Observability Analysis of Six Degree of Freedom Configuration Determination Using Vector Observations

Debo Sun*

John L. Crassidis†

Abstract

In this paper an observability analysis of the six degree of freedom attitude and position determination problem using line-of-sight observations is shown. This analysis involves decompositions of the associated error covariance matrix, derived from maximum likelihood, for a number of cases ranging from one vector observation to three or more vector observations. The covariance matrix is shown to be singular when one or two vector observations are used, leading to an unobservable system. For the one vector case the observable quantities involve a combination of both attitude and position information that cannot be decoupled. For the two vector case the covariance matrix has rank four, but only one axis of attitude and one axis of position is fully observable, with the other two observable quantities involving coupled attitude/position information. When three or more vector observations are present the covariance matrix has full rank, except for some special cases that are derived in this paper. This observability analysis is useful for the design and analysis of estimators using line-of-sight vector observations.

*Visiting Scholar at the Aerospace Engineering Department, Texas A&M University, College Station, TX 77843-3141, from Department of Control Science and Engineering, Harbin Institute of Technology, Harbin 150001, P. R. China.

†Assistant Professor, Department of Mechanical & Aerospace Engineering, University at Buffalo, State University of New York, Amherst, NY 14260-4400. Senior Member AIAA.

Introduction

Both the attitude and the position of a vehicle can be determined from line-of-sight (LOS) vector observations. One mechanism to accomplish this task involves a vision navigation (VISNAV) system based on Position Sensing Diodes (PSDs) in the focal plane of a camera, which allows the inherent centroiding of a LED beacon's incident light.[1] Other mechanisms may involve camera image measurements or laser reflector LOS measurements. The fundamental approach used to determine the attitude and position from LOS observations involves an object to image projective transformation, achieved through the *colinearity equations*.[2] These equations involve the angle of the body from the sensor boresight in two mutually orthogonal planes, which can be reconstructed into unit vector form. The most common approach to determine attitude and position using the colinearity equations involves a Gaussian Least Squares Differential Corrections (GLSDC) process, while a new estimation approach has been presented in Ref. [3] based on a predictive filter for nonlinear systems.

Determining attitude from LOS observations commonly involves finding a proper orthogonal matrix that minimizes the scalar weighted norm-error between sets of 3×1 body vector observations and 3×1 known reference vectors mapped (via the attitude matrix) into the body frame. This is known as Wahba's problem.[4] If the reference vectors are known, then at least two non-colinear unit vector observations are required to determine the attitude. Many methods have been developed that solve this problem efficiently and accurately.[5,6] Determining the position from LOS observations involves triangulation from known reference base points. If the attitude is known, then at least two non-colinear unit vector observations are required to establish a three-dimensional position. Determining both attitude and position from LOS observations is more complex since more than two non-colinear unit vector observations are required (as will be demonstrated in this paper), and, unlike Wahba's problem, the unknown attitude and position are interlaced in a highly nonlinear fashion.

In this paper an analysis is performed to study the observability of the coupled attitude and position determination problem from vector observations. In Ref. [3] an initial study has been performed for the two vector observation case, which showed that only one axis of attitude and one axis of position information can be determined for this case. Furthermore, an observability analysis using two vector observations indicates that the beacon that is closest to the target provides the most attitude information but has the least position information, and the beacon that is farthest to the target provides the most position information but has the least attitude information. This paper extends this initial result for the one and three or more vector observation cases, and also more fully quantifies the two vector observation case.

The organization of this paper proceeds as follows. First, a review of the colinearity equations is shown. Then, a generalized loss function derived from maximum likelihood for attitude and position determination is given. Next, the optimal estimate covariance is derived, which gives the Cramér-Rao lower bound. Then, an observability analysis is shown for cases involving one to three or more vector observations. This analysis is performed using an eigenvalue/eigenvector decomposition of the information matrix (i.e., the inverse of the covariance matrix). Finally, the trace and eigenvalues of the covariance matrix are studied.

The Colinearity Equations and Covariance

In this section the colinearity equations for attitude and position determination are shown. First, the observation model is reviewed. Then, the estimate (attitude and position) covariance matrix is derived using maximum likelihood.

Colinearity Equations

Photogrammetry is the technique of measuring objects (2D or 3D) from photographic images or LOS measurements. Photogrammetry can generally be divided into two categories: far range photogrammetry with camera distance settings to infinity (commonly used in star cameras[7]), and close range photogrammetry with camera distance settings to finite values. In general close range photogrammetry can be used to determine both the position and attitude of an object, while far range photogrammetry can only be used to determine attitude.

The relationship between the position/attitude and the observations used in photogrammetry involves a set of colinearity equations, which are reviewed in this section. Figure 1 shows a schematic of the typical quantities involved in basic photogrammetry from LOS measurements, derived from light beacons in this case. If we choose the z-axis of the sensor coordinate system to be directed outward along the boresight, then given object space (X, Y, Z) and image space (x, y, z) coordinate frames (see Fig. 1), the ideal object to image space projective transformation (noiseless) can be written as follows:[8]

$$x_i = -f\frac{A_{11}(X_i - X_c) + A_{12}(Y_i - Y_c) + A_{13}(Z_i - Z_c)}{A_{31}(X_i - X_c) + A_{32}(Y_i - Y_c) + A_{33}(Z_i - Z_c)}, \quad i = 1,\, 2,\, \ldots,\, N \tag{1a}$$

$$y_i = -f\frac{A_{21}(X_i - X_c) + A_{22}(Y_i - Y_c) + A_{23}(Z_i - Z_c)}{A_{31}(X_i - X_c) + A_{32}(Y_i - Y_c) + A_{33}(Z_i - Z_c)}, \quad i = 1,\, 2,\, \ldots,\, N \tag{1b}$$

where N is the total number of observations, (x_i, y_i) are the image space observations for the i^{th} line-of-sight, (X_i, Y_i, Z_i) are the known object space locations of the i^{th} beacon, (X_c, Y_c, Z_c) are the unknown object space location of the sensor, f is the known focal length, and A_{jk} are the unknown coefficients of the attitude matrix (A) associated to the orientation from the object plane to the image plane. The goal of the *inverse problem* is given observations (x_i, y_i) and object space locations (X_i, Y_i, Z_i), for $i = 1, 2, \ldots, N$, determine the attitude (A) and position (X_c, Y_c, Z_c). This can be accomplished by using a GLSDC process or by other methods.[3]

The observation can be reconstructed in unit vector form as

$$\mathbf{b}_i = A\mathbf{r}_i, \quad i = 1,\, 2,\, \ldots,\, N \tag{2}$$

where

$$\mathbf{b}_i \equiv \frac{1}{\sqrt{f^2 + x_i^2 + y_i^2}} \begin{bmatrix} -x_i \\ -y_i \\ f \end{bmatrix} \tag{3a}$$

$$\mathbf{r}_i \equiv \frac{1}{\sqrt{(X_i - X_c)^2 + (Y_i - Y_c)^2 + (Z_i - Z_c)^2}} \begin{bmatrix} X_i - X_c \\ Y_i - Y_c \\ Z_i - Z_c \end{bmatrix} \tag{3b}$$

When measurement noise is present, Shuster[5] has shown that nearly all the probability of the errors is concentrated on a very small area about the direction of $A\mathbf{r}_i$, so the sphere containing that point can be approximated by a tangent plane, characterized by

$$\tilde{\mathbf{b}}_i = A\mathbf{r}_i + \boldsymbol{\upsilon}_i, \quad \boldsymbol{\upsilon}_i^T A\mathbf{r}_i = 0 \tag{4}$$

where $\tilde{\mathbf{b}}_i$ denotes the i^{th} measurement, and the sensor error $\boldsymbol{\upsilon}_i$ is approximately Gaussian which satisfies

$$E\left\{\boldsymbol{\upsilon}_i\right\} = \mathbf{0} \tag{5a}$$

$$E\left\{\boldsymbol{\upsilon}_i \boldsymbol{\upsilon}_i^T\right\} = \sigma_i^2 \left[I - (A\mathbf{r}_i)(A\mathbf{r}_i)^T\right] \tag{5b}$$

and $E\{\,\}$ denotes expectation. Equation (5b) makes the small field-of-view assumption of Ref. [5]; however, for a large field-of-view lens with significant radial distortion, this covariance model should be modified appropriately.

Maximum Likelihood Estimation and Covariance

Attitude and position determination using LOS measurements involves finding estimates of the proper orthogonal matrix A and position vector $\mathbf{p} \equiv [X_c \; Y_c \; Z_c]^T$ that minimize the following loss function:

$$J(\hat{A},\,\hat{\mathbf{p}}) = \frac{1}{2}\sum_{i=1}^{N}\sigma_i^{-2}\|\tilde{\mathbf{b}}_i - \hat{A}\hat{\mathbf{r}}_i\|^2 \tag{6}$$

where $\wedge$ denotes estimate. An estimate error covariance can be derived from the loss function in Eq. (6). This is accomplished by using results from maximum likelihood estimation.[5,9] The Fisher information matrix for a parameter vector $\mathbf{x}$ is given by

$$F_{xx} = E\left\{\frac{\partial}{\partial\mathbf{x}\,\partial\mathbf{x}^T}J(\mathbf{x})\right\}_{\mathbf{x}_{\text{true}}} \tag{7}$$

where $J(\mathbf{x})$ is the negative log-likelihood function, which is the loss function in this case (neglecting terms independent of A and $\mathbf{p}$). Asymptotically, the Fisher information matrix tends to the inverse of the estimate error covariance so that $\lim_{N\to\infty} F_{xx} = P^{-1}$. The true attitude matrix is approximated by

$$A = e^{-[\boldsymbol{\delta\alpha}\times]}\hat{A} \approx (I_{3\times 3} - [\boldsymbol{\delta\alpha}\times])\,\hat{A} \tag{8}$$

where $\boldsymbol{\delta\alpha}$ represents a small angle error and $I_{3\times 3}$ is a 3×3 identity matrix. The 3×3 matrix $[\boldsymbol{\delta\alpha}\times]$ is referred to as a cross-product matrix because $\mathbf{a}\times\mathbf{b} = [\mathbf{a}\times]\,\mathbf{b}$, with

$$[\mathbf{a}\times] \equiv \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \tag{9}$$

The parameter vector is now given by $\mathbf{x} = [\boldsymbol{\delta\alpha}^T\ \ \hat{\mathbf{p}}^T]^T$, and the covariance is defined by $P = E\left\{\mathbf{x}\,\mathbf{x}^T\right\} - E\left\{\mathbf{x}\right\}E\left\{\mathbf{x}\right\}^T$. Substituting Eq. (8) into Eq. (6), and after taking the appropriate partials the following optimal error covariance can be derived:

$$P = \begin{bmatrix} -\sum_{i=1}^{N} \sigma_i^{-2} [A\,\mathbf{r}_i \times]^2 & \sum_{i=1}^{N} \sigma_i^{-2} \zeta_i A\,[\mathbf{r}_i \times] \\ \sum_{i=1}^{N} \sigma_i^{-2} \zeta_i [\mathbf{r}_i \times]^T A^T & -\sum_{i=1}^{N} \sigma_i^{-2} \zeta_i^2 [\mathbf{r}_i \times]^2 \end{bmatrix}^{-1} \equiv F^{-1} \tag{10}$$

with obvious definition for F, and where

$$\zeta_i \equiv \left[(X_i - X_c)^2 + (Y_i - Y_c)^2 + (Z_i - Z_c)^2 \right]^{-1/2} \tag{11}$$

The terms A and $\mathbf{r}_i$ are evaluated at their respective *true* values (although in practice the estimates are used). It should be noted that Eq. (10) gives the Cramér-Rao lower bound[9] (any estimator whose error covariance is equivalent to Eq. (10) is an *efficient*, i.e. optimal estimator). Also, Eq. (10) is directly used in the GLSDC process and predictive filter solution.[3]

The matrix F in Eq. (10) must have rank 6 in order for P to exist. The remainder of this paper is devoted to the analysis of the matrix F for a number of vector observation cases. We first prove that the rank of F is independent of the attitude matrix A. Since A is a proper orthogonal matrix, then $A\,A^T = A^T A = I_{3\times3}$. Also, the following identity is helpful:

$$[A\mathbf{r}\times] = A\,[\mathbf{r}\times]\,A^T \tag{12}$$

Next, a similarity transformation is performed using the following orthogonal matrix:

$$\mathcal{M} = \begin{bmatrix} A & 0_{3\times3} \\ 0_{3\times3} & I_{3\times3} \end{bmatrix} \tag{13}$$

where $0_{3\times3}$ is a 3×3 zero matrix. Defining $\mathcal{F} \equiv \mathcal{M}^T F\,\mathcal{M}$, and using the identity in Eq. (12)

gives

$$\mathcal{F} = \begin{bmatrix} -\sum_{i=1}^{N} \sigma_i^{-2} [\mathbf{r}_i \times]^2 & \sum_{i=1}^{N} \sigma_i^{-2} \zeta_i [\mathbf{r}_i \times] \\ \\ \sum_{i=1}^{N} \sigma_i^{-2} \zeta_i [\mathbf{r}_i \times]^T & -\sum_{i=1}^{N} \sigma_i^{-2} \zeta_i^2 [\mathbf{r}_i \times]^2 \end{bmatrix} \tag{14}$$

Since Rank($\mathcal{M}$) = 6, then Rank(F) = Rank($\mathcal{F}$), which indicates that the degree of observability (i.e., the rank of F) of the system is independent of the attitude matrix. This intuitively makes sense since the orientation of the body with respect to the beacon LOS sources does not affect the overall observability (it does however affect the relative degree of observability of each axis component).

One Vector Observation Case

In this section the one vector case is analyzed. Although in practice one observation would not be used, this case is worthy of study since as the range to multiple beacons becomes large, the angular separation decreases and the beacons ultimately approach co-location. The result is a geometric dilution of precision, and ultimately, a loss of observability analogous to the one beacon case. The rank of the information matrix is first investigated. For this case $\mathcal{F}$ is given by

$$\mathcal{F} = \sigma^{-2} \begin{bmatrix} -[\mathbf{r} \times]^2 & \zeta [\mathbf{r} \times] \\ \\ \zeta [\mathbf{r} \times]^T & -\zeta^2 [\mathbf{r} \times]^2 \end{bmatrix} \equiv \sigma^{-2} M \tag{15}$$

with obvious definition for M. Using the following matrix

$$\mathcal{N} = \begin{bmatrix} \zeta [\mathbf{r} \times]^T & I_{3\times3} \\ I_{3\times3} & 0_{3\times3} \end{bmatrix} \tag{16}$$

we have

$$\mathcal{N}^T M \mathcal{N} = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & -[\mathbf{r} \times]^2 \end{bmatrix} \tag{17}$$

where the following identities were used for any unit vector $\mathbf{r}$:

$$[\mathbf{r}\times]^3 = -[\mathbf{r}\times] \tag{18a}$$

$$[\mathbf{r}\times]^4 = -[\mathbf{r}\times]^2 \tag{18b}$$

Therefore, since Rank($\mathcal{N}$) $= 6$ then Rank(M), and ultimately the rank of F, is given by the Rank($-[\mathbf{r}\times]^2$). The matrix $-[\mathbf{r}\times]^2 = I_{3\times3} - \mathbf{r}\,\mathbf{r}^T$ is the projection matrix onto the space perpendicular to $\mathbf{r}$ and has rank 2. This indicates that only 2 pieces of information are given using one vector observation.

The eigenvalues of M are given by solving the following equation:

$$\det(\lambda I_{3\times3} - M) = \det \begin{bmatrix} \lambda I_{3\times3} + [\mathbf{r}\times]^2 & -\zeta[\mathbf{r}\times] \\ -\zeta[\mathbf{r}\times]^T & \lambda I_{3\times3} + \zeta^2[\mathbf{r}\times]^2 \end{bmatrix} = 0 \tag{19}$$

Performing the matrix determinant operation gives

$$\begin{aligned} \det(\lambda I_{3\times3} - M) &= \det\{(\lambda I_{3\times3} + [\mathbf{r}\times]^2)(\lambda I_{3\times3} + \zeta^2\,[\mathbf{r}\times]^2) + \zeta^2\,[\mathbf{r}\times]^2\} \\ &= \det\{\lambda^2 I_{3\times3} + \lambda(1+\zeta^2)\,[\mathbf{r}\times]^2 + \zeta^2\,[\mathbf{r}\times]^4 + \zeta^2\,[\mathbf{r}\times]^2\} \end{aligned} \tag{20}$$

Next, using the identity in Eq. (18b) yields

$$\det(\lambda I_{3\times3} - M) = \lambda^3 \det\{\lambda I_{3\times3} + (1+\zeta^2)[\mathbf{r}\times]^2\} \tag{21}$$

Clearly three eigenvalues of M are zero. The eigenvalues of $-(1+\zeta^2)[\mathbf{r}\times]^2$ are well known, which are given by 0 and twice repeated $(1+\zeta^2)$. Therefore, the eigenvalues of $\mathcal{F} = \sigma^{-2}M$ are given by

$$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0,\ \lambda_5 = \lambda_6 = \sigma^{-2}(1+\zeta^2) \tag{22}$$

Since the eigenvalues of a matrix are unaffected by a similarity transformation, Eq. (22) also gives the eigenvalues of F.

In order to calculate the eigenvectors of F we first state a well-known property of a symmetric matrix. Let Υ be an $n \times n$ symmetric matrix. There exists an orthogonal matrix Z such that $Z^T \Upsilon Z = D$, where D is a diagonal matrix with the characteristic roots of Υ. This also states that a symmetric matrix is similar to a diagonal matrix.[10] Note that if some eigenvalue has multiple-fold degeneracy (as in the present case), one can find an orthogonal basis in the subspace spanned by its eigenvectors. Therefore, $\mathcal{F} = W \mathrm{diag}\,[\lambda_1\ \lambda_2\ \lambda_3\ \lambda_4\ \lambda_5\ \lambda_6]\, W^T$, where $W = [\mathbf{w}_1\ \mathbf{w}_2\ \mathbf{w}_3\ \mathbf{w}_4\ \mathbf{w}_5\ \mathbf{w}_6]$ is an orthogonal matrix, and $\mathbf{w}_i$, $i = 1,\ 2,\ \ldots,\ 6$, are 6×1 orthogonal eigenvectors. We now calculate the eigenvectors $\mathbf{w}_5$ and $\mathbf{w}_6$, which correspond to the eigenvalues λ_5 and λ_6 in Eq. (22), respectively. For the eigenvalue λ_5 we have

$$\mathcal{F}\mathbf{w}_5 = \lambda_5 \mathbf{w}_5 = \sigma^{-2} \left(1 + \zeta^2\right) \mathbf{w}_5 \tag{23}$$

From Eq. (15), $\mathcal{F} = \sigma^{-2} M$; hence, $M\mathbf{w}_5 = (1 + \zeta^2)\mathbf{w}_5$. Let $\mathbf{w}_5$ and $\mathbf{w}_6$ be partitioned into

$$\mathbf{w}_5 \equiv \begin{bmatrix} \mathbf{w}_{51} \\ \mathbf{w}_{52} \end{bmatrix}, \quad \mathbf{w}_6 \equiv \begin{bmatrix} \mathbf{w}_{61} \\ \mathbf{w}_{62} \end{bmatrix} \tag{24}$$

where $\mathbf{w}_{51}$, $\mathbf{w}_{52}$, $\mathbf{w}_{61}$ and $\mathbf{w}_{62}$ are 3×1 partition vectors of $\mathbf{w}_5$ and $\mathbf{w}_6$, respectively. From the definition of M in Eq. (15) and using the partitioned eigenvector in Eq. (24), the following two equations are given:

$$-[\mathbf{r}\times]^2 \mathbf{w}_{51} + \zeta[\mathbf{r}\times]\mathbf{w}_{52} = (1 + \zeta^2)\mathbf{w}_{51} \tag{25a}$$

$$-\zeta[\mathbf{r}\times]\mathbf{w}_{51} - \zeta^2[\mathbf{r}\times]^2 \mathbf{w}_{52} = (1 + \zeta^2)\mathbf{w}_{52} \tag{25b}$$

Simultaneously solving Eqs. (25a) and (25b) gives

$$\mathbf{w}_{51} \perp \mathbf{r} \tag{26a}$$

$$\mathbf{w}_{52} = -\zeta[\mathbf{r}\times]\mathbf{w}_{51} \tag{26b}$$

This states that both $\mathbf{w}_{51}$ and $\mathbf{w}_{52}$ lie in the plane perpendicular to $\mathbf{r}$. Also, clearly $\mathbf{w}_{51} \perp \mathbf{w}_{52}$, which means that the vectors $\mathbf{w}_{51}$, $\mathbf{w}_{52}$ and $\mathbf{r}$ form an orthogonal set.

In order to determine the eigenvector $\mathbf{w}_5$, the vectors $\mathbf{r}$ and $\mathbf{w}_{51}$ are first given in component form by $\mathbf{r} = [r_1\ r_2\ r_3]^T$ and $\mathbf{w}_{51} = [w_{51}\ w_{52}\ w_{53}]^T$, respectively. At least one component of $\mathbf{r}$ must be nonzero. We assume that $r_1 \neq 0$, but the argument goes through with only minor modification for any nonzero component. Since $\mathbf{w}_{51}^T\mathbf{r} = 0$, and assuming $r_1 \neq 0$, then $w_{51} = -(w_{52}r_2 + w_{53}r_3)/r_1$. Next, without loss in generality we can assume that $w_{52} = 1$ and $w_{53} = 0$, so $\mathbf{w}_{51} = [-r_2/r_1\ 1\ 0]^T$. Therefore, using Eq. (26b) the normalized vector for $\mathbf{w}_5$ is given by

$$\mathbf{w}_5 = \mathbf{a}/||\mathbf{a}|| \tag{27}$$

where

$$\mathbf{a} \equiv \begin{bmatrix} -r_2/r_1 & 1 & 0 & \zeta r_3 & \zeta r_2 r_3/r_1 & -\zeta(r_2^2/r_1 + r_1) \end{bmatrix}^T \tag{28}$$

In a similar fashion, using $\mathbf{w}_5^T\mathbf{w}_6 = 0$ the normalized vector for $\mathbf{w}_6$ is given by

$$\mathbf{w}_6 = \mathbf{b}/||\mathbf{b}|| \tag{29}$$

where

$$\mathbf{b} \equiv \begin{bmatrix} r_1^2 r_3 & r_1 r_2 r_3 & -r_1(r_1^2 + r_2^2) & \zeta r_1 r_2 & -\zeta r_1^2 & 0 \end{bmatrix}^T \tag{30}$$

If $r_1 = 0$ then other eigenvectors can be found by using the non-zero component values of $\mathbf{r}$.

The next step involves determining the eigenvectors of F, which is decomposed as $F = V\text{diag}\,[\lambda_1\ \lambda_2\ \lambda_3\ \lambda_4\ \lambda_5\ \lambda_6]\,V^T$, where $V = [\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3\ \mathbf{v}_4\ \mathbf{v}_5\ \mathbf{v}_6]$ is an orthogonal matrix, and

$\mathbf{v}_i$, $i = 1, 2, \ldots, 6$, are 6×1 orthogonal eigenvectors. Using $F = \mathcal{M}\mathcal{F}\mathcal{M}^T$, where $\mathcal{M}$ is defined by Eq. (13), the eigenvectors $\mathbf{v}_5$ and $\mathbf{v}_6$ associated with the eigenvalues λ_5 and λ_6, respectively, are given by

$$\mathbf{v}_5 \equiv \begin{bmatrix} \mathbf{v}_{51} \\ \mathbf{v}_{52} \end{bmatrix} = \mathcal{M}\,\mathbf{w}_5 = \begin{bmatrix} A\,\mathbf{w}_{51} \\ \mathbf{w}_{52} \end{bmatrix} \tag{31a}$$

$$\mathbf{v}_6 \equiv \begin{bmatrix} \mathbf{v}_{61} \\ \mathbf{v}_{62} \end{bmatrix} = \mathcal{M}\,\mathbf{w}_6 = \begin{bmatrix} A\,\mathbf{w}_{61} \\ \mathbf{w}_{62} \end{bmatrix} \tag{31b}$$

The vectors $\mathbf{v}_5$ and $\mathbf{v}_6$ give information of the observable components for attitude and position. Each vector is equally observable since the eigenvalues are repeated. Position and attitude information cannot be decoupled since $||\mathbf{r}|| = 1 \neq 0$, which means that with one observation no useful information can be provided. This is in sharp contrast to standard attitude determination results using one vector observation, in which one vector observation provides 2-axis attitude information.[11] The analysis in this section also indicates that, for the multiple-beacon case, as the angular separation of the beacons decreases (approaching co-location) the physical meaning of the attitude and position results becomes skewed.

Two Vector Observation Case

In this section the two vector case is analyzed. We assume that the two vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ are non-colinear. Unlike the one vector case, the two vector case does provide some physical insights that are useful for beacon location studies. For the case the matrix $\mathcal{F}$ in Eq. (14) is given by

$$\mathcal{F} = \sum_{i=1}^{2} \mathcal{F}_i \tag{32}$$

where $\mathcal{F}_i$ is given by

$$\mathcal{F}_i = \sigma_i^{-2} L_i L_i^T \tag{33}$$

with

$$L_i \equiv \begin{bmatrix} -[\mathbf{r}_i\times] \\ \\ \zeta_i[\mathbf{r}_i\times]^2 \end{bmatrix} \tag{34}$$

Re-arranging the partitioned elements of L_i yields

$$\mathcal{F} = \sigma_1^{-2} L_1 L_1^T + \sigma_2^{-2} L_2 L_2^T \equiv L\, L^T \tag{35}$$

with

$$L \equiv \begin{bmatrix} -\sigma_1^{-1}[\mathbf{r}_1\times] & -\sigma_2^{-1}[\mathbf{r}_2\times] \\ \\ \sigma_1^{-1}\zeta_1[\mathbf{r}_1\times]^2 & \sigma_2^{-1}\zeta_2[\mathbf{r}_2\times]^2 \end{bmatrix} \tag{36}$$

where the identities in Eq. (18) were used in the above quantities. Clearly, we now have $\text{Rank}(\mathcal{F}) = \text{Rank}(L)$.

We now discuss the rank of the matrix L. Reference [12] shows that the rank of a $q \times n$ matrix C $(q \geq n)$ is $n - m$, where m is the maximum number of orthogonal vectors $\mathbf{y}$ that satisfy $C\mathbf{y} = \mathbf{0}$. For the two vector case consider the conditions for $L^T\mathbf{y} = \mathbf{0}$, with $\mathbf{y} \neq \mathbf{0}$, to be satisfied. Using the partitioned elements of L yields

$$[\mathbf{r}_1\times]\mathbf{y}_1 + \zeta_1[\mathbf{r}_1\times]^2\mathbf{y}_2 = \mathbf{0} \tag{37a}$$

$$[\mathbf{r}_2\times]\mathbf{y}_1 + \zeta_2[\mathbf{r}_2\times]^2\mathbf{y}_2 = \mathbf{0} \tag{37b}$$

where $\mathbf{y} \equiv \begin{bmatrix} \mathbf{y}_1^T & \mathbf{y}_2^T \end{bmatrix}^T$. The general relations for $\mathbf{y}_1$ and $\mathbf{y}_2$ that satisfy Eq. (37) are given by

$$\mathbf{y}_1 = -\zeta_1[\mathbf{r}_1\times]\mathbf{y}_2 + c_1\mathbf{r_1} \tag{38a}$$

$$\mathbf{y}_1 = -\zeta_2[\mathbf{r}_2\times]\mathbf{y}_2 + c_2\mathbf{r_2} \tag{38b}$$

where c_1 and c_2 are arbitrary constants. Subtracting (38a) from (38b) gives

$$\left(\zeta_1[\mathbf{r}_1\times] - \zeta_2[\mathbf{r}_2\times]\right)\mathbf{y}_2 = c_1\mathbf{r}_1 - c_2\mathbf{r}_2 \tag{39}$$

We first consider the case where $c_1\mathbf{r}_1 - c_2\mathbf{r}_2 = \mathbf{0}$. Since it is assumed that $\mathbf{r}_1$ and $\mathbf{r}_2$ are non-colinear then $c_1 = c_2 = 0$, so we have

$$\mathbf{y}_2 = \pm(\zeta_1\mathbf{r}_1 - \zeta_2\mathbf{r}_2) \tag{40}$$

Therefore, the vector $\mathbf{y}_2$ is contained in the plane given by $\mathbf{r}_1$ and $\mathbf{r}_2$. Next, we consider the case where $c_1\mathbf{r}_1 - c_2\mathbf{r}_2 \neq \mathbf{0}$. From Eq. (39) the quantity $c_1\mathbf{r}_1 - c_2\mathbf{r}_2$ must be perpendicular to both $\mathbf{y}_2$ and to $(\zeta_1\mathbf{r}_1 - \zeta_2\mathbf{r}_2)$. Therefore, another solution for $\mathbf{y}_2$, denoted by $\mathbf{y}_2'$, is given by

$$\mathbf{y}_2' = -\left(\zeta_1[\mathbf{r}_1\times] - \zeta_2[\mathbf{r}_2\times]\right)(c_1\mathbf{r}_1 - c_2\mathbf{r}_2) \tag{41}$$

Note that $\mathbf{y}_2$ and $\mathbf{y}_2'$ are orthogonal vectors. Equation (38) can be used to find $\mathbf{y}_1$ and $\mathbf{y}_1'$. Also, $\mathbf{y}$ and $\mathbf{y}'$, where $\mathbf{y}' \equiv \begin{bmatrix}\mathbf{y'}_1^T & \mathbf{y'}_2^T\end{bmatrix}^T$, are orthogonal vectors. Therefore, the maximum number of orthogonal vectors $\mathbf{y}$ that satisfy $L^T\mathbf{y} = \mathbf{0}$ is 2. Hence, $\text{Rank}(L) = \text{Rank}(F) = 6 - 2 = 4$. Therefore, four quantities are observable using two vector observations.

Reference [3] shows that out of these four observable quantities one axis of attitude and one axis of position information can be determined (the remaining two quantities must be a combination of attitude and position). This states that two out of the four observable eigenvectors of the matrix F can be decoupled in attitude and position. The results are summarized here for completeness. We first partition the information matrix F into 3×3 sub-matrices as

$$F = P^{-1} = \begin{bmatrix} F_{11} & F_{12} \\ F_{12}^T & F_{22} \end{bmatrix}, \quad P = \begin{bmatrix} \mathcal{P}_{11}^{-1} & \mathcal{P}_{12}^{-1} \\ \mathcal{P}_{12}^{-T} & \mathcal{P}_{22}^{-1} \end{bmatrix} \tag{42}$$

with obvious definitions for F_{11}, F_{12} and F_{22} from Eq. (10). The relationships between $\mathcal{P}_{11}$,

$\mathcal{P}_{12}$, $\mathcal{P}_{22}$ and F_{11}, F_{12}, F_{22} are given by[13]

$$\mathcal{P}_{11} = (F_{11} - F_{12}F_{22}^{-1}F_{12}^T) \tag{43a}$$

$$\mathcal{P}_{12} = F_{11}^{-1}F_{12}(F_{12}^T F_{11}^{-1}F_{12} - F_{22}) \tag{43b}$$

$$\mathcal{P}_{22} = (F_{22} - F_{12}^T F_{11}^{-1}F_{12}) \tag{43c}$$

The matrix $\mathcal{P}_{11}$ corresponds to the attitude information, and the matrix $\mathcal{P}_{22}$ corresponds to the position information. The matrix $\mathcal{P}_{11}$ can be shown to be given by

$$\mathcal{P}_{11} = A\mathcal{G}A^T \tag{44}$$

where

$$\mathcal{G} = \frac{1}{\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2}\,\mathbf{g}\,\mathbf{g}^T \tag{45}$$

with

$$\mathbf{g} = \begin{bmatrix} \pm\left\{(\rho_2^2 + \rho_3^2) - \|\boldsymbol{\rho}\|^2\gamma_1^2/\|\boldsymbol{\gamma}\|^2\right\}^{1/2} \\ \pm\left\{(\rho_1^2 + \rho_3^2) - \|\boldsymbol{\rho}\|^2\gamma_2^2/\|\boldsymbol{\gamma}\|^2\right\}^{1/2} \\ \pm\left\{(\rho_1^2 + \rho_2^2) - \|\boldsymbol{\rho}\|^2\gamma_3^2/\|\boldsymbol{\gamma}\|^2\right\}^{1/2} \end{bmatrix} \tag{46}$$

and

$$\boldsymbol{\rho} = \boldsymbol{\beta}_1 - \boldsymbol{\beta}_2 \tag{47a}$$

$$\boldsymbol{\gamma} = \boldsymbol{\beta}_1 \times \boldsymbol{\beta}_2 \tag{47b}$$

$$\boldsymbol{\beta}_i \equiv \begin{bmatrix} X_i - X_c \\ Y_i - Y_c \\ Z_i - Z_c \end{bmatrix}, \quad i = 1,\, 2 \tag{47c}$$

$$\tilde{\sigma}_i^2 \equiv \left[(X_i - X_c)^2 + (Y_i - Y_c)^2 + (Z_i - Z_c)^2\right]\sigma_i^2, \quad i = 1,\, 2 \tag{47d}$$

An eigenvalue/eigenvector decomposition of Eq. (44) can be used to assess the observability. The eigenvalues of Eq. (44) are given by $(0,\, 0,\, [\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2]^{-1}\|\boldsymbol{\rho}\|^2)$, and the eigenvector associated with the non-zero eigenvalue is given by $\mathbf{v} = A\mathbf{g}/\|\mathbf{g}\|$, which defines the axis of rotation for the observable attitude angle. The eigenvector can easily be shown to lie in the plane of the two body vector observations since $\mathbf{v}^T A(\boldsymbol{\beta}_1 \times \boldsymbol{\beta}_2) = 0$. This vector is in essence a weighted average of the body observations with

$$\|A\boldsymbol{\beta}_1\| \cos a_1 = \|A\boldsymbol{\beta}_2\| \cos a_2 \tag{48}$$

where a_1 is the angle between $A\boldsymbol{\beta}_1$ and $\mathbf{v}$, and a_2 is the angle between $A\boldsymbol{\beta}_2$ and $\mathbf{v}$, as shown in Fig. 2 ($a_1 + a_2$ is the angle between $A\boldsymbol{\beta}_1$ and $A\boldsymbol{\beta}_2$). Equation (48) indicates that the observable axis of rotation is closer to the vector with less length.

In a similar fashion, the position information matrix can be shown to be given by

$$\mathcal{P}_{22} = \frac{1}{\sigma_1^2 + \sigma_2^2}\, \mathbf{h}\, \mathbf{h}^T \tag{49}$$

with

$$\mathbf{h} = \begin{bmatrix} \pm \left\{ (\varrho_2^2 + \varrho_3^2) - \|\boldsymbol{\varrho}\|^2 \vartheta_1^2 / \|\boldsymbol{\vartheta}\|^2 \right\}^{1/2} \\ \pm \left\{ (\varrho_1^2 + \varrho_3^2) - \|\boldsymbol{\varrho}\|^2 \vartheta_2^2 / \|\boldsymbol{\vartheta}\|^2 \right\}^{1/2} \\ \pm \left\{ (\varrho_1^2 + \varrho_2^2) - \|\boldsymbol{\varrho}\|^2 \vartheta_3^2 / \|\boldsymbol{\vartheta}\|^2 \right\}^{1/2} \end{bmatrix} \tag{50}$$

and

$$\boldsymbol{\varrho} = \boldsymbol{\delta}_1 - \boldsymbol{\delta}_2 \tag{51a}$$

$$\boldsymbol{\vartheta} = \boldsymbol{\delta}_1 \times \boldsymbol{\delta}_2 \tag{51b}$$

$$\boldsymbol{\delta}_i = \boldsymbol{\beta}_i / \|\boldsymbol{\beta}_i\|^2, \quad i = 1,\, 2 \tag{51c}$$

The eigenvalues of Eq. (49) are given by $(0,\, 0,\, [\sigma_1^2 + \sigma_2^2]^{-1}\|\boldsymbol{\varrho}\|^2)$, and the eigenvector associated with the non-zero eigenvalue is given by $\mathbf{w} = \mathbf{h}/\|\mathbf{h}\|$, which defines the observable position axis. The eigenvector can be shown to lie in the plane of the two reference vectors since $\mathbf{w}^T(\boldsymbol{\beta}_1 \times \boldsymbol{\beta}_2) = 0$. The weighted average relationship for the observable position axis is given by

$$\|\boldsymbol{\beta}_1\|/\cos\alpha_1 = \|\boldsymbol{\beta}_2\|/\cos\alpha_2 \tag{52}$$

where α_1 is the angle between $\boldsymbol{\beta}_1$ and $\mathbf{w}$, and α_2 is the angle between $\boldsymbol{\beta}_2$ and $\mathbf{w}$ ($\alpha_1+\alpha_2$ is the angle between $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$). Equation (52) indicates that the observable position axis is closer to the vector with greater length, which intuitively makes sense because the position solution is more sensitive to the magnitude of the vectors. A slight change in the largest vector produces more change in the position than the same change in the smallest vector. Also, if $\|\boldsymbol{\beta}_1\| = \|\boldsymbol{\beta}_2\|$ or if $\boldsymbol{\beta}_1^T\boldsymbol{\beta}_2 = 0$, then the eigenvector reduces to $\mathbf{w} = \pm(\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2)/\|\boldsymbol{\beta}_1 + \boldsymbol{\beta}_2\|$, which is the bisector of the reference vectors. As before, the information given by the two observation vectors is used to calculate the part of the attitude needed to compute the observable position.

Comparing Eq. (48) to Eq. (52) indicates that the beacon that is closest to the target provides the most attitude information, but has the least position information (this is due to the inverse relationship between them). The converse is true as well, i.e., the beacon that is farthest from the target provides the most position information, but has the least attitude information (see Ref. [3] for more details). The covariance analysis can be useful to trade off the relative importance between attitude and position requirements with two vector observations.

Three Vector Observation Case

In this section the three vector case is analyzed. We assume that any two of the vectors $\mathbf{r}_1$, $\mathbf{r}_2$ or $\mathbf{r}_3$ are non-colinear. We will show that the covariance matrix in this case is full rank for most cases. In the three vector case the matrix $\mathcal{F}$ from Eq. (14) is given by

$$\mathcal{F} = L\,L^T \tag{53}$$

with

$$L \equiv \begin{bmatrix} -\sigma_1^{-1}[\mathbf{r}_1\times] & -\sigma_2^{-1}[\mathbf{r}_2\times] & -\sigma_3^{-1}[\mathbf{r}_3\times] \\ \sigma_1^{-1}\zeta_1[\mathbf{r}_1\times]^2 & \sigma_2^{-1}\zeta_2[\mathbf{r}_2\times]^2 & \sigma_3^{-1}\zeta_3[\mathbf{r}_3\times]^2 \end{bmatrix} \tag{54}$$

As before the rank of $\mathcal{F}$, and ultimately the rank of F, can be determined by considering the conditions for $L^T\mathbf{y} = \mathbf{0}$, with $\mathbf{y} \neq \mathbf{0}$, to be satisfied. Using the partitioned elements of L yields

$$[\mathbf{r}_1\times]\mathbf{y}_1 + \zeta_1[\mathbf{r}_1\times]^2\mathbf{y}_2 = \mathbf{0} \tag{55a}$$

$$[\mathbf{r}_2\times]\mathbf{y}_1 + \zeta_2[\mathbf{r}_2\times]^2\mathbf{y}_2 = \mathbf{0} \tag{55b}$$

$$[\mathbf{r}_3\times]\mathbf{y}_1 + \zeta_3[\mathbf{r}_3\times]^2\mathbf{y}_2 = \mathbf{0} \tag{55c}$$

where $\mathbf{y} \equiv \begin{bmatrix} \mathbf{y}_1^T & \mathbf{y}_2^T \end{bmatrix}^T$. The general relations for $\mathbf{y}_1$ and $\mathbf{y}_2$ that satisfy Eq. (55) are given by

$$\mathbf{y}_1 = -\zeta_1[\mathbf{r}_1\times]\mathbf{y}_2 + c_1\mathbf{r_1} \tag{56a}$$

$$\mathbf{y}_1 = -\zeta_2[\mathbf{r}_2\times]\mathbf{y}_2 + c_2\mathbf{r_2} \tag{56b}$$

$$\mathbf{y}_1 = -\zeta_3[\mathbf{r}_3\times]\mathbf{y}_2 + c_3\mathbf{r_3} \tag{56c}$$

where c_1, c_2 and c_3 are arbitrary constants. Eq. (56) can be written in matrix form as

$$D\mathbf{y} = \mathbf{z} \tag{57}$$

where

$$D \equiv \begin{bmatrix} I_{3\times3} & \zeta_1[\mathbf{r}_1\times] \\ I_{3\times3} & \zeta_2[\mathbf{r}_2\times] \\ I_{3\times3} & \zeta_3[\mathbf{r}_3\times] \end{bmatrix} \tag{58a}$$

$$\mathbf{z} = \begin{bmatrix} c_1\mathbf{r_1} \\ c_2\mathbf{r_2} \\ c_3\mathbf{r_3} \end{bmatrix} \tag{58b}$$

A solution to Eq. (57) exists if and only if the rank of the coefficient matrix D is equal to the rank of the augmented matrix $[D \ \ \mathbf{z}]$.[14] From this theorem the following scenarios are possible:

1. If $\text{Rank}(D) = 6$ and $\text{Rank}(D) = \text{Rank}\left([D \ \ \mathbf{z}]\right)$, where $\mathbf{z} \neq \mathbf{0}$, then a solution to Eq. (57) exists, and a nonzero $\mathbf{y}$ can be found such that $L^T\mathbf{y} = \mathbf{0}$. Therefore, $\text{Rank}(L) = \text{Rank}(F) < 6$.

2. If $\text{Rank}(D) = 6$ and $\text{Rank}(D) \neq \text{Rank}\left([D \ \ \mathbf{z}]\right)$, where $\mathbf{z} \neq \mathbf{0}$, then a solution to Eq. (57) cannot be determined, and a nonzero $\mathbf{y}$ cannot be found such that $L^T\mathbf{y} = \mathbf{0}$. Therefore, $\text{Rank}(L) = \text{Rank}(F) = 6$.

3. If $\text{Rank}(D) < 6$, then certainly a nonzero $\mathbf{y}$ exists such that Eq. (57) is satisfied, and $\text{Rank}(L) = \text{Rank}(F) < 6$.

We now discuss the properties of the matrix $[D \ \ \mathbf{z}]$. Through elementary row operations this matrix can be shown to be similar to

$$\begin{bmatrix} D & \mathbf{z} \end{bmatrix} \sim \begin{bmatrix} I_{3\times3} & \zeta_1[\mathbf{r}_1\times] & c_1\mathbf{r}_1 \\ 0_{3\times3} & [\mathbf{u}_1\times] & \boldsymbol{\eta}_1 \\ 0_{3\times3} & [\mathbf{u}_2\times] & \boldsymbol{\eta}_2 \end{bmatrix} \tag{59}$$

where

$$\mathbf{u}_1 \equiv \zeta_2\mathbf{r}_2 - \zeta_1\mathbf{r}_1 \tag{60a}$$

$$\mathbf{u}_2 \equiv \zeta_3\mathbf{r}_3 - \zeta_1\mathbf{r}_1 \tag{60b}$$

$$\boldsymbol{\eta}_1 \equiv c_2\mathbf{r}_2 - c_1\mathbf{r}_1 \tag{60c}$$

$$\boldsymbol{\eta}_2 \equiv c_3\mathbf{r}_3 - c_1\mathbf{r}_1 \tag{60d}$$

Define the lower partition of matrix in Eq. (59) by

$$Q \equiv \begin{bmatrix} [\mathbf{u}_1\times] & \boldsymbol{\eta}_1 \\ [\mathbf{u}_2\times] & \boldsymbol{\eta}_2 \end{bmatrix} \tag{61}$$

Also, let the following vectors be given in their components as $\mathbf{u}_1 = [u_{11}\ u_{12}\ u_{13}]^T$, $\mathbf{u}_2 = [u_{21}\ u_{22}\ u_{23}]^T$, $\boldsymbol{\eta}_1 = [\eta_{11}\ \eta_{12}\ \eta_{13}]^T$ and $\boldsymbol{\eta}_2 = [\eta_{21}\ \eta_{22}\ \eta_{23}]^T$. Assuming $u_{13} \neq 0$, $u_{23} \neq 0$ and $u_{11}u_{23} - u_{13}u_{21} \neq 0$ (if these conditions are not true then other nonzero elements can be used, which is discussed later), the matrix Q can be shown to be similar to

$$Q \sim \begin{bmatrix} V & \boldsymbol{\varpi}_1 \\ 0_{3\times 3} & \boldsymbol{\varpi}_2 \end{bmatrix} \tag{62}$$

where

$$V \equiv \begin{bmatrix} 0 & -u_{13} & u_{12} \\ u_{13} & 0 & -u_{11} \\ 0 & 0 & -u_{21} + u_{11}u_{23}/u_{13} \end{bmatrix} \tag{63a}$$

$$\boldsymbol{\varpi}_1 \equiv \begin{bmatrix} \eta_{11} \\ \eta_{12} \\ \eta_{22} - u_{23}\eta_{12}/u_{13} \end{bmatrix} \tag{63b}$$

$$\boldsymbol{\varpi}_2 \equiv \begin{bmatrix} \eta_{21} - u_{23}\eta_{11}/u_{13} + \dfrac{u_{12}u_{23} - u_{13}u_{22}}{u_{11}u_{23} - u_{13}u_{21}}(\eta_{22} - u_{23}\eta_{12}/u_{13}) \\ \eta_{13} + u_{12}\eta_{12}/u_{13} + u_{11}\eta_{11}/u_{13} \\ \eta_{23} + u_{22}\eta_{22}/u_{23} + u_{21}\eta_{21}/u_{23} \end{bmatrix} \tag{63c}$$

If $u_{13} \neq 0$ and $u_{11}u_{23} - u_{13}u_{21} \neq 0$, then $\text{Rank}(V) = 3$. If $\boldsymbol{\varpi}_2 = \mathbf{0}$, then $\text{Rank}(D) = \text{Rank}\left([D \;\; \mathbf{z}]\right)$. Therefore, if a set of nonzero c_1, c_2 and c_3 can be found such that $\boldsymbol{\varpi}_2 = \mathbf{0}$, then a nonzero $\mathbf{y}$ can be found such that $L^T\mathbf{y} = \mathbf{0}$ is true, so $\text{Rank}(L) = \text{Rank}(F) < 6$, which states that full observability in attitude and position is not possible. Also, if a set of nonzero c_1, c_2 and c_3 cannot be found such that $\boldsymbol{\varpi}_2 = \mathbf{0}$, then a nonzero $\mathbf{y}$ cannot be found such that $L^T\mathbf{y} = \mathbf{0}$ is true, so $\text{Rank}(L) = \text{Rank}(F) = 6$, which states that full observability in attitude and position is possible.

The condition $\boldsymbol{\varpi}_2 = \mathbf{0}$ can be restated as:

$$E\mathbf{c} = \mathbf{0} \tag{64}$$

where $\mathbf{c} = [c_1 \; c_2 \; c_3]^T$ and

$$E = \begin{bmatrix} -\mathbf{r}_1^T\mathbf{u}_1 & \mathbf{r}_2^T\mathbf{u}_1 & 0 \\ -\mathbf{r}_1^T\mathbf{u}_2 & 0 & \mathbf{r}_3^T\mathbf{u}_2 \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \tag{65}$$

The quantities e_{31}, e_{32} and e_{33} are given by

$$e_{31} = (u_{23} - u_{13})[(u_{11}u_{23} - u_{13}u_{21})r_{11} + (u_{12}u_{23} - u_{13}u_{22})r_{12}] \tag{66a}$$

$$e_{32} = -u_{23}[(u_{11}u_{23} - u_{13}u_{21})r_{21} + (u_{12}u_{23} - u_{13}u_{22})r_{22}] \tag{66b}$$

$$e_{33} = u_{13}[(u_{11}u_{23} - u_{13}u_{21})r_{31} + (u_{12}u_{23} - u_{13}u_{22})r_{32}] \tag{66c}$$

where $\mathbf{r}_1 = [r_{11}\ r_{12}\ r_{13}]^T$, $\mathbf{r}_2 = [r_{21}\ r_{22}\ r_{23}]^T$ and $\mathbf{r}_3 = [r_{31}\ r_{32}\ r_{33}]^T$. If $\mathbf{r}_2^T\mathbf{u}_1 \neq 0$ and $\mathbf{r}_3^T\mathbf{u}_2 \neq 0$, the following similarity condition can be obtained through elementary row operations:

$$E \sim \begin{bmatrix} -\mathbf{r}_1^T\mathbf{u}_1 & \mathbf{r}_2^T\mathbf{u}_1 & 0 \\ -\mathbf{r}_1^T\mathbf{u}_2 & 0 & \mathbf{r}_3^T\mathbf{u}_2 \\ \chi & 0 & 0 \end{bmatrix} \tag{67}$$

where

$$\chi = e_{31} + \frac{\mathbf{r}_1^T\mathbf{u}_1}{\mathbf{r}_2^T\mathbf{u}_1}e_{32} + \frac{\mathbf{r}_1^T\mathbf{u}_2}{\mathbf{r}_3^T\mathbf{u}_2}e_{33} \tag{68}$$

If $\chi \neq 0$ then $\text{Rank}(E) = 3$, and Eq. (64) can only be satisfied when $\mathbf{c} = \mathbf{0}$. Hence, $\text{Rank}(L) = \text{Rank}(F) = 6$, which gives an observable system. After some algebraic manipulations χ can also be shown to be given by

$$\chi = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} [(\mathbf{u}_1 \times \mathbf{u}_2) \times \mathbf{v}] \tag{69}$$

where

$$\mathbf{v} \equiv (u_{23} - u_{13})\mathbf{r}_1 - \frac{\mathbf{r}_1^T\mathbf{u}_1}{\mathbf{r}_2^T\mathbf{u}_1}u_{23}\mathbf{r}_2 + \frac{\mathbf{r}_1^T\mathbf{u}_2}{\mathbf{r}_3^T\mathbf{u}_2}u_{13}\mathbf{r}_3 \tag{70}$$

Therefore, $\chi = 0$ when the third component of $[(\mathbf{u}_1 \times \mathbf{u}_2) \times \mathbf{v}]$ is zero, which occurs when $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{v}$ lie in the same plane that is perpendicular to the object space plane given by $Z = 0$ (see Fig. 1). Hence, $\text{Rank}(E) < 3$ and a nonzero $\mathbf{c}$ can be found that satisfies Eq. (64), which means that the system is not observable since $\text{Rank}(L) = \text{Rank}(F) < 6$. From the matrix E in Eq. (65) the following cases can easily be proved:

1. If $\mathbf{r}_3^T\mathbf{u}_2 = 0$ and $e_{33} = 0$, then Rank(F) < 6.

2. If $\mathbf{r}_2^T\mathbf{u}_1 = 0$ and $e_{32} = 0$, then Rank(F) < 6.

3. If $\mathbf{r}_1^T\mathbf{u}_1 = \mathbf{r}_2^T\mathbf{u}_1 = 0$, then Rank($F$) < 6.

4. If $\mathbf{r}_1^T\mathbf{u}_2 = \mathbf{r}_3^T\mathbf{u}_2 = 0$, then Rank($F$) < 6.

5. If $\mathbf{r}_2^T\mathbf{u}_1 = 0$ and $e_{32} \neq 0$, with $\mathbf{r}_3^T\mathbf{u}_2 \neq 0$ and $\mathbf{r}_1^T\mathbf{u}_1 \neq 0$, then Rank($F$) $= 6$.

6. If $\mathbf{r}_3^T\mathbf{u}_2 = 0$ and $e_{33} \neq 0$, with $\mathbf{r}_2^T\mathbf{u}_1 \neq 0$ and $\mathbf{r}_1^T\mathbf{u}_2 \neq 0$, then Rank($F$) $= 6$.

The first two conditions are physically interesting cases since they can be satisfied even if the vectors $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{r}_3$ are not coplanar (e.g., consider the vectors $\mathbf{r}_1 = [1\ \ -2\ 1]^T/\sqrt{6}$, $\mathbf{r}_2 = [2\ \ -2\ 1]^T/3$ and $\mathbf{r}_3 = [3\ \ -1\ 1]^T/\sqrt{11}$, which gives a rank deficient F). Also, the third and fourth cases occur only when the three vectors $\zeta_1\mathbf{r}_1$, $\zeta_2\mathbf{r}_2$ and $\zeta_3\mathbf{r}_3$ are parallel to each other with equal magnitude, which violates the assumption made in this section. An obvious rank deficient condition for Q in Eq. (61) exists when $\mathbf{u}_1 \times \mathbf{u}_2 = \mathbf{0}$, which occurs when $\mathbf{u}_1$ and $\mathbf{u}_2$ are parallel.

In the previous derivations it has been assumed that $u_{13} \neq 0$, $u_{23} \neq 0$ and $u_{11}u_{23} - u_{13}u_{21} \neq 0$. If these conditions are not true, then the other nonzero elements of $\mathbf{u}_1$ and $\mathbf{u}_2$ can be used to derive similar conditions for unobservability. This yields a condition of unobservability that occurs when the endpoints of the position vectors ($\zeta_1\mathbf{r}_1$, $\zeta_2\mathbf{r}_2$ and $\zeta_3\mathbf{r}_3$) can be connected by a straight line (e.g., consider the vectors $\mathbf{r}_1 = [1\ 2\ 1]^T/\sqrt{6}$, $\mathbf{r}_2 = [1\ 2\ 2]^T/3$ and $\mathbf{r}_3 = [1\ 2\ 3]^T/\sqrt{14}$, which gives a rank deficient F). We should also note that even though F can be shown to have full rank using three vector observations under most conditions, a unique attitude and position cannot be determined due to a sign ambiguity in the solution. This is difficult to prove analytically, but can be shown by simulation. This scenario is similar to attitude determination results using angle observations.[15]

More Than Three Observations

In the four vector case the matrix $\mathcal{F}$ from Eq. (14) is given by

$$\mathcal{F} = L\,L^T \tag{71}$$

with

$$L \equiv \begin{bmatrix} -\sigma_1^{-1}[\mathbf{r}_1\times] & -\sigma_2^{-1}[\mathbf{r}_2\times] & -\sigma_3^{-1}[\mathbf{r}_3\times] & -\sigma_4^{-1}[\mathbf{r}_4\times] \\ \sigma_1^{-1}\zeta_1[\mathbf{r}_1\times]^2 & \sigma_2^{-1}\zeta_2[\mathbf{r}_2\times]^2 & \sigma_3^{-1}\zeta_3[\mathbf{r}_3\times]^2 & \sigma_4^{-1}\zeta_4[\mathbf{r}_4\times]^2 \end{bmatrix} \tag{72}$$

As before the rank of $\mathcal{F}$, and ultimately the rank of F, can be determined by considering the conditions for $L^T\mathbf{y} = \mathbf{0}$, with $\mathbf{y} \neq \mathbf{0}$, to be satisfied. In a similar fashion as the three vector case the conditions for $L^T\mathbf{y} = \mathbf{0}$ can be written as

$$D\mathbf{y} = \mathbf{z} \tag{73}$$

where

$$D \equiv \begin{bmatrix} I_{3\times 3} & \zeta_1[\mathbf{r}_1\times] \\ I_{3\times 3} & \zeta_2[\mathbf{r}_2\times] \\ I_{3\times 3} & \zeta_3[\mathbf{r}_3\times] \\ I_{3\times 3} & \zeta_4[\mathbf{r}_4\times] \end{bmatrix} \tag{74a}$$

$$\mathbf{z} = \begin{bmatrix} c_1\mathbf{r_1} \\ c_2\mathbf{r_2} \\ c_3\mathbf{r_3} \\ c_4\mathbf{r_4} \end{bmatrix} \tag{74b}$$

A condition for an unobservable system can be derived using the same procedure as in the three vector case. Similar to the three vector case, the four vector case is unobservable when the endpoints of the position vectors can be connected by a straight line. These results are also valid when more than four LOS vectors are used. Furthermore, a unique solution for the attitude and position exists when four beacons are present and the system is observable.[3]

Trace and Eigenvalues of the Covariance Matrix

In this section the trace of the covariance matrix, given in Eq. (10), is analyzed. The trace of this matrix is useful to quantify the overall performance of the solution for the attitude and position (i.e., a lower trace provides a more overall accurate solution). The matrix $\mathcal{F}$ in Eq. (14) can be written as

$$\mathcal{F} = \sum_{i=1}^{N} \mathcal{F}_i \tag{75}$$

where

$$\mathcal{F}_i = \begin{bmatrix} -\sigma_i^{-2}[\mathbf{r}_i\times]^2 & \sigma_i^{-2}\zeta_i[\mathbf{r}_i\times] \\ \\ \sigma_i^{-2}\zeta_i[\mathbf{r}_i\times]^T & -\sigma_i^{-2}\zeta_i^{\,2}[\mathbf{r}_i\times]^2 \end{bmatrix} \tag{76}$$

The eigenvalues of $\mathcal{F}_i$ are given by

$$\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0,\ \lambda_5 = \lambda_6 = \sigma_i^{-2}(1+\zeta_i^2) \tag{77}$$

Then the trace of the information matrix F is given by

$$\mathrm{tr}\,(F) = \mathrm{tr}\,(\mathcal{F}) = 2\sum_{i=1}^{N} \sigma_i^{-2}(1+\zeta_i^2) \tag{78}$$

where the invariance of the trace through a similarity transformation is used.

We now discuss the properties of the matrix P. First a useful theorem is shown. Given two real $n \times n$ symmetric matrices, A and B, with A positive definite and B positive semi-definite, there exists a nonsingular T such that $A = T\,T^T$ and $B = T\,\Upsilon\,T^T$, where Υ is a diagonal matrix with elements given by $\Upsilon = \mathrm{diag}\,[\mu_1\ \mu_2\ \cdots\ \mu_n]$.[16] The matrix T can be derived using the following procedure. Since A is symmetric and positive definite a singular value decomposition can be performed so that $A = U\,Q^2 U^T$, where U is an orthogonal matrix and Q is a diagonal matrix of the square roots of the eigenvalues of A. Then, compute $C = Q^{-1}U^T B\, U\, Q^{-1}$. Since C is symmetric a singular value decomposition can be performed so that $C = V\,\Upsilon\,V^T$. Then $T = U\,Q\,V$. Let $\bar{\mathcal{F}}_3 \equiv \sum_{i=1}^{3} \mathcal{F}_i$ have full rank, and define $\bar{\mathcal{F}}_4 \equiv \sum_{i=1}^{4} \mathcal{F}_i = \bar{\mathcal{F}}_3 + \mathcal{F}_4$ which also has full rank. Also, $\bar{\mathcal{F}}_3$ and $\bar{\mathcal{F}}_4$ are positive definite matrices, and $\mathcal{F}_4$ is positive semi-definite. This theorem can be shown to easily prove that if $\mathcal{F}$ has full rank for three vector observations, then $\mathcal{F}$ has full rank for more than three observations. Now let $\bar{\mathcal{F}}_3 = T\,T^T = U\,Q^2 U^T$ and let $C = Q^{-1}U^T \mathcal{F}_4 U\, Q^{-1} = V\,\Upsilon\,V^T$. So $\mathcal{F}_4 = T\,\Upsilon\,T^T$. Therefore, $\bar{\mathcal{F}}_4$ is given by

$$\begin{aligned} \bar{\mathcal{F}}_4 = \bar{\mathcal{F}}_3 + \mathcal{F}_4 &= T\,T^T + T\,\Upsilon\,T^T \\ &= T\,\mathrm{diag}\left[(1+\mu_1)\ (1+\mu_2)\ \cdots\ (1+\mu_6)\right]\,T^T \end{aligned} \tag{79}$$

After some algebraic manipulations $\bar{\mathcal{F}}_4^{-1}$ can be shown to be given by

$$\bar{\mathcal{F}}_4^{-1} = \bar{\mathcal{F}}_3^{-1} - \Delta\bar{\mathcal{F}} \tag{80}$$

where $\Delta\bar{\mathcal{F}}$ is a positive semi-definite matrix given by

$$\Delta\bar{\mathcal{F}} = U\,Q^{-1} V\,\mathrm{diag}\left[\frac{\mu_1}{1+\mu_1}\ \frac{\mu_2}{1+\mu_2}\ \cdots\ \frac{\mu_6}{1+\mu_6}\right] V^T Q^{-1} U^T \tag{81}$$

Using the fact that the trace of the sum of two matrices is given by sum of the trace of each matrix individually, we have $\mathrm{tr}\left(\bar{\mathcal{F}}_4^{-1}\right) = \mathrm{tr}\left(\bar{\mathcal{F}}_3^{-1}\right) - \mathrm{tr}\left(\Delta\bar{\mathcal{F}}\right)$. Therefore, since $\mathrm{tr}\left(\Delta\bar{\mathcal{F}}\right) > 0$ then $\mathrm{tr}\left(\bar{\mathcal{F}}_4^{-1}\right) < \mathrm{tr}\left(\bar{\mathcal{F}}_3^{-1}\right)$. Since the trace is invariant under a similarity transformation, then the trace of the covariance matrix P in Eq. (10) with four vector observations is always less than the trace of the covariance using three vector observations, which intuitively makes sense. This result can be further expanded to multiple observations (i.e., the trace of the covariance using N observations is always less than the trace using any number of observations less than N).

We now discuss the properties of the eigenvalues of P. Consider the following decomposition: $\bar{\mathcal{F}}_3\mathbf{x}_i = \lambda_i\mathbf{x}_i$ and $\bar{\mathcal{F}}_4\mathbf{y}_i = \alpha_i\mathbf{y}_i$ ($i = 1, 2, \ldots, 6$), where λ_i is an eigenvalue of the matrix $\bar{\mathcal{F}}_3$, $\mathbf{x}_i$ is the eigenvector of the matrix $\bar{\mathcal{F}}_3$ corresponding to λ_i, α_i is an eigenvalue of the matrix $\bar{\mathcal{F}}_4$, and $\mathbf{y}_i$ is the eigenvector of the matrix $\bar{\mathcal{F}}_4$ corresponding with α_i. Since the eigenvectors of a symmetric matrix are orthogonal $\bar{\mathcal{F}}_3\mathbf{y}_i$ is related by

$$\bar{\mathcal{F}}_3\mathbf{y}_i = \sum_{j=1}^{6} k_{ij}\lambda_j\mathbf{x}_j \tag{82}$$

where the k_{ij} are constants with $\sum_{j=1}^{6} k_{ij}^2 = 1$. Also, $\mathcal{F}_4\mathbf{y}_i$ is given by

$$\mathcal{F}_4\mathbf{y}_i = \sum_{j=1}^{6}(\alpha_i - \lambda_j)k_{ij}\mathbf{x}_j \tag{83}$$

Since the eigenvectors of $\bar{\mathcal{F}}_3$ are orthogonal and since $\mathcal{F}_4$ is symmetric positive semi-definite, then

$$\sum_{j=1}^{6}(\alpha_i - \lambda_j)k_{ij}^2 = \alpha_i\sum_{j=1}^{6} k_{ij}^2 - \sum_{j=1}^{6} k_{ij}^2\lambda_j = \alpha_i - \sum_{j=1}^{6} k_{ij}^2\lambda_j \geq 0 \tag{84}$$

Therefore, the following condition is true:

$$\alpha_i \geq \sum_{j=1}^{6} k_{ij}^2\lambda_j \tag{85}$$

Let $\lambda_{\min} = \min\,[\lambda_1\ \lambda_2\ \cdots\ \lambda_6]$. Then from Eq. (85) $\alpha_i > \lambda_{\min}$. We know that $1/\lambda_i$ is an eigenvalue of both $\bar{\mathcal{F}}_3^{-1}$ and P using three observations, and $1/\alpha_i$ is an eigenvalue of both $\bar{\mathcal{F}}_4^{-1}$ and P using four observations. The eigenvalue analysis can be extended to the N vector observation case, and indicates that each eigenvalue of P using N observations is less than the maximum eigenvalue of the matrix with less than N observations. This proves that as the number of vector observations (N) increases more information is provided, which again intuitively makes sense.

Examples

Observability examples using representative geometric scenarios are shown in this section. We first consider the VISNAV system configuration, shown in Figure 1, with the following 3 beacon locations:

$$X_1 = 1\text{m}, \quad Y_1 = 2\text{m}, \quad Z_1 = 1\text{m}$$
$$X_2 = 1\text{m}, \quad Y_2 = 2\text{m}, \quad Z_2 = 2\text{m}$$
$$X_3 = 1\text{m}, \quad Y_3 = 2\text{m}, \quad Z_3 = 3\text{m}$$

The variances of the measurement error processes are assumed to be equal for each observation, which subsequently do not affect the observability analysis. Therefore all measurement error variances can be set to $\sigma_i^2 = 1$ for $i = 1, 2, 3$. Also, the focal length can be set to $f = 1$ without loss in generality. The true vehicle motion is given by $X_c = 30\exp[-(1/300)t]$ m, $Y_c = 30 - (30/1800)t$ m and $Z_c = 10 - (10/1800)t$ m. A $1,800$ sec simulation has been performed to generate the Fisher information matrix, i.e., the inverse of the covariance matrix in Eq. (10). A plot of the eigenvalues of the Fisher information at each time is shown in Figure 3. Two of the eigenvalues are nearly equal (the top line in the plot represents these eigenvalues). For this example the Fisher information matrix is clearly rank deficient. Thus, this configuration leads to an unobservable system. This is due to the fact that the endpoints of the position vectors are connected by a straight line, as previously discussed.

For the second example we consider the following 3 beacon locations:

$$X_1 = 0.5\text{m}, \quad Y_1 = 0.5\text{m}, \quad Z_1 = 0.0\text{m}$$
$$X_2 = 0.5\text{m}, \quad Y_2 = -0.5\text{m}, \quad Z_2 = 0.0\text{m}$$
$$X_3 = 0.2\text{m}, \quad Y_3 = 0.0\text{m}, \quad Z_3 = 0.1\text{m}$$

A plot of the eigenvalues of the Fisher information at each time is shown in Figure 4. Once again two of the eigenvalues are nearly equal (the top line in the plot represents these eigenvalues). For this example the Fisher information matrix is now full rank at all times. Thus, this configuration leads to an observable system. A measure of the performance in the estimation algorithm is given by the condition number (the ratio of the largest eigenvalue of the information matrix to the smallest eigenvalue). For this example the performance improves as the vehicle approaches the beacons, since they now more completely span the focal place area. However as the vehicle moves past the beacons the performance degrades, which is more clearly seen in Figure 3. This is directly related to the variances of the attitude and position estimation errors (see Ref. [3] for more details). These examples indicate that the analysis shown in this paper can help to understand and assess the observability of the estimation process when using LOS measurements to determine attitude and position.

Conclusions

An observability analysis for six degree of freedom state determination using vector observations was performed. The observability analysis proved that when one vector observation is used, two pieces of information can be inferred. However, the observable quantities involve a combination of position and attitude information, which cannot be decoupled. When two vector observations are used the rank of the covariance matrix is four. However, only one axis of attitude and one axis of position can be determined physically while the other two pieces of information involve coupled attitude/position information. When three or more vector observations are used the covariance matrix has full rank in most cases, and a unique solution for attitude and position exists for four or more vector observations. Finally, a trace and eigenvalue analysis of the covariance matrix indicated that as the number of vector observations increases, more accurate attitude and position information is provided in general.

Acknowledgments

This work was supported under a NASA grant (NCC 5-448), under the supervision of F. Landis Markley at NASA-Goddard Space Flight Center. The authors wish to thank Dr. Markley for many helpful suggestions and comments. The authors also wish to acknowledge the significant contributions of John L. Junkins and Declan Hughes at Texas A&M University, for the invention of the VISNAV sensor, and for many useful discussions on attitude and position determination using vector observations. Finally, the authors wish to thank Jongrae Kim, graduate student at Texas A&M University, for checking all equations and derivations in the paper.

References

[1]Junkins, J. L., Hughes, D. C., Wazni, K. P., and Pariyapong, V., "Vision-Based Navigation for Rendezvous, Docking and Proximity Operations," *22nd Annual AAS Guidance and Control Conference*, Breckenridge, CO, February 1999, AAS 99-021.

[2]Junkins, J. L., *An Introduction to Optimal Estimation of Dynamical Systems*, chap. 3, Sijthoff & Noordhoff International Publishers, Alphen aan den Rijn, The Netherlands, 1978.

[3]Crassidis, J. L., Alonso, R., and Junkins, J. L., "Optimal Attitude and Position Determination from Line-of-Sight Measurements," *The Journal of the Astronautical Sciences*, Vol. 48, No. 2 and 3, 2001, pp. 391–408.

[4]Wahba, G., "A Least-Squares Estimate of Satellite Attitude," *SIAM Review*, Vol. 7, No. 3, 1965, pp. 409.

[5]Shuster, M. D., "Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, 1981, pp. 70–77.

[6]Markley, F. L., "Attitude Determination from Vector Observations: A Fast Optimal Matrix Algorithm," *The Journal of the Astronautical Sciences*, Vol. 41, No. 2, 1993, pp. 261–280.

[7]Ju, G., Pollack, T., and Junkins, J. L., "DIGISTAR II Micro-Star Tracker: Autonomous

On-Orbit Calibration and Attitude Estimation," *AAS/AIAA Astrodynamics Specialist Conference*, Girdwood, Alaska, August 1999, AAS 99-431.

[8]Light, D. L., "Satellite Photogrammetry," *Manual of Photogrammetry*, edited by C. C. Slama, chap. 17, American Society of Photogrammetry, Falls Church, VA, 4th ed., 1980.

[9]Sorenson, H. W., *Parameter Estimation, Principles and Problems*, chap. 5, Marcel Dekker, Inc., New York, NY, 1980.

[10]Golub, G. H. and Van Loan, C. F., *Matrix Computations*, chap. 4, The Johns Hopkins University Press, Baltimore, MD, 2nd ed., 1989.

[11]Shuster, M. D., "Maximum Likelihood Estimation of Spacecraft Attitude," *The Journal of the Astronautical Sciences*, Vol. 37, No. 1, 1989, pp. 79–88.

[12]Hildebrand, F. B., *Methods of Applied Mathematics*, chap. 1, Dover Publications, New York, NY, 2nd ed., 1965.

[13]Horn, R. A. and Johnson, C. R., *Matrix Analysis*, Cambridge University Press, Cambridge, 1985, p. 18.

[14]Graybill, F. A., *Introduction to Matrices with Applications in Statistics*, Wadsworth Publishing Company, Inc., Belmont, CA, 1969, p. 140.

[15]Crassidis, J. L. and Markley, F. L., "New Algorithm for Attitude Determination Using Global Positioning System Signals," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 5, 1997, pp. 891–896.

[16]Bellman, R., *Introduction to Matrix Analysis*, McGraw-Hill Book Company, New York, NY, 1960, pp. 58–59.

List of Figure Captions

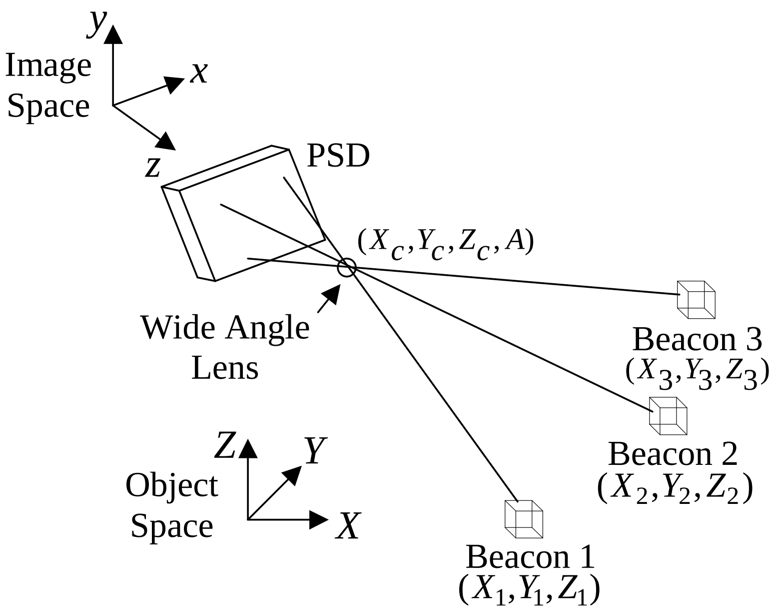

y
Image
Space
x
z
PSD
(X_c, Y_c, Z_c, A)
Wide Angle
Lens
Beacon 3
(X_3, Y_3, Z_3)
Beacon 2
(X_2, Y_2, Z_2)
Z
Y
Object
Space
X
Beacon 1
(X_1, Y_1, Z_1)

FIGURE 1:

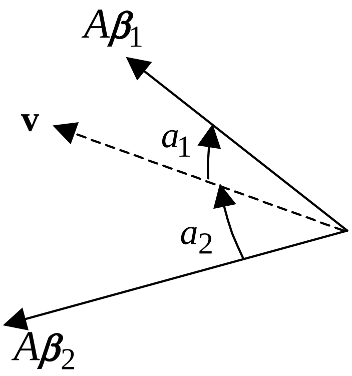
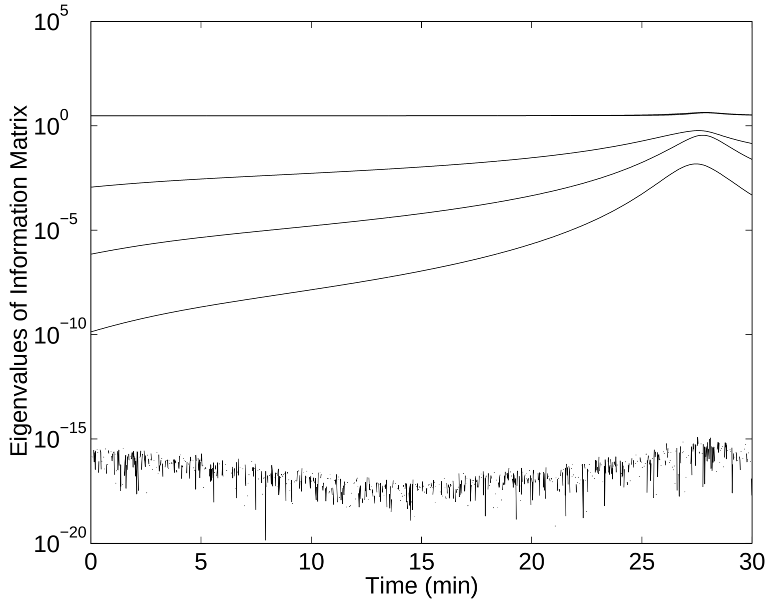
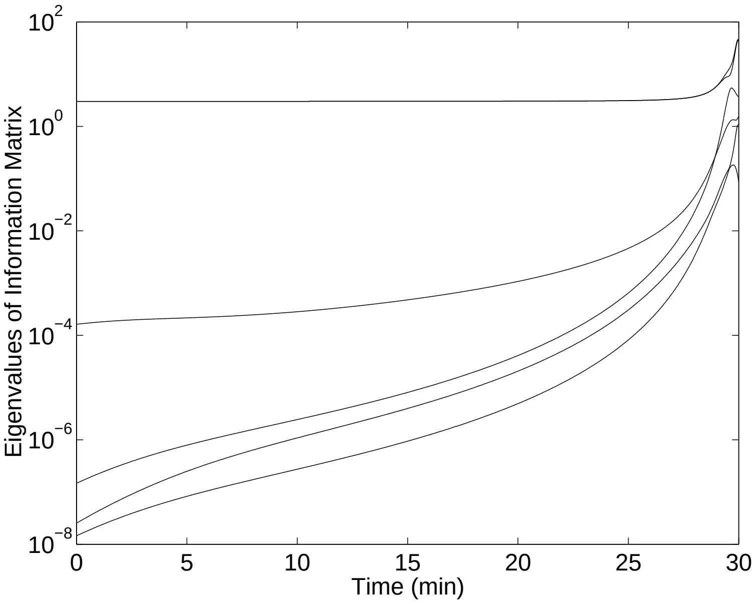


FIGURE 2:



Eigenvalues of Information Matrix
10^5
10^0
10^−5
10^−10
10^−15
10^−20
0
5
10
15
20
25
30
Time (min)

FIGURE 3:



Eigenvalues of Information Matrix
Time (min)
10^{2}
10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
0
5
10
15
20
25
30

FIGURE 4: