2014 American Control Conference (ACC)
June 4-6, 2014. Portland, Oregon, USA

SO(3)-Constrained Kalman Filtering with Application to Attitude Estimation

Anton H. J. de Ruiter[1]

***Abstract*— This paper presents a new continuous-time Kalman-like filter directly on $SO(3)$ for attitude estimation of a vehicle with vector and gyro measurements. The presented filter is inspired by a recently developed norm-constrained Kalman-like filter, which has successfully been applied to the attitude estimation problem when a quaternion parameterization is used. The filter is compared with a continuous-time multiplicative extended Kalman filter, which is also formulated directly on $SO(3)$.**

I. INTRODUCTION

A significant amount of effort has been spent on the development of different methods for the determination (or estimation) of the orientation of a vehicle, which is described by the special orthogonal group $SO(3)$ (the rotation matrices). References [1] to [7] are some examples of deterministic methods developed in the aerospace and robotics communities. References [8] and [9] provide an extensive overview of related state-estimation techniques.

While a rotation matrix contains nine components, only three of them are independent. This can cause computational difficulties, and practitioners have typically used unconstrained three-set parameterizations of the rotation matrix or the unit quaternion [10]. The disadvantage of three-set parameterizations is the presence of a singularity, and the disadvantage for the unit quaternion is non-uniqueness of representation. Consequently, there has been significant recent interest in performing state estimation and control directly on $SO(3)$ [11], [12].

In reference [13], a norm-constrained Kalman filter was presented, which allowed state estimation to be performed directly using the unit quaternion, unlike other approaches such as the multiplicative extended Kalman filter [14]. Inspired by [13], an $SO(3)$-constrained extended Kalman filter ($SO(3)$-EKF) is presented in this paper, allowing state estimation to be performed directly on $SO(3)$, using the rotation matrix. Unlike [13], which treats the vector discrete-time norm-constrained case, this paper treats the matrix continuous-time $SO(3)$-constrained case. This allows the comparison with a continuous-time multiplicative extended Kalman filter (MEKF) [14], which can also be formulated directly on $SO(3)$. It is demonstrated by a numerical example that the new $SO(3)$-EKF has similar performance to the MEKF, and is almost identical at steady-state. This presents the $SO(3)$-EKF as a viable option for attitude estimation problems. A disadvantage of the $SO(3)$-EKF is that it requires a 9 by 9 covariance matrix, while the MEKF requires only a 3 by 3 covariance matrix. However, a clear advantage of the $SO(3)$-EKF is that it is extendable to $SO(n)$, which has been recently completed in a separate work [15].

The remainder of the paper is organized as follows. Section II presents the problem formulation for attitude estimation of a vehicle using vector and gyro measurements. Section III presents a continuous-time MEKF. Section IV presents the $SO(3)$-constrained extended Kalman filter. Section V presents a numerical comparison between the two filters, and finally Section VI contains concluding remarks.

[1]Anton de Ruiter is with the Department of Aerospace Engineering, Ryerson University, Toronto, ON, Canada. aderuiter@ryerson.ca

II. PROBLEM FORMULATION

This paper considers the attitude determination problem for a vehicle with continuous-time vector and gyro measurements. The attitude kinematics and measurement models are presented in this section.

A. Attitude Kinematics and Gyro Model

The attitude kinematics obey Poisson's equations [16]

$$\dot{\mathbf{C}}(t) = -\boldsymbol{\omega}(t)^{\times}\mathbf{C}(t), \tag{1}$$

where $\mathbf{C}(t) \in SO(3)$ represents the vehicle's attitude (where $SO(3) = \{\mathbf{C} \in \mathbb{R}^{3\times 3} : \mathbf{C}^T\mathbf{C} = \mathbf{1}_{3\times 3}, \det \mathbf{C} = 1\}$), and $\boldsymbol{\omega}(t) \in \mathbb{R}^3$ is the vehicle angular velocity in vehicle body coordinates. Note that in this paper, $\mathbf{1}_{n\times n}$ denotes an n by n identity-matrix, and $\mathbf{0}_{n\times m}$ denotes an n by m matrix of zeros.

The gyro provides measurements of the vehicle's angular velocity according to the well-known model [17]

$$\boldsymbol{\omega}^m(t) = \boldsymbol{\omega}(t) + \mathbf{b}(t) + \mathbf{w}_\omega(t), \tag{2}$$

where $\boldsymbol{\omega}^m \in \mathbb{R}^3$ is the measured angular velocity, $\mathbf{b} \in \mathbb{R}^3$ is the gyro measurement bias and $\mathbf{w}_\omega(t) \in \mathbb{R}^3$ is the gyro measurement noise, which is modeled as a zero-mean white noise process. It is assumed that the bias obeys a random walk

$$\dot{\mathbf{b}}(t) = \mathbf{w}_b(t), \tag{3}$$

where $\mathbf{w}_b(t) \in \mathbb{R}^3$ is a zero-mean white noise process. Defining

$$\mathbf{w}(t) = \begin{bmatrix} \mathbf{w}_\omega(t) \\ \mathbf{w}_b(t) \end{bmatrix}, \tag{4}$$

the combined gyro noise $\mathbf{w}(t)$ is taken to have autocovariance

$$E\{\mathbf{w}(t)\mathbf{w}(t)^T\} = \mathbf{Q}(t), \tag{5}$$

where $\mathbf{Q}(t) \in \mathbb{R}^{6\times 6}$ is symmetric and positive definite, and $E\{\cdot\}$ denotes the expectation operator.

B. Vector Measurements

At time t, the availability of $m(t) \in \mathbb{N}$ vector measurements is assumed, described by

$$\mathbf{s}_i^b(t) = \mathbf{C}(t)\mathbf{s}_i^I(t) + \mathbf{v}_i(t), \;\; i = 1, ..., m(t), \tag{6}$$

where $\mathbf{s}_i^b \in \mathbb{R}^3$ is the vector measurement in vehicle body coordinates, $\mathbf{s}_i^I \in \mathbb{R}^3$ is the corresponding known reference vector in inertial coordinates, and $\mathbf{v}_i(t)$ is the measurement error, modeled as a zero-mean white noise process.

The measurement vector is then assembled as

$$\mathbf{y}(t) = \mathbf{h}(\mathbf{C}(t), t) + \mathbf{v}(t), \tag{7}$$

where

$$\mathbf{y}(t) = \underset{i=1,\ldots,m(t)}{\operatorname{col}}\{\mathbf{s}_i^b(t)\}, \;\; \mathbf{v}(t) = \underset{i=1,\ldots,m(t)}{\operatorname{col}}\{\mathbf{v}_i(t)\}, \tag{8}$$

and

$$\mathbf{h}(\mathbf{C}(t), t) = \underset{i=1,\ldots,m(t)}{\operatorname{col}}\{\mathbf{C}(t)\mathbf{s}_i^I(t)\}. \tag{9}$$

Note that for matrices $\mathbf{X}_i$ with the same number of columns for $i = 1, ..., m$, the following notation is used

$$\underset{i=1,\ldots,m}{\operatorname{col}}\{\mathbf{X}_i\} = \begin{bmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_m \end{bmatrix}.$$

Finally, it is assumed that the measurement noise has autocovariance

$$E\{\mathbf{v}(t)\mathbf{v}(t)^T\} = \mathbf{R}(t),$$

where $\mathbf{R}(t)$ is symmetric and positive-definite.

For brevity, from this point on, explicit time-dependence of variables will not be shown.

III. Multiplicative Extended Kalman Filter

This section describes a continuous-time MEKF formulation for the attitude determination problem using vector and gyro measurements. Note that this can be viewed as a limiting case of the discrete-time MEKF presented in [14]. The objective is to estimate the attitude $\mathbf{C} \in SO(3)$, together with the gyro bias $\mathbf{b} \in \mathbb{R}^3$. To this end, denote the attitude and gyro bias estimates by $\widehat{\mathbf{C}} \in SO(3)$ and $\widehat{\mathbf{b}} \in \mathbb{R}^3$, respectively.

A. Filter Structure

The following filter structure is imposed

$$\begin{aligned} \dot{\widehat{\mathbf{C}}} &= -\left(\boldsymbol{\omega}^m - \widehat{\mathbf{b}} + \mathbf{K}_\phi^M\left[\mathbf{y} - \mathbf{h}(\widehat{\mathbf{C}}, t)\right]\right)^\times \widehat{\mathbf{C}}, \quad (10) \\ \dot{\widehat{\mathbf{b}}} &= \mathbf{K}_b^M\left[\mathbf{y} - \mathbf{h}(\widehat{\mathbf{C}}, t)\right], \quad (11) \end{aligned}$$

where $\mathbf{K}_\phi^M, \mathbf{K}_b^M \in \mathbb{R}^{3\times 3m(t)}$ are to be determined gain matrices.

B. Error Dynamics and Gain Selection

Let $\mathbf{\Phi} \in \mathbb{R}^3$ be a rotation vector representing the attitude estimation error according to

$$\mathbf{C}_\phi(\mathbf{\Phi}) = \mathbf{C}\widehat{\mathbf{C}}^T, \tag{12}$$

where

$$\mathbf{C}_\phi(\mathbf{\Phi}) = \mathbf{1}_{3\times 3} + \frac{(1-\cos\phi)}{\phi^2}\mathbf{\Phi}^\times\mathbf{\Phi}^\times - \frac{\sin\phi}{\phi}\mathbf{\Phi}^\times, \tag{13}$$

with $\phi = \|\mathbf{\Phi}\|$, and for a given $\mathbf{a} = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}^T$, the corresponding skew-symmetric cross-product matrix $\mathbf{a}^\times \in \mathbb{R}^{3\times 3}$ is defined as

$$\mathbf{a}^\times \triangleq \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}.$$

Next, define the gyro bias estimation error according to

$$\widetilde{\mathbf{b}} = \mathbf{b} - \widehat{\mathbf{b}}. \tag{14}$$

Assuming small errors, the measurement residual in (10) and (11) can be approximated by

$$\mathbf{y} - \mathbf{h}(\widehat{\mathbf{C}}, t) = \mathbf{H}_\phi\mathbf{\Phi} + \mathbf{v}, \tag{15}$$

where

$$\mathbf{H}_\phi = \left.\frac{\partial \mathbf{h}}{\partial \mathbf{\Phi}}\right|_{\widehat{\mathbf{C}},t,} = \underset{i=1,\ldots,m(t)}{\operatorname{col}}\left\{(\widehat{\mathbf{C}}\mathbf{s}_i^I)^\times\right\}.$$

From equations (1), (10) and (12), one obtains

$$\begin{aligned} \frac{d}{dt}\left(\mathbf{C}_\phi(\mathbf{\Phi})\right) \;\; = \;\; &-\Big(\boldsymbol{\omega}^m - \widehat{\mathbf{b}} - \mathbf{w}_\omega - \widetilde{\mathbf{b}} - \mathbf{C}_\phi(\mathbf{\Phi})\,(\boldsymbol{\omega}^m \\ &-\widehat{\mathbf{b}} + \mathbf{K}_\phi^M(\mathbf{y} - \mathbf{h}(\widehat{\mathbf{C}}, t))\big)\Big)^\times \mathbf{C}_\phi(\mathbf{\Phi}). \end{aligned}$$

From this, one can immediately obtain [16]

$$\begin{aligned} \dot{\mathbf{\Phi}} \;\; = \;\; &\mathbf{S}_\phi(\mathbf{\Phi})^{-1}\Big(\boldsymbol{\omega}^m - \widehat{\mathbf{b}} - \mathbf{w}_\omega - \widetilde{\mathbf{b}} - \mathbf{C}_\phi(\mathbf{\Phi})\Big(\boldsymbol{\omega}^m - \widehat{\mathbf{b}} \\ &+\mathbf{K}_\phi^M(\mathbf{y} - \mathbf{h}(\widehat{\mathbf{C}}, t))\Big)\Big), \quad (16) \end{aligned}$$

where

$$\mathbf{S}_\phi(\mathbf{\Phi})^{-1} = \mathbf{1}_{3\times 3} + \frac{\mathbf{\Phi}^\times}{2} + \frac{1}{\phi^2}\left(1 - \frac{\phi/2}{\tan(\phi/2)}\right)\mathbf{\Phi}^\times\mathbf{\Phi}^\times,$$

is the inverse kinematic matrix associated with the rotation vector [16]. Next, it can be verified that

$$\mathbf{S}_\phi(\mathbf{\Phi})^{-1}\mathbf{C}_\phi(\mathbf{\Phi}) = \mathbf{S}_\phi(-\mathbf{\Phi})^{-1}.$$

Making use of this and (15), the attitude error kinematics in (16) can be approximated to first order as

$$\dot{\mathbf{\Phi}} = -\left(\boldsymbol{\omega}^m - \widehat{\mathbf{b}}\right)^\times \mathbf{\Phi} - \widetilde{\mathbf{b}} - \mathbf{w}_\omega - \mathbf{K}_\phi^M\mathbf{H}_\phi\mathbf{\Phi} - \mathbf{K}_\phi^M\mathbf{v}. \tag{17}$$

Next, the bias error dynamics are readily obtained as

$$\dot{\widetilde{\mathbf{b}}} = -\mathbf{K}_b^M\left[\mathbf{y} - \mathbf{h}(\widehat{\mathbf{C}}, t)\right] + \mathbf{w}_b,$$

which become to first order

$$\dot{\widetilde{\mathbf{b}}} = -\mathbf{K}_b^M\mathbf{H}_\phi\mathbf{\Phi} - \mathbf{K}_b{}^M\mathbf{v} + \mathbf{w}_b. \tag{18}$$

The error dynamics in (17) and (18) are now assembled as

$$\dot{\widetilde{\mathbf{x}}}^M = \left(\mathbf{A}^M - \mathbf{K}^M\mathbf{H}^M\right)\widetilde{\mathbf{x}}^M - \mathbf{K}^M\mathbf{v} + \mathbf{G}^M\mathbf{w}, \tag{19}$$

where

$$\widetilde{\mathbf{x}}^M = \begin{bmatrix} \mathbf{\Phi} \\ \widetilde{\mathbf{b}} \end{bmatrix},$$

is the error state vector,

$$\mathbf{A}^M = \begin{bmatrix} -\left(\boldsymbol{\omega}^m - \widehat{\mathbf{b}}\right)^\times & -\mathbf{1}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \end{bmatrix}, \ \mathbf{K}^M = \begin{bmatrix} \mathbf{K}_\phi^M \\ \mathbf{K}_b^M \end{bmatrix},$$

$$\mathbf{H}^M = \begin{bmatrix} \mathbf{H}_\phi & \mathbf{0}_{3m(t)\times3} \end{bmatrix}, \ \mathbf{G}^M = \begin{bmatrix} -\mathbf{1}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{1}_{3\times3} \end{bmatrix}.$$

Now, define the error covariance as

$$\mathbf{P}^M = E\{\widetilde{\mathbf{x}}^M\widetilde{\mathbf{x}}^{M,T}\}. \tag{20}$$

As is customary in the development of an extended Kalman filter, the variables $\widehat{\mathbf{C}}$, $\boldsymbol{\omega}^m$ and $\widehat{\mathbf{b}}$ are treated as deterministic [19]. Then, under the standard assumption that $\widetilde{\mathbf{x}}^M$, $\mathbf{w}$ and $\mathbf{v}$ are uncorrelated, one obtains [18], [19]

$$\begin{aligned}\dot{\mathbf{P}}^M &= \left(\mathbf{A}^M - \mathbf{K}^M\mathbf{H}^M\right)\mathbf{P}^M + \mathbf{P}^M\left(\mathbf{A}^M - \mathbf{K}^M\mathbf{H}^M\right)^T \\ &\quad + \mathbf{G}^M\mathbf{Q}\mathbf{G}^{M,T} + \mathbf{K}^M\mathbf{R}\mathbf{K}^{M,T}.\end{aligned} \tag{21}$$

Together, (19) and (21) are in the standard form of the state-error and covariance propagations for a continuous-time Kalman filter [19]. The gain $\mathbf{K}^M$ is therefore selected to minimize $J(\mathbf{K}^M) = \text{trace}[\dot{\mathbf{P}}^M]$ as in the continuous-time Kalman filter, which yields

$$\mathbf{K}^M = \mathbf{P}^M\mathbf{H}^{M,T}\mathbf{R}^{-1}. \tag{22}$$

C. MEKF Summary

To summarize, the continuous-time multiplicative Extended Kalman filter equations are given by equations (10), (11), (21) and (22). Note that the estimation of $\widehat{\mathbf{C}}$ is performed directly on $SO(3)$, without the need for any attitude parameterization.

IV. SO(3)-Constrained Kalman Filtering

In this section, similar in spirit to the norm-constrained Extended Kalman filter in [13], a continuous-time $SO(3)$-constrained Extended Kalman filter is developed for the attitude estimation problem using vector and gyro measurements.

A. Constraint Equations and Vectorization of $\mathbf{C}$

Let $\mathbf{c}_i \in \mathbb{R}^3$ denote the i^{th} column of $\mathbf{C}$, such that

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 \end{bmatrix}. \tag{23}$$

The constraint $\mathbf{C} \in SO(3)$, leads to the independent constraints

$$\mathbf{c}_i^T\mathbf{c}_j = \delta_{ij}, \ \ i = 1,2,3, \ \ j = i,...,3, \tag{24}$$

where

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

These constraints come from the orthonormality condition $\mathbf{C}^T\mathbf{C} = \mathbf{1}$. By symmetry, only the upper-triangular terms are needed. In addition to (24), there is also a positive determinant constraint. However, the orthonormality constraint $\mathbf{C}^T\mathbf{C} = \mathbf{1}$, yields $|\det\mathbf{C}| = 1$. Since the estimator is being developed in continuous-time, by continuity it is impossible for $\det\widehat{\mathbf{C}} = 1$ to jump to $\det\widehat{\mathbf{C}} = -1$ (where $\widehat{\mathbf{C}}$ denotes the estimate of $\mathbf{C}$). Therefore, the determinant constraint does not need to be included.

Differentiating (24), one obtains

$$\dot{\mathbf{c}}_i^T\mathbf{c}_j + \mathbf{c}_i^T\dot{\mathbf{c}}_j = 0, \ \ i = 1,2,3, \ \ j = i,...,3, \tag{25}$$

Clearly, if the differential constraints in (25) are satisfied for all $t \geq t_0$, and if the initial conditions $\mathbf{c}_i(t_0)$ for $i = 1,2,3$ satisfy (24), then (24) must also be satisfied for all time. Therefore, the differential constraints given in (25) are enforced, which amounts to forcing $\dot{\mathbf{C}}$ to lie in the tangent space of $SO(3)$.

The matrix $\mathbf{C}$ is now vectorized as

$$\mathbf{c}_{vec} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \mathbf{c}_3 \end{bmatrix}. \tag{26}$$

Consequently, (25) may be written as

$$\mathbf{c}_{vec}^T\mathbf{T}_{ij}\dot{\mathbf{c}}_{vec} = 0, \ \ i = 1,2,3, \ \ j = i,...,3, \tag{27}$$

where $\mathbf{T}_{ij} \in \mathbb{R}^{9\times9}$ is a square block matrix with each block belonging to $\mathbb{R}^{3\times3}$. Denoting the kl^{th} block in $\mathbf{T}_{ij}$ as $\mathbf{T}_{ij}^{kl} \in \mathbb{R}^{3\times3}$, it is defined as

$$\mathbf{T}_{ij}^{kl} = \begin{cases} \mathbf{1}_{3\times3}, & (k,l) = (i,j) \text{ or } (k,l) = (j,i), \\ \mathbf{0}_{3\times3}, & \text{otherwise.} \end{cases} \tag{28}$$

Equivalence of (25) and (27) is readily verified by direct multiplication.

Next, it may be verified by direct multiplication that

$$\mathbf{c}_{vec}^T\mathbf{T}_{kl}\mathbf{T}_{ij}\mathbf{c}_{vec} = \begin{cases} 1, & k = l = i = j, \\ 2, & k = i \text{ and } l = j \text{ with } k \neq l, \\ 0, & otherwise, \end{cases} \tag{29}$$

for $k = 1,2,3$, $l = k,...,3$ and $i = 1,2,3$, $j = i,...,3$. As a result, the vectors defined by $\mathbf{e}_{ij} = \mathbf{T}_{ij}\mathbf{c}_{vec}$ for $j = i,...,n$, $i = 1,...,n$, are orthogonal.

Finally, the rotational kinematics in (1) may be written in terms of $\mathbf{c}_{vec}$ as

$$\dot{\mathbf{c}}_{vec} = -\text{diag}\{\boldsymbol{\omega}^\times, \boldsymbol{\omega}^\times, \boldsymbol{\omega}^\times\}\mathbf{c}_{vec}. \tag{30}$$

B. Measurement Equation

Denoting $\mathbf{s}_i^I = \begin{bmatrix} s_{i,1}^i & s_{i,2}^I & s_{i,3}^I \end{bmatrix}^T$, it is easy to show that the vector measurement equation in (7) can be written as

$$\mathbf{y} = \mathbf{H}_c\mathbf{c}_{vec} + \mathbf{v}, \tag{31}$$

where

$$\mathbf{H}_c = \underset{i=1,...,m(t)}{\text{col}}\left\{\begin{bmatrix} s_{i,1}^I\mathbf{1}_{3\times3} & s_{i,2}^I\mathbf{1}_{3\times3} & s_{i,3}^I\mathbf{1}_{3\times3} \end{bmatrix}\right\}.$$

C. Filter Structure

Denoting the estimate of $\mathbf{c}_{vec}$ by $\widehat{\mathbf{c}}_{vec}$, the following filter structure is imposed

$$\begin{aligned}\dot{\widehat{\mathbf{c}}}_{vec} &= -\mathrm{diag}\{\bar{\boldsymbol{\omega}}^{\times},\bar{\boldsymbol{\omega}}^{\times},\bar{\boldsymbol{\omega}}^{\times}\}\widehat{\mathbf{c}}_{vec} \\ &\quad +\mathbf{K}_c^S\left[\mathbf{y}-\mathbf{H}_c\widehat{\mathbf{c}}_{vec}\right], \qquad (32)\\ \dot{\widehat{\mathbf{b}}} &= \mathbf{K}_b^S\left[\mathbf{y}-\mathbf{H}_c\widehat{\mathbf{c}}_{vec}\right], \qquad (33)\end{aligned}$$

where

$$\bar{\boldsymbol{\omega}}=\boldsymbol{\omega}^m-\widehat{\mathbf{b}}, \qquad (34)$$

and $\mathbf{K}_c^S \in \mathbb{R}^{9\times 3m(t)}$ and $\mathbf{K}_b^S \in \mathbb{R}^{3\times 3m(t)}$ are to be determined gain matrices.

D. Error Dynamics and Gain Selection

The state vector to be estimated is assembled as

$$\mathbf{x}^S=\begin{bmatrix}\mathbf{c}_{vec}\\ \mathbf{b}\end{bmatrix}, \qquad (35)$$

with corresponding estimate $\widehat{\mathbf{x}}^S$ and error $\widetilde{\mathbf{x}}^S=\mathbf{x}^S-\widehat{\mathbf{x}}^S$. Assuming small estimation errors $\widetilde{\mathbf{x}}^S$, and small process and measurement noises ($\mathbf{w}$ and $\mathbf{v}$, respectively), the error dynamics can be obtained to first order as (by using (30), (3), (31), (32) and (33))

$$\dot{\widetilde{\mathbf{x}}}^S=\left(\mathbf{A}^S-\mathbf{K}^S\mathbf{H}^S\right)\widetilde{\mathbf{x}}^S-\mathbf{K}^S\mathbf{v}+\mathbf{G}^S\mathbf{w}, \qquad (36)$$

where

$$\mathbf{A}^S=-\begin{bmatrix}\bar{\boldsymbol{\omega}}^{\times} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \widehat{\mathbf{c}}_1^{\times}\\ \mathbf{0}_{3\times3} & \bar{\boldsymbol{\omega}}^{\times} & \mathbf{0}_{3\times3} & \widehat{\mathbf{c}}_2^{\times}\\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \bar{\boldsymbol{\omega}}^{\times} & \widehat{\mathbf{c}}_3^{\times}\\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3}\end{bmatrix},\quad \mathbf{K}^S=\begin{bmatrix}\mathbf{K}_c^S\\ \mathbf{K}_b^S\end{bmatrix}$$

$$\mathbf{H}^S=\begin{bmatrix}\mathbf{H}_c & \mathbf{0}_{3m(t)\times3}\end{bmatrix},\quad \mathbf{G}^S=\begin{bmatrix}-\widehat{\mathbf{c}}_1^{\times} & \mathbf{0}_{3\times3}\\ -\widehat{\mathbf{c}}_2^{\times} & \mathbf{0}_{3\times3}\\ -\widehat{\mathbf{c}}_3^{\times} & \mathbf{0}_{3\times3}\\ \mathbf{0}_{3\times3} & \mathbf{1}_{3\times3}\end{bmatrix}.$$

Now, define the error covariance as

$$\mathbf{P}^S=E\{\widetilde{\mathbf{x}}^S\widetilde{\mathbf{x}}^{S,T}\}. \qquad (37)$$

As in the MEKF development, the variables $\widehat{\mathbf{c}}_{vec}$, $\boldsymbol{\omega}^m$ and $\widehat{\mathbf{b}}$ are treated as deterministic. Then, under the standard assumption that $\widetilde{\mathbf{x}}^S$, $\mathbf{w}$ and $\mathbf{v}$ are uncorrelated, one obtains [18], [19]

$$\begin{aligned}\dot{\mathbf{P}}^S &= \left(\mathbf{A}^S-\mathbf{K}^S\mathbf{H}^S\right)\mathbf{P}^S+\mathbf{P}^S\left(\mathbf{A}^S-\mathbf{K}^S\mathbf{H}^S\right)^T\\ &\quad +\mathbf{G}^S\mathbf{Q}\mathbf{G}^{S,T}+\mathbf{K}^S\mathbf{R}\mathbf{K}^{S,T}. \qquad (38)\end{aligned}$$

Now partition $\mathbf{P}^S$ as

$$\mathbf{P}^S=\begin{bmatrix}\mathbf{P}_{cc} & \mathbf{P}_{cb}\\ \mathbf{P}_{bc} & \mathbf{P}_{bb}\end{bmatrix}, \qquad (39)$$

and set

$$\mathbf{P}_1=\begin{bmatrix}\mathbf{P}_{cc}\\ \mathbf{P}_{bc}\end{bmatrix},\quad \mathbf{P}_2=\begin{bmatrix}\mathbf{P}_{cb}\\ \mathbf{P}_{bb}\end{bmatrix}. \qquad (40)$$

where $\mathbf{P}_{cc}\in\mathbb{R}^{9\times9}$, $\mathbf{P}_{bb}\in\mathbb{R}^{3\times3}$, $\mathbf{P}_{cb}\in\mathbb{R}^{9\times3}$, $\mathbf{P}_{bc}=\mathbf{P}_{cb}^T$. In addition, define

$$\mathbf{A}_c=-\begin{bmatrix}\mathrm{diag}\{\bar{\boldsymbol{\omega}}^{\times},\bar{\boldsymbol{\omega}}^{\times},\bar{\boldsymbol{\omega}}^{\times}\} & \underset{i}{\mathrm{col}}\{\widehat{\mathbf{c}}_i^{\times}\}\end{bmatrix},\quad \mathbf{A}_b=\mathbf{0}_{3\times12},$$

and partition $\mathbf{G}^S$ according to

$$\mathbf{G}^S=\begin{bmatrix}\mathbf{G}_c\\ \mathbf{G}_b\end{bmatrix}.$$

where $\mathbf{G}_c\in\mathbb{R}^{9\times6}$ and $\mathbf{G}_b\in\mathbb{R}^{3\times6}$.

Then, the block diagonal components of equation (38) can be written as

$$\begin{aligned}\dot{\mathbf{P}}_{cc} &= \left(\mathbf{A}_c-\mathbf{K}_c^S\mathbf{H}^S\right)\mathbf{P}_1+\mathbf{P}_1^T\left(\mathbf{A}_c-\mathbf{K}_c^S\mathbf{H}^S\right)^T\\ &\quad +\mathbf{G}_c\mathbf{Q}\mathbf{G}_c^T+\mathbf{K}_c^S\mathbf{R}\mathbf{K}_c^{S,T}, \qquad (41)\\ \dot{\mathbf{P}}_{bb} &= \left(\mathbf{A}_b-\mathbf{K}_b^S\mathbf{H}^S\right)\mathbf{P}_2+\mathbf{P}_2^T\left(\mathbf{A}_b-\mathbf{K}_b^S\mathbf{H}^S\right)^T\\ &\quad +\mathbf{G}_b\mathbf{Q}\mathbf{G}_b^T+\mathbf{K}_b^S\mathbf{R}\mathbf{K}_b^{S,T}. \qquad (42)\end{aligned}$$

The gains $\mathbf{K}_c^S$ and $\mathbf{K}_b^S$ are now found to minimize the cost function

$$J=\frac{1}{2}\mathrm{trace}[\dot{\mathbf{P}}]=\frac{1}{2}\mathrm{trace}[\dot{\mathbf{P}}_{cc}]+\frac{1}{2}\mathrm{trace}[\dot{\mathbf{P}}_{bb}]. \qquad (43)$$

subject to

$$\widehat{\mathbf{c}}_{vec}^T\mathbf{T}_{ij}\dot{\widehat{\mathbf{c}}}_{vec}=0,\ \ i=1,2,3,\ \ j=i,...,3. \qquad (44)$$

From (32) and (33), it can be seen that that the constraints in (44) only involve the gain $\mathbf{K}_c^S$. Likewise, by (41) and (42), it is clear that $\dot{\mathbf{P}}_{cc}$ only depends on $\mathbf{K}_c^S$, while $\dot{\mathbf{P}}_{bb}$ depends only on $\mathbf{K}_b^S$.

Therefore, the minimization of J in (43) subject to the constraint in (44) can be divided into two independent minimization problems

$$\begin{gathered}\min_{\mathbf{K}_c^S} J_c=\tfrac{1}{2}\mathrm{trace}[\dot{\mathbf{P}}_{cc}]\\ \text{subject to}\\ \widehat{\mathbf{c}}_{vec}^T\mathbf{T}_{ij}\dot{\widehat{\mathbf{c}}}_{vec}=0,\ \ i=1,2,3,\ \ j=i,...,3,\end{gathered} \qquad (45)$$

and

$$\min_{\mathbf{K}_b^S} J_b=\frac{1}{2}\mathrm{trace}[\dot{\mathbf{P}}_{bb}]. \qquad (46)$$

Problem (46) is an unconstrained minimization problem. Using (42), it has the unique solution

$$\mathbf{K}_b^S=\mathbf{P}_2^T\mathbf{H}^{S,T}\mathbf{R}^{-1}. \qquad (47)$$

Using (32), the Lagrangian [20] for the constrained problem (45) is

$$\begin{aligned}L &= \frac{1}{2}\mathrm{trace}[\dot{\mathbf{P}}_{cc}]+\sum_{i=1}^{3}\sum_{j=i}^{3}\lambda_{ij}\widehat{\mathbf{c}}_{vec}^T\mathbf{T}_{ij}\\ &\quad \times\left[-\mathrm{diag}\{\bar{\boldsymbol{\omega}}^{\times},\bar{\boldsymbol{\omega}}^{\times},\bar{\boldsymbol{\omega}}^{\times}\}\widehat{\mathbf{c}}_{vec}+\mathbf{K}_c^S\widetilde{\mathbf{y}}\right]. \qquad (48)\end{aligned}$$

where λ_{ij} are Lagrange multipliers, and

$$\widetilde{\mathbf{y}}=\mathbf{y}-\mathbf{H}_c\widehat{\mathbf{c}}_{vec}. \qquad (49)$$

Differentiating (48) with respect to $\mathbf{K}_c^S$, leads to

$$\frac{\partial L}{\partial\mathbf{K}_c^S}=-\mathbf{P}_1^T\mathbf{H}^{S,T}+\mathbf{K}_c^S\mathbf{R}+\sum_{i=1}^{3}\sum_{j=i}^{3}\lambda_{ij}\mathbf{T}_{ij}\widehat{\mathbf{c}}_{vec}\widetilde{\mathbf{y}}^T. \qquad (50)$$

In particular, note that the constraint gradients are given by

$$\mathbf{T}_{ij}\widehat{\mathbf{c}}_{vec}\widetilde{\mathbf{y}}^T,\ \ i=1,2,3,\ \ j=i,...,3. \qquad (51)$$

From (29), the vectors $\mathbf{T}_{ij}\widehat{\mathbf{c}}_{vec}$ are orthogonal. Therefore, assuming that $\widetilde{\mathbf{y}} \neq \mathbf{0}$ (which is true with probability one), the constraint gradients are linearly independent, and the first-order necessary conditions for an optimum apply [20].

Setting $\partial L / \partial \mathbf{K}_c$ in (50) to zero, the first-order necessary condition for a minimum is [20]

$$\mathbf{K}_c^S = \mathbf{K}_{c,unc} - \sum_{i=1}^{3}\sum_{j=i}^{3} \lambda_{ij} \mathbf{T}_{ij} \widehat{\mathbf{c}}_{vec} \widetilde{\mathbf{y}}^T \mathbf{R}^{-1} \tag{52}$$

where

$$\mathbf{K}_{c,unc} = \mathbf{P}_1^T \mathbf{H}^{S,T} \mathbf{R}^{-1}, \tag{53}$$

has been defined, which is the unconstrained optimal gain corresponding to (45) (obtained when the constraints are not enforced).

Note that $\mathbf{K}_b^S$ and $\mathbf{K}_{c,unc}$ in (47) and (53) can be obtained simultaneously from

$$\begin{bmatrix} \mathbf{K}_{c,unc} \\ \mathbf{K}_b^S \end{bmatrix} = \mathbf{P}^S \mathbf{H}^{S,T} \mathbf{R}^{-1}. \tag{54}$$

The Lagrange multipliers shall now be solved for. Substituting (52) into (44), and making use of (32), one has

$$\widehat{\mathbf{c}}_{vec}^T \mathbf{T}_{kl} \left[\boldsymbol{\Delta} - \sum_{i=1}^{3}\sum_{j=i}^{3} \lambda_{ij} \mathbf{T}_{ij} \widehat{\mathbf{c}}_{vec} \widetilde{\mathbf{y}}^T \mathbf{R}^{-1} \widetilde{\mathbf{y}} \right] = 0, \tag{55}$$

for $k = 1, 2, 3$ with $l = k, ..., 3$, where

$$\boldsymbol{\Delta} = -\mathrm{diag}\{\bar{\boldsymbol{\omega}}^\times, \bar{\boldsymbol{\omega}}^\times, \bar{\boldsymbol{\omega}}^\times\} \widehat{\mathbf{c}}_{vec} + \mathbf{K}_{c,unc} \widetilde{\mathbf{y}}. \tag{56}$$

Assuming that the residual $\widetilde{\mathbf{y}} \neq \mathbf{0}$, application of (29) yields

$$\lambda_{ii} = \frac{\widehat{\mathbf{c}}_{vec}^T \mathbf{T}_{ii} \boldsymbol{\Delta}}{\widetilde{\mathbf{y}}^T \mathbf{R}^{-1} \widetilde{\mathbf{y}}}, \ \ i = 1, 2, 3, \tag{57}$$

$$\lambda_{ij} = \frac{\widehat{\mathbf{c}}_{vec}^T \mathbf{T}_{ij} \boldsymbol{\Delta}}{2\widetilde{\mathbf{y}}^T \mathbf{R}^{-1} \widetilde{\mathbf{y}}}, \ \ i = 1, 2, \ \ j = i + 1, 3. \tag{58}$$

As such, the Lagrange multipliers are unique, and the constrained minimization problem has a unique stationary solution. The second-order sufficient condition for a minimum [20], which is that the Hessian of the Lagrangian with respect to $\mathbf{K}_c$ at the minimizing solution is positive-definite for all feasible directions of $\mathbf{K}_c$, can be readily verified.

Finally, substituting (52) together with (57) and (58) into (32), the estimator for $\widehat{\mathbf{c}}_{vec}$ becomes

$$\dot{\widehat{\mathbf{c}}}_{vec} = \boldsymbol{\Pi}(\widehat{\mathbf{c}}_{vec}) \boldsymbol{\Delta}, \tag{59}$$

where $\boldsymbol{\Delta}$ is given in (56), and is what the right-hand side of the estimator would be if the constraints were not enforced. The matrix $\boldsymbol{\Pi}(\widehat{\mathbf{c}}_{vec})$ is given by

$$\begin{aligned} \boldsymbol{\Pi}(\widehat{\mathbf{c}}_{vec}) \ \ = \ \ & \Bigg[\mathbf{1}_{9\times 9} - \sum_{i=1}^{3} \mathbf{T}_{ii} \widehat{\mathbf{c}}_{vec} \widehat{\mathbf{c}}_{vec}^T \mathbf{T}_{ii} \\ & - \frac{1}{2} \sum_{i=1}^{2} \sum_{j=i+1}^{3} \mathbf{T}_{ij} \widehat{\mathbf{c}}_{vec} \widehat{\mathbf{c}}_{vec}^T \mathbf{T}_{ij} \Bigg]. \end{aligned} \tag{60}$$

E. Computation of the Constrained Optimal Gain $\mathbf{K}_c^S$

An expression for $\mathbf{K}_c^S$ is required in order to be able to perform the covariance propagation in (38). Substituting the Lagrange multipliers in (57) and (58) into the gain expression in (52) and using (60), leads to

$$\mathbf{K}_c^S = \mathbf{K}_{c,unc} + \frac{(\boldsymbol{\Pi}(\widehat{\mathbf{c}}_{vec}) - \mathbf{1}_{9\times 9})}{\widetilde{\mathbf{y}}^T \mathbf{R}^{-1} \widetilde{\mathbf{y}}} \boldsymbol{\Delta} \widetilde{\mathbf{y}}^T \mathbf{R}^{-1}. \tag{61}$$

F. $SO(3)$-EKF Summary

In summary, the continuous-time $SO(3)$-constrained Extended Kalman filter is given by the state estimate propagation equations (59), (33) with (34), (49), (56) and (60), the covariance propagation equation (38), and the gain equations (54) and (61).

V. Numerical Example

This section presents the results of a numerical simulation comparing the presented MEKF and $SO(3)$-EKF filters. The simulated vehicle has true angular velocity

$$\boldsymbol{\omega}(t) = \begin{bmatrix} 2\sin(0.01t) \\ -3\cos(0.02t) \\ 4 + \sin(0.03t) \end{bmatrix} \text{ deg/s},$$

with initial attitude described by the principal axis and angle of rotation [16] given by $\mathbf{a}(0) = [1, -1, 2]^T / \sqrt{6}$ and $\phi(0) = 10$ degrees, respectively. The gyro bias is $\mathbf{b} = [-0.1, 0.1, 0.05]^T$ deg/s. Two vector measurements are used, with inertial references directions $\mathbf{s}_1^I(t) \equiv [1, 0, 0]^T$ and $\mathbf{s}_2^I(t) \equiv [0, 1, 0]^T$, respectively. The vector measurement noise covariances are $E\{\mathbf{v}_1(t)\mathbf{v}_1^T(t)\} = r_1 \mathbf{1}_{3\times 3}$ and $E\{\mathbf{v}_2(t)\mathbf{v}_1^T(t)\} = r_2 \mathbf{1}_{3\times 3}$, with $E\{\mathbf{v}_1(t)\mathbf{v}_2^T(t)\} = \mathbf{0}_{3\times 3}$ respectively, where $r_1 = 0.25 \text{ deg}^2$ and $r_2 = 0.0025 \text{ deg}^2$. The gyro measurement noise has covariance $E\{\mathbf{w}_\omega(t)\mathbf{w}_\omega^T(t)\} = q_\omega \mathbf{1}_{3\times 3}$, where $q_\omega = 2.5 \times 10^{-5}$ deg^2/s^2. For the purposes of the two filter implementations, the covariance of the random walk input is taken to be $E\{\mathbf{w}_b(t)\mathbf{w}_b^T(t)\} = q_b \mathbf{1}_{3\times 3}$, where $q_b = 10^{-6}$ rad^2/s^4, and it is assumed that $E\{\mathbf{w}_\omega(t)\mathbf{w}_b^T(t)\} = \mathbf{0}_{3\times 3}$. The initial bias estimate for both filters is $\widehat{\mathbf{b}}(0) = \mathbf{0}_{3\times 1}$. The attitude estimate for both filters is initialized using the least-squares solution to Wahba's problem, given the initial measurements $\mathbf{s}_1^b(0)$ and $\mathbf{s}_2^b(0)$, which is [6]

$$\widehat{\mathbf{C}}(0) = \mathbf{V}\mathrm{diag}\{1, 1, \det \mathbf{V} \det \mathbf{U}\}\mathbf{U}^T,$$

where $\mathbf{V}$ and $\mathbf{U}$ satisfy the singular value decomposition

$$\mathbf{V}\boldsymbol{\Sigma}\mathbf{U}^T = \frac{1}{r_1}\mathbf{s}_1^b(0)\mathbf{s}_1^{I,T}(0) + \frac{1}{r_2}\mathbf{s}_2^b(0)\mathbf{s}_2^{I,T}(0),$$

and $\boldsymbol{\Sigma}$ contains the singular values of the matrix on the right. Finally, the initial state error covariance matrices are taken to be $\mathbf{P}^M(0) = \mathrm{diag}\{(0.5\pi/180)^2 \mathbf{1}_{3\times 3}, (0.1\pi/180)^2 \mathbf{1}_{3\times 3}$ for the MEKF, and $\mathbf{P}^S(0) = \mathrm{diag}\{5 \times 10^{-4} \mathbf{1}_{9\times 9}, 5 \times 10^{-7} \mathbf{1}_{3\times 3}$ for the $SO(3)$-EKF.

For clarity of presentation, the estimation errors are presented as follows. The attitude estimation error is represented by the principal angle of rotation, ϕ_{err}, corresponding to the attitude estimation error $\mathbf{C}\widehat{\mathbf{C}}^T$, namely [16]

$$\phi_{err} = \cos^{-1}\left(\frac{\mathrm{trace}(\mathbf{C}\widehat{\mathbf{C}}^T) - 1}{2}\right).$$

The bias estimation error, is represented by its Euclidean norm, namely

$$b_{err} = \|\mathbf{b} - \widehat{\mathbf{b}}\|_2.$$

Figures 1 and 2 show the attitude and bias estimation errors for the two filters. It can be seen that the performances of both filters are very similar, and are almost identical at steady-state.

VI. Concluding Remarks

A new continuous-time Kalman-like filter has been presented for the attitude estimation problem using vector and gyro measurements, called the $SO(3)$-constrained extended Kalman filter ($SO(3)$-EKF). This filter operates directly on $SO(3)$, without the need for any parameterization. The filter has been compared with the continuous-time multiplicative extended Kalman filter (MEKF), which can also operate directly on $SO(3)$. The numerical performances of both filters are very similar, and almost identical at steady-state. This shows that the $SO(3)$-constrained extended Kalman filter is a viable option for estimation on $SO(3)$. The main disadvantage of the $SO(3)$-constrained extended Kalman filter is that the covariance matrix corresponding to the attitude estimation error is 9 by 9, while it is only 3 by 3 for the MEKF. On the other hand, the $SO(3)$-EKF can be directly generalized to $SO(n)$.

References

[1] M.D. Schuster, S.D. Oh, "Three-Axis Attitude Determination from Vector Observations," *AIAA Journal of Guidance, Control and Dynamics*, 4 (1), 1981, pp. 70–77.

[2] F.L. Markley, D. Mortari, "Quaternion Attitude Estimation Using Vector Observations," *Journal of the Astronautical Sciences*, 48 (2-3), 2000, pp. 359–280.

[3] R.H. Wessner, "A Least-Squares Estimate of Satellite Attitude," *SIAM Review*, 8 (3), 1966, pp. 386.

[4] J.E. Brock, "A Least-Squares Estimate of Satellite Attitude," *SIAM Review*, 8 (3), 1966, pp. 386.

[5] K.S. Arun, T.S. Huang, S.D. Blostein, "Least-Squares Fitting of Two Point Sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9 (5), 1987, pp. 698–700.

[6] F.L. Markley, "Attitude Determination Using Vector Observations and the Singular Value Decomposition," *Journal of the Astronautical Sciences*, 36 (3), July-September 1988, pp. 245-258.

[7] K. Kanatani, "Analysis of 3-D Rotation Fitting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16 (5), 1994, pp. 543–549.

[8] E.J. Lefferts, F.L. Markley, M.D. Shuster, "Kalman Filtering for Spacecraft Attitude Estimation", *Journal of Guidance and Control*, 5 (5), 1982. pp. 417–429.

[9] J.L. Crassidis, F.L. Markley, Y. Cheng, "Survey of Nonlinear Attitude Estimation Methods", *Journal of Guidance, Control and Dynamics*, 30 (1), 2007. pp. 12–28.

[10] Shuster, M.D., 1993. "A Survey of Attitude Representations," *The Journal of the Astronautical Sciences*, 41 (4), pp. 439–517.

[11] F.L. Markley, "Attitude Filtering on SO(3)," The Malcolm D. Shuster Symposium, Grand Island, New York, 12-15 June, 2006.

[12] Chaturvedi, N.A., Sanyal, A.K. and McClamroch, N-H., 2011. "Rigid-Body Attitude Control - Using Rotation Matrices for Continuous, Singularity-Free Control Laws," *IEEE Control Systems Magazine*, 31 (3), pp. 30–51.

[13] R. Zanetti, M. Majji, R.H. Bishop, D. Mortari, "Norm Constrained Kalman Filtering,"*AIAA Journal of Guidance, Control and Dynamics*, 32 (5), 2009, pp. 1458–1465.

[14] F.L. Markley, "Attitude Error Representations for Kalman Filtering," *Journal of Guidance, Control and Dynamics*, 26 (2), 2003. pp. 311–317.

[15] A. de Ruiter and J. Forbes, "Continuous-Time SO(n)-Constrained Kalman Filtering," Submitted to *IEEE Transactions on Automatic Control*, 2014.

[16] P.C. Hughes, *Spacecraft Attitude Dynamics,* Dover Publications, New York, 2004.

[17] Farrenkopf, R. L., 1978. "Analytic steady-state accuracy solutions for two common spacecraft attitude estimators," *Journal Guidance and Control*, 1 (4), pp. 282-284.

[18] Crassidis, J.L. and Junkins, J.L., 2011. *Optimal Estimation of Dynamic Systems, 2nd Ed.,'* CRC Press, New York.

[19] D. Simon, *Optimal State Estimation*, John Wiley and Sons, New Jersey, 2006.

[20] J. Nocedal, S.J. Wright, *Numerical Optimization*, Springer-Verlag, New York, 1999.

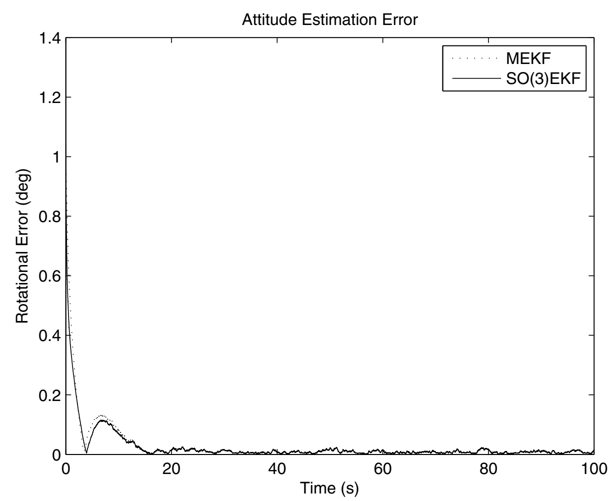

Fig. 1. Attitude Estimation Error

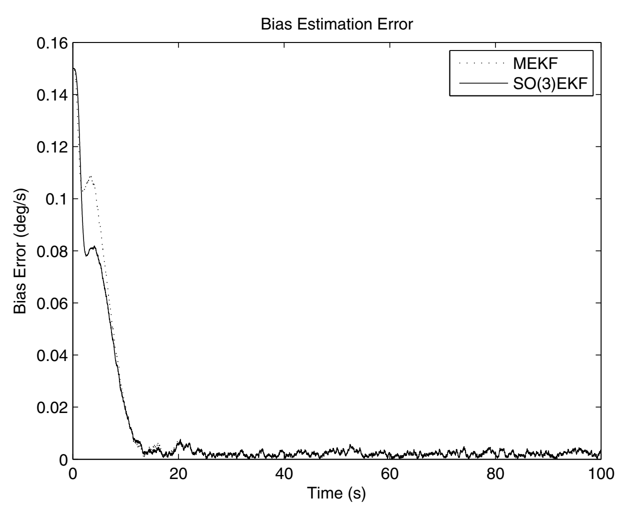

Fig. 2. Bias Estimation Error