Robustness to measurement noise of a globally convergent attitude observer with topological relaxations

Pedro Batista

Nonlinear Dynamics, vol. 98, no.1, pp. 589-600, October 2019

https://doi.org/10.1007/s11071-019-05214-z

Nonlinear Dyn manuscript No.
(will be inserted by the editor)

Robustness to measurement noise of a globally convergent attitude observer with topological relaxations

Pedro Batista

Received: September 14, 2018

Abstract In the past decade substantial effort has been put in the design of attitude observers with bias estimation that present both stability and convergence guarantees. However, most theoretical results consider a noiseless setting, deferring the analysis of the effect of noise to simulation and experimental results. This paper addresses the robustness of an existing solution to noise in all measurements by considering two settings: i) bounded noise; and ii) stochastic noise modeled by a Wiener process. The results are appealing in that they effectively show the robustness of the observer in both scenarios, thus complementing global exponential convergence in noiseless settings. In particular, for bounded noise, the estimation error remains bounded, whereas in the case of noise modeled by a Wiener process, the mean error converges to zero, with bounded covariance. Finally, an additional result regarding the computation of estimates on $SO(3)$ is also included.

Keywords attitude estimation · robustness · bounded noise · stochastic noise · navigation systems

Funding: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) through ISR under FCT [UID/EEA/50009/2013] and through the FCT project DECENTER [LISBOA-01-0145-FEDER-029605], funded by the Lisboa 2020 and PIDDAC programs.

P. Batista
Institute for Systems and Robotics,
Department of Electrical and Computer Engineering,
Instituto Superior Técnico,
Universidade de Lisboa, Portugal.
Tel.: +351-218418056
E-mail: pbatista@isr.tecnico.ulisboa.pt

1 Introduction

Attitude estimation is of paramount importance to the operation, whether remote or autonomous, of robotic platforms and vehicles. Classic solutions resorted to the EKF, see e.g. [8] for a large survey. More recently, in the last decade, the topic has witnessed great effort and progress, with the appearance of novel attitude estimation solutions that encompass theoretical guarantees of stability and performance. The seminal work [17] proposed nonlinear observers on the Special Orthogonal Group $SO(3)$ with guarantees of almost global stability and convergence, as well as local exponential stability. Attitude estimation in the context of hybrid systems was proposed in [6], which allows for global convergence. Global exponential stability was also achieved in [3] and [4], where the attitude rotation matrix is embedded in $\mathbb{R}^9$, albeit estimates on $SO(3)$ are also provided, maintaining exponential convergence for all initial conditions. Another successful approach of lifting the topological contraints of $SO(3)$ can be found in [11]. The problem of filtering on Lie groups was studied in [1], with application to attitude estimation on $SO(3)$. The problem of attitude estimation received further focus in [23], [13], [14], and [12] resorting to differential geometry tools and exploiting the manifold properties. Attitude estimation with multiple time-varying reference vectors was studied in [10], [2], and [19]. The last two also consider single vector measurements, which is a more demanding setup from a theoretical point of view, a problem that was also addressed in [22].

In most of the aforementioned references the attitude models and analysis assume noiseless scenarios. Moreover, robustness to noise is hardly ever considered, except for simulation results with noise and experiments. In [7] the robustness of an existing observer to bounded measurement disturbances was considered and a novel solution was proposed to overcome some limitations. A very interesting contribution is also presented in [24], where a novel attitude minimum-energy estimator was proposed that considers, in the cost function, unknown measurement errors.

In this paper, the robustness to noise in all measurements of the attitude observer proposed in [3] is characterized considering two different settings: i) in the first, measurement noise is assumed bounded, in a similar fashion to the analysis carried out in [7]; and ii) in the second, a stochastic framework is considered where the noise is modeled by a Wiener process. In order to address both problems, input-to-state stability results are extensively used, as well as known results for stochastic differential equations (SDEs) in the second case. In short, the attitude observer is shown to be robust to noise, as long as it is within certain bounds for the bounded-noise case or as long as the covariance is small enough in the stochastic setting modeled by a Wiener process. Finally, an additional result is also included regarding the computation of estimates directly on $SO(3)$.

1.1 Notation

The symbols $\mathbf{0}$ represents a matrix (or vector) of zeros and $\mathbf{I}_n$ denote the $n \times n$ identity matrix. When n is omitted, the matrices are of appropriate dimensions, which can be inferred from the context. A block diagonal matrix is represented by $\mathbf{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_\mathrm{n})$.

2 Problem statement

2.1 Attitude observer

The cascade attitude observer proposed in [3] is here briefly presented in order to set the grounds for the description of the problem addressed in this work. Let $\{I\}$ denote an inertial reference frame, $\{B\}$ a body-fixed reference frame, and $\mathbf{R} \in SO(3)$ the rotation matrix from $\{B\}$ to $\{I\}$. The attitude kinematics are $\dot{\mathbf{R}} = \mathbf{R}\mathbf{S}(\boldsymbol{\omega})$, where $\boldsymbol{\omega} \in \mathbb{R}^3$ is the angular velocity of $\{B\}$, expressed in $\{B\}$, and $\mathbf{S}(.)$ is the skew-symmetric matrix that encodes the cross product, i.e., $\mathbf{S}(\mathbf{x})\mathbf{y} = \mathbf{x} \times \mathbf{y}$, $\mathbf{x} \in \mathbb{R}^3$, $\mathbf{y} \in \mathbb{R}^3$.

In order to estimate $\mathbf{R}$, biased angular velocity measurements $\boldsymbol{\omega}_m = \boldsymbol{\omega} + \mathbf{b}_{\boldsymbol{\omega}} \in \mathbb{R}^3$, provided by rate gyros, are assumed available, where $\mathbf{b}_{\boldsymbol{\omega}} \in \mathbb{R}^3$ corresponds to the rate gyro bias, which is assumed constant. Additionally, there is available a set of $N \geq 2$ body-fixed vector observations $\{\mathbf{v}_i \in \mathbb{R}^3, i = 1, \ldots, N\}$ of known constant vectors in inertial coordinates $\{\mathbf{r}_i \in \mathbb{R}^3, i = 1, \ldots, N\}$, which thus satisfy $\mathbf{r}_i = \mathbf{R}\mathbf{v}_i,\ i = 1, \ldots, N$.

In this paper the following assumptions are considered.

Assumption 1 *There exist at least two non-collinear reference vectors, i.e., there exist i and j such that $\mathbf{r}_i \times \mathbf{r}_j \neq \mathbf{0}$.*

Assumption 2 *The signal $\boldsymbol{\omega}_m$ is bounded for all time.*

The first is a standard assumption, necessary for attitude estimation with constant vectors in inertial coordinates, and as such it adds no conservativeness whatsoever. The second is a standard technical assumption in attitude estimation, see e.g. [12], [17], and [22]. It is also verified for all systems in practice, since one cannot have arbitrarily large angular velocities.

The approach proposed in [3] for the observer design resorts to a sensor-based strategy, where the dynamics of the vector observations are explicitly considered. For the sake of completeness, the nominal dynamics of the vector observations are given by

$$\dot{\mathbf{v}}_i = -\mathbf{S}(\boldsymbol{\omega}_m)\mathbf{v}_i - \mathbf{S}(\mathbf{v}_i)\mathbf{b}_{\boldsymbol{\omega}},$$

$i = 1, \ldots, N$, see [3].

The attitude observer proposed in [3] has a cascade structure. The first block is given by

$$\begin{cases} \dot{\hat{\mathbf{v}}}_1 = -\mathbf{S}(\boldsymbol{\omega}_m)\hat{\mathbf{v}}_1 - \mathbf{S}(\mathbf{v}_1)\hat{\mathbf{b}}_{\boldsymbol{\omega}} + \alpha_1[\mathbf{v}_1 - \hat{\mathbf{v}}_1] \\ \quad\vdots \\ \dot{\hat{\mathbf{v}}}_N = -\mathbf{S}(\boldsymbol{\omega}_m)\hat{\mathbf{v}}_N - \mathbf{S}(\mathbf{v}_N)\hat{\mathbf{b}}_{\boldsymbol{\omega}} + \alpha_N(\mathbf{v}_N - \hat{\mathbf{v}}_N) \\ \dot{\hat{\mathbf{b}}}_{\boldsymbol{\omega}} = -\sum_{i=1}^{N} \beta_i \mathbf{S}(\mathbf{v}_i)\hat{\mathbf{v}}_i \end{cases}, \quad (1)$$

where $\hat{\mathbf{v}}_i \in \mathbb{R}^3$ is an estimate of $\mathbf{v}_i$, $i = 1, \ldots, N$, $\hat{\mathbf{b}}_{\boldsymbol{\omega}} \in \mathbb{R}^3$ is an estimate of the rate gyro bias, and $\alpha_i \in \mathbb{R}^+, \beta_i \in \mathbb{R}^+$, $i = 1, \ldots, N$, are observer parameters. Define now

$$\mathbf{C}_2 := \begin{bmatrix} r_{11}\mathbf{I}_3 & r_{12}\mathbf{I}_3 & r_{13}\mathbf{I}_3 \\ \vdots & \vdots & \vdots \\ r_{N1}\mathbf{I}_3 & r_{N2}\mathbf{I}_3 & r_{N3}\mathbf{I}_3 \end{bmatrix} \in \mathbb{R}^{3N \times 9},$$

with $\mathbf{r}_i = \begin{bmatrix} r_{i1} & r_{i2} & r_{i3} \end{bmatrix}^T \in \mathbb{R}^3$. The second observer block is given by

$$\dot{\hat{\mathbf{x}}}_2 = -\mathbf{S}_3\left(\boldsymbol{\omega}_m - \hat{\mathbf{b}}_{\boldsymbol{\omega}}\right)\hat{\mathbf{x}}_2 + \mathbf{C}_2^T\mathbf{Q}^{-1}[\mathbf{v} - \mathbf{C}_2\hat{\mathbf{x}}_2], \quad (2)$$

where $\mathbf{Q} = \mathbf{Q}^T \in \mathbb{R}^{3N \times 3N}$ is a tuning matrix,

$$\mathbf{S}_3(\mathbf{x}) := \mathbf{diag}(\mathbf{S}(\mathbf{x}), \mathbf{S}(\mathbf{x}), \mathbf{S}(\mathbf{x})) \in \mathbb{R}^{9 \times 9},\ \mathbf{x} \in \mathbb{R}^3,$$

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_1^T & \ldots & \mathbf{v}_N^T \end{bmatrix}^T \in \mathbb{R}^{3N},$$

and $\hat{\mathbf{x}}_2$ is an estimate of $\mathbf{x}_2$, which corresponds to a column representation of $\mathbf{R}$, as given by

$$\mathbf{x}_2 = \begin{bmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \\ \mathbf{z}_3 \end{bmatrix} \in \mathbb{R}^9,$$

where

$$\mathbf{R} = \begin{bmatrix} \mathbf{z}_1^T \\ \mathbf{z}_2^T \\ \mathbf{z}_3^T \end{bmatrix}, \ \mathbf{z}_i \in \mathbb{R}^3, \ i = 1, \ldots, 3.$$

Notice that the nominal dynamics of $\mathbf{x}_2$ are given by

$$\dot{\boldsymbol{x}}_2 = -\mathbf{S}_3\left(\boldsymbol{\omega}_m - \mathbf{b}_{\boldsymbol{\omega}}\right)\boldsymbol{x}_2$$

and that the vector observations can be expressed as

$$\mathbf{v} = \mathbf{C}_2\boldsymbol{x}_2,$$

see [3].

Define the estimation error of the first observer as

$$\tilde{\mathbf{x}}_1 := \begin{bmatrix} \tilde{\mathbf{v}}_1 \\ \vdots \\ \tilde{\mathbf{v}}_N \\ \tilde{\mathbf{b}}_{\boldsymbol{\omega}} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 - \hat{\mathbf{v}}_1 \\ \vdots \\ \mathbf{v}_N - \hat{\mathbf{v}}_N \\ \mathbf{b}_{\boldsymbol{\omega}} - \hat{\mathbf{b}}_{\boldsymbol{\omega}} \end{bmatrix} \in \mathbb{R}^{3(N+1)},$$

and the estimation error of the second observer as $\tilde{\mathbf{x}}_2 := \mathbf{x}_2 - \hat{\mathbf{x}}_2$. It is a matter of computation to show that the error dynamics can be written as

$$\dot{\tilde{\mathbf{x}}}_1 = \mathbf{A}_1\tilde{\mathbf{x}}_1 \tag{3}$$

and

$$\dot{\tilde{\mathbf{x}}}_2 = \mathbf{A}_2\tilde{\mathbf{x}}_2 + \mathbf{S}_3\left(\tilde{\mathbf{b}}_{\boldsymbol{\omega}}\right)\mathbf{x}_2 - \mathbf{S}_3\left(\tilde{\mathbf{b}}_{\boldsymbol{\omega}}\right)\tilde{\mathbf{x}}_2, \tag{4}$$

with

$$\mathbf{A}_1 = -\mathbf{diag}\left(\alpha_1\mathbf{I} + \mathbf{S}\left(\boldsymbol{\omega}_{\mathrm{m}}\right), \ldots, \alpha_{\mathrm{N}}\mathbf{I} + \mathbf{S}\left(\boldsymbol{\omega}_{\mathrm{m}}\right), \mathbf{0}\right) + \begin{bmatrix} \mathbf{0} & \ldots & \mathbf{0} & -\mathbf{S}\left(\mathbf{v}_1\right) \\ \vdots & & \vdots & \vdots \\ \mathbf{0} & \ldots & \mathbf{0} & -\mathbf{S}\left(\mathbf{v}_N\right) \\ -\beta_1\mathbf{S}\left(\mathbf{v}_1\right) & \ldots & -\beta_N\mathbf{S}\left(\mathbf{v}_N\right) & \mathbf{0} \end{bmatrix}$$

and $\mathbf{A}_2 = -\mathbf{C}_2^T\mathbf{Q}^{-1}\mathbf{C}_2 - \mathbf{S}_3\left(\boldsymbol{\omega}_m - \mathbf{b}_{\boldsymbol{\omega}}\right)$.

The following result, which will be required in this paper, is shown in [3].

Theorem 1 *Under Assumptions 1 and 2, consider the rate gyro bias observer* (1)*, where* $\alpha_i > 0$*,* $\beta_i > 0$*,* $i = 1, \ldots, N$*, are positive scalar parameters. Then, the origin of the observer error dynamics* (3) *is a globally exponentially stable equilibrium point.*

To establish global exponential stability of the error dynamics of the second observer, the following assumption is introduced.

Assumption 3 *The matrix* $[\mathbf{r}_1 \ldots \mathbf{r}_N] \in \mathbb{R}^{3\times 3N}$ *has full rank.*

It is important to stress that, in practice, this assumption does not impose the necessity of any additional measurements since it is possible and computationally inexpensive to obtain a set of reference vectors $\mathbf{r}_i$ and corresponding observations in body-fixed coordinates $\mathbf{v}_i$ such that Assumption 3 is satisfied, provided that Assumption 1 holds, see [3].

2.2 Observer framework in the presence of measurement noise

The analysis that was derived in [3] considers only a noiseless scenario. In practice, all measurements are obtained by sensors and are thus corrupted by noise. Here, noise on all measurements will be considered. For the rate gyros, instead of $\boldsymbol{\omega}_m$ in (1) and (2), one should consider $\boldsymbol{\omega}_m + \mathbf{n}_{\boldsymbol{\omega}}$, where $\mathbf{n}_{\boldsymbol{\omega}} \in \mathbb{R}^3$ accounts for the rate gyros measurement noise. Likewise, for the vector observations, instead of $\mathbf{v}_i$ in (1) and (2), $i = 1, \ldots, N$, one should consider $\mathbf{v}_i + \mathbf{n}_i$, where $\mathbf{n}_i \in \mathbb{R}^3$ accounts for the measurement noise of the body-fixed observations. Thus, the resulting observer dynamics in the presence of measurement noise are given by

$$\begin{cases} \dot{\hat{\mathbf{v}}}_1 = -\mathbf{S}\left(\boldsymbol{\omega}_m + \mathbf{n}_{\boldsymbol{\omega}}\right)\hat{\mathbf{v}}_1 - \mathbf{S}\left(\mathbf{v}_1 + \mathbf{n}_1\right)\hat{\mathbf{b}}_{\boldsymbol{\omega}} \\ \qquad +\alpha_1\left[\mathbf{v}_1 + \mathbf{n}_1 - \hat{\mathbf{v}}_1\right] \\ \vdots \\ \dot{\hat{\mathbf{v}}}_N = -\mathbf{S}\left(\boldsymbol{\omega}_m + \mathbf{n}_{\boldsymbol{\omega}}\right)\hat{\mathbf{v}}_N - \mathbf{S}\left(\mathbf{v}_N + \mathbf{n}_N\right)\hat{\mathbf{b}}_{\boldsymbol{\omega}} \\ \qquad +\alpha_N\left[\mathbf{v}_N + \mathbf{n}_N - \hat{\mathbf{v}}_N\right] \\ \dot{\hat{\mathbf{b}}}_{\boldsymbol{\omega}} = -\sum_{i=1}^{N}\beta_i\mathbf{S}\left(\mathbf{v}_i + \mathbf{n}_i\right)\hat{\mathbf{v}}_i \end{cases} \tag{5}$$

and

$$\dot{\hat{\mathbf{x}}}_2 = -\mathbf{S}_3\left(\boldsymbol{\omega}_m + \mathbf{n}_{\boldsymbol{\omega}} - \hat{\mathbf{b}}_{\boldsymbol{\omega}}\right)\hat{\mathbf{x}}_2 + \mathbf{C}_2^T\mathbf{Q}^{-1}\left[\mathbf{v} + \mathbf{n}_{\mathbf{v}} - \mathbf{C}_2\hat{\mathbf{x}}_2\right], \tag{6}$$

with

$$\mathbf{n}_{\mathbf{v}} = \begin{bmatrix} \mathbf{n}_1^T \\ \ldots \\ \mathbf{n}_N^T \end{bmatrix}^T \in \mathbb{R}^{3N}.$$

In turn, it is a matter of computation to show that the observer error dynamics in the presence of measurement noise can be written as

$$\dot{\tilde{\mathbf{x}}}_1 = \mathbf{A}_1\tilde{\mathbf{x}}_1 + \left[\mathbf{B}_{11} + \mathbf{B}_{12}\left(t, \tilde{\mathbf{x}}_1\right)\right]\mathbf{n} \tag{7}$$

and

$$\dot{\tilde{\mathbf{x}}}_2 = \mathbf{A}_2\tilde{\mathbf{x}}_2 + \mathbf{S}_3\left(\tilde{\mathbf{b}}_{\boldsymbol{\omega}}\right)\left[\mathbf{x}_2 - \tilde{\mathbf{x}}_2\right] + \mathbf{S}_3\left(\mathbf{n}_{\boldsymbol{\omega}}\right)\left[\mathbf{x}_2 - \tilde{\mathbf{x}}_2\right] - \mathbf{C}_2^T\mathbf{Q}^{-1}\mathbf{n}_{\mathbf{v}}, \tag{8}$$

where

$$\mathbf{n} = \begin{bmatrix} \mathbf{n}_{\mathbf{v}} \\ \mathbf{n}_{\boldsymbol{\omega}} \end{bmatrix} \in \mathbb{R}^{3(N+1)},$$

$$\begin{aligned}\mathbf{B}_{11} = &-\mathbf{diag}\left(\alpha_1 \mathbf{I} + \mathbf{S}\left(\mathbf{b}_{\boldsymbol{\omega}}\right), \ldots, \alpha_{\scriptscriptstyle N} \mathbf{I} + \mathbf{S}\left(\mathbf{b}_{\boldsymbol{\omega}}\right), \mathbf{0}\right) \\ &+ \begin{bmatrix} \mathbf{0} & \ldots & \mathbf{0} & -\mathbf{S}\left(\mathbf{v}_1\right) \\ \vdots & & \vdots & \vdots \\ \mathbf{0} & \ldots & \mathbf{0} & -\mathbf{S}\left(\mathbf{v}_{\scriptscriptstyle N}\right) \\ -\beta_1 \mathbf{S}\left(\mathbf{v}_1\right) & \ldots & -\beta_{\scriptscriptstyle N} \mathbf{S}\left(\mathbf{v}_{\scriptscriptstyle N}\right) & \mathbf{0} \end{bmatrix},\end{aligned}$$

and

$$\begin{aligned}\mathbf{B}_{12}\left(t, \tilde{\mathbf{x}}_1\right) = &\,\mathbf{diag}\left(\mathbf{S}\left(\tilde{\mathbf{b}}_{\boldsymbol{\omega}}\right), \ldots, +\mathbf{S}\left(\tilde{\mathbf{b}}_{\boldsymbol{\omega}}\right), \mathbf{0}\right) \\ &+ \begin{bmatrix} \mathbf{0} & \ldots & \mathbf{0} & \mathbf{S}\left(\tilde{\mathbf{v}}_1\right) \\ \vdots & & \vdots & \vdots \\ \mathbf{0} & \ldots & \mathbf{0} & \mathbf{S}\left(\tilde{\mathbf{v}}_{\scriptscriptstyle N}\right) \\ \beta_1 \mathbf{S}\left(\tilde{\mathbf{v}}_1\right) & \ldots & \beta_{\scriptscriptstyle N} \mathbf{S}\left(\tilde{\mathbf{v}}_{\scriptscriptstyle N}\right) & \mathbf{0} \end{bmatrix}.\end{aligned}$$

Notice that, in the absence of measurements noise, i.e., with $\mathbf{n} = \mathbf{0}$, (7) and (8) degenerate in (3) and (4), respectively.

Remark 1 Interestingly enough, the noise terms that are considered here can account for both high frequency noise and other disturbances, such as noise at low frequencies [7]. In the particular case of the rate gyros, the nominal bias, which is assumed constant, is already accounted for in the system dynamics and observer design. Nevertheless, it is common for the bias to change slowly due to, for instance, changes in the operating temperature. These small variations of the bias around the nominal value can be considered as perturbations and are also captured by the term $\mathbf{n}_{\boldsymbol{\omega}}$.

The problem considered in this paper is to analyze the robustness of the attitude observer (5)-(6) to measurement noise. In particular, two distinct settings are considered: i) bounded measurement noise; and ii) measurement noise modeled by a Wiener process.

3 Robustness to bounded measurement noise

The robustness of the observer (5)-(6) to bounded measurement noise is analyzed in this section. First, some preliminary results are provided.

The rate gyro bias observer error dynamics (3) are linear and globally exponentially stable. Moreover, $\mathbf{A}_1$ is norm-bounded under Assumption 2. Hence, there exists a symmetric, continuously differentiable matrix $\mathbf{Q}_1 \in \mathbb{R}^{3(N+1)\times 3(N+1)}$ such that

$$\mathbf{0} \prec \alpha_{11}\mathbf{I} \preceq \mathbf{Q}_1 \preceq \alpha_{12}\mathbf{I} \tag{9}$$

and

$$\mathbf{A}_1^T \mathbf{Q}_1 + \mathbf{Q}_1 \mathbf{A}_1 + \dot{\mathbf{Q}}_1 \preceq -\alpha_{13}\mathbf{I} \prec \mathbf{0}, \tag{10}$$

where α_{11}, α_{12}, and α_{13} are finite positive constants, see [20, Theorem 7.8].

The rate gyro bias is constant. Moreover, since $\mathbf{r}_i$ is constant, it follows that $\mathbf{v}_i$ also has constant norm for all $i = 1, \ldots, N$. Hence, there exists $\beta_{11} > 0$ such that

$$\left\|\mathbf{B}_{11}\right\| \leq \beta_{11} \tag{11}$$

for all $t \geq t_0$. Finally, notice that $\mathbf{B}_{12}\left(t, \tilde{\mathbf{x}}_1\right)$ is linear in the observer error. Hence, there exists $\beta_{12} > 0$ such that

$$\left\|\mathbf{B}_{12}\left(t, \tilde{\mathbf{x}}_1\right)\right\| \leq \beta_{12} \left\|\tilde{\mathbf{x}}_1\right\| \tag{12}$$

for all $t \geq t_0$ and $\tilde{\mathbf{x}}_1 \in \mathbb{R}^{3(N+1)}$.

The following theorem addresses the robustness of the rate gyro bias observer to bounded measurement noise.

Theorem 2 *Consider the rate gyro bias observer* (5) *in the conditions of Theorem 1 and choose* $\lambda_1 \in \mathbb{R}$ *such that*

$$0 < \lambda_1 < \alpha_{13}.$$

Then, the rate gyro bias observer error dynamics (7) *are locally input-to-state stable with respect to the measurement noise* $\mathbf{n}$*, with the domain of attraction characterized by* $\sup_{t \geq t_0} \left\|\mathbf{n}\right\| \leq r_n$ *and* $\mathbf{x}_1 \in \mathbb{R}^{3(N+1)}$*, with*

$$r_n := \frac{\alpha_{13} - \lambda_1}{2\alpha_{12}\beta_{12}}. \tag{13}$$

Proof Consider the Lyapunov-like function

$$V_1\left(t, \tilde{\mathbf{x}}_1\right) := \tilde{\mathbf{x}}_1^T \mathbf{Q}_1 \tilde{\mathbf{x}}_1.$$

Its time derivative is given by

$$\begin{aligned}\dot{V}_1\left(t, \tilde{\mathbf{x}}_1\right) = &\,\tilde{\mathbf{x}}_1^T \dot{\mathbf{Q}}_1 \tilde{\mathbf{x}}_1 + \tilde{\mathbf{x}}_1^T \left[\mathbf{A}_1^T \mathbf{Q}_1 + \mathbf{Q}_1 \mathbf{A}_1\right] \tilde{\mathbf{x}}_1 \\ &+ \tilde{\mathbf{x}}_1^T \mathbf{Q}_1 \left[\mathbf{B}_{11} + \mathbf{B}_{12}\left(t, \tilde{\mathbf{x}}_1\right)\right] \mathbf{n} \\ &+ \mathbf{n}^T \left[\mathbf{B}_{11} + \mathbf{B}_{12}\left(t, \tilde{\mathbf{x}}_1\right)\right]^T \mathbf{Q}_1 \tilde{\mathbf{x}}_1.\end{aligned} \tag{14}$$

Using (9)-(12) in (14) allows to write

$$\begin{aligned}\dot{V}_1\left(t, \tilde{\mathbf{x}}_1\right) \leq &-\left(\alpha_{13} - 2\alpha_{12}\beta_{12}\left\|\mathbf{n}\right\|\right)\left\|\tilde{\mathbf{x}}_1\right\|^2 \\ &+ 2\alpha_{12}\beta_{11}\left\|\tilde{\mathbf{x}}_1\right\|\left\|\mathbf{n}\right\|.\end{aligned}$$

Now, notice that

$$\dot{V}_1\left(t, \tilde{\mathbf{x}}_1\right) \leq -\lambda_1 \left\|\tilde{\mathbf{x}}_1\right\|^2 + 2\alpha_{12}\beta_{11}\left\|\tilde{\mathbf{x}}_1\right\|\left\|\mathbf{n}\right\| \tag{15}$$

for $\|\mathbf{n}\| \leq r_n$. Fix $0 < \theta < 1$ and rewrite (15) as

$$\dot{V}_1(t, \tilde{\mathbf{x}}_1) \leq -\lambda_1 (1-\theta) \|\tilde{\mathbf{x}}_1\|^2 - \lambda_1 \theta \|\tilde{\mathbf{x}}_1\| \left(\|\tilde{\mathbf{x}}_1\| - 2\frac{\alpha_{12}\beta_{11}}{\lambda_1 \theta} \|\mathbf{n}\| \right) \quad (16)$$

for $\|\mathbf{n}\| \leq r_n$. Then, it is clear from (16) that

$$\dot{V}_1(t, \tilde{\mathbf{x}}_1) \leq -\lambda_1 (1-\theta) \|\tilde{\mathbf{x}}_1\|^2$$

for all $\|\mathbf{n}\| \leq r_n$ and $\|\tilde{\mathbf{x}}_1\| \geq 2\frac{\alpha_{12}\beta_{11}}{\lambda_1\theta} \|\mathbf{n}\|$. The proof is completed invoking [15, Theorem 5.2].

Remark 2 This result allows to conclude that, for all initial conditions, as long as the measurement noise is within certain bounds, the estimation error of the rate gyro bias observer will also remain bounded. A simpler way to show local input-to-state stability would be to invoke [15, Lemma 5.4]. However, that would not give the explicit domain of attraction.

The attitude observer (6) depends not only on the sensor measurements but also on the estimates of the rate gyro bias provided by the rate gyro bias observer (5). As such, the observer error dynamics depends on the rate gyro bias estimation error $\tilde{\mathbf{b}}_{\boldsymbol{\omega}}$ and therefore, in order to address the robustness of the attitude observer (6) to bounded measurement noise, one must also consider the rate gyro bias estimation error. This is established in the following theorem.

Theorem 3 *Consider the attitude observer* (6) *and suppose that* $\mathbf{Q}$ *is a positive definite matrix. Assume also that Assumption 3 holds. Then, the attitude observer error dynamics* (8) *are globally input-to-state stable with respect to the input*

$$\mathbf{u}_2 = \begin{bmatrix} \mathbf{n} \\ \tilde{\mathbf{b}}_{\boldsymbol{\omega}} \end{bmatrix} \in \mathbb{R}^{3(N+2)}.$$

Proof Consider the Lyapunov-like function

$$V_2(t, \tilde{\mathbf{x}}_2) = \frac{1}{2} \|\tilde{\mathbf{x}}_2\|^2.$$

Recalling that $\mathbf{S}_3(.)$ is skew-symmetric gives

$$\dot{V}_2(t, \tilde{\mathbf{x}}_2) = -\tilde{\mathbf{x}}_2^T \mathbf{C}_2^T \mathbf{Q}^{-1} \mathbf{C}_2 \tilde{\mathbf{x}}_2 + \tilde{\mathbf{x}}_2^T \left[\mathbf{S}_3 \left(\tilde{\mathbf{b}}_{\boldsymbol{\omega}} + \mathbf{n}_{\boldsymbol{\omega}} \right) \mathbf{x}_2 - \mathbf{C}_2^T \mathbf{Q}^{-1} \mathbf{n}_{\mathbf{v}} \right]. \quad (17)$$

Under Assumption 3 one has that $\mathbf{C}_2$ is a constant matrix with full rank. Moreover, $\mathbf{Q}$ is a positive definite matrix. Hence, $\mathbf{C}_2^T \mathbf{Q}^{-1} \mathbf{C}_2$ is positive definite. Let $\lambda_2 > 0$ denote the minimum eigenvalue of $\mathbf{C}_2^T \mathbf{Q}^{-1} \mathbf{C}_2$. Then, using the Rayleigh-Ritz inequality allows to bound (17) by

$$\dot{V}_2(t, \tilde{\mathbf{x}}_2) \leq -\lambda_2 \|\tilde{\mathbf{x}}_2\|^2 + \tilde{\mathbf{x}}_2^T \mathbf{S}_3 \left(\tilde{\mathbf{b}}_{\boldsymbol{\omega}} \right) \mathbf{x}_2 + \tilde{\mathbf{x}}_2^T \mathbf{S}_3 (\mathbf{n}_{\boldsymbol{\omega}}) \mathbf{x}_2 - \tilde{\mathbf{x}}_2^T \mathbf{C}_2^T \mathbf{Q}^{-1} \mathbf{n}_{\mathbf{v}}. \quad (18)$$

Now recall that $\mathbf{x}_2$ corresponds to a column representation of a rotation matrix, hence $\|\mathbf{x}_2\| = \sqrt{3}$. Moreover, $\|\mathbf{S}_3(\mathbf{x})\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^3$. Finally, $\left\|\tilde{\mathbf{b}}_{\boldsymbol{\omega}}\right\| \leq \|\mathbf{u}_2\|$, $\|\mathbf{n}_{\mathbf{v}}\| \leq \|\mathbf{u}_2\|$, and $\|\mathbf{n}_{\boldsymbol{\omega}}\| \leq \|\mathbf{u}_2\|$. Then, using simple norm inequalities, is is possible to bound (18) by

$$\dot{V}_2(t, \tilde{\mathbf{x}}_2) \leq -\lambda_2 \|\tilde{\mathbf{x}}_2\|^2 + b_2 \|\tilde{\mathbf{x}}_2\| \|\mathbf{u}_2\|, \quad (19)$$

with $b_2 := 2\sqrt{3} + \left\|\mathbf{C}_2^T \mathbf{Q}^{-1}\right\| > 0$. Fix $0 < \theta < 1$ and rewrite (19) as

$$\dot{V}_2(t, \tilde{\mathbf{x}}_2) \leq -\lambda_2 (1-\theta) \|\tilde{\mathbf{x}}_2\|^2 - \lambda_2 \theta \|\tilde{\mathbf{x}}_2\| \left(\|\tilde{\mathbf{x}}_2\| - \frac{b_2}{\lambda_2 \theta} \|\mathbf{u}_2\| \right).$$

Thus, it follows that $\dot{V}_2(t, \tilde{\mathbf{x}}_2) \leq -\lambda_2 (1-\theta) \|\tilde{\mathbf{x}}_2\|^2$ for all $\|\tilde{\mathbf{x}}_2\| \geq \frac{b_2}{\lambda_2 \theta} \|\mathbf{u}_2\|$. The proof is completed invoking [15, Theorem 5.2].

Remark 3 The significance of this theorem is as follows: As long as the measurement noise and the rate gyro bias estimation error remain bounded, the attitude estimation error remains bounded. Moreover, if the first converges to zero, so thus the latter. Interestingly enough, a more elegant proof invoking [15, Lemma 5.5] is not possible since the relevant global Lipchitz property is not verified.

It is important to point out that, while the results may seem trivial at first, that is a consequence of the design and it does not diminish this first contribution: to show robustness of the solution to bounded measurement noise on all sensors. While the noise is often modeled as a Gaussian stochastic process, in practice the measurements are bounded, and so is the noise. Moreover, there exist globally exponentially systems that can be driven to infinity by arbitrarily small additive decaying exponentials, see [21]. Furthermore, in this case, the measurement noise lead not only to additive perturbations in the error dynamics but also multiplicative ones in (7) and (8).

4 Robustness to measurement noise modeled by a Wiener process

The analysis of the robustness of the observer (5)-(6) with the noise modeled by a Wiener process is detailed in this section. First, notice that from (7) it is possible to write the error dynamics of the rate gyro bias observer, in a stochastic setting with measurement noise, as the stochastic differential equation

$$d\tilde{\mathbf{x}}_1 = \mathbf{A}_1\tilde{\mathbf{x}}_1 dt + \left[\mathbf{B}_{11} + \mathbf{B}_{12}\left(t, \tilde{\mathbf{x}}_1\right)\right]\mathbf{W}d\boldsymbol{\xi}, \tag{20}$$

where $\boldsymbol{\xi} = \left[\xi_1 \; \ldots, \; \xi_{3(N+1)}\right]^T \in \mathbb{R}^{3(N+1)}$ is a Wiener process and $\mathbf{W} = \mathbf{diag}\left(w_1, \ldots, w_{3(N+1)}\right) \succ \mathbf{0}$ is a positive definite diagonal matrix that accounts for the intensity of the sensor noise. Given the structures of $\mathbf{B}_{11}$ and $\mathbf{B}_{12}$, it is clear that it is possible to rewrite (20) as the linear stochastic differential equation

$$d\tilde{\mathbf{x}}_1 = \mathbf{A}_1\tilde{\mathbf{x}}_1 dt + \sum_{i=1}^{3(N+1)} \left[\boldsymbol{b}_{1i} + \boldsymbol{\mathcal{B}}_{1i}\tilde{\mathbf{x}}_1\right] w_i d\xi_i, \tag{21}$$

where $\mathbf{b}_{1i} \in \mathbb{R}^{3(N+1)}$ are, under Assumption 2, norm bounded and $\boldsymbol{\mathcal{B}}_{1i} \in \mathbb{R}^{3(N+1)\times 3(N+1)}$ are constant matrices.

The following theorem establishes interesting results on the mean and second moment of $\tilde{\mathbf{x}}_1$.

Theorem 4 *Consider the linear SDE* (21) *under the conditions of Theorem 1. Then, there exists $W_{max} > 0$ such that, for all $\mathbf{W} \prec W_{max}\mathbf{I}$, the mean of $\tilde{\mathbf{x}}_1$ converges globally exponentially fast to zero and the second moment of $\tilde{\mathbf{x}}_1$ is bounded.*

Proof The evolution of the mean and second moment of $\tilde{\mathbf{x}}_1$ are well known since the evolution of $\tilde{\mathbf{x}}_1$ is described by a vector-valued linear SDE, see Appendix A for the general form. Let $\mathbf{m}_1 := \mathrm{E}\left\{\tilde{\mathbf{x}}_1\right\}$ and $\mathbf{P}_1 := \mathrm{E}\left\{\tilde{\mathbf{x}}_1\tilde{\mathbf{x}}_1^{\mathrm{T}}\right\}$ denote the mean and the second moment of $\tilde{\mathbf{x}}_1$, respectively. Then, the mean satisfies

$$\dot{\mathbf{m}}_1 = \mathbf{A}_1\mathbf{m}_1, \tag{22}$$

which corresponds to the same dynamics of (3). Hence, the results of Theorem 1 apply and as such it follows that the origin is a globally exponentially equilibrium point of (22), which establishes the first result. The evolution of the second moment of $\tilde{\mathbf{x}}_1$ is given by (30), with $\mathbf{m} = \mathbf{m}_1$, $\mathbf{P} = \mathbf{P}_1$, $\boldsymbol{\mathcal{A}} = \mathbf{A}_1$, $\boldsymbol{a} = \mathbf{0}$, $\boldsymbol{b}_i = w_i\boldsymbol{b}_{1i}$, and $\boldsymbol{\mathcal{B}}_i = w_i\boldsymbol{\mathcal{B}}_{1i}$, $i = 1, \ldots, 3(N+1)$. Consider, for a moment, the matrix differential equation

$$\dot{\mathbf{P}}_1 = \mathbf{A}_1\mathbf{P}_1 + \mathbf{P}_1\mathbf{A}_1^T. \tag{23}$$

Its solution is given by

$$\mathbf{P}_1(t) = \boldsymbol{\phi}\left(t, t_0\right)\mathbf{P}_1\left(t_0\right)\boldsymbol{\phi}^T\left(t, t_0\right), \tag{24}$$

where $\boldsymbol{\phi}\left(t, t_0\right)$ corresponds to the transition matrix associated with the system matrix $\mathbf{A}_1$ from t_0 to t. In the conditions of Theorem 1, there exist positive constants $c_1 > 0$ and $d_1 > 0$ such that

$$\left\|\boldsymbol{\phi}\left(t, t_0\right)\right\| \leq c_1 e^{-d_1(t-t_0)}, \; t \geq t_0,$$

see [20, Theorem 6.7]. Hence, using some norm inequalities in (24) allows to show that

$$\left\|\mathbf{P}_1\left(t\right)\right\| \leq c_1^2 \left\|\mathbf{P}_1\left(t_0\right)\right\| e^{-2d_1(t-t_0)},$$

which means that the origin of (23) is a globally exponentially stable equilibrium point. Now, let $\mathbf{p}_1 := \mathbf{vec}\left(\mathbf{P}_1\right)$ be a column representation of $\mathbf{P}_1$. Then, $\mathbf{p}_1$ also converges globally exponentially fast to zero. Since the dynamics of $\mathbf{p}_1$ are linear, it follows from [20, Theorem 7.8] that there exists a positive definite matrix $\boldsymbol{\mathcal{Q}}_1 \succ \mathbf{0}$ and constants $a_{11} > 0$, $a_{12} > 0$, and $a_{13} > 0$ such that

$$a_{11}\mathbf{I} \preceq \boldsymbol{\mathcal{Q}}_1 \preceq a_{12}\mathbf{I} \tag{25}$$

and

$$\boldsymbol{\mathcal{A}}_1^T\boldsymbol{\mathcal{Q}}_1 + \boldsymbol{\mathcal{Q}}_1\boldsymbol{\mathcal{A}}_1 + \dot{\boldsymbol{\mathcal{Q}}}_1 \preceq -a_{13}\mathbf{I}, \tag{26}$$

where $\boldsymbol{\mathcal{A}}_1$ is implicitly defined by

$$\boldsymbol{\mathcal{A}}_1\mathbf{p}_1 = \mathbf{vec}\left(\mathbf{A}_1\mathbf{P}_1 + \mathbf{P}_1\mathbf{A}_1^{\mathrm{T}}\right).$$

Now, recall the evolution of $\mathbf{P}_1$ described by (30) with the appropriate substitutions, as previously described, and consider the Lyapunov-like function

$$\mathcal{V}_1\left(t, \mathbf{p}_1\right) := \mathbf{p}_1^T\boldsymbol{\mathcal{Q}}_1\mathbf{p}_1.$$

Define

$$\mathbf{u}_1 := \mathbf{vec}\left(\sum_{\mathrm{i}=1}^{3(\mathrm{N}+1)} \mathrm{w}_{\mathrm{i}}^2\left[\boldsymbol{\mathcal{B}}_{1\mathrm{i}}\mathbf{m}_1\boldsymbol{b}_{1\mathrm{i}}^{\mathrm{T}} + \boldsymbol{b}_{1\mathrm{i}}\mathbf{m}_1^{\mathrm{T}}\boldsymbol{\mathcal{B}}_{1\mathrm{i}}^{\mathrm{T}} + \boldsymbol{b}_{1\mathrm{i}}\boldsymbol{b}_{1\mathrm{i}}^{\mathrm{T}}\right]\right),$$

which allows to write

$$\dot{\mathbf{p}}_1 = \boldsymbol{\mathcal{A}}_1\mathbf{p}_1 + \mathbf{vec}\left(\sum_{\mathrm{i}=1}^{3(\mathrm{N}+1)} \mathrm{w}_{\mathrm{i}}^2\boldsymbol{\mathcal{B}}_{1\mathrm{i}}\mathbf{P}_1\boldsymbol{\mathcal{B}}_{1\mathrm{i}}^{\mathrm{T}}\right) + \mathbf{u}_1.$$

Then, the time derivative of $\mathcal{V}_1\left(\mathbf{p}_1\right)$ is given by

$$\dot{\mathcal{V}}_1\left(t, \mathbf{p}_1\right) = \mathbf{p}_1^T\left[\boldsymbol{\mathcal{A}}_1^T\boldsymbol{\mathcal{Q}}_1 + \boldsymbol{\mathcal{Q}}_1\boldsymbol{\mathcal{A}}_1 + \dot{\boldsymbol{\mathcal{Q}}}_1\right]\mathbf{p}_1$$
$$+2\mathbf{p}_1^T\boldsymbol{\mathcal{Q}}_1\mathbf{vec}\left(\textstyle\sum_{\mathrm{i}=1}^{3(\mathrm{N}+1)} \mathrm{w}_{\mathrm{i}}^2\boldsymbol{\mathcal{B}}_{1\mathrm{i}}\mathbf{P}_1\boldsymbol{\mathcal{B}}_{1\mathrm{i}}^{\mathrm{T}}\right) + 2\mathbf{p}_1^{\mathrm{T}}\boldsymbol{\mathcal{Q}}_1\mathbf{u}_1.$$

Using (26) and the fact that $\boldsymbol{\mathcal{B}}_{1i}$ is constant allows to bound $\dot{\mathcal{V}}_1\left(t, \mathbf{p}_1\right)$ by

$$\dot{\mathcal{V}}_1\left(t, \mathbf{p}_1\right) \leq -a_{13}\left\|\mathbf{p}_1\right\|^2 + \sigma_1\overline{w}^2\left\|\mathbf{p}_1\right\|^2 + 2a_{12}\left\|\mathbf{p}_1\right\|\left\|\mathbf{u}_1\right\|$$

for some $\sigma_1 > 0$, where $\overline{w} := \max\left(w_1, \ldots, w_{3(N+1)}\right)$. Choose W_{max} such that $a_{13} - \sigma_1 W_{max}^2 \geq \lambda_3 > 0$. Then, for all $\mathbf{W} \prec W_{max}\mathbf{I}$, it follows that

$$\dot{\mathcal{V}}_1\left(t, \mathbf{p}_1\right) \leq -\lambda_3\left\|\mathbf{p}_1\right\|^2 + 2a_{12}\left\|\mathbf{p}_1\right\|\left\|\mathbf{u}_1\right\|.$$

Using (25) and setting $0 < \theta < 1$ gives

$$\dot{\mathcal{V}}_1\left(t, \mathbf{p}_1\right) \leq -\lambda_3\left(1 - \theta\right)\left\|\mathbf{p}_1\right\|^2$$
$$-\lambda_3\theta\left\|\mathbf{p}_1\right\|\left(\left\|\mathbf{p}_1\right\| - \frac{2a_{12}}{\lambda_3\theta}\left\|\mathbf{u}_1\right\|\right)$$

which means that $\dot{\mathcal{V}}_1(t,\mathbf{p}_1) \leq -\lambda_3(1-\theta)\|\mathbf{p}_1\|^2$ for all $\|\mathbf{p}_1\| \geq \frac{2a_{12}}{\lambda_3\theta}\|\mathbf{u}_1\|$. Hence, invoking [15, Theorem 5.2] one concludes that the dynamics of $\mathbf{p}_1$ are globally input-to-state stable with respect to $\|\mathbf{u}_1\|$. Now, notice that $\mathbf{u}_1$ corresponds to the sum of terms that converge globally exponentially fast to zero and a term that is norm-bounded for all time. Hence, $\|\mathbf{u}_1\|$ is bounded, which together with the input-to-state stability result, allows to conclude that $\mathbf{p}_1$ is norm-bounded, thus concluding the proof.

Remark 4 Loosely speaking, this result establishes that the estimates provided by the rate gyro bias estimator are asymptotically unbiased and their covariance is bounded. It is interesting to point out that this happens in spite of the presence of both multiplicative and additive noise in the rate gyro bias estimator dynamics. This is due to the fact that, in spite of the intrinsic nonlinear nature of the original estimation problem, the estimator that is here considered has a linear structure, which results in a particularly well-behaved linear stochastic differential equation.

The analysis for the second observer is more evolved since the dynamics (6) depend on the rate gyro bias estimates and as such the rate gyro bias estimation error appears in (8). In all truth, the rate gyro bias estimation error is a stochastic process that depends on the measurements noise, encoded by the Wiener process $\boldsymbol{\xi}$. Unfortunately, no closed-form solution exists. Therefore, in order to keep the analysis tractable but still meaningful, $\tilde{\mathbf{b}}_{\boldsymbol{\omega}}$ is here assumed to be equal to the sum of some mean value $\mathbf{m}_{\mathbf{b}} \in \mathbb{R}^3$ and a Wiener process. Thus, considering a stochastic setting, it is possible to see from (8) that, in differential form, the error dynamics of the second observer can be written as

$$\begin{aligned} d\tilde{\mathbf{x}}_2 = & \left[\mathbf{A}_2 - \mathbf{S}_3(\mathbf{m}_{\mathbf{b}})\right]\tilde{\mathbf{x}}_2 dt + \mathbf{S}_3(\mathbf{m}_{\mathbf{b}})\mathbf{x}_2 dt \\ & + \mathbf{S}_3(\mathbf{W}_{\mathbf{b}} d\boldsymbol{\xi}_{\mathbf{b}})\mathbf{x}_2 - \mathbf{S}_3(\mathbf{W}_{\mathbf{b}} d\boldsymbol{\xi}_{\mathbf{b}})\tilde{\mathbf{x}}_2 \\ & + \mathbf{S}_3(\mathbf{W}_{\boldsymbol{\omega}} d\boldsymbol{\xi}_{\boldsymbol{\omega}})\mathbf{x}_2 - \mathbf{S}_3(\mathbf{W}_{\boldsymbol{\omega}} d\boldsymbol{\xi}_{\boldsymbol{\omega}})\tilde{\mathbf{x}}_2 \\ & - \mathbf{C}_2^T \mathbf{Q}^{-1}\mathbf{W}_{\mathbf{v}} d\boldsymbol{\xi}_{\mathbf{v}}, \end{aligned} \tag{27}$$

where $\mathbf{diag}(\mathbf{W}_{\mathbf{v}}, \mathbf{W}_{\boldsymbol{\omega}}) = \mathbf{W}$ as previously defined, $\mathbf{W}_{\mathbf{b}}$ is a positive definite diagonal matrix that accounts for the covariance of the rate gyro bias estimation error, and

$$\boldsymbol{\xi}_a = \begin{bmatrix} \boldsymbol{\xi}_{\mathbf{v}} \\ \boldsymbol{\xi}_{\boldsymbol{\omega}} \\ \boldsymbol{\xi}_{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \xi_{a,1} \\ \vdots \\ \xi_{a,3(N+2)} \end{bmatrix}$$

is a Wiener process. Notice that Theorem 4 characterizes both $\mathbf{m}_{\mathbf{b}}$ and $\mathbf{W}_{\mathbf{b}}$. Hence, the trade-off that is here considered in order to keep the analysis tractable is to disregard the correlation between $\boldsymbol{\xi}_{\mathbf{b}}$ and the actual processes that account for the measurement noise, $\boldsymbol{\xi}_{\mathbf{v}}$ and $\boldsymbol{\xi}_{\boldsymbol{\omega}}$.

Before proceeding, it is convenient to rewrite (27) as

$$\begin{aligned} d\tilde{\mathbf{x}}_2 = & \left[\mathbf{A}_2 - \mathbf{S}_3(\mathbf{m}_{\mathbf{b}})\right]\tilde{\mathbf{x}}_2 dt + \mathbf{S}_3(\mathbf{m}_{\mathbf{b}})\mathbf{x}_2 dt \\ & + \sum_{i=1}^{3(N+2)} \left[\boldsymbol{b}_{2i} + \boldsymbol{\mathcal{B}}_{2i}\tilde{\mathbf{x}}_2\right] w_{a,i} d\xi_{a,i}, \end{aligned} \tag{28}$$

where $w_{a,i}$, $i = 1, \ldots, 3(N+2)$, correspond to the diagonal elements of $\mathbf{diag}(\mathbf{W}_{\mathbf{v}}, \mathbf{W}_{\boldsymbol{\omega}}, \mathbf{W}_{\mathbf{b}})$, and $\boldsymbol{b}_{2i}$ and $\boldsymbol{\mathcal{B}}_{2i}$, $i = 1, \ldots, 3(N+2)$, are implicitly defined by

$$\begin{aligned} & \textstyle\sum_{i=1}^{3(N+2)} \boldsymbol{b}_{2i} w_{a,i} d\xi_{a,i} = -\mathbf{C}_2^T \mathbf{Q}^{-1}\mathbf{W}_{\mathbf{v}} d\boldsymbol{\xi}_{\mathbf{v}} \\ & \quad + \mathbf{S}_3(\mathbf{W}_{\mathbf{b}} d\boldsymbol{\xi}_{\mathbf{b}})\mathbf{x}_2 + \mathbf{S}_3(\mathbf{W}_{\boldsymbol{\omega}} d\boldsymbol{\xi}_{\boldsymbol{\omega}})\mathbf{x}_2, \end{aligned}$$

and

$$\sum_{i=1}^{3(N+2)} \boldsymbol{\mathcal{B}}_{2i}\tilde{\mathbf{x}}_2 w_{a,i} d\xi_{a,i} = -\mathbf{S}_3(\mathbf{W}_{\mathbf{b}} d\boldsymbol{\xi}_{\mathbf{b}})\tilde{\mathbf{x}}_2 - \mathbf{S}_3(\mathbf{W}_{\boldsymbol{\omega}} d\boldsymbol{\xi}_{\boldsymbol{\omega}})\tilde{\mathbf{x}}_2.$$

The following theorem establishes interesting results on the mean and second moment of $\tilde{\mathbf{x}}_2$.

Theorem 5 *Consider the linear SDE* (28) *and suppose that* $\mathbf{Q}$ *is a positive definite matrix. Assume also that the mean of* $\tilde{\mathbf{x}}_1$ *converges globally exponentially fast to zero and the second moment of* $\tilde{\mathbf{x}}_1$ *is bounded. Furthermore, suppose that Assumption 3 holds. Then, there exists* $0 < w_{a,i} < \overline{w}$, $i = 1, \ldots, 3(N+2)$, *such that the mean of* $\tilde{\mathbf{x}}_2$ *converges globally exponentially fast to zero and the second moment of* $\tilde{\mathbf{x}}_2$ *is bounded.*

Proof The evolution of $\tilde{\mathbf{x}}_2$ is described by a vector-valued linear SDE. Hence, similarly to Theorem 4, let $\mathbf{m}_2 := \mathrm{E}\{\tilde{\mathbf{x}}_2\}$ and $\mathbf{P}_2 := \mathrm{E}\{\tilde{\mathbf{x}}_2\tilde{\mathbf{x}}_2^{\mathrm{T}}\}$ denote the mean and the second moment of $\tilde{\mathbf{x}}_2$, respectively. Then, the mean satisfies

$$\dot{\mathbf{m}}_2 = \left[\mathbf{A}_2 - \mathbf{S}_3(\mathbf{m}_{\mathbf{b}})\right]\mathbf{m}_2 + \mathbf{S}_3(\mathbf{m}_{\mathbf{b}})\mathbf{m}_2.$$

Since in the conditions of the theorem the mean $\mathbf{m}_{\mathbf{b}}$ converges globally exponentially fast to zero, it can be shown that $\mathbf{m}_2$ converges globally exponentially to zero using similar arguments to [3, Theorem 3]. The evolution of the second moment of $\tilde{\mathbf{x}}_2$ is given by (30), with $\mathbf{m} = \mathbf{m}_2$, $\mathbf{P} = \mathbf{P}_2$, $\boldsymbol{\mathcal{A}} = \mathbf{A}_2 - \mathbf{S}_3(\mathbf{m}_{\mathbf{b}})$, $\boldsymbol{a} = \mathbf{S}_3(\mathbf{m}_{\mathbf{b}})\mathbf{x}_2$, $\boldsymbol{b}_i = w_{a,i}\boldsymbol{b}_{2i}$, and $\boldsymbol{\mathcal{B}}_i = w_{a,i}\boldsymbol{\mathcal{B}}_{2i}$, $i = 1, \ldots, 3(N+2)$. The remaining analysis follows similar arguments to that of Theorem 4, in order to show that

$$\dot{\mathbf{p}}_2 = \boldsymbol{\mathcal{A}}_2\mathbf{p}_2 + \mathbf{u}_2 + \mathbf{vec}\left(\sum_{\mathrm{i=1}}^{\mathrm{3(n+2)}} \mathrm{w}_{\mathrm{a,i}}^2 \boldsymbol{\mathcal{B}}_{\mathrm{2i}}\mathbf{P}_2\boldsymbol{\mathcal{B}}_{\mathrm{2i}}^{\mathrm{T}}\right)$$

is input-to-state stable with respect to $\mathbf{u}_2$, where $\mathbf{p}_2 := \mathbf{vec}(\mathbf{P}_2)$,

$$\mathbf{u}_2 := \mathbf{vec}\left(\sum_{i=1}^{3(N+2)} w_{a,i}^2 \left[\boldsymbol{\mathcal{B}}_{2i}\mathbf{m}_2\boldsymbol{b}_{2i}^T + \boldsymbol{b}_{2i}\mathbf{m}_2^T\boldsymbol{\mathcal{B}}_{2i}^T + \boldsymbol{b}_{2i}\boldsymbol{b}_{2i}^T\right]\right) + \mathbf{m}_2\left[\mathbf{S}_3(\mathbf{m_b})\mathbf{x}_2\right]^T + \left[\mathbf{S}_3(\mathbf{m_b})\mathbf{x}_2\right]\mathbf{m}_2^T,$$

and $\boldsymbol{\mathcal{A}}_2$ is implicitly defined by

$$\boldsymbol{\mathcal{A}}_2 = \mathbf{vec}\left(\left[\mathbf{A}_2 - \mathbf{S}_3(\mathbf{m_b})\right]\mathbf{P}_2 + \mathbf{P}_2\left[\mathbf{A}_2 - \mathbf{S}_3(\mathbf{m_b})\right]^{\mathrm{T}}\right).$$

The only difference concerns the proof of global exponential stability of

$$\dot{\mathbf{P}}_2 = \left[\mathbf{A}_2 - \mathbf{S}_3(\mathbf{m_b})\right]\mathbf{P}_2 + \mathbf{P}_2\left[\mathbf{A}_2 - \mathbf{S}_3(\mathbf{m_b})\right]^T,$$

which nevertheless follows using similar arguments but also considering [16, Example 9.6]. With globally input-to-state stability established, the second result follows from the fact that $\mathbf{u}_2$ corresponds to the sum of a norm-bounded term plus exponentially decaying terms.

Remark 5 Loosely speaking, this result establishes that the attitude estimates provided by the cascade observer, under appropriate conditions, are asymptotically unbiased and, moreover, the error covariance is bounded. Thus, robustness of the solution to stochastic noise modeled by a Wiener process in all sensor measurements is established.

5 Estimates on $SO(3)$

In spite of the fact that the attitude estimates provided by the cascade observer (5)-(6) converge, in the absence of noise, to elements of $SO(3)$, and remain close to $SO(3)$ in the presence of realistic noise, it has been pointed out that in some cases one may be interested in an estimate on $SO(3)$. A possible solution was proposed in [3, Section 3.4.2]. Next, an adaptation of [2, Theorem 7] that also considers rate gyro bias, in the absence of measurement noise, is proposed.

Theorem 6 *Consider the estimate $\hat{\mathbf{R}}$ obtained from the cascade attitude observer (5)-(6) in the conditions of Theorem 1, considering null measurement noise. Further suppose that Assumption 3 holds and $\mathbf{Q}$ is a positive definite matrix. Fix $0 < \epsilon < 1$. Finally, consider that the initial estimate satisfies $\hat{\mathbf{R}}(t_0) \in SO(3)$ and define a new attitude estimate $\hat{\mathbf{R}}_f$ of the rotation matrix $\mathbf{R}$ as follows:*

- *if $\left\|\hat{\mathbf{R}}^T\hat{\mathbf{R}} - \mathbf{I}\right\| \leq \epsilon$, then the estimate $\hat{\mathbf{R}}_f$ is obtained by projecting $\hat{\mathbf{R}}$ on $SO(3)$, i.e.,*

$$\hat{\mathbf{R}}_f = \arg\min_{\mathbf{X}\in SO(3)} \left\|\mathbf{X} - \hat{\mathbf{R}}\right\|, \left\|\hat{\mathbf{R}}^T\hat{\mathbf{R}} - \mathbf{I}\right\| \leq \epsilon;$$

- *if $\left\|\hat{\mathbf{R}}^T\hat{\mathbf{R}} - \mathbf{I}\right\| > \epsilon$, propagate the estimate in open-loop, as given by*

$$\dot{\hat{\mathbf{R}}}_f = \hat{\mathbf{R}}_f\mathbf{S}\left(\boldsymbol{\omega}_m - \hat{\mathbf{b}}_{\boldsymbol{\omega}}\right).$$

Then,

1. *the projection of $\hat{\mathbf{R}}$ on $SO(3)$, for $\left\|\hat{\mathbf{R}}^T\hat{\mathbf{R}} - \mathbf{I}\right\| \leq \epsilon$, is unique;*
2. *$\hat{\mathbf{R}}_f \in SO(3)$;*
3. *there exists t_s such that $\left\|\hat{\mathbf{R}}^T(t)\hat{\mathbf{R}}(t) - \mathbf{I}\right\| \leq \epsilon$ for all $t \geq t_s$ and therefore $\hat{\mathbf{R}}_f$ corresponds to the projection on $SO(3)$ of $\hat{\mathbf{R}}$ for all $t \geq t_s$; and*
4. *the error $\tilde{\mathbf{R}}_f := \left\|\mathbf{R} - \hat{\mathbf{R}}_f\right\|$ is bounded and*

$$\lim_{t\to\infty} \left\|\tilde{\mathbf{R}}_f\right\| = 0.$$

Moreover, the convergence is exponentially fast.

Proof To show the first part of the theorem, suppose that $\left\|\hat{\mathbf{R}}^T\hat{\mathbf{R}} - \mathbf{I}\right\| \leq \epsilon$ and consider the singular value decomposition $\hat{\mathbf{R}} = \mathbf{U}\boldsymbol{\Xi}\mathbf{V}$, where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices and $\boldsymbol{\Xi}$ is a diagonal matrix whose diagonal elements are non-negative real numbers. Then, it follows that

$$\begin{aligned}\left\|\hat{\mathbf{R}}^T\hat{\mathbf{R}} - \mathbf{I}\right\| \leq \epsilon &\Leftrightarrow \left\|(\mathbf{U}\boldsymbol{\Xi}\mathbf{V})^T(\mathbf{U}\boldsymbol{\Xi}\mathbf{V}) - \mathbf{I}\right\| \leq \epsilon \\ &\Leftrightarrow \left\|\mathbf{V}^T\boldsymbol{\Xi}\mathbf{U}^T\mathbf{U}\boldsymbol{\Xi}\mathbf{V} - \mathbf{I}\right\| \leq \epsilon \\ &\Leftrightarrow \left\|\mathbf{V}^T\boldsymbol{\Xi}\boldsymbol{\Xi}\mathbf{V} - \mathbf{V}^T\mathbf{V}\right\| \leq \epsilon \\ &\Leftrightarrow \left\|\boldsymbol{\Xi}^2 - \mathbf{I}\right\| \leq \epsilon. \end{aligned} \tag{29}$$

Now, let $\sigma_i > 0$, $i = 1, 2, 3$, denote the diagonal elements of $\boldsymbol{\Xi}$. With $0 < \epsilon < 1$, it follows from (29) that

$$0 < \sqrt{1-\epsilon} \leq \sigma_i \leq \sqrt{1+\epsilon}, \; i = 1, 2, 3.$$

Hence, if $\left\|\hat{\mathbf{R}}^T\hat{\mathbf{R}} - \mathbf{I}\right\| \leq \epsilon$, then $\hat{\mathbf{R}}$ is nonsingular, which allows to conclude that its projection on $SO(3)$ is uniquely defined, see [5]. The second part of the theorem follows by construction, since the initial estimate belongs to $SO(3)$ and both the projection and the open-loop propagation yield estimates that belong to $SO(3)$. In the conditions of the theorem, in particular assuming a positive definite matrix $\mathbf{Q}$ and in the conditions of Theorem 1, it follows that the origin of the observer error dynamics is a globally exponentially stable equilibrium point, see [3, Theorem 3]. With that in mind, the proof of the last two parts of the theorem follow similar steps to [2, Theorem 7] and hence is omitted.

In practice, there is measurement noise. It is nevertheless clear that, after some time and for realistic noise, the estimate on $SO(3)$, $\hat{\mathbf{R}}_f$, will always correspond to the projection on $SO(3)$ of $\hat{\mathbf{R}}$, thus retaining the good properties in terms of robustness to noise. This is so due to the fact that, in practice, the sensor noise is bounded and sufficiently small to obtain good estimates of the attitude. To make this more clear, let $\tilde{\mathbf{R}} := \mathbf{R} - \hat{\mathbf{R}}$. This is directly related to the estimation error of the observer, since $\tilde{\mathbf{x}}_2$ is a column representation of $\tilde{\mathbf{R}}$. It is possible to show, by direct computation, that

$$\left\|\hat{\mathbf{R}}^T(t)\hat{\mathbf{R}}(t) - \mathbf{I}\right\| = \left\|\tilde{\mathbf{R}}^T\tilde{\mathbf{R}} - \mathbf{R}^T\tilde{\mathbf{R}} - \tilde{\mathbf{R}}^T\mathbf{R}\right\|.$$

Since $\|\mathbf{R}\| = 1$, it is clear that a bound for $\left\|\tilde{\mathbf{R}}\right\|$ results in a bound for $\left\|\hat{\mathbf{R}}^T(t)\hat{\mathbf{R}}(t) - \mathbf{I}\right\|$. The result then follows directly from Theorems 2, 3, and 6. Thus, for sufficiently small bounded noise, it is possible to choose $0 < \epsilon < 1$ such that $\left\|\hat{\mathbf{R}}^T(t)\hat{\mathbf{R}}(t) - \mathbf{I}\right\| \leq \epsilon$ for all $t \geq t_s$ and therefore $\hat{\mathbf{R}}_f$ corresponds to the projection on $SO(3)$ of $\hat{\mathbf{R}}$ for all $t \geq t_s$.

It is also important to remark that the projection on $SO(3)$ is computationally inexpensive and a numerically robust algorithm. Finally, given the covariance of $\hat{\mathbf{R}}$, the covariance of the projection on $SO(3)$ can also be characterized, see e.g. [18].

6 Numerical results

In Section 3 the robustness of the overall attitude observer (1)-(2) to bounded noise was assessed and the results can summarized as follows: the estimation error of the rate gyro bias observer (1) is bounded for all initial conditions provided that the measurement noise is sufficiently small, with an upper bound given by (13). Moreover, the smaller the measurement noise is, the smaller the estimation error is. For the attitude observer (2), the region of convergence is global. Moreover, the smaller the measurement noise and the bias estimation errors are, the smaller the attitude estimation error is. These are the results that are of greater interest since, in practice, the measurement noise is always bounded.

Nevertheless, it would be interesting to capture exactly the region of convergence, and a bound for the noise is given by (13), although this can be a conservative result. Unfortunately, the analytical computation of this bound proves to be difficult, if not impossible, since the closed-form eigenvalues of a positive definite solution $\mathbf{Q}_1$ that satisfies (9)-(10) is not available. Nevertheless, one can resort to numerical studies to evaluate whether or not this region of convergence is sufficiently large, in practice. In this section, Monte Carlo simulations are presented to evaluate the region of convergence of the rate gyro bias observer (1). The results for (2) are not shown since, as detailed before, the region of robust convergence of this observer is global.

The simulation setup replicates the one described in [3]. The observer parameters are $\alpha_1 = \frac{9.8}{0.008}10^{-3}$, $\alpha_2 = \frac{0.5}{0.0015}10^{-3}$, and $\beta_1 = \beta_2 = 10^{-3}$. These parameters were adjusted empirically to obtain good estimation results with the sensor suite described in [3]. Here, a total of 400 runs are carried out, 20 runs for 20 different noise levels, selected as follows. The noise of each sensor is assumed to follow a zero-mean uniform distribution, such that the noise is bounded. The bound on the noise of each sensor increases linearly along these 20 levels, with the minimum and maximum values described in Table 1. The initial estimate for the observer, in each run, follows a Gaussian distribution centered at the true value and with standard deviation of $2\ \mathrm{m/s^2}$ for the acceleration of gravity, 0.1 gauss for the magnetic field, and $0.5\ ^\circ/\mathrm{s}$ for the rate gyro bias. The distributions of initial states that are considered are centered at the true value because the goal here is to evaluate the robustness to sensor noise and the observer was shown to be robust for all initial conditions.

Table 1 Noise levels

Sensor	Noise level 1	Noise level 20
Rate gyros	$0.1\ ^\circ/\mathrm{s}$	$10\ ^\circ/\mathrm{s}$
Accelerometers	$0.098\ \mathrm{m/s^2}$	$1.96\ \mathrm{m/s^2}$
Magnetometers	5×10^{-3} gauss	0.1 gauss

For each simulation, the highest absolute value of the estimation error of each quantity was computed in steady-state, which was assumed only after 240 s, to eliminate possible slower convergence situations, even though in practice steady-state is usually achieved around 120 s. Then, for each noise level, the maximum steady-state error among all 20 runs was computed. This is depicted in Figs. 1, 2, and 3. From these figures it is clearly possible to identify the input-to-state stability result derived in Theorem 2. Indeed, the higher the noise level is, the higher the estimation error is, and the change is roughly linear in this case, as expected. Moreover, notice in Table 1 that the noise reaches extremely high values, which are not seen in practice. Even very cheap inertial measurement units (IMUs) have much better specifications. Hence, while the region of convergence r_n, given by (13), is not explicitly computed, it is possible to see that the observer is robust, in practice, to extremely high and very unlikely noise levels.

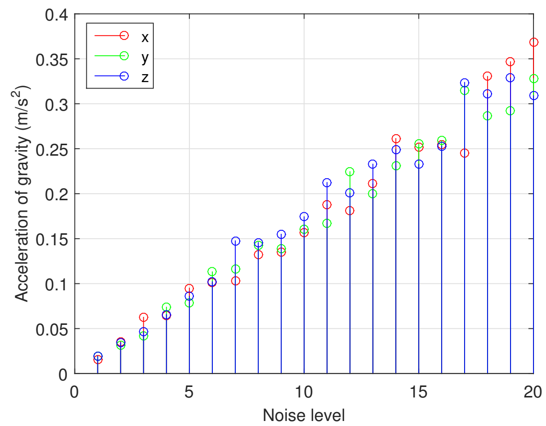

Fig. 1 Maximum error, in steady-state, of the acceleration of gravity estimate

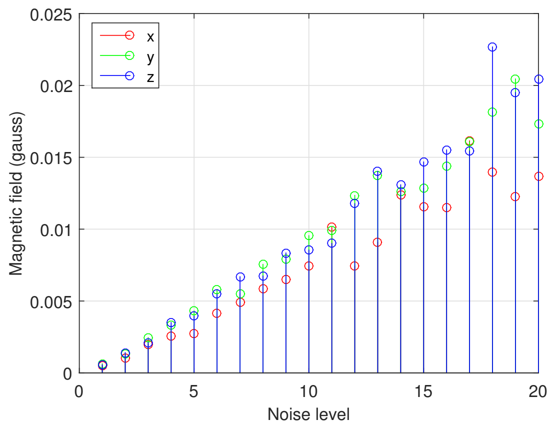

Fig. 2 Maximum error, in steady-state, of the magnetic field estimate

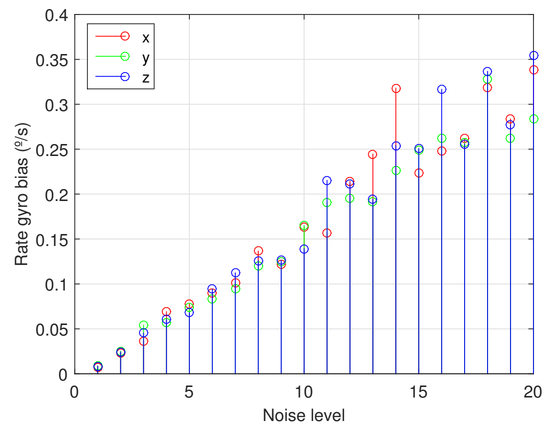

Fig. 3 Maximum error, in steady-state, of the rate gyro bias estimate

7 Conclusions

The robustness to measurement noise in all sensors of a cascade attitude observer with topological relaxations was addressed in this paper, considering both bounded noise and stochastic noise. The results can be summarized as follows: for all initial conditions, as long as the measurement noise (covariance) is within certain bounds for the bounded-noise setting (for the Wiener-based setting), the estimation error remains bounded in the bounded-noise setting (the mean error converges to zero and the error covariance remains bounded in the Wiener-based setting). The results are coherent with what could be expected: looking at the error dynamics (7)-(8), it is possible to see that if no limitations were imposed on the noise, one could always choose particular signals such that the system would become unstable. However, for reasonable values of the measurement noise (small enough), the system is robust to noise. An additional construct was provided that yields estimates directly on $SO(3)$. To conclude, it was shown with Monte Carlo simulations that, for practical parameters of the attitude observer, robustness is exhibited for very high noise levels, not even typical of the cheapest IMUs.

Compliance with Ethical Standards

Funding: This work was supported by the Fundação para a Ciência e a Tecnologia (FCT) through ISR under FCT [UID/EEA/50009/2013] and through the FCT project DECENTER [LISBOA-01-0145-FEDER-029605], funded by the Lisboa 2020 and PIDDAC programs.

Conflict of interest: The authors declare that they have no further conflict of interest concerning the publication of this manuscript.

References

1. Barrau, A., Bonnabel, S.: Intrinsic filtering on Lie groups with applications to attitude estimation. IEEE Transactions on Automatic Control **60**(2), 436–449 (2015)
2. Batista, P., Silvestre, C., Oliveira, P.: A GES Attitude Observer with Single Vector Observations. Automatica **48**(2), 388–395 (2012)
3. Batista, P., Silvestre, C., Oliveira, P.: Globally Exponentially Stable Cascade Observers for Attitude Estimation. Control Engineering Practice **20**(2), 148–155 (2012)
4. Batista, P., Silvestre, C., Oliveira, P.: Sensor-based Globally Asymptotically Stable Filters for Attitude Estimation: Analysis, Design, and Performance Evaluation. IEEE Transactions on Automatic Control **57**(8), 2095–2100 (2012)
5. Belta, C., Kumar, V.: An SVD-Based Projection Method for Interpolation on SE(3). IEEE Transactions on Robotics and Automation **18**(3), 334–345 (2002)

6. Berkane, S., Abdessameud, A., Tayebi, A.: Global Hybrid Attitude Estimation on the Special Orthogonal Group SO(3). In: Proceedings of the 2016 American Control Conference, pp. 113–118. Boston, USA (2016)
7. Berkane, S., Tayebi, A.: On Deterministic Attitude Observers on the Special Orthogonal Group SO(3). In: Proceedings of the 55th IEEE Conference on Decision and Control, pp. 1165–1170. Las Vegas, USA (2016)
8. Crassidis, J., Markley, F., Cheng, Y.: Survey of Nonlinear Attitude Estimation Methods. Journal of Guidance, Control and Dynamics **30**(1), 12–28 (2007)
9. Geering, H., Dondi, G., Herzog, F., Keel, S.: Stochastic Systems. Measurement and Control Laboratory, Swiss Federal Institute of Technology (ETH) (2011)
10. Grip, H., Fossen, T., Johansen, T., Saberi, A.: Attitude Estimation Using Biased Gyro and Vector Measurements with Time-Varying Reference Vectors. IEEE Transactions on Automatic Control **57**(5), 1332–1338 (2012)
11. Grip, H., Fossen, T., Johansen, T., Saberi, A.: Globally Exponentially Stable Attitude and Gyro Bias Estimation with Application to GNSS/INS Integration. Automatica **51**, 158–166 (2015)
12. Hua, M.D.: Attitude estimation for accelerated vehicles using GPS/INS measurements. Control Engineering Practice **18**(7), 723–732 (2010)
13. Izadi, M., Sanyal, A.: Rigid body attitude estimation based on the Lagrange-d'Alembert principle. Automatica **50**(10), 2570–2577 (2014)
14. Izadi, M., Viswanathan, S., Sanyal, A., Silvestre, C., Oliveira, P.: The Variational Attitude Estimator in the Presence of Bias in Angular Velocity Measurements. In: Proceedings of the 2016 American Control Conference, pp. 4065–4070. Boston, USA (2016)
15. Khalil, H.: Nonlinear Systems, 2nd edn. Prentice-Hall (1996)
16. Khalil, H.: Nonlinear Systems, 3rd edn. Prentice Hall (2001)
17. Mahony, R., Hamel, T., Pflimlin, J.M.: Nonlinear Complementary Filters on the Special Orthogonal Group. IEEE Transactions on Automatic Control **53**(5), 1203–1218 (2008)
18. Markley, F., Mortari, D.: How to Estimate Attitude from Vector Observations. Proceedings of the AAS/AIAA Astrodynamics Specialist Conference **103**(3), 1979–1996 (1999)
19. Namvar, M., Safaei, F.: Adaptive Compensation of Gyro Bias in Rigid-Body Attitude Estimation Using a Single Vector Measurement. IEEE Transactions on Automatic Control **58**(7), 1816–1822 (2013)
20. Rugh, W.: Linear system theory, 2nd edn. Prentice-Hall, Inc. (1995)
21. Teel, A.R., Hespanha, J.: Examples of GES Systems That can be Driven to Infinity by Arbitrarily Small Additive Decaying Exponentials. IEEE Transactions on Automatic Control **49**(8), 1407–1410 (2004)
22. Trumpf, J., Mahony, R., Hamel, T., Lageman, C.: Analysis of Non-Linear Attitude Observers for Time-Varying Reference Measurements. IEEE Transactions on Automatic Control **57**(11), 2789–2800 (2012)
23. Wu, T.H., Kaufman, E., Lee, T.: Globally Asymptotically Stable Attitude Observer on SO(3). In: Proceedings of the 54th IEEE Conference on Decision and Control, pp. 2164–2168. Osaka, Japan (2015)
24. Zamani, M., Trumpf, J., Mahony, R.: Minimum-Energy Filtering for Attitude Estimation. IEEE Transactions on Automatic Control **58**(11), 2917–2921 (2013)

A First and second moments of linear stochastic differential equations

For linear stochastic differential equations there exist ordinary differential equations for computing the first and second moments of the stochastic process, see e.g. [9, Section 3.4.3]. Consider the linear stochastic differential equation

$$d\boldsymbol{x} = \left[\boldsymbol{\mathcal{A}}\boldsymbol{x} + \boldsymbol{a}\right] dt + \sum_{i=1}^{n} \left[\boldsymbol{\mathcal{B}}_i \boldsymbol{x} + \boldsymbol{b}_i\right] d\xi_i ,$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{n\times n}$, $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{\mathcal{B}}_i \in \mathbb{R}^{n\times n}$, $\boldsymbol{b}_i \in \mathbb{R}^n$, and $\boldsymbol{\xi} = [\xi_1 \ldots \xi_n]^T \in \mathbb{R}^n$ is a n-dimensional Wiener process. Let $\mathbf{m} := \mathrm{E}\left\{\boldsymbol{x}\right\} \in \mathbb{R}^n$ and $\mathbf{P} = \mathrm{E}\left\{\boldsymbol{x}\boldsymbol{x}^T\right\} \in \mathbb{R}^{n\times n}$ denote the mean and second moment of $\boldsymbol{x}$, respectively. Then,

$$\dot{\mathbf{m}} = \boldsymbol{\mathcal{A}}\mathbf{m} + \boldsymbol{a}, \ \mathbf{m}\left(t_0\right) = \boldsymbol{x}_0 ,$$

and

$$\begin{aligned}\dot{\mathbf{P}} &= \boldsymbol{\mathcal{A}}\mathbf{P} + \mathbf{P}\boldsymbol{\mathcal{A}}^T + \mathbf{m}\boldsymbol{a}^T + \boldsymbol{a}\mathbf{m}^T \\ &\quad + \sum_{i=1}^{n} \left[\boldsymbol{\mathcal{B}}_i \mathbf{P} \boldsymbol{\mathcal{B}}_i^T + \boldsymbol{b}_i \boldsymbol{b}_i^T + \boldsymbol{\mathcal{B}}_i \mathbf{m} \boldsymbol{b}_i^T + \boldsymbol{b}_i \mathbf{m}^T \boldsymbol{\mathcal{B}}_i^T\right] , \qquad (30)\end{aligned}$$

with $\mathbf{P}\left(t_0\right) = \boldsymbol{x}_0 \boldsymbol{x}_0^T$.