Invariant filtering for Pose EKF-SLAM aided by an IMU

Axel Barrau, Silvère Bonnabel

***Abstract*— This paper presents a new method to combine measurements from motion sensors (wheel encoders or inertial measurements) with relative pose estimates inferred from scan matching of dense depth images taken at different time instants. These can optionally be completed with measurements of relative position to landmarks with known location. The fusion is performed through an invariant extended Kalman filter (IEKF) with augmented state, as an application of the stochastic cloning method. This results in a fusion algorithm allowing to include IMU information in pose EKF-SLAM and which inherits the desirable properties of IEKFs, that is, although the system is non-linear, the evolution of the error is independent of the trajectory, as in the linear case.**

I. Introduction

To control a moving robot, or to make it fully autonomous, it is important to accurately estimate its position and its orientation, that is, its pose. Ego-localization is usually performed by fusing information from proprioceptive motion sensors such as wheel encoders for ground vehicles or inertial measurement units (IMU) that consist of accelerometers and gyrometers, with information from other exteroceptive sensors such as GPS, cameras, and laser telemeters (LIDARS) for instance. When scans of the environment yield dense clouds of 3D points, as is the case with the Microsoft Kinect sensor of the Velodyne, the clouds obtained at different time instants can be matched to infer a relative displacement (pose) between the poses at those instants. The matching is typically handled by an iterative closest point algorithm (ICP) (see e.g., [21]). In turn, in order to feed an extended Kalman filter (EKF) or a particle filter or a maximum likelihood trajectory estimator, as in the smoothing problem, one must assess a probability distribution, or a covariance matrix, to the error on the relative pose returned by the ICP. Various papers have addressed this issue, see e.g. [10] and the references therein, as well as our more recent work [2].

Navigating with proprioceptive sensors and relative state measurements has also been the focus of various works (see e.g. [20] and the references therein), and the most recent approaches treat it as a variant of a simultaneous localization and mapping problem (SLAM) where one seeks to estimate only the robot trajectory and the environment is only used to provide relative poses measurements e.g. [16]. Pose SLAM is a sub-field of Graph SLAM [24] which is a smoothing algorithm in essence as it consists in seeking the maximum likelihood trajectory given measurements over a fixed time window. Each relative pose measurement is viewed as a (small) loop closure and adds a new constraint in the graph. Incoroporating inertial measurements and building the solution iteratively has recently been proposed in [17]. Figure 1 presents the problem of data fusion with multirates and relative state measurements as a factor graph.

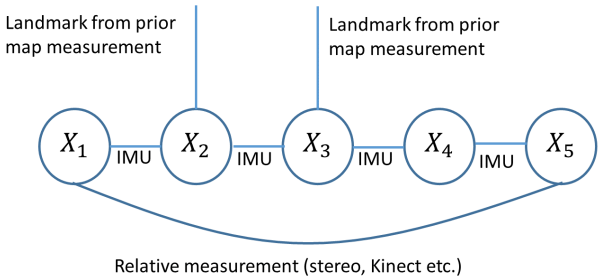


Fig. 1. Factor graph representation of the measurements relating the states (nodes). The IMU measurements relate any two consecutive states, whereas from time to time landmarks with known locations are observed, and matching images obtained at different time instants allows to relate some non-consecutive states.

In this paper we consider the problem of real-time localization, that is, the filtering and not the smoothing problem, for a robot equipped with proprioceptive sensors and a depth image sensor device, like the Kinect sensor. We devise an IEKF for the problem on the one hand for a mobile robot equipped with wheel encoders and on the other for a small Unmanned Aerial Vehicle (UAV) equipped with an IMU (see e.g. [15], [14] for an introduction to non-linear attitude estimation using IMUs). Due to a lack of space, we follow the methodology of [4] to the letter, and the reader is referred to this paper for more details. This yields a sensible fusion algorithm, as it is based on Kalman equations that explicitly account for the noises in the sensors (it is an EKF in essence), which inherits all the desirable properties of the IEKFs [8], [7], notably the nice autonomous structure of the error system, that reminds of the linear case albeit in the non-linear case, and potentially deterministic local convergence properties that conventional EKFs lack, and that stem from [4]. Note the advantages over the usual EKF have also notably been experimentally illustrated in [11], [1], [19], [3]. The latter results indicate the proposed IEKF for the considered pose SLAM problem should be very well behaved in practice. More results on invariant observers on Lie groups can be found in, e.g., [22], [25], [12], [9], [18], [23].

The contributions and organization of the present paper are as follows. In Section II, we first recap how covariance matrices associated to scan matching based pose estimates are derived, and then we review the stochastic cloning (SC)

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

EKF algorithm of [20]. In Section III we derive a novel (SC) IEKF for the fusion of wheel encoders and relative poses from depth images for a ground mobile robot. This is an extension of [1] which only considers the case of fusion with scans from a *known* 3D-map. In Section IV we derive a novel EKF, namely, a SC-IEKF, for fusion of IMU outputs with relative poses from depth images taken at different instants. Optionally, landmarks with known location in a prior map are also measured. Beyond the fact it provides a novel and simple algorithm in GPS-denied environments, the non-linear structure of the algorithm respects the underlying geometry of the state space, and possesses several striking properties. The main contributions of this paper are 1- to prove the pose SLAM problem with relative state measurements lends itself to invariant Kalman filtering through the trick of equation (2), and stochastic cloning, and 2- to derive a technogically meaningful non-linear filtering algorithm for fusion of IMU and depth images with nice mathematical properties.

II. Relative pose measurements

A. Covariance of relative displacement measurements

At time t_n, a snapshot depth image of the environment is taken (for instance using a Kinect sensor). A few seconds later, at time t_{n+1}, another depth image is taken. The comparison between those two images allows to compute, up to some errors, the relative pose between both images. Such relative pose is usually computed by a scan matching algorithm, the most popular being the Iterative Closest Point (ICP) algorithm, that computes a relative pose from dense depth scans by alternating an optimal matching step between both clouds for fixed relative pose, and an optimal relative pose calculation for fixed matching, iteratively, until convergence. Alternatively, the relative measurements can be performed through a stereo camera vision system, but this is not the focus of the paper.

The point-to-plane ICP algorithm works as follows. Consider two scans $\{p_i\}_{1\leq i\leq N}$ and $\{q_i\}_{1\leq i\leq M}$ obtained by scanning the environnement from different poses. Eeach point $p_i\in\mathbb{R}^3$ corresponds to a point of the environnement measured in the camera frame through the depth camera sensor. To compute the relative pose between both scans, one must find the right matching between the points (that is, which q_j in the second scan corresponds to the point p_i of the first scan) and one must also find the transfomation that maps the first scan to the second scan. When both scans are not too far from each other, the point-to-plane ICP iterates the following steps until convergence *a)* match each p_i of the first scan with its closest point q_j of the second scan *b)* relabel q_i the point q_j that was matched with p_i *c)* find the rotation R and translation x minimizing the cost $J(R,x)=\sum_i((q_i-Rp_i-x)^Tn_i)^2$ where n_i denotes the normal vector to the scanned surface at p_i *d)* transform the whole scan through $R\{p_i\}_{1\leq i\leq N}+x$ and for each i let p_i denote the transformed point Rp_i+x.

Of course even if both clouds are close to each other initially, so that there is a large overlap and the final matching is correct, the returned transformation (R,x) is not the exact relative pose between both scans due to sensor noise. There has been several attempts to assess a covariance matrix N to the obtained pose (see e.g. [5]). The methods are based on the following heuristic approach, mathematically justified for point-to-plane ICP in [2]: fix the matching between both clouds at its final convergence value. First consider the homogenous representation of $SE(3)$ by letting $X=\begin{pmatrix}R & x\\ 0_{1,3} & 1\end{pmatrix}$ denote the relative pose. Then, using the exponential map of $SE(3)$ (see the Appendix), compute $J(\exp(\xi)X)$ up to third order terms in $\xi\in\mathbb{R}^6$. The first order terms cancel as X is the argmin of J, and the second order terms (the Hessian) yield the covariance N through $N=\left(\frac{\partial^2 J}{\partial\xi^2}\right)^{-1}\frac{\partial^2 J}{\partial z\partial\xi}\mathrm{cov}(z)\frac{\partial^2 J}{\partial z\partial\xi}^T\left(\frac{\partial^2 J}{\partial\xi^2}\right)^{-1}$ [10] where z denotes the data. Let $v_n\in\mathbb{R}^6$ denote the noise in the Lie algebra $\mathfrak{se}(3)$ with covariance N. If we let $\{p_i\}$ be the cloud obtained by scanning the environnement at time t_n and $\{q_i\}$ the cloud obtained by scanning the environnement at time t_{n+1}, then the ICP returns the noisy relative pose (see [1] for more details):

$$E_{t_{n+1}}=V_{n+1}X_{t_{n+1}}^{-1}X_{t_n},\quad V_n=\exp_{\mathfrak{se}(3)}(v_{n+1}),\quad \mathrm{cov}(v_n)=N \tag{1}$$

B. Stochastic cloning for relative state measurements

The relative pose measurements taken from a depth sensor mounted on a vehicle relate the vehicle's state at different time instants (see Fig. 1). Such measurements go beyond the framework of standand Kalman filtering. To be handled properly in a probabilistic framework, the state needs to be duplicated at each time a depth image is taken. The method is called stochastic cloning (SC) and results in the so-called SC-KF as follows [20]. Suppose at time t_n when a first cloud is scanned the state is X_{t_n} with covariance matrix P_{t_n}. For the sake of conciseness assume linear evolution equations $\frac{d}{dt}X_t=A_tX_t+G_tw_t$ and relative measurements of the linear form $X_{t_n}-X_{t_{n+1}}$. At time t_n duplicate the state to form the augmented state vector

$$\check{X}_{t_n}=\begin{pmatrix}X_{t_n} & X_{t_n}\end{pmatrix}^T,\quad \check{P}_{t_n}=\begin{pmatrix}P_{t_n} & P_{t_n}\\ P_{t_n} & P_{t_n}\end{pmatrix}$$

Propagate by letting only the second variable of $\check{X}_t$ evolve. This results in a linear evolution equation with $\check{A}_t=\begin{pmatrix}0 & 0\\ 0 & A_t\end{pmatrix}$ and $\check{G}_t=\begin{pmatrix}0 & G_t\end{pmatrix}^T$. The corresponding augmented state estimate and error covariance matrix can be propagated through the usual Kalman equations in the augmented state space. At time t_{n+1}, the estimate $\check{X}_{t_{n+1}}$ is updated using standard Kalman equations (see e.g., eq (3) below) with $H=\begin{pmatrix}-I & I\end{pmatrix}$. This updates both X_{t_n} and $X_{t_{n+1}}$ into $X_{t_n}^+$ and $X_{t_{n+1}}^+$, which allows to retain the full correlation between both. Then, reduce the state to its second variable $X_{t_{n+1}}^+$ along with its covariance, i.e., the corresponding submatrix of $\check{P}_{t_{n+1}}$, and let it evolve through the non-duplicated equation. This must be done until the next scan measurement is available, where stochastic cloning will be needed again.

The paper [20] also investigates the effect of correlations between successive scans of the same environment. The IEKF approach being already non-trivial we will avoid this problem here by assuming each time an image is matched to the previous image, this image is discarded and not re-used for further matching.

III. IEKF Pose SLAM for wheeled robots

The robot is assumed to be equipped with a depth images sensor, such as a Kinect. At time t_n, a snapshot depth image of the environment is taken. A few seconds later, at time t_{n+1}, another depth image is taken. The comparison between those two images allow to compute up to some errors, the relative displacement between the two images. This provides a partial information about the state. The goal of the pose SLAM problem is to compute an approximation of the distribution of X_t conditioned on the past observations of the form (1). Note that, as a charateristics of SLAM problems, the pose X_t is *not* observable as only relative poses are measured. This inevitably induces a drift in the covariance of the computed trajectory (unless a loop closure is detected, but we do not consider this case in the present navigation oriented paper).

A. High-rate measurements: continuous-time dynamics

We consider a wheeled robot equipped with wheel encoders evolving on an approximatively flat ground (the so-called 2.5D case). Its pose can be parameterized by $X_t \in SE(3)$, that is, the transformation from the body frame to the fixed frame. Due to the fact the wheel encoders' measurements come at high frequency, the equations can be written in continuous time as follows. If we let $P = (x,y,z)^T$ denote the position in 3D of the robot, and $R \in SO(3)$ its orientation and we let $X_t = \begin{pmatrix} R_t & P_t \\ 0_{1\times 3} & 1 \end{pmatrix}$ (see the Appendix for more details), the kinematics equations write (e.g. [13]):

$$\frac{d}{dt}X_t = X_t(\Omega_t + w_t), \quad \Omega_t + w_t = \begin{pmatrix} (\omega + w_t^R)_\times & ve_1 + w_t^p \\ 0_{1,3} & 0 \end{pmatrix}$$

where v is the linear velocity measured by wheel encoders, u is a function of the steering angle (or alternatively the yaw angular velocity can be obtained by difference between wheel encoders), $\omega = (0,0,uv)^T$, or alternatively ω is measured by a gyroscope, where $e_1 = (1,0,0)^T$ is the heading direction of the robot (approximate rolling without slip is assumed) and where 1- the angular noise w_t^R accounts for the yaw angular velocity uncertainties, and also for the fact the ground may not be totally flat and 2- w_t^p is the velocity noise in the body frame accounting for longitudinal and lateral slip.

B. The complete model in matrix form

To address the problem above, we suppose there is a first depth image available at t_n and we consider the augmented state whose only the second part is let evolve until time t_{n+1} where another depth image is obtained and matched with the first one

$$\chi_t = \begin{pmatrix} X_{t_n} \\ X_t \end{pmatrix}, \quad t_n \le t \le t_{n+1}$$

which leads to the evolution equation

$$\frac{d}{dt}\chi_t = \begin{pmatrix} 0 \\ X_t(\Omega_t + w_t) \end{pmatrix}$$

The right invariant error is thus defined as

$$e_t = \begin{pmatrix} \eta_{t_n} \\ \eta_t \end{pmatrix} := \begin{pmatrix} \hat{X}_{t_n} X_{t_n}^{-1} \\ \hat{X}_t X_t^{-1} \end{pmatrix}$$

The outputs are compatible with this error in the sense of [6] as we can calculate the following modified measurement errors based on the ICP outputs (1)

$$\hat{X}_{t_{n+1}} E_{t_{n+1}} \hat{X}_{t_n}^{-1} = \hat{V}_{t_n} \eta_{t_{n+1}} \eta_{t_n}^{-1} \tag{2}$$

where $\hat{V}_t = \hat{X}_t V_t \hat{X}_t^{-1}$.

C. Invariant extended Kalman filter algorithm

We now follow the methodology of [4] to the letter. During the propagation step the RIEKF equations write

$$\frac{d}{dt}\hat{\chi}_t = \begin{pmatrix} 0 \\ \hat{X}_t \Omega_t \end{pmatrix}, \quad t_n \le t \le t_{n+1}$$

with update step based on measurements at times t_n, t_{n+1}

$$\hat{\chi}_{t_{n+1}}^+ = \begin{pmatrix} \exp[(L_{n+1}\exp^{-1}(\hat{X}_{t_{n+1}} E_{t_{n+1}} \hat{X}_{t_n}^{-1}))_{1:3}]\hat{X}_{t_n} \\ \exp[(L_{n+1}\exp^{-1}(\hat{X}_{t_{n+1}} E_{t_{n+1}} \hat{X}_{t_{n+1}}^{-1}))_{4:6}]\hat{X}_{t_{n+1}} \end{pmatrix}$$

where the exponential map of $SE(3)$ is defined in Appendix I-A. The right-invariant error evolution thus reads

$$\frac{d}{dt}e_t = \begin{pmatrix} 0 \\ -\hat{w}_t \eta_t \end{pmatrix}$$

where $\hat{w}_t = \hat{X}_t w_t \hat{X}_t^{-1}$and the update step yields

$$e_{t_{n+1}}^+ = \begin{pmatrix} \exp[(L_{n+1}\exp^{-1}(\hat{V}_{t_n}\eta_{t_{n+1}}\eta_{t_n}^{-1})_{4:6}]\eta_{t_n} \\ \exp[(L_{n+1}\exp^{-1}(\hat{V}_{t_n}\eta_{t_{n+1}}\eta_{t_n}^{-1})_{1:3}]\eta_{t_{n+1}} \end{pmatrix}$$

To linearize this equation we introduce the linearized error ξ_t defined by $\eta_t = I_4 + \mathscr{L}_{\mathfrak{se}(3)}(\xi_t)$ and the linearized observation noise defined as $V_n = I_4 + \mathscr{L}_{\mathfrak{se}(3)}(v_n)$. Introducing $\eta_t = I_4 + \mathscr{L}_{\mathfrak{se}(3)}(\xi_t)$, $\eta_t^+ = I_4 + \mathscr{L}_{\mathfrak{se}(3)}(\xi_t^+)$, $\eta_t^{-1} = I_4 - \mathscr{L}_{\mathfrak{se}(3)}(\xi_t)$, and $\exp^{-1}(ab) = \exp^{-1}(a) + \exp^{-1}(b)$ in the latter error equations and removing the terms of second order $O(\|\xi\|^2, \|w\|^2, \|v_n\|^2, \|\xi\|\|w\|, \|\xi\|\|v_n\|)$ we obtain:

$$\frac{d}{dt}\xi_t = -\hat{w}_t, \quad \begin{pmatrix} \xi_{t_n}^+ \\ \xi_{t_{n+1}}^+ \end{pmatrix} = \begin{pmatrix} \xi_{t_n} \\ \xi_{t_{n+1}} \end{pmatrix} - L_n(\xi_{t_n} - \xi_{t_{n+1}} - v_n)$$

The gains L_{n+1} are computed by initialization

$$\check{P}_{t_n} = \begin{pmatrix} P_{t_n} & P_{t_n} \\ P_{t_n} & P_{t_n} \end{pmatrix}, \quad \text{where} \quad \text{cov}(\xi_{t_n}) = P_{t_n}$$

and Riccati evolution

$$\begin{aligned} &\frac{d}{dt}\check{P}_t = \check{A}\check{P}_t + \check{P}_t\check{A}^T + \check{Q}_t, \quad S_{n+1} = H\check{P}_{t_{n+1}}H^T + \hat{N}_{n+1}, \\ &L_{n+1} = \check{P}_{t_{n+1}}H^T S^{-1}, \quad \check{P}_{t_{n+1}}^+ = (I - L_{n+1}H)\check{P}_{t_{n+1}} \end{aligned} \tag{3}$$

where the matrices $A, H, \check{Q}$ and $\hat{N}$ defined by:

$$A = 0_{6,6},\ H = \begin{pmatrix} -I_6 & I_6 \end{pmatrix},\ \check{Q}_t = \begin{pmatrix} 0 & 0 \\ 0 & \text{cov}(\hat{w}_t) \end{pmatrix}$$
$$\hat{N}_{n+1} = \text{cov}(\hat{v}_{n+1})$$

Until the next measurement time t_{n+2}, the estimated state boils down to $\hat{X}_t$ with covariance the bottom right 3×3 submatrix of the augmented state covariance matrix $\check{P}^+_{t_{n+1}}$.

D. Properties

As predictable from the theory of the IEKF, the obtained extended Kalman filter has striking properties, that stem from its non-linear structure anchored in the geometry of $SE(3)$.

Proposition 1: The pair (A,H) appearing in the Riccati equation and encoding the linearized error deterministic system, does *not* depend on the trajectory, and A is null, as in the *linear static* problem, although the problem at hand is non-linear, and time-varying as the robot commands (u,v) are totally free to take any values.

The insensitivity of the pair (A,H) to the estimate $\hat{X}_t$ has proved to lead to increased robustness of the filter, as demonstrated theoretically in [4] and experimentally in [11], [1].

E. Simulations

To briefly illustrate numerically what the filter does, we simulated a robot in a square 2D environment, and we implemented the proposed filter. From time to time the robot is assumed to observe three landmarks with known location. Note that, the corresponding update is not described in Section III-C due to space limitations, see [4]. We observed the filter manages to estimate correctly its pose. Figure 2 displays the 99% envelope computed by the filter for the heading error (red line), and the true heading error (blue line).

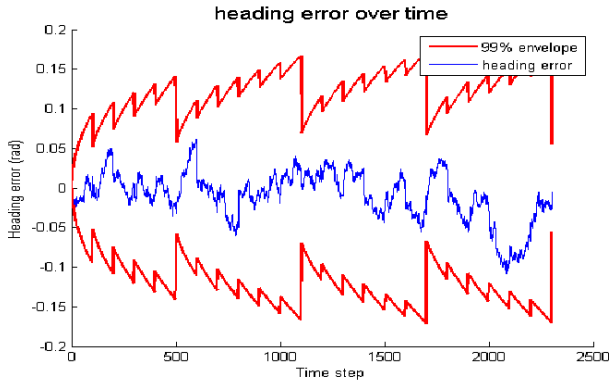
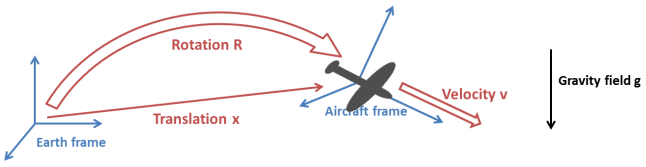


Fig. 2. 99% envelope computed by the filter for the heading error (red line), and the true heading error (blue line). The variance is reduced at each relative displacement observation but overall grows until landmarks with known locations are observed.

The figure illustrates the effect of the various discrete measurements on the filter's covariance, which is as expected. Indeed the variance is reduced at each relative displacement observation but overall grows until landmarks with known locations are observed. This is because the relative displacements measurements do not give an absolute information about the heading, so they can only limit the drift (as is typical of SLAM applications). On the other hand landmarks observations (for instance at time step 500) correspond to an absolute information in the fixed frame, it is thus logical that they yield a much larger decrease in the heading error variance as observed. Note also that the true heading error remains in the 99% envelope, which indicates consistency of the filter.

IV. IEKF POSE SLAM FOR UAVS



We consider here the more complicated model of an UAV evolving in the 3D space and whose state is parameterized by its attitude R_t, velocity v_t and position x_t. The vehicle is equipped with an IMU (accelerometers and gyroscopes) whose measurements are denoted respectively by u_t and ω_t. The continous time dynamics read:

$$\begin{aligned} &\frac{d}{dt}x_t = v_t \quad \frac{d}{dt}v_t = g + R_t(u_t + w^u_t) \\ &\frac{d}{dt}R_t = R_t(\omega_t + w^\omega_t)_\times \end{aligned} \tag{4}$$

Various kinds of observations can be considered, as illustrated on Fig. 1, but we here focus on the following:

$$(V^R_{n+1}R^T_{t_{n+1}}R_{t_n}, R^T_t(x_{t_n} - x_{t_{n+1}} - V^x_{n+1})) \tag{5}$$

where $V^R_{n+1}\in SO(3)$ and $V^x_{n+1}\in\mathbb{R}^3$ are noises. (5) is the measurement of relative pose computed by matching two depth images taken respectively at times t_n and t_{n+1}, corresponding to the measurement (1).

A. The complete model in matrix form

As explained in [4], the system (4) can be embedded in the group of double homogeneous matrices (see Appendix I-B) using the matrices X_t, w_t and function $f_{\omega,u}$:

$$\begin{aligned} X_t &= \begin{pmatrix} R_t & v_t & x_t \\ 0_{1,3} & 1 & 0 \\ 0_{1,3} & 0 & 1 \end{pmatrix}, \quad w_t = \begin{pmatrix} (w^\omega_t)_\times & w^f_t & 0_{3,1} \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{pmatrix}, \\ f_{\omega,u} &: \begin{pmatrix} R & v & x \\ 0_{1,3} & 1 & 0 \\ 0_{1,3} & 0 & 1 \end{pmatrix} \to \begin{pmatrix} R(\omega)_\times & g+Ru & v \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{pmatrix} \end{aligned} \tag{6}$$

The dynamics then become:

$$\frac{d}{dt}X_t = f_{\omega_t,u_t}(X_t) + X_t w_t \tag{7}$$

and the observations (5) are adapted as follows. Let

$$M(X_t) = \begin{pmatrix} R_t & 0 & x_t \\ 0_{1,3} & 1 & 0 \\ 0_{1,3} & 0 & 1 \end{pmatrix}$$

be the operator that puts the velocity to 0. Embedding the measurement space into the matrix representation of $SE_2(3)$ the relative pose measurements can thus be written

$$E_{t_{n+1}} = V_{n+1}M(X^{-1}_{t_{n+1}}X_{t_n}) \tag{8}$$

B. Invariant extended Kalman filter algorithm

1) Propagation step : To account for relative pose measurements, we clone the state as follows at time t_n to get for $t_n \leq t \leq t_{n+1}$

$$\chi_t = \begin{pmatrix} X_{t_n} \\ X_t \end{pmatrix}$$

and its evolution is given by

$$\frac{d}{dt}\chi_t = \begin{pmatrix} 0 \\ f_{\omega_t,u_t}(X_t) + X_t w_t \end{pmatrix} \tag{9}$$

The RIEKF for the system (9) merely reads during the propagation step:

$$\frac{d}{dt}\hat{\chi}_t = \begin{pmatrix} 0 \\ f_{\omega_t,u_t}(\hat{X}_t) \end{pmatrix} \tag{10}$$

Letting the (right-invariant) error be $\eta_t = \hat{X}_t X_t^{-1}$ and the corresponding invariant state error be for $t_n \leq t \leq t_{n+1}$

$$e_t = \begin{pmatrix} \eta_{t_n} \\ \eta_t \end{pmatrix} = \begin{pmatrix} \hat{X}_{t_n} X_{t_n}^{-1} \\ \hat{X}_t X_t^{-1} \end{pmatrix}$$

Its evolution during the propagation step (i.e. between two updates) reads

$$\frac{d}{dt}\begin{pmatrix} \eta_{t_n} \\ \eta_t \end{pmatrix} = \begin{pmatrix} 0 \\ g_{\omega_t,u_t}(\eta_t) - (\hat{X}_t w_t \hat{X}^{-1})\eta_t \end{pmatrix} \tag{11}$$

where the function g is defined by

$$g_{\omega,u} = \eta \to f_{\omega,u}(\eta) - \eta f_{\omega,u}(I_5)$$

as a direct application of the results of [4]. To linearize the equation verified by η_t we introduce the linearized error ξ_t defined by $\eta_t = I_3 + \mathscr{L}_{\mathfrak{se}(3),2}(\xi_t)$. Introducing $\eta_t = I_3 + \mathscr{L}_{\mathfrak{se}(3),2}(\xi_t)$ and $\eta_t^{-1} = I_3 - \mathscr{L}_{\mathfrak{se}(3),2}(\xi_t)$ in (11) and removing the terms of second order $O(\|\xi\|^2, \|w\|^2, \|\xi\|\|w\|)$ we get

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

This allows to define for the linearized system matrix

$$A = \begin{pmatrix} 0_{3,3} & 0_{3,3} & 0_{3,3} \\ (g)_\times & 0_{3,3} & 0_{3,3} \\ 0_{3,3} & I_3 & 0_{3,3} \end{pmatrix},$$

and the state noise covariance matrix

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

2) Update step: From the measurement $E_{t_{n+1}}$, obtained at time t_{n+1} one can compute

$$\hat{X}_{t_{n+1}} E_{t_{n+1}} \hat{X}_{t_n}^{-1} = M(\hat{X}_{t_{n+1}} V_{n+1} X_{t_{n+1}}^{-1} X_{t_n} \hat{X}_{t_n}^{-1}) = \hat{V}_{n+1} M(\eta_{t_{n+1}} \eta_{t_n}^{-1})$$

with $\hat{V}_{n+1} = \hat{X}_{t_{n+1}} V_{n+1} \hat{X}_{t_{n+1}}^{-1}$. The corresponding update for the IEKF writes

$$\hat{\chi}_{t_{n+1}}^+ = \begin{pmatrix} \exp[(L_n \exp^{-1}(\hat{X}_{t_{n+1}} E_{t_{n+1}} \hat{X}_{t_n}^{-1})_{1:9}]\hat{X}_{t_n} \\ \exp[(L_n \exp^{-1}(\hat{X}_{t_{n+1}} E_{t_{n+1}} \hat{X}_{t_n}^{-1})_{10:18}]\hat{X}_{t_{n+1}} \end{pmatrix} \tag{12}$$

where exp denotes the exponential map of the group $SE_2(3)$ as defined in Appendix I-B, and where $\exp^{-1}$ denotes its inverse which must be computed numerically, and the gain L_n is to be defined. The corresponding error update reads

$$e_{t_{n+1}}^+ = \begin{pmatrix} \exp[(L_n \exp^{-1}(\hat{V}_{n+1} M(\eta_{t_{n+1}} \eta_{t_n}^{-1}))_{1:9}]\eta_{t_n} \\ \exp[(L_n \exp^{-1}(\hat{V}_n M(\hat{V}_{n+1} M(\eta_{t_{n+1}} \eta_{t_n}^{-1}))_{10:18}]\eta_{t_{n+1}} \end{pmatrix}$$

This implies for the first order approximation of the error as defined above

$$\begin{pmatrix} \xi_{t_n}^+ \\ \xi_{t_{n+1}}^+ \end{pmatrix} = \begin{pmatrix} \xi_{t_n} \\ \xi_{t_{n+1}} \end{pmatrix} - L_{n+1}(\xi_{t_n} - \xi_{t_{n+1}} - v_{n+1})$$

3) IEKF final equations: Gathering all the equations obtained we get the following filter. During the propagation step between two consecutive depth images obtained at times t_n and t_{n+1} yielding a relative pose measurements (5), the IEKF is defined by the augmented state equation (10) and the corresponding covariance Riccati equation writes $\frac{d}{dt}\check{P}_t = \check{A}\check{P}_t + \check{P}_t\check{A}^T + \check{Q}_t$ with initial condition

$$\check{P}_{t_n} = \begin{pmatrix} P_{t_n} & P_{t_n} \\ P_{t_n} & P_{t_n} \end{pmatrix}, \quad \text{where} \quad \text{cov}(\xi_{t_n}) = P_{t_n}$$

and

$$\check{A} = \begin{pmatrix} \tilde{0} & 0 \\ 0 & A \end{pmatrix}, \text{ and } \check{Q}_t = \begin{pmatrix} 0 & 0 \\ 0 & \hat{Q}_t \end{pmatrix},$$

The update step is based on (12) with the gain defined by $L_{n+1} = \check{P}_{t_{n+1}} H^T S_{n+1}^{-1}$ with

$$S_{n+1} = H\check{P}_{t_{n+1}} H^T + \hat{N}_{n+1}, \quad \check{P}_{t_{n+1}}^+ = (I - L_{n+1}H)\check{P}_{t_{n+1}}$$

and where we let $H = \begin{pmatrix} -I_9 & I_9 \end{pmatrix}$, $\hat{N}_{n+1} = \text{cov}(\hat{v}_{n+1})$.

C. Properties

Note that, the relative state measurements only do not make the pose observable, but this is a usual feature of SLAM problems, and we just want the filter to reflect it.

Proposition 2: The IEKF which incorporates both types of measurements is such that the pair (A, H) which encodes the behavior of the deterministic error system does not depend on the trajectory, nor on the time, as in the *linear time-invariant* case, although the system is non-linear and time-varying.

V. Conclusion

In this paper we have proposed two novel non-linear filters for pose SLAM, allowing to process relative state measurements via stochastic cloning, both simple enough, and enjoying remarkable properties. The second filter allows to incorporate IMU measurements, which are generally not considered in the SLAM literature, but which are technologically highly relevant due to the rather recent burst of low-cost IMUs that now populate cell phones, cameras etc. In future work we would like to experimentally test those filters and compare them to state-of-the art SLAM approaches. Indeed, EKFs are generally disdained by the SLAM community due to their shortcomings (inconsistency). We anticipate our geometric filters may be better behaved due to the nice structure of the error system.

APPENDIX I
SOME FACTS ON MATRIX LIE GROUPS

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

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

If $\gamma(t)$ is a process taking values in G and verifying $\gamma(0) = Id$, then its derivative at $t = 0$ cannot take abitrary values in $\mathscr{M}_n(\mathbb{R})$. It is constrained in a subspace $\mathfrak{g}$ of $\mathscr{M}_n(\mathbb{R})$ called the "Lie algebra of G". As it proves useful to identify $\mathfrak{g}$ to $\mathbb{R}^{\dim \mathfrak{g}}$, a linear mapping $\mathscr{L}_{\mathfrak{g}} : \mathbb{R}^{\dim \mathfrak{g}} \to \mathfrak{g}$ is used in this paper. The vector space $\mathfrak{g}$ can be mapped to the matrix Lie group G through the classical matrix exponential $\exp_m$. As well, $\mathbb{R}^{\dim \mathfrak{g}}$ can be mapped to G through a function exp defined by $\exp(\xi) = \exp_m(\mathscr{L}_{\mathfrak{g}}(\xi))$.

A. Group of direct spatial isometries $SE(3)$

We have here $G = SE(3) = \{\begin{pmatrix} R & x \\ 0_{1,3} & 1 \end{pmatrix}, R \in SO(3), x \in \mathbb{R}^3\}$. The Lie algebra is $\mathfrak{se}(3) = \{\begin{pmatrix} (\xi)_\times & x \\ 0_{1,3} & 1 \end{pmatrix}, \xi, x \in \mathbb{R}^3\}$. An isomorphism between $\mathbb{R}^6$ and $\mathfrak{se}(3)$ is given by $\mathscr{L}_{\mathfrak{se}(3)} \begin{pmatrix} \xi \\ x \end{pmatrix} = \begin{pmatrix} (\xi)_\times & x \\ 0_{1,3} & 0 \end{pmatrix}$. The exponential mapping is given by the closed formula:

$$\exp \begin{pmatrix} \xi \\ x \end{pmatrix} = I_4 + S + \frac{1-\cos(||\xi||)}{||\xi||^2})S^2 + \frac{||\xi|| - \sin(||\xi||)}{||\xi||^3} S^3$$

where $S = \mathscr{L}_{\mathfrak{se}(3)} \begin{pmatrix} \xi \\ x \end{pmatrix}$.

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

We have here $G = SE_2(3) = \{\begin{pmatrix} R & v & x \\ 0_{1,3} & 1 & 0 \\ 0_{1,3} & 0 & 1 \end{pmatrix}, R \in SO(3), x \in \mathbb{R}^3\}$. The Lie algebra is $\mathfrak{se}_2(3) = \{\begin{pmatrix} (\xi)_\times & u & y \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{pmatrix}, \xi, u, y \in \mathbb{R}^3\}$. An isomorphism between $\mathbb{R}^9$ and $\mathfrak{se}_2(3)$ is given by $\mathscr{L}_{\mathfrak{se}_2(3)} \begin{pmatrix} \xi \\ u \\ y \end{pmatrix} = \begin{pmatrix} (\xi)_\times & u & y \\ 0_{1,3} & 0 & 0 \\ 0_{1,3} & 0 & 0 \end{pmatrix}$. The exponential mapping is given by the formula:

$$\exp \begin{pmatrix} \xi \\ u \\ y \end{pmatrix} = I_5 + S + \frac{1-\cos(||\xi||)}{||\xi||^2})S^2 + \frac{||\xi|| - \sin(||\xi||)}{||\xi||^3} S^3$$

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

REFERENCES

[1] Martin Barczyk, Silvere Bonnabel, Jean-Emmanuel Deschaud, and François Goulette. Invariant EKF design for scan matching-aided localization. In *IEEE Trans. on Systems Control Technology, DOI:10.1109/TCST.2015.2413933. Preprint http://arxiv.org/abs/1503.01407*, 2015.

[2] Martin Barczyk, Silvère Bonnabel, and François Goulette. Observability, covariance and uncertainty of icp scan matching. *arXiv preprint arXiv:1410.7632*, 2014.

[3] 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, May 2013.

[4] Axel Barrau and Silvère Bonnabel. The invariant extended kalman filter as a stable observer. *arXiv preprint arXiv:1410.1465*, 2014.

[5] Ola Bengtsson and Albert-Jan Baerveldt. Robot localization based on scan-matching — estimating the covariance matrix for the IDC algorithm. *Robotics and Autonomous Systems*, 44(1):29–40, July 2003.

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

[7] 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.

[8] Silvere Bonnabel. Left-invariant extended kalman filter and attitude estimation. In *IEEE conference on decision and control*, pages 1027–1032, 2007.

[9] Sérgio Brás, Rita Cunha, José F Vasconcelos, Carlos Silvestre, and Paulo Oliveira. A nonlinear attitude observer based on active vision and inertial measurements. *Robotics, IEEE Transactions on*, 27(4):664–677, 2011.

[10] Andrea Censi. An accurate closed-form estimate of ICP's covariance. In *Proceedings of the 2007 IEEE International Conference on Robotics and Automation*, pages 3167–3172, Roma, Italy, April 2007.

[11] Sébastien Diemer and Silvère Bonnabel. An invariant linear quadratic gaussian controller for a simplified car. *arXiv:1406.1619*, ICRA 2015.

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

[13] Denis Guillaume and Pierre Rouchon. Observation and control of a simplified car. In *IFAC International Workshop on Motion Control*, pages 63–67, Grenoble, France, September 1998.

[14] Minh-Duc Hua, Guillaume Ducard, Tarek Hamel, and Robert Mahony. Introduction to nonlinear attitude estimation for aerial robotic systems. *Aerospace Lab*, pages AL08–04, 2014.

[15] Minh-Duc Hua, Guillaume Ducard, Tarek Hamel, Robert Mahony, and Konrad Rudin. Implementation of a nonlinear attitude estimator for aerial robotic vehicles. *Control Systems Technology, IEEE Transactions on*, 22(1):201–213, 2014.

[16] Viorela Ila, Josep M Porta, and Juan Andrade-Cetto. Information-based compact pose slam. *Robotics, IEEE Transactions on*, 26(1):78–93, 2010.

[17] Vadim Indelman, Stephen Williams, Michael Kaess, and Frank Dellaert. Factor graph based incremental smoothing in inertial navigation systems. In *Information Fusion (FUSION), 2012 15th International Conference on*, pages 2154–2161. IEEE, 2012.

[18] C. Lagemann, J. Trumpf, and R. Mahony. Gradient-Like Observers for Invariant Dynamics on a Lie group. *IEEE Trans. on Automatic Control*, 55:2:367 – 377, 2010.

[19] 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.

[20] Anastasios I Mourikis, Stergios I Roumeliotis, and Joel W Burdick. Sc-kf mobile robot localization: a stochastic cloning kalman filter for processing relative-state measurements. *Robotics, IEEE Transactions on*, 23(4):717–730, 2007.

[21] Szymon Rusinkiewicz and Marc Levoy. Efficient variants of the ICP algorithm. In *Proceedings of the Third International Conference on 3-D Digital Imaging and Modeling*, pages 145–152, Quebec City, Canada, May 2001.

[22] A. K. Sanyal, T. Lee, M. Leok, and N. H. McClamroch. Global optimal attitude estimation using uncertainty ellipsoids. *Systems & Control Letters*, 57(3):236–245, 2008.

[23] Glauco Garcia Scandaroli and Pascal Morin. Nonlinear filter design for pose and imu bias estimation. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 4524–4530. IEEE, 2011.

[24] 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.

[25] 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.