Kalman filtering with a class of geometric state equality constraints

Paul Chauchat, A Barrau, S Bonnabel

Kalman filtering with a class of geometric state equality constraints

P. Chauchat, A. Barrau, S. Bonnabel

***Abstract*— In this paper we consider a noise free class of dynamics encompassing left- and right-invariant on a Lie group with noisy partial state measurements. We assume in addition that the covariance matrix of the state is initially rank-deficient. This, combined with the absence of process noise, keeps the system state within a (time-dependent) subset of the state space at all times. We prove mathematically that the invariant extended Kalman filter (IEKF) perfectly respects this kind of state constraints, contrarily to the standard EKF, or the unscented Kalman filter. This is a strong indication that the IEKF is particularly well suited to navigation when motion sensors are highly precise. The theory is applied to a non-holonomic car example on $SE(2)$, and to an attitude estimation example on $SO(3)$.**

I. INTRODUCTION

The Kalman filter (KF) and its extended version (EKF) have appeared in the 1960s, and played a big role in the guidance of spacecrafts during the space age. It has been the state of the art for industrial applications since the 1960s, notably for navigation. However, due to the nonlinear nature of the navigation equations, and in particular to the fact the orientation of the aircraft (i.e., the attitude) does not live in a vector space, the EKF may have some shortcomings. This has motivated the development of alternative filters, especially for attitude estimation, see e.g., [12], [19], [6], [15], [13], [10], [16], [20], [7].

The Invariant Extended Kalman Filter (IEKF) is a relatively recent variant of the EKF meant to account for the nonlinearities of the state space when devising EKFs on Lie groups, see [9], [8], [14], [2]. As such, it can be viewed as a variant of the multiplicative extended Kalman filter (MEKF) [11] for attitude estimation, and as an extension to it for more general state spaces. The two main arguments that advocate its use over other EKF variants are 1- its (local) guaranteed convergence properties [4], and 2- the fact it solves the well-studied inconsistency issues of the EKF for simultaneous localization and mapping (SLAM) and competes with state of the art SLAM algorithms, see [1].

The object of the present paper is to derive a novel general theory that advocates the use of the IEKF for attitude estimation, and more generally state estimation in navigation, when using very precise motion sensors. Indeed, the IEKF has been shown to literally outperform the EKF when the process noise is very small in simulations [4], and patented industrial results have shown great improvement over existing method for the so-called inertial measurement unit (IMU) alignment problem [5]. To be able to derive mathematical result, we need to consider the *limit case* where the motion sensors are ideal, that is, noise free, and when we have as well some prior deterministic information on the state (that is, the initial covariance matrix is rank deficient). Even if the mathematical guarantees obtained below do not strictly apply to the case of highly precise - but not perfect - sensors, they provide a strong indication that the IEKF is particularly suited for this setting, as previously observed [4], [5]. However, the impact of a rank-deficient initial covariance matrix was only studied in the very particular case of a non-holonomic car with GPS measurements [3].

In the present paper we consider mixed invariant dynamics on a Lie group. We assume the motion is noise free (i.e., no process noise) and the measurements are noisy. We furthermore assume the covariance matrix of the initial state to be rank deficient. Combined with noise-free dynamics this implies that at all times 1- the covariance matrix is rank deficient and 2- the state can only "reach" a well-characterized submanifold of the state space, that is, the "physical" state space is constrained at all times. Unfortunately, due to the fact it is based on approximations, the EKF (and the other Kalman variants such as the unscented Kalman filter) fail to capture those constraints. On the other hand, the IEKF embraces the Lie group structure of the state space, and is shown to perfectly ensure properties 1 and 2 above. Thus, the present paper allows to theorize and generalize our previous results on a non-holonomic car example in [3], where properties 1 and 2 have been shown "manually" on the studied example, and then leveraged to derive some global convergence properties of the filter for that example.

There has been prior authoritative work on Kalman filtering with state constraints, see [18], [17]. Indeed, in case of state constraints various methods exist to make the KF or the EKF respect them, ranging from the addition of perfect measurements to the addition of a projection step and probability density function truncation. But to our best knowledge, nothing has been done in this direction when the state constraints derive from symmetries. To this respect one could object it suffices to devise a state constrained EKF instead of using an IEKF. This might work, but then the obtained filter cannot be used with very small process noise, whereas the IEKF is still shown to outperform the EKF then [4] (but there is no more state constraint to preserve).

The paper is organized as follows. Section II presents the implications of deterministic dynamics in the linear case, and how the KF naturally encodes this information. Section III

Paul Chauchat and Silvere Bonnabel are with MINES ParisTech, PSL Reasearch University, Centre for Robotics, 60 bd Saint-Michel, 75006 Paris, France `paul.chauchat@mines-paristech.fr`, `silvere.bonnabel@mines-paristech.fr`

Axel Barrau is with SAFRAN TECH, Groupe Safran, Rue des Jeunes Bois - Châteaufort, 78772 Magny Les Hameaux CEDEX, France `axel.barrau@safrangroup.com`

describes the counterparts for systems devised on Lie groups and concisely summarizes the general derivation of the IEKF equations. Section IV contains the main theorem of this work stating the Invariant EKF naturally encodes the considered constraints. Finally, the implications of the theorem are illustrated on two examples, the non-holonomic car from [3] and an attitude estimation example.

II. Linear Kalman filtering with initial state constraint

We recall in this section a known property of linear Kalman filtering regarding its ability to handle partially deterministic information if no process noise is added. Although theoretical, this limit case is pivotal to filter robustness as will be illustrated on a simple example.

A. Considered system

Consider a classical continuous-discrete linear system in $\mathbb{R}^p$:

$$\dot{X}_t = A_t X_t, \tag{1}$$

$$Y_n = C_n X_{t_n} + V_n, \tag{2}$$

where $X_t \in \mathbb{R}^p$ is the state of the system at time t, $(t_n)_{n\geq 0}$ the sequence of observation times, $(Y_n)_{n\geq 0}$ the corresponding observations and V_n a Gaussian noise polluting observation n. A_t and C_n are the matrices defining the dynamics of the system and the function of the system observed through Y_n, respectively. Note that process noise has been deliberately removed in Equation (1).

We also make the additional assumption that the initial distribution of the state lies in an affine subspace V_0 of $\mathbb{R}^p$:

$$\exists a \in \mathbb{R}^p,\ X_0 \in a + V_0. \tag{3}$$

It can then be immediatley deduced from (3) that X_t lives in a known affine subspace at any time t:

$$X_t \in F_t a + F_t V_0, \tag{4}$$

where F_t is the matrix solution of the equation $F_0 = I_p$, $\frac{d}{dt}F_t = A_t F_t$ and the vector space defined as $F_t V_0 = \{F_t x, x \in V_0\}$. We will see now that the constraint (4) is automatically met, by the estimate of a Kalman filter.

Remark 1: In the framework of classical Kalman theory, system (1)-(2) is pathological, inasmuch as the state propagation is deterministic. This limit case is intended to give insight into the way a filter handles initial hard constraints such as (3) when the process noise is not sufficient to balance the ill-conditioning of the initial covariance matrix. Section V will show this situation can be extremely troublesome for non-linear systems.

B. The Linear Kalman Filter preserves the additional information

Initial linear information of the form (3) on a linear system is flawlessly captured by a Kalman filter (KF), as illustrated by the following proposition.

Proposition 1: Let $\hat{X}_0$ and P_0 be respectively the intial estimate and covariance matrix of a Kalman Filter tracking System (1)-(2), and assume they are consistent with condition (3) in the following sense:

$$\hat{X}_0 \in a + V_0, \qquad H_0 P_0 H_0^T = 0_{p\times p}, \tag{5}$$

where matrix H_0 is the orthogonal projection over the orthogonal complement of subspace V_0. Then, the state $\hat{X}_t$ and covariance matrix P_t returned at any time by the Kalman filter are consistent with (4), i.e., we have:

$$\hat{X}_t \in a + V_t, \qquad H_t P_t H_t^T = 0_{p\times p}, \tag{6}$$

Proof: Between two measurements, the Ricatti equation $\frac{d}{dt}P_t = A_t P_t + P_t A_t^T$ implies $H_t P_t H_t^T = 0$. Before an update, as P_{t_n} is symmetric, $H_{t_n} P_{t_n} = 0$ necessarily, and thus $H_{t_n} K_n = 0$. In turn, this implies $H_{t_n} P_{t_n}^+ H_{t_n}^T = 0$ and that $x_{t_n}^+$ remains in the subspace, as the update writes $P_{t_n}^+ = (I - K_n C_n) P_{t_n}$ and $\hat{x}_{t_n}^+ = \hat{x}_{t_n} + K_n (Y_n - C_n \hat{x}_{t_n})$. ∎

Conditions (5) and (6) are easily interpreted: initial error covariance over a direction orthogonal to V_0 (resp. V_t) is zero at time 0 (resp. t). Thus, Proposition 1 implies that at all times the estimate of the Kalman filter remains in the subspace the state lives in (if initialized in V_0), and P_t keeps reflecting the absence of dispersion of the probability distribution of the state orthogonally to this subspace.

C. Non-linear case

In the non-linear case, the initial subset in which the state lives may not be a vector space. But even if it is, it is distorted by the dynamics. Therefore, linearizations do not lead to updates that remain in that space. In turn, this leads the EKF to degradated performance, even in the presence of small process noise, as illustrated in the simulations of [3] and [4].

The aim of the present paper is to show that although the property above, along with Proposition 1, seems to be reserved for linear systems, it has in fact a counterpart for a class of non-linear dynamics and carries over to the Invariant EKF. This was already proved "manually" for a particular example in [3]. The results of the latter paper will prove to be a particular case of the general theorey developped herein.

III. Invariant Kalman filtering with geometric constraints

In this section, the general results will systematically be illustrated by an attitude estimation example, to help the reader grasp the theoretical ideas and concepts.

A. A short primer on matrix Lie groups

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

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

If $\gamma(t)$ is a curve over G with $\gamma(0) = Id$, then its derivative at $t = 0$ necessarily lies in a subset $\mathfrak{g}$ of $\mathscr{M}_N(\mathbb{R})$. $\mathfrak{g}$ is a vector space, called the Lie algebra of G and has same dimension as G. Thanks to a *linear* map from $\mathbb{R}^{\dim \mathfrak{g}} \to \mathfrak{g}$ denoted by $\xi \mapsto \xi^\wedge$, one can advantageously identify $\mathfrak{g}$ to $\mathbb{R}^q$ where $q = \dim G$. Besides, the vector space $\mathfrak{g}$ can be mapped to the matrix Lie group G through the classical matrix exponential $\exp_m$. Thus, $\mathbb{R}^q$ can be mapped to G through the Lie exponential map defined by $\exp(\xi) := \exp_m(\xi^\wedge)$ for $\xi \in \mathbb{R}^q$. We have thus $\exp(\xi) = Id + \xi^\wedge + O(\xi^2) \in \mathscr{M}_N(\mathbb{R})$.

B. Dynamics on a Lie group with equivariant state equality constraints

1) Considered nonlinear system: We first introduce a subclass of systems of [4]. Consider the following deterministic dynamics on a matrix Lie group $G \subset \mathbb{R}^{N\times N}$:

$$\frac{d}{dt}\chi_t = \chi_t u_t + v_t \chi_t, \tag{7}$$

where $\chi_t \in G$ is the state space, u_t and v_t are processes taking values in the Lie algebra $\mathfrak{g}$.

Note that, although the system might "look" linear, (7) are non-linear dynamics. Indeed, linearity may not even make sense in this context, as G is usually not a vector space.

2) Geometric constraints: Suppose that initially the state is known to satisfy k equivariant constraints of the form

$$\chi_0 b^i = c^i, \tag{8}$$

for some $(b^i, c^i)_{1\leq i\leq k} \subset \mathbb{R}^N$. That is, the system verifies a multiplicative counterpart of (3), as (8) is equivalent to

$$\chi_0 \in a \cdot G_0 = \{ax, x \in G_0\},$$

where a is an element of G verifying $a \cdot b^i = c^i$ and G_0 is the stabilizer subgroup of G with respect to b^i, i.e., $G_0 = \{x, xb = b\}$. Now, we derive in Proposition 2 the multiplicative counterpart of condition (4) of the linear case.

Proposition 2: Consider the dynamics (7) with initial condition (8). For all $1 \leq i \leq k$, let b_t^i and c_t^i be defined by the differential equations

$$b_0^i = b^i \qquad \frac{d}{dt}b_t^i = -u_t b_t^i \tag{9}$$

$$c_0^i = c^i \qquad \frac{d}{dt}c_t^i = v_t c_t^i \tag{10}$$

Then at all times we have necessarily

$$\forall t \geq 0,\ \chi_t \in \{\chi \in G \mid \forall i,\ \chi b_t^i = c_t^i\}. \tag{11}$$

Proof: It is straightforward to see that, for all i, $\frac{d}{dt}(\chi_t b_t^i - c_t^i) = u_t(\chi_t b_t^i - c_t^i)$. This linear system being equal to zero at $t = 0$ and linear, it is identically zero, which leads to $\forall t,\ \chi_t b_t^i - c_t^i = 0$. ∎

C. Illustration in terms of attitude estimation

As announced, let us present now a specific case of (7) modeling an attitude estimation problem. The state is represented by a rotation matrix $R_t \in SO(3)$, which maps the coordinates of a vector expressed in the body frame to those in the static frame. Letting ω_t denote the perfectly measured angular velocity, the dynamics read:

$$\frac{d}{dt}R_t = R_t(\omega_t)_\times, \tag{12}$$

with $(a)_\times$ the skew symmetric matrix associated to vector a.

A geometric constraint of the form of (8) for this system can mean that, when initialising, the vehicle was able to measure in its frame a known vector with certainty, say the direction of a distant star thanks to a high-definition camera. Denoting by s_{fixed} and s_0 the direction of the star in the fixed and the initial mobile frame respectively, this reads:

$$R_0^T s_{fixed} = s_0 \Leftrightarrow R_0 s_0 = s_{fixed} \tag{13}$$

Thus, Proposition 2 states that the true system always knows the true direction of the star, i.e., satisfies, for s_t such that $\frac{d}{dt}s_t = -(\omega_t)_\times s_t$,

$$\forall t,\ R_t^T s_{fixed} = s_t. \tag{14}$$

D. Invariant filtering on a Lie group

Consider a system with equivariant constraints such as in Section III-B, and left-equivariant noisy output Y^L:

$$Y_n^L = \chi_{t_n} \cdot d + V_n, \tag{15}$$

where $d \in \mathbb{R}^p$ and the V_n's are Gaussian independent noises. Of course the output can consist of several measurements of this type at each t_n considering various vectors $d_1, d_2, \cdots$.

1) Review of the L-IEKF equations for this system: The systems described in this section are suitable for the design of a Left-Invariant EKF (LIEKF), see e.g., [4]. It is defined through the usual propagation and update sequence. Assume discrete observations at times $(t_n)_{n>0}$, then it writes:

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

$$\hat{\chi}_{t_n}^+ = \hat{\chi}_{t_n} \exp\left[K_n(\hat{\chi}_{t_n}^{-1} Y_n - d)\right] \qquad \text{LIEKF Update} \tag{17}$$

where the function $K_n : \mathbb{R}^N \to \mathbb{R}^q$ ($q = \dim G$) is defined through linearizations as in the conventional EKF theory, using the following left-invariant state estimation error:

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

Such a non-linear error η_t can be associated to a vector $\zeta_t \in \mathbb{R}^q$ such that $\eta_t = \exp\zeta_t$. This form, along with the extensive use of the first-order linearization $\eta_t \approx Id + \zeta_t^\wedge$ allow us to derive the equations defining the observer's covariance, its propagation, and the gain used in (17). Indeed the linerizations of the left-invariant error system associated to (16), and (17) write:

$$\frac{d}{dt}\zeta_t = A_t \zeta_t \tag{19}$$

$$\zeta_{t_n}^+ = \zeta_{t_n} + K_n(C\zeta_{t_n} + \hat{V}_n) \tag{20}$$

where $\hat{V}_n = \hat{\chi}_{t_n}^{-1}$ represents the observation noise, A_t is the map defined by $(A_t\xi)^\wedge = \xi^\wedge u_t - u_t \xi^\wedge$ (see indeed Theorem 2 of [4]) and C is the matrix defined by $C\xi = -\xi^\wedge d$. Linearizations leading to (19)-(20) are detailed in the proof below.

The state error covariance P_t output by the IEKF is an approximation to $\mathbb{E}(\zeta_t \zeta_t^T)$. Thus Equations (19) and (20) lead to the following Kalman gain and covariance updates, where $\hat{N}$ denotes the covariance matrix of the observation noise $\hat{V}_n$:

$$\dot{P}_t = A_t P_t + P_t A_t^T \tag{21}$$

$$S_n = C P_{t_n} C^T + \hat{N}, \qquad K_n = P_{t_n} C^T S_n^{-1} \tag{22}$$

$$P_{t_n}^+ = (I - K_n C) P_{t_n} \tag{23}$$

IV. Mathematical results

In this section, we show our main result, stating that the equivariant equality constraints defined by Equation (8) are propagated by the LIEKF the same way as in Proposition 2, without having to incorporate them in the filter as hard constraints like (artificial) perfect measurements [17].

Theorem 1: Consider the LIEKF described by (16) and (17), associated to the dynamics (7), with initial equality constraint (8). It implies the constraint (11) at all times. Now, note that $\hat{\chi}_t b_t = c$ rewrites $\eta_t b_t = b_t$ due to definition (18). For small η_t the linearization of the latter equality writes $H_t \zeta_t = 0$, where $H_t : \xi \mapsto \xi^\wedge b_t$ since $\eta_t = \exp(\xi_t) = Id + \xi_t^\wedge + O(\xi_t^2)$. Suppose now that the filter is initialized such that

$$\eta_0 b_0 = b_0, \qquad H_0 P_0 H_0^T = 0 \tag{24}$$

Then the LIEKF estimate lies in the same submanifold (11) as the true system at all times since:

$$\forall t \geq 0,\ \hat{\chi}_t b_t = c_t, \qquad H_t P_t H_t^T = 0 \tag{25}$$

The second equality indicates the covariance output by the LIEKF correctly encodes an absence of dispersion orthogonally to the submanifold (11).

Proof: Let T_t denote the condition $H_t P_t H_t^T = 0$. To prove the theorem, it is enough to show the following four implications:

(i) $\eta_{t_n} b_{t_n} = b_{t_n} \Rightarrow \eta_t b_t = b_t$, for $t_n \leq t < t_{n+1}$
(ii) $T_{t_n}^+ \Rightarrow T_t$, for $t_n \leq t < t_{n+1}$
(iii) $T_{t_n} \Rightarrow T_{t_n}^+$ for $n > 0$
(iv) $[\eta_{t_n} b_{t_n} = b_{t_n}] \wedge T_{t_n} \Rightarrow \eta_{t_n}^+ b_{t_n} = b_{t_n}$

Proof of (i) : As the considered system has deterministic dynamics, this directly comes from Proposition 2 and (16).

Proof of (ii) : We have $\dot{\eta} = \eta u_t - u_t \eta$. Linearizing, it follows that $\dot{\zeta}^\wedge = \zeta^\wedge u_t - u_t \zeta^\wedge$ and we get $\dot{\zeta} = A_t \zeta$. Recall also that $\frac{d}{dt} P_t = A P_t + P_t A$. As P_t is symmetric, we can write $P_t = Q_t Q_t^T$, where $\frac{d}{dt} Q_t = A_t Q_t$. We will then prove that $H_t Q_t$ is identically zero, which implies $H_t P_t H_t^T = 0$ as wanted. We have:

$$\frac{d}{dt} H_t Q_t = (\dot{H}_t + H_t A_t) Q_t$$

From the definitions of H_t and A_t, we get that for all ζ, $\dot{H}_t \zeta = -\zeta^\wedge u_t b_t$, and $H_t A_t \zeta = (\zeta^\wedge u_t - u_t \zeta^\wedge) b_t$. Finally,

$$(\dot{H}_t + H_t A_t)\zeta = -u_t \zeta^\wedge b_t = -u_t H_t \zeta$$

Replacing ζ by the columns of Q_t, we get $\frac{d}{dt}[H_t Q_t] = -u_t [H_t Q_t]$, a linear system initialized at 0. Thus, it is identically zero, which proves $T_{t_n}^+ \Rightarrow T_t$ for $t_n \leq t < t_{n+1}$.

Proof of (iii) : Equation (17) rewrites in terms of error:

$$\eta_{t_n}^+ = \eta_{t_n} \exp\left[K_n(\eta_{t_n}^{-1} d - d + \hat{V}_n)\right]. \tag{26}$$

Linearizing then leads to $\zeta_{t_n}^+ \approx \zeta_{t_n} + K_n(\eta_{t_n}^{-1} d - d + \hat{V}_n) \approx \zeta_{t_n} + K_n(-\zeta_{t_n}{}^\wedge d + \hat{V}_n)$, that is, Equation (20). As $K_n = P_{t_n} C^T S_n^{-1}$ from (22), the image of K_n is included in that of P_{t_n}, which means that, thanks to T_{t_n}, we have Im $K_n \subset$ ker H_{t_n}. This directly leads to $H_{t_n} P_{t_n}^+ = H_{t_n} P_{t_n} - H_{t_n} K_n C P_{t_n} = 0$, i.e., $T_{t_n} \Rightarrow T_{t_n}^+$.

Proof of (iv): Suppose that $\eta_{t_n} b_{t_n} = b_{t_n}$. The matrix exponential map is defined by $\exp M = Id + \sum_{k\geq 1} \frac{M^k}{k!}$. By noting $z = \eta_{t_n}^{-1} d - d$, we thus have according to (26):

$$\begin{aligned}
\eta_{t_n}^+ b_{t_n} &= \eta_{t_n}(Id + \sum_{k\geq 1} \frac{(K_n z)^{\wedge k}}{k!}) b_{t_n} \\
&= \eta_{t_n} b_{t_n} + (\sum_{k\geq 0} \frac{(K_n z)^{\wedge k}}{k!}) H_{t_n}(K_n z) = b_{t_n}
\end{aligned}$$

which concludes the proof. ■

A. Direct corollary for the RIEKF

If the output is of the form

$$Y_n^R = \chi_{t_n}^{-1} \cdot d + V_n, \tag{27}$$

rather than (15), then it is said to be right-equivariant, and one should use a right-invariant EKF (RIEKF), see [4]. The results remain entirely valid then, by symmetry.

Theorem 2: If the hypotheses of Theorem 1 are satisfied, with output (27) instead of (15), then the RIEKF estimates also verify (25).

Proof: The dynamics (7) is neither right nor left invariant. It is such that χ_t^{-1} satisfies similar dynamics. Using $\chi_t^{-1} \in G$ as the state variable, the output (27) becomes left-equivariant as (15). And the RIEKF update for the variable χ_t^{-1} is exactly the LIEKF update. Theorem 1 then applies. ■

B. Graphical illustration of the theorem and discussion

As it was already known, part of what makes the IEKF work where a filter with linear update such as the EKF fails mostly is where the linearization is done. Indeed, the EKF tries to linearize on a non-linear space by embedding the state in the ambient vector space. Think again of $SO(3)$: there is no simple way of expressing a rotation as the sum of another rotation and some matrix. The IEKF however linearizes on the Lie algebra of the system, which is a linear space in its own right, the exponential map being just a translation between the two. When one writes $\chi = \exp(\xi)$, ξ is the axis of rotation of χ, the angle being the vector's norm, and summing rotation vectors makes perfect sense. It was already the idea behind the MEKF.

The second main argument which makes the proof work is the fact that the image of a Lie sub-algebra by the exponential map is a subgroup of the associated Lie group. This comes from the Baker-Campbell-Hausdorff formula, which states that if $X, Y \in \mathfrak{g}$, then $e^X e^Y = e^Z$ where Z is a series of X, Y and nested Lie bracket terms. Z thus stays in the sub-algebra.

This is illustrated by Figure 1, which gives a schematic view of the difference among the linear, the MEKF and IEKF updates, for an estimate lying on the subgroup represented by the circle. The IEKF update, through the exponential map, makes the estimate move along the circle. Since the covariance is expressed in the Lie algebra, its alignment stays consistent with the subgroup (11). On the contrary, the MEKF gives no guarantee that the estimate will stay in the subgroup. In the meantime, the EKF updates along a straight line in the direction of the covariance, so there are no guarantee that the estimate will even remain in G, or that the covariance will stay consistent with the curvature of the space.

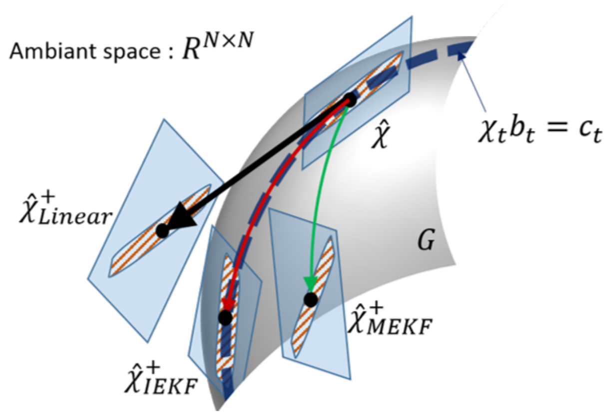

Fig. 1: Schematic difference among the linear, the MEKF and the IEKF updates, illustrating Theorem 1. The dotted curve represents the subgroup (11), and the hatched areas the covariances, originally consistent with (11). The IEKF stays on the subgroup, while the MEKF and the EKF update respectively leave the subgroup and even the group G.

As already said, the interest is that when process noise is low, the hard constraint becomes useless. However, the state will live near the manifold defined by (11), and so will the IEKF estimate due to the filters' very structure! It is easily understood that small process noise will lead to a situation being close to the one of Figure 1 indeed (the system smoothly depends on the process noise amplitude).

V. Examples

This section presents two examples illustrating the implications of Theorems 1 and 2. The first one shows that the result of [3] now appears as a direct application of Theorem 1. The second one presents the implications of Theorem 2 for the attitue estimation example of Section III-C, and illustrates what happens when noise is turned on.

A. Car position and heading estimation

1) Recall of the results of [3]: Consider the simple case of a non-holonomic car with perfect odometry, unknown heading and noisy position measurements. Suppose the initial position of the car is known. The dynamics are given by:

$$\frac{d}{dt}\theta_t = \omega_t, \qquad \frac{d}{dt}x_t = \begin{pmatrix} \cos(\theta_t)u_t \\ \sin(\theta_t)u_t \end{pmatrix}, \tag{28}$$

where θ_t is the heading of the car, x_t its position vector, and ω_t, u_t are the angular and linear velocities. Noisy position measurements $Y_n = x_{t_n} + V_n$ are acquired at discrete times $(t_n)_{n\in\mathbb{N}}$, corrupted by white noise V_n.

It was then proven in [3] that if $R(\theta)$ denotes the rotation matrix of angle θ, and $\hat{\theta}_t, \hat{x}_t$ denote the IEKF estimates, then $R(\theta_t)^T x_t = R(\hat{\theta}_t)^T \hat{x}_t = b_t$, where b_t is defined through the differential equation

$$b_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \qquad \frac{d}{dt}b_t = -\begin{pmatrix} 0 & -\omega_t \\ \omega_t & 0 \end{pmatrix} b_t + \begin{pmatrix} u_t \\ 0 \end{pmatrix}. \tag{29}$$

Figure 2, reproduced from [3], displays the trajectories of the true car and both EKF and IEKF estimates, for $\omega_t \equiv 0$. The IEKF car estimate is always traveling on a ray that passes through the origin, while that is not true for the EKF.

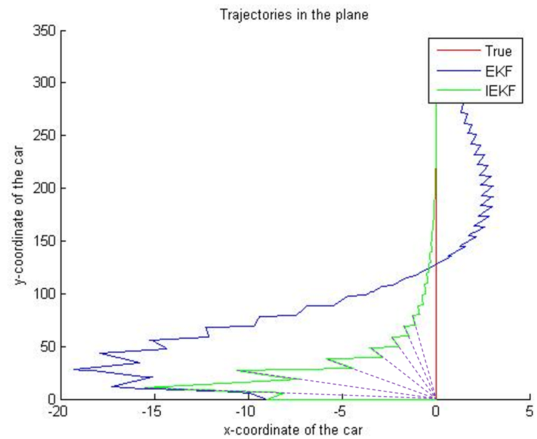

Fig. 2: First numerical example of Theorem 1: True, IEKF and EKF estimates trajectories of a car with perfect odometry, known initial position and uncertain initial heading. The car gets position measurement every 10 time steps. The odometry indicates $\omega_t \equiv 0$, thus the car is necessarily moving along a ray emanating from its original position. The IEKF encodes correctly this information, as its estimates always move along rays emanating from the initial position indeed.

2) Translation in the Lie group formalism: The system (28) can be seen as living in the Lie group $SE(2)$ with a state χ_t verifying

$$\chi_t = \begin{pmatrix} R(\theta_t) & x_t \\ 0_{1,2} & 1 \end{pmatrix}, \quad \dot{\chi}_t = \chi_t U_t, \quad U_t = \begin{pmatrix} 0 & -\omega_t & u_t \\ \omega_t & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{30}$$

Moreover, the position measurements and initial known position are respectively given by

$$Y_{t_n} = \chi_{t_n} \ (0\ 0\ 1)^T, \quad \chi_0 \ (b_0\ 1)^T = (0\ 0\ 1)^T \tag{31}$$

Thus, Proposition 2 implies that for $\beta_0 = \begin{pmatrix} b_0 & 1 \end{pmatrix}^T$ and $\dot{\beta}_t = -U_t\beta_t$, the state satisfies $\chi_t\beta_t = \begin{pmatrix} 0 & 0 & -1 \end{pmatrix}^T$ for all t, which is exactly what is stated in Section V-A.1.

In turn, the result of [3] showing that the LIEKF preserves the latter property is in fact a direct corollary of Theorem 1.

B. Attitude estimation

This example builds upon the work of [2]. Indeed, it did not consider the structure of the initial covariance, and in particular the impact of rank-deficiency. Consider that the system receives at times t_n, a measurement Y_n of the gravity

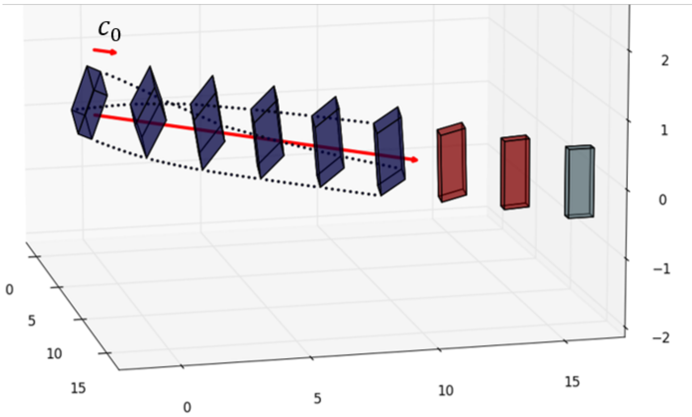

Fig. 3: Second numerical example. Bricks represent rotation matrices. Dots materialise rotations from one block to the next one, by highlighting the trajectory of each of the base vectors. The arrow represents the (known) direction of the distant star, c_0. The rotations induced by the first five updates of the IEKF are shown (blue), followed by the results after the 15th and 30th update (red) ,which show the convergence to the identity matrix (light blue). It can be clearly seen that they turn around c_0.

field g and the earth magnetic field b through a triplet of accelerometers and magnetometers, i.e.,

$$Y_n = (R_{t_n}^T g + V_n^g ; R_{t_n}^T b + V_n^b), \tag{32}$$

with V_n^g, V_n^b two centered noises in $\mathbb{R}^3$. It has thus right-equivariant outputs, and the system is suitable for the design of a RIEKF. The corresponding filter's equations for this problem were already written in [2]. They are similar to (17), with the update consisting of a left multiplication by a term of the form $\exp(K_n z)$, with $z \in \mathbb{R}^N$, that is, a rotation around the axis $K_n z$. Thus, theorem 2 implies that, if initialized correctly, the estimate will correctly estimate the direction of the star at all times whatever the motion. Figure 3 displays the results of numerical experiments. As stated, all the updates consist of rotations sharing the same axis, the direction of the star denoted c_0.

Impact of noise: We also performed simulations with process noise having standard deviation equal to 0.02 degrees/s. During all the experiment, the angle between c_0 and the axis of the updates' rotations was of course impacted and did not remain null. However, it never exceeded 0.01 degree, and the estimated star direction remained in a sharp cone, as could be expected from the IEKF's geometrical structure.

VI. Conclusion

This paper highlithgted the shortcomings of Extended Kalman Filtering for high-accuracy navigation problems through the degenerate situation of an infinitely accurate geometric prior. Without resorting to artificial process noise, the EKF fails to propagate this information. These situations are especially important in inertial navigation where they have to be handled carefully.

The novel result we proved shows the algebraic approach to filtering of the IEKF is a robust response to the issue of assimilating precise nonlinear prior information, ie rank-deficient initial covariance. In future work, this result will be extended to any system defined on a Lie group having group-affine dynamics as defined in [4], and to any information restricting the admissible state space to a subgroup.

VII. Acknowledgements

This work is supported by the company Safran through the CIFRE convention 2016/1444.

References

[1] A. Barrau and S. Bonnabel. An EKF-SLAM algorithm with consistency properties. *CoRR*, abs/1510.06263, 2015.

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

[3] A. Barrau and S. Bonnabel. Navigating with highly precise odometry and noisy GPS: a case study. *IFAC-PapersOnLine*, 49(18):618–623, 2016.

[4] A. Barrau and S. Bonnabel. The invariant extended Kalman filter as a stable observer. *IEEE Transactions On Automatic Control*, 2017.

[5] A. Barrau and S. Bonnabel. Alignment method for an inertial unit. *WIPO Patent Application WO/2015/075248*, 2014.

[6] P. Batista, C. Silvestre, and P. Oliveira. Attitude and earth velocity estimation-part i: Globally exponentially stable observer. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 121–126. IEEE, 2014.

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

[8] S. Bonnabel, Ph. Martin, and E. Salaun. 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.

[9] S. Bonnabel. Left-invariant extended Kalman filter and attitude estimation. In *IEEE Conference on Decision and Control (CDC)*, pages 1027–1032, 2007.

[10] G. Bourmaud, M. Arnaudon, R. Mégret, A. Giremus. Continuous-discrete extended Kalman filter on matrix Lie groups using concentrated Gaussian distributions. In *JMIV*, Vol 51(1), pp 209-228, 2015.

[11] J. L. Crassidis, F. Landis Markley, and Y. Cheng. Survey of nonlinear attitude estimation methods. *Journal of Guidance, Control, and Dynamics*, 30(1):12–28, 2007.

[12] M. D. Hua, G. Ducard, T. Hamel, R. Mahony, and K. Rudin. Implementation of a nonlinear attitude estimator for aerial robotic vehicles. *Control Systems Technology, IEEE Transactions on*, 22(1):201–213, 2014.

[13] M. Izadi and A. K. Sanyal. Rigid body attitude estimation based on the lagrange–dalembert principle. *Automatica*, 50(10):2570–2577, 2014.

[14] Ph. Martin, E. Salaün, et al. Generalized multiplicative extended kalman filter for aided attitude and heading reference system. In *Proc. AIAA Guid., Navigat., Control Conf*, pages 1–13, 2010.

[15] 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.

[16] A. K Sanyal and N. Nordkvist. Attitude state estimation with multirate measurements for almost global attitude feedback tracking. *Journal of Guidance, Control, and Dynamics*, 35(3):868–880, 2012.

[17] D. Simon. Kalman filtering with state constraints: a survey of linear and nonlinear algorithms. *IET Control Theory & Applications*, 4(8):1303–1318, 2010.

[18] D. Simon and T. L. Chia. Kalman filtering with state equality constraints. *IEEE transactions on Aerospace and Electronic Systems*, 38(1):128–136, 2002.

[19] K. C. Wolfe, M. Mashner, and G. S. Chirikjian. Bayesian fusion on Lie groups. *Journal of Algebraic Statistics*, 2(1):75–97, 2011.

[20] M. Zamani, J. Trumpf, and R. Mahony. On the distance to optimality of the geometric approximate minimum-energy attitude filter. In *ACC, 2014*, pages 4943–4948. IEEE, 2014.