

EM algorithm-based identification of a class of nonlinear Wiener systems with missing output data

Weili Xiong · Xianqiang Yang · Liang Ke · Baoguo Xu

Received: 6 March 2014 / Accepted: 17 December 2014 / Published online: 31 December 2014
© Springer Science+Business Media Dordrecht 2014

Abstract This paper is concerned with the problem of parameter estimation for nonlinear Wiener systems in the stochastic framework. Based on the expectation–maximization (EM) algorithm in dealing with the incomplete data, it is applied to estimate the parameters of nonlinear Wiener models considering the randomly missing outputs. By means of the EM approach, the parameters and the missing outputs can be estimated simultaneously. To obtain the noise-free output in the linear subsystem of the Wiener model, the auxiliary model identification idea is adopted here. The simulation results indicate the effectiveness of the proposed approach for identification of a class of nonlinear Wiener models.

Keywords Parameter estimation · Expectation–maximization algorithm · Missing output data · Wiener model

1 Introduction

Many industrial processes have the feature of nonlinearity and dynamic nature [1,2]. Some nonlinear systems are too complicated for researchers to study their corresponding performances. System identification is to find a system model based on measured data [3,4] and is basis for signal processing, process monitoring and optimization [5,6]. So-called block-oriented models, such as Wiener and Hammerstein models, can be used to approximate many nonlinear dynamic processes and have a simple structure as well [7,8]. The Wiener model can be represented by a dynamic linear subsystem followed by a nonlinear static block. It is a reasonable model for a distillation column, a pH control process, a linear system with a nonlinear measurement device, etc. [9]. In the field of system identification, many least-squares-based identification methods and their extension versions have been developed to cope with the identification issues for Wiener systems [10–12]. Wigren [13,14] proposed a recursive prediction error identification algorithm to identify the nonlinear Wiener model, and the convergence property of the algorithm was established. Wang et al. [15,16] proposed auxiliary model-based and gradient-

This work was supported by the National Natural Science Foundation of China (Nos. 21206053, 21276111, 61273131), the PAPD of Jiangsu Higher Education Institutions and the 111 Project (B12018).

W. Xiong (✉) · L. Ke · B. Xu
Key Laboratory of Advanced Process Control for Light Industry, Ministry of Education, Jiangnan University, Wuxi 214122, People's Republic of China
e-mail:greenpre@163.com

B. Xu
e-mail:xbg@jiangnan.edu.cn

X. Yang (✉)
Research Institute of Intelligent Control and Systems, Harbin Institute of Technology, Harbin 150080, Heilongjiang, People's Republic of China
e-mail:xianqiangyang@gmail.com

based iterative identification algorithms for Wiener or Hammerstein nonlinear systems.

Xiong et al. [17] derived an iterative numerical algorithm for modeling a class of output nonlinear Wiener systems. Westwick and Verhaegen [18] extended the multivariable output-error state space subspace model identification schemes to identify Wiener systems. Hagenblad et al. [9] proposed a maximum-likelihood method with general consistency property for identification of Wiener models. Among the literatures mentioned above, most of the contributions were derived from the same assumption that the input–output measurement data are available at every sampling instant. That is to say, the measurement data set for identification are complete.

Because of the growing scale and complexity of process industry, data missing problem is commonly encountered and should be handled carefully because of its negative effects imposed on the process identification and control [19]. There are many reasons for data missing such as a sudden mechanical fault, hardware measurement failures, data transmission malfunctions and losses in network communication [20, 21]. In such cases, the standard least-squares-based identification algorithm cannot be applied to estimate the system parameters directly. Ding et al. [22] presented an auxiliary model-based least-squares algorithm and hierarchical least-squares identification algorithm to identify the parameters of dual-rate systems, which can be seen as a special case, but may not be directly applied to identification with irregularly or randomly missing data. Then, the recursive least-squares algorithm combined with an auxiliary model was derived to cope with possibly irregularly missing outputs through output-error models and convergence properties were established simultaneously [23]. They derived the parameter estimation algorithm for systems with scarce measurements which are extension from dual-rate systems through gradient-based algorithm [24]. An output-error method is used [25] to identify systems with slowly and irregularly sampled output data. It was proven that when the system is in the model set, the consistence and minimum variance property of the output-error model can be obtained.

On the other hand, some works on irregularly or randomly missing data problems under the statistical framework have been paid great attention since 1990's. Isaksson [26] studied parameter estimation of an ARX model when the measurement information

may be incomplete by using several methods including the Kalman filtering and smoothing, maximum-likelihood estimation, and a new method so-called the expectation–maximization (EM) algorithm. A simplified iteration of data reconstruction and ARX parameter estimation were proposed in [27]. Raghavan and Gopaluni et al. [28] studied the EM-based state space model identification problems with irregular output sampling and presented some simulations, laboratory-scale and industrial case studies. Xie et al. [29] proposed a new EM algorithm-based approach to estimate an FIR model for multirate processes with random delays. Because of the feature of the statistical properties and the simplicity to realize, the EM algorithm has been used in linear parameter varying (LPV) soft sensor development and nonlinear parameter varying systems with irregularly missing output data [30–32].

The objective of this paper is to handle parameter identification and output estimation problems for nonlinear Wiener systems with randomly missing output data using the EM algorithm. The auxiliary model identification idea is utilized to estimate the noise-free output iteratively in the linear dynamic subsystem and the parameter estimation and missing output estimation are handled simultaneously in the EM algorithm.

The remainder of this paper is organized as follows. Section 2 introduces the identification model of nonlinear Wiener models and the data missing patterns. In Sect. 3, the auxiliary model identification idea is used to estimate the noise-free output of the dynamic linear subsystem in the nonlinear Wiener model. Based on this idea, the identification algorithm under the framework of the EM algorithm to deal with randomly missing output data is derived. Section 4 provides an illustrative example to show the effectiveness of the proposed algorithm. Finally, we draw some conclusions in Sect. 5.

2 Problem statement

Consider the stochastic Wiener model as shown in Fig. 1 with randomly missing output data. It is composed of a linear dynamic subsystem followed by a static nonlinear block. Assume that $\{u(t), t = 1, 2, \dots, N\}$ is the input sequences of the system, $\{y(t), t = 1, 2, \dots, N\}$ is the measurable output but randomly missing with certain percentage, $e(t)$ is a white noise sequence with zero mean and variance σ^2 ,

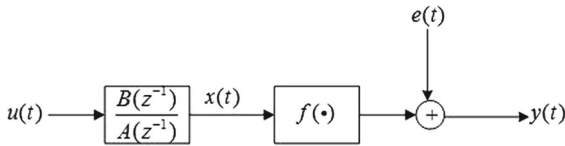


Fig. 1 The Wiener nonlinear system [23]

and $A(z^{-1})$ and $B(z^{-1})$ are polynomials in the unit backward shift operator, namely $z^{-1}y(t) = y(t - 1)$.

The linear dynamic subsystem takes the form,

$$\begin{aligned}
 x(t) &= \frac{B(z^{-1})}{A(z^{-1})}u(t) \\
 &= -a_1x(t - 1) + \dots - a_nx(t - a_n) \\
 &\quad + b_1u(t - 1) + \dots + b_nu(t - b_n) \\
 &= \varphi_p^T(t)\vartheta_p,
 \end{aligned} \tag{1}$$

where $A(z^{-1})$ and $B(z^{-1})$ are polynomials defined as

$$\begin{aligned}
 A(z^{-1}) &= 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_nz^{-a_n}, \\
 B(z^{-1}) &= b_1z^{-1} + b_2z^{-2} + \dots + b_nz^{-b_n}.
 \end{aligned} \tag{2}$$

For this class of Wiener systems, the static nonlinear block $f(\cdot)$ is generally assumed to be the sum of nonlinear basis functions based on a known basis $f = (f_1, f_2, \dots, f_n)$:

$$\begin{aligned}
 y(t) &= f(x(t)) + e(t) \\
 &= r_1f_1(x(t)) + r_2f_2(x(t)) \\
 &\quad + \dots + r_{n_r}f_{n_r}(x(t)) + e(t)
 \end{aligned} \tag{3}$$

Here, we assume that the nonlinear function $f(\cdot)$ can be represented by the polynomial with the order r :

$$f(x(t)) = r_1x(t) + r_2x^2(t) + \dots + r_{n_r}x^{n_r}(t). \tag{4}$$

As seen from Fig. 1, the linear noise-free block output $x(t)$ is the input of the nonlinear block in the nonlinear Wiener system. A direct substitution of $x(t)$ from Eqs. (1) to (4) would result in a very complex expression. Therefore, the key-term separation principle is incorporated to simplify this problem, namely the first coefficient of the nonlinear block is fixed to 1, i.e., $r_1 = 1$. Then, the system output $y(t)$ can be written as

$$\begin{aligned}
 y(t) &= x(t) + \sum_{i=2}^{n_r} r_i x^i(t) + e(t) \\
 &= \varphi_p^T(t)\vartheta_p + \varphi_r^T(t)\vartheta_r + e(t) \\
 &= \varphi^T(t)\vartheta + e(t),
 \end{aligned} \tag{5}$$

where the information vector $\varphi(t)$ includes $\varphi_p(t)$ is defined as:

$$\begin{aligned}
 \varphi_p(t) &= [-x(t - 1), -x(t - 2), \dots, -x(t - n_a) \\
 &\quad u(t - 1), u(t - 2), \dots, u(t - n_b)]^T \\
 &\quad \times \in \mathbb{R}^{n_a+n_b} \\
 \vartheta_p &= [a_1, a_2, \dots, a_{n_a} b_1, b_2, \dots, b_{n_b}]^T \\
 &\quad \times \in \mathbb{R}^{n_a+n_b} \\
 \varphi_r(t) &= [x^2(t), \dots, x^{n_r}(t)]^T \in \mathbb{R}^{n_r-1} \\
 \vartheta_r &= [r_2, \dots, r_{n_r}]^T \in \mathbb{R}^{n_r-1} \\
 \varphi(t) &= [\varphi_p^T(t), \varphi_r^T(t)]^T \in \mathbb{R}^{n_a+n_b+n_r-1} \\
 \vartheta &= [\vartheta_p, \vartheta_r]^T \in \mathbb{R}^{n_a+n_b+n_r-1}
 \end{aligned} \tag{6}$$

The missing data problem is very common in process industry. In this article, we assume that the causes for the missing outputs are unknown and believe that the occurrence of missing outputs does not depend on any input and output. This means part of the outputs is missing completely at random (MCAR) [20]. Thereafter, the data Y are divided into two parts, the randomly missing output $Y_{\text{mis}} = \{y_t\}_{t=m_1, \dots, m_\alpha}$ and the observed output sequence $Y_{\text{obs}} = \{y_t\}_{t=o_1, \dots, o_\beta}$. So, the identification problem considered under the EM framework is to estimate the parameters $\vartheta = \{\vartheta_p, \vartheta_r\}$ and the noise variance σ^2 based on the following data set:

$$C_{\text{obs}} = \{Y_{\text{obs}}, U\}, \tag{7}$$

$$C_{\text{mis}} = \{Y_{\text{mis}}\}. \tag{8}$$

3 The EM algorithm-based identification approach

3.1 The EM algorithm revisited

The EM algorithm is an ideal candidate for solving estimation problems for the maximum-likelihood estimate in the presence of missing data. The core idea behind the EM algorithm is to introduce hidden or missing variables to make the maximum-likelihood estimates

tractable [33]. The observed data set C_{obs} with missing data set C_{mis} performs a series of iterative optimizations. The steps including E-step and M-step proceed as follows [33]:

- 1) Initialization: initialize the value of the model parameter vector Θ^0 .
- 2) E-step: given the parameter estimate Θ^s obtained in the previous iteration, calculate the Q-function

$$Q(\Theta|\Theta^s) = E_{C_{\text{mis}}|C_{\text{obs}}, \Theta^s} \{\log p(C_{\text{obs}}, C_{\text{mis}}|\Theta)\},$$

- 3) M-step: calculate the new parameter estimate Θ^{s+1} by maximizing $Q(\Theta|\Theta^s)$ with respect to Θ . That is

$$\Theta^{s+1} = \arg \max_{\Theta} Q(\Theta|\Theta^s).$$

The procedure including E-step and M-step is carried out iteratively until the change in parameters after each iteration is within a specified tolerance level. The value of the Q-function is ensured to be non-decreasing at each iteration. The convergence of the EM algorithm has been proved by Wu [34].

3.2 The application of auxiliary model approach

Because $x(t)$ in the information vector $\varphi_p(t)$ are unknown and are also included in $\varphi_r(t)$ and $\varphi(t)$, the calculation of E-step cannot be applied to Eq. (5) directly. The solution is to construct an auxiliary model or reference model $B_a(z^{-1})/A_a(z^{-1})$ using the system input $u(t)$, where $B_a(z^{-1})$ and $A_a(z^{-1})$ have the same order with $B(z^{-1})$ and $A(z^{-1})$ [35]. The main idea of auxiliary model approach can be described as shown in Fig. 2.

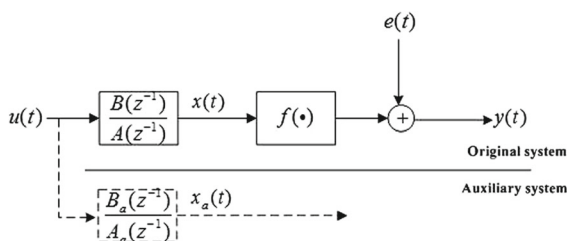


Fig. 2 The Wiener nonlinear system with the auxiliary model [23]

$$\begin{aligned} x_a(t) &= \frac{B_a(z^{-1})}{A_a(z^{-1})}u(t) \\ &= -a_1x_a(t-1) + \dots - a_nx_a(t-a_n) \\ &\quad + b_1u(t-1) + \dots + b_nu(t-b_n) \\ &= \varphi_a^T(t)\vartheta_a, \end{aligned} \tag{9}$$

where $\varphi_a(t)$ and ϑ_a are the information vector and the parameter vector of the auxiliary model, respectively. If we replace these unknown $x(t)$ in the information vector $\varphi_p(t)$ with output $x_a(t)$ of the auxiliary model, then the identification problem of ϑ can be solved by using $u(t)$, $y(t)$ and $x_a(t)$. It is noticed that the output $x_a(t)$ of the auxiliary model denoted by $\hat{x}(t)$ is used here as the estimate of $x(t)$. Define

$$\begin{aligned} \hat{\varphi}_p(t) &= [-\hat{x}(t-1), -\hat{x}(t-2), \dots, -\hat{x}(t-n_a), \\ &\quad u(t-1), u(t-2), \dots, u(t-n_b)]^T \\ \hat{\varphi}_r(t) &= [\hat{x}^2(t), \dots, \hat{x}^{n_r}(t)]^T \\ \hat{\varphi}(t) &= [\hat{\varphi}_p^T(t), \hat{\varphi}_r^T(t)]^T \end{aligned} \tag{10}$$

In identification, we use $\hat{\varphi}(t)$ to replace $\varphi(t)$, and based on the renewed and complete the information vectors, the EM algorithm can be carried out to identify the parameters of the Wiener model.

3.3 The mathematical formulation of the identification problem with EM algorithm

In this section, the EM algorithm is applied to solve the identification problem. The unknown parameters are $\Theta = \{\vartheta, \sigma^2\}$. The complete log likelihood function can be first decomposed using the probability chain rule as follows:

$$\log p(Y, U|\Theta) = \log p(Y|U, \Theta)p(U|\Theta) \tag{11}$$

The first term $p(Y|U, \Theta)$ can be decomposed into

$$\begin{aligned} p(Y|U, \Theta) &= p(y_{1:N}|u_{1:N}, \Theta) \\ &= p(y_N|y_{N-1:1}, u_{1:N}, \Theta)p(y_{N-1}|u_{1:N}, \Theta) \\ &= p(y_N|y_{N-1:1}, u_{1:N}, \Theta)p(y_{N-1}|y_{N-2:1}, u_{1:N}, \Theta) \\ &\quad \times \dots p(y_2|y_1, u_{1:N}, \Theta)p(y_1|u_{1:N}, \Theta) \end{aligned}$$

$$\begin{aligned}
&= \prod_{t=1}^N p(y_t|y_{t-1:1}, u_{1:N}, \Theta) \\
&= \prod_{t=1}^N p(y_t|u_{t-1:1}, \Theta). \quad (12)
\end{aligned}$$

Here, $\prod_{t=1}^N p(y_t|y_{t-1:1}, u_{1:N}, \Theta)$ can be simplified to $\prod_{t=1}^N p(y_t|u_{t-1:1}, \Theta)$ based on the fact that y_t only depends on the previous input sequence, namely $u_{t-1:1}$ and the parameter Θ . Since the input U of the system is the measurable data and is independent of the parameter Θ , the second term $p(U|\Theta)$ is constant defined as C . Therefore, the log likelihood function can be written as

$$\begin{aligned}
\log p(Y, U|\Theta) &= \log p(Y|U, \Theta)p(U|\Theta) \\
&= \sum_{t=1}^N \log p(y_t|u_{t-1:1}, \Theta) + \log C \\
&= \sum_{t=m_1}^{m_\alpha} \log p(y_t|u_{t-1:1}, \Theta) \\
&\quad + \sum_{t=o_1}^{o_\beta} \log p(y_t|u_{t-1:1}, \Theta) + \log C. \quad (13)
\end{aligned}$$

The Q-function can be obtained by calculating the expectation of the complete-data log likelihood function over the missing variable Y_{mis} ,

$$\begin{aligned}
Q(\Theta|\Theta^s) &= E_{C_{\text{mis}}|C_{\text{obs}}, \Theta^s} \{ \log p(U, Y|\Theta) \} \\
&= E_{Y_{\text{mis}}|C_{\text{obs}}, \Theta^s} \left\{ \sum_{t=m_1}^{m_\alpha} \log p(y_t|u_{t-1:1}, \Theta) \right. \\
&\quad \left. + \sum_{t=o_1}^{o_\beta} \log p(y_t|u_{t-1:1}, \Theta) + \log C \right\} \\
&= \sum_{t=m_1}^{m_\alpha} \int p(y_t|C_{\text{obs}}, \Theta^s) \log p(y_t|u_{t-1:1}, \Theta) dy_t \\
&\quad + \sum_{t=o_1}^{o_\beta} \log p(y_t|u_{t-1:1}, \Theta) + \log C. \quad (14)
\end{aligned}$$

Based on Eq. (5) and the Gaussian white noise assumption, we have

$$p(y_t|u_{t-1:1}, \Theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ \frac{-(y_t - \varphi^T(t)\vartheta)^2}{2\sigma^2} \right\}. \quad (15)$$

The problem left is to calculate the integral term $\int p(y_t|C_{\text{obs}}, \Theta^s) \log p(y_t|u_{t-1:1}, \Theta) dy_t$ in the Q-function. Based on the definitions of the first order and second order moments, the integral term can be calculated as

$$\begin{aligned}
&\int p(y_t|C_{\text{obs}}, \Theta^s) \log p(y_t|u_{t-1:1}, \Theta) dy_t \\
&= \int p(y_t|C_{\text{obs}}, \Theta^s) \log \frac{1}{\sqrt{2\pi\sigma^2}} \\
&\quad \times \exp \frac{-(y_t - \varphi^T(t)\vartheta)^2}{2\sigma^2} dy_t \\
&= -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \int p(y_t|C_{\text{obs}}, \Theta^s) \\
&\quad \times (y_t - \varphi^T(t)\vartheta)^2 dy_t \\
&= -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} ((\sigma^s)^2 + (\varphi^T(t)\vartheta^s)^2) \\
&\quad + \frac{1}{\sigma^2} (\varphi^T(t)\vartheta)(\varphi^T(t)\vartheta^s) - \frac{1}{2\sigma^2} (\varphi^T(t)\vartheta)^2 \\
&= -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} ((\sigma^s)^2 \\
&\quad + (\varphi^T(t)\vartheta - \varphi^T(t)\vartheta^s)^2) \quad (16)
\end{aligned}$$

Substituting Eqs. (15) and (16) into Eq. (14), we have

$$\begin{aligned}
Q(\Theta|\Theta^s) &= \sum_{t=m_1}^{m_\alpha} \left\{ -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} ((\sigma^s)^2 \right. \\
&\quad \left. + (\varphi^T(t)\vartheta - \varphi^T(t)\vartheta^s)^2 \right\} \\
&\quad + \sum_{t=o_1}^{o_\beta} \left\{ -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \right. \\
&\quad \left. \times (y_t - \varphi^T(t)\vartheta)^2 \right\} + \log C. \quad (17)
\end{aligned}$$

The following step is to obtain the estimates of all the unknown parameters, that is M-step. Taking the gradient of $Q(\Theta|\Theta^s)$ with respect to ϑ and σ^2 , respectively and setting them to zeros, the estimate of Θ can be derived as

$$\vartheta^{s+1} = \frac{\sum_{t=m_1}^{m_\alpha} \varphi(t)\varphi^T(t)\vartheta^s + \sum_{t=o_1}^{o_\beta} \varphi(t)y(t)}{\sum_{t=1}^N \varphi(t)\varphi^T(t)} \quad (18)$$

$$(\sigma^2)^{s+1} = \frac{\sum_{t=m_1}^{m_\alpha} ((\sigma^s)^2 + (\varphi^T(t)\vartheta^{s+1} - \varphi^T(t)\vartheta^s)^2) + \sum_{t=o_1}^{o_\beta} (y_t - \varphi^T(t)\vartheta^{s+1})^2}{N} \tag{19}$$

The detailed derivations of Eqs. (18) and (19) are given in Appendix.

Since the $\{x(t)\}_{t=1,\dots,N}$ in the information vector $\varphi(t)$ are unknown, they can be estimated by using the auxiliary model identification idea. Here, the auxiliary model in Eq. (9) is constructed based on the parameter estimates obtained in previous iteration. Therefore, the estimate of the information vector $\varphi(t)$ can be constructed based on Eq. (10). The new parameter estimates can be calculated by substituting $\varphi(t)$ with $\hat{\varphi}(t)$ in Eqs. (18) and (19):

$$\vartheta^{s+1} = \frac{\sum_{t=m_1}^{m_\alpha} \hat{\varphi}(t)\hat{\varphi}^T(t)\vartheta^s + \sum_{t=o_1}^{o_\beta} \hat{\varphi}(t)y(t)}{\sum_{t=1}^N \hat{\varphi}(t)\hat{\varphi}^T(t)}, \tag{20}$$

$$(\sigma^2)^{s+1} = \frac{\sum_{t=m_1}^{m_\alpha} ((\sigma^s)^2 + (\hat{\varphi}^T(t)\vartheta^{s+1} - \hat{\varphi}^T(t)\vartheta^s)^2) + \sum_{t=o_1}^{o_\beta} (y_t - \hat{\varphi}^T(t)\vartheta^{s+1})^2}{N}. \tag{21}$$

4 Simulation sample

Considering the following Wiener nonlinear system with the linear subsystem given as follows,

$$x(t) = \frac{B(z^{-1})}{A(z^{-1})}u(t),$$

$$A(z^{-1}) = 1 + a_1z^{-1} + a_2z^{-2} = 1 + 0.58z^{-1} + 0.41z^{-2}$$

$$B(z^{-1}) = b_1z^{-1} + b_2z^{-2} = -0.18z^{-1} + 0.44z^{-2} \tag{22}$$

and the nonlinearity is described by

$$f(x(t)) = r_1x(t) + r_2x^2(t) + \dots + r_nx^{n_r}(t)$$

$$= x(t) - 0.45x^2(t) + 0.25x^3(t) \tag{23}$$

3.4 The summary of the proposed identification algorithm

The proposed approach for nonlinear Wiener models taking the randomly missing outputs into account using the EM algorithm can be summarized as follows:

- 1) Set $s = 1$ and initialize the parameter vector ϑ and the variance σ^2 .
- 2) Calculate the estimates of $\{x(t)\}_{t=1,\dots,N}$ according to Eq. (9) with the parameters obtained in the previous iteration.
- 3) Update the estimates of the parameter ϑ^{s+1} and the variance $(\sigma^2)^{s+1}$ according to Eqs. (20) and (21), respectively.
- 4) Set $s = s + 1$ and repeat step 2 to step 3 until convergence.

The output of the Wiener system $y(t)$ can be expressed as

$$y(t) = f(x(t)) + e(t) \tag{24}$$

For this example system, the parameter vector of the Wiener model to be identified is $\vartheta = [0.58, 0.41, -0.18, 0.44, -0.45, 0.25]$. The input sequence $u(t)$ and output sequence $y(t)$ are generated by simulation, and $e(t)$ is a white noise process with zero mean and variance 0.001 added to the output. The input–output data of the system are given in Fig. 3. Setting the missing rate of the output data at around 12.5%, the proposed method is applied to identify the six parameters and noise variance simultaneously. The initial values of vector ϑ and variance σ^2 are $[0.45, 0.5, -0.2, 0.5, -0.41, 0.41]$ and 0.05, respec

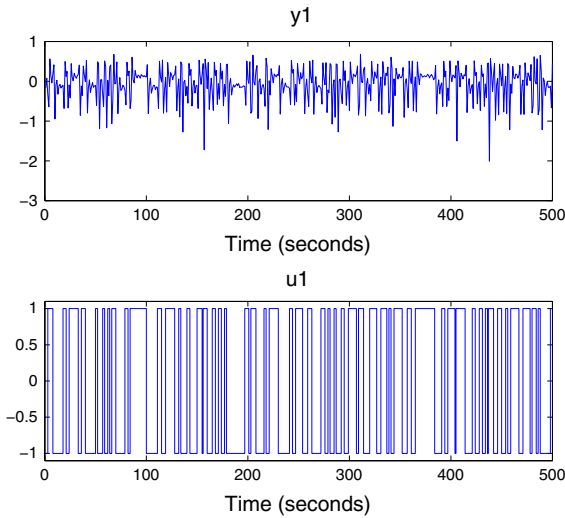


Fig. 3 The input and output data

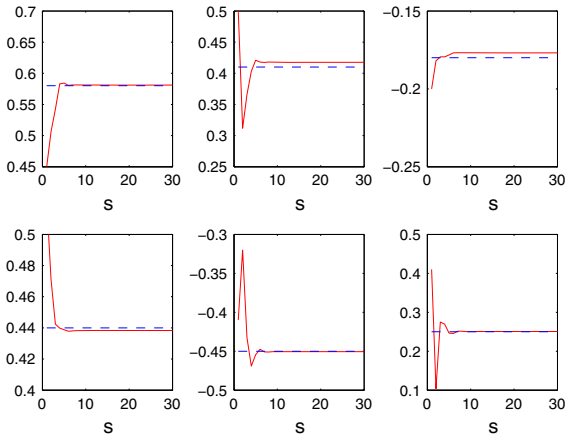


Fig. 4 The EM estimates of the Wiener model parameters with output missing 15%

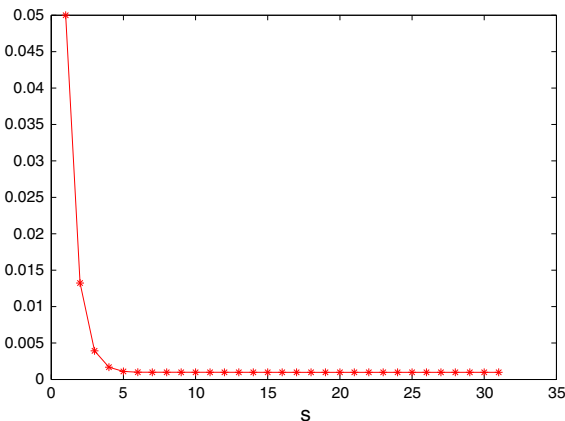


Fig. 5 The EM estimate of the noise with output missing 15%

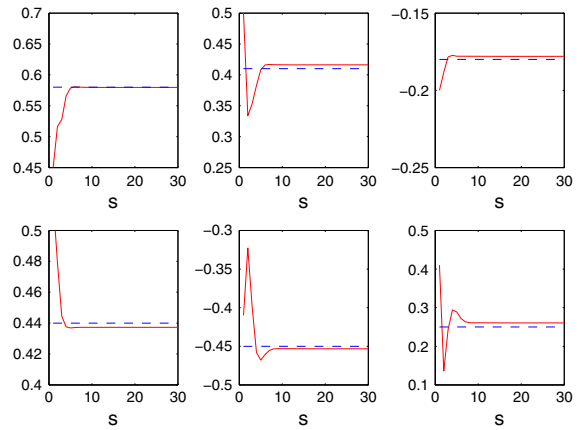


Fig. 6 The EM estimates of the Wiener model parameters with output missing 25%

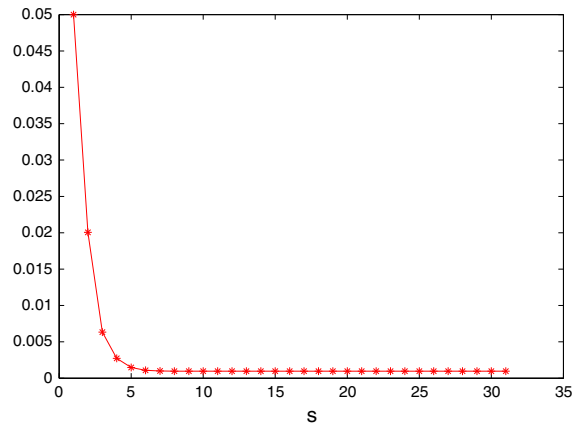


Fig. 7 The EM estimate of the noise with output missing 25%

tively. The estimated parameters versus iteration are shown in Fig. 4. It can be seen that the proposed EM-based identification algorithm has good performances since the parameter estimates approach the real ones after a few iterations. The noise variance trajectory is shown in Fig. 5. It is clear that almost all the parameters are close to the real value after 10 iterations.

To illustrate the effectiveness of the proposed method in dealing with the randomly data missing, the simulation is also carried out with the missing rate of the output at around 25 and 50%, respectively. The simulation results are shown from Figs. 6, 7, 8 and 9. It is noticed that the proposed approach keeps a good identification performance when the missing data are near the half of the whole output sequence. To evaluate the performance of the proposed algorithm, the relative

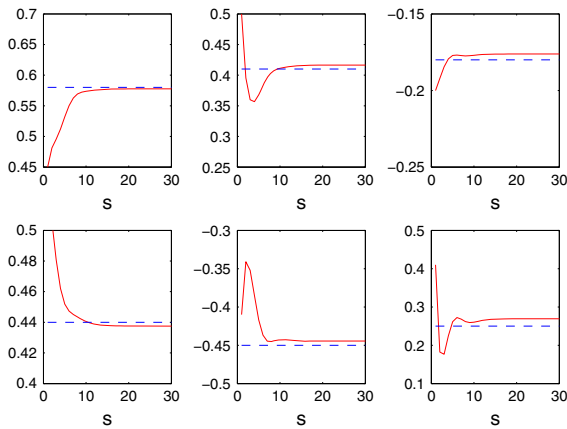


Fig. 8 The EM estimates of the Wiener model parameters with output missing 50%

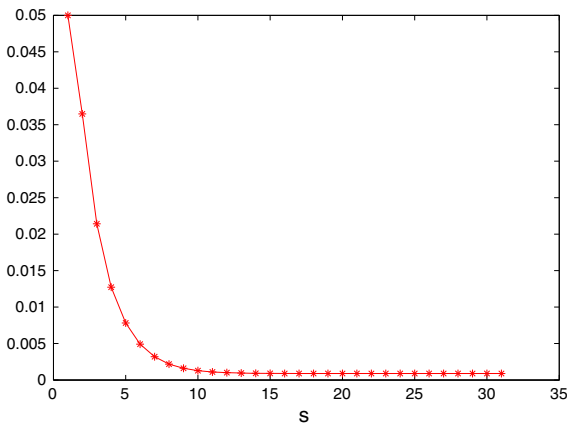


Fig. 9 The EM estimate of the noise with output missing 50%

error (RE) of the estimated parameter criterion can be used and is defined as:

$$RE = \sqrt{\frac{\|\hat{\Theta} - \Theta\|}{\|\Theta\|}} \tag{25}$$

From Table 1, we can see that with the increase in the missing rate, the relative error becomes larger.

5 Conclusions

This paper considers the parameter identification for a class of nonlinear Wiener models in the stochastic framework and takes the randomly missing out-

Table 1 Parameter estimates after 30 iterations under different missing rates

Para	No missing	12.5(%) missing	25(%) missing	50(%) missing
a_1	0.5817	0.5801	0.5849	0.5901
a_2	0.4175	0.4166	0.4180	0.4233
b_1	-0.1780	-0.1783	-0.1792	-0.1809
b_2	0.4390	0.4373	0.4412	0.4393
r_2	-0.4517	-0.4548	-0.4474	-0.4372
r_3	0.2437	0.2594	0.2289	0.2121
σ^2	0.0010	0.0010	0.0010	0.0010
RE	0.0104	0.0128	0.0233	0.0435

put problem into account. To deal with the missing outputs, the EM algorithm is employed to estimate the parameters and the noise variance simultaneously and the unknown noise-free outputs are estimated by using the auxiliary model identification idea [36,37]. Thereafter, the identification problem is formulated under the framework of the EM algorithm. A numerical example is provided to demonstrate the effectiveness of the proposed algorithm. The proposed algorithm can be extended to study identification problem of other linear systems [38–42] and nonlinear systems [43–45].

Appendix: Detailed derivation of Eqs.(18) and (19)

The Q-function in Eq. (17) can be further written as

$$\begin{aligned}
 Q(\Theta|\Theta^s) = & \sum_{t=m_1}^{m_\alpha} \left\{ -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} ((\sigma^s)^2 \right. \\
 & \left. + (\varphi^T(t)\vartheta - \varphi^T(t)\vartheta^s)^T (\varphi^T(t)\vartheta - \varphi^T(t)\vartheta^s) \right\} \\
 & + \sum_{t=o_1}^{o_\beta} \left\{ -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} (y_t - \varphi^T(t)\vartheta)^T \right. \\
 & \left. (y_t - \varphi^T(t)\vartheta) \right\} + \log C. \tag{26}
 \end{aligned}$$

Taking the gradient of $Q(\Theta|\Theta^s)$ with respect to ϑ and setting it to zero,

$$\begin{aligned} \frac{\partial Q(\Theta|\Theta^s)}{\partial \vartheta} &= \sum_{t=m_1}^{m_\alpha} \left\{ -\frac{1}{2\sigma^2} (2\varphi(t)\varphi^T(t)\vartheta \right. \\ &\quad \left. - 2\varphi(t)\varphi^T(t)\vartheta^s) \right\} \\ &\quad + \sum_{t=o_1}^{o_\beta} \left\{ -\frac{1}{2\sigma^2} (2\varphi(t)\varphi^T(t)\vartheta \right. \\ &\quad \left. - 2\varphi(t)y_t) \right\} \\ &= 0 \end{aligned} \quad (27)$$

Through keeping the terms that not related with ϑ at the right side, Eq. (27) can be written as

$$\begin{aligned} \sum_{t=m_1}^{m_\alpha} \varphi(t)\varphi^T(t)\vartheta + \sum_{t=o_1}^{o_\beta} \varphi(t)\varphi^T(t)\vartheta \\ = \sum_{t=m_1}^{m_\alpha} \varphi(t)\varphi^T(t)\vartheta^s + \sum_{t=o_1}^{o_\beta} \varphi_t y_t \end{aligned} \quad (28)$$

Then, the new estimate of parameter ϑ can be obtained as:

$$\vartheta^{s+1} = \frac{\sum_{t=m_1}^{m_\alpha} \varphi(t)\varphi^T(t)\vartheta^s + \sum_{t=o_1}^{o_\beta} \varphi(t)y(t)}{\sum_{t=1}^N \varphi(t)\varphi^T(t)} \quad (29)$$

Taking the gradient of $Q(\Theta|\Theta^s)$ with respect to σ^2 and setting it to zero,

$$\begin{aligned} \frac{\partial Q(\Theta|\Theta^s)}{\partial \sigma^2} &= \sum_{t=m_1}^{m_\alpha} \left\{ -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4} \left[(\sigma^s)^2 \right. \right. \\ &\quad \left. \left. + (\varphi^T(t)\vartheta - \varphi^T(t)\vartheta^s)^2 \right] \right\} \\ &\quad + \sum_{t=o_1}^{o_\beta} \left\{ -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4} (y_t - \varphi^T(t)\vartheta)^2 \right\} \\ &= 0. \end{aligned} \quad (30)$$

Through keeping the two terms including σ^2 at the left side, the Eq. (30) can be written as:

$$\begin{aligned} \sum_{t=m_1}^{m_\alpha} \sigma^2 + \sum_{t=o_1}^{o_\beta} \sigma^2 &= N \cdot \sigma^2 \\ &= \sum_{t=m_1}^{m_\alpha} \left\{ (\sigma^s)^2 + (\varphi^T(t)\vartheta - \varphi^T(t)\vartheta^s)^2 \right\} \\ &\quad + \sum_{t=o_1}^{o_\beta} (y_t - \varphi^T(t)\vartheta)^2 \end{aligned} \quad (31)$$

Then, the estimation of parameter σ^2 can be obtained as:

$$(\sigma^2)^{s+1} = \frac{\sum_{t=m_1}^{m_\alpha} ((\sigma^s)^2 + (\varphi^T(t)\vartheta^{s+1} - \varphi^T(t)\vartheta^s)^2) + \sum_{t=o_1}^{o_\beta} (y_t - \varphi^T(t)\vartheta^{s+1})^2}{N} \quad (32)$$

References

- Wang, C., Tang, T.: Several gradient-based iterative estimation algorithms for a class of nonlinear systems using the filtering technique. *Nonlinear Dyn.* **77**(3), 769–780 (2014)
- Rashid, M.T., Frasca, M., et al.: Nonlinear model identification for Artemia population motion. *Nonlinear Dyn.* **69**(4), 2237–2243 (2012)
- Ding, F.: *System Identification—New Theory and Methods*. Science Press, Beijing (2013)
- Ding, F.: *System Identification—Performances Analysis for Identification Methods*. Science Press, Beijing (2014)
- Yin, S., Ding, S., Haghani, A., Hao, H.: Data-driven monitoring for stochastic systems and its application on batch process. *Int. J. Syst. Sci.* **44**(7), 1366–1376 (2013)
- Sun, W., Gao, H., Kaynak, O.: Finite frequency H_{∞} control for vehicle active suspension systems. *IEEE Trans. Control Syst. Tech.* **19**(2), 416–422 (2011)
- Ding, F., Chen, T.: Identification of Hammerstein nonlinear ARMAX systems. *Automatica* **41**(9), 1479–1489 (2005)
- Wang, D.Q., Ding, F.: Hierarchical least squares estimation algorithm for Hammerstein–Wiener systems. *IEEE Signal Process. Lett.* **19**(12), 825–828 (2012)
- Hagenblad, A., Ljung, L., Wills, A.: Maximum likelihood identification of Wiener models. *Automatica* **44**(11), 2697–2705 (2008)
- Fan, D., Lo, K.: Identification for disturbed MIMO Wiener systems. *Nonlinear Dyn.* **55**, 31–42 (2009)
- Janczak, A.: Instrumental variables approach to identification of a class of MIMO Wiener system. *Nonlinear Dyn.* **48**, 275–284 (2007)
- Zhou, L., Li, X., Pan, F.: Gradient-based iterative identification for Wiener nonlinear systems with non-uniform sampling. *Nonlinear Dyn.* **76**, 627–634 (2014)
- Wigren, T.: Recursive prediction error identification using the nonlinear Wiener model. *Automatica* **29**(4), 1011–1025 (1993)
- Wigren, T.: Convergence analysis of recursive identification algorithm based on the nonlinear Wiener model. *IEEE Trans. Autom. Control* **39**(11), 2191–2206 (1994)
- Wang, D.Q., Ding, F.: Least squares based and gradient based iterative identification for Wiener nonlinear systems. *Signal Process.* **91**(5), 1182–1189 (2011)
- Ding, F., Shi, Y., Chen, T.: Auxiliary model-based least-squares identification methods for Hammerstein output-error systems. *Syst. Control Lett.* **56**(5), 373–380 (2007)
- Xiong, W.L., Ma, J.X., Ding, R.: An iterative numerical algorithm for a class of Wiener nonlinear system modeling. *Appl. Math. Lett.* **26**(4), 487–493 (2012)
- Westwick, D., Verhaegen, M.: Identifying MIMO Wiener systems using subspace model identification methods. *Signal Process.* **52**(2), 235–258 (1996)
- Ding, F.: State filtering and parameter identification for state space systems with scarce measurements. *Signal Process.* **104**, 369–380 (2014)
- Khatibisepehr, S., Huang, B.: Dealing with irregular data in soft sensors: Bayesian method and comparative study. *Ind. Eng. Chem. Res.* **47**(22), 8713–8723 (2008)
- Jin, X., Wang, S., Huang, B., Forbes, F.: Multiple model based LPV soft sensor development with irregular/missing process output measurement. *Control Eng. Pract.* **20**(2), 165–172 (2012)
- Ding, J., Ding, F., Liu, X.P., Liu, G.: Hierarchical least squares identification for linear SISO systems with dual-rate sampled-data. *IEEE Trans. Autom. Control* **56**(11), 2677–2683 (2011)
- Ding, F., Ding, J.: Least-squares parameter estimation for systems with irregularly missing data. *Int. J. Adapt. Control Signal Process.* **24**(7), 540–553 (2010)
- Ding, F., Liu, G., Liu, X.P.: Parameter estimation with scarce measurements. *Automatica* **47**(8), 1646–1655 (2011)
- Zhu, Y., Telkamp, H., Wang, J., Fu, Q.: System identification using slow and irregular output samples. *J. Process Control* **19**(1), 58–67 (2009)
- Isaksson, A.J.: Identification of ARX-models subject to missing data. *IEEE Trans. Autom. Control* **38**(5), 813–819 (1993)
- Wallin, R., Isaksson, A.J., Ljung, L.: An iterative method for identification of ARX models from incomplete data. In: *Proceedings of the 39th IEEE Conference Decision Control* 1, pp. 203–208 (2000)
- Gopaluni, R.B.: A particle filter approach to identification of nonlinear process under missing observations. *Can. J. Chem. Eng.* **86**(6), 1081–1092 (2008)
- Xie, L., Yang, H.Z., Huang, B.: FIR model identification of multirate processes with random delays using EM algorithm. *AIChE J.* **59**(11), 4124–4132 (2013)
- Deng, J., Huang, B.: Identification of nonlinear parameter varying systems with missing output data. *AIChE J.* **58**(11), 3454–3467 (2012)
- Xiong, W., Yang, X., Huang, B., Xu, B.: Multiple-model based linear parameter varying time-delay system identification with missing output data using an expectation-maximization algorithm. *Ind. Eng. Chem. Res.* **53**, 11074–11083 (2014)
- Yang, X., Gao, H.: Multiple model approach to linear parameter varying time-delay system identification with EM algorithm. *J. Frankl. Inst.* **351**(12), 5565–5581 (2014)
- Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc.* **39**(1), 1–38 (1977)
- Wu, J.: On the convergence properties of the EM algorithm. *Ann. Stat.* **11**(1), 95–103 (1983)
- Ding, F., Chen, T.: Combined parameter and output estimation of dual-rate systems using an auxiliary model. *Automatica* **40**(10), 1739–1748 (2004)
- Ding, F.: Hierarchical parameter estimation algorithms for multivariable systems using measurement information. *Inf. Sci.* **277**, 396–405 (2014)
- Ding, F., Wang, Y.J., Ding, J.: Recursive least squares parameter estimation algorithms for systems with colored noise using the filtering technique and the auxiliary model. *Digit. Signal Process.* (2015). <http://dx.doi.org/10.1016/j.dsp.2014.10.005>
- Hu, Y.B.: Iterative and recursive least squares estimation algorithms for moving average systems. *Simul. Model. Practice Theory* **34**, 12–19 (2013)
- Ding, J., Fan, C.X., Lin, J.X.: Auxiliary model based parameter estimation for dual-rate output error systems with colored noise. *Appl. Math. Model.* **37**(6), 4051–4058 (2013)

40. Ding, J., Lin, J.X.: Modified subspace identification for periodically non-uniformly sampled systems by using the lifting technique. *Circuits Syst. Signal Process.* **33**(5), 1439–1449 (2014)
41. Liu, Y.J., Ding, F., Shi, Y.: An efficient hierarchical identification method for general dual-rate sampled-data systems. *Automatica* **50**(3), 962–970 (2014)
42. Ding, F.: Combined state and least squares parameter estimation algorithms for dynamic systems. *Appl. Math. Model.* **38**(1), 403–412 (2014)
43. Wang, C., Tang, T.: Recursive least squares estimation algorithm applied to a class of linear-in-parameters output error moving average systems. *Appl. Math. Lett.* **29**, 36–41 (2014)
44. Hu, Y.B., Liu, B.L., Zhou, Q., Yang, C.: Recursive extended least squares parameter estimation for Wiener nonlinear systems with moving average noises. *Circuits Syst. Signal Process.* **33**(2), 655–664 (2014)
45. Hu, Y.B., Liu, B.L., Zhou, Q.: A multi-innovation generalized extended stochastic gradient algorithm for output nonlinear autoregressive moving average systems. *Appl. Math. Comput.* **247**, 218–224 (2014)