int. Symp. on Methods and Models in Automation and Robotics - MMAR'96, vol. 2, Międzyzdroje 1996, pp. 619-624
10. Hunter I. W., Korenberg M. J.: The identification of nonlinear biological systems: Wiener and Hammerstein cascade models, Biological Cybernetics, vol. 55, 1986, pp. 125-134
11. Kelly S., Kon M., Raphael L.: Pointwise convergence of wavelet expansions, Bull. Amer. Math. Soc., vol. 30, 1994, pp. 87 - 94
12. Mallat S.: A theory for multiresolution signal decomposition - the wavelet representation, IEEE Trans. Pattern Anal. Machine Intelig., vol. 11, 1989, pp. 674-693
13. Meyer Y.: Wavelets - Algorithms and Applications, SIAM Edition, Philadelphia, 1993
14. Sansone G.: Orthogonal Functions, Interscience Publishers Inc., 1959
15. Szegö G.: Orthogonal Polynomials, Amer. Math. Soc. Coll. Publ., 1959
16. Walter G.: Wavelets and Other Orthogonal Systems with Applications, CRC Press Inc., Boca Raton, 1994