A New Adaptive Extended Kalman Filter for a Class of Nonlinear Systems

Iyad Hashlamon

Department of Mechanical Engineering, Palestine Polytechnic University
Hebron, Palestine, Email: iyad@ppu.edu

Abstract. This paper proposes a new adaptive extended Kalman filter (AEKF) for a class of nonlinear systems perturbed by noise which is not necessarily additive. The proposed filter is adaptive against the uncertainty in the process and measurement noise covariances. This is accomplished by deriving two recursive updating rules for the noise covariances, these rules are easy to implement and reduce the number of noise parameters that need to be tuned in the extended Kalman filter (EKF). Furthermore, the AEKF updates the noise covariances to enhance filter stability. Most importantly, in the worst case, the AEKF converges to the conventional EKF. The AEKF performance is determined based on the mean square error (MSE) performance measure and the stability is proven. The results illustrate that the proposed AEKF has a dramatic improved performance over the conventional EKF, the estimates are more stable with less noise.

Keywords: Extended Kalman filter, Adaptive extended Kalman filter, Covariance matching, Quaternion.

1. Introduction

The Kalman filtering assumes the availability of the plant dynamic model, the process and the observation noises are white and independent [1]. The extended Kalman filter (EKF) is an extension for the linear Kalman filter and is one of the most famous estimation tools for nonlinear systems. The EKF uses noisy measurements to estimate the states of a dynamic system perturbed by noise [2-11]. However, the estimation process faces a problem related to the noise models. The structure of the EKF is composed of the plant dynamic nonlinear model and the noise stochastic models [12, 13]. The EKF uses the noise statistics to influence the EKF gain that is applied on the filter innovation error and then updates the process information to get the best estimate. Accordingly, the EKF performance, reliability and stability depend on the knowledge of the stochastic models parameters. Further, the EKF performance degrades or may even diverge with uncertain model parameters [14,15]. Therefore, improving the EKF such that it can adapt itself to the uncertainty in the noise statistical parameters and reduce their effects is of significant importance. This explains the interest of the researchers to develop several methods to overcome the noise uncertainty challenge. One method uses the Innovation-based Adaptive Estimation (IAE) method [16]. It assumes that the innovation sequence is white noise. Then, it estimates the process noise covariance matrix and/or the measurement covariance matrix using one of the following techniques: covariance matching, correlation and maximum likelihood techniques [17-25]. However, each of these techniques has its drawbacks, the first two techniques require large window of data which makes them impractical. The correlation technique has biased estimated covariances [26]. Maximum likelihood technique requires heavy computations and they can be implemented off-line. Another method is model based method called Multiple Model Adaptive Estimation (MMAE). It assumes the availability of the correct model among bank of different models. Then the probability of each model is computed using the measurements. At the end, the output of the highest probability model is considered [27]. However, it is hard to have the correct model for the uncertain dynamic systems [28]. An optimization-based Adaptive
Estimation (OAE) [29] is developed to handle the model uncertainty based on past data. Scaling the error state covariance matrix by a factor is reported too [18,30]. The factor calculation is either empirical or based on the filter innovations.

The main contribution of this paper is to obtain an AEKF in order to overcome the above mentioned drawbacks by adopting the recursive estimation approach. The author applied the idea of adaptive Kalman filter for linear systems [31] to nonlinear systems with a non-additive noise. By definition of recursive estimation and update, the AEKF will be able to adapt itself to the biased initial covariances, to increase the estimation accuracy, and to enhance the filter stability. In this paper, two recursive updating rules for the noise covariances are obtained. These rules are easy to implement and initialize. Each rule has a correction covariance error term calculated at each sample time by utilizing the advantage of the availability of the most recent measurements and innovations along with the available information about the state covariance error. The filter stability is proven.

The rest of the paper is organized as follows: section 2 introduces the conventional EKF and defines the problem. The adaptive EKF is derived in section 3. Section 4 explores the stability proof. The numerical example is presented in section 5 and the paper is concluded in section 6.

2. Conventional EKF and Problem Definition

Consider the discrete-time nonlinear state space model

\[ x_k = f(x_{k-1}, u_{k-1}) + \Gamma(x_{k-1})v_{k-1}, \]

\[ y_k = Hx_k + v_k, \]  

where \( x \in \mathbb{R}^n \) is the state vector, \( y \in \mathbb{R}^d \) is the measurement vector, \( u \) is the system input, \( k \) is the time index. \( u_{k-1} \in \mathbb{R}^m \) and \( v_k \in \mathbb{R}^d \) are the Gaussian process and measurement noises respectively. \( H \) is the output matrix. \( \Gamma \in \mathbb{R}^{n \times m} \) maps the noise to the states space. 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 expectations \( E(v_k) = E(v_k) = E(v v^T) = 0 \) and covariances \( Q = E(v v^T) \) and \( R = E(v v^T) \), where \( E(\cdot) \) stands for the expectation of \( (\cdot) \).

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

**Assumption 3:** The noise covariances are considered to be constant. Then for the given system in Eq. (1), the conventional EKF algorithm is composed of the prediction step

\[ \hat{x}_k = f(\hat{x}_{k-1}, u_{k-1}), \]  

\[ P_k = A_{k-1}P_{k-1}A_{k-1}^T + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^T, \]  

and the measurement update step

\[ S_k = HP_kH^T + R_k, \]  

\[ K_k = P_kH^TS_k^{-1}, \]  

\[ e_k = z_k - H\hat{x}_k, \]  

\[ \hat{x}_k = \hat{x}_k + K_ke_k, \]  

\[ P_k = (I - K_kH_k)P_k, \]  

where \( A_{k-1} = \frac{\partial f}{\partial x}|_{x_{k-1}, u_{k-1}} \).

In Eqs. (2)-(8) the following notation is employed. \((\cdot)^\circ\) and \((\cdot)^\bullet\) stand for the prior and posterior estimates, respectively. \( K \) is the Kalman gain, \( I \) is the identity matrix and \( P \) is the estimation error covariance matrix, is the estimated state and \( z \) is the measurement vector with the same dimension as \( y \).

Remark:
Kalman gain in (5) can be rewritten as \( K_k = P_kH^T R_k^{-1} \).

The noise covariances have significant importance on the estimation performance. Very small values or large values of \( Q \) with respect to the true value will result in a biased estimated \( \hat{x} \) or an oscillated estimated \( \hat{x} \) respectively [29]. Further,
A New Adaptive Extended Kalman Filter for a Class of Nonlinear Systems


remembering that the EKF performance degrades or may even diverge with uncertain model parameters [14,15]. Therefore, it is required to develop AEKF to overcome the uncertainty in the noise covariances.

3. Adaptive EKF

The values of $Q$ and $R$ have an important effect on the EKF performance. Too small or large values of these covariances with respect to the true value results in estimation degradation. Here two recursive updating rules $R_1$ and $R_2$ are developed to update both $Q$ and $R$ to form the AEKF. The AEKF are able to adapt itself to the noise covariance uncertainty in order to achieve better performance.

For the given system in Eq. (1), consider that the assumptions 1 to 3 hold. Then for a given initial value matrices $R_0$, $Q_0$ and selected positive constants $N_R$ and $N_Q$, there are noise covariance errors $Q\Delta$ and $R\Delta$ to update the observation and the process covariance matrices recursively as in (13) and (18) respectively. The AEKF is the same conventional filter with the following rules.

Initial values $\bar{\alpha}_0, \bar{\sigma}_0, \bar{P}_0 > 0, Q_0 > 0, R_0 > 0$  \hspace{1cm} (9)

$\alpha_k = \frac{N_R - 1}{N_R} $ \hspace{1cm} (10)

$\bar{\sigma}_k = \alpha_k \bar{\sigma}_{k-1} + \frac{1}{N_R} e_k $ \hspace{1cm} (11)

$\Delta R_k = \frac{1}{N_R - 1} (e_k - \bar{\sigma}_k) (e_k - \bar{\sigma}_k)^T - \frac{1}{N_R} H_k P_k H_k^T $ \hspace{1cm} (12)

$R_k = \text{diag}(\alpha_k R_{k-1} \Delta R_k) $ \hspace{1cm} (13)

$\alpha_k = \frac{N_Q - 1}{N_Q} $ \hspace{1cm} (14)

$\tilde{\omega}_k = \hat{x}_k - \bar{x}_k$ \hspace{1cm} (15)

$\tilde{\alpha}_k = \alpha_k \tilde{\alpha}_{k-1} + \frac{1}{N_Q} \tilde{\omega}_k$ \hspace{1cm} (16)

$\Delta Q_k = \left\{ \frac{1}{N_Q - 1} \Gamma^\perp_{k-1} (\omega_k - \bar{\omega}_k) (\omega_k - \bar{\omega}_k)^T \Gamma^\perp_{k-1} \right\} + \frac{1}{N_Q} \left[ \Gamma^\perp_{k-1} P_{k-1} \Gamma^\perp_{k-1}^T \right] - \left[ \Gamma^\perp_{k-1} A P_{k-1} A^T \Gamma^\perp_{k-1}^T \right] \right\}$ \hspace{1cm} (17)

$Q_k = \text{diag}(\alpha_k Q_{k-1} \Delta Q_k)$ \hspace{1cm} (18)

Remarks:
- The AEKF converges to the conventional EKF if the selected values of $N_R$ and $N_Q \rightarrow \infty$.
- The update rules keep the noise covariance matrices $Q$ and $R$ positive definite for all $k$.

3.1. The process noise covariance matrix proof

For this proof, we need to know the true value of the states which is not the case. Therefore, the approach of the best known states is acquired. The predicted state error covariance matrix $P^\perp$ is

$P^\perp_k = A P_{k-1} A^T + \Gamma^\perp_{k-1} Q_{k-1} \Gamma^\perp_{k-1}$ \hspace{1cm} (19)

where $Q$ is the assumed value of the process covariance and considered to be constant. Assume that there is uncertainty in $Q$ and it is called $\Delta Q$, then Eq. (19) can be written as

$\Gamma^\perp_{k-1} P^\perp_k \Gamma^\perp_{k-1} = \Gamma^\perp_{k-1} \tilde{P}_k A \tilde{T}_k A^T \Gamma^\perp_{k-1} + Q - \Delta Q$ \hspace{1cm} (20)

where $(\cdot)^\perp$ is the pseudo inverse. $\Delta Q$ requires the true values of the states which are not known, however, we can use an estimate as

$\tilde{\omega} = \hat{x} - \bar{x}$ \hspace{1cm} (21)
For a recorded number \( N_Q \) of measurements of the estimated states, the mean and sample covariance respectively are:

\[
\bar{\omega} = \frac{1}{N_Q} \sum_{i=1}^{N_Q} \omega_i
\]

\[
\Delta Q = \frac{1}{N_Q - 1} \sum_{i=1}^{N_Q} (\hat{\omega}_i - \bar{\omega})(\hat{\omega}_i - \bar{\omega})^T
\]

Taking the mean of Eq. (20) with Eq. (23) and after mathematical manipulation, we can obtain

\[
Q = \frac{1}{N_Q} \left( \sum_{i=1}^{N_Q} \left( \left( \Gamma_{k-1}^* P_k \Gamma_{k-1}^{*T} \right)_i \right) - \left( \Gamma_{k-1}^* A P_{k-1} A^T \Gamma_{k-1}^{*T} \right)_i \right) + \frac{1}{N_Q - 1} \sum_{i=1}^{N_Q} (\hat{\omega}_i - \bar{\omega})(\hat{\omega}_i - \bar{\omega})^T \Gamma_{k-1}^{*T}
\]

(24)

Since they are renewed each time, then the covariance has a strong relation with the previous covariance. To find it, the samples are divided into a group of all samples from \( i = k - N_Q \) up to \( i = k - 1 \) and a second group contains only the most recent sample arrived at the time instant \( k \). Then after some mathematical manipulation, with large \( N_Q \) to approximate the sample covariance of the first group with \( \left( Q_k - 1 \right)^2 \times \sum_{i=k-N_Q}^{k-1} (\hat{\omega}_k - \bar{\omega})(\hat{\omega}_k - \bar{\omega})^T \), Eq. (24) can be expanded as

\[
Q_k = \frac{N_Q - 1}{N_Q} Q_{k-1} + \Delta Q_k,
\]

where

\[
Q_{k-1} \approx \frac{1}{N_Q - 2} \sum_{i=k-N_Q}^{k-2} \Gamma_{k-1,i} (\hat{\omega}_i - \bar{\omega})(\hat{\omega}_i - \bar{\omega})^T \Gamma_{k-1,i}^* + \frac{1}{N_Q - 1} \sum_{i=k-N_Q}^{k-1} \left( \left( \Gamma_{k-1,i} P_k \Gamma_{k-1,i}^{*T} \right)_i - \left( \Gamma_{k-1,i} A P_{k-1} A^T \Gamma_{k-1,i}^{*T} \right)_i \right)
\]

(26)

and

\[
\Delta Q_k = \frac{1}{N_Q - 1} \Gamma_{k-1} (\hat{\omega}_k - \bar{\omega})(\hat{\omega}_k - \bar{\omega})^T \Gamma_{k-1}^* + \frac{1}{N_Q} \left( \left( \Gamma_{k-1} P_k \Gamma_{k-1}^{*T} \right)_k - \left( \Gamma_{k-1} A P_{k-1} A^T \Gamma_{k-1}^{*T} \right)_k \right)
\]

(27)

The same method is used to compute \( \bar{\omega}_k \) as

\[
\bar{\omega}_k = \frac{1}{N_Q} \sum_{i=k-N_Q}^{k} \hat{\omega}_i = \frac{1}{N_Q} \sum_{i=k-N_Q}^{k-1} \hat{\omega}_i + \frac{1}{N_Q} \hat{\omega}_k,
\]

(28)

this yields

\[
\bar{\omega}_k = \frac{N_Q - 1}{N_Q} \bar{\omega}_{k-1} + \frac{1}{N_Q} \hat{\omega}_k.
\]

(29)

3.2. The observation covariance matrix proof

Using the same approximation and starting from the innovation error Eq. (6) and the covariance Eq. (4), it will end up with Eq. (12).

4. Stability Enhancement Proof

The exponential behavior of the observer

\[
\dot{x}_k = f \left( \dot{x}_{k-1}, u_{k-1} \right)
\]

\[
\dot{x}_k = \dot{x}_k + K_s H \left( x_k - \dot{x}_k \right)
\]

is determined based on the exponential convergence of the dynamic error \( e_k = x_k - \dot{x}_k \) between the true state \( x \) and the estimated state \( \dot{x} \). Therefore, to analyze the exponential behavior, we first write the Taylor expansion for the observer and the given continuous system as with high order terms \( \partial_x \) and \( \partial_2 \) as,
A New Adaptive Extended Kalman Filter for a Class of Nonlinear Systems


Then after mathematical manipulation, this error \( \varepsilon \) can be expressed as

\[
\varepsilon_k = (A_{k-1} - K_kHA_{k-1})\varepsilon_{k-1} + \varphi_k,
\]

where

\[
\varphi_k = (I - K_kH)[\partial_1(x_{k-1}, u_{k-1}) - \partial_1(\hat{x}_{k-1}, u_{k-1})].
\]

The exponential stability is proven here based on Lyapunov function theory and follows the approach as in [32, 33]. The following definitions and lemmas are employed for the sake of completeness and proof.

**Definition 1:** The origin of the difference Eq. (34) is exponentially stable equilibrium point if there is a continuous differentiable positive definite function \( V(\varepsilon_k) \) such that [34]:

\[
\Delta V(\varepsilon_k) \leq -c_3\|\varepsilon_k\|^2,
\]

for positive constants \( c_1, c_2 \) and \( c_3 \) with \( \Delta V \) as the rate of change of \( V \) and defined by

\[
\Delta V = V(\varepsilon_k) - V(\varepsilon_{k-1}) .
\]

For sake of completeness, the exponential stability for discrete time systems is defined by the inequality \( \|\varepsilon_k\| \leq \beta \|\varepsilon_0\| Y^k \) for all \( k \geq 0 \) with \( \beta > 0 \) and \( 0 < Y < 1 \).

Satisfying the exponential stability for the origin of Eq. (34) implies that the observer in Eq. (31) is an exponential observer.

**Definition 2:** if \( A \) and \( \Gamma Q \Gamma^T \) are invertible matrices, and for the positive definite matrices \( P_k \) and \( P_{k-1} \), then

\[
P_k^{-1} \leq (I - K_kH)^{-1} A^{-T} \left[ P_{k-1}^{-1} - P_{k-1}^{-1} \Gamma_{k-1}Q_{k-1}^{-1} \Gamma_{k-1}^{-T} A \right] A^{-1} (I - K_kH)^{-1}
\]

for positive constants \( c_1, c_2 \) and \( c_3 \) with \( \Delta V \) as the rate of change of \( V \) and defined by

\[
\Delta V = V(\varepsilon_k) - V(\varepsilon_{k-1}) .
\]

\[
\Delta V = V(\varepsilon_k) - V(\varepsilon_{k-1}) .
\]

Proof: rewriting Eq. (8) as in [12]:

\[
P_k = (I - K_kH)P_k^{-1} (I - K_kH)^T + K_kR_kK_k^T ,
\]

then we have

\[
P_k \geq (I - K_kH)P_k^{-1} (I - K_kH)^T .
\]

Inverting Eq. (40) results in

\[
P_k^{-1} \leq (I - K_kH)^{-T} P_k^{-1} (I - K_kH)^{-1} ,
\]

The expression of \( (P_k^*)^{-1} \) is obtained by rearranging Eq. (3) as

\[
P_k^* = A (P_{k-1} + A^{-1} \Gamma_{k-1}Q_{k-1}^{-1} \Gamma_{k-1}^T A^{-T}) A^T ,
\]

and inverting Eq. (42) yields to

\[
(P_k^*)^{-1} = A^{-T} \left[ P_{k-1}^{-1} - P_{k-1}^{-1} \Gamma_{k-1}Q_{k-1}^{-1} \Gamma_{k-1}^T A \right] A^{-1} .
\]

As a result from Eq. (43), Eq. (41) is expressed by

\[
P_k^{-1} \leq (I - K_kH)^{-T} A^{-T} \left[ P_{k-1}^{-1} - P_{k-1}^{-1} \Gamma_{k-1}Q_{k-1}^{-1} \Gamma_{k-1}^T A \right] A^{-1} (I - K_kH)^{-1} ,
\]

which completes the proof \( \square \).
Lemma 1: Consider the real and bounded system states $x_k$, then the terms $\partial_1, \partial_2$ and the positive real numbers $\kappa_\sigma, \sigma, a > 0$ exist such that $\| A \| \leq a$ holds and

$$\| \partial_1 (x, u) - \partial_1 (\hat{x}, u) \| \leq \kappa_\sigma \| x - \hat{x} \|,$$  \hspace{1cm} (45)

holds for $\| x - \hat{x} \| \leq \sigma_\sigma < \frac{1}{2} \sigma$.

The stability Theorem: the given system in Eq. (30) with the proposed AEKF is exponentially stable if the following assumptions hold

A1) there are positive real numbers $a, p, p, \kappa_\sigma, \sigma_\sigma > 0$ such that $p \leq p_{k-1} \leq p$ and $\| A \| \leq a$ for every time instant $k$.

Further the inequality Eq. (45) holds for $\| x - \hat{x} \| \leq \sigma_\sigma < \frac{1}{2} \sigma$.

A2) the matrices $Q$ and $R$ are positive definite due to the updating rules for all $k$. The minimum eigenvalues of $\Gamma Q \Gamma^T$ is $q > 0$.

A3) the matrix $A_{k-1}$ is nonsingular.

Proof: the proof follows the approach as in [32]. Consider the positive definite Lyapunov function

$$V (x_{k-1}) = e_{k-1}^T P_{k-1}^{-1} e_{k-1},$$  \hspace{1cm} (46)

with $V (0) = 0$, Eq. (46) and $A_1$ imply that

$$\frac{1}{p} \| e_{k-1} \| \leq V (x_{k-1}) \leq \frac{1}{p} \| e_{k-1} \|^2.$$  \hspace{1cm} (47)

Then for $V (x_k)$ we obtain

$$V (x_k) = e_{k-1}^T P_{k-1}^{-1} e_{k-1}.$$  \hspace{1cm} (48)

Substituting Eq. (34) into Eq. (48) we get

$$V (x_k) = \left( (A_{k-1} - K \Sigma A_{k-1}) e_{k-1} + \varphi_k \right)^T P_{k-1}^{-1} \left( (A_{k-1} - K \Sigma A_{k-1}) e_{k-1} + \varphi_k \right).$$  \hspace{1cm} (49)

The assumption $A_1$ implies that $P_{k-1}$ is nonsingular along with $A_2$ and $A_3$ fulfill the requirement of Definition 2, then by using Eq. (38) together with Eq. (49) yield

$$V (x_k) \leq \left( e_{k-1} \right)^T \left( P_{k-1}^{-1} - P_{k-1}^{-1} (P_{k-1}^{-1} + A^T (R_{k-1} Q_k R_{k-1}^T)^{-1} A)^{-1} P_{k-1}^{-1} \right) \left( (A_{k-1} - K \Sigma A_{k-1}) e_{k-1} + \varphi_k \right)^T e_{k-1} + \left( (A_{k-1} - K \Sigma A_{k-1}) e_{k-1} + \varphi_k \right)^T P_{k-1}^{-1} \varphi_k.$$  \hspace{1cm} (50)

Now applying Lemma 1 along with $A_1$ on Eq. (50) we get

$$\Delta V \leq \frac{1}{p^2} \left( \frac{1}{p^2} + \frac{a^2}{q} \right) \| e_{k-1} \|^2 + 2 \frac{\kappa_\sigma}{a} \| e_{k-1} \|^2 \left( \frac{1}{p} - \frac{1}{p^2} \right) \| e_{k-1} \| + \frac{\kappa_\sigma^2}{a^2} \| e_{k-1} \|^3 + \frac{1}{p^2} \left( \frac{1}{p^2} + \frac{a^2}{q} \right) \| e_{k-1} \| \| e_{k-1} \|^2 \| e_{k-1} \| + \frac{1}{a^2} \sigma_\sigma \| e_{k-1} \|,$$  \hspace{1cm} (51)

with $\| x - \hat{x} \| \leq \sigma_\sigma$. Let $\Delta = \left( \frac{1}{p^2} \left( \frac{1}{p^2} + \frac{a^2}{q} \right) \right)^{-1}$, note that $0 < \Delta < 1 / p$, and define $\eta$ by

$$\eta = 2 \frac{\kappa_\sigma}{a} \left( \frac{1}{p} - \frac{1}{p^2} \right) + \kappa_\sigma^2 \left( \frac{1}{p^2} \right) + \frac{1}{a^2} \sigma_\sigma.$$  \hspace{1cm} (52)
where $\eta > 0$, then Eq. (51) can be reduced to
\[ \Delta V \leq -\eta \left( \frac{\Delta}{\eta} - \| \epsilon_{k+1} \| \right)^2 , \tag{53} \]
which holds for $\| x - \hat{x} \| \leq \sigma, \sigma < 2$. Let $\sigma = \Delta / \eta$, it follows that
\[ \Delta V \leq -\frac{\Delta}{2} \| \epsilon_{k+1} \| , \tag{54} \]
holds for $\| \epsilon_{k-1} \| \leq \sigma_g$. □
which satisfies Eq. (36), and thus the origin of Eq. (34) is exponentially stable. In terms of states and performance, using Eq. (51) and Eq. (47) we can write
\[ V(\epsilon_k) \leq \left( 1 - \frac{\Delta}{2} P \right) V(\epsilon_0) , \tag{55} \]
and then,
\[ \| \epsilon_k \| \leq \frac{P}{\sqrt{P}} \sqrt{1 - \frac{\Delta}{2}} \| \epsilon_0 \|, k \geq 0 \tag{56} \]
Recalling definition 1, we have
\[ \beta = \frac{P}{\sqrt{P}} > 0 , \tag{57} \]
and
\[ Y = \sqrt{1 - \frac{\Delta}{2}}, 0 < Y < 1 . \tag{58} \]

5. Numerical Example

The performances of the proposed AEKF is shown using the nonlinear function which uses the quaternion representation.

5.1. Model derivation

The quaternion vector has an important role in representing rotations of a rigid body with respect to a reference frame [35]. The quaternion time derivative forms a nonlinear dynamical model. However, this model has bias and contaminated noise. Therefore, the quaternion estimation is generally based on the EKF theory [36 to 40]. The quaternion vector $\mathbf{q}$ consists of four elements as $\mathbf{q} = [q_0, q_1, q_2, q_3]^T \in \mathbb{R}^4$. These elements are divided into two parts, scalar part $q_0 \in \mathbb{R}$ and vector part $\mathbf{n} \equiv [q_1, q_2, q_3] \in \mathbb{R}^3$. The normalized vector $\mathbf{q}$ is generally used with the readings of the gyroscope due to the direct relation between the quaternion time derivative $\dot{\mathbf{q}}$ and the gyro-meter angular velocity $\Omega \in \mathbb{R}^3$ as
\[ \mathbf{q} = \mathbf{q} \otimes \Omega \tag{59} \]
where $\otimes$ is the quaternion multiplication.
Equation (59) is nonlinear and the gyro-meter angular velocity has bias $\mathbf{b} \in \mathbb{R}^3$ and contaminated white zero mean noise $\mathbf{v} \in \mathbb{R}^3$ [41] as stated in,
\[ \Omega = \omega + \mathbf{b} + \mathbf{v} \tag{60} \]
where $\omega \in \mathbb{R}^3$ is the true angular velocity without bias or noise. The bias is modeled as an integrated white noise $\mathbf{v}_b \in \mathbb{R}^3$ as
\[ \mathbf{b}_k = \mathbf{b}_{k-1} + \mathbf{v}_{b,k-1} \tag{61} \]
Due to the bias and noise, the quaternion is estimated by employing the EKF. Defining the state vector as
and taking the discrete form of Eq. (59) along with Eq. (60) and Eq. (61), then the model Eq. (1) is obtained with

$$f(x_{k-1}, u_{k-1}) = \left[ I_4 + \frac{1}{2} UT(U_{k-1} - b_{k-1}) \right] q_{k-1}$$

(62)

$$u_{k-1} = \begin{bmatrix} v_{k-1} \\ v_{b,k-1} \end{bmatrix}$$

(63)

where

$$U(\chi) = \begin{bmatrix} 0 & -x_0 & -x_1 & -x_2 \\ x_0 & 0 & x_3 & -x_2 \\ x_1 & -x_3 & 0 & x_0 \\ x_2 & x_1 & -x_0 & 0 \end{bmatrix}$$

(64)

$$\Gamma_{k-1} = \frac{1}{2} T U(q_{k-1}) \begin{bmatrix} 0_{4 \times 3} \\ 0_{b_{k-1}} & I_5 \end{bmatrix}$$

(65)

$$\bar{U}(\chi) = \begin{bmatrix} -x_1 & -x_2 & -x_3 \\ x_0 & -x_3 & x_2 \\ x_3 & x_0 & -x_1 \\ -x_2 & x_1 & x_0 \end{bmatrix}$$

(66)

and $I_n$ is $n \times n$ identity matrix.

This model is used to study the performance of both the EKF and the proposed AEKF. The output of the EKF and the AEKF is $q$. The quaternion normalization constraint is not preserved by the EKF [42]. To overcome this problem, normalization is applied on the post-estimated quaternion to maintain its unity norm out of the structure of the EKF [38]. More structural methods were used by enforcing constraints in Kalman filtering [43-46]. Here the quaternion normalization is done as in [47].

5.2. Simulation environment and parameters

The simulation platform is MATLAB, both the EKF and the AEKF are tested and compared. The bias and process noises with the covariance $Q_{\text{inv}}$ are added to the angular velocity measured from the gyro-meter, this forms $\Omega$ which is the input $u$ for Eq. (1) and Eq. (2). The measurement noise with covariance $R_{\text{inv}}$ is added to the measured quaternion to form the measurement vector $z$ in Eq. (6). The Gaussian noise is generated by the MATLAB simulink Gaussian noise generator. Both filters have the same initial values $x_0$ and $P_0$. The model equations (62)-(66) are used. The true bias values are chosen to be time $t$ dependent as in Eq. (67).

$$b = \begin{cases} \begin{bmatrix} 0.5 & 0.1 & 0 \end{bmatrix}^T & t \leq 3 \\ \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}^T & 8 < t \leq 20 \\ \begin{bmatrix} 0 & 0 & 2 \end{bmatrix}^T & \text{else} \end{cases}$$

(67)

The simulation parameters, initializations and the corresponding numbers of $N_0$ and $N_r$ are listed in Table 1. Note that the values of $N_0$ and $N_r$ are different from each other since they don’t have to be the same in practice. This enhance more flexibility to the rules, the values of $N_0$ and $N_r$ are user defined based on the system noise characteristics. Without loss of generality, for a noisy system, select big $N_0$ and $N_r$. In the same context, small values of $N_0$ and $N_r$ are for less noisy systems. In this simulation, since the true observation noise is much smaller than the process noise, the value of $N_r$ is selected to be much smaller than $N_0$ as tabulated in Table 1. The following notations are employed: $Q_{\text{inv}}$ and $Q_{\text{true}}$ indicate that the considered process covariance noise in EKF is either smaller or larger than the true process covariance noise $Q_{\text{true}}$ respectively. The same definition goes for $R_{\text{inv}}$ and $R_{\text{true}}$. The notations $Q_{\text{inv}}$ and $R_{\text{inv}}$ refer to the
initial values of the covariances used in the AEKF. $I_n$ is an $n \times n$ identity matrix.

### Table 1. Initialization and simulation parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>0.01 sec</td>
</tr>
<tr>
<td>$R_{true}$</td>
<td>$10^{-6} I_4$</td>
</tr>
<tr>
<td>$Q_{true}$</td>
<td>$10^{-2} I_6$</td>
</tr>
<tr>
<td>$P_0$</td>
<td>$10 I_7$</td>
</tr>
<tr>
<td>$q_0$</td>
<td>$[0.5 \ 0.5 \ 0.5 \ 0.5]^T$</td>
</tr>
<tr>
<td>$b_0$</td>
<td>$[0 \ 0 \ 0]^T$</td>
</tr>
<tr>
<td>$N_0$</td>
<td>$10^4$</td>
</tr>
<tr>
<td>$N_0$</td>
<td>$3 \times 10^5$</td>
</tr>
<tr>
<td>$\sigma_0$</td>
<td>$0_{i \times 1}$</td>
</tr>
<tr>
<td>$\tau_0$</td>
<td>$0_{i \times 1}$</td>
</tr>
<tr>
<td>$R_{small}$</td>
<td>$10^{-10} I_4$</td>
</tr>
<tr>
<td>$R_{big}$</td>
<td>$10^{-2} I_6$</td>
</tr>
<tr>
<td>$Q_{small}$</td>
<td>$10^{-2} I_6$</td>
</tr>
<tr>
<td>$Q_{big}$</td>
<td>$I_6$</td>
</tr>
</tbody>
</table>

### 5.3. Results and discussions

Noise covariances have several scenarios, among them is that the used noise covariances are smaller or larger than the true covariances. For example, the values $Q_{small}$, $R_{small}$, $Q_{big}$ and $R_{big}$ in Table 1 are used with the EKF, the estimation performance is shown in Fig. 1. The performance show that the AEKF estimation has less noise than the EKF as clear in Fig. 1 (a,b) for the first bias element $b_1$, even the noise is decreasing with time as in Fig. 1(b). Fig. 1 (c,d) shows the MSE for the bias $b_2$ and $b_3$, respectively. The MSE for EKF with the values $Q_{small}$, $R_{small}$ is much larger than it for AEKF under the same conditions, i.e. the initial values of the AEKF covariances are the values $Q_{small}$, $R_{small}$. For $Q_{big}$ and $R_{big}$, at the beginning both filters are almost the same. However, the AEKF MSE decreases with time more than the EKF MSE. Moreover, whether the initial covariances are small or big they converge to the same result unlike the EKF. So we can claim that the AEKF adapts itself to biased initializations and has better performances than the EKF. Further, this AEKF is recursive and requires only the previous step data and thus overcomes the large window of data for other methods. And since the proposed AEKF adopts the recursively idea of the traditional EKF, it doesn't make any iterations inside the filter algorithm. Therefore it doesn't require heavy computations and can be implemented online the same as the online EKF.

![Fig. 1.](image-url)
Another scenario is when one of the covariances is small while the other one is big. For this case, the estimation error is shown in Fig. 2 with $Q_{\text{small}}$ and $R_{\text{big}}$. As clear the AEKF is still much better than the EKF. Furthermore, some values for the covariances may slow down the filter response or even cause divergence, for the selected values of $Q = 10^{-12}$ and $R = 10^{-2}$, the estimated bias using the EKF diverges as depicted in Fig. 3. However this problem is solved in the AEKF which keeps the stability of the filter and forces it to converge. This is because the AEKF gain is changing based on the estimation performance. This is clear since it depends on both the innovation error $\epsilon$ and the state error $\hat{\omega}$ respectively. In terms of the states, when $q \rightarrow 0$ then $\hat{\gamma} \rightarrow 1$ which decreases the convergence speed as in Eq. (58). This case is avoided in the AEKF, when the error increases, the value of $q$ increases too. As a result the value of $\hat{\gamma}$ decreases and hence the convergence speed increases as in Eq. (58). Thus we can claim that the proposed AEKF has better stability and convergence performances than the EKF.

![Fig. 2](image1)

**Fig. 2.** (a) The estimated bias $\hat{b}_1$ with filters and covariance’s as in the legend, (b) the estimated bias $\hat{b}_1$ with filters and covariance’s as in the legend (c) the MSE for $\hat{b}_2$ estimation with filters and covariance’s as in the legend, (d) the MSE for $\hat{b}_3$ estimation with filters and covariance’s as in the legend

![Fig. 3](image2)

**Fig. 3.** The stability enhancement of the AEKF

### 6. Conclusion

A new AEKF for a class of nonlinear systems with uncertain noise covariances is proposed. This AEKF can adjust itself recursively to achieve better performance for biased covariances. It relates the filter gain to the innovation and state errors through the noise covariance updating rules, these relations change the filter gain for better tracking and performance. Furthermore, its tuning parameters are less than the EKF, instead of tuning all of the diagonal elements of the noise covariance matrix, they can be initialized and then tuned using $N_\epsilon$ and $N_\xi$ only. The results show the dramatic improvements in the AEKF response compared with the conventional EKF under the same conditions.
Conflict of Interest

The author declared no potential conflicts of interest with respect to the research, authorship and publication of this article.

Funding

The author received no financial support for the research, authorship and publication of this article.

References


[27] X. Wang, Vehicle health monitoring system using multiple-model adaptive estimation, MSc Thesis, Electrical
Engineering, University of Hawaii at Manoa, Manoa, 2003.


[41] D. Roetenberg, "Inertial and magnetic sensing of human motion," PhD, University of Twente, Enschede, NL, 2006.


