Turkish Journal of Electrical Engineering and Computer Sciences

Volume 24 | Number 2 Article 14

1-1-2016

An improved real-time adaptive Kalman filter with recursive noise covariance updating rules

IYAD HASHLAMON

KEMALETTİN ERBATUR

Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2016) 24: 524 – 540

doi:10.3906/elk-1309-60

An improved real-time adaptive Kalman filter with recursive noise covariance updating rules

Iyad HASHLAMON*, **Kemalettin ERBATUR**
Mechatronics Research Laboratory, Faculty of Engineering and Natural Sciences, Sabancı University, İstanbul, Turkey

Received: 09.09.2013 • **Accepted/Published Online:** 17.12.2013 • **Final Version:** 05.02.2016

Abstract: The Kalman filter (KF) is used extensively for state estimation. Among its requirements are the process and observation noise covariances, which are unknown or partially known in real-life applications. Uncertain and biased values of the covariances result in KF performance degradation or divergence. Unlike previous methods, we are using the idea of the recursive estimation of the KF to develop two recursive updating rules for the process and observation covariances, respectively designed based on the covariance matching principles. Each rule has a tuning parameter that enhances its flexibility for noise adaptation. The proposed adaptive Kalman filter (AKF) proves itself to have an improved performance over the conventional KF and, in the worst case, it converges to the KF. The results show that the AKF estimates are more accurate, have less noise, and are more stable against biased covariances.

Key words: Kalman filter, adaptive Kalman filter, covariance matching

1. Introduction

The Kalman filter (KF) is among the most popular and famous estimation techniques. That is because it incorporates the observer theory and the Bayesian approach. It is a statistically optimal estimator that estimates the instantaneous state(s) of a dynamic system perturbed by noise using noisy observation(s) that are related to the state(s) [1]. Basically, the KF depends on two models: the plant dynamic model, which describes the system behavior over time, and the stochastic models, which describe both the process and observation noises properties .[2,3]. For the best performance of the KF, both the system dynamic model and the noise statistic model parameters must be known. However, in many applications, the stochastic model parameters may be unknown or partially known. As a result, the KF performance degrades or may even diverge [4,5].

The KF algorithm uses noise statistics to influence the KF gain that is applied to the error between the available process information and the most recent obtained measurements. The filter gain projects this error to the process information to get the best estimate. Thus, noise characteristics have a significant importance on KF performance, which motivates the research of developing and improving KF, such that it can adapt itself to the uncertainty in the noise statistical parameters, thus reducing their effects. This type of KF is well known as the adaptive Kalman filter (AKF).

The most often used AKF schemes in the literature can be categorized as innovation-based adaptive estimation (IAE), multiple model adaptive estimation (MMAE), and a scaled state covariance method. A summary of the first two methods can be found in [6].

*Correspondence: hashlamon@sabanciuniv.edu

The IAE method estimates the process noise covariance matrix Q and/or the measurement covariance matrix R assuming that the innovation sequence of KF is white noise. In this method, one of the covariance matching, correlation, and maximum likelihood techniques is used. In the covariance matching technique the sample covariance of the state error is computed based on a moving window or using past data, and Q is estimated using the calculated sample covariance. In the same way, R is computed, but using the sample covariance of the KF innovation [7–10]. In the correlation techniques, the covariance matrices are estimated based on the sample autocorrelations between the innovations by exploiting the relations between the estimation error covariance and the innovation covariance [11–14]. The drawbacks of this method are that it does not guarantee the positive definiteness of the matrices, the estimated covariances are biased [15], and the above techniques require a large window of data, which makes them impractical. Maximum likelihood methods estimate the covariances by maximizing the likelihood function of the innovations [16], but these methods need heavy computations and they can be implemented offline. A modified one using the expectation-maximization algorithm (EM) was reported in [17].

The MMAE method considers a bank of different models, assuming that one of them is the correct one. The Bayesian probability is computed for each model as it is the true model using the measurement sequence, and then the output to be considered is either the output of the model that has the highest probability or a weighting sum of all models [18]. However, the assumption that one of the models is the correct one makes it unsuitable for uncertain dynamic systems [19]. An optimization-based adaptive estimation [20] was developed to handle the model uncertainty.

Introducing a scaling method was reported to improve the filter stability and convergence performance [8,21], where a multiplication factor to the state error covariance matrix is introduced. The factor calculation is either empirical or based on the filter innovations.

Avoiding the aforementioned problems of using a moving window, excessive computations, exact models, and scaling the state error covariance matrix is the main aim of this paper in designing the AKF. Moreover, the AKF core concept is to be able to adapt itself to the biased initial covariances and parameter changes smoothly while running, to increase the estimation accuracy, and to stabilize the filter if the used noise covariance values cause filter divergence.

In this paper, a new method to design an AKF to achieve the above advantages is proposed. It adopts the idea of the recursive estimation of the KF to add two recursive unbiased updating rules for the noise covariances Q and R. The derivation of these rules is based on the covariance matching method. The rule is a scaled value of the previous covariance matrix and a correction covariance error term. This error is calculated at each sample time using the available information about the state covariance error and the most recent measurements and innovations. The updating rules have the capability of tuning the matrices Q and R to get the best performance.

The rest of the paper is organized as follows: Section 2 introduces some preliminaries and states the problem. The proposed AKF algorithm, modeling, derivation, and stability proof are given in Section 3. Section 4 presents the testing results. The paper is concluded in Section 5.

2. Preliminaries and problem statement

Consider the discrete-time linear state space model

$$\begin{aligned} x_k &= Ax_{k-1} + Bu_{k-1} + \upsilon_{k-1} \\ y_k &= Hx_k + v_k \end{aligned}, \tag{1}$$

where, $x_k \in R^n$ is an n-dimensional state vector with initial state value x_0 that has Gaussian distribution of mean m_0 and covariance P_0 (i.e. $x_0 \sim N(m_0, P_0)$), $A \in R^{n\times n}$ is the state matrix, $B \in R^{n\times m}$ is the input matrix, $u \in R^m$ is the system input, $\upsilon \in R^n$ is the Gaussian process noise with zero mean and constant covariance Q (i.e. $\upsilon \sim N(0, Q)$), $v \in R^d$ is the Gaussian measurement noise with zero mean and constant covariance R (i.e. $v \sim N(0, R)$), $y \in R^d$ is a d-dimensional measurement vector, $H \in R^{d\times n}$ is the output matrix, and $k$is the time index. For this system, the matrices $A, B,$ and H are considered to be known at the time instant k, and a random initial state mean m_0 and covariance P_0 are given before applying the KF. The state estimation is carried out under the following assumptions:

Assumption 1: The process and measurement noises are assumed to be independent and mutually uncorrelated with the given means and covariances.

$$\begin{array}{l} E(\upsilon_k) = E(v_k) = E\left(\upsilon_k v_i^T\right) = 0 \\ Q = \delta_{ki} E\left(\upsilon_k \upsilon_i^T\right); R = \delta_{ki} E\left(v_k v_i^T\right) \\ \delta_{ki} = \left\{ \begin{array}{ll} 1 & i = k \\ 0 & i \neq k \end{array} \right\} \end{array}, \tag{2}$$

where $E(\dot{})$ stands for the expectation of $(\dot{})$.

Assumption 2: The inputs are considered to be piecewise constant over the sampling time interval T, i.e. $u(t) = u_{k-1}, t_{k-1} \leq t < t_k = t_{k-1} + T$.

Assumption 3: The noise covariances are considered to be constant.

Assumption 4: The process and measurements have the same sampling time.

Then for the given system in Eq. (1), the conventional KF algorithm is composed of the prediction step as time invariant:

$$\begin{array}{l} \hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \\ P_k^- = AP_{k-1}A^T + Q_{k-1} \end{array}, \tag{3}$$

and the measurement update step:

$$\begin{array}{l} K_k = P_k^- H^T \left(HP_k^- H^T + R_k\right)^{-1} \\ \hat{x}_k = \hat{x}_k^- + K_k\left(z_k - H\hat{x}_k^-\right) \\ P_k = (I - K_k H) P_k^- \end{array}. \tag{4}$$

In Eq. (3) and Eq. (4) the following notation is employed. $(.)^-$ and $(.)$ stand for the prior and posterior estimates, respectively. P is the estimation error covariance matrix and K is the Kalman gain. I is the identity matrix, $\hat{x}$ is the estimated state, and z is the measurement vector with the same dimension as y.

The values of Q and R have an important effect on KF estimates; the estimated state $\hat{x}_k$ will be biased if the value of Q is too small with respect to the correct value, and $\hat{x}_k$ will oscillate around the true value if the value of Q is too large with respect to the correct value [20]. In this work, a real-time update for both Q and R is carried out. Time update rules for both of the matrices are derived as shown in Section 3.1 and Section 3.2.

3. Adaptive Kalman filter

In this part the proposed AKF is explained. It is based on developing two recursive updating rules, R1 and R2, for both of the noise covariances R and Q, respectively. As mentioned earlier, these rules are derived using the IAE method and in particular based on the covariance matching principles. Consider that Assumptions 1–4 hold for the discrete-time linear state space model given in Eq. (1); then, for given initial value matrices R_0 and Q_0, there are constants $0 < \alpha_1 < 1$ and $0 < \alpha_2 < 1$, positive constants N_R and N_Q, and noise covariance errors ΔQ and ΔR such that the KF performance is improved by updating the observation and the process covariance matrices as in the algorithm in Eqs. (12) and (20), respectively. For proof simplification, we will consider that $N_R = N_Q = N$.

The AKF algorithm is given in Eqs. (5)–(20).

$$\text{Initial values } \bar{\omega}_0, \bar{e}_0, \hat{x}_0, P_0, N_R, N_Q, Q_0 > 0, R_0 > 0, \tag{5}$$

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}, \tag{6}$$

$$P_k^- = AP_{k-1}A^T + Q_{k-1}, \tag{7}$$

$$e_k = z_k - H\hat{x}_k^-, \tag{8}$$

$$\alpha_1 = \frac{N_R - 1}{N_R}, \tag{9}$$

$$\bar{e}_k = \alpha_1 \bar{e}_{k-1} + \frac{1}{N_R} e_k, \tag{10}$$

$$\Delta R_k = \frac{1}{N_R - 1} \left(e_k - \bar{e}_k\right)\left(e_k - \bar{e}_k\right)^T - \frac{1}{N_R}\left(HP^- H^T\right)_k, \tag{11}$$

$$R_k = \left|diag\left(\alpha_1 R_{k-1} + \Delta R_k\right)\right|, \tag{12}$$

$$K_k = P_k^- H^T \left(HP_k^- H^T + R_k\right)^{-1}, \tag{13}$$

$$\hat{x}_k = \hat{x}_k^- + K_k e_k, \tag{14}$$

$$P_k = \left(I - K_k H\right) P_k^-, \tag{15}$$

$$\hat{\omega}_k = \hat{x}_k - \hat{x}_k^-, \tag{16}$$

$$\alpha_2 = \frac{N_Q - 1}{N_Q}, \tag{17}$$

$$\bar{\omega}_k = \alpha_2 \bar{\omega}_{k-1} + \frac{1}{N_Q}\hat{\omega}_k, \tag{18}$$

$$\Delta Q_k = \frac{1}{N_Q}\left(P_k - AP_{k-1}A^T\right) + \frac{1}{N_Q - 1}\left(\hat{\omega}_k - \bar{\omega}_k\right)\left(\hat{\omega}_k - \bar{\omega}_k\right)^T, \tag{19}$$

$$Q_k = \left|diag\left(\alpha_2 Q_{k-1} + \Delta Q_k\right)\right|. \tag{20}$$

The differences between this AKF and the sequential filter in [7] can be stated as follows: the sequential filter requires to store N_Q and N_R samples, which are used to calculate the noise covariances. Moreover, at each time a shifting operation is carried out to update the stored samples to include the most recent one. In the AKF, however, the updating rules require neither sample storage nor shifting operation so that they are different.

3.1. Observation covariance matrix update rule (R1) proof

Since the true values of the states are unknown, the observation covariance cannot be estimated using the observation-state relation in Eq. (1). Therefore, the estimation innovation $e_k = z_k - H\hat{x}_k^-$ and its covariance $\sum$ in Eq. (21) are used for this purpose.

$$\sum = HP_k^- H^T + R_k. \tag{21}$$

For the sake of proof only, let us assume that a number of samples N is available in a buffer as shown in Figure 1. Before arriving to the new sample, the samples are ordered from the oldest available sample in the buffer index $i = 1$ to the most recent sample stored at the end of the buffer with a buffer index $i = N$, or, in terms of time instants, they are ordered from the time index $k - N$ for the oldest sample to the time index k for the most recent sample. When a new sample arrives, it is stored in $i = N$ or time index k and each of the samples in the buffer is delayed by one unit delay, and hence the sample that was at $i = 1$ or $k - N$ becomes $k - N - 1$ and it is neglected. As stated in [7], the expected value of the sample covariance $\sum_s$ is the mean of Eq. (21). In mathematical terms, the value of $\sum_s$ for the samples up to k is

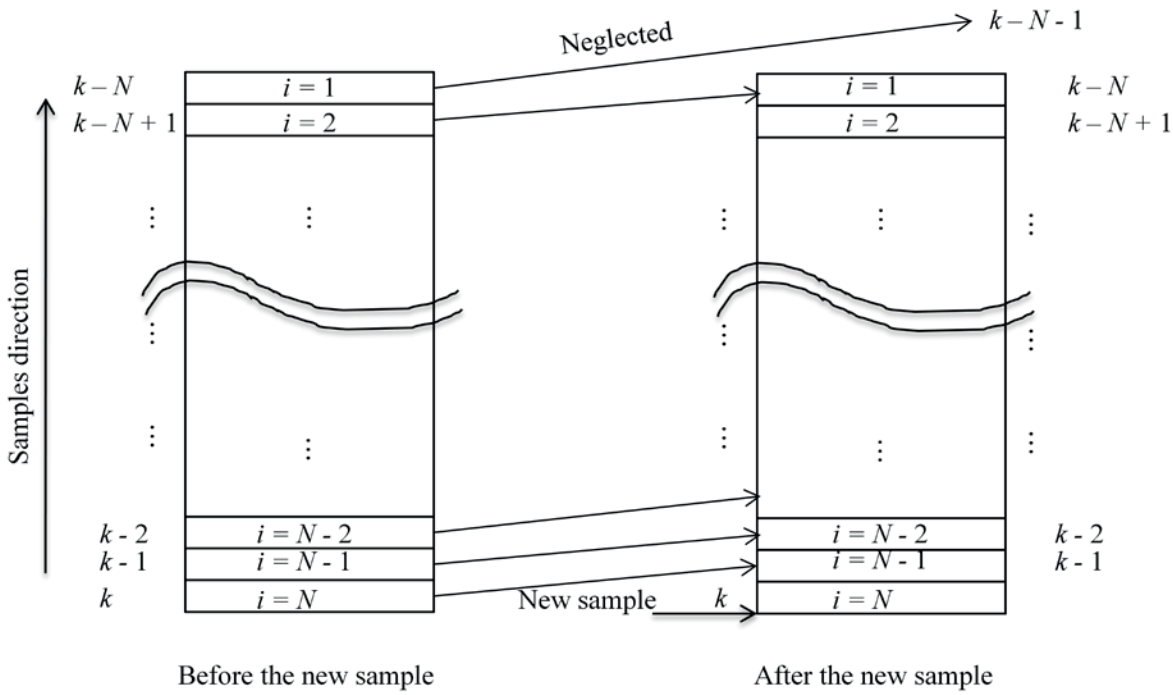

Figure 1. Samples' buffer.

$$\sum_{s,k} = \frac{1}{N-1} \sum_{i=k-N}^{k} (e_i - \bar{e})(e_i - \bar{e})^T, \tag{22}$$

where

$$\bar{e} = \frac{1}{N} \sum_{i=k-N}^{k} e_i \tag{23}$$

is the innovation mean, and the covariance mean $\bar{\sum}_k$ is

$$\bar{\sum}_k = E\left(\sum_{s,k}\right) = \frac{1}{N}\sum_{i=k-N}^{k}\left(HP^{-}H^{T}\right)_i + R_k. \tag{24}$$

Then, based on Assumption 3 and Eqs. (22) and (24), the estimated value of the observation noise can be written as [7]:

$$R_k = \frac{1}{N-1}\sum_{i=k-N}^{k}\left(e_i - \bar{e}\right)\left(e_i - \bar{e}\right)^T - \frac{1}{N}\left(\sum_{i=k-N}^{k}\left(HP^{-}H^{T}\right)_i\right). \tag{25}$$

The samples are divided into two groups; the first one contains all the samples up to sample $N-1$, which arrived at the time instant $k-1$, and the other group contains only the most recent sample that arrived at the time instant k. After some mathematical manipulation, Eq. (25) can be restated as

$$R_k = \frac{N-1}{N}X + \Delta R_k, \tag{26}$$

where

$$X = \frac{N}{(N-1)^2}\sum_{i=k-N}^{k-1}\left(e_i - \bar{e}\right)\left(e_i - \bar{e}\right)^T - \frac{1}{N-1}\sum_{i=k-N}^{k-1}\left(HP^{-}H^{T}\right)_i, \tag{27}$$

and

$$\Delta R_k = \frac{1}{N-1}\left(e_k - \bar{e}_k\right)\left(e_k - \bar{e}_k\right)^T - \frac{1}{N}\left(HP^{-}H^{T}\right)_k. \tag{28}$$

The term $\bar{e}_k$ is unknown and the term X in Eq. (27) depends on the samples up to sample $k-1$. As mentioned before we divided them into two groups. Taking the sample covariance of the first group as

$$\sum_{s,k-1} = \frac{1}{N-2}\sum_{i=k-N}^{k-1}\left(e_i - \bar{e}\right)\left(e_i - \bar{e}\right)^T, \tag{29}$$

and comparing Eq. (29) with the first term in X in Eq. (27), it is obvious that if the value of $N\Big/(N-1)^2$ can be approximated by $1/N-2$, then this term is the observation noise covariance of all the samples up to the previous arrived sample $k-1$. Let us define the error ε as

$$\varepsilon = \frac{N}{(N-1)^2} - \frac{1}{N-2}, \tag{30}$$

and then this error converges to zero as the number of samples increase, i.e. $\lim_{N\to\infty}\varepsilon \to 0$ as shown in Figure 2.

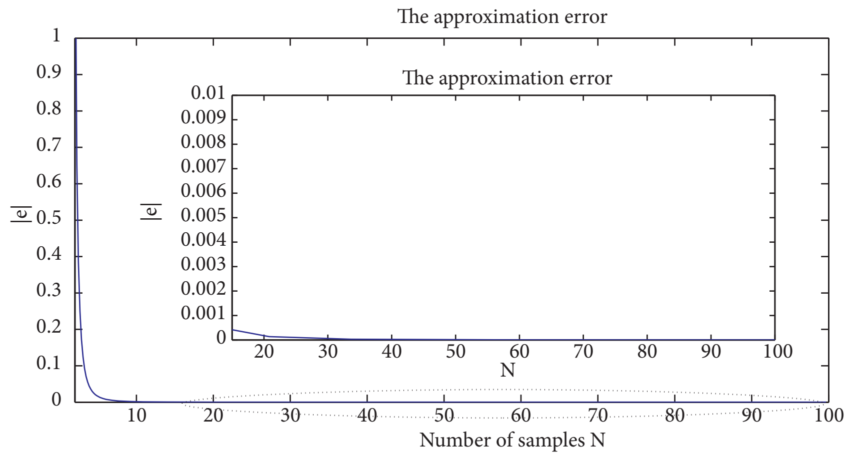

Figure 2. The approximation error.

Then, for large values of N, Eq. (27) can be approximated by

$$X \approx \frac{1}{N-2}\sum_{i=k-N}^{k-1}(e_i-\bar{e})(e_i-\bar{e})^T - \frac{1}{N-1}\sum_{i=k-N}^{k-1}\left(HP^-H^T\right)_i, \tag{31}$$

which is the observation covariance for the $N-1$ samples ordered from $i=k-N$ to $k-1$, i.e. R_{k-1}. Then by referring to Eq. (25), Eq. (31) yields:

$$R_{k-1} \approx \frac{1}{N-2}\sum_{i=k-N}^{k-1}(e_i-\bar{e})(e_i-\bar{e})^T - \frac{1}{N-1}\sum_{i=k-N}^{k-1}\left(HP^-H^T\right)_i. \tag{32}$$

The same method is used to compute $\bar{e}_k$ as follows:

$$\bar{e}_k = \frac{1}{N}\sum_{i=k-N}^{k} e_i = \frac{1}{N}\sum_{i=k-N}^{k-1} e_i + \frac{1}{N}e_k, \tag{33}$$

and this yields:

$$\bar{e}_k = \alpha_1\bar{e}_{k-1} + \frac{1}{N}e_k, \tag{34}$$

with

$$\alpha_1 = \frac{N-1}{N}. \tag{35}$$

This finalizes the proof of R1.

3.2. Process covariance matrix update rule proof

The true values of the states are unknown; the best known values are the estimated and updated values of the states. Considering the estimated and updated state $\hat{x}$ and the estimated but not updated state $\hat{x}^-$ as defined in Eq. (3), the state error can be written as

$$\hat{\omega}_k = \hat{x}_k - \hat{x}_k^-. \tag{36}$$

By using the state error covariance, and following the same procedure for proving R1, one can write Q_k as [7]:

$$Q_k = \frac{1}{N-1}\sum_{i=k-N}^{k}\left(\hat{\omega}_i - \bar{\omega}\right)\left(\hat{\omega}_i - \bar{\omega}\right)^T + \left(\frac{1}{N}\sum_{i=k-N}^{k} P_i - AP_{i-1}A^T\right), \tag{37}$$

When the same approximation in R1 is used here too, and the result is a proof R2, it ends up with Eq. (20).

3.3. Stability of AKF

In the observer theory and for an observable system, the exponential behavior of the AKF is determined based on the exponential convergence of the dynamic error between the state x and the estimated state $\hat{x}$. Therefore, to analyze the exponential behavior of the AKF, we first derive the error ε between x and $\hat{x}$. Then we recall the exponential stability of linear discrete time systems as stated in Definition 1. After that, we propose our theorem of the AKF stability and its proof.

Let the error ε be defined by $\varepsilon_k = x_k - \hat{x}_k$, where

$$x_k = Ax_{k-1} + Bu_{k-1}, \tag{38}$$

and

$$\hat{x}_k = A\hat{x}_{k-1} + Bu_{k-1} + K_k\left(Hx_k - H\hat{x}_k^-\right). \tag{39}$$

Then after mathematical manipulation, this error ε can be expressed as

$$\varepsilon_k = \left(A - K_k HA\right)\varepsilon_{k-1}. \tag{40}$$

Calling the eigenvalues of $(A - K_k HA)$ as $\lambda_i, i = 1, \cdots, n$, then the system is exponentially stable if $\|\lambda_i\| < 1$, i.e. they are inside the unit circle in the z-plane.

The exponentially stability is proven here based on Lyapunov function theory and follows the same approach as in [22,23]. The following definitions and lemmas are employed for the sake of completeness and proof.

Definition 1 [24] *The origin of the difference Eq. (40) is an exponentially stable equilibrium point if there is a continuous differentiable positive definite function $V(\varepsilon_k)$ such that*

$$c_1\|\varepsilon_k\|^2 \le V(\varepsilon_k) \le c_2\|\varepsilon_k\|^2, \Delta V(\varepsilon_k) \le -c_3\|\varepsilon_k\|^2, \tag{41}$$

for positive constants $c_1, c_2,$ and c_3 with ΔV as the rate of change of V and defined by

$$\Delta V = V(\varepsilon_k) - V(\varepsilon_{k-1}). \tag{42}$$

For the sake of completeness, the exponential stability for discrete time systems is defined by the inequality $\|\varepsilon_k\| \le \beta\|\varepsilon_0\|\Upsilon^k$ for all $k \ge 0$ with $\beta > 0$ and $0 < \Upsilon < 1$ [24].

Definition 2 *The observer in Eq. (39) is an exponential observer if Eq. (40) has an exponentially stable equilibrium at 0.*

Lemma 1 [25] *The matrix inequality lemma: For given nonsingular matrices* Γ *and* Λ *with suitable dimensions, and assuming that* $\Gamma^{-1}+\Lambda$ *is also nonsingular, then*

$$\left(\Gamma^{-1}+\Lambda\right)^{-1}=\Gamma-\Gamma\left(\Gamma+\Lambda^{-1}\right)^{-1}\Gamma. \tag{43}$$

Definition 3 *If* A *is an invertible matrix, and for the positive definite matrices* P_k^- *and* P_k*, then*

$$P_k^{-1}\leq\left(I-K_kH\right)^{-T}A^{-T}\left(P_{k-1}^{-1}-P_{k-1}^{-1}\left(P_{k-1}^{-1}+A^TQ_{k-1}^{-1}A\right)^{-1}P_{k-1}^{-1}\right)A^{-1}\left(I-K_kH\right)^{-1}. \tag{44}$$

Proof Rewriting Eq. (15) as in [2]:

$$P_k=\left(I-K_kH\right)P_k^-\left(I-K_kH\right)^T+K_kR_kK_k^T, \tag{45}$$

then we have

$$P_k\geq\left(I-K_kH\right)P_k^-\left(I-K_kH\right)^T. \tag{46}$$

Inverting Eq. (46) results in

$$P_k^{-1}\leq\left(I-K_kH\right)^{-T}\left(P_k^-\right)^{-1}\left(I-K_kH\right)^{-1}, \tag{47}$$

The expression of $\left(P_k^-\right)^{-1}$ is obtained by rearranging Eq. (7) as

$$P_k^-=A\left(P_{k-1}+A^{-1}Q_{k-1}A^{-T}\right)A^T, \tag{48}$$

and inverting Eq. (48) yields

$$\left(P_k^-\right)^{-1}=A^{-T}\left(P_{k-1}^{-1}-P_{k-1}^{-1}\left(P_{k-1}^{-1}+A^TQ_{k-1}^{-1}A\right)^{-1}P_{k-1}^{-1}\right)A^{-1}. \tag{49}$$

As a result from Eq. (49), Eq. (47) is expressed by:

$$P_k^{-1}\leq\left(I-K_kH\right)^{-T}A^{-T}\left(P_{k-1}^{-1}-P_{k-1}^{-1}\left(P_{k-1}^{-1}+A^TQ_{k-1}^{-1}A\right)^{-1}P_{k-1}^{-1}\right)A^{-1}\left(I-K_kH\right)^{-1}, \tag{50}$$

which completes the proof. □

The Stability Theorem The given system in Eq. (1) with the proposed AKF is exponentially stable if there are positive real numbers $a,p,\bar{p},q>0$ such that

$$pI\leq P_{k-1}\leq\bar{p}I, \tag{51}$$

$$\|A\|\leq a, \tag{52}$$

where q is the minimum eigenvalue of the update rule in Eq. (20).

Proof The proof is similar to [22]. Consider the positive definite Lyapunov function

$$V\left(\varepsilon_{k-1}\right)=\varepsilon_{k-1}^TP_{k-1}^{-1}\varepsilon_{k-1}, \tag{53}$$

with $V(0) = 0$. Eqs. (51) and (53) imply that

$$\frac{1}{\bar{p}}\left\|\varepsilon_{k-1}\right\|^2 \leq V\left(\varepsilon_{k-1}\right) \leq \frac{1}{p}\left\|\varepsilon_{k-1}\right\|^2. \tag{54}$$

Then for $V(\varepsilon_k)$ we obtain

$$V\left(\varepsilon_k\right) = \varepsilon_k^T P_k^{-1} \varepsilon_k. \tag{55}$$

From Eq. (44), Eq. (55) can be written as:

$$\begin{array}{c} V\left(\varepsilon_k\right) \leq \varepsilon_{k-1}^T A^T \left(I - K_k H\right)^T \left(I - K_k H\right)^{-T} A^{-T} \\ \times\left(P_{k-1}^{-1} - P_{k-1}^{-1}\left(P_{k-1}^{-1} + A^T Q_{k-1}^{-1} A\right)^{-1} P_{k-1}^{-1}\right) \\ \times A^{-1}\left(I - K_k H\right)^{-1}\left(I - K_k H\right) A \varepsilon_{k-1} \end{array}. \tag{56}$$

After some mathematical manipulation, the rate of change in the Lyapunov function is:

$$\Delta V = V\left(\varepsilon_k\right) - V\left(\varepsilon_{k-1}\right) \leq -\varepsilon_{k-1}^T P_{k-1}^{-1}\left(P_{k-1}^{-1} + A^T Q_{k-1}^{-1} A\right)^{-1} P_{k-1}^{-1} \varepsilon_{k-1}. \tag{57}$$

Applying the bounds in Eqs. (51) and (52), Eq. (57) is expressed as

$$\Delta V = V\left(\varepsilon_k\right) - V\left(\varepsilon_{k-1}\right) \leq -\frac{1}{\bar{p}^2\left(\frac{1}{p} + \frac{a^2}{q}\right)}\left\|\varepsilon_{k-1}\right\|^2, \tag{58}$$

which satisfies Eq. (41), and thus the origin of Eq. (40) is exponentially stable. This concludes that the AKF is an exponentially stable observer and $\|\lambda_i\| < 1, i = 1, \cdots, n$ is satisfied. □

In terms of states and degree of stability, using Eqs. (58) and (54) we can write:

$$V\left(\varepsilon_k\right) \leq \left(1 - \frac{p}{\bar{p}^2\left(\frac{1}{p} + \frac{a^2}{q}\right)}\right) V\left(\varepsilon_{k-1}\right), \tag{59}$$

and since

$$V\left(\varepsilon_{k-1}\right) \leq \left(1 - \frac{p}{\bar{p}^2\left(\frac{1}{p} + \frac{a^2}{q}\right)}\right) V\left(\varepsilon_{k-2}\right), \cdots, \tag{60}$$

we can obtain

$$V\left(\varepsilon_k\right) \leq \left(1 - \frac{p}{\bar{p}^2\left(\frac{1}{p} + \frac{a^2}{q}\right)}\right)^k V\left(\varepsilon_0\right), \tag{61}$$

and then,

$$\left\|\varepsilon_k\right\| \leq \sqrt{\frac{\bar{p}}{p}}\left(\sqrt{1 - \frac{p}{\bar{p}^2\left(\frac{1}{p} + \frac{a^2}{q}\right)}}\right)^k \left\|\varepsilon_0\right\|, k \geq 0. \tag{62}$$

Recalling Definition 1, we have

$$\beta = \sqrt{\frac{\bar{p}}{p}} > 0, \tag{63}$$

and

$$\Upsilon = \sqrt{1 - \frac{p}{\bar{p}^2\left(\frac{1}{p} + \frac{a^2}{q}\right)}}, 0 < \Upsilon < 1. \tag{64}$$

3.4. Implementation considerations

In numerical applications, the updated values of R_k and Q_k may become negative and may violate Assumption 1; this explains why we are taking the absolute value of the diagonal in the AKF algorithm in Eq. (12) and Eq. (20).

N is the tuning parameter; its value is user defined based on the system noise characteristics. For very big values of N, the AKF tends to be close to the KF. Basically, for a noisy system, it is much better to give more weight to the previous known values (R_{k-1}, Q_{k-1}) than the current noisy reading, and this is achieved by selecting a big N. In the same context, a small N gives more weight to the current reading $(\Delta R_k, \Delta Q_k)$ for less noisy systems.

4. Numerical results

In this part, the updating rules behavior of the KF performance is studied. For this analysis the following model is used:

$$\begin{aligned} x_k &= \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} = \begin{bmatrix} 1 & -T \\ 0 & 1 \end{bmatrix} x_{k-1} + \begin{bmatrix} T \\ 0 \end{bmatrix} u_{k-1} + \upsilon_{k-1} \\ y_k &= [1 \quad 0]\, x_k + v_k \end{aligned} . \tag{65}$$

This model is used to estimate the gyroscope bias x_2 and the rotation angle x_1 via integrating the gyroscope signal. The input u is the measured angular velocity from the gyroscope, and T is the integration time interval. The output is the angle that is considered available through some measurements or sensor fusion techniques. The Gaussian noise is generated by the MATLAB Simulink Gaussian noise generator. In this simulation, the simulation parameters, initializations, and constant true values Q_{true} and R_{true} along with the corresponding number of samples N_Q for R2 and N_R for R1 are listed in Table 1. Note that the values of N_Q and N_R are different from each other since they do not have to be the same in practice. The notation Q_{small} indicates that the used process covariance noise in the KF is smaller than the true process covariance noise Q_{True} and Q_{big} indicates that the used process covariance noise in the KF is larger than the true process covariance noise Q_{True}. The same definition holds for R_{small} and R_{big}. These notations are used with the KF while $Q_{initial}$ and $R_{initial}$ refer to the use of the AKF. I_n is an $n \times n$ identity matrix.

We chose the true value of x_2 to be a function of time t as in Eq. (66) to show the effect of its changing on the KF and the AKF performances.

$$x_2 = \left\{ \begin{array}{ll} 0.5 & 0 \le t \le 20 \\ 1 & 20 < t \le 50 \\ 0 & else \end{array} \right\}. \tag{66}$$

The aim of the new AKF is to improve the performance in terms of better convergence, stability, and noise adaptation, and to reduce the effect of biased initializations of noise covariance matrices. The results of the AKF are compared with the results of the KF running in the same conditions; in other words, the values of x_0, P_0, R_{True}, and Q_{True} are the same for both filters. The mean square error (MSE) is used as a performance measure.

Table 1. Model parameters and initializations.

Simulation parameters and common initializations		Experiment	AKF initializations		KF initializations	
			Parameter	Value	Parameter	Value
Parameter	Value	Q_{small} R_{small}	Q_0	10^{-8}	Q_0	10^{-8}
T	0.01 s		R_0	10^{-4}	R_0	10^{-4}
x_0	$\begin{bmatrix} 0 & 0 \end{bmatrix}^T$		N_Q	600	–	–
R_{true}	10^{-2}		N_R	2000	–	-
Q_{true}	$\begin{bmatrix} 10^{-4} & 0 \\ 0 & 10^{-6} \end{bmatrix}$	Q_{big} R_{big}	Q_0	1	Q_0	1
P_0	$100I_2$		R_0	1	R_0	1
$\bar{\omega}_0$	$\begin{bmatrix} 0 & 0 \end{bmatrix}^T$		N_Q	600	–	–
$\bar{e}_0$	0		N_R	2000	–	–
–	-	$N_{small} = 50$	Q_0	1	–	–
–	–		R_0	1	–	–
–	–	$N_{big} = 100000$	Q_0	1	–	–
–	–		R_0	1	–	–

Since the true observation noise is much higher than the process noise, the value of N_R is selected to be much bigger than N_Q for the experiment with covariance values Q_{small} and R_{small}, and the experiment with covariance values Q_{big} and R_{big}, as tabulated in Table 1.

For the cases when the noise covariance value is smaller or larger than the true values, the response is as depicted in Figure 3. The estimated bias $\hat{x}_2$ with small noise covariance R_{small} and Q_{small} is shown in Figure 3a; although the AKF starts noisy, it is smoothed as the update rules are converging, and thus both filters display similar performance. However, this is not the case with large noise covariance R_{big} and Q_{big}. The KF performance is noisy along the estimation interval while the AKF performance is smoothed as in Figure 3b. The same values of covariances are used in estimating $\hat{x}_1$ as in Figure 3c. The results show that the KF performance is highly sensitive to the changes in the bias for small covariances; it starts with a relatively low MSE until the changes in the bias take place. At time instant 20, the MSE increases with an effort by the KF to adapt its performance slowly; however, this goes less well at time instant 50. On the other hand, and under the same conditions, this sensitivity is dramatically decreased when using the AKF, as is clear from the values of the MSE at the same time instants. This shows the faster adaptation to the changes and better convergence of the AKF. For "big" covariance values, although the KF sensitivity is stable, the AKF still has lower MSE and thus better performance.

Furthermore, some values for the covariances may slow the KF response or even cause divergence. This is because the filter gain depends on Q and R. From Eq. (13), the filter gain increases with increasing Q and

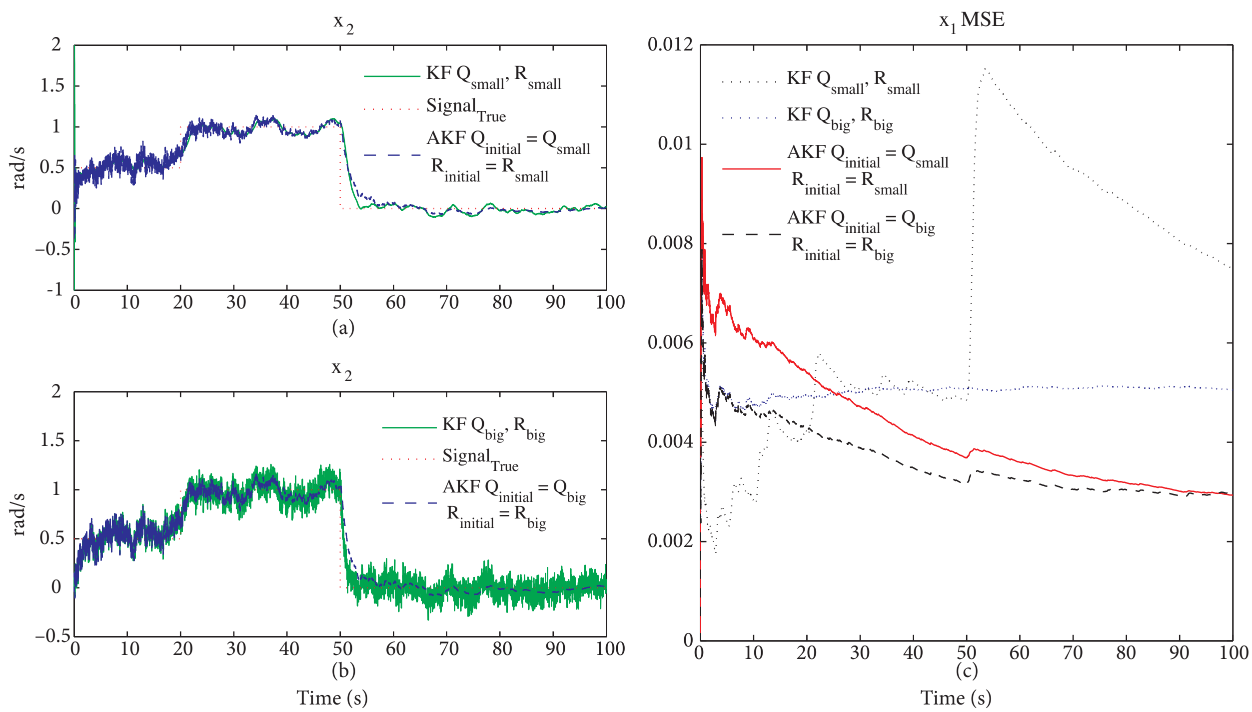

Figure 3. (a) The estimated $\hat{x}_2$ using KF and AKF with small noise covariance values, (b) the estimated $\hat{x}_2$ using KF and AKF with big noise covariance values, and (c) the MSE for the estimated state x1 using KF and AKF with small and big noise covariance values.

decreasing R. If the selected value of Q is very small compared to Q_{True}, then the resultant gain is small. In terms of convergence, if $q \to 0$, then the value of $\Upsilon \to 1$ and thus the convergence of Eq. (62) is slowed down. As an example, selecting $Q = 10^{-10}I$ and $R = 10^{-1}$ results in gain trajectories for the KF and AKF as is clear in Figure 4a for the first gain component and Figure 4b for the second gain component. KF gain (blue) converges to a very small value very fast and thus the estimated bias $\hat{x}_2$ tracking is very slow; however, the AKF estimated bias follows the true bias, as depicted in Figure 5a. Moreover, the estimated state $\hat{x}_1$ is diverging using KF, and this problem is solved with the AKF, which keeps the stability of the filter and forces it to converge, as in Figure 5b. The overall performance of divergence is plotted in Figure 5c and its zoomed version is given in Figure 5d.

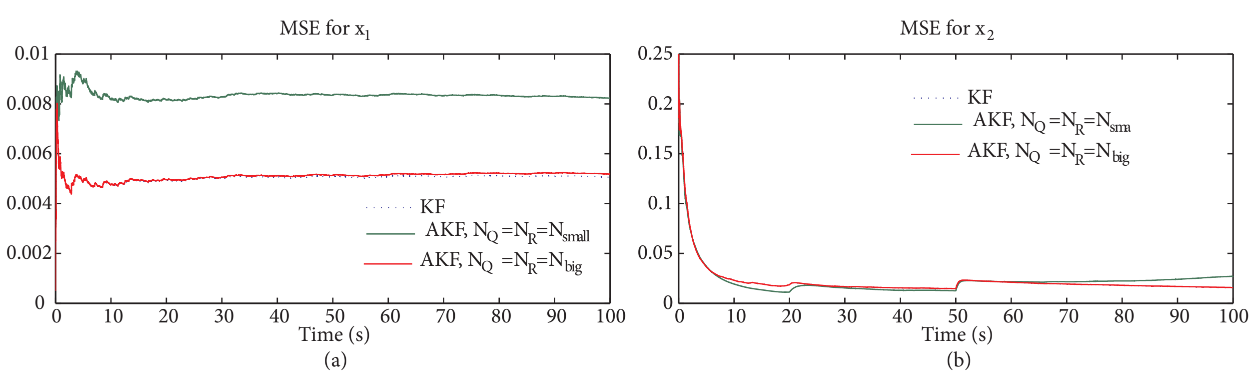

Figure 4. (a) The first component of the gain for KF and the AKF, and (b) the second component of the gain for KF and the AKF.

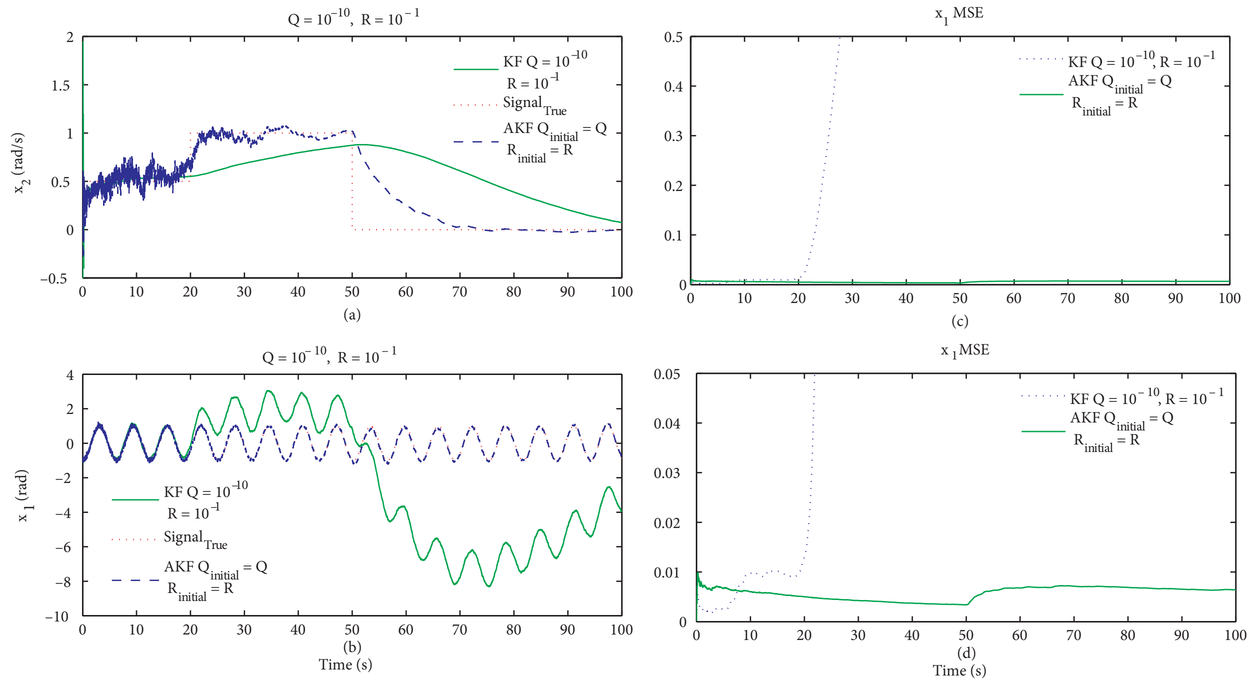

Figure 5. (a) The estimated $\hat{x}_2$ using KF and AKF with noise covariance values that lead to divergence, (b) the estimated $\hat{x}_1$ using KF and AKF with noise covariance values that lead to divergence, (c) the MSE for the estimated state $\hat{x}_1$ using KF and AKF with noise covariance values that lead to divergence, and (d) zoomed version of part (c).

The AKF having better tracking can be explained as follows: the AKF gain is changing based on the estimation performance. This is clear since it depends on both the innovation error e in Eq. (8) and the state error $\hat{\omega}$ in Eq. (16) through the covariances in Eq. (12) and Eq. (20), respectively. If the tracking is not satisfied then the gain will not converge to a very small value as with the KF. The reason is that the value of the noise covariance will increase to change the gain for better performance. In the same context, increasing the noise covariance increases q, which leads to better convergence, as in Eq. (62). This is clear from checking the trajectory of R in Figure 6a and the trajectories of $Q(1,1)$ and $Q(2,2)$ in Figures 6b and 6c, respectively; it is obvious that they are increasing before convergence with a larger rate of change and higher values for Q components than for R. This increases the gain and keeps the tracking ability of the filter. For the same dialog, we can see that big initial noise covariance values decrease without an overshoot to agree with the required gain value. As a final note, the noise covariance components converge, too. Thus, we can claim that the proposed AKF has better stability and convergence performances than the KF shows.

The effect of selecting big and small values of N_R and N_Q is studied too, as in Figure 7. The responses of the day AKF and KF are quite close when we select $N_R = N_Q = 10^5$, which is very big. This is shown in Figures 7a and 7b for both of the states where the MSE of the KF (blue line) and the MSE of the AKF (red line) display very similar behaviors. This supports our claim that in the worst case, selecting very big values of N_R and N_Q will lead the AKF to converge to the KF. However, when their values are very small, $N_R = N_Q = 50$, it leads to an increase in the MSE, as in Figure 7a, and it even may start diverging as in Figure 7b after 80 s.

As a numerical comparison, for all tested experiments in Table 1, the convergence values of the MSE in the case of convergence or the last value in case of divergence for both of the states are tabulated in Table 2. As is clear, for different values of noise covariances, the AKF has a lower MSE than the KF. In the case where the MSE is 0.0163 for the AKF, it is still close to the KF MSE value of 0.0159. Moreover, the MSE for the AKF is stable and smooth for different noise covariance values, while it has a considerable fluctuation in the case of the KF. The divergence in the KF is shown when the MSE is 16.1040. For the experiment where the values of N_Q and N_R are big, Table 2 tells us that both of the filters display the same behavior. However, for very small values of N_Q and N_R, the AKF is no longer applicable, especially when the MSE starts increasing after time instant 80 as mentioned before. The reason for this is that the small values of N_Q and N_R violate the approximation assumption in Eq. (30).

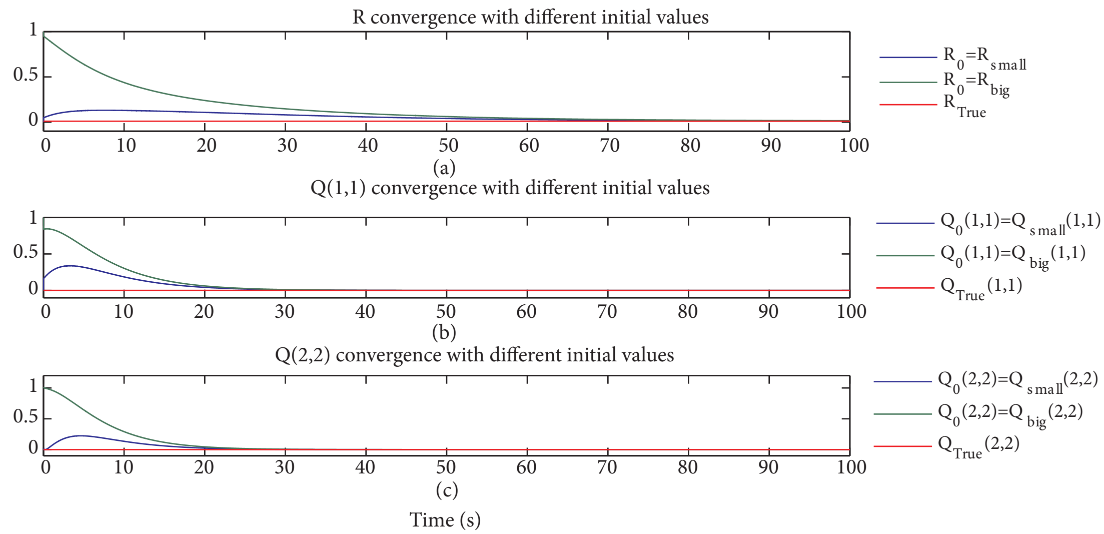

Figure 6. (a) Noise observation covariance trajectory as it is updated with different initial values, (b) the first element of the process noise trajectory as it is updated with different initial values, and (c) the last element of the process noise trajectory as it is updated with different initial values.

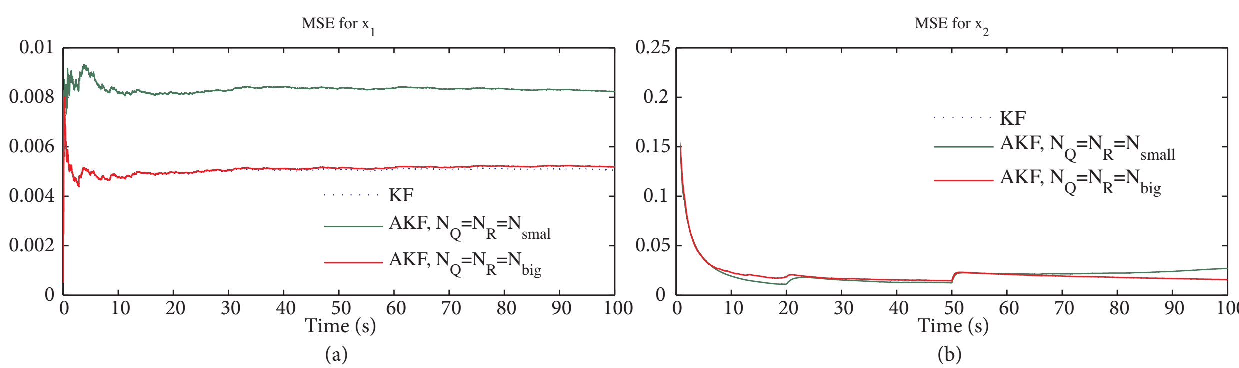

Figure 7. (a) The MSE for the estimated $\hat{x}_1$ using KF and AKF with different values of N_Q and N_R, and (b) the MSE for the estimated $\hat{x}_2$ using KF and AKF with different values of N_Q and N_R.

5. Conclusion

A new real-time AKF for systems with constant noise covariances is proposed. This AKF relates the filter gain to the innovation error through the noise covariance updating rules; this relation changes the gain for better tracking and performance. In terms of Q, it changes the minimum eigenvalue of Q to achieve better

Table 2. Numerical performance comparison between AKF and KF.

		MSE				
		Q_{small}, R_{small}	Q_{big}, R_{big}	$Q = 10^{-10}, R = 10^{-1}$	$N_{small} = 50$	$N_{big} = 100,000$
AKF	$\hat{x}_1$	0.0029	0.0030	0.0029	0.0083	0.0052
	$\hat{x}_2$	0.0171	0.0163	0.0170	0.0237	0.0158
KF	$\hat{x}_1$	0.0075	0.0051	16.1040	0.0051	0.0051
	$\hat{x}_2$	0.0222	0.0159	0.1871	0.0159	0.0159

convergence time. In this AKF, the tuning is not applied to the state covariance matrix. Furthermore, it has fewer tuning parameters; instead of tuning all of the diagonal elements of the noise covariance matrix, they can be initialized and then tuned using N_Q and N_R. Moreover, it requires neither sample storage nor shifting operation. However, the selection of values of N_Q and N_R requires some experience, and this motivates to look for a tuning mechanism as future research. The results show that the AKF has better performance than the KF as long as the values of N_Q and N_R are not very small. Moreover, in the worst case, the AKF converges to the KF for very big values of N_Q and N_R.

References

[1] Grewal MS, Andrews AP. Kalman Filtering: Theory and Practice Using MATLAB. 2nd ed. New York, NY, USA: John Wiley & Sons, 2001.

[2] Simon D. Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches. Hoboken, NJ, USA: Wiley & Sons, 2006.

[3] Maybeck PS. Stochastic Models, Estimation and Control. Vol. 3. New York, NY, USA: Academic Press, 1982.

[4] Heffes H. The effect of erroneous models on the Kalman filter response. IEEE T Automat Contr 1966; 11: 541–543.

[5] Fitzgerald RJ. Divergence of the Kalman filter. IEEE T Automat Contr 1971; 16: 736–747.

[6] Mohamed A, Schwarz K. Adaptive Kalman filtering for INS/GPS. J Geodesy 1999; 73: 193–203.

[7] Myers K, Tapley BD. Adaptive sequential estimation with unknown noise statistics. IEEE T Automat Contr 1976; 21: 520–523.

[8] Ding W, Wang J, Rizos C, Kinlyside D. Improving adaptive Kalman estimation in GPS/INS integration. J Navigation 2007; 60: 517.

[9] Shi E. An improved real-time adaptive Kalman filter for low-cost integrated GPS/INS navigation. In: IEEE 2012 International Conference on Measurement, Information and Control (MIC) 18–20 May 2012; Harbin, China. New York, NY, USA: IEEE. pp. 1093–1098.

[10] Hide C, Moore T, Smith M. Adaptive Kalman filtering for low-cost INS/GPS. J Navigation 2003; 56: 143–152.

[11] Mehra RK. On the identification of variances and adaptive Kalman filtering. IEEE T Automat Contr 1970; 15: 175–184.

[12] Bélanger PR. Estimation of noise covariance matrices for a linear time-varying stochastic process. Automatica 1974; 10: 267–275.

[13] Oussalah M, Schutter JD. Adaptive Kalman filter for noise identification. In: Proceedings of the International Seminar on Modal Analysis; 2001; Kissimmee, FL, USA. Leuven, Belgium: KU Leuven. pp. 1225–1232.

[14] Odelson BJ, Rajamani MR, Rawlings JB. A new autocovariance least-squares method for estimating noise covariances. Automatica 2006; 42: 303–308.

[15] Odelson BJ, Lutz A, Rawlings JB. The autocovariance least-squares method for estimating covariances: application to model-based control of chemical reactors. IEEE T Contr Syst T 2006; 14: 532–540.

[16] Raghavan H, Tangirala AK, Bhushan Gopaluni R, Shah SL. Identification of chemical processes with irregular output sampling. Control Eng Pract 2006; 14: 467–480.

[17] Bavdekar VA, Deshpande AP, Patwardhan SC. Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter. J Process Contr 2011; 21: 585–601.

[18] Wang X. Vehicle health monitoring system using multiple-model adaptive estimation. MSc, University of Hawaii at Manoa, Manoa, HI, USA, 2003.

[19] Rong Li X, Youmin Z. Multiple-model estimation with variable structure part V: Likely-model set algorithm. IEEE T Aero Elec Sys 2000; 36: 448–466.

[20] Karasalo M, Hu X. An optimization approach to adaptive Kalman filtering. Automatica 2011; 47: 1785–1793.

[21] Yang Y, Gao W. An optimal adaptive Kalman filter. J Geodesy 2006; 80: 177–183.

[22] Biçer C, Babacan EK, Özbek L. Stability of the adaptive fading extended Kalman filter with the matrix forgetting factor. Turk J Electr Eng Co 2012; 20: 819–833.

[23] Babacan EK, Ozbek L, Efe M. Stability of the extended Kalman filter when the states are constrained. IEEE T Automat Contr 2008; 53: 2707–2711.

[24] Khalil HK. Nonlinear Systems. 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2000.

[25] Reif K, Unbehauen R. The extended Kalman filter as an exponential observer for nonlinear systems. IEEE T Signal Proces 1999; 47: 2324–2328.