# A Robust Observer with Gyroscopic Bias Correction for Rotational Dynamics

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

# A Robust Observer with Gyroscopic Bias Correction for Rotational Dynamics

**Erjen Lefeber**[1]
TU Eindhoven

**Marcus Greiff**[2]
MERL

**Anders Robertsson**[3]
Lund University

April 7, 2023

## Abstract

We propose an observer for rotational dynamics subject to directional and gyroscopic measurements, which simultaneously estimates the gyroscopic biases and attitude rates. We show uniform almost global asymptotic and local exponential stability of the resulting error dynamics, implying robustness against bounded disturbances. This robustness is quantified with respect to a popular nonlinear complementary filter in quantitative simulation studies, and we explore how the measurement noise propagates to the asymptotic errors as a function of tuning. This is an extended version of a paper with the same title (to appear at IFAC WC 2023). Additional mathematical details are provided in this extended version.

## 1 Introduction

The inertial measurement unit (IMU) is a ubiquitous sensor in modern robotics, often used in conjunction with other sensing modalities to infer a system's rotational degrees of freedom. In applications such as micro quadrotor control, it is essential to acquire these estimates at high rates to implement controllers with sufficient bandwidth, necessitating computationally lightweight estimators.

Largely driven by aerospace applications, a significant body of work exists on how to fuse the IMU measurements into an accurate estimate of the rotation and gyroscopic biases, see, e.g., (Markley et al., 2005; Zamani et al., 2015; Ligorio and Sabatini, 2015; Caruso et al., 2021). In the context of attitude estimation, the early work of (Farrell, 1970) set the grounds for the myriad nonlinear Kalman filters since proposed. These Bayesian methods are often used in practice due to their simplicity and flexibility. However, while the extended, unscented, and other variant assumed Gaussian density filters revert to a standard Kalman filter in a linear setting for which convergence guarantees exist(see, e.g., (Särkkä, 2013)), little can be said about worst case performance, convergence, and robustness of these nonlinear filters (Arasaratnam and Haykin, 2009). It is worth noting that Bayesian particle filters (Arulampalam et al., 2002) are asymptotically optimal in the nonlinear setting as the number of particles (and implicitly, the computational burden) approaches infinity. These have also been considered for attitude estimation in (Cheng and Crassidis, 2004), but are not practical given how fast the estimates need to be computed. Due to the flexibility of these approaches, both *attitude kinematics* and *attitude dynamics* have been considered, often in conjunction with other modalities such as camera and GPS measurements (Johansen et al., 2017).

An alternative approach is to work with nonlinear stability theory, and not presuppose anything about the noise statistics, but rather design observers which are implicitly robust to disturbances. This method is used in the vast literature on nonlinear complementary filtering, culminating with the seminal works of (Mahony et al., 2005, 2008). Here, several observers are derived for *attitude kinematics* using Lyapunov theory, with subsequent applications in (Mahony

---

et al., 2012) and recent extensions in Mahony et al. (2022). A similar approach is taken in (Berkane and Tayebi, 2017), where the observer gains are made state dependent to further improve robustness.

However, when considering control applications, we are generally also interested in the attitude rates to compute the actuating torques. An appealing alternative is therefore to consider the *attitude dynamics*, making use of the torques to compute filtered estimates of the attitude, the gyroscopic biases, and the attitude rates. Nevertheless, the application of above mentioned methods to *attitude dynamics* is less explored. Some work has been done in, e.g., (Ng et al., 2020; Lu et al., 2016), but in these works gyroscopic measurements have been ignored. To the best knowledge of authors, there exist no works that show uniform local exponential stability and uniform almost global asymptotic stability of the error dynamics in this setting, producing filtered estimates of the attitude, the attitude rate, and the gyroscopic biases. We contribute such a solution, which is important for three reasons: it facilitates the derivation of filtered output feedback controllers for the *attitude dynamics* with explicit gyroscopic bias estimation, permitting extensions of (Lefeber et al., 2020). Secondly, the *uniform* stability property provides rigorous robustness guarantees in the sense of (Khalil, 2002, Lemma 9.3). Finally, the observer comes with almost global convergence guarantees in contrast to the nonlinear Kalman filters that are often considered for this problem.

## 1.1 Outline

The mathematical preliminaries are given in Sec. 2, before stating the problem formulation in Sec. 3. The main results are presented in Sec. 4 in four steps: we (i) start by presenting an observer for the angular momentum in the inertial frame; (ii) restate the seminal result by Mahony; (iii) combine these two observers with a convex combination of the innovation terms; and (iv) describe how the attitude rate estimates can be recovered in the body-fixed frame. This is illustrated by numerical results in Sec. 5, and the conclusion in Sec. 6 closes the paper. Some key steps in the proofs are elaborated upon in Appendix A, and a discrete-time implementation is provided as Matlab code in Appendix C.

## 2 Preliminaries

In this section we introduce the notation, definitions and theorems used in the remainder of this paper.

**Theorem 1** (Corollary of Loría et al. (2005, Theorem 1)). *Consider the dynamical system*

$$\dot{x} = f(t, x) \qquad\qquad x(t_0) = x_0 \qquad\qquad f(t, 0) = 0, \qquad (1)$$

*with $f : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$ locally bounded, continuous and locally uniformly continuous in t.*

*If there exist $j$ differentiable functions $V_i : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$, bounded in t, and continuous functions $Y_i : \mathbb{R}^n \to \mathbb{R}$ for $i \in \{1, 2, \ldots j\}$ such that*

- *$V_1$ is positive definite and radially unbounded,*

- *$\dot{V}_i(t, x) \leq Y_i(x)$, for all $i \in \{1, 2, \ldots, j\}$,*

- *$Y_i(x) = 0$ for $i \in \{1, 2, \ldots, k-1\}$ implies $Y_k(x) \leq 0$, for all $k \in \{1, 2, \ldots, j\}$,*

- *$Y_i(x) = 0$ for all $i \in \{1, 2, \ldots, j\}$ implies $x = 0$,*

*then the origin $x = 0$ of (1) is uniformly globally asymptotically stable (UGAS).*

For definitions of uniform global (or local) asymptotic (or exponential) stability (UGAS/UGES/ULES), refer to (Khalil, 2002).

**Definition 1.** *The origin of (1) is uniformly almost globally asymptotically stable (UaGAS) if it is UGAS, except for initial conditions in a set of measure zero.*

We consider rotations $R \in \mathrm{SO}(3) = \{R \in \mathbb{R}^{3 \times 3} \mid R^\top R = I, \det R = 1\}$, and define the skew-symmetric map

$$S(a) = -S(a)^\top = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \in \mathfrak{so}(3). \qquad (2)$$

As the cross product can be expressed $a \times b = S(a)b$, the following useful properties hold for $S : \mathbb{R}^3 \mapsto \mathbb{R}^{3 \times 3}$:

$$S(a)^\top = -S(a) \qquad\qquad \forall a \in \mathbb{R}^3 \tag{3a}$$

$$S(a)b = -S(b)a \qquad\qquad \forall a, b \in \mathbb{R}^3 \tag{3b}$$

$$a^\top S(b)a = 0 \qquad\qquad \forall a, b \in \mathbb{R}^3 \tag{3c}$$

$$RS(a) = S(Ra)R \qquad\qquad \forall R \in \text{SO}(3), \forall a \in \mathbb{R}^3 \tag{3d}$$

$$S(a)S(b) = ba^\top - (b^\top a)I_3 \qquad\qquad \forall a, b \in \mathbb{R}^3. \tag{3e}$$

We let $\|x\|_2 = (x^\top x)^{1/2}$, using the same notation referring to the induced two norm in the context of matrices. We also consider $\mathcal{L}_2$-norms over an interval $[a, b]$ defined in these norms, as $\|x\|_{\mathcal{L}_2([a,b])} = (\int_a^b \|x(t)\|_2^2 \mathrm{d}t)^{1/2}$.

**Lemma 1** (Lefeber et al. (2020, Lemma 5)). *Consider the dynamical systems $\dot{R}_1 = R_1 S(\omega_1)$ and $\dot{R}_2 = R_2 S(\omega_2)$. Let $R_{12} = R_1 R_2^\top$ and $\omega_{12} = \omega_1 - \omega_2$. Then*

$$\dot{R}_{12} = R_{12}S(R_2\omega_{12}) = S(R_1\omega_{12})R_{12}, \tag{4}$$

*and differentiating for some constant vector $v$*

$$V = \frac{1}{2}(R_{12}v - v)^\top (R_{12}v - v) = \frac{1}{2}\|R_{12}v - v\|_2^2,$$

*along solutions of* (4) *results in $\dot{V} = \omega_{12}^\top S(R_1^\top v)R_2^\top v$.*

**Lemma 2.** *Define $r_k = \sum_{i=1}^n k_i S(R^\top v_i)v_i$ with $k_i > 0$ and $v_i \in \mathbb{R}^3$ such that $M = \sum_{i=1}^n k_i v_i v_i^\top = U\Lambda U^\top$ with $U \in \text{SO}(3)$ and $\Lambda$ a diagonal matrix with distinct eigenvalues $\lambda_i$, i.e., $\lambda_3 > \lambda_2 > \lambda_1 > 0$. Then $r_k = 0$ implies that $U^\top RU \in \{I, D_1, D_2, D_3\}$, where $D_1 = \text{diag}(1, -1, -1)$, $D_2 = \text{diag}(-1, 1, -1)$, $D_3 = \text{diag}(-1, -1, 1)$. Furthermore, if in addition $\dot{R} = RS(\omega)$ and $\dot{r}_k = 0$, then also $\omega = 0$.*

**Proof 1.** *The first claim was shown in (Mahony et al., 2008). By defining $\bar{R} = URU^\top$, $\bar{\omega} = U\omega$, and $\bar{k}_i = k_i v_i^\top v_i$ it follows that without loss of generality, we can assume that $U = I$ and $v_i^\top v_i = 1$. Then $r_k = 0$ implies $R = \text{diag}(r_1, r_2, r_3) = \text{diag}(\pm 1, \pm 1, \pm 1)$. Let $\Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$. Then we have*

$$\begin{aligned}
\dot{r}_k &= -\sum_{i=1}^n k_i S(S(\omega)R^\top v_i)v_i \\
&= -\sum_{i=1}^n k_i S(v_i)S(R^\top v_i)\omega \\
&= -\text{diag}(r_2\lambda_2 + r_3\lambda_3, r_1\lambda_1 + r_3\lambda_3, r_1\lambda_1 + r_2\lambda_2)\omega, \tag{5}
\end{aligned}$$

*from which we can conclude that $\dot{r}_k = 0$ implies $\omega = 0$, since $\lambda_i$ are distinct and $r_i \in \{-1, 1\}$.*

## 3    Problem formulation

Let $R \in \text{SO}(3)$ denote the rotation matrix from the body-fixed frame to the inertial frame and let $\omega \in \mathbb{R}^3$ denote the body-fixed angular velocities. Then the *kinematics* of a rotating rigid body can be described by

$$\dot{R} = RS(\omega), \tag{6}$$

where $\omega$ is regarded as input. Consider the outputs

$$y_0 = \omega + b \qquad\qquad y_i = R^\top v_i \qquad\qquad i = 1, \ldots, n, \tag{7}$$

where $b$ is an unknown constant, and $v_i$ denote $n$ known inertial directions. That is, assume biased measurement of angular velocities and body-fixed frame observations of the fixed inertial directions $v_i$.

**Assumption 1.** *For attitude reconstruction $n \geq 2$ independent inertial directions are required. However, if we have two independent directions $v_1$ and $v_2$, then $v_3 = v_1 \times v_2 = S(v_1)v_2$ is a third independent direction. Therefore, in the remainder we assume without loss of generality that $n \geq 3$ instead.*

In this setting, a large number of observers exist, such as the filters in the seminal work of (Mahony et al., 2008):

**Theorem 2** (Mahony et al. (2008, Th. 5.1)). *Consider the explicit complementary filter with bias correction*

$$\dot{\hat{b}} = k_b \tilde{r}_k \qquad\qquad \dot{\hat{R}} = \hat{R}S(y_0 - \hat{b} - k_R \tilde{r}_k), \qquad (8)$$

*where $\tilde{r}_k = \sum_{i=1}^{n} k_i S(\hat{R}^\top v_i) y_i$, $k_R > 0$, and $k_b > 0$. Define the estimation errors $\tilde{R} = \hat{R}R^\top$ and $\tilde{b} = \hat{b} - b$. If $\omega(t)$ is a bounded absolutely continuous signal, the pair of signals $(\omega(t), \tilde{R})$ is asymptotically independent, and the weights $k_i > 0$ are chosen such that $M = \sum_{i=1}^{n} k_i v_i v_i^\top$ has distinct eigenvalues, then $(\tilde{R}, \tilde{b})$ is almost globally asymptotically stable and locally exponentially stable to $(I, 0)$.*

This explicit complementary filter with bias correction (8) has seen much use in practice. However, this filter only produces estimates for the attitude and bias, but not an estimate for the angular velocities. Clearly, from measurements $y_0$ and bias estimate $\hat{b}$ an unbiased estimate for the angular velocities is available, but for noisy $y_0$ this unbiased estimate for the angular velocities is also noisy and not a filtered signal. Therefore, the goal of this paper is to extend the explicit complementary filter with bias correction to the *dynamics* of a rotating body, producing not only filtered estimates for the attitude and bias, but also filtered unbiased estimates for the angular velocities. To be precise, we aim to solve the following problem.

**Problem 1.** *The motion of a rotating rigid body configured on $R \in \mathrm{SO}(3)$ is governed by the dynamics*

$$\dot{R} = RS(\omega) \qquad\qquad J\dot{\omega} = S(J\omega)\omega + \tau, \qquad (9)$$

*where $J = J^\top > 0$ denotes the inertia matrix with respect to the body-fixed frame and $\tau \in \mathbb{R}^3$ denotes the total moment vector in the body-fixed frame, is a known input.*

*Consider the outputs (7). Design an observer/filter which produces estimates $\hat{R}$, $\hat{\omega}$, and $\hat{b}$ such that the point $(I, 0, 0)$ of the estimation error dynamics $(\tilde{R}, \tilde{\omega}, \tilde{b})$, given by*

$$\tilde{R} = \hat{R}R^\top \qquad\qquad \tilde{\omega} = \hat{\omega} - \omega \qquad\qquad \tilde{b} = \hat{b} - b, \qquad (10)$$

*is almost globally and locally exponentially stable.*

## 4 Main results

The difficulty in almost globally solving Problem 1 is dealing with the Coriolis-terms, which contains quadratic expressions in the angular velocities. Our way around this difficulty is to first design an observer for the angular *momentum* expressed in the *inertial frame*. Next, our estimate for the attitude can be used to transform those estimates into estimates for the angular *velocities* expressed in the *body-fixed frame*.

As a first step, we consider the problem of designing an observer for both the attitude and the angular momentum expressed in the inertial frame without using the measurement of angular velocities. As a second step, we revisit the explicit complementary filter with bias correction by (Mahony et al., 2008) to prepare for our third step. In our third step we fuse the observers derived in the previous steps to produce an estimates for the attitude, the angular momentum expressed in the inertial frame, and a bias estimate. In our fourth and final step, the derived estimates are used to estimate the angular velocities in the body-fixed frame using only the measured outputs in (7).

### 4.1 Step 1: Angular momentum estimator

Our first goal is to design an observer for estimating the angular momentum expressed in the inertial frame using only the body-fixed frame observations of fixed inertial directions, that is without using measurement of angular velocities. To that end, define $\ell = RJ\omega$, so $\omega = J^{-1}R^\top \ell$. Then we get as resulting dynamics:

$$\dot{R} = RS(J^{-1}R^\top \ell) \qquad\qquad \dot{\ell} = R\tau. \qquad (11)$$

Consider only the outputs

$$y_i = R^\top v_i \qquad\qquad i = 1, \ldots, n. \qquad (12)$$

Our goal is to construct estimates $\hat{R}$ and $\hat{\ell}$ such that the estimation errors

$$\tilde{R} = \hat{R}R^\top \qquad\qquad \tilde{\ell} = \hat{\ell} - \ell \qquad (13)$$

converge to $I$ respectively $0$. Define the following observer:

$$\dot{\hat{R}} = \hat{R}S\left(J^{-1}R^\top\hat{\ell} - k_R\tilde{r}_k\right) \qquad\qquad \dot{\hat{\ell}} = R\tau - k_\ell RJ^{-1}\tilde{r}_k, \tag{14a}$$

where $k_R > 0$, $k_\ell > 0$, and

$$\tilde{r}_k = \sum_{i=1}^n k_i S(\hat{R}^\top v_i)R^\top v_i = \sum_{i=1}^n k_i S(\hat{R}^\top v_i)y_i. \tag{14b}$$

**Proposition 1.** *Consider the observer* (14) *in closed-loop with the dynamics* (11). *If $\omega$ and $\dot{\omega}$ are bounded and the weights $k_i$ are chosen such that $\sum_{i=1}^n k_i v_i v_i^\top$ has distinct eigenvalues $\lambda_i$, i.e., $\lambda_3 > \lambda_2 > \lambda_1 > 0$, then the estimation errors* (13) *are UaGAS and ULES towards $(I, 0)$.*

**Proof 2.** *Using Lemma 1, the estimation error dynamics can be written as*

$$\dot{\tilde{R}} = \tilde{R}S\left(R[J^{-1}R^\top\tilde{\ell} - k_R\tilde{r}_k]\right) \tag{15a}$$

$$\dot{\tilde{\ell}} = -k_\ell RJ^{-1}\tilde{r}_k. \tag{15b}$$

*Differentiating the Lyapunov function candidate*

$$V_1 = k_\ell \sum_{i=1}^n \frac{k_i}{2}\|\tilde{R}v_i - v_i\|_2^2 + \frac{1}{2}\tilde{\ell}^\top\tilde{\ell}, \tag{16}$$

*along* (15), *using Lemma 1, results in*

$$\dot{V}_1 = k_\ell(RJ^{-1}R^\top\tilde{\ell} - k_R\tilde{r}_k)^\top\tilde{r}_k + \tilde{\ell}^\top[-k_\ell RJ^{-1}R^\top\tilde{r}_k]$$
$$= -k_\ell k_R\|\tilde{r}_k\|_2^2 = Y_1, \tag{17}$$

*which is negative semi-definite. Differentiating $V_2 = -\tilde{r}_k^\top\dot{\tilde{r}}_k$ along* (15) *results in*

$$\dot{V}_2 = -\|\dot{\tilde{r}}_k\|_2^2 - \tilde{r}_k^\top\ddot{\tilde{r}}_k \tag{18a}$$

$$\leq -\|\dot{\tilde{r}}_k\|_2^2 + K\|\tilde{r}_k\|_2 \tag{18b}$$

$$\leq -\gamma\|\dot{\tilde{\ell}}\|_2^2 + \bar{K}\|\tilde{r}_k\|_2 = Y_2. \tag{18c}$$

*The first inequality follows from boundedness of $\ddot{\tilde{r}}_k$ which follows from $\dot{V}_1 \leq 0$ and* (15). *The second inequality follows from* (5) *and* (15a). *Applying Theorem 1 shows UGAS towards $\tilde{r}_k = 0$, $\dot{\tilde{r}}_k = 0$, which, using Lemma 2 implies UaGAS towards $\tilde{R} = I$, $\tilde{\ell} = 0$. Considering $V_1 + \epsilon V_2$, ULES can be shown along the lines of (Wu and Lee, 2016).* $\square$

## 4.2 Step 2: Gyroscopic bias estimator

As a second ingredient we need the observer of (8). Consider the kinematics (6) with outputs (7). Our goal is to obtain estimates $\hat{R}$ and $\hat{b}$ such that the errors

$$\tilde{b} = \hat{b} - b \qquad\qquad \tilde{R} = \hat{R}R^\top \tag{19}$$

converge to $0$ and $I$, respectively.

Define the following observer/filter:

$$\dot{\hat{b}} = k_b\tilde{r}_k \qquad\qquad \dot{\hat{R}} = \hat{R}S(y_0 - \hat{b} - k_R\tilde{r}_k) \tag{20}$$

with $k_b > 0$, $k_R > 0$, $J = J^\top > 0$ and $\tilde{r}_k$ as in (14b).

**Proposition 2.** *Consider the observer* (20) *in closed-loop with the kinematics* (6). *If $\omega$ and $\dot{\omega}$ are bounded and the weights $k_i$ are chosen such that $\sum_{i=1}^n k_i v_i v_i^\top$ has distinct eigenvalues $\lambda_i$, i.e., $\lambda_3 > \lambda_2 > \lambda_1 > 0$, then the estimation errors* (19) *are UaGAS and ULES towards $(I, 0)$.*

**Proof 3.** *The estimation error dynamics are given by*

$$\dot{\tilde{b}} = k_b\tilde{r}_k \qquad\qquad \dot{\tilde{R}} = \tilde{R}S(R[-\tilde{b} - k_R\tilde{r}_k]). \tag{21}$$

5

*Differentiating the Lyapunov function candidate*

$$V_1 = k_b \sum_{i=1}^{n} \frac{k_i}{2} \left\| \tilde{R} v_i - v_i \right\|_2^2 + \frac{1}{2} \tilde{b}^\top \tilde{b} \tag{22}$$

*along* (21) *results in*

$$\dot{V}_1 = k_b \left( -\tilde{b} - k_R \tilde{r}_k \right)^\top \tilde{r}_k + \tilde{b}^\top k_b \tilde{r}_k = -k_b k_R \left\| \tilde{r}_k \right\|_2^2, \tag{23}$$

*which is negative semi-definite. Differentiating* $V_2 = -\tilde{r}_k^\top \dot{\tilde{r}}_k$ *along* (15) *results in*

$$\dot{V}_2 = -\|\dot{\tilde{r}}_k\|_2^2 - \tilde{r}_k^\top \ddot{\tilde{r}}_k \tag{24a}$$

$$\leq -\|\dot{\tilde{r}}_k\|_2^2 + K \|\tilde{r}_k\|_2$$

$$\leq -\gamma \|\tilde{b}\|_2^2 + \bar{K} \|\tilde{r}_k\|_2 = Y_2. \tag{24b}$$

*The first inequality follows from boundedness of* $\ddot{\tilde{r}}_k$ *which follows from* $\dot{V}_1 \leq 0$, (5), *and boundedness of* $\omega$ *and* $\dot{\omega}$. *The second inequality follows from* (5) *and* (21). *The proof can be completed along the lines of that of Proposition 1.* $\square$

**Remark 1.** *Note that in our proof we do not require that the pair of signals* $(\omega(t), \tilde{R})$ *is asymptotically independent, which is difficult to check since* $\tilde{R}$ *is not an external signal (as it is generated in closed-loop with the observer). On the other hand, we need to assume that* $\dot{\omega}$ *is bounded, which is a slightly stronger condition than assuming that* $\omega$ *is absolutely continuous. However, this allows us to conclude uniform stability, which implies robustness against bounded disturbances by (Khalil, 2002, Lemma 9.3).*

## 4.3   Step 3: Fusing the two observers

Our next step is to fuse the two observers (14) and (20) into one. The observer (14) provides us with an estimate $\hat{l}$ for the angular momentum expressed in the inertial frame. Therefore, we can consider $J^{-1} R^\top \hat{\ell}$ as an estimate for the angular velocity. The observer (20) provides us with a bias estimate so that $y_0 - \hat{b}$ can also be considered as an estimate for the angular velocity. In our combined observer we fuse those to estimates, by using a fraction $\alpha$ of the first estimator, and a fraction $1 - \alpha$ of the second estimator.

With this intuition, consider the dynamics (11) together with the outputs (7). We propose the following observer

$$\dot{\hat{b}} = k_b \tilde{r}_k - \alpha k_b k_\alpha J \tilde{\delta}_L \tag{25a}$$

$$\dot{\hat{R}} = \hat{R} S \left( \alpha J^{-1} R^\top \hat{\ell} - (1-\alpha)(y_0 - \hat{b}) - k_R \tilde{r}_k \right) \tag{25b}$$

$$\dot{\hat{\ell}} = R\tau - k_\ell R J^{-1} \tilde{r}_k - (1-\alpha) k_\ell k_\alpha R \tilde{\delta}_L, \tag{25c}$$

where

$$\tilde{\delta}_L = R^\top \hat{\ell} - J(y_0 - \hat{b}) = R^\top \tilde{\ell} + J \tilde{b}. \tag{25d}$$

with $k_\alpha > 0$, $k_b > 0$, $k_R > 0$, $k_\ell > 0$, $0 < \alpha < 1$, and $\tilde{r}_k$ as defined in (14b).

**Remark 2.** *Note that,* $\tilde{\delta}_L$ *can be interpreted as the difference between two estimators for the angular momentum expressed in the body-fixed frame.*

**Proposition 3.** *Consider the observer* (25) *in closed-loop with the dynamics* (11). *If the weights* $k_i$ *are chosen such that* $\sum_{i=1}^{n} k_i v_i v_i^\top$ *has distinct eigenvalues* $\lambda_i$, *i.e.,* $\lambda_3 > \lambda_2 > \lambda_1 > 0$, *then the estimation errors*

$$\tilde{R} = \hat{R} R^\top \qquad\qquad \tilde{\ell} = \hat{\ell} - \ell \qquad\qquad \tilde{b} = \hat{b} - b, \tag{26}$$

*are UaGAS and ULES towards* $(I, 0, 0)$.

**Proof 4.** *The estimation error dynamics are given by*

$$\dot{\tilde{b}} = k_b \tilde{r}_k - \alpha k_b k_\alpha J \tilde{\delta}_L \tag{27a}$$

$$\dot{\tilde{R}} = \tilde{R} S \Big( R \big[ \alpha J^{-1} R^\top \tilde{\ell} - (1-\alpha) \tilde{b} - k_R \tilde{r}_k \big] \Big) \tag{27b}$$

$$\dot{\tilde{\ell}} = -k_\ell R J^{-1} \tilde{r}_k - (1-\alpha) k_\ell k_\alpha R \tilde{\delta}_L. \tag{27c}$$

*Differentiating the Lyapunov function candidate*

$$V_1 = k_\ell k_b \sum_{i=1}^{N} \frac{k_i}{2} \|\tilde{R} v_i - v_i\|_2^2 + \frac{k_\ell}{2}(1-\alpha)\tilde{b}^\top \tilde{b} + \frac{k_b}{2}\alpha \tilde{\ell}^\top \tilde{\ell}, \tag{28}$$

*along* (27) *results in*

$$\dot{V}_1 = k_\ell k_b \left( \alpha J^{-1} R^\top \tilde{\ell} - (1-\alpha)\tilde{b} - k_R \tilde{r}_k \right)^\top \tilde{r}_k \tag{29}$$

$$+ (1-\alpha)k_\ell \tilde{b}^\top [k_b \tilde{r}_k - \alpha k_b k_\alpha J \tilde{\delta}_L] \tag{30}$$

$$+ \alpha k_b \tilde{\ell}^\top [-k_\ell R J^{-1} \tilde{r}_k - (1-\alpha)k_\ell k_\alpha R \tilde{\delta}_L] \tag{31}$$

$$= -k_\ell k_b k_R \|\tilde{r}_k\|_2^2 - \alpha(1-\alpha)k_\ell k_b k_\alpha \|\tilde{\delta}_L\|_2^2, \tag{32}$$

*which is negative semi-definite. Here we used* (25d). *Differentiating* $V_2 = -\tilde{r}_k^\top \dot{\tilde{r}}_k$ *along* (27) *results in*

$$\dot{V}_2 = -\|\dot{\tilde{r}}_k\|_2^2 - \tilde{r}_k^\top \ddot{\tilde{r}}_k$$

$$\le -\|\dot{\tilde{r}}_k\|_2^2 + K\|\tilde{r}_k\|_2$$

$$\le -\gamma\| - R\tilde{b} + \alpha R J^{-1}\tilde{\delta}_L\|_2 + \bar{K}\|\tilde{r}_k\|_2^2$$

$$\le -\gamma\|\tilde{b}\|_2 + \bar{\bar{K}}(\|\tilde{\delta}_L\|_2 + \|\tilde{r}_k\|_2) = Y_2 \tag{33}$$

*The proof can be completed along the lines of that of Proposition* 1. □

**Remark 3.** *Note that, like in Proposition* 1, *there is no need for assuming that* $\omega$ *or* $\dot{\omega}$ *(or* $\tau$*) are bounded. From* $\dot{V}_1 \le 0$ *we have boundedness of the estimation errors, which is all we need to complete the proof.*

**Remark 4.** *Note that for* $\alpha = 0$ *or* $\alpha = 1$ *the observer* (25) *reduces to respectively* (14) *or* (20), *for which we obtained results in Proposition* 1 *respectively Proposition* 2.

## 4.4   Step 4: Final result

Our final step is to replace the estimate $\hat{\ell}$ for the angular momentum expressed in the inertial frame, obtained from the observer (25), by a filtered estimate $\hat{\omega}$ for the angular velocity expressed in the body-fixed frame. Furthermore, we need to overcome the problem that we do not know $R$, which is used in (25), as we only have (7) available for measurement, not $R$ itself.

The latter is actually less of a problem than it might seem at first glance. We assumed that the weights $k_i$ are chosen such that $M = \sum_{i=1}^{n} k_i v_i v_i^\top = U \Lambda U^\top$ has distinct eigenvalues $\lambda_i$, i.e., $\lambda_3 > \lambda_2 > \lambda_1 > 0$. Therefore, the matrix $M$ is invertible and we obtain

$$R = M^{-1} \sum_{i=1}^{n} k_i v_i v_i^\top R = M^{-1} \sum_{i=1}^{n} k_i v_i y_i^\top. \tag{34}$$

As a result, each occurrence of $R$ in (25) can be replaced by the right hand side of (34). We emphasize that (34) *is not* the attitude estimate, the attitude estimate $\hat{R}$ is still computed and updated through the ODEs in (25).

Our filtered estimate for the angular velocity expressed in the body-fixed frame is given by $\hat{\omega} = J^{-1}\hat{R}^\top \hat{\ell}$. We can now summarize our result in the following.

**Proposition 4.** *Consider the dynamics* (1) *and output* (7) *in closed-loop with the observer*

$$\dot{\hat{b}} = k_b \tilde{r}_k - \alpha k_b k_\alpha J \tilde{\delta}_L \tag{35a}$$

$$\dot{\hat{R}} = \hat{R} S \left( \alpha J^{-1} \tilde{\delta}_L + y_0 - \hat{b} - k_R \tilde{r}_k \right) \tag{35b}$$

$$\dot{\hat{\ell}} = \bar{R}[\tau - k_\ell J^{-1} \tilde{r}_k - (1-\alpha)k_\ell k_\alpha \tilde{\delta}_L] \tag{35c}$$

$$\hat{\omega} = J^{-1}\hat{R}^\top \hat{\ell} \tag{35d}$$

*where*

$$\tilde{r}_k = \sum_{i=1}^{n} k_i S(\hat{R}^\top v_i) y_i \tag{35e}$$

$$\tilde{\delta}_L = \bar{R}^\top \hat{\ell} - J(y_0 - \hat{b}) \tag{35f}$$

$$\bar{R} = \left(\sum_{i=1}^{n} k_i v_i v_i^\top\right)^{-1} \sum_{i=1}^{n} k_i v_i y_i^\top. \tag{35g}$$

*Let $k_\alpha > 0$, $k_b > 0$, $k_R > 0$, $k_\ell > 0$, $0 < \alpha < 1$. If in addition $k_i$ are chosen such that $M = \sum_{i=1}^{n} k_i v_i v_i^\top = U \Lambda U^\top$ has distinct eigenvalues $\lambda_i$, i.e., $\lambda_3 > \lambda_2 > \lambda_1 > 0$, the the observer errors (10) are UaGAS and ULES towards $(I, 0, 0)$, provided that $\omega$ is bounded.*

**Proof 5.** *From Proposition 3 we have that $\tilde{R}$, $\tilde{b}$ and $\tilde{\ell}$ are UaGAS and ULES towards $(I, 0, 0)$. Therefore, it only remains to show convergence of $\tilde{\omega}$. We have*

$$\tilde{\omega} = J^{-1} \hat{R}^\top \hat{\ell} - \omega = \underbrace{J^{-1} \hat{R}^\top \tilde{\ell}}_{\to 0} - \underbrace{J^{-1} R^\top [\tilde{R} - I] R J}_{\to 0} \omega, \tag{36}$$

*which explains the additional requirement that $\omega$ is bounded, in comparison with Proposition 3.* □

## 5 Numerical examples

In this section, we present three numerical examples. The first is a qualitative simulation to illustrate typical convergence behaviors of the estimators. Next, we give quantitative results showing the utility of combining the observers as in Propositions 3–4 by studying the statistics of the transient and stationary errors. Finally, we discuss how to tune the observers based on the asymptotic errors, and how these errors are affected by measurement noise.

### 5.1 Typical convergence in an ideal setting

In this ideal setting, we take the measurements to be noise-free and initialize a simulation with initial errors and parameters sampled from the distributions in Appendix B. The dynamical system in (1) is driven by a torque sequence

$$\tau(t) = (\sin(t + 1), \sin(2t + 2), \sin(3t + 3))^\top \in \mathbb{R}^3, \tag{37}$$

where the initial conditions and parameters are realized as

$$R(0) = \begin{bmatrix} 0.18 & 0.97 & -0.15 \\ 0.08 & 0.14 & 0.99 \\ 0.98 & -0.19 & -0.06 \end{bmatrix}, \qquad \omega(0) = \begin{bmatrix} -0.11 \\ 0.02 \\ -0.06 \end{bmatrix}, \qquad b = \begin{bmatrix} -0.12 \\ -2.54 \\ 0.28 \end{bmatrix},$$

$$\hat{R}(0) = \begin{bmatrix} 0.35 & 0.06 & 0.94 \\ 0.84 & 0.42 & -0.34 \\ -0.41 & 0.91 & 0.09 \end{bmatrix}, \qquad \hat{\ell}(0) = \begin{bmatrix} -1.12 \\ 0.05 \\ -1.24 \end{bmatrix}, \qquad \hat{b}(0) = \begin{bmatrix} -0.83 \\ 0.54 \\ 0.11 \end{bmatrix},$$

$$J = \begin{bmatrix} 0.91 & 0.03 & 0.14 \\ 0.03 & 0.73 & 0.15 \\ 0.14 & 0.15 & 0.64 \end{bmatrix}, \qquad [v_1 \quad v_2 \quad v_3] = \begin{bmatrix} 0.00 & -0.87 & -0.45 \\ 0.00 & -0.50 & 0.87 \\ -1.00 & -0.05 & 0.00 \end{bmatrix},$$

here rounded to two decimals to ease visualization. In this example, we tune the observer with

$$k_R = 2.0, \qquad k_l = 2.0, \qquad k_a = 1.0, \qquad k_b = 4.0, \tag{38a}$$
$$k_1 = 1.1, \qquad k_2 = 1.2, \qquad k_3 = 1.3, \qquad \alpha = 0.3. \tag{38b}$$

This results in a matrix $M$ in (34) with distinct eigenvalues $\Lambda = \text{diag}(1.07, 1.23, 1.30)$. The effects of the observer tuning are discussed later in Sec. 5.3. The resulting system response is shown in Fig. 1, where $\Psi(A, B) = \frac{1}{2}\text{Tr}(A^\top B - I)$. Despite initializing the estimator very away from the stable equilibrium point in this measure, we obtain a good estimate within seconds with relatively small transients in the attitude rate and bias estimates. For small errors, we observe a linear decay of the Lyapunov function $V_1$ in (28) in the 10-logarithm, as expected from the ULES property.
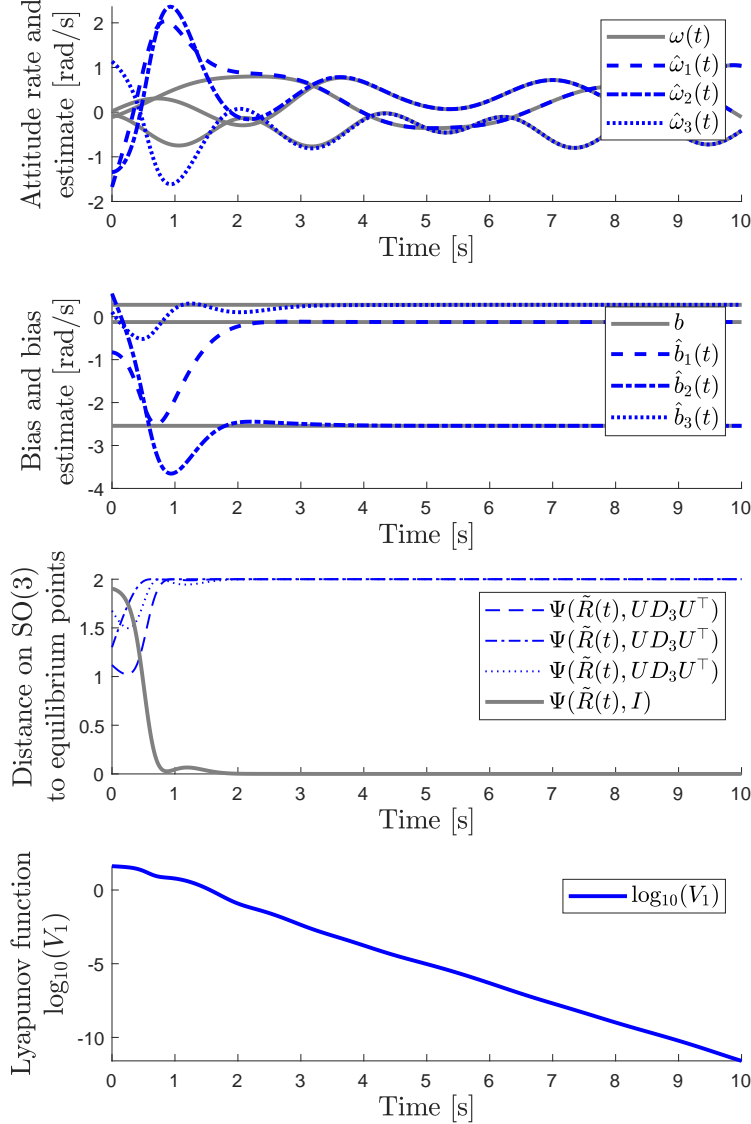
Figure 1: State trajectory (gray) and estimated states (blue), along with error signals. The gyroscopic biases, and attitude rates are shown in the top two subplots, respectively. The distance of the estimation error $\tilde{R}$ to the four equilibrium points is shown in the third subplot, with $\Psi(\tilde{R}, I)$ in gray. The Lyapunov function is depicted in the 10-logarithm in the bottom subplot.

## 5.2 Quantitative Monte Carlo results with noise

One of the more important effects of having $\alpha \in (0, 1)$ is that we effectively filter both the bias and the attitude rates, which reduces the impact of the noise in these estimates. To quantify and demonstrate this, we consider the same tuning as in Sec. 5.1, and compute the root mean-square error (RMSE) of the $\mathcal{L}_2$-norms in the signals $\|\tilde{\omega}(t)\|_2$, $\|\tilde{b}(t)\|_2$, and $\Psi(\tilde{R}(t), I)$. That is, we consider $N_{MC}$ realizations of the parameters in Appendix B, denote a trajectory from the $i$th simulation as $x^{(i)}(t)$, and let

$$\text{RMSE}_{\mathcal{L}_2([a,b])}(x) = \Big( \frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} \int_a^b \|x^{(i)}(t)\|_2^2 \mathrm{d}t \Big)^{1/2}. \tag{39}$$

Here, by considering this measure over the entire simulation time, $t \in [0, T]$, we capture the length of the initial transients, and by considering it over the last second of the simulation, $t \in [T-1, T]$, we capture the stationary errors primarily induced by the noise. These measures are shown in Table 1, as computed from $N_{MC} = 10^3$ realizations.

9

Table 1: RMSEs of transient and stationary errors categorized by signals and observers.

| Measure | $\text{RMSE}_{\mathcal{L}_2([0,T])}(x)$ | | | $\text{RMSE}_{\mathcal{L}_2([T-1,T])}(x)$ | | |
|---|---|---|---|---|---|---|
| Signal | $\Psi(\tilde{R}, I)$ | $\tilde{\omega}$ | $\tilde{b}$ | $\Psi(\tilde{R}, I)$ | $\tilde{\omega}$ | $\tilde{b}$ |
| Prop. 1 | 0.560 | 2.571 | 2.629 | $3.043 \cdot 10^{-5}$ | 0.022 | 0.178 |
| Prop. 2 | 0.577 | 2.463 | 2.401 | $2.718 \cdot 10^{-5}$ | 0.177 | 0.016 |
| Prop. 4 | 0.570 | 2.389 | 2.226 | $2.809 \cdot 10^{-5}$ | 0.021 | 0.016 |

**Remark 5.** *Here we note that there is significant variance in these measures when considered over the entire simulation time (i.e., with $[0, T]$), but the standard deviation of $\text{RMSE}_{\mathcal{L}_2([T-1,T])}(x)$ is in the order of $10^{-10}$ for $\Psi(\tilde{R}, I)$, and the order of $10^{-4}$ for $\tilde{\omega}$ and $\tilde{b}$, respectively. As such, there is a statistically significant difference in stationary performance between the observers when considering the parameter, noise, and error distributions in Appendix B.*

From these results, we note that the transient responses are similar in the three observers, but that the stationary noise levels differ greatly. In particular, the observer in Proposition 1 achieves low noise levels in the attitude rate errors, as the attitude rate estimate is filtered in the observer, but the stationary noise in the bias is relatively large. For the observer in Proposition 2, the relationship is the reverse. Finally, for the observer in Proposition 4, we filter both signals, resulting in low noise levels both in the attitude rate error and in the bias. In this simulation study, the asymptotic noise levels differ by almost one magnitude. If the observer is to be used for feedback control on the estimates $(\hat{R}, \hat{\omega})$ based on noisy measurements $\{y_i\}_{i=0}^n$, it is clear that the observers in Proposition 1 and Proposition 4 should be considered over Proposition 2 (the result of (Mahony et al., 2008)). Additionally, we note that there is clear merit to considering Proposition 4 over Proposition 1 if the asymptotic noise in the bias estimates are of concern.

## 5.3 Observer tuning

The tuning of the estimator is non-trivial, and somewhat counter intuitive. Some insight can be gained by following (Greiff, 2021, Section 5.4) and taking a local approximation of the attitude error close to the identity element, $\tilde{R} = I + S(\tilde{\epsilon}) + o(\|\tilde{\epsilon}\|_2^2)$. Here, we define measurement noise as an additive perturbation on $y_0$, and a multiplicative disturbance on $\{y_i\}_{i=1}^n$ perturbing the direction, with

$$y_0 = w + b + \delta_0, \qquad y_i = R^\top(I + S(\delta_i))v_i. \tag{40}$$

We then express the local error dynamics in (27) in $\tilde{x}^\top = (\tilde{\epsilon}^\top, \tilde{\omega}^\top, \tilde{b}^\top) \in \mathbb{R}^9$, driven by $\delta^\top = (\delta_0^\top, \delta_1^\top, \delta_2^\top, \delta_3^\top) \in \mathbb{R}^{12}$, and linearize the system about the origin, resulting in

$$\dot{\tilde{x}} = A\tilde{x} + B\delta. \tag{41}$$

Here, we compute $(A, B)$ using the automatic differentiation tool `CasADi` in (Andersson et al., 2012). This permits us to study how the tuning of the estimator affects the properties of the linear system in (41) governing the local estimation errors, and also facilitates reasoning about how certain noises affect the stationary errors by tools from linear systems theory, such as the singular-value plots from the inputs $\delta_i$ to the errors $\tilde{x}$.

In Fig. 2, we show how the spectrum of $A$, here denoted $\lambda(A)$, changes in the complex plane when varying the parameter $\alpha$ subject to the nominal tuning and realization in Sec. 5.1 and a stationary rotation $R$. Note, that the error dynamics are time invariant if and only if $R$ is time invariant. We also show the maximum singular value from the gyroscopic noise input $\delta_0$ to the local observation errors $\tilde{x}$. That is, with the transfer function $G(s) = (sI - A)^{-1}B$, we compute the singular values $\sigma(G(i\omega)) = \sqrt{\lambda(G(-i\omega)^\top G(i\omega))}$ as a function of the frequency $\omega$.

The location of the poles of the linearized error dynamics behave highly non-trivially as a function of the observer parameters $\{k_a, k_b, k_R, k_l, \alpha, J\}$, and that when fixing the nominal parameters and varying $\alpha$, we get a relatively balanced system with real-parts of the spectrum ranging from -1.5 to -2.5 (as expected from the ULES property). Importantly, when looking at the influence of the gyroscopic noise on the observation errors, we note that noise with DC characteristics will still affect the observation errors, but that this noise is greatly suppressed for higher frequencies. It is also interesting to note that we should pick a lower $\alpha$ if the noise has significant spectral density at higher frequencies, and that it should be picked higher if the noise is of a DC nature. For this tuning, we found that an $\alpha = 0.3$ yielded a good trade-off based on this (and several other) sigma plots. If using the estimator Proposition 2 without filtering, we would have unit amplification across the entire spectrum, whereas low-pass filtering would suppress the noise after a cutoff frequency, but introduce a phase lag in the attitude rate estimate. This is completely avoided
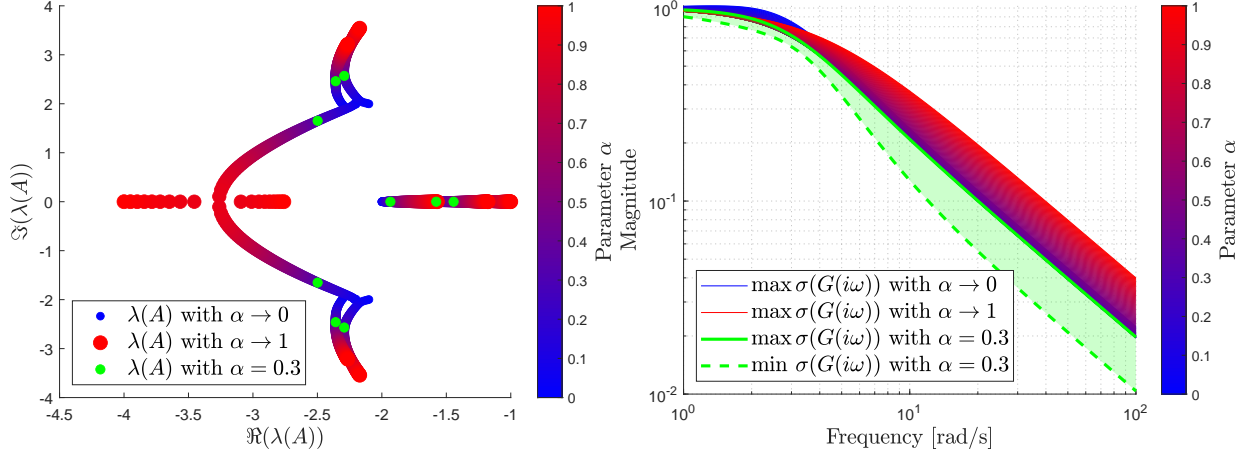
Figure 2: *Top:* Spectrum of the system matrix $A$ governing the local error dynamics as a function of $\alpha$ for the nominal tuning in Sec. 5.1, with $\alpha = 0.3$ marked in green. *Bottom:* Singular values of the error dynamics from the inputs $\delta_0$ to $\tilde{x}$, with the area between the smallest and largest singular value at $\alpha = 0.3$ in green.

with the observer in Proposition 4, where we get the best of both worlds: perfect tracking under ideal conditions, and suppression of the high-frequent measurement noise. This analysis, applied to all of the parameters in turn and selecting combinations yielding an attenuation of the noise-to-state gains gave rise to the tuning in Sec. 5.1.

# 6 Conclusions

In this paper, we first present an observer to estimate the angular momentum of the attitude dynamics without using measurements of angular velocities. We subsequently fuse this observer with a classical result of Mahony, generating an observer that is capable of estimating the attitude, attitude rate, and gyroscopic bias with UaGAS and ULES properties of the resulting error dynamics. Furthermore, we demonstrate that the combined observer has an edge over the two separate observers in terms of the asymptotic observer errors. Specifically, with the combined observer, we get good attenuation of high-frequent measurement noise, obtaining perfect tracking under ideal conditions, and having implicit robustness afforded by the uniform stability properties shown by the Matrosov result.

Importantly, this observer can be used to extend prior work on filtered output feedback in (Lefeber et al., 2020) to a setting in which the gyroscopic biases are estimated and accounted for. This will be done in our future work.

# 7 Acknowledgement

We thank Thor Inge Fossen for inspiring this paper during his visit to Lund University.

# References

Andersson, J., Åkesson, J., and Diehl, M. (2012). CasADi: A symbolic package for automatic differentiation and optimal control. In *Recent advances in algorithmic differentiation*, 297–307. Springer.

Arasaratnam, I. and Haykin, S. (2009). Cubature Kalman filters. *IEEE Trans. on Aut. Cont.*, 54(6), 1254–1269.

Arulampalam, M.S., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing*, 50(2), 174–188.

Berkane, S. and Tayebi, A. (2017). On the design of attitude complementary filters on SO(3). *IEEE Transactions on Automatic Control*, 63(3), 880–887.

Caruso, M., Sabatini, A.M., Laidig, D., Seel, T., Knaflitz, M., Della Croce, U., and Cereatti, A. (2021). Analysis of the accuracy of ten algorithms for orientation estimation using inertial and magnetic sensing under optimal conditions: One size does not fit all. *Sensors*, 21(7), 2543.

Cheng, Y. and Crassidis, J. (2004). Particle filtering for sequential spacecraft attitude estimation. In *AIAA guidance, navigation, and control conf. and exhibit*, 5337.

Farrell, J.L. (1970). Attitude determination by Kalman filtering. *Automatica*, 6(3), 419–430.

Greiff, M. (2021). *Nonlinear Control of Unmanned Aerial Vehicles: Systems With an Attitude*. Lund University.

Johansen, T.A., Hansen, J.M., and Fossen, T.I. (2017). Nonlinear observer for tightly integrated inertial navigation aided by pseudo-range measurements. *Journal of Dynamic Systems, Measurement, and Control*, 139(1), 011007.

Khalil, H. (2002). *Nonlinear Systems*. Prentice-Hall, Upper Saddle River, NJ, USA, 3rd edition.

Lefeber, E., Greiff, M., and Robertsson, A. (2020). Filtered output feedback tracking control of a quadrotor UAV. *IFAC-PapersOnLine*, 53, 5764–5770.

Ligorio, G. and Sabatini, A.M. (2015). A novel Kalman filter for human motion tracking with an inertial-based dynamic inclinometer. *IEEE Transactions on Biomedical Engineering*, 62(8), 2033–2043.

Loría, A., Panteley, E., Popovic, D., and Teel, A.R. (2005). A nested Matrosov theorem and persistency of excitation for uniform convergence in stable nonautonomous systems. *IEEE Trans. on Aut. Cont.*, 50(2), 183–198.

Lu, X., Jia, Y., and Matsuno, F. (2016). Gyro-free attitude observer of rigid body via only time-varying reference vectors. In *American Control Conference*, 4948–4953.

Mahony, R., Hamel, T., and Pflimlin, J.M. (2005). Complementary filter design on the special orthogonal group SO(3). In *Proceedings of the 44th IEEE Conference on Decision and Control*, 1477–1484. IEEE.

Mahony, R., Hamel, T., and Pflimlin, J.M. (2008). Nonlinear complementary filters on the special orthogonal group. *IEEE Trans. on Aut. Cont.*, 53(5), 1203–1218.

Mahony, R., Kumar, V., and Corke, P. (2012). Multirotor aerial vehicles: Modeling, estimation, and control of quadrotor. *IEEE Robotics and Aut. mag.*, 19(3), 20–32.

Mahony, R., van Goor, P., and Hamel, T. (2022). Observer design for nonlinear systems with equivariance. *Annual Review of Control, Robotics, and Autonomous Systems*, 5, 221–252.

Markley, F.L., Crassidis, J., and Cheng, Y. (2005). Nonlinear attitude filtering methods. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*, 5927.

Ng, Y., van Goor, P., Hamel, T., and Mahony, R. (2020). Equivariant systems theory and observer design for second order kinematic systems on matrix lie groups. In *59th Conf. on Decision and Control*, 4194–4199.

Särkkä, S. (2013). *Bayesian filtering and smoothing*. 3. Cambridge university press.

Wu, T.H. and Lee, T. (2016). Angular velocity observer for attitude tracking on SO(3) with the separation property. *International Journal of Control, Automation and Systems*, 14(5), 1289–1298.

Zamani, M., Trumpf, J., and Mahony, R. (2015). Nonlinear attitude filtering: A comparison study. *arXiv preprint arXiv:1502.03990*.

# A    Supplementary details for Proof 2

In this section, we give supplementary details for Proof 2, defining the constants of the proof as a function of the known parameters and assumed bounds. The maximum and minimum eigenvalues of a real symmetric matrix $J$ are denoted $\bar{\lambda}(J)$ and $\underline{\lambda}(J)$, respectively. Further, $M = \sum_{i=1}^{n} k_i v_i v_i^\top$ with distinct eigenvalues $\lambda_i$, i.e., $\lambda_3 > \lambda_2 > \lambda_1 > 0$. Let $D = \text{diag}(r_2\lambda_2 + r_3\lambda_3, r_1\lambda_1 + r_3\lambda_3, r_1\lambda_1 + r_2\lambda_2)$ with appropriate $r_i \in \{-1, 1\}$, and take $\bar{D} = R^\top D R$.

## A.1    The inequality in (18b)

We start by showing the first inequality in the context of Proof 2. Here: $\tilde{r}_k$ is bounded by definition; $\tilde{\ell}$ is bounded for all times as the Lyapunov function is negative semi-definite along the solutions of the error dynamics; the attitude rates $\omega$ and accelerations $\dot{\omega}$ are both bounded by assumption. In summary, $\exists K_i > 0$ for $i = 1, ..., 4$, such that

$$\|\tilde{r}_k\|_2 \leq \Big( \sum_{i=1}^{n} k_i^2 \Big)^{1/2} \triangleq K_1, \quad \|\tilde{\ell}\|_2 \leq \Big( 2V_1(0) - k_\ell \sum_{i=1}^{n} k_i \|\tilde{R}v_i - v_i\|_2^2 \Big)^{1/2} \triangleq K_2, \quad \|\omega\|_2 \leq K_3, \quad \|\dot{\omega}\|_2 \leq K_4.$$

Thus, $\dot{\tilde{\ell}}$ is bounded,

$$\|\dot{\tilde{\ell}}\|_2 \leq \|J\dot{\omega}\|_2 + \|S(J\omega)\omega)\|_2 + k_\ell \|J^{-1}\tilde{r}_k\|_2 \leq \bar{\lambda}(J)K_4 + \bar{\lambda}(J)K_3^2 + k_\ell \underline{\lambda}(J)^{-1}K_1 \triangleq K_5$$

Let $\tilde{r}_k = R^\top r_k$, then

$$\tilde{r}_k = \sum_{i=1}^n k_i S(\hat{R}^\top v_i) R^\top v_i \underset{(3d)}{=} R^\top \sum_{i=1}^n k_i S(R\hat{R}^\top v_i) v_i \underset{(10)}{=} R^\top \sum_{i=1}^n k_i S(\tilde{R}^\top v_i) v_i = R^\top r_k. \tag{42}$$

In light of Lemma 2, $\|D\|_2 = \|\bar{D}\|_2 = K_6$ and this constant is known in the observer tuning $\{(k_i, v_i)\}_{i=1}^N$. Now,

$$\dot{r}_k = \sum_{i=1}^n k_i S(\dot{\hat{R}}^\top v_i) v_i \underset{(5)}{=} -DR(J^{-1} R^\top \tilde{\ell} - k_R \tilde{r}_k) \qquad \Rightarrow \qquad \|\dot{r}_k\|_2 \le K_6(\underline{\lambda}(J)^{-1} K_2 + k_R K_1) \triangleq K_7. \tag{43}$$

By the chain rule

$$\dot{\tilde{r}}_k = S(\omega)^\top R^\top r_k - R^\top \dot{r}_k \quad \Rightarrow \quad \|\dot{\tilde{r}}_k\|_2 \le K_3 K_1 + K_7 \triangleq K_8. \tag{44}$$

Differentiating (44), once more

$$\ddot{\tilde{r}}_k = (S(\omega)^2 + S(\dot{\omega})^\top) R^\top r_k + S(\omega)^\top R^\top \dot{r}_k \tag{45}$$
$$+ [S(\omega)^\top \bar{D} + \bar{D} S(\omega)](J^{-1} R^\top \tilde{\ell} - k_R \tilde{r}_k) + \bar{D}(J^{-1} S(\omega)^\top R^\top \tilde{\ell} - k_R \dot{\tilde{r}}_k) \quad \Rightarrow$$
$$\|\ddot{\tilde{r}}_k\|_2 \le K_1(K_3^2 + K_4) + K_3 K_7 + 2K_3 K_6(\bar{\lambda}(J^{-1})K_2 + k_R K_1) + K_6(\bar{\lambda}(J^{-1})K_2 K_3 + k_R K_7) \triangleq K. \tag{46}$$

This is the constant $K > 0$ that appears in (18b), which is computable in the initial errors and tuning parameters.

## A.2  The inequality in (18c)

The second inequality follows directly from (44). We obtain

$$-\|\dot{\tilde{r}}_k\|_2^2 = -r_k^\top R S(\omega) S(\omega) R^\top r_k + 2r_k^\top R^\top \dot{r}_k - \|\dot{r}_k\|_2^2 \tag{47a}$$
$$\le K_3^2 \|\tilde{r}_k\|_2^2 + 2K_8 \|\tilde{r}_k\|_2 - \|DR(J^{-1} R^\top \tilde{\ell} - k_R \tilde{r}_k)\|_2^2 \tag{47b}$$
$$\le K_3^2 \|\tilde{r}_k\|_2^2 + 2K_8 \|\tilde{r}_k\|_2 - \tilde{\ell}^\top R J^{-1} R^\top D^2 R J^{-1} R^\top \tilde{\ell} + 2k_R K_2 \bar{\lambda}(J^{-1}) K_6 \|\tilde{r}_k\|_2 + k_R^2 \|\tilde{r}_k\|_2^2 \tag{47c}$$
$$\triangleq -\gamma \|\tilde{\ell}\|_2^2 + K_9 \|\tilde{r}_k\|_2 \tag{47d}$$

thus

$$-\|\dot{\tilde{r}}_k\|_2^2 + K\|\tilde{r}_k\|_2 \le -\gamma \|\tilde{\ell}\|_2^2 + K_9 \|\tilde{r}_k\|_2 + K\|\tilde{r}_k\|_2 \triangleq -\gamma \|\tilde{\ell}\|_2^2 + \bar{K}\|\tilde{r}_k\|_2, \tag{48}$$

where

$$\gamma = \inf_{R \in SO(3)} \underline{\lambda}(R J^{-1} R^\top D^2 R J^{-1} R^\top), \qquad \bar{K} = K_3^2 K_1 + 2K_8 + 2k_R K_2 \bar{\lambda}(J^{-1}) K_6 + k_R^2 K_1, \tag{49}$$

are the constants in (18c). Note that $D$ may be indefinite, but $D^2$ is positive definite, thus $\gamma > 0$ is a positive constant.

## A.3  Proof of ULES

The proof of ULES uses the main ideas in (Wu and Lee, 2016) and consists of three main steps:

(i) Show that for sufficiently small $\bar{\epsilon} > 0$, there exists positive constants $c_1, c_2$ for which

$$c_1 \|\tilde{r}_k\|_2^2 \le k_\ell \sum_{i=1}^n \frac{k_i}{2} \|\tilde{R} v_i - v_i\|_2^2 \le c_2 \|\tilde{r}_k\|_2^2, \qquad \forall \|\tilde{r}_k\|_2 \le \bar{\epsilon}. \tag{50}$$

(ii) As $\dot{V}_1 \le 0$, let $\tilde{\omega} = J^{-1} R^\top \tilde{\ell}$, and that $\exists \, \epsilon_1 > 0$ defining a set $\mathcal{S}(\epsilon_1) = \{(\tilde{r}_k, \tilde{\omega}) \in \mathbb{R}^6 \mid V_1 \le \epsilon_1\}$ for which

$$\sup_{(\tilde{r}_k, \tilde{\omega}) \in \mathcal{S}(\epsilon_1)} \|\tilde{r}_k\|_2 \le \bar{\epsilon}. \tag{51}$$

(iii) Finally, define $z^\top = (\tilde{r}_k^\top, \tilde{\omega}^\top) \in \mathbb{R}^6$ and consider a composite function $V_3 = V_1 + \epsilon V_2$ for some small $\epsilon > 0$. Show that for sufficiently small $\epsilon$, there exists positive definite matrices $M_1, M_2, W$, such that

$$z^\top M_1 z \le V_3 \le z^\top M_2 z, \qquad \dot{V}_3 \le -z^\top W z, \qquad \forall z \in \mathcal{S}(\epsilon). \tag{52}$$

ULES of the origin $z = 0$ then follows by application of Khalil (2002, Th. 4.10).

### A.3.1  Details on step (iii)

The first two steps are straight-forward, but some additional details are provided for step (iii). Recall that

$$\dot{\tilde{r}}_k = S(\omega)^\top \tilde{r}_k - R^\top \dot{r}_k = S(\omega)^\top \tilde{r}_k + \bar{D}(\tilde{\omega} - k_R \tilde{r}_k), \tag{53}$$

thus

$$V_2 = -\tilde{r}_k^\top [S(\omega)^\top \tilde{r}_k + \bar{D}(\tilde{\omega} - k_R \tilde{r}_k)] = -\tilde{r}_k^\top \bar{D}\tilde{\omega} + k_R \tilde{r}_k^\top \bar{D}\tilde{r}_k, \tag{54}$$

and we obtain

$$z^\top \underbrace{\left( \begin{bmatrix} c_1 I & 0 \\ 0 & J^2 \end{bmatrix} + \begin{bmatrix} \epsilon k_R \bar{D} & -\frac{1}{2}\epsilon\bar{D} \\ -\frac{1}{2}\epsilon\bar{D} & 0 \end{bmatrix} \right)}_{\triangleq M_1} z \le V_3 \le z^\top \underbrace{\left( \begin{bmatrix} c_2 I & 0 \\ 0 & J^2 \end{bmatrix} + \begin{bmatrix} \epsilon k_R \bar{D} & -\frac{1}{2}\epsilon\bar{D} \\ -\frac{1}{2}\epsilon\bar{D} & 0 \end{bmatrix} \right)}_{\triangleq M_2} z. \tag{55}$$

Sufficient conditions for $M_i > 0$ can be expressed as a bound on $\epsilon$ in $\{c_1, c_2, k_R, D, J\}$. Taking the Schur complement

$$M_i > 0 \Leftrightarrow c_i I + \epsilon k_R \bar{D} > 0 \ \wedge \ J^2 - \tfrac{1}{4}\epsilon\bar{D}(c_i I + \epsilon k_R \bar{D})^{-1}\epsilon\bar{D} > 0 \Leftarrow 4\lambda(J^2)(c_i I + \epsilon k_R D)D^{-2} - \epsilon^2 > 0, \tag{56}$$

as $D$ is a diagonal matrix and all constants are positive (definite). If we let $D = \text{diag}(d_1, d_2, d_3)$, the condition is

$$\epsilon < \epsilon_2 \triangleq \min_{\substack{i \in \{1,2\} \\ j \in \{1,2,3\}}} \left\{ \frac{2\lambda(J^2)k_R}{d_j} \left( 1 + \sqrt{1 + \frac{c_i}{\lambda(J^2)k_R^2}} \right) \right\}. \tag{57}$$

To find the quadratic form bounding $\dot{V}_3$, we first consider the terms of $\dot{V}_2 = -\|\dot{\tilde{r}}_k\|_2^2 - \tilde{r}_k^\top \ddot{\tilde{r}}_k$. It can be shown that

$$\|\dot{\tilde{r}}_k\|_2^2 = \begin{bmatrix} \tilde{r}_k \\ \tilde{\omega} \end{bmatrix}^\top \begin{bmatrix} k_R^2 I - S(\omega)^2 & (S(\omega) - 2k_R I)\bar{D} \\ \bar{D}(S(\omega)^\top - 2k_R I) & \bar{D}^2 \end{bmatrix} \begin{bmatrix} \tilde{r}_k \\ \tilde{\omega} \end{bmatrix}. \tag{58}$$

Furthermore, taking the time-derivative of (53), we get

$$\ddot{\tilde{r}}_k = S(\dot{\omega})^\top \tilde{r}_k + S(\omega)^\top \dot{\tilde{r}}_k + \dot{\bar{D}}(\tilde{\omega} - k_R \tilde{r}_k) + \bar{D}(\dot{\tilde{\omega}} - k_R \dot{\tilde{r}}_k), \tag{59}$$

and

$$\dot{\tilde{\omega}} = J^{-1}S(\omega)^\top J\tilde{\omega} - k_\ell J^{-2}\tilde{r}_k. \tag{60}$$

With these expressions, one can obtain

$$\tilde{r}_k^\top \ddot{\tilde{r}}_k = \begin{bmatrix} \tilde{r}_k \\ \tilde{\omega} \end{bmatrix}^\top \begin{bmatrix} S(\omega)^2 + \bar{W}_{11} & \bar{W}_{12} \\ \star & 0 \end{bmatrix} \begin{bmatrix} \tilde{r}_k \\ \tilde{\omega} \end{bmatrix}, \tag{61}$$

where bounds on $\|\bar{W}_{11}\|_2$ and $\|\bar{W}_{12}\|_2$ can be expressed $J$, $D$, and the assumed bound of $\omega$. Now, we have that

$$\tilde{r}_k^\top \ddot{\tilde{r}}_k + \|\dot{\tilde{r}}_k\|_2^2 = \begin{bmatrix} \tilde{r}_k \\ \tilde{\omega} \end{bmatrix}^\top \begin{bmatrix} k_R^2 I + \bar{W}_{11} & (S(\omega) - 2k_R I)\bar{D} + \bar{W}_{12} \\ \star & \bar{D}^2 \end{bmatrix} \begin{bmatrix} \tilde{r}_k \\ \tilde{\omega} \end{bmatrix} \triangleq \begin{bmatrix} \tilde{r}_k \\ \tilde{\omega} \end{bmatrix}^\top \begin{bmatrix} \bar{\bar{W}}_{11} & \bar{\bar{W}}_{12} \\ \star & \bar{D}^2 \end{bmatrix} \begin{bmatrix} \tilde{r}_k \\ \tilde{\omega} \end{bmatrix}, \tag{62}$$

where similarly $\|\bar{\bar{W}}_{11}\|_2$ and $\|\bar{\bar{W}}_{12}\|_2$ are bounded. Subsequently

$$\dot{V}_3 = \dot{V}_1 + \epsilon\dot{V}_2 = -k_\ell k_R\|\tilde{r}_k\|_2^2 - \epsilon(\tilde{r}_k^\top \ddot{\tilde{r}}_k + \|\dot{\tilde{r}}_k\|_2^2) = -\begin{bmatrix} \tilde{r}_k \\ \tilde{\omega} \end{bmatrix}^\top \underbrace{\epsilon\begin{bmatrix} \epsilon^{-1}k_\ell k_R I + \bar{\bar{W}}_{11} & \bar{\bar{W}}_{12} \\ \star & \bar{D}^2 \end{bmatrix}}_{\triangleq W} \begin{bmatrix} \tilde{r}_k \\ \tilde{\omega} \end{bmatrix}. \tag{63}$$

As such, we get a conservative but sufficient condition for $W > 0$ by picking a sufficiently small $\epsilon > 0$. Specifically,

$$W > 0 \Leftrightarrow \bar{D}^2 > 0 \ \wedge \ \epsilon^{-1}k_\ell k_R I + \bar{\bar{W}}_{11} - \bar{\bar{W}}_{12}^\top \bar{D}^{-2}\bar{\bar{W}}_{12} > 0 \Leftarrow \epsilon < \epsilon_3 \triangleq k_\ell k_R(\|\bar{\bar{W}}_{12}^\top\|_2\|\bar{D}^{-2}\|_2\|\bar{\bar{W}}_{12}\|_2 + \|\bar{\bar{W}}_{11}\|_2)^{-1}. \tag{64}$$

Using (Khalil, 2002, Th. 4.10) on $\mathcal{S}(\epsilon)$ with any $\epsilon < \min\{\epsilon_1, \epsilon_2, \epsilon_3\}$ from (51), (57) and (64) completes the proof.

## B   Parameter Distributions used in the Monte Carlo Simulations

In this section, we let $\mathcal{N}(x; \mu, \Sigma)$ denote a Gaussian probability density function (PDF) in $x$ with mean $\mu \in \mathbb{R}^n$ and covariance $\Sigma \in \mathbb{R}^{n \times n}$. We let $\mathcal{U}(x; I)$ be a uniform PDF in $x$ that samples every element of a closed interval $I \subset \mathbb{R}^n$, uniformly and independently in each dimension. In practice, we accomplish this for SO(3) by drawing an un-normalized quaternion $\mathcal{N}(q; 0, I)$, normalizing this, and embedding it in SO(3). We refer to this as $\mathcal{U}(R; \text{SO(3)})$.

The parameters $\theta = \{R_0, \omega_0, b, \hat{R}_0, \hat{\omega}_0, \hat{b}_0\}$ are sampled from a distribution with probability density function

$$p(\theta) = \mathcal{U}(R_0; \text{SO(3)}) \, \mathcal{N}(b_0; 0, I) \, \mathcal{N}(\omega_0; 0, 0.1I) \mathcal{U}(\hat{R}_0; \text{SO(3)}) \, \mathcal{N}(\hat{b}_0; 0, I) \, \mathcal{N}(\hat{\ell}_0; 0, I),$$

the inertia $J$ is constructed by sampling a random symmetric positive semi-definite matrix $J_A$ with spectrum $\{\lambda_1, \lambda_2, \lambda_3\}$, where $0 = \lambda_1 \leq \lambda_2 \leq \lambda_3 = 1$, and letting $J = 0.5(J_A + I)$. We let $v_1 = (0, 0, -1)^\top$, sample $\mathcal{N}(\bar{v}_2; 0, I)$, setting its last element to -0.1, and normalizing it, such that $v_2 = \bar{v}_2 / \|\bar{v}_2\|_1$. We then take $v_3 = v_1 \times v_2$.

In the ideal setting (Sec. 5.1), the outputs in (7) are sampled continuously without noise, and we run the simulation with a fixed-point RK4 solver over $t \in [0, 10]$ seconds.

In the Monte Carlo simulations, we add noise terms $n_i$, as

$$
\begin{aligned}
y_0(hk) &= \omega(hk) + b(hk) + n_0(hk) \\
\bar{y}_i(hk) &= R(hk)^\top v_i + n_i(hk) & i &= 1, \dots, n, \\
y_i(hk) &= \bar{y}_i(hk) / \|\bar{y}_i(hk)\|_2 & i &= 1, \dots, n,
\end{aligned}
$$

and we take this noise to be zero-mean Gaussian distributed, uncorrelated, sampled from $\mathcal{N}(n_i(hk); 0, 0.01I)$. We sample these outputs at a rate of 500 Hz (i.e. $h = 0.002$ s), but run the observer prediction at a rate of 1 kHz.

## C   An Equivalent Discrete-Time Quaternion Formulation

Just as in (Mahony et al., 2008, Appendix B), it is straightforward to give an equivalent representation of the filters when integrating the attitude as a quaternion. The set of quaternions is $\mathbb{H} = \{q = (q_w, q_v) \in \mathbb{R} \times \mathbb{R}^3 : |q| = 1\}$, and we use the Hamilton construction with quaternion representing a right-handed rotation (see, e.g., (Greiff, 2021)). The group $\mathbb{H}$ is 2-to-1 homomorphic to SO(3),

$$E : \mathbb{H} \mapsto \text{SO(3)}, \quad E(q) = (q_w^2 - q_v^\top q_v)I + 2q_v q_v^\top + 2q_w S(q_v).$$

The attitude kinematics of the quaternion, i.e., the differential equation preserving $q(t + \delta) \in \mathbb{H}$ for $\delta \geq 0$ is

$$\dot{q} = \frac{1}{2}Q(q)\begin{bmatrix} 0 \\ \omega \end{bmatrix}, \quad Q(q) = Iq_w + \begin{bmatrix} q_w & -q_v^\top \\ q_v & S(q_v) \end{bmatrix}.$$

As such, to implement an observer with an quaternion attitude representation $\hat{q}(t)$ such that $\hat{R}(t) = E(\hat{q}(t))$, we only need to replace (35b) in Proposition 4 by

$$\dot{\hat{q}} = \frac{1}{2}Q(\hat{q})\begin{bmatrix} 0 \\ \alpha J^{-1}\tilde{\delta}_L + y_0 - \hat{b} - k_R\tilde{r}_k \end{bmatrix},$$

and modify the computation of $\tilde{r}_k$ in (14b) as

$$\tilde{r}_k = \sum_{i=1}^{n} k_i S(E(\hat{q})^\top v_i) y_i.$$

In any practical implementation, the observer update would need to be discretized. Here, a sufficiently slow update rate with a sufficiently simple discretization will lead to numerical artifacts that may become a dominating factor in the noise floor of the error dynamics. Instead of a forward Euler scheme, as is commonly used in practice, we suggest the use of an RK scheme of higher order with a projection onto $\mathbb{H}$ on each time step, or a Crouch-Grossman integrator which performs the integration directly on $\mathbb{H}$ (refer to the discussion in (Greiff, 2021, Chapter 2.4)).

To simplify implementations of the theoretical results, we include the observer as Matlab code with a fixed-step RK4 integrator. Proposition 4 can be implemented as:

```matlab
function Xkp1 = obs_ODE(Xk,y0,y1,y2,y3,tau)
    % Define observer parameters
    v1 = ; v2 = ; v3 = ; k1 = ; k2 = ; k3 = ;
    J = ; kr = ; kl = ; ka = ; kb = ; alpha = ;
    % Required functions
    S = @(u) [    0,-u(3), u(2); u(3),    0,-u(1);  -u(2), u(1),    0];
    Q = @(q) eye(4).*q(1) + [q(1),-q(2:4)'; q(2:4),  S(q(2:4))]];
    E = @(q) (q(1)^2-q(2:4)'*q(2:4))*eye(3)+2*q(2:4)*q(2:4)'+2*q(1)*S(q(2:4));
    % Process arguments
    R = (k1*v1*v1' + k2*v2*v2' + k3*v3*v3') \ (k1*v1*y1' + k2*v2*y2' + k3*v3*y3');
    bhat = Xk(1:3);
    lhat = Xk(4:6);
    qhat = Xk(7:10);
    % Observer update
    rtilde = k1*S(E(qhat)'*v1)*y1 + k2*S(E(qhat)'*v2)*y2 + k3*S(E(qhat)'*v3)*y3;
    deltaL = R'*lhat - J*(y0 - bhat);
    deltaR = alpha * inv(J) * deltaL + y0 - bhat - kr*rtilde;
    bhatdot = kb*rtilde - alpha * kb * ka * J * deltaL;
    lhatdot = R * (tau - kl*  inv(J) * rtilde -(1-alpha) * kl * ka * deltaL);
    qhatdot = Q(qhat) * [0; deltaR/2];
    Xkp1    = [bhatdot; lhatdot; qhatdot];
end
```

The observer update with a fixed-step RK4 scheme is then:

```matlab
function Xkp1 = update(Xk,h,y0,y1,y2,y3,tau)
    % 4th order Runge-Kutta update
    k1 = obs_ODE(Xk           ,y0,y1,y2,y3,tau);
    k2 = obs_ODE(Xk + h/2*k1,y0,y1,y2,y3,tau);
    k3 = obs_ODE(Xk + h/2*k2,y0,y1,y2,y3,tau);
    k4 = obs_ODE(Xk + h/1*k3,y0,y1,y2,y3,tau);
    Xkp1 = Xk + h/6 * (k1 + 2*k2 + 2*k3 + k4);
    % Projection to a unit quaternion
    Xkp1(7:10) = Xkp1(7:10) /norm(Xkp1(7:10));
end
```

Here, we include a projection onto $\mathbb{H}$ which becomes necessary when tuning the observer with higher gains. For all of the simulations, the attitude was integrated directly on SO(3), but the RK-method above produces identical results and is well suited for practical implementations.