Parameter Identification of Nonlinearities in Hammerstein Systems with the Help of Nonparametric Regression Methods

Zygmunt Hasiewicz and Grzegorz Mzyk

Abstract— A combined, parametric-nonparametric routine for the identification of a static part of Hammerstein system is presented. Parameters of the input nonlinear characteristic of Hammerstein system are estimated for a wide range of random excitations and random noise, and without any knowledge of the parametric model of the output linear dynamics. The needed unmeasurable interaction inputs are estimated in a preliminary step by the nonparametric regression function estimation method. Next, they are used in the nonlinear optimization procedure for evaluating parameters of the static subsystem. Broad class of nonlinear characteristics including functions which are not linear in the parameters, as well as the infinite length impulse response of the linear dynamics are admitted. It is shown that the resulting parameter estimates are consistent for both white and colored noise. The analytical findings are validated using numerical simulation results.

Index Terms— Hammerstein system, nonparametric regression, kernel estimation, nonlinear least squares, Levenberg-Marquardt method.

I. INTRODUCTION

The decentralized approach to the identification of block-oriented complex systems seems to be most natural and desirable, as such an approach corresponds directly to the own nature of systems, composed of individual elements distinguished in the block structure. The Hammerstein system, built of a static non-linearity and a linear dynamics connected in a cascade (Fig. 1), is the simplest structure in the class and hence for the most part considered in the system identification literature (see e.g. [3] for the bibliography). Unfortunately, the popular, parametric, methods elaborated for Hammerstein system identification do not enable full decentralization of the system identification task, i.e. independent identification of a static nonlinearity and a linear dynamics in a completely decomposed manner – first of all, because of inaccessibility for measurements of the inner interconnection signal. They assume that the description of system components, i.e. of a static nonlinearity and a linear dynamics is known up to the parameters (a polynomial model along with the FIR dynamics representation is usually used) and these parameters are "glued" when using the measured input-output data of the overall system (e.g. [1], [2]). On the other hand, in a nonparametric setting (the second class of the existing identification methods, see, e.g. [5], [12]) no preliminary assumptions concerning the structure of subsystems are used and only the data decide about characteristics of the system components, but then any possible a priori knowledge about the description of subsystems is inevitably lost. We propose a method where the two approaches are combined. Namely, our idea is to join the results obtained in the nonparametric identification of nonlinear characteristics in Hammerstein systems (see the papers cited above), e.g. by using kernel regression or orthogonal series methods ([5], [9]), with parametric knowledge of subsystems and standard results concerning nonlinear optimization methods (see, for instance, [10]), taking advantages of both. The paper is an extension of [6], where the combined parametric-nonparametric algorithm was proposed for the identification of parameters appearing linearly in the static nonlinear element. We generalize the approach introduced in [6] in the sense that we admit the static characteristic $\mu()$ not linear in the parameters. Similarly as in [6], the presented identification algorithm is two stage. Stage 1 of the algorithm (nonparametric), consists in nonparametric estimation of interaction inputs $w_k = \mu(u_k)$ (Fig. 1) to cope which their inaccessibility for direct measurements. In stage 2 (parametric), using the obtained estimates $\widehat{w}_k$ of w_k, we identify parameters of the static characteristic, and no a priori knowledge of the linear dynamics is required. The latter may in particular be extremely inaccurate. In our consideration we assume that the system input is a random process, and that the random output noise can possess an arbitrary correlation structure. Due to the general form of subsystems and correlation of the output noise, the linear least squares approach presented in [6] fails, and the nonlinear least squares technique is applied instead of. We show that such an approach is computationally convenient and effective, i.e. leads to consistent models. Sufficient conditions for convergence in probability of the resulting parameter estimates are established. Performance of the method for moderate number of data illustrate simulation examples. We focus in the paper on the identification of a static nonlinear characteristic of a Hammerstein system because of the possible diversity of nonlinearities and that the nonlinearity is in fact the main distinguishing feature of each system. In turn, dynamical parts are each time linear.

The authors are with The Institute of Computer Engineering, Control and Robotics, Wrocław University of Technology, Janiszewskiego 11/17, 50-372 Wrocław, Poland, phone: (+48 71) 320-32-77, zygmunt.hasiewicz(grzegorz.mzyk)@pwr.wroc.pl

II. STATEMENT OF THE PROBLEM

A. Hammerstein system

We consider the discrete-time Hammerstein system as in Fig. 1 where u_k, y_k and z_k are respectively the input,

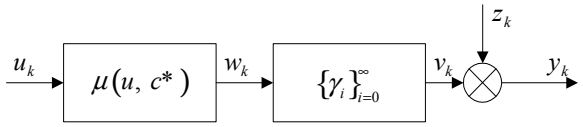

Fig. 1. The Hammerstein system.

output, and noise at time k, and w_k is the interaction input not available for measurement (see [2] for the discussion).

B. General assumptions

In what follows we assume the following.

Assumption 1: The nonlinear characteristic $\mu(u)$ is a Borel measurable function such that

$$|\mu(u)| \leqslant p_1 + p_2 |u| \tag{1}$$

where p_1 and p_2 are some positive constants (cf. [4]).

Assumption 2: The form of a static nonlinearity (1) is known up to the parameters, i.e. we are given the function $\mu(u, c)$ such that $\mu(u, c^*) = \mu(u)$ (Fig. 1), where $c^* = (c_1^*, c_2^*, ..., c_m^*)$ is a vector of unknown true parameters of the nonlinearity. The function $\mu(u, c)$ is by assumption differentiable with respect to c, and the gradient $\bigtriangledown_c \mu(u, c)$ is bounded

$$\|\bigtriangledown_c \mu(u, c)\| \leqslant G_{\max} < \infty$$

Assumption 3: The linear dynamics

$$v_k = \sum_{i=0}^{\infty} \gamma_i w_{k-i} \tag{2}$$

is asymptotically stable, with the unknown impulse response $\{\gamma_i\}_{i=0}^{\infty}$, (such that $\sum_{i=0}^{\infty} |\gamma_i| < \infty$).

Assumption 4: The input signal $\{u_k\}$ is an i.i.d. bounded random process $|u_k| < u_{\max}$, some $u_{\max} > 0$, and there exists (but is unknown) a probability density of u_k, say $\nu(u)$.

Assumption 5: The output noise $\{z_k\}$ is a random, in general correlated, process governed by the equation

$$z_k = \sum_{i=0}^{\infty} \omega_i \varepsilon_{k-i} \tag{3}$$

where $\{\varepsilon_k\}$ is a bounded stationary zero-mean white noise ($E\varepsilon_k = 0$, $|\varepsilon_k| < \varepsilon_{\max} < \infty$), independent of the input signal $\{u_k\}$; $\sum_{i=0}^{\infty} |\omega_i| < \infty$ and $\{\omega_i\}_{i=0}^{\infty}$ is unknown. Thus $\{z_k\}$ is a stationary zero-mean and bounded process $|z_k| \leqslant z_{\max}$, where $z_{\max} = \varepsilon_{\max} \sum_{i=0}^{\infty} |\omega_i|$.

Assumption 6: $\mu(u_0)$ is known at some point u_0 and $\gamma_0 = 1$.

Since, as was explained in detail in [6], Assumption 6 is taken only for technical reasons, hence we shall further assume for convenience that $u_0 = 0$ and $\mu(0) = 0$, without loss of generality.

The aim is to discover the true parameters $c^* = (c_1^*, c_2^*, ..., c_m^*)^T$ of the nonlinear subsystem using a set of input-output data $\{(u_k, y_k); k = 1, 2, ..., N\}$ obtained from the whole Hammerstein system in an identification experiment.

III. ESTIMATION OF NONLINEARITY PARAMETERS

For a collection of N_0 inputs $\{u_n; n = 1, 2, ..., N_0\}$ let us denote

$$W_{N_0} = (w_1, w_2, ..., w_{N_0})^T \tag{4}$$

where $w_n = \mu(u_n, c^*)$ and introduce the following vector

$$\overline{\mu}_{N_0}(c) = [\mu(u_1, c), \mu(u_2, c), ...\mu(u_{N_0}, c)]^T$$

Obviously, the square of Euclidean norm of the difference between $\overline{\mu}_{N_0}(c)$ and W_{N_0}, i.e. the index

$$Q_{N_0}(c) = \left\|\overline{\mu}_{N_0}(c) - W_{N_0}\right\|^2$$

is minimized by the vector c^* of true parameters of the nonlinearity, i.e.

$$c^* = \arg\min \left\|\overline{\mu}_{N_0}(c) - W_{N_0}\right\|^2$$

and hence the identification routine may be based on the minimization of the following (weighted "sum-of-squared errors") loss function

$$\overline{Q}_{N_0}(c) = \frac{1}{N_0} \sum_{n=1}^{N_0} [w_n - \mu(u_n, c)]^2 \tag{5}$$

However, for the sake of inaccessibility of interactions $\{w_n\}_{n=1}^{N_0}$ appearing in (5), such a direct minimization of (5) w.r.t. c is not possible, but instead an indirect two stage approach can be proposed where, in stage 1, a nonparametric estimation of w_n's is carried out yielding the estimates $\widehat{w}_n$, and next, in stage 2, a parametric minimization of $\overline{Q}_{N_0}(c)$ is completed using the obtained $\widehat{w}_n$ instead of w_n. The arising identification scheme is then the following.

Stage 1 (nonparametric): On the basis of M additional input-output measurements $\{(u_k, y_k)\}_{k=1}^{M}$, for the selected N_0 input points $\{u_n;\ n = 1, 2, ..., N_0\}$ appearing in the loss function (5), estimate corresponding interactions $\{w_n = \mu(u_n, c^*);\ n = 1, 2, ..., N_0\}$ as

$$\widehat{w}_{n,M} = \widehat{R}_M(u_n) - \widehat{R}_M(0), \tag{6}$$

where $\widehat{R}_M(u)$ is a consistent nonparametric estimate of the regression function $R(u) = E[y_k | u_k = u]$. Estimation points $\{u_n\}_{n=1}^{N_0}$ may be the measured data, or arbitrarily selected points (there is no formal difference between such choices; see [6]).

Stage 2 (parametric): Plug in the estimates $\widehat{w}_{n,M}$ obtained in stage 1 to the loss function (5) in place of w_n, and minimize the quality index

$$\widehat{Q}_{N_0,M}(c) = \frac{1}{N_0} \sum_{n=1}^{N_0} [\widehat{w}_{n,M} - \mu(u_n, c)]^2 \tag{7}$$

getting the solution $\widehat{c}_{N_0,M}$.

The idea behind stage 1 (already explained in detail in [6]) originates from the fact that under Assumptions 1÷5 it holds that

$$R(u) = \gamma_0 \mu(u) + d$$

where $d = E\mu(u_1)\sum_{i=1}^{\infty}\gamma_i$, and $\mu()$ is the Hammerstein system nonlinearity, and under Assumptions 6 and 2 along with the fact that by assumption $\mu(0) = 0$, we further get that (see [6])

$$\mu(u, c^*) = R(u) - R(0)$$

Such a routine is consistent, i.e. provides the estimate which converges (as $M \to \infty$) to the true parameter vector c^*. We emphasize that identification of c^* can be performed without any prior knowledge about a dynamic part of the system. The following theorem refers to this property.

Theorem 1: If in stage 1 $\widehat{R}_M(0) \to R(0)$ and $\widehat{R}_M(u_n) \to R(u_n)$, i.e. $\widehat{w}_{n,M} \to w_n$, in probability as $M \to \infty$ for $n = 1, 2, ..., N_0$, then also

$$\widehat{c}_{N_0,M} \to c^* \text{ as } M \to \infty \tag{8}$$

in probability.

Sketch of the proof. From triangle inequality we obtain (neglecting the scaling constant $\frac{1}{N_0}$)

$$\begin{aligned}
\widehat{Q}_{N_0,M}(c) &= \left\|\widehat{W}_{N_0,M} - \overline{\mu}_{N_0}(c)\right\|^2 = \\
&= \left\|\widehat{W}_{N_0,M} - W_{N_0} + W_{N_0} - \overline{\mu}_{N_0}(c)\right\|^2 \leqslant \\
&\leqslant 2\left\|\widehat{W}_{N_0,M} - W_{N_0}\right\|^2 + 2\left\|W_{N_0} - \overline{\mu}_{N_0}(c)\right\|^2
\end{aligned}$$

After some transformations, following the lines of the proof of Theorem 1 in [11], we get

$$\sup\left|\widehat{Q}_{N_0,M}(c) - Q_{N_0}(c)\right| \leqslant \left\|\widehat{W}_{N_0,M} - W_{N_0}\right\|^2$$

Hence if $M \to \infty$, then $\widehat{Q}_{N_0,M}(c)$ and $Q_{N_0}(c)$ become equivalent for each c and the convergence (8) holds. ■

The class of consistent nonparametric regression function estimates $\widehat{R}_M(u_n)$ which were elaborated up to now in the literature for Hammerstein systems and can be employed in stage 1 encompasses kernel and orthogonal series estimates, including that using wavelet functions. As an example we present below the kernel estimate and the orthogonal series regression estimate, very easy for computation.

Example 1) Kernel regression function estimate (studied in [5], [15]), has the form

$$\begin{aligned}
\widehat{R}_M(u) &= \frac{\sum_{k=1}^{M} y_k K(\frac{u-u_k}{h(M)})}{\sum_{k=1}^{M} K(\frac{u-u_k}{h(M)})} = \\
&= \sum_{k=1}^{M} \frac{K(\frac{u-u_k}{h(M)})}{\sum_{k=1}^{M} K(\frac{u-u_k}{h(M)})} \cdot y_k
\end{aligned} \tag{9}$$

where $K(u)$ is a kernel (weighting) function and $h(M)$ is a bandwidth parameter controlling the range of data used for estimating a regression function $R(u)$ at a given point u. Standard examples are $K(u) = I_{[-0.5,0.5]}(u)$, $(1 - |u|)I_{[-1,1]}(u)$ or $\left(1/\sqrt{2\pi}\right)e^{-u^2/2}$ and $h(M) = const \cdot M^{-\alpha}$ with $0 < \alpha < 1$. Owing to the convergence results worked out in [15], we find that for each of the above kernels $K(u)$ it holds that $\widehat{R}_M(u) \to R(u)$ pointwise in probability as $M \to \infty$, and that the convergence takes place at every point $u \in Cont(\mu, \nu)$, the set of continuity points of $\mu(u)$ and $\nu(u)$, at which $\nu(u) > 0$, where $\nu(u)$ is a probability density function of the system input (assumed to exist, cf. Assumption 4). Applying in particular the Gaussian kernel $K(u) = \left(1/\sqrt{2\pi}\right)e^{-u^2/2}$ and taking, according to the recommendation in [5], $h(M) \sim M^{-1/5}$ we attain the convergence rate $\left|\widehat{R}_M(u) - R(u)\right| = O(M^{-2/5})$ in probability, provided that $\mu(u)$ and $\nu(u)$ are at the estimation points $u \in \{0, u_n; n = 1, 2, ..., N_0\}$ (see (6)) at least two times continuously differentiable functions and $\nu(u) > 0$ there (cf. [5]).

Example 2) Orthogonal series regression estimate studied e.g. in [4] is based on the observation that

$$\mu(u) = \frac{g(u)}{\nu(u)}$$

where $g(u) = \mu(u)\nu(u)$, and for square integrable $g(u)$ and $\nu(u)$ and proper orthonormal basis $\{\phi_i\}_{i=0}^{\infty}$ we have

$$g(u) = \sum_{i=0}^{\infty} a_i \phi_i(u), \qquad f(u) = \sum_{i=0}^{\infty} b_i \phi_i(u)$$

where it easily appears that

$$a_i = E y_k \phi_i(u_k) \text{ and } b_i = E\phi_i(u_k)$$

The latter may be simply estimated in the following way

$$\widehat{a}_i = \frac{1}{M}\sum_{k=1}^{M} y_k \phi_i(u_k) \text{ and } \widehat{b}_i = \frac{1}{M}\sum_{k=1}^{M} \phi_i(u_k)$$

leading to the nonlinearity estimate of the form

$$\widehat{\mu}_M(u) = \frac{\sum_{i=0}^{q(M)} \widehat{a}_i \phi_i(u)}{\sum_{i=0}^{q(M)} \widehat{b}_i \phi_i(u)}$$

For assuring convergence of $\widehat{\mu}_M(u)$ to $\mu(u)$ in probability, the rate of increasing of the 'cut-off' factor $q(M)$ with M must be appropriately slow, e.g. $q(M) \to \infty$ as $M \to \infty$, but $q^2(M)/M \to 0$ for trigonometric or Legendre series, $q^6(M)/M \to 0$ for Laguerre series and $q^{5/3}(M)/M \to 0$ for Hermite series [9].

For a general treatment of nonparametric regression function estimation methods we send the reader to [8] and [15].

As regards the optimization task (unconstrained nonlinear least squares) to be solved in stage 2, we can use standard Levenberg-Marquardt method. Minimization of $\widehat{Q}_{N_0,M}(c)$ in stage 2 of the identification procedure is

then performed with the use of the following iterative routine:

$$\begin{aligned}\widehat{c}^{(i+1)}_{N_0,M} &= \widehat{c}^{(i)}_{N_0,M} - \\ &(J^T(\widehat{c}^{(i)}_{N_0,M})J(\widehat{c}^{(i)}_{N_0,M}) + \lambda_i I)^{-1} \cdot \\ &\cdot J^T(\widehat{c}^{(i)}_{N_0,M})r(\widehat{c}^{(i)}_{N_0,M})\end{aligned}$$

where $J(c)$ is the Jacobian matrix of the form $J(c)[n,j] = \frac{\partial r_n(c)}{\partial c_j}$ $(n = 1,2,...,N_0;\ j = 1,2,...,m)$, where $r_n(c) = \mu(u_n,c) - \widehat{w}_{n,M}$ and $r(c) = (r_1(c), r_2(c), ..., r_{N_0}(c))^T$. The λ_i's are weighting coefficients (continuation parameters) modified according to the rule

$$\lambda_{i+1} = \begin{cases} \lambda_i \cdot \nu, \text{ if } \widehat{Q}_{N_0,M}(\widehat{c}^{(i+1)}_{N_0,M}) \geqslant \widehat{Q}_{N_0,M}(\widehat{c}^{(i)}_{N_0,M}) \\ \lambda_i/\nu, \text{ otherwise} \end{cases}$$

where $\nu > 1$ (see [10] and [13] for various implementations and discussion of Levenberg-Marquardt method).

The optimization in stage 2 can be also performed with use of heuristic methods, e.g. simulated annealing, tabu search or genetic algorithms.

IV. SIMULATION EXAMPLES

A. Nonlinearity recovering under incorrect a priori knowledge of the linear dynamics

The Hammerstein system with the following static nonlinearity was simulated:

$\mu(u,c^*) = u + u^2 - u^3$ i.e. $\mu(u,c^*) = c^{*T}\phi(u)$, where $\phi(u) = (u, u^2, u^3)^T$ and $c^* = (1,1,-1)^T$;

We compared efficiency of our two-stage routine with the instrumental variables approach proposed by Söderström and Stoica in [14], assuming incorrect a priori knowledge about the linear dynamics structure. Namely, the true dynamic subsystem was in the experiment $v_k = 0.5v_{k-1} + w_k + 0.5w_{k-1}$ but the parametric model of the linear dynamics was assumed as $v_k = \beta_1 v_{k-1} + \alpha_0 w_k$. The system input and the output noise were generated as follows $u_k \sim U[-10,10]$ and $\varepsilon_k \sim U[-0.1,0.1]$, and the aim was to estimate true parameters c^* of the nonlinearity $\mu(u,c)$. To compare our two stage identification algorithm with the instrumental variables in [14], the overall input-output parametric model has been computed

$$\begin{aligned}v_k^{(m)} &= A_1 v_{k-1} + B_{01}u_k + B_{02}u_k^2 + B_{03}u_k^3 \triangleq \\ &\triangleq \phi_k^{(ss)T} p\end{aligned}$$

where

$$\phi_k^{(ss)} = (v_{k-1}, u_k, u_k^2, u_k^3)^T,\ p = (p_1,p_2,p_3,p_4)^T$$

and

$$\begin{aligned}p_1 &= A_1 = \beta_1 \qquad p_2 = B_{01} = \alpha_0 c_1 \\ p_3 &= B_{02} = \alpha_0 c_2 \qquad p_4 = B_{03} = \alpha_0 c_3\end{aligned}$$

and implemented along with the estimate borrowed from [14]:

$$\widehat{p}_N^{(ss)} = (\Psi_N^{(ss)T}\Theta_N)^{-1}\Psi_N^{(ss)T}Y_N \qquad (10)$$

with $\Psi_N^{(ss)} = (\psi_1^{(ss)}, \psi_2^{(ss)}, ..., \psi_N^{(ss)})^T$ and the instruments $\psi_k^{(ss)} = (u_{k-1}, u_k, u_k^2, u_k^3)^T$. In turn, in the nonparametric stage of our two-stage procedure we set $N_0 = 10$ and, for comparison purposes, assumed $M = N$ (with N the same as for (10)) in the kernel estimate of w_k's (see (6), (9)) taking due to [5] $h_{opt}(M) = 4.1M^{-0.2}$. In the parametric stage (stage 2) we used linear least squares, getting the estimate

$$\widehat{c}_{N_0,N} = \left(\Phi_{N_0}^T\Phi_{N_0}\right)^{-1}\Phi_{N_0}^T\widehat{W}_{N_0,N}$$

where $\Phi_{N_0} = (\phi_1, \phi_2, ..., \phi_{N_0})^T$, $\phi_k = (u_k, u_k^2, u_k^3)^T$ and $\widehat{W}_{N_0,N} = (\widehat{w}_{1,N}, \widehat{w}_{2,N}, ..., \widehat{w}_{N_0,N})^T$ with the computed estimates $\widehat{w}_{k,N}$. Next we compared appropriate estimation errors of the nonlinearity parameters produced by both methods

$$\Delta_c = \frac{\|\widehat{c}_{N_0,N} - c^*\|_2}{\|c^*\|_2} \cdot 100\%$$

and

$$\Delta_c^{(ss)} = \frac{\left\|\widehat{p'}_N^{(ss)} - c^*\right\|_2}{\|c^*\|_2} \cdot 100\%$$

where $\widehat{p'}_N^{(ss)} = (\frac{\widehat{p}_{2,N}^{(ss)}}{\alpha_0}, \frac{\widehat{p}_{3,N}^{(ss)}}{\alpha_0}, \frac{\widehat{p}_{4,N}^{(ss)}}{\alpha_0})^T$ for growing data length N. The results are presented in Fig. 2. We see that the combined parametric-nonparametric approach is truly robust to the incorrect knowledge of the parametric description of the linear subsystem, in contrast with the algorithm in [14], where the systematic estimation error (bias) appears.

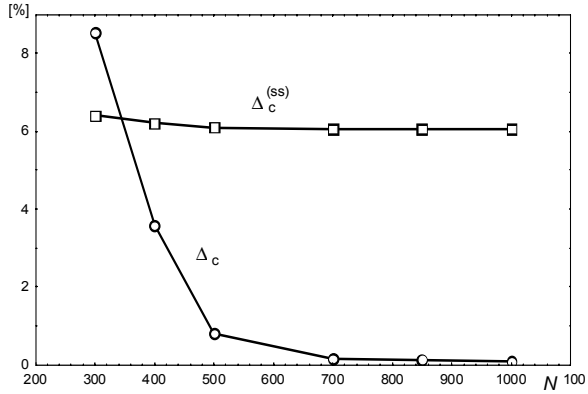

Fig. 2. Comparison between 2 step and S.-S. algorithm.

B. Estimation of non-linear in parameters static characteristic μ

The linear in parameters static characteristic of the Hammerstein system studied in Section 4.1 has been replaced by

$$\mu(u,c^*) = e^{c_1^* u} + \sin c_2^* u \text{ with } c_1^* = c_2^* = 1$$

i.e. by the function $\mu(u) = e^u + \sin u$. Since $\dim c^* = 2$, in Stage 1 of our procedure we can choose $N_0 = 2$ and select any $u_1, u_2 \neq 0$ $(u_1 \neq u_2)$ as the estimation points for computing the nonparametric pointwise estimates of $\mu(u_1,c^*)$ and $\mu(u_2,c^*)$. We set $u_1 = 1$, $u_2 = 2$ and used the kernel estimate, described in Example 1,

with Gaussian kernel. In turn, in stage 2, Levenberg-Marquardt methodology (sketched out in Section 3) has been applied for minimization of appropriate loss function (cf. (7))

$$\widehat{Q}_{N_0,M}(c) = \frac{1}{2}\sum_{n=1}^{2}[\widehat{w}_{n,M} - e^{c_1 u_n} - \sin c_2 u_n]^2 = \frac{1}{2}\sum_{n=1}^{2} r_n^2$$

being the empirical version of the index (5). The Jacobian matrix has the form

$$J(c) = \begin{bmatrix} \frac{\partial r_1}{\partial c_1} & \frac{\partial r_1}{\partial c_2} \\ \frac{\partial r_2}{\partial c_1} & \frac{\partial r_2}{\partial c2} \end{bmatrix} = \begin{bmatrix} e^{c_1} & \sin c_2 \\ 2e^{2c_1} & 2\sin 2c_2 \end{bmatrix}$$

For $\lambda_0 = 1/1024$ and $\nu = 8$ we obtained the estimation errors $\Delta_c(M) = \frac{\|\widehat{c}_{N_0,M} - c^*\|_2}{\|c^*\|_2} \cdot 100\%$ shown in Fig. 3 illustrating the dependence of the estimation accuracy on the number of data M used for nonparametric estimation in stage 1. Various plots correspond to various intensity δ of the noise $\varepsilon_k \sim U[-\delta, \delta]$.

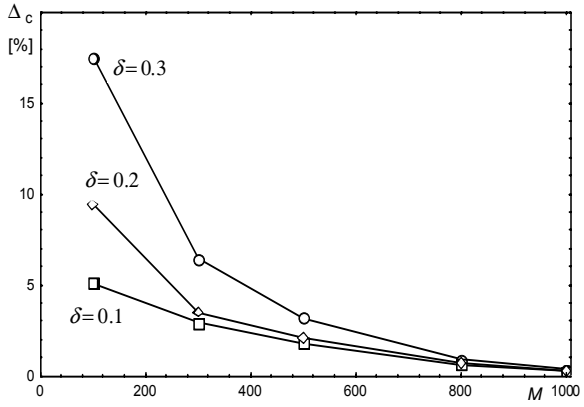

Fig. 3. Estimation error with the use of Levenberg-Marquardt or genetic algorithm.

The optimization in stage 2 has been also performed with use of the typical genetic method and provided the same results as in Fig. 3.

V. CONCLUSIONS

The approach proposed in the paper is distinguished by: 1) small a priori knowledge of the random signals needed for the identification scheme to work and generality, as any particular model of the noise is not assumed, 2) broad applicability, as vast class of non-linear characteristics (not necessarily linear in parameters) is admitted, 3) complete decomposition of the Hammerstein system identification task, and extraction of the nonlinearity recovering as identification of the static subsystem is performed independently of the system dynamics, 4) robustness to the partial inaccuracy of a priori knowledge, as successful identification of the static system component can be performed in spite of imprecise parametric information of the linear dynamics, 5) computational simplicity, as identification is performed by using standard identification routines well elaborated software, and nonparametric stage needs only elementary computations.

In the paper, general conditions are provided for obtaining a successful symbiosis of parametric and nonparametric methods, i.e. of getting hybrid parametric-nonparametric identification algorithms which guarantee achievement of consistent and effective estimates at no great expense.

References

[1] E. W. Bai, D. Li, "Convergence of the iterative Hammerstein system identification algorithm", IEEE Transactions on Automatic Control, vol. 49, pp. 1929-1940, 2004.

[2] S. A. Billings, S. Y. Fakhouri, "Identification of systems containing linear dynamic and static nonlinear elements", Automatica, vol. 18, No. 1, pp. 15-26, 1982.

[3] G. B. Giannakis, E. Serpedin, "A bibliography on nonlinear system identification", Signal Processing, vol. 81, pp. 533-580, 2001.

[4] W. Greblicki, "Nonparametric orthogonal series identification of Hammerstein systems", International Journal of Systems Science, vol. 20, No. 12, pp. 2355-2367, 1989.

[5] W. Greblicki, M. Pawlak, "Cascade non-linear system identification by a non-parametric method", International Journal of System Science, vol. 25, pp. 129-153, 1994.

[6] Z. Hasiewicz, G. Mzyk, "Combined parametric-nonparametric identification of Hammerstein systems", IEEE Transactions on Automatic Control, vol. 49, No. 8, pp. 1370-1375, August 2004.

[7] Z. Hasiewicz, G. Mzyk, "Hammerstein system identification by nonparametric instrumental variables", submitted to International Journal of Control.

[8] W. Härdle, Applied Nonparametric Regression, Cambridge University Press, Cambridge, 1990.

[9] A. Krzyżak, J. Sąsiadek, B. Kégl, "Identification of dynamic nonlinear systems using the Hermite series approach", International Journal of Systems Science, vol. 32, No. 10, pp. 1261-1285, 2001.

[10] J. J. Moré, The Levenberg-Marquardt Algorithm: Implementation and Theory, ser. Lecture Notes in Mathematics, New York: Springer-Verlag, pp. 105-116, 1977.

[11] W. K. Newey, J. L. Powell, "Instrumental variable estimation of nonparametric models", Econometrica, vol. 71, No. 5, pp. 1565-1578, 2003.

[12] M. Pawlak, Z. Hasiewicz, "Nonlinear system identification by the Haar multiresolution analysis", IEEE Transactions on Circuits and Systems, vol. 45, No. 9, pp. 945-961, 1998.

[13] W. H. Press, S. A. Teukolsky, W. Vetterling, B. Flannery, Numerical Recipes in C, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1992.

[14] T. Söderström, P. Stoica, "Instrumental-variable methods for identification of Hammerstein system", International Journal of Control, vol. 35, No. 3, pp. 459-476, 1982.

[15] M. P. Wand, H. C. Jones, Kernel Smoothing, Chapman and Hall, London, 1995.