Recursive identification of continuous-time Hammerstein systems

W. Greblicki

Institute of Engineering Cybernetics, Wrocław University of Technology
S. Wyspiańskiego 27, 50-370 Wrocław, Poland

A continuous-time Hammerstein system is identified. The characteristic of its nonlinear subsystem and the impulse response of the dynamic parts are estimated from observations taken at input and output of the whole system. All algorithms are of the on-line type. Their convergence is shown. Results of simulation examples are also presented.

printed in:
International Journal of Systems Science
Vol. 33, No. 12, pp. 969–977, 2002

1. Introduction

The history of composite nonlinear system identification seems to begin with Narendra and Gallman (1966) in which an iterative method for the identification of Hammerstein systems was presented. Such discrete-time systems have been the object of many studies, (Chang and Luus 1971, Haist *et al.* 1973, Thathachar and Ramaswamy 1973, Kaminskas 1975, Gallman 1976, Billings and Fakhouri 1977, 1978, 1979, 1982). The 1990's brought a new wave of papers (e.g., Lang 1993, 1994, 1997, Liao and Sethares 1995, Al-Duwaish and Karim 1997, Al-Duwaish *et al.* 1997, Giri *et al.* 2001). In all of them, the characteristic of the nonlinear subsystem is usually continuous of a polynomial form. In terms of mathematical statistics it means that, the *a priori* information about the system is parametric. When the characteristic is not a polynomial, e.g., has a discontinuity, the behavior of their algorithms is uncertain.

To overcome that drawback, Greblicki and Pawlak (1986, 1989a, b) have applied the nonparametric approach. They have proposed kernel algorithms which have then been studied by Krzyżak (1990, 1992), as well as Krzyżak and Partyka (1993). A class of kernel algorithms applying order statistics has been examined in Greblicki and Pawlak (1994) and Greblicki (1996). Those based on the idea of orthogonal expansions have been investigated in Greblicki (1989), Krzyżak (1989), Greblicki and Pawlak (1991, 1994b), as well as Pawlak (1991). In particular, wavelets expansions have been used by Pawlak and Hasiewicz (1998), as well as Hasiewicz (1999, 2001). In turn, Lang (1993, 1994, 1997) has proposed a polynomial method. In all those works, the unknown nonlinear characteristic belongs to a very wide class of admissible functions. Usually the class consists of all functions bounded by a first degree polynomial. No restriction has been imposed on their functional form, the characteristic can be, e.g., continuous or not.

Apart from only Billings and Fakhouri (1978, 1979), all papers mentioned above deal with discrete-time Hammerstein systems. Recently, however, Greblicki (2000) has identified continuous-time ones. He proposed a nonparametric kernel algorithm of an off-line type to recover the nonlinear characteristic. Contrary to Billings and Fakhouri (1978, 1979), the characteristic may have any functional form, may be continuous or not, a polynomial or not, *etc.* In this paper we develop the idea and present two recursive kernel algorithms. Their feature, which, in many circumstances, can be considered as an advantage, is that they can be calculated on-line. Moreover, we also examine a nonparametric estimate of the impulse response of the dynamic subsystem and show that its global error converges to zero. We also present results of numeric simulation.

We also want to mention that Hammerstein systems have already been used to describe processes in biology, (Korenberg and Hunter 1986), chemistry, (Eskinat and

Johnson 1991), (Patwardhan *et al.* 1998), and to model magnetic hysteresis, (Hsu and Ngo 1997). This, together with the fact that the *a priori* information is small, makes the nonparametric approach and, in particular, our recursive algorithms interesting not only for researchers but engineers, too.

2. Statement of the Problem

The continuous-time Hammerstein system shown in figure 1 consists of two subsystems connected in a cascade. A nonlinear memoryless part is followed by a linear dynamic one. The first subsystem has a characteristic m. Its input signal is a stationary white random process $\{U(t); t \in (-\infty, \infty)\}$ with zero mean and autocovariance function $\sigma_U^2 \delta(.)$, where δ is the Dirac impulse. The random variable $U(t)$ has an unknown probability density f. We assume that m is a Borel measurable function satisfying the following condition:

$$|m(u)| \leq c_1 + c_2|u| \tag{1}$$

with some $c_1, c_2 > 0$. Owing to that $m(U(t))$ is a random variable and $\{m(U(t)); t \in (-\infty, \infty)\}$ is a stationary white random process with autocovariance function denoted by $\operatorname{var}[m(U(t))]\delta(.)$. The fact that σ_U^2 is finite together with (1) imply $\operatorname{var}[m(U(t))] < \infty$.

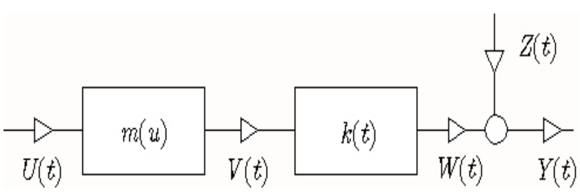

Figure 1: The identified Hammerstein system

The impulse response of the linear subsystem is denoted by k which means that

$$W(t) = \int_{-\infty}^{t} k(t-\xi)V(\xi)d\xi. \tag{2}$$

We assume that

$$\int_0^\infty k^2(t)dt < \infty. \tag{3}$$

Owing to (3), and the fact that $\operatorname{var}[m(U(t))] < \infty$, output $W(t)$ of the dynamic part is a random variable. Obviously, $\{W(t); t \in (-\infty, \infty)\}$ is a stationary correlated random process. Output of the dynamic subsystem is disturbed by additive stationary zero mean white random noise $\{Z(t); t \in (-\infty, \infty)\}$, i.e., a processes with autocovariance function $\sigma_Z^2 \delta(.)$. Therefore, $Y(t) = W(t) + Z(t)$, where $Y(t)$ is output of the whole system.

We identify both subsystems, i.e., estimate the nonlinear characteristic m and the impulse response k from observations $\{U(t), Y(t); t \in [0, \infty)\}$ taken at input and output of the whole system. Since the signal interconnecting the subsystems is not measured, both m and k can be estimated up to some constants only.

To simplify further considerations and formulas, we impose the following additional restrictions on the density f of $U(t)$ and the nonlinearity m:

$$f \text{ is an even function,} \tag{4}$$

$$m \text{ is an odd function.} \tag{5}$$

Restriction (4) implies $EU(t) = 0$. In turn, this, and (5) yield $EV(t) = EW(t) = EY(t) = 0$.

For the sake of simplicity, U and V, *etc.*, denotes random variables distributed like $U(t)$ and $V(t)$, *etc.* For convenience, for a fixed positive λ, we denote $\alpha = k(\lambda)$. Moreover, $\beta = E\{Um(U)\}$.

3. Nonlinear subsystem identification

3.1. Algorithms

To introduce the algorithm identifying the nonlinear subsystem, we need the following lemma which can be found in (Greblicki 2000):

Lemma 1 *In the system,*

$$E\{Y(t+\lambda)|U(t) = u\} = \alpha m(u).$$

Therefore, to recover $\alpha m(u)$ we estimate the regression $E\{Y(t+\lambda)|U(t) = u\}$ with the following algorithms:

$$\hat{\mu}(u;t) = \frac{\displaystyle\int_0^t Y(\lambda+\xi)\frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right)d\xi}{\displaystyle\int_0^t \frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right)d\xi}, \tag{6}$$

and

$$\tilde{\mu}(u;t) = \frac{\displaystyle\int_0^t Y(\lambda+\xi)K\left(\frac{u-U(\xi)}{h(\xi)}\right)d\xi}{\displaystyle\int_0^t K\left(\frac{u-U(\xi)}{h(\xi)}\right)d\xi}, \tag{7}$$

where, K and h are suitably selected kernel and bandwidth functions, respectively.

Depending on the estimate, the positive Borel measurable function h satisfies some of the following restrictions:

$$h(t) \to 0 \text{ as } t \to \infty, \tag{8}$$

$$\frac{1}{t^2}\int_0^t \frac{1}{h(\tau)}d\tau \to 0 \text{ as } t \to \infty, \tag{9}$$

$$\int_0^\infty h(\tau)d\tau = \infty. \tag{10}$$

The Borel measurable kernel K is selected to meet the following ones:

$$\sup_{-\infty<v<\infty} |K(v)| < \infty, \tag{11}$$

$$\int_{-\infty}^\infty |K(v)|\, dv < \infty, \tag{12}$$

$$vK(v) \to 0 \text{ as } |v| \to \infty. \tag{13}$$

Observe that for $h(t) = ct^{-\gamma}$, with any positive c, (8) and (9) hold for $0 < \gamma < 1$. Similarly, (8) and (10) hold also for $0 < \gamma < 1$. As the kernel, one can select, e.g., a rectangular function equal 1 or zero according to $|u|$ does not exceed or is greater than 1. Other examples are $e^{-|u|}$, e^{-u^2}, or $1/(1+u^2)$.

Algorithms (6) and (7) can be calculated recursively. Denoting

$$\hat{g}(t;u) = \frac{1}{t}\int_0^t Y(\lambda+\xi)\frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right)d\xi,$$

$$\hat{f}(u;t) = \frac{1}{t}\int_0^t \frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right)d\xi$$

and observing that $\hat{\mu}(u;t) = \hat{g}(t;u)/\hat{f}(u;t)$, we can write

$$\frac{d}{dt}\hat{g}(u;t) = -\frac{1}{t}\left[\hat{g}(u;t) - Y(\lambda+t)\frac{1}{h(t)}K\left(\frac{u-U(t)}{h(t)}\right)\right], \tag{14}$$

$$\frac{d}{dt}\hat{f}(u;t) = -\frac{1}{t}\left[\hat{f}(u;t) - \frac{1}{h(t)}K\left(\frac{u-U(t)}{h(t)}\right)\right], \tag{15}$$

with $\hat{g}(u;0) = \hat{f}(u;0) = 0$. In a similar way, we denote

$$\tilde{g}(t;u) = \frac{1}{\int_0^t h(\xi)d\xi}\int_0^t Y(\lambda+\xi)K\left(\frac{u-U(\xi)}{h(\xi)}\right)d\xi,$$

$$\tilde{f}(u;t) = \frac{1}{\int_0^t h(\xi)d\xi}\int_0^t K\left(\frac{u-U(\xi)}{h(\xi)}\right)d\xi$$

observe $\tilde{\mu}(u;t) = \tilde{g}(t;u)/\tilde{f}(u;t)$ and write

$$\frac{d}{dt}\tilde{g}(u;t) = -\gamma(t)\left[\tilde{g}(u;t) - Y(\lambda+t)\frac{1}{h(t)}K\left(\frac{u-U(t)}{h(t)}\right)\right], \tag{16}$$

$$\frac{d}{dt}\tilde{f}(u;t) = -\gamma(t)\left[\tilde{f}(u;t) - \frac{1}{h(t)}K\left(\frac{u-U(t)}{h(t)}\right)\right], \tag{17}$$

with $\gamma(t) = h(t)/\int_0^t h(\xi)d\xi$ and $\tilde{g}(u;0) = \tilde{f}(u;0) = 0$.

Both algorthms can be regarded as recursive versions of the following off-line one:

$$\bar{\mu}(u;t) = \frac{\displaystyle\int_0^t Y(\lambda+\xi)K\left(\frac{u-U(\xi)}{h(t)}\right)d\xi}{\displaystyle\int_0^t K\left(\frac{u-U(\xi)}{h(t)}\right)d\xi} \tag{18}$$

examined in (Greblicki 2000).

The idea standing behind our algorithms can be explained in the following way. From procedure (14) we expect that

$$E\left[\hat{g}(u;t) - Y(\lambda+t)\frac{1}{h(t)}K\left(\frac{u-U(t)}{h(t)}\right)\right] \to 0 \quad \text{as } t\to\infty$$

which is equivalent to

$$\left|E\hat{g}(u;t) - \alpha\frac{1}{h(t)}\int_{-\infty}^\infty m(u)K\left(\frac{u-v}{h(t)}\right)f(v)dv\right| \to 0 \quad \text{as } t \to \infty.$$

Owing to (11)–(13) and (8),

$$\frac{1}{h(t)}K\left(\frac{u-U(t)}{h(t)}\right)$$

gets close to $\delta(u-U(t))\int K(v)dv$ as $t \to \infty$, where $\delta(.)$ is the Dirac impulse. Thus,

$$\frac{1}{h(t)}\int_{-\infty}^\infty m(u)K\left(\frac{u-v}{h(t)}\right)f(v)dv \to m(u)f(u)\int_{-\infty}^\infty K(v)dv \text{ as } t \to \infty$$

and, consequently, $E\hat{g}(u;t) \to \alpha m(u)f(u)\int K(v)dv$ as $t \to \infty$. For similar reasons, $E\hat{f}(u;t) \to f(u)\int K(v)dv$ as $t \to \infty$. In the light of this, we can expect that (6), i.e. $\hat{g}(u;t)/\hat{f}(u;t)$, converges to $\alpha m(u)$. The reasoning holds also for (7).

3.2. Convergence

Our next theorem establishes convergence of the algorithm recovering the nonlinearity.

Theorem 1 *Let the kernel K satisfy (11)–(13). If h satisfy (8) and (9), then*

$$\hat{\mu}(u;t) \to \alpha m(u) \text{ as } t \to \infty \text{ in probability} \tag{19}$$

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

Proof Let u be a point at which both m and f are continuous. From Lemma 1, we get

$$\begin{aligned} E\hat{g}(u;t) &= \frac{1}{t}\int_0^t E\left\{Y(\lambda)\frac{1}{h(\xi)}K\left(\frac{u-U(0)}{h(\xi)}\right)\right\}d\xi \\ &= \alpha\frac{1}{t}\int_0^t E\left\{m(U)K\left(\frac{u-U}{h(\xi)}\right)\right\}d\xi. \end{aligned}$$

Since (8) holds and

$$\begin{aligned} &E\left\{m(U)K\left(\frac{u-U}{h}\right)\right\} \\ &\to m(u)f(u)\int_{-\infty}^{\infty}K(v)dv \text{ as } h \to 0, \end{aligned} \tag{20}$$

we have

$$E\hat{g}(u;t) \to \alpha f(u)m(u)\int_{-\infty}^{\infty}K(v)dv \text{ as } t \to \infty.$$

Since $Y(t) = W(t) + Z(t)$ with $W(t)$ as in (2), we have $\mathrm{var}[\hat{g}(u;t)] = R_1(u;t) + R_2(u;t)$, where

$$\begin{aligned} R_1(u;t) = \frac{1}{t^2}\int_0^t\int_0^t \mathrm{cov}\Bigg[&W(\lambda+\xi)\frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right), \\ &W(\lambda+\eta)\frac{1}{h(\eta)}K\left(\frac{u-U(\eta)}{h(\eta)}\right)\Bigg]d\xi d\eta, \end{aligned}$$

$$\begin{aligned} R_2(u;t) = \frac{1}{t^2}\int_0^t\int_0^t \mathrm{cov}\Bigg[&Z(\lambda+\xi)\frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right), \\ &Z(\lambda+\eta)\frac{1}{h(\eta)}K\left(\frac{u-U(\eta)}{h(\eta)}\right)\Bigg]d\xi d\eta. \end{aligned}$$

By virtue of Lemma 4 in Appendix,

$$R_1(u;t) = O\left(\frac{1}{t^2}\int_0^t \frac{1}{h(\tau)}d\tau\right).$$

Since processes $Z(.)$ and $U(.)$ are mutually independent,

$$\begin{aligned} R_2(u;t) &= \sigma_Z^2\frac{1}{t^2}\int_0^t \mathrm{var}\left[\frac{1}{h(\eta)}K\left(\frac{u-U(\eta)}{h(\eta)}\right)\right]d\eta \\ &\le \sigma_Z^2\frac{1}{t^2}\int_0^t \frac{1}{h^2(\eta)}E\left\{K^2\left(\frac{u-U(\eta)}{h(\eta)}\right)\right\}d\eta \\ &= \sigma_Z^2\frac{1}{t^2}O(1)\int_0^t \frac{1}{h(\tau)}d\tau. \end{aligned}$$

The last equality is a consequence of Lemma 6 in Appendix. Thus

$$\mathrm{var}[\hat{g}(u;t)] = O\left(\frac{1}{t^2}\int_0^t \frac{1}{h(\tau)}d\tau\right).$$

Finally,

$$\hat{g}(u;t) \to \alpha m(u)f(u)\int_{-\infty}^{\infty}K(v)dv \text{ as } t \to \infty \text{ in probability.}$$

Since, using similar arguments, one can verify that

$$\hat{f}(u;t) \to f(u)\int_{-\infty}^{\infty}K(v)dv \text{ as } t \to \infty$$

in probability at the point, the theorem has been verified. ■

Theorem 2 *Let the kernel K satisfy (11)–(13). If h satisfies (8) and (10), then*

$$\tilde{\mu}(u;t) \to \alpha m(u) \text{ as } t \to \infty \text{ in probability}$$

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

Proof Using Lemma 1, we get

$$\begin{aligned} &E\tilde{g}(u;t) \\ &= \frac{1}{\int_0^t h(\tau)d\tau}\int_0^t E\left\{Y(\lambda)K\left(\frac{u-U(0)}{h(\zeta)}\right)\right\}d\zeta \\ &= \alpha\frac{1}{\int_0^t h(\tau)d\tau}\int_0^t E\left\{m(U)K\left(\frac{u-U}{h(\zeta)}\right)\right\}d\zeta. \end{aligned}$$

Since (20) holds and (10) is satisfied,

$$E\tilde{g}(u;t) \to \alpha f(u)m(u)\int_{-\infty}^{\infty}K(v)dv \text{ as } t \to \infty.$$

Our analysis of the $\mathrm{var}[\tilde{g}(u;t)]$ is similar to that in Theorem 1. The only difference is that we use Lemma 5 rather than Lemma 4. ■

Arguing as in Greblicki (2000), one can verify that for m and f having three derivatives $\hat{\mu}(u;t) - \alpha m(u) = O(t^{-2/5})$ as $t \to \infty$ in probability provided that $h(t) \sim t^{-1/5}$. The kernel K is such that $\int K(v)dv = 0$ and $\int vK(v)dv = 0$. Similar result holds for $\tilde{\mu}(u;t)$. Details are left for the reader.

3.3. Global error

We now assume that m and f are continuous in the whole real line. In addition, $U(.)$ as well as disturbance $Z(.)$ are all bounded, i.e., that, for all t, $|U(t)| \leq c_1$ and $|Z(t)| \leq c_2$ with some c_1, c_2. From this and (1), it follows that $|\alpha m(u)| \leq c_4$, for all u and some c_4. Owing to that $|Y(t)| \leq c_5$ for some c_5. Thus, $|\hat{\mu}(u;t)| \leq c_6$ for all u and t. Therefore from (19), it follows that $E\left(\hat{\mu}(u;t) - \alpha m(u)\right)^2 \to 0$ as $t \to \infty$, all u. Hence, applying Lebesgue dominated convergence theorem, we get

$$\int_{-\infty}^{\infty} E\left(\hat{\mu}(u;t) - \alpha m(u)\right)^2 f(u) du \to 0 \text{ as } t \to \infty.$$

For similar reasons

$$\int_{-\infty}^{\infty} E\left(\tilde{\mu}(u;t) - \alpha m(u)\right)^2 f(u) du \to 0 \text{ as } t \to \infty.$$

3.4. Sampled data

We shall now show how to recover $m(u)$ from sampled input and output signals, i.e., from pairs

$$(U(\Delta), Y(\Delta)), (U(2\Delta), Y(2\Delta)), \ldots, (U(n\Delta), Y(n\Delta)),$$

where $\Delta > 0$. As our on-line algorithms recovering $k(\Delta)m(u)$, we now take

$$\hat{\mu}_n(u) = \frac{\displaystyle\sum_{i=1}^{n} Y((i+1)\Delta) \frac{1}{h_i} K\left(\frac{u - U(i\Delta)}{h_i}\right)}{\displaystyle\sum_{i=1}^{n} \frac{1}{h_i} K\left(\frac{u - U(i\Delta)}{h_i}\right)} \tag{21}$$

and

$$\tilde{\mu}_n(u) = \frac{\displaystyle\sum_{i=1}^{n} Y((i+1)\Delta) K\left(\frac{u - U(i\Delta)}{h_i}\right)}{\displaystyle\sum_{i=1}^{n} Y((i+1)\Delta) K\left(\frac{u - U(i\Delta)}{h_i}\right)}, \tag{22}$$

where $\{h_n\}$ is a sequence of positive numbers.

Denoting

$$\hat{g}_n(u) = \frac{1}{n} \sum_{i=1}^{n} Y((i+1)\Delta) \frac{1}{h_i} K\left(\frac{u - U(i\Delta)}{h_i}\right),$$

$$\hat{f}_n(u) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_i} K\left(\frac{u - U(i\Delta)}{h_i}\right),$$

we notice that $\hat{\mu}_n(u) = \hat{g}_n(u)/\hat{f}_n(u)$ and observe that the following recursive formulas can be applied:

$$\hat{g}_n(u) = \hat{g}_{n-1}(u) - \frac{1}{n}\left[\hat{g}_{n-1}(u) - Y((n+1)\Delta) \frac{1}{h_n} K\left(\frac{u - U(n\Delta)}{h_n}\right)\right],$$

$$\hat{f}_n(u) = \hat{f}_{n-1}(u) - \frac{1}{n}\left[\hat{f}_{n-1}(u) - \frac{1}{h_n} K\left(\frac{u - U(n\Delta)}{h_n}\right)\right]$$

with $\hat{g}_0(u) = \hat{f}_0(u) = 0$. One can also verify that $\tilde{\mu}_n(u) = \tilde{g}_n(u)/\tilde{f}_n(u)$, where

$$\tilde{g}_n(u) = \tilde{g}_{n-1}(u) - \gamma_n \left[\tilde{g}_{n-1}(u) - Y((n+1)\Delta) \frac{1}{h_n} K\left(\frac{u - U(n\Delta)}{h_n}\right)\right],$$

$$\tilde{f}_n(u) = \tilde{f}_{n-1}(u) - \gamma_n \left[\tilde{f}_{n-1}(u) - \frac{1}{h_n} K\left(\frac{u - U(n\Delta)}{h_n}\right)\right],$$

with $\gamma_n = h_n / \sum_{i=1}^{n} h_i$, $\tilde{g}_0(u) = \tilde{f}_0(u) = 0$.

Proofs of the following theorems are left to the reader:

Theorem 3 *Let the kernel K satisfy (11), (12), and (13). If,*

$$h_n \to 0 \text{ as } n \to \infty, \tag{23}$$

$$\frac{1}{n^2} \sum_{i} = 1^n \frac{1}{h_i} \to 0 \text{ as } n \to \infty,$$

then

$$\hat{\mu}_n(u) \to k(\Delta) m(u) \text{ as } n \to \infty \text{ in probability}$$

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

Theorem 4 *Let the kernel K satisfy (11), (12), and (13). If, in addition to (23),*

$$\sum_{n} = 1^{\infty} h_i = \infty,$$

then

$$\tilde{\mu}_n(u) \to k(\Delta) m(u) \text{ as } n \to \infty \text{ in probability}$$

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

4. Dynamic subsystem identification

4.1. Pointwise properties

The basis for recovering the impulse response is the following obvious equality

$$\operatorname{cov}\left[Y(t+\tau), U(t)\right] = \beta k(\tau).$$

Thus, as an estimate of $\beta k(\tau)$ we take

$$\hat{\kappa}(\tau;t) = \frac{1}{t}\int_0^t Y(\tau+\xi)U(\xi)d\xi \tag{24}$$

which can be calculated in the following recursive way:

$$\frac{d}{dt}\hat{\kappa}(\tau;t) = \frac{1}{t}\left[\hat{\kappa}(\tau;t) - Y(\tau+t)U(t)\right]$$

with $\hat{\kappa}(\tau;0) \equiv 0$.

Theorem 5 *Let $E\{U^2m^2(U)\} < \infty$. Then*

$$E(\hat{\kappa}(\tau;t) - \beta k(\tau))^2 \to 0 \text{ as } t \to \infty.$$

Proof Since $E\hat{\kappa}(\tau;t) = \beta k(\tau)$, it suffices to examine $\operatorname{var}[\hat{\kappa}(\tau;t)]$ which, owing to (2), equals $(1/t^2)[P_1(\tau;t) + P_2(\tau;t)]$ with

$$\begin{aligned} P_1(\tau;t) &= \operatorname{var}\left[\int_0^t W(\tau+\xi)U(\xi)d\xi\right] \\ &= \int_0^t\int_0^t \operatorname{cov}[W(\tau+\xi)U(\xi), W(\tau+\eta)U(\eta)]d\xi d\eta, \end{aligned}$$

$$\begin{aligned} P_2(\tau;t) &= \operatorname{var}\left[\int_0^t Z(\tau+\xi)U(\xi)d\xi\right] \\ &= \int_0^t\int_0^t \operatorname{cov}[Z(\tau+\xi)U(\xi), Z(\tau+\eta)U(\eta)]d\xi d\eta. \end{aligned}$$

Since, $EU = 0$, by virtue of Lemma 3 in Appendix,

$$\begin{aligned} P_1(\tau;t) &= \int_0^t\int_0^t \operatorname{cov}[W(\tau+\xi)U(\xi), W(\tau+\eta)U(\eta)]d\xi d\eta \\ &= k^2(\tau)t\operatorname{var}\left[Um(U)\right] + \sigma_V^2\sigma_U^2 t\int_0^\infty k^2(\xi)d\xi \\ &\quad + E^2\left\{Um(U)\right\} \\ &\quad \times \int_0^t\int_0^t k(\tau+\xi-\eta)k(\tau-\xi+\eta)d\xi d\eta. \end{aligned}$$

The double integral in the last term equals $2\int_0^t k(\tau+\lambda)k(\tau-\lambda)d\lambda$ and is bounded in absolute value by $2\int_0^\infty k^2(\xi)d\xi$. Therefore, $|P_1(t)| \le c_1 t$,some c_1 independent of t. Since

$$\operatorname{cov}[Z(\tau+\xi)U(\xi), Z(\tau+\eta)U(\eta)] = \sigma_U^2\sigma_Z^2\delta(\xi-\eta),$$

we get $P_2(\tau;t) = \sigma_U^2\sigma_Z^2 t$. Finally, $\operatorname{var}[\hat{\kappa}(\tau;t)] = O(t^{-1})$ which completes the proof. ∎

4.2. Global error

Going through the proof of Theorem 5, we observe that

$$E(\hat{\kappa}(\tau;t) - \beta k(\tau))^2 = (1+k^2(\tau))O(t^{-1}) \tag{25}$$

with $O(.)$ independent of τ, which does not guarantee that the global error $E\int_0^\infty(\hat{\kappa}(\tau;t) - \beta k(\tau))^2 d\tau$ is finite. Nevertheless, for any bounded function w such that $\int_0^\infty |w(\tau)|d\tau < \infty$, the weighted global error vanishes with t increasing to infinity, i.e.,

$$\int_0^\infty E(\hat{\kappa}(\tau;t) - \beta k(\tau))^2 w(\tau)d\tau = O(t^{-1}).$$

In particular, the convergence holds for $w(\tau)$ equal $e^{-\tau}$ or $k(\tau)$.

To avoid introducing a weighting function we can select a function $T(.)$ and define

$$\hat{\kappa}_T(\tau;t) = \begin{cases} \hat{\kappa}(\tau;t), & \text{for } \tau \le T(t) \\ 0, & \text{for } \tau > T(t), \end{cases}$$

as a new, truncated in the time domain, estimate of $\beta k(\tau)$. Owing to (25), its global error is equal or not greater than

$$\begin{aligned} &\int_0^{T(t)} (\hat{\kappa}_T(\tau;t) - \beta k(\tau))^2 d\tau + \beta^2\int_{T(t)}^\infty k^2(\tau)d\tau \\ &\le cT(t)/t + \beta^2\int_{T(t)}^\infty k^2(\tau)d\tau \end{aligned}$$

with some c independent of t. Thus,

$$\int_0^\infty E(\hat{\kappa}_T(\tau;t) - \beta k(\tau))^2 d\tau \to 0 \text{ as } t \to \infty$$

if $\lim_{t\to\infty} T(t)/t = 0$. Selecting $T(t) \sim \ln t$ and observing that, due to stability of the dynamic subsystem, $\int_{T(t)}^\infty k^2(\xi)d\xi = O(e^{-dT(t)})$ with some positive d, we obtain

$$\int_0^\infty E(\hat{\kappa}_T(\tau;t) - \beta k(\tau))^2 d\tau = O(t/\ln t).$$

5. Simulation Example

5.1. Nonlinear subsystem

In the example, the dynamic subsystem has a transfer function $e/(s+1)^2$. Moreover, $\lambda = 1$ and, consequently, $\alpha = 1$. We have applied the characteristic m shown in figure 2.

Input signal is uniformly distributed in the interval $(-2.5, 2.5)$ while Z_n has a Gaussian distribution with zero

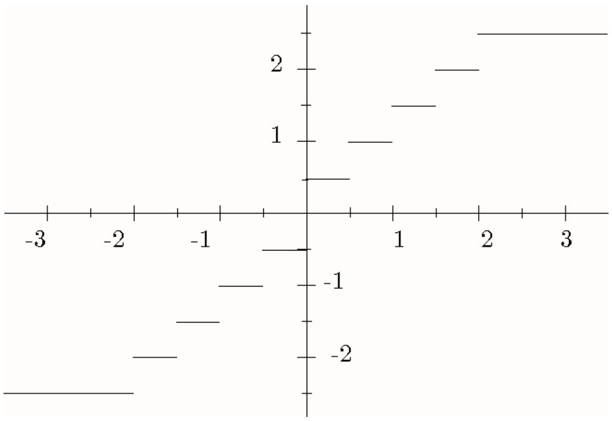

Figure 2: Characteristic m.

mean and variance 1. In both algorithms, the kernel $K(u)$ is a function equal 1 or zero for u in the interval $(-1,1)$ or outside, respectively. As the function h, we have selected $2.5t^{-0.5}$. The MISE has been defined as $E\int_{-2}^{2}(\mu(u)-m(u))^2du$, where μ is (6), (7) or (18), and calculated numerically. Results are shown in figure 3. Differences between algorithms are very small.

5.2. Dynamic subsystem

Now, the nonlinear subsystem has the characteristic $m(u) = \mathrm{sign}(u)$. The transfer function of the dynamic part is $e/(s+1)^2$ which means that its impulse response equals $k(t) = te^{-(t-1)}$. The input signal as well as disturbance have a normal distribution with zero mean and variance 1. For algorithm (24) the MISE defined as $E\int_0^\infty(\hat{\kappa}(\tau;t)-\beta k(\tau))^2 k(\tau)d\tau$ has been calculated and results presented in figure 4. It is not surprising that the error vanishes rapidly with t increasing to infinity.

6. Final remarks

On-line estimates presented here together with an off-line one examined in Greblicki (2000) constitute a set of nonparametric kernel algorithms recovering the characteristic of the nonlinear part of the Hammerstein system. The *a priori* information about the whole system, i.e., about its both subsystems is extremely small, in terms of statistics – nonparametric. The functional form of a nonlinear characteristic is completely unknown. It seems that this is often the case in real situations when our knowledge about identified systems is small and very much uncertain.

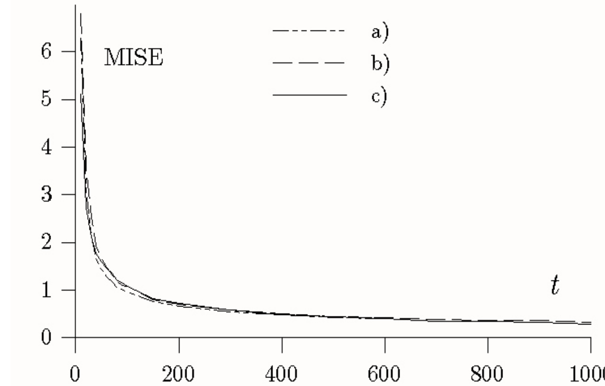

Figure 3: MISE versus t; a) (6), b) (7), c) (18).

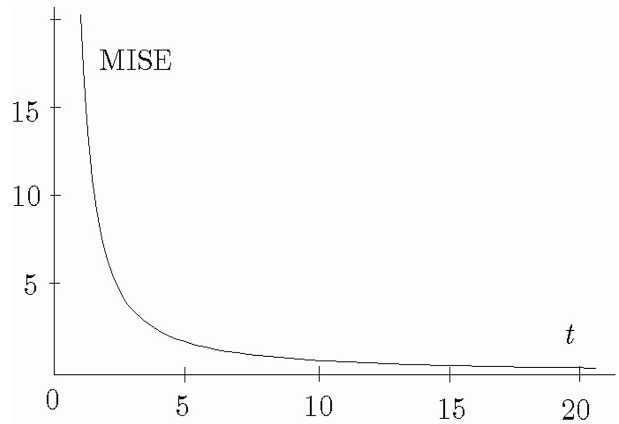

Figure 4: MISE versus t; algorithm (24).

Ackowledgement

The author thanks reviewers for suggestions concerning some improvements.

Appendix

A.1. The System

Using arguments applied in the proof of Lemma 3 in Greblicki (2000) one can easily verify the following lemma:

Lemma 2 *For any Borel functions ϕ and ψ,*

$$\begin{aligned}&\mathrm{cov}\left[W(\xi+\lambda)\phi(U(\xi)), W(\eta+\lambda)\psi(U(\eta))\right]\\&=\mathrm{cov}\left[m(U)\phi(U), m(U)\psi(U)\right]k^2(\lambda)\delta(\xi-\eta)\\&+\int_0^\infty k^2(v)dvE\left\{m^2(U)\right\}E\left\{\phi(U)\psi(U)\right\}\delta(\xi-\eta)\end{aligned}$$

$$+ E\left\{m^2(U)\phi(U)\right\}E\left\{\psi(U)\right\}k(\lambda)k(\eta-\xi+\lambda)$$
$$+ E\left\{\phi(U)\right\}E\left\{m^2(U)\psi(U)\right\}k(\lambda)k(\xi-\eta+\lambda)$$
$$+ E\left\{m(U)\phi(U)\right\}E\left\{m(U)\psi(U)\right\}$$
$$\times k(\xi-\eta+\lambda)k(\eta-\xi+\lambda).$$

Having the lemma we can prove our next one.

Lemma 3 *Let $W(t)$ be as in (2). For any Borel function $\varphi_t(.)$*

$$\int_0^t\int_0^t \operatorname{cov}[W(\xi+\lambda)\varphi_\xi(U(\xi)), W(\eta+\lambda)\varphi_\eta(U(\eta))]d\xi d\eta$$
$$= k^2(\lambda)\int_0^t \operatorname{var}\left[m(U)\varphi_\xi(U)\right]d\xi$$
$$+ \int_0^\infty k^2(v)dvE\left\{m^2(U)\right\}\int_0^t E\left\{\varphi_\xi^2(U)\right\}d\xi$$
$$+ k(\lambda)\int_0^t\int_0^t E\left\{m^2(U)\varphi_\xi(U)\right\}E\left\{\varphi_\eta(U)\right\}$$
$$\times k(\eta-\xi+\lambda)d\xi d\eta$$
$$+ k(\lambda)\int_0^t\int_0^t E\left\{\varphi_\xi(U)\right\}E\left\{m^2(U)\varphi_\eta(U)\right\}$$
$$\times k(\xi-\eta+\lambda)d\xi d\eta$$
$$+ \int_0^t\int_0^t E\left\{m(U)\varphi_\xi(U)\right\}E\left\{m(U)\varphi_\eta(U)\right\}$$
$$\times k(\xi-\eta+\lambda)k(\eta-\xi+\lambda)d\xi d\eta.$$

Lemma 4 *Let $W(t)$ be as in (2). Let K satisfy (11), (12). Then, at every point u at which m and f are continuous,*

$$\left|\int_0^t\int_0^t\left[\operatorname{cov} W(\xi+\lambda)\frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right),\right.\right.$$
$$\left.\left. W(\eta+\lambda)\frac{1}{h(\eta)}K\left(\frac{u-U(\eta)}{h(\eta)}\right)\right]d\xi d\eta\right|$$
$$\le \theta_1(u,\lambda)\int_0^t \frac{1}{h(\tau)}d\tau + \theta_2(u,\lambda)t$$

with some finite $\theta_1(u,\lambda)$ and $\theta_2(u,\lambda)$ independent of both t and h.

Proof Let u be a point at which m and f are continuous. The expression in the assertion equals $\sum_{i=1}^5 S_i(u,t)$ with

$$S_1(u,t) = k^2(\lambda)\int_0^t \operatorname{var}\left[m(U)\frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right)\right]d\xi,$$

$$S_2(u,t) = \frac{1}{t^2}\int_0^\infty k^2(v)dvE\left\{m^2(U)\right\}$$
$$\times\int_0^t \frac{1}{h(\xi)}E\left\{\frac{1}{h(\xi)}K^2\left(\frac{u-U(\xi)}{h(\xi)}\right)\right\}d\xi,$$

$$S_3(u,t) = k(\lambda)\int_0^t\int_0^t E\left\{m^2(U)\frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right)\right\}$$
$$\times E\left\{\frac{1}{h(\eta)}K\left(\frac{u-U(\eta)}{h(\eta)}\right)\right\}k(\eta-\xi+\lambda)d\xi d\eta,$$

$$S_4(u,t) = k(\lambda)\int_0^t\int_0^t E\left\{\frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right)\right\}$$
$$\times E\left\{\frac{1}{h(\eta)}m^2(U)K\left(\frac{u-U(\eta)}{h(\eta)}\right)\right\}$$
$$\times k(\eta-\xi+\lambda)d\xi d\eta,$$

$$S_5(u,t) = \int_0^t\int_0^t E\left\{m(U)\frac{1}{h(\xi)}K\left(\frac{u-U(\xi)}{h(\xi)}\right)\right\}$$
$$\times E\left\{m(U)\frac{1}{h(\eta)}K\left(\frac{u-U(\eta)}{h(\eta)}\right)\right\}$$
$$\times k(\xi-\eta+\lambda)k(\eta-\xi+\lambda)d\xi d\eta.$$

Denoting

$$\rho_1(u) = \sup_{h>0}\frac{1}{h}E\left|K\left(\frac{u-U}{h}\right)\right|,$$

$$\rho_2(u) = \sup_{h>0}\frac{1}{h}E\left|m(U)K\left(\frac{u-U}{h}\right)\right|,$$

and

$$\rho_3(u) = \sup_{h>0}\frac{1}{h}E\left|m^2(U)K\left(\frac{u-U}{h}\right)\right|,$$

we find $\rho_1(u)$, $\rho_2(u)$ and $\rho_3(u)$ all finite, see (27). Therefore we can write

$$|S_1(u,t)| \le k^2(\lambda)$$
$$\times\int_0^t \frac{1}{h^2(\xi)}E\left\{m^2(U)K^2\left(\frac{u-U(\xi)}{h(\xi)}\right)\right\}d\xi$$
$$\le \kappa k^2(\lambda)\rho_3(u)\int_0^t \frac{1}{h(\xi)}d\xi,$$

$$|S_2(u,t)| \le \kappa\int_0^\infty k^2(v)dvE\left\{m^2(U)\right\}\rho_1(u)\int_0^t \frac{1}{h(\xi)}d\xi,$$

$$|S_3(u,t)| \le |k(\lambda)|\int_0^\infty |k(\eta)|\,d\eta\rho_3(u)$$
$$\times\int_0^t E\left\{m^2(U)\frac{1}{h(\xi)}\left|K\left(\frac{u-U(\xi)}{h(\xi)}\right)\right|\right\}d\xi$$
$$\le |k(\lambda)|\int_0^\infty |k(\eta)|\,d\eta\rho_1(u)\rho_3(u)t,$$

$$|S_4(u,t)| \leq |k(\lambda)| \int_0^\infty |k(\eta)|\, d\eta \rho_1(u)\rho_3(u)t,$$

and

$$|S_5(u,t)| \leq \rho_5^2(u) \int_0^t \int_0^t |k(\xi-\eta+\lambda)k(\eta-\xi+\lambda)|\, d\xi d\eta \leq \rho_2^2(u) \int_0^\infty k^2(\tau) d\tau.$$

The proof has been completed. ■

Lemma 5 *Let K satisfy (11), (12). Then, at every point u at which m and f are continuous,*

$$\left| \int_0^t \int_0^t \operatorname{cov}\left[W(\xi+\lambda) K\left(\frac{u-U(\xi)}{h(\xi)}\right), W(\eta+\lambda) K\left(\frac{u-U(\eta)}{h(\eta)}\right)\right] d\xi d\eta \right| \leq \theta_3(u,\lambda) \int_0^t h(\tau) d\tau$$

with some finite $\theta_3(u,\lambda)$ independent of both t and h.

A.2. General result

In Wheeden and Zygmund (1977: Theorems 9.9 and 9.13) we find the following result:

Lemma 6 *Let f be a density function of a random variable U. Let φ be a Borel function such that $E|\varphi(U)| < \infty$. Let a Borel measurable kernel K satisfy (11)–(13). Then aaaaaa*

$$\frac{1}{h} E\left\{\varphi(U) K\left(\frac{u-U}{h}\right)\right\} \to \varphi(u) f(u) \int_{-} \infty^\infty K(v) dv \quad \text{as } h \to 0 \tag{26}$$

at every point u at which both f and φ are continuous and

$$\sup_h > 0 \frac{1}{h} E\left|\varphi(U) K\left(\frac{u-U}{h}\right)\right| < \infty \tag{27}$$

at the same points.

References

AL-DUWAISH, H., and KARIM, M. N., 1997, A new method for the identification of Hammerstein model. *Automatica*, 33, (10), 1871–1875.

AL-DUWAISH, H., KARIM, M. N., and CHANDRASEKAR, V., 1997, Hammerstein model identification by multilayer feedforward neural networks. *International Journal of Systems Science*, 28, (1), 49–54.

BILLINGS, S.A., and FAKHOURI, S. Y., 1977, Identification of nonlinear systems using the Wiener model. *Electronics Letters*, 13, (17), 502–504.

BILLINGS, S.A., and FAKHOURI, S. Y., 1978, Theory of separable processes with applications to the identification of nonlinear systems. *Proceedings of the IEE*, 125, (9), 1051–1058.

BILLINGS, S.A., and FAKHOURI, S. Y., 1979, Non-linear system identification using the Hammerstein model. *International Journal of Systems Science*, 10, (5), 567–578.

BILLINGS, S.A., and FAKHOURI, S. Y., 1982, Identification of systems containing linear dynamic and static nonlinear elements. *Automatica*, 18, (1), 15–26.

CHANG, F. H. I., and LUUS, R., 1971, A noniterative method for identification using Hammerstein model. *IEEE Transactions on Automatic Control*, AC-16, 465–468.

ESKINAT, E., and JOHNSON, S. H., 1991, Use of Hammerstein models in identification of nonlinear systems. *American Institute of Chemical Engineers Journal*, 37, (2), 255–268.

GALLMAN, P. G.,. 1976, A comparison of two Hammerstein model identification algorithms. *IEEE Transactions on Automatic Control*, AC-21, 125–127.

GIRI, F., CHAOUI, F. Z., and ROCHIDI, Y., 2001, Parameter identification of a class of Hammerstein plants. *Automatica*, 37, 749–756.

GREBLICKI, W., 1989, Nonparametric orthogonal series identification of Hammerstein systems. *International Journal of Systems Science*, 20, (12), 2355–2367.

GREBLICKI, W., 1996, Nonlinearity estimation in Hammerstein systems based on ordered observations. *IEEE Transactions on Signal Processing*, 44, (5), 1224–1233.

GREBLICKI, W., 2000, Continuous-time Hammerstein system identification. *IEEE Transactions on Automatic Control*, 45, (6), 1232–1236.

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

GREBLICKI, W., and PAWLAK, M., 1989a, Nonparametric identification of Hammerstein systems. *IEEE Transactions on Information Theory*, IT-35, (2), 409–418.

GREBLICKI, W., and PAWLAK, M., 1989b, Recursive identification of Hammerstein systems. *Journal of the Franklin Institute*, 326, (4), 461–481.

GREBLICKI, W., and PAWLAK, M., 1991, Nonparametric identification of a cascade nonlinear time series system. *Signal Processing,* 22, 61–75.

GREBLICKI, W., and PAWLAK, M., 1994a, Dynamic system identification with order statistics. *IEEE Transactions on Information Theory*, 40, (5), 1474–1489.

GREBLICKI, W., and PAWLAK, M., 1994b, Nonparametric recovering nonlinearities in block-oriented systems with the help of Laguerre polynomials. *Control-Theory and Advanced Technology*, 10, (4, Part A), 771–791.

HAIST, N. D., CHANG, F. H. I., and LUUS, R., 1973, Nonlinear identification in the presence of correlated noise using a Hammerstein model. *IEEE Transactions on Automatic Control*, AC-18, 553–555.

HASIEWICZ, Z., 1999, Hammerstein system identification by the Haar multiresolution approximation. *International Journal of Adaptive Control and Signal Processing*, 13, 691–717.

HASIEWICZ, Z., 2001, Non-parametric estimation of non-linearity in a cascade time-series system by multiscale approximation. *Signal Processing*, 81, 791–807.

HSU, J. T., and NGO, K. D. T., 1997, A Hammerstein-based dynamic model for hysteresis phenomenon. *IEEE Transactions on Power Electronics*, 12, (3), 406–413.

KAMINSKAS, V. A., 1975, Parameter estimation of discrete systems of the Hammerstein class. *Avtomatica i Telemekhanika*, (7), 63–69, (in Russian).

KORENBERG, M. J., AND HUNTER, I. W., 1986, The identification of nonlinear biological systems: LNL cascade models. *Biological Cybernetics*, 55, 125–134.

KRZYŻAK, A., 1989, Identification of discrete Hammerstein systems by the Fourier series regression estimate. *Internationa Journal of Systems Science,* 20, 1729–1744.

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

KRZYŻAK, A., 1992, Global convergence of the recursive kernel regression estimates with applications in classification and nonlinear system estimation. *IEEE Transactions on Information Theory*, 38, (4), 1323–1338.

KRZYŻAK, A., and PARTYKA, M. A., 1993, On identification of block orientated systems by non-parametric techniques. *International Journal of Systems Science*, 24, (6), 1049–1066.

LANG, Z. Q., 1993, Controller design oriented model identification method for Hammerstein system. *Automatica,* 39, (3), 767–771.

LANG, Z. Q., 1994, On identification of the controlled plants described by the Hammerstein system. *IEEE Transactions on Automatic Control*, 39, (3), 569–573.

LANG, Z. Q., 1997, A nonparametric polynomial identification algorithm for the Hammerstein system. *IEEE Transactions on Automatic Control*, 42, (10), 1435–1441.

LIAO, H. E., and SETHARES, W. A.,1995, Suboptimal identification of nonlinear ARMA models using an orthogonality approach. *IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications*, 42, (1), 14–22.

NARENDRA, K. S., and GALLMAN, P. G., 1966, An iterative method for identification of nonlinear systems using a Hammerstein model. *IEEE Transactions on Automatic Control*, AC-11, 546–550.

PATWARDHAN, R. S., LAKSHMINARAYANAN, S., and SHAH, S. L., 1998, Constrained nonlinear MPC using Hammerstein and Wiener models: PLS framework. *American Institute of Chemical Engineers Journal*, 44, (7), 1611–1622.

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

PAWLAK, M., and HASIEWICZ, Z., 1998, Nonlinear system identification by the Haar multiresolution analysis. *IEEE Transactions on Circuits and Systems, I-Fundamental Theory and Applications*, 45, (9), 945–961.

THATHACHAR, M. A. L., and RAMASWAMY, S., 1973, Identification of a class of non-linear systems. *International Journal of Control*, 18, (4), 741–752.

WHEEDEN, R. L., and ZYGMUND, A., 1977, *Measure and Integral.* (Dekker).

printed in: INTERNATIONAL JOURNAL OF SYSTEMS SCIENCE, Vol. 33, No. 12, pp. 969–977, 2002