ORTHOGONAL SERIES ALGORITHMS TO IDENTIFY HAMMERSTEIN SYSTEMS

WŁODZIMIERZ GREBLICKI*, PRZEMYSŁAW ŚLIWIŃSKI†

*†Institute of Engineering Cybernetics, Wrocław University of Technology,
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland,
*wgre@ict.pwr.wroc.pl, †slk@ict.pwr.wroc.pl

Abstract. An unified approach to non-linearity identification in Hammerstein systems by orthogonal series algorithms is presented. The conditions of convergence are given for orthogonal series algorithms examined in the literature.

Key Words. Hammerstein system, non-parametric identification, orthogonal series.

1. INTRODUCTION

In this work, we identify Hammerstein system, i.e. a non-linear memoryless subsystem followed by a linear dynamic one. We identify its nonlinear part from input-output observations of the whole system. The problem has received great attention from a number of authors, see Gallman [4], Chang and Luus [2], Tchatachar and Ramaswamy [19], Haist *et al.* [10], Billings and Fakhouri [1].

All authors mentioned above have assumed that the unknown non-linear characteristic can be represented in a parametric form, i.e. that it is a member of a class of functions which can be parameterized. Usually, the characteristic is a polynomial whose degree does not exceed a known number. So, in order to apply successfully their identification algorithms we must possess information which can often be unavailable.

In contradistinction to those authors, our *a priori* information about the nonlinear subsystem is much smaller. We only assume that the characteristic is square integrable. In the statistics, problems like ours are called non-parametric. The non-parametric approach to recovering the nonlinearity in the Hammerstein system has been studied by Greblicki and Pawlak [6, 7, 8, 9], Greblicki [5], Hasiewicz [11], Pawlak [14], Krzyżak [13], Pawlak and Hasiewicz [15], Śliwiński [18], Zi-Qiang [21], Vandersteen *et al.* [20]. Various algorithms have been proposed of which the kernel and orthogonal series look most effective.

In this paper we present an approach which unifies orthogonal series algorithms presented in the literature.

2. THE IDENTIFIED SYSTEM

The identified Hammerstein system with input U_n and output Y_n is shown on Fig. 1. In the present work $\{U_n; n = \cdots, -1, 0, 1, \cdots\}$ is a stationary white random process. The nonlinear memoryless subsystem has a characteristic m which means that

$$W_n = m(U_n).$$

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

$$\begin{aligned} X_{n+1} &= AX_n + bW_n \\ V_n &= c^T X_n, \end{aligned} \tag{1}$$

where A is a matrix, b and c are vectors, and ^{T} denotes transposition. Here A, b, and c are all unknown, but it is assumed that A is a stable matrix. Clearly $\{X_n; n = \cdots, -1, 0, 1, \cdots\}$ is a stationary random process; V_n is not accessible for measurement and we have only Y_n, where

$$Y_n = V_n + Z_n$$

and where $\{Z_n\}$ is an additive stationary noise with zero mean and finite variance. Moreover, processes $\{U_n\}$ and $\{Z_n\}$ are mutually independent. Our goal is to recover m from input-output observations of the whole system, i.e. from $(U_1, Y_1), \cdots, (U_n, Y_n)$.

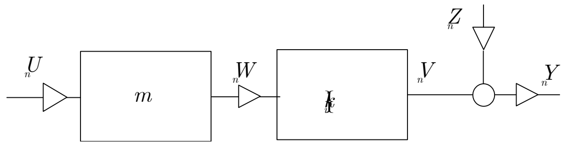

Fig. 1. The identified Hammerstein system, $k_0 = 0$ and $k_i = c^T A^{i-1} b$, $i = 1, 2, \cdots$.

We also assume that $U_n \in \Delta$ with some Δ. In the paper, we consider cases for which Δ is a finite interval, half real line or the whole real line.

By f we denote the probability density of U_n which is assumed to exist.

3. BASIC ALGORITHM

The idea of our algorithms is based on the fact that

$$\mathrm{E}\left\{Y_{n+1} \left| U_n = u\right.\right\} = \alpha m(u) + \beta,$$

where $\alpha = c^T b$, $\beta = c^T A \mathrm{E} X_n$. For the sake of simplicity, we denote

$$\mu(u) = \alpha m(u) + \beta.$$

Coefficients α and β are unknown and can't be estimated due to the fact that the inner system signal is not measured.

To identify the nonlinear subsystem, we apply orthogonal functions. Let $\varphi_k(u)$, $k = 0, 1, 2, \cdots$, be a sequence of such functions, i.e. a sequence of complete orthonormal functions in a set Δ. If Δ is a finite interval, one can apply the trigonometric or Legendre orthogonal functions. If Δ is a half real line, the Laguerre functions can be employed while for Δ being the whole real line, Hermite polynomials can be used. Compactly supported orthogonal wavelets can be exploited for finite and infinite interval Δ (see Appendix).

We assume that

$$\max_{u \in \Delta} |\varphi_k(u)| \leq c 2^{k/2} \tag{2}$$

for some orthogonal series, and

$$\max_{u \in \Delta} |\varphi_k(u)| \leq c k^{\delta} \tag{3}$$

for others.

Observe that

$$\mu(u) = \frac{g(u)}{f(u)},$$

where $g(u) = \mu(u) f(u)$. Expanding g and f, we get

$$g(u) \sim \sum_{k=0}^{\infty} a_k \varphi_k(u),$$

and

$$f(u) \sim \sum_{k=0}^{\infty} b_k \varphi_k(u),$$

where

$$a_k = \mathrm{E}\left\{Y_{n+1} \varphi_k(U_n)\right\} \text{ and } b_k = \mathrm{E}\varphi_k(U_n).$$

By $s_n(u; g)$ and $s_n(u, f)$ we denote partial expansions of g and f, respectively. It means that

$$s_n(u; g) = \sum_{k=0}^{n} a_k \varphi_k(u)$$

and

$$s_n(u; f) = \sum_{k=0}^{n} b_k \varphi_k(u).$$

Since both a_k and b_k can be easily estimated, our basic algorithm recovering μ has the following form:

$$\hat{\mu}(u) = \frac{\sum_{k=0}^{q(n)} \hat{a}_k \varphi_k(u)}{\sum_{k=0}^{q(n)} \hat{b}_k \varphi_k(u)},$$

where

$$\hat{a}_k = \frac{1}{n} \sum_{i=0}^{n} Y_{i+1} \varphi_k(U_i) \text{ and } \hat{b}_k = \frac{1}{n} \sum_{i=0}^{n} \varphi_k(U_i).$$

In the estimate $q(n)$ is a sequence of integers. We show that the proper selection of the sequence makes the algorithm converge to $\mu(u)$.

4. CONVERGENCE OF THE BASIC ALGORITHM

Denote

$$\hat{g}(u) = \sum_{k=0}^{q(n)} \hat{a}_k \varphi_k(u)$$

and

$$\hat{f}(u) = \sum_{k=0}^{q(n)} \hat{b}_k \varphi_k(u)$$

and observe that $\hat{\mu}(u) = \hat{g}(u)/\hat{f}(u)$. Clearly, $\mathrm{E}\hat{a}_k = a_k$ and $\mathrm{E}\hat{b}_k = b_k$. Owing to that,

$$\mathrm{E}\hat{g}(u) = \sum_{k=0}^{q(n)} a_k \varphi_k(u)$$

and

$$\mathrm{E}\hat{f}(u) = \sum_{k=0}^{q(n)} b_k \varphi_k(u).$$

We can now give the following

Lemma 1 *Let*

$$q(n) \to \infty \text{ as } n \to \infty. \quad (4)$$

Then

$$\lim_{n\to\infty} \mathrm{E}\hat{g}(u) = g(u) \text{ and } \lim_{n\to\infty} \mathrm{E}\hat{f}(u) = f(u)$$

at every point at which

$$\lim_{n\to\infty} s_n(u; g) = g(u) \text{ and } \lim_{n\to\infty} s_n(u; f) = f(u), \quad (5)$$

respectively.

We shall now examine variance of $\hat{g}(u)$ and $\hat{f}(u)$. Observe that

$$\hat{g}(u) = \frac{1}{n}\sum_{i=0}^{n} Y_{i+1}\Phi_{q(n)}(U_i, u)$$

where

$$\Phi_n(v, u) = \sum_{k=0}^{n} \varphi_k(v)\varphi_k(u).$$

Observe in passing that due to (2)

$$\max_{v\in\Delta} \left|\Phi_{q(n)}(v, u)\right| \le c^2 \sum_{k=0}^{q(n)} 2^k \le d2^{q(n)+1} \quad (6)$$

with some d. Thus

$$\begin{aligned} \mathrm{var}\left[\hat{g}(u)\right] &= \frac{1}{n^2}\,\mathrm{var}\left[\sum_{i=0}^{n} Y_{i+1}\Phi_{q(n)}(U_i, u)\right] \\ &= V_1(u) + V_2(u). \end{aligned}$$

where

$$V_1(u) = \frac{1}{n^2}\sum_{i=0}^{n} \mathrm{var}\left[Y_{i+1}\Phi_{q(n)}(U_i, u)\right]$$

and

$$V_2(u) = \frac{1}{n^2} \times \sum_{i=0}^{n}\sum_{\substack{j=0\\ j\ne i}}^{n} \mathrm{cov}\left[Y_{i+1}\Phi_{q(n)}(U_i, u), Y_{j+1}\Phi_{q(n)}(U_j, u)\right].$$

Owing to (6)

$$\begin{aligned} V_1(u) &\le \frac{d^2 2^{2q(n)+2}}{n^2}\sum_{i=0}^{n} \mathrm{E}Y_{i+1}^2 \\ &= O\left(\frac{2^{2q(n)+2}}{n}\right) \end{aligned}$$

In turn, because of (10)

$$\begin{aligned} V_2(u) &\le \frac{cd^2 2^{2q(n)+2}}{n^2}\sum_{i=1}^{n-1} (n-i)\left\|A^n\right\| \\ &= O\left(\frac{2^{2q(n)+2}}{n}\right). \end{aligned}$$

In this way we have verified the following

Lemma 2 *If*

$$\frac{2^{2q(n)+2}}{n} \to 0 \text{ as } n \to \infty, \quad (7)$$

then

$$\lim_{n\to\infty} \mathrm{var}\left[\hat{g}(u)\right] = 0 \text{ and } \lim_{n\to\infty} \mathrm{var}\left[\hat{f}(u)\right] = 0.$$

If, moreover, orthogonal functions satisfy (3), then one can show in a similar manner that

Lemma 3 *If*

$$\frac{q^{4\delta+2}(n)}{n} \to 0 \text{ as } n \to \infty, \quad (8)$$

then

$$\lim_{n\to\infty} \mathrm{var}\left[\hat{g}(u)\right] = 0 \text{ and } \lim_{n\to\infty} \mathrm{var}\left[\hat{f}(u)\right] = 0.$$

We are now in a position to present

Theorem 1 *If (4) and (7) hold for orthogonal series satisfying (2) (or if (4) and (8) hold for series satisfying (3)) then*

$$\hat{\mu}(u) \to \alpha m(u) + \beta \text{ as } n \to \infty \text{ in probability}$$

at every point $u \in \Delta$ at which (5) hold and $f(u) > 0$.

5. PARTICULAR ALGORITHMS

The following Corollaries are now simple consequence of our Theorem and facts given in Appendix.

Corollary 1 *If*

$$\lim_{n\to\infty} q(n) = \infty \text{ and } \lim_{n\to\infty} \frac{q^2(n)}{n} = 0,$$

then the trigonometric series algorithm converges at every point $u \in (-\pi, \pi)$ at which both f and g are differentiable and $f(u) > 0$.

Corollary 2 *If*

$$\lim_{n\to\infty} q(n) = \infty \text{ and } \lim_{n\to\infty} \frac{q^2(n)}{n} = 0,$$

then the Legendre series algorithm converges at every point $u \in (-1, 1)$ at which both f and g are differentiable and $f(u) > 0$.

Corollary 3 *If*

$$\lim_{n\to\infty} q(n) = \infty \text{ and } \lim_{n\to\infty} \frac{q^6(n)}{n} = 0,$$

then the Laguerre series algorithm converges at every point $u \in (0, \infty)$ at which both f and g are differentiable and $f(u) > 0$.

Corollary 4 *If*

$$\lim_{n\to\infty} q(n) = \infty \text{ and } \lim_{n\to\infty} \frac{q^{5/3}(n)}{n} = 0,$$

then the Hermite series algorithm converges at every point $u \in (-\infty, \infty)$ at which both f and g are differentiable and $f(u) > 0$.

Corollary 5 *If*

$$\lim_{n\to\infty} q(n) = \infty \text{ and } \lim_{n\to\infty} \frac{2^{2q(n)+2}}{n} = 0,$$

then the Daubechies wavelet series algorithm converges at every point $u \in (-\infty, \infty)$ at which both f and g are continuous and $f(u) > 0$.

CONCLUSIONS

In the paper we have presented the approach unifying the orthogonal series algorithms recovering the non-linearity in Hammerstein system. We have shown that the proper selection of the algorithm complexity, with relation to the employed orthogonal series and number of measurements, makes the algorithm converge. The convergence of the algorithm holds at every points of convergence of the respective orthogonal series.

APPENDIX

A ORTHOGONAL FUNCTIONS

Let Δ be a fixed set and let ρ be a function such that $\int_\Delta \rho^2(x)dx < \infty$. Let $\varphi_k(x), k = 0, 1, 2, \cdots$ be a complete set of functions orthonormal in Δ and let $s_n(x)$ be a partial expansion of ρ in the series, i.e. let $s_n(x) = \sum_{k=0}^{n} a_k \varphi_k(x)$, where $a_k = \int_\Delta \rho(x)\varphi_k(x)dx$. Below, we present various orthonormal series functions and examine the following convergence

$$\lim_{n\to\infty} s_n(x) = \rho(x). \tag{9}$$

Trigonometric series. Let $\Delta = [-\pi, \pi]$ and let

$$\varphi_k(x) = e^{jkx}, k = 0, \pm1, \pm2, \cdots.$$

It is clear that $\max_{-\pi\le x\le\pi} \left|e^{jkx}\right| \le 1$ and hence $\delta = 0$ in (3). Now the partial expansion has the following form $s_k(x) = \sum_{|i|=0}^{k} a_i\varphi_i(x)$. It is well known that (9) holds at every point $x \in (-\pi, \pi)$ at which ρ is differentiable, see [16].

Legendre series. Let $\Delta = [-1, 1]$ and let

$$\varphi_k(x) = (\sqrt{2k+1}/\sqrt{2})P_k(x),$$

where

$$P_k(x) = (1/2^k k!)(d^k/dx^k)\left(x^2-1\right)^2$$

is the kth Legendre polynomial. One can verify that $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = (3/2)x^2 - 1/2$ and so on. We have $\max_{-1\le x\le1} |\varphi_k(x)| \le 1$ and thus $\delta = 0$. Moreover (9) holds at every point $x \in [-1, 1]$ at which ρ is differentiable, see [16].

Laguerre series. Let $\Delta = [0, \infty)$ and let

$$\varphi_k(x) = e^{-x/2}L_k(x),$$

where

$$L_k(x) = (1/k!)e^x(d^k/dx^k)\left(x^k e^{-x}\right)$$

is the kth Laguerre polynomial. We have $L_0(x) = 1$, $L_1(x) = -x + 1$, $L_2(x) = x^2/2 - 2x + 1, \cdots$. Since it is known that $\max_{0\le x<\infty} |\varphi_k(x)| \le ck$ with some c then $\delta = 1$. Moreover, (9) holds at every point $x \in (0, \infty)$ at which ρ is differentiable, see [17].

Hermite series. Let $\Delta = (-\infty, \infty)$ and let

$$\varphi_k(x) = (1/\sqrt{2^k k!\sqrt{\pi}})H_k(x),$$

where

$$H_k(x) = (-1)^k e^{x^2}(d^k/dx^k)e^{-x^2}$$

is the kth Hermite polynomial. One can verify that $H_0(x) = 1$, $H_1(x) = -2x$, $H_2(x) = 4x^2 - 2, \cdots$. It is known that $\max_{-\infty<x<\infty} |\varphi_k(x)| \le ck^{-1/12}$, thus $\delta = -1/12$. Moreover (9) holds at every point $x \in (-\infty, \infty)$ at which ρ is differentiable, see [17].

Daubechies wavelet series. Let $\Delta = (-\infty, \infty)$ and

$$\varphi_0(x) = \sum_{n=n_{\min}(x;p)}^{n_{\max}(x;p)} D^p_{0,n}(x)$$
$$\varphi_k(x) = \sum_{l=l_{\min}(x;k,p)}^{l_{\max}(x;k,p)} W^p_{kl}(x),$$

for $k = 1, 2, \cdots$, and fixed $x \in \Delta$, and where

$$D^p_{0,n}(x) = D^p(x-n),$$
$$W^p_{kl}(x) = 2^{k/2}W^p\left(2^k x - l\right)$$

are translated dilations of functions $D^p(x)$ and $W^p(x)$, being the father and mother wavelet of the pth Daubechies wavelet family, see [3, p. 194]. The indices n and l run from $n_{\min}(x;p) = \lfloor x\rfloor - 2p + 2$ to $n_{\max}(x;p) = \lfloor x\rfloor$ and from $l_{\min}(x;k,p) = \lfloor 2^k x\rfloor - p + 1$ to $l_{\max}(x;k,p) = \lfloor 2^k x\rfloor + p - 1$, respectively.

Remark 1 *Alternatively, one can set*

$$\varphi_k(x) = \sum_{l=l_{\min}(x;k,p)}^{l_{\max}(x;k,p)} \psi^p_{kl}(x), k = 0, \pm1, \pm2, \cdots.$$

In this case the partial expansion turns to the form $s_k(x) = \sum_{|i|=0}^{k} a_i\varphi_i(x)$.

In both cases $\max_{-\infty<x<\infty} |\varphi_k(x)| \leq c2^{k/2}$. The convergence (9) holds at every point $x \in (-\infty, \infty)$ at which ρ is continuous, see [12].

B HAMMERSTEIN SYSTEM

The lemma below holds for any f and any m.

Lemma 4 *Let ϕ be a Borel measurable function. Then in the Hammerstein system*

$$\begin{aligned} & \operatorname{cov}[X_{n+1}\phi(U_n), X_1\phi(U_0)] \\ = \; & A^{n+1}\operatorname{cov}[X_0, X_0]\, A^T \mathrm{E}^2\{\phi(U_0)\} \\ & + A^n b \mathrm{E}\{X^T\} A^T \operatorname{cov}[m(U_0), \phi(U_0)]\, \mathrm{E}\{\phi(U_0)\} \\ & + A^n b b^T \mathrm{E}\{\phi(U_0)\} A^T \operatorname{cov}[m(U_0), m(U_0)\phi(U_0)] \end{aligned}$$

for $n = 1, 2, 3, \cdots$.

Proof. The proof is straightforward and is omitted. ■

Corollary 6 *Let ϕ be a bounded Borel measurable function. Then in the Hammerstein system*

$$\|\operatorname{cov}[X_{n+1}\phi(U_n), X_1\phi(U_0)]\| \leq \gamma d^2 \|A^n\| \quad (10)$$

for $n = 1, 2, 3, \cdots$, where $d = \max|\phi(u)|$ and γ is independent of n and ϕ.

REFERENCES

[1] S. Billings and S. Fakhouri. Non-linear system identification using the Hammerstein model. *International Journal of Systems Science*, 10(5):567–578, 1979.

[2] F. Chang and R. Luus. A noniterative method for identification using Hammerstein model. *IEEE Transactions on Automatic Control*, pages 465–468, October 1971.

[3] I. Daubechies. *Ten Lectures on Wavelets.* SIAM, Philadelphia, 1992.

[4] P. Gallman. A comparison of two Hammerstein model identification algorithms. *IEEE Transactions on Automatic Control*, AC-21:125–127, February 1976.

[5] W. Greblicki. Nonparametric orthogonal series identification of Hammerstein systems. *International Journal of Systems Science*, 20(12):2355–2367, 1989.

[6] W. Greblicki and M. Pawlak. Identification of discrete Hammerstein systems using kernel regression estimates. *IEEE Transactions on Automatic Control*, AC-31(1):74–77, January 1986.

[7] W. Greblicki and M. Pawlak. Nonparametric identification of a cascade nonlinear time series system. *Signal Processing*, 22:61–75, 1991.

[8] W. Greblicki and M. Pawlak. Nonparametric identification of a particular nonlinear time series system. *IEEE Transactions on Signal Processing*, 40(4):985–989, April 1992.

[9] W. Greblicki and M. Pawlak. Nonparametric recovering nonlinearities in block-oriented systems with the help of Laguerre polynomials. *Control-Theory and Advanced Technology*, 10(4, Part A):771–791, 1994.

[10] N. Haist, F. Chang, and R. Luus. Nonlinear identification in the presence of correlated noise using a Hammerstein model. *IEEE Transactions on Automatic Control*, AC-15:553–555, October 1973.

[11] Z. Hasiewicz, Non-parametric estimation of non-linearity in a cascade time series system by multiscale approximation, to appear in: *Signal Processing*, 81(4), 2001.

[12] S. Kelly, M. Kon and M. Raphael, Pointwise convergence of wavelet expansions, *Bulletin of The American Mathematical Society*, 30:87-94, 1994.

[13] A. Krzyżak. On estimation of a class of nonlinear systems by the kernel regression estimate. *IEEE Transactions on Information Theory*, 36(1):141–152, February 1990.

[14] M. Pawlak. On the series expansion approach to the identification of Hammerstein system. *IEEE Transactions on Automatic Control*, AC-36:763–767, 1991.

[15] M. Pawlak and Z. Hasiewicz. Nonlinear system identification by the Haar multiresolution analysis. *IEEE Transactions on Circuits and Systems, I-Fundamental Theory and Applications*, 45(9):945–961, September 1998.

[16] G. Sansone. *Orthogonal Functions.* Interscience, New York, 1959.

[17] G. Szegö. *Orthogonal Polynomials.* American Mathematical Society, Providence, R.I., third edition, 1974.

[18] P. Śliwiński, *Non-linear system identification using wavelet algorithms (Ph.D. thesis)*, Institute of Engineering Cybernetics, Wrocław University of Technology, Wrocław, 2000.

[19] M. Thathachar and S. Ramaswamy. Identification of a class of non-linear systems. *International Journal of Control*, 18(4):741–752, 1973.

[20] G. Vandersteen, Y. Rolain, and J. Schoukens. Non-parametric estimation of the frequency-response functions of the linear blocks of a Wiener-Hammerstein model. *Automatica*, 33(7):1351–1355, 1997.

[21] L. Zi-Qiang. A nonparametric polynomial identification algorithm for the Hammerstein system. *IEEE Transactions on Automatic Control*, 42(10):1435–1441, October 1997.