

Received March 29, 2016, accepted April 21, 2016, date of publication April 27, 2016, date of current version May 11, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2559445

A Simple Attitude Unscented Kalman Filter: Theory and Evaluation in a Magnetometer-Only Spacecraft Scenario

MURTY S. CHALLA¹, (Member, IEEE), JAY G. MOORE²,
AND DANIEL J. ROGERS³, (Member, IEEE)

¹Air and Missile Defense Sector, The Johns Hopkins University Applied Physics Laboratory, Laurel, MD 20723-6099, USA

²Asymmetric Operations Sector, The Johns Hopkins University Applied Physics Laboratory, Laurel, MD 20723-6099, USA

³Terbium Labs, Baltimore, MD 21230, USA

Corresponding author: M. S. Challa (e-mail: murty.challa@jhuapl.edu).

This work was supported by the Research and Exploratory Development Department and the Air and Missile Defense Sector of The Johns Hopkins University Applied Physics Laboratory.

ABSTRACT A quaternion-based attitude unscented Kalman filter is formulated with quaternion errors parameterized by small angle approximations and is applied to a filter with a state vector consisting of the attitude quaternion and the gyro bias vector. The filter is evaluated using extensive Monte Carlo data in a simulated lost-in-space scenario of a low-Earth orbiting spacecraft processing only three-axis magnetometer and gyro measurements. The filter is found to be robust, accurate, and rapidly convergent in this scenario for small true gyro biases and small initial uncertainties in their values, often converging in only one half of an orbit period to an attitude accuracy of 0.1 degrees. The filter convergence is found to depend significantly on the value of the true gyro biases as well as the initial gyro bias covariances. Monte Carlo results also indicate that this unscented Kalman filter is significantly less robust than an extended Kalman filter with the same attitude approach, but performs slightly better than another unscented Kalman filter with a generalized Rodrigues parameter approach to quaternion errors.

INDEX TERMS Attitude control, calibration, geomagnetism, gyroscopes, Kalman filters, spacecraft, unscented.

I. INTRODUCTION

The Kalman filter [1]–[3] has been a standard real-time estimation algorithm for over five decades now. The filter consists of two steps – processing of measurements to estimate the state vector, X , and propagating the updated state vector between measurements using dynamics. It optimizes the state estimation by propagating and updating the (error) covariance matrix, P , of the state using statistics of the noise in the measurement and propagation models. A crucial aspect of the filter is that the measurement models as well as the dynamic models are linear functions of X . However, a large number of phenomena in physics and engineering involve nonlinear dynamics and/or measurements. Soon after the development of the Kalman filter, these nonlinearities were handled in the Kalman filter framework via the Extended Kalman Filter (EKF) by linearizing the models [4], [5].

The linearizations impart a tendency to the EKF to underestimate P and thereby raise the possibility of

instability and divergent state estimates [6]. To counter this, a more recent development has been the Unscented Kalman Filter (UKF) by Julier and Uhlmann [7] and Julier *et al.* [8], who argued that it is safer to approximate P than to linearize the state vector dynamics or the measurements. Wan and van der Merwe [9] provide additional applications of the UKF and comparisons with the EKF. The UKF approach involves retaining the complete dynamics and measurement models on a small ensemble of state vector values called “sigma points” which are strategically chosen from the *a priori* P , and updating X and P from the statistics of this small ensemble. In particular, in the UKF, P is not propagated between measurements via a differential equation. This also eliminates the need to compute Jacobians at every time step, which itself can be a significant advantage of the UKF in large systems. (See, for example, the 9×9 matrix in [10], which results from an error analysis of an inertial navigation system. Significantly larger versions of this matrix equation

can result, for example, if we wish to estimate various errors in sensor models with an aim to calibrate the sensors.)

Attitude EKFs have been around since the 1960's [11]–[15] because attitude necessarily involves nonlinearities, especially when angular representations such as Euler angles are used. Recently, unscented attitude Kalman filters have also begun to be used [16]–[24] in a variety of fields such as robotics [16], inertial navigation [17], [23], and spacecraft [18]–[22], [24]. Attitude parameterization has mostly been in terms of quaternions where the treatments of attitude errors generally fall into two classes: small rotation angle approximations [16], [17], [20], [21], [24], and generalized Rodrigues parameters (GRPs) [18], [19], [22].

The UKF approach is promising and, because of the constant need for improved real-time attitude algorithms, we explore it additionally in this paper. We use quaternions to parameterize the attitude and treat the attitude errors as small rotation angles about the Body axes, which is the "Body-Fixed Covariance Representation" of Lefferts *et al.* [15]. The attitude errors thus have only three components and the purely attitude part of the covariance matrix is only 3×3 . By Euler's theorem, the attitude can be characterized by a single rotation angle, θ , about a unit vector, \mathbf{n} , with Body components $[n_x, n_y, n_z]^T$, and the corresponding quaternion can be given a more physical interpretation [25]–[27]:

$$\mathbf{q} = \begin{bmatrix} \mathbf{n} \sin(\theta/2) \\ \cos(\theta/2) \end{bmatrix}. \quad (1)$$

With " δ " indicating small differential rotations and $\delta\theta \equiv \delta\theta[n_x, n_y, n_z]^T \equiv [\delta\theta_x, \delta\theta_y, \delta\theta_z]^T$, (1) reduces to:

$$\delta\mathbf{q} = \begin{bmatrix} \delta\theta/2 \\ 1 \end{bmatrix}, \quad |\delta\theta| \ll 1 \quad (2)$$

(Note that vectors and matrices are in bold italic font, and that vectors are defined as column matrices for use in computations.) The "delta quaternion" in (2) is useful in treating the errors and updating the attitude, especially because attitude matrices corresponding to $\delta\theta$ commute approximately up to order $|\delta\theta|^2$, and the delta quaternion in (2) is normalized approximately to unity to the same order. We use this formalism in constructing the attitude innovation during the Kalman filter scheme. The normalization of \mathbf{q} to unity being a strict requirement for it to represent attitude, we explicitly normalize quaternions after propagation and measurement updates.

Our small angles treatment here is similar to that of Kraft [16] and Khoder *et al.* [17] but differs from the former by not having an iterative procedure for determining the mean sigma point and from the latter by not using either the "reduced" form of the UKF or spherical simplex scaling. Our approach is simpler than that of Cheon and Kim [24] who used an optimized approximate cost function for computing the mean quaternion. A further distinguishing characteristic of our approach is the augmentation of the covariance matrix and explicit use of process noise in constructing

the sigma points. Our approach also differs from that of Crassidis and Markley [18] who used GRPs to preserve quaternion normalization irrespective of the size of the initial errors.

We have three broad objectives in this paper: (i) Theoretical formulation of a simple attitude UKF which uses quaternions to parameterize the attitude, a small-angle formalism for quaternion updates, and a basic "first principles" Unscented Transform without complicated statistical artifacts or analyses; (ii) A comprehensive evaluation of the robustness and accuracy of the algorithm in a stressing scenario; (iii) Comparative evaluations of other Kalman filter approaches under identical conditions.

Our first objective is motivated by the fact that (2) rapidly approaches unit magnitude as the filter converges to the truth; for example, the cosine of 2 degrees (deg) differs from 1 by less than 0.1%. Of course, the small angles formalism will not eliminate all attitude errors if they are large during the initial stages of a Kalman filter run but the most important question really is whether an algorithm will systematically converge to the truth in spite of large initial errors. (As a trivial example to illustrate the importance of robustness, note that if an algorithm can consistently eliminate 10% of the errors at each time step, it would reach 1 deg accuracy in only 50 steps even after starting with an error of 180 deg.) In any case, explicit normalization during the attitude updates, besides being a quaternion requirement, prevents numerical errors from accumulating during a run and would thus expose any intrinsic weaknesses of the algorithm.

We address the second objective by applying this new UKF to the attitude determination of a low Earth orbit spacecraft. Since spacecraft usually also carry gyroscopes ("gyros") as part of the attitude sensor package and since the gyro biases can drift significantly over time, our state vector, \mathbf{X} , consists of the attitude quaternion as well as the three Body-frame components of the gyro bias vector, \mathbf{b} . The scenario is made stressing by making it "magnetometer-only, lost-in-space", wherein the spacecraft attitude as well as the gyro bias vector are unknown at the start and only measurements from a noisy three-axis magnetometer and a bias-corrupted noisy gyro are input to the Kalman filter. We assume that the magnetometer is well-calibrated and thus has no other errors. We use simulated data of NASA's Tropical Rainfall Measuring Mission (TRMM) spacecraft [28] and examine performance metrics such as robustness, accuracy, and speed of convergence by gathering statistics over many Monte Carlo runs with random initial errors.

We accomplish the final objective by extensive Monte Carlo studies comparing results from this UKF with two other Kalman filters. One of these filters is a standard attitude EKF with a similar small-angle attitude approach, and the other is a UKF which uses GRPs [18] instead of small angles to treat the attitude errors. This aspect of our work enable us to make a general UKF vs. EKF comparison as well as a comparison of two different attitude Kalman filter innovation approaches.

An important and unique feature of our work is the large volume of Monte Carlo data we use to evaluate the algorithms – *batches of 1000 Monte Carlo runs for each set of parameters or algorithm, with each run spanning at least 7 orbits*. A large amount of data such as this is needed to reach definitive conclusions for future use in attitude algorithms supporting autonomous systems such as spacecraft. Such extensive data is unfortunately missing in the current literature and conclusions on the relative merits of algorithms are often based on just a handful of runs. Another feature of our evaluations is also the consistency: once tuned, filter parameters were kept constant across all the Monte Carlo batches irrespective of the input parameters such as true initial gyro biases or *a priori* uncertainties. This is because any filter which needs re-tuning depending upon, for example, the location in orbit, will be less desirable. (We are not referring to adaptive filters in this context; all three filters studied here require manual tuning.) This mass of Monte Carlo data allows us to draw clear inferences besides also definitively pointing to future directions of study.

The choice of a spacecraft scenario for evaluation satisfies several secondary objectives. Historically, spacecraft have been one of the earliest “autonomous systems” and attitude determination has been an important navigational function because spacecraft need to satisfy a variety of requirements such as Earth pointing for communications or Sun avoidance for planetary observations. Hence, the results presented here will be of direct interest to the spacecraft community especially because the calibration of the gyro bias while simultaneously estimating attitude is commonly necessitated by requiring accurate gyro-based attitude maneuvers. The results will also be of interest to designers and users of a plethora of autonomous inertial navigational systems and unmanned vehicles these days, most of which use inertial measurement units comprising of gyros and accelerometers. This scenario also demonstrates to the broader technical community the “magnetometer-only” spacecraft attitude determination or attitude-and-rate determination capability which has evolved during the past three decades [24], [29]–[37]. Such a capability is possible on spacecraft because the magnetic field is constantly changing direction in the Body frame as the spacecraft orbits the Earth. As an example, Fig. 1 shows the variation of TRMM simulated measurements used in this paper. The magnetometer-only capability is attainable because, although at any instant the rotation angle about the magnetic field vector is unobservable, the propagated angle from an earlier instant can be used to complete the full attitude estimate in real-time if we have an algorithm such as a robust Kalman filter.

It will be seen below that “magnetometer + gyro” attitude accuracies of the order of 0.1 deg can be regularly achieved, often in as little as 1/3 of an orbit, which is about 1800 s for TRMM. Thus, this paper will also serve to demonstrate to the technical community

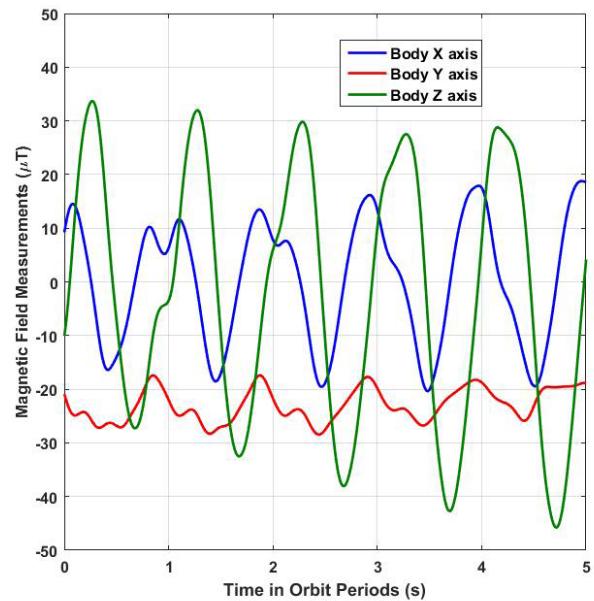


FIGURE 1. Simulated geomagnetic flux density measurements of the TRMM spacecraft.

that the “magnetometer-only” capability can lead to inexpensive spacecraft in two ways: by adding gyros or even eliminating the gyros altogether [29], [30], [34]. (A third use can be even more cost-saving – use of these techniques as contingency algorithms on expensive spacecraft.)

The rest of this paper is organized as follows. Section II deals with theory, and is made “standalone” via detailed subsections on quaternion algebra (IIA), general UKF approach (IIB), and our attitude UKF approach (IIC). Section III details results obtained using simulated data and has subsections dealing with scenario and data characteristics (IIIA), performance of our UKF using Monte Carlo data (IIIB), influence of initial conditions (IIIC), and comparisons of performances with two other Kalman filters – an EKF (IIID) and a GRP-based UKF (IIIE). Finally, the conclusions are summarized in Section IV.

II. THEORY

A. QUATERNION ESSENTIALS

We define here for ready reference some essential quaternion relationships invoked in this paper, beyond (1) and (2). More details on quaternions and attitude parameterizations are provided in [26] and [27].

An important requirement of an attitude quaternion, $\mathbf{q} = [q_1 \ q_2 \ q_3 \ q_4]^T$, is that it be normalized to unity. That is,

$$|\mathbf{q}|^2 = q_1^2 + q_2^2 + q_3^2 + q_4^2 = 1 \quad (3)$$

The attitude matrix (i.e., direction cosine matrix), \mathbf{A} , is related to \mathbf{q} , as follows:

$$A(\mathbf{q}) \equiv \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1 q_2 + q_3 q_4) & 2(q_1 q_3 - q_2 q_4) \\ 2(q_1 q_2 - q_3 q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2 q_3 + q_1 q_4) \\ 2(q_1 q_3 + q_2 q_4) & 2(q_2 q_3 - q_1 q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} \quad (4)$$

The product of two quaternions can be written compactly if the first three components of \mathbf{q} are denoted by a 3×1 vector, \mathbf{v} , and the fourth component by a scalar, s . (1) may then be expressed as:

$$\mathbf{q} = \begin{bmatrix} \mathbf{n} \sin(\theta/2) \\ \cos(\theta/2) \end{bmatrix} \equiv \begin{bmatrix} \mathbf{v} \\ s \end{bmatrix}, \quad (5)$$

and the quaternion multiplication, “ \otimes ”, between two quaternions \mathbf{q}_1 and \mathbf{q}_2 may be expressed as:

$$\mathbf{q} = \mathbf{q}_1 \otimes \mathbf{q}_2 = \begin{bmatrix} \mathbf{v}_1 \\ s_1 \end{bmatrix} \otimes \begin{bmatrix} \mathbf{v}_2 \\ s_2 \end{bmatrix} \equiv \begin{bmatrix} \mathbf{v}_1 \times \mathbf{v}_2 + s_1 \mathbf{v}_2 + s_2 \mathbf{v}_1 \\ s_1 s_2 - \mathbf{v}_1 \cdot \mathbf{v}_2 \end{bmatrix}. \quad (6)$$

The order in (6) is the reverse of the order of multiplication of the corresponding attitude matrices. (That is, the matrix analog of equation (6) is $A(\mathbf{q}) = A(\mathbf{q}_2) A(\mathbf{q}_1)$ after constructing the matrices from (4).) The identity quaternion, $[0\ 0\ 0\ 1]^T$ is generated if there is no rotation ($\theta = 0$). The inverse of the quaternion in (5) is defined by a rotation of equal magnitude in the opposite sense:

$$\mathbf{q}^{-1} = \begin{bmatrix} \mathbf{n} \sin(-\theta/2) \\ \cos(-\theta/2) \end{bmatrix} = \begin{bmatrix} -\mathbf{v} \\ s \end{bmatrix}. \quad (7)$$

If $\hat{\mathbf{q}}$ denotes an estimated quaternion and \mathbf{q}_t denotes the true quaternion, the error quaternion, $\delta\mathbf{q}$, represents an additional rotation in $\hat{\mathbf{q}}$ beyond that described by \mathbf{q}_t in correspondence with the matrix product, $A(\hat{\mathbf{q}}) = A(\delta\mathbf{q})A(\mathbf{q}_t)$:

$$\delta\mathbf{q} = \mathbf{q}_t^{-1} \otimes \hat{\mathbf{q}}. \quad (8)$$

Most calculations are done using matrices. To facilitate this, vector cross products are converted to matrix multiplication via the antisymmetric matrix $[\mathbf{a} \times]$ defined for a vector $\mathbf{a} \equiv (a_1, a_2, a_3)$ as:

$$[\mathbf{a} \times] = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}. \quad (9)$$

The matrix product $[\mathbf{a} \times] \mathbf{b}$ is then equivalent to the vector cross product $\mathbf{a} \times \mathbf{b}$.

B. GENERAL UKF FORMULATION

Let the measurement times be denoted t_{k-1} , t_k , etc. The $\ell \times 1$ state vector, X , evolves according to the following ordinary differential equation:

$$\frac{dX}{dt} = f(X) + \eta(t), \quad (10)$$

where $\eta(t)$ is a zero-mean white noise vector (“process noise”) and f is in general a nonlinear function of X . (10) is a simplified form of the general propagation equation used in control theory [3]; for example, it does not feature a control vector on the right hand side.) The covariance of the process

noise at t_k has a spectral density matrix, $\mathbf{Q}(t)$, defined as follows:

$$\langle \eta(t) \eta^T(t') \rangle = \mathbf{Q}(t) \delta(t - t'), \quad (11)$$

where $\delta(t - t')$ is the Dirac delta function.

In the (linear) Kalman filter, the propagation of X can be formally treated using a transition matrix [3], $\Phi(t_k, t_{k-1})$, so that the propagation of the $\ell \times \ell$ covariance matrix, \mathbf{P} , can be written formally as

$$\mathbf{P}_{k|k-1} = \Phi(t_k, t_{k-1}) \mathbf{P}_{k-1|k-1} \Phi^T(t_k, t_{k-1}) + \mathbf{Q}_{k-1}, \quad (12)$$

where the subscript “ $k|k-1$ ” and “ $k-1|k-1$ ” respectively denote *a priori* value at t_k and *a posteriori* value at t_{k-1} based on measurements up to t_{k-1} , and \mathbf{Q}_{k-1} is the $\ell \times \ell$ process noise covariance matrix related to $\mathbf{Q}(t)$ via:

$$\mathbf{Q}_{k-1} = \int_{t_{k-1}}^{t_k} \Phi(t_k, \tau) \mathbf{Q}(\tau) \Phi^T(t_k, \tau) d\tau. \quad (13)$$

Although the UKF is characterized by an absence of a propagation equation such as (12) for the covariance matrix, \mathbf{Q}_{k-1} will be used to account for the process noise during propagation.

The $m \times 1$ measurement vector, \mathbf{z}_k , at t_k is modeled via:

$$\mathbf{z}_k = \mathbf{h}(X_k) + \mathbf{v}_k \quad (14)$$

where \mathbf{v}_k is a zero-mean white noise vector, and \mathbf{h} is in general a nonlinear function of X . The $m \times m$ covariance matrix of the measurement noise at t_k , \mathbf{R}_k , is defined as follows:

$$\langle \mathbf{v}_k \mathbf{v}_k^T \rangle = \mathbf{R}_k \quad (15)$$

Consider first the UKF prediction of the measurement at t_k by propagating X from t_{k-1} . The augmented state vector and covariance matrix (with superscripts “ a ”) are constructed at t_{k-1} using the statistics of the process noise, with η_k being drawn from $N(\mathbf{0}, \mathbf{Q}_k)$:

$$X_{k-1|k-1}^a = \left[\hat{X}_{k-1|k-1}^T \langle \eta_k^T \rangle \right]^T = \left[\hat{X}_{k-1|k-1}^T \mathbf{0}_{l \times 1}^T \right]^T \quad (16)$$

$$\mathbf{P}_{k-1|k-1}^a = \begin{bmatrix} \mathbf{P}_{k-1|k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{k-1} \end{bmatrix}. \quad (17)$$

X^a will be $L \times 1$, \mathbf{P}^a will be $L \times L$, where $L = 2\ell$ if all elements of the state are independent, and all four submatrices in (17) are $\ell \times \ell$. A set of $2L + 1$ “sigma points”, $\{\chi\}$, each of dimensions $L \times 1$ are then constructed via the following formulas:

$$\chi_{k-1|k-1}^0 = X_{k-1|k-1}^a, \quad (18)$$

$$\chi_{k-1|k-1}^i = \chi_{k-1|k-1}^0 + \left(\sqrt{(L + \lambda) \mathbf{P}_{k-1|k-1}^a} \right)_i, \quad i = 1, \dots, L, \quad (19)$$

$$\chi_{k-1|k-1}^i = \chi_{k-1|k-1}^0 - \left(\sqrt{(L + \lambda) \mathbf{P}_{k-1|k-1}^a} \right)_{i-L}, \quad i = L + 1, \dots, 2L, \quad (20)$$

where λ is a user-defined parameter discussed later and $\left(\sqrt{(L + \lambda) \mathbf{P}_{k-1|k-1}^a} \right)_i$ denotes the i -th column of the matrix square root of $(L + \lambda) \mathbf{P}_{k-1|k-1}^a$. We used Cholesky decomposition to determine the matrix square root; thus, if S is the square root of a matrix, M , then $M = S S^T$. (Reference [23] evaluates various matrix square root algorithms in the context of developing an attitude UKF.)

Each sigma point is now propagated using the dynamics of the state vector in (10) as follows, with the process noise added in via the corresponding extra elements in the sigma points:

$$\begin{aligned} \chi_{k|k-1}^i(1 : \ell) &= \int f(\chi_{k-1|k-1}^i(1 : \ell)) dt \\ &\quad + \chi_{k-1|k-1}^i(\ell + 1 : L), \quad i = 0, 1, \dots, 2L \end{aligned} \quad (21)$$

((21) will be adapted to attitude quaternions in section IIIC because not all elements of X will then be independent and thus $L \neq 2\ell$.) The predicted state and covariance are constructed by taking weighted averages of the propagated sigma points as follows:

$$\hat{X}_{k|k-1} = \sum_{i=0}^{2L} W_s^i \chi_{k|k-1}^i(1 : \ell), \quad (22)$$

$$\begin{aligned} \mathbf{P}_{k|k-1} &= \sum_{i=0}^{2L} W_c^i \left[\chi_{k|k-1}^i(1 : \ell) - \hat{X}_{k|k-1} \right] \\ &\quad \times \left[\chi_{k|k-1}^i(1 : \ell) - \hat{X}_{k|k-1} \right]^T \end{aligned} \quad (23)$$

(An alternative procedure in the case of purely additive process noise is to add it via \mathbf{Q} to the predicted covariance in (23) as in [18].) The weights, $\{W\}$, and λ are defined below in terms of user-defined parameters α , β , and κ :

$$W_s^0 = \frac{\lambda}{L + \lambda}, \quad (24)$$

$$W_c^0 = W_s^0 + (1 - \alpha^2 + \beta), \quad (25)$$

$$W_s^i = W_c^i = \frac{1}{2(L + \lambda)}, \quad i \neq 0 \quad (26)$$

and

$$\lambda = \alpha^2(L + \kappa) - L. \quad (27)$$

The updates using measurements are accomplished as follows using the propagated state and covariance of (22) and (23), and the statistics of the measurement noise. The dimensions are in general different from the propagation case because the size, m , of the observation vector is usually different from that of X . The augmented quantities are generated via:

$$\mathbf{X}_{k|k-1}^a = \begin{bmatrix} \hat{X}_{k|k-1}^T & < \mathbf{v}_k^T > \end{bmatrix}^T = \begin{bmatrix} \hat{X}_{k|k-1}^T & \mathbf{0}_{m \times 1}^T \end{bmatrix}^T \quad (28)$$

$$\mathbf{P}_{k|k-1}^a = \begin{bmatrix} \mathbf{P}_{k|k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_k \end{bmatrix}. \quad (29)$$

(To avoid notational clutter, we retained the same notation for the augmented state vectors on the left hand side of (16) and (28), and (17) and (29) even though they have different lengths. Recognizing the context - propagation or measurement - will reduce any confusion.) Here $\mathbf{P}_{k|k-1}^a$ is of size $L' \times L'$, where $L' \equiv \ell + m$, and the sigma points, $\{\chi_{k|k-1}^i, i = 0, 1, \dots, 2L'\}$, are derived in a manner similar to (18) to (20). The predicted $m \times 1$ measurement vectors, $\{\gamma\}$, are generated using the model in (14) together with the noise generated by the $\ell + 1 : L'$ elements of the sigma points as follows:

$$\boldsymbol{\gamma}_k^i = \mathbf{h}(\chi_{k-1|k-1}^i(1 : \ell)) + \chi_{k-1|k-1}^i(\ell + 1 : L'), \quad i = 0, 1, \dots, 2L'. \quad (30)$$

As with (21), the addition of the noise in (30) will be adapted to quaternions in Section IIIC. (Alternatively, purely additive process noise can be added via \mathbf{R} to the covariance \mathbf{P}_{zz} defined below as in [18].)

The weighted average measurement, \hat{z}_k , and the $m \times m$ measurement covariance matrix, \mathbf{P}_{zz} , are generated from $\{\boldsymbol{\gamma}\}$ as follows:

$$\hat{z}_k = \sum_{i=0}^{2L'} W_s^i \boldsymbol{\gamma}_k^i, \quad (31)$$

$$\mathbf{P}_{zz} = \sum_{i=0}^{2L'} W_c^i [\boldsymbol{\gamma}_k^i - \hat{z}_k] [\boldsymbol{\gamma}_k^i - \hat{z}_k]^T. \quad (32)$$

The $\ell \times m$ state-measurement cross-covariance matrix, \mathbf{P}_{xz} , is computed via:

$$\mathbf{P}_{xz} = \sum_{i=0}^{2L'} W_c^i [\chi_{k|k-1}^i(1 : \ell) - \hat{X}_{k|k-1}] [\boldsymbol{\gamma}_k^i - \hat{z}_k]^T. \quad (33)$$

(32) and (33) are used to compute the UKF Kalman gain:

$$\mathbf{K}_k = \mathbf{P}_{xz} \mathbf{P}_{zz}^{-1}. \quad (34)$$

The state vector is updated using the Kalman gain times the measurement residual:

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + \mathbf{K}_k (z_k - \hat{z}_k), \quad (35)$$

and the covariance is updated via:

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{P}_{zz} \mathbf{K}_k^T. \quad (36)$$

C. FORMULATION OF THE ATTITUDE UKF

X is the following 7×1 state vector:

$$X = \begin{bmatrix} \mathbf{q} \\ \mathbf{b} \end{bmatrix}, \quad (37)$$

where \mathbf{q} is the 4×1 Earth Centered Inertial (ECI) -to-Body attitude quaternion and \mathbf{b} is the 3×1 vector of gyro biases. Although $\ell = 7$, \mathbf{P} is 6×6 because attitude errors are 3×1 vectors (2), and this leads to some additional bookkeeping. There is nothing special to discuss about \mathbf{b} , which obeys common additive algebra, and thus this

subsection will be mostly devoted to applying the general formulation of subsection IIB to the quaternion parts of \mathbf{X} and \mathbf{P} .

The governing equations for the propagation of \mathbf{X} are:

$$\frac{d\mathbf{q}}{dt} = \frac{1}{2}\Omega(\omega_g - \hat{\mathbf{b}})\mathbf{q} + \boldsymbol{\eta}_v \quad (38)$$

and

$$\frac{d\hat{\mathbf{b}}}{dt} = \boldsymbol{\eta}_u \quad (39)$$

where $\boldsymbol{\eta}_v$ and $\boldsymbol{\eta}_u$ are white noise vectors, ω_g is the angular velocity measured by the gyro, and $\Omega(\omega)$ is the following 4×4 antisymmetric matrix:

$$\Omega(\omega) \equiv \begin{bmatrix} -[\omega \times] & \omega \\ -\omega^T & 0 \end{bmatrix}. \quad (40)$$

The Farrenkopf [12], [13] models for $\boldsymbol{\eta}_v$ and $\boldsymbol{\eta}_u$ are used here: both have zero mean, isotropic variances denoted by σ_v^2 and σ_u^2 respectively, and result in the following process noise covariance matrix (to be used in (17)):

$$\mathbf{Q}_{k-1} = \begin{bmatrix} \left(\sigma_v^2 \Delta t_k + \frac{1}{3} \sigma_u^2 \Delta t_k^3\right) \mathbf{I}_{3 \times 3} & -\frac{1}{2} (\sigma_u^2 \Delta t_k^2) \mathbf{I}_{3 \times 3} \\ -\frac{1}{2} (\sigma_u^2 \Delta t_k^2) \mathbf{I}_{3 \times 3} & (\sigma_u^2 \Delta t_k) \mathbf{I}_{3 \times 3} \end{bmatrix}, \quad (41)$$

where $\Delta t_k = t_k - t_{k-1}$ is the time step.

The state estimates are propagated using the following differential equations where $\hat{\omega}_g \equiv \omega_g - \hat{\mathbf{b}}$:

$$\frac{d\hat{\mathbf{q}}}{dt} = \frac{1}{2}\Omega(\hat{\omega}_g)\hat{\mathbf{q}}, \quad (42)$$

$$\frac{d\hat{\mathbf{b}}}{dt} = \mathbf{0}. \quad (43)$$

Although (42) and (43) are, in principle, coupled, the fact that $\hat{\mathbf{b}}$ is treated as a constant during propagation means we can integrate $\hat{\mathbf{q}}$ separately. We integrated (42) via the following integral [38] which is exact if ω is constant during the time interval Δt :

$$\mathbf{q}(t + \Delta t) = \left[\cos(\omega \Delta t / 2) \mathbf{I}_{4 \times 4} + \frac{\sin(\omega \Delta t / 2)}{\omega} \Omega(\omega) \right] \mathbf{q}(t), \quad (44)$$

where $\omega = |\omega|$.

The construction of the quaternion sigma points in (19) and (20) follows readily after recognizing that they are constructed from χ^0 by adding and subtracting corresponding deviations obtained from the columns of $(\sqrt{(L + \lambda)} \mathbf{P}_{k-1|k-1}^a)$, where $L = 12$. For the state vector of (37), (16) and (18) yield the mean sigma point as:

$$\chi_{k-1|k-1}^0 \equiv X_{k-1|k-1}^a = \begin{bmatrix} \hat{\mathbf{q}}_{k-1|k-1}^T & \hat{\mathbf{b}}_{k-1|k-1}^T & \mathbf{0}_{6 \times 1}^T \end{bmatrix}^T \quad (45)$$

To construct the i -th and $i + L$ -th sigma points during the propagation phase, it is convenient to identify $(\sqrt{(L + \lambda)} \mathbf{P}_{k-1|k-1}^a)_i$ as being of the form $[\Delta\theta_i^T \Delta\mathbf{b}_i^T \boldsymbol{\eta}_{\theta,i}^T \boldsymbol{\eta}_{b,i}^T]^T$ where the $\boldsymbol{\eta}$'s are noise terms originating principally

from \mathbf{Q}_{k-1} . Thus the quaternion part of χ^i and χ^{i+L} are constructed by perturbing χ^0 as follows:

$$\chi_{k-1|k-1}^i (1 : 4) = \chi_{k-1|k-1}^0 (1 : 4) \otimes \begin{bmatrix} \frac{\Delta\theta_i^T}{2} \mathbf{1} \end{bmatrix}^T, \quad i = 1, \dots, L \quad (46)$$

$$\chi_{k-1|k-1}^i (5 : 7) = \chi_{k-1|k-1}^0 (5 : 7) + \Delta\mathbf{b}_i^T, \quad i = 1, \dots, L \quad (47)$$

$$\chi_{k-1|k-1}^{i+L} (1 : 4) = \chi_{k-1|k-1}^0 (1 : 4) \otimes \begin{bmatrix} -\frac{\Delta\theta_i^T}{2} \mathbf{1} \end{bmatrix}^T, \quad i = 1, \dots, L, \quad (48)$$

$$\chi_{k-1|k-1}^{i+L} (5 : 7) = \chi_{k-1|k-1}^0 (5 : 7) - \Delta\mathbf{b}_i^T, \quad i = 1, \dots, L. \quad (49)$$

The quaternion parts of the χ should be normalized to unity after the computations in (46) and (48). (Alternately, the normalizations can be performed on the incremental quaternions before constructing the sigma points.) The addition of the process noise to the sigma points after propagation, analogous to (21), is similar:

$$\chi_{k|k-1}^i (1 : 4) = \chi_{k|k-1}^i (1 : 4) \otimes \begin{bmatrix} \frac{\boldsymbol{\eta}_{\theta,i}^T}{2} & 1 \end{bmatrix}^T, \quad i = 1, \dots, L \quad (50)$$

$$\chi_{k|k-1}^i (5 : 7) = \chi_{k|k-1}^i (5 : 7) + \Delta\boldsymbol{\eta}_{b,i}^T, \quad i = 1, \dots, L, \quad (51)$$

$$\chi_{k|k-1}^{i+L} (1 : 4) = \chi_{k|k-1}^{i+L} (1 : 4) \otimes \begin{bmatrix} -\frac{\boldsymbol{\eta}_{\theta,i}^T}{2} & 1 \end{bmatrix}^T, \quad i = 1, \dots, L, \quad (52)$$

$$\chi_{k|k-1}^{i+L} (5 : 7) = \chi_{k|k-1}^{i+L} (5 : 7) - \boldsymbol{\eta}_{b,i}^T, \quad i = 1, \dots, L. \quad (53)$$

(No process noise is added to χ^0 because of the assumption of its zero mean.) As before, the quaternion parts of χ should be normalized to unity after the above computations.

The scheme for constructing \mathbf{X} and \mathbf{P} from the sigma points after propagation between measurements uses the fact that, from (24) and (26), the state vector weights, $\{W_s\}$, sum to unity as follows:

$$\sum_{i=0}^{2L} W_s^i = W_s^0 + \sum_{i=1}^{2L} W_s^i = \frac{\lambda}{L + \lambda} + \frac{2L}{2(L + \lambda)} = 1. \quad (54)$$

Thus, the weighted averages in (22) can be written as the sum of the mean sigma point and the weighted deviations of the other sigma points from it as follows:

$$\begin{aligned} \hat{\mathbf{X}} &= \sum_{i=0}^{2L} W_s^i \chi^i = \sum_{i=0}^{2L} W_s^i \chi^0 + \sum_{i=1}^{2L} W_s^i (\chi^i - \chi^0) \\ &= \chi^0 + \sum_{i=1}^{2L} W_s^i (\chi^i - \chi^0) \\ &= \chi^0 + \sum_{i=1}^{2L} W_s^i \delta \chi^i. \end{aligned} \quad (55)$$

In analogy to the sigma point constructions in (46), and using (2) and (8) which feature delta quaternions, the $\delta\chi^i$ corresponding to the quaternion parts of (55) can be identified as angular deviations, $\delta\theta^i$, via:

$$\frac{\delta\theta_i}{2} = \text{first three elements of } (\chi^{0^{-1}}(1:4) \otimes \chi^i(1:4)). \quad (56)$$

Defining a weighted- average angular deviation vector, $\langle \delta\theta \rangle$, via:

$$\langle \delta\theta \rangle = \sum_{i=1}^{2L} W_s^i \delta\theta_i, \quad (57)$$

allows the extension of (55) to quaternions via:

$$\hat{X}(1:4) = \hat{q} = \frac{\chi^0(1:4) \otimes \left[\frac{\langle \delta\theta^T \rangle}{2} 1 \right]^T}{\left| \chi^0(1:4) \otimes \left[\frac{\langle \delta\theta^T \rangle}{2} 1 \right]^T \right|^T}, \quad (58)$$

where normalization of the mean quaternion is explicitly indicated.

The post-propagation covariance matrix in (23) also can be computed in a similar way using 6×1 error vectors $\{\mathbf{x}^i\}$ (note the lower case used for the error vector) whose first three elements are small angular differences of the quaternion parts of $\{\chi^i\}$ from \hat{X} , the gyro bias elements being simple differences. That is,

$$\mathbf{x}_i = \begin{bmatrix} 2 \times \text{first three elements of } \hat{X}^{-1}(1:4) \otimes \chi^i(1:4) \\ \chi^i(5:7) - \hat{X}(5:7) \end{bmatrix}. \quad (59)$$

(59) is to be used in the covariance matrix calculation in (23), yielding:

$$\mathbf{P} = \sum_{i=0}^{2L} W_c^i \mathbf{x}_i \mathbf{x}_i^T. \quad (60)$$

(28) - (36) are applied to geomagnetic flux density measurements as follows, with the measured and modeled reference vectors denoted by \mathbf{B}_B and \mathbf{B}_I respectively. Here $m = 3$ and results in $L' = 9$. Predicted measurements are generated by using attitude matrices constructed from inserting the mean propagated quaternion into (4) which are then used to project \mathbf{B}_I into the estimated Body frame. The measurement model function, \mathbf{h} , thus involves the estimated attitude matrix and the \mathbf{B}_I at a given instant, and the predictions (with noise) are thereby generated via the following explicit version of (30):

$$\mathbf{B}_B^i = \mathbf{A}(\chi^i(1:4)) \mathbf{B}_I + \chi^i(8:10), \quad i = 0, 1, \dots, 2L', \quad (61)$$

where the subscript on χ^i has been dropped for clarity. (31) for the mean predicted measurement becomes:

$$\hat{\mathbf{B}}_B = \sum_{i=0}^{2L'} W_s^i \mathbf{B}_B^i, \quad (62)$$

and (32) now becomes:

$$\mathbf{P}_{zz} = \sum_{i=0}^{2L'} W_c^i [\mathbf{B}_B^i - \hat{\mathbf{B}}_B] [\mathbf{B}_B^i - \hat{\mathbf{B}}_B]^T. \quad (63)$$

The error vectors defined via (59) should be used in the first bracket on the right hand side in the cross-covariance equation, (33):

$$\mathbf{P}_{xz} = \sum_{i=0}^{2L'} W_c^i \mathbf{x}^i [\mathbf{B}_B^i - \hat{\mathbf{B}}_B]^T. \quad (64)$$

The state vector update of (35) is done in two parts, separately for the quaternion and the gyro bias because they follow different algebras. Introduce \mathbf{x}^* as the state update via:

$$\mathbf{x}^* = [\delta\theta^{*T} \quad \delta\mathbf{b}^{*T}]^T = \mathbf{K}_k (\mathbf{B}_B - \hat{\mathbf{B}}_B), \quad (65)$$

where \mathbf{B}_B is the measured field. Let $\hat{\mathbf{q}}_{k|k-1}$ and $\hat{\mathbf{b}}_{k|k-1}$ be the pre-update quaternion and gyro bias vector respectively, with $\hat{\mathbf{q}}_{k|k}$ and $\hat{\mathbf{b}}_{k|k}$ denoting the post-update quantities. $\delta\theta^*$ can be used to generate an incremental quaternion which updates the quaternion via:

$$\delta\mathbf{q}^* = \left[\left(\delta\theta^{*T}/2 \right) 1 \right]^T \quad (66)$$

and

$$\hat{\mathbf{q}}_{k|k} = \frac{\hat{\mathbf{q}}_{k|k-1} \otimes \delta\mathbf{q}^*}{|\hat{\mathbf{q}}_{k|k-1} \otimes \delta\mathbf{q}^*|}, \quad (67)$$

showing explicitly the normalization of the updated quaternion. The gyro bias vector is updated via:

$$\hat{\mathbf{b}}_{k|k} = \hat{\mathbf{b}}_{k|k-1} + \delta\mathbf{b}^*. \quad (68)$$

III. RESULTS USING SIMULATED TRMM DATA

A. SPACECRAFT SCENARIO AND DATA CHARACTERISTICS

Ten orbits of simulated TRMM data were generated with the following characteristics: data period of 10 s, altitude of 350 km, zero eccentricity, and inclination of 35 deg. The orbit period, τ , was 5492.3 s. A 10th order International Geomagnetic Reference Field [39] was generated in the Earth Centered Inertial frame at each data point in the 10 orbits, and the corresponding true Body frame components were generated using an Earth-pointing instantaneous Orbital frame which is also the nominal Body frame. In this system, the nadir vector is the Z axis, the negative orbit normal is the Y axis, and the forward (velocity) direction of this circular orbit is the X axis.

True gyro data were generated at the magnetic flux density measurement times, consisting of a constant pitch rate of negative one rotation per orbit which amounts to approximately -0.065546 deg/s. Zero-mean magnetometer noise of standard deviation (σ_B) 50 nT was added to the clean magnetometer “measurements”. The gyro data were corrupted with noise characteristics (see (38) and (39)) of $\sigma_u = 3.1623 \times 10^{-10}$ rad/s^{3/2} and $\sigma_v = 0.31623 \times 10^{-6}$ rad/s^{1/2}.

B. UKF PERFORMANCE USING MONTE CARLO DATA

There is some latitude in the numerical values of the UKF parameters in (24) – (27). One option was to use the following values suggested in [9]: $\alpha = 0.001$, $\beta = 2$, and $\kappa = 1$, which yielded $\lambda \approx -L$. But the value $\lambda = 1$ used in [18] appears to yield faster convergence and hence the following values were used in all of the results quoted here:

$$\alpha = \sqrt{3}, \quad \beta = 2, \quad \kappa = 1, \quad \lambda = 1. \quad (69)$$

This value of λ overrides (27) and thus the actual value of κ is immaterial.

We kept the initial state the same in every run in every batch with every Kalman filter. The initial attitude estimate was always the identity quaternion: $\hat{\mathbf{q}}_0 = [0\ 0\ 0\ 1]^T$, which invariably resulted in large initial root-sum-square (RSS) attitude errors. (If $\delta\theta \equiv [\delta\theta_x, \delta\theta_y, \delta\theta_z]^T$, the RSS value of $\delta\theta$ is $\sqrt{\delta\theta_x^2 + \delta\theta_y^2 + \delta\theta_z^2}$.) The initial bias estimate was always the zero vector: $\hat{\mathbf{b}}_0 = [0\ 0\ 0]^T$. A third quantity, the initial attitude covariance on each axis, P_{0a} , was kept similarly invariant throughout this study: it was always set to 180^2 deg^2 corresponding to a complete lack of attitude knowledge.

The effects of two gyro bias-related quantities were studied extensively. One was denoted $b_{0,\text{true}}$, which was the standard deviation of a zero-mean normal distribution from which the true initial gyro bias on each axis was drawn. We call this the “scale” of the true gyro bias, and set it between 0.1 deg/hr and 5 deg/hr depending on the Monte Carlo batch. The other quantity of importance was the gyro bias covariance on each axis, P_{0b} , which we set between $0.1^2 \text{ deg}^2/\text{hr}^2$ and $20 \text{ deg}^2/\text{hr}^2$, the same for every run in a given Monte Carlo batch. These are described on a case by case basis below. The “convergence time” was defined as the time since the start of the run when the RSS attitude errors first dropped below 0.1 deg.

A sizeable majority of the runs converged rapidly, in less than $\tau/2$ (≈ 2750 s). Various aspects of a typical convergent run are shown in Figs. 2 - 7, the results here being generated with $b_{0,\text{true}} = 0.1 \text{ deg/hr}$ and $P_{0b} = 1 \text{ deg}^2/\text{hr}^2$.

Figs. 2 and 3 respectively show the convergence of the state vector to the true values. Fig. 2 shows the attitude parameterized as {roll, pitch, yaw} angles which were defined as 1-2-3 Euler angles of the Body frame with respect to the Orbital frame defined earlier; thus the nominal values of {roll, pitch, yaw} are {0, 0, 0}. We see in Fig. 1 that all three angles converge to the zero nominal values starting from large *a priori* errors. (The scale of the top subplot in Fig. 2 is different from the other two to show the large initial errors.) Fig. 3 shows that the gyro biases also converged to the truth despite significant initial errors. Note that the true gyro bias in any run is not a constant in time because it undergoes a random walk. This is evident from Fig. 3.

Fig. 4 shows the magnetometer residuals ($\mathbf{B}_B - \hat{\mathbf{B}}_B$, on the right hand side of (65)) and also provides statistics for each component after convergence (i. e., during the time span 10000 s – 40000 s). The accuracy of the convergence

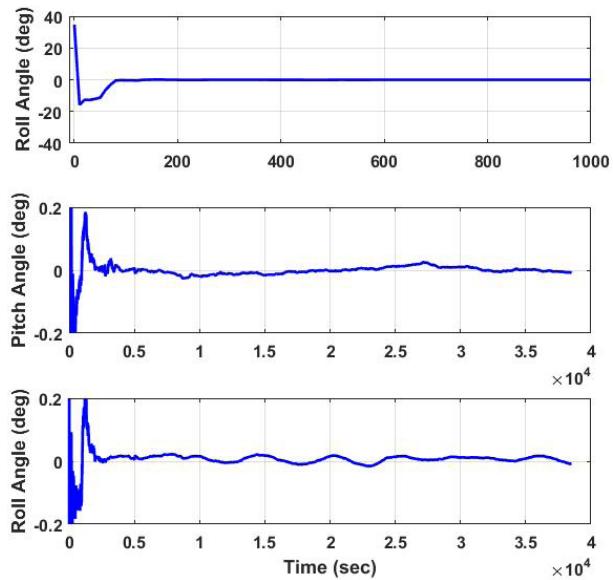


FIGURE 2. Convergence of roll, pitch, and yaw angles during a typical run.

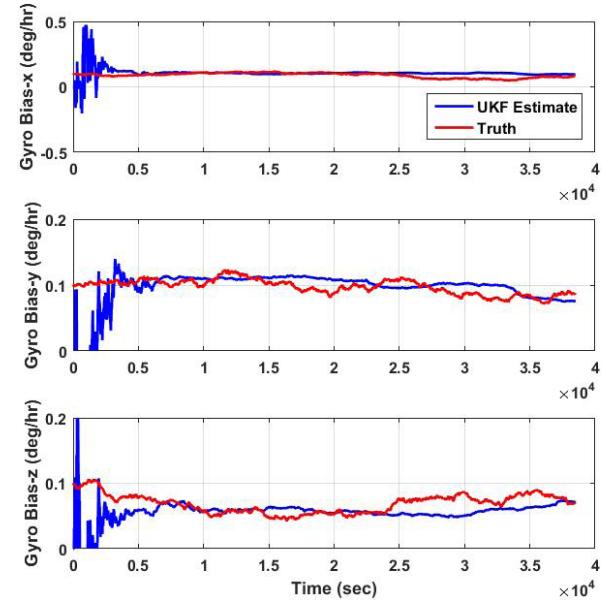


FIGURE 3. Convergence of gyro biases during the run of Fig. 2.

during this run is corroborated by the statistics of the residuals: the means and standard deviations of the residuals are close to those of the noise in the simulated data (mean = 0 nT and $\sigma_B = 50$ nT).

Fig. 5 shows the evolution of the attitude covariances. Of particular interest here are the periodic peaks recurring with the orbit period of about 5492.3 s. These peaks are characteristic of magnetometers and reflect the fact that at any given moment the rotation angle of the Body frame about the magnetic field vector is unobservable in a magnetometer-only situation thus leading to a rise in the corresponding covariance.

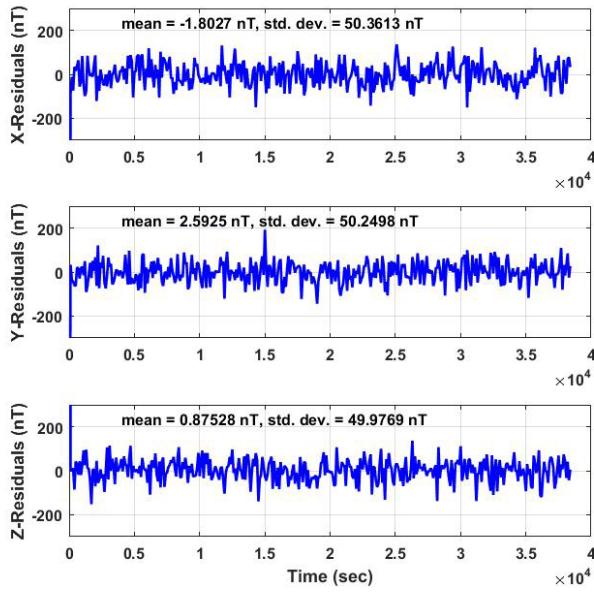


FIGURE 4. Time dependence of the magnetic flux density residuals during the run of Fig. 2.

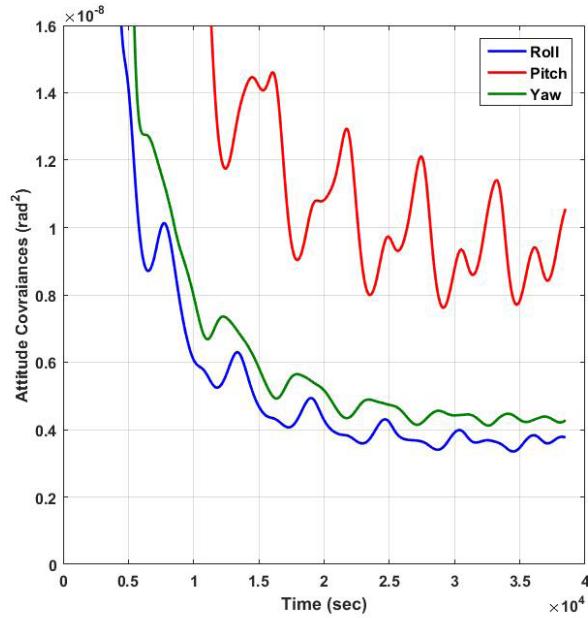


FIGURE 5. Time dependence of the attitude covariances during the run of Fig. 2. The variations of each component recur at the orbit period and are due to variations in observability as the spacecraft orbits the Earth.

Fig. 6 shows the evolution of the RSS errors, which are convenient scalars denoting the magnitude of the errors. The RSS attitude error drops to 0.1 deg in about 1800 s (“convergence time”) in this run, which is less than $\tau/3$, despite starting from an initial RSS attitude and gyro bias errors of about 77 deg and 0.17 deg/hr respectively. Corresponding individual axis errors during this run are shown in Fig. 7, where we see that the true errors fall within the 3σ bounds given by the UKF’s covariance matrix.

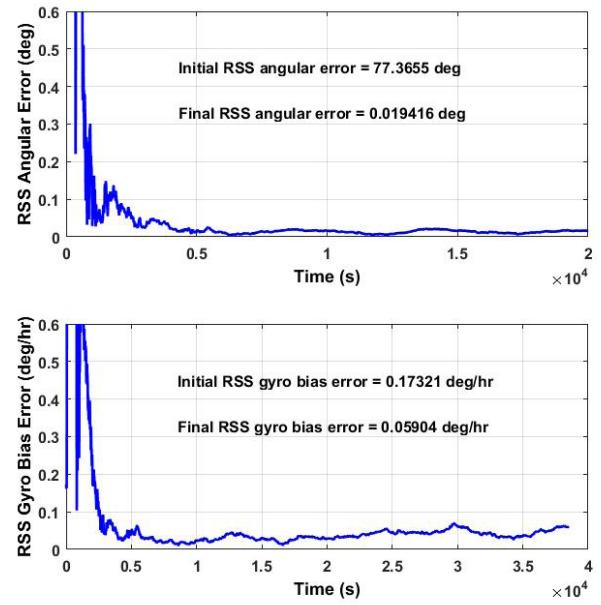


FIGURE 6. Time dependence of RSS errors during the run of Fig. 2 showing accurate convergence to the truth in spite of large initial errors.

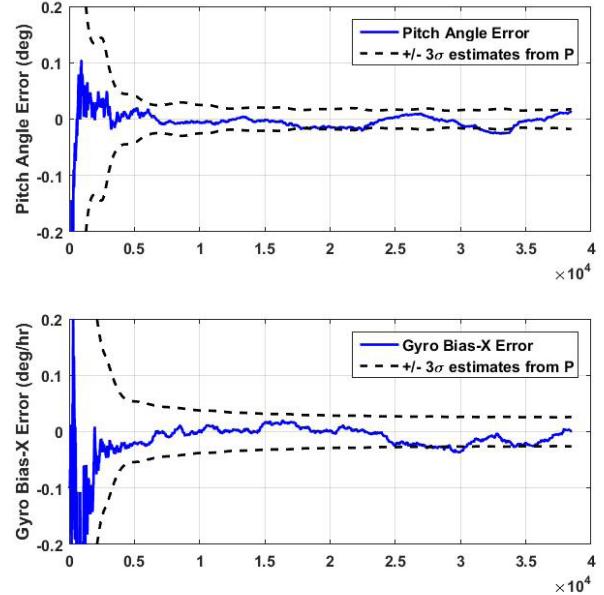


FIGURE 7. Comparison of single axis errors and covariances during the run of Fig. 2 showing that the covariances are commensurate with the true errors.

(For clarity, the errors corresponding to only one of the axes is chosen in each subplot.)

The above figures show results from one run out of a batch of 1000 Monte Carlo runs. To evaluate the filter in a comprehensive set of initial conditions, the starting point of each run were chosen randomly during the first three orbits, the true gyro biases were set randomly using $b_{0,true}$, and data were collected for each run for 38500 s (≈ 7 orbits). The convergence times were binned in half-orbit-period intervals

and the cumulative distribution function (CDF) of the data for this batch are plotted in Fig. 8.

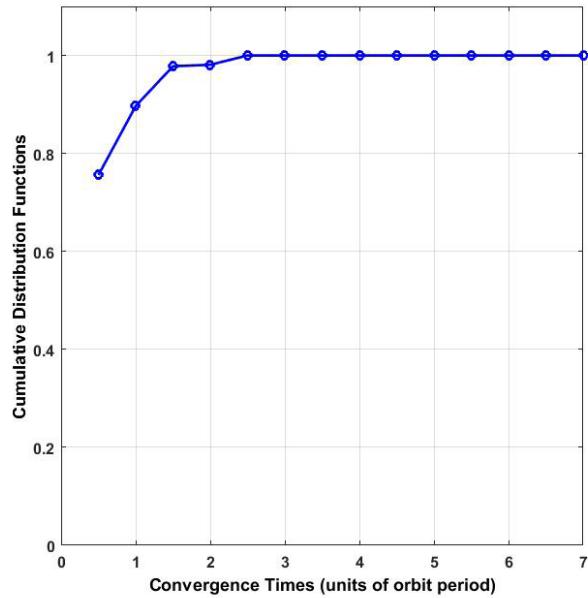


FIGURE 8. CDF of the convergence times of the UKF for a batch of 1000 runs with $b_{0,true} = 0.1 \text{ deg/hr}$ and $P_{0b} = 1 \text{ deg}^2/\text{hr}^2$.

Fig. 8 indicates that over 75% of the runs converged to 0.1 deg within 0.5 orbit (2750 s), about 90% converged within 1 orbit, and 100% converged within 2.5 orbits. The above results show that the UKF converges to the truth rapidly when $b_{0,true}$ was small and P_{0b} was commensurate with the true uncertainty.

C. EFFECTS OF INITIAL GYRO BIAS ERRORS AND COVARIANCES

The UKF sometimes *diverged* when P_{0b} was unreasonably large. For example, no divergence was seen when the run of Figs. 2 - 7 was repeated with $b_{0,true} = 0.1 \text{ deg/hr}$ as before but $P_{0b} = 20^2 \text{ deg}^2/\text{hr}^2$. A marked divergence is seen in a run starting at a different point in the first orbit. This divergence is shown in Fig. 9, with extremely large errors of over 100 deg and 400 deg/hr persist even after about 7 orbits.

Thus, somewhat unexpectedly, we find that it is possible for a UKF to diverge, a fact also noted by Perea *et. al.* [40], who conjectured that this could be a problem intrinsic to the UKF algorithm whereby the covariance reduction does not correspond to the information obtained from the measurements. One remedy suggested by them was to increase the noise covariance (\mathbf{R} in (15)). This can indeed change the convergence as can be seen in Fig. 10 when the standard deviation of the measurement noise is increased from the simulation value of 50 nT to the unreasonably large value of 500 nT. Fig. 10 also shows that the covariances are then in keeping with the true errors.

We studied the effect of P_{0b} on the convergent times of the UKF and found that it affected convergence profoundly.

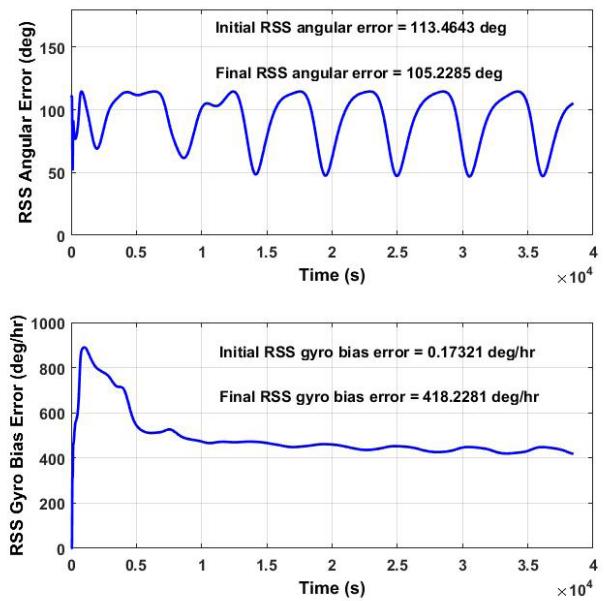


FIGURE 9. Divergence during a run when $b_{0,true} = 0.1 \text{ deg/hr}$ and P_{0b} was set at the unrealistic value of $20^2 \text{ deg}^2/\text{hr}^2$.

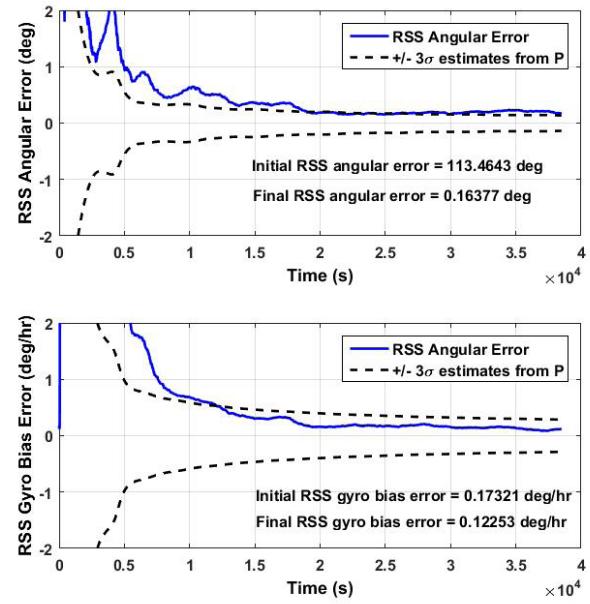


FIGURE 10. Restoration of convergence in the run of Fig. 9 by using large measurement uncertainties.

This is summarized by the systematic dependence of the CDFs on P_{0b} in Fig. 11. Fig 11 shows, in addition to the curve from Fig. 8, four other curves which are results from additional 1000-run batches with different combinations of $b_{0,true}$ and P_{0b} . We see that the probability of convergence drops significantly when $b_{0,true}$ is kept constant at 0.1 deg/hr and P_{0b} is increased unrealistically larger through $10 \text{ deg}^2/\text{hr}^2$ (red) and $20 \text{ deg}^2/\text{hr}^2$ (green). The last curve is masked by the magenta curve which also denotes P_{0b} at $20 \text{ deg}^2/\text{hr}^2$ but with a value of 5 deg/hr for $b_{0,true}$, thus showing that it was not

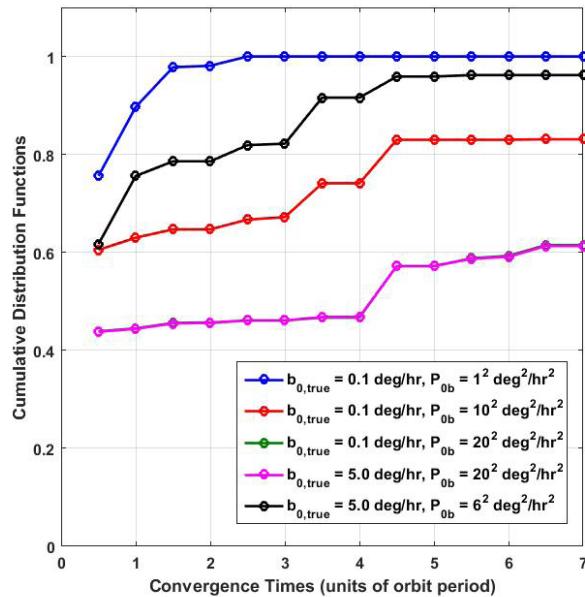


FIGURE 11. Effects of the scale of the initial true gyro bias, $b_{0,\text{true}}$, and initial gyro bias covariance, P_{0b} , on the CDFs of convergence times of the small-angle-based UKF using batches of 1000 Monte Carlo runs. The performance worsens as P_{0b} progressively deviates from $b_{0,\text{true}}$. (The two plots for the P_{0b} value of $20^2 \text{ deg}^2/\text{hr}^2$ are indistinguishable on the scale of the figure.)

the larger bias but the unrealistically large covariance which reduced the convergence from 100% within 2.5 orbits to barely 60% even after 7 orbits. Dramatically, the convergence once again rises when $b_{0,\text{true}}$ is fixed at 5 deg/hr but P_{0b} is reduced to 6 deg^2/hr^2 .

D. COMPARISON WITH EKF USING SMALL ANGLE APPROXIMATIONS

We compare the UKF here with an EKF which also uses the small angle approximations in treating attitude errors. The theory is cursorily listed here since it is discussed amply in [15]. One notable point is that we used the continuous-time form of EKF.

The state vector and its propagation equations are given by (37), (42), and (43) and much of the notation is analogous to that of Section II. The 6×1 state error vector, x , with small angles for the attitude errors obeys the following equation, which uses the white noise vectors in (38)-(39):

$$\frac{dx}{dt} = Fx + \begin{bmatrix} \eta_v \\ \eta_u \end{bmatrix}, \quad (70)$$

where

$$F \equiv \begin{bmatrix} -[\omega \times] & I_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} \end{bmatrix}. \quad (71)$$

The covariance matrix evolves through the following differential equation:

$$\frac{dP}{dt} = FP + PF^T + Q, \quad (72)$$

where Q is the spectral density matrix generating the noise covariance matrix of (41). (72) was integrated numerically using the *ODE45* function in MATLAB,¹ which is an adaptive step-size Runge-Kutta (4, 5) method. The magnetic flux density residuals, y , are given by differencing the measurements and the predictions using the propagated attitude:

$$y \equiv B_B - \hat{B}_B \equiv B_B - A(\hat{q})B_I. \quad (73)$$

The observation matrix, H , is given by:

$$H = \begin{bmatrix} -[\hat{B}_B \times] & 0_{3 \times 3} \end{bmatrix} \quad (74)$$

and is used along with the covariance propagated via (72), to compute the Kalman gain, K :

$$K = PH^T \left[HPH^T + R \right]^{-1}. \quad (75)$$

(65)-(68) are then used to update the state vector.

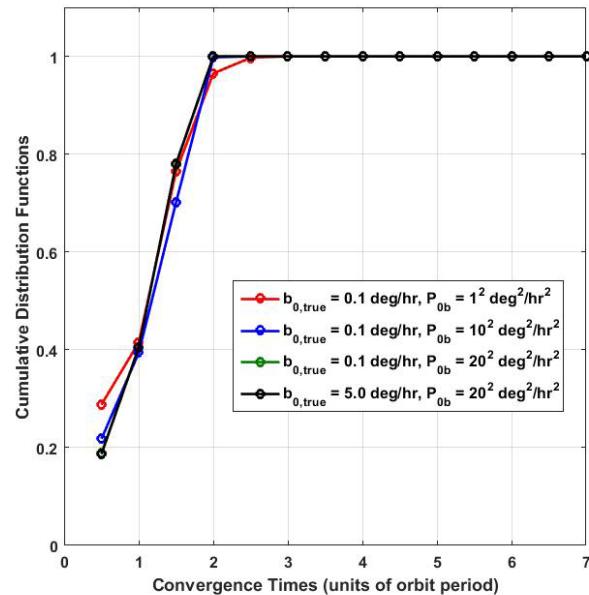


FIGURE 12. Effects of the scale of the initial true gyro bias, $b_{0,\text{true}}$, and initial gyro bias covariance, P_{0b} , on the CDFs of convergence times of the small-angle-based EKF using batches of 1000 Monte Carlo runs. In contrast to the UKF (Fig. 11), the EKF is remarkably robust.

The CDFs of the convergence times using the EKF are presented in Fig. 12. Three of the plots here (blue, red, and green) correspond to an initial true bias of 0.1 deg/hr but different values of P_{0b} . The last plot (black) shows results for initial true bias of 5.0 deg/hr and $P_{0b} = 20^2 \text{ deg}^2/\text{hr}^2$. We notice three striking features when these results are compared with the UKF results in Fig. 11: (1) Neither $b_{0,\text{true}}$ nor P_{0b} have little effect – all of the plots are practically identical and we do not have to worry about the reasonableness of P_{0b} ; (2) Far fewer EKF runs converge in 0.5, 1.0, or 1.5 orbits; (3) 100% of the EKF runs converged

¹ MATLAB is a software product of The Mathworks, Inc., Natick, MA, USA <http://www.mathworks.com/>

in 2.0 orbits. It appears that if we start the EKF with large P_{0b} , the EKF does take a while to converge, but the convergence is swift after that even when $b_{0,true}$ was 5 deg/hr. Thus, the UKF seems to converge faster provided the initial convergence is realistic. The EKF always converged, albeit slower, which is convenient as a practical matter since this robustness ensures that P_{0b} does not have to be tuned. Looking at all the plots in Fig. 12, we conclude that 100% of the EKF runs converged in 3 orbits.

E. COMPARISON WITH UKF USING GENERALIZED RODRIGUES PARAMETERS

We also examined the parameterization of angular errors via GRPs as suggested by Crassidis and Markley [18]. Given a delta quaternion, $\delta\mathbf{q} \equiv [\delta\rho^T \delta q_4]^T$, where $\delta\rho$ is the vector of the first three components of $\delta\mathbf{q}$, the corresponding vector of GRPs, $\delta\mathbf{p}$, is given by the following formula which involves the free parameters a and f :

$$\delta\mathbf{p} = \frac{f\delta\rho}{a + \delta q_4}. \quad (76)$$

The inverse process of constructing $\delta\mathbf{q} \equiv [\delta\rho^T \delta q_4]^T$ given $\delta\mathbf{p}$ is as follows [15]:

$$\begin{aligned} \delta q_4 &= \frac{-a|\delta\mathbf{p}|^2 + f\sqrt{f^2 + (1-a^2)|\delta\mathbf{p}|^2}}{f^2 + |\delta\mathbf{p}|^2} \\ \delta\rho &= \frac{(a + \delta q_4)\delta\mathbf{p}}{f}. \end{aligned} \quad (77)$$

Crassidis and Markley found $a = 1$ and $f = 4$ to be optimal values for the two free parameters, which then, by comparing with (2), leads to the identification that $\delta\mathbf{p} \rightarrow \delta\theta$ for $|\delta\rho| \rightarrow 0$.

We constructed a second version of our UKF which used (76) and (77) with $a = 1$ and $f = 4$ and input the simulated data with the identical conditions used by the small-angles version. To this end, several 1000-run batches were executed with the GRPs version and using identical Monte Carlo conditions used in the small angles version. We emphasize that the same starting orbital points, *a priori* values, and random number streams were input to the two versions, the only difference being the parameterization of attitude errors. The CDFs for the GRPs version are presented in Fig. 13 for several initial bias parameters similar to those in Fig. 11. It is clear that the small angles approximation performs a little better than the GRPs version. For example, when $b_{0,true} = 0.1$ deg/hr and $P_{0b} = 1$ deg²/hr², 100% of the runs in Fig. 11 converged in 2.5 orbits in contrast to 93.1% of runs here. In fact, nearly 2% of the GRP-based runs did not converge even after seven orbits for these conditions. The performance differences are amplified when $b_{0,true} = 5.0$ deg/hr and $P_{0b} = 6^2$ deg²/hr² (black): after 4.5 orbits, the small angles version saturates at about 96% convergence whereas the GRPs version does so at 87%.

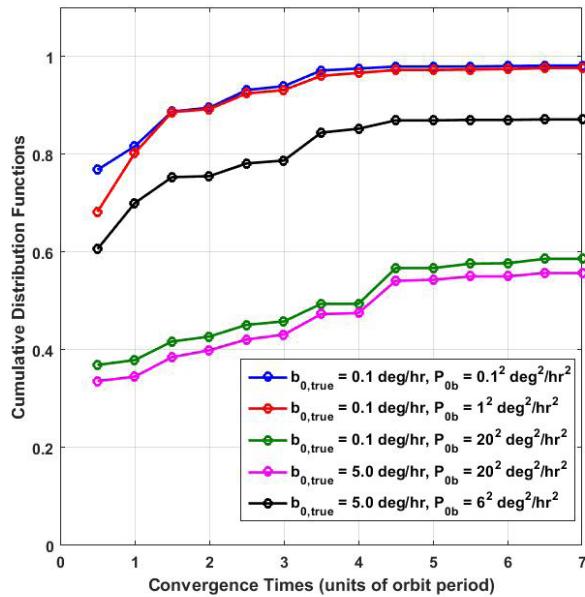


FIGURE 13. Effects of the scale of the initial true gyro bias, $b_{0,true}$, and initial gyro bias covariance, P_{0b} , on the CDFs of convergence times of the GRP-based UKF using batches of 1000 Monte Carlo runs. The performance is similar but slightly inferior to that of the small-angle-based UKF (Fig. 11).

IV. CONCLUSIONS

The theory for a simple quaternion-based attitude UKF using small angle approximations for attitude errors and explicitly using process and measurement noise in constructing the sigma points was presented, and applied to a state vector consisting of the quaternion and gyro biases. This UKF was evaluated using simulated TRMM spacecraft data in a magnetometer-only spacecraft scenario involving two crucial aspects of methodology: evaluation at different starting points in an orbit thereby facilitating the variation of the initial attitude observability, and using a large number of such starting points to obtain accurate statistics. The UKF was evaluated using several 1000-run Monte Carlo batches under different initial orbit locations, different initial gyro biases, no *a priori* knowledge (by always making the initial state an identity quaternion and a null gyro bias vector), large initial attitude covariance, and several different initial gyro bias covariances. An important aspect of this study is the vast amount of data – a total of 14,000 Monte Carlo runs – which allows some clear inferences.

The UKF was found to be robust, accurate, and rapidly convergent when the true bias and initial gyro bias uncertainty were small (less than 1 deg/hr): it converged to RSS attitude errors of 0.1 deg within 1/2 an orbit in over 75% of the time and within 2.5 orbits in 100% of the time. However, a significant finding here is that the performance depends strongly on the initial gyro bias covariance input to the UKF: nearly 40% of the runs did not converge when this initial gyro bias uncertainty was much larger than the true initial gyro bias errors. The performance improved significantly if this initial

uncertainty was comparable to the true bias errors even when the latter were as large as 5 deg/hr.

The performance of this UKF was compared to two other filters - an EKF which also uses the small angles parameterization of attitude errors, and another UKF which uses generalized Rodrigues parameters to treat attitude errors. The small-angle-based UKF converged faster than the small-angle-based EKF when the true bias and initial gyro bias uncertainty were less than 1 deg/hr. But, in general, the small-angle-based UKF performed poorly against the EKF, which converged to the same accuracies in 100% of several thousand Monte Carlo runs within 3.0 orbits. However, the small-angle-based UKF performed consistently better than the UKF using generalized Rodrigues parameters in identical conditions.

It is worth emphasizing that the superior robustness of the EKF, without sacrificing accuracy, was another significant finding of this study. This is contrary to popular belief that EKFs are more susceptible to divergences, which was in fact a motivation for the Unscented Transform. Thus there is now ample evidence from three different quaternion-based Kalman filters that the small angles approximation can be very robust – in the gyro-based UKF and EKF of this paper as well as a gyro-less EKF extensively studied in the past [30], [34]. Another important aspect of the robustness of the EKF is that it was *never re-tuned* irrespective of the run or batch. Such robustness obviously has important practical applications; for example, there is no need to constantly monitor the EKF or re-tune it when it has to be restarted for a variety of reasons.

The results here immediately point towards a couple of important future directions of study with UKFs. First, since some of the conclusions here may be specific to the magnetometer-only scenario where the attitude at any instant is not fully observable, it would be useful to examine if full attitude observability at every instant via, for example, simultaneous Sun sensor measurements, prevents divergences. Second, the finding that UKF performance can depend significantly on initial conditions will be of interest not just in spacecraft but also in a variety of autonomous inertial navigational systems such as those in unmanned air vehicles and terrestrial robots. In such systems, onboard inertial measurement units comprising of gyros and accelerometers are ubiquitous and the state vector often involves calibration parameters of these units during extended missions. It would thus be useful to conduct similar studies of attitude UKFs in such systems as well to determine the robustness against initial conditions and partial attitude observability.

REFERENCES

- [1] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960.
- [2] R. E. Kalman and R. S. Bucy, “New results in linear filtering and prediction theory,” *J. Basic Eng. D*, vol. 83, no. 1, pp. 95–108, 1961.
- [3] A. Gelb, Ed., *Applied Optimal Estimation*. Cambridge, MA, USA: MIT Press, 1974.
- [4] G. L. Smith, S. F. Schmidt, and L. A. McGee, “Application of statistical filter theory to the optimal estimation of position and velocity on board a circumlunar vehicle,” NASA, Washington, DC, USA, Tech. Rep. R-135, 1962.
- [5] B. A. McElhoe, “An assessment of the navigation and course corrections for a manned flyby of mars or venus,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-2, no. 4, pp. 613–623, Jul. 1966.
- [6] T. Lefebvre, H. Bruyninckx, and J. D. Schutter, “Kalman filters for nonlinear systems: A comparison of performance,” *Int. J. Control.*, vol. 77, pp. 639–653, 2004.
- [7] S. J. Julier and J. K. Uhlmann, “A new extension of the Kalman filter to nonlinear systems,” *Proc. SPIE, Signal Process., Sensor Fusion, Target Recognit. VI*, vol. 3068, pp. 182–193, Jul. 1997. doi: 10.1117/12.280797.
- [8] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, “A new method for the nonlinear transformation of means and covariances in filters and estimators,” *IEEE Trans. Autom. Control*, vol. AC-45, no. 3, pp. 477–482, Mar. 2000.
- [9] E. A. Wan and R. van der Merwe, “The unscented Kalman filter for nonlinear estimation,” in *Proc. IEEE Adapt. Syst. Signal Process., Commun., Control Symp. (AS-SPCC)*, Lake Louise, AB, Canada, Oct. 2000, pp. 153–158.
- [10] D. H. Titterton and J. L. Weston, *Strapdown Inertial Navigation Technology* (Progress in Astronautics and Aeronautics), vol. 207, 2nd ed. Washington, DC, USA: AIAA, 2004, p. 345.
- [11] N. F. Toda, J. L. Heiss, and F. H. Schlee, “SPARS: The system, algorithm, and test results,” in *Proc. Symp. Spacecraft Attitude Determination*, vol. 1, 1969, pp. 361–370.
- [12] R. L. Farrenkopf, “Generalized results for precision attitude reference systems using gyros,” in *Proc. AIAA Mech. Control Flight Conf.*, Anaheim, CA, USA, Aug. 1974, paper 74-903.
- [13] R. L. Farrenkopf, “Analytic steady-state accuracy solutions for two common spacecraft attitude estimators,” *J. Guid., Control, Dyn.*, vol. 1, no. 4, pp. 282–284, 1978.
- [14] J. W. Murrell, “Precision attitude determination for multimission spacecraft,” in *Proc. AIAA Guid. Control Conf.*, New York, NY, USA, 1978, pp. 70–87, doi: 10.2514/6.1978-1248.
- [15] E. J. Lefferts, F. L. Markley, and M. D. Shuster, “Kalman filtering for spacecraft attitude estimation,” *J. Guid., Control, Dyn.*, vol. 5, no. 5, pp. 417–429, 1982.
- [16] E. Kraft, “A quaternion-based unscented Kalman filter for orientation tracking,” in *Proc. 6th Int. Conf. Inf. Fusion*, Cairns, Australia, Jul. 2003, pp. 47–54.
- [17] W. Khoder, B. Fassinut-Mombot, and M. Benjelloun, “Quaternion unscented Kalman filtering for integrated inertial navigation and GPS,” in *Proc. 11th Int. Conf. Inf. Fusion*, Jun./Jul. 2008, pp. 1–8.
- [18] J. L. Crassidis and F. L. Markley, “Unscented filtering for spacecraft attitude estimation,” *J. Guid., Control, Dyn.*, vol. 26, no. 4, pp. 536–542, Jul. 2003.
- [19] K.-L. Lai, J. L. Crassidis, and R. R. Harman, “In-space spacecraft alignment calibration using the unscented filter,” in *Proc. AIAA Guid., Navigat., Control Conf. Exhibit*, Austin, TX, USA, Aug. 2003, pp. 1557–1567, paper AIAA 2003-5563, doi: 10.2514/6.2003-5563.
- [20] M. C. VanDyke, J. L. Schwartz, and C. D. Hall, “Unscented Kalman filtering for spacecraft attitude state and parameter estimation,” in *Proc. AAS/AIAA Space Flight Mech. Conf.*, 2004, paper AAS 04-115, pp. 1–13.
- [21] M. Alarfaj and F. Rogers-Marcovitz, “Interactions of pose estimation and online dynamic modeling for a small inspector spacecraft,” in *Proc. Small Satellite Syst. Services Symp.*, Madeira, Portugal, 2010, pp. 1–7.
- [22] H. E. Soken and S.-I. Sakai, “Residual based adaptive unscented Kalman filter for satellite attitude estimation,” in *Proc. AIAA Guid., Navigat., Control Conf.*, Minneapolis, MN, USA, Aug. 2012, paper AIAA 2012-4476.
- [23] M. Rhudy, Y. Gu, J. Gross, and M. R. Napolitano, “Evaluation of matrix square root operations for UKF within a UAV GPS/INS sensor fusion application,” *Int. J. Navigat. Obs.*, vol. 2011, 2011, Art. no. 416828, doi: 10.1155/2011/416828.
- [24] Y.-J. Cheon and J.-H. Kim, “Unscented filtering in a unit quaternion space for spacecraft attitude estimation,” in *Proc. IEEE Int. Symp. Ind. Electron.*, Jun. 2007, pp. 66–71.
- [25] H. Goldstein, *Classical Mechanics*, 2nd ed. Reading, MA, USA: Addison-Wesley, 1980, ch. 4.
- [26] J. R. Wertz, Ed., *Spacecraft Attitude Determination and Control*. Boston, MA, USA: Reidel, 1978.

- [27] M. D. Shuster, "A survey of attitude representations," *J. Astron. Sci.*, vol. 41, no. 4, pp. 439–517, Oct./Nov. 1993.
- [28] TRMM, accessed on Apr. 28, 2016. [Online]. Available: <http://trmm.gsfc.nasa.gov/>
- [29] M. L. Psiaki, F. Martel, and P. K. Pal, "Three-axis attitude determination via Kalman filtering of magnetometer data," *J. Guid. Control Dyn.*, vol. 13, no. 3, pp. 506–514, 1990.
- [30] M. S. Challa, G. A. Natanson, D. F. Baker, and J. K. Deutschmann, "Advantages of estimating rate corrections during dynamic propagation of spacecraft rates: Applications to real-time attitude determination of SAMPEX," in *Proc. Flight Mech./Est. Theory Symp.*, 1994, pp. 481–495.
- [31] Q. Lam, P. Pal, R. Welch, and W. Grossman, "Magnetometer based attitude determination system using a reduced order Kalman filter," in *Proc. AIAA/AAS Astrodyn. Conf.*, San Diego, CA, USA, Jul. 1996, pp. 535–543, paper 96–3629.
- [32] J. Hashmall and J. Sedlak, "The use of magnetometers for accurate attitude determination," in *Proc. 12th Int. Symp. Space Flight Dyn.*, Darmstadt, Germany, 1997, p. 179.
- [33] J. Deutschmann, R. Harman, and I. Y. Bar-Itzhack, "An innovative method for low cost, autonomous navigation for low Earth orbit satellites," in *Proc. AIAA/AAS Astrodyn. Specialist Conf. Exhibit*, Boston, MA, USA, 1998, paper AIAA-98-4183.
- [34] M. Challa, G. Natanson, and N. Ottenstein, "Magnetometer-only attitude and rate estimates for spinning spacecraft," in *Proc. AIAA/AAS Astrodyn. Specialist Conf.*, Denver, CO, USA, Aug. 2000, pp. 311–321, paper AIAA 2000-4241, doi: 10.2514/6.2000-4241.
- [35] J. K. Deutschmann and I. Y. Bar-Itzhack, "Evaluation of attitude and orbit estimation using actual earth magnetic field data," *J. Guid., Control, Dyn.*, vol. 24, no. 3, pp. 616–623, 2001.
- [36] M. L. Psiaki, "Global magnetometer-based spacecraft attitude and rate estimation," *J. Guid., Control, Dyn.*, vol. 27, no. 2, pp. 240–250, 2004.
- [37] J. D. Searcy and H. J. Pernicka, "Magnetometer-only attitude determination using novel two-step Kalman filter approach," *J. Guid., Control, Dyn.*, vol. 35, no. 6, pp. 1693–1701, 2012.
- [38] J. R. Wertz, Ed., *Spacecraft Attitude Determination and Control*. Boston, MA, USA: Reidel, 1978, p. 565.
- [39] IAGA V-MOD Geomagnetic Field Modeling: International Geomagnetic Reference Field IGRF-12, accessed on Apr. 29, 2016. [Online]. Available: <http://www.ngdc.noaa.gov/IAGA/vmod/igrf.html>
- [40] L. Perea, J. How, L. Breger, and P. Elosegui, "Nonlinearity in sensor fusion: Divergence issues in EKF, modified truncated GSF, and UKF," in *Proc. AIAA Guid., Navigat. Control Conf. Exhibit*, Hilton Head, SC, USA, Aug. 2007, paper AIAA 2007-6514.



JAY G. MOORE received the B.S. and M.S. degrees in aeronautics and astronautics from the Massachusetts Institute of Technology, Cambridge, MA, and the M.S. degree in computer science from The Johns Hopkins University, Baltimore, MD. He is an Applied Mathematician with the Advanced RF Capabilities Group, The Johns Hopkins University Applied Physics Laboratory. His career has included such widely-varying duties as aircraft battle damage repair for the A-10A in the United States Air Force, payload integration for the Predator and Reaper UAS's at General Atomics Aeronautical Systems, and science and technology portfolio management for the United States Marine Corps at the Office of Naval Research. Since joining JHU/APL, he has designed and implemented artificial intelligence algorithms for swarming autonomous vehicles and for simulation of manned air-to-air combat tactics, and signal processing algorithms for software-defined radios. His current focus is on highly non-linear, non-Gaussian estimation and optimization.



MURTY S. CHALLA received the B.Sc. degree in physics from Andhra University, Visakhapatnam, India, and the Ph.D. degree in physics from the University of Georgia, Athens, GA. He has been a Senior Professional Staff Member with The Johns Hopkins Applied Physics Laboratory (JHU/APL), Maryland, USA, since 2007. In a career spanning over three decades, his professional interests and activities span estimation and data fusion algorithms, such as Kalman filters, simultaneous localization and mapping; track batch estimators, and correlation/association; guidance, navigation, and control for spacecraft, unmanned aerial vehicles, and surface vessels; debris analyses, missile defense; quantum computing; statistical mechanics; and solid state physics and materials science. Prior to JHU/APL, he was a Senior Research Analyst with the Institute for Defense Analyses, Alexandria, VA, and Computer Sciences Corporation supporting NASA Goddard Space Flight Center, Greenbelt, MD. He has also served as a consultant to Iridium Satellite, LLC.



DANIEL J. ROGERS (M'10) received the bachelor's degree in mathematics and physics from Georgetown University, and the Ph.D. degree in chemical physics from the University of Maryland. He is the Co-Founder and CEO of Terbium Labs, an information security startup based in Baltimore, MD. He is a Computational Physicist with experience supporting Defense and Intelligence Community Cyber Operations, as well as startup experience in the defense, energy, and biotechnology sectors. He is an author and expert in the field of quantum cryptography and has published numerous patents and papers on that and other subjects. Prior to co-founding Terbium Labs, he managed a portfolio of physics and sensor research projects with The Johns Hopkins University Applied Physics Laboratory. He is a member of the Optical Society of America and was a member of the Institute of Navigation.

• • •