



HAL
open science

Invariant Kalman filtering

Axel Barrau, Silvère Bonnabel

► **To cite this version:**

Axel Barrau, Silvère Bonnabel. Invariant Kalman filtering. Annual Review of Control, Robotics, and Autonomous Systems , 2018, 1 (1), pp.237-257. <10.1146/annurev-control-060117-105010>. <hal-05088150>

HAL Id: hal-05088150

<https://hal.science/hal-05088150v1>

Submitted on 28 May 2025

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.



HAL Authorization

The following document is the preprint for the paper:

Invariant Kalman filtering

Axel Barrau and Silvere Bonnabel

Annual Reviews of Control, Robotics, and Autonomous Systems,

Vol 1, pp: 237-257, 2018.

Invariant Kalman filtering

Axel Barrau¹ and Silvère Bonnabel²

Contents

1	INTRODUCTION	3
1.1	Extended Kalman filtering (EKF)	3
1.2	Motivation for the use of geometry	5
1.3	Outline	5
2	LIE GROUPS AND PROBABILITY	5
2.1	Matrix Lie Groups	5
2.2	Uncertainty representation on matrix Lie groups	7
3	INVARIANT KALMAN FILTERING	7
3.1	Group affine systems	7
3.2	Invariant EKF methodology	8
3.3	Geometric insight	10
3.4	Unscented version	11
4	THEORETICAL CONVERGENCE GUARANTEES	11
4.1	The Invariant EKF as a non-linear stable observer	11
4.1.1	The conventional EKF as a non-linear observer	12
4.1.2	The invariant EKF as a non-linear observer	12
4.2	Non-linear observers on Lie groups	13
5	INDUSTRIAL APPLICATION: DRONE NAVIGATION	13
6	APPLICATION TO SIMULTANEOUS LOCALIZATION AND MAPPING	14
6.1	EKF-based SLAM inconsistency and benefits of the IEKF	15
6.2	Experimental results	17
7	CONCLUSION	18
8	SUMMARY POINTS	19

Abstract

The Kalman filter, or more precisely the extended Kalman filter (EKF), is a fundamental engineering tool that is pervasively used in control, robotics, and for various estimation tasks in autonomous systems. The recent field of Invariant extended Kalman filtering, aims at using the geometric structure of the state space and the dynamics to improve the EKF, notably in terms of mathematical guarantees. The methodology essentially applies in the field of localization, navigation, and simultaneous localization and mapping (SLAM). Albeit recent, its remarkable robustness properties have already motivated a real industrial implementation in the aerospace field. This review aims to provide an accessible introduction to the methodology of invariant Kalman filtering, and to allow the reader to gain insight into the relevance of the method, as well as what the important differences with the conventional EKF are. This should be of interest to readers intrigued by the application of mathematical theories to practical applications, and also to readers interested in finding simple to implement and robust filters for localization, navigation, and SLAM, notably for autonomous vehicle guidance.

1 INTRODUCTION

The goal of a filter is to estimate the state of a dynamical system, by combining an evolution model and some sensor measurements that bring partial information about the state. Unfortunately, models are inherently inaccurate and sensors are subject to noises that corrupt the measurements. The idea of filtering is to explicitly include both latter sources of **uncertainty** in the model, and to compute the best estimates of the state that can be inferred from the available information. Even though the mathematical theory is now well understood, it is still a challenge to design filters in practice for control, robotics, and autonomous systems. The extended Kalman filter (EKF) appeared in the 1960s with the advent of computers, and was first implemented by NASA in the Apollo program to estimate in real time the trajectory of the space capsule. The past two decades have also witnessed the burst of particle filters, with great advances both in theory and in practice. However, they rely on extensive numerical computations that are not always suited to real time on-board implementations, and there are theoretical caveats especially when process noise is low (typically for static parameter estimation). The robotics community has also more recently turned to optimization based techniques for filtering (the problem is typically formulated as nonlinear least squares) but the computation demands are extensive and robustness to erroneous initializations is not yet clearly established. As robots and many control systems are real-time, the amount of computation is limited and the EKF is still a very widespread tool in control and robotics along with its more recent variant the unscented Kalman filter (UKF). In the aerospace industry it remains the reference filter.

1.1 Extended Kalman filtering (EKF)

Consider a general dynamical system in discrete time whose state is described by the vector variable $X_n \in \mathbb{R}^d$. We associate a sequence of observations $(Y_n)_{n \geq 0} \in \mathbb{R}^p$ which are measurement data returned by sensors. The trusted evolution model is:

$$X_n = f(X_{n-1}, u_n, w_n), \quad (1)$$

where f is the function encoding the evolution of the system, w_n is the (unknown) process noise, that is a centered random variable with covariance matrix Q_n , the vector $u_n \in \mathbb{R}^m$ a control input, and the observation consists of partial measurements of the state at time n :

$$Y_n = h(X_n) + V_n, \quad (2)$$

with h the observation function and V_n the (unknown) measurement noise that accounts for sensors' limitations.

The **extended Kalman filter** (EKF) computes in real time an **approximation $\hat{X}_{n|n}$ to the best state estimate** given the observations. Let $Y_{1:n}$ denote the collection of past measurements Y_1, Y_2, \dots, Y_n and $u_{1:n}$

be similarly defined. To be more precise, the EKF represents the belief $\mathbb{P}(X_n | u_{1:n}, Y_{1:n})$, which assigns a probability to each possible value of the true state in the light of all the information collected so far, by a mean $\hat{X}_{n|n}$ and covariance matrix $P_{n|n}$. Indeed the rationale is to use the following Gaussian approximation: $\mathbb{P}(X_n | u_{1:n}, Y_{1:n}) \approx \mathcal{N}(\hat{X}_{n|n}, P_{n|n})$. To compute the mean and covariance, the EKF uses a two step procedure.

Step 1 - Propagation: The estimate $\hat{X}_{n-1|n-1}$ obtained after the observation Y_{n-1} , is propagated through the deterministic part of (1):

$$\hat{X}_{n|n-1} = f(\hat{X}_{n-1|n-1}, u_n, 0) \quad (3)$$

To compute the associated covariance, introduce the estimation errors defined as

$$e_{n-1|n-1} = X_{n-1} - \hat{X}_{n-1|n-1}, \quad e_{n|n-1} = X_n - \hat{X}_{n|n-1}. \quad (4)$$

The key idea underlying the EKF is to linearize the error system through a first-order Taylor expansion of the non-linear functions f and h at the estimate $\hat{X}_{n-1|n-1}$. Indeed, using the Jacobians $F_n = \frac{\partial f}{\partial X}(\hat{X}_{n-1|n-1}, u_n, 0)$, $G_n = \frac{\partial f}{\partial w}(\hat{X}_{n-1|n-1}, u_n, 0)$, and $H_n = \frac{\partial h}{\partial X}(\hat{X}_{n|n-1})$, the combination of equations (1), (2) and (3) yields the following first-order expansion of the error system:

$$e_{n|n-1} = F_n e_{n-1|n-1} + G_n w_n, \quad (5)$$

$$Y_n - h(\hat{X}_{n|n-1}) = H_n e_{n|n-1} + V_n, \quad (6)$$

where the second order terms, that is, terms of order $O(\|e\|^2, \|w\|^2, \|e\|\|w\|)$ have been removed, see e.g., [52]. $P_{n-1|n-1}$ is an approximation to the true covariance $\mathbb{E}(e_{n-1|n-1} e_{n-1|n-1}^T)$, and it is propagated through the linearized model (5) so that $P_{n|n-1} = F_n P_{n-1|n-1} F_n^T + G_n Q_n G_n^T$ is an approximation of $\mathbb{E}(e_{n|n-1} e_{n|n-1}^T)$, and we have $\mathbb{P}(X_n | u_{1:n}, Y_{1:n-1}) \approx \mathcal{N}(\hat{X}_{n|n-1}, P_{n|n-1})$.

Step 2 - Measurement update: To account for the measurement Y_n , we let $z_n = Y_n - h(\hat{X}_{n|n-1})$, and z_n is called the innovation. Assuming that $e_{n|n-1} \sim \mathcal{N}(0, P_{n|n-1})$ and the approximation (6) to be exact, the linear Kalman filter equations ensure that the updated error $e_{n|n} = X_n - \hat{X}_{n|n}$ satisfies $e_{n|n} \sim \mathcal{N}(0, P_{n|n})$ where

$$\hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n z_n, \quad \text{and} \quad P_{n|n} = [I - K_n H_n] P_{n|n-1} \quad (7)$$

with K_n , called the Kalman gain, computed in Algorithm 1. Of course the belief after update $\mathcal{N}(\hat{X}_{n|n}, P_{n|n})$ is only an approximation to $\mathbb{P}(X_n | u_{1:n}, Y_{1:n})$, due to the linearizations. In practice, those linearizations may lead the filter to inconsistencies and sometimes even divergence.

Algorithm 1 Extended Kalman Filter (EKF)

Choose an initial estimate $\hat{X}_{0|0}$ and uncertainty matrix $P_{0|0}$.

loop

Define F_n, G_n and H_n through (5) and (6).

Define Q_n as $\text{Cov}(w_n)$ and R_n as $\text{Cov}(V_n)$.

Propagation

$$\hat{X}_{n|n-1} = f(\hat{X}_{n-1|n-1}, u_n, 0)$$

$$P_{n|n-1} = F_n P_{n-1|n-1} F_n^T + G_n Q_n G_n^T$$

Measurement update

$$\text{Compute } z_n = Y_n - h(\hat{X}_{n|n-1}), \quad S_n = H_n P_{n|n-1} H_n^T + R_n, \quad K_n = P_{n|n-1} H_n^T S_n^{-1}$$

$$P_{n|n} = [I - K_n H_n] P_{n|n-1}$$

$$\hat{X}_{n|n} = \hat{X}_{n|n-1} + K_n z_n$$

end loop

1.2 Motivation for the use of geometry

The user who is facing the filtering problem defined by system (1)-(2) with f and h two non-linear maps, is free to choose a different coordinate system to design an EKF. For instance, in radar tracking, one can choose a frame attached to the target, or attached to the radar, but also range and bearing as an alternative to Cartesian coordinates. In general, it is unclear what the best coordinates for EKF design are. However, for some systems, there is a choice of coordinates that is natural. For instance, when facing a linear Gaussian system, the EKF boils down to the linear Kalman filter, which is optimal, and it would be nonsense to work with alternative coordinates that would make the system non-linear. In this paper we advocate that **for a large class of systems defined on matrix Lie groups, the machinery of geometry provides coordinates that are unarguably more suited to the problem.** In those cases, the Invariant EKF theory is useful, as the original problem is often formulated using coordinates that do not match the group structure, leading to degraded performance of the conventional EKF.

For systems on Lie groups, the Invariant extended Kalman filter (IEKF) was originally introduced in [11] and continued in [15, 44, 4, 16, 7]. The complete methodology along with the convergence properties of the IEKF can be found in the recent paper [7]. More generally, the use of Lie groups for state estimation dates back to the 1970s [18, 56, 28] and has recently spanned a range of applications and a rich stream of theoretical results, see e.g. [43, 53, 10, 9, 16, 40, 41].

In the robotics community, following the work of [22], it has increasingly been recognized that using probability distributions properly defined on Lie groups is of importance, notably for pose estimation, see e.g. [5, 21, 22, 23, 58, 42]. Moreover, the use of the IEKF over the conventional EKF has recently been shown to solve the inconsistency issues of the EKF based simultaneous localization and mapping (SLAM) in [8].

1.3 Outline

Section 2 consists of geometry preliminaries. Section 3 reviews the methodology of Invariant Kalman filtering (IEKF). Section 4 is concerned with the mathematical guarantees that come with the IEKF. Section 5 presents some real industrial applications to the field of inertial navigation. Section 6 reviews the inconsistency of EKF-based SLAM and the interest of IEKF-based SLAM.

The presentation of the present article is freely inspired by the very tutorial paper [6], that uses methods rooted in differential geometry to improve Monte Carlo schemes.

2 LIE GROUPS AND PROBABILITY

2.1 Matrix Lie Groups

In this section we provide the reader with the bare minimum of Lie group theory that is required to introduce the Invariant EKF methodology¹. A matrix **Lie group** G is a subset of square invertible $N \times N$ matrices $\mathcal{M}_N(\mathbb{R})$ verifying the following properties:

$$I_N \in G, \quad \forall \chi \in G, \chi^{-1} \in G, \quad \forall \chi_1, \chi_2 \in G, \chi_1 \chi_2 \in G,$$

where I_N is the identity matrix of \mathbb{R}^N . The subset G is generally *not* a vector space, and can thus be viewed as a curved space (see Fig. 1). To every point $\chi \in G$, one can associate a *vector space* $T_\chi G$ called the tangent space at χ , and defined as all the matrices that write $\frac{d}{dt}\gamma(0)$ where $\gamma: \mathbb{R} \rightarrow G$ is a smooth curve of

¹ **General Lie groups:** All the results carry over to general abstract Lie groups, but the (less general) matrix Lie groups are well suited to tutorial and computational purposes, and encompass all the applications discussed.

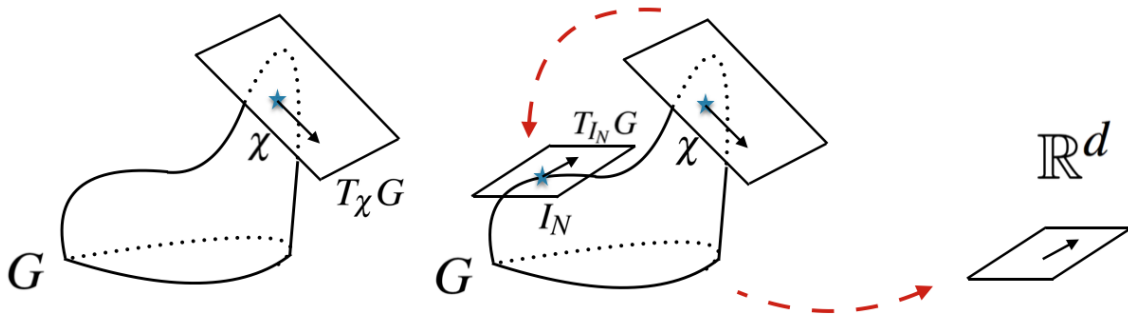


Figure 1: G is a curved space. Left and right multiplications offer two ways to identify the tangent space $T_\chi G$ at χ with the tangent space at Identity $T_{I_N} G$, called the Lie algebra \mathfrak{g} . In turn, the application $\xi \mapsto \xi^\wedge$ provides a linear bijection between the Euclidean space \mathbb{R}^d and \mathfrak{g} . (Remark: Lie groups are homogeneous spaces, which somehow “look the same everywhere”. As such, the figure may be slightly misleading, since the curved surface representing G seems irregular. Yet, it seems to us that representing a Lie group by, e.g., a sphere, would be an oversimplification.)

G that satisfies $\gamma(0) = \chi$. The elements of this space are called **tangent vectors**.

The tangent space $T_{I_N} G$ at the identity I_N is called the **Lie algebra**, and plays a very specific role. It is denoted \mathfrak{g} and its dimension d defines a dimension for the group G itself, where d is generally much smaller than the dimension N^2 of the ambient space. We have $\mathfrak{g} \subset \mathcal{M}_N(\mathbb{R})$, but there always is an invertible map $\mathbb{R}^d \rightarrow \mathfrak{g}$ that allows identifying \mathfrak{g} to \mathbb{R}^d . For $\xi \in \mathbb{R}^d$ we denote $\xi^\wedge \in \mathfrak{g}$ the corresponding element of \mathfrak{g} , and we recall that $\xi \mapsto \xi^\wedge$ is a linear map. There are then two canonical ways to identify \mathbb{R}^d and the tangent space $T_\chi G$ at any $\chi \in G$: through **left and right multiplications**, and they are generally different. Indeed for any $\xi \in \mathbb{R}^d$ the vectors $\chi(\xi^\wedge)$ and $(\xi^\wedge)\chi$ both are tangent vectors at χ . Of course they are different due to non-commutativity of matrix multiplication.

The usual matrix exponential map $\exp_m : \mathfrak{g} \rightarrow G$ constitutes a bijection from a neighborhood $V \subset \mathfrak{g}$ of 0 to a neighborhood of the identity I_N in G . In this paper, we will call the Lie **exponential map** the map $\exp : \mathbb{R}^d \rightarrow G$ defined by $\exp(\xi) = \exp_m(\xi^\wedge)$, that is a bijection in a neighborhood of $0 \in \mathbb{R}^d$, with $\exp(\xi^\wedge)^{-1} = \exp(-\xi^\wedge)$. Moreover, $\chi \exp(\cdot)$ and $\exp(\cdot)\chi$ provide two distinct bijections between a neighborhood of 0 in \mathbb{R}^d and a neighborhood of χ in G .

The **Baker-Campbell-Hausdorff (BCH)** formula gives a series expansion for the image in \mathfrak{g} of the product on G : $BCH(\xi, \zeta) = \exp^{-1}(\exp(\xi)\exp(\zeta))$. In particular it ensures $\exp(\xi)\exp(\zeta) = \exp(\xi + \zeta + T)$, where T is of the order $O(\|\xi\|^2, \|\zeta\|^2, \|\xi\|\|\zeta\|)$.

Example 1. Consider the group of **rotation matrices** $G = SO(3)$, that describes the orientation (attitude) of a body in space. It is the subset of matrices R of $\mathcal{M}_3(\mathbb{R})$ such that $RR^T = I_3$ and $\det(R) = 1$. Each $R \in G$ can be viewed as the rotation that maps vectors expressed in the body frame to vectors expressed in the earth-fixed frame. We have $d = 3$, and for any $\xi \in \mathbb{R}^3$, $\xi^\wedge \in \mathbb{R}^{3 \times 3}$ is the skew symmetric matrix associated with the “cross product” operator $\zeta \rightarrow \xi \times \zeta$. The Lie algebra is the set of skew symmetric matrices. Any tangent vector $U \in T_R G$ at R represents an infinitesimal shift of R in $SO(3)$. It can be written as $U = R\xi^\wedge$ for some vector $\xi \in \mathbb{R}^3$ or alternatively as $U = \zeta^\wedge R$ for some $\zeta \in \mathbb{R}^3$. Both ξ and ζ can be seen as angular velocity vectors that represent the same infinitesimal rotation, but ξ is a vector of the body frame whereas ζ is expressed in the earth-fixed frame. It is easy to prove indeed that $\zeta = R\xi$.

2.2 Uncertainty representation on matrix Lie groups

To define random variables on Lie groups, we cannot apply the usual approach of additive noise for $\chi \in G$ as G is not a vector space. In contrast, we define the probability distribution $\chi \sim \mathcal{N}_L(\bar{\chi}, P)$ as the probability law of the random variable $\chi \in G$ defined as

$$\chi = \bar{\chi} \exp(\xi), \quad \xi \sim \mathcal{N}(0, P), \quad (8)$$

where $\mathcal{N}(\cdot, \cdot)$ is the classical Gaussian distribution in Euclidean space \mathbb{R}^d - identified to the Lie algebra \mathfrak{g} and $P \in \mathbb{R}^{d \times d}$ is a covariance matrix. In (8), the original Gaussian vector ξ of the Lie algebra is moved over by left multiplication to be centered at $\bar{\chi} \in G$, and we similarly define the distribution $\chi \sim \mathcal{N}_R(\bar{\chi}, P)$ for right multiplication of $\bar{\chi}$ through the random variable

$$\chi = \exp(\xi) \bar{\chi}, \quad \xi \sim \mathcal{N}(0, P). \quad (9)$$

In (8) and (9), $\bar{\chi}$ is deterministic and can take any value, whereas P is the covariance of the small, noisy perturbation ξ . We stress that we have defined these probability density functions directly in the vector space \mathbb{R}^d and that both $\mathcal{N}_L(\cdot, \cdot)$ and $\mathcal{N}_R(\cdot, \cdot)$ are not Gaussian distributions. These kinds of distributions introduced to our best knowledge in [58, 23, 22] were advocated and studied in [5] for pose estimate and leveraged for Kalman filtering in [7, 16, 19]. They are sometimes referred to as concentrated Gaussians on Lie groups. Note that, χ in (8) and (9) is a well defined random variable of G . Computing its distribution on G is not trivial though, but for the purposes we pursue in invariant filtering it is absolutely not necessary.

Several approaches to filtering for systems possessing a geometric structure have been developed in previous literature. For stochastic processes on Riemannian manifolds [36] some results have been derived, see e.g., [46]. The specific situation where the process evolves in a vector space but the observations belong to a manifold has also been considered, see e.g. [28, 48] and more recently [49]. For systems on Lie groups powerful tools to study the filtering equations - such as harmonic analysis [57, 47, 55] - have been used, notably in the case of bilinear systems [56] and estimation of the initial condition of a Brownian motion in [27]. A somewhat different but related approach to filtering consists of finding the path that best fits the data in a deterministic setting. It is thus related to optimal control theory where geometric methods have long played an important role [18]. A certain class of least squares problems on the Euclidean group has been tackled in [32], see also [59].

3 INVARIANT KALMAN FILTERING

In invariant Kalman filtering, the state space is assumed to be a matrix Lie group. As a result it comes with the non-linear machinery of Subsection 2.1. However, we emphasize that it is the *coupling* of the choice of the Lie group structure and the dynamical model used in (1) that allows derivation of properties of the filter. Originally only dynamics that were invariant to the group action were considered - hence the name, the IEKF is rooted in [14, 12] - and the class of dynamics to be considered has recently been much extended, as presented in Subsection 3.1. The remainder of the section is as follows. Subsection 3.2 is devoted to the methodology of the IEKF, Subsection 3.3 to the geometric vision that underlies it, and Subsection 3.4 to its recent unscented version.

3.1 Group affine systems

This section reviews some results of [7], that are derived in the continuous time context, and we present here their (novel) discrete-time counterpart. Consider a *deterministic* dynamical system on a matrix Lie group G :

$$\chi_n = f(\chi_{n-1}, u_n) \quad (10)$$

with f a smooth map on the group, $u_n \in \mathbb{R}^m$, or possibly $u_n \in G$ (in which case we use a capital U_n). In invariant Kalman filtering, function f is required to verify

$$\forall \chi, v \in G, \forall u \quad f(\chi v, u) = f(\chi, u) f(I_N, u)^{-1} f(v, u). \quad (11)$$

Systems that satisfy this condition are called **group affine systems** in [7]. Indeed if the group is merely a vector space with addition as the group composition law, we recover the affine functions, as (11) then becomes $f(a+b, u) = f(a, u) - f(0, u) + f(b, u)$ which implies $f(x, u)$ is necessarily affine if it is smooth. Note that, the continuous version of condition (11) introduced in paper [7] is similar to the notion of linearity on a group defined in a different context in [3].

Theorem 1 (Fundamental property of invariant filtering). *We have the equivalence:*

$$f \text{ satisfies (11)} \Leftrightarrow \text{There exists a map } g \text{ such that } \forall \chi, v, u, f(v, u)^{-1} f(\chi, u) = g(v^{-1} \chi, u)$$

Moreover, in this case, for each $u \in \mathbb{R}^m$ there exists $F \in \mathbb{R}^{d \times d}$ such that $\forall \xi \in \mathbb{R}^d, g(\exp(\xi), u) = \exp(F\xi)$, that is, the function $g(\cdot, u)$ is wholly encoded in a simple matrix F .

This result is pivotal as will be explained in the sequel. Note that, the fact that $g(\cdot, u)$ may be described by a mere matrix F is not a first order approximation. It is true at all orders, and this is a quite surprising result.

Example 2. *Let us continue our simple example of rotation matrices. Consider a motion R_0, R_1, \dots on $SO(3)$ with $R_n = R_{n-1} U_n$. The matrix $U_n \in SO(3)$ represents the relative rotation undergone by a body in space between time steps $n-1$ and n , and can be measured by (flawless) gyroscopes: it can thus be viewed as an input. The dynamics can be cast into the form (10) letting $f(R, U) = RU$. Thus $f(\chi v, U) = \chi v U = f(\chi, U) f(I_3, U)^{-1} f(v, U)$ so condition (11) is satisfied. And we have:*

$$f(v, U)^{-1} f(\chi, u) = (vU)^{-1} \chi U = U^{-1} v^{-1} \chi U = U^{-1} (v^{-1} \chi) U.$$

We see that the last term is a function of $v^{-1} \chi$ and U indeed, as predicted by the theorem! This has been long known in the $SO(3)$ case, and was heavily exploited for attitude observer design, see e.g. [43, 53, 12, 10, 9, 16, 40, 41]. Moreover, it is well known from Lie group theory, see e.g. [5], that for any $\xi \in \mathbb{R}^d$, we have $U^{-1} \exp(\xi) U = \exp(\text{Ad}_U \xi)$ where $\text{Ad}_U \xi = U \xi$ in the $SO(3)$ case. Letting $F = U$, we see this agrees again with the second part of Theorem 1.

3.2 Invariant EKF methodology

This section is a summary of the IEKF methodology of [7]. Consider a general dynamical system $\chi_n \in G$ associated to a sequence of observations $(Y_n)_{n \geq 0} \in \mathbb{R}^p$ as follows :

$$\chi_n = f(\chi_{n-1}, u_n) \exp(w_n) \quad (12)$$

$$Y_n = h(\chi_n) + V_n, \quad (13)$$

where $u_n \in \mathbb{R}^d$ is a control input, $w_n \in \mathbb{R}^d$ is a (unknown) vector encoding the process noise, $V_n \in \mathbb{R}^p$ is the (unknown) measurement noise, and where f satisfies condition (11). The (left) **invariant EKF** (IEKF)² computes in real time an approximation to the posterior using the uncertainty representation (8), that is, it computes at each step two parameters $\hat{\chi}_{n|n} \in G$ and $P_{n|n} \in \mathbb{R}^{d \times d}$, and makes the approximation $\mathbb{P}(\chi_n | u_{1:n}, Y_{1:n}) \approx \mathcal{NL}(\hat{\chi}_{n|n}, P_{n|n})$.

² **Left versus right:** For simplicity of exposure the IEKF theory is presented using left-invariant error $\eta = \chi^{-1} \hat{\chi}$ and uncertainty representation (8). Swapping the role of left and right multiplications we can similarly define an alternative filter, the Right-Invariant EKF.

Step 1 - Propagation: The IEKF propagates an estimate obtained after the previous observation Y_{n-1} through the deterministic part of (12), i.e., by setting $w_n = 0$:

$$\hat{\chi}_{n|n-1} = f(\hat{\chi}_{n-1|n-1}, u_n) \quad (14)$$

To compute the associated covariance, introduce the left-invariant estimation errors defined as

$$\eta_{n-1|n-1} := \chi_{n-1}^{-1} \hat{\chi}_{n-1|n-1}, \quad \eta_{n|n-1} := \chi_n^{-1} \hat{\chi}_{n|n-1}. \quad (15)$$

This error is indeed invariant to left multiplications $(\chi, \hat{\chi}) \mapsto (\Gamma\chi, \Gamma\hat{\chi})$ for any group element $\Gamma \in G$, and right-invariant errors can be analogously defined. This is a natural way of measuring the discrepancy between the true state χ and the estimate $\hat{\chi}$ in a Lie group context where the usual linear error $\hat{\chi} - \chi$ is not even an element of G . As f satisfies (11), Theorem 1 yields:

$$\eta_{n|n-1} = [f(\chi_{n-1}, u_n) \exp(w_n)]^{-1} f(\hat{\chi}_{n-1|n-1}, u_n) = \exp(-w_n) g(\eta_{n-1|n-1}, u_n) \quad (16)$$

where we have also used that $\exp(w_n)^{-1} = \exp(-w_n)$. Along the lines of the EKF, we want to **linearize the error system** through a first-order Taylor expansion of the non-linear functions f and h at the estimate $\hat{\chi}$. However, in contrast to (4), the estimation errors (15) are elements of G , that is, square matrices, and not vectors. But when χ and $\hat{\chi}$ are close, the invariant error $\eta = \chi^{-1} \hat{\chi}$ is close to the identity matrix I_N . Using that the Lie exponential map provides a bijection between a neighborhood of \mathbb{R}^d and a neighborhood of I_N , the estimation error can be locally approximated by an element of \mathbb{R}^d , that is, we let $\xi_{n-1|n-1}, \xi_{n|n-1} \in \mathbb{R}^d$ be defined by

$$\eta_{n-1|n-1} = \exp(\xi_{n-1|n-1}), \quad \eta_{n|n-1} = \exp(\xi_{n|n-1}). \quad (17)$$

Let $F_n \in \mathbb{R}^{d \times d}$ be defined by $\forall \xi \in \mathbb{R}^d, g(\exp(\xi), u_n) = \exp(F_n \xi)$. This matrix exists thanks to Theorem 1, and it is not defined as through first-order approximation, as in the conventional EKF case. According to the BCH formula we have $\exp(-w_n) g(\exp(\xi_{n-1|n-1}), u_n) = \exp(F_n \xi_{n-1|n-1} - w_n)$ up to terms of order $\|\xi_{n-1|n-1}\|^2, \|w_n\|^2, \|\xi_{n-1|n-1}\| \|w_n\|$. Recalling (16) and (17), this means that up to the first order $\exp(\xi_{n|n-1}) = \exp(F_n \xi_{n-1|n-1} - w_n)$ so we get the following linearized equation, which is the Lie group counterpart of (5)

$$\xi_{n|n-1} = F_n \xi_{n-1|n-1} - w_n. \quad (18)$$

Remark 1. *Despite the fact that the state belongs to a non-linear space, and is a matrix and not a vector, the Lie exponential map allowed us to linearize the error system in \mathbb{R}^d as in the conventional EKF methodology. Moreover, note that F_n depends on u_n but not on $\hat{\chi}_{n-1|n-1}$. This is in contrast with the conventional EKF where F_n in (5) depends on $\hat{X}_{n-1|n-1}$. This is a consequence of requirement (11) and Thm. 1, and it will play an important role in the sequel.*

We have $\exp(\xi_{n|n-1}) = \eta_{n|n-1} = \chi_n^{-1} \hat{\chi}_{n|n-1}$, which implies that $\chi_n = \hat{\chi}_{n|n-1} \exp(-\xi_{n|n-1})$. Resorting to the uncertainty representation (8), and using (18), we have just proved that if $\mathbb{P}(\chi_{n-1} | u_{1:n-1}, Y_{1:n-1}) \approx \mathcal{N}_L(\hat{\chi}_{n-1|n-1}, P_{n-1|n-1})$ then we have approximately the propagated distribution $\mathbb{P}(\chi_n | u_{1:n}, Y_{1:n-1}) \approx \mathcal{N}_L(\hat{\chi}_{n|n-1}, P_{n|n-1})$ where $P_{n|n-1} = F_n P_{n-1|n-1} F_n^T + Q_n$, and where $Q_n = \text{Cov}(w_n) = \text{Cov}(-w_n)$.

Step 2 - Measurement update To account for the new measurement we let $z_n = Y_n - h(\hat{\chi}_{n|n-1})$ be the innovation. It is a vector of \mathbb{R}^p . We have $z_n = h(\chi_n) - h(\hat{\chi}_{n|n-1}) + V_n = h(\hat{\chi}_{n|n-1} \exp(\xi_{n|n-1})) - h(\hat{\chi}_{n|n-1}) + V_n$. As $\xi_{n|n-1}$ is assumed small, and as $\exp(0) = I_N$, a first-order Taylor expansion in $\xi \in \mathbb{R}^d$ arbitrary, allows defining H_n as follows

$$h(\hat{\chi}_{n|n-1} \exp(\xi)) - h(\hat{\chi}_{n|n-1}) := H_n \xi + O(\|\xi\|^2) \quad (19)$$

Now that we have obtained a linearized system in \mathbb{R}^d akin to (5)-(6), the conventional Kalman theory can be applied to derive the Kalman gain K_n , and the updated covariance matrix $P_{n|n}$. The term $K_n z_n$ is a corrective shift computed on the linearized system. Thus it should act on the linearized error $\xi_{n|n-1}$. As our estimation errors on the group are of the form $\exp(\xi) = \chi^{-1} \hat{\chi}$, that is, $\chi = \hat{\chi} \exp(\xi)$, an approximation to the best estimate of χ_n after observation Y_n which is consistent with (15), is obtained through the Lie group counterpart $\hat{\chi}_{n|n} = \hat{\chi}_{n|n-1} \exp(K_n z_n)$ of the linear update (7). The equations of the filter are detailed in Algorithm 2.

Algorithm 2 Invariant Extended Kalman Filter (IEKF)

Choose initial $\hat{\chi}_{0|0} \in G$ and $P_{0|0} \in \mathbb{R}^{d \times d} = \text{Cov}(\xi_{0|0})$

loop

Define H_n as in (19), and F_n through the expansion $g(\exp(\xi), u_n) = \exp(F_n \xi)$.
Define Q_n as $\text{Cov}(w_n)$ and R_n as $\text{Cov}(V_n)$.

Propagation

$$\hat{\chi}_{n|n-1} = f(\hat{\chi}_{n-1|n-1}, u_n)$$

$$P_{n|n-1} = F_n P_{n-1|n-1} F_n^{-1} + Q_n$$

Measurement update

$$\text{Compute } z_n = Y_n - h(\hat{\chi}_{n|n-1}), \quad S_n = H_n P_{n|n-1} H_n^T + R_n, \quad K_n = P_{n|n-1} H_n^T S_n^{-1}$$

$$P_{n|n} = [I - K_n H_n] P_{n|n-1}$$

$$\hat{\chi}_{n|n} = \hat{\chi}_{n|n-1} \exp(K_n z_n)$$

end loop

Remark 2. Many systems of interest are such that the observation (13) is of the form $Y_n = \chi_n(b + V_n)$ where $b \in \mathbb{R}^N$ is a known vector. Indeed, this means that one measures some combinations of entries of the matrix entries, i.e., partial measurements of the state. In this case, we use an alternative innovation $z_n = \hat{\chi}_{n|n-1}^{-1} Y_n - b$, which is equal to $\eta_{n|n-1}^{-1} b - b + V_n$. Then, since $\eta_{n|n-1}^{-1} b = \exp(-\xi_{n|n-1}) b$, the linearized output H_n corresponding to innovation z_n is defined similarly to (19), through the first-order expansion $z_n = \exp(-\xi)^{-1} b - b := H_n \xi + O(\|\xi\|^2)$. We then see that, along the lines of Remark 1, the matrix H_n does not depend on $\hat{\chi}_{n|n-1}$ either! Note that, the choice of the alternative innovation above is mostly tutorial though: a rigorous application of Algorithm 2 to the present case would in fact lead to entirely identical estimates.

3.3 Geometric insight

Table 1 compares the main features of both EKF and IEKF methodologies. The IEKF features admit a geometric interpretation illustrated by Figure 2.

Table 1: Differences between conventional EKF and Invariant EKF

Name	EKF Expression	EKF Nature	IEKF Expression	IEKF Nature
State	X_n	vector of \mathbb{R}^d	$\hat{\chi}_n$	matrix of $G \subset \mathbb{R}^{N \times N}$
Uncertainty repres.	$\mathcal{N}(\hat{X}, P)$	Gaussian in \mathbb{R}^d	$\mathcal{N}_L(\hat{\chi}, P)$	random matrix, see (8)
Linearized error	$e_{n n}$	vector of \mathbb{R}^d	$\xi_{n n}$	vector of \mathbb{R}^d
Non-linear error	$e_{n n}$	vector of \mathbb{R}^d	$\eta_{n n} = \exp(\xi_{n n})$	matrix of $G \subset \mathbb{R}^{N \times N}$
Covariance matrix	P	matrix of $\mathbb{R}^{d \times d}$	P	matrix of $\mathbb{R}^{d \times d}$
Correction term	$K_n z_n$	shift in \mathbb{R}^d	$K_n z_n$	shift in \mathbb{R}^d

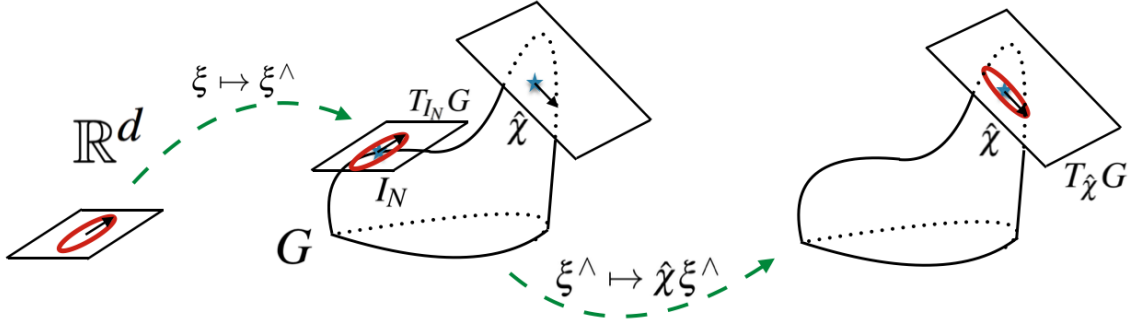


Figure 2: In the IEKF methodology the covariance matrix P represents a dispersion in \mathbb{R}^d (illustrated by the red 99% confidence ellipsoid on the left plot). In turn, using the bijection $\xi \mapsto \xi^\wedge$, it corresponds to a dispersion in the Lie algebra $T_{I_N}G = \mathfrak{g}$ (central plot). By adopting the uncertainty representation (8) we implicitly posit that the left-invariant error $\chi^{-1}\hat{\chi}$ is the exponential of a Gaussian. This means we implicitly use left multiplication to move the ellipsoid over from $T_{I_N}G$ to $T_{\hat{\chi}_{n|n}}G$ (right plot). Thus, the choice of a particular non-linear estimation error rules the way confidence ellipsoids “turn” when moved over in G . Using (9) instead, that is right-invariant errors (and thus right multiplication $\xi \mapsto \xi^\wedge \hat{\chi}_{n|n}$), or more generally the original coordinates in which the system is modeled, we would get a quite different ellipsoid in $T_{\hat{\chi}_{n|n}}G$!

3.4 Unscented version

When applying Algorithm 2 to a particular system, matrices F_n, H_n must be computed, which might prove a little difficult to some practitioners. This is one of the main reasons why the unscented extension of the Kalman filter (UKF) of [37] has become a very popular alternative to the EKF. In [19] an **unscented version of the Invariant EKF** was proposed, which spares the practitioner a computation of the Jacobians for systems modeled using discrete time dynamics.

There have been other attempts to use (partially) the Lie group structure for UKF design, see [25, 24] for symmetry-preserving observer design based on the unscented transform, [42] which uses the Lie group structure of $SE(3)$ for a drone navigation application, and [30] which exploits related ideas.

4 THEORETICAL CONVERGENCE GUARANTEES

Over the past decade there has been large body of literature devoted to *deterministic* non-linear observer design on Lie groups in the control community. We partially review this field in Subsection 4.2. But prior to that, we review the convergence properties of the invariant EKF (IEKF), when used as a candidate non-linear observer.

4.1 The Invariant EKF as a non-linear stable observer

Consider the model (1)-(2) and turn the noise off. This yields a *deterministic system* of the form $X_n = f(X_{n-1}, u_n), Y_n = h(X_n)$. A stable non-linear observer is a dynamical system of the form $\hat{X}_n = \hat{f}(\hat{X}_{n-1}, u_n, Y_n)$ where \hat{f} is designed to achieve asymptotic convergence of the estimation error, that is, $\hat{X}_n - X_n \rightarrow 0$ when $n \rightarrow \infty$. Although no noise perturbs the system, the state X_n is assumed unknown and the system’s equations only yield partial information about it. As a result, designing stable non-linear observers is often a great challenge.

4.1.1 The conventional EKF as a non-linear observer

When facing a non-linear observer problem, any EKF can readily be used as a candidate asymptotic observer, by choosing arbitrary matrices Q_n, R_n , that are then viewed as tuning parameters. Unfortunately, there are not many guarantees that the EKF when used as an observer will asymptotically converge. The main results [17, 51] rely on strong assumptions about the filter's behavior, and the EKF can actually fail to converge sometimes, even for small initial estimation errors. To understand why, consider the block diagram of Figure 3 that illustrates the architecture of the EKF (Algorithm 1).

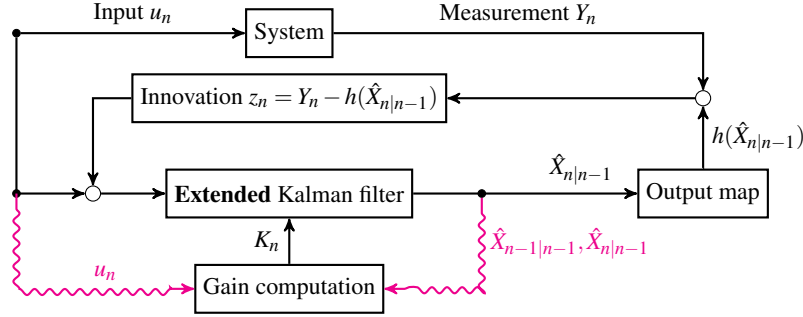


Figure 3: Architecture of the EKF (Algorithm 1). The gain depends on the estimated state.

The important point is that the computation of the gain K_n relies on the matrices F_n, H_n introduced at (5)-(6), since the system is linearized at the estimate. This creates a loop between the estimate and the gain computation that can destabilize the filter. Indeed if the estimate is not sufficiently close to the true state, the gain will be erroneous, and the correction applied by the filter (7) will amplify the estimation error. This positive feedback may lead to divergence indeed.

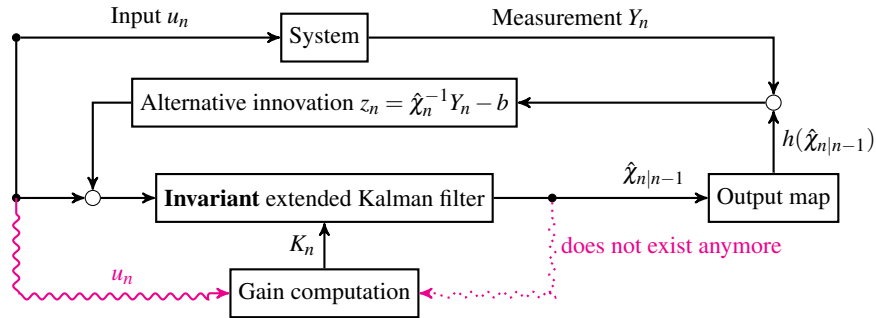


Figure 4: Architecture of the IEKF (Algorithm 2): the gain does *not* depend on the estimates.

4.1.2 The invariant EKF as a non-linear observer

The IEKF does not suffer from this drawback. Indeed, consider the system (12) with noise turned off, that is, $\chi_n = f(\chi_{n-1}, u_n)$ and assume f satisfies condition (11). As in Remark 2, *assume also* that the observation (with noise turned off) writes $Y_n = h(\chi_n) = \chi_n b$ with $b \in \mathbb{R}^N$ a known vector. As emphasized by Remarks 1 and 2, neither F_n nor H_n then depend on the estimates. This makes the gain computation independent from



Figure 5: Euroflir 410 (left), last generation of ultra-long-range electro-optical system commercialized by Safran. Its navigation system includes the first commercial implementation of an Invariant EKF. It is embedded in particular in the drone Patroller (right).

the estimate, and spares the filter potentially harmful positive feedback. Indeed, as illustrated by the block diagram of Figure 4, the inner loop of the EKF has disappeared.³

The independence of the gain computation to the trajectory means that if the system were linearized at the (unknown) *true* state, the gain K_n would be identical to the one obtained linearizing at the estimate! Thanks to this strong property, stability of the IEKF as an observer could be proved in [7]. Although the guaranteed convergence properties are only local, this is in sharp contrast with the conventional EKF where no such guarantees exist.

4.2 Non-linear observers on Lie groups

There has been a huge body of research devoted to nonlinear observers on Lie groups over the past decade, especially for attitude estimation. Although the literature is too broad to be covered here, we can point to a few significant references. Early developments on estimation on Lie groups were covered in Section 2.2. In the early 2000s, there are essentially two streams of research that have bolstered the developments of observers on Lie groups. The first one was initiated by Aghannan and Rouchon [2] and seeks to design non-linear observers that share the symmetries of the original system. The theory was formalized and developed in [14], and applied to estimation on Lie groups in [12]. At the same time, the complementary filter on $SO(3)$ for attitude estimation was introduced in [43]. It makes extensive use of left-invariant errors on $SO(3)$ and of the autonomy properties of the error equation. Due to its simplicity and its *global* convergence guarantees, it has become a very renowned attitude estimator, and has proved useful for quadrotor unmanned aerial vehicles (UAVs) control. Observers that are akin to the complementary filter on $SO(3)$ for attitude estimation include [53, 10, 9, 16, 40, 41], and similar ideas have been applied to pose estimation, e.g. in [50]. The idea to use similar techniques for noisy (instead of deterministic) systems on Lie groups and the subsequent theory of invariant Kalman filtering is slightly more recent, see e.g. [11, 44, 4, 25, 16, 7].

5 INDUSTRIAL APPLICATION: DRONE NAVIGATION

Hybrid inertial navigation has been amongst the first applications of the EKF in the 1960s, and a driver for scientific breakthroughs and reformulations that led from the original work of Kalman to a widespread industrial tool, see e.g., [45]. The company Safran Electronics & Defense (formerly known as Sagem), which

³ **Technical note:** In the general IEKF theory of [7], there are in fact situations where the gain may partially depend on the trajectory, but in a way that is not harmful.

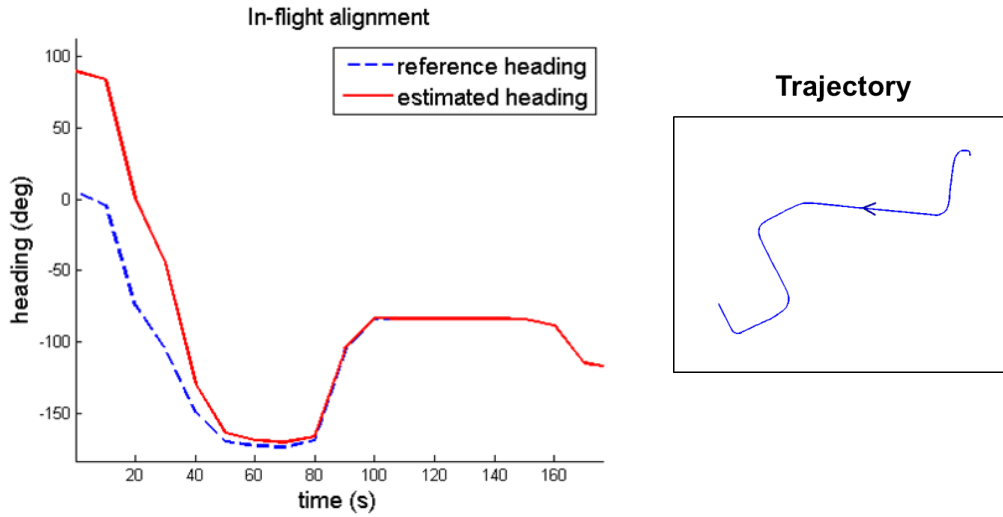


Figure 6: Heading estimated by an Invariant EKF based inertial navigation system started in-flight with extremely large initial error (90 degrees). Due to the the properties highlighted in Section 4, convergence is not perturbed despite a large initial error and a complicated trajectory.

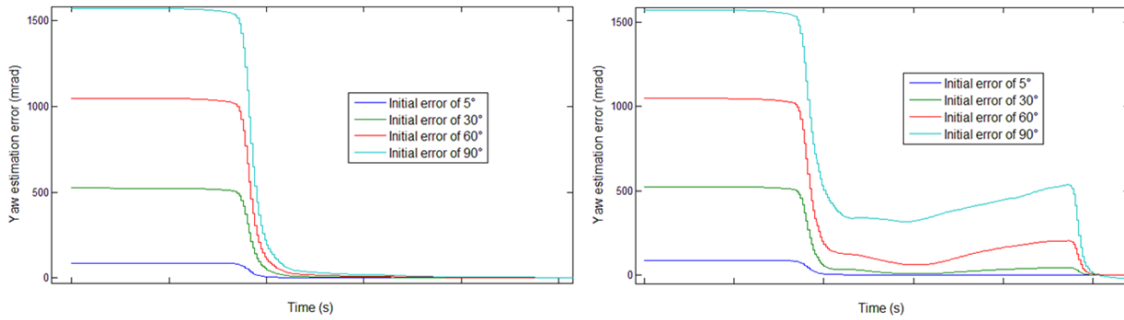


Figure 7: Heading error (milliradians) for invariant EKF (left) and conventional EKF (right) for GPS-INS hybrid navigation. The "small angles" hypothesis is crucial to EKF, but not to IEKF.

is n° 1 in Europe for inertial navigation systems (INS), chose to invest in invariant filtering. The Euroflir 410 (Fig. 5) is the first commercial product implementing invariant Kalman filtering. It is notably embedded in the recent drone Patroller, which is a long-endurance UAV used for military purposes (the French army has recently purchased fourteen Patrollers) but also for various tasks such as detection of forest fires.

Fig. 6 is obtained on *real* in-flight experimental data for hybrid INS-GPS navigation. The initial heading error has been set deliberately to an extreme value (90°) to obtain experimental confirmation that the convergence is not affected, and this whatever the actual trajectory followed (a fact which agrees with the insights of Section 4 indeed). The situation is wholly different when using a conventional EKF as shown by Fig. 7, also obtained on real experimental data. We see conventional and invariant EKF share a similar behavior for small angles (around 5°) indeed, but EKF rapidly deteriorates for larger angles while IEKF keeps converging almost as fast.

Remark 3. Commercial navigation systems such as the one equipping the Euroflir 410 have to model a very high-dimensional state which includes sensor bias, scale factors and several other uncertainty sources,

so that the state space is not a Lie group and the dynamics not strictly group affine. However, the IEKF methodology can easily be adapted and employed, by appending the other variables to the state, and treating them along the lines of the EKF methodology. Its success stems from the fact that navigation systems are clearly much “closer” to some group-affine system indeed, than to a linear system, as assumed by the conventional EKF methodology.

6 APPLICATION TO SIMULTANEOUS LOCALIZATION AND MAPPING

The problem of simultaneous localization and mapping (SLAM) has a rich history in robotics and autonomous navigation over the past two decades, see e.g. [26, 29]. It can be posed as a filtering problem, and the extended Kalman filter (EKF) based SLAM (the EKF-SLAM) one was of the first algorithm to be used in this field. It was mostly abandoned though, due to inconsistencies of its estimates. However, it was recently proved in [8] that invariant Kalman filtering resolves the inconsistency issues of the EKF.

The Invariant EKF-SLAM algorithm was introduced and studied in [8] and preliminary ideas date back to [13]. In [61] some complementary properties of the Invariant EKF-SLAM were derived, and it was successfully applied to visual-inertial SLAM in [60].

6.1 EKF-based SLAM inconsistency and benefits of the IEKF

The issue of EKF-SLAM inconsistency has been the object of many papers, see [38, 20, 35] to cite a few, where empirical evidence and theoretical explanations in various particular situations have been accumulated. In this context, inconsistency refers to the inability of the filter’s output covariance matrix $P_{n|n}$ to correctly reflect the true error dispersion $\mathbb{E}(X_n - \hat{X}_{n|n})(X_n - \hat{X}_{n|n})^T$. In particular, it was proved the orientation uncertainty is a key feature in the inconsistency, and this derives from the linearization process. A little later, [33, 34] have provided a complementary sound theoretical analysis of the EKF-SLAM inconsistency insisting on the EKF’s inability to correctly reflect the unobservabilities inherent to the SLAM problem.

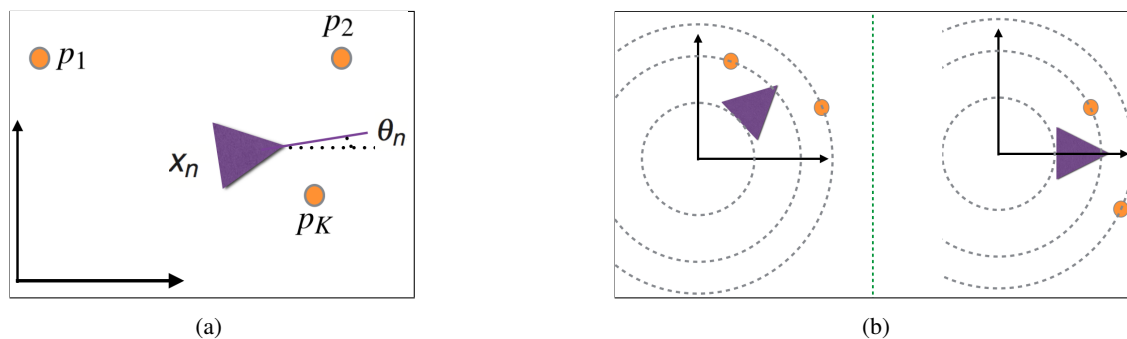


Figure 8: **(a)** In the SLAM problem a robot (triangle) is moving and makes measurements relative to some landmarks. The goal is to simultaneously estimate the robot’s trajectory and the position of the landmarks, i.e., the map. **(b)** The problem is unobservable as a global rotation of the robot and the landmarks is impossible to observe: the robot only makes relative measurements of landmarks. Thus it can’t distinguish between the left and the right configuration of picture (b).

Although all the results below carry over to 3D SLAM, we focus on the 2D case for simplicity. In the SLAM problem in 2D, a robot moves and observes some fixed features (or landmarks) of the environment. At time n , let $x_n \in \mathbb{R}^2$ be the position of the robot, θ_n its orientation, with respect to a global frame,

and let $p_1, p_2, \dots, p_K \in \mathbb{R}^2$ be the positions of the K landmarks in the global frame, see Fig. 8 (a). The state is the vector $X_n = (x_n, \theta_n, p_1, \dots, p_K)$ and the goal is to estimate it from observations of landmarks relative to the robot's frame through sensors (Lidars, cameras) attached to the robot. The collection of landmarks p_1, p_2, \dots, p_K constitutes the map, hence the name SLAM. Let $R(\alpha)$ denote a planar rotation of angle α and consider the transformation on the state space $\Psi_\alpha : (x_n, \theta_n, p_1, \dots, p_K) \mapsto (R(\alpha)x_n, \theta_n + \alpha, R(\alpha)p_1, \dots, R(\alpha)p_K)$. As simply illustrated on Fig. 8 (b), the state $X_n = (x_n, \theta_n, p_1, \dots, p_K)$ and the state $\Psi_\alpha(X_n)$ are undistinguishable, that is, departing from either one or the other configuration, whatever the motions of the robot, it is impossible for it to tell that its own position, orientation, and the landmarks' positions are actually different in both cases.

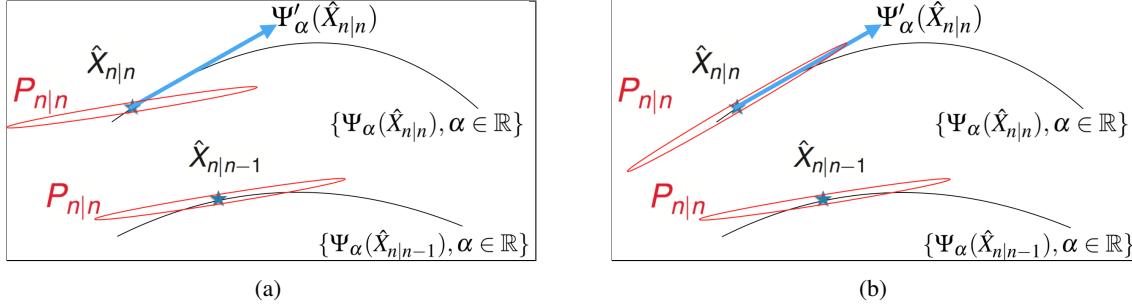


Figure 9: Difference between the conventional EKF and invariant EKF update steps with respect to unobservable directions. (a) **EKF-SLAM**: $\Psi_\alpha(\hat{X}_{n|n-1})$ and $\hat{X}_{n|n-1}$ correspond to the two undistinguishable configurations of Fig. 8 (b), so $\{\Psi_\alpha(\hat{X}_{n|n-1}), \alpha \in \mathbb{R}\}$ defines a continuous curve of states undistinguishable from $\hat{X}_{n|n-1}$ in the state space. In terms of linearized model, this means $\Psi'_\alpha(X_n) := \frac{d}{d\alpha} \Psi_\alpha(X_n)$ defines a direction in the state space at X_n along which no information can ever be gained through measurement data. The problem with the EKF is as follows. The covariance $P_{n|n}$ is computed with local information (that is, the linearized model) at the propagated step $\hat{X}_{n|n-1}$. Even if the updated belief correctly reflects the unobservability there, that is, the confidence ellipsoid defined by $P_{n|n}$ is very elongated in the unobservable direction $\Psi'_\alpha(\hat{X}_{n|n-1})$, it is then moved over to updated state $\hat{X}_{n|n}$ through a simple translation. But $\Psi'_\alpha(\hat{X}_{n|n-1}) \neq \Psi'_\alpha(\hat{X}_{n|n})$ which means the updated belief $\mathcal{N}(X_{n|n}, P_{n|n})$ is elongated in the direction $\Psi'_\alpha(\hat{X}_{n|n-1})$ that does not match with the actual unobservable direction $\Psi'_\alpha(\hat{X}_{n|n})$ at $\hat{X}_{n|n}$. (b) **IEKF-SLAM**: In the coordinates induced by the Lie group structure of the state space, the unobservable directions $\Psi'_\alpha(\hat{X})$ become independent of \hat{X} ! The effect, back in the original variables, is that confidence ellipsoids “turn” in a way that matches with unobservable directions.

The main source of inconsistency of the EKF-SLAM derives from this unobservability, as illustrated by Figure 9. The linearity of the update step (7) in the original variables combined with the uncertainty representation $\mathcal{N}(\hat{X}_{n|n}, P_{n|n-1})$ results in a mere translation of the confidence ellipsoid at the update step, that does not match with unobservabilities of the system. In contrast, the SLAM state space can be endowed with a Lie group structure as first noticed in [13]. As proved in [8] this defines local coordinates such that the unobservable directions become independent of the linearization point. As a result, an EKF-SLAM based on those coordinates, namely an Invariant EKF SLAM, computes beliefs that always match with the unobservable directions. And this is what allowed to prove consistency properties in [8].

6.2 Experimental results

The smoothing and mapping (SAM) approach is an alternative to the filtering approach for SLAM. The idea is roughly to formulate the problem of finding the most likely robot's trajectory and map as a big least

squares problem that involves all the past states and the past measurements up to current time step. The method finds its roots in [54] to our best knowledge. The problem can be solved incrementally, and one of the most commonly used such algorithm is the iSAM of [39], that can be implemented using the package available at <https://svn.csail.mit.edu/isam>. At each time step, it returns the maximum likelihood (MaxLik) value of the state given all past the measurements. When it converges, iSAM can be viewed as an optimal filter and provides a reference estimate. This is why over the past recent years, the SLAM community has gradually turned to optimization based techniques.

The conventional EKF, described in Algorithm 1, the Invariant EKF which is a right multiplication based version of Algorithm 2, and the iSAM, have been tested on the publicly available real Victoria Park dataset described in [31]. It is a large scale experiment, in which a vehicle drives through a Park where (hundreds of) trees constitute the landmarks. Estimated trajectories, as well as the GPS measured trajectory, are displayed in Figure 10. The SLAM algorithms use the odometers, steering angle, and laser sensors' measurements, whereas GPS measurements are solely used for evaluation purposes. Table 2 displays the error with respect to GPS measurements for various tunings of covariance matrices (as noise characteristics are not precisely known, various choices are possible indeed).

The results are very instructive regarding robustness of the filters: the results of IEKF and iSAM are almost insensitive to retained tuning parameters (RMSE stays between 5 and 11 *m*) while the EKF is very sensitive to tuning parameters. This can be a problem as the tuning of noise matrices when no ground truth is available is always difficult in practice, see e.g. [1]. Moreover, IEKF and iSAM results are always very close to each other. This is interesting, as there already is a large corpus of experience and method in the industry for EKF implementation in the field of navigation, which is not yet the case for optimization based filtering techniques. In particular, the Invariant EKF has been already industrially implemented on a flying device presented in Section 5 and invariant Kalman filtering may thus open up for novel robust industrial SLAM algorithms.

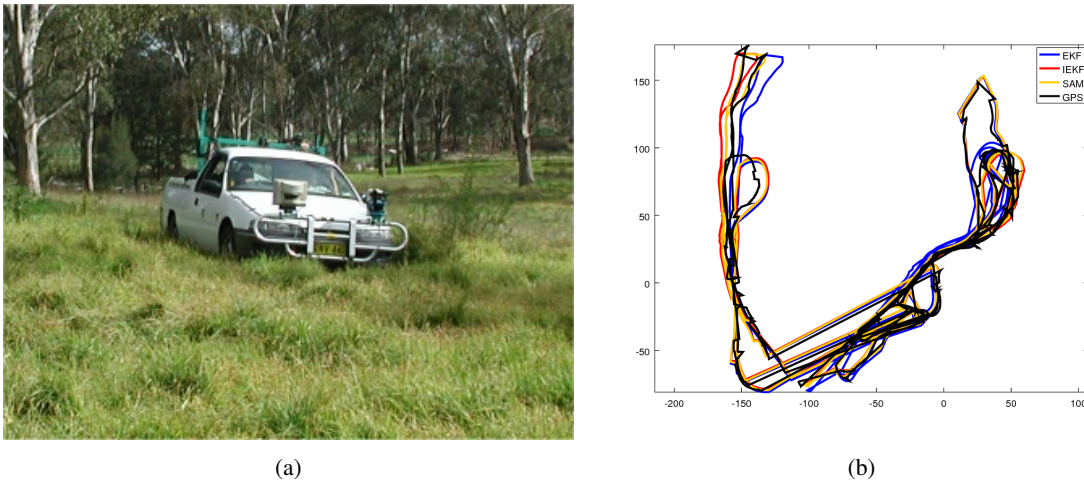


Figure 10: **(a)** The utility car used for the experiment is equipped with a Sick laser range and bearing sensor, linear variable differential transformer sensor for the steering and back wheel odometer. Image courtesy of J. Guivant and E. Nebot, Australian Centre for Field Robotics. **(b)** Trajectory (in meters) of the car estimated by several SLAM algorithms and GPS measurements (Robocentric mapping filter's estimates were removed to improve readability).

$\sigma_V \backslash \sigma$	1	4	8
1	6.56 m	9.61 m	10.50 m
10	5.52 m	5.99 m	6.29 m

$\sigma_V \backslash \sigma$	1	4	8
1	6.50 m	9.07 m	10.84 m
10	5.95 m	6.42 m	6.45 m

$\sigma_V \backslash \sigma$	1	4	8
1	6.54 m	123.01 m	34.5 m
10	6.32 m	10.04 m	13.2 m

Table 2: Position error (in meters) on the Victoria Park dataset for several noise tuning parameters. Parameter σ (in %) encodes model noise and σ_V (in m) observation noise. The trajectories returned by the filters are rotated and translated to match at best the GPS measurements. The residual errors in meters are reported. Note that, results achieved by IEKF and iSAM are very close, and weakly sensitive to (large) variations in the tuning parameters, contrarily to the EKF. This is another argument advocating the robustness of Invariant Kalman filtering.

7 CONCLUSION

The use of differential geometry and more precisely continuous symmetries for estimation emerged in the non-linear automatic control field decades ago. Attitude estimation, essentially for the control of unmanned aerial vehicles (UAVs), has been a strong driver for the development of Lie-group based estimators over the past ten years, the most popular attitude estimator probably being the non-linear complementary filter of [43].

Ten years ago, invariant Kalman filtering was introduced in [11] and has motivated many developments. The invariant EKF has since been mathematically shown to come with convergence properties, it has been implemented industrially for drone navigation, and it has recently been shown to resolve a major issue of the EKF when applied to simultaneous localization and mapping (SLAM). SLAM is an important application, that is sometimes considered as the “holy grail” of robotics, as it allows making a robot truly autonomous [29]. Invariant Kalman filtering opens up for novel implementations, as the EKF is a proven algorithm in terms of implementation and validation, but previous attempts to apply EKF to SLAM were not satisfactory.

This review has aimed to provide an accessible introduction to the methodology of invariant Kalman filtering, and to allow the reader to gain insight into the relevance of the method, as well as what the important differences with the conventional EKF are. This should be of interest to readers intrigued by the application of mathematical theories to practical applications, and also to readers interested in finding simple to implement and robust filters for localization, navigation, and SLAM, notably for autonomous vehicle guidance.

8 SUMMARY POINTS

1. Invariant Kalman filtering is a recent methodology to design extended Kalman filters (EKFs) based on alternative coordinates dictated by geometry.
2. The invariant EKF comes with convergence, stability, and robustness properties that the conventional EKF lacks. It is yet reserved for a class of systems on Lie groups, or systems that are close to this

class.

3. The invariant EKF is particularly suited to localization and navigation of autonomous vehicles, and has been successfully implemented in an industrial product of Safran Electronics & Defense, which is the n° 1 company in Europe for inertial navigation systems.
4. The invariant EKF has recently been proved to resolve the inconsistencies of the EKF for the important application of SLAM in robotics. Although the standard EKF is the historical algorithm for SLAM, it had mostly been abandoned due to its inconsistencies.

References

- [1] Pieter Abbeel, Adam Coates, Michael Montemerlo, Andrew Y Ng, and Sebastian Thrun. Discriminative training of kalman filters. In Robotics: Science and systems, volume 2, page 1, 2005.
- [2] Nasradine Aghannan and Pierre Rouchon. On invariant asymptotic observers. In Decision and Control, 2002, Proceedings of the 41st IEEE Conference on, volume 2, pages 1479–1484, 2002.
- [3] Victor Ayala and Juan Tirao. Linear control systems on Lie groups and controllability. In Proceedings of symposia in pure mathematics, volume 64, pages 47–64. AMERICAN MATHEMATICAL SOCIETY, 1999.
- [4] Martin Barczyk and Alan F Lynch. Invariant observer design for a helicopter uav aided inertial navigation system. IEEE Transactions on Control Systems Technology, 21(3):791–806, 2013.
- [5] T. D. Barfoot and P. T. Furgale. Associating Uncertainty with Three-Dimensional Poses for Use in Estimation Problems. IEEE Transactions on Robotics, 30(3):679–693, 2014.
- [6] Alessandro Barp, Francois-Xavier Briol, Anthony D Kennedy, and Mark Girolami. Geometry and dynamics for markov chain monte carlo. Annual Reviews of Statistics, 2018.
- [7] A. Barrau and S. Bonnabel. The Invariant Extended Kalman Filter as a Stable Observer. IEEE Transactions on Automatic Control, 62(4):1797–1812, 2017.
- [8] Axel Barrau and Silvere Bonnabel. An EKF-SLAM Algorithm with Consistency Properties. Arxiv preprint, 2015.
- [9] Pedro Batista, Carlos Silvestre, and Paulo Oliveira. A GES attitude observer with single vector observations. Automatica, 48(2):388–395, 2012.
- [10] Jan Bohn and Amit K Sanyal. Unscented State Estimation for Rigid Body Motion on SE(3). In Decision and Control (CDC), 2012 IEEE 51st Annual Conference on, pages 7498–7503. IEEE, 2012.
- [11] S. Bonnabel. Left-invariant extended Kalman filter and attitude estimation. In 2007 46th IEEE Conference on Decision and Control, pages 1027–1032, 2007.
- [12] 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.
- [13] Silvere Bonnabel. Symmetries in observer design: Review of some recent results and applications to EKF-based SLAM. In Robot Motion and Control 2011, pages 3–15. Springer, 2012.
- [14] Silvère Bonnabel, Philippe Martin, and Pierre Rouchon. Symmetry-Preserving Observers. IEEE Transactions on Automatic Control, 53(11):2514–2526, 2008.

- [15] Silvére Bonnabel, Philippe Martin, and Erwan Salaün. Invariant Extended Kalman Filter: theory and application to a velocity-aided attitude estimation problem. In 48th IEEE Conference on Decision and Control, pages 1297–1304, Shanghai, China, 2009.
- [16] Guillaume Bourmaud, Rémi Mégret, Marc Arnaudon, and Audrey Giremus. Continuous-Discrete Extended Kalman Filter on Matrix Lie Groups Using Concentrated Gaussian Distributions. Journal of Mathematical Imaging and Vision, 51(1):209–228, 2015.
- [17] 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(4):581–586, 1997.
- [18] R. W. Brockett. Lie algebras and Lie groups in control theory. In Geometric methods in system theory, pages 43–82. Springer, 1973.
- [19] M. Brossard, S. Bonnabel, and J.P. Condomines. Unscented Kalman Filtering on Lie Groups. In Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on, 2017.
- [20] José A Castellanos, José Neira, and Juan D Tardós. Limits to the consistency of EKF-based SLAM. 2004.
- [21] G. S. Chirikjian and M. Kobilarov. Gaussian Approximation of Non-linear Measurement Models on Lie Groups. In Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on, pages 6401–6406. IEEE, 2014.
- [22] Gregory S. Chirikjian. Stochastic Models, Information Theory, and Lie Groups, Volume 1: Classical Results and Geometric Methods. Applied and numerical harmonic analysis. Birkhäuser, 2009.
- [23] Gregory S Chirikjian. Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications. Springer Science & Business Media, 2011.
- [24] Jean-Philippe Condomines, Cédric Seren, and Gautier Hattenberger. Pi-invariant unscented Kalman filter for sensor fusion. In Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on, pages 1035–1040. IEEE, 2014.
- [25] Jean-Philippe Condomines, Cédric Seren, Gautier Hattenberger, et al. Nonlinear state estimation using an invariant unscented Kalman filter. In AIAA Guidance Navigation and Control Conference, pages 1–15, 2013.
- [26] G. Dissanayake, P. Newman, H.F. Durrant-Whyte, S. Clark, and M. Csobra. A solution to the simultaneous localisation and mapping (SLAM) problem. IEEE Trans. Robot. Automat., 17:229–241, 2001.
- [27] T.E. Duncan. An estimation problem in compact Lie groups. Systems & Control Letters, 10(4):257–263, 1988.
- [28] Tyrone E Duncan. Some filtering results in Riemann manifolds. Information and Control, 35(3):182–195, 1977.
- [29] Hugh Durrant-Whyte and Tim Bailey. Simultaneous localization and mapping: part i. Robotics & Automation Magazine, IEEE, 13(2):99–110, 2006.
- [30] J. Forbes and D. E. Zotnik. Sigma Point Kalman Filtering on Matrix Lie Groups Applied to the SLAM Problem. In Geometric Sciences of Information (GSI), 2017, Springer LNCS, 2017.

- [31] Jose Guivant, Eduardo Nebot, and Hugh Durrant-Whyte. Simultaneous localization and map building using natural features in outdoor environments. In Intelligent Autonomous Systems, volume 6, pages 581–586, 2000.
- [32] Y. Han and F.C. Park. Least squares tracking on the Euclidean group. Automatic Control, IEEE Transactions on, 46(7):1127–1132, 2001.
- [33] Guoquan P Huang, Anastasios I Mourikis, and Stergios I Roumeliotis. Analysis and improvement of the consistency of extended Kalman filter based SLAM. In Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on, pages 473–479. IEEE, 2008.
- [34] Guoquan P Huang, Anastasios I Mourikis, and Stergios I Roumeliotis. Observability-based rules for designing consistent EKF SLAM estimators. The International Journal of Robotics Research, 29(5):502–528, 2010.
- [35] Shoudong Huang and Gamini Dissanayake. Convergence and consistency analysis for extended Kalman filter based SLAM. Robotics, IEEE Transactions on, 23(5):1036–1049, 2007.
- [36] Kiyosi Itô. Stochastic differential equations in a differentiable manifold. Nagoya Mathematical Journal, 1:35–47, 1950.
- [37] Simon J Julier and Jeffrey K Uhlmann. A New Extension of the Kalman Filter to Nonlinear Systems. In AeroSense’97, pages 182–193, 1997.
- [38] Simon J Julier and Jeffrey K Uhlmann. A counter example to the theory of simultaneous localization and map building. In Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on, volume 4, pages 4238–4243. IEEE, 2001.
- [39] Michael Kaess, Ananth Ranganathan, and Frank Dellaert. isam: Incremental smoothing and mapping. IEEE Transactions on Robotics, 24(6):1365–1378, 2008.
- [40] Alireza Khosravian, Jochen Trunpf, Robert Mahony, and Christian Lageman. Observers for invariant systems on Lie groups with biased input measurements and homogeneous outputs. Automatica, 55:19–26, 2015.
- [41] Taeyoung Lee. Global Unscented Attitude Estimation via the Matrix Fisher Distributions on $SO(3)$. In American Control Conference (ACC), 2016, pages 4942–4947. IEEE, 2016.
- [42] G. Loianno, M. Watterson, and V. Kumar. Visual Inertial Odometry for Quadrotors on $SE(3)$. In 2016 IEEE International Conference on Robotics and Automation (ICRA), pages 1544–1551, 2016.
- [43] Robert Mahony, Tarek Hamel, and Jean-Michel Pflimlin. Nonlinear complementary filters on the special orthogonal group. IEEE Transactions on automatic control, 53(5):1203–1218, 2008.
- [44] Philippe Martin, Erwan Salaün, et al. Generalized multiplicative extended Kalman filter for aided attitude and heading reference system. In AIAA Guidance, Navigation, and Control Conference, page 8300, 2010.
- [45] Leonard A McGee and Stanley F Schmidt. Discovery of the Kalman filter as a practical tool for aerospace and industry. 1985.
- [46] SK Ng and PE Caines. Nonlinear filtering in Riemannian manifolds. IMA journal of mathematical control and information, 2(1):25–36, 1985.
- [47] W. Park, Y. Liu, Y. Zhou, M. Moses, G. S. Chirikjian, et al. Kinematic State Estimation and Motion Planning for Stochastic Nonholonomic Systems Using the Exponential Map. Robotica, 26(4):419–434, 2008.

- [48] M. Pontier and J. Szpirglas. Filtering on manifolds. In Stochastic Modelling and Filtering, pages 147–160. Springer, 1987.
- [49] Salem Said and Jonathan H Manton. On filtering with observation in a manifold: Reduction to a classical filtering problem. SIAM Journal on Control and Optimization, 51(1):767–783, 2013.
- [50] Glauco Garcia Scandaroli, Pascal Morin, and Geraldo Silveira. A nonlinear observer approach for concurrent estimation of pose, imu bias and camera-to-imu rotation. In Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on, pages 3335–3341, 2011.
- [51] Yongkyu Song and Jessy W Grizzle. The extended Kalman filter as a local asymptotic observer for nonlinear discrete-time systems. In American Control Conference, 1992, pages 3365–3369, 1992.
- [52] Robert F Stengel. Optimal Control and Estimation. Courier Corporation, 1986.
- [53] Abdelhamid Tayebi, Stephen McGilvray, A Roberts, and M Moallem. Attitude estimation and stabilization of a rigid body using low-cost sensors. In Decision and Control, 2007 46th IEEE Conference on, pages 6424–6429, 2007.
- [54] Sebastian Thrun and Michael Montemerlo. The graph SLAM algorithm with applications to large-scale mapping of urban structures. The International Journal of Robotics Research, 25(5-6):403–429, 2006.
- [55] A. S. Willsky. Some estimation problems on Lie groups. In D.Q. Mayne and R.W. Brockett, editors, Geometric Methods in System Theory, volume 3 of NATO Advanced Study Institutes Series, pages 305–314. Springer Netherlands, 1973.
- [56] A. S. Willsky and S. I. Marcus. Estimation for bilinear stochastic systems. In Variable Structure Systems with Application to Economics and Biology, pages 116–137. Springer, 1975.
- [57] A.S. Willsky. Dynamical Systems Defined on Groups: Structural Properties and Estimation. PhD thesis, MIT Dept. of Aeronautics and Astronautics, May 1973.
- [58] K Wolfe, Michael Mashner, and G. S. Chirikjian. Bayesian Fusion on Lie groups. Journal of Algebraic Statistics, 2(1):75–97, 2011.
- [59] Mohammad Zamani, Jochen Trumpf, and Robert Mahony. Minimum-energy filtering for attitude estimation. IEEE Transactions on Automatic Control, 58(11):2917–2921, 2013.
- [60] Teng Zhang, Kanzhi Wu, Jingwei Song, Shoudong Huang, and Gamini Dissanayake. Convergence and consistency analysis for a 3-d invariant-EKF SLAM. IEEE Robotics and Automation Letters, 2(2):733–740, 2017.
- [61] Teng Zhang, Kanzhi Wu, Daobilige Su, Shoudong Huang, and Gamini Dissanayake. An invariant-EKF VINS algorithm for improving consistency. arXiv preprint arXiv:1702.07920, 2017.