Near-Optimal Deterministic Attitude Filtering

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

***Abstract*— A deterministic attitude filter is derived based on the principles of minimum-energy filtering. The proposed filter is applied to the attitude kinematics of a rigid body in 3D space and is posed directly on the rotation group $\mathrm{SO}(3)$. The proposed filter generalizes recently published work by Coote et al. on deterministic filtering on the unit circle. The filter is near-optimal in the sense that it achieves a cost that is *close* to the cost incurred by a minimum-energy filter. We provide an explicit bound on the difference in cost occurred by the proposed filter vs. an optimal filter and show that this bound is small by means of simulations. We compare the performance of the proposed filter with a quaternion implementation of an Extended Kalman Filter (EKF). While achieving comparable results to an EKF, the proposed filter shows more robustness against a range of deterministic disturbance levels and initialization errors.**

I. INTRODUCTION

Optimal filtering in the sense of state reconstruction using optimization methods is a subject of active research. Driven by the needs of applications, the particular problem of attitude filtering has attracted much attention. Attitude filtering has many applications including rigid body motion planning in robotics and unmanned aerial vehicle (UAV) navigation in control.

According to a recent survey [1], a majority of attitude estimation applications are tackled using methods based on probabilistic modeling principles. Optimal probabilistic filtering for linear systems leads to the well known Kalman filter [2]. However, for nonlinear systems, like the attitude kinematics of a rigid body in 3D space, intricacies such as the lack of finite dimensional parameterizations of general stochastic processes have made it in general impossible to find a finite dimensional optimal probabilistic filter [3]. The most common approach to develop practical filters for nonlinear systems is to work with a linearization of the original system, e.g. the Extended Kalman Filter (EKF) [4]. Other methods such as particle filters [5] or the Unscented Kalman Filter (UKF) [6] approximate the infinite dimensional distributions by working with a finite sampled set.

Another common approach to attitude estimation is based on deterministic modeling principles (e.g. [7]–[10]), and the resulting estimators are commonly known as nonlinear attitude observers. The development of these observers usually focuses on global stability characteristics and they achieve guaranteed convergence from almost any initial condition [1, §X].

Minimum energy filtering [11] is a deterministic filtering method whose development is based on an optimization problem. This method leads to an optimal filter with the same filter formulas as in the Kalman filter when it is applied to a linear system [12]. In recent work, Aguiar et al. [13] applied this method to the pose (attitude and position) estimation problem and showed optimality for a system posed on the special Euclidian group $\mathrm{SE}(3)$ using a representation embedded in a vector space. This is similar to recent work by Choukroun et al. [14] that used an embedded representation of $\mathrm{SO}(3)$ (via the quaternions in $\mathbb{R}^4$) to obtain a filter based on probabilistic principles. In both cases constraining the filter solution to preserve the embedding constraints compromises the optimality of the approach leading to suboptimal estimates. In a recent paper, Coote et al. [15] applied the minimum energy filtering method directly to a nonlinear system on S^1. The outcome was a near-optimal filter with an explicit bound on the filter's distance from optimality.

In this paper, we generalize the work by Coote et al. [15] and apply minimum energy filtering to attitude kinematics on the rotation group $\mathrm{SO}(3)$. We derive the filter directly on $\mathrm{SO}(3)$ fully respecting the nonlinear geometry of the system. We show that once the rotation is confined to be around one fixed axis the proposed filter specializes to the filter on S^1 in Coote et al. [15]. The proposed filter is near-optimal, that is, it achieves a cost whose difference from the optimal cost is bounded by an explicitly given expression. We show that this bound is small under normal operating conditions and provide simulation results to support this statement. In comparison to an EKF, the proposed filter appears advantageous due to its robust performance, in a large range of experiments with different levels of deterministic disturbance and initialization errors, while not requiring any major calibration effort. The proposed method assumes that direct rotation measurements are available, similar to the early work in nonlinear observers e.g. [7]–[9]. See also the references in the survey by Crassidis et al. on attitude estimation methods [1, §X]. Although there is no sensor that can directly measure the 3D rotation, in practical situations our filter can be proceeded by a well known attitude reconstruction process, see e.g. [16]–[18] and the references therein.

The paper is structured as follows. Section II introduces some notation. The problem formulation is provided in Section III. Section IV describes the proposed filter, verifies its near-optimality and briefly shows how it specializes to the filter on S^1 due to Coote et al. A suite of simulation results is provided in Section V, comparing the performance of the proposed filter with a quaternion implementation of an Extended Kalman Filter (EKF) and checking the assumptions

The authors are with the 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.

we make. A short conclusion completes the paper. An extended version of this paper is currently under review for the IEEE Transactions on Automatic Control.

II. Notation

The rotation group is denoted by $\mathrm{SO}(3)$. The associated Lie algebra $\mathfrak{so}(3)$ is the set of skew-symmetric matrices

$$\mathfrak{so}(3) = \{A \in \mathbb{R}^{3\times 3} | A = -A^T\}.$$

The Frobenius norm of a matrix $X \in \mathbb{R}^{3\times 3}$ is given by $\|X\| := \sqrt{\mathrm{trace}(X^T X)}$. We define a cost $\phi_\Gamma : \mathrm{SO}(3) \longrightarrow \mathbb{R}^+$, by

$$\phi_\Gamma(R) := \mathrm{trace}\left[(R-I)^T \Gamma (R-I)\right],$$

where $\Gamma \in \mathbb{R}^{3\times 3}$ is symmetric positive definite. Note that $\phi_\Gamma(R)$ coincides with the squared Frobenius norm of $\Gamma^{\frac{1}{2}}(R-I)$ and hence is positive. The projection operator $\mathbb{P} : \mathbb{R}^{3\times 3} \longrightarrow \mathfrak{so}(3)$ is defined by $\mathbb{P}(M) := \dfrac{1}{2}(M - M^T)$.

III. Problem Formulation

Consider a system on $\mathrm{SO}(3)$

$$\begin{cases} \dot{R} &= R(A + g\delta), \\ Y &= R\epsilon \end{cases} \tag{1}$$

where R, Y and ϵ are $\mathrm{SO}(3)$ valued state, output and measurement disturbance signals, respectively. The signals A and δ denote the measured angular velocity and the process disturbance, respectively, and take values in $\mathfrak{so}(3)$. The scalar g acts as a scaling for the disturbance signal δ. Although scaling appears to be unnecessary given the arbitrary nature of δ, it will turn out that g provides an important degree of freedom in determining the main results of this paper. We assume all signals to be deterministic functions of time. Additionally, we assume enough regularity so that unique maximal solutions exist. Note that these will automatically exist for all time since $\mathrm{SO}(3)$ is compact. The following cost function is considered.

$$\begin{aligned} J_T &= \frac{1}{4}\phi_{K_0^{-1}}(R(0)) + \int_0^T \left(\frac{1}{2}\|\delta\|^2 + \frac{1}{4}\phi_I(\epsilon)\right) d\tau \\ &= \frac{1}{2}\int_0^T \left(\mathrm{trace}[\delta^T\delta] + \mathrm{trace}[I-\epsilon]\right) d\tau \\ &\quad + \frac{1}{4}\mathrm{trace}\left[(R(0)-I)^T K_0^{-1}(R(0)-I)\right], \end{aligned} \tag{2}$$

in which $K_0 \in \mathbb{R}^{3\times 3}$ is symmetric positive definite. The cost J_T is a nonnegative measure of the uncertainties of system (1) during the estimation period $[0, T]$. The cost function (2) measures the magnitude of the deterministic disturbance associated with a given trajectory of the system. By minimizing over all disturbance that are compatible with the model and the observed output the resulting system trajectory yields the minimum energy estimate. The problem considered is to find a recursive filter that estimates the minimum-energy state denoted by $\hat{R}(T)$, provided that all the past noisy state measurements $Y|_{[0,T]}$ and all the past applied inputs $A|_{[0,T]}$ are available to the filter.

IV. Main results

In this section we will discuss our solution to the problem sketched in the previous section. First we define a filter and then we show that this filter is a near-optimal solution to our filtering problem. Consider the filter

$$\dot{\hat{R}} = \hat{R}\left(A - \mathbb{P}\left(KY^T\hat{R}\right)\right), \tag{3a}$$

$$\dot{K} = \frac{1}{2}Q - \frac{1}{2}K(Y^T\hat{R} + \hat{R}^T Y)K + KA - AK, \tag{3b}$$

where $\hat{R}(0) := I_{3\times 3}$, $K(0) := K_0$, and $Q \in \mathbb{R}^{3\times 3}$ is symmetric positive definite. A and Y are defined by system (1). The filter in Equation (3) consists of two interconnected parts. Equation (3a) evolves on $\mathrm{SO}(3)$ and consists of a copy of system (1) plus an innovation term. The innovation term is a weighted distance between the (past) estimated signal and the noisy measured state signal projected on the tangent space. Note that $Y^T\hat{R}$ encodes the distance between Y and $\hat{R}$ on the group $\mathrm{SO}(3)$. That is, starting from Y and following this distance we will reach $\hat{R} = Y(Y^T\hat{R})$. This distance can be interpreted in terms of a rotation angle between $\hat{R}$ and Y. The weighting matrix K, dynamically generated by (3b), depends on estimates and measurements from the past. Equation (3b) is a time-varying Riccati differential equation. We briefly recall the following facts about the solutions of Riccati equations.

Proposition 1: Consider the time-varying matrix Riccati differential equation

$$\dot{K} = Q(t) + KS(t)K + KF(t) + F(t)^T K, \tag{4}$$

with the initial condition $K(0) = K_0$, where S and F are continuous functions of time.
(i) [19, p. 175] If $K_0 > 0$, S is symmetric and Q is symmetric positive definite the solution stays symmetric positive definite for $t > 0$ (as long as it exists).
(ii) If Q is symmetric positive semi-definite, S is symmetric negative semi-definite and K_0 is symmetric positive semi-definite then (4) has a solution for all times $t \geq 0$. This solution is unique, symmetric and positive semi-definite. If K_0 is positive definite then this solution is positive definite.
We now state our main result.

Theorem 1: Consider the system (1) and the cost (2). Given some measurements $Y_m(t)$ and their associated inputs $A_m(t)$ for $t \in [0, T]$, assume that unique solutions $\hat{R}(t)$ and $K(t)$ to (3a) and (3b) exist on $[0, T]$. Assuming further that

$$\begin{aligned} W(T) := \int_0^T \Bigg(&\mathrm{trace}\,\frac{1}{4}\left[\frac{1}{4}g^2K^{-2}\left((\hat{R}^T R)^2 - I\right)\right. \\ &+ K^{-1}QK^{-1}\left(I - \hat{R}^T R\right) \\ &\left.-\hat{R}^T Y_m K\left(\hat{R}^T RK^{-1} - K^{-1}\hat{R}^T R\right)\right]\Bigg) d\tau, \end{aligned} \tag{5}$$

is nonnegative, the filter (3) yields a near-optimal estimate $\hat{R}(T)$ of the state $R(T)$ in the sense that $J_T \leq J_T^* + W(T)$, where J_T^* denotes the optimal value for the cost (2) and $W(T)$ is the optimality gap.

Proof: Under the conditions listed in the theorem, a somewhat lengthy but not overly difficult calculation shows that

$$\begin{aligned} J_T = &\\ &\frac{1}{4}\operatorname{trace}\left[(R(T)-\hat{R}(T))K^{-1}(T)(R(T)-\hat{R}(T))^T\right] \\ &+\int_0^T\left(\frac{1}{2}\left\|\delta-\frac{1}{4}g(K^{-1}\hat{R}^TR-R^T\hat{R}K^{-1})\right\|^2\right. \\ &\left.+\frac{1}{4}\phi_I(Y_m^T\hat{R})\right)d\tau+W(T). \end{aligned} \tag{6}$$

According to Proposition 1, $K(T)$ is a positive definite matrix and hence its inverse $K^{-1}(T)$ exists. This matrix is also positive definite and hence the first term on the right hand side of (6) is a squared matrix Frobenius norm and is positive. We assumed $W(T)$ to be positive and hence the right hand side of Equation (6) consists of three positive terms. Therefore, the cost function J_T fulfills the inequality

$$J_T \geq \int_0^T\left(\frac{1}{4}\phi_I(Y_m^T\hat{R})\right)d\tau. \tag{7}$$

Observe that the right hand side of Equation (7) is independent of any specific choice of the variables $R(0)$, $\delta|_{[0,T]}$ and $\epsilon|_{[0,T]}$ and only depends on the measured data $Y_m(.)$ and the filter signals. This implies that the right hand side of Equation (7) is a lower bound for J_T. Therefore the minimum of the cost J_T is reachable once equality holds in Equation (7). Equivalently, the cost function will be minimized if among all the possible values of $R(0)$, $\delta|_{[0,T]}$ and $\epsilon|_{[0,T]}$ that together with $Y_m(.)$ and $A_m(.)$ fulfill the system equations (1), we choose a set that yields

$$\begin{cases} R(T) &= \hat{R}(T), \\ \delta &= \frac{1}{4}g\left(K^{-1}\hat{R}^TR-R^T\hat{R}K^{-1}\right), \\ W(T) &= 0. \end{cases} \tag{8}$$

We will show next that the first two conditions are achievable. However, we can not guarantee the third condition and hence only near-optimality will be concluded for this filter. Consequently, an optimality gap $W(T)>0$ remains by which the cost function can deviate from its optimal value during the estimation period.

Consider a trajectory $R_h : [0,T] \longrightarrow \mathrm{SO}(3)$ that is generated by

$$\dot{R_h} = R_h\left(A_m+\frac{1}{4}g\left(R_h^T\hat{R}K^{-1}-K^{-1}\hat{R}^TR_h\right)\right) \tag{9}$$

and is fixed by the final condition $R_h(T) := \hat{R}(T)$, where $\hat{R}$ and K^{-1} are solutions of the proposed filter in Equation (3). It is straightforward to show that (9) has a unique initial state $R_h(0)$ that produces the final condition $R_h(T) = \hat{R}(T)$. Also define a signal $\epsilon_h : [0,T] \longrightarrow \mathrm{SO}(3)$ by

$$\epsilon_h := R_h^TY_m, \tag{10}$$

and a signal $\delta_h : [0,T] \longrightarrow \mathfrak{so}(3)$ by

$$\delta_h := \frac{1}{4}\left(R_h^T\hat{R}K^{-1}-K^{-1}\hat{R}^TR_h\right). \tag{11}$$

Equations (9) and (10) show that $R_h(0)$, $\delta_h|_{[0,T]}$ and $\epsilon_h|_{[0,T]}$ together with $A_m(.)$ and $Y_m(.)$ satisfy the system equations (1). Moreover, Equations (9) and (11) show that R_h is a trajectory of the system given in Equation (1). Hence, $\hat{R} = R_h$ is the desired solution. ■

Remark 1: The previous theorem assumes that unique solutions exist for Equation (3) on $[0,T]$. However, it is not clear if (3b) will always admit a solution on this interval. This can potentially become problematic for (3a) since K appears in this equation as well. We know that according to Proposition 1, for Equation (3b) to have a unique solution the term $(Y^T\hat{R}+\hat{R}^TY)$ should be positive semi definite. This is the case if and only if the angle of rotation between Y and $\hat{R}$ is less than 90 degrees. However, this is not trivial to verify as $\hat{R}$ also depends on K. Nevertheless, in Section V we will show that in all the cases we tested by simulation, the assumption that a unique K exists for all time holds true.

We also assume that $W(T)$ is positive. This is certainly the case when we apply this filter to an S^1 system as our filter specializes to the work by Coote et al [15] and $W(T)$ is then fourth order in the error. On $SO(3)$, by choosing $Q = gI$ the first two terms in $W(T)$ (5) yield a positive value that specializes to the Gap incurred in the S^1 case. The third term in $W(T)$ is a curvature correction term that is zero in the S^1 case. In the general case, this term appears to evolve slowly and extensive simulation studies indicate that it is dominated by the first two terms.

Remark 2: Note that importance of Theorem 1 lies in providing a bound $W(T)$ for evaluating the performance of the filter against an optimal filter. Additionally, this bound is numerically quantifiable and is formulated in Equation (5). More importantly $W(T)$ is usually small. We demonstrate this later in Section V as well.

A. Filtering on S^1

Once the system (1) is confined to only involve rotations around a fixed axis, for example the z axis, it can be easily shown that the filter proposed in this paper coincides with the near-optimal filter on S^1 due to Coote et al [15]. A rotation of θ radians around the z axis can be parameterized as

$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{12}$$

Similarly we model the output and the measurement disturbance that are rotations of y and ϵ radians around the z axis, respectively. An angular velocity perpendicular to the plane of the rotation R and of w radians per time unit can be parameterized as

$$A = \begin{bmatrix} 0 & -w & 0 \\ w & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{13}$$

Similarly we model the process disturbance that is an angular velocity perpendicular to the plane of the rotation R of δ radians per time unit. Substituting these parameterized matrices into Equation (1) yields

$$\begin{cases} \dot{\theta} &= w + g\delta, \\ y &= \theta + \epsilon. \end{cases} \tag{14}$$

This is precisely the system on S^1 that Coote et al. [15] considered. Substituting in the cost given in Equation (2) and defining K_0^{-1} as

$$K_0^{-1} = \begin{bmatrix} k_0^{-1} & 0 & 0 \\ 0 & k_0^{-1} & 0 \\ 0 & 0 & h_0^{-1} \end{bmatrix} \tag{15}$$

yields

$$J_T = k_0^{-1} \left(1 - \cos\left(\theta_0\right)\right) + \int_0^T \left(\mid \delta \mid^2 + \left(1 - \cos\left(\epsilon\right)\right)\right) d\tau. \tag{16}$$

This the cost used in Coote et al [15]. We choose Q to be a multiple of the identity, $Q = gI$, and substitute parameterizations similar to the ones above into the filter equations (3). This yields

$$\begin{cases} \dot{\hat{\theta}} = w + k\sin(y - \hat{\theta}) \\ \dot{k} = \frac{1}{2}g^2 - k^2\cos(y - \hat{\theta}) \\ \dot{h} = \frac{1}{2}g^2 - h^2, \end{cases} \tag{17}$$

in harmony with the work by Coote et al [15]. Hence the proposed filter is a generalization of the filter introduced by Coote et al. Note that there is an extra equation (the third line in Equation (17)) that does not affect the other equations as far as it has a solution.

V. SIMULATIONS

In this section the performance of the proposed filter as well as the validity of its assumptions are demonstrated by a suite of simulation results.

A. Comparison to EKF

The proposed filter's performance is compared with a quaternion based implementation of the commonly applied Extended Kalman Filter (EKF). The process disturbance δ is generated in form of a skew symmetric matrix and the disturbance contaminated input is fed to both the EKF (after conversion to a quaternion) and the proposed filter. Note that we have adopted the concept that although the system on $\mathrm{SO}(3)$ works with an exact input Ω, the filter's measurement of Ω is noisy, that is $A = \Omega - g\delta$. The measurement disturbance ϵ is generated on an $\mathbb{R}^3$ vector and is transformed via the exponential map to a rotation and a quaternion, respectively, for the proposed filter and the EKF. The simulation step size is 0.001 time units and the stop time is at 30 time units. The angular velocity inputs $\Omega(t)$ considered are continuous skew symmetric matrices (18). Two cases are considered, firstly a periodically varying velocity and secondly a constant velocity. Counter-intuitively the constant angular velocity is actually more challenging for the filter since it generates system trajectories that "wrap around" the rotation group, exploring the full nonlinear nature and the global topology of $\mathrm{SO}(3)$, whereas the periodic velocity leads to trajectories that are confined to a local neighborhood in $\mathrm{SO}(3)$.

$$\begin{aligned} \Omega_1 &= \begin{bmatrix} 0 & -\cos(t) & 0 \\ \cos(t) & 0 & -3\sin(5t) \\ -0 & 3\sin(5t) & 0 \end{bmatrix}, \\ \Omega_2 &= \begin{bmatrix} 0 & -1 & 4 \\ 1 & 0 & -3 \\ -4 & 3 & 0 \end{bmatrix}. \end{aligned} \tag{18}$$

The initial state distribution considered is an exponential map of a normally distributed noise (with standard deviation of $\frac{\pi}{2}$) generated on the skew symmetric matrices. Note that this corresponds to potentially very large initialization errors.

The matrices Q and $K(0)$ are the only tuning parameters of the proposed filter and for the sake of simplicity they are fixed to $Q = I$ and $K(0) = 10I$ for all the simulation runs. As for tuning the EKF, we have tried to find a feasible set of filter parameters for every combination of input/noise levels in advance.

We have tested the two filters in five different treatments (choice of noise/input parameters), listed in Table I, with 40 repeats in each treatment. The angle of rotation between the true system trajectory R and the estimate $\hat{R}$ forms the basis of our comparison. The histogram in Figure 1 depicts the performance of the filters in terms of the tracking error obtained over all filtering steps for every repeat in the five treatments explained in Table I. Since both filters have the same starting point this histogram is indicative for both estimation speed and precision. The observed tracking error data has a mean, mode and standard deviation of (0.041,0.007,0.117) in

TABLE I

This table presents five different simulation treatments based on two different angular velocity inputs (18) *with two or three combinations of noise levels associated with each input. We have done simulations with* 40 *repeats in each treatment. The measurement noise ϵ is the dominant noise affecting the filters directly. Different levels of process noise δ are also considered for completeness.*

	Input $= \Omega_1$	
Parameters	$\log(\epsilon)$	$g\delta$
Treatment 1	$\sim \mathcal{N}(0, \frac{\pi}{15})$	$\sim \mathcal{N}(0, \frac{\pi}{22})$
Treatment 2	$\sim \mathcal{N}(0, \frac{\pi}{7})$	$\sim \mathcal{N}(0, \frac{\pi}{2})$
	Input $= \Omega_2$	
Treatment 3	$\sim \mathcal{N}(0, \frac{\pi}{300})$	$\sim \mathcal{N}(0, \frac{\pi}{30})$
Treatment 4	$\sim \mathcal{N}(0, \frac{\pi}{100})$	$\sim \mathcal{N}(0, \frac{\pi}{50})$
Treatment 5	$\sim \mathcal{N}(0, \frac{\pi}{13})$	$\sim \mathcal{N}(0, \frac{\pi}{20})$

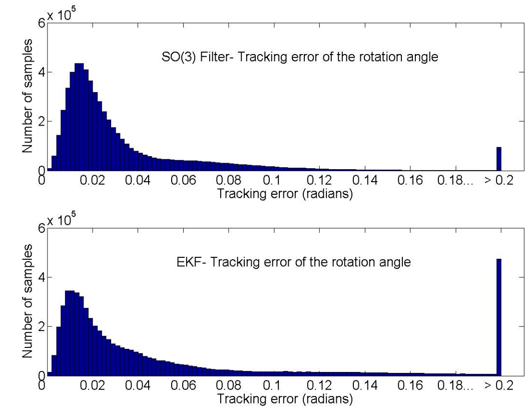

Fig. 1. Histogram of the two filters' angle of rotation tracking errors over all the simulation data. Note that angles greater than 0.2 radians are accumulated at 0.2 radians.

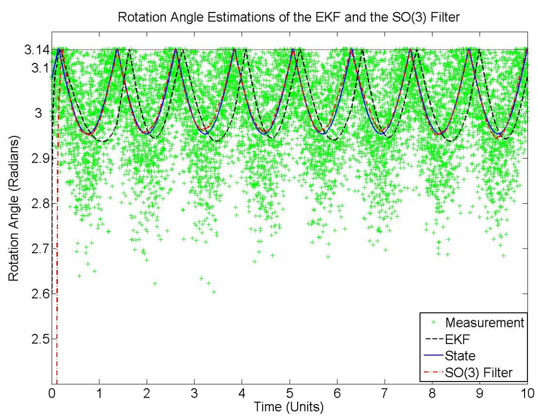

Fig. 2. The rotation angle tracking performance of the proposed filter and the EKF in a typical situation that the EKF performs badly. Note that the figure is zoomed to a neighborhood around an angle of π radians.

the proposed filter as opposed to (0.096,0.000042,0.27) in the EKF. These results indicate that the proposed filter achieves comparable or even slightly better estimates compared to the EKF in terms of consistency and repeatability of results. Moreover, the proposed filter demands less effort for the tuning process. Specifically, the lower mean and standard deviation associated with the proposed filter's tracking error indicate more robustness to the uncertainties of the system. The mode observed in the implemented EKF indicates that the EKF performs well in most of the situations where it is well tuned. However, in practice detailed knowledge of system noise processes is rarely available and tuning EKFs is known to be a challenging problem requiring a range of techniques [4, p. 204]. In particular, Figure 2 shows a typical situation where the implemented EKF fails. Here, the rotation angle of the true state R, that is shown in blue, is mainly oscillating in a small neighborhood close to π. Recall that the filters are initialized with $\hat{R}(0) := I$ that corresponds to an angle of 0 radians. In this situation, measurements Y with angles greater than π radians "wrap around" $\mathrm{SO}(3)$ in the sense that an angle of $\pi + \theta$ radians is the same as the angle of $-\pi + \theta$ radians. The EKF, since it is based on a linearization, hence encounters an ambiguity as to which is the right "direction" to correct the estimate. It might be possible - by modifying the EKF - to tackle this ambiguity if the axis of rotation was constant. However, our simulation results tell us that this is not always the case. On the other hand the proposed filter, whose rotation angle estimates are shown in blue, is posed directly on $\mathrm{SO}(3)$ and can correctly judge the nonlinear trend in the state and the measurements.

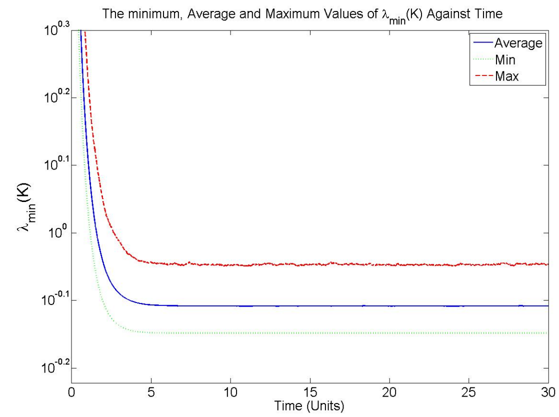

Fig. 3. Evolution of the smallest eigenvalue of K in time.

B. Results on Remark 1

In Remark 1 we discussed the assumptions we made in Theorem 1. Although there was no formal guarantee provided for when these assumptions hold, here we study their validity by means of simulations.

1) Simulation results on the existence of solutions: Figure 3 shows the evolution for the statistics of K against time. We observe that the minimal eigenvalue of K is positive in all the situations tested.

2) Simulation results on the smallness and positivity of $W(T)$: Remark 1 stated that $W(T)$ is a small positive quantity. Positivity is crucial in order for Theorem 1 to apply. From simulations we have observed that, relatively, $W(T)$ has a small positive value for a majority of the simulation runs listed in Table I. However, our results indicate that negative values for $W(T)$ also occur in some rare instances in each treatment. By increasing Q the number of situations where $W(T)$ has a negative value will be reduced and indeed will be eliminated for large enough Q. Such a choice will of course compromise the tracking performance of the filter. For example, Figure 4 shows the statistics of $W(T)$ over all simulations listed in Table I where $Q = 5I$. These results are organized in two different graphs according to the two inputs considered (18). With this choice of Q the negative instances for $W(T)$ occur only in treatments where the constant input Ω_2 is applied. By choosing Q even larger ($Q = 10I$), it was observed that $W(T)$ was always positive. To further see that $W(T)$ is only a small gap (recall that $W(T)$ denotes an upper bound on the difference between the cost the proposed filter achieves and the optimal cost) Figure 5 shows the cost

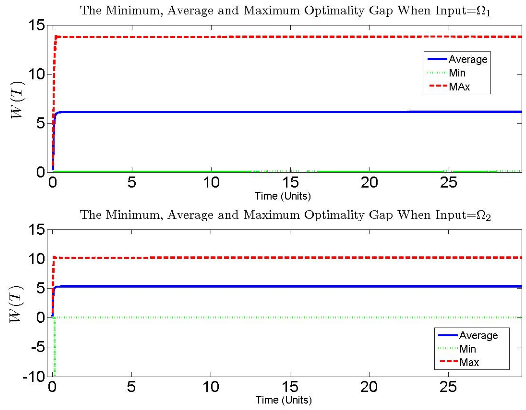

Fig. 4. Statistics of $W(T)$ over all simulation data when $Q = 5I$.

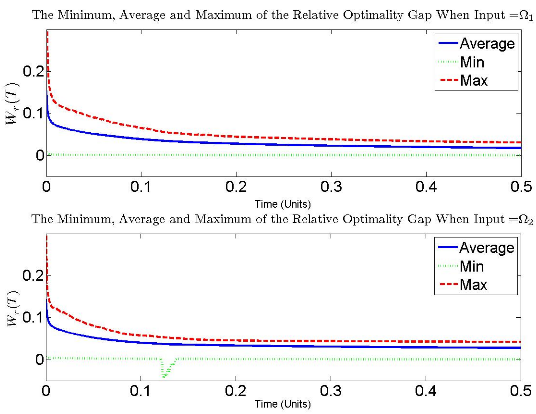

Fig. 5. Statistics of the relative optimality gap ($W_r(T) \coloneqq \frac{W(T)}{W(T)+J_{min}}$) over all simulation data when $Q = 5I$. Note the figure is zoomed to the first 0.5 units of time.

for the proposed filter relative to the optimal cost, i.e. the ratio $W_r(T) \coloneqq \frac{W(T)}{W(T)+J_{min}}$, plotted against time for each input (18) separately. Here we observe that the proposed filter's optimality gap $W(T)$ is *relatively* small compared to the optimal value. $W_r(T)$ appears to converge asymptotically to zero. This is because $W(T)$ in the numerator converges to a stable range (that is shown in Figure 4) although J_{min} keeps growing in the denominator. This is the case for both the constant and the periodic inputs (18).

VI. Conclusion

We proposed a near-optimal deterministic attitude filter posed directly on the rotation group $\mathrm{SO}(3)$. The proposed filter is a generalization of a near-optimal filter on S^1 due to Coote et al [15]. The proposed filter shows comparable but more robust performance compared to a quaternion based EKF filter while needing almost no calibration. Furthermore, it comes with an explicit bound on the distance to optimality and we showed that the optimality gap is asymptotically small, justifying the term "near-optimal".

VII. Acknowledgments

This research was supported by the Australian Research Council through discovery grant DP0987411 "State Observers for Control Systems with Symmetry". The authors would like to thank Ashkan Amirsadri for the EKF implementation.

References

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

[2] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering, Transactions of the ASME*, vol. 82, no. 1, pp. 35–45, 1960.

[3] S. I. Marcus, "Algebraic and geometric methods in nonlinear filtering," *SIAM Journal on Control and Optimization*, vol. 22, no. 6, pp. 817–844, 1984.

[4] B. DO. Anderson and J. Moore, *Optimal filtering*. Prentice Hall, Englewood Cliffs, NJ, 1979.

[5] A. Doucet and N. De Freitas, *Sequential Monte Carlo methods in practice*. Springer, 2001.

[6] S. J. Julier and J. K. Uhlmann, "Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations," in *American Control Conference, 2002. Proceedings of the 2002*, vol. 2, 2002, pp. 887–892.

[7] S. Salcudean, "A globally convergent angular velocity observer for rigid body motion," *IEEE Transactions on Automatic Control*, vol. 36, no. 12, pp. 1493–1497, 1991.

[8] 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.

[9] S. Bonnabel, P. Martin, and P. Rouchon, "A non-linear symmetry preserving observer for velocity-aided inertial navigation," in *American Control Conference (ACC06)*, 2006, pp. 2910–2914.

[10] R. Mahony, T. Hamel, and J. -M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *Automatic Control, IEEE Transactions on*, vol. 53, no. 5, pp. 1203–1218, June 2008.

[11] R. E. Mortensen, "Maximum-likelihood recursive nonlinear filtering," *Journal of Optimization Theory and Applications*, vol. 2, no. 6, pp. 386–394, 1968.

[12] A. H. Jazwinski, *Stochastic processes and filtering theory*. Academic Press, April 1970.

[13] 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.

[14] 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.

[15] P. Coote, J. Trumpf, R. Mahony, and J. C. Willems, "Near-optimal deterministic filtering on the unit circle," in *IEEE Conference on Decision and Control*, 2009, pp. 5490–5495.

[16] J. L. Farrell, J. C. Stuelpnagel, R. H. Wessner, J. R. Velman, and J. E. Brook, "A Least Squares Estimate of Satellite Attitude (Grace Wahba)," *SIAM Review*, vol. 8, no. 3, pp. 384–386, 1966.

[17] F. Markley, "Attitude determination using vector observations and the singular value decomposition," *Journal of the Astronautical Sciences*, vol. 36, no. 3, pp. 245–258, 1988.

[18] 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.

[19] H. Abou-Kandil, G. Freiling, V. Ionescu, and G. Jank, *Matrix Riccati equations in control and systems theory*. Systems and Control: Foundations and Applications, 2003.