Analysis of non-linear attitude observers for time-varying reference measurements

Jochen Trumpf, *Member, IEEE,* Robert Mahony, *Senior Member, IEEE,*
Tarek Hamel, *Member, IEEE,* and Christian Lageman.

***Abstract*—This paper provides a comprehensive observability and stability analysis of a suite of non-linear attitude observers that have been developed over the last few years. The observers considered are based on vectorial measurements of an *a-priori* known reference direction. By treating the reference direction and the measurement in the same analysis framework, and allowing time-variation of either, we are able to define general persistency of excitation criteria that incorporate and generalize convergence criteria used in prior work. A key outcome is conditions that ensure almost global asymptotic and local exponential stability of attitude observers based on a single vector measurement as long as the excitation conditions are met on the reference and system trajectory. The approach generalizes stability results provided in prior work, based on rank conditions, that required at least two or more vector measurements.**

I. Introduction

Attitude or orientation estimation is a core technology in a wide range of robotics and aerospace applications; for example, satellite systems, aerial, terrestrial and submersible vehicle systems, robotic toys, tangible human machine interfaces, etc. The attitude estimation problem is long standing and well established and there are many different algorithms that have been investigated; for example, Extended Kalman Filters (EKF) [1], [2], the Multiplicative Extended Kalman Filter (MEKF) [3], [4], unscented filters [5], particle filters [6], [7], and non-linear observers [8], [9], [10], [11], with an excellent review of work up to around 2005 provided in [12]. The design of non-linear observers for attitude estimation can be traced back to the seminal work by Salcudean [8] who exploited the unit quaternion representation of rotations to derive an almost globally convergent observer for attitude based on full state measurements. The quaternion representation has lead to a suite of algorithms [13], [14], [15], [16], [17], [18] that have strong robustness and asymptotic stability properties, however, until recently all such algorithms were limited by the requirement of a full orientation state measurement [12]. This limitation was overcome around 2005 [19], [20], [16], [21], [22] and the resulting body of work has lead to a renewed interest in non-linear observers for attitude and attitude heading reference systems [20], [17], [18], [23] as well as pose estimation [24], [25]. The principles of observer design for invariant kinematics on Lie groups are becoming well established [9], [10], [11], [26] with work extending to systems on homogeneous spaces [10], [27] and dynamic systems [28], [29].

In this paper, we consider the well established structure of a vector measurement based non-linear observer for attitude estimation posed on the special orthogonal group $\mathrm{SO}(3)$ [19], [20], [21], [22], [30], and via the standard equivalence, observers posed on the unit quaternions [14], [15], [16], [17], [18]. The goal of this paper is to analyse the observability and stability properties of these observers in detail, and especially to consider the case of time-varying vector reference directions. For left-invariant kinematics, we recognize two different categories of vector measurements; the more common *complementary* measurements, where the measured quantity is in the body-fixed frame and the reference is in the inertial frame, and the less common *compatible* measurements, where the measured quantity is in the inertial frame and the reference is in the body-fixed frame. The principle difference between these measurements is in the observability properties of the system, and this difference itself is due primarily to the implicit assumption made in prior work that the reference direction is constant. By allowing the reference direction to be time-varying it is possible to provide a coherent and comprehensive observability and stability analysis for a whole suite of non-linear observers defined for vector measurements [19], [20], [21], [22], [30], [14], [15], [16], [17], [18]. We provide a characterization of *sufficient excitation* for the vector measurements that ensures weak observability of the system considered and demonstrate the existence of a universal input that distinguishes all initial states. We show that the condition of sufficient excitation is implied by the 'sufficiently non-collinear' condition or 'full rank' conditions used in earlier work [19], [20], [21], [22], [30], [14], [15], [16], [17], [18]. The condition we provide is strictly weaker, however, and is satisfied by a single vector measurement as long as the reference direction is sufficiently exciting. We go on to provide a comprehensive stability analysis for non-linear attitude observers based on a definition of *persistently exciting* inputs. We show that the observers are locally exponential stable and almost globally asymptotically stable as long as the inputs are persistently exciting and satisfy some minimal continuity properties. Once again this generalizes the 'non-collinear' and 'full-rank' measurement assumptions that were available in prior work (where only constant reference directions were considered) and demonstrates that the attitude observers can

J. Trumpf and R. Mahony are with the Research School of Engineering, Australian National University, Canberra ACT 2601, Australia. e-mail: Jochen.Trumpf@anu.edu.au or Robert.Mahony@anu.edu.au.

T. Hamel is with I3S UNSA-CNRS, Nice-Sophia Antipolis, France. e-mail: thamel@i3s.unice.fr.

C. Lageman is with the Mathematical Institute, University of Würzburg, 97074 Würzburg, Germany. e-mail: christian.lageman@mathematik.uni-wuerzburg.de.

be exponentially stable even for single vectorial measurements (where time-varying reference directions are present). Time-varying reference directions were first considered in the authors' conference paper [31] where a persistency of excitation condition on the reference directions was introduced to prove stability of the observer error. Following this work, further related results have also appeared [32]. No prior work has considered state observability for time-varying reference directions, nor have they considered discontinuous time-variation of reference directions. The present paper provides a comprehensive observability analysis and extends the stability analysis of prior work by providing a unifying analysis framework that deals with discontinuous time-variation.

The body of the paper consists of three sections. Following the introduction, Section II introduces the concepts of *complementary* and *compatible* measurements and their motivation, discusses the role of the reference directions and introduces an understanding of the time-variation that will be considered. The results in this paper are presented using the special orthogonal group $\mathrm{SO}(3)$ representation for attitude estimation, however, the equivalence to the quaternion representation is provided in an appendix to the paper. Section III introduces the concept of *sufficiently exciting* inputs and provides an analysis of the observability properties of attitude kinematics with vectorial measurements. We show that the sufficient excitation condition introduced is a generalization of the 'full rank' type conditions introduced in prior works (the proofs of the algebraic details of rank correspondence in excitation conditions are deferred to a second appendix). Section IV introduces a *persistency of excitation* condition on the measurements and reference directions and provides a comprehensive stability analysis for non-linear attitude observers with vectorial measurements on $\mathrm{SO}(3)$. A short conclusion is provided followed by the two appendices mentioned above.

II. Problem formulation

The orientation of a rigid body in space can be encoded as a rotation that represents the coordinates of the body-fixed frame $\{B\}$ with respect to the coordinates of the inertial frame $\{A\}$. This rotation may be represented by an element R of the Special Orthogonal group $\mathrm{SO}(3)$. The system considered in this paper is the kinematics of this orientation matrix given by

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

where the dot denotes the derivative with respect to time (denoted by t) and the angular velocity $\Omega \in \{B\}$ is assumed known, e.g. measured by onboard gyrometers. The notation $\Omega_\times$ denotes the skew-symmetric matrix

$$\Omega_\times = \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{pmatrix} \in \mathfrak{so}(3),$$

constructed from an angular velocity vector $\Omega = (\Omega_1, \Omega_2, \Omega_3)^\top \in \mathbb{R}^3$. The set of skew-symmetric matrices $\mathfrak{so}(3) = \{A \in \mathbb{R}^{3\times 3} \,|\, A = -A^\top\}$ is identified with the Lie-algebra associated with the matrix representation of $\mathrm{SO}(3)$. In the following we will always assume that $\Omega\colon \mathbb{R} \longrightarrow \mathbb{R}^3$ is continuous, and hence unique continuously differentiable solutions for system (1) exist for all initial values and all initial times, and these solutions are defined for all time since $\mathrm{SO}(3)$ is compact.

System (1) is a left-invariant description of the orientation kinematics in the sense that it is invariant to multiplications of R from the left by another constant orientation matrix S, i.e. Equation (1) implies that $\frac{\mathrm{d}}{\mathrm{d}t}(SR) = (SR)\Omega_\times$, the same form as Equation (1). In the language of systems on Lie groups the right hand side of Equation (1) is called a left-invariant vector field.

Let $S^2_{\{B\}}$ denote the unit sphere in the body-fixed-frame $\{B\}$. Consider measurements $a_i \in S^2_{\{B\}}$ associated with one or several reference directions $\mathring{a}_i \in S^2_{\{A\}}$, $i = 1, 2, \dots n_1$, expressed as vectors on the unit sphere in the inertial frame $\{A\}$. That is

$$a_i = R^\top \mathring{a}_i, \quad i = 1, \dots, n_1. \tag{2}$$

We term such measurements as *complementary* measurements to distinguish them from compatible measurements that we introduce later. In contrast to the left-invariance of system (1), the complementary output equations (2) are right invariant (due to the matrix transpose): Transforming the reference direction $\mathring{a}_i$ first by $R^\top$, corresponding to an orientation matrix R, and then by $S^\top$, corresponding to another orientation matrix S, is equivalent to transforming by $(RS)^\top$ where S multiplies R from the right, i.e. $S^\top(R^\top \mathring{a}_i) = (RS)^\top \mathring{a}_i$. In the language of Lie group actions and homogeneous spaces this is called a right action (see [27] for further details). The term "complementary" refers to the left-right correspondence of invariance properties of system and measurement. In an analogous fashion, left invariant measurements for a right invariant system would also be referred to as complementary. Complementary measurements are the most common measurements encountered in the attitude estimation problem for small scale robotic vehicles and satellites, the motivating problems for much of the prior work in non-linear observer design for attitude estimation [8], [12]. For example, the body-fixed frame measurements of the earth's magnetic or gravitational field measured by a strap-down Inertial Measurement Unit (IMU) are complementary measurements for the body-fixed frame representation of orientation kinematics (1). Note that a complementary measurement is made relative to the reference $\mathring{a}_i \in \{A\}$ and is only useful for estimating attitude if the reference direction is known *a-priori*. For example, the inertial directions of the magnetic or gravitational fields must be known *a-priori* at the location in which a vehicle is operating in order to utilize accelerometer and magnetometer measurements from a strap down IMU for attitude estimation.

Consider additional measurements that are associated with one or several known body-fixed reference directions $\mathring{b}_j \in S^2_{\{B\}}$, $j = 1, 2, \dots, n_2$. The measurements $b_j \in S^2_{\{A\}}$ are now made in the inertial frame, i.e.

$$b_j = R\mathring{b}_j, \quad j = 1, \dots, n_2. \tag{3}$$

We term these as *compatible* measurements. Equation (3) has the same (left) invariance as the system equation (1).

That is, $S(R\mathring{b}_j) = (SR)\mathring{b}_j$ where S now multiplies R from the left on the right hand side of this equality. In the language of Lie group actions and homogeneous spaces this is called a left action. The term "compatible" refers to the left-left correspondence of invariance properties of system and measurement, and analogously, right invariant measurements for a right invariant system would also be referred to as compatible. Compatible measurements are less common than complementary measurement in robotic vehicle applications. An example of such a measurement is the differential vector derived from comparing two inertial position measurements obtained from separate GPS units attached to the robotic vehicle, with a known base-line in the body-fixed frame $\{B\}$ [30].

Both complementary and compatible measurements can only be incorporated in an estimation algorithm if the reference direction is known *a-priori* or separately measured in its appropriate frame of reference. Indeed, by re-arranging Equation (2) (resp. Equation (3)) it is easy to see that a compatible (resp. complementary) measurement is really just a complementary (resp. compatible) measurement, where now the measurement is treated as the reference and the reference is treated as the measurement. However, this simple correspondence conceals the complexity in the inherent observability properties of the underlying systems - the principal topic of this paper. For example, in early applications considered the known reference directions $\mathring{a}_i \in S^2_{\{A\}}$ or $\mathring{b}_i \in S^2_{\{B\}}$ are modelled as constant known vectors. If either measurement is viewed with respect to the reverse invariance then the reference becomes time-varying while the measurement becomes constant, violating the assumptions in the theorems presented in many of the earlier papers. By treating both the measurement and the reference as time-varying inputs to the observer we avoid the conceptual trap associated with treating the reference as a constant system parameter and can apply a general framework to the analysis of the observability and stability problem. Thus, we consider time-varying vector reference directions

$$\mathring{a}_i \colon \mathbb{R} \longrightarrow S^2_{\{A\}}, \quad \mathring{b}_i \colon \mathbb{R} \longrightarrow S^2_{\{B\}}.$$

In many applications the reference directions will vary continuously, and even smoothly, with time. However, a key contribution of this paper is to deal with some of the situations where the reference direction may be only piecewise continuous, as would be encountered for example if the active reference was switching between different candidate references depending on some sampling criteria. As such, we will be careful to explicitly state any continuity or integrability assumptions on the time-variation of the references in all results that follow. Note, however, that the modelling framework used in this paper does not allow us to directly incorporate *intermittent* measurements, where the *number* of reference signals varies with time due to unavailability of measurements. We assume the numbers n_1 and n_2 of measurements to be constant, and at least one of them must be nonzero.

We will study the combined system (1), (2) and (3) where we also allow for the cases $n_1 = 0$ or $n_2 = 0$ to indicate the absence of body-fixed or inertial measurements, respectively. To simplify notation we will use the convention that sets and sums with empty index set (such as $i = 1, \dots, n$ where $n = 0$) are treated as if they were not present in the respective expressions.

The results presented in this paper will be based on the standard matrix representation of $\mathrm{SO}(3)$. The unit quaternion representation of rotations, however, is commonly used for the realization of attitude estimation algorithms since it offers considerable efficiency in code implementation. Consequently, much of the work in this domain is presented directly in terms of the quaternion representations of the observer equations. Due to the homomorphic mapping between the group of unit quaternions, $\mathbb{Q}$, and $\mathrm{SO}(3)$ it is always possible to represent algorithms on $\mathrm{SO}(3)$ as an algorithm on $\mathbb{Q}$ but not vice versa (see e.g. [33]). As such the results presented in this paper are directly applicable to observers in the unit quaternion representation via the standard equivalence. The details of the equivalence are provided in the appendix along with the quaternion form of the principle observer equation.

III. Observability results

We start by stating the precise definitions of some well known concepts regarding observability of non-linear systems.

Definition 3.1: Two initial states of system (1), (2) and (3) are said to be *distinguished by* an admissible input if the respective outputs, resulting from applying the input while starting the system from the two initial states, are different at the initial time or at some time instance in the future. Equivalently, we say that the admissible input *distinguishes* the two initial states in this case. Two initial states of system (1), (2) and (3) are called *indistinguishable* if they are not distinguished by any admissible input. A system is called *strongly observable* if all pairs of distinct initial states are distinguished by all admissible inputs. It is called *weakly observable* if every pair of distinct initial states is distinguished by at least one admissible input.

The need to differentiate between strong observability, where every input will distinguish between distinct initial states, and weak observability, where this might not be the case, is well established in the non-linear systems literature [34]. It is maybe less obvious which of the various signals in system (1), (2) and (3) should be regarded as inputs; specifically, one could think of the reference directions $\mathring{a}_i$ and $\mathring{b}_j$ either as (potentially time-varying) parameters of the system or one could think of them as inputs. We opt for the latter point of view, i.e. we treat the reference directions as system inputs that just happen to not enter the dynamical part (1) but provide direct feedthrough to the output via (2) and (3).

We can now state our main observability result.

Theorem 3.2: System (1), (2) and (3) is weakly observable.

In order to prove this theorem, we need to provide system inputs $(\Omega, \{\mathring{a}_i\}, \{\mathring{b}_j\})$ that distinguish distinct initial states. It turns out that we can find *universal inputs* [34], where in our context this means inputs that distinguish *all* pairs of distinct initial states. We will express the criterion for these universal inputs in terms of a notion of excitation that is defined as follows.

Definition 3.3: A collection of locally integrable directions $v_i\colon \mathbb{R} \longrightarrow S^2$, $i = 1, \dots, n$ is called *sufficiently exciting* if there exists a time $T > 0$ such that

$$\lambda_2 \left(\int_0^T \sum_{i=1}^n v_i(\tau) v_i(\tau)^\top \mathrm{d}\tau \right) > 0, \tag{4}$$

where $\lambda_2(S)$ denotes the second largest eigenvalue of a symmetric matrix $S \in \mathbb{R}^{3\times 3}$.

The appearance of the second largest eigenvalue – rather than the smallest eigenvalue – in this definition is due to the geometry of the (two-dimensional) sphere S^2. This is better visible in the following characterization of sufficient excitation in terms of the projectors $(I - v_i v_i^\top)$ onto the tangent spaces to S^2 at v_i. Here, $I \in \mathbb{R}^{3\times 3}$ denotes the identity matrix.

Lemma 3.4: A collection of locally integrable directions $v_i\colon \mathbb{R} \longrightarrow S^2$, $i = 1, \dots, n$ is sufficiently exciting if and only if there exists a time $T > 0$ such that

$$\lambda_{\min} \left(\int_0^T \sum_{i=1}^n (I - v_i(\tau) v_i(\tau)^\top)\ \mathrm{d}\tau \right) > 0, \tag{5}$$

where $\lambda_{\min}(S)$ denotes the smallest eigenvalue of a symmetric matrix $S \in \mathbb{R}^{3\times 3}$.

The proof of this lemma is provided in the Appendix. We can now provide a sufficient criterion for distinguishing inputs as follows.

Proposition 3.5: An input $(\Omega, \{\mathring{a}_i\}, \{\mathring{b}_j\})$ distinguishes between the two initial states $R_1(0) \neq R_2(0)$ if the collection $\{\mathring{a}_i\} \cup \{b_{j,1}\} \cup \{b_{j,2}\}$ is sufficiently exciting. Here $b_{j,k}$, $k = 1, 2$ denotes the output (3) generated by running the system (1), (2) and (3) with the two possible different initial states $R_k(0)$, $k = 1, 2$.

Proof: Let $R_1(0), R_2(0) \in \mathrm{SO}(3)$, $R_1(0) \neq R_2(0)$ be two initial states and denote the corresponding outputs of system (1), (2) and (3) by $a_{i,1}$, $a_{i,2}$, $b_{j,1}$ and $b_{j,2}$, respectively. To arrive at a contradiction, assume that $R_1(0)$ and $R_2(0)$ are indistinguishable, i.e. assume that $a_{i,1}(t) = a_{i,2}(t)$ and $b_{j,1}(t) = b_{j,2}(t)$ for all i, j and $t \geq 0$. Denote the canonical right invariant error [26] $E_r(R_2, R_1) = R_2 R_1^\top$ by E_r. It follows that

$$\begin{aligned} E_r \mathring{a}_i &= R_2 R_1^\top \mathring{a}_i = R_2 a_{i,1} = R_2 a_{i,2} = \mathring{a}_i \\ E_r b_{j,1} &= R_2 R_1^\top b_{j,1} = R_2 \mathring{b}_j = b_{j,2} = b_{j,1} \quad \text{and} \\ E_r b_{j,2} &= R_2 R_1^\top b_{j,2} = R_2 R_1^\top b_{j,1} = R_2 \mathring{b}_j = b_{j,2} \end{aligned}$$

for all i and j (and $t \geq 0$). Moreover,

$$\begin{aligned} \dot{E}_r &= \dot{R}_2 R_1^\top + R_2 \dot{R}_1^\top = R_2 \Omega_\times R_1^\top + R_2 (R_1 \Omega_\times)^\top \\ &= R_2 \Omega_\times R_1^\top - R_2 \Omega_\times R_1^\top = 0 \end{aligned}$$

because $\Omega_\times$ is skew-symmetric, i.e. E_r is constant. It follows that for all $T > 0$

$$E_r Q_0 = Q_0$$

where

$$Q_0 := \int_0^T \left(\sum_{i=1}^{n_1} \mathring{a}_i(\tau) \mathring{a}_i(\tau)^\top + \sum_{j=1}^{n_2} b_{j,1}(\tau) b_{j,1}(\tau)^\top + \sum_{j=1}^{n_2} b_{j,2}(\tau) b_{j,2}(\tau)^\top \right) \mathrm{d}\tau.$$

Let $\{\mathring{a}_i\} \cup \{b_{j,1}\} \cup \{b_{j,2}\}$ be sufficiently exciting and let $Q_0 = U \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3) U^\top$ be an orthonormal diagonalization of Q_0 with $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Then $\lambda_2 > 0$ for some $T > 0$. But then λ_1 times the first column of U and λ_2 times the second column of U are two mutually orthogonal eigenvectors of E_r, both corresponding to the eigenvalue 1. However, E_r is a rotation matrix and hence its spectrum is of the form

$$(1,\ \cos(\theta) + i \sin(\theta),\ \cos(\theta) - i \sin(\theta)),$$

where θ denotes the rotation angle of E_r. It follows that $E_r = I$ and hence that $R_1(0) = R_2(0)$, a contradiction. This means that the input distinguishes $R_1(0)$ and $R_2(0)$ as required. ■

The previous proposition only provides an indirect criterion for distinguishing inputs in that the condition is phrased in terms of the resulting output signals. The following proposition provides a sufficient criterion for an input to be universal, i.e. to distinguish all pairs of distinct initial states, that is phrased entirely in terms of input signals.

Proposition 3.6: An input $(\Omega, \{\mathring{a}_i\}, \{\mathring{b}_j\})$ where the $\mathring{a}_i$, $i = 1, \dots, n_1$ are locally integrable and where the $\mathring{b}_j$, $j = 1, \dots, n_2$ are piecewise absolutely continuous distinguishes between all pairs of distinct initial states if there exists a time $T > 0$ such that

$$\lambda_2 \left(\int_0^T \sum_{i=1}^{n_1} \mathring{a}_i(\tau) \mathring{a}_i(\tau)^\top \mathrm{d}\tau \right) + \left\| \int_0^T \sum_{j=1}^{n_2} \left(\Omega_\times(\tau) \mathring{b}_j(\tau) + \frac{\mathrm{d}}{\mathrm{d}\tau} \mathring{b}_j(\tau) \right) \mathrm{d}\tau \right\| > 0. \tag{6}$$

Note that the derivative of $\mathring{b}_j$ exists almost everywhere.

Proof: Assume that condition (6) holds for some $T > 0$. Let $R_1(0), R_2(0) \in \mathrm{SO}(3)$, $R_1(0) \neq R_2(0)$ be two initial states and denote the corresponding outputs of system (1), (2) and (3) by $a_{i,1}$, $a_{i,2}$, $b_{j,1}$ and $b_{j,2}$, respectively. Assume further, to arrive at a contradiction, that the collection $\{\mathring{a}_i\} \cup \{b_{j,1}\} \cup \{b_{j,2}\}$ is not sufficiently exciting. By Lemma 3.4 this means that for all $T > 0$ there exists a vector $y \in S^2$ such that

$$\int_0^T y^\top \left(\sum_{i=1}^{n_1} (I - \mathring{a}_i(\tau) \mathring{a}_i(\tau)^\top) + \sum_{j=1}^{n_2} (I - b_{j,1}(\tau) b_{j,1}(\tau)^\top) + \sum_{j=1}^{n_2} (I - b_{j,2}(\tau) b_{j,2}(\tau)^\top) \right) y\ \mathrm{d}\tau \leq 0.$$

Since all the summands in the integrand are non-negative, this implies that

$$\begin{aligned} y^\top (I - \mathring{a}_i(\tau) \mathring{a}_i(\tau)^\top) y &= y^\top (I - b_{j,1}(\tau) b_{j,1}(\tau)^\top) y \\ &= y^\top (I - b_{j,2}(\tau) b_{j,2}(\tau)^\top) y = 0 \end{aligned}$$

for all i, all j and almost all $\tau \in [0, T]$. But then

$$\mathring{a}_i(\tau) = \pm y, \quad b_{j,1}(\tau) = \pm y \quad \text{and} \quad b_{j,2}(\tau) = \pm y$$

for all i, all j and almost all $\tau \in [0, T]$. If $n_1 \neq 0$ the first of these identities implies that

$$\operatorname{rank}\left(\int_0^T \sum_{i=1}^{n_1} \mathring{a}_i(\tau)\mathring{a}_i(\tau)^\top \mathrm{d}\tau\right) = \operatorname{rank}(yy^\top) = 1$$

and hence that

$$\lambda_2\left(\int_0^T \sum_{i=1}^{n_1} \mathring{a}_i(\tau)\mathring{a}_i(\tau)^\top \mathrm{d}\tau\right) = 0. \tag{7}$$

In the second and third identities the sign of y is piecewise constant since the $\mathring{b}_j$ are piecewise absolutely continuous and hence the $b_{j,k} = R_k\mathring{b}_j$ are piecewise continuous. But then the $b_{j,k}$ are even piecewise differentiable since they are piecewise constant. It follows that

$$\frac{\mathrm{d}}{\mathrm{d}\tau}b_{j,1}(\tau) = \frac{\mathrm{d}}{\mathrm{d}\tau}b_{j,2}(\tau) = 0$$

for all j and almost all $\tau \in [0, T]$. This in turn implies

$$R_k(\tau)\left(\Omega_\times(\tau)\mathring{b}_j(\tau) + \frac{\mathrm{d}}{\mathrm{d}\tau}\mathring{b}_j(\tau)\right) = 0$$

for all j, for $k = 1, 2$ and almost all $\tau \in [0, T]$, i.e. wherever the derivatives of $b_{j,k}$ and $\mathring{b}_j$ exist. Since $R_k(\tau) \in \mathrm{SO}(3)$ is invertible, we get

$$\int_0^T \sum_{j=1}^{n_2}\left(\Omega_\times(\tau)\mathring{b}_j(\tau) + \frac{\mathrm{d}}{\mathrm{d}\tau}\mathring{b}_j(\tau)\right)\mathrm{d}\tau = 0 \tag{8}$$

for the case $n_2 \neq 0$. Equations (8) and (7) yield a contradiction to Condition (6). Note that the contradiction occurs independent of whether n_1 or n_2 are zero. Applying Proposition 3.5 completes the proof. ■

Now all that remains to be done for the proof of Theorem 3.2 is to show that universal inputs exist, i.e. to show that Condition (6) can be fulfilled. By inspection, this is clearly the case but it is still instructive to discuss some of the special cases.

In the case where there is only a single *constant* reference direction in the inertial frame and a single measurement in the body-fixed frame, (that is $n_1 = 1$ and $n_2 = 0$ and $\mathring{a}_1 = \text{const.}$), it is known that the system is unobservable (see e.g. [10], [9]). In this case clearly the first integral in Condition (6) has rank 1 and the second integral is not present, i.e. the condition does not hold. As soon as the single reference direction takes on two different values over two time periods of non-zero measure, (for example $\mathring{a}_1(t) \in \{\mathring{a}_{1,1}, \mathring{a}_{1,2}\}$, $\mathring{a}_{1,1} \neq \mathring{a}_{1,2}$, with some persistent switching criterion), Condition (6) holds for some $T > 0$ and distinct initial states are distinguished.

For the case of more than one *constant* inertial reference direction the known criterion for observability is that the reference directions are *non-collinear* [10], [9]. This notion can easily be generalized to the time-varying setting as follows.

Definition 3.7: A collection of locally integrable directions $v_i \colon \mathbb{R} \longrightarrow S^2$, $i = 1, \ldots, n$ is called *sufficiently non-collinear* if there exists a time $T > 0$ such that

$$\int_0^T \sum_{i \neq j} |v_i(\tau) \times v_j(\tau)| \; \mathrm{d}\tau > 0. \tag{9}$$

The proof of the following lemma is provided in the Appendix.

Lemma 3.8: A collection of sufficiently non-collinear directions is sufficiently exciting.

It follows that a collection of (more than one) sufficiently non-collinear inertial reference directions distinguishes distinct initial states.

Finally, in the case where $n_1 = 0$, i.e. where all the measurements are made in the inertial frame of body-fixed reference directions, an inspection of the second integral in Condition (6) reveals that already a *single constant* reference direction $\mathring{b}$ can distinguish distinct initial states [31]. This is the case when

$$\left\|\int_0^T \Omega_\times(\tau)\mathring{b} \; \mathrm{d}\tau\right\| > 0$$

for some $T > 0$. The geometrical interpretation of this condition is that on average there is sufficient instantaneous rotational movement of the body-fixed frame around axes other than $\mathring{b}$.

Remark 3.9: The following simple example shows that Condition (6) is not necessary for distinguishability. Consider the case $n_1 = 0$ and $n_2 = 2$ and denote the i-th standard basis vector of $\mathbb{R}^3$ by e_i. Then the constant input $(\Omega, \mathring{b}_1, \mathring{b}_2) = (0, e_1, e_2)$ distinguishes between all pairs of distinct initial states since a rotation matrix is uniquely specified by two of its columns and the system trajectories are stationary. Condition (6) is not fulfilled in this example since the first integral is not present and the second integral evaluates to zero. Note that this example *does* satisfy the sufficient condition provided by Proposition 3.5.

IV. Observer analysis

Non-linear attitude observers have been proposed in prior literature for both the cases of constant reference directions [10], [9] and body-fixed reference directions [30], respectively. In this section we will combine these observers into a single observer for system (1), (2) and (3) and discuss its asymptotic properties in the fully general, time-varying setting. We recover all the known observer error convergence results as special cases of our general convergence result below which also slightly strengthens some of these previous results.

The combined observer takes the form

$$\dot{\hat{R}} = \hat{R}\left(\Omega + \sum_{i=1}^{n_1} k_i(a_i \times \hat{a}_i) + \hat{R}^\top \sum_{j=1}^{n_2} l_j(\hat{b}_j \times b_j)\right)_\times \tag{10}$$

where $\hat{a}_i := \hat{R}^\top \mathring{a}_i$, $\hat{b}_j := \hat{R}\mathring{b}_j$, $k_i > 0$ and $l_j > 0$ for $i = 1, \ldots, n_1$ and $j = 1, \ldots, n_2$. The k_i and l_j are arbitrary positive observer gains.

We will study the observer error in terms of the canonical right invariant error [26] $E_r := \hat{R}R^\top$, where no error corresponds to $E_r = I$. Clearly, some form of observability (or detectability) condition is necessary for observer error convergence. It turns out that the observability conditions provided in the previous section are also sufficient for observer error convergence if we replace *sufficient* excitation by *persistent* excitation.

Definition 4.1: A collection of locally integrable directions $v_i \colon \mathbb{R} \longrightarrow S^2$, $i = 1, \ldots, n$ is called *persistently exciting* if there exist $\delta > 0$ and $c > 0$ such that for all $t \geq 0$

$$\lambda_2 \left(\int_t^{t+\delta} \sum_{i=1}^n v_i(\tau) v_i(\tau)^\top \mathrm{d}\tau \right) > c. \tag{11}$$

As for the case of sufficient excitation (cf. Lemma 3.4), an alternative characterization in terms of projectors can be provided as follows. For convenience we also include arbitrary positive gains in the statement.

Lemma 4.2: Let $q_i > 0$, $i = 1, \ldots, n$. A collection of locally integrable directions $v_i \colon \mathbb{R} \longrightarrow S^2$, $i = 1, \ldots, n$ is persistently exciting if and only if there exist $\delta > 0$ and $c > 0$ such that for all $t \geq 0$

$$\lambda_{\min} \left(\int_t^{t+\delta} \sum_{i=1}^n q_i \left(I - v_i(\tau) v_i(\tau)^\top \right) \, \mathrm{d}\tau \right) > c. \tag{12}$$

A proof of this lemma is provided in the Appendix. We can now state our main observer error convergence result.

Theorem 4.3: Consider system (1), (2) and (3) and the observer (10). Let the angular velocity Ω be bounded and uniformly continuous, let the reference directions $\mathring{a}_i$ be uniformly continuous and let the reference directions $\mathring{b}_j$ be continuously differentiable with bounded and uniformly continuous derivatives. Assume that there exist $\delta > 0$ and $c > 0$ such that for all $t \geq 0$

$$\lambda_2 \left(\int_t^{t+\delta} \sum_{i=1}^{n_1} \mathring{a}_i(\tau) \mathring{a}_i(\tau)^\top \mathrm{d}\tau \right) + \left\| \int_t^{t+\delta} \sum_{j=1}^{n_2} \left(\Omega_\times(\tau) \mathring{b}_j(\tau) + \frac{\mathrm{d}}{\mathrm{d}\tau} \mathring{b}_j(\tau) \right) \mathrm{d}\tau \right\| > c. \tag{13}$$

Then the equilibrium $E_r = I$ is uniformly locally exponentially stable (ULES) and almost globally asymptotically stable (AGAS) with a basin of attraction including at least all initial conditions such that $\mathrm{tr}(I - E_r(0)) < 4$.

Remark 4.4: Note that $0 \leq \mathrm{tr}(I - E_r) \leq 4$ for $E_r \in \mathrm{SO}(3)$ and that $\{E_r \mid \mathrm{tr}(I - E_r) = 4\}$ has (Haar) measure zero in $\mathrm{SO}(3)$. Hence it is justified to speak of almost global convergence in the above theorem.

The proof of Theorem 4.3 starts from the observation that the observer (10) uses both the reference directions $\mathring{a}_i$ resp. $\mathring{b}_j$ and the system outputs a_i resp. b_j as its inputs while the way they appear in the observer equation is swapped between the two types of measurements. Specifically, if we denote

$$(q_i, \mathring{v}_i, v_i) := \begin{cases} (k_i, \mathring{a}_i, a_i), & i = 1, \ldots, n_1 \\ (l_{i-n_1}, b_{i-n_1}, \mathring{b}_{i-n_1}), & i = n_1 + 1, \ldots, n_1 + n_2 \end{cases}$$

then the system (1), (2) and (3) takes the form

$$\begin{aligned} \dot{R} &= R\Omega_\times, \\ v_i &= R^\top \mathring{v}_i, \quad i = 1, \ldots, n \end{aligned} \tag{14}$$

where $n := n_1 + n_2$, and the observer (10) takes the form

$$\dot{\hat{R}} = \hat{R} \left(\Omega + \sum_{i=1}^n q_i (v_i \times \hat{v}_i) \right)_\times, \tag{15}$$

where $\hat{v}_i = \hat{R}^\top \mathring{v}_i$. Implicitly, we partly swap the roles of reference directions and measurements in doing this, however, both of these are available to the observer and their labelling is secondary to the analysis.

Remark 4.5: The transformation above is the deeper reason why we consider the reference directions as part of the system inputs rather than as part of the system parameters. It captures the essence of how they are being treated in this process.

Adepts of the behavioral approach will of course notice that this is a very fine example of assigning different input/output splittings. Equations (1), (2) and (3) suggest $(\Omega, \mathring{a}_i, \mathring{b}_j)$ as input and (a_i, b_j) (and the state R) as output, while Equation (14) suggests $(\Omega, \mathring{a}_i, b_j)$ as input and $(a_i, \mathring{b}_j)$ (and the state R) as output. From our interpretation of these variables it is clear that the first of these choices is more natural, while the second choice makes it easier to analyze the observer error. From the behavioral point of view it doesn't matter too much which choice we make, the system itself is unchanged by that; only its representation has changed.

Analogous to the previous section on observability, we can now proceed in two steps as follows. Firstly, use a persistency of excitation condition on the collection $\{\mathring{v}_i\} = \{\mathring{a}_i\} \cup \{b_j\}$ to conclude observer error convergence and then, secondly, show that under additional regularity conditions this persistency of excitation condition is implied by Condition (13). The strong regularity conditions in Theorem 4.3 are owing to this latter step. If we are content working with the collection $\{\mathring{v}_i\}$ then the much weaker conditions of the next proposition will suffice. It uses an extension of Barbalat's Lemma to the piecewise setting [35].

Proposition 4.6: Consider the system (14) and the observer (15). Let the reference directions $\mathring{v}_i$ be piecewise continuous and let $\{\mathring{v}_i\}$ be persistently exciting. Then the equilibrium $E_r = I$ is uniformly locally exponentially stable (ULES). If, furthermore, the reference directions $\mathring{v}_i$ are piecewise uniformly continuous then the equilibrium $E_r = I$ is almost globally asymptotically stable (AGAS) with a basin of attraction including at least all initial conditions such that $\mathrm{tr}(I - E_r(0)) < 4$.

Proof: We compute

$$\begin{aligned} \dot{E}_r &= \sum_{i=1}^n q_i \hat{R} (v_i \times \hat{v}_i)_\times R^\top = \sum_{i=1}^n q_i \hat{R} ((R^\top \mathring{v}_i)_\times \hat{R}^\top \mathring{v}_i)_\times R^\top \\ &= \sum_{i=1}^n q_i \hat{R} R^\top ((\mathring{v}_i)_\times R \hat{R}^\top \mathring{v}_i)_\times = \sum_{i=1}^n q_i E_r ((\mathring{v}_i)_\times E_r^\top \mathring{v}_i)_\times . \end{aligned}$$

Next, we wish to linearize this system expressed in local coordinates around $E_r = I$. To this extent we approximate the local coordinates given by the exponential map $Z \mapsto \exp(Z)$

for $Z \in \mathfrak{so}(3)$ by $E_r \approx I+Z$, substitute in the above equation and neglect higher order terms in Z. This process yields

$$\begin{aligned}\dot{Z} &= \sum_{i=1}^{n} q_i \left[(I+Z)((\mathring{v}_i)_\times \hat{v}_i)_\times - ((\mathring{v}_i)_\times Z \hat{v}_i)_\times\right] \\ &= -\sum_{i=1}^{n} q_i ((\mathring{v}_i)_\times Z \mathring{v}_i)_\times = \sum_{i=1}^{n} q_i ((\mathring{v}_i)_\times (\mathring{v}_i)_\times \operatorname{vex}(Z))_\times\end{aligned}$$

where the operator $\operatorname{vex}\colon \mathfrak{so}(3) \longrightarrow \mathbb{R}^3$ denotes the inverse of the $(\cdot)_\times$ operator. Hence the linearization in local coordinates reads

$$\operatorname{vex}(\dot{Z}) = -\sum_{i=1}^{n} q_i (I - \mathring{v}_i \mathring{v}_i^\top) \operatorname{vex}(Z).$$

Defining the symmetric positive semi-definite time-varying matrix

$$P(t) := \sum_{i=1}^{n} q_i (I - \mathring{v}_i(t) \mathring{v}_i(t)^\top)$$

this system now takes the form studied by Morgan and Narendra in [36]. By our persistency of excitation assumption, and appealing to Lemma 4.2, we know that there exist $\delta > 0$ and $c > 0$ such that for all $t \geq 0$ and all $y \in S^2$

$$\int_t^{t+\delta} y^\top P(\tau) y \ \mathrm{d}\tau > c.$$

Hence, given an initial time $t_0 \geq 0$ and an arbitrary time $t \geq t_0$ we can "chop up" the time interval $[t_0, t]$ into pieces of length δ and a remainder, resulting in

$$\begin{aligned}\int_{t_0}^{t} y^\top P(\tau) y \ \mathrm{d}\tau &\geq \frac{c}{\delta} \lfloor \frac{t-t_0}{\delta} \rfloor \delta + \int_{t_0 + \lfloor \frac{t-t_0}{\delta} \rfloor \delta}^{t} y^\top P(\tau) y \ \mathrm{d}\tau \\ &\geq \frac{c}{\delta} \cdot \left[(t-t_0) - \left((t-t_0) - \lfloor \frac{t-t_0}{\delta} \rfloor \delta\right)\right] \\ &\geq \frac{c}{\delta} \cdot [(t-t_0) - \delta] = \frac{c}{\delta}(t-t_0) - c,\end{aligned}$$

where $\lfloor x \rfloor$ for $x \in \mathbb{R}$ denotes the largest integer smaller than or equal to x. This inequality is precisely the condition in [36, Theorem 1, Part 2] and it follows that our linearized system is uniformly asymptotically and hence [37, Theorem III.2.1] uniformly exponentially stable. By Theorem VII.1.3 in [38] it follows that the equlibrium $E_r = I$ of the nonlinear system is uniformly locally exponentially stable (note that the continuity assumption made in [38, Theorem VII.1.3] is actually not used in its proof).

Now consider the following candidate Lyapunov function

$$V = \operatorname{tr}(I - E_r).$$

As in the beginning of this proof we compute

$$\begin{aligned}\dot{V} &= -\operatorname{tr}\sum_{i=1}^{n} q_i \hat{R}(v_i \times \hat{v}_i)_\times R^\top = -\operatorname{tr}\sum_{i=1}^{n} q_i \hat{R}(\hat{v}_i v_i^\top - v_i \hat{v}_i^\top) R^\top \\ &= -\operatorname{tr}\sum_{i=1}^{n} q_i (\mathring{v}_i \mathring{v}_i^\top - E_r \mathring{v}_i \mathring{v}_i^\top E_r) = -\sum_{i=1}^{n} q_i (1 - \mathring{v}_i^\top E_r^2 \mathring{v}_i) \\ &\leq 0.\end{aligned}$$

Since $\dot{V}$ is negative semi-definite, the candidate Lyapunov function is bounded with $0 \leq V(t) \leq V(0)$ for $t \geq 0$. Additionally, $\dot{V}$ is piecewise uniformly continuous since the $\mathring{v}_i$ are bounded and piecewise uniformly continuous and E_r is bounded and absolutely continuous, hence also uniformly continuous. An application of Barbalat's lemma [35, Lemma 9] yields $\mathring{v}_i^\top E_r^2 \mathring{v}_i \to 1$ for all $i = 1, \ldots, n$. This implies that $E_r \mathring{v}_i \to \mathring{v}_i$ since the alternative possibilities, $E_r \mathring{v}_i \to U \operatorname{diag}(1, -1, -1) U^\top \mathring{v}_i$, $U \in \mathrm{SO}(3)$, are excluded by the constraint $\operatorname{tr}(I - E_r(0)) < 4$ on the initial condition and the fact that $V(t) \leq V(0)$ for $t \geq 0$. It follows that $\hat{v}_i \to v_i$ for all $i = 1, \ldots, n$ and hence $\dot{E}_r \to 0$. It remains to show that E_r converges to I. We already know that

$$E_r \mathring{v}_i - \mathring{v}_i \to 0.$$

Multiplying by $\mathring{v}_i^\top$, summing over i and integrating the result over a period of time of length δ we get

$$\int_t^{t+\delta} (E_r(\tau) - I) \sum_{i=1}^{n} \mathring{v}_i(\tau) \mathring{v}_i(\tau)^\top \mathrm{d}\tau \to 0.$$

Integrating by parts yields

$$\begin{aligned}&\int_t^{t+\delta} (E_r(\tau) - I) \sum_{i=1}^{n} \mathring{v}_i(\tau) \mathring{v}_i(\tau)^\top \mathrm{d}\tau = \\ &\qquad \left[(E_r(\tau) - I) \int_t^{\tau} \sum_{i=1}^{n} \mathring{v}_i(s) \mathring{v}_i(s)^\top \mathrm{d}s\right]_t^{t+\delta} \\ &\qquad - \int_t^{t+\delta} \dot{E}_r(\tau) \left(\int_t^{\tau} \sum_{i=1}^{n} \mathring{v}_i(s) \mathring{v}_i(s)^\top \mathrm{d}s\right) \mathrm{d}\tau.\end{aligned}$$

Observe that the term in brackets vanishes at $\tau = t$ and that the inner integral on the last line of this expression is bounded since its integrand is bounded (the reference directions have length 1) and the length of the integration interval is bounded by δ. Using the fact that $\dot{E}_r \to 0$ it follows that

$$(E_r(t+\delta) - I) \int_t^{t+\delta} \sum_{i=1}^{n} \mathring{v}_i(\tau) \mathring{v}_i(\tau)^\top \mathrm{d}\tau \to 0$$

and hence that

$$\bar{E}(t) U(t) \left(\int_t^{t+\delta} \sum_{i=1}^{n} \mathring{v}_i(\tau) \mathring{v}_i(\tau)^\top \mathrm{d}\tau\right) U(t)^\top \to 0,$$

where $\bar{E}(t) := U(t)(E_r(t+\delta) - I)U(t)^\top$ and $U(t)$ is a (time-varying) orthogonal transformation that diagonalises the integral to $\operatorname{diag}(\lambda_1(t), \lambda_2(t), \lambda_3(t))$ with $\lambda_1(t) \geq \lambda_2(t) \geq \lambda_3(t) \geq 0$ for all $t \geq 0$. It follows that $\bar{E}\lambda_1 e_1 \to 0$ and $\bar{E}\lambda_2 e_2 \to 0$, where $e_1 \in \mathbb{R}^3$ and $e_2 \in \mathbb{R}^3$ denote the first and second standard basis vector, respectively. Our persistency of excitation assumption implies that $\lambda_1(t) > c > 0$ and $\lambda_2(t) > c > 0$ for all $t \geq 0$ and hence it follows that $\bar{E} e_1 \to 0$ and $\bar{E} e_2 \to 0$. Since E_r is a rotation matrix, the spectrum of $\bar{E}$ is of the form

$$(0,\ \cos(\theta) - 1 + i\sin(\theta),\ \cos(\theta) - 1 - i\sin(\theta)),$$

where θ denotes the rotation angle of E_r. It follows that $\bar{E} \to 0$ or, equivalently, that $E_r \to I$. This completes the proof. ■

The following proposition now provides a condition for $\{\mathring{v}_i\}$ to be persistently exciting that is framed in terms of conditions on the reference directions $\mathring{a}_i$ and $\mathring{b}_j$ as well as the angular velocity Ω. We have not been able to provide such a condition in the piecewise setting, since the proof technique using the Arzelà-Ascoli theorem seems not to generalize without making unrealistic assumptions on switching patterns.

Proposition 4.7: Let Ω be bounded and uniformly continuous, let the $\mathring{a}_i$ be uniformly continuous and let the $\mathring{b}_j$ be continuously differentiable with bounded and uniformly continuous derivatives. Assume that there exist $\delta > 0$ and $c > 0$ such that Condition (13) holds for all $t \geq 0$. Then $\{\mathring{a}_i\} \cup \{b_j\}$ is persistently exciting.

Proof: Assume that condition (13) holds for some $\delta > 0$, some $c > 0$ and all $t \geq 0$. Assume further, to arrive at a contradiction, that $\{\mathring{v}_i\} = \{\mathring{a}_i\} \cup \{b_i\}$ is not persistently exciting. By Lemma 4.2 this means that for all $\tilde{\delta} > 0$ and all $\tilde{c} > 0$ there exists a time $t \geq 0$ and a constant vector $y \in S^2$ such that

$$\int_t^{t+\tilde{\delta}} y^\top \sum_{i=1}^n (I - \mathring{v}_i(\tau)\mathring{v}_i(\tau)^\top) y \, \mathrm{d}\tau \leq \tilde{c}.$$

Since the integrand is non-negative, and choosing $\tilde{\delta} = \delta$, it follows that there exist sequences $(t_k) \subset \mathbb{R}$ and $(y_k) \subset S^2$ such that

$$\lim_{k\to\infty} \int_{t_k}^{t_k+\delta} y_k^\top \sum_{i=1}^n (I - \mathring{v}_i(\tau)\mathring{v}_i(\tau)^\top) y_k \, \mathrm{d}\tau = 0.$$

Moreover, since $y_k^\top (I - \mathring{v}_i\mathring{v}_i^\top) y_k$ is always nonnegative, this implies

$$\lim_{k\to\infty} \int_{t_k}^{t_k+\delta} y_k^\top (I - \mathring{v}_i(\tau)\mathring{v}_i(\tau)^\top) y_k \, \mathrm{d}\tau = 0$$

for all $i = 1, \dots, n$. Define $\mathring{v}_{i,k}(s) := \mathring{v}_i(s + t_k)$ for $s \in [0, \delta]$ then the last equation can be rewritten as

$$\lim_{k\to\infty} \int_0^\delta y_k^\top (I - \mathring{v}_{i,k}(s)\mathring{v}_{i,k}(s)^\top) y_k \, \mathrm{d}s = 0$$

for all $i = 1, \dots, n$. By the Bolzano-Weierstrass theorem we can assume w.l.o.g. (passing to a subsequence) that $y_k \to \bar{y}$ for some $\bar{y} \in S^2$.

Now focus on the first n_1 directions $\mathring{v}_i = \mathring{a}_i$. Boundedness and uniform continuity of $\mathring{a}_i$ implies uniform boundedness and equicontinuity of $\mathring{v}_{i,k}(s) = \mathring{v}_i(s + t_k) = \mathring{a}_i(s + t_k)$ for $i = 1, \dots, n_1$. By the Arzelà-Ascoli theorem each of the sequences $(\mathring{v}_{i,k})$ contains a uniformly convergent subsequence with a continuous limit $\bar{v}_i \colon [0, \delta] \longrightarrow S^2$. For the remaining n_2 directions $\mathring{v}_{j+n_1} = b_j$ we apply the Arzelà-Ascoli theorem in a version for continuously differentiable functions (see e.g. the discussion in Chapter I.5 of [39], in particular Theorems 5.11 and 5.20 therein) which yields a uniformly convergent subsequence for each of the sequences $(\mathring{v}_{j+n_1,k})$ with a continuously differentiable limit $\bar{v}_{j+n_1} \colon [0, \delta] \longrightarrow S^2$ such that $\frac{\mathrm{d}}{\mathrm{d}s}\bar{v}_{j+n_1}(s)$ is the pointwise limit of the derivatives for all $s \in [0, \delta]$. Here we have used the hypotheses on the regularity of Ω and the $\mathring{b}_j$. In particular, the uniform regularity conditions on Ω and $\frac{\mathrm{d}}{\mathrm{d}s}\mathring{b}_j(s)$ will ensure equicontinuity of $\frac{\mathrm{d}}{\mathrm{d}s}\mathring{v}_{j+n_1,k}(s) = \frac{\mathrm{d}}{\mathrm{d}s}\mathring{v}_{j+n_1}(s + t_k) = \frac{\mathrm{d}}{\mathrm{d}s} b_j(s + t_k)$ for $j = 1, \dots, n_2$.

Passing to the above subsequences and applying the Lebesgue dominated convergence theorem yields

$$\int_0^\delta \bar{y}^\top (I - \bar{v}_i(s)\bar{v}_i(s)^\top) \bar{y} \, \mathrm{d}s = 0$$

and hence that $\bar{y}^\top (I - \bar{v}_i(s)\bar{v}_i^\top(s))\bar{y} = 0$ for almost all $s \in [0, \delta]$ and for all $i = 1, \dots, n$. Since the $\bar{v}_i$ are continuous this equality holds in fact for all $s \in [0, \delta]$ and consequently $\bar{v}_i(s) = \pm\bar{y}$ for all $s \in [0, \delta]$ and all $i = 1, \dots, n$ where the sign is fixed for each i due to continuity of $\bar{v}_i$.

Focussing on the first n_1 directions again this means that

$$\begin{aligned}
&\lim_{k\to\infty} \lambda_2 \left(\int_{t_k}^{t_k+\delta} \sum_{i=1}^{n_1} \mathring{a}_i(\tau)\mathring{a}_i(\tau)^\top \mathrm{d}\tau \right) \\
&\quad = \lim_{k\to\infty} \lambda_2 \left(\int_{t_k}^{t_k+\delta} \sum_{i=1}^{n_1} \mathring{v}_i(\tau)\mathring{v}_i(\tau)^\top \mathrm{d}\tau \right) \\
&\quad = \lim_{k\to\infty} \lambda_2 \left(\int_0^\delta \sum_{i=1}^{n_1} \mathring{v}_i(s + t_k)\mathring{v}_i(s + t_k)^\top \mathrm{d}s \right) \\
&\quad = \lim_{k\to\infty} \lambda_2 \left(\int_0^\delta \sum_{i=1}^{n_1} \mathring{v}_{i,k}(s)\mathring{v}_{i,k}(s)^\top \mathrm{d}s \right) \qquad (16) \\
&\quad = \lambda_2 \left(\int_0^\delta \sum_{i=1}^{n_1} \lim_{k\to\infty} \mathring{v}_{i,k}(s)\mathring{v}_{i,k}(s)^\top \mathrm{d}s \right) \\
&\quad = \lambda_2 \left(\int_0^\delta \sum_{i=1}^{n_1} \bar{v}_i(s)\bar{v}_i(s)^\top \mathrm{d}s \right) \\
&\quad = \lambda_2 \left(\int_0^\delta \sum_{i=1}^{n_1} \bar{y}\bar{y}^\top \mathrm{d}s \right) = 0.
\end{aligned}$$

For the remaining n_2 directions we compute

$$\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}s}\bar{v}_{j+n_1}(s) = \lim_{k\to\infty} \frac{\mathrm{d}}{\mathrm{d}s}\mathring{v}_{j+n_1,k}(s) \\
&= \lim_{k\to\infty} \frac{\mathrm{d}}{\mathrm{d}s}\mathring{v}_{j+n_1}(s + t_k) = \lim_{k\to\infty} \frac{\mathrm{d}}{\mathrm{d}s} b_j(s + t_k) \\
&= \lim_{k\to\infty} R \left[\Omega_x \mathring{b}_j + \frac{\mathrm{d}}{\mathrm{d}s}\mathring{b}_j \right] (s + t_k),
\end{aligned}$$

and using the fact that the singular values of rotation matrices are uniformly bounded away from zero it follows that

$$\lim_{k\to\infty} \left[\Omega_x \mathring{b}_j + \frac{\mathrm{d}}{\mathrm{d}s}\mathring{b}_j \right] (s + t_k) = 0$$

for all $s \in [0, \delta]$ and all $i = 1, \dots, n$. Summing over $j = 1, \dots, n_2$, integrating over $s \in [0, \delta]$, and once more applying the Lebesgue dominated convergence theorem yields

$$\lim_{k\to\infty} \| \int_{t_k}^{t_k+\delta} \sum_{j=1}^{n_2} \left(\Omega_\times(\tau)\mathring{b}_j(\tau) + \frac{\mathrm{d}}{\mathrm{d}\tau}\mathring{b}_j(\tau) \right) \mathrm{d}\tau \| = 0,$$

a contradiction to Condition (13) when combined with Equation (16). This completes the proof. ■

The proof of Theorem 4.3 now follows from a direct application of Propositions 4.6 and 4.7 noting that the conditions on the derivatives of the $\mathring{b}_j$ imply that they are (piecewise) uniformly continuous. A similar discussion as at the end of the previous section shows that Condition (13) can indeed be fulfilled. In particular, for the case of more than one inertial reference directions the usual non-collinearity condition for constant reference directions now generalizes as follows.

Definition 4.8: A collection of directions $v_i \colon \mathbb{R} \longrightarrow S^2$, $i = 1, \dots, n$ is called *uniformly non-collinear* if there exists a $c > 0$ such that for all times $t \geq 0$

$$\max_{i \neq j} |v_i(t) \times v_j(t)| > c.$$

The proof of the following lemma is provided in the Appendix.

Lemma 4.9: A collection of locally integrable uniformly non-collinear directions is persistently exciting.

It follows that a collection of (more than one) uniformly non-collinear and sufficiently regular inertial reference directions will lead to observer error convergence. Similarly, a *single constant* body-fixed reference direction $\mathring{b}$ will yield observer error convergence if there exist $\delta > 0$ and $c > 0$ such that for all $t \geq 0$

$$\| \int_t^{t+\delta} \Omega_\times(\tau) \mathring{b} \, \mathrm{d}\tau \| > c.$$

The geometric interpretation of this condition is the same as was discussed at the end of Section III.

V. Conclusion

In this paper, we introduced a general system that incorporates all non-linear attitude observers on $\mathrm{SO}(3)$ introduced in recent literature [19], [20], [21], [22], [30], and via the standard equivalence, the related observers posed on the unit quaternions [14], [15], [16], [17], [18]. By understanding the role of the reference direction as well as the measurement and considering time-varying reference and measurement directions as input to the filter, we have provided a comprehensive observability and stability analysis of these systems. A key outcome is conditions that ensure almost global asymptotic and local exponential stability of attitude observers based on a single vector measurement as long as the persistency of excitation conditions are met. This result generalizes the stability results available in prior work based on rank conditions that required at least two or more vector measurements to guarantee stability, even when persistency of excitation conditions were met.

Appendix

Quaternion representation of observers

The set of unit quaternions is denoted $\mathbb{Q} = \{q = (s, v) \in \mathbb{R} \times \mathbb{R}^3 : |q| = 1\}$. The set $\mathbb{Q}$ is a group under the operation

$$q_1 \otimes q_2 = \begin{bmatrix} s_1 s_2 - v_1^T v_2 \\ s_1 v_2 + s_2 v_1 + v_1 \times v_2 \end{bmatrix}$$

with identity element $\mathbf{1} = (1, 0, 0, 0)$ and inverse $q^{-1} = (s, -v)$. The group of unit quaternions is homomorphic to $\mathrm{SO}(3)$ via the map

$$F : \mathbb{Q} \to \mathrm{SO}(3), \quad F(q) := I_3 + 2 s v_\times + 2 v_\times^2$$

This map is a two to one mapping of $\mathbb{Q}$ onto $\mathrm{SO}(3)$ with kernel $\{(1, 0, 0, 0), (-1, 0, 0, 0)\}$, thus, $\mathbb{Q}$ is locally isomorphic to $\mathrm{SO}(3)$ via F. Given $R \in \mathrm{SO}(3)$ such that $R = \exp(\theta a_\times)$ then $F^{-1}(R) = \{\pm(\cos(\frac{\theta}{2}), \sin(\frac{\theta}{2}) a)\}$. Let $\Omega \in \{B\}$ denote a body-fixed frame velocity, then the pure quaternion $\mathbf{p}(\Omega) = (0, \Omega)$ is associated with a quaternion velocity. We use the notation $\mathbf{p}^\dagger(s, v) := v \in \mathbb{R}^3$ to be the projection onto the vector part of the quaternion. In particular, $\mathbf{p}^\dagger \circ \mathbf{p}(\Omega) = \Omega$. Consider the rotation kinematics on $\mathrm{SO}(3)$ given by Equation 1, then the associated quaternion kinematics are given by

$$\dot{q} = \frac{1}{2} q \otimes \mathbf{p}(\Omega) \tag{17}$$

For a vector $v \in S^2$ and a rotation $R = F(q)$, the transformation by a rotation, Rv, can be written in quaternion multiplication by

$$Rv = F(q) v = \mathbf{p}^\dagger \left(q \otimes \mathbf{p}(v) \otimes q^{-1} \right).$$

Indeed, one also has $\mathbf{p}(F(q)v) = q \otimes \mathbf{p}(v) \otimes q^{-1}$.

The combined observer (10) can be written in quaternion representation by

$$\dot{\hat{q}} = \frac{1}{2} \hat{q} \otimes \mathbf{p} \Bigg(\Omega + \sum_{i=1}^{n_1} k_i \left[a_i \times \mathbf{p}^\dagger \left(\hat{q}^{-1} \otimes \mathbf{p}(\mathring{a}_i) \otimes \hat{q} \right) \right] + \sum_{j=1}^{n_2} l_j \left[\mathring{b}_j \times \mathbf{p}^\dagger \left(\hat{q}^{-1} \otimes \mathbf{p}(b_j) \otimes \hat{q} \right) \right] \Bigg) \tag{18}$$

The solutions of this observer equation correspond to the solutions to Equation (10) via the homomorphism F. That is

$$\hat{R}(t) = F(\hat{q}(t)),$$

for $\hat{q}(t)$ a solution to Equation (18) and $\hat{R}(t)$ a solution to Equation (10) with initial conditions $\hat{R}(0) = F(\hat{q}(0))$.

Due to the two-to-one correspondence of unit quaternions to rotations, the estimates $+\hat{q}$ and $-\hat{q}$ correspond to the same rotation $\hat{R}$ and it is impossible to distinguish between them based on measurements. In this sense, the non-linear observer system on the unit quaternions is always unobservable, however, since the two estimates correspond to the same rotation, this indistinguishability is irrelevant in practice. It is straightforward to verify that Theorem 3.2 holds for the quaternion representation of an observer up to the equivalence between $\pm\hat{q}$. Similarly, it is easily verified that Theorem 4.3 holds subject to convergence to the set $\{(1, 0, 0, 0), -(1, 0, 0, 0)\}$, both isolated locally exponentially stable equilibria in the unit quaternions.

Proofs of Lemmas

In this appendix we provide the proofs of all lemmas that were stated in the main text. Some of these proofs depend on additional auxiliary lemmas that are stated and proved first.

Lemma A.1: Let $v_i \colon \mathbb{R} \longrightarrow S^2$, $i = 1, \dots, n$ be locally integrable and let $t_2 > t_1$. Then

$$\lambda_2 \left(\int_{t_1}^{t_2} \sum_{i=1}^{n} v_i(\tau) v_i(\tau)^\top \mathrm{d}\tau \right) > c \tag{19}$$

implies

$$\lambda_{\min}\left(\int_{t_1}^{t_2}\sum_{i=1}^{n}(I-v_i(\tau)v_i(\tau)^\top)\mathrm{d}\tau\right)>c, \tag{20}$$

and in turn Equation (20) implies

$$\lambda_2\left(\int_{t_1}^{t_2}\sum_{i=1}^{n}v_i(\tau)v_i(\tau)^\top\mathrm{d}\tau\right)>c/2. \tag{21}$$

Proof: Define $\Delta t := t_2 - t_1$ and

$$Q := \int_{t_1}^{t_2}\sum_{i=1}^{n}v_i(\tau)v_i(\tau)^\top\mathrm{d}\tau$$

and let $U\operatorname{diag}(\lambda_1,\lambda_2,\lambda_3)U^\top$ be an orthonormal diagonalization of Q with $\lambda_1\ge\lambda_2\ge\lambda_3\ge 0$. Note that $\operatorname{tr}(Q)=n\Delta t$ by the linearity of $\operatorname{tr}(\cdot)$ and the fact that $\operatorname{tr}(vv^\top)=1$ if $|v|=1$. On the other hand, $\operatorname{tr}(Q)=\lambda_1+\lambda_2+\lambda_3$ by the similarity invariance of $\operatorname{tr}(\cdot)$ and hence $\lambda_1=n\Delta t-\lambda_2-\lambda_3\le n\Delta t-\lambda_2$. But then $\lambda_2>c$ implies $\lambda_1<n\Delta t-c$. It now follows that

$$\begin{aligned}P &:= \int_{t_1}^{t_2}\sum_{i=1}^{n}(I-v_i(\tau)v_i(\tau)^\top)\mathrm{d}\tau = n\Delta t I - Q\\ &= U\operatorname{diag}(n\Delta t-\lambda_1, n\Delta t-\lambda_2, n\Delta t-\lambda_3)U^\top\end{aligned}$$

and hence $\lambda_{\min}(P)=n\Delta t-\lambda_1>c$. Conversely, $\lambda_{\min}(P)=n\Delta t-\lambda_1>c$ implies $2\lambda_2\ge\lambda_2+\lambda_3=n\Delta t-\lambda_1>c$ and hence $\lambda_2>c/2$. ■

Lemma A.2: Let $a,b,x\in S^2$ with $|a\times b|>0$. Then

$$x^\top(I-aa^\top+I-bb^\top)x\ge(1/2)\cdot|a\times b|^2. \tag{22}$$

Proof: Compute

$$\begin{aligned}x^\top(I-aa^\top+I-bb^\top)x &= x^\top\left((a_\times)^\top a_\times+(b_\times)^\top b_\times\right)x\\ &=(a_\times x)^\top(a_\times x)+(b_\times x)^\top(b_\times)=|a\times x|^2+|b\times x|^2.\end{aligned}$$

Since $|a\times b|>0$ then a, b and $a\times b$ form a basis of $\mathbb{R}^3$ and hence $x=c_1a+c_2b+c_3(a\times b)$ for some real numbers c_1, c_2 and c_3. Substituting the expansion of x in the above calculation yields

$$\begin{aligned}&x^\top(I-aa^\top+I-bb^\top)x\\ &=c_2^2|a\times b|^2+c_3^2|a\times(a\times b)|^2+c_1^2|b\times a|^2+c_3^2|b\times(a\times b)|^2\\ &=c_2^2|a\times b|^2+c_3^2|a|^2|a\times b|^2+c_1^2|a\times b|^2+c_3^2|b|^2|a\times b|^2\\ &=(c_1^2+c_2^2+2c_3^2)|a\times b|^2\ge(c_1^2+c_2^2+c_3^2)|a\times b|^2.\end{aligned}$$

Here we have used the fact that $a\times b$ is orthogonal to both a and b as well as the fact that $a,b\in S^2$. Furthermore,

$$\begin{aligned}1=x^\top x&=\left(\sum_{i=1}^{3}c_ie_i\right)^\top A^\top A\left(\sum_{i=1}^{3}c_ie_i\right)\\ &\le(c_1^2+c_2^2+c_3^2)\lambda_{\max}(A^\top A),\end{aligned}$$

where $e_i\in\mathbb{R}^3$ denotes the i-th standard basis vector, $A\in\mathbb{R}^{3\times3}$ is the matrix with columns a, b and $a\times b$ and $\lambda_{\max}(A^\top A)$ denotes the largest eigenvalue of the symmetric matrix $A^\top A$. The eigenvalues of $A^\top A$ are easily computed as $1\pm(a^\top b)=1\pm\cos(\alpha)$ and $|a\times b|^2=\sin^2(\alpha)$, where α denotes the angle between a and b, and hence $\lambda_{\max}(A^\top A)\le 2$. It follows that

$$x^\top(I-aa^\top+I-bb^\top)x\ge(1/2)\cdot|a\times b|^2.$$

■

Proof of Lemma 3.4: This follows directly from Lemma A.1 by setting $t_1:=0$, $t_2:=T$ and $c:=0$. ■

Proof of Lemma 3.8: Let $v_i\colon\mathbb{R}\longrightarrow S^2$, $i=1,\dots,n$ be sufficiently non-collinear then there exists a time $T>0$ such that

$$\int_0^T\sum_{i\ne j}|v_i(\tau)\times v_j(\tau)|\;\mathrm{d}\tau>0.$$

But then there exists a set $\mathfrak{T}\subset[0,T]$ of non-zero measure such that for every time instance $\tau\in\mathfrak{T}$ there are two indices i and j with $|v_i(\tau)\times v_j(\tau)|>0$. Setting $a(\tau):=v_i(\tau)$ and $b(\tau):=v_j(\tau)$ yields, at any time instance $\tau\in\mathfrak{T}$, a pair $a,b\in S^2$ to which Lemma A.2 can be applied. Using the fact that $x^\top(I-vv^\top)x\ge0$ for all $v,x\in S^2$, this implies

$$\begin{aligned}&x^\top\left(\int_0^T\sum_{i=1}^{n}(I-v_i(\tau)v_i(\tau)^\top)\mathrm{d}\tau\right)x\ge\\ &\int_{\mathfrak{T}}x^\top\left(\sum_{i=1}^{n}(I-v_i(\tau)v_i(\tau)^\top)\right)x\;\mathrm{d}\tau>0\end{aligned}$$

for all $x\in S^2$. An application of Lemma 3.4 completes the proof. ■

Proof of Lemma 4.2: Let $v_i\colon\mathbb{R}\longrightarrow S^2$, $i=1,\dots,n$ be persistently exciting then there exist $\delta>0$ and $c'>0$ such that for all $t\ge0$

$$\lambda_2\left(\int_t^{t+\delta}\sum_{i=1}^{n}v_i(\tau)v_i(\tau)^\top\mathrm{d}\tau\right)>c'.$$

But then for all $t\ge0$

$$\begin{aligned}&\lambda_{\min}\left(\int_t^{t+\delta}\sum_{i=1}^{n}q_i(I-v_i(\tau)v_i(\tau)^\top)\;\mathrm{d}\tau\right)\ge\\ &q_{\min}\cdot\lambda_{\min}\left(\int_t^{t+\delta}\sum_{i=1}^{n}(I-v_i(\tau)v_i(\tau)^\top)\;\mathrm{d}\tau\right)>q_{\min}\cdot c',\end{aligned}$$

where $q_{\min}:=\min\{q_i\,|\,i=1,\dots,n\}>0$ and the second inequality follows from Lemma A.1. The first half of the statement now follows with $c:=q_{\min}\cdot c'$. Conversely, assume that there exist $\delta>0$ and $c>0$ such that for all $t\ge0$

$$\lambda_{\min}\left(\int_t^{t+\delta}\sum_{i=1}^{n}q_i(I-v_i(\tau)v_i(\tau)^\top)\;\mathrm{d}\tau\right)>c.$$

Then for all $t\ge0$

$$\lambda_{\min}\left(\int_t^{t+\delta}\sum_{i=1}^{n}(I-v_i(\tau)v_i(\tau)^\top)\;\mathrm{d}\tau\right)>c/q_{\max},$$

where $q_{\max}:=\max\{q_i\,|\,i=1,\dots,n\}>0$. An application of Lemma A.1 completes the proof. ■

Proof of Lemma 4.9: Let $v_i\colon\mathbb{R}\longrightarrow S^2$, $i=1,\dots,n$ be locally integrable and uniformly non-collinear. Then for every time instance $\tau\ge0$ there exist two indices i and j

with $|v_i(\tau) \times v_j(\tau)| > c > 0$. Setting $a(\tau) := v_i(\tau)$ and $b(\tau) := v_j(\tau)$ yields, at any time instance $\tau \geq 0$, a pair $a, b \in S^2$ to which Lemma A.2 can be applied. Using the fact that $x^\top (I - vv^\top)x \geq 0$ for all $v, x \in S^2$, this implies

$$x^\top \left(\int_t^{t+\delta} \sum_{i=1}^n (I - v_i(\tau)v_i(\tau)^\top) \mathrm{d}\tau \right) x \geq$$
$$\int_t^{t+\delta} x^\top \left(\sum_{i=1}^n (I - v_i(\tau)v_i(\tau)^\top) \right) x \ \mathrm{d}\tau > (1/2) \cdot \delta c^2$$

for all $x \in S^2$ and where $t \geq 0$ and $\delta > 0$ are arbitrary. An application of Lemma 4.2 with $q_i := 1$, $i = 1, \ldots, n$ completes the proof. ■

Acknowledgments

This work was supported by the Australian Research Council through the ARC Discovery Project DP0987411 "State Observers for Control Systems with Symmetry" and the ANR-PSIRob project SCUAV "Sensory Control of Unmanned Aerial Vehicles".

References

[1] J. Farrell, "Attitude determination by Kalman filter," *Automatica*, vol. 6, no. 5, pp. 419–430, 1970.

[2] B. Barshan and H. Durrant-Whyte, "Inertial navigation systems for mobile robots," *IEEE Transactions on Robotics and Automation*, vol. 44, no. 4, pp. 751–760, 1995.

[3] E. Lefferts, F. Markley, and M. Shuster, "Kalman filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, vol. 5, no. 5, pp. 417–429, 1982.

[4] F. Markley, "Attitude error representations for Kalman filtering," *Journal of Guidance, Control, and Dynamics*, vol. 63, no. 2, pp. 311–317, 2003.

[5] J. Crassidis and F. Markley, "Unscented filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, vol. 26, no. 4, pp. 536–542, 2003.

[6] Y. Cheng and J. Crassidis, "Particle filtering for sequential spacecraft attitude estimation," in *AIAA Guidance, Navigation, and Control Conference*, 2004, paper no. 5337.

[7] Y. Oshman and A. Carmi, "Attitude estimation from vector observations using genetic-algorithm-embedded quaternion particle filter," *Journal of Guidance, Control, and Dynamics*, vol. 29, no. 4, pp. 879–891, 2006.

[8] S. Salcudean, "A globally convergent angular velocity observer for rigid body motion," *IEEE Transactions on Automatic Control*, vol. 46, no 12, pp. 1493–1497, 1991.

[9] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Non-linear complementary filters on the special orthogonal group," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1203–1218, 2008.

[10] S. Bonnabel, P. Martin, and P. Rouchon, "Symmetry-preserving observers," *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2514–2526, 2008.

[11] ——, "Nonlinear symmetry preserving observers on Lie groups," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1709–1713, 2009.

[12] J. Crassidis, F. Markley, and Y. Cheng, "Nonlinear attitude filtering methods," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 1, pp. 12–28, 2007.

[13] O. Egland and J. Godhaven, "Passivity-based adaptive attitude control of a rigid spacecraft," *IEEE Transactions on Automatic Control*, vol. 39, pp. 842–846, 1994.

[14] B. Vik and T. Fossen, "A nonlinear observer for GPS and INS integration," in *Proceedings of the 40th IEEE Conference on Decision and Control*, 2001, pp. 2956–2961.

[15] J. Thienel and R. M. Sanner, "A coupled nonlinear spacecraft attitude controller and observer with an unknow constant gyro bias and gyro noise," *IEEE Transactions on Automatic Control*, vol. 48, no. 11, pp. 2011–2015, 2003.

[16] A. Tayebi and S. McGilvray, "Attitude stabilization of a VTOL quadrotor aircraft," *IEEE Transactions on Control Systems Technology*, vol. 14, no. 3, pp. 562–571, 2006.

[17] P. Martin and E. Salaün, "Invariant observers for attitude and heading estimation from low-cost inertial and magnetic sensors," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007, pp. 1039–1045.

[18] P. Martin and E. Salaün, "An invariant observer for earth-velocity-aided attitude heading reference systems," in *Proceedings of the 17th World Congress The International Federation of Automatic Control*, 2008, pp. 9857–9864.

[19] T. Hamel and R. Mahony, "Attitude estimation on $\mathrm{SO}(3)$ based on direct inertial measurements," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2006, pp. 2170–2175.

[20] S. Bonnabel, P. Martin, and P. Rouchon, "A non-linear symmetry-preserving observer for velocity-aided inertial navigation," in *Proceedings of the American Control Conference*, 2006, pp. 2910–2914.

[21] D. Campoloa, F. Kellerb, and E. Guglielmellia, "Inertial/magnetic sensors based orientation tracking on the group of rigid body rotations with application to wearable devices," in *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 4762–4767.

[22] J. Vasconcelos, C. Silvestre, and P. Oliveira, "A nonlinear observer for rigid body attitude estimation using vector observations," in *Proceedings of the 17th World Congress The International Federation of Automatic Control*, 2008, pp. 8599–8604.

[23] M.-D. Hua, "Attitude estimation for accelerated vehicles using GPS/INS measurements," *Control Engineering Practice, Special Issue on Aerial Robotics*, vol. 18, no. 7, pp. 723–732, 2010.

[24] G. Baldwin, R. Mahony, and J. Trumpf, "A nonlinear observer for 6 DOF pose estimation from inertial and bearing measurements," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2009, pp. 2237–2242.

[25] J. Vasconcelos, R. Cunha, C. Silvestre, and P. Oliveira, "A nonlinear position and attitude observer on SE(3) using landmark measurements," *Systems & Control Letters*, vol. 59, no. 3-4, pp. 155–166, 2010.

[26] C. Lageman, J. Trumpf, and R. Mahony, "Gradient-like observers for invariant dynamics on a Lie group," *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 367–377, 2010.

[27] ——, "Observers for systems with invariant outputs," in *Proceedings of the European Control Conference*, 2009, pp. 4587–4592.

[28] N. Aghannan and P. Rouchon, "An intrinsic observer for a class of Lagrangian systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 936–945, 2003.

[29] D. Maithripala, W. Dayawansa, and J. Berg, "Intrinsic observer-based stabilization for simple mechanical systems on Lie-groups," *SIAM Journal on Control and Optimization*, vol. 44, no. 5, pp. 1691–1711, 2005.

[30] J. Vasconcelos, C. Silvestre, and P. Oliveira, "A nonlinear GPS/IMU based observer for rigid body attitude and position estimation," in *Proceedings of the 47th IEEE Conference on Decision and Control*, 2008, pp. 1255–1260.

[31] R. Mahony, T. Hamel, J. Trumpf, and C. Lageman, "Nonlinear attitude observers on $\mathrm{SO}(3)$ for complementary and compatible measurements: A theoretical study," in *Proceedings of the 48th IEEE Conference on Decision and Control*, 2009, pp. 6407–6412.

[32] A. Khosravian and M. Namvar, "Globally exponential estimation of satellite attitude using a single vector measurement and gyro," in *Proceedings of the 49th IEEE Conference on Decision and Control*, 2010, pp. 364–369.

[33] N. Chaturvedi, A. Sanyal, and N. McClamroch, "Rigid-body attitude control," *IEEE Control Systems Magazine*, vol. 31, no. 3, pp. 30–51, 2011.

[34] H. J. Sussmann, "Single-input observability of continuous-time systems," *Mathematical Systems Theory*, vol. 12, no. 4, pp. 371–393, 1979.

[35] F. Zhang and N. Leonard, "Coordinated patterns of unit speed particles on a closed curve," *Systems and Control Letters*, vol. 56, pp. 397–407, 2007.

[36] A. Morgan and K. Narendra, "On the uniform asymptotic stability of certain linear nonautonomous differential equations," *SIAM Journal on Control and Optimization*, vol. 15, no. 1, pp. 5–24, 1977.

[37] J. K. Hale, *Ordinary differential equations*, ser. Pure and Applied Mathematics, Vol. XXI. Wiley-Interscience, 1969.

[38] J. L. Daleckiĭ and M. G. Kreĭn, *Stability of solutions of differential equations in Banach space*, ser. Translations of Mathematical Monographs, Vol. 43. American Mathematical Society, 1974.

[39] J. Jost, *Postmodern analysis*, 3rd ed. Springer, 2005.

Jochen Trumpf (M'04) received the Dipl.-Math. and Dr. rer. nat. degrees in mathematics from the University of Würzburg, Germany, in 1997 and 2002, respectively. He is currently a Fellow in the Research School of Engineering at the Australian National University. His research interests include observer theory and design, linear systems theory and optimisation on manifolds with applications in robotics, computer vision and wireless communication.

Robert Mahony (S'92–M'95–SM'08) obtained a science degree majoring in applied mathematics and geology from the Australian National University (ANU) in 1989. After working for a year as a geophysicist processing marine seismic data he returned to study at ANU and obtained a Ph.D. in systems engineering in 1994. Between 1994 and 1997 he worked as a Research Fellow in the Cooperative Research Centre for Robust and Adaptive Systems in Australia. From 1997 to 1999 he was a post-doctoral fellow in the CNRS laboratory for Heuristics Diagnostics and complex systems (Heudiasyc), Compiegne University of Technology in France. In 1999 he returned to Australia to take up a Logan Fellowship at Monash University, Melbourne, Australia. In 2001 he took a faculty position in the Australian National University in Canberra, Australia, where he is now a professor. His research interests are in non-linear control theory with applications in mechanical systems and motion systems, mathematical systems theory and geometric optimisation techniques with applications in linear algebra and digital signal processing.

Tarek Hamel received his Bachelor of Engineering from the "Institut d'Electronique et d'Automatique d'Annaba", Algeria, in 1991. He conducted his Ph.D. research at the "Laboratoire Heuristique et Diagnostic des systèmes complexes de Compiègne", France, and received his doctorate degree in Robotics from the University of Technologie of Compiègne in 1995. After two years as a research assistant at the University of Technology of Compiègne, he joined the "Centre d'Etudes de Mècanique d'Iles de France" in 1997 as an assistant professor. In 2001/2002 he spent one year as CNRS researcher at the Heudiasyc Laboratory. Since 2003, he has been a Professor at the I3S UNSA-CNRS laboratory of the University of Nice-Sophia Antipolis, France. His research interests include control theory and robotics with particular focus on nonlinear control and vision-based control. He is involved in applications of these techniques to the control of Unmanned Aerial Vehicles and Mobile Robots.

Christian Lageman received the Dipl.-Inf. degree in computer science and the Dipl.-Math. and Dr. rer. nat. degrees in mathematics from the University of Würzburg, Germany, in 2001, 2002 and 2007, respectively. From 2007 to 2008 he worked as a Research Fellow at the Centre for Mathematics and its Applications, The Australian National University. From 2008 to 2009, he held a postdoctoral fellowship at the Department of Electrical Engineering and Computer Science, University of Liège, Belgium. Since September 2009, he has been with the Department of Mathematics, University of Würzburg, Germany. His current research interests are in nonlinear control theory, optimization, and estimation on manifolds.