Nonparametric System Identification Matlab Toolbox

Grzegorz Mzyk[a]
Department of Electronics, Wrocaw University of Science and Technology, W. Wyspiaskiego 27, 50-370 Wrocaw, Poland

Keywords: System Identification, Nonparametric Estimation, Kernel Regression, Orthogonal Expansion, Hammerstein System, Wiener System.

Abstract: In the paper the first version of Nonparametric System Identification Matlab Toolbox is presented. It is based on theoretical results concerning nonparametric identification method, achieved for the last four decades. The library includes both standard (kernel based or orthogonal expansion based) nonparametric methods and recent algorithms including combined (parametric-nonparametric) algorithms. Hammerstein and Wiener models and their serial connections are considered. Nonparametric estimates, usually run as a preliminary steps, play supporting role in the main procedure of estimating system parameters by the least squares method. Multi-level (hybrid) structure of algorithms, i.e. combining both parametric and nonparametric approaches allows to decompose the problem of identification of interconnected complex system into simpler local subproblems. Moreover, asymptotic consistency of all estimates was formally proved, even under existence of random and correlated noise.

1 INTRODUCTION

1.1 History

The need of having accurate models of relationships is of crucial meaning for decision making, system identification, forecasting, designing of optimal control, system identification, pattern recognition, simulation and many others. For ages, people wanted to explain the nature of real relationships to improve efficiency of production and organization, increase the level of safety or to forecast the future and adapt to changing conditions. Formally, the paper by Gauss ((Gauss and Davis, 1857)), from 19th century, which introduces the least squares method is treated as initiation of the field. In general, building models is based on two kind of knowledge:

- parametric, a priori, usually provided by experts or determined by laws of physics, i.e., we are given the formula with finite and known number of unknown parameters,
- nonparameric, i.e., the set of input-output data collected in the experiment (learning sequence).

As we feel intuitively, thanks to parametric knowledge we can significantly narrow the class of potential relationships taken into consideration, and consequently speed up the convergence rapidly. Nevertheless the risk of false parametric assumption cannot be neglected. If the assumed formula is not correct, the non-zero approximation error appears, which cannot be reduced even when the number of measurements tends to infinity.

Traditional approaches assumed linear dynamic models as the simplest (rough) approximation of the real system. If they turned out insufficient, the polynomial or bilinear models were applied. Assuming smoothness of nonlinear characteristics the Volterra kernel expansion approach has been proposed ((Boyd et al., 1984)). Nevertheless, the computational complexity of algorithms was not rewarding, owing to large number of parameters needed to be estimated, particularly for long-memory systems with irregular nonlinearities. Moreover, the theoretical analysis of statistical properties of the parametric estimates is relatively difficult in general case. As an alternative to Volterra representation, the concept of block-oriented models was proposed in 1960's ((Narendra and Gallman, 1966)). The system is modelled by interconnections of simple components of two types – linear dynamics and static nonlinearities. The most popular structures in this class are Hammerstein and Wiener models (see (Pintelon and Schoukens, 2004) and (Giri and Bai, 2010)).

[a] https://orcid.org/0000-0002-1701-5095

Mzyk, G.
Nonparametric System Identification Matlab Toolbox.
DOI: 10.5220/0007922306910698
In *Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2019)*, pages 691-698
ISBN: 978-989-758-380-3

In parallel, the theory of nonparametric regression function estimation was developed ((Nadaraya, 1964), (Stone, 1982), (Cristobal et al., 1987), (Härdle, 1990), (Wand and Jones, 1995), (Efromovich, 1999), (Gyorfi et al., 2002), (Ruppert et al., 2003)). First attempts to nonparametric identification of dynamic systems were made by Greblicki and Pawlak in 1980's (see e.g. (Greblicki and Pawlak, 1986)). The theory was developed towards relaxation of assumption concerning nonlinearities and restrictions imposed on the input process ((Hasiewicz et al., 2005), (Pawlak et al., 2007), (Greblicki and Pawlak, 2008), (Bai, 2010), (Rochdi et al., 2010), (Śliwiński, 2013)). The proposed algorithms recover true characteristics and are free of approximation error. Nevertheless, since they are based on measurements only (neglect risky prior knowledge about parametric representation), the rate of convergence is relatively slower and the obtained results are satisfactory only asymptotically.

Recent approaches to system identification tries to combine both parametric and nonparametric algorithms to inherit advantages of both philosophies, i.e. to achieve accurate estimates for moderate number of measurements and guarantee asymptotic consistency, when the number of data grows large. The idea was introduced in (Hasiewicz and Mzyk, 2004) and continued in (Hasiewicz and Mzyk, 2009), (Greblicki and Mzyk, 2009), (Mzyk, 2014) and (Mzyk and Wachel, 2017). In general, the nonparametric pilot kernel estimate supports least squares method in the sense that it censors the data to allow for decomposition of interconnected system identification task into simple (local) subproblems.

1.2 Paper Organization

The paper starts from recalling standard nonparametric estimates of probability density function (Section 2) and of the regression function (Section 3). Both kernel based and orthogonal expansion methods are reminded. Next, in Section 4, nonparametric algorithms are applied for identification of nonlinear static characteristic in Hammerstein system. Also the cross-correlation method is presented for nonparametric identification of linear dynamic element in Hammerstein system. Finally, the combined parametric-nonparametric method is shown, in which, kernel or orthogonal algorithms recover inaccessible interaction signal for independent modeling of individual blocks by the least squares. In Section 5, Wiener system identification problem is considered. It is relatively more difficult comparing to Hammerstein system, owing to correlated excitation of the nonlinear static component. Firstly, the traditional cross-correlation based method is shown under assumption of Gaussian excitation, and next, more sophisticated algorithms, based on input censoring or derivative estimation are shown. Finally, the multi-level (hybrid) strategies, elaborated for sandwich L-N-L (Wiener-Hammerstein) and N-L-N (Hammerstein-Wiener) systems are presented in sections 6 and 7, respectively. General information about *Nonparametric System Identification Toolbox* can be found in Section 8. Our goal is to provide the ready to use identification tools in accessible form, based on the theoretical results of nonparametric estimates, elaborated in the team over the last three decades.

2 ESTIMATION OF PROBABILITY DENSITY FUNCTION

Let us assume that we are given the sequence of N realizations $\{u_k\}_{k=1}^N$ of random variable u, and we need to recover the probability density function $f(u)$, without any prior assumptions concerning its parametric form.

2.1 Kernel Method

The kernel estimate of the probability density function has the form

$$\widehat{f}(u) = \frac{1}{Nh}\sum_{k=1}^{N} K\left(\frac{u_k - u}{h}\right), \tag{1}$$

where $K()$ is a kernel function, e.g.

$$K(v) = \begin{cases} 1, \text{ as } |v| \leq \frac{1}{2} \\ 0, \text{ otherwise} \end{cases}, \tag{2}$$

which selects measurements from neighbourhood of the point u, and $h = h(N)$ is a bandwidth parameter (radius of selection). It can be show that for $N \to \infty$, in all continuity points u it holds that

$$h(N) \quad \to \quad 0 \Longrightarrow E\widehat{f}(u) \to f(u), \tag{3}$$

$$Nh(N) \quad \to \quad \infty \Longrightarrow \text{var}\widehat{f}(u) \to 0, \tag{4}$$

i.e., the $\widehat{f}(u) \to f(u)$ in the mean squared sense, if both (3) and (4) are fulfilled.

2.2 Orthogonal Expansion Method

Alternatively, assuming that $f(u)$ is square integrable, i.e. $f(u) \subset L_2$, and using any complete set of orthonormal basis functions $\{\varphi_i(u)\}_{i=0}^{\infty}$, it can be represented as follows

$$f(u) = \sum_{i=1}^{\infty} a_i \varphi_i(u), \tag{5}$$

where

$$a_i = f(u) \circ \varphi_i(u) = \int_D f(u)\varphi_i(u)du = E\varphi_i(u). \tag{6}$$

The set D in (6) is specific for the orthonormal basis $\{\varphi_i(u)\}_{i=0}^{\infty}$ used in the identification algorithm. The most popular are trigonometric series, orthogonal polynomials (Laguerre, Lagrange, Hermite), or wavelets. The unknown coefficients a_i's can be recovered from experimental data $\{u_k\}_{k=1}^{N}$, as sample means

$$\widehat{a}_i = \frac{1}{N}\sum_{k=1}^{N} \varphi_i(u_k), \tag{7}$$

and owing to Parseval equality, arbitrary accuracy of the approximate

$$\widehat{f}(u) = \sum_{i=1}^{Q} \widehat{a}_i \varphi_i(u) \tag{8}$$

can be achieved, using appropriately selected scale (cut-off level) Q. In general, asymptotic consistency, i.e. the convergence $\widehat{f}(u) \to f(u)$ in the mean square sense is guaranteed as $Q(N) \to \infty$ and $\frac{Q(N)}{N} \to 0$, as $N \to \infty$.

3 ESTIMATION OF REGRESSION FUNCTION

In this section we consider the problem of nonparametric estimation of nonlinear characteristic $\mu()$ of the static system, with noise-corrupted output

$$y_k = \mu(u_k) + z_k. \tag{9}$$

Assuming that the noise z_k is zero mean, $Ez_k = 0$, has finite variance, var$<\infty$, and is independent of the excitation u_k, it can easily be shown that the input-output regression function is equivalent to characteristic $\mu()$, i.e.

$$R(u) = E\{y_k | u_k = u\} = \mu(u). \tag{10}$$

For all points u such that $f(u) > 0$ one can write

$$R(u) = \frac{g(u)}{f(u)}, \tag{11}$$

where $g(u) = R(u) f(u) = \mu(u) f(u)$.

3.1 Kernel Method

Since the noise-free output $\mu(u_k)$ is not accessible for measurement, assuming continuity of $\mu(u)$ in the point u, the natural idea in nonparametric estimation of $g(u)$ is to use selected measurements y_k's, for which respective inputs u_k's belong to the neighbourhood of u,

$$\widehat{g}(u) = \frac{1}{Nh}\sum_{k=1}^{N} y_k K\left(\frac{u_k - u}{h}\right). \tag{12}$$

It leads to kernel regression function estimate of the form

$$\widehat{R}(u) = \frac{\widehat{g}(u)}{\widehat{f}(u)} = \frac{\sum_{k=1}^{N} y_k K\left(\frac{u_k-u}{h}\right)}{\sum_{k=1}^{N} K\left(\frac{u_k-u}{h}\right)}. \tag{13}$$

The theoretical analysis of the limit properties of (13) and the issue of optimal selection of the bandwidth parameters $h(N)$, e.g. by the cross-validation method, is discussed in (Wand and Jones, 1995).

3.2 Orthogonal Expansion Method

Analogously, the numerator in (11) can be expanded as follows

$$g(u) = \sum_{i=1}^{\infty} b_i \varphi_i(u), \tag{14}$$

where the unknown coefficients b_i's can be estimated as

$$\widehat{b}_i = \frac{1}{N}\sum_{k=1}^{N} y_k \varphi_i(u_k), \tag{15}$$

and $\widehat{a}_i$'s are given by (7). Consequently,

$$\widehat{R}(u) = \frac{\sum_{i=1}^{Q} \widehat{b}_i \varphi_i(u)}{\sum_{i=1}^{Q} \widehat{a}_i \varphi_i(u)}. \tag{16}$$

For details concerning properties of the estimate (16), for various kinds of orthogonal basis, we refer the reader to (Śliwiński, 2013) and references cited therein.

4 HAMMERSTEIN SYSTEM IDENTIFICATION

The Hammerstein system (see Fig. 1) includes static nonlinear element with the characteristic $\mu()$ followed by the linear dynamic filter with the impulse response $\left\{\gamma_j\right\}_{j=0}^{\infty}$. The interaction signal w_k is hidden in the sense that it cannot be measured. The system is de-

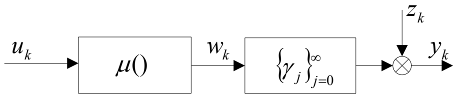

Figure 1: Hammerstein system.

scribed by the following equation

$$y_k = \sum_{j=0}^{\infty} \gamma_j \mu\left(u_{k-j}\right) + z_k. \tag{17}$$

Assuming that the linear dynamics is asymptotically stable, i.e. $\sum_{j=0}^{\infty} \left|\gamma_j\right| < \infty$, the goal is to estimate both characteristic $\mu()$ and the impulse response $\left\{\gamma_j\right\}_{j=0}^{\infty}$ from the input-output data $\{(u_k, y_k)\}_{k=1}^{N}$, collected in the experiment. The crucial meaning has the fact that for i.i.d. input sequence, it holds that

$$R(u) = \gamma_0 \mu(u) + c_1, \tag{18}$$

where $c_1 = E\mu(u_k)\sum_{j=1}^{\infty}\gamma_j = \text{const}$, i.e. standard regression $R(u)$ is Hammerstein system is scaled and shifted version of the nonlinear characteristic of its static component. Assuming that $\mu(0) = 0$, one can avoid the offset c_1, using the corrected nonparametric regression estimate

$$\widehat{\mu}(u) = \widehat{R}(u) - \widehat{R}(0). \tag{19}$$

The observation (18) allows to generalize nonparametric estimates (13) or (16) for dynamic system. It can be proved that under standard conditions concerning $h(N)$ or $Q(N)$ (see e.g. (Greblicki and Pawlak, 2008)), it holds that $\widehat{\mu}(u) \to \gamma_0\mu(u)$. The scale γ_0 is not identifiable independently of the identification method, owing inaccessibility of w_k. The Hammerstein systems $\mu(u) \times \left\{\gamma_j\right\}_{j=0}^{\infty}$ and $\gamma_0\mu(u) \times \left\{\frac{\gamma_j}{\gamma_0}\right\}_{j=0}^{\infty}$ are equivalent from the input-output point of view.

As regards identification of linear block, for i.i.d. input sequence one can apply standard cross-correlation analysis. It can easily be shown that the input-output cross-correlation coefficients

$$\varsigma_j = E\left\{(u_k - Eu_k)\, y_{k+j}\right\} \tag{20}$$

are proportional to the unknown impulse response elements, i.e.

$$\varsigma_j = c_2\gamma_j, \tag{21}$$

where $c_2 = \text{const}$ for all $j = 0, 1, 2....$ It leads to the following estimate

$$\widehat{\gamma}_j = \frac{1}{N-j}\sum_{k=1}^{N-j}(u_k - \overline{u})\, y_{k+j}, \tag{22}$$

where $\overline{u} = \frac{1}{N}\sum_{k=1}^{N} u_k$. Asymptotically, for $N \to \infty$, to assure consistency of the cut model $\left\{\widehat{\gamma}_j\right\}_{j=0}^{S(N)}$ of the stable linear subsystem, the order $S(N)$ should behave such that $S(N) \to \infty$, but $\frac{S(N)}{N} \to 0$.

Despite pure nonparametric estimates (19) and (22) guarantee asymptotic convergence to the true characteristic, the convergence rate is rather slow, and the results can be not satisfying for moderate number of measurements. Hence, the combined parametric-nonparametric algorithms proposed firstly in (Hasiewicz and Mzyk, 2004) are worth notifying. They allow to decompose complex system identification task into simpler local subproblems, and can be applied under partial or uncertain knowledge of individual components. In step 1 (nonparametric) we identify inaccessible interaction signal

$$\widehat{w}_k = \widehat{R}(u_k) - \widehat{R}(0) \tag{23}$$

with the use of kernel orthogonal regression estimators, and next, in step 2 (parametric) we incorporate least squares or instrumental variables approach for both static and dynamic subsystems using the pairs $\{(u_k, \widehat{w}_k)\}_{k=1}^{N}$, and $\{(\widehat{w}_k, y_k)\}_{k=1}^{N}$, respectively. For example, under parametric knowledge that $\mu(u) = \mu(u, \theta^*)$, the true vector of parameters, θ^*, is estimated as follows

$$\widehat{\theta} = \arg\min_{\theta}\sum_{k=1}^{N}\left(\widehat{w}_k - \mu(u_k, \theta^*)\right)^2. \tag{24}$$

Let us emphasize, that nonparametric estimate $\widehat{w}_k$ is plugged in to the definition of parameter estimate $\widehat{\theta}$. The formal proofs of consistency of $\widehat{\theta}$ and parametric-nonparametric estimates for IIR linear dynamics can be found in (Hasiewicz and Mzyk, 2009).

5 WIENER SYSTEM IDENTIFICATION

The Wiener system (Fig. 2) includes the components of Hammerstein system connected in reverse order. It

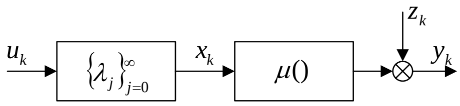

Figure 2: Wiener system.

is described by the equation

$$y_k = \mu\left(\sum_{j=0}^{\infty}\lambda_j u_{k-j}\right) + z_k. \tag{25}$$

The Wiener structure has very wide scope of potential applications (see (Giannakis and Serpedin, 2001)). Unfortunately, since the hidden nonlinearity input,

x_k, is correlated, the problem is much more difficult. In the contrary to Hammerstein system identification task, sufficient identifiability conditions for Wiener system were formulated in the literature only for some special cases. One of them is based on application of Gaussian excitation. In this specific situation, also the hidden process $\{x_k\}$ is normally distributed, and the Bussgang theorem holds. It allows to identify impulse response elements λ_j analogously to (22), i.e., $\widehat{\lambda}_j = \frac{1}{N-j}\sum_{k=1}^{N-j}(u_k - \overline{u})y_{k+j}$. Using the FIR approximate model of the linear dynamic block, the interaction signal x_k can be approximated as follows $\widehat{x}_k = \sum_{j=0}^{S}\widehat{\lambda}_j u_{k-j}$, and the characteristic of the nonlinear component can be estimated from the pairs $\{(\widehat{x}_k, y_k)\}_{k=1}^{N}$, analogously to (24), i.e. $\widehat{\theta} = \arg\min_{\theta}\sum_{k=1}^{N}(y_k - \mu(\widehat{x}_k, \theta^*))^2$. The input density restriction has been relaxed in (Mzyk, 2007), (Greblicki, 2010), (Pawlak et al., 2007) and (Wachel and Mzyk, 2016). For the survey of parametric and nonparametric methods for identification of Wiener system we refer the reader to (Mzyk, 2010).

6 WIENER-HAMMERSTEIN SYSTEM IDENTIFICATION

Although Hammerstein and Wiener models generalize the class of linear systems, they are still not sufficient in some practical applications. In this section we consider cascade connection of Wiener and Hammerstein system, with the L-N-L (sandwich) structure (see Fig. 3). The system is described by the equation

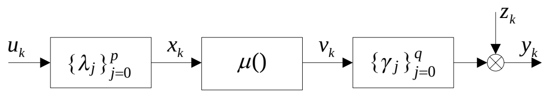

Figure 3: Wiener–Hammerstein system.

$$y_k = \sum_{j=0}^{q}\gamma_j\mu\left(x_{k-j}\right) + z_k, \tag{26}$$

where $x_k = \sum_{i=0}^{p}\lambda_i u_{k-i}$.

First attempts to parametric-nonparametric identification of Wiener-Hammerstein system were made in (Mzyk, 2012) and the proposed estimates were further analyzed in (Mzyk and Wachel, 2017). The algorithm consists of three steps.

Step 1. Nonparametric kernel identification on the nonlinear characteristic

$$\widehat{\mu}_N(x) = \frac{\sum_{k=1}^{N} y_k \cdot K\left(\frac{\delta_k(x)}{h}\right)}{\sum_{k=1}^{N} K\left(\frac{\delta_k(x)}{h}\right)}, \tag{27}$$

where

$$\delta_k(x) \triangleq \sum_{j=0}^{p+q}\left|u_{k-j} - x\right|. \tag{28}$$

Step 2. Estimation of the convolution of impulse response of linear dynamic objects

$$\varkappa_j = \lambda_j * \gamma_j = \sum_{i=0}^{j}\gamma_i\lambda_{j-i}, \tag{29}$$

by the local cross-correlation censored by the kernel technique

$$\widehat{\varkappa}_\tau = \frac{1}{N\eta^{p+q+3}}\sum_{k=p+q+1}^{N-(p+q)} u_k y_{k+\tau} K\left(\frac{\Delta_k}{\eta}\right), \tag{30}$$

where η is a bandwidth (analogously to h) and

$$\Delta_k = \max_{j=0,1,\ldots,p+q}\left|u_{k-j}\right|. \tag{31}$$

Step 3. Splitting the polynomial

$$\begin{aligned} W(d) &= \varkappa_{p+q}d^{p+q} + \varkappa_{p+q-1}d^{p+q-1} + \ldots + \varkappa_1 d + \varkappa_0 \\ &= \varkappa_{p+q}(d-d_1)(d-d_2)\ldots(d-d_{p+q}), \end{aligned} \tag{32}$$

where $\Omega = \left\{d_1, d_2, \ldots, d_{p+q}\right\}$ denotes the set of roots (generally complex), into two separate factors

$$W(d) = \varkappa_{p+q}\Lambda_\Theta(d)\cdot\Gamma_\Theta(d), \tag{33}$$

where $\Lambda_\Theta(d) = \prod_{d_i\in\Theta}(d-d_i)$, and $\Gamma_\Theta(d) = \prod_{d_i\in\Omega\backslash\Theta}(d-d_i)$, such that

$$\left\{\widehat{\lambda}, \widehat{\gamma}\right\} = \arg\min_{\Theta\in\Omega}\widehat{Q}(l_\Theta, g_\Theta), \tag{34}$$

where $\widehat{Q}(l_\Theta, g_\Theta) = \frac{1}{N}\sum_{k=1}^{N}\left[y_k - \overline{y}_k(l_\Theta, g_\Theta)\right]^2$ and $\overline{y}_k(l_\Theta, g_\Theta)$ is the model output for impulse responses l_Θ, and g_Θ, respectively.

Since the speed of convergence is sensitive on the orders p and q, the algorithm is rather devoted to FIR Wiener-Hammerstein systems with short memory.

7 HAMMERSTEIN-WIENER SYSTEM IDENTIFICATION

Serial connection of Hammerstein system with Wiener system leads to the N-L-N sandwich structure (see Fig. 4). The system is describes as follows

$$y_k = \eta\left(\sum_{j=0}^{q}\gamma_j\mu\left(u_{k-j}\right)\right) + z_k. \tag{35}$$

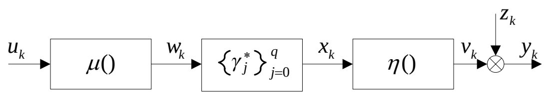

Figure 4: Hammerstein–Wiener system.

Our algorithm (see (Biegański, 2018)) uses both parametric and nonparametric system identification tools to recover parameters of each individual block and it estimates linear and nonlinear parts of the Hammerstein–Wiener system separately. We assume that nonlinear characteristics of the input and output static blocks are described by the linear combinations of *a priori* known base functions f and g

$$\mu(u) = \mu(u, a^*) = a^{*T} f(u), \tag{36}$$

$$a^* = (a_1^*, a_2^*, \ldots, a_m^*)^T, a^* \in \mathbb{R}^m,$$

$$f(u) = (f_1(u), f_2(u), \ldots, f_m(u))^T,$$

$$\eta(x) = \eta(x, b^*) = b^{*T} g(x), \tag{37}$$

$$b^* = (b_1^*, b_2^*, \ldots, b_n^*)^T, b^* \in \mathbb{R}^n,$$

$$g(x) = (g_1(x), g_2(x), \ldots, g_n(x))^T.$$

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 twice differentiable in arbitrarily small neighbourhoods of some points u_0 and $x_0 = \mu(u_0)\sum_{j=0}^{q}\gamma_j^*$ and $\mu'(u_0) \neq 0$, $\eta'(x_0) \neq 0$. Additionally, output characteristic is strictly monotonous, and therefore invertible. Hence the identification procedure is divided into four stages:

Stage 1. Direct identification of the finite impulse response parameters γ^* of linear dynamic subsystem in the presence of random input and random noise with the use of kernel-censored least squares method

$$\hat{\gamma} = \left(\sum_{k=1}^{N} \phi_k \phi_k^T K\left(\frac{\Delta_k}{h}\right) \right)^{-1} \left(\sum_{k=1}^{N} \phi_k y_k K\left(\frac{\Delta_k}{h}\right) \right), \tag{38}$$

where

$$\phi_k = \left(u_k^{(1)}, u_{k-1}^{(1)}, \ldots, u_{k-q}^{(1)}\right)^T, \tag{39}$$

and Δ_k is the infinity norm of the regression vector

$$\Delta_k = \|\phi_k\|_\infty = \max_{j=0,1,\ldots,q} \left|u_{k-j}^{(1)}\right|. \tag{40}$$

Stage 2. Estimation of parameter vector b^* of output nonlinear characteristic in active experiment (binary sequence excitation) with the use of kernel method

$$\hat{\eta}\left(x_{[i]}\right) = \frac{\sum_{k=1}^{N} y_k \delta(\phi_k, \varphi_i)}{\sum_{k=1}^{N} \delta(\phi_k, \varphi_i)}, \tag{41}$$

where

$$\delta(\phi_k, \varphi_i) = \begin{cases} 1, & \text{if } \phi_k = \varphi_i \\ 0, & \text{otherwise} \end{cases}. \tag{42}$$

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

$$\left\{ \left(x_{[i]}, \hat{\eta}\left(x_{[i]}\right)\right) \right\}_{i=1}^{N_0}. \tag{43}$$

Using this set of pairs, we can find the most suitable parameters with the least squares method

$$\hat{b} = \left(\Psi^T \Psi\right)^{-1} \Psi^T \zeta, \tag{44}$$

where Ψ and ζ are respectively

$$\begin{aligned} \Psi &= \left(g(x_{[1]}), g(x_{[2]}), \ldots, g(x_{[N_0]})\right), \\ \zeta &= \left(\hat{\eta}\left(x_{[1]}\right), \hat{\eta}\left(x_{[2]}\right), \ldots, \hat{\eta}\left(x_{[N_0]}\right)\right). \end{aligned}$$

Stage 3. Filtration of output signal y_k in order to generate additional process r_k with the same conditional expected value as non-measureable signal x_k. With

$$\varsigma(y) = E\{x_k | y_k = y\} = \int_{-\infty}^{\infty} \eta^{-1}(y - z) f(z) dz. \tag{45}$$

we can generate additional signal

$$r_k = \varsigma(y_k), \tag{46}$$

with the same conditional expected value as x_k, i.e.,

$$\begin{aligned} R(u) &= E\{r_k | u_k = u\} = E\{\varsigma(y_k) | u_k = u\} = \\ &= E\left\{ \int_{-\infty}^{\infty} \eta^{-1}(y_k - z) f(z) dz | u_k = u \right\} = \\ &= E\left\{ \int_{-\infty}^{\infty} \eta^{-1}(\eta(x_k)) f(z) dz | u_k = u \right\} = \\ &= E\left\{ x_k \cdot \int_{-\infty}^{\infty} f(z) dz | u_k = u \right\} = E\{x_k | u_k = u\}. \end{aligned}$$

Stage 4. Identification of input nonlinear characteristic

$$\widehat{R}(u) = \frac{\sum_{k=1}^{N} r_k K\left(\frac{u_k - u}{h(N)}\right)}{\sum_{k=1}^{N} K\left(\frac{u_k - u}{h(N)}\right)}. \tag{47}$$

The idea was to develop a procedure that would adapt itself to separate block-oriented structures, such as Hammerstein and Wiener systems, even without any additional *a priori* knowledge about the examined system. The problem of identification of such complicated structure is rather difficult, not only because of existence of Wiener part in which linear dynamics precedes static nonlinearity, but also because of the correlation between non-measurable signals. In the algorithm we benefit from multistage and two-experiment approaches to achieve specific, advantageous conditions in which linear and nonlinear parts of the system were less complicated to identify. Additionally, further steps of the algorithm profit from the former ones which significantly reduces the complexity and dimensionality of the identification problem. The uniqueness of the solution is strictly related to the impulse response fulfilling the given assumptions and for the output nonlinear characteristic satisfying Haar condition. The main drawback of the algorithm is that the effectiveness of the procedure is dependent on the length of the finite impulse response, which is characteristic for the whole class of Wiener-type systems ("course of dimensionality"). Therefore the proposed method is recommended for linear dynamic blocks with short impulse response and more sophisticated memoryless nonlinear characteristics.

Example 1. *Let's investigate the simple example of compensator building under knowledge of* $\eta(x)$ *and* $f(z)$. *Assume that nonlinear output block is described by* $\eta(x) = \sqrt[3]{x}$ *and the system is disturbed by additive random, uniformly distributed noise* $z_k \sim \mathcal{U}[-1,1]$. *Compensator can be determined as follows*

$$\xi(y) = E\{x_k|y_k = y\} = \int_{-1}^{1} (y-z)^3 \cdot \frac{1}{2} dz = y^3 + y,$$

i.e., $r_k = y_k^3 + y_k$. *Below, in Fig. 5, we present the estimate (47) of Hammerstein system nonlinearity*

$$\mu(u) = \begin{cases} u & , |u| \le 1 \\ sgn(u) & , |u| > 1 \end{cases}$$

for $\gamma = (\gamma_0, \gamma_1)^T = (1,1)^T$. *The result illustrates applicability of the proposed method.*

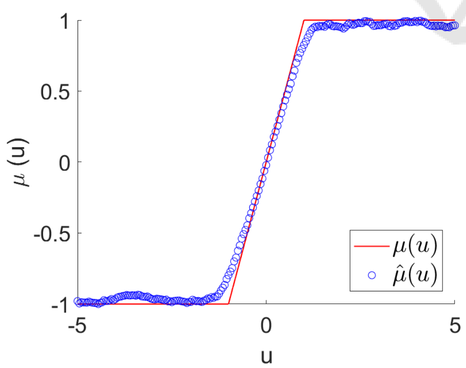

Figure 5: True characteristic $\mu(u)$ vs. the estimate $\hat{\mu}(u)$.

8 THE MATLAB TOOLBOX

The actual version of toolbox and its documentation can be accessed at the WWW page

http://staff.iiar.pwr.wroc.pl/grzegorz.mzyk/KIT

Below we present the list of names of selected functions:

cosineKernel() – returns value of the cosine kernel function

epanechnikovKernel() – returns value of the Epanechnikov kernel

gaussianKernel() – returns value of Gaussian kernel

triangularKernel() – returns value of triangular kernel

uniformKernel() – returns value of uniform (Parzen) kernel

kernelDensityEstimation() – computes probability density function estimate for a given point, by the kernel method

orthogonalRegressionEstimation() – computes model of the regression function using orthogonal expansion method

kernelRegressionEstimation() – computes model of the regression function using kernel method

hammerstein() – identifies both components of Hammerstein system using input-output data

wiener() – identifies both components of Wiener system using input-output data

trigonometricOrthonormalBasis() – supporting function generating trigonometric orthogonal basis functions

estimateDynamicSubsystem() – identification of linear dynamic block

crossValidation() – selection of optimal bandwidth parameter in kernel methods, or the scale in orthogonal expansion methods

9 SUMMARY

The methods presented in the paper combine the nonparametric and parametric tools. Such a strategy allows to solve various kinds of specific obstacles, which are difficult to be overcome in purely parametric or purely nonparametric approach. In particular, the global identification problem can be decomposed on simpler local problems, the measurement sequence can be pre-filtered in the nonparametric stage, or the rough parametric model can be refined by the nonparametric correction when the number of measurements is large enough. The schemes proposed in the paper can be used elastically and have a lot of degrees of freedom. In most of them we can obtain traditional parametric or nonparametric procedures by simple avoiding of the selected steps of combined algorithms. In this sense, the proposed ideas can be treated as generalizations of classical approaches to system identification.

ACKNOWLEDGEMENTS

The work was partially supported by the Grant No. 0401/0136/18, WUS&T, Poland.

REFERENCES

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

Biegański, M. (2018). Multi-level identification of Hammerstein-Wiener (N-L-N) system in active experiment. In *Proceedings of the 15th International Conference on Informatics in Control, Automation and Robotics, ICINCO, Porto, Portugal*, pages 355–361.

Boyd, S., Chua, L. O., and Desoer, C. (1984). Analytical foundations of Volterra series. *IMA Journal of Mathematical Control and Information*, 1(3):243–282.

Cristobal, J., Roca, P., and Manteiga, W. (1987). A class of linear regression parameter estimators constructed by nonparametric estimation. *The Annals of Statistics*, 15(2):603–609.

Efromovich, S. (1999). *Nonparametric Curve Estimation*. Springer, New York.

Gauss, C. F. and Davis, C. H. (1857). *Theory of the motion of the heavenly bodies moving about the sun in conic sections a translation of Gauss's*. Boston,Little, Brown and company,. https://www.biodiversitylibrary.org/bibliography/19023.

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

Giri, F. and Bai, E. W. (2010). *Block-Oriented Nonlinear System Identification*. Lecture Notes in Control and Information Sciences 404. Springer.

Greblicki, W. (2010). Nonparametric input density-free estimation of the nonlinearity in Wiener systems. *IEEE Transactions on Information Theory*, 56(7):3575–3580.

Greblicki, W. and Mzyk, G. (2009). Semiparametric approach to Hammerstein system identification. In *Proceedings of the 15th IFAC Symposium on System Identification, Saint-Malo, France*, pages 1680–1685.

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

Greblicki, W. and Pawlak, M. (2008). *Nonparametric System Identification*. Cambridge University Press.

Gyorfi, L., Kohler, M., Krzyżak, A., and Walk, H. (2002). *A Distribution-Free Theory of Nonparametric Regression*. Springer Verlag, New York.

Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press, Cambridge.

Hasiewicz, Z. and Mzyk, G. (2004). Combined parametric-nonparametric identification of Hammerstein systems. *IEEE Transactions on Automatic Control*, 48(8):1370–1376.

Hasiewicz, Z. and Mzyk, G. (2009). Hammerstein system identification by non-parametric instrumental variables. *International Journal of Control*, 82(3):440–455.

Hasiewicz, Z., Pawlak, M., and Śliwiński, 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.

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. (2010). Parametric versus nonparametric approach to Wiener systems identification. *Lecture Notes in Control and Information Sciences*, 404:111–125.

Mzyk, G. (2012). Nonparametric recovering of nonlinearity in Wiener-Hammerstein systems. In *Proceedings of the 9th International Conference on Informatics in Control, Automation and Robotics, ICINCO, Rome, Italy*, pages 439–445. IEEE.

Mzyk, G. (2014). *Combined Parametric-Nonparametric Identification of Block-Oriented Systems*. Lecture Notes in Control and Information Sciences 454. Springer.

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

Nadaraya, E. (1964). On estimating regression. *Teoriya Veroyatnostei i ee Primeneniya*, 9(1):157–159.

Narendra, K. and Gallman, P. G. (1966). An iterative method for the identification of nonlinear systems using the Hammerstein model. *IEEE Transactions on Automatic Control*, 11:546–550.

Pawlak, M., Hasiewicz, Z., and Wachel, P. (2007). On nonparametric identification of Wiener systems. *IEEE Transactions on Signal Processing*, 55(2):482–492.

Pintelon, R. and Schoukens, J. (2004). *System Identification: a Frequency Domain Approach*. Wiley-IEEE Press.

Rochdi, Y., Giri, F., Gning, J., and Chaoui, F. (2010). Identification of block-oriented systems in the presence of nonparametric input nonlinearities of switch and backlash types. *Automatica*, 46(5):864–877.

Ruppert, D., Wand, M., and Carroll, R. (2003). *Semiparametric Regression*, volume 12. Cambridge University Press.

Śliwiński, P. (2013). *Nonlinear System Identification by Haar Wavelets*. Lecture Notes in Statistics 210. Springer.

Stone, C. (1982). Optimal global rates of convergence for nonparametric regression. *The Annals of Statistics*, 10(4):1040–1053.

Wachel, P. and Mzyk, G. (2016). Direct identification of the linear block in Wiener system. *International Journal of Adaptive Control and Signal Processing*, 30(1):93–105.

Wand, M. and Jones, H. (1995). *Kernel Smoothing*. Chapman and Hall, London.