International Journal of Control
Vol. 00, No. 00, April 2011, 1–12

Regular Paper

Partial Attitude and Rate Gyro Bias Estimation: Observability Analysis, Filter Design, and Performance Evaluation

Pedro Batista*, Carlos Silvestre, and Paulo Oliveira

Instituto Superior Técnico / Institute for Systems and Robotics
Av. Rovisco Pais, 1, Lisboa, Portugal

(*received March 2011*)

This paper presents the analysis, design, and performance evaluation of a novel filter for partial attitude and rate gyro bias estimation. In addition to a single body-fixed vector observation of a constant reference vector in inertial coordinates, biased angular velocity measurements are available for filtering purposes, which occurs directly in the space of the vector observations. Appropriate observability conditions are derived and a filter with globally asymptotically stable (GAS) error dynamics is proposed. Simulation and experimental results are detailed, with ground truth data for comparison purposes, that illustrate the achievable performance of the envisioned solutions.

Keywords: attitude estimation; observability analysis; filter design.

1 Introduction

Attitude estimation has been a hot topic of research in the past decades, playing a key role in the development of navigation systems for autonomous vehicles and mobile platforms. In traditional solutions there exist, at least, two vector observations, in body-fixed coordinates, of known constant reference vectors in inertial coordinates, where at least two of them are non parallel. Extended Kalman Filters (EKFs) and variants have been widely used, see Farrell (1970) and Bar-Itzhack and Oshman (1985), for instance. In spite of the good performance achieved by EKF and EKF-like solutions, divergence due to the linearization of the system dynamics, see Crassidis et al. (2007), has led the scientific community to pursue different solutions, in particular nonlinear observers such as those presented in Metni et al. (2006), Sanyal et al. (2008), Vasconcelos et al. (2010), Tayebi et al. (2007), Rehbinder and Ghosh (2003), Batista et al. (2009), Mahony et al. (2008), Thienel and Sanner (2003) and references therein. The reader is referred to Crassidis et al. (2007) for a survey on this topic.

It is well known that, when there is a single vector observation, in body-fixed coordinates, of a single constant reference vector, in inertial coordinates, it is impossible to recover the whole attitude of the platform. However, it is still possible to partially recover some of the attitude information and, in addition, estimate the rate gyro bias. This last point is particularly important as, without bias compensation, the error of open-loop integration of inertial sensors diverges, see Savage (1998*a*) and Savage (1998*b*). By partially estimating the attitude and fully estimating the rate gyro bias, it is possible to aid inertial navigation systems, reducing the divergence of attitude estimation, in open-loop, to the effect of the integration of sensor-noise, but eliminating the overwhelming negative effect of rate gyro bias. In addition, for some applications, the entire

*Corresponding author. Email: pbatista@isr.ist.utl.pt

ISSN: 0020-7179 print/ISSN 1366-5820 online

DOI: 10.1080/0020717YYxxxxxxxx
http://www.informaworld.com

attitude is not required. As an example, for stabilization of quadrotors, the roll and pitch angles are quite important, but the yaw angle, with respect to an inertial frame, is not always required, particularly in vision-based approaches.

This paper presents the analysis, design, and performance evaluation of a novel filter for partial attitude and rate gyro bias estimation. In addition to a single body-fixed vector observation of a constant reference vector in inertial coordinates, biased angular velocity measurements are available for filtering purposes, which occurs directly in the space of the vector observations. An exact transformation of the system dynamics is at the core of the proposed design that allows to regard the system as linear time-varying for observer design purposes, even though it still is, in fact, nonlinear. Powerful results for linear systems are employed and appropriate observability conditions are derived, ending with the design of a Kalman filter with globally asymptotically stable (GAS) error dynamics. Preliminary work by the authors can be found in Batista et al. (2009), where the theoretical framework was first presented. The present paper elaborates in greater detail on the observability of the system and presents defining simulation and experimental results to evaluate the performance of the proposed solution.

The paper is organized as follows. Section 2 introduces the problem at hand, while the observability analysis is detailed in Section 3. The filter design is briefly addressed in Section 4 and simulation results are presented in Section 5. Experimental results are discussed in Section 6 and, finally, Section 7 presents the main conclusions and results of the paper.

1.1 *Notation*

Throughout the paper the symbol $\mathbf{0}$ denotes a matrix (or vector) of zeros and $\mathbf{I}$ an identity matrix, both of appropriate dimensions. A block diagonal matrix is represented as $\operatorname{diag}(\mathbf{A}_1, \ldots, \mathbf{A}_\mathrm{n})$. For $\mathbf{x} \in \mathbb{R}^3$ and $\mathbf{y} \in \mathbb{R}^3$, $\mathbf{x}\times\mathbf{y}$ represents the cross product. Finally, the Dirac delta function is denoted by $\delta(t)$.

2 Problem statement

Let $\{I\}$ denote a local inertial frame, $\{B\}$ the body-fixed frame, and $\mathbf{R}(t) \in SO(3)$ the rotation matrix from $\{B\}$ to $\{I\}$. The attitude kinematics are given by

$$\dot{\mathbf{R}}(t) = \mathbf{R}(t)\mathbf{S}\left[\boldsymbol{\omega}(t)\right],$$

where $\boldsymbol{\omega}(t) \in \mathbb{R}^3$ is the angular velocity of $\{B\}$, expressed in $\{B\}$, and $\mathbf{S}\left(\boldsymbol{x}\right)$ is the skew-symmetric matrix such that $\mathbf{S}\left(\boldsymbol{x}\right)\boldsymbol{y} = \boldsymbol{x} \times \boldsymbol{y}$. It is assumed that the angular velocity is a bounded continuous differentiable signal, with bounded derivative.

Suppose that measurements $\mathbf{y}(t) \in \mathbb{R}^3$ are available, in body-fixed coordinates, of known constant quantities in inertial coordinates, i.e.

$${}^{I}\mathbf{y} = \mathbf{R}(t)\mathbf{y}(t).$$

Further consider rate gyro measurements $\boldsymbol{\omega}_m(t) \in \mathbb{R}^3$ corrupted with bias $\mathbf{b}_{\boldsymbol{\omega}} \in \mathbb{R}^3$, i.e.,

$$\boldsymbol{\omega}_m(t) = \boldsymbol{\omega}(t) + \mathbf{b}_{\boldsymbol{\omega}}(t).$$

The problem considered here is that of obtaining filtered estimates of $\mathbf{y}(t)$ and estimating the rate gyro bias $\mathbf{b}_{\boldsymbol{\omega}}$. In addition, partial attitude reconstruction is carried out using ${}^{I}\mathbf{y}$ and the filtered estimates of $\mathbf{y}(t)$.

3 Observability analysis

It is well known that, with a single vector observation of a constant vector in inertial coordinates, it is impossible to recover the attitude. Indeed, one degree of freedom is unobservable, which is readily verified as $\mathbf{y}(t) = \mathbf{R}^T(t)^I\mathbf{y} = \mathbf{R}^T(t)\mathbf{R}_o^{T\,I}\mathbf{y}$ for all rotation matrices $\mathbf{R}_o$ about the inertial constant vector ${}^I\mathbf{y}$, while $\mathbf{R}_o\mathbf{R}(t)$ satisfies the corresponding differential equation

$$\frac{\mathrm{d}}{\mathrm{dt}}\left[\mathbf{R}_o\mathbf{R}(t)\right] = \left[\mathbf{R}_o\mathbf{R}(t)\right]\mathbf{S}\left[\boldsymbol{\omega}(t)\right].$$

Nevertheless, partial attitude reconstruction can still be performed and, in some cases, it is also possible to estimate the rate gyro bias. This is discussed in this section.

Considering a sensor-based approach, with the incorporation of the derivative of the vector observations in the system dynamics, gives

$$\begin{cases} \dot{\mathbf{x}}_1(t) = -\mathbf{S}\left[\boldsymbol{\omega}_m(t)\right]\mathbf{x}_1(t) + \mathbf{S}\left[\mathbf{x}_2(t)\right]\mathbf{x}_1(t) \\ \dot{\mathbf{x}}_2(t) = \mathbf{0} \\ \mathbf{y}(t) = \mathbf{x}_1(t) \end{cases}, \tag{1}$$

where $\mathbf{x}_1(t) = \mathbf{y}(t)$ and $\mathbf{x}_2(t) = \mathbf{b}_{\boldsymbol{\omega}}(t)$ are the system states and $\mathbf{y}(t)$ is the system output. Using the cross product property $\boldsymbol{x} \times \boldsymbol{y} = -\boldsymbol{y} \times \boldsymbol{x}$ and the relation $\mathbf{y}(t) = \mathbf{x}_1(t)$ allows to rewrite (1) as

$$\begin{cases} \dot{\mathbf{x}}_1(t) = -\mathbf{S}\left[\boldsymbol{\omega}_m(t)\right]\mathbf{x}_1(t) - \mathbf{S}\left[\mathbf{y}(t)\right]\mathbf{x}_2(t) \\ \dot{\mathbf{x}}_2(t) = \mathbf{0} \\ \mathbf{y}(t) = \mathbf{x}_1(t) \end{cases}$$

or, in compact form,

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\left(\mathbf{y}(t), \boldsymbol{\omega}_m(t)\right)\mathbf{x}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \end{cases}, \tag{2}$$

where $\mathbf{x}(t) = \left[\mathbf{x}_1^T(t)\ \mathbf{x}_2^T(t)\right]^T \in \mathbb{R}^6$ is the system state,

$$\mathbf{A}\left(\mathbf{y}(t), \boldsymbol{\omega}_m(t)\right) = \begin{bmatrix} -\mathbf{S}\left[\boldsymbol{\omega}_m(t)\right] & -\mathbf{S}\left[\mathbf{y}(t)\right] \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{6\times 6},$$

and $\mathbf{C} = \begin{bmatrix}\mathbf{I}\ \mathbf{0}\end{bmatrix} \in \mathbb{R}^{3\times 6}$. For simplicity of notation, the dependence of $\mathbf{A}\left(\mathbf{y}(t), \boldsymbol{\omega}_m(t)\right)$ on $\mathbf{y}(t)$ and $\boldsymbol{\omega}_m(t)$ is omitted in the sequel and the system matrix is simply denoted by $\mathbf{A}(t)$.

Although (2) is a nonlinear system, it can be regarded as linear time-varying because the output $\mathbf{y}(t)$ is available for observer design purposes. Before presenting the main results of this section, the following lemma is introduced.

Lemma 3.1: *Consider the nonlinear system*

$$\begin{cases} \dot{\boldsymbol{x}}(t) = \boldsymbol{\mathcal{A}}\left(t, \boldsymbol{u}(t), \boldsymbol{y}(t)\right)\boldsymbol{x}(t) + \boldsymbol{\mathcal{B}}(t)\boldsymbol{u}(t) \\ \boldsymbol{y}(t) = \boldsymbol{\mathcal{C}}(t)\boldsymbol{x}(t) \end{cases}. \tag{3}$$

If the observability Gramian $\boldsymbol{\mathcal{W}}\left(t_0, t_f\right)$ associated with the pair $\left(\boldsymbol{\mathcal{A}}\left(t, \boldsymbol{u}(t), \boldsymbol{y}(t)\right), \boldsymbol{\mathcal{C}}(t)\right)$ on $\mathcal{I} := [t_0, t_f]$ is invertible then the nonlinear system (3) is observable in the sense that, given the system input $\{\boldsymbol{u}(t), t \in \mathcal{I}\}$ and the system output $\{\boldsymbol{y}(t), t \in \mathcal{I}\}$, the initial condition $\boldsymbol{x}\left(t_0\right)$ is uniquely defined.

Proof Given the system input $\{\boldsymbol{u}(t), t \in \mathcal{I}\}$ and the system output $\{\boldsymbol{y}(t), t \in \mathcal{I}\}$, it is possible to compute the transition matrix associated with the system matrix $\boldsymbol{\mathcal{A}}(t, \boldsymbol{u}(t), \boldsymbol{y}(t))$

$$\begin{aligned}\boldsymbol{\phi}(t, t_0) = \mathbf{I} + &\int_{t_0}^{t} \boldsymbol{\mathcal{A}}(\sigma_1, \boldsymbol{u}(\sigma_1), \boldsymbol{y}(\sigma_1))\, d\sigma_1 \\ &+ \int_{t_0}^{t} \boldsymbol{\mathcal{A}}(\sigma_1, \boldsymbol{u}(\sigma_1), \boldsymbol{y}(\sigma_1)) \int_{t_0}^{\sigma_1} \boldsymbol{\mathcal{A}}(\sigma_2, \boldsymbol{u}(\sigma_2), \boldsymbol{y}(\sigma_2))\, d\sigma_2 d\sigma_1 + \ldots\end{aligned}$$

on $\mathcal{I}$, which clearly satisfies $\boldsymbol{\phi}(t_0, t_0) = \mathbf{I}$ and

$$\frac{\partial \boldsymbol{\phi}(t, t_0)}{\partial t} = \boldsymbol{\mathcal{A}}(t, \boldsymbol{u}(t), \boldsymbol{y}(t))\, \boldsymbol{\phi}(t, t_0).$$

Therefore, it is also possible to compute the observability Gramian

$$\boldsymbol{\mathcal{W}}(t_0, t_f) = \int_{t_0}^{t_f} \boldsymbol{\phi}^T(t, t_0) \boldsymbol{\mathcal{C}}^T(t) \boldsymbol{\mathcal{C}}(t) \boldsymbol{\phi}(t, t_0)\, dt.$$

Now, notice that it is possible to write the evolution of the state, given the system input and output (which allow to compute the transition matrix), as

$$\boldsymbol{x}(t) = \boldsymbol{\phi}(t, t_0)\, \boldsymbol{x}_0 + \int_{t_0}^{t} \boldsymbol{\phi}(t, \tau)\, \boldsymbol{\mathcal{B}}(\tau)\, \boldsymbol{u}(\tau)\, d\tau, \tag{4}$$

where $\boldsymbol{x}_0$ is the initial condition. This is easily verified as with $t = t_0$ in (4) gives $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ and taking the time derivative of (4) gives $\dot{\boldsymbol{x}}(t) = \boldsymbol{\mathcal{A}}(t, \boldsymbol{u}(t), \boldsymbol{y}(t))\, \boldsymbol{x}(t) + \boldsymbol{\mathcal{B}}(t) \boldsymbol{u}(t)$. The remainder of the proof follows as in classic theory. The output of the system can be written, from (4), as

$$\boldsymbol{y}(t) = \boldsymbol{\mathcal{C}}(t) \boldsymbol{\phi}(t, t_0)\, \boldsymbol{x}_0 + \boldsymbol{\mathcal{C}}(t) \int_{t_0}^{t} \boldsymbol{\phi}(t, \tau)\, \boldsymbol{\mathcal{B}}(\tau)\, \boldsymbol{u}(\tau)\, d\tau. \tag{5}$$

Multiplying (5) on both sides by $\boldsymbol{\phi}^T(t, t_0) \boldsymbol{\mathcal{C}}^T(t)$ and integrating on $\mathcal{I}$ yields

$$\boldsymbol{\mathcal{W}}(t_0, t_f)\, \boldsymbol{x}_0 = \int_{t_0}^{t_f} \boldsymbol{\phi}^T(t, t_0) \boldsymbol{\mathcal{C}}^T(t) \boldsymbol{y}(t) dt - \int_{t_0}^{t_f} \boldsymbol{\phi}^T(t, t_0) \boldsymbol{\mathcal{C}}^T(t) \boldsymbol{\mathcal{C}}(t) \int_{t_0}^{t} \boldsymbol{\phi}(t, \tau)\, \boldsymbol{\mathcal{B}}(\tau)\, \boldsymbol{u}(\tau)\, d\tau dt. \tag{6}$$

All quantities in (6) but the initial condition are known given the system input and output and therefore it corresponds to a linear algebraic equation on $\boldsymbol{x}_0$. If the observability Gramian $\boldsymbol{\mathcal{W}}(t_0, t_f)$ is invertible, then $\boldsymbol{x}_0$ is uniquely defined, which concludes the proof. □

Remark 1 : Notice that, even though the evolution of the state given in (4) seems to resemble the response of a linear system, the response of the system does not correspond to the superposition of the free response (due to the initial conditions) and the forced response (due to system input). This is so because the transition matrix in (4) depends explicitly on the system input. For observability purposes this is not a problem because both the input and output are available.

Remark 2 : Notice that Lemma 3.1 presents only a sufficient condition for the observability of the nonlinear system (3). For linear systems the present condition is also necessary. However, the fact that the underlying system is inherently nonlinear precludes that conclusion in this case.

The following result [Proposition 4.2, Batista et al. (2011)] is useful in the sequel.

Proposition 3.2: *Let* $\mathbf{f}(t) : [t_0, t_f] \subset \mathbb{R} \rightarrow \mathbb{R}^n$ *be a continuous and i-times continuously differentiable function on* $\mathcal{I} := [t_0, t_f]$, $T := t_f - t_0 > 0$, *and such that* $\mathbf{f}(t_0) = \dot{\mathbf{f}}(t_0) = \ldots =$

$\mathbf{f}^{(i-1)}(t_0) = \mathbf{0}$. *Further assume that* $\left\|\mathbf{f}^{(i+1)}(t)\right\| \leq C$ *for all* $t \in \mathcal{I}$. *If there exist* $\alpha > 0$ *and* $t_1 \in \mathcal{I}$ *such that* $\left\|\mathbf{f}^{(i)}(t_1)\right\| \geq \alpha$, *then there exist* $0 < \delta \leq T$ *and* $\beta > 0$ *such that* $\|\mathbf{f}(t_0 + \delta)\| \geq \beta$.

The following theorem addresses the observability of the nonlinear system (2).

Theorem 3.3 : *The nonlinear system* (2) *is observable on* $\mathcal{I} := [t_0, t_f]$, *in the sense that given* $\{\boldsymbol{\omega}_m(t), t \in \mathcal{I}\}$ *and* $\{\mathbf{y}(t), t \in \mathcal{I}\}$ *the initial condition* $\mathbf{x}(t_0)$ *is uniquely defined, if and only if for every unit vector* $\mathbf{d} \in \mathbb{R}^3$, *it is possible to choose* $t^* \in \mathcal{I}$ *such that* $\|\mathbf{y}(t^*) \times \mathbf{d}\| > 0$.

Proof The proof of sufficiency follows resorting to Lemma 3.1. The transition matrix associated with $\mathbf{A}(t)$ is given by

$$\boldsymbol{\phi}(t, t_0) = \begin{bmatrix} \mathbf{R}_m^T(t) & -\mathbf{R}_m^T(t) \int_{t_0}^t \mathbf{R}_m(\sigma)\mathbf{S}[\mathbf{y}(\sigma)]\, d\sigma \\ \mathbf{0} & \mathbf{I} \end{bmatrix},$$

where $\mathbf{R}_m(t) \in SO(3)$ is such that $\dot{\mathbf{R}}_m(t) = \mathbf{R}_m(t)\mathbf{S}[\boldsymbol{\omega}_m(t)]$, with $\mathbf{R}_m(t_0) = \mathbf{I}$. This can be easily verified as $\boldsymbol{\phi}(t_0, t_0) = \mathbf{I}$ and

$$\frac{\partial \boldsymbol{\phi}(t, t_0)}{\partial t} = \mathbf{A}(t)\boldsymbol{\phi}(t, t_0).$$

Let $\mathbf{d} = \left[\mathbf{d}_1^T \mathbf{d}_2^T\right]^T \in \mathbb{R}^6$ be a unit vector, with $\mathbf{d}_1$, $\mathbf{d}_2 \in \mathbb{R}^3$. Then, it follows that

$$\mathbf{d}^T \boldsymbol{\mathcal{W}}(t_0, t_f)\, \mathbf{d} = \int_{t_0}^{t_f} \|\mathbf{f}(\tau, t_0)\|^2\, d\tau,$$

where $\boldsymbol{\mathcal{W}}(t_0, t_f)$ denotes the observability Gramian associated with the pair $(\mathbf{A}(t), \mathbf{C})$ on $[t_0, t_f]$ and

$$\mathbf{f}(\tau, t_0) = \mathbf{d}_1 - \int_{t_0}^{\tau} \mathbf{R}_m(\sigma)\mathbf{S}[\mathbf{y}(\sigma)]\, \mathbf{d}_2 d\sigma \in \mathbb{R}^3.$$

The first derivative of $\mathbf{f}(\tau, t_0)$ with respect to τ is given by

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{f}(\tau, t_0) = \mathbf{R}_m(\tau)\mathbf{S}[\mathbf{y}(\tau)]\, \mathbf{d}_2$$

and it is straightforward to show that the second derivative with respect to τ is norm-bounded under the assumption of bounded angular velocities $\boldsymbol{\omega}(t)$. Now, notice that if $\mathbf{d}_1 \neq \mathbf{0}$ then $\|\mathbf{f}(t_0, t_0)\| = \|\mathbf{d}_1\| = \alpha_1 > 0$. On the other hand, if $\mathbf{d}_1 = \mathbf{0}$, it must be $\|\mathbf{d}_2\| = 1$, $\mathbf{f}(t_0, t_0) = \mathbf{0}$, and $\left\|\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{f}(\tau, t_0)\right\| = \|\mathbf{y}(\tau) \times \mathbf{d}_2\|$. Now, under the hypothesis of the theorem, for all $\|\mathbf{d}_2\| = 1$ it is possible to choose $t^* \in \mathcal{I}$ such that $\left\| \frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{f}(\tau, t_0)\Big|_{\tau = t^*}\right\| \geq \alpha_2 > 0$. But that means, using Proposition 3.2, that if $\|\mathbf{d}_2\| = 1$, there exist $t_2 \in \mathcal{I}$ such that $\|\mathbf{f}(t_2, t_0)\| = \alpha_3 > 0$. Thus, so far it is possible to conclude that, for all $\|\mathbf{d}\| = 1$, if is possible to choose $t_3 \in \mathcal{I}$ such that $\|\mathbf{f}(t_3, t_0)\|$. This, in turn, allows to conclude, using Proposition 3.2 again, that $\mathbf{d}^T \boldsymbol{\mathcal{W}}(t_0, t_f)\, \mathbf{d} > 0$ for all $\|\mathbf{d}\| = 1$, which means that the observability Gramian $\boldsymbol{\mathcal{W}}(t_0, t_f)$ is positive definite. As such, it follows, from Lemma 3, that under the hypothesis of the theorem the nonlinear system (2) is observable on $\mathcal{I}$.

The proof of necessity follows by simply showing that, if the hypothesis of the theorem is not verified, then there exist at least two initial conditions that explain the system output. Suppose that there exists a unit vector $\mathbf{d} \in \mathbb{R}^3$ such that $\mathbf{y}(t) \times \mathbf{d} = \mathbf{0}$ for all $t \in \mathcal{I}$ or, equivalently, $\mathbf{y}(t) = \mathbf{y}_0$ is a constant vector. Let $\mathbf{x}_1(t_0) = \mathbf{y}_0$ and $\mathbf{x}_2(t_0) = \mathbf{x}_{20}$ be the initial state of the nonlinear system (2). As the output is constant, it follows that

$$-\mathbf{S}\left[\boldsymbol{\omega}_m(t)\right]\mathbf{y}(t)-\mathbf{S}\left[\mathbf{y}(t)\right]\mathbf{x}_{20}=\mathbf{0}. \tag{7}$$

Now, notice that, considering different initial conditions

$$\begin{cases}\mathbf{x}_1\left(t_0\right)=\mathbf{y}_0\\ \mathbf{x}_2\left(t_0\right)=\mathbf{x}_{20}+\gamma\mathbf{y}_0\end{cases}, \tag{8}$$

where $\gamma\in\mathbb{R}\backslash\{0\}$, the same output is verified for all $t\in\mathcal{I}$, i.e., $\mathbf{y}(t)=\mathbf{y}_0$ is the output of the nonlinear system (2) with initial conditions (8). This can be readily verified: using (7) and $\mathbf{y}(t)\times\mathbf{y}_0=\mathbf{y}_0\times\mathbf{y}_0=\mathbf{0}$ gives

$$\dot{\mathbf{x}}_1(t)=-\mathbf{S}\left[\boldsymbol{\omega}_m(t)\right]\mathbf{y}(t)-\mathbf{S}\left[\mathbf{y}(t)\right]\mathbf{x}_{20}-\gamma\mathbf{S}\left[\mathbf{y}(t)\right]\mathbf{y}_0=\mathbf{0}$$

for all $t\in\mathcal{I}$, which means that $\mathbf{x}_1(t)=\mathbf{x}_1\left(t_0\right)=\mathbf{y}_0$ and as such $\mathbf{y}(t)=\mathbf{x}_1(t)=\mathbf{y}_0$. But this concludes the proof as it was shown that, if the hypothesis of the theorem does not hold, then the initial state is not uniquely defined. □

In general, stronger forms of observability are required in order to design observers (or filters) with globally asymptotically stable error dynamics. The following theorem provides a sufficient condition on uniform complete observability of the pair $(\mathbf{A}(t),\mathbf{C})$, a result that will be useful for the design of observers (filters) with globally asymptotically stable error dynamics.

Theorem 3.4: *The pair $(\mathbf{A}(t),\mathbf{C})$ is uniformly completely observable if and only if there exist scalars $\alpha>0$ and $\delta>0$ such that, for all $t\geq t_0$ and all unit vectors $\mathbf{d}\in\mathbb{R}^3$, it is possible to choose $t^*\in[t,t+\delta]$ such that $\left\|\mathbf{y}\left(t^*\right)\times\mathbf{d}\right\|\geq\alpha$.*

Proof The pair $(\mathbf{A}(t),\mathbf{C})$ is uniformly completely observable if and only if there exist scalars $\alpha_1>0$, $\alpha_2>0$, and $\delta^*>0$ such that

$$\alpha_1<\mathbf{d}^T\boldsymbol{\mathcal{W}}\left(t,t+\delta^*\right)\mathbf{d}<\alpha_2 \tag{9}$$

for all $t\geq t_0$ and all unit vectors $\mathbf{d}\in\mathbb{R}^6$, where $\boldsymbol{\mathcal{W}}\left(t,t+\delta^*\right)$ denotes the observability Gramian associated with the pair $(\mathbf{A}(t),\mathbf{C})$ on $[t,t+\delta^*]$. As all entries of the system matrices $\mathbf{A}(t)$ and $\mathbf{C}$ are continuous and bounded, it follows that, given $\delta^*>0$, it is always possible to choose $\alpha_2>0$ such that the right side of (9) is verified for all $t\geq t_0$ and all unit vectors $\mathbf{d}\in\mathbb{R}^6$. Therefore, in the remainder of the proof, only the left side of (9) is considered.

In order to prove sufficiency, suppose that there exist scalars $\alpha>0$ and $\delta>0$ such that, for all $t\geq t_0$ and all unit vectors $\mathbf{d}_2^{'}\in\mathbb{R}^3$, it is possible to choose $t^*\in[t,t+\delta]$ such that $\left\|\mathbf{y}\left(t^*\right)\times\mathbf{d}_2^{'}\right\|\geq\alpha$. Let $\delta^*=\delta$. As in Theorem 3.3, it is possible to write

$$\mathbf{d}^T\boldsymbol{\mathcal{W}}\left(t,t+\delta\right)\mathbf{d}=\int_t^{t+\delta}\left\|\mathbf{f}\left(\tau,t\right)\right\|^2d\tau,$$

where $\mathbf{d}=\left[\mathbf{d}_1^T\ \mathbf{d}_2^T\right]^T$, $\mathbf{d}_1\in\mathbb{R}^3$, $\mathbf{d}_2\in\mathbb{R}^3$, and

$$\mathbf{f}\left(\tau,t\right)=\mathbf{d}_1-\int_t^{\tau}\mathbf{R}_m(\sigma)\mathbf{S}\left[\mathbf{y}(\sigma)\right]\mathbf{d}_2d\sigma\in\mathbb{R}^3.$$

Suppose that $\mathbf{d}_1\neq\mathbf{0}$. Then, it is clear that there exists $c_1>0$ such that $\left\|\mathbf{f}\left(t,t\right)\right\|=\left\|\mathbf{d}_1\right\|=c_1$

for all $t \geq t_0$. On the other hand, if $\mathbf{d}_1 = \mathbf{0}$, then it must be $\|\mathbf{d}_2\| = 1$, $\mathbf{f}(t,t) = \mathbf{0}$ and

$$\left\| \frac{\mathrm{d}}{\mathrm{d}\tau} \mathbf{f}\left(\tau, t\right) \right\| = \left\| \mathbf{y}(\tau) \times \mathbf{d}_2 \right\|.$$

But under the hypothesis of the theorem, it is possible to choose $t^* \in [t, t+\delta]$ such that

$$\left\| \left. \frac{\mathrm{d}}{\mathrm{d}\tau} \mathbf{f}\left(\tau, t\right) \right|_{\tau = t^*} \right\| \geq \alpha$$

for all $t \geq t_0$ and $\|\mathbf{d}_2\| = 1$. Hence, resorting to Proposition 3.2, it is possible to conclude that there exists a scalar $c_2 > 0$ such that, for all $t \geq t_0$ and $\|\mathbf{d}_2\| = 1$, it is possible to choose $t' \in [t, t+\delta]$ such that $\left\|\mathbf{f}\left(t', t\right)\right\| \geq c_2$. Thus, it has been established that there exists a scalar $c > 0$ such that, for all $t \geq t_0$ and $\|\mathbf{d}\| = 1$, it is possible to choose $t_i \in [t, t+\delta]$ such that $\|\mathbf{f}\left(t_i, t\right)\| \geq c$. But this, in turn, allows to conclude, using Proposition 3.2 again, that there exists a scalar $\alpha_1 > 0$ such that, for all $t \geq t_0$ and $\|\mathbf{d}\| = 1$, $\mathbf{d}^T \boldsymbol{\mathcal{W}}\left(t, t+\delta\right) \mathbf{d} > \alpha_1$. This concludes the first part of the proof.

In order to establish necessity, it will be shown that if the hypothesis does not hold, then the pair $(\mathbf{A}(t), \mathbf{C})$ is not uniformly completely observable. Suppose then that, for all scalars $\alpha_1' > 0$ and $\delta > 0$, there exists $t^* \geq t_0$ and a unit vector $\mathbf{d}_2 \in \mathbb{R}^3$ such that, for all $t \in [t^*, t^*+\delta]$, it is true that $\|\mathbf{y}(t) \times \mathbf{d}_2\| < \alpha_1'$. Let $\mathbf{d} = \left[\mathbf{0}\, \mathbf{d}_2^T\right]^T \in \mathbb{R}^6$. Then, it is straightforward to conclude that

$$\begin{aligned}
\mathbf{d}^T \boldsymbol{\mathcal{W}}\left(t^*, t^*+\delta\right) \mathbf{d} &= \textstyle\int_{t_*}^{t^*+\delta} \left\|\mathbf{f}\left(\tau, t^*\right)\right\|^2 d\tau = \int_{t_*}^{t^*+\delta} \left\| \int_{t^*}^{\tau} \mathbf{R}_m(\sigma) \mathbf{S}\left[\mathbf{y}(\sigma)\right] \mathbf{d}_2 d\sigma \right\|^2 d\tau \\
&\leq \textstyle\int_{t_*}^{t^*+\delta} \int_{t^*}^{\tau} \left\|\mathbf{R}_m(\sigma) \mathbf{S}\left[\mathbf{y}(\sigma)\right] \mathbf{d}_2\right\|^2 d\sigma d\tau \leq \int_{t_*}^{t^*+\delta} \int_{t^*}^{\tau} \left\|\mathbf{y}(\sigma) \times \mathbf{d}_2\right\|^2 d\sigma d\tau \\
&\leq \textstyle\left\|\mathbf{y}\left[\xi\left(t^*\right)\right] \times \mathbf{d}_2\right\|^2 \int_{t_*}^{t^*+\delta} \int_{t^*}^{\tau} 1 d\sigma d\tau \leq \left\|\mathbf{y}\left[\xi\left(t^*\right)\right] \times \mathbf{d}_2\right\|^2 \frac{\delta^2}{2},
\end{aligned}$$

where $\xi\left(t^*\right) \in \left]t^*, t^*+\delta\right[$. As $\|\mathbf{y}(t) \times \mathbf{d}_2\| < \alpha_1'$ for all $t \in [t^*, t^*+\delta]$, it follows that $\mathbf{d}^T \boldsymbol{\mathcal{W}}\left(t^*, t^*+\delta\right) \mathbf{d} < \alpha_1$, with $\alpha_1 := \alpha_1' \delta^2/2$, which means that the pair $(\mathbf{A}(t), \mathbf{C})$ is not uniformly completely observable, therefore concluding the proof. □

Remark 3 : Theorem 3.3 has a clear simple physical interpretation. In simple words, the nonlinear system (2) is observable on $[t_0, t_f]$ if and only if the vector observation, in body-fixed coordinates, is not a constant vector on that time interval. Likewise, Theorem 3.4 corresponds to the same condition but imposing also a persistency of excitation requirement, meaning that the pair $(\mathbf{A}(t), \mathbf{C})$ is uniformly completely observable if and only if there exists a certain time interval length, δ, such that there is a minimum change of the vector observation on all time intervals $[t, t+\delta]$, $t \geq t_0$.

4 Filter design

This section presents an estimation solution for the nonlinear system (2). As (2) can be regarded as linear time-varying, even though it is a nonlinear system, it is possible to implement the well known Kalman filter to estimate the system state. This is so because all quantities are available to compute $\mathbf{A}(t)$, while $\mathbf{C}$ is a constant matrix. Its stability is readily characterized from Theorem 3.4, see Jazwinski (1970).

Considering additive system disturbances and sensor noise, the system dynamics are given by

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{w}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{n}(t) \end{cases}, \tag{10}$$

where $\mathbf{w}(t) \in \mathbb{R}^6$ and $\mathbf{n}(t) \in \mathbb{R}^3$ are uncorrelated zero-mean white Gaussian processes, with $\mathrm{E}\left[\mathbf{w}(\mathrm{t})\,\mathbf{w}^{\mathrm{T}}(\mathrm{t}-\tau)\right] = \boldsymbol{\Xi}\delta(\tau)$, and $\mathrm{E}\left[\mathbf{n}(\mathrm{t})\,\mathbf{n}^{\mathrm{T}}(\mathrm{t}-\tau)\right] = \boldsymbol{\Theta}\delta(\tau)$, $\boldsymbol{\Xi} \succ \mathbf{0}$, $\boldsymbol{\Theta} \succ \mathbf{0}$. Notice that, for filter design purposes, both $\mathbf{w}(t)$ and $\mathbf{n}(t)$ could have been modeled as the outputs of stable linear time invariant filters, which could be easily employed to model, e.g., colored noise. In this paper, and for the sake of clarity of presentation, the simplest white Gaussian noise version is presented.

The Kalman filter equations for the system dynamics (10) are standard, see Kalman and Bucy (1961), Gelb (1974). The state estimate evolves according to

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{A}(t)\hat{\mathbf{x}}(t) + \mathbf{K}(t)\left[\mathbf{y}(t) - \mathbf{C}\hat{\mathbf{x}}(t)\right]$$

while the covariance matrix propagation is described by

$$\dot{\mathbf{P}}(t) = \mathbf{A}(t)\,\mathbf{P}(t) + \mathbf{P}(t)\,\mathbf{A}(t) + \boldsymbol{\Xi} - \mathbf{P}(t)\,\mathbf{C}^T\boldsymbol{\Theta}^{-1}\mathbf{C}\mathbf{P}(t)\,.$$

The Kalman gain matrix is given by $\mathbf{K}(t) = \mathbf{P}(t)\,\mathbf{C}^T\boldsymbol{\Theta}^{-1}$.

The filter error dynamics are globally asymptotically stable under the conditions of Theorem 3.4 as the pair $(\mathbf{A}(t), \mathbf{C})$ is uniformly completely observable, see e.g. Jazwinski (1970), Kalman and Bucy (1961), and Anderson (1971).

Remark 4 : Notice that it is not possible to claim that the Kalman filter is the optimal solution because, in reality, there exists multiplicative noise. An alternative design to the Kalman filter is the $\mathcal{H}_\infty$ filter, considering that $\mathbf{w}(t)$ and $\mathbf{n}(t)$ are finite energy signals.

5 Simulation results

In order to evaluate the performance of the proposed solutions, simulations were first carried out considering that the vector observations are given by an accelerometer which, for low frequencies, is dominated by the acceleration of gravity, as discussed in Mahony et al. (2008). A rate gyro unit provides angular velocity measurements corrupted by bias, which was set to $\mathbf{b}_{\boldsymbol{\omega}} = [2\ \ {-3}\ 1]^T\ \pi/180$ rad s^{-1} in the simulations. Sensor noise was added both to the accelerometer and rate gyro measurements. Additive zero-mean white Gaussian processes were considered, with standard deviation of 0.05m s^{-2} for the acceleration measurements and $0.05\pi/180$rad s^{-1} for the rate gyro measurements. With nothing else but gravity readings, it is only possible to partially recover the attitude. In particular, two Euler angles may be recovered, the roll and pitch of a roll, pitch, and yaw Euler angle representation of the attitude. These angles correspond, in a different terminology, to the bank and elevation angles of air vehicles. The evolution of the Euler angles during the simulation is depicted in Fig. 1. Notice that the trajectory described by the platform is such that uniform complete observability is attained, according to Theorem 3.4.

In the simulations, the initial estimate of the gravity was set close to the true value, as it is always possible to initialize the filter according to the first vector observation. The rate gyro bias initial estimate was set to zero. The filter matrices were chosen as $\boldsymbol{\Xi} = \mathrm{diag}\,(0.05\mathbf{I}, 0.01\mathbf{I})$ and $\boldsymbol{\Theta} = 0.05\mathbf{I}$. The evolution of the filter errors is depicted in Fig. 2, while the evolution of the Euler angles error computed from the filtered estimates is shown in Fig. 3. The evolution of the error of the Euler angles computed directly from the vector observation is also depicted for comparison purposes. The results are quite good considering the low cost characteristics of the sensor suite and the fact that a single vector observation is available to estimate the rate gyro bias. The filtering effect on the vector observations is also quite evident.

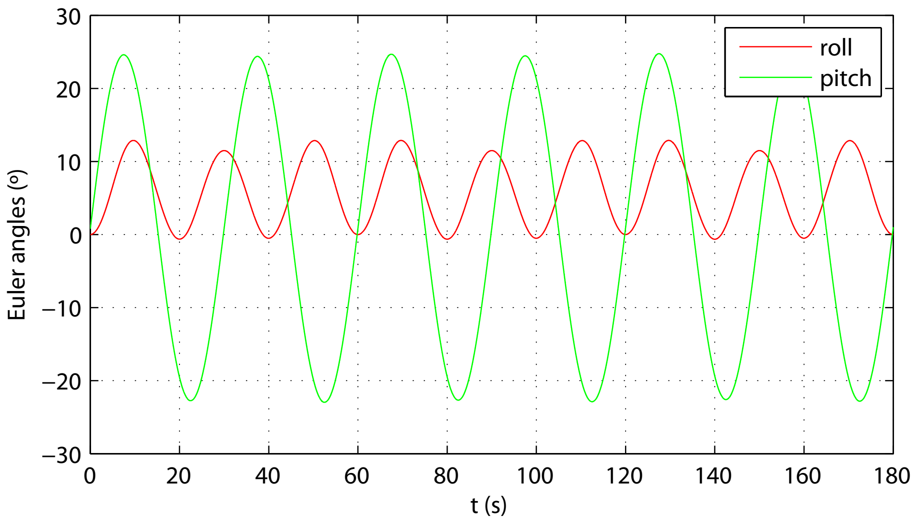

Figure 1. Evolution of the Euler angles

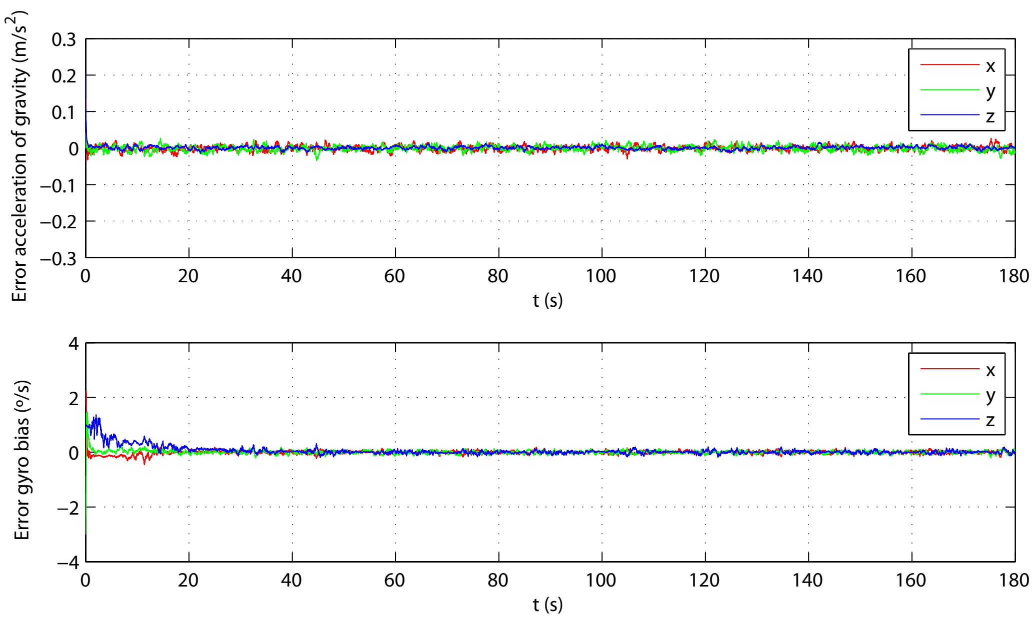

Figure 2. Evolution of the Kalman filter error

6 Experimental results

In order to evaluate the performance in real world applications, an experimental setup was developed resorting to a high precision motion rate table, Model 2103HT from Ideal Aerosmith. This table outputs, in a fixed-frequency profile, the angular position of the table with a resolution of 0.00025 °, considered as a ground truth signal for performance evaluation. The angular velocity and acceleration readings are obtained from an Inertial Measurement Unit (IMU), model NANO IMU NA02-0150F50, from MEMSENSE, which outputs data at a rate of 150 Hz. The worst case standard deviation values provided by the manufacturer of the IMU are $0.008\,\mathrm{m/s^2}$ for the accelerometers and $0.95\,°/s$ for the rate gyros. These performance specifications are typical of low cost units. Fig. 4 displays the motion rate table with the experimental setup mounted on the top.

The initial estimates of the filter were set in a similar fashion to those in the simulations: the estimate of the gravity was set close to the true value, according to the first vector observation, while the initial rate gyro bias estimate was set to zero. The system disturbances intensity matrix

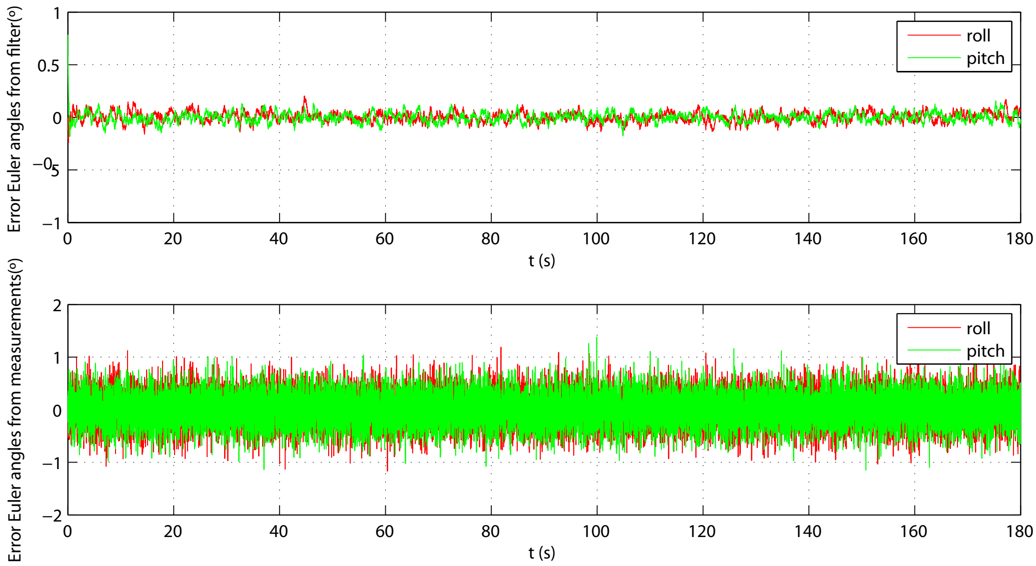

Figure 3. Evolution of the error of the Euler angles computed from the Kalman filter estimates

Figure 4. Experimental setup

was chosen as $\boldsymbol{\Xi} = 0.5\mathrm{diag}\left(0.008\mathbf{I}, 2 \times 10^{-8}\mathbf{I}\right)$, while the output noise intensity matrix was chosen as $\boldsymbol{\Theta} = 0.008\mathbf{I}$. No particular emphasis was given on the tuning process as the resulting performance with these simple parameters is very good. In practice, the spectral contents of the sensors noise may be experimentally approximated and frequency weights adjusted to improve the performance of the filter, see the examples provided in Batista et al. (2010). Moreover, correlation between the system disturbances $\mathbf{w}$ and the sensor noise $\mathbf{n}$ may also be considered.

The initial convergence of the filter error is depicted in Fig. 5, as well as the steady-state error. Notice that the convergence rate is still very fast and the filter copes well with the low specifications of the sensor suite. The initial convergence of the rate gyros biases estimates is shown in Fig. 6. Clearly, the filter copes well the slowly time-varying nature of the rate gyro biases.

7 Conclusions

This paper presented the design, analysis, and performance evaluation of a novel filter for partial attitude determination, in addition to rate gyro bias estimation, based on single vector observations. The observability of the system was thoroughly assessed and conditions with physical meaning derived. The resulting filter error dynamics are globally asymptotically stable. Simulation and experimental results were also detailed, with ground truth data for comparison

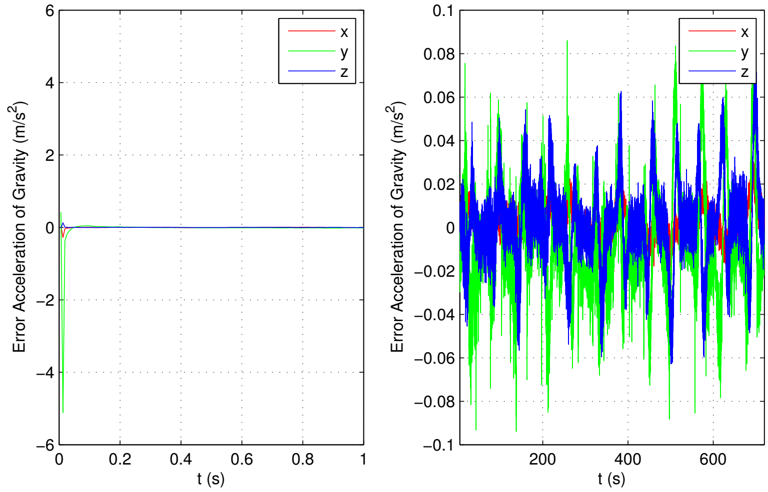

Figure 5. Evolution of the acceleration of gravity error: initial convergence and detailed evolution

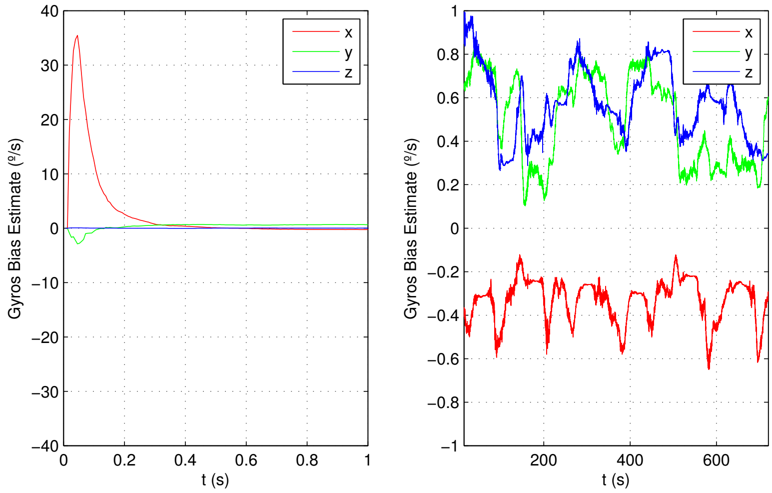

Figure 6. Evolution of the bias estimate: initial convergence and detailed evolution

purposes, that illustrate the achievable performance of the envisioned solutions.

References

Anderson, B. (1971), 'Stability properties of Kalman-Bucy filters', *Journal of Franklin Institute* **291**(2), 137–144.

Bar-Itzhack, I. and Y. Oshman (1985), 'Attitude Determination from Vector Observations: Quaternion Estimation', *IEEE Transactions on Aerospace and Electronic Systems* **321**(1), 128–136.

Batista, P., C. Silvestre and P. Oliveira (2009), Sensor-based Complementary Globally Asymptotically Stable Filters for Attitude Estimation, *in* 'Proceedings of the 48th IEEE Conference on Decision and Control', Shanghai, China, pp. 7563–7568.

Batista, P., C. Silvestre and P. Oliveira (2010), 'Optimal position and velocity navigation filters for autonomous vehicles', *Automatica* **46**(4), 767–774.

Batista, P., C. Silvestre and P. Oliveira (2011), 'On the Observability of Linear Motion Quantities in Navigation Systems', *Systems & Control Letters* **60**(2), 101–110.

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

Farrell, J. (1970), 'Attitude Determination by Kalman Filter', *Automatica* **6**(5), 419–430.

Gelb, A. (1974), *Applied Optimal Filtering*, The MIT Press.

Jazwinski, A. (1970), *Stochastic Processes and Filtering Theory*, Academic Press, Inc.

Kalman, R. and R. Bucy (1961), 'New Results in Linear Filtering and Prediction Theory', *Transactions of the ASME - Journal of Basic Engineering, Series D* **83**(3), 95–108.

Mahony, R., T. Hamel and J-M. Pflimlin (2008), 'Nonlinear Complementary Filters on the Special Orthogonal Group', *IEEE Transactions on Automatic Control* **53**(5), 1203–1218.

Metni, N., J.-M. Pflimlin, T. Hamel and P. Soueres (2006), 'Attitude and gyro bias estimation for a VTOL UAV', *Control Engineering Practice* **14**(12), 1511–1520.

Rehbinder, H. and B. Ghosh (2003), 'Pose Estimation Using Line-Based Dynamic Vision and Inertial Sensors', *IEEE Transactions on Automatic Control* **48**(2), 186–199.

Sanyal, A., T. Lee, M. Leok and N. McClamroch (2008), 'Global optimal attitude estimation using uncertainty ellipsoids', *Systems & Control Letters* **57**(3), 236–245.

Savage, P. (1998*a*), 'Strapdown Inertial Navigation Integration Algorithm Design Part 1: Attitude Algorithms', *Journal of Guidance, Control, and Dynamics* **21**(1), 19–28.

Savage, P. (1998*b*), 'Strapdown Inertial Navigation Integration Algorithm Design Part 2: Velocity and Position Algorithms', *Journal of Guidance, Control, and Dynamics* **21**(2), 208–221.

Tayebi, A., S. McGilvray, A. Roberts and M. Moallem (2007), Attitude estimation and stabilization of a rigid body using low-cost sensors, *in* 'Proceedings of the 46th IEEE Conference on Decision and Control', New Orleans, USA, pp. 6424 –6429.

Thienel, J. and R. Sanner (2003), 'A Coupled Nonlinear Spacecraft Attitude Controller and Observer With an Unknown Constant Gyro Bias and Gyro Noise', *IEEE Transactions on Automatic Control* **48**(11), 2011–2015.

Vasconcelos, J., R. Cunha, C. Silvestre and P. Oliveira (2010), 'A nonlinear position and attitude observer on $SE(3)$ using landmark measurements', *Systems & Control Letters* **59**(3-4), 155–166.