



Adaptive filtering-based recursive identification for time-varying Wiener output-error systems with unknown noise statistics

Zhu Wang*, Haoran An, Xionglin Luo

Received 7 November 2018; received in revised form 21 June 2019; accepted 4 November 2019
Available online 9 November 2019

Abstract

In area of control, model-based robust identification is rare, and studies in presence of unknown noise statistics are especially seldom. The robust estimation problem for time-varying Wiener output-error systems is considered in this paper. An adaptive filtering-based recursive identification scheme is proposed to distinguish nonlinear time-varying characteristics in complex noise environments. Firstly, a virtual equivalent state space model is constructed to achieve adaptive Kalman filtering. In filter design, a weighted noise estimator based on Sage-Husa principle is introduced, and is sensitive to noise changes. Secondly, the state estimates obtained by filters are used to form the unknown intermediate variables in information vectors. Then, a recursive estimation method based on multiple iterations is developed, and the convergence of identification is confirmed by martingale hyperconvergence theorem. Finally, the numerical simulation results verify the theoretical findings.

© 2019 The Franklin Institute. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Nonlinear block-oriented models can be used to describe many nonlinear dynamic behaviors [1]. Three typical nonlinear block-oriented models are Wiener models, Hammerstein models [2], and their combinations [3,4]. These types of models contain both static nonlin-

* Corresponding author.

wz21561@163.com (Z. Wang), anhaoran_9911@163.com (H. An), luoxl@cup.edu.cn (X. Luo).

ear blocks and linear dynamic blocks. Among them, a Wiener model has a linear dynamic part followed by a nonlinear part [5], and it is widely encountered in engineering practices. Some biological processes [6] and chemical processes [7] are usually modeled by Wiener systems. Meanwhile, these processes are considered time-varying [8] because of environmental or human influence. Besides, if system physical laws are accurate sufficiently [9], the absence of process noise is reasonable, and such systems are known as output-error systems on system identification. Hence, this paper considers the identification of time-varying Wiener output-error systems, and aims to grasp nonlinear and time-varying characteristics of systems.

The identification issues for Wiener systems have attracted great attention. Jvoros [10] studied the identification of Wiener models that have piecewise-linear or inverse functions of nonlinear parts. The identification of non-invertibility nonlinear parts has been also studied by Hu and Chen [11]. For FIR Wiener systems, Lacy and Berntein [12] exploited simultaneous direct estimation of non-invertible, polynomial non-linearities. From another perspective, existing identification methods for Wiener systems are roughly divided to several categories, e.g., blind approaches [13,14], maximum likelihood methods [15], iterative methods [16], and recursive methods [17]. Among these methods, recursive methods are popular because they are suitable for online identification and can be used for time-varying systems [18]. For Wiener systems whose output nonlinear function is continuous and invertible, Ding et al. proposed the auxiliary-model based recursive least squares algorithm [19]. Further, multiple iterations [20–22] in recursive identification have been introduced to improve robustness of parameter estimation in noisy environments.

Kalman filter is a widely used model-based state estimator [23–25], and is applied to linear systems with Gaussian noise. For time-varying output-error system identification, the stability of Kalman filter [26] has been ensured under uniform complete observability. For Hammerstein systems [27,28], the modified Kalman smoother was derived to estimate the unknown intermediate variables in systems. For the real-time estimation of 4WD vehicle states, the extended Kalman filter [29] was combined with minimum error criterion. For actual fault diagnosis of time varying systems, the adaptive Kalman filter [30] was proposed through joint state-parameter estimation. For soft sensor maintenance, data fusion technology was introduced based on Kalman filter [31]. However, under unknown noise statistics, estimators should be redesigned to achieve adaptive filtering. In this respect, Sage-Husa maximum a posteriori estimation principle [32] has been applied into Kalman filter, in order to give statistical noise properties. Unlike Nussbaum designs [33,34] to handle non-zero-mean nonlinearities, Sage-Husa Kalman filter can recursively estimates both mean values and variances of noise.

In this paper, a linear regression form of Wiener nonlinear systems is adopted for identification. The proposed recursive estimation is used to achieve identification of time-varying parameters, and the idea of multiple iterations is introduced to improve robustness. Meanwhile, the bounded convergence of time-varying systems is harvested by using martingale theorems [35,36]. Further, to estimate the unknown intermediate variables in information vectors, a virtual equivalent state space model is proposed, and an extended Kalman filter is implemented. This adaptive Kalman filter is designed for output-error time-varying systems, and is based on the results of [26]. Besides, Sage-Husa noise estimator is also introduced into adaptive filtering. Thus, the recursive identification based on adaptive filtering is able to integrate state estimation with system identification, and can provide an effective identification method for time-varying Wiener output-error systems under complex noise environments [37].

The rest of paper is organized as follows. Section 2 gives the system description. The adaptive filtering-based recursive estimation is exploited in Section 3. Section 4 shows main

The assumptions (A1) and (A2) shows that stochastic noise has properties of both zero mean and bounded variances. However, the noise can be either white or colored, and its statistical property can be non-Gaussian and time-varying. Specifically, the noise variances $\Sigma(t)$ can change with time. These assumptions reflect unknown noise statistics.

Assumptions 3 and 4. Define the parameter variation $\theta(t) = \theta(t) - \theta(t - 1)$. Assume that $\theta(t)$ satisfies the following assumptions

$$E[\theta(t) | \mathcal{Y}(t)] = 0, \quad \theta(t) \neq 0; E[v(t) | \mathcal{Y}(t)] = 0, \tag{A3}$$

$$E[\|\theta(t)\|^2] \leq \varepsilon \|\theta(t - 1)\|^2 \leq \sigma^2 < \infty, \tag{A4}$$

where $\varepsilon > 0$ is a very small number. That is, the parameter variation $\theta(t)$ remains tiny in $\theta(t - 1)$. The assumptions illustrate that the variation is uncorrelated with ambient noise.

3. Adaptive filtering-based recursive identification

This section aims to employ a weighted gradient identification algorithm to estimate parameters. Adaptive filtering is introduced to harvest the optimal estimates of true $\theta_1(t)$.

Referring to the works in [20], multiple iterations in recursive identification can improve robustness of parameter estimation in noisy environments, and thus improve model accuracy. Because of time-varying property of parameters, damping coefficients are also introduced to cost functions. That is, different coefficients enable fast forgetting of past information, and thus enable fast tracking of current parameter changes. Define the cost function as

$$J(\theta) = \frac{1}{2} \sum_{i=t-\tau+1}^t \left[\lambda^{-i} \|\theta(i) - \hat{\theta}(i)\|^2 \right], \tag{5}$$

where τ denotes the number of samples from epoch $t - \tau + 1$ to epoch t at each recursion. From (5), the damping coefficients fitted by negative exponential law are selected as $\{\lambda^{-1}, \dots, \lambda^{-\tau}\}$, where the scalar λ represents attenuation speed, and $\lambda = (1 - \alpha)/(1 - \alpha^\tau)$, $0 < \alpha < 1$. It should be noted that the values of λ and τ need to be adjusted, according to situations of parameter changes. The tradeoff between robust estimation and change track needs to be solved. Thus, the proposed cost function can not only guarantee robust estimation, but also keep track of parameter changes. The quadratic criterion in (5) can be minimized by adopting the well-known Gauss-Newton technique [20]. Define

$$\begin{aligned} \theta(t) &= [\theta(t - \tau + 1), \dots, \theta(t)]^T, \\ Y(t) &= [Y(t - \tau + 1), \dots, Y(t)]^T, \\ W(t) &= \text{diag}(\lambda^{-1}, \dots, \lambda^{-1}, \lambda^{-\tau}), \end{aligned} \tag{6}$$

where $\theta(t) \in \mathbb{R}^{n \times (\tau+1)}$ and $W(t) \in \mathbb{R}^{n \times n}$. Then, the following equation is yielded

$$-\frac{\partial}{\partial \theta} J(\theta) = -W(t)W(t)[Y(t) - \theta(t)]. \tag{7}$$

The Hessian of $J(\theta)$ is given by

$$\frac{\partial^2}{\partial \theta^2} J(\theta) = W(t)W(t)W(t). \tag{8}$$

The Gauss-Newton solution of minimizing (5) can be generated by

$$\begin{aligned} \hat{\theta}(t) &= \hat{\theta}(t-1) + \lambda(t) \left[-\frac{\partial^2 J(\theta)}{\partial \theta^2} \right]^{-1} \cdot \frac{\partial J(\theta)}{\partial \theta} \Big|_{\hat{\theta}(t-1)} \\ &= \hat{\theta}(t-1) + \lambda(t) \mathbf{H}^{-1}(t) \mathbf{W}(t) \left[\mathbf{Y}(t) - \mathbf{H}(t) \hat{\theta}(t-1) \right], \end{aligned} \tag{9}$$

where $\mathbf{H}(t) := \mathbf{W}(t) \mathbf{W}(t)^T$ and $\lambda(t)$ is a sequence of positive scalar. To enhance tracking capability of process changes, $\lambda(t)$ is usually given by

$$\lambda(t) = \gamma / (\lambda(t-1) + 1), \tag{10}$$

where initial term $\lambda(0)$ denotes a small number specified by the designer, and γ denotes forgetting factor. In addition, to prevent $\mathbf{H}(t)$ from being singular, $\mathbf{H}(t)$ can be modified according to Robbins-Monro [20] method

$$\mathbf{H}(t) = \mathbf{H}(t-1) + \lambda(t) (\mathbf{W}(t) \mathbf{W}(t)^T - \mathbf{H}(t-1)). \tag{11}$$

In summary, the proposed gradient-based recursive identification (RI) method is formulated as

Gradient-based recursive identification (RI) method

Begin

1. Initialize the vectors and matrices $\hat{\theta}(0), \mathbf{H}(0), \lambda(0)$ at epoch 0.
2. For $t \geq 1$, implement the proposed recursive identification algorithm

$$\mathbf{t}(t) = [t(t-1), \dots, t(t)]^T, \tag{12}$$

$$\mathbf{Y}(t) = [y(t-1), \dots, y(t)]^T, \tag{13}$$

$$\mathbf{W}(t) = \mathbf{W}(t-1) + \lambda(t) (\mathbf{t}(t) \mathbf{t}(t)^T - \mathbf{W}(t-1)), \tag{14}$$

$$\lambda(t) = \gamma / (\lambda(t-1) + 1), \tag{15}$$

$$\mathbf{H}(t) = \mathbf{H}(t-1) + \lambda(t) (\mathbf{W}(t) \mathbf{W}(t)^T - \mathbf{H}(t-1)), \tag{16}$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \lambda(t) \mathbf{H}^{-1}(t) \mathbf{W}(t) [\mathbf{Y}(t) - \mathbf{H}(t) \hat{\theta}(t-1)] \tag{17}$$

3. After getting $\hat{\theta}(t)$, the parameter estimates $\{\hat{\theta}(t)\}_{t=1}, \{\hat{\theta}(t)\}_{t=1}$ and $\{\hat{\theta}(t)\}_{t=1}$ can be further calculated. According to [21], the average method gives the following parameter estimates

$$\{\hat{\theta}(t)\}_{t=1} = [\hat{\theta}_1(t), \dots, \hat{\theta}_n(t)]^T, \tag{18}$$

$$\{\hat{\theta}(t)\}_{t=1} = \frac{1}{t} \sum_{i=1}^t \left[1, \frac{\hat{\theta}_1(i)}{\hat{\theta}_1(t)}, \frac{\hat{\theta}_2(i)}{\hat{\theta}_2(t)}, \dots, \frac{\hat{\theta}_{(n-1)}(i)}{\hat{\theta}_{(n-1)}(t)} \right]^T, \tag{19}$$

$$\{\hat{\theta}(t)\}_{t=1} = [\hat{\theta}_{n+1}(t), \dots, \hat{\theta}_{n+m}(t)]^T, \tag{20}$$

where $\hat{\theta}_1(t) = 1$ ensures uniqueness of the result [21].

4. Increase t by 1 and go to Step 2.

End

It should be mentioned that the proposed RI method in (12)–(20) is based on multiple iterations. This means that the parameter updates for epoch t considers multiple errors of previous moments. Hence, this kind of “memory” ensures more accurate grasp of system characteristics, and enhances robustness of identification. Actually, robustness of identification means the

ability that can distinguish system characteristics in complex noise environments. Moreover, in order to keep track of time-varying system characteristics, the attenuation speed $\bar{\alpha}$ and the length \bar{L} should be selected. That is, the balance between robustness and traceability needs to be guaranteed.

However, the difficulty in this recursive algorithm is that the information vector $\mathbf{y}(t)$ contains unknown components $y_1(t-1)$. Hence, in this paper, an adaptive filtering method is exploited to obtain the optimal estimates of $y_1(t)$. It is well known that the extended Kalman filter [29] has been widely used to harvest state estimates for nonlinear state space models. For the system model (1), its equivalent state space model is constructed as following

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{f}_{t-1}(\mathbf{x}(t-1)) + \mathbf{B}_{t-1}v(t-1), \\ y(t) &= \mathbf{C}\mathbf{x}(t) + v(t), \end{aligned} \quad (21)$$

where

$$\begin{aligned} \mathbf{x}(t) &= [y_1(t), y_2(t), \dots, y_n(t)]^T, \\ \mathbf{B}_{t-1} &= [y_1(t-1), y_2(t-1), \dots, y_n(t-1)]^T, \\ \mathbf{C} &= [1, 0, \dots, 0], \\ \mathbf{f}_{t-1}(\mathbf{x}(t-1)) &= [y_{1(t-1)}(\mathbf{x}(t-1)), y_{2(t-1)}(\mathbf{x}(t-1)), \dots, y_{(n-1)(t-1)}(\mathbf{x}(t-1))]^T, \\ y_{(t-1)}(\mathbf{x}(t-1)) &= y(t-1) \sum_{i=1}^t \alpha_i(t-1) (y_1(t-1)) + y_{+1}(t-1), \quad i=1, \dots, (n-1), \\ y_{(n-1)(t-1)}(\mathbf{x}(t-1)) &= y(t-1) \sum_{i=1}^t \alpha_i(t-1) (y_1(t-1)), \end{aligned}$$

It should be noted that $\mathbf{B}_{t-1} \in \mathbb{R}^{1 \times n}$, $\mathbf{f}_{t-1}(\mathbf{x}(t-1)) \in \mathbb{R}^n$ and $\mathbf{C} \in \mathbb{R}^{1 \times n}$. The states of this equivalent model in (21) includes true $y_1(t)$ and other virtual states $y_2(t), \dots, y_n(t)$. Thus, (21) can be used for extended Kalman filtering. In addition, for unknown noise statistics, estimators should be designed in order to achieve adaptive filtering. Here, the design approach of noise estimator is based on the principle of Sage-Husa maximum a posteriori estimation. From (21), Sage-Husa noise estimator [32] in the framework of Kalman is shown as below

$$\hat{y}(t) = \frac{1}{t} \sum_{i=1}^t [y(i) - \mathbf{C}\hat{\mathbf{x}}(i-1)], \quad (22)$$

$$\hat{\mathbf{P}}(t) = \frac{1}{t} \sum_{i=1}^t [\varepsilon^2(i) - \mathbf{C}\mathbf{P}(i-1)\mathbf{C}^T], \quad (23)$$

where $\varepsilon(i)$ is the innovation, i.e., $\varepsilon(i) := y(i) - \mathbf{C}\hat{\mathbf{x}}(i-1) - \hat{y}(i)$, $\hat{\mathbf{x}}(i-1) \in \mathbb{R}^n$ is the time update state estimate for epoch i , $\mathbf{P}(i-1) \in \mathbb{R}^{n \times n}$ is the time update error covariance estimate for epoch i , and $\hat{y}(i)$ and $\hat{\mathbf{P}}(i)$ denote estimated mean values and estimated variances, respectively. However, according to [32], (22) and (23) are only suitable to the situation where noise prior statistics remains unchanged. If ambient noise changes rapidly and continuously, it will cause the estimator performance to degrade significantly. Based on Sage-Husa estimation, different damping coefficients are also introduced to deal with the information of noise estimations at different moments. Since the initial epoch for adaptive filtering starts from epoch t_0 , the sequence length of damping coefficients should be $(t-t_0)$ instead of t . The sequence fitted by negative exponential law can be selected as $\{\tilde{\alpha}_1^{t-t_0-1}, \tilde{\alpha}_2^{t-t_0-1}, \dots, \tilde{\alpha}_n^{t-t_0-1}\}$, where $\tilde{\alpha}_i^{t-t_0-1} = (1-\tilde{\alpha}_i)/(1-\tilde{\alpha}_i^{t-t_0-1})$ and $0 < \tilde{\alpha}_i < 1$. Hence, the weighted expressions of (22)–(23) from

In conclusion, the adaptive filtering algorithm and the recursive identification algorithm both start at epoch $(t + 1)$. That is, the relationship can be shown as below

$$t = t + 1,$$

$$\hat{\mathbf{x}}_1(t-1|t-1), \hat{\mathbf{x}}(t-1) \Rightarrow \hat{\mathbf{x}}(t),$$

$$\hat{\mathbf{x}}(t-1|t-1), \hat{\mathbf{x}}(t-1) \Rightarrow \hat{\mathbf{x}}(t|t) \Rightarrow \hat{\mathbf{x}}_1(t|t),$$

From (44)–(45), it yields

$$\begin{aligned}
 \mathbf{G}(t) &= \gamma(t-1) - \tilde{Y}(t)(\mathbf{W}(t) - \mathbf{W}(t-))\mathbf{G}^{-1}(t) - \tilde{Y}(t)\mathbf{W}(t) \\
 &\quad + \tilde{Y}(t)\mathbf{W}(t)\mathbf{G}^{-1}(t) - \tilde{Y}(t)\mathbf{W}(t) + \tilde{Y}(t)\mathbf{G}(t) \\
 &\quad + 2\tilde{Y}(t)(\mathbf{W}(t) - \mathbf{W}(t-))\mathbf{G}^{-1}(t) - \tilde{Y}(t)\mathbf{W}(t) \Rightarrow 2\tilde{Y}(t-1)\mathbf{G}(t) \\
 &\quad - 2(-\tilde{Y}(t) + \tilde{Y}(t-))\mathbf{W}(t-).
 \end{aligned}
 \tag{46}$$

From (42), $\mathbf{G}(t)$ can be written as

$$\mathbf{G}(t) = \gamma^{t-1} \mathbf{G}(1) + \sum_{s=0}^{t-1} \gamma^{t-s-1} \mathbf{W}(s) \mathbf{G}^{-1}(s).
 \tag{47}$$

For $t \geq 2$, (A5) gives the following inequality

$$\mathbf{I} \left(\dots \right)$$

Consider the set $\mathcal{I}_\gamma = [\tilde{\theta}(t) : (1 - \gamma)(\theta - 1) \leq \eta_{\max}, \text{ a. s.}]$, and define \mathcal{I}_γ^c as the complementary set of \mathcal{I}_γ . Thus, for $\tilde{\theta}(t) \in \mathcal{I}_\gamma$, there exists $\delta(\tilde{\theta}(t)) > 0$. Applying martingale hyperconvergence theorem of [Lemma 1](#), it is known that

$$\lim_{t \rightarrow \infty} \tilde{\theta}(t) \in \mathcal{I}_\gamma, \text{ a. s.} \tag{54}$$

Therefore, we have

$$\lim_{t \rightarrow \infty} \tilde{\theta}(t) \leq \frac{\sigma_v^2}{(1 - \gamma)\alpha} + \frac{\gamma\sigma^2}{(1 - \gamma)} + \frac{\beta\sigma^2}{(1 - \gamma)^2}. \tag{55}$$

Then the parameter estimation error $\tilde{\theta}(t)$ satisfies

$$\begin{aligned} \lim_{t \rightarrow \infty} \|\tilde{\theta}(t)\|^2 &\leq \frac{\sigma_v^2}{(1 - \gamma)\alpha^2} + \frac{\gamma\sigma^2}{\alpha(1 - \gamma)} + \frac{\beta\sigma^2}{\alpha(1 - \gamma)^2} \\ &= \frac{1}{\alpha(1 - \gamma)} \left(\frac{\sigma_v^2}{\alpha} + \frac{\gamma\sigma^2}{1 - \gamma} + \frac{\beta\sigma^2}{1 - \gamma} \right). \end{aligned} \tag{56}$$

This proves [Theorem 1](#).

In [Theorem 1](#), the assumptions (A1)–(A4) specify conditions of both noise and parameter variation. The assumptions (A5) and (A6) show both persistent excitation property and boundedness of information vectors. By assuming that the information vector $\mathcal{W}(t)$ is known, [Theorem 1](#) shows the bounded convergence of time-varying Wiener output-error systems.

Corollary 1. Under (4)–(34), (35), (36), $\hat{\theta}(t) = \theta - \tilde{\theta}(t)(t - 1)$, (37)–(39), (A1)–(A2), (A3)–(A4), (A5)–(A6), i.e., $\alpha I \leq \hat{\mathcal{W}}(t)W(t)\hat{\mathcal{W}}(t) \leq \beta I$, $0 \leq \|\hat{\theta}(t)\|^2 \leq \dots < \infty$, (40) holds.

Proof. The proof of this corollary is similar to the proof of [Theorem 1](#), and is thus omitted.

Remark 2. By selecting the suitable value of $\tilde{\theta}$, $\hat{\theta}(t)$ and $\hat{\theta}(t)$ of Sage-Husa noise estimator can record noise changes correctly. The nonlinear gain $K(t)$ is sensitive to noise changes, and the adaptive Kalman filter in (26)–(33) can approximate the Kalman filter in [26]. In fact, under uniform complete observability, [26] proves that the error dynamics equation of Kalman filter is asymptotically stable for linear time-varying output-error systems. Meanwhile, recursive identification based on the Gauss-Newton technique can provide more and more accurate parameter estimates. Thus, with the improvement of parameter estimation accuracy, the proposed adaptive filter based on Sage-Husa principle can provide more accurate estimate of $\hat{\theta}(t)$. This leads to uncorrelation between $\hat{\theta}(t)$ and process data. Therefore, the conditions of [Corollary 1](#) are easy to be satisfied.

Remark 3. Both the theorem and the corollary in this paper are the extensions of convergence results in [35] and [36]. Specifically, the asymptotic convergence of linear time-invariant systems is given in [35], and the bounded convergence of linear time-varying systems is obtained in [36].

5. Numerical simulation

In this section, the time-varying Wiener output-error systems are considered for identification. Three examples are used to demonstrate the effectiveness of proposed AF-RI method. Meanwhile, by replacing unknown $y_1(t-1)$ in $y(t)$ with measurements $y_1(t-1)$, the proposed RI method in (12)–(20) is also implemented. These two identification methods are used to illustrate the following aspects: (i) the effectiveness of proposed recursive identification; (ii) the better performance of AF-RI method; (iii) the effectiveness of dealing with unknown noise statistics. Here, the same system structure is used for [examples 1 and 2](#), and it is taken the following form

$$y_1(t) = \sum_{i=1}^2 a_i(t-1) \sum_{j=1}^3 b_j(t-1) (y_1(t-1)) + \sum_{i=1}^2 c_i(t-1) y_1(t-1), \quad (57)$$

$$y(t) = y_1(t) + v(t), \quad (58)$$

$$y(t) = [y_1(t), y_2(t), y_3(t), y_4(t), y_5(t)]^T, \quad (59)$$

where the vector $[y_1(t-1), y_2(t-1), y_3(t-1)]^T$ contains nonlinear basis functions, and the vector $\theta(t)$ denotes the parameters of system. In simulation tests, the attenuation speed $\bar{\lambda} = \tilde{\lambda} = 0.9$, the sample number $N = 20$, and the

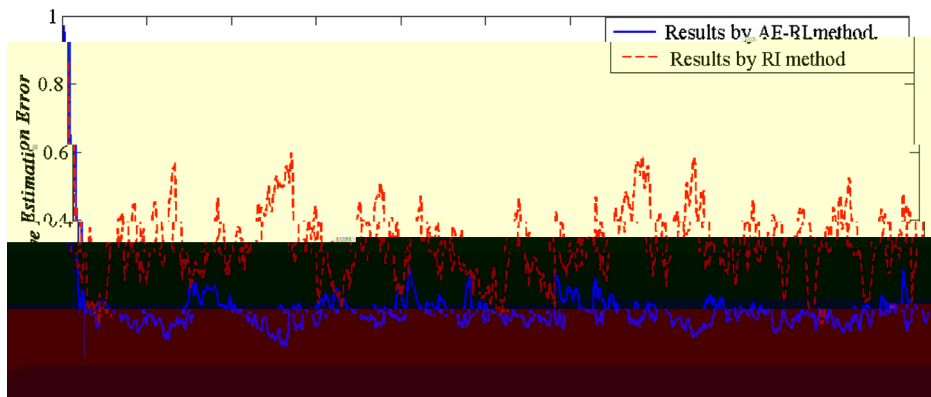


Fig. 4. The relative estimation error of parameters in Situation 1 of Example 1.

- i) The relative estimation errors are fluctuating.
- ii) AF-RI method can generate more accurate estimations of system parameters.
- iii) AF-RI method generates the more accurate model to track time-varying system.

Situation 2. Consider the following slow time-varying Wiener output-error system

$$y_1(t) = \sum_{i=1}^2 \beta_i [\beta_{1i} y_1(t - \tau_i) + \beta_{2i} (\beta_{1i} y_1(t - \tau_i)) + \beta_{3i} \cdot (\beta_{1i} y_1(t - \tau_i))] + \sum_{i=1}^2 \beta_i v_1(t - \tau_i), \quad (64)$$

$$y_1(t) = \beta_1(t) + v_1(t), \quad (65)$$

$$= \begin{bmatrix} \beta_1, \beta_2, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{21}, \beta_{22}, \beta_{23} \\ 0.25 + 0.0001 * t \\ 0.28 + 0.0001 * t \\ -0.3 - 0.0001 * t \\ 1.0 + 0.0001 * t \\ 1.0 \\ -0.5 + 0.0001 * t \\ -0.3317 + 0.0002 * t \end{bmatrix}, \quad (66)$$

where the parameters to be identified change slowly and linearly. In order to ensure uniqueness of identification, the parameter β_1 of nonlinear part is equal to 1, and does not participate in identification. The inputs $\{v_1(t)\}$ are taken as persistent excitation signals with zero mean and unit variance, and the noise $\{v_1(t)\}$ remains same as in (63). Thus, the assumptions (A1)–(A4) are approximately satisfied. With $N = 2000$, RI and AF-RI methods are applied to estimate parameters $\{\beta_1, \beta_2, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{21}, \beta_{22}, \beta_{23}\}$. The performances of tracking time-varying parameters are shown in Figs. 5–6. From the figures, the following conclusions are got

- i) The tracking performances of $\{\beta_1, \beta_2, \beta_{11}, \beta_{12}\}$ of linear part are poor by RI method.
- ii) The convergence speed of $\{\beta_{21}, \beta_{23}\}$ of nonlinear part is slow by RI method.
- iii) The convergence speed of parameter estimation is faster by AF-RI method.
- iv) The better performances of tracking time-varying parameters are got by AF-RI method.

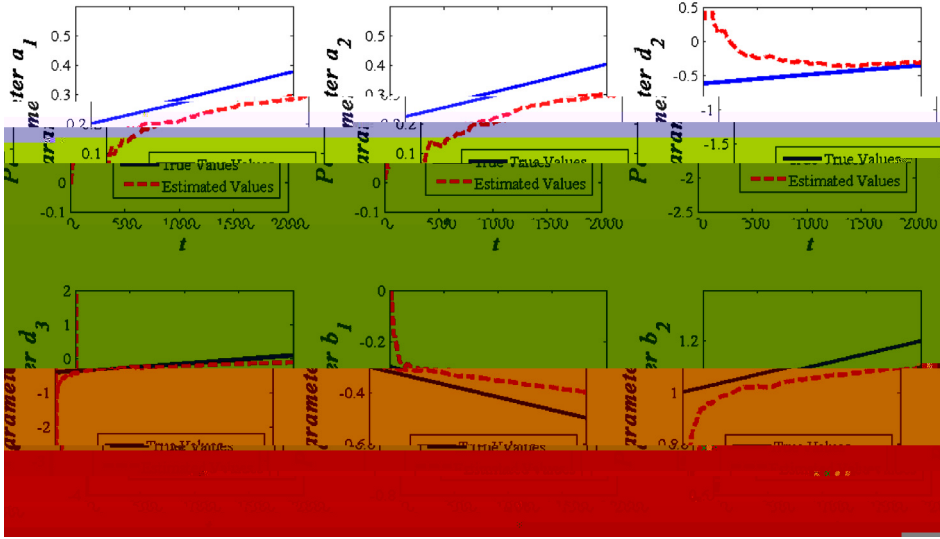


Fig. 5. The performances of tracking time-varying parameters by RI method.

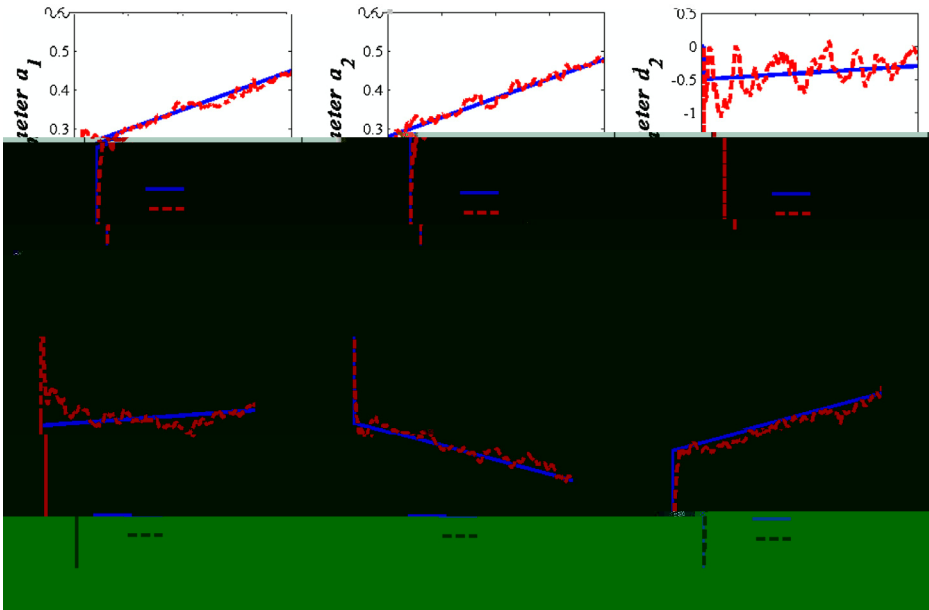


Fig. 6. The performances of tracking time-varying parameters by AF-RI method.

Example 2. In this example, robustness of identification is tested under different noise environments. The system structure adopts the same form described in (64)–(66). Both RI and AF-RI methods are applied into identification. The inputs $\{r(t)\}$ are taken as persistent excitation signals with zero mean and unit variance. Besides, the following root mean square

Table 1
Mean values and standard deviations of RMSE under colored impulsive noise.

Noise	Algorithms	RMSE
$v_2(t)$	RI	0.4714 ± 0.0071
	AF-RI	0.1937 ± 0.0064

error (RMSE) index is taken to evaluate robustness

$$\text{RMSE} = \sqrt{\sum_{t=1}^N (\hat{y}_1(t) - \hat{y}(t))^2 / N}, \quad (67)$$

where $\hat{y}_1(t)$ denotes the actual noise-free output,

proposes AF-RI method to achieve robust identification for time-varying Wiener output-error systems. Meanwhile, this paper tries to solve the problems below.

- The appropriate attenuation speed determines the sensitivity of proposed Sage-Husa Kalman filter, and the sensitive filter can keep track of noise changes.
- The virtual equivalent state space model is constructed to implement adaptive filtering.
- The unknown variables in information vectors are associated with optimal state estimates.
- The bounded convergence for time-varying nonlinear systems is exploited.

The proposed adaptive filtering-based recursive method can also be extended to the identification of other time-varying nonlinear systems under unknown noise statistics.

Acknowledgements

This work is supported by National Natural Science Foundation of China (No. 61703434) and Science Foundation of China University of Petroleum, Beijing (No. 2462016YJRC028).

References

- [1] M. Schoukens, K. Tiels, Identification of block-oriented nonlinear systems starting from linear approximations: a survey, *Automatica* 85 (2017) 272–292.
- [2] D. Wang, F. Ding, Parameter estimation algorithms for multivariable Hammerstein CARMA systems, *Inform. Sciences* 355-356 (2016) 237–248.
- [3] A. Wills, B. Ninness, Generalised Hammerstein-Wiener system estimation and a benchmark application, *Control Eng. Pract.* 20 (11) (2012) 1097–1108.
- [4] T. Katayama, H. Ase, Linear approximation and identification of MIMO Wiener-Hammerstein systems, *Automatica* 71 (2016) 118–124.
- [5] Q. Jin, J. Dou, F. Ding, L. Cao, A novel identification method for Wiener systems with the limited information, *Math. Comput. Model.* 58 (7–8) (2013) 1531–1539.
- [6] M.M.D. Silva, T. Wigren, T. Mendonca, Nonlinear identification of a minimal neuromuscular blockade model in anesthesia, *IEEE Trans. Control Syst. Technol.* 20 (1) (2012) 181–188.
- [7] G. Shafiee, M.M. Arefi, M.R. Jahed-Motlagh, A.A. Jalali, Nonlinear predictive control of a polymerization reactor based on piecewise linear Wiener model, *Chem. Eng. J.* 143 (1–3) (2008) 282–292.
- [8] Y. Zhang, B. Fidan, P.A. Ioannou, Backstepping control of linear time-varying systems with known and unknown parameters, *IEEE Trans. Autom. Control* 48 (11) (2003) 1908–1925.
- [9] B. Ni, Q. Zhang, Stability of the Kalman filter for continuous time11)11)1908–1925likelihood 7766/52611 U 6.6.3760 09

- [18] Z. Wang, Q. Jin, X. Liu, Recursive least squares identification of hybrid Box-Jenkins model structure in open-loop and closed-loop, *J. Franklin I.* 353 (2) (2016) 265–278.
- [19] F. Ding, X. Liu, M. Liu, The recursive least squares identification algorithm for a class of Wiener nonlinear systems, *J. Franklin I.* 353 (7) (2016) 1518–1526.
- [20] Z. Xu, Y. Zhu, K. Han, J. Qian, A multi-iteration pseudo-linear regression method and an adaptive disturbance model for MPC, *J. Process Contr.* 20 (4) (2010) 384–395.
- [21] S. Cheng, Y. Wei, D. Sheng, Y. Chen, Y. Wang, Identification for Hammerstein nonlinear ARMAX systems based on multi-innovation fractional order stochastic gradient, *Signal Process.* 142 (2018) 1–10.
- [22] F. Ding, P.X. Liu, G. Liu, Multi-innovation least-squares identification for system modeling, *IEEE Trans. Syst. Man Cybern. B* 40 (3) (2010) 767–778.
- [23] J. Wang, J. Wang, D. Zhang, X. Shao, G. Chen, Kalman filtering through the feedback adaption of prior error covariance, *Signal Process* 152 (2018) 47–53.
- [24] D. Marelli, M. Zamani, M. Fu, B. Ninness, Distributed Kalman filter in a network of linear systems, *Syst. Control Lett.* 116 (2018) 71–77.
- [25] M. Bohner, N. Wintz, The Kalman filter for linear systems on time scales, *J. Math. Anal. Appl.* 406 (2) (2013) 419–436.
- [26] Q. Zhang, On stability of the Kalman filter for discrete time output error systems, *Syst. Control Lett.* 107 (2017) 84–91.
- [27] J. Ma, O. Wu, B. Huang, F. Ding, Expectation maximization estimation for a class of input nonlinear state space systems by using the Kalman smoother, *Signal Process* 145 (2018) 295–303.
- [28] J. Ma, W. Xiong, J. Chen, F. Ding, Hierarchical identification for multivariate Hammerstein systems by using the modified Kalman filter, *Digit. Signal Process* 11 (6) (2017) 857–869.
- [29] W. Liu, H.W. He, F.C. Sun, Vehicle state estimation based on minimum model error criterion combining with extended Kalman filter, *J. Franklin I.* 353 (4) (2016) 834–856.
- [30] Q. Zhang, Adaptive Kalman filter for actuator fault diagnosis, *Automatica* 93 (1) (2018) 333–342.
- [31] Y. Wu, X. Luo, A novel calibration approach of soft sensor based on multi-rate data fusion technology, *J. Process Contr.* 10 (2010) 1252–1260.
- [32] A. Sage, G. Husa, Adaptive filtering with unknown prior statistics, in: *The 10th Joint Automatic Control Conference (JACC 1969)*, 1969, pp. 760–769.
- [33] C. Chen, Z. Liu, K. Xie, Y. Zhang, C.L. Philip