Unscented Kalman Filtering for Simultaneous Estimation of Attitude and Gyroscope Bias

Donghoon Kang, *Member, IEEE,* Cheongjae Jang, and Frank C. Park, *Fellow, IEEE*

***Abstract*—We present an unscented Kalman filtering (UKF) algorithm for simultaneously estimating attitude and gyroscope bias from an inertial measurement unit (IMU). The algorithm is formulated as a discrete-time stochastic nonlinear filter with state space given by the direct product matrix Lie group $SO(3) \times \mathbb{R}^3$, and observations in SO(3) reconstructed from measurements of the gravity and geomagnetic field vectors. Computationally efficient implementations of our filter are made possible by formulating the state space dynamics and measurement equations in a way that leads to closed-form equations for covariance propagation and correction. The resulting attitude estimates are invariant with respect to choice of fixed and moving reference frames. Numerical and hardware experiments involving both synthetic and real data demonstrate the performance advantages of our filter vis-à-vis existing state-of-the-art IMU attitude estimation algorithms.**

***Index Terms*—Attitude estimation, unscented Kalman filter, gyroscope bias, inertial measurement unit.**

I. Introduction

Estimating an object's orientation, or attitude, from an inertial measurement unit (IMU) attached to the object, arises in numerous applications, from vehicle and robot navigation [1]–[3] to human pose tracking [4]. A typical IMU consists of a gyroscope, accelerometer, and magnetometer. The gyroscope measures angular rates that are used to integrate the object's kinematic equations and calculate its attitude. Because gyroscope measurements are subject to a time-varying bias error, they are often augmented by other sensor measurements obtained from accelerometers (for measuring gravity and other accelerations) and magnetometers (for measuring the geomagnetic field vector). The practical and other benefits of simultaneously estimating the attitude and gyroscope bias from disparate sensor measurements have already been emphasized in the relevant literature (see, e.g., [5] and the references cited therein).

Notable among **deterministic filtering** methods that simultaneously estimate attitude and gyroscope bias are Mahony et al's series of nonlinear complementary filters (NCFs) [6]; these ensure almost global stability of the observer error, and their performance has been validated in numerous realistic scenarios [7], [8]. **Stochastic filtering** methods further take into account statistical characterizations of measurement noise, and include well-known and widely used methods such as the extended Kalman filter (EKF). More recently the unscented Kalman filter (UKF), despite its greater computational complexity, has been shown to outperform the EKF in a wide range of applications [9]–[11].

The problem of attitude estimation is usually formulated by nonlinear equations in two fundamental ways: both the underlying configuration space of rotations, and the state space dynamics and measurements, are described by a set of nonlinear equations. With respect to the curved nature of the space of rotations, the straightforward but naive approach of expressing a rotation in terms of some suitable local coordinates, e.g., Euler angles, is problematic at several levels: the local coordinates contain singularities that require special treatment, and the resulting estimates depend both on the choice of local coordinates as well as fixed and moving reference frames. The result is that filtering performance is highly uneven throughout different regions of the configuration space, and dependent on how reference frames are assigned.

Recent research has attempted to address the issue of coordinate dependency through the use of coordinate-invariant differential geometric methods. Although computationally more involved than standard vector space filtering algorithms, when correctly formulated these methods are independent of the choice of local coordinates used to parametrize the rotations, and also invariant with respect to the choice of fixed and moving reference frames. For estimation problems in which the underlying configuration space has the structure of a matrix Lie group (for example, the group of rotation matrices), coordinate-invariant versions of both the EKF [12]–[15], the UKF [16], [17], and also particle filtering methods [18] have been presented in the recent literature. Without exception, these general methods almost always include illustrative examples involving estimation on the rotation group, e.g., [14].

In this paper we address the problem of simultaneous estimation of attitude and gyroscope bias from a stochastic differential geometric perspective. When the assumed noise models are valid, the advantages of stochastic filtering methods over their deterministic counterparts are well-documented. For real-time applications, stochastic filtering methods require efficient calculation and propagation of covariances, which are often difficult to achieve for systems with complex nonlinear state dynamics and measurements. Our contribution takes advantage of the coordinate- and frame-invariant properties of geometric filtering, and at the same time leads to a robust and computationally efficient stochastic UKF algorithm that can be implemented in real-time. These improvements in

This work was supported in part by the SNU BMRR Center (DAPA-UD130070ID), SNU-IAMD, BK21+, MI Technology Innovation Program (10048320), Soft Robotics Research Center, and KIST institutional project (Corresponding author: Frank C. Park).

D. Kang is with the Imaging Media Center, Korea Institute of Science and Technology (KIST), Seoul, Korea (e-mail: kimbab.moowoo@gmail.com).

C. Jang and F.C. Park are with the Department of Mechanical and Aerospace Engineering, Seoul National University, Seoul, Korea (e-mail: jchastro@gmail.com; fcp@snu.ac.kr).

efficiency and robustness are achieved by formulating the state dynamics and measurements in a way that leads to closed-form equations for covariance propagation and correction, and also by drawing upon efficient geometric algorithms for key steps of the intrinsic UKF algorithm.

The paper is organized as follows. After a brief review of geometric preliminaries in Section II, our UKF algorithm for simultaneously estimating attitude and gyroscope bias is described in Section III. Section IV details the calculation of the measurement noise covariance. In Section V, we present extensive hardware and numerical simulation experiments involving both synthetic and real data, whose results are compared against other existing state-of-the-art estimators for attitude and gyroscope bias [6], [19], [20], to demonstrate the numerous performance advantages of our geometric UKF algorithm.

II. Geometric Preliminaries

We first recall some basic facts and useful formulas about the rotation group SO(3) [21], [22]. Elements of SO(3) are given by the 3×3 real matrices $\mathbf{R}$ satisfying $\mathbf{R}^T\mathbf{R} = \mathbf{I}$ and $\det \mathbf{R} = 1$, where $\mathbf{I}$ here denotes the 3×3 identity matrix. SO(3) is an example of a matrix Lie group, and its associated Lie algebra, denoted so(3), is given by the set of 3×3 real skew-symmetric matrices of the form

$$[\boldsymbol{\omega}] = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix},$$

where $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)^T \in \mathbb{R}^3$. An important connection between so(3) and SO(3) is via the matrix exponential map $\exp : \text{so(3)} \to \text{SO(3)}$:

$$\begin{aligned} \exp([\boldsymbol{\omega}]) &= \sum_{m=0}^{\infty} \frac{[\boldsymbol{\omega}]^m}{m!} \\ &= \mathbf{I} + \frac{\sin\|\boldsymbol{\omega}\|}{\|\boldsymbol{\omega}\|}[\boldsymbol{\omega}] + \frac{1-\cos\|\boldsymbol{\omega}\|}{\|\boldsymbol{\omega}\|^2}[\boldsymbol{\omega}]^2, \end{aligned}$$

where $\|\cdot\|$ represents the standard vector norm. The inverse of the exponential, or logarithm, of SO(3) is defined as follows: for any $\mathbf{R} \in \text{SO(3)}$ such that $\text{tr}(\mathbf{R}) \neq -1$,

$$\log \mathbf{R} = \frac{\theta}{2\sin\theta}(\mathbf{R} - \mathbf{R}^T),$$

where θ satisfies $1 + 2\cos\theta = \text{tr}(\mathbf{R}), |\theta| < \pi$ (here $\text{tr}(\cdot)$ denotes the trace of a matrix). If $\text{tr}(\mathbf{R}) = -1$, then the equation $\log \mathbf{R} = [\boldsymbol{\omega}]$ has two antipodal solutions $\pm\boldsymbol{\omega}$ that can be determined from the relation $\mathbf{R} = \mathbf{I} + (2/\pi^2)[\boldsymbol{\omega}]^2$. A straightforward calculation also establishes that $\|\log \mathbf{R}\|/\sqrt{2} = \theta$, where $\|\cdot\|$ denotes the Frobenius matrix norm.

The natural way to measure distances between two rotations $\mathbf{R}_1$ and $\mathbf{R}_2$ is via the formula

$$d(\mathbf{R}_1, \mathbf{R}_2) = \|\log(\mathbf{R}_1^T\mathbf{R}_2)\|.$$

The above distance metric is invariant with respect to left and right translations, or bi-invariant, in the sense that $d(\mathbf{R}_1, \mathbf{R}_2) = d(\mathbf{P}\mathbf{R}_1\mathbf{Q}, \mathbf{P}\mathbf{R}_2\mathbf{Q})$ for any $\mathbf{P}, \mathbf{Q} \in \text{SO(3)}$. The curve $\mathbf{R}(t)$ on SO(3) of shortest length (or minimal geodesic) that connects $\mathbf{R}_1 = \mathbf{R}(0)$ and $\mathbf{R}_2 = \mathbf{R}(1)$ is given by $\mathbf{R}(t) = \mathbf{R}_1\exp(\boldsymbol{\Omega}t)$, where $\boldsymbol{\Omega} = \log(\mathbf{R}_1^T\mathbf{R}_2) \in \text{so(3)}$.

Recalling that $\mathbb{R}^3$ is also trivially a Lie group under vector addition, the direct product $\text{SO(3)} \times \mathbb{R}^3$ can be given the structure of a Lie group via the product rule $(\mathbf{R}_1, \mathbf{b}_1) \cdot (\mathbf{R}_2, \mathbf{b}_2) = (\mathbf{R}_1\mathbf{R}_2, \mathbf{b}_1 + \mathbf{b}_2)$ and the inversion rule $(\mathbf{R}, \mathbf{b})^{-1} = (\mathbf{R}^T, -\mathbf{b})$.

Now define a random variable $\mathbf{X}$ on SO(3) as

$$\mathbf{X} := \exp([\boldsymbol{\eta}])\,\mathbf{X}_0, \tag{1}$$

where $\mathbf{X}_0 \in \text{SO(3)}$ is given and $\boldsymbol{\eta} \in \mathbb{R}^3$ is a zero-mean Gaussian, i.e. $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_\eta)$, where $\mathbf{P}_\eta$ denotes the covariance of $\boldsymbol{\eta}$. We refer to $\boldsymbol{\eta}$ as ***right-translated exponential noise*** with **right-invariant covariance** $\mathbf{P}_\eta$. Alternatively, defining the random variable $\mathbf{X}$ on SO(3) as $\mathbf{X} = \mathbf{X}_0\exp([\boldsymbol{\zeta}])$, where $[\boldsymbol{\zeta}] \in \text{so(3)}$ and $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_\zeta)$, we refer to $\boldsymbol{\zeta}$ as ***left-translated exponential noise*** with **left-invariant covariance** $\mathbf{P}_\zeta$. A straightforward calculation verifies that

$$\boldsymbol{\eta} = \mathbf{X}_0\boldsymbol{\zeta} \tag{2}$$

$$\mathbf{P}_\eta = \mathbf{X}_0\mathbf{P}_\zeta\mathbf{X}_0^T \tag{3}$$

Statistical and computational aspects of SO(3) exponential noise defined in this way are further discussed in [23], [24].

Now consider the element $(\mathbf{X}, \mathbf{b}) = (\exp([\boldsymbol{\eta}])\mathbf{X}_0, \mathbf{b}_0 + \mathbf{n}) \in \text{SO(3)} \times \mathbb{R}^3$, where $[\boldsymbol{\eta}] \in \text{so(3)}, \mathbf{X}_0 \in \text{SO(3)}$, and $\mathbf{b}_0, \mathbf{n} \in \mathbb{R}^3$, with $\mathbf{X}_0, \mathbf{b}_0$ constant and $\boldsymbol{\eta}, \mathbf{n}$ zero-mean Gaussian random vectors. Define the six-dimensional zero-mean Gaussian $\boldsymbol{\epsilon} = (\boldsymbol{\eta}, \mathbf{n}) \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_\epsilon)$, where $\mathbf{P}_\epsilon \in \mathbb{R}^{6\times6}$ is the covariance of $\boldsymbol{\epsilon}$. The six-dimensional covariance $\mathbf{P}_\epsilon$ will play a prominent role in our proposed UKF algorithm; in particular, the off-diagonal elements of $\mathbf{P}_\epsilon$ will typically be non-zero since $\mathbf{X}$ and $\mathbf{b}$ may be correlated.

III. UKF Algorithm for Estimating Attitude and Gyroscope Bias

Before describing our geometric UKF algorithm for IMU-based attitude and gyroscope bias estimation, we fix notation, describe the sensor models and their underlying assumptions, and review Wahba's problem [25] and its solutions.

Let $\{\mathcal{I}\}$ denote the inertial reference frame fixed to ground, and $\{\mathcal{B}\}$ denote the body frame fixed to the moving IMU. A typical IMU is equipped with gyroscopes, accelerometers, and magnetometers; gyroscopes measure the angular velocity vector $\boldsymbol{\omega}^m \in \mathbb{R}^3$, while the accelerometer measures the gravity vector (with other accelerations) and the magnetometer the local geomagnetic field; all measurements are with respect to the IMU frame $\{\mathcal{B}\}$. Assuming the IMU moves with low acceleration and is suitably calibrated, denote by $\mathbf{a}, \mathbf{m} \in \mathbb{R}^3$ the respective outputs of the accelerometer and magnetometer, and define the unit vectors $\mathbf{v}_1 := \mathbf{a}/\|\mathbf{a}\|$, $\mathbf{v}_2 := \mathbf{m}/\|\mathbf{m}\|$.

If the negative direction of gravity and the direction of the geomagnetic field expressed in the fixed frame $\{\mathcal{I}\}$ are respectively denoted by unit vectors $\mathbf{r}_1, \mathbf{r}_2 \in \mathbb{R}^3$, then assuming that $\mathbf{r}_1$ and $\mathbf{r}_2$ are not collinear and the measurements are perfect (i.e., no noise), $\mathbf{r}_i$ and $\mathbf{v}_i$ should then satisfy $\mathbf{r}_i = \mathbf{R}\mathbf{v}_i$, $i = 1, 2$, where $\mathbf{R} \in \text{SO(3)}$ is the orientation of the IMU frame $\{\mathcal{B}\}$ relative to the fixed frame $\{\mathcal{I}\}$. Since in practice

the measurements are noisy, $\mathbf{R}$ is typically estimated via the following optimization (referred to in the literature as Wahba's problem [5], [25], [26]:

$$\mathbf{R}^* = \underset{\mathbf{R}\in \mathrm{SO(3)}}{\arg\min} \sum_{i=1}^{2} w_i \|\mathbf{r}_i - \mathbf{R}\mathbf{v}_i\|^2, \tag{4}$$

where w_i denote positive weights. A common choice for w_i is $w_i = 1/\sigma_i^2$, where σ_i^2 denotes the variance of $\mathbf{v}_i$ along the direction normal to $\mathbf{R}^T\mathbf{r}_i$ [27]. (Equivalently, the normalized weights $w_i = \sigma_{\text{tot}}^2/\sigma_i^2$, where $1/\sigma_{\text{tot}}^2 = \sum_{i=1}^{2}(1/\sigma_i^2)$ are also commonly used [28].)

Wahba's problem as defined by Equation (4) admits the following closed-form solution [26]:

$$\mathbf{R}^* = \mathbf{V}\mathbf{D}\mathbf{U}^T, \tag{5}$$

where $\mathbf{U}$ and $\mathbf{V}$ are obtained from the singular value decomposition (SVD) of $\mathbf{F} := \sum_{i=1}^{2} w_i \mathbf{v}_i \mathbf{r}_i^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$. The matrix $\mathbf{D}$ in Equation (5) is of the form $\mathbf{D} = \mathrm{diag}(1, 1, \det(\mathbf{V}\mathbf{U}^T))$. (See [5] for a review of alternative solutions to (5) and a discussion of their robustness and computational efficiency.)

A. State Space Dynamics and Measurements

Estimates of $\mathbf{R}$ obtained via the static optimization procedure described above do not take into account the underlying motion dynamics or noise characteristics, and typically are inferior to estimates obtained via nonlinear stochastic filtering techniques. In this paper, we frame the estimation procedure as a discrete-time stochastic filtering problem. First, the angular rates $\boldsymbol{\omega}_k^m \in \mathbb{R}^3$ measured by the gyroscope at time step k are assumed to have the form

$$\boldsymbol{\omega}_k^m = \boldsymbol{\omega}_k + \mathbf{b}_k + \boldsymbol{\eta}_k, \tag{6}$$

where $\boldsymbol{\omega}_k$ denotes the ground truth angular rate vector, $\mathbf{b}_k \in \mathbb{R}^3$ is a time-varying bias term, and $\boldsymbol{\eta}_k$ is a zero-mean Gaussian noise. The state dynamics are then given by

$$\mathbf{R}_{k+1} = \mathbf{R}_k \exp([\boldsymbol{\omega}_k^m - \mathbf{b}_k - \boldsymbol{\eta}_k]h) \tag{7}$$

$$\mathbf{b}_{k+1} = \mathbf{b}_k + \mathbf{n}_k, \tag{8}$$

where h is the integration time interval, and $\boldsymbol{\eta}_k, \mathbf{n}_k$ are independent zero-mean Gaussians with the following distributions: $\boldsymbol{\eta}_k \sim \mathcal{N}(\mathbf{0}, c\mathbf{I})$, $\mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, d\mathbf{I})$, with $c, d > 0$.

We now derive an approximation of the state dynamics equation (7) that leads to a closed-form expression for the covariance of $\mathbf{R}_{k+1}$ that is geometrically consistent with (1). From the Baker-Campbell-Hausdorff formula [29], given $[x], [y] \in$ so(3), $\exp([\mathbf{x}])\exp([\mathbf{y}])$ can be written exactly in the form $\exp([\mathbf{x}])\exp([\mathbf{y}]) = \exp([\mathbf{z}])$, $[z] \in$ so(3), where

$$[\mathbf{z}] = \log(\exp([\mathbf{x}])\exp([\mathbf{y}])) \tag{9}$$

$$= [\mathbf{x}] + [\mathbf{y}] + \frac{1}{2}[[\mathbf{x}], [\mathbf{y}]] + \frac{1}{12}[[\mathbf{x}], [[\mathbf{x}], [\mathbf{y}]]] + \frac{1}{12}[[\mathbf{y}], [[\mathbf{y}], [\mathbf{x}]]] + \cdots, \tag{10}$$

with the Lie bracket operator $[\cdot, \cdot] : \mathrm{so(3)} \times \mathrm{so(3)} \to \mathrm{so(3)}$ defined by the matrix commutator, i.e., $[[\mathbf{a}], [\mathbf{b}]] = [\mathbf{a}][\mathbf{b}] - [\mathbf{b}][\mathbf{a}]$.

Let $\mathbf{x}' = \mathbf{z} - \mathbf{y} \in \mathbb{R}^3$ and rewrite Equation (9) in the form

$$\exp([\mathbf{x}' + \mathbf{y}]) = \exp([\mathbf{x}])\exp([\mathbf{y}]). \tag{11}$$

If we further assume that $\|\mathbf{x}\|$ is small, then by gathering only terms linear in $\mathbf{x}$ in Equation (10), the following approximation between $\mathbf{x}$ and $\mathbf{x}'$ holds [23]:

$$\mathbf{x} \approx \mathbf{J}_l(\mathbf{y})\mathbf{x}', \tag{12}$$

where $\mathbf{J}_l(\mathbf{y}) \in \mathbb{R}^{3\times 3}$ is given by

$$\mathbf{J}_l(\mathbf{y}) = \mathbf{I} + \left(\frac{1 - \cos\|\mathbf{y}\|}{\|\mathbf{y}\|^2}\right)[\mathbf{y}] + \left(\frac{\|\mathbf{y}\| - \sin\|\mathbf{y}\|}{\|\mathbf{y}\|^3}\right)[\mathbf{y}]^2. \tag{13}$$

The derivation of (12) is provided in Appendix A.

If $\|\boldsymbol{\eta}_k\| \ll 1$, then Equation (7) can be approximated by

$$\mathbf{R}_{k+1} \approx \mathbf{R}_k \exp([\boldsymbol{\eta}_k']h)\exp([\boldsymbol{\omega}_k^m - \mathbf{b}_k]h) \tag{14}$$

$$= \exp([\mathbf{l}_k])\mathbf{R}_k \exp([\boldsymbol{\omega}_k^m - \mathbf{b}_k]h), \tag{15}$$

where $\boldsymbol{\eta}_k' = -\mathbf{J}_l(\boldsymbol{\psi})\boldsymbol{\eta}_k$, $\mathbf{l}_k = \mathbf{R}_k\boldsymbol{\eta}_k' h$, $\boldsymbol{\psi} = (\boldsymbol{\omega}_k^m - \mathbf{b}_k)h$. In deriving (14), the first-order approximation given by Equation (12) is used. The relation $\mathbf{R}\exp([\boldsymbol{\omega}])\mathbf{R}^T = \exp([\mathbf{R}\boldsymbol{\omega}])$ for $\mathbf{R} \in$ SO(3), $[\boldsymbol{\omega}] \in$ so(3) is used to derive (15).

Note that $\mathbf{l}_k = -\mathbf{R}_k\mathbf{J}_l(\boldsymbol{\psi})\boldsymbol{\eta}_k h$ is also a random variable, as it is given as a function of random variables $\boldsymbol{\eta}_k, \mathbf{R}_k$, and $\boldsymbol{\psi}$. If we assume that $\|\boldsymbol{\psi}\|$ is small—this is a reasonable assumption provided h is sufficiently small—then $\mathbf{J}_l(\boldsymbol{\psi}) \approx \mathbf{I} + \frac{1}{2}[\boldsymbol{\psi}] \approx \exp(\frac{1}{2}[\boldsymbol{\psi}])$ holds from the first-order approximation. Note that $\mathbf{l}_k$ can be approximated as an isotropic Gaussian multiplied by rotation matrices, i.e., $\mathbf{l}_k \sim \mathcal{N}(\mathbf{0}, (ch^2)\mathbf{I})$.

The measurement function is taken to be of the form

$$\mathbf{Y}_{k+1} = \exp([\mathbf{w}_{k+1}])\mathbf{R}_{k+1}, \tag{16}$$

where $\mathbf{Y}_{k+1} \in$ SO(3) is calculated from the simultaneous measurements of the gravity and geomagnetic field vectors as a solution to Wahba's problem (4). In Equation (16), we assume that the measurement noise, $\mathbf{w}_{k+1} \in \mathbb{R}^3$ has the form of a zero-mean Gaussian vector, which implies that the statistics of measurement vectors are assumed to be rotationally symmetric about their true measurement vectors.

B. UKF Algorithm

We now present the geometric UKF algorithm for simultaneously estimating the attitude and gyroscope bias. The state at time step k is defined as $\mathbf{X}_k := (\mathbf{R}_k, \mathbf{b}_k) \in \mathrm{SO(3)} \times \mathbb{R}^3$, where $\mathbf{R}_k$ and $\mathbf{b}_k$ denote the attitude and the gyroscope bias, respectively.

1) Initialization: Let $\hat{\mathbf{X}}_{0|0} = (\hat{\mathbf{R}}_{0|0}, \hat{\mathbf{b}}_{0|0})$ be the initial state estimate. The right-invariant covariance of $\hat{\mathbf{X}}_{0|0}$, denoted $\hat{\mathbf{P}}_{0|0}$, is given. From Equation (5), $\hat{\mathbf{R}}_{0|0}$ can be estimated by solving Wahba's problem (4) from a pair of initial measurement vectors $(\mathbf{v}_1, \mathbf{v}_2)$.

2) Time Update:

- From the *a priori* state estimate $\hat{\mathbf{X}}_{k|k} = (\hat{\mathbf{R}}_{k|k}, \hat{\mathbf{b}}_{k|k})$ and its covariance $\mathbf{P}_{k|k}$, let us extract a set of sigma points, denoted $\mathcal{X}_k^{(i)} := (\mathcal{X}_{\mathbf{R},k}^{(i)}, \mathcal{X}_{\mathbf{b},k}^{(i)}) \in \text{SO(3)} \times \mathbb{R}^3$, $i = 0, \ldots, 12$, as follows:

$$\begin{aligned}\mathcal{X}_k^{(0)} &= (\hat{\mathbf{R}}_{k|k}, \hat{\mathbf{b}}_{k|k})\\ \mathcal{X}_k^{(i)} &= (\exp([\gamma \mathbf{s}_i^{(a)}])\hat{\mathbf{R}}_{k|k}, \hat{\mathbf{b}}_{k|k} + \gamma \mathbf{s}_i^{(b)}), i = 1, \ldots, 6\\ \mathcal{X}_k^{(i+6)} &= (\exp([-\gamma \mathbf{s}_i^{(a)}])\hat{\mathbf{R}}_{k|k}, \hat{\mathbf{b}}_{k|k} - \gamma \mathbf{s}_i^{(b)}), i = 1, \ldots, 6,\end{aligned}$$

where $\gamma = \sqrt{N_x + \lambda}$ is a design parameter with $\lambda = N_x(\alpha^2 - 1)$ and $0 < \alpha < 1$ [10]. Here N_x denotes the dimension of the state, i.e., $N_x = 6$. Let $\mathbf{s}_i \in \mathbb{R}^6$ denote the i^{th} column vector of the lower triangular matrix $\mathbf{S} \in \mathbb{R}^{6\times 6}$, where $\mathbf{S}$ is given by the Cholesky decomposition of $\mathbf{P}_{k|k} = \mathbf{S}\mathbf{S}^T$. In the above equations, $\mathbf{s}_i^{(a)}, \mathbf{s}_i^{(b)} \in \mathbb{R}^3$ refer to the upper and lower halves of $\mathbf{s}_i \in \mathbb{R}^6$, respectively.
- Setting $\mathbf{l}_k = \mathbf{0}$ in Equation (15) and $\mathbf{n}_k = \mathbf{0}$ in Equation (8) allows us to define a set of sigma points $\{(\Upsilon_{\mathbf{R},k+1}^{(i)}, \Upsilon_{\mathbf{b},k+1}^{(i)}) \in \text{SO(3)} \times \mathbb{R}^3 | i = 0, \ldots, 12\}$ as

$$\Upsilon_{\mathbf{R},k+1}^{(i)} = \mathcal{X}_{\mathbf{R},k}^{(i)} \exp([\boldsymbol{\omega}_k^m - \mathcal{X}_{\mathbf{b},k}^{(i)}]h) \tag{17}$$

$$\Upsilon_{\mathbf{b},k+1}^{(i)} = \mathcal{X}_{\mathbf{b},k}^{(i)}. \tag{18}$$

- Given a set of matrices $\{\Upsilon_{\mathbf{R},k+1}^{(0)}, \ldots, \Upsilon_{\mathbf{R},k+1}^{(12)}\}$ in SO(3), the weighted mean rotation, denoted $\bar{\Upsilon}_{\mathbf{R},k+1} \in \text{SO(3)}$, can be computed by using Algorithm 1. Since Algorithm 1 is known to converge rapidly in the general case [18], [30], we set $n = 3$ or 4, where n in line 2 of Algorithm 1 represents the number of iterations. The $w_m^{(i)} \in \mathbb{R}$ in line 3 represent normalized weights that satisfy $\sum_{i=0}^{12} w_m^{(i)} = 1$.
- The weighted mean of $\{\Upsilon_{\mathbf{b},k+1}^{(0)}, \ldots, \Upsilon_{\mathbf{b},k+1}^{(12)}\}$ in $\mathbb{R}^3$, denoted $\bar{\Upsilon}_{\mathbf{b},k+1} \in \mathbb{R}^3$ and representing the gyroscope bias term in terms of propagated sigma points, is trivially calculated as $\bar{\Upsilon}_{\mathbf{b},k+1} = \sum_{i=0}^{12} w_m^{(i)} \Upsilon_{\mathbf{b}}$. Thus, $\hat{\mathbf{X}}_{k+1|k} := (\hat{\mathbf{R}}_{k+1|k}, \hat{\mathbf{b}}_{k+1|k})$ is given by

$$(\hat{\mathbf{R}}_{k+1|k}, \hat{\mathbf{b}}_{k+1|k}) = (\bar{\Upsilon}_{\mathbf{R},k+1}, \bar{\Upsilon}_{\mathbf{b},k+1}). \tag{19}$$

Algorithm 1: Weighted Intrinsic Mean on SO(3)

Input: a set of rotations $\{\mathbf{Z}_0, \ldots, \mathbf{Z}_{12}\}$ in SO(3)

1 $\mathbf{T} \leftarrow \mathbf{Z}_0$
2 **for** $j \leftarrow 0$ ***to*** n **do**
3 $\quad \boldsymbol{\Lambda} \leftarrow \sum_{i=0}^{12} w_m^{(i)} \log(\mathbf{Z}_i \mathbf{T}^{-1})$
4 $\quad \mathbf{T} \leftarrow \exp(\boldsymbol{\Lambda})\, \mathbf{T}$
5 **return** $\mathbf{T}$

- Define the vectors $[\mathbf{q}_i^{(a)}] := \log(\Upsilon_{\mathbf{R},k+1}^{(i)} \bar{\Upsilon}_{\mathbf{R},k+1}^{-1}) \in \text{so(3)}$ and $\mathbf{q}_i^{(b)} := \Upsilon_{\mathbf{b},k+1}^{(i)} - \bar{\Upsilon}_{\mathbf{b},k+1}$. Concatenate the two vectors $\mathbf{q}_i^{(a)}, \mathbf{q}_i^{(b)}$ into a single vector $\mathbf{q}_i = (\mathbf{q}_i^{(a)}, \mathbf{q}_i^{(b)}) \in \mathbb{R}^6$. Then the predicted covariance is given by

$$\mathbf{P}_{k+1|k} = \sum_{i=0}^{12} w_c^{(i)} \mathbf{q}_i \mathbf{q}_i^T + \mathbf{N}_k, \tag{20}$$

where $w_c^{(i)} \in \mathbb{R}$ are the weights and $\mathbf{N}_k = \begin{bmatrix} {}^{(ch^2)}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & {}_{d}\mathbf{I} \end{bmatrix}$ is the process noise covariance.
- Let $\mathbf{u}_i \in \mathbb{R}^6$ denote the i^{th} column vector of the lower triangular matrix $\mathbf{U} \in \mathbb{R}^{6\times 6}$, where $\mathbf{U}$ can be calculated from the Cholesky decomposition of $\mathbf{P}_{k+1|k} = \mathbf{U}\mathbf{U}^T$. The upper and lower halves of $\mathbf{u}_i$ are respectively denoted $\mathbf{u}_i^{(a)} \in \mathbb{R}^3$ and $\mathbf{u}_i^{(b)} \in \mathbb{R}^3$.
A set of sigma points $\mathcal{X}_{k+1}^{(i)} := (\mathcal{X}_{\mathbf{R},k+1}^{(i)}, \mathcal{X}_{\mathbf{b},k+1}^{(i)})$, $(i = 0, \ldots, 12)$ is now redrawn from $\hat{\mathbf{X}}_{k+1|k}$ and $\mathbf{P}_{k+1|k}$:

$$\begin{aligned}\mathcal{X}_{k+1}^{(0)} &= (\hat{\mathbf{R}}_{k+1|k}, \hat{\mathbf{b}}_{k+1|k})\\ \mathcal{X}_{k+1}^{(i)} &= (\exp([\gamma \mathbf{u}_i^{(a)}])\hat{\mathbf{R}}_{k|k}, \hat{\mathbf{b}}_{k|k} + \gamma \mathbf{u}_i^{(b)}), i = 1, \ldots, 6\\ \mathcal{X}_{k+1}^{(i+6)} &= (\exp([-\gamma \mathbf{u}_i^{(a)}])\hat{\mathbf{R}}_{k|k}, \hat{\mathbf{b}}_{k|k} - \gamma \mathbf{u}_i^{(b)}), i = 1, \ldots, 6.\end{aligned}$$

3) Measurement Update:

- When the IMU moves with high acceleration or is subject to magnetic disturbances, the accelerometer and magnetometer measurements are often corrupted. Appendix C summarizes a subset of existing methods for addressing these disturbances.
- Setting $\mathbf{w}_{k+1} = \mathbf{0}$ in Equation (16), define the set of measurement sigma points $\mathcal{S}_{\mathcal{Y}} = \{\mathcal{Y}_{k+1}^{(i)} \in \text{SO(3)} | i = 0, \ldots, 12\}$ as

$$\mathcal{Y}_{k+1}^{(i)} = \mathcal{X}_{\mathbf{R},k+1}^{(i)} \quad (i = 0, \ldots, 12). \tag{21}$$

- The mean of $\mathcal{Y}_{k+1}^{(0)}, \ldots, \mathcal{Y}_{k+1}^{(12)}$, denoted $\hat{\mathbf{Y}}_{k+1}$, is simply

$$\hat{\mathbf{Y}}_{k+1} = \hat{\mathbf{R}}_{k+1|k}, \tag{22}$$

where $\hat{\mathbf{R}}_{k+1|k}$ is given by Equation (19). The covariance of $\mathcal{Y}_{k+1}^{(0)}, \ldots, \mathcal{Y}_{k+1}^{(12)}$ is computed by

$$\mathbf{P}_{\mathbf{yy}} = \sum_{i=0}^{12} w_c^{(i)} \mathbf{z}_i \mathbf{z}_i^T, \tag{23}$$

where $[\mathbf{z}_i] := \log(\mathcal{Y}_{k+1}^{(i)} \hat{\mathbf{Y}}_{k+1}^{-1}) \in \text{so(3)}$. The innovation covariance [9] is given by

$$\mathbf{P}_{\mathbf{vv}} = \mathbf{P}_{\mathbf{yy}} + \mathbf{W}_{k+1}, \tag{24}$$

where $\mathbf{W}_{k+1}$ is the right-invariant covariance of the solution to Wahba's problem. In the next section, we will show that $\mathbf{W}_{k+1}$ can be obtained in closed-form by using Equation (32).
- Define the vectors $[\mathbf{p}_i^{(a)}] := \log(\mathcal{X}_{\mathbf{R},k+1}^{(i)} \hat{\mathbf{R}}_{k+1|k}^{-1}) \in \text{so(3)}$ and $\mathbf{p}_i^{(b)} := \mathcal{X}_{\mathbf{b},k+1}^{(i)} - \hat{\mathbf{b}}_{k+1|k} \in \mathbb{R}^3$. The two vectors $\mathbf{p}_i^{(a)}, \mathbf{p}_i^{(b)}$ can be concatenated into a single vector $\mathbf{p}_i \in \mathbb{R}^6$. The associated covariance $\mathbf{P}_{\mathbf{xy}}$ is then calculated as

$$\mathbf{P}_{\mathbf{xy}} = \sum_{i=0}^{12} w_c^{(i)} \mathbf{p}_i \mathbf{z}_i^T. \tag{25}$$

- The Kalman gain is computed as $\mathbf{K} = \mathbf{P}_{\mathbf{xy}} \mathbf{P}_{\mathbf{vv}}^{-1}$. Define the innovation vector $\boldsymbol{\delta} \in \mathbb{R}^3$ as

$$[\boldsymbol{\delta}] := \log(\mathbf{Y}_{k+1} \hat{\mathbf{Y}}_{k+1}^{-1}) \in \text{so(3)}. \tag{26}$$

Recall that $\mathbf{Y}_{k+1}$ and $\hat{\mathbf{Y}}_{k+1}$ are given by Equations (16) and (22), respectively. Define $\boldsymbol{\phi}^{(a)} \in \mathbb{R}^3$ and $\boldsymbol{\phi}^{(b)} \in \mathbb{R}^3$

to be upper and lower halves of $\boldsymbol{\phi} := \mathbf{K}\boldsymbol{\delta} \in \mathbb{R}^6$. The state and covariance can now be updated according to

$$\hat{\mathbf{X}}_{k+1|k+1} = (\exp([\boldsymbol{\phi}^{(a)}])\hat{\mathbf{R}}_{k+1|k}, \hat{\mathbf{b}}_{k+1|k} + \boldsymbol{\phi}^{(b)}), \quad (27)$$

$$\mathbf{P}_{k+1|k+1} = \mathbf{M}(\boldsymbol{\phi}^{(a)})(\mathbf{P}_{k+1|k} - \mathbf{K}\mathbf{P}_{\mathbf{yy}}\mathbf{K}^T)\mathbf{M}(\boldsymbol{\phi}^{(a)})^T, \quad (28)$$

where $\mathbf{M}(\boldsymbol{\phi}^{(a)}) \in \mathbb{R}^{6\times 6}$ has the analytic form

$$\mathbf{M}(\boldsymbol{\phi}^{(a)}) = \begin{bmatrix} \mathbf{J}_l(\boldsymbol{\phi}^{(a)}) & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \quad (29)$$

The rationale for adopting $\mathbf{M}(\boldsymbol{\phi}^{(a)})$ in Equation (28) is provided in Appendix B.

IV. Measurement Noise Covariance

This section presents an algorithm for obtaining, from noisy unit vector measurements of the gravity and geomagnetic field vector, the measurement noise covariance with a full rank.

A. Covariances of the Solution to Wahba's Problem

In [27], Shuster provides the following first-order approximation to the left-invariant covariance of $\mathbf{R}$ in Equation (4) (i.e., the solution to Wahba's problem):

$$(\sum_{i=1}^{2} \frac{1}{\sigma_i^2}(\mathbf{I} - \bar{\mathbf{A}}\mathbf{r}_i\mathbf{r}_i^T\bar{\mathbf{A}}^T))^{-1}, \quad (30)$$

where $\bar{\mathbf{A}} \in$ SO(3) denotes the true value of $\mathbf{R}^T$, which is usually unknown. $\bar{\mathbf{A}}$ can be approximated by

$$\bar{\mathbf{A}} \approx \underset{\mathbf{A}\in \mathrm{SO(3)}}{\arg\min} \sum_{i=1}^{2} \frac{1}{\sigma_i^2}\|\mathbf{v}_i - \mathbf{A}\mathbf{r}_i\|^2. \quad (31)$$

We remark that in [27], it is asserted without rigorous proof that the left-invariant covariance of $\mathbf{R}$ is given by the inverse of the Fisher information matrix. Appendix E provides a more detailed and rigorous proof via the Cramer-Rao lower bound (CRLB).

Similarly, from Equation (30) the right-invariant covariance of $\mathbf{R}$ can be obtained as

$$(\sum_{i=1}^{2} \frac{1}{\sigma_i^2}(\mathbf{I} - \mathbf{r}_i\mathbf{r}_i^T))^{-1} \quad (32)$$

The above follows from a straightforward calculation combining Equations (3) and (30).

Note that the left-invariant covariance of $\mathbf{R}$ in Equation (30) is equivalent to the covariance of the solution to Wahba's problem represented with respect to the IMU body frame. In contrast, the right-invariant covariance of $\mathbf{R}$ in Equation (32) is the covariance of the solution to Wahba's problem represented with respect to the fixed ground frame.

If parameter values for $\sigma_i^2, \mathbf{r}_i$ are given, the right-invariant covariance of $\mathbf{R}$ in Equation (32) can be determined as a constant matrix, which is independent from $\bar{\mathbf{A}}$. However, the left-invariant covariance of $\mathbf{R}$ in Equation (30) requires both $\bar{\mathbf{A}}$ as well as σ_i^2 and $\mathbf{r}_i$.

When the IMU is moving, $\bar{\mathbf{A}}$ is also changing; in this situation the left-invariant covariance of $\mathbf{R}$ should be updated at every time step, by computing the inverse of some matrix that varies with $\bar{\mathbf{A}}$, while the right-invariant covariance is invariant. When the IMU motion involves both translation and rotation, estimates of $\bar{\mathbf{A}}$ can become even more inaccurate because of the inherently noisy measurements of the two direction vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. We remark that for fast and accurate computation of the measurement noise covariance in the UKF algorithm, the proposed formula given by Equation (32) is preferable to Shuster's formula (30).

B. Determination of Parameters in the Covariance of $\mathbf{R}$

In this section we present an offline algorithm for determining the parameters in Equation (32), i.e., σ_i^2 and $\mathbf{r}_i$, $i = 1, 2$, from actual accelerometer and magnetometer measurements.

1) Constant Vectors $(\mathbf{r}_1, \mathbf{r}_2)$*:* Without loss of generality, we can assign each axis of the inertial frame $\{\mathcal{I}\}$ as follows: The negative direction of gravity is set to be the y-axis of $\{\mathcal{I}\}$, while the x-axis of $\{\mathcal{I}\}$ is set to be orthogonal to both gravity and the geomagnetic field. In this setting, $\mathbf{r}_1 = (0, 1, 0)^T$ and

$$\mathbf{r}_2 = (0, \cos(\phi), \sin(\phi))^T, \quad (33)$$

where ϕ is unknown and to be determined.

We assume that the IMU is stationary, and multiple measurement pairs are collected. From Proposition 1 in Appendix D, $\hat{\mathbf{v}}_i := E(\mathbf{v}_i)$ can be calculated for $i = 1, 2$. Since $\mathbf{r}_1^T\mathbf{r}_2 \approx \hat{\mathbf{v}}_1^T\hat{\mathbf{v}}_2$, we can approximate

$$\phi \approx \cos^{-1}(\hat{\mathbf{v}}_1^T\hat{\mathbf{v}}_2). \quad (34)$$

2) Variances (σ_1^2, σ_2^2)*:* Let $\breve{\mathbf{v}}_i$ denote the true value of the measured unit vector $\mathbf{v}_i, (i = 1, 2)$, where $\|\breve{\mathbf{v}}_i\| = 1$. The covariance of $\mathbf{v}_i$ [31] is given by

$$\mathbf{M}_t = \sigma_i^2(\mathbf{I} - \breve{\mathbf{v}}_i\breve{\mathbf{v}}_i^T). \quad (35)$$

Let the SVD of $\mathbf{M}_t$ be given by $\mathbf{U}_t\mathbf{\Sigma}_t\mathbf{V}_t^T$, In principle we have $\mathbf{\Sigma}_t = \mathrm{diag}(\sigma_i^2, \sigma_i^2, 0)$ and $\breve{\mathbf{v}}_i$ is the corresponding direction for the singular value 0.

However, in practice the ground truth of $\breve{\mathbf{v}}_i$ is unavailable, so that σ_i^2 should be determined in an alternative way. We assume that the IMU is stationary and N measurements are available. The covariance of $\mathbf{v}_i$ can be estimated by

$$\mathbf{M}_a = \frac{1}{N}\sum_{j=1}^{N}(\mathbf{v}_i^{(j)} - \hat{\mathbf{v}}_i)(\mathbf{v}_i^{(j)} - \hat{\mathbf{v}}_i)^T, \quad (36)$$

where $\mathbf{v}_i^{(j)}$ denotes the j^{th} measurement vector coming from the i^{th} sensor (the first sensor is the accelerometer, while the second sensor is the magnetometer). Let the SVD of $\mathbf{M}_a$ be $\mathbf{U}_a\mathbf{\Sigma}_a\mathbf{V}_a^T$, where $\mathbf{\Sigma}_a = \mathrm{diag}(s_1, s_2, s_3)$ and $s_1 \geq s_2 \geq s_3 \approx 0$. $\mathbf{\Sigma}_a$ will typically be close to its theoretical value $\mathbf{\Sigma}_t$, in which case we can set

$$\sigma_i^2 = \frac{tr(\mathbf{M}_a)}{2}. \quad (37)$$

V. Experimental Results

In this section we compare the performance of the proposed algorithm ("UKF on SO(3)") against other state-of-the-art methods including "UKF on Quaternion" [19], "EKF on Quaternion" [20], and the passive complementary filter among Mahony's nonlinear complementary filters (NCF) on SO(3) ("NCF on SO(3)") [6] in terms of both convergence rate and the estimation accuracy of attitude and gyroscope bias, using both synthetic and real data in our experiments.

The ground-truths of the attitude and gyroscope bias at time step k are denoted $\breve{\mathbf{R}}_k$ and $\breve{\mathbf{b}}_k$, respectively. In both simulations and real experiments, the time interval of filter updates is equally set to $h_0 = 1/60$ seconds. Let us define

$$s_k := (180^\circ/\pi)\|\log \breve{\mathbf{R}}_k^{-1}\hat{\mathbf{R}}_{k|k}\| \tag{38}$$

$$d_k := \|\hat{\mathbf{b}}_{k|k} - \breve{\mathbf{b}}\|, \tag{39}$$

where s_k and d_k represent the estimation errors of the attitude and gyroscope bias at time step k, respectively.

The weighting factors $w_m^{(i)}$ and $w_c^{(i)}$ in Section III-B are given by

$$w_m^{(0)} = \frac{\lambda}{\lambda + N_x}, \quad w_c^{(0)} = \frac{\lambda}{\lambda + N_x} + (1 - \alpha^2 + \beta) \tag{40}$$

$$w_m^{(i)} = w_c^{(i)} = \frac{1}{2(\lambda + N_x)}, \quad (i = 1, \ldots, 2N_x). \tag{41}$$

In Equation (40), we set $\alpha = 0.9$. For a Gaussian prior, $\beta = 2$ [10].

A. Simulation

In the simulation, the vectors in Equation (4) are set to be $\mathbf{r}_1 = (0, 1, 0)^T$ and $\mathbf{r}_2 = (0, \cos(\phi_s), \sin(\phi_s))^T$, where $\phi_s = 2.4$ radian. We can randomly set the ground-truth, $\breve{\mathbf{R}}_1 \in$ SO(3) as an initial attitude.

For realistic simulation, we first collect a set of real angular rate data denoted $\breve{\boldsymbol{\omega}}_k$ from the actual gyroscope (L3G4200D) with a sampling rate, $1/h_0 = 60$ Hz. From $\breve{\mathbf{R}}_1$, true attitude matrices can be iteratively generated by

$$\breve{\mathbf{R}}_{k+1} = \breve{\mathbf{R}}_k \exp([\breve{\boldsymbol{\omega}}_k]h_0).$$

The ground-truth of the initial gyroscope bias is set to be $\breve{\mathbf{b}}_0 = (-0.06, 0.3, 0.3)^T$ radian/seconds. Then, we generate a set of synthetic data as follows:

$$\boldsymbol{\omega}_k^m = \breve{\boldsymbol{\omega}}_k + \breve{\mathbf{b}}_k + \boldsymbol{\eta}_{\omega,k} \tag{42}$$

$$\breve{\mathbf{b}}_k = \breve{\mathbf{b}}_{k-1} + \boldsymbol{\eta}_{\mathbf{b},k-1} \tag{43}$$

$$\mathbf{v}_{1,k} = (\breve{\mathbf{R}}_k^T\mathbf{r}_1 + \boldsymbol{\eta}_{\mathbf{v}_{1,k}})/\|\breve{\mathbf{R}}_k^T\mathbf{r}_1 + \boldsymbol{\eta}_{\mathbf{v}_{1,k}}\| \tag{44}$$

$$\mathbf{v}_{2,k} = (\breve{\mathbf{R}}_k^T\mathbf{r}_2 + \boldsymbol{\eta}_{\mathbf{v}_{2,k}})/\|\breve{\mathbf{R}}_k^T\mathbf{r}_2 + \boldsymbol{\eta}_{\mathbf{v}_{2,k}}\|, \tag{45}$$

where the Gaussian noise vectors have the following distributions: $\boldsymbol{\eta}_{\omega,k} \sim \mathcal{N}(\mathbf{0}, \sigma_0^2\mathbf{I}), \boldsymbol{\eta}_{\mathbf{b},k} \sim \mathcal{N}(\mathbf{0}, \sigma_1^2\mathbf{I}), \boldsymbol{\eta}_{\mathbf{v}_{1,k}} \sim \mathcal{N}(\mathbf{0}, \sigma_2^2\mathbf{I})$, and $\boldsymbol{\eta}_{\mathbf{v}_{2,k}} \sim \mathcal{N}(\mathbf{0}, \sigma_3^2\mathbf{I})$, $k = 1, \ldots, N$. Here $\sigma_0 = (1.1 \times 10^{-3}/h_0)$ radian/seconds, $\sigma_1 = (1.0 \times 10^{-5})$ radian/seconds, $\sigma_2 = 1.00 \times 10^{-2}$, and $\sigma_3 = 1.58 \times 10^{-2}$.

To simulate the large initial estimation errors of gyroscope bias and attitude, we set $\hat{\mathbf{b}}_{1|1} = \mathbf{0}$ and $\hat{\mathbf{R}}_{1|1} = \breve{\mathbf{R}}_1 \exp([\mathbf{a}_1])$, where $\mathbf{a}_1 = (3.13/\sqrt{3})(1, 1, 1)^T$. The noise covariances $\mathbf{N}_k$ in Equation (20) and $\mathbf{W}_{k+1}$ in (24) of the proposed attitude estimator ("UKF on SO(3)") are set as follows: $\mathbf{N}_k = \begin{bmatrix} (\sigma_0 h_0)^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_1^2\mathbf{I} \end{bmatrix}$ and $\mathbf{W}_{k+1} = (\frac{1}{\sigma_2^2}(\mathbf{I} - \mathbf{r}_1\mathbf{r}_1^T) + \frac{1}{\sigma_3^2}(\mathbf{I} - \mathbf{r}_2\mathbf{r}_2^T))^{-1}$.

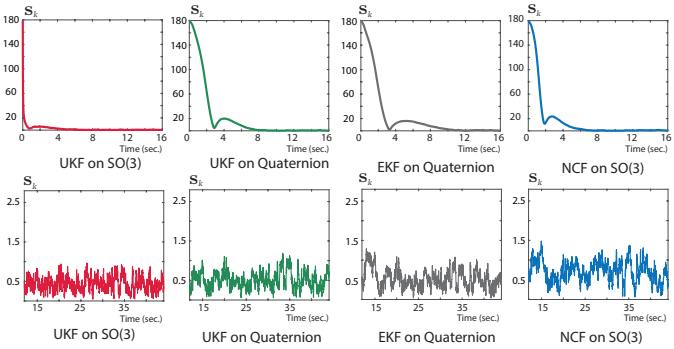


Fig. 1. Simulation: the errors of attitude estimates (in degrees) during time $t \in [0, 16]$ seconds (top); and during $t \in [12, 44]$ seconds (bottom).

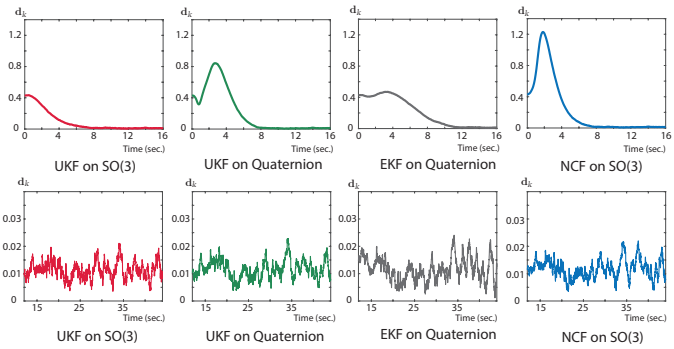


Fig. 2. Simulation: the errors of gyroscope bias estimates (in radian/seconds) during time $t \in [0, 16]$ seconds (top); and during $t \in [12, 44]$ seconds (bottom).

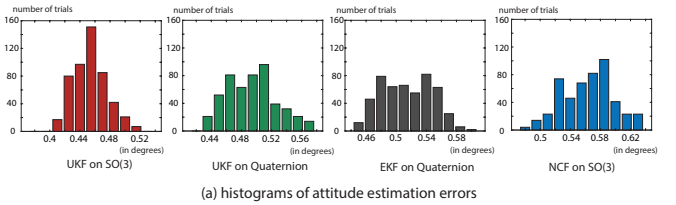
Figures 1 and 2 show the simulation results, from which one may find that the proposed algorithm ("UKF on SO(3)") converges most rapidly over the time interval $t \in [0, 14]$ seconds. To make a reliable accuracy analysis of each estimator, we generate 500 sets of synthetic data using Equations (42)-(45). Figure 3 shows the histograms of estimation errors of the attitudes and the slowly time-varying gyroscope biases. Tables I and II summarize the experimental results corresponding to Figure 3 (a) and (b), respectively. From Figure 3(b) and Table II, it can be seen that all estimators have similar performance in terms of the gyroscope bias estimates. In terms of attitude estimates, "UKF on SO(3)" is the most accurate among the estimators (see Figure 3(a) and Table I).

TABLE I
Average and standard deviation of attitude estimation errors (in degrees) during time $t \in [12, 44]$ seconds over 500 trials.

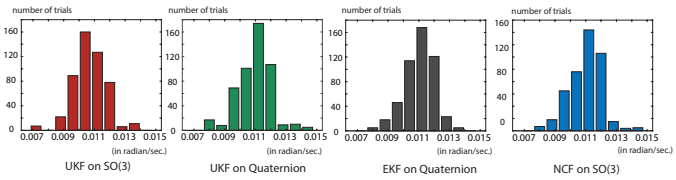
	UKF on SO(3)	UKF on Quaternion	EKF on Quaternion	NCF on SO(3)
Average	**0.45**	0.49	0.51	0.57
Standard deviation	**0.02**	0.03	0.03	0.03

B. Real Experiments

The IMU for real experiments consists of an L3G4200D gyroscope, LIS3LV02DQ accelerometer, HMC5883L magne-



(a) histograms of attitude estimation errors



(b) histograms of gyroscope bias estimation errors

Fig. 3. Simulation: histograms of estimation errors during time $t \in [12, 44]$ seconds over 500 trials.

TABLE II
AVERAGE AND STANDARD DEVIATION OF GYROSCOPE BIAS ESTIMATION ERRORS (IN RADIAN/SECONDS) DURING TIME INTERVAL $t \in [12, 44]$ SECONDS OVER 500 TRIALS.

	UKF on SO(3)	UKF on Quaternion	EKF on Quaternion	NCF on SO(3)
Average	**0.011**	0.011	0.011	0.011
Standard deviation	**0.001**	0.001	0.001	0.001

tometer, and Cortex-M3$^{\text{TM}}$ microcontroller. In real experiments, the ground-truth of the slowly time-varying gyroscope bias is unknown. Thus, we assume that the gyroscope bias is initially unknown but almost constant for a short time duration. If the IMU is static, then the gyroscope bias, denoted $\breve{\mathbf{b}}$, can be temporarily captured by averaging a set of gyroscope data within a certain time interval [32].

By putting the IMU in a static condition and using Equation (37), we can compute the variance σ_i^2 of a unit vector, $\mathbf{v}_{i,k}$, $i = 1, 2$. For this experiment, we obtain $\sigma_1^2 = 8.95 \times 10^{-5}$ and $\sigma_2^2 = 1.911 \times 10^{-4}$. Denoting by ϕ_r the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$: $\phi_r = \cos^{-1}(\mathbf{r}_1^T \mathbf{r}_2)$, in our experiments we find $\phi_r = 2.486$ radians via Proposition 1 of Appendix D and Equation (34). The noise covariances $\mathbf{N}_k$ in Equation (20) and $\mathbf{W}_{k+1}$ in (24) of the proposed attitude estimator ("UKF on SO(3)") are set as follows: $\mathbf{N}_k = \begin{bmatrix} (2.0\times10^{-9})\mathbf{I} & \mathbf{0} \\ \mathbf{0} & (3.0\times10^{-11})\mathbf{I} \end{bmatrix}$ and $\mathbf{W}_{k+1} = (\sum_{i=1}^{2} \frac{1}{\sigma_i^2}(\mathbf{I} - \mathbf{r}_i \mathbf{r}_i^T))^{-1}$.

To obtain the ground-truth of the attitude $\breve{\mathbf{R}}_k$ at time step k, we use the optical motion capture system OptiTrack$^{\text{TM}}$ consisting of multiple networked infrared cameras. The IMU and four reflective markers are first rigidly attached to a plastic plate. A set of real data $\{(\boldsymbol{\omega}_k^m, \mathbf{v}_{1,k}, \mathbf{v}_{2,k}) | k = 1, \ldots, N_r\}$ coming from the moving IMU and the ground-truth attitude $\breve{\mathbf{R}}_k$ obtained by the OptiTrack$^{\text{TM}}$ infrared camera system are synchronously saved into files with a sampling rate, $1/h_0 = 60$ Hz. Here the number of measurements $N_r = 3000$. For fair comparison among filters, we perform experiments with real data under the condition of negligible disturbances.

To evaluate the convergence rate and accuracy of each filter when the initial estimation errors of the gyroscope bias and attitude are large, we set the initial estimates as follows: $\hat{\mathbf{b}}_{1|1} = \breve{\mathbf{b}} + (1/h_0)(-0.001, 0.005, 0.005)^T = \breve{\mathbf{b}} + (-0.06, 0.3, 0.3)^T$ (radian/seconds) and $\hat{\mathbf{R}}_{1|1} \leftarrow \breve{\mathbf{R}}_1 \exp([\mathbf{a}_1])$, where $\mathbf{a}_1 = (3.13/\sqrt{3})(1, 1, 1)^T$. Recall that $\breve{\mathbf{b}}$ can be obtained under the stationary IMU assumption.

Like our earlier simulation results, Figures 4 and 5 show that the proposed method ("UKF on SO(3)") converges the most rapidly, whereas other methods show slow convergence rates and relatively large overshoots. To obtain reliable experimental results, we collected more data, i.e., additional 9 sets of real data. As shown in Table III, "UKF on SO(3)" demonstrates superior performance to existing methods in terms of the accuracy of attitude estimates.

At every time step, we measured the computation time of each filter implemented in C++, and executed on a desktop computer with CPU Intel$^{\text{TM}}$ i5-4670 (3.4GHz). The computation time of each estimator was averaged over N_r steps. Table IV shows that "NCF on SO(3)" is the fastest among the estimators. The computation time of "UKF on SO(3)" is similar to that of "Quaternion UKF".

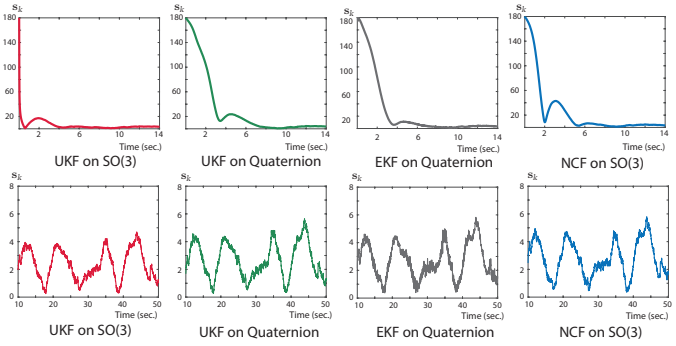


Fig. 4. Real experiments: the errors of attitude estimates (in degrees) during time $t \in [0, 14]$ seconds (top); and during $t \in [10, 50]$ seconds. (bottom).

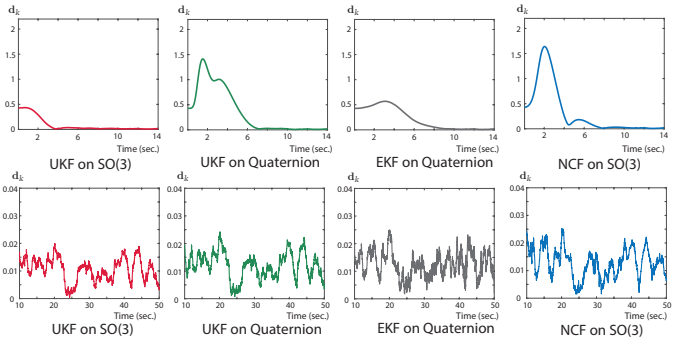


Fig. 5. Real experiments: the errors of gyroscope bias estimates (in radian/seconds) during time $t \in [0, 14]$ seconds (top); and during $t \in [10, 50]$ seconds (bottom).

TABLE III
RESULTS OF REAL EXPERIMENTS: AVERAGE ERRORS DURING TIME $t \in [10, 50]$ SECONDS OVER 10 EXPERIMENTS

Average of attitude errors (in degrees)				Average of gyroscope bias errors (in radian/seconds)			
UKF on SO(3)	UKF on Quaternion	EKF on Quaternion	NCF on SO(3)	**UKF on SO(3)**	UKF on Quaternion	EKF on Quaternion	NCF on SO(3)
2.60	2.69	2.71	2.76	**0.012**	0.012	0.012	0.012

TABLE IV
AVERAGE COMPUTATION TIME OF EACH FILTER (IN MICRO-SECONDS)

	UKF on SO(3)	UKF on Quaternion	EKF on Quaternion	NCF on SO(3)
Average time	**8.1**	7.9	6.8	0.2

VI. CONCLUSION

In this paper, we have presented a UKF algorithm for simultaneously estimating attitude and gyroscope bias. From the closed-form characterization of covariance propagation and correction, computationally efficient implementations of our filter are made possible, and the resulting attitude estimates are invariant with respect to the choice of fixed and moving reference frames. The parameters within the measurement noise covariance have been identified from actual accelerometer and magnetometer measurements. Through simulation and real experiments, the proposed algorithm has been shown to be superior to existing estimators in terms of convergence behavior and estimation accuracy.

APPENDIX A
FIRST ORDER APPROXIMATION OF EXPONENTIAL MAP

Given $[\mathbf{x}], [\mathbf{y}] \in$ so(3), let us consider $[\mathbf{z}] \in$ so(3) which satisfies

$$\exp([\mathbf{z}]) = \exp([\mathbf{x}])\exp([\mathbf{y}]). \tag{46}$$

From the Baker-Campbell-Housdorff formula [29], we have

$$\begin{aligned}[\mathbf{z}] &= \log(\exp([\mathbf{x}])\exp([\mathbf{y}])) \\ &= [\mathbf{x}] + [\mathbf{y}] + \frac{1}{2}[[\mathbf{x}],[\mathbf{y}]] + \frac{1}{12}[[\mathbf{x}],[[\mathbf{x}],[\mathbf{y}]]] \\ &\quad + \frac{1}{12}[[\mathbf{y}],[[\mathbf{y}],[\mathbf{x}]]] + \cdots .\end{aligned}$$

The Lie bracket operator $[\cdot,\cdot] : \text{so(3)} \times \text{so(3)} \to \text{so(3)}$ is defined as $[[\mathbf{a}],[\mathbf{b}]] = [\mathbf{a}][\mathbf{b}] - [\mathbf{b}][\mathbf{a}]$ for $[\mathbf{a}], [\mathbf{b}] \in$ so(3). Thus, $[\mathbf{c}] = [[\mathbf{a}],[\mathbf{b}]] \in$ so(3) admits the following vector representation $\mathbf{c} = [\mathbf{a}]\mathbf{b} \in \mathbb{R}^3$.

If we assume that $\|\mathbf{x}\|$ is small, then by gathering only terms linear in $\mathbf{x}$, the following approximation holds [23]

$$\mathbf{z} \approx \mathbf{y} + \sum_{n=0}^{\infty} \frac{B_n}{n!}[\mathbf{y}]^n \mathbf{x}, \tag{47}$$

where B_n are called the Bernoulli numbers ($B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, ...) in number theory. The Bernoulli numbers satisfy the following series expression: $\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n$ for a scalar $x \neq 0$.

Letting $[\mathbf{x}'] = [\mathbf{z}] - [\mathbf{y}] \in$ so(3), the relationship between $\mathbf{x}$ and $\mathbf{x}'$, which satisfies

$$\exp([\mathbf{x}'] + [\mathbf{y}]) = \exp([\mathbf{x}])\exp([\mathbf{y}]), \tag{48}$$

is given as follows:

$$\mathbf{x} \approx \mathbf{J}_l(\mathbf{y})\mathbf{x}', \tag{49}$$

where

$$\begin{aligned}\mathbf{J}_l(\mathbf{y}) &= (\sum_{n=0}^{\infty} \frac{B_n}{n!}[\mathbf{y}]^n)^{-1} && (50)\\ &= \sum_{n=0}^{\infty} \frac{1}{(n+1)!}[\mathbf{y}]^n && (51)\\ &= \int_0^1 \exp([\mathbf{y}]s)\, ds && (52)\end{aligned}$$

denotes the left Jacobian of SO(3) on $\mathbf{y}$ [23]. The closed-form formula of $\mathbf{J}_l(\mathbf{y})$ is given by

$$\mathbf{J}_l(\mathbf{y}) = \mathbf{I} + \left(\frac{1 - \cos\|\mathbf{y}\|}{\|\mathbf{y}\|^2}\right)[\mathbf{y}] + \left(\frac{\|\mathbf{y}\| - \sin\|\mathbf{y}\|}{\|\mathbf{y}\|^3}\right)[\mathbf{y}]^2. \tag{53}$$

APPENDIX B
COVARIANCE UPDATE IN UKF ON SO(3) $\times \mathbb{R}^3$

From Equation (1), a random variable $\mathbf{R} \in$ SO(3) can be modeled as

$$\mathbf{R} := \exp([\boldsymbol{\varphi}])\,\hat{\mathbf{R}}, \tag{54}$$

where $\boldsymbol{\varphi} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{\boldsymbol{\varphi}})$ is the right-translated exponential noise and $\hat{\mathbf{R}} \in$ SO(3) is the state estimate. In this paper we refer to $\mathbf{P}_{\boldsymbol{\varphi}}$ as the right-invariant covariance of $\mathbf{R}$.

The right-translated exponential noise after the time update as described in Section III-B2 is assumed to be zero-mean Gaussian, whose covariance, denoted $\mathbf{P}_{k+1|k}$, can be calculated by Equation (20). However, a special caution is required when computing $\mathbf{P}_{k+1|k+1}$, which is *a posteriori* right-invariant covariance of $(\mathbf{R}_{k+1}, \mathbf{b}_{k+1})$ after the measurement update. If one implements the measurement update like the UKF on the standard vector space, then state $(\mathbf{R}_{k+1}, \mathbf{b}_{k+1})$ may be given by

$$\begin{aligned}\mathbf{R}_{k+1} &= \exp([\boldsymbol{\xi}^{(a)}])\,\hat{\mathbf{R}}_{k+1|k}, && (55)\\ \mathbf{b}_{k+1} &= \hat{\mathbf{b}}_{k+1|k} + \boldsymbol{\xi}^{(b)} && (56)\end{aligned}$$

where $\boldsymbol{\xi}^{(a)}, \boldsymbol{\xi}^{(b)} \in \mathbb{R}^3$ refer to the upper and lower halves of $\boldsymbol{\xi} \sim \mathcal{N}(\boldsymbol{\phi}, \mathbf{P}_{k+1|k} - \mathbf{K}\mathbf{P}_{\mathbf{yy}}\mathbf{K}^T)$. However, since $\boldsymbol{\phi} \neq \mathbf{0}$ in general, there exists a discrepancy between the random variable models (54) and (55). Thus, we now reformulate (55) to conform to the model in (54) (in order to satisfy a "zero-mean" right-translated exponential noise). Let us assume that $(\mathbf{R}_{k+1}, \mathbf{b}_{k+1})$ can be represented as

$$\begin{aligned}\mathbf{R}_{k+1} &= \exp([\boldsymbol{\epsilon}'^{(a)}])\,\hat{\mathbf{R}}_{k+1|k+1}, && (57)\\ \mathbf{b}_{k+1} &= \hat{\mathbf{b}}_{k+1|k+1} + \boldsymbol{\epsilon}'^{(b)} && (58)\end{aligned}$$

where $\boldsymbol{\epsilon}' \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{\boldsymbol{\epsilon}'})$ and let $\mathbf{P}_{k+1|k+1} = \mathbf{P}_{\boldsymbol{\epsilon}'}$. The task of this section now is to find $\mathbf{P}_{\boldsymbol{\epsilon}'}$.

Let us define a vector $\boldsymbol{\epsilon} \in \mathbb{R}^6$ as $\boldsymbol{\epsilon} := \boldsymbol{\xi} - \boldsymbol{\phi}$. Thus, $\boldsymbol{\epsilon}$ has the following distribution: $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{\boldsymbol{\epsilon}})$, where

$$\mathbf{P}_{\boldsymbol{\epsilon}} = \mathbf{P}_{k+1|k} - \mathbf{K}\mathbf{P}_{\mathbf{yy}}\mathbf{K}^T. \tag{59}$$

Since $\boldsymbol{\xi} = \boldsymbol{\epsilon} + \boldsymbol{\phi}$, Equation (55) can be rewritten as

$$\mathbf{R}_{k+1} = \exp([\boldsymbol{\epsilon}^{(a)} + \boldsymbol{\phi}^{(a)}])\,\hat{\mathbf{R}}_{k+1|k}. \tag{60}$$

By substituting Equation (27) into (57), we have

$$\mathbf{R}_{k+1} = \exp([\boldsymbol{\epsilon}'^{(a)}])\, \exp([\boldsymbol{\phi}^{(a)}])\, \hat{\mathbf{R}}_{k+1|k}. \tag{61}$$

Combining Equations (60) and (61) leads to

$$\exp([\boldsymbol{\epsilon}^{(a)} + \boldsymbol{\phi}^{(a)}]) = \exp([\boldsymbol{\epsilon}'^{(a)}])\, \exp([\boldsymbol{\phi}^{(a)}]), \tag{62}$$

and $\boldsymbol{\epsilon}^{(b)} = \boldsymbol{\xi}^{(b)} - \boldsymbol{\phi}^{(b)} = \boldsymbol{\epsilon}'^{(b)}$ holds from equating (56) and (58) using Equation (27). If $\|\boldsymbol{\epsilon}\| \ll 1$, from the first order approximation derived from the Baker-Campbell-Hausdorff formula in Appendix A, we find

$$\boldsymbol{\epsilon}' \approx \mathbf{M}(\boldsymbol{\phi})\boldsymbol{\epsilon},$$

where

$$\mathbf{M}(\boldsymbol{\phi}) = \begin{bmatrix} \mathbf{J}_l(\boldsymbol{\phi}^{(a)}) & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{63}$$

and $\mathbf{J}_l(\boldsymbol{\phi}^{(a)})$ denotes the left Jacobian of SO(3) at $\boldsymbol{\phi}^{(a)}$ of which the closed form equation is given in (13).

Finally we have

$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{\boldsymbol{\epsilon}'} \approx \mathbf{M}(\boldsymbol{\phi})\mathbf{P}_{\boldsymbol{\epsilon}}\mathbf{M}(\boldsymbol{\phi})^T, \tag{64}$$

where $\mathbf{P}_{\boldsymbol{\epsilon}}$ is given by Equation (59). This justifies Equation (28) in Section III-B3.

- Reynolds [33] and Mueller [34] propose slightly different algorithms from the above Equation (64). More specifically, Reynolds proposes a method for covariance correction of the quaternion state, while Mueller takes a first-order approximation of both $\boldsymbol{\phi}^{(a)}$ and the noise vector $\boldsymbol{\epsilon}^{(a)}$ in his derivation. In contrast, Equation (64) is derived solely from the first-order approximation of $\boldsymbol{\epsilon}$.

Remark 1. *If the left-invariant noise is adopted [12], the right Jacobian should be used in the covariance update equation.*

APPENDIX C
MOTION AND MAGNETIC DISTURBANCES

If a triaxial accelerometer moves with large acceleration, it produces the vector sum of the negative direction of gravity and additional accelerations expressed in the body frame $\{\mathcal{B}\}$ fixed to the IMU. Harada [35] calls these additional acceleration terms as a motion disturbance. In magnetically disturbed environments, the measurement of a triaxial magnetometer is deviated from the local geomagnetic field expressed in $\{\mathcal{B}\}$.

To detect these disturbances, many reliability functions have been proposed [8], [35]. However, for simple and effective detection of disturbances, checking only the norms of calibrated outputs coming from accelerometers or magnetometers is enough [36]. Let $\tilde{\mathbf{v}}_i \in \mathbb{R}^3, (i = 1 \text{ or } 2)$ be the unnormalized calibrated output vector of three-axis accelerometer or magnetometer at a particular moment. If $|\|\tilde{\mathbf{v}}_i\| - 1| > \gamma_i$ where $\gamma_i > 0$ is a certain threshold, we call that a disturbance is detected. Otherwise, it is called that there is no disturbance.

When dealing with motion or magnetic disturbances in stochastic attitude filters, two methods are commonly used:

- Adaptation of noise covariances [37]: If disturbance is detected, then the noise covariance of Kalman filter is adjusted.
- Measurement reconstruction with a vector selector [38]: If disturbance is detected, then $\tilde{\mathbf{v}}_i$ is replaced by $\hat{\mathbf{R}}_{k+1|k}^T \mathbf{r}_i$. Here, $\hat{\mathbf{R}}_{k+1|k}$ is given by Equation (19).

In our estimator, the measurement reconstruction method with a vector selector is used to support fast response to disturbances.

APPENDIX D
EXTRINSIC MEAN OF UNIT VECTORS

Proposition 1. *Given a set of d-dimensional unit vectors, $\mathcal{S}_v = \{\mathbf{v}_i \in \mathbb{R}^d | \|\mathbf{v}_i\| = 1, i = 1, \ldots, N\}$, the extrinsic mean of $\mathcal{S}_v$ is defined as $\mathbf{v}^* := \arg\min_{\mathbf{v}} \sum_{i=1}^N \|\mathbf{v}_i - \mathbf{v}\|^2$ subject to $\|\mathbf{v}\| = 1$. If $\mathbf{m} := \sum_{i=1}^N \mathbf{v}_i \neq \mathbf{0}$, then $\mathbf{v}^* = \mathbf{m}/\|\mathbf{m}\|$.*

Proof. We construct the *Lagrangian* as $L(\mathbf{v}, \lambda) = \sum_{i=1}^N \|\mathbf{v}_i - \mathbf{v}\|^2 + \lambda(\mathbf{v}^T\mathbf{v} - 1)$ where $\lambda > 0$. The first order necessary conditions ($\frac{\partial L(\mathbf{v}^*, \lambda)}{\partial \mathbf{v}^*} = 0$ and $\frac{\partial L(\mathbf{v}^*, \lambda)}{\partial \lambda} = 0$) yield the result. □

APPENDIX E
PROOF OF EQUATION (30)

Given the inverse of a true attitude, denoted $\bar{\mathbf{A}} \in \text{SO(3)}$, let us consider a slightly modified form of the optimization problem in Equation (4) as follows:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta} \in \mathbb{R}^3} \sum_{i=1}^2 \frac{1}{\sigma_i^2} \|\mathbf{v}_i - \exp([\boldsymbol{\theta}])\, \bar{\mathbf{A}}\mathbf{r}_i\|^2, \tag{65}$$

where $\mathbf{v}_i = \bar{\mathbf{A}}\mathbf{r}_i + \Delta\mathbf{v}_i$, and $\Delta\mathbf{v}_i$ denotes the zero-mean measurement noise. The covariance of the random variable $\Delta\mathbf{v}_i$ is given by Equation (35), and $\exp([\boldsymbol{\theta}])\bar{\mathbf{A}}$ corresponds to the inverse of the optimization variable $\mathbf{R}$ in (4). Assuming that $\Delta\mathbf{v}_i$ is small, the solution $\boldsymbol{\theta}^*$ will be located near the origin. Under the first order approximation $\exp([\boldsymbol{\theta}]) \approx \mathbf{I} + [\boldsymbol{\theta}]$, the objective function can be approximated as

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta} \in \mathbb{R}^3} \sum_{i=1}^2 \frac{1}{\sigma_i^2} \|\Delta\mathbf{v}_i + [\bar{\mathbf{A}}\mathbf{r}_i]\boldsymbol{\theta}\|^2. \tag{66}$$

Equation (66) corresponds to a linear least-square estimation problem, and the estimate is represented as a linear function of $\Delta\mathbf{v}_i$:

$$\boldsymbol{\theta}^* = \sum_{i=1}^2 \mathbf{J}_i \Delta\mathbf{v}_i,$$

where

$$\mathbf{J}_i = \mathbf{M}^{-1}\left(\frac{1}{\sigma_i^2}[\mathbf{A}\mathbf{r}_i]\right), \tag{67}$$

and

$$\mathbf{M} := \sum_{i=1}^2 \frac{1}{\sigma_i^2}(\mathbf{I} - \bar{\mathbf{A}}\mathbf{r}_i\mathbf{r}_i^T\bar{\mathbf{A}}^T). \tag{68}$$

Here $\mathbf{M}$ denotes the Fisher information matrix [27]. Since Equation (66) has the form of a linear least-square estimation problem, the covariance of $\boldsymbol{\theta}^*$ achieves the CRLB [39]. Thus, the covariance of $\boldsymbol{\theta}^*$ is given by

$$E(\boldsymbol{\theta}^*\boldsymbol{\theta}^{*T}) = \sum_{i=1}^2 \mathbf{J}_i E(\Delta\mathbf{v}_i \Delta\mathbf{v}_i^T)\mathbf{J}_i^T \tag{69}$$

$$= \mathbf{M}^{-1}, \tag{70}$$

where $E(\boldsymbol{\theta}^*) = 0$ is used. Since $\mathbf{R} = \bar{\mathbf{A}}^{-1} \exp(-[\boldsymbol{\theta}])$ holds, the left invariant covariance of $\mathbf{R}$ in Equation (4) is the same as the covariance of $\boldsymbol{\theta}$. This completes the proof.

REFERENCES

[1] A. Bry, C. Richter, A. Bachrach, and N. Roy, "Aggressive flight of fixed-wing and quadrotor aircraft in dense indoor environments," *Int. J. Robot. Res.*, vol. 34, no. 7, pp. 969–1002, 2015.

[2] G. Loianno, C. Brunner, G. McGrath, and V. Kumar, "Estimation, control, and planning for aggressive flight with a small quadrotor with a single camera and IMU," *IEEE Trans. Robot. Autom. Lett.*, vol. 2, no. 2, pp. 404–411, 2017.

[3] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual-inertial odometry," *IEEE Trans. Robot.*, vol. 33, no. 1, pp. 1–21, 2017.

[4] S. Grzonka, A. Karwath, F. Dijoux, and W. Burgard, "Activity-based estimation of human trajectories," *IEEE Trans. Robot.*, vol. 28, no. 1, pp. 234–245, 2012.

[5] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *J. Guid. Control Dyn.*, vol. 30, no. 1, pp. 12–28, 2007.

[6] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Trans. Automat. Contr.*, vol. 53, no. 5, pp. 1203–1218, 2008.

[7] R. Mahony, V. Kumar, and P. Corke, "Multirotor aerial vehicles: Modeling, estimation, and control of quadrotor," *IEEE Robot. Automat. Mag.*, vol. 19, no. 3, pp. 20–32, 2012.

[8] R. Costanzi, F. Fanelli, N. Monni, A. Ridolfi, and B. Allotta, "An attitude estimation algorithm for mobile robots under unknown magnetic disturbances," *IEEE/ASME Trans. Mechatron.*, vol. 21, no. 4, pp. 1900–1911, 2016.

[9] J. L. Crassidis and F. L. Markley, "Unscented filtering for spacecraft attitude estimation," *J. Guid. Control Dyn.*, vol. 26, no. 4, pp. 536–542, 2003.

[10] C. Kim, R. Sakthivel, and W. K. Chung, "Unscented FastSLAM: A robust and efficient solution to the SLAM problem," *IEEE Trans. Robot.*, vol. 24, no. 4, pp. 808–820, 2008.

[11] A. Giannitrapani, N. Ceccarelli, F. Scortecci, and A. Garulli, "Comparison of EKF and UKF for spacecraft localization via angle measurements," *IEEE Tran. Aerosp. Electron. Syst.*, vol. 47, no. 1, pp. 75–84, 2011.

[12] G. Bourmaud, R. Mégret, M. Arnaudon, and A. Giremus, "Continuous-discrete extended Kalman filter on matrix Lie groups using concentrated Gaussian distributions," *J. Math. Imaging and Vision*, vol. 51, no. 1, pp. 209–228, 2015.

[13] G. Bourmaud, R. Mégret, A. Giremus, and Y. Berthoumieu, "From intrinsic optimization to iterated extended Kalman filtering on Lie groups," *J. Math. Imaging and Vision*, vol. 55, no. 3, pp. 284–303, 2016.

[14] A. Barrau and S. Bonnabel, "Intrinsic filtering on Lie groups with applications to attitude estimation," *IEEE Trans. Automat. Contr.*, vol. 60, no. 2, pp. 436–449, 2015.

[15] A. Barrau and S. Bonnabel, "The invariant extended Kalman filter as a stable observer," *IEEE Trans. Automat. Contr.*, vol. 62, no. 4, pp. 1797–1812, 2017.

[16] S. Hauberg, F. Lauze, and K. S. Pedersen, "Unscented Kalman filtering on Riemannian manifolds," *J. Math. Imaging and Vision*, vol. 46, no. 1, pp. 103–120, 2013.

[17] C. Hertzberg, R. Wagner, U. Frese, and L. Schröder, "Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds," *Information Fusion*, vol. 14, no. 1, pp. 57–77, 2013.

[18] J. Kwon, H. S. Lee, F. C. Park, and K. M. Lee, "A geometric particle filter for template-based visual tracking," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 36, no. 4, pp. 625–643, 2014.

[19] L. Chang, B. Hu, and G. Chang, "Modified unscented quaternion estimator based on quaternion averaging," *J. Guid. Control Dyn.*, vol. 37, no. 1, pp. 305–308, 2014.

[20] Y. S. Suh, "Orientation estimation using a quaternion–based indirect Kalman filter with adaptive estimation of external acceleration," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 12, pp. 3296–3305, 2010.

[21] F. C. Park, J. Bobrow, and S. R. Ploen, "A Lie group formulation of robot dynamics," *Int. J. Robot. Res.*, vol. 14, no. 6, pp. 609–618, 1995.

[22] J. Kim, S.-H. Lee, and F. C. Park, "Kinematic and dynamic modeling of spherical joints using exponential coordinates," *Proc. IMechE, Part C: J. Mech. Eng. Sci.*, vol. 228, no. 10, pp. 1777–1785, 2013.

[23] T. D. Barfoot and P. T. Furgale, "Associating uncertainty with three dimensional poses for use in estimation problems," *IEEE Trans. Robot.*, vol. 30, no. 3, pp. 679–693, 2014.

[24] T. D. Barfoot, *State Estimation for Robotics*. Cambridge: Cambridge University Press, 2017.

[25] G. Wahba, "A least–squares estimate of satellite attitude," *SIAM Review*, vol. 7, no. 3, p. 409, 1965.

[26] F. L. Markley, "Attitude determination using vector observations and the singular value decomposition," *J. Astronaut. Sci.*, vol. 36, no. 3, pp. 245–258, 1988.

[27] M. D. Shuster, "Maximum likelihood estimation of spacecraft attitude," *J. Astronaut. Sci.*, vol. 37, no. 1, pp. 79–88, 1989.

[28] M. D. Shuster, "The generalized Wahba problem," *J. Astronaut. Sci.*, vol. 54, no. 2, pp. 245–259, 2006.

[29] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction, 2nd ed.* Switzerland: Springer, 2016.

[30] R. Hartley, J. Trumpf, Y. Dai, and H. Li, "Rotation averaging," *Int. J. Comput. Vis.*, vol. 103, no. 3, pp. 267–305, 2013.

[31] M. D. Shuster and S. D. Oh, "Three-axis attitude determination from vector observations," *J. Guid. Control Dyn.*, vol. 4, no. 1, pp. 70–77, 1981.

[32] M. Hwangbo, J.-S. Kim, and T. Kanade, "Gyro-aided feature tracking for a moving camera: fusion, auto-calibration and GPU implementation," *Int. J. Robot. Res.*, vol. 30, no. 14, pp. 1755–1774, 2011.

[33] R. G. Reynolds, "Asymptotically optimal attitude filtering with guaranteed convergence," *J. Guid. Control Dyn.*, vol. 31, no. 1, pp. 114–122, 2008.

[34] M. W. Mueller, M. Hehn, and R. D'Andrea, "Covariance correction step for Kalman filtering with an attitude," *J. Guid. Control Dyn.*, vol. 40, no. 9, pp. 2301–2306, 2017.

[35] T. Harada, T. Mori, and T. Sato, "Development of a tiny orientation estimation device to operate under motion and magnetic disturbance," *Int. J. Robot. Res.*, vol. 26, no. 6, pp. 547–559, 2007.

[36] H. Rehbinder and X. Hu, "Drift-free attitude estimation for accelerated rigid bodies," *Automatica*, vol. 40, pp. 653–659, 2004.

[37] A. M. Sabatini, "Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 7, pp. 1346–1356, 2006.

[38] J. K. Lee and E. J. Park, "Minimum-order Kalman filter with vector selector for accurate estimation of human body orientation," *IEEE Trans. Robot.*, vol. 25, no. 5, pp. 1196–1201, 2009.

[39] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory, vol.1*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

Donghoon Kang (M'12) received his B.S. and M.S. degrees in mechanical engineering from Pohang University of Science and Technology (POSTECH), Pohang, Korea, in 1997 and 1999, respectively. He received the Ph.D. degree in mechanical and aerospace engineering from Seoul National University, Seoul, Korea in 2018.

Since 2000, he has been with Korea Institute of Science and Technology (KIST), where he is currently a senor researcher. His research interests are in computer vision and signal processing.

Cheongjae Jang received his B.S. degree in mechanical and aerospace engineering from Seoul National University in 2012. He is currently working toward the Ph.D. degree at Seoul National University, Seoul, Korea.

During his Ph.D. studies, he has been a research assistant with the Robotics Laboratory, school of mechanical and aerospace engineering at Seoul National University, Seoul, Korea. His research interests are in manifold learning and optimal control.

Frank C. Park (F'13) received his BS in electrical engineering from MIT in 1985, and Ph.D. in applied mathematics from Harvard University in 1991.

After joining the faculty of UC Irvine from 1991 to 1994, since 1995 he has been professor of mechanical and aerospace engineering at Seoul National University. His research interests are in robot mechanics, planning and control, vision and image processing, and related areas of applied mathematics.

He has been an IEEE Robotics and Automation Society Distinguished Lecturer, and has held adjunct faculty positions at the NYU Courant Institute, the Interactive Computing Department at Georgia Tech, and the HKUST Robotics Institute. He is a Fellow of the IEEE, current Editor-in-Chief of the IEEE Transactions on Robotics, developer of the EDX course Robot Mechanics and Control I, II, and co-author (with Kevin Lynch) of the textbook "Modern Robotics: Mechanics, Planning, and Control".