Kalman Filter Reinforced by Least Mean Square for Systems with Unknown Inputs

Mohammad Ali Majidi, Chien-Shu Hsieh & Hadi Sadoghi Yazdi

Circuits, Systems, and Signal Processing

ISSN 0278-081X
Volume 37
Number 11

Circuits Syst Signal Process (2018)
37:4955-4972
DOI 10.1007/s00034-018-0792-x

Springer

Circuits Syst Signal Process (2018) 37:4955–4972
https://doi.org/10.1007/s00034-018-0792-x

Kalman Filter Reinforced by Least Mean Square for Systems with Unknown Inputs

Mohammad Ali Majidi[1] · **Chien-Shu Hsieh**[2] · **Hadi Sadoghi Yazdi**[1]

Received: 3 August 2017 / Revised: 17 February 2018 / Accepted: 20 February 2018 /
Published online: 27 February 2018

Abstract This paper addresses the state estimation problem of linear discrete-time time-varying stochastic systems with unknown inputs (UIs). It is shown that the globally optimal unbiased minimum-variance filters may not satisfy the minimum-variance property, and hence they cannot eliminate noises appropriately. If this is the case, the well-known Kalman filter may give a better solution, which however may also not be the best one due to that the imbedded unknown input model may not be practical. To remedy the filtering degradation problem, a robust filter named as the KFLMS, which has good noise rejection property for such systems, is developed in this paper, where the UI estimates are obtained by using least mean square algorithm and the state estimation is achieved via the previous proposed two-stage Kalman filtering approach. Numerical examples are provided to show the effectiveness of the proposed results. Specifically, simulation results illustrate the goodness of the new method in the sense of lower root mean square error and better noise rejection property.

Keywords Unknown inputs · Kalman filter · Adaptive filter · Least mean square · Recursive state estimation · Minimum mean square error

✉ Hadi Sadoghi Yazdi
h-sadoghi@um.ac.ir

Mohammad Ali Majidi
m.ali.majidi@gmail.com

Chien-Shu Hsieh
cshsieh@tust.edu.tw

1 Department of Computer Engineering, Center of Excellence on Soft Computing and Intelligent Information Processing, Ferdowsi University of Mashhad, Mashhad 9177948974, Iran

2 Department of Electrical and Electronics Engineering, Ta Hwa University of Science and Technology, Qionglin, Taiwan, ROC

1 Introduction

It is well-known that if a known state space model (SSM) of a linear stochastic time-varying system is available, the Kalman filter (KF) gives the optimal solution to the state estimation of the system [9]. On the other hand, adaptive filtering, e.g., the least mean square (LMS) algorithm [10], or the weighted least-squares (WLS) [7,8] do not need to know the SSM in order to achieve the optimal state estimation, where the unmodeled state estimates are mainly obtained from the outputs. Note that, the LMS algorithm is an instantaneous gradient-based optimization technique, which can achieve a minimization of a certain cost function. As shown in [19], the LMS algorithm serves as an optimal solution to the unmodeled state estimation problem under the H^∞ criterion.

In many applications, e.g., fault detection and isolation [1], a model with some uncertainties may exist. These uncertainties often arise from linearization errors, parameter uncertainties, faults, and unmeasurable disturbances [11]. To remedy these uncertainties, many methods have been used such as recursive least squares, stochastic gradient algorithm, iterative algorithm and variational Bayes methods [2,3,21]. As mentioned in [18], the model uncertainties can be considered as an unknown input which affects known model of the system; in such cases, the KF usually gives a biased state estimation. In order to fix this biased estimation, unknown input filtering (UIF) techniques are often used to obtain unbiased minimum-variance (UMV) state estimates, and the obtained filters are named as the UMV estimators (UMVEs) [1,4,5,7,8,11,12,18], which makes a great evolution in incomplete model-based state estimation problems due to achieving globally optimal state estimation for uncertain SSMs with UIs [14]. However, the UMVE is restricted to have some assumptions on known matrices of the SSM to satisfy the unbiasedness condition, which although are completely determined [15] but are sometimes too restrictive to be applied. This is mainly due to the fact that the UMVE is in one of its degeneration forms, which are formed either by that the filter's unbiasedness condition not being satisfied [4,6,20] or the gain matrix is not fully connected to the Kalman gain [17]. A possible approach to alleviate the above-mentioned first degenerated case is by applying time-delayed state estimators [6,20]. Nevertheless, few literature results are focused on the second degenerated case. In [17], a hybrid minimum-variance filter named parameterized augmented robust two-stage Kalman filters (PARTSKF), which compromises between the unbiasedness and the minimum-variance estimation, has been proposed to remedy this filtering degradation problem. Although the PARTSKF has a promising improvement on the filtering performance, it may also fail to yield a satisfactory estimation due to that the assumed UI model may not work well.

In this paper, a new and innovative filtering structure is proposed to alleviate the deterioration problem of the UMVE and to relax the heuristic UI model in the PARTSKF. The obtained filter will be named as the Kalman filter reinforced by least mean square (KFLMS), which signifies that the main difference between this new filter and the Kalman Filter is that the UI is estimated by the LMS algorithm in the former. Note that, there are other adaptive filters which are also applicable to this new filtering structure, e.g., the recursive least-square (RLS) algorithm. However, for the sake of simplicity, computational complexity, and numerical stability [10], in this paper we only consider the LMS algorithm.

The rest of the paper is organized as follows. In Sect. 2, the statement of the problem is addressed. Section 3 recalls some preliminaries concerning the derivation of the KFLMS. Then, the proposed KFLMS is derived in Sect. 4. Two compact versions of the KFLMS are also presented to reduce the filters computational complexity. Some illustrative examples are given in Sect. 5 to show the effectiveness of the proposed methods. Finally, Sect. 6 has the conclusion.

2 Problem Formulation

Consider the following linear discrete-time time-varying stochastic system:

$$x_{k+1} = A_k x_k + B_k u_k + G_k d_k + \omega_k \tag{1}$$
$$y_k = C_k x_k + H_k d_k + \upsilon_k \tag{2}$$

where $x_k \in R^n$ is the state vector, $u_k \in R^l$ is the known input vector, $d_k \in R^p$ is an unknown input vector, and $y_k \in R^m$ is the measurement vector. The process noise ω_k and the measurement noise υ_k are assumed to be mutually uncorrelated, zero-mean, white random signals with known covariance matrices, $Q_k \geq 0$ and $R_k > 0$, respectively. The matrices A_k, B_k, G_k, C_k and H_k are known and it is assumed that (A_k, C_k) is observable and that x_0 is independent of υ_k and ω_k for all k. Moreover, an unbiased estimate $\check{x}_0$ of the initial state x_0 is available with covariance matrix P_0^x. The problem of interest is to design an optimal linear estimator of x_k without any information concerning the UI vector (d_k) given the measurements up to time k.
One of the UMVE that solves the addressed problem is given by the refined optimal unbiased minimum-variance filter (ROUMVF) [15], which is slightly rewritten as follows:

$$\hat{x}_k = \overline{x}_k + L_k\,(y_k - C_k \overline{x}_k) \tag{3}$$
$$P_k = (I - L_k C_k)\,\overline{P}_k (I - L_k C_k)^{\mathrm{T}} + L_k R_k L_k{}^{\mathrm{T}} \tag{4}$$
$$L_k = \bar{G}_k + (K_k - \bar{G}_k) W_k \tag{5}$$
$$\bar{x}_{k+1} = A_k \hat{x}_k + B_k u_k + G_k \hat{d}_k \tag{6}$$
$$\bar{P}_{k+1} = A_k P_k A_k{}^{\mathrm{T}} + A_k P_k^{xd} G_k{}^{\mathrm{T}} + G_k (A_k P_k^{xd})^{\mathrm{T}} + G_k\ P_k^d G_k{}^{\mathrm{T}} + Q_k \tag{7}$$

where $\bar{G}_k$, K_k, and W_k are named as the unbiasedness gain, the Kalman gain, and the gain weighting matrix, respectively, which are given as follows:

$$\bar{G}_k = [\,0\ \bar{U}_{k-1}\,] S_k{}^{+}, \quad \bar{U}_k = G_k (I - H_k{}^{+} H_k) \tag{8}$$
$$K_k = \bar{P}_k C_k{}^{\mathrm{T}} \tilde{R}_k^{-1}, \quad \tilde{R}_k = C_k \bar{P}_k C_k{}^{\mathrm{T}} + R_k \tag{9}$$
$$W_k = \tilde{R}_k \bar{T}_k^{\mathrm{T}} \left(\bar{T}_k \tilde{R}_k \bar{T}_k^{\mathrm{T}} \right)^{-1} \bar{T}_k, \quad \bar{T}_k = \alpha_k \left(I - S_k S_k{}^{+} \right) \tag{10}$$

in which $S_k = [H_k \quad C_k\bar{U}_{k-1}]$, A_k^+ denotes any generalized inverse of matrix A, and the matrix parameter α_k must be chosen so that $\bar{T}_k$ is of full-row rank. Here, $\hat{d}_k$, P_k^d and P_k^{xd} are the UI estimates and their corresponding covariances:

$$\hat{d}_k = M_k\left(S_k{}^{\mathrm{T}}\tilde{R}_k^{-1}S_k\right)^+ S_k{}^{\mathrm{T}}\tilde{R}_k^{-1}(y_k - C_k\bar{x}_k) \tag{11}$$

$$P_k^d = M_k(S_k^{\mathrm{T}}\tilde{R}_k^{-1}S_k)^+M_k^{\mathrm{T}} \tag{12}$$

$$P_k^{xd} = ([0 \quad \bar{U}_k - 1] - K_kS_k)(S_k^{\mathrm{T}}\tilde{R}_k^{-1}S_k)^+M_k^{\mathrm{T}} \tag{13}$$

where $M_k = [H_k^+H_k \quad 0]$. Note that the gain matrix L_k by (5) can also be expressed as follows:

$$L_k = \bar{G}_k(I - W_k) + K_kW_k$$

which clearly illustrates that the gain matrix of the ROUMVF is determined through a compromise between the unbiasedness gain and the Kalman gain, via a suitable gain weighting matrix. It should be stressed that the unbiasedness gain will promise that the obtained state estimates are unbiased, i.e., satisfying the unbiasedness condition $\bar{G}_kS_k = [0 \;\; \bar{U}_{k-1}]$, and the Kalman gain will promise that the obtained error covariance of the state estimate at each time instant is a minimum-variance one, which can be verified by checking the optimality condition: $W_k\tilde{R}_kW_k^{\mathrm{T}}K_k^{\mathrm{T}} = W_kC_k\bar{P}_k$. In the special case that UIs do not enter into the system, i.e., $W_k = I$, the gain L_k will be reduced to the Kalman gain K_k.

Based on the above argument, a special case illustrating the deterioration of the UMVE, i.e., $\bar{T}_k = 0$, was first observed in the previous study [13]. Recently, more general cases are highlighted by reducing the gain matrix L_k to the following form [17]:

$$L_k = \bar{G}_k + (I - \bar{G}_k\bar{G}_k^+)\Phi_k \tag{14}$$

where Φ_k is the only matrix in L_k that could contain the information of Q_k and R_k. Thus, using (3) and (14), one has:

$$\bar{G}_k^+\hat{x}_k = \bar{G}_k^+\bar{x}_k + \bar{G}_k^+\bar{G}_k(y_k - C_k\bar{x}_k) \tag{15}$$

Then, if there exists a matrix Y_k satisfying the following condition: $Y_k\bar{G}_k^+ = \mathrm{diag}\{I, 0\}$ then left-multiplying (15) by Y_k yields:

$$\mathrm{diag}\{I, 0\}\hat{x}_k = \mathrm{diag}\{I, 0\}(\bar{x}_k + \bar{G}_k(y_k - C_k\bar{x}_k))$$

which indicates that part of the state estimates will be completely irrelevant to the assumed noise covariances, and hence the obtained state estimates could not reject noises appropriately and may even just return the related measurements [17].

To the best of the authors knowledge, there is no feasible approach in the literature to solve the above-mentioned performance degradation problem of the UMVE. Hence, the main focus of this paper is to propose a new robust filtering framework in order to alleviate the filtering performance degradation problem.

3 Preliminaries

3.1 The Augmented State Kalman Filter

If the UI dynamics can be represented by a random walk process as below:

$$d_{k+1} = d_k + \omega_k^d \tag{16}$$

where ω_k^d is a zero-mean white Gaussian random process with covariance Q_k^d and is independent of both the process noise ω_k and the measurement noise ν_k, then the augmented state Kalman filter (ASKF) is the optimal solution for the joint state and UI estimation. The ASKF runs two KFs on system (1), (2), and (16) simultaneously, where the first KF (UI filter) estimates the UI and the second one (state filter) estimates the states. The two filters are given, respectively, as follows:

(1) UI filter:

$$\hat{d}_k = \hat{d}_{k-1} + K_k^d(y_k - C_k\hat{x}_k^- - H_k\hat{d}_{k-1}) \tag{17}$$

$$P_k^{d-} = P_{k-1}^d + Q_k^d \tag{18}$$

$$P_k^{xd-} = A_{k-1}P_{k-1}^{xd} + G_{k-1}P_{k-1}^d \tag{19}$$

$$K_k^d = ((C_k P_k^{xd-})^{\mathrm{T}} + P_k^{d-}H_k^{\mathrm{T}})\tilde{R}_k^{-1} \tag{20}$$

$$P_k^d = (I - K_k^d H_k)P_k^{d-} - K_k^d C_k P_k^{xd-} \tag{21}$$

where

$$\tilde{R}_k = C_k P_k^{x-} C_k^{\mathrm{T}} + H_k(C_k P_k^{xd-})^{\mathrm{T}} + C_k P_k^{xd-} H_k^{\mathrm{T}} + H_k P_k^{d-} H_k^{\mathrm{T}} + R_k \tag{22}$$

(2) State filter:

$$\hat{x}_k^- = A_{k-1}\hat{x}_{k-1} + B_{k-1}u_{k-1} + G_{k-1}\hat{d}_{k-1} \tag{23}$$

$$\hat{x}_k = \hat{x}_k^- + K_k^x(y_k + C_k\hat{x}_k^- - H_k\hat{d}_{k-1}) \tag{24}$$

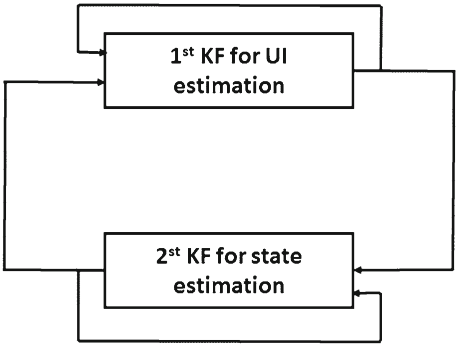
$$\begin{aligned} P_k^{x-} &= A_{k-1}P_{k-1}^x A_{k-1}^{\mathrm{T}} + G_{k-1}P_{k-1}^d G_{k-1}^{\mathrm{T}} + Q_{k-1} \\ &\quad + A_{k-1}P_{k-1}^{xd}G_{k-1}^{\mathrm{T}} + G_{k-1}(A_{k-1}P_{k-1}^{xd})^{\mathrm{T}} \end{aligned} \tag{25}$$

$$K_k^x = (P_k^{x-}C_k^{\mathrm{T}} + P_k^{xd-}H_k^{\mathrm{T}})\tilde{R}_k^{-1} \tag{26}$$

$$P_k^x = (I - K_k^x C_k)P_k^{x-} - K_k^x H_k(P_k^{xd-})^{\mathrm{T}} \tag{27}$$

$$P_k^{xd} = (I - K_k^x C_k)P_k^{xd-} - K_k^x H_k P_k^{d-} \tag{28}$$

Fig. 1 Flowchart for augmented state Kalman filter (ASKF)



So, the ASKF run two Kalman filters on system (1–2) and (16) simultaneously, the first Kalman filter estimate the unknown input, and the second one is for state estimation. So, we can draw ASKF flowchart as Fig. 1:

3.2 The Optimal Two-Stage Kalman Filter

It is known that the ASKF may not be feasible to be applied in practical applications due to that its computational cost increases with the UI vector dimension. If this is the case, the optimal two-stage Kalman filter (OTSKF) [16] can be used alternatively to maintain the computational complexity at a lower level while the filtering performance still be preserved. The OTSKF comprises two quasi-parallel filters, which are named as the modified bias-free filter and the bias filter, and are listed, respectively, as follows:

(1) Modified bias-free filter:

$$\bar{x}_k^- = A_{k-1}\hat{x}_{k-1} + B_{k-1}u_{k-1} + \bar{u}_{k-1} \tag{29}$$

$$\bar{x}_k = \bar{x}_k^- + \bar{K}_k^x(y_k + C_k\bar{x}_k^-) \tag{30}$$

$$\bar{P}_k^{x-} = A_{k-1}\bar{P}_k^x A_{k-1}^{\mathrm{T}} + Q_{k-1} + Q_{k-1}^{\bar{u}} \tag{31}$$

$$\bar{K}_k^x = \bar{P}_k^{x-} C_k^{\mathrm{T}}(C_k\bar{P}_k^{x-} C_k^{\mathrm{T}} + R_k)^{-1} \tag{32}$$

$$\bar{P}_k^x(I - \bar{K}_k^x C_k)\bar{P}_k^{x-} \tag{33}$$

where

$$\bar{u}_k = (G_k - U_{k+1})\hat{d}_k \tag{34}$$

$$Q_k^{\bar{u}} = (G_k - U_{k+1})\bar{P}_k^d\bar{U}_{k+1}^{\mathrm{T}} + (\bar{U}_{k+1} - G_k)\bar{P}_k^d G_k^{\mathrm{T}} \tag{35}$$

in which

$$U_k = \bar{U}_k(I - Q_{k-1}^d(\bar{P}_k^{d-})^{-1}) \tag{36}$$

$$\bar{U}_k = A_{k-1}V_{k-1} + G_{k-1} \tag{37}$$

(2) Bias filter:

$$\hat{d}_k = \hat{d}_{k-1} + \bar{K}_k^d(y_k - C_k\bar{x}_k^- - S_k\hat{d}_{k-1}) \tag{38}$$

$$\bar{P}_k^{d-} = \bar{P}_{k-1}^d + Q_{k-1}^d \tag{39}$$

$$\bar{K}_k^d = \bar{P}_k^{d-} S_k^{\mathrm{T}}(S_k\bar{P}_k^{d-} S_k^{\mathrm{T}} + \bar{R}_k)^{-1} \tag{40}$$

$$\bar{P}_k^d = (I - \bar{K}_k^d S_k)\bar{P}_k^{d-} \tag{41}$$

where

$$S_k = C_k U_k + H_k \tag{42}$$

Then, based on the above two filters, the state estimate $\hat{x}_k$ in (24) can be reconstructed as follows:

(3) State filter:

$$\hat{x}_k = \bar{x}_k + V_k\hat{d}_k, \quad P_k^x = \bar{P}_k^x + V_k\bar{P}_k^d V_k^{\mathrm{T}} \tag{43}$$

where

$$V_k = U_k - \bar{K}_k^x S_k \tag{44}$$

Comparing the OTSKF with the ASKF, it is clear that the former is more compact than the latter because the coupling terms P_k^{xd-} and P_k^{xd} in the latter are not used. Moreover, this compactness may help to derive a robust version of the KF, in which the dedicated UI model is not needed. An illustrative case showing this will be given in Sect. 3.

Remark 1 The above modified bias-free filter and bias filter can be viewed as standard KFs for the following respective subsystems:

(1) Bias-free subsystem:

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + u_k^d + \omega_k \\ y_k &= C_k x_k + v_k \end{aligned}$$

where u_k^d is viewed as a known external signal, independent of x_k and u_k, with the following statistics:

$$\begin{aligned} E[u_k^d] &= (A_k V_k + G_k - U_{k+1})\hat{d}_k \\ cov(u_k^d) &= A_k V_k \bar{P}_k^d (A_k V_k)^{\mathrm{T}} + Q_k^{\bar{u}} \end{aligned}$$

(2) UI subsystem

$$d_{k+1} = d_k + \omega_k^d$$

$$\tilde{y}_k = S_k d_k + \tilde{\nu}_k$$

where $\tilde{y}_k = y_k - x_k^-$ and $cov(\tilde{\nu}_k) = \bar{R}_k$.

3.3 The Least Mean Square Algorithm

Because the random walk assumption in (16) may not hold in practical applications, and hence the ASKF/OTSKF may not work well. In this subsection, we highlight the idea how to identify the UI estimates without resorting to the random walk assumption by applying the adaptive LMS algorithm, which is inspired by [19] as below.

First, we assume that measurement equation (2) can be expressed as follows:

$$D_k = X_k^{\mathrm{T}} W_k + E_k \tag{45}$$

where $D_k \in R^m$ is the desired output vector, $X_k \in R^{p\times m}$ is the input matrix, $E_k \in R^m$ is the error vector, and $W_k \in R^p$ is the weight vector to be determined to minimize the following cost function:

$$min_{W_k} \varepsilon_k = (||W_k - \bar{W}_k||^2_{\Pi_{k-1}} + ||D_k - X_k^{\mathrm{T}} W_k||_2^2)$$

where $\bar{W}_k$ is the initial guess of W_k, $||X||_{\Pi} = \sqrt{X_T \Pi^{-1} X}$, $\Pi_{k-1} = \infty I$, and $||X||_2 = ||X||_I$.

Second, assuming that ε_k, D_k and X_k are statistically stationary and $[X_k X_k^{\mathrm{T}}] \neq 0$, the optimal solution to identify the vector W_k from (45) is given as the following recursive least-squares (RLS) solution:

$$W_k^* = \bar{W}_k + (X_k X_k^{\mathrm{T}})^{-1} X_k (D_k - X_k^{\mathrm{T}} \bar{W}_k). \tag{46}$$

More practically, i.e., to further reduce the computational complexity of the RLS solution, we have the following more compact LMS solution:

$$\hat{W}_k^{(i)} = \hat{W}_k^{(i-1)} + \mu_k X_k^i (D_k^i - X_k^{i\,\mathrm{T}} \hat{W}_k^{(i-1)}) \tag{47}$$

where $\hat{W}_k^0 = \bar{W}_k$, $1 \leq i \leq m$, μ_k is a proper step-size parameter, which will be determined later, and matrices D_k, X_k and E_k are denoted, respectively, as follows:

$$D_k = \begin{bmatrix} D_k^1 \\ D_k^2 \\ \vdots \\ D_k^m \end{bmatrix}, \quad X_k = \begin{bmatrix} X_k^{1\,\mathrm{T}} \\ X_k^{2\,\mathrm{T}} \\ \vdots \\ X_k^{m\,\mathrm{T}} \end{bmatrix}^{\mathrm{T}}, \quad E_k = \begin{bmatrix} E_k^1 \\ E_k^2 \\ \vdots \\ E_k^m \end{bmatrix}$$

Based on the recursive weight update (47), we have the weight estimate as $\hat{W}_k = \hat{W}_k^{(m)}$.

Third, we note that, although the LMS solution is not a minimizer of the usual H^2 norm it builds itself as an optimal solution under the following H^∞ criterion:

$$\frac{\sum_{i=1}^{m} a_k^{i-1}|{X_k^i}^{\mathrm{T}} W_k - {X_k^i}^{\mathrm{T}} \hat{W}_k^{(i-1)}|^2}{\mu_k^{-1}||W_k - \bar{W}_k||_2^2 + \sum_{i=1}^{m} a_k^{i-1}|X_k^i|^2} \leq 1 \tag{48}$$

where $a_k \in (0, 1)$ is a weighting factor. To see this, we intend to verify the following inequality:

$$\begin{aligned} &\mu_k^{-1}||W_k - \hat{W}_k^{(i-1)}||_2^2 - a_k\mu_k^{-1}||W_k - \hat{W}_k^{(i)}||_2^2 + |E_k^i|^2 \\ &\geq ||X_k^i||_2^2.||W_k - \hat{W}_k^{(i-1)}||_2^2 \end{aligned} \tag{49}$$

Applying the LMS solution, we have

$$||W_k - \hat{W}_k^{(i)}||_2^2 \leq ||W_k - \hat{W}_k^{(i-1)}||_2^2 + ||\hat{W}_k^{(i)} - \hat{W}_k^{(i-1)}||_2^2 \leq \zeta \tag{50}$$

where

$$\zeta = ||W_k - \hat{W}_k^{(i-1)}||_2^2 + \mu_k^2||X_k^i||_2^2.|E_k^i|^2 + \mu_k^2||X_k^i||_2^4||W_k - \hat{W}_k^{(i-1)}||_2^2$$

Using (50) in (49) yields

$$((1 - a_k)\mu_k^{-1} - a_k\mu_k||X_k^i||_2^2) \times ||W_k - \hat{W}_k^{(i-1)}||_2^2 + (1 - a_k\mu_k||X_k^i||_2^2)|E_k^i|^2 \geq 0 \tag{51}$$

It is clear from (51) that if μ_k is chosen to satisfy the inequality:

$$0 \leq \mu_k \leq min\left(\frac{-1 + \sqrt{1 + 4a_k(1 - a_k)}}{2a_k l_k^i}, \frac{1}{a_k l_k^i}\right) = \frac{-1 + \sqrt{1 + 4a_k(1 - a_k)}}{2a_k l_k^i} \tag{52}$$

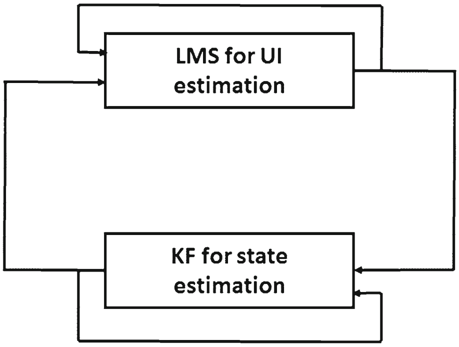
where $1 \leq i \leq m$ and $l_k^i = (X_k^i)^{\mathrm{T}} X_k^i$, inequality (49) will hold, which also promises the inequality:

$$\begin{aligned} &\mu_k^{-1}||W_k - \hat{W}_k^{(i-1)}||_2^2 - a_k\mu_k^{-1}||W_k - \hat{W}_k^{(i)}||_2^2 + |E_k^i|^2 \\ &\geq |(X_k^i)^{\mathrm{T}} W_k - (X_k^i)^{\mathrm{T}} \hat{W}_k^{(i-1)}|^2 \end{aligned} \tag{53}$$

Thus, summing (53) over i, eliminating the terms of $a_k\mu_k^{-1}||W_k - \hat{W}_k^{(i)}||_2^2$, and using $\hat{W}_k^{(0)} = \bar{W}_k$, we obtain the claim in (48). Note that, the inequality in (52) will facilitate the determination of the step-size parameter. This is addressed in Sect. 5.

Finally, we conclude from the above results that the robust behavior of the LMS algorithm is shown by guaranteeing that the energy of the weighted prediction error is

Fig. 2 Flowchart for KFLMS



always bounded by the energy of the initial uncertainty and the weighted disturbances, i.e., (48).

4 KFLMS Filters Design

4.1 Derivation of the KFLMS

The basic ideas to derive the KFLMS are)1) to develop a robust version of the modified bias-free filter of the OTSKF, which intends to be robust against the UI model, (2) to develop a modified bias filter of the OTSKF, which intends to compensate for the robust bias-free filter in order to obtain a robust state estimator, and (3) to replace the UI filter of the ASKF with the LMS algorithm, which intends to relax the imbedded random walk assumption in the UI subsystem. We can draw a flowchart for KFLMS by using the flowchart of ASKF (Fig. 1) as Fig. 2.

First, we modify Eqs. (29) and (31) as follows [12]:

$$\bar{x}_k^- = A_{k-1}\hat{x}_{k-1} + B_{k-1}u_{k-1} \tag{54}$$

$$\bar{P}_k^{x-} = A_{k-1}\bar{P}_{k-1}^x A_{k-1}^{\mathrm{T}} + Q_{k-1} \tag{55}$$

which is achieved by choosing the following blending matrix:

$$U_{k+1} = G_k \tag{56}$$

and ignoring the covariance of UI, i.e., $Q_k^d = 0$.

Second, using (54)–(56), the state filter of the OTSKF is given as follows:

$$\hat{x}_k = \bar{x}_k + V_k\check{d}_k, \quad P_k^x = \bar{P}_k^x + V_k\check{P}_k^d V_K^{\mathrm{T}} \tag{57}$$

where $\bar{x}_k$ and $\bar{P}_k^x$ are given by (30) and (33), respectively, $\check{d}_k$ is the modified bias filter, given as follows:

$$\check{d}_k = \check{d}_{k-1} + \hat{K}_k^d (y_k - C_k \bar{x}_k^- - F_k \hat{d}_{k-1}) \tag{58}$$

$$\check{P}_k^d = (I - \hat{K}_k^d F_k) P_{k-1}^d \tag{59}$$

$$F_k = C_k G_{k-1} + H_k \tag{60}$$

$$\hat{K}_k^d = P_{k-1}^d F_k^{\mathrm{T}} (F_k P_{k-1}^d F_k^{\mathrm{T}} + \bar{R}_k)^{-1} \tag{61}$$

In which $\hat{d}_{k-1}$ is the previous estimate of UI and P_{k-1}^d its error covariance, both of them are to be determined later, and the blending matrix V_k is given as follows:

$$V_k = G_{k-1} - \bar{K}_k^x F_k \tag{62}$$

Note that the state filter in (57) can be viewed as a special implementation of that of the ASKF, as shown in the following theorem.

Theorem 1 *If the following setting is used:*

$$P_k^{xd} = 0, \quad Q_{k-1}^d = 0 \tag{63}$$

the state subfilter of the ASKF is equivalent to that of the KFLMS, i.e., (57)–(62).

Proof Using (63), we obtain that Eqs. (25)–(27) of the state subfilter of the ASKF are rewritten, respectively, as follows:

$$P_k^{x^-} = A_{k-1} P_{k-1}^x A_{k-1}^{\mathrm{T}} + G_{k-1} P_{k-1}^d G_{k-1}^{\mathrm{T}} + Q_{k-1} \tag{64}$$

$$K_k^x = (P_k^{x^-} C_k^{\mathrm{T}} + G_{k-1} P_{k-1}^d H_k^{\mathrm{T}}) \tilde{R}_k^{-1} \tag{65}$$

$$P_k^x = (I - K_k^x C_k) P_k^{x^-} - K_k^x H_k (P_k^{xd^-})^{\mathrm{T}} \tag{66}$$

where

$$\tilde{R}_k = C_k P_k^{x^-} C_k^{\mathrm{T}} + H_k (C_k G_{k-1} P_{k-1}^d)^{\mathrm{T}} + C_k G_{k-1} P_{k-1}^d H_k^{\mathrm{T}} + H_k P_{k-1}^d H_k^{\mathrm{T}} + R_k \tag{67}$$

Using (32), (60)–(62), (65), and the following relationship:

$$P_k^{x^-} = \bar{P}_k^{x^-} + G_{k-1} P_{k-1}^d G_{k-1}^{\mathrm{T}} \tag{68}$$

we have

$$\tilde{R}_k = \bar{R}_k + F_k P_{k-1}^d F_k^{\mathrm{T}}$$

$$K_k^x = \bar{P}_k^{x^-} C_k^{\mathrm{T}} \tilde{R}_k^{-1} + G_{k-1} \hat{K}_k^d = \bar{K}_k^x - \bar{K}_k^x F_k \hat{K}_k^d + G_{k-1} \hat{K}_k^d = \bar{K}_k^x + V_k \hat{K}_k^d \tag{69}$$

Using (24), (30), (58), (60), (62), (69), and the following relationship:

$$\hat{x}_k^- = \bar{x}_k^- + G_{k-1}\hat{d}_{k-1} \tag{70}$$

we have the state estimates as:

$$\hat{x}_k = \bar{x}_k + V_k\hat{d}_{k-1} + V_k\hat{K}_k^d(y_k - C_k\bar{x}_k^- - F_k\hat{d}_{k-1}) = \bar{x}_k + V_k\check{d}_k \tag{71}$$

Using (59) and (61), we have

$$\hat{K}_k^d\bar{R}_k = (I - \hat{K}_k^d F_k)P_{k-1}^d F_k^{\mathrm{T}} = \check{P}_k^d F_k^{\mathrm{T}}$$

by which and using (32)–(33), (59)–(62), (66), and (68)–(69), we have the state estimation error covariance as:

$$P_k^x = \bar{P}_k^x + V_k\check{P}_k^d G_{k-1}^{\mathrm{T}} - V_k\hat{K}_k^d\bar{R}_k(\bar{K}_k^x)^{\mathrm{T}} = \bar{P}_k^x + V_k\check{P}_k^d V_k^{\mathrm{T}} \tag{72}$$

From (71) and (72), the theorem is proved.

Third, it remains to determine the UI estimates. Rewrite the innovation of the modified bias-free filter as follows:

$$D_k = y_k - C_k\bar{x}_k^- = F_k d_k + E_k \tag{73}$$

where

$$E_k = C_k(A_{k-1}\tilde{x}_{k-1} + G_{k-1}\epsilon_{k-1} + \omega_{k-1}) + \nu_k \tag{74}$$

in which $\epsilon_{k-1} = d_{k-1} - d_k$. Using the following substitutions in the LMS algorithm:

$$F_k^{\mathrm{T}} \to X_k, \quad d_k \to W_k$$

we can obtain the UI estimates as follows:

$$\begin{cases} \hat{d}_k^{(i)} = \hat{d}_k^{(i-1)} + \mu_k\left(F_k^i\right)^{\mathrm{T}}(y_k^i - C_k^i\bar{x}_k^- - F_k^i\hat{d}_k^{(i-1)}) \\ \hat{d}_k = \hat{d}_k^{(m)}. \quad \hat{d}_k^{(0)} = \hat{d}_{k-1} \end{cases} \tag{75}$$

where μ_k is determined by (52). From (58) and (59), we have the following heuristic estimation error covariance of $\hat{d}_k$:

$$P_k^d = \prod_{i=1}^{m}(I - \mu_k(F_k^i)^{\mathrm{T}}F_k^i)P_{k-1}^d \tag{76}$$

Finally, the KFLMS is given as follows:

(1) UI filter: (75)–(76).
(2) State filter: (30), (32)–(33), and (54)–(62). □

Remark 2 From Theorem 1, the state filter of the KFLMS can also be expressed as follows:

$$\hat{x}_k = \hat{x}_k^- + K_k^x(y_k - C_k\bar{x}_k^- - F_k\hat{d}_{k-1}) \tag{77}$$

$$K_k^x = (\bar{P}_k^{x^-}C_k^{\mathrm{T}} + G_{k-1}P_{k-1}^dF_k^{\mathrm{T}}) \times (C_k\bar{P}_k^{x^-}C_k^{\mathrm{T}} + F_kP_{k-1}^dF_k^{\mathrm{T}} + R_k)^{-1} \tag{78}$$

$$P_k^x = P_k^{x^-} - K_k^x(C_k\bar{P}_k^{x^-} + F_kP_{k-1}^dG_{k-1}^{\mathrm{T}}) \tag{79}$$

where x_k^-, $\hat{x}_k^-$, $\bar{P}_k^{x^-}$, and $P_k^{x^-}$ are given by (54), (70), (55), and (68), respectively.

Remark 3 Comparing (75)–(79) with (17)–(28), it is clear that the proposed KFLMS is in fact a special implementation of the ASKF that is obtained by using the specific setting in (63) and relaxing the random walk UI model via the adaptive LMS algorithm.

Remark 4 The UI filter (75)–(76) can be implemented more compactly if the following conditions hold:

$$F_k^i(F_k^{i-1})^{\mathrm{T}} = 0, \quad 2 \le i \le m$$

Then, the simplified filter is given as follows:

$$\begin{aligned} \hat{d}_k &= \hat{d}_{k-1} + \mu_kF_k^{\mathrm{T}}(y_k - C_k\bar{x}_k^- - F_k\hat{d}_{k-1}) \\ P_k^d &= (I - \mu_kF_k^{\mathrm{T}}F_k)P_{k-1}^d. \end{aligned}$$

4.2 Compact Versions of the KFLMS

As shown in Sect. 4.1, the KFLMS is composed of three subfilters: the modified bias-free filter $\bar{x}_k$, the modified bias filter $\check{d}_k$ and the UI filter $\hat{d}_k$; all of which can be integrated in a certain way to yield the simultaneous input and state estimates. Although the modified bias filter and the UI filter have similar estimator structures, the design methods embedded in them are different: the gain matrix of the former is derived by assuming a heuristic UI model, i.e., the random walk process (16), while that of the latter is obtained by using no UI model. In the conventional filter design for systems with unknown inputs, UI estimates are usually obtained by the modified bias filter instead of the UI filter; however, in this paper we intend to achieve the UI estimation by the latter.

Based on the above arguments, there are two ways to simplify the computational complexity of the KFLMS: one is to replace the modified bias filter by the UI filter, denoted by CKFLMS^1, and the other is to replace the UI filter by the modified bias filter, denoted by CKFLMS^2. The filtering performance of the above CKFLMSs is illustrated in Sect. 5.

5 Illustrative Examples

To illustrate the effectiveness of the proposed KFLMS filters, in this section we consider two numerical examples, where the first one satisfies the degenerated case (14) with $\Phi_k \neq 0$ and the other satisfies the specific condition $\bar{T}_k = 0$, i.e., the degenerated case (14) with $\Phi_k = 0$. In the simulations, the ASKF (17)–(28), the OTSKF (29)–(44), the KFLMS (75)–(79), the CKFLMS^1/KFLMS ($\hat{d}_k \to \check{d}_k$, $P_k^d \to \check{P}_k^d$), the CKFLMS^2/KFLMS ($\check{d}_k \to \hat{d}_k$, $\check{P}_k^d \to P_k^d$), and the ROUMVF (3)–(13) were considered. The simulation time is 1000 time steps with a Monte Carlo simulation of 50 runs. Note that in setting μ_k we use the upper bound in (52) by choosing a proper a_k that is determined via a tradeoff between the small upper bound and the rapid convergent rate. Moreover, the covariance Q_k^d used in the filter designs is given as $Q_k^d = \text{diag}\{0.025, 0.016\}$ and all the filters are initialized with $\hat{x}_0 = 0$ and $P_0 = I_{3\times 3}$.

The first case For the first system, the numerical example given by [1] was considered, which satisfies the condition in Eq. (14). The parameters of system (1)–(2) are given as follows:

$$A_k = \begin{bmatrix} 0.9944 & -0.1203 & -0.4302 \\ 0.0017 & 0.9902 & -0.0747 \\ 0 & 0.8187 & 0 \end{bmatrix}, \quad B_k = \begin{bmatrix} 0.4252 \\ -0.0082 \\ 0.1813 \end{bmatrix}, \quad G_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$
$$C = I_{3\times 3}, \quad H_k = 0_{3\times 2}$$

The covariance matrices are given as $Q_k = \text{diag}\{0.1^2, 0.1^2, 0.01^2\}$ and $R_k = I_{(3\times 3)}$. In simulation, we set $u_k = 10$, $x_0 = [1 \quad 1 \quad 1]^{\mathrm{T}}$, $P_0 = 0.1^2 I_{(3\times 3)}$, $a_k = 0.98$ which yield $\mu_k = 0.02$ and

$$d_k = \begin{bmatrix} \Delta a_{11} & \Delta a_{12} & \Delta a_{13} \\ \Delta a_{21} & \Delta a_{22} & \Delta a_{23} \end{bmatrix} x_k + \begin{bmatrix} \Delta b_1 \\ \Delta b_2 \end{bmatrix} u_k$$

where $\Delta a_{ij} = -0.5 a_{ij}$ and $\Delta b_j = 0.5 b_j$. a_{ij} and b_j are elements of matrix A_k and vector B_k, respectively.

The obtained rmses of the considered filters are given in Table 1, by which we have the following observations: (1) the ROUMVF has the worst filtering performance, which illustrates the possible shortcoming of the UMVE, (2) the ASKF and the OTSKF have the same filtering performance; both of them are superior to the ROUMVF, (3) the KFLMS has the best filtering performance, (4) the CKFLMS^1 has comparable filtering performance with the KFLMS, (5) the CKFLMS^2 is slightly worse than the ASKF.

Table 1 Performance of the ASKF, OTSKF, KFLMS, CKFLMS, and ROUMVF in the first case

Filter	rmse(x_k^1)	rmse(x_k^2)	rmse(x_k^3)	rmse(d_k^1)	rmse(d_k^2)
ASKF	0.6356	0.4680	0.3212	0.5778	0.1385
OTSKF	0.6356	0.4680	0.3212	0.5778	0.1385
KFLMS	0.5986	0.3076	0.2323	0.6756	0.0794
CKFLMS^1	0.6793	0.2929	0.2310	0.6477	0.0797
CKFLMS^2	0.7567	0.7100	0.4096	0.7112	0.4789
ROUMVF	0.9959	0.9998	0.6358	3.5210	0.1799

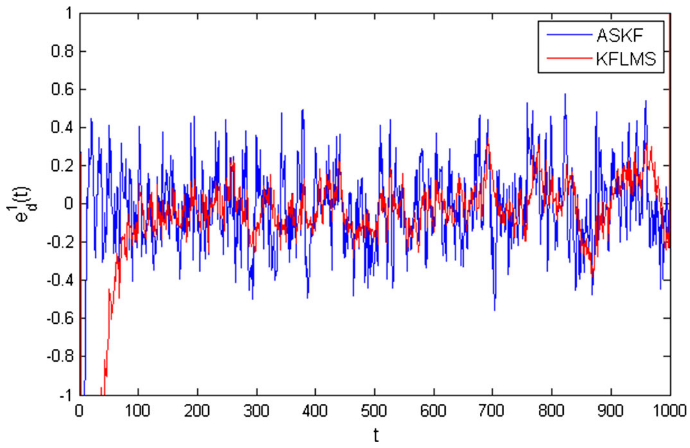


Fig. 3 UI (d_k^1) estimation errors of the ASKF and the KFLMS in the first case

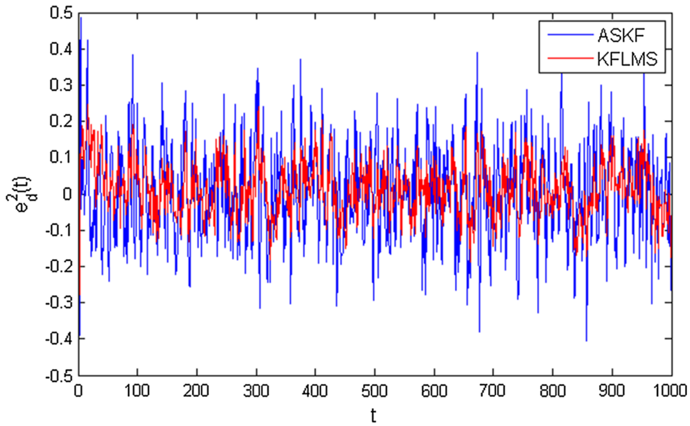


Fig. 4 UI (d_k^2) estimation errors of the ASKF and the KFLMS in the first case

The above observations illustrate the usefulness of the proposed results. Moreover, from the observations (3) and (4), it is shown that the UI filter is more effective than the modified bias filter in obtaining the UI estimates. Note that, in this simulation case the ROUMVF cannot yield the UI estimates due to $H_k = 0$.

To further show why the KFLMS is superior to the ASKF, we illustrate the estimation errors of the UI estimates of the ASKF and the KFLMS in Figs. 3 and 4. From the figures, it is clear that the UI estimates obtained by the adaptive LMS algorithm are more accurate than those obtained by using the random walk model.

The second case In this subsection, we consider the numerical system that satisfies $\bar{T}_k = 0$ and the unknown inputs affect both the state and the measurement equations. The parameters of system (1)–(2) are as follows:

$$A_k = \begin{bmatrix} -0.09 & -0.05 \\ 0.017 & 0.06 \end{bmatrix}, \quad B_k = 0_{2\times 1}, \quad G_k = \begin{bmatrix} 0.95 & 2 \\ -1 & 3 \end{bmatrix}$$

Table 2 Performance of the ASKF, OTSKF, KFLMS, CKFLMS, and ROUMVF in the second case

Filter	rmse(x_k^1)	rmse(x_k^2)	rmse(d_k^1)	rmse(d_k^2)
ASKF	0.1937	0.4061	0.1819	0.0974
OTSKF	0.1937	0.4061	0.1819	0.0974
KFLMS	0.1546	0.3930	0.1177	0.0550
CKFLMS^1	0.1329	0.3794	0.1176	0.0550
CKFLMS^2	0.3343	0.5229	0.3559	0.1531
ROUMVF	3.6336	3.6466	2.4870	1.7476

$$C_k = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad H_k = \begin{bmatrix} 1 & 0.7 \\ 0 & 0 \end{bmatrix}, \quad Q_k = \begin{bmatrix} 0.0036 & 0.0342 \\ 0.0342 & 0.3249 \end{bmatrix}, \quad R_k = \begin{bmatrix} 0.51 & 0 \\ 0 & 0.26 \end{bmatrix}$$

In the simulation, we set $x_0 = [1 \quad 1]^{\mathrm{T}}$ and $a_k = 0.85$ which yield $\mu_k = 0.0054$. The unknown input is considered as follows:

$$d_k = \begin{bmatrix} 0.5U_{k-200}^s - 0.5U_{k-500}^s \\ -0.4U_{k-500}^s + 0.4U_{k-700}^s \end{bmatrix}$$

where U_k^s is the unit-step function.

The obtained rmses of the considered filters are given in Table 2, by which it is clear that, same as case1, the ROUMVF has the worst performance owing to that its gain matrix L_k is in the degeneration form, i.e., $W_k = 0$, because the matrix S_k is of full-row rank, which yields $\bar{T}_k = 0$. Also, it is shown that the filtering performance of the ASKF is satisfactory, which indicates that the assumed UI model (16) works well for this simulation example. Nevertheless, it is also noticed that the performance of the proposed KFLMS is comparable and slightly better than that of the ASKF. The rationale behind this fact is due to that the UI estimates of the former are more accurate than those of the latter, which can also be justified by the simulation results that the CKFLMS^1 is slightly better than the KFLMS and the CKFLMS^2 is slightly worse than the ASKF. The simulation results show that the unknown system identification character of the KFLMS, a robust property inherited from the adaptive LMS algorithm, may be superior to those solely obtained from model-based filtering approaches, e.g., the conventional Kalman filtering approach, when state estimation is applied to systems with arbitrary UIs in the system dynamics. Furthermore, the estimation errors of the UI estimates of the ASKF and the KFLMS are depicted in Figs. 5 and 6, by which it once again shows that the UI estimates obtained by the adaptive LMS algorithm are more accurate than those obtained by using the random walk model.

6 Conclusion

This paper highlights some disadvantages of using unbiased minimum-variance filtering for systems with UIs, and further proposes a Kalman filter reinforced by least mean square method to remedy the problem. This is achieved by using the adaptive

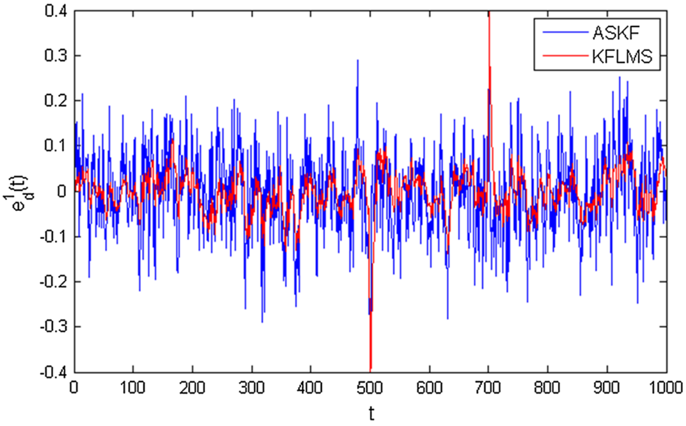


Fig. 5 UI (d_k^1) estimation errors of the ASKF and the KFLMS in the second case

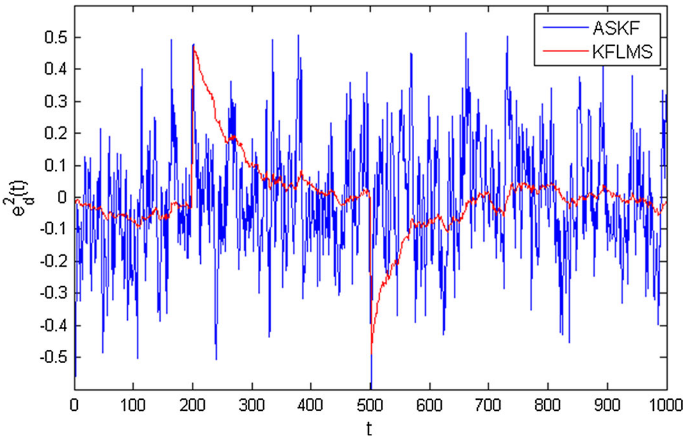


Fig. 6 UI (d_k^2) estimation errors of the ASKF and the KFLMS in the second case

LMS algorithm to estimate the UIs. In the sequel, a new filter named KFLMS is proposed to simultaneously estimate the state and UIs. It is shown by numerical examples that this new obtained robust filter is superior to the well-known Kalman filter. This research also shows that the UI estimates obtained by the adaptive LMS algorithm may be more accurate than those obtained by the conventionally used random walk model. To reduce computational complexity, a useful compact version of the KFLMS named CKFLMS is also proposed. It is shown by simulation results that this compact version is comparable with the KFLMS.

References

1. J. Chen, R.J. Patton, Optimal filtering and robust fault diagnosis of stochastic systems with unknown disturbances. IEE Proc. Control Theory Appl. **143**(1), 31–36 (1996)

2. J. Chen, Y. Liu, X. Wang, Recursive least squares algorithm for nonlinear dual-rate systems using missing-output estimation model. Circuits Syst. Signal Process. **36**(4), 1406–1425 (2017)
3. J. Chen, J. Ma, Y. Liu, F. Ding, Identification methods for time-delay systems based on the redundant rules. Signal Process. **137**, 192–198 (2017)
4. Y. Cheng, H. Ye, Y. Wang, D. Zhou, Unbiased minimum-variance state estimation for linear systems with unknown input. Automatica **45**(2), 485–491 (2009)
5. M. Darouach, M. Zasadzinski, Unbiased minimum variance estimation for systems with unknown exogenous inputs. Automatica **33**(4), 717–719 (1997)
6. T. Floquet, J.-P. Barbot, State and unknown input estimation for linear discrete-time systems. Automatica **42**(11), 1883–1889 (2006)
7. S. Gillijns, B. De Moor, Unbiased minimum-variance input and state estimation for linear discrete-time systems. Automatica **43**(1), 111–116 (2007)
8. S. Gillijns, B. De Moor, Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough. Automatica **43**(5), 934–937 (2007)
9. M.S. Grewal, Kalman filtering, in *International Encyclopedia of Statistical Science* (Springer, Berlin, 2011), pp. 705–708
10. S. Haykin, B. Widrow, *Least-Mean-square Adaptive Filters*, vol. 31 (Wiley, New York, 2003)
11. M. Hou, R.J. Patton, Optimal filtering for systems with unknown inputs. IEEE Trans. Autom. Control **43**(3), 445–449 (1998)
12. C.-S. Hsieh, Robust two-stage Kalman filters for systems with unknown inputs. IEEE Trans. Autom. Control **45**(12), 2374–2378 (2000)
13. C.-S. Hsieh, Optimal minimum-variance filtering for systems with unknown inputs, in *The Sixth World Congress on Intelligent Control and Automation, 2006 (WCICA 2006)*, volume 1 (IEEE, 2006), pp. 1870–1874
14. C.-S. Hsieh, On the global optimality of unbiased minimum-variance state estimation for systems with unknown inputs. Automatica **46**(4), 708–715 (2010)
15. C.-S. Hsieh, On the optimality of two-stage Kalman filtering for systems with unknown inputs. Asian J. Control **12**(4), 510–523 (2010)
16. C.-S. Hsieh, F.-C. Chen, Optimal solution of the two-stage Kalman estimator. IEEE Trans. Autom. Control **44**(1), 194–199 (1999)
17. C.-S. Hsieh, M.A. Majidi, Implementation issues of unbiased minimum-variance state estimation for systems with unknown inputs, in *2014 CACS International Automatic Control Conference (CACS)* (IEEE, 2014), pp. 323–328
18. K.P. Kitanidis, Unbiased minimum-variance linear state estimation. Automatica **23**(6), 775–778 (1987)
19. H.A. Sayed, T. Kailath, A state-space approach to adaptive RLS filtering. IEEE Signal Process. Mag. **11**(3), 18–60 (1994)
20. S. Sundaram, C.N. Hadjicostis, Delayed observers for linear systems with unknown inputs. IEEE Trans. Autom. Control **52**(2), 334–339 (2007)
21. D. Wang, Y. Gao, Recursive maximum likelihood identification method for a multivariable controlled autoregressive moving average system. IMA J. Math. Control Inf. **33**(4), 1015–1031 (2015)