Convergence and Consistency Analysis for An Invariant-EKF 3D SLAM

Teng Zhang, Kanzhi Wu, Jingwei Song, Shoudong Huang and Gamini Dissanayake

***Abstract*—In this paper, we investigate the convergence and consistency properties of an Invariant-Extended Kalman Filter (RI-EKF) based Simultaneous Localization and Mapping (SLAM) algorithm. Basic convergence properties of this algorithm are proven. These proofs do not require the restrictive assumption that the Jacobians of the motion and observation models need to be evaluated at the ground truth. It is also shown that the output of RI-EKF is invariant under any *stochastic rigid body transformation* in contrast to $\mathbb{SO}(3)$ based EKF SLAM algorithm ($\mathbb{SO}(3)$-EKF) that is only invariant under *deterministic rigid body transformation*. Implications of these invariance properties on the consistency of the estimator are also discussed. Monte Carlo simulation results demonstrate that RI-EKF outperforms $\mathbb{SO}(3)$-EKF, Robocentric-EKF and the "First Estimates Jacobian" EKF.**

***Index Terms*—Localization, Mapping, SLAM**

I. INTRODUCTION

EXTENDED Kalman filter (EKF) has been used extensively in solving Simultaneous Localization and Mapping (SLAM) problem in the past. However, a limitation of the traditional EKF based point feature SLAM is possible estimator inconsistency. Inconsistency here refers to the fact that the algorithm underestimates the uncertainty of the estimate leading to an overconfident result. This issue was recognized as early as in 2001 [1] and then discussed in detail later in [2][3]. Some research to enhance the consistency of EKF SLAM is reported in the literature. Robocentric EKF SLAM [4] estimates the location of landmarks in the robot local coordinate frame. As a result landmark positions to be estimated keep changing although landmarks are stationary in a fixed global coordinate frame. However, it has been shown that this robot-centric formulations lead to better performance in terms of estimator consistency. Guerreiro et al. [5] also reported a Kalman filter for SLAM problem formulated in a robocentric coordinate frame. Besides, it was shown in [6] that the inconsistency in EKF SLAM is closely related to the partial observability of SLAM problem [7][8]. This insight resulted in a number of EKF SLAM algorithms which significantly improve consistency, such as the "First Estimates Jacobian" EKF SLAM [6], observability-constrained EKF SLAM [9][10].

Manuscript received: *September, 10, 2016*; Revised *November, 27, 2016*; Accepted *December, 23, 2016.*

This paper was recommended for publication by Editor *Cyrill Stachniss* upon evaluation of the Associate Editor and Reviewers' comments.

The authors are with Center for Autonomous Systems, University of Technology Sydney, Australia {Teng.Zhang, Kanzhi.Wu, Jingwei.Song, Shoudong.Huang, Gamini.Dissanayake}@uts.edu.au

Digital Object Identifier (DOI): see top of this page.

On the other hand, a number of authors have addressed the behaviour of EKF SLAM to examine the convergence properties and derive bounds for the uncertainty of the estimate. In 2001, Dissanayake et. al. [11] proved three essential convergence properties of the algorithm under the assumption of linear motion and observation models, with theoretical achievable lower bounds on the resulting covariance matrix. In 2006, Mourikis and Roumeliotis [12] provided an analytical upper bound of the map uncertainty based on the observation noise level, the process noise level, and the size of the map. In 2007, Huang and Dissanayake [2] extended the proof of the convergence properties and the achievable lower bounds on covariance matrix in [11] to the nonlinear case, but under a restrictive assumption that the Jacobians are evaluated at the ground truth.

Recently, Lie group representation for three-dimensional orientation/pose has become popular in SLAM. (e.g., [13][14]), which can achieve better convergence and accuracy for both filter based algorithms (e.g., [15][16]) and the optimization based algorithms (e.g., [17][18]). Besides, the use of symmetry and Lie groups for observer design has gradually been recognized (e.g., [19]). The combination of the symmetry-preserving theory and the extended Kalman filter gives birth to the Invariant-EKF (I-EKF), which makes the traditional EKF possess the same invariance as the original system by using a geometrically adapted correction term. In [20], the I-EKF methodology is firstly applied into EKF-SLAM. And then an I-EKF based SLAM algorithm for 2D case, namely, the Right Invariant Error EKF (RI-EKF), is proposed in [21], which also intrinsically uses the Lie group representation, and the improved consistency is proven based on the linearized error-state model.

In this paper, we analyze the convergence and consistency properties of RI-EKF for 3D case. A convergence analysis for RI-EKF that does not require "Jacobians evaluated at the ground truth" assumption is presented. Furthermore, it is proven that the output of the filter is invariant under any *stochastic rigid body transformation* in contrast to $\mathbb{SO}(3)$ based EKF SLAM algorithm ($\mathbb{SO}(3)$-EKF) that is only invariant under *deterministic rigid body transformation*. We also discuss the relationship between these invariance properties and consistency and show that these properties have significant effect on the performance of the estimator via theoretical analysis and Monte Carlo simulations.

This paper is organized as follows. Section II recalls the motion model and the observation model of SLAM problem in 3D. Section III provides the RI-EKF SLAM algorithm. Section IV proves the convergence results of RI-EKF in two

fundamental cases. Section V proves the invariance property of RI-EKF and discusses the importance of this property. Section VI demonstrates RI-EKF outperforms $\mathbb{SO}(3)$-EKF, Robocentric-EKF and "First Estimates Jacobian" EKF SLAM algorithm through Monte Carlo simulations. Finally, Section VII outlines the main conclusions of this work.

Notations: Throughout this paper bold lower-case and upper-case letters are reserved for vectors and matrices/elements in manifold, respectively. The notation $S(\cdot)$ is the skew symmetric operator that transforms a 3-dimensional vector into a skew symmetric matrix: $S(\mathbf{x})\mathbf{y} = \mathbf{x} \times \mathbf{y}$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, where the notation $\times$ refers to the cross product.

II. Problem Statement

The EKF SLAM algorithms focus on estimating the current robot pose and the positions of all the observed landmarks with the given motion model and the observation model. In this work, SLAM problem in 3D scenarios is investigated and the state to be estimated is denoted by

$$\mathbf{X} = \left(\mathbf{R}, \mathbf{p}, \mathbf{f}^1, \cdots, \mathbf{f}^N\right), \tag{1}$$

where $\mathbf{R} \in \mathbb{SO}(3)$ and $\mathbf{p} \in \mathbb{R}^3$ are the robot orientation and robot position, $\mathbf{f}^i \in \mathbb{R}^3$ $(i = 1, \cdots, N)$ is the coordinate of the landmark i, all described in the fixed world coordinate frame.

A general motion model for moving robot and static landmarks in 3D scenarios can be represented by

$$\begin{aligned} &\mathbf{X}_{n+1} = f(\mathbf{X}_n, \mathbf{u}_n, \boldsymbol{\varepsilon}_n) \\ &= \left(\mathbf{R}_n \exp(\mathbf{w}_n + \boldsymbol{\varepsilon}^w_{\mathbf{n}}), \mathbf{p}_n + \mathbf{R}_n(\mathbf{v}_n + \boldsymbol{\varepsilon}^v_{\mathbf{n}}), \mathbf{f}^1_n, \cdots, \mathbf{f}^N_n\right), \end{aligned} \tag{2}$$

where $\mathbf{u}_n = \begin{bmatrix} \mathbf{w}_n^\intercal & \mathbf{v}_n^\intercal \end{bmatrix}^\intercal \in \mathbb{R}^6$ is the odometry, being $\mathbf{w}_n \in \mathbb{R}^3$ and $\mathbf{v}_n \in \mathbb{R}^3$ the angular increment and linear translation from time n to time $n+1$, $\exp(\cdot)$ is the exponential mapping of $\mathbb{SO}(3)$ defined in (15) and $\boldsymbol{\varepsilon}_n = \begin{bmatrix} (\boldsymbol{\varepsilon}^w_{\mathbf{n}})^\intercal & (\boldsymbol{\varepsilon}^v_{\mathbf{n}})^\intercal \end{bmatrix}^\intercal \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Phi}_n)$ is the odometry noise at time n.

As the robot is likely to observe different sets of landmarks in each time step, the notation O_{n+1} is used to represent the set that indicates the landmarks observed at time $n+1$. Also by assuming a 3D sensor which provides the coordinate of landmark i in $n+1$-th robot frame, the observation model is given as follows

$$\mathbf{z}_{n+1} = h_{n+1}(\mathbf{X}_{n+1}, \boldsymbol{\xi}_{n+1}), \tag{3}$$

where $h_{n+1}(\mathbf{X}_{n+1}, \boldsymbol{\xi}_{n+1})$ is a column vector obtained by stacking all entries $h^i(\mathbf{X}_{n+1}, \boldsymbol{\xi}^i_{n+1}) = \mathbf{R}^\intercal_{n+1}(\mathbf{f}^i_{n+1} - \mathbf{p}_{n+1}) + \boldsymbol{\xi}^i_{n+1} \in \mathbb{R}^3$ for all $i \in \mathrm{O}_{n+1}$, $\boldsymbol{\xi}_{n+1} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}_{n+1})$ is the observation noise vector obtained by stacking all entries $\boldsymbol{\xi}^i_{n+1} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi}^i_{n+1})$ $(i \in \mathrm{O}_{n+1})$. The covariance matrix $\boldsymbol{\Psi}_{n+1}$ of observation noise is a block diagonal matrix consisting of all $\boldsymbol{\Psi}^i_{n+1}$ $(i \in \mathrm{O}_{n+1})$.

III. The Invariant EKF SLAM Algorithm

In this section, RI-EKF based on the general EKF framework is briefly introduced. In the general EKF framework, the uncertainty of $\mathbf{X}$ is described by $\mathbf{X} = \hat{\mathbf{X}} \oplus \mathbf{e}$, where $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{P})$ is a white Gaussian noise vector and $\hat{\mathbf{X}}$ is the mean estimate of $\mathbf{X}$. The notation $\oplus$ is commonly called retraction in differentiable geometry [22] and it is designed as a smooth mapping such that $\mathbf{X} = \mathbf{X} \oplus \mathbf{0}$ and there exists the inverse mapping $\ominus$ of $\oplus$: $\mathbf{e} = \mathbf{X} \ominus \hat{\mathbf{X}}$. The process of propagation and update based on the general EKF framework has been summarized in Alg. 1, which is very similar to the standard EKF. Due to different uncertainty representation (compared to the standard EKF), the Jacobians of the general EKF framework in Alg. 1 are obtained by: $\mathbf{F}_n = \frac{\partial (f(\hat{\mathbf{X}}_n \oplus \mathbf{e}, \mathbf{u}_n, \mathbf{0}) \ominus f(\hat{\mathbf{X}}_n, \mathbf{u}_n, \mathbf{0}))}{\partial \mathbf{e}}|_{\mathbf{e}=\mathbf{0}}$, $\mathbf{G}_n = \frac{\partial (f(\hat{\mathbf{X}}_n, \mathbf{u}_n, \boldsymbol{\varepsilon}) \ominus f(\hat{\mathbf{X}}_n, \mathbf{u}_n, \mathbf{0}))}{\partial \boldsymbol{\varepsilon}}|_{\boldsymbol{\varepsilon}=\mathbf{0}}$, $\mathbf{H}_{n+1} = \frac{\partial h_{n+1}(\hat{\mathbf{X}}_{n+1|n} \oplus \mathbf{e}, \mathbf{0})}{\partial \mathbf{e}}|_{\mathbf{e}=\mathbf{0}}$.

Algorithm 1: The general EKF framework (RI-EKF)

Input: $\hat{\mathbf{X}}_n$, $\mathbf{P}_n$, $\mathbf{u}_n$, $\mathbf{z}_{n+1}$;
Output: $\hat{\mathbf{X}}_{n+1}$, $\mathbf{P}_{n+1}$;
Propagation:
$\hat{\mathbf{X}}_{n+1|n} \leftarrow f(\mathbf{X}_n, \mathbf{u}_n, \mathbf{0})$, $\mathbf{P}_{n+1|n} \leftarrow \mathbf{F}_n \mathbf{P}_n \mathbf{F}_n^\intercal + \mathbf{G}_n \Phi_n \mathbf{G}_n^\intercal$;
Update:
$\mathbf{S} \leftarrow \mathbf{H}_{n+1} \mathbf{P}_{n+1|n} \mathbf{H}_{n+1}^\intercal + \boldsymbol{\Psi}_{n+1}$, $\mathbf{K} \leftarrow \mathbf{P}_{n+1|n} \mathbf{H}_{n+1}^\intercal \mathbf{S}^{-1}$;
$\mathbf{y} \leftarrow h_{n+1}(\hat{\mathbf{X}}_{n+1|n}, \mathbf{0}) - \mathbf{z}_{n+1}$;
$\hat{\mathbf{X}}_{n+1} \leftarrow \hat{\mathbf{X}}_{n+1|n} \oplus \mathbf{K}\mathbf{y}$, $\mathbf{P}_{n+1} \leftarrow (\mathbf{I} - \mathbf{K}\mathbf{H}_{n+1})\mathbf{P}_{n+1|n}$;
(In RI-EKF, $(\mathbf{F}_n, \mathbf{G}_n, \mathbf{H}_n)$ are given in (4) and (5))

A. RI-EKF

RI-EKF follows the general EKF framework summarized in Alg. 1. The state space of RI-EKF is modeled as a Lie group $\mathcal{G}(N)$. The background knowledge about Lie group $\mathcal{G}(N)$ is provided in Appendix A.

1) The choice of $\oplus$: The retraction $\oplus$ of RI-EKF is chosen such that $\mathbf{X} = \hat{\mathbf{X}} \oplus \mathbf{e} := \exp(\mathbf{e})\hat{\mathbf{X}}$, where exp is the exponential mapping on the Lie group $\mathcal{G}(N)$[1], $\mathbf{X} \in \mathcal{G}(N)$ is the actual pose and landmarks, $\hat{\mathbf{X}} \in \mathcal{G}(N)$ is the *mean* estimate and the uncertainty vector $\mathbf{e} = \begin{bmatrix} \mathbf{e}_\theta^\intercal & \mathbf{e}_p^\intercal & (\mathbf{e}^1)^\intercal & \cdots & (\mathbf{e}^N)^\intercal \end{bmatrix}^\intercal \in \mathbf{R}^{3N+6}$ follows the Gaussian distribution $\mathcal{N}(\mathbf{0}, \mathbf{P})$.

2) Jacobian matrices: The Jacobians of the propagation step of RI-EKF are

$$\mathbf{F}_n = \mathbf{I}_{3N+6}, \ \mathbf{G}_n = ad_{\hat{\mathbf{X}}_n} \mathbf{B}_n, \tag{4}$$

where $\mathbf{B}_n = \begin{bmatrix} -J_r(-\mathbf{w}_n) & \mathbf{0}_{3,3} \\ -S(\mathbf{v}_n) J_r(-\mathbf{w}_n) & \mathbf{I}_3 \\ \mathbf{0}_{3N,3} & \mathbf{0}_{3N,3} \end{bmatrix}$. The adjoint operation ad and the right Jacobian $J_r(\cdot)$ are given in Appendix A. The Jacobian matrix $\mathbf{H}_{n+1}$ of the update step is obtained by stacking all matrices $\mathbf{H}^i_{n+1}$ for all $i \in \mathrm{O}_{n+1}$, where

$$\mathbf{H}^i_{n+1} = \begin{bmatrix} \mathbf{0}_{3,3} & \hat{\mathbf{R}}^\intercal_{n+1|n} & \cdots & -\hat{\mathbf{R}}^\intercal_{n+1|n} & \mathbf{0}_{3,3(N-i)} \end{bmatrix}. \tag{5}$$

For a general observation model that is a function of the relative position of the landmark, the Jacobian matrix $\mathbf{H}_{n+1}$ can be calculated by the chain rule.

[1]The exponential mapping is an overloaded function for Lie group and hence we also denote exp as the exponential mapping for the Lie group $\mathcal{G}(N)$ (given in (14)). More details and the Matlab code of the algorithms are available at "https://github.com/RomaTeng/EKF-SLAM-on-Manifold".

3) Landmark initialization: Here we provide the method to augment the state $\mathbf{X} \in \mathcal{G}(N)$ and adjust the covariance matrix $\mathbf{P}$ when the robot observes a new landmark with the observation $\mathbf{z} \in \mathbf{R}^3$. For brevity, the mathematical derivation is ignored here and the process to augment the state is summarized in Alg. 2, where $\mathbf{M}_N := \begin{bmatrix} \mathbf{0}_{3,3} & \mathbf{I}_3 & \mathbf{0}_{3,3N} \end{bmatrix}^\intercal$ and Ψ is the covariance matrix representing the noise level in the new landmark observation.

Algorithm 2: Landmark Initialization of RI-EKF

Input: $\hat{\mathbf{X}}$, $\mathbf{P}$, $\mathbf{z}$;
Output: $\hat{\mathbf{X}}_{new}$, $\mathbf{P}_{new}$;
Process:
$\hat{\mathbf{f}}^{N+1} = \hat{\mathbf{p}} + \hat{\mathbf{R}}\mathbf{z} \in \mathbb{R}^3$
$\hat{\mathbf{X}}_{new} \leftarrow (\hat{\mathbf{X}}, \hat{\mathbf{f}}^{N+1}) \in \mathcal{G}(N+1)$;
$\mathbf{P}_{new} \leftarrow \begin{bmatrix} \mathbf{P} & \mathbf{P}\mathbf{M}_N \\ \mathbf{M}_N^\intercal \mathbf{P} & \hat{\mathbf{R}}\Psi\hat{\mathbf{R}}^\intercal + \mathbf{M}_N^\intercal \mathbf{P}\mathbf{M}_N \end{bmatrix}$.

B. Discussion

The general EKF framework proposed in [21] allows more flexible uncertainty representation, compared to the standard EKF. From Alg. 1, one can see that **a general EKF framework based filter** can be designed via **a choice of retraction** $\oplus$. For example, the retraction $\oplus$ used in the 2D traditional EKF SLAM algorithm is the standard addition. A natural extension of the 2D traditional EKF SLAM algorithm is $\mathbb{SO}(3)$-EKF, in which the state space is modeled as $\mathbb{SO}(3) \times \mathbb{R}^{3+3N}$ and the retraction $\oplus$ is $\mathbf{X} = \hat{\mathbf{X}} \oplus \mathbf{e} = (\exp(\mathbf{e}_\theta)\hat{\mathbf{R}}, \mathbf{e}_p + \hat{\mathbf{p}}, \mathbf{e}^1 + \hat{\mathbf{f}}^1, \cdots, \mathbf{e}^N + \hat{\mathbf{f}}^N)$. Similarly, $\mathbb{SE}(3)$-EKF can be obtained by modeling the state space as $\mathbb{SE}(3) \times \mathbb{R}^{3N}$.

Another noticeable point is that two general EKF framework based filters with different choice of $\oplus$ may have the same Jacobians $(\mathbf{F}_n, \mathbf{G}_n, \mathbf{H}_n)$. For example, if the retraction $\oplus$ of RI-EKF is changed such that $\hat{\mathbf{X}} \oplus \mathbf{e} = (\exp(\mathbf{e}_\theta)\hat{\mathbf{R}}, \exp(\mathbf{e}_\theta)\hat{\mathbf{p}} + \mathbf{e}_p, \cdots, \exp(\mathbf{e}_\theta)\hat{\mathbf{f}}^N + \mathbf{e}^N)$, the resulting filter (Pseudo-RI-EKF) has the same Jacobians as that of RI-EKF but their performances are significantly different as shown in Section VI, showing that the choice of retraction $\oplus$ has significant effect on the behavior of the general EKF framework based filter. In the next section, we will compare the behavior of RI-EKF with $\mathbb{SO}(3)$-EKF via the theoretical proofs for the convergence property of RI-EKF and two simple examples.

IV. Convergence Analysis of RI-EKF SLAM Algorithm

The general expression for the covariance matrices evolution of RI-EKF cannot be easily obtained. Therefore, two representative scenarios are considered: (i) the robot is stationary, and (ii) the robot then moves one step. The convergence results of RI-EKF SLAM algorithm in the two scenarios are presented and proven, under the condition that Jacobians are evaluated at the latest estimate. Hence the results are significant extension to similar theorems in [2] where Jacobians evaluated at the true state are assumed to be available.

The general setting analyzed in the following subsections is as follows. The robot starts at point A with the initial condition $(\hat{\mathbf{X}}_0, \mathbf{P})$, where $\mathbf{P}$ is covariance matrix and $\hat{\mathbf{X}}_0 = (\hat{\mathbf{R}}, \hat{\mathbf{p}}, \hat{\mathbf{f}}^1, \cdots, \hat{\mathbf{f}}^N)$ (N landmarks have been observed). The covariance matrix of odometry noise is $\mathbf{\Phi}$ and the covariance matrix of one landmark observation noise is Ψ. In the following subsections, $\mathbf{M}_N := \begin{bmatrix} \mathbf{0}_{3,3} & \mathbf{I}_3 & \mathbf{0}_{3,3N} \end{bmatrix}^\intercal$, $\mathbf{L} := \mathbf{P}\mathbf{M}_N$ and $\mathbf{W} := \mathbf{M}_N^\intercal \mathbf{P}\mathbf{M}_N$. The odometry and the covariance of odometry noise are $\mathbf{0}_{6,1}$ and $\mathbf{0}_{6,6}$, respectively when robot remains stationary.

A. Scenario A: Robot remains stationary

Theorem 1: If the robot remains stationary at point A and does not observe any of the previously seen landmarks but observes a new landmark for k times, the *mean* estimate of robot pose and the N landmarks and their related uncertainty remain unchanged (via k times process of propagation and update of RI-EKF). The covariance matrix of the state when the new landmark is integrated becomes $\mathbf{P}_k = \begin{bmatrix} \mathbf{P} & \mathbf{L} \\ \mathbf{L}^\intercal & \frac{\hat{\mathbf{R}}\Psi\hat{\mathbf{R}}^\intercal}{k} + \mathbf{W} \end{bmatrix}$. When $k \to \infty$, the covariance matrix becomes

$$\mathbf{P}_\infty^A = \begin{bmatrix} \mathbf{P} & \mathbf{L} \\ \mathbf{L}^\intercal & \mathbf{W} \end{bmatrix}. \tag{6}$$

Proof 1: See Appendix B.

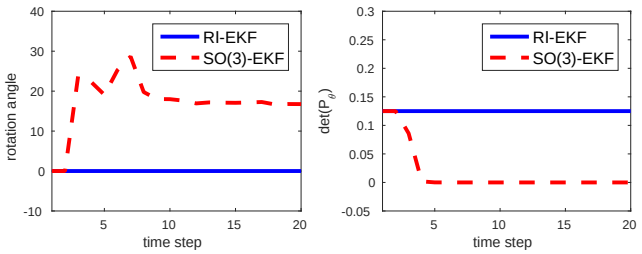

Figure 1: **Illustration of Theorem 1**. In this case, robot is stationary and always only observes the "new" landmark. Left: The error (unit: degree) in robot orientation w.r.t. the ground truth as estimated by RI-EKF and $\mathbb{SO}(3)$-EKF. Right: $\det(\mathbf{P}_\theta)$ estimated by RI-EKF and $\mathbb{SO}(3)$-EKF.

Theorem 1 can be interpreted as that the covariance matrix (w.r.t. robot pose) in RI-EKF will not be reduced by observing the "new" landmark when robot keeps stationary, which corresponds to a fact that the observations to previously unseen landmark do not convey any new information on the location of the robot. However, $\mathbb{SO}(3)$-EKF does not have this good convergence property.

We illustrate the results of Theorem 1 using the following scenario. The simulated robot remains stationary and always observes the "new" landmark (the covariance of observation noise is not zero). The "new" landmark is observed multiple times (a small loop closure) and the standard deviation of observation noise is set as 5% of robot-to-landmark distance along each axis. The initial covariance matrix $\mathbf{P}_\theta \in \mathbb{R}^{3\times3}$ of robot orientation is set as $\frac{1}{2}\mathbf{I}_3$. Fig. 1 presents results of a simulation of this scenario. The rotation angle relative to the initial orientation and $\det(\mathbf{P}_\theta)$ from RI-EKF correctly infers

that the robot remains stationary and the orientation uncertainty remains unchanged. In contrast, $\mathbb{SO}(3)$-EKF updates the robot orientation and furthermore predicts that the orientation uncertainty decreases as observations continue to be made, both of which are clearly erroneous and therefore leads to estimator inconsistency.

Theorem 1 can be easily extended to a multiple landmarks scenario.

Corollary 1: If the robot is stationary at point A and only observes m new landmark k times, the estimate of pose from RI-EKF does not change while the covariance matrix of the estimate becomes $\mathbf{P}_k = \begin{bmatrix} \mathbf{P} & \mathbf{L} & \mathbf{L} & \cdots & \mathbf{L} \\ \mathbf{L}^\intercal & \mathbf{Q}_k & \mathbf{W} & \cdots & \mathbf{W} \\ \mathbf{L}^\intercal & \mathbf{W} & \mathbf{Q}_k & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{W} \\ \mathbf{L}^\intercal & \mathbf{W} & \cdots & \mathbf{W} & \mathbf{Q}_k \end{bmatrix}$, where $\mathbf{Q}_k = \frac{\hat{\mathbf{R}}\boldsymbol{\Psi}\hat{\mathbf{R}}^\intercal}{k} + \mathbf{W}$. When $k \to \infty$, the covariance matrix becomes

$$\mathbf{P}_\infty^A = \begin{bmatrix} \mathbf{P} & \mathbf{L} & \mathbf{L} & \cdots & \mathbf{L} \\ \mathbf{L}^\intercal & \mathbf{W} & \mathbf{W} & \cdots & \mathbf{W} \\ \mathbf{L}^\intercal & \mathbf{W} & \mathbf{W} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{W} \\ \mathbf{L}^\intercal & \mathbf{W} & \cdots & \mathbf{W} & \mathbf{W} \end{bmatrix}. \tag{7}$$

B. Scenario B: Robot takes a step after a stationary period

Consider the condition that the robot moves one step after being stationary for a long period of time while observing new landmarks.

Theorem 2: Assume $\boldsymbol{\Psi} = \phi \mathbf{I}_3$ ($\phi \in \mathbb{R}^+$). If the robot remains stationary at point A, does not observe any of the previously seen landmarks but observes m new landmarks for $k = \infty$ times and then takes a step to B using control action $\mathbf{u} = \begin{bmatrix} \mathbf{w}^\intercal & \mathbf{v}^\intercal \end{bmatrix}^\intercal$ and observes the same set of landmarks l times, then the covariance matrix from RI-EKF becomes $\mathbf{P}_l^B = \mathbf{P}_\infty^A + \bar{\mathbf{P}}_l^B$, where $\mathbf{P}_\infty^A$ is given in (7), $\bar{\boldsymbol{\Psi}} = \phi \mathbf{I}_{3m}$ and

$$\bar{\mathbf{P}}_l^B = ad_{\hat{\mathbf{X}}_A}\mathbf{E}(\tilde{\boldsymbol{\Phi}}^{-1} + l\tilde{\mathbf{H}}^\intercal\bar{\boldsymbol{\Psi}}^{-1}\tilde{\mathbf{H}})^{-1}\mathbf{E}^\intercal ad_{\hat{\mathbf{X}}_A}^\intercal, \tag{8}$$

where $\bar{\boldsymbol{\Psi}} = \phi \mathbf{I}_{3m}$ and the covariance matrix of the odometry noise is $\boldsymbol{\Phi}$. In (8), $\hat{\mathbf{X}}_A$ is the estimated state at the point A, $\tilde{\boldsymbol{\Phi}} = \mathbf{B}\boldsymbol{\Phi}\mathbf{B}^\intercal$ is a positive definite matrix and

$$\begin{aligned} \mathbf{B} &= \begin{bmatrix} -J_r(-\mathbf{w}) & \mathbf{0}_{3,3} \\ -S(\mathbf{v})J_r(-\mathbf{w}) & \mathbf{I}_3 \end{bmatrix}, \\ \mathbf{E} &= \begin{bmatrix} \mathbf{I}_6 \\ \mathbf{0}_{3(N+m),6} \end{bmatrix}, \ \tilde{\mathbf{H}} = \mathbf{H}ad_{\hat{\mathbf{X}}_A}\mathbf{E}, \end{aligned} \tag{9}$$

where $\mathbf{H}$ is obtained by stacking all matrices $\mathbf{H}^i = \begin{bmatrix} \mathbf{0}_{3,3} & \mathbf{I}_3 & \mathbf{0}_{3,3(N+i-1)} & -\mathbf{I}_3 & \mathbf{0}_{3,3(m-i)} \end{bmatrix}$. When l tends to infinity, the covariance matrix becomes $\mathbf{P}_\infty^B = \mathbf{P}_\infty^A$ under the condition that there are three landmarks non-coplanar with the robot position.

Proof 2: See Appendix C.

From Theorem 2, one can see that the estimate of RI-EKF follows the expectation that "the only effect of the observations made at point B is to reduce the additional robot uncertainty generated from the odometry noise. The observations made at point B cannot reduce the uncertainty of the landmark further if the robot had already observed the landmark many times at point A. [2]"

We illustrate the results of Theorem 2 using the following scenario. Initially the robot is stationary at point A and continually observes ten previously unseen landmarks. It moves one step to point B after 200 such observations and then remains stationary for 200 more time steps while observing the same set of landmarks. The initial covariance matrix of robot pose is set as non-zero. In Fig. 2, we adopt $\log(\det(\mathbf{P}_r))$ as the extent of the uncertainty w.r.t. robot pose, where $\mathbf{P}_r \in \mathbb{R}^{6\times6}$ is the covariance matrix of the robot pose. In Fig. 2, the pose uncertainty from RI-EKF remains unchanged and increases at time 200 when robot moves one step due to odometry noise as expected. Further landmark observations at point B while remaining stationary gradually reduce the pose uncertainty. In contrast, the pose uncertainty from $\mathbb{SO}(3)$-EKF falls below the initial value indicating incorrect injection of information, leading to an overconfident estimate of uncertainty.

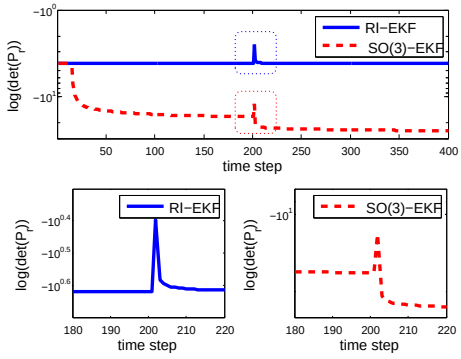

Figure 2: **Illustration of Theorem 2**. The y-axis is $\log(\det(\mathbf{P}_r))$ that represents the pose uncertainty. $\mathbf{P}_r$ is the covariance matrix of robot pose. Robot remains stationary from time 1 to time 200, moves one step at time 200 and then remains stationary.

V. Consistency Analysis

As seen in the previous section, RI-EKF SLAM algorithm meets the expectation that observing new landmarks does not help in reducing the robot pose uncertainty [3][5], while $\mathbb{SO}(3)$-EKF contradicts this. This section further investigates the reason for the phenomenon above.

A. Unobservability and invariance property

This subsection first reviews the unobservability of SLAM formulation (1)–(3), which is strongly related to the consistency issues of various SLAM estimation algorithms. The earliest concept of observability for nonlinear systems is proposed in [23]. From the viewpoint of nonlinear systems, SLAM formulation (as a system for the actual state $\mathbf{X}$) is not locally observable [23], as understood in [9][24]. In the following, we will mathematically describe the unobservability of SLAM formulation (1)–(3) in terms of stochastic rigid body transformation.

Definition 1: For SLAM problem formulation (1)–(3), a stochastic rigid body transformation $\mathcal{T}_{\mathbf{g}}$ is

$$\begin{aligned}\mathcal{T}_{\mathbf{g}}(\mathbf{X}) = (&\exp(\Theta_1)\bar{\mathbf{R}}\mathbf{R}, \exp(\Theta_1)\bar{\mathbf{R}}\mathbf{p}+\bar{\mathbf{T}}+\Theta_2,\\ &\exp(\Theta_1)\bar{\mathbf{R}}\mathbf{f}^1+\bar{\mathbf{T}}+\Theta_2,\cdots,\exp(\Theta_1)\bar{\mathbf{R}}\mathbf{f}^N+\bar{\mathbf{T}}+\Theta_2),\end{aligned} \tag{10}$$

where $\mathbf{X}$ is given in (1), $\mathbf{g}=(\bar{\mathbf{R}},\bar{\mathbf{T}},\Theta)$, $\bar{\mathbf{R}}\in\mathbb{SO}(3), \bar{\mathbf{T}}\in\mathbb{R}^3$ and $\Theta=\begin{bmatrix}\Theta_1^\intercal & \Theta_2^\intercal\end{bmatrix}^\intercal\in\mathbb{R}^6$ is white Gaussian noise with covariance $\bar{\mathbf{\Sigma}}$. When the covariance $\bar{\mathbf{\Sigma}}=\mathbf{0}_{6,6}$, this transformation degenerates into a **deterministic rigid body transformation**. When $\mathbf{g}=(\mathbf{I}_3,\mathbf{0}_{3,1},\Theta)$, this transformation degenerates into a **stochastic identity transformation**.

It can be easily verified that **the output (observations) of the system** (1)–(3) is invariant to any stochastic rigid body transformation $\mathcal{T}_{\mathbf{g}}$: for any two initial conditions, $\mathbf{X}_0$ and $\mathbf{Y}_0:=\mathcal{T}_{\mathbf{g}}(\mathbf{X}_0)$, we have $h_n(\mathbf{X}_n,\boldsymbol{\xi}_n)=h_n(\mathbf{Y}_n,\boldsymbol{\xi}_n)$ for all $n\geq 0$, where $\mathbf{X}_k=f(\mathbf{X}_{k-1},\mathbf{u}_{k-1},\boldsymbol{\varepsilon}_{k-1})$ and $\mathbf{Y}_k=f(\mathbf{Y}_{k-1},\mathbf{u}_{k-1},\boldsymbol{\varepsilon}_{k-1})$ $(k=1,\cdots,n-1)$. Therefore, SLAM formulation (1)–(3) is *unobservable* in terms of stochastic rigid body transformation. In the following, the invariance to stochastic rigid body transformation for the EKF framework based filter of SLAM formulation will be mathematically described.

Definition 2: The output (**estimated observations**) of a general EKF framework based filter is invariant under any stochastic rigid body transformation $\mathcal{T}_{\mathbf{g}}$ if for any two initial estimates $(\hat{\mathbf{X}}_0,\mathbf{P}_0)$ and $(\hat{\mathbf{Y}}_0,\mathbf{P}y_0)$, where $\hat{\mathbf{Y}}_0=\mathcal{T}_{\mathbf{g}}(\hat{\mathbf{X}}_0)$ and $\mathbf{P}y_0=\bar{\mathbf{Q}}_1\mathbf{P}_0\bar{\mathbf{Q}}_1^\intercal+\bar{\mathbf{Q}}_2\bar{\mathbf{\Sigma}}\bar{\mathbf{Q}}_2^\intercal$ in which

$$\begin{aligned}\bar{\mathbf{Q}}_1 &= \left.\frac{\partial(\mathcal{T}_{\hat{\mathbf{g}}}(\hat{\mathbf{X}}_0\oplus\mathbf{e})\ominus\mathcal{T}_{\hat{\mathbf{g}}}(\hat{\mathbf{X}}_0))}{\partial\mathbf{e}}\right|_{\mathbf{e}=\mathbf{0}},\\ \bar{\mathbf{Q}}_2 &= \left.\frac{\partial(\mathcal{T}_{\mathbf{g}}(\hat{\mathbf{X}}_0)\ominus\mathcal{T}_{\hat{\mathbf{g}}}(\hat{\mathbf{X}}_0))}{\partial\Theta}\right|_{\Theta=\mathbf{0}},\end{aligned} \tag{11}$$

and $\hat{\mathbf{g}}=(\bar{\mathbf{R}},\bar{\mathbf{T}},\mathbf{0})$, we have $h_n(\hat{\mathbf{X}}_n,\mathbf{0})=h_n(\hat{\mathbf{Y}}_n,\mathbf{0})$ for all $n>0$. The notations $\hat{\mathbf{X}}_n$ and $\hat{\mathbf{Y}}_n$ above represent the *mean* estimate of this filter at time n by using the same input (odometry and observations) from time 0 to n, from the initial conditions $(\hat{\mathbf{X}}_0,\mathbf{P}_0)$ and $(\hat{\mathbf{Y}}_0,\mathbf{P}y_0)$, respectively.

As shown in Def. 1 and Def. 2, the invariance to stochastic rigid body transformation can be divided into two properties: 1) **the invariance to deterministic rigid body transformation** and 2) **the invariance to stochastic identity transformation**. The results about the invariance of some general EKF framework based filters are summarized in Theorem 3 and TABLE I.

Theorem 3: The output of RI-EKF is invariant under stochastic rigid body transformation. The output of $\mathbb{SO}(3)$-EKF is only invariant under deterministic rigid body transformation. The output of Pseudo-RI-EKF is only invariant under stochastic identity transformation. The output of $\mathbb{SE}(3)$-EKF is *not* invariant under deterministic rigid body transformation or stochastic identity transformation.

Proof 3: See Appendix D.

Remark 1: From the proof of Theorem 3, one can see that the uncertainty representation of RI-EKF can be linearly (seamlessly) transformed under a deterministic rigid body transformation, which makes RI-EKF invariant under deterministic rigid body transformation. In addition, we also show that the invariance property to stochastic identity transformation directly depends on the Jacobians $\mathbf{F}_n$ and $\mathbf{H}_n$.

Table I: **The invariance property of the general EKF framework based filters.** DRBT/SRBT is short for "deterministic/stochastic rigid body transformation" and SIT is short for "stochastic identity transformation".

Filters	DRBT	SIT	SRBT
RI-EKF	Yes	Yes	Yes
Pseudo-RI-EKF	No	Yes	No
$\mathbb{SO}(3)$-EKF	Yes	No	No
$\mathbb{SE}(3)$-EKF	No	No	No

B. Consistency and invariance

The unobservability in terms of stochastic rigid body transformation is a fundamental property of SLAM formulation. Therefore a consistent filter (as a system for the estimated state $\hat{\mathbf{X}}$) should maintain this unobservability, i.e., **the (estimated) output of the estimator is invariant under any stochastic rigid body transformation.** Essentially speaking, the invariance to deterministic rigid body transformation can be interpreted that the estimate does *not* depend on the selection of the global frame and the invariance to stochastic identity transformation can be understood that the uncertainty w.r.t the global frame does not affect the estimate. Hence, **consistency for the general EKF framework based filter is tightly coupled with the invariance to stochastic rigid body transformation.** If a filter does not have this property, then unexpected information will be generated by the selection of the global frame or the uncertainty w.r.t. the global frame, which results in inconsistency (overconfidence). One can see that $\mathbb{SO}(3)$-EKF, not invariant to stochastic identity transformation, produces clearly illogical estimate (the pose uncertainty is reduced by the new landmarks) in the two cases of Section IV while RI-EKF, invariant to stochastic rigid body transformation, produces the expected estimate.

Remark 2: In [6] [9], a framework for designing an observability constrained filter is proposed. The keypoint of the observability constrained filter is evaluating the Jacobians $\mathbf{F}_i$ and $\mathbf{H}_i$ $(i\geq 0)$ at some selected points (instead of the latest estimate). In this way, the output of the filter would be invariant under the stochastic identity transformation. On the other hand, this filter models the state space as $\mathbb{SO}(2)\times\mathbb{R}^{2+2N}$ and hence the output is invariant under deterministic rigid body transformation (see the property of $\mathbb{SO}(3)$-EKF shown in Theorem 3). Finally, the resulting filter indeed has the invariance property to stochastic rigid body transformation.

Remark 3: In [21] the observability analysis is performed on the linearized error-state model from the viewpoint of information matrix. Our insight is in a different viewpoint that an estimator should mimic the unobservability (to stochastic rigid transformation) of the original system, which makes our analysis more intuitive and general.

VI. Simulation Results

In order to validate the theoretical results, we perform Monte Carlo simulations and compare RI-EKF to $\mathbb{SO}(3)$-EKF, Robotcentric-EKF, the First Estimates Jacobian EKF SLAM

algorithm (FEJ-EKF), Pseudo-RI-EKF and $\mathbb{SE}(3)$-EKF under conditions of different noise levels. The original Robocentric-EKF and FEJ-EKF are proposed in 2D. For comparison, we extend these into 3D.

A. Settings

Consider that a robot moves in a trajectory (contained in a $50m \times 40m \times 20m$ cubic) which allows sufficient 6-DOFs motion. In this environment, 300 landmarks are randomly generated around the specified robot trajectory. The observations and odometry with noises are randomly generated by this specific trajectory and the simulated robot always observes the landmarks in the sensor range (less than 20m and 120°FoV). In every simulation, the number of steps is 500 (about 8 loops), the landmarks are incrementally added into the state vector and the initial covariance of robot is set as zero matrix. For each condition (different noise level), 100 Monte Carlo simulations are performed. The simulation results are summarized in Fig. 3 and Table II, where σ_{od} is the odometry noise level and σ_{ob} is the observation noise level such that the covariance matrices of odometry and observation and is $\mathbf{\Phi}_n = \sigma_{od}^2 \text{diag}(|\mathbf{u}_{n,1}|^2, \cdots, |\mathbf{u}_{n,6}|^2)$ and $\mathbf{\Psi}_n^i = \sigma_{ob}^2 \text{diag}(|\mathbf{Z}_{n,1}^i|^2, \cdots, |\mathbf{Z}_{n,3}^i|^2)$, where $\mathbf{Z}_n^i = [\mathbf{Z}_{n,1}^i, \mathbf{Z}_{n,2}^i, \mathbf{Z}_{n,3}^i]^\intercal = \mathbf{R}_n^\intercal(\mathbf{f}^i - \mathbf{p}_n)$ is the ground truth of the coordinates of landmark i relative to the robot pose n. The root mean square (RMS) error and the average normalized estimation error squared (NEES) are used to evaluate accuracy and consistency, respectively.

B. Results and analysis

As shown in Table II, the estimate of $\mathbb{SE}(3)$-EKF diverges even under the condition of low noise ($\sigma_{od} = 1\%$, $\sigma_{ob} = 1\%$) and Pseudo-RI-EKF is also poor performing. These results can be understood because $\mathbb{SE}(3)$-EKF has no invariance property to deterministic rigid body transformation or stochastic identity transformation and Pseudo-RI-EKF is not invariant under deterministic rigid body transformation, which are proven in Theorem 3. $\mathbb{SO}(3)$-EKF, not invariant to stochastic identity transformation, is also not good performing in terms of consistency.

An interesting point in Table II is the performance of Robocentric-EKF. The uncertainty representation w.r.t landmarks in Robocentric-EKF refers to the uncertainty of the coordinates of landmarks relative to robot frame. In this way, Robocentric-EKF has the invariance property to stochastic rigid body transformation. However, Robocentric-EKF does not perform well under the condition of high noise ($\sigma_{od} = 5\%$, $\sigma_{ob} = 5\%$) because it incurs greater linearization errors in the propagation step due to the coordinate transformation applied to the landmarks, as compared to $\mathbb{SO}(3)$-EKF, FEJ-EKF and RI-EKF.

RI-EKF, invariant to stochastic rigid body transformation, is the best performing filter as shown in Table II and it is also consistent in terms of the 95% confidence bound as shown in Fig. 3. Similar results for 2D cases have been reported in [21]. On the other hand, it is discussed in *Remark* 2 of Section V that FEJ-EKF also has the invariance property to stochastic rigid body transformation but it performs less well than RI-EKF. It can be explained that FEJ-EKF uses a less accurate estimate as linearization point for evaluating the Jacobians while RI-EKF can always safely employ the latest estimate in Jacobians.

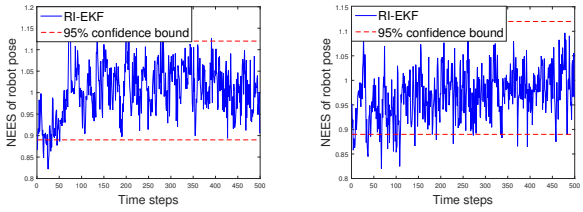

Figure 3: **Average NEES of robot pose by RI-EKF from 100 Monte Carlo results.** The 95% confidence bound is $[0.89, 1.12]$. Left: $\sigma_{od} = 1\%$, $\sigma_{ob} = 1\%$. Right: $\sigma_{od} = 5\%$, $\sigma_{ob} = 5\%$.

VII. Conclusion

In this work, the convergence properties and consistency of a Lie group EKF based SLAM algorithm (RI-EKF) are analyzed. For convergence, several theorems with proofs are provided for two fundamental cases. For consistency, we propose that consistency of the general EKF framework based filter is tightly coupled with the invariance property. We also proves that the output of RI-EKF is invariant under stochastic rigid body transformation while the output of $\mathbb{SO}(3)$-EKF is only invariant under deterministic rigid body transformation. Monte Carlo simulation results demonstrates that the invariance property has an important impact on the consistency and accuracy of the estimator and RI-EKF outperforms $\mathbb{SO}(3)$-EKF, $\mathbb{SE}(3)$-EKF, Robocentric-EKF and FEJ-EKF. Future work includes extensively comparing the performance of RI-EKF SLAM algorithm with the optimization based SLAM algorithms to identify situations under which RI-EKF is sufficient, as well as extending RI-EKF to the case of visual-inertial fusion.

Appendix

A. Lie Group $\mathcal{G}(N)$

The notation $\mathcal{G}(N)$ is a Lie group, defined as

$$\mathbf{G} = \{(\mathbf{R}, \mathbf{p}, \mathbf{f}^1, \cdots, \mathbf{f}^N) \,|\mathbf{R} \in \mathbb{SO}(3), \mathbf{p} \text{ and } \mathbf{f}^i \in \mathbb{R}^3\}. \tag{12}$$

The associated group operation of $\mathcal{G}(N)$ is

$$\mathbf{X}_1\mathbf{X}_2 = \left(\mathbf{R}_1\mathbf{R}_2, \mathbf{R}_1\mathbf{p}_2 + \mathbf{p}_1, \cdots, \mathbf{R}_1\mathbf{f}_2^N + \mathbf{f}_1^N\right), \tag{13}$$

where $\mathbf{X}_i = \left(\mathbf{R}_i, \mathbf{p}_i, \mathbf{f}_i^1, \cdots, \mathbf{f}_i^N\right) \in \mathcal{G}(N)$ for $i = 1, 2$. The associated Lie algebra of $\mathcal{G}(N)$ is homomorphic to $\mathbb{R}^{3N+6}$. The exponential mapping exp is represented as

$$\begin{aligned} &\exp(\mathbf{e}) \in \mathcal{G}(N) \\ &= \left(\exp(\mathbf{e}_\theta), J_r(-\mathbf{e}_\theta)\mathbf{e}_p, J_r(-\mathbf{e}_\theta)\mathbf{e}^1, \cdots, J_r(-\mathbf{e}_\theta)\mathbf{e}^N\right) \end{aligned} \tag{14}$$

for $\mathbf{e} = \begin{bmatrix} \mathbf{e}_\theta^\intercal & \mathbf{e}_p^\intercal & (\mathbf{e}^1)^\intercal & \cdots & (\mathbf{e}^N)^\intercal \end{bmatrix}^\intercal \in \mathbb{R}^{3N+6}$, where $\mathbf{e}_\theta$, $\mathbf{e}_p$ and $\mathbf{e}^i \in \mathbb{R}^3$ $(i = 1, \cdots, N)$, the notation exp in the right side of (14) and the mapping J_r are given:

$$\exp(\mathbf{y}) = \mathbf{I}_3 + \frac{\sin(\|\mathbf{y}\|)}{\|\mathbf{y}\|} S(\mathbf{y}) + \frac{1 - \cos(\|\mathbf{y}\|)}{\|\mathbf{y}\|^2} S^2(\mathbf{y}) \tag{15}$$

Table II: **Performance Evaluation**

$\sigma_{od}=1\%$, $\sigma_{ob}=1\%$	RI-EKF	FEJ-EKF	$\mathbb{SO}(3)$-EKF	Robocentric-EKF	Pseudo-RI-EKF	$\mathbb{SE}(3)$-EKF
RMS of position(m)	**0.25**	0.29	0.32	0.31	0.65	Diverge
RMS of orientation(rad)	**0.0058**	0.0071	0.0065	0.0060	0.0081	Diverge
NEES of orientation	**1.02**	1.12	1.34	1.04	2.91	Diverge
NEES of pose	**1.01**	1.14	1.35	1.15	10	Diverge
$\sigma_{od}=5\%$, $\sigma_{ob}=5\%$	RI-EKF	FEJ-EKF	$\mathbb{SO}(3)$-EKF	Robocentric-EKF	Pseudo-RI-EKF	$\mathbb{SE}(3)$-EKF
RMS of position(m)	**1.16**	1.24	2.0	2.4	3.90	Diverge
RMS of orientation(rad)	**0.027**	0.029	0.043	0.041	0.041	Diverge
NEES of orientation	**1.0**	1.05	3.7	3.0	1.77	Diverge
NEES of pose	**1.01**	1.13	3.1	7.5	92	Diverge

$$J_r(\mathbf{y}) = \mathbf{I}_3 - \frac{1-\cos(\|\mathbf{y}\|)}{\|\mathbf{y}\|^2}S(\mathbf{y}) + \frac{\|\mathbf{y}\|-\sin(\|\mathbf{y}\|)}{\|\mathbf{y}\|^3}S^2(\mathbf{y}) \quad (16)$$

for $\mathbf{y}\in\mathbb{R}^3$. The adjoint $\mathrm{ad}_{\mathbf{X}}$ is computed as

$$\mathrm{ad}_{\mathbf{X}} = \begin{bmatrix} \mathbf{R} & \mathbf{0}_{3,3} & \cdots & \cdots & \mathbf{0}_{3,3} \\ S(\mathbf{p})\mathbf{R} & \mathbf{R} & \ddots & & \vdots \\ S(\mathbf{f}^1)\mathbf{R} & \mathbf{0}_{3,3} & \mathbf{R} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \mathbf{0}_{3,3} \\ S(\mathbf{f}^N)\mathbf{R} & \mathbf{0}_{3,3} & \cdots & \mathbf{0}_{3,3} & \mathbf{R} \end{bmatrix}. \quad (17)$$

B. Proof of Theorem 1

In the following, we use mathematical induction to prove this theorem. Note that At the beginning, the estimate is $(\hat{\mathbf{X}},\mathbf{P})$ where $\hat{\mathbf{X}} = \left(\hat{\mathbf{R}},\hat{\mathbf{p}},\hat{\mathbf{f}}^1,\cdots,\hat{\mathbf{f}}^N\right)$. After the first observation, the mean estimate of state and covariance matrix are augmented as below via the method shown in Alg. 2: $\hat{\mathbf{X}}_1 = \left(\hat{\mathbf{R}},\hat{\mathbf{p}},\hat{\mathbf{f}}^1,\cdots,\hat{\mathbf{f}}^N,\hat{\mathbf{f}}^{N+1}\right)$ and $\mathbf{P}_1 = \begin{bmatrix} \mathbf{P} & \mathbf{L} \\ \mathbf{L}^\intercal & \hat{\mathbf{R}}\boldsymbol{\Psi}\hat{\mathbf{R}}^\intercal + \mathbf{W} \end{bmatrix}$. Obviously, after one observation, the mean estimate of robot pose and the previous "landmarks" does not change and the covariance matrix follows the proposed form. We now assume that after k times observations, the estimate becomes $\hat{\mathbf{X}}_k = \left(\hat{\mathbf{R}},\hat{\mathbf{p}},\hat{\mathbf{f}}^1,\cdots,\hat{\mathbf{f}}^N,\hat{\mathbf{f}}_k^{N+1}\right)$ and $\mathbf{P}_k = \begin{bmatrix} \mathbf{P} & \mathbf{L} \\ \mathbf{L}^\intercal & \frac{\hat{\mathbf{R}}\boldsymbol{\Psi}\hat{\mathbf{R}}^\intercal}{k} + \mathbf{W} \end{bmatrix}$. Now we discuss the case after k times observations of next propagation and update. Because the robot is always perfectly stationary, after propagation at time k, the mean estimate is $\hat{\mathbf{X}}_{k+1|k} = \hat{\mathbf{X}}_k$ and covariance matrix becomes $\mathbf{P}_{k+1|k} = \mathbf{P}_k$. According to Alg. 1, we have $\mathbf{S} = \mathbf{H}\mathbf{P}_{k+1|k}\mathbf{H}^\intercal + \boldsymbol{\Psi} = \frac{k+1}{k}\boldsymbol{\Psi}$ and $\mathbf{K} = \mathbf{P}_{k+1|k}\mathbf{H}^\intercal\mathbf{S}^{-1} = \begin{bmatrix} \mathbf{0}_{3,(3N+6)} & -\frac{1}{k+1}\hat{\mathbf{R}}^\intercal \end{bmatrix}^\intercal$, where $\mathbf{H} = \begin{bmatrix} \mathbf{0}_{3,3} & \hat{\mathbf{R}}^\intercal & \mathbf{0}_{3,3N} & -\hat{\mathbf{R}}^\intercal \end{bmatrix}$. Then it is easy to see that all elements from the vector $\mathbf{K}\mathbf{y}$ are zero except the last 3 elements, and hence the estimate of robot pose and the old landmarks after $k+1$ times observations are **the same** as that in the time step k. The covariance matrix at time $k+1$ is $\mathbf{P}_{k+1} = (\mathbf{I}-\mathbf{K}\mathbf{H})\mathbf{P}_{k+1|k} = \begin{bmatrix} \mathbf{P} & \mathbf{L} \\ \mathbf{L}^\intercal & \frac{\hat{\mathbf{R}}\boldsymbol{\Psi}\hat{\mathbf{R}}^\intercal}{k+1} + \mathbf{W} \end{bmatrix}$. When k converges to infinity, we have (6).

C. Proof of Theorem 2

By using result in Theorem 1 and the Jacobian matrices in (4), we have

$$\mathbf{P}_B^0 = \mathbf{P}_A^\infty + \Delta\mathbf{P}, \quad (18)$$

where $\mathbf{P}_A^\infty$ (given in (7)) is the covariance matrix before moving to the point B, $\Delta\mathbf{P} = ad_{\hat{\mathbf{X}}_A}\mathbf{E}\tilde{\boldsymbol{\Phi}}\mathbf{E}^\intercal ad_{\hat{\mathbf{X}}_A}^\intercal$ can be regarded as the incremental uncertainty caused by the odometry noise, and $\tilde{\boldsymbol{\Phi}} = \mathbf{B}\boldsymbol{\Phi}\mathbf{B}^\intercal$ is a positive definite matrix.

After l observations at point B, the information matrix $\boldsymbol{\Omega}_l^B$ (the inverse of $\mathbf{P}_l^B$) becomes $\boldsymbol{\Omega}_l^B = \boldsymbol{\Omega}_0^B + \sum_{j=1}^l \mathbf{H}_j^\intercal\bar{\boldsymbol{\Psi}}^{-1}\mathbf{H}_j$, where $\mathbf{H}_j$ is obtained by stacking all matrices $\mathbf{H}_j^i = \hat{\mathbf{R}}_j^\intercal\mathbf{H}^i$ $(i=1,\cdots,m)$, and $\hat{\mathbf{R}}_j$ is the estimated orientation after j times observations at point B. Note that $\bar{\boldsymbol{\Psi}}$ is isotropic, we have $\mathbf{H}_j^\intercal\bar{\boldsymbol{\Psi}}^{-1}\mathbf{H}_j = \mathbf{H}^\intercal\bar{\boldsymbol{\Psi}}^{-1}\mathbf{H}$ $(j=1,\cdots,l)$. Therefore, the information matrix is $\boldsymbol{\Omega}_l^B = \boldsymbol{\Omega}_0^B + l\mathbf{H}^\intercal\bar{\boldsymbol{\Psi}}^{-1}\mathbf{H}$. Via the matrix inversion lemma in [2], the covariance matrix after l observations at point B is

$$\mathbf{P}_B^l = (\boldsymbol{\Omega}_l^B)^{-1} = \mathbf{P}_B^0 - \mathbf{P}_B^0\mathbf{H}^\intercal\left(\frac{\bar{\boldsymbol{\Psi}}}{l} + \mathbf{H}\mathbf{P}_B^0\mathbf{H}^\intercal\right)^{-1}\mathbf{H}\mathbf{P}_B^0. \quad (19)$$

Note that $\mathbf{H}\mathbf{P}_A^\infty = \mathbf{0}$, we substitute (18) into (19):

$$\begin{aligned} \mathbf{P}_B^l &= \mathbf{P}_A^\infty + \Delta\mathbf{P} - \Delta\mathbf{P}\mathbf{H}^\intercal\left(\frac{\bar{\boldsymbol{\Psi}}}{l} + \mathbf{H}\Delta\mathbf{P}\mathbf{H}^\intercal\right)^{-1}\mathbf{H}\Delta\mathbf{P} \\ &= \mathbf{P}_A^\infty + ad_{\hat{\mathbf{X}}_A}\mathbf{E}\left(\tilde{\boldsymbol{\Phi}}^{-1} + l\tilde{\mathbf{H}}^\intercal\bar{\boldsymbol{\Psi}}^{-1}\tilde{\mathbf{H}}\right)^{-1}\mathbf{E}^\intercal ad_{\hat{\mathbf{X}}_A}^\intercal \\ &= \mathbf{P}_A^\infty + \bar{\mathbf{P}}_B^l. \end{aligned} \quad (20)$$

Furthermore, $\tilde{\mathbf{H}}^\intercal\bar{\boldsymbol{\Psi}}^{-1}\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{S}_1 & \mathbf{S}_2 \\ \mathbf{S}_2^\intercal & m\boldsymbol{\Psi}^{-1} \end{bmatrix}$ where $\mathbf{S}_1 = \sum_{i=1}^m S^\intercal(\tilde{\mathbf{f}}_i)\boldsymbol{\Psi}^{-1}S(\tilde{\mathbf{f}}_i)$, $\mathbf{S}_2 = (\sum_{i=1}^m S(\tilde{\mathbf{f}}_i))^\intercal\boldsymbol{\Psi}^{-1}$ and $\tilde{\mathbf{f}}_i = \hat{\mathbf{R}}^\intercal(\hat{\mathbf{p}}-\hat{\mathbf{f}}_i)$ $(i=1,\cdots,m)$. Generally speaking, $\tilde{\mathbf{H}}^\intercal\boldsymbol{\Psi}^{-1}\tilde{\mathbf{H}}$ is full rank when $m>3$ and there are three landmarks that are non-coplanar with the robot position. Under this condition, it is easy to see that $\mathbf{P}_B^l \to \mathbf{P}_A^\infty$ when $l\to\infty$.

D. Proof of Theorem 3

Here, we only prove that the invariance property of RI-EKF and $\mathbb{SO}(3)$-EKF. The invariance properties of the other algorithms can be easily proven in a similar way or through a counter example.

First, we prove that the outputs of $\mathbb{SO}(3)$-EKF and RI-EKF is invariant to deterministic rigid body transformation. Assume the estimate at time 0 is $(\hat{\mathbf{X}}_0,\mathbf{P}_0)$ in terms of the general EKF framework. After one step propagation via the odometry $\mathbf{u}_0$, the estimate becomes $(\hat{\mathbf{X}}_{1|0},\mathbf{P}_{1|0})$. Then after obtaining observations $\mathbf{z}_1$, the estimate becomes $(\mathbf{X}_1,\mathbf{P}_1)$. On the other hand, in $\mathbb{SO}(3)$-EKF and RI-EKF, there exists a matrix $\mathbf{Q}_{\mathcal{T}}$ for any rigid body transformation $\mathcal{T}$ such that

$$\mathcal{T}(\mathbf{X}\oplus\mathbf{Q}_{\mathcal{T}}^{-1}\mathbf{e}) = \mathcal{T}(\mathbf{X})\oplus\mathbf{e} \quad \forall\ \mathbf{X}. \quad (21)$$

Therefore, if a deterministic rigid body transformation $\mathcal{T}$ is applied at time 0, the estimate becomes $(\hat{\mathbf{Y}}_0,\mathbf{P}y_0)$, where

$\hat{\mathbf{Y}}_0 = \mathcal{T}(\hat{\mathbf{X}}_0)$ and $\mathbf{P}y_0 = \mathbf{Q}_\mathcal{T}\mathbf{P}_0\mathbf{Q}_\mathcal{T}^\intercal$. Now we calculate the new Jacobians $\mathbf{F}y_0$ and $\mathbf{G}y_0$ in propagation

$$\begin{aligned}
\mathbf{F}y_0 &= \left.\frac{\partial f(\hat{\mathbf{Y}}_0 \oplus \mathbf{e}, \mathbf{u}_0, \mathbf{0}) \ominus f(\hat{\mathbf{Y}}_0, \mathbf{u}_0, \mathbf{0})}{\partial \mathbf{e}}\right|_{\mathbf{0}} \\
&= \left.\frac{\partial f(\mathcal{T}(\hat{\mathbf{X}}_0) \oplus \mathbf{e}, \mathbf{u}_0, \mathbf{0}) \ominus f(\mathcal{T}(\hat{\mathbf{X}}_0), \mathbf{u}_0, \mathbf{0})}{\partial \mathbf{e}}\right|_{\mathbf{0}} \\
&\overset{(21)}{=} \left.\frac{\partial f(\mathcal{T}(\hat{\mathbf{X}}_0 \oplus \mathbf{Q}_\mathcal{T}^{-1}\mathbf{e}), \mathbf{u}_0, \mathbf{0}) \ominus f(\mathcal{T}(\hat{\mathbf{X}}_0), \mathbf{u}_0, \mathbf{0})}{\partial \mathbf{e}}\right|_{\mathbf{0}} \\
&= \left.\frac{\partial \mathcal{T}(f(\hat{\mathbf{X}}_0 \oplus \mathbf{Q}_\mathcal{T}^{-1}\mathbf{e}, \mathbf{u}_0, \mathbf{0})) \ominus \mathcal{T}(f(\hat{\mathbf{X}}_0, \mathbf{u}_0, \mathbf{0}))}{\partial \mathbf{e}}\right|_{\mathbf{0}} \\
&= \left.\frac{\partial \mathcal{T}(f(\hat{\mathbf{X}}_0, \mathbf{u}_0, \mathbf{0}) \oplus \mathbf{F}_0\mathbf{Q}_\mathcal{T}^{-1}\mathbf{e}) \ominus \mathcal{T}(f(\hat{\mathbf{X}}_0, \mathbf{u}_0, \mathbf{0}))}{\partial \mathbf{e}}\right|_{\mathbf{0}} \\
&\overset{(21)}{=} \mathbf{Q}_\mathcal{T}\mathbf{F}_0\mathbf{Q}_\mathcal{T}^{-1}.
\end{aligned} \tag{22}$$

Similarly, we have $\mathbf{G}y_0 = \mathbf{Q}_\mathcal{T}\mathbf{G}_0$. Hence, after one step propagation the estimate becomes $(\hat{\mathbf{Y}}_{1|0}, \mathbf{P}y_{1|0})$, where $\hat{\mathbf{Y}}_{1|0} = f(\hat{\mathbf{Y}}_0, \mathbf{u}_0, \mathbf{0}) = \mathcal{T}(\hat{\mathbf{X}}_{1|0})$ and $\mathbf{P}y_{1|0} = \mathbf{F}y_0\mathbf{P}y_0\mathbf{F}y_0^\intercal + \mathbf{G}y_0\mathbf{\Phi}_0\mathbf{G}y_0^\intercal = \mathbf{Q}_\mathcal{T}\mathbf{P}_{1|0}\mathbf{Q}_\mathcal{T}^\intercal$. The new Jacobians in update becomes $\mathbf{H}y_1 = \mathbf{H}_1\mathbf{Q}_\mathcal{T}^{-1}$. Then it is easy to obtain $\mathbf{K}_y = \mathbf{Q}_\mathcal{T}\mathbf{K}$, resulting in $\hat{\mathbf{Y}}_1 = \hat{\mathbf{Y}}_{1|0} \oplus \mathbf{K}_y\mathbf{y} = \mathcal{T}(\hat{\mathbf{X}}_{1|0}) \oplus \mathbf{Q}_\mathcal{T}\mathbf{K}\mathbf{y} = \mathcal{T}(\hat{\mathbf{X}}_{1|0} \oplus \mathbf{K}\mathbf{y}) = \mathcal{T}(\hat{\mathbf{X}}_1)$. The covariance matrix after update becomes $\mathbf{P}y_1 = (\mathbf{I} - \mathbf{K}_y\mathbf{H}y_1)\mathbf{P}y_{1|0} = \mathbf{Q}_\mathcal{T}\mathbf{P}_1\mathbf{Q}_\mathcal{T}^\intercal$. In all, $\hat{\mathbf{Y}}_1 = \mathcal{T}(\hat{\mathbf{X}}_1)$ and $\mathbf{P}y_1 = \mathbf{Q}_\mathcal{T}\mathbf{P}_1\mathbf{Q}_\mathcal{T}^\intercal$. By mathematical induction, we can see the outputs of $\mathbb{SO}(3)$-EKF (and RI-EKF) are invariant under deterministic rigid body transformation.

Secondly, we prove the invariance property of RI-EKF under stochastic identity body transformation $\mathcal{T}_\mathbf{g}$ ($\mathbf{g} = (\mathbf{I}_3, \mathbf{0}, \Theta)$) for all $\bar{\mathbf{\Sigma}}$ where $\bar{\mathbf{\Sigma}}$ is the covariance matrix of noise Θ. Consider the estimate at time 0 is $(\hat{\mathbf{X}}_0, \mathbf{P}_0)$ in RI-EKF. If the stochastic rigid body transformation $\mathcal{T}_\mathbf{g}$ is applied, the estimate becomes $(\hat{\mathbf{X}}_0, \mathbf{P}_0 + \Delta\mathbf{P})$ where $\Delta\mathbf{P} = \mathbf{C}\bar{\Sigma}\mathbf{C}^\intercal$ and

$$\mathbf{C} = \frac{\partial(\mathcal{T}_\mathbf{g}(\hat{\mathbf{X}}_0) \ominus (\hat{\mathbf{X}}_0))}{\partial \Theta}|_{\mathbf{0}} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0}_{3,3} \\ \mathbf{0}_{3,3} & \mathbf{I}_3 \\ \vdots & \vdots \\ \mathbf{0}_{3,3} & \mathbf{I}_3 \end{bmatrix}. \tag{23}$$

After propagation, the estimate becomes $(\hat{\mathbf{X}}_{1|0}, \mathbf{P}_{1|0} + \Delta\mathbf{P})$ due to $\mathbf{F}_n = \mathbf{I}$ given in (4). Note that $\mathbf{H}_1\Delta\mathbf{P} = \mathbf{0}$, it is easy to get the posterior estimate $(\hat{\mathbf{X}}_1, \mathbf{P}_1 + \Delta\mathbf{P})$. By mathematical induction, we can conclude that the output of RI-EKF is invariant under stochastic identity transformation.

REFERENCES

[1] S. J. Julier and J. K. Uhlmann, "A counter example to the theory of simultaneous localization and map building," in *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, vol. 4, 2001, pp. 4238–4243 vol.4.

[2] S. Huang and G. Dissanayake, "Convergence and consistency analysis for extended kalman filter based slam," *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 1036–1049, Oct 2007.

[3] T. Bailey, J. Nieto, J. Guivant, M. Stevens, and E. Nebot, "Consistency of the ekf-slam algorithm," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2006, pp. 3562–3568.

[4] J. A. Castellanos, J. Neira, and J. D. Tardós, "Limits to the consistency of ekf-based slam," 2004.

[5] P. Lourenço, B. J. Guerreiro, P. Batista, P. Oliveira, and C. Silvestre, "Simultaneous localization and mapping for aerial vehicles: a 3-d sensor-based gas filter," *Autonomous Robots*, vol. 40, no. 5, pp. 881–902, 2016.

[6] G. P. Huang, A. I. Mourikis, and S. I. Roumeliotis, "Analysis and improvement of the consistency of extended kalman filter based slam," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, May 2008, pp. 473–479.

[7] J. Andrade-Cetto and A. Sanfeliu, "The effects of partial observability in slam," in *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, vol. 1, April 2004, pp. 397–402 Vol.1.

[8] K. W. Lee, W. S. Wijesoma, and J. I. Guzman, "On the observability and observability analysis of slam," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct 2006, pp. 3569–3574.

[9] G. P. Huang, A. I. Mourikis, and S. I. Roumeliotis, "Observability-based rules for designing consistent ekf slam estimators," *The International Journal of Robotics Research*, vol. 29, no. 5, pp. 502–528, 2010.

[10] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, "Consistency analysis and improvement of vision-aided inertial navigation," *IEEE Transactions on Robotics*, vol. 30, no. 1, pp. 158–176, Feb 2014.

[11] G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (slam) problem," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 3, pp. 229–241, Jun 2001.

[12] A. Mourikis and S. Roumeliotis, "Analytical characterization of the accuracy of slam without absolute orientation measurements," in *Proceedings of Robotics: Science and Systems*, Philadelphia, USA, August 2006.

[13] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual–inertial odometry," *IEEE Transactions on Robotics*, vol. PP, no. 99, pp. 1–21, 2016.

[14] T. D. Barfoot and P. T. Furgale, "Associating uncertainty with three-dimensional poses for use in estimation problems," *IEEE Transactions on Robotics*, vol. 30, no. 3, pp. 679–693, June 2014.

[15] R. Mahony, T. Hamel, and J. M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1203–1218, June 2008.

[16] L. Carlone, V. Macchia, F. Tibaldi, and B. Bona, "Quaternion-based ekf-slam from relative pose measurements: observability analysis and applications," *Robotica*, vol. 33, no. 06, pp. 1250–1280, 2015.

[17] G. Dubbelman and B. Browning, "Cop-slam: Closed-form online pose-chain optimization for visual slam," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1194–1213, Oct 2015.

[18] C. Hertzberg, R. Wagner, U. Frese, and L. Schrder, "Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds," *Information Fusion*, vol. 14, no. 1, pp. 57 – 77, 2013.

[19] N. Aghannan and P. Rouchon, "On invariant asymptotic observers," in *Proceedings of the 41st IEEE Conference on Decision and Control, 2002.*, vol. 2, Dec 2002, pp. 1479–1484 vol.2.

[20] S. Bonnabel, "Symmetries in observer design: review of some recent results and applications to EKF-based SLAM," *ArXiv e-prints*.

[21] A. Barrau and S. Bonnabel, "An ekf-slam algorithm with consistency properties," *arXiv preprint arXiv:1510.06263*, 2015.

[22] P.-A. Absil, C. G. Baker, and K. A. Gallivan, "Trust-region methods on riemannian manifolds," *Foundations of Computational Mathematics*, vol. 7, no. 3, pp. 303–330, 2007.

[23] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Transactions on Automatic Control*, vol. 22, no. 5, pp. 728–740, Oct 1977.

[24] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, "Camera-imu-based localization: Observability analysis and consistency improvement," *The International Journal of Robotics Research*, vol. 33, no. 1, pp. 182–201, 2014.