53rd IEEE Conference on Decision and Control
December 15-17, 2014. Los Angeles, California, USA

Invariant Particle Filtering with application to localization

Axel Barrau and Silvère Bonnabel

Abstract— **The recently introduced Invariant Extended Kalman Filter (IEKF) is an extended Kalman filter designed for systems admitting symmetries, that possesses interesting convergence properties, and a relative independence of the filter behavior with respect to the system's trajectory. In the present paper, the ideas are extended to a broad class of systems introducing the notion of "conditional invariance", that is, invariance properties of the system once some of the state variables are known. We exploit this structure by devising an Invariant Rao-Blackwellized Particle Filter: those state variables are sampled, and the rest are marginalized out using IEKFs. The striking property of the obtained particle filter is that the Kalman gains are *identical* for all particles, leading to a drastic reduction of the computational burden. The strong potential of the method is illustrated by the challenging and realistic problem of localization from noisy inertial sensors and a noisy GPS having a randomly jumping bias.**

I. Introduction

The estimation of the state of a physical dynamical system using partial and noisy information plays a crucial role in a wide range of scientific fields such as mobile robotics, aeronautics, meteorology or hydrology. A popular approach, referred to as "filtering", consists of modeling the state evolution as a stochastic dynamical system and recursively compute its conditional probability density given a sequence of probabilistic measurements [11], [17]. The filtering problem is solved in closed form only in the particular case of linear Gaussian systems [20] by the well-known Kalman filter.

When the system is non-linear, there is no general method to solve the filtering problem, and the Kalman approach has given rise to multiple approximative methods where the system is linearized, and the state conditioned on the past inputs and outputs is approximated by a Gaussian, the mean yielding a relevant estimate of the system's true state, and the covariance matrix indicating the extent of uncertainty in the estimate. This is the rationale of the celebrated Extended Kalman Filter (EKF) and more recently the Unscented Kalman Filter (UKF) [19], where a deterministic sampling is meant to capture more precisely the second-order moments of the involved distributions.

Another wholly different route to attack non-linearities, that has gained immense popularity over the last twenty years, is to use sampling-based algorithms to numerically approximate the conditional distribution of the system's state. Notably, particle filters randomly sample a set of weighted particles supposed to accurately represent the desired density [16], [14]. Such methods are more flexible than Kalman filtering approaches as they allow to handle highly non Gaussian noise (as it is able to track multimodal distributions), or to incorporate negative information. This flexibility has been the key to their recent success for the simultaneous localization and mapping problem (SLAM) in robotics [24]. However, they come at the price of maintaining permanently a large number of particles in order to obtain a reasonable approximation of the true density. In particular, too a sparse sampling compared to the dimensionality of the problem leads to a type of degeneracy: no particle matches the data, even approximately, and the weights vanish. Several solutions have been proposed to tackle this issue, the two main ideas being to use the observations to smartly propagate the particles [25], and to sample only a reduced set of relevant variables, and use a classical filtering method, such as Kalman filtering, for each particle to marginalize out the remaining variables. The latter is known as the Rao-Blackwellized Particle Filter [15].

The present paper focuses on systems possessing a particular structure that can be related to the mathematical notion on symmetry. Indeed, for systems possessing symmetries, in a deterministic setting, the theory of symmetry-preserving observers [8] has recently introduced a class of observers such that the error between the estimated state and the true state follows an equation that depends on the trajectory only through a reduced number of variables. In the fully invariant case, i.e., the case of a left-invariant dynamics on a Lie group with a right-equivariant output, the error equation has been proved to be totally autonomous [5]. Several works have independently built upon this property to devise observers having strong convergence properties, see e.g., [7], [21], [26]. They are part of a more general effort to design observers specifically adapted to Lie groups [27] or more specifically rigid body movements [10], [22].

In the fully invariant case, and in a stochastic setting where the noise is explicitly accounted for, the Invariant Extended Kalman Filter, an EKF-like method with remarkable properties, has been introduced in continuous time in [4], [9], and in discrete time in [3], [2] where its properties are analyzed and related to the theory of Harris chains. Unfortunately, left-invariance with right-equivariant output is a strong structure that is reserved to a limited class of systems. In this paper, we show how IEKFs can be a applied to a much broader class of systems using approximate Rao-Blackwellized particle filters.

The organization and the contributions of the present paper are as follows. In Section II, we show that that the independence of the estimation error equation on the state is

A. Barrau and S. Bonnabel are with Centre de Robotique, MINES ParisTech, 75272 Paris, France [axel.barrau,silvere.bonnabel]@mines-paristech.fr

978-1-4673-6090-6/14/$31.00 ©2014 IEEE

not reserved to left-invariant systems on Lie groups. Indeed, we characterize a class of (not necessarily invariant) systems on Lie groups for which the estimation error equation does not depend on the followed trajectory but only on the system's inputs, leading to Kalman gains being independent of the trajectory and convergent as soon as the inputs are maintained constant. This is a novel result in itself extending the scope of invariant (Kalman) filtering. Then, in Section III, we combine this generalized invariant filtering theory to particle filters to extend its domain of application to a broad class of systems. Indeed, we introduce the notion of "conditional invariance", which essentially refers to invariance properties of a system conditioned on the fact that some state variables are known. We show how to exploit this structure by devising a particle filtering technique based on the Rao-Blackwellized Particle Filter, where those state variables are sampled, and the rest are marginalized out using IEKFs. The striking property of the obtained particle filter is that, as each particle is attached to an invariant system, the Kalman gains are *identical* for all particles, leading to a drastic reduction of the computational burden. Finally, in Section IV, an application of practical interest is considered: an Unmanned Aerial Vehicle (UAV) equipped with gyroscopes and accelerometers, and a GPS undergoing a non-Gaussian time-correlated noise modeled as the sum of a bias subject to random discrete jumps and a Gaussian residual error. The goal is to filter the noise out and recover the true state despite the jumping bias. The problem is of engineering interest, very challenging, and the model rather realistic. The proposed invariant particle filtering technique is shown to allow to correctly and efficiently estimate the true trajectory with only 100 particles, and to possess a much reduced computational complexity than any conventional approximate Rao-Blackwellized Particle Filter devised for this problem, due to the properties of the Invariant Kalman gains. The example is a contribution on itself and can be read (almost) independently.

II. Invariant filtering

A. Invariant update

We consider a class of filtering problems of the form:

$$\frac{d}{dt}\chi_t = f_{u_t}(\chi_t) \tag{1}$$

$$Y_{t_n} = \chi_{t_n} \cdot d \tag{2}$$

where the state χ is an element of a matrix Lie group G whose Lie algebra is noted $\mathfrak{g}$ and has dimension $N_{\mathfrak{g}}$, u_t is an input variable, d an element of the output space $\mathscr{Y} = \mathbb{R}^p$ and the symbol $\cdot$ an action of G on $\mathscr{Y}$. The classical multiplicative Kalman [12] filter reads:

$$\frac{d}{dt}\hat{\chi}_t = f_{u_t}(\hat{\chi}_t) \tag{3}$$

$$\hat{\chi}_{t_n}^+ = \hat{\chi}_{t_n}\phi[K_n(Y_{t_n} - \hat{\chi}_{t_n} \cdot d)] \tag{4}$$

where $\phi : \mathfrak{g} \to G$ is a retraction, i.e. a first-order approximation of the group exponential map [1]. The theory of invariant observers [6], [5] has brought to bear a slightly different form of updates that possess striking properties. Consider the "invariant" update:

$$\hat{\chi}_{t_n}^+ = \hat{\chi}_{t_n}\phi[K_n(\hat{\chi}_{t_n}^{-1}Y_{t_n} - d)] \tag{5}$$

K_n has merely be replaced by $K_n\hat{\chi}_{t_n}$ but the essential interest lies in the following result. Define the error variable η_t as:

$$\eta_t = \chi_t^{-1}\hat{\chi}_t \tag{6}$$

The evolution of this variable during the update reads:

$$\eta_{t_n}^+ = \chi_{t_n}^{-1}\hat{\chi}_{t_n}\phi[K_n(\hat{\chi}_{t_n}^{-1}\chi_{t_n} \cdot d - d)] = \eta_{t_n}\phi[K_n(\eta_{t_n}^{-1} \cdot d - d)] \tag{7}$$

The equation is state independent: it does not depend on the current estimation $\hat{\chi}_t$, a property usually reserved to the linear case.

B. State independent error equations

One can wonder in which cases the error equation is totally independent of the state's trajectory (including the propagation step, as in the linear case). This problem is solved by the following novel result:

Theorem 1: The evolution of the invariant error variable $\eta_t = \chi_t^{-1}\hat{\chi}_t$ is state trajectory independent, i.e. verifies an equation of the form $\frac{d}{dt}\eta_t = g_{u_t}(\eta_t)$, under the following necessary and sufficient condition:

$$\forall a, b \in G, f_{u_t}(ab) = f_{u_t}(a)b + af_{u_t}(b) - af_{u_t}(I_d)b.$$

Moreover $g_{u_t}(\eta) = f_{u_t}(\eta) - f_{u_t}(I_d)\eta$.

Proof: The condition reads $\frac{d}{dt}\eta_t = g_{u_t}(\eta_t)$ for a certain function g_{u_t}. Thus we have:

$$\begin{aligned} g_{u_t}(\chi_t^{-1}\hat{\chi}_t) &= \frac{d}{dt}\eta_t \\ &= -\chi_t^{-1}[\frac{d}{dt}\chi_t]\chi_t^{-1}\hat{\chi}_t + \chi_t^{-1}\frac{d}{dt}\hat{\chi}_t \\ &= -\chi_t^{-1}f_{u_t}(\chi_t)\eta_t + \chi_t^{-1}f_{u_t}(\hat{\chi}_t) \\ g_{u_t}(\eta_t) &= -\chi_t^{-1}f_{u_t}(\chi_t)\eta_t + \chi_t^{-1}f_{u_t}(\chi_t\eta_t) \end{aligned} \tag{8}$$

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

$$g_{u_t}(\eta_t) = f_{u_t}(\eta_t) - f_{u_t}(I_d)\eta_t \tag{9}$$

Reinjecting (9) in (8) we obtain:

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

The converse is trivial. ■

Remark 1: The particular cases of left-invariant and right-invariant dynamics, or their combinations, verify the hypothesis of Theorem 1. Let $f_{u_t}(\chi) = v_t\chi + \chi\omega_t$. We have indeed

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

Remark 2: In the particular case where G is a vector space the condition of theorem 1 reads $f_{u_t}(a+b) = f_{u_t}(a) + f_{u_t}(b) - f_{u_t}(0)$ and we recover the affine functions.

C. Invariant extended Kalman filtering

Assume the assumption of Theorem 1 is verified and introduce a noise in the process and the observation:

$$\frac{d}{dt}\chi_t = f_{u_t}(\chi_t) + \chi_t w_t \tag{10}$$

$$Y_{t_n} = \chi_{t_n} \cdot d + V_n \tag{11}$$

where w_t is a continuous white noise and V_n a discrete Gaussian noise. V_n only is assumed isotropic (under the action of G). The invariant Kalman filter is defined by the formulas:

$$\frac{d}{dt}\hat{\chi}_t = f_{u_t}(\hat{\chi}_t) \tag{12}$$

$$\hat{\chi}_{t_n}^{+} = \hat{\chi}_{t_n}\phi[K_n(\hat{\chi}_{t_n}^{-1}Y_{t_n} - d)] \tag{13}$$

The associated invariant error equation reads:

$$\frac{d}{dt}\eta_t = f_{u_t}(\eta_t) - f_{u_t}(I_d)\eta_t - w_t\eta_t \tag{14}$$

$$\eta_{t_n}^{+} = \eta_{t_n}\phi[K_n(\eta_{t_n}^{-1}\cdot d - d + V_n)] \tag{15}$$

where we have used the isotropy of V_n. The first-order optimal gains K_n can now be computed using a linearization of this error equation. Letting $F_t = \frac{\partial}{\partial \xi}[f_{u_t}(0,\phi(\xi)) - f_{u_t}(0,I_d)\phi(\xi)]$ and $H = \frac{\partial}{\partial \xi}[\phi(\xi)\cdot d]$ the linearized system reads:

$$\frac{d}{dt}\xi_t = F_t\xi_t + w_t$$
$$\xi_{t_n}^{+} = \xi_{t_n} - K_n(H\xi_{t_n} - V_n)$$

and the Kalman (Riccati) equation for the covariance matrix $P_t = E(\xi_t\xi_t^T)$ of the linearized error read:

$$\begin{aligned} \frac{d}{dt}P_t &= F_tP_t + P_tF_t^T + Q \\ S_n &= HP_{t_n}H^T + r^2I_p \\ K_n &= P_{t_n}H^TS_n^{-1} \\ P_{t_n}^{+} &= (I_{N_\mathfrak{g}} - K_nH)P_{t_n} \end{aligned} \tag{16}$$

III. Invariant Particle Filtering

In most filtering problems, either the assumptions of Theorem 1 are not verified or the observation cannot be defined as a group action. But a weaker property often holds: once a certain set of variables Θ is known the remaining marginal system has the desired properties and can be handled through Invariant Extended Kalman Filtering. If the associated gains are the same for any value of Θ the qualitative idea of Invariant Particle Filtering is simply to exploit this property to compute them once, and run several filters in parallel for free.

A. Conditional invariance

Assume the considered system can be divided into a set of variables gathered in a vector Θ_t, and an element χ_t of a matrix Lie group G, such (Θ,χ_t) verifies an equation of the form:

$$\frac{d}{dt}\Theta_t = \varphi(\Theta_t,\nu_t) \tag{17}$$

$$\frac{d}{dt}\chi_t = f_{u_t}(\Theta,\chi) + \chi_t w_t \tag{18}$$

$$Y_{t_n} = \psi_\Theta(\chi_{t_n}\cdot d + V_n) \tag{19}$$

where w_t is a Langevin noise, ν_t a Levy process, $\psi_\Theta : \mathscr{Y} \to \mathscr{Y}$ is invertible for any Θ and $f_.(\cdot,\cdot)$ verifies:

$$\forall u,\Theta,\eta, f_u(\Theta,\eta) - f_u(\Theta,I_d)\eta = f_u(0,\eta) - f_u(0,I_d)\eta \tag{20}$$

The additional parameter Θ can affect the observation as in the example of Section IV and/or the dynamics. An example of the latter case is given by the attitude χ_t (here a rotation matrix) of a body endowed with a gyroscope undergoing the combination of a bias Θ and an isotropic white noise, the stochastic evolution of the bias being possibly complicated.

B. Invariant Extended Kalman Filters running in parallel

The way we can take advantage of the property introduced in section III-A is illustrated by the situation where we consider two distinct trajectories $(\Theta_t^1)_{t\geq 0}$ and $(\Theta_t^2)_{t\geq 0}$ of (17) and let two Invariant Extended Kalman Filters run in parallel to filter the subsystem (18)-(19) where Θ is considered as an input.

Proposition 1: Let $(\Theta_t^1)_{t\geq 0}$ and $(\Theta_t^2)_{t\geq 0}$ be two solutions of (17). The equations (18)-(19) define then two different systems, one associated to Θ_t^1 and the other to Θ_t^2. If these systems are filtered by two IEKF's $(\hat{\chi}_t^i)_{i=1,2}$ of the form:

$$\frac{d}{dt}\hat{\chi}_t^i = f_{u_t}(\Theta_t^i,\hat{\chi}_t^i) \tag{21}$$

$$\hat{\chi}_{t_n}^{i+} = \hat{\chi}_{t_n}^i\phi(K_n^i[(\hat{\chi}_t^i)^{-1}\psi_{\Theta_t^i}^{-1}(Y_n) - d]) \tag{22}$$

then the Kalman gains K_n^i are *identical for both* systems.

Proof: It suffices to write the error equation associated to (21)-(22) and use (20) to obtain:

$$\frac{d}{dt}\eta_t^i = f_{u_t}(0,\eta_t^i) - f_{u_t}(0,I_d)\eta_t^i - w_t\eta_t^i \tag{23}$$

$$\eta_{t_n}^{i+} = \eta_{t_n}^i\phi[K_n((\eta_{t_n}^i)^{-1}\cdot d - d + V_n)] \tag{24}$$

which is independent of Θ_t^i. The linearization can be performed writing $\eta_t^i = \phi(\xi^i) \approx Id + D_\phi(\xi_t^i)$. ■

C. Invariant Particle Filtering algorithm

Inspired by the Rao-Blackwellized particle Filter, the Invariant Particle Filter consists of fully exploiting the conditional invariance property in the following way. The state being partitioned into (Θ_t,χ_t), the variable Θ_t can be sampled using particles $(\Theta_t^j)_{1\leq j\leq N_p}$, each particle having a weight w^j reflecting its likelihood assigned to it. There are numerous sampling methods that we will not detail here. Each time a new observation Y_n is available, the weight is updated and a resampling step may take place. The distribution of the state conditioned on the past outputs can then merely be factored using the chain rule as follows:

$$\mathbb{P}(\chi_t,\Theta_t|Y_1,...,Y_n) = \mathbb{P}(\Theta_t|Y_1,...,Y_n)\mathbb{P}(\chi_t|\Theta_t,Y_1,...,Y_n)$$

This distribution can be approximated combining particle filters and IEKFs to approximate $\mathbb{P}(\Theta_t|Y_1,...,Y_n)$ and $\mathbb{P}(\chi_t|\Theta_t,Y_1,...,Y_n)$ in the following way:

- $\mathbb{P}(\Theta_t|Y_1,...,Y_n)$ is computed using a particle filter, that is, is approximated by the distribution $\sum_{j=1}^{N_p} w^j \delta_{\Theta^j}(\Theta)$
- $\mathbb{P}(\chi_t|\Theta_t^j,Y_1,...,Y_n)$ is approximated by a normal law with mean and covariance output by an IEKF.

Using Proposition 1 we see that as soon as the system satisfies the conditional invariance property, the Kalman gains and covariances are identical for all particles, leading to a much decreased numerical complexity.

IV. Application

The method described above will now be used to integrate an advanced non-Gaussian and non-white GPS noise model into a localization algorithm. The problem of inertia/GPS fusion has been addressed using several kinds of Kalman filters [18], [23], [13], but most of them rely on Gaussian approximations. In [23] for instance, the GPS error is a Gaussian white noise, in [13] it is modeled as the sum of a bias following a small random walk and a white noise. Here, the GPS estimates are assumed to be polluted not only by white noise, but also by an offset that abruptly jumps at unknown times. Such a model reflects the realistic effects of highly correlated noises in GPS and the way we propose to handle it opens the door to the efficient integration of an even more sophisticated noise model.

A. Model

The inertial sensors delivering estimates at high rate, and the GPS at a much lower rate, we opt for the following continuous-time model with discrete observations, based on the equations of a rigid body moving on a flat earth:

$$\frac{d}{dt}R_t = R_t(\omega_t + w_t^{\omega})_{\times} \tag{25}$$

$$\frac{d}{dt}v_t = R_t(a_t + w_t^{a}) + g \tag{26}$$

$$\frac{d}{dt}x_t = v_t \tag{27}$$

$$Y_{t_n} = x_{t_n} + b_{t_n} + V_n \tag{28}$$

where for $u \in \mathbb{R}^3$, $(u)_\times$ is the skew-symmetric matrix such that $\forall z \in \mathbb{R}^3, (u)_\times z = u \times z$, R_t encodes the orientation of the body, v_t its velocity, x_t its position, b_t is the (jumping) GPS bias, Y_{t_n} the GPS measurement, $\omega_t \in \mathbb{R}^3$ the gyroscope measurement, $a_t \in \mathbb{R}^3$ the accelerometer measurement, w_t^{ω} the gyroscope noise, w_t^a the accelerometer noise (the covariance matrix of the concatenated noise (w_t^{ω}, w_t^a) will be denoted by Q), and V_n the GPS noise, assumed to be isotropic (see Remark 3) with variance $r^2 I_d$. The bias b_t is modeled as follows: there exist a sequence of times $(\tau_k)_{k\geqslant 0}$ independent of other variables and unknown to the user such that b_t is constant on $[\tau_k, \tau_{k+1}[$ and b_{τ_k} is a centered Gaussian variable with variance $\sigma_b^2 I_d$, independent of the past. The time elapsed between τ_k and τ_{k+1} follows an exponential law with known intensity, and we let p_J denote the probability of the bias to have jumped at leas once between two consecutive observations (p_J is constant here because we assume the GPS information arrives at a constant rate but all the computations hold if p_J is a function of $(t_n - t_{n-1})$).

Remark 3: The isotropic GPS noise assumption is not really restrictive as the method still works if the noise V_n is isotropic in the horizontal directions x, y only, and larger along axis z. This hypothesis is classical in GPS models.

B. Reformulation in a Lie group formalism

First we introduce the matrix Lie group of double homogeneous matrices:

$$G = \{ \begin{pmatrix} R & u_1 & u_2 \\ 0_{1\times 3} & 1 & 0 \\ 0_{1\times 3} & 0 & 1 \end{pmatrix}, R \in SO(3), u_1, u_2 \in \mathbb{R}^3 \}$$

and its Lie algebra:

$$\mathfrak{g} = \{ \begin{pmatrix} (\zeta)_\times & u_1 & u_2 \\ 0_{1\times 3} & 1 & 0 \\ 0_{1\times 3} & 0 & 1 \end{pmatrix}, \xi, u_1, u_2 \in \mathbb{R}^3 \}$$

The following linear mapping will be useful in the following:

$$\mathscr{L}\begin{pmatrix} \zeta \\ u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} (\zeta)_\times & u_1 & u_2 \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{pmatrix}$$

The variables are gathered in a double homogeneous matrix:

$$\chi_t = \begin{pmatrix} R_t & v_t & x_t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, w_t = \begin{pmatrix} w_t^{\omega} & w_t^a & 0_{3,1} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$d = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad V_n = \begin{pmatrix} B_n \\ 0 \end{pmatrix}$$

The equations become:

$$\frac{d}{dt}\chi_t = f_{u_t}(\chi_t) + \chi_t w_t \tag{29}$$

$$Y_{t_n} = \chi_{t_n} d + b_{t_n} + V_n \tag{30}$$

where $u_t = (\omega_t, a_t)$ and f_{u_t} is defined by:

$$\forall R \in SO(3), v, x \in \mathbb{R}^3,\ f_{u_t}\begin{pmatrix} R & v & x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} R\omega_t & Ra_t + g & v \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Besides, we have here $\psi_b(u) = b + u$. Elementary computations show f_{u_t} verifies the assumption of Theorem 1: $\forall a, b \in G, f_{u_t}(ab) = f_{u_t}(a)b + a f_{u_t}(b)$.

C. IEKF for known biases

In this preliminary subsection we assume the bias sequence $\tilde{b}_n = b_{t_1}, ..., b_{t_n}$ to be known. In this case the biases can be removed from the sequence of GPS measurements $\tilde{Y}_n = Y_{t_1}, ..., Y_{t_n}$ and the obtained system has the form (1)-(2). It can thus be filtered using an IEKF: the mean estimate follows then equations (3) and (5) with $\phi = \exp \circ \mathscr{L}$. Letting as before the error be $\eta_t = \chi_t^{-1}\hat{\chi}_t$ we have

$$\frac{d}{dt}\eta_t = g_{u_t}(\eta_t, \omega_t, a_t) = f_{u_t}(\eta_t) - f_{u_t}(I_5)\eta_t$$

To compute the Kalman gains K_n's we use the error covariance matrix equation (16), where F_t and H are obtained letting $\eta_t = I_5 + \mathscr{L}(\varepsilon_t)$ with $\varepsilon_t \in \mathbb{R}^9$, and using a first order Taylor expansion of $f_{u_t}(\eta_t) - f_{u_t}(I_5)\eta_t$, yielding:

$$F_t = \begin{pmatrix} -(\omega_t)_\times & 0_{3\times3} & 0_{3\times3} \\ -(a_t)_\times & -(\omega_t)_\times & 0_{3\times3} \\ 0_{3\times3} & I_3 & -(\omega_t)_\times \end{pmatrix}, \quad H = \begin{pmatrix} 0_{3\times3} \\ 0_{3\times3} \\ I_3 \end{pmatrix}.$$

For χ_t close to $\hat{\chi}_t$ the Gaussian error approximation reads:

$$\mathbb{P}(\phi^{-1}(\chi_{t_n}^{-1}\hat{\chi}_{t_n}) = \xi | \tilde{b}_n, \tilde{Y}_n) \approx \mathscr{N}(\xi, P_{t_n}, 0),$$

where $\mathscr{N}(x, \Sigma, \mu)$ is the value at x of the Gaussian density of mean μ and covariance matrix Σ.

Remark 4: There are two remarkable features of this filter. First, the Kalman gains do not depend on the bias b_{t_n}. This is logical but it will play an important role in the sequel. Then, around a large and relevant class of trajectories defined by constant inputs ω_t, a_t, the Riccati equation followed by the error covariance matrix is time-independent, and the Kalman gains are thus expected to converge.

D. Description of the proposed particle filter

The basis of the algorithm (and more generally of approximate Rao-Blackwellized Particle Filters) is the assumption that for a given sequence of bias values $\tilde{b}_n$ and an instant $t \in]t_n, t_{n+1}]$ the marginal density $\mathbb{P}(\chi_t | \tilde{b}_n, \tilde{Y}_n)$ is well approximated by a Gaussian with mean and covariance output by an IEKF. Thus, only the sequence of bias values has to be sampled, using N_p particles. As we are interested in the density of the current bias, each particle only stores the last bias b_t^j and conditional mean $\hat{\chi}_t^j$ (the conditional covariance matrix is computed using the invariant error equation). If the b_t^j are i.i.d sampled particles an empirical estimate of the posterior distribution $\mathbb{P}(b_t, \chi_t | \tilde{Y}_n)$ of the whole current state (b_t, χ_t) then writes:

$$\frac{1}{N_p} \sum_{j=1}^{N_p} \delta_{b_t^j}(b_t) \mathbb{P}(\chi_t | \tilde{b}_n^j)$$

As it is impossible to sample efficiently from the posterior distribution, we use importance sampling that weights the particles according to their likelihood. The following steps are thus repeated over time:

1) Propagation: the conditional estimates $\hat{\chi}_t^j$ evolve following the deterministic part of (29) and the Riccati equation (16) is integrated.
2) Resampling: when a measure Y_n is available the bias of each particle is resampled following the law $\mathbb{P}(b_n | \tilde{b}_{n-1}^j, \tilde{Y}_n)$.
3) Re-weighting: the weight attached to each particle is multiplied by $\mathbb{P}(Y_n | \tilde{b}_{n-1}^j, \tilde{Y}_{n-1})$, then the weights are normalized. A resampling can be performed.
4) Marginal density update: the conditional mean $\hat{\chi}_{t_n}^j$ associated to each particle is updated using (13) and the sampled bias b_n^j. The (particle independent) invariant covariance matrix P_{t_n} is updated.

The detailed particle filter is described by Algorithm 1 and the various derivations leading to it are described in the following subsection.

E. Main formulas derivation

To implement the procedure we should be able to:

1) Compute $\mathbb{P}(\chi_{t_n} | \tilde{b}_n^j, \tilde{Y}_n)$ from $\mathbb{P}(\chi_{t_{n-1}} | \tilde{b}_{n-1}^j, \tilde{Y}_{n-1})$, b_n, Y_n.
2) Compute $\mathbb{P}(Y_n | \tilde{b}_{n-1}^j, \tilde{Y}_{n-1})$ from P_{t_n}, b_{n-1}^j and $\hat{\chi}_{t_n}^j$.
3) Compute $\mathbb{P}(b_n | \tilde{b}_{n-1}^j, \tilde{Y}_n)$ from P_{t_n}, b_{n-1}^j and $\hat{\chi}_{t_n}^j$.

1) Computation of $\mathbb{P}(\chi_{t_n} | \tilde{b}_n^j, \tilde{Y}_n)$: We obtained in Section IV-C that $\mathbb{P}(\phi^{-1}(\chi_{t_n}^{-1}\hat{\chi}_{t_n}) = \xi | \tilde{b}_n, \tilde{Y}_n) \approx \mathscr{N}(\xi, P_{t_n}, 0)$. We can extract the relation $\mathbb{P}(x_{t_n} = x | \tilde{b}_n^j, \tilde{Y}_n) \approx \mathscr{N}(x, \hat{R}_{t_n}^j H P_{t_n} H^T \hat{R}_{t_n}^{jT}, \hat{x}_t^j)$ that will prove useful in the sequel.

2) Computation of $\mathbb{P}(Y_n | \tilde{b}_{n-1}^j, \tilde{Y}_{n-1})$: Let J_n be the event "the bias jumps between t_n and t_{n+1}". This event is independent of $\chi_{t_{n+1}}$. We let p_J denote its prior probability, and $\bar{J}$ its opposite. If J occurs, according to (28), Y_n is the sum of three independent Gaussian vectors: x_{t_n}, b_{t_n} and V_n. Thus

$$\mathbb{P}(Y_n | J, \tilde{b}_{n-1}^j, \tilde{Y}_{n-1}) = \mathscr{N}(Y_n,\ R_{t_n}^T H^T P_{t_n} H R_{t_n} + \sigma_b^2 I_3 + r^2 I_3, \hat{x}_{t_n}^j)$$

If $\bar{J}$ occurs, Y_n is the sum of two independent Gaussians, x_{t_n} and V_n, and a known vector b_{t_n}. Thus

$$\mathbb{P}(Y_n | \bar{J}, \tilde{b}_{n-1}^j, \tilde{Y}_{n-1}) = \mathscr{N}(Y_n,\ R_{t_n}^T H^T P_{t_n} H R_{t_n} + r^2 I_3, \hat{x}_{t_n}^j + b_{t_n}^j)$$

We can conclude introducing the quantities Π_1 and Π_2:

$$\Pi_1 = p_J \mathscr{N}(Y_n,\ R_{t_n}^T H^T P_{t_n} H R_{t_n} + \sigma_b^2 I_3 + r^2 I_3, \hat{x}_{t_n}^j)$$
$$\Pi_2 = (1 - p_J) \mathscr{N}(Y_n,\ R_{t_n}^T H^T P_{t_n} H R_{t_n} + r^2 I_3, \hat{x}_{t_n}^j + b_{t_n}^j)$$

yielding $\mathbb{P}(Y_n | \tilde{b}_{n-1}^j, \tilde{Y}_{n-1}) = \Pi_1 + \Pi_2$.

3) Computation of $\mathbb{P}(b_n | \tilde{b}_{n-1}^j, \tilde{Y}_n)$: This probability decomposes as follows:

$$\mathbb{P}(b_n | \tilde{b}_{n-1}^j, \tilde{Y}_n) = \mathbb{P}(J | \tilde{b}_{n-1}^j, \tilde{Y}_n) \mathbb{P}(b_n | J, \tilde{b}_{n-1}^j, \tilde{Y}_n) + \mathbb{P}(\bar{J} | \tilde{b}_{n-1}^j, \tilde{Y}_n) \mathbb{P}(b_n | \bar{J}, \tilde{b}_{n-1}^j, \tilde{Y}_n)$$

Conditioning on Y_n we have:

$$\mathbb{P}(J | \tilde{b}_{n-1}^j, \tilde{Y}_n) = \Pi_1 / (\Pi_1 + \Pi_2)$$
$$\mathbb{P}(\bar{J} | \tilde{b}_{n-1}^j, \tilde{Y}_n) = \Pi_2 / (\Pi_1 + \Pi_2)$$

If J occurs using the fact b_n is a Gaussian and $Y_n - \hat{x}_n^j$ a noisy measurement of it, the linear Kalman equations yield:

$$\mathbb{P}(b_n = b | J, \tilde{b}_{n-1}^j, \tilde{Y}_n) = \mathscr{N}(b, \hat{R}_{t_n}^{jT} P_n^b \hat{R}_{t_n}^j, \hat{R}_{t_n}^{jT} K_n^b \hat{R}_{t_n}^j (Y_n - \hat{x}_n^j))$$

with $K_n^b = \sigma_b^2 (H^T P_{t_n} H + \sigma_b^2 I_3 + r^2 I_3)^{-1}$ and $P_n^b = \sigma_b^2 (I_3 - K_n^b)$. If $\bar{J}$ occurs the bias does not change : $\mathbb{P}(b_n = b | \bar{J}, \tilde{b}_{n-1}^j, \tilde{Y}_n) = \delta_{b_{n-1}}^j(b)$. Finally $\mathbb{P}(b_n = b | \tilde{b}_{n-1}^j, \tilde{Y}_n)$ writes:

$$\Pi_1 \mathscr{N}(b, \hat{R}_{t_n}^{jT} P_n^b \hat{R}_{t_n}^j, \hat{R}_{t_n}^{jT} K_n^b \hat{R}_{t_n}^j (Y_n - \hat{x}_n^j)) + \Pi_2 \delta_{b_{n-1}}(b)$$

Remark 5: To sample from this law (see Algorithm 1) we first choose if the bias jumps (for instance comparing Π_1 to a uniform variable). Then, if it is the case, we use a Cholesky decomposition $P_n^b = (L_n^b)(L_n^b)^T$ and let $b_{n+1}^j = R_{t_n}^{jT} K_n^b R_{t_n}^j (Y_n - \hat{x}_{t_n}^j) + R_{t_n}^{jT} L_n^b X$ where X is sampled following a standard three dimensional normal law. Note that a direct Cholesky decomposition of the variance $R_{t_n}^{jT} L_n^b R_{t_n}$ would be a bad choice as the computation would have to be done for each particle.

Algorithm 1 Invariant Particle Filter

The prior estimation $\hat{\chi}_0$ and variance P_0 are supposed known
for $n = 1$ **to** T **do**
  Solve $R_0 = I_3$, $\frac{d}{dt} R_t = R_t \omega_{t_n+t}$ on $[0, \Delta t]$
  Compute $v_{\Delta t} = \int_0^{\Delta t} R_t a_{t_n+t}$ and $x_{\Delta t} = \int_0^{\Delta t} v_t$
  Solve $\frac{d}{dt} P_t = F_t P_t + P_t F_t^T + Q$
  Compute $S_n = H P_{t_n} H^T + r^2 I_3$ and $K_n = P_{t_n} H^T S_n^{-1}$
  Compute $K_n^b = \sigma^2 (H^T P_{t_n} H + \sigma^2 I_3 + r^2 I_3)^{-1}$, $P_n^b = \sigma^2 (I_3 - K_n^b)$
  Compute L_n^b such that $(L_n^b)(L_n^b)^T = P_n^{b+}$
  for $j = 1$ **to** N_p **do**
    $R_{t_{n+1}}^j = R_{t_n}^{j+} R_{\Delta t}$, $v_{t_{n+1}}^j = v_{t_n}^{j+} + R_{t_n}^{j+} v_{\Delta t} + (\Delta t) g$
    $x_{t_{n+1}}^j = x_{t_n}^{j+} + (\Delta t) v_{t_n}^j + R_{t_n}^{j+} x_{\Delta t} + \frac{1}{2} (\Delta t)^2 g$
    Compute $\Pi_1 = \mathbb{P}(Y|J_n)$ and $\Pi_2 = \mathbb{P}(Y|\bar{J}_n)$
    Sample $c \in [0,1]$ following an uniform law
    Update the weight: $w^{j+} = (\Pi_1 + \Pi_2) w^j$
    if $c < \Pi_1 / (\Pi_1 + \Pi_2)$ **then**
      Sample X following a 3-dimensional normal law
      $b^j = R_{t_n}^T K_b^n R_{t_n} Y_n + R_{t_n}^T L_n^b X$
    end if
    Build $\chi_{t_{n+1}}^j$ using $R_{t_{n+1}}^j$, $v_{t_{n+1}}^j$ and $x_{t_{n+1}}^j$
    $\chi_{t_{n+1}}^{j+} = \chi_{t_{n+1}}^j \exp(\mathscr{L}(K_n R_{t_{n+1}}^{jT} (Y_n - b^j - x_{t_{n+1}}^j)))$
    Extract $R_{t_{n+1}}^{j+}$, $v_{t_{n+1}}^{j+}$ and $x_{t_{n+1}}^{j+}$ from $\chi_{t_{n+1}}^{j+}$
  end for
  Compute $P_{t_n}^+ = (I_9 - K_n H) P_{t_n}$
  Normalize weights and possibly resample
  Compute the mean position $\tilde{x}_t = \sum_{j=1}^{N_p} w^{j+} x^{j+}$
end for

F. Some remarks about bias observability

Before the first GPS jump: A constant bias is not observable. Thus, the best one can do before the bias has jumped is to efficiently filter the GPS and inertial sensors noise to reconstruct the state up to a position offset.

After one jump: As the inertia gives an accurate short-term estimation of the trajectory, the bias jumps can be detected and the value updated using the difference between the estimated position and the GPS observation. But the residual centered noise V_n of the GPS (after bias correction) affects this first estimation and can be handled only using future GPS measurements. Hence the interest of sampling several values of the new bias then selecting those which best fit the observations to come.

After many jumps: As the bias is assumed centered, it can be recovered asymptotically. Indeed in the proposed algorithm, the re-weighting will eventually eliminate the particles whose sequence of bias has a mean far from zero. This prevents the position to diverge from the GPS trajectory in the long run, this discrepancy being erroneously explained explained by a large non-centered estimated bias.

TABLE I
PARAMETERS OF THE SIMULATION

Duration	$1000s$
Variance of the gyroscope noise	$10^{-8} rad/s^2$
Variance of the accelerometer noise	$10^{-6} m^2/s^5$
Probability of a GPS bias jump	$10^{-3} s^{-1}$
Frequency of the GPS observations	$10Hz$
Variance of the regular GPS observations	$1m^2$
Variance of the GPS bias	$100m^2$

TABLE II
RESULTS OF THE SIMULATION : RMSE OF THE ERROR

Number of particles	1	10	30	100	1000
RMSE (m)	105,62	96,31	75,73	70,37	66,59
Computation time (s)	1,29	2,61	5,64	15,59	150,52

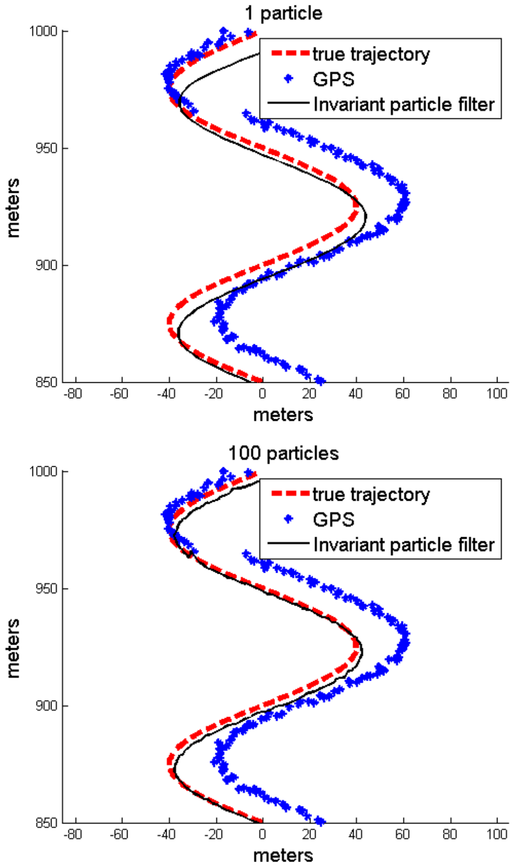
G. Results

Algorithm 1 has been implemented with the parameters given in table I. The initial state is known. Then the estimate deviates from the true trajectory due to gyroscope, accelerometer and GPS noise. As a small set of parameters is sampled (the components of the GPS bias) using moreover an optimal sampling the filter gives reasonable results even with one particle. The accuracy of the localization is measured using the total Root Mean Square Error (RMSE) of the position over the whole trajectory. Its performance improves as the number of particles grows but stabilizes fast (see table IV-G) as the sampled variable is only 3-dimensional. The computation time is also displayed in table IV-G, for a naive Matlab implementation using "for" loops and it can be seen that the cost of 10 particles is only around twice the cost of one particle. Considering Algorithm 1 in detail the reader could even expect a much smaller difference between the cost of 1 and 10 particles but the result we obtain is affected by Matlab's poor ability to deal with loops over the particles. The end of the trajectory estimated by the Invariant Particle Filter is drawn on Figure IV-G and compared to the true one. The apparent gap in the GPS observations is due to a bias change. We see that the filter is able to manage the non-Gausian noise model of the GPS with an accuracy improved as the number of particles increases.

V. CONCLUSION

This paper has introduced the idea of using some geometric properties of a system allowing to sample from a lower dimensional distribution in particle filtering. A novel property of invariant observers has also been derived, and directly applied to build a class of system for which the combination of invariant and particle filtering brings decisive properties in terms of computational burden. The realism of the proposed model has been proved by the implementation of the method on a very concrete non-linear localization problem with non-Gaussian and non-white noise. The strong convergence properties of the Invariant Extended Kalman

Fig. 1. Zoom on the end of the trajectory estimated by the Invariant Particle Filter with 1 and 100 particles, illustrating the usefulness of sampling. The beginning is not informative as the the inertia gives an accurate estimation on the short run in both cases. As an accurate near-optimal sampling is performed, a unique particle provides already a reasonable estimate. Sampling 100 particles improves the performance and, due to the properties of invariant filtering, is not much more expansive



Filter can be expected to have interesting consequences on the accuracy of the obtained sampling. No result in this direction is given here but the authors intent to address this issue in the future.

APPENDIX I
EXPONENTIAL MAP IN THE GROUP OF DOUBLE HOMOGENEOUS MATRICES

The exponential of a matrix of the form:

$$s = \begin{pmatrix} (\zeta)_\times & u_1 & u_2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

can be computed by the following closed formula:

$$\exp(s) = I_5 + s + \frac{1}{|\zeta|^2}(1 - \cos|\zeta|)s^2 + \frac{1}{|\zeta|^3}(|\zeta| - \sin|\zeta|)s^3$$

REFERENCES

[1] P-A Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[2] A. Barrau and S. Bonnabel. Intrinsic filtering on lie groups with applications to attitude estimation. *arXiv preprint arXiv:1310.2539*, 2013. conditionally accepted in IEEE Transactions on Automatic Control.

[3] A. Barrau and S. Bonnabel. Intrinsic filtering on SO(3) with discrete-time observations. In *IEEE Conference on Decision and Control*, 2013.

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

[5] 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.

[6] S. Bonnabel, P. Martin, and P. Rouchon. Symmetry-preserving observers. *IEEE Trans. Automat. Control*, 2009.

[7] S. Bonnabel, Ph. Martin, and P. Rouchon. A non-linear symmetry-preserving observer for velocity-aided inertial navigation. In *American Control Conference (ACC06)*, pages 2910–2914, June 2006.

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

[9] 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.

[10] S. Brás, P. Rosa, C. Silvestre, and P. Oliveira. Global attitude and gyro bias estimation based on set-valued observers. *Systems & Control Letters*, 62(10):937–942, 2013.

[11] D.L. Burkholder, É. Pardoux, A.S. Sznitman, and P.L. Hennequin. *École d'été de probabilités de Saint-Flour XIX, 1989*. Springer-Verlag, 1991.

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

[13] John L Crassidis. Sigma-point kalman filtering for integrated gps and inertial navigation. *Aerospace and Electronic Systems, IEEE Transactions on*, 42(2):750–756, 2006.

[14] A. Doucet. *Sequential monte carlo methods*. Wiley Online Library, 2001.

[15] A. Doucet, N. De Freitas, K. Murphy, and S. Russell. Rao-blackwellised particle filtering for dynamic bayesian networks. In *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, pages 176–183. Morgan Kaufmann Publishers Inc., 2000.

[16] A. Doucet, S. Godsill, and Ch. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and computing*, 10(3):197–208, 2000.

[17] T.E. Duncan. Some filtering results in Riemann manifolds. *Information and Control*, 35(3):182–195, 1977.

[18] Ch. Hide, T. Moore, and M. Smith. Adaptive kalman filtering for low-cost ins/gps. *The Journal of Navigation*, 56(01):143–152, 2003.

[19] S.J. Julier and J.K. Uhlmann. A new extension of the Kalman Filter to Nonlinear Systems. *Proc. of AeroSense : The 11th Int. Symp. on Aerospace/Defence Sensing, Simulation and Controls*, 1997.

[20] R. Kalman and R. Bucy. New results in linear filtering and prediction theory. *Basic Eng., Trans. ASME, Ser. D,*, 83(3):95–108, 1961.

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

[22] R. Mahony, T. Hamel, and J-M Pflimlin. Nonlinear complementary filters on the special orthogonal group. *IEEE-Trans. on Automatic Control*, 53(5):1203–1218, 2008.

[23] 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.

[24] M. Montemerlo, S. Thrun, D. Koller, B. Wegbreit, et al. Fastslam: A factored solution to the simultaneous localization and mapping problem. In *AAAI/IAAI*, pages 593–598, 2002.

[25] M Simandl and O. Straka. Sampling densities of particle filter: a survey and comparison. In *American Control Conference, 2007. ACC'07*, pages 4437–4442. IEEE, 2007.

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

[27] M. Zamani, J. Trumpf, and R. Mahony. Minimum-energy filtering on the unit circle. In *Australian Control Conference (AUCC), 2011*, pages 236–241. IEEE, 2011.