Proceedings of the 18th World Congress
The International Federation of Automatic Control
Milano (Italy) August 28 - September 2, 2011

Nonlinear attitude estimation with measurement decoupling and anti-windup gyro-bias compensation

Minh-Duc Hua [*], **Konrad Rudin** [**], **Guillaume Ducard** [*], **Tarek Hamel** [*], **Robert Mahony** [***]

[*] *I3S UNS–CNRS, Sophia Antipolis, France*
e-mails : minh.hua@polytechnique.org, ducard@i3s.unice.fr, thamel@i3s.unice.fr
[**] *Autonomous Systems Lab, ETH Zürich, Zürich, Switzerland*
e-mail : rudink@student.ethz.ch
[***] *Department of Engineering, ANU, Canberra, Australia*
e-mail: Robert.Mahony@anu.edu.au

Abstract: Any low-cost IMU is subject to biased measurement in its gyroscopes. This paper presents a nonlinear attitude observer which incorporates anti-windup gyro-bias compensation. It also provides an extension of the decoupling strategy of IMU measurement recently proposed in the literature by providing almost global stability proof along with a comprehensive study of the decoupling strategy. Simulation and experimental results are given to illustrate the concept.

1. INTRODUCTION

The development of aerial vehicles requires one to overcome several challenges including mechanical and geometrical conception, system modeling, state reconstruction, and control design. In the present paper, we focus more specifically on attitude estimation. Theoretically, it is possible to estimate the attitude just by integrating the rigid-body kinematic equation of rotation and using the angular velocity measurement data provided by gyroscopes. However, this solution is not safe for long-term applications due to the sensor noise and bias. Several alternative solutions have been proposed since the early 1960s'. The survey papers about attitude estimation methods Crassidis et al. (2007), Mahony et al. (2008), Hua (2009), Choukroun (2003) are useful references to start a research on the topic. In fact, the attitude can be algebraically determined by using the measurement, in the vehicle body-fixed frame, of at least two known non-collinear inertial directions (i.e. vectors) as explained in Wahba (1966), Davenport (1968), Shuster and Oh (1981), Markley and Mortari (2000). However, imperfect measurement and/or imprecise knowledge of the considered inertial directions can generate large errors on the extracted attitude. Thus, this rises the interest of fusing the measurement of known vectors with the angular velocity measurement provided by gyroscopes to obtain a more accurate and less noisy attitude estimation (cf. Lefferts et al. (1982), Vik and Fossen (2001), Thienel and Sanner (2003), Rehbinder and Ghosh (2003), Pflimlin et al. (2005), Hamel and Mahony (2006), Mahony et al. (2008), Bonnabel (2007), Martin and Salaun (2010), Vasconcelos et al. (2008), Tayebi et al. (2007), Sanyal et al. (2008)). In such approaches, not only the attitude but also the gyro-bias needs to be estimated. Nowadays, Micro-Electro-Mechanical systems (MEMS) components have revolutionized the robotics in general and aerial robotics in particular. Low-cost, light-weight MEMS inertial measurement units (IMUs) present an appealing solution for attitude estimation. The measurements of known inertial vectors are typically provided by accelerometers and magnetometers. Indeed, the accelerometers provide an approximate measurement of the gravity direction (with an assumption of weak accelerations of the body), whereas the magnetometers provide a measurement of the direction of the geomagnetic field. However, the signal output of low-cost IMUs are often corrupted by important noise and time-varying drift. Moreover, the magnetic measurement data can be influenced by important magnetic disturbances induced by the surrounding environment, electric motors and wires, etc. Therefore, an attitude estimation technique which effectively deals with gyro-bias and magnetic disturbances is of great interest in the context of low-cost IMUs. In this context, an attitude estimation algorithm recently proposed in Martin and Salaun (2010) ensures local decoupling of the gravity direction estimation from magnetic disturbances.

In the present paper, firstly, we adapt the nonlinear attitude estimation algorithm, called *"explicit complementary filter"*, proposed in Hamel and Mahony (2006), Mahony et al. (2008), for more effective gyro-bias compensation via the design of an anti-windup nonlinear integrator. In the context of IMUs, we extend the measurement decoupling strategy previously proposed in Martin and Salaun (2010) for our observer. Then, we show that this decoupling strategy ensures additional decoupling properties. Finally, improved performance has been confirmed through simulation and experimental results.

2. AN EXISTING ATTITUDE FILTER AND ITS ADAPTATION FOR MORE EFFECTIVE GYRO-BIAS COMPENSATION

2.1 Recalls on an existing attitude filter

The attitude of a rigid body can be represented by a rotation matrix $R \in SO(3)$ which describes the orientation

978-3-902661-93-7/11/$20.00 © 2011 IFAC **10.3182/20110828-6-IT-1002.00817**

of a frame $\mathcal{B}$ fixed to the body with respect to an inertial fixed frame $\mathcal{I}$. The kinematic equation of R satisfies the following equation

$$\dot{R} = R\Omega_\times \,, \tag{1}$$

where $\Omega \in \mathbb{R}^3$ is the body's angular velocity expressed in $\mathcal{B}$ and $(\cdot)_\times$ is the skew-symmetric matrix associated with the cross product $\times$, *i.e.* $x_\times y = x \times y$, $\forall x, y \in \mathbb{R}^3$.

In practice, the angular velocity vector Ω is typically measured by gyroscopes. For the sake of observer design and associated analyses, the measured angular velocity, denoted as Ω_y, is often modeled as the sum of the real angular velocity Ω with an unknown constant (or slowly time-varying in practice) bias vector $b \in \mathbb{R}^3$, *i.e.* $\Omega_y = \Omega + b$ (see, e.g., Mahony et al. (2008)).

Now, let us recall and discuss the *"explicit complementary filter"* proposed in Hamel and Mahony (2006), Mahony et al. (2008). Let $\{v_{0i}\}$ denotes a set of $n\,(\geq 2)$ *known non-collinear unit vectors* expressed in the inertial frame $\mathcal{I}$, and $\{v_i\}$ a set of measurement data of these vectors expressed in the body-fixed frame $\mathcal{B}$. By denoting $\hat{R}$ and $\hat{b}$ the estimates of R and b, the explicit complementary filter is written as

$$\begin{cases} \dot{\hat{R}} \;=\; \hat{R}\,(\Omega_y - \hat{b} + k_P\sigma)_\times, \quad R(0) \in SO(3) \\ \dot{\hat{b}} \;=\; -k_I\sigma, \quad \hat{b}(0) \in \mathbb{R}^3 \\ \sigma \;=\; \displaystyle\sum_{i=1}^{n} k_i v_i \times \hat{R}^\top v_{0i} \end{cases} \tag{2}$$

with k_P, k_I, k_i, $(i = 1, \cdots, n)$, denoting positive gains. As proved in Mahony et al. (2008), filter (2) ensures almost-global stability and local exponential stability of the equilibrium $(\tilde{R}, \tilde{b}) = (I_3, 0)$, with $\tilde{R} := \hat{R}^\top R$, $\tilde{b} := b - \hat{b}$, and I_3 the identity matrix. However, the usage of the integral correction term $\hat{b}$ may encounter some problems in practice, although it allows to compensate for the unknown constant bias vector b. For instance, the integral term $\hat{b}$ may grow arbitrarily large leading to slow desaturation (and slow convergence) and/or important overshoots of the estimation error variables (*i.e.* $\tilde{R}$ and $\tilde{b}$). Various sources of this phenomenon can be proposed such as : large initial errors, poor gain tuning, imperfect measurements of known vectors, and imprecise knowledge of the considered inertial vectors. Consequently, one may choose a very small value for the integral gain k_I to reduce overshoots, but this in turn degrades the estimation performance. In fact, this issue is well-known as the so-called integral wind-up effects, and has been largely studied in the literature Kothare et al. (1994), Seshagiri and Khalil (2005). However, in the context of attitude estimation design (on the group $SO(3)$), to our knowledge no solution to this issue is available in the literature.

2.2 Adaptation for more effective gyro-bias compensation

In what follows, we propose some modifications on the integral term $\hat{b}$ involved in the filter (2) so that *i)* the properties of convergence and stability of this filter are still ensured, and *ii)* integral wind-up effects can be limited.

Let $\mathrm{sat}_\Delta(\cdot)$ denotes the classical saturation function defined on $\mathbb{R}^3$, for some positive scalar Δ, as

$$\mathrm{sat}_\Delta(x) := x\,\min(1, \Delta/|x|)\,, \quad \forall x \in \mathbb{R}^3\,.$$

Then, instead of defining the correction term $\hat{b}$ as a solution to a pure integrator $\dot{\hat{b}} = -k_I\sigma$ we propose to use a "bounded" integral term $\hat{b}$ which is the solution to the following equation

$$\dot{\hat{b}} = -k_b\hat{b} + k_b\mathrm{sat}_\Delta(\,\hat{b}\,) - k_I\sigma, \quad |\hat{b}(0)| < \Delta\,, \tag{3}$$

with k_b, k_I denoting positive gains, and σ the innovation term specified in (2). Equation (3) indicates that $|\hat{b}(t)| \leq \bar{\Delta} := \Delta + k_I/k_b \sum_{i=1}^n k_i$, $\forall t \geq 0$, and that we have a pure integrator $\dot{\hat{b}} = -k_I\sigma$ as long as $|\hat{b}| \leq \Delta$. Moreover, equation (3) shows that k_b influences the rate of desaturation of $\hat{b}$ which can be observed, for instance, when $|\hat{b}|$ is larger than Δ and $\sigma = 0$. The larger the value of k_b the faster the rate of desaturation. The fact that *i)* the estimated gyro-bias $\hat{b}$ is bounded by the design threshold $\bar{\Delta}$ and *ii)* its rate of desaturation can be designed via the value of k_b, allows to reduce effectively integral wind-up effects.

From here, the main result of this section can be stated (see Mahony et al. (2008) for comparisons). The proof of the following proposition is similar to the proof of Theorem 5.1 in Mahony et al. (2008) and is only sketched in the present paper (see Appendix A).

Proposition 1. Consider the rotation kinematics (1). Consider the angular velocity measurement

$$\Omega_y = \Omega + b\,, \tag{4}$$

with $b \in \mathbb{R}^3$ an unknown constant bias. Assume that the set of $n\,(\geq 2)$ known non-collinear unit vectors of coordinates $\{v_{0i}, (i = 1, \cdots, n)\}$, expressed in the inertial frame $\mathcal{I}$, is measured to obtain the measurement set $\{v_i = R^\top v_{0i}\}$ expressed in the body-fixed frame $\mathcal{B}$. Consider the following observer

$$\begin{cases} \dot{\hat{R}} \;=\; \hat{R}\,(\Omega_y - \hat{b} + k_P\sigma)_\times, \quad R(0) \in SO(3) \\ \dot{\hat{b}} \;=\; -k_b\hat{b} + k_b\mathrm{sat}_\Delta(\,\hat{b}\,) - k_I\sigma, \quad |\hat{b}(0)| < \Delta \\ \sigma \;=\; \displaystyle\sum_{i=1}^{n} k_i v_i \times \hat{R}^\top v_{0i} \end{cases} \tag{5}$$

with k_P, k_I, k_i denoting positive gains, and $\Delta \geq |b|$. Choose k_i such that the matrix $M_0 := \sum_{i=1}^n k_i v_{i0} v_{i0}^\top$ has three distinct eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$. Define the error variables $\tilde{R} := \hat{R}^\top R$ and $\tilde{b} := b - \hat{b}$. Assume that Ω is a bounded and absolutely continuous signal. Then,

(1) There exist only three unstable equilibria defined by
$$(\hat{R}_{*i}, \hat{b}_{*i}) = (U_0 D_i U_0^\top R, b),\ (i = 1, 2, 3)\,,$$
with $D_1 = \mathrm{diag}(1, -1, -1)$, $D_2 = \mathrm{diag}(-1, 1, -1)$, $D_3 = \mathrm{diag}(-1, -1, 1)$, and $U_0 \in SO(3)$ such that $M_0 = U_0 \Lambda U_0^\top$, with $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$.
(2) The equilibrium $(\tilde{R}, \tilde{b}) = (I_3, 0)$ is locally exponentially stable.
(3) For any initial conditions such that $(\tilde{R}(0), \tilde{b}(0)) \neq (\hat{R}_{*i}^\top R, 0)$, $(i = 1, 2, 3)$, $(\hat{R}, \hat{b})$ converges to (R, b).

3. APPLICATION FOR IMUS

For analysis purposes, let us introduce Euler angles parametrization associated with the rotation matrices, knowing that singularities may occur for such a minimal parametrization. Let us denote ϕ, θ, ψ as the Euler angles corresponding to the parameters of roll, pitch, and yaw,

commonly used in the aerospace field. Then, the attitude matrix R can be written as

$$R=\begin{bmatrix} C\theta C\psi & S\phi S\theta C\psi - C\phi S\psi & C\phi S\theta C\psi + S\phi S\psi \\ C\theta S\psi & S\phi S\theta S\psi + C\phi C\psi & C\phi S\theta S\psi - S\phi C\psi \\ -S\theta & S\phi C\theta & C\phi C\theta \end{bmatrix}, \quad (6)$$

where C and S correspond to $\cos(\cdot)$ and $\sin(\cdot)$ operators.

In this section, we deal with practical problems concerning the application of the explicit complementary filters (2) and (5) for IMUs. There are two main problems which are widely discussed in various forums (see, e.g., Premerlani and Bizard (2009)):

(1) The dynamics of roll, pitch, and yaw angles are highly coupled. This implies that the estimation of the yaw angle strongly affects the estimation of the roll and pitch angles.
(2) Magnetic disturbances and bias influence the estimations of roll and pitch angles.

3.1 Coupling problem of the explicit complementary filters based on IMU measurement data

Let us assume that the IMU fixed to the body consists of a 3-axis gyroscope, a 3-axis accelerometer, and a 3-axis magnetometer.

- The 3-axis accelerometer measures the specific acceleration $a_{\mathcal{B}}$ expressed in the body frame $\mathcal{B}$. One has $a_{\mathcal{B}} = R^\top(\ddot{x} - ge_3)$, where the vehicle's acceleration expressed in the inertial frame $\mathcal{I}$ is $\ddot{x}$, and the gravitational acceleration expressed in the inertial frame $\mathcal{I}$ is ge_3, with $e_3 = (0,0,1)^\top$.
- The magnetometer measurement data is normalized to obtain $m_{\mathcal{B}} \in S^2$, the normalized magnetic field expressed in the body-fixed frame $\mathcal{B}$. If the magnetometer measurement data is not corrupted by magnetic disturbances, then one has $m_{\mathcal{B}} = R^\top m_{\mathcal{I}}$, with $m_{\mathcal{I}} \in S^2$ the normalized geomagnetic field expressed in the inertial frame $\mathcal{I}$. In Europe, the third component of $m_{\mathcal{I}}$ is dominant compared to others.

It is well known that the measurement of the gravity direction cannot be directly extracted from accelerometer measurement data due to the vehicle's acceleration $\ddot{x}$ involved in the relation $a_{\mathcal{B}} = R^\top(\ddot{x} - ge_3)$. In practice, an assumption that the vehicle's acceleration $\ddot{x}$ is negligible compared to the gravitational acceleration (*i.e.* $|\ddot{x}| \ll g$) [1] is often made so that the 3-axis accelerometer can be used as an inclinometer to measure the gravity direction, *i.e.* $a_{\mathcal{B}} \approx -gR^\top e_3$. From here, one can directly apply the explicit filter (2) or (5), with $n = 2$ and

$$v_1 = a_{\mathcal{B}}/|a_{\mathcal{B}}|, v_{01} = -e_3, v_2 = m_{\mathcal{B}}, v_{02} = m_{\mathcal{I}}\,.$$

Let us call these solutions as *"common observers"* to distinguish them from the *"enhanced observers"* described hereafter. In cases of weak accelerations of the vehicle and without magnetic disturbances, good performance can be expected. However, in many applications especially for small-size electric motorized aerial robots, important magnetic disturbances are almost unavoidable, leading to important difference between $m_{\mathcal{B}}$ and $R^\top m_{\mathcal{I}}$. This not only leads to large estimation errors of the yaw angle ψ, but also non-negligible errors in the roll and pitch estimation. Interestingly, in view of (6), ϕ and θ can be directly deduced from the vector $R^\top e_3$ which can be approximated by $a_{\mathcal{B}}$ in the case of weak accelerations of the vehicle. As a consequence, theoretically, estimating ϕ and θ can be done independently from magnetic measurement.

[1] To deal with strong acceleration motions, GPS velocity measurement can be fused with IMU measurement to estimate the vehicle's acceleration and thus improve the attitude estimation (see Martin and Salaun (2008), Hua (2010)).

3.2 Enhanced explicit complementary filters

A solution recently proposed in Martin and Salaun (2010) has been proved to *locally* decouple the estimation of pitch and roll angles from magnetic disturbances. Let us describe this strategy. Assume that $a_{\mathcal{B}} = -gR^\top e_3$ and compute the following orthogonal vectors

$$u_{\mathcal{B}} = \frac{a_{\mathcal{B}}}{|a_{\mathcal{B}}|}, v_{\mathcal{B}} = \frac{a_{\mathcal{B}} \times m_{\mathcal{B}}}{|a_{\mathcal{B}} \times m_{\mathcal{B}}|}, u_{\mathcal{I}} = -e_3, v_{\mathcal{I}} = \frac{u_{\mathcal{I}} \times m_{\mathcal{I}}}{|u_{\mathcal{I}} \times m_{\mathcal{I}}|} \quad (7)$$

Then, we apply the explicit filter (2) or (5), with $n = 2$, $v_1 = u_{\mathcal{B}}$, $v_{01} = u_{\mathcal{I}}$, $v_2 = v_{\mathcal{B}}$, $v_{02} = v_{\mathcal{I}}$, and name these solutions as *"enhanced observers"*. In what follows, we will show that these observers ensure additional decoupling properties. More precisely, if $a_{\mathcal{B}} = -gR^\top e_3$, then :

(1) In ideal situations, *i.e.* $m_{\mathcal{B}} = R^\top m_{\mathcal{I}}$, the (local) decoupling of roll/pitch estimation from yaw estimation is ensured.
(2) In the presence of a constant magnetic disturbance field or a magnetic bias, the estimated roll and pitch angles still converge to the real values for almost all initial conditions.

▷ ***Property 1: Local decoupling of roll/pitch estimation from yaw estimation***

In ideal situations, *i.e.* $m_{\mathcal{B}} = R^\top m_{\mathcal{I}}$, it is straightforward to deduce the convergence property of the "enhanced observers" since $v_i = R^\top v_{0i}$, with $i = 1, 2$. Furthermore, we will show that these "enhanced observers" ensure the local decoupling of roll and pitch angles estimation from yaw angle estimation, whereas aforementioned "common observers" do not. For the sake of clarity of analysis let us, for instance, neglect the dynamic of the gyro-bias estimation term $\hat{b}$ and suppose that $b = 0$. In that case, both filters (2) and (5) for the "common observers" and the "enhanced observers" write

$$\begin{cases} \dot{\hat{R}} = \hat{R}\,(\Omega + k_P\sigma)_\times, \\ \sigma = k_1 v_1 \times \hat{R}^\top v_{01} + k_2 v_2 \times \hat{R}^\top v_{02} \end{cases}$$

Then, one verifies that the dynamics of the new attitude error $\bar{R} := R\hat{R}^\top$ satisfies

$$\dot{\bar{R}} = -\bar{k}_1(v_{01} \times \bar{R}^\top v_{01})_\times \bar{R} - \bar{k}_2(v_{02} \times \bar{R}^\top v_{02})_\times \bar{R}\,, \quad (8)$$

with $\bar{k}_1 := k_P k_1$, $\bar{k}_2 := k_P k_2$. Denote the unit quaternion associated with the rotation matrix $\bar{R}$ as $\bar{q} := (\bar{q}_0, \bar{\mathbf{q}})^\top$, with $\bar{q}_0 \in \mathbb{R}$ and $\bar{\mathbf{q}} \in \mathbb{R}^3$ the real part and pure part of $\bar{q}$, respectively. From Rodrigues' rotation formula

$$\bar{R} = I_3 + 2\bar{q}_0\bar{\mathbf{q}}_\times + 2(\bar{\mathbf{q}}_\times)^2\,,$$

in first-order approximations about the equilibrium $\bar{R} = I_3$, one has $\bar{R} \approx I_3 + 2\bar{\mathbf{q}}_\times$. This and (8) implies that the linearized system of (8) about the equilibrium $\bar{R} = I_3$ can be written as follows

$$\begin{aligned} \dot{\bar{\mathbf{q}}} &= -\bar{k}_1\, v_{01} \times (\bar{\mathbf{q}} \times v_{01}) - \bar{k}_2\, v_{02} \times (\bar{\mathbf{q}} \times v_{02}) \\ &= \bar{k}_1(v_{01}v_{01}^\top - I_3)\bar{\mathbf{q}} + \bar{k}_2(v_{02}v_{02}^\top - I_3)\bar{\mathbf{q}}\,. \end{aligned} \quad (9)$$

As for the "common solutions", one has $v_{01} = -e_3$ and $v_{02} = m_{\mathcal{I}}$, where the third component of v_{02} satisfies

$v_{02,3} = m_{\mathcal{I},3} \neq 0$. From (9), it is straightforward to verify that the dynamics of all three components of $\bar{\mathbf{q}}$ are coupled. On the other hand, as for the "enhanced solutions", thanks to the specific construction of vectors (*i.e.* $v_1 = u_\mathcal{B}$, $v_{01} = u_\mathcal{I}$, $v_2 = v_\mathcal{B}$, $v_{02} = v_\mathcal{I}$, with $u_\mathcal{B}$, $u_\mathcal{I}$, $v_\mathcal{B}$, $v_\mathcal{I}$ defined by (7)) one verifies that the third components of v_{02} is null and that (9) can be decoupled into two independent stable subsystems

$$\begin{bmatrix}\dot{\bar{\mathbf{q}}}_1\\ \dot{\bar{\mathbf{q}}}_2\end{bmatrix} = \begin{bmatrix}-\bar{k}_1 - \bar{k}_2 v_{02,2}^2 & \bar{k}_2 v_{02,1} v_{02,2}\\ \bar{k}_2 v_{02,1} v_{02,2} & -\bar{k}_1 - \bar{k}_2 v_{02,1}^2\end{bmatrix}\begin{bmatrix}\bar{\mathbf{q}}_1\\ \bar{\mathbf{q}}_2\end{bmatrix} \tag{10a}$$

$$\dot{\bar{\mathbf{q}}}_3 = -\bar{k}_2 \bar{\mathbf{q}}_3 \tag{10b}$$

Remark 1. System (10a)–(10b) is a critically damped system for any choice of positive gains k_1, k_2, and k_P.

Denote $\hat{\phi}$, $\hat{\theta}$, $\hat{\psi}$ as the roll, pitch, and yaw angles associated with $\hat{R}$, and define $\tilde{\phi} := \phi - \hat{\phi}$, $\tilde{\theta} := \theta - \hat{\theta}$, $\tilde{\psi} := \psi - \hat{\psi}$ as the Euler angles estimation errors. Using the fact that $\bar{R} = R\hat{R}^\top \approx I_3 + 2\bar{\mathbf{q}}_\times$ and the relation (6), one can verify that in first-order approximations

$$\begin{cases}\bar{\mathbf{q}}_1 = (\tilde{\phi}\, C\theta\, C\psi - \tilde{\theta}\, S\psi)/2\\ \bar{\mathbf{q}}_2 = (\tilde{\phi}\, S\theta\, S\psi + \tilde{\theta}\, C\psi)/2\\ \bar{\mathbf{q}}_3 = (-\tilde{\phi}\, S\theta + \tilde{\psi})/2\end{cases}$$

From these relations, one remarks that $\tilde{\phi}$ and $\tilde{\theta}$ depend on $\bar{\mathbf{q}}_1$ and $\bar{\mathbf{q}}_2$, and do not depend on $\bar{\mathbf{q}}_3$, and that only $\tilde{\psi}$ depends on $\bar{\mathbf{q}}_3$. This and system (10a)–(10b) imply that the estimations of roll and pitch angles are locally decoupled with the yaw angle estimation.

▷ ***Property 2: Decoupling of roll/pitch estimation from magnetic bias***

Let us consider situations where magnetometer measurement data are influenced by a constant (or slowly time varying in practice) magnetic bias in the inertial frame $\mathcal{I}$. In this case, the vector of coordinates $m_\mathcal{I}^b$ of the biased magnetic field expressed in the inertial frame $\mathcal{I}$ is different from the geomagnetic field $m_\mathcal{I}$. Therefore, $m_\mathcal{B} = R^\top m_\mathcal{I}^b \neq R^\top m_\mathcal{I}$ and consequently $v_\mathcal{B} \neq R^\top v_\mathcal{I}$, with $v_\mathcal{B}$ and $v_\mathcal{I}$ defined in (7). Thus, the convergence of $\hat{R}$ to R is no longer ensured. In this case, instead of having $v_\mathcal{B} = R^\top v_\mathcal{I}$, one has $v_\mathcal{B} = \mathcal{R}^\top v_\mathcal{I}^b$, with $v_\mathcal{I}^b$ a unit constant vector orthogonal to $u_\mathcal{I}$ and different from $v_\mathcal{I}$. The fact that $v_\mathcal{I}^b$ is orthogonal to $u_\mathcal{I}$ is justified by the fact that $v_\mathcal{B}$ is orthogonal to $u_\mathcal{B}$ by construction. The fact that $v_\mathcal{I}^b$ is constant (or slowly time varying) is justified by the assumption that the biased magnetic field is constant (or slowly time varying).

As a consequence, there exists a unique constant angle $\Theta \in [0; \pi)$ such that

$$v_\mathcal{I}^b = \cos\Theta\, v_\mathcal{I} + \sin\Theta\, (u_\mathcal{I} \times v_\mathcal{I})\,.$$

Rodrigues' rotation formula allows to construct the following constant rotation matrix

$$R^b = I_3 + \sin\Theta\, u_{\mathcal{I}\times} + 2(\sin\frac{\Theta}{2})^2 (u_{\mathcal{I}\times})^2\,.$$

From here, one verifies that $R^b u_\mathcal{I} = u_\mathcal{I}$ and $R^b v_\mathcal{I} = v_\mathcal{I}^b$, where the latter relation is obtained as follows

$$\begin{aligned}R^b v_\mathcal{I} &= v_\mathcal{I} + 2(\sin\frac{\Theta}{2})^2 u_\mathcal{I} \times (u_\mathcal{I} \times v_\mathcal{I}) + \sin\Theta (u_\mathcal{I} \times v_\mathcal{I})\\ &= v_\mathcal{I} + 2(\sin\frac{\Theta}{2})^2 (u_\mathcal{I}(u_\mathcal{I}^\top v_\mathcal{I}) - v_\mathcal{I}|u_\mathcal{I}|^2) + \sin\Theta (u_\mathcal{I} \times v_\mathcal{I})\\ &= \cos\Theta v_\mathcal{I} + \sin\Theta (u_\mathcal{I} \times v_\mathcal{I}) = v_\mathcal{I}^b\,.\end{aligned}$$

By denoting $\hat{R}^b := R^b\, \hat{R}$ and using $\dot{R}^b = 0$, one verifies

$$\dot{\hat{R}}^b = \dot{R}^b \hat{R} + R^b \dot{\hat{R}} = \hat{R}^b (\Omega_y - \hat{b} + k_P \sigma)_\times\,, \tag{11}$$

where $\hat{b}$ is calculated from the second equation of either (2) or (5), and the innovation term σ is rewritten as

$$\begin{aligned}\sigma &= k_1 u_\mathcal{B} \times \hat{R}^\top u_\mathcal{I} + k_2 v_\mathcal{B} \times \hat{R}^\top v_\mathcal{I}\\ &= k_1 u_\mathcal{B} \times \hat{R}^{b\top} R^b u_\mathcal{I} + k_2 v_\mathcal{B} \times \hat{R}^{b\top} R^b v_\mathcal{I}\\ &= k_1 u_\mathcal{B} \times (\hat{R}^{b\top} u_\mathcal{I}) + k_2 v_\mathcal{B} \times (\hat{R}^{b\top} v_\mathcal{I}^b)\,,\end{aligned} \tag{12}$$

with $u_\mathcal{B} = R^\top u_\mathcal{I}$ and $v_\mathcal{B} = R^\top v_\mathcal{I}^b$. From here, one verifies that the system (11)–(12) has the same form as the filter (2) or (5), depending on the dynamics chosen for $\hat{b}$. As a consequence, one ensures that $(\hat{R}^b, \hat{b})$ converges to (R, b) for almost all initial conditions. Using the fact that $\hat{R}^b = R^b\, \hat{R}$ converges to R and that $R^b u_\mathcal{I} = u_\mathcal{I}$ one deduces that $\hat{R}^\top u_\mathcal{I} = \hat{R}^{b\top} R^b u_\mathcal{I} = \hat{R}^{b\top} u_\mathcal{I}$ converges to $R^\top u_\mathcal{I}$. More precisely,

$$\hat{R}^\top u_\mathcal{I} = \begin{bmatrix}S\hat{\theta}\\ -C\hat{\theta}C\hat{\phi}\\ -C\hat{\theta}C\hat{\phi}\end{bmatrix} \longrightarrow R^\top u_\mathcal{I} = \begin{bmatrix}S\theta\\ -C\theta C\phi\\ -C\theta C\phi\end{bmatrix},$$

which implies that the estimated roll and pitch angles $\hat{\theta}$, $\hat{\phi}$ converge to the real values θ, ϕ despite the magnetic bias. Furthermore, it is worth noting that the estimated gyro-bias vector $\hat{b}$ also converges to the real value b. In summary, the *almost global convergence* of the enhanced observers is still ensured in the presence of constant bias on the magnetic field.

Remark 2. The assumption that the magnetic field is biased by a constant value is well justified in practice because the dynamics of the vehicle is often slower than the filter dynamics. The decoupling of the estimations of roll and pitch angles with a time-varying magnetic field is not proved in the present paper. However, simulation results tend to support this idea.

4. SIMULATION RESULTS

In this section, we illustrate through simulation results the improved performance of the *"enhanced observer"* corresponding to filter (5), compared to the classical implementation (*i.e.* the *"common observer"*) of filter (2) proposed in Hamel and Mahony (2006), Mahony et al. (2008) for IMUs. To this purpose, simulations are carried on for the following two filters:

▷ **Filter A – enhanced observer**– corresponds to filter (5), with $n = 2$, $v_1 = u_\mathcal{B}$, $v_{01} = u_\mathcal{I}$, $v_2 = v_\mathcal{B}$, $v_{02} = v_\mathcal{I}$, and $u_\mathcal{B}$, $u_\mathcal{I}$, $v_\mathcal{B}$, $v_\mathcal{I}$ defined by (7), and with $k_1 = 1.4$, $k_2 = 0.8$, $k_P = 1$, $k_I = 0.1$, $k_b = 10$, $\Delta = 0.03$.

▷ **Filter B – common observer**– corresponds to filter (2), with $n = 2$, $v_1 = a_\mathcal{B}/|a_\mathcal{B}|$, $v_{01} = -e_3$, $v_2 = m_\mathcal{B}$, $v_{02} = m_\mathcal{I}$, and with $k_1 = 1.4$, $k_2 = 0.8$, $k_P = 1$, $k_I = 0.1$.

One notes that the sets of gains $\{k_1, k_2, k_P, k_I\}$ of the two filters are chosen equal. The gain k_1 is chosen larger than k_2 since we assume that the measurement of the gravity direction is more reliable than the measurement of the geomagnetic field. The integral gain k_I is chosen small compared the proportional gain k_P to reduce integral wind-up effects. The value of Δ chosen for Filter A corresponds to an estimated bound of each component of the gyro-bias vector b equal to 1deg/s. The gain k_b involved

in Filter A is chosen large (*i.e.* $k_b = 10$) in order to obtain a fast desaturation rate of $\hat{b}$.

Simulations are carried on for the following scenario: an IMU is fixed to a vertical take-off and landing (VTOL) vehicle which is in stationary flight so that its attitude matrix R is maintained at I_3, *i.e.* $\phi = \theta = \psi = 0$. The normalized geomagnetic field expressed in the inertial frame $\mathcal{I}$ is taken as $m_I = (0.434, -0.0091, 0.9008)^\top$. Two simulations are reported.

▷ **Simulation1** : allows to compare the performance of the two filters in the case of no sensor noise and good knowledge of the magnetic field, *i.e.* $m_\mathcal{I}$. In turn, a constant gyro-bias vector $b = (0.01, 0.005, -0.01)^\top$(rad/s) is introduced. The initial estimated Euler angles associated with the initial estimated attitude matrix $R(0)$ are rather large, *i.e.* $(\hat{\phi}(0), \hat{\theta}(0), \hat{\psi}(0)) = (-30, 30, 90)$(deg), and the initial estimated gyro-bias is taken as $\hat{b}(0) = (0, 0, 0)^\top$(rad/s). The results illustrated in Figures 1 and 2 show important performance differences between the proposed enhanced explicit complementary filter (*i.e.* Filter A) and the classical implementation of the explicit complementary filter (2) (*i.e.* Filter B). In particular, the latter would yield important overshoots and oscillations in both the estimated Euler angles and the estimated gyro-bias components. One can also observe, for Filter A, a very fast convergence of the estimated variables to the real value, and the quasi absence of overshoots of the estimated attitude despite the integral correction term (*i.e.* $\hat{b}$) and large initial estimation errors.

▷ **Simulation2** : allows to illustrate the influences of magnetic disturbances on the performance of the two filters. The same constant gyro-bias vector b and initial estimated gyro-bias $\hat{b}(0)$ as in Simulation 1 is introduced. The initial estimated Euler angles associated with the initial estimated attitude matrix $R(0)$ are now chosen small, *i.e.* $(\hat{\phi}(0), \hat{\theta}(0), \hat{\psi}(0)) = (-5, 5, 10)$(deg), for the sake of visualization. The magnetic field $m_\mathcal{I}$ given previously is corrupted as follows to obtained the normalized vector of coordinates $m^b_\mathcal{I}$, expressed in the inertial frame $\mathcal{I}$, of the *"true"* magnetic field:

$$m^b_\mathcal{I} = \frac{m_\mathcal{I} + m_{\text{periodic}} + m_{\text{bias}}}{|m_\mathcal{I} + m_{\text{periodic}} + m_{\text{bias}}|},$$

with $m_{\text{periodic}} = (0.2\sin(\pi t), 0.2\cos(\pi t), 0)^\top$, $m_{\text{bias}} = (0.4, -0.8, 0.2)^\top$. Furthermore, each component of the measurement vector $m_\mathcal{B}$ of the unit vector $R^\top m^b_\mathcal{I}$ is also corrupted by an additive white Gaussian noise of variance 0.01. Figures 3 and 4 show clearly a better performance of Filter A compared to that of Filter B. One can observe that, for the enhanced complementary explicit filter (*i.e* Filter A), the magnetic disturbances do not degrade the performance of the estimations of roll and pitch angles and the estimations of the first and second components of the gyro-bias vector b, whereas all the estimated variables of Filter B are strongly effected.

5. EXPERIMENTAL RESULTS

The experiments were performed on a quadricopter equipped with a low-cost IMU composed of a three-axis accelerometer (LIS344ALH) and three single-axis gyroscopes (ADXRS610), and a magnetometer (KMZ51 magnetic sensor). The attitude estimated by the algorithm of this paper is compared with "ground truth" measurement data acquired by a motion capture system from Vicon. This vision-based system is composed of 8 cameras mounted on the ceiling of the flying machine arena (FMA) of the ETH Zurich at the IDSC laboratory, where the experiments took place. The Vicon system provides the attitude of the flying vehicle in terms of Euler angles at a rate of 200 Hz.

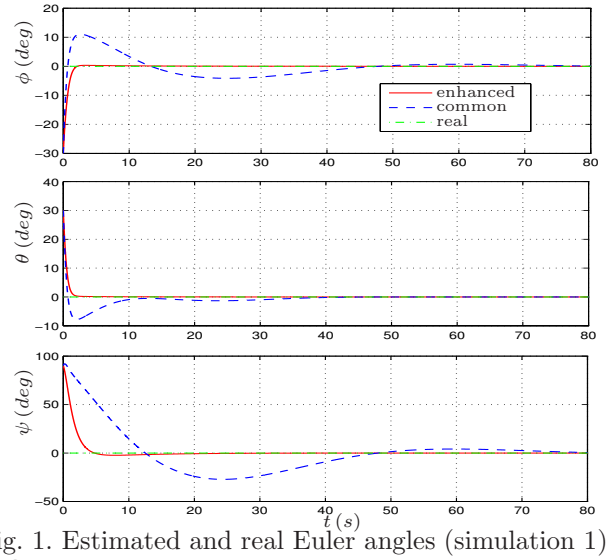

Fig. 1. Estimated and real Euler angles (simulation 1).

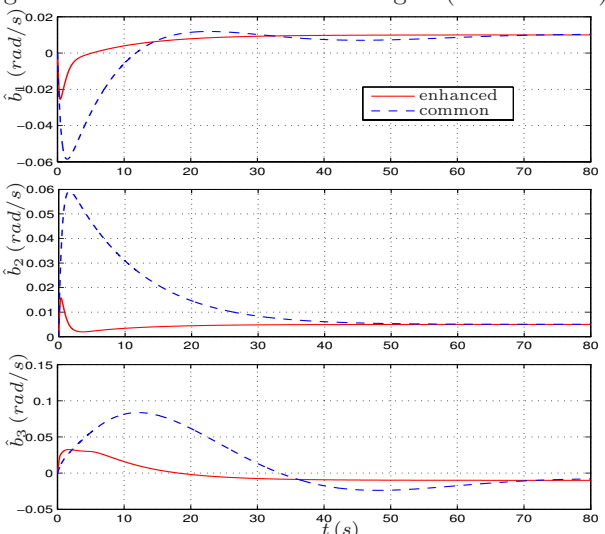

Fig. 2. Estimated gyro-bias (simulation 1).

A discretized version of the attitude estimator is implemented on the computer aboard the quadricopter with a sampling period of 200 Hz. Let us provide more details. Since it is difficult to preserve the evolution of the estimated matrix $\hat{R}$ on the group $SO(3)$, the quaternion is used here. Let $\hat{q}$ denote the unit quaternion associated with $\hat{R}$. By denoting $\hat{\Omega} := \Omega_y - \hat{b} + k_P\sigma$, the dynamics of $\hat{q}$ satisfies

$$\dot{\hat{q}} = \frac{1}{2}A(\hat{\Omega})\hat{q}\,, \text{ with } A(\hat{\Omega}) := \begin{bmatrix} 0 & -\hat{\Omega}^\top \\ \hat{\Omega} & -\hat{\Omega}_\times \end{bmatrix}.$$

Since the sample time τ is small enough, we assume that $\hat{\Omega}(t)$ approximately remains constant in every period of time $[k\tau, (k+1)\tau]$, $\forall k \in \mathbb{N}$, and denote this value as $\hat{\Omega}_k$. Then, by exact integration we obtain

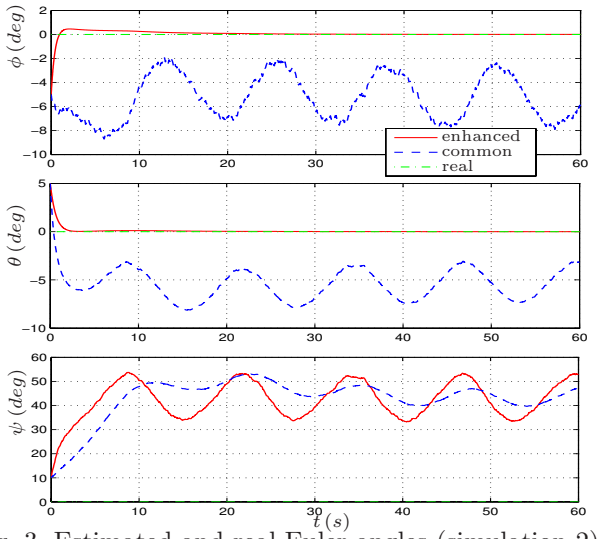

Fig. 3. Estimated and real Euler angles (simulation 2).

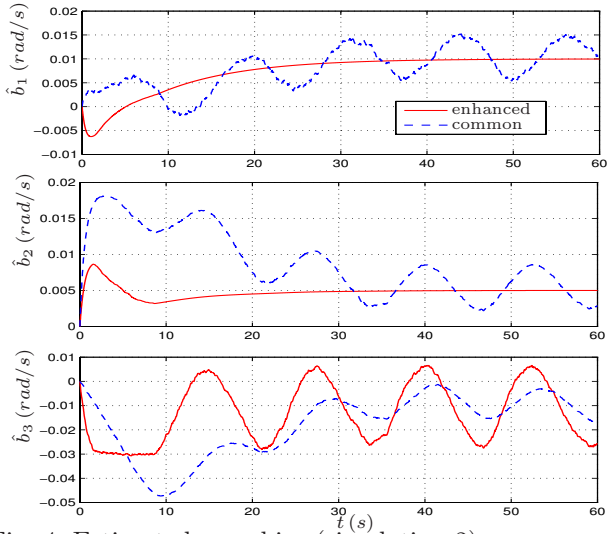

Fig. 4. Estimated gyro-bias (simulation 2).

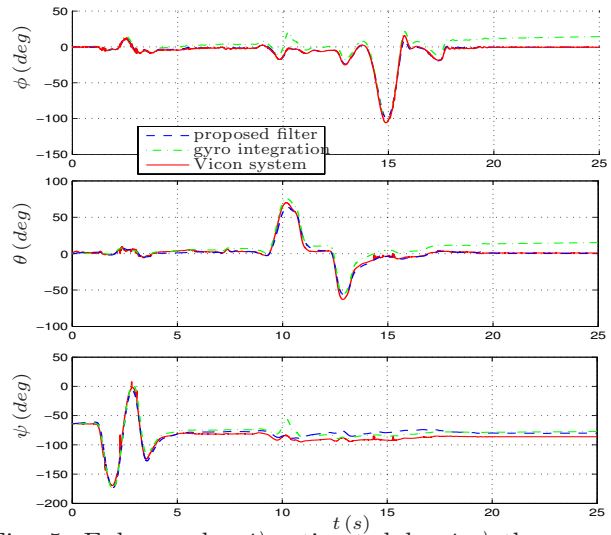

Fig. 5. Euler angles *i)* estimated by *i.a)* the proposed enhanced explicit complementary filter, *i.b)* simply integrating the kinematics equation of rotation using gyroscope measurement data, and *ii)* measured by the Vicon system

.

$$\hat{q}_{k+1} = \exp(\frac{\tau}{2}A(\hat{\Omega}_k))\hat{q}_k\,.$$

Using the fact that $A(\hat{\Omega}_k)^2 = -|\hat{\Omega}_k|^2 I_4$, with I_4 the identity matrix of $\mathbb{R}^{4\times4}$, one can verify that

$$\exp\left(\frac{\tau}{2}A(\hat{\Omega}_k)\right) = \cos\left(\frac{\tau}{2}|\hat{\Omega}_k|\right) I_4 + \frac{\tau}{2}\mathrm{sinc}\left(\frac{\tau}{2}|\hat{\Omega}_k|\right)A(\hat{\Omega}_k),$$

with $\mathrm{sinc(s)} := \sin(s)/s, \forall s \in \mathbb{R}$. As a consequence, the discretized version of the explicit filter (5) is given by

$$\begin{cases} \hat{q}_{k+1} = \left(\cos\left(\frac{\tau}{2}|\hat{\Omega}_k|\right) I_4 + \frac{\tau}{2}\mathrm{sinc}\left(\frac{\tau}{2}|\hat{\Omega}_k|\right) A(\hat{\Omega}_k)\right)\hat{q}_k \\ \hat{b}_{k+1} = \tau\left(-k_b\hat{b}_k + k_b\mathrm{sat}_\Delta(\hat{b}_k) - k_I\sigma_k\right) + \hat{b}_k \end{cases}$$

The gains and parameters involved in this filter correspond to the ones of Filter A defined in Section 4. The experimental results reported in Fig. 5 show a test where the platform was moved by hand. The estimated values for the pitch and roll angles are very closed to the ground truth, meaning that the bias has been properly estimated and removed. Regarding the estimation of the yaw angle, it can be seen the presence of a bias. This may be explained by the fact that the inertial magnetic field vector inside the FMA is not exactly known and might be slightly perturbed by the electrical equipment in this area. The presence of magnetic field disturbance does not prevent the estimates of the roll and pitch angles to converge to the true value, thus confirming experimentally the discussion about decoupling in Section 3.2.

6. PERSPECTIVES

For the proposed enhanced complimentary filter applied for IMUs, simulation and experimental results seem to validate the local decoupling property of the estimations of roll and pitch angles from the (time-varying) magnetic disturbances. However, analyzing this property, for the case of time-varying magnetic disturbances, is still an open problem. In order to ensure global decoupling, the following solution may be proposed

$$\begin{cases} \dot{\hat{R}} = \hat{R}\,(\Omega_y - \hat{b} + k_P\sigma)_\times, \quad R(0) \in SO(3) \\ \dot{\hat{b}} = -k_b\hat{b} + k_b\mathrm{sat}_\Delta(\hat{b}) - k_I\sigma, \quad |\hat{b}(0)| < \Delta \\ \sigma = k_1 u_\mathcal{B} \times \hat{R}^\top u_\mathcal{I} + k_2((m_\mathcal{B} \times \hat{R}^\top m_\mathcal{I})^\top u_\mathcal{B})u_\mathcal{B} \end{cases} \tag{13}$$

with $u_\mathcal{I}$ and $u_\mathcal{B}$ defined in (7). The sole difference of the filter (13) and the enhanced complementary filter corresponding to (5) (proposed in Section 3.2) lies in the definition of the innovation term σ. For the filter (13), we are able to prove the decoupling of roll and pitch angles estimation from a time-varying magnetic field for *i)* the case of gyro-bias free, *i.e.* $b \equiv \hat{b} \equiv 0$, and for *ii)* the case of constant gyro-bias b and no rotation motion, *i.e.* $\Omega \equiv 0$. The proof of convergence and decoupling of the filter (13) for other cases, however, remains an open problem.

7. ACKNOWLEDGMENTS

Special thanks go to the team of the Institute for Dynamic Systems and Control (IDSC) at the ETH Zurich for the possibility to use their vision-based motion capture Vicon system as the ground truth to validate the algorithm for the attitude estimation of this paper. The authors also acknowledge the French *Agence National de la Recherche* (ANR) within the ANR SCUAV project and the French *Fonds Unique Interministériel* (FUI) within the FUI ADOPIC project for the financial support.

REFERENCES

Bonnabel, S. (2007). Left-invariant extended Kalman filter and attitude estimation. In *IEEE Conference on Decision and Control*, 1027–1032.

Choukroun, D. (2003). *Novel methods for attitude determination using vector observations*. Ph.D. thesis, Israel Institute of Technology.

Crassidis, J.L., Markley, F.L., and Cheng, Y. (2007). Survey of nonlinear attitude estimation methods. *AIAA Journal of Guidance, Control, and Dynamics*, 30(1), 12–28.

Davenport, P.B. (1968). A vector approach to the algebra of rotations with applications. Technical Report TN D-4696, NASA.

Hamel, T. and Mahony, R. (2006). Attitude estimation on SO(3) based on direct inertial measurements. In *IEEE Conference on Robotics and Automation*, 2170–2175.

Hua, M.D. (2009). *Contributions to the automatic control of aerial vehicles*. Ph.D. thesis, Université de Nice-Sophia Antipolis. Available at http://tel.archives-ouvertes.fr/tel-00460801.

Hua, M.D. (2010). Attitude estimation for accelerated vehicles using GPS/INS measurements. *Control Engineering Practice*, 723–732.

Kothare, M.V., Campo, P.J., Morari, M., and Nett, C.N. (1994). A unified framework for the study of anti-windup designs. *Automatica*, 30(12), 1869–1883.

Lefferts, E., Markley, F., and Shuster, M. (1982). Kalman filtering for spacecraft attitude estimation. *AIAA Journal of Guidance, Control, Navigation*, 5, 417–429.

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

Markley, F.L. and Mortari, D. (2000). Quaternion attitude estimation using vector observations. *The Journal of the Astronautical Sciences*, 48(2, 3), 359–380.

Martin, P. and Salaun, E. (2008). An invariant observer for Earth-Velocity-Aided attitude heading reference systems. In *IFAC World Congress*, 9857–9864.

Martin, P. and Salaun, E. (2010). Design and implementation of a low-cost observer-based attitude and heading reference system. *Control Engineering Practice*, 712–722.

Pflimlin, J.M., Hamel, T., Souères, P., and Metni, N. (2005). Nonlinear attitude and gyroscopes bias estimation for a VTOL UAV. In *IFAC World Congress*.

Premerlani, W. and Bizard, P. (2009). Dcm imu theory : first draft. Technical report. Available at http://diydrones.ning.com/profiles/blogs/dcm-imu-theory-first-draft?id=705844.

Rehbinder, H. and Ghosh, B. (2003). Pose estimation using line-based dynamic vision and inertial sensors. *IEEE Transactions on Automatic Control*, 48(2), 186–199.

Sanyal, A., Lee, T., Leok, M., and McClamroch, N.H. (2008). Global optimal attitude estimation using uncertainty ellipsoids. *Systems & Control Letters*, 57, 236–245.

Seshagiri, S. and Khalil, H. (2005). Robust output feed back regulation of minimum-phase nonlinear systems using conditional integrators. *Automatica*, 41, 43–54.

Shuster, M.D. and Oh, S.D. (1981). Three-axis attitude determination from vector observations. *AIAA Journal of Guidance, Control, and Dynamics*, 4(1), 70–77.

Tayebi, A., McGilvray, S., Roberts, A., and Moallem, M. (2007). Attitude estimation and stabilization of a rigid body using low-cost sensors. In *IEEE Conference on Decision and Control*, 6424–6429.

Thienel, J. and Sanner, R.M. (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.F., Silvestre, C., and Oliveira, P. (2008). A nonlinear GPS/IMU based observer for rigid body attitude and position estimation. In *IEEE Conference on Decision and Control*, 1255–1260.

Vik, B. and Fossen, T. (2001). A nonlinear observer for GPS and INS integration. In *IEEE Conference on Decision and Control*, 2956–2961.

Wahba, G. (1966). A least squares estimate of satellite attitude. *SIAM Review*, 8(3), 384–386.

Appendix A. SKETCH OF PROOF OF PROPOSITION 1

The proof is based on a Lyapunov analysis. Consider the candidate Lyapunov function

$$\mathcal{L} := \sum_{i=1}^{n} k_i - \mathrm{tr}(\tilde{R}M) + \frac{1}{k_I}|\tilde{b}|^2\,, \tag{A.1}$$

with $M := \sum_{i=1}^{n} k_i R^\top v_{0i} v_{0i}^\top R$. Now, let us calculate the time-derivative of $\mathcal{L}$ and show that it is negative semi-definite. From (1), the first equation of (5), and the input (4), one verifies that

$$\dot{\tilde{R}} = (-\Omega_\times \tilde{R} + \tilde{R}\Omega_\times) - \tilde{b}_\times \tilde{R} - k_P \sigma_\times \tilde{R}\,, \tag{A.2}$$

$$\dot{\tilde{b}} = -\dot{\hat{b}} = -k_b \tilde{b} + k_b(b - \mathrm{sat}_\Delta(b - \tilde{b})) + k_I \sigma\,, \tag{A.3}$$

$$\dot{M} = -\Omega_\times M + M\Omega_\times\,. \tag{A.4}$$

From (A.1), (A.2), (A.4) one obtains

$$\begin{aligned}\dot{\mathcal{L}} &= -\mathrm{tr}((-\Omega_\times \tilde{R} + \tilde{R}\Omega_\times)M + (-\Omega_\times M + M\Omega_\times)\tilde{R}) \\ &\quad + \mathrm{tr}((\tilde{b} + k_P\sigma)_\times \tilde{R}M) + \frac{2}{k_I}\tilde{b}^\top \dot{\tilde{b}} \\ &= \mathrm{tr}((\tilde{b} + k_P\sigma)_\times \tilde{R}M) + \frac{2}{k_I}\tilde{b}^\top \dot{\tilde{b}}\,.\end{aligned}$$

Then, using (A.3) one gets

$$\begin{aligned}\dot{\mathcal{L}} &= k_P \mathrm{tr}(\sigma_\times \mathbb{P}_a(\tilde{R}M)) + \mathrm{tr}(\tilde{b}_\times \mathbb{P}_a(\tilde{R}M)) + 2\tilde{b}^\top \sigma \\ &\quad - \frac{2k_b}{k_I}|\tilde{b}|^2 + \frac{2k_b}{k_I}\tilde{b}^\top(b - \mathrm{sat}_\Delta(b - \tilde{b}))\,.\end{aligned} \tag{A.5}$$

Using the following equalities

$$\sigma_\times = \sum_{i=1}^{n} \frac{k_i}{2}(\hat{R}^\top v_{0i} v_i^\top - v_i v_{0i}^\top \hat{R}) = \mathbb{P}_a(\tilde{R}M)\,,$$

$$2\tilde{b}^\top \sigma = -\mathrm{tr}(\tilde{b}_\times \sigma_\times) = -\mathrm{tr}(\tilde{b}_\times \mathbb{P}_a(\tilde{R}M))\,,$$

and the inequality $|b - \mathrm{sat}_\Delta(b - \tilde{b})| \leq |\tilde{b}|$, $\forall \tilde{b} \in \mathbb{R}^3$, provided that $\Delta \geq |b|$ by assumption (see, e.g., (Hua, 2009, Chap.2, Sec. 2.8.14) for the proof), one deduces from (A.5) that

$$\begin{aligned}\dot{\mathcal{L}} &\leq k_P \mathrm{tr}(\mathbb{P}_a(\tilde{R}M)^2) - \frac{2k_b}{k_I}|\tilde{b}|^2 + \frac{2k_b}{k_I}|\tilde{b}||b - \mathrm{sat}_\Delta(b - \tilde{b})| \\ &\leq -k_P \mathrm{tr}(\mathbb{P}_a(\tilde{R}M)^\top \mathbb{P}_a(\tilde{R}M)) = -k_P \|\mathbb{P}_a(\tilde{R}M))\|_{\mathrm{F}}^2\,,\end{aligned}$$

with $\|\cdot\|_{\mathrm{F}}$ denoting the Frobenius norm. From here, the proof proceeds exactly like in the proof of Theorem 5.1 in Mahony et al. (2008).