HAL
open science

Invariant Smoothing on Lie Groups

Paul Chauchat, Axel Barrau, Silvere Bonnabel

To cite this version:

Paul Chauchat, Axel Barrau, Silvere Bonnabel. Invariant Smoothing on Lie Groups. IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2018, Oct 2018, Madrid, Spain. hal-01725847

HAL Id: hal-01725847

https://hal.science/hal-01725847v1

Submitted on 7 Mar 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Invariant smoothing on Lie Groups

Paul CHAUCHAT, Axel BARRAU and Silvère BONNABEL

***Abstract*—In this paper we propose a (non-linear) smoothing algorithm for group-affine observation systems, a recently introduced class of estimation problems on Lie groups that bear a particular structure. As most non-linear smoothing methods, the proposed algorithm is based on a maximum a posteriori estimator, determined by optimization. But owing to the specific properties of the considered class of problems, the involved linearizations are proved to have a form of independence with respect to the current estimates, leveraged to avoid (partially or sometimes totally) the need to relinearize. The method is validated on a robot localization example, both in simulations and on real experimental data.**

I. INTRODUCTION

Statistical state estimation based on sensor fusion has played a major role in localization for the last decades. Most of the proposed solutions revolved around filtering, especially under the framework of the standard Extended Kalman Filter (EKF) [20]. Many alternative filters have been introduced to overcome the inherent limitations of the EKF, such as particle filters or the Unscented Kalman Filter. In the robotics field, and especially in the Simultaneous Localization and Mapping (SLAM) community, the smoothing approach, based on the Gauss-Newton (GN) algorithm, was proposed in order to take advantage of the sparsity of the involved matrix, and reduce the consequences of wrong linearization points [18]. It had a tremendous impact and led to many of the state-of-the-art algorithms for SLAM and visual odometry systems [22,28,32], to cite a few. It was more recently applied to GPS aided inertial navigation, showing promising results [27,36]. In the meantime, the advantages of the Lie group structure for probabilistic robotics, and especially that of the configuration spaces $SE(d)$, have been widely recognized [4,16,17,19,26, 35]. This led to new non-linear filtering algorithms [7,10,13] which showed remarkable properties [5,7], and to geometrical methods at the core of all the recent solvers for smoothing problems involving relative poses [22,32].

The smoothing technique proposed in this paper builds upon the invariant filtering framework, whose main results can be found in [7,10]. An overview of the theory and its applications is made in [9], and a focus on the case of SLAM in [5]. Similar results were achieved with observers [30]. A more tutorial approach to the theory, along with some new results, can also be found in the recent paper [8] dedicated to the discrete-time case. In the present paper, we devise a smoothing framework adapted to robot localization, where the dependence of the whole information matrix of the system with respect to the linearization points is reduced, or even removed under mild conditions. This brings immediate benefits: 1) The number of iterations to convergence is reduced 2) The consistency of the smoother is improved. As usual with estimation methods derived from the theory of invariant filtering this comes at a price: the framework is adapted to a specific class of systems called *group-affine observation systems* defined in [7] (the equations of inertial navigation [7], as well as SLAM [5], fall within this framework).

Paul Chauchat and Silvere Bonnabel are with MINES ParisTech, PSL Reasearch University, Centre for Robotics, 60 bd Saint-Michel, 75006 Paris, France paul.chauchat@mines-paristech.fr, silvere.bonnabel@mines-paristech.fr

Axel Barrau is with SAFRAN TECH, Groupe Safran, Rue des Jeunes Bois - Châteaufort, 78772 Magny Les Hameaux CEDEX, France axel.barrau@safrangroup.com

A. Links and differences with the previous literature

Smoothing on manifolds is well established, and a thorough study of a Gauss-Newton algorithm on Riemannian manifolds can be found in [1]. In the SLAM community, a lot of work has been dedicated to studying the non-Euclidean structure of the state-space [3,16,17]. Kuemmerle et al. introduced pose-graph algorithms for general manifolds [24,29]. Most of the recent work on SLAM focused on the 2D and 3D Special Euclidean groups $SE(2)$ and $SE(3)$ which are used to represent position-orientation states, especially in the pose-SLAM framework. The majority uses the group structure to define the cost function, and we then recover a group synchronization problem [11]. To solve it, most methods revolve around using the manifold structure of the state-space and custom retractions [21] or using the matrix structure to relax the problem [31]–[33]. The recent work of Bourmaud et al [12,14] is, to the authors' knowledge, the only one which fully relies on the Lie group structure for the estimate. In a way, the proposed method can be seen as an instance of this algorithm, but where the choice of the uncertainty representation is commanded by the system instead of being free, and an additional trick regarding the observation factors provides new properties.

These approaches are focused on leveraging the intrinsic tools of a given structure of the state space (Riemannian manifold or Lie group). For instance in [14], the proposed algorithm only required the state to belong to a Lie group, but no assumption is made regarding the dynamics and observation functions. In the present paper, the equations of the system are assumed to have specific properties with respect to the chosen group structure. This reduces the application field, but ensures in turn striking properties.

B. Organization of the paper

The paper is divided as follows. Section II contains mathematical preliminaries. Section III recalls the standard framework of smoothing, both in the vector and manifold

settings. Section IV presents our smoothing algorithm for group-affine systems on Lie groups, which we call invariant smoothing, and which is the main contribution of the paper. Section V illustrates the theory with both a simulated and an experimental 2D robot localization problems, showing the advantage of the proposed method over classical smoothing algorithms, and also over the recently developed (invariant) Kalman filter based UKF-LG and IEKF (Invariant Extended Kalman Filter) [7,15].

II. Mathematical preliminaries

A. Lie Groups

In this section we recall the definitions and basic properties of matrix Lie groups, Lie algebra and random variables on Lie groups. A matrix Lie group $G \subset \mathbb{R}^{N\times N}$ is a set of square invertible matrices that is a group, i.e., the following properties hold:

$$I_N \in G;\ \forall \chi \in G, \chi^{-1} \in G;\ \forall \chi_1, \chi_2 \in G, \chi_1\chi_2 \in G \tag{1}$$

Locally about the identity matrix I_N, the group G can be identified with an vector space $\mathbb{R}^q$ using the matrix exponential map $\exp_m(.)$, where $q = \dim G$. Indeed, to any $\xi \in \mathbb{R}^q$ one can associate a matrix $\xi^\wedge$ of the tangent space of G at I_N, called the Lie algebra $\mathfrak{g}$. We then define the exponential map $\exp : \mathbb{R}^q \to G$ for Lie groups as

$$\exp(\xi) = \exp_m\left(\xi^\wedge\right), \tag{2}$$

Locally, it is a bijection, and one can define the Lie logarithm map $\log : G \to \mathbb{R}^q$ as its inverse: $\log(\exp(\xi)) = \xi$.

As Lie groups are not necessarily commutative, in general we have $\exp(x)\exp(y) \neq \exp(x+y)$ for $x, y \in \mathfrak{g}$. However, the vector z such that $\exp(x)\exp(y) = \exp(z)$ satisfies the Baker-Campbell-Hausdorff (BCH) formula [23] $z = BCH(x,y) = x + y + r(x,y)$, where $r(x,y)$ contains higher order terms.

A last classical tool is the (inner) automorphism $\Psi_a \in Aut(G)$ defined for each $a \in G$ as $\Psi_a :\quad g \mapsto aga^{-1}$. Its differential at the identity element Id of G is called *adjoint operator* and denoted by $Ad_a : \mathfrak{g} \mapsto \mathfrak{g}$. It satisfies :

$$\forall a \in G, u \in \mathfrak{g},\ a\exp(u)a^{-1} = \exp(Ad_a u).$$

B. Uncertainties on Lie Groups

To define random variables on Lie groups, we cannot apply the usual approach of additive noise for $\chi_1, \chi_2 \in G$ as G is not a vector space, i.e., generally $\chi_1 + \chi_2 \notin G$ does not hold. In contrast, we adopt the framework of [4], see also [6], which is slightly different from the pioneering approach of [16,17]. Indeed, we define the probability distribution $\chi \sim \mathcal{N}_L(\bar{\chi}, \mathbf{P})$ for the random variable $\chi \in G$ as

$$\chi = \bar{\chi}\exp(\xi),\ \xi \sim \mathcal{N}(\mathbf{0}, \mathbf{P}), \tag{3}$$

where $\mathcal{N}(.,.)$ is the classical Gaussian distribution in Euclidean space and $\mathbf{P} \in \mathbb{R}^{q\times q}$ is a covariance matrix. In the sequel, we will refer to (3) as the left-invariant Gaussian distribution on G, owing to the fact that the discrepancy $\bar{\chi}^{-1}\chi$ which is invariant to left multiplications $(\bar{\chi}, \chi) \mapsto (\Gamma\bar{\chi}, \Gamma\chi)$, is the exponential of a Gaussian ξ. In (3), the noise-free quantity $\bar{\chi}$ is viewed as the mean, and the dispersion arises through left multiplication with the exponential of a Gaussian random variable. Similarly, the distribution $\chi \sim \mathcal{N}_R(\bar{\chi}, \mathbf{P})$ can be defined through right multiplication as

$$\chi = \exp(\xi)\bar{\chi},\ \xi \sim \mathcal{N}(\mathbf{0}, \mathbf{P}). \tag{4}$$

We stress that we have defined these probability density functions directly in the vector space $\mathbb{R}^q$ such that both $\mathcal{N}_L(.,.)$ and $\mathcal{N}_R(.,.)$ are not Gaussian distributions.

III. Classical smoothing

A. Maximum a posteriori (MAP) estimate

Contrarily to filtering approaches, which only consider the latest state of the system, smoothing methods aim at recovering the maximum a posteriori estimate (MAP) of the full trajectory given a set of measurements $Z = (y_1, \cdots, y_m)$:

$$\chi_0^*, \cdots, \chi_n^* = \underset{(\chi_i)_{i\leq n}}{\operatorname{argmax}}\, p((\chi_i)_{i\leq n} \mid Z) \tag{5}$$

Under a markovian assumption on the trajectory and conditional independence of the measurements, and given a prior on the initial state χ_0, this becomes

$$\chi_0^*, \cdots, \chi_n^* = \underset{(\chi_i)_{i\leq n}}{\operatorname{argmax}}\, p(\chi_0)\prod_i p(\chi_{i+1}|\chi_i)\prod_k p(y_k|(\chi_i)_{i\leq n}) \tag{6}$$

In this paper, we consider partial measurements of a single state, i.e. we have for some instant t_k:

$$\forall k,\ p(y_k|(\chi_i)_{i\leq n}) = p(y_k|\chi_{t_k}) \tag{7}$$

B. Gaussian Smoothing on vector spaces

In the general case, (6) represents a difficult high dimensional optimization problem. However, if the noises are supposed to be additive and Gaussian, this boils down to a non-linear least-squares problem [3].

Consider the following prior, dynamics, and measurement model:

$$\chi_0 = \bar{\chi}_0 + z_0 \qquad \chi_{i+1} = f_i(\chi_i) + w_i \qquad y_k = h(\chi_{t_k}) + v_k \tag{8}$$

where z_0, w_i and v_k are gaussian white noises of respective covariances P_0, Q_i and R_k. Then the MAP estimates satisfies:

$$\chi_0^*, \cdots, \chi_n^* = \underset{(\chi_i)_{i\leq n}}{\operatorname{argmin}}\, \|\chi_0 - \bar{\chi}_0\|^2_{P_0} + \sum_i \|f_i(\chi_i) - \chi_{i+1}\|^2_{Q_i} + \sum_k \|h(\chi_{t_k}) - y_k\|^2_{R_k} \tag{9}$$

where we used the Mahalanobis norm $\|e\|^2_\Sigma = e^T\Sigma^{-1}e$.

Solving (9) is done through iterative algorithms involving sequential linearizations, such as Gauss-Newton (GN) [18, 36].

For $i \leq n$, denoting by $\hat{\chi}_i$ the current estimates and letting $\delta\chi_i$ be the search parameter for the current iteration, each iteration consists in solving the following least-squares problem, based on a first-order Taylor expansion of (9)

$$(\delta\chi_i)_{i\leq n} = \underset{(\delta\chi_i)_{i\leq n}}{\operatorname{argmin}}\, \|p_0 + \delta\chi_0\|^2_{P_0} + \sum_i \|a_i + F_i\delta\chi_i - \delta\chi_{i+1}\|^2_{Q_i} + \sum_k \|c_k + H_k\delta\chi_{t_k}\|^2_{R_k} \tag{10}$$

where $p_0 = \hat{\chi}_0 - \bar{\chi}_0$, $a_i = f_i(\hat{\chi}_i) - \hat{\chi}_{i+1}$, $c_k = h(\hat{\chi}_{t_k}) - y_k$, F_i and H_k are the Jacobians of f_i and h_{t_k} evaluated at $\hat{\chi}_i$ and $\hat{\chi}_{t_k}$ respectively. Each of the terms of the sum are referred to as factors. This name comes from the link between the MAP estimation problem and graphical models, in particular factor graphs [18], which lead to powerful solvers [28,29,34] still used in many recent applications [22].

In this paper, we focus on the non-linear least-squares problem which is to be solved, and argue that a better parametrization taking into account the geometry of the system will lead to better convergence properties, regardless of the solver used.

C. Smoothing on manifolds

A large body of literature was dedicated to smoothing of systems which do not live in vector spaces but on a manifold $\mathscr{M}$ [21,24]. The main idea is to replace the $-$ and $+$ operators of the vector space by adapted ones $\boxminus$, $\boxplus$, which highlight the fact that $\mathscr{M}$ is not a vector space and that linearization takes place on its tangent space $T\mathscr{M}$, where the $\delta\chi_i$ should live. Thus, these operators are defined such that $\boxminus : \mathscr{M} \times \mathscr{M} \to T\mathscr{M}$ and $\boxplus : \mathscr{M} \times T\mathscr{M} \to \mathscr{M}$.

Therefore, a smoothing algorithm can be divided into two main parts, the non-linear and the linear ones. The former dictates how the jacobians and errors are computed, and how the updates are handled, while the latter is the solver, only concerned with linear algebra, as illustrated on Figure 1. In this work, we focus on the modeling part, and propose a new framework which can be used with any existing solver.

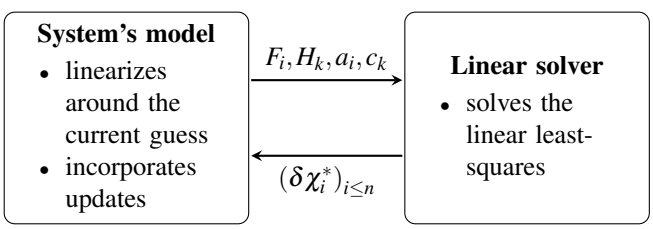

Fig. 1: Schematic representation of the two main components of a smoothing algorithm, the non linear model of the system which handles the transitions from the manifold to the tangent space and back, and the solver which inverts the obtained linear system.

When the manifold is a Lie group, the group multiplication and its intrinsic operators log and exp naturally yield a family of such operators, through the following definitions:

$$\boxplus_L(\chi,\xi) = \exp(\xi)\chi \qquad \boxminus_L(\chi_1,\chi_2) = \log(\chi_1\chi_2^{-1}) \quad (11)$$

$$\boxplus_R(\chi,\xi) = \chi\exp(\xi) \qquad \boxminus_R(\chi_1,\chi_2) = \log(\chi_2^{-1}\chi_1) \quad (12)$$

Note that, these operators were used in the definition of Section II. Smoothing methods on (Lie) groups were used extensively for the particular cases of pose-SLAM, visual odometry, and group synchronization problems in general [11,22,32], which only consider relative measurements of the full states. They were recently extended to a more general case by Bourmaud et al [14], but focusing on the Lie group structure essentially, without discussing the best choice of group structure with respect to the system's equations. In the next section, we focus our study on a particularly rich class of models on Lie groups.

IV. Invariant smoothing

In this section, we consider smoothing on Lie groups for a special class of systems. The remarkable result is that the dependency of the information matrix of smoothing algorithms (partially) disappears, owing to the interplay between the group structure and the system's properties.

A. Group-affine observation systems

In this work we argue that, just as vector spaces are particularly well-suited to linear systems, the Lie group formalism is beneficial to a particular class of dynamics called "group-affine observation systems". They were intially introduced for continuous-time dynamics [7,8] but are easily adapted to the discrete case [8,9]. Here we briefly introduce them and recall some of their properties.

Definition 1: A discrete time system (17) with state χ_i at time i being an element of a Lie group G is said to be "group-affine" if its dynamics/observation couple writes:

$$\forall i,k, \quad \begin{aligned} \chi_{i+1} &= f_i(\chi_i) \\ y_k &= \chi_{t_k} d \end{aligned} \quad (13)$$

$$\text{where } \forall a,b \in G, f_i(ab) = f_i(a) f_i(Id)^{-1} f_i(b) \quad (14)$$

The dynamics part of this definition is called group-affine dynamics and has a very powerful property, that will lead to interesting results in the sequel. Note that a continuous-time counterpart of these dynamics also exists in the theory of control on Lie groups [2].

Theorem 1 (from [7]): f defines a group-affine dynamics if and only if $g^L(\chi) =: f(Id)^{-1} f(\chi)$ is a group automorphism, i.e., satisfies

$$\forall a,b,\ g^L(ab) = g^L(a)g^L(b), \text{ and thus } g^L(a^{-1}) = g^L(a)^{-1} \quad (15)$$

In that case, the Lie-group/Lie algebra morphism correspondance ensures the existence of a $q \times q$ matrix G^L such that

$$\forall \xi, g^L(\exp(\xi)) = \exp(G^L\xi). \quad (16)$$

B. Noisy system

For the remainder of this paper we will focus on the following dynamics, observation and noise models :

$$\chi_{i+1} = f_i(\chi_i)\exp(w_i) \qquad y_k = \chi_{t_k} d + v_k \quad (17)$$

where w_i and v_k are white noise similar to those of (8), d is a given vector of $\mathbb{R}^q$, and f_i possesses the group-affine property (13) for all i.

In this model, the measurements are left-equivariant, a fact which according to [7] prompts the use of the left multiplication based uncertainty model (3). Therefore, given a current guess $\hat{\chi}$, the cost function will be linearized, noting ξ the searched parameter, using

$$\chi = \hat{\chi}\exp(\xi). \quad (18)$$

C. Factors definition

Under the Gaussian and conditional independence assumptions, defining the factors boils down to isolating the noise in (17). We have $\exp(w_i) = f_i(\chi_i)^{-1}\chi_{i+1}$. Let $\hat{\chi}_i, \hat{\chi}_{i+1}$ be some guesses, and let us write the true χ_i, χ_{i+1} using the Lie-group based error model (18). This yields

$$\exp(w_i) = f_i(\chi_i)^{-1}\chi_{i+1} \tag{19}$$
$$= f_i(\hat{\chi}_i \exp(\xi_i))^{-1}\hat{\chi}_{i+1}\exp(\xi_{i+1}) \tag{20}$$

Using (13), and then (16) this is equal to

$$= (f_i(\hat{\chi}_i) f_i(Id)^{-1} f_i(\exp(\xi_i)))^{-1}\hat{\chi}_{i+1}\exp(\xi_{i+1})$$
$$= (f_i(\hat{\chi}_i) g_i^L(\exp(\xi_i)))^{-1}\hat{\chi}_{i+1}\exp(\xi_{i+1})$$
$$= \exp(-G_i^L\xi_i) f_i(\hat{\chi}_i)^{-1}\hat{\chi}_{i+1}\exp(\xi_{i+1})$$

Let $a_i = \log(f_i(\hat{\chi}_i)^{-1}\hat{\chi}_{i+1})$ be the estimation error for the guess. Then, taking the logarithm of both hand sides and using the BCH formula, only keeping the first order terms in ξ *and* a_i leads to

$$w_i \approx a_i - G_i^L\xi_i + \xi_{i+1} \tag{21}$$

The major advantage (and very remarkable feature) of this parametrization is the fact that, if the dynamics is accurate enough the jacobians *do not* depend on the current estimate since G_i^L is independent of $\hat{\chi}$ from Theorem 1.

To get the measurement factors we propose to consider the following quantities

$$\hat{\chi}_{t_k}^{-1} v_k = \hat{\chi}_{t_k}^{-1} y_k - \exp(\xi_{t_k}) d. \tag{22}$$

By letting $c_k = \hat{\chi}_{t_k}^{-1} y_k$, linearization then leads to

$$\hat{\chi}_{t_k}^{-1} v_k \approx c_k - H\xi_{t_k} \tag{23}$$

where H is the matrix defined by $H\xi = \xi^\wedge d$, where we used a Taylor expansion of the matrix exponential in (2).

Finally, the factor associated to the prior must also be assessed. Indeed, the noise on the prior now writes

$$\exp(z_0) = \bar{\chi}_0^{-1}\hat{\chi}_0\exp(\xi_0) \tag{24}$$

Contrarily to the propagation factor, here the first-order terms of the BCH formula cannot be reduced to the sum of its arguments, as the term $\log(\bar{\chi}_0^{-1}\hat{\chi}_0)$ will tend to grow over the iterations, sometimes to a point where the jacobian of the Lie group [3] must be taken into account. We thus have in general

$$z_0 \approx p_0 + J_0\xi_0 \tag{25}$$

where $p_0 = \log(\bar{\chi}_0^{-1}\hat{\chi}_0)$, and J_0 is defined as $BCH(p_0, \xi) = p_0 + J_0\xi + o(\|\xi\|^2)$

Therefore, the linear least-squares problem to be solved at each iteration writes

$$(\xi_i)_{i\leq n} = \underset{(\xi_i)_{i\leq n}}{\operatorname{argmin}} \|p_0 + J_0\xi_0\|_{P_0}^2 + \sum_i \|a_i - G_i^L\xi_i + \xi_{i+1}\|_{Q_i}^2 + \sum_k \|c_k + H\xi_{t_k}\|_{\hat{R}_k}^2 \tag{26}$$

where $\hat{N}_k = \hat{\chi}_{t_k}^{-1} N_k \hat{\chi}_{t_k}^{-T}$ is the covariance of $\hat{\chi}_{t_k}^{-1} v_k$, see (23).

Algorithm 1: Smoothing for a group-affine system with a left multiplication based parametrization (18)

Input: $(\bar{\chi})_{1\leq i\leq n}, \mathbf{P}_0, (f_i)_i, (Q_i)_i, (y_k)_k, (N_k)_k$;
Initialization
Set $\hat{\chi}_i^0 = \bar{\chi}_i,\ i = 1, \cdots, n$
Until convergence do
Linearize around $(\hat{\chi}_i^k)_i$ according to (26) ;
Solve for $(\xi_i)_i$;
Update : $\hat{\chi}_i^{k+1} := \hat{\chi}_i \exp(\xi_i),\ i = 1, \cdots, n$
Set $\chi_i^* = \hat{\chi}_i^k,\ i = 1, \cdots, n$;
Output: $(\chi_i)_{i\leq n}^*$;

D. Final algorithm and benefits of the proposed approach

Injecting (26) into the smoothing framework, we get Algorithm 1 for the left multiplication based parametrization (18) (as opposed to $\chi = \exp(\xi)\hat{\chi}$). The main advantage of this framework is the fact that the current estimates do not appear in the definition of the propagation and observation jacobians, in (21) and (23) respectively (matrices G^L and H), although the latter comes at the expense of a modified covariance for the measurement factors. However, this dependency can often disappear, leaving only the prior jacobian of (24) to be a function of the estimate. If the initial guess is good, this can also be harmlessly relaxed by approximating $J_0 \approx Id$. This is summarized in the following proposition.

Proposition 1: We have the following properties: (i) Consider the system (17). If f_i is group-affine for all i, then all the propagation jacobians are independent from the current estimate. (ii) Moreover, if the measurement covariance N_k is such that $N_k = \chi N_k \chi^T$ for all χ, then the observation Jacobians are also independent from the estimates. (iii) Finally, if the initial guess is good, i.e. $J_0 \approx Id$, then the information matrix associated to (26) is independent from the current estimate.

Although the condition on the measurement covariance can seem restrictive at first, it often boils down to an isotropy assumption, as will be the case for the model of Section V.

V. Application to mobile robot localization

A. Considered problem: robot localization

To evaluate the performances of the developed approach, tests were conducted for a wheeled robot localization problem, using the standard non-linear equations of the 2D differential drive car modeling the position $x_i \in \mathbb{R}^2$ and heading $\theta_i \in \mathbb{R}$ of the robot. The odometer velocity is integrated between two time steps to give a position shift $u_i \in \mathbb{R}^2$ and the angular shift $\omega_i \in \mathbb{R}$ is measured through (differential) odometry and/or gyroscopes. The discrete noisy model writes:

$$\theta_{i+1} = \theta_i + \omega_i + w_i^\omega$$
$$x_{i+1} = x_i + R(\theta_i)(u_i + w_i^x) \tag{27}$$

where w_i^ω is the angular measurement error, w_i^x contains both the odometry and transversal shift errors, and $R(\theta) \in SO(2)$

denotes the planar rotation of angle θ. The vehicle also gets noisy position measurements (through e.g., GNSS) with standard deviation σ of the form

$$y_k = x_{t_k} + v_k, \qquad v_k \sim \mathcal{N}(\mathbf{0}, N_k = \sigma^2 I_2) \tag{28}$$

The state (and the increments) can be embedded in the matrix Lie group $SE(2)$, which is detailed in Appendix A, using the homogeneous matrix representation

$$\chi_i =: \begin{pmatrix} R(\theta_i) & x_i \\ 0_{1\times 2} & 1 \end{pmatrix}, \; U_i =: \begin{pmatrix} R(\omega_i) & u_i \\ 0_{1\times 2} & 1 \end{pmatrix}$$

and letting $\mathbf{w_i} = \begin{pmatrix} w_i^\omega & R(\omega_i)^T w_i^x \end{pmatrix}$ the stacked noise vector on the increments, we have using a first order expansion of the exponential map of $SE(2)$ (the increment noises are small if the time step is small)

$$\exp(\mathbf{w_i}) \approx \begin{pmatrix} R(w_i^\omega) & R(\omega_i)^T w_i^x \\ 0_{1\times 2} & 1 \end{pmatrix}.$$

Thus, the dynamics and the observations respectively rewrite in the desired form (8):

$$\chi_{i+1} = \chi_i U_i \exp(\mathbf{w_i}) = f_i(\chi_i)\exp(\mathbf{w_i}) \tag{29}$$

$$y_k = \chi_{t_k} \begin{pmatrix} 0_{2\times 1} \\ 1 \end{pmatrix} + \begin{pmatrix} v_k \\ 0 \end{pmatrix} \tag{30}$$

This system is group-affine and its measurement covariance is rotation-invariant (i.e., isotropic), therefore Proposition 1 (i) and (ii) hold. Moreover, we have, with the notations of Section IV and according to Section II,

$$g_i^L(\chi) = \Psi_{U_i^{-1}}(\chi) \tag{31}$$

B. Compared smoothing frameworks

In Sections V-C and V-D, the proposed smoothing method is compared with two non-invariant parametrizations, but which account for the non linear structure of the state space, from respectively [25] and [21], and a standard linear parametrization. The two former were developed for applications such as pose-SLAM or VIO, without directly considering navigation with absolute measurements. They were reimplemented in a batch setting, so that the focus was put exclusively on the parametrization. Writing the propagation linearized factor in the general form $a_i + F_i^i \xi_i + F_i^{i+1}$, the following jacobians were used, see [21,25] for the respective residuals' definitions :

	F_i^i	F_i^{i+1}
IS (ours)	$-Ad_{U_n^{-1}}$	Id
Lin	$\begin{bmatrix} -I_2 & -\hat{R}_i u_i \\ & -1 \end{bmatrix}$	Id
[25]	$\begin{bmatrix} -\Omega_i^T \hat{R}_i^T & J\Omega_i^T \hat{R}_i^T(\hat{x}_{i+1}-\hat{x}_i) \\ & -1 \end{bmatrix}$	$\begin{bmatrix} \Omega_i^T \hat{R}_i^T & \\ & 1 \end{bmatrix}$
[21]	$\begin{bmatrix} I_2 & J\hat{R}_i^T(\hat{x}_{i+1}-\hat{x}_i) \\ & 1 \end{bmatrix}$	$-\begin{bmatrix} \hat{R}_i^T \hat{R}_{i+1} & \\ & 1 \end{bmatrix}$

The observation factors are constructed using the following jacobians H_k, errors c_k and covariances :

	H_k	c_k	covariance
IS (ours)	$\begin{bmatrix} I_2 & 0 \end{bmatrix}$	$\hat{R}_{t_k}^T(y_k - \hat{x}_{t_k})$	$\hat{R}_{t_k}^T N_k \hat{R}_{t_k}$
Lin, [25]	$\begin{bmatrix} I_2 & 0 \end{bmatrix}$	$y_k - \hat{x}_{t_k}$	N_k
[21]	$\begin{bmatrix} \hat{R}_{t_k} & 0 \end{bmatrix}$	$y_k - \hat{x}_{t_k}$	N_k

C. Simulation results

The four smoothing methods of the previous section were compared. The vehicle is initialized at the true position but with a wrong heading of $-3\pi/4$, and an initial covariance of $diag(0.0025, 0.0025, (-3\pi/4)^2)$. The robot moves along a line at $7m/s$, where the odometry is polluted by a white noise of covariance $(0.1(m/s)^2, 0.01(rad/s)^2)$, and acquired at a rate of $10Hz$. Position measurements of covariance $0.01m^2$ are received every five steps. In this batch setting, Proposition 1 (iii) does not hold for the invariant framework.

The iterations of the four optimization schemes are displayed on Figure 2. As can be seen, the invariant parametrization converges faster than all the other ones, and much more smoothly. This is important for a sliding window smoothing setting, as restricting the number of solver iterations at each step will have a weaker impact on the estimation.

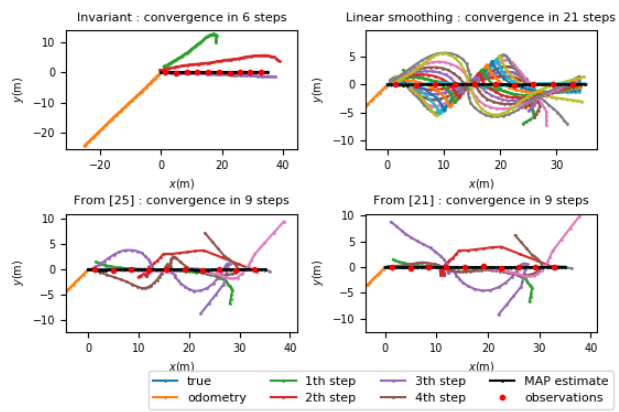

Fig. 2: Comparison of the iterations of the smoothing algorithms based on four parametrizations : invariant (top-left), linear (top-right), from [25] (bottom-left), from [21] (bottom-right). Thanks to its being linearized independently from the estimate, the invariant version converges faster and in a more "sensible" fashion.

D. Experimental setting

We compare the proposed method to various other methods based on data obtained in an experiment conducted at the Centre for Robotics, MINES ParisTech. A small wheeled robot, called Wifibot and photographed in Figure 3, is equipped with independent odometers on the left and right wheels. We made it follow an arbitrary trajectory for 80 seconds. The OptiTrack motion capture system, a set of seven highly precise cameras, provides the ground truth with sub-millimeter precision at a rate of 120 Hz. This choice allows us to directly compare our results with other recent algorithms based on the invariant framework, namely the Left-UKF-LG, which had already been tested on these data [15], the Invariant EKF, and both non-invariant smoothing methods from Section V-C. To keep the computational complexity acceptable, the smoothing is done in a sliding window, the oldest state being marginalized out when a new one is added. This choice makes Proposition 1 (iii) hold in this

case, leading to an information matrix fully independent from the estimate. The effect of the window's size is also studied here. The raw odometer inputs were provided to each method. Artificial position measurements are delivered by adding a Gaussian noise to the ground truth, at a rate of $1.35Hz$. For each setting, 100 Monte-Carlo simulations are run, each algorithm being initialized identically at each run. The trajectory followed by the robot, along with results of the various algorithms for one of the runs, are displayed in Figure 4. Three series of experiments were conducted, to study the evolution of the error with respect to the measurement covariance, the window size, and the initial error respectively. In the first case, the window size was fixed to 5, and the measurement covariance ranged from 10^{-1} to $10^{-5}m^2$, with noise turned on and off on the initial state. When on, $(\bar{x}_0, \bar{\theta}_0) \sim \mathcal{N}((0,0,0), \mathrm{diag}(1/8, 1/8, (\pi/4)^2)$. When off, it was put at $\bar{x}_0 = (1/4, 1/4)$ and $\bar{\theta}_0 = \pi/4$. In the second one, the measurement covariance was fixed to $10^{-1}m^2$, while the window size ranged from 5 to 13. For the last one, the window size was set to 5, the measurement covariance to $10^{-5}m^2$. The initial estimate was put at $\bar{x}_0 = (1/4, 1/4)$ and $\bar{\theta}_0$ spanned $]-\pi, \pi[$ with only one Gauss-Newton per iteration, then it was fixed at $\bar{\theta}_0 = 9\pi/10$ with one and seven GN iterations. Figures 5 and 6 show the averaged Root Mean Square Error (RMSE) of the position and heading for the two first experiments. Figure 7 shows the RMSE of the heading for the last one.

Fig. 3: Wifibot robot in its testing arena, surrounded by Optitrack cameras

Remark 1: Average computation times used to compute the full estimate, on a laptop with Intel i5-5300 2.3 GHz CPU, are given in the following table.

	Invariant	From [21]	From [25]	IEKF
Time (s)	0.82	0.81	0.83	0.62

It appears that the invariant smoothing, in a batch setting, does not induce extra computational load. Actually, it could prove more efficient in an incremental setting by taking advantage of the independence of the information matrix from the linearization points to prefactor the information matrix [28].

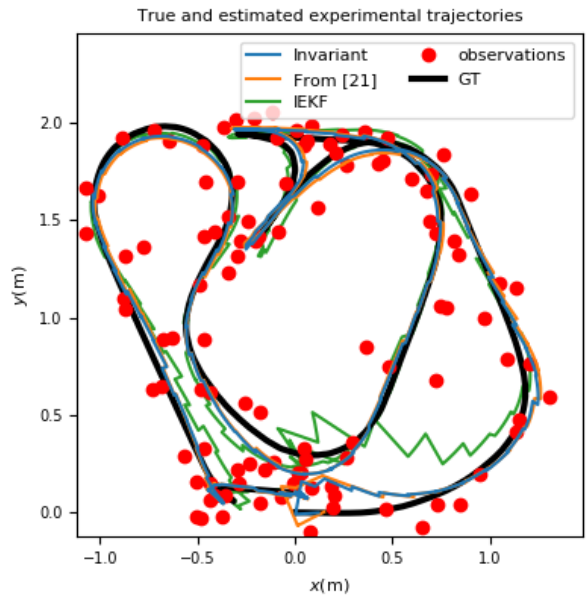

Fig. 4: Trajectories of the wifibot robot, as captured by the OptiTrack system and estimated by three of the studied methods : invariant smoothing, smoothing based on [21], both using a sliding window of size 5, and the IEKF, with artificial noisy measurements obtained by adding simulated moderate noise - $\sigma = 0.01(m/s)^2$ - to the ground truth position, showed by the red dots.

E. Experimental results

Figure 5 displays the evolution of the error with respect to the measurement noise covariance, when the initial state is fixed and when it is not. The first observation to be made is that, for a noise up to moderate ($10^{-2}m$), smoothing methods in general outperform the filtering ones, especially in terms of heading. This was expected, as the former explicitly contains the constraints between successive poses. If the initial state is fixed at $(1/4, 1/4, \pi/4)$, all smoothing approaches achieve almost identical results validating, in this case, the assumption made when deriving the propagation factor. However, invariant smoothing appeared to be the most robust to noise on the initial state, especially for low measurement covariance. This is further investigated in the third experiment.

The impact of the window size, illustrated in Figure 6, seems clear: increasing the size of the window is beneficial to all smoothing methods for strong noise. Still, none of them beats the UKF-LG in terms of position, highlighting its robustness. However, for lighter noise, the impact was negligible, as the odometry's uncertainty was quickly reached.

Finally, the invariant smoothing proved to be the more robust to initial heading errors such that $|\bar{\theta}_0| > \pi/2$, as shown by Figure 7, top, while both non invariant methods behaved identically. The two other graphs explain this difference : it appears that non-invariant smoothing methods fall into a local minimum where the estimate is completely turned around. Indeed, when increasing the number of GN iterations, the non-invariants experiments clearly form two clusters, one

with the expected heading RMSE, and one with a RMSE close to π. Note that it is not a problem of the non-invariant methods, as they were devised for problems relative to the original position, i.e. without such an influence of the initial error. Moreover, this confirms the results of Section V-C, as the invariant smoothing reaches the optimal RMSE quicker than the non-invariant ones, in the right cases.

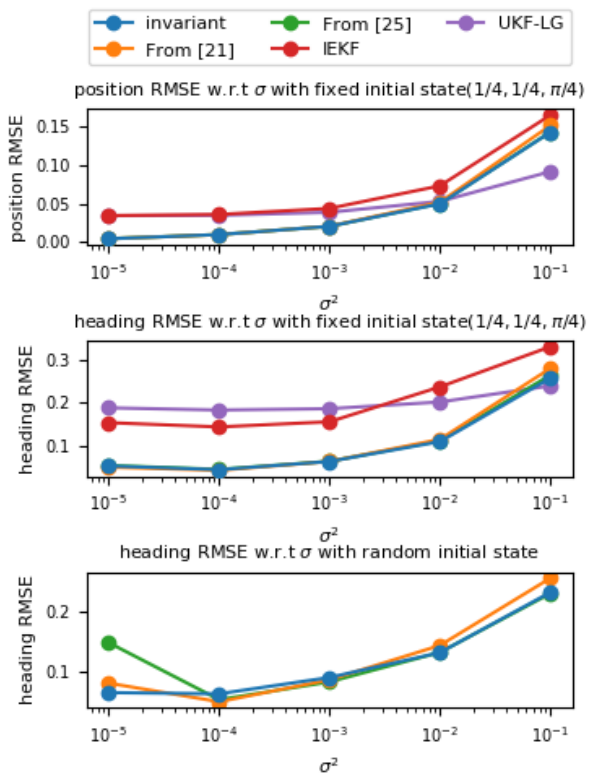

Fig. 5: Top and middle : average position and heading RMSE for the five estimation methods w.r.t σ, with fixed initial state. Down : heading RMSE w.r.t. σ, with random initial state. The window size was set to 5 for the smoothing methods.

F. Discussion

For this experiment, the proposed method managed to combine the advantages of the smoothing approach and of the invariant framework. Indeed, on the first hand it provided better estimates than the filtering methods thanks to the smoothing paradigm. On the other hand, owing to the mathematical invariances leveraged by the algorithm, their computation are based on the measured odometry inputs, and not the estimated ones. This reduces the need for relinearization, while ensuring more robustness to initial heading error, especially when non-invariant methods exhibited local minima which have no impact on the position estimate, but completely shift the heading one. However, this also means that the method will be dependent of the odometry's uncertainties.

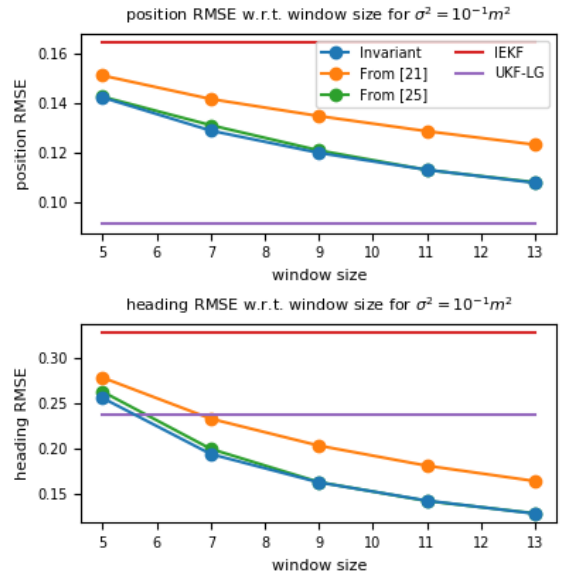

Fig. 6: Average position and heading RMSE for the smoothing methods w.r.t. the window size, for $\sigma^2 = 10^{-1}m^2$. The RMSE of the filtering methods are shown as a comparison.

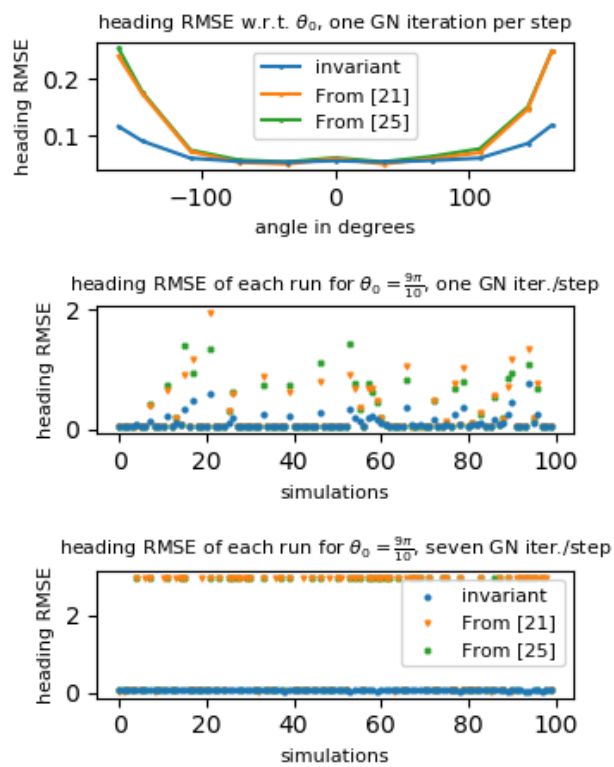

Fig. 7: Heading RMSE for the smoothing methods w.r.t the initial angle $\bar{\theta}_0$ for $\sigma^2 = 10^{-5}m^2$ and a window size of 5. Top : $\bar{\theta}_0$ spans $]-\pi,\pi]$ with one GN iteration per step. Middle : RMSE for each run with $\bar{\theta}_0 = 9\pi/10$ with one GN iteration. Down : RMSE for each run with $\bar{\theta}_0 = 9\pi/10$ with seven GN iterations.

VI. Conclusion

This paper presented the extension of the invariant Kalman filtering framework (see e.g., [9] for an overview) to smoothing. This resulted in Jacobians being independent from the current estimate, thus ensuring sound behavior of the estimate. To validate the results, this was applied to a robot localization problem, in simulations, and using experimental data. On the one hand, faster convergence to the MAP estimate than the other smoothing approaches was achieved in simulation. On the other hand, the invariant smoothing provided estimates similar to that of other smoothing methods, while being as fast to compute, and better than invariant filtering methods. It also appeared more robust to errors on the initial state, avoiding local minima of the non-invariant methods. Future work will include extending the framework to relative measurements of the state.

VII. Acknowledgements

The authors would like to thank Martin BARCZYK and Tony NOËL for their precious help with the experiments. This work is supported by the company Safran through the CIFRE convention 2016/1444.

Appendix A
Operators of the Special Euclidean 2D Group

The Special Euclidean 2D $SE(2)$ group represents rigid transformations in 2D space defined by their heading θ and position x. The exponential, logarithm and adjoint operators on $SE(2)$ are defined for $\chi \in SE(2)$ and $\xi \in \mathfrak{se}(2)$ as

$$\exp(\xi) = \begin{bmatrix} R(\xi^3) & \mathbf{V}_{\xi^3} \begin{bmatrix} \xi^1 \\ \xi^2 \end{bmatrix} \\ \mathbf{0} & 1 \end{bmatrix}, \log(\chi) = \begin{bmatrix} \mathbf{V}_\theta^{-1} x \\ \theta \end{bmatrix}, \quad (32)$$

$$\text{where } \mathbf{V}_\alpha = \frac{1}{\alpha} \begin{bmatrix} \sin(\alpha) & -1+\cos(\alpha) \\ 1-\cos(\alpha) & \sin(\alpha) \end{bmatrix} \quad (33)$$

$$Ad_\chi = \begin{bmatrix} R(\theta) & -Jx \\ \mathbf{0} & 1 \end{bmatrix}, \text{ where } J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (34)$$

References

[1] Pierre-Antoine Absil, Robert Mahony, and Rodolphe Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008.

[2] Victor Ayala and Juan Tirao. Linear control systems on lie groups and controllability. In *Proceedings of symposia in pure mathematics*, volume 64, pages 47–64. AMERICAN MATHEMATICAL SOCIETY, 1999.

[3] Timothy D. Barfoot. *State Estimation for Robotics*. Cambridge University Press, 2017.

[4] Timothy D. Barfoot and Paul. T. Furgale. Associating Uncertainty with Three-Dimensional Poses for Use in Estimation Problems. *IEEE Transactions on Robotics*, 30(3):679–693, 2014.

[5] Axel Barrau and Silvère Bonnabel. An EKF-SLAM algorithm with consistency properties. *CoRR*, abs/1510.06263, 2015.

[6] Axel Barrau and Silvère Bonnabel. Intrinsic filtering on lie groups with applications to attitude estimation. *IEEE Transactions on Automatic Control*, 60(2):436 – 449, 2015.

[7] Axel Barrau and Silvère Bonnabel. The invariant extended kalman filter as a stable observer. *IEEE Transactions on Automatic Control*, 62(4):1797–1812, April 2017.

[8] Axel Barrau and Silvere Bonnabel. Linear observation systems on groups (i). 2017.

[9] Axel Barrau and Silvère Bonnabel. Invariant kalman filtering. *Annual Review of Control, Robotics, and Autonomous Systems*, 1(1):null, 2018.

[10] Silvère Bonnabel, Philippe Martin, and Pierre Rouchon. Symmetry-Preserving Observers. *IEEE Transactions on Automatic Control*, 53(11):2514–2526, December 2008.

[11] Nicolas Boumal, Amit Singer, and Pierre-Antoine Absil. Robust estimation of rotations from relative measurements by maximum likelihood. In *52nd IEEE Conference on Decision and Control*, pages 1156–1161, Dec 2013.

[12] Guillaume Bourmaud. *Online Variational Bayesian Motion Averaging*, pages 126–142. Springer International Publishing, Cham, 2016.

[13] Guillaume Bourmaud, Rémi Mégret, Marc Arnaudon, and Audrey Giremus. Continuous-Discrete Extended Kalman Filter on Matrix Lie Groups Using Concentrated Gaussian Distributions. *Journal of Mathematical Imaging and Vision*, 51(1):209–228, 2015.

[14] Guillaume Bourmaud, Rémi Mégret, Audrey Giremus, and Yannick Berthoumieu. From intrinsic optimization to iterated extended kalman filtering on lie groups. *J. Math. Imaging Vis.*, 55(3):284–303, July 2016.

[15] Martin Brossard, Silvère Bonnabel, and J.P. Condomines. Unscented kalman filtering on lie groups. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2485–2491, Sept 2017.

[16] Gregory S. Chirikjian. *Stochastic Models, Information Theory, and Lie Groups, Volume 1: Classical Results and Geometric Methods*. Applied and numerical harmonic analysis. Birkhäuser, 2009.

[17] Gregory S Chirikjian. *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*. Springer Science & Business Media, 2011.

[18] Frank Dellaert and Michael Kaess. Square root sam: Simultaneous localization and mapping via square root information smoothing. *The International Journal of Robotics Research*, 25(12):1181–1203, 2006.

[19] Séebastien Diemer and Silvère Bonnabel. An Invariant Linear Quadratic Gaussian Controller for a Simplified Car. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pages 448–453. IEEE, 2015.

[20] Jay Farrell. *Aided Navigation: GPS with High Rate Sensors*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 2008.

[21] Christian Forster, Luca Carlone, Frank Dellaert, and Davide Scaramuzza. On-manifold preintegration for real-time visual–inertial odometry. *IEEE Transactions on Robotics*, 33(1):1–21, Feb 2017.

[22] Christian Forster, Zichao Zhang, Michael Gassner, Manuel Werlberger, and Davide Scaramuzza. Svo: Semidirect visual odometry for monocular and multicamera systems. *IEEE Transactions on Robotics*, 33(2):249–265, April 2017.

[23] J. Gallier. Notes on differential geometry and lie groups, 2016. University of Pennsylvannia.

[24] Giorgio Grisetti, Raimer Kümmerle, Cyrill Stachniss, Udo Frese, and Christoph Hertzberg. Hierarchical optimization on manifolds for online 2d and 3d mapping. In *2010 IEEE International Conference on Robotics and Automation*, pages 273–278, May 2010.

[25] Giorgio Grisetti, Ramier Kümmerle, Cyrill Stachniss, and Wolfram Burgard. A tutorial on graph-based slam. *IEEE Intelligent Transportation Systems Magazine*, 2(4):31–43, winter 2010.

[26] Christoph Hertzberg, René Wagner, Udo Frese, and Schröder Lutz. Integrating Generic Sensor Fusion Algorithms with Sound State Representations Through Rncapsulation of Manifolds. *Information Fusion*, 14(1):57–77, 2013.

[27] Vadim Indelman, Stephen Williams, Michael Kaess, and Frank Dellaert. Factor graph based incremental smoothing in inertial navigation systems. In *Information Fusion (FUSION), 2012 15th International Conference on*, pages 2154–2161. IEEE, 2012.

[28] Michael Kaess, Hordur Johannsson, Richard Roberts, Viorela Ila, John J. Leonard, and Frank Dellaert. iSAM2: Incremental smoothing and mapping using the Bayes tree. *The International Journal of Robotics Research*, 31:217–236, 2012.

[29] Rainer Kümmerle, Giorgio Grisetti, Hauke Strasdat, Kurt Konolige, and Wolfram Burgard. g2o: A general framework for graph optimization. In *2011 IEEE International Conference on Robotics and Automation*, pages 3607–3613, May 2011.

[30] Robert E. Mahony and Tarek Hamel. A geometric nonlinear observer for simultaneous localisation and mapping. In *56th IEEE Annual Conference on Decision and Control, CDC 2017, Melbourne, Australia, December 12-15, 2017*, pages 2408–2415, 2017.

[31] Daniel Martinec and Tomáš Pajdla. Robust rotation and translation estimation in multiview reconstruction. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2007.

[32] David M. Rosen, Luca Carlone, Afonso S. Bandeira, and John J. Leonard. Se-sync: A certifiably correct algorithm for synchronization over the special euclidean group. *CoRR*, abs/1612.07386, 2016.

[33] David M. Rosen, Charles DuHadway, and John J. Leonard. A convex relaxation for approximate global optimization in simultaneous localization and mapping. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5822–5829, May 2015.

[34] David M. Rosen, Michael Kaess, and John J. Leonard. Rise: An incremental trust-region method for robust online sparse least-squares estimation. 30(5):1091–1108, October 2014.

[35] Kevin C. Wolfe, Michael Mashner, and Gregory S. Chirikjian. Bayesian Fusion on Lie groups. *Journal of Algebraic Statistics*, 2(1):75–97, 2011.

[36] Sheng Zhao, Yiming Chen, Haiyu Zhang, and Jay A. Farrell. Differential gps aided inertial navigation: a contemplative realtime approach. *IFAC Proceedings Volumes*, 47(3):8959 – 8964, 2014. 19th IFAC World Congress.