JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 30, No. 6, November–December 2007

Optimal Linear Attitude Estimator

Daniele Mortari*
Texas A&M University, College Station, Texas 77843-3141
F. Landis Markley†
NASA Goddard Space Flight Center, Greenbelt, Maryland 20771
and
Puneet Singla‡
University at Buffalo, State University of New York, Buffalo New York 14260

DOI: 10.2514/1.29568

An optimal linear attitude estimator is presented for the case of a single-point real-time estimation of spacecraft attitude using the minimum-element attitude parameterization: Rodrigues (or Gibbs) vector g. The optimality criterion, which does not coincide with Wahba's constrained criterion, is rigorously quadratic and unconstrained. The singularity, which occurs when the principal angle is π, can easily be avoided by one rotation. The attitude accuracy tests show that the proposed method provides a precision comparable with those fully complying with the Wahba optimality definition. Finally, computational speed tests demonstrate that the proposed method belongs to the class of the fastest optimal attitude estimation algorithms.

Nomenclature

B	=	attitude profile matrix
b_i	=	observed ith unit vector
C	=	attitude matrix
e	=	principal axis
g	=	Rodrigues vector
L_m	=	optimal linear attitude estimator optimality cost function
L_w	=	Wahba optimality cost function
$P_{\theta\theta}$	=	covariance matrix
q	=	quaternion
r_i	=	reference ith unit vector
Φ	=	principal angle

I. Introduction

EVEN though spacecraft orientation depends on only three independent parameters, the spacecraft attitude is usually expressed by means of the singularity-free but redundant four-element quaternion (more precisely, the Euler–Rodrigues symmetric parameters). However, three-element attitude parameterizations have recently found increased interest in the attitude determination problem for two main reasons: 1) because they allow the development of faster algorithms and 2) because the method of sequential rotations [1–3] not only guarantees avoiding the geometric singularity, but also places the algorithm closer to its optimal working conditions.

It is a well-known fact that the attitude estimation problem consists of finding the proper orthogonal 3×3 matrix C such that the following is true:

$$b_i = C r_i \tag{1}$$

where $i = 1, \ldots, n$, and b_i and r_i are the ith observed and associated reference directions, respectively. These equations hold in the ideal case of perfect measurements (observations b_i). When the real case is considered, an optimality criterion must be introduced to estimate a spacecraft attitude as close as possible to the true attitude. In particular, the criterion introduced by Wahba in 1965 [4] has been widely adopted as the optimality definition par excellence, and a variety of "optimal" attitude estimators have been developed based on this criterion. These algorithms, which differ by small speed and robustness advantages/disadvantages, provide identical attitude accuracy because they fully comply with the same optimality definition. The method proposed in this paper, by contrast, does not make use of the standard Wahba optimality definition.

In this paper, a new optimal three-parameter attitude estimation algorithm (partially already introduced in [5]) is presented that reformulates the inherently nonlinear constrained attitude estimation problem as a rigorously *linear unconstrained* problem to solve the single-point optimal attitude estimation problem. This is achieved with the Cayley transformation that represents the direction cosine matrix by the Rodrigues (or Gibbs) vector g. The new optimality criterion is identical to Wahba's [4] optimality criterion in the absence of any errors in observed and reference vector directions.

The structure of the paper is as follows: First, an optimal linear attitude estimator (OLAE) is proposed for the determination of the single-point spacecraft attitude in terms of the Gibbs vector g. In the next section, the method of sequential rotations is discussed to avoid the singularity associated with the Gibbs vector for π rotation about the principal axis. Next, a detailed covariance analysis is presented to quantify the approximation error of the OLAE estimates. Finally, numerical tests are conducted to quantify and compare the accuracy of the OLAE with Wahba-compliant algorithms.

II. Optimal Linear Attitude Estimator

The proposed method is developed from relationships equivalent to Eq. (1), but written in terms of the Rodrigues vector. This is obtained by mapping the attitude matrix C to a minimum-element attitude parametrization, expressed by the skew-symmetric Rodrigues matrix G, using the Cayley transform.[6].

In general, the Cayley transformation represents an important and elegant mathematical tool in matrix analysis. In particular, for the attitude estimation problem, it allows the transformation of the most important attitude parametrization (the attitude matrix C) into the minimum number of elements describing the attitude (and vice versa). The Cayley transformation (Cayley conformal mapping) consists of the following forward

Presented as Paper 129 at the 10th Annual AIAA/AAS Space Flight Mechanics Meeting, Clearwater, FL, 23–26 January 2000; received 3 January 2007; revision received 29 June 2007; accepted for publication 5 July 2007.

*Associate Professor, Aerospace Engineering, 611C Bright Building. Member AIAA.

†Aerospace Engineer, Guidance, Navigation and Control Systems Engineering Branch, Mail Stop 591.0. Fellow AIAA.

‡Assistant Professor, Mechanical and Aerospace Engineering, 1010 Furnas Hall. Member AIAA.

$$C = (I - G)(I + G)^{-1} = (I + G)^{-1}(I - G) \tag{2}$$

and inverse

$$G = (I - C)(I + C)^{-1} = (I + C)^{-1}(I - C) \tag{3}$$

transformations, where G is a skew-symmetric 3×3 matrix, which is related to the elements of the Rodrigues vector g:

$$G = [g\times] = \begin{bmatrix} 0 & -g_3 & +g_2 \\ +g_3 & 0 & -g_1 \\ -g_2 & +g_1 & 0 \end{bmatrix} \tag{4}$$

The relation of the Cayley parameters G to the Rodrigues vector g is given in [6,7], even though the elements of the Rodrigues vector (Rodrigues parameters) were used by Rodrigues in [8] to describe rotations in 1840, long before Cayley, and Cayley apparently knew the relation to his transform in 1843.

Substituting the second term of Eq. (2) into Eq. (1), we obtain the following relationship:

$$[g\times](b_i + r_i) = -(b_i - r_i) \tag{5}$$

Defining the sum $s_i \triangleq b_i + r_i$ and difference $d_i \triangleq b_i - r_i$, we obtain

$$[g\times]s_i = -d_i \tag{6}$$

This equation, which is equivalent to Eq. (1), shows that the skew-symmetric matrix $[g\times]$ can take the place of the attitude matrix C. Because the matrix $[g\times]$ also represents the matrix performing the 3-D vector cross product, the previous equation can be rewritten as

$$[s_i\times]g = d_i \tag{7}$$

Note that the three elements of Eq. (7) are not all independent. In fact, it is easy to see that the third component can be obtained just by summing the first component multiplied by $-s_i(1)/s_i(3)$ with the second one multiplied by $-s_i(2)/s_i(3)$ and using the fact that $d_i^T s_i = 0 \quad \forall\ i$. This suggests that the problem can be simplified just by taking the two components of Eq. (7) with the greatest corresponding $d_i(k)$ values and disregarding the kth component corresponding to $\min_k(|d_i(k)|)$.

Note that to disregard the least significant component of Eq. (7), which speeds up the procedure, implies a small loss of information that, in turn, implies a small loss of accuracy. A further step toward the fastest algorithm would be to find a way to replace the observed and reference unit vectors with their polar coordinates.

It is a well-known fact that the skew-symmetric constraint for the G matrix

$$G + G^T = 0 \tag{8}$$

is sufficient to satisfy the orthogonality constraint $CC^T = I$, as well as the condition $\det(C) = +1$ required of attitude matrices. In other words, the off-diagonal elements of G (which are the only nonzero elements of the Rodrigues matrix) do not need to satisfy any particular constraint. This implies the well-known condition that the three Rodrigues parameters, which are the unknowns of Eq. (7), are all independent. It is important to outline the actual advantages of using Eq. (7) instead of Eq. (1):

1) The unknown attitude g does not have to satisfy any constraint.

2) The vectors s_i and d_i, which fully substitute the unit vectors r_i and b_i, are orthogonal ($s_i^T d_i = 0$, where $i = 1, \ldots, n$). [It is easy to show that $s_i^T s_i = 2(1 + b_i^T r_i)$ and $d_i^T d_i = 2(1 - b_i^T r_i)$, which implies that $s_i^T s_i + d_i^T d_i = 4$.]

An apparent disadvantage is the fact that the Rodrigues vector $g = e\tan(\Phi/2)$, where e and Φ are the principal axis and angle, respectively, is a singular attitude representation ($g \to \infty$ for $\Phi \to \pi$). However, as is explained in [2,3,9] and as will be explained in the next section, the use of only one sequential rotation completely avoids this singularity. Equation (7) cannot be satisfied by all the n measurements, due to the presence of sensor noise. In particular, we assume that the observed body vectors are given by the following expression:

$$\tilde{b}_i = b_i + \nu_i \tag{9}$$

where $i = 1, \ldots, n$, and ν_i is a zero-mean white-noise vector. Therefore, any given attitude (Rodrigues vector g) will affect the ith measurement by the following error:

$$e_i = \|[\tilde{s}_i\times]g - \tilde{d}_i\| \tag{10}$$

where $\tilde{s}_i = \tilde{b}_i + r_i$ and $\tilde{d}_i = \tilde{b}_i - r_i$. Therefore, any given solution g has the associated error vector $e = \{ e_1 \quad e_2 \quad \cdots \quad e_n \}^T$. This allows us to introduce a new optimality criterion for spacecraft attitude; we define the optimal attitude estimate as the unique Rodrigues parameter vector $\hat{g}$, which minimizes the following quadratic cost function:

$$L_m = \frac{1}{2} e^T \Xi e = \frac{1}{2}\sum_{i=1}^{n} \xi_i([\tilde{s}_i\times]g - \tilde{d}_i)^T([\tilde{s}_i\times]g - \tilde{d}_i) \tag{11}$$

where the diagonal matrix Ξ contains the weights $\xi_i = \Xi(i, i)$, where $i = 1, \ldots, n$, associated with the observations. It is important to mention that the *preceding introduced cost function L_m does not coincide with Wahba's cost function* [4]:

$$L_w = \frac{1}{2}\sum_{i=1}^{n} \alpha_i(Cr_i - \tilde{b}_i)^T(Cr_i - \tilde{b}_i) \tag{12}$$

This means that the optimal attitude that minimizes L_m generally differs from that minimizing L_w. This also implies that the associated attitude accuracy may be different and, therefore, the problem of how much they differ from each other has to be quantified by analysis or accuracy tests. The weights ξ_i, introduced in Eq. (11), are assumed here to be relative weights:

$$\sum_i \xi_i = 1$$

such as those used in the most classic attitude determination algorithms satisfying Eq. (12).

Because the unknown vector g consists of three independent parameters, the minimization of L_m is an unconstrained minimization problem. In contrast, the minimization of L_w is a constrained minimization problem, because $CC^T = I$ must be satisfied. Further, Eq. (11) can be rewritten as

$$\begin{aligned} 2L_m &= \sum_i \xi_i([\tilde{s}_i\times]g - \tilde{d}_i)^T([\tilde{s}_i\times]g - \tilde{d}_i) \\ &= \sum_i \xi_i(-g^T[\tilde{s}_i\times][\tilde{s}_i\times]g + g^T[\tilde{s}_i\times]\tilde{d}_i - \tilde{d}_i^T[\tilde{s}_i\times]g + \tilde{d}_i^T\tilde{d}_i) \\ &= \sum_i \xi_i\tilde{d}_i^T\tilde{d}_i - g^T\left[\sum_i \xi_i[\tilde{s}_i\times][\tilde{s}_i\times]\right]g - 2\left[\sum_i \xi_i\tilde{d}_i^T[\tilde{s}_i\times]\right]g \end{aligned} \tag{13}$$

and therefore,

$$L_m = c - g^T\tilde{M}_m g + 2\tilde{v}^T g \tag{14}$$

where

$$c = \frac{1}{2}\sum_{i=1}^{n}\xi_i\tilde{d}_i^T\tilde{d}_i, \quad \tilde{v} = \frac{1}{2}\sum_{i=1}^{n}\xi_i[\tilde{s}_i\times]\tilde{d}_i, \quad \tilde{M}_m = \frac{1}{2}\sum_{i=1}^{n}\xi_i[\tilde{s}_i\times][\tilde{s}_i\times] \tag{15}$$

Because this is an unconstrained minimization, and g is contained quadratically without approximation, the stationarity condition to minimize Eq. (11) yields

$$\frac{\mathrm{d}L_m}{\mathrm{d}g} = 2(\tilde{v} - \tilde{M}_m g) = 0 \tag{16}$$

and thus, the optimal least-squares estimate of g, denoted by $\hat{g}$, is found by solving the following *rigorously linear normal equation*:

$$\tilde{M}_m \hat{g} = \tilde{v} \tag{17}$$

where $\tilde{M}_m$ is a real, symmetric, negative definite matrix for $n \geq 2$. Hence, the second derivative of the cost function L_m with respect to the Rodrigues vector g, given as follows, is a positive definite matrix:

$$\frac{d^2 L_m}{\mathrm{d}g\,\mathrm{d}g^T} = -2\tilde{M}_m \tag{18}$$

Therefore, we can conclude that the OLAE estimate given by Eq. (17) corresponds to the global minimum of the cost function L_m.

It is a well-known fact that the Rodrigues vector represents the three vector components of Euler–Rodrigues symmetric parameters divided by the scalar component. Therefore, the optimal quaternion $\hat{q}_{\text{opt}}$ is simply obtained by normalizing $\hat{q} = \{\hat{g}^T \quad 1\}^T$; that is,

$$\hat{q}_{\text{opt}} = \frac{\hat{q}}{\sqrt{\hat{q}^T \hat{q}}} \tag{19}$$

This represents a new algorithm for attitude determination from vector observation, which is identified here as the OLAE. A fast variant of this algorithm, named OLAE-f, evaluates the symmetric $\tilde{M}_m$ matrix by disregarding the least significant line of Eq. (7), as explained earlier.

Some useful properties and relationships can be shown by analyzing the structure of the matrix $\tilde{M}_m$ and the vector $\tilde{v}$. The vector $\tilde{v}$ can be written as

$$\tilde{v} = \frac{1}{2}\sum_{i=1}^{n} \xi_i([\tilde{b}_i\times] + [r_i\times])(\tilde{b}_i - r_i) = \sum_{i=1}^{n} \xi_i [r_i\times]\tilde{b}_i = -\sum_{i=1}^{n} \xi_i [\tilde{b}_i\times] r_i \tag{20}$$

and the matrix $\tilde{M}_m$ can be expanded to obtain

$$2\tilde{M}_m = \sum_{i=1}^{n} \xi_i([\tilde{b}_i\times] + [r_i\times])([\tilde{b}_i\times] + [r_i\times]) = \sum_{i=1}^{n} \xi_i [\tilde{b}_i\times][\tilde{b}_i\times] + \sum_{i=1}^{n} \xi_i [r_i\times][r_i\times] + \sum_{i=1}^{n} \xi_i [\tilde{b}_i\times][r_i\times] + \sum_{i=1}^{n} \xi_i [r_i\times][\tilde{b}_i\times] \tag{21}$$

Using the property (valid for unit vectors) $[x\times][y\times] = xy^T - (x^T y)I$, where I indicates the 3×3 unit matrix, Eq. (21) becomes

$$2\tilde{M}_m = \sum_{i=1}^{n} \xi_i\left(\tilde{b}_i\tilde{b}_i^T + r_i r_i^T\right) - 2I\sum_{i=1}^{n} \xi_i + \sum_{i=1}^{n} \xi_i\left(\tilde{b}_i r_i^T + r_i \tilde{b}_i^T\right) - 2I\sum_{i=1}^{n} \xi_i r_i^T \tilde{b}_i \tag{22}$$

Under the assumption that the weights $\xi_i = \alpha_i$, where

$$\sum_i \alpha_i = 1$$

and introducing the attitude profile matrix B and the vector z as follows, we have

$$B = \sum_{i=1}^{n} \alpha_i \tilde{b}_i r_i^T, \qquad z = \sum_{i=1}^{n} \alpha_i [\tilde{b}_i\times] r_i = \begin{Bmatrix} B_{23} - B_{32} \\ B_{31} - B_{13} \\ B_{12} - B_{21} \end{Bmatrix} \tag{23}$$

Equations (20) and (22) become

$$\begin{cases} \tilde{v} = -z \\ \tilde{M}_m = \frac{1}{2}\sum_{i=1}^{n} \alpha_i\left(\tilde{b}_i\tilde{b}_i^T + r_i r_i^T\right) + \frac{1}{2}(B + B^T) - (\text{tr}[B] + 1)I \end{cases} \tag{24}$$

which implies that one needs to solve the following algebraic equations simultaneously for the OLAE algorithm:

$$\left[\frac{1}{2}\sum_{i=1}^{n} \alpha_i(\tilde{b}_i\tilde{b}_i^T + r_i r_i^T) + \frac{1}{2}(B + B^T) - (\text{tr}[B] + 1)I\right]\hat{g} = \tilde{M}_m \hat{g} = -z \tag{25}$$

We mention that although OLAE is a rigorously linear optimal attitude estimation algorithm, it suffers from the singularity associated with the π rotation about the principal axis e. In the next section, we discuss the method of sequential rotations to avoid this singularity.

A. Singularity

The singularity associated with the OLAE algorithm, due to the use of the Rodrigues vector, can be easily avoided by applying the method of sequential rotations, as introduced in [1] for the QUEST algorithm to avoid the singularity associated with π rotation about the principal axis. Recent developments given in [3,9,10] have demonstrated that because of some inequalities concerning the principal angle, it is possible to know in advance if (and which) sequential rotation is needed to place the algorithm far enough from its singular working position, occurring here at $\Phi = \pi$. The singularity [see Eq. (17)] occurs only when the determinant of the $\tilde{M}_m$ matrix becomes too small; the vector $\tilde{v}$ cannot diverge because it is a weighted sum of terms involving finite cross products.

In [9], it was demonstrated that the following inequality

$$\det(\tilde{M}_w) \leq f_w(\Phi) = c(1 + \cos\Phi) \tag{26}$$

where

$$c = \begin{cases} = -1 & \text{for } n = 2 \\ \approx -1.178 & \text{for } n > 2 \end{cases}$$

holds for the matrix $\tilde{M}_w$, where

$$\tilde{M}_w \hat{g} = [B + B^T - (\text{tr}[B] + 1)I]\hat{g} = -z \tag{27}$$

Analogously, the inequality

$$\det(\tilde{M}_m) \leq f_m(\Phi) = f_w(\Phi) + \frac{1 + \cos(2\Phi)}{8} \tag{28}$$

holds for matrix $\tilde{M}_m$. The inequalities given in Eq. (26) and (28) can be adopted to calculate an upper-bound value for the principal angle Φ, assuring us that $\Phi < \pi$. In particular, [9] showed that using the inequality given in Eq. (26) with the method of sequential rotations, the condition $\Phi \leq 2\pi/3$ is guaranteed.

Figures 1 and 2 plot the values of $\det(\tilde{M}_w)$ and $\det(\tilde{M}_m)$ (markers) and $f_w(\Phi)$ and $f_m(\Phi)$ (continuous), obtained in $N = 10,000$ random tests, each one constituted by n random observations, for which the random precision ranges from 0 to 0.3 deg (3σ) for the $n = 2$ and $2 \leq n \leq 5$ cases, respectively. These figures highlight that $f_m(\Phi)$ provides a lower bound for Φ lower than $f_w(\Phi)$. The random attitudes are created with a random principal axis and a principal angle Φ uniformly distributed from 0 to π. We mention that Eq. (27) can easily be obtained by partitioning the q-method solution equation $Kq_{\text{opt}} = \lambda_{\max} q_{\text{opt}}$ (see [11,12]), as explained in [3,13], and approximating the maximum eigenvalue $\lambda_{\max}$ to 1.

Note that Eq. (27) does not provide the same result for g as Eq. (25), because these equations are derived from two different optimality criteria. However, if measurements are error-free, then Eqs. (25) and (27) provide the same result for g. In particular,

$$\tilde{M}_m \hat{g} = -z = \tilde{M}_w \hat{g} \tag{29}$$

Solving for g from Eq. (27) and substituting into Eq. (25), we get

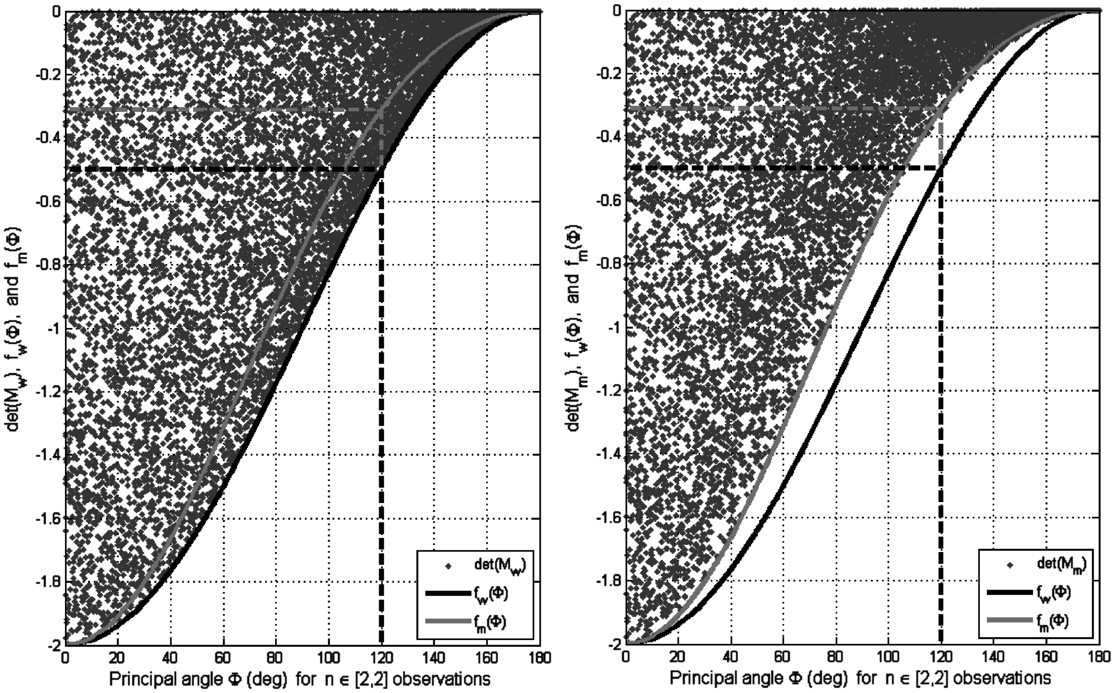


Fig. 1 Values of $\det(\tilde{M}_w)$ (left) and $\det(\tilde{M}_m)$ (right); $f_w(\Phi)$ and $f_m(\Phi)$ vs Φ for $n = 2$.

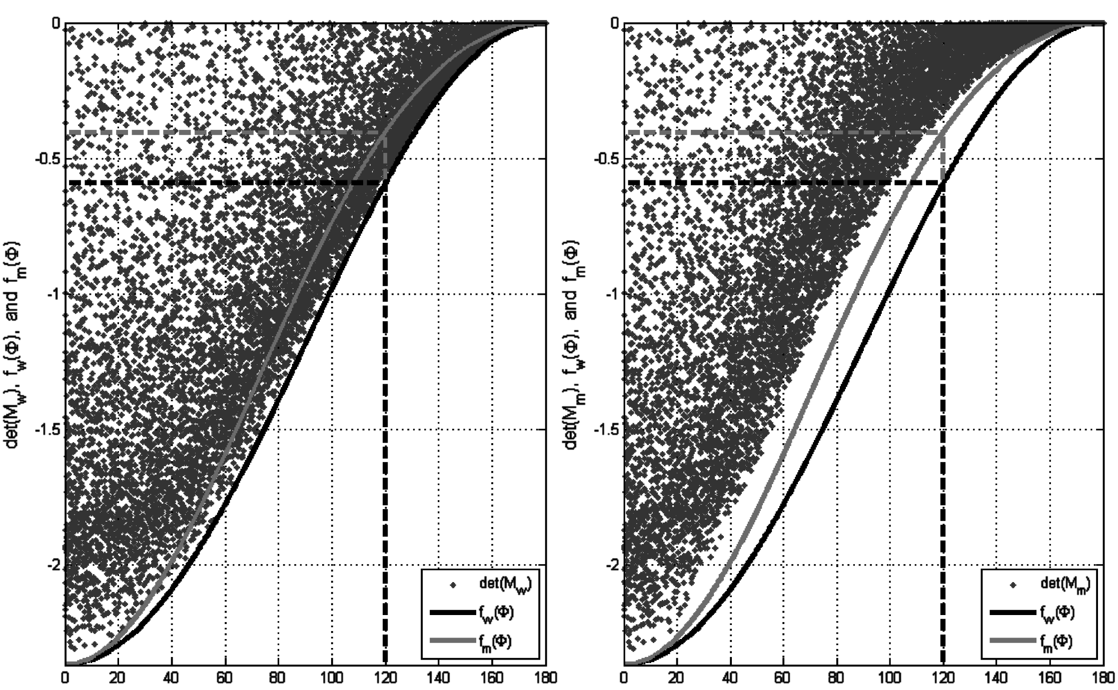


Fig. 2 Values of $\det(\tilde{M}_w)$ (left) and $\det(\tilde{M}_m)$ (right); $f_w(\Phi)$ and $f_m(\Phi)$ vs Φ for $2 \leq n \leq 5$.

$$\left\{\frac{1}{2}\sum_{i=1}^{n}\alpha_i\left(\tilde{b}_i\tilde{b}_i^T + r_i r_i^T\right) + \frac{1}{2}S - [\mathrm{tr}(B)+1]I\right\}\{S - [\mathrm{tr}(B)+1]I\}^{-1}z$$
$$= z, \quad S = B + B^T \tag{30}$$

Adding and subtracting $S/2$ inside the square bracket gives

$$\frac{1}{2}\left[\sum_{i=1}^{n}\alpha_i(\tilde{b}_i\tilde{b}_i^T + r_i r_i^T) - S\right]\hat{g} + z = z \tag{31}$$

Further, substituting for B from Eq. (23) into Eq. (31), we obtain the following condition for the OLAE and Wahba estimates to be identical:

$$\left(\sum_{i=1}^{n}\tilde{d}_i\tilde{d}_i^T\right)\hat{g} + z = z \tag{32}$$

It should be noticed that the two sides of the preceding equation will certainly be equal if $\tilde{d}_i^T g = 0$ for every measurement. But for perfect error-free measurements,

$$d_i^T\hat{g} = (b_i - r_i)^T\hat{g} = b_i^T(I - C^T)\hat{g} = r_i^T(\hat{g} - C\hat{g}) \tag{33}$$

This is zero because attitude matrix C changes neither the magnitude nor the direction of $\hat{g}$, which is along the rotation axis. As a consequence of this, Eqs. (25) and (27) provide the same result for $\hat{g}$ if measurements are error-free.

Further, from a practical point of view, the difference in two solutions is so small that it is possible to approximate $\tilde{M}_m$ by $\tilde{M}_w$. This approximation allows us an easy application of the sequential rotation technique because the *rotated* matrices $\bar{\tilde{M}}_w$ can easily be obtained from $\tilde{M}_w$ and z with no additional computation. In fact, setting $p = \text{tr}[B] + 1$, $m = \text{tr}[B] - 1$, and $S = B + B^T$, then Eq. (27) can be written as follows:

$$\tilde{M}_w \hat{g} = \begin{bmatrix} S_{11} - p & S_{12} & S_{13} \\ S_{12} & S_{22} - p & S_{23} \\ S_{13} & S_{23} & S_{33} - p \end{bmatrix} \hat{g} = -\begin{Bmatrix} z_1 \\ z_2 \\ z_3 \end{Bmatrix} \tag{34}$$

Using the sequential π rotations about the coordinate axes, the *rotated* matrices $\bar{\tilde{M}}_w$ and the *rotated* vectors $\bar{z}$ become

$$\bar{\tilde{M}}_w^{(x)} \hat{\bar{g}}^{(x)} = \begin{bmatrix} m & -z_3 & z_2 \\ -z_3 & (\tilde{M}_w)_{33} & -S_{23} \\ z_2 & -S_{23} & (\tilde{M}_w)_{22} \end{bmatrix} \hat{\bar{g}}^{(x)} = -\begin{Bmatrix} -z_1 \\ S_{13} \\ -S_{12} \end{Bmatrix} = -\bar{z}^{(x)} \tag{35}$$

$$\bar{\tilde{M}}_w^{(y)} \hat{\bar{g}}^{(y)} = \begin{bmatrix} (\tilde{M}_w)_{33} & z_3 & -S_{13} \\ z_3 & m & -z_1 \\ -S_{13} & -z_1 & (\tilde{M}_w)_{11} \end{bmatrix} \hat{\bar{g}}^{(y)} = -\begin{Bmatrix} -S_{23} \\ -z_2 \\ S_{12} \end{Bmatrix} = -\bar{z}^{(y)} \tag{36}$$

$$\bar{\tilde{M}}_w^{(z)} \hat{\bar{g}}^{(z)} = \begin{bmatrix} (\tilde{M}_w)_{22} & -S_{12} & -z_2 \\ -S_{12} & (\tilde{M}_w)_{11} & z_1 \\ -z_2 & z_1 & m \end{bmatrix} \hat{\bar{g}}^{(z)} = -\begin{Bmatrix} S_{23} \\ -S_{13} \\ -z_3 \end{Bmatrix} = -\bar{z}^{(z)} \tag{37}$$

which identify the transformed normal equations when a sequential rotation is applied to the x, y, or z coordinate axis, respectively. Note that, when approximate a priori attitude information is available, as in most common fast recursive applications, then the best sequential rotation can directly be read from the a priori quaternion $q = \{q_1 \quad q_2 \quad q_3 \quad q_4\}^T$. In fact, the smallest (the greatest) absolute element $|q_k|$ also indicates the greatest (the smallest) principal angle associated with the π sequential rotation applied about the kth coordinate axis ($k = 4$ means no rotation).

B. Covariance Analysis

In the previous section, we showed that in case of error-free measurements, the OLAE estimates are identical to any optimal estimator fully complying with Wabha's optimality definition. In this section, we present a detailed covariance analysis that helps us in quantifying the approximation error of the OLAE estimate.

Substituting $\tilde{s}_i = s_i + \nu_i$ into the expression for matrix $\tilde{M}_m$ gives

$$\tilde{M}_m = \frac{1}{2}\sum_i \alpha_i [\tilde{s}_i \times][\tilde{s}_i \times] = M_m + M_1(\nu) + M_2(\nu) \approx M_m + M_1(\nu) \tag{38}$$

where

$$M_m = \frac{1}{2}\sum_i \alpha_i [s_i \times][s_i \times] = \frac{1}{2}\sum_i \alpha_i (s_i s_i^T - s_i^T s_i I) \tag{39}$$

and where the two matrices

$$M_1(\nu) = \frac{1}{2}\sum_i \alpha_i [s_i \nu_i^T + \nu_i s_i^T - 2(s_i^T \nu_i) I] \tag{40}$$

$$M_2(\nu) = \frac{1}{2}\sum_i \alpha_i [\nu_i \nu_i^T - \nu_i^T \nu_i I] \tag{41}$$

contain only linear and quadratic terms in ν, respectively. A similar substitution into the expression for the vector $\tilde{v}$ gives

$$\tilde{v} = \frac{1}{2}\sum_i \alpha_i (\tilde{s}_i \times \tilde{d}_i) = v + v_1(\nu) \tag{42}$$

where

$$v = \frac{1}{2}\sum_i \alpha_i s_i \times d_i$$

and

$$v_1(\nu) = \frac{1}{2}\sum_i \alpha_i (s_i \times \nu_i + \nu_i \times d_i) = \sum_i \alpha_i r_i \times \nu_i \tag{43}$$

The exact vector of the Rodrigues parameters is given by

$$g = M_m^{-1} v \tag{44}$$

and the OLAE estimate is given (to first order in ν_i) by

$$\begin{aligned} \hat{g} = \tilde{M}_m^{-1}\tilde{v} &= (M_m + M_1)^{-1}\left(v + \sum_i \alpha_i r_i \times \nu_i\right) \\ &\approx (M_m^{-1} - M_m^{-1} M_1 M_m^{-1})\left(v + \sum_i \alpha_i r_i \times \nu_i\right) \\ &= (I - M_m^{-1} M_1)\left(g + M_m^{-1}\sum_i \alpha_i r_i \times \nu_i\right) \\ &\approx g - M_m^{-1}\left(M_1 g - \sum_i \alpha_i r_i \times \nu_i\right) = g + M_m^{-1}\sum_i \alpha_i W_i \nu_i \end{aligned} \tag{45}$$

where $W_i = \frac{1}{2}\{2 g s_i^T - s_i g^T - (s_i^T g) I + 2[r_i \times]\}$. In the small-error approximation, the error angle vector is given (to first order in ν_i) by

$$\begin{aligned} \theta = 2\hat{g} \otimes g^{-1} &= 2\frac{\hat{g} - g + \hat{g} \times g}{1 + \hat{g}^T g} \\ &= 2\frac{M_m^{-1}\sum_i \alpha_i W_i \nu_i + (M_m^{-1}\sum_i \alpha_i W_i \nu_i) \times g}{1 + g^2 + (M_m^{-1}\sum_i \alpha_i W_i \nu_i)^T g} \\ &= 2\frac{(I - [g\times]) M_m^{-1}\sum_i \alpha_i W_i \nu_i}{1 + g^2 + (M_m^{-1}\sum_i \alpha_i W_i \nu_i)^T g} \\ &\approx 2\frac{(I - [g\times]) M_m^{-1}\sum_i \alpha_i W_i \nu_i}{1 + g^2} \end{aligned} \tag{46}$$

Therefore, the covariance matrix $P_{\theta\theta} = E\{\theta\theta^T\}$ has the following expression:

$$\begin{aligned} P_{\theta\theta} = 4(1 + g^2)^{-2}(I - [g\times]) M_m^{-1} E\Bigg\{&\left(\sum_i \alpha_i W_i \nu_i\right) \\ &\times \left(\sum_j a_j W_j \nu_j\right)^T\Bigg\} M_m^{-1}(I - [g\times])^T \end{aligned} \tag{47}$$

Assuming the QUEST measurement model [1] for observations, that is,

$$E\{\nu_i \nu_j^T\} = \sigma_i^{-2}(I - b_i b_i^T) \tag{48}$$

and the relative weights to be $\alpha_i = \sigma_i^{-2}$ (which makes Wahba's optimal problem a maximum likelihood problem), we obtain

$$\begin{aligned} P_{\theta\theta} = {} & 4(1 + g^2)^{-2}(I - [g\times]) M_m^{-1} \\ & \times \left[\sum_i \alpha_i W_i (I - b_i b_i^T) W_i^T\right] M_m^{-1}(I - [g\times])^T \end{aligned} \tag{49}$$

Further, we can express the final result entirely in terms of the vectors g and s_i by using the following two identities:

$$r_i = \tfrac{1}{2}(I + [g\times])s_i \quad \text{and} \quad b_i = \tfrac{1}{2}(I - [g\times])s_i \tag{50}$$

that can be easily derived from Eq. (6). Substituting the first of these into the expression for W_i gives

$$\begin{aligned} W_i &= 2gs_i^T - s_i g^T - (s_i^T g)I + [s_i\times] + [g\times][s_i\times] \\ &= 2gs_i^T - s_i g^T - (s_i^T g)I + [s_i\times] + s_i g^T - gs_i^T \\ &= [s_i\times] + gs_i^T - (s_i^T g)I = [s_i\times](I + [g\times]) \end{aligned} \tag{51}$$

Now, using the preceding expression for W_i, we can write

$$W_i(I - b_i b_i^T)W_i^T = [s_i\times](I + [g\times])(I - b_i b_i^T)(I + [g\times])^T[s_i^T\times] \tag{52}$$

Also, using the following two identities

$$(I + [g\times])(I + [g\times])^T = I - [g\times][g\times] = (1 + g^2)I - gg^T \tag{53}$$

$$(I + [g\times])b_i = \tfrac{1}{2}(I + [g\times])(I - [g\times])s_i = \tfrac{1}{2}[(1 + g^2)I - gg^T]s_i \tag{54}$$

we can further simplify the expression in Eq. (52):

$$\begin{aligned} &W_i(I - b_i b_i^T)W_i^T \\ &= [s_i\times]\Big\{[(1 + g^2)I - gg^T] - \tfrac{1}{4}[(1 + g^2)I - gg^T]s_i s_i^T[(1 \\ &+ g^2)I - gg^T]\Big\}[s_i^T\times] = (1 + g^2)[s_i\times][s_i^T\times] \\ &- (s_i \times g)(s_i \times g)^T - \tfrac{1}{4}(s_i^T g)^2(s_i \times g)(s_i \times g)^T \\ &= (1 + g^2)(s_i^2 I - s_i s_i^T) - \left[1 + \tfrac{1}{4}(s_i^T g)^2\right](s_i \times g)(s_i \times g)^T \\ &= (1 + g^2)\left[(s_i^2 I - s_i s_i^T) - \tfrac{1}{4}s_i^2(s_i \times g)(s_i \times g)^T\right] \end{aligned} \tag{55}$$

where the last step uses the identity

$$\begin{aligned} s_i^2 &= |b_i + r_i|^2 = 2 + 2(b_i^T r_i) = 2 + \tfrac{1}{2}s_i^T(I - [g\times])^T(I + [g\times])s_i \\ &= 2 + \tfrac{1}{2}s_i^T(I + 2[g\times] + [g\times][g\times])s_i = 2 + \tfrac{1}{2}s_i^T[(1 - g^2)I + g^T g]s_i \\ &= 2 + \tfrac{1}{2}[(1 - g^2)s_i^2 + (g^T s_i)^2] \end{aligned}$$

from which we can derive the identity

$$(g^T s_i)^2 = 2s_i^2 - 4 - (1 - g^2)s_i^2 = (1 + g^2)s_i^2 - 4 \tag{56}$$

Finally, substituting Eq. (55) into Eq. (49), we obtain the final form for the covariance matrix:

$$P_{\theta\theta} = -(1 + g^2)^{-1}(I - [g\times])M_m^{-1}(4M_m + N)M_m^{-1}(I - [g\times])^T \tag{57}$$

where

$$N = \sum_i \alpha_i s_i^2(s_i \times g)(s_i \times g)^T \tag{58}$$

Note that $s_i \times g$ has a finite limit as g becomes infinitely large, and so N remains finite in this limit. Further, a major simplification happens for $g = 0$. In fact, in this case, $r_i = b_i$, $s_i = 2b_i$,

$$M_m = -4\sum_i \alpha_i(I - b_i b_i^T)$$

$N = 0$, and

$$P_{\theta\theta} = -4M_m^{-1} = \left[\sum_i \alpha_i(I - b_i b_i^T)\right]^{-1} \tag{59}$$

which is the optimal Wahba covariance. Thus, we see that OLAE complies with Wahba optimality at the identity attitude and we have already shown that, for error-free observations, the OLAE estimate is same as any Wahba-optimality-compliant attitude estimation algorithm.

C. Special Case, $n = 3$

In this section, we consider the special case that there are $n = 3$ measurements with equal weights along the spacecraft body axes. Therefore, b_i takes the three values $\{1 \quad 0 \quad 0\}^T$, $\{0 \quad 1 \quad 0\}^T$, and $\{0 \quad 0 \quad 1\}^T$. After a significant amount of algebra (that we do not include here), we obtain

$$M_m = -4\sigma^{-2}(1 + g^2)^{-1}[2I + (g^2 I - gg^T)] \tag{60}$$

from which we derive

$$M_m^{-1} = -\frac{1}{4}\sigma^2\frac{1 + g^2}{2 + g^2}\left(I + \frac{1}{2}gg^T\right) \tag{61}$$

and therefore

$$P_{\theta\theta} = \tfrac{1}{2}\sigma^2\{I + (2 + g^2)^{-2}[g^2(g^2 I - gg^T) - 2X]\} \tag{62}$$

where

$$X \triangleq \begin{bmatrix} 2g_2^2 g_3^2 & -g_3^2 g_1 g_2 & -g_2^2 g_3 g_1 \\ -g_3^2 g_1 g_2 & 2g_3^2 g_1^2 & -g_1^2 g_2 g_3 \\ -g_2^2 g_3 g_1 & -g_1^2 g_2 g_3 & 2g_1^2 g_2^2 \end{bmatrix} \tag{63}$$

Once again, for $g = 0$, the OLAE covariance matrix $P_{\theta\theta} = \frac{1}{2}\sigma^2 I$ (i.e., the OLAE estimates are compliant with Wahba optimality criteria for the identity attitude). Also notice that $P_{\theta\theta}g = \frac{1}{2}\sigma^2 g$, and so this is an eigenvector of the covariance with the optimal Wahba eigenvalue. The other two eigenvalues of X (associated with directions perpendicular to g) are

$$\lambda = g_1^2 g_2^2 + g_2^2 g_3^2 + g_3^2 g_1^2 \pm \sqrt{g_1^4 g_2^4 + g_2^4 g_3^4 + g_3^4 g_1^4 - g^2 g_1^2 g_2^2 g_3^2} \tag{64}$$

If g is parallel to $\{1 \quad 1 \quad 1\}^T$, then both of these eigenvalues are equal to $g^2/3$, and the eigenvalues of $P_{\theta\theta}$ perpendicular to g are found to be equal to $\frac{1}{2}\sigma^2[1 + \frac{1}{3}(2 + g^2)^{-2}g^4]$. In particular, for $g = \{1 \quad 1 \quad 1\}^T$, both the eigenvalues perpendicular to g are found to be $\frac{1}{2}\sigma^2(1.12)$. Further, if g is along a coordinate axis, then both of these eigenvalues of X are zero, and the eigenvalues of $P_{\theta\theta}$ perpendicular to g are found to be $\frac{1}{2}\sigma^2[1 + (2 + g^2)^{-2}g^4]$. This seems to be a worst (least optimal) case for the assumed measurement geometry. Note that as g becomes infinitely large, the covariance is never worse than twice the optimal value. We can use sequential rotations to limit the rotation angle to 120 deg, $g^2 \leq 3$, and the covariance to no greater than 1.36 times the optimal value.

In the following section, the accuracy provided by OLAE is numerically quantified and tested using MATLAB software [14].

III. Numerical Results

The accuracy tests compare OLAE with the solution provided by the QUEST algorithm [1] as one of the algorithms that fully comply with Wahba's definition of the optimal attitude. Any other optimal algorithm would provide identical accuracy results. This means that we are comparing the optimality definition of Eq. (11) with that of Eq. (12). The accuracy provided by the proposed method, which was already analytically estimated in the covariance analysis section, is quantified here with a Monte Carlo numerical analysis.

There are different ways to highlight the accuracy achieved by an attitude estimator:

1)_Attitude error angle Φ is defined as the principal angle of the corrective attitude matrix $C\hat{C}^T$, where C is the true attitude and $\hat{C}$ is its estimation. This angle, which represents the maximum direction error, can be easily evaluated [15] using quaternions $(\hat{q}, q)$ or Rodrigues parameters $(\hat{g}, g)$, as follows:

$$\cos\Phi = 2(\hat{q}^T q)^2 - 1 = \frac{2(\hat{g}^T g + 1)^2}{(1+\hat{g}^T\hat{g})(1+g^T g)} - 1 \tag{65}$$

2) The attitude error vector for the Rodrigues parameters is defined as

$$\delta\theta = 2\hat{g}\otimes g^{-1} = 2\frac{\hat{g} - g + \hat{g}\times g}{1+\hat{g}^T g} \tag{66}$$

3) Covariance matrix P (elements comparison) gives a detailed and complete picture of how the error is distributed in each of its components. The error, however, is defined here according to the error model selected. Unfortunately, most of the time, the error as defined in the error model is a residual and not the attitude error itself. This is important because it means that covariance matrix highlights the behavior of residuals and not the attitude error.

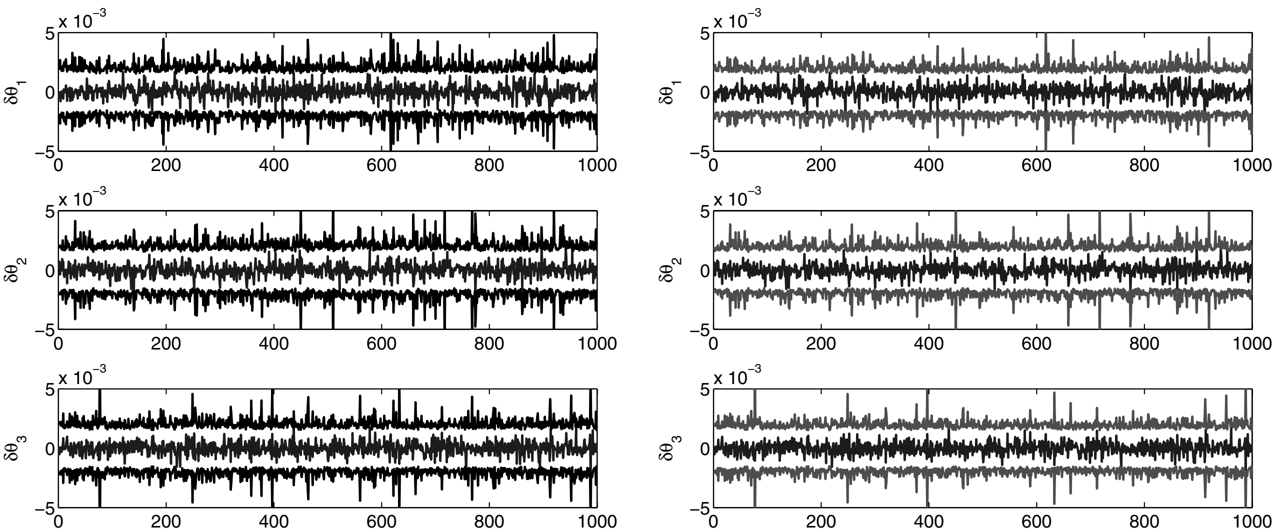
4.) The square root of the trace of the covariance matrix, $\sqrt{\mathrm{tr}(P)}$, gives a global picture of the selected error-model quantification. This parameter allows us to define upper bounds of the attitude accuracy for a specific selected error model.

For simulation purposes, the measurement data and attitudes are $N = 10,000$ times, randomly produced using sensor noise with 10^{-3} rad (3σ), and for a number of the observed directions n varying from 3 to 10. The results of these tests are plotted in Fig. 3. Figures 3a and 3b show the plots of $\delta\theta_m$ and $\delta\theta_w$, respectively, with corresponding 3-σ bounds for different Monte Carlo runs. From these results, it is clear that

1) The computed estimates are compatible with covariance analysis presented in the previous section.

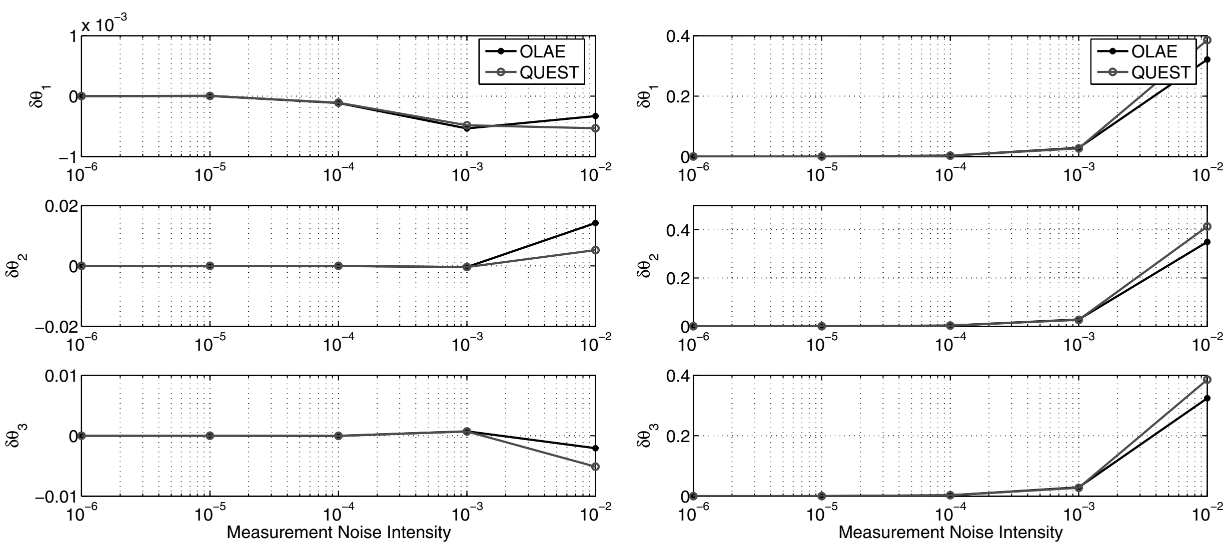
2) OLAE optimality criterion (11) provides an attitude accuracy almost identical with that derived from the Wahba cost function (12).

Further, to test the robustness of the algorithm with respect to sensor noise, the true measurement data are corrupted by zero-mean sensor noise with standard deviation between 10^{-1} to 10^{-6} rad. For simulation purposes, $N = 10,000$ sets of measurements for a number n of observed direction varying from 3 to 10 were randomly produced using sensor noise of 10^{-2} rad (3σ). In this case, the parameter used to quantify the accuracy is the mean value of attitude error



a) OLAE approximation error $\delta\theta_m$ b) QUEST approximation error $\delta\theta_w$

Fig. 3 Attitude accuracy plots for OLAE and QUEST.



a) Mean attitude error b) Standard deviation

Fig. 4 OLAE and QUEST estimate accuracy.

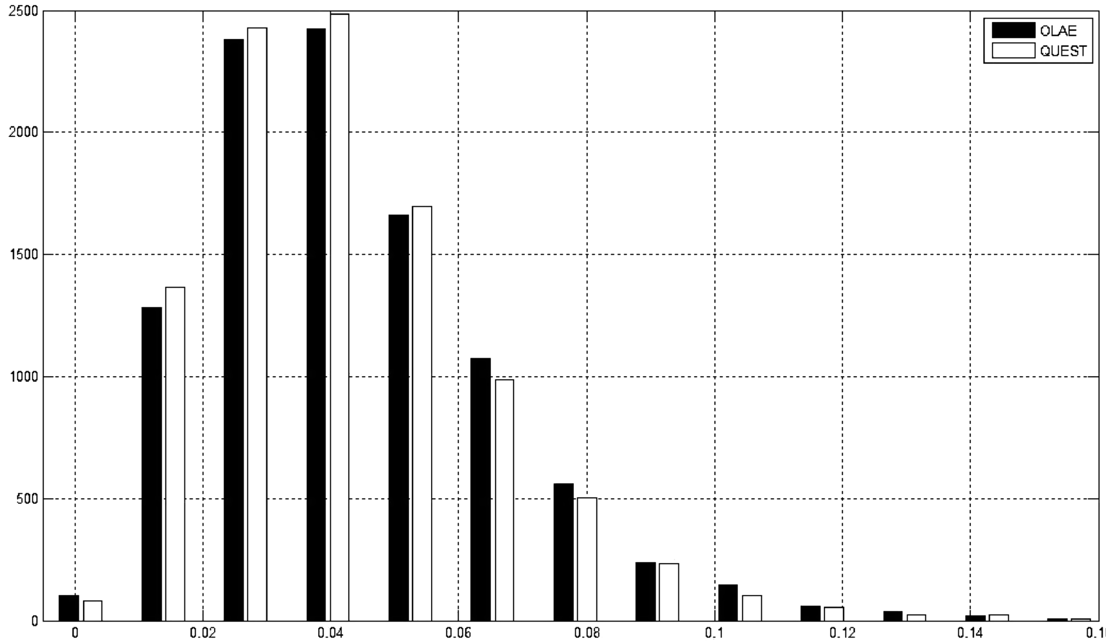


Fig. 5 Histogram of Φ, deg.

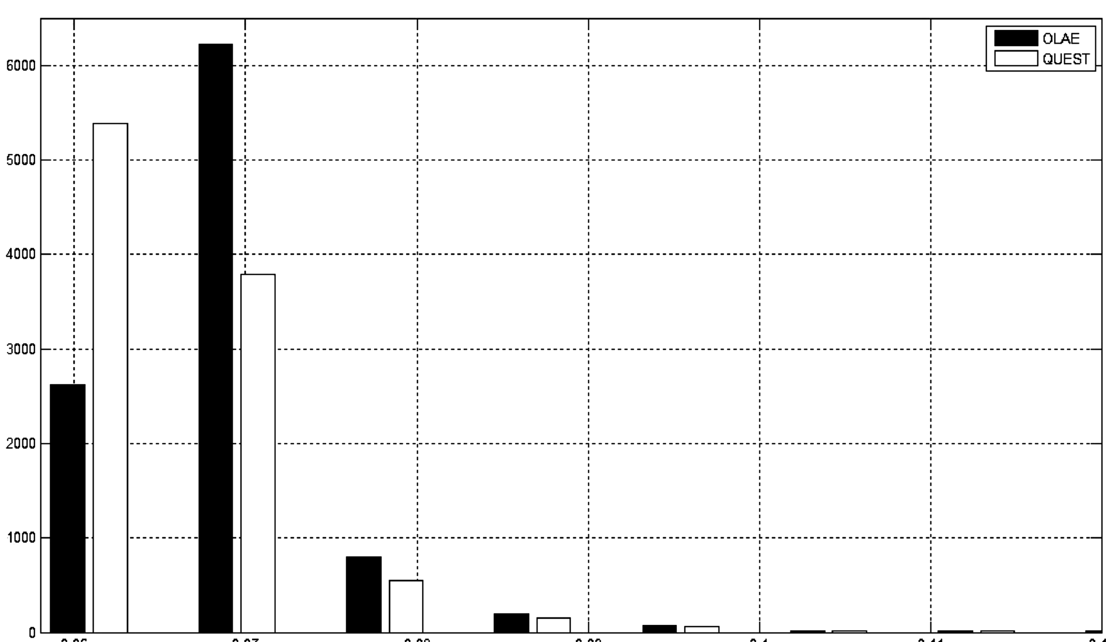


Fig. 6 Histogram of $\sqrt{\mathrm{tr}(P)}$, deg.

$$\varepsilon = \frac{1}{N}\sum_{i=1}^{N} \|\delta\theta\|$$

experienced in N random tests. Figures 4a and 4b show the plots of mean attitude error and corresponding standard deviation, respectively, for the OLAE and the QUEST algorithms. For the same set of $m = 10,000$ random tests, Φ is computed for OLAE and for the QUEST algorithm, giving Φ_m and Φ_w, respectively. The measurement data are randomly produced using sensor noise with 10^{-3} rad (3σ) and for a number of the observed directions n varying from 3 to 10. The results of these tests are plotted in Figs. 5 and 6, respectively.

As expected, the difference between OLAE estimates and Wahba-compliant QUEST estimates is negligible for all practical purposes. Thus, either criterion can be expected to give excellent practical results.

IV. Conclusions

This paper introduces the new optimal linear attitude estimator (OLAE) for spacecraft attitude that makes use of Cayley conformal mapping, a general mathematical tool relating orthogonal and skew-symmetric matrices. The relationship between observed and reference directions is transformed into a linear equation relating the sum and the difference between observed and reference directions (orthogonal to one another). The new relationship is shown to be linear if the unknown attitude is expressed in terms of the Rodrigues (or Gibbs) vector and allows us to define a cost function for which the minimization introduces a new optimality criterion for spacecraft attitude. The minimization is unconstrained, and the solution (the optimal attitude) only requires the inversion of a 3×3 symmetric matrix. The Rodrigues vector singularity, which occurs when the principal angle is close to π, is avoided using, at most, one sequential rotation. Further, it is shown that for error-free observations, the

OLAE estimate is equivalent to the Wahba-compliant estimates. However, in general, the OLAE attitude estimate does not mathematically comply with Wahba's optimality definition, and therefore the OLAE attitude estimation accuracy is numerically quantified by using a Monte Carlo approach. The accuracy provided is compared with the optimal attitude fully complying with Wahba condition, and the differences between the two estimates are insignificant in a practical attitude estimation application.

References

[1] Shuster, M. D., and Oh, S. D., "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance, Control, and Dynamics*, Vol. 4, No. 1, Jan.–Feb. 1981, pp. 70–77.

[2] Shuster, M. D., and Natanson, G. A., "Quaternion Computation from a Geometric Point of View," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, Oct.–Dec. 1993, pp. 545–556.

[3] Mortari, D., Angelucci, M., and Markley, F. L., "Singularity and Attitude Estimation," *Advances in the Astronautical Sciences*, Vol. 105, No. 1, 2000, pp. 479–496.

[4] Wahba, G., "Problem 65-1: A Least Squares Estimate of Spacecraft Attitude," *SIAM Review*, Vol. 7, No. 3, July 1965, p. 409. doi:10.1137/1007077

[5] Mortari, D., Markley, F. L., and Junkins, J. L., "Optimal Linear Attitude Estimator," *Advances in the Astronautical Sciences*, Vol. 105, No. 1, 2000, pp. 465–478.

[6] Junkins, J. L., and Kim, Y., *Introduction to Dynamics and Control of Flexible Structures*, AIAA Education Series, AIAA, Washington, DC, 1993.

[7] Shuster, M. D., "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, Oct.–Dec. 1993, pp. 439–517.

[8] Rodrigues, M. O., "Des Lois Géométriques Qui Régissent les Déplacement d'un Systéme Solide dans l'Espace, et de la Variation des Coordonnées Provenant de ces Déplacements Considérés Independamment des Causes qui Peuvent les Produire," *Journal de Mathématique Pures et Appliquées*, Vol. 5, 1840, pp. 380–440.

[9] Markley, F. L., and Mortari, D., "How to Compute Attitude from Vector Observations," Astrodynamics Specialist Conference, Girdwood, AK, American Astronautical Society, Paper 99-427, 15–19 Aug. 1999.

[10] Markley, F. L., and Mortari, D., "New Developments in Quaternion Estimation from Vector Observations," *Advances in the Astronautical Sciences*, Vol. 106, 20–21 Mar. 2000, pp. 373–393; also American Astronautical Society Paper 00-266.

[11] Keat, J., "Analysis of Least-Squares Attitude Determination Routine DOAOP," Computer Sciences Corp., Rept. CSC/TM-77/6034, Silver Spring, MD, Feb. 1977.

[12] Wertz, J. R., *Spacecraft Attitude Determination and Control*, D. Reidel Publishing, Dordrecht, Holland, The Netherlands, 1978, Sec. 12.2.3, pp. 426–428.

[13] Mortari, D., "Second Estimator of the Optimal Quaternion," *Journal of Guidance, Control, and Dynamics*, Vol. 23, No. 5, Sept.–Oct. 2000, pp. 885–888.

[14] "MATLAB Reference Guide," The MathWorks, Inc., Natick, MA, Oct. 1992.

[15] Mortari, D., Scuro, S. R., and Bruccoleri, C., "Attitude and Orbit Error in *n*-Dimensional Spaces," *Journal of the Astronautical Sciences*, Vol. 54, Nos. 3–4, July–Dec. 2006, pp. 467–484.