Attitude Estimation by Multiplicative eXogenous Kalman Filter ⋆

Bård N. Stovner [a], Tor A. Johansen [b], Thor I. Fossen [b], Ingrid Schjølberg [a]

[a] *Department of Marine Technology, Norwegian University of Science and Technology, Trondheim, Norway*

[b] *Center for Autonomous Marine Operations and Systems (NTNU-AMOS), Department of Engineering Cybernetics, Norwegian University of Science and Technology, Trondheim, Norway*

Abstract

This paper presents a novel attitude estimator called the multiplicative exogenous Kalman filter. The estimator inherits the stability properties of a nonlinear observer and the near-optimal steady-state performance of the linearized Kalman filter for estimation in nonlinear systems. The multiplicative exogenous Kalman filter is derived in detail, and its error dynamics is shown to be globally exponentially stable, which provides guarantees on robustness and transient performance. It is shown in simulations and experiments to yield similar steady-state performance as the multiplicative extended Kalman filter, which is the workhorse for attitude estimation today. The filter assumes biased angular rate measurements and two or more time-varying vector measurements, and it estimates the attitude represented by the quaternion and the angular rate sensor bias.

Key words: Nonlinear observer and filter design; Global exponential stability; Kalman filter

1 Introduction

Estimation of the attitude of a rigid body is an essential part of many navigation systems, whether it is in marine, terrestrial, aerial, or extraterrestrial applications. Solutions typically involve comparing nonparallel vector measurements in the rigid body's body-fixed frame to the corresponding known inertial vectors. Examples of this can be found in satellites, which often navigate by tracking known stars, and in marine, terrestrial, and aerial applications, in which body-fixed measurements of Earth's gravitational and magnetic field are commonly used. The principle behind attitude estimation from vector measurements is the relationship $r^i = R(q_b^i)r^b$ where a unit reference vector r is known in two frames, an inertial frame denoted $\{i\}$ and a body-fixed frame denoted $\{b\}$. The two vectors only differ by the rotation $R(q_b^i)$, here parametrized by the quaternion q_b^i. With at least two nonparallel vector pairs (r_1^i, r_1^b) and (r_2^i, r_2^b), the attitude can be determined.

One of the first attitude estimation algorithms was the TRIAD algorithm presented by Black [4], which finds the rotation matrix explicitly for two nonparallel vector pairs. The main weakness of the TRIAD algorithm is its sensitivity to noise. Bar-Itzhack and Harman [2] improves on this by calculating a weighted average of several different TRIAD solutions, but the achieved estimate is still not optimal in the sense of minimum variance. Wahba [31] posed the problem of finding the rotation matrix R_i^b that minimizes the cost function $J = \sum_{j=1}^{N} \|r_j^b - R_i^b r_j^i\|_2^2$ for N measurements. This problem has received great interest and inspired many solutions. For a review and comparison of several of these solutions, see Markley and Mortari [22].

Attitude estimation often combine vector measurements with angular rate measurements from an angular rate sensor (ARS). These measurements are often corrupted by a biased noise, which requires the estimation of an ARS bias. The workhorse of nonlinear state estimation is the extended Kalman filter (EKF). For attitude estimation, both the additive EKF (AEKF) and the multiplicative EKF (MEKF) has received great attention [20,33,34,17]. The AEKF expresses the quaternion cor-

⋆ This work was partly supported by the Norwegian Research Council (grants no. 234108 and 225259) through the Center of Autonomous Marine Operations and Systems (NTNU-AMOS), grant no. 223254.

Email addresses: bard.b.stovner@ntnu.no (Bård N. Stovner), tor.arne.johansen@itk.ntnu.no (Tor A. Johansen), thor.fossen@ntnu.no (Thor I. Fossen), ingrid.schjolberg@ntnu.no (Ingrid Schjølberg).

rection as an addition, which violates the quaternion norm constraint if implemented naively. The MEKF expresses the correction as a quaternion product, thus maintaining the unit norm constraint. With the correct accomodations, however, Shuster [26] showed that the two methods are identical. In order to gain a unique representation of the attitude error, the four-dimensional multiplicative correction is often mapped to a representation of *minimal degree*, i.e. using three variables to represent the three rotations. Other Kalman filter (KF) solutions to the attitude estimation problem include the unscented KF (UKF) [7], the the invariant EKF (IEKF) [5], and the geometric EKF (GEKF) [1]. These have been shown improve upon the performance of the MEKF.

Nonlinear observers (NLOs) have in recent years received increasing attention for the attitude estimation problem. They often come with global or semi-global stability properties that can be verified a priori, which generally lacks for the EKF-based methods. Salcudean [25] presented an angular velocity observer with global convergence properties. Thienel and Sanner [30] proposed a nonlinear observer with globally exponentially stable (GES) estimation of attitude and bias, provided constant reference vectors, under a persistency-of-excitation (POE) requirement. Later, this requirement was lifted by e.g. Mahoney, Hamel, and Pfimlin [19] and Batista, Silvestre, and Oliveira [3]. Observers for time-varying reference vectors with semi-global stability results were developed by Hua [12] and Grip et al. [10,11]. Batista, Silvestre, and Oliveira [3] presented an attitude observer using a single time-varying and persistently nonconstant reference vector.

In Johansen and Fossen [13], a new way of using the linearized KF is presented, called the exogenous KF (XKF). This idea has been used in Johansen and Fossen [14,15] on the position estimation problem with biased range measurements, in Stovner et al. [28], Jørgensen et al. [16], and Stovner, Johansen, and Schjølberg [29] for underwater position estimation, and in Stovner and Johansen [27] for joint position and attitude estimation. The XKF is explained conceptually in Section 2.1.

Scope and Contribution

In this paper, a novel GES quaternion-based attitude and ARS bias filter is developed. Using a GES NLO and building on the results of Johansen and Fossen [13], the filter is developed and its global stability is proven. Beyond the theoretically guaranteed robustness and transient performance, the MXKF is both in simulations and experiments shown to outperform the NLO and an MEKF with identical tuning; The MXKF is shown to have better steady-state performance than the NLO, while better transient performance than, and at least as good steady-state performance as, the MEKF.

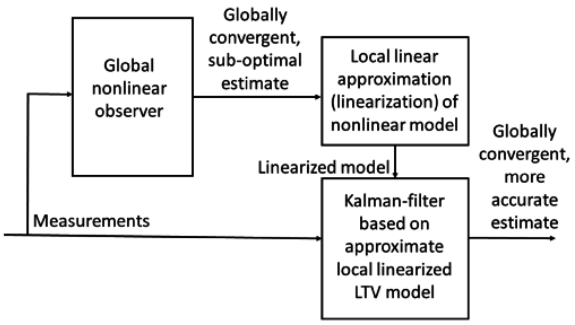

Fig. 1. The general XKF structure [13].

2 Models and Preliminaries

2.1 Exogenous Kalman Filter

The XKF is in many ways similar to the EKF. The EKF linearizes a nonlinear model about its own estimate of the state of the system, and employs the linear time-varying (LTV) Kalman filter (KF) on the linearized model. When the estimate is close to the true state, the linearization approximates the nonlinear model accurately, which often yields near-optimal performance. This, along with its simplicity, largely explains the success of the EKF, which is proven by its extensive use in various applications. A problem with the EKF is the lack of proven stability in the general case. This is caused by the feedback of the state estimate to the linearization, which may increase rather than decrease the estimation error in the correction step. The XKF remediates this by linearizing the nonlinear model about an exogenous signal, thus replacing the potentially destabilizing feedback with a feedforward from an auxiliary estimator. The exogenous signal should be a globally stable, but possibly suboptimal, state estimate. This cascade gains two desired properties: the global stability of the auxiliary estimator and the near-optimality of the linearized KF w.r.t. noise. The general XKF structure is shown in Figure 1.

2.2 State and Error Representation

The north-east-down (NED) frame, denoted $\{n\}$, is assumed to be an inertial frame. This implies that Earth's rotation is neglected and that the vehicle stays within a limited area such that a local flat Earth approximation is accurate. The vehicle's body-fixed (BF) frame, denoted $\{b\}$, is the frame in which the ARS and vector measurements are gathered. The interest of this paper is to estimate the rotation from the BF to NED frame. This attitude estimate defines a third coordinate frame, the estimated BF (EBF) frame, denoted $\{\hat{b}\}$. Lastly, the MXKF relies on an exogenous signal, which is the state estimate of the NLO. The NLO's attitude estimate defines the exogenous BF (XBF) frame, denoted $\{\bar{b}\}$.

In this paper, the attitude is respresented by the unit quaternion $q = [\eta; \varepsilon]$ where η is the scalar part, and $\varepsilon = [\varepsilon_x, \varepsilon_y, \varepsilon_z]^\top$ is the vector part. For column vectors a and b, the ;-notation denotes the vertical concatenation $[a; b] = [a^\top, b^\top]^\top$. For a quaternion q, the corresponding rotation matrix is $R(q) = I + 2\eta S(\varepsilon) + 2S^2(\varepsilon)$ [9] where $S(x)$ is the skew symmetric matrix

$$S(\varepsilon) = \begin{bmatrix} 0 & -\varepsilon_z & \varepsilon_y \\ \varepsilon_z & 0 & -\varepsilon_x \\ -\varepsilon_y & \varepsilon_x & 0 \end{bmatrix} \tag{1}$$

and I is the 3×3 identity matrix. By 0_n and $0_{n\times m}$, we denote $n \times n$ and $n \times m$ matrices of zeros, respectively.

For the coordinate frames denoted a, b, and c, a rotation from c to a can be described by $R(q_c^a) = R(q_b^a)R(q_c^b) = R(q_b^a \otimes q_c^b)$, where $\otimes$ denotes the Hamilton product

$$q_1 \otimes q_2 = \begin{bmatrix} \eta_1\eta_2 - \varepsilon_1^\top \varepsilon_2 \\ \eta_1\varepsilon_2 + \eta_2\varepsilon_1 - S(\varepsilon_1)\varepsilon_2 \end{bmatrix} = q_1\eta_2 + \Psi(q_1)\varepsilon_2 \tag{2}$$

and $\Psi(q_1) = [-\varepsilon_1^\top; \eta_1 I - S(\varepsilon_1)]$ for $q_1 = [\eta_1; \varepsilon_1]$ and $q_2 = [\eta_2; \varepsilon_2]$. The quaternions representing the rotations from the BF, XBF, and EBF frames to the NED frame are denoted q_b^n, $q_{\check{b}}^n$, and $q_{\hat{b}}^n$, respectively. The true ARS bias is denoted b^b when it is decomposed in the BF frame. When decomposed in their respective frames, the MXKF's and NLO's bias estimates are denoted $\hat{b}^{\hat{b}}$ and $\bar{b}^{\bar{b}}$, respectively.

The true attitude q_b^n and the estimates of it, $q_{\hat{b}}^n$ and $q_{\bar{b}}^n$, can be expressed by their scalar (η) and vector (ε) parts $q_b^n = [\eta; \varepsilon]$, $q_{\hat{b}}^n = [\hat{\eta}; \hat{\varepsilon}]$, and $q_{\bar{b}}^n = [\bar{\eta}; \bar{\varepsilon}]$, respectively. We define the *additive quaternion estimation error* and the bias estimation error of the MXKF as

$$\tilde{q} \triangleq q_b^n - q_{\hat{b}}^n \tag{3a}$$
$$\tilde{b} \triangleq b^b - \hat{b}^{\hat{b}} \tag{3b}$$

respectively. Let the true state and the MXKF's and NLO's state estimates be given by $z = [q_b^n; b^b]$, $\hat{z} = [q_{\hat{b}}^n; \hat{b}^{\hat{b}}]$, and $\bar{z} = [q_{\bar{b}}^n; \bar{b}^{\bar{b}}]$, respectively. Finally, the *additive estimation error* of the MXKF is defined as $\tilde{z} \triangleq z - \hat{z}$.

Next, we define the *multiplicative quaternion estimation error* of the MXKF as

$$\delta q = \begin{bmatrix} \delta\eta \\ \delta\varepsilon \end{bmatrix} \triangleq (q_{\hat{b}}^n)^{-1} \otimes q_b^n = q_b^{\hat{b}} \tag{4}$$

where $(q_{\hat{b}}^n)^{-1} = q_n^{\hat{b}} = [\hat{\eta}; -\hat{\varepsilon}]$ is the unit quaternion inverse. Notice that definition (4) differs from the more commonly used multiplicative quaternion error definition $\delta q' \triangleq q_b^n \otimes (q_{\hat{b}}^n)^{-1}$ [21]. This is done because $\delta q = q_b^{\hat{b}}$ directly represents the goal of minimization, i.e. the difference between the EBF and BF frames. It is argued that using $\delta q = q_b^{\hat{b}}$ is more intuitive than using $\delta q'$, though it should neither influence the stability nor the performance.

Representing three rotations by a 4-dimensional vector is ambiguous and gives multiple solutions. Therefore, the 3-dimensional vector representation called the Modified Rodrigues Parameter (MRP) is used. It is defined as [21]

$$\delta u \triangleq \frac{\delta\varepsilon}{1 + \delta\eta} \tag{5}$$

and is widely used in MEKFs because it is a three-parameter representation of the attitude that is only singular when $\delta\eta = q_{\hat{b}}^{n\top} q_b^n = -1$, i.e. when $q_{\hat{b}}^n = -q_b^n$. This corresponds to a 360° estimation error. Therefore, by ensuring that the absolute estimation error is always less than 360°, i.e. $\delta\eta > -1$, one achieves a globally non-singular attitude error representation. Also, as will be shown later, the MRP relates linearly to the additive quaternion error $\tilde{q}$. These two properties motivate the choice of the MRP as the attitude error representation of the MXKF. Lastly, we define the *multiplicative estimation error* of the MXKF as $\delta x \triangleq [\delta u; \tilde{b}]$.

2.3 Kinematic Model

The kinematics of the quaternion and bias are [21]

$$\dot{q}_b^n = f_q(z,t) = \frac{1}{2}\Xi(q_b^n)\omega_{nb}^b = \frac{1}{2}\Omega(\omega_{nb}^b)q_b^n \tag{6a}$$
$$\dot{b}^b = f_b(t) = \epsilon_b \tag{6b}$$

where ω_{nb}^b is the true angular rate, b^b is the slowly-varying ARS bias, ϵ_b is a random process, and

$$\Xi(q) = \begin{bmatrix} -\varepsilon^\top \\ \eta I + S(\varepsilon) \end{bmatrix}, \Omega(\omega_{nb}^b) = \begin{bmatrix} 0 & -{\omega_{nb}^b}^\top \\ \omega_{nb}^b & -S(\omega_{nb}^b) \end{bmatrix} \tag{7}$$

We put the full state dynamics (6) in the compact form

$$\dot{z} = f_z(z,t) = \begin{bmatrix} f_q(z,t) \\ f_b(t) \end{bmatrix} \tag{8}$$

2.4 Measurement Equation

The measurement equation is

$$y_j^b = h_j(z, y_j^n) + \epsilon_{yj} = R^\top(q_b^n) y_j^n + \epsilon_{yj} \tag{9}$$

for vector measurements $j = 1, ..., M$, where ϵ_{yj} is the noise on measurement j. We concatenate (9) for all j by writing $y^b = [y_1^b; ...; y_M^b]$, $y^n = [y_1^n; ...; y_M^n]$, $\epsilon_y = [\epsilon_{y1}; ...; \epsilon_{yM}]$, and $h(z, y^n) = [h_1(z, y_1^n); ...; h_M(z, y_M^n)]$ which yields

$$y^b = h(z, y^n) + \epsilon_y \tag{10}$$

As mentioned in Section 1, two nonparallel vector measurements are needed to determine the attitude at an instant. This is stated by the following assumption.

Assumption 1 *At least two reference vectors are nonparallel, i.e., there exists a positive constant γ such that $|y_1^n \times y_2^n| \geq \gamma > 0$.*

The above assumption is sufficient for observability in continuous time, but could nevertheless be relaxed. In particular for a discrete-time implementation of the nonlinear observer and linearized KF, one can use standard modifications to the gain and covariance updates to handle multirate sampling and missing observations, see e.g. Farrell [8].

The angular rate sensor provides the measurement $\omega_{nb,m}^b = \omega_{nb}^b + b^b + \epsilon_\omega$ where ϵ_ω is a random process. Now, we define the process noise vector $\epsilon_z \triangleq [\epsilon_\omega; \epsilon_b]$.

3 Multiplicative eXogenous Kalman Filter

3.1 Nonlinear Observer

Figure 2 shows the role of the NLO in the structure of the MXKF. Any NLO estimating the attitude and ARS bias with strong convergence properties can be used. Here, the attitude observer from Grip et al. [11] is used, i.e.

$$\dot{\bar{R}}_{\bar{b}}^n = \bar{R}_{\bar{b}}^n S(\omega_{nb,m}^b - \bar{b}^{\bar{b}}) + \sigma K_P J(\bar{R}_{\bar{b}}^n, t) \tag{11a}$$

$$\dot{\bar{b}}^{\bar{b}} = \text{Proj}(\bar{b}^{\bar{b}}, -k_I \text{vex}(\mathbb{P}(\bar{R}_{\bar{b}}^{n\top} K_P J(\bar{R}_{\bar{b}}^n, t)))) \tag{11b}$$

$$J(R_b^n, t) = \sum_{j=1}^{3} (w_j^n - R_b^n w_j^b) w_j^{b\top} \tag{11c}$$

$$w_1^\iota = \frac{y_1^\iota}{\|y_1^\iota\|_2}, w_2^\iota = \frac{S(y_1^\iota) y_2^\iota}{\|S(y_1^\iota) y_2^\iota\|_2}, w_3^\iota = \frac{S(y_1^\iota)^2 y_2^\iota}{\|S(y_1^\iota)^2 y_2^\iota\|_2} \tag{11d}$$

where $\iota \in (n, b)$, K_P is a symmetric positive-definite gain matrix, k_I is a strictly positive scalar gain, $\sigma \geq 1$ is a stability tuning factor, $\bar{R}_{\bar{b}}^n s$ is the matrix $\bar{R}_{\bar{b}}^n$ with all its elements saturated between ± 1, Proj is a projection function that ensures $\|\bar{b}^{\bar{b}}\| \leq \bar{M}$ for $\bar{M} > M$ where M is an a priori known upper bound on the ARS bias, i.e. $\|b^b\| < M$, $\mathbb{P}(X) = 0.5(X + X^\top)$ for any square matrix X, and $\text{vex}(S(x)) = x$. Note that $\bar{R}_{\bar{b}}^n$ is not always on SO(3), but it *converges to* SO(3). When $\bar{R}_{\bar{b}}^n$ is projected onto SO(3), the result is denoted $R_{\bar{b}}^n$.

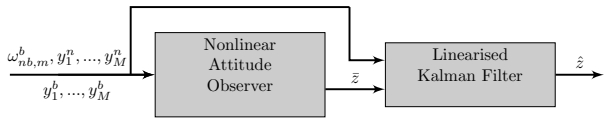

Fig. 2. The structure of the MXKF

Denote by Σ_1 the dynamics of the estimation error $R(q_b^n) - \bar{R}_{\bar{b}}^n$ and $b^b - \bar{b}^{\bar{b}}$. Let $q_{\bar{b}}^n$ be extracted from $R_{\bar{b}}^n$ in such a way that it forms a continuous signal, and does not jump between the two representation $q_{\bar{b}}^n$ and $-q_{\bar{b}}^n$.

Proposition 2 *The origin $R(q_b^n) - \bar{R}_{\bar{b}}^n = 0_3$ and $b^b - \bar{b}^{\bar{b}} = 0$ of Σ_1 is GES.*

PROOF. Grip et al. [11] □

3.2 Linearized Kalman Filter

In this section, the linearized KF of the MXKF is derived. In Section 3.2.2, the dynamics and measurement equations, expressed by the quaternion, are linearized. The error dynamics of the linearized system expressed by the quaternion is formed, which is of seventh order. In Section 3.2.3, the seventh order error dynamics is transformed to a sixth order representation. The linearized KF is formulated with this representation. This makes the KF less computationally demanding and makes it possible to show stability of the system with straightforward time-varying observability analysis. For convenience, the analysis is conducted in continuous time.

The internal structure of the linearized KF part of the MXKF is shown in Figure 3.

When the attitude is correctly estimated, i.e. $q_{\bar{b}}^n = \pm q_b^n$, this leaves two possible additive quaternion errors $\tilde{q} = 0$ and $\tilde{q} = 2q_b^n$. Below, we resolve this amgibuity.

Definition 3 *Define $q_b^{n\star}$ as the one of q_b^n and $-q_b^n$ that is closest to $q_{\bar{b}}^n$, i.e.*

$$q_b^{n\star} \triangleq \begin{cases} q_b^n, & \text{if } q_{\bar{b}}^{n\top} q_b^n \geq 0 \\ -q_b^n, & \text{if } q_{\bar{b}}^{n\top} q_b^n < 0 \end{cases} \tag{12}$$

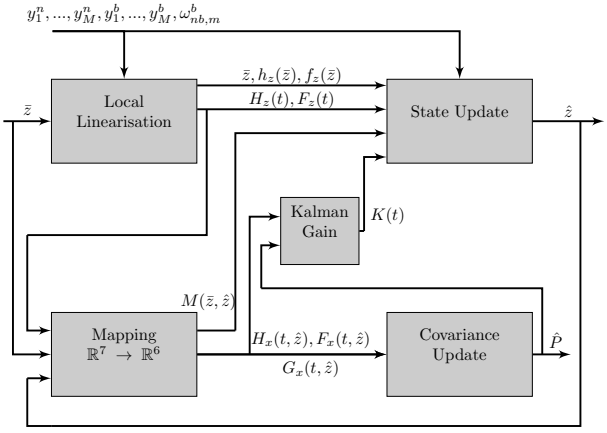

Fig. 3. A closer look at the linearized KF part of the MXKF. The reset operation (26) is excluded for ease of presentation.

Now, we must also redefine $\tilde{q} \triangleq q_b^{n\star} - q_{\hat{b}}^n$, $z \triangleq [q_b^{n\star}; b^b]$, and $\delta q \triangleq (q_{\hat{b}}^n)^{-1} \otimes q_b^{n\star}$.

3.2.1 Mapping from Additive to Multiplicative Error State

Lemma 4 *The additive error state $\tilde{z}$ can be mapped to the multiplicative error state δx by the mapping*

$$\tilde{z} = M(z, \hat{z})\delta x = (M(\bar{z}, \hat{z}) + \tilde{M}(z, \bar{z}))\delta x \tag{13}$$

where

$$M(\bar{z}, \hat{z}) = \begin{bmatrix} \Psi(q_{\bar{b}}^n) + \Psi(q_{\hat{b}}^n) & 0_{4\times 3} \\ 0_3 & I_3 \end{bmatrix} \tag{14}$$

$$\tilde{M}(z, \bar{z}) = \begin{bmatrix} \Psi(q_b^{n\star}) - \Psi(q_{\bar{b}}^n) & 0_{4\times 3} \\ 0_3 & 0_3 \end{bmatrix} \tag{15}$$

Moreover, the mapping approximation converges, i.e. $M(\bar{z}, \hat{z}) \rightarrow M(z, \hat{z})$, as the NLO converges.

PROOF. In order to verify the mapping $\tilde{z} = M(z, \hat{z})\delta x$, we insert (4) in (3a) twice to get $\tilde{q} = q_b^{n\star} - q_{\hat{b}}^n = q_{\hat{b}}^n \otimes \delta q - q_b^{n\star} \otimes (\delta q)^{-1}$. Further, we apply (2) twice to get

$$\tilde{q} = q_{\hat{b}}^n \delta\eta + \Psi(q_{\hat{b}}^n)\delta\varepsilon - q_b^{n\star}\delta\eta + \Psi(q_b^{n\star})\delta\varepsilon \tag{16}$$

$$(1 + \delta\eta)\tilde{q} = (\Psi(q_{\hat{b}}^n) + \Psi(q_b^{n\star}))\delta\varepsilon \tag{17}$$

$$\tilde{q} = (\Psi(q_{\hat{b}}^n) + \Psi(q_b^{n\star}))\delta u \tag{18}$$

Noticing that $\tilde{q} = (\Psi(q_{\hat{b}}^n + q_{\bar{b}}^n) + \Psi(q_b^{n\star} - q_{\bar{b}}^n))\delta u$ verifies the mapping using (13) and (15). From Proposition 2, we know that $R(q_b^n) - R(q_{\bar{b}}^n) \rightarrow 0$, which together with (12) means that $q_{\bar{b}}^n \rightarrow q_b^{n\star}$. This proves that $\tilde{M}(z, \bar{z}) \rightarrow 0$ exponentially. □

3.2.2 Additive Estimation Error Model

As discussed in Section 2.1, the model given by (8) and (10) is linearized about the exogenous signal $\bar{z}$

$$\begin{aligned} f_z(z,t) =& f_z(\bar{z},t) + F_z(t)(z - \bar{z}) \\ &+ G_z(t)\epsilon_z + \chi_z(t) \end{aligned} \tag{19a}$$

$$h_z(z,t) = h_z(\bar{z},t) + H_z(t)(z - \bar{z}) + \epsilon_y + \chi_y(t) \tag{19b}$$

where $\chi_z(t)$ and $\chi_y(t)$ are the linearization errors, i.e., higher order terms in the linearization. From (6a) and (6b), we see that $F_z(t)$ and $G_z(t)$ become

$$F_z(t) = \left.\frac{\mathrm{d}f_z(z,t)}{\mathrm{d}z}\right|_{z=\bar{z}} = \frac{1}{2}\begin{bmatrix} \Omega(\omega_{nb,m}^b - \bar{b}^{\bar{b}}) & -\Xi(q_{\bar{b}}^n) \\ 0_{3\times 4} & 0_3 \end{bmatrix} \tag{20}$$

$$G_z(t) = \left.\frac{\mathrm{d}f_z(z,t)}{\mathrm{d}\epsilon_z}\right|_{z=\bar{z}} = \begin{bmatrix} -\frac{1}{2}\Xi(q_{\bar{b}}^n) & 0_{4\times 3} \\ 0_3 & I_3 \end{bmatrix} \tag{21}$$

Using equation (A.44) from Markley and Crassidis [21], we find $H_z(\bar{z})$ to be

$$H_z(t) = \begin{bmatrix} 2S(y_1^{\bar{b}})\Xi^\top(q_{\bar{b}}^n) & 0_3 \\ \vdots & \\ 2S(y_M^{\bar{b}})\Xi^\top(q_{\bar{b}}^n) & 0_3 \end{bmatrix} \tag{22}$$

where $y_j^{\bar{b}} = R^\top(q_{\bar{b}}^n)y_j^n$. Notice that the dependency of known bounded time-varying signals have been replaced by dependency of time, t, in $F_z(t)$, $G_z(t)$, $H_z(t)$, $\chi_z(t)$, and $\chi_y(t)$.

We define an estimator

$$\begin{aligned} \dot{\hat{z}} \triangleq& f_z(\bar{z},t) + F_z(t)(\hat{z} - \bar{z}) \\ &+ M(\bar{z},\hat{z})K(t)(y - \hat{y}) \end{aligned} \tag{23a}$$

$$\hat{y} \triangleq h_z(\bar{z},t) + H_z(t)(\hat{z} - \bar{z}) \tag{23b}$$

where the gain matrix $K(t)$ is specified later.

Define $\tilde{y} \triangleq y - \hat{y}$. Now, subtracting (23) from (19) yields the additive estimation error model

$$\dot{\tilde{z}} = F_z(t)\tilde{z} - M(\bar{z},\hat{z})K(t)\tilde{y} + G_z(t)\epsilon_z + \chi_z(t) \tag{24a}$$

$$\tilde{y} = H_z(t)\tilde{z} + \epsilon_y + \chi_y(t) \tag{24b}$$

Inserting (24b) into (24a) yields the additive error dynamics

$$\begin{aligned} \Sigma_A : \dot{\tilde{z}} =& \left(F_z(t) - M(\bar{z},\hat{z})K(t)H_z(t)\right)\tilde{z} + G_z(t)\epsilon_z \\ &+ \chi_z(t) - M(\bar{z},\hat{z})K(t)\left(\epsilon_y + \chi_y(t)\right) \end{aligned} \tag{25}$$

3.2.3 Transformation from Additive to Multiplicative Error Dynamics

In this section, we transform the error dynamics from $\mathbb{R}^7$ to $\mathbb{R}^6$. We begin by defining $\delta\check{q} = [\delta\check{\eta}; \delta\bar{\varepsilon}] \triangleq (q_{\hat{b}}^n)^{-1} \otimes q_{\bar{b}}^n$ and $\delta\bar{q} = [\delta\bar{\eta}; \delta\bar{\varepsilon}] \triangleq (q_{\bar{b}}^n)^{-1} \otimes q_b^{n\star}$. Later, it will be required that $q_{\hat{b}}^n \neq -q_{\bar{b}}^n$, and from (5) we must have $\delta\eta > -1$. This is guaranteed by the following Lemma.

Lemma 5 *The reset rule*

$$\hat{z} \leftarrow \bar{z}, \ \textit{if}\ q_{\hat{b}}^{n\top} q_{\bar{b}}^n \leq \epsilon \tag{26}$$

guarantees both that $q_{\hat{b}}^n \neq -q_{\bar{b}}^n$ *and* $\delta\eta > -1 + \Delta(\epsilon)$ *where* $\Delta(\epsilon) > 0$ *for an input* $\epsilon > 0$.

PROOF. Define $\delta\bar{q} = [\delta\bar{\eta}; \delta\bar{\varepsilon}] \triangleq (q_{\bar{b}}^n)^{-1} \otimes q_b^{n\star}$, where by (12) we have that $\delta\bar{\eta} \geq 0$. From $\delta q = \delta\check{q} \otimes \delta\bar{q}$, we get $\delta\eta = \delta\bar{\eta}\delta\check{\eta} - \delta\bar{\varepsilon}^\top\delta\check{\varepsilon}$. Since $\delta\bar{\eta} \geq 0$ from (12) and $\delta\check{\eta} > \epsilon$ and $\|\delta\check{\varepsilon}\|_2 = \sqrt{1-\delta\check{\eta}} < \sqrt{1-\epsilon}$ from (26), we have that $\delta\eta \geq -\delta\bar{\varepsilon}^\top\delta\check{\varepsilon} \geq -\|\delta\bar{\varepsilon}\|_2\|\delta\check{\varepsilon}\|_2 > -\sqrt{1-\epsilon} = -1 + \Delta(\epsilon)$, where $\Delta(\epsilon) > 0$ for $\epsilon > 0$. □

Lemma 6 *The seventh order additive error dynamics* Σ_A *in* (25) *can be transformed by the approximate linear mapping* $M(\bar{z}, \hat{z})$ *to the sixth order multiplicative error dynamics*

$$\begin{aligned}\Sigma_2 : \delta\dot{x} =& (F_x(\hat{z},t) - K(t)H_x(\hat{z},t))\delta x + \chi_x(\hat{z},t) \\ &+ G_x(\hat{z},t)\epsilon_z - K(t)\epsilon_y + \zeta(\hat{z},\delta x,t)\end{aligned} \tag{27}$$

where

$$F_x(\hat{z},t) \triangleq M^\dagger(\bar{z},\hat{z})(F_z(t)M(\bar{z},\hat{z}) - \dot{M}(\bar{z},\hat{z})) \tag{28a}$$
$$G_x(\hat{z},t) \triangleq M^\dagger(\bar{z},\hat{z})G_z(t) \tag{28b}$$
$$\chi_x(\hat{z},t) \triangleq M^\dagger(\bar{z},\hat{z})\chi_z(t) - K(t)\chi_y(t) \tag{28c}$$
$$H_x(\hat{z},t) \triangleq H_z(t)M(\bar{z},\hat{z}) \tag{28d}$$

and $(\cdot)^\dagger$ *denotes the Moore-Penrose pseudo-inverse operation* $C^\dagger = (C^\top C)^{-1}C^\top$. *Furthermore,*

$$\chi_x(\hat{z},t) \leq k_\chi ||z - \bar{z}||_2^2 \tag{29a}$$
$$\zeta(\hat{z},\delta x,t) \leq k_\zeta \|z - \bar{z}\|_2^2 \tag{29b}$$

PROOF. Differentiating (13) w.r.t. time yields

$$\dot{\tilde{z}} = \dot{M}(\bar{z},\hat{z})\delta x + M(\bar{z},\hat{z})\delta\dot{x} + \dot{\tilde{M}}(z,\bar{z},t)\delta x + \tilde{M}(z,\bar{z})\delta\dot{x} \tag{30}$$

$$\delta\dot{x} = M^\dagger(\bar{z},\hat{z})(\dot{\tilde{z}} - \dot{M}(\bar{z},\hat{z})\delta x - \dot{\tilde{M}}(z,\bar{z},t)\delta x - \tilde{M}(z,\bar{z})\delta\dot{x}) \tag{31}$$

Defining $M_1 \triangleq M_1(\hat{z},t) = I + M^\dagger(\bar{z},\hat{z})\tilde{M}(z,\bar{z})$ and collecting $\delta\dot{x}$ yields

$$\delta\dot{x} = M_1^{-1}M^\dagger(\bar{z},\hat{z})\left(\dot{\tilde{z}} - \dot{M}(\bar{z},\hat{z})\delta x - \dot{\tilde{M}}(z,\bar{z},t)\delta x\right). \tag{32}$$

It is straightforward to show that

$$\det(M_1) = \frac{1 + \delta\eta + \delta\bar{\eta} + \delta\check{\eta}}{2(1+\delta\check{\eta})}\det\left(I + \frac{S(\delta\varepsilon + \delta\bar{\varepsilon} + \delta\check{\varepsilon}))}{(1 + \delta\eta + \delta\bar{\eta} + \delta\check{\eta})}\right) \tag{33}$$

Since $1 + \delta\eta + \delta\bar{\eta} + \delta\check{\eta} > \epsilon + \Delta(\epsilon)$ and $x^\top(I + A)x = x^\top x > 0 \ \forall\ x \neq 0$ for any skew-symmetric matrix A, $\det(M_1) \neq 0$ which means that M_1 is invertible. By the Sherman-Morrison-Woodbury formula [32], we have that $M_1^{-1} = I - M_2$, where $M_2 \triangleq M_2(\hat{z},t) = (I + M^\dagger(\bar{z},\hat{z})\tilde{M}(z,\bar{z}))^{-1}M^\dagger(\bar{z},\hat{z})\tilde{M}(z,\bar{z})$. Now, inserting (28) and expanding M_1^{-1}, we can rewrite (32) as

$$\begin{aligned}\delta\dot{x} =& M_1^{-1}\Big((F_x(\hat{z},t) - K(t)H_x(\hat{z},t))\delta x + \chi_x(\hat{z},t) && (34)\\ &+ G_x(\hat{z},t)\epsilon_z - K(t)\epsilon_y - \dot{\tilde{M}}(z,\bar{z},t)\delta x\Big) && (35)\\ \delta\dot{x} =& (F_x(\hat{z},t) - K(t)H_x(\hat{z},t))\delta x + \chi_x(\hat{z},t) && (36)\\ &+ G_x(\hat{z},t)\epsilon_z - K(t)\epsilon_y + \zeta(\hat{z},\delta x,t) && (37)\end{aligned}$$

where $\zeta(\hat{z},\delta x,t) \triangleq -M_2((F_x(\hat{z},t) - K(t)H_x(\hat{z},t))\delta x + \chi_x(\hat{z},t) + G_x(\hat{z},t)\epsilon_z - K(t)\epsilon_y) - M_1^{-1}\dot{\tilde{M}}(z,\bar{z},t)\delta x$. Using that $\Psi^\top(q_{\bar{b}}^n)\Psi(q_{\hat{b}}^n) = I\delta\check{\eta} - S(\delta\bar{\varepsilon})$, we find

$$(M^\top(\bar{z},\hat{z})M(\bar{z},\hat{z}))^{-1} = \begin{bmatrix}\frac{1}{2(1+\delta\check{\eta})}I & 0\\ 0 & I\end{bmatrix} \tag{38}$$

and we see that $M^\dagger(\bar{z},\hat{z})$ exists whenever $\delta\check{\eta} = q_{\hat{b}}^{n\top}q_{\bar{b}}^n \neq -1$, i.e., when $q_{\hat{b}}^n \neq -q_{\bar{b}}^n$, which is guaranteed by Lemma 5.

Due to the boundedness of $\bar{z}$ and $q_{\hat{b}}^n$ and the smoothness of f_z, h_z, G_z, and $\tilde{M}$, there exist constants $k_\zeta, k_\chi > 0$ such that $\chi_x(\hat{z},t) < k_\chi||z - \bar{z}||_2^2$, $\zeta(\hat{z},\delta x,t) < k_\zeta\|z - \bar{z}\|_2^2$. □

In the following, explicit expressions for the matrices (28a), (28b), and (28d) are derived. From Markley and Crassidis [21] we have that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Psi(q_{\hat{b}}^n) = \frac{1}{2}\Omega(\omega_{nb,m}^b - \hat{b}^{\hat{b}})\Psi(q_{\hat{b}}^n) \tag{39}$$

and similarly for $\Psi(q^n_{\bar{b}})$. This is used in order to find

$$F_x(\hat{z},t) = \begin{bmatrix} F_1(\hat{z},t) & F_2(\hat{z},t) \\ 0_3 & 0_3 \end{bmatrix} \tag{40}$$

$$F_1(\hat{z},t) = \frac{\Psi^\top(q^n_{\bar{b}}) + \Psi^\top(q^n_{\hat{b}})}{4(1+\delta\check{\eta})}\Omega(\hat{b}^{\hat{b}} - \bar{b}^{\bar{b}})\Psi(q^n_{\hat{b}}) \tag{41}$$

$$F_2(\hat{z},t) = -\frac{\Psi^\top(q^n_{\bar{b}}) + \Psi^\top(q^n_{\hat{b}})}{4(1+\delta\check{\eta})}\Xi(q^n_{\bar{b}}) = -\frac{1}{4}D_1 R(q^n_{\bar{b}}) \tag{42}$$

Here, $\Psi^\top(q^n_{\bar{b}})\Xi(q^n_{\bar{b}}) = R(q^n_{\bar{b}})$, $\Xi(q^n_{\bar{b}}) = \Psi(q^n_{\bar{b}})R(q^n_{\bar{b}})$, and $\Psi^\top(q^n_{\bar{b}})\Psi(q^n_{\bar{b}}) = I\delta\check{\eta} - S(\delta\check{\varepsilon})$ have been used to find the useful identities

$$(\Psi^\top(q^n_{\bar{b}}) + \Psi^\top(q^n_{\hat{b}}))\Xi(q^n_{\bar{b}}) = DR(q^n_{\bar{b}}) \tag{43}$$

$$D \triangleq D(\hat{z},t) = (1+\delta\check{\eta})I - S(\delta\check{\varepsilon}) \tag{44}$$

$$D_1 \triangleq D_1(\hat{z},t) = \frac{1}{1+\delta\check{\eta}}D = I - S(\delta\check{u}) \tag{45}$$

where $\delta\check{u} = \delta\check{\varepsilon}/(1+\delta\check{\eta})$. $G_x(\hat{z},t)$ and $H_x(\hat{z},t)$ are found to be

$$G_x(\hat{z},t) = \begin{bmatrix} -\frac{1}{4}D_1 R(q^n_{\bar{b}}) & 0_{4\times 3} \\ 0_3 & I_3 \end{bmatrix} \tag{46}$$

$$H_x(\hat{z},t) = \begin{bmatrix} 2R^\top(q^n_{\bar{b}})S(y^n_1)D^\top & 0_3 \\ \vdots & \vdots \\ 2R^\top(q^n_{\bar{b}})S(y^n_M)D^\top & 0_3 \end{bmatrix} \tag{47}$$

and $S(y^{\bar{b}}_i) = S(R^\top(q^n_{\bar{b}})y^n_i) = R^\top(q^n_{\bar{b}})S(y^n_i)R(q^n_{\bar{b}})$ has been used to find $H_x(\hat{z},t)$.

$K(t)$ introduced in (23a) is the solution of the Riccati equation with $F_x(\hat{z},t)$, $H_x(\hat{z},t)$, and $G_x(\hat{z},t)$, i.e.

$$K(t) = P(t)H_x^\top(\hat{z},t)\mathcal{R}^{-1}(t) \tag{48a}$$

$$\dot{P}(t) = F_x(\hat{z},t)P(t) + P(t)F_x^\top(\hat{z},t) + G_x(\hat{z},t)\mathcal{Q}(t)G_x^\top(\hat{z},t) - P(t)H_x^\top(\hat{z},t)\mathcal{R}^{-1}(t)H_x(\hat{z},t)P(t) \tag{48b}$$

where $\mathcal{Q}(t) = \mathrm{E}(\epsilon_z\epsilon_z^\top)$, $\mathcal{R}(t) = \mathrm{E}(\epsilon_y\epsilon_y^\top)$, and $P(0) = \mathrm{E}(\delta x(0)\delta x(0)^\top)$ are covariance matrices.

3.2.4 Stability Analysis

Lemma 7 *The pair $(F_x(\hat{z},t), H_x(\hat{z},t))$ is uniformly completely observable (UCO) and $(F_x(\hat{z},t), G_x(\hat{z},t))$ is uniformly completely controllable (UCC).*

PROOF. First, we show observability of the pair $(F_x(\hat{z},t), H_x(\hat{z},t))$ by employing Theorem 6.O12 in Chen [6], which states that the pair $(F_x(\hat{z},t), H_x(\hat{z},t))$ is observable if and only if the observability codistribution $\mathrm{d}\mathcal{O} = [\mathcal{M}_0(\hat{z},t); ...; \mathcal{M}_{n-1}(\hat{z},t)]$ has full rank for all t, where $\mathcal{M}_0(\hat{z},t) = H_x(\hat{z},t)$, $\mathcal{M}_m(\hat{z},t) = \mathcal{M}_{m-1}(\hat{z},t)F_x(\hat{z},t) + (\mathrm{d}/\mathrm{d}t)\mathcal{M}_{m-1}(\hat{z},t)$ for $m = 1, ..., n-1$, and n is the dimension of the state space, i.e. $n = 6$. Since both $H_x(\hat{z},t)$ and $F_x(\hat{z},t)$ are continuously differentiable, the matrix $\mathrm{d}\mathcal{O}$ can be formed. It suffices to examine the column rank of $\mathrm{d}\mathcal{O}_1 = [\mathcal{M}_0(t); \mathcal{M}_1(t)]$ because if it has full column rank, $\mathrm{d}\mathcal{O}$ must have full rank. We find $\mathrm{d}\mathcal{O}_1$ to be

$$\mathrm{d}\mathcal{O}_1 = \begin{bmatrix} \mathcal{O}_1 & 0_{3M\times 3} \\ \mathcal{O}_2 & \mathcal{O}_3 \end{bmatrix} \tag{49}$$

$$\mathcal{O}_1 = 2(1+\delta\check{\eta})R^\top(q^n_{\bar{b}})S_{1:M}D_1^\top \tag{50}$$

$$\mathcal{O}_3 = -\frac{1+\delta\check{\eta}}{2}R^\top(q^n_{\bar{b}})S_{1:M}D_1^\top D_1 R(q^n_{\bar{b}}) \tag{51}$$

where $S_{1:M} = [S(y^n_1); ...; S(y^n_M)]$. In order to verify that $\mathrm{d}\mathcal{O}_1$ has full rank, it suffices to verify that $\mathcal{O}_1$ and $\mathcal{O}_3$ have full rank, as proven by Meyer [23]. Therefore, we need not find $\mathcal{O}_2$. Through the identity $\mathrm{rank}(\mathcal{O}) = \mathrm{rank}(\mathcal{O}^\top\mathcal{O})$ for any matrix $\mathcal{O}$, we examine the ranks of $\mathcal{O}_1$ and $\mathcal{O}_3$:

$$\mathrm{rank}(\mathcal{O}_1) = \mathrm{rank}\Big(-4(1+\delta\check{\eta})^2 D_1 S_{1:M}^\top S_{1:M} D_1^\top\Big) \tag{52}$$

$$\mathrm{rank}(\mathcal{O}_3) = \mathrm{rank}\Big(-\frac{(1+\delta\check{\eta})^2}{4}R^\top(q^n_{\bar{b}})D_1^\top D_1 S_{1:M}^\top S_{1:M} D_1^\top D_1 R(q^n_{\bar{b}})\Big) \tag{53}$$

The determinant of D_1 is $\det(D_1(\hat{z},t)) = 1 + \delta\check{u}^\top\delta\check{u} > 0$, meaning that D_1 always has full rank. Under Assumption 1, $S_{1:M}^\top S_{1:M} = \sum_{j=1}^M S^2(y^n_j)$ also has full rank. Since the product of square full rank matrices has full rank and $\delta\check{\eta} > \epsilon$ as guaranteed by (26), we know that both $\mathcal{O}_1$ and $\mathcal{O}_3$ has full rank. Therefore, we know that $\mathrm{rank}(\mathrm{d}\mathcal{O}_1) = 6$, which proves that the observability codistribution $\mathrm{d}\mathcal{O}$ has full column rank. Thus, $(F_x(\hat{z},t), H_x(\hat{z},t))$ is UCO.

In order to show that $(F_x(\hat{z},t), G_x(\hat{z},t))$ is UCC, we employ Theorem 6.12 of Chen [6] which states that if the controllability codistribution $\mathrm{d}\mathcal{C} = [\mathcal{K}_0(\hat{z},t), ..., \mathcal{K}_{n-1}(\hat{z},t)]$, where $\mathcal{K}_m(\hat{z},t) = F_x(t)\mathcal{K}_{m-1}(\hat{z},t) + (\mathrm{d}/\mathrm{d}t)\mathcal{K}_{m-1}(\hat{z},t)$ and $\mathcal{K}_0(\hat{z},t) = G_x(\hat{z},t)$, has full rank, then the pair $(F_x(\hat{z},t), G_x(\hat{z},t))$ is UCC. Since $F_x(\hat{z},t)$ and $G_x(\hat{z},t)$ are continuously differentiable and $\mathrm{rank}(G_x(\hat{z},t)) = 6$, we have shown that $(F_x(\hat{z},t), G_x(\hat{z},t))$ is UCC. □

Proposition 8 *Consider the estimator defined by* (11), (23), *and* (26) *and assume $\epsilon_z = 0$, $\epsilon_y = 0$. The equilibrium points $R(q^n_{\bar{b}}) - \bar{R}^n_{\bar{b}} = 0_3$, $b^b - \bar{b}^{\bar{b}} = 0$, and $\delta x = 0$ of the cascaded error dynamics Σ_1–Σ_2 is GES for positive definite symmetric matrices $P(0)$, $\mathcal{Q}$, and $\mathcal{R}$. Consequently, the equilibrium point $\tilde{z} = 0$ is GES.*

PROOF. First we examine the case when no resets of the form (26) occur. Under Proposition 2, Lemma 7, (29), and choosing $\mathcal{Q}$, $\mathcal{R}$, and $P(0)$ to be symmetric and positive definite, Theorem 2.1 of Johansen and Fossen [13] proves that the origin $R(q_b^n) - R_{\bar{b}}^n = 0$, $b^b - \bar{b}^{\bar{b}} = 0$, and $\delta x = 0$ of the cascaded error dynamics Σ_1–Σ_2 is GES.

If resets occur, it is sufficient to prove that there will only be a finite number of them, after the last of which the above result will be true.

There exists a threshold ξ_1 such that when $\delta\bar{u} < \xi_1$, there exists a $\xi_2(\|\delta\bar{u}\|_2)$ such that if $\|\delta u\|_2 < \xi_2(\|\delta\bar{u}\|_2)$, no resets can occur. That such a threshold exists is apparent when $\|\delta\bar{u}\|_2 = 0 \Rightarrow \delta u = \delta\check{u}$, and $\|\delta u\|_2^2 \geq (1-\epsilon)/(1+\epsilon)$ in order for a reset to occur. Therefore, $\xi_2(0) = (1-\epsilon)/(1+\epsilon)$ and $\xi_2(\|\delta\bar{u}\|_2) > 0 \ \forall \ \|\delta\bar{u}\|_2 < \xi_1$. Let resets occur at times $t_k, k = 0, 1, ...$ where $k = 0$ is the index of initialization and $k > 0$ are the indices of resets. By Theorem 4.14 in Khalil [18], we have between resets for the GES error dynamics Σ_1–Σ_2 that $c_1\|\delta x(t)\|_2^2 \leq V(\delta x(t)) \leq c_2\|\delta x(t)\|_2^2$ and $\dot{V}(\delta x(t), t) \leq -c_3\|\delta x(t)\|_2^2$. From the latter and $-V(\delta x(t))/c_2 \geq -\|\delta x(t)\|_2^2$ we find

$$\int \frac{\mathrm{d}V(\delta x(t))}{V(\delta x(t))} \leq \int -\frac{c_3}{c_2}\mathrm{d}t \tag{54a}$$

$$V(\delta x(t)) \leq V(\delta x(t_k))e^{-\frac{c_3}{c_2}(t-t_k)} \tag{54b}$$

$$\|\delta x(t)\|_2 \leq c\|\delta x(t_k)\|_2 e^{-\frac{c_3}{2c_2}(t-t_k)} \tag{54c}$$

where $c = \sqrt{c_2/c_1}$. Since c_1 and c_2 always can be chosen strictly positive and bounded, respectively, c always exists and is bounded. In order to prove that the number of resets is finite, we explore a claim about an infinite number of resets. Since $\|\delta\bar{x}\|_2$ is exponentially decaying, there must then come a k such that $\|\delta\bar{x}(t_k)\|_2 = \|\delta x(t_k)\|_2 \leq \xi_2(\|\delta\bar{x}(t)\|_2)/c$. Since $c \geq 1$, this means that $\|\delta x(t_k)\|_2 \leq \xi_2(\|\delta\bar{x}(t)\|_2)$, and clearly, the last reset has occured. The claim is therefore false, and only a finite number of resets can occur. Consequently, $\delta x \to 0$ exponentially.

Since $\|\delta u\|_2 = 0 \Rightarrow \delta q = [1, 0, 0, 0]^\top$ and $q_{\hat{b}}^n = q_b^{n\star} \otimes (\delta q)^{-1}$, we know that $\|\delta u\|_2 \to 0$ implies $q_{\hat{b}}^n \to q_b^{n\star}$, and consequently, that $\|\delta x\|_2 \to 0$ implies $\|\tilde{z}\|_2 \to 0$. This concludes the proof. □

4 Simulations and Experiments

Both the MXKF and MEKF discussed hereafter are implemented as discrete time KFs [1] . The number of scalar operations performed in the NLO, MXKF, and MEKF implementations are represented in Table 1.

[1] Code implementation of the MXKF can be found at http://folk.ntnu.no/bardbakk/MXKF-Automatica/

Table 1
The number of scalar addition (A.), multiplication (M.), division (D.) subtraction (S.), and square root (Sq.) operations performed in one time and measurement update of each estimator in their respective implementations. Available normalized measurements have been assumed here.

	A.+S.	M.	D.	Sq.
NLO	149	220	12	4
MXKF [2]	1998	2542	38	2
MEKF [2]	1736	2204	30	2
NLO+MXKF [2]	2147	2762	50	6

In this section, Euler angle errors are used to calculate MAE values and display attitude error trajectories. These have been found by extracting the Euler angles from $(q_{\hat{b}}^n)^{-1}q_b^n$, where $q_{\hat{b}}^n$ is a placeholder for the NLO, MXKF, and MEKF estimates and q_b^n is the true attitude.

In the following comparison study, four estimators are included: one aggressively tuned NLO (denoted $\overline{\text{NLO}}$), one conservatively tuned NLO, one MXKF linearized about the former NLO, and a standard MEKF using the MRP formulation. The former NLO serves as a linearization point with fast convergence and the latter is included to provide a fair comparison of steady-state performance.

4.1 Simulations

The simulated scenario is a rotating vehicle with no translative motion. Accelerometer and magnetometer measurements, $f_{nb,m}^b$ and m_m^b respectively, form vector pairs with the NED counterparts $-g^n = [0; 0; -9.818]$ and $m^n = [0.3197; 0; 0.6926]$, respectively. We define $y_1^n \triangleq -g^n/\|g^n\|_2$, $y_1^b \triangleq f_{nb,m}^b/\|f_{nb,m}^b\|_2$, $y_2^n \triangleq m^n/\|m^n\|_2$, and $y_2^b \triangleq m_m^b/\|m_m^b\|_2$ where $y_1^b = R^\top(q_b^n)y_1^n + \epsilon_1$, $y_2^b = R^\top(q_b^n)y_2^n + \epsilon_2$, $\epsilon_1 \sim \mathcal{N}(0, \sigma_1^2)$ and $\epsilon_2 \sim \mathcal{N}(0, \sigma_2^2)$.

100 different 600 seconds long scenarios have been simulated. The same attitude trajectory was used in all simulations, which was generated by $\omega_{nb}^b = [-0.1\cos(0.15t), 0.1\sin(0.10t), -0.1\cos(0.05t)]^\top$. Furthermore, $b_{ars}^b = [0.012; -0.021; 0.014]$rad/s, $\bar{b}_{ars}^{\bar{b}}(0) = \hat{b}_{ars}^{\hat{b}}(0) = [0; 0; 0]$rad/s, $\sigma_1 = 2 \cdot 10^{-3}$, $\sigma_2 = 4 \cdot 10^{-3}$, and $\sigma_\omega = 10^{-3}$rad/s. The NLO tuning parameters were set to $K_p = 10$, $k_I = 0.02$, and $\sigma = 1$ for the aggressively tuned one and $K_p = 1.5$, $k_I = 0.02$, and $\sigma = 1$ for the conservatively tuned one. For the MXKF and MEKF, the true values of σ_1, σ_2, and σ_ω were used in addition to $\sigma_b = 10^{-4}$, and the initial covariance matrix was $P(0) = \text{blockdiag}(I_3, I_3 \cdot 10^{-7})$, where blockdiag$(\cdot)$

[2] The 6×6 matrix inversion in the calculation of the Kalman gain is assumed to be performed by equations (13.65)–(13.66a) of [24].

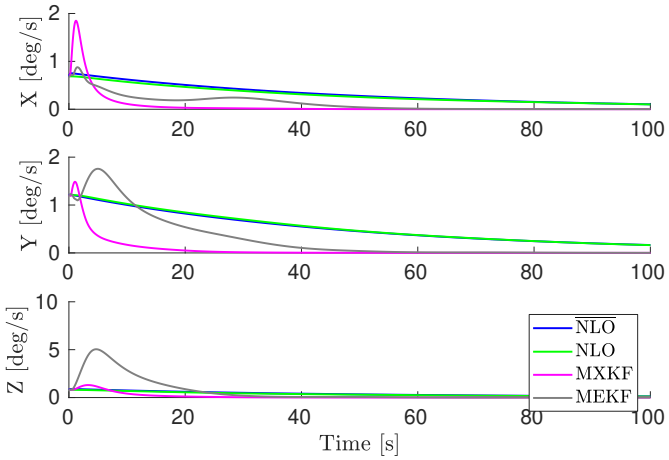

Fig. 4. The 100 first seconds of the ARS bias estimation errors averaged over the 100 simulations.

Table 2
MAE values describing the steady-state performance of the estimators.

	Roll (deg)	Pitch (deg)	Yaw (deg)
$\overline{\text{NLO}}$	0.029	0.032	0.147
NLO	0.021	0.026	0.073
MXKF	0.007	0.007	0.021
MEKF	0.007	0.007	0.022

Table 3
MAE values describing the transient performance of the estimators.

	Roll (deg)	Pitch (deg)	Yaw (deg)
$\overline{\text{NLO}}$	0.065	0.062	0.174
NLO	0.410	0.161	0.583
MXKF	0.065	0.051	0.323
MEKF	0.173	0.092	1.357

forms a block diagonal matrix of its inputs. The estimators were given the same initial estimate of attitude and bias in each simulation, and the initial attitude was randomly drawn from a uniform distribution between $-180°$ and $180°$. Both the IMU and estimators were updated with 100Hz.

In Table 2 and 3, the steady-state and transient performances are represented by mean absolute error (MAE) values. The transient MAEs are calculated from the first 200 seconds of the simulations, and the steady-state MAEs from the last 300 seconds. In Figure 4, the bias trajectories for each estimator are shown. The above mentioned attitude and bias errors are averaged over the 100 simulations.

Table 4
MAE values calculated from 50 seconds to the end of the experiments.

	Roll (deg)	Pitch (deg)	Yaw (deg)
$\overline{\text{NLO}}$	0.705	0.714	4.276
NLO	0.484	0.497	4.109
MXKF	0.311	0.407	3.023
MEKF	0.330	0.417	2.748

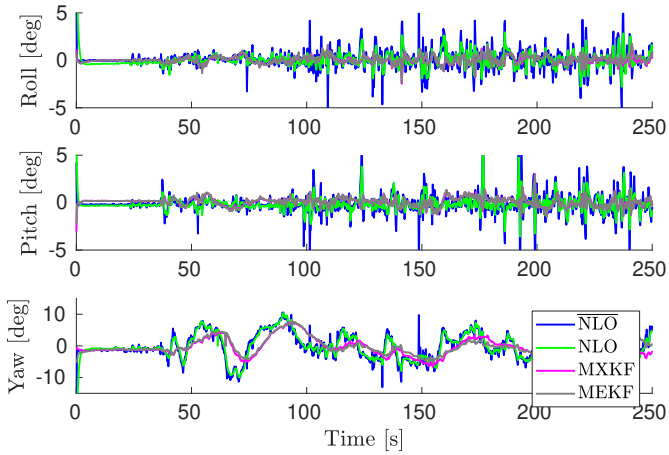

Fig. 5. This figure shows the error between the attitude trajectories from Qualisys and NLO, MXKF, and MEKF.

4.2 Experiments

The experiments were conducted in the Marine Cybernetics Lab (MC-Lab) at NTNU, which contains a water tank for experimental testing. It is equipped with Qualisys Oqus Underwater camera system providing a ground truth attitude estimate. The IMU used in the experiments was an Xsens MTi-3, delivering accelerometer, ARS, and magnetometer measurements at 25Hz, the same rate with which the estimators were updated. IMU calibration yielded the values g^n, m^n, σ_1, σ_2, and σ_ω described in Section 4.1, which naturally are also used here. The tuning of the estimators in the experiments were identical to the tuning in the simulations.

In Figure 5–6, the attitude, attitude error, and bias estimates from the experimental data are shown, respectively.

5 Discussion

The small angle approximation commonly used in the development of the MEKF is not used for the MXKF. Instead, the XKF method ensures that all linearization errors vanish as the NLO converges, yielding a global stability result. However, this also increases the computational burden of the algorithm, as F_x, G_x, and H_x become more complex and an NLO is needed. Table

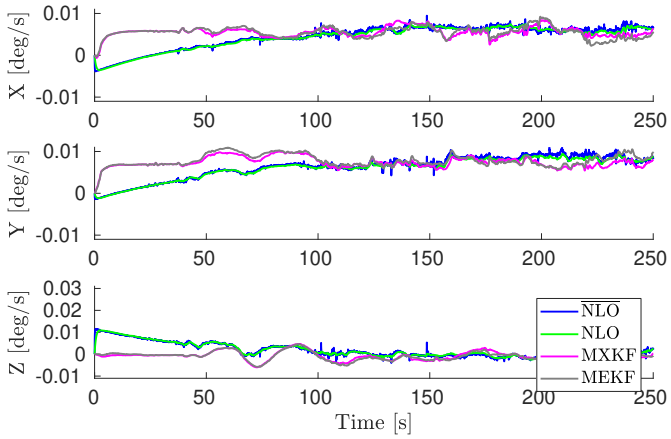

Fig. 6. The bias estimates from experimental data.

1 shows an increase of approximately 25% in computational complexity when using the MXKF linearized about the estimate from an NLO relative to the MEKF.

The linearization point provided by the NLO is used directly in the MXKF design. Thus, the uncertainty of a noisy linearization point is neglected. If the noise on the NLO estimate is large, it might be beneficial to account for the added uncertainty of the noisy linearization point by scaling of $\mathcal{Q}$ and $\mathcal{R}$ or adding appropriate noise terms in the model. This is an interesting topic in the design of XKFs generally, but outside the scope of this work.

The aggressively tuned NLO can be seen to achieve significantly improved transient performance relative to the conservatively tuned NLO without a similar deterioration of steady-state performance. The largest performance decrease can be seen in yaw. This is because there is relatively little information about yaw in the measurements as all of g^n and most of m^n are vertical components, and thus, noise sensitivity increases. Still, the performance difference is small enough to not affect the performance of the MXKF, which is apparent from the identical steady-state MAE values of the MXKF and the MEKF. This makes the NLO of [11] a suitable auxiliary estimator in the MXKF.

The improved transient performance of the aggressively tuned NLO is seen to yield significantly better transient performance of the MXKF than of the MEKF. Combined with the identical steady-state performance of the MXKF and MEKF, this is an important result. It shows that the tuning of the MXKF w.r.t. transient and steady-state performance can be decoupled; tune the NLO for fast convergence and the linearized KF for steady-state performance.

The MXKF and MEKF inherit the near-optimality of the linearized KF, which from Table 2 can be seen to yield significantly better steady-state performance than the NLO. Also, with the linearized KF, the reference vector measurements are used directly and their noises are tuned for individually. The NLO, on the other hand, calculates (11c) and does not offer as intuitive tuning w.r.t. measurement noise.

The experimental data is primarily meant to verify that the estimators work in practice, as seen in Figure 5–6, and not to provide a qualitative comparison of the estimators. It is still interesting to see that the MAE values of the MXKF and MEKF are lower than that of the NLO, and that the MXKF does not seem to suffer greatly from a less accurate linearization point. This confirms the conclusions drawn from the simulations study.

Though higher performing KFs than the MEKF exists, e.g. the UKF, IEKF, and GEKF, these have not been included in this paper since a broad comparison study is outside the scope of this work.

6 Conclusion

In this paper, the MXKF is presented. It is a novel KF-based attitude filter employing an attitude representation of minimal degree and with global exponential stability. The MXKF was shown both in simulations and experimentally to have better transient performance than and identical steady-state performance as the MEKF with identical tuning. This is an important result, as it improves the stability and transient performance of the MEKF.

References

[1] M. S. Andrle and J. L. Crassidis. Attitude Estimation Employing Common Frame Error Representations. *Journal of Guidance, Control, and Dynamics*, 38(9):1614–1624, 2015.

[2] I. Y. Bar-Itzhack and R. R. Harman. Optimized TRIAD Algorithm for Attitude Determination. *Journal of Guidance, Control, and Dynamics*, 20(1):208–211, 1997.

[3] P. Batista, C. Silvestre, and P. Oliveira. A GES attitude observer with single vector observations. *Automatica*, 48(2):339–388, 2012.

[4] H. D. Black. A passive system for determining the attitude of a satellite. *AIAA Journal*, 2(7):1350–1351, 1964.

[5] S. Bonnabel, P. Martin, and E. Salaun. Invariant Extended Kalman Filter: theory and application to a velocity-aided attitude estimation problem. In *Proceedings of the 48th IEEE Conference on Decision and Control, 2009*, volume 6, pages 1297–1304, 2009.

[6] C. Chen. *Linear System Theory and Design*. Oxford University Press, Inc., 3rd edition, 1998.

[7] J. L. Crassidis and F. L. Markley. Unscented Filtering for Spacecraft Attitude Estimation. *Journal of Guidance, Control, and Dynamics*, 26(4):536–542, 2003.

[8] J. A. Farrell. *Aided navigation : GPS with high rate sensors*. McGraw-Hill, 2008.

[9] T. I. Fossen. *Handbook of Marine Craft Hydrodynamics and Motion Control*. 2011.

[10] H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi. Attitude Estimation Using Biased Gyro and Vector Measurements With Time-Varying Reference Vectors. *IEEE Transactions on Automatic Control*, 57(5):1332–1338, 2012.

[11] H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi. Globally exponentially stable attitude and gyro bias estimation with application to GNSS/INS integration. *Automatica*, 51(June):158–166, 2015.

[12] Minh D Hua. Attitude estimation for accelerated vehicles using GPS/INS measurements. *Control Engineering Practice*, 18(7):723–732, 2010.

[13] T. A. Johansen and T. I. Fossen. Nonlinear Filtering with eXogenous Kalman Filter and Double Kalman Filter. In *Proceedings of European Control Conference*, 2016.

[14] Tor A Johansen and Thor I Fossen. Nonlinear Observer for Tightly Coupled Integration of Pseudorange and Inertial Measurements. *IEEE Transactions on Control Systems Technology*, pages 1–8, 2016.

[15] Tor A Johansen, Thor I Fossen, and Graham C Goodwin. Three-stage filter for position estimation using pseudo-range measurements. *IEEE Transactions on Aerospace and Electronic Systems*, pages 1–12, 2016.

[16] E. K. Jørgensen, T. A. Johansen, and I. Schjølberg. Enhanced Hydroacoustic Range Robustness of Three-Stage Position Filter based on Long Baseline Measurements with Unknown Wave Speed. In *Conference on Control Applications in Marine Systems*, 2016.

[17] C. Karlgaard and H. Schaub. Nonsingular attitude filtering using Modified Rodrigues Parameters. *The Journal of the Astronautical Sciences*, 57:777–791, 2010.

[18] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, 3 edition, 2000.

[19] R. Mahony, T. Hamel, and J. M. Pflimlin. Nonlinear complementary filters on the special orthogonal group. *IEEE Transactions on Automatic Control*, 53(5):1203–1218, 2008.

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

[21] F. L. Markley and J. L. Crassidis. *Fundamentals of Spacecraft Attitude Determination and Control*. 2014.

[22] F. L. Markley and D. Mortari. How to estimate attitude from vector observations. *Advances in the Astronautical Sciences*, 103(PART III):1979–1996, 1999.

[23] Carl D MEYER. Generalized Inverses and Ranks of Block Matrices. *SIAM Review*, 25(4):597–602, 1973.

[24] S. Puntanen, G. P. H. Styan, and J. Isotalo. Block-Diagonalization and the Schur Complement. In *Matrix Tricks for Linear Statistical Models*, chapter 13, pages 291–304. Springer Berlin Heidelberg, 2011.

[25] S. Salcudean. A Globally Convergent Angular Velocity Observer for Rigid Body Motion. *IEEE Transactions on Automatic Control*, 36(12):1493–1497, 1991.

[26] M. D. Shuster. The Quaternion in Kalman Filtering. *Advances in the Astronautical Sciences*, 85(pt 1):25–37, 1993.

[27] B. B. Stovner and T. A Johansen. Hydroacoustically Aided Inertial Navigation for Joint Position and Attitude Estimation in Absence of Magnetic Field Measurements. *Proc. of the American Contr. Conf.*, 37(1):1211–1218, 2017.

[28] B. B. Stovner, T. A. Johansen, T. I. Fossen, and I. Schjølberg. Three-stage Filter for Position and Velocity Estimation from Long Baseline Measurements with Unknown Wave Speed. *Proc. of the American Contr. Conf.*, pages 4532–4538, 2016.

[29] B. B. Stovner, T. A. Johansen, and I. Schjølberg. Globally Exponentially Stable Aided Inertial Navigation with Hydroacoustic Measurements from One Transponder. *Proc. of the American Contr. Conf.*, 37(1):1219–1226, 2017.

[30] J Thienel and R M Sanner. A Coupled Nonlinear Spacecraft Attitude Controller and Observer with an Unknown Constant Gyro Bias and Gyro Noise. *IEEE Transactions on Automatic Control*, 48(11):2011–2014, 2003.

[31] G. Wahba. A Least Squares Estimate of Satellite Attitude. *SIAM Review*, 7(3):409, 1965.

[32] X. Xu. Generalization of the Sherman-Morrison-Woodbury formula involving the Schur complement. *Applied Mathematics and Computations*, 309:183–191, 2017.

[33] R. Zanetti and R. H. Bishop. Quaternion estimation and norm constrained Kalman filtering. *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, (August):1–15, 2006.

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