Multiplicative extended Kalman filter ignoring initial conditions for attitude estimation using vector observations

Lubin Chang

College of Electrical Engineering, Naval University of Engineering, Wuhan, Hubei, P.R. China.
E-mail: changlubin@163.com

Received: 25 May 2021; Accepted: 19 July 2022; First published online: 12 January 2023

Keywords: inertial navigation; lie group; multiplicative extended Kalman filter; attitude estimation

Abstract
In this paper, the well-known multiplicative extended Kalman filter (MEKF) is re-investigated for attitude estimation using vector observations. From the Lie group theory, it is shown that the attitude estimation model is group-affine and its error state model should be trajectory-independent. Moreover, with such a trajectory-independent error state model, the linear Kalman filter is still effective for large initialisation errors. However, the measurement model of the traditional MEKF is dependent on the attitude prediction, which is therefore trajectory-dependent. This is also the main reason why the performance of traditional MEKF is degraded for large initialisation errors. Through substitution of the attitude prediction related term with vector observations in the body frame, a trajectory-independent measurement model is derived for MEKF. Meanwhile, MEKFs with reference attitude error definition and with global state formulating on special Euclidean group have also been studied, with the main focus on derivation of the trajectory-independent measurement models. Extensive Monte Carlo simulations of spacecraft attitude estimation implementations demonstrate that the performance of MEKFs can be much improved with trajectory-independent measurement models.

1. Introduction

Extended Kalman filter (EKF) has been widely used for attitude estimation using vector observations (Lefferts et al., 1982; Crassidis and Markley, 2003; Markley, 2003; Crassidis et al., 2007). With quaternion as the attitude parameterisation, the resultant EKF is known as multiplicative extended Kalman filter (MEKF) (Lefferts et al., 1982; Markley, 2003). In MEKF, the quaternion is used for the global attitude propagation and a local unconstrained three-dimensional parameterisation is used for local attitude error estimation in filtering recursion. The MEKF has been widely approved as the workhorse of real-time attitude estimation using vector observations (Crassidis et al., 2007). However, the MEKF may become very difficult to tune and even prone to divergence when it has not been well initialised. In order to relax restrictions on valid initialisation region, some advanced filtering methods have been proposed (Crassidis and Markley, 2003; Cheng and Crassidis, 2010). These advanced filtering methods can indeed outperform MEKF for large initialisation errors, but at the cost of increasing computational burden. In Andrle and Crassidis (2015), a new frame-consistent EKF, termed geometric EKF (GEKF), is proposed for spacecraft attitude estimation. In GEKF, an attitude error related gyroscope bias error is defined based on common coordinates perspective. The GEKF can outperform MEKF for large initialisation errors with similar computational cost. In recent years, the invariant extended Kalman filter (IEKF) rooted in the matrix Lie group theory has been investigated for attitude or pose estimation in inertial based applications (Barrau and Bonnabel, 2014, 2015, 2017; Barrau, 2015; Wang et al., 2018, 2019; Cui et al., 2021). Barrau and Bonnabel (2014) argued that IEKF exhibits superior convergence...
and robustness properties over MEKF, due to its error state model not depending on the state estimate. Gui and de Ruiter (2018) investigated two quaternion IEKFs for spacecraft attitude estimation based on the invariance theory and found that the MEKF can be viewed as a minor variant of the quaternion left IEKF (QLIEKF). According to Chang (2020), both the GEKF in Andrle and Crassidis (2015) and the quaternion right IEKF (QRIEKF) in Gui and de Ruiter (2018) can be derived by formulating the attitude and gyroscope drift bias together as elements of a special Euclidean group, $\mathbb{SE}(3)$.

It is generally recognised that the performance degradation of MEKF with large initial attitude error owes much to local linearisation. In other words, the local linearised error state model is not applicable for cases with large attitude error. However, such a conclusion may not hold true universally. From the Lie group theory, if the group state dynamic model is group-affine, then the corresponding linear vector error state model will be trajectory-independent. Moreover, the nonlinear group error can be recovered exactly from the linear vector error (Barrau and Bonnabel, 2017; Hartley et al., 2020). In other words, the linear approximation for group-affine will not lose accuracy. In this paper, it is demonstrated that the attitude kinematics model formulated on the special orthogonal group $\mathbb{SO}(3)$ is group-affine and, therefore, the MEKF has the potential to handle large initialisation errors. Based on the aforementioned discussion, with a small modification of the measurement transition matrix of the existing MEKF, its performance can be much improved. Moreover, the MEKF with reference attitude error and MEKF with global state formulated on $\mathbb{SE}(3)$ have also been revisited and investigated from the perspective of group-affine.

The remainder of this paper is organised as follows. Section 2 presents the dynamic model for attitude estimation using vector observations. The group-affine property of the model has also been analysed, which provides the theoretical basis for the probability of handling large initialisation errors with the linear state-space models. In Section 3, the state-space models with both body frame and reference frame attitude errors have been presented. A modified measurement model which is trajectory-independent is proposed for body frame attitude error definition. Meanwhile, through transformation of the innovation, the measurement model can also be transformed to trajectory-independent for reference frame attitude error definition. In Section 4, through formulating the attitude and gyroscope drift bias as a compact element of $\mathbb{SE}(3)$, the corresponding vector error state-space models are presented. In Section 5, the explicit filtering procedures have been presented with different state-space models. The focus of this section is on the global state update procedures corresponding to their respective error definitions. Numerical simulations are given in Section 6 to evaluate the performance of these different state-space models involved in the MEKF, and conclusions are drawn in Section 7.

2. Group-affine analysis of attitude estimation model

2.1. System model

The attitude kinematics model in terms of attitude matrix is given by Crassidis and Junkins (2012)

$$\dot{A}(t) = -[\omega(t) \times] A(t)$$  (1)

where $\omega$ is the angular rate measured by the gyroscopes and governed by

$$\ddot{\omega} = \omega + \beta + \eta_v$$

$$\dot{\beta} = \eta_u$$  (2)

where $\beta$ is the constant gyroscope bias. $\eta_v$ and $\eta_u$ are assumed to be zero-mean, Gaussian white-noise processes.

The measurement model corresponding to the vector observation is given by

$$b_k = A(t_k) r_k + v_k$$  (3)
where $r_k$ is the reference vector observation at time instant $t_k$ and $b_k$ is the observation vector with measurement noise $v_k$. The measurement can be provided by sun sensors, horizon sensor, magnetometers and so on; for more details, refer to Markley and Crassidis (2014). Here only one vector observation is considered and the generalisation to case with multiple vector observations is straightforward.

### 2.2. Group-affine analysis of attitude estimation model

Given a group state model

$$\dot{\chi} = f_u(\chi)$$

(4)

where $\chi \in G$ is the state on group $G$. $u$ are the model inputs. From the Lie group theory, if the group state model in Equation (4) satisfies the following equation, then it is group-affine (Barrau and Bonnabel, 2017; Hartley et al., 2020).

$$f_u(\chi_1 \chi_2) = f_u(\chi_1) \chi_2 + \chi_1 f_u(\chi_2) - \chi_1 f_u(I_{nd}) \chi_2$$

(5)

where $\chi_1, \chi_2 \in G$.

It is known that attitude matrix $A(t)$ is on group $SO(3)$. Next, we will demonstrate that the attitude kinematics in Equation (1) is group-affine. In the following demonstration, the measurement errors of gyroscopes are not considered and the involved variables’ time dependences are omitted temporarily for brevity.

For the model in Equation (1), the group state is $A$ and the input is $\omega$. According to Equation (1), we can get

$$f_u(A_1 A_2) = -(\omega \times) A_1 A_2$$

(6a)

$$f_u(A_1) A_2 = -(\omega \times) A_1 A_2$$

(6b)

$$A_1 f_u(A_2) = -A_1 (\omega \times) A_2$$

(6c)

$$A_1 f_u(I_{3 \times 3}) A_2 = -A_1 (\omega \times) A_2$$

(6d)

It can be easily checked that the attitude kinematics from Equation (1) satisfy Equation (5) and therefore, it is group-affine. According to Lie group theory (Barrau and Bonnabel, 2017; Hartley et al., 2020), the vector error state models corresponding to the group-affine model are trajectory-independent. Moreover, if the observations take the following particular forms, then the linearised observation model will also be trajectory-independent.

Left-Invariant Observation: $y = \chi \mu$

Right-Invariant Observation: $y = \chi^{-1} \mu$

(7)

where $\mu$ is a constant vector. It can be checked that the measurement model in Equation (3) is left-invariant and the linearised observation model corresponding to the left error definition will be trajectory-independent. However, at a certain time instant, both the observation vector $b_k$ and the reference vector $r_k$ can be viewed as known values. Therefore, Equation (3) can be rewritten as

$$r_k = A(t_k)^{-1} b_k - A(t_k)^{-1} v_k$$

(8)

That is to say, the vector observation model can also be viewed as right-invariant. Therefore, the linearised observation model corresponding to the right error definition will also be trajectory-independent.

In the next section, we will present the explicit error state-space models for the attitude estimation problem using vector observations and their trajectory-independence will be clearly shown. This is also the starting point of the modification of the traditional MEKF.
3. Error state-space models on $\mathbb{SO}(3) + \mathbb{R}^3$

3.1. Error state-space model with body frame attitude error

In the traditional MEKF, the attitude error is defined as

$$ A(\delta q) = A(q)A(\hat{q})^{-1} \tag{9} $$

Here, we add the quaternion dependence of each attitude matrix because, in the MEKF framework, the global attitude is propagated in quaternion form. The attitude error in Equation (9) is in the right error form. From the frame transformation perspective, it denotes the attitude error in the body frame. This is a little different from those in Li and Mourikis (2012), Barrau and Bonnabel (2014) and Gui and de Ruiter (2018). The difference is caused by the fact that the global attitude matrix in Equations (1) and (3) denotes attitude transformation from reference frame to body frame while the attitude matrix in Li and Mourikis (2012), Barrau and Bonnabel (2014) and Gui and de Ruiter (2018) denotes attitude transformation from body frame to reference frame.

We term the error in Equation (9) as group error. The corresponding vector form is denoted as $\delta a^b$. With the assumption of small attitude error, their relationship is given by

$$ A(\delta q) = I_{3\times3} - (\delta a^b \times) \tag{10} $$

We denote vector state as $\delta x_{body} = [\delta a^b_T \delta \beta^T]^T$, where $\delta \beta = \beta - \hat{\beta}$. The subscript 'body' means that the involved attitude error is expressed in the body frame. The process model corresponding to Equations (1) and (2) is given by

$$ \delta \dot{x}_{body} = F_{body}(\dot{x}(t), t)\delta x_{body} + G_{body}(t)w(t) \tag{11a} $$

where

$$ F_{body}(\dot{x}(t), t) = \begin{bmatrix} -\hat{\omega} \times & -I_{3\times3} \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix} \tag{11b} $$

$$ G_{body}(t) = \begin{bmatrix} -I_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & I_{3\times3} \end{bmatrix} \tag{11c} $$

In Equation (11a,b), $w(t) = [\eta^T_u(t) \eta^T_v(t)]^T$, $\hat{\omega} = \tilde{\omega} - \hat{\beta}$. It can be seen from Equation (11) that if the gyroscope bias is not considered, the differential model corresponding to $\delta a^b$ is independent of $A(\hat{q})$. This is consistent with the group-affine analysis of Equation (1). Unfortunately, as noted in Barrau and Bonnabel (2017) and Hartley et al. (2020), there is no Lie group that includes the bias terms while also having dynamics that satisfy the group-affine property. However, much of the group-affine structure of attitude can be preserved even with bias augmentation.

With the attitude error definition in Equation (9), the linearised measurement transition matrix corresponding to Equation (3) is given by Crassidis and Junkins (2012)

$$ H_{body,k}(\hat{x}_k) = \begin{bmatrix} A(\hat{q}_k) r_k \times \end{bmatrix} \theta_{3\times3} \tag{12} $$

It appears that Equation (12) is dependent on the global attitude prediction $A(\hat{q}_k)$, which is not consistent with the invariance analysis of the vector observation model. Essentially, this is just the reason why the traditional MEKF does not perform so well with large initialisation errors in previous works. However, according to Equations (3) and (12), a modified measurement transition matrix can be given by

$$ H_{body,k} = \begin{bmatrix} (\hat{b}_k \times) \end{bmatrix} \theta_{3\times3} \tag{13} $$
where \( \tilde{b}_k \) denotes the measured vector observation with noise. It is now shown that the measurement transition matrix in Equation (13) is trajectory-independent. According to the Lie group theory, it is possible to recover the nonlinear attitude error \( A(\delta \dot{q}) \) with the linear state-space model in Equations (11) and (13). Here we have not claimed that the nonlinear attitude error can be recovered ‘exactly’ from its linear vector error state estimate. This is because the gyroscope bias augmentation has made the process model not totally trajectory-independent, as shown in Equation (11).

### 3.2. Error state-space model with reference frame attitude error

For MEKF, the attitude error can also be defined as

\[
A(\delta q) = A(\dot{q})^{-1}A(q)
\]  

(14)

The attitude error in Equation (14) is in the left error form. From the frame transformation perspective, it denotes the attitude error in the reference (inertial) frame. The MEKF with reference attitude error definition was first studied by Gai et al. (1985). The attitude quaternion in Gai et al. (1985) defines the transition matrix in Equation (13) is trajectory-independent. According to the Lie group theory, it is possible to recover the nonlinear attitude error

\[
\delta \dot{x}_{\text{ref}} = F_{\text{ref}}(\dot{x}(t), t)\delta x_{\text{ref}} + G_{\text{ref}}(t)w(t)
\]

(16a)

where

\[
F_{\text{ref}}(\dot{x}(t), t) = \begin{bmatrix} 0_{3 \times 3} & -A(\dot{q})^T \\ 0_{3 \times 3} & 0_{3 \times 3} \end{bmatrix}
\]

(16b)

\[
G_{\text{ref}}(t) = \begin{bmatrix} -A(\dot{q})^T \\ 0_{3 \times 3} \end{bmatrix}
\]

(16c)

It can be seen from Equation (16) that, if the gyroscope bias is not considered, the differential model corresponding to \( \delta \alpha^r \) is independent of \( A(\dot{q}) \). This is consistent with the group-affine analysis of Equation (1). The augmentation of gyroscope bias introduces model dependence of the global attitude estimate \( A(\dot{q}) \).

Similarly to Equation (10), the attitude error in Equation (14) can also be approximated as

\[
A(\delta q) = I_{3 \times 3} - (\delta \alpha^r \times)
\]  

(15)

We denote the vector state as \( \delta x_{\text{ref}} = [\delta \alpha^r \ T \delta \beta]^T \), where \( \delta \beta = \beta - \tilde{\beta} \). The subscript ‘ref’ means that the involved attitude error is expressed in the reference frame. The process model of \( \delta x_{\text{ref}} \) is given by

\[
\delta \dot{x}_{\text{ref}} = F_{\text{ref}}(\dot{x}(t), t)\delta x_{\text{ref}} + G_{\text{ref}}(t)w(t)
\]

(16a)

It is shown that the measurement transition matrix in Equation (17) is dependent on the global attitude prediction \( A(\dot{q}_k) \), which is not consistent with the invariance analysis of the vector observation model. However, Equation (17) can be reorganised as

\[
H_{\text{ref},k}(\dot{x}_k) = [A(\dot{q}_k)A(\dot{q}_k)^{-1}[(A(\dot{q}_k)r_k)\times]A(\dot{q}_k) \ 0_{3 \times 3}]
\]  

(17)

\[
H_{\text{ref},k}(\dot{x}_k) = [A(\dot{q}_k)(A(\dot{q}_k)^{-1}(A(\dot{q}_k)r_k)\times) \ 0_{3 \times 3}]
\]  

(18)

where

\[
F_{\text{ref}}(\dot{x}(t), t) = \begin{bmatrix} 0_{3 \times 3} & -A(\dot{q})^T \\ 0_{3 \times 3} & 0_{3 \times 3} \end{bmatrix}
\]

(16b)

\[
G_{\text{ref}}(t) = \begin{bmatrix} -A(\dot{q})^T \\ 0_{3 \times 3} \end{bmatrix}
\]

(16c)
Although Equation (18) is now still dependent on the global attitude, through operation in the following theorem, it is possible to derive a global attitude independent measurement transition matrix.

Theorem 1 (Barrau, 2015): Applying the same linear function to the linearised measurement transition matrix and innovation term of an EKF before computing the gains does not change the results of the filter.

The explicit proof is presented in Barrau (2015). Based on Theorem 1, the matrix in Equation (18) can be transformed as

$$ H_{ref,k,trans} = A(\hat{q}^-_k)^{-1}H_k(\hat{x}^-_k) = [(r_k \times) \ 0_{3x3}] $$

(19)

Accordingly, the innovation should be transformed as

$$ n_{k,trans} = A(\hat{q}^-_k)^{-1}[b_k - A(\hat{q}^-_k)r_k] = A(\hat{q}^-_k)^{-1}b_k - r_k $$

(20)

The measurement noise covariance should also be transformed as

$$ R_{k,trans} = A(\hat{q}^-_k)^{-1}R_k A(\hat{q}^-_k) $$

(21)

where $R_k$ is the covariance corresponding to $u_k$ in Equation (3).

Now, the measurement transition matrix $H_{ref,k,trans}$ is trajectory-independent. It should be noted that making use of the trajectory-independent matrix of Equation (19) and the transformed innovation of Equation (20) offers the same performance as that making use of Equation (17) or (18) and traditional innovation. The transformation procedures from Equations (18) to (21) are only used to demonstrate that the invariant measurement of Equation (3) can result in a trajectory-independent measurement transition matrix corresponding to reference attitude error definition.

Remark 1: It is known from Section 2 that the attitude kinematics from Equation (1) is group-affine and therefore its error state model is trajectory-independent. It can be seen from Equations (11) and (16) that the attitude error state model is indeed trajectory-independent. Moreover, these trajectory-independent attitude error state models will also be valid when the initial attitude error is large. However, this statement is seemingly contradictory with the approximation in Equations (10) and (15) when deriving the attitude error state model. In the Appendix, we have derived the attitude error state model without any approximation. This can further support that the linear model can be used to accomplish the attitude estimation with large initial attitude error, as will be shown in Section 6.

4. Error state-space models on $\mathbb{SE}(3)$

In Chang (2020), the attitude estimation has also been investigated making use of $\mathbb{SE}(3)$. Through formulating the attitude and gyroscope bias as a compact state on $\mathbb{SE}(3)$, both the GEKF and QRIEKF can be derived. In this section, the explicit error state-space model is directly presented with group state definition on $\mathbb{SE}(3)$.

For attitude estimation, the group state on $\mathbb{SE}(3)$ is defined as

$$ \chi = \begin{bmatrix} A(q) & \beta \\ 0_{1x3} & 1 \end{bmatrix} $$

(22)

It can also be demonstrated that with the group state from Equation (22), its dynamic model does not satisfy Equation (5). Therefore, the resulting error state model will be dependent on the global state estimate, which will be shown in Equations (25) and (28).
4.1. Error state-space model with right error

The right group error corresponds to Equation (9) and it is defined as

$$\delta \chi_{\text{right}} = \dot{\chi} \chi^{-1} = \begin{bmatrix} A(q)A(\hat{q})^{-1} & \beta - A(q)A(\hat{q})^{-1}\hat{\beta} \\ 0_{1 \times 3} & 1 \end{bmatrix}$$

(23)

The vector form corresponding to $\delta \chi_{\text{right}}$ is denoted as $d\chi_{\text{right}} = [d\alpha^b T \ d\beta^b T]^T$. The subscript ‘right’ means that the error state is corresponding to the right group error. $d\alpha^b$ is corresponding to the attitude error $A(q)A(\hat{q})^{-1}$. It is clearly shown that $d\alpha^b$ is just the attitude error in Equation (9). $d\beta^b$ is corresponding to the gyroscope drift bias error and is given by

$$d\beta^b = \beta - A(q)A(\hat{q})^{-1}\hat{\beta} = \beta - [I_{3 \times 3} - (d\alpha \times)]\hat{\beta} = \delta \beta - (\hat{\beta} \times) d\alpha$$

(24)

where we have made the same approximation for $d\alpha^b$ as Equation (10).

The process model of $d\chi_{\text{right}}$ is given by

$$d\chi_{\text{right}} = F_{\text{right}}(\dot{\chi}(t), t)d\chi_{\text{right}} + G_{\text{right}}(\dot{\chi}(t), t)w(t)$$

(25a)

where

$$F_{\text{right}}(\dot{\chi}(t), t) = \begin{bmatrix} -(\hat{\omega} \times) & -I_{3 \times 3} \\ (\hat{\beta} \times)(\hat{\omega} \times) & (\hat{\beta} \times) \end{bmatrix}$$

(25b)

$$G_{\text{right}}(\dot{\chi}(t), t) = \begin{bmatrix} -I_{3 \times 3} \\ 0_{3 \times 3} \end{bmatrix}$$

(25c)

It is clearly shown that Equation (25b) is just the process transition matrix used in GEKF. The defined error state $d\chi_{\text{right}}$ is a little different from that in Chang (2020). This is because we intend to derive a consistent result with GEKF. It is shown in Equation (25) that, with the right error definition, the process transition matrix is now dependent on the global state estimate $\hat{\beta}$. This is also the result of the fact that the dynamic model of the group state in Equation (22) is not group-affine.

Since the attitude error element of $d\chi_{\text{right}}$ has the same definition as that of $\delta \chi_{\text{body}}$, the measurement transition matrix corresponding to $d\chi_{\text{right}}$ is the same Equation (12) or (13). Based on the invariance theory, Equation (13) will be preferable to Equation (12).

4.2. Error state-space model with left error

The left group error corresponds to Equation (14) and is defined as

$$\delta \chi_{\text{left}} = \dot{\chi} \chi^{-1} = \begin{bmatrix} A(\hat{q})^{-1}A(q) & A(\hat{q})^{-1}(\beta - \hat{\beta}) \\ 0_{1 \times 3} & 1 \end{bmatrix}$$

(26)

The vector form corresponding to $\delta \chi_{\text{left}}$ is denoted as $d\chi_{\text{left}} = [d\alpha^r T \ d\beta^r T]^T$. The subscript ‘left’ means that the error state is corresponding to the left group error. $d\alpha^r$ is corresponding to the attitude error $A(\hat{q})^{-1}A(q)$. It is clearly shown that $d\alpha^r$ is just the attitude error in Equation (14). $d\beta^r$ is corresponding to the gyroscope bias error and is given by

$$d\beta^r = A(\hat{q})^{-1}(\beta - \hat{\beta}) = A(\hat{q})^{-1}\delta \beta$$

(27)
From the attitude transformation perspective, Equation (27) can be viewed as transforming the gyroscope bias error from body frame to reference frame. This is also the insight in Gui and de Ruiter (2018), that ‘the group actions can be intuitively viewed as some coordinate transformations’.

The process model of $\mathbf{d}x_{\text{left}}$ is given by

$$
\mathbf{d}x_{\text{left}} = F_{\text{left}}(\dot{\mathbf{x}}(t), t)\mathbf{d}x_{\text{left}} + G_{\text{left}}(\dot{\mathbf{x}}(t), t)\mathbf{w}(t)
$$

(28a)

where

$$
F_{\text{left}}(\dot{\mathbf{x}}(t), t) = \begin{bmatrix} 0_{3\times3} & -I_{3\times3} \\ 0_{3\times3} & [(A(\hat{q})^T\hat{\omega}) \times] \end{bmatrix}
$$

(28b)

$$
G_{\text{left}}(\dot{\mathbf{x}}(t), t) = \begin{bmatrix} -A(\hat{q})^T \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & A(\hat{q})^T \end{bmatrix}
$$

(28c)

It is clearly shown that Equation (28b) is just the process transition matrix used in QRIEKF in Gui and de Ruiter (2018). The defined error state $\mathbf{d}x_{\text{left}}$ is a little different from that in Chang (2020). This is because we intend to derive the consistent result with QRIEKF. In this paper, the process model with the form as in Equation (28) is corresponding to left group state error while such form is corresponding to the right group state error in Gui and de Ruiter (2018). This is also caused by the difference between the meanings of the attitude matrix. Specifically, the attitude matrix in this paper denotes the attitude transformation from reference frame to body frame, while the attitude matrix in Gui and de Ruiter (2018) denotes attitude transformation from body frame to reference frame. It is shown that Equation (28b) is also dependent on the global state estimate. Compared with Equation (16), the dependence has been moved from attitude error model to gyroscope bias error model. Since the attitude error element of $\mathbf{d}x_{\text{left}}$ has the same definition as that of $\mathbf{d}x_{\text{ref}}$, the measurement transition matrix corresponding to $\mathbf{d}x_{\text{left}}$ is the same as Equation (17) or (19). In QRIEKF, the measurement transition matrix in Equation (19) is used. In Gui and de Ruiter (2018), such a trajectory-independent measurement transition matrix is derived from the invariance theory. In this paper, we present another perspective on the derivation of the trajectory-independent measurement transition matrix as shown in Equations (17)–(19).

5. Application aspects

The explicit MEKF procedures are summarised in Table 1. The transition matrices $F(t)$, $G(t)$ and $H_k$ in Table 1 take different forms, as in Equations (11)–(13), (16)–(19), (25) and (28), according to different error state definitions. GEKF and QRIEKF have also been viewed as types of MEKF. The essential difference between the different algorithms lies in the error state definition. Different error state definitions result in different state-space models and different attitude update procedures.

The retraction operation $\bowtie$ denotes how to update the global state using the local error state estimate. Specifically, for error state definition $\mathbf{d}x_{\text{body}}$, $\Delta \hat{x}_k = \begin{bmatrix} \delta \hat{\alpha}_k^{b+T} & \delta \beta_k^{b+T} \end{bmatrix}^T$, the retraction operation is given by

$$
\hat{q}_k^+ = exp_q(\delta \hat{\alpha}_k^{b+}) \odot \hat{q}_k^-
$$

(29a)

$$
\hat{\beta}_k^+ = \hat{\beta}_k^- + \delta \hat{\beta}_k^+
$$

(29b)

where

$$
exp_q(\delta \alpha) = \begin{bmatrix} \frac{\delta \alpha}{||\delta \alpha||} \sin(||\delta \alpha||/2) \\ \cos(||\delta \alpha||/2) \end{bmatrix}
$$

(30)
Table 1. MEKF for attitude estimation using vector observations.

| Initialise | $\mathbf{q}(t_0) = \mathbf{q}_0, \mathbf{\beta}(t_0) = \hat{\mathbf{\beta}}_0, \mathbf{P}(t_0) = \mathbf{P}_0$ |
|------------|---------------------------------------------------------------------------------------------------|
| Gain       | $\mathbf{K}_k = \mathbf{P}_k^{-1} \mathbf{H}_k^T [\mathbf{H}_k \mathbf{P}_k^{-1} \mathbf{H}_k^T + \mathbf{R}_k]^{-1}$ |
| Update     | $\mathbf{P}_k^- = [\mathbf{I}_{6\times 6} - \mathbf{K}_k \mathbf{H}_k] \mathbf{P}_k^-$ |
| Propagation| $\mathbf{\hat{\omega}}(t) = \tilde{\mathbf{\omega}}(t) - \hat{\mathbf{\beta}}(t)$ |
|            | $\mathbf{\dot{q}}(t) = \frac{1}{2} \mathbf{\Xi}(\mathbf{\tilde{q}}(t)) \mathbf{\tilde{\omega}}(t)$ |
|            | $\mathbf{\dot{P}}(t) = \mathbf{F}(t) \mathbf{P}(t) + \mathbf{P}(t) \mathbf{F}^T(t) + \mathbf{G}(t) \mathbf{Q}(t) \mathbf{G}^T(t)$ |

For error state definition $\delta \mathbf{x}_\text{ref}$, $\Lambda \mathbf{\tilde{x}}_k = \left[ \mathbf{\delta} \dot{\mathbf{x}}_k^T \mathbf{\delta} \dot{\mathbf{x}}_k^T \right]^T$, the retraction operation is given by

$$\mathbf{\hat{q}}_k^+ = \mathbf{\hat{q}}_k^- \odot \exp_{\mathbf{q}}(\mathbf{\delta} \dot{\mathbf{x}}_k^+), \quad (31a)$$

$$\mathbf{\hat{\beta}}_k^+ = \mathbf{\hat{\beta}}_k^- + \mathbf{\delta} \dot{\mathbf{x}}_k^+ \odot \mathbf{\delta} \dot{\mathbf{x}}_k^+ \quad (31b)$$

For error state definition $\mathbf{d} \mathbf{x}_\text{right}$, $\Lambda \mathbf{\tilde{x}}_k = \left[ \mathbf{d} \dot{\mathbf{x}}_k^T \mathbf{d} \dot{\mathbf{x}}_k^T \right]^T$, the retraction operation is given by

$$\mathbf{\hat{q}}_k^+ = \exp_{\mathbf{q}}(\mathbf{d} \dot{\mathbf{x}}_k^+) \odot \mathbf{\hat{q}}_k^-, \quad (32a)$$

$$\mathbf{\hat{\beta}}_k^+ = \mathbf{\hat{\beta}}_k^- + \mathbf{\delta} \dot{\mathbf{x}}_k^+ + (\mathbf{\hat{\beta}}_k^- \times \mathbf{\delta} \dot{\mathbf{x}}_k^+) \quad (32b)$$

For error state definition $\mathbf{d} \mathbf{x}_\text{left}$, $\Lambda \mathbf{\tilde{x}}_k = \left[ \mathbf{d} \dot{\mathbf{x}}_k^T \mathbf{d} \dot{\mathbf{x}}_k^T \right]^T$, the retraction operation is given by

$$\mathbf{\hat{q}}_k^+ = \mathbf{\hat{q}}_k^- \odot \exp_{\mathbf{q}}(\mathbf{d} \dot{\mathbf{x}}_k^+), \quad (33a)$$

$$\mathbf{\hat{\beta}}_k^+ = \mathbf{\hat{\beta}}_k^- + \mathbf{\mathbf{\mathbf{A}}}(\mathbf{\hat{\mathbf{\beta}}}_k^-) \mathbf{\mathbf{d}} \dot{\mathbf{x}}_k^+ \quad (33b)$$

It should be noted that for error state $\delta \mathbf{x}_\text{ref}$ and $\mathbf{d} \mathbf{x}_\text{left}$, if making use of the global state-independent transition matrix in Equation (19), the transformed innovation in Equation (20) and measurement noise covariance in Equation (21) should be used accordingly.

It should be noted that, in the aforementioned methods, the covariance update has not been incorporated for the attitude update. One can refer to Mueller et al. (2017) for the corresponding covariance update methods.

6. Numerical example

In this section, MEKF with different state error definitions and different state-space models are evaluated via the example of simulated spacecraft attitude estimation. The following six algorithms are evaluated and compared.

1. The traditional MEKF, that is, with state-space model from Equations (11) and (12). In this algorithm, the rigorous error quaternion construction from Equation (30) has been used. In this respect, here the MEKF is just the QLIEKF in Gui and de Ruiter (2018).

2. MEKF with state-space model from Equations (11) and (13), denoted as invariant MEKF (IMEKF). Here we make use of the terminology ‘invariant’ to express that the measurement transition matrix is independent of the global state estimate.
3. MEKF with state-space model from Equations (25) and (12), denoted as GEKF. This state-space model is just the same as that in GEKF. So, we have made use of the same terminology ‘GEKF’ in Andrle and Crassidis (2015).

4. MEKF with state-space model from Equations (25) and (13), denoted as invariant GEKF (IGEKF).

5. MEKF with state-space model from Equations (16) and (19), denoted as MEKF-ref. In this algorithm, the attitude error is expressed in reference frame.

6. MEKF with state-space model from Equations (28) and (19), denoted as QRIEKF. In this algorithm, the attitude error is expressed in reference frame. This is just the filter developed in Gui and de Ruiter (2018), so we make use of the same terminology.

The numerical example used is the same as that in Gui and de Ruiter (2018). The simulation parameters are all the same as those in Gui and de Ruiter (2018) for a freely tumbling spacecraft in a circular orbit. The spacecraft angular velocity is measured by the gyroscopes at a frequency of 0.1 Hertz. The vector observations are provided by the sun sensor and magnetometer. The sampling frequency of the two sensors is 1 Hz. The 12th Generation (IGRF-12) model is used to provide the reference magnetic field vector.

First, a scenario with large initial estimation errors is considered. The noise parameters for the gyroscope measurements are given by $\eta_u = \sqrt{10} \times 10^{-10}$ and $\eta_v = \sqrt{10} \times 10^{-7}$. The covariance of sun sensor noise is set to $0.0175^2 \mathbf{I}_{3 \times 3}$ and the covariance of magnetometer noise is set to $0.0873^2 \mathbf{I}_{3 \times 3}$. The true initial attitude and gyroscope bias are generated randomly according to $\alpha(0) \sim N(\mathbf{0}_{3 \times 1}, (150^\circ)^2 \mathbf{I}_{3 \times 3})$ and $\beta(0) \sim N(\mathbf{0}_{3 \times 1}, (20^\circ/\text{h})^2 \mathbf{I}_{3 \times 3})$. For each filter, the initial attitude error covariance is set as $(150^\circ)^2$ for each axis and the initial bias error covariance is set as $(20^\circ/\text{h})^2$. The initial bias estimate is set to zero. The initial attitude quaternion estimate is set to unit quaternion. A total of 100 Monte Carlo runs are conducted, each of which spans 60 min. The root mean squared errors (RMSEs) of the attitude and gyroscope bias estimates by different filters are shown in Figures 1 and 2, respectively. It is shown that all the filters with trajectory-independent measurement models have a good performance in terms of both convergence speed and steady-state estimation accuracy. All of them demonstrate significantly better performance than the traditional MEKF and GEKF. In this simulation scenario, the superiority of GEKF over MEKF has not been observed. It can be said that the positive effect provided by the common frame treatment of the gyroscope bias has been counteracted by the negative effect of the
According to the difference between MEKF and IMEKF, it can be concluded that the state prediction-dependent measurement model in MEKF has a negative effect on the filtering performance. For the four filters with trajectory-independent measurement models, their process models are all trajectory-dependent due to the incorporation of the gyroscope bias as state element. However, regarding the attitude part, the corresponding model is trajectory-independent. This is just the primary reason why these filters can still work well for large initialisation errors with the linear Kalman filtering framework. There are slight differences in the steady-state estimates between the four filters with trajectory-independent measurement models. MEKF-ref and QRIEKF perform a little better than IMEKF and IGEKF. This difference is caused by the incorporation form of the gyroscope drift bias estimate in the process models. For MEKF-ref and QRIEKF, the relation between the attitude error and gyroscope bias error is linear as shown in Equations (16) and (28), while it is nonlinear for IMEKF and IGEKF as shown in Equations (11) and (25). However, the performance of MEKF-ref and QRIEKF is of the same grade, so it cannot be concluded which one of $SE(3)$ and $SO(3)$ is better than the other. In previous works, the $SE(3)$ state formulation is always argued to be better than the $SO(3)$ formulation. This is because with the $SE(3)$ formulation, the resultant state-space model is trajectory-independent, while this is not the case for the $SO(3)$ formulation. However, for the spacecraft attitude estimate problem, no matter the $SE(3)$ or $SO(3)$ formulation, the attitude error models are both trajectory-independent if the gyroscope bias is not considered. Although IGEKF performs better than IMEKF for attitude estimation in this case, IMEKF performs better than IGEKF in the next case. Therefore, we cannot judge which one is superior or inferior.

Second, the severe case with a small initial error covariance but extremely large initial state error is considered. The noise parameters for the gyroscope measurements are given by $\eta_u = \sqrt{10 \times 10^{-8}}$ and $\eta_v = \sqrt{10 \times 10^{-5}}$. The covariance of sun sensor noise and magnetometer noise are the same as the first case. The initial attitude quaternion estimate is still set to unit quaternion and the initial bias estimate is still set to zero. But the initial true gyroscope bias is set as $\beta(0) = [100, 10, 10]^{\circ}/h$ and the initial true attitude quaternion is set as $q(0) = [1, 0, 0, 0]^T$, which means that the initial attitude error is $180^{\circ}$. For each filter, the initial attitude error covariance is set as $(10^{\circ})^2$ for each axis and the initial bias error covariance is set as $(5^{\circ}/h)^2$. A total of 100 Monte Carlo runs are conducted, each of which spans 80 min. The RMSEs of the attitude and gyroscope bias estimates by these filters are shown.

Figure 2. RMSEs of gyroscope drift bias estimate errors for large initial estimation errors.
in Figures 3 and 4, respectively. The results are consistent with those in the last case. All the filters with trajectory-independent measurement models perform much better than the traditional MEKF. This further demonstrates the importance of the trajectory-independent measurement models. In this case, MEKF-ref and QRIEKF still perform a little better than IMEKF and IGEKF for attitude estimation. However, for gyroscope bias estimation, IMEKF performs the best. For attitude estimation, IMEKF performs a little better than IGEKF, which is opposite to the last case. This further indicates that $\mathbb{SE}(3)$
or $\mathbb{SO}(3)$ are only two different state formulations and there is no definitive conclusion as to which one is better. Their performance in filtering is problem-dependent.

7. Conclusion

In this paper, the MEKF is revisited for attitude estimation with vector observations making use of group Lie theory. It is pointed out that the dynamic model of the attitude estimation is group-affine and, therefore, the corresponding error state model will be trajectory-independent. With the trajectory-independent linear state-space models, it is possible to cope with large initialisation errors even within the linear Kalman filtering framework. However, the measurement model used in existing MEKF is dependent on the attitude prediction, which is inconsistent with invariant theory. Substituting the attitude prediction related vectors with the vector observations in the body frame, the resultant measurement model is now trajectory-independent. Meanwhile, the measurement model with reference frame attitude error definition can also be transformed into trajectory-independent form. Monte Carlo simulations were conducted to evaluate the MEKF variants with different state-space models. The results have shown that the performance of MEKF can be much improved with the trajectory-independent measurement models in some severe scenarios such as large initial estimation errors or erroneous initial error covariance.

Funding statement. This research was supported by the National Natural Science Foundation of China under the project 61873275.

Acknowledgements. The author would like to thank Prof. Haichao Gui from the School of Astronautics, Beihang University, for providing the numerical examples codes.

References

Andrle, M. S. and Crassidis, J. L. (2015). Attitude estimation employing common frame error representations. *Journal of Guidance Control and Dynamics*, 38(9), 1614–1624.

Barrau, A. (2015). Non-linear state error based extended Kalman filters with applications to navigation. Thesis, MINES Paristech, Paris, France.

Barrau, A. and Bonnabel, S. (2014). Intrinsic filtering on Lie groups with applications to attitude estimation. *IEEE Transactions on Automatic Control*, 60(2), 436–449.

Barrau, A. and Bonnabel, S. (2015). An EKF-SLAM algorithm with consistency properties. *arXiv preprint*, arXiv:1510.06263.

Barrau, A. and Bonnabel, S. (2017). The invariant extended Kalman filter as a stable observer. *IEEE Transactions on Automatic Control*, 62(4), 1797–1812.

Chang, L. B. (2020). SE(3) based extended Kalman filter for spacecraft attitude estimation. *arXiv preprint*, arXiv:2003.12978.

Cheng, Y. and Crassidis, J. L. (2010). Particle filtering for attitude estimation using a minimal local-error representation. *Journal of Guidance, Control, and Dynamics*, 33(4), 1305–1310.

Crassidis, J. L. and Junkins, J. L. (2012). *Optimal Estimation of Dynamic Systems* (2nd ed.). Boca Raton, FL: CRC Press, Chap. 3, pp. 153–154, Chap. 7, pp. 421–424.

Crassidis, J. L. and Markley, F. L. (2003). Unscented filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 26(4), 536–542.

Crassidis, J. L., Markley, F. L. and Cheng, Y. (2007). Survey of nonlinear attitude estimation methods. *Journal of Guidance, Control, and Dynamics*, 30(1), 12–28.

Cui, J., Wang, M., Wu, W. and He, X. (2021). Lie group based nonlinear state errors for MEMS-IMU/GNSS/magnetometer integrated navigation. *The Journal of Navigation*, 1–14. doi:10.1017/S037346332100014X

Gai, E., Daly, K., Harrison, J. and Lemos, L. (1985). Star-sensor-based satellite attitude/attitude rate estimator. *Journal of Guidance, Control, and Dynamics*, 8(5), 560–565.

Gui, H. and de Ruiter, A. H. J. (2018). Quaternion invariant extended Kalman filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 41(4), 863–878.

Hartley, R., Ghaifari, M., Eustice, R. M. and Grizzle, J. W. (2020). Contact-aided invariant extended Kalman filtering for robot state estimation. *The International Journal of Robotics Research*, 39(4), 402–430.

Lefferts, E. J., Markley, F. L. and Shuster, M. D. (1982). Kalman filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 5(5), 417–429.

Li, M. and Mourtakis, A. I. (2012). Improving the Accuracy of EKF-Based Visual-Inertial Odometry. *2012 IEEE International Conference on Robotics and Automation*, St. Paul, MN, 828–835.

Markley, F. L. (2003). Attitude error representations for Kalman filtering. *Journal of Guidance, Control, and Dynamics*, 26(2), 311–317.
Appendix

In this Appendix, we will derive the attitude error state model without any approximation as that in Equations (10) and (15). First, we will derive the attitude error state model in the body frame. The attitude error in body frame is defined in Equation (9). Taking derivative of both sides of Equation (9) gives

\[
\frac{dA(q)}{dt} = \dot{A}(q)A(q)^{-1} - A(q)\dot{A}(q)^{-1}A(q)
\]

\[
= - (\omega \times)A(q)A(q)^{-1} + A(q)\dot{A}(q)^{-1}(\omega \times)
\]

\[
= - (\omega \times)A(\delta q) + A(\delta q)(\omega \times) \tag{A1}
\]

Suppose the initial vector-form error corresponds to \(A(\delta q(0))\) as \(\delta \alpha^b(0)\). They are related to each other through the exponential map as

\[
A(\delta q(0)) = \exp[\delta \alpha^b(0)] = \exp_m[\delta \alpha^b(0)\times] \tag{A2}
\]

Define the following attitude error

\[
A(\delta q) = A(q)A(\delta q(0))A(q)^{-1} \tag{A3}
\]

Taking the derivative of both sides of Equation (A3) gives

\[
\frac{dA(\delta q)}{dt} = \dot{A}(q)A(\delta q(0))A(q)^{-1} - A(q)A(\delta q(0))A(q)^{-1}\dot{A}(q)A(q)^{-1}
\]

\[
= -(\omega \times)A(q)A(\delta q(0))A(q)^{-1} + A(q)A(\delta q(0))A(q)^{-1}(\omega \times)A(q)A(q)^{-1}
\]

\[
= -(\omega \times)A(\delta q) + A(\delta q)(\omega \times) \tag{A4}
\]

which means that the defined attitude error in Equation (A2) is a solution of the attitude error model in Equation (A1).

Suppose the initial vector-form error corresponds to \(A(\delta q)\) as \(\delta \alpha^b\). According to the relationship between the matrix Lie group and its associated Lie algebra, Equation (A3) can be rewritten as

\[
\exp(\delta \alpha^b) = A(q)\exp[\delta \alpha^b(0)]A(q)^{-1} = \exp[A(q)\delta \alpha^b(0)] \tag{A5}
\]

That is, we have

\[
\delta \alpha^b = A(q)\delta \alpha^b(0) \tag{A6}
\]

Taking the derivative of both sides of Equation (A6) gives

\[
\delta \dot{\alpha}^b = -(\omega \times)A(q)\delta \alpha^b(0) = -(\omega \times)\delta \alpha^b \tag{A7}
\]

which is just the attitude error model involved in Equation (11). From Equation (A1) to (A7), it is shown that there is no approximation when deriving the attitude error model. That is to say, the attitude error model is also applicable with large initial attitude error.
Similarly, we now derive the attitude error model in reference frame. Taking the derivative of both sides of Equation (14) gives

\[
\frac{dA(\delta q)}{dt} = -A(\dot{\hat{q}})A(\dot{\hat{q}})A^{-1}(q) + A(\dot{\hat{q}})A^{-1}(q)
\]

\[
= A(\dot{\hat{q}})(\omega \times)A(\dot{\hat{q}})A(\delta q) - A(\dot{\hat{q}})(\omega \times)A(q)
\]

\[
= A(\dot{\hat{q}})(\omega \times)A(\dot{\hat{q}})A(\delta q) - A(\dot{\hat{q}})(\omega \times)A(\dot{\hat{q}})A(\dot{\hat{q}})A^{-1}(q)
\]

\[
= 0_{3\times3}
\]

Suppose the initial vector-form error corresponding to \( A(\delta q) \) is \( \delta \alpha^r \). They are related to each other through the exponential map as

\[
A(\delta q) = \exp(\alpha^r) = \exp_m(\alpha^r \times)
\]

It can be easily deduced from Equation (A8) and (A9) that

\[
\dot{\alpha}^r = 0_{3\times1}
\]

which is just the attitude error model involved in Equation (16). From Equation (A8) to (A10), it is shown that there is no approximation when deriving the attitude error model. That is to say, the attitude error model is also applicable with large initial attitude error.