Nonlinearity Recovering in Hammerstein System from Short Measurement Sequence

Grzegorz Mzyk

***Abstract*—The problem of data pre-filtering for nonparametric identification of Hammerstein system from short (finite) data set is considered. The two-stage method is proposed. First, the linear dynamic block is identified using instrumental variables technique, and the inverse of the obtained model is used for output filtering. Next, the standard procedure of nonparametric regression function estimation (kernel-based, or using orthogonal series expansion) is applied, involving the filtered output sequence instead of the original one. It is shown, that for small and moderate number of data, the estimation error can be significantly reduced in comparison with standard nonparametric methods. The asymptotic properties of the method (consistency and rate of convergence) remain the same as in the classical versions of nonparametric algorithms.**

***Index Terms*—Hammerstein system, inverse filtering, kernel regression, nonparametric identification, orthogonal series expansion.**

I. Introduction

THIS paper addresses the problem of estimation of the nonlinear static characteristic in Hammerstein system. It has fundamental meaning in practice, particularly in signal processing [6], [15], and automatic control [2], [10]. In the classical, i.e., *parametric*, approach to system identification it is assumed that the nonlinear static and linear dynamic block in Hammerstein system (see Fig. 1) can be described with the use of finite number of unknown parameters. Since the internal signal w_k is not available, the parameters of both subsystems are aggregated (see e.g., [1]) and jointly estimated. Usually it leads to complicated and badly conditioned numerical procedures, and the nonlinearity estimates have systematic approximation error connected with improper model selection. *Nonparametric* approach to Hammerstein system identification, proposed in 1980's (see e.g., [7], [8]) and intensively elaborated till now [9], [10], [13], [17], is based on the regression function estimation. Algorithms involve only learning sequence and hence are free of a risk of false parametric *a priori* knowledge. Nonparametric methods enable moreover decentralization of the system identification task. When the nonlinear characteristic is recovered, the Hammerstein system is treated as a static system with specific disturbance and properties of the linear dynamics are, in some sense, ignored. In spite of their simplicity and good limit properties, the nonparametric estimates in standard versions are however inefficient for small and moderate number of data. The aim of this paper is to propose the modified versions of nonparametric routines, more effective in the use with short data sequence. The inverse filtering approach (see e.g., [4], [5], [19]) is used for improving the performance of the nonparametric estimate of the Hammerstein system nonlinearity. In particular, it is shown that for a finite number of data, the variance of the nonparametric estimate can be significantly reduced without affecting its limit properties. The two stage algorithm for nonlinearity recovering, exploiting the new idea of the combined approach to system identification, is proposed. The parameter form of the static characteristic is not known (in the contrary to [11] and [12]) and the input signal is assumed to be a random process (in contrast to e.g., [16]). The strategy allows to incorporate additional knowledge about the linear dynamic block to improve small sample size properties of the nonparametric estimates of the static characteristic.

Manuscript received April 02, 2009; revised May 20, 2009. First published June 10, 2009; current version published July 01, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Kainam Thomas Wong.

The author is with the Institute of Computer Engineering, Control and Robotics, Wroclaw University of Technology, 50-372 Wroclaw, Poland (e-mail: grzegorz.mzyk@pwr.wroc.pl).

Digital Object Identifier 10.1109/LSP.2009.2024795

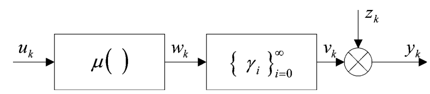

Fig. 1. Hammerstein system.

II. Statement of the Problem

A. Hammerstein System

Denote a static non-linearity as $\mu()$ and the impulse response of a linear dynamics as $\{\gamma_i\}_{i=0}^{\infty}$. The Hammerstein system can now thus be described by the following set of equations: $y_k = v_k + z_k,\ v_k = \sum_{i=0}^{\infty} \gamma_i w_{k-i},\ w_k = \mu(u_k)$, or equivalently

$$y_k = \sum_{i=0}^{\infty} \gamma_i \mu(u_{k-i}) + z_k \tag{1}$$

where u_k and y_k denote the system input and output at time k, respectively, and z_k is the output noise (see Fig. 1).

Assumptions

(A1) The nonlinear characteristic $\mu(u)$ is a Lipschitz function, i.e., it exists a positive constant $l < \infty$, such that for each $u, x \in R$ it holds $|\mu(u) - \mu(x)| \leq l\,|u - x|$.

(A2) The linear dynamics with the unknown impulse response $\{\gamma_i\}_{i=0}^{\infty}$ is stable, i.e., $\sum_{i=0}^{\infty} |\gamma_i| < \infty$, and can be described by the $\mathrm{AR}(p)$ difference equation, i.e., $v_k = \sum_{i=1}^{p} a_i v_{k-i} + b_0 w_k$. The order p is finite and known *a priori*.

(A3) The input $\{u_k\}$ and the noise $\{z_k\}$ are mutually independent i.i.d. random processes, $\sigma_u^2 = \mathrm{var} u_k < \infty$, $\sigma_z^2 = \mathrm{var} z_k < \infty$ and $Ez_k = 0$. It exists the input probability density $f(u)$.

The objective is to recover $\mu(u)$ using input-output measurements $\{(u_k, y_k)\}_{k=1}^{N}$ of the *whole* Hammerstein system.

B. Regression Function

The fundamental meaning in nonparametric estimation of nonlinear block in Hammerstein system has the following equivalence between the regression function $R(u)$ and the characteristic $\mu(u)$

$$R(u) \triangleq E\{y_k|u_k = u\} = \gamma_0\mu(u) + \delta \tag{2}$$

where $\delta = E\mu(u_k) \cdot \sum_{i=1}^{\infty}\gamma_i$. By virtue of (2), the regression function $R(u)$ is the scaled and shifted version of the true static characteristic $\mu(u)$. Since the signal $\{w_k\}$ cannot be measured, the constant γ_0 is not identifiable. Hence, for clarity of exposition and without any loss of generality we can further assume that $\gamma_0 = b_0 = 1$, i.e., $R(u) = \mu(u) + \delta$. The additive constant δ can be determined only under additional knowledge, e.g., when the parametric model of the linear block is given, or when the static characteristic is known at least in one point. In Section III we assume that $\mu(u) = 0$ and hence $\mu(u) = R(u) - R(0)$ (for discussion see [11]). This restriction can be omitted in Section IV.

III. Nonparametric Estimation of the Regression

As was mentioned above, the standard nonparametric methods ([7]–[10]) for nonlinearity recovering work completely independently of the shape of impulse response of the linear dynamics. The cost paid for simplicity, robustness and universality is a high variance of estimates. In the standard approach the Hammerstein system is treated in fact as a nonlinear static element corrupted by a correlated noise. Namely, one can specify three components of the output, i.e.,

$$y_k = \mu(u_k) + \sum_{i=1}^{\infty}\gamma_i\mu(u_{k-i}) + z_k. \tag{3}$$

The most part of the signal y_k is in a sense wasted, because the "system noise" $\xi_k \triangleq \sum_{i=1}^{\infty}\gamma_i\mu(u_{k-i})$ produced by linear dynamics is treated as an additional disturbance. In the Sections III-A and III-B the most important properties of the kernel regression estimates and orthogonal series expansion estimation methods are reminded.

A. Kernel Method

The kernel regression estimate has the form [10], [14]

$$\hat{\mu}_N(u) = \hat{R}_N(u) - \hat{R}_N(0),$$
$$\hat{R}_N(u) = \frac{\sum_{k=1}^{N} y_k K\left(\frac{u_k - u}{h_N}\right)}{\sum_{k=1}^{N} K\left(\frac{u_k - u}{h_N}\right)} \tag{4}$$

where h_N is a bandwidth parameter, which fulfils the following conditions

$$h_N \to 0 \text{ and } Nh_N \to \infty, \text{ as } N \to \infty \tag{5}$$

and $K()$ is a kernel function, such that

$$K(x) \geq 0, \sup K(x) < \infty \text{ and } \int K(x)dx < \infty. \tag{6}$$

Standard examples are shown in the equation at the bottom of the page, and $h_N = h_0 N^{-\alpha}$ with $0 < \alpha < 1$ and positive $h_0 = \text{const}$.

Remark 1: [10] Under (A1)–(A3), (5) and (6), it holds that $\hat{\mu}_N(u) \to \mu(u)$ in probability, as $N \to \infty$, at every u, at which $\mu(u)$ and $f(u)$ are continuous, and $f(u) > 0$. If moreover $\mu(u)$ and $f(u)$ are at least two times continuously differentiable at u, then for $h_N = h_0 N^{-1/5}$ the convergence rate is $|\hat{\mu}_N(u) - \mu(u)| = O(N^{-2/5})$ in probability.

B. Orthogonal Series Expansion Method

Denoting $g(u) \triangleq R(u)f(u)$ one can write $R(u) = g(u)/f(u)$. Let $\{\varphi_i(u)\}_{i=0}^{\infty}$ be the complete set of orthonormal functions in the input domain. If $g()$ and $f()$ are square integrable, then $g(u) = \sum_{i=0}^{\infty}\alpha_i\varphi_i(u)$, $f(u) = \sum_{i=0}^{\infty}\beta_i\varphi_i(u)$, where $\alpha_i = Ey_k\varphi_i(u_k)$ and $\beta_i = E\varphi_i(u_k)$, are orthogonal series representations of $g(u)$ and $f(u)$ in the basis $\{\varphi_i(u)\}_{i=0}^{\infty}$. The standard estimates of the coefficients α_i's and β_i's are $\hat{\alpha}_{i,N} = 1/N\sum_{k=1}^{N} y_k\varphi_i(u_k)$, and $\hat{\beta}_{i,N} = 1/N\sum_{k=1}^{N}\varphi_i(u_k)$, which leads to the following ratio estimate of $\mu(u)$

$$\hat{\mu}_N(u) = \hat{R}_N(u) - \hat{R}_N(0),$$
$$\hat{R}_N(u) = \frac{\sum_{i=0}^{q(N)} \hat{\alpha}_{i,N}\varphi_i(u)}{\sum_{i=0}^{q(N)} \hat{\beta}_{i,N}\varphi_i(u)} \tag{7}$$

where $q(N)$ is some cutoff level [7].

Remark 2: [10] To assure vanishing of the approximation error, the scale $q(N)$ must behave so that $\lim_{N\to\infty} q(N) = \infty$. For the convergence of $\hat{\mu}_N(u)$ to $\mu(u)$, the rate of $q(N)$-increasing must be appropriately slow, e.g., $\lim_{N\to\infty} q^2(N)/N = 0$ for trigonometric or Legendre series, $\lim_{N\to\infty} q^6(N)/N = 0$ for Laguerre series, $\lim_{N\to\infty} q^{5/3}(N)/N = 0$ for Hermite series. Optimal choice of $q(N)$ with respect to the rate of convergence is considered in [7] and [13].

IV. The Proposed Algorithm

Under Assumption (A2) and the fact that $v_k = y_k - z_k$ we obtain

$$\begin{aligned} y_k &= a_1 y_{k-1} + a_2 y_{k-2} + \cdots + a_p y_{k-p} + w_k + \tilde{z}_k \\ &= a_1 y_{k-1} + \cdots + a_p y_{k-p} + Ew_k + \tilde{w}_k + \tilde{z}_k \end{aligned} \tag{8}$$

where $\tilde{w}_k \triangleq w_k - Ew_k$, and $\tilde{z}_k \triangleq z_k - a_1 z_{k-1} - a_2 z_{k-2} - \cdots - a_p z_{k-p}$ are zero-mean stationary random processes. Equation (8) may be rewritten in the form

$$y_k = a_1 y_{k-1} + a_2 y_{k-2} + \cdots + a_p y_{k-p} + c + d_k \tag{9}$$

$$K(x) = I_{[-0.5,0.5]}(x) \triangleq \begin{cases} 1, \text{as } |x| \leq 0.5, \\ 0, \text{elsewhere} \end{cases} \quad (1 - |x|)I_{[-1,1]}(x) \text{ or } \left(\frac{1}{\sqrt{2\pi}}\right)e^{-x^2/2}$$

where $c = Ew_k$ is some (not informative) constant, and

$$d_k = \tilde{z}_k + \tilde{w}_k \tag{10}$$

may be interpreted as a zero-mean correlated random disturbance. Introducing the regressor $\phi_k = (y_{k-1}, y_{k-2}, \ldots, y_{k-p}, 1)^T \in R^{(p+1)\times 1}$, and the vector of unknown parameters

$$\theta = (a_1, a_2, \ldots, a_p, c)^T \in R^{(p+1)\times 1} \tag{11}$$

we get $y_k = \phi_k^T \theta + d_k$, $k = 1, 2, \ldots, N$, or in the compact, matrix-vector version $Y_N = \Phi_N \theta + D_N$, where $Y_N = (y_1, y_2, \ldots, y_N)^T \in R^{N\times 1}$, $\Phi_N = (\phi_1, \phi_2, \ldots, \phi_N)^T \in R^{N\times(p+1)}$, and $D_N = (d_1, d_2, \ldots, d_N)^T \in R^{N\times 1}$.

A. The Two Stage Method

The scheme of the proposed procedure is presented as follows.

Step 1. Identify the parameter vector (11) by the instrumental variables method:

$$\hat{\theta}_N = (\hat{a}_{1,N}, \hat{a}_{2,N}, \ldots, \hat{a}_{p,N}, \hat{c}_N)^T = (\Psi_N^T \Phi_N)^{-1} \Psi_N^T Y_N,$$

where

$$\Psi_N = (\psi_1, \psi_2, \ldots, \psi_N)^T \in R^{N\times(p+1)},$$
$$\psi_k = (\psi_{k,1}, \psi_{k,2}, \ldots, \psi_{k,p}, \psi_{k,p+1})^T \in R^{(p+1)\times 1} \tag{12}$$

is additional matrix including instruments ψ_k, which fulfill the following two standard conditions [12]

$$\det E\psi_k \phi_k^T \neq 0 \text{ and } E\psi_k d_k = 0 \tag{13}$$

and perform the following FIR output filtering

$$y_k^f = y_k - \hat{a}_{1,N} y_{k-1} - \hat{a}_{2,N} y_{k-2} - \cdots - \hat{a}_{p,N} y_{k-p}. \tag{14}$$

Step 2. Using the filtered data $\{u_k, y_k^f\}_{k=1}^N$, compute the nonparametric estimate

$$\hat{\mu}_N^f(u) = \frac{\sum_{k=1}^N y_k^f K\left(\frac{u_k - u}{h_N}\right)}{\sum_{k=1}^N K\left(\frac{u_k - u}{h_N}\right)} \tag{15}$$

or

$$\hat{\mu}_N^f(u) = \frac{\sum_{i=0}^{q(N)} \hat{\alpha}_{i,N}^f \varphi_i(u)}{\sum_{i=0}^{q(N)} \hat{\beta}_{i,N} \varphi_i(u)}, \text{ where}$$

$$\hat{\alpha}_{i,N}^f = \frac{1}{N} \sum_{k=1}^N y_k^f \varphi_i(u_k). \tag{16}$$

Remark 3: Step 1 of the procedure can be replaced by any deconvolution method [4], and generalized for a class of invertible ARMA models. We assume AR dynamics and present the instrumental variables method, because each invertible filter can be approximated with arbitrarily small error by $\mathrm{AR}(p)$ model, when p grows large [3].

Remark 4: Conditions (13) mean that the instruments ψ_k should be correlated with output and simultaneously not correlated with the noise. We refer the reader to [12], where the universal method of generation of instruments for Hammerstein system is introduced. In this paper we assume, for shortness, that we know a priori the function $m()$, such that $m \triangleq E\mu(u_k)m(u_k) \neq 0$. In particular, we can choose $m(u) = u, u^3, \ldots$ if $\mu()$ is odd, or $m(u) = |u|, u^2, \ldots$ if $\mu()$ is even (see Remark 1 in [8]). In Section IV-B we show that the instruments of the form $\psi_k = (m(u_{k-1}), m(u_{k-2}), \ldots, m(u_{k-p}), 1)^T$ fulfill (13).

B. Limit Properties

The following theorem holds.

Theorem 1: If $|\hat{\mu}_N(u) - \mu(u)| = O(N^{-\tau})$, $0 < \tau < 1/2$, in probability as $N \to \infty$, then

$$\left|\hat{\mu}_N^f(u) - \mu(u)\right| = O(N^{-\tau}) \tag{17}$$

in probability as $N \to \infty$.

Proof: It is obvious that the system with the input u_k and the filtered output y_k^f also belongs to the class of Hammerstein systems, and has the same static characteristic $\mu(u)$. To prove (17) it remains to show that the resulting linear dynamics (i.e., $\{\gamma_i\}_{i=0}^\infty$ in a cascade with the filter (14)) and the resulting output noise (i.e., $\{z_k\}$ transferred as in (14)) fulfill (A2) and (A3) as $N \to \infty$. Let us emphasize that the parameters of the filter (14) are random. The estimation error in Step 1 has the form $\Delta_N \triangleq \hat{\theta}_N - \theta = \left(1/N \Psi_N^T \Phi_N\right)^{-1} \left(1/N \Psi_N^T D_N\right)$ and under ergodicity of the processes $\{m(u_k)\}$, $\{y_k\}$ and $\{d_k\}$ it holds that $1/N \Psi_N^T \Phi_N = 1/N \sum_{k=1}^N \phi_k \psi_k^T \to E\phi_k \psi_k^T$ and $1/N \Psi_N^T D_N = 1/N \sum_{k=1}^N d_k \psi_k^T \to E d_k \psi_k^T$ with probability 1, as $N \to \infty$. Since $\{d_k\}$ is zero-mean and d_{k_1} is independent of u_{k_2} for any time instants k_1 and k_2 such that $k_1 > k_2$, it holds that $E d_k \psi_k^T = E\left(d_k \cdot (m(u_{k-1}), m(u_{k-2}), \ldots, m(u_{k-p}), 1)\right) = 0$. Using the results presented in [9], concerning dependence between input-output cross-correlation in Hammerstein system and the terms of impulse response we have $\sigma_{m(u)} \sigma_y \gamma_{|i-j|} = E\left\{(y_{k-i} - Ey)\left(m(u_{k-j}) - Em(u)\right)\right\} = E\left\{y_{k-i}\left(m(u_{k-j}) - Em(u)\right)\right\} - EyE\left(m(u_{k-j}) - Em(u)\right) = Ey_{k-i} m(u_{k-j}) - Ey \cdot Em(u)$, where $\sigma_{m(u)}^2 = \mathrm{var} m(u)$, and hence

$$E\phi_k \psi_k^T = \begin{bmatrix} \Gamma & \overline{y} \\ & I \end{bmatrix}$$

where $\overline{y} = (Ey, Ey, \ldots, Ey)^T$, $I = (0, 0, \ldots, 0, 1)$, and

$$\Gamma = \begin{bmatrix} \bar{\gamma}_0 & \bar{\gamma}_1 & \cdots & \bar{\gamma}_{p-1} \\ \bar{\gamma}_1 & \bar{\gamma}_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bar{\gamma}_1 \\ \bar{\gamma}_{p-1} & \cdots & \bar{\gamma}_1 & \bar{\gamma}_0 \end{bmatrix}$$

where $\bar{\gamma}_l = m\sigma_{m(u)} \sigma_y \gamma_l + Ey \cdot Em(u)$. For $\mathrm{AR}(p)$ process $\{y_k\}$ the matrix Γ is of full rank [3], [8]. Since $\det E\phi_k \psi_k^T \neq 0$, the estimate (12) is well defined with probability 1 as $N \to \infty$. Consequently, for $i = 1, 2, \ldots, p$ it holds that, $MSE\left[\hat{a}_{i,N}\right] \triangleq E\left(\hat{a}_{i,N} - a_i\right)^2 = O(1/N)$ and $\left|\sum_{i=1}^p (a_i - \hat{a}_{i,N}) y_{k-i}\right| = O\left(1/\sqrt{N}\right)$ in probability, as $N \to \infty$. Inserting (8) to (14) we get

$$y_k^f = w_k + \tilde{z}_k + \sum_{i=1}^p (a_i - \hat{a}_{i,N}) y_{k-i}. \tag{18}$$

Since for $N \to \infty$ it holds that $\left|\sum_{i=1}^p (a_i - \hat{a}_{i,N}) y_{k-i}\right| = 0$ with probability 1, we obtain $y_k^f = w_k + \tilde{z}_k$ with probability 1, as $N \to \infty$, which guarantees fulfillment of (A2) and (A3). ■

C. Properties for $N < \infty$

Owing to (3) and (18) we have respectively $y_k = w_k + \delta_k$, and $y_k^f = w_k + \delta_k^f$, where $\delta_k \triangleq \sum_{i=1}^{\infty} \gamma_i w_{k-i} + z_k$, and $\delta_k^f \triangleq \tilde{z}_k + \sum_{i=1}^{p}(a_i - \hat{a}_{i,N}) y_{k-i}$. One can show that the variance of the nonparametric regression function estimate is bounded from above as follows

$$\mathrm{var}\hat{\mu}_N(u) \leq c_0 n(N) \mathrm{var}\delta_k \tag{19}$$

where c_0 is some constant dependent of the used kernel or basis functions, and $n(N) = 1/Nh_N$ for kernel estimate or $n(N) = q(N)/N$ for orthogonal methods (for details, see, e.g., [13, App. I–II], and [9, Sec. VI]). The variances of δ_k and δ_k^f which influence the upper bounds [see (19)] of the variances of (4), (7), (15), and (16) have the form

$$\mathrm{var}\delta_k = \sigma_z^2 + \sigma_w^2 \sum_{i=1}^{\infty} \gamma_i^2, \tag{20}$$

$$\mathrm{var}\delta_k^f = \mathrm{var}\left(\tilde{z}_k + \sum_{i=1}^{p}(a_i - \hat{a}_{i,N}) y_{k-i}\right) \leq \leq \sigma_z^2 \left(1 + \sum_{i=1}^{p} a_i^2\right) + \frac{c_v}{N} \tag{21}$$

where c_v is some constant. Equations (20) and (21) illustrate the effect of filtration. If $\sigma_z^2 \left(1 + \sum_{i=1}^{p} a_i^2\right) < \sigma_z^2 + \sigma_w^2 \sum_{i=1}^{\infty} \gamma_i^2$, which is often the case in Hammerstein system, then it exists N_0 such that for all $N > N_0$ it holds that $\mathrm{var}\delta_k^f < \mathrm{var}\delta_k$. The variance of measurement noise is slightly amplified, but the method gets rid of the harmful influence of the dynamics. For example, if $\sigma_z^2 = 1$, $\sigma_w^2 = 100$, $p = 1$, and $a_1 = 0.8$, then $\mathrm{var}\delta_k = 401$ whereas $\mathrm{var}\delta_k^f \simeq 1.64$.

V. Numerical Example

The input u_k and the noise z_k are uniformly distributed on $[-2\pi, 2\pi]$ and $[-0.1, 0.1]$, respectively. We took the AR(1) linear dynamics: $v_k = 0.8 v_{k-1} + w_k$, and the nonlinear characteristic: $w_k = \mu(u_k) = (1/2 + \mathrm{sgn}(u_k))\,(||u_k| - \pi| - \pi) + 2\sin u_k$. In Step 1 we set $\psi_k = (u_{k-1}, u_{k-2}, \ldots, u_{k-p}, 1)^T$. In the kernel-type estimation algorithm (15) we applied the window kernel $K(x) = \begin{cases} 1, & \text{as } |x| < 1 \\ 0, & \text{elsewhere} \end{cases}$, and set $h(N) \sim N^{-1/5}$. In the orthogonal series expansion method (16) we used trigonometric orthonormal system $1/\sqrt{4\pi}$, $1/\sqrt{2\pi}\cos u/2$, $1/\sqrt{2\pi}\sin u/2$, $1/\sqrt{2\pi}\cos 2u/2$, $1/\sqrt{2\pi}\sin 2u/2$, $1/\sqrt{2\pi}\cos 3u/2$, $\ldots$ and set $q(N) \sim N^{1/5}$. Both methods have been compared with their classical versions (4) and (7). The mean integrated squared error has been computed numerically, according the rule $MISE\hat{e}_N(u) = \int_{-2\pi}^{2\pi} \left(\hat{e}_N(u) - \mu(u)\right)^2 du$, where $\hat{e}_N(u)$ stands for $\hat{\mu}_N^f(u)$ and $\hat{\mu}_N(u)$ respectively. The results are presented in Table I. For small and moderate sample sizes the estimation error has been reduced about 15 times.

TABLE I
$MISE$ of the Kernel/Orthogonal Expansion Estimates

N	(4)	(15)	(7)	(16)
50	37.37	2.81	41.90	3.93
100	27.11	2.18	34.19	3.01
300	15.48	0.89	17.30	2.81
500	9.77	0.57	12.12	2.67
1000	7.33	0.46	9.06	2.20
3000	4.42	0.22	4.87	1.25

VI. Conclusions

Additional prior knowledge about the linear dynamic block allows to speed up the convergence of nonparametric regression-type estimates of nonlinearity in Hammerstein system. The proper output filtering reduces the estimation error for small and moderate number of measurements, and does not affect the limit properties even if the assumed model of the linear dynamics is not correct. In the light of this, nonparametric methods with data pre-filtering are worth further studies.

References

[1] E. W. Bai, "An optimal two-stage identification algorithm for Hammerstein–Wiener nonlinear systems," *Automatica*, vol. 34, no. 3, pp. 333–338, 1998.
[2] E. W. Bai and M. Fu, "A blind approach to Hammerstein model identification," *IEEE Trans. Signal Process.*, vol. 50, pp. 1610–1619, 2002.
[3] G. Box and G. Jenkins, *Time Series Analysis—Forecasting and Control*. San Francisco, CA: Holden-Day, 1976.
[4] J. A. Cadzow, "Blind deconvolution via cumulant extrema," *IEEE Signal Process. Mag.*, vol. 13, pp. 24–42, May 1996.
[5] C. C. Feng and C. Y. Chi, "Performance of cumulant based inverse filters for blind deconvolution," *IEEE Trans. Signal Process.*, vol. 47, pp. 1922–1935, 2002.
[6] G. B. Giannakis and E. Serpedin, "A bibliography on nonlinear system identification," *Signal Process.*, vol. 81, pp. 533–580, 2001.
[7] W. Greblicki, "Non-parametric orthogonal series identification of Hammerstein systems," *Int. J. Syst. Sci.*, vol. 20, no. 12, pp. 2355–2367, 1989.
[8] W. Greblicki and M. Pawlak, "Identification of discrete Hammerstein systems using kernel regression estimates," *IEEE Trans. Automat. Control*, vol. 31, pp. 74–77, 1986.
[9] W. Greblicki and M. Pawlak, "Cascade non-linear system identification by a non-parametric method," *Int. J. Syst. Sci.*, vol. 25, no. 1, pp. 129–153, 1994.
[10] W. Greblicki and M. Pawlak, *Nonparametric System Identification*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
[11] Z. Hasiewicz and G. Mzyk, "Combined parametric–nonparametric identification of Hammerstein systems," *IEEE Trans. Automat. Control*, vol. 49, pp. 1370–1376, 2004.
[12] Z. Hasiewicz and G. Mzyk, "Hammerstein system identification by non-parametric instrumental variables," *Int. J. Control*, vol. 82, pp. 440–455, 2009.
[13] Z. Hasiewicz, M. Pawlak, and P. Śliwiński, "Nonparametric identification of nonlinearities in block-oriented systems by orthogonal wavelets with compact support," *IEEE Trans. Circuits Syst. I—Regular Papers*, vol. 52, no. 2, pp. 427–442, 2005.
[14] W. Härdle, *Applied Nonparametric Regression*. Cambridge, U.K.: Cambridge Univ. Press, 1990.
[15] I. W. Hunter and M. J. Korenberg, "The identification of nonlinear biological systems: Wiener and Hammerstein cascade models," *Biol. Cybern.*, vol. 55, pp. 135–144, 1986.
[16] J. C. Jeng and H. P. Huang, "Nonparametric identification for control of MIMO Hammerstein systems," *Ind. Eng. Chem. Res.*, vol. 47, pp. 6640–6647, 2008.
[17] G. Mzyk, "Generalized kernel regression estimate for the identification of Hammerstein systems," *Int. J. Appl. Math. Comput. Sci.*, vol. 17, no. 2, pp. 101–109, 2007.
[18] P. Stoica and T. Söderström, "Instrumental variable methods for system identification," *Circuits Syst. Signal Process.*, vol. 21, no. 1, pp. 1–9, 2002.
[19] J. K. Tugnait, "Estimation of linear parametric models using inverse filter criteria and higher order statistics," *IEEE Trans. Signal Process.*, vol. 41, pp. 3196–3199, 1993.