Extended Information Filter on Matrix Lie Groups ⋆

Josip Ćesić, Ivan Marković, Mario Bukal, Ivan Petrović [a]

[a] *University of Zagreb Faculty of Electrical Engineering and Computing, Unska 3, HR-10000, Zagreb, Croatia*

Abstract

In this paper we propose a new state estimation algorithm called the extended information filter on Lie groups. The proposed filter is inspired by the extended Kalman filter on Lie groups and exhibits the advantages of the information filter with regard to multisensor update and decentralization, while keeping the accuracy of stochastic inference on Lie groups. We present the theoretical development and demonstrate its performance on multisensor rigid body attitude tracking by forming the state space on the $\mathsf{SO}(3) \times \mathbb{R}^3$ group, where the first and second component represent the orientation and angular rates, respectively. The performance of the filter is compared with respect to the accuracy of attitude tracking with parametrization based on Euler angles and with respect to execution time of the extended Kalman filter formulation on Lie groups. The results show that the filter achieves higher performance consistency and smaller error by tracking the state directly on the Lie group and that it keeps smaller computational complexity of the information form with respect to high number of measurements.

Key words: Extended Kalman filters; Information filter; Lie groups.

1 Introduction

The information filter (IF) is the dual of the Kalman filter (KF) relying on the state representation by a Gaussian distribution [1], and hence is the subject of the same assumptions underlying the KF. Whereas the KF family of algorithms is represented by the first two moments involving the mean and covariance, the IF relies on the canonical parametrization consisting of an information matrix and information vector [2]. Both the KF and IF operate cyclically in two steps: the prediction and update step. The advantages of the IF lie in the update step, especially when the number of measurements is significantly larger than the size of the state space, since this step is additive for the IF. For the KF, the opposite applies; it is the prediction step which is additive and computationally less complex. What is computationally complex in one parametrization turns out to be simple in the other (and vice-versa) [3].

Given this duality, the IF has proven its mettle in a number of applications facing large number of measurements, features or demanding a decentralized filter form. For example, if the system is linear and the state is modeled as Gaussian, then multisensor fusion can be performed with the decentralized KF proposed in [4], which enables fusion of not only the measurements, but also of the local independent KFs. Therein, the inverse covariance form is utilized, thus resulting in additive fusion equations, which can further be elegantly translated to the IF form as shown in [5]. In [6] an IF is presented for robust decentralized estimation based on the robustness property of the H_∞ filter with respect to noise statistics, whereas in [7] stability of consensus extended Kalman filter for distributed state estimation was investigated. In [8] collaborative target tracking is developed for wireless sensor networks and a mutual-information-based sensor selection is adopted for participation in the IF form fusion process. In [9] the IF form is used in multitarget tracking sensor allocation based on solving a constrained optimization problem. In [10] a sigma-point IF was used for decentralized target tracking, in [11] a square root form of the same filter was used for cooperative

⋆ This paper was not presented at any IFAC meeting.
Corresponding author I. Marković, Tel. +3851-6129-577. Fax +3851-6129-795.
This work has been supported from the Unity Through Knowledge Fund under the project Cooperative Cloud based Simultaneous Localization and Mapping in Dynamic Environments (cloudSLAM) and the research has also been carried out within the activities of the Centre of Research Excellence for Data Science and Cooperative Systems supported by the Ministry of Science, Education and Sports of the Republic of Croatia.

Email addresses: `josip.cesic@fer.hr` (Josip Ćesić), `ivan.markovic@fer.hr` (Ivan Marković), `mario.bukal@fer.hr` (Mario Bukal), `ivan.petrovic@fer.hr` (Ivan Petrović).

tracking with unmanned aerial vehicles, and in [12, 13] square-root information filtering was further explored with respect to numerical stability. The unscented IF was presented in [14] for tracking of a re-entry vehicle entering into an atmosphere from space, and in [15] the square root cubature IF was proposed and demonstrated on the example of speed and rotor position estimation of a two phase permanent magnet synchronous motor.

Another important aspect of estimation is the state space geometry, hence many works have been dedicated to dealing with uncertainty and estimation techniques accounting for it. For example, Lie groups are natural ambient (state) spaces for description of the dynamics of rigid body mechanical systems [16, 17]. Furthermore, error propagation on the $\mathsf{SE}(3)$ group with applications to manipulator kinematics was presented in [18] by developing closed-form solutions for the convolution of the concentrated Gaussian distributions on $\mathsf{SE}(3)$. Furthermore, in [19] the authors propose a solution to Bayesian fusion on Lie groups by assuming conditional independence of observations on the group, thus setting the fusion result as a product of concentrated Gaussian distributions, and finding the single concentrated Gaussian distribution parameters which are closest to the starting product. Uncertainty association, propagation and fusion on $\mathsf{SE}(3)$ was investigated in [20] along with sigma point method for uncertainty propagation through a nonlinear camera model. In [21] the authors preintegrated a large number of inertial measurement unit measurements for visual-inertial navigation into a single relative motion constraint by respecting the structure of the $\mathsf{SO}(3)$ group and defining the uncertainty thereof in the pertaining tangent space. A state estimation method based on an observer and a predictor cascade for invariant systems on Lie groups with delayed measurements was proposed in [22]. Recently, some works have also addressed the uncertainty on the $\mathsf{SE}(2)$ group proposing new distributions [23, 24]; however, these approaches do not yet provide a closed-form Bayesian recursion framework (involving both the prediction and update) that can include higher order motion and non-linear models. A least squares optimization and nonlinear KF on manifolds in the vein of the unscented KF was proposed in [25] along with an accompanying software library. Therein the authors demonstrate the filter on a synthetic dataset addressing the problem of trajectory estimation by posing the system state to reside on the manifold $\mathbb{R}^3 \times \mathsf{SO}(3) \times \mathbb{R}^3$, i.e., the position, orientation and velocity. In the end, the authors also demonstrate the approach on real-world simultaneous localization and mapping (SLAM) data and perform pose relation graph optimization. In the vein of the extended Kalman filter (EKF) a nonlinear continuous-discrete extended Kalman filter on Lie groups (LG-EKF) was proposed in [26]. Therein, the prediction step is presented in the continuous domain, while the update step is discrete. The authors have demonstrated the efficiency of the filter on a synthetic camera pose filtering problem by forming the system state to reside on the $\mathsf{SO}(3) \times \mathbb{R}^9$ group, i.e. the camera orientation, position, angular and radial velocities. In an earlier publication [27], the authors have presented a discrete version of the LG-EKF, which servers as the inspiration for the filter proposed in the present paper. In [28] we have explored modeling of the pose of tracked objects on the $\mathsf{SE}(2)$ group within the LG-EKF framework, and applied it on the problem of multitarget tracking by fusing a radar sensor and stereo vision. Given the advantages of the IF and filtering on Lie groups, a natural question arises; Can LG-EKF be cast in the information form and will the corresponding information filter on Lie groups keep the additivity and computational advantages of the update step?

A quite prominent example of an application where the need arises for computational benefits of the IF and the geometric accuracy of Lie groups is SLAM. SLAM is of great practical importance in many robotic and autonomous system applications and the earliest solutions were based on the EKF. However, EKF in practice can handle maps that contain a few hundred features, while in many applications maps are orders of magnitude larger [29]. Therefore, the extended information filter (EIF) is often employed and widely accepted for SLAM [30], and has reached its zenith with sparsification approaches resulting with sparse EIF (SEIF) [29] and exactly sparse delayed-state filter (ESDF) [31]. However, the localization component of SLAM conforms the pose estimation problem as arising on Lie groups, i.e., describing the pose in the special euclidean group $\mathsf{SE}(3)$ [20]. Furthermore, the mapping part of SLAM consists of landmarks whose position, as well, arises on $\mathsf{SE}(3)$. Therefore, some recent SLAM solutions approached the problem by respecting the geometry of the state space [32, 33], since significant cause of error in such application was determined to stem from the state space geometry approximations. However, these SLAM solutions, although able to account for the geometry of the state space, exclusively rely on graph optimization [34, 35], but not on filtering approaches. By using the herein proposed algorithm, one can extend the SLAM filtering approaches, such as SEIF or ESDF, and at the same time respect the geometry of the state space via formulation on Lie groups.

The main contribution of this paper is a new state estimation algorithm called the extended information filter on Lie groups (LG-EIF), which exhibits the advantages of the IF with regard to multisensor update and decentralization, while keeping the accuracy of the LG-EKF for stochastic inference on Lie groups. We present the theoretical development of the LG-EIF recursion equations and the applicability of the proposed approach is demonstrated on a rigid body attitude tracking problem with multiple sensors. In the experiments we define the state space to reside on the Cartesian product of the special orthogonal group $\mathsf{SO}(3)$ and $\mathbb{R}^3$, with the first component representing the attitude of the rigid body and the second component representing the pertaining angular rates. Given that, the model of the system is then set as a constant angular rate model acting on the state space $\mathsf{SO}(3) \times \mathbb{R}^3$. Note that, just like the LG-EKF, the proposed filter can be applied on any matrix Lie group or combination thereof. In the end, we compare the proposed LG-EIF to an EIF based on Euler angles, and we analyze the computational complexity of the LG-EIF multisensor update with respect to the LG-EKF. The results show that the proposed filter achieves higher performance consistency and smaller error by tracking the state directly on the Lie group and that it keeps smaller computational complexity of the information form with respect to large number of measurements.

The rest of the paper is organized as follows. In Section 2 we present the theoretical preliminaries addressing Lie groups and uncertainty definition in the form of the concentrated Gaussian distribution. In Section 3 we derive the proposed LG-EIF, while in Section 4 we present the experimental results. In the end, Section 5 concludes the paper.

2 Preliminaries

2.1 Lie groups and Lie algebras

Generally, a Lie group is a group which has also the structure of a differentiable manifold and the group operations (product and inversion) are differentiable. In this paper we restrict our attention to a special class of Lie groups, the matrix groups over the field of reals, where the group operations are matrix multiplication and inversion, with the identity matrix I^d being the identity element of the group. These groups are frequently called, especially in the engineering literature, matrix Lie groups. The name emphasizes the fact that every matrix group is a Lie group, as well as the differential geometric viewpoint that is regularly employed. A matrix Lie group G can be characterized as a closed subgroup of a general linear group $\mathsf{GL}(d;\mathbb{R})$, in the sense that: if (A_n) is a sequence of matrices in G and A_n converges to a matrix A, with respect to a norm on $\mathbb{R}^{d\times d}$, then $A \in \mathsf{G}$ or $A \notin \mathsf{GL}(d;\mathbb{R})$, i.e., A is not invertible [36]. Ubiquitous examples of real matrix Lie groups are the general linear group $\mathsf{GL}(d;\mathbb{R})$, special linear group $\mathsf{SL}(d;\mathbb{R})$, orthogonal $\mathsf{O}(d)$ and special orthogonal $\mathsf{SO}(d)$ groups, etc. For an introductory, but rigorous mathematical treatment of matrix Lie groups, the interested reader is advised to confer [36].

To every Lie group G, there is an associated Lie algebra $\mathfrak{g}$ — a linear space (of the same dimension as G) endowed with a binary operation $[\cdot,\cdot]$ called the Lie bracket. From the differential geometric point of view, it is an open neighborhood of the origin in the tangent space of G at the identity element. A local diffeomorphism between a Lie group (manifold) and associated Lie algebra (tangent space) is established through the exponential mapping $\exp : \mathfrak{g} \to \mathsf{G}$ and its inverse $\log : \mathsf{G} \to \mathfrak{g}$ called the logarithm. This is a crucial mechanism for transfer of information between the group and its algebra. In case of matrix Lie groups, the exponential mapping is simply the matrix exponential

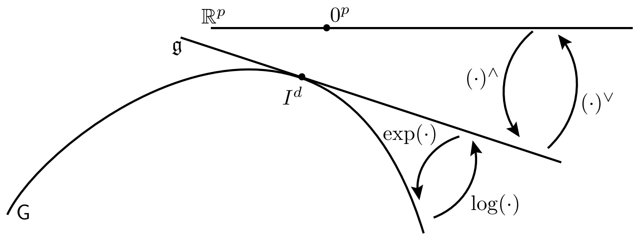

Fig. 1. An illustration of mappings within the triplet of Lie group G – Lie algebra $\mathfrak{g}$ – Euclidean space $\mathbb{R}^p$.

$$\exp(X) = \sum_{n=0}^{\infty} \frac{1}{n!} X^n \,,$$

and its inverse is of course the matrix logarithm defined for all $d\times d$ matrices A satisfying $\|A - I^d\| < 1$. Moreover, the matrix exponential can even be used to characterize the matrix Lie algebra — if G is a matrix Lie group, then its Lie algebra, denoted by $\mathfrak{g}$, is the set of all matrices X such that $\exp(tX) \in \mathsf{G}$ for all $t \in \mathbb{R}$ [36]. Being a linear space, a (real) p-dimensional matrix Lie algebra $\mathfrak{g}$ is naturally related to the Euclidean space $\mathbb{R}^p$ through a linear isomorphism $(\cdot)^\vee : \mathfrak{g} \to \mathbb{R}^p$ and its inverse denoted by $(\cdot)^\wedge : \mathbb{R}^p \to \mathfrak{g}$. An illustration of these concepts is given in Fig. 1 [26].

The adjoint representation of a matrix Lie group G is the map $\mathrm{Ad} : \mathsf{G} \to \mathsf{GL}(\mathfrak{g})$ defined by $A \mapsto \mathrm{Ad}\, A$, where $\mathrm{Ad}\, A$ is a linear invertible operator $\mathrm{Ad}\, A : \mathfrak{g} \to \mathfrak{g}$ given by

$$\mathrm{Ad}\, A(X) = AXA^{-1} \,, \quad X \in \mathfrak{g} \,.$$

Due to the natural isomorphism between $\mathfrak{g}$ and $\mathbb{R}^p$, essentially $\mathsf{GL}(\mathfrak{g}) = \mathsf{GL}(p;\mathbb{R})$ and Ad is to be understood as a group homomorphism. Therefore, there exists a unique linear map $\mathrm{ad} : \mathfrak{g} \to \mathsf{GL}(\mathfrak{g})$, called the adjoint representation of the Lie algebra $\mathfrak{g}$, defined by $X \mapsto \mathrm{ad}\, X$, where $\mathrm{ad}\, X$ is a linear operator on $\mathfrak{g}$ given by

$$\mathrm{ad}\, X(Y) = [X, Y] = XY - YX \,, \quad Y \in \mathfrak{g} \,.$$

In fact, from the differential geometric point of view, ad is the differential of Ad at the identity of G, and they are related through the following [36]

$$\mathrm{Ad} \exp(X) = \exp(\mathrm{ad}\, X) \,, \quad \text{for all } X \in \mathfrak{g} \,. \qquad (1)$$

The action of both adjoints can be further transferred from $\mathfrak{g}$ to $\mathbb{R}^p$ by the above isomorphism, and we denote them by $\mathrm{Ad}^\vee$ and $\mathrm{ad}^\vee$, respectively.

2.2 Concentrated Gaussian Distribution

Let G be a connected *unimodular* real matrix group. Unimodular means that its integration (Haar) measure ζ is both left and right translation invariant, i.e., $\zeta(A\mathcal{E}) = \zeta(\mathcal{E}A) = \zeta(\mathcal{E})$ for all $A \in \mathsf{G}$ and all Borel subsets $\mathcal{E}$ of G. Prominent examples like $\mathsf{SO}(3)$ and $\mathsf{SE}(3)$ are unimodular matrix groups [37]. Let us assume that a random variable X taking values in G has the probability distribution with the probability density function (pdf) of the following form [38]

$$p(X;\Sigma) = \beta \exp\left(-\frac{1}{2}(\log(X)^\vee)^{\mathrm{T}}\Sigma^{-1}\log(X)^\vee\right), \quad (2)$$

where β is a normalizing constant such that (2) integrates to unity (over G with respect to ζ), and Σ is a positive definite $p \times p$ matrix. Seemingly, in notation $\varepsilon = \log(X)^\vee \in \mathbb{R}^p$, density (2) has the structure of a zero mean Gaussian with covariance matrix Σ. However, observe that the normalizing constant β differs from $(2\pi)^{-p/2}(\det\Sigma)^{-1/2}$ and, in the sense of $\mathbb{R}^p$, it is only defined on an open neighborhood of the origin, which is the image of the $\log^\vee$ map. Random variables on G having the probability distribution given by density (2) are therefore called *normally (or Gauss) distributed* with mean I^d and covariance Σ. Additionally, we will assume that all eigenvalues of Σ are small, thus, almost all the mass of the distribution is concentrated in a small neighborhood around the mean value, and such a distribution is called a *concentrated Gaussian distribution* [38]. Furthermore, we say that a random variable X has a concentrated Gaussian distribution of mean $M \in \mathsf{G}$ and covariance matrix Σ, written $X \sim \mathcal{G}(M,\Sigma)$, if $M^{-1}X$ has the concentrated Gaussian distribution of mean I^d and covariance Σ [38], i.e., the density of $\mathcal{G}(M,\Sigma)$ is given by

$$\begin{aligned} &p(X;M,\Sigma) \\ &\quad = \beta \exp\left(-\frac{1}{2}(\log(M^{-1}X)^\vee)^{\mathrm{T}}\Sigma^{-1}\log(M^{-1}X)^\vee\right). \end{aligned} \quad (3)$$

An illustration of the concentrated Gaussian distribution is provided in Fig. 2.

It is well known that in the Euclidean setting multivariate Gaussian distributions $\mathcal{G}(m,\Sigma)$ form an exponential family [39] and in the canonical representation source parameters (m,Σ) are replaced by the corresponding natural parameters $(y,Y) = (\Sigma^{-1}m, \frac{1}{2}\Sigma^{-1})$, which also uniquely determine the Gaussian distribution. Canonical representation has many advantages, in particular, it is very useful for implementation of the standard IF. In the present paper we pursue the same idea for concentrated Gaussian distribution $\mathcal{G}(M,\Sigma)$ defined on matrix Lie groups. Using the BCH expansion (A.2) we have

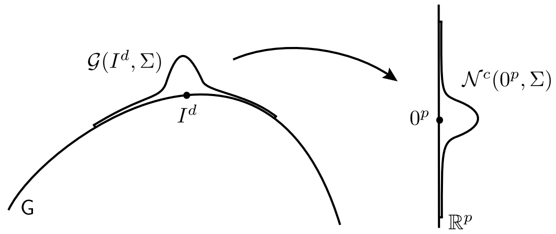

Fig. 2. An illustration of the concentrated Gaussian distribution $\mathcal{G}(I^d,\Sigma)$. The mean value I^d resides on the group $\mathsf{G} \subset \mathsf{GL}(d;\mathbb{R})$ while the covariance matrix Σ belongs to $\mathsf{GL}(p;\mathbb{R})$. On the right we depict the truncated or compactly supported corresponding $\mathcal{N}^c$ Gaussian in $\mathbb{R}^p$ with mean value 0^p and covariance matrix Σ.

$$\log(M^{-1}X) = -\log M + \log X - \frac{1}{2}[\log X, \log M] + \dots, \quad (4)$$

thus, according to (3), $\mathcal{G}(M,\Sigma)$ is also completely determined by the so called information vector-matrix pair (y,Y) given by

$$y = \Sigma^{-1}(\log M)^\vee \text{ and } Y = \Sigma^{-1}. \quad (5)$$

Given that, we have formed the basis for the derivation of the LG-EIF.

3 The Extended Information Filter on Matrix Lie Groups

Just as the standard KF, the LG-EIF recursion is divided in two steps: prediction and update and in the sequel we derive the equations of the proposed information form of the LG-EKF. First, we start with the prediction step where the same logic applies as in the case of the standard IF; namely, the computational burden is increased since, in order to apply the motion model, we need to convert the information vector to the mean. Second, the update step of the filter is derived where the advantages of the information form are kept, thus facilitating updates with multiple sensors or opening the way for decentralization approaches.

3.1 Motion and measurement models

Let G be a matrix Lie group and $X_k \in \mathsf{G}$ denote a system state at time step $k \geq 0$. We assume that the motion model of the system (the state equation) is described by a non-linear twice continuously differentiable function [1] $\Omega : \mathcal{U} \supset \mathsf{G} \to \mathbb{R}^p$ and the left action of the current state as follows [26]

$$X_{k+1} = X_k \exp\left(\Omega(X_k)^\wedge + n_k^\wedge\right), \quad k \geq 0, \qquad (6)$$

where $n_k \sim \mathcal{N}_{\mathbb{R}^p}\left(0_{p\times 1}, Q_k\right)$ is Gaussian noise in $\mathbb{R}^p$. For example, such models have appeared in [26, 28, 40] modeling motion as constant velocity on $\mathsf{SE}(2)$ and constant acceleration on $\mathsf{SO}(2)$, $\mathsf{SO}(3)$ and $\mathsf{SE}(3)$, respectively.

The discrete measurement model on the matrix Lie group is modeled by a continuously differentiable function $h : \mathcal{U} \supset \mathsf{G} \to \mathsf{G}'$ and the group perturbation as [26]

$$Z_{k+1} = h(X_{k+1}) \exp\left(r_{k+1}^\wedge\right), \qquad (7)$$

where $r_{k+1} \sim \mathcal{N}_{\mathbb{R}^q}\left(\mathbf{0}^{q\times 1}, R_{k+1}\right)$ is a Gaussian noise in $\mathbb{R}^q$ and exp denotes the exponential mapping on a q-dimensional matrix Lie group G'.

3.2 LG-EIF prediction

We assume that the posterior distribution at time step k is given by the concentrated Gaussian distribution $\mathcal{G}(M_k, \Sigma_k)$, shortly $\mathcal{G}_k$. In fact, we assume that $\mathcal{G}_k$ is known through the canonical parameters (y_k, Y_k), for which we aim to derive the filter recursions. Note that, according to relation (5), $\Sigma_k = Y_k^{-1}$ and $M_k = \exp((Y_k^{-1} y_k)^\wedge)$.

Following the idea proposed in [26], we first consider the covariance propagation under the motion model. For that purpose the Lie algebraic error, defined by $\varepsilon_k^\wedge = \log(M_k^{-1} X_k)$, is propagated under the motion model according to

$$\exp\left(\varepsilon_{k+1|k}^\wedge\right) = M_{k+1|k}^{-1} X_{k+1},$$

where $M_{k+1|k} = M_k \exp(\Omega(M_k)^\wedge)$. Therefore, the predicted state error on G can be expressed as

$$\exp\left(\varepsilon_{k+1|k}^\wedge\right) = \exp\left(-\Omega_k^\wedge\right) \exp\left(\varepsilon_k^\wedge\right) \exp\left(\Omega(X_k)^\wedge + n_k^\wedge\right),$$

where $\Omega_k^\wedge = \Omega(M_k)^\wedge$. Linearizing Ω in M_k and using the BCH expansion (A.2), defined in the Appendix A, one obtains the following propagated Lie algebraic error

$$\varepsilon_{k+1|k} = \mathcal{F}_k \varepsilon_k + \Psi(\Omega_k) n_k + \mathcal{O}\left(|\varepsilon_k, n_k|^2\right), \qquad (8)$$

[1] For the ease of differentiability requirement, we assume that Ω is defined on $\mathcal{U}$, which is an open subset of $\mathbb{R}^{d\times d}$ containing the group G.

where $\mathcal{O}(|\varepsilon_k, n_k|^2)$ is short for $\mathcal{O}(|\varepsilon_k|^2) + \mathcal{O}(|n_k|^2) + \mathcal{O}(|\varepsilon_k n_k|)$. Operators $\mathcal{F}_k$, the matrix Lie group equivalent to the Jacobian of the nonlinearity of the motion model, and Ψ are given by the following formulae:

$$\mathcal{F}_k = \mathrm{Ad}^\vee\left(\exp\left(-\Omega_k^\wedge\right)\right) + \Psi(\Omega_k)\mathscr{C}_k, \qquad (9)$$

$$\Psi(\boldsymbol{v}) = \sum_{m=0}^{\infty} \frac{(-1)^m}{(m+1)!} \mathrm{ad}^\vee(\boldsymbol{v})^m, \ \boldsymbol{v} \in \mathbb{R}^p, \qquad (10)$$

$$\mathscr{C}_k = \frac{\partial}{\partial \varepsilon} \Omega\left(M_k \exp\left(\varepsilon^\wedge\right)\right)|_{\varepsilon=0}. \qquad (11)$$

Operator Ψ is called the right Jacobian of G [20], while $\mathscr{C}_k$ denotes the linearization of the motion model (6) at M_k. The above formulae can be found in [26, 27]; however, without a detailed derivation, which we provide for the reader's convenience in Appendix A. Neglecting the second-order terms in (8) and using the fact that $\mathbb{E}(\varepsilon_k) = 0$, which is satisfied by the construction of the concentrated Gaussian distribution (see (3)), the expectation of $\varepsilon_{k+1|k}$ becomes

$$\mathbb{E}(\varepsilon_{k+1|k}) = \mathcal{F}_k \mathbb{E}(\varepsilon_k) = 0.$$

The predicted covariance matrix $\Sigma_{k+1|k}$ is the covariance matrix of the predicted Lie algebraic error $\varepsilon_{k+1|k}$ and due to the linear equation (8) it evaluates to

$$\begin{aligned}\Sigma_{k+1|k} &= \mathbb{E}\left[\varepsilon_{k+1|k}\varepsilon_{k+1|k}^{\mathrm{T}}\right] \\ &= \mathcal{F}_k \Sigma_k \mathcal{F}_k^{\mathrm{T}} + \Psi(\Omega_k) Q_k \Psi(\Omega_k)^{\mathrm{T}}.\end{aligned}$$

Applying the Woodbury's matrix identity [41], $Y_{k+1|k} = \Sigma_{k+1|k}^{-1}$ evaluates to

$$Y_{k+1|k} = \tilde{Q}_k^{-1} - \tilde{Q}_k^{-1}\mathcal{F}_k\left(Y_k + \mathcal{F}_k^{\mathrm{T}}\tilde{Q}_k^{-1}\mathcal{F}_k\right)^{-1}\mathcal{F}_k^{\mathrm{T}}\tilde{Q}_k^{-1},$$

where $\tilde{Q} = \Psi_k Q_k \Psi_k^{\mathrm{T}}$, $\Psi_k = \Psi(\Omega_k)$, and all inverse matrices are assumed to exist. Finally, the predicted information vector $y_{k+1|k} = Y_{k+1|k}(\log M_{k+1|k})^\vee$ amounts to

$$y_{k+1|k} = Y_{k+1|k} \log\left(\exp((Y_k^{-1}y_k)^\wedge)\exp(\Omega_k^\wedge)\right)^\vee. \qquad (12)$$

Remark 1 *Assuming that M_k and $\exp(\Omega_k^\wedge)$ are such that according to the BCH expansion (A.2)*

$$\begin{aligned}&\log\left(\exp((Y_k^{-1}y_k)^\wedge)\exp(\Omega_k^\wedge)\right)^\vee \\ &\qquad\approx Y_k^{-1}y_k + \Omega_k + \frac{1}{2}[Y_k^{-1}y_k, \Omega_k]^\vee,\end{aligned}$$

then the prediction formula (12) simplifies to

$$y_{k+1|k} = Y_{k+1|k}\Big(Y_k^{-1}y_k + \Omega_k + \frac{1}{2}[Y_k^{-1}y_k, \Omega_k]^\vee\Big). \qquad (13)$$

3.3 LG-EIF update

Let us define the innovation term as

$$z_{k+1} = \log\left(h\left(M_{k+1|k}\right)^{-1} Z_{k+1}\right)^{\vee}$$

Again, applying the BCH formula (A.2) and linearizing the nonlinear terms at $M_{k+1|k}$, we obtain [26]

$$z_{k+1} = \mathcal{H}_{k+1}\varepsilon_{k+1|k} + r_{k+1} + \mathcal{O}\left(|\varepsilon_{k+1|k} r_{k+1}|\right) , \quad (14)$$

with

$$\mathcal{H}_{k+1} = \frac{\partial}{\partial \varepsilon}\left[\log\left(h(M_{k+1|k})^{-1} h\left(M_{k+1|k}\exp\left(\varepsilon^{\wedge}\right)\right)\right)\right]\Big|^{\vee}_{\varepsilon=0} . \quad (15)$$

Since (14) is linear in the Lie algebraic error $\varepsilon_{k+1|k}$, we assert that the standard update equations of the IF [3] can be applied. From the previous section we know that $\varepsilon_{k+1|k}$ is distributed according to the truncated zero mean Gaussian with covariance matrix $\Sigma_{k+1|k}$, which is assumed to be well approximated by the Gaussian of the same parameters. Given that, the updated Lie algebraic error ε^{-}_{k+1} will be Gaussian distributed with natural parameters

$$\begin{aligned} y^{-}_{k+1} &= \mathcal{H}^{\mathrm{T}}_{k+1} R^{-1}_{k+1} z_{k+1} , \\ Y^{-}_{k+1} &= Y_{k+1|k} + \mathcal{H}^{\mathrm{T}}_{k+1} R^{-1}_{k+1}\mathcal{H}_{k+1} . \end{aligned} \quad (16)$$

However, we have not completed the update step for the following reasons. Namely, from the definition of ε^{-}_{k+1}, the conditional random variable $X_{k+1|k+1} := X_{k+1}|\{Z_1, \ldots, Z_{k+1}\}$ has the form

$$X_{k+1|k+1} = M_{k+1|k}\exp(\varepsilon^{-\wedge}_{k+1}) , \quad (17)$$

but, the mean value of ε^{-}_{k+1} now equals $m^{-}_{k+1} = (Y^{-}_{k+1})^{-1} y^{-}_{k+1}$, which in general differs from the zero vector, and (17) is not in the form suitable for description by the concentrated Gaussian distribution. To overcome that issue, the state reparametrization, as proposed in [26], is performed. Let us define $\xi_{k+1} = \varepsilon^{-}_{k+1} - m^{-}_{k+1}$, then $\mathbb{E}(\xi_{k+1}) = 0$ and using formula (A.3) from the Appendix A we obtain (up to $\mathcal{O}(|\xi_{k+1}|^2)$ terms)

$$\begin{aligned} X_{k+1|k+1} &= M_{k+1|k}\exp\left(m^{-\wedge}_{k+1} + \xi^{\wedge}_{k+1}\right) \\ &= M_{k+1|k}\exp\left(m^{-\wedge}_{k+1}\right)\exp\left(\Psi(m^{-}_{k+1})\xi^{\wedge}_{k+1}\right) . \end{aligned}$$

Now defining $M_{k+1} = M_{k+1|k}\exp\left(m^{-\wedge}_{k+1}\right)$ and $\varepsilon_{k+1} = \Psi(m^{-}_{k+1})\xi_{k+1}$, we have $X_{k+1|k+1}$ in a more suitable form

$$X_{k+1|k+1} = M_{k+1}\exp(\varepsilon^{\wedge}_{k+1}) , \quad (18)$$

Algorithm 1 The pseudocode of the LG-EIF

Require: $\mathcal{G}_k = \mathcal{G}(y_k, Y_k)$, $\Omega(X)$, Q_k

Prediction

Evaluate Ω_k, $\mathscr{C}_k$ and $\tilde{Q}_k$

$\mathcal{F}_k = \mathrm{Ad}^{\vee}\Big(\exp\left(-\Omega^{\wedge}_k\right)\Big) + \Psi(\Omega_k)\mathscr{C}_k$

$Y_{k+1|k} = \tilde{Q}^{-1}_k - \tilde{Q}^{-1}_k\mathcal{F}_k\big(Y_k + \mathcal{F}^{\mathrm{T}}_k\tilde{Q}^{-1}_k\mathcal{F}_k\big)^{-1}\mathcal{F}^{\mathrm{T}}_k\tilde{Q}^{-1}_k$

$y_{k+1|k} = Y_{k+1|k}\log\Big(\exp\left((Y^{-1}_k y_k)^{\wedge}\right)\exp\left(\Omega^{\wedge}_k\right)\Big)^{\vee}$

Require: $\mathcal{G}_{k+1|k} = \mathcal{G}(y_{k+1|k}, Y_{k+1|k})$, $h(X)$, R_{k+1}

Update

Evaluate $\mathcal{H}_{k+1}$

$Y^{-}_{k+1} = Y_{k+1|k} + \mathcal{H}^{\mathrm{T}}_{k+1}R^{-1}_{k+1}\mathcal{H}_{k+1}$

$y^{-}_{k+1} = \mathcal{H}^{\mathrm{T}}_{k+1}R^{-1}_{k+1}z_{k+1}$

$m^{-}_{k+1} = \left(Y^{-}_{k+1}\right)^{-1}y^{-}_{k+1}$

$Y_{k+1} = \Psi\left(m^{-}_{k+1}\right)^{-\mathrm{T}}Y^{-}_{k+1}\Psi\left(m^{-}_{k+1}\right)^{-1}$

$y_{k+1} = Y_{k+1}\Bigg(\log\Big(\exp\left((Y^{-1}_k y_k)^{\wedge}\right)\exp\left(\Omega^{\wedge}_k\right)\exp\left(m^{-\wedge}_{k+1}\right)\Big)\Bigg)^{\vee}$

return $\mathcal{G}_{k+1} = \mathcal{G}(y_{k+1}, Y_{k+1})$

from which the posterior distribution can be plainly read off. By definition

$$\begin{aligned} \Sigma_{k+1} &= \mathbb{E}\left[\varepsilon_{k+1}\varepsilon^{\mathrm{T}}_{k+1}\right] = \mathbb{E}\left[\Psi(m^{-}_{k+1})\xi_{k+1}\xi^{\mathrm{T}}_{k+1}\Psi(m^{-}_{k+1})^{\mathrm{T}}\right] \\ &= \Psi\left(m^{-}_{k+1}\right)\left(Y^{-}_{k+1}\right)^{-1}\Psi\left(m^{-}_{k+1}\right)^{\mathrm{T}} , \end{aligned}$$

and therefore, the finally updated information matrix equals

$$Y_{k+1} = \Psi(m^{-}_{k+1})^{-\mathrm{T}}Y^{-}_{k+1}\Psi(m^{-}_{k+1})^{-1} . \quad (19)$$

Concerning the information vector, we find

$$\begin{aligned} y_{k+1} &= Y_{k+1}(\log M_{k+1})^{\vee} \\ &= Y_{k+1}\left(\log\left(\exp((Y^{-1}_k y_k)^{\wedge})\exp(\Omega^{\wedge}_k)\exp(m^{-\wedge}_{k+1})\right)\right)^{\vee} . \end{aligned} \quad (20)$$

The update step is illustratively summarized in Figure 3. Note that, in comparison to the standard EIF, we cannot calculate the final information vector update in (20) by using just the information form. However, this does not preclude an advantage in computational complexity with respect to the LG-EKF (as shown in Section 4). The pseudocode of the LG-EIF is given in Algorithm 1.

Remark 2 *Recall that one of the main advantages of the IF lies in the simultaneous update of multiple measurements in the same time step. In case that N measurements are available at time step $k+1$ through* *different measurement models h_i and measurement noise $r^i_{k+1} \sim \mathcal{N}_{\mathbb{R}^q}(0, R_{i,k+1})$, the updated information vector and matrix (prior to the reparametrization step) become*

$$\begin{aligned} y^-_{k+1} &= \sum_{i=1}^{N} \mathcal{H}^{\mathrm{T}}_{i,k+1} R^{-1}_{i,k+1} z_{i,k+1}\,, \\ Y^-_{k+1} &= Y_{k+1|k} + \sum_{i=1}^{N} \mathcal{H}^{\mathrm{T}}_{i,k+1} R^{-1}_{i,k+1} \mathcal{H}_{i,k+1}\,. \end{aligned} \tag{21}$$

Remark 3 *Difficulties that could be encountered in the filter design are twofold. First, the evaluation of operators $\mathcal{C}_k$ and $\mathcal{H}_k$, which arise in the linearization of $\Omega(\cdot)$ and $h(\cdot)$, could be mathematically involved. And second,* $\exp(\cdot)$*,* $\log(\cdot)$*,* $\mathrm{Ad}(\cdot)$*,* $\mathrm{ad}(\cdot)$*, and $\Psi(\cdot)$ might not allow closed-form expressions for some Lie groups. In that case, it is necessary to apply a truncated Taylor series expansion. However, many Lie groups that are significant for engineering applications allow for closed form expressions for the majority of aforementioned maps.*

4 Experiments

In this section we demonstrate the effectiveness and applicability of the LG-EIF on the problem of rigid body attitude tracking in 3D. We pose the experiment as a multisensor estimation problem of a state residing on the group $\mathsf{G} = \mathsf{SO}(3) \times \mathbb{R}^3$, where the first group, $\mathsf{SO}(3)$, represents the rigid body orientation in 3D, while the second group, $\mathbb{R}^3$, represents pertaining angular rates. This is a slight abuse of notation intended for clarity, since when talking about $\mathbb{R}^p$ in the framework of groups, we are actually referring to their matrix representation. $\mathbb{R}^3$ can be thought of as a three-dimensional matrix Lie group through the following identification

$$\mathbb{R}^3 \ni (a_1, a_2, a_3) \mapsto \begin{pmatrix} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & a_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathsf{GL}(4;\mathbb{R})\,, \tag{22}$$

which transfers the addition of vectors, as the group operation in $\mathbb{R}^3$, to the multiplication of matrices. Hence, G can be thought of as a subgroup of $\mathsf{GL}(7;\mathbb{R})$, whose elements are block diagonal matrices where the first 3×3 block belongs to $\mathsf{SO}(3)$, while the second 4×4 block is of the form (22). In that setting G is a unimodular matrix Lie group with the Haar measure being the tensor product of the Haar measure on $\mathsf{SO}(3)$ and essentially the Lebesgue measure on $\mathbb{R}^3$. Thus, the LG-EIF methodology developed in previous sections is applicable on G.

4.1 Filtering on $\mathsf{SO}(3) \times \mathbb{R}^3$

Note that we designate the system state as $X_k \in \mathsf{SO}(3) \times \mathbb{R}^3$ which consists of the orientation component $\Phi_k \in \mathsf{SO}(3)$ and the angular rate component $\dot{\Phi}_k \in \mathbb{R}^3$.

4.1.1 Prediction

We propose to model the motion (6) by a constant angular rate motion model

$$\begin{aligned} \Omega(X_k) &= \left[T\dot{\phi}_{1,k}\ T\dot{\phi}_{2,k}\ T\dot{\phi}_{3,k}\ 0\ 0\ 0\right]^{\mathrm{T}} \\ n_k &= \left[\tfrac{T^2}{2} n_{1,k}\ \tfrac{T^2}{2} n_{2,k}\ \tfrac{T^2}{2} n_{3,k}\ T n_{1,k}\ T n_{2,k}\ T n_{3,k}\right]^{\mathrm{T}}, \end{aligned} \tag{23}$$

where T is the discretization time. With such a defined motion model, the system is corrupted with white noise over three separated components, i.e., $n_{1,k}$ the noise in local ϕ_1 direction, $n_{2,k}$ the noise in local ϕ_2 direction and $n_{3,k}$ the noise in local ϕ_3 direction. Given that, the components can be seen as resembling a Wiener process over the associated axes.

The uncertainty propagation can be challenging, since it requires the calculation of (11), which needs to be patiently evaluated for each considered problem. However, for the Lie algebraic error $\varepsilon = \left[\varepsilon_1\ \varepsilon_2\ \varepsilon_3\ \dot{\varepsilon}_1\ \dot{\varepsilon}_2\ \dot{\varepsilon}_3\right]^{\mathrm{T}}$, and the motion model given by (23), which extracts only the Euclidean part of the state, we obtain

$$\mathscr{C}_k = \begin{bmatrix} 0_{3\times3} & T \cdot I_3 \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix}. \tag{24}$$

Now, we have all the ingredients for applying the motion model to predict the state in an LG-EIF manner.

4.1.2 Update

The measurement function is the map $h : \mathsf{SO}(3) \times \mathbb{R}^3 \to \mathsf{SO}(3)$, and although we have N measurements we use the expression (21) for the update, hence mapping dimensions correspond as if having a single measurements. The element that specifically needs to be derived is the measurement matrix $\mathcal{H}_{k+1}$, which in the vein of (15) requires evaluating partial derivatives and multivariate limits. With having the Lie algebraic error defined, the function to be partially derived is

$$\begin{aligned} &\Big[\log\left(h(M_{k+1|k})^{-1}\right. \\ &\qquad \left. h\left(M_{k+1|k}\exp\left(\varepsilon^\wedge\right)\right)\right)\Big]^\vee = \left[\varepsilon_1\ \varepsilon_2\ \varepsilon_3\right]^{\mathrm{T}}. \end{aligned} \tag{25}$$

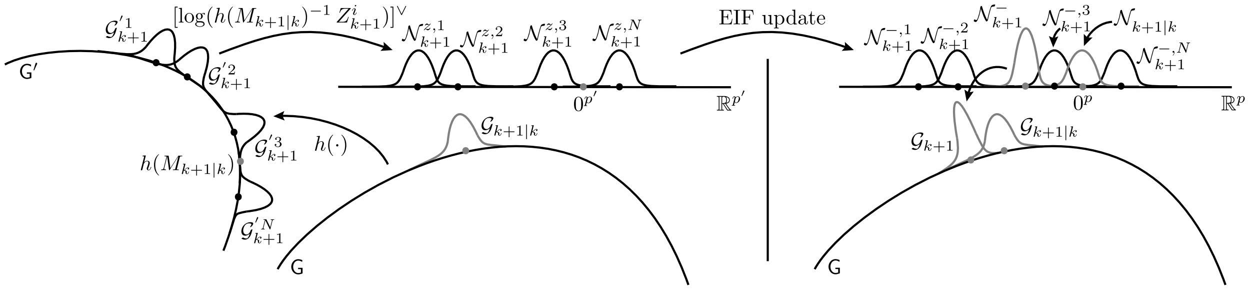

Fig. 3. An illustration of the LG-EIF update step. The predicted $\mathcal{G}_{k+1|k}$ on the group G is updated with N measurements $\mathcal{G}'^{i}_{k+1}$ on group G'. First, the predicted mean value $M_{k+1|k}$ is mapped to G' via measurement function $h(\cdot)$ and innovations $\mathcal{N}^{z,i}_{k+1}$ are calculated in the tangent space $\mathbb{R}^{p'}$ of $h(M_{k+1|k})$. Then, the innovations are mapped to the tangent space $\mathbb{R}^{p}$ of $M_{k+1|k}$, where the $\mathcal{N}_{k+1|k}$ corresponding to $\mathcal{G}_{k+1|k}$ is at the origin. The predicted distribution $\mathcal{N}_{k+1|k}$ is updated with the mapped measurements $\mathcal{N}^{-,i}_{k+1}$ resulting with a distribution displaced from the origin $\mathcal{N}^{-}_{k+1}$ which needs to be reparametrized and mapped back to G as the finally updated $\mathcal{G}_{k+1}$.

The final measurement matrix, for this case, is obtained by taking the partial derivatives of (25) with respect to the Lie algebraic error. Finally, the measurement matrix evaluates to $\mathcal{H}_{k+1} = [I_3 \; 0_{3\times 3}]$. Now, we have all the ingredients to update the filter in the LG-EIF manner.

4.2 Evaluation

In order to demonstrate the performance of the proposed filter we have simulated a rotating rigid body with the constant angular rate model. First, the initial orientation of the rigid body in $\mathsf{SO}(3)$ and initial angular rates are defined. Note that the angular rates are defined in the $\mathbb{R}^p$ isomorphic to the $\mathfrak{so}(3)$, i.e., the Euler axes representation (see Appendix B). Then, under the assumption of the constant angular rate model, random disturbances are added via accelerations in the pertaining Euclidean space $\mathbb{R}^p$. Measurements are generated by corrupting the true orientation of the body in $\mathbb{R}^p$ with white Gaussian noise, and then mapping the result via the exponential map back to the $\mathsf{SO}(3)$.

In Fig. 4 we can see the result of LG-EIF and Euler angles EIF comparison on 100 randomly generated trajectories measured with $N = 5$ sensors for $k = 100$ steps. The initial state of the system was set to $[\log X_0]^\vee = [0_{1\times 6}]^{\mathrm{T}}$, the standard deviation of random accelerations over the three axes acting as disturbances was $\sigma_p = 10°/\mathrm{s}^2$ and standard deviation of measurement noise over the three axes ranged from $\sigma_m = 0.1°$ to $\sigma_m = 20°$. The estimated orientation of the LG-EIF is defined in $\mathsf{SO}(3)$, and in Fig. 4 we show the the attitude error calculated as the cosine angle between two rotation matrices

$$\Phi_{\mathrm{err}} = \arccos\left(\frac{1}{2}(\mathrm{Tr}[\Phi_{\mathrm{t}}^{\mathrm{T}}\Phi_{\mathrm{e}}] - 1)\right), \qquad (26)$$

where Φ_{t} is the true orientation and Φ_{e} is estimated orientation. We can see from Fig. 4 that for measurement noise standard deviation larger than 2° on average the LG-EIF achieves smaller attitude root-mean-square-error (RMSE) and has significantly smaller variation (not noticeable in the figure) in the results compared to the Euler angles EIF. In Fig. 5 we show three examples (with different measurement noise intensity) of time behavior of the attitude estimation error for different filters, where the smaller variation for LG-EIF can be noticed. Furthermore, it could be argued that other filtering methods inlieu of EIF could be used which can better handle nonlinearities. However, the EIF system and measurement equations are linear in this case and we assert that the main reason behind larger errors in EIF comes from the suboptimal state space parametrization, rather than linearization errors in state and measurement equations.

The main advantage of the IF form is the computational efficiency of the update step with respect to a large number of measurements. To verify that the same advantage holds also for the Lie group EKF we have compared the execution time of the LG-EKF and LG-EIF on 100 examples of 100 step long simulated trajectories. Fig. 6 shows the execution time ratio of the LG-EKF and LG-EIF. We can see that for a large number of sensors or features this difference is prominent.

5 Conclusion

In this paper we have proposed a new state estimation algorithm on Lie groups. We have embedded the LG-EKF with an EIF form for non-linear systems, thus endowing the filter with the information form advantages with regard to multisensor update and decentralization, while keeping the accuracy of the LG-EKF for stochastic inference on Lie groups. The theoretical development of the LG-EIF recursion equations was presented and the applicability of the proposed approach demonstrated on the problem of

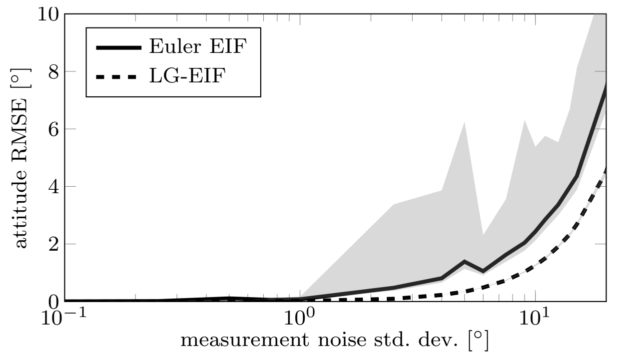

Fig. 4. Comparison of attitude RMSE with respect to increase in measurement noise standard deviation. The results represent the mean value of the RMSE and one standard deviation of 100 MC runs. We can see that the LG-EIF exhibits smaller error and has more consistent performance over various trajectories.

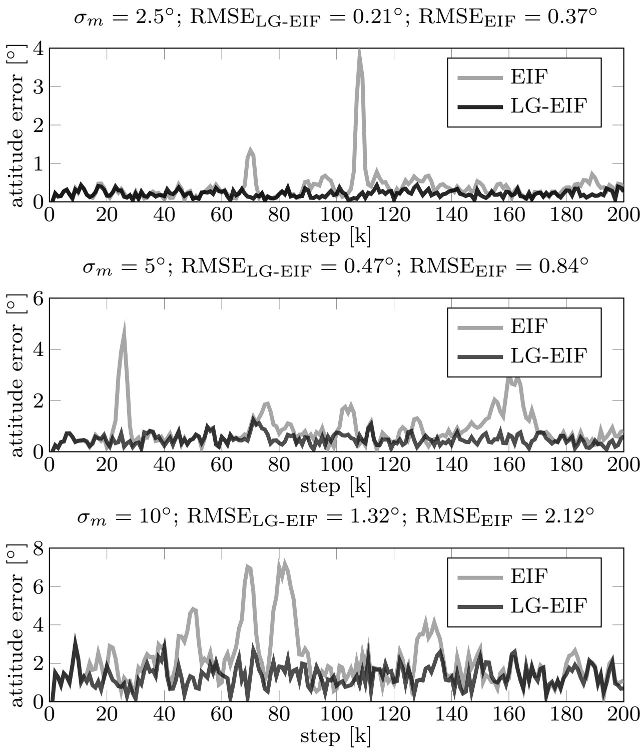

Fig. 5. Three examples of time behavior of the attitude estimation error through 200 steps. The standard deviation for measurement noise was set to $\sigma_m = 2.5°$ (top), $\sigma_m = 5°$ (middle), $\sigma_m = 10°$ (bottom). The attitude RMSE for each filter is given in the subfigure titles.

rigid body attitude tracking with multiple sensors by setting the state on the Lie group $\mathsf{SO}(3) \times \mathbb{R}^3$. The first component of the state represented the rigid body orientation in 3D, while the second component represented the pertaining angular rates. The system model was then set as a constant angular rate model acting on the state space $\mathsf{SO}(3) \times \mathbb{R}^3$. The results have shown that the filter can accurately track the rigid

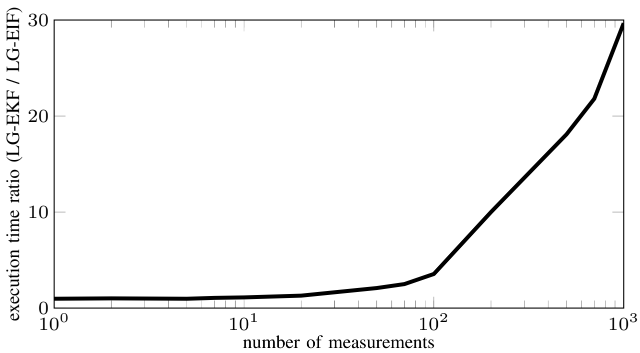

Fig. 6. Comparison of LG-EKF and LG-EIF time execution with respect to the number of measurements in the update step. We can see that after 100 measurements the difference becomes extremely prominent. The figure represents mean value of the execution time ratio for 50 Monte Carlo runs.

body attitude and that on average it exhibits lower RMSE and more consistent performance than the Euler angles based EIF. Furthermore, the information form of the LG-EIF keeps the multisensor or decentralization computational advantage of the update step with respect to the LG-EKF.

References

[1] P. S. Maybeck, *Stochastic models, estimation and control. Volume I.*, A. Press, Ed., 1979.

[2] B. Grocholsky, A. Makarenko, and H. Durrant-Whyte, “Information-theoretic coordinated control of multiple sensor platforms,” in *International Conference on Robotics and Automation (ICRA)*, vol. 1. IEEE, 2003, pp. 1521–1526.

[3] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics.* The MIT Press, 2006.

[4] B. S. Y. Rao, H. F. Durrant-Whyte, and J. A. Sheen, “A fully decentralized multi-sensor system for tracking and surveillance,” *The International Journal of Robotics Research*, vol. 12, no. 1, pp. 20–44, 1993.

[5] E. Nettleton, H. F. Durrant-Whyte, and S. Sukkarieh, “A robust architecture for decentralised data fusion,” in *International Conference on Advanced Robotics (ICAR)*, 2003.

[6] Y. Zhang, Y. Chai Soh, and W. Chen, “Robust information filter for decentralized estimation,” *Automatica*, vol. 41, no. 12, pp. 2141–2146, 2005.

[7] G. Battistelli and L. Chisci, “Stability of consensus extended Kalman filter for distributed state estimation,” *Automatica*, vol. 68, pp. 169–178, 2016.

[8] T. Onel, C. Ersoy, and H. Delic, “Information content-based sensor selection and transmission power adjustment for collaborative target tracking,” *IEEE Transactions on Mobile Computing*, vol. 8, no. 8, pp. 1103–1116, 2009.

[9] Y. Fu, Q. Ling, and Z. Tian, “Distributed sensor allocation for multi-target tracking in wireless sensor networks,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3538–3553, 2012.

[10] T. Vercauteren and X. Wang, “Decentralized sigma-point information filters for target tracking in collaborative sensor networks,” *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2997–3009, 2005.

[11] M. E. Campbell and W. W. Whitacre, “Cooperative tracking using vision measurements on SeaScan UAVs,” *IEEE Transactions on Control Systems Technology*, vol. 15, no. 4, pp. 613–626, 2007.

[12] G. Liu, F. Wörgötter, and I. Markelić, “Square-root sigma-point information filtering,” *IEEE Transactions on Automatic Control*, vol. 57, no. 11, pp. 2945–2950, 2012.

[13] S. Wang, J. Feng, and C. K. Tse, “A class of stable square-root nonlinear information filters,” *IEEE Transactions on Automatic Control*, vol. 59, no. 7, pp. 1893–1898, 2014.

[14] D.-J. Lee, “Nonlinear estimation and multiple sensor fusion using unscented information filtering,” *IEEE Signal Processing Letters*, vol. 15, no. 2, pp. 861–864, 2008.

[15] K. Pakki, B. Chandra, and I. Postlethwaite, “Square root cubature information filter,” *IEEE Sensors Journal*, vol. 13, no. 2, pp. 750–758, 2013.

[16] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. Ann Arbor, MI:CRC, 1994.

[17] J. M. Selig, *Geometrical Methods in Robotics*. New York: Springer-Verlag, 1996.

[18] Y. Wang and G. S. Chirikjian, “Error propagation on the Euclidean group with applications to manipulator kinematics,” *IEEE Transactions on Robotics*, vol. 22, no. 4, pp. 591–602, 2006.

[19] K. C. Wolfe, M. Mashner, and G. S. Chirikjian, “Bayesian fusion on Lie groups,” *Journal of Algebraic Statistics*, vol. 2, no. 1, pp. 75–97, 2011.

[20] 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, 2014.

[21] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, “IMU preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation,” in *Robotics: Science and Systems*, 2015, p. 9.

[22] A. Khosravian, J. Trumpf, R. Mahony, and T. Hamel, “State estimation for invariant systems on Lie groups with delayed output measurements,” *Automatica*, vol. 68, pp. 254–265, 2015.

[23] G. Kurz, I. Gilitschenski, and U. D. Hanebeck, “The partially wrapped normal distribution for SE(2) estimation,” in *International Conference on Multisensor Fusion and Information Integration for Intelligent Systems*, no. 2, 2014, p. 8.

[24] I. Gilitschenski, G. Kurz, S. J. Julier, and U. D. Hanebeck, “A new probability distribution for simultaneous representation of uncertain position and orientation,” in *International Conference on Information Fusion (FUSION)*, 2014, p. 7.

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

[26] G. Bourmaud, R. Mégret, M. Arnaudon, and A. Giremus, “Continuous-discrete extended Kalman filter on matrix Lie groups using concentrated Gaussian distributions,” *Journal of Mathematical Imaging and Vision*, vol. 51, no. 1, pp. 209–228, 2015.

[27] G. Bourmaud, R. Mégret, A. Giremus, and Y. Berthoumieu, “Discrete extended Kalman filter on Lie groups,” in *European Signal Processing Conference (EUSIPCO)*, 2013, pp. 1–5.

[28] J. Ćesić, I. Marković, I. Cvišić, and I. Petrović, “Radar and stereo vision fusion for multitarget tracking on the special Euclidean group,” *Robotics and Autonomous Systems*, 2016.

[29] S. Thrun, Y. Liu, D. Koller, A. Y. Ng, Z. Ghahramani, and H. Durrant-Whyte, “Simultaneous localization and mapping with sparse extended information filters,” *The International Journal of Robotics Research*, vol. 23, no. 7-8, pp. 693–716, 2004.

[30] T. Bailey, B. Upcroft, and H. Durrant-Whyte, “Validation gating for non-linear non-Gaussian target tracking,” in *International Conference on Information Fusion*, 2006, pp. 1–6.

[31] R. M. Eustice, H. Singh, and J. J. Leonard, “Exactly sparse delayed-state filters for view-based SLAM,” *IEEE Transactions on Robotics*, vol. 22, no. 6, pp. 1100–1114, 2006.

[32] G. Ros, J. Guerrero, A. D. Sappa, D. Ponsa, and A. M. Lopez, “VSLAM pose initialization via Lie groups and Lie algebras optimization,” in *International Conference on Robotics and Automation (ICRA)*. IEEE, 2013, pp. 5740 – 5747.

[33] R. Kümmerle, G. Grisetti, K. Rainer, H. Strasdat, K. Konolige, and W. Burgard, “g2o: A general framework for graph optimization,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2011, pp. 3607–3613.

[34] J. Engel, T. Sch, and D. Cremers, “LSD-SLAM: Direct monocular SLAM,” in *European Conference on Computer Vision (ECCV)*, 2014, pp. 834–849.

[35] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, “ORB-SLAM: a versatile and accurate monocular slam system,” *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.

[36] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer-Verlag, 2003.

[37] G. S. Chirikjian and A. B. Kyatkin, *Engineering Applications of Noncommutative Harmonic Analysis: With Emphasis on Rotation and Motion Groups*. CRC Press, 2000.

[38] Y. Wang and G. Chirikjian, “Error propagation on the Euclidean group with applications to manipulator kinematics,” *IEEE Transactions on Robotics*, vol. 22, no. 4, pp. 591–602, Aug. 2006.

[39] F. Nielsen and V. Garcia, “Statistical exponential families: A digest with flash cards,” *Computing Research Repository*, vol. abs/0911.4, 2009.

[40] J. Ćesić, V. Joukov, I. Petrović, and D. Kulić, “Full Body Human Motion Estimation on Lie Groups Using 3D Marker Position Measurements,” in *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2016.

[41] M. A. Woodbury, *Inverting Modified Matrices*, ser. Statistical Research Group Memorandum Reports. Princeton, NJ: Princeton University, 1950, no. 42.

A Lie algebraic error prediction

Proposition 4 (Baker-Campbell-Hausdorff) *Given a Lie algebra $\mathfrak{g}$, for all $a, b \in \mathfrak{g}$ such that $|a^\vee|$ and $|b^\vee|$ are sufficiently small, the following identity holds [36]:*

$$\log(\exp(a)\exp(b)) = a + \int_0^1 \psi(\exp(\mathrm{ad}(a))\exp(t\,\mathrm{ad}(b)))(b)\mathrm{d}t\,, \quad \text{(A.1)}$$

where $\psi(z) = z\log z/(z-1)$.

This is an integral version of the famous Baker-Campbell-Hausdorff (BCH) formula, which is better known in an expanded form

$$\begin{aligned}&\log(\exp(a)\exp(b))\\ &\quad= a+b+\frac{1}{2}[a,b]+\frac{1}{12}([a,[a,b]]+[b,[b,a]])+\dots\,.\end{aligned}\tag{A.2}$$

Another useful identity used in the derivation of the predicted Lie algebraic error is a first-order approximation relation between additive and multiplicative perturbations on matrix Lie groups. Namely, for every $a,\delta\in\mathfrak{g}$ and $|\delta^\vee|$ small, i.e., neglecting the second-order terms in $|\delta^\vee|$, it holds

$$\exp(a+\delta)\approx\exp(a)\exp(\Psi(a)\delta)\,,\tag{A.3}$$

where $\Psi(a)$ denotes the right Jacobian of the Lie group defined by (10). The latter identity implies the approximation formula

$$\log(\exp(-a)\exp(a+\delta))=\Psi(a)\delta\,.\tag{A.4}$$

Now we proceed with the derivation of the Lie algebraic error prediction. In Section 3.2 we computed the exponential of the predicted Lie algebraic error $\varepsilon_{k+1|k}$ at time step k

$$\exp(\varepsilon^\wedge_{k+1|k})=\exp\left(-\Omega^\wedge_k\right)\exp\left(\varepsilon^\wedge_k\right)\exp\left(\Omega(X_k)^\wedge+n^\wedge_k\right)\,,$$

where we recall $\Omega_k=\Omega(M_k)$ and ε_k is the Lie algebraic error at time step k. Linearizing the map Ω at M_k we have

$$\begin{aligned}&\exp(\varepsilon^\wedge_{k+1|k})\\ &\quad=\exp\left(-\Omega^\wedge_k\right)\exp\left(\varepsilon^\wedge_k\right)\exp\left(\Omega^\wedge_k+(\mathscr{C}_k\varepsilon_k)^\wedge+n^\wedge_k\right)\,,\end{aligned}\tag{A.5}$$

where

$$\mathscr{C}_k=\frac{\partial}{\partial\varepsilon}\Omega\left(M_k\exp\left(\varepsilon^\wedge\right)\right)|_{\varepsilon=0}\,.$$

Using the BCH formula (A.2) by considering only the first four members from the expansion and neglecting $\mathcal{O}(|\varepsilon^\wedge_k,n^\wedge_k|^2)$ terms, we obtain

$$\begin{aligned}z^\wedge_k&=\log(\exp\left(\varepsilon^\wedge_k\right)\exp\left(\Omega^\wedge_k+(\mathscr{C}_k\varepsilon_k)^\wedge+n^\wedge_k\right))\\ &=\varepsilon^\wedge_k+\Omega^\wedge_k+(\mathscr{C}_k\varepsilon_k)^\wedge+n^\wedge_k+\frac{1}{2}[\varepsilon^\wedge_k,\Omega^\wedge_k]\\ &\quad+\frac{1}{12}[\Omega^\wedge_k,[\Omega^\wedge_k,\varepsilon^\wedge_k]]\,.\end{aligned}\tag{A.6}$$

Inserting (A.6) into (A.5) one has

$$\exp(\varepsilon^\wedge_{k+1|k})=\exp\left(-\Omega^\wedge_k\right)\exp(z^\wedge_k)\,,$$

thus, using the approximation identity (A.4), the following expression holds

$$\begin{aligned}\varepsilon^\wedge_{k+1|k}=\Psi(\Omega_k)\Big(&\varepsilon^\wedge_k+(\mathscr{C}_k\varepsilon_k)^\wedge+n^\wedge_k+\frac{1}{2}[\varepsilon^\wedge_k,\Omega^\wedge_k]\\ &+\frac{1}{12}[\Omega^\wedge_k,[\Omega^\wedge_k,\varepsilon^\wedge_k]]\Big)\,.\end{aligned}$$

Recognizing terms $\Psi(\Omega_k)(\mathscr{C}_k\varepsilon_k)^\wedge$ and $\Psi(\Omega_k)n^\wedge_k$ in the prediction formula (8), it remains to discuss terms

$$\Psi(\Omega_k)\Big(\varepsilon^\wedge_k+\frac{1}{2}[\varepsilon^\wedge_k,\Omega^\wedge_k]+\frac{1}{12}[\Omega^\wedge_k,[\Omega^\wedge_k,\varepsilon^\wedge_k]]\Big)\,.\tag{A.7}$$

Evaluating (A.7):

$$\begin{aligned}&\varepsilon^\wedge_k+\frac{1}{2}[\varepsilon^\wedge_k,\Omega^\wedge_k]+\frac{1}{12}[\Omega^\wedge_k,[\Omega^\wedge_k,\varepsilon^\wedge_k]]\\ &\quad-\frac{1}{2}\Big[\Omega^\wedge_k,\varepsilon^\wedge_k+\frac{1}{2}[\varepsilon^\wedge_k,\Omega^\wedge_k]+\frac{1}{12}[\Omega^\wedge_k,[\Omega^\wedge_k,\varepsilon^\wedge_k]]\Big]\\ &\quad+\frac{1}{6}\Big[\Omega^\wedge_k,\Big[\Omega^\wedge_k,\varepsilon^\wedge_k+\frac{1}{2}[\varepsilon^\wedge_k,\Omega^\wedge_k]+\frac{1}{12}[\Omega^\wedge_k,[\Omega^\wedge_k,\varepsilon^\wedge_k]]\Big]\Big]\\ &\quad-\frac{1}{24}\Big[\Omega^\wedge_k,\Big[\Omega^\wedge_k,\Big[\Omega^\wedge_k,\varepsilon^\wedge_k+\frac{1}{2}[\varepsilon^\wedge_k,\Omega^\wedge_k]+\frac{1}{12}[\Omega^\wedge_k,[\Omega^\wedge_k,\varepsilon^\wedge_k]]\Big]\Big]\Big]\\ &\quad+\dots\end{aligned}$$

leads to the expression

$$\begin{aligned}&\varepsilon^\wedge_k+[-\Omega^\wedge_k,\varepsilon^\wedge_k]+\frac{1}{2}[-\Omega^\wedge_k,[-\Omega^\wedge_k,\varepsilon^\wedge_k]]\\ &\quad+\frac{1}{6}[-\Omega^\wedge_k,[-\Omega^\wedge_k,[-\Omega^\wedge_k,\varepsilon^\wedge_k]]]\\ &\quad+\frac{5}{144}[-\Omega^\wedge_k,[-\Omega^\wedge_k,[-\Omega^\wedge_k,[-\Omega^\wedge_k,\varepsilon^\wedge_k]]]]\\ &\quad+\frac{1}{288}[-\Omega^\wedge_k,[-\Omega^\wedge_k,[-\Omega^\wedge_k,[-\Omega^\wedge_k,[-\Omega^\wedge_k,\varepsilon^\wedge_k]]]]]+\dots\,,\end{aligned}$$

which can be finally recognized as an approximation of

$$\exp(\mathrm{ad}(-\Omega^\wedge_k))\varepsilon^\wedge_k=\mathrm{Ad}(\exp(-\Omega^\wedge_k)))\varepsilon^\wedge_k\,.$$

This finishes the derivation of the Lie algebraic error prediction.

B The Special Orthogonal group $\mathsf{SO}(3)$

The $\mathsf{SO}(3)$ group is a set of orthogonal matrices with determinant one, whose elements geometrically represent rotations. Rotations in 3D can also be represented with an Euler vector (also called the axis-angle notation), where a vector $\phi=[\phi_1\ \phi_2\ \phi_3]^T\in\mathbb{R}^3$ denotes a rotation about the unit vector $\phi/|\phi|$ by the angle $|\phi|$. An interesting notion is that the Lie algebra

$\mathfrak{so}(3)$ is given as the skew symmetric matrix of the Euler vector

$$\phi^\wedge = \begin{bmatrix} 0 & -\phi_3 & \phi_2 \\ \phi_3 & 0 & -\phi_1 \\ -\phi_2 & \phi_1 & 0 \end{bmatrix} \in \mathfrak{so}(3)\,, \tag{B.1}$$

where $(\cdot)^\wedge : \mathbb{R}^3 \to \mathfrak{so}(3)$ and its inverse, $(\cdot)^\vee : \mathfrak{so}(3) \to \mathbb{R}^3$, follow trivially. The exponential map $\exp : \mathfrak{so}(3) \to \mathsf{SO}(3)$ is given as [20]

$$\begin{aligned}\exp(\phi^\wedge) &= \cos(|\phi|)I^3 \\ &\quad + (1-\cos(|\phi|))\frac{\phi\phi^T}{|\phi|^2} + \sin(|\phi|)\frac{\phi^\wedge}{|\phi|}.\end{aligned} \tag{B.2}$$

Furthermore, for an $\Phi \in \mathsf{SO}(3)$, the matrix logarithm, performing mapping $\log : \mathsf{SO}(3) \to \mathfrak{so}(3)$, is given as

$$\log(\Phi) = \begin{cases} \dfrac{\gamma}{2\sin(\gamma)}(\Phi - \Phi^T), & \text{if } \gamma \neq 0 \\ 0, & \text{if } \gamma = 0\,, \end{cases} \tag{B.3}$$

where $1 + 2\cos\gamma = \mathrm{Tr}(\Phi)$ and $\mathrm{Tr}(\,\cdot\,)$ designates the matrix trace. The adjoint operators Ad and ad for $\mathsf{SO}(3)$ are respectively given as

$$\mathrm{Ad}(\Phi) = \Phi \text{ and } \mathrm{ad}(\phi^\wedge) = \phi^\wedge\,. \tag{B.4}$$

Given the above definitions, we have all the needed ingredients to use the $\mathsf{SO}(3)$ group within the LG-EIF.