RICE UNIVERSITY

Spacecraft Attitude Estimation Integrating the Q-Method into an Extended Kalman Filter

by

Thomas G. Ainscough, 2d Lt, USAF

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Master of Science

APPROVED, THESIS COMMITTEE:

Pol D. Spanos, Chair
Lewis B. Ryon Professor of Mechanical
Engineering & Materials Science

Renato Zanetti
Group Leader Vehicle Dynamics and
Control, C.S. Draper Laboratory

Andrew J. Dick
Assistant Professor Mechanical
Engineering & Materials Science

Rouzbeh Shahsavari
Assistant Professor Civil &
Environmental Engineering

HOUSTON, TEXAS

APRIL, 2013

Abstract

Spacecraft Attitude Estimation Integrating the Q-Method into an Extended Kalman Filter

by

Thomas G. Ainscough, 2d Lt, USAF

A new algorithm is proposed that smoothly integrates the nonlinear estimation of the attitude quaternion using Davenport's q-method and the estimation of non-attitude states within the framework of an extended Kalman filter. A modification to the q-method and associated covariance analysis is derived with the inclusion of an *a priori* attitude estimate. The non-attitude states are updated from the nonlinear attitude estimate based on linear optimal Kalman filter techniques. The proposed filter is compared to existing methods and is shown to be equivalent to second-order in the attitude update and exactly equivalent in the non-attitude state update with the Sequential Optimal Attitude Recursion filter. Monte Carlo analysis is used in numerical simulations to demonstrate the validity of the proposed approach. This filter successfully estimates the nonlinear attitude and non-attitude states in a single Kalman filter without the need for iterations.

Acknowledgments

I would like to thank Dr. Pol Spanos and Rice University for providing me with the opportunity to obtain a master's degree. I am grateful for the mentoring and enthusiasm Dr. Spanos provided as my advisor. He ensured that I was able to develop the requisite aptitude for conducting research and scrutinizing its scientific findings. I owe a special thanks to Dr. Renato Zanetti for his invaluable guidance in support of my research and thesis. His help in everything from building a strong foundation of the relevant theory, understanding its potential applications, and all the way down to debugging stubborn code has been extraordinary. I also wish to thank Dr. Andre Dick and Dr. Rouzbeh Shahsavari for eargerly serving on the committee. I am grateful to the Charles Stark Draper Laboratory for the great experiences and opportunity to pursue my passion in astronautical engineering.

I would like to thank my family for their continual support. I owe so much of what I have accomplished to their encouragement and example. I owe a special thanks to my wonderful girlfriend. Her love and support has been so much more than I can ever repay. I am a better person because of her.

I would also like to thank all my fellow Air Force LTs for the great times and the memories, especially my roommate. They are some of the greatest people that I have met and our shared experience is a unique bond. I will miss them all and look forward to crossing paths again throughout our careers.

Finally, I would like to thank the U.S. Air Force for the opportunity to pursue my graduate studies as my first assignment.

Contents

List of Figures

Chapter 1

Introduction

Virtually all spacecraft require some sort of attitude determination. While accuracy requirements vary based on mission requirements, a certain level of attitude information is vital. Often accurate attitude estimation is essential for the primary mission objective such as in remote sensing satellites or space based telescopes like the well known Hubble Space Telescope. Furthermore, some degree of attitude determination is necessary for successful operation of satellite secondary subsystems such as orientation with respect to the sun in order to optimize solar panel efficiency or proper orientation of antenna for effective communications. As a result, a variety of sensors and algorithms have been developed over the decades in order to accurately estimate attitude with a wide range of complexity and cost. Highly accurate sensors such as star trackers are capable of determining the spacecraft attitude with very high accuracy, but come at a correspondingly high cost which could exceed the budget for some applications. Cost savings on spacecraft subsystems such as the attitude determination system provide a greater margin for the primary mission. It is therefore desirable to employ attitude estimation algorithms that can sufficiently determine the spacecraft attitude from noisy measurements.

The basic problem of spacecraft attitude determination is to ascertain the spacecraft's orientation by comparing measurements from attitude sensors in a spacecraft body-fixed frame to a known reference frame. By determining the appropriate rotation from the known reference frame to the measurements in the spacecraft body frame the attitude can be determined. Typical spacecraft attitude sensors consist of sun sensors, star trackers, horizon sensors, magnetometers, and GPS receivers as well as inertial sensors such as various forms of gyroscopes and accelerometers. Attitude determination methods can be classified into the two main classes of single-point attitude determination methods and recursive attitude estimation methods. Single-point solutions utilize two or more vector measurements to calculate the attitude at a single point in time. Recursive estimation algorithms combine attitude measurements over time with kinematic and dynamic models to estimate the spacecraft attitude. The advantages and shortcomings of both classes has led to the development of many different attitude determination methods.

This thesis presents a new recursive algorithm for attitude estimation which incorporates a nonlinear attitude estimation method into the framework of the extended Kalman filter. The proposed filter uses Davenport's q-method [1] to solve for the attitude without any small angle approximations and uses the nonlinear attitude solution to update the non-attitude states using the optimal gain from the Kalman filter. The standard q-method solution is modified to incorporate *a priori* attitude information using the method of averaging quaternions [2] and the corresponding error covariance analysis is performed. It is shown that the method of first updating the attitude and subsequently updating the non-attitude states from the attitude update is equivalent to the standard Kalman filter for the linear measurement case. The proposed q-method extended Kalman filter (qEKF) is presented where the non-attitude states are updated according to the linear case. The qEKF is compared with the Sequential Optimal Attitude Recursion (SOAR) filter [3] and shown to be equivalent. Pertinent numerical simulations are used to verify the proposed algorithm.

1.1 Single-Point Methods

Single-point attitude determination methods are also known as point-by-point methods or batch estimation as they determine the attitude of a spacecraft at a single point in time from at least two vector observations. Spacecraft attitude sensors typically output unit vector measurements, $\mathbf{y}$, which can be compared with known reference vectors, $\mathbf{n}$, in order to identify the attitude often represented as the orthogonal attitude matrix, $\mathbf{T}$. Specifically,

$$\mathbf{y}_i = \mathbf{T}\mathbf{n}_i \qquad \text{for } i = 1, ..., n. \tag{1.1}$$

With single-point methods, knowledge of the spacecraft dynamics is not necessary. This results in reduced complexity and potentially negating the need for expensive sensors such as accelerometers and gyroscopes. *A priori* knowledge of the attitude is also not required enabling single-point methods to be effective when such information is lacking or very poor. One of the oldest and often used single-point methods is the TRIAD algorithm also known as the Algebraic Method [4, 5]. In this method exactly two unit vector measurements and the corresponding unit vectors in the reference frame are used to calculate the attitude. In the absence of noise, the two unit vector measurements provide sufficient information to determine the attitude as represented in Eq. (1.1). However, in the case of noisy measurements a solution does not typically exist. The TRIAD algorithm combines the unit vector measurements in such a way that discards part of the less accurate measurement in order to obtain a solution for the attitude in the presence of noise. This method is very simple and quickly determines the spacecraft attitude with very little computational cost. However, the TRIAD algorithm is only capable of processing exactly two unit vector measurements and therefore incapable of incorporating additional attitude measurements which may be available. Nevertheless, the simplicity and success of the TRIAD algorithm has led to its use in several satellites over the years such as in the Navy Navigation Satellite System [6].

The foundation of most single-point attitude determination algorithms from vector observations is the well known *Wahba problem* [7]. While interning with IBM Federal which was supporting NASA attitude activities, Grace Wahba posed the problem in a 1965 issue of *SIAM Review* [7]. The Wahba problem is simply a nonlinear, weighted least-squares problem to determine the optimal attitude matrix from a set of at least two independent vector measurements. The resulting performance index to be minimized is given by

$$\mathcal{J} = \frac{1}{2}\sum_{i=1}^{n} a_i \left\|\mathbf{y}_i - \mathbf{T}\mathbf{n}_i\right\|^2, \tag{1.2}$$

where a_i are scalar, positive weights associated with each vector pair. The Wahba problem is capable of processing any number of synchronized, noisy vector measurements to produce the optimal attitude in the sense of minimizing the weighted residual between the reference and measurement vectors. It has received much attention over the years because of its ability to provide a globally optimal solution for the attitude without making any linearization or small angle assumptions [1, 5, 8, 9, 10, 11, 12, 13, 14].

Over the years many solutions have been developed to solve the Wahba problem, some of which are purely mathematical and others much more relevant to practical applications. Most solutions to the Wahba problem rewrite the performance index in some manner. A typical approach is to rewrite the Wahba performance index as a function of the attitude quaternion. Paul Davenport [1] showed that this approach results in a quadratic performance index and the optimal solution is obtained by solving an eigenvalue problem.

As will be shown in a later section the loss function of the Wahba problem can be rewritten as

$$\mathcal{J} = \sum_{i=1}^{n} a_i - \text{trace}\left[\mathbf{T}\mathbf{B}^{\mathrm{T}}\right], \tag{1.3}$$

where

$$\mathbf{B} \equiv \sum_{i=1}^{n} a_i \mathbf{y}_i \mathbf{n}_i^{\mathrm{T}}, \tag{1.4}$$

and trace[$\cdot$] signifies the matrix trace. A solution developed by Markley takes advantage of the fact that the Wahba problem rewritten as in Eq. (1.3) is a special case of the *Orthogonal Procrustes Problem* [15] and then solved using Singular Value Decomposition (SVD) [16]. In SVD the optimal attitude is obtained directly from decomposing the **B** matrix to its singular values. While computationally intensive, SVD utilizes mathematically rigorous matrix algorithms and is very robust [17]. In order to reduce the computational burden Markley developed a numerical extension of SVD in the Fast Optimal Attitude Matrix algorithm (FOAM) which is much more efficient [12]. In FOAM the singular values are used to develop an expression for the optimal attitude matrix that does not require the singular value decomposition, but rather computes the necessary coefficients by means of iteration from relationships derived from the singular values.

Davenport's solution (also known as the q-method) calculates the attitude quaternion rather than the orthogonal attitude matrix [1]. The attitude quaternion is subject to a unit normal constraint and the resulting loss function is an eigenvalue problem where the largest eigenvalue of the Davenport matrix minimizes the loss function and the corresponding unit eigenvector is the optimal attitude quaternion. A more detailed derivation of the q-method is included in a subsequent section. Davenport's q-method is also very robust and solves the nonlinear Wahba problem exactly without any linearization or simplifying assumptions.

While mathematically rigorous, solving the eigenvalue problem of the q-method is computationally burdensome and not ideal for on-board attitude determination. As a result, numerous numerical techniques have been developed to estimate a solution to the q-method in a more efficient manner. The foundation for such numerical solutions is the Quaternion Estimator (QUEST)[8]. The most computationally burdensome part of the q-method is solving the eigenvalue problem. Shuster noted that when the value of the performance index in Eq. (1.3) is small (which is a valid assumption as the attitude is chosen such that the performance index is minimized) the maximum eigenvalue is very close to $\sum_{i=1}^{n} a_i$ which in turn may be used as the starting value for a Newton-Rhapson iteration that quickly converges to the maximum eigenvalue. After some manipulation in which the quaternion is factored in terms of Rodrigues parameters [18] the eigenvector can be computed. The introduction of Rodrigues parameters also adds a singularity for rotations of π radians which Shuster avoids by employing a method of sequential rotations [8]. The computational requirements for QUEST are significantly reduced compared to the q-method which has made the algorithm much more appealing for real-time on board attitude estimation. With the increased speed QUEST sacrifices robustness, but performs well as long as the measurement noise does not vary excessively between measurements [17].

Another numerical solution to the q-method is the Estimator of the Optimal Quaternion (ESOQ) algorithm [13]. Mortari uses the same iterative method to calculate the eigenvalue as in QUEST, but avoids the the singularity introduced by the Rodrigues parameters by instead computing the quaternion as a four-dimensional vector cross product. This arises as a result of the eigenvalue problem. The optimal quaternion or eigenvector must be orthogonal to all columns of the matrix $\mathbf{K} - \lambda_{\max}\mathbf{I}$ where $\mathbf{K}$ is the Davenport matrix, $\lambda_{\max}$ the corresponding maximum eigenvalue, and $\mathbf{I}$ is the identity matrix. Therefore the eigenvector is computed for the four-dimensional cross product of any three columns of the aforementioned matrix. ESOQ2 is a follow-on algorithm which parametrizes the quaternion in terms of the Euler axis/angle representation of the attitude [14]. This parametrization calculates the Euler axis from the null space of a 3×3 matrix that is derived from the Davenport matrix. However, it also introduces a singularity for a zero angle rotation which is also resolved by using successive rotations. Like QUEST, ESOQ and ESOQ2 are much faster than the q-method, but with reduced robustness.

1.2 Recursive Estimation Methods

In contrast with single-point methods which process batches of measurements at a single time in order to produce an attitude measurement, recursive estimation methods take into account the measurements of all previous times while accounting for vehicle dynamics in order to produce an accurate estimate of the attitude at the current time. They combine previous measurements and propagate the estimate to the current time while also providing an estimate of the accuracy of the current state. These techniques can process any type or number of attitude measurements and can be used to filter out measurement noise. Stochastic processes are used in order to combine the measurements and previous estimate in some statistically optimal manner. If properly tuned, recursive methods are capable of producing very accurate estimates of the attitude in real time. However, they often require *a priori* knowledge of the state and can be sensitive to initial conditions, possibly diverging for poor initial estimates.

The workhorse of recursive estimation is the extended Kalman filter [19, 20, 21]. The Kalman filter is essentially a recursive form of linear least squares estimation. Provided a measurement model relating the states to the measurements a Kalman filter minimizes the residual between the observed measurements and the expected measurements based on the measurement model and the current state. For linear, Gaussian systems the Kalman filter provides the optimal minimum mean square error estimate of the state at the current time given all previous measurements. The algorithm is split into two main steps: a propagation step which moves the state estimate and covariance forward in time and a update step which incorporates new measurements into the current state estimate. Because few systems of interest a linear, the extended Kalman filter (EKF) applies the method to nonlinear systems. In EKF the system is linearized by a first-order Taylor series expansion about the nominal or current state estimate. Instead of actually estimating the state itself, EKF estimates the differential correction to the state which is then added to the nominal state. As a result of the linearization the extended Kalman filter is no longer optimal and can diverge for highly nonlinear systems or large errors in the initial condition. However, the extended Kalman filter has proved sufficiently accurate for a wide number of applications and continues to be a highly successful recursive estimation method.

In the context of attitude estimation the preferred attitude representation is the quaternion for its efficiency in computations and lack of singularities. However, the quaternion is also subject to a unit norm constraint which makes direct implementation into a Kalman filter more difficult. As opposed to a traditional Kalman filter which features an additive update, the Multiplicative Extended Kalman Filter (MEKF) maintains the unity constraint on the quaternion and makes use of the fact that an additional rotation is more accurately represented as a multiplication of the quaternion instead of an addition [22]. Recall that EKF uses differential correction to the state and in the case of attitude amounts to a small rotation. For the quaternion this amounts to the multiplication of the initial quaternion by an incremental quaternion. Recognizing that the incremental rotation will necessarily by a small angle, the scalar component of the quaternion will be approximately one and the incremental rotation may be accurately expressed as the three-dimensional error angles expressed in the body frame. These error angles may be viewed as the small deviations of the

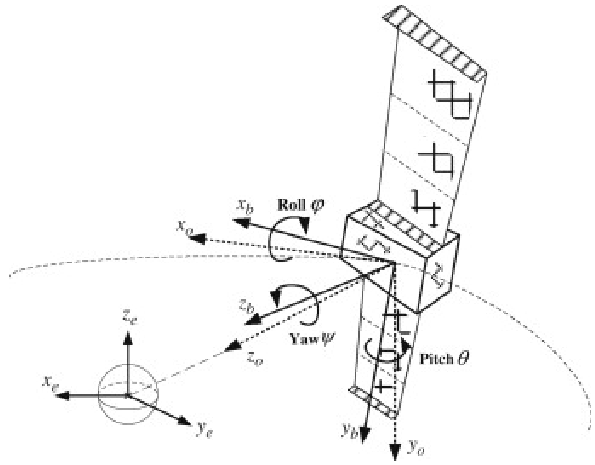

Figure 1.1: Spacecraft Attitude Representation

traditional roll, pitch, and yaw angles as shown in Figure 1.1 [23]. In the MEKF this three-dimensional error vector is used to represent the attitude portion of the state and error covariance. As a result, the complications that arise from the unit norm constraint on the quaternion can be avoided. The updated quaternion is obtained by composing the error angle vector with the associated *a priori* quaternion.

As the Wahba problem provides a globally optimal nonlinear attitude estimate, several methods have been developed that seek to mechanize the Wahba problem into a recursive algorithm for continuous attitude estimation. Shuster proved that when the scalar weights of the Wahba problem are chosen according to the QUEST measurement model that the resulting attitude estimate is a maximum likelihood estimate [8, 9]. Recognizing that the Kalman filter is a sequential mechanization of maximum likelihood estimation, Shuster extended QUEST into a Kalman filter to produce Filter QUEST [24]. The attitude profile matrix defined in Eq. (1.4) contains all the necessary information to compute the maximum likelihood attitude estimate. In [9] Shuster also shows that the attitude covariance matrix may also be extracted from the attitude profile matrix using the relationship

$$\mathbf{B} \simeq \left(\frac{1}{2}\mathrm{trace}\left[\mathbf{P}_{\boldsymbol{\theta\theta}}^{-1}\right]\mathbf{I}_{3\times 3} - \mathbf{P}_{\boldsymbol{\theta\theta}}^{-1}\right)\hat{\mathbf{T}}, \tag{1.5}$$

where $\mathbf{P}_{\boldsymbol{\theta\theta}}$ is the attitude covariance, $\hat{\mathbf{T}}$ is the estimated attitude matrix, and the approximation is accurate to the lowest order in the standard deviation of the vector measurements. As a result, updating the attitude profile matrix provides both the attitude and the covariance updates. The attitude matrix is propagated forward in time until the next measurement at which point it is updated following the definition of the attitude profile matrix by simply adding the new measurements to the propagated matrix as the ith measurements. The attitude and covariance are not directly computed during the course of the filter, but can readily be obtained from the attitude profile matrix at any point using QUEST to determine the attitude and the equation

$$\mathbf{P}_{\boldsymbol{\theta\theta}}^{-1} \simeq \mathrm{trace}\left[\hat{\mathbf{T}}\mathbf{B}\right]\mathbf{I}_{3\times 3} - \hat{\mathbf{T}}\mathbf{B} \tag{1.6}$$

to calculate the corresponding covariance. As opposed to an extended Kalman filter, Filter QUEST does not require linearization and updates the full attitude in the form of the attitude profile matrix instead of using differential correction. However, process noise is not included in Filter QUEST and its effects are approximated by the addition of a fading memory factor.

An alternative recursive algorithm based off of QUEST and dubbed REQUEST is developed by Bar-Itzhack [25]. In REQUEST the entire **K** matrix of the Davenport solution to the Wahba problem is recursively updated as opposed to only the attitude profile matrix in Filter QUEST. Like Filter QUEST, REQUEST accounts for process noise during state propagation by adding a fading memory factor which is determined heuristically. Shuster shows that REQUEST is mathematically equivalent to Filter QUEST [26]. Both of these methods represent sub-optimal attitude estimation filters which account for process noise using fading memory factors as weights to appropriately scale the states, but do not account for measurement noise. The sub-optimality is addressed and accounted for to produce an optimal attitude estimation filter in Optimal-REQUEST [27]. Optimal-REQUEST uses Kalman filtering techniques to solve for the optimal fading memory factor of REQUEST accounting for both process and measurement noise. All of these filters only estimate attitude and are incapable of estimating any additional states.

In order to accommodate the inclusion of non-attitude states into the attitude estimation filters based on the Wahba problem, additional derivations of the above filters have been developed. In [28] Markley first demonstrated how to incorporate other parameters such as sensor biases into an attitude estimation algorithm based on the Wahba problem. He reformulates the loss function of the Wahba problem given in Eq. (1.3) to include other parameters. The resulting performance index is given by

$$\mathcal{J} = \sum_{i=1}^{n} a_i - \mathrm{tr}\left[\mathbf{T}\mathbf{B}\left(\mathbf{x}\right)^{\mathrm{T}}\right] + \frac{1}{2}\left(\mathbf{x}-\mathbf{x}_0\right)^{\mathrm{T}}\mathbf{W}_0\left(\mathbf{x}-\mathbf{x}_0\right), \tag{1.7}$$

where the reference and measurement vectors of the attitude profile matrix, $\mathbf{B}$, are allowed to be functions of non-attitude states, $\mathbf{x}$, and $\mathbf{W}_0$ is a weighting matrix corresponding to the *a priori* estimates. The attitude is obtained by using the q-method and the non-attitude states are determined using an iterative procedure. Extended QUEST expands the QUEST-based filters to include non-attitude states using square-root information filtering techniques [29]. Where Markley's algorithm uses batch iteration to determine the non-attitude states, Extended QUEST more closely resembles the form of a Kalman filter by using stage-to-stage iterations. In Extended QUEST the loss function is also modified to include a series of additional terms that penalize the differences between the attitude, non-attitude states, and process noise.

Another recursive algorithm, the Sequential Optimal Attitude Recursion (SOAR) filter is proposed by Christian and Lightsey [3]. In SOAR the Wahba problem is recast into the framework of maximum likelihood estimation which allows for the straightforward inclusion of other parameters. It is developed as the information matrix formulation of the extended Kalman filter. Using Bayesian estimation the performance index for SOAR is given by

$$\mathcal{J} = \frac{1}{2}\left(\mathbf{x}-\hat{\mathbf{x}}_0\right)^{\mathrm{T}}\mathbf{P}_0^{-1}\left(\mathbf{x}-\hat{\mathbf{x}}_0\right)+\frac{1}{2}\sum_{i=1}^{n}a_i\left\|\mathbf{y}_i-\mathbf{T}\mathbf{n}_i\right\|^2, \tag{1.8}$$

where $\hat{\mathbf{x}}_0$ is the initial estimate of the full state with covariance $\mathbf{P}_0$. In SOAR the state vector is partitioned into the same three-dimensional attitude error angle vector, $\boldsymbol{\theta}$, as MEKF and a non-attitude state vector, $\boldsymbol{\beta}$. The optimal attitude is solved using the q-method and the optimal update for the non-attitude states is derived. The *a priori* attitude is incorporated by calculating an *a priori* attitude profile matrix from the quaternion and attitude covariance using the same relationship as identified by Shuster [9]

$$\mathbf{B}^{-} = \left(\frac{1}{2}\text{trace}\left[\left(\mathbf{P}^{-}_{\boldsymbol{\theta}\boldsymbol{\theta}}\right)^{-1}\right]\mathbf{I}_{3\times 3} - \left(\mathbf{P}^{-}_{\boldsymbol{\theta}\boldsymbol{\theta}}\right)^{-1}\right)\mathbf{T}\left(\hat{\bar{\mathbf{q}}}^{-}\right), \tag{1.9}$$

where the superscript $[\cdot]^{-}$ represents *a priori* quantities. The measurement attitude profile matrix is computed according to the q-method with both attitude profile matrices used to form the Davenport matrix $\mathbf{K}$ which results in the standard eigenvalue problem for the optimal attitude. The optimal non-attitude state update arises from partitioning the state in Eq. (1.8) and minimizing the cost function subject to the unity constraint on the quaternion according to standard optimal control theory. The resulting filter avoids the assumptions surrounding the fading memory factor of Filter QUEST and REQUEST to produce an optimal state update that achieves equivalent performance to MEKF when errors are small and superior performance in the presence of large errors.

The q-method extended Kalman filter presented in this thesis, like SOAR, incorporates the q-method into a Kalman filter in order to perform the optimal attitude update and the non-attitude states are updated based on the attitude update. The main difference between SOAR and qEKF is that the former utilizes the information formulation of the Kalman filter whereas the qEKF is derived using the covariance formulation. This is advantageous when the state vector is large as the required matrix inversions in the algorithm will generally be smaller for the covariance formulation. In the qEKF the method of averaging quaternions [2] is used to incorporate the initial attitude and covariance whereas SOAR uses the information matrix relationship identified by Shuster [9] and discussed above. A comparison of the two approaches shows that the attitude update is equivalent at least to second-order while the non-attitude state update is identical.

1.3 Thesis Outline

This thesis begins with an introduction to the Wahba problem and a derivation of the Davenport solution in Chapter 2. Attitude covariance analysis is also performed based on the Wahba problem and a discussion on the appropriate selection of the Wahba problem weights is included. Chapter 2 concludes with the inclusion of an initial attitude estimate into the framework of the Wahba problem with a solution for the optimal quaternion and associated attitude covariance. Chapter 3 begins with a brief introduction to the Kalman filter. The appropriate equations for a partitioned state where the measurements are only a function of part of the state are derived. The update equations of the q-method extended Kalman filter are derived performing the attitude update first followed by the non-attitude state update and shown to be equivalent. The qEKF algorithm is presented and compared to Filter Quest and SOAR. Several numerical simulations are presented in Chapter 4 along with a comparison to the SOAR filter using the same simulation. Concluding remarks and recommendations for future work close out the thesis in Chapter 5.

Chapter 2

Attitude Estimation

The first task of the algorithm proposed in this thesis is to accurately estimate the spacecraft attitude. As outlined in the introduction, the Wahba problem provides an advantageous approach to attitude estimation and is the foundation for this work. As a result, the Wahba problem and associated Davenport solution are derived, with a slight modification to the definition of the Davenport matrix, along with the associated attitude error covariance matrix. A clarification on the selection of scalar weights proposed by Shuster in [8] is also provided. A significant feature of this work is the inclusion of an initial attitude estimate into the framework of the Wahba problem. This thesis integrates the performance index proposed in [2] to include the initial condition and resulting modifications to the q-method and covariance analysis are derived.

2.1 Wahba Problem and Davenport Solution

The Wahba problem is a nonlinear, weighted least squares problem composed of the attitude matrix and vector measurements [7]. Solutions to the Wahba problem seek an optimal attitude estimate from vector measurements in the sense of minimizing the least squares residual. Davenport's solution, the so called q-method, provides a global solution for the Wahba problem parametrized with the attitude quaternion without any simplifying assumptions [1]. It is this globally optimal, nonlinear attitude update which the present work seeks to integrate into a Kalman filter. This approach differs from the linearization which is required for a traditional extended Kalman filter algorithm. As it forms the foundation of the proposed filter, a derivation of the q-method is presented here. The attitude matrix $\mathbf{T}$ is obtained from observation unit vectors $\mathbf{y}$ in the spacecraft body frame and corresponding reference unit vectors $\mathbf{n}$ in the inertial frame. Solving the Wahba problem requires the minimization of the performance index

$$\mathcal{J} = \frac{1}{2}\sum_{i=1}^{n} a_i \left\|\mathbf{y}_i - \mathbf{T}\mathbf{n}_i\right\|^2 , \tag{2.1}$$

where a_i are scalar, positive weights associated with each vector pair [7]. The Wahba problem can be rewritten as

$$\begin{aligned}
\mathcal{J} &= \frac{1}{2}\sum_{i=1}^{n} a_i \left[\mathbf{y}_i - \mathbf{T}\mathbf{n}_i\right]^{\mathrm{T}} \left[\mathbf{y}_i - \mathbf{T}\mathbf{n}_i\right] \\
&= \frac{1}{2}\sum_{i=1}^{n} a_i \left(\mathbf{y}_i^{\mathrm{T}}\mathbf{y}_i + \mathbf{n}_i^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{T}\mathbf{n}_i - 2\mathbf{y}_i^{\mathrm{T}}\mathbf{T}\mathbf{n}_i\right) \\
&= \sum_{i=1}^{n} a_i \left(1 - \mathbf{y}_i^{\mathrm{T}}\mathbf{T}\mathbf{n}_i\right)
\end{aligned} \tag{2.2}$$

recognizing that the observed and reference vectors have unit length and that the attitude matrix is orthonormal. The Davenport solution to the Wahba problem parametrizes the attitude matrix by the quaternion $\bar{\mathbf{q}} = \begin{bmatrix} \mathbf{q}_v^{\mathrm{T}} & q_4 \end{bmatrix}^{\mathrm{T}}$ consisting of a vector component $\mathbf{q}_v$ and a scalar component q_4 [1]. The attitude matrix is determined from the quaternion by [10]

$$\mathbf{T}(\bar{\mathbf{q}}) = \mathbf{I}_{3\times 3} - 2q_4\left[\mathbf{q}_v\times\right] + 2\left[\mathbf{q}_v\times\right]^2, \tag{2.3}$$

where the skew symmetric cross product matrix is defined as

$$\left[\mathbf{q}_v\times\right] = \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix}. \tag{2.4}$$

From Eq. (2.2) it is clear that minimizing the performance index $\mathcal{J}$ is equivalent to maximizing the auxiliary problem

$$\max_{\bar{\mathbf{q}}} \mathcal{G} = \sum_{i=1}^{n} a_i \mathbf{y}_i^{\mathrm{T}} \mathbf{T}(\bar{\mathbf{q}}) \mathbf{n}_i. \tag{2.5}$$

Introducing

$$\mathbf{B} \equiv \sum_{i=1}^{n} a_i \mathbf{y}_i \mathbf{n}_i^{\mathrm{T}}, \tag{2.6}$$

which is commonly referred to as the attitude profile matrix because it contains all the necessary information to compute the attitude, and using matrix trace properties, the auxiliary performance index is rewritten as

$$\begin{aligned} \mathcal{G} &= \mathrm{trace}[\mathcal{G}] = \mathrm{trace}\left[\sum_{i=1}^{n} a_i \mathbf{y}_i^{\mathrm{T}} \mathbf{T} \mathbf{n}_i\right] = \mathrm{trace}\left[\sum_{i=1}^{n} a_i \mathbf{T}^{\mathrm{T}} \mathbf{y}_i \mathbf{n}_i^{\mathrm{T}}\right] \\ &= \mathrm{trace}\left[\mathbf{T}^{\mathrm{T}}\mathbf{B}\right] = \mathrm{trace}\left[\mathbf{B}^{\mathrm{T}}\mathbf{T}(\bar{\mathbf{q}})\right]. \end{aligned} \tag{2.7}$$

By substituting Eq. (2.3) the auxiliary function can be rewritten as

$$\mathcal{G} = \text{trace}\left[\mathbf{B}^{\mathrm{T}}\right] - 2q_4\text{trace}\left[\mathbf{B}^{\mathrm{T}}\left[\mathbf{q}_v\times\right]\right] + 2\text{trace}\left[\mathbf{B}^{\mathrm{T}}\left[\mathbf{q}_v\times\right]^2\right]. \tag{2.8}$$

Using matrix trace properties it is possible to rewrite Eq. (2.8) in a more useful format [8]. The first term can be rewritten as

$$\text{trace}\left[\mathbf{B}^{\mathrm{T}}\right] = \text{trace}\left[\mathbf{B}\right] \equiv \sigma. \tag{2.9}$$

The second term of Eq. (2.8) is rewritten as the following

$$\begin{aligned}
-2q_4\text{trace}\left[\mathbf{B}^{\mathrm{T}}\left[\mathbf{q}_v\times\right]\right] &= -2q_4\text{trace}\left[\left[\mathbf{q}_v\times\right]^{\mathrm{T}}\sum_{i=1}^{n} a_i\mathbf{y}_i\mathbf{n}_i^{\mathrm{T}}\right] \\
&= -2q_4\text{trace}\left[\sum_{i=1}^{n} a_i\mathbf{y}_i^{\mathrm{T}}\left[\mathbf{q}_v\times\right]\mathbf{n}_i\right] \\
&= 2q_4\text{trace}\left[\sum_{i=1}^{n} a_i\mathbf{y}_i^{\mathrm{T}}\left[\mathbf{n}_i\times\right]\mathbf{q}_v\right] \\
&= 2q_4\text{trace}\left[\sum_{i=1}^{n} a_i\mathbf{q}_v^{\mathrm{T}}\left(\mathbf{y}_i\times\mathbf{n}_i\right)\right] \\
&= 2q_4\mathbf{z}^{\mathrm{T}}\mathbf{q}_v,
\end{aligned} \tag{2.10}$$

where

$$\mathbf{z} \equiv \sum_{i=1}^{n} a_i\left(\mathbf{y}_i\times\mathbf{n}_i\right). \tag{2.11}$$

The third term of Eq. (2.8) is expanded as

$$\begin{aligned}
2\text{trace}\left[\mathbf{B}^{\mathrm{T}}\left[\mathbf{q}_v\times\right]^2\right] &= 2\text{trace}\left[\sum_{i=1}^{n} a_i \mathbf{y}_i \mathbf{n}_i^{\mathrm{T}} \left[\mathbf{q}_v\times\right]\left[\mathbf{q}_v\times\right]\right] \\
&= 2\text{trace}\left[\sum_{i=1}^{n} a_i \left[\mathbf{q}_v\times\right] \mathbf{n}_i \mathbf{y}_i^{\mathrm{T}} \left[\mathbf{q}_v\times\right]\right] \\
&= 2\text{trace}\left[\sum_{i=1}^{n} a_i \left[\mathbf{y}_i\times\right] \mathbf{q}_v \mathbf{q}_v^{\mathrm{T}} \left[\mathbf{n}_i\times\right]\right] \\
&= \text{trace}\left[\sum_{i=1}^{n} a_i \mathbf{q}_v^{\mathrm{T}} \left[\mathbf{y}_i\times\right]\left[\mathbf{n}_i\times\right] \mathbf{q}_v\right] + \text{trace}\left[\sum_{i=1}^{n} a_i \mathbf{q}_v^{\mathrm{T}} \left[\mathbf{n}_i\times\right]\left[\mathbf{y}_i\times\right] \mathbf{q}_v\right] \\
&= \sum_{i=n}^{n} a_i \mathbf{q}_v^{\mathrm{T}} \left(\left[\mathbf{y}_i\times\right]\left[\mathbf{n}_i\times\right] + \left[\mathbf{n}_i\times\right]\left[\mathbf{y}_i\times\right]\right) \mathbf{q}_v \\
&= \mathbf{q}_v^{\mathrm{T}} \mathbf{H} \mathbf{q}_v
\end{aligned} \tag{2.12}$$

using matrix trace and cross product properties where the matrix $\mathbf{H}$ is defined as

$$\mathbf{H} \equiv \sum_{i=1}^{n} a_i \left(\left[\mathbf{y}_i\times\right]\left[\mathbf{n}_i\times\right] + \left[\mathbf{n}_i\times\right]\left[\mathbf{y}_i\times\right]\right). \tag{2.13}$$

Take note that the matrix is defined in this way such that it is symmetric which will be important in the upcoming eigenvalue problem. The auxiliary function can now be written as

$$\mathcal{G} = \sigma + 2q_4 \mathbf{z}^{\mathrm{T}} \mathbf{q}_v + \mathbf{q}_v^{\mathrm{T}} \mathbf{H} \mathbf{q}_v, \tag{2.14}$$

which is maximized with respect to $\bar{\mathbf{q}}$ therefore σ can be ignored. By defining

$$\mathbf{K} \equiv \begin{bmatrix} \mathbf{H} & \mathbf{z} \\ \mathbf{z}^{\mathrm{T}} & 0 \end{bmatrix}, \tag{2.15}$$

the auxiliary problem is equivalent to maximizing[30]

$$\mathcal{G} = \bar{\mathbf{q}}^{\mathrm{T}} \mathbf{K} \bar{\mathbf{q}} \tag{2.16}$$

subject to the constraint $\bar{\mathbf{q}}^{\mathrm{T}}\bar{\mathbf{q}} = 1$ from the definition of the quaternion. This maximization problem can be solved by the method of Lagrange multipliers. That is,

$$\max_{\bar{\mathbf{q}}} \mathcal{G}^{\star} = \bar{\mathbf{q}}^{\mathrm{T}}\mathbf{K}\bar{\mathbf{q}} - \lambda\left(\bar{\mathbf{q}}^{\mathrm{T}}\bar{\mathbf{q}} - 1\right) \tag{2.17a}$$

$$\frac{\partial \mathcal{G}^{\star}}{\partial \bar{\mathbf{q}}^{\mathrm{T}}} = \mathbf{0} = \left(\mathbf{K} + \mathbf{K}^{\mathrm{T}}\right)\bar{\mathbf{q}} - 2\lambda\bar{\mathbf{q}} = \mathbf{K}\bar{\mathbf{q}} - \lambda\bar{\mathbf{q}} \tag{2.17b}$$

$$\mathbf{K}\bar{\mathbf{q}} = \lambda\bar{\mathbf{q}}. \tag{2.17c}$$

The result of the maximization is the familiar eigenvalue problem where the optimal quaternion is the eigenvector associated with the largest eigenvalue of the matrix $\mathbf{K}$ [10].

$$\mathcal{G} = \bar{\mathbf{q}}^{\mathrm{T}}\mathbf{K}\bar{\mathbf{q}} = \bar{\mathbf{q}}^{\mathrm{T}}\lambda\bar{\mathbf{q}} = \bar{\mathbf{q}}^{\mathrm{T}}\bar{\mathbf{q}}\lambda = \lambda. \tag{2.18}$$

Clearly the largest eigenvalue maximizes the performance index and therefore produces the optimal attitude estimate. Note that the $\mathbf{K}$ matrix derived in this work is slightly different from the one proposed by Davenport and commonly referred to as the Davenport matrix. In [1] Davenport includes σ from Eq. (2.9) in the formulation of the $\mathbf{K}$ matrix. In this work the σ is factored out of the matrix. The resulting solution to the eigenvalue problem is identical using both methods due to the maximization as σ is independent of the quaternion.

2.2 Wahba Problem Covariance Analysis

The estimated attitude matrix may be decomposed into the true attitude and an attitude error $\delta\bar{\mathbf{q}}$ which defined in the body frame rotates from the estimated attitude to the true attitude

$$\mathbf{T}\left(\hat{\bar{\mathbf{q}}}\right) = \mathbf{T}\left(\delta\bar{\mathbf{q}}^{\star}\right)\mathbf{T}\left(\bar{\mathbf{q}}\right), \tag{2.19}$$

where the quaternion conjugate is defined by $\bar{\mathbf{q}}^{\star} = \begin{bmatrix} -\mathbf{q}_v^{\mathrm{T}} & q_4 \end{bmatrix}^{\mathrm{T}}$. With this substitution the performance index for the Wahba problem Eq. (2.1) becomes

$$\mathcal{J}\left(\delta\bar{\mathbf{q}}^{\star}\right) = \frac{1}{2}\sum_{i=1}^{n} a_i \left\|\mathbf{y}_i - \mathbf{T}\left(\delta\bar{\mathbf{q}}^{\star}\right)\left(\mathbf{T}\left(\bar{\mathbf{q}}\right)\mathbf{n}_i\right)\right\|^2, \tag{2.20}$$

the minimization of which produces the attitude error given knowledge of the true attitude. In the ideal case of perfect measurements defined as $\mathbf{y}_i = \mathbf{T}\mathbf{n}_i$, where $\mathbf{T}$ is the true attitude and the vectors are free of error, the minimization of Eq. (2.20) through the q-method again yields

$$\begin{bmatrix} \mathbf{H}_{true} & \mathbf{z}_{true} \\ \mathbf{z}_{true}^{\mathrm{T}} & 0 \end{bmatrix} \delta\bar{\mathbf{q}}^{\star} = \lambda\delta\bar{\mathbf{q}}^{\star}, \tag{2.21}$$

where now

$$\begin{aligned} \mathbf{H}_{true} &= \sum_{i=1}^{n} a_i \left([\mathbf{y}_i\times][\mathbf{T}\mathbf{n}_i\times] + [\mathbf{T}\mathbf{n}_i\times][\mathbf{y}_i\times]\right) \\ &= \sum_{i=1}^{n} a_i \left(\mathbf{y}_i^{\mathrm{T}}\mathbf{y}_i - \mathbf{I}_{3\times3}\right) = 2\sum_{i=1}^{n} a_i [\mathbf{y}_i\times]^2 \end{aligned} \tag{2.22}$$

$$\mathbf{z}_{true} = \sum_{i=1}^{n} a_i \left(\mathbf{y}_i \times \mathbf{T}\mathbf{n}_i\right) = \mathbf{0}. \tag{2.23}$$

$\mathbf{H}_{true}$ always has non-positive eigenvalues and the Davenport matrix is now singular. Therefore, the maximum eigenvalue is zero and the corresponding optimal attitude is given by the identity quaternion defined as $\mathbf{i}_{\bar{\mathbf{q}}} = \begin{bmatrix} \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix}^{\mathrm{T}}$. This result is intuitive as it would be expected that in the absence of errors the attitude error would simply be zero (represented by the identity quaternion) and the estimated attitude would be equivalent to the true attitude.

If errors in the measurements are now reintroduced the measurement model becomes

$$\tilde{\mathbf{y}}_i = \mathbf{T}\mathbf{n}_i + \delta\mathbf{y}_i \tag{2.24a}$$

$$\text{and} \quad \tilde{\mathbf{n}}_i = \mathbf{n}_i + \delta\mathbf{n}_i. \tag{2.24b}$$

The performance index is given by

$$\mathcal{J}\left(\delta\bar{\mathbf{q}}^{\star}\right) = \frac{1}{2}\sum_{i=1}^{n} a_i \left\|\tilde{\mathbf{y}}_i - \mathbf{T}\left(\delta\bar{\mathbf{q}}^{\star}\right)\left(\mathbf{T}\left(\bar{\mathbf{q}}\right)\tilde{\mathbf{n}}_i\right)\right\|^2. \tag{2.25}$$

The auxiliary performance index becomes

$$\mathcal{G} = \delta\bar{\mathbf{q}}^{\star\mathrm{T}} \begin{bmatrix} \mathbf{H}_{\boldsymbol{\theta}} & -\delta\mathbf{z} \\ -\delta\mathbf{z}^{\mathrm{T}} & 0 \end{bmatrix} \delta\bar{\mathbf{q}}^{\star} = \delta\bar{\mathbf{q}}^{\mathrm{T}} \begin{bmatrix} \mathbf{H}_{\boldsymbol{\theta}} & \delta\mathbf{z} \\ \delta\mathbf{z}^{\mathrm{T}} & 0 \end{bmatrix} \delta\bar{\mathbf{q}}, \tag{2.26}$$

where the conjugate error quaternion has been replaced with the error quaternion in the body frame (estimated to true) and

$$\mathbf{H}_{\boldsymbol{\theta}} = \sum_{i=1}^{n} a_i \left(\left[\tilde{\mathbf{y}}_i \times\right]\left[\mathbf{T}\tilde{\mathbf{n}}_i \times\right] + \left[\mathbf{T}\tilde{\mathbf{n}}_i \times\right]\left[\tilde{\mathbf{y}}_i \times\right]\right) \tag{2.27}$$

$$\delta\mathbf{z} = -\sum_{i=1}^{n} a_i \left(\tilde{\mathbf{y}}_i \times \mathbf{T}\tilde{\mathbf{n}}_i\right). \tag{2.28}$$

Note the addition of the negative in Eq. (2.28) arises as a result of replacing the conjugate error quaternion with the error quaternion. Forming the eigenvalue problem as before results in the equation

$$\mathbf{H}_{\boldsymbol{\theta}}\delta\mathbf{q}_v + \delta q_4 \delta\mathbf{z} = \delta\lambda\delta\mathbf{q}_v \tag{2.29}$$

where the optimal eigenvalue $\delta\lambda$ is a small quantity recognizing that in the case of perfect measurements the optimal eigenvalue is zero and the addition of noise results in a nearly singular Davenport matrix. Noting that the errors are small quantities, $\delta q_4 \approx 1$ and $\delta\lambda\delta\mathbf{q}_v \approx 0$, and solving for the estimation error in the form of the vector component of the error quaternion yields

$$\delta\mathbf{q}_v = -\mathbf{H}_{\boldsymbol{\theta}}^{-1}\delta\mathbf{z}. \tag{2.30}$$

The error angle vector is related to the vector component of the error quaternion by $\delta\boldsymbol{\theta} = 2\delta\mathbf{q}_v$. Therefore, the attitude error covariance is given by

$$\mathbf{P}_{\boldsymbol{\theta\theta}} = \mathrm{E}\left\{\delta\boldsymbol{\theta}\delta\boldsymbol{\theta}^{\mathrm{T}}\right\} = 4\mathrm{E}\left\{\delta\mathbf{q}_v\delta\mathbf{q}_v^{\mathrm{T}}\right\} = 4\mathbf{H}_{\boldsymbol{\theta}}^{-1}\mathrm{E}\left\{\delta\mathbf{z}\delta\mathbf{z}^{\mathrm{T}}\right\}\mathbf{H}_{\boldsymbol{\theta}}^{-\mathrm{T}}. \tag{2.31}$$

This expression for the attitude error covariance is equivalent to the result obtained by Shuster [8]. In order to calculate the covariance $\mathrm{E}\left\{\delta\mathbf{z}\delta\mathbf{z}^{\mathrm{T}}\right\}$ the measurement model Eqs. (2.24a) and (2.24b) are substituted for the measurement vectors and a first order approximation is made

$$\begin{aligned}
\delta\mathbf{z} &= -\sum_{i=1}^{n} a_i\left(\tilde{\mathbf{y}}_i \times \mathbf{T}\tilde{\mathbf{n}}_i\right) = -\sum_{i=1}^{n} a_i\left[\left(\mathbf{y}_i + \delta\mathbf{y}_i\right) \times \mathbf{T}\left(\mathbf{n}_i + \delta\mathbf{n}_i\right)\right] \\
&= -\sum_{i=1}^{n} a_i\left(\mathbf{y}_i \times \mathbf{T}\mathbf{n}_i + \mathbf{y}_i \times \mathbf{T}\delta\mathbf{n}_i + \delta\mathbf{y}_i \times \mathbf{T}\mathbf{n}_i + \delta\mathbf{y}_i \times \mathbf{T}\delta\mathbf{n}_i\right) \\
&= -\sum_{i=1}^{n} a_i\left(\mathbf{y}_i \times \mathbf{T}\delta\mathbf{n}_i + \delta\mathbf{y}_i \times \mathbf{T}\mathbf{n}_i\right).
\end{aligned} \tag{2.32}$$

By making the common assumption that each source of error is uncorrelated the resulting covariance is given by

$$\mathrm{E}\left\{\delta\mathbf{z}\delta\mathbf{z}^{\mathrm{T}}\right\} = \sum_{i=1}^{n} a_i^2 \left\{ [\mathbf{y}_i\times] \mathbf{T}\mathrm{E}\left\{\delta\mathbf{n}_i\delta\mathbf{n}_i^{\mathrm{T}}\right\} \mathbf{T}^{\mathrm{T}} [\mathbf{y}_i\times]^{\mathrm{T}} + [\mathbf{T}\mathbf{n}_i\times] \mathrm{E}\left\{\delta\mathbf{y}_i\delta\mathbf{y}_i^{\mathrm{T}}\right\} [\mathbf{T}\mathbf{n}_i\times]^{\mathrm{T}} \right\}. \tag{2.33}$$

In the actual filter the true attitude and measurement vectors are unknown and must be approximated with the estimated and measured values $\hat{\mathbf{T}}$, $\tilde{\mathbf{y}}_i$, and $\tilde{\mathbf{n}}_i$. It is important to note that the minimization of the Wahba problem minimizes the square of the residual $\tilde{\mathbf{y}}_i - \mathbf{T}\left(\hat{\bar{\mathbf{q}}}\right)\tilde{\mathbf{n}}_i$ and not the error quaternion. As a result there is no guarantee that the error quaternion obtained through this method will be minimal, however, it is an effective method for providing an optimal attitude estimate.

2.3 Wahba Problem Weights and the QUEST Measurement Model

The scalar weights of the Wahba problem may be arbitrarily selected as any positive non-zero value. It is therefore desirable to determine an optimal selection for the scalar weights in some sense. A logical selection would be to choose the weights a_i that minimize the trace of the attitude error covariance matrix. The primary goal of the algorithm is to accurately determine the attitude and such a choice would minimize a measure of the total attitude error. However, as noted by Shuster [8], such a choice would have a rather complex dependence on the observation vectors and in the case of the proposed q-method extended Kalman filter an analytic solution does not appear possible. As a simpler alternative Shuster recommends selecting a_i that minimize the performance index of the original Wahba problem when evaluated at the true attitude [8]. This procedure may be developed as follows where the measured or estimated vectors ($\tilde{\mathbf{y}}_i$, $\tilde{\mathbf{n}}_i$) are composed of the true ($\mathbf{y}_i$, $\mathbf{n}_i$) and the error vectors ($\delta\mathbf{y}_i$, $\delta\mathbf{n}_i$). The Wahba problem is first expanded and evaluated at the true attitude $\mathbf{T}$.

$$\begin{aligned}
\mathcal{J} &= \frac{1}{2}\sum_{i=1}^{n} a_i \left\|\tilde{\mathbf{y}}_i - \mathbf{T}\tilde{\mathbf{n}}_i\right\|^2 \\
&= \frac{1}{2}\sum_{i=1}^{n} a_i \left\|\mathbf{y}_i + \delta\mathbf{y}_i - \mathbf{T}\mathbf{n}_i - \mathbf{T}\delta\mathbf{n}_i\right\|^2 \\
&= \frac{1}{2}\sum_{i=1}^{n} a_i \left\|\delta\mathbf{y}_i - \mathbf{T}\delta\mathbf{n}_i\right\|^2 \\
&= \frac{1}{2}\sum_{i=1}^{n} a_i \left(\delta\mathbf{y}_i - \mathbf{T}\delta\mathbf{n}_i\right)^{\mathrm{T}} \left(\delta\mathbf{y}_i - \mathbf{T}\delta\mathbf{n}_i\right). \qquad (2.34)
\end{aligned}$$

Utilizing properties of the trace

$$\mathcal{J} = \mathrm{trace}\left[\mathcal{J}\right] = \frac{1}{2}\mathrm{trace}\left[\sum_{i=1}^{n} a_i \left(\delta\mathbf{y}_i - \mathbf{T}\delta\mathbf{n}_i\right)\left(\delta\mathbf{y}_i - \mathbf{T}\delta\mathbf{n}_i\right)^{\mathrm{T}}\right], \qquad (2.35)$$

expanding and taking the expectation of both sides (assume errors are uncorrelated) gives

$$\mathcal{J} = \mathrm{E}\left\{\mathcal{J}\right\} = \frac{1}{2}\mathrm{trace}\left[\sum_{i=1}^{n} a_i \left(\mathrm{E}\left\{\delta\mathbf{y}_i\delta\mathbf{y}_i^{\mathrm{T}}\right\} + \mathbf{T}\mathrm{E}\left\{\delta\mathbf{n}_i\delta\mathbf{n}_i^{\mathrm{T}}\right\}\mathbf{T}^{\mathrm{T}}\right)\right] \qquad (2.36)$$

The QUEST measurement model [8, 31]defines the measurement error covariances as

$$\mathbf{R}_{\mathbf{yy}} = \sigma_{\mathbf{y}_i}^2\left(\mathbf{I}_{3\times3} - \mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\right) \qquad (2.37)$$

$$\text{and} \quad \mathbf{R}_{\mathbf{nn}} = \sigma_{\mathbf{n}_i}^2\left(\mathbf{I}_{3\times3} - \mathbf{n}_i\mathbf{n}_i^{\mathrm{T}}\right), \qquad (2.38)$$

which arises from the assumption that the error has an axially symmetric distribution about the vector. Shuster explains that for vector sensors with limited fields of view such a distribution is typically accurate and even for sensors with large fields of view the approximation is generally sufficient. Substituting the measurement error the performance index reduces to

$$\begin{aligned}
\mathcal{J} &= \frac{1}{2}\text{trace}\left[\sum_{i=1}^{n} a_i\left(\sigma_{\mathbf{y}_i}^2\left(\mathbf{I}_{3\times3}-\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\right)+\sigma_{\mathbf{n}_i}^2\mathbf{T}\left(\mathbf{I}_{3\times3}-\mathbf{n}_i\mathbf{n}_i^{\mathrm{T}}\right)\mathbf{T}^{\mathrm{T}}\right)\right] \\
&= \frac{1}{2}\text{trace}\left[\sum_{i=1}^{n} a_i\left(\sigma_{\mathbf{y}_i}^2\left(\mathbf{I}_{3\times3}-\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\right)+\sigma_{\mathbf{n}_i}^2\left(\mathbf{I}_{3\times3}-\mathbf{T}\mathbf{n}_i\mathbf{n}_i^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\right)\right)\right] \\
&= \frac{1}{2}\text{trace}\left[\sum_{i=1}^{n} a_i\left(\sigma_{\mathbf{y}_i}^2\left(\mathbf{I}_{3\times3}-\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\right)+\sigma_{\mathbf{n}_i}^2\left(\mathbf{I}_{3\times3}-\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\right)\right)\right] \\
&= \frac{1}{2}\text{trace}\left[\sum_{i=1}^{n} a_i\left(\sigma_{\mathbf{y}_i}^2+\sigma_{\mathbf{n}_i}^2\right)\left(\mathbf{I}_{3\times3}-\mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\right)\right].
\end{aligned} \tag{2.39}$$

By defining the complete measurement error variance as $\sigma_i^2 = \sigma_{\mathbf{y}_i}^2 + \sigma_{\mathbf{n}_i}^2$ and recognizing that for unit vector measurements trace $\left[\mathbf{I}_{3\times3} - \mathbf{y}_i\mathbf{y}_i^{\mathrm{T}}\right] = 3 - 1 = 2$, the performance index may be reduced to a simple function of the scalar weights and the measurement error variance. That is,

$$\mathcal{J} = \sum_{i=1}^{n} a_i\sigma_i^2. \tag{2.40}$$

The performance index of the Wahba problem can be arbitrarily scaled without impacting the attitude estimation. As a result, Shuster employs a common additional constraint by normalizing the scalar weights such that $\sum_{i=1}^{n} a_i = 1$ [8]. Together with the constraint that the weights must be positive, Shuster identifies the selection of a_i

that minimizes the performance index as

$$a_i = \frac{\sigma_{tot}^2}{\sigma_i^2}, \quad \frac{1}{\sigma_{tot}^2} = \sum_{i=1}^{n} \frac{1}{\sigma_i^2}, \tag{2.41}$$

where σ_{tot}^2 is simply a scaling parameter determined by the relationship above in order to meet the constraint. However, the minimization of $\mathcal{J}$ with respect to a_i is ill-posed and the selection posed by Shuster does not actually minimize the performance index. This may be readily shown with a simple example of two measurements. In this case the performance index is given by

$$\mathcal{J} = a_1\sigma_1^2 + a_2\sigma_2^2, \tag{2.42}$$

subject to the constraints

$$a_1 + a_2 = 1 \tag{2.43a}$$

$$\text{and} \quad a_1,\, a_2 > 0. \tag{2.43b}$$

The constraints, may be recast as

$$a_2 = 1 - a_1 \tag{2.44a}$$

$$\text{and} \quad 0 < a_1 < 1. \tag{2.44b}$$

Thus, substituting into the performance index

$$\mathcal{J} = a_1\sigma_1^2 + (1 - a_1)\,\sigma_2^2 \tag{2.45}$$

it becomes clear that as a_1 approaches zero, $\mathcal{J}$ approaches σ_2^2 and as a_1 approaches one, $\mathcal{J}$ approaches σ_1^2 with a linear dependence on a_1. Therefore, assuming $\sigma_1^2 < \sigma_2^2$

the performance index is minimized as a_1 approaches one and not with Shuster's selection where $a_1 = {}^{(\sigma_1^2+\sigma_2^2)}/_{\sigma_1^2}$. While Shuster's selection for the scalar weights does not minimize the performance index of the Wahba problem when evaluated at the true attitude, the selection does intuitively make sense and is not without reason. Shuster also proves that this selection of the scalar weights corresponds to a maximum likelihood estimate of the spacecraft attitude under certain assumptions [9].

This selection is also intuitively pleasing when viewed from the perspective of the general weighted least squares cost function given by

$$\mathcal{J} = \left(\mathbf{y} - \mathbf{H}\mathbf{x}\right)^{\mathrm{T}} \mathbf{W} \left(\mathbf{y} - \mathbf{H}\mathbf{x}\right), \tag{2.46}$$

with measurements $\mathbf{y}$, states $\mathbf{x}$, measurement sensitivity matrix $\mathbf{H}$, and weighting matrix $\mathbf{W}$. It is known that the optimal selection for the weighting matrix is the inverse of the covariance. Recalling the covariance for the QUEST measurement model, it is clear that the selection $a_i = {}^1/_{\sigma_i^2}$ is reminiscent of the inverse of the covariance.

2.4 Incorporation of Initial Attitude Estimate

Often an initial estimate of the attitude is available and is desirable to incorporate it into the attitude determination method. Markley *et al.* [2] developed a method for averaging quaternions that is equivalent in form to the Davenport solution of the

Wahba problem [1]. The average of n quaternions $\bar{\mathbf{q}}_i$ is given by

$$\bar{\mathbf{q}}_{avg} = \min_{\bar{\mathbf{q}}} \sum_{i=1}^{n} \bar{\mathbf{q}}^{\mathrm{T}} \Xi(\bar{\mathbf{q}}_i) \mathbf{A}_i \Xi^{\mathrm{T}}(\bar{\mathbf{q}}_i) \bar{\mathbf{q}}, \tag{2.47}$$

where

$$\Xi(\bar{\mathbf{q}}) = \begin{bmatrix} q_4 \mathbf{I}_{3\times 3} + [\mathbf{q}_v \times] \\ -\mathbf{q}_v^{\mathrm{T}} \end{bmatrix}, \tag{2.48}$$

and $\mathbf{A}$ is a weighting matrix. For the case with a single *a priori* quaternion $\bar{\mathbf{q}}_\circ$ the minimization of Eq. (2.47) is completed as

$$\mathcal{J} = \bar{\mathbf{q}}^{\mathrm{T}} \Xi(\bar{\mathbf{q}}_\circ) \mathbf{A}_\circ \Xi^{\mathrm{T}}(\bar{\mathbf{q}}_\circ) \bar{\mathbf{q}} \tag{2.49a}$$

$$\text{and} \quad \frac{\mathrm{d}\mathcal{J}}{\mathrm{d}\bar{\mathbf{q}}^{\mathrm{T}}} = 0 = \Xi(\bar{\mathbf{q}}_\circ) \mathbf{A}_\circ \Xi^{\mathrm{T}}(\bar{\mathbf{q}}_\circ) \bar{\mathbf{q}}. \tag{2.49b}$$

Noting the property $\Xi^{\mathrm{T}}(\bar{\mathbf{q}}_\circ)\bar{\mathbf{q}}_\circ = 0$ the minimization criterion is met when $\bar{\mathbf{q}} = \bar{\mathbf{q}}_\circ$ and simply returns the *a priori* quaternion.

The original Wahba problem Eq. (2.1) can be augmented with the *a priori* quaternion in order to incorporate the initial attitude information. The resulting augmented performance index to be minimized is given by

$$\mathcal{J}' = \bar{\mathbf{q}}^{\mathrm{T}} \Xi(\bar{\mathbf{q}}_\circ) \mathbf{A}_\circ \Xi^{\mathrm{T}}(\bar{\mathbf{q}}_\circ) \bar{\mathbf{q}} + \frac{1}{2} \sum_{i=1}^{n} a_i \left\| \tilde{\mathbf{y}}_i - \mathbf{T}(\bar{\mathbf{q}}) \tilde{\mathbf{n}}_i \right\|^2. \tag{2.50}$$

The term from the Wahba problem can be rewritten as in Eq. (2.14) and the term σ can be ignored in the minimization as it is not a function of the quaternion. The resulting performance index poses an equivalent maximization problem

$$\mathcal{G}' = -\bar{\mathbf{q}}^{\mathrm{T}} \Xi(\bar{\mathbf{q}}_\circ) \mathbf{A}_\circ \Xi^{\mathrm{T}}(\bar{\mathbf{q}}_\circ) \bar{\mathbf{q}} + \bar{\mathbf{q}}^{\mathrm{T}} \mathbf{K} \bar{\mathbf{q}}, \tag{2.51}$$

where $\mathbf{K}$ is defined by Eq. (2.15). The optimal attitude incorporating an *a priori* attitude estimate is determined in the same manner as the q-method by now solving for the corresponding unit eigenvector of the maximum eigenvalue of the augmented Davenport matrix defined as

$$\mathbf{K}_{aug} = -\Xi(\bar{\mathbf{q}}_\circ)\mathbf{A}_\circ\Xi^{\mathrm{T}}(\bar{\mathbf{q}}_\circ) + \mathbf{K} \tag{2.52}$$

instead of the matrix in Eq. (2.15).

Covariance analysis is applied to the augmented problem in the same manner as for the Wahba problem discussed previously. In the covariance analysis of the Wahba problem the reference vector is replaced with $\mathbf{T}_{true}\mathbf{n}_i$ (where the true quaternion is approximated by the estimated quaternion) so that the minimization produces the error quaternion. In a similar manner the initial quaternion estimate of the first term of Eq. (2.51) is composed with the true quaternion to produce the error quaternion upon minimization

$$\delta\bar{\mathbf{q}}_\circ = \bar{\mathbf{q}}_{true} \otimes \bar{\mathbf{q}}_\circ^\star, \tag{2.53}$$

where the quaternion product is defined as

$$\bar{\mathbf{q}} \otimes \bar{\mathbf{p}} = \begin{bmatrix} q_4\mathbf{p}_v + p_4\mathbf{q}_v - \mathbf{q}_v \times \mathbf{p}_v \\ q_4p_4 - \mathbf{q}_v \cdot \mathbf{p}_v \end{bmatrix}. \tag{2.54}$$

Note that in this manner the quaternion product is defined in the same order as multiplying attitude matrices. In the actual calculation of $\delta\bar{\mathbf{q}}_\circ$ the true quaternion is approximated to first order by the estimated quaternion. Using this substitution the initial error quaternion is the identity quaternion

$$\delta\bar{\mathbf{q}}_\circ = \bar{\mathbf{q}}_\circ \otimes \bar{\mathbf{q}}_\circ^\star = \bar{\mathbf{i}}_\mathbf{q} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^\mathrm{T}. \tag{2.55}$$

Recalling the definition in Eq. (2.48), the first term of Eq. (2.51) can be rewritten as

$$\begin{aligned} -\delta\bar{\mathbf{q}}^\mathrm{T} \begin{bmatrix} \delta q_{4_\circ}\mathbf{I}_{3\times3} + [\delta\mathbf{q}_{v\circ}\times] \\ -\delta\mathbf{q}_{v\circ}^\mathrm{T} \end{bmatrix} \mathbf{A}_\circ \begin{bmatrix} \delta q_{4_\circ}\mathbf{I}_{3\times3} - [\delta\mathbf{q}_{v\circ}\times] & -\delta\mathbf{q}_{v\circ} \end{bmatrix} \delta\bar{\mathbf{q}} & \\ = -\delta\bar{\mathbf{q}}^\mathrm{T} \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2 \\ \mathbf{M}_3 & \mathbf{M}_4 \end{bmatrix} \delta\bar{\mathbf{q}}, & \qquad (2.56a) \end{aligned}$$

$$\begin{aligned} \text{where } \mathbf{M}_1 &= \delta q_{4_\circ}^2\mathbf{A}_\circ + \delta q_{4_\circ}[\delta\mathbf{q}_{v\circ}\times]\mathbf{A}_\circ - \delta q_{4_\circ}\mathbf{A}_\circ[\delta\mathbf{q}_{v\circ}\times] & (2.56b) \\ &\quad - [\delta\mathbf{q}_{v\circ}\times]\mathbf{A}_\circ[\delta\mathbf{q}_{v\circ}\times], & (2.56c) \\ \mathbf{M}_2 &= -\delta q_{4_\circ}\mathbf{A}_\circ\delta\mathbf{q}_{v\circ} - [\delta\mathbf{q}_{v\circ}\times]\mathbf{A}_\circ\delta\mathbf{q}_{v\circ}, & (2.56d) \\ \mathbf{M}_3 &= \mathbf{M}_2^\mathrm{T} = -\delta q_{4_\circ}\delta\mathbf{q}_{v\circ}^\mathrm{T}\mathbf{A}_\circ + \delta\mathbf{q}_{v\circ}^\mathrm{T}\mathbf{A}_\circ[\delta\mathbf{q}_{v\circ}\times], & (2.56e) \\ \text{and } \mathbf{M}_4 &= \delta\mathbf{q}_{v\circ}^\mathrm{T}\mathbf{A}_\circ\delta\mathbf{q}_{v\circ}. & (2.56f) \end{aligned}$$

Again applying a first order approximation $\delta\bar{\mathbf{q}}_\circ = \begin{bmatrix} \delta\mathbf{q}_{v\circ}^\mathrm{T} & 1 \end{bmatrix}^\mathrm{T}$ and higher order terms may be neglected. Eq. (2.51) can now be expressed as

$$g'(\delta\bar{\mathbf{q}}) = \delta\bar{\mathbf{q}}^\mathrm{T} \begin{bmatrix} -\mathbf{A}_\circ + \mathbf{A}_\circ[\delta\mathbf{q}_{v\circ}\times] - [\delta\mathbf{q}_{v\circ}\times]\mathbf{A}_\circ + \mathbf{H}_{\boldsymbol{\theta}} & \mathbf{A}_\circ\delta\mathbf{q}_{v\circ} + \delta\mathbf{z} \\ \delta\mathbf{q}_{v\circ}^\mathrm{T}\mathbf{A}_\circ + \delta\mathbf{z}^\mathrm{T} & 0 \end{bmatrix} \delta\bar{\mathbf{q}}. \tag{2.57}$$

Eq. (2.57) is maximized subject to the constraint $\delta\bar{\mathbf{q}}^\mathbf{T}\delta\bar{\mathbf{q}} = 1$ to produce the familiar eigenvalue problem

$$\begin{bmatrix} -\mathbf{A}_\circ + \mathbf{A}_\circ[\delta\mathbf{q}_{v\circ}\times] - [\delta\mathbf{q}_{v\circ}\times]\mathbf{A}_\circ + \mathbf{H}_{\boldsymbol{\theta}} & \mathbf{A}_\circ\delta\mathbf{q}_{v\circ} + \delta\mathbf{z} \\ \delta\mathbf{q}_{v\circ}^\mathrm{T}\mathbf{A}_\circ + \delta\mathbf{z}^\mathrm{T} & 0 \end{bmatrix} \begin{bmatrix} \delta\mathbf{q}_v \\ 1 \end{bmatrix} = \delta\lambda \begin{bmatrix} \delta\mathbf{q}_v \\ 1 \end{bmatrix}. \tag{2.58}$$

and again recognizing that to first order $\delta\lambda\delta\mathbf{q}_v \approx 0$, the vector component of the error quaternion is given by

$$\delta\mathbf{q}_v = \left(-\mathbf{A}_\circ + \mathbf{A}_\circ\left[\delta\mathbf{q}_{v\circ}\times\right] - \left[\delta\mathbf{q}_{v\circ}\times\right]\mathbf{A}_\circ + \mathbf{H}_{\boldsymbol{\theta}}\right)^{-1}\left(\delta\mathbf{z} - \mathbf{A}_\circ\delta\mathbf{q}_{v\circ}\right) \tag{2.59}$$

which is approximated to first order by

$$\delta\mathbf{q}_v \approx \left(-\mathbf{A}_\circ + \mathbf{H}_{\boldsymbol{\theta}}\right)^{-1}\left(\delta\mathbf{z} - \mathbf{A}_\circ\delta\mathbf{q}_{v\circ}\right). \tag{2.60}$$

Assuming that the errors $\delta\mathbf{q}_{v\circ}$ and $\delta\mathbf{z}$ are uncorrelated and defining

$$\mathbf{K}_\theta = \left(-\mathbf{A}_\circ + \mathbf{H}_{\boldsymbol{\theta}}\right)^{-1} \tag{2.61}$$

the attitude error covariance matrix of the attitude angles expressed in the body frame is obtained in the same manner as Eq. (2.31).

$$\mathbf{P}_{\boldsymbol{\theta\theta}} = \mathbf{K}_\theta\mathbf{A}_\circ\mathbf{P}_{\boldsymbol{\theta\theta}_\circ}\mathbf{A}_\circ^{\mathrm{T}}\mathbf{K}_\theta^{\mathrm{T}} + \mathbf{K}_\theta\mathbf{R}\mathbf{K}_\theta^{\mathrm{T}} \tag{2.62}$$

where

$$\mathbf{R} = 4\mathbb{E}\left\{\delta\mathbf{z}\delta\mathbf{z}^{\mathrm{T}}\right\}. \tag{2.63}$$

Using the following rearrangement

$$\begin{aligned}
\left(-\mathbf{A}_\circ + \mathbf{H}_{\boldsymbol{\theta}}\right)^{-1}\left(-\mathbf{A}_\circ\right) &= \left(-\mathbf{A}_\circ + \mathbf{H}_{\boldsymbol{\theta}}\right)^{-1}\left(-\mathbf{A}_\circ + \mathbf{H}_{\boldsymbol{\theta}} - \mathbf{H}_{\boldsymbol{\theta}}\right) \\
&= \mathbf{I}_{3\times3} - \left(-\mathbf{A}_\circ + \mathbf{H}_{\boldsymbol{\theta}}\right)^{-1}\mathbf{H}_{\boldsymbol{\theta}} \\
&= \mathbf{I} - \mathbf{K}_\theta\mathbf{H}_{\boldsymbol{\theta}} \qquad (2.64)
\end{aligned}$$

Eq. (2.60) is equivalently expressed as

$$\delta\mathbf{q}_v = \left(\mathbf{I} - \mathbf{K}_\theta\mathbf{H}_{\boldsymbol{\theta}}\right)\delta\mathbf{q}_{v\circ} + \mathbf{K}_\theta\delta\mathbf{z} \tag{2.65}$$

The resulting error covariance matrix is given by

$$\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}} = (\mathbf{I} - \mathbf{K}_{\theta}\mathbf{H}_{\boldsymbol{\theta}})\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}_{\circ}}(\mathbf{I} - \mathbf{K}_{\theta}\mathbf{H}_{\boldsymbol{\theta}})^{\mathrm{T}} + \mathbf{K}_{\theta}\mathbf{R}\mathbf{K}_{\theta}^{\mathrm{T}} \tag{2.66}$$

which is in the same form as the Joseph formula for the covariance update, which is used for Kalman filters [32].

Up to this point no choices have been made as to the appropriate selection of the weighting matrix $\mathbf{A}_0$. The goal is to select a weighting matrix so that the initial condition in Eq. (2.50) is added to the Wahba problem term in an equivalent manner. Begin with the measurement model including noise

$$\tilde{\mathbf{y}}_i = \mathbf{T}\mathbf{n}_i + \delta\mathbf{y}_i \tag{2.67a}$$

$$\tilde{\mathbf{n}}_i = \mathbf{n}_i + \delta\mathbf{n}_i \tag{2.67b}$$

and substitute into the Wahba problem term. For simplicity assume that there is no error in the reference measurements ($\delta\mathbf{n}_i = 0$)

$$\begin{aligned}
\frac{1}{2}\sum_{i=1}^{n} a_i \left\|\tilde{\mathbf{y}}_i - \mathbf{T}\tilde{\mathbf{n}}_i\right\|^2 &= \frac{1}{2}\sum_{i=1}^{n} a_i \left\|\mathbf{y}_i + \delta\mathbf{y}_i - \mathbf{T}\mathbf{n}_i\right\|^2 \\
&= \frac{1}{2}\sum_{i=1}^{n} a_i \left\|\delta\mathbf{y}_i\right\|^2 \\
&= \frac{1}{2}\sum_{i=1}^{n} \delta\mathbf{y}_i^{\mathrm{T}} a_i \mathbf{I}_{3\times 3}\delta\mathbf{y}_i
\end{aligned} \tag{2.68}$$

Recall that the scalar weights are selected such that $a_i\mathbf{I}_{3\times 3}$ is the inverse of the measurement covariance $\mathbf{P}_{\mathbf{yy}}^{-1}$. Therefore, recognizing that $\Xi^{\mathrm{T}}(\bar{\mathbf{q}}_{\circ})\bar{\mathbf{q}} = \delta\mathbf{q}_{v\circ}$ Eq. (2.50) may now be expressed as

$$\begin{aligned}\mathcal{J}' &= \delta\mathbf{q}_{v\circ}^{\mathrm{T}}\mathbf{A}_{\circ}\delta\mathbf{q}_{v\circ} + \frac{1}{2}\sum_{i=1}^{n}\delta\mathbf{y}_i^{\mathrm{T}}\mathbf{P}_{\mathbf{yy}}^{-1}\delta\mathbf{y}_i \\ &= \frac{1}{4}\delta\boldsymbol{\theta}^{\mathrm{T}}\mathbf{A}_0\delta\boldsymbol{\theta} + \frac{1}{2}\sum_{i=1}^{n}\delta\mathbf{y}_i^{\mathrm{T}}\mathbf{P}_{\mathbf{yy}}^{-1}\delta\mathbf{y}_i \end{aligned} \tag{2.69}$$

Selecting the weighting matrix as

$$\mathbf{A}_0 = 2\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}_0}^{-1} \tag{2.70}$$

results in an appropriately scaled method of incorporating the initial condition to the Wahba problem.

Chapter 3

The q-Method Extended Kalman Filter

The Kalman filter is the primary tool for recursive estimation and serves as the framework for the proposed algorithm of this thesis. For background, a basic review of the linear Kalman filter update equations is derived. Recall that the q-method extended Kalman filter seeks to update the attitude using the q-method as derived in the previous chapter and then update the non-attitude states using standard Kalman filter techniques. The linear optimal state and covariance update are first derived for a partitioned state where the measurements are linear and only a function of one partition of the state. This linear optimal case is then compared to a Kalman filter where the attitude update is used as the new measurement and is subject to correlated measurement and process noise. The q-method Kalman filter non-attitude state and covariance updates are the extension of the linear optimal case. The qEKF algorithm is then compared to the SOAR filter [3] and shown to be equivalent in the attitude update to second-order and identical in the non-attitude state update.

3.1 Kalman Filter Review

The Kalman filter is the workhorse of on-board estimation and provides the foundation for this thesis. A review of the basic formulation is beneficial in illustrating the feasibility of the q-Method Extended Kalman Filter as it relates to traditional filtering techniques. It is sufficient to begin with linear measurements $\mathbf{y}$ given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta}, \tag{3.1}$$

with state vector $\mathbf{x}$, measurement sensitivity or observation partials matrix $\mathbf{H}$ which maps the states to the measurements, and zero-mean white noise $\boldsymbol{\eta}$ with a corresponding covariance $\mathbf{R}$. In the case of nonlinear measurements where

$$\mathbf{y} = \mathbf{h}\left(\mathbf{x}\right) + \boldsymbol{\eta} \tag{3.2}$$

the measurements are linearized about a nominal state using a first order Taylor series expansion in what is known as the extended Kalman filter which is explained in greater detail in [19, 20, 21]. All Kalman filters are built on a linear assumption so the linear case is sufficient to illustrate the necessary formulation. Define $\hat{\mathbf{x}}^-$ as the unbiased *a priori* estimate of the state with estimation error covariance $\mathbf{P}^-$ and $\hat{\mathbf{x}}^+$

represent the *a posteriori* state estimate with resulting estimation error covariance $\mathbf{P}^+$. The *a priori* and *a posteriori* estimation errors are given by

$$\mathbf{e}^- = \mathbf{x} - \hat{\mathbf{x}}^- \tag{3.3a}$$

$$\text{and} \quad \mathbf{e}^+ = \mathbf{x} - \hat{\mathbf{x}}^+ \tag{3.3b}$$

respectively. The objective of the Kalman filter is to combine an *a priori* estimate of the state and new measurements in some optimal manner in order to produce the best estimate of the state at the current time. For illustrative purposes assume that the *a priori* estimate and the measurement are combined in a linear weighted average to produce the *a posteriori* estimate

$$\hat{\mathbf{x}}^+ = \mathbf{K}_1 \hat{\mathbf{x}}^- + \mathbf{K}\mathbf{y} \tag{3.4}$$

where $\mathbf{K}_1$ and $\mathbf{K}$ are constant matrices to be determined. In the case of a perfect (true) *a priori* state and measurement, such that $\hat{\mathbf{x}}^- = \mathbf{x}$ and $\mathbf{y} = \mathbf{H}\mathbf{x}$, the update should also produce the true state. Making the appropriate substitutions and solving for $\mathbf{K}_1$

$$\mathbf{x} = \mathbf{K}_1 \mathbf{x} + \mathbf{K}\mathbf{H}\mathbf{x} \tag{3.5a}$$

$$\text{and} \quad \mathbf{K}_1 = \mathbf{I} - \mathbf{K}\mathbf{H} \tag{3.5b}$$

where $\mathbf{I}$ is the identity matrix. The state update equation may now be rewritten

$$\begin{aligned} \hat{\mathbf{x}}^+ &= (\mathbf{I} - \mathbf{K}\mathbf{H})\,\hat{\mathbf{x}}^- + \mathbf{K}\mathbf{y} \\ &= \hat{\mathbf{x}}^- + \mathbf{K}\left(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}^-\right). \end{aligned} \tag{3.6}$$

The *a posteriori* estimation error may now be expressed in terms of the *a priori* values.

$$\begin{aligned}\mathbf{e}^{+} &= \mathbf{x}-\hat{\mathbf{x}}^{+}=\mathbf{x}-\hat{\mathbf{x}}^{-}-\mathbf{K}\left(\mathbf{H}\mathbf{x}+\boldsymbol{\eta}-\mathbf{H}\hat{\mathbf{x}}^{-}\right)\\ &= \mathbf{e}^{-}-\mathbf{K}\left(\mathbf{H}\mathbf{e}^{-}+\boldsymbol{\eta}\right)\\ &= \left(\mathbf{I}-\mathbf{K}\mathbf{H}\right)\mathbf{e}^{-}-\mathbf{K}\boldsymbol{\eta} \end{aligned} \tag{3.7}$$

Recall that the estimation errors are zero mean and the *a priori* estimation error covariance is defined as

$$\mathbf{P}^{-}=\mathrm{E}\left\{\mathbf{e}^{-}\left(\mathbf{e}^{-}\right)^{\mathrm{T}}\right\} \tag{3.8}$$

It is typically assumed that the estimation errors and measurement noise are uncorrelated therefore $\mathrm{E}\left\{\mathbf{e}^{-}\boldsymbol{\eta}^{\mathrm{T}}\right\}=0$, as well as for the transpose. Therefore the *a posteriori* estimation error covariance is determined by

$$\mathbf{P}^{+}=\mathrm{E}\left\{\mathbf{e}^{+}\left(\mathbf{e}^{+}\right)^{\mathrm{T}}\right\}=\left(\mathbf{I}-\mathbf{K}\mathbf{H}\right)\mathbf{P}^{-}\left(\mathbf{I}-\mathbf{K}\mathbf{H}\right)^{\mathrm{T}}+\mathbf{K}\mathbf{R}\mathbf{K}^{\mathrm{T}}, \tag{3.9}$$

which is known as the Joseph formula for the covariance update [32]. At the expense of extra operations the Joseph formula guarantees a symmetric positive semi-definite covariance as opposed to alternative derivations of the covariance update. It is important to note that the Kalman gain, $\mathbf{K}$, has not yet be defined and the Joseph covariance update is valid for any choice of $\mathbf{K}$. This fact is key to the development of the qEKF. In a traditional Kalman filter the optimal gain $\mathbf{K}$ is selected so as to minimize the estimation error. This is accomplished by minimizing the trace of the estimation error covariance update with respect to the gain $\mathbf{K}$.

$$\min_{\mathbf{K}} \text{trace}\left[\mathbf{P}^{+}\right] = \text{trace}\left[\left(\mathbf{I}-\mathbf{KH}\right)\mathbf{P}^{-}\left(\mathbf{I}-\mathbf{KH}\right)^{\mathrm{T}}+\mathbf{KRK}^{\mathrm{T}}\right] \tag{3.10}$$

$$\begin{aligned}
\frac{\partial \text{trace}\left[\mathbf{P}^{+}\right]}{\partial \mathbf{K}} &= 0 = \frac{\partial}{\partial \mathbf{K}}\text{trace}\left[\left(\mathbf{I}-\mathbf{KH}\right)\mathbf{P}^{-}\left(\mathbf{I}-\mathbf{KH}\right)^{\mathrm{T}}+\mathbf{KRK}^{\mathrm{T}}\right] \\
0 &= \frac{\partial}{\partial \mathbf{K}}\text{trace}\left[\mathbf{P}^{-}\right]+\text{trace}\left[-\mathbf{KHP}^{-}\right]+\text{trace}\left[-\mathbf{P}^{-}\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\right] \\
&\quad +\text{trace}\left[\mathbf{KHP}^{-}\mathbf{H}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\right]+\text{trace}\left[\mathbf{KRK}^{\mathrm{T}}\right] \\
0 &= -2\mathbf{P}^{-}\mathbf{H}^{\mathrm{T}}+2\mathbf{KHP}^{-}\mathbf{H}^{\mathrm{T}}+2\mathbf{KR} \qquad (3.11) \\
\mathbf{K}\left(\mathbf{HP}^{-}\mathbf{H}^{\mathrm{T}}+\mathbf{R}\right) &= \mathbf{P}^{-}\mathbf{H}^{\mathrm{T}} \\
\mathbf{K}_{opt} &= \mathbf{P}^{-}\mathbf{H}^{\mathrm{T}}\left(\mathbf{HP}^{-}\mathbf{H}^{\mathrm{T}}+\mathbf{R}\right)^{-1}. \qquad (3.12)
\end{aligned}$$

Note that $\mathbf{P}^{-}$ and $\mathbf{R}$ are both covariance matrices and therefore symmetric. The expression in Eq. (3.12) is given as the optimal Kalman gain in references of Kalman filtering as it results in the optimal updated state by minimizing the *a posteriori* estimation error.

3.2 q-Method EKF

3.2.1 Partitioned State Update Equations

In the development of the qEKF it is necessary to separate the attitude from the non-attitude states in the Kalman filter. For reference, the linear optimal update equations for a standard Kalman filter are derived for a partitioned state

$$\mathbf{x} = \begin{bmatrix} \boldsymbol{\theta} \\ \mathbf{s} \end{bmatrix}, \tag{3.13}$$

where $\boldsymbol{\theta}$ corresponds to the three component attitude state and $\mathbf{s}$ are the remaining non-attitude states. Again assuming linear measurements $\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta}$, the observation partials, gain, and covariance matrices are similarly partitioned

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{\boldsymbol{\theta}} & \mathbf{H}_{\mathbf{s}} \end{bmatrix}, \tag{3.14}$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_{\boldsymbol{\theta}} \\ \mathbf{K}_{\mathbf{s}} \end{bmatrix}, \tag{3.15}$$

$$\text{and} \quad \mathbf{P} = \begin{bmatrix} \mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}} & \mathbf{P}_{\boldsymbol{\theta}\mathbf{s}} \\ \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}} & \mathbf{P}_{\mathbf{ss}} \end{bmatrix}. \tag{3.16}$$

For this work, it is assumed that the measurements are only a function of the attitude, $\boldsymbol{\theta}$, and independent of all non-attitude states, $\mathbf{s}$. Recalling that the observation partials relates the measurements to the states by $\mathbf{H} = {}^{\partial \mathbf{y}}/_{\partial \mathbf{x}}$, $\mathbf{H}_{\mathbf{s}} = \mathbf{0}$ by definition. The optimal Kalman gain of Eq. (3.12) is partitioned as

$$\mathbf{K}_{opt} = \begin{bmatrix} \mathbf{K}_{\boldsymbol{\theta},opt} \\ \mathbf{K}_{\mathbf{s},opt} \end{bmatrix} = \begin{bmatrix} \mathbf{P}^{-}_{\boldsymbol{\theta}\boldsymbol{\theta}}\mathbf{H}^{\mathrm{T}}_{\boldsymbol{\theta}} \\ \mathbf{P}^{-}_{\mathbf{s}\boldsymbol{\theta}}\mathbf{H}^{\mathrm{T}}_{\boldsymbol{\theta}} \end{bmatrix} \mathbf{W}^{-1}, \tag{3.17}$$

where

$$\mathbf{W} = \mathbf{H}_{\boldsymbol{\theta}}\mathbf{P}^{-}_{\boldsymbol{\theta}\boldsymbol{\theta}}\mathbf{H}_{\boldsymbol{\theta}}{}^{\mathrm{T}} + \mathbf{R}. \tag{3.18}$$

Following Eq. (3.6) the attitude partition of the update is given by

$$\hat{\boldsymbol{\theta}}^{+} = \hat{\boldsymbol{\theta}}^{-} + \mathbf{K}_{\boldsymbol{\theta}}\left(\mathbf{y} - \mathbf{H}_{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}^{-}\right) \tag{3.19}$$

and define the residual as

$$\boldsymbol{\epsilon} = \mathbf{y} - \mathbf{H}_{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}^{-}. \tag{3.20}$$

The residual may be determined as a function of the attitude update and the Kalman gain (for this work the optimal gain of Eq. (3.17) is used)

$$\boldsymbol{\epsilon} = \mathbf{K}_{\boldsymbol{\theta},opt}^{-1}\Delta\boldsymbol{\theta} \tag{3.21}$$

where $\Delta\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^{+} - \hat{\boldsymbol{\theta}}^{-}$ is the optimal attitude update. Switching now to the non-attitude states partition, the state update is

$$\hat{\mathbf{s}}^{+} = \hat{\mathbf{s}}^{-} + \mathbf{K}_{\mathbf{s}}\left(\mathbf{y} - \mathbf{H}_{\boldsymbol{\theta}}\hat{\boldsymbol{\theta}}^{-}\right). \tag{3.22}$$

Substitute the optimal gain from Eq. (3.17) and the alternative expression of the residual from Eq. (3.21) to produce an alternative equation for the optimal non-attitude state update in the form

$$\begin{aligned} \hat{\mathbf{s}}^{+} &= \hat{\mathbf{s}}^{-} + \mathbf{K}_{\mathbf{s},opt}\mathbf{K}_{\boldsymbol{\theta},opt}^{-1}\Delta\boldsymbol{\theta} && (3.23a) \\ &= \hat{\mathbf{s}}^{-} + \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}^{-}\left(\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-}\right)^{-1}\Delta\boldsymbol{\theta}. && (3.23b) \end{aligned}$$

For the covariance analysis the same method as given in section (3.1) may be used. The attitude and non-attitude state estimation errors are given by

$$\begin{aligned} \mathbf{e}_{\boldsymbol{\theta}}^{+} &= \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}}\mathbf{H}_{\boldsymbol{\theta}}\right)\mathbf{e}_{\boldsymbol{\theta}}^{-} - \mathbf{K}_{\boldsymbol{\theta}}\boldsymbol{\eta} && (3.24a) \\ \text{and} \quad \mathbf{e}_{\mathbf{s}}^{+} &= \left(\mathbf{I} - \mathbf{K}_{\mathbf{s}}\mathbf{H}_{\mathbf{s}}\right)\mathbf{e}_{\mathbf{s}}^{-} - \mathbf{K}_{\mathbf{s}}\boldsymbol{\eta}. && (3.24b) \end{aligned}$$

The error covariance update for the full state is given by the Joseph formula given in

Eq. (3.9) and repeated

$$\mathbf{P}^{+} = \mathrm{E}\left\{\mathbf{e}^{+}\left(\mathbf{e}^{+}\right)^{\mathrm{T}}\right\} = \left(\mathbf{I} - \mathbf{K}\mathbf{H}\right)\mathbf{P}^{-}\left(\mathbf{I} - \mathbf{K}\mathbf{H}\right)^{\mathrm{T}} + \mathbf{K}\mathbf{R}\mathbf{K}^{\mathrm{T}}$$

Partitioning the full state covariance update using Eqs. (3.14), (3.15), and (3.16) results in

$$\mathbf{P}^{+}_{\boldsymbol{\theta\theta}} = \mathbf{P}^{-}_{\boldsymbol{\theta\theta}} - \mathbf{K}_{\boldsymbol{\theta}}\mathbf{H}_{\boldsymbol{\theta}}\mathbf{P}^{-}_{\boldsymbol{\theta\theta}} - \mathbf{P}^{-}_{\boldsymbol{\theta\theta}}\mathbf{H}^{\mathrm{T}}_{\boldsymbol{\theta}}\mathbf{K}^{\mathrm{T}}_{\boldsymbol{\theta}} + \mathbf{K}_{\boldsymbol{\theta}}\mathbf{W}\mathbf{K}^{\mathrm{T}}_{\boldsymbol{\theta}}, \tag{3.25a}$$

$$\mathbf{P}^{+}_{\boldsymbol{\theta}\mathbf{s}} = \mathbf{P}^{-}_{\boldsymbol{\theta}\mathbf{s}} - \mathbf{K}_{\boldsymbol{\theta}}\mathbf{H}_{\boldsymbol{\theta}}\mathbf{P}^{-}_{\boldsymbol{\theta}\mathbf{s}} - \mathbf{P}^{-}_{\boldsymbol{\theta\theta}}\mathbf{H}^{\mathrm{T}}_{\boldsymbol{\theta}}\mathbf{K}^{\mathrm{T}}_{\mathbf{s}} + \mathbf{K}_{\boldsymbol{\theta}}\mathbf{W}\mathbf{K}^{\mathrm{T}}_{\mathbf{s}}, \tag{3.25b}$$

$$\mathbf{P}^{+}_{\mathbf{s}\boldsymbol{\theta}} = \mathbf{P}^{-}_{\mathbf{s}\boldsymbol{\theta}} - \mathbf{K}_{\mathbf{s}}\mathbf{H}_{\boldsymbol{\theta}}\mathbf{P}^{-}_{\boldsymbol{\theta\theta}} - \mathbf{P}^{-}_{\mathbf{s}\boldsymbol{\theta}}\mathbf{H}^{\mathrm{T}}_{\boldsymbol{\theta}}\mathbf{K}^{\mathrm{T}}_{\boldsymbol{\theta}} + \mathbf{K}_{\mathbf{s}}\mathbf{W}\mathbf{K}^{\mathrm{T}}_{\boldsymbol{\theta}}, \tag{3.25c}$$

$$\text{and} \quad \mathbf{P}^{+}_{\mathbf{ss}} = \mathbf{P}^{-}_{\mathbf{ss}} - \mathbf{K}_{\mathbf{s}}\mathbf{H}_{\boldsymbol{\theta}}\mathbf{P}^{-}_{\boldsymbol{\theta}\mathbf{s}} - \mathbf{P}^{-}_{\mathbf{s}\boldsymbol{\theta}}\mathbf{H}^{\mathrm{T}}_{\boldsymbol{\theta}}\mathbf{K}^{\mathrm{T}}_{\mathbf{s}} + \mathbf{K}_{\mathbf{s}}\mathbf{W}\mathbf{K}^{\mathrm{T}}_{\mathbf{s}}. \tag{3.25d}$$

Substituting the optimal gains from Eq. (3.17) further simplifies the partitioned covariance updates to

$$\begin{aligned}\mathbf{P}^{+}_{\boldsymbol{\theta\theta}} &= \mathbf{P}^{-}_{\boldsymbol{\theta\theta}} - \mathbf{K}_{\boldsymbol{\theta},opt}\mathbf{W}\mathbf{K}^{\mathrm{T}}_{\boldsymbol{\theta},opt} \\ &= \mathbf{P}^{-}_{\boldsymbol{\theta\theta}}\left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta},opt}\mathbf{H}_{\boldsymbol{\theta}}\right)^{\mathrm{T}}, \end{aligned} \tag{3.26a}$$

$$\mathbf{P}^{+}_{\boldsymbol{\theta}\mathbf{s}} = \mathbf{P}^{-}_{\boldsymbol{\theta}\mathbf{s}} - \mathbf{P}^{-}_{\boldsymbol{\theta\theta}}\mathbf{H}^{\mathrm{T}}_{\boldsymbol{\theta}}\mathbf{K}^{\mathrm{T}}_{\mathbf{s},opt}, \tag{3.26b}$$

$$\mathbf{P}^{+}_{\mathbf{s}\boldsymbol{\theta}} = \mathbf{P}^{-}_{\mathbf{s}\boldsymbol{\theta}} - \mathbf{K}_{\mathbf{s},opt}\mathbf{H}_{\boldsymbol{\theta}}\mathbf{P}^{-}_{\boldsymbol{\theta\theta}}, \tag{3.26c}$$

$$\text{and} \quad \mathbf{P}^{+}_{\mathbf{ss}} = \mathbf{P}^{-}_{\mathbf{ss}} - \mathbf{K}_{\mathbf{s},opt}\mathbf{W}\mathbf{K}^{\mathrm{T}}_{\mathbf{s},opt}. \tag{3.26d}$$

3.2.2 Attitude Measurement Update Equations

Returning to the main objective of this thesis, the qEKF seeks to integrate the Wahba problem into the framework of a Kalman filter capable of estimating the attitude and other states. In Section 2.4 an optimal update for the attitude and associated covariance in the presence of new measurements was derived as the nonlinear solution to the Wahba problem incorporating the *a priori* attitude information. After updating the attitude, the remaining task for the Kalman filter under development is to determine the appropriate update for the non-attitude states. This is accomplished by updating the non-attitude states from the corresponding attitude update which appropriately incorporates the new information gained from the new measurements. As a result, consider the case of partitioned update equations above except now using the updated attitude, $\hat{\boldsymbol{\theta}}^+$, as the measurement. With this new pseudo-measurement (pseudo so as not to be confused with processing the updated attitude as an entirely new measurement), the attitude observation partial is identity $\mathbf{H}_{\boldsymbol{\theta}}^\star = \mathbf{I}_{3\times 3}$ and the non-attitude state observation partial is zero, $\mathbf{H}_{\mathbf{s}}^\star = \mathbf{0}$. Therefore, the pseudo-measurement is given by

$$\mathbf{y}^\star = \hat{\boldsymbol{\theta}}^+ = \boldsymbol{\theta} + \boldsymbol{\eta}^\star \tag{3.27}$$

$$\text{and} \quad \boldsymbol{\eta}^\star = -\mathbf{e}_{\boldsymbol{\theta}}^+. \tag{3.28}$$

Note that for the remainder of this section quantities with the superscript $(\cdot)^\star$ correspond to the pseudo-measurement.

In the standard Kalman filter it is assumed that the process noise and the measurement noise are uncorrelated. However, in the case of this pseudo-measurement the measurement noise is the *a posteriori* attitude estimation error, $-\mathbf{e}_{\boldsymbol{\theta}}^+$, which is certainly correlated with the *a priori* estimation errors, $\mathbf{e}_{\boldsymbol{\theta}}^-$ and $\mathbf{e}_{\mathbf{s}}^-$. This additional correlation must be properly accounted for in the derivation of the update equations and for the partitioned state the correlation covariance matrix is defined as

$$\mathbf{C}^{\star} = \begin{bmatrix} \mathbf{C}^{\star}_{\boldsymbol{\theta}} \\ \mathbf{C}^{\star}_{\mathbf{s}} \end{bmatrix} = \begin{bmatrix} \mathrm{E}\left\{\mathbf{e}^{-}_{\boldsymbol{\theta}} \left(\boldsymbol{\eta}^{\star}\right)^{\mathrm{T}}\right\} \\ \mathrm{E}\left\{\mathbf{e}^{-}_{\mathbf{s}} \left(\boldsymbol{\eta}^{\star}\right)^{\mathrm{T}}\right\} \end{bmatrix}. \tag{3.29}$$

Substitute for $\mathbf{e}^{+}_{\boldsymbol{\theta}}$ from Eq. (3.24a) for $\boldsymbol{\eta}^{\star}$

$$\begin{aligned} \mathbf{C}^{\star}_{\boldsymbol{\theta}} &= -\mathrm{E}\left\{\mathbf{e}^{-}_{\boldsymbol{\theta}} \left(\mathbf{e}^{+}_{\boldsymbol{\theta}}\right)^{\mathrm{T}}\right\} \\ &= -\mathrm{E}\left\{\mathbf{e}^{-}_{\boldsymbol{\theta}} \left(\mathbf{e}^{-}_{\boldsymbol{\theta}}\right)^{\mathrm{T}}\right\} \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}}\mathbf{H}_{\boldsymbol{\theta}}\right)^{\mathrm{T}} \\ &= -\mathbf{P}^{-}_{\boldsymbol{\theta\theta}} \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}}\mathbf{H}_{\boldsymbol{\theta}}\right)^{\mathrm{T}} \end{aligned} \tag{3.30a}$$

and

$$\begin{aligned} \mathbf{C}^{\star}_{\mathbf{s}} &= -\mathrm{E}\left\{\mathbf{e}^{-}_{\mathbf{s}} \left(\mathbf{e}^{+}_{\boldsymbol{\theta}}\right)^{\mathrm{T}}\right\} \\ &= -\mathrm{E}\left\{\mathbf{e}^{-}_{\mathbf{s}} \left(\mathbf{e}^{-}_{\boldsymbol{\theta}}\right)^{\mathrm{T}}\right\} \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}}\mathbf{H}_{\boldsymbol{\theta}}\right)^{\mathrm{T}} \\ &= -\mathbf{P}^{-}_{\mathbf{s}\boldsymbol{\theta}} \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}}\mathbf{H}_{\boldsymbol{\theta}}\right)^{\mathrm{T}}, \end{aligned} \tag{3.30b}$$

where the process noise and the measurement noise, $\boldsymbol{\eta}$ not to be confused with $\boldsymbol{\eta}^{\star}$, remain uncorrelated. Recalling the optimal attitude covariance update from Eq. (3.26a)

$$\left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta},opt}\mathbf{H}_{\boldsymbol{\theta}}\right)^{\mathrm{T}} = \left(\mathbf{P}^{-}_{\boldsymbol{\theta\theta}}\right)^{-1} \mathbf{P}^{+}_{\boldsymbol{\theta\theta}}, \tag{3.31}$$

with the assumption of an optimal attitude gain, $\mathbf{K}_{\boldsymbol{\theta},opt}$, the correlation covariance matrices are simplified to

$$\mathbf{C}^{\star}_{\boldsymbol{\theta}} = -\mathbf{P}^{+}_{\boldsymbol{\theta\theta}} \tag{3.32a}$$

$$\text{and} \quad \mathbf{C}^{\star}_{\mathbf{s}} = -\mathbf{P}^{-}_{\mathbf{s}\boldsymbol{\theta}} \left(\mathbf{P}^{-}_{\boldsymbol{\theta\theta}}\right)^{-1} \mathbf{P}^{+}_{\boldsymbol{\theta\theta}}. \tag{3.32b}$$

Taking into account this additional correlation, the same procedure as in Section 3.1 can be used to derive the covariance update equation as well as the optimal gain which minimizes the trace of the updated covariance. A detailed derivation of these equations is presented in [20]. The resulting optimal gain is given by

$$\mathbf{K}^\star = \left(\mathbf{P}^-\mathbf{H}^{\star\mathrm{T}} + \mathbf{C}^\star\right)\left(\mathbf{H}^\star\mathbf{P}^-\mathbf{H}^{\star\mathrm{T}} + \mathbf{R}^\star + \mathbf{H}^\star\mathbf{C}^\star + \mathbf{C}^{\star\mathrm{T}}\mathbf{H}^{\star\mathrm{T}}\right)^{-1}. \tag{3.33}$$

For the partitioned case with $\mathbf{H}^\star$ selected as previously determined according to the pseudo-measurement model

$$\mathbf{H}^\star = \begin{bmatrix}\mathbf{H}^\star_{\boldsymbol{\theta}} & \mathbf{H}^\star_{\mathbf{s}}\end{bmatrix} = \begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix} \tag{3.34}$$

and the pseudo-measurement noise

$$\mathbf{R}^\star = \mathrm{E}\left\{\boldsymbol{\eta}^\star\left(\boldsymbol{\eta}^\star\right)^{\mathrm{T}}\right\} = \mathbf{P}^+_{\boldsymbol{\theta\theta}}. \tag{3.35}$$

The partitioned optimal gains are given by

$$\mathbf{K}^\star = \begin{bmatrix}\mathbf{K}^\star_{\boldsymbol{\theta}} \\ \mathbf{K}^\star_{\mathbf{s}}\end{bmatrix}, \tag{3.36a}$$

$$\text{where}\quad \mathbf{K}^\star_{\boldsymbol{\theta}} = \left(\mathbf{P}^-_{\boldsymbol{\theta\theta}} + \mathbf{C}^\star_{\boldsymbol{\theta}}\right)\left(\mathbf{P}^-_{\boldsymbol{\theta\theta}} + \mathbf{P}^+_{\boldsymbol{\theta\theta}} + \mathbf{C}^\star_{\boldsymbol{\theta}} + \mathbf{C}^{\star\mathrm{T}}_{\boldsymbol{\theta}}\right)^{-1} \tag{3.36b}$$

$$\text{and}\quad \mathbf{K}^\star_{\mathbf{s}} = \left(\mathbf{P}^-_{\mathbf{s}\boldsymbol{\theta}} + \mathbf{C}^\star_{\mathbf{s}}\right)\left(\mathbf{P}^-_{\boldsymbol{\theta\theta}} + \mathbf{P}^+_{\boldsymbol{\theta\theta}} + \mathbf{C}^\star_{\boldsymbol{\theta}} + \mathbf{C}^{\star\mathrm{T}}_{\boldsymbol{\theta}}\right)^{-1}. \tag{3.36c}$$

By substituting for $\mathbf{C}^\star_{\boldsymbol{\theta}}$ from Eq. (3.32a), $\mathbf{K}^\star_{\boldsymbol{\theta}}$ is reduced to

$$\begin{aligned}\mathbf{K}^\star_{\boldsymbol{\theta}} &= \left(\mathbf{P}^-_{\boldsymbol{\theta\theta}} - \mathbf{P}^+_{\boldsymbol{\theta\theta}}\right)\left(\mathbf{P}^-_{\boldsymbol{\theta\theta}} + \mathbf{P}^+_{\boldsymbol{\theta\theta}} - \mathbf{P}^+_{\boldsymbol{\theta\theta}} - \mathbf{P}^+_{\boldsymbol{\theta\theta}}\right)^{-1} \\ &= \mathbf{I}.\end{aligned} \tag{3.37}$$

This result is intuitively pleasing as the attitude has already been updated through the q-method. The use of $\hat{\boldsymbol{\theta}}^+$ as a pseudo-measurement provides no additional information pertaining to the attitude and should result in no additional update from the Kalman filter. The resulting attitude state update from the Kalman filter is given by

$$\begin{aligned}
\hat{\boldsymbol{\theta}}^\star &= \hat{\boldsymbol{\theta}}^- + \mathbf{K}_{\boldsymbol{\theta}}^\star\left(\mathbf{y}^\star - \hat{\boldsymbol{\theta}}^-\right) \\
&= \hat{\boldsymbol{\theta}}^- + \mathbf{K}_{\boldsymbol{\theta}}^\star\left(\hat{\boldsymbol{\theta}}^+ - \hat{\boldsymbol{\theta}}^-\right) \\
&= \hat{\boldsymbol{\theta}}^+.
\end{aligned} \tag{3.38}$$

The optimal $\mathbf{K}_{\mathbf{s}}^\star$ is reduced in an identical manner by substituting the appropriate values for $\mathbf{C}_{\boldsymbol{\theta}}^\star$, $\mathbf{C}_{\mathbf{s}}^\star$, and $\mathbf{P}_{\boldsymbol{\theta\theta}}^+$ from Eqs. (3.32a), (3.32b), and (3.26a).

$$\begin{aligned}
\mathbf{K}_{\mathbf{s}}^\star &= \left(\mathbf{P}_{\boldsymbol{\theta\theta}}^- - \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}^-\left(\mathbf{P}_{\boldsymbol{\theta\theta}}^-\right)^{-1}\mathbf{P}_{\boldsymbol{\theta\theta}}^+\right)\left(\mathbf{P}_{\boldsymbol{\theta\theta}}^- + \mathbf{P}_{\boldsymbol{\theta\theta}}^+ - \mathbf{P}_{\boldsymbol{\theta\theta}}^+ - \mathbf{P}_{\boldsymbol{\theta\theta}}^+\right)^{-1} \\
&= \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}^-\left[\mathbf{I} - \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta},opt}\mathbf{H}_{\boldsymbol{\theta}}\right)^{\mathrm{T}}\right]\left[\mathbf{P}_{\boldsymbol{\theta\theta}}^-\left(\mathbf{I} - \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta},opt}\mathbf{H}_{\boldsymbol{\theta}}\right)^{\mathrm{T}}\right)\right]^{-1} \\
&= \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}^-\mathbf{H}_{\boldsymbol{\theta}}^{\mathrm{T}}\mathbf{K}_{\boldsymbol{\theta},opt}^{\mathrm{T}}\left(\mathbf{P}_{\boldsymbol{\theta\theta}}^-\mathbf{H}_{\boldsymbol{\theta}}^{\mathrm{T}}\mathbf{K}_{\boldsymbol{\theta},\mathbf{opt}}^{\mathrm{T}}\right)^{-1} \\
&= \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}^-\left(\mathbf{P}_{\boldsymbol{\theta\theta}}^-\right)^{-1}.
\end{aligned} \tag{3.39}$$

The non-attitude state update is given by

$$\begin{aligned}
\hat{\mathbf{s}}^\star &= \hat{\mathbf{s}}^- + \mathbf{K}_{\mathbf{s}}^\star\left(\mathbf{y}^\star - \hat{\boldsymbol{\theta}}^-\right) \\
&= \hat{\mathbf{s}}^- + \mathbf{K}_{\mathbf{s}}^\star\left(\hat{\boldsymbol{\theta}}^+ - \hat{\boldsymbol{\theta}}^-\right) \\
&= \hat{\mathbf{s}}^- + \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}^-\left(\mathbf{P}_{\boldsymbol{\theta\theta}}^-\right)^{-1}\Delta\boldsymbol{\theta}.
\end{aligned} \tag{3.40}$$

Assuming that the residual for this case is the optimal attitude update, the non-attitude state update for the pseudo-measurement is identical to the linear optimal state update given in Eq. (3.23b). In summary, the pseudo-measurement state update Eqs. (3.38) and (3.40) are equivalent to the linear optimal update Eqs. (3.19) and (3.22) when the gains are chosen as

$$\mathbf{K}_{\boldsymbol{\theta}}^{\star} = \mathbf{I} \tag{3.41a}$$

$$\text{and} \quad \mathbf{K}_{\mathbf{s}}^{\star} = \mathbf{K}_{\mathbf{s},opt}\mathbf{K}_{\boldsymbol{\theta},opt}^{-1}. \tag{3.41b}$$

The estimation error is calculated in like manner as before, this time beginning with the full state

$$\mathbf{e}^{\star} = \left(\mathbf{I} - \mathbf{K}^{\star}\mathbf{H}^{\star}\right)\mathbf{e}^{-} - \mathbf{K}^{\star}\boldsymbol{\eta}^{\star} \tag{3.42}$$

where

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_{\boldsymbol{\theta}} \\ \\ \mathbf{e}_{\mathbf{s}} \end{bmatrix}. \tag{3.43}$$

The partitioned equations are

$$\mathbf{e}_{\boldsymbol{\theta}}^{\star} = \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta}}^{\star}\right)\mathbf{e}_{\boldsymbol{\theta}}^{-} + \mathbf{K}_{\boldsymbol{\theta}}^{\star}\mathbf{e}_{\boldsymbol{\theta}}^{+} \tag{3.44a}$$

$$\text{and} \quad \mathbf{e}_{\mathbf{s}}^{\star} = \mathbf{e}_{\mathbf{s}}^{-} - \mathbf{K}_{\mathbf{s}}^{\star}\mathbf{e}_{\boldsymbol{\theta}}^{-} + \mathbf{K}_{\mathbf{s}}^{\star}\mathbf{e}_{\boldsymbol{\theta}}^{+}. \tag{3.44b}$$

Note that the measurement noise for this case is $-\mathbf{e}_{\boldsymbol{\theta}}^{+}$ and unlike the previous examples, this measurement noise is clearly correlated with the process noise $\left(\mathbf{e}_{\boldsymbol{\theta}}^{-} \text{ and } \mathbf{e}_{\mathbf{s}}^{-}\right)$. As a result, the covariance update calculated in the manner of the Joseph formula, $\mathrm{E}\left\{\mathbf{e}^{\star}\left(\mathbf{e}^{\star}\right)^{\mathrm{T}}\right\}$, yields

$$\begin{aligned}\mathbf{P}^{\star} = {} & \left(\mathbf{I} - \mathbf{K}^{\star}\mathbf{H}^{\star}\right)\mathbf{P}^{-}\left(\mathbf{I} - \mathbf{K}^{\star}\mathbf{H}^{\star}\right)^{\mathrm{T}} + \mathbf{K}^{\star}\mathbf{R}^{\star}\mathbf{K}^{\star\mathrm{T}} \\ & - \left(\mathbf{I} - \mathbf{K}^{\star}\mathbf{H}^{\star}\right)\mathbf{C}^{\star}\mathbf{K}^{\star\mathrm{T}} - \mathbf{K}^{\star}\mathbf{C}^{\star\mathrm{T}}\left(\mathbf{I} - \mathbf{K}^{\star}\mathbf{H}^{\star}\right)^{\mathrm{T}}\end{aligned} \tag{3.45}$$

where $\mathbf{C}^\star$, $\mathbf{H}^\star$, $\mathbf{R}^\star$, and $\mathbf{K}^\star$ are defined by Eqs. (3.29), (3.34), (3.35), and (3.36a). Recall that this definition of the covariance update is independent of the choice of the Kalman gain $\mathbf{K}^\star$ and a detailed derivation of the update equations for a Kalman filter with correlated noise is given in [20].

Using Eqs. (3.29), (3.34), (3.35), and (3.36a) the full state covariance update is partitioned just as with the linear optimal case. The resulting attitude covariance update is

$$\begin{aligned}\mathbf{P}^\star_{\boldsymbol{\theta\theta}} &= \mathbf{P}^-_{\boldsymbol{\theta\theta}} - \mathbf{K}^\star_{\boldsymbol{\theta}}\mathbf{P}^-_{\boldsymbol{\theta\theta}} - \mathbf{P}^-_{\boldsymbol{\theta\theta}}\mathbf{K}^{\star\mathrm{T}}_{\boldsymbol{\theta}} + \mathbf{K}^\star_{\boldsymbol{\theta}}\mathbf{P}^-_{\boldsymbol{\theta\theta}}\mathbf{K}^{\star\mathrm{T}}_{\boldsymbol{\theta}} + \mathbf{K}^\star_{\boldsymbol{\theta}}\mathbf{P}^+_{\boldsymbol{\theta\theta}}\mathbf{K}^{\star\mathrm{T}}_{\boldsymbol{\theta}} \\ &\quad -\mathbf{C}^\star_{\boldsymbol{\theta}}\mathbf{K}^{\star\mathrm{T}}_{\boldsymbol{\theta}} + \mathbf{K}^\star_{\boldsymbol{\theta}}\mathbf{C}^\star_{\boldsymbol{\theta}}\mathbf{K}^{\star\mathrm{T}}_{\boldsymbol{\theta}} - \mathbf{K}^\star_{\boldsymbol{\theta}}\mathbf{C}^{\star\mathrm{T}}_{\boldsymbol{\theta}} + \mathbf{K}^\star_{\boldsymbol{\theta}}\mathbf{C}^{\star\mathrm{T}}_{\boldsymbol{\theta}}\mathbf{K}^{\star\mathrm{T}}_{\boldsymbol{\theta}} \end{aligned} \tag{3.46}$$

and substituting for $\mathbf{C}^\star_{\boldsymbol{\theta}}$ and $\mathbf{K}^\star_{\boldsymbol{\theta}}$ from Eqs. (3.32a) and (3.41a) respectively reduces the attitude error covariance to

$$\mathbf{P}^\star_{\boldsymbol{\theta\theta}} = \mathbf{P}^+_{\boldsymbol{\theta\theta}}. \tag{3.47}$$

This result indicates the attitude covariance update for the pseudo-measurement is equivalent to the linear optimal case of Eq. (3.25a) and is intuitively pleasing. Just as for the attitude state update where no additional attitude update is desired from the Kalman filter using the pseudo-measurement, no additional update to the attitude covariance is expected or desired.

The non-attitude state covariance update partition is given by

$$\begin{aligned}\mathbf{P}^\star_{\mathbf{ss}} &= \mathbf{P}^-_{\mathbf{ss}} - \mathbf{K}^\star_{\mathbf{s}}\mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}} - \mathbf{P}^-_{\mathbf{s}\boldsymbol{\theta}}\mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} + \mathbf{K}^\star_{\mathbf{s}}\mathbf{P}^-_{\boldsymbol{\theta\theta}}\mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} + \mathbf{K}^\star_{\mathbf{s}}\mathbf{P}^+_{\boldsymbol{\theta\theta}}\mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} \\ &\quad -\mathbf{C}^\star_{\mathbf{s}}\mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} + \mathbf{K}^\star_{\mathbf{s}}\mathbf{C}^\star_{\boldsymbol{\theta}}\mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} - \mathbf{K}^\star_{\mathbf{s}}\mathbf{C}^{\star\mathrm{T}}_{\mathbf{s}} + \mathbf{K}^\star_{\mathbf{s}}\mathbf{C}^{\star\mathrm{T}}_{\boldsymbol{\theta}}\mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} \end{aligned} \tag{3.48}$$

and substituting for $\mathbf{C}^\star_{\boldsymbol{\theta}}$, $\mathbf{C}^\star_{\mathbf{s}}$, and $\mathbf{K}^\star_{\mathbf{s}}$ from Eqs. (3.32a), (3.32b), and (3.41b) respectively reduces to

$$\mathbf{P}^\star_{\mathbf{ss}} = \mathbf{P}^-_{\mathbf{ss}} - \mathbf{P}^-_{\mathbf{s}\boldsymbol{\theta}} \left(\mathbf{P}^-_{\boldsymbol{\theta\theta}}\right)^{-1} \mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}} + \mathbf{P}^-_{\mathbf{s}\boldsymbol{\theta}} \left(\mathbf{P}^-_{\boldsymbol{\theta\theta}}\right)^{-1} \mathbf{P}^+_{\boldsymbol{\theta\theta}} \left(\mathbf{P}^-_{\boldsymbol{\theta\theta}}\right)^{-1} \mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}}. \tag{3.49}$$

Now substitute for the optimal attitude update, $\mathbf{P}^+_{\boldsymbol{\theta\theta}}$, using Eq. (3.26a) and recall the definitions of the optimal gains $\mathbf{K}_{\boldsymbol{\theta},opt}$ and $\mathbf{K}_{\mathbf{s}}$ from Eq. (3.17) to produce

$$\begin{aligned}
\mathbf{P}^\star_{\mathbf{ss}} &= \mathbf{P}^-_{\mathbf{ss}} - \mathbf{P}^-_{\mathbf{s}\boldsymbol{\theta}} \mathbf{H}^{\mathrm{T}}_{\boldsymbol{\theta}} \mathbf{W}^{-1} \mathbf{H}_{\boldsymbol{\theta}} \mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}} \\
&= \mathbf{P}^-_{\mathbf{ss}} - \mathbf{K}_{\mathbf{s},opt} \mathbf{W} \mathbf{K}^{\mathrm{T}}_{\mathbf{s},opt}
\end{aligned} \tag{3.50}$$

which is identical to the non-attitude state covariance update Eq. (3.26d) from the general partitioned case.

Finally, the cross-covariance partition of the covariance update is

$$\begin{aligned}
\mathbf{P}^\star_{\boldsymbol{\theta}\mathbf{s}} &= \mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}} - \mathbf{K}^\star_{\boldsymbol{\theta}} \mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}} - \mathbf{P}^-_{\boldsymbol{\theta\theta}} \mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} + \mathbf{K}^\star_{\boldsymbol{\theta}} \mathbf{P}^-_{\boldsymbol{\theta\theta}} \mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} + \mathbf{K}^\star_{\boldsymbol{\theta}} \mathbf{P}^+_{\boldsymbol{\theta\theta}} \mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} \\
&\quad - \mathbf{C}^\star_{\boldsymbol{\theta}} \mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} + \mathbf{K}^\star_{\boldsymbol{\theta}} \mathbf{C}^\star_{\boldsymbol{\theta}} \mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}} - \mathbf{K}^\star_{\boldsymbol{\theta}} \mathbf{C}^{\star\mathrm{T}}_{\mathbf{s}} + \mathbf{K}^\star_{\boldsymbol{\theta}} \mathbf{C}^{\star\mathrm{T}}_{\boldsymbol{\theta}} \mathbf{K}^{\star\mathrm{T}}_{\mathbf{s}}
\end{aligned} \tag{3.51}$$

and again substituting for $\mathbf{P}^+_{\boldsymbol{\theta\theta}}$, $\mathbf{C}^\star_{\boldsymbol{\theta}}$, $\mathbf{C}^\star_{\mathbf{s}}$, $\mathbf{K}^\star_{\boldsymbol{\theta}}$, and $\mathbf{K}^\star_{\mathbf{s}}$ from Eqs. (3.26a), (3.32a), (3.32b), (3.41a), and (3.41b) respectively reduces to

$$\begin{aligned}
\mathbf{P}^\star_{\boldsymbol{\theta}\mathbf{s}} &= \mathbf{P}^+_{\boldsymbol{\theta\theta}} \left(\mathbf{P}^-_{\boldsymbol{\theta\theta}}\right)^{-1} \mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}} \\
&= \left(\mathbf{I} - \mathbf{K}_{\boldsymbol{\theta},opt} \mathbf{H}_{\boldsymbol{\theta}}\right) \mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}} \\
&= \mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}} - \mathbf{P}^-_{\boldsymbol{\theta\theta}} \mathbf{H}^{\mathrm{T}}_{\boldsymbol{\theta}} \mathbf{W}^{-1} \mathbf{H}_{\boldsymbol{\theta}} \mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}} \\
&= \mathbf{P}^-_{\boldsymbol{\theta}\mathbf{s}} - \mathbf{P}^-_{\boldsymbol{\theta\theta}} \mathbf{H}^{\mathrm{T}}_{\boldsymbol{\theta}} \mathbf{K}_{\mathbf{s},opt}.
\end{aligned} \tag{3.52}$$

This result is likewise identical to the general partitioned case given in Eq. (3.26b). The same procedure can by applied to $\mathbf{P}^{\star}_{\mathbf{s}\boldsymbol{\theta}}$, but it is more straightforward to note that it is by definition the transpose of $\mathbf{P}^{\star}_{\boldsymbol{\theta}\mathbf{s}}$. Therefore, the corresponding covariance update is also identical to the general partitioned case.

Section 3.2.1 outlines the optimal update for the case of linear measurements where the measurements are only a function of part of the state. Recall that it is optimal in the sense of minimizing the trace of the error covariance matrix which can alternatively be derived as the minimum mean square error (MMSE) best estimate. This section demonstrates that it is equivalent to first update the attitude and subsequently use this updated portion of the state as a pseudo-measurement in order to update the remainder of the state. While presented in the context of spacecraft attitude, this method is valid for any linear measurement model where the state is correspondingly partitioned. It is also important to note that while the Joseph form of the covariance update equations are valid for any choice of the gains, in the derivation of $\mathbf{K}^{\star}$ as selected as in Eq. (3.41a) and Eq. (3.41b), the assumption of linear optimal Kalman gains was made from Eq. (3.17). As a result, the covariance update equations from the pseudo-measurement case of this section may be equivalently expressed as the optimal covariance update equations given by Eqs. (3.26a), (3.26b), (3.26c), and (3.26d). Alternatively, with this selection of the gains, the correlation covariance terms of Eq. (3.45) are zero and the covariance update due to the pseudo-measurement may be obtained directly from the standard Joseph formula.

$$\mathbf{P}^{+} = \left(\mathbf{I} - \mathbf{K}^{\star}\mathbf{H}^{\star}\right)\mathbf{P}^{-}\left(\mathbf{I} - \mathbf{K}^{\star}\mathbf{H}^{\star}\right)^{\mathrm{T}} + \mathbf{K}^{\star}\mathbf{R}^{\star}\mathbf{K}^{\star\mathrm{T}} \tag{3.53}$$

Here the methodology is expanded to the attitude estimation case where the measurements are solely a function of the attitude and the q-method is used to update the attitude partition of the state as derived in section 2.4. Recall that for this case the measurements are related to the attitude as described in the Wahba problem in Eq. (2.1) which is nonlinear. Consequently, the non-attitude state update derived in this equation is no longer guaranteed to be optimal in the MMSE sense (in fact it will almost certainly not be optimal). However, the q-method is an optimal nonlinear solution to the nonlinear Wahba problem and the resulting non-attitude state update will be near-optimal. This consequence is consistent with the well established and ubiquitous extended Kalman filter. While the Kalman filter is only optimal for linear measurements, most real-world applications are nonlinear. The extended Kalman filter linearizes the measurements by a Taylor series expansion about a nominal state thereby sacrificing the optimality of the Kalman filter for application to real-world systems. In like manner, the qEKF sacrifices optimality in the non-attitude state update for a nonlinear update to the attitude. Whether the choice of the qEKF or other methods are better suited to a particular application is a trade to be accomplished by the designer depending on the dynamics and priorities of the application.

In summary the qEKF filter has a propagation phase exactly the same as in the MEKF and an update phase as follows

1. Calculate the Davenport matrix $\mathbf{K}$ from Eq. (2.15) associated with all attitude vector measurements

2. Calculate $\mathbf{A}_0 = 2\left(\mathbf{P}_{\theta\theta}^{-}\right)^{-1}$

3. Calculate the updated attitude quaternion as the unit eigenvector associated with the maximum eigenvalue of

$$\mathbf{K}_{aug} = -\mathbf{\Xi}\left(\hat{\mathbf{q}}^{-}\right)\mathbf{A}_0\mathbf{\Xi}\left(\hat{\mathbf{q}}^{-}\right)^{\mathrm{T}} + \mathbf{K}$$

4. Calculate the updated attitude covariance partition $\mathbf{P}_{\boldsymbol{\theta\theta}}^{+}$ of the full covariance $\mathbf{P}$ from Eqs. (2.27), (2.33), (2.61), (2.63), and (2.66)

5. Update the non-attitude states using

$$\hat{\mathbf{s}}^{+} = \hat{\mathbf{s}}^{-} + \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}^{-}\left(\mathbf{P}_{\boldsymbol{\theta\theta}}^{-}\right)^{-1}\Delta\boldsymbol{\theta} \tag{3.54a}$$

$$\Delta\boldsymbol{\theta} = 2\mathbf{\Xi}\left(\hat{\mathbf{q}}^{-}\right)^{\mathrm{T}}\hat{\mathbf{q}}^{+} \tag{3.54b}$$

6. Calculate the total covariance update using Eq. (3.45), (3.36a), (3.34), and (3.35)

3.3 Comparison with Filter QUEST

Recall that Filter QUEST [24] is a recursive implementation of the QUEST algorithm [8] into a Kalman filter where QUEST is simply a numerical solution of the q-method. It is among the first methods which seek to implement the q-method in a recursive algorithm. Filter QUEST relies on the fundamental property that the attitude profile matrix Eq. (1.4), contains all the information necessary to compute the attitude and attitude covariance. In Filter QUEST only the attitude profile matrix is propagated and updated by adding new measurements as they become available. The attitude or the covariance is not available at each time step, but may be obtained by forming the Davenport matrix from the attitude profile matrix and solving the eigenvalue problem using the desired method such as QUEST. Like Filter QUEST, qEKF also solves for the attitude estimate without making any linearization assumptions. QUEST or any other numerical method may also be used to solve the eigenvalue problem in the qEKF. However, Filter QUEST is only capable of processing vector measurements and does not include *a priori* attitude estimates other than what is available from previous measurements. Furthermore, Filter QUEST does not incorporate process noise which is accounted for by using a fading memory factor. The computation of the optimal fading memory factor is difficult. Finally, Filter Quest is only capable of estimating the attitude and no other states. The proposed qEKF resolves all of these shortcomings by incorporating an initial attitude estimate using the method of averaging quaternions [2], allowing for the inclusion of process noise just as in an extended Kalman filter, and expanding the capability to estimating non-attitude states. Both methods rely on the q-method to obtain a nonlinear estimate of the attitude, but the q-method extended Kalman filter presents an algorithm capable of meeting a much expanded range of attitude estimation requirements in a single filter.

3.4 Comparison With The SOAR Filter

The proposed qEKF filter is most similar to the SOAR filter developed by Christian and Lightsey [3]. The SOAR filter also integrates the q-method to update the attitude into the framework of the Kalman filter and is capable of estimating other states. The initial condition is incorporated into SOAR in a different manner than qEKF and SOAR is derived from the information matrix formulation of the Kalman filter whereas qEKF is derived from the covariance formulation. Otherwise, the next section demonstrates that the qEKF and SOAR filters are equivalent to second-order in the attitude update and identical in the non-attitude update.

3.4.1 Equivalence of the Attitude Update

In the SOAR filter, the *a priori* attitude is incorporated into the Davenport matrix by the creation and addition of an *a priori* Davenport matrix, $\mathbf{K}^-$. Recall that this is the method proposed by Shuster [9]. The *a priori* attitude profile matrix is given by

$$\mathbf{B}^- = \left(\frac{1}{2}\text{trace}\left[\left(\mathbf{P}^-_{\boldsymbol{\theta\theta}}\right)^{-1}\right]\mathbf{I}_{3\times 3} - \left(\mathbf{P}^-_{\boldsymbol{\theta\theta}}\right)^{-1}\right)\mathbf{T}\left(\hat{\bar{\mathbf{q}}}^-\right) \tag{3.55}$$

from which the *a priori* Davenport matrix is calculated resulting in the following term from the objective function

$$-\,\text{trace}\left[\mathbf{T}\left(\mathbf{B}^-\right)^{\mathrm{T}}\right] = -\hat{\bar{\mathbf{q}}}^{\mathrm{T}}\mathbf{K}^-\hat{\bar{\mathbf{q}}}. \tag{3.56}$$

In [3] it is also shown that after a second-order expansion of the matrix exponential of $[\delta\boldsymbol{\theta}\times]$ about the *a priori* attitude, this objective function may be rewritten as

$$-\hat{\bar{\mathbf{q}}}^{\mathrm{T}}\mathbf{K}^{-}\hat{\bar{\mathbf{q}}} = -\left(\hat{\bar{\mathbf{q}}}^{-}\right)^{\mathrm{T}}\mathbf{K}^{-}\hat{\bar{\mathbf{q}}}^{-} + \frac{1}{2}\delta\boldsymbol{\theta}^{\mathrm{T}}\mathcal{F}_{\boldsymbol{\theta}\boldsymbol{\theta}}\delta\boldsymbol{\theta}. \tag{3.57}$$

In the computation of the optimal attitude, the first differential is taken with respect to the *a posteriori* attitude. As a result the first term is a constant and may be ignored and the *a priori* attitude is incorporated into the SOAR filter by $1/2\,\delta\boldsymbol{\theta}^{\mathrm{T}}\mathcal{F}_{\boldsymbol{\theta}\boldsymbol{\theta}}\delta\boldsymbol{\theta}$ to second-order.

Recall from the *a priori* attitude term given in the objective function for the qEKF from Eq. (2.50)

$$\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}^{-}\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}} = \delta\mathbf{q}_{v} = \sin\left(\frac{\delta\boldsymbol{\theta}}{2}\right). \tag{3.58}$$

Taking the Taylor Series expansion of $\sin\left(\delta\boldsymbol{\theta}/2\right)$ and approximating to second-order

$$\delta\mathbf{q}_{v} = \sin\left(\frac{\delta\boldsymbol{\theta}}{2}\right) = \frac{\delta\boldsymbol{\theta}}{2} - \frac{1}{3}\left(\frac{\delta\boldsymbol{\theta}}{2}\right)^{3} + \frac{1}{5}\left(\frac{\delta\boldsymbol{\theta}}{2}\right)^{5}\ldots \approx \frac{\delta\boldsymbol{\theta}}{2}. \tag{3.59}$$

Therefore, the first term in Eq. (2.50) may be rewritten as,

$$\hat{\bar{\mathbf{q}}}^{\mathrm{T}}\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}_{0}\right)\mathbf{A}_{0}\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}_{0}\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}} \approx \frac{1}{4}\delta\boldsymbol{\theta}^{\mathrm{T}}\mathbf{A}_{0}\delta\boldsymbol{\theta}. \tag{3.60}$$

Noting from before that $\mathbf{A}_0$ was chosen as $\mathbf{A}_0 = 2\left(\mathbf{P}^{-}_{\boldsymbol{\theta}\boldsymbol{\theta}}\right)^{-1} \approx 2\mathcal{F}_{\boldsymbol{\theta}\boldsymbol{\theta}}$, this directly yields

$$\hat{\bar{\mathbf{q}}}^{\mathrm{T}}\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}_{0}\right)\mathbf{A}_{0}\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}_{0}\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}} \approx \frac{1}{2}\delta\boldsymbol{\theta}^{\mathrm{T}}\mathcal{F}_{\boldsymbol{\theta}\boldsymbol{\theta}}\delta\boldsymbol{\theta}. \tag{3.61}$$

Therefore, the *a priori* attitude is added to the objective functions of both SOAR and qEKF in an equivalent manner to second-order.

3.4.2 Equivalence of the Non-Attitude Update

The non-attitude state update in the SOAR filter is identical to that derived in the qEKF. In SOAR the optimal update of the non-attitude states is given by [3]

$$\mathbf{s}^{+} = \mathbf{s}^{-} - 2\left(\mathbf{F}_{\mathbf{ss}}^{-}\right)^{-1}\mathbf{F}_{\mathbf{s}\boldsymbol{\theta}}^{-}\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}^{-}\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}}^{+} \tag{3.62a}$$

$$\approx \mathbf{s}^{-} - \left(\mathbf{F}_{\mathbf{ss}}^{-}\right)^{-1}\mathbf{F}_{\mathbf{s}\boldsymbol{\theta}}^{-}\delta\boldsymbol{\theta}. \tag{3.62b}$$

Recall that the Fisher information matrix is approximately equal to the inverse of the covariance matrix as the number of observations becomes large. The partitioned Fisher information and covariance matrices are related by

$$\mathbf{P}^{-1} = \begin{bmatrix} \mathbf{P}_{\mathbf{ss}} & \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}} \\ \mathbf{P}_{\boldsymbol{\theta}\mathbf{s}} & \mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}} \end{bmatrix}^{-1} \approx \mathbf{F}_{\mathbf{xx}} = \begin{bmatrix} \mathbf{F}_{\mathbf{ss}} & \mathbf{F}_{\mathbf{s}\boldsymbol{\theta}} \\ \mathbf{F}_{\boldsymbol{\theta}\mathbf{s}} & \mathbf{F}_{\boldsymbol{\theta}\boldsymbol{\theta}} \end{bmatrix}. \tag{3.63}$$

Using the relationships for the inversion of a partitioned matrix,

$$\mathbf{F}_{\mathbf{ss}} = \left(\mathbf{P}_{\mathbf{ss}} - \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}\mathbf{P}_{\boldsymbol{\theta}\mathbf{s}}\right)^{-1}, \tag{3.64a}$$

$$\mathbf{F}_{\mathbf{s}\boldsymbol{\theta}} = -\left(\mathbf{P}_{\mathbf{ss}} - \mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}\mathbf{P}_{\boldsymbol{\theta}\mathbf{s}}\right)^{-1}\mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}, \tag{3.64b}$$

$$\mathbf{F}_{\mathbf{s}\boldsymbol{\theta}} = -\mathbf{F}_{\mathbf{ss}}\mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}, \tag{3.64c}$$

$$\text{and}\quad \mathbf{F}_{\mathbf{ss}}^{-1}\mathbf{F}_{\mathbf{ss}} = -\mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}, \tag{3.64d}$$

a simple substitution is used to show the equivalence of the SOAR and the qEKF non-attitude state update. Substitute the above relationship from the partitioned Fisher information and covariance matrices into the SOAR update Eq. (3.62a)

$$\mathbf{s}^{+} = \mathbf{s}^{-} + 2\mathbf{P}_{\mathbf{s}\boldsymbol{\theta}}^{-}\left(\mathbf{P}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-}\right)^{-1}\boldsymbol{\Xi}\left(\hat{\bar{\mathbf{q}}}^{-}\right)^{\mathrm{T}}\hat{\bar{\mathbf{q}}}^{+} \tag{3.65}$$

which is exactly the same as the qEKF non-attitude state update given in Eq. (3.54b).

In summary, both qEKF and SOAR select the non-attitude state Kalman gain according to the linear optimal value and produce identical non-attitude state updates where qEKF uses the covariance formulation and SOAR uses the information matrix approach. However, the covariance formulation used in qEKF is advantageous over the information matrix formulation when the state vector becomes large. Regardless of the number of states, the qEKF algorithm requires the inversion of 3×3 matrices related to the attitude. The computation of the information matrix for the SOAR algorithm requires the inversion of the full state covariance. Therefore, for an estimation algorithm consisting of n total states, the SOAR algorithm requires the inversion of an $n \times n$ matrix while the qEKF algorithm only requires the inversion of a 3×3 matrix. This results in an increasing computational savings as the state vector becomes large.

Chapter 4

Numerical Simulation

Numerical simulations are used to verify the approach proposed by the q-method extended Kalman filter. A simple spacecraft model is used with magnetometers and sun sensors for attitude determination and a rate gyro which eliminates the need for modeling the spacecraft attitude dynamics. Vector measurements are simulated by perturbing the true values with simulated measurement noise. In order to demonstrate the capability of the qEKF to estimate attitude as well as non-attitude states, the state vector used by the Kalman filter for the numerical simulations consists of the spacecraft attitude $\boldsymbol{\theta}$, and the gyro bias, $\boldsymbol{\beta}$. That is,

$$\mathbf{x} = \begin{bmatrix} \boldsymbol{\theta} \\ \boldsymbol{\beta} \end{bmatrix}. \tag{4.1}$$

Here the attitude is again represented by the three-component angles (roll, pitch, and yaw). Monte Carlo analysis is used to validate the convergence and performance of qEKF. SOAR is also implemented into the same simulations in order to demonstrate the equality between SOAR and qEKF.

4.1 Selection of Orbit

For the numerical examples in this work a common orbit is used. The only orbital requirements for this work are such that the orbit provides sufficient observability properties. For simplicity, a circular orbit with a semi-major axis of 7,000 km is selected. The inclination is selected as 45 degrees so that the spacecraft will experience a sufficient variation in Earth's magnetic field vector throughout the course of its orbit such that there is adequate variation in the magnetometer measurements. The rest of the orbital parameters are arbitrarily assigned for simplicity. At the beginning of the simulation the Earth is at vernal equinox, 20 March 2012, and the spacecraft is located at the ascending node. The simulation spans a time period of 6,000 seconds which is slightly more than one orbital period for the selected orbit. For a circular orbit the dynamics is sufficiently modeled by a simple rotation

$$\mathbf{r} = \mathbf{C}\mathbf{r}_0, \tag{4.2}$$

where $\mathbf{r}$ is the current spacecraft position vector, $\mathbf{r}_0$ is the *a priori* position vector, and $\mathbf{C}$ is the rotation matrix given by

$$\mathbf{C} = \mathbf{I}_{3\times3}\cos\left(n\cdot\Delta t\right) + \mathbf{e}\mathbf{e}^{\mathrm{T}}\left(1-\cos\left(n-\Delta t\right)\right) - \left[\mathbf{e}\times\right]\sin\left(n\cdot\Delta t\right), \tag{4.3}$$

where n is the orbital mean motion, Δt is the time step, and $\mathbf{e}$ is the axis of rotation. This dynamic model is only valid for circular orbits according to the restricted two-body equation where the mass of the orbiting body is insignificant compared to the mass of the primary body and no other perturbational effects are considered. Throughout its orbit the spacecraft is oriented such that the body-fixed X axis is directed in track and the Z axis is Earth-pointing with the Y axis following a right handed coordinate system. As a result the spacecraft has a constant angular velocity equal in magnitude and opposite in direction to the orbital mean motion as expressed in the spacecraft body frame. That is,

$$\boldsymbol{\omega} = \begin{bmatrix} 0 & -n & 0 \end{bmatrix}^{\mathrm{T}}. \tag{4.4}$$

The orbital mean motion, n, for a circular orbit is simply the average orbital angular velocity which is given by 2π divided by the orbital period. This attitude orientation, also referred to as nadir pointing, is common for many satellites such as telecommunication or remote sensing satellites.

4.2 Attitude Sensors

Two types of vector measurements are considered for this thesis, sun sensor and magnetometer measurements. Sun sensors provide a unit vector measurement from the spacecraft body frame (related through a fixed sensor frame) to the sun. For Earth orbits, the sun is sufficiently approximated as a point source which greatly simplifies sensor design and data processing. Varying types of sun sensors may have a field of view as large as 128 degrees and have typical performance accuracies between

0.005 and 3 degrees for attitude determination. Furthermore, the sun is of significant interest with regards to mission design for the vast majority of spacecraft, particularly for power generation and thermal constraints. All of these factors contribute to the selection of sun sensors as the most widely used sensor type [5, 33]. For this example with the Earth at vernal equinox, the inertial reference vector for the sun sensor is given in the Earth-centered inertial (ECI) frame by

$$\mathbf{n}_{sun} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^{\mathrm{T}}, \tag{4.5}$$

and may sufficiently be assumed to remain constant for the simulation time span of 6,000 seconds. Sun sensor measurements in the spacecraft body frame are generated in 1 second intervals by rotating the reference sun vector according to the true attitude and corrupted by simulated sensor noise.

Magnetometers are vector sensors that provide both the magnitude and direction of Earth's magnetic field at the spacecraft. They are simple, inexpensive, lightweight, and reliable sensors with no moving parts and low power requirements. As a result, they are widely used, especially with spacecraft that rely on magnetorquers or torque rods as an attitude control actuator or small, inexpensive spacecraft which lack tight attitude determination tolerances. However, magnetometers are often less accurate for attitude determination than alternative attitude sensors largely due to the difficulty of modeling Earth's magnetic field and yield a typical accuracy range of 0.5 to 3 degrees. While often approximated as a simple dipole, Earth's magnetic field is time varying and not completely known resulting in errors in the inertial reference model. Higher fidelity models of Earth's magnetic field use spherical harmonics based on empirical data, but can become considerably more complex which results in an increased demand on the computing resources of the spacecraft's attitude determination system. Furthermore, the strength of Earth's magnetic field drops according to an inverse cubed relationship with increased distance from Earth restricting the practical use of magnetometer measurements in most cases to orbits less than 1,000 kilometers in altitude [5, 33].

Earth's magnetic field remains relatively constant in magnitude and direction around the equator and experiences the most variation over the poles. The orbital inclination is selected as 45 degrees in order to increase the variation of Earth's magnetic field throughout the spacecraft's orbit, thus increasing the observability. In the examples for this thesis, the inertial magnetic field vectors are obtained in 1 sec intervals from the orbital position, time, and date using the World Magnetic Model (WMM) produced by the National Geospatial-Intelligence Agency (NGA) of the United States and the Defence Geographic Center (DGC) of the United Kingdom. The magnetometer measurements are also simulated by rotating the reference field vector according to the true spacecraft attitude and then corrupting with simulated measurement noise.

Another common sensor is the rate gyroscope or gyro. Rate gyros are not used to measure the spacecraft's attitude, but rather its angular velocity. Gyros come in a large range of accuracies, complexity, size, and capabilities in order to meet the

Error Source	Symbol	Value
Sun-sensor Noise ($\boldsymbol{\eta}_{sun}$)	σ_{sun}	0.1 deg
Magnetometer Noise $\left(\boldsymbol{\eta}_{mag}\right)$	σ_{mag}	220 nT
Angular Random Walk ($\boldsymbol{\eta}_v$)	σ_v	$\sqrt{10} \times 10^{-7}\ {}^{\text{rad}}\!/_{\text{sec}^{1/2}}$
Gyro Bias Random Walk ($\boldsymbol{\eta}_u$)	σ_u	$\sqrt{10} \times 10^{-10}\ {}^{\text{rad}}\!/_{\text{sec}^{3/2}}$

Table 4.1: Sensor Errors

requirements for a particular mission. Rate gyros represent the most simple and least expensive type. The use of gyros eliminates the need for modeling complicated attitude dynamics resulting in significant computational savings for the attitude determination and control system. Gyros operate at a much higher sampling frequency than the attitude sensors and can therefore be used to accurately propagate the state and covariance between attitude sensor measurements. Rate gyros suffer from errors due to nonlinearity, drift, and hysteresis [5, 33]. For this work, the rate gyro is defined according to the following sensor model [34]

$$\boldsymbol{\omega} = \tilde{\boldsymbol{\omega}} - \boldsymbol{\beta} - \boldsymbol{\eta}_v \tag{4.6a}$$

$$\text{and} \quad \dot{\boldsymbol{\beta}} = \boldsymbol{\eta}_u, \tag{4.6b}$$

where $\boldsymbol{\omega}$ is the true angular velocity, $\tilde{\boldsymbol{\omega}}$ is the measured angular velocity, $\boldsymbol{\beta}$ is the gyro bias vector, and $\boldsymbol{\eta}_v$ and $\boldsymbol{\eta}_u$ are zero-mean, Gaussian, white-noise processes. The noise parameters used in the simulations are summarized in Table 4.1.

4.3 Propagation

The Kalman filter is a two-step process. One step consists of updating the state and covariance estimates by incorporating information from measurements at the current time. For optimal Kalman filters, this update is done in a statistically optimal manner to produce the minimum mean square error estimate of the state vector and associated covariance. The update phase has been discussed extensively in this work. The second step of the Kalman filter is the propagation phase. This phase moves the state and covariance forward in time with a penalty on the covariance due to the process noise which causes it to grow as time goes on. This thesis uses a propagation step identical to MEKF with more detailed derivations available in literature [22, 11]. The attitude kinematics equation for quaternions is given by

$$\dot{\bar{\mathbf{q}}} = \frac{1}{2}\boldsymbol{\Xi}\left(\bar{\mathbf{q}}\right)\boldsymbol{\omega}. \tag{4.7}$$

Recall that for MEKF, the three component error angle vector $\delta\boldsymbol{\theta}$ is the attitude portion of the state which is passed into the Kalman filter as opposed to the full quaternion. This substitution is valid for small angles which is accomplished by using a sufficiently small step size, 1 second for this thesis. Including the gyro model from Eq. (4.6a), the attitude dynamics model is given by

$$\delta\dot{\boldsymbol{\theta}} = \left[-\hat{\boldsymbol{\omega}}\times\right]\delta\boldsymbol{\theta} - \delta\boldsymbol{\beta}, \tag{4.8}$$

where the estimated angular velocity, $\hat{\boldsymbol{\omega}} = \tilde{\boldsymbol{\omega}} - \hat{\boldsymbol{\beta}}$, is obtained directly from the gyro measurement estimated bias. As a result, integration of the dynamics is not necessary.

Using Eqs. (4.6b) and (4.8) The state resulting full state dynamic model in state is given by

$$\delta\dot{\mathbf{x}} = \mathbf{F}\delta\mathbf{x} + \mathbf{G}\boldsymbol{\nu} \tag{4.9a}$$

$$\begin{bmatrix} \delta\dot{\boldsymbol{\theta}} \\ \delta\dot{\boldsymbol{\beta}} \end{bmatrix} = \begin{bmatrix} [-\hat{\boldsymbol{\omega}}\times] & -\mathbf{I}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{0}_{3\times3} \end{bmatrix} \begin{bmatrix} \delta\boldsymbol{\theta} \\ \delta\boldsymbol{\beta} \end{bmatrix} + \begin{bmatrix} -\mathbf{I}_{3\times3} & \mathbf{0}_{3\times3} \\ \mathbf{0}_{3\times3} & \mathbf{I}_{3\times3} \end{bmatrix} \begin{bmatrix} \boldsymbol{\eta}_v \\ \boldsymbol{\eta}_u \end{bmatrix}, \tag{4.9b}$$

where $\mathbf{F}$ is the Jacobian for the state space model. The process noise, $\boldsymbol{\nu}$, is used to calculate the covariance parameter, $\mathbf{Q}$, which is related to the process noise covariance matrix by

$$\mathbf{Q}_k = \mathbf{G}\mathbf{Q}\mathbf{G}^{\mathrm{T}}\Delta t = \begin{bmatrix} \sigma_u^2\mathbf{I}_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & \sigma_v^2\mathbf{I}_{3\times3} \end{bmatrix} \tag{4.10}$$

noting that a unit step size is used for this thesis. The process noise defined in Eqs. (4.6a) and (4.6b) are assigned values as in Table 4.1. The propagated covariance matrix is therefore given by

$$\mathbf{P}_{k+1} = \boldsymbol{\Phi}\mathbf{P}_k\boldsymbol{\Phi}^{\mathbf{T}} + \mathbf{Q}_k, \tag{4.11}$$

where $\boldsymbol{\Phi}$ is the state transition matrix that maps from time t_k to time t_{k+1}. The state transition matrix is obtained using the dynamics Jacobian, $\mathbf{F}$, which is assumed constant during the interval of the time step and approximated to first-order by

$$\boldsymbol{\Phi} = e^{\mathbf{F}\Delta t} \approx \mathbf{I}_{6\times6} + \mathbf{F}\Delta t. \tag{4.12}$$

Since gyro measurements are available and are typically sampled at a high rate, the estimated angular velocity may be assumed to remain constant for the duration of the step size. Using the small angle assumption, the quaternion a the next time step is obtained by

$$\Delta\boldsymbol{\theta} = \hat{\boldsymbol{\omega}}\Delta t \tag{4.13a}$$

$$\text{and} \quad \hat{\bar{\mathbf{q}}}_{k+1} = \begin{bmatrix} \Delta\boldsymbol{\theta}/2 \\ 1 \end{bmatrix} \otimes \hat{\bar{\mathbf{q}}}_k. \tag{4.13b}$$

Detailed derivations of the state and covariance propagation for extended Kalman filters are available in literature [20, 11, 21].

4.4 Simulation Results

For the each simulation the initial attitude error covariance is $0.1^2 \deg^2$ in each axis and the initial gyro bias error covariance is $0.2^2\,(deg/hr)^2$ in each axis. There is no initial cross covariance and the initial gyro bias is assumed to be zero mean. A 100 run Monte Carlo simulation is performed for each test case varying the measurement noise, process noise, and initial state estimate. Each figure displays the estimation error for the roll, pitch, and yaw axes. The estimation error for each Monte Carlo run is shown in red for qEKF and green for SOAR. The $3-\sigma$ bounds of the estimated covariance for each axis is shown in blue with the corresponding statistical covariance from the Monte Carlo estimation errors shown in black.

4.4.1 Test Case #1: Synchronized Sun Sensor and Magnetometer Measurements

Initially a simulation is performed for a simple and observable test case. In this first simulation synchronized sun sensor and magnetometer measurements are available for the entire orbit. This is not the case in reality for the selected orbit as it will be behind Earth's shadow for approximately one third of the orbit, but it provides a good test case to ensure the qEKF converges properly. The SOAR filter is run simultaneously using the exact same inputs for comparison.

Figure 4.3 displays the attitude estimation error as expressed in the spacecraft body frame for the roll, pitch, and yaw axes using the qEKF and SOAR algorithms. As expected, both filters yield identical results. Note that the attitude converges to an accurate estimate very quickly. One of the primary advantages of the q-method based qEKF is the globally optimal nonlinear attitude update. In contrast, standard extended Kalman filter techniques provide a linear best estimate and can be subject to convergence issues for nonlinear systems. Figure 4.2 shows the corresponding gyro bias estimation error which also converges to an appropriate steady state value. Notice that the performance in the pitch axis exceeds that of the roll and yaw axes. This is a direct consequence of the dynamics of the system. Recall that the modeled satellite remains Earth-pointing throughout its orbit and as a result, only rotates about the spacecraft pitch axis. Therefore the greatest variation in measurements is in the pitch axis making it the most observable which leads to the increased performance.

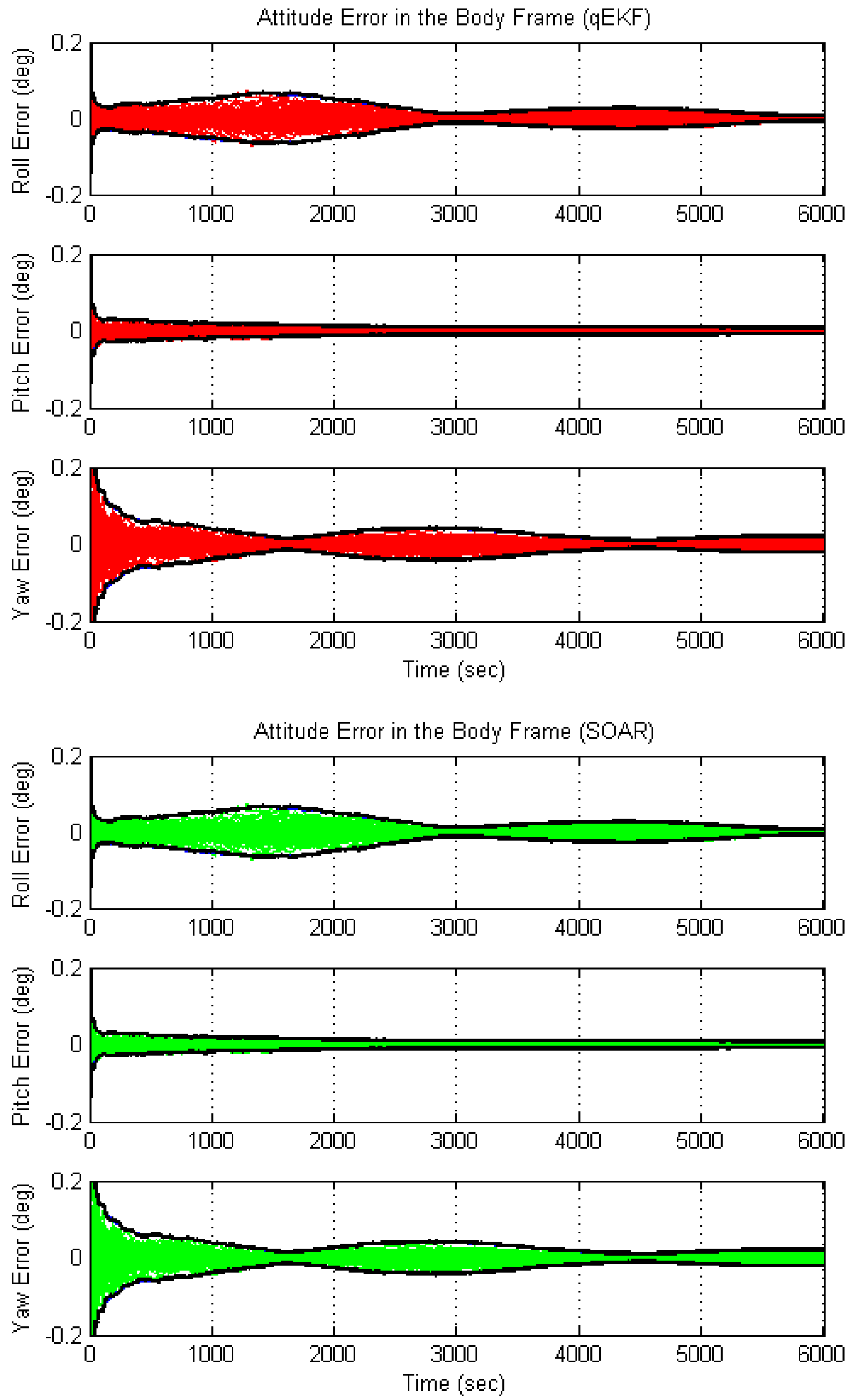

Figure 4.1: Attitude Estimation Error 100 Run Monte Carlo

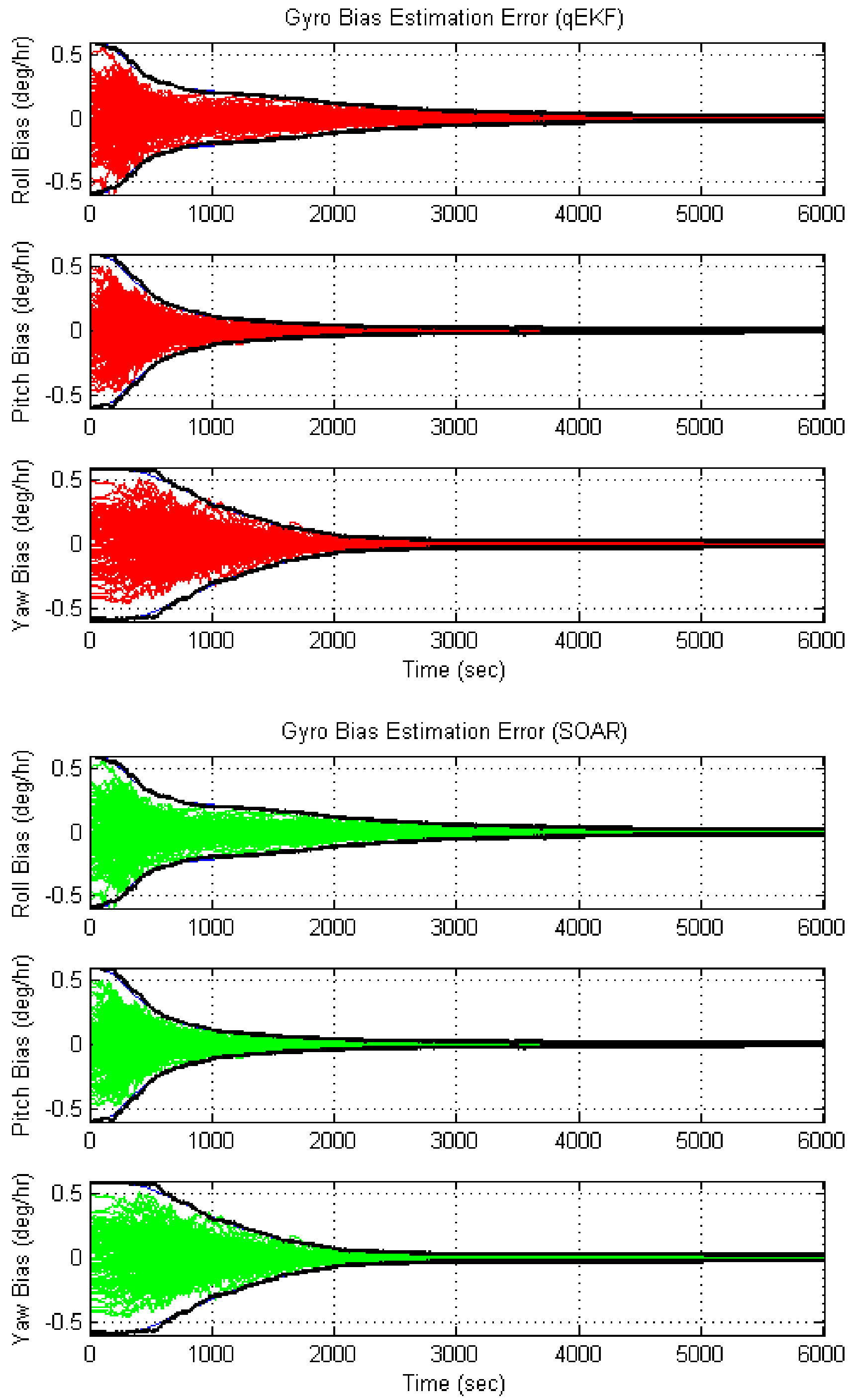

Figure 4.2: Gyro Bias Estimation Error 100 Run Monte Carlo

The sinusoidal fluctuations in the attitude error for the roll and yaw axes are also consequences of the spacecraft dynamics.

4.4.2 Test Case #2: Only Magnetometer Measurements

Another case of interest consists of attitude estimation using only magnetometer measurements. An advantage of the qEKF is that by incorporating the *a priori* attitude estimate, only a single vector measurement is required to determine the attitude from the q-method as opposed to the customary requirement of at least two non-collinear vector measurements. As a result, the qEKF is capable of estimating the attitude and gyro bias from magnetometer measurements alone. This case has practical applications for small, inexpensive satellites, such as CubeSats, or as a backup attitude estimation method for larger and more complex satellites. The results are again compared with the results from the SOAR filter subject to the same measurement inputs. Figure 4.3 presents the attitude estimation error using only magnetometer measurements and Figure 4.4 presents the associated gyro bias estimation error.

The qEKF again produces identical results to SOAR and both the attitude and gyro bias converge to appropriate values. Comparing with the results from the case using sun sensor and magnetometer measurements, the magnetometer only case takes longer to converge and results in a larger steady state estimation error. This behavior is expected as attitude information is lost from the lack of sun sensor measurements resulting in larger errors in the system.

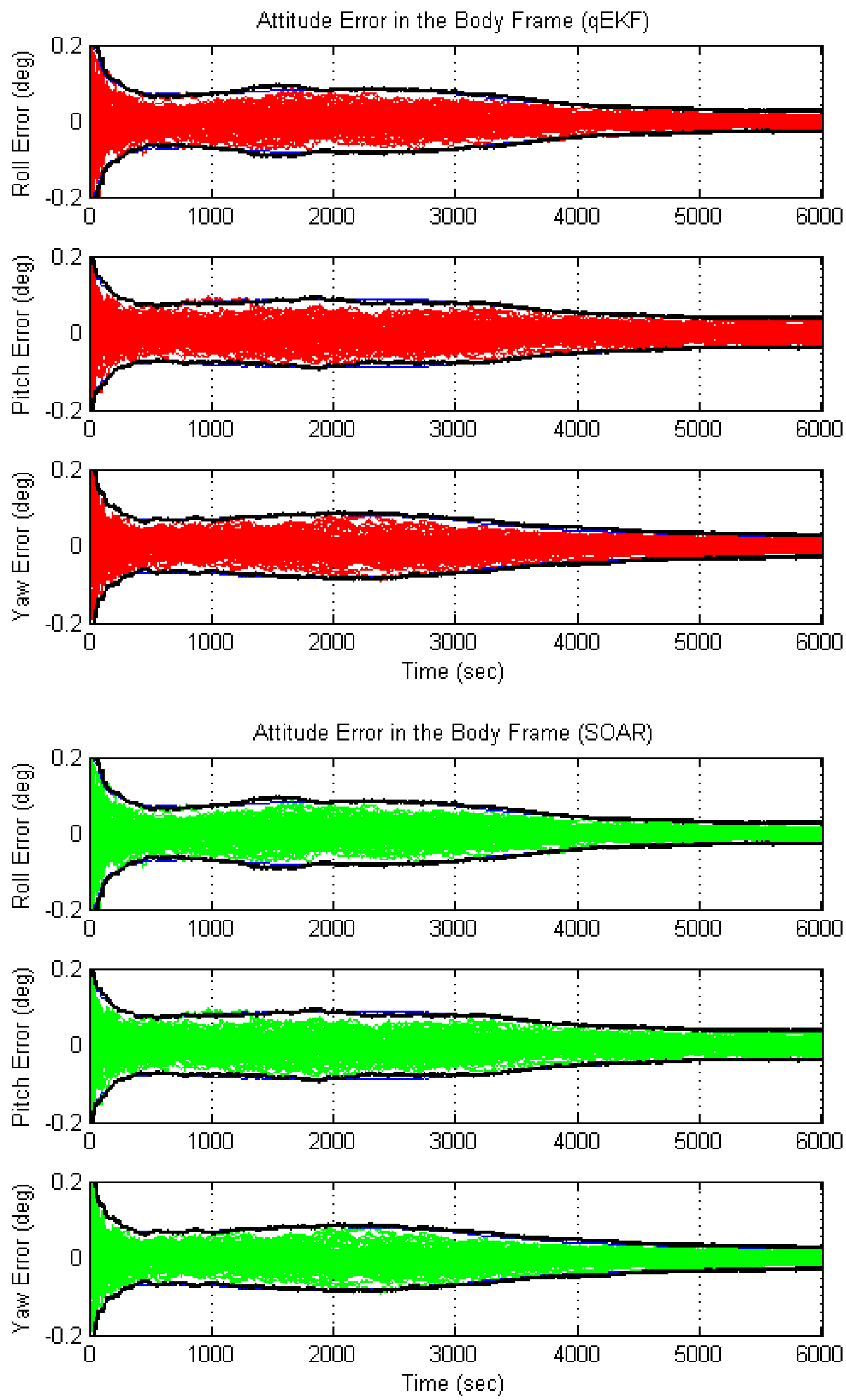

Figure 4.3: Attitude Estimation Error for Only Magnetometer Measurements 100 Run Monte Carlo

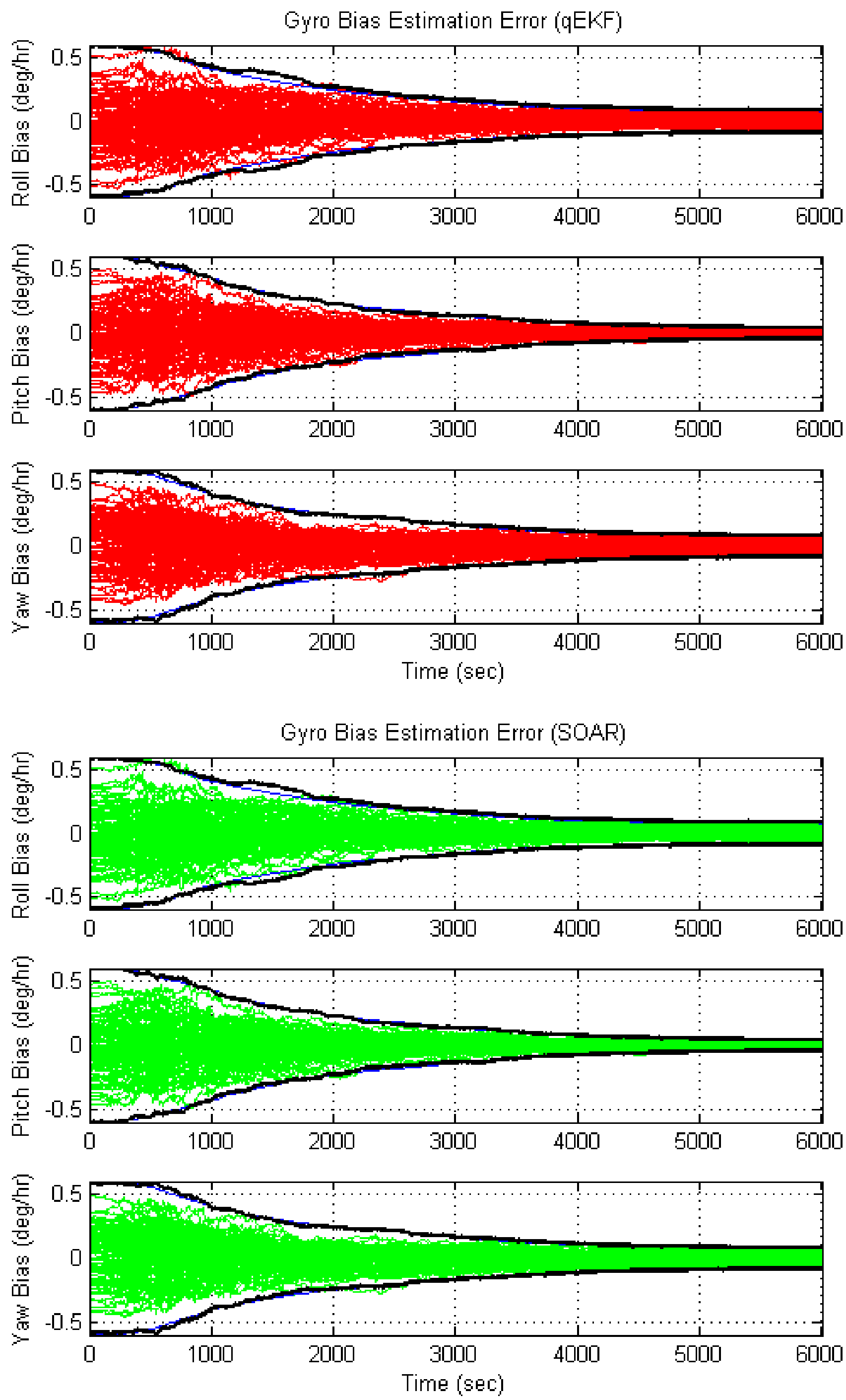

Figure 4.4: Gyro Bias Estimation Error for Only Magnetometer Measurements 100 Run Monte Carlo

4.4.3 Test Case #3: Asynchronous Sun Sensor and Magnetometer Measurements with Eclipse

In actuality, the measurements from different sensors on a spacecraft are never obtained at the exact same time. Recall that two independent vector measurements are required to determine the attitude. For single-point solutions the measurements are either propagated to a mutual time or assumed to be close enough together that the difference may be ignored. The qEKF filter incorporates an initial attitude estimate, therefore, the attitude may be determined from a single vector measurement. As a result, it is advantageous to process each measurement as soon as it becomes available. To represent this capability, in this test case the sun sensor and magnetometer measurements are alternately available at each step in the simulation. During the period of eclipse, when the satellite is behind the Earth's shadow and sun sensor measurements are unavailable, they are replaced with magnetometer measurements instead. The eclipse period occurs between simulation time 2,000 and 3,800 seconds.

The attitude estimation error for the case including eclipse is shown in Figure 4.5 and the corresponding gyro bias estimation error is shown in Figure 4.6. Again, both the attitude and non-attitude states converge to appropriate steady state errors with a performance that fits in between the magnetometer only case and the case with sun sensor and magnetometer measurements available simultaneously for the entire orbit. The sun sensor provides additional attitude information that improves the performance over the magnetometer only case, but less than the first test case as

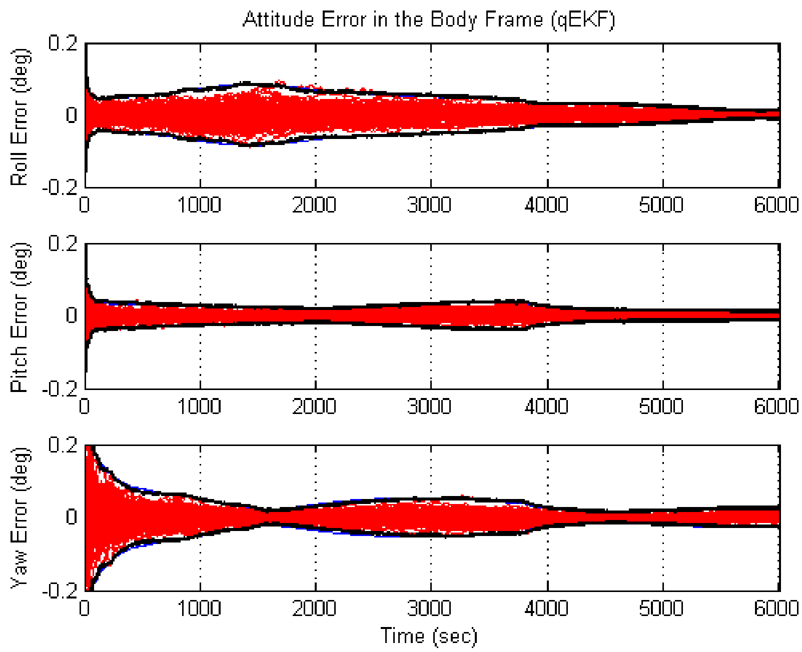

Figure 4.5: Attitude Estimation Error for Eclipse 100 Run Monte Carlo

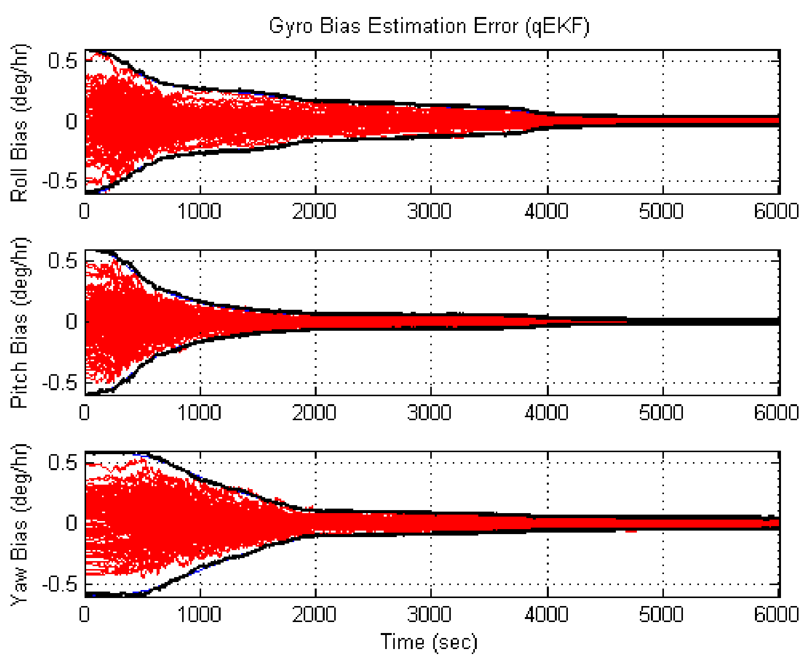

Figure 4.6: Gyro Bias Estimation Error for Eclipse 100 Run Monte Carlo

the measurements are staggered every other second. During the period of eclipse in between 2,000 and 3,800 seconds when the sun sensor measurements are not available, the performance mirrors the magnetometer only case and tighter convergence is shown once the sun sensor measurements again become available.

4.4.4 Test Case #4: Only Magnetometer Measurements with Large Initial Errors

Recall that the primary advantage of the q-method extended Kalman filter over standard Kalman filtering techniques is that qEKF provides a nonlinear attitude estimate. In the previous test cases, a standard method such as the multiplicative extended Kalman filter (MEKF) yields identical results. However, the linearization assumptions implicit in the extended Kalman filter become less appropriate for highly nonlinear dynamics or poor initial estimates. In order to demonstrate the advantage of the nonlinear attitude update of qEKF over MEKF, this case uses large initial errors. For this example, all parameters are the same as in test case #2 using only magnetometer measurements except that the sensor errors from Table 4.1 have been increased by one order of magnitude, the initial attitude error covariance is now $200^2\,\mathrm{deg}^2$ in each axis and the initial gyro bias error covariance is $20^2\,(deg/hr)^2$ in each axis. The same simulation is performed using the qEKF and the MEKF in order to compare the effects of the nonlinear attitude update.

The attitude estimation error for the case of large initial errors is shown in Figure

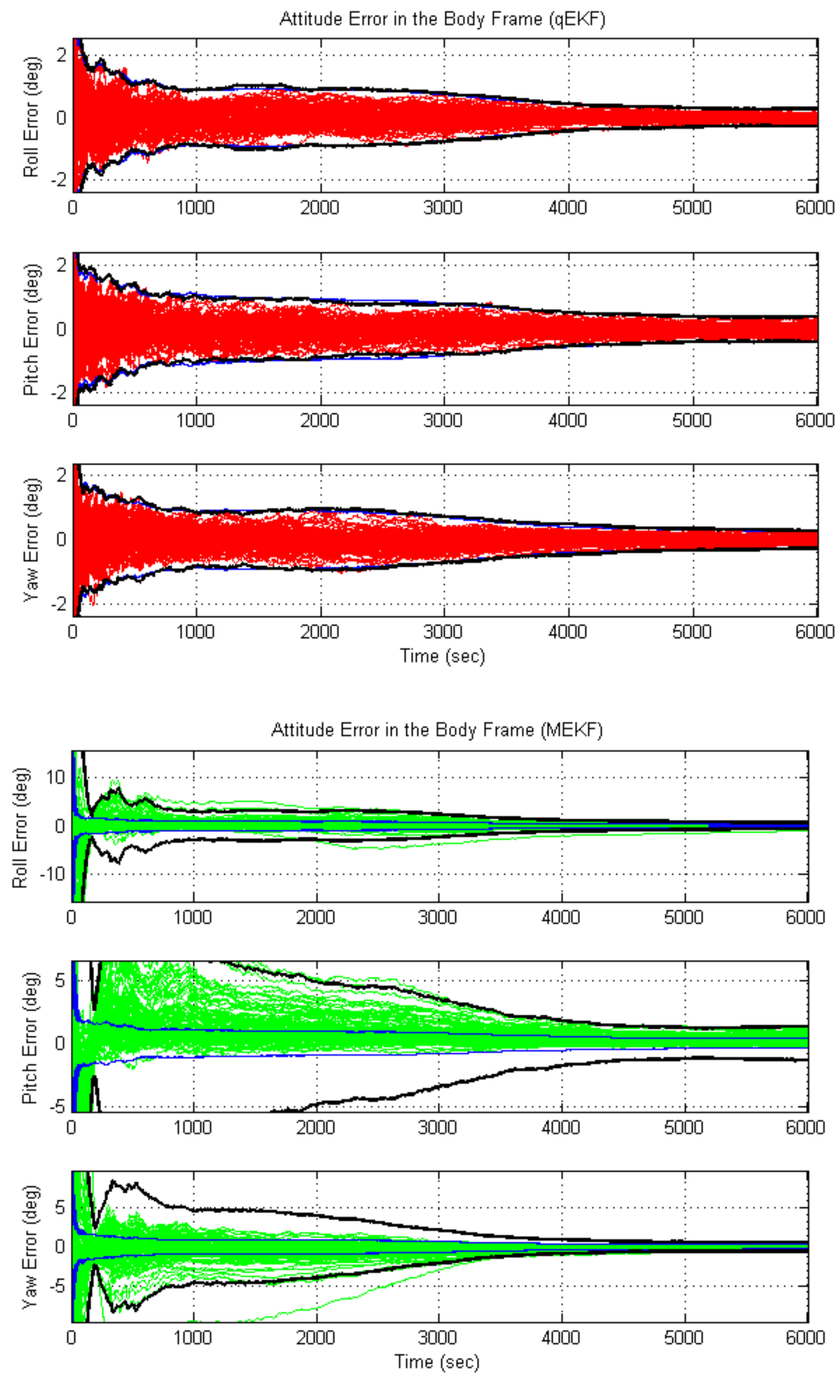

Figure 4.7: Attitude Estimation Error for Only Magnetometer and Large Initial Errors 100 Run Monte Carlo

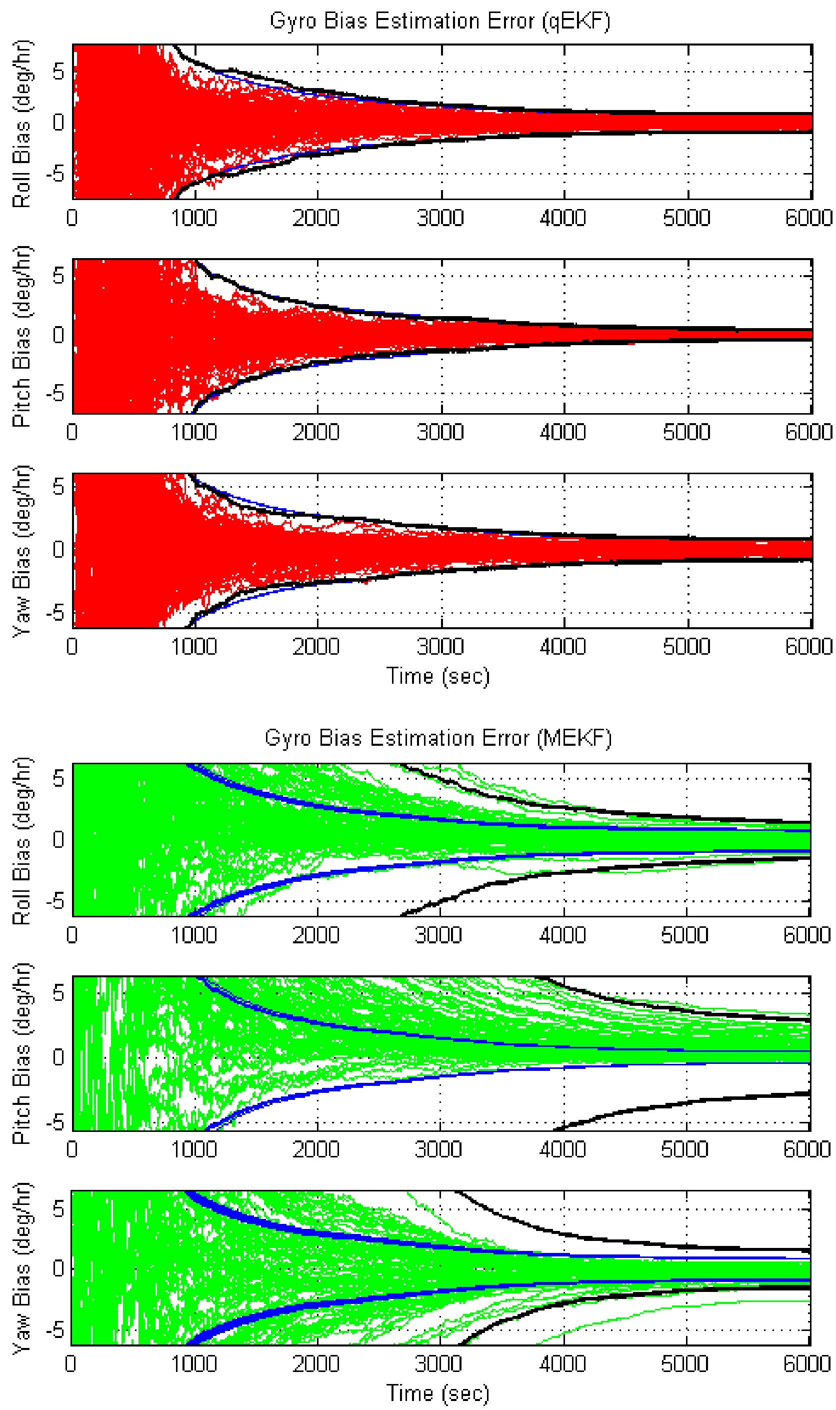

Figure 4.8: Gyro Bias Estimation Error for Only Magnetometer and Large Initial Errors 100 Run Monte Carlo

4.7 and the gyro bias estimation error is shown in Figure 4.8. For this case the results from the qEKF are shown in red and from the MEKF in green. For both the attitude and the gyro bias the qEKF converges quickly to achieve an appropriate steady state error. The resulting steady state error remains larger than the previous test cases as the sensor errors are an order of magnitude larger for this case. As a result, the qEKF is able to accurately estimate the attitude even with an extremely poor initial estimate. While it does eventually converge, the MEKF significantly under-performs when presented with the same large errors. The overshoot visible in Figure 4.7 is a direct consequence of the linearization in MEKF which attempts to over-correct based on the poor initial estimate. Furthermore, a significant number of the Monte Carlo runs for both the attitude and gyro bias estimation errors remain outside of the $3-\sigma$ bounds of the predicted covariance. The practical implication is that in a real system where the true values are unavailable, MEKF would predict a more accurate estimation of the states than is true for this case. Finally, the steady state estimation error is approximately twice as large for the MEKF as opposed to the qEKF. As a result, the nonlinear attitude estimation provided by qEKF performs much better than MEKF in the presence of large initial errors.

Chapter 5

Concluding Remarks

5.1 Conspectus

In this thesis the q-method is successfully integrated into the extended Kalman filter in order to provide the capability of nonlinear attitude estimation along with the estimation of non-attitude states. The *a priori* attitude is combined with the Wahba problem performance index and the q-method solution to the Wahba problem is appropriately modified. Covariance analysis is performed to determine the corresponding attitude error covariance update from the q-method solution. For the linear case, it is shown to be equivalent to update the attitude first and subsequently update the non-attitude states as to using standard linear Kalman filter to update the entire state. Just as with the extended Kalman filter, the qEKF sacrifices optimality in the non-attitude update from the linear case in order to process nonlinear systems.

However, a nonlinear attitude update is preserved through the q-method.

The first proposed methods of recursive attitude estimation based on the Wahba problem are only capable of estimating the attitude and cannot include other states. Subsequent extensions add this capability, but require iteration to estimate the non-attitude states. In contrast, extended Kalman filter methods such as MEKF are capable of estimating any number of observable states. However, due to the required linearization extended Kalman filters are sensitive to initial conditions and proper tuning in order to converge properly and can fall victim to divergence issues in the presence of large errors or large nonlinearities in the system. The q-method Kalman filter provides an ideal algorithm for on-board attitude estimation by merging the best of these two classes of filters. A globally optimal, nonlinear attitude update is obtained using the q-method and non-attitude states are simultaneously estimated with a Kalman filter. By integrating the q-method into the framework of the Kalman filter, the *a priori* attitude information is maintained. For increased speed, any number of numerical solutions to the eigenvalue problem from the q-method such as QUEST may be used to calculate the optimal attitude.

The q-method extended Kalman filter is compared with the Sequential Optimal Attitude Recursion filter and shown to be equivalent to second-order in the attitude update and identical in the non-attitude state update. The qEKF may be viewed as the covariance formulation of integrating the q-method into an extended Kalman filter whereas SOAR is the information matrix formulation. This distinction becomes important as the size of the state vector increases because the covariance formulation will typically require smaller matrix inversions than the information matrix formulation. Numerical simulations of a satellite equipped with sun sensors, magnetometers, and a rate gyro demonstrate the ability of the qEKF algorithm to accurately estimate the attitude and gyro bias. Three test cases are examined consisting of synchronized magnetometer and sun sensor measurements available throughout the orbit, only magnetometer measurements, and a more realistic case with asynchronous sun sensor and magnetometer measurements accounting for a period of eclipse. In all cases the qEKF converges to an accurate estimate with identical results as obtained by the SOAR filter for the same test cases. A final test case demonstrates the advantage of the qEKF over a standard extended Kalman filter represented by MEKF when the initial errors are large.

5.2 Future Work

The most significant task that remains is the expansion of the q-method extended Kalman filter to allow the vector measurements, $\mathbf{y}_i$ and $\mathbf{n}_i$, to be functions of non-attitude states. Examples of interest include attitude sensor biases and misalignments which affect the measurement vectors, $\mathbf{y}_i$, and the spacecraft orbital position which is used to obtain the the magnetometer reference vector, $\mathbf{n}_{mag}$. First, the attitude covariance obtained from the q-method and given in Eq. (2.62) must be modified to incorporate the additional uncertainty added to the measurements by the associated non-attitude states. This is accomplished by appropriately modifying the measurement error given by Eqs. (2.28) and (2.33). In the derivations of Sections 3.2.1 and 3.2.2 it was assumed that the measurements are only a function of the attitude and not any other states. This restriction results in the measurement sensitivity or observation partial matrix for the non-attitude states equal to the zero matrix, $\mathbf{H_s} = \mathbf{0}$. For the case of measurements dependent on non-attitude states, this assumption is no longer true and the observation partial matrix, $\mathbf{H_s} \neq \mathbf{0}$. As a result, additional terms will be present in the derivation of the linear optimal Kalman gains from Eq. (3.17). This necessitates appropriate modifications to the non-attitude state Kalman gain, $\mathbf{K_s}$, and the full state covariance update. Finally, a new selection criteria for the scalar weights of the Wahba problem, a_i, must be derived as the QUEST measurement model does not account for the additional uncertainty in the measurements due to the related non-attitude states. Attempts to select the weights which minimize the trace of the attitude error covariance do not appear feasible. Potential solutions include bounding the maximum error from the maximum eigenvalue of the measurement covariance. A trade study may then be performed by ranging the scalar weight from zero to this maximum value in order to identify the choice that gives best performance on a case by case basis. These extensions to the q-method extended Kalman filter will provide the capability of estimating the nonlinear attitude and any observable state without the need for iteration and in a single Kalman filter.

Bibliography

[1] J. E. Keat, "Analysis of least-squares attitude determination routine doaop," Tech. Rep. Tech. Rep. CSC/TM-77/6034, Computer Sciences Corporation, February 1977.

[2] F. L. Markley, Y. Cheng, J. L. Crassidis, and Y. Oshman, "Averaging quaternions," *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 30, pp. 1193–1197, July-August 2007.

[3] J. A. Christian and E. G. Lightsey, "Sequential optimal attitude recursion filter," *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 33, pp. 1787–1800, November-December 2010.

[4] H. D. Black, "A passive system for determining the attitude of a satellite," *AIAA Journal*, vol. 2, pp. 1350–1351, July 2004.

[5] J. R. Wertz, ed., *Spacecraft Attitude Determination and Control*. D. Reidel Publishing Company, 1978.

[6] H. D. Black, “Early development of transit, the navy navigation satellite system,” *Journal of Guidance, Control, and Dynamics*, vol. 13, pp. 577–585, July 1990.

[7] G. Wahba, “A least squares estimate of satellite attitude,” *SIAM Review*, vol. 7, p. 409, July 1965. Problem 65-1.

[8] M. Shuster and S. Oh, “Three-axis attitude determination from vector observiations,” *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 4, no. 1, pp. 70–77, 1981.

[9] M. D. Shuster, “Maximum likelihood estimation of spacecraft attitude,” *The Journal of the Astronautical Sciences*, vol. 37, pp. 79–88, January-March 1989.

[10] G. M. Lerner, *Three-Axis Attitude Determination.* Dordrecht, The Netherlands: Reidel Publishing Company, 1978.

[11] J. Crassidis and J. Junkins, *Optimal Estimation of Dynamic Systems.* Chapman and Hall/CRC Applied Mathematics and Nonlinear Science Series, CRC PressINC, 2011.

[12] F. L. Markley, “Attitude determination using vector observiations: A fast optimal matrix algorithm,” *The Journal of the Astronautical Sciences*, vol. 41, pp. 261–280, April-June 1993.

[13] D. Mortari, “Esoq: A closed-form solution to the wahba problem,” *The Journal of the Astronautical Sciences*, vol. 45, pp. 195–204, April-June 1997.

[14] D. Mortari, "Esoq-2 single-point algorithm for fast optimal spacecraft attitude determination," *Advances in the Astronautical Sciences*, vol. 95, no. 2, pp. 817–826, 1997.

[15] J. C. Gower and G. B. Dijksterhuis, *Orthogonal Procrustes problems.* Oxford University Press, 2004.

[16] F. L. Markley, "Attitude determination using vector observiations and the singular value decomposition," *The Journal of the Astronautical Sciences*, vol. 36, pp. 245–258, July-September 1988.

[17] F. L. Markley and D. Mortari, "How to estimate attitude from vector observiations," in *AAS/AIAA Astrodynamics Specialist Conference*, no. 99-427, August 1999.

[18] H. Schaub and J. Junkins, *Analytical mechanics of space systems.* AIAA Education Series, American Institute of Aeronautics and Astronautics, 2003.

[19] A. Gelb, J. F. K. Jr., R. A. N. Jr., C. F. Price, and A. A. S. Jr., *Applied Optimal Estimation.* The M.I.T. Press, 1974.

[20] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals adn Applied Kalman Filtering.* John Wiley & Sons, third ed., 1997.

[21] P. Zarchan and H. Musoff, *Fundamentals Of Kalman Filtering: A Practical Approach*. No. v. 208 in Progress in Astronautics and Aeronautics Series, American Institute of Aeronautics and Astronautics, 2005.

[22] E. Lefferts, F. Markley, and M. Shuster, "Kalman filtering for spacecraft attitude estimation," *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 5, pp. 417–429, September-October 1982.

[23] C. Dong, L. Xu, Y. Chen, and Q. Wang, "Networked flexible spacecraft attitude maneuver based on adaptive fuzzy sliding mode control," *Acta Astronautica*, vol. 65, no. 11, pp. 1561 – 1570, 2009.

[24] M. D. Shuster, "A simple kalman filter and smoother for spacecraft attitude," *The Journal of the Astronautical Sciences*, vol. 37, pp. 89–106, January-March 1989.

[25] I. Bar-Itzhack, "Request: A recursive quest algorithm for sequential attitude determination," *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 19, pp. 1034–1038, September-October 1996.

[26] M. D. Shuster, "Filter quest or request," *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 32, pp. 643–645, March-April 2009.

[27] D. Choukroun, I. Bar-Itzhack, and Y. Oshman, "Optimal-request algorithm for attitude determination," *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 27, pp. 418–425, May-June 2004.

[28] F. L. Markley, “Attitude determination and parameter estimation using vector observiations: Theory,” *The Journal of the Astronautical Sciences*, vol. 37, pp. 41–58, January-March 1989.

[29] M. L. Psiaki, “Attitude-determination filtering via extended quaternion estimation,” *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 23, pp. 206–214, March-April 2000.

[30] M. D. Shuster, “The generalized wahba problem,” *The Journal of the Astronautical Sciences*, vol. 54, pp. 245–259, April-June 2006 2006.

[31] M. D. Shuster, “Kalman filtering of spacecraft attitude and the quest model,” *The Journal of the Astronautical Sciences*, vol. 38, pp. 377–393, July-September 1990.

[32] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes With Applications to Guidance.* American Mathematical Soc., 2 ed., 1987.

[33] J. Wertz and W. Larson, *Space Mission Analysis and Design.* Space Technology Library, Microcosm Press, 1999.

[34] R. Farrenkopf, “Analytic steady state accuracy solutions for two common spacecraft attitude estimators,” *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 1, pp. 282–284, July 1978.

[35] M. Abdelrahman and S.-Y. Park, "Simultaneous spacecraft attitude and orbit estimation using magnetic field vector measurements," *Aerospace Science and Technology*, vol. 15, pp. 653–669, 2011.

[36] I. Bar-Itzhack and Y. Oshman, "Attitude determination from vector observations: Quaternion estimation," *IEEE Transactions on Aerospace and Electrical Systems*, vol. 21, pp. 128–135, January 1985.

[37] Y. Cheng, J. Crassidis, and F. L. Markley, "Attitude estimation for large field-of-view sensors," *The Journal of the Astronautical Sciences*, vol. 54, pp. 433–448, July-Dec 2006.

[38] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *Journal of Guidance, Control, and Dynamics*, vol. 30, pp. 12–28, Jan. 2007.

[39] P. Hughes, *Spacecraft Attitude Dynamics*. Dover Publications, Incorporated, 2012.

[40] D. Hull, *Optimal Control Theory for Applications*. Mechanical Engineering Series, Springer, 2003.

[41] A. Jazwinski, *Stochastic Processes and Filtering Theory*. Mathematics in Science and Engineering, Elsevier Science, 1970.

[42] J. Kuipers, *Quaternions and Rotation Sequences: A Primer With Applications to Orbits, Aerospace, and Virtual Reality.* Princeton University Press, 1999.

[43] C. Lawson and R. Hanson, *Solving Least Squares Problems.* Classics in applied mathematics, Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 1974.

[44] F. L. Markley, "Attitude error representations for kalman filtering," *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 26, pp. 311–317, March-April 2003.

[45] F. L. Markley, "Attitude error representations for kalman filtering," *Journal of Guidance, Control, and Dynamics*, vol. 26, pp. 311–317, March-April 2003.

[46] F. L. Markley, "Attitude determination and parameter estimation using vector observiations," *The Journal of the Astronautical Sciences*, vol. 39, pp. 367–381, July-September 1991.

[47] F. L. Markley and D. Mortari, "Quaternion attitude estimation using vector observiations," *The Journal of the Astronautical Sciences*, vol. 48, pp. 359–380, April-September 2000.

[48] D. Mortari, "Second estimator of the optimal quaternion," *Journal of Guidance, Control, and Dynamics*, vol. 23, no. 5, pp. 885–888, 2000.

[49] D. Mortari and M. Majji, “Multiplicative measurement model and single-point attitude estimation,” in *F. Landis Markley Astronautics Symposium*, (Cambridge, MD), American Astronautical Society, June 29 - July 2 2008.

[50] A. Papoulis and S. Pillai, *Probability, random variables, and stochastic processes.* McGraw-Hill electrical and electronic engineering series, McGraw-Hill, 2002.

[51] M. L. Psiaki, F. Martel, and P. K. Pal, “Three-axis attitude determination via kalman filtering of magnetometer data,” *Journal of Guidance, Navigation, Control, and Dynamics*, vol. 13, pp. 506–514, May-June 1990.

[52] M. D. Shuster, “New quests for better attitudes,” in *Flight Mechanics/Estimation Theory Symposium*, (Greenbelt, MD), NASA Goddard Space Flight Center, 21-23 May 1991.

[53] H. W. Sorenson, *Parameter Estimation: Principles and Problems.* New York, NY: Macrel Dekker, Inc, 1980.

[54] D. Vallado and W. McClain, *Fundamentals of Astrodynamics and Applications.* Space Technology Library, Springer, 2001.

[55] R. Zanetti, “A multiplicative residual approach to attitude kalman filtering with unit-vector measurements,” in *AAS Flight Mechanics Conference*, (San Diego, CA), 14-17 February 2010.

[56] R. Zanetti, T. Ainscough, J. A. Christian, and P. Spanos, "Q-method extended kalman filter," in *AAS/AIAA Spaceflight Mechanics Meeting*, February 10-14 2013.