On the Observability of Attitude with Single Direction Measurements

Weixin Wang, Kanishke Gamagedara, and Taeyoung Lee

***Abstract*—The attitude of a rigid body evolves on the three-dimensional special orthogonal group, and it is often estimated by measuring reference directions, such as gravity or magnetic field, using onboard sensors. As a single direction measurement provides a two-dimensional constraint, it has been widely accepted that at least two non-parallel reference directions should be measured, or the reference direction should change over time, to determine the attitude completely. This paper uncovers an intriguing fact that the attitude can actually be estimated by using multiple measurements of a single, fixed reference direction, provided that the angular velocity and the direction measurements are resolved in appropriate frames, respectively. More specifically, after recognizing that the attitude uncertainties propagated over the left-trivialized stochastic kinematics are distinct from those over the right-trivialized one, stochastic attitude observability with single direction measurements is formulated by an information theoretic analysis. These are illustrated by experiments.**

I. Introduction

Rigid body attitude estimation using angular velocity and direction measurements is fundamental for numerous aerospace and robotics applications. A wide variety of attitude estimators have been developed, including the multiplicative extended Kalman filter (MEKF) [1], a deterministic attitude observer [2], an invariant extended Kalman filter on matrix Lie groups [3], and an attitude estimator based on the matrix Fisher distribution [4], [5]. It has been understood that at least two non-parallel reference directions are required to determine the attitude completely. For a single direction measurement, the rotation about the reference vector is unobservable, except in a few special cases, for example, when the single reference direction is time-varying in the inertial frame [6], [7], [8], and when the gyroscope can capture the rotation of the earth [9].

This paper presents attitude observability in stochastic formulations, where it is discovered that there are two additional cases in which attitude can be completely estimated using a single fixed reference direction (Table I). In particular, it is shown that the attitude uncertainties are propagated in different ways, depending on whether the angular velocity is resolved in the body-fixed frame or in the inertial frame. For the former, the direction of one-dimensional ambiguity caused by a single direction measurement remains fixed in the inertial frame, but for the latter, it is fixed in the body-fixed frame. This explains the fundamental reason why the attitude is unobservable with a single inertial reference direction and the angular velocity measured by a gyroscope: the direction of ambiguity caused by the inertial reference direction measurement remains unchanged by the angular velocity resolved in the body-fixed frame. This leads to two strategies to achieve full attitude observability with single direction measurements, namely utilizing the angular velocity resolved in the inertial frame such that the direction of ambiguity is rotated over propagation, or measuring a reference direction fixed to the body such that the next measurement can resolve the ambiguity.

W. Wang, K. Gamagedara, and T. Lee are with the Department of Mechanical and Aerospace Engineering, The George Washington University, Washington DC, USA. {wwang442,kanishkegb,tylee}@gwu.edu

This research has been supported in part by NSF under the grants CMMI-1760928 and CNS-1837382, and by AFOSR under the grant FA9550-18-1-0288.

TABLE I
ATTITUDE OBSERVABILITY WITH SINGLE DIRECTION MEASUREMENTS

ref. vec. \ ang. vel.	body-fixed frame	inertial frame
body-fixed frame	observable	unobservable
inertial frame	unobservable	observable

This discovery is more rigorously studied by introducing two formulations of stochastic attitude observability. For a given probability density function on $\mathsf{SO}(3)$, the attitude that minimizes the mean square error may not be unique [10], [11]. This gives a characterization of the attitude observability, since if multiple attitudes can minimize the mean square error, it indicates deficiency of information to distinguish them. Alternatively, the inverse of Fisher information gives a lower bound for the variance of all unbiased estimators, known as the Cramér–Rao bound, and the observability can be characterized by the positive-definiteness of the Fisher information matrix [12]. We adopt the method in [13] which generalizes the Fisher information to Riemannian manifold to calculate the Fisher information of the mean attitude of a matrix Fisher distribution. It is further shown that the two attitude observability criteria are consistent with each other, and they also agree with a classic deterministic observability analysis when only a single direction measurement is available. Finally, the presented two cases of observability are illustrated by experiments. An extended version of this paper with more detailed proofs and numerical examples is available in [14].

II. Mathematical Preliminary

A. Attitude Kinematics

We define an inertial frame $\mathcal{I} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, where $\mathbf{e}_i$ denotes the vector for the i-th axis of $\mathcal{I}$. Throughout this paper, we distinguish a vector from its coordinates resolved in a selected basis. For example, the coordinates of $\mathbf{e}_i$ in $\mathcal{I}$ is denoted by $e_i \in \mathbb{R}^3$, e.g., $e_1 = (1, 0, 0)$. Similarly, we define a body-fixed frame, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$. The attitude of the rigid body is the orientation of $\mathcal{B}$ relative to $\mathcal{I}$, and it can be defined by a rotation matrix $R \in \mathbb{R}^{3\times3}$ in the special orthogonal group

$$\mathsf{SO}(3) = \{R \in \mathbb{R}^{3\times3} \,|\, R^T R = I_{3\times3},\ \det[R] = 1\}.$$

The matrix R transforms the coordinates of a vector from $\mathcal{B}$ to $\mathcal{I}$.

Next, let $\mathbf{w}$ be the angular velocity vector of the rigid body, or equivalently, $\dot{\mathbf{b}}_i = \mathbf{w} \times \mathbf{b}_i$ for all $i \in \{1, 2, 3\}$. As the coordinates of $\mathbf{b}_i$ in $\mathcal{I}$ are Re_i, it implies $\dot{R}e_i = \omega \times Re_i$, where $\omega \in \mathbb{R}^3$ is the coordinates of $\mathbf{w}$ in $\mathcal{I}$. Thus,

$$\dot{R} = \hat{\omega} R, \tag{1}$$

where the hat map $\wedge : \mathbb{R}^3 \to \mathfrak{so}(3)$ is defined such that $\hat{x}y = x \times y$ for any $x, y \in \mathbb{R}^3$, and $\mathfrak{so}(3)$ denotes the Lie algebra composed of 3×3 skew-symmetric matrices, i.e., $\mathfrak{so}(3) = \{S \in \mathbb{R}^{3\times3} \,|\, S^T = -S\}$. Let $\Omega = R^T\omega \in \mathbb{R}^3$ be the coordinates of $\mathbf{w}$ in $\mathcal{B}$. As $\widehat{Rx} = R\hat{x}R^T$ for any $R \in \mathsf{SO}(3)$ and $x \in \mathbb{R}^3$, (1) can be rewritten as

$$\dot{R} = R\hat{\Omega}. \tag{2}$$

Both of (1) and (2) ensure that their solutions evolve on $\mathsf{SO}(3)$, or $\dot{R}$ belongs to the tangent space of $\mathsf{SO}(3)$ at R. Right-multiplying (1) with R^T, we obtain $\hat{\omega} = \dot{R}R^T$, which is referred to as *right-trivialization* of the tangent space. Similarly, $\hat{\Omega} = R^T\dot{R}$ is referred to as *left-trivialization*. Given $\Omega = R^T w$, (1) is equivalent to (2).

B. Matrix Fisher Distribution

The matrix Fisher distribution is an exponential density formulated for random matrices [15], [16], and its stochastic properties on $\mathsf{SO}(3)$ are presented in [4]. A random matrix $R \in \mathsf{SO}(3)$ is distributed according to the matrix Fisher distribution with the matrix parameter $F \in \mathbb{R}^{3\times 3}$, or $R \sim \mathcal{M}(F)$, if its probability density is given by

$$p(R) = \frac{1}{c(F)} \exp(\mathrm{tr}[F^T R]), \tag{3}$$

where $c(F) = \int_{R\in\mathsf{SO}(3)} \exp(\mathrm{tr}[F^T R])dR$ is the normalizing constant [4].

The role of F in specifying the shape and dispersion of the distribution can be described after decomposing it into the proper singular value decomposition (pSVD) [17] as follows:

$$F = USV^T, \tag{4}$$

where $U, V \in \mathsf{SO}(3)$ and $S = \mathrm{diag}[s_1, s_2, s_3] \in \mathbb{R}^{3\times 3}$ is a diagonal matrix with $s_1 \geq s_2 \geq |s_3| \geq 0$. This is a variation of the common singular value decomposition, defined to ensure $U, V \in \mathsf{SO}(3)$. These provide the following interpretation of $\mathcal{M}(F)$:

- $UV^T \in \mathsf{SO}(3)$: mean attitude in the sense of the least Frobenius mean square error, and it is denoted by $\mathrm{M}[F] \in \mathsf{SO}(3)$.
- $U \in \mathsf{SO}(3)$: the columns represent the coordinates of principal axes in the inertial frame.
- $V \in \mathsf{SO}(3)$: the columns represent the coordinates of principal axes in the body-fixed frame of the mean attitude.
- $s_1, s_2, s_3 \in \mathbb{R}^3$: degree of concentration; the rotation about the i-th principal axis is more concentrated as $s_j + s_k$ is increased, for $(i, j, k) \in \{(1, 2, 3), (2, 3, 1), (3, 1, 2)\}$.

The mean and principal axes are interpreted similarly as the Gaussian distribution, and $s_j + s_k$ can be interpreted in a vague sense as the inverse of variance along the i-th principal axis. More specifically, the rotations about principal axes are uncorrelated. The first and the last principal axes represent the most and the least uncertain directions, respectively.

The first moment of the matrix Fisher distribution is

$$\mathrm{E}[R] = UDV^T, \tag{5}$$

where $D = \mathrm{diag}[d_1, d_2, d_3]$ is given by

$$d_i = \frac{1}{c(S)} \frac{\partial c(S)}{\partial s_i}, \qquad i = 1, 2, 3. \tag{6}$$

The relationship between S and D is summarized as follows.

Lemma II.1. *Suppose a random rotation matrix $Q \in \mathsf{SO}(3)$ is distributed according to $Q \sim \mathcal{M}(S)$ for $S = \mathrm{diag}[s_1, s_2, s_3]$. Based on* (5)*, $\mathrm{E}[Q] = D = \mathrm{diag}[d_1, d_2, d_3]$ is diagonal. Then, the following properties hold:*

1) $s_i + s_j = 0$ *if and only if* $d_i + d_j = 0$*;*
2) $s_i = s_j = 0$ *if and only if* $d_i = d_j = 0$*;*
3) $d_i + d_j$ *is monotonically increasing with* $s_i + s_j$*,*

for any $(i, j, k) \in \{(1, 2, 3), (2, 3, 1), (3, 1, 2)\}$.

Proof. The proof uses the one dimensional integral formula of $c(S)$ given in [4], and it is available in [14]. □

According to Lemma II.1, S and D have a one-to-one correspondence, which is denoted by $\mathcal{E}[S] = D$ as given in (6). This implies F and $\mathrm{E}[R]$ carry the same information, and we can construct a matrix Fisher distribution from $\mathrm{E}[R]$. Let $\mathrm{E}[R] = UDV^T$ be its pSVD, the maximum likelihood estimation (MLE) for F is given by $F = USV^T$, where $S = \mathcal{E}^{-1}[D]$ [4], [16].

III. Attitude Uncertainty Propagation

In this section, we consider a stochastic version of the attitude kinematics equations (1) and (2), and we study how the uncertainty distribution of R evolves over time.

A. Stochastic Attitude Kinematics

The stochastic attitude kinematics equations corresponding to (1) and (2) are given by

$$dR = (\omega(t)dt + H(t)dW)^\wedge R, \tag{7}$$

$$dR = R(\Omega(t)dt + H(t)dW)^\wedge, \tag{8}$$

respectively. The angular velocity $\omega(t)$ is resolved in $\mathcal{I}$, and $\Omega(t)$ is resolved in $\mathcal{B}$. Both are assumed to be given as deterministic functions of time. The angular velocity is perturbed by the additive noise $H(t)dW$, which is a Wiener process $W \in \mathbb{R}^3$ scaled by a matrix $H(t) \in \mathbb{R}^{3\times 3}$. The above stochastic differential equations are defined according to the Stratonovich sense so that the random matrix R evolves on $\mathsf{SO}(3)$ [18]. As presented in Section II, we call (7) and (8) right-trivialized and left-trivialized, respectively. In the deterministic case, the right-trivialized (1) and the left-trivialized (2) are equivalent. One of the fundamental discoveries of this paper is that such equivalence does not hold in the stochastic case.

To make the argument more accessible, we study the discretized versions of (7) and (8). Let the time be discretized by a sequence $\{t_k\}_{k=0}^\infty$ with a fixed time step $h = t_{k+1} - t_k$ for all $k \in \mathbb{N}$. Then as $h \to 0$, the solutions of

$$R_{k+1} = \exp\big(h\hat{\omega}_k + (H_k\Delta W)^\wedge\big) R_k \tag{9}$$

$$R_{k+1} = R_k \exp\big(h\hat{\Omega}_k + (H_k\Delta W)^\wedge\big) \tag{10}$$

converge in probability to the solutions of the continuous equations (7) and (8), respectively [18], where ΔW is the stochastic increment of W over a time step h. We further assume the noise is isotropic [18], i.e., $H_k = \gamma I_{3\times 3}$ for $\gamma > 0$. Unfortunately, there is no explicit analytical solution to the stochastic difference equations (9) and (10). Instead, we study how the first moment of R evolves, and interpret its uncertainty utilizing the matrix Fisher distribution.

B. Propagation of First Moments

The moments $\mathrm{E}[R_k]$ for the right-trivialized (9) and left-trivialized (10) are propagated into $\mathrm{E}[R_{k+1}]$ according to the following theorem.

Theorem III.1. *Suppose R follows the right-trivialized stochastic difference equation* (9)*, then*

$$\mathrm{E}[R_{k+1}]_R = (1 - h\gamma^2)\exp(h\hat{\omega}_k)\mathrm{E}[R_k] + \mathcal{O}(h^2). \tag{11}$$

If instead R follows the left-trivialized (10)*, then*

$$\mathrm{E}[R_{k+1}]_L = (1 - h\gamma^2)\mathrm{E}[R_k]\exp(h\hat{\Omega}_k) + \mathcal{O}(h^2). \tag{12}$$

Proof. By the Baker–Campbell–Hausdorff formula, (9) can be written as

$$R_{k+1} = \exp(h\hat{\omega}_k)\exp\big((H_k\Delta W)^\wedge + \mathcal{O}(h\Delta W)\big)R_k.$$

Expanding the exponential term in the middle of the right hand side, and noting that $\mathrm{E}[\Delta W] = 0$, we have

$$\mathrm{E}\Big[\exp\big((H_k\Delta W)^\wedge + \mathcal{O}(h\Delta W)\big)\Big]$$

$$
\begin{aligned}
&= I_{3\times 3} + \tfrac{1}{2}\mathrm{E}\Big[\big((H_k\Delta W)^\wedge + \mathcal{O}(h\Delta W)\big)^2\Big] + \mathcal{O}(h^2) \\
&= (1 - h\gamma^2) I_{3\times 3} + \mathcal{O}(h^2),
\end{aligned}
$$

where we used $\mathrm{E}[\Delta W \Delta W^T] = h I_{3\times 3}$, and the fact that for any $x \in \mathbb{R}^3$, $\hat{x}^2 = -x^T x I_{3\times 3} + x x^T$ in the second equality. The left-trivialized (12) can be derived similarly. □

In (11) and (12), the propagation of the first moment of R from t_k to t_{k+1} is composed of two parts: the exponential term corresponds to the advection, or the rotation of the distribution, due to the deterministic angular velocity, which acts on the left or on the right of the original expectation depending on whether the right- or left-trivialized kinematics is used; and the scalar $1 - h\gamma^2$ represents the diffusion due to noise.

C. Difference Between Right- and Left-Trivialization

The expectation of R carries the similar information as the first two moments of random vectors in Euclidean space, and it can be interpreted using the matrix Fisher distribution. Suppose $R_k \sim \mathcal{M}(F_k)$ with $F_k = U_k S_k V_k^T \in \mathbb{R}^{3\times 3}$. From (5), its first moment is given by $\mathrm{E}[R_k] = U_k D_k V_k^T$. Therefore, (11) and (12) are rewritten as

$$
\begin{aligned}
\mathrm{E}[R_{k+1}]_R &= \exp(h\hat{\omega}_k) U_k \times (1 - h\gamma^2) D_k \times V_k^T, \\
\mathrm{E}[R_{k+1}]_L &= U_k \times (1 - h\gamma^2) D_k \times (\exp(-h\hat{\Omega}_k) V_k)^T
\end{aligned}
$$

respectively, after omitting the higher order terms of h. We assume $R_{k+1} \sim \mathcal{M}(F_{R_{k+1}})$ for (9), and $R_{k+1} \sim \mathcal{M}(F_{L_{k+1}})$ for (10), with the pSVD of the parameters given by $F_{R_{k+1}} = U_{R_{k+1}} S_{R_{k+1}} V_{R_{k+1}}^T$, and $F_{L_{k+1}} = U_{L_{k+1}} S_{L_{k+1}} V_{L_{k+1}}^T$. Then using the MLE for matrix Fisher distribution, for (9) we have

$$
U_{R_k+1} = \exp(h\hat{\omega}_k) U_k, \quad V_{R_{k+1}} = V_k, \tag{13}
$$

and for (10) we have

$$
U_{L_{k+1}} = U_k, \quad V_{L_{k+1}} = \exp(-h\hat{\Omega}_k) V_k. \tag{14}
$$

Next, for both of (9) and (10),

$$
S_{R_{k+1}} = S_{L_{k+1}} = \mathcal{E}^{-1}((1 - h\gamma^2) D_k) \triangleq S_{k+1}. \tag{15}
$$

Now, we consider how the uncertainty of R is propagated in three aspects: the mean attitude, the degree of dispersion, and the principal axes. First, the mean attitude is rotated from $\mathrm{M}[F_k] = U_k V_k^T$ into

$$
\mathrm{M}[F_{R_{k+1}}] = U_{R_{k+1}} V_{R_{k+1}}^T = \exp(h\hat{\omega}_k) U_k V_k^T, \tag{16}
$$

$$
\mathrm{M}[F_{L_{k+1}}] = U_{L_{k+1}} V_{L_{k+1}}^T = U_k V_k^T \exp(h\hat{\Omega}_k), \tag{17}
$$

for (9) and (10) respectively, which is rotated by the deterministic angular velocity as expected. If the angular velocities are transformed to each other by the mean attitude at t_k, then the propagated mean attitudes are identical, i.e., $\Omega_k = (U_k V_k^T)^T \omega_k$ implies $\mathrm{M}[F_{R_{k+1}}] = \mathrm{M}[F_{L_{k+1}}]$. Second, the uncertainty becomes more dispersed in the same manner for (9) and (10), as S_{k+1} is reduced from S_k according to (15) and Lemma II.1.

Finally, the most notable distinction between (9) and (10) is how the principal axes are rotated. For (9), since $U_{R_{k+1}} = \exp(h\hat{\omega}_k) U_k$, the principal axes are rotated by the rotation vector $h\omega_k$ when perceived in the inertial frame. However, as $V_{R_{k+1}} = V_k$, the principal axes remain unchanged when observed from the body-fixed frame. For (10), it is exactly the opposite: since $V_{L_{k+1}} = \exp(-h\hat{\Omega}_k) V_k$, the principal axes are rotated by the rotation vector $-h\Omega_k$ when perceived in the body-fixed frame, and they remain unchanged when observed from the inertial frame. In other words, for (9), the shape of uncertainty represented by the most and least uncertain directions remains fixed in the body-fixed frame, but it is rotated in the inertial frame. On the other hand, for (10), the most and least uncertain directions are rotated in the body-fixed frame, but they are fixed in the inertial frame. These are illustrated in Figure 1.

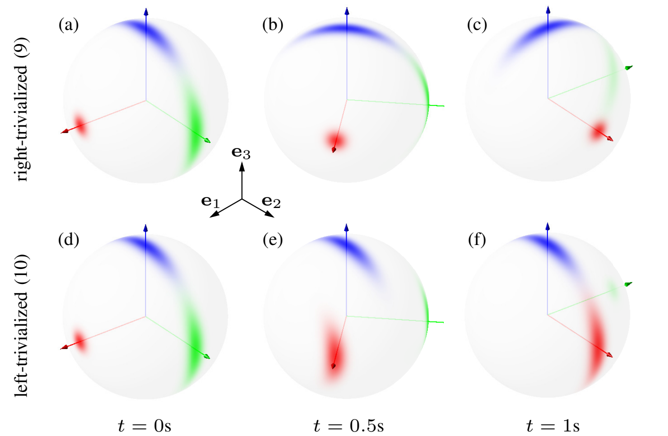


Fig. 1. Illustration of propagated uncertainties: (a-c) by the right-trivialized (9); (d-f) by the left-trivialized (10). The initial distribution is $R_0 \sim \mathcal{M}(F_0)$ with $F_0 = \mathrm{diag}[150, 10, 0]$. The angular velocities are $\omega = \Omega = \frac{\pi}{2} e_3 \mathrm{rad\,s}^{-1}$ without any noise, which can be transformed to each other by the initial mean attitude $I_{3\times 3}$. The red, green and blue arrows represent the mean directions of the body-fixed $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ axes, and the corresponding shades represent the marginal distribution for each body-fixed axis. It can be observed that the mean attitude and the degree of dispersion are consistent between (9) and (10). However, for (9), the most uncertain direction is fixed along the body-fixed $\mathbf{b}_1$ axis (red), but it is rotated in the inertial frame; and the most uncertain direction for (10) is rotated in the body-fixed frame (from red to green), but it remains fixed along the inertial $\mathbf{e}_1$ axis, thereby causing all of the shades circular about $\mathbf{e}_1$.

IV. Single Direction Measurements

In this section, we present two types of direction measurements, and we show how they differ in characterizing attitude uncertainties.

A. Measurement Update

The first type of direction measurement, referred to as *inertial direction measurement*, is when the reference direction is known in the inertial frame, and the measurement output is resolved in the body-fixed frame. For example, this corresponds to a magnetometer that measures the direction of magnetic field, or an accelerometer that measures the direction of gravity. Alternatively, in the less common second type, referred to as *body-fixed direction measurement*, the reference direction is known in the body-fixed frame, and the measurement output is resolved in the inertial frame. For instance, a differential GPS has been utilized in attitude determination of an aircraft, where two GPS antennas are attached to the left wing-tip and the right-wing tip, and GPS measurements provide the direction from the left wing to the right wing in the inertial frame.

For the inertial direction measurement, let $\mathbf{a}$ be a reference vector fixed in the inertial frame, and $a \in \mathsf{S}^2 = \{q \in \mathbb{R}^3 \,|\, \|q\| = 1\}$ be its coordinates in the inertial frame. The reference direction is measured with a sensor fixed to the body, which provides the measurement $x \in \mathsf{S}^2$ resolved in the body-fixed frame. It is assumed that the sensor measurement for a given attitude is distributed according to the von Mises-Fisher distribution on S^2 [19] as follows:

$$
p(x|R) = \frac{\kappa}{4\pi \sinh \kappa} \exp(\kappa a^T R x), \tag{18}
$$

for a parameter $\kappa > 0$. This distribution is centered at $R^T a$ and it is more concentrated as κ is increased, implying a more accurate sensor.

Similarly, for the body-fixed direction measurement, a reference vector $\mathbf{b}$ fixed to the body with the coordinates $b \in \mathsf{S}^2$ in the body-fixed frame, is measured by a sensor as $y \in \mathsf{S}^2$ resolved in the inertial frame. And for $\kappa > 0$, $y|R$ is assumed to be distributed by

$$p(y|R) = \frac{\kappa}{4\pi \sinh \kappa} \exp(\kappa b^T R^T y). \tag{19}$$

The additional information in the direction measurement is injected into the prior uncertainty by calculating the conditional distribution $R|x$ or $R|y$, using Bayes' rule.

Theorem IV.1. *Let $R \sim \mathcal{M}(F^-)$ be a prior distribution for a given $F^- \in \mathbb{R}^{3\times 3}$. Then, for the inertial direction measurement, the posterior distribution of the attitude conditioned by x is also matrix Fisher with $R|x \sim \mathcal{M}(F_I^+)$, where the posterior matrix parameter $F_I^+ \in \mathbb{R}^{3\times 3}$ is given by*

$$F_I^+ = F^- + \kappa a x^T. \tag{20}$$

And for the body-fixed direction measurement, the posterior distribution is $R|y \sim \mathcal{M}(F_B^+)$, where $F_B^+ \in \mathbb{R}^{3\times 3}$ is given by

$$F_B^+ = F^- + \kappa y b^T. \tag{21}$$

Proof. The proof for inertial direction measurement is given in Theorem 3.2 of [4], and the body-fixed direction measurement can be derived similarly. □

B. Difference Between Inertial and Body-Fixed Measurement

Now, we study the implication of (20) and (21). Suppose that the attitude is completely unknown before the measurement, i.e., $F^- = 0_{3\times 3}$. For the inertial direction measurement, the matrix parameter (20) for the posterior distribution is decomposed into

$$F_I^+ = \begin{bmatrix} a & a' & a'' \end{bmatrix} \mathrm{diag}[\kappa, 0, 0] \begin{bmatrix} x & x' & x'' \end{bmatrix}^T, \tag{22}$$

where $a', a'' \in \mathsf{S}^2$ are arbitrarily chosen such that the matrix $[a, a', a''] \in \mathsf{SO}(3)$, and $x', x'' \in \mathsf{S}^2$ are defined similarly. Therefore, F_I^+ is written in the form of pSVD with $S = \mathrm{diag}[\kappa, 0, 0]$. The first principal axis is a when resolved in the inertial frame, or x when resolved in the body-fixed frame. Also, the rotation about the first principal axis is completely unknown as $s_2 + s_3 = 0$. More intuitively, the marginal distribution of each body-fixed axis makes a circle normal to $\mathbf{a}$ (the top row of Figure 2), which implies the rotation about $\mathbf{a}$ cannot be determined. Since $\mathbf{a}$ is fixed in the inertial frame, the direction of this ambiguity is also fixed in the inertial frame regardless of the measurement.

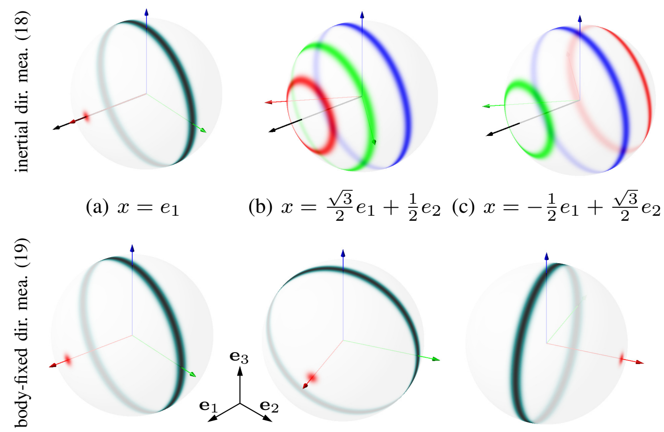
Next, the matrix parameter (21) for the posterior distribution of a body-fixed direction measurement is decomposed into

$$F_B^+ = \begin{bmatrix} y & y' & y'' \end{bmatrix} \mathrm{diag}[\kappa, 0, 0] \begin{bmatrix} b & b' & b'' \end{bmatrix}^T, \tag{23}$$

where $y', y'' \in \mathsf{S}^2$ and $b', b'' \in \mathsf{S}^3$ are chosen such that the corresponding matrices in brackets belong to $\mathsf{SO}(3)$. The resulting first principal axis is b when resolved in the body-fixed frame, or y when resolved in the inertial frame. The marginal distribution of each body-fixed axis makes a circle normal to $\mathbf{b}$ (the bottom row of Figure 2), about which the rotation cannot be determined. Since $\mathbf{b}$ is fixed in the body-fixed frame, the direction of this ambiguity is also fixed to the body.

V. Stochastic Attitude Observability

Here, we introduce two stochastic attitude observability criteria. Since the posterior distribution of R conditioned by measurements contains all the information available to determine the attitude, we study: (i) whether there is a *unique* attitude that minimizes the mean square error for the posterior distribution, and (ii) whether the Fisher information matrix for the mean attitude of R is positive-definite when R is distributed according to the matrix Fisher distribution.



(a) $x = e_1$ (b) $x = \frac{\sqrt{3}}{2}e_1 + \frac{1}{2}e_2$ (c) $x = -\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2$

(d) $y = e_1$ (e) $y = \frac{\sqrt{3}}{2}e_1 + \frac{1}{2}e_2$ (f) $y = -\frac{1}{2}e_1 + \frac{\sqrt{3}}{2}e_2$

Fig. 2. Posterior distribution with single direction measurements ($\kappa = 500$): (a-c) inertial direction measurements with $a = e_1$; (d-f) body-fixed direction measurements with $b = e_1$.

A. Uniqueness of Attitude Estimate

One of the methods to estimate an attitude from a density function on $\mathsf{SO}(3)$ is to solve an optimization problem that minimizes the mean square Frobenius norm [20], [21].

Definition V.1. *Let $p(R)$ be the probability density function for a random $R \in \mathsf{SO}(3)$. Its minimum mean square estimate (MMSE) is defined as*

$$\mathrm{M}_{\mathrm{MMSE}}[R] = \operatorname*{arg\,min}_{Q \in \mathsf{SO}(3)} \{\mathrm{E}[\|R - Q\|_F^2]\}, \tag{24}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

Finding MMSE can be addressed in terms of pSVD as follows.

Lemma V.1. *Suppose $R \in \mathsf{SO}(3)$ is a random rotation matrix. Let the pSVD of its first moment be $\mathrm{E}[R] = UDV^T$ with $D = \mathrm{diag}[d_1, d_2, d_3]$. Depending on D, the MMSE of R is given by*

1) *$d_2 + d_3 > 0$: UV^T (unique),*
2) *$d_1 \neq d_2$ and $d_2 + d_3 = 0$: $\{U \exp(\theta \hat{e}_1) V^T \,|\, \theta \in [-\pi, \pi)\}$ (1D),*
3) *$d_1 = d_2 = -d_3 > 0$: $\{U \exp(\theta \hat{a}) V^T \,|\, a \in \mathsf{S}^2,\ a_3 = 0, \theta \in [-\pi, \pi)\}$ (2D),*
4) *$d_1 = d_2 = d_3 = 0$: $\mathsf{SO}(3)$ (3D),*

where the number in the parentheses indicates the dimension of the set corresponding to the solution of MMSE.

Proof. The proof uses the equivalent definition [4]

$$\mathrm{M}_{\mathrm{MMSE}}[R] = \operatorname*{arg\,max}_{Q \in \mathsf{SO}(3)} \{\mathrm{tr}[Q^T \mathrm{E}[R]]\},$$

and is based on a careful analysis of the uniqueness of pSVD depending on the multiplicity of singular values. The detailed procedure is available in [14]. □

For all cases, the set of MMSE contains UV^T. Nevertheless, only when $d_2 + d_3 > 0$, the MMSE is unique. Otherwise, it can only be determined up to a rotation, where the dimension of the set representing the solution of MMSE is equivalent to 3 minus the rank of $\mathrm{tr}[D] I_{3\times 3} - D = \mathrm{diag}[d_2 + d_3, d_1 + d_3, d_1 + d_2]$. Therefore, we claim that the attitude is completely observable given a density function on $\mathsf{SO}(3)$ if $\mathrm{tr}[D] I_{3\times 3} - D$ is positive-definite, i.e., when the MMSE is unique.

B. Information Theoretic Observability Analysis

The above observability criterion is solely based on the first moment, and therefore, it can be applied to an unknown attitude following an arbitrary distribution. Here, we assume that the attitude is distributed by a matrix Fisher distribution to present an alternative information theoretic observability criterion.

Suppose $R \sim \mathcal{M}(F)$ where the pSVD of $F \in \mathbb{R}^{3\times 3}$ is given by $F = USV^T$. As discussed in Section II-B, the MMSE of $R \sim \mathcal{M}(USV^T)$ is given by the mean attitude $R^* = UV^T$, and we want to calculate its Fisher information $\mathbb{I}(R^*)$. We first study a more general problem of estimating U, S, and V from the given samples of R. The log-likelihood is

$$l(R|U,S,V) = \mathrm{tr}[USV^T R^T] - \log c(S). \tag{25}$$

And the corresponding Fisher information matrix is calculated as follows [13].

Lemma V.2. *The Fisher information matrix of* (25)*, namely* $\mathbb{I}(U,S,V) : \mathbb{R}^9 \times \mathbb{R}^9 \to \mathbb{R}$ *is constructed as*

$$\begin{aligned}&\mathbb{I}(U,S,V) = -\mathrm{E}[\nabla^2 l(R|U,S,V)]\\ &= \begin{bmatrix} \mathrm{tr}[DS]I_{3\times 3} - DS & 0 & \sum_{i=1}^3 e_i^T s\hat{e}_i D\hat{e}_i \\ 0 & \frac{\partial^2 \log c(S)}{\partial s^2} & 0 \\ \sum_{i=1}^3 e_i^T s\hat{e}_i D\hat{e}_i & 0 & \mathrm{tr}[DS]I_{3\times 3} - DS \end{bmatrix}, \end{aligned} \tag{26}$$

where $D \in \mathbb{R}^{3\times 3}$ *is the diagonal matrix composed of the proper singular values of* $\mathrm{E}[R|U,S,V]$*,* $s = \begin{bmatrix} s_1 & s_2 & s_3 \end{bmatrix}^T$*, and* ∇^2 *is the covariant Hessian on* $\mathsf{SO}(3) \times \mathbb{R}^3 \times \mathsf{SO}(3)$*.*

Proof. Let $\mathsf{Q} = \mathsf{SO}(3) \times \mathbb{R}^3 \times \mathsf{SO}(3)$, and $q = (U,S,V) \in \mathsf{Q}$. The tangent space $\mathsf{T}_q\mathsf{Q}$ is identified with $\mathsf{T}_q\mathsf{Q} \simeq \mathbb{R}^9$ through the hat map, and the cotangent space is also identified with $\mathbb{R}^9$ using the dot product. More specifically, for $\xi = (u, \varsigma, v) \in \mathbb{R}^9$, the corresponding tangent vector is given by $(U\hat{u}, \varsigma, V\hat{v}) \in \mathsf{T}_q\mathsf{Q}$.

Since l is a real-valued function on Q, its covariant derivative $\nabla_\xi l$ along ξ is equivalent to the differential $dl(\xi)$ given by

$$\begin{aligned}\nabla_\xi l &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} l(R|U\exp(\epsilon\hat{u}), S + \epsilon\mathrm{diag}[\varsigma], V\exp(\epsilon\hat{v}))\\ &= \mathrm{tr}[(U\hat{u}SV^T + U\mathrm{diag}[\varsigma]V^T - US\hat{v}V^T)^T R] - \frac{\partial \log c(S)}{\partial s}\cdot\varsigma\\ &= \begin{bmatrix} (QS - SQ^T)^\vee \\ \mathrm{diag}[Q] - \frac{1}{c(S)}\frac{\partial c(S)}{\partial s} \\ (Q^T S - SQ)^\vee \end{bmatrix}\cdot\xi,\end{aligned}$$

where $Q = U^T RV$. Because $\mathrm{E}[Q] = D$ is diagonal, it is straightforward to show $\mathrm{E}[dl(\xi)] = 0$ for any $\xi \in \mathbb{R}^9$.

The covariant Hessian of l along ξ_1 and ξ_2 is given by $\nabla^2_{\xi_1,\xi_2} l = \xi_2(\xi_1 l) - (\nabla_{\xi_2}\xi_1)l = \xi_2(dl(\xi_1)) - dl(\nabla_{\xi_2}\xi_1)$, where the second term vanishes after taking expectation. The first term is bi-linear in ξ_1 and ξ_2, thus it can be written as a matrix as in (26). Suppose $\xi_1 = (u_1, 0, 0)$ and $\xi_2 = (u_2, 0, 0)$. We have

$$\begin{aligned}\xi_2(dl(\xi_1)) &= (-\hat{u}_2 QS - SQ^T\hat{u}_2)^\vee \cdot u_1\\ &= u_1^T\left\{\frac{1}{2}(QS + SQ^T) - \mathrm{tr}[QS]\, I_{3\times 3}\right\}u_2.\end{aligned}$$

Taking the expectation of the expression in the braces with $\mathrm{E}[Q] = D$ and multiplying it with -1 yield the upper-left 3-by-3 block of (26). The remaining blocks can be obtained similarly. ☐

Next, we calculate the Fisher information $\mathbb{I}(R^*)$ with the above information matrix. Since the variations of U and V are written as $\delta U = U\hat{u}$ and $\delta V = V\hat{v}$ for $u, v \in \mathbb{R}^3$, we have

$$\delta R^* = U\hat{u}V^T - U\hat{v}V^T.$$

Let $\eta = u - v \in \mathbb{R}^3$ so that $\delta R^* = U\hat{\eta}V^T$. Thus, the Fisher information matrix for the mean attitude R^* is constructed by left-multiplying (26) with the matrix $\frac{1}{2}[I_{3\times 3}; 0_{3\times 3}; -I_{3\times 3}]$, and by right-multiplying (26) with its transpose, to obtain

$$\mathbb{I}(R^*) = \frac{1}{2}\mathrm{diag}\begin{bmatrix} (d_2+d_3)(s_2+s_3) \\ (d_3+d_1)(s_3+s_1) \\ (d_1+d_2)(s_1+s_2) \end{bmatrix}. \tag{27}$$

According to the Cramér–Rao inequality, the inverse of Fisher information $\mathbb{I}(R^*)$ is a lower bound of the variance of all unbiased estimates, up to additional curvature terms. Therefore, its positive-definiteness can be used to define observability [12]. Interestingly, by Lemma II.1, the positive-definiteness of $\mathbb{I}(R^*)$ is equivalent to the uniqueness of MMSE presented in Lemma V.1. Based on these results, we formulate stochastic attitude observability for an arbitrary density as follows.

Definition V.2. *A random rotation matrix* $R \sim p(R)$ *is stochastically observable if* $d_2 + d_3 > 0$*, or equivalently*

$$\mathcal{O} = \mathrm{tr}[D]I_{3\times 3} - D \succ 0, \tag{28}$$

where $D = \mathrm{diag}[d_1, d_2, d_3]$ *is the proper singular values of* $\mathrm{E}[R]$*. The corresponding measure of observability is*

$$\rho(R) = \det[\mathcal{O}] = (d_1 + d_2)(d_3 + d_1)(d_2 + d_3). \tag{29}$$

When $R \sim \mathcal{M}(USV^T)$, it is straightforward to show that (28) is equivalent to $\mathrm{tr}[S]I_{3\times 3} - S \succ 0$ with Lemma II.1. Note that these are readily applied to the stochastic observability considered in this paper, as the posterior distribution conditioned by direction measurements is assumed to be a matrix Fisher distribution according to Theorem IV.1, which is presented in the next section.

VI. Attitude Observability with Single Direction Measurements

The attitude uncertainty propagation in Section III, and the measurement update in Section IV constitute a Bayesian estimator, which provides the posterior distribution of the attitude conditioned by the history of direction measurements. The posterior distribution can then be used in Definition V.2 to determine attitude observability. As there are two cases for each of uncertainty propagation and measurement update, we have four possible combinations as summarized in Table I. This section identifies two combinations that yield unobservability, and two other cases resulting in observability with single direction measurements. The same results can also be derived in a deterministic sense as presented in Appendix.

A. Combinations with Unobservability

We first discuss why the common IMU cannot estimate the full attitude with single direction measurements. In a typical IMU, the angular velocity is measured in the body-fixed frame using a gyroscope, and the reference direction in the inertial frame, such as the direction of gravity, is measured in the body-fixed frame. As such it is a combination of the left-trivialized (10) and the inertial direction measurement (18). Looking at the bottom row of Figure 1 and the top row of Figure 2, it is clear why the attitude cannot be determined in this case: the direction about which the rotation cannot be determined, namely the first principal axis, remains unchanged in the inertial frame for both (10) and (18). This is formulated in the next theorem.

Theorem VI.1. *Consider the two Bayesian attitude estimators for*

- *right-trivialized angular velocity in the inertial frame* (9) *and body-fixed direction measurement* (19)

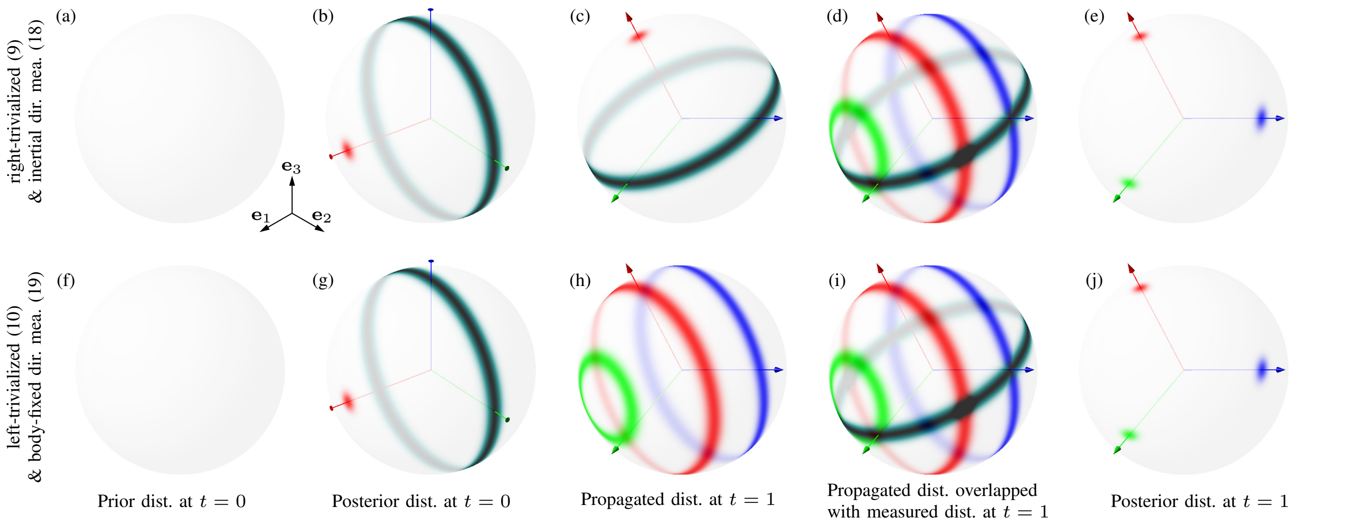


Fig. 3. Combinations when the attitude is observable. (a-e) right-trivialized (9) and inertial direction measurement (18); (f-j) left-trivialized (10) and body-fixed direction measurement (19).

- *left-trivialized angular velocity in the body-fixed frame* (10)*, and inertial direction measurement* (18)

with the initial distribution $F_0 = 0_{3\times3}$. For both cases, the attitude is not observable.

Proof. First consider the filter given by (13), (15), and (21). The propagated uncertainty before the first measurement is $F_1^- = 0_{3\times3}$, thus $F_1 = \kappa y_1 b^T$ after conditioning the first measurement. As shown in (23), $(s_2)_1 = (s_3)_1 = 0$, and $V_1 e_1 = b$. We proceed with induction. Suppose $(s_2)_k = (s_3)_k = 0$, and $V_k e_1 = b$. Then by (13), (15) and Lemma II.1, the propagated parameters before the next measurement still satisfy $(s_2)_{k+1}^- = (s_3)_{k+1}^- = 0$ and $V_{k+1}^- e_1 = b$. Next consider the update $F_{k+1} = F_{k+1}^- + \kappa y_{k+1} b^T$, which can be written as

$$F_{k+1} = \left((s_1)_{k+1}^- U_{k+1}^- e_1 + \kappa y_{k+1}\right) b^T \triangleq u b^T,$$

where $u \in \mathbb{R}^3$. Let $U_{k+1} = \begin{bmatrix}\frac{u}{\|u\|} & u' & u''\end{bmatrix}$, where $u', u'' \in \mathsf{S}^2$ are arbitrarily chosen such that $U_{k+1} \in \mathsf{SO}(3)$. Also let $S_{k+1} = \mathrm{diag}[\|u\|, 0, 0]$, and $V_{k+1} = \begin{bmatrix}b & b' & b''\end{bmatrix}$ as in (23). Then $F_{k+1} = U_{k+1} S_{k+1} V_{k+1}^T = u b^T$ is the pSVD of F_{k+1}, and we have shown that $(s_2)_{k+1} = (s_3)_{k+1} = 0$ and $V_{k+1} e_1 = b$. Therefore, $(s_2)_k = (s_3)_k = 0$ for all $k \in \mathbb{N}$, and by Lemma II.1, the attitude is not observable. The proof for the second case is similar. $\square$

B. Combinations with Observability

Next, we present examples illustrating that the other two combinations yield observability. Consider the combination of right-trivialized (9) and inertial direction measurement (18) for a specific case where the true attitude evolves according to

$$R(t) = R_0 \exp(\hat{\Omega} t) = \exp(\hat{\omega} t) R_0, \tag{30}$$

with $R_0 = I_{3\times3}$ and $\Omega = \omega = -\frac{\pi}{2\sqrt{3}}[1, 1, 1] \in \mathbb{R}^3$. Initially, it is assumed that the attitude is completely unknown, i.e., $F_0 = 0_{3\times3}$, as illustrated in Figure 3(a). Then after updated by an inertial direction measurement with $a = e_1$, the rotation about the reference direction is unobservable and the resulting distribution is axially symmetric about $\mathbf{e}_1$ (Figure 3(b)). For the right-trivialized (9), the distribution rotates in the inertial frame over the propagation, and therefore, the direction of ambiguity is no longer along $\mathbf{e}_1$ (Figure 3(c)). The next inertial direction measurement is fixed in the inertial frame along $\mathbf{e}_1$ (Figure 3(d)). Thus, it resolves the ambiguity of the propagated density to determine the attitude completely (Figure 3(e)).

The other combination of left-trivialized (10) and body-fixed direction measurement (19) is similar. The direction of ambiguity caused by the first body-fixed direction measurement with $b = e_1$ is fixed in the inertial frame after propagation (Figure 3(h)). However, the ambiguous direction of the next direction measurement is rotated in the inertial frame (Figure 3(i)), and this resolves the previous ambiguity (Figure 3(j)). The above intuition is formally presented in the next theorem.

Theorem VI.2. *Consider the two Bayesian attitude estimators for*

- *right-trivialized angular velocity in the inertial frame* (9) *and inertial direction measurement* (18)
- *left-trivialized angular velocity in the body-fixed frame* (10)*, and body-fixed direction measurement* (19)

with the initial distribution $F_0 = 0_{3\times3}$. Suppose there is some k_0 such that $\omega_{k_0} \times a \neq 0$ for the first case, and $\Omega_{k_0} \times b \neq 0$ for the second case. Then the attitude is observable with probability one for both cases.

Proof. Consider the first case. The posterior distribution after the first measurement is given by $F_1 = \kappa a x_1^T$. Suppose $k_0 = 1$ and denote $\exp(h\hat{\omega}_1) = \delta R$, then by (22), (13), (15) and Lemma II.1, $U_2^- = \delta R \begin{bmatrix}a & a' & a''\end{bmatrix}$, $S_2^- = \mathrm{diag}([(s_1)_2^-, 0, 0])$, $V_2^- e_1 = x_1$, where $(s_1)_2^- > 0$ satisfies

$$\frac{1}{c(S_2^-)} \frac{\partial c(S_2^-)}{\partial (s_1)_2^-} = (1 - h\gamma^2) \frac{1}{c(S_1)} \frac{\partial c(S_1)}{\partial (s_1)_1}.$$

Thus, the posterior distribution after the second measurement is

$$F_2 = (s_1)_2^- \delta R a x_1^T + \kappa a x_2^T. \tag{31}$$

Let $\delta R a = \alpha a + \alpha' a' + \alpha'' a''$ for some $\alpha, \alpha', \alpha'' \in \mathbb{R}$. Since $\omega_1 \times a \neq 0$, α' and α'' cannot both be zeros. Then

$$\begin{aligned} F_2 &= (s_1)_2^- (\alpha a + \alpha' a' + \alpha'' a'') x_1^T + \kappa a x_2^T \\ &= a((s_1)_2^- \alpha x_1 + \kappa x_2)^T + (s_1)_2^- \alpha' a' x_1^T + (s_1)_2^- \alpha'' a'' x_1^T. \end{aligned}$$

Let $v = (s_1)_2^- \alpha_1 x_1 + \kappa x_2$, and v', v'' be arbitrarily chosen such that $\begin{bmatrix}\frac{v}{\|v\|} & v' & v''\end{bmatrix} \in \mathsf{SO}(3)$. Also, let $x_1 = \beta \frac{v}{\|v\|} + \beta' v' + \beta'' v''$. Note that β_2 and β_3 cannot both be zeros almost surely. Using these, F_2 is written as

$$F_2 = \begin{bmatrix} a & a' & a'' \end{bmatrix} \Lambda \begin{bmatrix} \frac{v}{\|v\|} & v' & v'' \end{bmatrix}^T,$$

where $\Lambda \in \mathbb{R}^{3\times 3}$ is

$$\Lambda = \begin{bmatrix} \|v\| & 0 & 0 \\ (s_1)_2^- \alpha'\beta & (s_1)_2^- \alpha'\beta' & (s_1)_2^- \alpha'\beta'' \\ (s_1)_2^- \alpha''\beta & (s_1)_2^- \alpha''\beta' & (s_1)_2^- \alpha''\beta'' \end{bmatrix}.$$

Since F_2 is obtained by multiplying rotation matrices to Λ, it is straightforward to see that F_2 and Λ share the same proper singular values. We have $\det(\Lambda) = 0$, so there is at least one zero singular value. However, the rank of Λ is two almost surely, as at least one element of the right bottom 2-by-2 block is nonzero. Thus, Λ has only one zero singular value. By the definition of proper singular value decomposition, this concludes $\mathrm{tr}[S_2]I_{3\times 3} - S_2$ is positive-definite, and therefore the attitude is observable with probability one.

Next suppose $k_0 > 1$. By (31), $F_2 = a((s_2)_1^- x_1 + \kappa x_2)^T$ since δR does not rotate a, which means both the uncertainty propagation and update steps leave $U_k e_1 = a$, $(s_1)_k > 0$, and $(s_2)_k = (s_3)_k = 0$. Thus, the argument in the last paragraph still applies at time $t = t_{k_0}$. The proof for the other estimator is similar. □

Finally, it should be noted that although we have assumed the attitude follows matrix Fisher distribution, it is not restrictive in the sense that if the initial distribution is uniform, and the angular velocity noise in (9) and (10) is zero, i.e., $\gamma = 0$, then the attitude conditioned by single direction measurements exactly follows the matrix Fisher distribution. Therefore, we suppose that the result presented in Table I is not specific to the estimator. Instead, it is inherent to the observed stochastic dynamical system given by (9), (10) and (18), (19). Indeed, it is shown by experiments in the next section that multiplicative extended Kalman filter also exhibits the same observability.

VII. Experiments

In this section, the attitude observability presented above is validated through experiments. We use a custom-made hardware platform to collect measurements while moving it with hands. A VICON motion capture system detects reflective markers attached to the platform to determine its attitude, which is used as the ground truth. An IMU (VectorNav VN100) is attached to the platform, and the onboard gyroscope provides the angular velocity measurement in the body-fixed frame, which is also transformed into the inertial frame using the true attitude. For the inertial direction measurement, the direction of gravity is measured by the accelerometer on IMU. And for the body-fixed direction measurement, two additional markers are attached to the platform as the reference direction in the body-fixed frame, which is measured by the Vicon motion system in the inertial frame. All Vicon and IMU measurements are synchronized and sampled at 100 Hz. The platform was rotated by hands about its roll, pitch and yaw axes during the data collection.

The matrix Fisher estimator and MEKF are executed off-board using the collected experimental data, with the single direction measurement update applied at 20 Hz. The initial attitude is set as the true attitude rotated about the body-fixed $\mathbf{b}_1$ axis by 180 deg, and the initial uncertainty is set as uniform, i.e., $F_0 = 0_{3\times 3}$. The noise parameters are chosen as $\gamma = 10\,\mathrm{deg}/\sqrt{\mathrm{s}}$, and $\kappa = 200$. The four combinations of angular velocity and direction measurements are labeled as:

ref. vec. \ ang. vel.	body-fixed frame	inertial frame
body-fixed frame	**AVB_RVB**	AVI_RVB
inertial frame	AVB_RVI	**AVI_RVI**

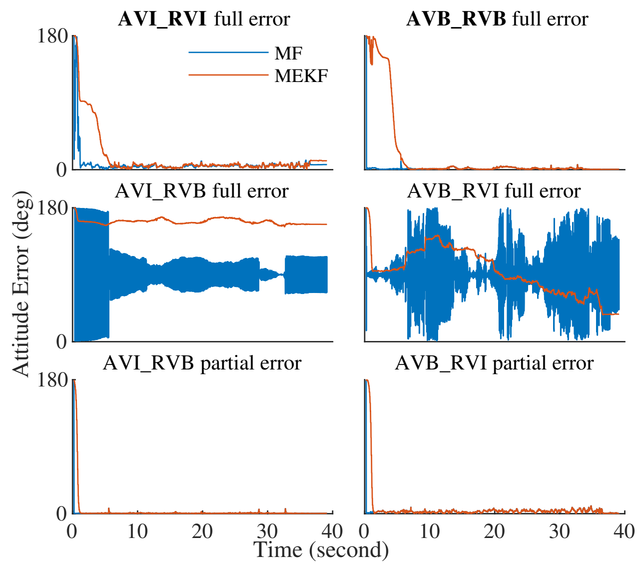


Fig. 4. Attitude errors for the matrix Fisher estimator (MF) and MEKF.

where the boldface font indicates the cases with observability. The full attitude error denotes the angle between the estimated and true attitude. For the two unobservable cases, the partial attitude error is also calculated, which is the angle between the reference vector resolved in the measurement frame using the true and estimated attitudes, neglecting the rotation about the reference vector.

The attitude errors for the four combinations are presented in Figure 4. It is shown that for the two observable cases (AVI_RVI and AVB_RVB), the full attitude error converges, whereas it does not converge for the two unobservable cases (AVI_RVB and AVB_RVI). However, the partial attitude error for the unobservable cases converges to around zero, indicating *only* the rotation about the reference direction cannot be estimated. The attitude uncertainty is presented in Fig. 5. For all combinations, the uncertainty along the directions perpendicular to the reference direction (blue and red lines) remains very low at around 3 deg. On the other hand, the uncertainty along the reference direction (yellow lines) is low only for the two observable cases, and it is almost a uniform distribution (around 10^7deg) for the two unobservable cases. An exception is for MEKF, the uncertainty along the reference vector for AVB_RVI is also low. This is caused by that the attitude error is formulated in the body-fixed frame, rather than in the inertial frame. And if the reference vector is known in the inertial frame, MEKF has been shown to apply some slight but erroneous corrections to the rotation about the reference vector due to the linearization of the measurement function [22].

VIII. Conclusions

This paper addresses the fundamental question whether the attitude of a rigid body is observable with angular velocity and single direction measurements. By observing that the attitude uncertainties are propagated distinctively depending on how the angular velocity measurements are resolved, this paper has discovered two particular cases where the attitude is observable with multiple measurements of a single, fixed reference direction, which has been widely accepted to be impossible. This is further studied by formulating stochastic attitude observability through information-theoretic analysis, and it is also validated by experimental results.

Appendix

In this appendix, attitude observability with single direction measurements is presented in a deterministic sense, which is consistent

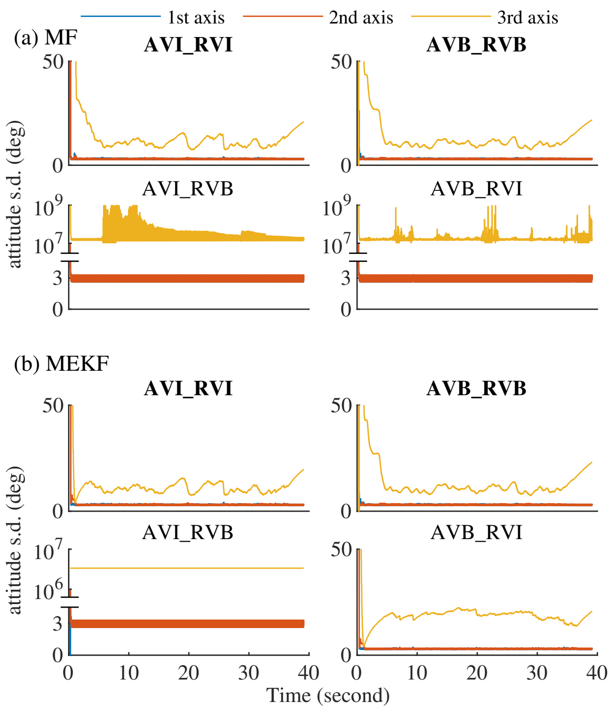


Fig. 5. Attitude uncertainty for the matrix Fisher estimator (MF) and MEKF. In the "RVI" cases, the attitude covariance matrix is expressed in the inertial frame; and in the "RVB" cases, it is expressed in the b'-b''-b frame, where b is the reference vector when resolved in the body-fixed frame, and b', b'' are perpendicular to b. For the MF filter, $(\mathrm{tr}[S]\, I_{3\times 3} - S)^{-1}$ is used as the attitude covariance matrix in the principal axes frame.

with Theorem VI.1 and Theorem VI.2.

Theorem A.1. *Let the deterministic inertial and body-fixed direction measurements be*

$$x = R^T a, \tag{32}$$
$$y = Rb, \tag{33}$$

respectively, where $a, b \in \mathsf{S}^2$ are reference vectors. Then the system (1) *and* (32)*, and the system* (2) *and* (33) *are weakly locally observable. Conversely, the system* (1) *and* (33)*, and the system* (2) *and* (32) *are unobservable.*

Proof. The proof is based on the results of [23], and we adopt notations therein without reintroducing them here for brevity. Without loss of generality, we assume $a = b = e_1$.

For (1), $\mathcal{F}^0(R)$ is spanned by $\{\hat{e}_i R\}_{i=1}^3$. For any $R \in \mathsf{SO}(3)$ and $\hat{\eta} \in \mathfrak{so}(3)$, the Lie derivative of $x(R)$ along $\hat{\eta}R$ is

$$(L_{(\hat{\eta}R)}x)(R) = \frac{d}{dt}\Big|_{t=0} R^T \exp(t\hat{\eta})^T a = R^T \hat{\eta}^T a. \tag{34}$$

Let $\tilde{\mathcal{G}} = \{L_{(\hat{e}_2 R)}x, L_{(\hat{e}_3 R)}x\} \subset \mathcal{G}$. For any $R_0 \in \mathsf{SO}(3)$, define a local coordinate $\theta \in \mathbb{R}^3$ with $R(\theta) = \exp(\hat{\theta})R_0$. Then $dR = \widehat{d\theta}R_0$, and we have

$$d(L_{(\hat{e}_2 R_0)}x)(R_0) = R_0^T \widehat{d\theta}^T \hat{e}_2^T e_1 = R_0^T \hat{e}_3 d\theta,$$
$$d(L_{(\hat{e}_3 R_0)}x)(R_0) = R_0^T \widehat{d\theta}^T \hat{e}_3^T e_1 = -R_0^T \hat{e}_2 d\theta.$$

Thus, $d\tilde{\mathcal{G}}(R_0) = R_0^T[\hat{e}_3, -\hat{e}_2]d\theta$. Since $\mathrm{rank}[\hat{e}_3, -\hat{e}_2] = 3$, it follows that the dimension of $d\mathcal{G}(R_0)$ is three. Therefore, the system (1) and (32) is weakly locally observable [23, Theorem 3.2].

Next, for any $R \in \mathsf{SO}(3)$ and $\hat{\eta}_1, \hat{\eta}_2 \in \mathfrak{so}(3)$, we have

$$d\big(L_{(\hat{\eta}_1 R)}(L_{(\hat{\eta}_2 R)}y)\big)(R_0) = \hat{\eta}_1\hat{\eta}_2\widehat{d\theta}R_0 b = -\hat{\eta}_1\hat{\eta}_2\widehat{R_0 b}d\theta.$$

As such, any higher-order Lie derivative of y along (1) would include the factor $\widehat{R_0 b}$, which has rank two. Thus, the dimension of $d\mathcal{G}(R_0)$ is at most two. Because the system (1) is locally controllable, the system of (1) and (33) is unobservable according to [23, Theorem 3.12]. The remaining cases with (2) can be shown similarly. $\square$

References

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

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

[3] A. Barrau and S. Bonnabel, "The invariant extended Kalman filter as a stable observer," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1797–1812, 2016.

[4] T. Lee, "Bayesian attitude estimation with the matrix Fisher distribution on SO(3)," *IEEE Transactions on Automatic Control*, vol. 63, no. 10, pp. 3377–3392, 2018.

[5] W. Wang and T. Lee, "Matrix Fisher–Gaussian distribution on $\mathsf{SO}(3) \times \mathbb{R}^n$ for Bayesian attitude estimation," *IEEE Transactions on Automatic Control*, vol. 67, no. 5, pp. 2175–2191, 2022.

[6] P. Batista, C. Silvestre, and P. Oliveira, "A GES attitude observer with single vector observations," *Automatica*, vol. 48, no. 2, pp. 388–395, 2012.

[7] 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 Transactions on automatic control*, vol. 57, no. 5, pp. 1332–1338, 2011.

[8] T. Lee, M. Leok, N. McClamroch, and A. Sanyal, "Global attitude estimation using single direction measurements," in *Proceedings of the American Control Conference*, New York, NY, Jul. 2007, pp. 3659–3664.

[9] J. Reis, P. Batista, P. Oliveira, and C. Silvestre, "Nonlinear observer on SO(3) for attitude estimation on rotating earth using single vector measurements," *IEEE Control Systems Letters*, vol. 3, no. 2, pp. 392–397, 2018.

[10] M. Moakher, "Means and averaging in the group of rotations," *SIAM journal on matrix analysis and applications*, vol. 24, no. 1, pp. 1–16, 2002.

[11] X. Pennec, "Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements," *Journal of Mathematical Imaging and Vision*, vol. 25, no. 1, p. 127, 2006.

[12] R. Mohler and C. Hwang, "Nonlinear data observability and information," *Journal of the Franklin Institute*, vol. 325, no. 4, pp. 443–464, 1988.

[13] S. T. Smith, "Covariance, subspace, and intrinsic Cramér-Rao bounds," *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1610–1630, 2005.

[14] W. Wang, K. Gamagedara, and T. Lee, "On the observability of attitude with single direction measurements," *arXiv preprint arXiv:2008.13067*, 2020.

[15] T. Downs, "Orientation statistics," *Biometrika*, vol. 59, pp. 665–676, 1972.

[16] C. Khatri and K. Mardia, "The von Mises-Fisher matrix distribution in orientation statistics," *Journal of the Royal Statistical Society. Series B*, vol. 39, no. 1, pp. 95–106, 1977.

[17] F. Markley, "Attitude determination using vector observations and the singular value decomposition," *The Journal of the Astronautical Sciences*, vol. 36, no. 3, pp. 245–259, 1988.

[18] A. Barrau and S. Bonnabel, "Stochastic observers on Lie groups: a tutorial," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 1264–1269.

[19] K. Mardia and P. Jupp, *Directional Statistics*. Wiley, 1999.

[20] Y. Oshman and A. Carmi, "Estimating attitude from vector observations using a genetic algorithm-embedded quaternion particle filter," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2004, p. 5340.

[21] F. Markley, "Attitude filtering on SO(3)," in *Proceedings of the AAS Malcolm D. Shuster Astronautics Symposium*, 2005, AAS 05-460.

[22] M. Li and A. I. Mourikis, "High-precision, consistent EKF-based visual-inertial odometry," *The International Journal of Robotics Research*, vol. 32, no. 6, pp. 690–711, 2013.

[23] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Transactions on automatic control*, vol. 22, no. 5, pp. 728–740, 1977.