# Inertial-based Navigation by Polynomial Optimization: Inertial-Magnetic Attitude Estimation 

Maoran Zhu, Student Member, IEEE, Yuanxin Wu, Senior Member, IEEE


#### Abstract

Inertial-based navigation refers to the navigation methods or systems that have inertial information or sensors as the core part and integrate a spectrum of other kinds of sensors for enhanced performance. Through a series of papers, the authors attempt to explore information blending of inertial-based navigation by a polynomial optimization method. The basic idea is to model rigid motions as finite-order polynomials and then attacks the involved navigation problems by optimally solving their coefficients, taking into considerations the constraints posed by inertial sensors and others. In the current paper, a continuoustime attitude estimation approach is proposed, which transforms the attitude estimation into a constant parameter determination problem by the polynomial optimization. Specifically, the continuous attitude is first approximated by a Chebyshev polynomial, of which the unknown Chebyshev coefficients are determined by minimizing the weighted residuals of initial conditions, dynamics and measurements. We apply the derived estimator to the attitude estimation with the magnetic and inertial sensors. Simulation and field tests show that the estimator has much better stability and faster convergence than the traditional extended Kalman filter does, especially in the challenging large initial state error scenarios.


Index Terms-Attitude Estimation, Chebyshev Polynomial, Inertial Sensor, Extended Kalman Filter, Polynomial Optimization

## I. InTRODUCTION

Attitude estimation of a rigid body with low-cost MEMS inertial measurement unit (MIMU) sensors has received numerous applications, including but not limited to robotics [1], aerospace [2], unmanned aerial vehicle [3] and human motion tracking [4, 5]. The MIMU usually integrates the three-axis gyroscope, accelerometer and magnetometer into a module or chip. The magnetometer measures the local magnetic field of the Earth, while the gyroscope and accelerometer measure the angular velocity and specific force of the rigid body with respect to the inertial frame, respectively [6].

The inertial and magnetic attitude estimation has witnessed about forty years' endeavors in trying to optimally fuse gyroscope measurements with the valid accelerometer and magnetometer measurements [7, 8]. Generally, the fusion

[^0]algorithms can be classified into three categories: complementary filter-based methods [9-12], Kalman filterbased methods [7, 13-16] and optimization-based methods [1719].

The complementary filter [20] utilizes the accelerometer and magnetometer vector measurements to determine the attitude and then combines with the time integration of gyroscope by a proper gain. The complementary filter is a straightforward frequency domain approach with low computation burden. Though quite popular in low cost platforms [10], it does not consider the statistical characteristics of measurements.

As for the Kalman filter-based attitude estimation, the most celebrated tool is the multiplicative extended Kalman filter (EKF) [21], which approximates the attitude nonlinearity by the first-order linearization. However, the local linearization at inaccurate estimate will degrade the filter performance, especially in the case of the large initial state error. The invariant extended Kalman filter [22, 23] utilizes the Lie group affine property and derives the error state model independent of the attitude to improve the stability and robustness of the estimation. To address the linearization issue, the unscented Kalman filter (UKF) is also applied for attitude estimation [24], which employs a set of sigma points to approximate the nonlinear probability distribution and propagates the statistical information of posteriori density. The Monte Carlo-based particle filter is also used for attitude estimation [25] by drawing a large number of random samples to approximate and propagate the probability distribution, which leads to a great computation burden.

There are several early attempts to solve the attitude estimation by the optimization-based method. The works [17, 18,26] determine the attitude by first optimizing the objective function of the magnetometer and accelerometer measurements and then incorporating gyroscope measurements through a fusion algorithm. Instead of numerically integrating the gyroscope measurements, a more accurate and direct method is to optimize all the available inertial and magnetic measurements simultaneously [19, 27]. The paper [19] proposes the moving-horizon estimation method, which utilizes all measurements in a time window to solve the spacecraft attitude estimation. However, approximating the attitude

[^1]derivative by the first-order difference therein is not accurate enough, especially for the high dynamic applications. A continuous-time batch attitude estimation using B-splines curve to approximate the attitude is proposed in [27], which is not friendly to real time applications.

Inspired by the collocation method, we quite recently come up with a continuous-time state estimation method by the Chebyshev polynomial optimization for general problems [28]. Through a series of papers, we attempt to explore information blending of inertial-based navigation by the polynomial optimization method. In [29], we have introduced the Chebyshev collocation method [30, 31] into the strapdown inertial navigation computation. This paper will further exploit the idea of the collocation method to derive polynomial optimization-based attitude estimators, which transforms the attitude estimation into a problem of Chebyshev coefficient optimization in a sliding window. Specifically, the attitude profile in the current window is represented by the Chebyshev polynomial and the unknown coefficients are then determined by minimizing the weighted residuals of initial conditions, dynamics and measurements. We apply the derived estimators to the attitude and bias estimation with the magnetic and inertial sensors and find that they have much better stability and faster convergence than the traditional EKF does, especially in the large initial state error scenarios.

The rest of this paper is organized as follows. The inertialmagnetic sensor measurement models are first introduced in Section II. Section III derives the polynomial optimizationbased inertial-magnetic attitude estimation in terms of the quaternion and Rodrigues vector. A linear initial algorithm is also proposed to provide good initial values for the optimization-based methods. The simulations and field tests are conducted in Section IV. Finally, the conclusion is drawn in Section V.

## II. Inertial Sensor Measurement Model

The gyroscope measures the angular velocity of the body with respect to the inertial axes as

$$
\begin{equation*}
\mathbf{y}_{g}=\boldsymbol{\omega}_{i b}^{b}+\mathbf{b}_{g}+\mathbf{n}_{g} \tag{1}
\end{equation*}
$$

where the subscripts $i$ and $b$ denote the inertial frame ( $i$-frame) and body frame ( $b$-frame), respectively. $\mathbf{y}_{g}$ denotes the gyroscope measurement and $\boldsymbol{\omega}_{i b}^{b}$ denotes the true angular velocity in $b$-frame. $\mathbf{n}_{g} \sim N\left(0, \mathbf{R}_{g}\right)$ denotes the gyroscope measurement white noise, and $\mathbf{b}_{g}$ denotes the gyroscope bias that is modeled as

$$
\begin{equation*}
\dot{\mathbf{b}}_{g}=\mathbf{n}_{b g} \tag{2}
\end{equation*}
$$

where $\mathbf{n}_{b g}$ is Gaussian white noise. When the non-gravitational acceleration such as motion disturbance is small, When the nongravitational acceleration such as motion disturbance is small, the accelerometer measurement roughly reflects the opposite direction of local gravity in the $b$-frame, i.e.,

$$
\begin{equation*}
\mathbf{y}_{a} \approx-\mathbf{q}_{n}^{b^{*}} \circ \boldsymbol{\gamma}^{n} \circ \mathbf{q}_{n}^{b}+\mathbf{b}_{a}+\mathbf{n}_{a} \tag{3}
\end{equation*}
$$

where the subscript $n$ represents the navigation frame ( $n$-frame). Without the loss of generality, the navigation frame in this paper takes the definition of North-Up-East. $\mathbf{n}_{a} \sim N\left(0, \mathbf{R}_{a}\right)$ denotes the measurement Gaussian noise and $\mathbf{b}_{a}$ denotes the measurement noise that is modeled as

$$
\begin{equation*}
\dot{\mathbf{b}}_{a}=\mathbf{n}_{b a} \tag{4}
\end{equation*}
$$

where $\mathbf{n}_{b a}$ is a Gaussian white noise. $\boldsymbol{\gamma}^{n}=\left[\begin{array}{lll}0 & -g & 0\end{array}\right]^{T}$ denotes the gravity vector in $n$-frame and $g$ denotes the magnitude of the local gravity. $\mathbf{q}_{n}^{b} \triangleq\left[\begin{array}{ll}s & \boldsymbol{\eta}^{T}\end{array}\right]^{T}$ encodes the unit attitude quaternion of $b$-frame relative to $n$-frame, where $S$ is the scale part and $\boldsymbol{\eta}$ is the vector part of $\mathbf{q}_{n}^{b}$. The operator $\circ$ denotes the quaternion multiplication, defined as

$$
\mathbf{q}_{1} \circ \mathbf{q}_{2}=\left[\stackrel{+}{\mathbf{q}_{1}}\right]\left[\begin{array}{l}
s_{2}  \tag{5}\\
\mathbf{\eta}_{2}
\end{array}\right]=\left[\stackrel{-}{\mathbf{q}_{2}}\right]\left[\begin{array}{l}
s_{1} \\
\mathbf{\eta}_{1}
\end{array}\right]
$$

The two quaternion multiplication matrices, $\left[\begin{array}{c}+ \\ \mathbf{q}\end{array}\right]$ and $[\overline{\mathbf{q}}]$, are respectively defined by

$$
[\stackrel{+}{\mathbf{q}}] \triangleq\left[\begin{array}{cc}
s & -\boldsymbol{\eta}^{T}  \tag{6}\\
\boldsymbol{\eta} & s \mathbf{I}_{3}+\boldsymbol{\eta} \times
\end{array}\right], \quad[\overline{\mathbf{q}}] \triangleq\left[\begin{array}{cc}
s & -\boldsymbol{\eta}^{T} \\
\boldsymbol{\eta} & s \mathbf{I}_{3}-\boldsymbol{\eta} \times
\end{array}\right]
$$

where $\mathbf{I}_{n}$ denotes an $n \times n$ identity matrix. The skew symmetric matrix $(\cdot \times)$ is defined that the cross product $x \times y=(x \times) y$ is satisfied for arbitrary two vectors. The conjugate of the unit quaternion $\mathbf{q}_{n}^{b^{*}}$ in (3) is defined as $\mathbf{q}_{n}^{b^{*}}=\left[\begin{array}{ll}s & -\boldsymbol{\eta}^{T}\end{array}\right]^{T}$.
The magnetometer triad measures the total magnetic flux density in $b$-frame, given as

$$
\begin{equation*}
\mathbf{y}_{m}=\mathbf{q}_{n}^{b^{*}} \circ \mathbf{m}^{n} \circ \mathbf{q}_{n}^{b}+\mathbf{n}_{m} \tag{7}
\end{equation*}
$$

where $\quad \mathbf{y}_{m}$ and $\mathbf{n}_{m} \sim N\left(0, \mathbf{R}_{m}\right)$ denote the normalized magnetometer triad measurements (unit norm) and its noise, respectively. $\mathbf{m}^{n}$ is the normalized magnetic field in $n$-frame, expressed as

$$
\mathbf{m}^{n}=\left[\begin{array}{c}
\cos \alpha_{m} \cos \gamma_{m}  \tag{8}\\
-\sin \gamma_{m} \\
\sin \alpha_{m} \cos \gamma_{m}
\end{array}\right]
$$

where $\alpha_{m}$ and $\gamma_{m}$ are, respectively, the magnetic declination and inclination [6]. It is noted that other sensor parameters, such as the scale factor and misalignment, are assumed having been well calibrated [32, 33]. In this paper, the accelerometer and magnetometer measurements are only used when the external acceleration and magnetic disturbance retain small. Specifically, the norm-based detectors are applied to select the feasible measurements [16]

$$
\begin{align*}
\left\|\left\|\mathbf{y}_{a}\right\|-g\right\| & <\varepsilon_{a}  \tag{9}\\
\|1-\| \mathbf{y}_{m}\| \| & <\varepsilon_{m} \tag{10}
\end{align*}
$$

where the first inequality represents the accelerometer detector and the second inequality represents the magnetometer detector,
$\|\cdot\|$ denotes the norm of a vector, and $\varepsilon_{a}$ and $\varepsilon_{m}$ denote the given thresholds of the two detectors.

Considering the Earth rotation rate, the dynamics of the attitude quaternion satisfies [6]

$$
\begin{equation*}
\dot{\mathbf{q}}_{n}^{b}=\mathbf{q}_{n}^{b} \circ \boldsymbol{\omega}_{n b}^{b} / 2=\left(\mathbf{q}_{n}^{b} \circ \boldsymbol{\omega}_{i b}^{b}-\boldsymbol{\omega}_{i n}^{n} \circ \mathbf{q}_{n}^{b}\right) / 2 \tag{11}
\end{equation*}
$$

where $\boldsymbol{\omega}_{n b}^{b}$ denotes the body angular velocity vector with respect to $n$-frame expressed in $b$-frame. $\boldsymbol{\omega}_{i n}^{n} \approx \boldsymbol{\omega}_{i e}^{n}$ for the lowspeed vehicles and $\boldsymbol{\omega}_{i e}^{n}=\left[\begin{array}{lll}\Omega \cos L & \Omega \sin L & 0\end{array}\right]^{T}$ denotes the Earth's angular velocity vector expressed in $n$-frame, among which $\Omega$ is the norm of the Earth rotation and $L$ is the latitude of the location. Substituting (1) into (11) and multiplying the conjugate of $\mathbf{q}_{n}^{b}$ on both sides, the angular velocity measurements can be represented as a function of quaternion and gyroscope bias, as

$$
\begin{equation*}
\mathbf{y}_{g}=2 \mathbf{q}_{n}^{b^{*}} \circ \dot{\mathbf{q}}_{n}^{b}+\mathbf{q}_{n}^{b^{*}} \circ \boldsymbol{\omega}_{i e}^{n} \circ \mathbf{q}_{n}^{b}+\mathbf{b}_{g}+\mathbf{n}_{g} \tag{12}
\end{equation*}
$$

Another popular parameter to present the attitude is the Rodrigues vector. It has only three components and no inherent constraint, but is singular for a $180^{\circ}$ rotation [21]. Therefore, the Rodrigues vector is usually used to parameterize the attitude update in a short time interval. In this regard, the attitude in the time interval can be written as

$$
\begin{equation*}
\mathbf{q}_{n}^{b}(\tau)=\mathbf{q}_{n, 0}^{b} \circ \Delta \mathbf{q}(\tau) \tag{13}
\end{equation*}
$$

where $\mathbf{q}_{n, 0}^{b}$ denotes the initial attitude quaternion at the start of the current time window and $\Delta \mathbf{q}$ denotes the quaternion update, which is related with the Rodrigues vector update as

$$
\Delta \mathbf{q}=\frac{\left[\begin{array}{cc}
2 & \Delta \mathbf{g}^{T} \tag{14}
\end{array}\right]^{T}}{\sqrt{4+\|\Delta \mathbf{g}\|^{2}}}
$$

reversely,

$$
\begin{equation*}
\Delta \mathbf{g}=\frac{2 \Delta \boldsymbol{\eta}}{\Delta s} \tag{15}
\end{equation*}
$$

where $\Delta s$ is the scale part and $\Delta \boldsymbol{\eta}$ is the vector part of $\Delta \mathbf{q}$. Substituting (13) and (14) into (12), the angular velocity measurements can also be represented by the Rodrigues vector update

$$
\begin{equation*}
\mathbf{y}_{g}=\frac{4 \Delta \dot{\mathbf{g}}-2 \Delta \mathbf{g} \times \Delta \dot{\mathbf{g}}}{4+\|\Delta \mathbf{g}\|^{2}}+\Delta \mathbf{C}(\Delta \mathbf{g}) \mathbf{C}_{n, 0}^{b} \boldsymbol{\omega}_{i e}^{n}+\mathbf{b}_{g}+\mathbf{n}_{g} \tag{16}
\end{equation*}
$$

where $\mathbf{C}_{n, 0}^{b}$ is the initial rotation matrix corresponding to the initial attitude quaternion $\mathbf{q}_{n, 0}^{b}$ and $\Delta \mathbf{C}(\Delta \mathbf{g})$ means a function that transforms the Rodrigues vector to the corresponding rotation matrix.

## III. Inertial-Magnetic Attitude Estimation by Chebyshev Polynomial Optimization (AttestPO)

## A. Formulation of Attitude Estimation

Without the loss of the generality, consider the attitude estimation on the time interval $\left[\begin{array}{ll}t_{0} & t_{M}\end{array}\right]$, in which $M+1$
inertial-magnetic samples are available at $t_{k}, k=0, \cdots M$. Because the gyroscope and accelerometer biases are slowly changing, it is reasonable to treat them as constants in a short time window. Therefore, the state to be estimated in the time interval includes the continuous attitude and the constant gyroscope and accelerometer biases, which are denoted as $\mathbf{x}(t) \triangleq\left[\mathbf{q}_{n}^{b}(t)^{T}, \mathbf{b}_{a}^{T}, \mathbf{b}_{g}^{T}\right]^{T}$. Assume the initial state is given as $\mathbf{x}_{0}=\left[\mathbf{q}_{n, 0}^{b}{ }^{T}, \mathbf{b}_{a_{0}}^{T}, \mathbf{b}_{g_{0}}^{T}\right]^{T}$, where $\mathbf{q}_{n, 0}^{b}, \mathbf{b}_{a_{0}}$ and $\mathbf{b}_{g_{0}}$ denote the quaternion, accerlerometer and gyroscope biases at the initial time $t_{0}$, respectively.

In the least squares sense [34], an optimal continuousdiscrete attitude estimation is to minimize the measurement residuals with respect to $\mathbf{x}(t)$

$$
\begin{equation*}
\min _{\mathbf{x}(t)}\left(J_{\mathbf{x}_{0}}+J_{\mathbf{v}}+J_{\mathbf{z}}\right) \quad \text { s.t }\|\mathbf{q}(t)\|=1 \tag{17}
\end{equation*}
$$

where the objective functions, $J_{\mathrm{x}_{0}}, J_{\mathrm{v}}$ and $J_{\mathrm{z}}$, respectively denote the prior, dynamics and measurement terms. They are explicitly given as follows

$$
\begin{align*}
& J_{\mathbf{x}_{0}}=\mathbf{e}_{\mathbf{x}_{0}}^{T} \mathbf{e}_{\mathbf{x}_{0}} \\
& J_{\mathbf{v}}=\int_{t_{0}}^{t_{M}} \mathbf{e}_{\mathbf{v}}^{T}(\tau) \mathbf{e}_{\mathbf{v}}(\tau) d \tau  \tag{18}\\
& J_{\mathbf{z}}=\sum_{k=1}^{p} \mathbf{e}_{a_{k}}^{T} \mathbf{e}_{a_{k}}+\sum_{k=1}^{s} \mathbf{e}_{m_{k}}^{T} \mathbf{e}_{m_{k}}
\end{align*}
$$

where $p$ and $s$ respectively denote the number of valid accelerometer and magnetometer measurements in the current time window that are selected by the detectors in (9) and (10). With the measurement models (3), (7) and (12), the weighted residuals $\mathbf{e}_{\mathbf{x}_{0}}, \mathbf{e}_{\mathbf{v}}, \mathbf{e}_{a}$ and $\mathbf{e}_{m}$ are given as

$$
\begin{align*}
& \mathbf{e}_{\mathbf{x}_{0}}=\mathbf{W}_{\mathbf{x}_{0}}^{T}\left[\delta \boldsymbol{\Psi}_{0}^{T} \mathbf{b}_{a}^{T}-\mathbf{b}_{a_{0}}^{T} \mathbf{b}_{g}^{T}-\mathbf{b}_{g_{0}}^{T}\right] \\
& \mathbf{e}_{\mathbf{v}}=\mathbf{W}_{g}^{T}\left(\mathbf{y}_{g}-2 \mathbf{q}_{n}^{b^{*}} \circ \dot{\mathbf{q}}_{n}^{b}-\mathbf{q}_{n}^{b^{*}} \circ \boldsymbol{\omega}_{i e}^{n} \circ \mathbf{q}_{n}^{b}-\mathbf{b}_{g}\right) \\
& \mathbf{e}_{a}=\mathbf{W}_{a}^{T}\left(\mathbf{y}_{a}+\mathbf{q}_{n}^{b^{*}} \circ \mathbf{g}^{n} \circ \mathbf{q}_{n}^{b}-\mathbf{b}_{a}\right)  \tag{19}\\
& \mathbf{e}_{m}=\mathbf{W}_{m}^{T}\left(\mathbf{y}_{m}-\mathbf{q}_{n}^{b^{*}} \circ \mathbf{m}^{n} \circ \mathbf{q}_{n}^{b}\right)
\end{align*}
$$

which are related to the prior, attitude dynamics and valid accelerometer and magnetometer measurements, respectively, through their corresponding weight matrices. Specifically, regarding the residual of initial state, $\delta \psi_{0}$ is the initial threedimensional attitude error that is roughly related to the attitude quaternion as [6]

$$
\begin{equation*}
\delta \boldsymbol{\psi} \approx 2\left[\mathbf{q}_{n}^{b} \circ \hat{\mathbf{q}}_{n}^{b^{*}}\right]_{2: 4} \tag{20}
\end{equation*}
$$

where $\mathbf{q}_{n}^{b}$ and $\hat{\mathbf{q}}_{n}^{b}$ denote the true and error-contaminated quaternions, respectively. The operator $[\cdot]_{2: 4}$ extracts the second to the fourth rows of a matrix. The weight matrices are obtained from the Cholesky factorization of the inverse covariance matrix, namely, $\mathbf{P}_{\mathbf{x}_{0}}^{-1}=\mathbf{W}_{\mathbf{x}_{0}} \mathbf{W}_{\mathbf{x}_{0}}^{T}, \mathbf{R}_{g}^{-1}=\mathbf{W}_{g} \mathbf{W}_{g}^{T}$, $\mathbf{R}_{a}^{-1}=\mathbf{W}_{a} \mathbf{W}_{a}^{T}, \mathbf{R}_{m}^{-1}=\mathbf{W}_{m} \mathbf{W}_{m}^{T}$. It should be noted that the covariance matrix of the initial state is defined as
$\mathbf{P}_{\mathbf{x}_{0}}^{-1} \triangleq \operatorname{diag}\left[\mathbf{P}_{\boldsymbol{\psi}_{0}}, \mathbf{P}_{b_{a}, 0}, \mathbf{P}_{b_{g}, 0}\right]^{-1}$, where $\mathbf{P}_{\psi_{0}}, \mathbf{P}_{b_{a}, 0}$ and $\mathbf{P}_{b_{g}, 0}$ are the initial covariances of the attitude, accelerometer and gyroscope biases, respectively.

## B. Attitude Estimation by Chebyshev Polynomial Optimization

To solve the infinite-dimension optimization problem in (17), the Chebyshev collocation method is introduced to represent the attitude by a finite-order Chebyshev polynomial and transform the continuous-time attitude estimation into a constant parameter optimization problem.

The Chebyshev polynomial of the first kind is defined over the internal $\left[\begin{array}{ll}-1 & 1\end{array}\right]$ by the recurrence relation as [30]

$$
\begin{align*}
& F_{0}(\tau)=1, F_{1}(\tau)=\tau \\
& F_{i+1}(\tau)=2 \tau F_{i}(\tau)-F_{i-1}(\tau) \text { for } i \geq 1 \tag{21}
\end{align*}
$$

where $F_{i}(\tau)$ is the $i^{\text {th }}$-degree Chebyshev polynomial. In order to apply the Chebyshev polynomial, the optimization in (17) on $t \in\left[\begin{array}{ll}t_{0} & t_{M}\end{array}\right]$ needs to be mapped into $\tau \in\left[\begin{array}{ll}-1 & 1\end{array}\right]$ by the affine transformation as follows

$$
\begin{equation*}
\tau=\frac{2}{t_{M}-t_{0}} t-\frac{t_{M}+t_{0}}{t_{M}-t_{0}} \tag{22}
\end{equation*}
$$

The attitude quaternion in the time interval is then approximated by a Chebyshev polynomial up to order $N_{q}$ as

$$
\begin{equation*}
\mathbf{q}_{n}^{b}(\tau) \approx \sum_{i=0}^{N_{q}} \mathbf{d}_{i} F_{i}(\tau) \triangleq \mathbf{D F}(\tau) \tag{23}
\end{equation*}
$$

where $\mathbf{d}_{i}$ denote the $i^{\text {th }}$-degree Chebyshev coefficient. The matrices $\quad \mathbf{F}(\tau) \triangleq\left[\begin{array}{llll}F_{0}(\tau) & F_{1}(\tau) & \cdots & F_{N_{q}}(\tau)\end{array}\right]^{T}$ and $\mathbf{D} \triangleq\left[\mathbf{d}_{0}, \cdots, \mathbf{d}_{N_{q}}\right]$ are defined for compact denotation.

Then, the derivative of the attitude quaternion is given by

$$
\begin{equation*}
\dot{\mathbf{q}}_{n}^{b}(\tau) \approx \sum_{i=0}^{N_{q}} \mathbf{d}_{i} \dot{F}_{i}(\tau) \triangleq \mathbf{D} \dot{\mathbf{F}}(\tau) \tag{24}
\end{equation*}
$$

where $\dot{\mathbf{F}}(\tau) \triangleq\left[\begin{array}{llll}\dot{F}_{0}(\tau) & \dot{F}_{1}(\tau) & \cdots & \dot{F}_{N_{q}}(\tau)\end{array}\right]^{T}$ and $\dot{F}_{i}$ is the time derivative of $F_{i}$ that is directly obtained by taking the derivative of $(21)$. That is to say

$$
\begin{align*}
& \dot{F}_{0}(\tau)=0, \dot{F}_{1}(\tau)=1 \\
& \dot{F}_{i+1}(\tau)=2 F_{i}(\tau)+2 \tau \dot{F}_{i}(\tau)-\dot{F}_{i-1}(\tau) \text { for } i \geq 1 \tag{25}
\end{align*}
$$

The reason that the Chebyshev polynomial is selected as the basis function is for its high accuracy and efficacy in functional approximation. As a matter of fact, the Chebyshev polynomial is very close to the best polynomial approximation in the $\infty-$ norm [30]. The integral term $J_{v}$ in (18) can be numerically solved by the Clenshaw-Curtis quadrature formula as [30]

$$
\begin{equation*}
J_{\mathbf{v}} \approx \sum_{i=0}^{N} w_{i} \mathbf{e}_{\mathbf{v}}^{T}\left(\tau_{i}\right) \mathbf{e}_{\mathbf{v}}\left(\tau_{i}\right) \tag{26}
\end{equation*}
$$

where $\tau_{i}$ denotes the Chebyshev points

$$
\begin{equation*}
\tau_{i}=-\cos (i \pi / N), \quad i=0,1, \cdots, N \tag{27}
\end{equation*}
$$

and $w_{i}$ denotes the weight that is determined by the integrals
of the Lagrange polynomials [35] and $N+1$ is the number of Chebyshev points. We see that (26) requires the angular velocity at the Chebyshev points. In this paper, it is achieved by the way of the extended Floater and Hormann (EFH) interpolation method [36], which reconstructs the angular velocity from the equally-spaced time sampled angular velocity measurements [37]. Readers are referred to [29] for the details of the angular velocity reconstruction.

Following the commonly used strategy to handle the constraints in the collocation-based optimal control [38], the continuous unit quaternion constraint in (17) is discretized at the Chebyshev points as

$$
\begin{equation*}
\left\|\mathbf{q}_{n}^{b}\left(\tau_{k}\right)\right\|=1 \quad k=0,1, \cdots, N \tag{28}
\end{equation*}
$$

Substituting the Chebyshev approximations (23), (24), (26) and the discretized constraints (28) into (17), we can reformulate the estimation task for the time window of interest as

$$
\begin{align*}
& \min _{\left\{\mathbf{p}, \mathbf{b}_{a}, \mathbf{b}_{g}\right\}}\left(J_{\mathbf{x}_{0}}+J_{\mathbf{v}}+J_{\mathbf{z}}\right)  \tag{29}\\
& \text { s.t. }\left\|\mathbf{q}_{n}^{b}\left(\tau_{k}\right)\right\|=1 \quad k=0,1, \cdots, N
\end{align*}
$$

where the estimation parameters include the Chebyshev coefficient and the accelerometer/gyroscope biases. Equation (29) is a constrained nonlinear squares problem, which can be transformed to an unconstrained nonlinear least squares by the augmented Lagrangian method and then solved by the Levenberg-Marquardt algorithm [39]. Once the coefficients are determined, the continuous attitude quaternion as a function of time will be finally acquired by (23).

It should be noted that (29) is a constrained optimization that is generally more time-consuming than an unconstrained optimization. On the other hand, the above optimization problem can be reformulated as an unconstrained one using the Rodrigues vector. To avoid its singularity at a $180^{\circ}$ rotation, however, we consider the Rodrigues vector update in the shorttime window. Specifically, the Rodrigues vector update is approximated by a Chebyshev polynomial up to order $N_{g}$ as

$$
\begin{equation*}
\Delta \mathbf{g}(\tau) \approx \sum_{i=0}^{N_{g}} \mathbf{h}_{i} F_{i}(\tau) \triangleq \mathbf{H F}(\tau) \tag{30}
\end{equation*}
$$

where $\mathbf{h}_{i}$ denote the $i^{\text {th }}$-degree Chebyshev coefficient, and $\mathbf{H} \triangleq\left[\mathbf{h}_{0}, \cdots, \mathbf{h}_{N_{g}}\right]$ denotes the combined matrix of Chebyshev coefficients. Then, the derivative of the attitude quaternion is given by

$$
\begin{equation*}
\Delta \dot{\mathbf{g}}(\tau) \approx \sum_{i=0}^{N_{g}} \mathbf{h}_{i} \dot{F}_{i}(\tau) \triangleq \mathbf{H} \dot{\mathbf{F}}(\tau) \tag{31}
\end{equation*}
$$

By analogy with the quaternion estimation, the Rodrigues vector update estimation can be formulated as

$$
\begin{equation*}
\min _{\left\{\mathbf{H}, \mathbf{b}_{a}, \mathbf{b}_{\mathbf{z}}\right\}}\left(J_{\mathbf{x}_{0}}+J_{\mathbf{v}}+J_{\mathbf{z}}\right) \tag{32}
\end{equation*}
$$

where the estimation parameters include the Chebyshev coefficient and the accelerometer/gyroscope biases. The weighted residuals of initial state and accelerometer/magnetometer measurements are obtained by


Fig. 1. Estimated state and its covariance propagation (the state and covariance at the end of Window $k$ are served as the initial condition for Window $k+1$ ).
substituting (13) and (14) into (19), and the weighted residual of attitude dynamics is obtained from (16), given as

$$
\begin{equation*}
\mathbf{e}_{\mathbf{v}}=\mathbf{W}_{g}^{T}\left(\mathbf{y}_{g}-\frac{4 \Delta \dot{\mathbf{g}}-2 \Delta \mathbf{g} \times \Delta \dot{\mathbf{g}}}{4+\|\Delta \mathbf{g}\|^{2}}-\Delta \mathbf{C}(\Delta \mathbf{g}) \mathbf{C}_{n, 0}^{b} \boldsymbol{\omega}_{i e}^{n}-\mathbf{b}_{g}\right) \tag{33}
\end{equation*}
$$

With the optimized Chebyshev coefficients, the attitude quaternion in the window is readily obtained by substituting (30) into (13) and (14).

In summary, the approximations involved in the proposed batch estimators include the attitude profile approximation by the Chebyshev polynomial in (23) or (30), the Clenshaw-Curtis quadrature in (26) and the angular velocity interpolant by the EFH strategy. With the increased order of Chebyshev polynomial, the errors incurred by Chebyshev approximation and Clenshaw-Curtis quadrature will decay towards zero [30]. Additionally, the EFH interpolant is also likely very close to optimality for time-equispaced samples [40]. In this regard, the proposed attitude estimator by polynomial optimization (AttEstPO) is nearly optimal in the least squares sense.

To meet the requirement of real-time applications, the above optimizations are expected to be solved recursively on consecutive short-time windows. The estimation result at the end of the current window can serve as the prior for the next window, but the required covariance estimation is not directly available. In view of the strategy in the moving horizon estimation [41], the multiplicative EKF, with the linearized dynamics and measurements at the AttEstPO estimate, can be resorted to compute the state covariance. The multiplicative EKF, also known as the error-state EKF, estimates the error state (instead of the original state) to avoid the overparameterization issue of the attitude [6]. The error state is defined as the estimate subtracting the truth, i.e., $\delta \mathbf{x}=\hat{\mathbf{x}}-\mathbf{x}$, except that the definition of attitude error follows (20). The linearized error-state dynamics is expressed as

$$
\begin{equation*}
\delta \dot{\mathbf{x}}=\mathbf{B} \delta \mathbf{x}+\mathbf{G w} \tag{34}
\end{equation*}
$$

where the error state is a 9-dimensional vector $\delta \mathbf{x} \triangleq\left[\begin{array}{lll}\delta \boldsymbol{\psi} & \delta \mathbf{b}_{a} & \delta \mathbf{b}_{g}\end{array}\right]^{T}$, the dynamic noise is $\mathbf{w} \triangleq\left[\begin{array}{lll}\mathbf{n}_{g} & \mathbf{n}_{b a} & \mathbf{n}_{b g}\end{array}\right]^{T}$ and the matrices are

$$
\mathbf{B}=\left[\begin{array}{ccc}
\mathbf{0}_{3} & \mathbf{0}_{3} & -\mathbf{C}_{b}^{n}  \tag{35}\\
\mathbf{0}_{3} & \mathbf{0}_{3} & \mathbf{0}_{3}
\end{array}\right] \quad \mathbf{G}=\left[\begin{array}{ccc}
-\mathbf{C}_{b}^{n} & \mathbf{0}_{3} & \mathbf{0}_{3} \\
\mathbf{0}_{3} & \mathbf{I}_{3} & \mathbf{0}_{3} \\
\mathbf{0}_{3} & \mathbf{0}_{3} & \mathbf{I}_{3}
\end{array}\right]
$$

where $\mathbf{0}_{n}$ denotes an $n \times n$ zero matrix. The multiplicative EKF covariance prediction from $t_{k-1}$ to $t_{k}$ is given as

$$
\begin{equation*}
\mathbf{P}_{k}^{-}=\left(\mathbf{I}_{9}+\mathbf{B}_{k}\right) \mathbf{P}_{k-1}\left(\mathbf{I}_{9}+\mathbf{B}_{k}\right)^{T}+\mathbf{G}_{k} \mathbf{Q}_{k} \mathbf{G}_{k}^{T} T \tag{36}
\end{equation*}
$$

where the covariance update interval is set to $T \triangleq t_{k}-t_{k-1}$. When the measurement at time $t_{k}$ comes, the covariance is updated by

$$
\begin{equation*}
\mathbf{P}_{k}=\mathbf{P}_{k}^{-}-\mathbf{P}_{k}^{-} \mathbf{H}_{k}^{T}\left(\mathbf{H}_{k} \mathbf{P}_{k}^{-} \mathbf{H}_{k}^{T}+\mathbf{R}_{k}\right)^{-1} \mathbf{H}_{k} \mathbf{P}_{k}^{-} \tag{37}
\end{equation*}
$$

where $\mathbf{R}_{k}$ and $\mathbf{H}_{k}$ denote the measurement noise covariance and linearized measurement matrices, respectively. For the accelerometer and magnetometer, the matrix $\mathbf{H}_{k}$ is respectively written as

$$
\begin{align*}
\mathbf{H}_{k, m} & =\left[\begin{array}{lll}
\mathbf{C}_{n}^{b} \mathbf{m}^{n} \times & \mathbf{0}_{3} & \mathbf{0}_{3}
\end{array}\right] \\
\mathbf{H}_{k, a} & =\left[\begin{array}{lll}
-\mathbf{C}_{n}^{b} \boldsymbol{\gamma}^{n} \times & \mathbf{I}_{3} & \mathbf{0}_{3}
\end{array}\right] \tag{38}
\end{align*}
$$

Note that the matrices $\mathbf{G}_{k}, \mathbf{H}_{k}$ and $\mathbf{B}_{k}$ in (36) and (37) are calculated using the AttEstPO estimate. In contrast, the covariance propagation of EKF is performed at the current state estimate, which is usually inferior to that by AttEstPO because the latter is a kind of local smoothing using all information during the current time window and EKF only uses information up to the current time. Taken as an example, Fig. 1 illustrates the state and covariance propagation between two adjacent windows, where the state estimate and its covariance at the end time of window $k$ serve as the initial condition for window $k+1$.

From the theoretical perspective, the proposed AttEstPO algorithm only makes a Gaussian assumption of the prior and does not introduce any approximations of nonlinear dynamics and measurements in the time interval of interest. This feature is a significant advantage over the well-known linearizationbased EKF that approximates the nonlinearity by successive Taylor expansion at current estimate. In this regard, the AttEstPO is promising to obtain a better estimation at the price of time delay, as shown in Section IV.

## C. Chebyshev Coefficient Initialization

A fine initialization of Chebyshev coefficients is required for the above nonlinear optimization. In this subsection, we present a linear algorithm to acquire the initial attitude Chebyshev coefficients by assuming approximately known biases and correct initial attitude. Multiplying $\mathbf{q}_{n}^{b}$ on both sides of the attitude dynamics (12) and the accelerometer/magnetometer measurement in (3) and (7), ignoring the measurement noises and using the quaternion multiplication property in (5), we have

$$
\begin{align*}
& \boldsymbol{\rho}_{a}(t) \mathbf{q}_{n}^{b}(t)=\mathbf{0}_{4 \times 1} \\
& \boldsymbol{\rho}_{m}(t) \mathbf{q}_{n}^{b}(t)=\mathbf{0}_{4 \times 1}  \tag{39}\\
& \boldsymbol{\rho}_{g}(t) \mathbf{q}_{n}^{b}(t)-2 \dot{\mathbf{q}}_{n}^{b}(t)=\mathbf{0}_{4 \times 1}
\end{align*}
$$

where the matrices are readily defined as

$$
\begin{align*}
& \boldsymbol{\rho}_{a}(t) \triangleq\left[\left(\mathbf{y}_{a}(t)-\mathbf{b}_{a}\right)\right]+\left[\boldsymbol{\gamma}^{+}\right] \\
& \boldsymbol{\rho}_{m}(t) \triangleq\left[\mathbf{y}_{m}(t)\right]-\left[\mathbf{m}^{+}\right]  \tag{40}\\
& \boldsymbol{\rho}_{g}(t) \triangleq\left[\left(\mathbf{y}_{g}(t)-\mathbf{b}_{g}\right)\right]-\left[\boldsymbol{\omega}_{i e}^{n}(t)\right]
\end{align*}
$$

TABLE I
Inertial-Magnetic Attitude estimation by Polynomial Optimization

|  | Qua-AttEstPO Rod-AttEstPO |
| :---: | :---: |
| Input: | Chebyshev order $N_{q}$ or $N_{g}$, initial state $\mathbf{x}_{0}$ with associated covariance $\mathbf{P}_{0}$, inertial and magnetic measurements $\mathbf{y}_{g}\left(t_{i}\right), \mathbf{y}_{a}\left(t_{i}\right), \mathbf{y}_{m}\left(t_{i}\right) i=0,1, \ldots, M$ and the measurement noise covariances $\mathbf{R}_{g}, \mathbf{R}_{a}, \mathbf{R}_{m}$ |
| Step 1: | Interpolate angular velocity measurement by EFH and obtain angular velocity at Chebyshev points $\mathbf{y}_{g}\left(\tau_{i}\right) \quad i=0,1, \cdots, N$ |
| Step 2: | Compute Chebyshev polynomial coefficients of quaternion $\mathbf{q}_{n}^{b}$ by homogenous least squares (Eq. (45)) $\min _{\mathbf{D}}\\|\mathbf{A} \operatorname{vec}(\mathbf{D})\\| \text { s.t }\\|\mathbf{B} \operatorname{vec}(\mathbf{D})\\|=1$ |
| Step 3: | Acquire initial coefficients of the Rodrigues vector $\Delta \mathbf{g}$ using quaternion Chebyshev coefficients (Eq. (47)) <br> Obtain quaternion Chebyshev coefficients and sensor biases by solving constrained nonlinear least squares (Eq. (29)) $\mathbf{h}_{i} \approx \frac{2-\delta_{0 i}}{P_{N_{g}}} \sum_{k=0}^{P_{N_{N_{g}}-1}} \Delta \mathbf{g}\left(\Delta \mathbf{q}\left(\cos \frac{(k+1 / 2) \pi}{P_{N_{g}}}\right)\right) \cos \frac{i(k+1 / 2) \pi}{P_{N_{g}}}$ $\begin{aligned} & \min _{\left\{\mathbf{D}, \mathbf{b}_{a}, \mathbf{b}_{g}\right\}}\left(J_{\mathbf{x}_{0}}+J_{\mathbf{v}}+J_{\mathbf{z}}\right) \\ & \text { s.t. }\left\\|\mathbf{q}_{n}^{b}\left(\tau_{k}\right)\right\\|=1 \quad k=0,1, \cdots, N \end{aligned}$ <br> Compute Rodrigues vector Chebyshev polynomial coefficients and sensor biases by nonlinear least squares (Eq. (32)) $\min _{\left\{\mathbf{H}, \mathbf{b}_{a}, \mathbf{b}_{g}\right\}}\left(J_{\mathbf{x}_{0}}+J_{\mathbf{v}}+J_{\mathbf{z}}\right)$ |
| Step 4: | Obtain attitude quaternion (Eq. (23)) <br> Obtain Rodrigues vector update and transform it to quaternion $\mathbf{q}_{n}^{b}(\tau)=\sum_{i=0}^{N_{q}} \mathbf{d}_{i} F_{i}(\tau)$ <br> (Eqs. (30), (13) and (14)) $\Delta \mathbf{g}(\tau)=\sum_{i=0}^{N_{s}} \mathbf{h}_{i} F_{i}(\tau) \quad \mathbf{q}_{n}^{b}(\tau)=\mathbf{q}_{n, 0}^{b} \circ \Delta \mathbf{q}(\Delta \mathbf{g}(\tau))$ |

Compute estimate covariance by multiplicative EKF using AttEstPO estimated attitude (Eqs. (36)-(37))
Step 5:

$$
\mathbf{P}_{k}^{-}=\left(\mathbf{I}_{9}+\mathbf{B}_{k}\right) \mathbf{P}_{k-1}\left(\mathbf{I}_{9}+\mathbf{B}_{k}\right)^{T}+\mathbf{G}_{k} \mathbf{Q}_{k} \mathbf{G}_{k}^{T} T \quad \mathbf{P}_{k}=\mathbf{P}_{k}^{-}-\mathbf{P}_{k}^{-} \mathbf{H}_{k}^{T}\left(\mathbf{H}_{k} \mathbf{P}_{k}^{-} \mathbf{H}_{k}^{T}+\mathbf{R}_{k}\right)^{-1} \mathbf{H}_{k} \mathbf{P}_{k}^{-}
$$

Substituting the quaternion Chebyshev polynomial (23) into (39) yields

$$
\begin{align*}
& \mathbf{a}_{a}(t) \operatorname{vec}(\mathbf{D})=\mathbf{0}_{4 \times 1} \\
& \mathbf{a}_{m}(t) \operatorname{vec}(\mathbf{D})=\mathbf{0}_{4 \times 1}  \tag{41}\\
& \mathbf{a}_{g}(t) \operatorname{vec}(\mathbf{D})=\mathbf{0}_{4 \times 1}
\end{align*}
$$

where $\operatorname{vec}(\mathbf{D})$ is formed by stacking the columns of $\mathbf{D}$ and the involved matrices are defined as

$$
\begin{align*}
& \mathbf{a}_{a}(t) \triangleq\left[\mathbf{F}^{T}(t) \otimes \boldsymbol{\rho}_{a}(t)\right] \\
& \mathbf{a}_{m}(t) \triangleq\left[\mathbf{F}^{T}(t) \otimes \boldsymbol{\rho}_{m}(t)\right]  \tag{42}\\
& \mathbf{a}_{g}(t) \triangleq\left[\mathbf{F}^{T}(t) \otimes \boldsymbol{\rho}_{g}(t)-2 \dot{\mathbf{F}}^{T}(t) \otimes \mathbf{I}_{4}\right]
\end{align*}
$$

where $\otimes$ denotes the Kronecker product. Note that the identity property $\operatorname{vec}\left(\mathbf{A}_{1} \mathbf{A}_{2} \mathbf{A}_{3}\right)=\left(\mathbf{A}_{3}{ }^{T} \otimes \mathbf{A}_{1}\right) \operatorname{vec}\left(\mathbf{A}_{2}\right)$ has been used in deriving (41), where $\mathbf{A}_{1}, \mathbf{A}_{2}, \mathbf{A}_{3}$ are matrices of appropriate dimensions. Using the quaternion multiplication property and the definition of the attitude error in (20), a correct initial attitude means

$$
\delta \Psi_{0} \approx 2\left[\mathbf{q}_{n}^{b}\left(\tau_{0}\right) \otimes \mathbf{q}_{n, 0}^{b *}\right]_{2: 4}=2\left[\left[\mathbf{q}_{n, 0}^{-b}\right] \mathbf{q}_{n}^{b}\left(\tau_{0}\right)\right]_{2: 4}=\mathbf{0}_{3 \times 1}
$$

(43)

Approximating $\mathbf{q}_{n}^{b}$ by Chebyshev polynomial, (43) is written as

$$
\begin{equation*}
\mathbf{A}_{i n i t} \operatorname{vec}(\mathbf{D})=\mathbf{0}_{3 \times 1} \tag{44}
\end{equation*}
$$

where $\quad \mathbf{A}_{\text {init }} \triangleq 2\left[\mathbf{F}^{T}\left(\tau_{0}\right) \otimes\left[\mathbf{q}_{n, 0}^{b}{ }^{*}\right]^{-}\right]_{2: 4}$. Assume the initial biases are roughly known, the coefficient $\mathbf{D}$ can be obtained by using all available measurements and the initial attitude in the current time window as

$$
\begin{equation*}
\min _{\mathbf{D}}\|\mathbf{A} \operatorname{vec}(\mathbf{D})\| \text { s.t. }\|\mathbf{B} \operatorname{vec}(\mathbf{D})\|=1 \tag{45}
\end{equation*}
$$

where the constraint term denotes the unit norm of the initial quaternion, which helps ensure a non-zero solution. The matrices $\mathbf{A}$ and $\mathbf{B}$ are known matrices as follows

$$
\begin{align*}
& \mathbf{A}=\left[\begin{array}{llll}
\mathbf{A}_{\text {init }}{ }^{T} & \mathbf{A}_{g}{ }^{T} & \mathbf{A}_{a}{ }^{T} & \mathbf{A}_{m}{ }^{T}
\end{array}\right]^{T}  \tag{46}\\
& \mathbf{B}=\mathbf{F}^{T}\left(\tau_{0}\right) \otimes \otimes \mathbf{I}_{4}
\end{align*}
$$



Fig. 2. Euler angle errors of Qua-AttEstPO, Rod-AttEstPO and EKF across 100 Monte-Caro runs.


Fig. 3. Gyroscope bias errors of Qua-AttEstPO, Rod-AttEstPO and EKF across 100 Monte-Caro runs.
where $\mathbf{A}_{g} \triangleq\left[\begin{array}{lll}\mathbf{a}_{g}^{T}\left(\tau_{0}\right) & \ldots & \mathbf{a}_{g}^{T}\left(\tau_{N}\right)\end{array}\right]^{T}$ is the matrix formed by the fitted angular velocity at Chebyshev points,

$$
\mathbf{A}_{a} \triangleq\left[\begin{array}{lll}
\mathbf{a}_{a}^{T}\left(t_{j}\right) & \mathbf{a}_{a}^{T}\left(t_{k}\right) & \cdots
\end{array}\right]^{T} \quad \text { and } \quad \mathbf{A}_{m} \triangleq\left[\begin{array}{lll}
\mathbf{a}_{m}^{T}\left(t_{l}\right) & \mathbf{a}_{m}^{T}\left(t_{p}\right) & \cdots
\end{array}\right]^{T}
$$

respectively denote the matrices from accelerometer and magnetometer measurements and the timestamp in $\mathbf{a}_{m}$ and $\mathbf{a}_{a}$ denotes all the feasible measurement times that are detected by (9) and (10). The minimization in (45) is a homogenous least squares problem of the quaternion Chebyshev coefficients, which can be solved by the generalized singular value decomposition [42].

As for the attitude estimation problem in the form of the Rodrigues vector in (32), the initial coefficients can be readily obtained by transforming the obtained quaternion coefficients in (45) to the corresponding Rodrigues vector update coefficients as


Fig. 4. Accelerometer bias errors of Qua-AttEstPO, Rod-AttEstPO and EKF across 100 Monte-Caro runs.


Fig. 5. Average attitude errors across 100 Monte-Caro runs for Qua-AttEstPO, Rod-AttEstPO and EKF with the associated 2-standard deviation bound.

$$
\mathbf{h}_{i} \approx \frac{2-\delta_{0 i}}{P_{N_{g}}} \sum_{k=0}^{P_{N_{g}}-1} \Delta \mathbf{g}\left(\Delta \mathbf{q}\left(\cos \frac{(k+1 / 2) \pi}{P_{N_{g}}}\right)\right) \cos \frac{i(k+1 / 2) \pi}{P_{N_{g}}}
$$

(47)
where $\delta_{0 i}$ is the Kronecker delta function, yielding 1 for $i=1$ and zero otherwise. The exact coefficients could be obtained only if the number of summation terms, $P_{N_{g}}$, approaches infinity [35]. $\Delta \mathbf{g}(\Delta \mathbf{q}(\tau))$ transforms the quaternion update to the Rodrigues vector update as shown in (15), among which $\Delta \mathbf{q}(\tau)$ is obtained from (13) using the resultant $\mathbf{q}_{n}^{b}(\tau)$ by (45).

Hereafter, we name the proposed attitude quaternion and Rodrigues vector estimation by polynomial optimization as Qua-AttEstPO and Rod-AttEstPO, respectively. Table I lists their main steps. The algorithm input includes the user-defined time-window size, the Chebyshev polynomial order for attitude approximation, the feasible sensor measurements in the current window and the initial state with the associated covariance. After the estimation is completed, the state and covariance at


Fig. 6. Average gyroscope and accelerometer bias errors across 100 MonteCaro runs for Qua-AttEstPO, Rod-AttEstPO and EKF with the associated 2standard deviation bounds.


Fig. 7. Average attitude errors across 100 Monte-Caro runs for Qua-AttEstPO with different window sizes.
the end of current time window are exploited as the initial state and covariance for the next time window.

## IV. Simulations and Experiments

## A. Simulation Results

In this section, simulations are conducted to evaluate the proposed Qua-AttEstPO and Rod-AttEstPO algorithms, against the popular multiplicative EKF. The estimate state in EKF includes the attitude, gyroscope and accelerometer biases. Interested readers are referred to [7] for more details about the EKF implementation.

Assume a 9-axis MIMU rotating at a fix position (longitude: 112 deg, latitude: 28 deg height: 0 m ) under the classic coning motion for 20 seconds. The attitude time trajectory is described by the quaternion as $\mathbf{q}_{n}^{b}=\cos (\alpha / 2)+\sin (\alpha / 2)\left[\begin{array}{lll}0 & \cos (\varsigma t) & \sin (\varsigma t)\end{array}\right]^{T}$, where the coning frequency $\varsigma=0.74 \pi \mathrm{rad} / \mathrm{s}$ and the coning angle


Fig. 8. Average gyroscope and accelerometer bias errors across 100 MonteCaro runs for Qua-AttEstPO with different window sizes.


Fig. 9. Average attitude errors across 100 Monte-Caro runs for Qua-AttEstPO, Rod-AttEstPO and EKF with the associated 2-standard deviation bounds for small initial attitude error.

TABLE II
RMSE of Final Attitude Estimates

|  | Roll (deg) | Yaw (deg) | Pitch (deg) |
| :---: | :---: | :---: | :---: |
| EKF | 4.884 | 12.364 | 2.233 |
| Qua-AttEstPO | 0.026 | 0.082 | 0.026 |
| Rod-AttEstPO | 0.026 | 0.077 | 0.015 |

$\alpha=10 \mathrm{deg}$. The sampling frequency of the MIMU is set to 100 Hz . The biases of gyroscopes and accelerometers are assumed to be $\left[\begin{array}{lll}0.5 & 0.3 & -0.2\end{array}\right]^{T} \mathrm{deg}$ and $\left[\begin{array}{lll}0.1 & 0.2 & -0.2\end{array}\right]^{T} \mathrm{~m} / \mathrm{s}^{2}$, respectively. The root power spectral density of the gyroscope measurement noise of each axis is set to $1 \% / \sqrt{\mathrm{h}}$, and the noise standard deviations of the normalized magnetometer and accelerometer measurements in each axis are set to 0.02 and $0.01 \mathrm{~m} / \mathrm{s}^{2}$, respectively.

For all algorithms, the initial gyroscope and accelerometer biases are set to zeros and the statistics of the random noises are

TABLE III
Comparison of Computation Time

| Algorithm <br> $($ Window Size $)$ | Rod-AttEstPO <br> $(0.1 \mathrm{~s})$ | Qua-AttEstPO <br> $(0.1 \mathrm{~s})$ | Qua-AttEstPO <br> $(1 \mathrm{~s})$ | Qua-AttEstPO <br> $(5 \mathrm{~s})$ | Qua-AttEstPO <br> $(20 \mathrm{~s})$ |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Time Cost | 0.91 s | 3.37 s | 4.93 s | 25.4 s | 52.5 s |



Fig. 10. The MIMU installation in the experiment. The MIMU is placed in the center of the dodecahedron and five Vicon markers are attached on the edges to acquire the true attitude of the dodecahedron.
assumed to be known. The Chebyshev polynomial orders of both quaternion and Rodrigues vector approximation are set to $N_{q}=N_{g}=6$, the time window size is set to 0.1 seconds and the number of Chebyshev points in the window is set to $N+1=7$. In practical scenarios, the initial level angle (roll and pitch) can be readily obtained by the accelerometer leveling, but the yaw is more difficult to determine by the magnetometer due to the magnetic disturbance nearby. Thus, a 180 deg initial yaw error is assumed to fully assess the performance of different algorithms. Specifically, the initial orientation errors of roll, yaw and pitch are assumed as Gaussian noises with zero mean and the standard derivation of $\left[\begin{array}{lll}5 & 180 & 5\end{array}\right]$ deg.

Figures 2-4 plot the attitude, gyroscope and accelerometer bias errors across 100 Monte Carlo runs. We see that the QuaAttEstPO and Rod-AttEstPO are much superior to EKF in terms of stability and convergence speed. The average of absolute estimation errors is defined to quantity the estimation accuracy as

$$
\begin{equation*}
\varepsilon_{i}(k)=\frac{1}{L} \sum_{l=1}^{L}\left\|\hat{\mathbf{x}}_{(i), k}^{(l)}-\mathbf{x}_{(i), k}^{(l)}\right\| \tag{48}
\end{equation*}
$$

where $L$ denotes the number of Monte Carlo runs, and $i$ and $k$ denote the $i$-th state component at time $k$. For brevity, Figs. 5-6 respectively plot the norm of the average errors for attitude and gyroscope/accelerometer biases. The $2 \sigma$ bounds are calculated by twice the averaged square root of diagonal elements of the covariance matrix. For a consistent filter, the estimation error should stay below the $2 \sigma$ derivation bounds with a possibility of $95 \%$ [43]. Figures 5-6 indicate that both Qua-AttEstPO and Rod-AttEstPO are consistent estimators, while the EKF is too optimistic. The RMSE of the estimate at the end of simulation is listed in Table II. It shows that the RodAttEstPO is marginally better than Qua-AttEstPO, which may be arguably owed to the fact that the parameter dimension of Qua-AttEstPO is higher than that of Rod-AttEstPO and solving the constrained optimization in Qua-AttEstPO may be less


Fig. 11. The rough attitude trajectory obtained from the integration of the calibrated gyroscope measurements.
accurate than the unconstrained optimization in Rod-AttEstPO. However, due to the singularity of the Rodrigues vector at a $180^{\circ}$ rotation, it is not recommended to utilize the RodAttEstPO with a large window size. Therefore, we only use Qua-AttEstPO for a large window computation in the sequel. Figures 7-8 plot the average attitude and sensor bias errors of Qua-AttEstPO with different window sizes, namely, 0.1/1/5/20 seconds. The Chebyshev polynomial orders are respectively set to $6,40,150$, and 300 . We see that the estimation accuracy improves along with the increased window size. Larger window size means the algorithm employs more sensor information to estimate the state, which causes a larger time delay. Since the algorithm cannot execute until all the measurements in the current window come in. Note that the generalized singular value decomposition of a high-dimension matrix as in (45) is time-consuming. Therefore, the initialization of Chebyshev coefficients for AttEstPO in a larger window is carried out by fitting the states acquired by AttEstPO of 0.1 s window size.

All of the proposed algorithms are implemented with the function 'lsqnonlin' on the MATLAB platform. The average time cost across 100 Monte Carlo runs is listed in Table III. The time costs of Rod-AttEstPO and Qua-AttEstPO with 0.1 s window size are respectively about 3 and 10 times larger than that of EKF, which can ensure a real-time algorithm execution for 20 -second data. Although the computation burden of QuaAttEstPO increases dramatically along with the window size, it can still be used for batch estimation.

The average attitude error across 100 Monte Caro simulation runs for small initial attitude error is plotted in Fig. 9. The initial attitude error is presumed as Gaussian noise with zero mean and $\left[\begin{array}{lll}5 & 10 & 5\end{array}\right]$ deg standard deviation, the window size is 0.1 s and the Chebyshev polynomial order is 6 , and other settings are the


Fig. 12. Relative rotation angle of the reference Vicon system in the experiment (Upper left is the details of the first 10 seconds).


Fig.13. Relative rotation angle errors across 100 Monte-Caro runs for QuaAttEstPO, Rod-AttEstPO and EKF in the experiment.
same with the above. The three algorithms achieve similar final accuracy, but the convergence speed of Qua-AttEstPO and Rod-AttEstPO is faster than that of EKF and the EKF appears inconsistent either in the first several seconds.

## B. Experiment Results

The experiment is conducted using the Xsens MTW unit with the sample frequency of 100 Hz . The bias stabilities of gyroscope and accelerometer are $10 \mathrm{deg} / \mathrm{h}$ and 0.1 mg , respectively. The power spectrum densities of gyroscope, accelerometer and magnetometer noises are respectively $0.01 \mathrm{deg} / \mathrm{s} / \sqrt{\mathrm{Hz}}, 20 \mu \mathrm{~g} / \sqrt{\mathrm{HZ}}$ and $0.2 \mathrm{mGauss} / \sqrt{\mathrm{Hz}}$. A Vicon motion capture system is used to provide the true attitude relative to its inherently-defined reference frame. As shown in Fig. 10, the MIMU is fixed at the center of a regular dodecahedron, with five Vicon markers attached on the edges of the dodecahedron. The Vicon system tracks the attitude change by the markers. In view of the magnetic material in the dodecahedron, the magnetometer is calibrated before the experiment [32, 33]. In the experiment, the unit stays stationary


Fig. 14. Gyroscope bias estimate results across 100 Monte-Caro runs for QuaAttEstPO, Rod-AttEstPO, EKF and the reference in the experiment.


Fig. 15. Average of relative rotation angle errors across 100 Monte Carlo runs for Qua-AttEstPO, Rod-AttEstPO and EKF in the experiment.
on a bench for about 90 seconds, then is picked up and rotated around itself for about 40 seconds, and finally put back on the bench at the same attitude. Two datasets are collected to test the proposed algorithms. The two datasets produce similar results, so we only show one dataset result in this paper for brevity. The rough attitude time trajectory is plotted in Fig. 11, which is obtained by integrating the gyroscope measurements with the gyroscope bias roughly calibrated by averaging the static gyroscope outputs before the rotation.

Due to the attitude misalignment between the MIMU and dodecahedron, their attitudes are not exactly the same. However, the relative rotation angle of the MIMU body frame, denoted by $\alpha_{b(0)}^{b(t)}$, should be identical with that of the Vicon. Specifically, the attitude quaternion estimate $\mathbf{q}_{n(t)}^{b(t)}$ of the MIMU is transformed to the relative attitude quaternion $\mathbf{q}_{b(0)}^{b(t)}$ by multiplying the conjugate of the MIMU initial quaternion $\mathbf{q}_{n, 0}^{b}$,
i.e., $\mathbf{q}_{b(0)}^{b(t)}=\mathbf{q}_{n, 0}^{b}{ }^{*}{ }^{\circ} \mathbf{q}_{n}^{b}(t)$. Then, the relative rotation angle can be computed by

$$
\begin{equation*}
\alpha_{b(0)}^{b(t)}=2 \arccos \left(\left[\mathbf{q}_{b(0)}^{b(t)}\right]_{1}\right) \tag{49}
\end{equation*}
$$

where $\left[\mathbf{q}_{b(0)}^{b(t)}\right]_{1}$ denotes the first element of $\mathbf{q}_{b(0)}^{b(t)}$ and the relative angle is defined on $\left[\begin{array}{ll}0 & \pi\end{array}\right]$. The error of the relative rotation angle is defined as

$$
\begin{equation*}
\Delta \alpha=\left\|\alpha_{v i c o n}-\alpha_{b(0)}^{b(t)}\right\| \tag{50}
\end{equation*}
$$

where $\alpha_{v i c o n}$ is the relative rotation angle of the Vicon system, which is similarly computed from the Vicon's attitude output. Figure 12 plots the relative rotation angle from the Vicon system, where the variation during the first 10 -second static outputs indicates that the accuracy of relative rotation angle is about 0.3 deg .

The window size of Rod-AttEstPO and Qua-AttEstPO is 0.1 s and the Chebyshev polynomial order is 6 . A set of Monte Carlo runs are also performed in the experiment by intentionally randomizing the initial conditions. Specifically, the initial gyroscope and accelerometer biases are set to zeros and the initial attitude error is set to Gaussian noise with zero mean and $\left[\begin{array}{lll}5 & 180 & 5\end{array}\right]$ deg standard deviation. Figures 13-14 plot the relative rotation angle error and the gyroscope bias across 100 Monte Carlo runs, where the reference gyroscope bias is roughly obtained by averaging the static gyroscope outputs before the rotation. It can be seen that the results of QuaAttEstPO and Rod-AttEstPO for different initial state values are quite similar in each run, which demonstrates the superior stability of AttEstPO in face of large initial attitude errors. And, the average of relative rotation angle errors in Fig. 15 also highlights the accuracy superiority of Qua-AttEstPO and RodAttEstPO over EKF. The RMSEs of relative rotation angle at the end of experiment are about 1.1, 0.03, 0.02 degrees for EKF, Qua-AttEstPO and Rod-AttEstPO, respectively.

## V. Conclusion

Attitude estimation by way of inertial-magnetic sensors is vital for many applications. Traditional linearization-based EKF is prone to divergence in the challenging scenarios with large initial state error. To improve the estimation stability and convergence speed, this paper introduces the Chebyshev collocation method into the attitude estimation, resulting in the so-called Qua-AttEstPO and Rod-AttEstPO algorithms. They essentially transform the quaternion/Rodrigues vector estimation into constrained/unconstrained least squares problem in a time window of interest. Due to the inherent singularity of the Rodrigues vector, Rod-AttEstPO is only suitable for a short time window. Simulation and field test results show that the Qua-AttEstPO and Rod-AttEstPO algorithms have similar accuracy with each other and perform quite well under large initial errors. Increasing the size of time window in Qua-AttEstPO can further improve the estimation accuracy yet at the price of time delay. Therefore, QuaAttEstPO with a large window size can be treated as a batch
attitude estimation method. Our future work will explore the advantage of the polynomial optimization-based estimation in other navigation applications, e.g., INS/GNSS integrated navigation.

## AcKNOWLEDGMENT

Thanks to Miss Miaomiao Yan and Dr. Boying Li, at Shanghai Jiao Tong University, for their test assistance with the Vicon tracking system.

## REFERENCES

[1] H. Rehbinder, and X. Hu, "Drift-free attitude estimation for accelerated rigid bodies," Automatica, vol. 40, no. 4, pp. 653-659, 2004.
[2] G. Vukovich, and H. Gui, "Robust adaptive tracking of rigid-body motion with applications to asteroid proximity operations," IEEE Transactions on Aerospace and Electronic Systems, vol. 53, no. 1, pp. 419-430, 2017.
[3] M. Hua, G. Ducard, T. Hamel, R. Mahony, and K. Rudin, "Implementation of a nonlinear attitude estimator for aerial robotic vehicles," IEEE Transactions on Control Systems Technology, vol. 22, no. 1, pp. 201-213, 2013.
[4] D. Roetenberg, H. Luinge, and P. Slycke, "Xsens MVN: Full 6DOF human motion tracking using miniature inertial sensors,"Xsens Motion Technologies BV, Tech. Rep, vol. 1, 2009.
[5] M. Zhu, Y. Wu, and S. Luo, "f2IMU-R: pedestrian navigation by low-cost foot-mounted dual IMUs and interfoot ranging," IEEE Transactions on Control Systems Technology, vol. 30, no. 1, pp. 247-260, 2022.
[6] P. D. Groves, Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems: Artech house, 2013.
[7] J. Farrell, Aided Navigation: GPS with High Rate Sensors: McGraw-Hill, Inc., 2008.
[8] M. Nazarahari, and H. Rouhani, "40 years of sensor fusion for orientation tracking via magnetic and inertial measurement units: Methods, lessons learned, and future challenges," Information Fusion, vol. 68, pp. 67-84, 2021.
[9] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," IEEE Transactions on Automatic Control, vol. 53, no. 5, pp. 1203-1217, 2008.
[10] J. Wu, Z. Zhou, F. H, R. Li, and M. Liu, "Generalized linear quaternion complementary filter for attitude estimation from multisensor observations: an optimization approach," IEEE Transactions on Automation Science and Engineering, vol. 16, no. 3, pp. 1330-1343, 2019.
[11] J. F. Vasconcelos, B. Cardeira, C. Silvestre, P. Oliveira, and P. Batista, "Discrete-time complementary filters for attitude and position estimation: Design, analysis and experimental validation," IEEE Transactions on Control Systems Technology, vol. 19, no. 1, pp. 181-198, 2010.
[12] P. Marantos, Y. Koveos, and K. J. Kyriakopoulos, "UAV state estimation using adaptive complementary filters," IEEE Transactions on Control Systems Technology, vol. 24, no. 4, pp. 1214-1226, 2015.
[13] A. M. Sabatini, "Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing," IEEE Transactions on Biomedical Engineering, vol. 53, no. 7, pp. 1346-1356, 2006.
[14] J. L. Marins, X. Yun, E. R. Bachmann, R. B. McGhee, and M. J. Zyda, "An extended Kalman filter for quaternion-based orientation estimation using MARG sensors," in Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2001, pp. 2003-2011.
[15] X. Yun, and E. R. Bachmann, "Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking," IEEE Transactions on Robotics, vol. 22, no. 6, pp. 1216-1227, 2006.
[16] M. Zhu, W. Ouyang, and Y. Wu, "Orientation estimation by partial-state updating Kalman filter and vectorial magnetic interference detection," IEEE Transactions on Aerospace and Electronic Systems, vol. 57, no. 3, pp. 1815-1826, 2021.
[17] S. O. Madgwick, A. J. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in 2011 IEEE international conference on rehabilitation robotics, 2011, pp. 1-7.
[18] S. Wilson, H. Eberle, Y. Hayashi, S. O. Madgwick, A. McGregor, X. Jing, and R. Vaidyanathan, "Formulation of a new gradient descent MARG
orientation algorithm: Case study on robot teleoperation," Mechanical Systems and Signal Processing, vol. 130, pp. 183-200, 2019.
[19] J. Vandersteen, M. Diehl, C. Aerts, and J. Swevers, "Spacecraft attitude estimation and sensor calibration using moving horizon estimation," Journal of Guidance Control and Dynamics, vol. 36, no. 3, pp. 536-542, 2013.
[20] W. Higgins, "A comparison of complementary and Kalman filtering," IEEE Transactions on Aerospace and Electronic Systems, vol. 11, no. 3, pp. 321-325, 1975.
[21] F. L. Markley, and J. L. Crassidis, Fundamentals of Spacecraft Attitude Determination and Control: Springer, 2014.
[22] A. Barrau, and S. Bonnabel, "Intrinsic filtering on Lie groups with applications to attitude estimation," IEEE Transactions on Automatic Control vol. 60, no. 2, pp. 436-449, 2015.
[23] H. Gui, and A. H. J. d. Ruiter, "Quaternion invariant extended Kalman filtering for spacecraft attitude estimation," Journal of Guidance Control and Dynamics, vol. 41, no. 4, pp. 863-878, 2018.
[24] J. L. Crassidis, and F. L. Markley, "Unscented filtering for spacecraft attitude estimation," Journal of Guidance Control and Dynamics, vol. 26, no. 4, pp. 536-542, 2003.
[25] Y. Cheng, and J. L. Crassidis, "Particle filtering for attitude estimation using a minimal local-error representation," Journal of Guidance Control and Dynamics, vol. 33, no. 4, pp. 1305-1310, 2010.
[26] H. Fourati, N. Manamanni, L. Afilal, and Y. Handrich, "Complementary observer for body segments motion capturing by inertial and magnetic sensors," IEEE/ASME Transactions on Mechatronics, vol. 19, no. 1, pp. 149-157, 2014.
[27] H. Sommer, J. R. Forbes, R. Siegwart, and P. Furgale, "Continuous-time estimation of attitude using B-splines on Lie groups," Journal of Guidance Control and Dynamics, vol. 39, no. 2, 2016.
[28] M. Zhu, and Y. Wu, "Continuous-time State Estimation by Chebyshev Polynomial Optimization," IEEE Transactions on Signal Processing, vol. 70, pp. 3136-3147, 2022.
[29] M. Zhu, and Y. Wu, "Lightweight precision inertial computation based on chebyshev polynomial optimization," in 29th Saint Petersburg International Conference on Integrated Navigation Systems, 2022, pp. 19.
[30] L. N. Trefethen, Approximation Theory and Approximation Practice: SIAM, 2019.
[31] J. P. Boyd, Chebyshev and Fourier Spectral Methods, 2 ed.: Courier Corporation, 2001.
[32] M. Zhu, Y. Wu, and W. Yu, "An efficient method for gyroscope-aided full magnetometer calibration," IEEE Sensors Journal, vol. 19, no. 15, pp. 6355-6361, 2019.
[33] Y. Wu, D. Zou, P. Liu, and W. Yu, "Dynamic magnetometer calibration and alignment to inertial sensors by Kalman filtering," IEEE Transactions on Control Systems Technology, vol. 26, no. 2, pp. 716-723, 2017.
[34] A. H. Jazwinski, Stochastic Process and Filtering Theory, NY, USA: Academic Press, 1970.
[35] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Numerical Recipes : the Art of Scientific Computing, 3 ed.: Cambridge University Press, 2007.
[36] G. Klein, "An extension of the Floater-Hormann family of barycentric rational interpolants," Mathematics of Computation, vol. 82, pp. 22732292, 2013.
[37] Y. Wu, and M. Zhu, "Attitude reconstruction from inertial measurement: mitigating Runge effect for dynamic applications," IEEE Transactions on Aerospace and Electronic Systems, 2021.
[38] I. M. Ross, and M. Karpenko, "A review of pseudospectral optimal control: From theory to flight," Annual Reviews in Control, vol. 36, no. 2, pp. 182-197, 2012.
[39] S. Boyd, and L. Vandenberghe, Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares: Cambridge University Press, 2018.
[40] J. Berrut, and G. Klein, "Recent advances in linear barycentric rational interpolation," Journal of Computational and Applied Mathematics, vol. 259, pp. 95-107, 2014.
[41] C. V. Rao, J. B. Rawlings, and D. Q. Mayne, "Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations," IEEE Transactions on Automatic Control, vol. 48, no. 2, pp. 246-258, 2003.
[42] K. Inkilä, "Homogeneous least squares problem," Photogrammetric Journal of Finland, vol. 19, no. 2, pp. 34-42, 2005.
[43] Y. Wu, D. Hu, M. Wu, and X. Hu, "A numerical-integration perspective on Gaussian filters," IEEE Transactions on Signal Processing, vol. 54, no. 8, pp. 2910-2921, 2006.


[^0]:    This paper was supported in part by National Natural Science Foundation of China (62273228) and Shanghai Jiao Tong University Scientific and Technological Innovation Funds.

[^1]:    Author's address: M. Zhu and Y. Wu are with Shanghai Key Laboratory of Navigation and Location-based Services, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (email: zhumaoran@sjtu.edu.cn, yuanx_wu@hotmail.com).

