Noname manuscript No.
(will be inserted by the editor)

From Intrinsic Optimization to Iterated Extended Kalman Filtering on Lie Groups

Guillaume Bourmaud · Rémi Mégret · Audrey Giremus · Yannick Berthoumieu

Received: date / Accepted: date

Abstract In this paper, we propose a new generic filter called Iterated Extended Kalman Filter on Lie Groups. It allows to perform parameter estimation when the state and the measurements evolve on matrix Lie groups. The contribution of this work is threefold: 1) the proposed filter generalizes the Euclidean Iterated Extended Kalman Filter to the case where both the state and the measurements evolve on Lie groups, 2) this novel filter bridges the gap between the minimization of intrinsic non linear least squares criteria and filtering on Lie groups, 3) in order to detect and remove outlier measurements, a statistical test on Lie groups is proposed.

In order to demonstrate the efficiency of the proposed generic filter, it is applied to the specific problem of relative motion averaging, both on synthetic and real data, for Lie groups $SE(3)$ (rigid body motions), $SL(3)$ (homographies) and $Sim(3)$ (3D similarities). Typical applications of these problems are camera network calibration, image mosaicing and partial 3D reconstruction merging problem. In each of these three applications, our approach significantly outperforms the state of the art algorithms.

Keywords Filtering · Optimization · Manifolds · Lie Groups · Matrix Transformations · Kalman · Gauss-Newton · Motion · Homography · Camera Pose · 3D Similarity

1 Introduction

During the last decade, parameter estimation through optimization on matrix manifolds has been extensively studied [1] and employed in a wide range of applications [46]. This is due to the fact that taking intrinsically into account the geometry of the manifold increases the rate of convergence of algorithms [54] and helps to avoid singularities.

When the parameters follow a dynamical system and/or when the measurements are acquired sequentially, it is important to be able to perform the estimation by filtering [35].

The link between optimization and filtering on Euclidean spaces, in the context of non linear least squares, has already been studied [5, 9]. It led to the Iterated Extended Kalman Filter (IEKF) that produces better results in practice than the Extended Kalman Filter (EKF).

Even if several works successfully extended Euclidean filtering algorithms to manifolds (see Table 1 in [15]), to the best of our knowledge, the link between optimization and filtering has not been established for parameter estimation on manifolds yet.

In this work, we focus on bridging the gap between the formulation of intrinsic non linear least squares criteria and Kalman filtering on matrix Lie groups [32]

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement 288199 - Dem@Care

G. Bourmaud
University of Bordeaux, IMS Laboratory CNRS UMR 5218,
E-mail: guillaume.bourmaud@ims-bordeaux.fr

R. Mégret
University of Bordeaux, IMS Laboratory CNRS UMR 5218,
Dept of Mathematical Sciences, University of Puerto-Rico at Mayaguez,
E-mail: remi.megret@ims-bordeaux.fr

A. Giremus
University of Bordeaux, IMS Laboratory CNRS UMR 5218,
E-mail: audrey.giremus@ims-bordeaux.fr

Y. Berthoumieu
University of Bordeaux, IMS Laboratory CNRS UMR 5218,
E-mail: yannick.berthoumieu@ims-bordeaux.fr

that form an important kind of manifolds. Typical examples include 3D rotation matrices $SO(3)$, unitary quaternions $SU(2)$, rigid-body motion $SE(3)$, 3D similarities $Sim(3)$, homographies $SL(3)$. More specifically, we are interested in generalizing the IEKF to the case of a state and measurements evolving on Lie groups.

1.1 Related Work

On the one hand, a large amount of works proposes to minimize intrinsic non linear least squares to estimate parameters evolving on a Lie group. Most of them apply a modified Gauss-Newton algorithm (GN) [10], or closely related algorithms such as Levenberg-Marquardt or Reweighted Non Linear Least Squares, to take into account the geometry of the Lie group. [62] was the first to propose a modified GN to estimate a 3D orientation. Simultaneous Localization and Mapping approaches (SLAM) [31, 36, 41, 42] are based on a GN like algorithm to estimate 3D points as well as camera poses from a video sequence and [44] uses a GN to estimate homographies. In the more specific context of estimating global transformations from relative transformation measurements, several works derived specific GN like algorithms: [20] and [30] tackle the synchronization of rotations problem, [55] considers the consistent pose registration problem, [60] covers the partial 3D reconstruction merging problem and [50] performs image mosaicking.

On the other hand, a large number of Kalman-like filters has been proposed to estimate a state evolving on a Lie group: [23, 48, 43, 52] deal with the estimation of a 3D orientation, [47] dynamically estimates homographies and [24] performs SLAM. In [12, 11] an extended Kalman filter on Lie groups is derived, which is dedicated to continuous systems possessing symmetries. In our previous papers [18, 15], we generalize the extended Kalman filter to Lie groups both for discrete and continuous prediction models. However, these approaches are not related to optimization on Lie groups.

To the best of our knowledge, only a few works, that are specifically devoted to SLAM, have related filtering to optimization on Lie groups. [61] proposes an information-filter like algorithm while [39, 38] derive efficient square root information like filters. However, none of these approaches deal with the case of measurements evolving on a Lie group.

In preliminary works ([17] and [16]), we propose an Iterated Extended Kalman Filter on Lie groups specifically dedicated to relative motion averaging. In this paper, we are interested in proposing a generic Iterated Extended Kalman Filter on Lie groups. Thus, in the next section, we recall the (Euclidean) IEKF formalism.

1.2 Reminder: The Iterated Extended Kalman Filter

The IEKF [5, 9] is a filter dedicated to non-linear systems. The objective of this filter is to recursively approximate the posterior distribution $p\left(x_k|z_1,\ldots,z_k\right)$ by a Gaussian distribution, where $x_k \in \mathbb{R}^p$ is the state we wish to estimate at time k and $z_l \in \mathbb{R}^q$ is a measurement available at time l. Then the state estimate is taken as the mean of this approximated posterior distribution. More specifically, the filter is composed of two steps, a prediction step and an update step.

1.2.1 Prediction

The prediction step consists in approximating the following distribution:

$$\begin{aligned} & p\left(x_k|z_1,\ldots,z_{k-1}\right) \\ &= \int p\left(x_k|x_{k-1}\right) p\left(x_{k-1}|z_1,\ldots,z_{k-1}\right) d_L x_{k-1} && (1) \\ &\approx \mathcal{N}_{\mathbb{R}^p}\left(x_k;\mu_{k|k-1},P_{k|k-1}\right) && (2) \end{aligned}$$

where $d_L x_{k-1}$ corresponds to the Lebesgue measure on $\mathbb{R}^p$, $p\left(x_{k-1}|z_1,\ldots,z_{k-1}\right) = \mathcal{N}_{\mathbb{R}^p}\left(x_{k-1};\mu_{k-1|k-1},P_{k-1|k-1}\right)$ and the state transition is defined as:

$$x_k = f\left(x_{k-1}\right) + n_k \qquad (3)$$

where $n_k \sim \mathcal{N}_{\mathbb{R}^p}\left(n_k;\mathbf{0}_{p\times 1},R_k\right)$ is a white Gaussian noise and $f : \mathbb{R}^p \to \mathbb{R}^p$ is a differentiable function. Classically, $\mu_{k|k-1}$ is obtained by propagating the previous mean $\mu_{k-1|k-1}$ through (3) without noise, while $P_{k|k-1}$ is computed by linearizing (3) in $x_{k-1} = \mu_{k-1|k-1}$ and propagating $P_{k-1|k-1}$ through this linearized version of the state transition model.

The same prediction step can be seen alternatively as the fitting of a Gaussian distribution to the integrand in (1), using a Gauss-Laplace approximation (see Section 3), followed by the marginalization of x_{k-1}. In this case, $\mu_{k-1|k-1}$ is propagated by minimizing the negative log-likelihood of the integrand in (1):

$$\left\{\hat{x}_k,\hat{x}_{k-1}\right\} = \underset{x_k\in\mathbb{R}^p,x_{k-1}\in\mathbb{R}^p}{\operatorname{argmin}} \begin{pmatrix} \left\|x_k - f\left(x_{k-1}\right)\right\|^2_{R_k} \\ + \left\|x_{k-1} - \mu_{k-1|k-1}\right\|^2_{P_{k-1|k-1}} \end{pmatrix} \qquad (4)$$

where $\left\|\cdot\right\|^2_{\Sigma} = \left(\cdot\right)^T \Sigma^{-1}\left(\cdot\right)$ is the squared Mahalanobis distance. It is trivial to show that $\hat{x}_{k-1} = \mu_{k-1|k-1}$ and $\hat{x}_k = f\left(\mu_{k-1|k-1}\right)$. The predicted mean $\mu_{k|k-1}$, which coincides with the mode in the Gaussian case, is taken as $\hat{x}_k$. Finally, $P_{k|k-1}$ is computed by applying the Gauss-Laplace covariance approximation equation (see Section 3), followed by a marginalization of x_{k-1}, which results in a closed form covariance prediction equation. This alternative view will help us in generalizing the IEKF prediction step to Lie groups.

1.2.2 Update

The update step consists in approximating the following distribution:

$$\begin{aligned} & p\left(x_k|z_1,\dots,z_k\right) \\ & \propto p\left(z_k|x_k\right)p\left(x_k|z_1,\dots,z_{k-1}\right) && (5) \\ & \approx \mathcal{N}_{\mathbb{R}^p}\left(x_k;\mu_{k|k},P_{k|k}\right) && (6) \end{aligned}$$

assuming $p\left(x_k|z_1,\dots,z_{k-1}\right) = \mathcal{N}_{\mathbb{R}^p}\left(x_k;\mu_{k|k-1},P_{k|k-1}\right)$ and a measurement model defined as:

$$z_k = h\left(x_k\right) + w_k \tag{7}$$

where $w_k \sim \mathcal{N}_{\mathbb{R}^q}\left(w_k;\mathbf{0}_{q\times 1},Q_k\right)$ is a white Gaussian noise, $h(\cdot) : \mathbb{R}^p \to \mathbb{R}^q$ is a differentiable function. w_k and the state model noise n_k are assumed to be independent.

Once again, this update step can be seen as the fitting of a Gaussian distribution to (5), using a Gauss-Laplace approximation (see Section 3).

In this case, the updated mean $\mu_{k|k} = \hat{x}_k$ is defined as the minimizer of the negative log-likelihood of (5):

$$\hat{x}_k = \underset{x_k\in\mathbb{R}^p}{\operatorname{argmin}}\left(\left\|z_k - h\left(x_k\right)\right\|^2_{Q_k} + \left\|x_k - \mu_{k|k-1}\right\|^2_{P_{k|k-1}}\right) \tag{8}$$

Employing a GN algorithm to solve (8) allows one to compute $\mu_{k|k}$, while $P_{k|k}$ can be obtained by applying the Gauss-Laplace covariance approximation equation (see Section 3). Fortunately, by exploiting the specific structure of the GN applied to (8), it is possible to derive the IEKF update equations that allow to compute both $\mu_{k|k}$ and $P_{k|k}$ very efficiently [5].

1.3 Contribution and Outline of the paper

In this paper, we propose a generic Iterated Extended Kalman Filter on Lie Groups (LG-IEKF) that extends the link between the minimization of non linear least squares criteria and the Iterated Extended Kalman Filter [5, 9] to the case where the state and the observations evolve on Lie groups.

For this purpose, we first present a fitting approach, called intrinsic Gauss-Laplace approximation, that allows to fit a concentrated Gaussian distribution on Lie groups to the probability density of a random variable evolving on a Lie group.

Since this fitting technique requires to find the minimizer of an intrinsic non linear least squares criterion, we present a generic intrinsic GN algorithm on Lie groups, called LG-GN, which has the advantage of taking intrinsically into account the geometry of the Lie group on which the parameters evolve.

Then, we show that a generalization of the IEKF, to the case where the state and the observations evolve on Lie groups, can be obtained by employing intrinsic Gauss-Laplace approximations both to derive the LG-IEKF prediction step and the LG-IEKF update step (which involves the LG-GN algorithm).

Finally, because of the Gaussian noise assumption, the LG-IEKF is sensitive to outlier measurements. Thus, a statistical test on Lie groups is derived to detect and remove them.

In order to demonstrate the efficiency of the proposed generic LG-IEKF, it is applied to the specific problem of relative motion averaging, both on synthetic and real data for Lie groups $SE(3)$ (rigid body motions), $SL(3)$ (homographies) and $Sim(3)$ (3D similarities).

The rest of the paper is organized as follows: Section 2 introduces the formalism of Lie groups and the concentrated Gaussian distribution on Lie groups. Section 3 describes the intrinsic Gauss-Laplace approximation while Section 4 deals with the intrinsic GN on Lie groups algorithm. The theory behind the proposed filter is described in Section 5. In Section 6, our formalism is applied to the relative motion averaging problem, and evaluated experimentally. Finally, the conclusion is provided in Section 7.

2 Preliminaries

2.1 Introduction to Lie groups

2.1.1 Definitions

In this section we give the definitions and basic properties of (matrix) Lie groups and Lie algebra. For a detailed description of these notions the reader is referred to [21]. A Lie group G is a group which has also the structure of a smooth manifold such that group composition and inversion are smooth operations. If G is a matrix Lie group, then $X \in G \subset \mathbb{R}^{n\times n}$ and its operations are matrix multiplication and inversion with the $n \times n$ identity matrix as identity element Id.

The matrix exponential $\exp_G$ and matrix logarithm $\log_G$ mappings establish a local diffeomorphism between an open neighborhood of $\mathbf{0}_{n\times n}$ in the tangent space at the identity $T_{Id}G$, called the *Lie Algebra* $\mathfrak{g}$, and an open neighborhood of Id in G. The Lie Algebra $\mathfrak{g} \subset \mathbb{R}^{n\times n}$ associated to a p-dimensional Lie group is a p-dimensional vector space defined by a basis consisting of real matrices E_i for $i = 1 \dots p$. Hence there is a linear isomorphism between $\mathfrak{g}$ and $\mathbb{R}^P$ that we denote as follows: $[\cdot]_G^\vee : \mathfrak{g} \to \mathbb{R}^P$ and $[\cdot]_G^\wedge : \mathbb{R}^P \to \mathfrak{g}$. For example let $\mathfrak{a} \in \mathfrak{g} \subset \mathbb{R}^{n\times n}$, then we have $[\mathfrak{a}]_G^\vee = a \in \mathbb{R}^P$. Thus we can define a basis $[E_i]_G^\vee = e_i$ such that $\{e_i\}_{i=1\dots p}$ is the natural basis of $\mathbb{R}^P$ and $\mathfrak{a} = \sum_{i=1}^{p} a_i E_i$ with $a = (a_1 \dots a_p)^T$. The previous notions are summarized in Fig. 1.

In order to lighten the notations, we define[1]: $\exp_G^\wedge(\cdot) = \exp_G\left([\cdot]_G^\wedge\right)$ and $\log_G^\vee(\cdot) = [\log_G(\cdot)]_G^\vee$.

For the sake of brevity, in the rest of the paper, when we mention the Lie algebra $\mathfrak{g}$ of a p-dimensional Lie group G, we implicitly refer to its isomorphic Euclidean space $\mathbb{R}^p$. Moreover, when we write $A = \exp_G^\wedge(a)$ we assume that $\log_G^\vee(A) = a$, i.e we work only on sets where $\exp_G^\wedge(\cdot)$ and $\log_G^\vee(\cdot)$ are bijective functions.

Let us introduce the linearized Baker-Campbell-Hausdorff formula which expresses the group product directly in $\mathbb{R}^p$:

$$\log_G^\vee\left(\exp_G^\wedge(a)\exp_G^\wedge(b)\right) = b + \varphi_G(b)\,a + O\left(\|a\|^2\right) \tag{9}$$

where

$$\varphi_G(b) = \sum_{n=0}^{\infty} \frac{B_n \mathrm{ad}_G(b)^n}{n!} = Id - \frac{1}{2}\mathrm{ad}_G(b) + \cdots \tag{10}$$

is the inverse of the left Jacobian[2] of G, the B_n are the Bernoulli numbers and

$$\mathrm{ad}_G(b)\,a = \left[[b]_G^\wedge [a]_G^\wedge - [a]_G^\wedge [b]_G^\wedge\right]_G^\vee \tag{11}$$

2.1.2 Product of Lie groups

Most of the time, we wish to estimate several parameters evolving on different Lie groups, at once. Since the product of Lie groups is a Lie group [56], the algorithms presented in this paper can be applied to "concatenations" of Lie groups. They can even be applied to Euclidean parameters since an Euclidean space $\mathbb{R}^p$ is a trivial Lie group by taking the matrix embedding:

$$x \in \mathbb{R}^p \mapsto X = \begin{bmatrix} Id & x \\ \mathbf{0}_{1\times p} & 1 \end{bmatrix} \subset \mathbb{R}^{(p+1)\times(p+1)} \tag{12}$$

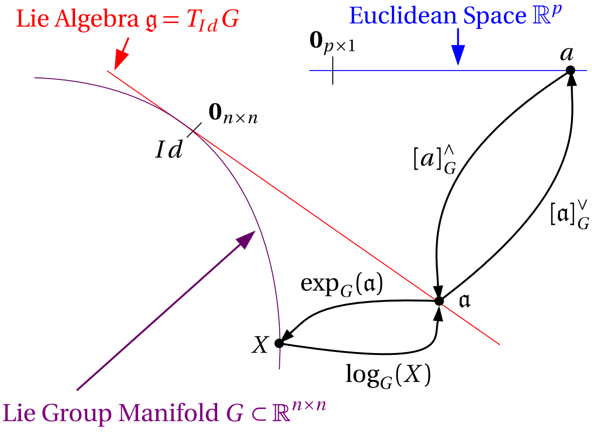


Fig. 1: Illustration of the geometry of a (matrix) Lie group

The simplest way to "concatenate" several components is to consider the Lie group formed by their direct product (e.g: $SO(3)\times\mathbb{R}^3$). Note that other ways to "concatenate" Lie groups exist [56] (e.g: semi-direct product, twisted product).

2.2 Additional notations

In the rest of the paper, G and G' are Lie groups of intrinsic dimensions p and q respectively. $d_H X$ is the right invariant Haar measure of G.

2.3 Concentrated Gaussian distribution on Lie groups

Here, we introduce the concept of concentrated Gaussian distribution on Lie groups, which was initially proposed in [59] and then studied in [66, 67], as a generalization of the normal distribution in Euclidean space. The distribution of $X \in G$ is called a right[3] concentrated Gaussian distribution on G of "mean" μ and "covariance" P, denoted $X \sim \mathcal{N}_G^R(X;\mu,P)$ (the superscript R stands for "right"), if:

$$X = \exp_G^\wedge(\epsilon)\,\mu \tag{13}$$

where $\epsilon \sim \mathcal{N}_{\mathbb{R}^p}(\epsilon;\mathbf{0}_{p\times 1},P)$ and $P \subset \mathbb{R}^{p\times p}$ is a symmetric positive-definite matrix. Note that (13) depends on the choice of the Lie algebra basis $\{E_i\}_{i=1\dots p}$.

When the maximum of the eigenvalues of P is sufficiently "small", the probability mass is concentrated

[1] For several Lie groups of interest, such as $SO(3)$, $SE(3)$, $Sim(3)$, analytical expressions of $\exp_G^\wedge(\cdot)$ and $\log_G^\vee(\cdot)$ exist [57]. However, for $SL(3)$ for example, matrix exponential and logarithm have to be computed.

[2] A closed form expression of $\varphi_G(b)$ was recently derived for $SE(3)$ in [4].

[3] In this paper, we consider quantities that are invariant to the right action of the Lie group on itself. Similar results could be obtained by considering the left action, leading to a left concentrated Gaussian distribution on G, which is the modelization used for instance in [18].

around μ and the probability density of X, w.r.t the right invariant Haar measure of G (denoted $d_H X$) [26], can be approximated as follows:

$$p(X) \approx \frac{1}{\sqrt{(2\pi)^p \det(P)}} e^{-\frac{1}{2}\left\|\log_G^\vee\left(X\mu^{-1}\right)\right\|_P^2} \tag{14}$$

Such a distribution allows us to describe the covariance in $\mathbb{R}^p$ and hence to use Euclidean tools while being invariant w.r.t the right action of the group on itself: $\exp_G^\wedge(\epsilon) = X\mu^{-1} = XX'(\mu X')^{-1}$ with $X' \in G$.

In Fig. 2, we provide an example of a concentrated Gaussian distribution on the Lie group

$$SE(2) = \left\{ X = \begin{bmatrix} R & t \\ \mathbf{0}_{1\times 2} & 1 \end{bmatrix} \middle| R \in SO(2), t = \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^2 \right\}$$

with Lie algebra

$$\mathfrak{se}(2) = \left\{ \mathfrak{x} = \begin{bmatrix} 0 & -\theta & x \\ \theta & 0 & y \\ 0 & 0 & 0 \end{bmatrix} \middle| \theta, x, y \in \mathbb{R} \right\}$$

Note that the "banana" shape [45] comes from the nonlinearity[4] of $\exp_{SE(2)}$ and the correlations between θ, x and y. Such a distribution is a better representation of the uncertainty of a robot negotiating a bend compared to a Euclidean Gaussian distribution.

3 Intrinsic Gauss-Laplace approximation

A classical way to tackle a Bayesian filtering problem is to fit, at each time instant, a parametric distribution to the posterior distribution of the state. In this section, we propose a fitting method which will be employed, in the rest of the paper, to fit concentrated Gaussian distributions on Lie groups.

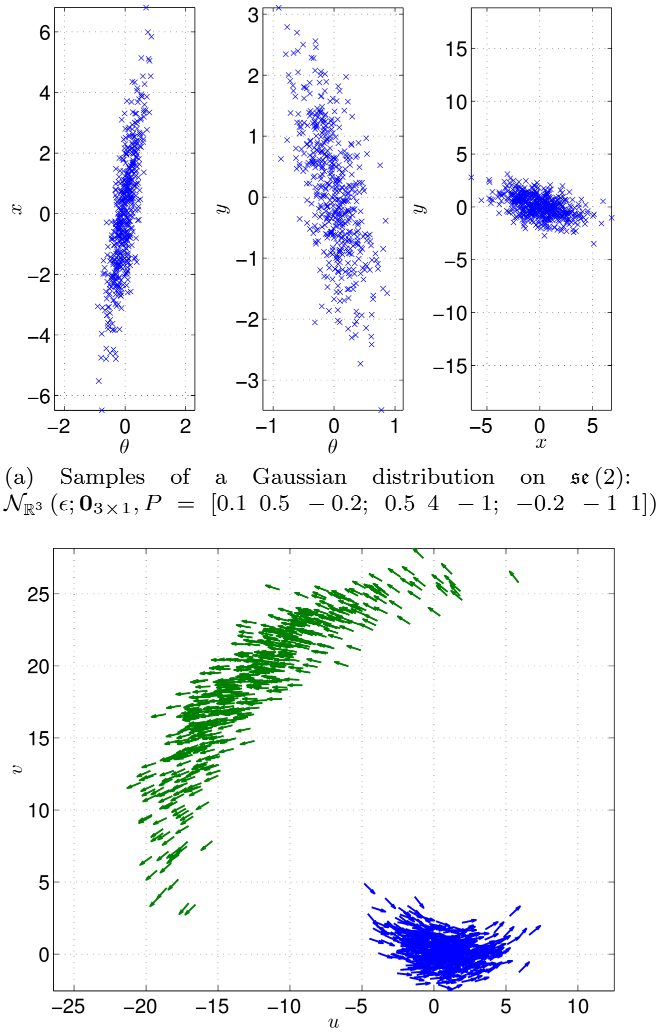
3.1 Problem

Let us consider the probability distribution of a random variable $X \in G$:

$$p(X) = \alpha e^{-\|\phi(X)\|_\Sigma^2} \tag{15}$$

where $\phi : G \to \mathbb{R}^m$ is a differentiable function and $p < m$.

The objective is to propose a method to fit a concentrated Gaussian distribution to $p(X)$.



(a) Samples of a Gaussian distribution on $\mathfrak{se}(2)$: $\mathcal{N}_{\mathbb{R}^3}(\epsilon; \mathbf{0}_{3\times 1}, P = [0.1\ \ 0.5\ \ -0.2;\ \ 0.5\ \ 4\ \ -1;\ \ -0.2\ \ -1\ \ 1])$

(b) Samples mapped to $SE(2)$ using $\exp_{SE(2)}^\wedge(\cdot)$ (blue arrows) and moved around the mean $\mu = \exp_{SE(2)}^\wedge\left([\pi\ \ 30\ \ 20]^T\right)$ using the right action of $SE(2)$ on itself (green arrows)

Fig. 2: Illustration of a concentrated Gaussian distribution on $SE(2)$. An arrow represents the position and the orientation of a robot in a 2D plane.

3.2 Proposed solution

First of all, let us define the minimizer of the cost function $\|\phi(X)\|_\Sigma^2$:

$$\hat{X} = \underset{X \in G}{\operatorname{argmin}} \|\phi(X)\|_\Sigma^2 \tag{16}$$

A first order Taylor expansion of $\phi(\cdot)$ around $\hat{X}$ gives:

$$\phi(X) = \phi\left(\exp_G^\wedge(\delta)\hat{X}\right) \approx \phi\left(\hat{X}\right) + J\delta \tag{17}$$

[4] Of course, this shape is emphasized by the action of μ on $\exp_G^\wedge(\epsilon)$.

where $\delta = \log_G^{\vee}\left(X\hat{X}^{-1}\right)$ and

$$J = \left.\frac{d\phi\left(\exp_G^{\wedge}(s)\,\hat{X}\right)}{ds}\right|_{s=\mathbf{0}} \tag{18}$$

We choose to take $q(X)$ as an approximation of $p(X)$:

$$q(X) = \beta e^{-\left\|\phi(\hat{X})+J\log_G^{\vee}(X\hat{X}^{-1})\right\|_{\Sigma}^{2}} \tag{19}$$

Finally, it is possible to show that $q(X)$ has the form of a concentrated Gaussian distribution:

$$q(X) = \mathcal{N}_G^R\left(X; \hat{X}, P = \left(J^T\Sigma^{-1}J\right)^{-1}\right) \tag{20}$$

We call this fitting method "intrinsic Gauss-Laplace approximation". This method is a generalization of the Euclidean Gauss-Laplace approximation which is used for instance in [5] to derive the IEKF update equations. In fact, if G is a Euclidean space, the method comes down to fitting a Gaussian distribution.

This method will be employed in Section 5 to derive both the prediction and update steps of the novel filter we present in this paper.

4 Intrinsic Gauss-Newton on Lie groups

The intrinsic Gauss-Laplace approximation introduced in the previous section assumes that we are capable of finding the minimizer $\hat{X}$ of (16).

However, when the minimizer does not have a closed-form expression, it is common to employ an iterative optimization algorithm.

In this section, we derive an intrinsic GN algorithm that allows one to estimate a matrix X of parameters evolving on G.

4.1 Introduction

When performing an iterative optimization on an Euclidean space, such as a gradient descent, the parameters $x \in \mathbb{R}^n$ are iteratively updated as follows:

$$x^{l+1} = x^l + \delta^{l+1/l} \tag{21}$$

where $\delta^{l+1/l} \in \mathbb{R}^n$ is an increment that corrects the previous parameter vector x^l to get x^{l+1}. In order to perform intrinsic optimization on a Lie group G, the update equation (21) cannot be applied since the operator "+" does not allow the parameters to remain on the manifold. The key ingredient to intrinsically take into account the geometry of the Lie group is to replace (21) with:

$$X^{l+1} = \exp_G^{\wedge}\left(\delta^{l+1/l}\right)X^l \tag{22}$$

where $X^l, X^{l+1} \in G$ and $\delta^{l+1/l} \in \mathbb{R}^p$. One can see that the update equation (22) is tightly related to the concept of concentrated Gaussian distribution on Lie groups (13).

4.2 Algorithm

A common iterative method for solving the problem (16), when G is a Euclidean space, is the Gauss-Newton method. In the following, we extend this formalism to Lie groups and call it intrinsic Gauss-Newton on Lie groups (LG-GN).

A Riemannian GN algorithm can be found in [1]. As a consequence, the algorithm we present can fit into their formalism (since a Lie group is a Riemannian manifold) and the convergence proof proposed in [1] applies. However, the LG-GN we introduce is dedicated to Lie groups (as the one already proposed in [63]) and leverages the "one parameter subgroup" [21] structure of the Lie group, i.e the Riemannian structure of the Lie group is not explicitly exploited.

As in the Euclidean case, the convergence of the algorithm depends on the starting point X^0. X^{l+1} is obtained by linearizing $\phi(\cdot)$ at the previous value X^l and solving the following problem:

$$\delta^{l+1/l} = \underset{\delta\in\mathbb{R}^p}{\operatorname{argmin}} \left\|\phi\left(X^l\right) - J_l\delta\right\|_{\Sigma}^{2} \tag{23}$$

where J_l is defined as:

$$J_l = -\left.\frac{d\phi\left(\exp_G^{\wedge}(s)\,X^l\right)}{ds}\right|_{s=\mathbf{0}} \tag{24}$$

Assuming that J_l has full column rank, the minimizer of (23) is:

$$\delta^{l+1/l} = \left(J_l{}^T\Sigma^{-1}J_l\right)^{-1}J_l{}^T\Sigma^{-1}\phi\left(X^l\right) \tag{25}$$

The current value of the parameters is finally updated as follows:

$$X^{l+1} = \exp_G^{\wedge}\left(\alpha^l\delta^{l+1/l}\right)X^l \tag{26}$$

where $0 < \alpha^l \leq 1$ is a step size.

At convergence, we take $\hat{X} = X^l$.

The approach described here will be employed in the next section to derive the update step of the novel filter we present in this paper.

5 Iterated Extended Kalman Filter on Lie Groups

From the concept of concentrated Gaussian distribution on Lie groups, the intrinsic Gauss-Laplace approximation and the LG-GN algorithm introduced in the previous sections, we derive a new filter called Iterated Extended Kalman Filter on Lie Groups (LG-IEKF).

5.1 Proposed System Model

In order to derive our novel filter, we first introduce a prediction model that describes the dynamical behavior of the state $X_k \in G$ we wish to estimate, where k stands for the time. Then we propose a measurement model that relates X_k to the measurement $Z_k \in G'$.

5.1.1 Prediction Model

Let the system state be modeled as satisfying the following equation:

$$X_k = \exp_G^\wedge(n_k)\, f(X_{k-1}) \tag{27}$$

where $X_k \in G$, $X_{k-1} \in G$ and $n_k \sim \mathcal{N}_{\mathbb{R}^p}(n_k; \mathbf{0}_{p\times 1}, R_k)$ is a white Gaussian noise. $f : G \to G$ is a differentiable function. The prediction model that we consider induces the following conditional distribution:

$$p(X_k|X_{k-1}) = \mathcal{N}_G^R(X_k; f(X_{k-1}), R_k) \tag{28}$$

This model is more generic than the one studied in [18] and thus allows us to deal with a larger class of problems (see section 6).
In order to simplify the notations, we have assumed w.l.o.g that the Lie groups G_k and G_{k-1}, on which X_k and X_{k-1} evolve respectively, are the same Lie group G. In practice, we will allow them to be different in order to augment the size of the state during prediction[5] (see Section 6.2.2).

5.1.2 Measurement Model

We consider discrete measurements Z_k on G' related to X_k as follows:

$$Z_k = \exp_{G'}^\wedge(w_k)\, h(X_k) \tag{29}$$

where $w_k \sim \mathcal{N}_{\mathbb{R}^q}(w_k; \mathbf{0}_{q\times 1}, Q_k)$ is a white Gaussian noise and $h : G \to G'$ is a differentiable function. Moreover, n_k and w_k are assumed to be independent.
The measurement model, that we consider, induces the following conditional distribution:

$$p(Z_k|X_k) = \mathcal{N}_{G'}^R(Z_k; h(X_k), Q_k) \tag{30}$$

5.2 Objective

We propose to approximate the state posterior distribution with a concentrated Gaussian distribution on Lie groups: $p(X_k|Z_1, \ldots, Z_l) \approx \mathcal{N}_G^R\left(X_k; \mu_{k|l}, P_{k|l}\right)$. We focus on $l = k-1$ (prediction) and $l = k$ (update). Therefore, the aim of the LG-IEKF is to predict and update the distribution parameters $\mu_{k|k}$ and $P_{k|k}$. In our formalism, $\mu_{k|k}$ is taken as the state estimate at time k.

5.3 LG-IEKF Prediction

We assume that the state posterior distribution at time $k-1$ is represented by:

$$p(X_{k-1}|Z_1, \ldots, Z_{k-1}) = \mathcal{N}_G^R\left(X_{k-1}; \mu_{k-1|k-1}, P_{k-1|k-1}\right) \tag{31}$$

The aim of this section is to show how to fit a concentrated Gaussian distribution to the posterior distribution of the predicted state:

$$\begin{aligned} & p(X_k|Z_1, \ldots, Z_{k-1}) \\ =& \int p(X_k|X_{k-1})\, p(X_{k-1}|Z_1, \ldots, Z_{k-1})\, d_H X_{k-1} \qquad (32) \\ \approx& \mathcal{N}_G^R\left(X_k; \mu_{k|k-1}, P_{k|k-1}\right) \qquad (33) \end{aligned}$$

In order to estimate $\mu_{k|k-1}$ and $P_{k|k-1}$, we propose to apply an intrinsic Gauss-Laplace approximation (see Section 3) to the integrand of (32) and then to marginalize X_{k-1}.

5.3.1 Mean Prediction

In order to predict the mean, we minimize the negative log-likelihood of (32):

$$\left\{\hat{X}_k, \hat{X}_{k-1}\right\} = \underset{X_k \in G, X_{k-1} \in G}{\operatorname{argmin}} \left(\begin{array}{c} \left\| \log_G^\vee\left(X_k f(X_{k-1})^{-1}\right)\right\|^2_{R_k} \\ + \left\| \log_G^\vee\left(X_{k-1}\mu^{-1}_{k-1|k-1}\right)\right\|^2_{P_{k-1|k-1}} \end{array} \right) \tag{34}$$

A trivial minimizer of this problem is $\hat{X}_{k-1} = \mu_{k-1|k-1}$ and $\hat{X}_k = f\left(\mu_{k-1|k-1}\right)$. We finally take:

$$\mu_{k|k-1} = \hat{X}_k = f\left(\mu_{k-1|k-1}\right) \tag{35}$$

[5] Augmenting the size of the state during the prediction step is sometimes called “smoothing” and not “filtering” in the literature.

5.3.2 Covariance Prediction

Concerning the covariance prediction, we apply the intrinsic Gauss-Laplace approximation formula (20).

By linearizing the error function inside the squared Mahalanobis norm in (34) around its minimizer, we obtain:

$$\begin{aligned}&-\log\left(p\left(X_k|X_{k-1}\right)p\left(X_{k-1}|Z_1,\ldots,Z_{k-1}\right)\right)\\&\approx\left\|\begin{bmatrix}Id & -F_k\\ \mathbf{0} & Id\end{bmatrix}\begin{bmatrix}\delta_k\\ \delta_{k-1}\end{bmatrix}\right\|^2_{\Sigma}\end{aligned}\tag{36}$$

where $X_{k-1} = \exp^\wedge_G\left(\delta_{k-1}\right)\mu_{k-1|k-1}$ and $X_k = \exp^\wedge_G\left(\delta_k\right)f\left(\mu_{k-1|k-1}\right)$. Moreover,

$$\Sigma = \begin{bmatrix}R_k & \\ & P_{k-1|k-1}\end{bmatrix}\tag{37}$$

and

$$F_k = -\left.\frac{d\log^\vee_G\left(f(\mu_{k-1|k-1})f(\exp^\wedge_G\left(s\right)\mu_{k-1|k-1})^{-1}\right)}{ds}\right|_{s=\mathbf{0}}\tag{38}$$

Thus, according to equation (20), the covariance P can be approximated by:

$$P = \left(\begin{bmatrix}Id & -F_k\\ \mathbf{0} & Id\end{bmatrix}^T\Sigma^{-1}\begin{bmatrix}Id & -F_k\\ \mathbf{0} & Id\end{bmatrix}\right)^{-1}\tag{39}$$

$$= \begin{bmatrix}R_k^{-1} & -R_k^{-1}F_k\\ -F_k^TR_k^{-1} & F_k^TR_k^{-1}F_k + P_{k-1|k-1}^{-1}\end{bmatrix}^{-1}\tag{40}$$

$$= \begin{bmatrix}F_kP_{k-1|k-1}F_k^T + R_k & F_kP_{k-1|k-1}\\ P_{k-1|k-1}F_k^T & P_{k-1|k-1}\end{bmatrix}\tag{41}$$

where we employed the blockwise matrix inversion [8]. Under the concentrated Gaussian approximation, the top left block of P in (41) corresponds to $P_{k|k-1}$. Thus, we obtain the following covariance prediction formula:

$$P_{k|k-1} = R_k + F_kP_{k-1|k-1}F_k^T\tag{42}$$

5.3.3 Prediction step summary

The prediction step consists in propagating the mean $\mu_{k-1|k-1}$ and the covariance $P_{k-1|k-1}$ by using the prediction model (27). At the end of the prediction step, the concentrated Gaussian approximation of the posterior is:

$$p\left(X_k|Z_1,\ldots,Z_{k-1}\right) \approx \mathcal{N}^R_G\left(X_k;\mu_{k|k-1},P_{k|k-1}\right)\tag{43}$$

5.4 LG-IEKF Update

The aim of this section is to demonstrate how to fit a concentrated Gaussian distribution to the posterior distribution of the state, after having received the measurement Z_k, using an intrinsic Gauss-Laplace approximation:

$$\begin{aligned}&p\left(X_k|Z_1,\ldots,Z_k\right)\\&\propto p\left(Z_k|X_k\right)p\left(X_k|Z_1,\ldots,Z_{k-1}\right) &(44)\\&\approx\mathcal{N}^R_G\left(X_k;\mu_{k|k},P_{k|k}\right) &(45)\end{aligned}$$

5.4.1 Mean Update

In order to update the mean, we minimize the negative log-likelihood of (44):

$$\hat{X}_k = \underset{X_k\in G}{\operatorname{argmin}}\begin{pmatrix}\left\|\log^\vee_{G'}\left(Z_kh\left(X_k\right)^{-1}\right)\right\|^2_{Q_k}\\ +\left\|\log^\vee_G\left(X_k\mu^{-1}_{k|k-1}\right)\right\|^2_{P_{k|k-1}}\end{pmatrix}\tag{46}$$

To minimize this function, we propose to apply the LG-GN algorithm described in Section 4. We introduce the following notations: $\delta^{l+1/l} = \log^\vee_G\left(X^{l+1}\left(X^l\right)^{-1}\right)$, $\delta^l = \log^\vee_G\left(X^l\mu^{-1}_{k|k-1}\right)$ and $\delta^{l+1} = \log^\vee_G\left(X^{l+1}\mu^{-1}_{k|k-1}\right)$ where X^l denotes the parameters value at iteration l of the LG-GN.

At iteration l, we seek the minimizer of the following problem:

$$\begin{aligned}\delta^{l+1/l} &= \underset{\delta\in\mathbb{R}^p}{\operatorname{argmin}}\begin{pmatrix}\left\|\log^\vee_{G'}\left(Z_kh\left(X^l\right)^{-1}\right) - H_l\delta\right\|^2_{Q_k}\\ +\left\|\delta^l + \varphi_G\left(\delta^l\right)\delta\right\|^2_{P_{k|k-1}}\end{pmatrix}\\ &= \underset{\delta\in\mathbb{R}^p}{\operatorname{argmin}}\left\|\psi\left(X^l\right) - \Psi_l\delta\right\|^2_{\Xi_k} &(47)\end{aligned}$$

where we used (9). The matrix $\varphi_G\left(\delta^l\right)$ is defined in (10),

$$H_l = -\left.\frac{d\log^\vee_{G'}\left(Z_kh\left(\exp^\wedge_G\left(s\right)X^l\right)^{-1}\right)}{ds}\right|_{s=\mathbf{0}}\tag{48}$$

$$\psi\left(X^l\right) = \left[\log^\vee_{G'}\left(Z_kh\left(X^l\right)^{-1}\right)^T\ \left(\delta^l\right)^T\right]^T\tag{49}$$

$$\Psi_l = \left[H_l^T\ -\varphi_G\left(\delta^l\right)^T\right]^T\tag{50}$$

and

$$\Xi_k = \begin{bmatrix}Q_k & \mathbf{0}\\ \mathbf{0} & P_{k|k-1}\end{bmatrix}\tag{51}$$

For the sake of brevity, the inverse of the left Jacobian of G is denoted $\varphi_G\left(\delta^l\right) \equiv \varphi_l$ and $\Phi_l = \varphi_l^{-1}$.

The minimizer of (47) is given by:

$$\begin{aligned}\delta^{l+1/l} &= \left(\Psi_l^T \Xi_k^{-1} \Psi_l\right)^{-1} \Psi_l^T \Xi_k^{-1} \psi\left(X^l\right) \\ &= \left(H_l^T Q_k^{-1} H_l + \varphi_l^T P_{k|k-1}^{-1} \varphi_l\right)^{-1} \left\{H_l^T Q_k^{-1}\right. \\ &\left.\log_{G'}^{\vee}\left(Z_k h\left(X^l\right)^{-1}\right) - \varphi_l^T P_{k|k-1}^{-1} \delta^l\right\} \end{aligned} \tag{52}$$

We now demonstrate (see Appendix A.1) that, neglecting second order terms in $\delta^{l+1/l}$, it is possible to obtain a generalization of the IEKF mean update equation [5] to Lie groups:

$$\delta^{l+1} = K_l \left\{\log_{G'}^{\vee}\left(Z_k h\left(X^l\right)^{-1}\right) + H_l \Phi_l \delta^l\right\} \tag{53}$$

where K_l is a gain that we call Lie-Kalman gain (see Appendix A.2) and is defined as:

$$K_l = P_{k|k-1} \Phi_l^T H_l^T \left(H_l \Phi_l P_{k|k-1} \Phi_l^T H_l^T + Q_k\right)^{-1} \tag{54}$$

Note that this Lie-Kalman gain is a generalization of the Kalman gain [5] to Lie groups. To the best of our knowledge, it is the first time that this expression is derived.

Moreover (53) can be simplified:

$$\Phi_l \delta^l = \Phi_G\left(\delta^l\right) \delta^l = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} \mathrm{ad}_G\left(\delta^l\right)^n \delta^l = \delta^l \tag{55}$$

since

$$\mathrm{ad}_G\left(\delta^l\right) \delta^l = \mathbf{0} \tag{56}$$

Finally, the current value of the parameters can be updated as follows:

$$\begin{aligned} X^{l+1} &= \exp_G^{\wedge}\left(\delta^{l+1/l}\right) X^l \qquad (57) \\ &\approx \exp_G^{\wedge}\left(K_l \left\{\log_{G'}^{\vee}\left(Z_k h\left(X^l\right)^{-1}\right) + H_l \delta^l\right\}\right) \mu_{k|k-1} \qquad (58) \end{aligned}$$

At convergence, we take $\mu_{k|k} = \hat{X}_k$.

5.4.2 Covariance Update

In order to estimate the updated covariance, we apply the intrinsic Gauss-Laplace approximation formula (20). To do so, we linearize the error function inside the squared Mahalanobis norm in (46) around its minimizer $\mu_{k|k}$:

$$\begin{aligned} &-\log\left(p\left(Z_k|X_k\right) p\left(X_k|Z_1, \ldots, Z_{k-1}\right)\right) \\ &\approx \left\| \begin{bmatrix} \log_{G'}^{\vee}\left(Z_k h\left(\mu_{k|k}\right)^{-1}\right) - H_l \delta \\ \delta^l + \varphi_l \delta \end{bmatrix} \right\|_{\Sigma}^2 \end{aligned} \tag{59}$$

where $\Sigma = \begin{bmatrix} Q_k & \mathbf{0} \\ \mathbf{0} & P_{k|k-1} \end{bmatrix}$, $X_k = \exp_G^{\wedge}(\delta) \mu_{k|k}$ and the subscript l corresponds to the last iteration of the LG-GN employed to update the mean. Thus, according to equation (20), the covariance $P_{k|k}$ can be approximated by:

$$P_{k|k} = \left(H_l^T Q_k^{-1} H_l + \varphi_l^T P_{k|k-1}^{-1} \varphi_l\right)^{-1} \tag{60}$$

However, it is possible to show that (see Appendix A.3):

$$P_{k|k} = \Phi_l \left(Id - K_l H_l \Phi_l\right) P_{k|k-1} \Phi_l^T \tag{61}$$

which is a generalization of the IEKF covariance update equation [5] to Lie groups.

5.4.3 Update step summary

The update step consists in updating the mean $\mu_{k|k-1}$ and the covariance $P_{k|k-1}$ by incorporating the information coming from the measurement Z_k, by using the measurement model (29). At the end of the update step, the concentrated Gaussian approximation of the posterior is:

$$p\left(X_k|Z_1, \ldots, Z_k\right) \approx \mathcal{N}_G^R\left(X_k; \mu_{k|k}, P_{k|k}\right) \tag{62}$$

5.5 Inlier test before update

The novel filter we propose is based on the classical assumption that the measurement noise w_k (see Section 5.1) is a white Gaussian noise. This assumption is convenient and allowed us to derive an efficient algorithm. However, it also makes the LG-IEKF not resilient to outlier measurements (this is also a limitation of Kalman filtering on Euclidean spaces [53]). Consequently, we propose a statistical test on Lie groups to detect and remove outliers.

A measurement Z_k is an inlier if and only if:

$$\left\|\log_{G'}^{\vee}\left(Z_k h\left(X_k\right)^{-1}\right)\right\|_{Q_k}^2 < t \tag{63}$$

where t is a threshold to define. However, the true value X_k is unknown. We only have an approximation of the posterior distribution of the state:

$$p\left(X_k|Z_1, \ldots, Z_{k-1}\right) \approx \mathcal{N}_G^R\left(X_k; \mu_{k|k-1}, P_{k|k-1}\right) \tag{64}$$

which can also be expressed as follows:

$$X_k = \exp_G^{\wedge}\left(\epsilon_{k|k-1}\right) \mu_{k|k-1} \tag{65}$$

where $\epsilon_{k|k-1} \sim \mathcal{N}_{\mathbb{R}^p}\left(\epsilon; \mathbf{0}_{p\times 1}, P_{k|k-1}\right)$. Thus, we propose a statistical inlier test on Lie groups w.r.t the current estimate of X_k:

$$\left\|\log_{G'}^{\vee}\left(Z_k h\left(\mu_{k|k-1}\right)^{-1}\right)\right\|_{Q_{err}}^2 < t \tag{66}$$

where

$$Q_{err} = \mathbb{E}\left(\left(H_{k|k-1}\epsilon_{k|k-1} + w_k\right)\left(H_{k|k-1}\epsilon_{k|k-1} + w_k\right)^T\right)$$
$$= H_{k|k-1}P_{k|k-1}H_{k|k-1}^T + Q_k \tag{67}$$

Indeed, from (29) and neglecting second order terms in w_k and $\epsilon_{k|k-1}$, we have:

$$\begin{aligned}\mathbf{0} &= \log_{G'}^{\vee}\left(Z_k h\left(X_k\right)^{-1}\exp_{G'}^{\wedge}\left(-w_k\right)\right)\\ &\simeq \log_{G'}^{\vee}\left(Z_k h\left(\exp_G^{\wedge}\left(\epsilon_{k|k-1}\right)\mu_{k|k-1}\right)^{-1}\exp_{G'}^{\wedge}\left(-w_k\right)\right)\\ &\simeq \log_{G'}^{\vee}\left(Z_k h\left(\mu_{k|k-1}\right)^{-1}\right) - H_{k|k-1}\epsilon_{k|k-1} - w_k \end{aligned} \tag{68}$$

where

$$H_{k|k-1} = -\left.\frac{d\log_{G'}^{\vee}\left(Z_k h\left(\exp_G^{\wedge}\left(s\right)\mu_{k|k-1}\right)^{-1}\right)}{ds}\right|_{s=\mathbf{0}} \tag{69}$$

Thus, under the concentrated Gaussian assumption, the LHS of (66) is distributed according to the chi-squared distribution with q degrees of freedom, since G' is a q-dimensional Lie group. Consequently, one way to decide whether Z_k is an inlier is to define a threshold based on the p-value of $\chi^2\left(q\right)$ [27]. Note that since we neglected second order terms, this theoretical threshold is possibly restrictive.

5.6 Summary of the LG-IEKF and remarks

The LG-IEKF algorithm is summarized in Alg.1.

*Remark 1*Employing the LG-GN allowed us to iteratively refine the linearization point during the update step contrary to the Extended Kalman Filter on Lie Groups (LG-EKF), proposed in [18], that performs only one linearization. In this sense, the LG-EKF is a special case of the LG-IEKF proposed in this paper. Note that the prediction model we use in this paper is more generic than the one employed in [18].

*Remark 2*The models (27) and (29) as well as the algorithm Alg.1 reduce to the traditional models and algorithm of the IEKF in the case of a state and measurements evolving on Euclidean spaces. In this sense, the LG-IEKF generalizes the IEKF to Lie groups.

Algorithm 1 LG-IEKF Algorithm

```
Inputs : $\mu_{k-1|k-1}$, $P_{k-1|k-1}$, $Z_k$, $Q_k$, $R_k$, $t$ (optional)
Outputs : $\mu_{k|k}$, $P_{k|k}$

1) Prediction
$\mu_{k|k-1} = f\left(\mu_{k-1|k-1}\right)$
$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + R_k$
where $F_k = -\left.\frac{d\log_G^{\vee}\left(f\left(\mu_{k-1|k-1}\right)f\left(\exp_G^{\wedge}(s)\mu_{k-1|k-1}\right)^{-1}\right)}{ds}\right|_{s=\mathbf{0}}$

2) Inlier Test (optional)
$\left\|\log_{G'}^{\vee}\left(Z_k h\left(\mu_{k|k-1}\right)^{-1}\right)\right\|_{Q_{err}}^2 < t$
where $Q_{err} = H_{k|k-1}P_{k|k-1}H_{k|k-1}^T + Q_k$
and $H_{k|k-1} = -\left.\frac{d\log_{G'}^{\vee}\left(Z_k h\left(\exp_G^{\wedge}(s)\mu_{k|k-1}\right)^{-1}\right)}{ds}\right|_{s=\mathbf{0}}$

3) Update (if 2) is satisfied)
Set $X^0 = \mu_{k|k-1}$ and $\delta^0 = \mathbf{0}_{p\times 1}$
Iterate until convergence:
    $K_l = P_{k|k-1}\Phi_l^T H_l^T\left(H_l\Phi_l P_{k|k-1}\Phi_l^T H_l^T + Q_k\right)^{-1}$
    $\delta^{l+1} = K_l\left\{\log_{G'}^{\vee}\left(Z_k h\left(X^l\right)^{-1}\right) + H_l\delta^l\right\}$
    $X^{l+1} = \exp_G^{\wedge}\left(\delta^{l+1}\right)\mu^{k|k-1}$
    where $H_l = -\left.\frac{d\log_{G'}^{\vee}\left(Z_k h\left(\exp_G^{\wedge}(s)X^l\right)^{-1}\right)}{ds}\right|_{s=\mathbf{0}}$
    $\varphi_l \equiv \varphi_G\left(\delta^l\right)$ is the inverse of the left Jacobian of $G$ and
    $\Phi_l = \varphi_l^{-1}$.
At convergence:
    $\mu_{k|k} = X^l$
    $P_{k|k} = \Phi_l\left(Id - K_l H_l \Phi_l\right)P_{k|k-1}\Phi_l^T$

$F_k$, $H_{k|k-1}$ and $H_l$ are Jacobian matrices which
expressions depend on the application.
```

6 Application To Relative Motion Averaging

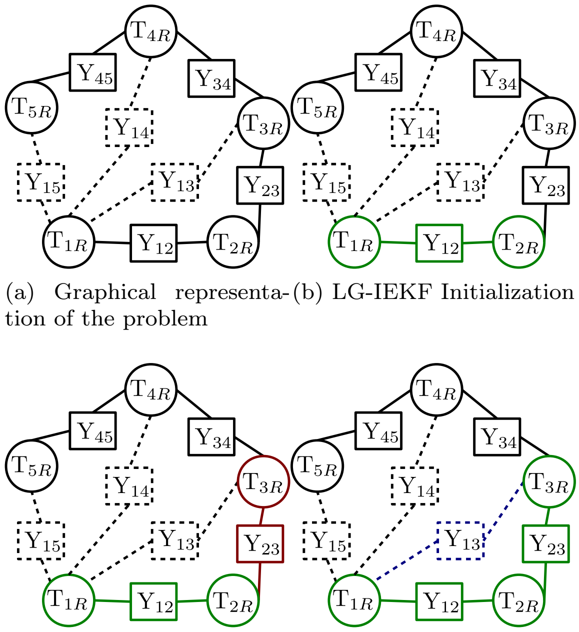
Here we apply the generic LG-IEKF algorithm to the problem of estimating global transformations from noisy relative transformation measurements[6], a.k.a relative motion averaging.

Such a problem occurs for instance in the context of consistent pose registration [2] encountered in 3D localization, structure from motion or camera network calibration. In this case, a motion or transformation is a rigid body transformation matrix ($SE(3)$). Thus, the

[6] The Matlab code is available at https://sites.google.com/site/guillaumebourmaud/

relative measurements correspond to the rigid transformations between two camera poses and the global motions we wish to estimate are the rigid transformation matrices between a reference camera pose and all the other camera poses.

In our context, a transformation is defined as an element of a Lie group. Consequently, the solution we propose can be applied to several other problems such as the synchronization of rotations problem ($SO(3)$) [13], image mosaicing ($SL(3)$) [50] or the partial 3D reconstruction merging problem ($Sim(3)$) [64].



(a) Graphical representation of the problem (b) LG-IEKF Initialization

(c) LG-IEKF First Prediction (d) LG-IEKF First Update

Fig. 3: Illustration of the relative motion averaging problem: dashed observations (Y_{ij} with $j > i+1$) are used as measurements during the LG-IEKF update step. Temporally consecutive observations ($Y_{i\,(i+1)}$ for $i = 1...N-1$) are used as control inputs during the LG-IEKF prediction step in order to guide the estimation process.

6.1 Objective

We aim at estimating global transformations $\{T_{iR}\}_{i=1:N}$, where each global transformation $T_{iR} \in G'$ is defined as the transformation between a main reference frame (RF) R and a RF i, and G' is a q-dimensional Lie group.

We consider the case where the noises on the relative transformation measurements $\{Y_{ij}\}_{1\leq i<j\leq N}$ are mutually independent. Each $Y_{ij} \in G'$ denotes a noisy relative transformation between a RF j and a RF i expressed as follows:

$$Y_{ij} = \exp^{\wedge}_{G'}\left(b^i_{ij}\right) T_{iR} T_{jR}^{-1} \tag{70}$$

where $b^i_{ij} \sim \mathcal{N}_{\mathbb{R}^q}\left(b^i_{ij}; \mathbf{0}_{q\times 1}, \Sigma^i_{ij}\right)$ is a white Gaussian noise. The problem is illustrated in Fig. 3a. Note that (70) is invariant w.r.t the right action of G'. Indeed, $T_{iR} M \left(T_{jR} M\right)^{-1} = T_{iR} T_{jR}^{-1}$ for any $M \in G'$. In the context of consistent pose registration, it simply means that rotating and translating all the camera poses does not affect the relative measurements.

6.2 Case 1: Outlier Free Estimation

In this section, we consider the case where the measurements are outlier free.

6.2.1 State of the Art

A large amount of works addressing the problem tackled in this section has been previously published. However, they usually do not intrinsically take into account the Lie group structure and are tailored to specific applications such as:
-Relative orientation averaging (Lie group $SO(3)$, or $SU(2)$ for unitary quaternion), a.k.a multiple rotation averaging [34], a.k.a synchronization of rotations [14]. In [29] and [49], in order to obtain a closed-form solution that only requires solving a large linear system of equations, the Lie group geometry of $SU(2)$ and $SO(3)$ respectively are overlooked. [28] formulates the problem as [29], but the Lie group constraints of $SU(2)$ are extrinsically taken into account using Lagrangian duality.
-Relative Euclidean motion (Lie group $SE(3)$) averaging, a.k.a camera pose registration problem. In [2], a Euclidean motion is parametrized as an element of the Lie algebra $\mathfrak{se}(3)$ of $SE(3)$, which is a vector space, in order to formulate the problem into a classical nonlinear unconstrained Euclidean minimization problem. However, this parametrization causes problems at the boundary of the Lie algebra. [30] proposes an iterative algorithm that intrinsically takes into account the Lie group structure of $SE(3)$. Nevertheless, the minimized criterion is not invariant (w.r.t the left and right action of the Lie group on itself) while the measurement model is invariant to the action of $SE(3)$.
-Relative homography (Lie group $SL(3)$) averaging, a.k.a image mosaicing. In [19], an Euclidean Extended Kalman Filter is derived for which the Lie group constraints are extrinsically taken into account using a reprojection after each iteration. [50] derives an iterative algorithm based on a matrix exponential update rule. However

they do not completely take advantage of the Lie group structure of $SL(3)$. In order to obtain a closed-form solution to the problem, [40] performs a relaxation where the Lie group constraints of $SL(3)$ are not considered.

6.2.2 Implementation of the LG-IEKF

In order to apply the LG-IEKF to the relative motion averaging problem, we need to specify the prediction and update models.

However, before detailing these models, let us describe the way the observations $\{Y_{ij}\}_{1\leq i<j\leq N}$ are used.

We distinguish the temporally consecutive observations ($Y_{i\,(i+1)}$ for $i=1...N-1$) from the other observations (Y_{ij} with $j>i+1$). Indeed, in order to guide the estimation process, the temporally consecutive observations are used as control inputs in the prediction steps (see Section 6.2.3). The other observations are used as measurements in the update steps (see Section 6.2.4). A graphical example of the prediction and update steps is proposed in Fig.3.

6.2.3 Prediction model

As explained in Section 5.1.1, during the LG-IEKF derivation, we have assumed w.l.o.g that the Lie groups G_k and G_{k-1}, on which X_k and X_{k-1} evolve respectively, are the same Lie group G. For this application, we allow them to be different in order to augment the size of the state during prediction. Thus, at time instant $k-1$, the state X_{k-1} contains the global transformations $T_{1R}, T_{2R}, ..., T_{(k-1)R}, T_{kR}$, while at time k, the state X_k contains the global transformations $T_{1R}, T_{2R}, ..., T_{(k-1)R}, T_{kR}, T_{(k+1)R}$.

Consequently, $G_k = G' \times G' \times \cdots \times G'$, i.e $k+1$ direct products of G'.

More precisely, the prediction step that we propose augments the size of the state by duplicating the global motion T_{kR} and propagating it using the control input $Y_{(k+1)\,k} = Y^{-1}_{k\,(k+1)}$ in order to predict $T_{(k+1)R}$. Thus we consider the following prediction model:

$$X_k = \begin{bmatrix} \exp^{\wedge}_{G'}\left(b^{k+1}_{(k+1)k}\right) Y_{(k+1)\,k}\left(X_{k-1}\right)_{kR} & \mathbf{0} \\ \mathbf{0} & X_{k-1} \end{bmatrix} \tag{71}$$

$$= \exp^{\wedge}_{G_k}\left(\begin{bmatrix} b^{k+1}_{(k+1)k} \\ \mathbf{0} \end{bmatrix}\right) \begin{bmatrix} Y_{(k+1)\,k} & \mathbf{0} \\ \mathbf{0} & Id \end{bmatrix} \begin{bmatrix} \left(X_{k-1}\right)_{kR} & \mathbf{0} \\ \mathbf{0} & X_{k-1} \end{bmatrix} \tag{72}$$

where the notation $\left(X_{k-1}\right)_{kR}$ corresponds to extracting the global motion component T_{kR} from the state X_{k-1}. $b^{k+1}_{(k+1)k} \sim \mathcal{N}_{\mathbb{R}^q}\left(b^{k+1}_{(k+1)k}; \mathbf{0}_{q\times 1}, \Sigma^{k+1}_{(k+1)k}\right)$ is a white Gaussian noise.

As can be seen, eq.(72) has the form of (27) where $f\left(X_{k-1}\right) = \begin{bmatrix} Y_{(k+1)\,k} & \mathbf{0} \\ \mathbf{0} & Id \end{bmatrix} \begin{bmatrix} \left(X_{k-1}\right)_{kR} & \mathbf{0} \\ \mathbf{0} & X_{k-1} \end{bmatrix}$ and $n_k = \begin{bmatrix} b^{k+1}_{(k+1)k} \\ \mathbf{0} \end{bmatrix}$. As a consequence, the LG-IEKF prediction equations can be applied.

6.2.4 Update model

The measurement model we consider corresponds to.

$$Z_k = Y_{i\,(k+1)} = \exp^{\wedge}_{G'}\left(b^i_{i(k+1)}\right)\left(X_k\right)_{iR}\left(X_k\right)^{-1}_{(k+1)R} \tag{73}$$

where $i<k$ and $b^i_{i(k+1)} \sim \mathcal{N}_{\mathbb{R}^q}\left(b^i_{i(k+1)}; \mathbf{0}_{q\times 1}, \Sigma^i_{i(k+1)}\right)$ is a white Gaussian noise.

As can be seen, eq.(73) has the form of (29), where $w_k = b^i_{i(k+1)}$ and $h\left(X_k\right) = \left(X_k\right)_{iR}\left(X_k\right)^{-1}_{(k+1)R}$. As a consequence, the LG-IEKF update equations can be applied.

Note that if several measurements $Y_{i\,(k+1)}$, for all $i<k$, are available, then we concatenate them in order to have a single measurement equation.

6.2.5 Application of the LG-IEKF prediction step

In order to apply the LG-IEKF prediction step to (72), we need to compute F_k (see eq.(38)). We obtain (see Appendix B.1):

$$F_k = \begin{bmatrix} \mathrm{Ad}_{G'}\left(Y_{(k+1)\,k}\right) & \mathbf{0} \\ \mathbf{0} & Id \end{bmatrix} \begin{bmatrix} \mathbf{0} & Id & \mathbf{0} \\ & Id & \end{bmatrix} \tag{74}$$

where we introduced the adjoint representation $\mathrm{Ad}_{G'}(\cdot) \subset \mathbb{R}^{q\times q}$ of G' on $\mathbb{R}^q$ that enables us to transform an increment $\epsilon^i_{ij} \in \mathbb{R}^q$, that acts onto an element Y_{ij} through left multiplication, into an increment $\epsilon^j_{ij} \in \mathbb{R}^q$, that acts through right multiplication:

$$\exp^{\wedge}_{G'}\left(\epsilon^i_{ij}\right) Y_{ij} = Y_{ij}\exp^{\wedge}_{G'}\left(\mathrm{Ad}_{G'}\left(Y^{-1}_{ij}\right)\epsilon^i_{ij}\right) \tag{75}$$

6.2.6 Application of the LG-IEKF update step

In order to apply the LG-IEKF update step to (73), we need to compute H_l (see eq.(48)). We obtain (see Appendix B.2):

$$H_l = \begin{bmatrix} \mathbf{0} \; Id \; \mathbf{0} \; -\mathrm{Ad}_{G'}\left(\left(X^l\right)_{iR}\left(X^l\right)^{-1}_{(k+1)R}\right) \; \mathbf{0} \end{bmatrix} \tag{76}$$

In our implementation, φ_l and Φ_l are approximated by the identity matrix. Also, since we consider the case where the measurements are outlier free, the inlier test (Step 2 in Alg.1) is skipped.

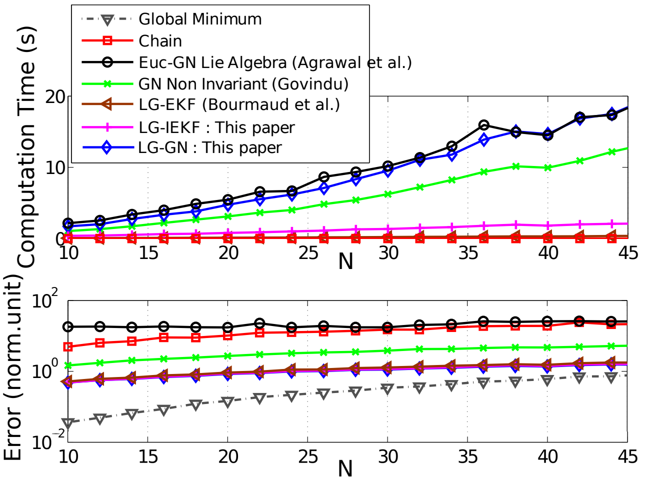


Fig. 4: Computational time and Residual errors for different number of global motions (N). The number of relative motions is fixed to $N+n$ (in our experiments $n=30$).

6.2.7 Results on Simulated Data for the Camera Pose Registration Problem

We compare the proposed approach to state of the art algorithms on a camera pose registration problem, i.e global Euclidean motion (Lie group $SE(3)$) estimation from relative measurements.

The explicit definitions of $\exp^{\wedge}_{SE(3)}$, $\log^{\vee}_{SE(3)}$ and $\mathrm{Ad}_{SE(3)}$ can be found in [57].

The data are simulated using the generative model proposed in (70).

We consider 6 different algorithms:

-"Chain": it composes the temporally consecutive relative measurements to obtain the global motions

-"Euc-GN Lie Algebra": algorithm proposed in [2] (Agrawal et al.)

-"GN Non Invariant": algorithm proposed in [30] (Govindu)

-"LG-GN": LG-GN derived in Section 4 minimizing

$$\sum_{i,j}\left\|\log^{\vee}_{SE(3)}\left(Y_{ij}T_{jR}T^{-1}_{iR}\right)\right\|^2_{\Sigma^i_{ij}} \qquad (77)$$

-"LG-IEKF": LG-IEKF derived in Section 5

-"LG-EKF": modified version of the algorithm proposed in [18] (Bourmaud et al.) which actually corresponds to applying the LG-IEKF with only one iteration at each update step

In order to fairly compare the algorithms [30], [2] and the LG-GN to the LG-IEKF, they are applied incrementally since it guides the estimation process and reduces the chance of falling in a poor local minimum. All the algorithms are coded in Matlab and tested with the following configuration: core I5 4x2.27GHz, 4GB, Linux 64 bits.

We simulate circular camera trajectories (see Fig.6) with N cameras. In order to compare the results of each approach to the true global motions, we need to add a step to align the estimated global motions with the true global motions. For that purpose, we choose to employ another LG-GN to minimize the sum of the following error function: $\left\|\log^{\vee}_{SE(3)}\left(\mu_{iR}T_{RR_{True}}T^{-1}_{iR_{True}}\right)\right\|^2$. The final error of this LG-GN obtained for each approach as well as the computational time are presented in Fig.4.

Results show that the LG-IEKF, which considers invariant errors, performs significantly better than the state of the art algorithms [2] and [30].

Indeed, the parametrization used in [2] causes issues at the boundary of the Lie algebra, while [30] does not consider an invariant error function and thus does not fully exploit the geometry of the problem. For these reasons, [2] and [30] frequently fall into poor local minima.

As expected, the LG-GN, which is a batch optimization, performs slightly better than the LG-IEKF. However, the LG-IEKF has a much lower computational time. Finally, one can see that the LG-IEKF, that iteratively refines its linearization point, outperforms the LG-EKF.

6.3 Case 2: Estimation in the presence of outliers

In this section, we consider the case where the measurements are corrupted with outliers and compare the performances of the proposed LG-IEKF against the state of the art algorithms able to deal with outliers. As we will see in the experimentations on real data, the outlier measurements are usually due to duplicated structures in the environment [55].

6.3.1 State of the Art

A large amount of works has been recently devoted to specifically dealing with multiple rotation averaging in the presence of outliers. This problem is also known as synchronization of rotations in the mathematics community and is usually tackled by minimizing a given criterion. In [3] and [58], spectral relaxations of the problem are proposed while [13] uses their results as initialization for a second order Riemannian trust-region algorithm to compute a local maximizer. [65] derives

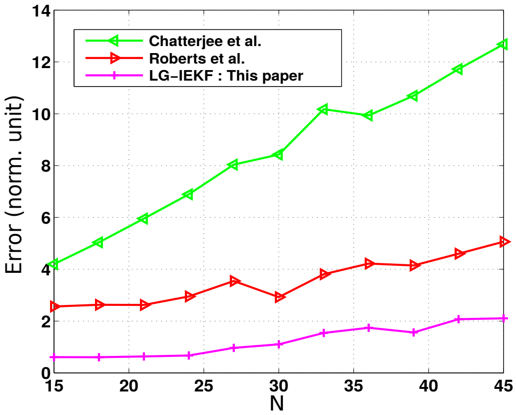


Fig. 5: Estimation error of the global motions for our approach, Chatterjee et al. [20] and Roberts et al. [55] on a camera pose estimation problem ($\lambda = \frac{1}{10}$ and number of relative motions fixed to $5N + n$ with $n = 60$).

an algorithm that exactly estimates the global rotations when a subset of the measurements are perfect and outperforms [58]. In [33] and [20], two robust iterative algorithms, based on L1 and L1-L2 minimization criterion, respectively, are devised. However, the considered error functions are not convex and consequently need a good initialization such as [65] to avoid poor local minima. Finally, [22] proposes a discretization of $SO(3)$ to apply a loopy belief propagation algorithm on the resulting Markov random field.

The works [25], [37] and [51] are also relevant for the multiple rotation averaging problem. However, they are specifically tailored for $SO(3)$ and it is not straightforward to apply them to other Lie groups.

To the best of our knowledge, only one approach [55] was proposed to deal with the generic problem of global motion estimation from relative measurements in the presence of outliers. An Expectation Maximization (EM) algorithm is proposed, introducing latent variables to classify the measurements as inliers or outliers.

6.3.2 Implementation of the LG-IEKF

In order to apply the LG-IEKF, we assume that the temporally consecutive observations ($Y_{i\,(i+1)}$ for $i = 1...N-1$) are not corrupted with outliers. This assumption might appear restrictive, however, for image sequences for example, it is usually satisfied (see Sections 6.3.4 and 6.3.5). Indeed, relative transformations computed between consecutive images usually do not produce outlier measurements.

The LG-IEKF we apply is the same than the one derived in Section 6.2.2. Nevertheless, this time the inlier test (Step 2 in Alg.1) is not skipped because of the presence of outlier measurements.

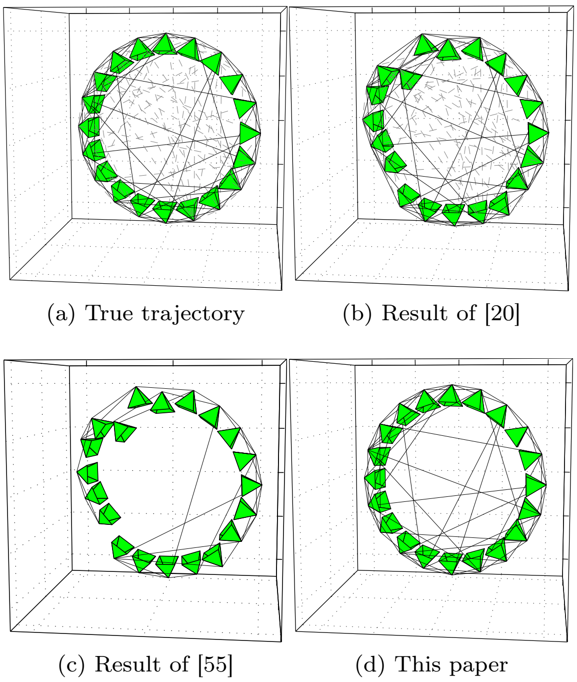

(a) True trajectory
(b) Result of [20]
(c) Result of [55]
(d) This paper

Fig. 6: Camera pose registration problem results, a cone represents a camera pose, a black line is an inlier measurement and a gray dashed line is an outlier

6.3.3 Results on Simulated Data for the Camera Pose Registration Problem

We compare the performance of the LG-IEKF to two state of the art algorithms [55] and [20] on a camera pose registration problem (Lie group $SE(3)$) with outliers. [20] was developed to deal with $SO(3)$ but its extension to $SE(3)$ is straightforward. We simulate circular camera trajectories (see Fig.6) with N cameras where each camera $T_{iR_{True}}$ has a timestamp t_i and we generate noisy relative motions as follows: first of all, a measurement can be either an inlier or an outlier (except temporally consecutive relative measurements that are always inliers). We model the probability of a measurement to be an inlier as $P\left(Z_{ij} \text{ is inlier}\right) = \exp\left(-\lambda \left|t_i - t_j\right|\right)$ where λ is a user-chosen parameter, i.e a larger time difference increases the chance to produce an outlier. After having drawn the label of a measurement (inlier or outlier), we sample the measurement. The distribution of the independent outliers is modeled as a centered Gaussian distribution on Lie groups with a large covariance matrix (the large covariance is not a problem in this case since $\log^{\vee}_{SE(3)}$

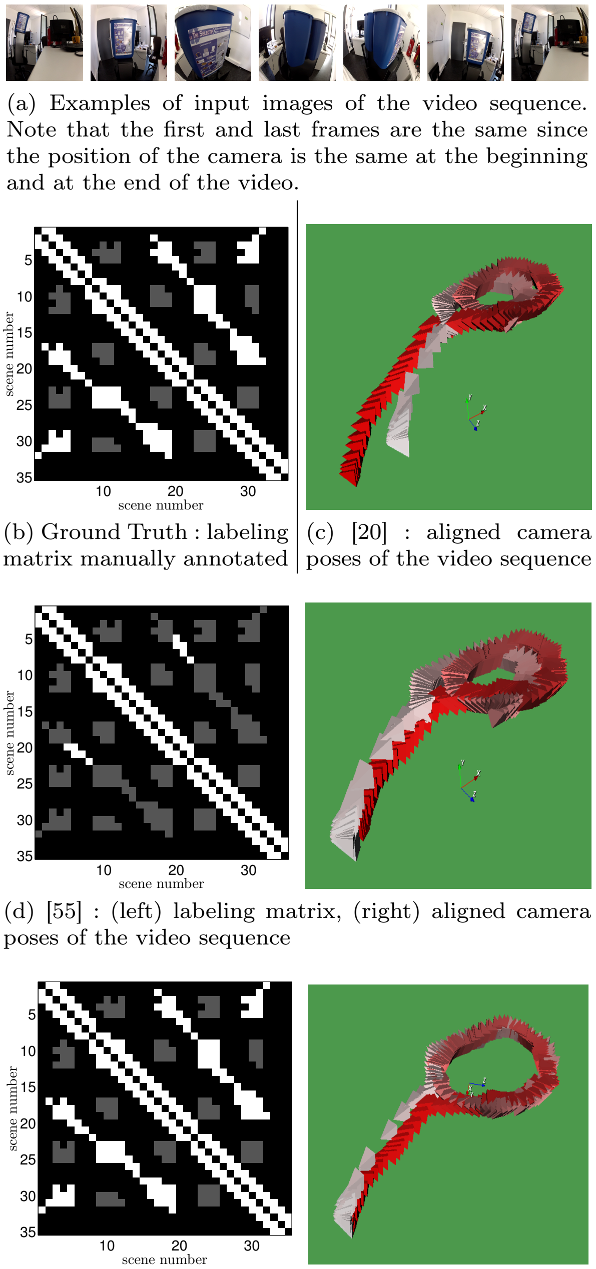
is defined on the whole group) while an inlier can be sampled using (70).

In Fig.5, we compare the proposed method (LG-IEKF with inlier test) against the robust approach proposed in [20] and the Expectation Maximization algorithm (EM) of [55]. [20] and [55] are initialized by composing the temporally consecutive relative measurements as it is proposed by the authors of those papers. As in Section 6.2.7, in order to compare the results of each approach to the true global motions, we align the estimated global motions with the true global motions using an LG-GN. The error obtained, for each approach, at convergence of the LG-GN is presented in Fig.5. We show that our method outperforms both [20] and [55]. Indeed, [20] is based on a robust convex L2-L1 norm to mitigate the influence of the outliers. However, because of the Lie group curvature, the complete functional is not convex. Therefore, the algorithm usually gets stuck in a poor local minimum. [55] introduces latent variables to classify the relative motions as inliers or outliers. However, the labels obtained at the initialization of the global motions are very difficult to modify. Indeed, the E-step does not take into account the estimation errors of the current global motion estimates which is negligible only when N is small. Therefore, a lot of inliers remain classified as outliers. In comparison, our approach incrementally rejects outliers, taking into account the current uncertainty of the global motions, and refines its estimate with the inliers. Consequently, the global motions are correctly recovered. An example of recovered global motions with the three different approaches is presented in Fig.6.

6.3.4 Results on Real Data for the Partial 3D reconstruction merging problem

In this section, the LG-IEKF is applied to a partial 3D reconstruction merging problem (Lie group $Sim(3)$). The experimental setup is the following: the camera is hand-held and evolves around two duplicated objects (the position of the camera is the same at the beginning and at the end of the video). We split the video in several (half-overlapping) parts. For each part, we applied a SLAM algorithm similar to [41] and obtained a 3D point cloud of the scene as well as the camera poses. We estimated the 3D similarities (Lie group $Sim(3)$) between every pair of point clouds using a RANSAC algorithm followed by a LG-GN algorithm on $Sim(3)$. The covariance matrix of each relative motion is obtained by applying an intrinsic Gauss-Laplace approximation.

In Fig.7, we compare the results of our approach to the EM algorithm of [55] and the method of [20]. For



(a) Examples of input images of the video sequence. Note that the first and last frames are the same since the position of the camera is the same at the beginning and at the end of the video.

(b) Ground Truth : labeling matrix manually annotated

(c) [20] : aligned camera poses of the video sequence

(d) [55] : (left) labeling matrix, (right) aligned camera poses of the video sequence

(e) This paper : (left) labeling matrix, (right) aligned camera poses of the video sequence

Fig. 7: Wearable camera experiment: in the labeling matrices a white pixel corresponds to an inlier, a black pixel corresponds to an unavailable measurement, a gray pixel corresponds to an outlier. In the results, a cone represents the pose of a camera. The first camera position of the video sequence is dark red and the last camera position is bright red.

each algorithm, we provide the camera trajectory obtained by aligning the camera poses of each reconstruction with the estimated global 3D similarities. Both [55] and [20] are initialized by composing the control inputs relative similarities.

In order to qualitatively compare the results, we manually annotated the 3D similarity measurements either as inlier or as outlier. Additionally, the position of the camera is the same at the beginning and at the end of the video, thus the first and the last estimated camera pose should be perfectly superimposed.

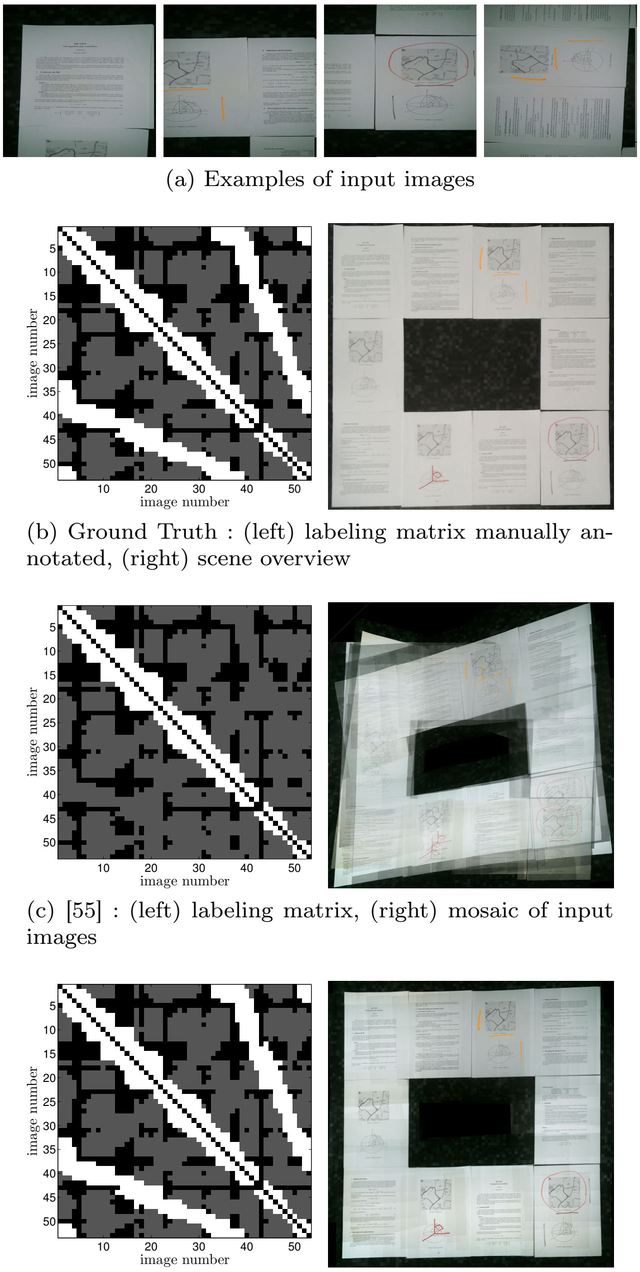
On the one hand, both the EM algorithm of [55] and the method of [20] remains stuck in a poor local minimum. Indeed, the first and the last camera poses are far from being superimposed and a lot of inliers are classified as outliers by [55] (see Fig.7d). On the other hand, our approach perfectly infers the set of inliers (see Fig.7e) while the first camera pose and the last one are almost perfectly superimposed.

6.3.5 Results on Real Data for the Automatic planar image mosaicking problem

In this section, the LG-IEKF presented in this paper is applied to an automatic planar image mosaicking problem. We took 53 photos of a planar scene (see Fig.8a) with a smartphone, detected points of interest and estimated the homographies (Lie group $SL(3)$) between every pair of images using a RANSAC algorithm followed by a LG-GN algorithm on $SL(3)$. The covariance matrix of each relative motion is obtained by applying an intrinsic Gauss-Laplace approximation. In our implementation, we use the Lie algebra basis of $\mathfrak{sl}(3)$ given in [6]. In this dataset, there are 65% of outliers (see Fig.8b) due to the ambiguity of the scene (some paper sheets are almost identical). Note that the relative homography between image 43 and image 44 is not available. Instead, we employ the relative homography between image 43 and image 45 as control input.

In Fig.8, we compare the results of our approach against the EM algorithm of [55] which is initialized by composing the control inputs relative homographies. On the one hand, once again, the proposed EM of [55] classifies a lot of inliers as outliers since the estimation errors of the global motions estimates is not taken into account during the E-step. Consequently, [55] is not able to correctly recover the global motions (see Fig.8c). On the other hand, our approach perfectly infers the set of inliers and produces a mosaic visually close to the ground truth (see Fig.8d).

We could not apply [20] because $\log^{\vee}_{SL(3)}$ is not defined on the whole group.



(a) Examples of input images

(b) Ground Truth : (left) labeling matrix manually annotated, (right) scene overview

(c) [55] : (left) labeling matrix, (right) mosaic of input images

(d) This paper : (left) labeling matrix, (right) mosaic of input images

Fig. 8: Image mosaicking results: in the labeling matrices a white pixel corresponds to an inlier, a black pixel corresponds to an unavailable measurement, a gray pixel corresponds to an outlier.

7 Conclusion

In this paper, we proposed a generic Iterated Extended Kalman Filter on Lie Groups (LG-IEKF) that allows to perform parameter estimation when the state and the measurements evolve on Lie groups. This novel filter extends the link between the minimization of non linear least squares criteria and the Iterated Extended Kalman Filter [5, 9] to Lie groups.

For this purpose, we first presented a fitting approach, called intrinsic Gauss-Laplace approximation, that allows to fit a concentrated Gaussian distribution on Lie groups to the probability density of a random variable evolving on a Lie group.

Since this fitting technique requires to find the minimizer of an intrinsic non linear least squares criterion, we presented a generic intrinsic GN algorithm on Lie groups, called LG-GN. This optimization algorithm has the advantage of taking intrinsically into account the geometry of the Lie group on which the parameters evolve.

Then, we showed that a generalization of the IEKF, to the case where the state and the observations evolve on Lie groups, can be obtained by employing intrinsic Gauss-Laplace approximations both to derive the LG-IEKF prediction step and the LG-IEKF update step (which involves the LG-GN algorithm). For each of these two steps, we were able to obtain a computationally efficient algorithm, by exploiting the specific structure of the problem.

Finally, we derived a statistical test on Lie groups to detect and remove outlier measurements.

In order to demonstrate the efficiency of the LG-IEKF, it has been applied to relative motion averaging problems for three Lie groups of interest. In each of these applications, the proposed approach significantly outperforms the state of the art algorithms.

Since this novel filter was derived for generic prediction and measurement models on Lie groups, it can be applied to a large class of problems.

Future work will consider its use in localization from a wearable camera in a known 3D environment as well as in visual odometry [7].

A Derivation of the LG-IEKF

A.1 Derivation of δ^{l+1}

Here, we derive the expression of the δ^{l+1} in (53):

$$
\begin{aligned}
&\delta^{l+1} = \log_G^{\vee}\left(X^{l+1}\left(X^l\right)^{-1}X^l\mu_{k|k-1}^{-1}\right)\\
&= \log_G^{\vee}\left(\exp_G^{\wedge}\left(\delta^{l+1/l}\right)\exp_G^{\wedge}\left(\delta^l\right)\right)\\
&\simeq \varphi_l\delta^{l+1/l} + \delta^l\\
&= \varphi_l\left(H_l^TQ_k^{-1}H_l + \varphi_l^TP_{k|k-1}^{-1}\varphi_l\right)^{-1}\\
&\left\{H_l^TQ_k^{-1}\ \log_{G'}^{\vee}\left(Z_kh\left(X^l\right)^{-1}\right) - \varphi_l^TP_{k|k-1}^{-1}\delta^l\right\} + \delta^l\\
&= \varphi_l\left(H_l^TQ_k^{-1}H_l + \varphi_l^TP_{k|k-1}^{-1}\varphi_l\right)^{-1}\\
&\left\{H_l^TQ_k^{-1}\log_{G'}^{\vee}\left(Z_kh\left(X^l\right)^{-1}\right) - \varphi_l^TP_{k|k-1}^{-1}\delta^l\right.\\
&\left.+\left(H_l^TQ_k^{-1}H_l + \varphi_l^TP_{k|k-1}^{-1}\varphi_l\right)\Phi_l\delta^l\right\}\\
&= \varphi_l\left(H_l^TQ_k^{-1}H_l + \varphi_l^TP_{k|k-1}^{-1}\varphi_l\right)^{-1}\\
&H_l^TQ_k^{-1}\left\{\log_{G'}^{\vee}\left(Z_kh\left(X^l\right)^{-1}\right) + H_l\Phi_l\delta^l\right\}\\
&= K_l\left\{\log_{G'}^{\vee}\left(Z_kh\left(X^l\right)^{-1}\right) + H_l\Phi_l\delta^l\right\} \qquad (78)
\end{aligned}
$$

where K_l is the Lie-Kalman gain derived in Appendix A.2.

A.2 Lie-Kalman Gain Derivation

Here, we derive the expression of the Lie-Kalman gain (54) (the superscripts and underscripts are omitted):

$$
\begin{aligned}
&K = \varphi\left(H^TQ^{-1}H + \varphi^TP^{-1}\varphi\right)^{-1}H^TQ^{-1}\\
&= \varphi\left(H^TQ^{-1}H + \varphi^TP^{-1}\varphi\right)^{-1}\\
&\left(H^TQ^{-1}\left(H\Phi P\Phi^TH^T + Q\right)\left(H\Phi P\Phi^TH^T + Q\right)^{-1}\right)\\
&= \varphi\left(H^TQ^{-1}H + \varphi^TP^{-1}\varphi\right)^{-1}\\
&\left(\left(H^TQ^{-1}H\Phi P\Phi^TH^T + H^T\right)\left(H\Phi P\Phi^TH^T + Q\right)^{-1}\right)\\
&= \varphi\left(H^TQ^{-1}H + \varphi^TP^{-1}\varphi\right)^{-1}\\
&\left(\left(\left(H^TQ^{-1}H + \varphi^TP^{-1}\varphi\right)\Phi P\Phi^TH^T\right)\left(H\Phi P\Phi^TH^T + Q\right)^{-1}\right)\\
&= P\Phi^TH^T\left(H\Phi P\Phi^TH^T + Q\right)^{-1} \qquad (79)
\end{aligned}
$$

A.3 Covariance Update Derivation

Here, we derive the expression of the $P_{k|k}$ in (61) (the superscripts and underscripts are omitted):

$$
\begin{aligned}
&P_{k|k} = \left(H^TQ^{-1}H + \varphi^TP^{-1}\varphi\right)^{-1}\\
&= \left(H^TQ^{-1}H + \varphi^TP^{-1}\varphi\right)^{-1}\\
&\left\{\left(H^TQ^{-1}H + \varphi^TP^{-1}\varphi\right)\Phi P\Phi^T - H^TQ^{-1}H\Phi P\Phi^T\right\}\\
&= \Phi P\Phi^T - \left(H^TQ^{-1}H + \varphi^TP^{-1}\varphi\right)^{-1}H^TQ^{-1}H\Phi P\Phi^T\\
&= \Phi P\Phi^T - \Phi KH\Phi P\Phi^T\\
&= \Phi\left(Id - KH\Phi\right)P\Phi^T \qquad (80)
\end{aligned}
$$

where K is defined in (79).

B Relative Motion Averaging

B.1 Derivation of F_k

From (38) and (72), we have:

$$\begin{aligned}
&\log^{\vee}_{G_k}\left(f\left(\mu_{k-1|k-1}\right)f\left(\exp^{\wedge}_{G_{k-1}}\left(\delta\right)\mu_{k-1|k-1}\right)^{-1}\right)\\
=&\log^{\vee}_{G_k}\left(\begin{bmatrix}Y_{(k+1)\,k} & \mathbf{0}\\ \mathbf{0} & Id\end{bmatrix}\begin{bmatrix}\left(\mu_{k-1|k-1}\right)_{kR} & \mathbf{0}\\ \mathbf{0} & \mu_{k-1|k-1}\end{bmatrix}\right.\\
&\left.\left(\begin{bmatrix}Y_{(k+1)\,k} & \mathbf{0}\\ \mathbf{0} & Id\end{bmatrix}\begin{bmatrix}\exp^{\wedge}_{G'}\left(\delta_{kR}\right)\left(\mu_{k-1|k-1}\right)_{kR} & \mathbf{0}\\ \mathbf{0} & \exp^{\wedge}_{G_{k-1}}\left(\delta\right)\mu_{k-1|k-1}\end{bmatrix}\right)^{-1}\right)\\
=&\log^{\vee}_{G_k}\left(\begin{bmatrix}Y_{(k+1)\,k} & \mathbf{0}\\ \mathbf{0} & Id\end{bmatrix}\begin{bmatrix}\left(\mu_{k-1|k-1}\right)_{kR} & \mathbf{0}\\ \mathbf{0} & \mu_{k-1|k-1}\end{bmatrix}\right.\\
&\left.\left(\begin{bmatrix}Y_{(k+1)\,k} & \mathbf{0}\\ \mathbf{0} & Id\end{bmatrix}\exp^{\wedge}_{G_k}\left(\begin{bmatrix}\delta_{kR}\\ \delta\end{bmatrix}\right)\begin{bmatrix}\left(\mu_{k-1|k-1}\right)_{kR} & \mathbf{0}\\ \mathbf{0} & \mu_{k-1|k-1}\end{bmatrix}\right)^{-1}\right)\\
=&\log^{\vee}_{G_k}\left(\begin{bmatrix}Y_{(k+1)\,k} & \mathbf{0}\\ \mathbf{0} & Id\end{bmatrix}\exp^{\wedge}_{G_k}\left(-\begin{bmatrix}\delta_{kR}\\ \delta\end{bmatrix}\right)\begin{bmatrix}Y^{-1}_{(k+1)\,k} & \mathbf{0}\\ \mathbf{0} & Id\end{bmatrix}\right)\\
=&\log^{\vee}_{G_k}\left(\begin{bmatrix}Y_{(k+1)\,k} & \mathbf{0}\\ \mathbf{0} & Id\end{bmatrix}\exp^{\wedge}_{G_k}\left(-\begin{bmatrix}\mathbf{0} & Id & \mathbf{0}\\ & Id & \end{bmatrix}\delta\right)\begin{bmatrix}Y^{-1}_{(k+1)\,k} & \mathbf{0}\\ \mathbf{0} & Id\end{bmatrix}\right)\\
=&-\begin{bmatrix}\mathrm{Ad}_{G'}\left(Y_{(k+1)\,k}\right) & \mathbf{0}\\ \mathbf{0} & Id\end{bmatrix}\begin{bmatrix}\mathbf{0} & Id & \mathbf{0}\\ & Id & \end{bmatrix}\delta
\end{aligned}\tag{81}$$

where we introduced the adjoint representation $\mathrm{Ad}_{G'}\left(\cdot\right)\subset\mathbb{R}^{q\times q}$ of G' on $\mathbb{R}^q$ that enables us to transform an increment $\epsilon^i_{ij}\in\mathbb{R}^q$, that acts onto an element Y_{ij} through left multiplication, into an increment $\epsilon^j_{ij}\in\mathbb{R}^q$, that acts through right multiplication:

$$\exp^{\wedge}_{G'}\left(\epsilon^i_{ij}\right)Y_{ij}=Y_{ij}\exp^{\wedge}_{G'}\left(\mathrm{Ad}_{G'}\left(Y^{-1}_{ij}\right)\epsilon^i_{ij}\right)\tag{82}$$

Consequently, from (81), we obtain:

$$\begin{aligned}
F_k=&-\left.\frac{d\log^{\vee}_{G_k}\left(f\left(\mu_{k-1|k-1}\right)f\left(\exp^{\wedge}_{G_{k-1}}(\delta)\,\mu_{k-1|k-1}\right)^{-1}\right)}{d\delta}\right|_{\delta=\mathbf{0}}\\
=&\begin{bmatrix}\mathrm{Ad}_{G'}\left(Y_{(k+1)\,k}\right) & \mathbf{0}\\ \mathbf{0} & Id\end{bmatrix}\begin{bmatrix}\mathbf{0} & Id & \mathbf{0}\\ & Id & \end{bmatrix}
\end{aligned}\tag{83}$$

B.2 Derivation of H_l

From (48) and (73), we have:

$$\begin{aligned}
&\log^{\vee}_{G'}\left(Z_k h\left(\exp^{\wedge}_{G}\left(\delta\right)X^l\right)^{-1}\right)\\
=&\log^{\vee}_{G'}\left(Y_{i\,(k+1)}\exp^{\wedge}_{G'}\left(\delta_{(k+1)R}\right)\left(X^l\right)_{(k+1)R}\left(X^l\right)^{-1}_{iR}\exp^{\wedge}_{G'}\left(-\delta_{iR}\right)\right)\\
=&\log^{\vee}_{G'}\left(Y_{i\,(k+1)}\left(X^l\right)_{(k+1)R}\left(X^l\right)^{-1}_{iR}\right.\\
&\left.\exp^{\wedge}_{G'}\left(\mathrm{Ad}_{G'}\left(\left(X^l\right)_{iR}\left(X^l\right)^{-1}_{(k+1)R}\right)\delta_{(k+1)R}\right)\exp^{\wedge}_{G'}\left(-\delta_{iR}\right)\right)\\
\simeq&\log^{\vee}_{G'}\left(Y_{i\,(k+1)}\left(X^l\right)_{(k+1)R}\left(X^l\right)^{-1}_{iR}\right)\\
&-\delta_{iR}+\mathrm{Ad}_{G'}\left(\left(X^l\right)_{iR}\left(X^l\right)^{-1}_{(k+1)R}\right)\delta_{(k+1)R}
\end{aligned}\tag{84}$$

where we approximated $\varphi\left(\cdot\right)$ (see (10)) by Id.

Consequently, from (84), we obtain:

$$\begin{aligned}
H_l=&-\left.\frac{d\log^{\vee}_{G'}\left(Z_k h\left(\exp^{\wedge}_{G}\left(\delta\right)X^l\right)^{-1}\right)}{d\delta}\right|_{\delta=\mathbf{0}}\\
\simeq&\begin{bmatrix}\mathbf{0} & Id\,\mathbf{0} & -\mathrm{Ad}_{G'}\left(\left(X^l\right)_{iR}\left(X^l\right)^{-1}_{(k+1)R}\right) & \mathbf{0}\end{bmatrix}
\end{aligned}\tag{85}$$

References

1. Absil, P.A., Mahony, R., Sepulchre, R.: Optimization algorithms on matrix manifolds. Princeton University Press (2009)
2. Agrawal, M.: A Lie algebraic approach for consistent pose registration for general Euclidean motion. In: IROS, pp. 1891–1897 (2006)
3. Bandeira, A.S., Singer, A., Spielman, D.A.: A Cheeger inequality for the graph connection laplacian. arXiv preprint arXiv:1204.3873 (2012)
4. Barfoot, T.D., Furgale, P.T.: Associating uncertainty with three-dimensional poses for use in estimation problems. IEEE Trans. Robot. **30**(3), 679–693 (2014). DOI 10.1109/tro.2014.2298059
5. Bell, B.M., Cathey, F.W.: The iterated Kalman filter update as a Gauss-Newton method. Automatic Control, IEEE Transactions on **38**(2), 294–297 (1993)
6. Benhimane, S., Malis, E.: Homography-based 2d visual tracking and servoing. The International Journal of Robotics Research **26**(7), 661–676 (2007)
7. Berger, J., Lenzen, F., Becker, F., Neufeld, A., Schnörr, C.: Second-order recursive filtering on the rigid-motion lie group se (3) based on nonlinear observations. arXiv preprint arXiv:1507.06810 (2015)
8. Bernstein, D.S.: Matrix mathematics: theory, facts, and formulas. Princeton University Press (2009)
9. Bertsekas, D.P.: Incremental least squares methods and the extended Kalman filter. SIAM Journal on Optimization **6**(3), 807–822 (1996). DOI 10.1137/S1052623494268522
10. Björck, A.: Numerical methods for least squares problems. Siam (1996)
11. Bonnabel, S.: Left-invariant extended Kalman filter and attitude estimation. In: IEEE Conference on Decision and Control (2007)
12. Bonnabel, S., Martin, P., Salaun, E.: Invariant extended Kalman filter : theory and application to a velocity-aided attitude estimation problem. In: IEEE Conference on Decision and Control and 28th Chinese Control Conference (2009)
13. Boumal, N., Singer, A., Absil, P.A.: Robust estimation of rotations from relative measurements by maximum likelihood. In: Proceedings of the 52nd Conference on Decision and Control, CDC. (2013)
14. Boumal, N., Singer, A., Absil, P.A., Blondel, V.D.: Cramer-Rao bounds for synchronization of rotations. Information and Inference (2013). DOI 10.1093/imaiai/iat006. URL http://dx.doi.org/10.1093/imaiai/iat006
15. Bourmaud, G., Mégret, R., Arnaudon, M., Giremus, A.: Continuous-discrete extended Kalman filter on matrix Lie groups using concentrated Gaussian distributions. J Math Imaging Vis (2014). DOI 10.1007/s10851-014-0517-0
16. Bourmaud, G., Mégret, R., Giremus, A., Berthoumieu, Y.: Global motion estimation from relative measurements in the presence of outliers. ACCV 2014

17. Bourmaud, G., Mégret, R., Giremus, A., Berthoumieu, Y.: Global motion estimation from relative measurements using iterated extended Kalman filter on matrix Lie groups. ICIP 2014
18. Bourmaud, G., Mégret, R., Giremus, A., Berthoumieu, Y.: Discrete extended Kalman filter on Lie groups. In: Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European (2013)
19. Caballero, F., Merino, L., Ferruz, J., Ollero, A.: Homography based Kalman filter for mosaic building. applications to UAV position estimation. In: Robotics and Automation, 2007 IEEE International Conference on, pp. 2004–2009. IEEE (2007)
20. Chatterjee, A., Govindu, V.M.: Efficient and robust large-scale rotation averaging. In: ICCV (2013)
21. Chirikjian, G.S.: Stochastic Models, Information Theory, and Lie Groups, Volume 2. Springer-Verlag (2012). DOI 10.1007/978-0-8176-4944-9
22. Crandall, D.J., Owens, A., Snavely, N., Huttenlocher, D.: Discrete-continuous optimization for large-scale structure from motion. In: CVPR, pp. 3001–3008 (2011)
23. Crassidis, J., Markley, F.: Unscented filtering for spacecraft attitude estimation. Journal of Guidance, Control, and Dynamics **26**, 536–542 (2003)
24. Davison, A.J., Reid, I.D., Molton, N.D., Stasse, O.: Monoslam: Real-time single camera slam (2007)
25. Enqvist, O., Kahl, F., Olsson, C.: Non-sequential structure from motion. In: ICCV Workshops, pp. 264–271 (2011)
26. Faraut, J.: Analysis on Lie groups: an introduction, vol. 110. Cambridge University Press (2008)
27. Fisher, R.A., Yates, F., et al.: Statistical tables for biological, agricultural and medical research. Statistical tables for biological, agricultural and medical research. (Ed. 3.) (1949)
28. Fredriksson, J., Olsson, C.: Simultaneous multiple rotation averaging using Lagrangian duality. In: Asian Conference of Computer Vision (2012)
29. Govindu, V.M.: Combining two-view constraints for motion estimation. In: CVPR (2), pp. 218–225 (2001)
30. Govindu, V.M.: Lie-algebraic averaging for globally consistent motion estimation. In: CVPR (1), pp. 684–691 (2004)
31. Grisetti, G., Kuemmerle, R., Ni, K.: Robust optimization of factor graphs by using condensed measurements. In: IEEE International Conference on Intelligent Robots and Systems (IROS) (2012)
32. Hall, B.: Lie Groups, Lie Algebras, and Representations An Elementary Introduction. Springer (2003)
33. Hartley, R., Aftab, K., Trumpf, J.: L1 rotation averaging using the Weiszfeld algorithm, pp. 3041–3048. Institute of Electrical and Electronics Engineers (2011). DOI 10.1109/CVPR.2011.5995745
34. Hartley, R., Trumpf, J., Dai, Y., Li, H.: Rotation averaging. International Journal of Computer Vision pp. 1–39 (2013). DOI 10.1007/s11263-012-0601-0
35. Haykin, S.S., et al.: Kalman filtering and neural networks. Wiley Online Library (2001)
36. Jeong, Y., Nister, D., Steedly, D., Szeliski, R., Kweon, I.: Pushing the envelope of modern methods for bundle adjustment. In: CVPR (2010)
37. Jiang, N., Cui, Z., Tan, P.: A global linear method for camera pose registration. 2013 IEEE International Conference on Computer Vision DOI 10.1109/iccv.2013.66
38. Kaess, M., Johannsson, H., Roberts, R., Ila, V., Leonard, J.J., Dellaert, F.: isam2: Incremental smoothing and mapping using the bayes tree. I. J. Robotic Res. **31**(2), 216–235 (2012)
39. Kaess, M., Ranganathan, A., Dellaert, F.: iSAM: Incremental smoothing and mapping. IEEE Trans. on Robotics (TRO) **24**(6), 1365–1378 (2008)
40. Khurd, P., Grady, L., Oketokoun, R., Sundar, H., Gajera, T., Gibbs-Strauss, S., Frangioni, J.V., Kamen, A.: Global error minimization in image mosaicing using graph connectivity and its applications in microscopy. Journal of Pathology Informatics **2, S8** (2012)
41. Klein, G., Murray, D.: Parallel tracking and mapping for small ar workspaces. In: International Symposium on Mixed and Augmented Reality (ISMAR) (2007)
42. Konolige, K.: Sparse sparse bundle adjustment. In: BMVC, pp. 1–11 (2010)
43. Lefferts, E., Markley, F., Shuster, M.: Kalman filtering for spacecraft attitude estimation. Journal of Guidance, Control and Dynamics (1982)
44. Li, G., Liu, Y., Yin, J., Shi, Z.: Newton geodesic optimization on special linear group. Proceedings of the 48h IEEE Conference on Decision and Control (CDC) (2009). DOI 10.1109/cdc.2009.5400115
45. Long, A.W., Wolfe, K.C., Mashner, M., Chirikjian, G.S.: The banana distribution is Gaussian: A localization study with exponential coordinates. In: Robotics: Science and Systems (2012)
46. Lui, Y.: Advances in matrix manifolds for computer vision. Image and Vision Computing **30**, 380–388 (2011)
47. Malis, E., T.Hamel, Mahony, R., Morin, P.: Dynamic estimation of homography transformations on the special linear group of visual servo control. In: IEEE Conference on Robotics and Automation (2009)
48. Markley, F.: Attitude error representation for Kalman filtering. Journal of Guidance, Control and Dynamics (2003)
49. Martinec, D., Pajdla, T.: Robust rotation and translation estimation in multiview reconstruction. In: Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, pp. 1–8. IEEE (2007)
50. Meidow, J.: Efficient video mosaicking by multiple loop closing. In: Photogrammetric Image Analysis, pp. 1–12. Springer (2011)
51. Moulon, P., Monasse, P., Marlet, R.: Global fusion of relative motions for robust, accurate and scalable structure from motion. In: Computer Vision (ICCV), 2013 IEEE International Conference on, pp. 3248–3255. IEEE (2013)
52. Persson, S., Sharf, I.: Invariant momentum-tracking Kalman filter for attitude estimation. In: Robotics and Automation (ICRA), 2012 IEEE International Conference on, pp. 592–598 (2012). DOI 10.1109/ICRA.2012.6224562
53. Ramachandra, K.: Kalman filtering techniques for radar tracking. CRC Press (2000)
54. Ring, W., Wirth, B.: Optimization methods on riemannian manifolds and their application to shape space. SIAM Journal on Optimization **22**(2), 596–627 (2012)
55. Roberts, R., Sinha, S.N., Szeliski, R., Steedly, D.: Structure from motion for scenes with large duplicate structures. In: Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, pp. 3137–3144. IEEE (2011)
56. Rudkovskii, M.: Twisted product of Lie groups. Siberian Mathematical journal pp. 969–977 (1997)
57. Selig, J.M.: Lie groups and Lie algebras in robotics. NATO Science Series II: Mathematics, Physics and Chemistry pp. 101–125 (2005)

58. Singer, A., Shkolnisky, Y.: Three-dimensional structure determination from common lines in cryo-em by eigenvectors and semidefinite programming. SIAM journal on imaging sciences **4**(2), 543–572 (2011)
59. Smith, P., Drummond, T., Roussopoulos, K.: Computing map trajectories by representing, propagating and combining pdfs over groups. In: ICCV, pp. 1275–1282 (2003)
60. Strasdat, H., Davison, A., Montiel, J., Konolige, K.: Double window optimisation for constant time visual SLAM. In: IEEE International Conference on Computer Vision (ICCV) (2011)
61. Strasdat, H., Montiel, J.M.M., Davison, A.J.: Real-time monocular slam: Why filter? In: ICRA, pp. 2657–2664 (2010)
62. Taylor, C.J., Kriegman, D.J.: Minimization on the Lie group SO(3) and related manifolds (1994)
63. Vercauteren, T., Pennec, X., Malis, E., Perchant, A., Ayache, N.: Insight into efficient image registration techniques and the demons algorithm. In: Information Processing in Medical Imaging, pp. 495–506. Springer (2007)
64. Wachinger, C., Wein, W., Navab, N.: Registration strategies and similarity measures for three-dimensional ultrasound mosaicing. Academic radiology **15**(11), 1404–1415 (2008)
65. Wang, L., Singer, A.: Exact and stable recovery of rotations for robust synchronization. Information and Inference **2**(2), 145–193 (2013). DOI 10.1093/imaiai/iat005
66. Wang, Y., Chirikjian, G.: Error propagation on the Euclidean group with applications to manipulators kinematics. IEEE Transactions on Robotics **22** (2006)
67. Wolfe, K., Mashner, M., Chirikjian, G.: Bayesian fusion on Lie groups. Journal of Algebraic Statistics **2**, 75–97 (2011)