AAS 08-284

DECENTRALIZED ATTITUDE ESTIMATION USING A QUATERNION COVARIANCE INTERSECTION APPROACH

John L. Crassidis,* Yang Cheng,†
Adam M. Fosbury,‡ and Christopher K. Nebelecky§

This paper derives an approach to combine estimates and covariances for decentralized attitude estimation using a quaternion parameterization. The approach is based on the covariance intersection method, which is modified to maintain quaternion normalization in the combination process. A simple simulation result is provided where local extended Kalman filters are used on two star trackers, each running with common gyro measurements. The covariance intersection approach is shown to provide more accurate estimates than either of the local filters.

INTRODUCTION

Decentralized estimation is an important topic in a data fusion system composed of several processing nodes. The key to a decentralized approach is that, even though communication links may exist between some of the nodes, none of the nodes has knowledge about the overall network topology [1]. This has the advantage on not relying on a common communication system, which upon failure can cause the whole node structure to also be inoperable. Another advantage of decentralized estimation is that nodes can easily be added or deleted in the network without requiring drastic changes to the overall topology. The main disadvantage of decentralized estimation is that since some of the nodes may be using redundant information, their respective state estimates may be correlated and the fusion process cannot assume independence.

A simple example of a decentralized estimation approach involves a spacecraft system that has two star trackers, each running an extended Kalman filter using common gyro measurements. The state vector involves the overall spacecraft attitude and gyro biases. The star observations between the two trackers are clearly independent processes, but since each filter uses common gyro measurements, correlations will exist. The correlations are automatically accounted for in the calculation of the Kalman gain through the cross-correlation covariance terms when a single centralized filter is processing all star observations and gyro measurements simultaneously. However maintaining consistent cross covariances is not possible in a decentralized system where estimates using redundant data are combined. This can yield a covariance that will underestimate the actual errors.

An elegant solution to the consistency problem is the covariance intersection (CI) approach [2]. The authors of this work describe the approach using a geometric interpretation of the Kalman filter,

*Professor, Department of Mechanical & Aerospace Engineering, University at Buffalo, State University of New York, Amherst, NY 14260-4400. E-mail: johnc@eng.buffalo.edu.

†Research Assistant Professor, Department of Mechanical & Aerospace Engineering, University at Buffalo, State University of New York, Amherst, NY 14260-4400. E-mail: cheng3@buffalo.edu.

‡Aerospace Engineer, Air Force Research Laboratory, Space Vehicles Directorate, Kirtland Air Force Base, NM 87117. E-mail: afrl.rvsv@kirtland.af.mil.

§Graduate Student, Department of Mechanical & Aerospace Engineering, University at Buffalo, State University of New York, Amherst, NY 14260-4400. E-mail: ckn@buffalo.edu.

considering the covariance ellipses of a two-dimensional state vector. When the cross covariance is known exactly, the fused estimate's covariance always lays within the intersection of the individual covariances. An analogy here is that the form of the estimate and covariance is identical to the standard Kalman filter when independence is given and generalizes to a colored-noise Kalman filter [3] when there are known nonzero cross correlations. When the cross covariance is unknown, a consistent estimate is one whose covariance encloses the intersection region. Note that a family of solutions is possible and one can be chosen by minimizing the expected errors by some means, such as minimizing the trace of the combined covariance matrix. In the CI approach a scalar weighted average of the covariance matrices is used. When combining two estimates, only a one-dimensional search is required verses one that involves the whole parameter space in the matrix weighted case. Fortunately, the standard CI approach is found to be optimal, in the trace minimization sense, even in the general weighted case [4]. The CI is however conservative in that its error ellipse is larger than the true one. The largest ellipsoid algorithm [5] avoids this by creating the largest ellipse that will fit within the intersection of the covariances, which is always more optimistic than the CI algorithm [5, 6], but consistency is yet to be established for this approach.

An important technology that benefits from decentralized estimation involves spacecraft formation flying, which uses a set of smaller and generally cheaper spacecraft working in cooperation to achieve a mission objective. The first known recorded study of formation flying involves the now well-known concept of using multiple spacecraft to form an interferometer for synthetic aperture applications [7]. References [8] and [9] provide an in-depth survey of guidance and control issues in early and modern day spacecraft formation flying applications. By using decentralized schemes, the entire formation is less vulnerable to individual spacecraft failures both at the estimation and control levels [10]. Most decentralized spacecraft formation flying applications focus on relative position information. For example, [11] shows a study using GPS signals in a decentralized setting and lays the foundation for hierarchic clustering to mitigate scaling problems for larger fleets. Reference [12] employs the CI approach to develop relative spacecraft position and velocity estimates using relative position measurements. A centralized Kalman filter for full state estimation, which includes both relative position and attitude, using only relative line-of-sight observations is shown in [13].

For spacecraft attitude estimation, the four-dimensional quaternion [14] is the attitude parameterization of choice for several reasons: 1) it is free of singularities, 2) the attitude matrix is quadratic in the quaternion components, and 3) the kinematics equations is bilinear and an analytic solution exists for the propagation. However, since a four-dimensional vector is used to describe three dimensions, the quaternion components cannot be independent of each other, which is shown by the fact that the quaternion must have unit norm. This leads to problems when attempting to average a set of quaternions, which is further compounded by the 2:1 mapping of the rotation group. Reference [15] presents an approach for determining the average norm-preserving quaternion from a set of weighted quaternions, which is accomplished by performing an eigenvalue/eigenvector decomposition of a matrix composed of the given quaternions and weights. Independence is inherently implied in the solution. In this paper, the quaternion averaging algorithm is extended to handle appended state vectors. In particular, a new CI combination approach is derived that preserves quaternion normalization during the solution process. The basic idea is to perform the CI operation over the nonlinear manifold of the unit sphere.

The organization of this paper is as follows. First, the CI approach is summarized and then re-derived from a loss-function point of view. Next, a CI approach that fuses a single quaternion and other quantities is derived that maintains normalization of the fused quaternion. A square root version is also derived that provides a better conditioned approach from a numerical viewpoint. Theory to construct a CI approach to handle the case of $n > 1$ quaternions is then developed. Finally, simulation results for the single quaternion case are shown using a two star-tracker system, with each tracker incorporating common gyro measurements in their decentralized nodes.

COVARIANCE INTERSECTION

This section summarizes the CI approach (see [2] for more details), which is rooted in the concept of Gaussian intersection [16]. Consider two estimate-covariance pairs, $\{\mathbf{a}, P_{aa}\}$ and $\{\mathbf{b}, P_{bb}\}$. The true values of each are denoted with an overbar, with

$$\bar{P}_{aa} = E\{\tilde{\mathbf{a}}\,\tilde{\mathbf{a}}^T\}, \quad \bar{P}_{ab} = E\{\tilde{\mathbf{a}}\,\tilde{\mathbf{b}}^T\}, \quad \bar{P}_{bb} = E\{\tilde{\mathbf{b}}\,\tilde{\mathbf{b}}^T\} \tag{1}$$

where $\tilde{\mathbf{a}} \triangleq \mathbf{a} - \bar{\mathbf{a}}$ and $\tilde{\mathbf{b}} \triangleq \mathbf{b} - \bar{\mathbf{b}}$, which are the state errors and $E\{\cdot\}$ is the expectation operator. It is assumed that the estimates for $\mathbf{a}$ and $\mathbf{b}$ are consistent, so that $P_{aa} - \bar{P}_{aa} \geq 0$ and $P_{bb} - \bar{P}_{bb} \geq 0$. This means that $P_{aa} - \bar{P}_{aa}$ and $P_{bb} - \bar{P}_{bb}$ are positive semi-definite matrices. A consistent estimate formed by fusing $\mathbf{a}$ and $\mathbf{b}$ is given by

$$P_{cc}^{-1} = \omega P_{aa}^{-1} + (1-\omega) P_{bb}^{-1} \tag{2a}$$

$$\mathbf{c} = \omega P_{cc} P_{aa}^{-1}\mathbf{a} + (1-\omega) P_{cc} P_{bb}^{-1}\mathbf{b} \tag{2b}$$

where $\omega \in [0,\, 1]$ is a scalar weight. The requirement for ω ensures that the covariance $P_{cc} \geq 0$, $P_{aa} \geq P_{cc}$, and $P_{bb} \geq P_{cc}$. Reference [2] proves that $P_{cc} - \bar{P}_{cc} \geq 0$ for all P_{ab} and ω, where $\bar{P}_{cc} = E\{\tilde{\mathbf{c}}\,\tilde{\mathbf{c}}^T\}$ with $\tilde{\mathbf{c}} \triangleq \mathbf{c} - \bar{\mathbf{c}}$. The weight can be found using a simple optimization scheme that minimizes the trace or the determinant of P_{cc}. The trace and the determinant of P_{cc} characterize the size of the Gaussian uncertainty ellipsoid associated with P_{cc}. In two-dimensional cases, the former is approximately proportional to the squared perimeter of the ellipse and the latter is proportional to the squared area of the ellipse. Consider the identity $\log(\det P_{cc}) = \mathrm{tr}(\log P_{cc})$, where tr is the matrix trace operator. Using the fact that the logarithm function is monotonic, it can be seen that minimizing the determinant of P_{cc} is equivalent to minimizing the trace of the matrix logarithm of P_{cc}, not to minimizing the trace of P_{cc}. Minimizing the trace or the determinant of P_{cc} is a convex optimization problem. This means that the cost function has only one local optimum of ω in the range of $[0, 1]$, which is also the global optimum.

Loss Function Point of View

The CI solution can be determined from a loss function point of view. The usefulness of this perspective will be made clear in the next section. Consider minimizing the following loss function:

$$J(\mathbf{c}) = \omega(\mathbf{c}-\mathbf{a})^T P_{aa}^{-1}(\mathbf{c}-\mathbf{a}) + (1-\omega)(\mathbf{c}-\mathbf{b})^T P_{bb}^{-1}(\mathbf{c}-\mathbf{b}) \tag{3}$$

The loss function is identical to that of fusing two uncorrelated estimates with dilated covariances P_{aa}/ω and $P_{bb}/(1-\omega)$, respectively. Minimizing Eq. (3) with respect to $\mathbf{c}$ results in

$$\omega(\mathbf{c}-\mathbf{a})^T P_{aa}^{-1} + (1-\omega)(\mathbf{c}-\mathbf{b})^T P_{bb}^{-1} = \mathbf{0} \tag{4}$$

Taking the transpose and rearranging yields

$$\left(\omega P_{aa}^{-1} + (1-\omega) P_{bb}^{-1}\right)\mathbf{c} = \omega P_{aa}^{-1}\mathbf{a} + (1-\omega) P_{bb}^{-1}\mathbf{b} \tag{5}$$

Using the definition of P_{cc} from Eq. (2a) we obtain

$$\mathbf{c} = P_{cc}[\omega P_{aa}^{-1}\mathbf{a} + (1-\omega)P_{bb}^{-1}\mathbf{b}] \tag{6}$$

which is identical to Eq. (2b). Note that when $\omega = 0.5$ the loss function is equivalent to maximum likelihood estimation with the assumed independence property applied.

Fusion of Multiple Estimates

It is straightforward to apply the CI approach to fuse multiple estimates. The CI algorithm closely resembles an electrical resistance calculation within a parallel architecture. Given a set of n estimates $\{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}\}$ and associated covariances $\{P_1, P_2, \ldots, P_n\}$, a consistent estimate of the fused estimate and covariance is given by

$$\mathbf{c} = P_{cc}\sum_{i=1}^{n}\omega_i P_i^{-1}\mathbf{x}_i \tag{7a}$$

$$P_{cc}^{-1} = \omega_1 P_1^{-1} + \omega_2 P_2^{-1} + \cdots + \omega_n P_n^{-1} \tag{7b}$$

where the weights satisfy $\sum_{i=1}^{n}\omega_i = 1$ and $\omega_i \in$ [0,1]. The weights ω_i can be found by minimizing the trace or the determinant of P_{cc} subject to the aforementioned constraints. The bounded optimization problem is convex and can be solved efficiently using, for example, CVX, the MATLAB software for disciplined convex programming [17].

ATTITUDE ESTIMATION VIA COVARIANCE INTERSECTION

In this section the CI approach is extended to attitude estimation. The objective is to fuse n attitude estimates with unknown correlations to yield a single quaternion estimate. It is assumed that the i^{th} state vector, $\mathbf{x}_i$, is composed of a quaternion, $\mathbf{q}_i$, and other quantities, $\mathbf{b}_i$, such as gyro biases. A standard multiplicative quaternion Kalman filter is employed, where the covariance matrix, denoted by P_i, is the reduced order form for the small half-attitude errors and errors for the remainder quantities [18]. Clearly, Eq. (7a) cannot be directly employed in this case because the resulting quaternion will not be guaranteed to have unit norm. For simplicity, we assume that the covariance after the CI update is of the form Eq. (7b), independent of the updated state estimate. The optimal weights used in Eq. (7a) are determined by minimizing the trace or the determinant of the covariance of the assumed form.

A method to average quaternions is presented in [15], which also shows its relation to maximum likelihood estimation. The loss function is given by

$$J(\mathbf{q}) = \sum_{i=1}^{n}\mathbf{q}^T\Xi(\mathbf{q}_i)\,P_{qq_i}^{-1}\Xi^T(\mathbf{q}_i)\mathbf{q} \tag{8}$$

subject to the constraint $1 - \mathbf{q}^T\mathbf{q} = 0$. The matrix $\Xi(\mathbf{q})$ is defined by

$$\Xi(\mathbf{q}) \triangleq \begin{bmatrix} q_4 I_{3\times3} + [\boldsymbol{\rho}\times] \\ -\boldsymbol{\rho}^T \end{bmatrix} \tag{9}$$

where $\boldsymbol{\rho}$ denotes the vector part of the quaternion and q_4 is the scalar part, i.e. $\mathbf{q} \triangleq [\boldsymbol{\rho}^T \;\; q_4]^T$. The magnitude of $\Xi^T(\mathbf{q}_i)\mathbf{q}$ is the absolute value of the sine of the half-error angle [15]. The matrix

P_{qq_i} is the 3×3 covariance matrix of the vector part of the error quaternion corresponding to $\mathbf{q}_i$. The solution approach uses a Lagrange multiplier to handle the equality constraint. The average quaternion is given by finding the eigenvector corresponding to the maximum eigenvalue of the matrix

$$\mathcal{M} = -\sum_{i=1}^{n} \Xi(\mathbf{q}_i)\, P_{qq_i}^{-1} \Xi^T(\mathbf{q}_i) \tag{10}$$

A straightforward implementation of the quaternion averaging algorithm cannot be applied to the problem with appended state vectors, i.e. state vectors that include quantities other than the quaternion. To overcome this issue the following function is maximized:

$$J(\Delta\mathbf{x}) = -\sum_{i=1}^{n} \omega_i\, \Delta\mathbf{x}_i^T P_i^{-1} \Delta\mathbf{x}_i \tag{11}$$

where $\sum_{i=1}^{n} \omega_i = 1$, $\omega_i \in$ [0,1] and

$$\mathbf{x} \triangleq \overset{1}{\begin{bmatrix} \mathbf{q} \\ \mathbf{b} \end{bmatrix}} \begin{matrix} 4 \\ n_b \end{matrix} \tag{12a}$$

$$\Delta\mathbf{x}_i \triangleq \overset{1}{\begin{bmatrix} \Xi^T(\mathbf{q}_i)\mathbf{q} \\ \\ \mathbf{b} - \mathbf{b}_i \end{bmatrix}} \begin{matrix} 3 \\ \\ n_b \end{matrix} \tag{12b}$$

$$P_i^{-1} \triangleq \begin{matrix} \quad 3 & \quad n_b \end{matrix} \begin{bmatrix} \mathcal{P}_{qq_i} & \mathcal{P}_{qb_i} \\ \\ \mathcal{P}_{qb_i}^T & \mathcal{P}_{bb_i} \end{bmatrix} \begin{matrix} 3 \\ \\ n_b \end{matrix} \tag{12c}$$

It has been assumed that P_i^{-1} is nonsingular. Note that the vector $\mathbf{b}$ can be of any dimension, denoted by n_b. For spacecraft attitude estimation applications with gyros, this vector may contain a combination of gyro biases, scale factors and misalignment parameters. It is known that $\mathbf{q}_i$ and $-\mathbf{q}_i$ represent the same attitude. However, changing $\mathbf{q}_i$ (or $\mathbf{q}$) to $-\mathbf{q}_i$ (or $-\mathbf{q}$) in Eq. (11) alters the value of $J(\mathbf{x})$ unless $\mathcal{P}_{qb_i}$ is a null matrix. Care therefore needs to be taken in preparing the attitude data. The basic idea is to have all the $\mathbf{q}_i$ point largely to the same direction (here $\mathbf{q}_i$ are treated the same way as the line-of-sight vectors).

The quaternion constraint is handled using the method of Lagrange multipliers. The appended objective function is now

$$J(\Delta\mathbf{x}) = -\sum_{i=1}^{n} \omega_i \Delta\mathbf{x}_i^T P_i^{-1} \Delta\mathbf{x}_i + \lambda(1 - \mathbf{q}^T\mathbf{q}) \tag{13}$$

The necessary conditions for maximization of Eq. (13) are

$$\frac{\partial J}{\partial \mathbf{b}} = -2\sum_{i=1}^{n} \omega_i \{\mathbf{q}^T \Xi(\mathbf{q}_i)\mathcal{P}_{qb_i} + (\mathbf{b} - \mathbf{b}_i)^T \mathcal{P}_{bb_i}\} = \mathbf{0} \tag{14a}$$

$$\frac{\partial J}{\partial \mathbf{q}} = -2\sum_{i=1}^{n} \omega_i \{\mathbf{q}^T \Xi(\mathbf{q}_i)\mathcal{P}_{qq_i}\Xi^T(\mathbf{q}_i) + \Xi(\mathbf{q}_i)\mathcal{P}_{qb_i}(\mathbf{b} - \mathbf{b}_i)\} - 2\lambda\mathbf{q}^T = \mathbf{0} \tag{14b}$$

$$\frac{\partial J}{\partial \lambda} = 1 - \mathbf{q}^T\mathbf{q} = \mathbf{0} \tag{14c}$$

Expanding Eq. (14a), and taking the transpose and solving for **b** yields

$$\mathbf{b} = \mathcal{B}_{bb}^{-1}(\mathbf{d} - \mathcal{B}_{qb}^{T}\mathbf{q}) \tag{15}$$

where the following definitions have been used:

$$\mathcal{B}_{bb} \equiv \sum_{i=1}^{n} \omega_i \mathcal{P}_{bb_i} \tag{16a}$$

$$\mathbf{d} \equiv \sum_{i=1}^{n} \omega_i \mathcal{P}_{bb_i} \mathbf{b}_i \tag{16b}$$

$$\mathcal{B}_{qb} \equiv \sum_{i=1}^{n} \omega_i \Xi(\mathbf{q}_i) \mathcal{P}_{qb_i} \tag{16c}$$

Substituting Eq. (15) into Eq. (14b) with similar manipulations yields

$$(\mathcal{B}_{qq} - \mathcal{B}_{qq}\mathcal{B}_{bb}^{-1}\mathcal{B}_{qq} + \lambda I_{4\times4})\mathbf{q} = \mathbf{c} - \mathcal{B}_{qb}\mathcal{B}_{bb}^{-1}\mathbf{d} \tag{17}$$

where

$$\mathcal{B}_{qq} \equiv \sum_{i=1}^{n} \omega_i \Xi(\mathbf{q}_i) \mathcal{P}_{qq_i} \Xi^{T}(\mathbf{q}_i) \tag{18a}$$

$$\mathbf{c} \equiv \sum_{i=1}^{n} \omega_i \Xi(\mathbf{q}_i) \mathcal{P}_{qb_i} \mathbf{b}_i \tag{18b}$$

The definitions presented in Eqs. (16) and (18) are formed such that **x** can be expressed as

$$\begin{bmatrix} \mathcal{B}_{qq} + \lambda I_{4\times4} & \mathcal{B}_{qb} \\ \\ \mathcal{B}_{qb}^{T} & \mathcal{B}_{bb} \end{bmatrix} \begin{bmatrix} \mathbf{q} \\ \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{c} \\ \\ \mathbf{d} \end{bmatrix} \tag{19}$$

subject to the constraint

$$\mathbf{q}^{T}\mathbf{q} = 1 \tag{20}$$

The matrix $\mathcal{B}$ formed by the elements in Eq. (19) is

$$\mathcal{B} = \begin{bmatrix} \mathcal{B}_{qq} & \mathcal{B}_{qb} \\ \\ \mathcal{B}_{qb}^{T} & \mathcal{B}_{bb} \end{bmatrix} \triangleq \begin{bmatrix} \sum_{i=1}^{n} \omega_i \Xi(\mathbf{q}_i) \mathcal{P}_{qq_i} \Xi^{T}(\mathbf{q}_i) & \sum_{i=1}^{n} \omega_i \Xi(\mathbf{q}_i) \mathcal{P}_{qb_i} \\ \\ \sum_{i=1}^{n} \omega_i \mathcal{P}_{qb_i}^{T} \Xi^{T}(\mathbf{q}_i) & \sum_{i=1}^{n} \omega_i \mathcal{P}_{bb_i} \end{bmatrix} \tag{21}$$

which is a positive semi-definite matrix. It is singular only when all of the $\mathbf{q}_i$ are identical. As will be seen, the motivation for expressing the maximization in the form of Eq. (19) is that it is easily extendible when fusing more then one quaternion.

From Eq. (17), define the following:

$$Z \triangleq \mathcal{B}_{qq} - \mathcal{B}_{qb}\mathcal{B}_{bb}^{-1}\mathcal{B}_{qb}^{T} \tag{22a}$$

$$\mathbf{g} \triangleq \mathbf{c} - \mathcal{B}_{qb}\mathcal{B}_{bb}^{-1}\mathbf{d} \tag{22b}$$

Note that Z is a positive semi-definite matrix. Maximizing the objective function has now been reduced to the solution of the following set of consistent Lagrange equations:

$$(Z + \lambda I_{4\times4})\mathbf{q} = \mathbf{g} \tag{23a}$$

$$\mathbf{q}^{T}\mathbf{q} = 1 \tag{23b}$$

SOLUTION TO THE LAGRANGE EQUATIONS

The Lagrange equations in Eq. (23) have been studied in detail. In this section we consider several solutions.

Secular Equation

First consider an eigenvalue decomposition of $Z = QVQ^T$ where V is a diagonal matrix of eigenvalues, $V \triangleq \mathrm{diag}(\delta_1, \ldots, \delta_4)$, and Q is the associated matrix of eigenvectors satisfying $Q^TQ = QQ^T = I$. Substituting the eignvalue decomposition for Z in Eq. (23a) and rearranging yields

$$QVQ^T\mathbf{q} = -\lambda QQ^T\mathbf{q} + \mathbf{g} \tag{24}$$

If Eq. (24) is pre-multiplied by Q^T and defining the following $[4 \times 1]$ vectors:

$$\mathbf{u} \equiv Q^T\mathbf{q} \tag{25a}$$

$$\mathbf{a} \equiv Q^T\mathbf{g} \tag{25b}$$

then Eq. (23a) becomes

$$V\mathbf{u} = -\lambda\mathbf{u} + \mathbf{a} \tag{26}$$

Because V is diagonal we can now solve for each of the u_i values:

$$u_i = \frac{a_i}{\delta_i + \lambda} \tag{27}$$

Using $\mathbf{q} = Q\mathbf{u}$, the normalization constraint becomes

$$\mathbf{q}^T\mathbf{q} = \mathbf{u}^T\mathbf{u} = \sum_{i=1}^{4}\left(\frac{a_i}{\delta_i + \lambda}\right)^2 = 1 \tag{28}$$

Equation (28) represents an explicit secular function in λ. The explicit secular function is an 8^{th} degree polynomial in λ which must be solved. In [19] it is shown that the optimal λ is the maximum real zero of Eq. (28). In order to solve Eq. (28), a robust root finder is necessary. Once $\lambda_{\max}$ is determined, the quaternion and vector **b** are determined by

$$\mathbf{q} = (Z + \lambda_{\text{max}} I_{4\times4})^{-1}\mathbf{g} \tag{29a}$$

$$\mathbf{b} = \mathcal{B}_{bb}^{-1}(\mathbf{d} - \mathcal{B}_{qb}^T\mathbf{q}) \tag{29b}$$

Note that the preceding approach is fundamentally the same as that used in the extended QUEST algorithm [21].

When λ is solved iteratively, a good initial guess is important for convergence and computational efficiency. A good approximate solution for λ can be found if the correlations between the quaternion and other states are small, i.e. $\mathcal{B}_{qb} = \sum_{i=1}^{n} \omega_i\, \Xi(\mathbf{q}_i)\mathcal{P}_{qb_i}$ is smaller than the other terms in Eqs. (23). An approximate quaternion, denoted by $\mathbf{q}_{\text{app}}$, is given by finding the eigenvector associated with the maximum eigenvalue of the matrix $\mathcal{M} = -\mathcal{B}_{qq}$. Pre-multiplying Eq. (23a) by $\mathbf{q}_{\text{app}}$ and solving for λ gives the approximation

$$\lambda_{\text{app}} = \mathbf{q}_{\text{app}}^T\,\mathbf{g} - \mathbf{q}_{\text{app}}^T\,Z\,\mathbf{q}_{\text{app}} \tag{30}$$

which can be used as a starting guess for the actual λ in an iterative scheme. Note that because the quaternion and its negative represent the same rotation, then Eq. (30) should be checked using both $\mathbf{q}_{\text{app}}$ and $-\mathbf{q}_{\text{app}}$ to see which one produces a higher value of λ_{app}. In many cases, $\lambda_{\text{app}} = 0$ is a good initial guess as well [22]. Other iterative schemes can be found in [23–25].

Quadratic Eigenvalue Problem

Rather than solving an explicit secular function in λ, the Lagrange equations can be reduced to a quadratic eigenvalue problem (QEP) [19]. This is due to the fact that the Lagrange equations are consistent (equality in the norm constraint). If the Lagrange equations are inconsistent, the QEP could still be used in order to define the spectrum for which the solution lays. The QEP is well known because of its many applications to dynamic systems and structural analysis [28]. In many cases one can then reduce the QEP to a standard eigenvalue problem (SEP), for which solution techniques are well known. Begin by solving Eq. (23a) for $\mathbf{q}$ and substituting the result into Eq. (23b), which gives

$$\mathbf{g}^T (Z + \lambda I_{4\times 4})^{-2} \mathbf{g} = 1 \tag{31}$$

Define a new $[4 \times 1]$ vector $\boldsymbol{\gamma}$ as

$$\boldsymbol{\gamma} \triangleq (Z + \lambda I_{4\times 4})^{-2} \mathbf{g} \tag{32}$$

Equation (31) can then be written as

$$\mathbf{g}^T \boldsymbol{\gamma} = 1 \tag{33}$$

Pre-multiplying Eq. (32) by $(Z + \lambda I_{4\times 4})^2$ gives

$$(Z + \lambda I_{4\times 4})^2 \boldsymbol{\gamma} = \mathbf{g} \tag{34}$$

Finally multiplying each side of Eq. (34) by unity using Eq. (33) gives

$$(Z + \lambda I_{4\times 4})^2 \boldsymbol{\gamma} = \mathbf{g}\mathbf{g}^T \boldsymbol{\gamma} \tag{35}$$

Equation (35) is the associated QEP for the Lagrange equations of Eq. (23). Reference [19] goes through several rigorous proofs to show that the maximum eigenvalue of the associated QEP is the unique solution for the Lagrange equations. As stated, the QEP can be transformed into a SEP with relative ease. Define the $[4 \times 1]$ vector $\boldsymbol{\eta}$ as

$$\boldsymbol{\eta} \triangleq (Z + \lambda I_{4\times 4}) \boldsymbol{\gamma} \tag{36}$$

Substituting $\boldsymbol{\eta}$ into Eq. (35) and rearranging slightly yields

$$Z\boldsymbol{\eta} - \mathbf{g}^T \mathbf{g}\, \boldsymbol{\gamma} = -\lambda \boldsymbol{\eta} \tag{37}$$

Rearranging Eq. (36) results in

$$Z\boldsymbol{\gamma} - \boldsymbol{\eta} = -\lambda \boldsymbol{\gamma} \tag{38}$$

Defining the vector $\boldsymbol{\xi} \triangleq [\boldsymbol{\gamma}^T \ \ \boldsymbol{\eta}^T]^T$ allows Eqs. (37) and (38) to be written as

$$-\begin{bmatrix} Z & -I_{4\times 4} \\ -\mathbf{g}^T \mathbf{g} & Z \end{bmatrix} \boldsymbol{\xi} = \lambda \boldsymbol{\xi} \quad \Rightarrow \quad \mathcal{A}\boldsymbol{\xi} = \lambda \boldsymbol{\xi} \tag{39}$$

Equation (39) is an SEP and (λ, ξ) are an associated right eigenpair of $\mathcal{A}$. Because we have an augmented $[8 \times 1]$ vector $\boldsymbol{\xi}$, there will be 8 eigenpairs. This is consistent with the results of the secular equation. Again the correct value for λ is the largest real eigenvalue. After determining the largest eigevalue, Eq. (29) can be used directly to find **q** and **b**. Note that determination of the fused covariance P_{cc} is done prior to determination of **q** and **b**, and has no effect other then the weight ω to the CI algorithm.

Square Root Formulation of CI

If all the $\mathbf{q}_i$ vectors are close to each other, then the matrix $(Z + \lambda I_{4\times 4})$ is close to being singular. To alleviate this problem a square root formulation is derived in this section based on the techniques in [29]. First consider that the error-state vector $\Delta\mathbf{x}_i$ can be written as

$$\Delta\mathbf{x} \triangleq \begin{bmatrix} \Xi^T(\mathbf{q}_i)\mathbf{q} \\ \mathbf{b} - \mathbf{b}_i \end{bmatrix} = \begin{bmatrix} \Xi(\mathbf{q}_i) & 0_{3\times n_b} \\ 0_{n_b\times 4} & I_{n_b\times n_b} \end{bmatrix}^T \begin{bmatrix} \mathbf{q} \\ \mathbf{b} - \mathbf{b}_i \end{bmatrix} \tag{40}$$

Using the definition for **x** from Eq. (12) and defining

$$\mathbf{z}_i \triangleq \begin{bmatrix} \mathbf{0}_{4\times 1} \\ \mathbf{b}_i \end{bmatrix} \tag{41a}$$

$$\Psi_i \triangleq \begin{bmatrix} \Xi(\mathbf{q}_i) & 0_{4\times n_b} \\ 0_{n_b\times 3} & I_{n_b\times n_b} \end{bmatrix} \tag{41b}$$

allows the objective function in Eq. (11) to be written as

$$J(\mathbf{x}) = -\sum_{i=1}^{n} \omega_i \left(\mathbf{x} - \mathbf{z}_i\right)^T \Psi_i P_i^{-1} \Psi_i^T \left(\mathbf{x} - \mathbf{z}_i\right) \tag{42}$$

Define the positive semi-definite matrix $\mathcal{W}_i$ as

$$\mathcal{W}_i = \omega_i \Psi_i P_i^{-1} \Psi_i^T \tag{43}$$

Because $\mathcal{W}_i$ is positive semi-definite we can compute its matrix square root as

$$\mathcal{W}_i = C_i^T C_i \tag{44}$$

The matrix square root is assisted noting that $\mathcal{W}_i$ is symmetric. Computing the eigenvalue decomposition of $\mathcal{W}_i$ gives

$$\mathcal{W}_i = Q_i \Sigma_i^2 Q_i^T \tag{45}$$

where Σ_i is a diagonal matrix of the singular values of $\mathcal{W}_i$. Comparing Eqs. (44) and (45) yields

$$C_i = \Sigma_i Q_i^T \tag{46}$$

Distributing C_i into Eq. (42), the objective function is

$$J(\mathbf{x}) = -\sum_{i=1}^{n} \left(C_i\mathbf{x} - C_i\mathbf{z}_i\right)^T \left(C_i\mathbf{x} - C_i\mathbf{z}_i\right) \tag{47}$$

The cost function

$$J(\mathbf{x}) = -\sum_{i=1}^{n} \left(C_i\mathbf{x} - C_i\mathbf{z}_i\right)^T \left(C_i\mathbf{x} - C_i\mathbf{z}_i\right) \tag{48}$$

can be written as

$$J(\mathbf{x}) = -\left(\mathcal{S}\mathbf{x} - \mathbf{z}\right)^T\left(\mathcal{S}\mathbf{x} - \mathbf{z}\right) - r^2 \tag{49}$$

for some $\mathcal{S}$, **z** and r. The proof of this relationship is now shown. The summation in Eq. (47) can be rewritten as

$$-\sum_{i=1}^{n}\left(C_i\mathbf{x} - C_i\mathbf{z}_i\right)^T\left(C_i\mathbf{x} - C_i\mathbf{z}_i\right) = -\left|\left|\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix}\mathbf{x} - \begin{bmatrix} C_1\mathbf{z}_1 \\ C_2\mathbf{z}_2 \\ \vdots \\ C_n\mathbf{z}_n \end{bmatrix}\right|\right|_2^2 \tag{50}$$

where $||\cdot||_2$ denotes the 2-norm. Note for any vector **y** we have $||\mathbf{y}||_2^2 = \mathbf{y}^T\mathbf{y}$. Now suppose we have an orthogonal matrix U such that $UU^T = U^TU = I$. The 2-norm of the vector **y** is unaffected by multiplication with U as

$$\left|\left|U^T\mathbf{y}\right|\right|_2^2 = \left(U^T\mathbf{y}\right)^T\left(U^T\mathbf{y}\right) = \mathbf{y}^TUU^T\mathbf{y} = \mathbf{y}^T\mathbf{y} = ||\mathbf{y}||_2^2 \tag{51}$$

Following the results of Eq. (51), pre-multiply the argument of Eq. (50) by some orthogonal matrix Q^T:

$$-\sum_{i=1}^{n}\left(C_i\mathbf{x} - C_i\mathbf{z}_i\right)^T\left(C_i\mathbf{x} - C_i\mathbf{z}_i\right) = -\left|\left|Q^T\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix}\mathbf{x} - Q^T\begin{bmatrix} C_1\mathbf{z}_1 \\ C_2\mathbf{z}_2 \\ \vdots \\ C_n\mathbf{z}_n \end{bmatrix}\right|\right|_2^2 \tag{52}$$

Now consider the $[n(n_b + 4) \times (n_b + 5)]$ matrix $\mathcal{G}$

$$\mathcal{G} \triangleq \begin{bmatrix} C_1 & C_1\mathbf{z}_1 \\ C_2 & C_2\mathbf{z}_2 \\ \vdots & \vdots \\ C_n & C_n\mathbf{z}_n \end{bmatrix} \tag{53}$$

A QR decomposition of $\mathcal{G}$ results in an $[n(n_b + 5) \times (n_b + 5)]$ orthogonal matrix Q and an upper triangular matrix R partitioned as

$$R = \begin{array}{c} \begin{array}{cc} {\scriptstyle (n_b+4)} & {\scriptstyle 1} \end{array} \\ \begin{bmatrix} \mathcal{S} & \mathbf{z} \\ 0 & r \end{bmatrix} \end{array} \begin{array}{c} {\scriptstyle (n_b+4)} \\ {\scriptstyle 1} \end{array} \tag{54}$$

where $\mathcal{S}$ is upper triangular. Multiplying $QR = \mathcal{G}$ by Q^T results in

$$R = Q^T\begin{bmatrix} C_1 & C_1\mathbf{z}_1 \\ C_2 & C_2\mathbf{z}_2 \\ \vdots & \vdots \\ C_n & C_n\mathbf{z}_n \end{bmatrix} \tag{55}$$

Because Q^T left-multiplies $\mathcal{G}$ we have

$$\begin{bmatrix} \mathcal{S} \\ 0 \end{bmatrix} = Q^T \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix} \tag{56a}$$

$$\begin{bmatrix} \mathbf{z} \\ r \end{bmatrix} = Q^T \begin{bmatrix} C_1\mathbf{z}_1 \\ C_2\mathbf{z}_2 \\ \vdots \\ C_n\mathbf{z}_n \end{bmatrix} \tag{56b}$$

Substituting Eq. (56) into Eq. (52) yields

$$J(\mathbf{x}) = -\left|\left| \begin{bmatrix} \mathcal{S} \\ 0 \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{z} \\ r \end{bmatrix} \right|\right|_2^2 \tag{57}$$

Equation (57) can also be written as

$$\begin{aligned} J(\mathbf{x}) = -\left|\left| \begin{bmatrix} \mathcal{S}\mathbf{x} - \mathbf{z} \\ r \end{bmatrix} \right|\right|_2^2 &= -\,||(\mathcal{S}\mathbf{x} - \mathbf{z})||_2^2 - r^2 \\ &= -\left(\mathcal{S}\mathbf{x} - \mathbf{z}\right)^T \left(\mathcal{S}\mathbf{x} - \mathbf{z}\right) - r^2 \end{aligned} \tag{58}$$

Once again the quaternion norm constraint is handled using the method of Lagrange multipliers. Here the constraint is defined as

$$\mathbf{x}^T \begin{bmatrix} I_{4\times 4} & 0_{4\times n_b} \\ 0_{n_b\times 4} & 0_{n_b\times n_b} \end{bmatrix} \mathbf{x} \triangleq \mathbf{x}^T I_q \mathbf{x} = 1 \tag{59}$$

which is equivalent to Eq. (14c). The appended objective function is

$$J(\mathbf{x}) = -\left(\mathcal{S}\mathbf{x} - \mathbf{z}\right)^T \left(\mathcal{S}\mathbf{x} - \mathbf{z}\right) - r^2 + \lambda(1 - \mathbf{x}^T I_q \mathbf{x}) \tag{60}$$

Taking the partial derivatives result in the following conditions to maximize the objective function:

$$\frac{\partial J(\mathbf{x})}{\partial \mathbf{x}} : \left(\mathcal{S}^T\mathcal{S} + \lambda I_q\right)\mathbf{x} = \mathcal{S}^T\mathbf{z} \tag{61a}$$

$$\frac{\partial J(\mathbf{x})}{\partial \lambda} : \mathbf{x}^T I_q \mathbf{x} = 1 \tag{61b}$$

The solution for the square root approach hinges on the knowledge that the two representations of the objective function, Eqs. (60) and (13) must be equivalent. It then follows that their respective necessary conditions, Eqs. (61) and (14) must also be equivalent. Therefore,

$$\mathcal{B} = \mathcal{S}^T\mathcal{S} \tag{62a}$$

$$\begin{bmatrix} \mathbf{c} \\ \mathbf{d} \end{bmatrix} = \mathcal{S}^T\mathbf{z} \tag{62b}$$

In what follows we will relate certain quantities of the square root approach with the standard approach shown previously. First, start off by defining

$$\mathcal{S}^{-1} \triangleq S = \begin{array}{c} \begin{array}{cc} 4 & n_b \end{array} \\ \begin{bmatrix} S_{qq} & S_{qb} \\ 0 & S_{bb} \end{bmatrix} \end{array} \begin{array}{c} 4 \\ n_b \end{array} \tag{63}$$

The matrix Z can be written in terms of the partitions of S as

$$Z^{-1} = S_{qq}S_{qq}^T + S_{qb}S_{qb}^T \tag{64}$$

To prove this relationship, recall from the previous sections the following:

$$\mathcal{B} = \begin{bmatrix} \mathcal{B}_{qq} & \mathcal{B}_{qb} \\ \mathcal{B}_{qb}^T & \mathcal{B}_{bb} \end{bmatrix} \tag{65a}$$

$$Z = \mathcal{B}_{qq} - \mathcal{B}_{qb}\mathcal{B}_{bb}^{-1}\mathcal{B}_{qb}^T \tag{65b}$$

Now consider blockwise inversion of the following generic matrix

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{bmatrix} \tag{66}$$

Taking the inverse of $\mathcal{B}$ and consulting Eq. (66) we find that Z^{-1} is the top-left block of $\mathcal{B}^{-1}$. That is

$$\mathcal{B}^{-1} = \begin{bmatrix} Z^{-1} & \cdot \\ \cdot & \cdot \end{bmatrix} \tag{67}$$

where the dot ($\cdot$) simply represents some quantity. From Eq. (62a) $\mathcal{B}^{-1}$ can also be expressed as

$$\mathcal{B}^{-1} = \left(\mathcal{S}^T\mathcal{S}\right)^{-1} = \mathcal{S}^{-1}\mathcal{S}^{-T} = SS^T \tag{68}$$

Carrying out the blockwise matrix multiplication from Eq. (68) using Eq. (63) results in

$$\mathcal{B}^{-1} = \begin{bmatrix} S_{qq}S_{qq}^T + S_{qb}S_{qb}^T & S_{qb}S_{bb}^T \\ S_{bb}S_{qb}^T & S_{bb}S_{bb}^T \end{bmatrix} \tag{69}$$

Then from Eq. (67) it follows that Eq. (64) is true.

The vector $\mathbf{g}$ can be written as

$$\mathbf{g} = Z\left[S_{qq} \quad S_{qb}\right]\mathbf{z} \tag{70}$$

This is proven by first recalling from our previous derivation, that $\mathbf{g}$ is independent of λ, shown from Eq. (22). With this knowledge we set $\lambda = 0$. Solving Eq. (23a) directly yields

$$\check{\mathbf{q}} = Z^{-1}\mathbf{g} \tag{71}$$

where $\check{\mathbf{q}}$ is the optimal quaternion estimate to the unconstrained problem. If we consider Eq. (61) in the unconstrained case then we have

$$\mathbf{x} = \mathcal{S}^{-1}\mathbf{z} \tag{72}$$

Substituting Eq. (63) yields

$$\begin{bmatrix} \check{\mathbf{q}} \\ \check{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} S_{qq} & S_{qb} \end{bmatrix} \mathbf{z} \\ S_{bb}\,\mathbf{z} \end{bmatrix} \tag{73}$$

Comparing top block of Eq. (73) with Eq. (71) we conclude

$$\mathbf{g} = Z \begin{bmatrix} S_{qq} & S_{qb} \end{bmatrix} \mathbf{z}$$

which completes the proof.

The matrix Z^{-1} can be expressed as the product of a triangular matrix and its transpose

$$Z^{-1} = R_z^T R_z \tag{74}$$

where R_z is a triangular matrix. Consider that Z^{-1} from Eq. (64) can be written as

$$Z^{-1} = M^T M \tag{75}$$

where

$$M \triangleq Q^T \begin{bmatrix} S_{qq}^T \\ \\ S_{qb}^T \end{bmatrix} \tag{76}$$

and Q is an orthogonal matrix. Pre-multiplying Eq. (76) by Q results in

$$QM = \begin{bmatrix} S_{qq}^T \\ \\ S_{qb}^T \end{bmatrix} \tag{77}$$

From Eq. (77) we see that Q and M are the results of a QR decomposition of $\begin{bmatrix} S_{qq} & S_{qb} \end{bmatrix}^T$. With $R_z \triangleq M$ the proof is complete.

For the secular-equation based approach, the eigenvalue decomposition of Z is needed. We look to exploit the fact that we can now represent Z as a square root factor. The eigenvalue decomposition of Z is $Z = QVQ^T$. Because Z is positive definite we can take the square root of the eigenvalue matrix V:

$$Z = Q\Sigma\Sigma Q^T \tag{78}$$

where Σ is the diagonal matrix of singular values, $\Sigma \triangleq \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$. If there exists some orthogonal matrix U, then Eq. (78) is equivalent to

$$Z = Q\Sigma U^T U \Sigma Q^T \tag{79}$$

Comparing Eqs. (74) and (79) we see that U, Σ and V are the results of a singular value decomposition of R_z^{-1}. Substituting Eq. (78) into Eq. (70) gives

$$\mathbf{g} = Q\Sigma^2 Q^T \begin{bmatrix} S_{qq} & S_{qb} \end{bmatrix} \mathbf{z} \tag{80}$$

From Eq. (25b)

$$\mathbf{a} = Q^T \mathbf{g} = \Sigma^2 Q^T \begin{bmatrix} S_{qq} & S_{qb} \end{bmatrix} \mathbf{z} \tag{81}$$

Given **a** and the values of $\delta_i \triangleq \sigma_i^2$, one can now solve the secular equation, Eq. (28) for λ. Once the the optimal value of λ is determined, the optimal **x** is computed using

$$\mathbf{x} = \left(\mathcal{S}^T\mathcal{S} + \lambda I_q\right)^{-1}\mathcal{S}^T\mathbf{z} \tag{82}$$

from Eq. (61a). The matrix inverse can be computed effectively as follows. Define

$$\mathcal{Z} \triangleq \mathcal{S}^T\mathcal{S} + \lambda I_q \tag{83}$$

When $\mathcal{Z}$ is symmetric, positive definite, it can be characterized by a Cholesky factorization

$$\mathcal{Z} = \mathcal{L}^T\mathcal{L} \tag{84}$$

where $\mathcal{L}$ is defined by four rank-one Cholesky updates [31] of $\mathcal{S}$ with the four update vectors being the columns of

$$\begin{bmatrix} \text{sign}(\lambda)\sqrt{|\lambda|}I_{4\times 4} \\ 0_{n_b\times 4} \end{bmatrix}$$

After the Cholesky updates then the optimal **x** is computed as

$$\mathbf{x} = \mathcal{L}^{-1}\mathcal{L}^{-T}\mathcal{S}\mathbf{z} \tag{85}$$

The square-root formulation presented in the preceding sections can be summarized as follows:

1. Calculate C_i and $\mathbf{z}_i$ from Eqs. (46) and (41a), respectively.
2. Form the matrix $\mathcal{G}$ using Eq. (53).
3. Compute $\mathcal{S}$ and **z** based on the QR decomposition of $\mathcal{G}$.
4. Compute and $S = \mathcal{S}^{-1}$ and partition as in Eq. (63).
5. Compute R_z using a QR decomposition of $\begin{bmatrix} S_{qq} & S_{qb} \end{bmatrix}^T$.
6. Compute Q, Σ and U from a singular value decomposition of R_z^{-1}.
7. Compute **a** from Eq. (81) and $\delta_i = \sigma_i^2$ from Σ.
8. Solve the secular equation, Eq. (28) for λ.
9. Compute $\mathcal{L}$ basked on four rank-one Cholesky updates of $\mathcal{S}$.
10. Compute the optimal **x** using Eq. (85).

The matrix inverse needs to be replaced by the Penrose-Moore pseudo-inverse when the matrix is singular.

Practical Issues

All the approaches are derived under the assumption that $(Z + \lambda_{\text{max}} I)$ is nonsingular and that the optimal quaternion satisfies $\mathbf{q} = (Z + \lambda_{\text{max}} I_{4\times4})^{-1}\mathbf{g}$. When $(Z + \lambda_{\text{max}} I)$ is singular, which occurs when λ_{max} equals the negative of the minimum eigenvalue of Z, the optimal quaternion may take a more complex form. Define $\bar{\mathbf{q}} = (Z + \lambda_{\text{max}} I_{4\times4})^{\dagger}\mathbf{g}$, where † denotes the Penrose-Moore pseudo-inverse. The following observations help determine the optimal quaternion [22]:

1. If $l = \|\bar{\mathbf{q}}\| = 1$, then $\mathbf{q} = \bar{\mathbf{q}}$.
2. If $l = \|\bar{\mathbf{q}}\| < 1$, then $(Z + \lambda_{\text{max}} I_{4\times4})$ is singular, and $\mathbf{q} = \bar{\mathbf{q}} + \mathbf{t}$, where $\mathbf{t}$ is in the null space of Z and of magnitude $\sqrt{1 - l^2}$. Note that in this case the solution is non-unique because of the ambiguity in the sign of $\mathbf{t}$.

In some instances it is also found that the norm of **q** does not equal unity. To alleviate this problem in implementation of any of the aforementioned routines, a simple normalization of the quaternion was performed. Note that this will not change the result of the fusion as the method of Lagrange multipliers was used to conserve the quaternion norm. In some instances deviation from unity can be caused by numerical issues and in other cases, the quaternion multiplied by some scale factor.

MULTIPLE ATTITUDE ESTIMATION OF A SPACECRAFT FORMATION

In decentralized attitude estimation of a formation of three or more spacecraft, estimates consisting of spacecraft attitude quaternions and other (unconstrained) quantities are combined by the CI approach. The challenge for the CI approach comes from the multiple attitude quaternion constraints. In this section, a case of relative attitude estimation of a formation of three spacecraft is considered. The more general case can be treated similarly. The state vector is composed of two relative attitudes $\mathbf{q}^{(1)}$ and $\mathbf{q}^{(2)}$, with $\mathbf{q}^{(1)T}\mathbf{q}^{(1)} = 1$ and $\mathbf{q}^{(2)T}\mathbf{q}^{(2)} = 1$, and unconstrained quantities $\mathbf{b}$. It is assumed that the estimates originate from n sources. Define the estimate, error, and associated covariance of the i^{th} source as

$$\mathbf{x}_i \triangleq \overset{1}{\begin{bmatrix} \mathbf{q}_i^{(1)} \\ \mathbf{q}_i^{(2)} \\ \mathbf{b}_i \end{bmatrix}} \begin{matrix} 4 \\ 4 \\ n_b \end{matrix}, \quad \Delta\mathbf{x}_i \triangleq \overset{1}{\begin{bmatrix} \Xi^T(\mathbf{q}_i^{(1)})\mathbf{q}^{(1)} \\ \Xi^T(\mathbf{q}_i^{(2)})\mathbf{q}^{(2)} \\ \mathbf{b} - \mathbf{b}_i \end{bmatrix}} \begin{matrix} 3 \\ 3 \\ n_b \end{matrix} \tag{86a}$$

$$P_i^{-1} \triangleq \begin{matrix} & 3 & 3 & n_b & \\ & \left[\begin{matrix} \mathcal{P}_{q^{(1)}q_i^{(1)}} \\ \mathcal{P}^T_{q^{(1)}q_i^{(2)}} \\ \mathcal{P}^T_{q^{(1)}b_i} \end{matrix}\right. & \begin{matrix} \mathcal{P}_{q^{(1)}q_i^{(2)}} \\ \mathcal{P}_{q^{(2)}q_i^{(2)}} \\ \mathcal{P}^T_{q^{(2)}b_i} \end{matrix} & \left.\begin{matrix} \mathcal{P}_{q^{(1)}b_i} \\ \mathcal{P}_{q^{(2)}b_i} \\ \mathcal{P}_{bb_i} \end{matrix}\right] & \begin{matrix} 3 \\ 3 \\ n_b \end{matrix} \end{matrix} \tag{86b}$$

The objective function to be maximized is of the same form as that used in the single quaternion case, Eq. (13). Augmenting the constraint function with the two quaternion constraints gives

$$J(\Delta x) = -\sum_{i=1}^{n} \omega_i \Delta\mathbf{x}_i^T P_i^{-1} \Delta\mathbf{x}_i + \lambda_1 \left[1 - \mathbf{q}^{(1)T}\mathbf{q}^{(1)}\right] + \lambda_2 \left[1 - \mathbf{q}^{(2)T}\mathbf{q}^{(2)}\right] \tag{87}$$

with λ_1 and λ_2 being Lagrange multipliers. The necessary conditions are found, after taking the required partial derivatives and can be arranged into the following form

$$\begin{bmatrix} \mathcal{B}_{q^{(1)}q^{(1)}} + \lambda_1 I_{4\times 4} & \mathcal{B}_{q^{(1)}q^{(2)}} & \mathcal{B}_{q^{(1)}b} \\ \mathcal{B}^T_{q^{(1)}q^{(2)}} & \mathcal{B}_{q^{(2)}q^{(2)}} + \lambda_2 I_{4\times 4} & \mathcal{B}_{q^{(2)}b} \\ \mathcal{B}^T_{q^{(1)}b} & \mathcal{B}^T_{q^{(2)}b} & \mathcal{B}_{bb} \end{bmatrix} \begin{bmatrix} \mathbf{q}^{(1)} \\ \mathbf{q}^{(2)} \\ \mathbf{b} \end{bmatrix} = \begin{bmatrix} \mathbf{c}^{(1)} \\ \mathbf{c}^{(2)} \\ \mathbf{d} \end{bmatrix} \tag{88a}$$

$$\mathbf{q}^{(1)T}\mathbf{q}^{(1)} = 1 \tag{88b}$$

$$\mathbf{q}^{(2)T}\mathbf{q}^{(2)} = 1 \tag{88c}$$

where

$$\mathcal{B}_{q^{(1)}q^{(1)}} \triangleq \sum_{i=1}^{n} \omega_i \Xi(\mathbf{q}_i^{(1)}) \mathcal{P}_{q^{(1)}q_i^{(1)}} \Xi^T(\mathbf{q}_i^{(1)}) \tag{89a}$$

$$\mathcal{B}_{q^{(1)}q^{(2)}} \triangleq \sum_{i=1}^{n} \omega_i \Xi(\mathbf{q}_i^{(1)}) \mathcal{P}_{q^{(1)}q_i^{(2)}} \Xi^T(\mathbf{q}_i^{(2)}) \tag{89b}$$

$$\mathcal{B}_{q^{(1)}b} \triangleq \sum_{i=1}^{n} \omega_i \Xi(\mathbf{q}_i^{(1)}) \mathcal{P}_{q^{(1)}b_i} \tag{89c}$$

$$\mathcal{B}_{q^{(2)}q^{(2)}} \triangleq \sum_{i=1}^{n} \omega_i \Xi(\mathbf{q}_i^{(2)}) \mathcal{P}_{q^{(2)}q_i^{(2)}} \Xi^T(\mathbf{q}_i^{(2)}) \tag{89d}$$

$$\mathcal{B}_{q^{(2)}b} \triangleq \sum_{i=1}^{n} \omega_i \Xi(\mathbf{q}_i^{(2)}) \mathcal{P}_{q^{(2)}b_i} \tag{89e}$$

$$\mathcal{B}_{bb} \triangleq \sum_{i=1}^{n} \omega_i \mathcal{P}_{bb_i} \tag{89f}$$

and

$$\mathbf{c}^{(1)} \triangleq \sum_{i=1}^{n} \omega_i \, \Xi(\mathbf{q}_i^{(1)}) \mathcal{P}_{q^{(1)}b_i} \mathbf{b}_i \tag{90a}$$

(90b)

$$\mathbf{c}^{(2)} \triangleq \sum_{i=1}^{n} \omega_i \, \Xi(\mathbf{q}_i^{(2)}) \mathcal{P}_{q^{(2)}b_i} \mathbf{b}_i \tag{90c}$$

(90d)

$$\mathbf{d} \triangleq \sum_{i=1}^{n} \omega_i \, \mathcal{P}_{bb_i} \mathbf{b}_i \tag{90e}$$

The goal is to solve the above equations for the optimal λ_1, λ_2, $\mathbf{q}^{(1)}$, $\mathbf{q}^{(2)}$, and $\mathbf{b}$. As in the one-quaternion case in the previous section, there are many critical points satisfying the necessary conditions. The optimal set maximizes the objective function. One strategy for finding the optimal solution is to use a general-purpose root-finder to solve for λ_1, λ_2, $\mathbf{q}^{(1)}$, $\mathbf{q}^{(2)}$, and $\mathbf{b}$ simultaneously. However, that does not make use of the structure of the problem and is computationally expensive. A more theoretically sound solution follows the procedures in the previous section.

1. Express the optimal $\mathbf{b}$ in terms of the optimal $\mathbf{q}^{(1)}$ and $\mathbf{q}^{(2)}$ from Eq. (88):

$$\mathbf{b} = \mathcal{B}_{bb}^{-1} \left(\mathbf{d} - \mathcal{B}_{q^{(1)}b}^T \mathbf{q}^{(1)} - \mathcal{B}_{q^{(2)}b}^T \mathbf{q}^{(2)} \right) \tag{91}$$

2. Express $\mathbf{q}^{(1)}$ and $\mathbf{q}^{(2)}$ in terms of known quantities:

$$\begin{bmatrix} \mathbf{q}^{(1)} \\ \mathbf{q}^{(2)} \end{bmatrix} = (\mathcal{D} + \Lambda)^{-1} \mathbf{h} \tag{92}$$

   where

$$\mathcal{D} \triangleq \begin{bmatrix} \mathcal{D}_{11} & \mathcal{D}_{12} \\ \mathcal{D}_{21} & \mathcal{D}_{22} \end{bmatrix} = \begin{bmatrix} \mathcal{B}_{q^{(1)}q^{(1)}} & \mathcal{B}_{q^{(1)}q^{(2)}} \\ \mathcal{B}_{q^{(1)}q^{(2)}}^T & \mathcal{B}_{q^{(2)}q^{(2)}} \end{bmatrix} - \begin{bmatrix} \mathcal{B}_{q^{(1)}b} \\ \mathcal{B}_{q^{(2)}b} \end{bmatrix} \mathcal{B}_{bb}^{-1} \begin{bmatrix} \mathcal{B}_{q^{(1)}b}^T & \mathcal{B}_{q^{(2)}b}^T \end{bmatrix} \tag{93}$$

$$\Lambda = \begin{bmatrix} \lambda_1 I_{4\times4} & 0_{4\times4} \\ 0_{4\times4} & \lambda_2 I_{4\times4} \end{bmatrix} \tag{94}$$

$$\mathbf{h} \triangleq \begin{bmatrix} \mathbf{h}^{(1)} \\ \mathbf{h}^{(2)} \end{bmatrix} = \begin{bmatrix} \mathbf{c}^{(1)} - \mathcal{B}_{q^{(1)}b} \mathcal{B}_{bb}^{-1} \mathbf{d} \\ \mathbf{c}^{(2)} - \mathcal{B}_{q^{(2)}b} \mathcal{B}_{bb}^{-1} \mathbf{d} \end{bmatrix} \tag{95}$$

3. Solve the following two equations for λ_1 and λ_2 using, for example, SOLVE of MATLAB or NSolve of Mathematica:

$$\mathbf{h}^T (\mathcal{D} + \Lambda)^{-1} \begin{bmatrix} I_{4\times4} & 0_{4\times4} \\ 0_{4\times4} & 0_{4\times4} \end{bmatrix} (\mathcal{D} + \Lambda)^{-1} \mathbf{h} = 1 \tag{96a}$$

$$\mathbf{h}^T (\mathcal{D} + \Lambda)^{-1} \begin{bmatrix} 0_{4\times4} & 0_{4\times4} \\ 0_{4\times4} & I_{4\times4} \end{bmatrix} (\mathcal{D} + \Lambda)^{-1} \mathbf{h} = 1 \tag{96b}$$

4. Compute the optimal $\mathbf{q}^{(1)}$ and $\mathbf{q}^{(2)}$ in terms of the optimal λ_1 and λ_2 using Eq. (92).

5. Compute $\mathbf{b}$ using Eq. (91).

The optimum set of zeros will now be defined. The optimal set of zeros $\{\lambda_1, \lambda_2\}$ are now shown to be the ones with the maximum sum. Because $\mathbf{q}^{(1)}$, $\mathbf{q}^{(2)}$, and $\mathbf{b}$ are functions of λ_1 and λ_2, the objective function is a function of λ_1 and λ_2. Substituting Eq. (91) into the objective function Eq. (87), and with the use of the definitions in Eqs. (89) and (93) yields

$$J^s(\lambda_1, \lambda_2) = -\begin{bmatrix} \mathbf{q}^{(1)T} & \mathbf{q}^{(2)T} \end{bmatrix} \mathcal{D} \begin{bmatrix} \mathbf{q}^{(1)} \\ \mathbf{q}^{(2)} \end{bmatrix} + 2 \begin{bmatrix} \mathbf{q}^{(1)T} & \mathbf{q}^{(2)T} \end{bmatrix} \mathbf{h} \tag{97}$$

where J^s is the state-dependent part of the objective function. Substituting

$$\mathbf{h} = (\mathcal{D} + \Lambda) \begin{bmatrix} \mathbf{q}^{(1)} \\ \mathbf{q}^{(2)} \end{bmatrix} \tag{98}$$

from Eq. (92) allows the objective function to be written as

$$J^s(\lambda_1,\lambda_2) = \begin{bmatrix}\mathbf{q}^{(1)T} & \mathbf{q}^{(2)T}\end{bmatrix}\mathcal{D}\begin{bmatrix}\mathbf{q}^{(1)}\\ \mathbf{q}^{(2)}\end{bmatrix} + 2(\lambda_1+\lambda_2) \tag{99}$$

Next consider two sets of zeros $\{\lambda_1,\lambda_2\}$ and $\{\lambda_1',\lambda_2'\}$ which correspond to quaternion estimates $[\mathbf{q}^{(1)}\ \ \mathbf{q}^{(2)}]$ and $[\mathbf{q}^{(1)'T}\ \ \mathbf{q}^{(2)'T}]$, respectively. From Eq. (98) it follows that

$$\mathbf{h} = (\mathcal{D}+\Lambda)\begin{bmatrix}\mathbf{q}^{(1)}\\ \mathbf{q}^{(2)}\end{bmatrix} = (\mathcal{D}+\Lambda')\begin{bmatrix}\mathbf{q}^{(1)'}\\ \mathbf{q}^{(2)'}\end{bmatrix} \tag{100}$$

Equation (100) can also be written as

$$\mathbf{0} = (\mathcal{D}+\Lambda)\begin{bmatrix}\mathbf{q}^{(1)}\\ \mathbf{q}^{(2)}\end{bmatrix} - (\mathcal{D}+\Lambda')\begin{bmatrix}\mathbf{q}^{(1)'}\\ \mathbf{q}^{(2)'}\end{bmatrix} \tag{101}$$

Pre-multiplying both sides by $[\mathbf{q}^{(1)'T}-\mathbf{q}^{(1)T}\ ,\ \mathbf{q}^{(2)'T}-\mathbf{q}^{(2)T}]$ and rearranging yields

$$\begin{bmatrix}\mathbf{q}^{(1)T} & \mathbf{q}^{(2)T}\end{bmatrix}\mathcal{D}\begin{bmatrix}\mathbf{q}^{(1)}\\ \mathbf{q}^{(2)}\end{bmatrix} - \begin{bmatrix}\mathbf{q}^{(1)'T} & \mathbf{q}^{(2)'T}\end{bmatrix}\mathcal{D}\begin{bmatrix}\mathbf{q}^{(1)'}\\ \mathbf{q}^{(2)'}\end{bmatrix} = (\lambda_1'+\lambda_2') - (\lambda_1+\lambda_2) \tag{102}$$

Equation (99) can then be written as

$$\begin{bmatrix}\mathbf{q}^{(1)T} & \mathbf{q}^{(2)T}\end{bmatrix}\mathcal{D}\begin{bmatrix}\mathbf{q}^{(1)}\\ \mathbf{q}^{(2)}\end{bmatrix} = J^s(\lambda_1,\lambda_2) - 2(\lambda_1+\lambda_2) \tag{103}$$

Using Eq. (103) for each respective set of zeros in Eq. (102) leads to

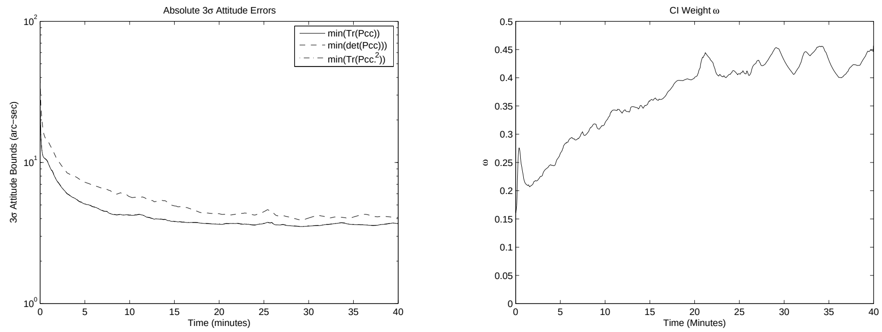
$$\begin{aligned} J^s(\lambda_1,\lambda_2) - J^s(\lambda_1',\lambda_2') &= [(\lambda_1'+\lambda_2') - (\lambda_1+\lambda_2)] + 2[(\lambda_1+\lambda_2) - (\lambda_1'+\lambda_2')] \\ &= (\lambda_1+\lambda_2) - (\lambda_1'+\lambda_2') \end{aligned} \tag{104}$$

From Eq. (104) it is clear that the optimal set of zeros (λ_1,λ_2) is the set with the maximum sum.

STAR TRACKER SIMULATION RESULTS

In this section results using two star trackers with gyros are shown. The spacecraft is assumed to be in low-Earth orbit with zero inclination. The trackers are pointed ± 45 degrees facing away from the Earth. Each tracker is assumed to have an 8 degree field-of-view and can observe stars down to magnitude 6 with a maximum of 10 stars at any time. The $+45$ degree (north) tracker observations are corrupted with zero-mean Gaussian white noise using a standard deviation of 3.5 arc-sec, while the -45 degree (south) tracker observations have noise with a standard deviation of 35 arc-sec. A sampling interval of 1 second is assumed for the star observations and gyro measurements. Each tracker is running its own extended Kalman filter using a common gyro. Details of the Kalman filter employed can be found in [3]. The estimated quantities are the spacecraft's attitude and three gyro biases, i.e. $\mathbf{x} = [\mathbf{q}^T\ \ \mathbf{b}^T]^T$.

The CI algorithm was used as previously derived. The weight ω was found using a simple 1-D bounded optimization routine to minimize the trace of P_{cc} as defined in Eq. (2a). It was found that minimizing the trace of P_{cc} provides superior results to minimizing the determinant of P_{cc}, see



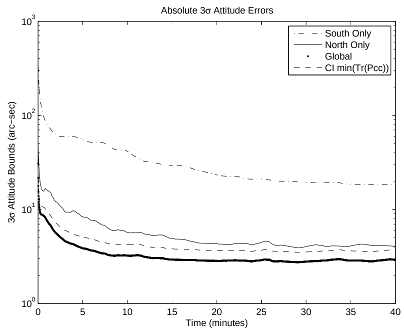
(a) Comparison of 3σ Bounds for Different Minimizations of P_{cc} (b) CI Weights when Minimizing Trace of P_{cc}

Figure 1 Comparisons of Minimization Routines and Optimal CI Weights

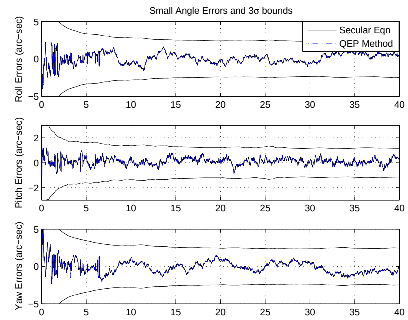
Figure 1(a). Minimizing the sum of the trace elements squared was also investigated but found to yield no improvement over minimizing $\mathrm{tr}(P_{cc})$. The weights associated with minimizing the tr(P_{cc}) can be seen in Figure 1(b). During the transient stage, the CI estimate relies more on the north tracker, which is more accurate than the south tracker. As the filter converges each filter's estimate is weighted nearly equally.

The estimated error results can be seen in Figure 2. Figure 2(a) shows the 3σ attitude bounds for the north only, south only and global filters as compared to the CI solution. The global filter represents a centralized extended Kalman filter which processes all available star and gyro measurements. Clearly, the CI bounds are lower than either tracker alone but greater then that of the centralized filter. These simulation results confirm that the CI approach is somewhat conservative in the computation of the fused covariance. The 3-axis attitude errors and respective 3σ bounds can be seen in Figure 2(b). The results from the secular equation algorithm (the square root implementation is identical to the standard implementation of the secular equation algorithm for this example) is juxtaposed with those from the QEP algorithm. The results are identical without increased magnification.

Simulations are also run assuming that both star trackers have the same noise standard deviation of 3.5 arc-sec. This was done in order to test the numerical properties of the proposed algorithms. Figure 3 shows the estimation results. As seen the CI weights the results from each filter almost equally. Also, note the rather large improvement in the 3σ attitude bounds obtained by fusing the two estimates as opposed to the more subtle improvement when the noise parameters were unequal. From the 3-axis attitude errors we see that the QEP algorithm has some instances where the errors are outside the 3σ bounds. The jumps are instantaneous and last for only one time step. These are due to a singular $(Z + \lambda I_{4\times 4})$ matrix as discussed previously. This occurs when the attitude estimates from the north and south local filters are almost identical. It should also be noted that in this case, the maximum λ contained an insignificant imaginary component. When this occurred only the real part of λ was used.

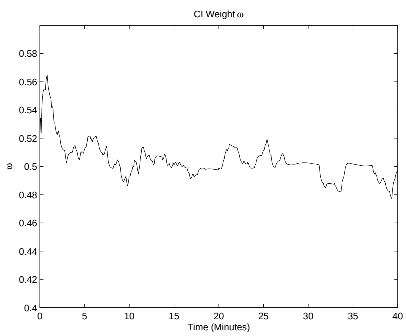


(a) 3σ Attitude Bounds

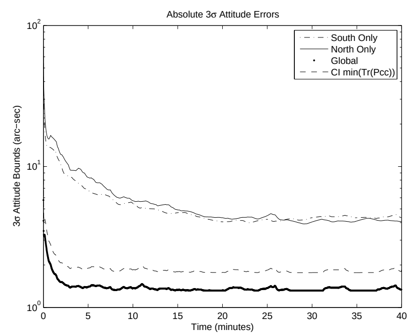


(b) CI Errors and 3σ Bounds

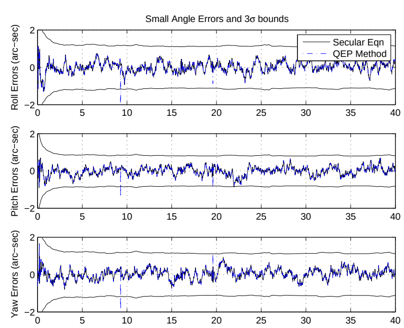
Figure 2 Estimation Errors and Bounds for CI



(a) CI Weights ω



(b) 3σ Attitude Bounds



(c) CI Errors and 3σ Bounds

Figure 3 Hard Case Simulation. Both Star Trackers have Noise with Standard Deviation of 3.5 arc-sec

CONCLUSIONS

A challenge for fusing estimates from multiple sources of a decentralized system arises from the correlation of the estimates. Tracking the correlation of the estimates may be impractical or impossible in certain applications. The CI method is a simple yet effective approach in fusing multiple estimates of unknown correlation. It guarantees that the updated estimate is consistent, although the estimate tends to be conservative. When applied to attitude estimation, the CI method involves solving a quadratic programming problem with one or more quadratic equality constraints that the attitude quaternions must have unity norm. The Lagrange multipliers are used to augment the objective function with the equality constraints. When there is only one quaternion constraint, the solution is obtained by solving the secular equation or a quadratic eigenvalue problem. A square root formulation of the secular-equation based approach is also derived. The star tracker simulation results illustrate the effectiveness of the CI method for single attitude quaternion estimation.

REFERENCES

[1] S. J. Julier and J. K. Uhlmann, "General Decentralized Data Fusion with Covariance Intersection (CI)," *Handbook of Multisensor Data Fusion* (D. L. Hall and J. Llinas, eds.), ch. 12, Boca Raton, FL: CRC Press, 2001.

[2] S. J. Julier and J. K. Uhlmann, "A Non-Divergent Estimation Algorithm in the Presence of Unknown Correlations," *Proceedings of the American Control Conference*, Vol. 4, Albuquerque, NM, June 1997, pp. 2369–2373.

[3] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems*. Boca Raton, FL: Chapman & Hall/CRC, 2004.

[4] L. Chen, P. O. Arambel, and R. K. Mehra, "Estimation Under Unknown Correlation: Covariance Intersection Revisited," *IEEE Transactions on Automatic Control*, Vol. 47, Nov. 2002, pp. 1879–1882.

[5] A. R. Benaskeur, "Consistent Fusion of Correlated Data Sources," *Proceedings of the 28th Annual Conference of the IEEE Industrial Electronics Society*, Vol. 4, Sevilla, Spain, Nov. 2002, pp. 2652–2656.

[6] D. Smith and S. Singh, "Approaches to Multisensor Data Fusion in Target Tracking: A Survey," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 18, Dec. 2006, pp. 1696–1710.

[7] G. B. Sholomitsky, O. F. Prilulsky, and V. G. Rodin, "Infra-Red Space Interferometer," *28th International Astronautical Congress of the International Astronautical Federation*, Prague, Czechoslovakia, Sept. 1977. IAF-77-68.

[8] D. P. Scharf, F. Y. Hadaegh, and S. R. Ploen, "A Survey of Spacecraft Formation Flying Guidance and Control (Part I): Guidance," *Proceedings of the American Control Conference*, Vol. 2, Denver, CO, June 2004, pp. 1733–1739.

[9] D. P. Scharf, F. Y. Hadaegh, and S. R. Ploen, "A Survey of Spacecraft Formation Flying Guidance and Control (Part II): Control," *Proceedings of the American Control Conference*, Vol. 4, Boston, MA, June-July 2004, pp. 2976–2985.

[10] J. R. Carpenter, "Decentralized Control of Satellite Formations," *International Journal of Robust and Nonlinear Control*, Vol. 12, No. 2-3, 2002, pp. 141–161.

[11] P. Ferguson and J. How, "Decentralized Estimation Algorithms for Formation Flying Spacecraft," *AIAA Guidance, Navigation, and Control Conference*, Austin, TX, Aug. 2003. AIAA-2003-5442.

[12] P. O. Arambel, C. Rago, and R. K. Mehra, "Covariance Intersection Algorithm for Distributed Spacecraft State Estimation," *Proceedings of the American Control Conference*, Vol. 6, Arlington, VA, June 2001, pp. 4398–4403.

[13] S.-G. Kim, J. L. Crassidis, Y. Cheng, A. M. Fosbury, and J. L. Junkins, "Kalman Filtering for Relative Spacecraft Attitude and Position Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 30, Jan.-Feb. 2007, pp. 133–143.

[14] M. D. Shuster, "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, Oct.-Dec. 1993, pp. 439–517.

[15] F. L. Markley, Y. Cheng, J. L. Crassidis, and Y. Oshman, "Averaging Quaternions," *Journal of Guidance, Control, and Dynamics*, Vol. 30, July-Aug. 2007, pp. 1193–1196.

[16] J. K. Uhlmann, "General Data Fusion for Estimates with Unknown Cross Covariances," *SPIE AeroSense Conference*, Vol. 2755, 1996, pp. 165–173.

[17] M. Grant and S. Boyd, "CVX: MATLAB software for disciplined convex programming," (web page and software). `http://stanford.edu/~boyd/cvx`.

[18] E. J. Lefferts, F. L. Markley, and M. D. Shuster, "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 5, Sept.-Oct. 1982, pp. 417–429.

[19] W. Gander, G. H. Golub, and U. von Matt, "A Constrained Eigenvalue Problem," *Linear Algebra and its Applications*, 1989, pp. 815 – 839.

[20] D. C. Sorenson, "Newton's Method with a Model Trust Region Modification," *SIAM Journal on Numerical Analysis*, Vol. 19, No. 2, 1982, pp. 409–426.

[21] M. L. Psiaki, "Attitude-Determination Filtering via Extended Quaternion Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 23, March-April 2000, pp. 206–214.

[22] Z. Zhang and Y. Huang, “A Projection Method for Least Squares Problems with A Quadratic Equality Constraint,” *SIAM Journal of Matrix Analysis and Applications*, Vol. 25, No. 1, 2003, pp. 188–212.

[23] D. M. Sima, S. V. Huffel, and G. H. Golub, “Regularized Total Least Squares Based on Quadratic Eivenvalue Problem Solvers,” *BIT Numerical Mathematics*, Vol. 44, December 2004, pp. 793–812.

[24] G. H. Golub and U. von Matt, “Quadratically Constrained Least Squares ad Quadratic Problems,” *Numerische Mathematik*, Vol. 59, 1991, pp. 561 – 180.

[25] M. R. Abdel-Aziz and M. M. El-Amen, “Solving Large-Scale Constrained Least Squares Problems,” *Applied Mathematics and Computation*, Vol. 137, No. 2-3, 2003, pp. 571–587.

[26] W. Gander, “Least Squares with A Quadratic Constraint,” *Numerische Mathematik*, Vol. 36, 1981, pp. 291–307.

[27] G. H. Golub, “Some Modified Matrix Eigenvalue Problems,” *SIAM Review*, Vol. 15, No. 2, 1973, pp. 318 – 334.

[28] F. Tisseur and K. Meerbergen, “The Quadratic Eigenvalue Problem,” *SIAM Review*, Vol. 43, No. 2, 2001, pp. 235–286.

[29] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*, ch. 4. Orlando, FL: Academic International Press, 1977.

[30] D. C. Lay, *Linear Algebra and Its Applications*. Boston, MA: Pearson/Addison-Wesley, 2003.

[31] The Mathworks, *MATLAB Help: Cholesky Update (cholupdate)*.