IET Control Theory & Applications

Table of Contents

22 September 2011
Volume 5, Issue 14, p. 1577-1675

---

www.ietdl.org

Published in IET Control Theory and Applications
Received on 21st July 2010
Revised on 21st February 2011
doi: 10.1049/iet-cta.2010.0416

ISSN 1751-8644

Brief Paper

Least squares-based recursive and iterative estimation for output error moving average systems using data filtering

D.Q. Wang

College of Automation Engineering, Qingdao University, Qingdao 266071, Peoples' Republic of China
E-mail: dqwang64@163.com

Abstract: For parameter estimation of output error moving average (OEMA) systems, this study combines the auxiliary model identification idea with the filtering theory, transforms an OEMA system into two identification models and presents a filtering and auxiliary model-based recursive least squares (F-AM-RLS) identification algorithm. Compared with the auxiliary model-based recursive extended least squares algorithm, the proposed F-AM-RLS algorithm has a high computational efficiency. Moreover, a filtering and auxiliary model-based least squares iterative (F-AM-LSI) identification algorithm is derived for OEMA systems with finite measurement input–output data. Compared with the F-AM-RLS approach, the proposed F-AM-LSI algorithm updates the parameter estimation using all the available data at each iteration, and thus can generate highly accurate parameter estimates.

1 Introduction

Much work on the identification of output error (OE)-type systems has been reported [1–9], including OE systems, output error moving average (OEMA) systems and Box–Jenkins systems etc. For example, Forssell and Ljung estimated parameters of OE model structures of unstable systems by reparameterising the OE model structures to guarantee stability [1]; Söderström *et al.* statistically analysed the estimation accuracy of the time-domain maximum likelihood method and the sample maximum likelihood method for OE models [2]; Rosenqvist and Karlström derived a prediction-error minimisation method for estimating parameters of piecewise-linear OE models [3]; Zheng used the bias compensation least squares algorithms to identify OE systems [4, 5], and also Zhang and Cui presented the bias compensation recursive least squares identification algorithm for OE systems with coloured noises [6]. Based on the uniformly and non-uniformly sampled data, Gillberg and Ljung presented different approaches to obtain the frequency-domain estimation of continuous-time transfer functions by frequency function approximation and B-spline output approximation [7, 8]; Zhu *et al.* proposed an OE method to estimate the parameters of this kind of system [9]; Ding and Ding developed a least squares parameter estimation method with irregularly missing data [10].

This paper considers identification problems for a class of OE systems – the OEMA systems [i.e. the OE system with moving average (MA) noises]. In this literature, much early work exists, for example, Dugard and Landau derived several recursive parametric identification algorithms for OE models using the model reference adaptive system techniques in detail and analysed their performances in the deterministic and stochastic environment using the equivalent feedback representation and ordinary differential equation methods, respectively, [11]; Landau and Karimi studied the recursive OE identification for closed-loop systems and made the stability and convergence analysis in the deterministic and stochastic environments [12, 13]. Recently, based on the OE method, Ding and Chen presented an auxiliary model idea and developed a recursive combined parameter and intersample output estimation for dual-rate systems [14, 15] and Ding *et al.* proposed the multi-innovation least squares/stochastic gradient identification methods using the auxiliary model [16–18]. The auxiliary model identification idea is very useful in the identification of OE-type systems. The auxiliary model-based least squares algorithms have fast convergence rate but need computing the covariance matrices, leading to a large computational burden. The auxiliary model-based stochastic gradient algorithms require less computational load but have slow convergence rate.

The filtering technique have been used in many fields, such as fault detection [19], parameter estimation [20, 21], ground target tracking [22], switched systems [23] etc. The objective of this paper is to study data filtering and auxiliary model-based least squares algorithms for OEMA systems by combining the filtering technique [24] with the auxiliary model identification idea [14, 15] to estimate the system parameters from available input–output data and to evaluate the accuracy of the parameter estimates by simulations on computers.

This paper extends the identification algorithms [14, 15] from a simple OE system with white noises to an OEMA system with MA noises. The basic idea is, by means of filtering input–output data of an OEMA system with a linear filter and by the variable substitution, to obtain two identification models: an OE model and a MA model, and then to present a filtering and auxiliary model-based recursive least squares (F-AM-RLS) identification algorithm and a filtering and auxiliary model-based least squares iterative (F-AM-LSI) identification algorithm for OEMA systems. Compared with the auxiliary model-based recursive extended least squares (AM-RELS) algorithm, the proposed F-AM-RLS algorithm has a high computational efficiency, because the dimensions of the covariance matrices of the decomposed OE and MA models become smaller than that of the original OEMA model. Compared with the F-AM-RLS approach, the proposed F-AM-LSI algorithm updates the parameter estimation using all the available data at each iteration, and thus can produce highly accurate parameter estimation.

The paper is organised as follows. Section 2 gives an auxiliary model-based least squares algorithm for the OEMA system. Section 3 derives a data filtering and auxiliary model-based recursive least squares algorithm for the OEMA system. Section 4 presents an F-AM-LSI algorithm. Section 5 provides an illustrative example for the results in this paper. Finally, concluding remarks are given in Section 6.

2 Auxiliary model-based least squares algorithm

Consider an OEMA system in Fig. 1 [25]

$$y(t) = \frac{B(z)}{A(z)}u(t) + D(z)v(t) \tag{1}$$

where $u(t)$ and $y(t)$ are the system input and output, respectively, $v(t)$ is a stochastic white noise with zero mean and variance σ^2, $A(z)$, $B(z)$ and $D(z)$ are polynomials in the unit backward shift operator z^{-1}, and defined by

$$A(z) = 1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_{n_a}z^{-n_a}$$

$$B(z) = b_1z^{-1} + b_2z^{-2} + \cdots + b_{n_b}z^{-n_b}$$

$$D(z) = 1 + d_1z^{-1} + d_2z^{-2} + \cdots + d_{n_d}z^{-n_d}$$

Assume that the degrees n_a, n_b and n_d are known, and $n{:} = n_a + n_b + n_d$, $y(t) = 0$, $u(t) = 0$ and $v(t) = 0$ for $t \le 0$. The disturbance $e(t){:} = D(z)v(t)$ is an MA process.

The objective of this paper is to present new identification algorithms to estimate the parameters a_i, b_i and d_i from available input–output data $\{u(t), y(t)\}$.

For comparison purposes, we simply give the auxiliary model-based recursive least squares algorithm in this section.

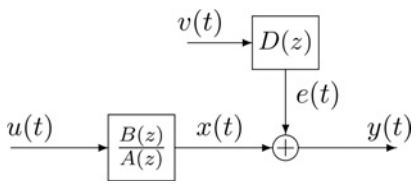

Fig. 1 *OE systems with coloured noises*

Define the intermediate variable

$$x(t) := \frac{B(z)}{A(z)}u(t) \tag{2}$$

and the parameter vector $\boldsymbol{\theta}$ and the information vector $\boldsymbol{\varphi}(t)$ as

$$\boldsymbol{\theta} := \begin{bmatrix} \boldsymbol{\theta}_{\mathrm{s}} \\ \boldsymbol{\theta}_{\mathrm{n}} \end{bmatrix} \in \mathbb{R}^{n_a+n_b+n_d}$$

$$\boldsymbol{\theta}_{\mathrm{s}} := [a_1, a_2, \ \ldots, a_{n_a}, b_1, b_2, \ \ldots, b_{n_b}]^{\mathrm{T}} \in \mathbb{R}^{n_a+n_b}$$

$$\boldsymbol{\theta}_n := [d_1, d_2, \ \ldots, d_{n_d}]^{\mathrm{T}} \in \mathbb{R}^{n_d}$$

$$\boldsymbol{\varphi}(t) := \begin{bmatrix} \boldsymbol{\varphi}_{\mathrm{s}}(t) \\ \boldsymbol{\varphi}_{\mathrm{n}}(t) \end{bmatrix} \in \mathbb{R}^{n_a+n_b+n_d}$$

$$\begin{aligned}\boldsymbol{\varphi}_{\mathrm{s}}(t) := [&-x(t-1), \ -x(t-2), \ \ldots, \ -x(t-n_a), u(t-1), \\ &u(t-2), \ \ldots, u(t-n_b)]^{\mathrm{T}} \in \mathbb{R}^{n_a+n_b}\end{aligned}$$

$$\varphi_n(t) := [v(t-1), v(t-2), \ \ldots, v(t-n_d)]^{\mathrm{T}} \in \mathbb{R}^{n_d}$$

Equations (2) and (1) can be written as

$$x(t) = \boldsymbol{\varphi}_{\mathrm{s}}^{\mathrm{T}}(t)\boldsymbol{\theta}_{\mathrm{s}} \tag{3}$$

$$\begin{aligned} y(t) &= x(t) + D(z)v(t) \\ &= \boldsymbol{\varphi}_{\mathrm{s}}^{\mathrm{T}}(t)\boldsymbol{\theta}_{\mathrm{s}} + \boldsymbol{\varphi}_{\mathrm{n}}^{\mathrm{T}}(t)\boldsymbol{\theta}_{\mathrm{n}} + v(t) \\ &= [\boldsymbol{\varphi}_{\mathrm{s}}^{\mathrm{T}}(t), \boldsymbol{\varphi}_{\mathrm{n}}^{\mathrm{T}}(t)]\begin{bmatrix} \boldsymbol{\theta}_{\mathrm{s}} \\ \boldsymbol{\theta}_{\mathrm{n}} \end{bmatrix} + v(t) \\ &= \varphi^{\mathrm{T}}(t)\boldsymbol{\theta} + v(t) \end{aligned} \tag{4}$$

Equation (4) is the identification model for the OEMA system.

Let $\boldsymbol{I}$ be an identity matrix of appropriate size. Minimising the cost function

$$J(\theta) := \sum_{i=1}^{t} [y(i) - \varphi^{\mathrm{T}}(i)\theta]^2$$

gives the following recursive algorithm of estimating $\boldsymbol{\theta}$ [26]

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \boldsymbol{L}(t)[y(t) - \boldsymbol{\varphi}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}(t-1)] \tag{5}$$

$$\boldsymbol{L}(t) = \frac{\boldsymbol{P}(t-1)\varphi(t)}{1 + \varphi^{\mathrm{T}}(t)\boldsymbol{P}(t-1)\varphi(t)} \tag{6}$$

$$\boldsymbol{P}(t) = [\boldsymbol{I} - \boldsymbol{L}(t)\varphi^{\mathrm{T}}(t)]\boldsymbol{P}(t-1), \quad \boldsymbol{P}(0) = p_0\boldsymbol{I} \tag{7}$$

However, the algorithm in (5)–(7) is impossible to realise because the information vector $\boldsymbol{\varphi}(t)$ on the right-hand sides contains the unknown true outputs $x(t-i)$ and noise terms $v(t-i)$. The solution here is based on the auxiliary model identification idea [14, 15]: these unknown variables $x(t-i)$ and $v(t-i)$ in $\boldsymbol{\varphi}(t)$ are replaced with the outputs $x_a(t-i)$ of an auxiliary model (or reference model) and estimated residual $\hat{v}(t-i)$, respectively. Define

$$\hat{\boldsymbol{\varphi}}_{\rm s}(t) := [-x_a(t-1), -x_a(t-2), \ldots, -x_a(t-n_a), u(t-1), u(t-2), \ldots, u(t-n_b)]^{\rm T} \in \mathbb{R}^{n_a+n_b}$$

$$\hat{\boldsymbol{\varphi}}_{\rm n}(t) := [\hat{v}(t-1), \hat{v}(t-2), \ldots, \hat{v}(t-n_d)]^{\rm T} \in \mathbb{R}^{n_d}$$

$$\hat{\boldsymbol{\varphi}}(t) := [\hat{\boldsymbol{\varphi}}_{\rm s}^{\rm T}(t), \hat{\boldsymbol{\varphi}}_{\rm n}^{\rm T}(t)]^{\rm T} \in \mathbb{R}^{n_a+n_b+n_d}$$

Let

$$\hat{\theta}(t) := \begin{bmatrix} \hat{\theta}_s(t) \\ \hat{\theta}_n(t) \end{bmatrix}$$

be the estimate of

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_s \\ \boldsymbol{\theta}_n \end{bmatrix}$$

at time t. If $\boldsymbol{\varphi}_{\rm s}(t)$ and $\boldsymbol{\theta}_s$ in (3) are replaced with the estimates $\hat{\boldsymbol{\varphi}}_{\rm s}(t)$ and $\hat{\boldsymbol{\theta}}_{\rm s}(t)$, respectively, then the outputs $x_a(t)$ of the auxiliary model or estimated outputs can be computed by

$$x_a(t) = \hat{\boldsymbol{\varphi}}_{\rm s}^{\rm T}(t)\hat{\boldsymbol{\theta}}_{\rm s}(t)$$

Similarly, $\boldsymbol{\varphi}(t)$ and $\boldsymbol{\theta}$ in (4) are replaced with the estimates $\hat{\boldsymbol{\varphi}}(t)$ and $\hat{\boldsymbol{\theta}}(t)$, respectively, then the estimated residual $\hat{v}(t)$ can be computed by

$$\hat{v}(t) = y(t) - \hat{\boldsymbol{\varphi}}^{\rm T}(t)\hat{\boldsymbol{\theta}}(t)$$

Note that $\hat{\boldsymbol{\varphi}}(t)$ is known at time t. Replacing $\boldsymbol{\varphi}(t)$ in (5)–(7) with $\hat{\boldsymbol{\varphi}}(t)$ leads to the following AM-RELS algorithm for identifying $\boldsymbol{\theta}$ in (4) [27]

$$\hat{\boldsymbol{\theta}}(t) = \hat{\boldsymbol{\theta}}(t-1) + \boldsymbol{L}(t)[y(t) - \hat{\boldsymbol{\varphi}}^{\rm T}(t)\hat{\boldsymbol{\theta}}(t-1)] \tag{8}$$

$$\boldsymbol{L}(t) = \frac{\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)}{1 + \hat{\boldsymbol{\varphi}}^{\rm T}(t)\boldsymbol{P}(t-1)\hat{\boldsymbol{\varphi}}(t)} \tag{9}$$

$$\boldsymbol{P}(t) = [\boldsymbol{I} - \boldsymbol{L}(t)\hat{\boldsymbol{\varphi}}^{\rm T}(t)]\boldsymbol{P}(t-1), \quad \boldsymbol{P}(0) = p_0\boldsymbol{I} \tag{10}$$

$$\hat{\boldsymbol{\theta}}(t) = \begin{bmatrix} \hat{\boldsymbol{\theta}}_{\rm s}(t) \\ \hat{\boldsymbol{\theta}}_{\rm n}(t) \end{bmatrix}, \quad \hat{\boldsymbol{\varphi}}(t) = \begin{bmatrix} \hat{\boldsymbol{\varphi}}_{\rm s}(t) \\ \hat{\boldsymbol{\varphi}}_{\rm n}(t) \end{bmatrix} \tag{11}$$

$$\hat{\boldsymbol{\varphi}}_{\rm s}(t) = [-x_a(t-1), -x_a(t-2), \ldots, -x_a(t-n_a), u(t-1), u(t-2), \ldots, u(t-n_b)]^{\rm T} \tag{12}$$

$$\hat{\boldsymbol{\varphi}}_n(t) = [\hat{v}(t-1), \hat{v}(t-2), \ldots, \hat{v}(t-n_d)]^{\rm T} \tag{13}$$

$$x_a(t) = \hat{\boldsymbol{\varphi}}_{\rm s}^{\rm T}(t)\hat{\boldsymbol{\theta}}_{\rm s}(t) \tag{14}$$

$$\hat{v}(t) = y(t) - \hat{\boldsymbol{\varphi}}^{\rm T}(t)\hat{\boldsymbol{\theta}}(t) \tag{15}$$

3 Data filtering-based least squares algorithm

For the OEMA model in (1), if the input–output data $\{u(t), y(t)\}$ are filtrated by an estimated rational function $\hat{D}^{-1}(t, z)$ a (an estimated linear filter), then the model in (1) will become an 'output error model'. By using the auxiliary model identification idea [14, 15], the unknown filtered and no-filtered true outputs are replaced with the outputs of auxiliary models, and then the recursive least squares algorithm can be applied.

Multiplying both sides of (1) and (2) by $[1/D(z)]$ gives

$$\frac{1}{D(z)}y(t) = \frac{B(z)}{A(z)}\frac{1}{D(z)}u(t) + v(t) \tag{16}$$

$$\frac{1}{D(z)}x(t) = \frac{B(z)}{A(z)}\frac{1}{D(z)}u(t) \tag{17}$$

Define the filtered input $u_{\rm f}(t)$, the filtered true output $x_{\rm f}(t)$ and the filtered output $y_{\rm f}(t)$ as

$$u_{\rm f}(t) := \frac{1}{D(z)}u(t), \quad x_{\rm f}(t) := \frac{1}{D(z)}x(t), \quad y_{\rm f}(t) := \frac{1}{D(z)}y(t) \tag{18}$$

Then we have

$$y_{\rm f}(t) = \frac{B(z)}{A(z)}u_{\rm f}(t) + v(t) \tag{19}$$

$$x_{\rm f}(t) = \frac{B(z)}{A(z)}u_{\rm f}(t) \tag{20}$$

Define the related information vector

$$\boldsymbol{\varphi}_{\rm f}(t) := [-x_{\rm f}(t-1), -x_{\rm f}(t-2), \ldots, -x_{\rm f}(t-n_a), u_{\rm f}(t-1), u_{\rm f}(t-2), \ldots, u_{\rm f}(t-n_b)]^{\rm T} \in \mathbb{R}^{n_a+n_b}$$

Equation (20) can be written as

$$x_{\rm f}(t) = \boldsymbol{\varphi}_{\rm f}^{\rm T}(t)\boldsymbol{\theta}_{\rm s} \tag{21}$$

Define the inner variable

$$e(t) := D(z)v(t) \tag{22}$$

then we have

$$e(t) = \boldsymbol{\varphi}_n^{\rm T}(t)\boldsymbol{\theta}_n + v(t) \tag{23}$$

From (1)–(3) and (19)–(23), we have

$$\begin{aligned} y(t) &= x(t) + e(t) \\ &= \boldsymbol{\varphi}_{\rm s}^{\rm T}(t)\boldsymbol{\theta}_{\rm s} + e(t) \end{aligned} \tag{24}$$

$$= \boldsymbol{\varphi}_s^{\rm T}(t)\boldsymbol{\theta}_s + \boldsymbol{\varphi}_n^{\rm T}(t)\boldsymbol{\theta}_n + v(t) \tag{25}$$

$$y_{\rm f}(t) = x_{\rm f}(t) + v(t) = \boldsymbol{\varphi}_{\rm f}^{\rm T}(t)\boldsymbol{\theta}_s + v(t) \tag{26}$$

Applying the least squares principle to (23) and (26) gives the two least squares algorithms for estimating $\boldsymbol{\theta}_{\rm s}$ and $\boldsymbol{\theta}_{\rm n}$ as

IET Control Theory Appl., 2011, Vol. 5, Iss. 14, pp. 1648–1657
doi: 10.1049/iet-cta.2010.0416

follows

$$\hat{\boldsymbol{\theta}}_s(t) = \hat{\boldsymbol{\theta}}_s(t-1) + \boldsymbol{L}_{\mathrm{f}}(t)[y_{\mathrm{f}}(t) - \boldsymbol{\varphi}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_s(t-1)] \tag{27}$$

$$\boldsymbol{L}_{\mathrm{f}}(t) = \frac{\boldsymbol{P}_{\mathrm{f}}(t-1)\boldsymbol{\varphi}_{\mathrm{f}}(t)}{1 + \boldsymbol{\varphi}_{\mathrm{f}}^{\mathrm{T}}(t)\boldsymbol{P}_{\mathrm{f}}(t-1)\boldsymbol{\varphi}_{\mathrm{f}}(t)} \tag{28}$$

$$\boldsymbol{P}_{\mathrm{f}}(t) = [\boldsymbol{I} - \boldsymbol{L}_{\mathrm{f}}(t)\boldsymbol{\varphi}_{\mathrm{f}}^{\mathrm{T}}(t)]\boldsymbol{P}_{\mathrm{f}}(t-1) \tag{29}$$

$$\hat{\boldsymbol{\theta}}_n(t) = \hat{\boldsymbol{\theta}}_n(t-1) + \boldsymbol{L}_n(t)[e(t) - \boldsymbol{\varphi}_n^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_n(t-1)] \tag{30}$$

$$\boldsymbol{L}_n(t) = \frac{\boldsymbol{P}_n(t-1)\boldsymbol{\varphi}_n(t)}{1 + \boldsymbol{\varphi}_n^{\mathrm{T}}(t)\boldsymbol{P}_n(t-1)\boldsymbol{\varphi}_n(t)} \tag{31}$$

$$\boldsymbol{P}_n(t) = [\boldsymbol{I} - \boldsymbol{L}_n(t)\boldsymbol{\varphi}_n^{\mathrm{T}}(t)]\boldsymbol{P}_n(t-1) \tag{32}$$

The polynomial $D(z)$ is unknown, and so the filtered variables $u_{\mathrm{f}}(t-i)$ and $x_{\mathrm{f}}(t-i)$ in $\boldsymbol{\varphi}_{\mathrm{f}}(t)$, and $y_{\mathrm{f}}(t)$ are unknown, and $x(t-i)$ in $\boldsymbol{\varphi}_{\mathrm{s}}(t)$, $v(t-i)$ in $\boldsymbol{\varphi}_{\mathrm{n}}(t)$ and $e(t)$ are also unknown. Here, we adopt the idea of replacing the unknown $x(t-i)$ and $x_{\mathrm{f}}(t-i)$ with the outputs of auxiliary models, and unknown $u_{\mathrm{f}}(t-i)$, $y_{\mathrm{f}}(t)$, $v(t-i)$ and $e(t)$ with their estimates, to derive a filtering-based identification algorithm.

Since the transfer function relating $u(t)$ to $x(t)$ is the same as that from $u_{\mathrm{f}}(t)$ to $x_{\mathrm{f}}(t)$, that is

$$u_{\mathrm{f}}(t) := \frac{1}{D(z)}u(t), \quad x_{\mathrm{f}}(t) := \frac{1}{D(z)}x(t)$$

we use an identity auxiliary model $P_a(z)$ to generate the estimates $\hat{x}_{\mathrm{f}}(t)$ and $x_a(t)$ in the following. Construct two identical auxiliary models to obtain the estimates $x_a(t)$ of $x(t)$ and $\hat{x}_{\mathrm{f}}(t)$ of $x_{\mathrm{f}}(t)$ from $u(t)$ and $u_{\mathrm{f}}(t)$, respectively, shown in Fig. 2, where $P_a(z) := \{B_a(z)/A_a(z)\}$ is the transfer function of auxiliary models [$A_a(z)$ and $B_a(z)$ are the polynomials of same orders as $A(z)$ and $B(z)$], $\hat{x}_{\mathrm{f}}(t)$ and $x_a(t)$ are the outputs of the auxiliary models with and without filtering, respectively. The unknown $x_{\mathrm{f}}(t-i)$ in $\boldsymbol{\varphi}_{\mathrm{f}}(t)$ and $x(t-i)$ in $\boldsymbol{\varphi}_{\mathrm{s}}(t)$ are replaced with the outputs $\hat{x}_{\mathrm{f}}(t-i)$ and $x_a(t-i)$ of the auxiliary models, respectively, and the unmeasurable term $u_{\mathrm{f}}(t-i)$ in $\boldsymbol{\varphi}_{\mathrm{f}}(t)$ is replaced with its estimate $\hat{u}_{\mathrm{f}}(t-i)$, $v(t-i)$ and $e(t)$ are replaced with their estimates $\hat{v}(t-i)$ and $\hat{e}(t)$, respectively. Define the information vectors

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t) := [-\hat{x}_{\mathrm{f}}(t-1), -\hat{x}_{\mathrm{f}}(t-2), \ldots, -\hat{x}_{\mathrm{f}}(t-n_a), \hat{u}_{\mathrm{f}}(t-1), \hat{u}_{\mathrm{f}}(t-2), \ldots, \hat{u}_{\mathrm{f}}(t-n_b)]^{\mathrm{T}}$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{s}}(t) := [-x_a(t-1), -x_a(t-2), \ldots, -x_a(t-n_a), u(t-1), u(t-2), \ldots, u(t-n_b)]^{\mathrm{T}}$$

$$\hat{\boldsymbol{\varphi}}_n(t) := [\hat{v}(t-1), \hat{v}(t-2), \ldots, \hat{v}(t-n_d)]^{\mathrm{T}}$$

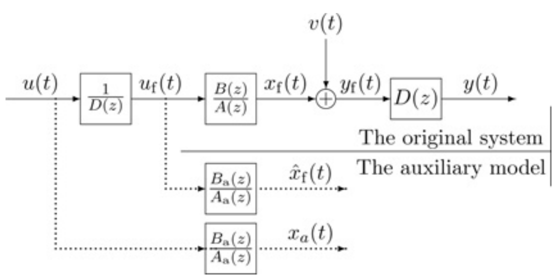

Fig. 2 *Stochastic system with auxiliary models*

From Fig. 2, we obtain

$$x_a(t) = \frac{B_a(z)}{A_a(z)}u(t), \quad \hat{x}_{\mathrm{f}}(t) = \frac{B_a(z)}{A_a(z)}u_{\mathrm{f}}(t)$$

$$x_a(t) = \boldsymbol{\varphi}_a^{\mathrm{T}}(t)\boldsymbol{\theta}_a, \quad \hat{x}_{\mathrm{f}}(t) = \boldsymbol{\varphi}_{\mathrm{f}a}^{\mathrm{T}}(t)\boldsymbol{\theta}_a$$

where $\boldsymbol{\varphi}_a(t)$ and $\boldsymbol{\varphi}_{\mathrm{f}a}(t)$ are the no-filtered and filtered information vectors, and $\boldsymbol{\theta}_a$ is the parameter vector of the auxiliary models. There are several ways of choosing the information vectors and the parameter vector of the auxiliary models [14, 15]. Here, we take $\hat{\boldsymbol{\varphi}}_s(t)$ and $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$ as the information vectors $\boldsymbol{\varphi}_a(t)$ and $\boldsymbol{\varphi}_{\mathrm{f}a}(t)$ of the auxiliary models, and $\hat{\boldsymbol{\theta}}_{\mathrm{s}}(t)$ as the parameter vector $\boldsymbol{\theta}_a$ of the auxiliary models, and thus we have

$$x_a(t) := \hat{\boldsymbol{\varphi}}_s^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_s(t) \tag{33}$$

$$\hat{x}_{\mathrm{f}}(t) := \hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_s(t) \tag{34}$$

Using $\hat{\boldsymbol{\varphi}}_s(t)$ at instant t and $\hat{\boldsymbol{\theta}}_s(t-1)$ at instant $t-1$ to replace $\boldsymbol{\varphi}_{\mathrm{s}}(t)$ and $\boldsymbol{\theta}_{\mathrm{s}}$. From (24), we obtain

$$\hat{e}(t) = y(t) - \hat{\boldsymbol{\varphi}}_s^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_s(t-1)$$

Using $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$, $\hat{\boldsymbol{\varphi}}_n(t)$ and $\hat{y}_{\mathrm{f}}(t)$ to replace $\boldsymbol{\varphi}_{\mathrm{f}}(t)$, $\boldsymbol{\varphi}_n(t)$ and $y_{\mathrm{f}}(t)$; Let $\hat{\boldsymbol{\theta}}_s(t)$ and $\hat{\boldsymbol{\theta}}_n(t)$ be the estimates of $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_{\mathrm{n}}$. From (23) and (26), we obtain

$$\hat{v}(t) = \hat{e}(t) - \hat{\boldsymbol{\varphi}}_n^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_n(t), \quad \text{or} \quad \hat{v}(t) = \hat{y}_{\mathrm{f}}(t) - \hat{\boldsymbol{\varphi}}_{\mathrm{f}}^{\mathrm{T}}(t)\hat{\boldsymbol{\theta}}_s(t)$$

Using the parameter estimates of the noise model

$$\hat{\boldsymbol{\theta}}_n(t) = [\hat{d}_1(t), \hat{d}_2(t), \ldots, \hat{d}_{n_d}(t)]^{\mathrm{T}}$$

to construct the estimate of $D(z)$

$$\hat{D}(t, z) = 1 + \hat{d}_1(t)z^{-1} + \hat{d}_2(t)z^{-2} + \cdots + \hat{d}_{n_d}(t)z^{-n_d}$$

Filtering $u(t)$ and $y(t)$ with $\hat{D}^{-1}(t, z)$ to obtain the estimates of $u_{\mathrm{f}}(t)$ and $y_{\mathrm{f}}(t)$ as follows

$$\hat{u}_{\mathrm{f}}(t) := \hat{D}^{-1}(t, z)u(t)$$

$$\hat{y}_{\mathrm{f}}(t) := \hat{D}^{-1}(t, z)y(t)$$

From the above equations, we can recursively compute $\hat{u}_{\mathrm{f}}(t)$ and $\hat{y}_{\mathrm{f}}(t)$ by the following equations

$$\hat{u}_{\mathrm{f}}(t) = u(t) - \hat{d}_1(t)\hat{u}_{\mathrm{f}}(t-1) - \hat{d}_2(t)\hat{u}_{\mathrm{f}}(t-2) - \cdots - \hat{d}_{n_d}(t)\hat{u}_{\mathrm{f}}(t-n_d)$$

$$\hat{y}_{\mathrm{f}}(t) = y(t) - \hat{d}_1(t)\hat{y}_{\mathrm{f}}(t-1) - \hat{d}_2(t)\hat{y}_{\mathrm{f}}(t-2) - \cdots - \hat{d}_{n_d}(t)\hat{y}_{\mathrm{f}}(t-n_d)$$

Replacing the unknown information vector $\boldsymbol{\varphi}_{\mathrm{f}}(t)$ in (27)–(29) with its estimate $\hat{\boldsymbol{\varphi}}_{\mathrm{f}}(t)$, $\boldsymbol{\varphi}_n(t)$ in (30)–(32) with $\hat{\boldsymbol{\varphi}}_n(t)$, and the unknown filtered output $y_{\mathrm{f}}(t)$ in (27) and the noise term $e(t)$ in (30) with their estimates $\hat{y}_{\mathrm{f}}(t)$ and $\hat{e}(t)$, respectively, we obtain the F-AM-RLS algorithm of estimating the

doi: 10.1049/iet-cta.2010.0416

parameter vectors $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_n$ for the OEMA system

$$\hat{\boldsymbol{\theta}}_s(t) = \hat{\boldsymbol{\theta}}_s(t-1) + \boldsymbol{L}_f(t)[\hat{y}_f(t) - \hat{\boldsymbol{\varphi}}_f^T(t)\hat{\boldsymbol{\theta}}_s(t-1)] \quad (35)$$

$$\boldsymbol{L}_f(t) = \frac{\boldsymbol{P}_f(t-1)\hat{\boldsymbol{\varphi}}_f(t)}{1 + \hat{\boldsymbol{\varphi}}_f^T(t)\boldsymbol{P}_f(t-1)\hat{\boldsymbol{\varphi}}_f(t)} \quad (36)$$

$$\boldsymbol{P}_f(t) = [\boldsymbol{I} - \boldsymbol{L}_f(t)\hat{\boldsymbol{\varphi}}_f^T(t)]\boldsymbol{P}_f(t-1), \quad \boldsymbol{P}_f(0) = p_0\boldsymbol{I} \quad (37)$$

$$\hat{\boldsymbol{\varphi}}_f(t) = [-\hat{x}_f(t-1), -\hat{x}_f(t-2), \ldots, -\hat{x}_f(t-n_a), \hat{u}_f(t-1), \hat{u}_f(t-2), \ldots, \hat{u}_f(t-n_b)]^T \quad (38)$$

$$\hat{x}_f(t) = \hat{\boldsymbol{\varphi}}_f^T(t)\hat{\boldsymbol{\theta}}_s(t) \quad (39)$$

$$\hat{y}_f(t) = y(t) - \hat{d}_1(t)\hat{y}_f(t-1) - \hat{d}_2(t)\hat{y}_f(t-2) - \cdots - \hat{d}_{n_d}(t)\hat{y}_f(t-n_d) \quad (40)$$

$$\hat{u}_f(t) = u(t) - \hat{d}_1(t)\hat{u}_f(t-1) - \hat{d}_2(t)\hat{u}_f(t-2) - \cdots - \hat{d}_{n_d}(t)\hat{u}_f(t-n_d) \quad (41)$$

$$\hat{\boldsymbol{\theta}}_n(t) = \hat{\boldsymbol{\theta}}_n(t-1) + \boldsymbol{L}_n(t)[\hat{e}(t) - \hat{\boldsymbol{\varphi}}_n^T(t)\hat{\boldsymbol{\theta}}_n(t-1)] \quad (42)$$

$$\boldsymbol{L}_n(t) = \frac{\boldsymbol{P}_n(t-1)\hat{\boldsymbol{\varphi}}_n(t)}{1 + \hat{\boldsymbol{\varphi}}_n^T(t)\boldsymbol{P}_n(t-1)\hat{\boldsymbol{\varphi}}_n(t)} \quad (43)$$

$$\boldsymbol{P}_n(t) = [\boldsymbol{I} - \boldsymbol{L}_n(t)\hat{\boldsymbol{\varphi}}_n^T(t)]\boldsymbol{P}_n(t-1), \quad \boldsymbol{P}_n(0) = p_0\boldsymbol{I} \quad (44)$$

$$\hat{\boldsymbol{\varphi}}_s(t) = [-x_a(t-1), -x_a(t-2), \ldots, -x_a(t-n_a), u(t-1), u(t-2), \ldots, u(t-n_b)]^T \quad (45)$$

$$\hat{\boldsymbol{\varphi}}_n(t) = [\hat{v}(t-1), \hat{v}(t-2), \ldots, \hat{v}(t-n_d)]^T \quad (46)$$

$$\hat{e}(t) = y(t) - \hat{\boldsymbol{\varphi}}_s^T(t)\hat{\boldsymbol{\theta}}_s(t-1) \quad (47)$$

$$x_a(t) = \hat{\boldsymbol{\varphi}}_s^T(t)\hat{\boldsymbol{\theta}}_s(t) \quad (48)$$

$$\hat{v}(t) = \hat{y}_f(t) - \hat{\boldsymbol{\varphi}}_f^T(t)\hat{\boldsymbol{\theta}}_s(t) \quad (49)$$

$$\hat{\boldsymbol{\theta}}_s(t) = [\hat{a}_1(t), \hat{a}_2(t), \ldots, \hat{a}_{n_a}(t), \hat{b}_1(t), \hat{b}_2(t), \ldots, \hat{b}_{n_b}(t)]^T \quad (50)$$

$$\hat{\boldsymbol{\theta}}_n(t) = [\hat{d}_1(t), \hat{d}_2(t), \ldots, \hat{d}_{n_d}(t)]^T \quad (51)$$

To initialise the F-AM-RLS algorithm, we take

$$\hat{\boldsymbol{\theta}}_s(i) = \mathbf{1}_{n_a+n_b}/p_0, \quad \hat{\boldsymbol{\theta}}_n(i) = \mathbf{1}_{n_d}/p_0, \quad i \le 0 \quad (52)$$

$$\boldsymbol{P}_f(0) = p_0\boldsymbol{I}_{n_a+n_b}, \quad \boldsymbol{P}_n(0) = p_0\boldsymbol{I}_{n_d}, \quad p_0 = 10^6 \quad (53)$$

where $\mathbf{1}_n$ represents an n-dimensional column vector whose entries are all 1.

The proposed algorithm named F-AM-RLS can be summarised as

1. Let $t = 1$, set the initial values of the parameter estimation vectors and covariance matrices according to (52) and (53), and $\hat{y}_f(i) = 1/p_0$, $\hat{x}_f(i) = 1/p_0$, $\hat{u}_f(i) = 1/p_0$, $x_a(i) = 1/p_0$, $\hat{e}(i) = 1/p_0$ and $\hat{v}(i) = 1/p_0$ for $i \le 0$.
2. Collect the input–output data $u(t)$ and $y(t)$, construct information vectors $\hat{\boldsymbol{\varphi}}_s(t)$ by (45), $\hat{\boldsymbol{\varphi}}_f(t)$ by (38) and $\hat{\boldsymbol{\varphi}}_n(t)$ by (46).
3. Compute $\hat{e}(t)$ by (47), the gain vector $\boldsymbol{L}_n(t)$ by (43), the covariance matrix $\boldsymbol{P}_n(t)$ by (44).
4. Update the parameter estimate $\hat{\boldsymbol{\theta}}_n(t)$ by (42).
5. Compute $\hat{y}_f(t)$ by (40) and $\hat{u}_f(t)$ by (41).
6. Compute the gain vector $\boldsymbol{L}_f(t)$ by (36) and the covariance matrix $\boldsymbol{P}_f(t)$ by (37).
7. Update the parameter estimate $\hat{\boldsymbol{\theta}}_s(t)$ by (35).
8. Compute $\hat{x}_f(t)$ by (39), $x_a(t)$ by (48) and $\hat{v}(t)$ by (49).
9. Compare

$$\hat{\boldsymbol{\theta}}(t) = \begin{bmatrix} \hat{\boldsymbol{\theta}}_s(t) \\ \hat{\boldsymbol{\theta}}_n(t) \end{bmatrix}$$

with $\hat{\boldsymbol{\theta}}(t-1)$, until a specified fitting threshold of $\varepsilon > 0$ is satisfied with $\|\hat{\boldsymbol{\theta}}(t) - \hat{\boldsymbol{\theta}}(t-1)\| \le \varepsilon$, terminate the procedure, and obtain the parameter estimates $\hat{\boldsymbol{\theta}}(t)$; otherwise, increase t by 1, go to step 2.

Table 1 lists the computation cost of the AM-RELS and F-AM-RLS algorithms at each recursion, where the numbers in the brackets denote the computation cost for $n_a = 5$, $n_b = 5$ and $n_d = 5$. From Table 1, we can see that the F-AM-RLS algorithm requires less computation than the AM-RELS algorithm because the dimensions of the covariance matrices $\boldsymbol{P}_f(t) \in \mathbb{R}^{(n_a+n_b)\times(n_a+n_b)}$ and $\boldsymbol{P}_n(t) \in \mathbb{R}^{n_d\times n_d}$ in the F-AM-RLS algorithm are smaller than those of the covariance matrix $\boldsymbol{P}(t) \in \mathbb{R}^{(n_a+n_b+n_d)\times(n_a+n_b+n_d)}$.

4 Least squares-based iterative algorithm

The F-AM-RLS algorithm can generate the parameter estimation of the OEMA model; however, when computing the parameter estimation vectors $\hat{\boldsymbol{\theta}}_s(t)$ and $\hat{\boldsymbol{\theta}}_n(t)$ at time t $(1 \le t \le L)$, the F-AM-RLS algorithm uses only the measured data $\{u(i), y(i): i = 0, 1, 2, \ldots, t\}$ up to time t, not including the data $\{u(i), y(i): i = t+1, t+2, \ldots, L\}$. That is, the F-AM-RLS algorithm fails to make sufficient use of all the measured information $\{u(i), y(t): i = 1, 2, 3, \ldots, L\}$. In contrast, an iterative algorithm is able to make full use of all the measured data in each iteration so that the parameter estimation accuracy can be greatly improved. Hence, we will investigate the iterative identification approach for the OEMA system.

Suppose that the data length $L \gg n_a + n_b + n_d$. Define the stacked filtered output vector $\boldsymbol{Y}_f(L)$, the inner vector $\boldsymbol{E}(L)$, the noise vector $\boldsymbol{V}(L)$, the information matrices $\boldsymbol{\Phi}_f(L)$ and $\boldsymbol{\Phi}_n(L)$

Table 1 Comparison of computational efficiency ($n := n_a + n_b + n_d$)

Algorithms	Number of multiplications	Number of additions
AM-RELS	$2n^2 + 4n + n_a + n_b$ [520]	$2n^2 + 3n + n_a + n_b - 1$ [504]
F-AM-RLS	$2(n_a + n_b)^2 + 2n_d^2 + 8(n_a + n_b) + 6n_d$ [360]	$2(n_a + n_b)^2 + 2n_d^2 + 5(n_a + n_b) + 3n_d - 2$ [313]

IET Control Theory Appl., 2011, Vol. 5, Iss. 14, pp. 1648–1657
doi: 10.1049/iet-cta.2010.0416

as

$$\boldsymbol{Y}_{\rm f}(L) = \begin{bmatrix} y_{\rm f}(1) \\ y_{\rm f}(2) \\ \vdots \\ y_{\rm f}(L) \end{bmatrix} \in \mathbb{R}^L, \quad \boldsymbol{E}(L) = \begin{bmatrix} e(1) \\ e(2) \\ \vdots \\ e(L) \end{bmatrix} \in \mathbb{R}^L$$
$$\boldsymbol{V}(L) = \begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(L) \end{bmatrix} \in \mathbb{R}^L \tag{54}$$

$$\boldsymbol{\Phi}_{\rm f}(L) = \begin{bmatrix} \boldsymbol{\varphi}_{\rm f}^{\rm T}(1) \\ \boldsymbol{\varphi}_{\rm f}^{\rm T}(2) \\ \vdots \\ \boldsymbol{\varphi}_{\rm f}^{\rm T}(L) \end{bmatrix} \in \mathbb{R}^{L\times(n_a+n_b)},$$
$$\boldsymbol{\Phi}_n(L) = \begin{bmatrix} \boldsymbol{\varphi}_n^{\rm T}(1) \\ \boldsymbol{\varphi}_n^{\rm T}(2) \\ \vdots \\ \boldsymbol{\varphi}_n^{\rm T}(L) \end{bmatrix} \in \mathbb{R}^{L\times n_d} \tag{55}$$

From (23) and (26), we have

$$\boldsymbol{Y}_{\rm f}(L) = \boldsymbol{\Phi}_{\rm f}(L)\boldsymbol{\theta}_s + \boldsymbol{V}(L) \tag{56}$$

$$\boldsymbol{E}(L) = \boldsymbol{\Phi}_n(L)\boldsymbol{\theta}_n + \boldsymbol{V}(L) \tag{57}$$

Notice that $\boldsymbol{V}(L)$ is a white noise vector with zero mean and define two quadratic criterion functions

$$J(\boldsymbol{\theta}_s) = \|\boldsymbol{Y}_{\rm f}(L) - \boldsymbol{\Phi}_{\rm f}(\boldsymbol{L})\boldsymbol{\theta}_s\|^2 \tag{58}$$

$$J(\boldsymbol{\theta}_n) = \|\boldsymbol{E}(L) - \boldsymbol{\Phi}_n(\boldsymbol{L})\boldsymbol{\theta}_n\|^2 \tag{59}$$

Let $k = 1, 2, 3, \ldots$ be an iterative variable and $\hat{\boldsymbol{\theta}}_{s,k}$ and $\hat{\boldsymbol{\theta}}_{n,k}$ denote the estimates of $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_n$ at iteration k, respectively. Provided that $\boldsymbol{\varphi}_{\rm f}(t)$ and $\boldsymbol{\varphi}_{\rm n}(t)$ are persistently exciting, minimising $J(\boldsymbol{\theta}_s)$ and $J(\boldsymbol{\theta}_n)$ in (58) and (59) gives the least squares estimates of $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_n$

$$\hat{\boldsymbol{\theta}}_{s,k} = [\boldsymbol{\Phi}_{\rm f}^{\rm T}(L)\Phi_{\rm f}(L)]^{-1}\boldsymbol{\Phi}_{\rm f}^{\rm T}(L)\boldsymbol{Y}_{\rm f}(L) \tag{60}$$

$$\hat{\boldsymbol{\theta}}_{n,k} = [\boldsymbol{\Phi}_n^{\rm T}(L)\boldsymbol{\Phi}_n(L)]^{-1}\boldsymbol{\Phi}_n^{\rm T}(L)\boldsymbol{E}(L) \tag{61}$$

A problem arises, that is: the polynomial $D(z)$ is unknown, the filtered variables $u_{\rm f}(t-i)$, $x_{\rm f}(t-i)$ and $y_{\rm f}(t)$, and variables $x(t-i)$, $v(t-i)$ and $e(t)$ are unknown, and then $\boldsymbol{\varphi}_{\rm f}(t)$ in $\boldsymbol{\Phi}_{\rm f}(L)$ and $\boldsymbol{\varphi}_{\rm n}(t)$ in $\boldsymbol{\Phi}_{\rm n}(L)$ are unknown, and so the iterative solution $\hat{\boldsymbol{\theta}}_{s,k}$ and $\hat{\boldsymbol{\theta}}_{n,k}$ at iteration k are impossible to compute.

The approach here is based on the interactive estimation theory [28].

Let $\hat{x}_{a,k}(t-i)$ be the estimate of $x(t-i)$ at iteration k, replacing the unknown $\boldsymbol{\varphi}_{\rm s}(t)$ and $\boldsymbol{\theta}_{\rm s}$ in (3) with the estimates $\hat{\boldsymbol{\varphi}}_{s,k}(t)$ and $\hat{\boldsymbol{\theta}}_{s,k}$, the estimate $\hat{x}_{a,k}(t)$ [or $\hat{x}_{a,k}(t-i)$] can be computed by

$$\hat{x}_{a,k}(t) = \hat{\boldsymbol{\varphi}}_{s,k}^{\rm T}(t)\hat{\boldsymbol{\theta}}_{s,k}, \quad \hat{x}_{a,k}(t-i) = \hat{\boldsymbol{\varphi}}_{s,k}^{\rm T}(t-i)\hat{\boldsymbol{\theta}}_{s,k} \tag{62}$$

where $\hat{\boldsymbol{\varphi}}_{s,k}(t)$ denotes the estimate of the information vector $\boldsymbol{\varphi}_{\rm s}(t)$ obtained by replacing $x(t-i)$ in $\boldsymbol{\varphi}_{\rm s}(t)$ with $\hat{x}_{a,k-1}(t-i)$, that is

$$\hat{\boldsymbol{\varphi}}_{s,k}(t) = [-\hat{x}_{a,k-1}(t-1), \ldots, -\hat{x}_{a,k-1}(t-n_a), u(t-1), \ldots, u(t-n_b)]^{\rm T} \in \mathbb{R}^{n_a+n_b} \tag{63}$$

Let $\hat{e}_k(t)$ and $\hat{v}_k(t)$ be the estimates of $e(t)$ and $v(t)$ at iteration k, replacing the unknown $\boldsymbol{\varphi}_s(t)$ and $\boldsymbol{\theta}_s$ in (24) with the estimates $\hat{\boldsymbol{\varphi}}_{s,k}(t)$ at iteration k and $\hat{\boldsymbol{\theta}}_{s,k-1}$ at iteration $k-1$, the estimate $\hat{e}_k(t)$ can be computed by

$$\hat{e}_k(t) = y(t) - \hat{\boldsymbol{\varphi}}_{s,k}^{\rm T}(t)\hat{\boldsymbol{\theta}}_{s,k-1} \tag{64}$$

Replacing $y_{\rm f}(t)$, $\boldsymbol{\varphi}_{\rm f}(t)$ and $\boldsymbol{\theta}_s$ in (26) with $\hat{y}_{{\rm f},k}(t)$, $\hat{\boldsymbol{\varphi}}_{{\rm f},k}(t)$ and $\hat{\boldsymbol{\theta}}_{s,k}$, the estimate $\hat{v}_k(t)$ [or $\hat{v}_k(t-i)$] can be computed by

$$\hat{v}_k(t) = \hat{y}_{{\rm f},k}(t) - \hat{\boldsymbol{\varphi}}_{{\rm f},k}^{\rm T}(t)\hat{\boldsymbol{\theta}}_{s,k},$$
$$\hat{v}_k(t-i) = \hat{y}_{{\rm f},k}(t-i) - \hat{\boldsymbol{\varphi}}_{{\rm f},k}^{\rm T}(t-i)\hat{\boldsymbol{\theta}}_{s,k} \tag{65}$$

Let $\hat{\boldsymbol{\varphi}}_{n,k}(t)$ denotes the estimate of the information vector $\boldsymbol{\varphi}_n(t)$ obtained by replacing $v(t-i)$ in $\boldsymbol{\varphi}_n(t)$ with $\hat{v}_{k-1}(t-i)$, that is

$$\hat{\boldsymbol{\varphi}}_{n,k}(t) = [\hat{v}_{k-1}(t-1), \ldots, \hat{v}_{k-1}(t-n_d)]^{\rm T} \tag{66}$$

Using the parameter estimates of the noise model

$$\hat{\boldsymbol{\theta}}_{n,k} = [\hat{d}_{1,k}, \hat{d}_{2,k}, \ldots, \hat{d}_{n_d,k}]^{\rm T}$$

to construct the estimate of $D(z)$

$$\hat{D}_k(t, z) = 1 + \hat{d}_{1,k}z^{-1} + \hat{d}_{2,k}z^{-2} + \cdots + \hat{d}_{n_d,k}z^{-n_d}$$

Filtering $u(t)$ and $y(t)$ with $\hat{D}_k^{-1}(t, z)$ to obtain the estimates of $u_{\rm f}(t)$ and $y_{\rm f}(t)$ as follows

$$\hat{u}_{{\rm f},k}(t) = \hat{D}_k^{-1}(t, z)u(t), \quad \hat{y}_{{\rm f},k}(t) = \hat{D}_k^{-1}(t, z)y(t)$$

Then $\hat{u}_{{\rm f},k}(t)$ and $\hat{y}_{{\rm f},k}(t)$ can be recursively computed by

$$\hat{u}_{{\rm f},k}(t) = u(t) - \hat{d}_{1,k}\hat{u}_{{\rm f},k}(t-1) - \hat{d}_{2,k}\hat{u}_{{\rm f},k}(t-2) - \cdots - \hat{d}_{n_d,k}\hat{u}_{{\rm f},k}(t-n_d) \tag{67}$$

$$\hat{y}_{{\rm f},k}(t) = y(t) - \hat{d}_{1,k}\hat{y}_{{\rm f},k}(t-1) - \hat{d}_{2,k}\hat{y}_{{\rm f},k}(t-2) - \cdots - \hat{d}_{n_d,k}\hat{y}_{{\rm f},k}(t-n_d) \tag{68}$$

Let $\hat{x}_{{\rm f},k}(t-i)$ be the estimate of $x_{\rm f}(t-i)$ at iteration k, replacing the unknown $\boldsymbol{\varphi}_{\rm f}(t)$ and $\boldsymbol{\theta}_{\rm s}$ in (21) with their estimates $\hat{\boldsymbol{\varphi}}_{{\rm f},k}(t)$ and $\hat{\boldsymbol{\theta}}_{s,k}$ at iteration k, the estimate $\hat{x}_{{\rm f},k}(t)$ [or $\hat{x}_{{\rm f},k}(t-i)$] can be computed by

$$\hat{x}_{{\rm f},k}(t) = \hat{\boldsymbol{\varphi}}_{{\rm f},k}^{\rm T}(t)\hat{\boldsymbol{\theta}}_{s,k}, \quad \hat{x}_{{\rm f},k}(t-i) = \hat{\boldsymbol{\varphi}}_{{\rm f},k}^{\rm T}(t-i)\hat{\boldsymbol{\theta}}_{s,k} \tag{69}$$

where $\hat{\boldsymbol{\varphi}}_{{\rm f},k}(t)$ denotes the the estimate of information vector $\boldsymbol{\varphi}_{\rm f}(t)$ obtained by replacing $x_{\rm f}(t-i)$ and $u_{\rm f}(t-i)$ in $\boldsymbol{\varphi}_{\rm f}(t)$ with $\hat{x}_{{\rm f},k-1}(t-i)$ at iteration $k-1$ and $\hat{u}_{{\rm f},k}(t-i)$ at iteration k,

doi: 10.1049/iet-cta.2010.0416

that is

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t) = [-\hat{x}_{\mathrm{f},k-1}(t-1),\ \cdots,\ -\hat{x}_{\mathrm{f},k-1}(t-n_a),\ \hat{u}_{\mathrm{f},k}(t-1),\ \ldots,\ \hat{u}_{\mathrm{f},k}(t-n_b)]^{\mathrm{T}} \quad (70)$$

Define the following stacked vectors and matrices (at iteration k) as

$$\hat{\boldsymbol{Y}}_{\mathrm{f},k}(L) = \begin{bmatrix} \hat{y}_{\mathrm{f},k}(1) \\ \hat{y}_{\mathrm{f},k}(2) \\ \vdots \\ \hat{y}_{\mathrm{f},k}(L) \end{bmatrix}, \quad \hat{\boldsymbol{E}}_k(L) = \begin{bmatrix} \hat{e}_k(1) \\ \hat{e}_k(2) \\ \vdots \\ \hat{e}_k(L) \end{bmatrix} \quad (71)$$

$$\hat{\boldsymbol{\Phi}}_{\mathrm{f},k}(L) = \begin{bmatrix} \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},k}(1) \\ \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},k}(2) \\ \vdots \\ \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},k}(L) \end{bmatrix}, \quad \hat{\boldsymbol{\Phi}}_{n,k}(L) = \begin{bmatrix} \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{n,k}(1) \\ \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{n,k}(2) \\ \vdots \\ \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{n,k}(L) \end{bmatrix} \quad (72)$$

Replacing the unknown information vector $\boldsymbol{\Phi}_{\mathrm{f}}(L)$ and $\boldsymbol{Y}_{\mathrm{f}}(L)$ in (60) with $\hat{\boldsymbol{\Phi}}_{\mathrm{f},k}(L)$ and $\hat{\boldsymbol{Y}}_{\mathrm{f},k}(L)$, and $\boldsymbol{\Phi}_{\mathrm{n}}(L)$ and $\boldsymbol{E}(L)$ in (61) with $\hat{\boldsymbol{\Phi}}_{n,k}(L)$ and $\hat{\boldsymbol{E}}_k(L)$, we obtain the F-AM-LSI algorithm of estimating the parameter vectors $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_n$ for the OEMA system

$$\hat{\boldsymbol{\theta}}_{s,k} = [\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{\mathrm{f},k}(L)\hat{\boldsymbol{\Phi}}_{\mathrm{f},k}(L)]^{-1}\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{\mathrm{f},k}(L)\hat{\boldsymbol{Y}}_{\mathrm{f},k}(L) \quad (73)$$

$$\hat{\boldsymbol{\theta}}_{n,k} = [\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{n,k}(L)\hat{\boldsymbol{\Phi}}_{n,k}(L)]^{-1}\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{n,k}(L)\hat{\boldsymbol{E}}_k(L) \quad (74)$$

From (62)–(74), we can summarise the F-AM-LSI algorithm as follows

$$\hat{\boldsymbol{\theta}}_{s,k} = [\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{\mathrm{f},k}(L)\hat{\boldsymbol{\Phi}}_{\mathrm{f},k}(L)]^{-1}\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{\mathrm{f},k}(L)\hat{\boldsymbol{Y}}_{\mathrm{f},k}(L) \quad (75)$$

$$\hat{\boldsymbol{Y}}_{\mathrm{f},k}(L) = \begin{bmatrix} \hat{y}_{\mathrm{f},k}(1) \\ \hat{y}_{\mathrm{f},k}(2) \\ \vdots \\ \hat{y}_{\mathrm{f},k}(L) \end{bmatrix}, \quad \hat{\boldsymbol{\Phi}}_{\mathrm{f},k}(L) = \begin{bmatrix} \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},k}(1) \\ \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},k}(2) \\ \vdots \\ \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},k}(L) \end{bmatrix} \quad (76)$$

$$\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t) = [-\hat{x}_{\mathrm{f},k-1}(t-1),\ \ldots,\ -\hat{x}_{\mathrm{f},k-1}(t-n_a),\ \hat{u}_{\mathrm{f},k}(t-1),\ \ldots,\ \hat{u}_{\mathrm{f},k}(t-n_b)]^{\mathrm{T}} \quad (77)$$
$$t = 1, 2,\ \ldots,\ L$$

$$\hat{u}_{\mathrm{f},k}(t) = u(t) - \hat{d}_{1,k}\hat{u}_{\mathrm{f},k}(t-1) - \hat{d}_{2,k}\hat{u}_{\mathrm{f},k}(t-2) - \cdots - \hat{d}_{n_d,k}\hat{u}_{\mathrm{f},k}(t-n_d) \quad (78)$$

$$\hat{y}_{\mathrm{f},k}(t) = y(t) - \hat{d}_{1,k}\hat{y}_{\mathrm{f},k}(t-1) - \hat{d}_{2,k}\hat{y}_{\mathrm{f},k}(t-2) - \cdots - \hat{d}_{n_d,k}\hat{y}_{\mathrm{f},k}(t-n_d) \quad (79)$$

$$\hat{x}_{\mathrm{f},k}(t) = \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},k}(t)\hat{\boldsymbol{\theta}}_{s,k} \quad (80)$$

$$\hat{\boldsymbol{\theta}}_{n,k} = [\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{n,k}(L)\hat{\boldsymbol{\Phi}}_{n,k}(L)]^{-1}\hat{\boldsymbol{\Phi}}^{\mathrm{T}}_{n,k}(L)\hat{\boldsymbol{E}}_k(L) \quad (81)$$

$$\hat{\boldsymbol{E}}_k(L) = \begin{bmatrix} \hat{e}_k(1) \\ \hat{e}_k(2) \\ \vdots \\ \hat{e}_k(L) \end{bmatrix}, \quad \hat{\boldsymbol{\Phi}}_{\mathrm{n},k}(L) = \begin{bmatrix} \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{n,k}(1) \\ \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{n,k}(2) \\ \vdots \\ \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{n,k}(L) \end{bmatrix} \quad (82)$$

$$\hat{\boldsymbol{\varphi}}_{n,k}(t) = [\hat{v}_{k-1}(t-1),\ \ldots,\ \hat{v}_{k-1}(t-n_d)]^{\mathrm{T}} \quad (83)$$

$$\hat{\boldsymbol{\varphi}}_{s,k}(t) = [-\hat{x}_{a,k-1}(t-1),\ \ldots,\ -\hat{x}_{a,k-1}(t-n_a),\ u(t-1),\ \ldots,\ u(t-n_b)]^{\mathrm{T}} \quad (84)$$

$$\hat{x}_{a,k}(t) = \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{s,k}(t)\hat{\boldsymbol{\theta}}_{s,k} \quad (85)$$

$$\hat{e}_k(t) = y(t) - \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{s,k}(t)\hat{\boldsymbol{\theta}}_{s,k-1} \quad (86)$$

$$\hat{v}_k(t) = \hat{y}_{\mathrm{f},k}(t) - \hat{\boldsymbol{\varphi}}^{\mathrm{T}}_{\mathrm{f},k}(t)\hat{\boldsymbol{\theta}}_{s,k} \quad (87)$$

$$\hat{\boldsymbol{\theta}}_{s,k} = [\hat{a}_{1,k},\ \hat{a}_{2,k},\ \ldots,\ \hat{a}_{n_a,k},\ \hat{b}_{1,k},\ \hat{b}_{2,k},\ \ldots,\ \hat{b}_{n_b,k}]^{\mathrm{T}} \quad (88)$$

$$\hat{\boldsymbol{\theta}}_{n,k} = [\hat{d}_{1,k},\ \hat{d}_{2,k},\ \ldots,\ \hat{d}_{n_d,k}]^{\mathrm{T}} \quad (89)$$

In order to obtain highly accurate parameter estimates, the data length L should be large and is at least far greater that the number of the parameters.

To summarise, we list the steps involved in the F-AM-LSI algorithm to compute $\hat{\boldsymbol{\theta}}_{s,k}$ and $\hat{\boldsymbol{\theta}}_{n,k}$ as k increases

1. To initialise, let $k = 1$, $\hat{\boldsymbol{\theta}}_{s,0} = \mathbf{1}_n/p_0$, $\hat{\boldsymbol{\theta}}_{n,0} = \mathbf{1}_n/p_0$, $\hat{y}_{f,0}(t) = 1/p_0$, $\hat{x}_{\mathrm{f},0}(t) = 1/p_0$, $\hat{u}_{\mathrm{f},0}(t) = 1/p_0$, $\hat{x}_{\mathrm{a},0}(t) = 1/p_0$, $\hat{e}_0(t) = 1/p_0$, $\hat{v}_0(t) = 1/p_0$, $p_0 = 10^6$.
2. Collect the input/output data $\{u(i), y(i)\colon i = 1, 2, \ldots, L\}$, form $\hat{\boldsymbol{\varphi}}_{s,k}(t)$ by (84), $\hat{\boldsymbol{\varphi}}_{n,k}(t)$ by (83) and $\hat{\boldsymbol{\Phi}}_{n,k}(L)$ by (82). Compute $\hat{\boldsymbol{e}}_k(t)$ by (86), and form $\hat{\boldsymbol{E}}_k(L)$ by (82).
3. Update the parameter estimate $\hat{\boldsymbol{\theta}}_{n,k}$ by (81).
4. Compute $\hat{u}_{\mathrm{f},k}(t)$ and $\hat{y}_{\mathrm{f},k}(t)$ by (78) and (79), form $\hat{\boldsymbol{\varphi}}_{\mathrm{f},k}(t)$ by (77), $\hat{\boldsymbol{Y}}_{\mathrm{f},k}(L)$ and $\hat{\boldsymbol{\Phi}}_{\mathrm{f},k}(L)$ by (76).
5. Update the parameter estimate $\hat{\boldsymbol{\theta}}_{s,k}$ by (75).
6. Compute $\hat{x}_{a,k}(t)$ by (85), $\hat{x}_{\mathrm{f},k}(t)$ by (80), and $\hat{v}_k(t)$ by (87).
7. Compare

$$\hat{\boldsymbol{\theta}}_k = \begin{bmatrix} \hat{\boldsymbol{\theta}}_{s,k} \\ \hat{\boldsymbol{\theta}}_{n,k} \end{bmatrix}$$

with $\hat{\boldsymbol{\theta}}_{k-1}$: if they are sufficiently close, or for some pre-set small ε, if

$$\|\hat{\boldsymbol{\theta}}_k - \hat{\boldsymbol{\theta}}_{k-1}\| \le \varepsilon$$

then terminate the procedure and obtain the iterative times k and estimate $\hat{\boldsymbol{\theta}}_k$; otherwise, increment k by 1 and go to step 2.

Like the identification approaches for OE systems in [11], the proposed recursive and iterative algorithms for OEMA systems with coloured noises are convergent under certain conditions which are difficult to find, for example, the positive reality assumptions [26]. The iterative algorithm

IET Control Theory Appl., 2011, Vol. 5, Iss. 14, pp. 1648–1657
doi: 10.1049/iet-cta.2010.0416

Table 2 AM-RELS and F-AM-RLS estimates and their errors ($\sigma^2 = 0.40^2$, $\delta_{ns} = 50.48\%$)

Algorithms	t	$\hat{a}_1(t)$	$\hat{a}_2(t)$	$\hat{b}_1(t)$	$\hat{b}_2(t)$	$\hat{d}_1(t)$	δ (%)
AM-RELS	100	0.61056	0.27102	0.37097	−0.54178	−0.74116	9.95379
	200	0.52384	0.26909	0.43163	−0.57088	−0.73660	10.26228
	500	0.57169	0.34867	0.46827	−0.56241	−0.68327	9.57333
	1000	0.56809	0.32953	0.45643	−0.56005	−0.73773	5.79080
	2000	0.57312	0.34168	0.45749	−0.56900	−0.76409	3.91788
F-AM-RLS	100	0.56921	0.31505	0.43924	−0.56372	−0.76931	4.58094
	200	0.55912	0.33073	0.45043	−0.57322	−0.76002	5.06792
	500	0.57110	0.33764	0.46012	−0.57485	−0.70070	8.43092
	1000	0.57269	0.33001	0.45421	−0.56456	−0.74214	5.38721
	2000	0.58180	0.33957	0.45369	−0.56384	−0.76707	3.25982
true values		0.60000	0.35000	0.45000	−0.55000	−0.80000	–

uses the finite data length for parameter estimation, and the estimation errors cannot converge to zero as k increases. A large number of simulations indicate that the estimation errors of the iterative algorithm converge to a constant for large data length and the fluctuation of the estimation errors is caused for large k mainly owing to the stationarity of noise.

5 Example

An example is given to demonstrate the effectiveness of the proposed algorithm. Consider an OEMA system

$$y(t) = \frac{B(z)}{A(z)}u(t) + D(z)v(t)$$

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} = 1 + 0.60z^{-1} + 0.35z^{-2}$$

$$B(z) = b_1 z^{-1} + b_2 z^{-2} = 0.45z^{-1} - 0.55z^{-2}$$

$$D(z) = 1 + d_1 z^{-1} = 1 - 0.80z^{-1}$$

$$\begin{aligned}\theta &= [a_1, a_2, b_1, b_2, d_1]^{\mathrm{T}} \\ &= [0.60, 0.35, 0.45, -0.55, -0.80]^{\mathrm{T}}\end{aligned}$$

The input $\{u(t)\}$ is taken as an uncorrelated persistent excitation signal sequence with zero mean and unit variance, and $\{v(t)\}$ as a white noise sequence with zero mean and variance $\sigma^2 = 0.40^2$, and the noise-to-signal ratio is $\delta_{ns} = 50.48\%$. Applying the AM-RELS and the F-AM-RLS algorithms to estimate the parameters of the system, the parameter estimates and their errors are shown in Table 2 and the estimation errors $\delta := \| \hat{\theta}(t) - \theta \| / \| \theta \|$ against t are shown in Fig. 3. Applying the F-AM-LSI algorithms to estimate the parameters of the system, the parameter estimates and their errors are shown in Table 3 and the estimation errors $\delta_k := \| \hat{\theta}_k - \theta \| / \| \theta \|$ against k with different lengths are shown in Fig. 4. The data length L depends on the number of the gathered data and is generally far greater than the number of the parameters of the system. Here, we take $L = 1000$ and $L = 2000$ for the iterative algorithms, respectively.

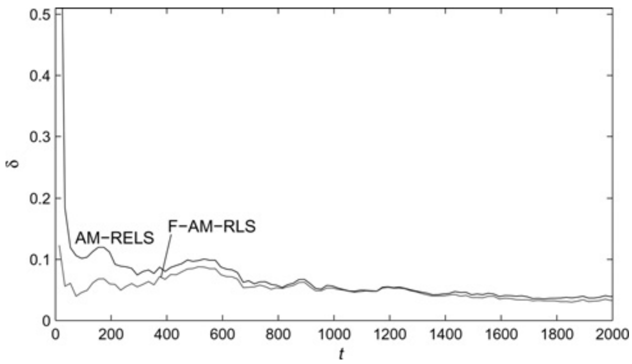

Fig. 3 *AM-RELS and F-AM-RLS estimation errors δ against t ($\sigma^2 = 0.40^2$)*

Table 3 F-AM-LSI estimates and their errors ($\sigma^2 = 0.40^2$, $\delta_{ns} = 50.48\%$)

Data length	k	$\hat{a}_{1,k}$	$\hat{a}_{2,k}$	$\hat{b}_{1,k}$	$\hat{b}_{2,k}$	$\hat{d}_{1,k}$	δ_k (%)
$L = 1000$	1	−0.03247	0.18246	0.44914	−0.85840	0.03600	86.65453
	2	0.41982	−0.00620	0.45826	−0.64587	−0.30296	50.53624
	5	0.58445	0.33321	0.45968	−0.55886	−0.79125	2.17830
	10	0.58869	0.33474	0.45901	−0.55392	−0.78622	1.99434
	20	0.58870	0.33476	0.45900	−0.55389	−0.78620	1.99271
	30	0.58870	0.33476	0.45900	−0.55389	−0.78620	1.99271
$L = 2000$	1	−0.06836	0.22739	0.45397	−0.82308	−0.00142	84.93439
	2	0.43155	−0.03136	0.45312	−0.63245	−0.34056	49.05898
	5	0.59488	0.35310	0.45396	−0.55569	−0.81461	1.35175
	10	0.59927	0.35307	0.45392	−0.54952	−0.81071	0.92833
	20	0.59934	0.35319	0.45392	−0.54944	−0.81070	0.93023
	30	0.59934	0.35319	0.45392	−0.54944	−0.81070	0.93023
true values		0.60000	0.35000	0.45000	−0.55000	−0.80000	–

doi: 10.1049/iet-cta.2010.0416

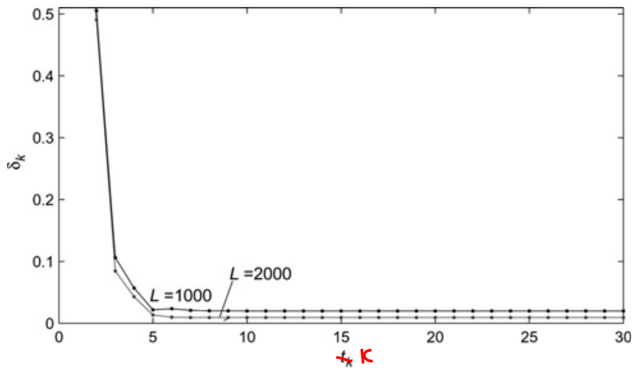

Fig. 4 *F-AM-LSI estimation errors δ_k against k ($\sigma^2 = 0.40^2$)*

From Tables 2 and 3, and Figs. 3 and 4, we can draw the following conclusions:

- The parameter estimates of the F-AM-LSI algorithm converge fast compared with the AM-RELS and F-AM-RLS algorithms, and only needs about five iterations so as to obtain high accurate estimates see Tables 2 and 3. The fast convergence rates partly compensate for the computational load.
- The parameter estimation errors become (generally) smaller and smaller with the data length t increasing. This shows that the proposed F-AM-RLS algorithm is effective.

6 Conclusions

Combining the auxiliary model identification idea with the filtering technique, this paper derives the F-AM-RLS algorithm and the F-AM-LSI algorithm for OEMA systems. Compared with the the AM-RELS algorithm, the F-AM-RLS algorithm requires less computational effort and the F-AM-LSI algorithm has a fast convergence rate and only needs about five iterations so as to obtain highly accurate estimates. The proposed methods can be extended to dual-rate or multirate systems [29–35], non-stationary or non-uniformly sampled-data systems [36–39] and non-linear systems [33, 40, 41].

7 Acknowledgments

This research was supported by the Shandong Provincial Natural Science Foundation (ZR2010FM024), the Shandong Province Higher Educational Science and Technology Program (J10LG12) and the Postdoctoral Innovation Program Foundation of Shandong Province of China (201002002).

8 References

1 Forssell, U., Ljung, L.: 'Identification of unstable systems using output error and Box–Jenkins model structures', *IEEE Trans. Autom. Control*, 2000, **45**, (1), pp. 137–141

2 Söderström, T., Hong, M., Schoukens, J., Pintelon, R.: 'Accuracy analysis of time domain maximum likelihood method and sample maximum likelihood method for errors-in-variables and output error identification', *Automatica*, 2010, **46**, (4), pp. 721–727

3 Rosenqvist, F., Karlström, A.: 'Realisation and estimation of piecewise-linear output error models', *Automatica*, 2005, **41**, (3), pp. 545–551

4 Zheng, W.X.: 'A bias correction method for identification of linear dynamic errors-in-variables models', *IEEE Trans. Autom. Control*, 2002, **47**, (7), pp. 1142–1147

5 Zheng, W.X.: 'On indirect identification of feedback-control systems via the instrumental variables methods', *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, 2003, **50**, (9), pp. 1232–1238

6 Zhang, Y., Cui, G.M.: 'Bias compensation methods for stochastic systems with colored noise', *Appl. Math. Model.*, 2011, **35**, (4), pp. 1709–1716

7 Gillberg, J., Ljung, L.: 'Frequency domain identification of continuous-time output error models, part I: uniformly sampled data and frequency function approximation', *Automatica*, 2010, **46**, (1), pp. 1–10

8 Gillberg, J., Ljung, L.: 'Frequency domain identification of continuous-time output error models, part II: non-uniformly sampled data and B-spline output approximation', *Automatica*, 2010, **46**, (1), pp. 11–18

9 Zhu, Y.C., Telkamp, H., Wang, J.H., Fu, Q.L.: 'System identification using slow and irregular output samples', *J. Process Control*, 2009, **19**, (1), pp. 58–67

10 Ding, F., Ding, J.: 'Least squares parameter estimation with irregularly missing data', *Int. J. Adapt. Control Signal Process.*, 2010, **24**, (7), pp. 540–553

11 Dugard, L., Landau, I.D.: 'Recursive output error identification algorithms theory and evaluation', *Automatica*, 1980, **16**, (5), pp. 443–462

12 Landau, I.D., Karimi, A.: 'An output error recursive algorithm for unbiased identification in closed loop', *Automatica*, 1997, **33**, (5), pp. 933–938

13 Landau, I.D., Karimi, A.: 'Recursive algorithms for identification in closed loop: a unified approach and evaluation', *Automatica*, 1997, **33**, (8), pp. 1499–1523

14 Ding, F., Chen, T.: 'Combined parameter and output estimation of dual-rate systems using an auxiliary model', *Automatica*, 2004, **40**, (10), pp. 1739–1748

15 Ding, F., Chen, T.: 'Parameter estimation of dual-rate stochastic systems by using an output error method', *IEEE Trans. Autom. Control*, 2005, **50**, (9), pp. 1436–1441

16 Ding, F., Liu, P.X., Liu, G.: 'Auxiliary model based multi-innovation extended stochastic gradient parameter estimation with colored measurement noises', *Signal Process.*, 2009, **89**, (10), pp. 1883–1890

17 Ding, F., Liu, P.X., Liu, G.: 'Multi-innovation least squares identification for system modelling', *IEEE Trans. Syst. Man Cybern. B, Cybern.*, 2010, **40**, (3), pp. 767–778

18 Xie, L., Liu, Y.J., Yang, H.Z., Ding, F.: 'Modeling and identification for non-uniformly periodically sampled-data systems', *IET Control Theory Appl.*, 2010, **4**, (5), pp. 784–794

19 Mao, Z., Jiang, B., Shi, P.: 'H_∞ fault detection filter design for networked control systems modelled by discrete Markovian jump systems', *IET Control Theory Appl.*, 2007, **1**, (5), pp. 1336–1343

20 Li, P., Goodall, R., Kadirkamanathan, V.: 'Estimation of parameters in a linear state space model using a Rao-Blackwellised particle filter', *IET Control Theory Appl.*, 2004, **151**, (6), pp. 727–738

21 Qiu, H.Z., Zhang, H.Y., Sun, X.F.: 'Solution of two-stage Kalman filter', *IET Control Theory Appl.*, 2005, **152**, (2), pp. 152–156

22 Cong, S., Hong, L., Layne, J.R.: 'Iterative robust filtering for ground target tracking', *IET Control Theory Appl.*, 2007, **1**, (1), pp. 372–380

23 Qiu, J., Feng, G., Yang, J.: 'Robust mixed H_2/H_∞ filtering design for discrete-time switched polytopic linear systems', *IET Control Theory Appl.*, 2008, **2**, (5), pp. 420–430

24 Wang, D.Q., Ding, F.: 'Input–output data filtering based recursive least squares identification for CARARMA systems', *Digit. Signal Process.*, 2010, **20**, (4), pp. 991–999

25 Ding, F., Liu, P.X., Liu, G.: 'Gradient based and least-squares based iterative identification methods for OE and OEMA systems', *Digit. Signal Process.*, 2010, **20**, (3), pp. 664–677

26 Goodwin, G.C., Sin, K.S.: 'Adaptive filtering prediction and control' (Prentice-Hall, Englewood Cliffs, NJ, 1984)

27 Wang, D.Q.: 'Recursive extended least squares identification method based on auxiliary models', *Control Theory Appl.*, 2009, **26**, (1), pp. 51–56

28 Ding, F., Chen, T.: 'Hierarchical least squares identification methods for multivariable systems', *IEEE Trans. Autom. Control*, 2005, **50**, (3), pp. 397–402

29 Ding, F., Chen, T.: 'Identification of dual-rate systems based on finite impulse response models', *Int. J. Adapt. Control Signal Process.*, 2004, **18**, (7), pp. 589–598

30 Ding, F., Chen, T.: 'Least squares based self-tuning control of dual-rate systems', *Int. J. Adapt. Control Signal Process.*, 2004, **18**, (8), pp. 697–714

31 Ding, F., Chen, T.: 'Hierarchical identification of lifted state-space models for general dual-rate systems', *IEEE Trans. Circuits Syst. I, Regul. Pap.*, 2005, **52**, (6), pp. 1179–1187

32 Ding, F., Chen, T.: 'A gradient based adaptive control algorithm for dual-rate systems', *Asian J. Control*, 2006, **8**, (4), pp. 314–323

doi: 10.1049/iet-cta.2010.0416

33 Ding, F., Chen, T., Iwai, Z.: 'Adaptive digital control of Hammerstein nonlinear systems with limited output sampling', *SIAM J. Control Optim.*, 2006, **45**, (6), pp. 2257–2276

34 Ding, F., Liu, P.X., Yang, H.Z.: 'Parameter identification and intersample output estimation for dual-rate systems', *IEEE Trans. Syst. Man Cybern. A, Syst. Humans*, 2008, **38**, (4), pp. 966–975

35 Ding, J., Shi, Y., Wang, H.G., Ding, F.: 'A modified stochastic gradient based parameter estimation algorithm for dual-rate sampled-data systems', *Digital Signal Process.*, 2010, **20**, (4), pp. 1238–1249

36 Ding, F., Shi, Y., Chen, T.: 'Performance analysis of estimation algorithms of non-stationary ARMA processes', *IEEE Trans. Signal Process.*, 2006, **54**, (3), pp. 1041–1053

37 Liu, Y.J., Xie, L., Ding, F.: 'An auxiliary model based recursive least squares parameter estimation algorithm for non-uniformly sampled multirate systems', *Proc. Inst. Mech. Eng. I, J. Syst. Control Eng.*, 2009, **223**, (4), pp. 445–454

38 Ding, F., Qiu, L., Chen, T.: 'Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems', *Automatica*, 2009, **45**, (2), pp. 324–332

39 Xie, L., Yang, H.Z., Ding, F.: 'Modeling and identification for non-uniformly periodically sampled-data systems', *IET Control Theory Appl.*, 2010, **4**, (5), pp. 784–794

40 Ding, F., Liu, X.P., Liu, G.: 'Identification methods for Hammerstein nonlinear systems', *Digital Signal Process.*, 2011, **21**, (2), pp. 215–238

41 Wang, D.Q., Ding, F.: 'Least squares based and gradient based iterative identification for Wiener nonlinear systems', *Signal Process.*, 2011, **91**, (5), pp. 1182–1189

doi: 10.1049/iet-cta.2010.0416