Int. J. Appl. Math. Comput. Sci., 2019, Vol. 29, No. 3, 571–580
DOI: 10.2478/amcs-2019-0042

TWO–STAGE INSTRUMENTAL VARIABLES IDENTIFICATION OF POLYNOMIAL WIENER SYSTEMS WITH INVERTIBLE NONLINEARITIES

ANDRZEJ JANCZAK [a,*], JÓZEF KORBICZ [a]

[a]Institute of Control and Computation Engineering
University of Zielona Góra, ul. Szafrana 2, 65-516 Zielona Góra, Poland
e-mail: {A.Janczak,J.Korbicz}@issi.uz.zgora.pl

A new two-stage approach to the identification of polynomial Wiener systems is proposed. It is assumed that the linear dynamic system is described by a transfer function model, the memoryless nonlinear element is invertible and the inverse nonlinear function is a polynomial. Based on these assumptions and by introducing a new extended parametrization, the Wiener model is transformed into a linear-in-parameters form. In Stage I, parameters of the transformed Wiener model are estimated using the least squares (LS) and instrumental variables (IV) methods. Although the obtained parameter estimates are consistent, the number of parameters of the transformed Wiener model is much greater than that of the original one. Moreover, there is no unique relationship between parameters of the inverse nonlinear function and those of the transformed Wiener model. In Stage II, based on the assumption that the linear dynamic model is already known, parameters of the inverse nonlinear function are estimated uniquely using the IV method. In this way, not only is the parameter redundancy removed but also the parameter estimation accuracy is increased. A numerical example is included to demonstrate the practical effectiveness of the proposed approach.

Keywords: nonlinear systems, parameter estimation, dynamic models, polynomial models.

1. Introduction

The Wiener model is a block-oriented nonlinear system, in which a linear time-invariant dynamic subsystem is connected in series with a nonlinear memoryless block. A large class of nonlinear systems can be approximated by Wiener models with arbitrarily high accuracy, and Wiener models are widely used in different areas of science and engineering. This stems from the fact that, as was shown by Boyd and Chua (1985), any time-invariant continuous nonlinear operator can be approximated by a Volterra series operator and this approximating operator can be realized as a finite-dimensional linear dynamic system with a nonlinear readout map, i.e., a Wiener-like model.

It is well known that many industrial processes are inherently nonlinear and when the operating point changes it is difficult to represent adequately a given process by means of a linear model. Therefore, to achieve the required system performance, advanced control methods based on nonlinear process models are applied. Wiener models are very useful to represent many real nonlinear systems, such as the pH neutralization process in continuous flow reactors, distillation columns, pneumatic valves, continuous-time and discrete-time chaotic systems (Janczak, 2005). For example, Ipanaqu and Manrique (2011) applied the nonlinear model predictive control based on a piecewise linear Wiener model to an experimental control of pH. A design procedure based on nonlinear internal model control with application to a pH neutralization case study was presented by Kwang *et al.* (2012). The application of Wiener models for the model predictive control was also studied by Ławryńczuk (2013; 2015; 2016). A robust model predictive control method for a distillation column using its Wiener-like model was proposed by Figueroa *et al.* (2013). Using Wiener model in a feedforward control scheme for a simulated continuous stirred tank reactor (CSTR) was demonstrated by Rollins *et al.* (2016).

Wiener system identification has been studied for several decades (see Schoukens and Tiels, 2017). A large number of papers have been published on the

*Corresponding author

identification of Wiener systems, and many different identification methods have been developed that are based on correlation analysis (e.g., Billings and Fakhouri, 1987; 1982; Van Vaerenbergh *et al.*, 2013), frequency analysis (Giri *et al.*, 2014; Brouri and Slassi, 2015), nonlinear optimization (Wigren, 1993; Al-Duwaish *et al.*, 1996; Janczak, 2005; Vörös, 2007; 2015; Ławryńczuk, 2013; Zhou *et al.*, 2013), linear regression (Janczak, 2005; 2007; 2018; Stanisławski *et al.*, 2014), nonparametric regression (Greblicki, 1997; 2001), and subspace approach (Westwick and Verhaegen, 1996; Gómez and Baeyens, 2002, 2005; Ase and Katayama, 2015).

Polynomials (Westwick and Verhaegen, 1996; Janczak, 2005; Stanisławski *et al.*, 2014; Tiels and Schoukens, 2014; Ding *et al.*, 2015; Jansson and Medvedev, 2015; Xiong *et al.*, 2015; Mahataa *et al.*, 2016; Bottegal *et al.*, 2017; Kazemi and Arefi, 2017; Schoukens and Tiels, 2017; Janczak, 2018), Legendre polynomials (Ase and Katayama, 2015), piecewise-linear functions (Dong *et al.*, 2009; Fan and Lo, 2009), cubic splines (Aljamaan *et al.*, 2016), least-squares support vector machine models (Ławryńczuk, 2016), sets of basis functions (Gómez and Baeyens, 2002; 2005; Schoukens and Tiels, 2011; Yang *et al.*, 2017), multilayer perceptrons (Al-Duwaish *et al.*, 1996; Janczak, 2005; Ławryńczuk, 2013), kernel expansions (Van Vaerenbergh *et al.*, 2013), or nonparametric representations (Greblicki, 1997; 2001) are commonly used for modeling the static nonlinear element or its inverse. The linear dynamic subsystem is usually represented by a transfer function (Janczak, 2005; Dong *et al.*, 2009; Schoukens and Tiels, 2011; 2017; Ławryńczuk, 2013; 2016; Ding *et al.*, 2015; Xiong *et al.*, 2015; Mahataa *et al.*, 2016; Bottegal *et al.*, 2017; Kazemi and Arefi, 2017; Yang *et al.*, 2001), impulse response model (Greblicki, 1997; 2001), a finite impulse response model (Fan and Lo, 2009; Van Vaerenbergh *et al.*, 2013), a state-space model (Westwick and Verhaegen, 1996; Gómez and Baeyens, 2002; 2005), Laguerre orthonormal basis functions (Stanisławski *et al.*, 2014; Jansson and Medvedev, 2015), or generalized orthonormal basis functions (Tiels and Schoukens, 2014).

In polynomial Wiener models, the linear part is represented by a transfer function and the nonlinear element or its inverse is described by a polynomial of a known degree. An online algorithm for identification of polynomial Wiener systems in the presence of output colored noise was proposed by Kazemi and Arefi (2017). Their method is based on the extended recursive least squares algorithm, and an unknown intermediate signal is estimated by using an inner iterative algorithm. Zhou *et al.* (2013) proposed a gradient based iterative identification algorithm to estimate parameters of a discrete transfer model and a polynomial model of nonlinearity using the key-term separation principle. The identification of Wiener time-delay systems with the output data contaminated with outliers was considered by Yang *et al.* (2017). In this approach, the parameter estimation problem is solved in the framework of the expectation maximization (EM) algorithm and the time-delay and outliers are handled simultaneously. Both the time-delay and other unknown model parameters are estimated iteratively. The EM-based identification algorithm was also applied by Xiong *et al.* (2015) for the identification of Wiener systems with missing output data.

System parameters of a nonlinear cascade system with output hysteresis were estimated by Vörös (2015) based on the key-term separation principle and using a least squares based iterative algorithm with internal variable estimation. The polynomial representation of an inverse nonlinear function and a transfer function model of a linear dynamic system were considered by Ding *et al.* (2015). In their approach, system parameters are estimated using the recursive least squares method, in which unmeasurable variables are replaced by their estimates but the prediction error is still nonlinear in parameters. The strategy for modeling and identification of fractional-order discrete-time Wiener systems was proposed by Stanisławski *et al.* (2014). In this approach, discrete-time Laguerre filters are uniquely embedded in the modeling of the fractional-order dynamics, resulting in a linear regression formulation of parameter estimation problem.

It was shown by Janczak (2005; 2007; 2018) that by introducing a new extended parametrization it is possible to transform the polynomial Wiener model with inverse nonlinearity described by a polynomial into a linear-in-parameters form. Unfortunately, in this case, some regressors are correlated with the system disturbance and the least squares (LS) parameter estimates are inconsistent. To obtain consistent parameter estimates, the instrumental variables (IV) method is used. This simple non-iterative approach, however, has two serious inconveniences. One is an increased parameter variance error, i.e., the part of the model error that is due to a deviation of estimated parameters from their optimal values, as the number of parameters of the transformed Wiener model is much greater than that of parameters of the original one. The other is parameter redundancy, as there is no unique relationship between the transformed Wiener model and the original one, i.e., there are many combinations of the original inverse nonlinear function model parameters that can be obtained from transformed Wiener model parameters. The parameter redundancy problem was addressed by Janczak (2018) in a single stage estimation procedure in which parameters of an inverse nonlinear function are calculated via solving an overdetermined system of equations, calculating the average values of redundant parameters or calculating parameters from sums of transformed Wiener model parameters.

This paper presents a new two-stage approach that outperforms the single stage solution as it contains an additional estimation step, in which parameters of the inverse nonlinear function are determined uniquely using the IV method based on the assumption that parameters of the linear dynamic model are already known. The Wiener model determined in this way is not overparameterized and a higher parameter estimation accuracy can be achieved.

Many common Wiener systems have invertible nonlinearities that cannot be described accurately by a polynomial, and the assumption that the inverse nonlinear function can be expressed by the polynomial of a given degree is a very restrictive one. However, there are many inverse nonlinear functions that are both continuous and invertible on a given closed interval. In this case, according to the Stone–Weierstrass approximation theorem, they can be uniformly approximated as closely as desired by a polynomial function. Therefore, using the proposed approach, a polynomial approximation of the inverse nonlinear function can be obtained with an arbitrary high accuracy.

The structure of the paper is as follows. The polynomial representation of the true Wiener system is introduced in Section 2. Then the transformation of the polynomial Wiener model into a linear-in-parameters form is shown in Section 3, which also contains details of the LS and IV estimation of transformed model parameters. The IV estimation of nonlinear function parameters is described in Section 4. The effectiveness of the proposed approach is illustrated with a numerical example in Section 5. Finally, conclusions are presented in Section 6.

2. Polynomial Wiener system representation

We assume that the Wiener system to be identified (Fig. 1) is composed of a linear dynamic system described by the transfer function followed by an invertible static nonlinear element. The output v_i of the linear dynamic system to the input u_i at time i,

$$v_i = \frac{B_0(q^{-1})}{A_0(q^{-1})} u_i + \varepsilon_i, \tag{1}$$

is the input to the memoryless nonlinear element

$$y_i = f_0(v_i), \tag{2}$$

where

$$A_0(q^{-1}) = 1 + a_1^0 q^{-1} + \cdots + a_{n_a}^0 q^{-n_a}, \tag{3}$$

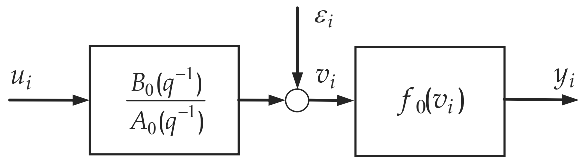

Fig. 1. Wiener system.

$$B_0(q^{-1}) = b_1^0 q^{-1} + \cdots + b_{n_b}^0 q^{-n_b}, \tag{4}$$

and $f_0(\cdot)$ is the nonlinear element characteristic, q^{-1} is the backward shift operator, $a_1^0, \ldots, a_{n_a}^0$, $b_1^0, \ldots, b_{n_b}^0$ are the unknown parameters of linear dynamic system, and ε_i is the system disturbance.

The following assumptions are made about the Wiener system and the system input:

(A1) The linear dynamic system is casual and asymptotically stable.

(A2) $f_0(v_i)$ is a real, continuous, bounded, and strictly monotonic nonlinear function over the restricted domain $[v_{\min}, v_{\max}]$.

(A3) The inverse function of $f_0(v_i)$ is a polynomial of degree r,

$$f_0^{-1}(y_i) = \gamma_0^0 + \gamma_1^0 y_i + \gamma_2^0 y_i^2 + \cdots + \gamma_r^0 y_i^r. \tag{5}$$

(A4) The first degree term in (5) is nonzero. Without loss of generality, we assume $\gamma_1^0 = 1$.

(A5) The polynomial degrees n_a, n_b, and r are known.

(A6) The input u_i is bounded and persistently exciting of order n_b.

(A7) The input u_i is chosen in a such way that v_i covers the full domain of $f_0(v_i)$.

The identification problem is formulated as follows. Given the sequence of the Wiener system input and output measurements $\{u_i, y_i\}$, $i = 1, 2, \ldots, N$, estimate parameters of the linear dynamic system and the inverse nonlinear element minimizing the following criterion:

$$J_N(\theta) = \frac{1}{2} \sum_{j=1}^{N} e_i^2(\theta), \tag{6}$$

where

$$e_i(\theta) = y_i - \hat{y}_i \tag{7}$$

is the equation error, $\hat{y}_i$ is the Wiener model output, and θ is the model parameter vector.

3. Stage I: Least squares and instrumental variables parameter estimation of a transformed Wiener system

The polynomial Wiener system is nonlinear in parameters, and to estimate its parameters, nonlinear optimization methods are commonly used (e.g., Bottegal *et al.*, 2017). In this section, we show that by introducing—a new extended parametrization, the Wiener model (1)–(5) can be transformed into a linear-in-parameters form. Then consistent parameter estimates of this transformed model can be obtained using the combined least squares and instrumental variables identification procedure.

3.1. Least squares parameter estimation. Assume that the following Wiener model is given:

$$\hat{y}_i = f\left[\frac{B(q^{-1})}{A(q^{-1})}u_i\right], \tag{8}$$

with

$$A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}, \tag{9}$$

$$B(q^{-1}) = b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}, \tag{10}$$

where $f(\cdot)$ is the model nonlinear function, $A(q^{-1})$, $B(q^{-1})$ are the linear model polynomials, and $a_1, \ldots, a_{n_a}$, $b_1, \ldots, b_{n_b}$ are the parameters of a linear dynamic model. If $f(\cdot)$ is invertible, (8) can be written as

$$f^{-1}(\hat{y}_i) = \frac{B(q^{-1})}{A(q^{-1})}u_i. \tag{11}$$

We assume that the function $f^{-1}(\cdot)$ is a polynomial of the r-th degree, i.e.,

$$f^{-1}(\hat{y}_i) = \gamma_0 + \gamma_1 \hat{y}_i + \gamma_2 \hat{y}_i^2 + \cdots + \gamma_r \hat{y}_i^r. \tag{12}$$

Assume also that $\gamma_1 \neq 0$. Without the loss of generality, we can assume that $\gamma_1 = 1$. Then substituting (12) into (11), the model output can be expressed as

$$\hat{y}_i = \frac{B(q^{-1})}{A(q^{-1})}u_i - \Delta f^{-1}(\hat{y}_i), \tag{13}$$

where

$$\Delta f^{-1}(\hat{y}_i) = \gamma_0 + \gamma_2 \hat{y}_i^2 + \gamma_3 \hat{y}_i^3 + \cdots + \gamma_r \hat{y}_i^r. \tag{14}$$

The model (13) can be written as

$$\begin{aligned}\hat{y}_i = {} & \big[1 - A(q^{-1})\big]\hat{y}_i + B(q^{-1})u_i \\ & - A(q^{-1})\Delta f^{-1}(\hat{y}_i).\end{aligned} \tag{15}$$

Replacing $\hat{y}_i$ with y_i on the right-hand side of (15), the following modified series-parallel model can be obtained (Janczak, 2005):

$$\begin{aligned}\hat{y}_i = {} & \big[1 - A(q^{-1})\big]y_i + B(q^{-1})u_i \\ & - A(q^{-1})\Delta f^{-1}(y_i)\end{aligned} \tag{16}$$

or, equivalently,

$$\begin{aligned}\hat{y}_i = {} & -\sum_{m=1}^{n_a} a_m y_{i-m} + \sum_{m=1}^{n_b} b_m u_{i-m} - \gamma_0 \\ & - \sum_{j=2}^{r}\gamma_j y_i^j - \sum_{m=1}^{n_a} a_m\Big(\gamma_0 + \sum_{j=2}^{r}\gamma_j y_{i-m}^j\Big).\end{aligned} \tag{17}$$

Introducing a new extended parametrization, the model (17) can be transformed into the following linear-in-parameters form:

$$\begin{aligned}\hat{y}_i = {} & -\sum_{m=1}^{n_a} a_m y_{i-m} + \sum_{m=1}^{n_b} b_m u_{i-m} - \alpha_{0,0} \\ & - \sum_{j=2}^{r}\sum_{m=0}^{n_a}\alpha_{j,m} y_{i-m}^r,\end{aligned} \tag{18}$$

where

$$\alpha_{0,0} = \big(1 + \sum_{m=1}^{n_a} a_m\big)\gamma_0 \tag{19}$$

and

$$\begin{aligned}& \alpha_{j,m} \\ & = \begin{cases}\gamma_j, & j = 2,3,\ldots,r,\ m = 0, \\ a_m\gamma_j, & j = 2,3,\ldots,r,\ m = 1,2,\ldots,n_a.\end{cases}\end{aligned} \tag{20}$$

Thus, the model (18) can be written as

$$\hat{y}_i = x_i^T\theta, \tag{21}$$

with the model parameter vector

$$\begin{aligned}\theta = {} & \big[a_1 \ldots a_{n_a}\ b_1 \ldots b_{n_b}\ \alpha_{0,0}\ \alpha_{2,0} \ldots \alpha_{2,n_a} \\ & \ldots \alpha_{r,0} \ldots \alpha_{r,n_a}\big]^T\end{aligned} \tag{22}$$

and the regressor vector

$$\begin{aligned}x_i = {} & \big[-y_{i-1} \ldots -y_{i-n_a}\ u_{i-1} \ldots u_{i-n_b}\ \ 1 \\ & -y_i^2 \ldots -y_{i-n_a}^2 \ldots -y_i^r \ldots -y_{i-n_a}^r\big]^T.\end{aligned} \tag{23}$$

Minimizing (6) using the LS method, the following parameter vector estimate can be obtained:

$$\hat{\theta}_N^{LS} = \left[\frac{1}{N}\sum_{i=1}^{N} x_i x_i^T\right]^{-1}\left[\frac{1}{N}\sum_{i=1}^{N} x_i y_i\right]. \tag{24}$$

Note that the number of parameters in (21) is $n_a + n_b + (r-1)(n_a+1) + 1$, while the number of parameters of $A(q^{-1})$, $B(q^{-1})$, and $f(\cdot)$ is $n_a + n_b + r$. Moreover, there are $n_a + 1$ values of parameter estimates γ_j, $j = 2, 3, \ldots, r$, that can be calculated from (20). The overall number of combinations of parameter estimates γ_j that can be obtained from parameter estimates $\alpha_{j,k}$, $j = 2,\ 3, \ldots, r.$, $m = 0, 1, \ldots, n_a$ is $(n_a+1)^{r-1}$.

3.2. Inconsistency of the least squares Wiener system parameter estimator. It can be shown that the LS estimate of θ_0 is inconsistent, i.e., $\hat{\theta}_N^{LS}$ does not converge (with probability 1) to the true parameter vector θ_0 even if the linear part is the autoregressive with exogenous input (ARX) system and the additive disturbance ε_i is

$$\varepsilon_i = \frac{\epsilon_i}{A_0(q^{-1})}, \tag{25}$$

where ϵ_i is the discrete white noise.

The output y_i of the Wiener system, defined by (1)–(5) and (25), is

$$\begin{aligned} y_i = &\big[1 - A_0(q^{-1})\big]y_i + B_0(q^{-1})u_i \\ &- A_0(q^{-1})\Delta f_0^{-1}(y_i) + \epsilon_i, \end{aligned} \tag{26}$$

where $\Delta f_0^{-1}(y_i) = f_0^{-1}(y_i) - y_i$.

Introducing the true parameter vector θ_0,

$$\begin{aligned} \theta^0 = \big[&a_1^0 \dots a_{n_a}^0 \; b_1^0 \dots b_{n_b}^0 \; \alpha_{0,0}^0 \; \alpha_{2,0}^0 \dots \alpha_{2,n_a}^0 \\ &\dots \alpha_{r,0}^0 \dots \alpha_{r,n_a}^0\big]^T, \end{aligned} \tag{27}$$

where

$$\alpha_{0,0}^0 = \big(1 + \sum_{m=1}^{n_a} a_m^0\big)\gamma_0^0, \tag{28}$$

$$\alpha_{j,m}^0 = \begin{cases} \gamma_j^0, & j = 2,3,\dots,r,\ m = 0, \\ a_m^0\gamma_j^0, & j = 2,3,\dots,r,\ m = 1,2,\dots,n_a, \end{cases} \tag{29}$$

the system output can be expressed as

$$y_i = x_i^T\theta_0 + \epsilon_i. \tag{30}$$

From (24) and (30), it follows that

$$\hat{\theta}_N^{LS} = \theta_0 + \left[\frac{1}{N}\sum_{i=1}^{N} x_i x_i^T\right]^{-1}\left[\frac{1}{N}\sum_{i=1}^{N} x_i\epsilon_i\right]. \tag{31}$$

It follows immediately for the estimate $\hat{\theta}_N^{LS}$ to be consistent, i.e., to converge in probability to θ_0, that the following two conditions should be fulfilled:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} x_i x_i^T \text{ must be nonsingular,} \tag{32}$$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} x_i\epsilon_i = 0. \tag{33}$$

Condition (32) requires the persistency of excitation and it can be fulfilled easily. Condition (33) is not fulfilled, as the regressors y_i^2, y_i^3, ..., y_i^r are correlated with ϵ_i, i.e., $\mathrm{E}\big[y_i^2\epsilon_i\big] \neq 0$,...,$\mathrm{E}\big[y_i^r\epsilon_i\big] \neq 0$, and thus $\mathrm{E}\big[x_i\epsilon_i\big] \neq 0$.

3.3. Instrumental variables parameter estimation. To obtain consistent parameter estimates, the regression vector should be uncorrelated with the noise. This is not the case if we use a transformed model (18). The instrumental variables method is a well-known remedy for such a situation. Using the IV method, the parameter estimation can be performed according to the following four step scheme:

1. Estimate parameters using the LS method.
2. Simulate the linear dynamic model
$$\hat{s}_i = \frac{B(q^{-1})}{A(q^{-1})}u_i. \tag{34}$$
3. Choose instrumental variables vector z_i uncorellated with the noise.
4. Estimate parameters using the IV method with the instrumental variables vector z_i.

The choice of instrumental variables is an important design problem in any IV approach. Clearly, the best choice would be the undisturbed system outputs, but these are not available for measurement. Instead of the system outputs in (23), we can use the outputs of the linear model obtained using the LS method, and define the instrumental variables vector as

$$\begin{aligned} z_i = \big[&-\hat{s}_{i-1} \dots -\hat{s}_{i-n_a} \; u_{i-1} \dots u_{i-n_b} \;\; 1 \\ &-\hat{s}_i^2 \dots -\hat{s}_{i-n_a}^2 \dots -\hat{s}_i^r \dots -\hat{s}_{i-na}^r\big]^T. \end{aligned} \tag{35}$$

The instrumental variables vector z_i is uncorrelated with the system noise, i.e., $E[z_i\epsilon_i] = 0$. The IV estimate of θ_0 is given by

$$\hat{\theta}_N^{IV} = \left[\frac{1}{N}\sum_{i=1}^{N} z_i x_i^T\right]^{-1}\left[\frac{1}{N}\sum_{i=1}^{N} z_i y_i\right]. \tag{36}$$

The IV parameter estimates are calculated based on the simulated linear dynamic model output (34), which is uncorrelated with the noise but can be far from the linear dynamic system output as $\hat{\theta}_N^{LS}$ is inconsistent. Therefore, the estimation process can be further improved by repeating Steps 2–4 and using in Step 2 of each iteration the IV parameter estimates from the previous iteration. As such a procedure converges very fast, two or three iterations are commonly enough.

4. Stage II: Estimation of nonlinear system parameters

It is clear that there is no unique relationship between the parameters of the transformed model and those of the inverse nonlinear function as

$$\hat{\alpha}_{j,m}^{IV} = \begin{cases} \hat{\gamma}_j, & j = 2,3,\dots,r,\ m = 0, \\ \hat{a}_m^{IV}\hat{\gamma}_j, & j = 2,3,\dots,r,\ m = 1,2,\dots,n_a, \end{cases} \tag{37}$$

where $\hat{\alpha}_{j,m}^{IV}$ and $\hat{a}_m^{IV}$ denote IV parameter estimates obtained in Stage I. In other words, there is a parameter redundancy as $n_a + 1$ different values of $\hat{\gamma}_j,\ j = 2, 3, \ldots, r$ can be calculated from $\hat{\alpha}_{j,m}^{IV},\ m = 0, 1, \ldots, n_a$. Note that the parameter estimate $\hat{\gamma}_0$ can be calculated uniquely from

$$\hat{\alpha}_{0,0}^{IV} = \big(1 + \sum_{m=1}^{n_a} \hat{a}_m^{IV}\big)\hat{\gamma}_0. \tag{38}$$

To avoid this inconvenience, an additional estimation step is proposed in which parameter estimates of the inverse nonlinear function are calculated uniquely. We assume that parameters of the linear dynamic model (17) are already known and they are equal to their IV estimates, i.e., $a_m = \hat{a}_m^{IV}$ and $b_m = \hat{b}_m^{IV}$.

Define the following parameter vector:

$$\hat{\theta}_1^{IV} = \big[\hat{a}_1^{IV} \ldots \hat{a}_{n_a}^{IV}\ \hat{b}_1^{IV} \ldots \hat{b}_{n_b}^{IV}\big]^T, \tag{39}$$

the vector

$$\tilde{x}_i = \big[-y_{i-1} \ldots - y_{i-n_a}\ u_{i-1} \ldots u_{i-n_b}\big]^T \tag{40}$$

and the auxiliary variable

$$\tilde{s}_i = \tilde{x}_i^T \hat{\theta}_1^{IV}, \tag{41}$$

where $\hat{a}_1^{IV}, \ldots, \hat{a}_{n_a}^{IV}$ and $\hat{b}_1^{IV}, \ldots, \hat{b}_{n_b}^{IV}$ are IV parameter estimates obtained in Stage I. Define also the parameter vector

$$\theta_2 = \big[\gamma_0\ \ \gamma_2\ \ \gamma_3 \ldots \gamma_r\big]^T, \tag{42}$$

the regressor vector

$$\begin{aligned}\bar{x}_i = \Big[1 + \sum_{m=1}^{n_a} \hat{a}_m^{IV} \quad y_i^2 + \sum_{m=1}^{n_a} \hat{a}_m^{IV} y_{i-m}^2 \quad y_i^3 \\ + \sum_{m=1}^{n_a} \hat{a}_m^{IV} y_{i-m}^3 \ \ldots\ y_i^r + \sum_{m=1}^{n_a} \hat{a}_m^{IV} y_{i-m}^r\Big]^T\end{aligned} \tag{43}$$

and the instrumental variables vector

$$\begin{aligned}\bar{z}_i = \Big[1 + \sum_{m=1}^{n_a} \hat{a}_m^{IV} \quad \tilde{s}_i^2 + \sum_{m=1}^{n_a} \hat{a}_m^{IV} \tilde{s}_{i-m}^2 \quad \tilde{s}_i^3 \\ + \sum_{m=1}^{n_a} \hat{a}_m^{IV} \tilde{s}_{i-m}^3 \ \ldots\ \tilde{s}_i^r + \sum_{m=1}^{n_a} \hat{a}_m^{IV} \tilde{s}_{i-m}^r\Big]^T.\end{aligned} \tag{44}$$

Then the nonlinear component of (17) can be expressed as

$$\bar{v}_i = \bar{x}_i^T \theta_2. \tag{45}$$

Define the criterion

$$J_N(\theta_2) = \frac{1}{2} \sum_{j=1}^{N} e_i^2(\theta_2), \tag{46}$$

where

$$e_i(\theta_2) = y_i^* - \bar{v}_i \tag{47}$$

and $y_i^* = y_i - \tilde{s}_i$. The consistent estimate of θ_2 can be obtained using the IV method,

$$\hat{\theta}_{2,N}^{IV} = \Bigg[\frac{1}{N} \sum_{i=1}^{N} \bar{z}_i \bar{x}_i^T\Bigg]^{-1} \Bigg[\frac{1}{N} \sum_{i=1}^{N} \bar{z}_i y_i^*\Bigg]. \tag{48}$$

5. Simulation example

The Wiener system composed of the linear dynamic system:

$$\frac{B_0(q^{-1})}{A_0(q^{-1})} = \frac{0.3631q^{-1} - 0.3160q^{-2}}{1 - 1.6253q^{-1} + 0.6592q^{-2}}, \tag{49}$$

and the nonlinear function

$$y_i = \sqrt[5]{5v_i} + 1 \tag{50}$$

is used in the simulation example. The function (50) is strictly monotonic, and it has neither undefined points nor domain constraints. The inverse nonlinear function of (50) is the following polynomial:

$$f_0^{-1}(y_i) = -0.2 + y_i - 2y_i^2 + 2y_i^3 - y_i^4 + 0.2y_i^5. \tag{51}$$

The input sequence $\{u_i\}$ contains 10^5 pseudo-random numbers uniformly distributed in $[-1, 1]$. The additive system disturbance is given by

$$\varepsilon_i = \frac{1}{A_0(q^{-1})}\epsilon_i, \tag{52}$$

where $\{\epsilon_i\}$ is the Gaussian pseudo-random sequence $\mathcal{N}(0, 0025)$.

The LS and IV parameter estimation was performed assuming $\gamma_1 = 1$, $n_a = 2$, $n_b = 2$ and $r = 5$. First, in Stage I, the LS and IV parameter estimates of the transformed Wiener model were calculated. The LS and IV parameter estimates of transformed Wiener model (18) given in Table 1, the inverse nonlinear function and its LS and IV estimates shown in Fig. 2, and the estimation error of the inverse nonlinear function shown in Fig. 3 confirm both the inconsistency of LS estimates and the consistency of IV estimates. Then, Stage II was performed to estimate parameters of the inverse nonlinear function uniquely. For comparison, the unique values of inverse nonlinear function parameters were also calculated solving the overdetermined set of equations (37) (Janczak, 2018).

Parameter estimates $\hat{\gamma}_j,\ j = 2, 3, \ldots, r$, of the inverse nonlinear function obtained using the LS method, applying the IV approach and solving the overdetermined set of equations (OVIV), and performing a two-stage estimation procedure (2SIV), are given in

Table 2. To compare the estimation accuracy, 50 different input-output data sets were generated, and the identification experiment was performed for each input-output data set. For the parameter estimation accuracy comparison, the following mean square error index is used:

$$\mathrm{MSE} = \frac{1}{5}\Big[(\gamma_0 - \hat{\gamma}_0)^2 + \sum_{j=2}^{5}(\gamma_j - \hat{\gamma}_j)^2\Big]. \tag{53}$$

The MSE values for the LS, OVIV, and 2SIV parameter estimates are shown in Fig. 4, and the corresponding mean values and variances of the MSE index are given in Table 3. The obtained results demonstrate that the two-stage estimation procedure makes it possible to achieve a higher parameter estimation accuracy in comparison with the one-stage IV procedure.

Table 1. Parameter estimates of the transformed Wiener model.

Parameter	True	LS	IV
$\hat{b}_2$	0.3168	0.2294	0.3160
$\hat{a}_1$	-1.6253	-1.4469	-1.6276
$\hat{a}_2$	0.6592	0.5441	0.6615
$\hat{\alpha}_{2,0}$	-2.0000	-1.5129	-1.9940
$\hat{\alpha}_{2,1}$	3.2506	2.4444	3.2374
$\hat{\alpha}_{2,2}$	-1.3185	-0.9867	-1.3098
$\hat{\alpha}_{3,0}$	2.0000	1.1838	1.9981
$\hat{\alpha}_{3,1}$	-3.2506	-2.0777	-3.2467
$\hat{\alpha}_{3,2}$	1.3185	0.8843	1.3165
$\hat{\alpha}_{4,0}$	-1.0000	-0.5500	-1.0010
$\hat{\alpha}_{4,1}$	1.6253	0.9709	1.6284
$\hat{\alpha}_{4,2}$	-0.6592	-0.4153	-0.6619
$\hat{\alpha}_{5,0}$	0.2000	0.1146	0.2004
$\hat{\alpha}_{5,1}$	-0.3251	-0.1966	-0.3261
$\hat{\alpha}_{5,2}$	0.1318	0.0826	0.1326

Table 2. Parameter estimates of the inverse nonlinear function.

Parameter	LS	OVIV	2SIV
$\hat{\gamma}_0$	-0.1151	-0.2139	-0.1994
$\hat{\gamma}_2$	-1.5129	-1.9893	-1.9986
$\hat{\gamma}_3$	1.1838	1.9951	1.9998
$\hat{\gamma}_4$	-0.5500	-1.0007	-1.0009
$\hat{\gamma}_5$	0.1146	0.20004	0.2003

Table 3. Comparison of inverse nonlinear function parameter estimation accuracy.

Method	Mean(MSE)	Var(MSE)
LS	1.21×10^{0}	4.41×10^{-5}
OVIV	1.37×10^{-3}	3.44×10^{-6}
2SIV	2.49×10^{-4}	7.00×10^{-8}

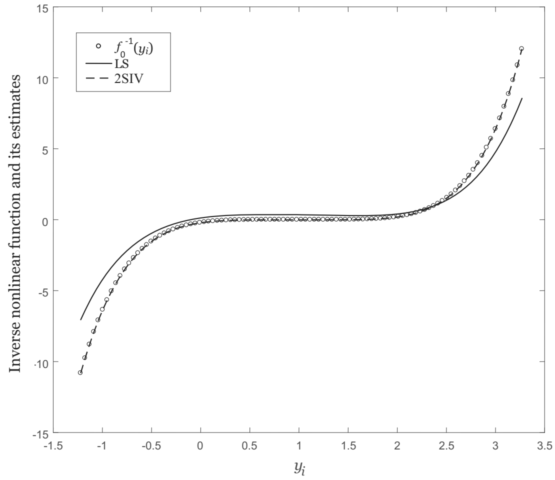

Fig. 2. Inverse nonlinear function $f_0^{-1}(y_i)$ and its LS and two-stage (2SIV) estimates.

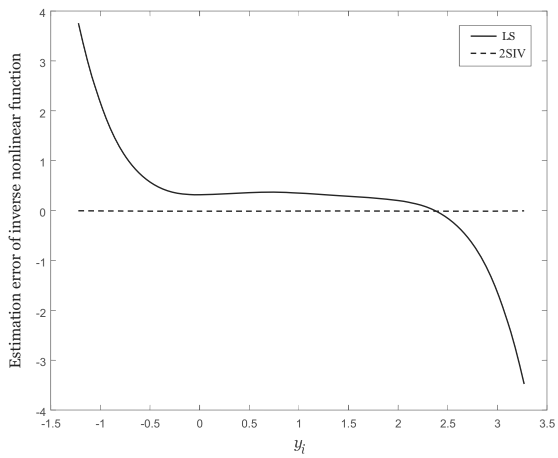

Fig. 3. Estimation error of $\hat{f}_0^{-1}(y_i)$.

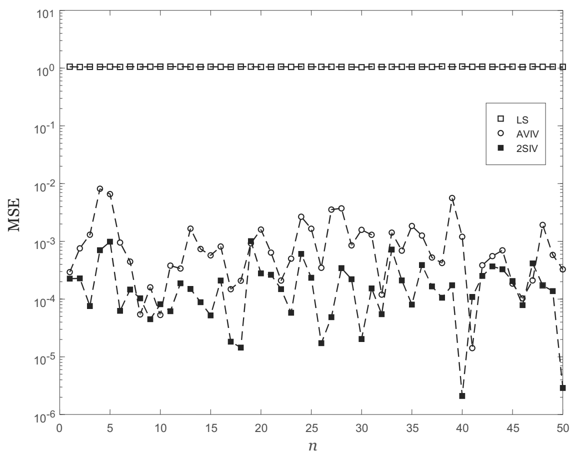

Fig. 4. MSE vs. the trial number.

6. Conclusions

This paper described a two-stage instrumental variables approach to identification of polynomial Wiener systems. It was assumed that the linear dynamic part is given by a transfer function model and the inverse nonlinear element is described by a polynomial. In a more practical case, in which the nonlinear function is not a polynomial but a continuous and invertible function on a given closed interval, parameters of polynomial approximation of the inverse nonlinear function can be obtained.

The proposed identification procedure consists of two stages. In Stage I, introducing a new extended parametrization, the polynomial Wiener model is transformed into a liner-in-parameters form and its parameters are estimated using the combined LS and IV method. It is shown that the LS parameter estimates of the transformed Wiener model are inconsistent. To avoid the inconsistency problem, the IV method is employed, in which the model of the linear dynamic system obtained with the LS method is used to generate instrumental variables. An obvious inconvenience of such an approach is that there is no unique relationship between parameters of transformed model and those of the inverse nonlinear function. Therefore, Stage II is proposed, in which parameters of the inverse nonlinear function are estimated uniquely using the IV method and assuming that linear dynamic model parameters are equal to their IV estimates. The two-stage identification procedure not only produces unique parameter estimates of those inverse nonlinear function but also reduces the model variance error.

The presented algorithm can be extended to the identification of a specific class of MIMO Wiener systems in which the MIMO linear dynamic system is described by a matrix polynomial form and the inverse characteristics of SISO nonlinear elements are polynomials (Janczak, 2007). An extension of this approach to the two-stage case will be considered in future work.

A simulation example was included in the paper to illustrate practical effectiveness of the proposed identification procedure.

References

Al-Duwaish, H., Karim, M. and Chandrasekar, V. (1996). Use of multilayer feedforward neural networks in identification and control of Wiener model, *IEE Proceedings: Control Theory and Applications* **143**(3): 255–258, DOI: 10.1049/ip-cta:19960376.

Aljamaan, I., Westwick, D., Foley, M. and Chandrasekar, V. (2016). Identification of Wiener models in the presence of ARIMA process noise, *IFAC-PapersOnLine* **49**(7): 1008–1013, DOI: 10.1016/j.ifacol.2016.07.334.

Ase, H. and Katayama, T. (2015). A subspace-based identification of two-channel Wiener systems, *IFAC-PapersOnLine* **48**(28): 638–643, DOI: 10.1016/j.ifacol.2015.12.201.

Billings, S. and Fakhouri, S. (1978). Theory of separable processes with applications to the identification of nonlinear systems, *Proceedings of the Institution of Electrical Engineers* **125**(10): 1051–1058, DOI: 10.1049/piee.1978.0241.

Billings, S. and Fakhouri, S. (1982). Identification of systems containing linear dynamic and static nonlinear elements, *Automatica* **18**(1): 15–26, DOI: 0.1016/0005-1098(82)90022-X.

Bottegal, G., Castro-Garcia, R. and Suykens, J. (2017). On the identification of Wiener systems with polynomial nonlinearity, *2017 IEEE 56th Annual Conference on Decision and Control (CDC), Melbourne, Australia*, DOI: 10.1109/CDC.2017.8264635.

Boyd, S. and Chua, L.O. (1985). Fading memory and the problem of approximating nonlinear operators with Volterra series, *IEEE Transactions on Circuits and Systems* **32**(11): 1150–1161, DOI: 10.1109/TCS.1985.1085649.

Brouri, A. and Slassi, S. (2015). Frequency identification approach for Wiener systems, *International Journal of Computational Engineering Research (IJCER)* **5**(8): 12–16.

Ding, F., Liu, X. and Liu, M. (2015). The recursive least squares identification algorithm for a class of Wiener nonlinear systems, *Journal of the Franklin Institute* **353**(7): 1518–1526, DOI: 10.1016/j.jfranklin.2016.02.013.

Dong, R., Tan, Q. and Tan, Y. (2009). Recursive identification algorithm for dynamic systems with output backlash and its convergence, *International Journal of Applied Mathematics and Computer Science* **19**(4): 631–638, DOI: 10.2478/v10006-009-0050-2.

Fan, D. and Lo, K. (2009). Identification for disturbed MIMO Wiener systems, *Nonlinear Dynamics* **55**(1): 31–42, DOI: 10.1007/s11071-008-9342-6.

Figueroa, J., Biagiola, S., Alvarez, M., Castro, L. and Agamennoni, O.E. (2013). Robust model predictive control of a Wiener-like system, *Journal of the Franklin Institute* **350**(3): 556–574, DOI: 10.1016/j.jfranklin.2012.12.016.

Giri, F., Radouane, A., Brouri, A. and Chaoui, F. (2014). Combined frequency-prediction error identification approach for Wiener systems with backlash and backlash-inverse operators, *Automatica* **50**(3): 768–783, DOI: 10.1016/j.automatica.2013.12.030.

Gómez, M. and Baeyens, E. (2002). Subspace identification of multivariable Hammerstein and Wiener models, *IFAC Proceedings Volumes* **35**(1): 55–60, DOI: 10.3182/20020721-6-ES-1901.00420.

Gómez, M. and Baeyens, E. (2005). Subspace-based identification algorithms for Hammerstein and Wiener models, *European Journal of Control* **11**(2): 127–136, DOI: 10.3166/ejc.11.127-136.

Greblicki, W. (1997). Nonparametric approach to Wiener system identification, *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* **44**(6): 538–545, DOI: 10.1109/81.586027.

Greblicki, W. (2001). Recursive identification of Wiener systems, *International Journal of Applied Mathematics and Computer Science* **11**(4): 977–991.

Ipanaqué, W. and Manrique, J. (2011). Identification and control of pH using optimal piecewise linear Wiener model, *IFAC Proceedings Volumes* **44**(41): 12301–12306, DOI: 10.3182/20110828-6-IT-1002.03695.

Janczak, A. (2005). *Identification of Nonlinear Systems Using Neural Networks and Polynomial Models. A Block-Oriented Approach*, Springer Verlag, Berlin/Heidelberg/New York,NY.

Janczak, A. (2007). Instrumental variables approach to identification of a class of MIMO Wiener systems, *Nonlinear Dynamics* **48**(3): 275–284, DOI: 10.1007/s11071-006-9088-y.

Janczak, A. (2018). Least squares and instrumental variables identification of polynomial Wiener systems, *23rd International Conference on Methods and Models in Automation and Robotics (MMAR'2018), Międzyzdroje, Poland*, DOI: 10.1109/MMAR.2018.8486049.

Jansson, D. and Medvedev, A. (2015). Identification of polynomial Wiener systems via Volterra–Laguerre series with model mismatch, *IFAC-PapersOnLine* **48**(11): 831–836, DOI: 10.1016/j.ifacol.2015.09.293.

Kazemi, M. and Arefi, M. (2017). A fast iterative recursive least squares algorithm for Wiener model identification of highly nonlinear systems, *ISA Transactions* **67**: 382–388, DOI: 10.1016/j.isatra.2016.12.002.

Kim, K.-K., Rios-Patronc, E. and Braatz, R. (2012). Robust nonlinear internal model control of stable Wiener systems, *Journal of Process Control* **22**(8): 1468–1477, DOI: 10.1016/j.jprocont.2012.01.019.

Ławryńczuk, M. (2013). Practical nonlinear predictive control algorithms for neural Wiener models, *Journal of Process Control* **22**(5): 696–714, DOI: 10.1016/j.jprocont.2013.02.004.

Ławryńczuk, M. (2015). Nonlinear state-space predictive control with on-line linerisation and state estimation, *International Journal of Applied Mathematics and Computer Science* **25**(4): 833–847, DOI: 10.1515/amcs-2015-0060.

Ławryńczuk, M. (2016). Modelling and predictive control of a neutralisation reactor using sparse support vector machine Wiener models, *Neurocomputing* **205**: 311–328, DOI: 10.1016/j.neucom.2016.03.066.

Mahataa, K., Schoukens, J. and Cock, A.D. (2016). Information matrix and D-optimal design with Gaussian inputs for Wiener model identification, *Automatica* **69**: 65–77, DOI: 10.1016/j.automatica.2016.02.026.

Rollins, D., Mei, Y., Loveland, S. and Bhandari, N. (2016). Block-oriented feedforward control with demonstration to nonlinear parameterized Wiener modeling, *Chemical Engineering Research and Design* **109**: 397–404, DOI: 10.1016/j.cherd.2016.02.016.

Schoukens, M. and Tiels, K. (2011). Parametric MIMO parallel Wiener identification, *2011 50th IEEE Conference on Decision and Control/European Control Conference, Orlando, FL, USA*, DOI: 10.1109/CDC.2011.6160230.

Schoukens, M. and Tiels, K. (2017). Identification of block-oriented nonlinear systems starting from linear approximations: A survey, *Automatica* **85**: 272–292, DOI: 10.1016/j.automatica.2017.06.044.

Stanisławski, R., Latawiec, K., Gałek, M. and Łukaniszyn, M. (2014). Modeling and identification of a fractional-order discrete-time SISO Laguerre–Wiener system, *19th International Conference on Methods and Models in Automation and Robotics (MMAR'2014), Międzyzdroje, Poland*, DOI: 10.1109/MMAR.2014.6957343.

Tiels, K. and Schoukens, J. (2014). Wiener system identification with generalized orthonormal basis functions, *Automatica* **50**(12): 3147–3154, DOI: 10.1016/j.automatica.2014.10.010.

Van Vaerenbergh, S., Via, J. and Santamaria, I. (2013). Blind identification of SIMO Wiener systems based on kernel canonical correlation analysis, *IEEE Transactions on Signal Processing* **61**(9): 2219–2230, DOI: 10.1109/TSP.2013.2248004.

Vörös, J. (2007). Parameter identification of Wiener systems with multisegment piecewise-linear nonlinearities, *Systems & Control Letters* **56**(2): 99–105, DOI: 10.1016/j.sysconle.2006.08.001.

Vörös, J. (2015). Identification of nonlinear cascade systems with output hysteresis based on the key term separation principle, *Applied Mathematical Modelling* **39**(18): 5531–5539, DOI: 10.1016/j.apm.2015.01.018.

Westwick, D. and Verhaegen, M. (1996). Identifying MIMO Wiener systems using subspace model identification methods, *Systems & Control Letters* **52**(2): 235–258, DOI: 10.1016/0165-1684(96)00056-4.

Wigren, T. (1993). Recursive prediction error identification algorithm using the nonlinear Wiener model, *Automatica* **29**(4): 1011–1025, DOI: 10.1016/0005-1098(93)90103-Z.

Xiong, W., Yang, X., Ke, L. and Xu, B. (2015). EM algorithm-based identification of a class of nonlinear Wiener systems with missing output data, *Nonlinear Dynamics* **80**(1–2): 329–339, DOI: 10.1007/s11071-014-1871-6.

Yang, X., Xiong, W., Ma, J. and Wang, Z. (2017). Robust identification of Wiener time-delay system with expectation-maximization algorithm, *Journal of the Franklin Institute* **354**(13): 5678–5693, DOI: 10.1016/j.jfranklin.2017.05.023.

Zhou, L., Li, X. and Pan, F. (2013). Gradient based iterative parameter identification for Wiener nonlinear systems, *Applied Mathematical Modelling* **37**(16–17): 8203–8209, DOI: 10.1016/j.apm.2013.03.005.

Andrzej Janczak received the MSc degree in electronics engineering from the Gdańsk University of Technology in 1976 and the PhD degree in control engineering from the Silesian University of Technology in 1986. He obtained the DSc degree in electrical engineering from the University of Zielona Góra in 2005. He is currently an associate professor of automatic control and the head of the Division of Control Systems and Robotics at the Faculty of Computer, Electrical and Control Engineering, University of Zielona Góra, Poland. His main research interests cover system identification, including identification of linear repetitive processes using subspace methods, as well as application of artificial neural networks in system control and identification.

Józef Korbicz has been a full-rank professor of automatic control at the University of Zielona Góra, Poland, since 1994. In 2007 he was elected a corresponding member of the Polish Academy of Sciences (PAS). His current research interests include fault detection and isolation, control theory and computational intelligence. He has published more than 370 scientific publications, authored or co-authored 8 books and co-edited 7 books. He served as the IPC chairman of the IFAC *SAFEPROCESS* Symposium held in Beijing, China (2006), and as the chairman of the NOC for *SAFEPROCESS* held in Warsaw, Poland (2018). Together with Prof. J.M. Kościelny, he founded the Polish conferences on *Diagnostics of Processes and Systems, DPS*, organized every two years. He is currently the chair of the Committee on Automatic Control and Robotics of the PAS, a senior member of the IEEE and a member of the IFAC SAFEPROCESS TC. More at `http://www.uz.zgora.pl/~jkorbicz/`.

Received: 13 December 2018
Revised: 3 April 2019
Accepted: 10 April 2019