# Minimum-Energy Filtering for Attitude Estimation

Mohammad Zamani, Student Member, IEEE, Jochen Trumpf, Member, IEEE, and Robert Mahony, Senior Member, IEEE

Abstract—In this work we study minimum-energy filtering for attitude kinematics with vectorial measurements using Mortensen's approach. The exact form of a minimum-energy attitude observer is derived and is shown to depend on the Hessian of the value function of an associated optimal control problem. A suitably chosen matrix representation of the Hessian operator leads to a Riccati equation that approximates a minimumenergy attitude filter. An extended version of the proposed approximate filter is included for a situation where there is slowly time-varying bias in the gyro measurements. A unit quaternion version of the proposed filter is derived and shown to outperform the multiplicative extended Kalman filter (MEKF) for situations with large initialization errors or large measurement errors.

## I. INTRODUCTION

Consider the problem of deriving an optimal filter for the attitude of a rigid-body using vectorial measurements in 3D space. The natural state formulation for attitude kinematics of a rigid-body evolves on the special orthogonal Lie group SO(3) [1]. State of the art stochastic methods apply modifications to the extended Kalman filter (EKF) and the unscented Kalman filter (UKF) equations in order to preserve the group structure of the estimates. See for example the multiplicative extended Kalman filter (MEKF) [2], the unscented quaternion estimator (USQUE) [3] or the invariant extended Kalman filter (IEKF) [4]. In a different approach, Choukroun et al [5] use an embedded representation of SO(3), namely as unit quaternions in  $\mathbb{R}^4$ , to obtain an 'optimal' attitude filter in  $\mathbb{R}^4$ . However, the optimal filter estimate needs to be re-projected onto the rotation group to obtain feasible estimates.

In the late 1960s, Mortensen [6] has introduced a systematic approach to deriving filtering algorithms for deterministic nonlinear systems in Euclidean space. This approach, known as minimum energy

The authors are with the Research School of Engineering, Australian National University, Canberra, ACT, Australia. Mohammad.Zamani@anu.edu.au, Jochen.Trumpf@anu.edu.au and Robert.Mahony@anu.edu.au. filtering, was further explored by Hijab [7]. Convergence of minimum-energy filters was studied by Krener [8] who proved that under some conditions including the uniform observability of the system, a minimum-energy estimate converges exponentially fast to the true state. Aguiar et al. [9] applied minimum-energy filtering to systems with perspective outputs by embedding the nonlinear geometry in an overarching Euclidean space. The resulting estimates need to be projected back to the special Euclidean group SE(3). A simple quaternion normalization as the projection process could arguably yield a sub-optimal filter unless proven otherwise mathematically. Geometric nonlinear observers for attitude kinematics have been heavily studied in the past few years [10]–[14]. Moreover, different approaches to deterministic optimal attitude filtering can be found in [15], [16] that are based on uncertainty ellipsoids and non-integral cost functions, respectively. In recent work [17] a near-optimal deterministic filter was derived for a system defined on the unit circle  $S^1$ . The authors extended this work to attitude kinematics on SO(3) for the case of full attitude measurements [18]. The earlier work by the authors was based on ad-hoc methods and to the authors knowledge there is no prior work that uses the more structured Mortensen's approach to derive minimum energy filters on SO(3).

In this paper we derive a deterministic attitude filter using vectorial measurements by adapting Mortensen's approach [6] to the geometric structure of SO(3). This approach yields an elegant representation of the filter equation as a gradient flow of the measurement error cost associated with the optimal control Lagrangian, where the gradient is taken with respect to a Riemannian metric based on the Hessian of the value function. The Hessian of the value function that is required to implement the filter can itself be computed by repeated differentiation. Using a matrix representation of the Hessian we obtain a Riccati equation that approximates a minimumenergy attitude filter by neglecting the third order

derivative of the value function. We provide an extended version of the proposed filter that deals with bias in the gyro measurements. For numerical analysis, we provide a quaternion implementation of the proposed filter. We observe that the proposed filter has the same observer part as a unit quaternion attitude MEKF [2]. The proposed Riccati equation however augments the Riccati equation of the MEKF with additional terms that are associated with the geometric structure of the system that is not captured by the MEKF construction. By means of simulations we compare the proposed filter to the MEKF. The proposed filter achieves faster convergence and lower estimation error in the bias estimates and consequently in the quaternion estimates than the MEKF in a suit of simulations where there are large initialization or large measurement errors.

The remainder of the paper is organized as follows. Section II introduces our notation and some mathematical identities used later in the paper. In Section III we formally present the problem of minimum-energy attitude filtering for a system defined on SO(3) and characterize the optimal filter estimate as the minimum of a value function derived from an associated optimal control problem. Section IV contains the filter derivation using Mortensen's approach. We provide a quaternion version of the proposed filter along with a quaternion implementation of the attitude MEKF in Section V. We demonstrate simulation-based comparisons between the two filters. Lastly, Section VI concludes the paper.

## II. NOTATION

The rotation group is denoted by  $SO(3) = \{X \in \mathbb{R}^{3\times3} | X^{\top}X = I, \det(X) = 1\}$ , where *I* is the 3 by 3 identity matrix. The associated Lie algebra  $\mathfrak{so}(3) = \{A \in \mathbb{R}^{3\times3} | A = -A^{\top}\}$ , is the set of skew-symmetric matrices. For  $\Omega = [a, b, c]^{\top} \in \mathbb{R}^3$ , the lower index operator  $(.)_{\times} : \mathbb{R}^3 \longrightarrow \mathfrak{so}(3)$  yields the skew-symmetric matrix  $\Omega_{\times} = [\Omega_1 \ \Omega_2 \ \Omega_3]$  where  $\Omega_1 = [0 \ c \ -b]^{\top}$ ,  $\Omega_2 = [-c \ 0 \ a]^{\top}$  and  $\Omega_3 = [b \ -a \ 0]^{\top}$ . Inversely, the operator vex :  $\mathfrak{so}(3) \longrightarrow \mathbb{R}^3$  extracts the skew coordinates,  $\operatorname{vex}(\Omega_{\times}) = \Omega$ . The cost  $\|.\|_R : \mathbb{R}^{3\times3} \longrightarrow \mathbb{R}_0^+$  is given by

$$\|M\|_R \coloneqq \sqrt{\frac{1}{2}\operatorname{trace}(M^\top R M)}$$

where  $R \in \mathbb{R}^{3\times 3}$  is symmetric positive definite. Note that  $||M||_R$  coincides with the Frobenius norm of  $R^{1/2}M$ . The symmetric projector  $\mathbb{P}_s$  is defined by  $\mathbb{P}_s(M) \coloneqq 1/2(M+M^{\top})$ . The skew-symmetric projector

 $\mathbb{P}_a$  is defined by  $\mathbb{P}_a(M) \coloneqq 1/2(M - M^{\top})$ . It is easily verified that the vector product of the two vectors  $\gamma, \psi \in \mathbb{R}^3$  satisfies

$$(\gamma \times \psi) = \operatorname{vex}(2\mathbb{P}_a(\psi \gamma^{\top})) = 2\operatorname{vex}(\mathbb{P}_a(\gamma_{\times}\psi_{\times})).$$
(1)

Let  $L_X : SO(3) \longrightarrow SO(3)$ ,  $L_X S = XS$ , be the left translation and let  $TL_X : TSO(3) \longrightarrow TSO(3)$  denote the associated tangent map. Let  $\mathcal{D}_1 F(X, Y) \cdot TL_X \Gamma$  denote the derivative of the function F with respect to the first argument  $X \in SO(3)$  in the tangent direction  $TL_X \Gamma = X\Gamma \in T_X SO(3)$ , where  $\Gamma \in \mathfrak{so}(3)$ . Recall the relationship between a directional derivative  $\mathcal{D}$  and a gradient  $\nabla$  with respect to a Riemannian metric  $\langle ., . \rangle_X : T_X SO(3) \times T_X SO(3) \longrightarrow \mathbb{R}$ :

$$\mathcal{D}_1 F(X, Y) \cdot \mathrm{TL}_X \Gamma = \langle \nabla_1 F(X, Y), \mathrm{TL}_X \Gamma \rangle_X = \langle \mathrm{TL}_X^* \nabla_1 F(X, Y), \Gamma \rangle_I.$$
(2)

The asterisk denotes the adjoint with respect to the given Riemannian metric. We use the standard left-invariant Riemannian metric on SO(3). That is, for  $\Gamma, \Omega \in \mathfrak{so}(3)$  and  $X \in SO(3)$ 

$$\langle \mathrm{TL}_X \, \Gamma, \mathrm{TL}_X \, \Omega \rangle_X = \langle \Gamma, \Omega \rangle_I \coloneqq \frac{1}{2} \operatorname{trace}(\Gamma^\top \Omega).$$
 (3)

One has

$$\langle \operatorname{TL}_X \Gamma, \operatorname{TL}_X \Omega \rangle_X = \langle \operatorname{vex}(\Gamma), \operatorname{vex}(\Omega) \rangle = \operatorname{vex}(\Gamma)^\top \operatorname{vex}(\Omega).$$
 (4)

For the sake of simplicity, in the reminder of the paper we will omit the subscript notation from the Riemannian metrics.

## **III. PROBLEM FORMULATION**

In this section, we present the system governing the kinematics of the attitude of a rigid body and an associated measurement model that yields vectorial sensor measurements. We formulate the problem of minimum-energy filtering for this system.

The following equation is a model for the attitude kinematics of a rigid body.

$$\begin{cases} \dot{X}(t) = X(t)\Omega_{\times}(t), \ X(0) = X_{0}, \\ u(t) = \Omega(t) + Bv(t), \\ y_{i}(t) = X(t)^{\top}\dot{y}_{i} + D_{i}w_{i}(t), \ i = 1, \cdots, n , \end{cases}$$
(5)

where X is an SO(3)-valued state signal representing the attitude of a body-fixed frame, i.e. a frame attached to a moving rigid body, relative to a reference frame, i.e. a frame fixed at a reference point. The signal  $\Omega \in \mathbb{R}^3$  represents the angular velocity of the moving body expressed in the bodyfixed frame. The signals  $u \in \mathbb{R}^3$  and  $v \in \mathbb{R}^3$  denote the body-fixed frame measured angular velocity input and the input measurement error, respectively. The coefficient matrix  $B \in \mathbb{R}^{3\times3}$  allows for different weightings for the components of the unknown input measurement error v. We assume that B is

full rank and hence that  $Q := BB^{\top}$  is positive definite. The unit norm vectors  $\mathring{y}_i \in S^2 \subset \mathbb{R}^3$ are known vector directions in the reference frame. The measurements  $y_i \in \mathbb{R}^3$  are measurements of the  $\hat{y}_i$  in the body-fixed frame and the signals  $w_i \in \mathbb{R}^3$  are the unknown output measurement errors. The coefficient matrix  $D_i \in \mathbb{R}^{3 \times 3}$  allows for different weightings of the components of the output measurement error  $w_i$ . Again we assume that  $D_i$  is full rank and hence  $R_i \coloneqq D_i D_i^{\top}$  is positive definite.

Consider the cost

$$J(t; X_0, v|_{[0, t]}, \{w_i|_{[0, t]}\}) = \frac{1}{2} \int_0^t \left(v^\top v + \sum_i w_i^\top w_i\right) d\tau + \frac{1}{2} \operatorname{trace}\left[(X_0 - \chi)^\top K_0^{-1} (X_0 - \chi)\right],$$
(6)

in which  $K_0 \in \mathbb{R}^{3 \times 3}$  is symmetric positive definite and  $\chi \in SO(3)$  is an initial guess for the unknown initial state  $X_0$ . Note that  $\chi = I$  in case no reliable candidate for the value of  $\chi$  is available. The cost (6) can be thought of as a measure of the aggregate energy stored in the unknown signals of system (5).

The principle of minimum-energy filtering is as follows. At each time t, given the measurements  $\{y_i|_{[0, t]}\}$  and  $u|_{[0, t]}$ , the goal is to obtain an estimate X(t) of the true state X(t) by minimizing the cost (6). In order to obtain X(t), one seeks a combination of the unknowns  $(X_0, v|_{[0, t]}, \{w_i|_{[0, t]}\})$ that is compatible with the measurements  $\{y_i|_{[0, t]}\}$ and  $u|_{[0, t]}$  in fulfilling the system equations (5). Note that in general, infinitely many combinations of these unknowns are compatible with the measurements. By minimizing the cost (6) a triplet  $(X_0^*, v^*|_{[0, t]}, \{w_i^*|_{[0, t]}\})$  is chosen that contains minimum collective energy.

The minimizing unknowns  $(X_0^*, v^*|_{[0, t]},$  $\{w_i^*|_{[0, t]}\}$ ) replaced in the system equations (5) yield the optimal state trajectory  $X_{[0, t]}^*$ . The subscript [0, t] indicates that the optimization takes place on the interval [0, t]. We pick the final optimal state  $X_{[0, t]}^{*}(t)$  as our minimum-energy estimate at time t,  $\hat{X}(t) \coloneqq X^*_{[0, t]}(t)$ .

A naive approach to the minimum energy filtering problem leads to an infinite dimensional optimization problem for each time interval [0, t]. To obtain a practical algorithm we will seek to derive a recursive filter that at each time t yields the minimumenergy estimate as its state value.

Note that the cost (6) depends on the unknowns  $X_0, v|_{[0, t]}$  and  $\{w_i|_{[0, t]}\}$ , but given  $X_0$  and  $v|_{[0, t]}$ , the known  $\{\hat{y}_i\}$ , and the measurements  $u|_{[0, t]}$  and  $\{y_i|_{[0, t]}\}$ , the  $w_i|_{[0, t]}$  are uniquely determined by (5). Hence, the cost (6) is equivalent to

$$J(t; X_0, v|_{[0, t]}) = \frac{1}{2} \operatorname{trace} \left[ (X_0 - \chi)^\top K_0^{-1} (X_0 - \chi) \right] + \int_0^t \frac{1}{2} \left( v^\top v + \sum_i (X^\top \mathring{y}_i - y_i)^\top R_i^{-1} (X^\top \mathring{y}_i - y_i) \right) d\tau$$
(7)

where the cost (7) depends only on the signals  $X_0$ and  $v|_{[0, t]}$ . Minimizing (7) over these two arguments is simplified by first assuming that  $X_0$  is known and minimizing over  $v|_{[0, t]}$ , then later optimizing over  $X_0$ . The problem of minimizing (7) subject to (5) can be seen as an optimal control problem where the signal  $v|_{[0, t]}$  is considered as a control input.

As in the maximum-principle [19] the pre-Hamiltonian for the above optimal control problem İS

$$\mathcal{H}^{-}(X, \mu_{\times}, v, t) \coloneqq \frac{1}{2} [v^{\top} v + \sum_{i} (X^{\top} \mathring{y}_{i} - y_{i})^{\top} R_{i}^{-1} (X^{\top} \mathring{y}_{i} - y_{i})] - \mu^{\top} (u - Bv),$$
(8)

where  $\mu \in \mathbb{R}^3$  represents a costate variable  $\Theta \in$  $\mathfrak{so}^*(3)$  via  $\langle \mu_{\times}, \Gamma \rangle = \Theta(\Gamma)$  for all  $\Gamma \in \mathfrak{so}(3)$ . In the following the identification of  $\Theta \in \mathfrak{so}^*(3)$  with  $\mu_{\times} \in \mathfrak{so}(3)$  will be used without further reference. Since the pre-Hamiltonian (8) is quadratic in v its minimum is given by the differential condition

$$\mathcal{D}_{v}\mathcal{H}^{-}\cdot\gamma=0,\;\forall\gamma\in\mathbb{R}^{3}.$$
(9)

Solving for v yields the optimal  $v^* = -B^{\top}\mu$ . Substituting  $v^*$  in (8) yields the optimal Hamiltonian

$$\mathcal{H}(X,\mu_{\times},t) = \frac{1}{2} [-\mu^{\top} Q \mu + \sum_{i} (X^{\top} \mathring{y}_{i} - y_{i})^{\top} R_{i}^{-1} (X^{\top} \mathring{y}_{i} - y_{i})] - \mu^{\top} u.$$
(10)

In order to apply the dynamic programming principle [19] to this problem the following value function is defined

$$V(X,t) \coloneqq \min_{v|_{[0,t]}} J(t; X_0, v|_{[0,t]}), \qquad (11)$$

where J is the cost (7) and the minimization is constrained by the system equations (5). This is well defined because from (5),  $X_0$  and  $v|_{[0,t]}$  uniquely determine X(t) and vice versa X(t) and  $v|_{[0,t]}$ uniquely determine X(0). The Hamilton-Jacobi-Bellman equation is then [19]

$$\mathcal{H}(X, \mathrm{TL}_X^* \, \nabla_1 V(X, t), t) - \frac{\partial V}{\partial t}(X, t) = 0.$$
(12)

From (7) the initial time boundary condition is

$$V(X_0, 0) = \frac{1}{2} \operatorname{trace} \left[ (X_0 - \chi)^\top K_0^{-1} (X_0 - \chi) \right].$$
(13)

Preprint submitted to IEEE Transactions on Automatic Control. Received: February 21, 2013 21:46:04 PST

I

Up to here we have used the dynamic programming principle to address the optimal control part of the problem (by minimizing (7) over v). To complete the optimal filtering problem, we also need to optimize V over  $X_0$ . This can be equivalently posed as optimization over the final condition X(t) since the initial and final conditions are deterministically coupled by the optimal input  $v|_{[0, t]}$ . Assuming that the value function is strictly convex, its minimum is characterized by the condition

$$\nabla_1 V(X,t)|_{X=\hat{X}(t)} = 0.$$
 (14)

This equation is also referred to as the 'final condition' since it is evaluated at time t. Recall that the minimum-energy estimate  $\hat{X}(t)$  is defined as the final value of the minimizing argument  $X_{[0,t]}^*(t)$ . Solving Equation (14) characterizes  $\hat{X}(t)$ . However, this still requires an explicit solution to a potentially infinite dimensional nonlinear optimization problem and must be repeated at every time t. To overcome this issue we will use Mortensen's approach [6] to derive a recursive solution to this problem.

#### **IV. FILTER DERIVATION**

In this section we apply Mortensen's approach [6] to the minimum-energy filtering problem presented in Section III. Note that the final condition (14) is equivalent to the following equation. For all  $\Gamma \in \mathfrak{so}(3)$ 

$$\{\mathcal{D}_1 V(X,t) \cdot X\Gamma\}_{X=\hat{X}(t)} = 0.$$
(15)

Equation (15) characterizes the solution  $\hat{X}(t)$ . Next we compute the total time derivative of (15) in order to derive a dynamic filter updating the solution [6].

$$\frac{d}{dt} \{ \mathcal{D}_1 V(X, t) \cdot X\Gamma \}_{X = \hat{X}(t)} = 0.$$
(16)

Applying the chain rule to Equation (16) yields

$$\{\mathcal{D}_1^2 V(X,t) \cdot (X\Gamma, \dot{\hat{X}}(t)) + \mathcal{D}_1 \frac{\partial V}{\partial t}(X,t) \cdot X\Gamma\}_{X=\hat{X}(t)} = 0.$$
(17)

The second order derivative of the value function is related to the Hessian of the value function as an operator acting on a tangent direction. In order to obtain a matrix formulation we represent this by a matrix  $K \in \mathbb{R}^{3\times 3}$  such that for all  $\Gamma, \Psi \in \mathfrak{so}(3)$ , with vector representations  $\gamma := \operatorname{vex}(\Gamma)$  and  $\psi := \operatorname{vex}(\Psi)$ ,

$$\mathcal{D}_1^2 V(X,t) \cdot (X\Gamma, X\Psi) = \langle K\gamma, \psi \rangle = \langle \gamma, K\psi \rangle.$$

(18) The second term in (17) is given by replacing the partial time derivative in (17) from (12) and calculating the derivative with respect to X.

$$\mathcal{D}_1 \frac{\partial V}{\partial t} (X, t) \cdot X\Gamma = -\langle \sum_i ((X^\top \mathring{y}_i) \times (R_i^{-1} (X^\top \mathring{y}_i - y_i))_{\times}, \Gamma \rangle.$$
(19)

Denote

$$l := \sum_{i} (R_i^{-1}(\hat{y}_i - y_i)) \times \hat{y}_i$$
 (20)

where  $\hat{y} \coloneqq \hat{X}^{\top} \hat{y}$ . Equation (17) yields

$$\dot{\hat{X}} = \hat{X}(u - Pl)_{\times}, \tag{21}$$

where  $P := K^{-1}$  and  $\hat{X}(0) = \chi$  is obtained by evaluating the final condition (14) at time 0 using the boundary condition (13).

Note that Equation (21) is the *exact* form of a minimum-energy (optimal) observer for system (5) where energy is measured by the cost (6). The observer contains the innovation term l which is a weighted sum proportional to the information contained in the error  $\hat{y}_i - y_i$ . The matrix P acts as the gain for this innovation term.

To fully solve (21), the matrix P needs to be updated on-line. We follow Mortensen's approach [6] to compute the total time derivative  $\frac{d}{dt}\langle P\gamma,\psi\rangle$ . A rather tedious and lengthy calculation yields the following Riccati differential equation that dynamically updates P. Note that in order to obtain this equation one needs to neglect the third order derivatives of the value function with respect to X. In other words we assume that the value function (11) is quadratic in the state X. This provides a second order approximation of the infinite dimensional minimum-energy filter.

$$\dot{P} = Q + 2\mathbb{P}_s(P(u - Pl)_{\times}) - PSP + PEP, \qquad (22)$$

where

$$S \coloneqq \sum_{i} (\hat{y}_{i})_{\times}^{\top} R_{i}^{-1} (\hat{y}_{i})_{\times}, \ E \coloneqq \operatorname{trace}(C)I - C,$$
  
$$C \coloneqq \sum_{i} \mathbb{P}_{s} (R_{i}^{-1} (\hat{y} - y_{i})\hat{y}_{i}^{\top}).$$
  
(23)

The initial condition  $P(0) = (\operatorname{trace}(K_0^{-1})I - K_0^{-1})^{-1}$  is given by calculating the second order derivative of the value function (11) and evaluating using the boundary condition (13) at time 0. Note that  $K_0^{-1}$  was given in the cost (7).

The Riccati equation (22) along with the minimum-energy observer (21) approximates a minimum-energy filter for system (5) up to the second order. The approximation is due to neglecting the third order derivative of the value function (11). One could continue to apply Mortensen's approach to obtain higher order approximations of a minimum-energy filter but that would require tensor

algebra. In a recent paper [17] the authors provide expressions for the third order terms for the case of rotations in a plane  $SO(2) \equiv S^1$  where the Hessian and all higher order derivatives of the value function are scalars.

Note that one can also consider bias in the angular velocity measurements (5). This yields the following modified angular velocity measurement model.

$$u(t) = \Omega(t) + B_{\Omega}v_{\Omega}(t) + b(t), \qquad (24)$$

where  $b(t) \in \mathbb{R}^3$  is an unknown slowly timevarying bias signal generated from

$$b(t) = B_b v_b(t), \quad b(0) = b_0,$$
 (25)

where  $B_b \in \mathbb{R}^{3\times3}$  is a weighting matrix known from the model,  $v_b \in \mathbb{R}^3$  is a small unknown perturbation and  $b_0 \in \mathbb{R}^3$  is an unknown initial bias. Using this model and a similar derivation as we used in this section, but on the state manifold  $SO(3) \times \mathbb{R}^3$  using the natural metric, one can solve the filtering problem in Section III by modifying the cost function (6) to include the initial bias  $b_0$  and the model perturbation  $v_b$ . This yields the new filter

$$\hat{X} = \hat{X}(u - \hat{b} - P_a l)_{\times}, \quad X(0) = \chi,$$
 (26a)

where  $\hat{b}$  is the estimate of the bias b given from

$$\hat{b} = -P_c^{\top} l, \quad \hat{b}(0) = \zeta, \tag{26b}$$

and where l is the same innovation term as (20) and  $\zeta \in \mathbb{R}^3$  is a guess for the initial condition  $b_0$ that can be chosen as a zero vector in case there is no reliable guess available. The gains  $P_a$  and  $P_c$  are updated from the following equations.

$$P_{a} = Q_{\Omega} + 2\mathbb{P}_{s}(P_{a}(2(u-b) - P_{a}l)_{\times}) + P_{a}(E-S)P_{a} - P_{c}^{'} - P_$$

where  $Q_{\Omega} \coloneqq BB^{\top}$  and  $Q_b \coloneqq B_b B_b^{\top}$ .

### V. SIMULATIONS

In this Section we provide a formulation of the proposed filter (26) using unit quaternions. Using this formulation we compare the performance of our filter against the quaternion-based multiplicative extended Kalman filter (MEKF) [2].

A unit quaternion belongs to the set  $\mathbb{Q} = \{q = (s, v) \in \mathbb{R} \times \mathbb{R}^3 : |q| = 1\}$ . The set  $\mathbb{Q}$  is a group under the operation  $q_1 \otimes q_2 = (s_1s_2 - v_1^{\top}v_2, s_1v_2 + s_2v_1 + v_1 \times v_2)$ , with identity element 1 = (1, 0, 0, 0) and inverse  $q^{-1} = (s, -v)$ . The unit quaternion version of the attitude observer (26a) is given by

$$\dot{\hat{q}}(t) = \frac{1}{2}\hat{q} \otimes \mathbf{p}(u - \hat{b} - P_a l) , \qquad (27)$$

where  $\hat{q}(0) = 1$  and  $\mathbf{p}(\gamma) \coloneqq (0, \gamma)$  for  $\gamma \in \mathbb{R}^3$ . The signal  $\hat{b}$  is generated from (26b) and the gains are given from (26c). The MEKF [2] which in this context is a special case of the IEKF [4] is given by the same observer equations (26a) and (26b). However, in the Riccati equations of the MEKF, the term  $u - \hat{b}l$  is used rather than  $2(u - \hat{b}) - P_a l$ , and the term E is not present. The following simulation study shows that the extra terms in the proposed Riccati equations (26c) facilitate faster convergence and lower root mean square (RMS) estimation error of the proposed filter, specially when the initialization and measurement errors are large, as is the case in attitude filtering for low cost unmanned aerial vehicles (UAVs).

The two filters are simulated using zero initial bias estimates and identity unit quaternion initial quaternion estimates. Also, the identity matrix is used as the initial gain matrix of the two filters. A sinusoidal input  $\Omega = [0.2 \sin(\frac{\pi}{6}t) - \cos(\frac{\pi}{6}t) 2 \cos(\frac{\pi}{6}t)]$ drives the true trajectory X. The input measurement errors v and  $v_b$  are Gaussian zero mean random processes with unit standard deviation. The coefficient matrix B is chosen so that the signal Bv has a standard deviation of 25 degrees per 'second'. The bias variation is adjusted by  $B_b$  such that  $B_b v_b$  has a standard deviation of 0.0004 radians per 'second' squared. The system is initialized with a unit quaternion representing a rotation through 120 degrees and an initial bias of 20 degrees per 'second'. We assume that two orthogonal unit reference vectors are available. We also consider Gaussian zero mean measurement noise signals  $w_i$  with unit standard deviations. The coefficient matrices  $D_i$  are chosen so that the signals  $D_i w_i$  have standard deviations of 30 degrees. Although the two filters do not have access to the noise signals  $v_{\Omega}$ ,  $v_b$  and  $w_i$  themselves, they have access to the matrices  $Q_{\Omega} = BB^{\top}$ ,  $Q_b = B_b B_b^{\top}$  and  $R_i = D_i D_i^{\top}$ . We have tested the two filters in a suit of simulations involving different levels of initialization errors, measurement errors and different reference vectors. Figures 1 show a situation that is typical for our simulations where there is large initialization or large measurement

errors, as in the case of using low cost MEMS gyros such as the popular InvenSense MPU-3000 family. We have performed a Monte-Carlo simulation and the RMS of the estimation errors of the two filters are demonstrated for 100 repeats. Figure 1 indicates that the RMS of the rotation angle estimation error of the proposed filter converges faster towards zero and remains lower than the error of the MEKF. This is due to the fact that the proposed filter performs better in estimating the bias (Figure 1).



Fig. 1. The rotation angle and bias estimation performance of the proposed filter compared against the MEKF.

#### VI. CONCLUSIONS

A new attitude filter was proposed that is derived using Mortensen's approach to minimum-energy filtering. The filter is posed directly on the special orthogonal group SO(3) and uses vectorial measurements. The proposed method yields the exact form of a minimum-energy attitude observer which is shown to depend on the Hessian of the value function of an associated optimal control problem. An extended version of the filter is proposed that also deals with gyro bias. The proposed algorithm outperforms the industry standard attitude filter, the MEKF, in simulations for the case of UAVs where there are large initialization or large measurement errors.

## REFERENCES

- [1] Abraham, R. and Marsden, J.E. and Ratiu, T.S., *Manifolds, tensor analysis, and applications.* Springer, 1988.
- [2] F. Markley, "Attitude error representations for Kalman filtering," *Journal of guidance, control, and dynamics*, vol. 26, no. 2, pp. 311–317, 2003.
- [3] J. Crassidis and F. Markley, "Unscented filtering for spacecraft attitude estimation," *Journal of Guidance Control and Dynamics*, vol. 26, no. 4, pp. 536–542, 2003.
- [4] Bonnabel, S. and Martin, P. and Salaun, E., "Invariant Extended Kalman Filter: theory and application to a velocity-aided attitude estimation problem," in *Proceedings of the 48th IEEE Conference on Decision and Control*, 2009, pp. 1297–1304.
- [5] D. Choukroun, I. Bar-Itzhack, and Y. Oshman, "A novel quaternion Kalman filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 174–190, 2006.
- [6] R. E. Mortensen, "Maximum-likelihood recursive nonlinear filtering," *Journal of Optimization Theory and Applications*, vol. 2, no. 6, pp. 386–394, 1968.
- [7] O. B. Hijab, "Minimum energy estimation," Ph.D. dissertation, University of California, Berkeley, 1980.
- [8] A. Krener, "The convergence of the minimum energy estimator," in *Lecture Notes in Control and Information Sciences*, W. Kang, Ed. Springer, 2004, pp. 187–208.
- [9] A. Aguiar and J. Hespanha, "Minimum-energy state estimation for systems with perspective outputs," *IEEE Transactions on Automatic Control*, vol. 51, no. 2, pp. 226–241, 2006.
- [10] J. Thienel and R. M. Sanner, "A coupled nonlinear spacecraft attitude controller and observer with an unknown constant gyro bias and gyro noise," *IEEE Transactions on Automatic Control*, vol. 48, no. 11, pp. 2011–2015, 2003.
- [11] J. Vasconcelos, R. Cunha, C. Silvestre, and P. Oliveira, "Landmark based nonlinear observer for rigid body attitude and position estimation," in *Proceedings of the 46th IEEE Conference* on Decision and Control, 2007, pp. 1033–1038.
- [12] R. Mahony, T. Hamel, and J. -M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1203– 1218, 2008.
- [13] S. Bonnabel, P. Martin, and P. Rouchon, "Non-linear symmetrypreserving observers on Lie groups," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1709–1713, 2009.
- [14] C. Lageman, J. Trumpf, and R. Mahony, "Gradient-like observers for invariant dynamics on a Lie group," *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 367–377, 2010.
- [15] A. Sanyal, T. Lee, M. Leok, and N. McClamroch, "Global optimal attitude estimation using uncertainty ellipsoids," *Systems & Control Letters*, vol. 57, no. 3, pp. 236–245, 2008.
- [16] J. Crassidis, R. Alonso, and J. Junkins, "Optimal attitude and position determination from line-of-sight measurements." *Journal of Astronautical Sciences*, vol. 48, no. 2, pp. 391–408, 2000.
- [17] M. Zamani, J. Trumpf, and R. Mahony, "Minimum-energy filtering on the unit circle," in *Proceedings of the Australian Control Conference (AUCC)*, 2011, pp. 236–241.
- [18] M. Zamani, J. Trumpf, and R. Mahony, "Near-Optimal deterministic filtering on the rotation group," *IEEE Transactions on Automatic Control*, vol. 56, no. 6, pp. 1411–1414, 2011.
- [19] M. Athans and P. Falb, Optimal control: an introduction to the theory and its applications. McGraw-Hill, 1966.