Final Report
Submarine Attitude and Pose Estimation using Invariant Extended Kalman filter on SE(3)

Nikita Rudin, McGill University, April 2018

Abstract—**This report presents the project component of the MECH 600 course at McGill University. The purpose of the project was to familiarize with the theory of the Invariant Extended Kalman Filter (IEKF). A step by step methodology method is presented.**

***Index Terms*—SE(3), Estimation, IEKF.**

I. Introduction

THE Extended Kalman filter is one of the principal tools used in the industry to solve various estimation problems. It has been developed by NASA in the 1960 for the Apollo program and is now specifically popular in control and robotics. This popularity is due to its ease of implementation and tuning. It is an extension of the linear Kalman filter obtained by linearizing the dynamics of the system and the observation function. This leads to a lack of mathematically provable convergence. In real world examples the Kalman filter can diverge if the error becomes to large.

A recent idea to improve the performance of the EKF is to use the geometrical symmetry of the system to our advantage. This leads to the so called Invariant Extended Kalman filter (IEKF). The purpose of this project is to study the IEKF and apply it to a submarine attitude and pose estimation problem.

II. Problem setting and Notation

Wed define the the state of the submarine as follows. It includes the Direction Cosine Matrix (DCM) $\mathbf{C}_{b_k a} \in SO(3)$ transforming the body frame at step k $\underrightarrow{\mathcal{F}}_{b_k}$ into a reference frame $\underrightarrow{\mathcal{F}}_{a}$, and the vector $\mathbf{r}_a^{z_k w}$ representing the position of a given point (z) of the body relative to a fixed point (w) in the inertial frame at step k. The relationship between $\underrightarrow{\mathcal{F}}_{a}, \underrightarrow{\mathcal{F}}_{b_k}$ and $\underrightarrow{\mathcal{F}}_{b_{k+1}}$ is showed in figure 1. The relationship between $\mathbf{r}_a^{z_k w}$ and $\mathbf{r}_a^{z_{k+1} w}$ is :

$$\mathbf{r}_a^{z_{k+1} w} = \mathbf{r}_a^{z_k w} + \mathbf{r}_a^{z_k z_{k+1}} \tag{1}$$

Throughout this report we will have to mix matrices, Lie groups, Lie algebras and elements of the so called "Tangent space". We will use capital letters for Lie groups, bold for matrices and vectors. Lower cases are used for the Lie algebra and the tangent space. $\exp_m$ is the matrix exponential function,

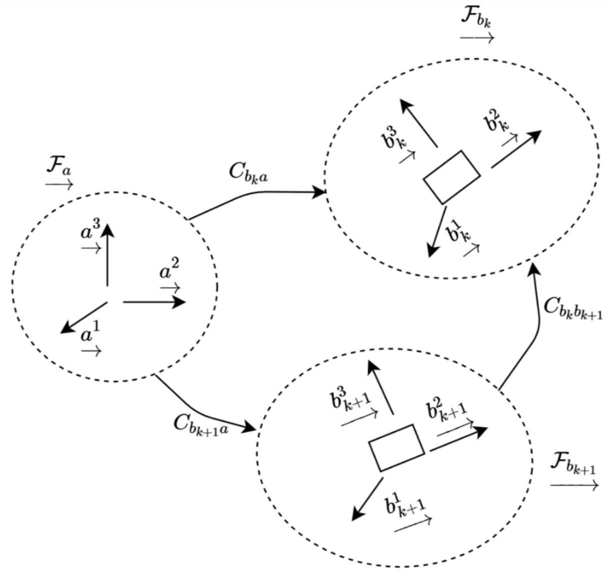

Fig. 1. Relationship between the reference frames at step k and k+1.

while exp is the exponential map for the particular Lie Group. This can be summarized by the equation below:

$$X_k = \mathbf{T}_{ab_k} = \begin{bmatrix} \mathbf{C}_{ab_k} & \mathbf{r}_a^{z_k w} \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix} = \begin{bmatrix} \exp_m(-\boldsymbol{\phi}^\times) & \mathbf{r}_a^{z_k w} \\ \mathbf{0}_{3\times 1} & \mathbf{1} \end{bmatrix}$$
$$= \exp_m(\begin{bmatrix} \phi \\ \mathbf{V}\mathbf{r}_a^{z_k w} \end{bmatrix}^\times) \equiv \exp_m(x^\times) \equiv \exp(x)$$

The notation is slightly different for the errors, η is an element of the Lie Group and ξ an element of the tangent space. Also note that in our case elements of the Lie Group are also matrices, but they obey particular constraints. To make the distinction clear X_k is not bold.

III. SE(3) kinematics

Let us denote our state as:

$$X_k = \mathbf{T}_{ab_k} = \begin{bmatrix} \mathbf{C}_{ab_k} & \mathbf{r}_a^{z_k w} \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix}$$

In order to simlify the notation let (b) represent the body frame at step k and (c) the body frame at step k+1. Let also $\mathbf{r}_a$ be the position of a fixed point(z) of the body relative to a fixed point (w) in the inertial frame and $\delta\mathbf{r}_b$ be the relative displacement between z_{k+1} and z_k resolved in the body frame. We define:

$$\mathbf{T}_{ab} = \begin{bmatrix} \mathbf{C}_{ab} & \mathbf{r}_a \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix}; \mathbf{T}_{bc} = \begin{bmatrix} \mathbf{C}_{bc} & \delta\mathbf{r}_b \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix} \quad (2)$$

We therefore have:

$$\mathbf{T}_{ab}\mathbf{T}_{bc} = \mathbf{T}_{ac} = \begin{bmatrix} \mathbf{C}_{ab}\mathbf{C}_{bc} & \mathbf{r}_a + \mathbf{C}_{ab}\delta\mathbf{r}_b \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{ac} & \mathbf{r}_a + \delta\mathbf{r}_a \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix}$$

$\mathbf{T}_{bc}$ is the transition matrix between steps k and k+1. It has the following general form:

$$\mathbf{T}_{bc} = \begin{bmatrix} \exp_m(\boldsymbol{\omega}_b^{\times} dt) & \mathbf{v}_b dt \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix}$$

where $\boldsymbol{\omega}_b, \mathbf{v}_b \in \mathbb{R}^3$ and therefore:

$$\mathbf{T}_{ac} = \begin{bmatrix} \exp_m(\boldsymbol{\omega}_b^{\times} dt) & \mathbf{v}_b dt \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{C}_{ab}\mathbf{C}_{bc} & \mathbf{r}_a + \mathbf{C}_{ab}\delta\mathbf{r}_b \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix} \quad (3)$$

We will now get an expression for ω_b and $\mathbf{v}_b$. In order to this, we will take the derivative of both (2) and (4). First let's start with (4).

We need to remember that $\mathbf{T}_{ac}$ is our state at step k+1 and $\mathbf{T}_{ab}$ is the same state at step k. Going to continuous time step k is time t and step k+1 is time t+δt. We can then take the derivative of $\mathbf{T}_{ab}$ as follows:

$$\dot{\mathbf{T}}_{ab} = \lim_{dt\to 0} \frac{\mathbf{T}_{ac} - \mathbf{T}_{ab}}{dt} = \lim_{dt\to 0} \frac{\mathbf{T}_{ab}(\mathbf{T}_{bc} - \mathbf{I})}{dt} =$$
$$\lim_{dt\to 0} \frac{\mathbf{T}_{ab}(\begin{bmatrix} \exp_m(\boldsymbol{\omega}_b^{\times} dt) & \mathbf{v}_b dt \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix} - \mathbf{I})}{dt} =$$
$$\lim_{dt\to 0} \frac{\mathbf{T}_{ab}(\begin{bmatrix} \mathbf{I} + \boldsymbol{\omega}_b^{\times} dt & \mathbf{v}_b dt \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix} - \mathbf{I})}{dt} =$$
$$\lim_{dt\to 0} \frac{\mathbf{T}_{ab}(\begin{bmatrix} \boldsymbol{\omega}_b^{\times} dt & \mathbf{v}_b dt \\ \mathbf{0}_{3\times 1} & 0 \end{bmatrix})}{dt} =$$
$$\mathbf{T}_{\mathbf{ab}}(\begin{bmatrix} \boldsymbol{\omega}_b^{\times} & \mathbf{v}_b \\ \mathbf{0}_{3\times 1} & 0 \end{bmatrix} =$$
$$\begin{bmatrix} \mathbf{C}_{ab}\mathbf{C}_{bc} & \mathbf{r}_a + \mathbf{C}_{ab}\delta\mathbf{r}_b \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{\omega}_b^{\times} & \mathbf{v}_b \\ \mathbf{0}_{3\times 1} & 0 \end{bmatrix} =$$
$$\begin{bmatrix} \mathbf{C}_{ab}\boldsymbol{\omega}_b^{\times} & \mathbf{C}_{ab}\mathbf{v}_b \\ \mathbf{0}_{3\times 1} & 0 \end{bmatrix}$$

Now we take the derivative of (4) directly by applying kinematics principles:

$$\dot{\mathbf{T}}_{ab} = \begin{bmatrix} \mathbf{C}_{ab} & \mathbf{r}_a \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix} = \begin{bmatrix} \dot{\mathbf{C}}_{ab} & \dot{\mathbf{r}}_a \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix}$$

$$\dot{\mathbf{C}}_{ab} = -\boldsymbol{\omega}_a^{ba\times}\mathbf{C}_{ab} = (-\mathbf{C}_{ab}\boldsymbol{\omega}_b^{ab})^{\times}\mathbf{C}_{ab} = -\mathbf{C}_{ab}\boldsymbol{\omega}_b^{ba\times} = \mathbf{C}_{ab}\boldsymbol{\omega}_b^{ba\times}$$

$$\dot{\mathbf{r}}_a = \mathbf{C}_{ab}(\dot{\mathbf{r}}_b + \boldsymbol{\omega}_b^{ba^{\times}}\mathbf{r}_b)$$

By equating the two results we must have:

$$\boldsymbol{\omega}_b = \boldsymbol{\omega}_b^{ba}$$
$$\mathbf{v}_b = \dot{\mathbf{r}}_b + \boldsymbol{\omega}_b^{ba\times}\mathbf{r}_b$$

IV. IEKF PRINCIPLE

The invariant Extended Kalmman filter is applied to elements of a Lie Group. It takes its name from the fact that it uses the invariance of the system under the group action. In this report, the theory will be developed for elements of $\mathbf{SE(3)}$, which is a Lie group under the matrix multiplication action. Remember that X_k is defined as $\mathbf{T}_{ab_k}$.
While the EKF uses the error defined as:

$$\eta_k = X_k - \hat{X}_k$$

where X is the correct state and $\hat{X}$ is the estimated state. The IEKF error is given by:

$$\eta_k^L = X_k^{-1}\hat{X}_k = \mathbf{T_{ab_k}}\mathbf{T_{\hat{b}_k a}}$$

This error is left-invariant to a transformation $S \in SE(3)$

$$\eta^L = (SX)^{-1}S\hat{X} = X^{-1}S^{-1}S\hat{X} = X^{-1}\hat{X}^{-1} = \eta^L$$

We will see that an other (right-invariant) error will be useful:

$$\eta_k^R = \hat{X}_k X_k^{-1}$$

this time we have:

$$\eta^R = \hat{X}S(XS)^{-1} = \hat{X}SS^{-1}X^{-1} = \hat{X}_k X_k^{-1} = \eta^R$$

Therefore it is right-invariant.
Another remarkable property is that when to SE(3) the error is itself is an element of SE(3). As such we have:

$$\eta^L = \exp_m(\boldsymbol{\xi}^{L\times}) \equiv \exp(\xi^L)$$
$$\eta^R = \exp_m(\boldsymbol{\xi}^{R\times}) \equiv \exp(\xi^R)$$

with ξin se(3) $\forall T_1, T_2 \in SE(3)$ the product $T_1 T_2$ is still an element of SE(3), while $T_1 - T_2 \notin SE(3)$ Based on this nonlinear error the IEKF methodology and its properties were derived by [**?**]. The following section is a summary of the methodology.

V. IEKF METHODOLOGY

By analogy with the EKF we need to develop a method that would give us the linearized error matrices $\mathbf{A}_k, \mathbf{L}_k, \mathbf{C}_k, \mathbf{M}_k$. First, we define the following system:

$$\hat{X}_{k+1} = f(\hat{X}_k, u_k)$$

$$X_{k+1} = f(X_k, u_k)\exp(w_k)$$

where the process noise is modeled as $w_k \in se(3), w_k \sim \mathcal{N}(0, \mathbf{Q}_k)$.

A. Propagation

1) left-invariant: The left invariant error at step k+1 becomes:

$$\eta^L_{k+1} = X^{-1}_{k+1}\hat{X}_{k+1} =$$
$$[f(X_k,u_k)\exp(w_k)]^{-1} f(\hat{X}_k,u_k) =$$
$$\exp(-w_k)[f(X_k,u_k)]^{-1} f(\hat{X}_k,u_k)$$

According to Theorem 1 of [?], there exists (given that the system is group affine) a function g and a matrix $\mathbf{A_k}$ such that:

$$[f(X_k,u_k)]^{-1} f(\hat{X}_k,u_k) = g(X_k^{-1}\hat{X}_k,u_k) =$$
$$g(\eta_k,u_k) = g(\exp(\xi_k),u_k) = \exp(\mathbf{A}_k\xi_k)$$

We have:

$$\eta^L_{k+1} = \exp(\xi^L_{k+1}) = \exp(-w_k)\exp(\mathbf{A}_k\xi^L_k)$$
$$= \exp(\mathbf{A}_k\xi^L_k - w_k)$$
$$\xi^L_{k+1} = \mathbf{A}_k\xi^L_k - w_k$$

Where we have used BCH formula [?] to first order since both ξ^L_k and w_k are assumed small.
In the case where $f(X_k,u_k)$ can be defined by a transformation S_k such that $f(X_k,u_k) = X_kS_k$, we have a closed form for $\mathbf{A}_k$.

$$\eta^L_{k+1} = \exp(\xi^L_{k+1}) = [f(X_k,u_k)\exp(w_k)]^{-1} f(\hat{X}_k,u_k) =$$
$$\exp(-w_k)S_k^{-1}X_k^{-1}\hat{X}_kS_k =$$
$$\exp(-w_k)S_k^{-1}\eta^L_kS_k = \exp(-w_k)S_k^{-1}\exp(\xi^L_k)S_k =$$
$$\exp(Adj_{S_k^{-1}}\xi^L_k - w_k)$$
$$\xi^L_{k+1} = Adj_{S_k^{-1}}\xi^L_k - w_k$$

$$\mathbf{A}_k = Adj_{S_k^{-1}}$$
$$\mathbf{L}_k = -\mathbf{I}$$

2) right-invariant: If we use the right-invariant error, at step k+1 it becomes:

$$\eta^R_{k+1} = \hat{X}_{k+1}X^{-1}_{k+1} =$$
$$f(\hat{X}_k,u_k)[f(X_k,u_k)\exp(w_k)]^{-1} =$$
$$f(\hat{X}_k,u_k)\exp(-w_k)[f(X_k,u_k)]^{-1} =$$
$$f(\hat{X}_k,u_k)\exp(-w_k)[f(\hat{X}_k,u_k)]^{-1} f(\hat{X}_k,u_k)[f(X_k,u_k)]^{-1}$$

Since $f(\hat{X}_k,u_k)$ must be an element of SE(3) we can define:

$$f(\hat{X}_k,u_k)\exp(-w_k)[f(\hat{X}_k,u_k)]^{-1} = exp(Adj_{f(\hat{X}_k,u_k)} - w_k)$$

We then get:

$$\eta^R_{k+1} =$$
$$f(\hat{X}_k,u_k)\exp(-w_k)[f(\hat{X}_k,u_k)]^{-1} f(\hat{X}_k,u_k)[f(X_k,u_k)]^{-1}$$
$$\exp(Adj_{f(\hat{X}_k,u_k)}(-w_k)) f(\hat{X}_k,u_k)[f(X_k,u_k)]^{-1}$$

We can use the same theorem as in the left-invariant case to say $f(\hat{X}_k,u_k)[f(X_k,u_k)]^{-1} = \exp(\mathbf{A}_k\xi_k)$. Then:

$$\eta^R_{k+1} = \exp(\xi_{k+1})$$
$$\exp(-Adj_{f(\hat{X}_k,u_k)}w_k)\exp(\mathbf{A}_k\xi_k)$$

Finnally using BCH formula to first order since both eps_k and w_k are assumed small:

$$\xi_{k+1} = \mathbf{A}_k\xi_k - Adj_{f(\hat{X}_k,u_k)}w_k$$

In the case where $f(X_k,u_k)$ can be defined by a transformation S such that $f(X_k,u_k) = X_kS_k$, we have a closed form for $\mathbf{A}_k$.

$$\eta^R_{k+1} = \exp(\xi^R_{k+1}) = f(\hat{X}_k,u_k)[f(X_k,u_k)\exp(w_k)]^{-1} =$$
$$\hat{X}_kS_k\exp(-w_k)S_k^{-1}X_k^{-1} =$$

Since $-w_k$ and S_k are assumed small, $S_k\exp(-w_k)S_k^{-1} = \exp(s_k - w_k - s_k) = \exp(-w_k)$

$$\exp(\xi^R_{k+1}) \approx \hat{X}_k\exp(-w_k)X_k^{-1} =$$
$$\hat{X}_k\exp(-w_k)\hat{X}_k^{-1}\eta^R_k =$$
$$\exp(-Adj_{\hat{X}_k}w_k)\exp(\xi^R_k)$$

$$\xi^R_{k+1} = \xi^R_k - Adj_{\hat{X}_k}w_k$$
$$\mathbf{A_k} = \mathbf{I}$$
$$\mathbf{L_k} = -\mathbf{Adj_{\hat{X}_k}}$$

B. Measurement

We will consider two specific measurement types:

1) Measurement of a vector of the body frame in the inertial frame: The first type is a series of measurements of the following form:

$$\mathbf{Y}^L_{k,i} = X_k\mathbf{d}^i_b$$
$$\hat{\mathbf{Y}}^L_{k,i} = \hat{X}_k\mathbf{d}^i_b$$

where $\mathbf{d}^i_b$ are known. We will now show that this measurement must be used with the left-invariant error. The derivation will be done neglecting measurement noise for simplicity.
First we define the innovation term as (dropping the subscripts L,i):

$$\mathbf{K}_k\mathbf{z}_k = \mathbf{K}_k\hat{X}_k^{-1}(\mathbf{Y}_k - \hat{\mathbf{Y}}_k) = \mathbf{K}_k(\hat{X}_k^{-1}\mathbf{Y}_k - \mathbf{d}_b)$$

The updated estimate is defined as follows:

$$\hat{X}^+_k = \hat{X}_k\exp(\mathbf{K}_k(\hat{X}_k^{-1}\mathbf{Y}_k - \mathbf{d}_b)$$

We will now derive and linearize the updated error, in order to get an equation resembling the standard EKF.

$$\eta^+_k = X_k^{-1}\hat{X}^+_k =$$
$$X_k^{-1}\hat{X}\exp(\mathbf{K}_k(\hat{X_k}^{-1}\mathbf{Y}_k - \mathbf{d}_b) =$$
$$\eta_k\exp(\mathbf{K}_k(\hat{X}_k^{-1}\mathbf{Y}_k - \mathbf{d}_b)$$

linearize assuming a small ξ_k and $\mathbf{z}_k$, neglecting $O(\xi_k \mathbf{z}_k)$.

$$\eta_k \exp(\mathbf{K}_k(\hat{X}_k^{-1}\mathbf{Y}_k - \mathbf{d}_b) \approx$$

$$\mathbf{I} + \boldsymbol{\xi}_k^{+\times} = \mathbf{I} + (\mathbf{K}_k(\hat{X}_k^{-1}\mathbf{Y}_k - \mathbf{d}_b))^\times(\mathbf{I} + \boldsymbol{\xi}_k^\times) =$$

$$\mathbf{I} + \boldsymbol{\xi}_k^{+\times} = \mathbf{I} + (\mathbf{K}_k(\hat{X}_k^{-1}\mathbf{Y}_k - \mathbf{d}_b))^\times + \boldsymbol{\xi}_k^{+\times}$$
$$\boldsymbol{\xi}_k^{+\times} = (\mathbf{K}_k(\hat{X}_k^{-1}\mathbf{Y}_k - \mathbf{d}_b))^\times + \boldsymbol{\xi}_k^\times$$

plug in the value of Y_k.

$$\boldsymbol{\xi}_k^{+\times} = (\mathbf{K}_k(\hat{X}_k^{-1}X_k\mathbf{d}_b - \mathbf{d}_b))^\times + \boldsymbol{\xi}_k^\times =$$
$$(\mathbf{K}_k(\eta_k\mathbf{d}_b - \mathbf{d}_b))^\times + \boldsymbol{\xi}_k^\times \approx$$
$$(\mathbf{K_k}((\mathbf{I} + \boldsymbol{\xi_k}^\times)\mathbf{d_b} - \mathbf{d_b}))^\times + \boldsymbol{\xi_k}^\times =$$
$$(\mathbf{K}_k\boldsymbol{\xi}_k^\times\mathbf{d}_b)^\times + \boldsymbol{\xi}_k^\times$$

$$\xi_k^+ = \xi_k + \mathbf{K}_k\boldsymbol{\xi}_k^\times\mathbf{d}_b$$
$$\xi_k^+ = \xi_k + \mathbf{K}_k\mathbf{C}_k\xi_k$$

Where we have define $\mathbf{C}_k$ such that:

$$\mathbf{C}_k\xi_k = \boldsymbol{\xi}_k^\times\mathbf{d}_b$$

Remember that $\times$ is the map from the tangent space to the Lie algebra, which is a generalised cross operator. As such we can not simply say that $\boldsymbol{\xi}_k^\times\mathbf{d}_b = -\mathbf{d}_b^\times\boldsymbol{\xi}_k$. We will show later how to get $\mathbf{C}_k$ in a specific example.
If we now add noise by replacing $\mathbf{Y}_k$ by $\mathbf{Y}_k + \mathbf{V}_k$ where:

$$\mathbf{V}_k = \begin{bmatrix} \mathbf{v}_k \\ 0 \end{bmatrix}$$
$$\mathbf{v}_k \sim \mathcal{N}(0, \mathbf{R_k})$$

We can easily follow the same steps and get:

$$\xi_k^+ = \xi_k + \mathbf{K}_k(\mathbf{C}_k\xi_k + X_k^{-1}\mathbf{V}_k)$$

As such we get the last linearized matrix:

$$\mathbf{M}_k = X_k^{-1}$$

2) Example: GPS: A GPS can be modeled as a direct measurement of $\mathbf{r}_a^{zw}$. If we set:

$$\mathbf{d}_b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

We get:

$$\mathbf{Y}_k^L = X_k\mathbf{d}_b = \begin{bmatrix} \mathbf{C}_{ab_k} & \mathbf{r}_a^{z_kw} \\ \mathbf{0}_{3\times1} & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_a^{z_kw} \\ 1 \end{bmatrix}$$

Discarding the last element we have the GPS measurement.

3) Measurement of a vector of the inertial frame in the body frame: The second type of measurements has the following form:

$$\mathbf{Y}_{k,i}^R = X_k^{-1}\mathbf{d}_a^i$$
$$\hat{\mathbf{Y}}_{k,i}^R = \hat{X}_k^{-1}\mathbf{d}_a^i$$

where $\mathbf{d}_a^i$ are known. We will now show that this measurement must be used with the right-invariant error. Again, the derivation will be done neglecting measurement noise for simplicity. First we define the innovation term as (dropping the subscripts R,i):

$$\mathbf{K}_k\mathbf{z}_k = \mathbf{K}_k\hat{X}_k(\mathbf{Y}_k - \hat{\mathbf{Y}}_k) = \mathbf{K}_k(\hat{X}_k\mathbf{Y}_k - \mathbf{d}_a)$$

The updated estimate is defined as follows:

$$\hat{X_k^+} = \exp(\mathbf{K}_k(\hat{X}_k\mathbf{Y}_k - \mathbf{d}_a)\hat{X}_k$$

Once again, we will derive and linearize the updated error.

$$\eta_k^+ = \hat{X}_k^+X_k^{-1} =$$
$$\exp(\mathbf{K}_k(\hat{X}_k\mathbf{Y}_k - \mathbf{d}_a)\hat{X}_kX_k^{-1} =$$
$$\exp(\mathbf{K}_k(\hat{X}_k\mathbf{Y}_k - \mathbf{d}_a)\eta_k$$

From here the steps of the previous section can be followed. Finnaly we get:

$$\boldsymbol{\xi}_k^+ = \boldsymbol{\xi}_k + \mathbf{K}_k\mathbf{C}_k\boldsymbol{\xi_k}$$

Where we have defined $\mathbf{C}_k$ such that:

$$\mathbf{C}_k\boldsymbol{\xi}_k = \boldsymbol{\xi}_k^\times\mathbf{d}_a$$

Once again adding noise we follow the same process and get:

$$\xi_k^+ = \xi_k + \mathbf{K}_k(\mathbf{C}_k\xi_k + X_k\mathbf{V}_k)$$
$$\mathbf{M}_k = X_k$$

Using $\mathbf{A}_k, \mathbf{L}_k, \mathbf{C}_k, \mathbf{M}_k$ matrices in "standard" EKF equations provides us with the Kalman gain, covariance propagation and update.

4) Example: Landmarks measurement: We will show that the measurement of landmarks can be assotiated with the right-invariant case. The measurement is assumed to be the vector from a known feature in the environment to the body. It is measured and therefore resolved in the body frame. Figure 2 illustrates this.

We measure:

$$\mathbf{r}_b^{l_iz} = \mathbf{r}_b^{l_iw} - \mathbf{r}_b^{zw}$$

If we set:

$$\mathbf{d}_a^i = \begin{bmatrix} \mathbf{r}_b^{l_iw} \\ 1 \end{bmatrix}$$

We get:

$$\mathbf{Y_k^L} = X_k^{-1}\mathbf{d}_a = \begin{bmatrix} \mathbf{C}_{ab_k}^{-1} & \mathbf{C}_{ab_k}^{-1}\mathbf{r}_a^{z_kw} \\ \mathbf{0}_{3\times1} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}_a^{l_iw} \\ 1 \end{bmatrix} =$$
$$\begin{bmatrix} \mathbf{C}_{b_ka} & \mathbf{r}_b^{z_kw} \\ \mathbf{0_{3\times1}} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}_a^{l_iw} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{b_ka}\mathbf{r}_a^{l_iw} - \mathbf{r}_b^{z_kw} \\ 1 \end{bmatrix} =$$
$$\begin{bmatrix} \mathbf{r}_b^{l_iw} - \mathbf{r}_b^{z_kw} \\ 1 \end{bmatrix}$$

Discarding the last element we have the wanted measurement.

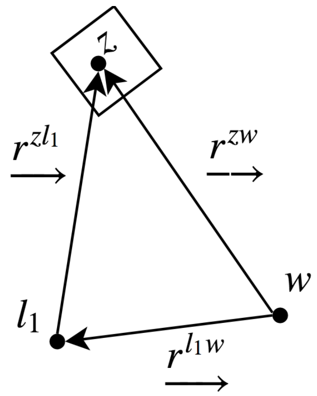

Fig. 2. setup of landmarks measurement

VI. Practical considerations

A. Adjoint representation

The Adjoint representation for different lie groups can easily be found in the litterature(for example [?]) It is important to note there are two Adjoint representations usually denoted Adj and adj. Adj is defined by the lie group:

$$S = \exp(s) S \exp(\xi) S^{-1} = \exp(Adj_S \xi) = \exp(Adj_S \xi)$$

adj is defined through the lie algebra by:

$$Adj_S \xi = \exp_m(adj_s)\xi$$

To make things more complicated there two ways to define the tangent space for se(3):

$$\mathbf{x} = \begin{bmatrix} \phi \\ \mathbf{r} \end{bmatrix}$$
$$or$$
$$\mathbf{x} = \begin{bmatrix} \mathbf{r} \\ \phi \end{bmatrix}$$

These two definitions each have their own version of Adj and adj. The easiest way to determine if the form of Adj is right is to directly check the definition with a numerical example.

B. Update matrices order reduction

It can be shown that the 4th row of the innovation term will always be zero. Therefore both $\mathbf{C}_k$ and $\mathbf{M}_k$ can be reduced by removing their last row.
In the SE(3), left invariant case $\mathbf{M_k}$ becomes :

$$\mathbf{M}_k = \mathbf{C}_{ab_k}^{-1}$$

And in the right invariant case it is:

$$\mathbf{M}_k = \mathbf{C}_{ab_k}$$

C. Closed form for $\mathbf{C_k}$

If $\mathbf{d}_{a,b}$ can be written as:

$$\mathbf{d} = \begin{bmatrix} \mathbf{r} \\ k \end{bmatrix}$$

we can derive a closed form for $\mathbf{C}_k$

$$\mathbf{C}_k \boldsymbol{\xi}_k = \boldsymbol{\xi}_k^{\times} \mathbf{d}_a = \begin{bmatrix} \boldsymbol{\xi}_r \\ \boldsymbol{\xi}_\phi \end{bmatrix}^{\times} \mathbf{d}_a = \begin{bmatrix} \boldsymbol{\xi}_\phi^{\times} & \boldsymbol{\xi}_r \\ \mathbf{0} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{r} \\ k \end{bmatrix} =$$
$$\begin{bmatrix} \boldsymbol{\xi}_\phi^{\times} \mathbf{r} + k\boldsymbol{\xi}_r \\ 0 \end{bmatrix} = \begin{bmatrix} -\mathbf{r}^{\times} \boldsymbol{\xi}_\phi + k\boldsymbol{\xi}_r \\ 0 \end{bmatrix} =$$
$$\begin{bmatrix} k\mathbf{I} & -\mathbf{r}^{\times} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}_r \\ \boldsymbol{\xi}_\phi \end{bmatrix}$$

$$\mathbf{C}_k = \begin{bmatrix} k\mathbf{I} & -\mathbf{r}^{\times} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

As mentioned above we can reduce $\mathbf{C}_k$ by removing the last row. We get:

$$\mathbf{C}_k = \begin{bmatrix} k\mathbf{I} & -\mathbf{r}^{\times} \end{bmatrix}$$

VII. Simulation results

The example of a 3D submarine was chosen to demonstrate the IEKF with a SE(3) state. Two different measurements are used to compare the RIEKF and LIEKF. The left-invariant measurement is a noisy GPS with a covariance $\mathbf{R}_k$. The GPS can be represented by a measurement of the vector $\mathbf{d}_b = [0; 0; 0; 1]^T$. In the right-invariant case the submarine measures the direction and distance to 3 known landmarks. This can be done by a LIDAR. The measurement is 3 stacked vectors, each with a covariance $\mathbf{R}_k$. In both cases the inputs are linear and angular velocities $\mathbf{v}_b$, $\boldsymbol{\omega}_b^{ba}$. The LIEKF is compared to a more classical MEKF.

A. Common equations

$$\mathbf{v}_b = \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix}; \boldsymbol{\omega}_b^{ba} = \begin{bmatrix} 0 \\ 0 \\ 0.5 \end{bmatrix}$$
$$X_k = \mathbf{T}_{ab_k} = \begin{bmatrix} \mathbf{C}_{ab_k} & \mathbf{r}_a^{z_k w} \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix}$$
$$S_k = \begin{bmatrix} \exp_m(\boldsymbol{\omega}_b^{ba\times} dt) & \mathbf{v}_b \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix}$$

Initialization:

$$\hat{X}_0 = \begin{bmatrix} \exp_m(\boldsymbol{\phi}_{offset}^{\times})\mathbf{C}_{ab_0,true} & \mathbf{r}_a^{z_0 w,true} + \mathbf{r}_{offset} \\ \mathbf{0}_{3\times 1} & 1 \end{bmatrix}$$

$$\mathbf{P}_0 = \begin{bmatrix} \pi/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \pi/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \pi/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Propagation:

$$\hat{X}_{k+1} = \hat{X}_k S_k$$
$$\mathbf{P}_{k+1} = \mathbf{A}_k \mathbf{P}_k \mathbf{A}_k^T + \mathbf{L}_k \mathbf{P}_k \mathbf{L}_k^T$$

Update:

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{C}_k (\mathbf{C}_k \mathbf{P}_k \mathbf{C}_k^T + \mathbf{R}_k)^{-1}$$
$$\mathbf{P}_k^+ = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k)\mathbf{P}_k$$

B. LIEKF

The following equations are specific to the LIEKF:

$$\mathbf{A}_k = Adj_{S_k^{-1}}$$
$$\mathbf{L}_k = \mathbf{I}$$
$$\mathbf{C}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$
$$\mathbf{M}_k = \hat{\mathbf{C}}_{ab_k}$$
$$\mathbf{R}_k = 0.1\mathbf{I}_{3\times3}$$
$$\mathbf{Q}_k = 0.1\mathbf{I}_{6\times6}$$

Update:

$$\hat{\mathbf{Y}}_k = \hat{X}_k \mathbf{d}_b = \hat{\mathbf{r}}_a^{z_k w}$$
$$\hat{X}_k^+ = \hat{X}_k \exp(\mathbf{K}_k(\hat{\mathbf{C}}_{ab}^{-1}(\mathbf{r}_a^{z_k w} - \hat{\mathbf{r}}_a^{z_k w})$$

The following results were obtained for $\mathbf{r}_{offset} = \mathbf{0}$ and $\phi_{offset} = [0; \pi/2; 0]^T$ with GPS and propagation noise respectively off and then on.

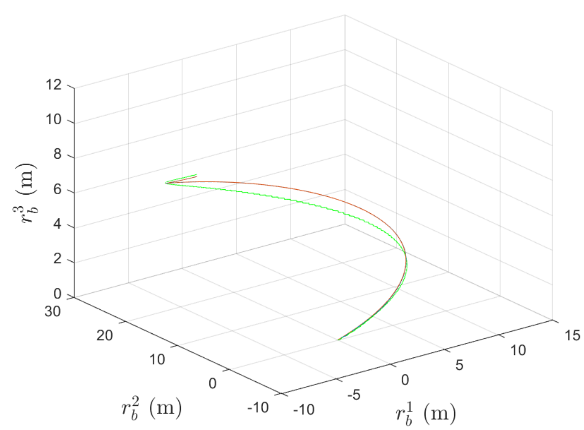

Fig. 3. True and estimated Position in 3D without noise. Red:true trajectory, Blue:LIEKF, Green: MEKF

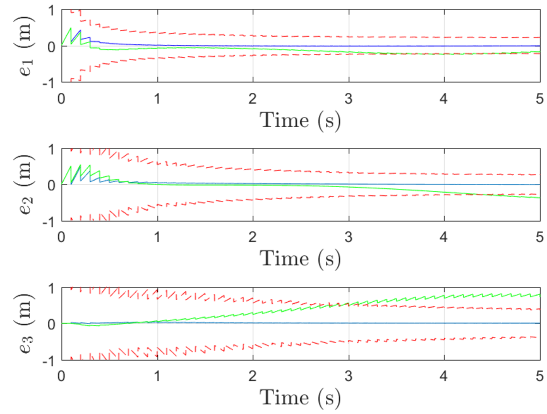

Fig. 4. Position error without noise. Red:LIEKF 3σ bounds, Blue:LIEKF, Green: MEKF

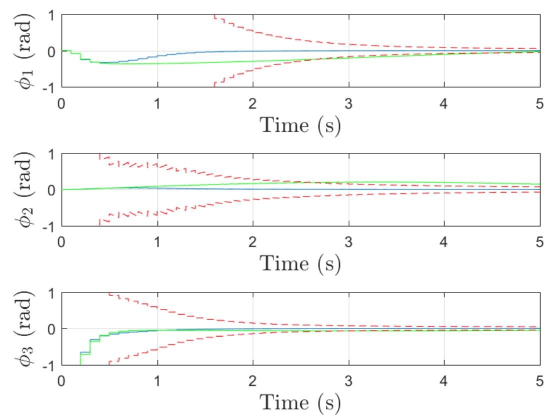

Fig. 5. Attitude error represented by a vector without noise. Red:LIEKF 3σ bounds, Blue:LIEKF, Green: MEKF

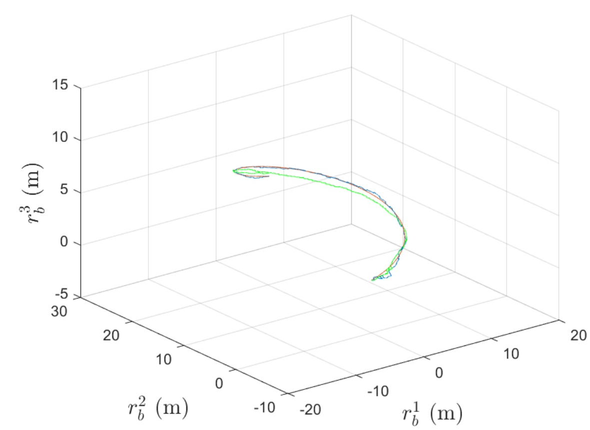

Fig. 6. True and estimated Position in 3D with noise. Red:true trajectory, Blue:LIEKF, Green: MEKF

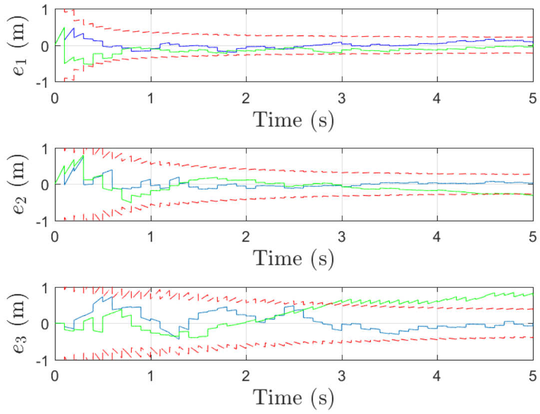

Fig. 7. Position error with noise. Red:LIEKF 3σ bounds, Blue:LIEKF, Green: MEKF

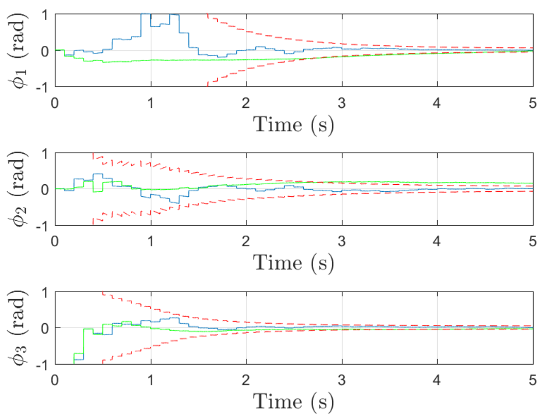

Fig. 8. Attitude error represented by a vector with noise. Red:LIEKF 3σ bounds, Blue:LIEKF, Green: MEKF

C. RIEKF

The following equations are specific to the RIEKF:

$$\mathbf{A}_k = \mathbf{I}$$
$$\mathbf{L}_k = Adj_{\hat{X}_k}$$
$$\mathbf{C}_k^i = \begin{bmatrix} \mathbf{I} & -\mathbf{r}_a^{l_i w \times} \end{bmatrix}$$
$$\mathbf{C}_k = \begin{bmatrix} \mathbf{C}_k^1 \\ \mathbf{C}_k^2 \\ \mathbf{C}_k^3 \end{bmatrix}$$
$$\mathbf{M}_k = \hat{\mathbf{C}}_{ab_k}^{-1}$$
$$\mathbf{R}_k = 0.1\mathbf{I}_{9\times 9}$$
$$\mathbf{Q}_k = 0.1\mathbf{I}_{6\times 6}$$

Update:

$$\mathbf{d}_a^i = \begin{bmatrix} \mathbf{r}_b^{l_i w} \\ 1 \end{bmatrix}$$
$$\mathbf{Y}_k^i = X_k \mathbf{d}_a = \begin{bmatrix} \mathbf{r}_b^{l_i w} - \mathbf{r}_b^{z_k w} \\ 1 \end{bmatrix}$$
$$\mathbf{z}_k = \begin{bmatrix} X_k^{-1}\hat{\mathbf{Y}}_k^1 - \mathbf{d}_a^1 \\ X_k^{-1}\hat{\mathbf{Y}}_k^2 - \mathbf{d}_a^2 \\ X_k^{-1}\hat{\mathbf{Y}}_k^3 - \mathbf{d}_a^3 \end{bmatrix}$$
$$\hat{X}_k^+ = \exp(\mathbf{K}_k \mathbf{z})\hat{X}_k$$

The following results were obtained for $\mathbf{r}_{offset} = \mathbf{0}$ and $\phi_{offset} = [0; \pi/2; 0]^T$ with GPS and propagation noise respectively off and then on.

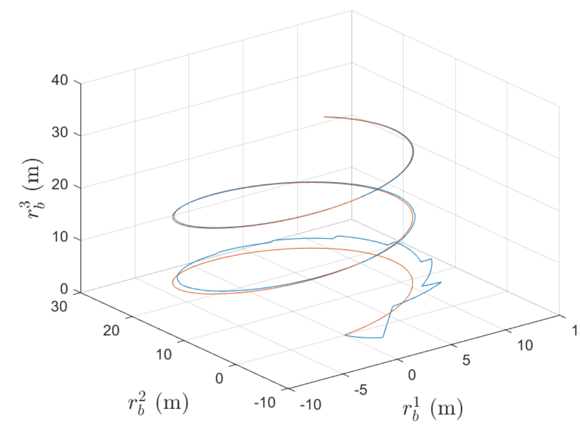

Fig. 9. True and estimated Position in 3D without noise. Red:true trajectory, Blue:RIEKF

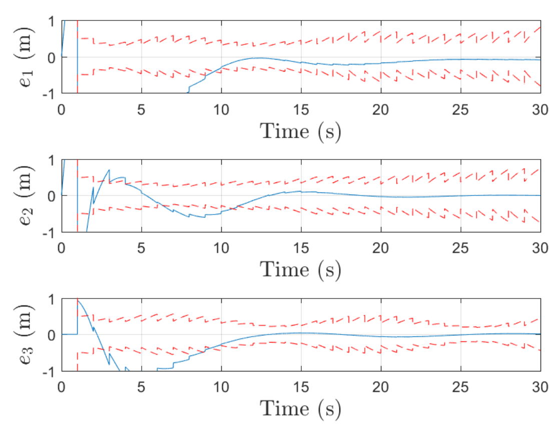

Fig. 10. Position error without noise. Red:RIEKF 3σ bounds, Blue:RIEKF

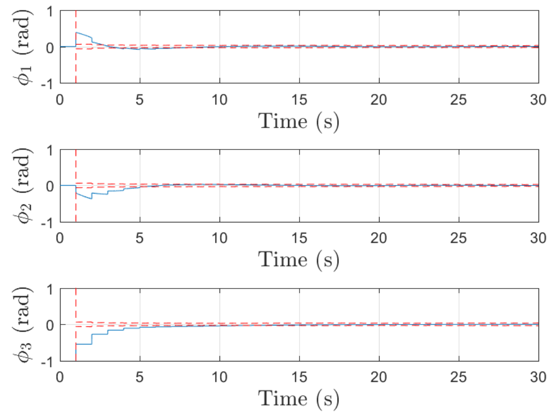

Fig. 11. Attitude error represented by a vector without noise. Red:RIEKF 3σ bounds, Blue:RIEKF

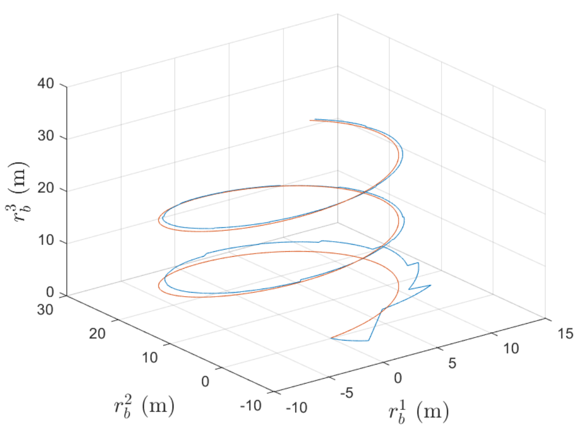

Fig. 12. True and estimated Position in 3D with noise. Red:true trajectory, Blue:RIEKF

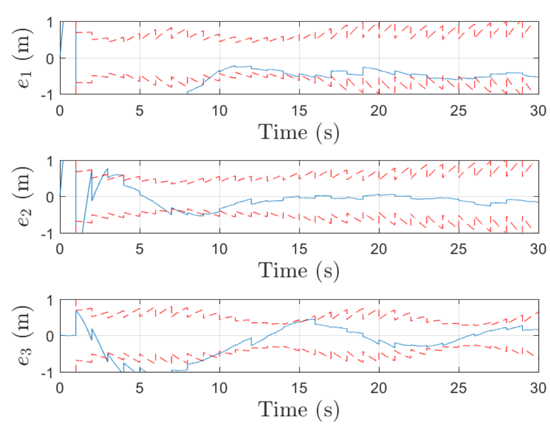

Fig. 13. Position error with noise. Red:RIEKF 3σ bounds, Blue:RIEKF

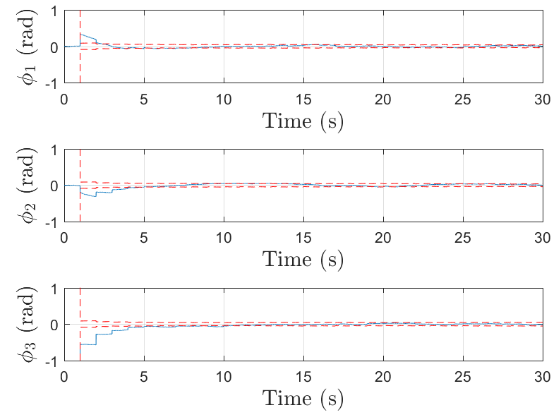

Fig. 14. Attitude error represented by a vector with noise. Red:RIEKF 3σ bounds, Blue:RIEKF

D. Discussion

The results provided above show that both LIEKF and RIEKF converge despite a large initial error. In the left-invariant case, we can see that the MEKF fails to converge for such an inital error. Actually, the MEKF and LIEKF behave very similarly for small errors. Their reaction to noise is also very similar. The main difference comes from the fact the LIEKF's $\mathbf{A}_k$ matrix is not a function of the state estimate and as such it converges even if the estimate is really bad. The right-invariant version takes much more time to converge. This can be understood by the fact when the attitude error is large, the measurement of landmarks in the body frame will be completely different from what is expected. As such the innovation term will be large, which breaks the assumption we have made in the update derivation. It would seem that the RIEKF is inferior to the LIEKF, but it is not the case. The LIEKF is more robust to attitude errors while the RIEKF is actually much more robust to position errors. The first case was demonstrated in the plots above. The second case can be seen in the figurres below. Interestingly the MEKF seems to perform better than the LIEKF at the beginning, but e_3 fails to converge. This time $\mathbf{r}_{offset} = [0; 5; 5]^T$ and $\boldsymbol{\phi}_{offset} = \mathbf{0}$. Note that the update frequency of the RIEKF is 1Hz and it still converges much faster (mostly in a single update).

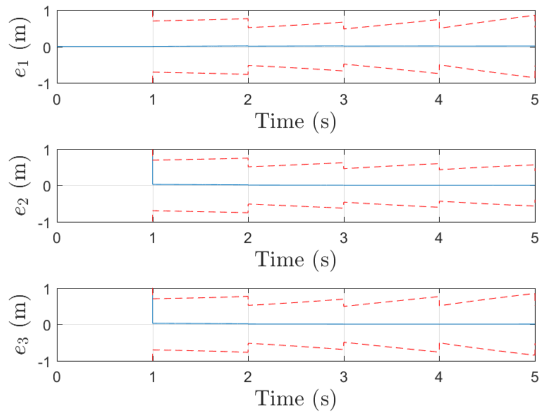

Fig. 15. Postion error represented by a vector. Red:RIEKF 3σ bounds, Blue:RIEKF

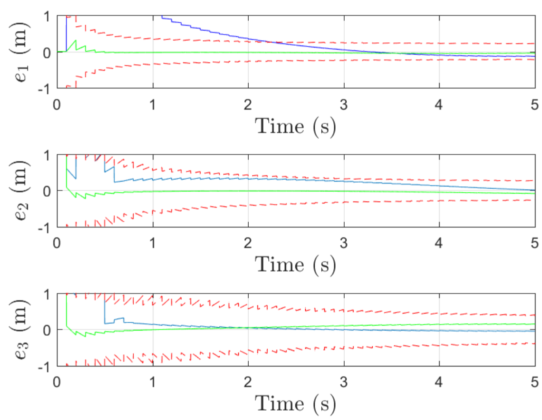

Fig. 16. Postion error represented by a vector. Red:LIEKF 3σ bounds, Blue:LIEKF, Green: MEKF

VIII. Conclusion

This report has developped the theory behind the Invariant Extended Kalman Filter. All derivations were based on basic linear algebra principles and known Lie group results. The only external result is the theorem proposed by [**?**]. Simulation results implementing the developed theory were then presented. It looks like the IEKF is a good candidate for attitude and pose estimation. In particular circumstances, it performs much better than the MEKF. An intersting extension to this report would be to prove the stability properties of this filter (this is done in [**?**]).

A logical progression would be to develop an invariant observer for more complex states. For example estimate velocity and/or bias in the measurements. This theory is general enough to encompass these cases as long as they can represented by a matrix Lie Group. Therefore, the procedure for the development of a more complex observer should start by showing that the state can be embedded in a Lie Group, then the properties of this Lie Group (exponential map and Adjoint) should be derived. Finally the steps shown in this report can be followed.

References

[1] A. Barrau, S. Bonnabel, "The Invariant Extended Kalman filter as a stable observer", *IEEE Transactions on Automatic Control*, Volume 62, Issue 4, April 2017.

[2] A. Barrau, S. Bonnabel, "Navigating with highly precise odometry and noisy GPS: a case study", *10th IFAC Symposium on Nonlinear Control Systems NOLCOS 2016*, : Monterey, California, USA, 2325 August 2016 pp. 70–77, 1981.

[3] E. Eade "Lie Groups for Computer Vision",http://ethaneade.com/lie.pdf.