Attitude Estimation from Vector Measurements: Necessary and Sufficient Conditions and Convergent Observer Design

Bowen Yi, Lei Wang, and Ian R. Manchester

***Abstract*—The paper addresses the problem of attitude estimation for rigid bodies using (possibly time-varying) vector measurements, for which we provide a *necessary and sufficient* condition of distinguishability. Such a condition is shown to be strictly weaker than those previously used for attitude observer design. Thereafter, we show that even for the single vector case the resulting condition is sufficient to design almost globally convergent attitude observers, and two explicit designs are obtained. To overcome the weak excitation issue, the first design employs to make full use of historical information, whereas the second scheme dynamically generates a virtual reference vector, which remains non-collinear to the given vector measurement. Simulation results illustrate the accurate estimation despite noisy measurements.**

***Index Terms*—Nonlinear system, observer design, observability, attitude estimation**

I. Introduction

The attitude of a rigid body is its orientation with respect to an inertial reference frame. Attitude estimation is an essential element in a wide range of robotics and aerospace applications, in particular control, navigation, and localization tasks. Many common sensor types, e.g. magnetometers, accelerometers, and monocular cameras, provide body-fixed measurements of quantities with a known inertial value, e.g. the earth's magnetic field and gravitational force, or the bearing to certain known landmarks. These are known as *complementary* measurements [20]. In some less common scenarios a set of known vectors in the body-fixed frame are measured in the inertial frame, e.g. measurements from two GPS receivers attached to the rigid body with a known base-line. These are known as *compatible* measurements [20].

Estimation of attitude from multiple non-collinear vector measurements was formulated as a total least-squares problem over rotation matrices by Wahba [21]. Several efficient algorithms exist for its solution, including singular value decomposition methods, TRIAD, and QUEST [19].

However, when estimating a time-varying attitude it often is beneficial to fuse the vector measurements with information from gyroscopes using a dynamical model. The resulting dynamic estimator is commonly known as a filter or observer. These approaches can significantly reduce the impact of high-frequency measurement noise. Furthermore, in many applications there is only a single vector available for attitude estimation and in this case the attitude is not completely determined at a single moment. Applications for estimation from a single vector measurement include Sun sensors in eclipse periods [12], improving reliability with redundant measurements and simplifying designs [16], as well visual-inertial navigation with only two feature points visible in some periods.

This work was supported by Australian Research Council.

B. Yi and I. R. Manchester are with the Australian Centre for Field Robotics, and Sydney Institute for Robotics and Intelligent Systems, The University of Sydney, Sydney, NSW 2006, Australia (email: bowen.yi, ian.manchester@sydney.edu.au)

L. Wang is with State Key Laboratory of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou 310027, China (email: lei.wangzju@zju.edu.cn, corresponding author)

Among filtering approaches, extended Kalman filter is the most widely-applied for attitude estimation. However the domain of attraction is intrinsically local since the filter is based on first-order linearization; see [7] for a recent review. Alternatively, interest in nonlinear attitude observers was spurred by Salcudean's seminal work [17], and has achieved significant progress since then. There are many nonlinear attitude observers making direct use of vector measurements, e.g., with multiple measurements [10], [20], [26] or single vector measurements [1], [2], [8], [9]. The latter works impose a persistently non-constant condition on the single reference vector, or similar conditions in which the *uniformity* of excitation with respect to time plays an essential role to guarantee asymptotic convergence. In [20], the authors provide a comprehensive treatment of observability of a rigid-body attitude kinematic model with vectorial outputs. However, as illustrated in [20, Remark 3.9], the condition is only sufficient but *not* necessary for distinguishability, a specific type of observability for nonlinear dynamical systems [4], [5]. In this paper, we revisit the problem of observability analysis and propose two novel attitude observers. To be precise, the main contributions of the note are two-fold:

- **C1** For the problem of attitude estimation from vector measurements, we provide the *necessary and sufficient* condition to distinguishability of the associated dynamical model, which is known as the necessity to reconstruct attitude over time in any deterministic estimators;
- **C2** We show that the resulting distinguishability condition is also sufficient to design a continuous-time attitude observer. By focusing on single vector measurements, we provide two novel almost globally convergent attitude observers, which require significantly weaker conditions than existing methods.

The constructive tool we adopt in observer design is the parameter estimation-based observer (PEBO), which was recently proposed in Euclidean space [13], [14], and extended to matrix Lie groups by the authors in [23], [24]. Its basic idea is to translate system state observation into the estimation of certain constant quantities. The interested reader may refer to [25] for the geometric interpretation to PEBOs. In contrast to the case with at least two non-collinear vectors in [23], [24], in this paper we consider a more challenging scenario with only a single vector measurement available under a weak excitation condition. We are unaware of any previous results capable to deal with such a case. In the first observer design, after translating the problem into on-line parameter identification, we propose a mechanism to integrate both the historical and current information to achieve uniform convergence. The second proposed scheme uses a filter to generate a "virtually" measurable vector, which remains non-collinear with respect to the given reference vector.

Notation. $I_n \in \mathbb{R}^{n\times n}$ represents the identity matrix of dimension n, and $0_n \in \mathbb{R}^n$ and $0_{n\times m} \in \mathbb{R}^{n\times m}$ denote the zero column vector of dimension n and the zero matrix of dimension $n\times m$, respectively. We use $\mathbb{N}$ to represent the set of all natural integers, and $\mathbb{N}_+$ for the set of positive integers. We also define the skew-symmetric matrix

$$\mathcal{J} := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Given a square matrix $A \in \mathbb{R}^{n\times n}$ and a vector $x \in \mathbb{R}^n$, the Frobenius norm is defined as $\|A\| = \sqrt{\mathrm{tr}(A^\top A)}$, and $|x|$ represents the standard Euclidean norm. The n-sphere is defined as $\mathbb{S}^n := \{x \in \mathbb{R}^{n+1} : |x| = 1\}$, and we use $SO(3)$ to represent the special orthogonal group, and $\mathfrak{so}(3)$ is the associated Lie algebra as the set of skew-symmetric matrices satisfying $SO(3) = \{R \in \mathbb{R}^{3\times 3} | R^\top R = I_3,\ \det(R) = 1\}$. Given a variable $R \in SO(3)$, we use $|R|_I$ to represent the normalized distance to I_3 on $SO(3)$ with $|R|_I^2 := \frac{1}{4}\mathrm{tr}(I_3 - R)$. The operator $\mathrm{skew}(\cdot)$ is defined as $\mathrm{skew}(A) := \frac{1}{2}(A - A^\top)$ for a square matrix A. Given $a \in \mathbb{R}^3$, we define the operator $(\cdot)_\times$ as

$$a_\times := \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \in \mathfrak{so}(3),$$

and its inverse operator is defined as $\mathrm{vex}(a_\times) = a$.

The paper is organized as follows. In Section II, we recall the dynamical models and some basic definitions used in the paper. It is followed by the necessary and sufficient condition for observability in Section III. Based on the proposed condition, we introduce two nonlinear attitude observer design in Section IV, the simulation results of which are presented in Section VI. Some discussions with practical considerations may be found in Section V. The paper is wrapped up by some concluding remarks in Section VII. A preliminary version of this paper has been submitted to the 2022 IFAC Symposium on Nonlinear Control Systems (NOLCOS).

II. Problem Formulation

The aim of this note is to study the observability and observer design of the rotation matrix representing the coordinates of the body-fixed frame $\{B\}$ with respect to the coordinates of the inertial frame $\{I\}$, which lives in the group $SO(3)$. Its dynamics is given by

$$\dot{R} = R\omega_\times, \quad R(0) = R_0 \tag{1}$$

with the rotational velocity $\omega \in \mathbb{R}^3$ measured in the body-fixed coordinate. Assume there is a vector $g \in \mathbb{S}^2$, known in the inertial frame, being measured in the body-fixed frame, and the output is

$$y_{\mathtt{B}} = R^\top g \tag{2}$$

with $y_{\mathtt{B}} \in \mathbb{S}^2$, which is known as complementary measurement. We also consider the compatible measurement $y_{\mathtt{I}}$, i.e., a known vector $b \in \mathbb{S}^2$ in the body-fixed frame is measured in the inertial frame

$$y_{\mathtt{I}} = Rb \tag{3}$$

with $y_{\mathtt{I}} \in \mathbb{S}^2$. It is referred to [20, Sec. II] for more details about "complementary" and "compatible" measurements and their practical motivation.

Before closing this section, let us recall some definitions used throughout the paper.

Definition 1: (*Distinguishability [4]*) Consider an open set $\mathcal{X} \subset \mathbb{R}^n$ and a complete nonlinear system

$$\dot{x} = f(x,t), \quad y = h(x,t) \tag{4}$$

with state $x \in \mathbb{R}^n$ and output $y \in \mathbb{R}^m$. The system (4) is distinguishable on $\mathcal{X}$ if for all $(x_a, x_b) \in \mathcal{X} \times \mathcal{X}$,

$$h\big(X(t;t_0,x_a),t\big) = h\big(X(t;t_0,x_b),t\big),\ \forall t \ge t_0 \implies x_a = x_b,$$

in which $X(t;t_0,x_a)$ represents the solution at time t of (4) through x_0 at time t_0. In this paper, we focus on the particular case $t_0 = 0$.

Definition 2: (*Persistent and interval excitation [15]*) Given a bounded signal $\phi : \mathbb{R}_+ \to \mathbb{R}^n$, it is persistently excited (PE) if

$$\int_t^{t+T} \phi(s)\phi^\top(s)ds \ge \delta I_n,\ \forall t \ge 0$$

for some $T > 0, \delta > 0$; or intervally excited (IE), if there exists $T \ge 0$ such that for some $\delta > 0$

$$\int_0^T \phi(s)\phi^\top(s)ds \ge \delta I_n.$$

III. Necessary and Sufficient Conditions to Observability

First, we consider the observability for the case with multiple measurements

$$\begin{aligned} y_{\mathtt{B},i} &= R^\top g_i, & i \in \ell_1 := \{1,\ldots,n_1\} \\ y_{\mathtt{I},j} &= Rb_j, & j \in \ell_2 := \{1,\ldots,n_2\} \end{aligned} \tag{5}$$

with $n_1, n_2 \in \mathbb{N}$.[1] It is clear that the single measurement is corresponding to the case $n_1 + n_2 = 1$, for which we will construct two asymptotically convergent observers in the next section.

In the following proposition, we uncover a necessary and sufficient condition to the distinguishability for attitude estimation.

Proposition 1: The time-varying system (1) with the output (5), and $n := n_1 + n_2 \ge 1$, is distinguishable if and only if there exist two moments $t_1, t_2 \ge 0$ such that

$$\begin{aligned}\sum_{i,l\in\ell_1,j,k\in\ell_2} \Big|g_i(t_1)\times g_l(t_2)\Big| \;+\; \Big|g_i(t_1)_\times R_0 \Phi(0,t_2) b_j(t_2)\Big| \\ +\; \Big|b_j(t_1)_\times \Phi(t_1,t_2) b_k(t_2)\Big| > 0,\end{aligned} \tag{6}$$

in which $\Phi(s,t)$ is the state transition matrix of the time-varying system matrix $\omega_\times(t)$ from s to t.

Proof: The state transition matrix $\Phi(s,t)$ of the linear time-varying (LTV) system

$$\dot{x} = \omega_\times x$$

with $x \in \mathbb{R}^3$ is defined as

$$\begin{aligned} \frac{\partial}{\partial t}\Phi(s,t) &= \omega_\times(t)\Phi(s,t) \\ \Phi(s,s) &= I_3. \end{aligned}$$

It is equivalent to define $\Phi(s,t) = Q(s)^{-1}Q(t)$, in which $Q \in SO(3)$ is generated by the dynamics

$$\dot{Q} = Q\omega_\times, \quad Q(0) = I_3 \tag{7}$$

with $Q \in SO(3)$. From

$$\dot{\overbrace{RQ^\top}} = \dot{R}Q^\top - RQ^\top \dot{Q} Q^\top = 0,$$

we have for all $t, s \ge 0$

$$\begin{aligned} R(t)Q(t)^\top = R(0)Q(0)^\top \iff R(t) &= R_0 Q(t) \\ \iff R(t) &= R(s)\Phi(s,t), \end{aligned}$$

with $R_0 := R(0)$.

Now we collect all the measured outputs in the vector

$$\bar{y} = \mathrm{col}(y_{\mathtt{B},1}, \ldots, y_{\mathtt{B},n_1}, y_{\mathtt{I},1}, \ldots, y_{\mathtt{I},n_2}).$$

With a slight abuse of notation, we denote the output signal $\bar{y}$ from the initial condition $R_0 \in SO(3)$ as $\bar{y}(t;R_0)$. In terms of Definition 1, the system is distinguishable from $t = 0$ if and only if

$$\bar{y}(t;R_a) \not\equiv \bar{y}(t;R_b) \quad \implies \quad R_a \neq R_b \tag{8}$$

[1] If $n_i = 0$ $(i = 1,2)$, then the set ℓ_i is defined as the empty set $\emptyset$.

for any $R_a, R_b \in SO(3)$. Clearly, the above condition (8) is equivalent to the *identifiability* of the constant matrix $R_0 \in SO(3)$ from the nonlinear regressor equation

$$\bar{y} = h(R_0, t) \tag{9}$$

with the equation

$$h(R_0,t) := \begin{bmatrix} Q^\top(t) R_0^\top g_1(t) \\ \vdots \\ Q^\top(t) R_0^\top g_{n_1}(t) \\ R_0 Q(t) b_1(t) \\ \vdots \\ R_0 Q(t) b_{n_2}(t) \end{bmatrix}.$$

The regressor equation (9) can be equivalently rewritten as

$$Y(t) = R_0^\top \phi(t), \quad R_0 \in SO(3) \tag{10}$$

with $Y \in \mathbb{R}^{3\times n}$ and $\phi \in \mathbb{R}^{3\times n}$ given by

$$\begin{aligned} Y &:= Q\left[y_{\mathtt{B},1}, \ldots, y_{\mathtt{B},n_1}, b_1, \ldots, b_{n_2}\right] \\ \phi &:= \left[g_1, \ldots, g_{n_1}, y_{\mathtt{I},1}, \ldots, y_{\mathtt{I},n_2}\right]. \end{aligned}$$

Hence, the identifiability of the constant matrix R_0 on $SO(3)$ from the nonlinear regression model (9) is equivalent to the solvability of the Wahba problem for the regression model (10) over time [21] – invoking that (10) holds for all $t \geq 0$. That is the existence of moments $t_1, t_2 \geq 0$ such that

$$\phi_i(t_1) \times \phi_j(t_2) \neq 0 \tag{11}$$

for some $i, j \in \{1, \ldots, n\}$, with ϕ_i representing the i-th column vector of ϕ.

The last step of the proof is to show that the condition (11) is equivalent to (6). There are totally three possible cases when (11) holds true: 1) $i, j \in \{1, \ldots, n_1\}$, 2) $i, j \in \{n_1+1, \ldots, n\}$, and $i \in \{1, \ldots, n_1\}, j \in \{n_1+1, \ldots, n\}$.[2] For the first case, the condition (11) is equivalent to

$$\sum_{i,l\in\ell_1} \Big| g_i(t_1) \times g_l(t_2) \Big| > 0. \tag{12}$$

The second case is equivalent to for some $j, k \in \ell_2$

$$\begin{aligned} & y_{\mathtt{I},j}(t_1) \times y_{\mathtt{I},k}(t_2) \neq 0 \\ \iff & [R(t_1) b_j(t_1)]_\times R(t_2) b_k(t_2) \neq 0 \\ \iff & R(t_1)[b_j(t_1)]_\times R(t_1)^\top R(t_2) b_k(t_2) \neq 0 \\ \iff & [b_j(t_1)]_\times R(t_1)^\top R(t_2) b_k(t_2) \neq 0 \\ \iff & [b_j(t_1)]_\times \Phi(t_1,t_2) b_k(t_2) \neq 0 \end{aligned} \tag{13}$$

where in the second implication we use the identity $(Rb)_\times = Rb_\times R^\top$, the full-rankness of $R(t_1)$ in the third implication, and in the last

$$\begin{aligned} R(t_1)^\top R(t_2) &= Q(t_1)^\top R_0^\top R_0 Q(t_2) \\ &= \Phi(t_1,t_2). \end{aligned}$$

Note that the last line of the condition (13) can be compactly written as

$$\sum_{j,k\in\ell_2} \Big| b_j(t_1)_\times \Phi(t_1,t_2) b_k(t_2) \Big| > 0. \tag{14}$$

Similarly, we get that for the third case the condition (11) is equivalent to

$$\sum_{i\in\ell_1,\ j\in\ell_2} \Big| g_i(t_1)_\times R_0 \Phi(0,t_2) b_j(t_2) \Big| > 0. \tag{15}$$

[2]We do not distinguish the order of i and j.

Combining these three cases, it is sufficient to obtain (6). On the other hand, since each term in (6) is non-negative, if the condition (6) holds, at least one of the above cases should be satisfied. We complete the proof. □

For the case with *only* complementary or compatible measurements ($n_1 \cdot n_2 = 0$), then the distinguishability condition becomes

$$\sum_{i,l\in\ell_1,j,k\in\ell_2} \Big| g_i(t_1) \times g_l(t_2) \Big| + \Big| b_j(t_1)_\times \Phi(t_1,t_2) b_k(t_2) \Big| > 0.$$

If there are two types of measurements, the identifiability is dependent of the initial attitude R_0, and this implies that some region in $SO(3)$ may be not distinguishable for a given specific trajectory. However, the following corollary shows that such a region has zero Lebesgue measure in the group $SO(3)$. Note that the condition below does not rely on the initial attitude R_0.

Corollary 1: If the condition (6) is replaced by the initial attitude-independent term

$$\begin{aligned} \sum_{i,l\in\ell_1,j,k\in\ell_2} \Big| g_i(t_1) \times g_l(t_2) \Big| \; + \; \Big| g_i(t_1)_\times \Phi(0,t_2) b_j(t_2) \Big| \\ + \; \Big| b_j(t_1)_\times \Phi(t_1,t_2) b_k(t_2) \Big| > 0, \end{aligned} \tag{16}$$

the distinguishability is guaranteed almost surely.[3]

Proof: It is given in the appendix. □

Remark 1: In [20] the authors propose the following *sufficient* (but not necessary, c.f. [20, Remark 3.9]) condition for distinguishability of the given system.

$$\begin{aligned} \lambda_2 \left(\sum_{i\in\ell_1} \int_0^T g_i(s) g_i^\top(s) ds \right) \\ + \left\| \int_0^T \sum_{j\in\ell_2} \left(\omega_\times b_j(s) + \frac{d}{ds} b_j(s) \right) ds \right\| > 0, \end{aligned} \tag{17}$$

for some $T > 0$, with $\lambda_2(\cdot)$ representing the second largest eigenvalue of a square matrix. Note that in the above condition it is necessary to impose (piece-wise) smoothness of the signals b_j. In the following corollary, we show that the above condition is sufficient to the proposed necessary and sufficient condition (6).

Corollary 2: Consider the time-varying system (1) with the output (5), and $n := n_1 + n_2 \geq 1$. If (17) holds, then the condition in Proposition 1 is also verified.

Proof: It is given in the appendix. □

IV. Attitude Observer for A Single Vector Measurement

In this section, we show that the distinguishability condition – identified in Proposition 1 – is adequate to design a continuous-time observer with almost globally asymptotically convergent estimate to the unknown attitude.

Since the scenario with only a single vector measurement is more challenging than the multiple vector case, we focus on the former in this section. The main results can be extended to the case with multiple vector measurements in a straightforward manner.

A. Attitude Observer Using Integral Correction Term

Let us consider the observer design with a single complementary measurement (2). In the first observer design, we construct a dynamic extension – following the PEBO methodology [14] – in order to

[3]We refer to that the set of initial attitudes making the system lose distinguishability has zero Lebesgue measure in the entire state space.

reformulate attitude estimation as an on-line consistent parameter identification problem. By adding an elaborated construction of "integral"-type correction term, we are able to achieve asymptotic stability of the observer.

Proposition 2: For the system (1) with the complementary output (2), we assume that all signals are piece-wisely continuous and the reference satisfies the distinguishability condition, i.e.,

$$\exists t_1, t_2 > 0, \quad |g(t_1) \times g(t_2)| > 0, \tag{18}$$

with a *known* bound $T > 0$ on the distinguishability interval.[4] The attitude observer

$$\dot{Q} = Q\omega_\times \tag{19}$$

with $Q(0) \in SO(3)$ and

$$\dot{\hat{Q}}_c = \eta_\times \hat{Q}_c, \quad \hat{R} = \hat{Q}_c^\top Q \tag{20}$$

with

$$\begin{aligned} \eta &= \gamma_{\mathtt{P}}(\hat{Q}_c g) \times (Q y_{\mathtt{B}}) + \gamma_{\mathtt{I}} \xi \\ \xi &= 2\mathrm{vex}\big(\mathrm{skew}(A\hat{Q}_c^\top)\big) \\ \dot{A} &= \begin{cases} Q y_{\mathtt{B}} g^\top, & t \in [0, T) \\ 0_{3\times 3}, & t \geq T. \end{cases} \end{aligned}$$

with the gains $\gamma_{\mathtt{P}}, \gamma_{\mathtt{I}} > 0$ and $A(0) = 0_{3\times 3}$, guarantees $\hat{R}(t) \in SO(3)$ for all $t \geq 0$ and the convergence

$$\lim_{t\to\infty} \|\hat{R}(t) - R(t)\| = 0 \tag{21}$$

almost globally.

Proof: Let us consider the dynamic extension $\dot{Q} = Q\omega_\times$ for the initial condition $Q(0) \in SO(3)$. By defining a variable $E(R, Q) = QR^\top$ – which also lives in $SO(3)$ – we have

$$\dot{E} = \dot{Q}R^\top - QR^\top \dot{R} R^\top = 0.$$

Therefore, there exists a constant matrix $Q_c \in SO(3)$ such that

$$Q(t)R^\top(t) = Q_c, \quad \forall t \in [0, +\infty). \tag{22}$$

Note that $Q(t)$ is an available signal by construction, and Q_c is unknown. Invoking (22) and the full-rankness of Q, the estimation of R is equivalent to the one of Q_c.

Based on the above idea, we construct the following auxiliary system

$$\Sigma_c : \begin{cases} \dot{Q}_c = Q_c(\omega_c)_\times \\ y_c = Q_c b_c, \end{cases} \tag{23}$$

in which $Q_c \in SO(3)$ is constant thus $\omega_c = 0_3$, the output

$$y_c(t) := Q(t) y_{\mathtt{B}}(t),$$

and the "body-fixed coordinate" reference $b_c := g$. It is clear that the system Σ_c is exactly in the same form as the kinematic model with a compatible measurement (1) and (3).

We now define the estimation error of $\tilde{Q}_c := \hat{Q}_c Q_c^\top$, the dynamics of which is given by

$$\dot{\tilde{Q}}_c = \dot{\hat{Q}}_c Q_c^\top - \hat{Q}_c Q_c^\top \dot{Q}_c Q_c^\top = \eta_\times \tilde{Q}_c. \tag{24}$$

The term η satisfies

$$\begin{aligned} \eta_\times &= \gamma_{\mathtt{P}}[(\hat{Q}_c g) \times (Q y_{\mathtt{B}})]_\times + \gamma_{\mathtt{I}} \xi_\times \\ &= \gamma_{\mathtt{P}}\big[Q y_{\mathtt{B}} (\hat{Q}_c g)^\top - \hat{Q}_c g (Q y_{\mathtt{B}})^\top\big] + \gamma_{\mathtt{I}} \xi_\times \\ &= \gamma_{\mathtt{P}}\big(y_c y_c^\top \tilde{Q}_c^\top - \tilde{Q}_c y_c y_c^\top\big) + \gamma_{\mathtt{I}} \xi_\times \end{aligned} \tag{25}$$

[4]Namely, there exists a known $T > 0$ such that $0 \leq t_1 < t_2 \leq T$.

in which for $t \in [0, T]$

$$\begin{aligned} \xi_\times &= \int_0^t \Big[Q(s) y_{\mathtt{B}}(s)\big(\hat{Q}_c(t) g(s)\big)^\top - \hat{Q}_c(t) g(s)\big(Q(s) y_{\mathtt{B}}(s)\big)^\top\Big] ds \\ &= 2\mathrm{skew}\left(\int_0^t y_c(s) y_c^\top(s) ds \cdot \tilde{Q}_c^\top\right), \end{aligned}$$

and for $t > T$ we have $\xi(t) = \xi(T)$.

Consider the candidate Lyapunov function $V(\tilde{Q}_c) = 3 - \mathrm{tr}(\tilde{Q}_c)$, which has its minimal value zero has at $\tilde{Q}_c = I_3$. It yields

$$\begin{aligned} \dot{V} &= -\mathrm{tr}(\eta_\times \tilde{Q}_c) \\ &= -\gamma_{\mathtt{P}} \mathrm{tr}\Big(y_c y_c^\top - \tilde{Q}_c y_c y_c^\top \tilde{Q}_c\Big) \\ &\quad - \gamma_{\mathtt{I}} \int_0^t \mathrm{tr}\Big(y_c(s) y_c^\top(s) - \tilde{Q}_c y_c(s) y_c^\top(s) \tilde{Q}_c\Big) ds \\ &= -\gamma_{\mathtt{P}} y_c^\top (I - \tilde{Q}_c^2) y_c - \gamma_{\mathtt{I}} \int_0^t y_c^\top(s)(I - \tilde{Q}_c^2) y_c(s) ds \\ &= -2\mathrm{vex}\big(\mathrm{skew}(\tilde{Q}_c)\big)^\top \Gamma \mathrm{vex}\big(\mathrm{skew}(\tilde{Q}_c)\big) \\ &\leq -\lambda_{\mathtt{min}}(\Gamma) \|\mathrm{skew}(\tilde{Q}_c)\|^2, \end{aligned}$$

where in the fourth equation we have used $2|v|^2 = \|v_\times\|^2$ for any $v \in \mathbb{R}^3$, with the definition of Γ as

$$\Gamma = \Gamma_{\mathtt{P}} + \Gamma_{\mathtt{I}} \tag{26}$$

with

$$\begin{aligned} \Gamma_{\mathtt{P}}(t) &:= \gamma_{\mathtt{P}}(I - y_c(t) y_c^\top(t)) \\ \Gamma_{\mathtt{I}}(t) &:= \begin{cases} \gamma_{\mathtt{I}} \displaystyle\int_0^t \Big(I - y_c(s) y_c^\top(s)\Big) ds, & t \in [0, T] \\ \Gamma_{\mathtt{I}}(T), & t > T. \end{cases} \end{aligned}$$

Let us study the property of the matrix $\Gamma \in \mathbb{R}^{3\times 3}$. From the assumption $|g(t_1) \times g_2(t_2)| > 0$ for some $t_1, t_2 \leq T$, we have

$$\begin{aligned} |y_c(t_1) \times y_c(t_2)| &= |(Q_c g(t_1))_\times (Q_c g(t_2))| \\ &= |Q_c g(t_1) \times g(t_2)| \\ &> 0. \end{aligned}$$

It implies that

$$2I - y_c(t_1) y_c^\top(t_1) - y_c(t_2) y_c^\top(t_2) > 0, \tag{27}$$

in which we have used the fact that for $a, b \in \mathbb{S}^2$, $|a \times b| > 0$ implies the positiveness of $(2I - aa^\top - bb^\top)$; see [20, Lemma A.2] for its proof. On the other hand, in terms of the continuity of y_c and (27), we have

$$\int_{t_1}^{t_1+\varepsilon} I - y_c(s) y_c^\top(s) ds + \int_{t_2}^{t_2+\varepsilon} I - y_c(s) y_c^\top(s) ds > 0$$

for some sufficiently small $\varepsilon > 0$, and thus

$$\int_0^T I - y_c(s) y_c^\top(s) ds > 0 \quad \Longrightarrow \quad \lambda_{\mathtt{min}}(\Gamma(t)) > 0, \ \forall t \geq T.$$

From $\dot{V} \leq -\lambda_{\mathtt{min}}(\Gamma) \|\mathrm{skew}(\tilde{Q}_c(s))\|^2$, we get that $0 \leq V(\tilde{Q}_c(t)) \leq V(\tilde{Q}_c(0))$ and

$$V(\tilde{Q}_c(t)) - V(\tilde{Q}_c(0)) \leq -\int_0^t \lambda_{\mathtt{min}}(\Gamma(s)) \|\mathrm{skew}(\tilde{Q}_c(s))\|^2 ds.$$

By taking $t \to \infty$, in terms of Barbalat's lemma and the boundedness of the time derivative of $\tilde{Q}_c$, we have $\|\mathrm{skew}(\tilde{Q}_c)\| \to 0$. The set $\{\tilde{Q}_c \in SO(3) : \|\mathrm{skew}(\tilde{Q}_c)\| = 0\}$ only contains the stable equilibrium I_3 and the unstable equilibria $U^\top \mathrm{diag}(1, -1, 1)U$. For the latter case, the Lyapunov $V(\tilde{Q}_c)$ equals to its maximal value, thus having zero Lebesgue measure. Therefore, the dynamics (24) is almost globally asymptotically stable on $SO(3)$. Invoking the algebraic relation $R = Q_c^\top Q$, we complete the proof. $\square$

Remark 2: In the above attitude observer design, the error term η contains two parts

$$\eta = \underbrace{\gamma_{\mathtt{P}}(\hat{Q}_c g) \times (Qy)}_{\text{current}} + \underbrace{\gamma_{\mathtt{I}}\xi}_{\text{historical}},$$

which may be viewed as an observer design using a "proportional + integral"-type error term. The first term only utilizes the current information, making it behave as an on-line design. The second "integral" term, which may be written as

$$\xi(t) = \begin{cases} \int_0^t [\hat{Q}_c(t)g(s)] \times [Q(s)y_{\mathtt{I}}(s)]ds, & t \in [0, T] \\ \xi(T), & t \geq T \end{cases}$$

enables to achieve asymptotic convergence of the estimation error under the extremely weak identifiability condition (18). The gain parameters $\gamma_{\mathtt{P}}$ and $\gamma_{\mathtt{I}}$ can be used as the weights on how we trust the current and historical data.

Remark 3: The bound $T > 0$ is used in the dynamics of the variable A in order to be able to guarantee its boundedness. Indeed, the bound T is not necessarily known apriori, since the distinguishability condition (18) is an easily-checkable condition on measured quantities. The proposed scheme may be modified as an *adaptive* design in which such a condition is checked online continuously, and the dynamics of A simply changes until the condition holds. It is also natural to replace the condition (18) by $|g(t_1) \times g(t_2)| > \delta$ for some $\delta > 0$, to deal with sensor noise.

Remark 4: As is shown above, the "integral" term only accumulates information in the interval $[0, T]$, which, however, does not have the sort of "fading memory" property on past measurements. As long as the excitation condition, which is easily monitored on-line, the observer performance can be improved considering the moving interval $[t - T, t]$ rather than $[0, T]$ in Proposition 2.

B. Attitude Observer Using Virtual Vectors

In this subsection, we provide an alternative observer design, which does not need the information of T. The basic idea is to generate a new "*virtual*" vector measurement

$$y_v = Q_c b_v \tag{28}$$

from the real measurement (23), such that

$$b_v \times b_c \neq 0 \tag{29}$$

uniformly after some moment with $b_c = g$. Then, it becomes the well-studied attitude observer design problem with (not less than) two non-collinear vectors, which has been well addressed in the literature.

Proposition 3: For the system (1) with the output (2), we assume that all signals are continuous and satisfy $\exists t_i > 0 \ (i = 1, 2, 3)$

$$\det\Big(\begin{bmatrix} g(t_1) & g(t_2) & g(t_3) \end{bmatrix}\Big) \neq 0 \tag{30}$$

Consider the dynamic extension (19) and the LTV filter

$$\begin{aligned} \dot{Z} &= \gamma_z g(y_{\mathtt{B}}^\top Q^\top - g^\top Z) \\ \dot{\Omega} &= -\gamma_z g g^\top \Omega, \ \Omega(0) = I_3 \\ \dot{P} &= \gamma(I_3 - \Omega)^\top [Z - \Omega Z_0 - (I_3 - \Omega)P] \end{aligned} \tag{31}$$

with $\gamma, \gamma_z > 0$, $Z_0 := Z(0)$, and the filtering outputs

$$b_v = Ug, \quad y_v = P^\top U g$$

in which $U := \mathrm{diag}(\mathcal{J}, 1)\mathrm{diag}(1, \mathcal{J})$. Then, the observer (20) with

$$\eta = \gamma_c(\hat{Q}_c g) \times (Q y_{\mathtt{B}}) + \gamma_v(\hat{Q}_c b_v) \times y_v \tag{32}$$

and the gains $\gamma_c, \gamma_v > 0$ guarantees $\hat{R}(t) \in SO(3)$ for all $t \in [0, \infty)$ and the convergence (21) almost globally.

Proof: First, let us study the property of the LTV filter (31). Note that U can be decomposed as the product of three basic (element) rotations $U = R_x(\theta_1)R_y(\theta_2)R_z(\theta_3)$ with $\theta_1 = \theta_3 = \frac{\pi}{2}$ and $\theta_2 = 0$. Hence, $g \times (Ug) \neq 0$, verifying the equation (29). Then, we need to verify (28) in an asymptotic sense.

According to the proof in Proposition 2, we may reformulate the estimation of R as the one of Q_c. From the dynamics of $\dot{Z}$, one has

$$\begin{aligned} \frac{d}{dt}(Z - Q_c^\top) &= \gamma_z g(y_c^\top - g^\top Z) \\ &= -\gamma_z g g^\top (Z - Q_c^\top), \end{aligned}$$

and thus

$$Z - Q_c^\top = \Omega(Z_0 - Q_c^\top) \implies Z - \Omega Z_0 = (I - \Omega)Q_c^\top.$$

Then, it yields

$$\frac{d}{dt}(P - Q_c^\top) = -\gamma\phi\phi^\top(P - Q_c^\top). \tag{33}$$

with $\phi := I_3 - \Omega^\top$.

From the condition (30), for *any* non-zero $x \in \mathbb{R}^3$ it may always be represented as

$$x = c_1 g(t_1) + c_2 g(t_2) + c_3 g(t_3), \tag{34}$$

with at least one of $c_i \ (i = 1, 2, 3)$ non-zero. Hence

$$\begin{aligned} x^\top \sum_{i=1}^3 g(t_i)g(t_i)^\top x > 0 &\implies \sum_{i=1}^3 g(t_i)g(t_i)^\top > 0 \\ &\implies \int_{t_i}^{t_i+\varepsilon} \sum_{i=1}^3 g(s)g(s)^\top ds > 0 \\ &\implies \int_0^T g(s)g(s)^\top ds > 0. \end{aligned}$$

for some $T > t_i$ and some sufficiently small $\varepsilon > 0$. It implies that the vector signal g is intervally excited. As a result, the matrix $\phi = I_3 - \Omega^\top$ is PE [22, Proposition 2]. Following [18, Thm. 2.5.1], we are able to derive

$$\lim_{t\to\infty} \|P - Q_c^\top\| = 0 \quad \implies \quad \lim_{t\to\infty} |P^\top U g - Q_c g| = 0 \tag{35}$$

exponentially fast. Hence, the algebraic equation (28) is guaranteed asymptotically, i.e.,

$$y_v = Q_c b_v + \varepsilon_t$$

with an exponentially decaying term ε_t.

The last step is to study the convergence of the estimation error of Q_c, which is defined as $\tilde{Q}_c := Q_c \hat{Q}_c^\top$. The observer $\dot{\hat{Q}}_c = \eta_\times \hat{Q}_c$ with the output error term (32) may be written as the "standard" compatible observer

$$\dot{\hat{Q}}_c = \hat{Q}_c \Big[\hat{Q}_c\Big(\gamma_c(\hat{Q}_c b_c) \times b_c + \gamma_v(\hat{Q}_c b_v + \varepsilon_t) \times b_v\Big)\Big]_\times$$

for the auxiliary dynamics (23) with $\omega_c = 0_3$. If the term ε_t was zero, in terms of [20, Thm 4.3] and the uniform non-collinear relation $b_c \times b_v \neq 0$, we would obtain $\hat{Q}_c \to Q_c$ as $t \to \infty$ almost globally. However, the term ε_t is exponentially decaying to zero, and we may follow the perturbation analysis in [23, Proposition 6] using a time-varying Lyapunov function to obtain the almost global asymptotic stability. We omit the details here. Invoking the algebraic relation (22), we complete the proof. $\square$

Remark 5: The condition (30) with three non-collinear moments is slightly stronger than the one (18). However, it removes the necessity of having a priori known bound $T > 0$ for the observer design in

Proposition 3. The key idea to construct a new "virtual" non-collinear reference vector b_v is similar to some recent results on generation of PE regressors from those only satisfying IE [6], [22], [23]. We refer the reader to the monograph [5] for a discussion of excitation in observer design.

Remark 6: For the case with a single compatible measurement (3), we may still get the auxiliary model (23) by designing the dynamic extension $\dot{Q} = Q\omega_\times$, but with the new definitions of $y_c := y_{\mathtt{I}}$ and $g_c := Qb$. Then, the above two designs are capable to solve the problem with slight modifications accordingly.

V. Discussions

In this section, we show some practical issues in attitude estimation – intermittent and delayed measurements [1], [3] – can be easily (and even trivially) tackled by the proposed methodology.

Remark 7: (*Delayed measurement*) Time delay in attitude estimation is generally unavoidable, which is usually caused by low-quality data sampling and poor sensors [1]. A common scenario is that a known delay τ appears in the single vector measurement, i.e. $y(t) = y_{\mathtt{I}}(t-\tau) = R^\top(t-\tau)g(t-\tau)$, in which $\tau(t)$ may be constant or time-varying. After designing the dynamic extension (19), the delayed output can be rewritten as $y(t) = Q^\top(t-\tau)Q_c g(t-\tau)$. Then, we are still able to get the auxiliary model (23) with

$$y_c := Q(t-\tau)y(t), \quad b_c := g(t-\tau). \tag{36}$$

Since Q_c is a *constant* matrix on the special orthogonal group, the observers in Propositions 2-3 can provide asymptotically convergent attitude estimation by modifying the "reference vector" b_c and the vector y_c as (36).

Remark 8: (*Intermittent measurement*) Some types of sensors only provide intermittent measurement $y_{\mathtt{I}}$ at some instants of time t_k ($k \in \mathbb{N}_+$). Let sequence $\{t_k\}_{k\in\mathbb{N}_+}$ be strictly increasing, and $|t_1|$ and $|t_{k+1} - t_k|$ ($k \geq 1$) are upper and lower bounded by two positive constants. In the proposed attitude PEBO framework, we have translated the estimation of the variable $R(t)$ into that of constant Q_c, and thus the intermittent measurement does not bring any difficulty in observer design. By defining the vectors

$$y_c(t) := Q(t_i)y(t_i), \quad b_c(t) := g(t_i), \quad \forall t \in [t_i, t_{i+1}]. \tag{37}$$

Again, we get the auxiliary model (23) with the modified reference vectors in (37). The proposed two *continuous-time* observers can solve attitude determination.

Remark 9: (*Robustness*) In practice, the measurement vector is perturbed by sensor noise, i.e., $y_{\mathtt{M}} = y_{\mathtt{B}} + n_y$ with a bounded term n_y, and $y_{\mathtt{M}}, y_{\mathtt{B}} \in \mathbb{S}^2$. Then, the correction term η of the observer in Proposition 2 becomes $\eta_{\mathtt{M}} := \eta + \Delta_\eta$, in which η is the nominal part defined in Proposition 2 and Δ_η is the additive term stemmed from the noise term n_y. Since $\hat{Q}_c, Q, g$ and $y_{\mathtt{B}}$ all live in some compact sets, as well as the variable A being the integral over a *finite* interval $[0,T]$, there exists a constant $k>0$ such that $|\Delta_\eta(t)| \leq k\|n_y\|_\infty$. For this case, the time derivative of the Lyapunov function becomes $\dot{V} \leq -\lambda_{\min}(\Gamma)\|\text{skew}(\tilde{Q}_c)\|^2 + |\text{tr}((\Delta_\eta)_\times \tilde{Q}_c)|$. Hence, we are able to establish the robustness of the observer in the bounded-input-bounded-output (BIBO) sense. The is also verified via noisy simulations in the coming section.

VI. Simulations

Example 1: We consider a single time-varying inertial vector

$$g(t) = \begin{cases} e_1, & t \in [0,5)\text{s} \\ e_3, & t \geq 5\text{s}, \end{cases} \tag{38}$$

in which e_i represents the i-th standard Euclidean basis in $\mathbb{R}^3$. Clearly, it satisfies the sufficient excitation condition (18), but not for the persistently non-constant reference vector assumption in many works [2]. The attitude of the rigid-body starts from the initial condition $R(0) = \text{diag}(-1,-1,1)$ under the rotational velocity $\omega = [0.23\ -0.5\ 0.15]^\top$. We added noise to both the angular velocity readings and the vector measurements.

First, we evaluate the performance of the scheme in Proposition 2. The observer is initialized from $Q(0) = \hat{Q}_c(0) = I_3$, with the gains $\gamma_{\mathtt{P}} = 3$ and $\gamma_{\mathtt{I}} = 1$. It corresponds to the initial yaw, pitch and roll estimates all being 0°. The results of simulations are shown in Fig. 1 in the form of Euler angles, and also see the norm of the estimation error $|\tilde{R}|_I$ in Fig. 5, which is drawn in a logarithmic scale for the y-axis. During $[0,5]$ s, the error $\tilde{R}$ is converging to some non-zero constant under a constant vector measurement. This is because a single vector output makes two of three Euler angles partially observable [11]. After 5s the model satisfies the distinguishability, and then all Euler angles converge to their true values. Note that the proposed scheme is robust *vis-à-vis* measurement noise. Then, we test the second observer design in Proposition 3, with the simulation results presented in the same figure. Though the reference vector g in (38) does not satisfy the sufficient condition (30), it is interesting to observe that all the Euler angles converge to zero asymptotically for the same reference vector g. This implies that the condition (30) is not necessary for the convergence of the second observer design. At the end, let us compare the proposed schemes to the complementary attitude observer in [20], whose convergence is guaranteed by a persistent excitation condition. Clearly, this is not satisfied by the inertial reference vector in (38). We show the simulation results for in Fig. 5. As expected, the estimate $\hat{R}$ from the observer in [20] fails to converge to its true values. Besides, we note that the first design in Proposition 2 is less sensitive to measurement noise.

Example 2: In the second example, we consider the problem set in [8], i.e., the vector being the highly time-varying acceleration of a helicopter. To be precise, a remotely controlled helicopter is equipped with accelerometers to detect the acceleration $y_{\mathtt{B}} = {}^{\mathtt{B}}a$ in the body-fixed frame, and the corresponding inertial acceleration ${}^{\mathtt{I}}a$ can be calculated from the GPS velocity v using the relation $\dot{v} = {}^{\mathtt{I}}a$. Clearly, the relation ${}^{\mathtt{B}}a = R^\top {}^{\mathtt{I}}a$ holds true. Note that the "dirty derivative" is usually provided by the filtered approximation $H_1(p)[v] = \frac{\alpha p}{\alpha + p}[v]$ with the differential operator $p := \frac{d}{dt}$ and selecting $\alpha > 0$ some large parameter. It is widely known that such an operation yields the phase shift. In our problem set, in terms of Proposition 3, we have the identity $Q(t){}^{\mathtt{B}}a(t) = Q_c\dot{v}(t)$. Applying the filter $H_2(p)[\cdot] = \frac{\alpha}{\alpha+p}[\cdot]$ to the both sides – thanks to Q_c being constant – we have

$$H_2\big[Q{}^{\mathtt{B}}a\big] \;=\; Q_c H_1[v]. \tag{39}$$

By injecting the above into the proposed observer, we are able to overcome the phase-shifting issue caused by the filter $H_1(p)$, and there is no need to require the parameter α large to approximate the time derivative. Instead, a small α makes $H_2(p)$ behave as a low-pass filter, which is capable to attenuate noise significantly.

We consider the helicopter trajectory as shown in Fig. 4. To make the simulation more realistic, the gyros and the accelerometer provide data at 100 Hz, and the GPS receiver is at 10 Hz – all with high-frequency noise. Besides, in order to evaluate robustness of the proposed design, we consider acceleration bias from the senor, but do not make any compensation. The first observer design was implemented at 1000 Hz using the solver "ODE 4 (Runge-Kutta)" in Matlab/Simulink™, with $\gamma_{\mathtt{I}} = 1, \gamma_{\mathtt{P}} = 5, \alpha = 1$, and the initial conditions $Q(0) = \hat{Q}_c(0) = I_3$. The simulation results are given in Fig. 3, which illustrates its good robustness, though it brings additional errors in the steady-state stage. We compare it to the design

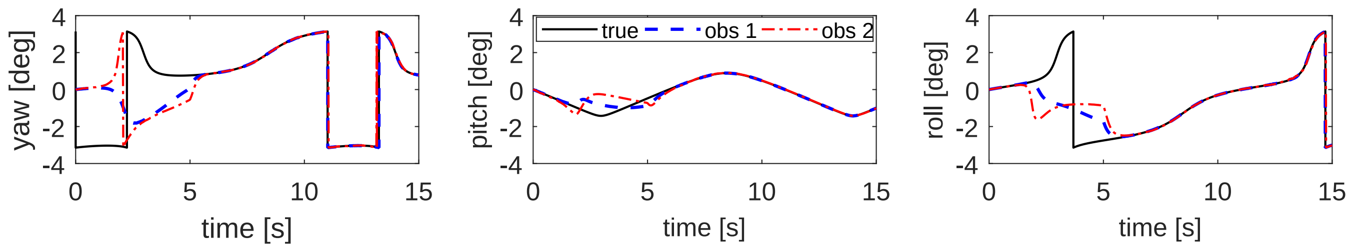

Fig. 1: Performance of the attitude observer in Proposition 2 (obs 1) and 3 (obs 2) with Euler angles (Example 1)

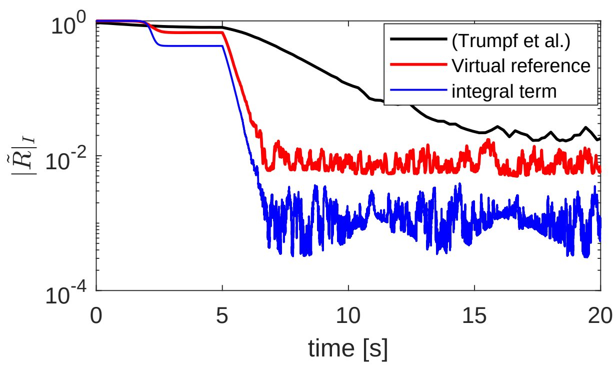

Fig. 2: Comparison of the norms of estimation errors $|\tilde{R}|_I$ among the proposed observers and [20] (Example 1)

in [8] using $H_1(p)$ with $\alpha = 8$ approximate the differentiator. The phase lag from the filter leads to the offset in estimates observed in Fig. 3. This effect can be reduced by increasing α, but at the expense of higher sensitivity to noise. As discussed above, using (39) the proposed design does not suffer from this issue.

VII. Concluding Remarks

In this paper, we studied the observability and observer design for the attitude estimation problem with vectorial measurements. By translating the observation problem into one of on-line parameter identification, we provided the necessary and sufficient condition to the distinguishability for the dynamical model on $SO(3)$, which is complementary to the existing necessary conditions in the literature. As is shown later, though the resulting distinguishability condition is quite weak, we are still able to use it to derive a continuous-time attitude observer with almost global asymptotic stability guaranteed for the single vector case. Finally, simulation results demonstrated accurate estimation performance in the presence of measurement noise.

Appendix

Proof of Corollary 1

Proof: Compared to Proposition 1, the only difference relies on the second term, which corresponds to the third case in the proof of Proposition 1.

The modified condition assumes the existence of two indices $i \in \ell_1, j \in \ell_2$ such that $|g_i(t_1) \times [\Phi(0,t_2)b_j(t_2)]| > 0$. Since $\Phi(0,t_2)b_j(t_2) \in \mathbb{S}^2$, we have

$$\Big| g_i(t_1) \times \big[R_0 \Phi(0,t_2) b_j(t_2) \big] \Big| > 0 \tag{40}$$

if R_0 is not in an inadmissible initial set

$$\mathcal{E} := \{ R_0 \in SO(3) \mid R_0 \mathbf{v} = \pm \mathbf{w} \} \tag{41}$$

with $\mathbf{v} := \Phi(0,t_2)b_j(t_2)$ and $\mathbf{w} = g_i(t_1)$ both in $\mathbb{S}^2$. For a given rotational velocity ω and the references b_j, g_i, two of three Euler angles of the initial rotation matrix R_0 is uniquely determined by the equality in (41). Hence, the inadmissible initial set $\mathcal{E}$ has zero Lebesgue measure in the group $SO(3)$. As a result, we guarantee the condition (15) from the modified assumption almost surely. $\square$

Proof of Corollary 2

Proof: Since those two terms in (17) are non-negative, is should satisfy at least one of the following cases:

$$\text{(i)} \qquad \lambda_2 \left(\sum_{i \in \ell_1} \int_0^T g_i(s) g_i^\top(s) ds \right) > 0 \tag{42}$$

$$\text{(ii)} \qquad \left\| \int_0^T \sum_{j \in \ell_2} \left(\omega_\times b_j(s) + \frac{d}{ds} b_j(s) \right) ds \right\| > 0. \tag{43}$$

In the case (i), note that for a single vector $g_i(t) \in \mathbb{S}^2$, the matrix $g_i(t) g_i^\top(t)$ has rank one at any instance $t \geq 0$. Hence, a necessary condition to (42) is the existence of $t_1, t_2 \geq 0$ and $i, l \in \ell_1$ (i and l may be the same) such that

$$\lambda_2 \Big(g_i(t_1) g_i^\top(t_1) + g_l(t_2) g_l^\top(t_2) \Big) > 0, \tag{44}$$

which implies $\sum_{i,l \in \ell_1} |g_i(t_1) \times g_l(t_2)| > 0$, thus guaranteeing the condition (6).

For the case (ii), we consider (piecewisely) smooth outputs $y_{\mathtt{I},i}$ with $i \in \ell_2$. Its dynamics is given by

$$\dot{y}_{\mathtt{I},i} \;\; = \;\; R(\omega_\times b_i + \dot{b}_i). \tag{45}$$

A necessary condition to (43) is that there exist j and a moment $t_1 > 0$ such that

$$\omega(t_1)_\times b_j(t_1) + \dot{b}_j(t_1) \neq 0 \quad \implies \quad \dot{y}_{\mathtt{I},j}(t_1) \neq 0.$$

Let us select a sufficiently small $\Delta t > 0$, and define $t_2 := t_1 + \Delta t$. It yields

$$y_{\mathtt{I},j}(t_2) \;\; = \;\; y_{\mathtt{I},j}(t_1) + \dot{y}_{\mathtt{I},j}(t_1)\Delta t + o(\Delta t^2)$$

in which $o(\Delta t^2)$ represents the high-order remainder term, with the *constraint* $y_{\mathtt{I},j}(t_2) \in \mathbb{S}^2$. Now, we show $y_{\mathtt{I},j}(t_1) \times \frac{d}{dt} y_{\mathtt{I},j}(t_1) \neq 0$ by contradiction. If this cross product is equal to zero, invoking $\frac{d}{dt} y_{\mathtt{I},j}(t_1) \neq 0$, we have $\dot{y}_{\mathtt{I},j}(t_1) = a y_{\mathtt{I},j}(t_1)$ for some non-zero $a \in \mathbb{R}$. Then, we have

$$\begin{aligned} |y_{\mathtt{I},j}(t_2)| \;\; &= \;\; \Big| (1 + a\Delta t) y_{\mathtt{I},j}(t_2) + o(\Delta t^2) \Big| \\ &= \;\; |1 + a\Delta t| + o(\Delta t^2), \end{aligned}$$

which contradicts with the fact $y_{\mathtt{I},j}(t_2) \in \mathbb{S}^2$. As a consequence, we obtain that $y_{\mathtt{I},j}(t_1) \times y_{\mathtt{I},j}(t_2) \neq 0$. Invoking the equivalence (13), we have

$$(43) \quad \implies \quad \sum_{j,k \in \ell_2} \Big| b_j(t_1)_\times \Phi(t_1,t_2) b_k(t_2) \Big| > 0,$$

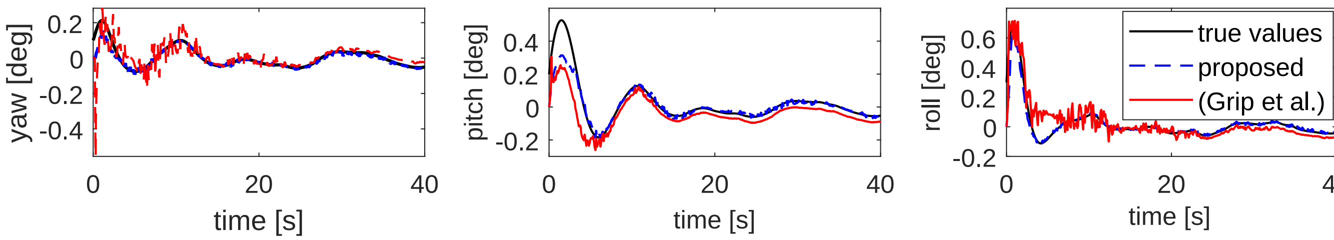

Fig. 3: Simulation results for helicopter attitude estimation using the observers in Proposition 2 and in [8] (Example 2)

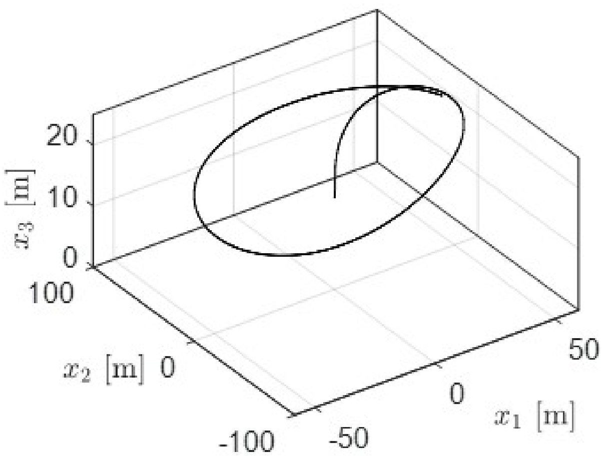

Fig. 4: The trajectory of the helicopter (Example 2)

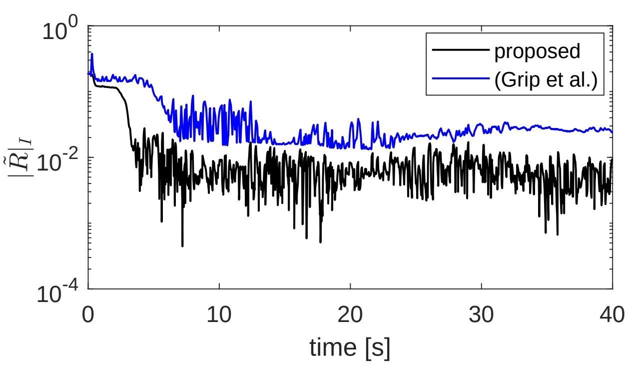

Fig. 5: Comparison of the norms of estimation errors $|\tilde{R}|_I$ among the proposed observer and the design in [8] (Example 2)

thus verifying (6). It completes the proof. $\square$

References

[1] S. Bahrami and M. Namvar. Global attitude estimation using single delayed vector measurement and biased gyro. *Automatica*, 75:88–95, 2017.

[2] P. Batista, C. Silvestre, and P. Oliveira. A GES attitude observer with single vector observations. *Automatica*, 48(2):388–395, 2012.

[3] S. Berkane and A. Tayebi. Attitude estimation with intermittent measurements. *Automatica*, 105:415–421, 2019.

[4] P. Bernard. *Observer Design for Nonlinear Systems*. Lecture Notes in Control and Information Sciences. Springer International Publishing, 2019.

[5] G. Besançon. *Nonlinear Observers and Applications*, volume 363. Springer, 2007.

[6] A. Bobtsov, B. Yi, R. Ortega, and A. Astolfi. Generation of new exciting regressors for consistent on-line estimation of unknown constant parameters. *IEEE Trans. Autom. Control*, 2022.

[7] J. L. Crassidis, F. L. Markley, and Y. Cheng. Survey of nonlinear attitude estimation methods. *J. Guid. Control Dyn.*, 30(1):12–28, 2007.

[8] H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi. Attitude estimation using biased gyro and vector measurements with time-varying reference vectors. *IEEE Trans. Autom. Control*, 57(5):1332–1338, 2011.

[9] J. C. Kinsey and L. L. Whitcomb. Adaptive identification on the group of rigid-body rotations and its application to underwater vehicle navigation. *IEEE Trans. Robot.*, 23(1):124–136, 2007.

[10] R. Mahony, T. Hamel, and J.-M. Pflimlin. Nonlinear complementary filters on the special orthogonal group. *IEEE Trans. Autom. Control*, 53(5):1203–1218, 2008.

[11] P. Martin and I. Sarras. Partial attitude estimation from a single measurement vector. In *Proc. IEEE Conf. Control Tech. Appl.*, pages 1325–1331, 2018.

[12] M. Namvar and F. Safaei. Adaptive compensation of gyro bias in rigid-body attitude estimation using a single vector measurement. *IEEE Trans. Autom. Control*, 58(7):1816–1822, 2013.

[13] R. Ortega, A. Bobtsov, N. Nikolaev, J. Schiffer, and D. Dochain. Generalized parameter estimation-based observers: Application to power systems and chemical–biological reactors. *Automatica*, 129, 2021. Art. no. 109635.

[14] R. Ortega, A. Bobtsov, A. Pyrkin, and S. Aranovskiy. A parameter estimation approach to state observation of nonlinear systems. *Syst. Control Lett.*, 85:84–94, 2015.

[15] R. Ortega, V. Nikiforov, and D. Gerasimov. On modified parameter estimators for identification and adaptive control: A unified framework and some new schemes. *Annu. Rev. Control*, 50:278–293, 2020.

[16] J. Reis, P. Batista, P. Oliveira, and C. Silvestre. Attitude, body-fixed earth rotation rate, and sensor bias estimation using single observations of direction of gravitational field. *Automatica*, 125, 2021. Art. no. 109475.

[17] S. Salcudean. A globally convergent angular velocity observer for rigid body motion. *IEEE Trans. Autom. Control*, 36(12):1493–1497, 1991.

[18] S. Sastry and M. Bodson. *Adaptive Control: Stability, Convergence And Robustness*. Courier Corporation, 2011.

[19] M. D. Shuster and S. D. Oh. Three-axis attitude determination from vector observations. *J. Guid. Control Dyn.*, 4(1):70–77, 1981.

[20] J. Trumpf, R. Mahony, T. Hamel, and C. Lageman. Analysis of non-linear attitude observers for time-varying reference measurements. *IEEE Trans. Autom. Control*, 57(11):2789–2800, 2012.

[21] G. Wahba. A least squares estimate of satellite attitude. *SIAM Review*, 7(3):409–409, 1965.

[22] L. Wang, R. Ortega, A. Bobtsov, J. G. Romero, and B. Yi. Identifiability implies robust, globally exponentially convergent on-line parameter estimation: Application to model reference adaptive control. *arXiv preprint arXiv:2108.08436*, 2021.

[23] B. Yi, C. Jin, and I. R. Manchester. Globally convergent visual-feature range estimation with biased inertial measurements. *arXiv preprint arXiv:2112.12325*, 2021.

[24] B. Yi, C. Jin, L. Wang, G. Shi, and I. R. Manchester. An almost globally convergent observer for visual SLAM without persistent excitation. In *60th Proc. IEEE Conf. Decis. Control*, pages 5441–5446, 2021.

[25] B. Yi, R. Ortega, and W. Zhang. On state observers for nonlinear systems: A new design and a unifying framework. *IEEE Trans. Autom. Control*, 64(3):1193–1200, 2018.

[26] D. E. Zlotnik and J. R. Forbes. Nonlinear estimator design on the special orthogonal group using vector measurements directly. *IEEE Trans. Autom. Control*, 62(1):149–160, 2016.