

# Minimum-Energy Filtering on the Unit Circle Using Velocity Measurements with Bias and Vectorial state Measurements

Mohammad Zamani, Minh-Duc Hua, Jochen Trumpf, and Robert Mahony

**Abstract**— We consider minimum-energy filtering of a system defined on the unit circle when the angular velocity measurements are contaminated with deterministic measurement error and bias. We propose a second order approximate minimum-energy filter for this system using vectorial measurements. This work extends prior work in two aspects; Firstly, by including a model for slowly time varying angular velocity measurement bias, we estimate and reject the bias in order to estimate the state of the system more accurately. Secondly, rather than using full state measurements we use vector measurements to derive the filter. Both of these two innovations make the filter more practical in real world applications. In simulations we show that the proposed filter is globally convergent and robust to different levels of measurement error and bias.

## I. INTRODUCTION

Optimal filtering in the context of linear state space models was introduced in the early 1960s by Kalman [1]. The framework for this work is stochastic where the observations of the input and the output signals of the system are assumed to be contaminated with zero mean Gaussian noise signals. Based on the observations, the Kalman filter estimates the state of that system by minimizing the covariance of the estimation error. Optimal filtering theory has also been developed in a deterministic system modeling framework where, unlike in the stochastic framework, the observations are contaminated with disturbance signals modeled as unknown deterministic functions of time. The deterministic optimal filtering problem, called the minimum-energy filtering problem, is to find an estimate for the state that is compatible with the observations and that minimizes an energy cost functional in the associated disturbance signals. For linear systems, the minimum-energy filtering framework yields the same Kalman filtering formulas when using a least squares cost (cf. [2]–[5]).

In the late 1960's Mortensen [6] proposed a systematic approach to computing minimum-energy filters for general nonlinear systems modeled on vector spaces. The method was further explored by Hijab [7]. In this approach the optimal filtering problem is broken down into two steps. The first step involves applying the maximum principle of optimal control and dynamic programming to optimize an energy functional in the system disturbances. In the second step a further optimization takes place over the initial state of the system. Krener [8] proved that under some conditions

including the uniform observability of the system and the presence of a time-dependent “forgetting factor” in the cost, a minimum-energy estimate converges exponentially fast to the true state.

Recently, minimum-energy filtering has been employed in robotic applications where the system nonlinearity is partly due to the geometric structure of the underlying state space. For instance, Aguilar *et al.* [9] proposed a minimum-energy nonlinear filter for systems with state space embedded in  $\mathbb{R}^n$  with perspective outputs. They account for the geometry of the system using algebraic state constraints. Coote *et al.* [10] designed a near-optimal minimum-energy filter directly considering the geometric structure of the unit circle  $S^1$ . In follow up work [11], the authors extended this work by designing a near-optimal minimum-energy filter posed directly on the Special Orthogonal Group  $SO(3)$ . These two filters are derived through identification of a suitable “Lyapunov” function for the optimality analysis and include explicit bounds on their distance to optimality. In [12] the authors extended Mortensen’s minimum-energy filtering approach to a system defined directly on  $S^1$  proving that Coote’s filter on  $S^1$  [10] is a second-order approximation of a minimum-energy filter. Due to the systematic nature of the method used in that work [12], it was straightforward to derive higher order approximations of the minimum-energy filter on  $S^1$ . The authors subsequently also extended this work providing second order approximate minimum-energy filters on  $SO(3)$  [13] and also on the special Euclidean group  $SE(3)$  [14].

In this paper we consider geometric minimum-energy filtering on  $SO(2) \triangleq S^1$  when bias is present in the angular velocity measurements. Filtering on  $S^1$  has many applications in communications, coding theory and power networks. The  $S^1$  system is also interesting from the theoretical point of view as it is a training ground for the attitude system  $SO(3)$  and more complicated systems defined on other Lie groups. In this work we propose the exact form of a minimum-energy observer on  $SO(2) \times \mathbb{R}$  for estimating the rotation and the angular velocity bias. We show that the observer gains are related to the second order derivatives of a value function of the associated optimization problem. We provide Riccati equations that dynamically update these gains based on a second-order approximation of the minimum-energy filter. In this paper we formulate the state kinematics on  $SO(2)$  to facilitate using vector measurements and we show that this formulation is equivalent to the state kinematics on  $S^1$  considered in [12]. Furthermore, in this work we use a vector measurement model with vector measurement error along

Mohammad Zamani, Jochen Trumpf, and Robert Mahony are with the Research School of Engineering, ANU, Australia. E-mails: mohammad.zamani@anu.edu.au, jochen.trumpf@anu.edu.au, robert.mahony@anu.edu.au and Minh-Duc Hua is with the laboratory ISIR UPMC, France. hua@isir.upmc.fr

with a vector norm cost function rather than measuring the state directly with measurement error modeled on  $S^1$  and a nonlinear cost. These measurements make the filter more practical for applications such as vehicle navigation where the angular velocity is measured by a noisy biased gyroscope and the state measurement is a vector measurement such as the horizontal components of the magnetic field, measured by a magnetometer [15], [16]. While the current measurement model is preferable in some applications we show that the resulting filter is the same as the one in our previous work [12] in the absence of bias. A simulation study is provided showing that the proposed filter converges quickly in the presence of large measurement errors and bias variation and initialization errors.

The rest of the paper is organized as follows. Section II formally introduces the system on  $SO(2)$  and the problem of minimum-energy filtering on this system. The derivations and the formulas of the proposed filter are given in Section III. Section IV is a simulation study of the performance of the proposed filter. Finally Section V concludes the paper.

## II. PROBLEM FORMULATION

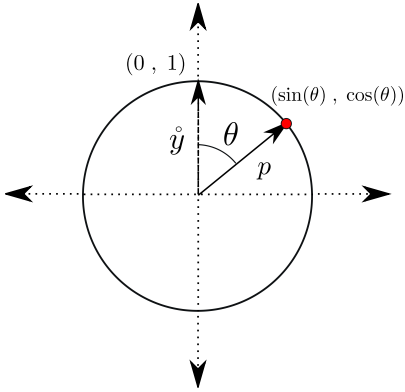


Fig. 1. An object moving around the unit circle

Consider an object moving on the unit circle as shown in Figure 1. Let  $\omega \in \mathbb{R}$  be the instantaneous angular velocity of the object. The rotation  $X$  satisfies the following kinematics equation

$$\dot{X} = X\omega_{\times}, X(0) = X_0, \quad (1)$$

where

$$X = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \in SO(2) \quad (2)$$

is the state of the system and represents the rotation with respect to the angle  $\theta$ . The initial rotation  $X_0$  is unknown. The cross notation  $(\cdot)_{\times} : \mathbb{R} \rightarrow \mathfrak{so}(2)$  is defined as

$$\omega_{\times} := \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix}. \quad (3)$$

Conversely the notation  $\text{vex}(\cdot) : \mathfrak{so}(2) \rightarrow \mathbb{R}$  extracts the scalar part,  $\text{vex}(\omega_{\times}) = \omega$ .

*Lemma 1:* The system (1) can be equivalently written in terms of the angle  $\theta$  with kinematics

$$\dot{\theta} = \omega, \theta(0) = \theta_0. \quad (4)$$

*Proof:* Note that from (1)

$$X^{-1}\dot{X} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \omega, \quad (5)$$

Also from (1) and (2)

$$\begin{aligned} X^{-1}\dot{X} &= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} -\sin(\theta) & -\cos(\theta) \\ \cos(\theta) & -\sin(\theta) \end{bmatrix} \dot{\theta} \\ &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \dot{\theta}, \end{aligned} \quad (6)$$

and the result (4) follows.  $\blacksquare$

Although system (4) was used in previous works on  $S^1$  [10], [12], we continue our derivation using system (1) to facilitate using vector measurements introduced later in Equation (9). The angular velocity  $\omega$  is measured as

$$u = \omega + B_{\omega}v_{\omega} + b, \quad (7)$$

where  $v_{\omega} \in \mathbb{R}$  is the measurement error signal and  $B_{\omega} \in \mathbb{R}$  is a scaling known from the model. The signal  $b \in \mathbb{R}$  is a constant or slowly time varying bias in the measurement  $u \in \mathbb{R}$  that we model as

$$\dot{b} = B_b v_b, b(0) = b_0, \quad (8)$$

where  $b_0$  is unknown,  $v_b \in \mathbb{R}$  is an unknown rate of change of the bias  $b$  and  $B_b \in \mathbb{R}$  is a scaling known from the model. Note that in our measurement model (7) the measurement error  $v_{\omega}$  and the bias  $b$  are modeled separately although we consider a general deterministic measurement model. This is due to the fact that we are going to minimize a cost (12) in the size of the measurement error  $v_{\omega}$  to derive our filter. The assumption behind this is that the measurement error should not, in general, be large. However, bias, in general, needs not to be small but rather slowly time-varying. Therefore, we use an independent bias model (8) with  $v_b$  modeling a small unknown rate of change for  $b$ .

Assume a vector measurement of the current position of the object is given by

$$y = X^{\top} \hat{y} + Dw, \quad (9)$$

where  $\hat{y} = (1, 0)^{\top}$ ,  $w \in \mathbb{R}^2$  is an unknown measurement error signal and  $D \in \mathbb{R}^{2 \times 2}$  is a full rank scaling matrix known from the model.

*Remark 1:* Depending on the application one might have a different state measurement model such as

$$y' = (WX)^{\top} \hat{y}, \quad (10)$$

where  $y' \in S^1$  is a vector measurement of the state  $X$  and  $W \in SO(2)$  represents the measurement error. In case the full angle measurement is available the following output map is suitable.

$$y'' = \theta + w', \quad (11)$$

where  $y' \in \mathbb{R}$  is the measured state angle  $\theta \in \mathbb{R}$  and  $w' \in \mathbb{R}$  is the measurement error. The output measurement (11) was previously considered in [10], [12].

Consider the cost functional

$$J(t; X_0, b_0, v_\omega|_{[0, t]}, v_b|_{[0, t]}, w|_{[0, t]}) = \frac{\text{trace}(I - X_0)}{K_1} + \frac{b_0^2}{2K_2} + \frac{1}{2} \int_0^t (v_\omega^2 + v_b^2 + \|w\|^2) d\tau, \quad (12)$$

where  $K_1, K_2 \in \mathbb{R}^+$  are given. Later on, these parameters will appear in the initial conditions of the proposed filter. If available, a priori information on the expected size of the initial state and the initial bias can be used to tune  $K_1$  and  $K_2$  relative to each other and the other unknowns in the cost function. The trace function is used to measure the size of the initial state  $X_0$  relative to the identity matrix  $I \in \mathbb{R}^{2 \times 2}$  (that is the identity element for the group  $\text{SO}(2)$ ).

*Remark 2:* In case the measurement model (11) is used, the cost function (12) is modified to allow for measuring the measurement error  $w'$  as an angle.

$$J'(t; X_0, b_0, v_\omega|_{[0, t]}, v_b|_{[0, t]}, w'|_{[0, t]}) = \frac{\text{trace}(I - X_0)}{K_1} + \frac{b_0^2}{2K_2} + \int_0^t \left( \frac{1}{2} (v_\omega^2 + v_b^2) + 1 - \cos(w') \right) d\tau. \quad (13)$$

The following is the formal statement of the minimum-energy filtering problem for the system (1).

*Problem 1:* Given the system (1), the measurement models (7), (9), the bias model (8) and the past measurements  $u|_{[0, t]}$  and  $y|_{[0, t]}$  find the state estimate  $\hat{X}(t)$  for the current state  $X(t)$  and the bias estimate  $\hat{b}(t)$  for the current bias  $b(t)$  such that the cost (12) is minimized over the unknowns  $X_0, b_0, v_\omega|_{[0, t]}, v_b|_{[0, t]}$  and  $w|_{[0, t]}$ .

Note that the cost (12) encodes the total energy associated with the unknowns  $X_0, b_0, v_\omega|_{[0, t]}, v_b|_{[0, t]}$  and  $w|_{[0, t]}$ . In a sense, by minimizing (12) the goal is to find unknowns of minimum energy that together with the measurements  $u|_{[0, t]}$  and  $y|_{[0, t]}$  satisfy the model equations (1), (7), (8) and (9). Note that in general one might find infinitely many possible combinations of these unknowns that together with the measurements satisfy the model equations. However, by minimizing the cost (12) a set of minimizing unknowns is singled out that collectively has minimal energy. Substituting the minimizing unknowns and the measurements into equations (1), (7), (8) and (9) yields the minimum-energy state trajectory  $X_{[0, t]}^*$  and the minimum-energy bias  $b_{[0, t]}^*$ . The subscript  $[0, t]$  indicates that the optimization takes place on the interval  $[0, t]$ . The final values  $X_{[0, t]}^*(t)$  and  $b_{[0, t]}^*(t)$  are then assigned as the minimum-energy estimates at time  $t$ ,  $\hat{X}(t) := X_{[0, t]}^*(t)$  and  $\hat{b}(t) := b_{[0, t]}^*(t)$ . In the following, rather than resolving this infinite dimensional optimization problem at each time instance  $t$ , we seek recursive filters that directly update the estimates.

We proceed to solve Problem 1 in the spirit of our previous work [12] by making the measurements constraint (9) explicit in the cost. Hence, substituting  $w$  in (12) from (9)

yields the simplified cost

$$J(t; X_0, b_0, v_\omega|_{[0, t]}, v_b|_{[0, t]}) = \frac{\text{trace}(I - X_0)}{K_1} + \frac{1}{2} \frac{b_0^2}{K_2} + \frac{1}{2} \int_0^t (v_\omega^2 + v_b^2 + \|y - X^\top \hat{y}\|_{R^{-1}}^2) d\tau, \quad (14)$$

where  $R := DD^\top$  is positive definite and

$$\|y - X^\top \hat{y}\|_{R^{-1}}^2 := (y - X^\top \hat{y})^\top R^{-1} (y - X^\top \hat{y}). \quad (15)$$

Now, similar to optimal control problems, the cost (14) is to be minimized over  $v_\omega|_{[0, t]}, v_b|_{[0, t]}$ . For now consider  $X_0$  and  $b_0$  as fixed but later to fully solve the problem we will need to further optimize over these two initial values. In order to apply Hamilton Jacobi Bellman theory [17], define the following pre-Hamiltonian function,

$$\mathcal{H}^-(X, b, \mu_\omega, \mu_b, v_\omega, v_b, t) := \frac{1}{2} (v_\omega^2 + v_b^2 + \|y - X^\top \hat{y}\|_{R^{-1}}^2) - \mu_\omega (u - b - B_\omega v_\omega) - \mu_b B_b v_b, \quad (16)$$

where  $\mu_\omega \in \mathbb{R}$  represents the costate variable  $\Theta \in \mathfrak{so}^*(2)$  via  $\langle (\mu_\omega)_\times, \Gamma \rangle = \Theta(\Gamma)$  for all  $\Gamma \in \mathfrak{so}(2)$ . This algebraic representation will be used in the following without further reference. The variable  $\mu_b \in \mathbb{R}$  is the costate variable associated with the bias  $b$ . The optimal Hamiltonian  $H$  is obtained by minimizing the pre-Hamiltonian  $\mathcal{H}^-$  over the signals  $v_\omega$  and  $v_b$  that yields  $v_\omega^* = -B_\omega \mu_\omega$  and  $v_b^* = B_b \mu_b$ .

$$\mathcal{H}(X, b, \mu_\omega, \mu_b, t) = \frac{1}{2} \left( -\mu_\omega^2 Q_\omega - \mu_b^2 Q_b + \|y - X^\top \hat{y}\|_{R^{-1}}^2 \right) - \mu_\omega (u - b), \quad (17)$$

where  $Q_\omega := B_\omega^2$  and  $Q_b := B_b^2$  are positive definite. Define the value function

$$V(X, b, t) := \min_{v_\omega|_{[0, t]}, v_b|_{[0, t]}} J(t; X_0, b_0, v_\omega|_{[0, t]}, v_b|_{[0, t]}, w|_{[0, t]}), \quad (18)$$

where  $J$  is the cost (14) and the minimization is subject to the equations (1) and (8). The initial boundary condition for the value function (18) is obtained from (14)

$$V(X(0), b(0), 0) = \frac{\text{trace}(I - X_0)}{K_1} + \frac{1}{2} \frac{b_0^2}{K_2}. \quad (19)$$

In the following we write down the Hamilton-Jacobi-Bellman equation [17] relating the optimal Hamiltonian (17) and the value function (18)

$$\mathcal{H}(X, b, \nabla_X V(X, b, t), \nabla_b V(X, b, t), t) - \nabla_t V(X, b, t) = 0. \quad (20)$$

Up to here we have only optimized over the signals  $v_\omega|_{[0, t]}$  and  $v_b|_{[0, t]}$ . To complete the optimal filtering problem, we also need to optimize  $V$  over the initial values  $X_0$  and  $b_0$ . This is equivalent to further optimizations over  $X(t)$  and  $b(t)$  since (the minimizing)  $X(t)$  and  $b(t)$  are uniquely determined given the measurements  $u|_{[0, t]}$ , the equations (1), (7) and (8), the (minimizing) signals  $v_\omega|_{[0, t]}$  and  $v_b|_{[0, t]}$  and the (minimizing) initial values  $X_0$  and  $b_0$ . Hence, similar

to Mortensen's approach [6], the minimum-energy estimates  $\hat{X}(t)$  and  $\hat{b}(t)$  are characterized by the criticality conditions

$$\begin{aligned}\nabla_X V(X, b, t)|_{X=\hat{X}(t), b=\hat{b}(t)} &= 0, \\ \nabla_b V(X, b, t)|_{X=\hat{X}(t), b=\hat{b}(t)} &= 0.\end{aligned}\quad (21)$$

Solving Equations (21) is clearly a way to obtain the minimum-energy estimates  $\hat{X}(t)$  and  $\hat{b}(t)$ , minimizing the cost (14) at every time  $t$ . However, in the following, rather than solving this optimization problem for each time  $t$ , the goal is to find differential equations (filters) that dynamically update these estimates when new measurements are obtained as time evolves.

### III. FILTER DERIVATION AND RESULTS

In this section, we propose a filter that estimates the state  $X$  and the bias  $b$  by approximating the solution to Problem (1). The proposed filter is derived using Mortensen's approach [6] albeit modified to the geometric structure of the state space  $\text{SO}(2) \times \mathbb{R}$ . In order to introduce geometry in our algebra, first we define the gradients in (21).

In the following the gradients  $\nabla_X V(X, b, t) \in \text{TSO}(2)$  and  $\nabla_b V(X, b, t) \in \mathbb{R}$  are defined in terms of directional derivatives. For all  $\alpha, \gamma \in \mathbb{R}$

$$\begin{aligned}\mathcal{D}_X V(X, b, t) \circ X\alpha_\times &= \langle \nabla_X V(X, b, t), X\alpha_\times \rangle, \\ \mathcal{D}_b V(X, b, t)\gamma &= \langle \nabla_b V(X, b, t), \gamma \rangle,\end{aligned}\quad (22)$$

where the cross notation was defined in (3). The scalar inner product  $\langle \cdot, \cdot \rangle : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  is defined as

$$\langle \alpha, \gamma \rangle := \alpha\gamma, \quad (23)$$

and the left invariant inner product  $\langle \cdot, \cdot \rangle : \text{TSO}(2) \times \text{TSO}(2) \rightarrow \mathbb{R}$  is defined as

$$\langle X\alpha_\times, X\gamma_\times \rangle = \langle \alpha_\times, \gamma_\times \rangle := \text{trace}\left(\frac{1}{2}\alpha_\times^\top \gamma_\times\right) = \langle \alpha, \gamma \rangle. \quad (24)$$

Similarly the following second order directional derivatives are defined.

$$\begin{aligned}\mathcal{D}_X^2 V(X, b, t) \circ (X\alpha_\times, X\gamma_\times) &= \langle \nabla_X^2 V(X, b, t) \circ X\alpha_\times, X\gamma_\times \rangle \\ &= \langle X\alpha_\times, \nabla_X^2 V(X, b, t) \circ X\gamma_\times \rangle, \\ \mathcal{D}_b^2 V(X, b, t) \circ (\alpha, \gamma) &= \langle \nabla_b^2 V(X, b, t) \circ \alpha, \gamma \rangle \\ &= \langle \alpha, \nabla_b^2 V(X, b, t) \circ \gamma \rangle, \\ \mathcal{D}_b(\mathcal{D}_X V(X, b, t) \circ X\alpha_\times)\gamma &= \mathcal{D}_X(\mathcal{D}_b V(X, b, t) \circ \alpha) \circ X\gamma_\times = \\ &= \langle \nabla_b \nabla_X V(X, b, t) \circ \alpha, \gamma \rangle = \langle X\alpha_\times, \nabla_X \nabla_b V(X, b, t) \circ X\gamma_\times \rangle.\end{aligned}\quad (25)$$

Note that (25) shows that the second order derivatives are symmetric bi-linear mappings to  $\mathbb{R}$ , in the directions  $\alpha$  and  $\gamma$ . Therefore, we use the following parametric representations to serve as their values.

$$\begin{aligned}\mathcal{D}_X^2 V(X, b, t) \circ (X\alpha_\times, X\gamma_\times) &:= P'_1 \alpha\gamma, \\ \mathcal{D}_b^2 V(X, b, t) \circ (\alpha, \gamma) &:= P'_2 \alpha\gamma, \\ \mathcal{D}_b(\mathcal{D}_X V(X, b, t) \circ X\alpha_\times)\gamma &:= P'_{12} \alpha\gamma, \\ \mathcal{D}_X(\mathcal{D}_b V(X, b, t)\gamma) \circ X\alpha_\times &:= P'_{12} \alpha\gamma,\end{aligned}\quad (26)$$

where  $P'_1, P'_2, P'_{12} \in \mathbb{R}$ .

Now we can rewrite the final conditions (21) in terms of directional derivatives. For all  $\alpha, \gamma \in \mathbb{R}$ ,

$$\begin{aligned}\mathcal{D}_X V(X, b, t) \circ X\alpha_\times|_{X=\hat{X}(t), b=\hat{b}(t)} &= 0, \\ \mathcal{D}_b V(X, b, t)\gamma|_{X=\hat{X}(t), b=\hat{b}(t)} &= 0.\end{aligned}\quad (27)$$

Condition (27) holds for every time  $t$  and therefore the total time derivative of (27) satisfies

$$\begin{aligned}\frac{d}{dt}\{\mathcal{D}_X V(X, b, t) \circ X\alpha_\times\}|_{X=\hat{X}(t), b=\hat{b}(t)} &= 0, \\ \frac{d}{dt}\{\mathcal{D}_b V(X, b, t)\gamma\}|_{X=\hat{X}(t), b=\hat{b}(t)} &= 0.\end{aligned}\quad (28)$$

Applying the chain rule and changing the order of the derivatives gives

$$\begin{aligned}\{\mathcal{D}_X^2 V(X, b, t) \circ (\dot{X}, X\alpha_\times) + \mathcal{D}_X(\mathcal{D}_b V(X, b, t)\dot{b}) \circ X\alpha_\times + \\ \mathcal{D}_X(\nabla_t V(X, b, t)) \circ X\alpha_\times\}|_{X=\hat{X}(t), b=\hat{b}(t)} &= 0, \\ \{\mathcal{D}_b^2 V(X, b, t)\dot{b}\gamma + \mathcal{D}_b(\mathcal{D}_X V(X, b, t) \circ \dot{X})\gamma + \\ \mathcal{D}_b(\nabla_t V(X, b, t))\gamma\}|_{X=\hat{X}(t), b=\hat{b}(t)} &= 0.\end{aligned}\quad (29)$$

Next, the derivatives  $\mathcal{D}_X(\nabla_t V(X, b, t)) \circ X\alpha_\times$  and  $\mathcal{D}_b(\nabla_t V(X, b, t))\gamma$  are calculated by first replacing the time gradient  $\nabla_t V(X, b, t)$  with  $\mathcal{H}(X, b, \nabla_X V(X, b, t), \nabla_b V(X, b, t), t)$  using (20). First, using (17) and (22) yields

$$\begin{aligned}\mathcal{H}(X, b, \nabla_X V(X, b, t), \nabla_b V(X, b, t), t) = \\ \frac{1}{2}\left(-Q_\omega \mathcal{D}_X V(X, b, t) \circ \nabla_X V(X, b, t) \right. \\ \left. - Q_b \mathcal{D}_b V(X, b, t) \nabla_b V(X, b, t) + \|y - X^\top \hat{y}\|_{R^{-1}}^2 \right) \\ - \mathcal{D}_X V(X, b, t) \circ X(u - b)_\times.\end{aligned}\quad (30)$$

Therefore,

$$\begin{aligned}\mathcal{D}_X(\mathcal{H}(X, b, \nabla_X V(X, b, t), \nabla_b V(X, b, t), t)) \circ X\alpha_\times = \\ -Q_\omega \mathcal{D}_X^2 V(X, b, t) \circ (\nabla_X V(X, b, t), X\alpha_\times) \\ -Q_b \mathcal{D}_b(\mathcal{D}_X V(X, b, t) \circ X\alpha_\times) \nabla_b V(X, b, t) \\ + 2\alpha \text{vex} \mathbb{P}_a(R^{-1}(y - X^\top \hat{y}))^\top X \\ - \mathcal{D}_X^2 V(X, b, t) \circ (X(u - b)_\times, X\alpha_\times) \\ - \mathcal{D}_X V(X, b, t) \circ X\mathbb{P}_a(\alpha_\times(u - b)_\times), \\ \mathcal{D}_b(\mathcal{H}(X, b, \nabla_X V(X, b, t), \nabla_b V(X, b, t), t))\gamma = \\ -Q_\omega \mathcal{D}_X(\mathcal{D}_b V(X, b, t)\gamma) \circ \nabla_X V(X, b, t) \\ -Q_b \mathcal{D}_b^2 V(X, b, t) \nabla_b V(X, b, t)\gamma \\ - \mathcal{D}_X(\mathcal{D}_b V(X, b, t)\gamma) \circ X(u - b)_\times + \mathcal{D}_X V(X, b, t) \circ X\gamma_\times,\end{aligned}\quad (31)$$

where the anti-symmetric projection operator  $\mathbb{P}_a(\cdot) : \mathbb{R}^{2 \times 2} \rightarrow \mathfrak{so}(2)$  for all  $M \in \mathbb{R}^{2 \times 2}$  is defined as

$$\mathbb{P}_a(M) := \frac{1}{2}(M - M^\top). \quad (32)$$

Next, we substitute (31) in (29). Note that the first order derivatives, that appear in (31), evaluated at  $X = \hat{X}(t)$ ,  $b = \hat{b}(t)$ , yield zero. This is due to the final conditions (21) and (27). Replacing the second order derivatives from (26)

and canceling the arbitrary directions  $\alpha$  and  $\gamma$  from both sides of (29) yields

$$\begin{aligned} P_1 \text{vex}(\hat{X}^\top \hat{X}) + P_{12} \dot{\hat{b}} + 2 \text{vex} \mathbb{P}_a(R^{-1}(y - \hat{y})\hat{y}^\top) \\ - P_1(u - \hat{b}) = 0, \\ P_2 \dot{\hat{b}} + P_{12} \text{vex}(\hat{X}^\top \hat{X}) - P_{12}(u - \hat{b}) = 0, \end{aligned} \quad (33)$$

where

$$\begin{aligned} \hat{y} &:= \hat{X}^\top \hat{y}, \\ P_1 &:= P'_1|_{X=\hat{X}(t), b=\hat{b}(t)}, \\ P_2 &:= P'_2|_{X=\hat{X}(t), b=\hat{b}(t)}, \\ P_{12} &:= P'_{12}|_{X=\hat{X}(t), b=\hat{b}(t)}. \end{aligned} \quad (34)$$

Rearranging (33) yields the observer equations

$$\begin{aligned} \dot{\hat{X}} &= \hat{X} \left( u - \hat{b} - P_1^{-1} \left( l + P_{12} \dot{\hat{b}} \right) \right)_{\times}, \\ \dot{\hat{b}} &= -\frac{P_{12}}{P_2} \left( -u + \hat{b} + \text{vex}(\hat{X}^\top \hat{X}) \right), \end{aligned} \quad (35)$$

where  $l := 2 \text{vex} \mathbb{P}_a(R^{-1}(y - \hat{y})\hat{y}^\top)$ . The initial conditions  $\hat{X}(0) = I$  and  $\hat{b}(0) = 0$ , are obtained using (19) and (27).

The proposed observers (35) yield the exact dynamical equations that will update the value of the minimum-energy estimates  $\hat{X}(t)$  and  $\hat{b}(t)$  (that are the solutions to Problem 1), for every time  $t$ . Note that the two observers (35) are interconnected and use the current measurements  $u(t)$  and  $y(t)$  to update their estimates. The measurements are weighted by dynamic gains  $P_1$ ,  $P_2$  and  $P_{12}$  that are related to the value function defined in (18) through (26).

In order to implement these observers we also need dynamical equations that concurrently update the gains  $P_1$ ,  $P_2$  and  $P_{12}$ . This can be done using the definitions (26) and by calculating the total time derivatives

$$\begin{aligned} \dot{P}_1 \alpha \gamma &= \frac{d}{dt} \{ \mathcal{D}_X^2 V(X, b, t) \circ (X \alpha_{\times}, X \gamma_{\times}) \}_{X=\hat{X}(t), b=\hat{b}(t)}, \\ \dot{P}_2 \alpha \gamma &= \frac{d}{dt} \{ \mathcal{D}_b^2 V(X, b, t) \circ (\alpha, \gamma) \}_{X=\hat{X}(t), b=\hat{b}(t)}, \\ \dot{P}_{12} \alpha \gamma &= \frac{d}{dt} \{ \mathcal{D}_b(\mathcal{D}_X V(X, b, t) \circ X \alpha_{\times}) \gamma \}_{X=\hat{X}(t), b=\hat{b}(t)}. \end{aligned} \quad (36)$$

The calculation details for (36) are similar to our observer derivations and are omitted due to space limitations. It is worth noting that the right hand sides of (36) are going to depend on the third order derivatives of the value function. However, in this work, we assume that the third order derivatives of the value function are negligible. Subsequently, the following Riccati equations are obtained that update the gains of the observers on-line.

$$\begin{aligned} \dot{P}_1 &= -Q_\omega P_1^2 - Q_b P_{12}^2 - \hat{y}_1^\top S R^{-1} S \hat{y} + (y - \hat{y})^\top R^{-1} \hat{y}, \\ \dot{P}_2 &= -Q_b P_2^2 - Q_\omega P_{12}^2 + 2P_{12}, \\ \dot{P}_{12} &= -Q_\omega P_{12} P_1 - Q_b P_2 P_1 + P_1, \end{aligned} \quad (37)$$

where

$$S := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \quad (38)$$

The initial conditions  $P_1(0) = K_1^{-1}$ ,  $P_2(0) = K_2^{-1}$  and  $P_{12}(0) = 0$  are obtained using (19) and (26). Note that these

Riccati equations (37) provide a second-order approximation of the minimum-energy dynamics of the observer gains. We can continue deriving dynamics of the higher order derivatives of the value function using calculations similar to our previous workings. However, as was suggested by simulations in previous work [12], the third order derivatives of the value function are going to be small and there is no advantage in adding more computations to model their small effect. Therefore, we restrict this work to a second order approximation of the minimum-energy filter.

In summary the following filter is proposed

$$\begin{aligned} \dot{\hat{X}} &= \hat{X} \left( u - \hat{b} - \frac{P_2}{P_1 P_2 - P_{12}^2} l \right)_{\times}, \\ \dot{\hat{b}} &= \frac{P_{12}}{P_1 P_2 - P_{12}^2} l, \\ l &:= 2 \text{vex} \mathbb{P}_a(R^{-1}(y - \hat{y})\hat{y}^\top), \\ \dot{P}_1 &= -Q_\omega P_1^2 - Q_b P_{12}^2 - \hat{y}_1^\top S R^{-1} S \hat{y} + (y - \hat{y})^\top R^{-1} \hat{y}, \\ \dot{P}_2 &= -Q_b P_2^2 - Q_\omega P_{12}^2 + 2P_{12}, \\ \dot{P}_{12} &= -Q_\omega P_{12} P_1 - Q_b P_2 P_1 + P_1, \end{aligned} \quad (39)$$

where  $\hat{X}(0) = I$ ,  $\hat{b}(0) = 0$ ,  $P_1(0) = K_1^{-1}$ ,  $P_2(0) = K_2^{-1}$  and  $P_{12}(0) = 0$ . Note that the equations (35) have been rearranged into a new form in which the kinematics of the state and the bias estimates are in cascade form. This form is more useful for implementing the filter by discretization.

*Lemma 2:* If the measurement model (11) along with the cost (13) is considered instead of (9) and (12), the resulting filter equations are

$$\begin{aligned} \dot{\hat{\theta}} &= u - \hat{b} - \frac{P_2}{P_1 P_2 - P_{12}^2} \sin(y'' - \hat{\theta}), \quad \hat{\theta}(0) = 0, \\ \dot{\hat{b}} &= \frac{P_{12}}{P_1 P_2 - P_{12}^2} \sin(y'' - \hat{\theta}), \quad \hat{b}(0) = 0, \\ \dot{p}_1 &= -Q_\omega p_1^2 - Q_b p_{12}^2 + \cos(y'' - \hat{\theta}), \quad p_1(0) = K_1^{-1}, \\ \dot{p}_2 &= -Q_b p_2^2 - Q_\omega p_{12}^2 + 2p_{12}, \quad p_2(0) = K_2^{-1}, \\ \dot{p}_{12} &= -Q_\omega p_{12} p_1 - Q_b p_2 p_1 + p_1, \quad p_{12}(0) = 0. \end{aligned} \quad (40)$$

Moreover, equations (40) are equivalent to the proposed filter equations (39) if the matrix  $D$ , and hence  $R$ , is equal to the identity matrix.

The proof of Lemma 2 involves an argument, similar to the proof of Lemma 1, to show the first equation in (40). Furthermore, all the calculations that involve the measurements  $y$  need to be done using  $y''$  instead.

#### IV. SIMULATIONS

In this section we provide a simulation study of the estimation performance of the proposed filter (39), in the presence of measurement errors, initialization errors and bias in the angular velocity measurements. We have tested the proposed filter in many situations involving different levels of input, initialization, measurement errors and bias variations. The proposed filter proves to be robust producing consistent results in all these tests. In this section we demonstrate simulation results that are typical for the general performance of the proposed filter.

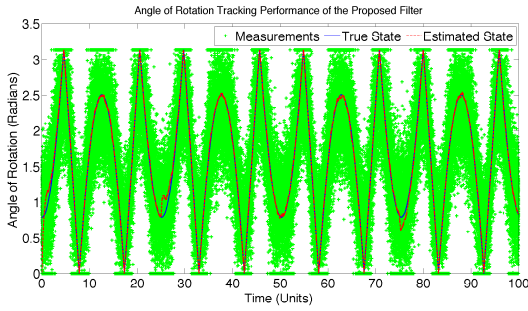


Fig. 2. The state tracking performance of the proposed filter is shown in the presence of bias and measurement errors. Note that the state measurements are shown in green with the “+” markers and are scattered around the true state trajectory. Note that there three transient periods, in the beginning, first quarter and the third quarter of the simulation time, before the filter converges. These are due to the initial bias error and the two bias jumps.

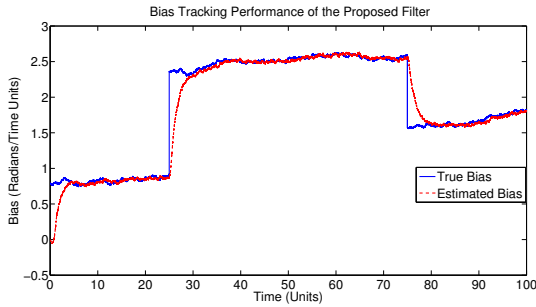


Fig. 3. The performance of the proposed filter in tracking the bias is shown. Note that the value of the true bias has two jumps at times 25 time units and 75 time units.

Consider the following simulation parameters. The system (1) is initialized to  $\theta_0 = 45^\circ$  and an initial bias of 45 degrees per time units is considered. The angular velocity measurement error  $v_\omega$  and the bias variation  $v_b$  are considered as zero mean random variables with standard deviations of  $B_\omega = 17$  degrees per time units and  $B_b = 52$  degrees per squared time units. The output measurement error  $w$  is set to a zero mean random vector with standard deviation of  $D = 0.4I$ . A sinusoidal angular velocity input  $\omega(t) = \sin(.25t)$  drives the system. We consider a situation in which there are three sudden jumps in the bias value.

As can be seen in Figure 2, after a short transient period, the proposed filter (39) is able to track the bias very well. Figure 3 shows the performance of the proposed filter in tracking the angle of rotation of the system state. As can be seen, the estimated angle converges to the true angle after a transient that is slightly longer than the bias estimation transient period.

## V. CONCLUSIONS

In this work we proposed a second-order nonlinear minimum-energy filter for a kinematics model of an object moving on the unit circle using angular velocity measurements that include an unknown small disturbance and slowly

time-varying bias. We model the system on the Lie group  $SO(2)$  to allow for using vector measurements rather than full state measurements and also to facilitate future extension of the method to the Lie group  $SO(3)$ . The proposed filter provides excellent on-line estimates of the bias and the state and is robust to different levels of bias and measurement and initialization errors as shown in the simulations.

## VI. ACKNOWLEDGMENTS

This research was supported by the Australian Research Council through Discovery Grant DP120100316 “Geometric observer theory for mechanical control systems”.

## REFERENCES

- [1] 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.
- [2] E. D. Sontag, *Mathematical control theory: deterministic finite dimensional systems*. Springer, 1998.
- [3] A. H. Jazwinski, *Stochastic processes and filtering theory*. Academic Press, April 1970.
- [4] J. C. Willems, “Deterministic least squares filtering,” *Journal of Econometrics*, vol. 118, no. 1-2, pp. 341–373, 2004.
- [5] A. Krener, “Minimum covariance, minimax and minimum energy linear estimators,” *Stochastic Control Theory and Stochastic Differential systems*, pp. 490–495, 1979.
- [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. P. Aguiar and J. P. Hespanha, “Minimum-energy state estimation for systems with perspective outputs and state constraints,” in *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, vol. 5. IEEE, 2003, pp. 4908–4913.
- [10] P. Coote, J. Trumpf, R. Mahony, and J. C. Willems, “Near-optimal deterministic filtering on the unit circle,” in *Proceedings of the 48th IEEE Conference on Decision and Control*, 2009, pp. 5490–5495.
- [11] 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.
- [12] 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.
- [13] —, “A second order minimum-energy filter on the special orthogonal group,” in *In Proceedings of the 2012 American Control Conference (ACC)*, to appear. Accepted for publication 28 January, 2012.
- [14] —, “Minimum-energy pose filtering on the special Euclidean group,” in *the 20th International Symposium on Mathematical Theory of Networks and Systems (MTNS)*, to appear. Accepted for publication 16 March.
- [15] S. Bonnabel and E. Salaün, “Design and prototyping of a low-cost vehicle localization system with guaranteed convergence properties,” *Control Engineering Practice*, 2011.
- [16] R. Mahony, T. Hamel, and J. -M. Pfimlin, “Nonlinear complementary filters on the special orthogonal group,” *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1203–1218, 2008.
- [17] M. Athans and P. Falb, *Optimal control: an introduction to the theory and its applications*. McGraw-Hill, 1966.