The Journal of the Astronautical Sciences, Vol. 51, No. 3, July–September 2003, pp. 279–290

# Batch, Sequential and Hybrid Approaches to Spacecraft Sensor Alignment Estimation

**Maria Celília Zanardi[1] and Malcolm D. Shuster[2]**

**Abstract**

Two Kalman-filter formulations are presented for the estimation of spacecraft sensor misalignments from inflight data. In the first the sensor misalignments are part of the filter state variable; in the second, which we call HYLIGN, the state vector contains only dynamical variables, but the sensitivities of the filter innovations to the misalignments are calculated within the Kalman filter. This procedure permits the misalignments to be estimated in batch mode as well as a much smaller dimension for the Kalman filter state vector. This results not only in a significantly smaller computational burden but also in a smaller sensitivity of the misalignment estimates to outliers in the data. Numerical simulations of the filter performance are presented.

## Introduction

Alignment estimation forms an important part of many missions since the alignment estimation accuracy directly affects the accuracy with which the payload attitude can be determined. A complete treatment of the batch estimation of spacecraft sensor alignments from flight data has been presented previously [1, 2]. The use of these batch techniques, however, requires that the data be arranged in repeated frames of simultaneous measurements. The attitude sensors, however, are typically sampled asynchronously.

The present work presents two filter approaches. The first approach is that of the standard, or "naive" Kalman filter, in which any parameter to be estimated is made a component of the state vector. Since a spacecraft may easily have ten sensors, this leads to a state vector of dimension at least 36 when one considers also the minimum number of dynamical variables. This high dimensionality, coupled with the inherent nonlinearity of the dependence of the measurements on the attitude can lead to poor convergence of the filter in addition to a large computational burden.

[1]Assistant Professor, Orbital Dynamics and Planetology Group, Department of Mathematics, Universidade Estadual Paulista (UNESP), Campus de Guaratinguetá, Guaratinguetá (SP), Brazil.
[2]Director of Research, Acme Spacecraft Company, 13017 Wisteria Drive, Box 328, Germantown, Maryland 20874. email: m.shuster@ieee.org.

The solution to this dilemma was provided by Gupta and Mehra [3], who noted that the innovations sequence in the filter was white and therefore provided the appropriate uncorrelated sequence of effective measurement sequence. This made it possible to apply MLE techniques directly to the innovations sequence. In fact, since the Kalman filter itself may be treated as an MLE algorithm for the case of Gaussian measurement and process noise [4], this meant that the entire estimation process could be treated entirely within the framework of MLE. Friedland [5] has shown how the estimation of the deterministic parameters (i.e., parameters which are not random variables) may be carried out efficiently in a second Kalman filter, using sensitivity matrices computed in the Kalman filter for the stochastic dynamical variables (in this work, the attitude and the gyro biases). Since alignments are purely static, however, these can be estimated via batch lease squares using Friedland's sensitivity matrices with still greater computational savings. This is the second method presented in this work, which we call the HYLIGN algorithm (for HYbrid aLIGNment algorithm), since it combines the best qualities of the Kalman filter and batch least-square estimation.

## Basic Definitions

*Sensor Referenced Measurements*

A spacecraft line-of-sight sensor such as a vector Sun sensor or star tracker measures a direction $\hat{\mathbf{U}}_{\rm i,k}$ in sensor coordinates, defined to be directed outward from the sensor, which is describable statistically as

$$\hat{\mathbf{U}}_{\rm i,k} = \hat{\mathbf{U}}_{\rm i,k}^{\rm true} + \Delta\hat{\mathbf{U}}_{\rm i,k} \tag{1}$$

where $\hat{\mathbf{U}}_{\rm i,k}^{\rm true}$ is the true value of the direction and $\Delta\hat{\mathbf{U}}_{\rm i,k}$ is the measurement noise. Here $i$ is the sensor index, $i = 1, \ldots, n$, and $k$ is the temporal index, $k = 1, \ldots, N$. We assume that $\Delta\hat{\mathbf{U}}_{\rm i,k}$ is Gaussian, zero-mean, and white, with covariance matrix $R_{\hat{\rm U}_{\rm i,k}}$. Because the observations are constrained to be unit vectors, $R_{\hat{\rm U}_{\rm i,k}}$ must be singular with $\hat{\mathbf{U}}_{\rm i,k}^{\rm true}$ a null vector.

*Body-Referenced Vectors and Alignments*

If $\hat{\mathbf{W}}_{\rm i,k}$ denotes the measured direction in the spacecraft body frame, then the alignment matrix, $S_i$, is the proper orthogonal matrix defined by

$$\hat{\mathbf{W}}_{\rm i,k} = S_i\,\hat{\mathbf{U}}_{\rm i,k} \tag{2}$$

and, therefore

$$\hat{\mathbf{W}}_{\rm i,k} = S_i\,\hat{\mathbf{U}}_{\rm i,k}^{\rm true} + S_i\,\Delta\hat{\mathbf{U}}_{\rm i,k} \tag{3}$$

$$\equiv \hat{\mathbf{W}}_{\rm i,k}^{\rm true} + \Delta\hat{\mathbf{W}}_{\rm i,k} \tag{4}$$

Thus, the body-referenced observations have an error covariance matrix given by

$$R_{\hat{\rm W}_{\rm i,k}} = S_i\,R_{\hat{\rm U}_{\rm i,k}}\,S_i^{\rm T} \tag{5}$$

*Misalignments*

In general, the alignment matrix $S_i$ is not known exactly. Instead, what is known is $S_i^o$, the alignment matrix determined by the prelaunch alignment calibration. Thus, we are led to define the misalignment matrix, $M_i$, according to

$$S_i = M_i\,S_i^o \tag{6}$$

$M_i$ is necessarily orthogonal. Therefore, we define the misalignment vectors, $\boldsymbol{\theta}_{\rm i}$, according to

$$M_i \equiv e^{[\![\boldsymbol{\theta}_{\rm i}]\!]} = I + \left(\frac{\sin|\boldsymbol{\theta}_{\rm i}|}{|\boldsymbol{\theta}_{\rm i}|}\right)[\![\boldsymbol{\theta}_{\rm i}]\!] + \left(\frac{1 - \cos|\boldsymbol{\theta}_{\rm i}|}{|\boldsymbol{\theta}_{\rm i}|^2}\right)[\![\boldsymbol{\theta}_{\rm i}]\!]^2 \tag{7}$$

where $e^{\{\cdot\}}$ denotes the matrix exponentiation, and $[\![\boldsymbol{\theta}]\!]$ denotes the usual antisymmetric matrix

$$[\![\boldsymbol{\theta}]\!] \equiv \begin{bmatrix} 0 & \theta_3 & -\theta_2 \\ -\theta_3 & 0 & \theta_1 \\ \theta_2 & -\theta_1 & 0 \end{bmatrix} \tag{8}$$

Equation (7) is just Euler's formula for the rotation matrix recast as a function of the rotation vector. The angles $\theta_1$, $\theta_2$, $\theta_3$ are misalignment angles or simply the misalignments. Since the misalignment matrix is generally a very small rotation, the misalignments will be small and we can write

$$M_i = I + [\![\boldsymbol{\theta}_{\rm i}]\!] + O(|\boldsymbol{\theta}_{\rm i}|^2) \tag{9}$$

As a rule, we will keep only first-order terms. The measurement equation now becomes finally

$$\hat{\mathbf{U}}_{\rm i,k} = S_i^{o\rm T} M_i^{\rm T} \hat{\mathbf{W}}_{\rm i,k}^{\rm true} + \Delta\hat{\mathbf{U}}_{\rm i,k} \tag{10}$$

*Dependence of the Measurements on the Attitude*

If $\hat{\mathbf{V}}_{\rm i,k}$ denotes the reference vector, i.e., the representation of the measured vector in the primary reference system (for example, geocentric inertial), then the attitude matrix $A_k$ is defined according to

$$\hat{\mathbf{W}}_{\rm i,k}^{\rm true} = A_k \hat{\mathbf{V}}_{\rm i,k} \tag{11}$$

whence

$$\hat{\mathbf{W}}_{\rm i,k} = A_k \hat{\mathbf{V}}_{\rm i,k} + \Delta\hat{\mathbf{W}}_{\rm i,k} \tag{12}$$

We assume that $\hat{\mathbf{V}}_{\rm i,k}$ is free of error. From this it follow that the actual sensor measurements are related to the reference vectors by

$$\hat{\mathbf{U}}_{\rm i,k} = S_i^{\rm T} A_k \hat{\mathbf{V}}_{\rm i,k} + \Delta\hat{\mathbf{U}}_{\rm i,k} \tag{13}$$

We note immediately from equation (13) that the values of the measurement vectors are unchanged by the simultaneous transformations

$$S_i \to TS_i, \qquad i = 1, \ldots, n \tag{14a}$$

$$A_k \to TA_k, \qquad k = 1, \ldots, N \tag{14b}$$

where $T$ is an arbitrary proper orthogonal matrix. Thus, it is impossible from inflight sensor measurements to distinguish a common misalignment of the sensors from a change in the attitude. It is, therefore, impossible to estimate the sensor alignments and the attitude unambiguously from the spacecraft sensor measurements alone, and some additional measurement, e.g., the prelaunch alignment calibration, is needed in order to obtain separate estimates of these quantities. In terms of the misalignments, equation (13) becomes

$$\hat{\mathbf{U}}_{\rm i,k} = S_i^{o\rm T} M_i^{\rm T} A_k \hat{\mathbf{V}}_{\rm i,k} + \Delta\hat{\mathbf{U}}_{\rm i,k} \tag{15}$$

This is the point of departure of the alignment estimation algorithms which we will now consider.

## Estimation of Alignments as Kalman-Filter State Variables

Since the Kalman filter can be formulated as a maximum likelihood estimator [4], the Kalman filter estimate of the alignments is also a maximum likelihood estimate, the one which takes account of the spacecraft dynamical degrees of freedom.

Assume that the spacecraft is equipped with $n$ vector sensors for which we wish to estimate alignments using the Kalman filter. The complete state vector, $\mathbf{X}(t)$, in the context of combined attitude and alignment estimation is

$$\mathbf{X}(t) = [\bar{q}^{\mathrm{T}}(t), \boldsymbol{\varepsilon}^{\mathrm{T}}(t), \bar{q}_1^{\mathrm{T}}(t), \ldots, \bar{q}_n^{\mathrm{T}}(t)]^{\mathrm{T}} \tag{16}$$

where $\bar{q}(t)$ is the attitude quaternion, $\bar{q}_i(t)$, $i = 1, \ldots, n$, are the alignment quaternia, which have the same relation to the respective alignment matrices as the attitude quaternion has to the attitude matrix, and $\varepsilon(t)$ is the gyro bias vector. The inclusion of additional degrees of freedom in the state vector is straightforward but needlessly complicates the present discussion.

The state equations for the attitude and the gyro biases are usually modeled [6, 7] as

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{q}(t) = \frac{1}{2}\Omega(\mathbf{g}(t) - \boldsymbol{\varepsilon}(t) - \boldsymbol{\eta}_1(t))\bar{q}(t) \tag{17a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\varepsilon}(t) = \boldsymbol{\eta}_2(t) \tag{17b}$$

where $\mathbf{g}(t)$ is the gyro reading, and $\boldsymbol{\eta}_1(t)$ and $\boldsymbol{\eta}_2(t)$ are white Gaussian processes with power spectral density matrices $Q_1(t)$ and $Q_2(t)$, respectively, $\Omega(\boldsymbol{\omega})$ is the $4 \times 4$ matrix

$$\Omega(\boldsymbol{\omega}) \equiv \begin{bmatrix} 0 & \omega_3 & -\omega_2 & \omega_1 \\ -\omega_3 & 0 & \omega_1 & \omega_2 \\ \omega_2 & -\omega_1 & 0 & \omega_3 \\ -\omega_1 & -\omega_2 & -\omega_3 & 0 \end{bmatrix} \tag{18}$$

The gyro-referenced attitude $\bar{q}_{\mathrm{ref}}(t)$ in the interval $(t_{k-1}, t_k)$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{q}_{\mathrm{ref}}(t) = \frac{1}{2}\Omega(\mathbf{g}(t) - \boldsymbol{\varepsilon}^*_{\mathrm{k-1|k-1}})\bar{q}(t) \tag{19}$$

and the complete state vector based on the gyro-referenced attitude and the prelaunch alignments is, in obvious notation

$$\mathbf{X}^{\mathrm{ref}}(t) = [\bar{q}_{\mathrm{ref}}^{\mathrm{T}}(t), \boldsymbol{\varepsilon}^{\mathrm{T}}(0), \bar{q}_1^{o\mathrm{T}}(t), \ldots, \bar{q}_n^{o\mathrm{T}}(t)]^{\mathrm{T}} \tag{20}$$

where for uniformity we have written time arguments for the alignment quaternia.

The incremental attitude quaternion is given by

$$\delta\bar{q}(t) = \bar{q}(t) \otimes (\bar{q}_{\mathrm{ref}}(t))^{-1} \tag{21}$$

In general, $\delta\bar{q}(t)$ will be the quaternion of an infinitesimal rotation, which we may write as

$$\delta\bar{q}(t) = \begin{bmatrix} \boldsymbol{\xi}(t)/2 \\ 1 \end{bmatrix} + O(|\boldsymbol{\xi}(t)|^2) \tag{22}$$

Defining the gyro-bias increment vector by

$$\Delta\boldsymbol{\varepsilon}(t) = \boldsymbol{\varepsilon}^*(t) - \boldsymbol{\varepsilon}^*_{\mathrm{k-1|k-1}} \tag{23}$$

leads to the incremental equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\xi}(t) = -(\mathbf{g}(t) - \boldsymbol{\varepsilon}(0)) \times \boldsymbol{\xi}(t) - \Delta\boldsymbol{\varepsilon}(t) - \boldsymbol{\eta}_1(t) \tag{24a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta\boldsymbol{\varepsilon}(t) = \boldsymbol{\eta}_2(t) \tag{24b}$$

Likewise, we assume that the spacecraft is rigid and the misalignments satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\theta}_{\mathrm{i}}(t) = 0, \qquad i = 1, \ldots, n \tag{24c}$$

Thus, we define the incremental state vector as

$$\mathbf{x}(t) \equiv [\boldsymbol{\xi}^{\mathrm{T}}(t), \Delta\boldsymbol{\varepsilon}^{\mathrm{T}}(t), \boldsymbol{\theta}_1^{\mathrm{T}}(t), \ldots, \boldsymbol{\theta}_n^{\mathrm{T}}(t)]^{\mathrm{T}} \tag{25}$$

The complete state vector has dimension $4(n + 1) + 3$, while the incremental state vector has dimension $3(n + 1) + 3$. The composition of the reference complete state vector with the incremental state vector is not simple addition. Note that in the above formulation of the Kalman filter, the gyro measurements have replaced the dynamical model and the gyro measurement noise has become state process noise.

The discretised incremental state vector satisfies a state equation of the form

$$\mathbf{x}_{\mathrm{k+1}} = \Phi_k \mathbf{x}_{\mathrm{k}} + \mathbf{w}_{\mathrm{k}} \tag{26}$$

where $\mathbf{w}_{\mathrm{k}}$ is a discrete white noise process calculated from $\boldsymbol{\eta}_1(t)$ and $\boldsymbol{\eta}_2(t)$ and with covariance matrix $Q_k$. $\Phi_k$ and $\mathbf{w}_{\mathrm{k}}$ must be such that

$$\boldsymbol{\theta}_{\mathrm{i,k+1}} = \boldsymbol{\theta}_{\mathrm{i,k}}, \qquad i = 1, \ldots, n \tag{27}$$

The state covariance is defined in terms of the incremental state vector

$$P_k = E\{\mathbf{x}_{\mathrm{k}}\mathbf{x}_{\mathrm{k}}^{\mathrm{T}}\} - E\{\mathbf{x}_{\mathrm{k}}\}E\{\mathbf{x}_{\mathrm{k}}^{\mathrm{T}}\} \tag{28}$$

and the Kalman filter is mechanized in terms of $\mathbf{x}_{\mathrm{k}}$. The prediction equations are

$$\mathbf{x}_{\mathrm{k|k-1}} = \Phi_{k-1}\mathbf{x}_{\mathrm{k-1|k-1}} \tag{29a}$$

$$P_{k|k-1} = \Phi'_{k-1} P_{k-1|k-1} \Phi'^{\mathrm{T}}_{k-1} + Q_{k-1} \tag{29b}$$

To simplify the notation, we do not write an asterisk to denote the estimate or estimator when the subscript makes this identification clear.

The prediction of the misalignments as given by equation (29a) is necessarily

$$\boldsymbol{\theta}_{\mathrm{i,k|k-1}} = \boldsymbol{\theta}_{\mathrm{i,k-1|k-1}} \tag{30}$$

The primes on the transition matrices in equation (29b) are a result of the basic nonlinearity of the combined attitude-gyro-bias dynamics, which leads to different transition matrices for the incremental state vectors and the incremental state errors [7]. Note that $\boldsymbol{\xi}_{\mathrm{k}}$ is related to the attitude matrix according to

$$A_k = e^{[\![\boldsymbol{\xi}_{\mathrm{k}}]\!]} A_k^{\mathrm{ref}} \tag{31}$$

which is similar to the equivalent relation for the misalignment vectors

$$S_{i,k} = e^{[\![\boldsymbol{\theta}_{\mathrm{i,k}}]\!]} S_i^o \tag{32}$$

Thus, we may write the measurement equation as

$$\begin{aligned}\hat{\mathbf{U}}_{\mathrm{i,k}} &= S_{i,k}^{\mathrm{T}} A_k \hat{\mathbf{V}}_{\mathrm{i,k}} + \Delta\hat{\mathbf{U}}_{\mathrm{i,k}} \\ &= \hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}} + C_{i,k}^{\xi}\boldsymbol{\xi}_{\mathrm{k}} + C_{i,k}^{\theta}\boldsymbol{\theta}_{\mathrm{k}} + \Delta\hat{\mathbf{U}}_{\mathrm{i,k}} \end{aligned} \tag{33}$$

where

$$\hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}} \equiv S_i^{o\mathrm{T}} \hat{\mathbf{W}}_{\mathrm{i,k}}^{\mathrm{ref}} \tag{34}$$

$$\hat{\mathbf{W}}_{\mathrm{i,k}}^{\mathrm{ref}} \equiv A_k^{\mathrm{ref}} \hat{\mathbf{V}}_{\mathrm{i,k}} \tag{35}$$

are the sensor-referenced and body-referenced measurements for the reference trajectory

$$C_{i,k}^{\theta} = -C_{i,k}^{\xi} = S_i^{o\mathrm{T}} [\![\hat{\mathbf{W}}_{\mathrm{i,k}}^{\mathrm{ref}}]\!] = [\![\hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}}]\!] S_i^{o\mathrm{T}} \tag{36}$$

are the measurement sensitivity matrices, and

$$\hat{\mathbf{U}}_{\mathrm{i,k}} = \hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}} + C_{i,k}\mathbf{x}_{\mathrm{k}} + \Delta\hat{\mathbf{U}}_{\mathrm{i,k}} \tag{37}$$

where the submatrices of $C_{i,k}$ vanish except for those which multiply $\boldsymbol{\xi}$ and $\boldsymbol{\theta}_{\mathrm{i}}$.

In general, the measurements are not the $\hat{\mathbf{U}}_{\mathrm{i,k}}$ themselves but scalar functions of the $\hat{\mathbf{U}}_{\mathrm{i,k}}$, which we denote by $f_{i,k}(\hat{\mathbf{U}}_{\mathrm{i,k}})$. Thus, we define the equivalent scalar measurements as

$$z_k = f_{i,k}(\hat{\mathbf{U}}_{\mathrm{i,k}}) - f_{i,k}(\hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}}) \tag{38}$$

$$\simeq H_k \mathbf{x}_{\mathrm{k}} + \upsilon_k \tag{39}$$

where, we have expanded equation (38) in a Taylor series about $\hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}}$ to obtain

$$H_k = \left(\frac{\partial f_k}{\partial \hat{\mathbf{U}}_{\mathrm{i,k}}}(\hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}})\right)^{\mathrm{T}} C_{i,k} \tag{40}$$

$$\upsilon_k = \left(\frac{\partial f_k}{\partial \hat{\mathbf{U}}_{\mathrm{i,k}}}(\hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}})\right)^{\mathrm{T}} \Delta\hat{\mathbf{U}}_{\mathrm{i,k}} \tag{41}$$

We use throughout the convention that the matrix of partial derivatives of a scalar with respect to a column vector is again a column vector; hence, the transpose superscript in equations (40) and (41).

The measurement equation is now in a form familiar to us. In principal, we can neglect the index $i$ in labeling the measurements, as we have done in equations (38) through (41) if we choose the temporal index so that each scalar measurement corresponds to a different value of $k$ (the order of truly simultaneous measurements is unimportant). Thus, ideally, we should write $i_k$ in place of $i$ and be aware that $t_k$ will sometimes assume the same value for successive values of $k$. In predicting between equal times the transition matrices will be identity matrices and no process noise will be accumulated.

Thus, we may write the Kalman filter equations for the update step as

$$B_k = H_k P_{k|k-1} H_k^{\mathrm{T}} + R_k \tag{42}$$

$$K_k = P_{k|k-1} H_k^{\mathrm{T}} B_k^{-1} \tag{43}$$

$$\nu_k = z_k - H_k \mathbf{x}_{\mathrm{k|k-1}} \tag{44}$$

$$\mathbf{x}_{\mathrm{k|k}} = \mathbf{x}_{\mathrm{k|k-1}} + K_k \nu_k \tag{45}$$

$$P_{k|k} = (I - K_k H_k) P_{k|k-1} (I - K_k H_k)^{\mathrm{T}} + K_k R_k K_k^{\mathrm{T}} \tag{46}$$

and $R_k$ is the variance of $\nu_k$. The *a posteriori* estimate of the misalignment vector and its covariance are given by

$$\mathbf{\Theta}^*(+) = \mathbf{\Theta}_{\mathrm{N|N}}, \qquad P_{\Theta\Theta}(+) = (P_{N|N})_{\Theta\Theta} \tag{47}$$

## Kalman-Filter-Based Batch Estimation of Sensor Alignments

Even if the measurement and process noise are small, the Kalman filter for the attitude and alignments may converge slowly because of the nonlinear dependence of the measurements on the attitude. Also, the filter will be very sensitive to outliers at the beginning of a data segment. Batch algorithms, which process all of the data at once, are less sensitive to outliers and to the nonlinear dependence of the negative-log-likelihood function on the parameters being estimated. However, from equation (39) we see that all of the measurements are correlated with one another through the correlation in $\mathbf{x}_k$. Thus, not only will the parameter set in a batch estimation procedure be very large because of the large number of attitudes to be computed, but the measurement covariance matrix, if all of the measurements were stacked into one large measurement vector, would be very large and nondiagonal, hence, very difficult to invert.

A method of removing this difficulty was developed by Gupta and Mehra [3]. These authors noted that although the measurements, $z_k$, are correlated, the innovations, $\nu_k$ computed by the Kalman filter are always a white sequence. Hence, instead of finding the value of $\mathbf{\Theta}$ which minimizes

$$J(z_1, \ldots, z_N; \mathbf{\Theta})$$

it is sufficient to find the value which minimizes

$$J(\nu_1, \ldots, \nu_N; \mathbf{\Theta})$$

Gupta and Mehra noted also that the Jacobian determinant of the (very high-dimensional) transformation matrix which transforms the column vector containing all the $z_k$ into the column vector containing the corresponding $\nu_k$ will be unity. Hence, the two negative-log-likelihood functions will yield the same Fisher information matrix. Thus, we are led to estimate $\mathbf{\Theta}$ by minimizing the *a posteriori* negative-log-likelihood function

$$\begin{aligned} J(\nu_1, \ldots, \nu_N; \mathbf{\Theta}) = {} & \frac{1}{2}\, \mathbf{\Theta}^{\mathrm{T}} P_{\Theta\Theta}^{-1}(-)\, \mathbf{\Theta} \\ & + \frac{1}{2} \sum_{k=1}^{N} \{\nu_k^{\mathrm{T}}(\mathbf{\Theta}) B_k^{-1}(\mathbf{\Theta})\, \nu_k(\mathbf{\Theta}) + \log\det B_k(\mathbf{\Theta}) + \log 2\pi\} \end{aligned} \tag{48}$$

instead of the negative-log-likelihood function given directly in terms of the $z_k$, although the two are formally equivalent.

In the present instance the total alignment vector, $\mathbf{\Theta}$, is no longer a state variable but a *constant* parameter of the system. The state vector, therefore, is now much reduced in dimension and simply

$$\mathbf{X}_{\mathrm{k}} = \begin{bmatrix} \bar{q}_k \\ \boldsymbol{\varepsilon}_{\mathrm{k}} \end{bmatrix} \qquad \text{and} \qquad \mathbf{x}_{\mathrm{k}} = \begin{bmatrix} \boldsymbol{\xi}_{\mathrm{k}} \\ \Delta\boldsymbol{\varepsilon}_{\mathrm{k}} \end{bmatrix} \tag{49}$$

Thus the gradient of the *a posteriori* negative-log-likelihood function in terms of the innovations process and the residual covariance $B_k$ is [3]

$$\frac{\partial J}{\partial \Theta_m} = (P_{\Theta\Theta}^{-1}(-)\,\boldsymbol{\Theta})_m + \sum_{k=1}^{N}\left\{\frac{\partial \nu_k^{\mathrm{T}}(\boldsymbol{\Theta})}{\partial \Theta_m} B_k^{-1}(\boldsymbol{\Theta})\nu_k(\boldsymbol{\Theta}) - \frac{1}{2}\nu_k^{\mathrm{T}}(\boldsymbol{\Theta})B_k^{-1}(\boldsymbol{\Theta})\frac{\partial B_k(\boldsymbol{\Theta})}{\partial \Theta_m}B_k^{-1}(\boldsymbol{\Theta})\nu_k(\boldsymbol{\Theta}) + \frac{1}{2}\,\mathrm{tr}\left[B_k^{-1}(\boldsymbol{\Theta})\frac{\partial B_k(\boldsymbol{\Theta})}{\partial \Theta_m}\right]\right\} \tag{50}$$

and the corresponding Fisher information matrix is given by

$$F_{lm} \equiv E\left\{\frac{\partial^2 J}{\partial \Theta_l\,\partial \Theta_m}\right\} = (P_{\Theta\Theta}^{-1}(-))_{lm} + \sum_{k=1}^{N}\left\{\frac{1}{2}\,\mathrm{tr}\left[B_k^{-1}(\boldsymbol{\Theta})\frac{\partial B_k(\boldsymbol{\Theta})}{\partial \Theta_l}B_k^{-1}(\boldsymbol{\Theta})\frac{\partial B_k(\boldsymbol{\Theta})}{\partial \Theta_m}\right] + E\left\{\left[\frac{\partial \nu_k^{\mathrm{T}}(\boldsymbol{\Theta})}{\partial \Theta_l}B_k^{-1}(\boldsymbol{\Theta})\frac{\partial \nu_k(\boldsymbol{\Theta})}{\partial \Theta_m}\right]\right\}\right\} \tag{51}$$

(Note that Gupta and Mehra make an error in their derivation of equation (51) leading them to include an extraneous term.) The mechanization of the filter now proceeds as before without the components related to the misalignments, which are now simply constant parameters in the measurements. Equation (32) is now replaced by

$$S_i = M_i S_i^o = e^{[\![\boldsymbol{\theta}_i]\!]} S_i^o \tag{52}$$

which is the same as in the batch estimator presented earlier. Equations (38) and (39) now become

$$z_k = f_{i,k}(\hat{\mathbf{U}}_{\mathrm{i,k}}) - f_{i,k}(\hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}}) \tag{53}$$

$$\simeq H_k^I \mathbf{x}_{\mathrm{k}} + C_k\Theta + \upsilon_k \tag{54}$$

where

$$H_k^I = \left(\frac{\partial f_k}{\partial \hat{\mathbf{U}}_{\mathrm{i,k}}}(\hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}})\right)^{\mathrm{T}}[C_{i,k}^{\xi} \,\vdots\, O_{3\times 3}] \tag{55}$$

$$C_k = \left(\frac{\partial f_k}{\partial \hat{\mathbf{U}}_{\mathrm{i,k}}}(\hat{\mathbf{U}}_{\mathrm{i,k}}^{\mathrm{ref}})\right)^{\mathrm{T}}[O_{3\times 3} \cdots C_{i,k}^{\theta} \cdots O_{3\times 3}] \tag{56}$$

with $C_{i,k}^{\theta}$ and $C_{i,k}^{\xi}$ given still by equation (36), and the nonzero entries in $C_k$ occur in the submatrix which multiplies $\boldsymbol{\theta}_{\mathrm{i}}$. The superscript $I$, distinguishes the measurement sensitivity matrix in equation (54) from the related quantity in equation (39) *et seq.* and denotes that it represents that component of the measurement which is insensitive to the alignments.

To calculate the dependence of $\nu_k$ on $\boldsymbol{\Theta}$ we note that because the Kalman filter consists only of linear operations on the state variables we may write

$$\mathbf{x}_{\mathrm{k|k-1}} = \mathbf{x}_{\mathrm{k|k-1}}^{\mathrm{I}} - T_{k|k-1}\boldsymbol{\Theta} \tag{57a}$$

$$\mathbf{x}_{\mathrm{k|k}} = \mathbf{x}_{\mathrm{k|k}}^{\mathrm{I}} - T_{k|k}\boldsymbol{\Theta} \tag{57b}$$

where $\mathbf{x}^{\mathrm{I}}_{\mathrm{k|k-1}}$ and $\mathbf{x}^{\mathrm{I}}_{\mathrm{k|k}}$ are independent of $\boldsymbol{\Theta}$. To determine these alignment-independent state estimates and the alignment sensitivity matrices $T_{k|k-1}$ and $T_{k|k}$ we substitute these expressions into the Kalman filter equations to obtain new filter equations of the form

$$\mathbf{x}^{\mathrm{I}}_{\mathrm{k|k-1}} = \Phi^{I}_{k-1}\mathbf{x}^{\mathrm{I}}_{\mathrm{k-1|k-1}} \tag{58}$$

$$P^{I}_{k|k-1} = \Phi'^{I}_{k-1} P^{I}_{k|k-1}(\Phi'^{I}_{k-1})^{\mathrm{T}} + Q^{I}_{k-1} \tag{59}$$

$$B^{I}_{k} = H^{I}_{k}P^{I}_{k|k-1}(H^{I}_{k})^{\mathrm{T}} + R_{k} \tag{60}$$

$$K^{I}_{k} = P^{I}_{k|k-1}(H^{I}_{k})^{\mathrm{T}}(B^{I}_{k})^{-1} \tag{61}$$

$$\nu^{I}_{k} = z_{k} - H^{I}_{k}\mathbf{x}^{\mathrm{I}}_{\mathrm{k|k-1}} \tag{62}$$

$$\mathbf{x}^{\mathrm{I}}_{\mathrm{k|k}} = \mathbf{x}^{\mathrm{I}}_{\mathrm{k|k-1}} + K^{I}_{k}\nu^{I}_{k} \tag{63}$$

$$P^{I}_{k|k} = (I - K^{I}_{k}H^{I}_{k})P^{I}_{k|k-1}(I - K^{I}_{k}H^{I}_{k})^{\mathrm{T}} + K^{I}_{k}R_{k}(K^{I}_{k})^{\mathrm{T}} \tag{64}$$

and the superscript $I$ on $\Phi_k$ and $Q_k$ distinguishes these quantities from related matrices of larger dimension in the previous Kalman-filter implementation. The truncation amounts to simply deleting zeros and ones corresponding to the alignment components. The alignment sensitivity matrices are given by the recursion relations

$$T_{o|o} = 0 \tag{65a}$$

$$T_{k|k-1} = \Phi^{I}_{k-1}T_{k-1|k-1} \tag{65b}$$

$$T_{k|k} = (I - K^{I}_{k}H^{I}_{k})T_{k|k-1} - K^{I}_{k}C_{k} \tag{65c}$$

The innovation is thus given by

$$\nu_{k} = z_{k} - H^{I}_{k}\mathbf{x}_{\mathrm{k|k-1}} - C_{k}\boldsymbol{\Theta} \tag{66}$$

$$= \nu^{I}_{k} - F_{k}\boldsymbol{\Theta} \tag{67}$$

where

$$F_{k} = H^{I}_{k}T_{k|k-1} + C_{k} \tag{68}$$

Thus, the prior-free negative-log-likelihood function for the misalignments is given by

$$J^{\text{prior-free}}(\boldsymbol{\Theta}) = \frac{1}{2}\sum_{k=1}^{N}\{(\nu^{I}_{k} - F_{k}\boldsymbol{\Theta})^{\mathrm{T}}(B^{I}_{k})^{-1}(\nu^{I}_{k} - F_{k}\boldsymbol{\Theta}) + \log\det B^{I}_{k} + \log 2\pi\} \tag{69}$$

For clarity we write equation (69) in matrix form even though the three factors are each scalars. From this negative-log-likelihood function we may estimate the prior-free relative alignments and the launch-shock error levels. The *a posteriori* estimates of the alignments taking into account both the *a priori* estimate and equation (69) is then given by the usual normal equations

$$P^{-1}_{\Theta\Theta}(+)\boldsymbol{\Theta}^{*}(+) = \sum_{k=1}^{N}F^{\mathrm{T}}_{k}(B^{I}_{k})^{-1}\nu^{I}_{k} \tag{70a}$$

$$P^{-1}_{\Theta\Theta}(+) = P^{-1}_{\Theta\Theta}(-) + \sum_{k=1}^{N}F^{\mathrm{T}}_{k}(B^{I}_{k})^{-1}F_{k} \tag{70b}$$

which are equivalent to equations (49) and (50) if we note that $B_k$ is independent of $\boldsymbol{\Theta}$. The values $\boldsymbol{\Theta}^{*}(+)$ and $P^{-1}_{\Theta\Theta}(+)$ from equations (70) correspond exactly to $\boldsymbol{\Theta}_{\mathrm{N|N}}$ and $(P_{N|N})_{\Theta\Theta}$ which would have been obtained using the larger Kalman filter presented in the previous section.

## Numerical Results

To illustrate the efficacy of the hybrid HYLIGN method we have computed the misalignments for spacecraft with 3, 5, 10, 15, and 20 sensors, oriented at random over the spacecraft. The configuration for three sensors has been taken from [1]. In each case 100 frames of data were simulated, with each sensor active in each frame. Table 1 shows the results for the two sequential algorithms presented here together with the purely batch algorithm of reference [1] (marked SPB). In order to make a comparison possible, we have simulated the sensor data as being simultaneous. Otherwise, the batch method could not be applied at all. The three results are seen to be of equal accuracy. In the simulations we have estimated not the individual misalignments $\theta_1$, $\theta_2$, and $\theta_3$, but rather the relative misalignments [1] $\boldsymbol{\psi}_2 = \boldsymbol{\theta}_2 - \boldsymbol{\theta}_1$ and $\boldsymbol{\psi}_3 = \boldsymbol{\theta}_3 - \boldsymbol{\theta}_1$. These quantities have been estimated without using the *a priori*

**TABLE 1. Comparison of the Three Methods**

**A. Batch Alignment Estimation with the SPB Algorithm**

| $\psi_a^{\text{true}}$ | $\psi_a^*$ | $\sigma_{\psi_a}$ | $\Delta\psi_a^*/\sigma_{\psi_a}$ |
|---|---|---|---|
| −73. | −70.947 | 1.426 | 1.439 |
| −40. | −39.372 | 6.926 | 0.090 |
| 63. | 54.038 | 11.282 | −0.794 |
| −14. | −17.143 | 7.192 | −0.437 |
| −43. | −43.145 | 1.425 | −0.094 |
| 131. | 116.938 | 10.966 | −1.282 |

**B. Naive Kalman Filter Alignment Estimation**

| $\psi_a^{\text{true}}$ | $\psi_a^*$ | $\sigma_{\psi_a}$ | $\Delta\psi_a^*/\sigma_{\psi_a}$ |
|---|---|---|---|
| −73. | −71.011 | 1.424 | 1.396 |
| −40. | −37.903 | 6.487 | 0.323 |
| 63. | 55.238 | 10.845 | −0.716 |
| −14. | −17.288 | 6.766 | −0.486 |
| −43. | −43.086 | 1.423 | −0.061 |
| 131. | 119.475 | 10.675 | −1.080 |

**C. Alignment Estimation with the HYLIGN Algorithm**

| $\psi_a^{\text{true}}$ | $\psi_a^*$ | $\sigma_{\psi_a}$ | $\Delta\psi_a^*/\sigma_{\psi_a}$ |
|---|---|---|---|
| −73. | −71.970 | 1.426 | 1.417 |
| −40. | −39.627 | 6.964 | 0.054 |
| 63. | 58.048 | 11.524 | −0.430 |
| −14. | −17.783 | 7.195 | −0.526 |
| −43. | −43.156 | 1.426 | −0.110 |
| 131. | 120.105 | 11.185 | −0.974 |

information, which is of low quality because of launch shock. The large launch shock error levels would, in fact, obscure the comparison of the different methods. In the table, the first column gives the true value of the misalignment in arc seconds, the second gives the estimated misalignment, the third the standard deviation, and the fourth column the normalized error. If the estimates of the different misalignments were uncorrelated, these last would have a Gaussian distribution with mean zero and variance unity. Although this is not true exactly, the result permits us to assess roughly the consistency of the estimation process. Note that two relative misalignments show much larger error levels than the others. These are the relative misalignments about the sensor boresights. Hence, the model variances are larger due to the well-known geometric dilution of precision.

A comparison of the computational burden (in flops) for the three algorithms is given in Table 2. The results have been normalized so that the computational burden is unity for the hybrid algorithm with three sensors. The comparison for the two sequential estimation techniques only is shown in Table 3. As expected, the elimination of the alignment parameters from the the Kalman filter update leads to a considerable savings in the computational burden, which accelerates as the number of sensors increases.

Space limitations prohibit our presenting the detailed results for the configurations with larger numbers of sensors. The agreement of the two sequential approaches is indistinguishable in general within the computed estimation error, and root-mean-square values of the differences in the normalized errors for the two methods are typically much less than one. For the case of three sensors, the differences between each of the two sequential methods and the batch method are smaller than that between the two sequential methods, as is to be expected owing to the significantly greater computational burden of each relative to the batch method.

**TABLE 2. Relative Timings for the Case of Three Sensors**

| Algorithm | Relative Computation Time (flops) |
|---|---|
| SPB Batch Method | 0.2 |
| Naive Kalman Filter | 2.0 |
| HYLIGN Algorithm | 1.0 |

**TABLE 3. Relative Timings for the Filter-Based Algorithms**

| Number of Sensors | Naive KF | HYLIGN Algorithm |
|---|---|---|
| 3 | 2.0 | 1.0 |
| 5 | 7.8 | 2.9 |
| 10 | 57.7 | 17.9 |
| 15 | 191.6 | 56.5 |
| 20 | 451.1 | 130.2 |

## Summary

We have presented two sequential methods for estimating spacecraft sensor misalignments from flight data. The first was a naive Kalman filter which included the sensor alignments among the state variables, the second was a hybrid approach, HYLIGN, based on the ideas of Gupta and Mehra [3] and Friedland [5], which excluded the alignment parameters from the filter. The HYLIGN method has been shown to be more efficient computationally. It is also expected to be less sensitive to outliers, and is to be preferred in general.

### Acknowledgment

One of the authors (M.C.Z.) wishes to thank the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for its generous support and the Department of Aerospace Engineering, Mechanics and Engineering Science of the University of Florida, Gainesville, Florida, for its kind hospitality during the academic year 1995–1996.

## References

[1] SHUSTER, M. D. and PITONE, D. S. "Batch Estimation of Spacecraft Sensor Alignments, I. Relative Alignment Estimation," *The Journal of the Astronautical Sciences,* Vol. 39, No. 4, pp. 519–546, October–December, 1991.

[2] SHUSTER, M. D., BIERMAN, G. J., and PITONE, D. S. "Batch Estimation of Spacecraft Sensor Alignments, II. Absolute Alignment Estimation," *The Journal of the Astronautical Sciences,* Vol. 39, No. 4, pp. 547–571, October–December, 1991.

[3] GUPTA, N. K. and MEHRA, R. K., "Computational Aspects of Maximum Likelihood Estimation and Reduction in Sensitivity Function Calculations," *IEEE Transactions in Automatic Control,* Vol. AC-19, No. 6, pp. 774–783, 1974.

[4] RAUCH, H. E., TUNG, F., and STRIEBEL, C. T. "Maximum Likelihood Estimation of Linear Dynamic Systems," *AIAA Journal,* Vol. 3, No. 8, pp. 1445–1450, 1965.

[5] FRIEDLAND, B. "Treatment of Bias in Recursive Filtering," *IEEE Transactions on Automatic Control,* Vol. AC-17, No. 1, pp. 359–367, 1969.

[6] MURRELL, J. W. "Precision Attitude Determination for Multimission Spacecraft," Paper No. 78-1248, *Proceedings, AIAA Guidance and Control Conference,* Palo Alto, California, August 7–9, 1978.

[7] LEFFERTS, E. J., MARKLEY, F. L., and SHUSTER, M. D. "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control and Dynamics,* Vol. 5, No. 5, Sept.–Oct. 1982, pp. 417–429.