The Invariant Extended Kalman Filter as a Stable Observer

Axel Barrau and Silvère Bonnabel

***Abstract*—We analyze the convergence aspects of the invariant extended Kalman filter (IEKF), when the latter is used as a deterministic nonlinear observer on Lie groups, for continuous-time systems with discrete observations. One of the main features of invariant observers for left-invariant systems on Lie groups is that the estimation error is autonomous. In this paper we first generalize this result by characterizing the (much broader) class of systems for which this property holds. For those systems, the Lie logarithm of the error turns out to obey a *linear* differential equation. Then, we leverage this "log-linear" property of the error evolution, to prove for those systems the local stability of the IEKF around *any* trajectory, under the standard conditions of the linear case. One mobile robotics example and one inertial navigation example illustrate the interest of the approach. Simulations evidence the fact that the EKF is capable of diverging in some challenging situations, where the IEKF with identical tuning keeps converging.**

***Index Terms*—Asymptotic observers, Kalman filters, nonlinear filters, observers.**

I. Introduction

THE design of nonlinear observers is always a challenge, as except for a few classes of systems (e.g., [15]), no general method exists. Of course, the grail of nonlinear observer design is to achieve global convergence to zero of the state estimation error, but this is a very ambitious property to pursue. As a first step, a general method is to use standard linearization techniques, such as the extended Kalman filter (EKF) that makes use of Kalman equations to stabilize the linearized estimation error, and then attempt to derive local convergence properties around *any* trajectory. This is yet a rare property to obtain in a nonlinear setting (see, e.g., [1]), due to the fact that the linearized estimation error equation is time varying, and contrarily to the linear case it generally depends on the unknown true state we seek to estimate. The EKF, the most popular observer in the engineering world, provides an "off the shelf" candidate observer, potentially able to deal with the time-varying nature of the linearized error equation, due to its adaptive gain tuning through a Riccati equation. However, the EKF does not possess any optimality guarantee, and its efficiency is aleatory. Indeed, its main flaw lies in its very nature: the Kalman gain is computed assuming the estimation error is sufficiently small to be propagated analytically through a first-order linearization of the dynamics about the *estimated* trajectory. When the estimate is actually far from the true state variable, the linearization is not valid, and results in an unadapted gain that may amplify the error. In turn, such positive feedback loop may lead to divergence of the filter. This is the reason why most of the papers dealing with the stability of the EKF (see [10], [11], [24], [25]) rely on the highly nontrivial assumption that the eigenvalues of the Kalman covariance matrix P_t computed about the *estimated* trajectory are lower and upper bounded by strictly positive scalars. To the authors' knowledge, only a few papers deal with the stability of the EKF without invoking this assumption [19]. It is then replaced by second-order properties whose verification can prove difficult in practical situations, and also by the assumption that the process noise covariance matrix is positive definite. This lack of guarantee is also due to the fact the filter *can* diverge indeed in a number of applications. Note that, beyond the general theory, there are not even that many engineering examples where the EKF is proved to be (locally) stable.

The present paper builds upon the theory of symmetry preserving observers [7], [8] and notably the theory of invariant Kalman filtering [5], [6], [9], [23] in a purely *deterministic* context. As such, it is a contribution to the theory of nonlinear observer design on Lie groups that has lately attracted considerable interest, notably for attitude estimation, see, e.g., [3], [16]–[18], [20], [22], [26]. The detailed contributions and organization of the paper are as follows.

In Section II, we recall the main contribution of [7], [8] is to evidence the fact that for left-invariant systems on Lie groups, nonlinear observers may be designed in such a way that the left-invariant estimation error obeys an autonomous equation, a key property for observer design on Lie groups. We show here this property of the error equation can actually be obtained for a *much* broader class of systems, and we characterize this class. Very surprisingly, it turns out that, up to a suitable nonlinear change of variables, the error evolution (in the absence of measurements) obeys a *linear* differential equation.

In Section III, we focus on the invariant extended Kalman filter (IEKF) [6] when applied to the broad class of systems of Section II. We consider continuous-time models with discrete observations, which best suits navigation systems where high rate sensors governing the dynamics are to be combined with low rate sensors [14]. We change a little the IEKF equations to cast them into a matrix Lie group framework, more handy to use

Manuscript received September 12, 2015; revised September 14, 2015, April 12, 2016, and June 15, 2016; accepted July 14, 2016. Date of publication July 27, 2016; date of current version March 27, 2017. Recommended by Associate Editor L. Menini.

A. Barrau is with Safran Tech, Groupe Safran, Rue des Jeunes Bois—Châteaufort 78772 Magny Les Hameaux Cedex France (e-mail: axel.barrau@safrangroup.com).

S. Bonnabel is with the MINES ParisTech, PSL Research University, Centre for Robotics, 75006 Paris, France (e-mail: silvere.bonnabel@mines-paristech.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2016.2594085

than the usual abstract Lie group formulation of [6]. We then prove, that under the standard convergence conditions of the linear case [13], applied to the linearized model around the *true* state, the IEKF is an asymptotic observer around *any* trajectory of the system, a rare to obtain property. This way, we produce a generic observer with guaranteed local convergence properties under natural assumptions, for a broad class of systems on Lie groups, whereas this property has so far only been reserved to specific examples on Lie groups. This also allows putting on firm theoretical ground the good behavior of the IEKF in practice, as already noticed in a few papers, see e.g., [2], [3], [5].

In Section IV, we consider a mobile robotics example, where a unicycle robot (or simplified car) tries to estimate its position and orientation from GPS position (only) measurements, or alternatively landmarks range and bearing measurements. On this example of engineering interest, the IEKF is proved to converge around *any* trajectory using the results of the paper, which is a contribution in itself. Simulations indicate the IEKF is always superior to the EKF and may even outperform the latter in challenging situations.

In Section V, we consider the highly relevant problem of an unmanned aerial vehicle (UAV) navigating with accelerometers and gyrometers, and range and bearing measurements of known landmarks. Although the system is not invariant in the sense of [7], [8], it is proved to fit into our framework so that the autonomous error equation property of [7] holds, a fact never noticed before to our best knowledge (except in our preliminary conference paper [4]). The IEKF is shown to converge around *any* trajectory using the results of the paper, which is a contribution in itself. Moreover, it is shown to outperform the EKF which even diverges when, as in high precision navigation, the user has way more trust in the inertial sensors than in the landmark measurements.

The main contributions can be summarized as follows:

- The class of systems, for which the key result of [7] about the (state) error equation autonomy holds, is completely characterized, and actually shown to be much broader than left-invariant systems.
- The autonomy of the error equation is proved to come with a very intriguing property: a well-chosen nonlinear function of the nonlinear error is proved to obey a linear differential equation.
- In turn, this property allows proving that, for the introduced class of systems, the IEKF used in a deterministic context possesses powerful local convergence guarantees that the standard EKF lacks.
- Two examples of navigation illustrate the results, and simulations indicate indeed the IEKF is always superior to the EKF, and may turn out to literally outperform the latter when confronted with some challenging situations—the EKF being capable of diverging.

II. A Special Class of Multiplicative Systems

A. An Introductory Example

Consider a linear (deterministic) system $(d/dt)z_t = A_t z_t$. Consider two trajectories of this system, say, a reference trajectory $(\bar{z}_t)_{t\geq 0}$ and another one $(z_t)_{t\geq 0}$. The discrepancy between both trajectories $\Delta z_t := z_t - \bar{z}_t$ satisfies the linear equation $(d/dt)\Delta z_t = A_t \Delta z_t$. This is a key property for the design of linear convergent observers, as during the propagation step, the evolution of the error between the true state and the estimate does *not* depend on the true state's trajectory.

Consider now the following nonlinear model of a two-dimensional simplified car. Its state is defined by three parameters: heading θ_t and position $x_t = (x_t^{(1)}, x_t^{(2)})$. The velocity $v_t \in \mathbb{R}$ is measured by an odometer and the angular velocity $\omega_t \in \mathbb{R}$ is measured by differential odometry. Assume the motion is modelled by the unicycle equations [12]: $(d/dt)\theta_t = \omega_t$, $(d/dt)x_t^{(1)} = \cos(\theta_t)v_t$, $(d/dt)x_t^{(2)} = \sin(\theta_t)v_t$. Now consider a reference trajectory $(\bar{\theta}_t, \bar{x}_t)$ and another trajectory (θ_t, x_t) with different initial conditions but same inputs. The exact propagation of the "error" $(\Delta\theta_t, \Delta x_t) = (\theta_t - \bar{\theta}_t, x_t - \bar{x}_t)$, satisfies

$$\frac{d}{dt}\Delta\theta_t = 0,\ \frac{d}{dt}\Delta x_t^{(1)} = \left[\cos(\theta_t) - \cos(\bar{\theta}_t)\right] v_t$$
$$\frac{d}{dt}\Delta x_t^{(2)} = \left[\sin(\theta_t) - \sin(\bar{\theta}_t)\right] v_t \tag{1}$$

where we let $\Delta x_t = (\Delta x_t^{(1)}, \Delta x_t^{(2)})$. We see the time derivative of $(\Delta\theta_t, \Delta x_t)$ is not a function of $(\Delta\theta_t, \Delta x_t)$ only: it also involves $\bar{\theta}_t$ and θ_t individually. Moreover, the equation is nonlinear. These features, characteristic of nonlinear systems, make the design of observers way more complicated in the nonlinear case. Now, let us introduce the following nonlinear error, where $R(\theta) = \cos(\theta)I_2 + \sin(\theta)J_2$ denotes the planar rotation matrix of angle θ (see definition of J_2 below):

$$\xi_t := \begin{pmatrix} (\theta_t - \bar{\theta}_t) \\ \frac{1}{2}\left[a(\theta_t - \bar{\theta}_t)I_2 - (\theta_t - \bar{\theta}_t)J_2\right] R(-\bar{\theta}_t)(x_t - \bar{x}_t) \end{pmatrix}$$
$$\text{with } J_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ and } a(s) = \frac{s\sin(s)}{1-\cos(s)} \tag{2}$$

which is equal to 0 indeed if and only if both trajectories coincide (assuming we set $a(0) = 2$ by continuity). We are about to prove by elementary means a surprising property that will be generalized by Theorem 2.

Proposition 1: Contrarily to the linear error obeying (1), the alternative nonlinear error (2) obeys the following *linear* and autonomous equation although the system, and the error, are totally nonlinear:

$$\frac{d}{dt}\xi_t = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \omega_t \\ -v_t & -\omega_t & 0 \end{pmatrix} \xi_t. \tag{3}$$

Proof: We will use the notations $\delta\theta_t = (\theta_t - \bar{\theta}_t)$ and $\delta x_t = (1/2)[a(\delta\theta_t)I_2 - \delta\theta_t J_2]R(-\bar{\theta}_t)(x_t - \bar{x}_t)$. Since we have $(d/dt)\delta\theta_t = 0$ as in the linear case above, only the factor $R(-\bar{\theta}_t)(x_t - \bar{x}_t)$ in δx_t changes over time. Differenciating it, we get $(d/dt)\delta x_t = (1/2)[a(\delta\theta_t)I_2 - \delta\theta_t J_2](-\omega_t J_2)R(-\bar{\theta}_t)(x_t - \bar{x}_t) + (1/2)[a(\delta\theta_t)I_2 - \delta\theta_t J_2](R(\delta\theta_t) - I_2)(v_t, 0)^T$. The first term is equal to $-\omega_t J_2 \delta x_t$, and we can use the relation $R(s) - I_2 = (\cos(s) - 1)I_2 + \sin(s)J_2$ to expand the second one yielding the term $(1/2)[(a(\delta\theta_t)(\cos(\delta\theta_t) - 1) + \delta\theta_t \sin(\delta\theta_t))I_2 - (\delta\theta_t(\cos(\delta\theta_t) - 1) + a(\delta\theta_t)\sin(\delta\theta_t))J_2](v_t, 0)$. The coefficient before I_2 vanishes, and using the equality $\sin(s)^2/(1-\cos(s))=1+\cos(s)$ the coefficient before J_2 becomes $2\delta\theta_t$. The second term of $(d/dt)\delta x_t$ then boils down to $\delta\theta_t J_2(v_t,0)^T$. We have finally $(d/dt)\delta x_t=-\omega_t J_2\delta x_t+\delta\theta_t(0,-v_t)^T$, that is, (3). ■

The present section provides a novel geometrical framework—encompassing this example—to characterize systems on Lie groups for which such a property holds. In turn, such a property will simplify the convergence analysis of nonlinear observers, namely the IEKF, due to the implied similarities with the linear case.

B. Systems on Lie Groups With State Trajectory Independent Error Propagation Property

Let $G\subset\mathbb{R}^{N\times N}$ be a matrix Lie group whose Lie algebra is denoted $\mathfrak{g}$ and has dimension dim $\mathfrak{g}$. We consider a class of dynamical systems

$$\frac{d}{dt}\chi_t=f_{u_t}(\chi_t) \tag{4}$$

where the state χ_t lives in the Lie group G and u_t is an input variable. Consider two distinct trajectories χ_t and $\bar{\chi}_t$ of (4). Define the left-invariant and right-invariant errors η_t^L and η_t^R between the two trajectories as

$$\eta_t^L=\chi_t^{-1}\bar{\chi}_t \qquad \text{(left invariant)} \tag{5}$$

$$\eta_t^R=\bar{\chi}_t\chi_t^{-1} \qquad \text{(right invariant)}. \tag{6}$$

The terminology stems from the invariance of, e.g., (5) to (left) multiplications $(\chi,\bar{\chi})\to(\Gamma\chi,\Gamma\bar{\chi})$ for $\Gamma\in G$.

Definition 1: The left-invariant and right-invariant errors are said to have a state-trajectory independent propagation if they satisfy a differential equation of the form $(d/dt)\eta_t=g_{u_t}(\eta_t)$.

Note that, in general the time derivative of η_t is a complicated function depending on u_t and both χ_t and $\bar{\chi}_t$ in a way that does not boil down to a function of η_t, see for instance eq (1) above. The following result allows characterizing the class of systems of the form (4) for which the property holds.

Theorem 1: The three following conditions are equivalent for the dynamics (4):

i The left-invariant error (5) is state trajectory independent
ii The right-invariant error (6) is state trajectory independent
iii For all $t>0$ and $a,b\in G$ we have (in the tangent space at ab)

$$f_{u_t}(ab)=f_{u_t}(a)b+af_{u_t}(b)-af_{u_t}(I_d)b \tag{7}$$

where I_d denotes the identity matrix. Moreover, if one of these conditions is satisfied we have

$$\frac{d}{dt}\eta_t^L=g_{u_t}^L\left(\eta_t^L\right) \quad \text{where} \quad g_{u_t}^L(\eta)=f_{u_t}(\eta)-f_{u_t}(I_d)\eta \tag{8}$$

$$\frac{d}{dt}\eta_t^R=g_{u_t}^R\left(\eta_t^R\right) \quad \text{where} \quad g_{u_t}^R(\eta)=f_{u_t}(\eta)-\eta f_{u_t}(I_d). \tag{9}$$

Proof: Assume we have $(d/dt)\eta_t^L=g_{u_t}(\eta_t^L)$ for a certain function g_{u_t} and any $\eta_t^L=\chi_t^{-1}\bar{\chi}_t$, where χ_t and $\bar{\chi}_t$ are solutions of (4). We have

$$\begin{aligned} g_{u_t}\left(\chi_t^{-1}\bar{\chi}_t\right) &= \frac{d}{dt}\left(\chi_t^{-1}\bar{\chi}_t\right)=-\chi_t^{-1}\left[\frac{d}{dt}\chi_t\right]\chi_t^{-1}\bar{\chi}_t+\chi_t^{-1}\frac{d}{dt}\bar{\chi}_t \\ &= -\chi_t^{-1}f_{u_t}(\chi_t)\eta_t^L+\chi_t^{-1}f_{u_t}(\bar{\chi}_t) \end{aligned}$$

$$\text{i.e. } g_{u_t}\left(\eta_t^L\right)=-\chi_t^{-1}f_{u_t}(\chi_t)\eta_t^L+\chi_t^{-1}f_{u_t}\left(\chi_t\eta_t^L\right). \tag{10}$$

This has to hold for any χ_t and η_t^L. In the particular case where $\chi_t=I_d$ we obtain

$$g_{u_t}\left(\eta_t^L\right)=f_{u_t}\left(\eta_t^L\right)-f_{u_t}(I_d)\eta_t^L. \tag{11}$$

Reinjecting (11) in (10), we obtain

$$f_{u_t}\left(\chi_t\eta_t^L\right)=f_{u_t}(\chi_t)\eta_t^L+\chi_t f_{u_t}\left(\eta_t^L\right)-\chi_t f_{u_t}(I_d)\eta_t^L.$$

The converse is trivial and the proof is analogous for right-invariant errors. ■

Remark 1: The particular cases of left-invariant and right-invariant dynamics, or the combination of both as follows, verify (7). Let $f_{v_t,\omega_t}(\chi)=v_t\chi+\chi\omega_t$. We have indeed

$$\begin{aligned} &f_{v_t,\omega_t}(a)b+af_{v_t,\omega_t}(b)-af_{v_t,\omega_t}(I_d)b \\ &=(v_t a+a\omega_t)b+a(v_t b+b\omega_t)-a(v_t+\omega_t)b=f_{v_t,\omega_t}(ab). \end{aligned}$$

Remark 2: In the particular case where G is a vector space with standard addition as the group composition law, the condition (7) boils down to $f_{u_t}(a+b)=f_{u_t}(a)+f_{u_t}(b)-f_{u_t}(0)$ and we recover the affine functions, that is, linear systems! We thus propose to name systems (4) with condition (7) *group affine systems*.

In the next section, we show that group affine systems have striking properties generalizing those of linear systems.

C. Log-Linear Property of the Error Propagation

In the sequel, we will systematically consider group affine systems (4)–(7), i.e. systems on Lie groups defined by

$$\frac{d}{dt}\chi_t=f_{u_t}(\chi_t)$$

$$\text{where } \forall\,(u,a,b) \quad f_u(ab)=af_u(b)+f_u(a)b-af_u(Id)b. \tag{12}$$

For such systems, Theorem 1 proves that the left (resp. right) invariant error is a solution to the equation $(d/dt)\eta_t=g_{u_t}(\eta_t)$ where g_{u_t} is given by (8) [respectively, (9)]. We have the following novel and striking property.

Theorem 2—[Log-Linear Property of the Error]: Consider the left or right invariant error η_t^i as defined by (5) or (6) between two arbitrarily far trajectories of (12), the superscript i denoting indifferently L or R. Let $\mathscr{L}_{\mathfrak{g}}$ and $\exp(.)$ be defined as in Appendix A. Let $\xi_0^i\in\mathbb{R}^{\dim\mathfrak{g}}$ be such that initially $\exp(\xi_0^i)=\eta_0^i$. Let A_t^i be defined by $g_{u_t}(\exp(\xi))=\mathscr{L}_{\mathfrak{g}}(A_t^i\xi)+O(\|\xi\|^2)$. The existence of A_t^i is proved in Appendix B. If ξ_t^i is defined for $t>0$ by the *linear* differential equation in $\mathbb{R}^{\dim\mathfrak{g}}$

$$\frac{d}{dt}\xi_t^i=A_t^i\xi_t^i \tag{13}$$

then, we have for the true nonlinear error η_t^i, the correspondence *at all times* and for arbitrarily large errors

$$\forall t \geq 0 \quad \eta_t^i = \exp\left(\xi_t^i\right).$$

The latter result, whose proof has been moved to the Appendix, shows that a wide range of nonlinear problems (see examples below) can lead to linear error equations provided the error variable is correctly chosen. We also see the results displayed in the previous introductory example of Section II-A are mere applications of the latter theorem, as the nonholonomic car example turns out to perfectly fit into our framework (see Section IV) and ξ_t in (2) actually merely is the Lie logarithm of the left-invariant error. This will be extensively used in Section III, and in the examples to prove stability properties of IEKFs.

III. Invariant Extended Kalman Filtering

In this section, we first recap the equations of the Invariant EKF (IEKF), a variant of the EKF devoted to Lie groups space states, that has been introduced in continuous time in [6] and [9]. We derive the equations in continuous time with discrete observations here, which has already been done in a restricted setting in [5], and we propose a novel matrix (Lie group) framework to simplify the design. We then show that for the class of systems introduced in Section II, under observability conditions, and painless conditions on the covariance matrices considered here as design parameters, the IEKF is a (deterministic) nonlinear observer with local convergence properties around *any* trajectory, a feature extremely rare to obtain in the field of nonlinear observers, due to the dependency of the estimation error to the true unknown trajectory. The notions necessary to follow Section III are given in Appendix A.

A. Full System and IEKF General Structure

We consider in this section an equation on a matrix Lie group $G \subset \mathbb{R}^{N\times N}$ of the form:

$$\frac{d}{dt}\chi_t = f_{u_t}(\chi_t) \tag{14}$$

with $f_u(ab) = af_u(b) + f_u(a)b - af_u(Id)b$ for all $(u,a,b) \in U \times G \times G$. This system will be associated to two different kinds of observations.

1) Left-Invariant Observations: The first family of outputs we are interested in write:

$$Y_{t_n}^1 = \chi_{t_n} d^1, \ldots, Y_{t_n}^k = \chi_{t_n} d^k \tag{15}$$

where $(d^i)_{i\leq k}$ are known vectors. The Left-Invariant Extended Kalman Filter (LIEKF) is defined in this setting through the following propagation and update steps:

$$\frac{d}{dt}\hat{\chi}_t = f_{u_t}(\hat{\chi}_t), \quad t_{n-1} \leq t < t_n, \qquad \text{Propagation} \tag{16}$$

$$\hat{\chi}_{t_n}^+ = \hat{\chi}_{t_n} \exp\left[L_n\begin{pmatrix}\hat{\chi}_{t_n}^{-1}Y_{t_n}^1 - d^1\\ \ldots \\ \hat{\chi}_{t_n}^{-1}Y_{t_n}^k - d^k\end{pmatrix}\right], \quad \text{Update} \tag{17}$$

where the function $L_n : \mathbb{R}^{kN} \to \mathbb{R}^{\dim \mathfrak{g}}$ is to be defined in the sequel using error linearizations. A left-invariant error between true state χ_t and estimated state $\hat{\chi}_t$ can be associated to this filter:

$$\eta_t^L = \chi_t^{-1}\hat{\chi}_t. \tag{18}$$

During the Propagation step, χ_t and $\hat{\chi}_t$ are two trajectories of the system (14). Thus, the error (18) is independent from the true state trajectory from Theorem 1 and (8)! We have thus

$$\frac{d}{dt}\eta_t^L = g_{u_t}^L\left(\eta_t^L\right), \quad t_{n-1} \leq t < t_n. \tag{19}$$

Consider now the following linear differential equation in $\mathbb{R}^{\dim \mathfrak{g}}$:

$$\frac{d}{dt}\xi_t = A_t\xi_t \tag{20}$$

where A_t is defined by $g_{u_t}^L(\exp(\xi)) = \mathscr{L}_{\mathfrak{g}}(A_t\xi) + O(\|\xi_t\|^2)$. Theorem 2 implies the unexpected result:

Proposition 2: If ξ_t is defined as a solution to the *linear* system (20) and η_t is defined as the solution to the *nonlinear* error system (19), then if at time t_{n-1} we have $\eta_{t_{n-1}} = \exp(\xi_{t_{n-1}})$ then the equality $\eta_t = \exp(\xi_t) = \exp_m(\mathscr{L}_{\mathfrak{g}}(\xi_t))$ is verified at all times $t_{n-1} \leq t < t_n$, even for arbitrarily large initial errors.

Besides, at the update step, the evolution of the invariant error variable (18) merely writes

$$\left(\eta_{t_n}^L\right)^+ = \chi_{t_n}^{-1}\hat{\chi}_{t_n}^+ = \eta_{t_n}^L \exp\left[L_n\begin{pmatrix}\left(\eta_{t_n}^L\right)^{-1}d^1 - d^1\\ \ldots \\ \left(\eta_{t_n}^L\right)^{-1}d^k - d^k\end{pmatrix}\right]. \tag{21}$$

We see that the nice geometrical structure of the LIEKF allows the updated error $\left(\eta_{t_n}^L\right)^+$ to be here again only a function of the error just before update $\eta_{t_n}^L$, i.e., to be independent from the true state χ_{t_n}.

2) Right-Invariant Observations: The second family of observations we are interested in have the form

$$Y_{t_n}^1 = \chi_{t_n}^{-1}d^1, \ldots, Y_{t_n}^k = \chi_{t_n}^{-1}d^k \tag{22}$$

with the same notation as in the previous section. The Right-Invariant EKF (RIEKF) is defined here as

$$\frac{d}{dt}\hat{\chi}_t = f_{u_t}(\hat{\chi}_t) \tag{23}$$

$$\hat{\chi}_{t_n}^+ = \exp\left[L_n\begin{pmatrix}\hat{\chi}_{t_n}Y_{t_n}^1 - d^1\\ \ldots \\ \hat{\chi}_{t_n}Y_{t_n}^k - d^k\end{pmatrix}\right]\hat{\chi}_{t_n}. \tag{24}$$

A right-invariant error can be associated to this filter

$$\eta_t^R = \hat{\chi}_t\chi_t^{-1}. \tag{25}$$

Once again, Theorem 1 proves the evolution of the error does not depend on the state of the system. The analog of Proposition 2 is thus easily derived for the error (25) and we skip it due to space limitations.

At the update step, the evolution of the invariant error variable reads

$$\left(\eta_{t_n}^R\right)^+ = \hat{\chi}_{t_n}^+ \chi_{t_n}^{-1} = \exp\left[L_n \begin{pmatrix} \eta_{t_n}^R d^1 - d^1 \\ \cdots \\ \eta_{t_n}^R d^k - d^k \end{pmatrix}\right] \eta_{t_n}^R$$

so that the error update does not depend on the true state either.

B. IEKF Gain Tuning

In the standard theory of Kalman filtering, EKFs are designed for "noisy" systems associated with the deterministic considered system. In a deterministic context, the covariance matrices Q and N of the noises are left free to tune by the user, and are design parameters for the EKF used as a nonlinear observer. Yet, in the spirit of [25], it is nevertheless convenient to associate a "noisy" system with the considered deterministic system consisting of dynamics (14) with outputs (15) or (22). The obtained error equations can be linearized, and the standard Kalman equations applied to make this error decrease. This way, the matrices Q and N can be *interpreted* as covariance matrices. And in many engineering applications, the characteristics of the noises of the sensors are approximately known, so that the engineer can use the corresponding covariance matrices as a useful guide to tune (or design) the nonlinear observer, that is here the IEKF. This provides him with (at least) a first sensible tuning which is consistent with the trust he has in each sensor. Indeed, the IEKF being an EKF variant, it is the *optimal* filter for the *linearized* noisy problem. Moreover, in the same spirit, the IEKF viewed as a nonlinear observer remedies a common weakness shared by numerous nonlinear observers on Lie groups, as it conveys an information about its own accuracy through the computed covariance matrix P_t. Although it comes with no rigorous interpretation in a deterministic context, the information conveyed by P_t may prove useful in applications.

Note that, in mobile robotics and navigation, the sensors are attached to the earth-fixed frame (e.g., a GPS) or to the body frame (e.g., a gyrometer). To *interpret* them as covariance matrices in the IEKF framework (see below) those matrices Q and N may have to undergo a change of frame yielding trajectory dependent tuning matrices $\hat{Q}$ and $\hat{N}$, such as in the application examples in the sequel. This does not weaken the results, but however comes at the price of making the stability analysis a little more complicated.

1) Associated "Noisy" System: To tune the IEKF (16), (17) or (23), (24), we associate to the system (14) the following "noisy" system:

$$\frac{d}{dt}\chi_t = f_{u_t}(\chi_t) + \chi_t w_t \tag{26}$$

where w_t is a continuous white noise belonging to $\mathfrak{g}$ whose covariance matrix is denoted by Q_t (for a proper discussion on multiplicative noise for systems defined on Lie groups, see e.g., [5]).

In the same way, we associate to the family of left-invariant observations (15) the following family of "noisy" outputs:

$$Y_{t_n}^1 = \chi_{t_n}\left(d^1 + B_n^1\right) + V_n^1, \ldots, Y_{t_n}^k = \chi_{t_n}\left(d^k + B_n^k\right) + V_n^k \tag{27}$$

where the $(V_n^i)_{i\le k}$, $(B_n^i)_{i\le k}$ are noises with known characteristics. To the family of right-invariant observations (22) we associate the following family of "noisy" outputs

$$Y_{t_n}^1 = \chi_{t_n}^{-1}\left(d^1 + V_n^1\right) + B_n^1, \ldots, Y_{t_n}^k = \chi_{t_n}^{-1}\left(d^k + V_n^k\right) + B_n^k. \tag{28}$$

2) Linearized "Noisy" Estimation Error Equation: As in a conventional EKF, we assume the error to be small (here close to Id as it is equal to Id if $\hat{\chi}_t = \chi_t$) so that the error system can be linearized to compute the gains L_n. By definition, the Lie algebra $\mathfrak{g}$ represents the infinitesimal variations around Id of an element of G. Thus the natural way to define a vector error variable ξ_t in $\mathbb{R}^{\dim \mathfrak{g}}$ is (see Appendix A):

$$\eta_t = \exp(\xi_t) = \exp_m\left(\mathscr{L}_{\mathfrak{g}}(\xi_t)\right). \tag{29}$$

During the Propagation step, that is for $t_{n-1} \le t < t_n$, elementary computations based on the results of Theorem 1 show that for the noisy model (26) we have

$$\begin{aligned} \frac{d}{dt}\eta_t^L &= g_{u_t}^L\left(\eta_t^L\right) - w_t \eta_t^L \\ \frac{d}{dt}\eta_t^R &= g_{u_t}^R\left(\eta_t^R\right) - \left(\hat{\chi}_t w_t \hat{\chi}_t^{-1}\right)\eta_t^R. \end{aligned} \tag{30}$$

Defining $\hat{w}_t \in \mathbb{R}^{\dim \mathfrak{g}}$ by $\mathscr{L}_{\mathfrak{g}}(\hat{w}_t) = -w_t$ in the first case and $\mathscr{L}_{\mathfrak{g}}(\hat{w}_t) = -\hat{\chi}_t w_t \hat{\chi}_t^{-1}$ (i.e. $\hat{w}_t = -Ad_{\hat{\chi}_t}\mathscr{L}_{\mathfrak{g}}^{-1}(w_t)$) in the second case, and using the superscript i to denote indifferently L or R we end up with the linearized error equation in $\mathbb{R}^{\dim \mathfrak{g}}$:

$$\frac{d}{dt}\xi_t = A_t^i \xi_t + \hat{w}_t \tag{31}$$

where A_t^i is defined by $g_{u_t}^i(\exp(\xi)) = \mathscr{L}_{\mathfrak{g}}(A_t^i \xi) + O(\|\xi_t\|^2)$ and where we have neglected terms of order $O(\|\xi_t\|^2)$ as well as terms of order $O(\|\hat{w}_t\|\|\xi_t\|)$. The latter approximation, as well as the fact that $\hat{w}_t$ is considered as white noise, are approximations that may require a proper justification in a stochastic setting. However, they are part of the standard nonadditive noise formulation of the EKF, see e.g. [21], and are difficult to avoid. Besides, the emphasis is put here on deterministic properties of the observer.

Regarding the output, we consider for instance the case of left-invariant observations, and define ξ_{t_n} through the exponential mapping (29), i.e. $\exp(\xi_{t_n}) = \eta_{t_n}^L$. Moreover, for $1 \le i \le k$ let $\hat{V}_n^i$ denote $\hat{\chi}_{t_n}^{-1} V_n^i$. The error update (21), when the LIEKF update (17) is fed with the "noisy" measurements (27) becomes

$$\begin{aligned} \left(\eta_{t_n}^L\right)^+ &= \chi_{t_n}^{-1}\hat{\chi}_{t_n}^+ \\ &= \eta_{t_n}^L \exp\left[L_n \begin{pmatrix} \left(\eta_{t_n}^L\right)^{-1} d^1 - d^1 + \hat{V}_n^1 + \left(\eta_{t_n}^L\right)^{-1} B_n^1 \\ \cdots \\ \left(\eta_{t_n}^L\right)^{-1} d^k - d^k + \hat{V}_n^k + \left(\eta_{t_n}^L\right)^{-1} B_n^k \end{pmatrix}\right]. \end{aligned} \tag{32}$$

To linearize it we proceed as follows. For $1 \le i \le k$, we have

$$\begin{aligned} &\left(\eta_{t_n}^L\right)^{-1} d^i - d^i + \hat{V}_n^i + \left(\eta_{t_n}^L\right)^{-1} B_n^i \\ &\quad = \exp_m\left(\mathscr{L}_{\mathfrak{g}}\left(\xi_{t_n}\right)\right)^{-1}\left(d^i + B_n^i\right) - d^i + \hat{V}_n^i \\ &\quad = \left(Id - \mathscr{L}_{\mathfrak{g}}(\xi)_{t_n}\right)\left(d^i + B_n^i\right) - d^i + \hat{V}_n^i + O\left\|\xi_{t_n}\right\|^2 \\ &\quad = -\mathscr{L}_{\mathfrak{g}}(\xi)_{t_n} d^i + \hat{V}_n^i + B_n^i + O\Big(\left\|\xi_{t_n}\right\|^2\Big) + O\big(\left\|\xi_{t_n}\right\| \left\|B_n^i\right\|\big) \end{aligned}$$

using a simple Taylor expansion of the matrix exponential map. Expanding similarly (32) yields

$$Id+\mathscr{L}_{\mathfrak{g}}(\xi_{t_n}^+)=Id+\mathscr{L}_{\mathfrak{g}}\left(L_n\begin{pmatrix}-\mathscr{L}_{\mathfrak{g}}(\xi_{t_n})\,d^1+\hat{V}_n^1+B_n^1\\ \cdots\\ -\mathscr{L}_{\mathfrak{g}}(\xi_{t_n})\,d^k+\hat{V}_n^k+B_n^k\end{pmatrix}\right)+T \quad (33)$$

with T terms of order $O(\|\xi_{t_n}\|^2)+O(\|\xi_{t_n}\|\|B_n\|)$. Neglecting them we finally get the following linearized error equation in $\mathbb{R}^{\dim\mathfrak{g}}$

$$\xi_{t_n}^+ = \xi_{t_n} + L_n\left(H\xi_{t_n} + \hat{V}_n + B_n\right) \quad (34)$$

where $H\in\mathbb{R}^{kN\times\dim\mathfrak{g}}$, $\hat{V}_n\in\mathbb{R}^{kN}$ and $B_n\in\mathbb{R}^{kN}$ are defined by

$$H\xi=\begin{pmatrix}-\mathscr{L}_{\mathfrak{g}}(\xi)d^1\\ \cdots\\ -\mathscr{L}_{\mathfrak{g}}(\xi)d^k\end{pmatrix},\quad \hat{V}_n=\begin{pmatrix}\hat{V}_n^1\\ \cdots\\ \hat{V}_n^k\end{pmatrix},\quad B_n=\begin{pmatrix}B_n^1\\ \cdots\\ B_n^k\end{pmatrix}.$$

Now, let $\hat{Q}_t$ reflect the trusted covariance of the modified process noise $\hat{w}_t$, and $\hat{N}_n$ the trusted covariance of the modified measurement noise $\hat{V}_n+B_n$. Note that, equations (31) and (34) mimick those of a Kalman filter designed for the following auxiliary linear system with discrete measurements: $(d/dt)x_t=A_tx_t+\hat{w}_t$, $y_n=Hx_{t_n}+\hat{V}_n+B_n$. The standard Kalman theory thus suggests to compute L_n through the Riccati equation

$$\frac{d}{dt}P_t=A_tP_t+P_tA_t^T+\hat{Q}_t,\quad P_{t_n}^+=(I-L_nH)P_{t_n},$$
$$\text{with } S_n=HP_{t_n}H^T+\hat{N}_n,\ L_n=P_{t_n}H^TS^{-1}. \quad (35)$$

C. Summary of IEFK Equations

In a deterministic context, the IEKF equations can be compactly recapped as follows:

$$\frac{d}{dt}\hat{\chi}_t=f_{u_t}(\hat{\chi}_t),\quad t_{n-1}\le t<t_n$$
$$\hat{\chi}_{t_n}^+=\hat{\chi}_{t_n}\exp\left[L_n\begin{pmatrix}\hat{\chi}_{t_n}^{-1}Y_{t_n}^1-d^1\\ \cdots\\ \hat{\chi}_{t_n}^{-1}Y_{t_n}^k-d^k\end{pmatrix}\right],\ \text{(LIEKF)}$$
$$\underline{\text{or}}\quad \hat{\chi}_{t_n}^+=\exp\left[L_n\begin{pmatrix}\hat{\chi}_{t_n}Y_{t_n}^1-d^1\\ \cdots\\ \hat{\chi}_{t_n}Y_{t_n}^k-d^k\end{pmatrix}\right]\hat{\chi}_{t_n},\ \text{(RIEKF)} \quad (36)$$

where the LIEKF (respectively, RIEKF) is to be used in the case of left (respectively, right) invariant outputs. The gain L_n is obtained in each case through the following Riccati equation:

$$\frac{d}{dt}P_t=A_tP_t+P_tA_t^T+\hat{Q}_t,\ S_n=HP_{t_n}H^T+\hat{N}_n$$
$$L_n=P_{t_n}H^TS^{-1},\ P_{t_n}^+=(I-L_nH)P_{t_n}. \quad (37)$$

As concerns the LIEKF, A_t is defined by $g_{u_t}^L(\exp(\xi))=\mathscr{L}_{\mathfrak{g}}(A_t\xi)+O(\|\xi\|^2)$, and $H\in\mathbb{R}^{kN\times\dim\mathfrak{g}}$ is defined by $H\xi=\left(-(\mathscr{L}_{\mathfrak{g}}(\xi)d^1)^T,\dots,-(\mathscr{L}_{\mathfrak{g}}(\xi)d^k)^T\right)^T$. The design matrix parameters $\hat{Q}_t,\hat{N}_n$ are freely assigned by the user. When sensor noise characteristics are known, they can provide the user with a first sensible tuning of those matrices by considering the associated "noisy" system (26), (27) of Section III-B. In this case, the matrices can be interpreted in the following way: $\hat{Q}_t\in\mathbb{R}^{\dim\mathfrak{g}\times\dim\mathfrak{g}}$ denotes the covariance of the modified process noise $\hat{w}_t=-\mathscr{L}_{\mathfrak{g}}^{-1}(w_t)$ and $\hat{N}_n$ the covariance matrix of the noise $\hat{V}_n+B_n$, $\hat{V}_n$ and B_n being defined as: $\hat{V}_n=\left(\hat{\chi}_{t_n}^{-1}V_n^1,\dots,\hat{\chi}_{t_n}^{-1}V_n^k\right)^T$, $B_n=(B_n^1,\dots,B_n^k)^T$.

As concerns the RIEKF implementation, A_t is defined by $g_{u_t}^R(\exp(\xi))=\mathscr{L}_{\mathfrak{g}}(A_t\xi)+O(\|\xi\|^2)$, and $H\in\mathbb{R}^{kN\times\dim\mathfrak{g}}$ by $H\xi=\left((\mathscr{L}_{\mathfrak{g}}(\xi)d^1)^T,\dots,(\mathscr{L}_{\mathfrak{g}}(\xi)d^k)^T\right)^T$. The design matrix parameters $\hat{Q}_t,\hat{N}_n$ are freely assigned by the user. Considering the associated "noisy" system (26)–(28) of Section III-B they can be interpreted as follows. $\hat{Q}_t$ denotes the covariance of the modified process noise $\hat{w}_t=-Ad_{\hat{\chi}_t}\mathscr{L}_{\mathfrak{g}}^{-1}(w_t)$ and $\hat{N}_n$ the covariance matrix of the noise $V_n+\hat{B}_n$, V_n and $\hat{B}_n$ being defined as: $V_n=(V_n^1,\dots,V_n^k)^T$, $\hat{B}_n=(\hat{\chi}_{t_n}B_n^1,\dots,\hat{\chi}_{t_n}B_n^k)^T$.

D. Stability Properties

The aim of the present section is to study the stability properties of the IEKF as a *deterministic* observer for the system (14), (15), or alternatively (14)–(22). We are about to prove the IEKF has *guaranteed* (local) stability properties, that rely on the error equation properties the EKF does *not* possess. The stability of an observer is defined as its ability to recover from a perturbation or an erroneous initialization:

Definition 2: Let $(x_0,t_0,t)\to X_{t_0}^t(x_0)$ denote a continuous flow on a space $\mathscr{X}$ endowed with a distance d. A flow $(z,t_0,t)\to\hat{X}_{t_0}^t(z)$ is an asymptotically stable observer of X about the trajectory $(X_{t_0}^t(x))_{t\ge t_0}$ if there exists $\varepsilon>0$ such that

$$d(x,\tilde{x})<\varepsilon\Rightarrow d\left(X_{t_0}^t(x),\hat{X}_{t_0}^t(\tilde{x})\right)\to 0 \text{ when } t\to+\infty.$$

Theorem 4 below is the main result of the paper. It is a consequence of Theorem 2 of Section II. Deyst and Price have shown in [13] the following theorem, stating sufficient conditions for the Kalman filter to be a stable observer for *linear* (time-varying) deterministic systems.

Theorem 3 (Deyst and Price, 1968): Consider the linear system $(d/dt)x_t=A_tx_t$, $y_{t_n}=Hx_{t_n}$ with $x_t\in\mathbb{R}^p$ and let $\Phi_{t_0}^t$ denote the square matrix defined by $\Phi_{t_0}^{t_0}=I_p$, $(d/dt)\Phi_{t_0}^t=A_t\Phi_{t_0}^t$. If there exist $\alpha_1,\alpha_2,\beta_1,\beta_2,\delta_1,\delta_2,\delta_3,M$ such that

i $(\Phi_{t_n}^{t_{n+1}})^T\Phi_{t_n}^{t_{n+1}}\succeq\delta_1I_p\succeq 0$,
ii $\exists q\in\mathbb{N}^*$, $\forall s>0$, $\exists G_s\in\mathbb{R}^{p\times q}$, $Q_s=G_sQ'G_s^T$ where $Q'\succeq\delta_2I_q\succeq 0$,
iii $N_n\succeq\delta_3I\succeq 0$,
iv $\alpha_1I_p\le\int_{s=t_{n-M}}^{t_n}(\Phi_s^{t_n})Q_s(\Phi_s^{t_n})^T\le\alpha_2I_p$,
v $\beta_1I_p\le\sum_{i=n-M}^{n-1}(\Phi_{t_{i+1}}^{t_n})^TH^TN_n^{-1}H(\Phi_{t_{i+1}}^{t_n})\le\beta_2I_p$.

Then the linear Kalman filter tuned with covariance matrices Q_t and N_n is an asymptotically stable observer for the Euclidean distance. More precisely there exist $\gamma_{\min},\gamma_{\max}>0$ such that $\gamma_{\min}I\preceq P_tI\preceq\gamma_{\max}I$ for all t and $(\hat{x}_t-x_t)^TP_t^{-1}(\hat{x}_t-x_t)$ has exponential decay.

The main theorem of the present paper is the extension of this linear result to the nonlinear case when the Invariant Extended Kalman Filter is used for systems of Section II.

Theorem 4: Consider the system (14), (15) [respectively (14)–(22)]. Suppose the stability conditions of the linear Kalman filter given in Theorem 3 are verified about the *true* system's trajectory χ_t (i.e. are verified for the linear system obtained by linearizing the system (14) with left (respectively, right) invariant output about χ_t). Then the Left (respectively, Right) Invariant Extended Kalman Filter estimate $\hat{\chi}_t$ defined at eq. (36) is an asymptotically stable observer of χ_t in the sense of Definition 2. Moreover, the convergence radius $\varepsilon > 0$ is valid over the whole trajectory (i.e., is independent of the initialization time t_0).

Proof: The full proof is technical and has been moved to Appendix C. The rationale is to compare the evolution of the logarithmic error ξ_t defined as $\eta_t = \exp(\xi_t)$, with its linearization. For the general EKF, the control of second and higher order terms in the error equation is difficult because: 1-they depend on the inputs u_t 2-they depend on the linearization point $\hat{\chi}_t$ 3-the estimation error impacts the gain matrices. For the IEKF, the main difficulties vanish as during the propagation step the IEKF is built for the logarithmic error ξ_t whose evolution $(d/dt)\xi_t = A_t\xi_t$ is in fact *exact* (no higher order terms) due to Theorem 2. At the update step, due to the specific form of the IEKF update, second-order terms can be *uniformly* bounded over n. And finally, due to the error equation of the IEKF, the Riccati equation depends on the estimate *only* through the matrices $\hat{Q}_t$ and $\hat{N}_t$ which affect stability in a minor way, as shown by Theorem 3. ■

The result displayed in Theorem 4 is in sharp contrast with the usual results available, that make the *highly nontrivial assumption* that the linearized system around the *estimated* trajectory is well-behaved [10], [11], [24], [25]. But this fact is almost impossible to predict as when the estimate is (even slightly) away from the true state, the Kalman gain becomes erroneous, which can in turn amplify the discrepancy between estimate and true state so that there is *no* reason the assumption should keep holding. On the other hand, when considering an actual system undergoing a realistic physical motion (see the examples below), if sufficiently many sensors are available, one can generally assert *in advance* the linearized system around the *true* trajectory possesses all the desired properties. Note also, that contrarily to [19], Q_t need not be positive definite here. The following consequence proves useful in practice.

Theorem 5: Assume the system linearized around the *true* trajectory has the following properties: the propagation matrix A_t is constant, there exist matrices B, D such that $\hat{Q} = B\bar{Q}B^T$ and $\hat{N} = D\bar{N}D^T$ with $\bar{Q}$ and $\bar{N}$ upper- and lower-bounded, with (A, H, B, D) detectable and reachable. Then the conditions of Theorem 4 are satisfied and the IEKF is asymptotically stable.

IV. Simplified Car Example

A. Considered Model

Consider a (nonholonomic) car evolving in the 2D plane. Its heading is denoted by an angle $\theta_t \in [-\pi, \pi]$ and its position by a vector $x_t \in \mathbb{R}^2$. We model the motion using the unicycle equations (see, e.g., [12])

$$\frac{d}{dt}\theta_t = \omega_t, \quad \frac{d}{dt}x_t^{(1)} = \cos(\theta_t)v_t, \quad \frac{d}{dt}x_t^{(2)} = \sin(\theta_t)v_t \tag{38}$$

where v_t is the velocity measured by an odometer and ω_t is measured through differential odometry or a single axis gyrometer. Two kinds of observations are considered

$$\bar{Y}_n = x_{t_n} \tag{39}$$

$$\underline{\text{or}}\ \bar{Y}_n^k = R(\theta_{t_n})^T (x_{t_n} - p_k), \quad k \in [1, K] \tag{40}$$

where $R(\theta)$ is a planar rotation of angle θ. Equation (39) represents a position measurement (GPS for instance) whereas (40) represents a range-and-bearing observation of a sequence of known features located at $p_k \in \mathbb{R}^2$ for $k \in [1, K]$.

B. IEKF Gain Tuning

To derive and tune the IEKF equations, we follow the methodology of Section III-B which amounts to 1-associate a "noisy "system to the original considered system, just because it allows obtaining a sensible tuning of the design matrices from an engineering viewpoint, 2-transform it into a system defined on a matrix Lie group to make it fit into our framework, 3-linearize the "noisy" equations, and 4- use the Kalman equations to tune the observer gain.

1) Associated "Noisy" System: Taking into account the possible noise in the measurements we get

$$\frac{d}{dt}\theta_t = \omega_t + w_t^\theta, \quad \frac{d}{dt}x_t^{(1)} = \cos(\theta_t)\big(v_t + w_t^l\big) - \sin(\theta_t)w_t^{tr}$$

$$\frac{d}{dt}x_t^{(2)} = \sin(\theta_t)\left(v_t + w_t^l\right) + \cos(\theta_t)w_t^{tr} \tag{41}$$

with w_t^θ the differential odometry error, w_t^l the longitudinal odometry error and w_t^{tr} the transversal shift. By letting noise enter the measurement equations we get the following two kinds of measurements:

$$\bar{Y}_n = x_{t_n} + V_n \tag{42}$$

$$\text{or}\ \bar{Y}_n^k = R(\theta_{t_n})^T (x_{t_n} - p_k) + \bar{V}_n^k, \quad k \in [1, K]. \tag{43}$$

2) Matrix Form: This system can be embedded in the matrix Lie group $SE(2)$ (see Appendix A-A) using the matrices:

$$\chi_t = \begin{pmatrix} \cos(\theta_t) & -\sin(\theta_t) & x_t^{(1)} \\ \sin(\theta_t) & \cos(\theta_t) & x_t^{(2)} \\ 0 & 0 & 1 \end{pmatrix}$$

$$\nu_t = \begin{pmatrix} 0 & -\omega_t & v_t \\ \omega_t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad w_t = \begin{pmatrix} 0 & -w_t^\theta & w_t^l \\ w_t^\theta & 0 & w_t^{tr} \\ 0 & 0 & 0 \end{pmatrix}.$$

The equation (41) governing the "noisy" system evolution writes

$$\frac{d}{dt}\chi_t = \chi_t(\nu_t + w_t) \tag{44}$$

and the observations (42) and (43), respectively, have the equivalent form

$$Y_n = \begin{pmatrix} x_{t_n} + V_n \\ 1 \end{pmatrix} = \chi_{t_n} \begin{pmatrix} 0_{2\times 1} \\ 1 \end{pmatrix} + \begin{pmatrix} V_n \\ 0 \end{pmatrix} \tag{45}$$

$$\begin{aligned} Y_n^k &= \begin{pmatrix} R\left(\theta_{t_n}\right)^T \left(x_{t_n} - p_k\right) \\ 1 \end{pmatrix} + \begin{pmatrix} \bar{V}_n^k \\ 0 \end{pmatrix} \\ &= -\,\chi_{t_n}^{-1} \begin{pmatrix} p_k \\ 1 \end{pmatrix} + \begin{pmatrix} \bar{V}_n^k \\ 0 \end{pmatrix}. \end{aligned} \tag{46}$$

The reader can verify relation (7) letting $f_{\nu_t}(\chi_t) = \chi_t \nu_t$.

3) IEKF Equations for the Left-Invariant Output (39): The LIEKF equations (36) for the associated "noisy" system (41), (42), or (44), (45) in matrix form, write

$$\frac{d}{dt}\hat{\chi}_t = \hat{\chi}_t \nu_t, \quad \hat{\chi}_{t_n}^+ = \hat{\chi}_{t_n} \exp\left(L_n \left[\hat{\chi}_{t_n}^{-1} Y_n - \begin{pmatrix} 0_{2\times 1} \\ 1 \end{pmatrix}\right]\right). \tag{47}$$

As the bottom element of $\left[\hat{\chi}_{t_n}^{-1} Y_n - \begin{pmatrix} 0_{2\times 1} \\ 1 \end{pmatrix}\right]$ is always zero we can conveniently use a reduced-dimension gain matrix $\tilde{L}_n$ defined by $L_n = \tilde{L}_n \tilde{p}$ with $\tilde{p} = (I_2, 0_{2,1})$. The gains are computed as follows. Using that $(d/dt)\chi_t^{-1} = -\chi_t^{-1}((d/dt)\chi_t)\chi_t^{-1}$ and the product rule, the left-invariant error $\eta_t = \chi_t^{-1}\hat{\chi}_t$ satisfies

$$\begin{aligned} \frac{d}{dt}\eta_t &= \eta_t \nu_t - \nu_t \eta_t - w_t \eta_t \\ \eta_{t_n}^+ &= \eta_{t_n} \exp\left(\tilde{L}_n \tilde{p}\left[\eta_{t_n}^{-1} \begin{pmatrix} 0_{2\times 1} \\ 1 \end{pmatrix} - \begin{pmatrix} 0_{2\times 1} \\ 1 \end{pmatrix} + \hat{\chi}_{t_n}^{-1} \begin{pmatrix} V_n \\ 0 \end{pmatrix}\right]\right). \end{aligned}$$

To linearize we use Appendix A. Let $\eta_t := \exp(\xi_t)$ with $\|\xi_t\| \ll 1$. We have $\nu_t = \mathscr{L}_{\mathfrak{se}(2)}(\mu_t)$ where $\mu_t = (\omega_t, v_t, 0)^T$ and $w_t = \mathscr{L}_{\mathfrak{se}(2)}(\beta_t)$ where $\beta_t = (w_t^\theta, w_t^l, w_t^{tr})^T$. Replacing η_t with $I_3 + \mathscr{L}_{\mathfrak{se}(2)}(\xi_t)$, assuming $\|\beta_t\|\|\xi_t\| \ll \|\beta_t\|$ as in the nonadditive noise standard EKF methodology, and using (58) the first equation becomes $(d/dt)\mathscr{L}_{\mathfrak{se}(2)}(\xi_t) = \mathscr{L}_{\mathfrak{se}(2)}(-ad_{\mu_t}\xi_t - \beta_t)$. Using $(\exp(u))^{-1} = \exp(-u)$ and $\exp(\cdot) \approx I_3 + \mathscr{L}_{\mathfrak{se}(2)}(\cdot)$ in the second equation and assuming $\|\xi_t\|\|V_n\| \ll \|V_n\|$ we obtain

$$\begin{aligned} \frac{d}{dt}\xi_t &= -\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\omega_t \\ -v_t & \omega_t & 0 \end{pmatrix} \xi_t - \begin{pmatrix} w_t^\theta \\ w_t^l \\ w_t^{tr} \end{pmatrix} \\ \xi_{t_n}^+ &= \xi_{t_n} - \tilde{L}_n \left[(0_{2,1}, I_2)\xi_t - R\left(\hat{\theta}_{t_n}\right)^T V_n\right] \end{aligned}$$

as the linearized error system. The gains $\tilde{L}_n$ are thus finally computed using the Riccati equation (37) with

$$A_t = -\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\omega_t \\ -v_t & \omega_t & 0 \end{pmatrix}, \quad H = (0_{2,1}, I_2)$$

$$\hat{Q}_t = \operatorname{Cov}\left[\left(w_t^\theta, w_t^l, w_t^{tr}\right)^T\right], \; \hat{N} = R\left(\hat{\theta}_{t_n}\right)\operatorname{Cov}(V_n) R\left(\hat{\theta}_{t_n}\right)^T.$$

4) IEKF Equations for the Right-Invariant Output (40): The RIEKF equations (36) for the associated "noisy" system (41)–(43), or (44), (46) in matrix form, write $(d/dt)\hat{\chi}_t = \hat{\chi}_t \nu_t$, and

$$\hat{\chi}_{t_n}^+ = \exp\left(L_n \left[\hat{\chi}_{t_n} Y_n^1 + \begin{pmatrix} p_1 \\ 1 \end{pmatrix}; \ldots; \hat{\chi}_{t_n} Y_n^K + \begin{pmatrix} p_K \\ 1 \end{pmatrix}\right]\right) \hat{\chi}_{t_n}. \tag{48}$$

As the bottom element of $\left[\hat{\chi}_{t_n}^{-1} Y_n^k + \begin{pmatrix} p_k \\ 1 \end{pmatrix}\right]$ is always zero we can conveniently use a reduced-dimension gain matrix $\tilde{L}_n$ defined by $L_n = \tilde{L}_n \tilde{p}$ with $\tilde{p} = \begin{pmatrix} [I_2, 0_{2,1}] & & \\ & \ddots & \\ & & [I_2, 0_{2,1}] \end{pmatrix}$.

To compute the gains $\tilde{L}_n$, we derive the evolution of the right-invariant error variable $\eta_t = \hat{\chi}_t \chi_t^{-1}$ between the estimate and the state of the associated "noisy" system

$$\begin{aligned} \frac{d}{dt}\eta_t &= -\left(\hat{\chi}_t w_t \hat{\chi}_t^{-1}\right)\eta_t \\ \eta_{t_n}^+ &= \exp\left(\tilde{L}_n \tilde{p} \begin{pmatrix} -\eta_{t_n}\begin{pmatrix} p_1 \\ 1 \end{pmatrix} + \begin{pmatrix} p_1 \\ 1 \end{pmatrix} + \hat{\chi}_{t_n}\begin{pmatrix} V_n^1 \\ 0 \end{pmatrix} \\ \cdots \\ -\eta_{t_n}\begin{pmatrix} p_K \\ 1 \end{pmatrix} + \begin{pmatrix} p_K \\ 1 \end{pmatrix} + \hat{\chi}_{t_n}\begin{pmatrix} V_n^K \\ 0 \end{pmatrix} \end{pmatrix}\right)\eta_{t_n}. \end{aligned} \tag{49}$$

To linearize this equation, we introduce the linearized error ξ_t defined as $\eta_t = I_3 + \mathscr{L}_{\mathfrak{se}(2)}(\xi_t)$. Introducing $\eta_t = I_3 + \mathscr{L}_{\mathfrak{se}(2)}(\xi_t)$, $\eta_t^+ = I_3 + \mathscr{L}_{\mathfrak{se}(2)}(\xi_t^+)$, $\exp(u) = I_3 + \mathscr{L}_{\mathfrak{se}(2)}(u)$ and $\eta_t^{-1} = I_3 - \mathscr{L}_{\mathfrak{se}(2)}(\xi_t)$ in (49) and removing the second-order terms in ξ_t, V_n, and w_t (as we did with the LIEKF) we obtain

$$\begin{aligned} \frac{d}{dt}\xi_t &= -\begin{pmatrix} 1 & 0_{1,2} \\ \hat{x}_t^{(2)} & \\ -\hat{x}_t^{(1)} & R(\hat{\theta}_t) \end{pmatrix} \begin{pmatrix} w_t^\theta \\ w_t^l \\ w_t^{tr} \end{pmatrix} \\ \xi_{t_n}^+ &= \xi_{t_n} - \tilde{L}_n \left[\begin{pmatrix} \begin{pmatrix} -p_1^{(2)} & 1 & 0 \\ p_1^{(1)} & 0 & 1 \end{pmatrix} \\ \cdots \\ \begin{pmatrix} -p_K^{(2)} & 1 & 0 \\ p_K^{(1)} & 0 & 1 \end{pmatrix} \end{pmatrix} \xi_{t_n} - \begin{pmatrix} R\left(\hat{\theta}_{t_n}\right) V_n^1 \\ \cdots \\ R\left(\hat{\theta}_{t_n}\right) V_n^K \end{pmatrix}\right]. \end{aligned}$$

The gains are thus computed using the Riccati equation (37) with A_t, H, $\hat{Q}$, and $\hat{N}$ defined as

$$A_t = 0_{3,3}, \quad H_n = \begin{pmatrix} \begin{pmatrix} -p_1^{(2)} & 1 & 0 \\ p_1^{(1)} & 0 & 1 \end{pmatrix} \\ \cdots \\ \begin{pmatrix} -p_K^{(2)} & 1 & 0 \\ p_K^{(1)} & 0 & 1 \end{pmatrix} \end{pmatrix}$$

$$\hat{Q}_t = \begin{pmatrix} 1 & 0_{1,2} \\ \hat{x}_t^{(2)} & \\ -\hat{x}_t^{(1)} & R(\hat{\theta}_t) \end{pmatrix} \operatorname{Cov}\begin{pmatrix} w_t^\theta \\ w_t^l \\ w_t^{tr} \end{pmatrix} \begin{pmatrix} 1 & 0_{1,2} \\ \hat{x}_t^{(2)} & \\ -\hat{x}_t^{(1)} & R(\hat{\theta}_t) \end{pmatrix}^T$$

$$\hat{N}_n = \begin{pmatrix} R\left(\hat{\theta}_{t_n}\right)\operatorname{Cov}(N^1) R\left(\hat{\theta}_{t_n}\right)^T & & 0 \\ & \ddots & \\ 0 & & R\left(\hat{\theta}_{t_n}\right)\operatorname{Cov}(N^K) R\left(\hat{\theta}_{t_n}\right)^T \end{pmatrix}.$$

C. Stability Properties of the IEKF Viewed as a Nonlinear Observer for the Simplified Car

1) Stability of the IEKF for the Left-Invariant Output (39):

Proposition 3: If there exists $v_{\max}, v_{\min} > 0$ such that the displacement satisfies $\|x_{t_{n+1}} - x_{t_n}\| \geq v_{\min} > 0$ and the input velocity satisfies $\nu_t \leq v_{\max}$ then the LIEKF (47) derived in Section IV-B3 is an asymptotically stable observer in the sense of Definition 2 about *any* trajectory of system (38)–(39).

The proof is a verification of the hypotheses of Theorem 4 and has been moved to Appendix D. Note that it seems very difficult to improve on the assumptions: if the car is at the same place each time its position is measured, the heading θ_t becomes unobservable. And regarding the upper bound, the velocity is bounded in practice.

2) Stability of the IEKF for the Right-Invariant Output (40):

Proposition 4: If at least two distinct points are observed then the RIEKF (48) derived in Section IV-B-4, is an asymptotically stable observer in the sense of Definition 2 about *any* bounded trajectory of the system (38)–(40).

Proof: According to Theorem 5 it is sufficient to show that in this case the observation matrix H is full-rank, i.e., of rank 3. This is obvious as the position and the heading are easily computed from the observation of two vectors at known locations. ■

D. Simulations

The IEKF described in Section IV-B has been implemented and compared to a standard EKF for the experimental setting described by Fig. 1. The car drives along a 10-meter diameter circle for 40 seconds with high rate odometer measurements (100 Hz) and low rate GPS measurements (1 Hz).

The equations of the (L)IEKF can be found above. The conventional EKF is based on the linear error $e_t = \begin{pmatrix} \theta_t - \hat{\theta}_t \\ x_t^{(1)} - \hat{x}_t^{(1)} \\ x_t^{(2)} - \hat{x}_t^{(2)} \end{pmatrix}$ yielding the linearized matrices $F_t = \begin{pmatrix} 0 & 0 & 0 \\ -\sin(\hat{\theta}_t)v_t & 0 & 0 \\ \cos(\hat{\theta}_t)v_t & 0 & 0 \end{pmatrix}$ and $H_n = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Both filters are tuned with the same design parameters (which can be interpreted as odometer and GPS noise covariances) $N = I_2$, and $Q = \mathrm{diag}((\pi/180)^2, 10^{-4}, 10^{-4})$ i.e., moderate angular velocity uncertainty and highly precise linear velocity. The simulation is performed for two initial values of the heading error: 1° and 45° while the initial position is always assumed known. The covariance matrix P_0 is consistent with the initial error (it encodes a standard deviation of the heading of 1° and 45° respectively).

The results are displayed on Fig. 1. We see that for small initial errors both filters behave similarly for a long time, but for larger errors they soon behave differently, and we see the IEKF, whose design has been adapted to the specific structure of the system, completely outperforms the EKF.

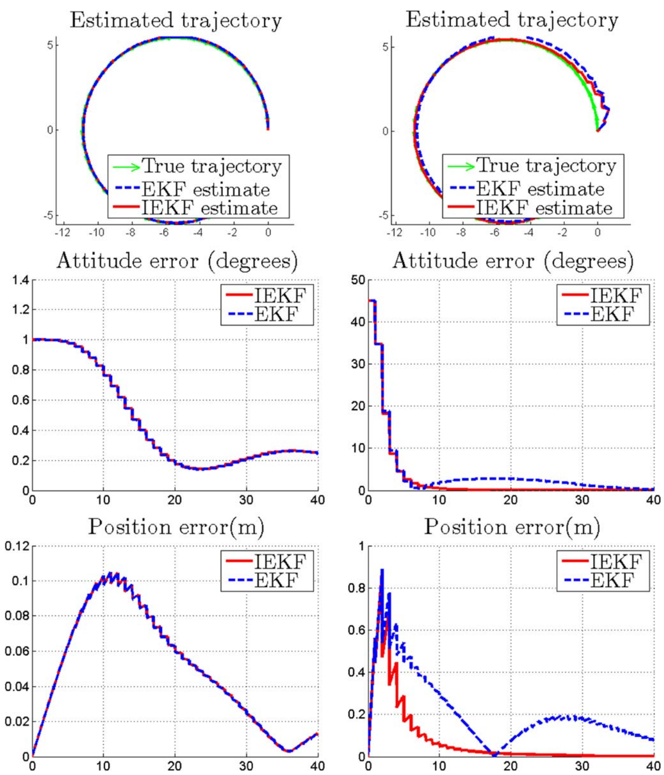

Fig. 1. The heading and position of the car are estimated through EKF and IEKF with high rate odometry and low rate GPS measurements. Top plots illustrate the experimental setting and display the estimated trajectories, middle plots display the heading errors and bottom plots the position errors. As the starting point is assumed known in this simulation, the initial values of the latter are zero. But it increases afterwards due to initial heading error. **Left column**: small initial angle error (1°). We see EKF and IEKF behave similarly (at least for a long time) as propagation steps are identical. **Right column**: large initial angle error (45°). The behaviors rapidly become different, and the EKF is outperformed. Due to its righteous use of the system's nonlinearities, the IEKF keeps ensuring rapid estimation error decrease.

V. Navigation on Flat Earth

In this example, we estimate the orientation, velocity, and position of a rigid body in space from inertial sensors and relative observations of points having known locations (the setting of [26] but with the state including the position). To our knowledge, this is the first time the invariant observer on Lie groups based approach is applied to this full navigation problem with landmarks, except for our preliminary conference paper [4]. Indeed, this example does not fit into the usual framework leading to autonomous errors (unless we discard the position estimate as in [26]) but thanks to Theorem 1 we see it still leads to an autonomous error equation. This allows the IEKF observer to possess provable convergence properties. Note the problem at hand is different from the navigation problems using magnetometers, and velocity and position measurements of the GPS [16], [17].

Of course, the EKF, to be more precise its appropriate variant the multiplicative (M)EKF [21], is the state of the (industrial) art for this navigation example, due to its good performances, and easy tuning based on sensors' noise covariances. But to our best knowledge it is nowhere proved to possess stability properties as a nonlinear observer, and simulations below even indicate it may diverge in some situations whereas the IEKF converges. The computations require basic formulas recalled in Appendix A.

A. Considered Model

We consider here the more complicated model of a vehicle evolving in the 3D space and characterized by its attitude R_t, velocity v_t and position x_t. The vehicle is endowed with accelerometers and gyroscopes whose measures are denoted, respectively, by u_t and angular velocity ω_t. The dynamics read

$$\frac{d}{dt}R_t = R_t(\omega_t)_\times, \quad \frac{d}{dt}v_t = g + R_t u_t, \quad \frac{d}{dt}x_t = v_t \tag{50}$$

where $(\omega)_\times$ denotes the 3×3 skew-symmetric matrix associated with the cross product with ω, that is, for any $b \in \mathbb{R}^3$ we have $(\omega)_\times b = \omega \times b$. Observations of the relative position of known features (using for instance a depth camera) are considered

$$\left(Y_n^1, \ldots, Y_n^k\right) = \left(R_{t_n}^T\left(p_1 - x_{t_n}\right), \ldots, R_{t_n}^T\left(p_k - x_{t_n}\right)\right) \tag{51}$$

where $(p_1, \ldots, p_k)$ denote the (assumed known) position of the features in the earth-fixed frame.

B. IEKF Gain Tuning

To derive and tune the IEKF equations, we follow the methodology of Section III-B which amounts to 1-associate a "noisy "system to the original considered system, just because it allows obtaining a sensible tuning of the design matrices from an engineering viewpoint, 2-transform it into a system defined on a matrix Lie group to make it fit into our framework, 3-linearize the "noisy" equations, and 4-use the Kalman equations to tune the observer gain.

1) Associated "Noisy" System: By merely introducing noise in the accelerometers' and gyrometers' measurements we obtain the well-known equations [14]

$$\frac{d}{dt}R_t = R_t\left(\omega_t + w_t^\omega\right)_\times, \ \frac{d}{dt}v_t = g + R_t\left(u_t + w_t^u\right), \ \frac{d}{dt}x_t = v_t. \tag{52}$$

Letting additive noise pollute the observations (the sensor being in the body frame), we get

$$\left(Y_n^1, \ldots, Y_n^k\right) = \left(R_{t_n}^T(p_1 - x_{t_n}) + V_n^1, \ldots, R_{t_n}^T(p_k - x_{t_n}) + V_n^k\right) \tag{53}$$

where $V_n, V_n^1, \ldots, V_n^k$ are noises in $\mathbb{R}^3$.

2) Matrix Form: As already noticed in the preliminary work [4], the system (52) can be embedded in the group of double homogeneous matrices $SE_2(3)$ (see Appendix A-B) using the matrices χ_t, w_t and function $f_{\omega,u}$

$$\chi_t = \begin{pmatrix} R_t & v_t & x_t \\ 0_{3,1} & 1 & 0 \\ 0_{3,1} & 0 & 1 \end{pmatrix}, \quad w_t = \begin{pmatrix} (w_t^\omega)_\times & w_t^u & 0_{3,1} \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{pmatrix}$$

$$f_{\omega,u} : \begin{pmatrix} R & v & x \\ 0_{3,1} & 1 & 0 \\ 0_{3,1} & 0 & 1 \end{pmatrix} \to \begin{pmatrix} R(\omega)_\times & g + Ru & v \\ 0_{3,1} & 0 & 0 \\ 0_{3,1} & 0 & 0 \end{pmatrix}.$$

The equation of the dynamics becomes

$$\frac{d}{dt}\chi_t = f_{\omega_t,u_t}(\chi_t) + \chi_t w_t \tag{54}$$

and the observations (53) have the equivalent forms

$$\left(Y_n^1, \ldots, Y_n^k\right) = \left(\chi_{t_n}^{-1}\begin{pmatrix} p_1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} V_n^1 \\ 0 \\ 0 \end{pmatrix}, \ldots, \chi_{t_n}^{-1}\begin{pmatrix} p_k \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} V_n^k \\ 0 \\ 0 \end{pmatrix}\right). \tag{55}$$

Proposition 5: The matricial function f_{ω_t,u_t} is neither left nor right invariant. However, the reader can verify relation (7) which is easy to derive.

3) IEKF Equations: The RIEKF (36) for the associated "noisy" system (52) with right-invariant "noisy" observations (53), or (54), (55) in matrix form, reads

$$\frac{d}{dt}\hat{\chi}_t = f_{\omega_t,u_t}(\hat{\chi}_t), \quad \hat{\chi}_t^+ = \exp\left(L_n \begin{pmatrix} \hat{\chi}_t Y_n^1 \\ \ldots \\ \hat{\chi}_t Y_n^k \end{pmatrix}\right)\hat{\chi}_t.$$

As the two last entries of each matrix $\hat{\chi}_t^{-1}Y_n^j$ are always zero, one we can conveniently use a reduced-dimension gain matrix $\tilde{L}_n$ defined by $L_n = \tilde{L}_n\tilde{p}$ with $\tilde{p} = \begin{pmatrix} [I_3, 0_{3,2}] & & \\ & \ddots & \\ & & [I_3, 0_{3,2}] \end{pmatrix}$. The right-invariant error is $\eta_t = \hat{\chi}_t\chi_t^{-1}$ and its evolution reads

$$\frac{d}{dt}\eta_t = f_{\omega_t,u_t}(\eta_t) - \eta_t f_{\omega_t,u_t}(I_5) - (\hat{\chi}_t w_t \hat{\chi}^{-1})\eta_t \tag{56}$$

$$\eta_{t_n}^+ = \exp\left(\tilde{L}_n\tilde{p}\begin{pmatrix} \eta_{t_n}\begin{pmatrix} p_1 \\ 0 \\ 1 \end{pmatrix} + \hat{\chi}_{t_n}\begin{pmatrix} V_n^1 \\ 0_{2,1} \end{pmatrix} \\ \ldots \\ \eta_{t_n}\begin{pmatrix} p_k \\ 0 \\ 1 \end{pmatrix} + \hat{\chi}_{t_n}\begin{pmatrix} V_n^k \\ 0_{2,1} \end{pmatrix} \end{pmatrix}\right)\eta_{t_n}. \tag{57}$$

To linearize this equation, we introduce the linearized error ξ_t by replacing η_t with $I_3 + \mathscr{L}_{\mathfrak{se}_2(3)}(\xi_t)$. Using the first-order approximations $\eta_t = I_3 + \mathscr{L}_{\mathfrak{se}_2(3)}(\xi_t)$, $\eta_t^+ = I_3 + \mathscr{L}_{\mathfrak{se}_2(3)}(\xi_t^+)$, $\exp(u) = I_3 + \mathscr{L}_{\mathfrak{se}_2(3)}(u)$ and $\eta_t^{-1} = I_3 - \mathscr{L}_{\mathfrak{se}_2(3)}(\xi_t)$ in (56), (57), assuming $\|\xi_t\|\|w_t\| \ll \|w_t\|$, $\|\xi_t\|\|V_n\| \ll \|V_n\|$ we get

$$\frac{d}{dt}\xi_t = \begin{pmatrix} 0_{3,3} & 0_{3,3} & 0_{3,3} \\ (g)_\times & 0_{3,3} & 0_{3,3} \\ 0_{3,3} & I_3 & 0_{3,3} \end{pmatrix}\xi_t - \begin{pmatrix} \hat{R}_t & 0_{3,3} & 0_{3,3} \\ (\hat{v}_t)_\times\hat{R}_t & \hat{R}_t & 0_{3,3} \\ (\hat{x}_t)_\times\hat{R}_t & 0_{3,3} & \hat{R}_t \end{pmatrix} w_t$$

$$\xi_{t_n}^+ = \xi_{t_n} - \tilde{L}_n\left[\begin{pmatrix} (p_1)_\times & 0_{3,3} & -I_3 \\ & \ldots & \\ (p_k)_\times & 0_{3,3} & -I_3 \end{pmatrix}\xi_{t_n} - \begin{pmatrix} \hat{R}_{t_n}V_n^1 \\ \ldots \\ \hat{R}_{t_n}V_n^k \end{pmatrix}\right].$$

The gains $\tilde{L}_n$ are computed using the Riccati equation (37) and matrices $A_t, H, \hat{Q}$, and $\hat{N}$ defined as

$$A_t = \begin{pmatrix} 0_{3,3} & 0_{3,3} & 0_{3,3} \\ (g)_\times & 0_{3,3} & 0_{3,3} \\ 0_{3,3} & I_3 & 0_{3,3} \end{pmatrix}, \quad H = \begin{pmatrix} (p_1)_\times & 0_{3,3} & -I_3 \\ & \ldots & \\ (p_k)_\times & 0_{3,3} & -I_3 \end{pmatrix}$$

$$\hat{Q} = \begin{pmatrix} \hat{R}_t & 0_{3,3} & 0_{3,3} \\ (\hat{v}_t)_\times\hat{R}_t & \hat{R}_t & 0_{3,3} \\ (\hat{x}_t)_\times\hat{R}_t & 0_{3,3} & \hat{R}_t \end{pmatrix}\mathrm{Cov}(w_t)\begin{pmatrix} \hat{R}_t & 0_{3,3} & 0_{3,3} \\ (\hat{v}_t)_\times\hat{R}_t & \hat{R}_t & 0_{3,3} \\ (\hat{x}_t)_\times\hat{R}_t & 0_{3,3} & \hat{R}_t \end{pmatrix}^T$$

$$\hat{N} = \begin{pmatrix} \hat{R}_{t_n}\mathrm{Cov}\left(V_n^1\right)\hat{R}_{t_n}^T & & \\ & \ddots & \\ & & \hat{R}_{t_n}\mathrm{Cov}\left(V_n^k\right)\hat{R}_{t_n}^T \end{pmatrix}.$$

C. Stability Properties of the IEKF Viewed as a Nonlinear Observer for the Navigation Example

Theorem 6: If three noncollinear points are observed, then the IEKF whose equations are derived in Section V-B3 is an asymptotically stable observer for the system (50), (51) in the sense of Definition 2 about *any* bounded trajectory.

Proof: According to Theorem 5 we only have to ensure the couple (A, H) is observable. Integrating the propagation on one step we obtain the discrete propagation matrix $\Phi = \begin{pmatrix} I_3 & 0_{3\times 3} & 0_{3\times 3} \\ t(g)_\times & I_3 & 0_{3\times 3} \\ (1/2)t^2(g)_\times & tId_{3\times 3} & I_3 \end{pmatrix}$. The observation matrix is denoted H. We will show that $[H; H\Phi]$ has rank 9. We can keep only the raws corresponding to the observation of three noncollinear features p_1, p_2, p_3 and denote the remaining matrix by $\mathscr{H}_1$. Matrices $\mathscr{H}_2$ and $\mathscr{H}_3$, obtained using elementary operations on the columns of $\mathscr{H}_1$, have a rank inferior or equal to the rank of $\mathscr{H}_1$

$$\mathscr{H}_1 = \begin{pmatrix} (p_1)_\times & 0_3 & -I_3 \\ (p_2)_\times & 0_3 & -I_3 \\ (p_3)_\times & 0_3 & -I_3 \\ (p_1)_\times - \frac{1}{2}t^2(g)_\times & -tI_3 & -I_3 \\ (p_2)_\times - \frac{1}{2}t^2(g)_\times & -tI_3 & -I_3 \\ (p_3)_\times - \frac{1}{2}t^2(g)_\times & -tI_3 & -I_3 \end{pmatrix}$$

$$\mathscr{H}_2 = -\begin{pmatrix} -(p_1)_\times & 0_3 & I_3 \\ -(p_2)_\times & 0_3 & I_3 \\ -(p_3)_\times & 0_3 & I_3 \\ \frac{1}{2}t^2(g)_\times & tI_3 & 0_3 \\ \frac{1}{2}t^2(g)_\times & tI_3 & 0_3 \\ \frac{1}{2}t^2(g)_\times & tI_3 & 0_3 \end{pmatrix}$$

$$\mathscr{H}_3 = \begin{pmatrix} (p_1 - p_3)_\times & 0_3 & 0_3 \\ (p_2 - p_3)_\times & 0_3 & 0_3 \\ \frac{1}{2}t^2(g)_\times & tI_3 & 0_3 \\ (p_3)_\times & 0_3 & -I_3 \end{pmatrix}.$$

The diagonal blocks $\begin{pmatrix} -(p_1 - p_3)_\times & 0_{3\times 3} \\ -(p_2 - p_3)_\times & 0_{3\times 3} \end{pmatrix}$, tI_3 and I_3 have rank 3 thus the full matrix has rank 9. ■

D. Simulations

The IEKF described in Section V-B3 has been implemented and compared to a state of the art multiplicative EKF [21] for the experimental setting described by Fig. 2 (top plots). The vehicle drives a 10-meter diameter circle (green arrows) in 30 seconds and observes three features (black circles) every second while receiving high-frequency inertial measurements (100 Hz). The equations of the IEKF have already been detailed. The error variable to be linearized for the multiplicative (M)EKF is $e_t = (\hat{R}_t R_t^{-1}, \hat{v}_t - v_t, \hat{x}_t - x_t)$. As $\hat{R}_t R_t^{-1}$ is not a vector variable, it is linearized using the first-order expansion $\hat{R}_t R_t^{-1} = I_3 + (\zeta_t)_\times, \zeta_t \in \mathbb{R}^3$ [21]. The linearized error variable is thus a vector $\varepsilon_t = (\zeta, dx, dv) \in \mathbb{R}^9$. Expanding the propagation and observation steps up to the first order in ε_t give the classical F_t and H_n matrices used in the Riccati Equation

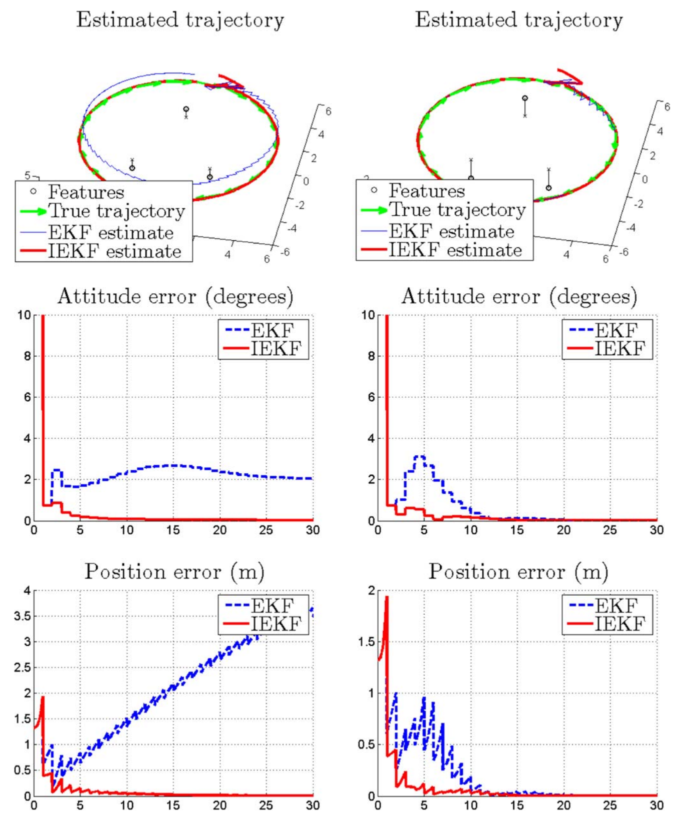

Fig. 2. Aided inertial navigation based on high rate accelerometers' and gyrometers' measurements and low rate observation of known landmarks. We also displayed the orthogonal projection of the landmarks on the plane containing the trajectory (black crosses) to help visualizing the 3D position of the landmarks, which is the same in both experiments. Top plots illustrate the experimental setting and display the EKF and IEKF estimates. Middle plots display the attitude errors and bottom plots the position errors. **Left column**: the tuning of Q is tight $(Q = Q_1)$ due to highly precise inertial sensors. This causes robustness issues: the gains of the EKF decrease obviously faster than they should during the transitory phase. When the position estimate is impacted, the gains have become too small to correct the errors, leading to filter's divergence. IEKF ensures rapid decrease to 0 of the estimation error with identical tuning. **Right plot**: Q is inflated $(Q = Q_2)$. This classical engineering trick prevents the EKF to diverge but IEKF still prevails in terms of time of convergence.

of the MEKF

$$F_t = \begin{pmatrix} 0_3 & 0_3 & 0_3 \\ -(\hat{R}_t u_t)_\times & 0_3 & 0_3 \\ 0_3 & I_3 & 0_3 \end{pmatrix}$$

$$H_n = \begin{pmatrix} -\hat{R}_{t_n}^T \left(p^1 - \hat{x}_{t_n}\right)_\times & 0_3 & \hat{R}_{t_n}^T \\ & \dots & \\ -\hat{R}_{t_n}^T \left(p^k - \hat{x}_{t_n}\right)_\times & 0_3 & \hat{R}_{t_n}^T \end{pmatrix}.$$

We use the following design parameters in two distinct simulations, with same N but two different matrices Q:

$$N = \begin{pmatrix} 10^{-2}I_3 & 0_3 & 0_3 \\ 0_3 & 10^{-2}I_3 & 0_3 \\ 0_3 & 0_3 & 10^{-2}I_3 \end{pmatrix}$$

$$Q_1 = \begin{pmatrix} 10^{-8}I_3 & 0_3 & 0_3 \\ 0_3 & 10^{-8}I_3 & 0_3 \\ 0_3 & 0_3 & 0_3 \end{pmatrix}$$

$$Q_2 = \begin{pmatrix} 10^{-4}I_3 & 0_3 & 0_3 \\ 0_3 & 10^{-4}I_3 & 0_3 \\ 0_3 & 0_3 & 0_3 \end{pmatrix}.$$

The initial errors are the same for both simulations: 15 degrees for attitude and 1 meter for position standard deviations. The small "process noise" matrix Q_1, yet consistent with the actual noise amplitude of high-precision inertial sensors, has been deliberately chosen here to challenge EKF-like methods: the corresponding gains are small so the errors introduced during the transitory phase due to nonlinearities in the initial errors can never be corrected. Note this would not be an issue if the system was linear: the estimation errors and filter gains would decrease simultaneously. The problem is that the error does not decrease as fast as predicted by the *linear* Kalman theory. As shown by the plots of the left column of Fig. 2 (top plots) it makes the EKF even diverge! This is probably the simplest way to make the EKF fail in a navigation problem and this is purely a problem of nonlinearity as no noise has been added whatsoever. Still on the left column, we see the IEKF is not affected by the problem, due to its appropriate nonlinear structure. In particular, the attitude and position errors go to zero in accordance with Theorem 6.

Usually, engineers get around those convergence problems by artificially inflating the "process noise" matrix Q (see also [24]). This classical solution, sometimes referred to as robust tuning, is illustrated here by using Q_2 instead. The results are displayed on the right column of Fig. 2. They illustrate the fact the EKF, as an observer, can be improved through a proper tuning, although still much slower to converge than the IEKF. But this raises issues: Q and N have been chosen for a specific trajectory with no guarantee regarding robustness. Moreover, these matrices admit a physical interpretation (the accuracy of the sensors) and arbitrarily changing them by several orders of magnitude is a renouncement to use this precious information when available. In turn, this makes the matrix P_t loose its interpretability as an indication of the observer's accuracy in response to the sensors' trusted accuracy. For this relevant problem, we thus see the IEKF turns out to be a viable alternative to the EKF thanks to its *guaranteed* properties, and to its convincing experimental behavior reflecting way better performances than the EKF, even for challenging Q and N.

VI. Conclusion

The Invariant EKF, when used as a deterministic observer for a novel class of problems on Lie groups, is shown to possess theoretical stability guarantees under the simple and natural hypotheses of the linear case, a feature the EKF has never been proved to share so far. Simulations confirm the IEKF is an appealing alternative indeed, as it is always superior to the EKF and outperforms it in challenging situations, while remaining similar to the EKF in terms of tuning, implementation, and computational load. By construction, it also shares its first-order optimality properties in the presence of noise.

Appendix A
Matrix Lie Groups Useful Formulas

A matrix Lie group G is a subset of square invertible $N \times N$ matrices $\mathscr{M}_N(\mathbb{R})$ verifying the following properties

$$Id \in G, \quad \forall g \in G,\ g^{-1} \in G, \quad \forall a,\ b \in G,\ ab \in G.$$

If $\gamma(t)$ is a curve over G with $\gamma(0) = I_N$, then its derivative at $t = 0$ necessarily lies in a subset $\mathfrak{g}$ of $\mathscr{M}_N(\mathbb{R})$. $\mathfrak{g}$ is a vector space and it is called the Lie algebra of G and has same dimension as G. Thanks to a *linear* map denoted by $\mathscr{L}_{\mathfrak{g}} : \mathbb{R}^{\dim \mathfrak{g}} \to \mathfrak{g}$, one can adantageously identify $\mathfrak{g}$ to $\mathbb{R}^{\dim \mathfrak{g}}$. Besides, the vector space $\mathfrak{g}$ can be mapped to the matrix Lie group G through the classical matrix exponential $\exp_m$. Thus, $\mathbb{R}^{\dim \mathfrak{g}}$ can be mapped to G through the Lie exponential map defined by $\exp(\xi) := \exp_m(\mathscr{L}_{\mathfrak{g}}(\xi))$ for $\xi \in \mathbb{R}^{\dim \mathfrak{g}}$. We have thus $\exp(\xi) = I_N + \mathscr{L}_{\mathfrak{g}}(\xi) + O(\xi^2) \in \mathscr{M}_N(\mathbb{R})$. For all $\zeta \in \mathbb{R}^{\dim \mathfrak{g}}$ the adjoint matrix $ad_\zeta \in \mathbb{R}^{\dim \mathfrak{g} \times \dim \mathfrak{g}}$ is very useful to linearize equations over Lie groups. It satisfies $ad_\zeta \xi = -ad_\xi \zeta$ and

$$\mathscr{L}_{\mathfrak{g}}(\xi)\mathscr{L}_{\mathfrak{g}}(\zeta) - \mathscr{L}_{\mathfrak{g}}(\zeta)\mathscr{L}_{\mathfrak{g}}(\xi) = \mathscr{L}_{\mathfrak{g}}(-ad_\zeta \xi). \tag{58}$$

A. Group of Direct Planar Isometries $SE(2)$

$G = SE(2) := \left\{ \begin{pmatrix} R(\theta) & x \\ 0_{1,2} & 1 \end{pmatrix} ; \theta \in \mathbb{R},\ x \in \mathbb{R}^2 \right\}$, where $R(\theta)$ is a planar rotation matrix of angle θ. $\mathfrak{g} = \mathfrak{se}(2) = \left\{ \begin{pmatrix} 0 & -\alpha & u_1 \\ \alpha & 0 & u_2 \\ 0 & 0 & 0 \end{pmatrix} ; \begin{pmatrix} \alpha \\ u_1 \\ u_2 \end{pmatrix} \in \mathbb{R}^3 \right\}$. Let $\zeta = \begin{pmatrix} \alpha \\ u_1 \\ u_2 \end{pmatrix}$.

$\mathscr{L}_{\mathfrak{se}(2)}(\zeta) = \begin{pmatrix} 0 & -\alpha & u_1 \\ \alpha & 0 & u_2 \\ 0 & 0 & 0 \end{pmatrix}$, $ad_\zeta = \begin{pmatrix} 0 & 0 & 0 \\ u_2 & 0 & -\omega_t \\ -u_1 & \omega_t & 0 \end{pmatrix}$.

The Lie exponential writes $\exp(\zeta) = \begin{pmatrix} R(\alpha) & x \\ 0_{1,2} & 1 \end{pmatrix}$ where $x = \begin{pmatrix} \sin(\alpha)/\alpha & -(1-\cos(\alpha)/\alpha) \\ (1-\cos(\alpha))/\alpha & \sin(\alpha)/\alpha \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$.

B. Group of Double Direct Spatial Isometries $SE_2(3)$

$$G = SE_2(3) = \left\{ \begin{pmatrix} R & v & x \\ 0_{1,3} & 1 & 0 \\ 0_{1,3} & 0 & 1 \end{pmatrix}, R \in SO(3), v, x \in \mathbb{R}^3 \right\},$$

$\mathfrak{g} = \mathfrak{se}_2(3) = \left\{ \begin{pmatrix} (\xi)_\times & u & y \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{pmatrix}, \xi, u, y \in \mathbb{R}^3 \right\}$. $\mathscr{L}_{\mathfrak{se}_2(3)} \begin{pmatrix} \xi \\ u \\ y \end{pmatrix} = \begin{pmatrix} (\xi)_\times & u & y \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{pmatrix}$. $\exp((\xi^T, u^T, y^T)^T) = I_5 + S + ((1 - \cos(\|\xi\|))/\|\xi\|^2)S^2 + ((\|\xi\| - \sin(\|\xi\|))/\|\xi\|^3)S^3$, where $S = \mathscr{L}_{\mathfrak{se}_2(3)}(\xi, u, y)^T$.

Appendix B
Further Explanation and Proof of the Log-Linear Property

The definition of A_t through $g_{u_t}(\exp(\xi)) = \mathscr{L}_{\mathfrak{g}}(A_t\xi) + O(\|\xi\|^2)$ is very convenient in practice as shown in Sections IV and V, but its existence requires a quick theoretical explanation as follows. For any $\xi \in \mathfrak{g}$ we have $\exp(\xi)^{-1} g_{u_t}(\exp(\xi)) \in \mathfrak{g}$ thus the function $t_L : \xi \to \mathscr{L}_{\mathfrak{g}}^{-1}(\exp(\xi)^{-1} g_{u_t}(\exp(\xi)))$ goes from $\mathbb{R}^{\dim \mathfrak{g}}$ to $\mathbb{R}^{\dim \mathfrak{g}}$. Let A_t denote its differential at zero. As $t_L(0) = 0$ we have: $t_L(\xi) = A_t\xi + \circ(\|\xi\|)$. Thus $g_{u_t}(\exp(\xi)) = \exp(\xi)\mathscr{L}_{\mathfrak{g}}(A_t\xi + \circ(\|\xi\|)) = (I_d + \mathscr{L}_{\mathfrak{g}}(\xi))\mathscr{L}_{\mathfrak{g}}(A_t\xi + \circ(\|\xi\|)) = \mathscr{L}_{\mathfrak{g}}(A_t\xi) + \circ(\|\xi\|)$.

The proof of Theorem 2 is based upon three lemmas.

Lemma 1: Consider the system (12) and let $\bar{\chi}_t$ denote a particular solution. Consider the condition

$$\forall (u, a, b) \in G \quad g_u(ab) = a g_u(b) + g_u(a) b. \tag{59}$$

We have the following properties:

- The function $g_u^R(\eta) = f_{u_t}(\eta) - \eta f_{u_t}(Id)$ verifies (59) and all the solutions of (12) have the form $\chi_t = \eta_t^R \bar{\chi}_t$, where η_t^R verifies $(d/dt)\eta_t^R = g_{u_t}^R(\eta_t^R)$.
- The function $g_{u_t}^L(\eta) = f_{u_t}(\eta) - f_{u_t}(Id)\eta$ verifies (59) and all the solutions of (12) have the form $\chi_t = \bar{\chi}_t \eta_t^L$, where η_t^L verifies $(d/dt)\eta_t^L = g_{u_t}^L(\eta_t^L)$.

The verification of these two properties is trivial. The functions g_{u_t} governing the errors propagation turn out to possess an intriguing property.

Lemma 2: Let Φ_t be the flow (that is the solution at time t associated to a given initial condition) associated to the system $(d/dt)\eta_t = g_{u_t}(\eta_t)$, where g_{u_t} verifies (59). Then

$$\forall \eta_0, \eta_0' \in G, \ \Phi_t(\eta_0 \eta_0') = \Phi_t(\eta_0)\Phi_t(\eta_0').$$

Proof: We simply have to see that $\Phi_t(\eta_0)\Phi_t(\eta_0')$ is solution of the system $(d/dt)\eta_t = g_{u_t}(\eta_t)$

$$\frac{d}{dt}[\Phi_t(\eta_0)\Phi_t(\eta_0')] = g_{u_t}(\Phi_t(\eta_0))\Phi_t(\eta_0') + \Phi_t(\eta_0) g_{u_t}(\Phi_t(\eta_0'))$$
$$= g_{u_t}(\Phi_t(\eta_0)\Phi_t(\eta_0')).$$

■

An immediate recursion gives then:

Lemma 3: We have furthermore

$$\forall \eta_0 \in G, \quad \forall p \in \mathbb{Z}, \ \Phi_t(\eta_0^p) = \Phi_t(\eta_0)^p.$$

Lemmas 2 and 3 indicate the behavior of the flow infinitely close to Id dictates its behavior arbitrarily far from it, as the flow commutes with exponentiation. The use of the exponential thus allows deriving an infinitesimal version of the Lemma 3, which is an equivalent formulation of Theorem 2.

Theorem 7: Let Φ_t be the flow associated to the system $(d/dt)\eta_t = g_{u_t}(\eta_t)$ satisfying (59). We have

$$\Phi_t(\exp(\xi_0)) = \exp(F_t \xi_0) \tag{60}$$

where F_t is the solution of the matrix equation $F_0 = Id$, $(d/dt)F_t = A_t F_t$.

Proof: The function $\exp^{-1} \circ \Phi_t \circ \exp$ goes from $\mathbb{R}^{\dim \mathfrak{g}}$ to $\mathbb{R}^{\dim \mathfrak{g}}$. Its differental with respect to ξ, defined in the classical sense, is denoted by F_t. The Taylor rest of first-orer expansion of $\exp^{-1} \circ \Phi_t \circ \exp$ is denoted by $r_t(\xi)$, with $r_t(\xi) = \circ(\|\xi\|)$

$$\exp^{-1} \circ \Phi_t \circ \exp = F_t \xi + r_t(\xi). \tag{61}$$

Thanks to Lemma 3 we have, for any $n \in \mathbb{N}$, $\Phi_t(e^{\xi_0}) = \Phi_t([e^{(1/n)\xi_0}]^n) = \Phi_t(e^{(1/n)\xi_0})^n = [e^{(1/n)F_t\xi_0 + r_t((1/n)\xi_0)}]^n$, where (61) has been used in the last equality. This ensures in turn $\Phi_t(e^{\xi_0}) = e^{F_t\xi_0 + n r_t((1/n)\xi_0)}$ and letting $n \to \infty$ we get (60). We still have to show $(d/dt)F_t = A_t F_t$. Differentiating both sides of (60) with respect to time we obtain $g_{u_t}(\Phi_t(e_0^{\xi})) = (d/dt)\exp(F_t\xi_0)$, by definition of the flow Φ_t. Reintroducing (60) we get $g_{u_t}(\exp(F_t\xi_0)) = (d/dt)\exp(F_t\xi_0) = D\exp_{F_t\xi_0}((d/dt)F_t\xi_0)$, where $D\exp_{F_t\xi_0} : \mathfrak{g} \to \mathscr{M}_N(\mathbb{R})$ is the differential of the exp function at $F_t\xi_0$, exp being considered a function from $\mathfrak{g}$ to $\mathscr{M}_n(\mathbb{R})$ (note that, this differential actually takes values in the smaller space $TG_{\exp(F_t\xi_0)} \subset \mathscr{M}_n(\mathbb{R})$, i.e., the tangent space to G at $\exp(F_t\xi_0)$). A first-order expansion in ξ_0 using matrix A_t gives: $\mathscr{L}_\mathfrak{g}(A_t F_t \xi_0) + \circ(\|\xi_0\|) = \mathscr{L}_\mathfrak{g}((d/dt)F_t\xi_0) + \circ(\|\xi_0\|)$, then $A_t F_t \xi_0 = (d/dt)F_t\xi_0$ for any $\xi_0 \in \mathbb{R}^{\dim \mathfrak{g}}$ and finally $A_t F_t = (d/dt)F_t$. ■

Appendix C
Proof of Theorem 4

A. Proof Rationale

We define the rest r_n, here for the left-invariant filter only, as follows: $\exp[(I - L_n H)\xi_n - r_n(\xi)] = \exp[\xi]\exp[-L_n H(\exp[\xi]b - b)]$. We then introduce the flow $\Psi_{t_0}^t$ of the linear part of the equations governing ξ_t (that is, $\Psi_{t_0}^{t_0} = Id$, $(d/dt)\Psi_{t_0}^t = A_t\Psi_{t_0}^t$, $\Psi_{t_0}^{t+} = (I - L_n H)\Psi_{t_0}^t$) and decompose the solution ξ_t as

$$\forall t \geq 0, \quad \xi_t = \Psi_{t_0}^t \xi_0 + \sum_{t_n < t} \Psi_{t_n}^t r_n(\xi_{t_n}). \tag{62}$$

All we have to verify is that the appearance of the second-order terms $r_n(\xi_{t_n})$ at each update is compensated by the exponential decay of $\Psi_{t_0}^t$ (Theorem 3).

B. Review of Existing Linear Results

Consider a linear time-varying Kalman filter and let $\Psi_{t_0}^t$ denote the flow of the error variable ξ_s. It is proved in [13] that if the parameters of the Riccati equation verify conditions (i)–(v) then there exist $\gamma_{\max} > 0$ and $\gamma_{\min} > 0$ such that $\gamma_{\max} I \succeq P_t \succeq \gamma_{\min} I$. This pivotal property allows proving the solution of the *linear* error equation $\Psi_s^t \xi_s$ verifies for $V(P, \xi) = \xi^T P^{-1} \xi$:

$$V\left(P_{t_{n+M}}^+, \Psi_{t_n}^{t_{n+M}}\left(\xi_{t_n}^+\right)\right) \leq V\left(P_{t_n}^+, \xi_{t_n}^+\right) - \beta_3 \left\| \Psi_{t_n}^{t_{n+M}}(\xi_{t_n}^+) \right\|^2 \tag{63}$$

where β_3 only depends on $\alpha_1, \alpha_2, \beta_1, \beta_2, \delta_1, \delta_2, \delta_3, M$. Of course, the proof given in [13] holds if the inequalities are only verified on an interval $[0, T]$. We will also use the classical result (also mentioned in [13] but easily derived)

$$V\left(P_{t_{n+1}}^+, \Psi_{t_n}^{t_{n+1}}\left(\xi_{t_n}^+\right)\right) \leq V\left(P_{t_n}^+, \xi_{t_n}^+\right). \tag{64}$$

C. Preliminary Lemmas

The proof of Theorem 4 is displayed in the next subsection. It relies on the final Lemma 7, which is proved step by step in this section through lemmas 4, 5, and 6. The time interval between two successive observations will be denoted Δt. $\tilde{P}_t$ will denote the Kalman covariance about the true state trajectory.

Lemma 4—[Modified Constants for Closeby Trajectories]: If the conditions (i) to (v) are satisfied about the *true* trajectory, then for any $k > 1$ there exists a radius ε such that the bound $\forall\, s \in [0, t], \|\xi_{t_0+s}\| < \varepsilon$ ensures the conditions (i) to (v) are also verified on $[t_0, t_0 + t]$ about the *estimated* trajectory, with the modified constants $\hat{\delta}_1 = \delta_1$, $\hat{\delta}_2 = (1/k^2)\delta_2$, $\hat{\delta}_3 = (1/k^2)\delta_3$, $\hat{\alpha}_1 = (1/k^2)\alpha_1$, $\hat{\alpha}_2 = k^2\alpha_2$, $\hat{\beta}_1 = (1/k^2)\beta_1$, $\hat{\beta}_2 = k^2\beta_2$. Moreover, if $(1/k)\tilde{P}_{t_0} \le P_{t_0} \le k\tilde{P}_{t_0}$ then $(1/k)\tilde{P}_{t_0+s} \le P_{t_0+s} \le k\tilde{P}_{t_0+s}$ holds on $[0, t]$.

Proof: We consider the LIEKF, the proof works the same way for the RIEKF. Matrices $\hat{Q}_t$ and $\hat{N}_n$ depend on the estimate $\hat{\chi}_t$, this is why this lemma is needed. So we replace them by their values: $\hat{Q}_t = \mathrm{Cov}(w_t)$ if the noise term has the form $\chi_t w_t$, $\hat{Q}_t = Ad_{\hat{\chi}_t^{-1}}\mathrm{Cov}(w_t)Ad_{\hat{\chi}_t^{-1}}^T$ if the noise term has the form $w_t\chi_t$, and $\hat{N}_n = \hat{\chi}_{t_n}^{-1}\mathrm{Cov}(V_n)\hat{\chi}_{t_n}^{-T} + \mathrm{Cov}(B_n)$. All these situations are covered if we assume there exist four (possibly time-dependent) matrices Q_1, Q_2, N_1 and N_2 such that $\hat{Q}_t = Q_1 + Ad_{\hat{\chi}_t^{-1}}Q_2Ad_{\hat{\chi}_t^{-1}}^T$ and $\hat{N}_n = N_1 + \hat{\chi}_{t_n}^{-1}N_2\hat{\chi}_{t_n}^{-T}$. These notations will be used in the sequel but they hold only for this proof: they are not related to matrices Q_1 and Q_2 defined in the simulations sections. The Riccati equation computed about the true trajectory reads

$$\frac{d}{dt}\tilde{P}_t = A_t\tilde{P}_t + \tilde{P}_tA_t^T + Q_1 + Ad_{\chi_t}Q_2Ad_{\chi_t}^T$$
$$\tilde{P}_{t_n}^+ = \tilde{P}_{t_n} - \tilde{P}_{t_n}H^T\left(H\tilde{P}_{t_n}H^T + N_1 + \chi_{t_n}^{-1}N_2\chi_{t_n}^{-T}\right)^{-1}H\tilde{P}_{t_n}.$$

The Riccati equation computed on the estimated trajectory is obtained replacing χ_t with $\hat{\chi}_t$. Recalling the error η_t and the properties of the Ad, the idea of the proof is simply to rewrite the Riccati equation computed about $\hat{\chi}_t$ as a perturbation of the Riccati equation computed about χ_t

$$\frac{d}{dt}P_t = A_tP_t + P_tA_t^T + Q_1 + Ad_{\eta_t^{-1}}\left[Ad_{\chi_t^{-1}}Q_2Ad_{\chi_t^{-1}}^T\right]Ad_{\eta_t^{-1}}^T$$
$$P_{t_n}^+ = P_{t_n} - P_{t_n}H^T\big(HP_{t_n}H^T + N_1 + \eta_{t_n}^{-1}\left[\chi_{t_n}^{-1}N_2\chi_{t_n}^{-T}\right]\eta_{t_n}^{-T}\big)^{-1}HP_{t_n}.$$

Controlling the perturbation is easy: matrix-valued functions $\xi \to e^{-\xi}$ and $\xi \to Ad_{e^{-\xi}}$ are continuous and equal to I_d for $\xi = 0$, thus there exists a real $\varepsilon > 0$ depending only on k such that $\|\xi_{t_0+s}\| \le \varepsilon$ ensures $(1/k)N_2 \preceq (e^{-\xi_{t_0+s}})N_2(e^{-\xi_{t_0+s}})^T \preceq kN_2$ and $(1/k)Q_2 \preceq Ad_{e^{-\xi_{t_0+s}}}Q_2Ad_{e^{-\xi_{t_0+s}}}^T \preceq kQ_2$. It ensures consequently $(1/k)(Q_1 + Ad_{\chi_{t_0+s}^{-1}}Q_2Ad_{\chi_{t_0+s}^{-1}}^T) \preceq Q_1 + Ad_{\hat{\chi}_{t_0+s}^{-1}}Q_2Ad_{\hat{\chi}_{t_0+s}^{-1}}^T \preceq k(Q_1 + Ad_{\chi_{t_0+s}^{-1}}Q_2Ad_{\chi_{t_0+s}^{-1}}^T)$ and $(1/k)(N_1 + \chi_{t_0+s}^{-1}N_2\chi_{t_0+s}^{-T}) \preceq N_1 + \hat{\chi}_{t_0+s}^{-1}N_2\hat{\chi}_{t_0+s}^{-T} \preceq k(N_1 + \chi_{t_0+s}^{-1}N_2\chi_{t_0+s}^{-T})$, and a mere look at the definitions of the constants of Theorem 3 yields the modified constants.

The inequality $(1/k)\tilde{P}_{t_0+s} \le P_{t_0+s} \le k\tilde{P}_{t_0+s}$ follows from the matrix inequalities above on the covariance matrices, by writing the Riccati equation verified by kP_t and $(1/k)P_t$ and using simple matrix inequalities. ■

Lemma 5—[First-Order Control of Growth]: Under the same conditions as in Lemma 4 (including $(1/k)\tilde{P}_{t_0} \le P_{t_0} \le k\tilde{P}_{t_0}$) and $\|\xi_{t_0+s}\|$ bounded by the same ε for $s \in [0, 2M\Delta T]$ (i.e. over $2M$ time steps, where M is defined as in Theorem 3), there exists a continuous function l_1 depending *only* on k ensuring $\|\xi_{t_0+s}\| \le l_1(\|\xi_{t_0}\|)$ for any $s \in [0, 2M\Delta T]$ and $l_1(x) = O(x)$.

Proof: Using Lemma 4 and then Theorem 3 we know there exist two constants $\gamma_{\min} > 0$ and $\gamma_{\max} > 0$ such that $\gamma_{\max}I \succeq P_t \succeq \gamma_{\min}I$. The nonlinear rest $r_{t_n}(\xi)$ introduced in (62) is defined by $\exp(\xi)\exp(L_n(e^{\xi}b - b - H\xi)) = \exp((I - L_nH)\xi + r_{t_n}(\xi))$. The Baker-Campbell-Hausdorff (BCH) formula gives $r_{t_n}(\xi) = O(\|\xi\|.\|L_n\xi\|)$ but L_n is uniformly bounded over time by $(\gamma_{\max}\|H\|)/\delta_3$ as an operator. Thus $\|r_n\|$ is *uniformly* dominated over time by a second order: there exists a continuous function $\tilde{l}^k$ (depending only on k and on the *true* trajectory) such that $\tilde{l}^k(x) = O(x^2)$ and $\|r_{t_n}(\xi)\| \le \tilde{l}^k(\|\xi\|)$ for any n such that $t_n \le 2M\Delta T$.

Now we can control the evolution of the error using $\tilde{l}^k$. The propagation step is linear, thus we have the classical result $(d/dt)V_t(\xi_t) < 0$. It ensures $\|\xi_{t_0+s}\| < \sqrt{\gamma_{\max}/\gamma_{\min}}\|\xi_{t_0}\|$ as long as there is no update on $[t_0, t_0 + s]$. At each update step we have $V_{t_n}^+(\xi_{t_n}^+)^{1/2} = V_{t_n}^+([I - L_nH]\xi_{t_n} + r_n(\xi_{t_n})^{1/2} \le V_{t_n}^+([I - L_nH]\xi_{t_n})^{1/2} + V_{t_n}^+(r_n(\xi_{t_n})^{1/2} \le V_{t_n}(\xi_{t_n})^{1/2} + V_{t_n}^+(r_n(\xi_{t_n}))^{1/2}$ using the triangular inequality. Thus: $\|\xi_{t_n}^+\| \le \sqrt{\gamma_{\max}/\gamma_{\min}}(\|\xi_{t_n}\| + \|r_n(\xi_{t_n})\|) \le \sqrt{\gamma_{\max}/\gamma_{\min}}(\|\xi_{t_n}\| + \tilde{l}^k(\|\xi_{t_n}\|))$. Reiterating over successive propagations and updates over $[0, 2M\Delta T]$, we see $\|\xi_{t_0+s}\|$ is uniformly bounded by a function $l_1(\|\xi_{t_0}\|)$ that is first order in $\|\xi_{t_0}\|$. ■

Lemma 6—[Second-Order Control of the Lyapunov Function]: Under the same conditions as in Lemma 4 (including $(1/k)\tilde{P}_{t_0} \le P_{t_0} \le k\tilde{P}_{t_0}$) and for $\|\xi_{t_0+s}\|$ bounded by the same ε for $s \in [0, 2M\Delta T]$ ($2M$ time steps, see Theorem 3 for the definition of M), there exists a continuous function l_2 depending *only* on k ensuring $V_{t_0+s}(\xi_{t_0+s}) \le V_{t_0+s}(\Psi_{t_0}^{t_0+s}\xi_{t_0}) + l_2(\|\xi_{t_0}\|) \le V_{t_0}(\xi_{t_0}) + l_2(\|\xi_{t_0}\|)$ for any $s \in [0, 2M\Delta t]$ with $l_2(x) = O(x^2)$. We also have $V_{t_n}(\xi_{t_n}^+) \le V_{t_n}^+((\Psi_{t_0}^{t_n})^+\xi_{t_0}) + l_2(\|\xi_{t_0}\|) \le V_{t_0}(\xi_{t_0}) + l_2(\|\xi_{t_0}\|)$ for $t_0 \le t_n \le t_0 + 2M\Delta T$.

Proof: The result stems from the decomposition (62) as

$$\begin{aligned}V_{t_0+s}(\xi_{t_0+s})^{\frac{1}{2}} &= V_{t_0+s}\left(\Psi_{t_0}^{t_0+s}\xi_{t_0} + \sum_{t_0<t_n<t_0+s}\Psi_{t_n}^{t_0+s}r_n(\xi_{t_n})\right)^{\frac{1}{2}}\\ &\le V_{t_0+s}\big(\Psi_{t_0}^{t_0+s}\xi_{t_0+s}\big)^{\frac{1}{2}} + \sum_{t_0<t_n<t_0+s}V_{t_0+s}\big(\Psi_{t_n}^{t_0+s}r_n(\xi_{t_n})\big)^{\frac{1}{2}}.\end{aligned}$$

Using (64) and the definition of V this is less than $V_{t_0+s}(\Psi_{t_0}^{t_0+s}\xi_{t_0+s})^{1/2} + \sum_{t_0<t_n<t_0+s}\sqrt{\gamma_{\max}/\gamma_{\min}}\|r_n(\xi_{t_n})\|^{1/2}$. But from Lemma 5 $\|r_n(\xi_{t_n})\| \le (\tilde{l}^k \circ l_1^k)(\|\xi_{t_0}\|)$. As we have $(\tilde{l}^k \circ l_1^k)(x) = O(x^2)$, we obtain the result squaring the inequality and using $V_{t_0+s}(\Psi_{t_0}^{t_0+s}\xi_{t_0+s}) \le V_{t_0}(\xi_{t_0}) \le (\gamma_{\max}/\gamma_{\min})\|\xi_{t_0}\|$ to control the crossed terms. ■

Lemma 7—[Final Second Order Growth Control]: Under the same conditions as in Lemma 4 (including $(1/k)\tilde{P}_{t_0} \le P_{t_0} \le k\tilde{P}_{t_0}$) and for $\|\xi_{t_0+s}\|$ bounded by the same ε for $s \in [0, t]$, there exist two functions $l_1^k(\xi) = O(\|\xi\|^2)$ and $l_2^k = o(\|\xi\|^2)$ and a constant β^k ensuring the relation

$$V_{t_0+s}(\xi_{t_0+s}) \le V_{t_0}(\xi_{t_0}) + l_1^k(\xi_{t_0}) - \sum_{i=0}^{J-1}\left[\beta^k\|\xi_{t_{n_0}+iM\Delta t}\|^2 - l_2^k(\xi_{t_{n_0}+iM\Delta t})\right] + l_1^k\left(\left\|\xi_{t_{n_{\max}}}^+\right\|\right) \tag{65}$$

where $n_{\max}$ is the last update before t_0+s (i.e., $n_{\max}=\max\{n, t_n \leq t_0+s\}$), J is the number of successive sequences of M updates in $[t_0+M\Delta t, t_{n_{\max}}]$ (i.e., $J=\max\{j, t_{n_{\max}-jM} \geq t_0\}-1$) and $n_0=n_{\max}-JM$. If $t_0+s=t_{n_{\max}}$ the last term can be removed.

Proof: For l_1^k we choose the same function as in Lemma 6. There is nothing more to prove for $s<2M\Delta t$. Let $s \geq 2M\Delta t$. We have $V_{t_0+s}(\xi_{t_0+s})-V_{t_0}^+(\xi_{t_0}^+)=(V_{t_0+s}(\xi_{t_0+s})-V_{t_{n_{\max}}}^+(\xi_{t_{n_{\max}}}^+))+(V_{t_{n_{\max}}}^+(\xi_{t_{n_{\max}}}^+)-V_{t_{n_0}}^+(\xi_{t_{n_0}}^+))+(V_{t_{n_0}}^+(\xi_{t_{n_0}}^+)-V_{t_0}(\xi_{t_0}))$. The first and third terms are upper bounded using Lemma 6. The second term $V_{t_{n_{\max}}}^+(\xi_{t_{n_{\max}}}^+)-V_{t_{n_0}}^+(\xi_{t_{n_0}}^+)$ is equal to

$$\sum_{i=0}^{J-1}\left[V_{t_{n_0+(i+1)M}}^+\left(\xi_{t_{n_0+(i+1)M}}^+\right)-V_{t_{n_0+iM}}^+\left(\xi_{t_{n_0+iM}}^+\right)\right]$$
$$\leq \sum_{i=0}^{J-1}\left[V_{t_{n_0+(i+1)M}}^+\left(\Psi_{t_{n_0+iM}}^{t_{n_0+(i+1)M}+}\xi_{t_{n_0+iM}}^+\right)\right.$$
$$\left.-V_{t_{n_0+iM}}^+\left(\xi_{t_{n_0+iM}}^+\right)+l_2\left(\xi_{t_{n_0+iM}}^+\right)\right].$$

And we conclude using [see [13]]

$$\begin{aligned}V_{t_{n_0+(i+1)M}}^+&\left(\Psi_{t_{n_0+iM}}^{t_{n_0+(i+1)M}+}\xi_{t_{n_0+iM}}^+\right)-V_{t_{n_0+iM}}^+\left(\xi_{t_{n_0+iM}}^+\right)\\ &\leq -\tilde{\beta}^k\left\|\Psi_{t_{n_0+iM}}^{t_{n_0+(i+1)M}+}\xi_{t_{n_0+iM}}^+\right\|^2\\ &\leq -\tilde{\beta}^k\left(\frac{\gamma_{\min}}{\gamma_{\max}}\delta_1\right)^M\left\|\xi_{t_{n_0+iM}}^+\right\|^2\end{aligned}$$

for a $\tilde{\beta}^k$ depending only on the modified constants of Lemma 4. The last inequality is obtained using $\Psi_{t_0}^{t_n+}=(P_n^+P_n^{-1})\Psi_{t_0}^{t_n}$ and an obvious recursion over M time steps. We finally set $\beta=\tilde{\beta}^k((\gamma_{\min}/\gamma_{\max})\delta_1)^M$. ∎

The control we have obtained on ξ_{t_0+s} is verified if $\|\xi_{t_0+s}\|$ is *already* less than ε over the interval $[t_0, t_0+t]$. We now prove the result holds assuming *only* that ξ_{t_0} is sufficiently small.

D. Proof of Theorem 4

Applying Lemma 7 with $t_0+s=t_{n_{\max}}$ yields for $(1/k)\tilde{P}_{t_0} \leq P_{t_0} \leq k\tilde{P}_{t_0}$ and $\|\xi_{t_0+s}\|<\varepsilon$ on $[0,t]$ $\|\xi_{t_{n_{\max}}}^+\|^2 \leq (\gamma_{\max}/\gamma_{\min})\|\xi_{t_0}\|^2+\gamma_{\max}l_1^k(\|\xi_{t_0}\|)-(\gamma_{\max}/\gamma_{\min})\sum_{i=0}^{J-1}[\beta^k\|\xi_{t_{n_0+iM}}^+\|^2-l_2(\xi_{t_{n_0+iM}}^+)]$. There exist $K>0$ and $\varepsilon'>0$ such that for $x<\varepsilon'$, we have $l_2(x)<(\beta^k/2)x$ and $\gamma_{\max}l_1^k(x)<Kx$ (as $l_2(x)$ and $l_1^k(x)$ are $O(x^2)$) and thus $\|\xi_{t_{n_{\max}}}^+\|^2 \leq ((\gamma_{\max}/\gamma_{\min})+K)\|\xi_{t_0}\|^2$. Thus, for $\|\xi_{t_0}\|<(\varepsilon'/\sqrt{(\gamma_{\max}/\gamma_{\min})+K})$

$$\|\xi_{t_0+s}\|^2 \leq \left(\frac{\gamma_{\max}}{\gamma_{\min}}+K+K\left(\frac{\gamma_{\max}}{\gamma_{\min}}+K\right)\right)\|\xi_{t_0}\|^2 - \frac{\gamma_{\max}}{\gamma_{\min}}\sum_{i=0}^{J-1}\frac{\beta^k}{2}\left\|\xi_{t_{n_0+iM}}^+\right\|^2 \quad (66)$$

which finally ensures

$$\|\xi_{t_0}\|<\frac{1}{2}\varepsilon'\Bigg/\left(\frac{\gamma_{\max}}{\gamma_{\min}}+K+K\left(\frac{\gamma_{\max}}{\gamma_{\min}}+K\right)\right) \Rightarrow \|\xi_{t_0+s}\| \leq \varepsilon'/2.$$

Reducing ε' if necessary to have $\varepsilon' \leq \varepsilon$, we have obtained $\|\xi_{t_0+s}\|<\varepsilon'$ for $s\in[0,t] \Rightarrow \|\xi_{t_0+s}\| \leq \varepsilon'/2$ for sufficiently small $\|\xi_{t_0}\|$ (as Lemma 7 applies). Letting $t=\inf\{s, \|\xi_{t_0+s}\| \geq (3/4)\varepsilon'\}$ for sufficiently small $\|\xi_{t_0}\|$ we end up with a contradiction if we suppose $t<+\infty$, which proves $t=+\infty$ (the results holding *only* for sufficiently small $\|\xi_{t_0}\|$).

Moreover, (66) shows that $\sum_{i=0}^{J-1}(\beta^k/2)\|\xi_{t_{n_0+iM}}^+\|^2$ is bounded and has positive terms thus $\|\xi_{t_{n_0+iM}}^+\|^2$ goes to zero. Note also that $\|P_t-\tilde{P}_t\| \underset{t\to+\infty}{\longrightarrow} 0$ as a byproduct.

Appendix D
Proof of Proposition 3

Only conditions (i) and (v) are nontrivial. Let Φ denote the flow of the dynamics. We have

$$\begin{aligned}\frac{d}{dt}\left[\left(\Phi_{t_n}^t\right)^T\Phi_{t_n}^t\right]&=\left(\Phi_{t_n}^t\right)^T\begin{pmatrix}0&0&v_t\\0&0&0\\-v_t&0&0\end{pmatrix}\Phi_{t_n}^t\\&\succeq -v_{\max}\left(\Phi_{t_n}^t\right)^T\Phi_{t_n}^t\end{aligned}$$

as the eigenvalues of the matrix in the center are $-v_t(1,-1,0)$. Thus, $\forall z\in\mathbb{R}^3$, $(d/dt)\log(z^T(\Phi_{t_n}^t)^T\Phi_{t_n}^t z) \geqslant -v_{\max}$ and finally $z^T(\Phi_{t_n}^{t_{n+1}})^T\Phi_{t_n}^{t_{n+1}}z \geqslant \exp(-v_{\max}(t_{n+1}-t_n))\|z\|^2$ as $(\Phi_{t_n}^{t_n})^T\Phi_{t_n}^{t_n}=I_3$. Thus $(\Phi_{t_n}^{t_{n+1}})^T\Phi_{t_n}^{t_{n+1}} \succeq \exp(-v_{\max}(t_{n+1}-t_n))I_3$ and (i) is verified. The difficult part of (v) is the lower bound. Denoting $\mathrm{Cov}(V_n)$ by N we will show

$$\begin{aligned}&\exists\beta_1 \quad \forall n\in\mathbb{N}, \quad \beta_1 I_3 \leq \hat{R}_{t_n}^T N^{-1}\hat{R}_{t_n}\\&+\left(\Phi_{t_n}^{t_{n+1}}\right)^T\begin{pmatrix}0_{1,2} & I_2\end{pmatrix}^T\hat{R}_{t_{n-1}}^T N^{-1}\hat{R}_{t_{n-1}}\begin{pmatrix}0_{1,2} & I_2\end{pmatrix}\Phi_{t_n}^{t_{n+1}}.\end{aligned}$$

That is to say that we want a lower bound on the quadratic form

$$\begin{aligned}M(\theta,u)&=(\theta,u)\hat{R}_{t_n}^T N^{-1}\hat{R}_{t_n}(\theta,u)^T\\&\quad+\begin{pmatrix}\theta\\u\end{pmatrix}^T\left(\Phi_{t_n}^{t_{n+1}}\right)^T\begin{pmatrix}0_{1,2} & I_2\end{pmatrix}^T\hat{R}_{t_{n-1}}^T N^{-1}\hat{R}_{t_{n-1}}\\&\quad\times\begin{pmatrix}0_{1,2} & I_2\end{pmatrix}\Phi_{t_n}^{t_{n+1}}\begin{pmatrix}\theta\\u\end{pmatrix}.\end{aligned}$$

We decompose $\Phi_{t_n}^{t_{n+1}}$ as $\Phi_{t_n}^{t_{n+1}}=\begin{pmatrix}1 & 0 \quad 0\\ \delta V_n & T_n\end{pmatrix}$. To simplify the writing we introduce the norms $\|x\|_N^2=x^TN^{-1}x$ and the associated scalar product $\langle .,.\rangle_N$. There exists $\alpha>0$ such that $\forall x\in\mathbb{R}^2, \|x\|_N \geq \alpha\|x\|$. For any $\begin{pmatrix}\theta\\u\end{pmatrix}\in\mathbb{R}^3$, we have $M(\theta,u)=\|\hat{R}_{t_n}u\|_N^2+\|\theta\hat{R}_{t_{n-1}}\delta V_n+\hat{R}_{t_{n-1}}T_nu\|_N^2=\|\hat{R}_{t_n}u\|_N^2+\theta^2\|\hat{R}_{t_{n-1}}\delta V_n\|_N^2+2\theta\langle\hat{R}_{t_{n-1}}\delta V_n,\hat{R}_{t_{n-1}}T_nu\rangle_N+\|\hat{R}_{t_{n-1}}T_nu\|_N^2$ and for $\lambda\in]0,1]$ we have that $M(\theta,u)$ equals

$$\begin{aligned}&\left\|\hat{R}_{t_n}u\right\|_N^2+(1-\lambda^2)\theta^2\left\|\hat{R}_{t_{n-1}}\delta V_n\right\|_N^2+\lambda^2\theta^2\left\|\hat{R}_{t_{n-1}}\delta V_n\right\|_N^2\\&\quad+2\theta\left\langle\hat{R}_{t_{n-1}}\delta V_n,\hat{R}_{t_{n-1}}T_nu\right\rangle_N+\left\|\hat{R}_{t_{n-1}}T_nu\right\|_N^2\\&=\left\|\hat{R}_{t_n}u\right\|_N^2+(1-\lambda^2)\theta^2\left\|\hat{R}_{t_{n-1}}\delta V_n\right\|_N^2\\&\quad+\left\|\lambda\theta\hat{R}_{t_{n-1}}\delta V_n+\frac{1}{\lambda}\hat{R}_{t_{n-1}}T_nu\right\|_N^2+\left(1-\frac{1}{\lambda^2}\right)\left\|\hat{R}_{t_{n-1}}T_nu\right\|_N^2.\end{aligned}$$

Regrouping terms and using the fact rotations are isometries, yields as a lower bound $\alpha(2-(1/\lambda^2))(\|u\|^2+\theta^2+[((1-\lambda^2)/(2-(1/\lambda^2)))v_{\min}^2-1]\theta^2)$. As $((1-\lambda^2)/(2-(1/\lambda^2)))\longrightarrow +\infty$ when $\lambda \to (1/\sqrt{2})^-$ there exists λ_0 such that: $M(\theta,u)^T \geq \alpha(2-(1/\lambda_0^2))\|(\theta,u)^T\|^2$ and the result is true for $\beta_1=\alpha(2-(1/\lambda_0^2))$.

References

[1] N. Aghannan and P. Rouchon, "An intrinsic observer for a class of Lagrangian systems," *IEEE Trans. Autom. Control*, vol. 48, no. 6, pp. 936–945, 2003.

[2] M. Barczyk, S. Bonnabel, J. E. Deschaud, and F. Goulette, "Invariant EKF design for scan matching-aided localization," *IEEE Trans. Syst. Control Technol.*, vol. 23, no. 6, pp. 2440–2448, 2015.

[3] M. Barczyk and A. F. Lynch, "Invariant observer design for a helicopter UAV aided inertial navigation system," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 3, pp. 791–806, 2013.

[4] A. Barrau and S. Bonnabel, "Invariant particle filtering with application to localization," in *Proc. IEEE 53rd Annu. CDC*, pp. 5599–5605. *IEEE*, 2014.

[5] A. Barrau and S. Bonnabel, "Intrinsic filtering on Lie groups with applications to attitude estimation," *IEEE Trans. Autom. Control*, vol. 60, no. 2, pp. 436–449, 2015.

[6] S. Bonnabel, "Left-invariant extended Kalman filter and attitude estimation," in *Proc. 46th IEEE Conf. Decision Control*, 2007, pp. 1027–1032.

[7] S. Bonnabel, P. Martin, and P. Rouchon, "Non-linear symmetry-preserving observers on Lie groups," *IEEE Trans. Autom. Control*, vol. 54, no. 7, pp. 1709–1713, 2009.

[8] S. Bonnabel, P. Martin, and P. Rouchon, "Symmetry-preserving observers," *IEEE Trans. Autom. Control*, vol. 53, no. 11, pp. 2514–2526, 2008.

[9] S. Bonnabel, P. Martin, and E. Salaün, "Invariant extended Kalman filter: theory and application to a velocity-aided attitude estimation problem," in *Proc. 48th IEEE Conf. Decision Control*, 2009, pp. 1297–1304.

[10] S. Bonnabel and J. J. Slotine, "A contraction theory-based analysis of the stability of the deterministic extended Kalman filter," *IEEE Trans. Autom. Control*, vol. 60, no. 2, pp. 565–569, 2015.

[11] M. Boutayeb, H. Rafaralahy, and M. Darouach, "Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems," *IEEE Trans. Autom. Control*, vol. 42, 1997.

[12] A. De Luca, G. Oriolo, and C. Samson, "Feedback control of a nonholonomic car-like robot," in *Robot Motion Planning and Control*, pp. 171–253. Springer, 1998.

[13] J. Deyst and C. Price, "Conditions for asymptotic stability of the discrete minimum-variance linear estimator," *IEEE Trans. Autom. Control*, vol. 13, pp. 702–705, 1968.

[14] J. Farrell, *Aided Navigation: GPS With High Rate Sensors*. New York, NY, USA: McGraw-Hill, 2008.

[15] J. P. Gauthier and I. Kupka, *Deterministic Observation Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2001.

[16] H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi, "Globally exponentially stable attitude and gyro bias estimation with application to GNSS/INS integration," *Automatica*, vol. 51, pp. 158–166, 2015.

[17] M. D. Hua, "Attitude estimation for accelerated vehicles using GPS/INS measurements," *Control Eng. Pract.*, vol. 18, no. 7, pp. 723–732, 2010.

[18] M. Izadi and A. K. Sanyal, "Rigid body attitude estimation based on the Lagrange-d—Alembert principle," *Automatica*, vol. 50, no. 10, pp. 2570–2577, 2014.

[19] A. J. Krener, "The convergence of the extended Kalman filter," in *Directions in Mathematical Systems Theory and Optimization*. New York, NY, USA: Springer, 2003, pp. 173–182.

[20] C. Lageman, J. Trumpf, and R. Mahony, "Gradient-like observers for invariant dynamics on a Lie group," *IEEE Trans. Autom. Control*, vol. 55, no. 2, pp. 367–377, 2010.

[21] E. J. Lefferts, F. Landis Markley, and M. D. Shuster, "Kalman filtering for spacecraft attitude estimation," *J. Guid. Control, Dynam.*, vol. 5, no. 5, pp. 417–429, 1982.

[22] R. Mahony, T. Hamel, and J-M. Pflimlin, "Complementary filter design on the special orthogonal group SO(3)," in *Proc. 44th IEEE CDC-ECC'05*, 2005, pp. 1477–1484.

[23] P. Martin and E. Salaün, "Generalized multiplicative extended Kalman filter for aided attitude and heading reference system," in *Proc. AIAA Guid., Navigat., Control Conf.*, 2010, p. 8300.

[24] F. Sonnemann Reif, K. and R. Unbehauen, "An EKF-based nonlinear observer with a prescribed degree of stability," *Automatica*, vol. 34, pp. 1119–1123, 1998.

[25] Y. K. Song and J. W. Grizzle, "The extended Kalman filter as a local asymptotic observer," *Estimat. Control*, vol. 5, pp. 59–78, 1995.

[26] J. F. Vasconcelos, R. Cunha, C. Silvestre, and P. Oliveira, "A nonlinear position and attitude observer on SE(3) using landmark measurements," *Syst. Control Lett.*, vol. 59, pp. 155–166, 2010.

Axel Barrau received the Eng. degree in applied mathematics from Ecole Polytechnique, Palaiseau, France, in 2013, and the Ph.D. degree in computer science from Ecole des Mines ParisTech, Paris, France, in 2015.

He is currently a Research Engineer with Safran, Magny Les Hameaux Cedex, France. His fields of interest are nonlinear filtering, inertial navigation, and computer vision.

Silvère Bonnabel received the Eng. and Ph.D. degrees in mathematics and control from Mines ParisTech, Paris, France, in 2004 and 2007, respectively.

In 2008, he held a postdoctoral position at the University of Liège. He is currently a professor at Mines Paristech. His research includes nonlinear state estimation, optimization on manifolds, and industrial applications.

Dr. Bonnabel was awarded the IEEE/SEE Glavieux Prize in 2015. He serves an Associate Editor for *Systems and Control Letters*.