

# STOCHASTIC MODELING AND FILTERING OF THE ATTITUDE QUATERNION

Daniel Choukroun\*

A novel continuous-time stochastic differential equation (SDE) for spacecraft attitude quaternion kinematics with state-multiplicative noise and a novel continuous-time *exact* optimal quaternion filter are developed in the framework of Itô (mean-square) calculus. The quaternion Itô SDE contains dissipative terms that ensure the mean-square stability of the process. The filter gain computations, which include coupled Riccati equations of the estimation error matrix and of the quaternion second-order moment, are not estimate-dependent and can therefore be performed *off-line*. The case of gyro errors including white noise with independent identically distributed components and additive constant biases is considered. Extensive Monte-Carlo simulations show that, for high signal to noise ratio, the novel approach can increase the accuracy of a conventional Kalman filter by orders of magnitudes.

## INTRODUCTION

Much effort has been produced in order to develop better stochastic models of spacecraft (SC) attitude quaternions<sup>1</sup> and of optimal stochastic estimators in the realm of Kalman filtering.<sup>2,3,4,5,6</sup> There is an advantage of working in the continuous-time setting since continuous-time analysis usually provides upper bounds on the performance of the discrete-time equivalent filters. Also, analysis in continuous-time usually yields less cumbersome expressions. For that purpose, Kalman-Bucy filtering theory<sup>7</sup> and its extension to non-linear, non-Gaussian systems are the classical tools.

One central issue, towards the development of quaternion Kalman filters, lies in the elaboration of a continuous-time model for the quaternion process equation, in particular, in the passage from the physical process equation to its mathematical approximation. The drawback that stems from previously proposed models is briefly explained next. Consider the Langevin<sup>8,9</sup> stochastic differential equation (SDE) for the quaternion,  $\mathbf{q}_t$ , featuring a nominal deterministic angular velocity,  $\boldsymbol{\omega}_t$ , with an additive error with respect to the true value,  $\boldsymbol{\epsilon}_t$ \*:

$$\dot{\mathbf{q}}_t = \frac{1}{2} \boldsymbol{\Omega}(t) \mathbf{q}_t - \frac{1}{2} \boldsymbol{\Xi}(\mathbf{q}_t) \boldsymbol{\epsilon}_t \quad (1)$$

where  $\boldsymbol{\Omega}$  is a skew-symmetric matrix functions of  $\boldsymbol{\omega}_t$  and  $\boldsymbol{\Xi}(\mathbf{q}_t)$  is a linear matrix function of  $\mathbf{q}_t$ . Since the measured angular velocity, acquired via e.g. a triad of gyroscopes, is usually more precise than the nominal angular velocity, the commonly-used practice is to substitute the measured velocity to its nominal value. Henceforth, following the common practice,  $\boldsymbol{\omega}_t$  will denote the measured angular velocity and  $\boldsymbol{\epsilon}_t$  will denote a zero-mean additive Gaussian white noise. Assuming that  $\boldsymbol{\epsilon}_t$  is white ensures the Markov property for the  $\mathbf{q}_t$ -process. It is known, however, that Eq. (1) is not well-defined:<sup>8,10</sup> the  $\boldsymbol{\epsilon}_t$ -process being delta-correlated is not mean-square (*m.s.*) integrable and, since its sample paths are delta functions, it is not integrable almost everywhere (*a.e.*). Furthermore, since white noise is an abstraction and not a physical process, what one really means by Eq. (1) in practice is probably “an equation driven by a stationary Gaussian process with a spectral density that is flat over a wide range of frequencies” [10, p. 156]. In that case however, the  $\mathbf{q}_t$ -process loses the Markov property.<sup>†</sup>

The contribution of this work is twofold. First, it introduces a novel continuous-time stochastic model for the quaternion process equation. The quaternion modelling is developed using Itô (mean-square) calculus, which is the mathematical background for the Kalman-Bucy filter. This model provides a precise meaning to the  $\mathbf{q}_t$ -process while retaining the crucial Markov property. The Itô SDE that approximates Eq. (1) presents dissipative terms in its drift term. The special case of white noise in the gyro error is first considered but, continuing an early work,<sup>11</sup> the case of constant additive biases is also developed for the numerical study. Although the  $\mathbf{q}_t$ -process is not Gaussian, standard

\*Ben-Gurion University, Mechanical Engineering Dept., POB 653, Beer-Sheva, 84105, Israel. danielch@bgu.ac.il. AIAA, Member. Pearlstone Center for Aeronautical Engineering Studies, Member. This work is supported in part by ISF–The Israeli Science Foundation.

\*The term  $\boldsymbol{\epsilon}_t$  is usually referred to as environment dispersion

†As opposed to constant biases or slowly-varying gyro drifts, modeling the  $\boldsymbol{\epsilon}_t$ -process as a first-order Markov process is not desired for it would yield a dimension augmentation, it would require guessing the correlation-time and its estimates would not be of physical interest.

techniques<sup>12,13,14</sup> yield closed-form deterministic equations for the second-order moment of  $\mathbf{q}_t$ . It will be seen that the dissipative term of the Itô equation ensures the mean-square stability of the  $\mathbf{q}_t$ -process. Depending on the quality of the gyroscopes, neglecting the dissipative terms might have serious impact on the filter performance during long duration time-propagation stages.

In addition, based on the proposed quaternion state-space model in Itô form, an *exact* best linear unbiased (BLU) filter is developed. It is seen that the Itô dissipative terms are necessary to the filter stability. The implementation of the filter must be consistent with the Itô dynamics. The main features of this filter are that it is free from approximations and the gain computations are deterministic. The drawback is that it requires computing simultaneously the estimation error covariance matrix as well as the second-order moment of the  $\mathbf{q}_t$ -process. These computations, however, can be performed off-line *since they are not estimate-dependent*.

The quaternion stochastic modelling is presented in the next section. Then the best linear unbiased filter is developed. The subsequent section includes a numerical study. Conclusions are proposed in the last section.

## QUATERNION STOCHASTIC MODELING

### Physical Process Equation

Consider a rigid-body spacecraft (SC) rotating with respect to some reference Cartesian coordinates frame,  $\mathcal{R}$ , with an angular velocity vector  $\boldsymbol{\omega}_t^o$ , as resolved along the SC Cartesian frame,  $\mathcal{B}_t$ . Let  $\mathbf{q}_t^T = [e^T, q]$  denote the quaternion of rotation from  $\mathcal{R}$  to  $\mathcal{B}_t$ . The physical process equation modelling the quaternion kinematics is the following well-known deterministic ordinary differential equation [1, p. 511]:

$$\dot{\mathbf{q}}_t = \frac{1}{2} \Omega^o(t) \mathbf{q}_t; \quad \mathbf{q}(0) = \mathbf{q}_0; \quad t \in [t_0, T] \quad (2)$$

where

$$\Omega^o(t) = \begin{bmatrix} -[\boldsymbol{\omega}_t^o \times] & \boldsymbol{\omega}_t^o \\ -\boldsymbol{\omega}_t^{oT} & 0 \end{bmatrix} \quad (3)$$

and  $[\boldsymbol{\omega}_t^o \times]$  denotes the cross-product matrix. The true angular velocity,  $\boldsymbol{\omega}_t^o$  is in general only known with an additive error  $\boldsymbol{\epsilon}_t$ ; that is,

$$\boldsymbol{\omega}_t = \boldsymbol{\omega}_t^o + \boldsymbol{\epsilon}_t \quad (4)$$

where  $\boldsymbol{\epsilon}_t$  may be a measurement error vector from gyroscopes outputs, including scale-factors, constant biases, slowly varying drifts and rapidly fluctuating disturbances. In the following, we will focus on the latter case since the proposed results of this work are related to the modelling of the wide-band disturbances. The treatment of the more general case is however straightforward using state-augmentation techniques. For analytical purposes, the usual approach consists in modelling  $\boldsymbol{\epsilon}_t$  as a zero-mean Gaussian white noise process, thus assuming that  $E\{\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_\tau^T\} = Q_t \delta(t - \tau)$ , where  $\delta(\cdot)$  denotes the Dirac delta function and  $E\{\cdot\}$  denotes the unconditional expectation operator over all the underlying random variables. Inserting Eq. (4) into Eq. (2) yields the known quaternion stochastic differential equation<sup>2,3</sup> in Langevin form<sup>9</sup> [8, p. 94]:

$$(L) \quad \dot{\mathbf{q}}_t = \frac{1}{2} \Omega(t) \mathbf{q}_t - \frac{1}{2} \Xi(\mathbf{q}_t) \boldsymbol{\epsilon}_t; \quad \mathbf{q}(0) \stackrel{a.e.}{=} \mathbf{q}_0; \quad t \in [t_0, T] \quad (5)$$

$$\Xi(\mathbf{q}_t) = \begin{bmatrix} [\mathbf{e} \times] + q I_3 \\ -\mathbf{e}^T \end{bmatrix} \quad (6)$$

where *a.e.* stands for “almost everywhere”. As mentioned in the Introduction, the  $\boldsymbol{\epsilon}_t$  process is delta-correlated, so  $\boldsymbol{\epsilon}_t$  is not mean square integrable. The sample functions are delta functions, so  $\boldsymbol{\epsilon}_t$  is not integrable with probability one (w.p.1). “Consequently, [Eq. (5)] has no mathematical meaning” [8, p. 122]. On the other hand, modeling  $\boldsymbol{\epsilon}_t$  as a wide-band process (a colored noise) would imply for Eq. (5) to loose the convenient Markov property.

### Itô Stochastic Differential Equation

**Proposition 1:** *From the theory of stochastic differential equations,<sup>15,16</sup> the Itô differential form for Eq. (5) is*

$$(I) \quad d\mathbf{q}_t = \frac{1}{2} \left[ \Omega(t) - \frac{tr Q}{4} I_4 \right] \mathbf{q}_t dt - \frac{1}{2} \Xi(\mathbf{q}_t) d\boldsymbol{\beta}_t; \quad \mathbf{q}_t(0) \stackrel{a.e.}{=} \mathbf{q}_0; \quad t \in [t_0, T] \quad (7)$$

where  $d\boldsymbol{\beta}_t$  denotes infinitesimal independent increments of a Brownian motion  $\boldsymbol{\beta}_t$  such that  $E\{\boldsymbol{\beta}_t \boldsymbol{\beta}_\tau^T\} = Q(t - \tau)$ , and  $Q$  is assumed diagonal with elements  $\sigma_1^2, \sigma_2^2, \sigma_3^2$ .

*Proof.* Equation (7) is obtained via Wong and Zakai’s theorem.<sup>15</sup> See a detailed proof in Ref. [11].  $\square$

From the viewpoint of analysis, Itô's equation is preferable to the physical process equation with wide-band noise since it allows a precise meaning of the process  $\mathbf{q}_t$  while retaining the Markov property. The process  $\mathbf{q}_t$  is the solution in the mean-square sense to Eq. (7) on  $[t_0, T]$  and is uniquely determined, in the mean-square sense, by the initial conditions  $\mathbf{q}_0$ . Let  $F_I(t)$  denote the dynamics matrix in Eq. (7). It can be shown that its spectrum is given as follows:

$$\text{Sp}(F_I) = \left\{ -\frac{\text{tr}Q}{8}; -\frac{\text{tr}Q}{8}; -\frac{\text{tr}Q}{8} \pm j \frac{1}{2} \|\boldsymbol{\omega}_t\| \right\} \quad (8)$$

The diagonal terms in  $F_I$  are interpreted as dissipative terms in the Itô formulation. Without them,  $\mathbf{q}_t$  would diverge in the mean square sense. This is seen by examining the propagation equation for the second-order moment of  $\mathbf{q}_t$ .

### Second-Order Moment

As seen from Eq. (7), and as recognized in earlier works,<sup>2,3</sup> the quaternion process equation is perturbed by a state multiplicative noise, which is linear in  $\mathbf{q}_t$ . Thanks to this linear-in- $\mathbf{q}_t$  property, and in spite of the non-Gaussian nature of  $\mathbf{q}_t$ , the evolution equation for the second-order moment of  $\mathbf{q}_t$ ,  $X_t = E\{\mathbf{q}_t \mathbf{q}_t^T\}$ , can be developed in closed-form. The general case can be found e.g. in Refs. [13] or [14], from which we use the formalism. The following lemma is central for subsequent developments.

**Itô Lemma for linear systems with multiplicative noise [14, p. 212]:** Consider

$$d\mathbf{x}_t = A_t \mathbf{x}_t dt + D_t \mathbf{x}_t d\beta_t \quad (9)$$

where  $\beta_t$  is a scalar valued standard Wiener process, such that  $E\{d\beta_t^2\} = dt$ . Defining  $X_t = E\{\mathbf{x}_t \mathbf{x}_t^T\}$ , the propagation equation for  $X_t$  is

$$\dot{X}_t = A_t X_t + X_t A_t^T + D_t X_t D_t^T; \quad X(0) = X_0 \quad (10)$$

**Proposition 2:** Consider the case where the components of  $\beta_t$  are independently identically distributed with variance parameter  $\sigma^2$ ; that is,  $Q = \sigma^2 I_3$ . Let  $X$  denote the second-order moment of  $\mathbf{q}_t$ ,  $X_t = E\{\mathbf{q}_t \mathbf{q}_t^T\}$ , then

$$\dot{X}_t = \frac{1}{2} \left[ \Omega(t) - \frac{3\sigma^2}{4} I_4 \right] X_t + X_t \frac{1}{2} \left[ \Omega(t) - \frac{3\sigma^2}{4} I_4 \right]^T + \frac{\sigma^2}{4} [(\text{tr} X_t) I_4 - X_t]; \quad X(0) = X_0 \quad (11)$$

*Proof.* See Ref. [11]. □

Notice that without the dissipative terms,  $-\frac{3\sigma^2}{4} I_4$ , Eq. (11) would have unstable poles. Conditions for boundedness of linear matrix equations are given in Refs. [20]. Furthermore, Eq. (11) provides an exact and deterministic computation of  $X_t$ . These results are central in the implementation of the Kalman filter. It will indeed allow *off-line computation of the filter gains*, which is usually not the case in typical quaternion Extended Kalman filters. Furthermore, the existence of the gains will be ensured via the convergence of Eq. (11).

### Length of $\mathbf{q}_t$

**Proposition 3:** Let the  $\varphi_t(\mathbf{q}_t)$  denote the squared length of  $\mathbf{q}_t$  solution of Eq. (7); that is

$$\varphi_t = \mathbf{q}_t^T \mathbf{q}_t; \quad \varphi(0) \stackrel{a.e.}{=} \|\mathbf{q}_0\|^2 = 1 \quad (12)$$

Then,

$$\varphi_t \stackrel{m.s.s.}{=} \|\mathbf{q}_0\|^2 = 1; \quad t \in [t_0, T] \quad (13)$$

*Proof.* It is straightforward to check that  $\varphi_t$  satisfies the continuous and differentiability conditions needed in order to apply the Itô differential rule [8, p. 112]. The above result stems from a direct application of Itô's rule. □

The above proposition states that, once the quaternion is initialized with a given length, which must be equal to one in order for a quaternion of rotation, it keeps the same length, in the mean square sense, at any future time. Hence the stochastic formulation of the quaternion dynamics addresses, via the dissipative terms, the issue of quaternion length preservation in the process model equation.

Let  $x_t$  denote the expected value of  $\|\mathbf{q}_t\|^2$ ; that is,  $x_t = E\{\|\mathbf{q}_t\|^2\}$  and assume that the stochastic correction does not appear in Eq. (11). Then a straightforward application of Itô's rule yields the following propagation equation for  $x_t$ :

$$\dot{x}_t = \frac{3\sigma^2}{4} x_t; \quad x(0) = E\{\|\mathbf{q}(0)\|^2\} = 1 \quad (14)$$

Clearly, Eq. (14) is diverging. With inertial rate gyroscopes, this divergence rate may not be a practical issue. For low grade gyroscopes however, with  $\sigma \sim 0.1 \text{ deg/sec}^{3/2}$ , and read-out rates of the order of KHz, the error growth rate may yield the length of  $\mathbf{q}_t$  to double within a couple of weeks.

### Measurement Model

New sensor packages that uses star trackers and yield the SC attitude in terms of the attitude quaternion became available recently. Therefore, it became possible to use the quaternion supplied by such sensors as measurements. This approach was indeed used in Ref. [17] in discrete-time. The continuous-time measurement model equation, in Itô form, is assumed to be as follows:

$$d\mathbf{z}_t = \mathbf{q}_t dt + d\mathbf{n}_t \quad (15)$$

where  $\mathbf{n}_t$  is a Brownian motion in  $\mathbb{R}^4$  with intensity matrix  $R_t$ . If the measured quaternion is not constrained to be unit-norm, which implies that the noise may be assumed independent of the true quaternion, then the intensity matrix,  $R_t$ , may be assumed full rank. Notice that if the measured quaternion is constrained to be of norm one, previous works about unit-norm vector measurement probability distribution (see e.g. [18]) showed that  $R_t$ , while not being rank deficient, would however be badly conditioned. Both cases for  $R_t$  will be tested in the numerical study section. In addition to the simple proposed model of Eq. (15), other types of quaternion measurement equations, possibly including state-dependent non-white noises, can be dealt with along the approach previously adopted for the process model equation.

## QUATERNION BEST LINEAR UNBIASED (BLU) FILTERING

The filter is developed as the best linear unbiased (BLU) filter using the general results from Refs. [13, 14]. The summary is described in the following and the detailed development is provided in Ref. [11].

### Unconstrained BLU Filter Summary

The continuous-time equations of the unconstrained BLU quaternion filter are

$$F_t(t) = \frac{1}{2} \left[ \Omega(t) - \frac{3\sigma^2}{4} I_4 \right] \quad (16)$$

$$d\hat{\mathbf{q}}_t = F_t(t) \hat{\mathbf{q}}_t dt + K_t(d\mathbf{z}_t - \hat{\mathbf{q}}_t dt) \quad (17)$$

$$\dot{P}_t = (F_t - K_t)P_t + P_t(F_t - K_t)^T + \frac{\sigma^2}{4} [(tr X_t) I_4 - X_t] + K_t R_t K_t^T \quad (18)$$

$$\dot{X}_t = F_t X_t + X_t F_t^T + \frac{\sigma^2}{4} [(tr X_t) I_4 - X_t] \quad (19)$$

where the optimal gain is computed as

$$K_t = P_t R_t^{-1} \quad (20)$$

with  $\hat{\mathbf{q}}(0) = E\{\mathbf{q}(0)\}$  (for unbiasedness),  $X(0) = E\{\mathbf{q}(0)\mathbf{q}^T(0)\}$  and  $P(0) = \text{cov}\{\mathbf{q}(0)\}$ .

*Remark 1: If the unconditional moments are unknown and the filter is initialized with a priori values, for instance  $\hat{\mathbf{q}}(0)^T = [0001]$ , the error process becomes asymptotically unbiased as  $t \rightarrow \infty$ .*

*Remark 2: The BLU estimate is obtained uniformly in time. It is exact and enjoys therefore the orthogonality property: If  $E\{\tilde{\mathbf{q}}(t_0)\tilde{\mathbf{q}}(t_0)^T\} = 0$  then  $E\{\tilde{\mathbf{q}}(t)\tilde{\mathbf{q}}(t)^T\} = 0 \forall t \in [t_0, T]$ . Furthermore,  $E\{\tilde{\mathbf{q}}(t)Z_t\} = 0 \forall t \in [t_0, T]$  where  $\tilde{\mathbf{q}}(t) = \mathbf{q}_t - \hat{\mathbf{q}}_t$  and  $Z_t$  denotes any function of the past measurement history.*

*Remark 3: These filtering equations do not require any approximation. They are similar to the Kalman-Bucy filter<sup>7</sup> but require that the second-order moment of  $\mathbf{q}_t$  be calculated. This is the computational penalty for including the state-dependent noise. Furthermore, the stability of this filter, different from the Kalman filter, depends upon the boundedness of  $X_t$ .<sup>20</sup>*

*Remark 4: The gain computation, similarly to the KF, are deterministic. For this reason the Langevin equation is obtained from the Itô form, Eq. (17), without stochastic correction; that is,*

$$\dot{\hat{\mathbf{q}}}_t = F_t(t) \hat{\mathbf{q}}_t + K_t(\dot{\mathbf{z}}_t - \hat{\mathbf{q}}_t); \quad \hat{\mathbf{q}}(0) = E\{\mathbf{q}(0)\} \quad (21)$$

*In simulation, we will use the Langevin equation (21).*

*Remark 5: As opposed to a classical extended Kalman filter, the gain computations are independent from the value of the estimate and can be therefore performed off-line. Thus, in practice, the computational burden inherited from the computation of  $X_t$  is not much of a drawback.*

## Discussion: Approximate quaternion filtering using nonlinear quaternion measurements

The previous analysis and results assumed that a full quaternion measurement was acquired along the time. Frequently, however, spacecraft attitude sensors only deliver measurements that are nonlinear functions of the attitude quaternion. This happens, for instance, when the continuously sampled quantity is the Earth Magnetic field, i.e., the three components of the field in the SC Cartesian coordinates frame. The state-space equations for the quaternion  $\mathbf{q}$  are then consisting of Eq. (7) and of the following nonlinear measurement model equation:

$$d\mathbf{z}_t = \mathbf{h}(\mathbf{q}_t) dt + d\mathbf{n}_t \quad (22)$$

where  $\mathbf{h}(\mathbf{q}_t)$  is defined as:<sup>2</sup>

$$\mathbf{h}(\mathbf{q}_t) = A(\mathbf{q}_t) \mathbf{r}_t \quad (23)$$

and, in Eq. (23),  $A(\mathbf{q}_t)$  denotes the Direction Cosine Matrix (DCM) as a function of the quaternion,  $\mathbf{r}_t$  denotes the components of, say, the Earth Magnetic field along a reference Cartesian coordinates frame, and  $\mathbf{n}_t$  is a Brownian motion with intensity matrix  $R_t$ . Notice that, providing that the measurements are not of unit norm,  $R_t$  is not ill-conditioned.

For this type of measurements, the (linear) analysis that brought an exact BLU filter of  $\mathbf{q}$  fails because the DCM is a quadratic function of  $\mathbf{q}$ . Nevertheless, standard approaches for developing approximate nonlinear filters, which approximate the *conditional* mean and variance of the state, can then be applied. As an example, it is straightforward to develop the quaternion Extended Kalman filtering equations for the conditional mean and variance [8, p. 338]:

$$d\hat{\mathbf{q}}_t = F_t \hat{\mathbf{q}}_t dt + P_t \widehat{H}_t^T R_t^{-1} [d\mathbf{z}_t - \mathbf{h}(\hat{\mathbf{q}}_t) \mathbf{r}_t dt] \quad (24)$$

$$\dot{P}_t = F_t P_t + P_t F_t^T + \frac{\sigma^2}{4} [(tr X_t) I_4 - X_t] - P_t \widehat{H}_t^T R_t^{-1} \widehat{H}_t P_t \quad (25)$$

$$\dot{X}_t = F_t X_t + X_t F_t^T + \frac{\sigma^2}{4} [(tr X_t) I_4 - X_t] \quad (26)$$

where  $F_t$  includes the Itô correction term, as given in Eq. (16), and  $\widehat{H}_t$  is the following gradient matrix:<sup>21</sup>

$$\widehat{H}_t = 2 \left[ \widehat{\mathbf{e}}^T \mathbf{r} I_3 + \widehat{\mathbf{e}} \mathbf{r}^T - \mathbf{r} \widehat{\mathbf{e}}^T + 2 \widehat{q} [\mathbf{r} \times], \quad \widehat{q} \mathbf{r} + [\mathbf{r} \times] \widehat{\mathbf{e}} \right] \quad (27)$$

where  $\widehat{\mathbf{e}}$  and  $\widehat{q}$  are the vector and scalar part of  $\widehat{\mathbf{q}}$ .

The differential equation describing the approximate nonlinear filter, Eq. (24), is an (Itô) stochastic differential equation. The coefficient of  $d\mathbf{z}_t$  is random since the gradient matrix  $\widehat{H}_t$  is function of the random estimate  $\widehat{\mathbf{q}}_t$ . Therefore, care must be taken in simulating Eq. (24) on a computer. Unlike the BLU filtering equation previously obtained [Eq. (21)], Eq. (24) needs to be modified. One way of simulating the filter is first to transform the filter equation by subtracting the Itô correction term from the drift term in Eq. (24) and, then, to use standard integration techniques.

Notice that the resulting filter *still* embeds the improvements in the quaternion model that were brought by Itô's calculus tools. In particular, the gain computations involve the *exact* expression for the gyro quaternion-dependent intensity matrix. This matrix might be computed beforehand, via the propagation of the matrix  $X_t$  in Eq. (26), and its values stored *off-line*. On the other hand, as is well known, the filter gain computations are coupled with the estimate computations and have to be performed on-line.

## NUMERICAL STUDY

Consider a spacecraft rotating around its center of mass at an angular rate,  $\boldsymbol{\omega}^o$ , given as follows:

$$\boldsymbol{\omega}^o = \sin(2\pi t/150) \begin{bmatrix} 1 & -1 & 1 \end{bmatrix}^T \quad deg/sec$$

It is assumed that the spacecraft measures the angular velocity via a three-axis rate gyro with an output error that includes three additive biases as well as a white Gaussian vector noise with standard deviation  $\sigma/\sqrt{\Delta t}$ , where  $\Delta t$  is the time interval between two consecutive gyro readouts. The simulation runs are performed over a time interval of 6000 seconds, which corresponds to one revolution of a typical Low-Earth-Orbit (LEO) satellite.

### Model validation: Process equation

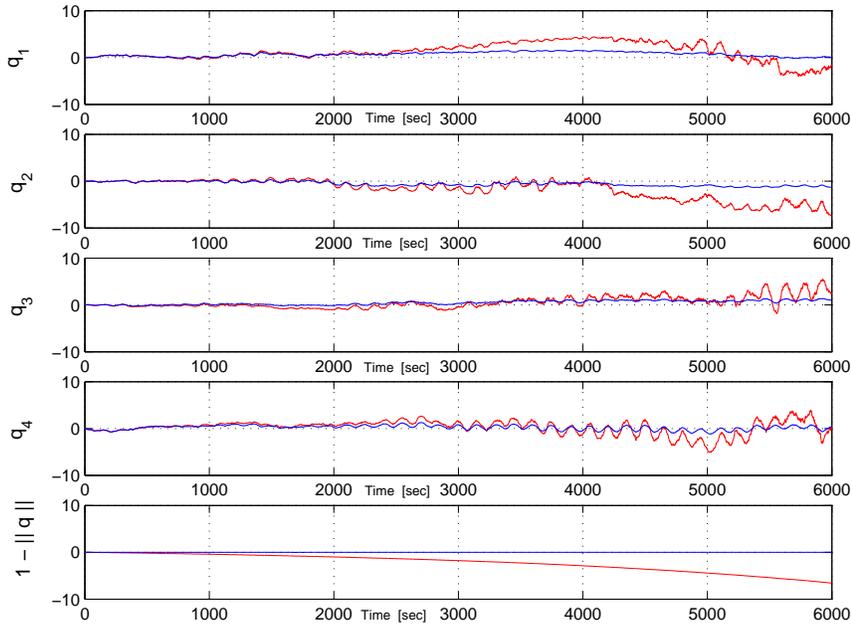
A first set of simulations were performed in order to numerically investigate the validity of the Langevin quaternion design model, Eq. (5), and of the Itô design model, Eq. (7), and to compare between them. For that purpose, in order to focus on the white-noise simulation issue, the value of the biases were set equal to zero. The errors in both design

models were computed as the difference between the model quaternion ( $\mathbf{q}_t^L$  for Langevin or  $\mathbf{q}_t^I$  for Itô) and the truth-model (deterministic) quaternion, as given by Eq. (2). The following error indices were defined in order to quantify the modeling errors:  $\delta_q^I = \max_t \|\mathbf{q}_t - \mathbf{q}_t^I\|$  and  $\delta_q^L = \max_t \|\mathbf{q}_t - \mathbf{q}_t^L\|$ . Table 1 is a comparative table that presents the values of the ratio  $\delta_q^I/\delta_q^L$  for a range of values in  $\sigma$  and  $\Delta t$ . Notice that these values were chosen such as to preserve the validity condition related to the Itô correction term<sup>15</sup>[p. 162], that is, such that  $\Delta t B \ll 1$ , where  $B$  corresponds to the bandwidth of the modeled equation. Adapting the expression for  $B$  to the current equation yields the following necessary condition on  $\sigma$  and  $\Delta t$ :  $\sigma^2 \Delta t \ll 1$ . As expected, the modeling errors,  $\delta_q^L$  and  $\delta_q^I$ , are growing with  $\sigma$  and  $\Delta t$ , with common rapid growths around  $\sigma \simeq 10^{-3} \text{ rad}/\sqrt{\text{sec}}$ . The table shows that, for small enough values of  $\sigma$  and  $\Delta t$ , both models yield very similar error levels. Nevertheless, the error  $\delta_q^L$  is systematically bigger than the error  $\delta_q^I$ , for any value of the parameters  $\sigma$  and  $\Delta t$ . Moreover, the error in the Langevin model grows unbounded for  $\sigma \geq 10^{-2}$ , while the error in the Itô model remains bounded. Figure 1 depicts the time histories of the four components in the

**Table 1 Ratios of the modeling quaternion errors,  $\delta_q^I/\delta_q^L$ , between the Itô model and the Langevin model.**

$\Delta t [\text{sec}] \setminus \sigma [\frac{\text{rad}}{\sqrt{\text{sec}}}]$	$10^{-4}$	$5.10^{-4}$	$10^{-3}$	$5.10^{-3}$	$10^{-2}$	$10^{-1}$
$10^{-3}$	$\frac{.007}{.007}$	$\frac{.17}{.19}$	$\frac{.54}{.60}$	$\frac{1.7}{7.6}$	$\frac{1.9}{144}$	$\frac{2.1}{1077}$
$10^{-2}$	$\frac{.009}{.009}$	$\frac{.17}{.19}$	$\frac{.63}{.75}$	$\frac{1.8}{8.4}$	$\frac{1.9}{377}$	$\frac{2.2}{2310}$
$10^{-1}$	$\frac{.009}{.009}$	$\frac{.18}{.20}$	$\frac{.64}{.80}$	$\frac{1.9}{9.9}$	$\frac{2.1}{681}$	$\frac{2.2}{3181}$

quaternion modeling error and the time history of the norm modeling error, in both models, for  $\sigma\sqrt{\Delta t} = 10^{-2} \text{ rad}$ . The norm modeling error is computed as the difference between the norm of the modeled quaternion and unity. The red curves are related to the Langevin model and the blue curves depict the errors using the Itô model. The Itô model consistently shows a better accuracy than the Langevin model, which displays severe instabilities, according to the level of the noise intensity. The later fact is emphasized on the plot of the norm modeling error, where the Langevin model is unable to preserve the squared norm of the quaternion, and shows an exponential divergence.



**Figure 1 Modeling errors in the quaternion (Langevin-red curve and Itô-blue). Single run.  $\sigma\sqrt{\Delta t} = 10^{-2} \text{ rad}$ . The Langevin error grows exponentially.**

### ***Model validation: Second-order moment equation***

*Preservation of the trace and Statistical Consistency:* Simulations were performed in order to illustrate the properties of the propagation equations for the second-order moment matrices,  $X$ , of the modeled quaternion in both models. Particular attention was given to the trace of  $X$ . Ideally, the trace of  $X$  should equal one at all times. The error induced in the Itô model is developing much slower than in the Langevin model. The bigger the noise intensity, the higher are these errors. However, while both models have the same accuracy for  $\sigma = 10^{-7} \text{ rad}/\sqrt{\text{sec}}$ , the discrepancy in their accuracy grows by several order of magnitudes when  $\sigma$  increases. For instance, at  $\sigma = 10^{-4} \text{ rad}/\sqrt{\text{sec}}$ , the final error in the Langevin model is  $5 \cdot 10^6$  times larger than the error in the Itô model. This result is an illustration of the exponential divergence of the trace of  $X$  in the Langevin model, as given in Eq. (14). Indeed, at  $t_f = 6000 \text{ sec}$  that equation predicts for  $\text{tr}X$  a value of  $\exp(3/4\sigma^2 t_f) \simeq 5 \cdot 10^{-5}$ . Graphical results, which are not shown here for the sake of brevity, can be found in Ref. [11]. An additional set of simulations was performed in order to check the statistical consistency of the Itô model, that is, to check that the matrix  $X$  correctly predicts the second-order moment of the quaternion, as modeled via the Itô process equation. This was done for  $\sigma = 10^{-4} \text{ rad}/\sqrt{\text{sec}}$  and  $\Delta t = 0.1 \text{ sec}$ . A Monte-Carlo (MC) simulation (200 runs) was performed in order to compute the MC variances for each of the components of the quaternion, as well as for its squared norm. It comes out that that these differences remain at  $10^{-4}$  for the quaternion components, and at  $10^{-7}$  for the squared norm. The accuracy in the latter result stems from the corrective term in the Itô formulation of the quaternion model.

### ***Estimation performance: Comparison with the Additive Extended Kalman Filter***

The previous theoretical analysis showed that the proposed quaternion filter (SQKF), which includes the Ito stochastic correction and the augmented second moment computations, provides a quaternion estimate with a clear mathematical significance. The following numerical study will illustrate, via extensive Monte-Carlo simulations, under what conditions the SQKF can advantageously replace the standard Additive Extended Kalman Filter. For the sake of simplicity the quaternion measurement noise covariance,  $R_t$ , has been assumed of the following forms:

$$R_t = \sigma_q^2 I_4 \quad \text{and} \quad R_t = \sigma_q^2 (I_4 - \bar{\mathbf{q}}_t \bar{\mathbf{q}}_t^T) + \alpha I_4 \quad (28)$$

where  $\bar{\mathbf{q}}$  is the measured quaternion,  $\sigma_q$  is the variance parameter,  $I_4$  is the four dimensional identity matrix, and  $\alpha$  is a regularization parameter used in order to avoid singularity (typically of order  $10^{-15}$ ). The second expression in Eq. (28) is inspired from previous results on unit-norm vector measurements, which, usually, are 3-dimensional vector measurements.<sup>18</sup> Both expressions were used without noticing significant discrepancies in the performance. as given in the left equation of Eq. (28).

### ***Estimation performance: Attitude estimation***

Several MC simulations (200 runs) were performed while varying the signal to noise ratio (SNR), which is here defined as follows:  $SNR = \frac{\sigma \sqrt{\Delta t}}{\sigma_q}$ . For each value of the SNR, the SQKF and the AEKF were implemented and the time averages of the associated angular estimation errors were computed. Table 2 depicts the ratio of the SQKF angular error time average over the AEKF angular error time average. In parenthesis appears the order of magnitude of the angular error in SQKF. These results illustrate that, as expected, on the whole range of values for the SNR, the SQKF performance are either similar or superior to the AEKF performance. This directly stems from the optimality of the SQKF. The higher the SNR, the clearer the result becomes: for  $SNR = 10^{-5}$ , the performance increase may reach tens of percents. The configuration of very high SNR corresponds to very reliable sensors (quaternion units delivering arcsecond accuracy) jointly used with very poor gyros, e.g., MEMS-based gyros, which are known for their high-valued noises.<sup>22</sup>

In Figure 2 are plotted the time variations of the quaternion estimation error, for each component, on a single run (green solid lines), together with the  $\pm 1\sigma$  envelope, as computed in the SQKF (green dotted lines), and the MC averages of the estimation errors (blue solid lines). It appears that, without any need for tuning, the filter shows statistical consistency, i.e., the single run realization is bounded by the  $\pm 1\sigma$  envelope. Furthermore, the MC averages illustrate the unbiasedness of the SQKF. Additional Monte-Carlo simulations were performed showing that the MC estimation error variances were close to the filter computed variances.

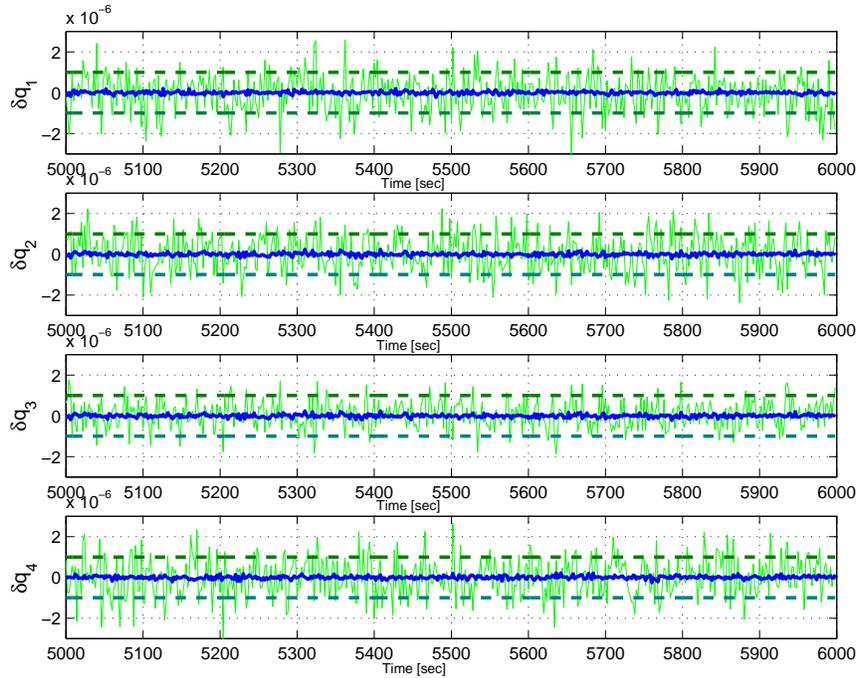
### ***Estimation performance: Attitude and gyro bias estimation***

In this subsection, the proposed approach for quaternion estimation is extended to the case where the gyros measurements are corrupted by additive random biases,  $\mu_t$ , modeled as random walks, as well as by additive white noise:

$$\omega_t = \omega_t^o + \mu_t + \epsilon_t \quad (29)$$

**Table 2** Ratio of angular error MC averages of SQKF over AEKF for various SNR  $\sigma\sqrt{\Delta t}/\sigma_q$ . The error in SQKF is systematically smaller than in AEKF and can increase the filter performance by tens of percents.

$\sigma_q$ [rad] \ $\sigma\sqrt{\Delta t}$ [rad]	$5.10^{-3}$	$1.10^{-2}$	$5.10^{-2}$	$7.10^{-2}$	$1.10^{-1}$
$10^{-7}$	$(10^{-5} \text{ deg})$ 0.999	$(10^{-5} \text{ deg})$ 0.994	$(10^{-5} \text{ deg})$ 0.101	$(10^{-5} \text{ deg})$ 0.066	$(10^{-6} \text{ deg})$ 0.014
$10^{-6}$	$(10^{-4} \text{ deg})$ 0.999	$(10^{-4} \text{ deg})$ 0.999	$(10^{-4} \text{ deg})$ 0.665	$(10^{-4} \text{ deg})$ 0.347	$(10^{-4} \text{ deg})$ 0.127
$10^{-5}$	$(10^{-3} \text{ deg})$ 1.0	$(10^{-3} \text{ deg})$ 0.999	$(10^{-3} \text{ deg})$ 0.995	$(10^{-3} \text{ deg})$ 0.866	$(10^{-3} \text{ deg})$ 0.730



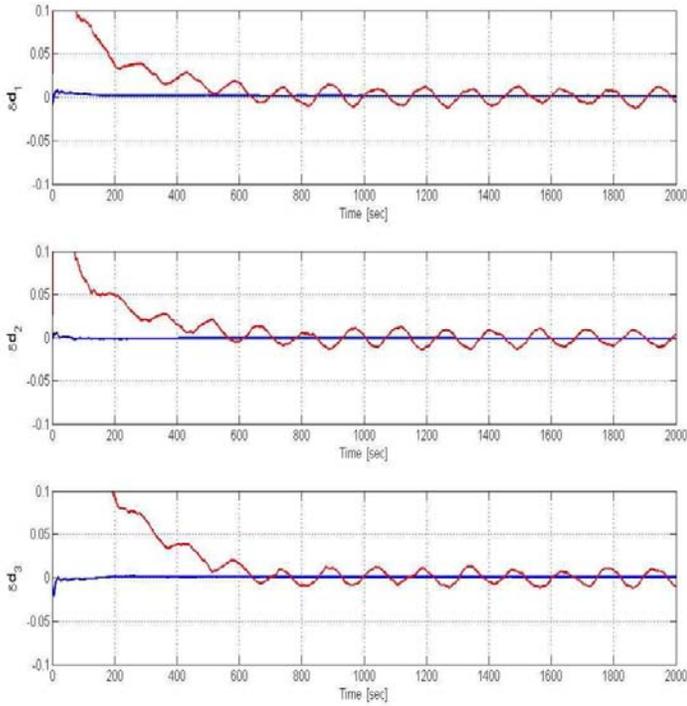
**Figure 2** Time variations of the estimation errors in the quaternion components for a typical run of the SQKF (solid green lines). The dotted green lines depict the  $\pm 1\sigma$  envelopes as computed by the SQKF. The solid blue lines are MC averages (200 runs). Without tuning, the SQKF computes estimates that are consistent with their predicted statistics. The estimation errors are unbiased.  $\sigma\sqrt{\Delta t}/\sigma_q = 10^5$ .

The standard approach consists in using the State Augmentation technique in order to apply the filtering theory to the augmented following system:

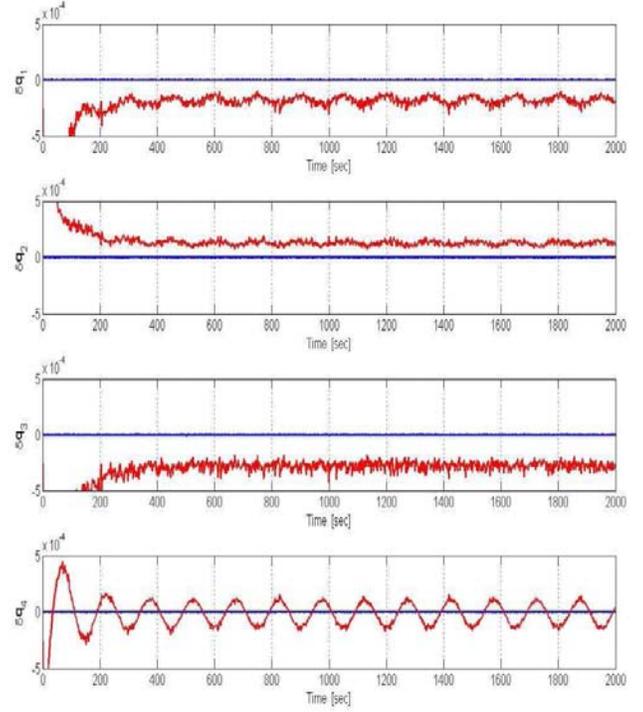
$$d\mathbf{q}_t = \frac{1}{2} \left[ \Omega(t) - M(t) - \frac{3\sigma^2}{4} I_4 \right] \mathbf{q}_t dt - \frac{1}{2} \Xi(\mathbf{q}_t) d\boldsymbol{\beta}_t; \quad \mathbf{q}_t(0) \stackrel{a.e.}{=} \mathbf{q}_0 \quad (30)$$

$$d\boldsymbol{\mu}_t = d\boldsymbol{\eta}_t, \quad \boldsymbol{\mu}(0) \stackrel{a.e.}{=} \boldsymbol{\mu}_0; \quad t \in [t_0, T] \quad (31)$$

where  $\boldsymbol{\eta}_t$  is a standard Brownian motion with intensity  $\sigma_\eta^2$ , and  $M(t)$  is defined by Eq. (3) with  $\boldsymbol{\mu}_t$  substituted to  $\boldsymbol{\omega}_t$ . The development of the augmented filter follows well-known steps (see e.g.[2]) and needs no further description.



(a) Biases estimation errors in rad/sec



(b) Quaternion estimation errors

**Figure 3 Attitude and bias estimation. MC-mean (200runs) of the gyro biases and of the quaternion estimation errors for  $\sigma\sqrt{\Delta t}/\sigma_q = 10^5$ . (SQKF in blue, AEKF in red). SQKF produces unbiased estimates in spite of the gyro-quaternion system's nonlinearities.**

Notice that the  $\hat{I}t\hat{o}$  corrective term remains identical since the multiplicative noise only depends on the quaternion. The resulting estimator is not the optimal unbiased linear filter any more since the augmented system is nonlinear with respect to the augmented state. Nevertheless it is still possible to use, with some approximation, the analysis related to the quaternion second-order moment propagation. Typical biases in MEMS rate gyros are of order  $5 \text{ deg/hr}$ .<sup>22</sup> This numerical study assumed more severe cases where  $\|\boldsymbol{\mu}(0)\| \simeq 200 \text{ deg/hr}$ . Figure 3 presents results from MC simulation (200 runs over 6000 seconds) where  $\sigma = 10^{-1} \text{ rad}/\sqrt{\text{sec}}$ ,  $\sigma_q = 10^{-6} \text{ rad}$ ,  $\sigma_\eta = 10^{-6} \text{ rad/sec}^{1.5}$ ,  $\boldsymbol{\mu}_0 = 10^{-3} \text{ rad/sec}$ , and  $\Delta t = 0.1 \text{ sec}$ . The quaternion measurement noise level is chosen very low on purpose. This choice, together with the fact that the measurement equation is linear, allows a comparison of the filters performances solely based on how they handle the process equation nonlinearities and noise.

Figure 3-a shows the MC averages of the biases estimation errors in the SQKF and in the AEKF. It should be noticed that the SQKF did not require any tuning. On the other hand, the AEKF performances were achieved after multiplying the filter intensity parameter  $\sigma_\eta$  by a factor 10 in order to avoid divergence of the drifts estimate and in order to increase the convergence rate. In addition, the diagonal elements of the initial filter covariance matrix,  $P_0$ , that are associated with the drift estimate were multiplied by a factor  $10^3$  with regards to unity nominal values. Both filters provide drifts estimates that are of the same order of magnitude as the true values. The AEKF estimates are, however, depicting SC dynamics induced oscillations, which is the consequence of an estimate-dependent filter covariance computation. In spite of the tuning, the AEKF still shows a slower transient than the SQKF transient, which is practically immediate. The advantage of the SQKF in steady state is clear when plotting the quaternion estimation errors. Figure 3-b shows the MC averages of these errors over time. The AEKF presents biased errors in the first three components and an undamped oscillatory error in the fourth component. On the other hand, the SQKF produces practically unbiased estimates ( $\simeq 10^{-9}$ ) in all four components. The latter is remarkable in view of the nonlinearities that are present in the bias-quaternion system's dynamics. This last result illustrates the benefit of properly computing the quaternion-dependent process noise covariance in the filter.

## CONCLUSION

A novel continuous-time Itô stochastic differential equation for the quaternion of rotation,  $\mathbf{q}_t$ , was developed assuming angular velocity measurements corrupted by zero-mean gaussian white noise and slowly varying drifts. The quaternion process equation presents dissipative terms in the dynamics matrix that ensures mean square stability. A best (minimum-variance) linear unbiased quaternion estimator was developed, where the filtering equations include the dissipative terms, in accordance with the Itô form of the  $\mathbf{q}_t$ -process. It was shown that neglecting these dissipative terms would result in an exponentially diverging bias in the quaternion estimate length at a rate that is proportional to the gyro noise variance parameter. Extensive Monte-Carlo simulation results illustrate the performance of the proposed quaternion modeling and filtering approaches in the case of a rotating spacecraft, when compared with a standard additive extended Kalman filter (AEKF). For high signal-to-noise ratios, the novel filter clearly outperforms the AEKF, while maintaining, without the need for tuning, statistical consistency in its estimates. Further work remains to be done in order to investigate the convergence properties of the Riccati second-order moment equations and to check whether quaternion normalization may improve the filter performance.

## REFERENCES

- [1] Wertz, J.R. (ed.), *Spacecraft Attitude Determination and Control*, D. Reidel, Dordrecht, The Netherlands, 1984.
- [2] Lefferts, E.J., Markley, F.L., and Shuster, M.D., "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 5, Sept.-Oct. 1982, pp. 417–429.
- [3] Bar-Itzhack, I.Y., and Oshman, Y., "Attitude Determination from Vector Observations: Quaternion Estimation," *IEEE Transactions on Aerospace and Electronic Systems*, Vol-AES-21, Jan. 1985, pp. 128–136.
- [4] Bar-Itzhack, I.Y., Deutschmann, J., and Markley, F.L., "Quaternion Normalization in Additive EKF for Spacecraft Attitude Determination," AIAA Paper 91-2706, *Guidance, Navigation, and Control Conference*, New Orleans, LA, August 1991.
- [5] Zanetti, R., Bishop, R.H., "Quaternion Estimation and Norm Constrained Kalman Filtering," AIAA Paper 06-6164, *Guidance, Navigation, and Control Conference*, Keystone, CO, August 2006.
- [6] Calise, A., "Enforcing an Algebraic Constraint in Extended Kalman Filter Design," AIAA Paper 07-6515, *Guidance, Navigation, and Control Conference*, Hilton Head, SC, August 2007.
- [7] Kalman, R.E., and Bucy, R.S., "New Results in Linear Filtering and Prediction Theory," *Trans. ASME, J. Basic Eng.*, Ser. D, Vol. 83, March 1961, pp. 95–108.
- [8] Jazwinski, A.H., *Stochastic Processes and Filtering Theory*, New York: Academic, 1970.
- [9] Wonham, W.M., Random Differential Equations in Control Theory, In A. Bharucha-Reid, editor, *Probabilistic Methods in Applied Mathematics*, Volume 2, Academic Press, 1970, pp. 192–194.
- [10] Wong, E. *Stochastic Processes in Information and Dynamical Systems*, McGraw-Hill Book Company, 1971.
- [11] Choukroun, D., "Novel Stochastic Modeling and Filtering of the Attitude Quaternion," Paper AAS 08-288, *F. Landis Markley Astronautics Symposium*, Cambridge MD, June 2008.
- [12] McLane, P.J., "Optimal Linear Filtering for Linear Systems with State-dependent Noise," *Int. J. Contr.*, Vol. 10, No. 1, 1969, pp. 45–51.
- [13] Gustafson, D., Speyer, J.L., "Linear Minimum Variance Filters Applied to Carrier Tracking," *IEEE Transactions on Automatic Control*, AC-21, No. 1, Feb. 1976, pp. 65–73.
- [14] Gershon, E., Shaked, U., Yaesh, I.,  *$H_\infty$  Control and Estimation of State-multiplicative Linear Systems*, LNCIS-Lecture Notes in Control and Information Sciences, Vol. 318, 2005, Springer.
- [15] Wong, E., and Zakai, M., "On the Convergence of the Solutions of Ordinary Integrals to Stochastic Integrals," *Ann. Math. Stat.*, 36, 1965, pp. 1560–1564.
- [16] Clark, J.M.C. *The Representation of Non-Linear Stochastic Systems with Applications to Filtering*, Ph.D. Thesis, Electrical Engineering Dept., Imperial College, London, England, 1965.
- [17] Bar-Itzhack, I.Y., Harman, R., Choukroun, D., "State-Dependent Pseudo-Linear Filters for Spacecraft Attitude and Rate Estimation," *Proceedings of the AIAA Guidance, Navigation and Control Conference*, AIAA-2002-4461, Monterey, California, Aug. 5-8, 2002.
- [18] Markley, F.L., "Attitude Determination Using Vector Observations and the Singular Value Decomposition," *Journal of the Astronautical Sciences*, Vol. 36, No. 3, July-Sept. 1988, pp. 245–258.
- [19] Athans, M., and Tse, E., "A Direct Derivation of the Optimal Linear Filter Using the Maximum Principle," *IEEE Transactions on Automatic Control*, Vol. AC-12, No. 6, 1967, pp. 690–698.
- [20] Kalman, R.E., Bertram, J.E., "Control System Analysis and Design via the Second-Method of Lyapunov-Part I: Continuous-time Systems," *Trans. ASME (J. Basic Eng.)*, Vol. 82, pp. 394–400, 1960.
- [21] Choukroun, D., Oshman, Y., and Bar-Itzhack, I.Y., "Novel Quaternion Kalman Filter," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AC-42, No. 1, Jan. 2006, pp. 174–190.
- [22] Durrant, D., Dussy, S., Shackleton, B., Malvern, A. "MEMS Rate Sensors in Space Becomes Reality," *Proceedings of the AIAA Guidance, Navigation and Control Conference*, AIAA-2007-6547, Hilton Head, South Carolina, Aug. 2007.