JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 28, No. 4, July–August 2005

Estimation of a Spacecraft's Attitude Dynamics Parameters by Using Flight Data

Mark L. Psiaki*
Cornell University, Ithaca, New York 14853-7501

An algorithm that uses flight data to estimate the parameters in an attitude dynamics model of a spacecraft has been developed. The new algorithm's estimates can enhance the fidelity of Euler equation models that are used to implement attitude determination and control functions. The algorithm's estimation equation is an integrated version of Euler's equation expressed in inertial coordinates. It uses three-axis attitude data and three-axis rate-gyro data to yield a set of linear equations in the unknown dynamics parameters, which include moments and products of inertia and scale factors, alignments, and biases for all reaction wheels and magnetic torque rods. The estimation problem statement includes the statistics of unmodeled torques and sensor errors, and it incorporates a scalar quadratic constraint to overcome the unobservability of the parameters' overall scaling. The sensor errors enter the model equation in a multiplicative fashion, which yields a total least-squares problem. The solution algorithm employs a guarded Newton iteration and recursive factorizations that deal efficiently with the problem's dynamic structure. The algorithm has been applied to the Wilkinson Microwave Anisotropy Probe spacecraft, and the resulting parameter estimates reduce torque modeling errors by a factor of 5–7.

I. Introduction

THE goal of this research is to develop improved Euler attitude dynamics models for spacecraft by using flight data to estimate model parameters.

The primary motivation for this work is to improve the performance of backup mode attitude determination systems. They must be able to operate without rate-gyro data or with incomplete rate-gyro information. This type of system relies on Euler's equation to propagate attitude rate estimates as in Refs. 1–6. Improved Euler models also could be used in slewing and pointing controllers that function with a reduced actuator set, with long open-loop intervals, or with feed-forward terms as in the normal scanning mode of the Wilkinson Microwave Anisotropy Probe (WMAP) spacecraft.[7]

Backup mode attitude determination and control systems that rely on an Euler equation model can exploit the property of observability or controllability to achieve three-axis performance without the need for full three-axis sensing or actuation. Their stability and accuracy are dependent on the accuracy with which the rate dynamics can be modeled by Euler's equations. It would be good to have better parameter estimates in this situation.

This paper envisions the availability of highly accurate sensor data during a time period before backup mode attitude determination is required. Its estimation algorithm uses these data to develop attitude parameter estimates that could be used as part of a backup system that could operate later in the mission if any of the primary sensors failed.

A number of previous works have dealt with estimation of moment-of-inertia matrix elements and additional dynamics parameters. Reference 2, 3, 5, and 6 estimated corrections to five of the six moment-of-inertia matrix elements in attempts to improve Euler-based attitude and rate estimation filters. These efforts considered both off-line estimation of moment-of-inertia parameters based on the nutational properties of a spinning spacecraft[2] and on-line estimation via inclusion of inertia matrix elements as states in an extended Kalman filter, whose primary goal was to determine attitude, rate, or both.[3,5,6] A set of batch estimation methods that are based on angular velocity measurements in spacecraft coordinates has been developed. One strategy develops measurement equations for a spacecraft's moment-of-inertia matrix, center-of-mass location, and mass by considering the sudden angular rate changes caused by thruster firings or caused by abrupt angular momentum changes in reaction wheels or control moment gyros (CMGs).[8–10] Another approach develops estimation equations based on the time-varying spacecraft kinetic energy.[11] It works for a spinning spacecraft and is able to consider coasting periods as well as the effects of control inputs. Yet another set of techniques develops measurement equations based on the angular momentum dynamics in inertial coordinates,[12] or based on projections of the angular momentum vector onto measured inertial directions.[13] One effort dealt with estimation of the alignments of reaction/momentum wheels and single-axis CMGs.[14]

This paper's principal contributions are to define a new, more general attitude dynamics parameter estimation problem and to develop an algorithm to solve it. The estimation problem is based on angular momentum dynamics in inertial coordinates as in Ref. 12. The estimated quantities are a general set of Euler dynamics parameters that include moment-of-inertia matrix elements, reaction-wheel scale factors, alignments, and biases, and magnetic torque rod scale factors, alignments, and biases. These parameters enter the angular momentum equation linearly. The problem formulation presumes that accurate three-axis attitude and rate sensor time histories are available along with reaction-wheel angular momenta, torque rod dipole moments, and magnetometer outputs, if relevant. The measurement errors, such as rate-gyro errors and errors in the reported reaction-wheel angular momenta, enter as perturbations to the coefficient matrix that multiplies the unknown parameter vector in the estimation equation. This bilinear form of the model equation is known as a total least-squares problem.[15,16] The solution algorithm simultaneously estimates the unknown parameter vector and the multiplicative measurement errors.

The remaining five sections of this paper describe the parameter estimation problem, the solution algorithm, and the results that have been obtained for a test case. Section II sets up the total least-squares estimation problem. Section III develops the solution algorithm. Section IV presents a recursive orthogonal-triangular (QR) factorization algorithm that gets used to efficiently implement parts of Sec. III's algorithm. Section V applies the algorithm to data from the WMAP spacecraft, and Sec. VI presents the paper's conclusions.

Received 31 December 2003; revision received 20 May 2004; accepted for publication 24 May 2004.

*Associate Professor, Sibley School of Mechanical and Aerospace Engineering. Associate Fellow AIAA.

II. Attitude Parameter Estimation Problem Based on Euler's Equation

Euler's equation for the attitude dynamics of a rigid spacecraft with reaction/momentum wheels equates the time rate of change of angular momentum in inertial coordinates to the applied external torque. A trapezoidally integrated version of Euler's equation takes on the following form when expressed in inertial coordinates:

$$\begin{aligned} A_{k+1}^T \Big\{ & I_m \boldsymbol{\omega}_{k+1} + C_{\mathrm{rw}} \boldsymbol{h}_{\mathrm{wk+1}} + \boldsymbol{h}_{\mathrm{bias}} - \tfrac{1}{2}(t_{k+1} - t_k)[-\boldsymbol{b}_{\mathrm{sck+1}} \\ & \times (C_{\mathrm{mt}} \boldsymbol{m}_{k+1} + \boldsymbol{m}_{\mathrm{bias}}) + \boldsymbol{n}_{\mathrm{gg}}(A_{k+1}, G_{\mathrm{ink+1}}, I_m)] \Big\} \\ = A_k^T \Big\{ & I_m \boldsymbol{\omega}_k + C_{\mathrm{rw}} \boldsymbol{h}_{\mathrm{wk}} + \boldsymbol{h}_{\mathrm{bias}} + \tfrac{1}{2}(t_{k+1} - t_k)[-\boldsymbol{b}_{\mathrm{sck}} \times (C_{\mathrm{mt}} \boldsymbol{m}_k \\ & + \boldsymbol{m}_{\mathrm{bias}}) + \boldsymbol{n}_{\mathrm{gg}}(A_k, G_{\mathrm{ink}}, I_m)] \Big\} + \Delta \boldsymbol{h}_{\mathrm{ink}} \end{aligned} \tag{1}$$

The subscripts k and $k+1$ refer to values at the sample times of the beginning and end of the numerical integration interval, t_k and t_{k+1}. Other quantities in Eq. (1) are A, the direction cosines matrix for the transformation from inertial coordinates to spacecraft coordinates; I_m, the moment-of-inertia matrix of the rigid spacecraft body in body coordinates; $\boldsymbol{\omega}$, the angular velocity of the spacecraft with respect to inertial coordinates expressed in body coordinates; C_{rw}, a scale-factor/alignment matrix for the reaction wheels; $\boldsymbol{h}_w$, the vector of nominal reaction-wheel angular momenta caused by their rotation rates with respect to the spacecraft; $\boldsymbol{h}_{\mathrm{bias}}$, the net bias angular momentum in spacecraft coordinates of the reaction wheels and any other spacecraft instruments when $\boldsymbol{h}_w = 0$; $\boldsymbol{b}_{\mathrm{sc}}$, the magnetic field vector in spacecraft coordinates; C_{mt}, a scale-factor/alignment matrix for the magnetic torque rods; $\boldsymbol{m}$, the vector of nominal torque rod magnetic dipole moments; $\boldsymbol{m}_{\mathrm{bias}}$, the net bias dipole moment in spacecraft coordinates of the torque rods and any other spacecraft instruments when $\boldsymbol{m} = 0$; $\boldsymbol{n}_{\mathrm{gg}}(A, G_{\mathrm{in}}, I_m)$, the gravity-gradient torque in spacecraft coordinates; and $\Delta \boldsymbol{h}_{\mathrm{in}}$, the net inertial impulse caused by unmodeled torques during the integration interval. The gravity-gradient torque is a function of the attitude matrix, the gravity-gradient tensor in inertial coordinates G_{in}, and the moment-of-inertia matrix.

The unknown parameters that get estimated are I_m, C_{rw}, $\boldsymbol{h}_{\mathrm{bias}}$, C_{mt}, and $\boldsymbol{m}_{\mathrm{bias}}$. It is assumed that the other quantities in the equation are all known, possibly with some measurement error. Star trackers provide A_k and A_{k+1}, and rate gyros provide $\boldsymbol{\omega}_k$ and $\boldsymbol{\omega}_{k+1}$. A magnetometer measures $\boldsymbol{b}_{\mathrm{sck}}$ and $\boldsymbol{b}_{\mathrm{sck+1}}$. The reaction-wheels output $\boldsymbol{h}_{\mathrm{wk}}$ and $\boldsymbol{h}_{\mathrm{wk+1}}$, and the torque rod controller reports $\boldsymbol{m}_k$ and $\boldsymbol{m}_{k+1}$. The orbital ephemerides and an Earth gravity model yield G_{ink} and $G_{\mathrm{ink+1}}$.

The unknown parameters can be stacked into an estimation vector:

$$\begin{aligned} x = \Big[& I_{m11}, I_{m12}, I_{m13}, I_{m22}, I_{m23}, I_{m33}, \boldsymbol{c}_{\mathrm{rw1}}^T, \boldsymbol{c}_{\mathrm{rw2}}^T, \\ & \ldots, \boldsymbol{c}_{\mathrm{rwnw}}^T, \boldsymbol{h}_{\mathrm{bias}}^T, \boldsymbol{c}_{\mathrm{mt1}}^T, \ldots, \boldsymbol{c}_{\mathrm{mtnt}}^T, \boldsymbol{m}_{\mathrm{bias}}^T \Big]^T \end{aligned} \tag{2}$$

where I_{mij} is the ijth element of I_m, $\boldsymbol{c}_{\mathrm{rwj}}$ is the jth column of C_{rw}, and $\boldsymbol{c}_{\mathrm{mtj}}$ is the jth column of C_{mt}. There are n_w reaction wheels and n_t magnetic torque rods. The vector of unknown parameters has $n_x = 12 + 3(n_w + n_t)$ elements.

The estimation algorithm determines the parameter values that minimize the weighted sum of the squared errors in Euler's equation and in the measurements. The Euler equation cost is the weighted sum of the squared magnitudes of the residual unmodeled angular impulse time history $\Delta \boldsymbol{h}_{\mathrm{in0}}$, $\Delta \boldsymbol{h}_{\mathrm{in1}}$, $\Delta \boldsymbol{h}_{\mathrm{in2}}, \ldots \Delta \boldsymbol{h}_{\mathrm{in}N-1}$. Each $\| \Delta \boldsymbol{h}_{\mathrm{ink}} \|^2$ term gets weighted through division by its per-axis variance $\sigma_{\Delta \mathrm{hk}}^2 = \Delta t_k q_{\Delta h}$, where $\Delta t_k = t_{k+1} - t_k$. The quantity $q_{\Delta h}$ is the continuous-time white-noise torque intensity and is expressed in N^2-m^2-sec units. The $\Delta \boldsymbol{h}_{\mathrm{ink}}$ impulses represent the unmodeled effects of solar and albedo radiation pressure, aerodynamic drag, and thermal radiation pressure.

An integrated version of Euler's equation is used to avoid differentiation of the measured angular velocity time history $\boldsymbol{\omega}_k, \boldsymbol{\omega}_{k+1}, \ldots$. Trapezoidal integration yields a reasonable approximation if the rotation $0.5(\boldsymbol{\omega}_k + \boldsymbol{\omega}_{k+1})\Delta t_k$ has a magnitude much smaller than 1 and if Δt_k is a small fraction of the orbital period. Inertial coordinates are used because trapezoidal integration is exact in inertial coordinates if the gravity-gradient and magnetic torques are negligible.

Equation (1) can be generalized to include articulating appendages, such as solar arrays. The estimator needs articulation angle data in this case, and it can estimate corrections to a subset of the appendage's moment-of-inertia elements. If the spacecraft dynamics include significant flexibility or fuel slosh motions, then Eq. (1) will be a poor model, and this paper's estimation algorithm will not produce useful results.

All of the unknown parameters enter Eq. (1) linearly. This is obvious from Eq. (1) and from the gravity-gradient torque formula:

$$\boldsymbol{n}_{\mathrm{gg}}(A, G_{\mathrm{in}}, I_m) = \begin{bmatrix} \{G_{\mathrm{sc23}}(I_{m22} - I_{m33}) + (G_{\mathrm{sc33}} - G_{\mathrm{sc22}}) I_{m23} + G_{\mathrm{sc13}} I_{m12} - G_{\mathrm{sc12}} I_{m13}\} \\ \{G_{\mathrm{sc13}}(I_{m33} - I_{m11}) + (G_{\mathrm{sc11}} - G_{\mathrm{sc33}}) I_{m13} + G_{\mathrm{sc12}} I_{m23} - G_{\mathrm{sc23}} I_{m12}\} \\ \{G_{\mathrm{sc12}}(I_{m11} - I_{m22}) + (G_{\mathrm{sc22}} - G_{\mathrm{sc11}}) I_{m12} + G_{\mathrm{sc23}} I_{m13} - G_{\mathrm{sc13}} I_{m23}\} \end{bmatrix} \tag{3}$$

where $G_{\mathrm{sc}} = A G_{\mathrm{in}} A^T$ is the gravity-gradient tensor in spacecraft coordinates.

Measurement errors contribute to errors in Eq. (1). Suppose that the measured values, the true values, and the measurement errors for the angular rate vector, the nominal wheel angular momenta, the magnetic field vector, and the nominal magnetic torque rod dipole strengths are as follows:

$$\boldsymbol{\omega}_k = \boldsymbol{\omega}_{\mathrm{meask}} + \Delta \boldsymbol{\omega}_k \tag{4a}$$

$$\boldsymbol{h}_{\mathrm{wk}} = \boldsymbol{h}_{\mathrm{wmeask}} + \Delta \boldsymbol{h}_{\mathrm{wk}} \tag{4b}$$

$$\boldsymbol{b}_k = \boldsymbol{b}_{\mathrm{meask}} + \Delta \boldsymbol{b}_k \tag{4c}$$

$$\boldsymbol{m}_k = \boldsymbol{m}_{\mathrm{meask}} + \Delta \boldsymbol{m}_k \tag{4d}$$

The quantities on the left-hand sides of these equations are the unknown true values that appear in Eq. (1). The terms with the meas subscript on the right-hand sides are the known measured quantities, and the terms with the Δ prefix are the unknown measurement errors. The effects of star tracker measurement errors on A_k are not considered because they are normally insignificant.

The estimation problem formulation lumps all of the measurement errors into a single error vector and rescales them:

$$v_k = \left[\Delta \boldsymbol{\omega}_k^T / \sigma_\omega, \Delta \boldsymbol{h}_{\mathrm{wk}}^T / \sigma_{\mathrm{hw}}, \Delta \boldsymbol{b}_k^T / \sigma_b, \Delta \boldsymbol{m}_k^T / \sigma_m \right]^T \tag{5}$$

where σ_ω, σ_{hw}, σ_b, and σ_m are the per-axis measurement error standard deviations of the four measured vectors. The vector $\boldsymbol{v}_k$ is modeled as a zero-mean, identity-covariance, discrete-time, white-noise Gaussian random process, that is, $\boldsymbol{v}_k \sim N(0, I)$ and $E\{\boldsymbol{v}_j \boldsymbol{v}_k^T\} = 0$ if $j \neq k$. The dimension of $\boldsymbol{v}_k$ is $n_v = 6 + n_w + n_t$.

The definitions in Eqs. (2) and (4a–5) and the gravity-gradient torque formula in Eq. (3) can be used to rewrite Eq. (1) in the following generalized form:

$$\left[E_{0k+1} + \sum_{i=1}^{n_v} (\boldsymbol{v}_{k+1})_i E_{ik+1} \right] x = \left[D_{0k} + \sum_{i=1}^{n_v} (\boldsymbol{v}_k)_i D_{ik} \right] \boldsymbol{x} + \boldsymbol{\eta}_k \tag{6}$$

where the subscript i refers to the ith element of the vector in question and where $\boldsymbol{\eta}_k = \Delta \boldsymbol{h}_{\mathrm{ink}} / \sigma_{\Delta \mathrm{hk}}$ is the normalized angular impulse process noise vector. The $3 \times n_x$ matrices D_{0k}, D_{ik}, E_{0k+1}, and $E_{\mathrm{ik+1}}$ are computed from the measurements. Their formulas can be derived based on Eqs. (1–5). The formulas for several columns of

these matrices are given here as examples:

$$\text{column} \quad 2 \quad \text{of} \quad D_{0k} = \left(\frac{1}{\sigma_{\Delta \text{hk}}}\right) A_k^T \left\{ \begin{bmatrix} (\omega_{\text{meask}})_2 \\ (\omega_{\text{meask}})_1 \\ 0 \end{bmatrix} + \frac{\Delta t_k}{2} \begin{bmatrix} (G_{\text{sck}})_{13} \\ -(G_{\text{sck}})_{23} \\ (G_{\text{sck}})_{22} - (G_{\text{sck}})_{11} \end{bmatrix} \right\} \tag{7a}$$

$$\text{columns} \quad 7\text{–}9 \quad \text{of} \quad D_{4k} = \left(\frac{1}{\sigma_{\Delta \text{hk}}}\right) A_k^T \sigma_{\text{hw}} \tag{7b}$$

$$\text{columns} \quad 11 \quad \text{of} \quad E_{0k+1} = \left(\frac{1}{\sigma_{\Delta \text{hk}}}\right) A_{k+1}^T \begin{bmatrix} 0 \\ (\boldsymbol{h}_{\text{wmeask+1}})_2 \\ 0 \end{bmatrix} \tag{7c}$$

$$\text{columns} \quad 6 \quad \text{and} \quad 7 \quad \text{of} \quad E_{3k+1} = \left(\frac{1}{\sigma_{\Delta \text{hk}}}\right) A_{k+1}^T \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \sigma_\omega & 0 \end{bmatrix} \tag{7d}$$

where the subscript ij refers to the ijth element of the matrix in question. The remaining columns of the Eq. (6) matrices are straightforward to compute. Their formulas have been omitted for the sake of brevity.

Equation (6) fails to be an exact representation of Eq. (1) on one point. It does not include the products of magnetic field measurement error vectors $\Delta \boldsymbol{b}$ and magnetic torque rod dipole moment measurement errors $\Delta \boldsymbol{m}$. These error product terms have been neglected because their inclusion would complicate estimation equation (6) and because they are small relative to the corresponding measured values, which makes their product negligible.

A lumped total least-squares parameter estimation problem can be defined by considering Eq. (6) for the N different intervals, $k = 0, 1, 2, \ldots, N-1$. Suppose that one lumps the measurement errors into one large vector and the torque errors into another large vector:

$$\boldsymbol{v}_{\text{big}} = \left[\boldsymbol{v}_0^T, \boldsymbol{v}_1^T, \boldsymbol{v}_2^T, \ldots, \boldsymbol{v}_N^T\right]^T \tag{8a}$$

$$\boldsymbol{\eta}_{\text{big}} = \left[\boldsymbol{\eta}_0^T, \boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T, \ldots, \boldsymbol{\eta}_{N-1}^T\right]^T \tag{8b}$$

The length of the large measurement error vector is $n_{\text{vbig}} = n_v(N+1)$, and the length of the large torque error vector is $n_{\eta\text{big}} = 3N$. One can lump the N copies of Eq. (6) into the following large system of measurement equations:

$$\left[H_0 + \sum_{j=1}^{n_{\text{vbig}}} (\boldsymbol{v}_{\text{big}})_j H_j\right] \boldsymbol{x} = \boldsymbol{\eta}_{\text{big}} \tag{9}$$

where the large H matrices are

$$H_0 = \begin{bmatrix} (E_{01} - D_{00}) \\ (E_{02} - D_{01}) \\ (E_{03} - D_{02}) \\ \vdots \\ (E_{0N} - D_{0N-1}) \end{bmatrix}, \qquad H_1 = \begin{bmatrix} -D_{10} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$H_2 = \begin{bmatrix} -D_{20} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad H_{\text{nv}} = \begin{bmatrix} -D_{\text{nv0}} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$H_{\text{nv+1}} = \begin{bmatrix} E_{11} \\ -D_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad H_{\text{nv+2}} = \begin{bmatrix} E_{21} \\ -D_{21} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\ldots, \quad H_{2\text{nv}} = \begin{bmatrix} E_{\text{nv1}} \\ -D_{\text{nv1}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad H_{\text{nvbig}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ E_{\text{nvN}} \end{bmatrix} \tag{10}$$

Lumped measurement equation. (9) and the statistical models of the random error vectors $\boldsymbol{v}_{\text{big}}$ and $\boldsymbol{\eta}_{\text{big}}$ can be used to define the following nonlinear least-squares cost function:

$$J(\boldsymbol{x}, \boldsymbol{v}_{\text{big}}) = \frac{1}{2} \boldsymbol{x}^T \left[H_0 + \sum_{j=1}^{n_{\text{vbig}}} (\boldsymbol{v}_{\text{big}})_j H_j\right]^T \left[H_0 + \sum_{j=1}^{n_{\text{vbig}}} (\boldsymbol{v}_{\text{big}})_j H_j\right] \boldsymbol{x} + \frac{1}{2} \boldsymbol{v}_{\text{big}}^T \boldsymbol{v}_{\text{big}} \tag{11}$$

This paper's algorithm estimates $\boldsymbol{x}$ and $\boldsymbol{v}_{\text{big}}$ by finding the values that minimize this cost function subject to a constraint. This cost function equals a constant plus the negative log of the probability density function for the error vectors $\boldsymbol{v}_{\text{big}}$ and $\boldsymbol{\eta}_{\text{big}}$ conditioned on the parameters given in $\boldsymbol{x}$. The minimizing $\boldsymbol{x}$ value is a maximum likelihood estimate, and the minimizing $\boldsymbol{v}_{\text{big}}$ value is a maximum a posteriori estimate.

Extra information must get included in the estimation problem for it to generate a nontrivial solution. The unconstrained global minimum of $J(\boldsymbol{x}, \boldsymbol{v}_{\text{big}})$ occurs at $\boldsymbol{x} = 0, \boldsymbol{v}_{\text{big}} = 0$ because Eq. (9) is homogeneous. The physical reason for this is that all of the parameters in the $\boldsymbol{x}$ vector can be scaled up or down without affecting the dynamic response of the system. A large spacecraft and a small spacecraft will undergo the same dynamic response if the ratios of the parameters in the $\boldsymbol{x}$ vector are preserved.

This paper's problem definition adds scaling information by including a single quadratic constraint on $\boldsymbol{x}$:

$$0 = 1 - (L\boldsymbol{x})^T (L\boldsymbol{x}) \tag{12}$$

where L is a nonzero matrix with n_x columns and at least one row. It is necessary to choose L wisely in order to get reasonable estimates from the solution of the estimation problem. The best choice of L is problem dependent. Most of the problems solved in this paper use an L matrix that has $3n_w$ rows. All of its columns are zero except for columns 7 through $(6 + 3n_w)$. These columns are set equal to the identity matrix divided by $\sqrt{(n_w)}$. This L definition constrains the square of the Frobenius norm of C_{rw} to equal n_w, which forces the average of the squares of the reaction-wheels' recalibrated scale factors to equal one. If the spacecraft does not have reaction wheels, then a similar constraint can be applied to the magnetic torque rod scale-factor/alignment matrix C_{mt} or to the moment-of-inertia matrix I_m. In the latter case, L should be scaled so that the a priori I_m estimate satisfies Eq. (12).

Such L choices act equally for all three spacecraft axes. A constraint based on a single axis should be avoided because it can yield poor estimates when there is only weak dynamic coupling between the axes. In this situation, the axis with the constraint has a reasonable scaling of its parameter estimates, but the other two axes have scalings that are too small. The estimator reduces the other two axes' contributions to $J(\boldsymbol{x}, \boldsymbol{v}_{\text{big}})$ simply by scaling down the $\boldsymbol{x}$ components associated with these axes. The resulting estimates have nonsensical relative scalings between the axes, which can produce moment-of-inertia estimates that fail to obey physical constraints. If the L matrix weights all three axes equally, however, then the quadratic

constraint will tend to balance the relative scaling between the axes in a reasonable manner.

The estimation problem can be augmented to incorporate a priori parameter information. This information takes the form of an a priori square-root information equation[17]:

$$R_{\text{ap}}\boldsymbol{x} + \boldsymbol{z}_{\text{ap}} = \boldsymbol{\eta}_{\text{ap}} \tag{13}$$

where R_{ap} and $\boldsymbol{z}_{\text{ap}}$ are the a priori square-root information matrix and vector, respectively, and where $\boldsymbol{\eta}_{\text{ap}}$ is a zero-mean, identity-covariance Gaussian random vector $\boldsymbol{\eta}_{\text{ap}} \sim N(0, I)$.

The final form of the estimation problem, after the incorporation of the quadratic constraint and the a priori information is as follows.

Find:

$$\boldsymbol{x}, \qquad \boldsymbol{v}_{\text{big}} \tag{14a}$$

To minimize:

$$J(\boldsymbol{x}, \boldsymbol{v}_{\text{big}}) = \frac{1}{2}\boldsymbol{x}^T\left[H_0 + \sum_{j=1}^{n_{\text{vbig}}}(\boldsymbol{v}_{\text{big}})_j H_j\right]^T\left[H_0 + \sum_{j=1}^{n_{\text{vbig}}}(\boldsymbol{v}_{\text{big}})_j H_j\right]\boldsymbol{x}$$

$$+ \frac{1}{2}[R_{\text{ap}}\boldsymbol{x} + \boldsymbol{z}_{\text{ap}}]^T[R_{\text{ap}}\boldsymbol{x} + \boldsymbol{z}_{\text{ap}}] + \frac{1}{2}\boldsymbol{v}_{\text{big}}^T\boldsymbol{v}_{\text{big}} \tag{14b}$$

Subject to:

$$0 = 1 - (L\boldsymbol{x})^T(L\boldsymbol{x}) \tag{14c}$$

The a priori information equation can be used to enforce an additional constraint to make the problem observable. In this situation, the constraining row of R_{ap} and the corresponding row of $\boldsymbol{z}_{\text{ap}}$ will take on very large values, normally orders of magnitude larger than the corresponding rows of the H matrices. This large row acts as a soft constraint on the solution. Note, however, that one must use numerically robust square-root-based techniques, as are developed in Secs. III and IV, if one wants to mitigate the potential for adverse affects from the resulting wide variations in the magnitudes of different problem matrices.

One must be careful to apply an extra linear constraint in the form of an a priori information equation only where appropriate. One might try to substitute a linear overall scaling constraint for the quadratic scaling constraint in Eq. (14c). This might not be a good idea because a linear scaling constraint can bias the relative scalings between the different axes towards equality even when the true scalings are not equal.

An appropriate situation for adding a linear constraint occurs when one is also estimating the moment-of-inertia matrix of a rotating flexible appendage. Suppose that the appendage's center of mass lies on the rotation axis, that $[\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3]$ is an orthonormal triad in the main spacecraft body's coordinate system, and that $\boldsymbol{q}_3$ is directed along the appendage articulation axis. It is possible to make simultaneous unobservable changes to the moment-of-inertia matrix of the main spacecraft body I_m and the moment-of-inertia matrix of the appendage I_a. Given any scalar β, the unobservable change adds $\beta(\boldsymbol{q}_1\boldsymbol{q}_1^T + \boldsymbol{q}_2\boldsymbol{q}_2^T)$ to I_m while subtracting the same quantity from I_a. The following a priori linear constraint removes this ambiguity:

$$\left(\boldsymbol{q}_1^T I_m \boldsymbol{q}_1 + \boldsymbol{q}_2^T I_m \boldsymbol{q}_2\right) - \gamma\left(\boldsymbol{q}_1^T I_a \boldsymbol{q}_1 + \boldsymbol{q}_2^T I_a \boldsymbol{q}_2\right) = 0 \tag{15}$$

where γ is the positive number that makes the a priori estimates of I_m and I_a satisfy this constraint.

Summarizing the use of hard and soft constraints, at least one constraint must be added to the problem in order to provide a priori information about the overall dynamic scaling of the spacecraft, which is unobservable. Normally, the a priori scaling is best specified through a single quadratic constraint that acts equally about all three axes. It might be necessary to include additional soft constraints that provide a priori information about unobservable subspaces of the parameter space if such unobservable subspaces exist for the particular problem. It is acceptable to do this using linear constraints. The best choice of constraints is an art rather than a science.

Given a reasonable constraint choice and enough data from sufficiently rich attitude maneuvers, the resulting parameter estimates will be the only self-consistent parameters that will describe the observed spacecraft motion. This will be true because the total least-squares estimation problem will have a unique global minimum.

The explicit inclusion of the multiplicative measurement errors of Eqs. (4a–4d) is an important feature that distinguishes the present problem formulation from that of Refs. 8–14. The preceding methods assumed that quantities such as the wheel angular momenta were measured perfectly. This can be a poor assumption, as is demonstrated in Sec. V. They also ignored the error correlations that are caused by the fact that each measurement error affects Eq. (1) for two neighboring sample intervals.

III. Attitude Parameter Estimation Algorithm

A. Algorithm Overview

The solution algorithm for the problem in Eqs. (14a–14c) exploits its bilinear structure. For a fixed $\boldsymbol{x}$, the cost function is a linear least-squares cost in the unconstrained variable $\boldsymbol{v}_{\text{big}}$. Matrix factorizations can be used in an inner-loop optimization to solve for the global minimizer $\boldsymbol{v}_{\text{big}}$ as a function of $\boldsymbol{x}$. Call this solution $\boldsymbol{v}_{\text{bigopt}}(\boldsymbol{x})$. This solution can be substituted into the original cost function in Eq. (14b) to define a new reduced-order estimation problem.

Find:

$$\boldsymbol{x} \tag{16a}$$

To minimize:

$$J_{\text{ro}}(\boldsymbol{x}) = J[\boldsymbol{x}, \boldsymbol{v}_{\text{bigopt}}(\boldsymbol{x})] \tag{16b}$$

Subject to:

$$0 = 1 - (L\boldsymbol{x})^T(L\boldsymbol{x}) \tag{16c}$$

This reduced-order problem is then solved numerically by using the iterative Newton method in an outer optimization loop.

Newton's method applied to problem (16a–16c) starts with a guess $\boldsymbol{x}_0$ that satisfies the quadratic constraint in Eq. (16c). Next, it defines a nonlinear function $\boldsymbol{x}_{\text{curve}}(\boldsymbol{x}_0 + \Delta\boldsymbol{x})$ that is guaranteed to remain on the quadratic constraint as $\Delta\boldsymbol{x}$ varies from zero and that takes on the initial value $\boldsymbol{x}_{\text{curve}}(\boldsymbol{x}_0) = \boldsymbol{x}_0$. It uses this function to define the quadratic cost function approximation:

$$J_{\text{newt}}(\Delta\boldsymbol{x}, \Delta\boldsymbol{v}_{\text{big}}) \cong J[\boldsymbol{x}_{\text{curve}}(\boldsymbol{x}_0 + \Delta\boldsymbol{x}), \boldsymbol{v}_{\text{bigopt}}(\boldsymbol{x}_0) + \Delta\boldsymbol{v}_{\text{big}}]$$

$$= J[\boldsymbol{x}_0, \boldsymbol{v}_{\text{bigopt}}(\boldsymbol{x}_0)] + \boldsymbol{g}_x^T\Delta\boldsymbol{x} + \left[\Delta\boldsymbol{x}^T, \Delta\boldsymbol{v}_{\text{big}}^T\right]$$

$$\times \begin{bmatrix} W_{xx} & W_{xv} \\ W_{xv}^T & W_{vv} \end{bmatrix}\begin{bmatrix} \Delta\boldsymbol{x} \\ \Delta\boldsymbol{v}_{\text{big}} \end{bmatrix} \tag{17}$$

where the matrices W_{xx}, W_{xv}, and W_{vv} are blocks of the Hessian matrix. This approximate cost function gets minimized subject to the linearized approximation of the constraint in Eq. (16c): $(L\boldsymbol{x}_0)^T L\Delta\boldsymbol{x} = 0$. The solution is $[\Delta\boldsymbol{x}_{\text{opt}}; \Delta\boldsymbol{v}_{\text{bigopt}}]$, and $\Delta\boldsymbol{x}_{\text{opt}}$ is the Newton step toward the solution of problem (16a–16c).

The Newton iteration finishes by performing a line search to approximately minimize the one-dimensional cost function $J_{1D}(\alpha) = J_{\text{ro}}[\boldsymbol{x}_{\text{curve}}(\boldsymbol{x}_0 + \alpha\Delta\boldsymbol{x}_{\text{opt}})]$. A step length of $\alpha = 1$ is tried in order to get the superlinear convergence of Newton's method when $\boldsymbol{x}_0$ is near the solution, but α gets decreased from one if necessary when far from the solution in order to guarantee a cost decrease: $J_{1D}(\alpha_{\text{opt}}) < J_{1D}(0)$. This forced decrease of the cost safeguards the Newton algorithm so that it is guaranteed to converge at least to a local minimum. Enforcement of the cost-decrease constraint can be accomplished by an appropriate search process.[18] The improved guess of the solution becomes the next Newton iterate, $\boldsymbol{x}_1 = \boldsymbol{x}_{\text{curve}}(\boldsymbol{x}_0 + \alpha_{\text{opt}}\Delta\boldsymbol{x}_{\text{opt}})$, and the process can be repeated until it converges to a local minimum.

It is possible, in theory, to reverse the roles of $\boldsymbol{x}$ and $\boldsymbol{v}_{\text{big}}$ in the algorithm, but this approach works poorly in practice. For a fixed guess of $\boldsymbol{v}_{\text{big}}$, one can solve exactly for the $\boldsymbol{x}$ that globally minimizes

the cost function in Eq. (14b) subject to the constraint in Eq. (14c). The global minimization can be performed by using an algorithm found in Ref. 19, or if $z_{ap} = 0$, by using two singular-value decompositions (SVDs) and several QR factorizations. This reversed method has been tried, but it does not work well. One problem with this approach is the difficulty of finding a reasonable initial guess of $\boldsymbol{v}_{big}$. Seemingly reasonable values, such as $\boldsymbol{v}_{big} = 0$, can produce very strange results when the measurement error standard deviations are large. Another problem is that the reduced-order cost function, $J_{ro}(\boldsymbol{v}_{big}) = J[\boldsymbol{x}_{opt}(\boldsymbol{v}_{big}), \boldsymbol{v}_{big}]$, can have strong nonlinearities that retard convergence.

Note that there exist closed-form global solutions to certain total least-squares problems,[15] but such techniques cannot handle the special form of the measurement error model in the problem of Eqs. (14a–14c). In other words, they are not applicable to the attitude dynamics parameter estimation problem because their measurement error models are not general enough. The philosophy taken here is that it is better to solve the desired problem using a less general, less robust algorithm than to use a more general, more robust algorithm that forces unrealistic modeling assumptions. The exact solution of the wrong problem is not very useful!

B. Inner-Loop Optimization of v_{big}

The inner-loop optimization of $\boldsymbol{v}_{big}$ for $\boldsymbol{x}$ fixed at $\boldsymbol{x}_0$ minimizes the linear least-squares cost function

$$J_{inner}(\boldsymbol{v}_{big}) = \frac{1}{2}\left\{\begin{bmatrix}\boldsymbol{e}_v\\0\end{bmatrix} + \begin{bmatrix}H_v\\I\end{bmatrix}\boldsymbol{v}_{big}\right\}^T\left\{\begin{bmatrix}\boldsymbol{e}_v\\0\end{bmatrix} + \begin{bmatrix}H_v\\I\end{bmatrix}\boldsymbol{v}_{big}\right\} \tag{18}$$

which is equivalent to the cost in Eq. (14b) without the a priori terms if $\boldsymbol{e}_v$ and H_v are

$$\boldsymbol{e}_v = H_0\boldsymbol{x}_0 \qquad \text{and}$$

$$H_v = [(H_1\boldsymbol{x}_0), (H_2\boldsymbol{x}_0), (H_3\boldsymbol{x}_0), \ldots, (H_{nvbig}\boldsymbol{x}_0)] \tag{19}$$

The linear least-squares optimization problem in Eq. (18) gets solved by using standard QR factorization techniques.[18] An orthogonal matrix Q_v and a square, upper-triangular, nonsingular matrix R_{vv} get computed via QR factorization to satisfy

$$Q_v\begin{bmatrix}R_{vv}\\0\end{bmatrix} = \begin{bmatrix}H_v\\I\end{bmatrix} \tag{20}$$

and the optimal $\boldsymbol{v}_{big}$ gets computed as

$$\boldsymbol{v}_{bigopt}(\boldsymbol{x}_0) = -\begin{bmatrix}R_{vv}^{-1} & 0\end{bmatrix}Q_v^T\begin{bmatrix}\boldsymbol{e}_v\\0\end{bmatrix} \tag{21}$$

The matrices involved in this calculation are large and sparse and have a dynamic programming structure. Section IV develops equivalent calculations that use an efficient recursion which exploits this structure.

C. Enforcement of the Curved Constraint

The purpose of the function $\boldsymbol{x}_{curve}(\boldsymbol{x})$ is to generate a vector that satisfies the constraint in Eq. (16c) while being close to $\boldsymbol{x}$. It must have the property that $\boldsymbol{x}_{curve}(\boldsymbol{x}_0) = \boldsymbol{x}_0$ if $\boldsymbol{x}_0$ satisfies the constraint, and the difference $\boldsymbol{x}_{curve}(\boldsymbol{x}) - \boldsymbol{x}$ must be small if $\boldsymbol{x}$ is near the constraint. These properties enable the approximate cost function in Eq. (17) to accurately model changes that occur in the true cost during line searches that follow the curved constraint in Eq. (16c).

A suitable function $\boldsymbol{x}_{curve}(\boldsymbol{x})$ has been developed based on the SVD of L:

$$[U_{L1} \quad U_{L2}]\begin{bmatrix}\Sigma_L & 0\\0 & 0\end{bmatrix}\begin{bmatrix}V_{L1}^T\\V_{L2}^T\end{bmatrix} = L \tag{22}$$

where the matrices U_{L1}, U_{L2}, V_{L1}, and V_{L2} are sets of columns of the orthogonal factors of the SVD and where Σ_L is the positive-definite diagonal matrix that contains the nonzero singular values of L. The matrix V_{L1}^T projects $\boldsymbol{x}$ onto a subspace in which every component affects the quadratic constraint, and the matrix V_{L2}^T projects $\boldsymbol{x}$ onto the complementary subspace in which none of the components affect the constraint. Given these matrices, a suitable $\boldsymbol{x}_{curve}(\boldsymbol{x})$ function is

$$\boldsymbol{x}_{curve}(\boldsymbol{x}) = \left[\frac{V_{L1}V_{L1}^T}{\sqrt{(L\boldsymbol{x})^T(L\boldsymbol{x})}} + V_{L2}V_{L2}^T\right]\boldsymbol{x} \tag{23}$$

This function splits $\boldsymbol{x}$ into its constrained and unconstrained components and rescales the constrained component by $1/\sqrt{[(L\boldsymbol{x})^T(L\boldsymbol{x})]}$ to force satisfaction of the constraint.

D. Quadratic Problem and Solution for Newton Increment to x

The following quadratic approximate cost function is used for determining the Newton increment to $\boldsymbol{x}$:

$$J_{newt}(\Delta\boldsymbol{x}, \Delta\boldsymbol{v}_{big}) = \frac{1}{2}\left\{\begin{bmatrix}\boldsymbol{e}\\\boldsymbol{v}_{bigopt}\\\boldsymbol{e}_{ap}\end{bmatrix} + \begin{bmatrix}H_v\\I\\0\end{bmatrix}\Delta\boldsymbol{v}_{big} + \begin{bmatrix}H_x\\0\\R_{ap}\end{bmatrix}\Delta\boldsymbol{x}\right\}^T$$

$$\times\left\{\begin{bmatrix}\boldsymbol{e}\\\boldsymbol{v}_{bigopt}\\\boldsymbol{e}_{ap}\end{bmatrix} + \begin{bmatrix}H_v\\I\\0\end{bmatrix}\Delta\boldsymbol{v}_{big} + \begin{bmatrix}H_x\\0\\R_{ap}\end{bmatrix}\Delta\boldsymbol{x}\right\}$$

$$+ \Delta\boldsymbol{x}^T B\,\Delta\boldsymbol{v}_{big} - \frac{1}{2}\Delta\boldsymbol{x}^T\{\lambda L^T L\}\Delta\boldsymbol{x} \tag{24}$$

where $\boldsymbol{e} = H_x\boldsymbol{x}_0$, $\boldsymbol{e}_{ap} = R_{ap}\boldsymbol{x}_0 + \boldsymbol{z}_{ap}$, $H_x = H_0 + (\boldsymbol{v}_{bigopt})_1 H_1 + (\boldsymbol{v}_{bigopt})_2 H_2 + \cdots + (\boldsymbol{v}_{bigopt})_{nvbig} H_{nvbig}$, and $B = [(H_1^T\boldsymbol{e}), (H_2^T\boldsymbol{e}), \ldots, (H_{nvbig}^T\boldsymbol{e})]$. Note, $\boldsymbol{e}$ is the error in the concatenated set of Euler equations, and $\boldsymbol{e}_{ap}$ is the error in the a priori information equation. Both errors are evaluated at the current solution guess $[\boldsymbol{x}_0; \boldsymbol{v}_{bigopt}(\boldsymbol{x}_0)]$.

The quantity λ in Eq. (24) is like a Lagrange multiplier. It predicts the cost effects of the curvature in $\boldsymbol{x}_{curve}(\boldsymbol{x}_0 + \Delta\boldsymbol{x})$ if $\boldsymbol{x}_0$ satisfies the quadratic constraint in Eq. (16c) and if $\Delta\boldsymbol{x}$ satisfies the linearized approximation of the constraint: $\boldsymbol{a}^T\Delta\boldsymbol{x} = 0$, where $\boldsymbol{a} = L^T L\boldsymbol{x}_0$. The proper value of this multiplier depends on the actual form of the function $\boldsymbol{x}_{curve}(\boldsymbol{x})$. For the function defined in Eq. (23), the proper value is

$$\lambda = \boldsymbol{x}_0^T V_{L1}V_{L1}^T\left[H_x^T\boldsymbol{e} + R_{ap}^T\boldsymbol{e}_{ap}\right] \tag{25}$$

The solution for the minimum of Eq. (24) subject to the constraint $\boldsymbol{a}^T\Delta\boldsymbol{x} = 0$ starts with a transformation of the first cost term. Suppose that one performs the following QR factorization and associated transformation:

$$Q_{big}\begin{bmatrix}R_{vv} & R_{vx}\\0 & R_{xx}\\0 & 0\end{bmatrix} = \begin{bmatrix}H_v & H_x\\I & 0\\0 & R_{ap}\end{bmatrix} \text{ and } \begin{bmatrix}0\\\boldsymbol{z}_x\\\boldsymbol{z}_r\end{bmatrix} = Q_{big}^T\begin{bmatrix}\boldsymbol{e}\\\boldsymbol{v}_{bigopt}\\\boldsymbol{e}_{ap}\end{bmatrix} \tag{26}$$

where Q_{big} is an orthogonal transformation, R_{vv}, R_{vx}, and R_{xx} are block matrices, with R_{vv} and R_{xx} square and upper-triangular and R_{vv} nonsingular; and $\boldsymbol{z}_x$ and $\boldsymbol{z}_r$ are vectors. The upper zero in the large vector that contains $\boldsymbol{z}_x$ and $\boldsymbol{z}_r$ is there by virtue of the fact that the cost function in Eq. (24) is an expansion about an optimized value of $\boldsymbol{v}_{big}$. The Q_{big} matrix can be used to transform each of the factors in the first cost term of Eq. (24) so that the block matrices and vectors on the left-hand sides of Eqs. (26) replace those on the right. After performing this transformation, the minimizing $\Delta\boldsymbol{v}_{big}$ can be determined by solving the necessary condition that results from setting the partial derivative of $J_{newt}(\Delta\boldsymbol{x}, \Delta\boldsymbol{v}_{big})$ with respect to $\Delta\boldsymbol{v}_{big}$ equal to zero. The result is

$$\Delta\boldsymbol{v}_{big} = -R_{vv}^{-1}\left[R_{vx} + \left(BR_{vv}^{-1}\right)^T\right]\Delta\boldsymbol{x} \tag{27}$$

Note that the matrices in Eqs. (26) and (27) are large, sparse, and structured. Section IV presents techniques that exploit the structure

of these matrices to carry out the calculations of Eqs. (26) and (27) in an efficient recursive manner.

The optimal solution for $\Delta \boldsymbol{v}_{\text{big}}$ can be used to eliminate it from the transformed version of the cost function in Eq. (24). This leaves a linearly constrained reduced-order quadratic optimization problem whose solution is the Newton search direction.

Find:

$$\Delta \boldsymbol{x} \tag{28a}$$

To minimize:

$$J_{\text{newtro}}(\Delta \boldsymbol{x}) = \tfrac{1}{2}\Delta \boldsymbol{x}^T \left\{ R_{xx}^T R_{xx} - \left(B R_{vv}^{-1}\right)\left(B R_{vv}^{-1}\right)^T - \left(B R_{vv}^{-1}\right) R_{vx} - R_{vx}^T \left(B R_{vv}^{-1}\right)^T - \lambda L^T L \right\} \Delta \boldsymbol{x} + \left(z_x^T R_{xx}\right) \Delta \boldsymbol{x} + \tfrac{1}{2}\left(z_x^T z_x + z_r^T z_r\right) \tag{28b}$$

Subject to:

$$0 = \boldsymbol{a}^T \Delta \boldsymbol{x} = (L^T L \boldsymbol{x}_0)^T \Delta \boldsymbol{x} \tag{28c}$$

The problem in Eqs. (28a–28c) can be solved using QR factorizations and Cholesky factorizations. The first QR factorization is used to determine the null space of the constraint:

$$[\boldsymbol{q}_{a1} \quad Q_{a2}] \begin{bmatrix} r_a \\ 0 \end{bmatrix} = \boldsymbol{a} \tag{29}$$

where $\boldsymbol{q}_{a1}$ is the first column of the resulting orthogonal matrix, Q_{a2} constitutes the last $n_x - 1$ columns, and r_a is a nonzero scalar. This factorization can be used to transform the perturbation vector $\Delta \boldsymbol{x}$ into a constrained component $\Delta \boldsymbol{x}_c$ and an unconstrained component $\Delta \boldsymbol{x}_u$:

$$\begin{bmatrix} \Delta \boldsymbol{x}_c \\ \Delta \boldsymbol{x}_u \end{bmatrix} = \begin{bmatrix} \boldsymbol{q}_{a1}^T \\ Q_{a2}^T \end{bmatrix} \Delta \boldsymbol{x} \tag{30}$$

One next uses Q_{a2} to project the square-root weighting matrix R_{xx} onto the unconstrained $\Delta \boldsymbol{x}_u$ subspace, and one QR factorizes the result to retrieve a square upper-triangular matrix:

$$Q_u \begin{bmatrix} R_{uu} \\ 0 \end{bmatrix} = R_{xx} Q_{a2} \tag{31}$$

where Q_u is an orthogonal matrix and R_{uu} is a square, upper-triangular, nonsingular matrix of dimension $n_x - 1$.

Next, one defines the change of coordinates $\Delta \boldsymbol{z}_u = R_{uu} \Delta \boldsymbol{x}_u$. The following unconstrained $\Delta \boldsymbol{z}_u$ optimization problem is then equivalent to the estimation problem in Eqs. (28a–28c).

Find:

$$\Delta \boldsymbol{z}_u \tag{32a}$$

To minimize:

$$J_{\text{newtz}}(\Delta \boldsymbol{z}_u) = \tfrac{1}{2}\Delta \boldsymbol{z}_u^T \left\{ \rho I - \Delta W_{zz} \right\} \Delta \boldsymbol{z}_u + \boldsymbol{g}_z^T \Delta \boldsymbol{z}_u + \tfrac{1}{2}\left(z_x^T z_x + z_r^T z_r\right) \tag{32b}$$

where

$$\Delta W_{zz} = \left(R_{uu}^{-1}\right)^T Q_{a2}^T \left\{ \left(B R_{vv}^{-1}\right)\left(B R_{vv}^{-1}\right)^T + \left(B R_{vv}^{-1}\right) R_{vx} + R_{vx}^T \left(B R_{vv}^{-1}\right)^T + \lambda L^T L \right\} Q_{a2} R_{uu}^{-1} \tag{33a}$$

$$\boldsymbol{g}_z = \left(R_{uu}^{-1}\right)^T Q_{a2}^T R_{xx}^T z_x \tag{33b}$$

and where the positive scalar ρ is normally set equal to one, and ρ will get increased above one if necessary in order to ensure that the cost Hessian matrix $\rho I - \Delta W_{zz}$ is positive definite. This modification option is part of the guarding technique that ensures the convergence of Newton's method to a local minimum. Its effect on $\Delta \boldsymbol{x}$ guarantees that there exists a positive α, which yields a decrease of the line-search cost function $J_{\text{ro}}[\boldsymbol{x}_{\text{curve}}(\boldsymbol{x}_0 + \alpha \Delta \boldsymbol{x})]$.

The Cholesky factorization of the Hessian produces R_{zz} such that

$$R_{zz}^T R_{zz} = \rho I - \Delta W_{zz} \tag{34}$$

where R_{zz} is a square, upper-triangular, nonsingular matrix of dimension $n_x - 1$. The Cholesky factorization process can be used to monitor the positive definiteness of the Hessian. One begins using the value $\rho = 1$. If the Cholesky factorization process fails because the matrix on the right-hand side of Eq. (34) is not positive definite, then ρ gets increased, and the Cholesky factorization gets reevaluated. A simple geometric progression of the form $\rho_{\text{new}} = 1.05\rho$ normally yields a positive definite Hessian matrix after a few Cholesky factorizations. The matrix dimensions involved are not very large. Therefore, it is acceptable to determine a reasonable value of ρ via this brute-force heuristic iteration. One might try a larger increase factor if one finds that too many iterations are required to determine an acceptable ρ value.

The final problem transformation from $\Delta \boldsymbol{x}_u$ to $\Delta \boldsymbol{z}_u$ is motivated by numerical considerations. One should avoid the squaring of matrices because squaring amplifies the adverse effects of poor matrix conditioning. The transformation to $\Delta \boldsymbol{z}_u$ allows one to avoid squaring the matrix R_{uu} before computing the Cholesky factorization of the Hessian. For a small residuals estimation problem, that is, for small $\boldsymbol{e}$, the magnitude of ΔW_{zz} will be small. In this case, $\rho = 1$ will yield a positive definite, well-conditioned Hessian matrix for the $\Delta \boldsymbol{z}_u$ problem, and the only squaring that occurs, the squaring in Eq. (33a), will have no adverse effects on the solution.

The matrix factorizations and problem transformations in Eqs. (29–34) can be used to solve the original problem in Eqs. (28a–28c). The solution is

$$\Delta \boldsymbol{x} = -Q_{a2} R_{uu}^{-1} R_{zz}^{-1} \left(R_{zz}^{-1}\right)^T \boldsymbol{g}_z \tag{35}$$

These matrix factorizations also provide a means of calculating the estimation error covariance matrix:

$$P_{xx} = Q_{a2} R_{uu}^{-1} R_{zz}^{-1} \left(Q_{a2} R_{uu}^{-1} R_{zz}^{-1}\right)^T \tag{36}$$

In reality, this is only an approximation of the Cramer–Rao lower bound of the estimation error covariance. This covariance matrix is only valid at the termination of the optimization process, when $\boldsymbol{x}_0$ is the final solution to the estimation problem and when $\rho = 1$. Note that P_{xx} is singular. All of its eigenvalues are positive except for one that is zero. It corresponds to the eigenvector in the direction $\boldsymbol{a} = L^T L \boldsymbol{x}_0$, which is the normal to the quadratic constraint in Eq. (16c). It makes sense for P_{xx} to have a zero eigenvalue in this direction because, according to the constraint, this component of $\boldsymbol{x}$ is known exactly.

E. Warning About Total Least-Squares Problems

It is well known that a total least-squares problem can fail to have sensible solutions.[15] If the attitude parameter estimation problem in Eqs. (14a–14c) does not have a sensible solution, then this fact will manifest itself in a failure of the guarded Newton method to converge. It will continue to decrease its cost at every iteration, but the usual algorithm termination criterion of $\|\Delta \boldsymbol{x}\| \to 0$ will not be achieved. If the algorithm fails to converge even after hundreds of iterations, then one should consider whether the problem is well posed. It is often possible to transform an ill-posed problem into a well-posed problem by adding a priori information through R_{ap} and z_{ap} or by decreasing the measurement error standard deviations σ_ω, σ_{hw}, σ_b, and σ_m.

IV. Recursive QR Matrix Factorizations and Inversions

The QR factorizations in Eqs. (20) and (26) and the inversions of R_{vv} in Eqs. (21) and (27) can be carried out in an efficient recursive manner. The following analysis concentrates on recursive calculations for the large matrix operations in Eqs. (26) and (27).

The methods for Eqs. (20) and (21) are almost identical and are discussed only in brief at the end of this section.

Suppose that one interchanges the order of the rows on the right-hand sides of the two equations in Eq. (26). Then the error equations whose squares form the first term on the right-hand side of Eq. (24) can be rewritten in the form

$$\begin{bmatrix} I & 0 & 0 & \dots & 0 & 0 & 0 \\ -D_{v0} & E_{v1} & 0 & \dots & 0 & 0 & F_0 \\ 0 & I & 0 & \dots & 0 & 0 & 0 \\ 0 & -D_{v1} & E_{v2} & \dots & 0 & 0 & F_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -D_{vN-1} & E_{vN} & F_{N-1} \\ 0 & 0 & 0 & \dots & 0 & I & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & R_{\text{ap}} \end{bmatrix} \times \begin{bmatrix} \Delta \boldsymbol{v}_0 \\ \Delta \boldsymbol{v}_1 \\ \Delta \boldsymbol{v}_2 \\ \vdots \\ \Delta \boldsymbol{v}_{N-1} \\ \Delta \boldsymbol{v}_N \\ \Delta \boldsymbol{x} \end{bmatrix} + \begin{bmatrix} \boldsymbol{v}_0 \\ \boldsymbol{e}_0 \\ \boldsymbol{v}_1 \\ \boldsymbol{e}_1 \\ \vdots \\ \boldsymbol{e}_{N-1} \\ \boldsymbol{v}_N \\ \boldsymbol{e}_{\text{ap}} \end{bmatrix} = 0 \tag{37}$$

where, based on Eq. (6) and associated definitions,

$$D_{vk} = [D_{1k}\boldsymbol{x}_0,\, D_{2k}\boldsymbol{x}_0,\, \dots,\, D_{nvk}\boldsymbol{x}_0] \qquad \text{for} \qquad k = 0, \dots, N-1 \tag{38a}$$

$$E_{vk} = [E_{1k}\boldsymbol{x}_0,\, E_{2k}\boldsymbol{x}_0,\, \dots,\, E_{nvk}\boldsymbol{x}_0] \qquad \text{for} \qquad k = 1, \dots, N \tag{38b}$$

$$F_k = \left[E_{0k+1} + \sum_{i=1}^{n_v} (\boldsymbol{v}_{k+1})_i E_{ik+1} \right] - \left[D_{0k} + \sum_{i=1}^{n_v} (\boldsymbol{v}_k)_i D_{ik} \right]$$

$$\boldsymbol{e}_k = F_k \boldsymbol{x}_0 \qquad \text{for} \qquad k = 0, \dots, N-1 \tag{38c}$$

An iterative QR factorization is then used to orthogonally transform Eq. (37) into the following form:

$$\begin{bmatrix} R_{00} & R_{01} & 0 & \dots & 0 & R_{0x} \\ 0 & R_{11} & R_{12} & \dots & 0 & R_{1x} \\ 0 & 0 & R_{22} & \dots & 0 & R_{2x} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & R_{NN} & R_{Nx} \\ 0 & 0 & 0 & \dots & 0 & R_{xx} \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta \boldsymbol{v}_0 \\ \Delta \boldsymbol{v}_1 \\ \Delta \boldsymbol{v}_2 \\ \vdots \\ \Delta \boldsymbol{v}_{N-1} \\ \Delta \boldsymbol{v}_N \\ \Delta \boldsymbol{x} \end{bmatrix} + \begin{bmatrix} \boldsymbol{z}_0 \\ \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ \vdots \\ \boldsymbol{z}_N \\ \boldsymbol{z}_x \\ \boldsymbol{z}_r \end{bmatrix} = 0 \tag{39}$$

Each R_{kk}, R_{kk+1}, and R_{kx} is a matrix block, with R_{kk} square, upper-triangular, and nonsingular, and each $\boldsymbol{z}_k$ is a vector. The matrix R_{xx} and the vectors $\boldsymbol{z}_x$ and $\boldsymbol{z}_r$ are exactly as defined in Eq. (26). The upper right-hand blocks involving R_{kk} for $k = 0, \dots, N$ and R_{kk+1} for $k = 0, \dots, N-1$ constitute the R_{vv} matrix of Eq. (26). The far-left blocks involving R_{kx} for $k = 0, \dots, N$ constitute the R_{vx} matrix of Eq. (26). In accordance with Eq. (26), the vectors $\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_N$ will be zero because Eq. (26) is a linearization about optimized $\boldsymbol{v}_k$ values.

The transformation from Eq. (37) to Eq. (39) proceeds as follows. It temporarily stores intermediate results in three matrices, R_{t11k}, R_{t12k}, and R_{t22k}, and in two vectors, $\boldsymbol{z}_{t1k}$ and $\boldsymbol{z}_{t2k}$. At $k = 0$ these temporary matrices and vectors get initialized to be empty arrays. Their values at later stages get computed by the factorization process. The transformation starts by recursively performing the following QR factorization and transformation for stages $k = 0, \dots, N-1$:

$$Q_k \begin{bmatrix} R_{kk} & R_{kk+1} & R_{kx} \\ 0 & R_{t11k+1} & R_{t12k+1} \\ 0 & 0 & R_{t22k+1} \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} R_{t11k} & 0 & R_{t12k} \\ 0 & 0 & R_{t22k} \\ I & 0 & 0 \\ -D_{vk} & E_{vk+1} & F_k \end{bmatrix}$$

$$\begin{bmatrix} \boldsymbol{z}_k \\ \boldsymbol{z}_{t1k+1} \\ \boldsymbol{z}_{t2k+1} \\ \boldsymbol{z}_{rk} \end{bmatrix} = Q_k^T \begin{bmatrix} \boldsymbol{z}_{t1k} \\ \boldsymbol{z}_{t2k} \\ \boldsymbol{v}_k \\ \boldsymbol{e}_k \end{bmatrix} \tag{40}$$

At stage N the transformation takes the modified form

$$Q_N \begin{bmatrix} R_{NN} & R_{Nx} \\ 0 & R_{xx} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} R_{t11N} & R_{t12N} \\ 0 & R_{t22N} \\ I & 0 \\ 0 & R_{\text{ap}} \end{bmatrix}$$

$$\begin{bmatrix} \boldsymbol{z}_N \\ \boldsymbol{z}_x \\ \boldsymbol{z}_{rN} \end{bmatrix} = Q_N^T \begin{bmatrix} \boldsymbol{z}_{t1N} \\ \boldsymbol{z}_{t2N} \\ \boldsymbol{v}_N \\ \boldsymbol{e}_{\text{ap}} \end{bmatrix} \tag{41}$$

The vectors $\boldsymbol{z}_{r0}, \boldsymbol{z}_{r1}, \boldsymbol{z}_{r2}, \dots, \boldsymbol{z}_{rN}$ are components of the large residual error vector $\boldsymbol{z}_r$.

Given transformed Eq. (39), the two multiplications by R_{vv}^{-1} in Eq. (27) can be performed as follows. Suppose that the large matrix B in Eq. (27) is broken down into blocks $B = [B_0, B_1, B_2, \dots, B_N]$, and suppose that the matrix BR_{vv}^{-1} is also broken down into blocks $BR_{vv}^{-1} = [C_0, C_1, C_2, \dots, C_N]$. Then the blocks of this latter matrix can be calculated by using the following recursion:

$$C_0 = B_0 R_{00}^{-1}, \qquad C_k = [B_k - C_{k-1}R_{k-1k}]R_{kk}^{-1} \qquad \text{for} \qquad k = 1, \dots, N \tag{42}$$

Next, define the vector sequence $\Delta \boldsymbol{z}_{vk} = (R_{kx} + C_k^T)\Delta \boldsymbol{x}$ for $k = 0, \dots, N$. This vector sequence is used in the following backwards recursion in order to complete the evaluation of Eq. (27):

$$\Delta \boldsymbol{v}_N = -R_{NN}^{-1} \Delta \boldsymbol{z}_{vN}$$

$$\Delta \boldsymbol{v}_k = -R_{kk}^{-1}[\Delta \boldsymbol{z}_{vk} + R_{kk+1}\Delta \boldsymbol{v}_{k+1}]$$

$$\text{for} \qquad k = N-1, N-2, N-3, \dots, 0 \tag{43}$$

Note that the C_k values from Eq. (42) can also be used to develop efficient summations that compute the terms involving BR_{vv}^{-1} in Eq. (33a)'s formula for ΔW_{zz}.

The evaluation of Eqs. (20) and (21) proceeds in a similar manner. There is no need to include the columns associated with $\boldsymbol{x}$ in the equivalents of Eqs. (37) and (39–41). The $\boldsymbol{v}_k$ vectors in the equivalents of Eqs. (37), (40), and (41) are all set to zero, and the vectors $\boldsymbol{e}_k = (E_{0k+1} - D_{0k})\boldsymbol{x}_0$ in these equations because they are linearized about $\boldsymbol{v}_k = \boldsymbol{v}_{k+1} = 0$. This causes the vectors $\boldsymbol{z}_0, \boldsymbol{z}_1, \boldsymbol{z}_2, \dots, \boldsymbol{z}_N$ in the equivalent of Eq. (39) to be nonzero. They replace the $\Delta \boldsymbol{z}_{vk}$ vectors in the equivalent of Eq. (43), and the outputs of that equation's recursion are $\boldsymbol{v}_{0\text{opt}}(\boldsymbol{x}_0), \boldsymbol{v}_{1\text{opt}}(\boldsymbol{x}_0), \boldsymbol{v}_{2\text{opt}}(\boldsymbol{x}_0), \dots, \boldsymbol{v}_{N\text{opt}}(\boldsymbol{x}_0)$.

V. Results Using Data from the WMAP Spacecraft

A. Overview of Spacecraft and Data Sets

The new parameter estimation algorithm has been tested by using flight data from the WMAP spacecraft. WMAP was launched on 30 June 2001 and reached its final orbit near the L2 Lagrange

point of the sun/Earth–moon system in early October 2001 after using gravity assists from the moon.[7] The spacecraft carries two star trackers and a three-axis rate-gyro package for attitude sensing, and it uses three nearly orthogonal reaction wheels to control attitude. Dynamics parameters have been estimated by using data that were collected on 1 and 2 July 2001. Gravity-gradient and magnetic torques have been neglected because the spacecraft was far enough from the Earth during this time to render G_{sc} and b_{sc} negligible.

The estimation problem has been defined to estimate an 18-element $\boldsymbol{x}$ vector. It includes the six independent elements of the moment-of-inertia matrix I_m, the nine elements of the reaction-wheel scale-factor/alignment matrix C_{rw}, and the three elements of the angular momentum bias vector $\boldsymbol{h}_{bias}$. The quadratic constraint in Eq. (14c) is defined to constrain the sum of the norms squared of the three columns of C_{rw} to equal three.

The white-noise intensity of the unmodeled torque and the measurement error standard deviations that have been used in the estimator are $q_{\Delta h} = 2.3 \times 10^{-11}$ N^2-m^2-s, $\sigma_\omega = 2.5 \times 10^{-6}$ rad/s, and $\sigma_{hw} = 0.0244$ N-m-s. Note that σ_{hw} is relatively large, on the order of 2.5% of the total system angular momentum, which was about 1 N-m-s during the time when flight data were collected. The limited resolution of h_{wmeask} provided the most significant noise source in the estimation problem. Its largeness stems from the limited number of bits that are used to digitize WMAP's analog reaction-wheel tachometer data.

Three different data sets have been used to do estimation. The first is about 6 h long and starts at 15:19 UT on 1 July 2001. It includes 10 large-angle slew maneuvers that are controlled by the reaction wheels. The second data set is about 5 h long and starts at 12:57 UT on 2 July 2001. It also includes 10 slew maneuvers that are similar in magnitude, but they are more abrupt. The third data set is only 1.6 h long and starts at 19:22 UT on 2 July 2001. It operates in the WMAP sky scanning mode: The WMAP spacecraft rotates about its z axis at a nominal rate of 0.047 rad/s (spin period = 134 s), and superimposed on the spin is a nutation that has a body-axis period of 129 s and a coning half angle of 22.5 deg. This scanning motion is controlled by the reaction wheels. They operate at a nominal angular momentum magnitude of $\|h_{wmeask}\| = 29.3$ N-m-s to produce this motion, and the range of the dynamic variations of the wheel angular momentum for each axis is no more than 0.34 N-m-s. Thus, this last data set lacks significant variability of its wheel speeds and its body-axis angular rates. The nominal sample interval for all data sets is $\Delta t_k = 10$ s. Data are available at 1 Hz, but a 0.1-Hz sampling rate has been used in the interests of conserving computation time.

B. Representative Results

Consider the results of a typical estimation case. This case uses all three data sets to define its estimation problem. Different data sets can be concatenated by zeroing out the D_{0k}, D_{ik}, E_{0k+1}, and E_{ik+1} matrices for the stage that falls on the boundary between two data sets. The estimation problem does not include any a priori information in the form of R_{ap} and z_{ap}. The algorithm converged to a well-defined minimum of the least-squares cost function in about a dozen Newton iterations starting from the preflight estimates of the dynamic parameters.

The algorithm is able to improve its attitude dynamics parameter estimates significantly. Although there are no known truth values against which to measure success, there are several metrics that point to a significant improvement. One indication is that the final parameter estimates' least-squares cost is smaller than the preflight estimates' cost by a factor of nine even though the latter cost was calculated using the corresponding optimized estimates of the measurement error vector $\boldsymbol{v}_{big}$. This large difference in cost occurs mostly in the third data set.

Consider the modeling errors shown in Fig. 1. Figure 1a plots norm time histories for the estimated wheel angular momentum measurement error $\|\Delta h_{wk}\|$, and Fig. 1b plots norm time histories for the unmodeled disturbance torque $\|\Delta h_{ink}/\Delta t_k\|$. The solid grey curves on the two graphs correspond to the nominal prelaunch dynamic parameter estimates, and the dotted black curves correspond

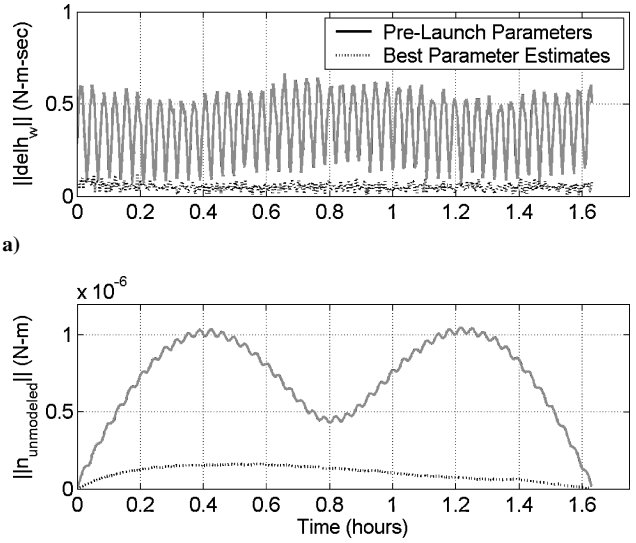

Fig. 1 Comparison of a) wheel angular momentum error magnitudes and b) unmodeled disturbance torque magnitudes for two different parameter estimates, third data set.

to the new algorithm's best estimates of the parameters. Both sets of curves use the optimal estimates of the measurement errors that correspond to the specified parameter estimates. The new estimates of the dynamic parameters yield measurement errors and unmodeled torques that, on average, are about 7 to 7.5 times smaller than those associated with the preflight parameters. Thus, the new parameter estimates model the WMAP spacecraft's dynamics much better than do the preflight parameters. Note that the residual unmodeled torque for the best parameter estimates is on the order of 0.2 μN-m. This is about what had been expected for the solar radiation pressure torque, which is the only significant disturbance torque on WMAP when it is far from the Earth and the moon.

The differences between the new parameter estimates and the preflight values are statistically significant. The maximum difference between the old and new estimates of the elements of I_m is 7.3% of the maximum principal inertia, but the maximum calculated estimation error standard deviation for an I_m element is only 0.28% of the maximum principal inertia. Similarly, the maximum change in an element of the C_{rw} matrix is 0.035, whereas the maximum standard deviation for the estimation error of an element of C_{rw} is 0.0026. The differences between the preflight parameter estimates and the new estimates are small enough to seem reasonable, given the usual methods by which preflight estimates are determined, yet they are significant if one needs a very accurate model of the attitude dynamics.

A more comprehensive method of assessing the statistical significance of the differences between the preflight and optimal parameters is to calculate the scalar test statistic $(\boldsymbol{x}_{opt} - \boldsymbol{x}_{preflight})^T Q_{a2} R_{uu}^T R_{zz}^T R_{zz} R_{uu} Q_{a2}^T (\boldsymbol{x}_{opt} - \boldsymbol{x}_{preflight})$. This is the difference squared of the estimates normalized by the inverse of the estimation error covariance matrix as projected onto the local tangent space of the quadratic constraint in Eq. (14c). This statistic will be a sample from a χ^2 distribution of degree $n_x - 1$ if the difference between the two estimates is the result only of random errors. The value of this statistic is 1.74×10^5 for this case. Given that $n_x - 1 = 17$, there is virtually zero probability that these parameter differences are the result of random noise. Put differently, the average difference between the preflight parameter estimates and the optimal estimates is $\sqrt{(1.74 \times 10^5/17)} = 101$ standard deviations, which is very significant.

One might question how differences of I_m and C_{rw} that are on the order of 7% or less can produce the factor-of-seven differences in the error magnitudes shown on Fig. 1. The answer has to do with how WMAP was being controlled during the third data set. The magnitude of the actual wheel angular momentum is 27 N-m-s, and the

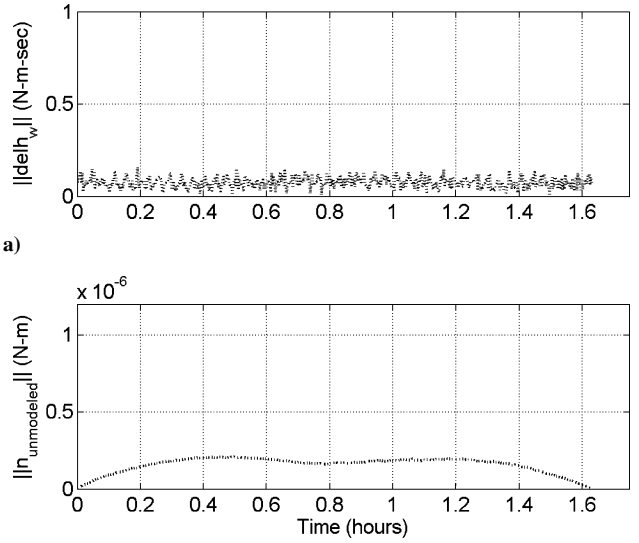

Fig. 2 Magnitudes for a) wheel angular momentum error and b) unmodeled disturbance torque for the third data set when parameter estimates based on the first two data sets are used.

magnitude of the rotational angular momentum of the main spacecraft body is 26 N-m-s, but the total spacecraft angular momentum is only 1 N-m-s. Thus, the large wheel angular momentum and the large rotational angular momentum almost cancel each other to yield a total angular momentum that is less than 4% of either component. This cancellation has the potential to amplify the significance of parameter errors by a factor of 26 to 27, which is why a 7% change in an element of I_m can have such a large impact on the unmodeled torque magnitudes.

C. Prediction Capability

Another case has been run to more clearly illustrate the usefulness of the algorithm's parameter estimates for purposes of predicting the dynamic response of the WMAP spacecraft. In this test, only the first two data sets have been used to estimate parameters, and the resulting parameter values have been used to estimate unmodeled torque errors and measurement errors for the third data set without further correction of the parameter estimates. The resulting wheel angular momentum errors and unmodeled torques are shown in Fig. 2. Note that the axis scales of Fig. 2 are the same as in Fig. 1. It is obvious from a comparison of the two figures that the parameter estimates from the first two data sets do a much better job of predicting the WMAP attitude dynamics than do the preflight parameter estimates—compare the solid gray curves in Fig. 1 with the curves in Fig. 2. The model errors when using the parameter estimates from the first two data sets are between 4.5 and 5 times smaller, on average, than the model errors when using the preflight parameters. Also obvious is the fact that the parameter estimates that use all three data sets do a slightly better job of modeling the attitude dynamics during the third data set—compare the dotted black curves in Fig. 1 with the curves in Fig. 2. This makes sense because the optimal parameter estimates from Fig. 1 interpolate into the third data set.

D. Degradation of Estimation Error with Limited Data

Another case has been run in an attempt to estimate parameters based only on the third data set. Recall that the third data set does not have a rich dynamic response. It just nutates, and its reaction-wheel speeds and angular velocity vector are almost constant. If the only scaling information added to the problem is the quadratic constraint, then the estimation algorithm takes many iterations to converge, 117, and its estimates are nonsensical. The reaction-wheel scale factors differ from the prelaunch estimates by as much as 47%, and the I_m estimate has two negative eigenvalues! This confirms a previously known fact: the estimation accuracy is dependent on the richness of the control inputs.[9]

An attempt has been made to rescue the situation by adding a priori information. The R_{ap} and z_{ap} values have been set up to indicate that the a priori reaction-wheel scale factors are accurate to within 0.2% $1-\sigma$. This estimation run yields results that are more reasonable. All of the eigenvalues of the I_m estimate are positive, and the estimates of the elements of C_{rw} differ from their prelaunch values by no more than 0.086. Nevertheless, the overall estimated parameter vector is not very accurate. For example, the minimum eigenvalue of the I_m estimate is only 10% of the preflight value. It is unrealistic to suppose that the preflight estimate would be in error by such a large amount.

The calculated estimation error standard deviations are very large even when the algorithm thinks that it has been given highly accurate a priori estimates for the scalings of the C_{rw} columns. The best estimation error standard is two times larger than the corresponding standard deviation for the case that uses all three data sets. The worst one is almost 5000 times larger than its counterpart for the three-data-set case. This is particularly telling when one considers that the case with three data sets did not even include a priori information beyond the quadratic constraint. Thus, it is critically important to use data with rich dynamic variations in the control inputs and the attitude response if one wants to get good parameter estimates from this paper's algorithm.

E. Miscellaneous Results Discussion

A test has been made of what happens when one tries to pose the attitude estimation problem as a regular least-squares problem with a quadratic constraint. One might be tempted to model the problem in this way because the global solution can be determined simply by performing two SVDs and several QR factorizations. What one sacrifices is model fidelity. One must eliminate the multiplicative errors from the model, which eliminates the errors in $\boldsymbol{\omega}_k$ and $\boldsymbol{h}_{wk}$. The difficulty with this approach is that these are the problem's most significant error sources. If one eliminates them, then the problem model becomes dubious. Although the problem is easy to solve, it is obvious that the resulting estimates contain significant errors. The minimum eigenvalue of the I_m estimate is only 22% of the preflight estimate, and the columns of the C_{rw} estimate differ in alignment from their preflight values by 20–27 deg. This failure highlights the importance of proper error modeling in parameter estimation problems.

As mentioned already, the dominant errors in the WMAP attitude dynamics data are those in the reaction-wheel telemetry stream $\boldsymbol{h}_{\text{wmeask}}$. This fact suggests two ideas for future consideration. The first is that it would be good for attitude estimation and control engineers to ask wheel subsystem engineers to deliver more accuracy in the $\boldsymbol{h}_{\text{wmeask}}$ measurement on future missions. The current level of accuracy might suffice for normal "housekeeping" purposes, but increased accuracy could be critical to any attempt to implement backup mode attitude determination or control functions in the event of a hardware failure.

The second idea has to do with the design of attitude estimation algorithms based on Euler's equations. This paper's parameter estimation algorithm obtains good results when it estimates corrections to the telemetered reaction-wheel angular momentum time histories. This suggests that one should include estimation of reaction-wheel corrections in an Euler-based attitude and rate estimation Kalman filter if the reaction wheel data are of low accuracy. Such an augmentation can make the difference between success and failure of the attitude filter.

VI. Conclusions

A new algorithm has been developed for estimating the parameters of a spacecraft's attitude dynamics model by using telemetered attitude and rate measurements. The estimated quantities include the moment-of-inertia matrix and, if present, alignments, scale factors, and biases for reaction wheels and magnetic torque rods. The input data include three-axis attitude from star trackers, three-axis rate from rate gyros, nominal reaction-wheel angular momenta, nominal torque rod dipole moment strengths, and magnetic field as measured by a magnetometer.

The algorithm uses a trapezoidally integrated version of Euler's equation in inertial coordinates as its basic estimation equation. The resulting estimation problem includes additive disturbance torque errors and multiplicative sensor measurement errors and is in the general form known as a total least-squares problem. A scalar quadratic constraint is added to the least-squares problem statement in order to make it well posed by constraining the overall scaling of the parameter estimates.

The algorithm uses inner and outer least-squares optimizations to estimate the dynamic model parameters. The inner optimization uses linear least-squares techniques to compute estimates of the sensor measurement errors for given values of the model parameters. The outer optimization uses a guarded Newton iterative numerical procedure to estimate the most likely model parameters.

The algorithm has been tested using data from the Wilkinson Microwave Anisotropy Probe spacecraft. The algorithm's best estimates of the moment-of-inertia matrix and the reaction-wheel scale-factor/alignment matrix differ from the preflight estimates by 7% or less, but they can reduce the level of torque modeling error by a factor of 5 to 7 in certain modes of operation where spacecraft angular momentum and reaction wheel angular momentum almost cancel each other.

Acknowledgments

This work was supported in part by NASA Grant NAG5-11919. Richard Harman was the Grant Monitor and supplied the Wilkinson Microwave Anisotropy Probe data that have been used.

References

[1]Psiaki, M. L., Martel, F., and Pal, P. K., "Three-Axis Attitude Determination via Kalman Filtering of Magnetometer Data," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 3, 1990, pp. 506–514.

[2]Psiaki, M. L., Theiler, J., Bloch, J., Ryan, S., Dill, R. W., and Warner, R. E., "ALEXIS Spacecraft Attitude Reconstruction with Thermal/Flexible Motions due to Launch Damage," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 5, 1997, pp. 1033–1041.

[3]Psiaki, M. L., Klatt, E. M., Kintner, P. M., Jr., and Powell, S. P., "Attitude Estimation for a Flexible Spacecraft in an Unstable Spin," *Journal of Guidance, Control, and Dynamics*, Vol. 25, No. 1, 2002, pp. 88–95.

[4]Challa, M., Natanson, G., and Ottenstein, N., "Magnetometer-Only Attitude and Rates for Spinning Spacecraft," *Proceedings of the AIAA/AAS Astrodynamics Specialists Conference*, AIAA, Reston, VA, 2000, pp. 311–321.

[5]Psiaki, M. L., and Oshman, Y., "Spacecraft Attitude Rate Estimation from Geomagnetic Field Measurements," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 2, 2003, pp. 244–252.

[6]Psiaki, M. L., "Global Magnetometer-Based Spacecraft Attitude and Rate Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 2, 2004, pp. 240–250.

[7]Markley, F. L., Andrews, S. F., O'Donnell, J. R., Jr., and Ward, D. K., "The Microwave Anisotropy Probe (MAP) Mission," *Proceedings of the AIAA Guidance, Navigation, and Control Conference* [CD-ROM], AIAA, Reston, VA, 2002.

[8]Bergmann, E. V., Walker, B. K., and Levy, D. R., "Mass Property Estimation for Control of Asymmetrical Satellites," *Journal of Guidance, Control, and Dynamics*, Vol. 10, No. 5, 1987, pp. 483–491.

[9]Richfield, R. F., Walker, B. K., and Bergmann, E. V., "Input Selection for a Second-Order Mass Property Estimator," *Journal of Guidance, Control, and Dynamics*, Vol. 11, No. 3, 1988, pp. 207–212.

[10]Bergmann, E., and Dzielski, J., "Spacecraft Mass Property Identification with Torque-Generating Control," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 1, 1990, pp. 99–103.

[11]Tanygin, S., and Williams, T., "Mass Property Estimation Using Coasting Maneuvers," *Journal of Guidance, Control, and Dynamics*, Vol. 20, No. 4, 1997, pp. 625–632.

[12]Peck, M. A., "Mass-Properties Estimation for Spacecraft with Powerful Damping," *Proceedings of the AAS/AIAA Astrodynamics Specialist Conference*, American Astronautical Society, San Diego, CA, 1999, pp. 2005–2024.

[13]Peck, M. A., "Attitude Determination for Gyrostats in Non-Equilibrium Spins from Infrequent Vector Observations," *Proceedings of the AIAA Guidance, Navigation, and Control Conference* [CD-ROM], AIAA, Reston, VA, 2000.

[14]Peck, M. A., "Estimation of Wheel and CMG Alignments from on-Orbit Telemetry," *Proceedings of the Flight Mechanics Symposium*, edited by J. P. Lynch, NASA Center for Aerospace Information, Hanover, MD, 2001, pp. 187–201.

[15]Golub, G. H., and Van Loan, C. F., "An Analysis of the Total Least Squares Problem," *SIAM Journal on Numerical Analysis*, Vol. 17, No. 6, 1980, pp. 883–893.

[16]Van Huffel, S., and Vandewalle, J., "Analysis and Properties of the Generalized Total Least Squares Problem AX $\approx$ B When Some or All Columns in A Are Subject to Error," *SIAM Journal on Matrix Analysis and Applications*, Vol. 10, No. 3, 1989, pp. 294–315.

[17]Bierman, G. J., *Factorization Methods for Discrete Sequential Estimation*, Academic Press, New York, 1977, pp. 69–76, 115–122.

[18]Gill, P. E., Murray, W., and Wright, M. H., *Practical Optimization*, Academic Press, New York, 1981, pp. 37–40, 88–92.

[19]Psiaki, M. L., "Attitude-Determination Filtering via Extended Quaternion Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 23, No. 2, 2000, pp. 206–214.