2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

A General Iterative Extended Kalman Filter Framework for State Estimation on Matrix Lie Groups

Ben Liu, Hua Chen, Wei Zhang

***Abstract*— In this paper, we focus on state estimation problem for nonlinear systems on joint matrix Lie group $\mathcal{G}$ and Euclidean space $\mathbb{R}^n$. We propose a general iterative Kalman filter, aiming to integrate the prediction step into the iteration scheme, which is not considered in the conventional iterative extended Kalman filter framework. Such an extra iteration scheme in the prediction step helps improving the accuracy of probability density function propagation through nonlinearities, which can further lead to more accurate estimations of the system states. In addition, the proposed framework unifies the Kalman filter based estimation schemes on studied manifold by adopting the perspective from Gaussian Bayesian inference. The improvement of the proposed framework is illustrated by the ES-GIKF algorithm that is instantiated from the proposed framework in a numerical simulation.**

I. Introduction

State estimation studies the problem of obtaining an estimate of the current system state given historical observations from the beginning to the present time [1], which plays an essential role in modern control systems. For applications such as control of robotic systems [2], the system states involve not only those lying on Euclidean spaces, but also those lying on a manifold. In such scenarios, the orientation of the robot, which belongs to *matrix Lie groups*, is indispensable for planning and control.

Some research has been focused on the state estimation problem on matrix Lie groups. *Full smoother* [3] and *fix-lag smoother* [4] have been applied to matrix Lie groups, by formulating and solving a large nonlinear optimization over manifolds. However, they are hard to be applied in high real-time scenarios due to computational complexity. To alleviate the computational consuming, *generalized Gaussian filter* on matrix Lie groups was proposed with concise form and computational efficiency [1] such as the *extended Kalman filter* (EKF) [5], [6]. This type of filter is also known as the Kalman filter (KF) family, which comes from the Gaussian Bayesian inference and comprises the *prediction* step and the *correction* step. The EKF [7] and *unscented Kalman filter* (UKF or SPKF) [8] were respectively implemented on matrix Lie groups using *linearization* and *sigma point transformation* to propagate the *probability density function* (PDF). But they were developed using distinct concepts, with the *error state* and random variables on matrix Lie groups, respectively. *Error state* and random variables on matrix Lie groups respectively. In order to deduce the linearization error in propagating PDF in EKF, the *iterative extended Kalman filter* (IEKF) on general manifold has been proposed in [9], which revalues the linearization point to improve the accuracy of the linearization and has been successfully applied in practice [10]. For a special class of *group affine* system on Lie groups, the *invariant extended Kalman filter* (InEKF) [11] was proposed with the help of error state and outperformed the EKF [12], [6].

However, when addressing state estimation on matrix Lie groups using KF families, two issues remain unaddressed. One is that the existing IEKF framework only applies the iterative strategy in the correction step, neglecting its application in the prediction step. The iterative strategy aims to propagate the PDF more precisely through nonlinearity to yield more accurate estimates [13], [14]. However, the PDF propagation between matrix Lie groups is more complicated, which occurs in the prediction step. This is because two independent Gaussian variables are no longer Gaussian after compounding [15], but this holds for Euclidean space. Moreover, in the case where the InEKF degenerates to the "imperfect" InEKF due to the breaking of group affine properties [12], the invariant property in the prediction step is corrupted by the nonlinearity. Therefore, the PDF propagation in the prediction step for matrix Lie groups is more difficult to keep accurate, which should be paid the same attention as the correction step. Another issue is the non-unified formulation involving utilization of the error state rather than the state itself [7], this indirect way contrasts with the Euclidean case. Only a few papers focus on the state itself [8], [16]. This non-unified formulation creates a gap for interpreting EKF on Lie groups through Bayesian inference from the perspective of random variables.

In this paper, we focus on the state estimation problem on joint Lie groups and Euclidean space $\mathcal{G} \times \mathbb{R}^n$ and aim to address aforementioned deficiencies by introducing a general iterative Kalman filter (GIKF) framework. We unify the KF families into a single state estimation problem formulation and subsequently extend the iterative strategy from the conventional IEKF to the prediction step in order to address the proposed state estimation problem.

The main contributions are summarized as follows. First, we present a novel formulation of state estimation problem on $\mathcal{G} \times \mathbb{R}^n$, which directly works with random variables on $\mathcal{G} \times \mathbb{R}^n$. This formulation unifies the KF family methods

This work was supported in part by National Natural Science Foundation of China under Grant No. 62073159 and Grant No. 62003155, in part by the Shenzhen Science and Technology Program under Grant No. JCYJ20200109141601708, and in part by the Shenzhen Key Laboratory of Control Theory and Intelligent Systems, under grant No.ZDSYS20220330161800001.

The authors are with Shenzhen Key Laboratory of Control Theory and Intelligent Systems, Southern University of Science and Technology, Shenzhen 518055, China; and with the School of System Design and Intelligent Manufacturing, Southern University of Science and Technology, Shenzhen, 518055, China. (Emails: 12131156@mail.sustech.edu.cn, chenh6@sustech.edu.cn, zhangw3@sustech.edu.cn)

on $\mathcal{G} \times \mathbb{R}^n$, and further offers a unified perspective on KF family for state estimation problem over Euclidean space and matrix Lie groups. Second, a general iterative Kalman filter framework is developed, which integrates the prediction step in an iterative framework while the conventional IEKF framework does not. Such an additional iteration in the prediction step provides a more accurate PDF propagation, and is consistent with the conventional IEKF in the correction step. Third, we instantiate the proposed framework and synthesize an explicit ES-GIKF algorithm. Such an algorithm performs an additional re-linearization in the prediction step in an efficient way and leads to enhanced estimation accuracy compared to conventional IEKF.

II. Preliminaries and Problem Formulation

To establish a comprehensive formulation of the estimation problem on $\mathcal{G} \times \mathbb{R}^n$, it is necessary to first specify the definition of Gaussian random variables on $\mathcal{G} \times \mathbb{R}^n$. Let $\xi \in \mathbb{R}^d$ be the zero-mean Gaussian variable, then the corresponding Gaussian variable on manifold $\mathcal{G}$, denoted by $X \in \mathcal{G}$ can be expressed as [15]:

$$X = \bar{X}\mathrm{Exp}(\xi) \qquad \xi \sim \mathcal{N}(0, \Sigma_\xi) \tag{1}$$

where $\bar{X} \in \mathcal{G}$ is some noise-free element. $\mathrm{Exp} : \mathbb{R}^d \to \mathfrak{g} \to \mathcal{G}$ is a vectorized version of the matrix exponential map [3]. It can be shown that $\bar{X}$ and Σ_ξ are the mean and covariance [1] of $p(X)$, which represents the PDF of X [15]. Based on (1), we can further consider the case where another random variable $x \in \mathbb{R}^n$ is jointly Gaussian with ξ. This joint distribution is referred to as the *standard Gaussian* for X and x, and is denoted as:

$$\mathbf{z} \sim \mathcal{N}_L(\bar{\mathbf{z}}, P) \tag{2}$$

where $\bar{\mathbf{z}} = (\bar{X}, \bar{x})$, $\bar{x}$ is the mean of x, and P is the covariance of $[\xi^T, x^T]^T$. With this definition, we can fit the random variables on $\mathcal{G} \times \mathbb{R}^n$ in Gaussian distribution and easily calculate the expectation similar to $\mathbb{R}^n$.

In this paper, we consider the following discrete-time nonlinear system on $\mathcal{G} \times \mathbb{R}^n$ as the underlying dynamical system

$$\begin{aligned} \mathbf{z}_k &= f_{k-1}(\mathbf{z}_k, w_k) \\ y_k &= h_k(\mathbf{z}_k, v_k) \end{aligned} \tag{3}$$

where $\mathbf{z}_k \triangleq (X_k, x_k) \in \mathcal{G} \times \mathbb{R}^n$ is the system state, $y_k \in \mathbb{R}^m$ is the observation and they both are random variables. The associated *Lie algebra* [18] $\mathfrak{g}$ of $\mathcal{G}$ has dimension d. The subscript k represents the time step which means $\mathbf{z}_k = \mathbf{z}(k\Delta t)$, where Δt is the sampling period. f_{k-1} is the dynamics model which is a nonlinear function that maps $\mathcal{G} \times \mathbb{R}^n$ to $\mathcal{G} \times \mathbb{R}^n$. h_k is the observation model which is a nonlinear function that maps $\mathcal{G} \times \mathbb{R}^n$ to $\mathbb{R}^m$. Both f_{k-1} and h_k are assumed to be smooth. $w_k \sim \mathcal{N}(0, \Sigma_{w_k})$ and $v_k \sim \mathcal{N}(0, \Sigma_{v_k})$ are the independent Gaussian noise. This type of system is common in practice.

With the definition of the standard Gaussian in (1), we consider the estimation problem of this system using the generalized Gaussian filter can be formulated as:

[1] See [17] for detailed definitions of mean and covariance on Lie groups

Problem 1.

- *(Prediction) Given $\mathbf{z}_{k-1} \sim \mathcal{N}_L(\hat{\mathbf{z}}_{k-1}, \hat{P}_{k-1})$, fit $p(\mathbf{z}_k)$ as $\mathcal{N}_L(\check{\mathbf{z}}_k, \check{P}_k)$ by dynamics model in* (3).
- *(Correction) Subsquently fit $p(\mathbf{z}_k|y_k)$ as $\mathcal{N}_L(\hat{\mathbf{z}}_k, \hat{P}_k)$ by observation model in* (3).

The conditional expectation of $p(\mathbf{z}_k|y_k)$ is $\hat{\mathbf{z}}_k$, which gives the estimate of $\mathbf{z}_k$. And the conditional PDF $p(\mathbf{z}_k|y_k)$ can be used for the prediction step at the next time step.

Problem 1 is very challenging particularly due to the fact that the standard Gaussian distribution will not preserve through nonlinearity in general. The key issue lies in the difficulty of accurately using only the mean and covariance to characterize the non-Gaussian PDF. Various algorithms belonging to KF families, including EKF, UKF, and IEKF, all strive to solve this issue.

III. General Iterative Kalman Filter Framework

In this section, we propose a general iterative Kalman filter (GIKF) framework to address the problem 1, with further development of Gaussian Bayesian inference on the standard Gaussian, which serves for the GIKF framework.

A. Framework

The proposed framework contains two re-processing steps, namely the prediction step and the correction step. The prediction step forward propagates the PDF of the system state distribution from time step $k-1$ to time step k through the system model $f_{k-1}(\cdot)$. The correction step first utilizes the observation model of y_k in (3) to fit $p(\mathbf{z}_k, y_k|Y_{0:k-1})$ as a standard Gaussian based on the prior PDF of $\mathbf{z}_k$. Then the conditional PDF $p(\mathbf{z}_k|y_k, Y_{0:k-1})$ is approximated as a standard Gaussian using $p(\mathbf{z}_k, y_k|Y_{0:k-1})$. The schematic diagram of the overall framework is given in Fig. 1.

In the prediction step, we aim to approximate the resulting PDF $p(\mathbf{z}_k|Y_{0:k-1})$ for time step k as a standard Gaussian distribution $\mathcal{N}_L(\check{\mathbf{z}}_k, \check{P}_k)$, where $Y_{0:k-1}$ represents the observations from time step 0 to $k-1$. This step can be represented as

$$\mathcal{N}_L(\check{\mathbf{z}}_k, \check{P}_k) = \Psi_{f_{k-1}}(\hat{\mathbf{z}}_{k-1}, \hat{P}_{k-1}, \tilde{\mathbf{z}}_{k-1}) \tag{4}$$

where $\mathcal{N}_L(\hat{\mathbf{z}}_{k-1}, \hat{P}_{k-1})$ is the PDF $\mathbf{z}_{k-1}$ at time step $k-1$. $\Psi_{f_{k-1}}$ represents different mapping that passes PDF into the standard Gaussian by f_{k-1} such as linearization in EKF, and the sigma point transformation in UKF. $\tilde{\mathbf{z}}_{k-1}$ is the operating point, which is desired to be close to the true state. Commonly, they are chosen as the mean $\hat{\mathbf{z}}_{k-1}$ in conventional IEKF.

For the correction step, $p(\mathbf{z}_k, y_k|Y_{0:k-1})$ is always obtained by Bayesian inference $p(y_k|\mathbf{z}_k, Y_{0:k-1})p(\mathbf{z}_k|Y_{0:k-1})$ and this process can be represented as

$$\mathcal{N}_L((\check{\mathbf{z}}_k, \check{y}_k), \check{\Sigma}_A) = \Upsilon_{h_k}(\check{\mathbf{z}}_k, \check{P}_k, \tilde{\mathbf{z}}_k) \tag{5}$$

where $\mathcal{N}_L((\check{\mathbf{z}}_k, \check{y}_k), \check{\Sigma}_A)$ represents the standard Gaussian of X_k with $[x_k^T, y_k^T]^T$, $\tilde{z}_k$ are some operating points and Υ_{h_k} represents different methods of passing PDF for $(\mathbf{z}_k, y_k)$ like (4). With the standard Gaussian in (5), using Gaussian Bayesian inference, the conditional posterior $p(\mathbf{z}_k, |y_k, Y_{0:k-1})$ can be easily approximated by

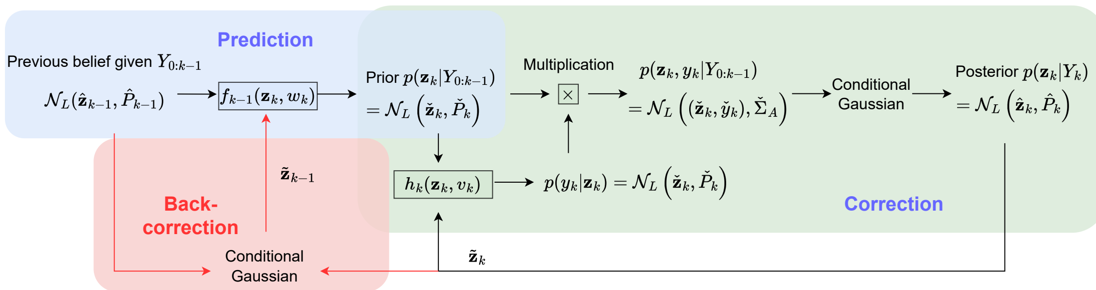

Fig. 1. The general iterative Kalman filter framework.

$\mathcal{N}_L(\hat{\mathbf{z}}_k, \hat{P}_k)$. Note that the Gaussian Bayesian inference gives a concise form on the Euclidean space and is the core theoretical foundation for the Kalman filter, Bayesian inference for the standard Gaussian distribution can also give a simple form, which will be presented in Sec. III-B.

In the above process, (5) and (4) are iteratively recomputed by setting the operating points as

$$\tilde{\mathbf{z}}_k^i = E(\mathbf{z}_k | y_k, \tilde{\mathbf{z}}_k^{i-1}), \tag{6}$$

and

$$\tilde{\mathbf{z}}_{k-1}^i = E(\mathbf{z}_{k-1} | \mathbf{z}_k = \tilde{\mathbf{z}}_k^i, \tilde{\mathbf{z}}_{k-1}^{i-1}), \tag{7}$$

respectively. Note that the result of (6) is used in (7).

The superscript i is the iteration index. The initial condition is set to be $\tilde{\mathbf{z}}_{k-1}^0 = \hat{z}_{k-1}$, $\tilde{\mathbf{z}}_k^0 = \check{\mathbf{z}}_k^0$. The re-evaluation process (6) involves utilizing the operating point from the previous iteration to generate the posterior distribution $p(\mathbf{z}_k | y_k, Y_{0:k-1})$. The obtained posterior distribution is then used to derive the estimator for $\mathbf{z}_k$, which subsequently serves as the new operating point for the next iteration.

The re-evaluation process (7) is similar to the case in (6). The main difference lies in the consideration of posterior distribution $p(\mathbf{z}_{k-1} | \mathbf{z}_k, Y_{0:k-1})$, which takes the estimate of $\mathbf{z}_k$ in (6) by incorporating the information of y_k as the "pseudo-observation" for $\mathbf{z}_k$. The process (7) is essentially a simplification of calculating the PDF of $p(\mathbf{z}_{k-1} | y_k, Y_{0:k-1})$, which stems from $E(x_{k-1} | y_k, Y_{0:k-1}) = E(x_{k-1} | x_k = \tilde{x}_k, Y_{0:k-1})$ for the Euclidean case in KF families. With the re-evaluations (7) and (6), iteratively processing the prediction step and the correction step leads to our GIKF framework. In fact, (6) can be considered as just processing the correction step, but (7) is the extra step separated from either the prediction or correction step. We refer to this extra step as *back-correction* step, as shown in Fig 1.

The proposed framework unifies the KF families on $\mathcal{G} \times \mathbb{R}^n$, all different methods such as EKF, UKF can be interpreted as applying the different methods of passing PDF through the nonlinearity. Despite most EKF-based methods on $\mathcal{G} \times \mathbb{R}^n$ focus on the error state dynamics in an indirect way [6], which is different from the standard EKF on $\mathbb{R}^n$, the error state dynamics can still be integrated into PDF propagation in a direct way. This unifies can be seen in Sec IV, where error state dynamics is utilized for propagating PDF using the linearization method.

In particular, the iterative strategy is integrated into this framework, which embraces conventional iterative algorithms in KF families such as IEKF and *iterative sigma point Kalman filter* (ISPKF) on $\mathcal{G} \times \mathbb{R}^n$. The proposed framework differs from conventional IEKF/ISPKF framework primarily in the inclusion of a back-correction step. The back-correction step extends the iterative strategy of the update step in the conventional IEKF to the prediction step, which can give an operating point closer to the true state in the prediction step. This more accurate operating point has been illustrated to lead to more accurate PDF propagation [13]. Moreover, the back-correction step provides a more concise form of (7) instead of directly calculating $E(\mathbf{z}_{k-1} | y_k, Y_{0:k-1})$, this formulation improves computational efficiency by using intermediate variables in other steps such that the GIKF framework can be implemented for real-time tasks that are compatible with the conventional IEKF framework.

B. Bayesian Inference on Standard Gaussian

In the previous subsection, we assume that the Bayesian inference for the standard Gaussian has a simple form similar to the Euclidean case. In this subsection, we outline the main idea for implementing this.

Following the observation that $X = \bar{X}\text{Exp}(\zeta)$ is a function of ζ, we consider the case that ζ is not a zero-mean Gaussian variable, i.e., $\zeta \sim \mathcal{N}(\bar{\zeta}, \Sigma_\zeta)$. In this case, it can be transformed into "zero-mean" form in (1) using (29a):

$$X = \bar{X}\text{Exp}(\bar{\zeta} + \varepsilon) \approx \bar{X}\text{Exp}(\bar{\zeta})\text{Exp}(J_r(\bar{\zeta})\varepsilon) \tag{8}$$

where $\varepsilon = \zeta - \bar{\zeta}$ and $J_r(\bar{\zeta})\varepsilon \sim \mathcal{N}(0, J_r(\bar{\zeta})\Sigma_\zeta J_r(\bar{\zeta})^T)$. Thus we can fit X in (8) as a Gaussian distribution. Furthermore, note that ζ and $J_r(\bar{\zeta})\varepsilon$ are affine. If $l \in \mathbb{R}^m$ is jointly Gaussian with either of them, the relationship holds for the other. This means we can apply the Gaussian Bayesian inference on l and the Gaussian variable in the tangent space that represents X and then use (8) to make this PDF into a standard Gaussian. For example, consider the correction step, when X_k in $\mathbf{z}_k$ is written as $X_k = \tilde{X}_k\text{Exp}(\zeta_k)$, and ζ_k is jointly Gaussian with x_k. If the measurement y_k can be approximated as jointly Gaussian with $[\zeta_k^T, x_k^T]^T$, once we get the posterior estimation $([\zeta_k^T, x_k^T]^T | y_k) \sim \mathcal{N}([\delta_X^T, \hat{x}_k^T]^T, \Sigma')$ under the measurement y_k by Gaussian Bayesian inference on Euclidean space, we can immediately fit the posterior distribution for $p(\mathbf{z}_k | y_k)$ as

$$\mathcal{N}_L((\tilde{X}_k\text{Exp}(\delta_X), \hat{x}_k), \Phi\Sigma'\Phi^T) \tag{9}$$

where $\Phi \triangleq \text{diag}\{J_r(\delta_X), I\}$. There exists a little approximation for applying Gaussian Bayesian inference on the standard Gaussian variables.

IV. Algorithm for GIKF framework

The GIKF framework has been proposed in the last section, in this section, we instantiate this framework by choosing linearization to propagate the PDF through the nonlinearity. The linearization on $\mathcal{G}$ is implemented with the help of the error state like many works before, however, with further consideration of the standard Gaussian fitting in the GIKF framework. We call this instantiation as *error state general iterative Kalman filter* (ES-GIKF) algorithm.

Some operators are defined for notation simplicity:

$$\begin{aligned}\mathbf{z} \boxminus \tilde{\mathbf{z}} &\triangleq \begin{bmatrix}\text{Log}(\tilde{X}^{-1}X)\\ x-\tilde{x}\end{bmatrix} \in \mathbb{R}^{d+n}\\ \tilde{\mathbf{z}} \boxplus \begin{bmatrix}e^X\\ e^x\end{bmatrix} &\triangleq (\tilde{X}\text{Exp}(e^X), \tilde{x}+e^x)\end{aligned} \tag{10}$$

where Log is the local inverse of Exp.

A. Prediction

This step aims to propagate the PDF from time step $k-1$ to k. With the prior PDF $\mathcal{N}_L(\hat{\mathbf{z}}_{k-1}, \hat{P}_{k-1})$ for $\mathbf{z}_{k-1}$, let $\tilde{\mathbf{z}}_{k-1}$ be the operating point, i.e. linearization point, then we define an auxiliary equation (also the "error states dynamics")

$$\begin{aligned}\alpha(\mathbf{z}_{k-1}) &\triangleq f_{k-1}(\mathbf{z}_{k-1}, w_k) \boxminus f_{k-1}(\tilde{z}_{k-1}, 0)\\ &\approx F_{k-1}e_{k-1} + w_k'\end{aligned} \tag{11}$$

where $e_{k-1} = \mathbf{z}_{k-1} \boxminus \tilde{\mathbf{z}}_{k-1}$ and F_{k-1} and w_k' have the formulation of

$$F_{k-1} = \left.\frac{\partial(f_{k-1}(\tilde{\mathbf{z}}_{k-1} \boxplus \phi, 0) \boxminus f_{k-1}(\tilde{\mathbf{z}}_{k-1}, 0))}{\partial \phi}\right|_{\phi=0} \tag{12}$$

and

$$w_k' = \left.\frac{\partial f_{k-1}(\tilde{\mathbf{z}}_{k-1}, w) \boxminus f_{k-1}(\tilde{\mathbf{z}}_{k-1}, 0)}{\partial w}\right|_{w=0} w_k \sim \mathcal{N}(0, Q_k) \tag{13}$$

According to the auxiliary equation (11), we have

$$\begin{aligned}\mathbf{z}_k &= f_{k-1}(\mathbf{z}_{k-1}, w_k)\\ &\approx f_{k-1}(\tilde{z}_{k-1}, 0) \boxplus (F_{k-1}e_{k-1} + w_k')\end{aligned} \tag{14}$$

Further, e_{k-1} can be approximated as a Gaussian variable $e_{k-1} \sim \mathcal{N}(\tilde{e}_{k-1}, \tilde{P}_{k-1})$ using the standard Gaussian of $\mathbf{z}_{k-1}$ according to (29b), where (Φ is defined in (9))

$$\begin{aligned}\tilde{e}_{k-1} &= \hat{\mathbf{z}}_{k-1} \boxminus \tilde{\mathbf{z}}_{k-1}\\ \tilde{P}_{k-1} &= \Phi(\tilde{e}_{k-1})^{-1}\hat{P}_{k-1}\Phi(\tilde{e}_{k-1})^{-T}\end{aligned} \tag{15}$$

Therefore $\mathbf{z}_k$ in (14) is a non-zero-mean form standard Gaussian, which we can be fitted as a standard Gaussian $\mathcal{N}_L(\check{\mathbf{z}}_k, \check{P}_k)$ following the idea in Sec. III-B, where

$$\begin{aligned}\check{\mathbf{z}}_k &= f_{k-1}(\tilde{\mathbf{z}}_{k-1}, 0) \boxplus \check{\alpha}\\ \check{P}_k &= \Phi(\check{\alpha})(F_{k-1}\tilde{P}_{k-1}F_{k-1}^T + Q_k)\Phi(\check{\alpha})^T\end{aligned} \tag{16}$$

where $\check{\alpha} \triangleq F_{k-1}\tilde{e}_{k-1}$.

B. Correction

This step aims to get the posterior standard Gaussian of $p(\mathbf{z}_k|y_k)$ by the standard Gaussian $p(\mathbf{z}_k, y_k)$. Let $\tilde{\mathbf{z}}_k$ be the linearization point, y_k can be approximated as

$$y_k = h_k(\mathbf{z}_k, v_k) \approx h_k(\tilde{\mathbf{z}}_k, 0) + H_k e_k + v_k' \tag{17}$$

where $e_k = \mathbf{z}_k \boxminus \tilde{\mathbf{z}}_k$. H_k and v_k' are

$$H_k = \left.\frac{\partial(h_k(\tilde{\mathbf{z}}_k \boxplus \phi, 0) - h_k(\tilde{\mathbf{z}}_k, 0))}{\partial \phi}\right|_{\phi=0} \tag{18}$$

$$v_k' = \left.\frac{\partial h_k(\tilde{\mathbf{z}}_k, v) \boxminus h_k(\tilde{\mathbf{z}}_k, 0)}{\partial v}\right|_{v=0} v_k \sim \mathcal{N}(0, R_k) \tag{19}$$

Under the prior (16) and similar to (15), it can be get that $e_k \sim \mathcal{N}\left(\tilde{e}_k, \tilde{P}_k\right)$, where $\tilde{P}_k = \Phi(\tilde{e}_k)^{-1}\check{P}_k\Phi(\tilde{e}_k)^{-T}$ and $\tilde{e}_k = \check{\mathbf{z}}_k \boxminus \tilde{\mathbf{z}}_k$. Therefore y_k and e_k are the jointly Gaussian as

$$\begin{bmatrix}e_k\\ y_k\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\tilde{e}_k\\ \tilde{y}_k + H_k\tilde{e}_k\end{bmatrix}, \begin{bmatrix}\tilde{P}_k & \tilde{P}_k H_k^T\\ H_k\tilde{P}_k & H_k\tilde{P}_k H_k^T + R_k\end{bmatrix}\right) \tag{20}$$

where $\tilde{y}_k = h_k(\tilde{\mathbf{z}}_k, 0)$. By the Bayesian rule, we can get $(e_k|y_k) \sim \mathcal{N}(\hat{e}_k, \hat{\Sigma}_{e_k})$, where

$$\begin{aligned}\hat{e}_k &= \tilde{e}_k + K_k(y_k - \tilde{y}_k - H_k\tilde{e}_k)\\ \hat{\Sigma}_{e_k} &= (I - K_k H_k)\tilde{P}_k\end{aligned} \tag{21}$$

where $K_k = \tilde{P}_k H_k^T(H_k\tilde{P}_k H_k^T + R_k)^{-1}$ is the *Kalman gain*. Similar to (16), the standard Gaussian of $p(\mathbf{z}_k|y_k)$ is given by $\mathcal{N}_L(\tilde{\mathbf{z}}_k \boxplus \hat{e}_k, \Phi(\hat{e}_k)\hat{\Sigma}_{e_k}\Phi(\hat{e}_k)^T)$. Therefore we can get the estimator of $\mathbf{z}_k$, which is also the next operating point by (6):

$$\tilde{\mathbf{z}}_k^{i+1} = \tilde{\mathbf{z}}_k^i \boxplus \hat{e}_k^i \tag{22}$$

The algorithm terminates when $\hat{e}_k^i$ is sufficiently small. If we regard $\hat{e}_k^i$ as zero when converging, then $\Phi(\hat{e}_k^i) = I$, which leads to $\hat{\Sigma}_{e_k}^i$ is the covariance for $p(\mathbf{z}_k|y_k)$.

C. Back-correction

This step aims to get the posterior standard Gaussian of $p(\mathbf{z}_{k-1}|\mathbf{z}_k = \tilde{\mathbf{z}}_k^{i+1})$. Substitute $\mathbf{z}_k = \tilde{\mathbf{z}}_k \boxplus e_k$ into (14), then ignore the second and higher terms of $e_{k-1} - \tilde{e}_{k-1} + w_k'$ and $e_k - \tilde{e}_k$, then we have

$$e_k = \underbrace{\Phi(\tilde{e}_k)^{-1}\Phi(\check{\alpha})}_{\Pi}(F_{k-1}(e_{k-1} - \tilde{e}_{k-1}) + w_k') + \tilde{e}_k \tag{23}$$

Therefore e_k and e_{k-1} are jointly Gaussian. Note the condition $\mathbf{z}_k = \tilde{\mathbf{z}}_k^{i+1}$ means $e_k = \hat{e}_k$ by (22), thus we can apply Bayesian inference to get

$$E(e_{k-1}|\hat{e}_k) = \tilde{e}_{k-1} + \tilde{P}_{k-1}F_{k-1}^T\Pi^T\tilde{P}_k^{-1}(\hat{e}_k - \tilde{e}_k) \tag{24}$$

Therefore $p(\mathbf{z}_{k-1}|\mathbf{z}_k = \tilde{\mathbf{z}}_k^{i+1})$ can be obtained, whose mean which is also the next operating point for $\tilde{\mathbf{z}}_{k-1}$ according to (7), is given by

$$\tilde{\mathbf{z}}_{k-1}^{i+1} = \tilde{z}_{k-1}^i \boxplus E(e_{k-1}|e_k = \hat{e}_k)^i \tag{25}$$

Remark 1. *In the back-correction step, we directly manipulate $p(\mathbf{z}_{k-1}|\mathbf{z}_k = \tilde{\mathbf{z}}_k^{i+1})$ to avoid deal with $p(\mathbf{z}_{k-1}|y_k)$, resulting in a more concise formulation. In fact, they result* *in the same results, which can be verified by developing the jointly Gaussian of y_k and e_{k-1} using* (23) *and* (17).

The ES-GIKF algorithm is summarized in Algorithm 1. The significant improvement of this algorithm is integrating the prediction step into the iterative loop by the back-correction step, which allows more accurate linearization points to Propagate PDF in the prediction step. Moreover, if we get rid of the back-correction step and only have the prediction for only once, then only the correction step in Algorithm 1 needs iteration and this becomes the conventional IEKF framework, which denotes C-IEKF.

Algorithm 1 ES-GIKF Algorithm

Input: $\hat{\mathbf{z}}_{k-1}, \hat{P}_{k-1}$
Initialization: $\tilde{\mathbf{z}}_{k-1}^0 = \hat{\mathbf{z}}_{k-1}$, $\tilde{\mathbf{z}}_k^0 = f_{k-1}(\hat{\mathbf{z}}_{k-1}, 0)$
for $i = 0 : N$ **do** ▷ N is the maximum ietration
  Prediction:
  Calculate F_{k-1}, Q_k, $\tilde{P}_{k-1}$ in (12), (13), (15).
  Calculate $\check{\mathbf{z}}_k$ and $\check{P}_k$ in (16).
  Correction:
  Calculate H_k, R_k, $\tilde{P}_k$ in (18), (19), (20).
  $K_k = \tilde{P}_k H_k^T (H_k \tilde{P}_k H_k^T + R_k)^{-1}$
  $\hat{u}_k = K_k(y - h_k(\tilde{\mathbf{z}}_k^i, 0) - H_k \tilde{e}_k^i)$
  $\tilde{\mathbf{z}}_k^{i+1} = \tilde{\mathbf{z}}_k^i \boxplus (\tilde{e}_k^i + \hat{u}_k)$
  Back-correction:
  Calculate Π as (23)
  $\tilde{\mathbf{z}}_{k-1}^{i+1} = \tilde{\mathbf{z}}_{k-1}^i \boxplus (\tilde{e}_{k-1}^i + \tilde{P}_{k-1} F_{k-1}^T \Pi^T \tilde{P}_k^{-1} \hat{u}_k)$
  if $(\tilde{e}_k^i + \hat{u}_k)$ sufficient small **then**
    break
  end if
end for
$\hat{\mathbf{z}}_k = \tilde{\mathbf{z}}_k^i$ and $\hat{P}_k = (I - K_k H_k)\tilde{P}_k$
Output: $\hat{\mathbf{z}}_k$, $\hat{P}_k$

V. Numerical Simulation

Here we apply the ES-GIKF algorithm 1 to the nonlinear system (26). We take the results from EKF as the benchmark. In order to compare the GIKF framework with the conventional IKEF framework, we also applied the C-IEKF algorithm to the problem and compared the results. Since ES-GIKF and C-IEKF are derived from the perspective of the conditional expectation based on the standard Gaussian, we also apply the existing method in [19] as the reference for the conventional IKEF framework, which is derived from *Maximum A Posterior* perspective, denotes MAP-IEKF.

A. System model

The system is modified from a benchmark problem [20]:

$$\begin{aligned} R_k &= R_{k-1}\text{Exp}\left(0.01[1, x_{k-1} - 7, x_{k-1}^2 - 49]^T)\text{Exp}(w_k^R)\right) \\ x_k &= 0.5x_{k-1} + \frac{25x_{k-1}}{1 + x_{k-1}^2} + w_k^x \\ y_k &= \frac{x_k^2}{20} + v_k \end{aligned} \tag{26}$$

where $R \in SO(3)$ and $x \in \mathbb{R}$ are the states, $y \in \mathbb{R}$ is the observation. w_k^R, w_k^x and v_k are independent Gaussian noises. The Jacobian matrix can be calculated as

$$\begin{aligned} F_{k-1} &= \begin{bmatrix} \text{Exp}(-\omega) & 0.01J_r(\omega)[0, 1, 2x_{k-1}]^T \\ 0 & 0.5 + \frac{25(1-x_{k-1}^2)}{(1+x_{k-1}^2)^2} \end{bmatrix} \\ w_k' &= \begin{bmatrix} w_k^R \\ w_k^x \end{bmatrix} \quad H_k = \begin{bmatrix} 0 & 0 & 0 & x_k/10 \end{bmatrix} \end{aligned} \tag{27}$$

where $\omega = 0.01[1, x_{k-1} - 7, x_{k-1}^2 - 49]^T$.

B. Numerical Results

In all the estimation algorithms, the initial conditions and covariance are the same in different algorithms. We use mean squared error (MSE) as the evaluation criterion:

$$\text{MSE} = \frac{1}{N}\sum_{k=0}^{N} \|(\hat{R}_k, \hat{x}_k) \boxminus (R_{k,t}, x_{k,t})\|^2 \tag{28}$$

where $\hat{(\cdot)}_k$ is the estimation and $(\cdot)_{k,t}$ is the true state.

Fig. 2 depicts the relationship between the relative mean squared error (MSE) and the initial error. The MSE is calculated for fixed-length trajectories using different methods, with the relative MSE plotted in Fig. 2 based on the EKF's MSE. The improved performance of the ES-GIKF algorithm in the prediction step is demonstrated in Fig. 3, where the relative MSE of the final linearization point in the prediction step after convergence is shown. This emphasizes the capability of the proposed algorithm to attain a more precise linearization point.

In general, the proposed method ES-GIKF exhibits reduced error. The two methods employed within the conventional IEKF framework, namely C-IEKF, and MAP-IEKF, yield similar results, which is reasonable given that they solely utilize an iteration strategy during the correction step. The fact that C-IEKF and MAP-IEKF demonstrate lower MSE compared to the EKF indicates that the iterative strategy can lead to a more accurate linearization point, which serves as an estimate for the state. Moreover, ES-GIKF outperforms C-IEKF and MAP-IEKF as the initial error increases. This can be attributed to the fact that with increasing initial error, the linearization point for propagating the PDF becomes increasingly inaccurate in both the prediction step and the correction step. ES-GIKF, on the other hand, incorporates an additional iterative strategy in the prediction step, resulting in a more accurate linearization point. This improvement is evident in Fig. 3, where the MSE of the linearization point in the prediction step, which is also the estimate for $\mathbf{z}_{k-1}$ using y_k, is depicted. The figure reveals that the estimate of $\mathbf{z}_{k-1}$ exhibits closer proximity to the true state subsequent to the iteration in the prediction step for estimating $\mathbf{z}_k$. This significant improvement shows proposed framework can propagate PDF through nonlinearity for the prediction better than the conventional IEKF framework. Based on the fact that $\mathbf{z}_{k-1}$ is also estimated at time step k, the proposed framework can be regarded as a "local smoother", but embedded into the KF family framework.

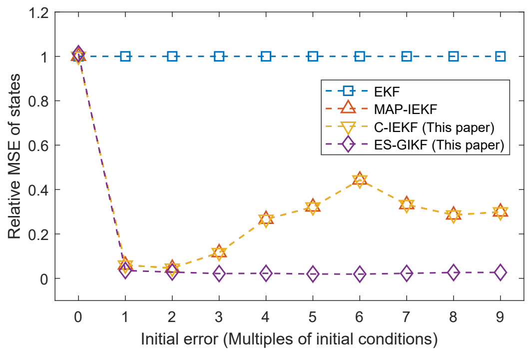

Fig. 2. The relative percentage of MSE for states of different methods based on the MSE of EKF. Lower is better. The initial error means the initial derivation between the mean of the PDF of the first state and the first state's true value, which is denoted by the 'initial condition'.

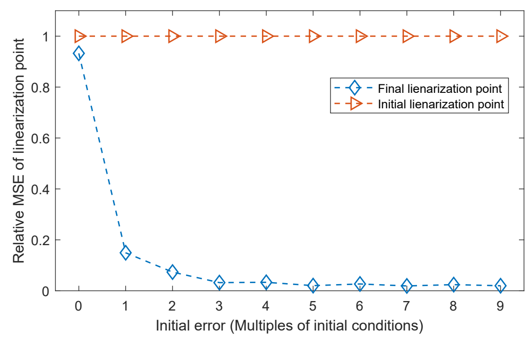

Fig. 3. The MSE of the linearization point for the prediction step. Lower is better. 'Initial linearization point' and 'Final linearization point' respectively represent the first and the convergent result of the back-correction step in ES-GIKF.

The computational cost is intricately linked to the number of iterations employed. Thus, we present the percentage of calculation time consumed by the prediction step and the correction step for each iteration, offering a comparative measure of the computational overhead relative to the conventional IEKF. The calculation time proportions of the prediction step and the correction step are respectively 63.65% and 36.35% overall. The higher computational burden in the prediction step can be attributed to the recurrent covariance propagation and the increased complexity of the prediction model as defined by equation (27).

VI. Conclusion

In this paper, we formulate the estimation problem for $\mathcal{G} \times \mathbb{R}^n$ to unify the KF families and propose a general iterative Kalman filter framework to address the problem. The GIKF integrates the iterative strategy into the prediction step in order to improve the accuracy of the PDF propagation. The numerical result illustrates more accurate PDF propagation in the prediction step leads to more accurate estimates than the conventional IEKF. This framework has the potential to improve the accuracy of the estimate for the system that has high nonlinearity in the motion model. Moreover, the proposed framework can enhance observation accuracy for systems that simplify observations to rely solely on the current state, even though certain observations depend on both the current and previous states [10]. This is achieved by estimating not only the current state but also the state at the previous time in the proposed framework.

Appendix

A. Useful Approximation

$$\mathrm{Exp}(\phi_1+\phi_2) \approx \mathrm{Exp}(\phi_1)\mathrm{Exp}(J_r(\phi_1)\phi_2) \tag{29a}$$

$$\mathrm{Log}(\mathrm{Exp}(\phi_1)\mathrm{Exp}(\phi_2)) \approx \phi_1 + J_r^{-1}(\phi_1)\phi_2 \tag{29b}$$

holds for small ϕ_2. See [1, equation (7.75)] for details.

References

[1] T. D. Barfoot, *State estimation for robotics*. Cambridge University Press, 2017.

[2] D. Wisth, M. Camurri, and M. Fallon, "Vilens: Visual, inertial, lidar, and leg odometry for all-terrain legged robots," *IEEE Transactions on Robotics*, vol. 39, no. 1, pp. 309–326, 2023.

[3] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, "On-manifold preintegration for real-time visual–inertial odometry," *IEEE Transactions on Robotics*, vol. 33, no. 1, pp. 1–21, 2017.

[4] J.-H. Kim, S. Hong, G. Ji, S. Jeon, J. Hwangbo, J.-H. Oh, and H.-W. Park, "Legged robot state estimation with dynamic contact event information," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6733–6740, 2021.

[5] M. Bloesch, M. Hutter, M. A. Hoepflinger, S. Leutenegger, C. Gehring, C. D. Remy, and R. Siegwart, "State estimation for legged robots-consistent fusion of leg kinematics and imu," *Robotics*, vol. 17, pp. 17–24, 2013.

[6] R. Hartley, M. Ghaffari, R. M. Eustice, and J. W. Grizzle, "Contact-aided invariant extended kalman filtering for robot state estimation," *The International Journal of Robotics Research*, vol. 39, no. 4, pp. 402–430, 2020.

[7] J. Sola, "Quaternion kinematics for the error-state kalman filter," *arXiv preprint arXiv:1711.02508*, 2017.

[8] M. Brossard, S. Bonnabel, and J.-P. Condomines, "Unscented kalman filtering on lie groups," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 2485–2491.

[9] D. He, W. Xu, and F. Zhang, "Symbolic representation and toolkit development of iterated error-state extended kalman filters on manifolds," *IEEE Transactions on Industrial Electronics*, pp. 1–10, 2023.

[10] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "Fast-lio2: Fast direct lidar-inertial odometry," *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2053–2073, 2022.

[11] A. Barrau and S. Bonnabel, "The invariant extended kalman filter as a stable observer," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1797–1812, 2017.

[12] A. Barrau, "Non-linear state error based extended kalman filters with applications to navigation," Ph.D. dissertation, Mines Paristech, 2015.

[13] B. Bell and F. Cathey, "The iterated kalman filter update as a gauss-newton method," *IEEE Transactions on Automatic Control*, vol. 38, no. 2, pp. 294–297, 1993.

[14] G. Sibley, G. S. Sukhatme, and L. H. Matthies, "The iterated sigma point kalman filter with applications to long range stereo." in *Robotics: Science and Systems*, vol. 8, no. 1. Philadelphia, Pennsylvania, USA, 2006, pp. 235–244.

[15] 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.

[16] 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, pp. 209–228, 2015.

[17] K. C. Wolfe and M. Mashner, "Bayesian fusion on lie groups," *Journal of Algebraic Statistics*, vol. 2, no. 1, 2011.

[18] B. C. Hall, *Lie Groups, Lie Algebras, and Representations*. Springer Cham, 2015.

[19] G. Bourmaud, R. Mégret, A. Giremus, and Y. Berthoumieu, "From intrinsic optimization to iterated extended kalman filtering on lie groups," *Journal of Mathematical Imaging and Vision*, vol. 55, no. 3, pp. 284–303, 2016.

[20] J. Havlík and O. Straka, "Performance evaluation of iterated extended kalman filter with variable step-length," *Journal of Physics: Conference Series*, vol. 659, no. 1, p. 012022, nov 2015.