GENERAL INSTRUCTION

- **Authors: We cannot accept new source files as corrections for your paper. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.**

Queries

Q1. Author: Please confirm or add details for any funding or financial support for the research of this article.
Q2. Author: Acronym UKF has been used for both "unscented Kalman filter" and "unscented Kalman filtering." Please check.
Q3. Author: Please provide the complete page range for Ref. [25].

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 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 IMU measurements of gravity and the earth's magnetic field. 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 update. The resulting attitude estimates are invariant with respect to choice of fixed and moving reference frames. The performance advantages of our filter vis-à-vis existing state-of-the-art IMU attitude estimation algorithms are validated via numerical and hardware experiments involving both synthetic and real data.**

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

I. Introduction

ESTIMATING an object's orientation, or attitude, from an inertial measurement unit (IMU) attached to the object arises in applications ranging 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 velocities (which can be integrated to calculate the attitude), the accelerometer measures accelerations due to gravity and other external forces, and the magnetometer measures the earth's magnetic field. Gyroscopic measurements contain a time-varying bias error, and accelerometer and magnetometer measurements can be used to identify and compensate for this gyroscope bias. More generally, the challenges and benefits of simultaneously estimating the attitude and gyroscope bias from disparate sensor measurements are detailed in [5] and the cited references.

Notable among deterministic filtering methods for simultaneously estimating attitude and gyroscope bias are Mahony *et al.*'s series of nonlinear complementary filters (NCFs) [6]–[8]; these filters ensure almost global stability of the observer error, and their performance has been validated in numerous experimental scenarios. Stochastic filtering methods further take into account statistical characterizations of measurement and process 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 com- Q2
plexity, has been shown to outperform the EKF in a wide range of applications [9]–[11].

Because the underlying configuration space of rotations, represented by the group SO(3) of 3×3 real orthogonal matrices with unit determinant, is not a vector space but a curved space, the attitude estimation problem is fundamentally a nonlinear one. The straightforward but naive approach of expressing a rotation in terms of some suitable local coordinates (e.g., roll-pitch-yaw angles, Euler angles) is problematic at several levels: the local coordinates contain singularities that require special treatment (for example, when the pitch angle is $90°$), and the resulting estimates depend both on the choice of local coordinates as well as fixed and moving reference frames. If standard vector space filters are naively adapted to local coordinate representations of the attitude, not only are the equations for the state space dynamics and measurements highly nonlinear and dependent on the choice of reference frames, but filtering performance is highly uneven throughout different regions of the configuration space.

Recent research has attempted to address the issue of coordinate and reference frame dependency through the use of differential geometric methods. Although computationally more involved than standard vector space filtering algorithms, when correctly formulated, these methods are invariant with respect to the choice of fixed and moving reference frames, and also independent of the choice of local coordinates used to parameterize the rotations. For estimation problems in which the underlying configuration space has the structure of a matrix Lie group like SO(3), 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

Manuscript received February 23, 2018; revised July 1, 2018 and
Q1 September 6, 2018; accepted December 31, 2018. Recommended by Technical Editor T. Singh. This work was supported in part by the NAVER LABS, in part by the ADD-ICMTC, in part by the SNU-IAMD, in part by the BK21+, in part by the MI Technology Innovation Program 10048320, in part by the National Research Foundation of Korea under Grant NRF-2016R1A5A1938472, in part by MOTIE Technology Innovation Program under Grant 2017-10069072, in part by SNU BMRR Center under Grant DAPA-UD160027ID, and in part by the KIST Institutional Project. *(Corresponding author: Donghoon Kang.)*

D. Kang is with the Imaging Media Center, Korea Institute of Science and Technology, Seoul 136-791, South 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 151-744, South Korea (e-mail: jchastro@gmail.com; fcp@snu.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMECH.2019.2891776

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 often prove to be difficult 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 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 update, and also by drawing upon Lie-theoretic properties in key steps of our geometric UKF algorithm.

This 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. Section V compares the performance of our geometric UKF algorithm against other existing state-of-the-art estimators for attitude and gyroscope bias [6], [19], [20], with detailed experiments involving both synthetic and real data validating the 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 represented by the 3×3 real matrices $\mathbf{R}$ satisfying $\mathbf{R}^T\mathbf{R} = \mathbf{I}$ and $\det \mathbf{R} = 1$, where $\mathbf{I}$ denotes the 3×3 identity matrix. SO(3) is an example of a matrix Lie group; 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 $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)^T \in \mathbb{R}^3$. A fundamental connection between so(3) and SO(3) is the matrix exponential map $\exp : \mathrm{so}(3) \rightarrow \mathrm{SO}(3)$, given as

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

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

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

where θ satisfies $1 + 2\cos\theta = \mathrm{tr}(\mathbf{R})$, $|\theta| < \pi$ [here, $\mathrm{tr}(\cdot)$ denotes the trace of a matrix]. If $\mathrm{tr}(\mathbf{R}) = -1$, 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 aforementioned 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 $\mathbf{P}, \mathbf{Q} \in \mathrm{SO}(3)$. With this notion of distance, 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 $\boldsymbol{\Omega} = \log(\mathbf{R}_1^T\mathbf{R}_2) \in \mathrm{so}(3)$.

Recalling that $\mathbb{R}^3$ is also trivially a Lie group under vector addition, the direct product $\mathrm{SO}(3) \times \mathbb{R}^3$ 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 $\mathbf{X}_0 \in \mathrm{SO}(3)$ is given and $\boldsymbol{\eta} \in \mathbb{R}^3$ is a zero-mean Gaussian with covariance $\mathbf{P}_\eta$, i.e., $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_\eta)$. 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 $[\boldsymbol{\zeta}] \in \mathrm{so}(3)$ 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] and [24].

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

III. UKF Algorithm for Estimating Attitude and Gyroscope Bias

Before describing our geometric UKF algorithm, we fix notation, describe the sensor models and their underlying assumptions, and review Wahba's Problem [25] and its solutions.

Let $\{\mathcal{I}\}$ be the inertial reference frame fixed to ground, and let $\{\mathcal{B}\}$ denote the body frame fixed to the moving IMU. Let $\boldsymbol{\omega}^m \in \mathbb{R}^3$ be the angular velocity measured by the IMU gyroscope with respect to frame $\{\mathcal{B}\}$. Denote by $\mathbf{a}, \mathbf{m} \in \mathbb{R}^3$ the IMU accelerometer and magnetometer measurements, respectively; like $\boldsymbol{\omega}^m$, both $\mathbf{a}$ and $\mathbf{m}$ are assumed measured with respect to the

IMU frame $\{\mathcal{B}\}$. Further define the unit vectors $\mathbf{v}_1 := \mathbf{a}/\|\mathbf{a}\|$, $\mathbf{v}_2 := \mathbf{m}/\|\mathbf{m}\|$.

In what follows, we assume that the IMU is suitably calibrated, and that the gravitational acceleration is dominant in the accelerometer measurement $\mathbf{a}$. Let $\mathbf{r}_1 \in \mathbb{R}^3$ be the unit vector in the opposite direction of gravity, and $\mathbf{r}_2 \in \mathbb{R}^3$ be the unit vector in the direction of the earth's magnetic field. If $\mathbf{r}_1$ and $\mathbf{r}_2$ are not collinear, then $\mathbf{r}_i$ and $\mathbf{v}_i$ should satisfy $\mathbf{r}_i = \mathbf{R}\mathbf{v}_i$, $i = 1, 2$, for some rotation $\mathbf{R} \in \mathrm{SO}(3)$ representing the orientation of the IMU frame $\{\mathcal{B}\}$ relative to the fixed frame $\{\mathcal{I}\}$. Since in practice IMU measurements are noisy, $\mathbf{R}$ is typically estimated as the solution to the following optimization problem (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 the w_i are positive weights. A popular choice for w_i is $w_i = 1/\sigma_i^2$, where σ_i^2 denotes the variance of $\mathbf{v}_i$ in 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 widely used [28].)

Wahba's Problem as defined by (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 (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 earlier do not take into account the state space dynamics of the object or process and measurement noise characteristics, and typically are inferior to estimates obtained via nonlinear stochastic filtering techniques. We now formulate the overall problem in a discrete-time stochastic filtering setting. 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 zero-mean Gaussian noise. The state dynamics are then assumed to be of the form

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

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

where h is the integration time step, 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 a first-order linear approximation of the state dynamics (7) that leads to a closed-form expression for the covariance of $\mathbf{R}_{k+1}$ consistent with (1). From the Baker–Campbell–Hausdorff formula [29], given $[x], [y] \in \mathrm{so}(3)$, $\exp([\mathbf{x}])\exp([\mathbf{y}])$ can be written exactly in the form $\exp([\mathbf{x}])\exp([\mathbf{y}]) = \exp([\mathbf{z}])$, $[z] \in \mathrm{so}(3)$, 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 (9) in the form

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

Gathering only terms linear in $\mathbf{x}$ in (10), the following approximation between $\mathbf{x}$ and $\mathbf{x}'$ holds for $\|\mathbf{x}\|$ sufficiently small [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 (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 (12) is used. The relation $\mathbf{R}\exp([\boldsymbol{\omega}])\mathbf{R}^T = \exp([\mathbf{R}\boldsymbol{\omega}])$ for $\mathbf{R} \in \mathrm{SO}(3)$, $[\boldsymbol{\omega}] \in \mathrm{so}(3)$ is used in the derivation of (15).

Note that $\mathbf{l}_k = -\mathbf{R}_k\mathbf{J}_l(\boldsymbol{\psi})\boldsymbol{\eta}_k h$ is itself a random variable, since it is 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 equations are assumed to be of the form

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

where $\mathbf{Y}_{k+1} \in \mathrm{SO}(3)$ is calculated as a solution to Wahba's Problem (4) using IMU gravitational acceleration and magnetic field measurements. The measurement noise $\mathbf{w}_{k+1} \in \mathbb{R}^3$ is assumed to be zero-mean Gaussian, implying that the measurement vector statistics are rotationally symmetric about their true measurement vectors.

B. UKF Algorithm

We now present the geometric UKF algorithm for simultaneous attitude and gyroscope bias estimation. Let $\mathbf{R}_k$ and $\mathbf{b}_k$, respectively, denote the attitude and the gyroscope bias at time step k, and $\mathbf{X}_k := (\mathbf{R}_k, \mathbf{b}_k) \in \mathrm{SO}(3) \times \mathbb{R}^3$.

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

Algorithm 1: Weighted Intrinsic Mean on SO(3).

Input: 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 $\quad \mathbf{\Lambda} \leftarrow \sum_{i=0}^{12} w_m^{(i)} \log(\mathbf{Z}_i \mathbf{T}^{-1})$
4 $\quad \mathbf{T} \leftarrow \exp(\mathbf{\Lambda})\,\mathbf{T}$
5 **return** $\mathbf{T}$

given. From (5), $\hat{\mathbf{R}}_{0|0}$ is estimated by solving Wahba's Problem (4) from a pair of initial measurement vectors $(\mathbf{v}_1, \mathbf{v}_2)$.

2) Time Update:

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

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

where following the work presented in [10], the parameter γ is chosen as $\gamma = \sqrt{N_x + \lambda}$, with N_x set to the state dimension (six) and $\lambda = N_x(\alpha^2 - 1)$, $0 < \alpha < 1$; $\mathbf{s}_i \in \mathbb{R}^6$ is the ith column vector of the lower-triangular matrix $\mathbf{S} \in \mathbb{R}^{6\times 6}$ in the Cholesky decomposition $\mathbf{P}_{k|k} = \mathbf{S}\mathbf{S}^T$, and $\mathbf{s}_i^{(a)}, \mathbf{s}_i^{(b)} \in \mathbb{R}^3$ are, respectively, the upper and lower halves of $\mathbf{s}_i$.

2) Setting $\mathbf{l}_k = \mathbf{0}$ in (15) and $\mathbf{n}_k = \mathbf{0}$ in (8), 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

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

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

3) Given the set of rotations $\{\Upsilon_{\mathbf{R},k+1}^{(0)}, \ldots, \Upsilon_{\mathbf{R},k+1}^{(12)}\}$ in SO(3), evaluate the weighted mean rotation $\bar{\Upsilon}_{\mathbf{R},k+1} \in \text{SO}(3)$ using Algorithm 1. Taking advantage of the rapid convergence of Algorithm 1 [18], [30], set the number of iterations in line 2 of the algorithm to $n = 3$ or 4. The weights $w_m^{(i)} \in \mathbb{R}$ in line 3 satisfy $\sum_{i=0}^{12} w_m^{(i)} = 1$.

4) The gyroscope bias estimate $\bar{\Upsilon}_{\mathbf{b},k+1} \in \mathbb{R}^3$ is given by the weighted mean of $\{\Upsilon_{\mathbf{b},k+1}^{(0)}, \ldots, \Upsilon_{\mathbf{b},k+1}^{(12)}\}$ in $\mathbb{R}^3$, i.e., $\bar{\Upsilon}_{\mathbf{b},k+1} = \sum_{i=0}^{12} w_m^{(i)} \Upsilon_{\mathbf{b}}$. $\hat{\mathbf{X}}_{k+1|k} := (\hat{\mathbf{R}}_{k+1|k}, \hat{\mathbf{b}}_{k+1|k})$ is then

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

5) 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 $\mathbf{q}_i = (\mathbf{q}_i^{(a)}, \mathbf{q}_i^{(b)}) \in \mathbb{R}^6$. 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 $\mathbf{N}_k = \left[\begin{smallmatrix} (ch^2)\mathbf{I} & \mathbf{0} \\ \mathbf{0} & d\mathbf{I} \end{smallmatrix}\right]$ is the process noise covariance.

6) Let $\mathbf{u}_i \in \mathbb{R}^6$ denote the ith column vector of the lower-triangular matrix $\mathbf{U} \in \mathbb{R}^{6\times 6}$ in the Cholesky decomposition $\mathbf{P}_{k+1|k} = \mathbf{U}\mathbf{U}^T$. The upper and lower halves of $\mathbf{u}_i$ are, respectively, denoted $\mathbf{u}_i^{(a)} \in \mathbb{R}^3$ and $\mathbf{u}_i^{(b)} \in \mathbb{R}^3$. Redraw the sigma points $\mathcal{X}_{k+1}^{(i)} := (\mathcal{X}_{\mathbf{R},k+1}^{(i)}, \mathcal{X}_{\mathbf{b},k+1}^{(i)})$, $(i = 0, \ldots, 12)$ from $\hat{\mathbf{X}}_{k+1|k}$ and $\mathbf{P}_{k+1|k}$ as follows:

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

3) Measurement Update:

1) If the IMU moves with high acceleration or is subject to magnetic disturbances, the accelerometer and magnetometer measurements may be corrupted and not satisfy our earlier assumptions. Appendix C summarizes some existing methods for addressing these disturbances.

2) Setting $\mathbf{w}_{k+1} = \mathbf{0}$ in (16), define the set of measurement sigma points $\mathcal{S}_{\mathcal{Y}} = \{\mathcal{Y}_{k+1}^{(i)} \in \text{SO}(3) \mid i = 0, \ldots, 12\}$ as follows:

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

3) The mean $\hat{\mathbf{Y}}_{k+1}$ of $\{\mathcal{Y}_{k+1}^{(0)}, \ldots, \mathcal{Y}_{k+1}^{(12)}\}$ is given by

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

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

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

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 Section IV, we derive a closed-form expression for $\mathbf{W}_{k+1}$ from (32).

4) Define $[\mathbf{p}_i^{(a)}] := \log(\mathcal{X}_{\mathbf{R},k+1}^{(i)} \hat{\mathbf{R}}_{k+1|k}^{-1}) \in \text{so}(3)$ and $\mathbf{p}_i^{(b)} := \mathcal{X}_{\mathbf{b},k+1}^{(i)} - \hat{\mathbf{b}}_{k+1|k} \in \mathbb{R}^3$, and $\mathbf{p}_i = (\mathbf{p}_i^{(a)}, \mathbf{p}_i^{(b)}) \in$

$\mathbb{R}^6$. The associated covariance $\mathbf{P}_{\mathbf{xy}}$ is then calculated as

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

5) 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

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

where $\mathbf{Y}_{k+1}$ and $\hat{\mathbf{Y}}_{k+1}$ are, respectively, given by (16) and (22). Define $\boldsymbol{\phi}^{(a)} \in \mathbb{R}^3$ and $\boldsymbol{\phi}^{(b)} \in \mathbb{R}^3$ to be the upper and lower halves of $\boldsymbol{\phi} := \mathbf{K}\boldsymbol{\delta} \in \mathbb{R}^6$. The state and covariance are now updated according to

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

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

where $\mathbf{M}(\boldsymbol{\phi}^{(a)}) \in \mathbb{R}^{6\times 6}$ is given by

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

The justification for $\mathbf{M}(\boldsymbol{\phi}^{(a)})$ in (28) is given in Appendix B.

IV. Measurement Noise Covariance

This section presents an algorithm for obtaining, from a set of noisy unit vector measurements of the gravity and magnetic field vectors, a full-rank measurement noise covariance matrix.

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 the solution to Wahba's Problem (4):

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

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

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

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

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

Equation (32) follows from a straightforward calculation combining (3) and (30).

Note that the left-invariant covariance of $\mathbf{R}$ in (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 (32) is the covariance of the solution to Wahba's Problem represented with respect to the fixed ground frame. If values for $\sigma_i^2, \mathbf{r}_i$ are given, the right-invariant covariance of $\mathbf{R}$ in (32) can be determined to be a constant matrix, independent of $\bar{\mathbf{A}}$. However, the left-invariant covariance of $\mathbf{R}$ in (30) requires $\bar{\mathbf{A}}$, σ_i^2, and $\mathbf{r}_i$.

When the IMU is moving, $\bar{\mathbf{A}}$ is also changing, and the left-invariant covariance of $\mathbf{R}$ needs to be updated at every time step. The left-invariant covariance can be evaluated as the inverse of a matrix that varies with $\bar{\mathbf{A}}$, while the right-invariant covariance remains invariant. When the IMU motion involves both translation and rotation, measurements of the two direction vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are subject to greater errors, leading to less accurate estimates of $\bar{\mathbf{A}}$. For the reasons outlined earlier, our measurement noise covariance formula of (32) is preferable to Shuster's formula (30) in the geometric UKF algorithm.

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

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

1) Constant Vectors $(\mathbf{r}_1, \mathbf{r}_2)$: Assign each axis of the inertial reference frame $\{\mathcal{I}\}$ as follows: The direction opposite to gravity is set to be the y-axis of $\{\mathcal{I}\}$, while the x-axis of $\{\mathcal{I}\}$ is orthogonal to both gravity and the earth's magnetic field. With these assignments, $\mathbf{r}_1 = (0, 1, 0)^T$ and

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

where ϕ is unknown and to be determined.

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

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

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

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

Let the SVD of $\mathbf{M}_t$ be $\mathbf{M}_t = \mathbf{U}_t\mathbf{\Sigma}_t\mathbf{V}_t^T$, where in principle $\mathbf{\Sigma}_t = \mathrm{diag}(\sigma_i^2, \sigma_i^2, 0)$ and $\check{\mathbf{v}}_i$ is the corresponding direction for the singular value 0. Since in practice ground-truth values of $\check{\mathbf{v}}_i$ are unavailable, an alternative method of determining σ_i^2 is needed. Assuming that the IMU is stationary and N measurements are available, the covariance of $\mathbf{v}_i$ can be estimated by

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

where $\mathbf{v}_i^{(j)}$ denotes the jth measurement vector obtained from the ith sensor (sensor 1 is the accelerometer, while sensor 2 is the magnetometer). Let the SVD of $\mathbf{M}_a$ be $\mathbf{M}_a = \mathbf{U}_a\mathbf{\Sigma}_a\mathbf{V}_a^T$, where $\boldsymbol{\Sigma}_a = \mathrm{diag}(s_1, s_2, s_3)$ and $s_1 \geq s_2 \geq s_3$, $s_3 \approx 0$. $\boldsymbol{\Sigma}_a$ will typically be close to its theoretical value $\boldsymbol{\Sigma}_t$, in which case we can set

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

V. Experimental Results

In this section, we compare the performance of our geometric UKF algorithm [UKF on SO(3)] against other state-of-the-art methods (UKF on Quaternion [19], EKF on Quaternion [20], and the passive NCF (NCF on SO(3)) [6]). Using both synthetic and real data in our experiments, both the convergence rate and accuracy of the attitude and gyroscope bias estimates are compared.

Ground-truth values 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 filter update time step is set to $h_0 = 1/60$ seconds. 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 set to

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

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

α in (40) is set to 0.9, and β is set to two for a Gaussian prior [10].

A. Synthetic Data

In our numerical simulation experiments, the vectors in (4) are set to $\mathbf{r}_1 = (0, 1, 0)^T$ and $\mathbf{r}_2 = (0, \cos(\phi_s), \sin(\phi_s))^T$, where $\phi_s = 2.4$ rad. The ground-truth value $\breve{\mathbf{R}}_1 \in \mathrm{SO}(3)$ is set randomly to be the initial attitude.

For realistic simulation, we first collect a set of real angular rate measurements $\breve{\boldsymbol{\omega}}_k$ from an actual gyroscope (L3G4200D) at the 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 value of the initial gyroscope bias is set to be $\breve{\mathbf{b}}_0 = (-0.06, 0.3, 0.3)^T$ rad/s. We then 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, $\sigma_0 = (1.1 \times 10^{-3}/h_0)$ rad/s, $\sigma_1 = (1.0 \times 10^{-5})$ rad/s, $\sigma_2 = 1.00 \times 10^{-2}$, and $\sigma_3 = 1.58 \times 10^{-2}$.

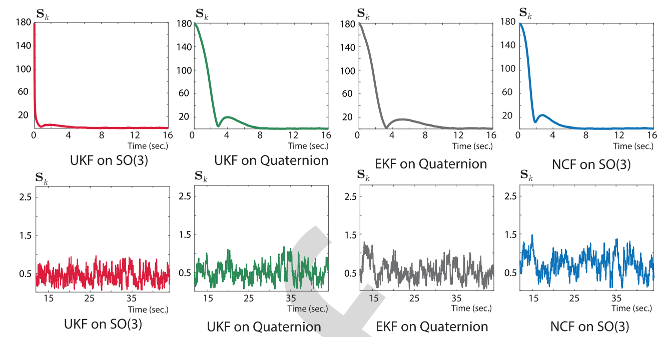


Fig. 1. Simulation experiments: Attitude estimation errors (in degrees) over the time intervals $t \in [0, 16]$ s (top) and $t \in [12, 44]$ s (bottom).

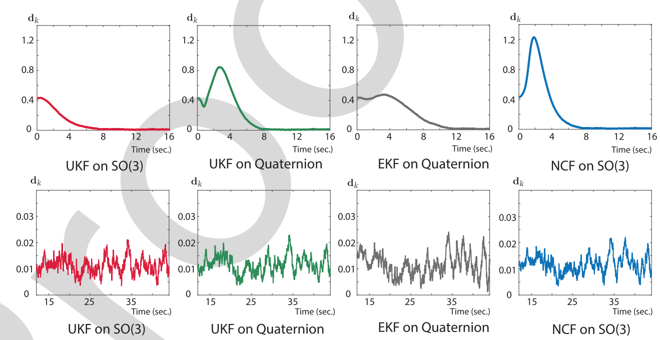


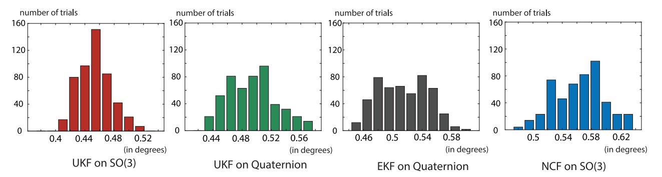
Fig. 2. Simulation experiments: Gyroscope bias estimate errors (in radian/seconds) over the time intervals $t \in [0, 16]$ s (top) and $t \in [12, 44]$ s (bottom).

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 (20) and $\mathbf{W}_{k+1}$ in (24) of the proposed attitude estimator [UKF on SO(3)] are set as follows: $\mathbf{N}_k = \left[\begin{smallmatrix} (\sigma_0 h_0)^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_1^2\mathbf{I} \end{smallmatrix}\right]$ and $\mathbf{W}_{k+1} = (\frac{1}{\sigma_2^2}(\mathbf{I} - \mathbf{r}_1\mathbf{r}_1^T) + \frac{1}{\sigma_3^2}(\mathbf{I} - \mathbf{r}_2\mathbf{r}_2^T))^{-1}$.

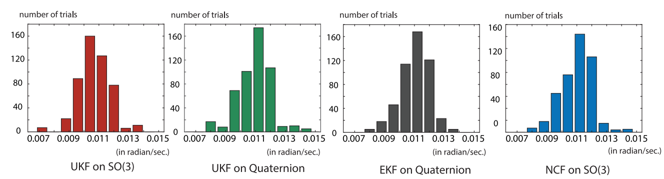
From the simulation results shown in Figs. 1 and 2, it can be seen that the proposed algorithm [UKF on SO(3)] converges most rapidly over the time interval $t \in [0, 14]$ s. To more reliably assess the accuracy of each estimator, we generate 500 sets of synthetic data using (42)–(45). Fig. 3 shows the histograms of estimation errors of the attitudes and the slowly time-varying gyroscope biases. Tables I and II summarize the experimental results corresponding to Fig. 3(a) and (b). From Fig. 3(b) and Table II, it can be seen that the gyroscope bias estimates show similar performance for all estimators. In terms of attitude estimates, "UKF on SO(3)" is the most accurate among the estimators [see Fig. 3(a) and Table I].

B. Real Experiments

The IMU for real experiments consists of an L3G4200D gyroscope, LIS3LV02DQ accelerometer, HMC5883L magnetometer, and Cortex-M3 microcontroller. In real experiments, ground-truth values of the slowly time-varying gyroscope bias



(a)



(b)

Fig. 3. Simulation experiments: Histograms of estimation errors over the time interval $t \in [12, 44]$ s (averaged over 500 trials). (a) Histograms of attitude estimation errors. (b) Histograms of gyroscope bias estimation errors.

TABLE I
AVERAGE AND STANDARD DEVIATION OF ATTITUDE ESTIMATION ERRORS (IN DEGREES) OVER THE TIME INTERVAL $t \in [12, 44]$ S (AVERAGED 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

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

are unknown. We therefore assume that the gyroscope bias is initially unknown, but near-constant over short time durations. If the IMU is stationary, then the gyroscope bias, denoted $\breve{\mathbf{b}}$, can be temporarily captured by averaging a set of gyroscope data over a certain time interval [32].

Keeping the IMU stationary, the variance σ_i^2 of the unit vector $\mathbf{v}_{i,k}$, $i = 1, 2$, can be calculated from (37); in our experiments we obtain the values $\sigma_1^2 = 8.95 \times 10^{-5}$ and $\sigma_2^2 = 1.911 \times 10^{-4}$. Denoting by ϕ_r the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$, i.e., $\phi_r = \cos^{-1}(\mathbf{r}_1^T \mathbf{r}_2)$, we obtain $\phi_r = 2.486$ rad using Proposition 1 of Appendix D and (34). The noise covariances $\mathbf{N}_k$ in (20) and $\mathbf{W}_{k+1}$ in (24) of the proposed attitude estimator [UKF on SO(3)] are set as follows: $\mathbf{N}_k = \left[\begin{smallmatrix}(2.0\times10^{-9})\mathbf{I} & \mathbf{0}\\ \mathbf{0} & (3.0\times10^{-11})\mathbf{I}\end{smallmatrix}\right]$ 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 value of the attitude $\breve{\mathbf{R}}_k$ at time step k, we use the optical motion capture system OptiTrack 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 $\{(\boldsymbol{\omega}_k^m, \mathbf{v}_{1,k}, \mathbf{v}_{2,k}) \mid k = 1, \ldots, N_r\}$ obtained from the moving IMU, and the ground-truth attitude $\breve{\mathbf{R}}_k$ obtained from the OptiTrack infrared camera system, are synchronously saved into files at a sampling rate $1/h_0 = 60$ Hz. Here, the number of measurements N_r is set to 3000. For fair comparison among filters, we perform experiments with real data under the condition of negligible disturbances.

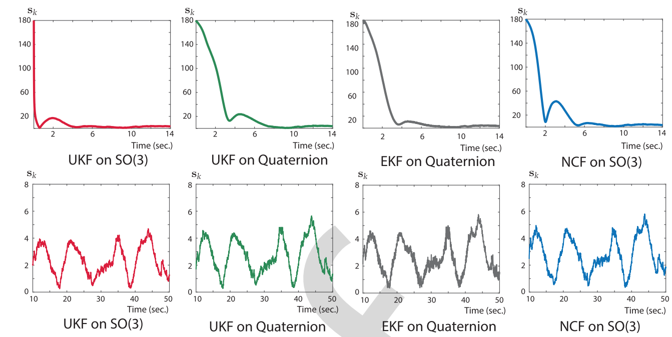


Fig. 4. Real experiments: Attitude estimate errors (in degrees) over the time interval $t \in [0, 14]$ s (top) and $t \in [10, 50]$ s (bottom).

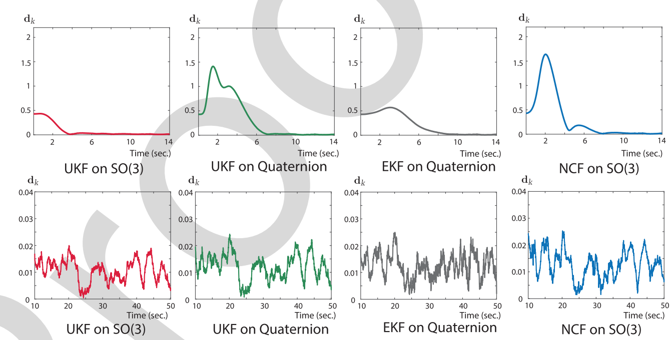


Fig. 5. Real experiments: Gyroscope bias estimate errors (in radian/seconds) over the time intervals $t \in [0, 14]$ s (top) and $t \in [10, 50]$ s (bottom).

TABLE III
RESULTS OF REAL EXPERIMENTS: AVERAGE ERRORS OVER THE TIME INTERVAL $t \in [10, 50]$ S (AVERAGED OVER TEN 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

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: $\hat{\mathbf{b}}_{1|1} = \breve{\mathbf{b}} + (1/h_0)(-0.001, 0.005, 0.005)^T = \breve{\mathbf{b}} + (-0.06, 0.3, 0.3)^T$ (rad/s) 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, Figs. 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 further experimentally verify these results, we collect nine additional sets of real data. As shown in Table III, "UKF on SO(3)" demonstrates superior performance compared to existing methods in terms of the accuracy of attitude estimates.

TABLE IV
AVERAGE COMPUTATION TIMES FOR 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

We also measure, at every time step, the computation times for each filter—all implemented in C++ and executed on a desktop computer with Intel i5-4670 (3.4 GHz) CPU. The computation times for each estimator are averaged over N_r steps. From Table IV it can be seen that "NCF on SO(3)" is the fastest among the estimators. Computation times for "UKF on SO(3)" are similar to those for "Quaternion UKF."

VI. CONCLUSION

This paper has presented a geometric unscented Kalman filtering algorithm for simultaneously estimating attitude and gyroscope bias from an inertial measurement unit. Drawing upon the Lie group properties of the set of rotation matrices SO(3), we derive a discrete-time stochastic nonlinear filtering algorithm evolving on SO(3) $\times \mathbb{R}^3$. One of the key features of our algorithm is to express observations as elements of SO(3), by determining the rotation corresponding to the IMU's gravitational acceleration and magnetic field vector measurements as a solution to Wahba's Problem. By doing so, first-order linear approximations of the state dynamics and measurement equations lead to closed-form equations for covariance propagation and update. These in turn lead to computationally efficient implementations of our filter, with the resulting attitude estimates invariant with respect to the choice of fixed and moving reference frames. Extensive numerical simulation and hardware experiments have demonstrated the superior convergence behavior and estimation accuracy of our proposed algorithm compared to existing state-of-the-art IMU estimators for attitude and gyroscope bias.

APPENDIX A
FIRST-ORDER APPROXIMATION OF EXPONENTIAL MAP

Given $[\mathbf{x}], [\mathbf{y}] \in \mathrm{so}(3)$, let $[\mathbf{z}] \in \mathrm{so}(3)$ satisfy

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

From the Baker–Campbell–Hausdorff 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 $[\cdot,\cdot] : \mathrm{so}(3) \times \mathrm{so}(3) \to \mathrm{so}(3)$ is defined as $[[\mathbf{a}],[\mathbf{b}]] = [\mathbf{a}][\mathbf{b}] - [\mathbf{b}][\mathbf{a}]$ for $[\mathbf{a}],[\mathbf{b}] \in \mathrm{so}(3)$. $[\mathbf{c}] = [[\mathbf{a}],[\mathbf{b}]] \in \mathrm{so}(3)$ also admits the 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 the Bernoulli numbers ($B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}, \ldots$). The Bernoulli numbers satisfy the following series expression: $\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} x^n$ for any scalar $x \neq 0$.

Letting $[\mathbf{x}'] = [\mathbf{z}] - [\mathbf{y}] \in \mathrm{so}(3)$, we have

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

with

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

where

$$\mathbf{J}_l(\mathbf{y}) = \left(\sum_{n=0}^{\infty} \frac{B_n}{n!}[\mathbf{y}]^n\right)^{-1} \tag{50}$$

$$= \sum_{n=0}^{\infty} \frac{1}{(n+1)!}[\mathbf{y}]^n \tag{51}$$

$$= \int_0^1 \exp([\mathbf{y}]s)\, ds \tag{52}$$

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

From (1), a random variable $\mathbf{R} \in \mathrm{SO}(3)$ can be defined 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 $\hat{\mathbf{R}} \in \mathrm{SO}(3)$ is the state estimate. 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, with covariance $\mathbf{P}_{k+1|k}$ calculated by (20). Special caution is required when computing $\mathbf{P}_{k+1|k+1}$, which is the *a posteriori* right-invariant covariance of $(\mathbf{R}_{k+1}, \mathbf{b}_{k+1})$ after the measurement update. If one implements the measurement update as in standard vector space UKF, the state $(\mathbf{R}_{k+1}, \mathbf{b}_{k+1})$ is given by

$$\mathbf{R}_{k+1} = \exp([\boldsymbol{\xi}^{(a)}])\,\hat{\mathbf{R}}_{k+1|k} \tag{55}$$

$$\mathbf{b}_{k+1} = \hat{\mathbf{b}}_{k+1|k} + \boldsymbol{\xi}^{(b)} \tag{56}$$

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). Equation (55) is therefore reformulated to conform to (54) (i.e., to satisfy the property of "zero-mean" right-translated exponential noise). Assume that $(\mathbf{R}_{k+1}, \mathbf{b}_{k+1})$ can be represented as

$$\mathbf{R}_{k+1} = \exp([\boldsymbol{\epsilon}'^{(a)}])\,\hat{\mathbf{R}}_{k+1|k+1} \tag{57}$$

$$\mathbf{b}_{k+1} = \hat{\mathbf{b}}_{k+1|k+1} + \boldsymbol{\epsilon}'^{(b)} \tag{58}$$

where $\boldsymbol{\epsilon}' \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{\boldsymbol{\epsilon}'})$ and $\mathbf{P}_{k+1|k+1} = \mathbf{P}_{\boldsymbol{\epsilon}'}$. We now find $\mathbf{P}_{\boldsymbol{\epsilon}'}$.

Define the vector $\boldsymbol{\epsilon} \in \mathbb{R}^6$ by $\boldsymbol{\epsilon} := \boldsymbol{\xi} - \boldsymbol{\phi}$. $\boldsymbol{\epsilon}$ 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 $\boldsymbol{\xi} = \boldsymbol{\epsilon} + \boldsymbol{\phi}$, (55) can be rewritten as

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

Substituting (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 (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 by equating (56) and (58) using (27). If $\|\boldsymbol{\epsilon}\| \ll 1$, from the first-order approximation derived from the Baker–Campbell–Hausdorff formula in Appendix A, it follows that

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

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 $\boldsymbol{\phi}^{(a)}$, with corresponding closed-form equation given by (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 (59). This justifies (28) in Section III-B3. ([33] and [34] propose slightly different algorithms from (64): the former proposes a method for covariance correction of the quaternion state, while the latter takes a first-order approximation of both $\boldsymbol{\phi}^{(a)}$ and the noise vector $\boldsymbol{\epsilon}^{(a)}$ in the derivation. In contrast, (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 is subject to large accelerations, it outputs the vector sum of the negative gravitational acceleration vector and other accelerations due to external forces; the resulting acceleration vector measurement is expressed in the moving frame $\{\mathcal{B}\}$ attached to the IMU. In [35], these additional acceleration terms are referred to as motion disturbances. In magnetically disturbed environments, the measurement of a triaxial magnetometer deviates from the local magnetic field expressed in frame $\{\mathcal{B}\}$ coordinates.

To detect these disturbances, a number of reliability functions have been proposed [8], [35]. In [36], it is claimed that checking only the norms of the calibrated outputs of the accelerometers and magnetometers is in many cases sufficient for practical purposes. Let $\tilde{\mathbf{v}}_i \in \mathbb{R}^3$, $i = 1, 2$ be the unnormalized calibrated output vector of the three-axis accelerometer or magnetometer at a particular instant. If $|\|\tilde{\mathbf{v}}_i\| - 1| > \gamma_i$ for some positive threshold value γ_i, the disturbance is regarded as detected; otherwise no disturbance is presumed to exist.

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

1) Adaptation of noise covariances [37]: If a disturbance is detected, then the noise covariance of the Kalman filter is adjusted.
2) Measurement reconstruction with a vector selector [38]: If a 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 (19).

In our estimator, the measurement reconstruction method with a vector selector is used.

Appendix D
Extrinsic Mean of Unit Vectors

Proposition 1: Given a set of N unit vectors in $\mathbb{R}^d$, denoted $\mathcal{S}_v = \{\mathbf{v}_i \in \mathbb{R}^d \mid \|\mathbf{v}_i\| = 1, i = 1, \ldots, N\}$, 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: Defining $L(\mathbf{v}, \lambda) = \sum_{i=1}^N \|\mathbf{v}_i - \mathbf{v}\|^2 + \lambda(\mathbf{v}^T\mathbf{v} - 1)$ where $\lambda > 0$, the first-order necessary conditions for optimality ($\frac{\partial L(\mathbf{v}^*,\lambda)}{\partial \mathbf{v}^*} = 0$ and $\frac{\partial L(\mathbf{v}^*,\lambda)}{\partial \lambda} = 0$) yield the result.

Appendix E
Proof of (30)

Given the inverse $\bar{\mathbf{A}} \in$ SO(3) of the true attitude, consider the following slightly modified version of the optimization problem of (4):

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

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 (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-squares estimation problem, with the optimal estimate given as a linear function of

$\Delta \mathbf{v}_i$ as follows:

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

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 (66) has the form of a linear least-squares estimation problem, the covariance of $\boldsymbol{\theta}^*$ achieves the CRLB [39]. The covariance of $\boldsymbol{\theta}^*$ is therefore 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 (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. Automat. Lett.*, vol. 2, no. 2, pp. 404–411, Apr. 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, Feb. 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, Feb. 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. Autom. Control*, vol. 53, no. 5, pp. 1203–1218, Jun. 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, Sep. 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. Mechatronics*, vol. 21, no. 4, pp. 1900–1911, Aug. 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, Aug. 2008.

[11] A. Giannitrapani, N. Ceccarelli, F. Scortecci, and A. Garulli, "Comparison of EKF and UKF for spacecraft localization via angle measurements," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 1, pp. 75–84, Jan. 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. Imag. Vis.*, 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. Imag. Vis.*, 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. Autom. Control*, vol. 60, no. 2, pp. 436–449, Feb. 2015.

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

[16] S. Hauberg, F. Lauze, and K. S. Pedersen, "Unscented Kalman filtering on Riemannian manifolds," *J. Math. Imag. Vis.*, 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," *Inf. 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. Mach. Intell.*, vol. 36, no. 4, pp. 625–643, Apr. 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, Dec. 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, Jun. 2014.

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

[25] G. Wahba, "A least–squares estimate of satellite attitude," *SIAM Rev.*, vol. 7, no. 3, p. 409, 1965. Q3

[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. Basel, 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, Jul. 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, Oct. 2009.

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

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

Since 2000, he has been with the Korea Institute of Science and Technology, Seoul, South Korea, where he is currently a Senior Researcher. His research interests include computer vision and signal processing.

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

His research interests include robot planning and control, and manifold learning from high-dimensional data.

Frank C. Park (F'13) received the B.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1985, and the Ph.D. degree in applied mathematics from Harvard University, Cambridge, MA, USA, in 1991.

After joining the faculty of UC Irvine in 1991, since 1995, he has been a Professor in mechanical and aerospace engineering with Seoul National University, Seoul, South Korea, where he is currently the Department Head. He has been an IEEE Robotics and Automation Society Distinguished Lecturer, and has held Adjunct Faculty positions with the New York University Courant Institute; the Interactive Computing Department, Georgia Tech; and the HKUST Robotics Institute. His research interests include robot mechanics, planning and control, vision and image processing, and related areas of applied mathematics.

Dr. Park is a former Editor-in-Chief for 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."

GENERAL INSTRUCTION

- **Authors: We cannot accept new source files as corrections for your paper. If possible, please annotate the PDF proof we have sent you with your corrections and upload it via the Author Gateway. Alternatively, you may send us your corrections in list format. You may also upload revised graphics via the Author Gateway.**

Queries

Q1. Author: Please confirm or add details for any funding or financial support for the research of this article.
Q2. Author: Acronym UKF has been used for both "unscented Kalman filter" and "unscented Kalman filtering." Please check.
Q3. Author: Please provide the complete page range for Ref. [25].

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 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 IMU measurements of gravity and the earth's magnetic field. 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 update. The resulting attitude estimates are invariant with respect to choice of fixed and moving reference frames. The performance advantages of our filter vis-à-vis existing state-of-the-art IMU attitude estimation algorithms are validated via numerical and hardware experiments involving both synthetic and real data.**

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

I. Introduction

ESTIMATING an object's orientation, or attitude, from an inertial measurement unit (IMU) attached to the object arises in applications ranging 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 velocities (which can be integrated to calculate the attitude), the accelerometer measures accelerations due to gravity and other external forces, and the magnetometer measures the earth's magnetic field. Gyroscopic measurements contain a time-varying bias error, and accelerometer and magnetometer measurements can be used to identify and compensate for this gyroscope bias. More generally, the challenges and benefits of simultaneously estimating the attitude and gyroscope bias from disparate sensor measurements are detailed in [5] and the cited references.

Notable among deterministic filtering methods for simultaneously estimating attitude and gyroscope bias are Mahony *et al.*'s series of nonlinear complementary filters (NCFs) [6]–[8]; these filters ensure almost global stability of the observer error, and their performance has been validated in numerous experimental scenarios. Stochastic filtering methods further take into account statistical characterizations of measurement and process 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]. Q2

Because the underlying configuration space of rotations, represented by the group SO(3) of 3×3 real orthogonal matrices with unit determinant, is not a vector space but a curved space, the attitude estimation problem is fundamentally a nonlinear one. The straightforward but naive approach of expressing a rotation in terms of some suitable local coordinates (e.g., roll-pitch-yaw angles, Euler angles) is problematic at several levels: the local coordinates contain singularities that require special treatment (for example, when the pitch angle is 90°), and the resulting estimates depend both on the choice of local coordinates as well as fixed and moving reference frames. If standard vector space filters are naively adapted to local coordinate representations of the attitude, not only are the equations for the state space dynamics and measurements highly nonlinear and dependent on the choice of reference frames, but filtering performance is highly uneven throughout different regions of the configuration space.

Recent research has attempted to address the issue of coordinate and reference frame dependency through the use of differential geometric methods. Although computationally more involved than standard vector space filtering algorithms, when correctly formulated, these methods are invariant with respect to the choice of fixed and moving reference frames, and also independent of the choice of local coordinates used to parameterize the rotations. For estimation problems in which the underlying configuration space has the structure of a matrix Lie group like SO(3), 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

Q1 Manuscript received February 23, 2018; revised July 1, 2018 and September 6, 2018; accepted December 31, 2018. Recommended by Technical Editor T. Singh. This work was supported in part by the NAVER LABS, in part by the ADD-ICMTC, in part by the SNU-IAMD, in part by the BK21+, in part by the MI Technology Innovation Program 10048320, in part by the National Research Foundation of Korea under Grant NRF-2016R1A5A1938472, in part by MOTIE Technology Innovation Program under Grant 2017-10069072, in part by SNU BMRR Center under Grant DAPA-UD160027ID, and in part by the KIST Institutional Project. *(Corresponding author: Donghoon Kang.)*

D. Kang is with the Imaging Media Center, Korea Institute of Science and Technology, Seoul 136-791, South 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 151-744, South Korea (e-mail: jchastro@gmail.com; fcp@snu.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMECH.2019.2891776

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 often prove to be difficult 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 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 update, and also by drawing upon Lie-theoretic properties in key steps of our geometric UKF algorithm.

This 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. Section V compares the performance of our geometric UKF algorithm against other existing state-of-the-art estimators for attitude and gyroscope bias [6], [19], [20], with detailed experiments involving both synthetic and real data validating the 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 represented by the 3×3 real matrices $\mathbf{R}$ satisfying $\mathbf{R}^T\mathbf{R} = \mathbf{I}$ and $\det \mathbf{R} = 1$, where $\mathbf{I}$ denotes the 3×3 identity matrix. SO(3) is an example of a matrix Lie group; 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 $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)^T \in \mathbb{R}^3$. A fundamental connection between so(3) and SO(3) is the matrix exponential map $\exp$: so(3) $\rightarrow$ SO(3), given as

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

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

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

where θ satisfies $1 + 2\cos\theta = \mathrm{tr}(\mathbf{R})$, $|\theta| < \pi$ [here, tr$(\cdot)$ denotes the trace of a matrix]. If 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 aforementioned 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 $\mathbf{P}, \mathbf{Q} \in$ SO(3). With this notion of distance, 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 $\boldsymbol{\Omega} = \log(\mathbf{R}_1^T\mathbf{R}_2) \in$ so(3).

Recalling that $\mathbb{R}^3$ is also trivially a Lie group under vector addition, the direct product 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

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

where $\mathbf{X}_0 \in$ SO(3) is given and $\boldsymbol{\eta} \in \mathbb{R}^3$ is a zero-mean Gaussian with covariance $\mathbf{P}_\eta$, i.e., $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_\eta)$. 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 $[\boldsymbol{\zeta}] \in$ 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

$$\boldsymbol{\eta} = \mathbf{X}_0\boldsymbol{\zeta} \quad (2)$$

$$\mathbf{P}_\eta = \mathbf{X}_0\mathbf{P}_\zeta\mathbf{X}_0^T. \quad (3)$$

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

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

III. UKF Algorithm for Estimating Attitude and Gyroscope Bias

Before describing our geometric UKF algorithm, we fix notation, describe the sensor models and their underlying assumptions, and review Wahba's Problem [25] and its solutions.

Let $\{\mathcal{I}\}$ be the inertial reference frame fixed to ground, and let $\{\mathcal{B}\}$ denote the body frame fixed to the moving IMU. Let $\boldsymbol{\omega}^m \in \mathbb{R}^3$ be the angular velocity measured by the IMU gyroscope with respect to frame $\{\mathcal{B}\}$. Denote by $\mathbf{a}, \mathbf{m} \in \mathbb{R}^3$ the IMU accelerometer and magnetometer measurements, respectively; like $\boldsymbol{\omega}^m$, both $\mathbf{a}$ and $\mathbf{m}$ are assumed measured with respect to the

IMU frame $\{\mathcal{B}\}$. Further define the unit vectors $\mathbf{v}_1 := \mathbf{a}/\|\mathbf{a}\|$, $\mathbf{v}_2 := \mathbf{m}/\|\mathbf{m}\|$.

In what follows, we assume that the IMU is suitably calibrated, and that the gravitational acceleration is dominant in the accelerometer measurement $\mathbf{a}$. Let $\mathbf{r}_1 \in \mathbb{R}^3$ be the unit vector in the opposite direction of gravity, and $\mathbf{r}_2 \in \mathbb{R}^3$ be the unit vector in the direction of the earth's magnetic field. If $\mathbf{r}_1$ and $\mathbf{r}_2$ are not collinear, then $\mathbf{r}_i$ and $\mathbf{v}_i$ should satisfy $\mathbf{r}_i = \mathbf{R}\mathbf{v}_i$, $i = 1, 2$, for some rotation $\mathbf{R} \in$ SO(3) representing the orientation of the IMU frame $\{\mathcal{B}\}$ relative to the fixed frame $\{\mathcal{I}\}$. Since in practice IMU measurements are noisy, $\mathbf{R}$ is typically estimated as the solution to the following optimization problem (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 the w_i are positive weights. A popular choice for w_i is $w_i = 1/\sigma_i^2$, where σ_i^2 denotes the variance of $\mathbf{v}_i$ in the direction normal to $\mathbf{R}^T\mathbf{r}_i$ [27]. (Equivalently, the normalized weights $w_i = \sigma_{\mathrm{tot}}^2/\sigma_i^2$, where $1/\sigma_{\mathrm{tot}}^2 = \sum_{i=1}^2 (1/\sigma_i^2)$ are also widely used [28].)

Wahba's Problem as defined by (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 $\mathbf{F} := \sum_{i=1}^2 w_i \mathbf{v}_i \mathbf{r}_i^T = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$. The matrix $\mathbf{D}$ in (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 earlier do not take into account the state space dynamics of the object or process and measurement noise characteristics, and typically are inferior to estimates obtained via nonlinear stochastic filtering techniques. We now formulate the overall problem in a discrete-time stochastic filtering setting. 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 zero-mean Gaussian noise. The state dynamics are then assumed to be of the form

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

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

where h is the integration time step, 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 a first-order linear approximation of the state dynamics (7) that leads to a closed-form expression for the covariance of $\mathbf{R}_{k+1}$ 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 $[\cdot, \cdot]$: so(3) × so(3) → 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 (9) in the form

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

Gathering only terms linear in $\mathbf{x}$ in (10), the following approximation between $\mathbf{x}$ and $\mathbf{x}'$ holds for $\|\mathbf{x}\|$ sufficiently small [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 (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 (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 in the derivation of (15).

Note that $\mathbf{l}_k = -\mathbf{R}_k\mathbf{J}_l(\boldsymbol{\psi})\boldsymbol{\eta}_k h$ is itself a random variable, since it is 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 equations are assumed 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 as a solution to Wahba's Problem (4) using IMU gravitational acceleration and magnetic field measurements. The measurement noise $\mathbf{w}_{k+1} \in \mathbb{R}^3$ is assumed to be zero-mean Gaussian, implying that the measurement vector statistics are rotationally symmetric about their true measurement vectors.

B. UKF Algorithm

We now present the geometric UKF algorithm for simultaneous attitude and gyroscope bias estimation. Let $\mathbf{R}_k$ and $\mathbf{b}_k$, respectively, denote the attitude and the gyroscope bias at time step k, and $\mathbf{X}_k := (\mathbf{R}_k, \mathbf{b}_k) \in$ SO(3) × $\mathbb{R}^3$.

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

Algorithm 1: Weighted Intrinsic Mean on SO(3).

```
   Input: Set of rotations {Z_0, ..., Z_12} in SO(3)
1  T ← Z_0
2  for j ← 0 to n do
3  |  Λ ← Σ_{i=0}^{12} w_m^(i) log(Z_i T^{-1})
4  |_ T ← exp(Λ) T
5  return T
```

given. From (5), $\hat{\mathbf{R}}_{0|0}$ is estimated by solving Wahba's Problem (4) from a pair of initial measurement vectors $(\mathbf{v}_1, \mathbf{v}_2)$.

2) Time Update:

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

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

where following the work presented in [10], the parameter γ is chosen as $\gamma = \sqrt{N_x + \lambda}$, with N_x set to the state dimension (six) and $\lambda = N_x(\alpha^2 - 1)$, $0 < \alpha < 1$; $\mathbf{s}_i \in \mathbb{R}^6$ is the ith column vector of the lower-triangular matrix $\mathbf{S} \in \mathbb{R}^{6\times 6}$ in the Cholesky decomposition $\mathbf{P}_{k|k} = \mathbf{S}\mathbf{S}^T$, and $\mathbf{s}_i^{(a)}, \mathbf{s}_i^{(b)} \in \mathbb{R}^3$ are, respectively, the upper and lower halves of $\mathbf{s}_i$.

2) Setting $\mathbf{l}_k = \mathbf{0}$ in (15) and $\mathbf{n}_k = \mathbf{0}$ in (8), define a set of sigma points $\{(\Upsilon_{\mathbf{R},k+1}^{(i)}, \Upsilon_{\mathbf{b},k+1}^{(i)}) \in \mathrm{SO(3)} \times \mathbb{R}^3 | i = 0, \ldots, 12\}$ as

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

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

3) Given the set of rotations $\{\Upsilon_{\mathbf{R},k+1}^{(0)}, \ldots, \Upsilon_{\mathbf{R},k+1}^{(12)}\}$ in SO(3), evaluate the weighted mean rotation $\bar{\Upsilon}_{\mathbf{R},k+1} \in \mathrm{SO(3)}$ using Algorithm 1. Taking advantage of the rapid convergence of Algorithm 1 [18], [30], set the number of iterations in line 2 of the algorithm to $n = 3$ or 4. The weights $w_m^{(i)} \in \mathbb{R}$ in line 3 satisfy $\sum_{i=0}^{12} w_m^{(i)} = 1$.

4) The gyroscope bias estimate $\bar{\Upsilon}_{\mathbf{b},k+1} \in \mathbb{R}^3$ is given by the weighted mean of $\{\Upsilon_{\mathbf{b},k+1}^{(0)}, \ldots, \Upsilon_{\mathbf{b},k+1}^{(12)}\}$ in $\mathbb{R}^3$, i.e., $\bar{\Upsilon}_{\mathbf{b},k+1} = \sum_{i=0}^{12} w_m^{(i)} \Upsilon_{\mathbf{b}}$. $\hat{\mathbf{X}}_{k+1|k} := (\hat{\mathbf{R}}_{k+1|k}, \hat{\mathbf{b}}_{k+1|k})$ is then

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

5) Define the vectors $[\mathbf{q}_i^{(a)}] := \log(\Upsilon_{\mathbf{R},k+1}^{(i)} \bar{\Upsilon}_{\mathbf{R},k+1}^{-1}) \in \mathrm{so(3)}$ 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 $\mathbf{q}_i = (\mathbf{q}_i^{(a)}, \mathbf{q}_i^{(b)}) \in \mathbb{R}^6$. The predicted covariance is given by

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

where $w_c^{(i)} \in \mathbb{R}$ are the weights and $\mathbf{N}_k = \left[\begin{smallmatrix} (ch^2)\mathbf{I} & \mathbf{0} \\ \mathbf{0} & d\mathbf{I} \end{smallmatrix}\right]$ is the process noise covariance.

6) Let $\mathbf{u}_i \in \mathbb{R}^6$ denote the ith column vector of the lower-triangular matrix $\mathbf{U} \in \mathbb{R}^{6\times 6}$ in the Cholesky decomposition $\mathbf{P}_{k+1|k} = \mathbf{U}\mathbf{U}^T$. The upper and lower halves of $\mathbf{u}_i$ are, respectively, denoted $\mathbf{u}_i^{(a)} \in \mathbb{R}^3$ and $\mathbf{u}_i^{(b)} \in \mathbb{R}^3$. Redraw the sigma points $\mathcal{X}_{k+1}^{(i)} := (\mathcal{X}_{\mathbf{R},k+1}^{(i)}, \mathcal{X}_{\mathbf{b},k+1}^{(i)})$, $(i = 0, \ldots, 12)$ from $\hat{\mathbf{X}}_{k+1|k}$ and $\mathbf{P}_{k+1|k}$ as follows:

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

3) Measurement Update:

1) If the IMU moves with high acceleration or is subject to magnetic disturbances, the accelerometer and magnetometer measurements may be corrupted and not satisfy our earlier assumptions. Appendix C summarizes some existing methods for addressing these disturbances.

2) Setting $\mathbf{w}_{k+1} = \mathbf{0}$ in (16), define the set of measurement sigma points $\mathcal{S}_{\mathcal{Y}} = \{\mathcal{Y}_{k+1}^{(i)} \in \mathrm{SO(3)} \mid i = 0, \ldots, 12\}$ as follows:

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

3) The mean $\hat{\mathbf{Y}}_{k+1}$ of $\{\mathcal{Y}_{k+1}^{(0)}, \ldots, \mathcal{Y}_{k+1}^{(12)}\}$ is given by

$$\hat{\mathbf{Y}}_{k+1} = \hat{\mathbf{R}}_{k+1|k} \quad (22)$$

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

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

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

$$\mathbf{P}_{\mathbf{vv}} = \mathbf{P}_{\mathbf{yy}} + \mathbf{W}_{k+1} \quad (24)$$

where $\mathbf{W}_{k+1}$ is the right-invariant covariance of the solution to Wahba's Problem. In Section IV, we derive a closed-form expression for $\mathbf{W}_{k+1}$ from (32).

4) Define $[\mathbf{p}_i^{(a)}] := \log(\mathcal{X}_{\mathbf{R},k+1}^{(i)} \hat{\mathbf{R}}_{k+1|k}^{-1}) \in \mathrm{so(3)}$ and $\mathbf{p}_i^{(b)} := \mathcal{X}_{\mathbf{b},k+1}^{(i)} - \hat{\mathbf{b}}_{k+1|k} \in \mathbb{R}^3$, and $\mathbf{p}_i = (\mathbf{p}_i^{(a)}, \mathbf{p}_i^{(b)}) \in$

$\mathbb{R}^6$. The associated covariance $\mathbf{P_{xy}}$ is then calculated as

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

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

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

where $\mathbf{Y}_{k+1}$ and $\hat{\mathbf{Y}}_{k+1}$ are, respectively, given by (16) and (22). Define $\boldsymbol{\phi}^{(a)} \in \mathbb{R}^3$ and $\boldsymbol{\phi}^{(b)} \in \mathbb{R}^3$ to be the upper and lower halves of $\boldsymbol{\phi} := \mathbf{K}\boldsymbol{\delta} \in \mathbb{R}^6$. The state and covariance are now updated according to

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

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

where $\mathbf{M}(\boldsymbol{\phi}^{(a)}) \in \mathbb{R}^{6\times 6}$ is given by

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

The justification for $\mathbf{M}(\boldsymbol{\phi}^{(a)})$ in (28) is given in Appendix B.

IV. Measurement Noise Covariance

This section presents an algorithm for obtaining, from a set of noisy unit vector measurements of the gravity and magnetic field vectors, a full-rank measurement noise covariance matrix.

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 the solution to Wahba's Problem (4):

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

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

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

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

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

Equation (32) follows from a straightforward calculation combining (3) and (30).

Note that the left-invariant covariance of $\mathbf{R}$ in (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 (32) is the covariance of the solution to Wahba's Problem represented with respect to the fixed ground frame. If values for $\sigma_i^2, \mathbf{r}_i$ are given, the right-invariant covariance of $\mathbf{R}$ in (32) can be determined to be a constant matrix, independent of $\bar{\mathbf{A}}$. However, the left-invariant covariance of $\mathbf{R}$ in (30) requires $\bar{\mathbf{A}}$, σ_i^2, and $\mathbf{r}_i$.

When the IMU is moving, $\bar{\mathbf{A}}$ is also changing, and the left-invariant covariance of $\mathbf{R}$ needs to be updated at every time step. The left-invariant covariance can be evaluated as the inverse of a matrix that varies with $\bar{\mathbf{A}}$, while the right-invariant covariance remains invariant. When the IMU motion involves both translation and rotation, measurements of the two direction vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are subject to greater errors, leading to less accurate estimates of $\bar{\mathbf{A}}$. For the reasons outlined earlier, our measurement noise covariance formula of (32) is preferable to Shuster's formula (30) in the geometric UKF algorithm.

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

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

1) Constant Vectors $(\mathbf{r}_1, \mathbf{r}_2)$*:* Assign each axis of the inertial reference frame $\{\mathcal{I}\}$ as follows: The direction opposite to gravity is set to be the y-axis of $\{\mathcal{I}\}$, while the x-axis of $\{\mathcal{I}\}$ is orthogonal to both gravity and the earth's magnetic field. With these assignments, $\mathbf{r}_1 = (0, 1, 0)^T$ and

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

where ϕ is unknown and to be determined.

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

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

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

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

Let the SVD of $\mathbf{M}_t$ be $\mathbf{M}_t = \mathbf{U}_t\boldsymbol{\Sigma}_t\mathbf{V}_t^T$, where in principle $\boldsymbol{\Sigma}_t = \mathrm{diag}(\sigma_i^2, \sigma_i^2, 0)$ and $\check{\mathbf{v}}_i$ is the corresponding direction for the singular value 0. Since in practice ground-truth values of $\check{\mathbf{v}}_i$ are unavailable, an alternative method of determining σ_i^2 is needed. Assuming that the IMU is stationary and N measurements are available, the covariance of $\mathbf{v}_i$ can be estimated by

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

where $\mathbf{v}_i^{(j)}$ denotes the jth measurement vector obtained from the ith sensor (sensor 1 is the accelerometer, while sensor 2 is the magnetometer). Let the SVD of $\mathbf{M}_a$ be $\mathbf{M}_a = \mathbf{U}_a\boldsymbol{\Sigma}_a\mathbf{V}_a^T$, where $\boldsymbol{\Sigma}_a = \mathrm{diag}(s_1, s_2, s_3)$ and $s_1 \geq s_2 \geq s_3$, $s_3 \approx 0$. $\boldsymbol{\Sigma}_a$ will typically be close to its theoretical value $\boldsymbol{\Sigma}_t$, in which case we can set

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

V. Experimental Results

In this section, we compare the performance of our geometric UKF algorithm [UKF on SO(3)] against other state-of-the-art methods (UKF on Quaternion [19], EKF on Quaternion [20], and the passive NCF (NCF on SO(3)) [6]). Using both synthetic and real data in our experiments, both the convergence rate and accuracy of the attitude and gyroscope bias estimates are compared.

Ground-truth values 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 filter update time step is set to $h_0 = 1/60$ seconds. 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 set to

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

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

α in (40) is set to 0.9, and β is set to two for a Gaussian prior [10].

A. Synthetic Data

In our numerical simulation experiments, the vectors in (4) are set to $\mathbf{r}_1 = (0, 1, 0)^T$ and $\mathbf{r}_2 = (0, \cos(\phi_s), \sin(\phi_s))^T$, where $\phi_s = 2.4$ rad. The ground-truth value $\breve{\mathbf{R}}_1 \in \mathrm{SO}(3)$ is set randomly to be the initial attitude.

For realistic simulation, we first collect a set of real angular rate measurements $\breve{\boldsymbol{\omega}}_k$ from an actual gyroscope (L3G4200D) at the 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 value of the initial gyroscope bias is set to be $\breve{\mathbf{b}}_0 = (-0.06, 0.3, 0.3)^T$ rad/s. We then 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, $\sigma_0 = (1.1 \times 10^{-3}/h_0)$ rad/s, $\sigma_1 = (1.0 \times 10^{-5})$ rad/s, $\sigma_2 = 1.00 \times 10^{-2}$, and $\sigma_3 = 1.58 \times 10^{-2}$.

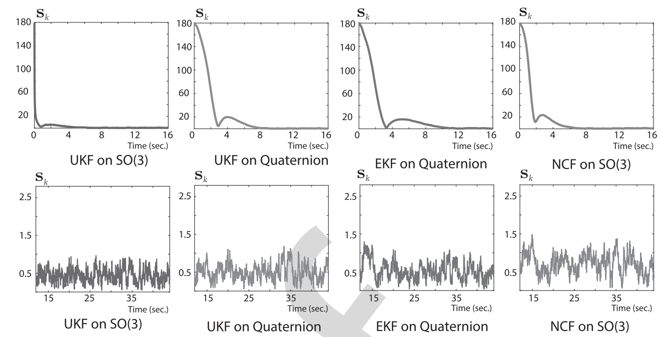


Fig. 1. Simulation experiments: Attitude estimation errors (in degrees) over the time intervals $t \in [0, 16]$ s (top) and $t \in [12, 44]$ s (bottom).

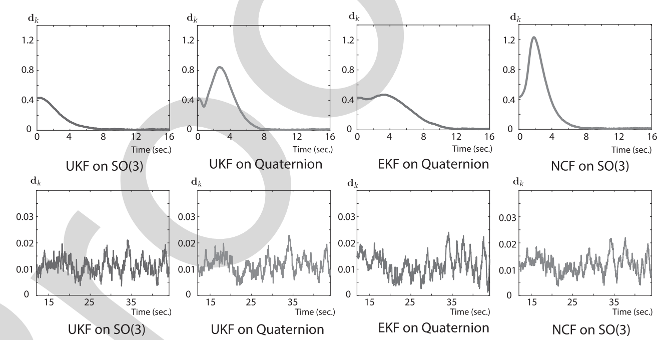


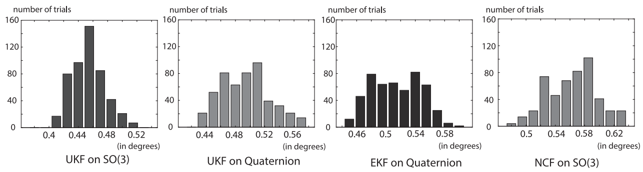
Fig. 2. Simulation experiments: Gyroscope bias estimate errors (in radian/seconds) over the time intervals $t \in [0, 16]$ s (top) and $t \in [12, 44]$ s (bottom).

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 (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 $\mathbf{W}_{k+1} = (\frac{1}{\sigma_2^2}(\mathbf{I} - \mathbf{r}_1\mathbf{r}_1^T) + \frac{1}{\sigma_3^2}(\mathbf{I} - \mathbf{r}_2\mathbf{r}_2^T))^{-1}$.

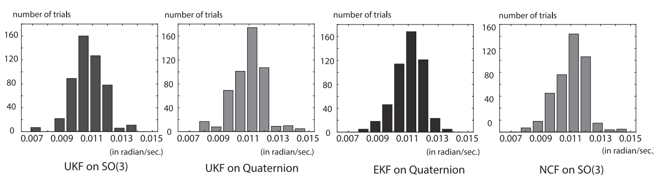
From the simulation results shown in Figs. 1 and 2, it can be seen that the proposed algorithm [UKF on SO(3)] converges most rapidly over the time interval $t \in [0, 14]$ s. To more reliably assess the accuracy of each estimator, we generate 500 sets of synthetic data using (42)–(45). Fig. 3 shows the histograms of estimation errors of the attitudes and the slowly time-varying gyroscope biases. Tables I and II summarize the experimental results corresponding to Fig. 3(a) and (b). From Fig. 3(b) and Table II, it can be seen that the gyroscope bias estimates show similar performance for all estimators. In terms of attitude estimates, "UKF on SO(3)" is the most accurate among the estimators [see Fig. 3(a) and Table I].

B. Real Experiments

The IMU for real experiments consists of an L3G4200D gyroscope, LIS3LV02DQ accelerometer, HMC5883L magnetometer, and Cortex-M3 microcontroller. In real experiments, ground-truth values of the slowly time-varying gyroscope bias



(a)



(b)

Fig. 3. Simulation experiments: Histograms of estimation errors over the time interval $t \in [12, 44]$ s (averaged over 500 trials). (a) Histograms of attitude estimation errors. (b) Histograms of gyroscope bias estimation errors.

TABLE I
AVERAGE AND STANDARD DEVIATION OF ATTITUDE ESTIMATION ERRORS (IN DEGREES) OVER THE TIME INTERVAL $t \in [12, 44]$ S (AVERAGED 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

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

are unknown. We therefore assume that the gyroscope bias is initially unknown, but near-constant over short time durations. If the IMU is stationary, then the gyroscope bias, denoted $\breve{\mathbf{b}}$, can be temporarily captured by averaging a set of gyroscope data over a certain time interval [32].

Keeping the IMU stationary, the variance σ_i^2 of the unit vector $\mathbf{v}_{i,k}$, $i = 1, 2$, can be calculated from (37); in our experiments we obtain the values $\sigma_1^2 = 8.95 \times 10^{-5}$ and $\sigma_2^2 = 1.911 \times 10^{-4}$. Denoting by ϕ_r the angle between $\mathbf{r}_1$ and $\mathbf{r}_2$, i.e., $\phi_r = \cos^{-1}(\mathbf{r}_1^T \mathbf{r}_2)$, we obtain $\phi_r = 2.486$ rad using Proposition 1 of Appendix D and (34). The noise covariances $\mathbf{N}_k$ in (20) and $\mathbf{W}_{k+1}$ in (24) of the proposed attitude estimator [UKF on SO(3)] are set as follows: $\mathbf{N}_k = \left[\begin{smallmatrix} (2.0\times10^{-9})\mathbf{I} & \mathbf{0} \\ \mathbf{0} & (3.0\times10^{-11})\mathbf{I} \end{smallmatrix}\right]$ 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 value of the attitude $\breve{\mathbf{R}}_k$ at time step k, we use the optical motion capture system OptiTrack 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 $\{(\boldsymbol{\omega}_k^m, \mathbf{v}_{1,k}, \mathbf{v}_{2,k}) \mid k = 1, \ldots, N_r\}$ obtained from the moving IMU, and the ground-truth attitude $\breve{\mathbf{R}}_k$ obtained from the OptiTrack infrared camera system, are synchronously saved into files at a sampling rate

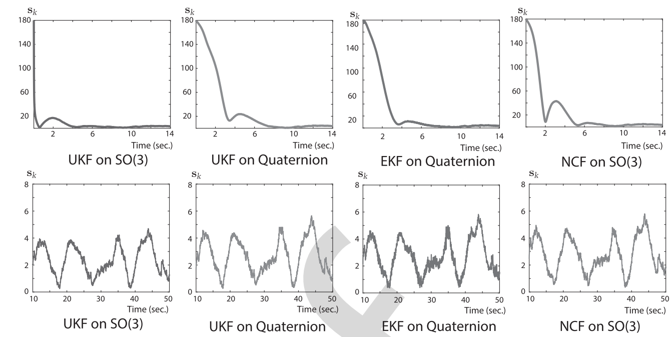


Fig. 4. Real experiments: Attitude estimate errors (in degrees) over the time interval $t \in [0, 14]$ s (top) and $t \in [10, 50]$ s (bottom).

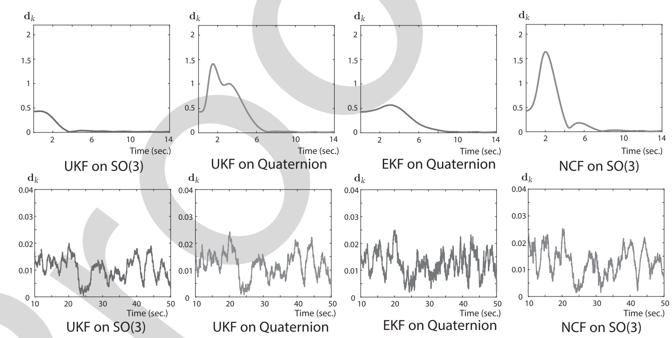


Fig. 5. Real experiments: Gyroscope bias estimate errors (in radian/seconds) over the time intervals $t \in [0, 14]$ s (top) and $t \in [10, 50]$ s (bottom).

TABLE III
RESULTS OF REAL EXPERIMENTS: AVERAGE ERRORS OVER THE TIME INTERVAL $t \in [10, 50]$ S (AVERAGED OVER TEN 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

$1/h_0 = 60$ Hz. Here, the number of measurements N_r is set to 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: $\hat{\mathbf{b}}_{1|1} = \breve{\mathbf{b}} + (1/h_0)(-0.001, 0.005, 0.005)^T = \breve{\mathbf{b}} + (-0.06, 0.3, 0.3)^T$ (rad/s) 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, Figs. 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 further experimentally verify these results, we collect nine additional sets of real data. As shown in Table III, "UKF on SO(3)" demonstrates superior performance compared to existing methods in terms of the accuracy of attitude estimates.

TABLE IV
AVERAGE COMPUTATION TIMES FOR 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

We also measure, at every time step, the computation times for each filter—all implemented in C++ and executed on a desktop computer with Intel i5-4670 (3.4 GHz) CPU. The computation times for each estimator are averaged over N_r steps. From Table IV it can be seen that "NCF on SO(3)" is the fastest among the estimators. Computation times for "UKF on SO(3)" are similar to those for "Quaternion UKF."

VI. CONCLUSION

This paper has presented a geometric unscented Kalman filtering algorithm for simultaneously estimating attitude and gyroscope bias from an inertial measurement unit. Drawing upon the Lie group properties of the set of rotation matrices SO(3), we derive a discrete-time stochastic nonlinear filtering algorithm evolving on SO(3) $\times\mathbb{R}^3$. One of the key features of our algorithm is to express observations as elements of SO(3), by determining the rotation corresponding to the IMU's gravitational acceleration and magnetic field vector measurements as a solution to Wahba's Problem. By doing so, first-order linear approximations of the state dynamics and measurement equations lead to closed-form equations for covariance propagation and update. These in turn lead to computationally efficient implementations of our filter, with the resulting attitude estimates invariant with respect to the choice of fixed and moving reference frames. Extensive numerical simulation and hardware experiments have demonstrated the superior convergence behavior and estimation accuracy of our proposed algorithm compared to existing state-of-the-art IMU estimators for attitude and gyroscope bias.

APPENDIX A
FIRST-ORDER APPROXIMATION OF EXPONENTIAL MAP

Given $[\mathbf{x}], [\mathbf{y}] \in$ so(3), let $[\mathbf{z}] \in$ so(3) satisfy

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

From the Baker–Campbell–Hausdorff formula [29], we have

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

The Lie bracket operator $[\cdot,\cdot]$: so(3) $\times$ so(3) $\to$ 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). $[\mathbf{c}] = [[\mathbf{a}],[\mathbf{b}]] \in$ so(3) also admits the 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 the Bernoulli numbers ($B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}, \ldots$). The Bernoulli numbers satisfy the following series expression: $\frac{x}{e^x-1} = \sum_{n=0}^{\infty}\frac{B_n}{n!}x^n$ for any scalar $x \neq 0$.

Letting $[\mathbf{x}'] = [\mathbf{z}] - [\mathbf{y}] \in$ so(3), we have

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

with

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

where

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

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

From (1), a random variable $\mathbf{R} \in$ SO(3) can be defined 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. 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, with covariance $\mathbf{P}_{k+1|k}$ calculated by (20). Special caution is required when computing $\mathbf{P}_{k+1|k+1}$, which is the *a posteriori* right-invariant covariance of $(\mathbf{R}_{k+1}, \mathbf{b}_{k+1})$ after the measurement update. If one implements the measurement update as in standard vector space UKF, the state $(\mathbf{R}_{k+1}, \mathbf{b}_{k+1})$ is given by

$$\mathbf{R}_{k+1} = \exp([\boldsymbol{\xi}^{(a)}])\,\hat{\mathbf{R}}_{k+1|k} \tag{55}$$

$$\mathbf{b}_{k+1} = \hat{\mathbf{b}}_{k+1|k} + \boldsymbol{\xi}^{(b)} \tag{56}$$

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). Equation (55) is therefore reformulated to conform to (54) (i.e., to satisfy the property of "zero-mean" right-translated exponential noise). Assume that $(\mathbf{R}_{k+1}, \mathbf{b}_{k+1})$ can be represented as

$$\mathbf{R}_{k+1} = \exp([\boldsymbol{\epsilon}'^{(a)}])\, \hat{\mathbf{R}}_{k+1|k+1} \tag{57}$$

$$\mathbf{b}_{k+1} = \hat{\mathbf{b}}_{k+1|k+1} + \boldsymbol{\epsilon}'^{(b)} \tag{58}$$

where $\boldsymbol{\epsilon}' \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_{\boldsymbol{\epsilon}'})$ and $\mathbf{P}_{k+1|k+1} = \mathbf{P}_{\boldsymbol{\epsilon}'}$. We now find $\mathbf{P}_{\boldsymbol{\epsilon}'}$.

Define the vector $\boldsymbol{\epsilon} \in \mathbb{R}^6$ by $\boldsymbol{\epsilon} := \boldsymbol{\xi} - \boldsymbol{\phi}$. $\boldsymbol{\epsilon}$ 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 $\boldsymbol{\xi} = \boldsymbol{\epsilon} + \boldsymbol{\phi}$, (55) can be rewritten as

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

Substituting (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 (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 by equating (56) and (58) using (27). If $\|\boldsymbol{\epsilon}\| \ll 1$, from the first-order approximation derived from the Baker–Campbell–Hausdorff formula in Appendix A, it follows that

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

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 $\boldsymbol{\phi}^{(a)}$, with corresponding closed-form equation given by (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 (59). This justifies (28) in Section III-B3. ([33] and [34] propose slightly different algorithms from (64): the former proposes a method for covariance correction of the quaternion state, while the latter takes a first-order approximation of both $\boldsymbol{\phi}^{(a)}$ and the noise vector $\boldsymbol{\epsilon}^{(a)}$ in the derivation. In contrast, (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 is subject to large accelerations, it outputs the vector sum of the negative gravitational acceleration vector and other accelerations due to external forces; the resulting acceleration vector measurement is expressed in the moving frame $\{\mathcal{B}\}$ attached to the IMU. In [35], these additional acceleration terms are referred to as motion disturbances. In magnetically disturbed environments, the measurement of a triaxial magnetometer deviates from the local magnetic field expressed in frame $\{\mathcal{B}\}$ coordinates.

To detect these disturbances, a number of reliability functions have been proposed [8], [35]. In [36], it is claimed that checking only the norms of the calibrated outputs of the accelerometers and magnetometers is in many cases sufficient for practical purposes. Let $\tilde{\mathbf{v}}_i \in \mathbb{R}^3$, $i = 1, 2$ be the unnormalized calibrated output vector of the three-axis accelerometer or magnetometer at a particular instant. If $|\|\tilde{\mathbf{v}}_i\| - 1| > \gamma_i$ for some positive threshold value γ_i, the disturbance is regarded as detected; otherwise no disturbance is presumed to exist.

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

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

In our estimator, the measurement reconstruction method with a vector selector is used.

APPENDIX D
EXTRINSIC MEAN OF UNIT VECTORS

Proposition 1: Given a set of N unit vectors in $\mathbb{R}^d$, denoted $\mathcal{S}_v = \{\mathbf{v}_i \in \mathbb{R}^d \mid \|\mathbf{v}_i\| = 1, i = 1, \ldots, N\}$, 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: Defining $L(\mathbf{v}, \lambda) = \sum_{i=1}^N \|\mathbf{v}_i - \mathbf{v}\|^2 + \lambda(\mathbf{v}^T\mathbf{v} - 1)$ where $\lambda > 0$, the first-order necessary conditions for optimality ($\frac{\partial L(\mathbf{v}^*, \lambda)}{\partial \mathbf{v}^*} = 0$ and $\frac{\partial L(\mathbf{v}^*, \lambda)}{\partial \lambda} = 0$) yield the result.

APPENDIX E
PROOF OF (30)

Given the inverse $\bar{\mathbf{A}} \in$ SO(3) of the true attitude, consider the following slightly modified version of the optimization problem of (4):

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

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 (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-squares estimation problem, with the optimal estimate given as a linear function of

$\Delta \mathbf{v}_i$ as follows:

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

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 (66) has the form of a linear least-squares estimation problem, the covariance of $\boldsymbol{\theta}^*$ achieves the CRLB [39]. The covariance of $\boldsymbol{\theta}^*$ is therefore 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 (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. Automat. Lett.*, vol. 2, no. 2, pp. 404–411, Apr. 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, Feb. 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, Feb. 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. Autom. Control*, vol. 53, no. 5, pp. 1203–1218, Jun. 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, Sep. 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. Mechatronics*, vol. 21, no. 4, pp. 1900–1911, Aug. 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, Aug. 2008.

[11] A. Giannitrapani, N. Ceccarelli, F. Scortecci, and A. Garulli, “Comparison of EKF and UKF for spacecraft localization via angle measurements,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 1, pp. 75–84, Jan. 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. Imag. Vis.*, 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. Imag. Vis.*, 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. Autom. Control*, vol. 60, no. 2, pp. 436–449, Feb. 2015.

[15] A. Barrau and S. Bonnabel, “The invariant extended Kalman filter as a stable observer,” *IEEE Trans. Autom. Control*, vol. 62, no. 4, pp. 1797–1812, Apr. 2017.

[16] S. Hauberg, F. Lauze, and K. S. Pedersen, “Unscented Kalman filtering on Riemannian manifolds,” *J. Math. Imag. Vis.*, 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,” *Inf. 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. Mach. Intell.*, vol. 36, no. 4, pp. 625–643, Apr. 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, Dec. 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, Jun. 2014.

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

[25] G. Wahba, “A least–squares estimate of satellite attitude,” *SIAM Rev.*, vol. 7, no. 3, p. 409, 1965. Q3

[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. Basel, 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, Jul. 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, Oct. 2009.

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

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

Since 2000, he has been with the Korea Institute of Science and Technology, Seoul, South Korea, where he is currently a Senior Researcher. His research interests include computer vision and signal processing.

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

His research interests include robot planning and control, and manifold learning from high-dimensional data.

Frank C. Park (F'13) received the B.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1985, and the Ph.D. degree in applied mathematics from Harvard University, Cambridge, MA, USA, in 1991.

After joining the faculty of UC Irvine in 1991, since 1995, he has been a Professor in mechanical and aerospace engineering with Seoul National University, Seoul, South Korea, where he is currently the Department Head. He has been an IEEE Robotics and Automation Society Distinguished Lecturer, and has held Adjunct Faculty positions with the New York University Courant Institute; the Interactive Computing Department, Georgia Tech; and the HKUST Robotics Institute. His research interests include robot mechanics, planning and control, vision and image processing, and related areas of applied mathematics.

Dr. Park is a former Editor-in-Chief for 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."