Nonlinear Dyn (2009) 55: 31–42
DOI 10.1007/s11071-008-9342-6

ORIGINAL PAPER

Identification for disturbed MIMO Wiener systems

Dan Fan · Kueiming Lo

Received: 16 April 2007 / Accepted: 12 February 2008 / Published online: 3 March 2008

Abstract The identification of Multi-input Multi-output (MIMO) Wiener systems is concerned in this paper. The system presented is comprised of a multi-dimensional linear subsystem and a memory-less nonlinear block which is made of discontinuous asymmetric piece-wise linear functions. A recursive algorithm is proposed to estimate all the unknown parameters of the system with interference noises. It is shown that the recursive algorithm for the disturbed MIMO Wiener system is convergent. Finally, some simulation results illustrate the identification accuracy and the convergence rate.

Keywords Convergence rate · Interference noise · MIMO Wiener systems · Parameter identification · Recursive estimation

This work is supported by the Funds 2004CB719400, NSFC60672110, NSFC60474026 and the JSPS Foundation.

D. Fan (✉) · K. Lo
School of Software, Key Lab for ISS of MOE, Tsinghua University, Beijing 100084, China
e-mail: fd05@mails.tsinghua.edu.cn

K. Lo
Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan

1 Introduction

In telecommunications, biotechnology, industrial chemistry, automatic control, and many other fields, nonlinear models are needed to describe real systems [1–3]. However, due to their complexity and variety, generally it is hard to identify the structures of these models. Among them, a class of models called Wiener models [4, 5] are composed of dynamic linear subsystems in series with static nonlinear functions.

Because of the particular structure of this kind of models, it is possible to simulate the structure and behavior of whole systems on the basis of the input–output measurements, that is, to estimate the unknown coefficients of the linear subsystems and the characteristic curve of the nonlinear subsystem. Moreover, systems with a finite Volterra series representation can be approximated to arbitrary accuracy by a finite (parallel) sum of Wiener systems [6, 7].

By far, Wiener models have been used successfully in many applications, such as modeling the pH neutralization process [8, 9], automobile [10], distillation columns [11, 12], chromatographic separation process [13], adaptive precompensation of nonlinear distortions [14], visual systems [15], biological systems [16], chaotic systems [17], and so on. For example, the pH system with the input reagent flow-rate and output pH, can be approximated by a Wiener model. The dynamic linear block is used to describe the mixing state of the reacting streams in the reactor, while the relationship between the reaction species and the pH inside the reactor can be represented by the static nonlinear block [8, 9].

In the above models, however, the signals between their two subsystems cannot be measured. It makes the identification of such systems difficult. In order to overcome this problem, several approaches have been proposed for the identification and control of the Wiener models (iterative algorithms [16], neural networks [13], correlation analysis [18, 19], parametric regression [8, 20, 21], and nonparametric regression [22]. However, most of these methods are built for the Single-input Single-output (SISO) systems. Since the systems in practical applications could be abstracted as multi-dimensional models [16, 23], great attention has been paid to the study of the MIMO Wiener models in recent years. It is also the focus of the research presented in this paper. In [13], the neural network and discrete Laguerre filters are used to model a MIMO Wiener system. In [24], the MIMO Wiener system based on a state space formulation of the model and the subspace identification algorithm is studied. Recently, a combined least squares and instrumental variables method is presented to identify MIMO Wiener systems in [25], which uses inverse nonlinear blocks.

In this paper, the nonlinear functions of the MIMO Wiener model treated are assumed to be discontinuous asymmetric piece-wise linear functions, whose discontinuity points need not symmetric about the grid origin, since some previous studies have indicated that this kind of system is very important in engineering applications [26, 27], whereas most of the previous methods are not suitable for this kind of models. Given this condition, the identification of the nonlinear block of the system is reduced to the estimation of the unknown parameters in these piece-wise linear functions. Recently, similar kinds of studies have been carried out by some scholars [26–28]. In [26], a key term separation principle is introduced, which is used with an iterative algorithm for the identification of ARMA (Autoregressive Moving Average) Wiener systems. Although it has a good result when the sample size is fixed, especially for linear subsystems, there is no proof of the convergence of this method. Moreover, in this method, the discontinuity points of the nonlinear block must be symmetric about the grid origin, which is not satisfied in many practical applications. The same method is used in [28] for the identification of a similar Wiener model with multi-segment piecewise-linear nonlinearities. For better results in estimating the nonlinearity, a recursive estimation algorithm is proposed in [27], and most importantly, its convergence has been proven without no restrictive conditions except the structural assumptions. However, both of the two methods are not for MIMO systems, which consequently can be further improved. Therefore, this paper presents a brief study of this issue in which the recursive estimation algorithm in [27] is extended for MIMO case. Moreover, real systems will inevitably be affected by noise, which is not considered in [27] whereas is also discussed in this paper.

The paper is organized as follows. In Sect. 2, we define our identification problem explicitly. Section 3 gives the assumptions for the discussion of the identification method and its convergency. Then the detailed identification algorithms for the nonlinear and linear subsystems are given in Sects. 4 and 5, respectively, and their convergence are proved at the same time. Section 6 illustrates the identification accuracy and the convergence rate by means of simulation studies. At last, a conclusion is given in Sect. 7.

2 Problem description

To facilitate the following research, let us only take system with the same input and output dimension into consideration, as shown in Fig. 1.

In Fig. 1, we can see that the input vector

$$U(k) = \left[u_1(k), u_2(k), \ldots, u_n(k)\right]^T$$

first goes through a linear subsystem and creates an intermediate vector

$$V(k) = \left[v_1(k), v_2(k), \ldots, v_n(k)\right]^T$$

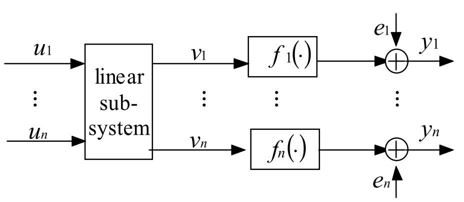

Fig. 1 MIMO Wiener system

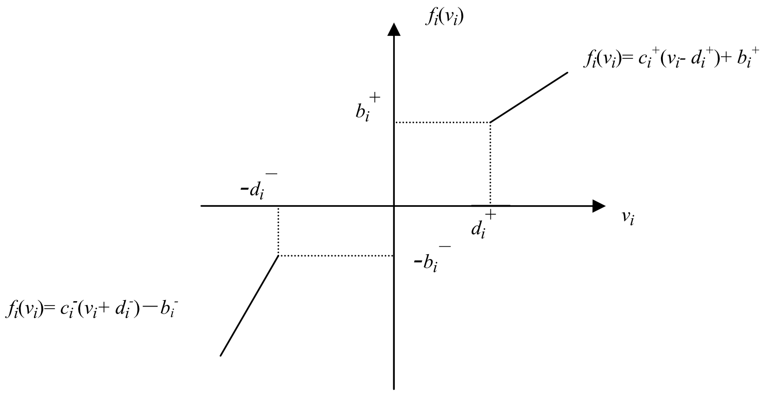

Fig. 2 Nonlinear block

that is not measurable and then becomes the input signals of the nonlinearity of the Wiener model (each dimension goes into one nonlinear block).

$$Y(k) = \left[y_1(k), y_2(k), \dots, y_n(k)\right]^T$$

is the output of the whole system which is measurable.

Assume that the difference equation of the linear subsystem is given as:

$$V(k) = B(z)U(k), \tag{1}$$

where

$$B(z) = I + B_1 z^{-1} + \dots + B_q z^{-1},$$

$$B_j = \begin{bmatrix} b_{j11} & b_{j12} & \dots & b_{j1n} \\ \dots & \dots & \dots & \dots \\ b_{jn1} & b_{jn2} & \dots & b_{jnn} \end{bmatrix} \ (j = 1, 2, \dots, q),$$

and I is the n-order identity matrix, z is the unit delay operator, and q is the upper boundary of the real order. The regressor vector and the unknown coefficient matrix of the linear subsystem are expressed respectively as

$$\phi(k) = \left[U^T(k-1), U^T(k-2), \dots, U^T(k-q)\right]^T,$$

and

$$\Theta^T = [B_1, B_2, \dots, B_q] = [\theta_1, \theta_2, \dots, \theta_n]^T,$$

where for $i = 1, 2, \dots, n$,

$$\theta_i = [b_{1i1}, \dots, b_{1in}, b_{2i1}, \dots, b_{2in}, \dots, b_{qi1}, \dots, b_{qin}]^T.$$

Then (1) can be rewritten as:

$$V(k) = U(k) + \Theta^T \phi(k).$$

Furthermore, the nonlinear blocks are assumed to be characterized by the same form of piece-wise linear function; the only dissimilarities are the different values of the unknown parameters. Following is the ith nonlinear block which is shown in Fig. 2.

$$f_i\big(v_i(k)\big) = \begin{cases} c_i^+(v_i(k) - d_i^+) + b_i^+, \\ \quad v_i(k) > d_i^+, c_i^+ \ge 0, \\ 0, \quad -d_i^- \le v_i(k) \le d_i^+, \\ c_i^-(v_i(k) + d_i^-) - b_i^-, \\ \quad v_i(k) < -d_i^-, c_i^- \ge 0, \end{cases} \tag{2}$$

where $b_i^+, b_i^-, c_i^+, c_i^-$, and d_i^+, d_i^- are the corresponding preloads, slopes, and dead zones, which need to be estimated. It is worth mentioning that these are different from the ones in [26], in which the preloads and dead zones are symmetric, which leading to only four unknown parameters.

Real systems are always subject to varying degrees of random interference [1, 29], which is not considered in [27]. Generally speaking, in our studies an adaptive output noise vector

$$e(k) = \left[e_1(k), e_2(k), \dots, e_n(k)\right]^T$$

can substitute for this interference, as shown in Fig. 1.

If we denote the output of the nonlinear subsystem by

$$F\big(V(k)\big) = \left[f_1\big(v_1(k)\big), f_2\big(v_2(k)\big), \dots, f_n\big(v(k)\big)\right]^T,$$

the output of the Wiener model can be written as:

$$Y(k) = F\big(V(k)\big) + e(k). \tag{3}$$

For each dimension, the relationship of the input-output signals is:

$$v_i(k) = u_i(k) + \theta_i^T \phi(k), \\ y_i(k) = f_i\big(v_i(k)\big) + e_i(k). \tag{4}$$

From the equations above, we can see that in order to identify this kind of Wiener model, the problem that needs to be solved is to choose a suitable kind of input vector and estimate the unknown parameters in both the linear (Θ) and nonlinear subsystems ($b_i^+, b_i^-, c_i^+, c_i^-, d_i^+, d_i^-$) on the basis of the measurement of the input-output vectors.

3 Assumptions

In addition to the assumptions about the structure of the model, we should also assume the following conditions:

(A.1) The input vector $U(k)$ is an n-dimensional iid normal random vector, and $U(k) \sim (\mathbf{0}, I)$;
(A.2) For $i = 1, 2, \dots, n$ (except where otherwise specified, in this paper, $i = 1, 2, \dots, n$), $b_i^+ > m_b$, and $b_i^- > m_b$, where $m_b > 0$;
(A.3) Noise vector $e(k)$ is an n-dimensional iid random vector, and the mathematical expectation and variance of each dimension are 0 and σ_{e_i}, respectively, and $|e_i(k)| < m_e$, where m_e is known and $0 < m_e \le m_b/2$.

4 Estimation of nonlinearity

By the characteristics of the normal random vector [30] we know that under (A.1),

$$v_i(k) = u_i(k) + \theta_i^T \phi(k)$$

must be Gaussian stationary and ergodic, and its mathematical expectation, variance, and marginal density function should be:

$$\mu_{v_i} = 0, \qquad \sigma_{v_i} = 1 + \|\theta_i\|^2, \\ p_i(v_i) = \frac{1}{\sqrt{2\pi}\sigma_{v_i}} e^{-\frac{v_i^2}{2\sigma_{v_i}^2}}. \tag{5}$$

For the sake of convenience, we predefine several interim variables that will be estimated before the unknown parameters of the model:

$$\alpha_i^+ = \frac{d_i^+}{\sigma_{v_i}}, \qquad \alpha_i^- = \frac{d_i^-}{\sigma_{v_i}}, \\ \beta_i^+ = c_i^+ \sigma_{v_i}, \qquad \beta_i^- = c_i^- \sigma_{v_i}, \\ h_i^+ = c_i^+ d_i^+ - b_i^+, \qquad h_i^- = c_i^- d_i^- - b_i^-. \tag{6}$$

4.1 Estimation of interim variables

To estimate α_i^+ and α_i^- let us define

$$g_i^+(k) = (1-k) g_i^+(k-1) + \frac{1}{k} I_{[y_i(k) > m_e]}, \\ g_i^-(k) = (1-k) g_i^-(k-1) + \frac{1}{k} I_{[y_i(k) < -m_e]}. \tag{7}$$

It is clear that with the arbitrary initial values $g_i^+(0)$ and $g_i^-(0)$, $g_i^+(k)$, and $g_i^-(k)$ can be obtained. Then according to the characteristics of the normal random vector we can figure out that

$$g_i^+(k) = 1 - \Phi\big(\alpha_i^+(k)\big), \\ g_i^-(k) = \Phi\big(-\alpha_i^-(k)\big). \tag{8}$$

Lemma 1 *For the system described by (1)–(4), $y_i(k) > m_e$ is equivalent to $v_i(k) > d_i^+$, and similarly, $y_i(k) < -m_e$ is equivalent to $v_i(k) < -d_i^-$, where $i = 1, 2, \dots, n$.*

Proof On the one hand, supposing $v_i(k) > d_i^+$, by the definition of $f(\cdot)$ and Fig. 1 together with (A.2), we can see that it must be correct that

$$f_i\big(v_i(k)\big) > b^+ > m_b.$$

Considering (A.3), we know that $|e_i| < m_e$, and then it is clear that

$$y_i(k) = f_i\big(v_i(k)\big) + e_i(k) > m_b + (-m_e) \\ = m_b - m_e > m_e.$$

On the other hand, let us assume that $y_i(k) > m_e$. Since $c^+ \ge 0$ and $c^- \ge 0$, which is to say, in the intervals $(-\infty, -d_i^-)$ and (d_i^+, ∞), respectively, $f(\cdot)$ is an increasing function, so it is inevitably right that $v_i(k) > d_i^+$. Because if not, we should have

$$y_i(k) = f_i\big(v_i(k)\big) + e_i(k) < 0 + m_e = m_e,$$

which contradicts the assumption $y_i(k) > m_e$.

Through these two aspects, we can conclude that $y_i(k) > m_e$ is equivalent to $v_i(k) > d_i^+$. In the same way it can be obtained that $y_i(k) < -m_e$ is equivalent to $v_i(k) < -d_i^-$. □

Lemma 2 *For the system described by* (1)–(4)*, if* (A.1)–(A.3) *are satisfied, when* $k \to \infty$*, we have*

$$\alpha_i^+(k) \to \alpha_i^+, \qquad \alpha_i^-(k) \to \alpha_i^-, \tag{9}$$

where $\alpha_i^+(k)$ *and* $\alpha_i^-(k)$ *are calculated by* (8).

Proof As mentioned earlier, $V(k)$ is a Gaussian stationary and ergodic normal random vector, and the noise vector $e(k)$ is also ergodic. Therefore, the sum of the two vectors which is the output of the whole system $Y(k)$, is ergodic also. Then we have

$$\begin{aligned} g_i^+(k) &= (1-k)g_i^+(k-1) + \frac{1}{k} I_{[y_i(k)>m_e]} \\ &= \frac{1}{k}\sum_{l=1}^{k} I_{[y_i(l)>m_e]} \overset{k\to\infty}{\longrightarrow} E I_{[y_i(1)>m_e]}, \\ g_i^-(k) &= (1-k)g_i^-(k-1) + \frac{1}{k} I_{[y_i(k)<-m_e]} \\ &= \frac{1}{k}\sum_{l=1}^{k} I_{[y_i(l)<-m_e]} \overset{k\to\infty}{\longrightarrow} E I_{[y_i(1)<-m_e]}. \end{aligned} \tag{10}$$

Furthermore, from Lemma 1, we know that $y_i(1) > m_e$ is equivalent to $v_i(1) > d_i^+$, and $y_i(1) < -m_e$ is equivalent to $v_i(1) < -d_i^-$. Then according to the characteristics of the normal random vector, we have:

$$\begin{aligned} E I_{[y_i(1)>m_e]} &= P\big(y_i(1) > m_e\big) = P\big(v_i(1) > d_i^+\big) \\ &= \int_{d_i^+}^{+\infty} p_i(v_i)\,dv_i = 1 - \Phi(\alpha_i^+), \\ E I_{[y_i(1)<-m_e]} &= P\big(y_i(1) < -m_e\big) \\ &= P\big(v_i(1) < -d_i^-\big) \\ &= \int_{-\infty}^{d_i^-} p_i(v_i)\,dv_i = \Phi(-\alpha_i^-). \end{aligned} \tag{11}$$

Since $\Phi(x)$ is continuous and increasing, then from Lemma 2 and (10) and (11), we know that

$$\alpha_i^+(k) \overset{k\to\infty}{\longrightarrow} \alpha_i^+, \qquad \alpha_i^-(k) \overset{k\to\infty}{\longrightarrow} \alpha_i^- \quad \text{a.s.}$$

Let us continue to estimate for β_i^+, β_i^-, h_i^+, and h_i^-. We need to define:

$$\begin{aligned} \bar{y}_i^+(k) &= (1-k)\bar{y}_i^+(k-1) + \frac{1}{k} y_i(k) I_{[y_i(k)>m_e]}, \\ \bar{y}_i^-(k) &= (1-k)\bar{y}_i^-(k-1) + \frac{1}{k} y_i(k) I_{[y_i(k)<-m_e]}, \\ \underline{y}_i^+(k) &= (1-k)\underline{y}_i^+(k-1) + \frac{1}{k} y_i^2(k) I_{[y_i(k)>m_e]}, \\ \underline{y}_i^-(k) &= (1-k)\underline{y}_i^-(k-1) + \frac{1}{k} y_i^2(k) I_{[y_i(k)<-m_e]}. \end{aligned} \tag{12}$$

Similarly, with arbitrary initial values, by solving the following equations in (13), β_i^+, β_i^-, h_i^+, and h_i^- can be obtained as shown in (14):

$$\begin{aligned} \bar{y}_i^+(k) &= \frac{\beta_i^+(k)}{\sqrt{2\pi}} e^{-\frac{(\alpha_i^+(k))^2}{2}} - h_i^+(k) g_i^+(k), \\ \underline{y}_i^+(k) &= \big(\beta_i^+(k)\big)^2 \left(\frac{\alpha_i^+(k)}{\sqrt{2\pi}} e^{-\frac{(\alpha_i^+(k))^2}{2}} + g_i^+(k)\right) \\ &\quad - \frac{2}{\sqrt{2\pi}} \beta_i^+(k) h_i^+(k) e^{-\frac{(\alpha_i^+(k))^2}{2}} \\ &\quad + \big(\big(h_i^+(k)\big)^2 + \sigma_{ei}^2\big) g_i^+(k), \\ \bar{y}_i^-(k) &= \frac{\beta_i^-(k)}{\sqrt{2\pi}} e^{-\frac{(\alpha_i^-(k))^2}{2}} - h_i^-(k) g_i^-(k), \\ \underline{y}_i^-(k) &= \big(\beta_i^-(k)\big)^2 \left(\frac{\alpha_i^-(k)}{\sqrt{2\pi}} e^{-\frac{(\alpha_i^-(k))^2}{2}} + g_i^-(k)\right) \\ &\quad - \frac{2}{\sqrt{2\pi}} \beta_i^-(k) h_i^-(k) e^{-\frac{(\alpha_i^-(k))^2}{2}} \\ &\quad + \big(\big(h_i^-(k)\big)^2 + \sigma_{ei}^2\big) g_i^-(k), \end{aligned} \tag{13}$$

$$\begin{aligned} h_i^+(k) &= \frac{1}{g_i^+(k)} \big(\beta_i^+(k)\gamma_i^+(k) - \bar{y}_i^+(k)\big), \\ \beta_i^+(k) &= \sqrt{\frac{\underline{y}_i^+(k) - (\bar{y}_i^+(k))^2/g_i^+(k) - \sigma_{ei}^2 g_i^+(k)}{\alpha_i^+(k)\gamma_i^+(k) + g_i^+(k) - (\gamma_i^+(k))^2/g_i^+(k)}}, \\ h_i^-(k) &= \frac{1}{g_i^-(k)} \big(\beta_i^-(k)\gamma_i^-(k) - \bar{y}_i^-(k)\big), \\ \beta_i^-(k) &= \sqrt{\frac{\underline{y}_i^-(k) - (\bar{y}_i^-(k))^2/g_i^-(k) - \sigma_{ei}^2 g_i^-(k)}{\alpha_i^-(k)\gamma_i^-(k) + g_i^-(k) - (\gamma_i^-(k))^2/g_i^-(k)}}, \end{aligned} \tag{14}$$

where

$$\gamma_i^+(k) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\alpha_i^+(k))^2}{2}},$$

$$\gamma_i^-(k) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(\alpha_i^-(k))^2}{2}}. \qquad \square$$

Lemma 3 *Under the conditions of Lemma 2, when $k \to \infty$, $\beta_i^+(k)$, $h_i^+(k)$ and $\beta_i^-(k)$, $h_i^-(k)$ calculated by (14) have the following characteristic*:

$$h_i^+(k) \to h_i^+, \qquad \beta_i^+(k) \to \beta_i^+,$$
$$h_i^-(k) \to h_i^-, \qquad \beta_i^-(k) \to \beta_i^- \quad a.s.$$

Proof Since we have determined that $Y(k)$ is ergodic, from the definition of (12), we have:

$$\begin{aligned} \bar{y}_i^+(k) &= \frac{1}{k}\sum_{l=1}^{k} y_i(l) I_{[y_i(l)>m_e]} \\ &\overset{k\to\infty}{\longrightarrow} E y_i(1) I_{[y_i(1)>m_e]} \quad \text{a.s.}, \\ \underline{y}_i^+(k) &= \frac{1}{k}\sum_{l=1}^{k} y_i^2(l) I_{[y_i(l)>m_e]} \\ &\overset{k\to\infty}{\longrightarrow} E y_i^2(1) I_{[y_i(1)>m_e]} \quad \text{a.s.} \end{aligned} \tag{15}$$

Note that $e(k)$ and $V(k)$ are independent of each other, so we have:

$$\begin{aligned} &E y_i(1) I_{[y_i(1)>m_e]} \\ &\quad = E\big(c_i^+ v_i(1) - h_i^+ + e_i(1)\big) I_{[v_i(1)>d^+]} \\ &\quad = \frac{\beta_i^+}{\sqrt{2\pi}} e^{-\frac{(\alpha_i^+)^2}{2}} - h_i^+\big(1-\Phi(\alpha_i^+)\big), \\ &E y_i^2(1) I_{[y_i(1)>m_e]} \\ &\quad = E\big(c_i^+ v_i(1) - h_i^+ + e_i(1)\big)^2 I_{[v_i(1)>d^+]} \\ &\quad = \big(\beta_i^+\big)^2 \left[\frac{\alpha_i^+ e^{-\frac{(\alpha_i^+)^2}{2}}}{\sqrt{2\pi}} + \big(1-\Phi(\alpha_i^+)\big)\right] \\ &\qquad - \frac{2}{\sqrt{2\pi}} \beta_i^+ h_i^+ e^{-\frac{(\alpha_i^+)^2}{2}} \\ &\qquad + \big((h_i^+)^2 + \sigma_{e_i}^2\big)\big(1-\Phi(\alpha_i^+)\big). \end{aligned} \tag{16}$$

Then from (15) and (16) and Lemma 2, it is clear that

$$\beta_i^+(k) \overset{k\to\infty}{\longrightarrow} \beta_i^+, \qquad h_i^+(k) \overset{k\to\infty}{\longrightarrow} h_i^+ \quad \text{a.s.}$$

In the same way we can obtain

$$\beta_i^-(k) \overset{k\to\infty}{\longrightarrow} \beta_i^-, \qquad h_i^-(k) \overset{k\to\infty}{\longrightarrow} h_i^- \quad \text{a.s.} \qquad \square$$

4.2 Estimation of b_i^+ and b_i^-

Now we have completed the estimation for all the interim variables. Then from (7), it is clear that

$$\begin{aligned} b_i^+(k) &= \alpha_i^+(k)\beta_i^+(k) - h_i^+(k), \\ b_i^-(k) &= \alpha_i^-(k)\beta_i^-(k) - h_i^-(k). \end{aligned} \tag{17}$$

Theorem 1 *According to Lemmas 2 and 3, it is clear that $b_i^+(k)$ and $b_i^-(k)$ calculated by (17) are the consistent estimation of b_i^+ and b_i^-, respectively; that is, to say, it is true that*

$$b_i^+(k) \overset{k\to\infty}{\longrightarrow} b_i^+, \qquad b_i^-(k) \overset{k\to\infty}{\longrightarrow} b_i^- \quad a.s.$$

(The proof is omitted.)

4.3 Estimation of σ_{v_i}

Since we have obtained the consistent estimation of all the interim variables, to estimate the unknown parameters of the nonlinearity we only need to know σ_{v_i}. So, next let us estimate it with the help of the kernel function described in [27]. First, we define the kernel function as follows:

$$\omega_i(k) = k^{2\varepsilon} e^{-k^{4\varepsilon} u_i^2(k)}, \quad \varepsilon \in \left(0, \frac{1}{4}\right). \tag{18}$$

Define

$$\begin{aligned} G_i(k) &= \frac{1}{k}\sum_{l=1}^{k} \omega_i(l) y_i(l) I_{[y_i(l)>m_e]} \\ &= (1-k)G_i(k-1) \\ &\quad + \frac{1}{k}\omega_i(k) y_i(k) I_{[y_i(k)>m_e]}. \end{aligned} \tag{19}$$

Lemma 4 *Under the conditions of Lemma 2, we have*

$$\begin{aligned} \lim_{k\to\infty} G_i(k) &= \frac{-h_i^+}{\sqrt{2}}\left(1-\Phi\left(\frac{d_i^+}{\|\theta_i\|}\right)\right) \\ &\quad + \frac{c_i^+\|\theta_i\|}{2\sqrt{\pi}} e^{-\frac{1}{2}(\frac{d_i^+}{\|\theta_i\|})^2} \quad a.s. \end{aligned} \tag{20}$$

Proof Denoting the right hand of (20) by G_i, by Lemma 1 and (2), (4), and (18), we have

$$\omega_i(k)y_i(k)I_{[y_i(k)>m_e]} = k^{2\varepsilon}e^{-k^{4\varepsilon}u_i^2(k)}\big[c_i^+u_i(k)+c_i^+\theta_i^T\phi_k - h_i^+ + e_i(k)\big]I_{[v_i(k)>d^+]}. \quad (21)$$

Since

$$\frac{1}{k}\sum_{l=1}^{k} l^{2\varepsilon}e^{-l^{4\varepsilon}u_i^2(l)}\big[c_i^+u_i(l)+c_i^+\theta_i^T\phi_k - h_i^+\big]I_{[v_i(l)>d^+]} \xrightarrow{k\to\infty} G_i \quad \text{a.s.} \quad (22)$$

which has been proven in the Appendix of [27], then we only need to prove that

$$\frac{1}{k}\sum_{l=1}^{k} l^{2\varepsilon}e^{-l^{4\varepsilon}u_i^2(l)}e_i(l)I_{[v_i(l)>d^+]} \xrightarrow{k\to\infty} 0 \quad \text{a.s.} \quad (23)$$

Suppose that Φ_k is generated by the σ-algebra $\{u_1, u_2, \dots, u_k, e_1, e_2, \dots, e_k\}$. Since

$$\begin{aligned}&\sup_k E\big[\big(k^{\varepsilon}e^{-k^{4\varepsilon}u_i^2(k)}e_i(k)\big)^2 I_{[u_i(k)>d^+-\theta_i^T\phi_k]} \mid F_{k-1}\big]\\ &= \frac{k^{2\varepsilon}}{2\pi\sigma_{e_i}}\int_{d_i^+-\theta_i^T\phi_k}^{+\infty}\int_{-\infty}^{+\infty} e^{-k^{4\varepsilon}x^2}e^{-\frac{x^2}{2}}e^{-\frac{y^2}{2\sigma_{e_i}^2}}y\,dx\,dy\\ &\le \frac{1}{4\pi\sigma_{e_i}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{s^2}{2}}e^{-\frac{k^{-4\varepsilon}s^2}{4}}e^{-\frac{y^2}{2\sigma_{e_i}^2}}y\,ds\,dy\\ &< \infty \end{aligned} \quad (24)$$

and $\sum_{k=1}^{\infty}(1/k^{1-\varepsilon})^2 < \infty$, then by the convergence theorem for Martingale difference sequences [31], we have

$$\begin{aligned}&\sum_{k=1}^{\infty}\frac{1}{k^{1-\varepsilon}}\Big\{k^{\varepsilon}e^{-k^{4\varepsilon}u_i^2(k)}e_i(k)I_{[u_i(k)>d_i^+-\theta_i^T\phi_k]}\\ &\quad - E\big[k^{\varepsilon}e^{-k^{4\varepsilon}u_i^2(k)}e_i(k)I_{[u_i(k)>d^+-\theta_i^T\phi_k]} \mid F_{k-1}\big]\Big\}\\ &< \infty.\end{aligned}$$

Thus, by the Kronecker Lemma [31], we know

$$\begin{aligned}&\frac{1}{n}\sum_{k=1}^{n}k^{2\varepsilon}\Big\{e^{-k^{4\varepsilon}u_i^2(k)}e_i(k)I_{[u_i(k)>d_i^+-\theta_i^T\phi_k]}\\ &\quad - E\big[e^{-k^{4\varepsilon}u_i^2(k)}e_i(k)I_{[u_i(k)>d^+-\theta_i^T\phi_k]} \mid F_{k-1}\big]\Big\}\\ &\xrightarrow{k\to\infty} 0 \quad \text{a.s.}\end{aligned} \quad (25)$$

Noting that $e_i(k)$ is independent of $\theta_i^T\phi(k)$ and $u_i(k)$, we have

$$E\big[k^{2\varepsilon}e^{-k^{4\varepsilon}u_i^2(k)}e_i(k)I_{[u_i(k)>d^+-\theta_i^T\phi_k]} \mid F_{k-1}\big] \xrightarrow{k\to\infty} 0, \quad (26)$$

so by (25) and (26), (23) can be obtained, then with (22), (20) follows.

Then according to Lemmas 2–4, we know that $G_i(k)$ converges to the G_i defined in (20), and if we can find the root $x(k)$ of (27) with respect to x, it surely converges to $\sqrt{1/\|\theta_i\|^2+1}$, so the consistent estimation of $\|\theta_i\|$ and $\sigma_{v_i}(k) = \sqrt{1+\|\theta_i\|^2}$ can be obtained in succession.

$$G_i(k) + \frac{h_i^+}{\sqrt{2}}\big(1-\Phi\big(\alpha_i^+(k)\big)x\big) - \frac{\beta_i^+(k)}{2\sqrt{\pi}x}e^{-\frac{(\alpha_i^+(k))^2x^2}{2}} = 0. \quad (27)$$

The existence and uniqueness of the solution $x(k)$ has been given in [27], and we can use a numerical method to find the solution $x(k)$. □

4.4 Estimation of c_i^+, c_i^- and d_i^+, d_i^-

Now that we have consistently estimated σ_{v_i}, it is easy to calculate the values of $c_i^+(k)$, $c_i^-(k)$ and $d_i^+(k)$, $d_i^-(k)$ with the estimation of the interim variables:

$$\begin{aligned} c_i^+(k) &= \frac{\beta_i^+(k)}{\sigma_{v_i}}, \qquad d_i^+(k) = \alpha_i^+(k)\sigma_{v_i},\\ c_i^-(k) &= \frac{\beta_i^-(k)}{\sigma_{v_i}}, \qquad d_i^-(k) = \alpha_i^-(k)\sigma_{v_i}. \end{aligned} \quad (28)$$

Theorem 2 *By Lemmas 2–4, we can easily conclude that under the conditions of Lemma 2, $c_i^+(k)$, $c_i^-(k)$, $d_i^+(k)$, $d_i^-(k)$ calculated by (28) are the consistent estimation of c_i^+, c_i^-, d_i^+, d_i^-, respectively.*

(The proof is omitted.)

5 Estimation of linear subsystem

At this point, we have obtained all the unknown parameters in the nonlinearity; now let us continue by estimating the linear subsystem using the least squares algorithm.

Since $V(k)$ can not be measured directly, we must calculate it with the help of the output $Y(k)$ and the estimation of the nonlinearity. As in [27], define

$$\hat{v}_i(k) = \begin{cases} \frac{1}{\bar{c}_i^+(k)}(h_i^+(k) + y_i(k)), & y_i(k) > m_e, \\ 0, & -m_e \le y_i(k) \le m_e, \\ \frac{1}{\bar{c}_i^-(k)}(y_i(k) - h_i^-(k)), & y_i(k) < -m_e \end{cases} \tag{29}$$

and

$$\begin{aligned} z_i(k) &= \big(\hat{v}_i(k) - u_i(k)\big) I_{[y_i(k)>m_e \cup y_i(k)<-m_e]} \\ &= \big[\theta_i^T \phi(k) + \varepsilon_i(k)\big] I_{[y_i(k)>m_e \cup y_i(k)<-m_e]} \\ &= \theta_i^T \hat{\phi}(k) + \varepsilon_i(k) I_{[y_i(k)>m_e \cup y_i(k)<-m_e]}, \end{aligned} \tag{30}$$

where

$$\bar{c}_i^+(k) = c_i^+(k) \vee 1/k, \qquad \bar{c}_i^-(k) = c_i^-(k) \vee 1/k$$

are the modifications of $c_i^+(k)$ and $c_i^-(k)$ and have the same limits as them, respectively, while

$$\hat{\phi}(k) = \phi(k) I_{[y_i(k)>m_e \cup y_i(k)<-m_e]},$$
$$\varepsilon_i(k) = \hat{v}_i(k) - v_i(k).$$

With an arbitrary initial $\theta_i(0)$ and $P(0) > 0$, the coefficient vector θ_i can be estimated as:

$$\begin{aligned} \theta_i(k) &= \theta_i(k-1) + a(k)P(k)\hat{\phi}(k) \\ &\quad \times \big(z_i(k) - \theta_i^T(k-1)\hat{\phi}(k)\big), \\ P(k+1) &= P(k) - a(k)P(k)\hat{\phi}(k)\hat{\phi}^T(k)P(k), \\ a(k) &= (1 - \hat{\phi}^T(k)P(k)\hat{\phi}(k))^{-1}. \end{aligned} \tag{31}$$

6 Simulation

To illustrate the validity of the method described above, simulation trials were conducted with the help of Matlab, and in the following, we take a simple one for example. Considering a 2-dimensional Wiener system:

$$V(k) = U(k) + \Theta^T \phi(k) = U(k) + [\theta_1, \theta_2]^T \begin{bmatrix} U(k-1) \\ U(k-2) \\ U(k-3) \\ U(k-4) \end{bmatrix},$$

where $U(k) = [u_1(k), u_2(k)]^T$,

$$\begin{aligned} \theta_1 &= [b_{111} \quad b_{112} \quad b_{211} \quad b_{212} \quad b_{311} \quad b_{312} \quad b_{411} \quad b_{412}] \\ &= [0.95 \quad 0.80 \quad 0.50 \quad 0.50 \quad 0 \quad -0.10 \quad -0.45 \quad -0.60]^T, \\ \theta_2 &= [b_{121} \quad b_{122} \quad b_{221} \quad b_{222} \quad b_{321} \quad b_{322} \quad b_{421} \quad b_{422}] \\ &= [0.85 \quad 0.75 \quad 0.45 \quad 0.20 \quad -0.15 \quad -0.35 \quad -0.30 \quad -0.80]^T, \end{aligned}$$

and

$$Y(k) = F\big(V(k)\big) + e(k) = \begin{bmatrix} f_1(v_1(k)) + e_1(k) \\ f_2(v_2(k)) + e_2(k) \end{bmatrix},$$

where for $i = 1, 2$,

$$\begin{aligned} f_i\big(v_i(k)\big) &= \begin{cases} c_i^+(v_i(k) - d_i^+) + b_i^+, & v_i(k) > d_i^+, c_i^+ \ge 0, \\ 0, & -d_i^- \le v_i(k) \le d_i^+, \\ c_i^-(v_i(k) + d_i^-) - b_i^-, & v_i(k) < -d_i^-, c_i^- \ge 0 \end{cases} \\ &= \begin{cases} 1 \times (v_i(k) - 2) + 5, & v_i(k) > 2, \\ 0, & -(-1.75) \le v_i(k) \le 2, \\ 3.6(v_i(k) + (-1.75)) - 1.36, & v_i(k) < -(-1.75). \end{cases} \end{aligned}$$

That is to say the nonlinear blocks have the same parameters.

For the noises, we let $\sigma_{e_i} = 1, m_e = 0.5$, which are mentioned in (A.3). With regard to the ε used in (18), let us choose the same value used in [27], that is, 1/13500.

The estimation results for this system are shown in Figs. 3–9, where the solid lines are the real values and the dotted lines are the estimated values. Figures 3–6 show the simulation results of the linear subsystem. To appear more clearly, each θ_i is shown in two figures. Figure 3 and Fig. 4 show the estimates for θ_1, where the true values are $(0.95, 0.50, 0, -0.45)^T$ and $(0.80, 0.50, -0.10, -0.60)^T$, respectively; Fig. 5 and Fig. 6 show the estimates for θ_2, where the true values are $(0.85, 0.45, -0.15, -0.30)^T$ and $(0.75, 0.20, -0.35, -0.80)^T$, respectively. The experiment sample size is 2,000. We can see that the estimates are all

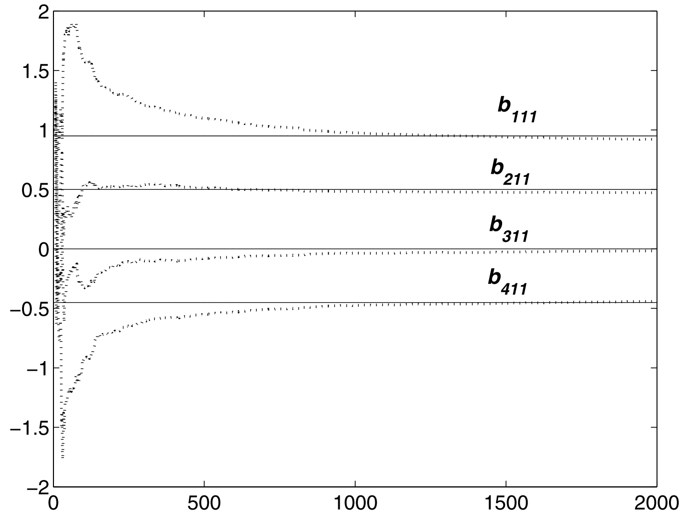

Fig. 3 Estimates for part of θ_1: $b_{111} \sim b_{411}$

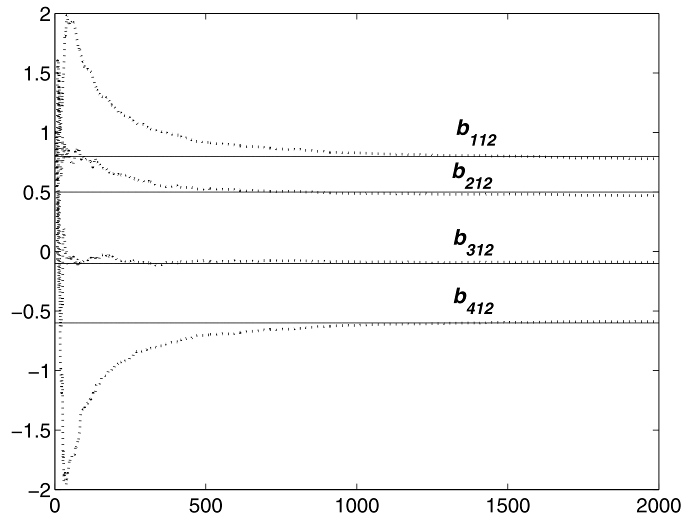

Fig. 4 Estimates for the rest of θ_1: $b_{112} \sim b_{412}$

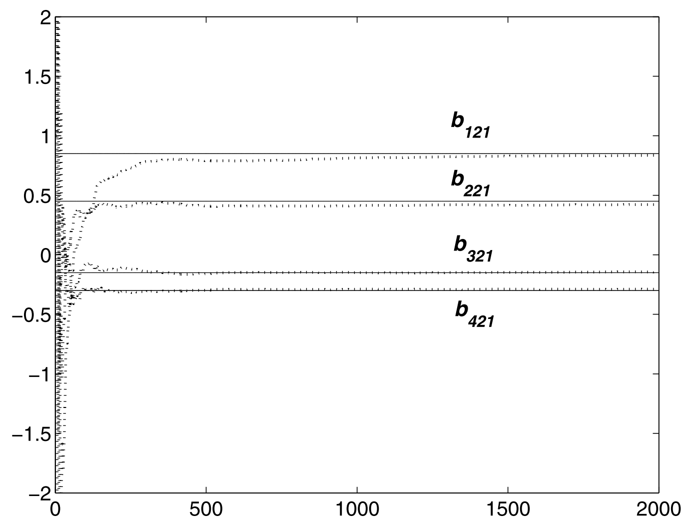

Fig. 5 Estimates for part of θ_2: $b_{121} \sim b_{421}$

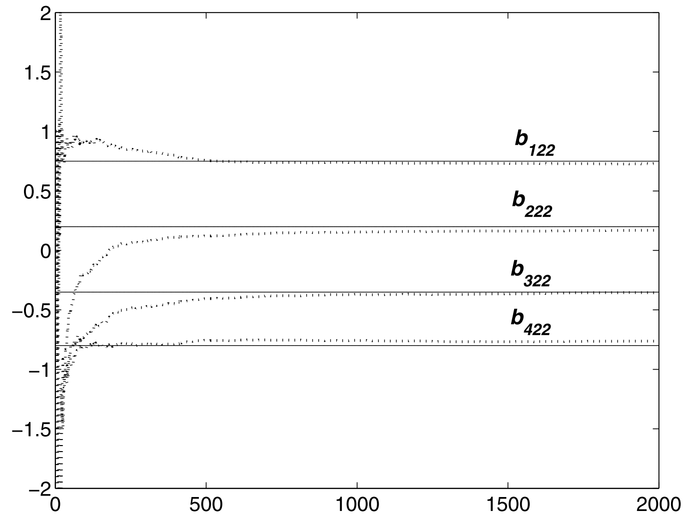

Fig. 6 Estimates for the rest of θ_2: $b_{122} \sim b_{422}$

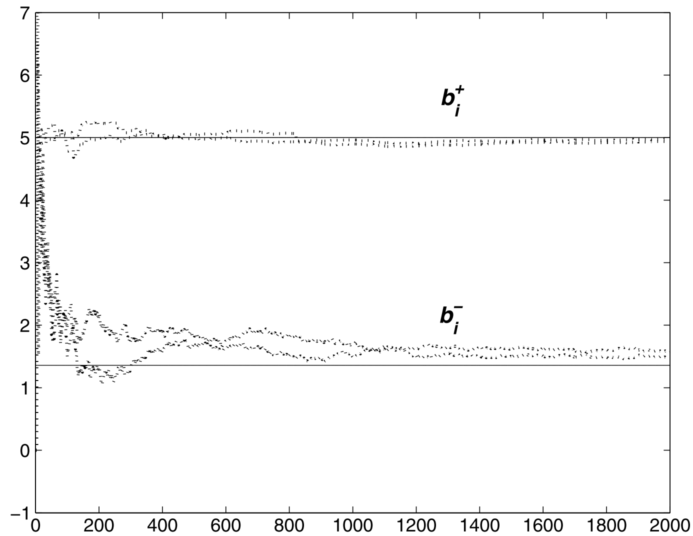

Fig. 7 Estimates for preloads in nonlinearity: b_i^+ and b_i^-

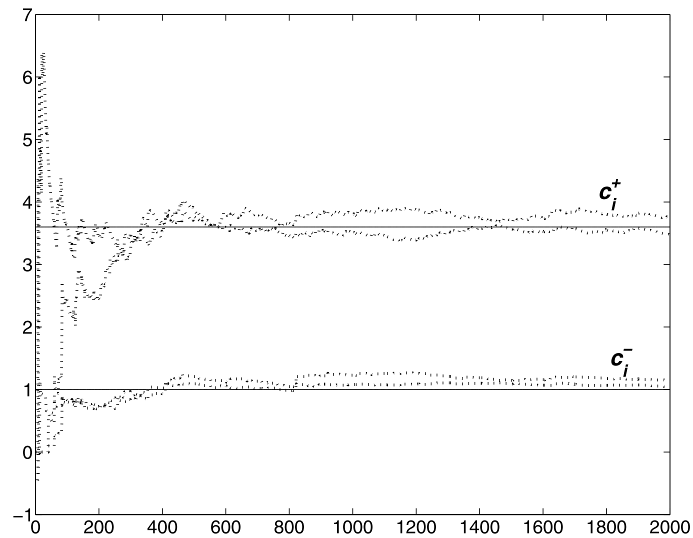

Fig. 8 Estimates for slopes in nonlinearity: c_i^+ and c_i^-

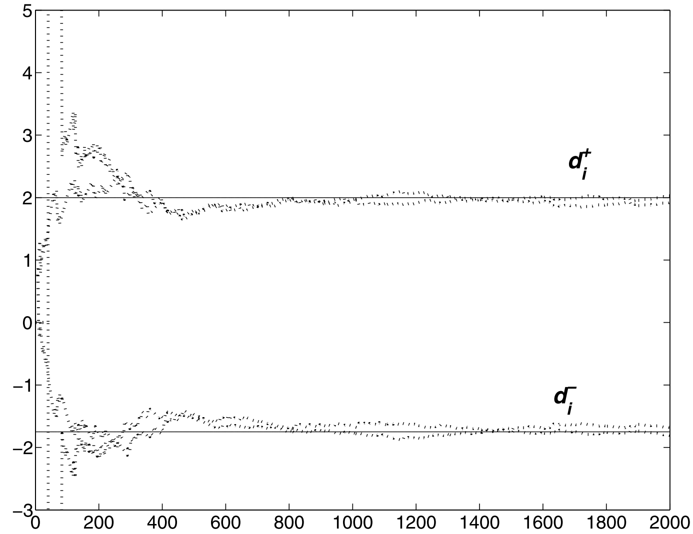

Fig. 9 Estimates for dead zones in nonlinearity: d_i^+ and d_i^-

perfect after about 1,000 samples. The estimations of the nonlinear blocks (the preloads, slopes, and dead zones) are shown in Figs. 7–9. From these figures, we can see clearly that the estimated values in both the nonlinear and linear subsystems all converge to their corresponding real values fast. It confirms the validity of this recursive algorithm for the estimation of MIMO Wiener systems in noisy environments.

7 Conclusion

Based on [27], this paper proposes a recursive estimation algorithm which proved to have good results for the identification of MIMO Wiener systems with the nonlinearity being discontinuous asymmetric piecewise linear functions in noisy environments. Although the system we studied had the same dimensions in the input and output terminals, but this method can also be applied to systems with different input and output dimensions. However, there is a strong restriction on the noise signals. From (A.3), they are required Gaussian and bounded, which is unrealistic. This problem will be studied in the future research work.

References

1. Greblicki, W.: Nonparametric approach to Wiener system identification. IEEE Trans. Circuits Syst. I: Fund. Theory Appl. **44**(6), 538–545 (1997)
2. Greblicki, W.: Nonparametric identification of Wiener systems. IEEE Trans. Inf. Theory **38**(5), 1478–1493 (1992)
3. Greblicki, W.: Nonparametric identification of Wiener systems by orthogonal series. IEEE Trans. Autom. Control **39**(10), 2077–2086 (1994)
4. Ljung, L.: System Identification: Theory for the User. Prentice-Hall, Upper Saddle River (1999), pp. 141–143
5. Hsia, T.C.: System Identification Least Squares Methods. Lexington Books, Toronto (1977), pp. 112–113
6. Korenberg, M.J.: Parallel cascade identification and kernel estimation for nonlinear systems. Ann. Biomed. Eng. **19**, 429–455 (1991)
7. Boyd, S., Chua, L.O.: Fading memory and the problem of approximating nonlinear operators with Volterra series. IEEE Trans. Circuits Syst. **32**(11), 1150–1161 (1985)
8. Kalafatis, A.D., Arifin, N., Wang, L., Cluett, W.R.: A new approach to the identification of pH processes based on the Wiener model. Chem. Eng. Sci. **50**, 3693–3701 (1995)
9. Kalafatis, A.D., Wang, L., Cluett, W.R.: Linearizing feedforward-feedback control of pH processes based on Wiener model. J. Process Control **15**, 103–112 (2005)
10. Ralston, J.C., Zoubir, A.M., Boashash, B.: Identification of a class of time-varying nonlinear system based on the Wiener model with application to automotive engineering. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **E78-A**(9), 1192–1200 (1995)
11. Bloemen, H.H.J., Chou, C.T., Boom, T.J.J., Verdult, V., Verhaegen, M., Backx, T.C.: Wiener model identification and predictive control for dual composition control of a distillation column. J. Process Control **11**, 601–620 (2001)
12. Ling, W.M., Rivera, D.: Nonlinear black-box identification of distillation column models-design variable selection for model performance enhancement. Int. J. Appl. Math. Comput. Sci. **8**, 794–813 (1998)
13. Visala, A., Pitkanen, H., Aarne, H.: Modeling of chromatographic separation process with Wiener-MLP representation. J. Process Control **78**, 443–458 (2001)
14. Kang, H.W., Cho, Y.S., Youn, D.H.: Adaptive precompensation of Wiener systems. IEEE Trans. Signal Process. **46**, 2825–2829 (1998)

15. den Brinker, A.C.: A comparison of results from parameter estimations of impulse responses of the transient visual system. Biol. Cybern. **61**, 139–151 (1989)
16. Hunter, I.W., Korenberg, M.J.: The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. Biol. Cybern. **55**, 135–144 (1986)
17. Chen, G., Chen, Y., Ogmen, H.: Identifying chaotic systems via a Wiener-type cascade model. IEEE Control Syst. Mag. **17**, 29–36 (1997)
18. Billings, S.A., Fakhouri, S.Y.: Theory of separable processes with applications to the identification of nonlinear systems. Proc. IEE **125**, 1051–1058 (1977)
19. Billings, S.A., Fakhouri, S.Y.: Identification of systems containing linear dynamic and static nonlinear elements. Automatica **18**, 15–26 (1982)
20. Bai, E.W.: Frequency domain identification of Wiener models. Automatica **39**, 1521–1530 (2003)
21. Gomez, J.C., Baeyens, E.: Identification of block-oriented nonlinear systems using orthonormal bases. J. Process Control **14**, 685–697 (2004)
22. Greblicki, W.: Recursive identification of Wiener systems. Int. J. Appl. Math. Comput. Sci. **11**, 977–991 (2001)
23. Chan, K.H., Bao, J., Whiten, W.J.: A new approach to control of MIMO processes with static nonlinearities using an extended IMC framework. Comput. Chem. Eng. **30**, 329–342 (2005)
24. Verhaegen, M., Westwick, D.: Identifying MIMO Wiener systems using subspace model identification methods. Signal Process. **52**, 235–258 (1996)
25. Janczak, A.: Instrumental variables approach to identification of a class of MIMO Wiener systems. Nonlinear Dyn. **48**, 275–284 (2007)
26. Vörös, J.: Parameter identification of Wiener systems with discontinuous nonlinearities. Syst. Control Lett. **44**(5), 363–372 (2001)
27. Chen, H.F.: Recursive identification for Wiener model with discontinuous piece-wise linear function. IEEE Trans. Autom. Control **51**(3), 390–400 (2006)
28. Vörös, J.: Parameter identification of Wiener systems with multisegment piecewise-linear nonlinearities. Syst. Control Lett. **56**(2), 99–105 (2007)
29. Kalafatis, A.D., Wang, L., Cluett, W.R.: Identification of Wiener-type nonlinear systems in a noisy environment. Int. J. Control **66**(6), 923–941 (1997)
30. Rotar, V.: Probability Theory. World Scientific, London (1997), pp. 135–144
31. Chen, H.F., Guo, L.: Identification and Stochastic Adaptive Control. Birkhäuser, Boston (1991), pp. 24–43