HAL
open science

The Invariant Extended Kalman Filter as a Stable Observer

Silvere Bonnabel, Axel Barrau

▶ **To cite this version:**

Silvere Bonnabel, Axel Barrau. The Invariant Extended Kalman Filter as a Stable Observer. IEEE Transactions on Automatic Control, 2017, 62 (4), pp.1797-1812. 10.1109/TAC.2016.2594085 . hal-01692380

HAL Id: hal-01692380

https://minesparis-psl.hal.science/hal-01692380v1

Submitted on 25 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:1410.1465v4 [cs.SY] 19 Oct 2015

The Invariant Extended Kalman filter as a stable observer

Axel Barrau, Silvère Bonnabel *

Abstract

We analyze the convergence aspects of the invariant extended Kalman filter (IEKF), when the latter is used as a deterministic non-linear 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. Then, we leverage the result 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.

1 Introduction

The design of non-linear 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 non-linear 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 non-linear 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

*A. Barrau and S. Bonnabel are with MINES ParisTech, PSL Research University, Centre for robotics, 60 Bd St Michel 75006 Paris, France. [axel.barrau,silvere.bonnabel]@mines-paristech.fr

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 [11, 25, 24, 10]) rely on the highly non-trivial 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. 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 [8, 7] and notably the theory of invariant Kalman filtering [6, 9, 23, 5] in a purely *deterministic* context. As such, it is a contribution to the theory of non-linear observer design on Lie groups that has lately attracted considerable interest, notably for attitude estimation, see, e.g., [22, 26, 3, 17, 20, 16, 18]. The detailed contributions and organization of the paper are as follows.

In Section 2, we recall the main contribution of [8, 7] is to evidence the fact that for left-invariant systems on Lie groups, non-linear 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 non-linear change of variables, the error evolution (in the absence of measurements) obeys a *linear* differential equation.

In Section 3, we focus on the invariant extended Kalman filter (IEKF) [6] when applied to the broad class of systems of Section 2. 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 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., [3, 5, 2].

In Section 4 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 5 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 [8, 7], 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 non-linear function of the non-linear 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 applicability of 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 even capable to diverge.

2 A special class of multiplicative systems

2.1 An introductory example

Consider a linear (deterministic) system $\frac{d}{dt}x_t = A_t x_t$. Consider two trajectories of this system, say, a reference trajectory $(\bar{x}_t)_{t\geq 0}$ and another one $(x_t)_{t\geq 0}$. The discrepancy between both trajectories $\Delta x_t := x_t - \bar{x}_t$ satisfies the linear equation $\frac{d}{dt}\Delta x_t = A_t \Delta x_t$. 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 non-linear standard model of the two-dimensional non-holonomic car. Its state is defined by three parameters : heading θ_t and position $X_t = (x_t^1, x_t^2)$. The velocity $u_t \in \mathbb{R}$ is given by an odometer, the angular velocity $\omega_t \in \mathbb{R}$ is measured by a differential odometer or a gyrometer. The equations read (see, e.g., [12]):

$$\frac{d}{dt}\theta_t = \omega_t, \qquad \frac{d}{dt}x_t^1 = \cos(\theta_t)u_t, \qquad \frac{d}{dt}x_t^2 = \sin(\theta_t)u_t.$$

Now consider a reference trajectory $(\bar{\theta}_t, \bar{X}_t)$ and a second 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:

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

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 non-linear. These features, characteristic of non-linear systems, make the design of observers way more complicated in the non-linear case. Now, let us introduce the following non-linear 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):

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

which is equal to 0 indeed if and only if both trajectories coincide. 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 non-linear error* (2) *obeys the following* linear *and autonomous equation although the system, and the error, are totally non-linear:*

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

Proof. We will use the notations $\delta x_t = \frac{1}{2}\left(\theta_t - \bar{\theta}_t\right)\left[a(\theta_t - \bar{\theta}_t)I_2 - J_2\right] R(-\bar{\theta}_t)(X_t - \bar{X}_t)$. Since we have $\frac{d}{dt}\delta\theta_t = 0$ as in the linear case above, only the two last terms of δx_t change over time. Moreover, J_2 commutes with I_2, J_2 and $R(-\bar{\theta}_t)$ and $\frac{d}{dt}R(-\bar{\theta}_t) = -J_2\omega_t R(-\bar{\theta}_t)$. Thus we have:

$$\begin{aligned}
\frac{d}{dt}\delta x_t &= -J_2\omega_t\delta x_t\\
&\quad + \frac{1}{2}\delta\theta_t\left[a(\delta\theta_t)I_2 - J_2\right] R(-\bar{\theta}_t)\left[R(\theta_t) - R(\bar{\theta}_t)\right](u_t, 0)^T\\
&= -J_2\omega_t\delta x_t\\
&\quad + \frac{1}{2}\delta\theta_t\left[a(\delta\theta_t)I_2 - J_2\right]\left[(\cos(\delta\theta_t) - 1)I_2 + \sin(\delta\theta_t)J_2\right](u_t, 0)^T\\
&= -J_2\omega_t\delta x_t + \delta\theta_t J_2(u_t, 0)^T\\
&= -(\omega_t J_2)\delta x_t + \delta\theta_t(0, -u_t)^T.
\end{aligned}$$

In the second to last equality we used the relation $\frac{\sin(s)^2}{1-\cos(s)} = 1 + \cos(s)$. Equation (3) is proved. □

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 non-linear observers, namely the IEKF, due to the implied similarities with the linear case.

2.2 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* $\frac{d}{dt}\eta_t = g_{u_t}(\eta_t)$.

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*

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

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

Proof. Assume we have $\frac{d}{dt}\eta_t^L = g_{u_t}(\eta_t^L)$ 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:

$$\begin{aligned} g_{u_t}(\chi_t^{-1}\bar{\chi}_t) &= \frac{d}{dt}(\chi_t^{-1}\bar{\chi}_t) = -\chi_t^{-1}[\frac{d}{dt}\chi_t]\chi_t^{-1}\bar{\chi}_t + \chi_t^{-1}\frac{d}{dt}\bar{\chi}_t \\ &= -\chi_t^{-1} f_{u_t}(\chi_t)\eta_t^L + \chi_t^{-1} f_{u_t}(\bar{\chi}_t), \\ \text{i.e. } g_{u_t}(\eta_t^L) &= -\chi_t^{-1} f_{u_t}(\chi_t)\eta_t^L + \chi_t^{-1} f_{u_t}(\chi_t\eta_t^L). \end{aligned} \tag{10}$$

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

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

Reinjecting (11) in (10) we obtain:

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

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:*

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

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. We thus see the class of system introduced here appears as a generalization of the linear case.*

In the next section we show that the dynamics of the form (4) with additional property (7) have striking properties generalizing those of linear systems.

2.3 Log-linear property of the error propagation

In the sequel, we will systematically consider systems of the form (4) with the additional property (7), i.e. systems on Lie groups defined by

$$\begin{aligned} \frac{d}{dt}\chi_t &= f_{u_t}(\chi_t), \\ \text{where } \forall(u,a,b) \quad f_u(ab) &= af_u(b) + f_u(a)b - af_u(Id)b. \end{aligned} \tag{12}$$

For such systems, Theorem 1 proves that the left (resp. right) invariant error is a solution to the equation $\frac{d}{dt}\eta_t = g_{u_t}(\eta_t)$ where g_{u_t} is given by (8) (resp. (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 $A_{u_t}^i$ be defined by $g_{u_t}(\exp(\xi)) = \mathscr{L}_{\mathfrak{g}}(A_{u_t}^i \xi) + O(\|\xi\|^2)$. 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 non-linear error η_t^i, the correspondence* at all times *and for arbitrarily large errors*

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

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 2.1 are mere applications of the latter theorem, as the non-holonomic car example turns out to perfectly fit into our framework (see Section 4) and ξ_t in eq (2) actually merely is the Lie logarithm of the left-invariant error. This will be extensively used in Section 3, and in the examples to prove stability properties of IEKFs.

3 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, 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 2, under observability conditions, and painless conditions on the covariance matrices considered here as design parameters, the IEKF is a (deterministic) non-linear observer with local convergence properties around *any* trajectory, a feature extremely rare to obtain in the field of non-linear observers, due to the dependency of the estimation error to the true unknown trajectory. The notions necessary to follow Section 3 are given in Appendix A.

3.1 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.

3.1.1 Left-invariant observations

The first family of outputs we are interested in write:

$$Y^1_{t_n} = \chi_{t_n} d^1 \quad , \quad \dots \quad , \quad Y^k_{t_n} = \chi_{t_n} d^k, \tag{15}$$

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:

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

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

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:

$$\eta^L_t = \chi_t^{-1} \hat{\chi}_t. \tag{18}$$

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 eq. (8) ! We have thus

$$\frac{d}{dt}\eta^L_t = g^L_{u_t}(\eta^L_t), \quad t_{n-1} \leq t < t_n. \tag{19}$$

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

$$\frac{d}{dt}\xi_t = A_{u_t}\xi_t, \tag{20}$$

where A_{u_t} is defined by $g^L_{u_t}(\exp(\xi)) = \mathscr{L}_{\mathfrak{g}}(A_{u_t}\xi) + O(\|\xi_t\|^2)$. 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:

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

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

3.1.2 Right-invariant observations

The second family of observations we are interested in have the form:

$$Y^1_{t_n} = \chi_{t_n}^{-1} d^1 \quad , \quad \dots \quad , \quad Y^k_{t_n} = \chi_{t_n}^{-1} d^k, \tag{22}$$

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

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

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

A right-invariant error can be associated to this filter:

$$\eta^R_t = \hat{\chi}_t \chi_t^{-1}. \tag{25}$$

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:

$$(\eta^R_{t_n})^+ = \hat{\chi}^+_t \chi_t^{-1} = \exp\left[L_n \begin{pmatrix} \eta^R_{t_n} d^1 - d^1 \\ \dots \\ \eta^R_{t_n} d^k - d^k \end{pmatrix} \right] \eta^R_{t_n},$$

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

3.2 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 non-linear 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 non-linear observer, that is here the IEKF. This provides him with (at least) a first sensible tuning of the parameter matrices, which is consistent with the trust he has in each sensor. Moreover, in the same spirit, the IEKF viewed as a non-linear observer remedies a common weakness shared by numerous non-linear 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.

3.2.1 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:

$$Y^1_{t_n} = \chi_{t_n}\left(d^1 + B^1_n\right) + V^1_n \ , \ \dots \ , \ Y^k_{t_n} = \chi_{t_n}\left(d^k + B^k_n\right) + V^k_n, \tag{27}$$

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

$$Y^1_{t_n} = \chi^{-1}_{t_n}\left(d^1 + V^1_n\right) + B^1_n \ , \ \dots \ , \ Y^k_{t_n} = \chi^{-1}_{t_n}\left(d^k + V^k_n\right) + B^k_n. \tag{28}$$

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

$$\begin{aligned} \frac{d}{dt}\eta^L_t &= g^L_{u_t}\left(\eta^L_t\right) - w_t\eta^L_t, \\ \frac{d}{dt}\eta^R_t &= g^R_{u_t}\left(\eta^R_t\right) - \left(\hat{\chi}_t w_t \hat{\chi}_t^{-1}\right)\eta^R_t. \end{aligned} \tag{30}$$

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 $\mathbb{R}^{\dim \mathfrak{g}}$:

$$\frac{d}{dt}\xi_t = A^i_{u_t}\xi_t + \hat{w}_t, \tag{31}$$

where $A^i_{u_t}$ is defined by $g^i_{u_t}(\exp(\xi)) = \mathscr{L}_{\mathfrak{g}}(A^i_{u_t}\xi) + O(\|\xi_t\|^2)$ 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 $\hat{w}_t$ can be considered as white noise, are approximations that would require a proper justification in a stochastic setting. However, they are part of the standard EKF methodology, see e.g., [21] and justifying them is not the object of the present paper. 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. $\exp(\xi_{t_n}) = \eta^L_{t_n}$. Moreover, for $1 \le i \le k$ let $\hat{V}^i_n$ denote $\hat{\chi}^{-1}_{t_n}V^i_n$. The error update (21), when the LIEKF update (17) is fed with the "noisy" measurements (27) becomes

$$\begin{aligned}\left(\eta^L_{t_n}\right)^+ &= \chi^{-1}_{t_n}\hat{\chi}^+_{t_n} \\ &= \eta^L_{t_n}\exp\left[L_n\begin{pmatrix}\left(\eta^L_{t_n}\right)^{-1}d^1 - d^1 + \hat{V}^1_n + \left(\eta^L_{t_n}\right)^{-1}B^1_n \\ \dots \\ \left(\eta^L_{t_n}\right)^{-1}d^k - d^k + \hat{V}^k_n + \left(\eta^L_{t_n}\right)^{-1}B^k_n\end{pmatrix}\right].\end{aligned} \tag{32}$$

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

$$\begin{aligned}&(\eta_{t_n})^{-1}d^i - d^i + \hat{V}^i_n + (\eta^L_{t_n})^{-1}B^i_n \\ &= \exp_m(\mathscr{L}_{\mathfrak{g}}(\xi_{t_n}))^{-1}\left(d^i + B^i_n\right) - d^i + \hat{V}^i_n \\ &= \left(Id - \mathscr{L}_{\mathfrak{g}}(\xi)_{t_n}\right)\left(d^i + B^i_n\right) - d^i + \hat{V}^i_n + O\|\xi_{t_n}\|^2 \\ &= -\mathscr{L}_{\mathfrak{g}}(\xi)_{t_n}d^i + \hat{V}^i_n + B^i_n + O\left(\|\xi_{t_n}\|^2\right) + O\left(\|\xi_{t_n}\|\|B^i_n\|\right),\end{aligned}$$

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

$$Id + \mathscr{L}_{\mathfrak{g}}(\xi^+_{t_n}) = Id + \mathscr{L}_{\mathfrak{g}}\left(L_n\begin{pmatrix}-\mathscr{L}_{\mathfrak{g}}(\xi_{t_n})d^1 + \hat{V}^1_n + B^1_n \\ \dots \\ -\mathscr{L}_{\mathfrak{g}}(\xi_{t_n})d^k + \hat{V}^k_n + B^k_n\end{pmatrix}\right) + T. \tag{33}$$

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

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

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

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

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: $\frac{d}{dt}x_t = A_{u_t}x_t + \hat{w}_t$, $y_n = Hx_{t_n} + \hat{V}_n + B_n$. The standard Kalman theory thus suggests to compute L_n through the Riccati equation

$$\begin{aligned} \frac{d}{dt}P_t = A_{u_t}P_t + P_tA_{u_t}^T + \hat{Q}_t, \qquad P_{t_n}^+ = (I - L_nH)P_{t_n}, \\ \text{with } S_n = HP_{t_n}H^T + \hat{N}_n,\ L_n = P_{t_n}H^TS^{-1}. \end{aligned} \tag{35}$$

3.3 Summary of IEFK equations

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

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

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

$$\begin{aligned} \frac{d}{dt}P_t = A_{u_t}P_t + P_tA_{u_t}^T + \hat{Q}_t, \\ S_n = HP_{t_n}H^T + \hat{N}_n, \quad L_n = P_{t_n}H^TS^{-1}, \quad P_{t_n}^+ = (I - L_nH)P_{t_n}. \end{aligned} \tag{37}$$

As concerns the LIEKF, A_{u_t} is defined by $g_{u_t}^L(\exp(\xi)) = \mathscr{L}_{\mathfrak{g}}(A_{u_t}\xi) + O(\|\xi\|^2)$, and $H \in \mathbb{R}^{kN\times \dim\mathfrak{g}}$ is defined by $H\xi = \left(-\left(\mathscr{L}_{\mathfrak{g}}(\xi)d^1\right)^T, \dots, -\left(\mathscr{L}_{\mathfrak{g}}(\xi)d^k\right)^T\right)^T$. The design matrix parameters $\hat{Q}_t, \hat{N}_n, \hat{V}_n, B_n$ 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 3.2. 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: $\hat{V}_n = (\hat{\chi}_{t_n}^{-1}V_n^1, \dots, \hat{\chi}_{t_n}^{-1}V_n^k)^T$, $B_n = (B_n^1, \dots, B_n^k)^T$.

As concerns the RIEKF implementation, A_{u_t} is defined by $g_{u_t}^R(\exp(\xi)) = \mathscr{L}_{\mathfrak{g}}(A_{u_t}\xi) + O(\|\xi\|^2)$, and $H \in \mathbb{R}^{kN\times\dim\mathfrak{g}}$ by $H\xi = \left(\left(\mathscr{L}_{\mathfrak{g}}(\xi)d^1\right)^T, \dots, \left(\mathscr{L}_{\mathfrak{g}}(\xi)d^k\right)^T\right)^T$. The design matrix parameters $\hat{Q}_t, \hat{N}_n, \hat{V}_n, B_n$ are freely assigned by the user. Considering the associated "noisy" system (26)-(28) of Section 3.2 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: $V_n = \left(V_n^1, \dots, V_n^k\right)^T$, $\hat{B}_n = \left(\hat{\chi}_{t_n}B_n^1, \dots, \hat{\chi}_{t_n}B_n^k\right)^T$.

3.4 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 $\left(X_{t_0}^t(x)\right)_{t\geq t_0}$ if there exists $\varepsilon > 0$ such that:*

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

Theorem 4 below is the main result of the paper . It is a consequence of Theorem 2 of Section 2. J. J. Deyst and C. F. 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 $\frac{d}{dt}x_t = A_t x_t$, $y_{t_n} = Hx_{t_n}$ with $x_t \in \mathbb{R}^p$ and let $\Phi_{t_0}^t$ denote the square matrix defined by $\Phi_{t_0}^{t_0} = I_p, \frac{d}{dt}\Phi_{t_0}^t = A_t\Phi_{t_0}^t$. If there exist $\alpha_1,\alpha_2,\beta_1,\beta_2,\delta_1,\delta_2,\delta_3,M$ such that:*

i $(\Phi_{t_n}^{t_{n+1}})^T\Phi_{t_n}^{t_{n+1}} \succeq \delta_1 I_p \succeq 0,$

ii $\exists q \in \mathbb{N}^*, \forall s > 0, \exists G_s \in \mathbb{R}^{p\times q}, Q_s = G_s Q' G_s^T$ *where* $Q' \succeq \delta_2 I_q \succeq 0,$

iii $N_n \succeq \delta_3 I \succeq 0,$

iv $\alpha_1 I_p \leq \int_{s=t_{n-M}}^{t_n} \left(\Phi_s^{t_n}\right) Q_s \left(\Phi_s^{t_n}\right)^T \leq \alpha_2 I_p,$

v $\beta_1 I_p \leq \sum_{i=n-M}^{n-1} \left(\Phi_{t_{i+1}}^{t_n}\right)^T H^T N_n^{-1} H \left(\Phi_{t_{i+1}}^{t_n}\right) \leq \beta_2 I_p.$

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

The main theorem of the present paper is the extension of this linear result to the non-linear case when the Invariant Extended Kalman Filter is used for systems of Section 2.

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 (resp. right) invariant output about χ_t). Then the Left (resp. 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 $\frac{d}{dt}\xi_t = A_{u_t}\xi_t$ 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 non-trivial assumption* that the linearized system around the *estimated* trajectory is well-behaved [11, 25, 24, 10]. 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. 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.*

4 Simplified car example

The computations require only basic knowledge about matrix Lie groups as recalled in Appendix A.

4.1 Considered model

Consider a (non-holonomic) car evolving on the 2D plane. Its heading is denoted by an angle $\theta_t \in [-\pi, \pi]$ and its position by a vector $x_t \in \mathbb{R}^2$. They follow the classical equations (see, e.g., [12]):

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

where v_t is the velocity measured by an odometer and u_t (a function of) the steering angle. Two kinds of observations are considered:

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

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

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 $k \in [1, K]$.

4.2 IEKF gain tuning

To derive and tune the IEKF equations, we follow the methodology of Section 3.2 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.

4.2.1 Associated "noisy" system

Taking into account the possible noise in the measurements we get

$$\begin{gathered} \frac{d}{dt}\theta_t = u_t v_t + w_t^\theta, \qquad \frac{d}{dt}x_t^1 = \cos(\theta_t)(v_t + w_t^l) - \sin(\theta_t) w_t^{tr}, \\ \frac{d}{dt}x_t^2 = \sin(\theta_t)(v_t + w_t^l) + \cos(\theta_t) w_t^{tr}, \end{gathered} \tag{41}$$

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:

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

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

4.2.2 Matrix form

This system can be embedded in the matrix Lie group $SE(2)$ (see Appendix A.1) using the matrices:

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

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

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:

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

$$Y_n^k = \begin{pmatrix} R(\hat{\theta}_{t_n})^T (x_{t_n} - p_k) \\ 1 \end{pmatrix} + \begin{pmatrix} \bar{V}_n^k \\ 0 \end{pmatrix} = -\chi_{t_n}^{-1} \begin{pmatrix} p_k \\ 1 \end{pmatrix} + \begin{pmatrix} \bar{V}_n^k \\ 0 \end{pmatrix}. \tag{46}$$

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

4.2.3 IEKF equations for the left-invariant output (39)

The LIEKF equations (36) for the associated "noisy" system (44), (45) write:

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

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 $\tilde{p} = (I_2, 0_{2,1})$. To compute the gains, we write the left-invariant error $\eta_t = \chi_t^{-1}\hat{\chi}_t$ whose evolution is:

$$\begin{aligned} \frac{d}{dt}\eta_t &= \eta_t \nu_t - \nu_t \eta_t - w_t \eta_t, \\ \eta_{t_n}^+ &= \eta_{t_n} \exp\left(\tilde{L}_n \tilde{p} \left[\eta_{t_n}^{-1} \begin{pmatrix} 0_{2\times 1} \\ 1 \end{pmatrix} - \begin{pmatrix} 0_{2\times 1} \\ 1 \end{pmatrix} + \hat{\chi}_{t_n}^{-1} \begin{pmatrix} V_n \\ 0 \end{pmatrix} \right] \right). \end{aligned} \tag{47}$$

To linearize this equation we introduce the linearized error ξ_t defined replacing η_t with $I_3 + \mathscr{L}_{\mathfrak{se}(2)}(\xi_t)$. Introducing the first-order approximations $\eta_t = I_3 + \mathscr{L}_{\mathfrak{se}(2)}(\xi_t)$, $\eta_t^+ = I_3 + \mathscr{L}_{\mathfrak{se}(2)}(\xi_t^+)$, $\exp(u) = I_3 + \mathscr{L}_{\mathfrak{se}(2)}(u)$ and $\eta_t^{-1} = I_3 - \mathscr{L}_{\mathfrak{se}(2)}(\xi_t)$ in (47) and removing the second-order terms in ξ_t, V_n and w_t we obtain:

$$\frac{d}{dt}\xi_t = -\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -u_t v_t \\ -v_t & u_t v_t & 0 \end{pmatrix} \xi_t - \begin{pmatrix} w_t^\theta \\ w_t^l \\ w_t^{tr} \end{pmatrix}$$

$$\text{and} \qquad \xi_{t_n}^+ = \xi_{t_n} - \tilde{L}_n \left[(0_{2,1}, I_2)\xi_t - R\left(\hat{\theta}_{t_n}\right)^T V_n \right].$$

The gains $\tilde{L}_n$ are thus finally computed using the Riccati equation (37) with:

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

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

4.2.4 IEKF equations for the right-invariant output (43)

The RIEKF equations (36) for the associated "noisy" system (44), (46) write:

$$\frac{d}{dt}\hat{\chi}_t = \hat{\chi}_t \nu_t \ , \ \hat{\chi}_{t_n}^+ = \exp\left(L_n\left[\hat{\chi}_{t_n} Y_n^1 + \begin{pmatrix} p_1 \\ 1 \end{pmatrix}; \ldots; \hat{\chi}_{t_n} Y_n^K + \begin{pmatrix} p_K \\ 1 \end{pmatrix}\right]\right)\hat{\chi}_{t_n}.$$

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:

$$\begin{aligned} \frac{d}{dt}\eta_t &= -(\hat{\chi}_t w_t \hat{\chi}_t^{-1})\eta_t, \\ \eta_{t_n}^+ &= \exp\left(\tilde{L}_n \tilde{p} \begin{pmatrix} -\eta_{t_n}\begin{pmatrix} p_1 \\ 1 \end{pmatrix} + \begin{pmatrix} p_1 \\ 1 \end{pmatrix} + \hat{\chi}_{t_n}\begin{pmatrix} V_n^1 \\ 0 \end{pmatrix} \\ \ldots \\ -\eta_{t_n}\begin{pmatrix} p_K \\ 1 \end{pmatrix} + \begin{pmatrix} p_K \\ 1 \end{pmatrix} + \hat{\chi}_{t_n}\begin{pmatrix} V_n^K \\ 0 \end{pmatrix} \end{pmatrix}\right)\eta_{t_n}. \end{aligned} \tag{48}$$

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 (48) and removing the second-order terms in ξ_t, V_n and w_t we obtain:

$$\begin{aligned} \frac{d}{dt}\xi_t &= -\begin{pmatrix} 1 & 0_{1,2} \\ \hat{x}_t^2 & R(\hat{\theta}_t) \\ -\hat{x}_t^1 & \end{pmatrix}\begin{pmatrix} w_t^\theta \\ w_t^l \\ w_t^{tr} \end{pmatrix}, \\ \xi_{t_n}^+ &= \xi_{t_n} - \tilde{L}_n\left[\begin{pmatrix} \begin{pmatrix} -(p_1)_2 & 1 & 0 \\ (p_1)_1 & 0 & 1 \end{pmatrix} \\ \ldots \\ \begin{pmatrix} -(p_K)_2 & 1 & 0 \\ (p_K)_1 & 0 & 1 \end{pmatrix} \end{pmatrix}\xi_{t_n} - \begin{pmatrix} R(\hat{\theta}_{t_n})V_n^1 \\ \ldots \\ R(\hat{\theta}_{t_n})V_n^K \end{pmatrix}\right]. \end{aligned}$$

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

$$A_t = 0_{3,3}, \quad H_n = \begin{pmatrix} \begin{pmatrix} -(p_1)_2 & 1 & 0 \\ (p_1)_1 & 0 & 1 \end{pmatrix} \\ \ldots \\ \begin{pmatrix} -(p_K)_2 & 1 & 0 \\ (p_K)_1 & 0 & 1 \end{pmatrix} \end{pmatrix},$$

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

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

4.3 Stability properties of the IEKF viewed as a non-linear observer for the simplified car

4.3.1 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 $u_t \leq v_{\max}$ then the LIEKF derived at Section 4.2.4 is an asymptotically stable observer in the sense of Definition 2 about* any *trajectory.*

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 in practice an arbitrary high velocity is unfeasible.

4.3.2 Stability of the IEKF for the right-invariant output (40)

Proposition 4. *If at least two distinct points are observed then the IEKF is an asymptotically stable observer in the sense of Definition 2 about* any *bounded trajectory.*

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. □

4.4 Simulations

The IEKF described in Section 4.2 has been implemented and compared to a classical EKF for the experimental setting described by Figure 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 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 $Q = diag((\pi/180)^2, 10^{-4}, 10^{-4})$ 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 Figure 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.

5 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 [17, 16].

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 non-linear 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.

5.1 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:

$$\frac{d}{dt}R_t = R_t(\omega_t)_\times \quad , \quad \frac{d}{dt}v_t = g + R_t u_t \quad , \quad \frac{d}{dt}x_t = v_t, \tag{49}$$

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:

$$(Y_n^1, ..., Y_n^k) = \left(R_{t_n}^T(p_1 - x_{t_n}), ..., R_{t_n}^T(p_k - x_{t_n})\right), \tag{50}$$

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

5.2 IEKF gain tuning

To derive and tune the IEKF equations, we follow the methodology of Section 3.2 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.

5.2.1 Associated "noisy" system

By merely introducing noise in the accelerometers' and gyrometers' measurements we obtain the well-known equations [14]:

$$\frac{d}{dt}R_t = R_t(\omega_t + w_t^{\omega})_\times, \quad \frac{d}{dt}v_t = g + R_t(u_t + w_t^u), \quad \frac{d}{dt}x_t = v_t. \tag{51}$$

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

$$(Y_n^1, ..., Y_n^k) = \left(R_{t_n}^T(p_1 - x_{t_n}) + V_n^1, ..., R_{t_n}^T(p_k - x_{t_n}) + V_n^k\right), \tag{52}$$

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

5.2.2 Matrix form

As already noticed in the preliminary work [4], the system (51) can be embedded in the group of double homogeneous matrices (see Appendix A.2) using the matrices χ_t, w_t and function $f_{\omega,u}$:

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

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

The equation of the dynamics becomes:

$$\frac{d}{dt}\chi_t = f_{\omega_t,u_t}(\chi_t) + \chi_t w_t, \tag{53}$$

and the observations (52) have the equivalent forms:

$$(Y_n^1, ..., Y_n^k) = \left(\chi_{t_n}^{-1}\begin{pmatrix} p_1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} V_n^1 \\ 0 \\ 0 \end{pmatrix}, ..., \chi_{t_n}^{-1}\begin{pmatrix} p_k \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} V_n^k \\ 0 \\ 0 \end{pmatrix}\right). \tag{54}$$

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.*

5.2.3 IEKF equations

The RIEKF (36) for the associated "noisy" system (53), with right-invariant "noisy" observations (54) reads:

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

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:

$$\frac{d}{dt}\eta_t = f_{\omega_t,u_t}(\eta_t) - \eta_t f_{\omega_t,u_t}(I_5) - (\hat{\chi}_t w_t \hat{\chi}^{-1})\eta_t, \tag{55}$$

$$\eta_{t_n}^+ = \exp\left(\tilde{L}_n \tilde{p} \begin{pmatrix} \eta_{t_n}\begin{pmatrix} p_1 \\ 0 \\ 1 \end{pmatrix} + \hat{\chi}_{t_n}\begin{pmatrix} V_n^1 \\ 0_{2,1} \end{pmatrix} \\ \dots \\ \eta_{t_n}\begin{pmatrix} p_k \\ 0 \\ 1 \end{pmatrix} + \hat{\chi}_{t_n}\begin{pmatrix} V_n^k \\ 0_{2,1} \end{pmatrix} \end{pmatrix}\right)\eta_{t_n}. \tag{56}$$

To linearize this equation we introduce the linearized error ξ_t by replacing η_t with $I_3 + \mathscr{L}_{\mathfrak{se}(3),2}(\xi_t)$. Using the first order approximations $\eta_t = I_3 + \mathscr{L}_{\mathfrak{se}(3),2}(\xi_t)$, $\eta_t^+ = I_3 + \mathscr{L}_{\mathfrak{se}(3),2}(\xi_t^+)$, $\exp(u) = I_3 + \mathscr{L}_{\mathfrak{se}(3),2}(u)$ and $\eta_t^{-1} = I_3 - \mathscr{L}_{\mathfrak{se}(3),2}(\xi_t)$ in (55), (56) and removing the second-order terms in ξ_t, V_n and w_t we obtain:

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

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

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

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

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

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

5.3 Stability properties of the IEKF viewed as a non-linear observer for the navigation example

Theorem 6. *If three non-collinear points are observed, then the IEKF whose equations are derived in Section 5.2.3 is an asymptotically stable observer for the system* (49)-(50) *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

$$\Phi = \begin{pmatrix} I_3 & 0_{3\times 3} & 0_{3\times 3} \\ t(g)_\times & I_3 & 0_{3\times 3} \\ \frac{1}{2}t^2(g)_\times & tId_{3\times 3} & I_3 \end{pmatrix}.$$

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 non-collinear 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 $\mathscr{H}_1$:

$$\mathscr{H}_1 = \begin{pmatrix} (p_1)_\times & 0_{3\times 3} & -I_3 \\ (p_2)_\times & 0_{3\times 3} & -I_3 \\ (p_3)_\times & 0_{3\times 3} & -I_3 \\ (p_1)_\times - \frac{1}{2}t^2(g)_\times & -tI_3 & -I_3 \\ (p_2)_\times - \frac{1}{2}t^2(g)_\times & -tI_3 & -I_3 \\ (p_3)_\times - \frac{1}{2}t^2(g)_\times & -tI_3 & -I_3 \end{pmatrix}, \qquad \mathscr{H}_2 = \begin{pmatrix} (p_1)_\times & 0_{3\times 3} & -I_3 \\ (p_2)_\times & 0_{3\times 3} & -I_3 \\ (p_3)_\times & 0_{3\times 3} & -I_3 \\ -\frac{1}{2}t^2(g)_\times & -tI_3 & 0_{3\times 3} \\ -\frac{1}{2}t^2(g)_\times & -tI_3 & 0_{3\times 3} \\ -\frac{1}{2}t^2(g)_\times & -tI_3 & 0_{3\times 3} \end{pmatrix},$$

$$\mathscr{H}_3 = \begin{pmatrix} (p_1 - p_3)_\times & 0_{3\times 3} & 0_{3\times 3} \\ (p_2 - p_3)_\times & 0_{3\times 3} & 0_{3\times 3} \\ \frac{1}{2}t^2(g)_\times & tI_3 & 0_{3\times 3} \\ (p_3)_\times & 0_{3\times 3} & -I_3 \end{pmatrix}.$$

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. □

5.4 Simulations

The IEKF described in Section 5.2.3 has been implemented and compared to a state of the art multiplicative EKF [21] for the experimental setting described by Figure 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 of the MEKF:

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

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

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

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

$$Q_1 = \begin{pmatrix} 10^{-8} I_3 & 0_3 & 0_3 \\ 0_3 & 10^{-8} I_3 & 0_3 \\ 0_3 & 0_3 & 0_3 \end{pmatrix}, \quad Q_2 = \begin{pmatrix} 10^{-4} I_3 & 0_3 & 0_3 \\ 0_3 & 10^{-4} I_3 & 0_3 \\ 0_3 & 0_3 & 0_3 \end{pmatrix}.$$

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, although reasonable in the context of high-precision inertial navigation, has been deliberately chosen to challenge EKF-like methods: the corresponding gains are small so the errors introduced during the transitory phase due to non-linearities 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 Figure 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 non-linearity 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 non-linear 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 that 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 choices of Q and N.

6 Conclusion

The Invariant EKF, when used as a deterministic observer for an introduced and well characterized 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 EKF in terms of tuning, implementation, and computational load.

A Matrix Lie groups useful formulas

A matrix Lie group G is a set of square invertible matrices of size $N \times N$ verifying the following properties:

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

If $\gamma(t)$ is a process taking values in G and verifying $\gamma(0) = Id$, then its derivative at $t = 0$ cannot take any value in the set of squared matrices $\mathscr{M}_N(\mathbb{R})$. It is constrained to lie in a vector subspace $\mathfrak{g}$ of $\mathscr{M}_N(\mathbb{R})$ called the "Lie algebra of G". As it proves useful to identify $\mathfrak{g}$ to $\mathbb{R}^{\dim \mathfrak{g}}$, a linear mapping $\mathscr{L}_{\mathfrak{g}} : \mathbb{R}^{\dim \mathfrak{g}} \to \mathfrak{g}$ is used in this paper. The vector space $\mathfrak{g}$ can be mapped to the matrix Lie group G through the classical matrix exponential $\exp_m$. As well, $\mathbb{R}^{\dim \mathfrak{g}}$ can be mapped to G through a function exp defined by $\exp(\xi) = \exp_m(\mathscr{L}_{\mathfrak{g}}(\xi))$. For any $g \in G$ the operator $Ad_g : \mathbb{R}^{\dim \mathfrak{g}} \to \mathbb{R}^{\dim \mathfrak{g}}$ is defined by $g\mathscr{L}_{\mathfrak{g}}(\xi)g^{-1} = \mathscr{L}_{\mathfrak{g}}(Ad_g\xi)$. For any $x \in \mathbb{R}^{\dim \mathfrak{g}}$ the operator $Ad_x : \mathbb{R}^{\dim \mathfrak{g}} \to \mathbb{R}^{\dim \mathfrak{g}}$ is defined by $\mathscr{L}_{\mathfrak{g}}(x)\mathscr{L}_{\mathfrak{g}}(\xi) - \mathscr{L}_{\mathfrak{g}}(\xi)\mathscr{L}_{\mathfrak{g}}(x) = \mathscr{L}_{\mathfrak{g}}(Ad_x\xi)$. These operators are very handy in practical computations. For all matrix Lie groups considered in this paper, no matrix exponentiation is actually needed as there exist closed formulas, given thereafter. We give now a short description of the matrix Lie groups appearing in the present paper.

A.1 Group of direct planar isometries $SE(2)$

We have here $G = SE(2)$ and $\mathfrak{g} = \mathfrak{se}(2)$, where

$$SE(2) = \left\{ \begin{pmatrix} R(\theta) & x \\ 0_{1,2} & 1 \end{pmatrix}, \begin{pmatrix} \theta \\ x \end{pmatrix} \in \mathbb{R}^3 \right\},$$

$$\mathfrak{se}(2) = \left\{ \begin{pmatrix} 0 & -\theta & u_1 \\ \theta & 0 & u_2 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} \theta \\ u_1 \\ u_2 \end{pmatrix} \in \mathbb{R}^3 \right\},$$

$$\mathscr{L}_{\mathfrak{se}(2)} \begin{pmatrix} \theta \\ u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 & -\theta & u_1 \\ \theta & 0 & u_2 \\ 0 & 0 & 0 \end{pmatrix},$$

where $R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ denotes the rotation of angle θ, and the exponential mapping is:

$$\exp\left((\theta, u_1, u_2)^T\right) = \begin{pmatrix} R(\theta) & x \\ 0_{1,2} & 1 \end{pmatrix},$$

where

$$x = \begin{pmatrix} \frac{\sin(\theta)}{\theta} & -\frac{1-\cos(\theta)}{\theta} \\ \frac{1-\cos(\theta)}{\theta} & \frac{\sin(\theta)}{\theta} \end{pmatrix}.$$

A.2 Group of double direct spatial isometries $SE_2(3)$

We have here:

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

$$\mathfrak{g} = \mathfrak{se}_2(3) = \left\{ \begin{pmatrix} (\xi)_\times & u & y \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{pmatrix}, \xi, u, y \in \mathbb{R}^3 \right\}.$$

An isomorphism between $\mathbb{R}^9$ and $\mathfrak{se}_2(3)$ is given by $\mathscr{L}_{\mathfrak{se}_2(3)} \begin{pmatrix} \xi \\ u \\ y \end{pmatrix} = \begin{pmatrix} (\xi)_\times & u & y \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{pmatrix}$. The exponential mapping is given by the formula:

$$\begin{aligned} &\exp\left(\left(\xi^T, u^T, y^T\right)^T\right) \\ &\quad = I_5 + S + \frac{1-\cos(||\xi||)}{||\xi||^2})S^2 + \frac{||\xi|| - \sin(||\xi||)}{||\xi||^3} S^3, \end{aligned}$$

where $S = \mathscr{L}_{\mathfrak{se}_2(3)}(\xi, u, y)^T$.

B Further explanation and proof of the log-linear property

The definition of A_t through $g_{u_t}(\exp(\xi)) = \mathscr{L}_{\mathfrak{g}}(A_{u_t}\xi) + O(\|\xi_t\|^2)$ is very convenient in practice as shown in Sections 4 and 5, but it requires a quick theoretical explanation as follows. On an abstract Lie group, vector $g_{u_t}(\exp(\xi))$ belongs to $T_{\exp(\xi)}G$ (tangent space at $\exp(\xi)$) whereas the image of $\mathscr{L}_{\mathfrak{g}}$ is $\mathfrak{g} = T_{Id}G$. But as $g_{u_t}(Id) = 0$ it is true that the first order terms of $g_{u_t}(\exp(\xi))$ belong to $\mathfrak{g} \subset \mathbb{R}^{N\times N}$. Indeed as $\exp(\xi)^{-1} g(\exp(\xi)) \in \mathfrak{g}$ we have $g(\exp(\xi)) = \exp_m(\mathscr{L}_{\mathfrak{g}}(\xi))\mathscr{L}_{\mathfrak{g}}(\omega(\xi))$ for some function ω with $\omega(0) = 0$. Thus $g(\exp(\xi)) = (Id + \mathscr{L}_{\mathfrak{g}}(\xi))\mathscr{L}_{\mathfrak{g}}(\omega(\xi)) + O(\|\xi\|^2) = \mathscr{L}_{\mathfrak{g}}(\omega(\xi)) + O(\|\xi\|^2)$.

The proof of Theorem 2 is based upon the following lemmas.

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

$$\forall (u,a,b) \in G \quad g_u(ab) = ag_u(b) + g_u(a)b \tag{57}$$

We have the following properties:

- *The function* $g_u^R(\eta) = f_{u_t}(\eta) - \eta f_{u_t}(Id)$ *verifies* (57) *and all the solutions of* (12) *have the form* $\chi_t = \eta_t^R \bar{\chi}_t$*, where* η_t^R *verifies* $\frac{d}{dt}\eta_t^R = g_{u_t}^R(\eta_t^R)$.
- *The function* $g_{u_t}^L(\eta) = f_{u_t}(\eta) - f_{u_t}(Id)\eta$ *verifies* (57) *and all the solutions of* (12) *have the form* $\chi_t = \bar{\chi}_t \eta_t^L$*, where* η_t^L *verifies* $\frac{d}{dt}\eta_t^L = g_{u_t}^L(\eta_t^L)$.

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* $\frac{d}{dt}\eta_t = g_{u_t}(\eta_t)$*, where* g_{u_t} *verifies* (57)*. 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 $\frac{d}{dt}\eta_t = g_{u_t}(\eta_t)$:

$$\begin{aligned}\frac{d}{dt}[\Phi_t(\eta_0)\Phi_t(\eta_0')] &= g_{u_t}(\Phi_t(\eta_0))\Phi_t(\eta_0') + \Phi_t(\eta_0)g_{u_t}(\Phi_t(\eta_0')) \\ &= g_{u_t}\left(\Phi_t(\eta_0)\Phi_t(\eta_0')\right)\end{aligned}$$

□

An immediate recursion gives then:

Lemma 3. *We have furthermore*

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

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* $\frac{d}{dt}\eta_t = g_{u_t}(\eta_t)$ *satisfying* (57)*. We have:*

$$\Phi_t(\exp(\xi_0)) = \exp(F_t\xi_0)$$

where F_t *is the solution of the matrix equation* $F_0 = Id, \frac{d}{dt}F_t = A_tF_t$ *.*

Proof. Thanks to Lemma 3 we have, for any $n \in \mathbb{N}$, $\Phi_t\left(e^{\xi_0}\right) = \Phi_t\left(\left[e^{\frac{1}{n}\xi_0}\right]^n\right) = \Phi_t\left(e^{\frac{1}{n}\xi_0}\right)^n = \left[e^{\frac{1}{n}F_t\xi_0 + r_t\left(\frac{1}{n}\xi_0\right)}\right]^n$, where $F_t = D\left(\exp^{-1}\circ\Phi_t\circ\exp\right)\big|_0$ and $r_t\left(\frac{1}{n}\xi_0\right)$ is a quadratic term, which ensures in turn $\Phi_t\left(e^{\xi_0}\right) = e^{F_t\xi_0 + nr_t\left(\frac{1}{n}\xi_0\right)}$. Letting $n \to \infty$ we get $\Phi_t\left(e^{\xi_0}\right) = e^{F_t\xi_0}$. Differentiating both sides of the latter equality we obtain $g_{u_t}\left(\Phi_t\left(e_0^{\xi}\right)\right) = D\exp_{F_t\xi_0}\left[\frac{d}{dt}F_t\xi_0\right]$. A first-order expansion in ξ_0 using matrix A_t gives: $\mathscr{L}_{\mathfrak{g}}\left(A_tF_t\xi_0\right) + \circ\left(||\xi_0||\right) = \mathscr{L}_{\mathfrak{g}}\left(\frac{d}{dt}F_t\xi_0\right) + \circ\left(||\xi_0||\right)$, then $A_tF_t\xi_0 = \frac{d}{dt}F_t\xi_0$ for any $\xi_0 \in \mathbb{R}^{\dim\mathfrak{g}}$ and finally $A_t = \frac{d}{dt}F_t$. □

C Proof of theorem 4

C.1 Proof rationale

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

$$\forall t \geq 0, \qquad \xi_t = \Psi^t_{t_0}\xi_0 + \sum_{t_n<t}\Psi^t_{t_n} r_n(\xi_{t_n}) \tag{58}$$

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 $\Psi^t_{t_0}$ (Theorem 3).

C.2 Review of existing linear results

Consider a linear time-varying Kalman filter and let $\Psi^t_{t_0}$ 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 $\Psi^t_s\xi_s$ verifies for $V(P,\xi) = \xi^T P^{-1}\xi$:

$$V\left(P^+_{t_{n+N}}, \Psi^{t_{n+N}}_{t_n}(\xi^+_{t_n})\right) \leq V(P^+_{t_n}, \xi^+_{t_n}) - \beta_3||\Psi^{t_{n+N}}_{t_n}(\xi^+_{t_n})||^2 \tag{59}$$

where β_3 only depends on $\alpha_1, \alpha_2, \beta_1, \beta_2, \delta_1, \delta_2, \delta_3, N$. Of course, the proof given in [13] holds if the inequalities are only verified on an interval $[0,T]$. We will also use the direct consequence:

$$V\left(P^+_{t_{n+1}}, \Psi^{t_{n+1}}_{t_n}(\xi^+_{t_n})\right) < V\left(P^+_{t_n}, \xi^+_{t_n}\right) \tag{60}$$

C.3 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 $\hat{\delta}_1 = \delta_1, \hat{\delta}_2 = \frac{1}{k^2}\delta_2, \hat{\delta}_3 = \frac{1}{k^2}\delta_3, \hat{\alpha}_1 = \frac{1}{k^2}\alpha_1, \hat{\alpha}_2 = k^2\alpha_2, \hat{\beta}_1 = \frac{1}{k^2}\beta_1, \hat{\beta}_2 = k^2\beta_2$. Moreover, if $\frac{1}{k}\tilde{P}_{t_0} \leq P_{t_0} \leq k\tilde{P}_{t_0}$ then $\frac{1}{k}\tilde{P}_{t_0+s} \leq P_{t_0+s} \leq k\tilde{P}_{t_0+s}$ 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: $\hat{Q}_t = Cov(w_t)$ if the noise term has the form $\chi_t w_t$, $\hat{Q}_t = Ad_{\hat{\chi}_t^{-1}} Cov(w_t) Ad^T_{\hat{\chi}_t^{-1}}$ if

the noise term has the form $w_t\chi_t$, and $\hat{N}_n = \hat{\chi}_{t_n}^{-1} Cov\,(V_n)\,\hat{\chi}_{t_n}^{-T} + Cov\,(B_n)$. 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 $\hat{Q}_t = Q_1 + Ad_{\hat{\chi}_t^{-1}} Q_2 Ad^T_{\hat{\chi}_t^{-1}}$ 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:

$$\frac{d}{dt}\tilde{P}_t = A_t\tilde{P}_t + \tilde{P}_t A_t^T + Q_1 + Ad_{\chi_t} Q_2 Ad^T_{\chi_t},$$
$$\tilde{P}^+_{t_n} = \tilde{P}_{t_n} - \tilde{P}_{t_n} H^T \left(H\tilde{P}_{t_n} H^T + N_1 + \chi_{t_n}^{-1} N_2 \chi_{t_n}^{-T}\right)^{-1} H\tilde{P}_{t_n}.$$

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:

$$\frac{d}{dt}P_t = A_t P_t + P_t A_t^T + Q_1 + Ad_{\eta_t^{-1}}\left[Ad_{\chi_t^{-1}} Q_2 Ad^T_{\chi_t^{-1}}\right] Ad^T_{\eta_t^{-1}},$$

$$P^+_{t_n} = P_{t_n} - P_{t_n} H^T \left(HP_{t_n} H^T + N_1 + \eta_{t_n}^{-1}\left[\chi_{t_n}^{-1} N_2 \chi_{t_n}^{-T}\right]\eta_{t_n}^{-T}\right)^{-1} HP_{t_n}.$$

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 $||\xi_{t_0+s}|| \leq \varepsilon$ ensures $\frac{1}{k}N_2 \preceq \left(e^{-\xi_{t_0+s}}\right) N_2 \left(e^{-\xi_{t_0+s}}\right)^T \preceq kN_2$ and $\frac{1}{k}Q_2 \preceq Ad_{e^{-\xi_{t_0+s}}} Q_2 Ad^T_{e^{-\xi_{t_0+s}}} \preceq kQ_2$. It ensures consequently

$$\frac{1}{k}\left(Q_1 + Ad_{\chi_{t_0+s}^{-1}} Q_2 Ad^T_{\chi_{t_0+s}^{-1}}\right) \preceq Q_1 + Ad_{\hat{\chi}_{t_0+s}^{-1}} Q_2 Ad^T_{\hat{\chi}_{t_0+s}^{-1}} \preceq k\left(Q_1 + Ad_{\chi_{t_0+s}^{-1}} Q_2 Ad^T_{\chi_{t_0+s}^{-1}}\right)$$

and

$$\frac{1}{k}\left(N_1 + \chi_{t_0+s}^{-1} N_2 \chi_{t_0+s}^{-T}\right) \preceq N_1 + \hat{\chi}_{t_0+s}^{-1} N_2 \hat{\chi}_{t_0+s}^{-T} \preceq k\left(N_1 + \chi_{t_0+s}^{-1} N_2 \chi_{t_0+s}^{-T}\right),$$

and a mere look at the definitions of the constants of Theorem 3 yields the modified constants.

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

□

Lemma 5. [first-order control of growth] *Under the same conditions as in Lemma 4 (including $\frac{1}{k}\tilde{P}_{t_0} \leq P_{t_0} \leq k\tilde{P}_{t_0}$) 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 $||\xi_{t_0+s}|| \leq l_1(||\xi_{t_0}||)$ 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 non-linear rest $r_{t_n}(\xi)$ introduced in (58) is defined by $\exp(\xi)\exp(L_n(e^{\xi}b - b - H\xi)) = \exp((I - L_n H)\xi + r_{t_n}(\xi))$. 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 $\frac{\gamma_{\max}||H||}{\delta_3}$ 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 $||r_{t_n}(\xi)|| \leq \tilde{l}^k(||\xi||)$ for any n such that $t_n \leq 2M\Delta T$.

Now we can control the evolution of the error using $\tilde{l}^k$. The propagation step is linear, thus we have the classical result $\frac{d}{dt}V_t(\xi_t) < 0$. It ensures $||\xi_{t_0+s}|| < \sqrt{\frac{\gamma_{\max}}{\gamma_{\min}}}||\xi_{t_0}||$ as long as there is no update on $[t_0, t_0+s]$. At each update step we have $V_{t_n}^+(\xi_{t_n}^+)^{1/2} = V_{t_n}^+([I - L_nH]\xi_{t_n} + r_n(\xi_{t_n}))^{1/2}$ $\leq V_{t_n}^+([I-L_nH]\xi_{t_n})^{1/2} + V_{t_n}^+(r_n(\xi_{t_n}))^{1/2} \leq V_{t_n}(\xi_{t_n})^{1/2} + V_{t_n}^+(r_n(\xi_{t_n}))^{1/2}$ using the triangular inequality. Thus: $||\xi_{t_n}^+|| \leq \sqrt{\frac{\gamma_{\max}}{\gamma_{\min}}}(||\xi_{t_n}|| + ||r_n(\xi_{t_n})||) \leq \sqrt{\frac{\gamma_{\max}}{\gamma_{\min}}}(||\xi_{t_n}|| + \tilde{l}^k(||\xi_{t_n}||))$. 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 $\frac{1}{k}\tilde{P}_{t_0} \leq P_{t_0} \leq k\tilde{P}_{t_0}$) 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 $V_{t_0+s}(\xi_{t_0+s}) \leq V_{t_0+s}(\Psi_{t_0}^{t_0+s}\xi_{t_0}) + l_2(||\xi_{t_0}||) \leq V_{t_0}(\xi_{t_0}) + l_2(||\xi_{t_0}||)$ for any $s \in [0, 2M\Delta t]$ with $l_2(x) = O(x^2)$. We also have $V_{t_n}(\xi_{t_n}^+) \leq V_{t_n}^+((\Psi_{t_0}^{t_n})^+\xi_{t_0}) + l_2(||\xi_{t_0}||) \leq V_{t_0}(\xi_{t_0}) + l_2(||\xi_{t_0}||)$ for $t_0 \leq t_n \leq t_0 + 2M\Delta T$.*

Proof. The result stems from the decomposition (58) as:

$$V_{t_0+s}(\xi_{t_0+s})^{1/2} = V_{t_0+s}\left(\Psi_{t_0}^{t_0+s}\xi_{t_0} + \sum_{t_0<t_n<t_0+s}\Psi_{t_n}^{t_0+s}r_n(\xi_{t_n})\right)^{1/2}$$

$$\leq V_{t_0+s}\left(\Psi_{t_0}^{t_0+s}\xi_{t_0+s}\right)^{1/2} + \sum_{t_0<t_n<t_0+s}V_{t_0+s}\left(\Psi_{t_n}^{t_0+s}r_n(\xi_{t_n})\right)^{1/2}$$

(triangular inequality)

$$\leq V_{t_0+s}\left(\Psi_{t_0}^{t_0+s}\xi_{t_0+s}\right)^{1/2} + \sum_{t_0<t_n<t_0+s}V_{t_n}(r_n(\xi_{t_n}))^{1/2} \qquad \text{(using (60))}$$

$$\leq V_{t_0+s}\left(\Psi_{t_0}^{t_0+s}\xi_{t_0+s}\right)^{1/2} + \sum_{t_0<t_n<t_0+s}\sqrt{\frac{\gamma_{\max}}{\gamma_{\min}}}||r_n(\xi_{t_n})||^{1/2}$$

(from the def. of V)

$$\leq V_{t_0+s}\left(\Psi_{t_0}^{t_0+s}\xi_{t_0+s}\right)^{1/2} + \sum_{t_0<t_n<t_0+s}\sqrt{\frac{\gamma_{\max}}{\gamma_{\min}}}(\tilde{l}^k \circ l_1^k)(||\xi_{t_0}||)^{1/2}$$

(from Lemma 5).

As we have $(\tilde{l}^k \circ l_1^k)(x) = O(x^2)$, we obtain the result squaring the inequality and using $V_{t_0+s}\left(\Psi_{t_0}^{t_0+s}\xi_{t_0+s}\right)$ $\leq V_{t_0}(\xi_{t_0}) \leq \frac{\gamma_{\max}}{\gamma_{\min}}||\xi_{t_0}||$ to control the crossed terms. □

Lemma 7. [final second order growth control] *Under the same conditions as in Lemma 4 (including $\frac{1}{k}\tilde{P}_{t_0} \leq P_{t_0} \leq k\tilde{P}_{t_0}$) 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:

$$V_{t_0+s}(\xi_{t_0+s}) \leq V_{t_0}(\xi_{t_0}) + l_1^k(\xi_{t_0}) \tag{61}$$

$$- \sum_{i=0}^{J-1} \left[\beta^k ||\xi_{t_{n_0}+iM\Delta t}||^2 - l_2^k(\xi_{t_{n_0}+iM\Delta t}) \right] + l_1^k(||\xi^+_{t_{n_{\max}}}||), \tag{62}$$

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. $J = \max\left\{j, t_{n_{\max}-jM} \geq t_0\right\} - 1$) 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 $V_{t_0+s}(\xi_{t_0+s}) - V^+_{t_0}(\xi^+_{t_0}) = \left(V_{t_0+s}(\xi_{t_0+s}) - V^+_{t_{n_{\max}}}(\xi^+_{t_{n_{\max}}})\right) + \left(V^+_{t_{n_{\max}}}(\xi^+_{t_{n_{\max}}}) - V^+_{t_{n_0}}(\xi^+_{t_{n_0}})\right) + \left(V^+_{t_{n_0}}(\xi^+_{t_{n_0}}) - V_{t_0}(\xi_{t_0})\right)$. The first and third terms are upper bounded using Lemma 6. The second term is controlled as follows:

$$\begin{aligned}
&V^+_{t_{n_{\max}}}(\xi^+_{t_{n_{\max}}}) - V^+_{t_{n_0}}(\xi^+_{t_{n_0}}) \\
&= \sum_{i=0}^{J-1} \left[V^+_{t_{n_0+(i+1)M}}(\xi^+_{t_{n_0+(i+1)M}}) - V^+_{t_{n_0+iM}}(\xi^+_{t_{n_0+iM}}) \right] \\
&\leq \sum_{i=0}^{J-1} \left[V^+_{t_{n_0+(i+1)M}}(\Psi^{t_{n_0+(i+1)M}+}_{t_{n_0+iM}} \xi^+_{t_{n_0+iM}}) - V^+_{t_{n_0+iM}}(\xi^+_{t_{n_0+iM}}) + l_2(\xi^+_{t_{n_0+iM}}) \right].
\end{aligned}$$

And we conclude using (see [13]):

$$\begin{aligned}
&V^+_{t_{n_0+(i+1)M}}(\Psi^{t_{n_0+(i+1)M}+}_{t_{n_0+iM}} \xi^+_{t_{n_0+iM}}) - V^+_{t_{n_0+iM}}(\xi^+_{t_{n_0+iM}}) \leq \\
&\qquad - \tilde{\beta}^k ||\Psi^{t_{n_0+(i+1)M}+}_{t_{n_0+iM}} \xi^+_{t_{n_0+iM}}||^2 \leq -\tilde{\beta}^k (\frac{\gamma_{\min}}{\gamma_{\max}} \delta_1)^M ||\xi^+_{t_{n_0+iM}}||^2,
\end{aligned}$$

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

Remark 3. *The control we have obtained on ξ_{t_0+s} is verified if $||\xi_{t_0+s}||$ is* already *in a ball of radius ε over the whole interval $[t_0, t_0+t]$. We now prove the result holds assuming* only *that ξ_{t_0} is sufficiently small.*

C.4 Proof of theorem 4

Applying Lemma 7 with $t_0+s = t_{n_{\max}}$ gives for $\frac{1}{k}\tilde{P}_{t_0} \leq P_{t_0} \leq k\tilde{P}_{t_0}$ and $||\xi_{t_0+s}|| < \varepsilon$ on $[0,t]$:

$$\begin{aligned}
||\xi^+_{t_{n_{\max}}}||^2 \leq &\frac{\gamma_{\max}}{\gamma_{\min}} ||\xi_{t_0}||^2 + \gamma_{\max} l_1^k(||\xi_{t_0}||) \\
&- \frac{\gamma_{\max}}{\gamma_{\min}} \sum_{i=0}^{J-1} \left[\beta^k ||\xi^+_{t_{n_0+iM}}||^2 - l_2(\xi^+_{t_{n_0+iM}}) \right].
\end{aligned}$$

There exist $K>0$ and $\varepsilon'>0$ such that for $x<\varepsilon'$, we have $l_2(x)<\frac{\beta^k}{2}x$ and $\gamma_{\max}l_1^k(x)<Kx$ (as $l_2(x)=O(x^2)$ and $l_1^k(x)=O(x^2)$) which gives: $||\xi^+_{t_{n_{\max}}}||^2\leq\left(\frac{\gamma_{\max}}{\gamma_{\min}}+K\right)||\xi_{t_0}||^2$. Thus, for $||\xi_{t_0}||<\frac{\varepsilon'}{\sqrt{\frac{\gamma_{\max}}{\gamma_{\min}}+K}}$:

$$||\xi_{t_0+s}||^2\leq\left(\frac{\gamma_{\max}}{\gamma_{\min}}+K+K\left(\frac{\gamma_{\max}}{\gamma_{\min}}+K\right)\right)||\xi_{t_0}||^2 - \frac{\gamma_{\max}}{\gamma_{\min}}\sum_{i=0}^{J-1}\frac{\beta^k}{2}||\xi^+_{t_{n_0+iM}}||^2, \tag{63}$$

which finally ensures

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

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 $t=\inf\left\{s,||\xi_{t_0+s}||\geq\frac{3}{4}\varepsilon'\right\}$ for sufficiently small $||\xi_{t_0}||$ we end up with a contradiction if we suppose $t<+\infty$, which proves $t=+\infty$. All the previous results thus hold *only* for sufficiently small $||\xi_{t_0}||$.

Moreover, (63) shows that $\sum_{i=0}^{J-1}\frac{\beta^k}{2}||\xi^+_{t_{n_0+iM}}||^2$ 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.

D Proof of proposition 3

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

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

as the eigenvalues of $\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix}$ are $(1,-1,0)$. Thus, $\forall z\in\mathbb{R}^3,\frac{d}{dt}\log\left(z^T\left(\Phi^t_{t_n}\right)^T\Phi^t_{t_n}z\right)\geqslant-v_{\max}$ and finally $z^T(\Phi^{t_{n+1}}_{t_n})^T\Phi^{t_{n+1}}_{t_n}z\geqslant\exp\left(-v_{\max}(t_{n+1}-t_n)\right)||z||^2$ as $(\Phi^{t_n}_{t_n})^T\Phi^{t_n}_{t_n}=I_3$. Thus $(\Phi^{t_{n+1}}_{t_n})^T\Phi^{t_{n+1}}_{t_n}\succeq\exp\left(-v_{\max}(t_{n+1}-t_n)\right)I_3$ and (i) is verified. The difficult part of (v) is the lower bound. Denoting $Cov(V_n)$ by N we will show:

$$\exists\beta_1,\forall n\in\mathbb{N},\quad\beta_1I_3\leq\hat{R}^T_{t_n}N^{-1}\hat{R}_{t_n} + (\Phi^{t_{n+1}}_{t_n})^T\begin{pmatrix}0_{1,2}&I_2\end{pmatrix}^T\hat{R}^T_{t_{n-1}}N^{-1}\hat{R}_{t_{n-1}}\begin{pmatrix}0_{1,2}&I_2\end{pmatrix}\Phi^{t_{n+1}}_{t_n}.$$

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

$$M\begin{pmatrix}\theta\\u\end{pmatrix}=\begin{pmatrix}\theta\\u\end{pmatrix}^T\hat{R}^T_{t_n}N^{-1}\hat{R}_{t_n}\begin{pmatrix}\theta\\u\end{pmatrix} + \begin{pmatrix}\theta\\u\end{pmatrix}^T(\Phi^{t_{n+1}}_{t_n})^T\begin{pmatrix}0_{1,2}&I_2\end{pmatrix}^T\hat{R}^T_{t_{n-1}}N^{-1}\hat{R}_{t_{n-1}}\begin{pmatrix}0_{1,2}&I_2\end{pmatrix}\Phi^{t_{n+1}}_{t_n}\begin{pmatrix}\theta\\u\end{pmatrix}.$$

We decompose $\Phi_{t_n}^{t_{n+1}}$ as $\Phi_{t_n}^{t_{n+1}} = \begin{pmatrix} 1 & 0 & 0 \\ \delta V_n & & T_n \end{pmatrix}$. To simplify the writing we introduce the norms $||x||_N^2 = x^T N^{-1} x$ 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 $M \begin{pmatrix} \theta \\ u \end{pmatrix} = ||\hat{R}_{t_n} u||_N^2 + ||\theta \hat{R}_{t_{n-1}} \delta V_n + \hat{R}_{t_{n-1}} T_n u||_N^2 = ||\hat{R}_{t_n} u||_N^2 + \theta^2 ||\hat{R}_{t_{n-1}} \delta V_n||_N^2 + 2\theta \left\langle \hat{R}_{t_{n-1}} \delta V_n, \hat{R}_{t_{n-1}} T_n u \right\rangle_N + ||\hat{R}_{t_{n-1}} T_n u||_N^2$ and for $\lambda \in]0,1]$ we have:

$$
\begin{aligned}
M \begin{pmatrix} \theta \\ u \end{pmatrix} =& ||\hat{R}_{t_n} u||_N^2 + (1-\lambda^2)\theta^2 ||\hat{R}_{t_{n-1}} \delta V_n||_N^2 + \lambda^2 \theta^2 ||\hat{R}_{t_{n-1}} \delta V_n||_N^2 \\
&+ 2\theta \left\langle \hat{R}_{t_{n-1}} \delta V_n, \hat{R}_{t_{n-1}} T_n u \right\rangle_N + ||\hat{R}_{t_{n-1}} T_n u||_N^2 \\
=& ||\hat{R}_{t_n} u||_N^2 + (1-\lambda^2)\theta^2 ||\hat{R}_{t_{n-1}} \delta V_n||_N^2 + ||\lambda \theta \hat{R}_{t_{n-1}} \delta V_n \\
&+ \frac{1}{\lambda} \hat{R}_{t_{n-1}} T_n u||_N^2 + (1 - \frac{1}{\lambda^2}) ||\hat{R}_{t_{n-1}} T_n u||_N^2 \\
\geq& \alpha \Bigg[||\hat{R}_{t_n} u||^2 + (1-\lambda^2)\theta^2 ||\hat{R}_{t_{n-1}} \delta V_n||^2 \\
&+ ||\lambda \theta \hat{R}_{t_{n-1}} \delta V_n + \frac{1}{\lambda} \hat{R}_{t_{n-1}} T_n u||^2 + (1 - \frac{1}{\lambda^2}) ||\hat{R}_{t_{n-1}} T_n u||^2 \Bigg] \\
\geq& \alpha \left[||u||^2 + (1-\lambda^2)\theta^2 ||\delta V_n||^2 + (1 - \frac{1}{\lambda^2}) ||u||^2 \right] \\
\geq& \alpha \left[(2 - \frac{1}{\lambda^2}) ||u||^2 + \theta^2 + [\frac{1-\lambda^2}{2 - \frac{1}{\lambda^2}} ||\delta V_n||^2 - 1] \theta^2 \right] \\
\geq& \alpha (2 - \frac{1}{\lambda^2}) \left(||u||^2 + \theta^2 + [\frac{1-\lambda^2}{2 - \frac{1}{\lambda^2}} v_{\min}^2 - 1] \theta^2 \right).
\end{aligned}
$$

As $\frac{1-\lambda^2}{2-\frac{1}{\lambda^2}} \underset{\lambda \to \frac{1}{\sqrt{2}}^-}{\longrightarrow} +\infty$ there exists λ_0 such that: $M \begin{pmatrix} \theta \\ u \end{pmatrix} \geq \alpha \left(2 - \frac{1}{\lambda_0^2}\right) \left|\left| \begin{matrix} \theta \\ u \end{matrix} \right|\right|^2$ and the result is true for $\beta_1 = \alpha(2 - \frac{1}{\lambda_0^2})$.

References

[1] N. Aghannan and P. Rouchon. An intrinsic observer for a class of Lagrangian systems. *IEEE Trans. on Automatic Control*, 48(6):936–945, 2003.

[2] Martin Barczyk, Silvere Bonnabel, Jean-Emmanuel Deschaud, and François Goulette. Invariant EKF design for scan matching-aided localization. *IEEE Trans. on Systems Control Technology. DOI 10.1109/TCST.2015.2413933*, 2015.

[3] Martin Barczyk and Alan F Lynch. Invariant observer design for a helicopter UAV aided inertial navigation system. *Control Systems Technology, IEEE Transactions on*, 21(3):791–806, 2013.

[4] Axel Barrau and Silvere Bonnabel. Invariant particle filtering with application to localization. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 5599–5605. IEEE, 2014.

[5] Axel Barrau and Silvere Bonnabel. Intrinsic filtering on Lie groups with applications to attitude estimation. *Automatic Control, IEEE Transactions on*, 60(2):436–449, 2015.

[6] S. Bonnabel. Left-invariant extended Kalman filter and attitude estimation. In *Decision and Control, 2007 46th IEEE Conference on*, pages 1027–1032. IEEE, 2007.

[7] S. Bonnabel, P. Martin, and P. Rouchon. Non-linear symmetry-preserving observers on Lie groups. *IEEE Transactions on Automatic Control*, 54(7):1709–1713, 2009.

[8] S. Bonnabel, Ph. Martin, and P. Rouchon. Symmetry-preserving observers. *Automatic Control, IEEE Transactions on*, 53(11):2514–2526, 2008.

[9] Silvere Bonnabel, Philippe Martin, and Erwan Salaün. Invariant extended Kalman filter: theory and application to a velocity-aided attitude estimation problem. In *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pages 1297–1304. IEEE, 2009.

[10] Silvere Bonnabel and Jean-Jacques Slotine. A contraction theory-based analysis of the stability of the deterministic extended Kalman filter. *Automatic Control, IEEE Transactions on*, 60(2):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 transactions on automatic control*, 42, 1997.

[12] Alessandro De Luca, Giuseppe Oriolo, and Claude Samson. Feedback control of a nonholonomic car-like robot. In *Robot motion planning and control*, pages 171–253. Springer, 1998.

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

[14] Jay Farrell. *Aided navigation: GPS with high rate sensors*. McGraw-Hill, Inc., 2008.

[15] J.P. Gauthier and I. Kupka. *Deterministic Observation Theory and Applications*. Cambridge University Press, 2001.

[16] Håvard Fjær Grip, Thor I Fossen, Tor A Johansen, and Ali Saberi. Globally exponentially stable attitude and gyro bias estimation with application to GNSS/INS integration. *Automatica*, 51:158–166, 2015.

[17] Minh-Duc Hua. Attitude estimation for accelerated vehicles using GPS/INS measurements. *Control Engineering Practice*, 18(7):723–732, 2010.

[18] Maziar Izadi and Amit K Sanyal. Rigid body attitude estimation based on the Lagrange–d- Alembert principle. *Automatica*, 50(10):2570–2577, 2014.

[19] Arthur J Krener. The convergence of the extended Kalman filter. In *Directions in mathematical systems theory and optimization*, pages 173–182. Springer, 2003.

[20] Ch. Lageman, J. Trumpf, and R. Mahony. Gradient-like observers for invariant dynamics on a Lie group. *Automatic Control, IEEE Transactions on*, 55(2):367–377, 2010.

[21] E. J. Lefferts, F Landis Markley, and M. D. Shuster. Kalman filtering for spacecraft attitude estimation. *Journal of Guidance, Control, and Dynamics*, 5(5):417–429, 1982.

[22] Robert Mahony, Tarek Hamel, and J-M Pflimlin. Complementary filter design on the special orthogonal group SO(3). In *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05. 44th IEEE Conference on*, pages 1477–1484. IEEE, 2005.

[23] Philippe Martin and Erwan Salaün. Generalized multiplicative extended Kalman filter for aided attitude and heading reference system. In *AIAA Guidance, Navigation, and Control Conference*, page 8300, 2010.

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

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

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

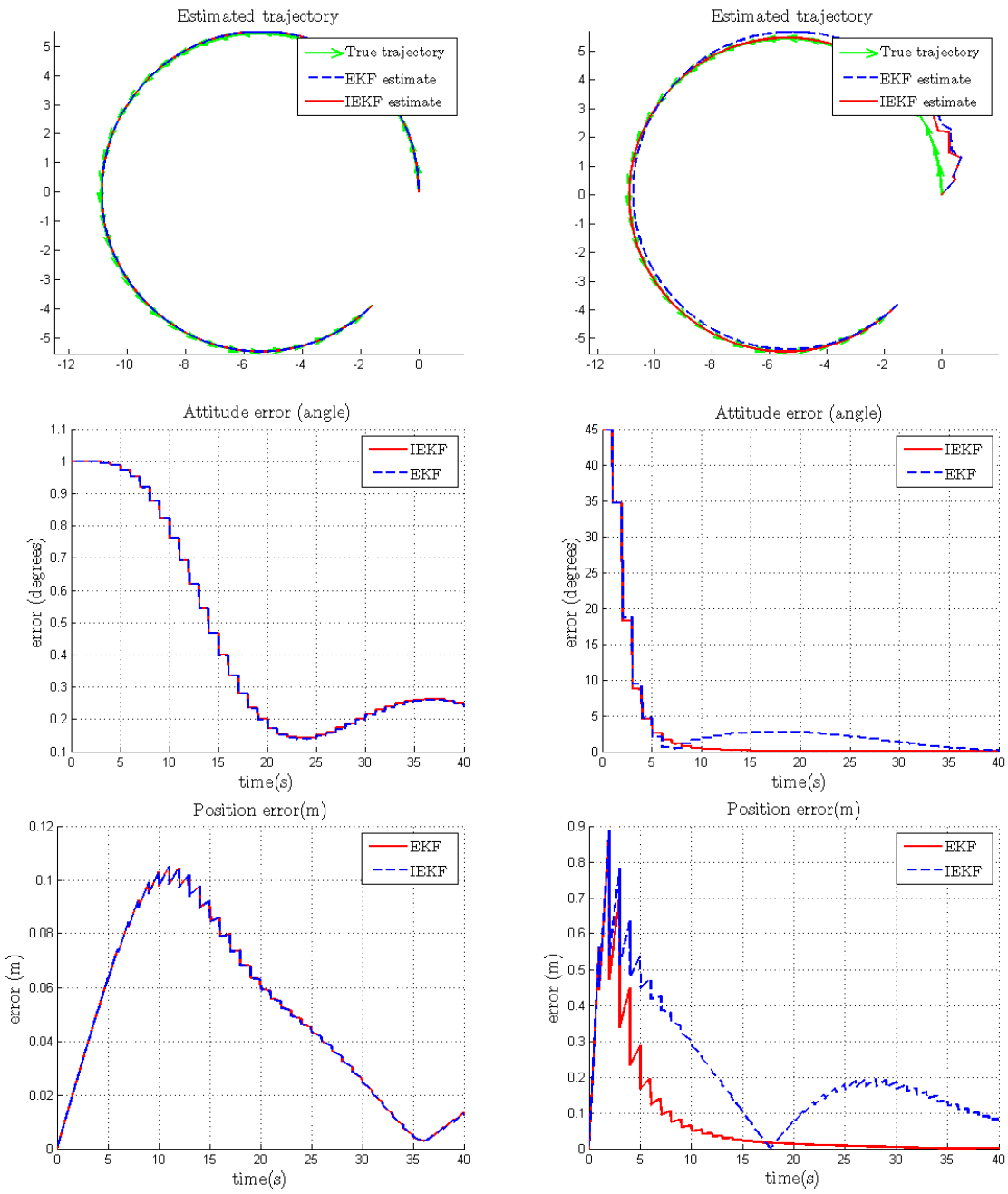

Figure 1: The heading and position of the car are estimated through EKF and IEKF with high rate odometry and low rate GPS measurements. Top plots illustrate the experimental setting and display the estimated trajectories, middle plots display the heading errors and bottom plots the position errors. As 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 non-linearities, the IEKF keeps ensuring rapid estimation error decrease.

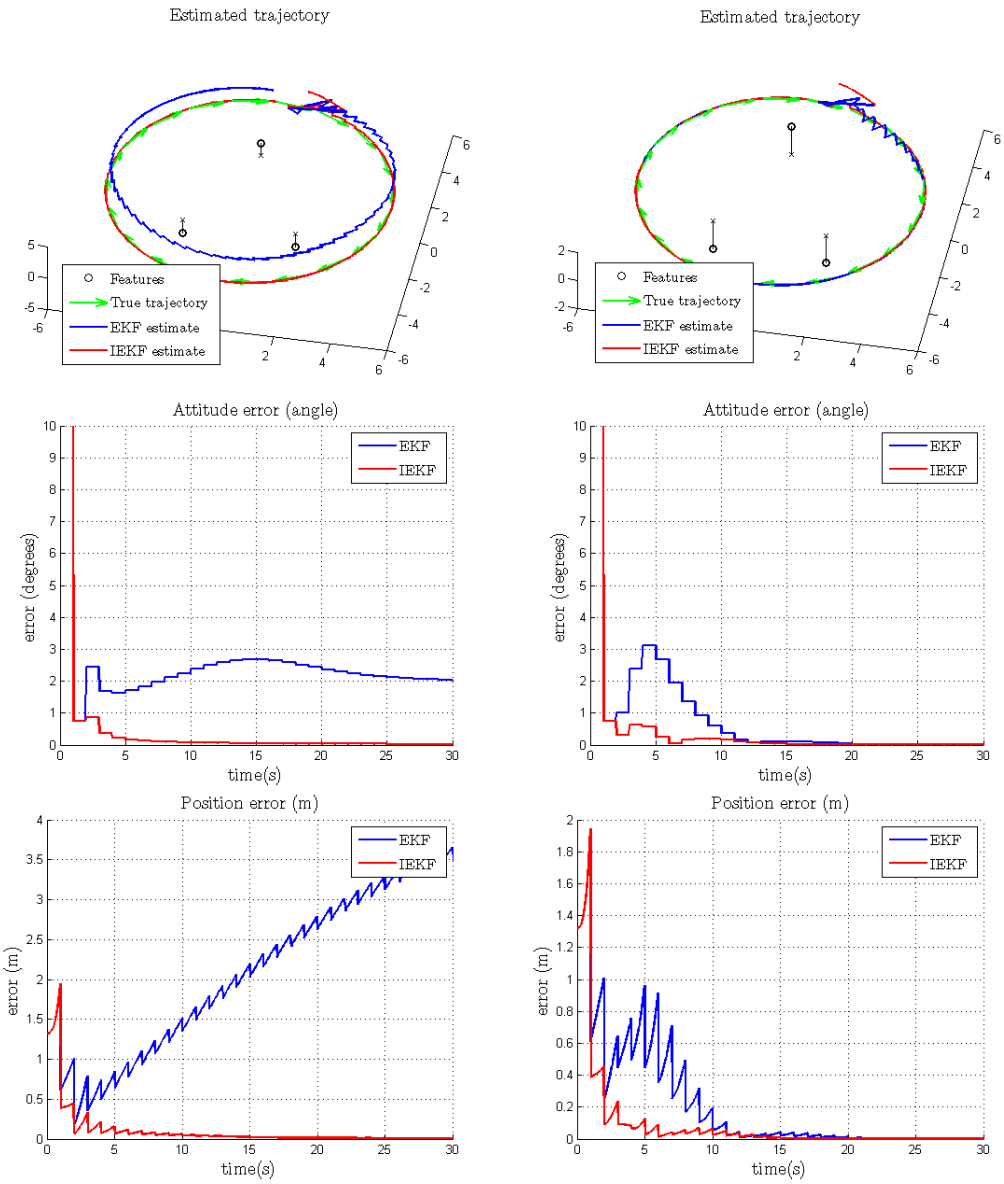

Figure 2: Aided inertial navigation based on high rate accelerometers' and gyrometers' measurements and low rate observation of known landmarks. We also displayed the orthogonal projection of the landmarks on the plane containing the trajectory (black crosses) to help imagining the 3D position of the landmarks. This shows the disposition of the landmarks 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 creates robustness issues: the gains of the EKF decrease rapidly during the transitory phase while the attitude error is not reduced enough due to non-linearities. 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.