A Total Least-Squares Estimate for Attitude Determination

Yang Cheng*
Mississippi State University, Mississippi State, MS 39762

John L. Crassidis†
University at Buffalo, State University of New York, Amherst, New York, 14260-4400

This paper presents the general total least-squares formulation for the problem of attitude determination from vector observations. Except for a special case, the total least-squares attitude estimation problem cannot be converted to Wahba's problem. Two iterative solutions are presented that determine both the attitude matrix and the unit or non-unit vector observations. The associated covariance matrices are also derived under the small error assumption.

I. Introduction

Most methods for attitude determination from vector observations are exact or approximate solutions of Wahba's problem, which is to find the least-squares estimate of the attitude matrix A that minimizes the loss function [1]

$$L(A) = \frac{1}{2}\sum_{i=1}^{n} w_i \|\tilde{\mathbf{b}}_i - A\tilde{\mathbf{r}}_i\|^2, \quad n \geq 2 \tag{1}$$

subject to the constraint that A is a three-dimensional proper orthogonal matrix

$$AA^T = A^T A = I_{3\times 3}, \quad \det A = 1 \tag{2}$$

where $\tilde{\mathbf{b}}_i$ are n vectors observed in the body frame, $\tilde{\mathbf{r}}_i$ are the corresponding vectors in the reference frame, and w_i are non-negative scalar weights accounting for the total measurement errors of the $\tilde{\mathbf{b}}_i$ and $\tilde{\mathbf{r}}_i$ pairs. The vectors and the attitude matrix satisfy

$$\tilde{\mathbf{b}}_i \approx A\tilde{\mathbf{r}}_i \tag{3}$$

The approximation is due to measurement errors. In satellite attitude determination applications, $\tilde{\mathbf{b}}_i$ and $\tilde{\mathbf{r}}_i$ are usually unit vectors representing line-of-sight directions. In other applications, for example, ground vehicle attitude determination by measuring the local magnetic field vector and local gravity vector, the vectors are not necessarily unit vectors. The loss function has an equivalent form

$$L(A) = \frac{1}{2}\sum_{i=1}^{n} w_i \left(\tilde{\mathbf{b}}_i^T\tilde{\mathbf{b}}_i + \tilde{\mathbf{r}}_i^T\tilde{\mathbf{r}}_i\right) - \sum_{i=1}^{n} w_i \tilde{\mathbf{b}}_i^T A\tilde{\mathbf{r}}_i \tag{4}$$

Because $L(A)$ appears linearly in A, Wahba's problem can be solved in closed form using an eigenvalue decomposition or a singular value decomposition [2, 3]. Many other solutions can be found in [3] and the references therein.

Another equivalent form of the loss function is

$$L(A) = \|A - B\|_F^2 \tag{5}$$

with $B = \sum_{i=1}^n w_i \tilde{\mathbf{b}}_i \tilde{\mathbf{r}}_i^T$ and $\|\cdot\|_F$ denotes the Frobenius norm. This means that Wahba's problem is equivalent to the orthogonal Procrustes problem [4].

Wahba's problem can also be related to a pose (position and attitude) determination problem [5], where the attitude matrix A and position vector $\mathbf{p}$ are simultaneously determined from the vector measurement pairs $\tilde{\mathbf{b}}_i$ and $\tilde{\mathbf{r}}_i$, $i = 1, \ldots, n$. The geometry may be assumed to be

$$\tilde{\mathbf{b}}_i \approx A(\tilde{\mathbf{r}}_i - \mathbf{p}) \tag{6}$$

*Associate Professor, Department of Aerospace Engineering. Email: cheng@ae.msstate.edu. Associate Fellow AIAA.
†Samuel P. Capen Chair Professor, Department of Mechanical & Aerospace Engineering. Email: johnc@buffalo.edu. Fellow AIAA.

The approximation is caused by measurement errors. In a LIDAR-based vehicle navigation problem, for example, $\tilde{\mathbf{r}}_i$ may represent the locations of the feature points in the reference frame and $\tilde{\mathbf{b}}_i$ the positions of the feature points relative to the vehicle in the body frame. The loss function for the least-squares pose estimation problem is

$$L(A, \mathbf{p}) = \frac{1}{2}\sum_{i=1}^{n} w_i \|\tilde{\mathbf{b}}_i - A(\tilde{\mathbf{r}}_i - \mathbf{p})\|^2 \tag{7}$$

In terms of the optimal attitude matrix $\hat{A}$, the optimal position estimate $\hat{\mathbf{p}}$ is

$$\hat{\mathbf{p}} = \hat{A}^T \bar{\mathbf{b}} - \bar{\mathbf{r}} \tag{8}$$

with

$$\bar{\mathbf{b}} = \frac{\sum_{i=1}^{n} w_i \tilde{\mathbf{b}}_i}{\sum_{i=1}^{n} w_i}, \quad \bar{\mathbf{r}} = \frac{\sum_{i=1}^{n} w_i \tilde{\mathbf{r}}_i}{\sum_{i=1}^{n} w_i} \tag{9}$$

Substituting Eq. (8) into Eq. (7) gives the loss function of the attitude matrix

$$L(A) = \frac{1}{2}\sum_{i=1}^{n} w_i \left\| \left(\tilde{\mathbf{b}}_i - \bar{\mathbf{b}}\right) - A\left(\tilde{\mathbf{r}}_i - \bar{\mathbf{r}}\right)\right\|^2 \tag{10}$$

The pose estimation problem now reduces to a Wahba's problem with non-unit vectors $(\tilde{\mathbf{b}}_i - \bar{\mathbf{b}})$ and $(\tilde{\mathbf{r}}_i - \bar{\mathbf{r}})$.

Although proposed as a least-squares attitude estimation problem, Wahba's problem can be interpreted as a special total least-squares problem. The total least-squares estimate is usually more accurate and less biased than the least-squares estimate [6]. The loss function of the total least-squares problem is [7]

$$L(A, \mathbf{r}_1, \ldots, \mathbf{r}_n) = \frac{1}{2}\sum_{i=1}^{n} w_{b_i} \|\tilde{\mathbf{b}}_i - A\mathbf{r}_i\|^2 + \frac{1}{2}\sum_{i=1}^{n} w_{r_i} \|\tilde{\mathbf{r}}_i - \mathbf{r}_i\|^2 \tag{11}$$

where two sets of scalars w_{b_i} and w_{r_i} are used to weight the measurement errors in the body frame and the reference frame, respectively. The objective of this total least-squares problem is to find the optimal estimates $\hat{A}$ and $\hat{\mathbf{r}}_i$ that minimize $L(A, \mathbf{r}_1, \ldots, \mathbf{r}_n)$. The optimal vector estimates $\hat{\mathbf{r}}_i$ and $\hat{\mathbf{b}}_i = \hat{A}\hat{\mathbf{r}}_i$ are not required to be unit vectors. The attitude part of the problem is a Wahba's problem because the optimal attitude estimate $\hat{A}$ minimizes [7]

$$\begin{aligned} L(A) &= \frac{1}{2}\sum_{i=1}^{n} \frac{1}{w_{b_i}^{-1} + w_{r_i}^{-1}} \|\tilde{\mathbf{b}}_i - A\tilde{\mathbf{r}}_i\|^2 \\ &= \frac{1}{2}\sum_{i=1}^{n} \frac{w_{b_i} w_{r_i}}{w_{b_i} + w_{r_i}} \|\tilde{\mathbf{b}}_i - A\tilde{\mathbf{r}}_i\|^2 \end{aligned} \tag{12}$$

where the effective scalar weight for each pair of vector measurements is $1/(w_{b_i}^{-1} + w_{r_i}^{-1})$. It is also known that the total least-squares estimation problem subject to a unitary constraint is identical to the orthogonal Procrustes problem [8].

This paper formulates, analyzes, and solves two general total least-squares problems for attitude determination without the restriction that the weights in the loss function are scalars. For coarse line-of-sight vectors and non-unit vectors such as $(\tilde{\mathbf{b}}_i - \bar{\mathbf{b}})$ and $(\tilde{\mathbf{r}}_i - \bar{\mathbf{r}})$ in pose estimation, fully populated weighting matrices can better reflect the anisotropic measurement error properties than scalars. The two formulations differ only in whether $\hat{\mathbf{r}}_i$ are unit vectors. The unity norm constraints are motivated by the requirement that the estimates of unit vectors be unit vectors. An important result of the paper is that the total least-squares problems do not reduce to Wahba's problem except for the special case of [7], where the weights are scalars and the vector estimates are not subject to any norm constraint.

The remainder of the paper presents the two problem formulations, the covariance analysis, the solutions of the two problems, and a simple numerical example with coarse line-of-sight vector measurements.

II. Problem Statement

A. Total Least-Squares Estimation

Suppose that the relationship between the parameter vector $\mathbf{x}$ and the output $\mathbf{y}$ is described by a linear model

$$\mathbf{y} = H\mathbf{x} \tag{13}$$

where H is the coefficient matrix. Given H and the measurement $\tilde{\mathbf{y}}$ of $\mathbf{y}$, the weighted least-squares estimate of $\mathbf{x}$ is the minimizer of the loss function

$$L(\mathbf{x}) = \frac{1}{2}(\tilde{\mathbf{y}} - H\mathbf{x})^T W_y (\tilde{\mathbf{y}} - H\mathbf{x}) \tag{14}$$

where W_y is a weighting matrix accounting for the measurement errors in $\tilde{\mathbf{y}}$.

The total least-squares estimation problem arises when errors are present in both $\mathbf{y}$ and H [6]. Given the measurements $\tilde{H}$ as well as $\tilde{\mathbf{y}}$, the total least-squares estimate of $\mathbf{x}$ and H (or of $\mathbf{y}$ and H) minimizes the loss function

$$L(\mathbf{x}, H) = \frac{1}{2}(\tilde{\mathbf{y}} - H\mathbf{x})^T W_y (\tilde{\mathbf{y}} - H\mathbf{x}) + \frac{1}{2}\mathrm{vec}^T(\tilde{H} - H)\, W_H \,\mathrm{vec}(\tilde{H} - H) \tag{15}$$

where $\mathrm{vec}(\cdot)$ denotes a column vector formed by stacking the consecutive columns of the input matrix and W_H is the weighting matrix for the coefficient matrix. The loss function can be extended to account for the cross correlation between the measurement errors in $\tilde{\mathbf{y}}$ and $\tilde{H}$ [9]. The estimate of $\mathbf{y}$ satisfies $\hat{\mathbf{y}} = \hat{H}\hat{\mathbf{x}}$. The total least-squares estimate $\hat{\mathbf{x}}$ minimizes a loss function of the form

$$L(\mathbf{x}) = \frac{1}{2}(\tilde{\mathbf{y}} - H\mathbf{x})^T W_y' (\tilde{\mathbf{y}} - H\mathbf{x}) \tag{16}$$

However, W_y' is a nonlinear function in $\mathbf{x}$ and not identical to W_y.

When the measurement errors obey the Gaussian distribution and the weighting matrices W_y and W_H are the inverse of the covariance matrices, the solutions to the least-squares and total least-squares problems are the maximum likelihood estimates.

B. Linear Attitude Measurement Model

The attitude measurement model appears linearly in the attitude matrix:

$$\mathbf{b}_i = A\mathbf{r}_i \tag{17}$$

It can be written as $\mathbf{y} = H\mathbf{x}$, with

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_n \end{bmatrix}, \quad H = \begin{bmatrix} H_1 \\ \vdots \\ H_n \end{bmatrix} \tag{18a}$$

$$\mathbf{y}_i = \mathbf{b}_i = \begin{bmatrix} \mathbf{r}_i^T & 0_{1\times 3} & 0_{1\times 3} \\ 0_{1\times 3} & \mathbf{r}_i^T & 0_{1\times 3} \\ 0_{1\times 3} & 0_{1\times 3} & \mathbf{r}_i^T \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix} = H_i\mathbf{x} \tag{18b}$$

and

$$\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \end{bmatrix}, \quad A = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{bmatrix} \tag{19}$$

In error analyses, another model of the form $\mathbf{y} = H\mathbf{x}$ is used. The small attitude error between A and a reference attitude $\bar{A}$ close to A approximately satisfies a linear model

$$\mathbf{y}_i = \mathbf{b}_i - \bar{A}\mathbf{r}_i = [\bar{A}\mathbf{r}_i\times]\,\boldsymbol{\delta\alpha} = H_i\mathbf{x} \tag{20}$$

where

$$\mathbf{x} = \boldsymbol{\delta\alpha}, \quad A = (I_{3\times 3} - [\boldsymbol{\delta\alpha}\times])\bar{A} \tag{21}$$

Here $\boldsymbol{\delta\alpha}$ denotes a vector of small angles. Because many elements of H are known to be zero, the vector $\mathrm{vec}(\tilde{H} - H)$ in the loss function of the total least-squares attitude estimation problem is replaced by $[\mathbf{r}_1^T, \ldots, \mathbf{r}_n^T]^T$.

C. Total Least-Squares Formulations for Attitude Determination

For sake of simplicity, the measurement errors in $\tilde{\mathbf{b}}_i$ and $\tilde{\mathbf{r}}_i$ are assumed to be uncorrelated. The goal is to find $\hat{A}$ and $\hat{\mathbf{r}}_i$. The vectors in the body frame are not estimated but computed using $\hat{\mathbf{b}}_i = \hat{A}\hat{\mathbf{r}}_i$. Since $\hat{A}$ is an attitude matrix, which leaves the length of a vector unchanged, then $\|\hat{\mathbf{b}}_i\| = \|\hat{\mathbf{r}}_i\|$. Two problems are formulated, different only in whether $\hat{\mathbf{r}}_i$ obeys the unity norm constraint.

Problem 1

Problem 1 is to minimize the loss function

$$L(A, \mathbf{r}_1, \ldots, \mathbf{r}_n) = \frac{1}{2}\sum_{i=1}^{n}(\tilde{\mathbf{b}}_i - A\mathbf{r}_i)^T W_{b_i}(\tilde{\mathbf{b}}_i - A\mathbf{r}_i) + \frac{1}{2}\sum_{i=1}^{n}(\tilde{\mathbf{r}}_i - \mathbf{r}_i)^T W_{r_i}(\tilde{\mathbf{r}}_i - \mathbf{r}_i) \tag{22}$$

subject to

$$AA^T = A^T A = I_{3\times3}, \quad \det A = 1 \tag{23}$$

Note that $\hat{\mathbf{r}}_i$ and $\hat{\mathbf{b}}_i$ are free of constraints in Problem 1. The weighting matrices W_{b_i} and W_{r_i} can be singular.

Problem 2

Problem 2 requires that $\mathbf{b}_i$, $\mathbf{r}_i$, $\tilde{\mathbf{b}}_i$, $\tilde{\mathbf{r}}_i$, $\hat{\mathbf{b}}_i$, $\hat{\mathbf{r}}_i$ are all unit vectors. It minimizes the same loss function as Problem 1

$$L(A, \mathbf{r}_1, \ldots, \mathbf{r}_n) = \frac{1}{2}\sum_{i=1}^{n}(\tilde{\mathbf{b}}_i - A\mathbf{r}_i)^T W_{b_i}(\tilde{\mathbf{b}}_i - A\mathbf{r}_i) + \frac{1}{2}\sum_{i=1}^{n}(\tilde{\mathbf{r}}_i - \mathbf{r}_i)^T W_{r_i}(\tilde{\mathbf{r}}_i - \mathbf{r}_i) \tag{24}$$

but has more constraints

$$AA^T = A^T A = I_{3\times3}, \quad \det A = 1 \tag{25a}$$

$$\mathbf{r}_i^T \mathbf{r}_i = 1 \tag{25b}$$

A third problem can be formulated where only some of the vectors are unit vectors. It can be solved the same way Problem 2 is solved.

III. Covariance Analysis

The covariance matrices of the estimates are derived for both problem formulations using a first-order perturbation approach. Under the small error assumption, the quadratic equality constraints in the attitude matrix and the unit vectors are eliminated or replaced with linear equality constraints.

The true and estimated quantities satisfy

$$\mathbf{b}_i = A\mathbf{r}_i \tag{26a}$$

$$\hat{\mathbf{b}}_i = \hat{A}\hat{\mathbf{r}}_i \tag{26b}$$

The measurement models are

$$\tilde{\mathbf{b}}_i = \mathbf{b}_i + \mathbf{\Delta b}_i \tag{27a}$$

$$\tilde{\mathbf{r}}_i = \mathbf{r}_i + \mathbf{\Delta r}_i \tag{27b}$$

where $\mathbf{\Delta b}_i$ and $\mathbf{\Delta r}_i$ are zero-mean, additive measurement errors with the associated covariances given by

$$E\{\mathbf{\Delta b}_i \mathbf{\Delta b}_i^T\} = R_{b_i} \tag{28a}$$

$$E\{\mathbf{\Delta r}_i \mathbf{\Delta r}_i^T\} = R_{r_i} \tag{28b}$$

It is further assumed that $\mathbf{\Delta b}_i$ and $\mathbf{\Delta r}_i$ are uncorrelated. The estimation errors are denoted by $\boldsymbol{\delta\alpha}$, $\boldsymbol{\delta}\mathbf{r}_i$, and $\boldsymbol{\delta}\mathbf{b}_i$:

$$\hat{A} \approx (I_{3\times3} - [\boldsymbol{\delta\alpha}\times])A \tag{29a}$$

$$\hat{\mathbf{r}}_i = \mathbf{r}_i + \boldsymbol{\delta}\mathbf{r}_i \tag{29b}$$

$$\hat{\mathbf{b}}_i = \mathbf{b}_i + \boldsymbol{\delta}\mathbf{b}_i \tag{29c}$$

Equation (29a) is valid under the small attitude error assumption. The measurement error $\mathbf{\Delta b}_i$ is linear in $\boldsymbol{\delta\alpha}$ and $\boldsymbol{\delta}\mathbf{r}_i$ to first order:

$$\mathbf{\Delta b}_i \approx [\mathbf{b}_i\times]\boldsymbol{\delta\alpha} + A\boldsymbol{\delta}\mathbf{r}_i \tag{30}$$

The unity norm constraint is approximated by

$$\mathbf{r}_i^T \boldsymbol{\delta}\mathbf{r}_i \approx 0 \tag{31}$$

A. Problem 1

With the attitude matrix replaced by the attitude error, Problem 1 becomes an unconstrained minimization problem. Approximately, the estimation errors $\boldsymbol{\delta\alpha}$ and $\boldsymbol{\delta}\mathbf{r}_i$ minimize the following loss function

$$L(\boldsymbol{\delta\alpha},\boldsymbol{\delta}\mathbf{r}_1,\ldots,\boldsymbol{\delta}\mathbf{r}_n) = \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\Delta}\mathbf{b}_i-[\mathbf{b}_i\times]\boldsymbol{\delta\alpha}-A\boldsymbol{\delta}\mathbf{r}_i)^T W_{b_i}(\boldsymbol{\Delta}\mathbf{b}_i-[\mathbf{b}_i\times]\boldsymbol{\delta\alpha}-A\boldsymbol{\delta}\mathbf{r}_i)+\frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\Delta}\mathbf{r}_i-\boldsymbol{\delta}\mathbf{r}_i)^T W_{r_i}(\boldsymbol{\Delta}\mathbf{r}_i-\boldsymbol{\delta}\mathbf{r}_i) \tag{32}$$

The optimality conditions can be written as

$$M\boldsymbol{\Delta}\mathbf{x} = N\boldsymbol{\Delta}\mathbf{y} \tag{33}$$

where

$$\boldsymbol{\Delta}\mathbf{x} = \begin{bmatrix}\boldsymbol{\delta\alpha}\\ \boldsymbol{\delta}\mathbf{r}_1\\ \vdots\\ \boldsymbol{\delta}\mathbf{r}_n\end{bmatrix}, \quad \boldsymbol{\Delta}\mathbf{y} \triangleq \boldsymbol{\Delta}\mathbf{y}(\boldsymbol{\Delta}\mathbf{b}_1,\boldsymbol{\Delta}\mathbf{r}_1,\ldots,\boldsymbol{\Delta}\mathbf{b}_n,\boldsymbol{\Delta}\mathbf{r}_n) = \begin{bmatrix}\boldsymbol{\Delta}\mathbf{b}_1\\ \boldsymbol{\Delta}\mathbf{r}_1\\ \vdots\\ \boldsymbol{\Delta}\mathbf{b}_n\\ \boldsymbol{\Delta}\mathbf{r}_n\end{bmatrix} \tag{34a}$$

$$M \triangleq M(A,\mathbf{b}_1,\ldots,\mathbf{b}_n) = \left[\begin{array}{c|ccc} -\sum_{i=1}^{n}[\mathbf{b}_i\times]W_{b_i}[\mathbf{b}_i\times] & -[\mathbf{b}_1\times]W_{b_1}A & \cdots & -[\mathbf{b}_n\times]W_{b_n}A \\ \hline A^T W_{b_1}[\mathbf{b}_1\times] & A^T W_{b_1}A + W_{r_1} & 0_{3\times3(n-2)} & 0_{3\times3} \\ \vdots & 0_{3(n-2)\times3} & \ddots & 0_{3(n-2)\times3} \\ A^T W_{b_n}[\mathbf{b}_n\times] & 0_{3\times3} & 0_{3\times3(n-2)} & A^T W_{b_n}A + W_{r_n} \end{array}\right] \tag{34b}$$

(34c)

$$N \triangleq N(A,\mathbf{b}_1,\ldots,\mathbf{b}_n) = \left[\begin{array}{cc|c|cc} -[\mathbf{b}_1\times]W_{b_1} & 0_{3\times3} & \cdots & -[\mathbf{b}_n\times]W_{b_n} & 0_{3\times3} \\ \hline A^T W_{b_1} & W_{r_1} & 0_{3\times6(n-2)} & \multicolumn{2}{c}{0_{3\times6}} \\ \multicolumn{2}{c|}{0_{3(n-2)\times6}} & \ddots & \multicolumn{2}{c}{0_{3(n-2)\times6}} \\ \multicolumn{2}{c|}{0_{3\times6}} & 0_{3\times6(n-2)} & A^T W_{b_n} & W_{r_n} \end{array}\right] \tag{34d}$$

Solving Eq. (33) gives

$$\boldsymbol{\Delta}\mathbf{x} = M^{-1}N\boldsymbol{\Delta}\mathbf{y} \tag{35}$$

Hence, the covariance matrix is given by

$$P_{\boldsymbol{\Delta}\mathbf{x}} = E\{\boldsymbol{\Delta}\mathbf{x}\,\boldsymbol{\Delta}\mathbf{x}^T\} = M^{-1}NRN^T M^{-1} \tag{36}$$

where

$$R = \text{blkdiag}(R_{b_1},R_{r_1},\ldots,R_{b_n},R_{r_n}) \tag{37}$$

The covariance expression is valid for any weights and any measurement error covariances.

Special Cases

If the weighting matrices are chosen as the inverse of the measurement error covariance matrices

$$W_{b_i} = R_{b_i}^{-1}, \quad W_{r_i} = R_{r_i}^{-1} \tag{38}$$

then it can be shown that the covariance takes a simple form

$$P_{\boldsymbol{\Delta}\mathbf{x}} = M^{-1} \tag{39}$$

The attitude error covariance is given by

$$
\begin{aligned}
P_{\delta\alpha} &= \left(\sum_{i=1}^{n} -[\mathbf{b}_i\times]W_{b_i}[\mathbf{b}_i\times] + [\mathbf{b}_i\times]W_{b_i}A(A^T W_{b_i}A + W_{r_i})^{-1}A^T W_{b_i}[\mathbf{b}_i\times]\right)^{-1} \\
&= -\sum_{i=1}^{n}[\mathbf{b}_i\times]\left(W_{b_i}^{-1} + AW_{r_i}^{-1}A^T\right)^{-1}[\mathbf{b}_i\times] \\
&= -\sum_{i=1}^{n}[\mathbf{b}_i\times]\left(R_{b_i} + AR_{r_i}A^T\right)^{-1}[\mathbf{b}_i\times]
\end{aligned}
\tag{40}
$$

If the measurement error covariances are a scalar times the identity matrix

$$
W_{b_i} = R_{b_i}^{-1} = \frac{1}{\sigma_{b_i}^2} I_{3\times3}, \quad W_{r_i} = R_{r_i}^{-1} = \frac{1}{\sigma_{r_i}^2} I_{3\times3} \tag{41}
$$

then the attitude error covariance reduces to

$$
P_{\delta\alpha} = -\sum_{i=1}^{n}\frac{1}{\sigma_{b_i}^2 + \sigma_{r_i}^2}[\mathbf{b}_i\times]^2 \tag{42}
$$

which has the same form as in Wahba's problem [10].

B. Problem 2

Under the small error approximation, the estimation errors minimize the loss function

$$
L(\boldsymbol{\delta\alpha}, \boldsymbol{\delta}\mathbf{r}_1, \ldots, \boldsymbol{\delta}\mathbf{r}_n) = \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\Delta}\mathbf{b}_i - [\mathbf{b}_i\times]\boldsymbol{\delta\alpha} - A\boldsymbol{\delta}\mathbf{r}_i)^T W_{b_i}(\boldsymbol{\Delta}\mathbf{b}_i - [\mathbf{b}_i\times]\boldsymbol{\delta\alpha} - A\boldsymbol{\delta}\mathbf{r}_i) + \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{\Delta}\mathbf{r}_i - \boldsymbol{\delta}\mathbf{r}_i)^T W_{r_i}(\boldsymbol{\Delta}\mathbf{r}_i - \boldsymbol{\delta}\mathbf{r}_i) \tag{43}
$$

subject to the linear equality constraints

$$
\mathbf{r}_i^T \boldsymbol{\delta}\mathbf{r}_i = 0 \tag{44}
$$

The augmented loss function with the Lagrange multipliers λ_i is given by

$$
L^a(\boldsymbol{\delta\alpha}, \boldsymbol{\delta}\mathbf{r}_1, \ldots, \boldsymbol{\delta}\mathbf{r}_n) = L(\boldsymbol{\delta\alpha}, \boldsymbol{\delta}\mathbf{r}_1, \ldots, \boldsymbol{\delta}\mathbf{r}_n) + \sum_{i=1}^{n}\lambda_i \mathbf{r}_i^T \boldsymbol{\delta}\mathbf{r}_i \tag{45}
$$

and the optimality conditions are

$$
M^a \boldsymbol{\Delta}\mathbf{x}^a = N^a \boldsymbol{\Delta}\mathbf{y} \tag{46}
$$

with

$$
\boldsymbol{\Delta}\mathbf{x}^a = \begin{bmatrix} \boldsymbol{\Delta}\mathbf{x} \\ \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix} \tag{47a}
$$

$$
\tag{47b}
$$

$$
M^a \triangleq M^a(A, \mathbf{b}_1, \ldots, \mathbf{b}_n) = \begin{bmatrix} M & C^T \\ C & 0_{n\times n} \end{bmatrix} \tag{47c}
$$

$$
C = \left[\begin{array}{c|ccc} & \mathbf{r}_1^T & 0_{1\times3(n-2)} & 0_{1\times3} \\ 0_{n\times3} & 0_{(n-2)\times3} & \ddots & 0_{(n-2)\times3} \\ & 0_{1\times3} & 0_{1\times3(n-2)} & \mathbf{r}_n^T \end{array}\right] \tag{47d}
$$

$$N^a \triangleq N^a(A, \mathbf{b}_1, \ldots, \mathbf{b}_n) = \begin{bmatrix} N \\ 0_{n\times 6n} \end{bmatrix} \tag{47e}$$

So, the error vector and the associated covariance matrix are respectively given by

$$\mathbf{\Delta x}^a = (M^a)^{-1} N^a \mathbf{\Delta y} \tag{48}$$

and

$$P_{\mathbf{\Delta x}^a} = E\{\mathbf{\Delta x}^a\,(\mathbf{\Delta x}^a)^T\} = (M^a)^{-1} N^a R (N^a)^T (M^a)^{-1} \tag{49}$$

Note that the covariances of $\boldsymbol{\delta}\mathbf{r}_i$ are singular because of the equality constraints.

Special Cases

If $W_{b_i} = R_{b_i}^{-1}$ and $W_{r_i} = R_{r_i}^{-1}$, then the covariance matrix becomes

$$P_{\mathbf{\Delta x}^a} = \begin{bmatrix} M & C^T \\ C & 0_{n\times n} \end{bmatrix}^{-1} \begin{bmatrix} M & 0_{3(n+1)\times n} \\ 0_{n\times(3(n+1)} & 0_{n\times n} \end{bmatrix} \begin{bmatrix} M & C^T \\ C & 0_{n\times n} \end{bmatrix}^{-1} \tag{50}$$

If

$$W_{b_i} = R_{b_i}^{-1} = \frac{1}{\sigma_{b_i}^2} I_{3\times 3}, \quad W_{r_i} = R_{r_i}^{-1} = \frac{1}{\sigma_{r_i}^2} I_{3\times 3} \tag{51}$$

then it can be shown that the attitude error covariance is

$$P_{\delta\alpha} = -\sum_{i=1}^{n} \frac{1}{\sigma_{b_i}^2 + \sigma_{r_i}^2} [\mathbf{b}_i\times]^2 \tag{52}$$

So, in the special case of scalar weights, Problems 1 and 2 have the same attitude error covariance (to first order only).

IV. Algorithm Development

In both Problem 1 and Problem 2, the optimal vector estimates $\hat{\mathbf{r}}_i \triangleq \hat{\mathbf{r}}_i(\hat{A}, \tilde{\mathbf{b}}_i, \tilde{\mathbf{r}}_i)$ have an closed-form expression in terms of the optimal $\hat{A}$. The resultant loss function for the attitude matrix is

$$\begin{aligned} L(A) &\triangleq L\left(A, \hat{\mathbf{r}}_1(A, \tilde{\mathbf{b}}_1, \tilde{\mathbf{r}}_1), \ldots, \hat{\mathbf{r}}_n(A, \tilde{\mathbf{b}}_n, \tilde{\mathbf{r}}_n)\right) \\ &= \frac{1}{2}\sum_{i=1}^{n} (\tilde{\mathbf{b}}_i - A\hat{\mathbf{r}}_i)^T W_{b_i} (\tilde{\mathbf{b}}_i - A\hat{\mathbf{r}}_i) + \frac{1}{2}\sum_{i=1}^{n} (\tilde{\mathbf{r}}_i - \hat{\mathbf{r}}_i)^T W_{r_i} (\tilde{\mathbf{r}}_i - \hat{\mathbf{r}}_i) \end{aligned} \tag{53}$$

It will be shown that except for Problem 1 with scalar weights, the loss function is not equivalent to that of Wahba's problem. Deriving the gradient and/or Hessian of $L(A)$ required by an gradient based algorithm is complicated, especially for Problem 2. Problems 1 and 2 are therefore solved iteratively based on the linear error models in the covariance analysis section.

A. Solution of Problem 1

The solution $\hat{\mathbf{r}}_i(\hat{A}, \tilde{\mathbf{b}}_i, \tilde{\mathbf{r}}_i)$ is given by

$$\hat{\mathbf{r}}_i = (\hat{A}^T W_{b_i} \hat{A} + W_{r_i})^{-1} \left(\hat{A}^T W_{b_i} \tilde{\mathbf{b}}_i + W_{r_i} \tilde{\mathbf{r}}_i\right) \tag{54}$$

Note that $\hat{\mathbf{r}}_i$ is nonparallel to $(\hat{A}^T W_{b_i} \tilde{\mathbf{b}}_i + W_{r_i} \tilde{\mathbf{r}}_i)$ in general. Substituting Eq. (54) into Eq. (22) gives

$$L(A) = \frac{1}{2}\sum_{i=1}^{n} (\tilde{\mathbf{b}}_i - A\tilde{\mathbf{r}}_i)^T \left(W_{b_i}^{-1} + A W_{r_i}^{-1} A^T\right)^{-1} (\tilde{\mathbf{b}}_i - A\tilde{\mathbf{r}}_i) \tag{55}$$

An iterative algorithm is used to solve Problem 1:

1) Solve Wahba's problem with the following loss function for the initial estimate $\hat{A}^{(0)}$

$$L(A) = \frac{1}{2} \sum_{i=1}^{n} \frac{1}{\mathrm{tr}\left(W_{b_i}^{-1} + W_{r_i}^{-1}\right)} \|\tilde{\mathbf{b}}_i - A\tilde{\mathbf{r}}_i\|^2 \tag{56}$$

(Replace matrix inversion with pseudo-matrix inversion if the weighting matrices are singular.)
2) Compute $\hat{\mathbf{r}}_i^{(k)}$ from $\hat{A}^{(k)}$ using Eq. (54)
3) Compute $\hat{\mathbf{b}}_i^{(k)} = \hat{A}^{(k)}\,\hat{\mathbf{r}}_i^{(k)}$
4) Compute $\mathbf{\Delta b}_i^{(k)}$ and $\mathbf{\Delta r}_i^{(k)}$

$$\mathbf{\Delta b}_i^{(k)} = \tilde{\mathbf{b}}_i - \hat{\mathbf{b}}_i^{(k)}, \quad \mathbf{\Delta r}_i^{(k)} = \tilde{\mathbf{r}}_i - \hat{\mathbf{r}}_i^{(k)} \tag{57}$$

5) Construct $M^{(k)}$, $N^{(k)}$, and $\mathbf{\Delta y}^{(k)}$ using $\hat{A}^{(k)}$, $\hat{\mathbf{b}}_i^{(k)}$, $\mathbf{\Delta b}_i^{(k)}$, and $\mathbf{\Delta r}_i^{(k)}$
6) Compute $\mathbf{\Delta x}^{(k)}$ using Eq. (35)
7) Retrieve $\boldsymbol{\delta\alpha}^{(k)}$ from $\mathbf{\Delta x}^{(k)}$
8) Increase k by one and compute $\hat{A}^{(k+1)}$

$$\hat{A}^{(k+1)} = \exp([-\boldsymbol{\delta\alpha}^{(k)}\times])\,\hat{A}^{(k)} \tag{58}$$

9) Repeat steps 2 through 8 until $\mathbf{\Delta x}^{(k)}$ is sufficiently small

The final estimates are given by $\hat{A}^{(k)}$ and $\hat{\mathbf{r}}_i^{(k)}$.

B. Solution of Problem 2

Solving Problem 2 for the optimal $\hat{\mathbf{r}}_i$ under the quadratic equality constraints gives

$$\hat{\mathbf{r}}_i = (\hat{A}^T W_{b_i} \hat{A} + W_{r_i} + \lambda_i I_{3\times3})^{-1} \left(\hat{A}^T W_{b_i} \tilde{\mathbf{b}}_i + W_{r_i} \tilde{\mathbf{r}}_i\right) \tag{59}$$

where the Lagrange multipliers λ_i are such that the norm of $\hat{\mathbf{r}}_i$ is one:

$$\left(\hat{A}^T W_{b_i} \tilde{\mathbf{b}}_i + W_{r_i} \tilde{\mathbf{r}}_i\right)^T (\hat{A}^T W_{b_i} \hat{A} + W_{r_i} + \lambda_i I_{3\times3})^{-2} \left(\hat{A}^T W_{b_i} \tilde{\mathbf{b}}_i + W_{r_i} \tilde{\mathbf{r}}_i\right) = 1 \tag{60}$$

Note that the Lagrange multipliers obtained by solving Eq. (46) only satisfy the above equation approximately. In general, $\hat{\mathbf{r}}_i$ is nonparallel to $(\hat{A}^T W_{b_i} \tilde{\mathbf{b}}_i + W_{r_i} \tilde{\mathbf{r}}_i)$. The $L(A)$ has an even more complex form in Problem 2 than Problem 1.

An iterative algorithm is used to solve Problem 2:
1) Solve Wahba's problem with the following loss function for the initial estimate $\hat{A}^{(0)}$

$$L(A) = \frac{1}{2} \sum_{i=1}^{n} \frac{1}{\mathrm{tr}\left(W_{b_i}^{-1} + W_{r_i}^{-1}\right)} \|\tilde{\mathbf{b}}_i - A\tilde{\mathbf{r}}_i\|^2 \tag{61}$$

(Replace matrix inversion with pseudo matrix inversion if the weighting matrices are singular.)
2) Solve Eq. (60) for $\lambda^{(k)}$
3) Compute $\hat{\mathbf{r}}_i^{(k)}$ from $\hat{A}^{(k)}$ and $\lambda^{(k)}$ using Eq. (59)
4) Compute $\hat{\mathbf{b}}_i^{(k)} = \hat{A}^{(k)}\,\hat{\mathbf{r}}_i^{(k)}$
5) Compute $\mathbf{\Delta b}_i^{(k)}$ and $\mathbf{\Delta r}_i^{(k)}$

$$\mathbf{\Delta b}_i^{(k)} = \tilde{\mathbf{b}}_i - \hat{\mathbf{b}}_i^{(k)}, \quad \mathbf{\Delta r}_i^{(k)} = \tilde{\mathbf{r}}_i - \hat{\mathbf{r}}_i^{(k)} \tag{62}$$

6) Construct $M^{a(k)}$, $N^{a(k)}$, and $\mathbf{\Delta y}^{(k)}$ using $\hat{A}^{(k)}$, $\hat{\mathbf{b}}_i^{(k)}$, $\mathbf{\Delta b}_i^{(k)}$, and $\mathbf{\Delta r}_i^{(k)}$
7) Compute $\mathbf{\Delta x}^{a(k)}$ using Eq. (48)
8) Retrieve $\boldsymbol{\delta\alpha}^{(k)}$ from $\mathbf{\Delta x}^{a(k)}$
9) Increase k by one and compute $\hat{A}^{(k+1)}$

$$\hat{A}^{(k+1)} = \exp([-\boldsymbol{\delta\alpha}^{(k)}\times])\,\hat{A}^{(k)} \tag{63}$$

10) Repeat steps 2 through 9 until $\mathbf{\Delta x}^{(k)}$ is sufficiently small

The final estimates are given by $\hat{A}^{(k)}$ and $\hat{\mathbf{r}}_i^{(k)}$.

C. Scalar Weights

If all the weights are scalars,

$$W_{b_i} = w_{b_i} I_{3\times 3}, \quad W_{r_i} = w_{r_i} I_{3\times 3} \tag{64}$$

Problem 1 is equivalent to Wahba's problem because the loss function becomes

$$L(A) = \frac{1}{2}\sum_{i=1}^{n} \frac{1}{w_{b_i}^{-1} + w_{r_i}^{-1}} \|\tilde{\mathbf{b}}_i - A\tilde{\mathbf{r}}_i\|^2 \tag{65}$$

It can be solved without iteration. In this case, the corrections $\mathbf{\Delta x}^{(k)}$ in step 6) of the solution vanish.

The optimal estimates for $\hat{\mathbf{r}}_i$ are given by

$$\hat{\mathbf{r}}_i = \frac{w_{b_i}\hat{A}^T\tilde{\mathbf{b}}_i + w_{r_i}\tilde{\mathbf{r}}_i}{w_{b_i} + w_{r_i}} \tag{66}$$

If $\hat{A}^T\tilde{\mathbf{b}}_i$ and $\tilde{\mathbf{r}}_i$ are nonparallel unit vectors, $\hat{\mathbf{r}}_i$ are not unit vectors.

Problem 2 is not equivalent to Wahba's problem. The Lagrange multipliers and $\hat{\mathbf{r}}_i$ are given by

$$\lambda_i = \|w_{b_i}\hat{A}^T\tilde{\mathbf{b}}_i + w_{r_i}\tilde{\mathbf{r}}_i\| - (w_{b_i} + w_{r_i}) \tag{67a}$$

$$\hat{\mathbf{r}}_i = \frac{w_{b_i}\hat{A}^T\tilde{\mathbf{b}}_i + w_{r_i}\tilde{\mathbf{r}}_i}{\|w_{b_i}\hat{A}^T\tilde{\mathbf{b}}_i + w_{r_i}\tilde{\mathbf{r}}_i\|} \tag{67b}$$

Equation (67b) can be written as

$$\hat{\mathbf{r}}_i = \tilde{w}_{b_i}\hat{A}^T\tilde{\mathbf{b}}_i + \tilde{w}_{r_i}\tilde{\mathbf{r}}_i \tag{68}$$

with

$$\tilde{w}_{b_i} = \frac{w_{b_i}}{\|w_{b_i}\hat{A}^T\tilde{\mathbf{b}}_i + w_{r_i}\tilde{\mathbf{r}}_i\|}, \quad \tilde{w}_{r_i} = \frac{w_{r_i}}{\|w_{r_i}\hat{A}^T\tilde{\mathbf{b}}_i + w_{r_i}\tilde{\mathbf{r}}_i\|} \tag{69}$$

Note that $\tilde{w}_{b_i}$ and $\tilde{w}_{r_i}$ are attitude dependent and $\tilde{w}_{b_i} + \tilde{w}_{r_i} \neq 1$ unless $\hat{A}^T\tilde{\mathbf{b}}_i$ and $\tilde{\mathbf{r}}_i$ are identical unit vectors. Hence, the loss function given by

$$L(A) = \frac{1}{2}\sum_{i=1}^{n} w_{b_i}\|(1-\tilde{w}_{b_i})\tilde{\mathbf{b}}_i - \tilde{w}_{r_i}A\tilde{\mathbf{r}}_i\|^2 + \frac{1}{2}\sum_{i=1}^{n} w_{r_i}\|\tilde{w}_{b_i}\tilde{\mathbf{b}}_i - (1-\tilde{w}_{r_i})A\tilde{\mathbf{r}}_i\|^2 \tag{70}$$

cannot reduce to that of Wahba's problem.

On the other hand, if the measurement and attitude estimation errors are small, then $\hat{A}^T\tilde{\mathbf{b}}_i \approx \tilde{\mathbf{r}}_i$ because $A^T\mathbf{b}_i = \mathbf{r}_i$. In this case,

$$\tilde{w}_{b_i} \approx \frac{w_{b_i}}{w_{b_i} + w_{r_i}}, \quad \tilde{w}_{r_i} \approx \frac{w_{r_i}}{w_{b_i} + w_{r_i}} \tag{71}$$

and

$$L(A) \approx \frac{1}{2}\sum_{i=1}^{n} \frac{1}{w_{b_i}^{-1} + w_{r_i}^{-1}} \|\tilde{\mathbf{b}}_i - A\tilde{\mathbf{r}}_i\|^2 \tag{72}$$

V. Numerical Example

An example of attitude determination from two coarse line-of-sight measurements is used to compare Problem 1 and Problem 2 when all the weights are scalars. The true attitude matrix is $A = I_{3\times 3}$. The true vectors in the reference and body frames are

$$\mathbf{b}_1 = \mathbf{r}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \quad \mathbf{b}_2 = \mathbf{r}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix} \tag{73}$$

The measurements (unit vectors) are

$$\tilde{\mathbf{b}}_1 = \begin{bmatrix}0.9940\\0.0868\\-0.0664\end{bmatrix}, \quad \tilde{\mathbf{r}}_1 = \begin{bmatrix}0.9906\\-0.1197\\-0.0666\end{bmatrix}, \quad \tilde{\mathbf{b}}_2 = \begin{bmatrix}0.1186\\0.9886\\0.0924\end{bmatrix}, \quad \tilde{\mathbf{r}}_2 = \begin{bmatrix}-0.1232\\0.9923\\0.0126\end{bmatrix} \tag{74}$$

The angular measurement errors are

$$\tilde{\mathbf{b}}_1^T\mathbf{b}_1 = 6.2755°, \quad \tilde{\mathbf{r}}_1^T\mathbf{r}_1 = 7.8747°, \quad \tilde{\mathbf{b}}_2^T\mathbf{b}_2 = 8.6481°, \quad \tilde{\mathbf{r}}_2^T\mathbf{r}_2 = 7.1164° \tag{75}$$

The scalar weights are chosen as

$$w_{b_1} = \frac{1}{(2°)^2}, \quad w_{r_1} = \frac{1}{(2°)^2}, \quad w_{b_2} = \frac{1}{(3°)^2}, \quad w_{r_2} = \frac{1}{(3°)^2} \tag{76}$$

The solution of Problem 1 is

$$\hat{A} = \begin{bmatrix} 0.9979 & -0.0647 & 0.0085 \\ 0.0652 & 0.9927 & -0.1019 \\ -0.0018 & 0.1022 & 0.9948 \end{bmatrix} \tag{77}$$

The solution of Problem 2 is

$$\hat{A} = \begin{bmatrix} 0.9980 & -0.0629 & 0.0085 \\ 0.0635 & 0.9928 & -0.1018 \\ -0.0020 & 0.1021 & 0.9948 \end{bmatrix} \tag{78}$$

The difference between the two attitude estimates is 0.1017°, small compared with the estimation errors of approximately 7°, but the two solutions are clearly not identical.

A Monte Carlo analysis is done to assess how well the computed 3σ bounds using the unconstrained covariance in Eq. (39) and the constrained covariance in Eq. (50) bound the actual respective errors. The true vectors in the reference and body frames for this analysis are

$$\mathbf{b}_1 = \mathbf{r}_1 = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{b}_2 = \mathbf{r}_2 = \frac{\sqrt{2}}{2}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \tag{79}$$

Noise is generated using a zero-mean Gaussian noise process with standard deviations of 2 degrees for both $\mathbf{b}_1$ and $\mathbf{r}_1$, and standard deviations of 3 degrees for both $\mathbf{b}_2$ and $\mathbf{r}_2$. Five thousand Monte Carlo runs are executed. Plots of the attitude errors along with their respective 3σ bounds for the unconstrained and constrained solutions are shown in Figs. 1(a) and 1(b). The 3σ bounds are equivalent in this case because the weighting matrices are isotropic. But, the actual errors are not the same, which reinforces the previous example results. Figures 1(c) and 1(d) show the errors in the $\mathbf{r}_1$ estimates along with their respective 3σ bounds for the unconstrained and constrained solutions. All the errors are well bounded by their respective 3σ bounds. Comparing Figs. 1(c) and 1(d) shows that the bounds and errors for the constrained solution are smaller than the unconstrained solution. This shows that forcing the constraint in the solution produces more accurate results.

VI. Conclusions

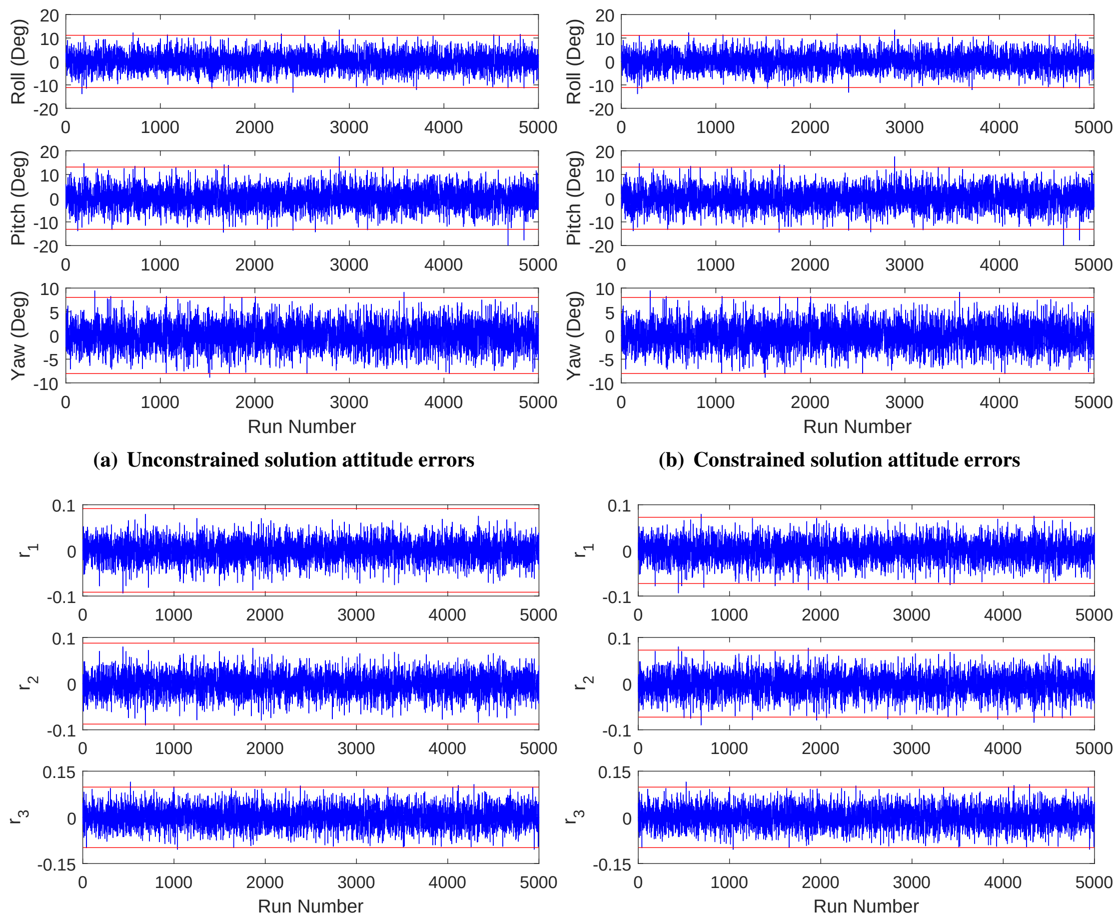
Most methods for attitude determination from vector observations are solutions of Wahba's problem, a least-squares estimation problem with a scalar weight for each pair of vectors. This paper proposes to solve the attitude estimation problem as a total least-squares problem. Two total least-squares problems are formulated without the restriction that the vector weights are scalars and solved using an iterative procedure. When all the weights are scalars, the total least-squares problem does not reduce to Wahba's problem if the estimated vectors need to obey the unity norm constraint.

References

[1] Wahba, G., "A Least Squares Estimate of Satellite Attitude," *SIAM Review*, Vol. 7, No. 3, 1965, p. 409. d0i:10.1137/1007077.

[2] Markley, F. L., and Mortari, D., "Quaternion Attitude Estimation Using Vector Observations," *The Journal of the Astronautical Sciences*, Vol. 48, No. 2/3, 2000, pp. 359–380.

[3] Markley, F. L., and Crassidis, J. L., *Fundamentals of Spacecraft Attitude Determination and Control*, Springer, New York, NY, 2014, Chap. 5.



(a) Unconstrained solution attitude errors

(b) Constrained solution attitude errors

(c) Unconstrained $\mathbf{r}_1$ errors

(d) Constrained $\mathbf{r}_2$ errors

Fig. 1 Monte Carlo results.

[4] Schönemann, P., "A Generalized Solution of the Orthogonal Procrustes Problem," *Psychometrika*, Vol. 31, No. 1, 1966, pp. 1–10.

[5] Eggert, D. W., Lorusso, A., and Fisher, R. B., "Estimating 3-D Rigid Body Transformations: A Comparison of Four Major Algorithms," *Machine Vision and Applications*, Vol. 9, No. 5, 1997, pp. 272–290.

[6] Van Huffel, S., and Vandewalle, J., *The Total Least Squares Problem: Computational Aspects and Analysis*, Society for Industial and Applied Mathematics, Philadelphia, PA, 1991, Chap. 2.

[7] Chang, G., "Total Least-Squares Formulation of Wahba's Problem," *Electronics Letters*, Vol. 51, No. 17, 2015, pp. 334–1335.

[8] Arun, K., "A Unitarily Constrained Total Least-Squares Problem in Signal-Processing," *SIAM Journal on Matrix Analysis and Applications*, Vol. 13, 1992, pp. 729–745.

[9] Crassidis, J. L., and Cheng, Y., "Maximum Likelihood Analysis of the Total Least Squares Problem with Correlated Errors," *AIAA Guidance, Navigation, and Control Conference*, San Diego, CA, 2019.

[10] Shuster, M. D., and Oh, S. D., "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, 1981, pp. 70–77. doi:10.2514/3.19717.