Multi–level Identification of Hammerstein–Wiener (N–L–N) System in Active Experiment

Marcin Biegański
Department of Control Systems and Mechatronics, Faculty of Electronics, Wrocaw University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

Keywords: Hammerstein–Wiener System, Nonlinear System Identification, Nonparametric Identification.

Abstract: The paper addresses the problem of Hammerstein–Wiener (N–L–N) system identification. The proposed strategy embraces two-experiment approach to the system identification, in which system is excited with random process in passive experiment and with binary process in active experiment. The proposed approach uses both parametric (least squares) and nonparametric (kernel estimates) identification tools. It consists of four consecutive stages, where linear dynamic and nonlinear static parts of the system are identified separately. Output nonlinearity estimation is executed under active experiment. The consistency of the estimate is analyzed and simple simulation example is presented.

1 INTRODUCTION

Hammerstein and Wiener systems are the most popular structures in a class of so-called block-oriented nonlinear systems (BONL) (Giri and Bai, 2010) – the systems that consist of the interaction of linear time-invariant dynamic subsystems and static nonlinear elements. The scope of applications of these models is relatively wide and expands on signal processing (Hasiewicz et al., 2005), biocybernetics, automatic control (Giannakis and Serpedin, 2001), medicine, artificial neural networks (Rubio and Yu, 2007), physical, biological, chemical processes (Gómez and Jutan, 2003) and so on. Despite the fact, that Hammerstein and Wiener systems have high flexibility and provide remarkable ability to capture large class of complex and nonlinear systems, there are still processes that need more complex structures with higher modelling capabilities. And that is why scientists started working on a series combination of Hammerstein and Wiener models (Bai, 1998), (Sjöberg et al., 2012), (Wills et al., 2013).

The paper addresses the problem of SISO Hammerstein–Wiener system identification, that is, identification of object being a cascade connection of two nonlinear static characteristics sandwiched by dynamic linear block (Figure 1). Hammerstein–Wiener system is more convenient, when both actuator and sensor nonlinearities are present, but it has been also successfully applied to modelling several physical processes, such as polymerase reactors (Lee et al., 2004), pH processes (Kalafatis et al., 2005), magnetospheric dynamics, among many others. Unfortunately, despite such extensive interest, the problem of Hammerstein–Wiener system identification still remains completely open. The fundamental difficulty in identification is caused by the presence of Wiener part in which dynamic linear block precedes static nonlinearity. Hence, the input of nonlinearity is inaccessible for measurement and correlated. All in all, the state of the art in Wiener and Hammerstein–Wiener systems identification is still not satisfying.

The main goal of the paper is to propose the algorithm for Hammerstein–Wiener system identification that would adapt itself both to Hammerstein and Wiener systems separately without any additional knowledge about the examined system. The proposed procedure is performed in two-experiment approach. In a passive experiment the system is excited and disturbed by ordinary random processes, whereas in an active experiment input signal takes shape of a binary process.

The paper is organized as follows. In Section 2 the identification problem is introduced and formally described. In Section 3 the proposed steps of the algorithm are described and estimates of the separate blocks are presented. In Section 4 the consistency of the output nonlinearity estimate is discussed and proved. Simple simulation example is presented in Section 5, and in Section 6 final remarks and conclusions are given.

Biegański, M.
Multi–level Identification of Hammerstein–Wiener (N–L–N) System in Active Experiment.
DOI: 10.5220/0006863703550361
In *Proceedings of the 15th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2018) - Volume 1*, pages 355-361
ISBN: 978-989-758-321-6

2 PROBLEM STATEMENT

We consider a SISO Hammerstein–Wiener system, i.e. a block-oriented sandwich structure shown in Figure 1, where u_k and y_k are measurable input and output signals at time k, respectively and z_k is a random noise. Signals w_k, x_k and v_k are inaccessible for direct measurement.

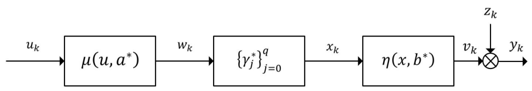

Figure 1: Hammerstein–Wiener system.

Functions $\mu()$ and $\eta()$ denote the unknown input and output nonlinear characteristics, whereas $\{\gamma_j^*\}_{j=0}^q$ are the true and unknown parameters of the finite impulse response of the linear dynamic block. The system is described by the following input-output equation:

$$y_k = \eta\left(\sum_{j=0}^{q}\left(\gamma_j^*\mu(u_{k-j})\right)\right) + z_k. \tag{1}$$

The output of the whole system is a sum of output of second nonlinearity and additive noise.

Regarding the system we assume that:

A. Input and output characteristics of the static blocks are described by the linear combinations of known base functions f and g:

$$\mu(u) = \mu(u, a^*) = a^{*T} f(u), \tag{2}$$
$$a^* = (a_1^*, a_2^*, \ldots, a_m^*)^T, a^* \in \mathbb{R}^m,$$
$$f(u) = (f_1(u), f_2(u), \ldots, f_m(u))^T,$$
$$\eta(x) = \eta(x, b^*) = b^{*T} g(x), \tag{3}$$
$$b^* = (b_1^*, b_2^*, \ldots, b_n^*)^T, b^* \in \mathbb{R}^n,$$
$$g(x) = (g_1(x), g_2(x), \ldots, g_n(x))^T.$$

Dimensions of the parameters vectors a^* and b^* are fixed and known. Moreover it is assumed that static nonlinear characteristics are both Lipschitz functions, i.e. are uniformly continuous with bounded first derivatives. Characteristics are two times differentiable in arbitrarily small neighbourhoods of some points u_0 and $x_0 = \mu(u_0)\sum_{j=0}^q \gamma_j^*$ and $\mu'(u_0) \neq 0$, $\eta'(x_0) \neq 0$. For ease of presentation let us assume $u_0 = 0$, though the method can be generalized for $u_0 \neq 0$. Additionally, output characteristic is strictly monotonous.

B. The dynamic subsystem has the impulse response $\{\gamma_j^*\}_{j=0}^q$, where q – the length of the system memory – is assumed to be finite and known:

$$\gamma^* = \left(\gamma_0^*, \gamma_1^*, \ldots, \gamma_q^*\right), \gamma^* \in \mathbb{R}^{q+1}.$$

The above assumptions state that *a priori* knowledge about the system is purely parametric, as the system is described by $n+m+q+1$ parameters. Our further assumptions concern input and noise signals and are given below:

C. System is excited in two separate experiments, with two different types of signals. The input $u_k^{(1)}$ is an *i.i.d.* random process with Lipschitz probability density function $\nu(u)$ and $\nu(0) \neq 0$. The input $u_k^{(2)}$ is a random binary process.

D. The signals u_k and z_k are mutually independent and have finite variances $\sigma_u^2 < \infty$ and $\sigma_z^2 < \infty$, respectively. Furthermore $E[z_k] = 0$.

Moreover, input characteristic $\mu(u_k^{(2)})$ is presumed as follows: $\mu(0) = 0$, $\mu(1) \neq 0$. The steady-state gain of the linear dynamic block is not identifiable regardless of the identification method, since the internal signals w_k and x_k cannot be measured. Hence, for clarity of presentation and without any loss of generality we assume $G^* = \sum_{j=0}^q \gamma_j^* = 1$ and $\mu(1) = 1$.

The aim is to estimate the unknown characteristic of the output nonlinearity $\eta()$ only on a basis of the input–output measurements of the whole Hammerstein–Wiener system $\{u_k, y_k\}_{k=1}^N$.

3 IDENTIFICATION ALGORITHM

Similarly to (Mzyk and Wachel, 2017) and (Mzyk et al., 2017), the identification algorithm estimates linear and nonlinear parts of the Hammerstein–Wiener system separately. Hence the identification procedure is divided into four stages:

- direct identification of the impulse response parameters γ^* of linear dynamic block in the presence of random input and random noise with the use of least squares method censored by the box kernel selector (Section 3.1),
- recovery of parameters b^* of output nonlinearity in active experiment by kernel-based estimate on the grid of deterministic points determined by the binary excitation (Section 3.2),
- output process filtration in order to generate additional signal r_k with exactly the same expected value as non-measureable signal x_k (Section 3.3),
- identification of parameters a^* of input nonlinearity analogously to the simpler Hammerstein system identification method (Hasiewicz and Mzyk, 2004) (Section 3.4).

Every single stage of the identification algorithm must be executed in a specific order as further stages benefit from the results of former ones.

As system parameters are extracted only from the measurement data $\{(u_k, y_k)\}_{k=1}^{N}$, the purpose of the algorithm is to minimize the following mean squared criterion $Q\left(\hat{\gamma}, \hat{a}, \hat{b}\right) = E(y_k - \hat{y}_k)^2 \rightarrow \min_{\hat{\gamma}, \hat{a}, \hat{b}}$, where $y_k = y_k(\gamma^*, a^*, b^*)$ is the system output and $\hat{y}_k = \hat{y}_k(\hat{\gamma}, \hat{a}, \hat{b})$ is the model output dependent on estimated parameters $\hat{\gamma}$, $\hat{a}$ and $\hat{b}$.

3.1 Identification of Impulse Response

Consider the regression vector that consists of $q+1$ consecutive inputs of a system excited with random process

$$\phi_k = \left(u_k^{(1)}, u_{k-1}^{(1)}, \ldots, u_{k-q}^{(1)}\right)^T.$$

Assuming linear local behaviour of the system around point $u_0 = 0$ and using the following form of the box kernel selector:

$$K(v) = \begin{cases} 1, & \text{if } |v| \leq 1 \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

we propose the following least squares based estimate of the impulse response parameters γ^* (cf. (Bai, 2010), (Mzyk et al., 2017)):

$$\hat{\gamma} = \left(\sum_{k=1}^{N} \phi_k \phi_k^T K\left(\frac{\Delta_k}{h}\right)\right)^{-1} \left(\sum_{k=1}^{N} \phi_k y_k K\left(\frac{\Delta_k}{h}\right)\right), \tag{5}$$

where Δ_k is the infinity norm of the regression vector:

$$\Delta_k = \|\phi_k\|_\infty = \max_{j=0,1,\ldots,q} \left|u_{k-j}^{(1)}\right|.$$

We assume persistent excitation of the input process $u_k^{(1)}$ for the matrix $\sum_{k=1}^{N} \phi_k \phi_k^T K\left(\frac{\Delta_k}{h}\right)$ to be invertible (see (Söderström and Stoica, 1988)). For selecting bandwidth parameter h there exist dedicated methods such as cross-validation method (Wand and Jones, 1994), that establish a good trade-off between the bias and the variance of the estimate.

3.2 Estimation of Output Nonlinearity

Second stage of the algorithm is performed under active experiment, where system is excited with binary random process $u_k^{(2)}$. This active experiment blinds the input nonlinearity so the signal w_k also takes shape of binary process.

Let us introduce the binary representation of all possibilities of the regression vector ϕ_k:

$$\varphi_1 = (1, 0, \ldots, 0)^T,$$
$$\varphi_2 = (0, 1, \ldots, 0)^T,$$
$$\vdots$$
$$\varphi_{N_0} = (1, 1, \ldots, 1)^T.$$

There are N_0 such vectors, where $N_0 = 2^{q+1}$. With this representation and the knowledge about parameters γ^* of the impulse response of the linear dynamic block we are able to form the grid of deterministic points

$$x_{[i]} = \varphi_i^T \gamma^*, i = 1, 2, \ldots, N_0 \tag{6}$$

that represent all possible realizations of non-measurable signal x_k. In these points we can estimate output nonlinearity with the proposed kernel-based estimate:

$$\hat{\eta}\left(x_{[i]}\right) = \frac{\sum_{k=1}^{N} y_k \delta(\phi_k, \varphi_i)}{\sum_{k=1}^{N} \delta(\phi_k, \varphi_i)}, \tag{7}$$

where

$$\delta(\phi_k, \varphi_i) = \begin{cases} 1, & \text{if } \phi_k = \varphi_i \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

is the kernel-like selector (the regression vector must exactly match one of the vectors φ_i) and with $\frac{0}{0}$ understood as 0. Denominator in the proposed estimator averages measurements in local clusters. In general denominator may be equal to 0, so with the introduction of denotement of all output measurements, that have been selected by the kernel technique for given estimation point $x_{[i]}$: $y_{(1)}, y_{(2)}, \ldots, y_{(L)}$ where L is a random number of measurements, the estimate takes the following form:

$$S\left(x_{[i]}\right) = \{y_k : \phi_k = \varphi_i\}, \tag{9}$$

$$\hat{\eta}\left(x_{[i]}\right) = Avg\left(S\left(x_{[i]}\right)\right) = \begin{cases} \frac{1}{L}\sum_{l=1}^{L} y_{(l)}, & \text{if } L > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

Probability of the perfect match $(\phi_k = \varphi_i)$ is constant and equals

$$P\{\delta(\phi_k, \varphi_i) = 1\} = \frac{1}{N_0} = \frac{1}{2^{q+1}}.$$

The result of this step is given by the set of N_0 pairs

$$\left\{\left(x_{[i]}, \hat{\eta}\left(x_{[i]}\right)\right)\right\}_{i=1}^{N_0}. \tag{11}$$

Using this set of pairs, we can easily find the best fitting parameters by the least squares method:

$$\hat{b} = \left(\Psi^T \Psi\right)^{-1} \Psi^T \zeta, \tag{12}$$

where Ψ and ζ are respectively (for $g(x)$ see (3) in Section 2) :

$$\Psi = \left(g(x_{[1]}), g(x_{[2]}), \ldots, g(x_{[N_0]})\right),$$

$$\zeta = \left(\hat{\eta}\left(x_{[1]}\right), \hat{\eta}\left(x_{[2]}\right), \ldots, \hat{\eta}\left(x_{[N_0]}\right)\right).$$

Invertibility of the matrix $\Psi^T\Psi$ (as well as matrix $\Lambda_N^T\Lambda_N$ in (19)) depends on the excitation, input probability density function and the shape of nonlinear base functions (Söderström and Stoica, 1988). Formulation of sufficient general conditions for invertibility still remains open, and only some special cases can be given, at the present state of research.

Additionally there is a byproduct of this stage of the identification procedure. For every single estimation point $x_{[i]}$ we obtain a set of output measurements and its variety depends solely on the presence of the disturbance process z_k. So probability density function of the noise signal can be estimated based on the value of deviation from the average value $\hat{\eta}\left(x_{[i]}\right)$ e.g. with the kernel-based method (with $L > 0$):

$$\hat{f}(z) = \frac{1}{Lh_z}\sum_{l=1}^{L} K\left(\frac{z_{(l)} - z}{h_z}\right) \tag{13}$$

and it can be done N_0 times.

3.3 Output Signal Filtration

After extraction of parameters b^* of output nonlinearity parameters a^* of the first nonlinear block still remain unidentified. If we could identify input nonlinearity first, there would be no problem with identification of the second nonlinear static block. But the other way around, our idea is to estimate signal x_k so we can identify parameters of first nonlinearity in the way it is done with simpler Hammerstein system. Asssuming strict monotonicity of the output nonlinear characteristic we obtain reversible function that can be used to recover x_k process. The proposed approach is shown in Figure 2.

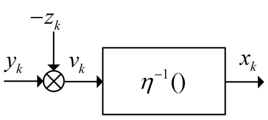

Figure 2: Reverse flow of output nonlinear block.

Estimation is described by the following equations

$$y_k = \eta(x_k) + z_k, \tag{14}$$

$$x_k = \eta^{-1}(y_k - z_k). \tag{15}$$

The problem of non-accessible input noise is known in the literature as the error-in problem (Chen and Zhao, 2014). But with the knowledge of probability density function of the disturbance process $f(z)$ we can form additional function $\varsigma()$ (Figure 3):

$$\varsigma(y) = E\left\{x_k | y_k = y\right\} = \int_{-\infty}^{\infty} \eta^{-1}(y - z) f(z) dz. \tag{16}$$

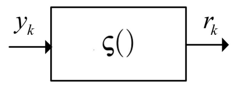

Figure 3: Output signal filtration.

With $\varsigma(y)$ we can generate additional signal r_k with the same expected value as x_k i.e.

$$\begin{aligned} Er_k = E\varsigma(y_k) = E\left[\int_{-\infty}^{\infty} \eta^{-1}(y_k - z) f(z) dz\right] = \\ = E\left[\int_{-\infty}^{\infty} \eta^{-1}\left(\eta(x_k)\right) f(z) dz\right] = \\ = E\left[x_k \cdot \int_{-\infty}^{\infty} f(z) dz\right] = Ex_k. \end{aligned} \tag{17}$$

3.4 Identification of Input Nonlinearity

In the final section, due to multistage approach and with r_k signal generated, identification of the first nonlinear block is reduced to the Hammerstein system identification problem (Hasiewicz and Mzyk, 2004). Signal x_k can be presented as

$$x_k = \lambda_k^T \theta^*, \tag{18}$$

where

$$\lambda_k = \Big(f_1(u_k^{(1)}), \ldots, f_m(u_k^{(1)}), \ldots, f_1(u_{k-q}^{(1)}), \ldots, f_m(u_{k-q}^{(1)})\Big)^T$$

is the regression vector and

$$\theta^* = \left(\gamma_0^* a_1^*, \ldots, \gamma_0^* a_m^*, \ldots, \gamma_q^* a_1^*, \ldots, \gamma_q^* a_m^*\right)$$

is the vector of mixed products. After introducing generic vectors $\Lambda_N = \left(\lambda_1^T, \lambda_2^T, \ldots, \lambda_N^T\right)^T$ and $R_N = \left(r_1, r_2, \ldots, r_N\right)^T$, mixed products estimate $\hat{\theta}$ can be represented with the use of signal r_k:

$$\hat{\theta} = \left(\Lambda_N^T \Lambda_N\right)^{-1} \Lambda_N^T R_N. \tag{19}$$

With the knowledge about impulse response parameters (Section 3.1) we can benefit from singular value decomposition to extract parameters $\hat{a}$ (for details see (Kincaid and Cheney, 1991)).

With parameters $\hat{a}$ extracted, the estimate of the input nonlinearity takes the following shape:

$$\hat{\mu}(u) = \hat{\mu}(u, \hat{a}) = \hat{a}^T f(u), \tag{20}$$

where

$$\hat{a} = \left(\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_m\right)^T.$$

4 STATISTICAL PROPERTIES

Theorem 1. *Let assumptions A. — D. be in force. Then, for the Hammerstein–Wiener system and* $h \sim N^{-\alpha}$, *where* $\alpha \in \left(0, \frac{1}{d}\right)$ *and* $d = q+3$, *it holds that*

$$\hat{\gamma}_j \to \gamma_j^*, \ j = 0,1,\ldots,q \tag{21}$$

in probability as $N \to \infty$, *provided that* $c = \mu'(u_0)\eta'(x_0) \neq 0$.

Sketch of the proof. Under Assumption A. let's take Taylor series expansion and apply it to input and output nonlinearities around points u_0 and x_0:

$$\mu(u_k) = \mu(u_0) + c_1(u_k - u_0) + \rho(u_k), \tag{22}$$

$$\eta(x_k) = \eta(x_0) + c_2(x_k - x_0) + \xi(x_k). \tag{23}$$

Parameters c_1 and c_2 are non-zero constants and are equal to first derivatives of functions $\mu()$ and $\eta()$ in points u_0 and x_0. Observing that $|\rho(u_k)| = o(h)$ and $|\xi(x_k)| = o(h)$, the theorem can be proved analogously to Theorem 3 in (Mzyk and Wachel, 2017). □

Theorem 2. *Under Assumptions A. — D., it holds that*

$$E\left[\eta\left(x_{[i]}\right) - \hat{\eta}\left(x_{[i]}\right)\right]^2 \to 0, \ as\ N \to \infty, \tag{24}$$

in each estimation point $x_{[i]} = \varphi_i^T \gamma^*, i = 1,2,\ldots,N_0$, *such that* $x_{[i]} \in cont(\eta(), g())$, *where* $cont(\eta(), g())$ *is a set of all points of continuity of* $\eta()$ *and* $g()$.

Proof. For each estimation point $x_{[i]}$ we have

$$\begin{aligned}
E\left[\hat{\eta}(x_{[i]})\right] &= E\left[\frac{\sum_{k=1}^N y_k \delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)}\right] = \\
&= E\left[\frac{\sum_{k=1}^N (v_k + z_k)\cdot\delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)}\right] = \\
&= E\left[\frac{\sum_{k=1}^N \eta\left(x_{[i]}\right)\cdot\delta(\phi_k, \varphi_i) + z_k\cdot\delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)}\right] = \\
&= E\left[\eta\left(x_{[i]}\right) + \frac{\sum_{k=1}^N z_k\cdot\delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)}\right] = \\
&= \eta\left(x_{[i]}\right) + E\left[\frac{\sum_{k=1}^N z_k\cdot\delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)}\right].
\end{aligned} \tag{25}$$

Let $y_{(1)}, y_{(2)}, \ldots, y_{(L)}$ be the output measurements selected in (9). The number of measurements L is random, but its expected value tends to infinity with increasing number of measurements:

$$EL = P(\phi_k = \varphi_i)\cdot N = \frac{1}{N_0}\cdot N = \frac{N}{2^{q+1}}, \tag{26}$$

so

$$P(L=0) \to 0, \text{ as } N \to \infty. \tag{27}$$

As a result, using Wald identity, bias of the estimate asymptotically tends to zero:

$$\begin{aligned}
E\left[\frac{\sum_{k=1}^N z_k\cdot\delta(\phi_k, \varphi_i)}{\sum_{k=1}^N \delta(\phi_k, \varphi_i)}\right] &= \frac{E\left[\sum_{k(l)=1}^L z_{k(l)}\right]}{EL} = \\
&= \frac{EL\cdot Ez_1}{EL} = 0.
\end{aligned} \tag{28}$$

The variance of the output measurement depends solely on the variance of noise signal and equals

$$\begin{aligned}
var\left[y_{(l)}\right] &= var\left[\eta\left(x_{[i]}\right) + z_{(l)}\right] = \\
&= var\left[\eta\left(x_{[i]}\right)\right] + var\left[z_{(l)}\right] = \sigma_z^2.
\end{aligned} \tag{29}$$

Assuming that the random number of measurement is greater than zero, the conditional variance of the estimate takes the following form:

$$\begin{aligned}
var\left[\hat{\eta}\left(x_{[i]}\right)\right] &= \sum_{k=1}^N P(L=k)\cdot var\left[\frac{1}{k}\sum_{l=1}^k y_{(l)}\right] = \\
&= \sum_{k=1}^N P(L=k)\cdot\frac{var\left[y_{(l)}\right]}{k} = \\
&= \sum_{k=1}^N P(L=k)\cdot\frac{\sigma_z^2}{k} = \\
&= \frac{N_0\cdot\sigma_z^2}{N} = c\cdot\frac{1}{N} \sim N^{-1}.
\end{aligned} \tag{30}$$

Finally, from (28) and (4), the asymptotical convergence of the estimate is proven (cf. (Mzyk, 2007) or (Mzyk and Wachel, 2017)).

□

5 NUMERICAL EXAMPLE

In this section we describe a simple and intuitive simulation example that illustrates the first two stages of the proposed algorithm. In the experiment, we simulated Hammerstein–Wiener system with nonlinear static characteristics chosen as:

$$\mu(u, a^*) = a_1^* u + a_2^* u^2, \ a^* = (a_1^*, a_2^*)^T = (0.8, 0.2)^T,$$

$$\eta(x, b^*) = b_1^* x + b_2^* x^2, \ b^* = (b_1^*, b_2^*)^T = (0.7, 0.4)^T,$$

and with the following impulse response of the dynamic filter:

$$\gamma^* = (0.6, 0.3, 0.1)^T.$$

The system was excited by two types of random processes. The first one used in the passive experiment was chosen as uniformly distributed random process $u_k^{(1)} \sim \mathcal{U}(-1,1)$. The second one used in the active experiment $\left(u_k^{(2)}\right)$ was a random binary process with equal probabilities of 0 and 1. Moreover, the system was disturbed by random process $z_k \sim \mathcal{U}(-0.5,0.5)$ (50% noise with respect to excitation signal). Bandwidth parameter h was selected with the cross-validation method (see Figure 4) and set to 0.66. To illustrate the asymptotic (i.e. for $N \to \infty$) behaviour of the proposed method, the parameters were recovered based on $N = 10^5$ input–output measurement pairs $\{(u_k, y_k))\}_{k=1}^{N}$.

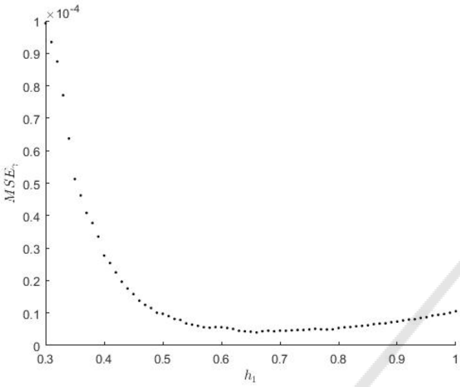

Figure 4: Cross-validation method for bandwidth parameter h.

As a result of the first stage, kernel-censored least squares estimate recovered the following parameters of the finite impulse response (all parameters are rounded to the third decimal place): $\hat{\gamma} = (0.613, 0.302, 0.085)^T$, so the mean square error between real and estimated parameters was equal to $MSE_\gamma = \frac{1}{q+1}\sum_{i=0}^{q}(\gamma_i^* - \hat{\gamma}_i)^2 = 1.38 \cdot 10^{-4}$.

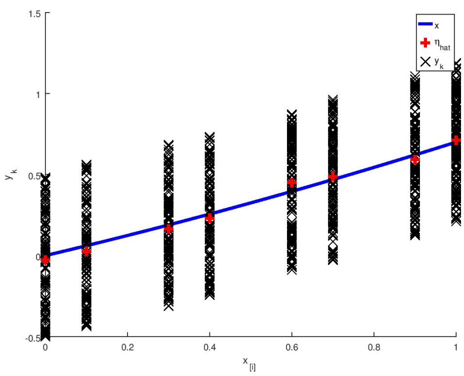

Figure 5: The nonparametric estimates of output nonlinearity (crosses) compared to real characterisic $\eta(x)$ (line). Output measurements presented in the background (X marks).

In the second stage, the output nonlinearity was estimated on the grid of deterministic points $\{x_{[i]}\}_{i=1}^{N_0}$, where $N_0 = 2^{q+1} = 8$. In these points, output nonlinearity was estimated with the use of kernel-based method. In the Figure 5 we can see the output measurements grouped in clusters and results of the estimate compared to the real output nonlinearity. Finally, we recovered the following parameters $\hat{b} = (0.699, 0.401)^T$ with the mean squared error $MSE_b = \frac{1}{2}\sum_{i=1}^{2}\left(b_i^* - \hat{b}_i\right)^2 = 1.1 \cdot 10^{-6}$. Furthermore, as a byproduct of this step we could estimate the probability density function of the noise signal in each estimation point $x_{[i]}$. Results of this estimation are presented in Figure 6.

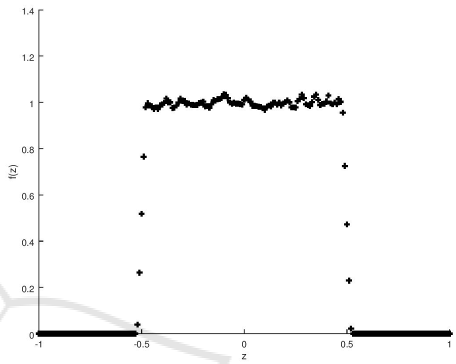

Figure 6: Kernel-based identification of probability density function $f(z)$.

The next step of the algorithm – creation of additional signal r_k as a covolution of inversion of output nonlinearity and probability density function of the disturbance process – can be processed simulatively with the following methods: fast Fourier transform, numerical integration with Riemann sum or Monte Carlo method. Lastly, the input nonlinearity can be recovered as in simpler Hammerstein system.

6 CONCLUSIONS

In the paper, a new method of identification of the the Hammerstein–Wiener (N–L–N) system has been proposed and analyzed under assumptions presented at the beginning of the article. The idea of identification routine consists of both parametric (least squares) and nonparametric (kernel estimate) techniques, and is divided into separate identification of linear dynamic, and nonlinear static parts of the system. Identification of the output nonlinearity is done under active experiment, with random binary excitation. The system is identifiable and the solution is unique for the impulse response fulfilling the given assumptions and for output nonlinearity satisfying Haar condition. Dividing the problem into four separate stages significantly reduces dimensionality of the problem. Effectiveness of the method is strictly dependent on the length of the impulse response which is specific for the whole class of Wiener-type systems, where linear dynamic block is followed by static nonlinearity ("course of dimensionality"). In consequence, proposed strategy is rather recommended for short memory dynamic filters, but the class of admissible nonlinearities is relatively broad. More general cases are remained for further research.

ACKNOWLEDGEMENTS

The work was supported by the National Science Centre, Poland, Grant No. 2016/21/B/ST7/02284.

REFERENCES

Bai, E.-W. (1998). An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems. *Automatica*, 34(3):333–338.

Bai, E.-W. (2010). Non-parametric nonlinear system identification: An asymptotic minimum mean squared error estimator. *IEEE Transactions on Automatic Control*, 55(7):1615–1626.

Chen, H. and Zhao, W. (2014). *Recursive Identification and Parameter Estimation*. CRC Press.

Giannakis, G. B. and Serpedin, E. (2001). A bibliography on nonlinear system identification. *Signal Processing*, 81(3):533 – 580. Special section on Digital Signal Processing for Multimedia.

Giri, F. and Bai, E. (2010). *Block-oriented Nonlinear System Identification*. Springer-Verlag London, London.

Gómez, J. C. and Jutan, A. (2003). Identification and model predictive control of a pH neutralization process based on linear and Wiener models. *IFAC Proceedings Volumes*, 36(16):1507 – 1512. 13th IFAC Symposium on System Identification (SYSID 2003), Rotterdam, The Netherlands, 27-29 August, 2003.

Hasiewicz, Z. and Mzyk, G. (2004). Combined parametric-nonparametric identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, 49(8):1370–1375.

Hasiewicz, Z., Pawlak, M., and Śliwinski, P. (2005). Nonparametric identification of nonlinearities in block-oriented systems by orthogonal wavelets with compact support. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 52(2):427–442.

Kalafatis, A. D., Wang, L., and Cluett, W. R. (2005). Identification of time-varying pH processes using sinusoidal signals. *Automatica*, 41(4):685 – 691.

Kincaid, D. and Cheney, W. (1991). *Numerical Analysis: Mathematics of Scientific Computing*. Brooks/Cole Publishing Co., Pacific Grove, CA, USA.

Lee, Y. J., Sung, S. W., Park, S., and Park, S. (2004). Input test signal design and parameter estimation method for the Hammerstein-Wiener processes. *Industrial & Engineering Chemistry Research*, 43(23):7521–7530.

Mzyk, G. (2007). A censored sample mean approach to nonparametric identification of nonlinearities in Wiener systems. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 54(10):897–901.

Mzyk, G., Biegański, M., and Kozdraś, B. (2017). Multi-stage identification of an N-L-N Hammerstein-Wiener system. In *2017 22nd International Conference on Methods and Models in Automation and Robotics (MMAR)*, pages 343–346.

Mzyk, G. and Wachel, P. (2017). Kernel-based identification of Wiener-Hammerstein system. *Automatica*, 83:275 – 281.

Rubio, J. and Yu, W. (2007). Stability analysis of nonlinear system identification via delayed neural networks. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 54(2):161–165.

Sjöberg, J., Lauwers, L., and Schoukens, J. (2012). Identification of Wiener-Hammerstein models: Two algorithms based on the best split of a linear model applied to the SYSID'09 benchmark problem. *Control Engineering Practice*, 20(11):1119 – 1125. Special Section: Wiener-Hammerstein System Identification Benchmark.

Söderström, T. and Stoica, P., editors (1988). *System Identification*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Wand, M. and Jones, M. (1994). *Kernel Smoothing*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis.

Wills, A., Schön, T. B., Ljung, L., and Ninness, B. (2013). Identification of Hammerstein-Wiener models. *Automatica*, 49(1):70–81.