AAS 99-427

HOW TO ESTIMATE ATTITUDE FROM VECTOR OBSERVATIONS

F. Landis Markley* and Daniele Mortari†

The most robust estimators minimizing Wahba's loss function are Davenport's q method and the Singular Value Decomposition method. The q method, which computes the optimal quaternion as the eigenvector of a symmetric 4×4 matrix with the largest eigenvalue, is somewhat faster. The fastest algorithms, the QUaternion ESTimator (QUEST) and the EStimators of the Optimal Quaternion (ESOQ and ESOQ2), are less robust since they solve the characteristic polynomial equation for the maximum eigenvalue. This is only an issue for measurements with widely differing accuracies, so these estimators are well suited to star trackers that track multiple stars with comparable accuracies.

WAHBA'S PROBLEM

In many spacecraft attitude systems, the attitude observations are naturally represented as unit vectors. Typical examples are the unit vectors giving the direction to the sun or a star and the unit vector in the direction of the Earth's magnetic field. Almost all algorithms for estimating spacecraft attitude from vector measurements are based on minimizing a loss function proposed in 1965 by Grace Wahba[1]:

Wahba's problem is to find the orthogonal matrix A with determinant +1 that minimizes the loss function

$$L(A) \equiv \tfrac{1}{2}\sum_i a_i \left|\mathbf{b}_i - A\mathbf{r}_i\right|^2 . \tag{1}$$

where $\{\mathbf{b}_i\}$ is a set of unit vectors measured in a spacecraft's body frame, $\{\mathbf{r}_i\}$ are the corresponding unit vectors in a reference frame, and $\{a_i\}$ are non-negative weights. In this paper we choose the weights to be inverse variances, $a_i = \sigma_i^{-2}$, in order to relate Wahba's problem to Maximum Likelihood Estimation[2]. This choice differs from that of Wahba and many other authors, who assumed the weights normalized to unity.

The purpose of this paper is to give an overview of the most popular and most promising algorithms in a unified notation, and to provide accuracy and speed comparisons.

ORTHOGONAL PROCRUSTES PROBLEM

It is possible and has proven very convenient to write the loss function as

$$L(A) = \sum_i a_i - \mathrm{tr}(AB^{\mathrm{T}}) \tag{2}$$

with

$$B \equiv \sum_i a_i \mathbf{b}_i \mathbf{r}_i^{\mathrm{T}} . \tag{3}$$

Now it is clear that $L(A)$ is minimized when the trace, $\mathrm{tr}(AB^{\mathrm{T}})$, is maximized.

This has a close relation to the orthogonal Procrustes problem, which is to find the orthogonal matrix A that is closest to B in the sense of the Frobenius (or Euclidean, or Schur, or Hilbert-Schmidt) norm[3]

$$\left\| M \right\|_{\mathrm{F}}^2 \equiv \sum_{i,j} M_{ij}^2 = \mathrm{tr}(MM^{\mathrm{T}}) . \tag{4}$$

* Guidance, Navigation, and Control Center, NASA's Goddard Space Flight Center, Greenbelt, MD

† Aerospace Engineering School, University of Rome "*La Sapienza*," Rome, Italy

Now

$$\|A-B\|_F^2 = \|A\|_F^2 + \|B\|_F^2 - 2\,\mathrm{tr}(AB^T) = 3 + \|B\|_F^2 - 2\mathrm{tr}(AB^T)\ , \tag{5}$$

so Wahba's problem is equivalent to the orthogonal Procrustes problem with the proviso that the determinant of A must be +1.

FIRST SOLUTIONS

J. L. Farrell and J. C. Stuelpnagel presented the first solution of Wahba's problem [4]. They noted that any real matrix, including B, has the polar decomposition

$$B = WH, \tag{6}$$

where W is orthogonal and H is symmetric and positive semidefinite. Then H can be diagonalized by

$$H = VDV^T, \tag{7}$$

where V is orthogonal and D is diagonal with elements arranged in decreasing order. The optimal attitude estimate is then given by

$$A_{opt} = WV\,\mathrm{diag}[1\ \ 1\ \ \det W]\ V^T. \tag{8}$$

In most cases, $\det W$ is positive and $A_{opt} = W$, but this is not guaranteed.

R. H. Wessner proposed the alternate solution[5]:

$$A_{opt} = (B^T)^{-1}(B^TB)^{1/2}, \tag{9}$$

which is equivalent to

$$A_{opt} = B(B^TB)^{-1/2}. \tag{10}$$

Equations (9) and (10) have the disadvantage of requiring B to be non-singular, which means that a minimum of three vector observations must be available, although it is well known that two vector observations are sufficient to determine the attitude.

J. R. Velman[6], J. E. Brock[7], R. Desjardins, and Wahba also provided solutions to Wahba's problem.

UNCONSTRAINED LEAST-SQUARES

Wahba's loss function can be minimized without requiring the orthogonality constraint by

$$A_{unconstrained} = B\left(\sum_i a_i \mathbf{r}_i \mathbf{r}_i^T\right)^{-1}. \tag{11}$$

This gives the representation

$$B = A_{unconstrained}\left(\sum_i a_i \mathbf{r}_i \mathbf{r}_i^T\right), \tag{12}$$

which looks like the polar decomposition, but isn't really, because $A_{unconstrained}$ is only approximately orthogonal. Note that Eq. (11), like Eqs. (9) and (10), requires three vectors, while only two are really necessary. This solution was proposed by Brock[8] and has been analyzed by Markley and Bar-Itzhack[9].

DAVENPORT'S *q* METHOD

None of the early solutions of Wahba's problem was widely applied, to our knowledge. Paul Davenport provided the real breakthrough in applying Wahba's problem to spacecraft attitude determination[10,11].

We can parameterize A by a unit quaternion[12,13]

$$q = \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix}, \quad \text{where} \quad |q|^2 = 1, \tag{13}$$

as

$$A = (q_4^2 - |\mathbf{q}|^2)I + 2\mathbf{q}\mathbf{q}^T - 2q_4[\mathbf{q}\times]\ . \tag{14}$$

This representation of the attitude matrix is a homogenous quadratic function of q, so we can write

$$\mathrm{tr}(AB^{\mathrm{T}}) = q^{\mathrm{T}} K q \tag{15}$$

where K is the symmetric traceless matrix

$$K \equiv \begin{bmatrix} S - I\,\mathrm{tr}B & \mathbf{z} \\ \mathbf{z}^{\mathrm{T}} & \mathrm{tr}B \end{bmatrix} \tag{16}$$

with

$$S \equiv B + B^{\mathrm{T}} \tag{17}$$

and

$$\mathbf{z} \equiv \begin{bmatrix} B_{23} - B_{32} \\ B_{31} - B_{13} \\ B_{12} - B_{21} \end{bmatrix} = \sum\nolimits_i a_i \mathbf{b}_i \times \mathbf{r}_i \;. \tag{18}$$

It is easy to prove that the optimal unit quaternion is the normalized eigenvector of K with the largest eigenvalue, *i.e.*, the solution of

$$K q_{\mathrm{opt}} \equiv \lambda_{\max} q_{\mathrm{opt}} \;. \tag{19}$$

Very robust algorithms exist to solve the symmetric eigenvalue problem[3,14]. They tend to be slow but are trivial to implement (and fast) in MATLAB. There is no unique solution if the two largest eigenvalues of K are equal. This is not a failure of the q method; it means that the data aren't sufficient to determine the attitude uniquely.

QUATERNION ESTIMATOR (QUEST)

Equation (19) is equivalent to the two equations

$$[(\lambda_{\max} + \mathrm{tr}B)I - S]\mathbf{q} = q_4 \mathbf{z} \tag{20}$$

and

$$(\lambda_{\max} - \mathrm{tr}B) q_4 = \mathbf{q}^{\mathrm{T}} \mathbf{z} \tag{21}$$

Equation (20) gives

$$\mathbf{q} = q_4 [(\lambda_{\max} + \mathrm{tr}B)I - S]^{-1} \mathbf{z} = q_4 \{\mathrm{adj}[(\lambda_{\max} + \mathrm{tr}B)I - S]\mathbf{z}\} / \det[(\lambda_{\max} + \mathrm{tr}B)I - S] \,. \tag{22}$$

The Cayley-Hamilton theorem for a general 3×3 matrix G states that[3]

$$G^3 - (\mathrm{tr}\,G)G^2 + [\mathrm{tr}(\mathrm{adj}G)]G - (\det G)I = 0 \,, \tag{23}$$

where adj$G$ is the classical adjoint (adjugate) of G. This can be used to express the adjoint as

$$\mathrm{adj}\,G = G^2 - (\mathrm{tr}G)\,G + [\mathrm{tr}(\,\mathrm{adj}G)]\,I \,. \tag{24}$$

In particular

$$\mathrm{adj}\,[(\lambda_{\max} + \mathrm{tr}B)I - S] = \alpha I + \beta S + S^2 \,, \tag{25}$$

where

$$\alpha \equiv \lambda_{\max}^2 - (\mathrm{tr}\,B)^2 + \mathrm{tr}(\mathrm{adj}\,S) \tag{26}$$

and

$$\beta \equiv \lambda_{\max} - \mathrm{tr}B. \tag{27}$$

We also have

$$\gamma \equiv \det[(\lambda_{\max} + \mathrm{tr}B)I - S] = \alpha(\lambda_{\max} + \mathrm{tr}\,B) - \det S \,. \tag{28}$$

The optimal quaternion is then given by[15,16]

$$q_{\mathrm{opt}} = \frac{1}{\sqrt{\gamma^2 + |\mathbf{x}|^2}} \begin{bmatrix} \mathbf{x} \\ \gamma \end{bmatrix}, \tag{29}$$

where

$$\mathbf{x} \equiv (\alpha I + \beta S + S^2)\mathbf{z}. \tag{30}$$

All these computations require knowledge of $\lambda_{\max}$. This is obtained by substituting Eq. (22) into Eq. (21), which yields the equation:

$$0 = \psi(\lambda_{\max}) \equiv \gamma(\lambda_{\max} - \mathrm{tr}B) - \mathbf{z}^{\mathrm{T}}(\alpha I + \beta S + S^2)\mathbf{z}. \tag{31}$$

Substituting Eqs. (26–28) gives a fourth-order equation for $\lambda_{\max}$. This is simply the characteristic equation $\det(K - \lambda_{\max} I) = 0$, which can be solved analytically. Shuster observed, however, that $\lambda_{\max}$ is very close to

$$\lambda_0 \equiv \sum_i a_i \tag{32}$$

if the optimized loss function

$$L(A_{\mathrm{opt}}) = \lambda_{\mathrm{o}} - \lambda_{\max} \tag{33}$$

is small, so that $\lambda_{\max}$ can be easily obtained by Newton-Raphson iteration, starting from λ_0 as the initial estimate[15,16]. In fact, a single iteration is generally sufficient. But numerical analysts know that solving the characteristic equation is one of the *worst* ways to find eigenvalues, in general, so QUEST is in principle less robust than Davenport's original q method.

The optimal quaternion is not defined by Eq. (29) if

$$\gamma^2 + |\mathbf{x}|^2 = 0, \tag{34}$$

so Shuster devised the method of sequential rotations to handle this case[15–17]. It is desirable to have a precise criterion for limiting the number of sequential rotations performed, since these are somewhat expensive computationally. Substituting Eq. (30) into Eq. (34) and applying the Cayley-Hamilton theorem twice to eliminate S^4 and S^3 gives, after some tedious algebra,

$$\gamma^2 + |\mathbf{x}|^2 = \gamma(\mathrm{d}\psi / \mathrm{d}\lambda), \tag{35}$$

where $\psi(\lambda)$ is the quartic function defined implicitly by Eq. (31). It can be shown that $\mathrm{d}\psi / \mathrm{d}\lambda$ is invariant under rotations, and this quantity must be nonzero for the Newton-Raphson iteration for $\lambda_{\max}$ to be successful. The singular condition of Eq. (34) is thus seen to be equivalent to $\gamma = 0$, which means that $(q_{\mathrm{opt}})_4 = 0$ and the optimal attitude represents a 180° rotation. We can always use sequential rotations to find a γ such that

$$(q_{\mathrm{opt}})_4 > q_{\min} \tag{36}$$

for any $q_{\min}$ in (0, 1/2), by insisting that

$$\gamma > q_{\min}^2(\mathrm{d}\psi / \mathrm{d}\lambda). \tag{37}$$

In practice, $q_{\min} = 0.1$ is adequate to avoid loss of significance in the computation.

Shuster also provided an estimate of the covariance of the rotation angle error vector in the body frame,

$$P = \left[\sum_i a_i (I - \mathbf{b}_i \mathbf{b}_i^{\mathrm{T}})\right]^{-1}. \tag{38}$$

and showed that the optimized loss function $L(A_{\mathrm{opt}})$ obeys a chi-square probability distribution to a good approximation, assuming Gaussian measurement errors. This provides a useful data quality check. QUEST, first applied in the MAGSAT mission in 1979, is the most widely used algorithm for Wahba's problem.

SINGULAR VALUE DECOMPOSITION (SVD) METHOD

The matrix B has the Singular Value Decomposition[3]:

$$B = U\Sigma V^{\mathrm{T}} = U \,\mathrm{diag}[\Sigma_{11} \;\; \Sigma_{22} \;\; \Sigma_{33}]\, V^{\mathrm{T}}, \tag{39}$$

where U and V are orthogonal, and the singular values obey the inequalities $\Sigma_{11} \geq \Sigma_{22} \geq \Sigma_{33} \geq 0$. Then

$$\operatorname{tr}(AB^{\mathrm{T}}) = \operatorname{tr}(A\,V\,\operatorname{diag}[\Sigma_{11}\;\;\Sigma_{22}\;\;\Sigma_{33}]\,U^{\mathrm{T}}) = \operatorname{tr}(U^{\mathrm{T}}A\,V\,\operatorname{diag}[\Sigma_{11}\;\;\Sigma_{22}\;\;\Sigma_{33}]\,). \tag{40}$$

The trace is maximized, consistent with the constraint $\det A = 1$, by

$$U^T A_{\mathrm{opt}}\,V = \operatorname{diag}[1\quad 1\quad (\det U)(\det V)]\;, \tag{41}$$

which gives the optimal attitude matrix[18,19]:

$$A_{\mathrm{opt}} = U\,\operatorname{diag}[1\quad 1\quad (\det U)(\det V)]\,V^{\mathrm{T}}. \tag{42}$$

The SVD solution is completely equivalent to the original solution by Farrell and Stuelpnagel, since Eq. (42) is identical to Eq. (8) with $U = WV$. The difference is that robust SVD algorithms exist now[3,14]. In fact, computing the SVD is one of the most robust numerical algorithms.

It is convenient to define

$$s_1 \equiv \Sigma_{11},\; s_2 \equiv \Sigma_{22},\;\text{ and } s_3 \equiv (\det U)(\det V)\,\Sigma_{33}, \tag{43}$$

so that $s_1 \geq s_2 \geq |s_3|$. The attitude error covariance is given by

$$P = U\,\operatorname{diag}[(s_2+s_3)^{-1}\;\;(s_3+s_1)^{-1}\;\;(s_1+s_2)^{-1}]\,U^{\mathrm{T}}. \tag{44}$$

The eigenvalues of Davenport's K matrix, $\lambda_{\max} \equiv \lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$, are related to the singular values by[10]

$$\lambda_1 = s_1 + s_2 + s_3,\quad \lambda_2 = s_1 - s_2 - s_3,\quad \lambda_3 = -\,s_1 + s_2 - s_3,\quad \lambda_4 = -\,s_1 - s_2 + s_3. \tag{45}$$

The eigenvalues sum to zero because K is traceless. The singularity (unobservability) condition, which is the condition of infinite covariance, is

$$s_2 + s_3 = 0. \tag{46}$$

This is equivalent to $\lambda_1 = \lambda_2$, the previously-stated unobservability condition for Davenport's q method.

FAST OPTIMAL ATTITUDE MATRIX (FOAM)

The SVD decomposition of B gives a convenient representation for adj$B$, det$B$, and $\|B\|_{\mathrm{F}}^2$. These can be used to write the optimal attitude matrix as[20,21]

$$A_{\mathrm{opt}} = (\kappa\lambda_{\max} - \det B)^{-1}[(\kappa + \|B\|_{\mathrm{F}}^2)B + \lambda_{\max}\operatorname{adj}B^{\mathrm{T}} - BB^{\mathrm{T}}B]\,, \tag{47}$$

where

$$\kappa \equiv \tfrac{1}{2}(\lambda_{\max}^2 - \|B\|_{\mathrm{F}}^2). \tag{48}$$

It's important to note that all the quantities in Eqs. (47) and (48) can be computed without performing the SVD of B. In this method, $\lambda_{\max}$ is found from

$$\lambda_{\max} = \operatorname{tr}(A_{\mathrm{opt}}B^{\mathrm{T}}) = (\kappa\lambda_{\max} - \det B)^{-1}\Big[\big(\kappa + \|B\|_{\mathrm{F}}^2\big)\|B\|_{\mathrm{F}}^2 + 3\lambda_{\max}\det B - \operatorname{tr}\big(BB^{\mathrm{T}}BB^{\mathrm{T}}\big)\Big], \tag{49}$$

or

$$0 = \psi(\lambda_{\max}) \equiv (\lambda_{\max}^2 - \|B\|_{\mathrm{F}}^2)^2 - 8\lambda_{\max}\det B - 4\|\operatorname{adj}B\|_{\mathrm{F}}^2, \tag{50}$$

which follows from some matrix algebra. Equations (31) and (50) for $\psi(\lambda_{\max})$ would be numerically identical with infinite-precision computations, but the FOAM form of the coefficients is less subject to errors arising from cancellations in finite-precision computations.

The FOAM algorithm gives the convenient form for the error covariance:

$$P = (\kappa\lambda_{\max} - \det B)^{-1}(\kappa I + BB^{\mathrm{T}}). \tag{51}$$

ESTIMATOR OF THE OPTIMAL QUATERNION (ESOQ)

Davenport's eigenvalue equation, Eq. (19), says that the optimal quaternion is orthogonal to all the columns of the matrix $K - \lambda_{\max} I$, which means that it must be in the one-dimensional subspace orthogonal to the subspace spanned by any three columns of $K - \lambda_{\max} I$. The optimal quaternion is conveniently computed as the generalized four-dimensional cross-product of any three columns of this matrix[22-24]. This is possible, of course, because the four columns of $K - \lambda_{\max} I$ are not linearly independent.

Another way of seeing this result is to examine the classical adjoint of $K - \lambda_{\max} I$. Representing K in terms of its eigenvalues and eigenvectors gives

$$\operatorname{adj}(K - \lambda I) = \operatorname{adj}\left[\sum_{i=1}^{4} (\lambda_i - \lambda) q_i q_i^{\mathrm{T}}\right] = \sum_{i=1}^{4} (\lambda_j - \lambda)(\lambda_k - \lambda)(\lambda_l - \lambda) q_i q_i^{\mathrm{T}}, \tag{52}$$

for any scalar λ, where $\{i, j, k, l\}$ is a permutation of $\{1, 2, 3, 4\}$. Setting $\lambda = \lambda_{\max} \equiv \lambda_1$ causes all the terms in this sum to vanish except the first, with the result

$$\operatorname{adj}(K - \lambda_{\max} I) = (\lambda_2 - \lambda_{\max})(\lambda_3 - \lambda_{\max})(\lambda_4 - \lambda_{\max}) q_{\mathrm{opt}} q_{\mathrm{opt}}^{\mathrm{T}} \tag{53}$$

Thus q_{opt} can be computed by normalizing any non-zero column (indexed by k) of $\operatorname{adj}(K - \lambda_{\max} I)$:

$$(q_{\mathrm{opt}})_i = c(-1)^{k+i} \det[(K - \lambda_{\max} I)_{ki}] \quad i = 1, \cdots, 4, \tag{54}$$

where $(K - \lambda_{\max} I)_{ki}$ is the 3×3 matrix obtained by deleting the kth row and ith column from $K - \lambda_{\max} I$, and c is a multiplicative factor determined by normalizing the quaternion. It is desirable to choose the column with the maximum Euclidean norm. Because of the symmetry of K, it is only necessary to examine the diagonal elements of the adjoint to determine which column to use.

SECOND ESTIMATOR OF THE OPTIMAL QUATERNION (ESOQ2)

The relation of the quaternion q_{opt} to the rotation axis $\mathbf{e}$ and rotation angle ϕ is

$$q_{\mathrm{opt}} = \begin{bmatrix} \mathbf{e} \sin(\phi/2) \\ \cos(\phi/2) \end{bmatrix}. \tag{55}$$

Inserting this into Eqs. (20) and (21) gives

$$(\lambda_{\max} - \operatorname{tr} B) \cos(\phi/2) = \mathbf{z}^{\mathrm{T}} \mathbf{e} \sin(\phi/2) \tag{56}$$

and

$$\mathbf{z} \cos(\phi/2) = [(\lambda_{\max} + \operatorname{tr} B) I - S] \mathbf{e} \sin(\phi/2) \tag{57}$$

Multiplying Eq. (57) by $(\lambda_{\max} - \operatorname{tr} B)$ and substituting Eq. (56) gives

$$M \mathbf{e} \sin(\phi/2) = \mathbf{0}, \tag{58}$$

where

$$M \equiv (\lambda_{\max} - \operatorname{tr} B)[(\lambda_{\max} + \operatorname{tr} B) I - S] - \mathbf{z}\mathbf{z}^{\mathrm{T}} = [\mathbf{m}_1 \vdots \mathbf{m}_2 \vdots \mathbf{m}_3]. \tag{59}$$

These computations lose numerical significance if $(\lambda_{\max} - \operatorname{tr} B)$ and $\mathbf{z}$ are close to zero, which would be the case for zero rotation angle. We can always avoid this singular condition by using one of the sequential reference system rotations to ensure that $\operatorname{tr} B$ is less than or equal to zero[15–17]. Then Eq. (58) says that the rotation axis is a null vector of M. The columns of adj M are the cross products of the columns of M:

$$\operatorname{adj} M = [\mathbf{m}_2 \times \mathbf{m}_3 \vdots \mathbf{m}_3 \times \mathbf{m}_1 \vdots \mathbf{m}_1 \times \mathbf{m}_2]. \tag{60}$$

Because M is singular, all these columns are parallel, and all are parallel to the rotation axis $\mathbf{e}$. Thus we set

$$\mathbf{e} = \mathbf{y} / |\mathbf{y}|, \tag{61}$$

where $\mathbf{y}$ is the column of $\operatorname{adj} M$ (*i.e*., the cross product) with maximum norm. It is only necessary to examine the diagonal elements of the adjoint matrix to determine which column to use. The rotation angle is found from Eq. (56) or one of the components of Eq. (57). The use of a rotated reference system to ensure a non-positive tr$B$ makes Eq. (56) the best choice. With Eq. (61), this can be written

$$(\lambda_{\max} - \operatorname{tr} B)|\mathbf{y}|\cos(\phi/2) = (\mathbf{z}\cdot\mathbf{y})\sin(\phi/2), \tag{62}$$

which means that there is some scalar h for which

$$\cos(\phi/2) = h(\mathbf{z}\cdot\mathbf{y}) \tag{63}$$

and

$$\sin(\phi/2) = h(\lambda_{\max} - \operatorname{tr} B)|\mathbf{y}|. \tag{64}$$

Substituting into Eq. (55) gives the optimal quaternion as[25,26]

$$q_{\text{opt}} = \frac{1}{\sqrt{|(\lambda_{\max} - \operatorname{tr} B)\,\mathbf{y}|^2 + (\mathbf{z}\cdot\mathbf{y})^2}} \begin{bmatrix} (\lambda_{\max} - \operatorname{tr} B)\,\mathbf{y} \\ \mathbf{z}\cdot\mathbf{y} \end{bmatrix}. \tag{65}$$

Note that it is not necessary to normalize the rotation axis. The rotation of the reference frame is trivially "undone" by permuting quaternion components, as in QUEST[15–17]. ESOQ2 does not define the rotation axis uniquely if M has rank less than two. This includes the usual case of unobservable attitude and also the case of zero rotation angle. Requiring tr$B$ to be non-positive avoids zero rotation angle singularity, however.

TWO-OBSERVATION CASE

If only two observations are used, B is of rank two, so det B = 0. It is clear from Eq. (50) that the characteristic equation is a quadratic equation in $\lambda_{\max}^2$ in this case. The solution can be written as[16]

$$\lambda_{\max} = \sqrt{a_1^2 + a_2^2 + 2a_1 a_2[(\mathbf{b}_1\cdot\mathbf{b}_2)(\mathbf{r}_1\cdot\mathbf{r}_2) + |\mathbf{b}_1\times\mathbf{b}_2||\mathbf{r}_1\times\mathbf{r}_2|]}\ . \tag{66}$$

This speeds up all methods that require a solution of the characteristic equation: namely QUEST, FOAM, ESOQ, and ESOQ2. The simplification in FOAM is especially nice, since it gives an explicit solution for the optimal attitude estimate[20,21]

$$A_{\text{opt}} = \mathbf{b}_3\mathbf{r}_3^{\mathrm{T}} + (a_1/\lambda_{\max})[\mathbf{b}_1\mathbf{r}_1^{\mathrm{T}} + (\mathbf{b}_1\times\mathbf{b}_3)(\mathbf{r}_1\times\mathbf{r}_3)^{\mathrm{T}}] + (a_2/\lambda_{\max})[\mathbf{b}_2\mathbf{r}_2^{\mathrm{T}} + (\mathbf{b}_2\times\mathbf{b}_3)(\mathbf{r}_2\times\mathbf{r}_3)^{\mathrm{T}}], \tag{67}$$

where

$$\mathbf{b}_3 \equiv (\mathbf{b}_1\times\mathbf{b}_2)/|\mathbf{b}_1\times\mathbf{b}_2| \tag{68}$$

and

$$\mathbf{r}_3 \equiv (\mathbf{r}_1\times\mathbf{r}_2)/|\mathbf{r}_1\times\mathbf{r}_2|. \tag{69}$$

It should be emphasized that the minimized loss function, Eq. (1), contains only two observations. The "pseudo-observation" represented by $\mathbf{b}_3$ and $\mathbf{r}_3$ arises automatically from the term $\operatorname{adj} B^{\mathrm{T}}$ in FOAM. Equation (67) goes over to the TRIAD[11,16,27,28] solution for $a_1 = 0$, $a_2 = 0$, or $a_1 = a_2$.

SEQUENTIAL METHODS: Filter QUEST

When observations are obtained over a range of times, it is often convenient to employ a filter that propagates the attitude information from the past to the current time and then adds the information from current measurements. Shuster pointed out that the nine components of the "attitude profile matrix" B contain full information about the attitude (with three degrees of freedom) and the angular error covariance (with six independent components)[2]. He proposed the Filter QUEST algorithm[29], based on propagating and updating B:

$$B(t_k) = \mu\,\Phi_{3\times3}(t_k, t_{k-1})B(t_{k-1}) + \sum_i a_i\mathbf{b}_i\mathbf{r}_i^{\mathrm{T}}\ , \tag{70}$$

where $\Phi_{3\times3}(t_k, t_{k-1})$ is the state transition matrix for the attitude matrix, $\mu < 1$ is a "fading memory" factor, and the sum is over observations at time t_k.

Recursive QUEST

An alternative sequential algorithm, Recursive QUEST or REQUEST, propagates and updates Davenport's K matrix by[30]

$$K(t_k) = \mu \Phi_{4\times4}(t_k, t_{k-1}) K(t_{k-1}) \Phi^{\mathrm{T}}_{4\times4}(t_k, t_{k-1}) + \sum_i a_i \tilde{K}_i \,, \tag{71}$$

where $\Phi_{4\times4}(t_k, t_{k-1})$ is the quaternion state transition matrix and $\tilde{K}_i$ is the Davenport matrix for a single observation,

$$\tilde{K}_i \equiv \begin{bmatrix} \mathbf{b}_i\mathbf{r}_i^{\mathrm{T}} + \mathbf{r}_i\mathbf{b}_i^{\mathrm{T}} - (\mathbf{b}_i \cdot \mathbf{r}_i)I & (\mathbf{b}_i \times \mathbf{r}_i) \\ (\mathbf{b}_i \times \mathbf{r}_i)^{\mathrm{T}} & \mathbf{b}_i \cdot \mathbf{r}_i \end{bmatrix}. \tag{72}$$

Filter QUEST and REQUEST are mathematically equivalent, but Filter QUEST requires fewer computations. Neither has been competitive with an extended Kalman filter in practice, largely due to the suboptimality of the fading memory approximation to the effect of process noise.

The Quaternion Projection Algorithm of Reynolds

Reynolds has proposed a sequential algorithm based on projections in quaternion space, which begins with the observation that the quaternion using the minimum-angle rotation to map the reference vector $\mathbf{r}_i$ into the body frame vector $\mathbf{b}_i$ is

$$q_1 \equiv \frac{1}{\sqrt{2(1+\mathbf{b}_i \cdot \mathbf{r}_i)}} \begin{bmatrix} \mathbf{b}_i \times \mathbf{r}_i \\ 1 + \mathbf{b}_i \cdot \mathbf{r}_i \end{bmatrix}. \tag{73}$$

The most general quaternion that maps $\mathbf{r}_i$ into $\mathbf{b}_i$ is

$$q = \begin{bmatrix} \mathbf{b}_i \sin(\vartheta_b / 2) \\ \cos(\vartheta_b / 2) \end{bmatrix} \otimes q_1 \otimes \begin{bmatrix} \mathbf{r}_i \sin(\vartheta_r / 2) \\ \cos(\vartheta_r / 2) \end{bmatrix} = \left(\cos\frac{\vartheta}{2}\right) q_1 + \left(\sin\frac{\vartheta}{2}\right) q_2 \,, \tag{74}$$

where ϑ_b and ϑ_r are arbitrary angles of rotation about $\mathbf{b}_i$ and $\mathbf{r}_i$, respectively, $\vartheta = \vartheta_b + \vartheta_r$, and

$$q_2 \equiv \frac{1}{\sqrt{2(1+\mathbf{b}_i \cdot \mathbf{r}_i)}} \begin{bmatrix} \mathbf{b}_i + \mathbf{r}_i \\ 0 \end{bmatrix}. \tag{75}$$

The quaternion q_2 maps $\mathbf{r}_i$ into $\mathbf{b}_i$ by means of a 180° rotation about the bisector of $\mathbf{b}_i$ and $\mathbf{r}_i$. The order of quaternion multiplication in Eq. (74) results from using Shuster's quaternion product convention[13,31]:

$$p \otimes q = \begin{bmatrix} \mathbf{p} \\ p_4 \end{bmatrix} \otimes \begin{bmatrix} \mathbf{q} \\ q_4 \end{bmatrix} = \begin{bmatrix} q_4\mathbf{p} + p_4\mathbf{q} - \mathbf{p} \times \mathbf{q} \\ p_4 q_4 - \mathbf{p} \cdot \mathbf{q} \end{bmatrix}. \tag{76}$$

Equation (74) expresses q as a linear combination of the two orthogonal quaternions q_1 and q_2, which constitute an orthogonal basis for the two-dimensional subspace of four-dimensional quaternion space consistent with the measurement. The projection matrix onto this two-dimensional subspace is

$$q_1 q_1^{\mathrm{T}} + q_2 q_2^{\mathrm{T}} = \tfrac{1}{2}(I + \tilde{K}_i) \,. \tag{77}$$

where $\tilde{K}_i$ is defined by Eq. (72). Note that $\tilde{K}_i$ is nonsingular if $\mathbf{b}_i \cdot \mathbf{r}_i = -1$, even though both q_1 and q_2 are undefined in this case. The Reynolds algorithm updates the quaternion estimate by[32,33]

$$q^+ = \left[(I + \eta\tilde{K}_i)q^-\right] \Big/ \left|(I + \eta\tilde{K}_i)q^-\right|, \tag{78}$$

where q^- is the pre-update quaternion estimate propagated to the time of the current measurement, and η is an update gain. We set $\eta = 1$ for perfect measurements to project onto the desired subspace, and $0 < \eta < 1$ for filtering. An extended exercise in quaternion algebra shows that Eq. (78) is equivalent to

$$q^+ = \delta q \otimes q^- \,, \tag{79}$$

where

$$\delta q \equiv \frac{1}{\sqrt{1+2\eta\,\mathbf{b}_i\cdot[A(q^-)\mathbf{r}_i]+\eta^2}}\begin{bmatrix}\eta\mathbf{b}_i\times[A(q^-)\mathbf{r}_i]\\ 1+\eta\mathbf{b}_i\cdot[A(q^-)\mathbf{r}_i]\end{bmatrix}. \tag{80}$$

For $\eta = 1$, δq is the quaternion that takes $A(q^-)\mathbf{r}_i$ into $\mathbf{b}_i$ using the minimal rotation. In the opposite limit of small η, Eq. (79) looks like a Kalman filter update with a cross-product measurement model[31]. We note, though, that Eqs. (72) and (78) are significantly less burdensome computationally than Eqs. (79) and (80).

ACCURACY

We test the accuracy of MATLAB implementations of the q method, the SVD method, QUEST, ESOQ, ESOQ2, and FOAM, using simulated data. The q and SVD methods use the MATLAB functions `eig` and `svd`, respectively, and the others were coded using the equations in this paper, with the approximation $\lambda_{\max} = \lambda_0$ (no iterations) or with one or two Newton-Raphson iterations of the quartic equation for $\lambda_{\max}$. We will analyze three test scenarios. The first scenario simulates a single star tracker with a narrow field of view and boresight at $[1, 0, 0]^T$. This is an application for which the QUEST algorithm has been widely used. We assume that the tracker is tracking five stars at

$$\mathbf{b}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix},\quad \mathbf{b}_2 = \begin{bmatrix}0.99712\\0.07584\\0\end{bmatrix},\quad \mathbf{b}_3 = \begin{bmatrix}0.99712\\-0.07584\\0\end{bmatrix},\quad \mathbf{b}_4 = \begin{bmatrix}0.99712\\0\\0.07584\end{bmatrix},\quad \text{and}\quad \mathbf{b}_5 = \begin{bmatrix}0.99712\\0\\-0.07584\end{bmatrix}. \tag{81}$$

We simulate 1000 test cases with uniformly distributed random attitude matrices. We use these attitude matrices to map the five observation vectors to the reference frame, add Gaussian random noise with equal standard deviations of 6 arcseconds per axis to the reference vectors, and then normalize them. The errors are unconventionally applied to the reference vectors rather than the observation vectors so that Eq. (81) will remain valid in the presence of noise.

The loss function was computed with measurement variances in (radians)^2, since this results in $2L(A_{\text{opt}})$ approximately obeying a χ^2 distribution with $2n_{\text{obs}} - 3$ degrees of freedom, where n_{obs} is the number of vector observations[34]. The minimum and maximum values of the loss function in the 1000 test runs were 0.23 and 12.1, respectively. The probability distribution of the loss function is plotted as the solid line in Figure 1, and several values of $P(\chi^2 \mid \nu)$ for $\chi^2 = 2L(A_{\text{opt}})$ and $\nu = 7$ are plotted as circles[35]. The agreement is seen to be excellent.

The estimation errors for the star tracker scenario are presented in Table 1. The roll error (x axis rotation) and pitch/yaw error (rotation about an axis in the y-z plane) in arcseconds are presented separately, since the estimates of pitch and yaw transverse to the star tracker boresight are more accurate than the estimate of the roll rotation about the boresight. The q method and the SVD method should both give the truly optimal solution, since they are based on robust matrix analysis algorithms[3,14]. The q method was taken as optimal by definition, so no estimated-to-optimal differences are presented for that algorithm, and the differences between the SVD and q methods provide an estimate of the computational errors in both methods. In particular, the loss function is computed exactly by both methods, in principle, which means in practice that it is computed to about one part in 10^5. No estimate of the loss function is provided when no update of $\lambda_{\max}$ is performed, accounting for the lack of entries in the loss function column in Table 1 for these cases.

Not all the decimals exhibited are significant, since the results of 1000 different random cases would not agree to more than two decimal places. The extra decimal places are shown to emphasize the fact that although the different algorithms give results that are closer or farther from the optimal estimate, all the algorithms provide estimates that are equally close to the true attitude. The differences between the estimated and optimal values also show that one Newton-Raphson iteration for $\lambda_{\max}$ is always sufficient; a second iteration provides no significant improvement in the estimate for this scenario.

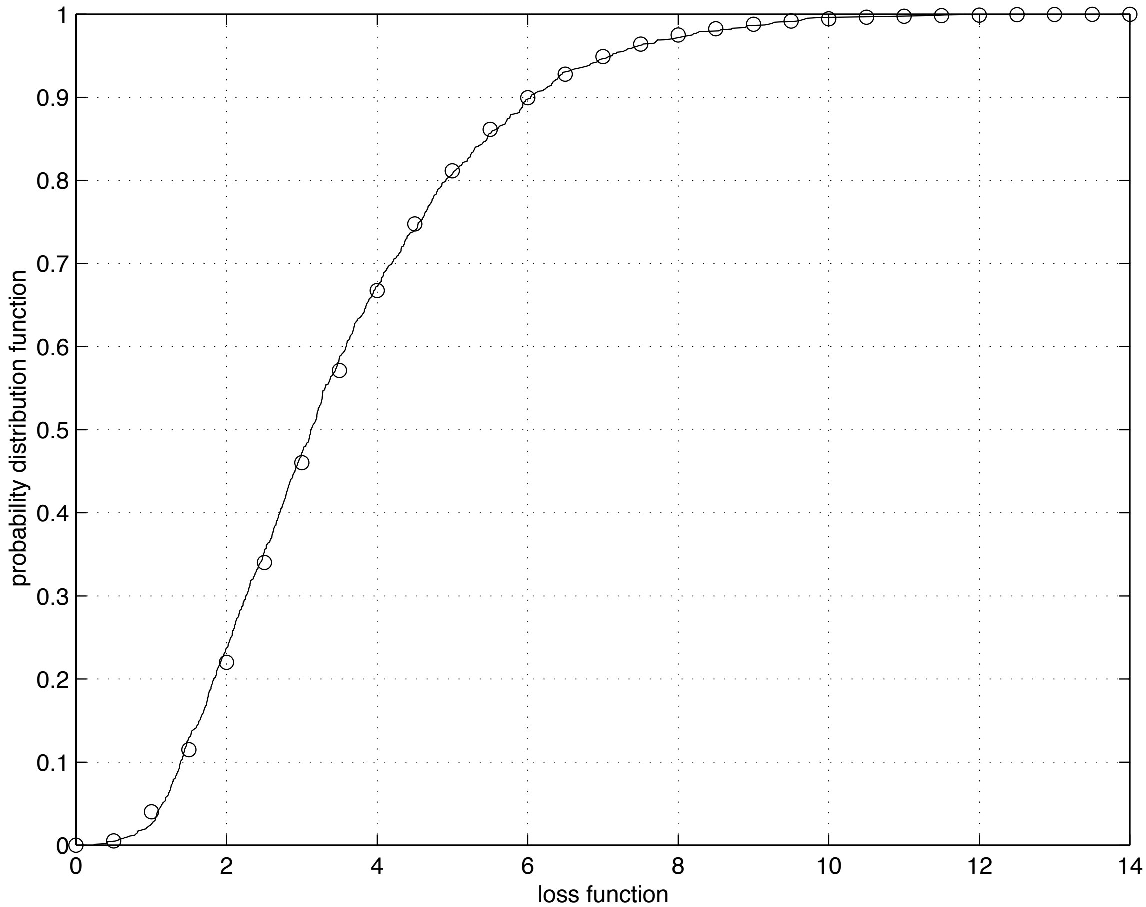

Figure 1: Empirical (solid line) and Theoretical (dots) Loss Function Distribution for the Seven-Degree-of-Freedom Star Tracker Scenario

Table 1: Estimation Errors for Star Tracker Scenario

Algorithm	λ_{max} iterations	RSS (max) estimated-to-optimal			RSS (max) estimated-to-true	
		loss function	x (arcsec)	yz (arcsec)	x (arcsec)	yz (arcsec)
q	–	—	—	—	39.328 (152.83)	3.7895 (13.392)
SVD	–	$0.4\ (1.8) \times 10^{-5}$	$1.5\ (9.4) \times 10^{-8}$	$0.8\ (3.4) \times 10^{-10}$	39.328 (152.83)	3.7895 (13.392)
QUEST	0	—	$1.5\ (12) \times 10^{-2}$	$9.8\ (54) \times 10^{-5}$	39.328 (152.87)	3.7895 (13.392)
	1	$2.5\ (7.2) \times 10^{-5}$	$9.6\ (47) \times 10^{-8}$	$0.6\ (2.5) \times 10^{-9}$	39.328 (152.83)	3.7895 (13.392)
	2	$2.9\ (8.9) \times 10^{-5}$	$11\ (62) \times 10^{-8}$	$0.7\ (3.2) \times 10^{-9}$	39.328 (152.83)	3.7895 (13.392)
ESOQ	0	—	$1.5\ (12) \times 10^{-2}$	$9.8\ (54) \times 10^{-5}$	39.328 (152.87)	3.7895 (13.392)
	1	$2.5\ (7.3) \times 10^{-5}$	$9.4\ (48) \times 10^{-8}$	$1.1\ (6.6) \times 10^{-9}$	39.328 (152.83)	3.7895 (13.392)
	2	$2.9\ (8.9) \times 10^{-5}$	$11\ (55) \times 10^{-8}$	$1.1\ (5.5) \times 10^{-9}$	39.328 (152.83)	3.7895 (13.392)
ESOQ2	0	—	$1.5\ (12) \times 10^{-2}$	$9.8\ (54) \times 10^{-5}$	39.328 (152.87)	3.7895 (13.392)
	1	$2.5\ (7.2) \times 10^{-5}$	$9.5\ (48) \times 10^{-8}$	$0.7\ (2.8) \times 10^{-9}$	39.328 (152.83)	3.7895 (13.392)
	2	$2.9\ (9.3) \times 10^{-5}$	$11\ (55) \times 10^{-8}$	$0.7\ (3.3) \times 10^{-9}$	39.328 (152.83)	3.7895 (13.392)
FOAM	0	—	$1.5\ (12) \times 10^{-2}$	$9.8\ (54) \times 10^{-5}$	39.328 (152.87)	3.7895 (13.392)
	1	$0.4\ (1.9) \times 10^{-5}$	$1.5\ (9.6) \times 10^{-8}$	$2.5\ (12) \times 10^{-9}$	39.328 (152.83)	3.7895 (13.392)
	2	$0.4\ (1.9) \times 10^{-5}$	$1.5\ (9.6) \times 10^{-8}$	$2.6\ (11) \times 10^{-9}$	39.328 (152.83)	3.7895 (13.392)

Equation (38) gives the covariance for the star tracker scenario as

$$P = (6\,\text{arcsec})^2 [5I - \sum_{i=1}^{5} \mathbf{b}_i \mathbf{b}_i^{\mathrm{T}}]^{-1} = \operatorname{diag}[1565, 7.2, 7.2]\,\text{arcsec}^2, \tag{82}$$

which gives the standard deviations of the attitude estimation errors as

$$\sigma_x = \sqrt{1565}\,\text{arcsec} = 40\,\text{arcsec} \quad \text{and} \quad \sigma_{yz} = \sqrt{7.2 + 7.2}\,\text{arcsec} = 3.8\,\text{arcsec}\,. \tag{83}$$

It is apparent that this covariance estimate is quite accurate.

The second scenario uses three observations with widely varying accuracies to provide a difficult test case for the algorithms under consideration. The three observation vectors are

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} -0.99712 \\ 0.07584 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{b}_3 = \begin{bmatrix} -0.99712 \\ -0.07584 \\ 0 \end{bmatrix} \quad . \tag{84}$$

We simulate 1000 test cases as in the star tracker scenario, but with Gaussian noise of one arcsecond per axis on the first observation, and one degree per axis on the other two. This would be the case if the first observation were obtained from an onboard astronomical telescope, and the other two observations were from a coarse sun sensor and a magnetometer, for example. A very accurate estimate of pitch and yaw is required in such an application, but the roll attitude determination is expected to be fairly coarse.

The minimum and maximum values of the loss function computed by the q method in the 1000 test runs were 0.003 and 8.5, respectively. The probability distribution of the loss function is plotted as the solid line in Figure 2, and several values of the χ^2 distribution with three degrees of freedom are plotted as circles. The agreement is almost as good as the seven-degree-of-freedom case.

The estimation errors for this scenario are presented in Table 2, which is similar to Table 1 except that the roll errors are given in degrees. The agreement of the q and SVD method computations is virtually identical to their agreement for the star tracker scenario, but the other algorithms show varying performance. The iterative computation of $\lambda_{\max}$ in QUEST, ESOQ, and ESOQ2 is extraordinarily poor. These algorithms all use the Eq. (50), but with different reference frames, to solve for $\lambda_{\max}$. Surprisingly, this has very little effect on the pitch and yaw determination, but roll determination is affected by an inaccurate computation of $\lambda_{\max}$. The best results of QUEST, ESOQ, and ESOQ2 are all obtained by not performing any iterations for $\lambda_{\max}$; roll determination using the updated $\lambda_{\max}$ is basically useless, with errors attaining their maximum possible value of 180°. The iterative computation of $\lambda_{\max}$ in FOAM is reliable and improves the agreement with the optimal estimate, but does not result in noticeably better agreement with the true attitude.

The failure of the iterative solution for $\lambda_{\max}$ suggests that the analytic solution[35] might be preferable. This is not the case, though, since the errors arise from inaccurate values of the coefficients of the quartic characteristic equation, not from the solution method. In fact, the analytic formulas often give complex solutions in this scenario, which is theoretically impossible for the eigenvalues of a real symmetric matrix. The FOAM computation is more reliable because its characteristic equation coefficients are more accurate.

The predicted covariance in this scenario is, to a very good approximation,

$$P = \operatorname{diag}[\tfrac{1}{2}(1 - 0.99712^2)^{-1}\,\text{deg}^2, 1\,\text{arcsec}^2, 1\,\text{arcsec}^2], \tag{85}$$

which gives

$$\sigma_x = 9.3\,\text{deg} \quad \text{and} \quad \sigma_{yz} = 1.4\,\text{arcsec}\,, \tag{86}$$

in agreement with the best results in Table 2.

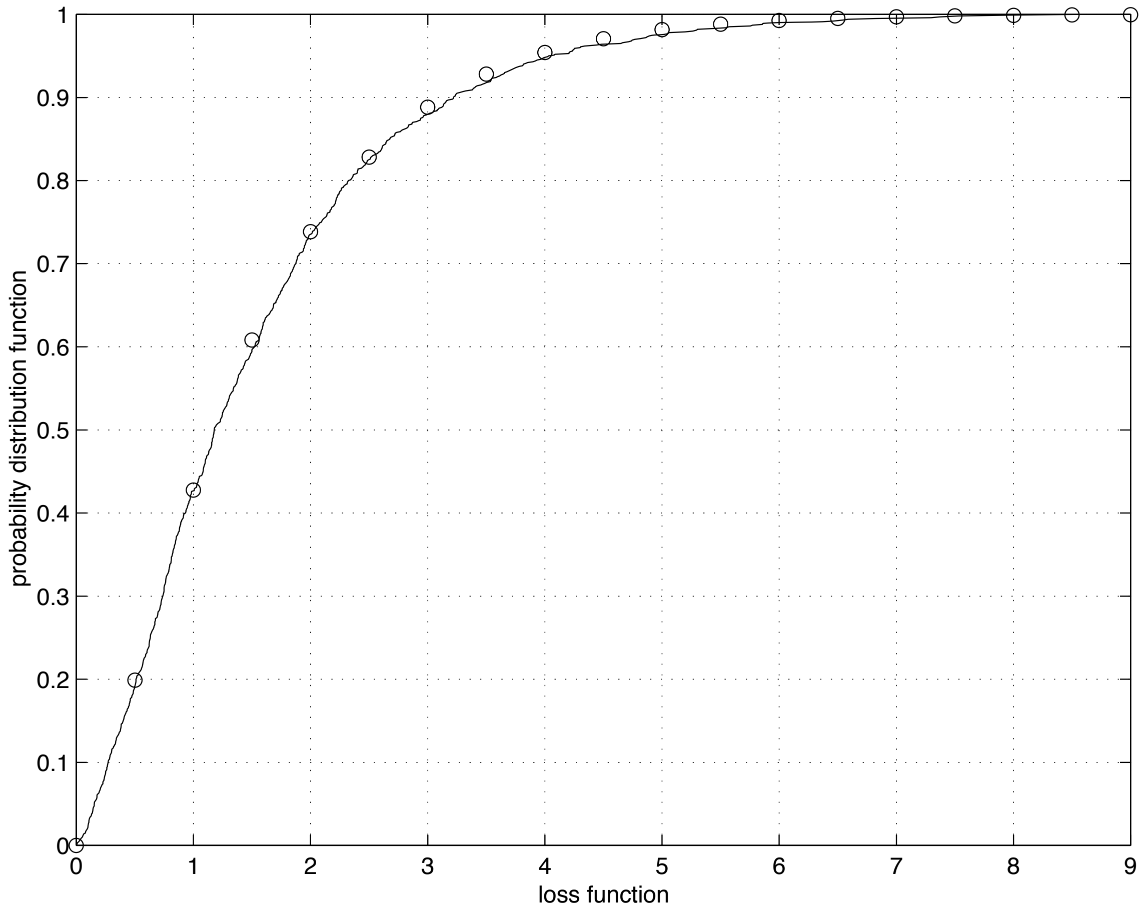

Figure 2: Empirical (solid line) and Theoretical (dots) Loss Function Distribution for the Three-Degree-of-Freedom Unequal Measurement Weight Scenario

Table 2: Estimation Errors for Unequal Measurement Weight Scenario

Algorithm		RSS (max) estimated-to-optimal			RSS (max) estimated-to-true	
$\lambda_{\max}$ iterations		loss function	x (deg)	yz (arcsec)	x (deg)	yz (arcsec)
q	–	—	—	—	9.6 (40)	1.424 (4.519)
SVD	–	$1.7\ (9.3) \times 10^{-5}$	$1.5\ (8.7) \times 10^{-5}$	$0.8\ (3.4) \times 10^{-10}$	9.6 (40)	1.424 (4.519)
QUEST	0	—	1.5 (16)	0.0003 (0.0047)	9.4 (39)	1.424 (4.519)
	1	766 (2752)	55 (180)	0.0024 (0.013)	44 (91)	1.424 (4.520)
	2	11387 (1034977)	55 (180)	0.035 (3.1)	44 (91)	1.425 (4.520)
ESOQ	0	—	1.5 (16)	0.0048 (0.029)	9.4 (39)	1.424 (4.520)
	1	733 (2759)	56 (180)	0.0046 (0.12)	44 (91)	1.424 (4.520)
	2	8847 (786903)	55 (180)	0.026 (2.2)	44 (91)	1.425 (4.520)
ESOQ2	0	—	1.5 (13)	0.0011 (0.011)	9.7 (44)	1.424 (4.519)
	1	734 (2759)	104 (180)	0.0022 (0.011)	65 (92)	1.424 (4.521)
	2	6488 (381771)	104 (180)	0.020 (1.0)	65 (93)	1.425 (4.520)
FOAM	0	—	1.5 (13)	0.0076 (0.042)	9.7 (44)	1.424 (4.512)
	1	0.096 (2.3)	0.088 (1.7)	0.0082 (0.050)	9.6 (40)	1.424 (4.518)
	2	0.0018 (0.14)	0.0015 (0.11)	0.0082 (0.039)	9.6 (40)	1.424 (4.516)

The third scenario investigates the effect of measurement noise mismodeling, illustrating problems that first appeared in analyzing data from the Upper Atmosphere Research Satellite[36]. Of course, no one would intentionally use erroneous models, but it can be very difficult to determine an accurate noise model for real data, and the assumption of any level of white noise is often a poor approximation to real measurement errors. This scenario uses the same three observation vectors as the second scenario, given by Eq. (84). We again simulate 1000 test cases, but with Gaussian white noise of one degree per axis on the first observation and 0.1 degrees per axis on the other two. The estimator, however, incorrectly assumes measurement errors of 0.1 degrees per axis on all three observations, so it weights the measurements equally.

The minimum and maximum values of the loss function computed by the q method in the 1000 test runs were 0.07 and 453, respectively. The probability distribution of the loss function is plotted in Figure 3. The theoretical three-degree-of-freedom distribution is not plotted, since it would be a very poor fit to the data. More than 95% of the values of $L(A_{\mathrm{opt}})$ are theoretically expected to lie below 4, according to the χ^2 distribution plotted in Figure 2, but almost half of the values of the loss function plotted in Figure 3 have values greater than 50. Shuster has emphasized that large values of the loss function are an excellent indication of measurement mismodeling or simply of bad data.

The estimation errors for this scenario are presented in Table 3, which is similar to Tables 1 and 2 except that all the angular errors are given in degrees. The truly optimal q and SVD methods agree even more closely than in the other scenarios. In this scenario, the iterative computation of λ_{max} works well, and both iterations improve the agreement of the loss function and attitude estimates with optimal values. The first order refinement is reflected in a reduction of the roll attitude errors, but no visible improvement in pitch and yaw. All the estimators with first-order updates give estimates equally good as those of the q and SVD methods, and the second order update to λ_{max} provides no significant improvement.

The three scenarios taken together show that the most robust, reliable, and accurate estimators are Davenport's q method and the SVD method. This is not surprising, since these methods are based on robust and well-tested general-purpose matrix algorithms. The FOAM algorithm with one or more iterative refinements of λ_{max} gives equally accurate results in these scenarios. It does not inspire the same level of confidence, though, because it is based on finding a matrix eigenvalue as a solution of the characteristic equation, an operation that is mathematically suspect.

The other algorithms, QUEST, ESOQ, and ESOQ2, perform as well as the more robust algorithms when measurement weights do not vary too widely and are reasonably well modeled. This includes most of the cases for which vector observations are used to compute spacecraft attitude, in particular the case of an attitude solution from multiple stars. In such an application, the largest eigenvalue of the K matrix can either be approximated by λ_0 or else computed by a single Newton-Raphson iteration of the characteristic quartic polynomial. If the measurement uncertainties are not well represented by white noise, however, an update is required, and this update can lead to large errors if the measurement weights span a wide range.

SPEED

There are two caveats to make with regard to timing comparisons. First, absolute speed numbers are not very important for ground computations, since the actual estimation algorithm is only a small part of the overall attitude determination data processing effort. Absolute speed was more important in the past, when thousands of attitude solutions had to be computed by slower machines, which is why QUEST was so important for the MAGSAT mission. Second, the longest time required by a computation may be more important than the average time, since a real-time computer in a spacecraft attitude control system or a star tracker must finish all its required tasks in a limited time. For this reason, we present maximum execution times, which would appear to penalize QUEST for real-time applications, because of its use of sequential rotations. In order to identify the stars in its field of view, however, a star tracker must have a fairly good *a priori* attitude estimate, which can determine the optimal rotated coordinate frame for QUEST in a star tracker application. We will present timing for QUEST both with and without use of *a priori* information to eliminate sequential rotations.

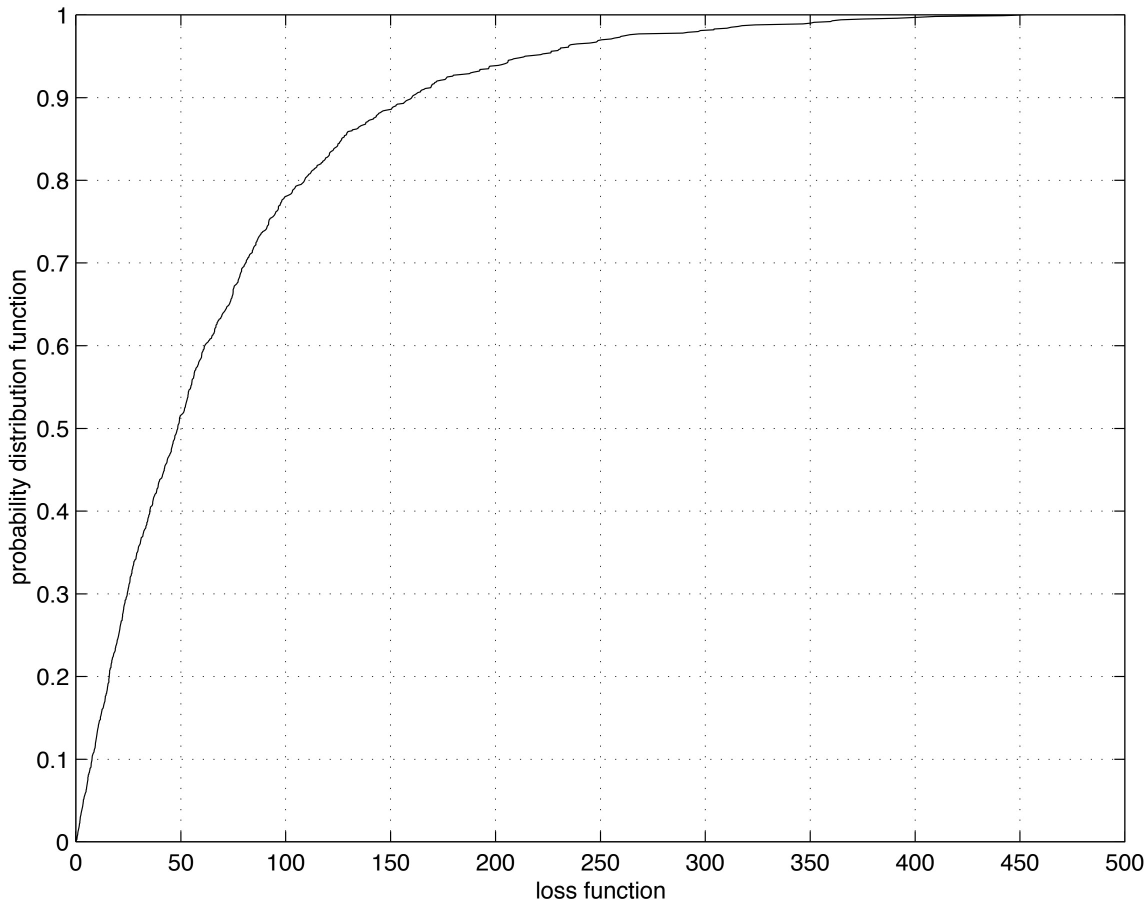

Figure 3: Empirical Loss Function Distribution for the Mismodeled Measurement Weight Scenario

Table 3: Estimation Errors for Mismodeled Measurement Weight Scenario

Algorithm		RSS (max) estimated-to-optimal			RSS (max) estimated-to-true	
$\lambda_{\max}$ iterations		loss function	x (deg)	yz (deg)	x (deg)	yz (deg)
q	–	—	—	—	0.93 (3.64)	0.476 (1.53)
SVD	–	$4.0\ (25) \times 10^{-10}$	$3.7\ (20) \times 10^{-12}$	$2.4\ (8.3) \times 10^{-14}$	0.93 (3.64)	0.476 (1.53)
QUEST	0	—	0.69 (6.2)	0.0038 (0.034)	1.15 (5.60)	0.476 (1.53)
	1	2.7 (57)	0.020 (0.54)	$1.0\ (22) \times 10^{-4}$	0.93 (3.65)	0.476 (1.53)
	2	0.0083 (0.43)	$6.2\ (414) \times 10^{-5}$	$2.4\ (108) \times 10^{-7}$	0.93 (3.64)	0.476 (1.53)
ESOQ	0	—	0.69 (6.2)	0.0038 (0.034)	1.15 (5.60)	0.476 (1.53)
	1	2.7 (57)	0.020 (0.54)	$1.0\ (22) \times 10^{-4}$	0.93 (3.65)	0.476 (1.53)
	2	0.0083 (0.43)	$6.2\ (414) \times 10^{-5}$	$2.4\ (108) \times 10^{-7}$	0.93 (3.64)	0.476 (1.53)
ESOQ2	0	—	0.69 (6.2)	0.0038 (0.034)	1.15 (5.60)	0.476 (1.53)
	1	2.7 (57)	0.020 (0.54)	$1.0\ (22) \times 10^{-4}$	0.93 (3.65)	0.476 (1.53)
	2	0.0083 (0.43)	$6.2\ (414) \times 10^{-5}$	$2.4\ (108) \times 10^{-7}$	0.93 (3.64)	0.476 (1.53)
FOAM	0	—	0.69 (6.2)	0.0038 (0.034)	1.15 (5.60)	0.476 (1.53)
	1	2.7 (57)	0.020 (0.54)	$1.0\ (22) \times 10^{-4}$	0.93 (3.65)	0.476 (1.53)
	2	0.0083 (0.43)	$6.2\ (414) \times 10^{-5}$	$2.4\ (108) \times 10^{-7}$	0.93 (3.64)	0.476 (1.53)

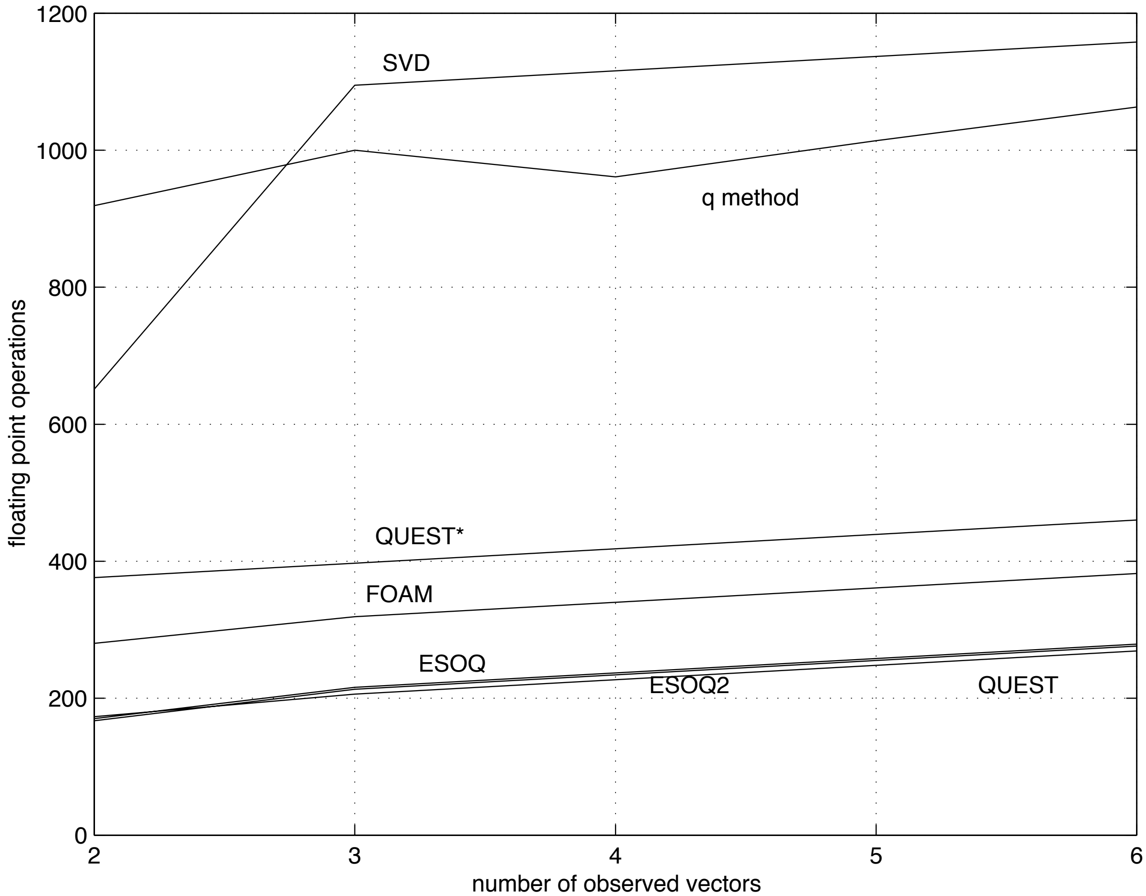

Figure 4: Execution Times for Estimation Algorithms

Figure 4 shows the maximum number of MATLAB floating-point operations (flops) to compute an attitude using two to six reference vectors; the times to process more than six vectors follow the trends seen in the figure. The inputs for the timing tests are the n_{obs} normalized reference and observation vector pairs and their n_{obs} weights. One thousand test cases with random attitudes and random reference vectors with Gaussian measurement noise were simulated for each number of reference vectors. It is clear that the q method and the SVD method require significantly more computational effort than the other algorithms, as expected. The q method is more efficient than the SVD method, except in the least interesting two-observation case. The line labeled QUEST plots the times for the QUEST algorithm with *a priori* input to eliminate the need for sequential rotations, and QUEST* denotes the algorithm without this information, so that the difference shows the maximum time required for sequential rotations. In these tests, QUEST, ESOQ, ESOQ2, and FOAM perform one Newton-Raphson iteration for λ_{max} when processing more than two observations. All these algorithms except QUEST* use the exact quadratic expression for λ_{max} rather than Newton-Raphson iteration in the two-observation case, accounting for the break in the lines at $n_{obs} = 3$ for these methods. QUEST* performs a Newton-Raphson iteration for any number of observations, since it uses the denominator of the update in Eq. (37) to test for sequential rotations. Figure 4 shows that QUEST is less efficient than FOAM when sequential rotations are required, even though times for the more robust FOAM algorithm include 13 flops to compute quaternion output, which is preferable to the nine-component attitude matrix[37]. For this reason, we only consider the version of QUEST that uses *a priori* information in the remaining speed comparisons. One conclusion from these tests is that QUEST, ESOQ, and ESOQ2 are the fastest algorithms, with nearly equal speeds. The relative speeds of these three algorithms can be seen more clearly in the next figure.

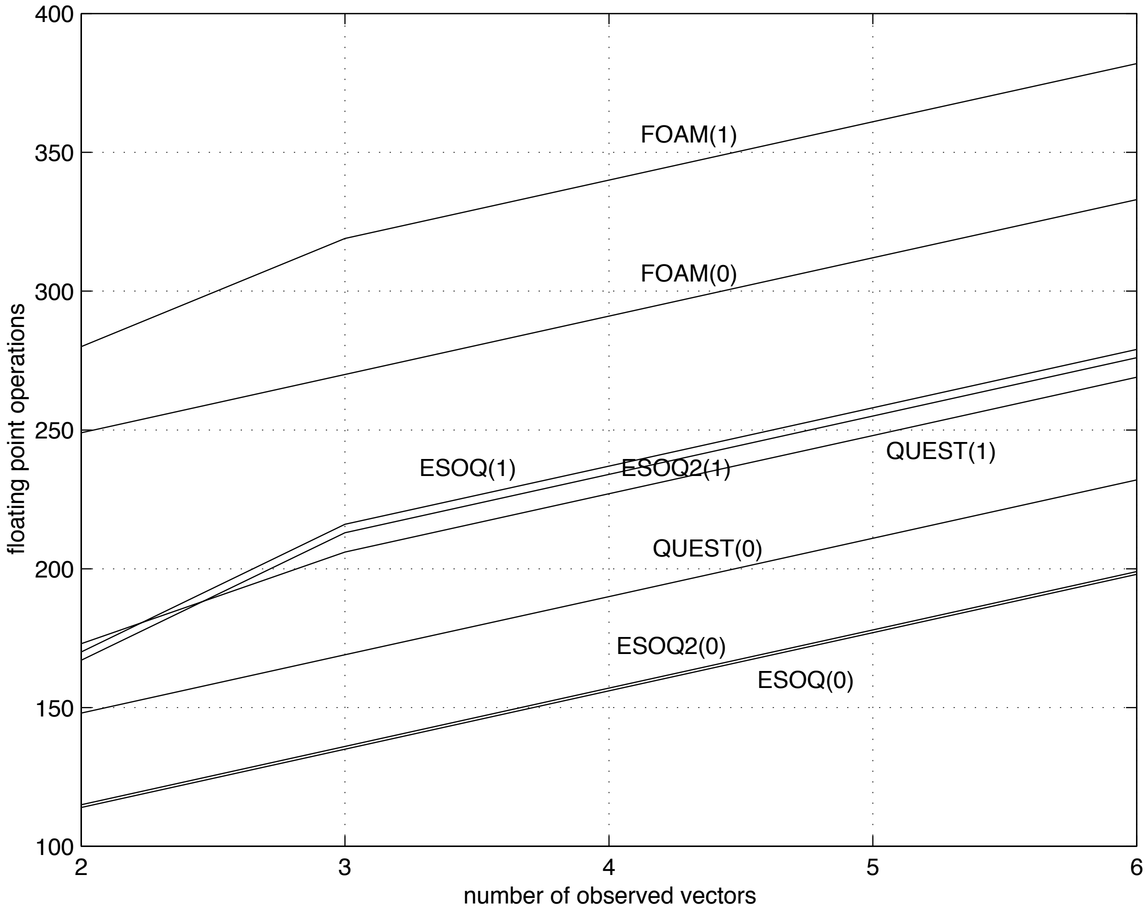

Figure 5: Execution Times for Fastest Estimation Algorithms
The numbers in parentheses indicate the number of Newton-Raphson iterations

Figure 5 compares the timing of QUEST, ESOQ, ESOQ2, and FOAM with one iteration for λ_{max} and with the zeroth-order approximation λ_0, *i.e.* without any iteration. Additional iterations would cost only 11 flops each for any of these algorithms; the main computational cost is in computing the coefficients of the characteristic equation. FOAM is seen to be significantly slower than the other algorithms; in fact FOAM with the zeroth-order approximation for λ_{max} is slower than the other methods with first order updates. Examination of Figure 5 also reveals that the additional cost of the λ_{max} update in QUEST is significantly less than the cost for ESOQ or ESOQ2. This is because QUEST was designed to use some of the same computations in the eigenvalue update and the quaternion computation; ESOQ and ESOQ2 are penalized by using the QUEST update for λ_{max} rather than updates specifically tailored to these algorithms.

SUMMARY

This paper has examined the most useful of algorithms that estimate spacecraft attitude from vector measurements by minimizing Wahba's loss function. The most robust estimators are the *q* method and the Singular Value Decomposition (SVD) method, which incorporate well-tested and mathematically rigorous matrix algorithms. The *q* method, which computes the optimal quaternion as the eigenvector of a symmetric 4×4 matrix with the largest eigenvalue, is somewhat faster than the SVD method. Several algorithms are significantly less burdensome computationally than the *q* and SVD methods. These methods are less robust, since they solve the quartic characteristic polynomial equation for the maximum eigenvalue, a procedure that is potentially numerically unreliable.

The most robust, but not the fastest, of these other algorithms is the Fast Optimal Attitude Matrix (FOAM) algorithm, which performed as well as the *q* and SVD methods in the tests carried out here. The fastest methods are the QUaternion ESTimator (QUEST) and EStimator of the Optimal Quaternion (ESOQ and ESOQ2) algorithms. These algorithms perform quite well when the measurement noise of the observations is well characterized and does not vary too widely from measurement to measurement. They may require Newton-Raphson iteration to solve the quartic characteristic polynomial equation, and this computation can produce large errors when measurements with greatly differing accuracies are combined.

The examples in the paper show that these robustness concerns are only an issue when processing measurements with widely differing accuracies, which is not the case for star trackers that track multiple stars with comparable accuracies, the most common application of Wahba's loss function. Thus the fastest algorithms, QUEST, ESOQ, and ESOQ2, are well suited to star tracker attitude determination applications. In general-purpose applications where weights may vary greatly or the measurement errors are poorly modeled as white noise, the more robust *q* method or FOAM may be preferred.

REFERENCES

1. Wahba, Grace, "A Least Squares Estimate of Spacecraft Attitude," *SIAM Review*, Vol. 7, No. 3, July 1965, p. 409.

2. Shuster, Malcolm D., "Maximum Likelihood Estimate of Spacecraft Attitude," *Journal of the Astronautical Sciences*, Vol. 37, No. 1, January-March 1989, pp. 79–88.

3. Horn, Roger A. and Charles R. Johnson, *Matrix Analysis*, Cambridge, UK, Cambridge University Press, 1985.

4. Farrell, J. L. and J. C. Stuelpnagel, "A Least Squares Estimate of Spacecraft Attitude," *SIAM Review*, Vol. 8, No. 3, July 1966, pp. 384-386.

5. Wessner, R. H., *ibid* .

6. Velman, J. R., *ibid*.

7. Brock, J. E., *ibid*.

8. Brock, John E., "Optimal Matrices Describing Linear Systems," *AIAA Journal*, Vol. 6, No. 7, July 1968, pp. 1292-1296.

9. Markley, F. Landis and Itzhack Y. Bar-Itzhack, "Unconstrained Optimal Transformation Matrix," *IEEE Transactions on Aerospace and Electronic Systems* Vol. 34, No. 1, January 1998, pp. 338-340.

10. Keat, J., "Analysis of Least-Squares Attitude Determination Routine DOAOP," CSC Report CSC/TM-77/6034, February 1977.

11. Lerner, Gerald M., "Three-Axis Attitude Determination," in *Spacecraft Attitude Determination and Control*, ed. by James R. Wertz, Dordrecht, Holland, D. Reidel, 1978.

12. Markley, F. Landis, "Parameterizations of the Attitude," in *Spacecraft Attitude Determination and Control*, ed. by James R. Wertz, Dordrecht, Holland, D. Reidel, 1978.

13. Shuster, Malcolm D., "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, October-December 1993, pp. 439-517.

14. Golub, Gene H. and Charles F. Van Loan, *Matrix Computations*, Baltimore, MD, The Johns Hopkins University Press, 1983.

15. Shuster, M. D. "Approximate Algorithms for Fast Optimal Attitude Computation," AIAA Paper 78-1249, AIAA Guidance and Control Conference, Palo Alto, CA, August 7–9, 1978.

16. Shuster, M. D. and S. D. Oh, "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, January-February 1981, pp. 70-77.

17. Shuster, Malcolm D. and Gregory A. Natanson, "Quaternion Computation from a Geometric Point of View," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, October-December 1993, pp. 545-556.

18. Markley, F. Landis, “Attitude Determination Using Vector Observations and the Singular Value Decomposition,” AAS Paper 87-490, AAS/AIAA Astrodynamics Specialist Conference, Kalispell, MT, August 1987.

19. Markley, F. Landis, “Attitude Determination Using Vector Observations and the Singular Value Decomposition,” *Journal of the Astronautical Sciences*, Vol. 36, No. 3, July-Sept. 1988, pp. 245-258.

20. Markley, F. Landis, “Attitude Determination Using Vector Observations: a Fast Optimal Matrix Algorithm,” Flight Mechanics/Estimation Theory Symposium, Goddard Space Flight Center, Greenbelt, MD, May 1992, NASA Conference Publication 3186.

21. Markley, F. Landis, “Attitude Determination Using Vector Observations: a Fast Optimal Matrix Algorithm,” *Journal of the Astronautical Sciences*, Vol. 41, No. 2, April-June 1993, pp. 261-280.

22. Mortari, Daniele, “ESOQ: A Closed-Form Solution to the Wahba Problem,” Paper AAS 96-173, AAS/AIAA Space Flight Mechanics Meeting, Austin, TX, February 11–15, 1996.

23. Mortari, Daniele, “ESOQ: A Closed-Form Solution to the Wahba Problem,” *Journal of the Astronautical Sciences*, Vol. 45, No.2 April-June 1997, pp. 195-204.

24. Mortari, Daniele, “*n*-Dimensional Cross Product and its Application to Matrix Eigenanalysis,” *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 3, May-June 1997, pp. 509-515.

25. Mortari, Daniele, “ESOQ2 Single-Point Algorithm for Fast Optimal Attitude Determination,” Paper AAS 97–167, AAS/AIAA Space Flight Mechanics Meeting, Huntsville, AL, February 10-12, 1997.

26. Mortari, Daniele, “Second Estimator of the Optimal Quaternion,” *Journal of Guidance, Control, and Dynamics* (in press).

27. Black, H. D. “A Passive System for Determining the Attitude of a Satellite,” *AIAA Journal*, Vol. 2, No. 7, July 1964, pp. 1350-1351.

28. Markley, F. Landis, “Attitude Determination Using Two Vector Measurements,” Flight Mechanics Symposium, Goddard Space Flight Center, Greenbelt, MD, May 1999, NASA Conference Publication NASA/CP-19989-209235, pp. 39-52.

29. Shuster, Malcolm D., “A Simple Kalman Filter and Smoother for Spacecraft Attitude,” *Journal of the Astronautical Sciences*, Vol. 37, No. 1, January-March 1989, pp. 89–106.

30. Bar-Itzhack, Itzhack Y., “REQUEST: A Recursive QUEST Algorithm for Sequential Attitude Determination,” *Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 5, September-October 1996, pp. 1034–1038.

31. Lefferts, E. J., F. L. Markley, and M. D. Shuster, “Kalman Filtering for Spacecraft Attitude Estimation,” *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, Sept.-Oct. 1982, pp. 417-429.

32. Reynolds, R. G., “Quaternion Parameterization and Global Attitude Estimation,” Flight Mechanics/Estimation Theory Symposium, Goddard Space Flight Center, Greenbelt, MD, May 1997, NASA Conference Publication 3345, pp. 347–355.

33. Reynolds, R. G., “Quaternion Parameterization and a Simple Algorithm for Global Attitude Estimation,” *Journal of Guidance, Control, and Dynamics*, Vol. 21, No. 4, July-August 1998, pp. 669-671.

34. Shuster, Malcolm D., private communication.

35. Abramowitz, Milton, and Irene A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, New York, NY, Dover Publications, Inc., 1965, Chapter 26.

36. Deutschmann, J., “Comparison of Single Frame Attitude Determination Methods,” Goddard Space Flight Center Memo to Thomas H. Stengle, July 26, 1993.

37. Markley, F. L., “New Quaternion Attitude Estimation Method,” *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 2, March-April 1994, pp. 407-409.