Automatica 125 (2021) 109464

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Extended invariant-EKF designs for state and additive disturbance estimation☆

Kevin Coleman [a], He Bai [a,*], Clark N. Taylor [b]

[a] School of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, OK, 74078, USA
[b] Department of Electrical and Computer Engineering, Air Force Institute of Technology, WPAFB, OH, 45433, USA

ARTICLE INFO

Article history:
Received 11 March 2020
Received in revised form 5 October 2020
Accepted 11 December 2020
Available online 15 January 2021

Keywords:
Invariant systems
Extended Kalman filters
Mobile robots
Disturbance estimation

ABSTRACT

受动态加性扰动影响的不变非线性系统的估计问题

In this paper, we consider an estimation problem of invariant nonlinear systems subject to dynamic additive disturbances. We identify two sets of sufficient conditions that preserve the invariant properties of the systems under the disturbances. We apply the conditions to a unicycle model under linear dynamic disturbances and design two different Invariant Extended Kalman filters (IEKFs). Both IEKFs estimate the state of the unicycle and the disturbances based on position measurements. We also propose a correction to the IEKF covariances to better represent uncertainties in the invariant frame. The benefit of including the covariance correction and the performances of the two IEKF designs are demonstrated through Monte-Carlo simulations.

1. Introduction

Estimation and filtering of nonlinear systems is an important problem in research and in industry. There are several filters capable of dealing with nonlinear systems, such as the extended Kalman filter (EKF) (Gelb, 1974), unscented Kalman filter (UKF) (Julier & Uhlmann, 2004) and particle filters (Arulampalam, Maskell, Gordon, & Clapp, 2002). When applied to robotic applications, these filters provide simple, 'off-the-shelf' solutions. However, they do not take advantage of properties present in robotic dynamics, such as symmetries. There has been much interest lately in designing observers that can leverage symmetries of certain nonlinear dynamics to improve estimation performance. These are known more generally as symmetry preserving observers.

对称性

The theory behind symmetry preserving observers is given in Bonnabel, Martin, and Rouchon (2008). When the EKF equations are used to compute the gain matrix of a symmetry preserving observer, it is referred to as an invariant EKF (IEKF) (Bonnabel, Martin, & Salaün, 2009). More recently, the IEKF has gained attention as a tool well suited for applications in localization of mobile robots and sensor fusion for navigation of unmanned aerial vehicles. In Bonnable et al. (2009) the authors apply the IEKF to the problem of estimating the attitude and velocity of an aircraft using GPS velocity and measurements from on board gyroscopes and accelerometers. Martin and Salaun (2008) designs a symmetry preserving observer for fusing measurements from several sensors in different coordinate frames for attitude heading systems for aircraft. Barczyk, Bonnabel, Deschaud, and Goulette (2015) develops an IEKF for use with a low cost Kinect depth camera to perform Scan-Matching aided localization of a mobile ground robot. They compare the performance of the IEKF to the Multiplicative EKF (MEKF) and show that the IEKF has better performance. In De Silva, Mann, and Gosine (2014) the authors apply the IEKF to the problem of relative localization for multiple mobile robots. Wu, Zhang, Su, Huang, and Dissanayake (2017) uses the IEKF in a visual inertial navigation system. In Zhang, Wu, Song, Huang, and Dissanayake (2017) the authors show that an IEKF based SLAM (simultaneous localization and mapping) algorithm has better consistency and convergence properties over other EKF based SLAM techniques. Trumpf, Mahony, and Hamel (2018) provides checkable sufficient conditions on kinematic systems with symmetries to determine whether considered systems can be lifted to invariant systems on symmetry groups. More recently, Barrau and Bonnabel (2017) proposes a matrix Lie group framework for the IEKF and show that it possesses local stability properties.

In this paper, we consider invariant systems subject to additive dynamic disturbances and extend IEKF designs to simultaneously estimate the state and disturbances. The first contribution of this paper is the identification of two scenarios under which the

☆ The work of the first two authors was supported in part by the National Science Foundation (NSF), USA under award number 1925147 and AFOSR, USA Grant FA8650-15-D-1845. The material in this paper was partially presented at the 2020 American Control Conference, July 1–3, 2020, Denver, CO, USA. This paper was recommended for publication in revised form by Associate Editor Gianluigi Pillonetto under the direction of Editor Torsten Söderström.

* Corresponding author.
E-mail addresses: kevin.coleman10@okstate.edu (K. Coleman), he.bai@okstate.edu (H. Bai), Clark.Taylor@afit.edu (C.N. Taylor).

https://doi.org/10.1016/j.automatica.2020.109464

extended state and disturbance system remains invariant. In the first scenario, we ensure the invariant dynamics by transforming the output matrix of the disturbance model while keeping the disturbance dynamics unchanged. In the second scenario, we prove that the extended system remains invariant if certain 'commutation' conditions of the output matrix and the disturbance dynamics are satisfied.

The second contribution of this paper is the application of the extended IEKF designs to a unicycle model subject to disturbances modeled as the output of a linear time-invariant system. A unicycle robotic model is widely used to model the kinematics of a differential drive mobile vehicle, underwater vehicle motion (Petrich, Woolsey, & Stilwell, 2009), and the simplified kinematics of a fixed wing aerial vehicle in planar flight (Beard & McLain, 2012). Furthermore, some applications include estimating the states of these types of vehicles for the purpose of localization (Betke & Gurvits, 1997), trajectory tracking (Kolmanovsky & Harris McClamroch, 1995), or flow field reconstruction (Bai, 2018; Palanthandalam-Madapusi, Girard, & Bernstein, 2008). The linear disturbance models can represent uniform flow and sinusoidal wave disturbances with known frequencies.

We design two extended IEKFs for the unicycle to estimate both its heading and disturbance based on position information. The two designs are based on the two identified scenarios where the extended dynamics are invariant. We show that the first design is applicable to general linear dynamic disturbances while the second design is restricted to a class of systems satisfying 'rotational invariance' conditions on the dynamics and the output matrices. Our simulation examples demonstrate that when applicable, the second design yields better transient performance than the first design. We also establish connections of the second design with the matrix IEKF formulation in Barrau and Bonnabel (2017). Indeed, for the unicycle example, the 'rotational invariance' conditions are the same conditions that ensure the 'group affine condition' in Barrau and Bonnabel (2017).

As the third contribution of the paper, we improve the IEKF performance by better characterizing the statistics of the sensor noise in the invariant frame. In particular, we improve the result in Barrau and Bonnabel (2017) and derive a first order approximation for the covariance of the transformed noise. Using Monte Carlo simulations, we show that the introduction of the first order approximation improves the performance of the IEKF, particularly for non-isotropic sensor noise.

Compared with our preliminary work (Coleman, Bai, & Taylor, 2020) that considered only the IEKF design in Section 4, this paper provides new and significant contributions, including the generalized invariant conditions and illustrative examples in Section 2, another IEKF design in Section 5, and a numerical comparison between EKF and the two IEKF designs. The second IEKF design is shown to perform better than the design in Coleman et al. (2020).

The rest of the paper is organized as follows. In Section 2 we develop the theory for invariant systems to remain invariant in the presence of dynamic disturbances. In Section 3 we pose an estimation problem of interest. In Sections 4 and 5, we propose two different IEKF designs for the problem of interest. We derive correction terms for the covariances of the general IEKF in Section 6. Simulation results are discussed in Section 7. Conclusions and future work are presented in Section 8.

2. Invariant systems with disturbances

Consider the following nonlinear system

$$\begin{aligned} \dot{x} &= f(x,u) \\ y &= h(x,u) \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^q$ is the control, and $y \in \mathbb{R}^p$ is the measured output. Let G be an n dimensional Lie group. Given $g \in G$, define local transformations on the state and input as $\varphi_g(x)$ and $\psi_g(u)$, respectively. By definition, the system (1) is *invariant* with respect to G if $f(\varphi_g(x), \psi_g(u)) = \frac{\partial}{\partial x}\varphi_g(x) f(x,u)$ for all g, x and u. The output is said to be *equivariant* with respect to G if there exists a transformation of the output, $\varrho_g(y)$, such that $h(\varphi_g(x), \psi_g(u)) = \varrho_g(h(x,u))$ (Bonnabel et al., 2008).

Assumption 1. The system (1) is invariant with respect to the transformations $\varphi_g(x)$ and $\psi_g(u)$.

Consider (1) cascaded with nonlinear dynamic disturbances d

$$\begin{aligned} \dot{x} &= f(x,u) + Cd \\ \dot{d} &= J(d) \\ y &= h(x,u) \end{aligned} \tag{2}$$

where $d \in \mathbb{R}^m$, $C \in \mathbb{R}^{n\times m}$ and $J(\cdot)$ is a smooth nonlinear function. Note that we choose to write the disturbances affecting the states as Cd instead of an arbitrary nonlinear function $g(d)$. The nonlinear disturbance model ($\dot{d} = J(d)$, $z = Cd$) is general since a nonlinear dynamical system with nonlinear outputs of full row-rank can be converted to a system with linear outputs through a nonlinear coordinate transformation, e.g., based on its normal forms (Khalil, 2002, Section 13.2) (Schwartz, Isidori, & Tarn, 1999).

Define two transformations $\beta_g(C) : G \times \mathbb{R}^{n\times m} \mapsto \mathbb{R}^{n\times m}$ and $\xi_g(d) : G \times \mathbb{R}^m \mapsto \mathbb{R}^m$. We next derive sufficient conditions on $\beta_g(C)$ and $\xi_g(d)$ such that the cascaded system (2) remains invariant under the group actions $(\varphi_g(x), \psi_g(u), \beta_g(C), \xi_g(d))$. We do this by proposing two different approaches outlined in Propositions 3 and 4. Proposition 3 takes $\xi_g(\cdot)$ to be the identity operator and examines invariance conditions on $\beta_g(\cdot)$. Proposition 4 takes $\beta_g(\cdot)$ to be the identity operator and examines invariance conditions on $\xi_g(\cdot)$. Our results rely on the following assumption.

Assumption 2. $\varphi_g(x)$ and $\xi_g(d)$ are linear in x and d, respectively.

Assumption 2 allows us to define

$$\alpha(g) = \frac{\partial}{\partial x}\varphi_g(x) \in \mathbb{R}^{n\times n}, \ \kappa(g) = \frac{\partial}{\partial d}\xi_g(d) \in \mathbb{R}^{m\times m}. \tag{3}$$

Proposition 3. *Suppose that Assumptions 1 and 2 hold. Then (2) is invariant with respect to G if $\beta_g(C)$ and $\xi_g(d)$ are selected as $\beta_g(C) = \alpha(g)C$ and $\xi_g(d) = d$, respectively.*

Proof. It follows from the definition of invariance that (2) is invariant if the following two equations hold:

$$f(\varphi_g(x), \psi_g(u)) + \beta_g(C)\xi_g(d) = \frac{\partial}{\partial x}\varphi_g(x)\left(f(x,u) + Cd\right) \tag{4}$$

$$J(\xi_g(d)) = \frac{\partial}{\partial d}\xi_g(d) J(d). \tag{5}$$

Let $\xi_g(d) = d$. Then (5) is trivially satisfied. Since $f(x,u)$ is invariant with respect to G, (4) reduces to

$$\beta_g(C)d = \frac{\partial}{\partial x}\varphi_g(x)Cd = \alpha(g)Cd. \tag{6}$$

Since $\beta_g(C)$ can only be a function of g, it follows from (6) that $\frac{\partial}{\partial x}\varphi_g(x)$ cannot be a function of x, which means that $\varphi_g(x)$ is linear in x. Therefore, invariance with respect to G is preserved by leaving the disturbances (d) unchanged and transforming C with a transformation defined by $\beta_g(C) = \alpha(g)C$. □

From the proof, we see that invariance can be preserved in the augmented system (2) by performing a transformation on the system parameter C, instead of on the disturbances d. In Proposition 4, we preserve the invariance property by performing a transformation directly on d instead of on C.

Proposition 4. *Suppose that Assumptions 1 and 2 hold. Then (2) is invariant with respect to G if $\beta_g(C)$ and $\xi_g(d)$ satisfy $\beta_g(C) = C$, $C\kappa(g) = \alpha(g)C$ and $J(\kappa(g)d) = \kappa(g)J(d)$.*

Proof. Let $\beta_g(C) = C$. From Assumption 1 it follows that $\varphi_g(x) = \alpha(g)x$ and $\xi_g(d) = \kappa(g)d$. Then (4) reduces to $C\kappa(g)d = \alpha(g)Cd$ which implies that $C\kappa(g) = \alpha(g)C$ must be satisfied. The second equation (5) becomes $J(\kappa(g)d) = \kappa(g)J(d)$. □

Propositions 3 and 4 provide two approaches to defining transformations that preserve invariance of a nonlinear system when state dynamic disturbances are included. Both approaches assume that the original group action is linear with respect to the states. Motivated by internal model control and disturbance rejection literature (see e.g., Isidori & Byrnes, 1990; Isidori, Marconi, & Serrani, 2012), we next focus on disturbances resulting from a linear dynamic model, i.e., $J(d) = Ad$, $A \in \mathbb{R}^{m\times m}$. In this case, the conditions in Proposition 4 become $C\kappa(g) = \alpha(g)C$ and $A\kappa(g) = \kappa(g)A$, the second signifying that the Lie bracket of the vector fields Ad and $\xi_g(d)$ must be zero.

Note that Proposition 4 does not provide specific transformations of $\kappa(g)$ and $\alpha(g)$ to ensure invariance. In Proposition 5, we consider a special case where the disturbances affecting the individual elements of x share the same linear generating model $(\mathcal{A}, \mathcal{C})$. We provide explicit expressions of $\kappa(g)$ and $\alpha(g)$ that ensure the invariance conditions in Proposition 4. We assume that the first s elements of x are affected by the disturbance, $0 < s \le n$. Denote by I_m the m-dimensional identity matrix and by $\otimes$ the Kronecker product.

Proposition 5. *Suppose that Assumptions 1 and 2 hold. Suppose that $J(d) = Ad = (I_s \otimes \mathcal{A})d$ and $C = \left(I_s \otimes \mathcal{C}^\top, 0_{m\times(n-s)}\right)^\top$, $0 < s \le n$, where $\mathcal{A} \in \mathbb{R}^{r\times r}$, $\mathcal{C} \in \mathbb{R}^{1\times r}$ and $r \cdot s = m$. Assume that $\alpha(g)$ satisfies*

$$\alpha(g) = \begin{pmatrix} \alpha_1(g) & \alpha_2(g) \\ 0 & \alpha_3(g) \end{pmatrix}, \tag{7}$$

where $\alpha_1(g) \in \mathbb{R}^{s\times s}$, $\alpha_2(g) \in \mathbb{R}^{s\times(n-s)}$ and $\alpha_3(g) \in \mathbb{R}^{(n-s)\times(n-s)}$. Then (2) is invariant with respect to G by choosing $\beta_g(C) = C$ and $\xi_g(d) = \kappa(g)d$, where

$$\kappa(g) = \alpha_1(g) \otimes I_r. \tag{8}$$

Proof. Using Kronecker product properties and the forms of C and $\alpha(g)$, we have

$$\alpha(g)C = \begin{pmatrix} \alpha_1(g)(I_s \otimes \mathcal{C}) \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_1(g) \otimes \mathcal{C} \\ 0 \end{pmatrix}, \tag{9}$$

which can be further rewritten as

$$\begin{pmatrix} \alpha_1(g) \otimes \mathcal{C} \\ 0 \end{pmatrix} = \begin{pmatrix} (I_s \otimes \mathcal{C})(\alpha_1(g) \otimes I_r) \\ 0 \end{pmatrix} = C\kappa(g). \tag{10}$$

Similarly, we verify that $(I_s \otimes \mathcal{A})\kappa(g) = (I_s \otimes \mathcal{A})(\alpha_1(g) \otimes I_r) = \alpha_1(g) \otimes \mathcal{A} = \kappa(g)(I_s \otimes \mathcal{A})$. Thus, the invariant conditions in Proposition 4 are satisfied with the choice of $\kappa(g)$ in (8). □

In Proposition 5, the disturbance affecting each element of x is generated from the same dynamic system specified by $(\mathcal{A}, \mathcal{C})$ with possibly different initial conditions. When $s = n$, Proposition 5 holds for any $\alpha(g)$. When $s < n$, $\alpha(g)$ needs to satisfy (7) to ensure invariance. The condition (7) means that after the transformation $\varphi_g(x)$, the last $n-s$ elements of x remain unaffected by the disturbances.

2.1. Illustrative examples

Assumption 2 holds in a number of applications, including the unicycle example in Section 3, chemical reactor dynamics (Bonnabel et al., 2008), and attitude dynamics (Phogat & Chang, 2020). We next briefly discuss how our theoretical results can be applied to the chemical reactor dynamics and the attitude dynamics and then focus on demonstrating invariant EKF designs for the unicycle example.

The chemical reactor dynamics in Bonnabel et al. (2008) are given by

$$\begin{aligned} \frac{d}{dt}X^{in} &= 0 \\ \frac{d}{dt}X &= D(t)(X^{in} - X) - k\exp\left(-\frac{E_A}{RT}\right)X \\ \frac{d}{dt}T &= D(t)(T^{in}(t) - T) + c\exp\left(-\frac{E_A}{RT}\right)X + v(t) \end{aligned} \tag{11}$$

where X and T are the reactor composition and temperature respectively and X^{in} is the inlet composition. E_A, R, k and c are known positive constant parameters. $D(t)$, $T^{in}(t)$ and $v(t)$ are known functions of time and $D(t) \ge 0$. The dynamics in (11) are invariant with respect to

$$\varphi_g(x) = \begin{pmatrix} gX^{in} \\ gX \\ T \end{pmatrix} \qquad \psi_g(u) = \begin{pmatrix} c/g \\ D(t) \\ T^{in}(t) \\ v(T) \end{pmatrix} \tag{12}$$

where $g \in \mathbb{R}_+$. This transformation represents a positive scaling of the reactor and inlet compositions. Note that $\varphi(\cdot)$ is linear in the state X^{in}, X, T. Thus, Assumptions 1 and 2 are satisfied.

When the disturbance Cd is added to (11), Proposition 3 provides the transformation $\beta_g(C) = \alpha(g)C$ where

$$\alpha(g) = \begin{bmatrix} g & 0 & 0 \\ 0 & g & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{13}$$

This transformation scales the rows of C corresponding to the transformed states. A second approach is to take $\beta_g(C) = C$ and $\xi_g(d) = \kappa(g)d$ where $\kappa(g) = (\alpha(g) \otimes I_{\frac{m}{3}})$ with $\alpha(g)$ defined in (13). This transformation satisfies Proposition 4 if (*i*) $C(\alpha(g) \otimes I_{\frac{m}{3}}) = \alpha(g)C$ and (*ii*) $J\left((\alpha(g) \otimes I_{\frac{m}{3}})d\right) = (\alpha(g) \otimes I_{\frac{m}{3}})J(d)$ are both satisfied. If $J(d) = Ad$ and (A, C) satisfies $A = I_3 \otimes \mathcal{A}$ and $C = I_3 \otimes \mathcal{C}$, Proposition 5 provides the transformation $\xi_g(d) = (\alpha(g) \otimes I_{\frac{m}{3}})d$ with $\alpha(g)$ in (13) that satisfies (*i*) and (*ii*), ensuring invariance of the disturbed system.

We next consider the undisturbed attitude dynamics with a unit quaternion $q = (q_0, q_v^\top)^\top$ representation, which is given by

$$\dot{q} = \begin{pmatrix} -\frac{1}{2}q_v^\top \\ \frac{1}{2}(q_0 I + \widehat{q_v}) \end{pmatrix}\omega, \quad J\dot{\omega} + \widehat{\omega}J\omega = \tau, \tag{14}$$

where $\widehat{w}v = w \times v$ and ω, J, and τ are the angular velocity, moment of inertia and control torque in the body frame. The dynamics in (14) is invariant with respect to

$$\varphi_g(q, \omega) = \begin{pmatrix} (q_0, (R_g q_v)^\top)^\top \\ R_g\omega \end{pmatrix} \qquad \psi_g(u) = \begin{pmatrix} R_g J R_g^\top \\ R_g\tau \end{pmatrix} \tag{15}$$

for any $R_g \in SO(3)$. Thus, $\varphi(\cdot)$ is linear in the state q and ω, thereby satisfying Assumptions 1 and 2.

When the disturbance Cd is added to the right hand side of the ω dynamics in (14), applying Proposition 3 results in the transformation $\beta_g(C) = \alpha(g)C$, where $\alpha(g) = R_g$, which corresponds to a rotation of the columns of C. Another choice of transformations on the disturbances can be $\xi_g(d) = \kappa(g)d$, where $\kappa(g) = (R_g \otimes I_{\frac{m}{3}})$. By Proposition 4, invariance of the disturbed system is preserved with the transformation $\xi_g(d)$ if (*i*) $C(R_g \otimes I_{\frac{m}{3}}) = R_g C$ and (*ii*) $J((R_g \otimes I_{\frac{m}{3}})d) = (R_g \otimes I_{\frac{m}{3}})J(d)$. These two conditions are an additional requirement not needed for Proposition 3. The example in Section 5 demonstrates a similar requirement in $SO(2)$. When $J(d) = Ad$, Proposition 5 ensures (*i*) and (*ii*) if the pair (A, C) satisfies $A = I_3 \otimes \mathcal{A}$ and $C = I_3 \otimes \mathcal{C}$.

In the rest of this paper, we will demonstrate and compare invariant EKF designs based on Propositions 3 and 4 for the unicycle example in Section 3. We will also introduce better characterizations on noise covariance matrices that lead to improved estimation performance.

3. Unicycle model under linear disturbances

Consider a unicycle robot subject to velocity disturbances. The kinematic model of the robot is given by

$$\begin{aligned} \dot{x} &= v\cos\theta + C_x d \\ \dot{y} &= v\sin\theta + C_y d \\ \dot{\theta} &= \omega, \end{aligned} \tag{16}$$

where (x, y) is the position of the robot, θ is the heading, v is the linear velocity and ω is the turning rate. We assume that $(C_x d, C_y d)$ are outputs from a linear system given by

$$\dot{d} = Ad \tag{17}$$

where $d \in \mathbb{R}^{m\times 1}$, $A \in \mathbb{R}^{m\times m}$ and $C_x, C_y \in \mathbb{R}^{1\times m}$. The matrices A, C_x, and C_y are assumed known and constant. For example, $C_x d$ and $C_y d$ can represent constant disturbances and sinusoidal disturbances with known frequencies.

The robot is equipped with a positioning device, such as a GPS or a suite of range and bearing sensors, measuring its position (x, y). The position measurement can be in a global frame or with respect to a known landmark. In the latter case, without loss of generality, we assume that the landmark is at the origin. Then (x, y) represents the relative position between the robot and the landmark. The measurement equation of the system is

$$Y = [x \;\; y]^\top. \tag{18}$$

In Bonnabel et al. (2008), it was shown that the undisturbed form of (16) is invariant with respect to actions of the special Euclidean group $SE(2)$, the group of translations and rotations in 2 dimensions. With the additive disturbances, our objective is to design an IEKF to estimate both the states and the disturbances. In the following two sections, we design two invariant IEKFs that correspond with Propositions 3 and 4 introduced in Section 2.

4. IEKF design 1

Let G be the group $SE(2)$. Any element g of G can be represented by (x_g, y_g, θ_g). Let $X = [x, y, \theta]^\top$ and define two transformations as

$$\varphi_g(X) = \begin{pmatrix} x\cos\theta_g - y\sin\theta_g + x_g \\ x\sin\theta_g + y\cos\theta_g + y_g \\ \theta + \theta_g \end{pmatrix} \tag{19}$$

$$\xi_g(d) = d, \tag{20}$$

where (x_g, y_g, θ_g) represent the parameters that define the group action on the state space and (x, y, θ) represent the components of the original non-transformed state. Notice that the disturbances d remain unchanged by the transformation. We now use the result of Proposition 3 to find the transformations on C_x and C_y. Since the disturbances do not affect θ, we concatenate C_x and C_y with a row of zeros and define $\beta_g(\cdot)$ as

$$\begin{aligned} \beta_g\left(\begin{bmatrix} C_x \\ C_y \\ 0 \end{bmatrix}\right) &= \frac{\partial}{\partial X}\varphi_g(X)\begin{bmatrix} C_x \\ C_y \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta_g & -\sin\theta_g & 0 \\ \sin\theta_g & \cos\theta_g & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} C_x \\ C_y \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} C_x\cos\theta_g - C_y\sin\theta_g \\ C_x\sin\theta_g + C_y\cos\theta_g \\ 0 \end{bmatrix}. \end{aligned} \tag{21}$$

Let $U = (v, \omega, A)$ and define a transformation of U as $\psi_g(U) = U$.

Corollary 6. *The dynamics in (16) and (17) is invariant with respect to $SE(2)$.*

Proof. As shown in Bonnabel et al. (2008), the undisturbed system (16) without (d_x, d_y) is invariant with respect to $SE(2)$. With the transformations defined in (20) and (21), it follows from Proposition 3 that the augmented system in (16)–(17) is invariant with respect to $SE(2)$. □

Following the methods outlined in Bonnabel et al. (2008), $\varphi_g(X)$ can be split into $\varphi_g^a(X)$ and $\varphi_g^b(X)$ such that $\varphi_g^a(X)$ is invertible with respect to g. Setting $\varphi_g^a(X) = 0$ gives the normalization equation

$$\begin{pmatrix} x_g \\ y_g \\ \theta_g \end{pmatrix} = \gamma\begin{pmatrix} x \\ y \\ \theta \end{pmatrix} = \begin{pmatrix} -x\cos\theta - y\sin\theta \\ x\sin\theta - y\cos\theta \\ -\theta \end{pmatrix} \tag{22}$$

where γ is called the moving frame, which is a mapping from the state space to the group G. For more details on the moving frame refer to Olver (1999). The invariants are

$$\begin{aligned} &I(\hat{X}, U) = \left(\varphi^b_{\gamma(\hat{X})}\left(\hat{X}\right), \psi_{\gamma(\hat{X})}(U)\right) = \\ &\left(v, \omega, C_x\cos\hat{\theta} + C_y\sin\hat{\theta}, -C_x\sin\hat{\theta} + C_y\cos\hat{\theta}, A\right), \end{aligned} \tag{23}$$

where $\hat{X}$ is the estimate of X. The invariant output error is given by

$$\begin{aligned} E &= \varrho_g(\hat{x}, \hat{y}) - \varrho_g(x, y) \\ &= \begin{pmatrix} \hat{x}\cos\theta_g - \hat{y}\sin\theta_g + x_g - x\cos\theta_g + y\sin\theta_g - x_g \\ \hat{x}\sin\theta_g + \hat{y}\cos\theta_g + y_g - x\sin\theta_g - y\cos\theta_g - y_g \end{pmatrix} \\ &= T(\hat{\theta})\begin{bmatrix} \hat{x} - x \\ \hat{y} - y \end{bmatrix}, \end{aligned} \tag{24}$$

where

$$T(\hat{\theta}) = \begin{bmatrix} \cos\hat{\theta} & \sin\hat{\theta} \\ -\sin\hat{\theta} & \cos\hat{\theta} \end{bmatrix}. \tag{25}$$

The invariant frame is given by

$$W(\hat{\theta}) = \begin{bmatrix} T(\hat{\theta})^\top & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & I_m \end{bmatrix}. \tag{26}$$

Thus, the observer equation has the following form

$$\dot{\hat{X}} = f(\hat{X}) + W(\hat{\theta}) \cdot L \cdot T(\hat{\theta})\left(Y - \hat{Y}\right), \tag{27}$$

where L is a gain matrix to be designed. For notation convenience, we let

$$L = \begin{bmatrix} L_{11} & L_{21} & L_{31} & L_{d1}^\top \\ L_{21} & L_{22} & L_{32} & L_{d2}^\top \end{bmatrix}^\top, \tag{28}$$

where L_{ij} are scalars for $i = 1, 2, 3, j = 1, 2$, and $L_{d1}, L_{d2} \in \mathbb{R}^{m\times 1}$. The invariant state error is given by

$$\begin{aligned}\sigma(\hat{X}, X) &= \varphi_{\gamma(\hat{X})}(X) - \varphi_{\gamma(\hat{X})}(\hat{X}) \\ &= W(\hat{\theta})^\top \begin{bmatrix} x-\hat{x} \\ y-\hat{y} \\ \theta-\hat{\theta} \\ d-\hat{d}\end{bmatrix}. \end{aligned} \tag{29}$$

To find the invariant error dynamics, we differentiate (29) and obtain

$$\begin{aligned}\dot{\sigma} &= W(\hat{\theta})^\top \left[\begin{pmatrix} v\cos\theta + C_x d - v\cos\hat{\theta} - C_x\hat{d} \\ v\sin\theta + C_y d - v\sin\hat{\theta} - C_y\hat{d} \\ 0 \\ Ad - A\hat{d}\end{pmatrix} \right. \\ &\left. - W(\hat{\theta}) L \begin{bmatrix} \cos\hat{\theta} & \sin\hat{\theta} \\ -\sin\hat{\theta} & \cos\hat{\theta}\end{bmatrix} \begin{bmatrix} x-\hat{x} \\ y-\hat{y}\end{bmatrix}\right] + \begin{bmatrix}\dot{\hat{\theta}}\sigma_y \\ -\dot{\hat{\theta}}\sigma_x \\ 0 \\ 0\end{bmatrix}, \end{aligned} \tag{30}$$

which yields

$$\begin{aligned}\dot{\sigma}_x &= v\,(\cos\sigma_\theta - 1) + \omega\sigma_y + \left(C_x\cos\hat{\theta} + C_y\sin\hat{\theta}\right)\sigma_d \\ &\quad + L_{11}\sigma_x + L_{12}\sigma_y + L_{31}\sigma_x\sigma_y + L_{32}\sigma_y^2 \\ \dot{\sigma}_y &= v\sin\sigma_\theta - \omega\sigma_x + \left(-C_x\sin\hat{\theta} + C_y\cos\hat{\theta}\right)\sigma_d \\ &\quad + L_{21}\sigma_x + L_{22}\sigma_y - L_{31}\sigma_x^2 - L_{32}\sigma_x\sigma_y \\ \dot{\sigma}_\theta &= L_{31}\sigma_x + L_{32}\sigma_y \\ \dot{\sigma}_d &= A\sigma_d + L_{d1}\sigma_x + L_{d2}\sigma_y. \end{aligned} \tag{31}$$

Note that the invariant error dynamics (31) depend only on σ and the invariants $I(\hat{X}, U)$ in (23).

Linearizing (31) around $\sigma = 0$ yields the state matrix needed for implementing the IEKF at time step k:

$$A_k = \begin{bmatrix} 0 & \omega_k & 0 & C_x\cos\hat{\theta}_k + C_y\sin\hat{\theta}_k \\ -\omega_k & 0 & v_k & -C_x\sin\hat{\theta}_k + C_y\cos\hat{\theta}_k \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & A\end{bmatrix}. \tag{32}$$

The A_k matrix is used in the IEKF algorithm to propagate the state covariance matrix. Before presenting the IEKF algorithm, we illustrate a second IEKF design for (16)–(17) based on Proposition 4 in the next section.

5. IEKF design 2

Compared with the design in Section 4, this design assumes the same state transformation $\varphi_g(X)$ in (19) and introduces transformations on the disturbances. We define

$$\beta_g\left(\begin{bmatrix} C_x \\ C_y \\ 0\end{bmatrix}\right) = \begin{bmatrix} C_x \\ C_y \\ 0\end{bmatrix} \tag{33}$$

$$\xi_g(d) = \left(T(\theta_g)^\top \otimes I_{\frac{m}{2}}\right) d \tag{34}$$

where $T(\cdot)$ is given in (25). Notice that $\xi_g(d)$ is linear in d. Applying Proposition 4, we note that to preserve the invariance property, we need

$$\begin{bmatrix} C_x \\ C_y \\ 0\end{bmatrix}\left(T(\theta_g)^\top \otimes I_{\frac{m}{2}}\right) = \begin{bmatrix} \cos\theta_g & -\sin\theta_g & 0 \\ \sin\theta_g & \cos\theta_g & 0 \\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} C_x \\ C_y \\ 0\end{bmatrix} \tag{35}$$

and

$$A\left(T(\theta_g)^\top \otimes I_{\frac{m}{2}}\right) = \left(T(\theta_g)^\top \otimes I_{\frac{m}{2}}\right) A. \tag{36}$$

Proposition 7. *Eqs. (35) and (36) are satisfied if A and $[C_x^\top\ C_y^\top]^\top$ satisfy*

$$A = \begin{bmatrix} \mathcal{M} & \mathcal{N} \\ -\mathcal{N} & \mathcal{M}\end{bmatrix} \text{ and } \begin{bmatrix} C_x \\ C_y\end{bmatrix} = \begin{bmatrix} \mathcal{D} & \mathcal{E} \\ -\mathcal{E} & \mathcal{D}\end{bmatrix}, \tag{37}$$

where $\mathcal{M}, \mathcal{N} \in \mathbb{R}^{\frac{m}{2}\times\frac{m}{2}}$ and $\mathcal{D}, \mathcal{E} \in \mathbb{R}^{1\times\frac{m}{2}}$ are arbitrary matrices.

Proof. Let

$$\begin{bmatrix} C_x \\ C_y\end{bmatrix} = \begin{bmatrix} \mathcal{C}_1 & \mathcal{C}_2 \\ \mathcal{C}_3 & \mathcal{C}_4\end{bmatrix}, \tag{38}$$

where $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4 \in \mathbb{R}^{1\times\frac{m}{2}}$. Then (35) becomes

$$\begin{bmatrix} C_x \\ C_y\end{bmatrix} = T(\theta_g)\begin{bmatrix} \mathcal{C}_1 & \mathcal{C}_2 \\ \mathcal{C}_3 & \mathcal{C}_4\end{bmatrix}\begin{bmatrix} I_{\frac{m}{2}}\cos\theta_g & -I_{\frac{m}{2}}\sin\theta_g \\ I_{\frac{m}{2}}\sin\theta_g & I_{\frac{m}{2}}\cos\theta_g\end{bmatrix}.$$

Multiplying the matrices together and simplifying the 2 independent equations lead to

$$\begin{bmatrix} -\sin^2\theta_g & -\cos\theta_g\sin\theta_g \\ \cos\theta_g\sin\theta_g & -\sin^2\theta_g \\ \cos\theta_g\sin\theta_g & -\sin^2\theta_g \\ \sin^2\theta_g & \cos\theta_g\sin\theta_g\end{bmatrix}^\top \begin{bmatrix} \mathcal{C}_1 \\ \mathcal{C}_2 \\ \mathcal{C}_3 \\ \mathcal{C}_4\end{bmatrix} = \begin{bmatrix} 0 \\ 0\end{bmatrix}. \tag{39}$$

Thus, $[\mathcal{C}_1^\top, \mathcal{C}_2^\top, \mathcal{C}_3^\top, \mathcal{C}_4^\top]^\top$ must lie in the non-trivial null spaces spanned by $[0\ -1\ 1\ 0]^\top$ and $[1\ 0\ 0\ 1]^\top$, which means that $\mathcal{C}_1 = \mathcal{C}_4$ and $\mathcal{C}_2 = -\mathcal{C}_3$, verifying the form of C in (37). A similar analysis of (36) shows that A must have the specific form in (37). □

Thus, the cascaded system (16)–(17) remains invariant under the transformations given in (33)–(34) if A and $\left[C_x^\top \quad C_y^\top\right]^\top$ satisfy (37). Characterizing what linear systems can be transformed to satisfy (37) is beyond the scope of this paper. However, we note that an important case where (37) is satisfied is when $\mathcal{N}$ and $\mathcal{E}$ are zero matrices, which means that the disturbances along the x and y directions are decoupled and share the same dynamic model. The case where $\mathcal{N}$ and $\mathcal{E}$ are zero can also be proved using Proposition 5. Note that Proposition 5 is applicable to any group operations satisfying (7). However, because Proposition 7 is specific to $\varphi_g(X)$ in (19) and the rotation operation (34), it allows $\mathcal{N}$ and $\mathcal{E}$ to be nonzero, thereby encompassing a wider class of disturbance systems than Proposition 5.

Barrau and Bonnabel (2017) provides a novel matrix IEKF design that is applicable to this problem. Such a design requires that the state be written as an element of a matrix Lie group and that a certain 'group affine' condition be satisfied. It turns out that the conditions on the A, C matrices given in (37) are the exact same conditions required for the matrix implementation to be group affine. Additional details are provided in the Appendix.

For the remainder of the section, we assume that the disturbance subsystem is in the form of (37). Applying the same process as in Section 4, we obtain the observer equation as

$$\dot{\hat{X}} = f(\hat{X}) + W(\hat{\theta})\cdot L\cdot T(\hat{\theta})\left(Y - \hat{Y}\right), \tag{40}$$

where $T(\hat{\theta})$ is the same as (25). The invariant frame is now given by

$$W\left(\hat{\theta}\right) = \begin{bmatrix} T(\hat{\theta})^\top & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & T(\hat{\theta})^\top \otimes I_{\frac{m}{2}}\end{bmatrix}. \tag{41}$$

The invariant error is given in (29) with the invariant frame now defined by (41), where the error of the disturbances is also rotated. This results in the following invariant error dynamics

$$\begin{aligned}\dot{\sigma}_x &= v\,(\cos\sigma_\theta - 1) + \omega\sigma_y + C_x\sigma_d \\ &\quad + L_{11}\sigma_x + L_{12}\sigma_y + L_{31}\sigma_x\sigma_y + L_{32}\sigma_y^2\end{aligned}$$

$$\begin{aligned}\dot{\sigma}_y &= v\sin\sigma_\theta - \omega\sigma_x + C_y\sigma_d \\ &\quad + L_{21}\sigma_x + L_{22}\sigma_y - L_{31}\sigma_x^2 - L_{32}\sigma_x\sigma_y \\ \dot{\sigma}_\theta &= L_{31}\sigma_x + L_{32}\sigma_y \\ \dot{\sigma}_d &= A_\omega\sigma_d + L_{d1}\sigma_x + L_{d2}\sigma_y\end{aligned} \tag{42}$$

where

$$A_\omega = A + \left(\begin{bmatrix} 0 & \omega \\ -\omega & 0\end{bmatrix} \otimes I_{\frac{m}{2}}\right). \tag{43}$$

Linearizing (42) around $\sigma = 0$ results in the state matrix needed for implementing this IEKF design:

$$A_k = \begin{bmatrix} 0 & \omega_k & 0 & C_x \\ -\omega_k & 0 & v_k & C_y \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & A_\omega \end{bmatrix}. \tag{44}$$

Note that unlike (32), the state matrix given in (44) is not a function of the estimated state $\hat{\theta}$.

6. Filter covariance transformation

To fully derive the IEKF algorithm, we note that the invariant state error (29) rotates the conventional estimation error to another frame. Thus, the initial state covariance, process noise, and measurement noise matrices, P, Q and R, respectively, can no longer accurately represent the uncertainty in the transformed system. We propose that these matrices be transformed to ensure the IEKF operates at its full potential for different cases of sensor noise and initial error. We next discuss how to rotate the covariance to the invariant error frame.

For the system in (16)–(17), we define

$$\mathcal{X} = \begin{bmatrix} X^\top & d^\top \end{bmatrix}^\top \tag{45}$$

and denote the invariant state error and invariant output error as:

$$\sigma = W(\hat{\theta})^\top(\mathcal{X} - \hat{\mathcal{X}}), \quad E = T(\hat{\theta})(\hat{Y} - Y), \tag{46}$$

respectively, where $W(\hat{\theta})^\top$ is given in either (26) or (41) and $T(\hat{\theta})$ is given in (25). We now derive the transformation rule for the measurement noise matrix R. We use the notation $\mathcal{N}(\mu, \Sigma)$ to denote the Gaussian distribution with mean μ and covariance Σ.

Proposition 8. *Let $\epsilon = \hat{Y} - Y \sim \mathcal{N}(0, R)$. Let $\hat{\theta}$ be the estimate of θ such that $\delta_\theta = \hat{\theta} - \theta \sim \mathcal{N}(0, q_\theta)$. Suppose that δ_θ and ϵ are uncorrelated. Then*

$$\mathrm{cov}(T(\hat{\theta})\epsilon) \approx T(\theta)RT(\theta)^\top + q_\theta \frac{\partial T}{\partial\theta} R \frac{\partial T}{\partial\theta}^\top \tag{47}$$

for a sufficiently small q_θ.

Proof. Since $\delta_\theta \sim \mathcal{N}(0, q_\theta)$ with q_θ sufficiently small and $T(\hat{\theta})\epsilon = T(\theta + \delta_\theta)\epsilon$, we use the first order approximation to obtain

$$\mathbb{E}(T(\theta+\delta_\theta)\epsilon) \approx \mathbb{E}\left(\left(T(\theta) + \frac{\partial T}{\partial\theta}\delta_\theta\right)\epsilon\right) = \frac{\partial T}{\partial\theta}\mathbb{E}(\delta_\theta\epsilon). \tag{48}$$

When δ_θ is uncorrelated with ϵ, $\mathbb{E}(\delta_\theta\epsilon) = 0$, implying $\mathbb{E}(T(\theta + \delta_\theta)\epsilon) \approx 0$.

The covariance of $T(\hat{\theta})\epsilon$ is computed as

$$\begin{aligned}\mathrm{cov}(T(\hat{\theta})\epsilon) &\approx \mathbb{E}\left(\left(T(\theta) + \frac{\partial T}{\partial\theta}\delta_\theta\right)\epsilon\epsilon^\top\left(T(\theta) + \frac{\partial T}{\partial\theta}\delta_\theta\right)^\top\right) \\ &\quad - \frac{\partial T}{\partial\theta}\mathbb{E}(\delta_\theta\epsilon)\mathbb{E}(\delta_\theta\epsilon^\top)\frac{\partial T}{\partial\theta}^\top \\ &= T(\theta)RT(\theta)^\top + \frac{\partial T}{\partial\theta}\mathbb{E}\left(\delta_\theta\epsilon\epsilon^\top\delta_\theta\right)\left(\frac{\partial T}{\partial\theta}\right)^\top \\ &\quad + \frac{\partial T}{\partial\theta}\mathbb{E}\left(\delta_\theta\epsilon\epsilon^\top\right)T(\theta)^\top + T(\theta)\mathbb{E}\left(\epsilon\epsilon^\top\delta_\theta\right)\left(\frac{\partial T}{\partial\theta}\right)^\top \\ &= T(\theta)RT(\theta)^\top + \frac{\partial T}{\partial\theta}\mathbb{E}(\delta_\theta^2\epsilon\epsilon^\top)\frac{\partial T}{\partial\theta}^\top\end{aligned} \tag{49}$$

where the expectation of the third order term $\delta_\theta\epsilon\epsilon^\top$ is zero (Triantafyllopoulos, 2002, Theorem 3.1). Because δ_θ is zero mean with a variance q_θ and δ_θ is uncorrelated with ϵ, it follows that $\mathbb{E}(\delta_\theta^2\epsilon\epsilon^\top) = q_\theta R$ and thus (47) follows from (49). □

Similarly, the process noise matrix Q can be transformed in the same way. Suppose that the process noise is given by $\nu \sim \mathcal{N}(0, Q)$. Since ν and δ_θ are uncorrelated, it follows from Proposition 8 that

$$\mathrm{cov}(W(\hat{\theta})^\top\nu) \approx W(\theta)^\top QW(\theta) + q_\theta\frac{\partial W}{\partial\theta}^\top Q\frac{\partial W}{\partial\theta}. \tag{50}$$

To transform the initial state covariance matrix P, we let the initial state error be $\eta_0 = \hat{\mathcal{X}}_0 - \mathcal{X}_0 \sim \mathcal{N}(0, P_0)$. Note that δ_θ is the same as the initial state error of θ in η_0. We assume that the other elements in η_0 are uncorrelated with δ_θ. Thus, η_0 and δ_θ are correlated only in the θ element. Due to the specific form of $W(\theta)$ for this problem, $\frac{\partial W}{\partial\theta}^\top\mathbb{E}(\delta_\theta\eta_0) = 0$. It then follows from (49) that

$$\mathrm{cov}(W(\hat{\theta})^\top\eta_0) \approx W(\theta_0)^\top P_0 W(\theta_0) + \frac{\partial W}{\partial\theta}^\top\mathbb{E}(\delta_\theta^2\eta_0\eta_0^\top)\frac{\partial W}{\partial\theta}. \tag{51}$$

In implementation, we replace θ with its estimate $\hat{\theta}$, assuming that they are close. Note that Barrau and Bonnabel (2017) uses only the first term in (47) in their examples (Section IV-B-3), which corresponds to the zeroth-order approximation of the covariance. Through simulations in Section 7, we demonstrate the significant improvement due to the second term when the measurement noise is non-isotropic.

Having found the rotated covariances, we present the IEKF algorithm in Algorithm 1. Algorithm 1 follows the standard steps of an EKF except line 2, 7, and 9 where the covariance matrices are modified, line 5 where the linearized A_k is computed based on the invariant error dynamics (see A_k in (32) and (44) for the two IEKF designs), and line 11 where the update equation is modified with transformations of the innovation.

Algorithm 1 The IEKF

1: Initialize $\hat{\mathcal{X}}_0$, P_0 in the original coordinates.
2: $P = W(\theta_0)^\top P_0 W(\theta_0) + \frac{\partial W}{\partial\theta}^\top\mathbb{E}(\delta_\theta^2\eta_0\eta_0^\top)\frac{\partial W}{\partial\theta}$
3: **for** $k = 1$ **to** n **do**
4: $\quad\hat{\mathcal{X}}_k^- = f(\hat{\mathcal{X}}_{k-1}^+, U)$
5: $\quad$Compute A_k
6: $\quad$Compute H_k
7: $\quad Q_{rot} = W(\hat{\theta})^\top QW(\hat{\theta}) + P_{k-1}^\theta\frac{\partial W}{\partial\hat{\theta}}^\top Q\frac{\partial W}{\partial\hat{\theta}}$
8: $\quad P_k^- = A_k P_{k-1}^+ A_k^\top + Q_{rot}$
9: $\quad R_{rot} = T(\hat{\theta})RT(\hat{\theta})^\top + P_k^\theta\frac{\partial T}{\partial\hat{\theta}}R\frac{\partial T}{\partial\hat{\theta}}^\top$
10: $\quad L_k = P_k^- H_k^\top\left(H_k P_k^- H_k^\top + R_{rot}\right)^{-1}$
11: $\quad\hat{\mathcal{X}}_k^+ = \hat{\mathcal{X}}_k^- + W(\hat{\theta})L_k T(\hat{\theta})(Y - h(\hat{\mathcal{X}}_k^-, U))$
12: $\quad P_k^+ = (I - L_k H_k)P_k^-$
13: **end for**

7. Simulations

In this section, we compare the performances of the proposed IEKF designs against the EKF in a simulation environment. Each graph represents a Monte Carlo simulation with 100 trials. The simulations were run with the robot maintaining a constant linear velocity and constant turning rate, collecting measurements

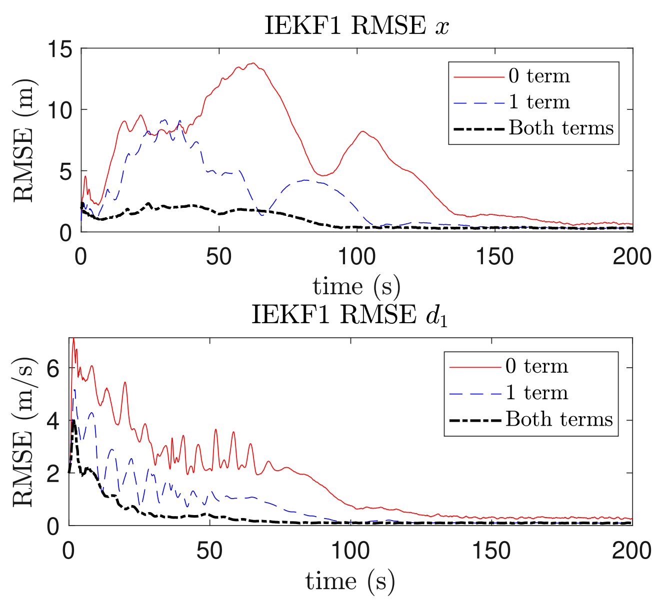

Fig. 1. Effect of rotated noise terms for IEKF1. *Top:* RMSE of x estimate. *Bottom:* RMSE of d_1 estimate.

at a rate of 10 Hz. The following parameters were used in all the simulations: $v = 13$ m/s, $\omega = 4$ deg/s, $\mu_0 = \mathbf{0}_{n\times 1}$, $P_0 = \text{diag}(10^2, 10^2, (\pi/2)^2, 2^2, 2^2, 2^2, 2^2)$, $\mathcal{X}_0 \sim \mathcal{N}(\mu_0, P_0)$, $\hat{\mathcal{X}}_0 = \mu_0$. The simulated measurement noise was generated from a zero mean Gaussian distribution with a non-isotropic covariance given by $R = \begin{bmatrix} 9 & 8 \\ 8 & 9 \end{bmatrix}$. The disturbances were generated from a linear time-invariant model with

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix}. \tag{52}$$

The outputs from the linear system are two signals containing both sinusoidal oscillations plus a constant offset. Note that (52) satisfies the specific form given in Proposition 7. Thus, both IEKF designs can be applied.

Our metric of performance is the root mean square error (RMSE) of each filter's estimate with simulated 'truth' data, calculated at every time step. Let $x_i(t)$ be the ith element of $\mathcal{X}$ at time t. The RMSE of $x_i(t)$ is given by

$$RMSE_i(t) = \sqrt{\frac{\sum_{j=1}^{n} \left(x_i(t) - \hat{x}_i(t)\right)^2}{n}} \tag{53}$$

where n is the number of trials.

Effect of Transformed Noise We demonstrate the effect of different rotated noise terms on the performance of the IEKF designs. In the simulation the filters are run using 3 different approaches to handling the covariance matrices. The first approach, denoted as '0 term' in Figs. 1 and 2, does not transform the P, Q, and R matrices, i.e., in Algorithm 1, $P = P_0$, $Q_{rot} = Q$, and $R_{rot} = R$. The second approach, referred to as '1 term', includes only the first term on the right side of (47), (50) and (51), excluding the first derivative terms. The '1 term' approach corresponds to noise covariance used in Barrau and Bonnabel (2017, Section IV-B-3). Lastly, 'both terms' refers to using the transformations in (47), (50) and (51).

From Fig. 1 we see this comparison for IEKF1. Using both terms from Eqs. (47), (50) and (51) results in the best transient performance and fastest convergence rate for estimating x and d_1. The rest of the states all have similar trends.

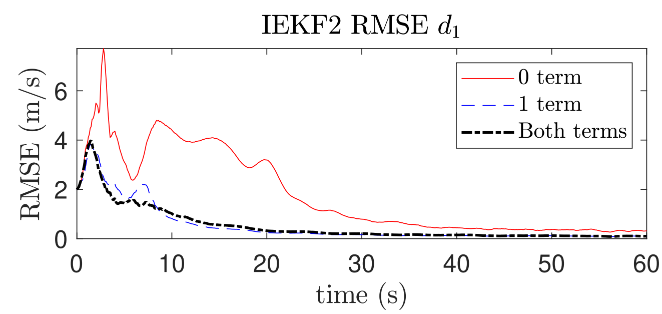

Fig. 2. Effect of rotated noise terms for IEKF2. RMSE of d_1 estimate.

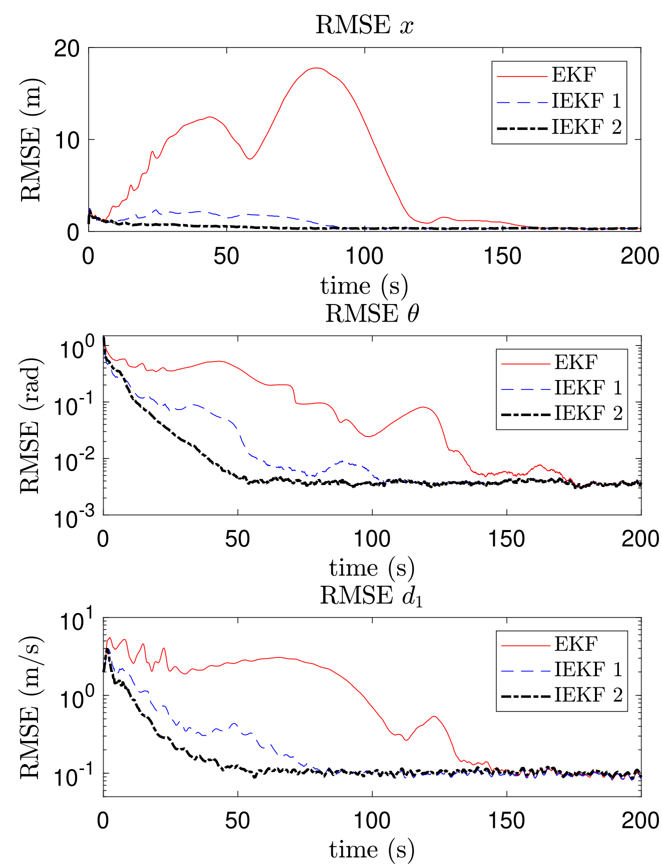

Fig. 3. RMSE comparison of EKF, IEKF1 and IEKF2. *Top:* RMSE of x estimate. *Middle:* RMSE of θ estimate. *Bottom:* RMSE of d_1 estimate.

Fig. 2 shows the same comparison for IEKF2. From Fig. 2, we see that both changes to the covariances improve the transient performance over the nominal case. For IEKF2, the addition of the first order correction term has a less significant impact than it does for IEKF1. However, it still improves the performance at the beginning of the simulation. Only the graph of the state d_1 is provided, however, similar trends extend to the other states. Since our simulation results show that adding the full noise correction given in (47), (50) and (51) improves the performance of both IEKF1 and IEKF2, this implementation is included in the performance comparisons for the remainder of the section.

EKF/IEKF Comparison We now compare the performances of the IEKF designs with that of the traditional EKF. Fig. 3 shows the comparison of the RMSE for the EKF, IEKF1 and IEKF2 for the states x, θ and d_1. In all graphs, both IEKF1 and IEKF2 show superior transient performance over the EKF. IEKF2 has the best performance of the two IEKF designs. Therefore, it is clear that if the disturbance model can be represented in a form that satisfies Proposition 7, the design IEKF2 should be used.

If the disturbance model does not satisfy (37), IEKF2 is not applicable and we propose IEKF1 as another option. Through numerous numerical simulations, which are not included due to the space limit, we have observed that IEKF1 usually produces a performance that is comparable to or better than that of the EKF. The nature of this improvement is related to the form of A and C, which will be studied in our future work.

8. Conclusions and future work

In this paper, we have extended the theory of invariant nonlinear systems by analyzing the requirements for invariant systems to remain invariant when dynamic additive disturbances are applied. Two sets of invariant conditions are developed. We show three examples where these conditions can be utilized, including an attitude dynamics, a chemical reactor, and a unicycle robot model. We focus specifically on the unicycle model to develop and compare two IEKFs designed to estimate both the unicycle state and the disturbance. An additional correction for the IEKF filter covariance is proposed and its contribution is demonstrated through Monte Carlo simulations. Finally, the performance of both proposed IEKF designs is shown to be superior to the performance of the EKF. Our future work includes further investigation of connections of the proposed IEKF designs with matrix IEKF designs for different dynamical systems and extending the unicycle example to 3-D disturbance estimation.

Appendix. Matrix implementation

We now show that the condition obtained in Proposition 7 is equivalent to the group affine condition in Barrau and Bonnabel (2017) for the matrix IEKF design for (16)–(18). Let $\mathcal{G}$ be the matrix Lie Group of double direct spatial isometries and define the system state $M \in \mathcal{G}$ as

$$M = \begin{bmatrix} \cos\theta & -\sin\theta & x & d_1^\top \\ \sin\theta & \cos\theta & y & d_2^\top \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & I_{\frac{m}{2}} \end{bmatrix}, \tag{A.1}$$

where $d_1, d_2 \in \mathbb{R}^{\frac{m}{2}\times 1}$ such that $d = [d_1^\top \ d_2^\top]^\top$. We rewrite the augmented dynamics (16)–(17) as $\dot{M} = F(M)$, where

$$F(M) = \begin{bmatrix} -\omega\sin\theta & -\omega\cos\theta & v\cos\theta + C_x d & d^\top A_1^\top \\ \omega\cos\theta & -\omega\sin\theta & v\sin\theta + C_y d & d^\top A_2^\top \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{A.2}$$

in which A_1 and A_2 are defined as the top and bottom $\frac{m}{2}$ rows of A, respectively, i.e., $A = [A_1^\top \ A_2^\top]^\top$. The measurements given in (18) are $Y = Mq$, where $q = [0, 0, 1, 0, \dots, 0]^\top$. Algebraic manipulations then show that the group affine condition

$$F(ab) = F(a)b + aF(b) - aF(I)b, \quad \forall a, b \in \mathcal{G} \tag{A.3}$$

results in the same condition (37) on A and C. Thus, for this problem, if Proposition 7 is satisfied, a matrix IEKF can be designed accordingly. The resulting matrix IEKF has a linear error dynamics with the same state matrix as in (44).

References

Arulampalam, M. Sanjeev, Maskell, Simon, Gordon, Neil, & Clapp, Tim (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, *50*(2), 174–188.

Bai, He (2018). Motion-dependent estimation of a spatial vector field with multiple vehicles. In *2018 IEEE conference on decision and control (CDC)* (pp. 1379–1384). IEEE.

Barczyk, Martin, Bonnabel, Silvère, Deschaud, Jean-Emmanuel, & Goulette, François (2015). Invariant EKF design for scan matching-aided localization. *IEEE Transactions on Control Systems Technology*, *23*(6), 2440–2448.

Barrau, Axel, & Bonnabel, Silvère (2017). The invariant extended Kalman filter as a stable observer. *IEEE Transactions on Automatic Control*, *62*(4), 1797–1812.

Beard, Randal W., & McLain, Timothy W. (2012). *Small unmanned aircraft: Theory and practice*. Princeton university press.

Betke, Margrit, & Gurvits, Leonid (1997). Mobile robot localization using landmarks. *IEEE Transactions on Robotics and Automation*, *13*(2), 251–263.

Bonnabel, Silvère, Martin, Philippe, & Rouchon, Pierre (2008). Symmetry-preserving observers. *IEEE Transactions on Automatic Control*, *53*(11), 2514–2526.

Bonnable, Silvère, Martin, Philippe, & Salaün, Erwan (2009). Invariant extended Kalman filter: theory and application to a velocity-aided attitude estimation problem. In *Proceedings of the 48h IEEE conference on decision and control (CDC) held jointly with 2009 28th Chinese control conference* (pp. 1297–1304). IEEE.

Coleman, Kevin, Bai, He, & Taylor, Clark N. (2020). Invariant-EKF design for a unicycle robot under linear disturbances. In *2020 American control conference (ACC)* (pp. 3479–3484). IEEE.

De Silva, Oscar, Mann, George K. I., & Gosine, Raymond G. (2014). Relative localization with symmetry preserving observers. In *2014 IEEE 27th Canadian conference on electrical and computer engineering (CCECE)* (pp. 1–6). IEEE.

Gelb, Arthur (1974). *Applied optimal estimation*. MIT press.

Isidori, Alberto, & Byrnes, Christopher I. (1990). Output regulation of nonlinear systems. *IEEE Transactions on Automatic Control*, *35*(2), 131–140.

Isidori, Alberto, Marconi, Lorenzo, & Serrani, Andrea (2012). *Robust autonomous guidance: an internal model approach*. Springer Science & Business Media.

Julier, Simon J., & Uhlmann, Jeffrey K. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, *92*(3), 401–422.

Khalil, Hassan K. (2002). Nonlinear systems. *Upper Saddle River*.

Kolmanovsky, Ilya, & Harris McClamroch, N. (1995). Developments in nonholonomic control problems. *IEEE Control Systems Magazine*, *15*(6), 20–36.

Martin, Philippe, & Salaun, Erwan (2008). A general symmetry-preserving observer for aided attitude heading reference systems. In *2008 47th IEEE conference on decision and control* (pp. 2294–2301). IEEE.

Olver, Peter J. (1999). *Classical invariant theory, Vol. 44*. Cambridge University Press.

Palanthandalam-Madapusi, Harish J, Girard, Anouck, & Bernstein, Dennis S (2008). Wind-field reconstruction using flight data. In *2008 American control conference* (pp. 1863–1868). IEEE.

Petrich, Jan, Woolsey, Craig A., & Stilwell, Daniel J. (2009). Planar flow model identification for improved navigation of small AUVs. *Ocean Engineering*, *36*(1), 119–131.

Phogat, Karmvir Singh, & Chang, Dong Eui (2020). Invariant extended Kalman filter on matrix Lie groups. *Automatica*, *114*, Article 108812.

Schwartz, B., Isidori, Alberto, & Tarn, Tzyh Jong (1999). Global normal forms for MIMO nonlinear systems, with applications to stabilization and disturbance attenuation. *Mathematics of Control, Signals, and Systems*, *12*(2), 121–142.

Triantafyllopoulos, Kostas (2002). *Moments and cumulants of the multivariate real and complex Gaussian distributions, Vol. 12*. Department of Mathematics, University of Bristol, Version.

Trumpf, Jochen, Mahony, Robert, & Hamel, Tarek (2018). On the structure of kinematic systems with complete symmetry. In *2018 IEEE conference on decision and control (CDC)* (pp. 1276–1280). IEEE.

Wu, Kanzhi, Zhang, Teng, Su, Daobilige, Huang, Shoudong, & Dissanayake, Gamini (2017). An invariant-EKF VINS algorithm for improving consistency. In *2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (pp. 1578–1585). IEEE.

Zhang, Teng, Wu, Kanzhi, Song, Jingwei, Huang, Shoudong, & Dissanayake, Gamini (2017). Convergence and consistency analysis for a 3-d invariant-EKF SLAM. *IEEE Robotics and Automation Letters*, *2*(2), 733–740.

Kevin Coleman received his B.S. from the University of Missouri and his M.S. in Mechanical and Aerospace Engineering from Oklahoma State University. He currently works as a mechanical engineer for Lockheed Martin Corporation. His research interests include nonlinear estimation and control, adaptive control and robotic systems.

He Bai received his B.Eng. degree from the Department of Automation at the University of Science and Technology of China, Hefei, China, in 2005, and the M.S. and Ph.D. degrees in Electrical Engineering from Rensselaer Polytechnic Institute (RPI) in 2007 and 2009, respectively. From 2009 to 2010, he was a Post-doctoral Researcher at Northwestern University, Evanston, IL. From 2010 to 2015, he was a Senior Research and Development Scientist at UtopiaCompression Corporation. In 2015, he joined the Mechanical and Aerospace Engineering Department at Oklahoma State University as an assistant professor. He has published over eighty peer-reviewed journal and conference papers related to control and robotics and a research monograph "Cooperative control design: a systematic passivity-based approach" in Springer. He holds one patent on monocular passive ranging. His research interests include nonlinear control and estimation, distributed optimization and learning, reinforcement learning control, multi-agent systems, trajectory optimization, and human–robot interaction.

Clark N. Taylor is currently an Assistant Professor in Computer Engineering at the Air Force Institute of Technology. He received his Ph.D. degree in electrical and computer engineering from the University of California, San Diego, in 2004. From 2004–2010, he was an assistant professor in electrical and computer engineering at Brigham Young University, and from 2010–2018, he was a research electronics engineer with the Air Force Research Laboratory (AFRL). He has published over 90 papers in the fields of video processing for unmanned aerial vehicles (UAVs), GPS-denied navigation, estimation theory, video communication, and digital systems design.