An Homogeneous Space Geometry for Simultaneous Localisation and Mapping

Robert Mahony[a], Tarek Hamel[b] and Jochen Trumpf[c]

ARTICLE INFO

Keywords:
Simultaneous Localisation and Mapping
Equivariant Systems Theory
Observer
Lie Group

ABSTRACT

Simultaneous Localisation and Mapping (SLAM) is the archetypal chicken and egg problem: Localisation of a robot with respect to a map requires an estimate of the map, while mapping an environment from data acquired by a robot requires an estimate of the robot localisation. The nonlinearity and co-dependence of the SLAM problem has made it an ongoing research problem for more than thirty years. The present paper details recent advances in understanding the SLAM problem, specifically the existence of an underlying geometry and symmetry structure that provides significant insight into the difficulties that have plagued many SLAM algorithms. To demonstrate the power of the geometric insight we derive a constant gain observer for the SLAM problem that; that does not depend on linearisation, has globally asymptotically stable error dynamics, is very robust, and operates in dynamic environments (estimating the landmark velocities as states in the observer).

Contents

Robert.Mahony@anu.edu.au (R. Mahony); thamel@i3s.unice.fr (T. Hamel); jochen.trumpf@anu.edu.au. (J. Trumpf)
ORCID(s):

1. Introduction

Simultaneous Localisation and Mapping (SLAM) has been a core problem in robotics for the last thirty years (Stachniss, Thrun and Leonard (2016); Cadena, Carlone, Carrillo, Latif, Scaramuzza, Neira, Reid and Leonard (2016)) and remains one of the key technologies for automation in the 21st century. The problem involves estimation of an environment "map" (consisting of landmark feature point positions) while simultaneously estimating robot pose. The co-dependence of the map/pose estimation processes make the problem non-linear, difficult, and of considerable theoretical interest (Durrant-Whyte and Bailey (2006); Bailey and Durrant-Whyte (2006)). In comparison, if a robot's pose is known *a-priori* then estimating the map is straightforward (Thrun (2002)). If a map is available *a-priori*, the problem of pose-estimation is well understood in the robotics community (Scaramuzza and Fraundorfer (2011)), and has attracted attention in the non-linear observer community (Vasconcelos, Cunha, Silvestre and Oliveira (2010); Hua, Zamani, Trumpf, Mahony and Hamel (2011); Guerreiro, Batista, Silvestre and Oliveira (2013)).

The term Simultaneous Localisation and Mapping (SLAM) was coined in the "classical" period of SLAM research – lasting from the late 90's through to around 2010-2012 (Cadena et al. (2016)). This period was characterised by applications of extended Kalman and particle filtering techniques (Durrant-Whyte and Bailey (2006); Bailey and Durrant-Whyte (2006); Aulinas, Petillot, Salvi and Lladó. (2008)). Filtering approaches exploit the temporal nature of the problem to fuse measurements sequentially into an "information" state which is then updated as new data is obtained (Strasdat, Montiel and Davison (2012)). A key advantage is that only recent measurements are considered, minimizing the complexity of maintaining large data sets of measurements. Information from old measurements is captured in an information state, typically in the form of a covariance matrix associated with a Gaussian approximation of the information state distribution, that propagates according to a Ricatti equation. Classical filtering techniques suffer from a number of weaknesses: For large maps the computational complexity of the covariance update can become an issue (Durrant-Whyte and Bailey (2006); Bailey and Durrant-Whyte (2006); Aulinas et al. (2008); Strasdat et al. (2012)). More importantly, classical SLAM EKF filters encountered issues with stochastic consistency (Dissanayake, Huang, Wang and Ranasinghe. (2011)). In particular, when the linearisation point of the EKF is incorrect (as is always the case when the true state is unknown), the Ricatti equation does not correctly propagate the information state covariance estimate, introducing errors that degrade the filter performance. State-of-the-art EKF algorithms address these issues by using sliding windows of data, carefully choosing linearisation points, and working in sub-maps (Huang, Mourikis and Roumeliotis (2010, 2011); andJ.A. Hesch, Bowman and Roumeliotis. (2012); Hesch, Kottas, Bowman and Roumeliotis (2014)). These algorithms remain highly competitive in dynamic environments and for odometry problems but are less competitive for pure mapping problems. The modern era of SLAM – dating from around 2011 (Cadena et al. (2016)) – is based on optimization principles and was heavily influenced by bundle adjustment techniques drawn from the computer vision literature (Triggs, McLauchlan, Hartley and Fitzgibbon (2000)). This approach maintains 'all' the data and then optimizes a nonlinear least squares cost over the full time sequence to find the best estimates of landmark positions and robot pose trajectory (Grisetti, Kummerle, Stachniss and Burgard (2010)). The cost is linearised and graph factorization methods are exploited to guarantee sparsity of the Hessian in order that the resulting optimization problem is tractable (Ila, Porta and Andrade-Cetto (2010); Dellaert (2012); Kaess, Johannsson, Roberts, Ila, Leonard and Dellaert (2012)). Modern SLAM methods overcome the fundamental consistency problem that plagues classical SLAM by updating the linearisation point during the optimization. However, this additional performance comes at a cost. Every time the linearisation point is updated, all the (relevant) Hessians in the cost must be recomputed using all the relevant data. State-of-the-art optimization based methods (Kaess et al. (2012); Mur-Artal, Montiel and Tardós (2015); Engel J. (2014)) depend heavily on sparsity of data and this encodes a number of restrictions on scenarios that can easily be considered. For example, they are ill-suited to dynamic environments (Dayoub, Cielniak and Duckett (2011); Walcott-Bryant, Kaess, Johannsson and Leonard (2012); Krajn/'ıık, Fentanes, Mozos, Duckett, Ekekrantz and Hanheide (2014)) although recent work has been undertaken in this direction (Yang and Scherer (2019); Judd and Gammell (2019); Zhang, Henein, Mahony and Ila (2020)).

The non-linear observer community has contributed strongly to state-estimation problems in robotics in the last ten years. Nonlinear observers for attitude estimation played a key role in aerial robotics (Mahony, Hamel and Pflimlin (2008); Bonnabel, Martin and Rouchon (2008); Vasconcelos, Silvestre and Oliveira (2008)). Full pose estimation has been studied (Vik and Fossen (2001); Baldwin, Mahony and Trumpf (2009); Vasconcelos et al. (2010)) and is used in head tracking systems amongst other applications. The SLAM problem has been considered by a number of authors: Direct application of geometric non-linear observer methods was undertaken by Zlotnik and Forbes (2018). An alternative approach taken by Johansen and Brekke (2016); Bjorne, Johansen and Brekke (2017) is to use other sensors to obtain a rotation estimate and then apply Kalman filtering to the linear map-position estimation problem. When the rotation part of robot pose is assumed known then it is also possible to consider bearing only measurements (Hamel and Samson (2016); Le Bras, Hamel, Mahony and Samson (2017); Hamel and Samson (2018)). A more complete approach to the SLAM problem was presented in recent work by Guerreiro et al. (2013) and Lourenço, Guerreiro, Batista, Oliveira and Silvestre (2016) who consider a "robo-centric" nonlinear observer that estimates environment points expressed in the body-fixed-frame of the robot. Robo-centric and relative SLAM formulations have been considered in the robotics community (Castellanos, Martinez-Cantin, Tardós and Neira (2007); Williams and Reid (2010); Mei, Sibley, Cummins, Newman and Reid (2011)), however, mostly in the context of addressing consistency issues in EKF based methods rather than studying robustness and asymptotic stability (Guerreiro et al. (2013); Lourenço et al. (2016)). A robot-centric representation, however, requires propagation of the ego-motion of the landmark points, introducing additional error into the map particularly in the case of significant rotational motion of the robot, or requires the storing of the map in reference coordinates anyway. Recent work by Barrau and Bonnabel (2016) proposes a Lie-group structure $\mathbf{SE}_{n+1}(3)$ for the SLAM configuration state. The group structure is used to apply the Invariant Kalman Filter algorithm (Bonnabel (2007); Bonnabel, Martin and Salaun (2009); Barrau and Bonnabel (2017)) and the resulting filter overcomes the consistency issues that plagued the EKF algorithms from the classical SLAM era. This work is starting to have impact back in the robotics community (Zhang, Wu, Song, Huang and Dissanayake (2017)).

In this paper, we present a highly robust, simple, and computationally cheap nonlinear observer for the general landmark SLAM problem. The approach is based on a novel formulation of the SLAM state-space as a principal (fibre) bundle of landmark and pose configurations that we term the SLAM manifold. The inherent gauge transform invariance of the SLAM problem (Kanatani and Morris (2001)) forms the fibres of the quotient manifold structure. The Lie group $\mathbf{SE}_{n+1}(3)$ first proposed by Barrau and Bonnabel (2016, 2017) is shown to act transitively on the SLAM manifold and the intrinsic kinematics of robot and landmarks are shown to be equivariant with respect to this action. The authors believe that the SLAM manifold will provide an elegant geometric model in which to formulate and understand robo-centric coordinates, gauge invariance, as well as issues with consistency of EKF SLAM and indefiniteness of the cost Hessian in optimisation based methods. The symmetry action of $\mathbf{SE}_{n+1}(3)$ allows direct application of the authors' previous work (Mahony, Trumpf and Hamel (2013); Mahony, Hamel and Trumpf (2020)) to yield a novel constant gain observer for continuous-time SLAM. The formulation naturally allows for a dynamic environment and we include a simple integral estimator in the observer that estimates constant velocities of moving landmark features. The proposed algorithm is fully non-linear and no linearisation is required. Theorem 5.1 proves global asymptotic stability and local exponential stability of the error coordinates. The inherent symmetry of the approach ensures high levels of robustness to non-Gaussian disturbances such as data association errors (mislabeling of feature point correspondences between image frames). These errors are common in real world systems (Tombari, Salti and Di Stefano (2013)) where feature detection and data association algorithms that identify and remove outliers (Tombari et al. (2013); Chin, Kee, Eriksson and Neumann (2016); Bustos and Chin (2017)) consume a major part of the computational resources required for modern SLAM algorithms. We provide evidence through a simulation study that the proposed algorithm outperforms classical extended Kalman filter algorithms in both robustness and performance in the presence of data association errors. Moreover, the computational cost (both CPU and memory requirements) of the algorithm is far lower than for a classical EKF based SLAM algorithm. The target applications for the proposed algorithm are in IoT (Internet of Things) applications and small mobile robotic systems that require local spatial awareness and have limited compute available. In such situations, there is a sweet spot for a low complexity SLAM algorithm that is highly robust and provides good enough map and pose estimates.

The paper is organised in seven sections including this introduction. In Section 2 we introduce notation and definitions. In Section 3 we introduce the SLAM manifold and develop the necessary theory to define the kinematics of the system. In Section 4 we define the symmetry group actions and demonstrate that the SLAM kinematics are equivariant. Section 5 introduces the observer design and proves the main result (Theorem 5.1); local exponential and global asymptotic stability of the observer error. Section 6 provides simulations demonstrating the functionality, robustness and computational advantages of the proposed algorithm. Concluding remarks are provided in Section 7.

2. Notation

We will make extensive use of smooth Lie-group actions α, φ, ψ, ρ. A 'right' action $\beta : \mathbf{G} \times \mathcal{M} \to \mathcal{M}$ of a Lie group $\mathbf{G}$ on a smooth manifold $\mathcal{M}$ is a smooth mapping with properties

$$\beta(A, \beta(B, \xi)) = \beta(B \cdot A, \xi), \qquad \beta(\mathrm{id}, \xi) = \xi$$

for all $A, B \in \mathbf{G}$ and $\xi \in \mathcal{M}$, where id $\in \mathbf{G}$ is the identity element and '$\cdot$' is the group operation. We use the partial map notation $\beta_A : \mathcal{M} \to \mathcal{M}$ and $\beta_\xi : \mathbf{G} \to \mathcal{M}$ where

$$\beta_A(\xi) := \beta(A, \xi) =: \beta_\xi(A)$$

for $A \in \mathbf{G}$ and $\xi \in \mathcal{M}$. For a matrix group, left and right actions $L, R : \mathbf{G} \times \mathbf{G} \to \mathbf{G}$ are defined by

$$L_A P = AP, \qquad R_A P = PA,$$

where the righthand side of the equations is matrix multiplication. Denote the Lie-algebra of $\mathbf{G}$ by $\mathfrak{g}$ and identify $\mathfrak{g}$ with $T_{\mathrm{id}}\mathbf{G}$. For a matrix group, $\mathfrak{g}$ is a linear matrix space closed under the matrix Lie-bracket $[U, V] = UV - VU$. In this case, the identity element id is the identity matrix I and the differentials $\mathrm{d}L_A, \mathrm{d}R_A : T\mathbf{G} \to T\mathbf{G}$ evaluated at the identity are

$$\mathrm{d}L_A(I)V = AV, \qquad \mathrm{d}R_A(I)V = VA.$$

The adjoint map is $\mathrm{Ad}_A V = AVA^{-1}$.

A point $p \in \mathbb{R}^3$ in Euclidean space has coordinates (p_1, p_2, p_3) representing the position of p with respect to a reference frame $\{0\}$. We use homogeneous coordinates

$$\overline{p} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ 1 \end{pmatrix}$$

and write ${}^0\overline{p}$ when the reference frame is not clear from context. We abuse notation to write $\overline{p} \in \mathbb{R}^3$ or ${}^0\overline{p} \in \mathbb{R}^3$ to indicate that the underlying point $p \in \mathbb{R}^3$.

We use the notation $\mathbf{R}^3$ to denote elements of the real 3-dimensional additive group and we distinguish between $\mathbf{R}^3$ and Euclidean space $\mathbb{R}^3$. Elements $a \in \mathbf{R}^3$ are used to represent translations of Euclidean space. We use the homogeneous "free" vector notation[1]

$$\mathring{\overline{a}} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ 0 \end{pmatrix}$$

and allow ourselves to write $\mathring{\overline{a}} \in \mathbf{R}^3$ to indicate that the underlying $a \in \mathbf{R}^3$. The "bar-circle" notation ($\mathring{\overline{a}}$) is similar to the homogeneous coordinates "bar" notation ($\overline{p}$) except that the fourth entry of the vector is zero rather than one. Exploiting this notation, we express the action of the additive group $\mathbf{R}^3$ in translating Euclidean space $\mathbb{R}^3$ as a coordinate addition

$$\overline{p} + \mathring{\overline{a}} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ 1 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ 0 \end{pmatrix} = \begin{pmatrix} p_1 + a_1 \\ p_2 + a_2 \\ p_3 + a_3 \\ 1 \end{pmatrix} = \overline{(p + a)}. \tag{1}$$

This action is a commutative (both left and right) group action of $\mathbf{R}^3$ on $\mathbb{R}^3$:

$$(\overline{p} + \mathring{\overline{a}}) + \mathring{\overline{b}} = \overline{p} + (\mathring{\overline{a}} + \mathring{\overline{b}}) = (\overline{p} + \mathring{\overline{b}}) + \mathring{\overline{a}}, \tag{2a}$$

$$\overline{p} + \mathring{\overline{0}} = \overline{p}. \tag{2b}$$

Note that the group $\mathbf{R}^3$ is not 'physical' and hence there is no concept of reference frame or coordinates for elements $a \in \mathbf{R}^3$. Elements of the additive group $\mathbf{R}^3$ can be added and subtracted in the natural manner corresponding to Abelian vector addition. One may never add two elements of $\mathbb{R}^3$, $\overline{p} + \overline{q} \neq \overline{(p + q)}$, however, the difference between

[1] We will use the homogeneous "free" vector notation for both additive group elements and linear velocities of points in $\mathbb{R}^3$.

two elements of $\mathbb{R}^3$, $\overline{p} - \overline{q} = \overset{\circ}{\overline{(p - q)}}$ is well defined as the unique element $(p - q) \in \mathbf{R}^3$ of the additive group such that $\overline{p} = \overline{q} + \overset{\circ}{\overline{(p - q)}}$.

The special Euclidean group $\mathbf{SE}(3)$ is the set of rigid-body transformations of $\mathbb{R}^3$. For $A \in \mathbf{SE}(3)$ we use the notation $R_A \in \mathbf{SO}(3)$ and $x_A \in \mathbf{R}^3$ to denote the rotation and translation components respectively. The classical (homogeneous) matrix representation of A is

$$A = \begin{pmatrix} R_A & x_A \\ 0 & 1 \end{pmatrix}, \tag{3}$$

and the rigid-body transformation on homogeneous coordinates is matrix multiplication $\overline{q} = A\overline{p} = \overline{R_A p + x_A}$. One has $R_{AB} = R_A R_B$, and $R_{A^{-1}} = R_A^{-1} = R_A^\top$. A key property of the action (2) that we exploit is

$$A(\overline{p} + \overset{\circ}{\overline{a}}) = A\overline{p} + A\overset{\circ}{\overline{a}} = A\overline{p} + \overset{\circ}{\overline{R_A a}}. \tag{4}$$

The pose of a vehicle moving in Euclidean space is represented by a moving frame: that is, a location in space and a set of orthonormal axes directions. We use the frame bundle notation $\mathcal{F}(\mathrm{T}\mathbb{R}^3)$ to denote the set of all poses[2] and distinguish strongly between poses in the frame bundle and elements of the special Euclidean transformation group. Coordinates for a pose are given by the location $x_P \in \mathbb{R}^3$ of the frame with coordinates relative to a reference frame and an orthonormal matrix R_P that encodes the relative orientation of the frame axes relative to the same reference frame. We use the notation $P \in \mathcal{F}(\mathrm{T}\mathbb{R}^3)$ to denote the homogeneous matrix constructed according to (3) from x_P and R_P and thought of as coordinates for a robot pose. For a reference frame $\{0\}$ and a moving frame $\{P\}$ we write the homogeneous coordinates of the same physical point $p \in \mathbb{R}^3$ in Euclidean space as ${}^0\overline{p}$ and ${}^P\overline{p}$, respectively, where the top left index indicates the coordinates used. The coordinate transform that preserves the physical location of a point in Euclidean space $\mathbb{R}^3$ is given by the homogeneous matrix coordinates of the pose $P \in \mathcal{F}(\mathrm{T}\mathbb{R}^3)$, that is

$$P\,{}^P\overline{p} = \overline{R_P\,{}^Pp + x_P} = {}^0\overline{p}.$$

The velocity of a moving point $\overline{p} \in \mathbb{R}^3$, relative to a (stationary) reference frame ($\overline{p} = {}^0\overline{p}$), can be written as a homogeneous free vector

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{p} = \begin{pmatrix} \dot{p} \\ 0 \end{pmatrix} = \overset{\circ}{\overline{(\frac{\mathrm{d}}{\mathrm{d}t}p)}}. \tag{5}$$

Consider a landmark point $\overline{p}_i = {}^0\overline{p}_i \in \mathbb{R}^3$ written in homogeneous reference frame coordinates. We denote its velocity, with respect to the (stationary) reference frame, but expressed in the body-fixed frame $\{P\}$ by $v_i \in \mathbb{R}^3$. The coordinate transform P operates as a rotation on free homogeneous vectors and thus the homogeneous reference frame coordinates for the landmark velocity are given by ${}^0\overset{\circ}{\overline{v}}_i = P\overset{\circ}{\overline{v}}_i = \overset{\circ}{\overline{R_P v_i}}$. Thus,

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{p}_i = P\overset{\circ}{\overline{v}}_i. \tag{6}$$

For a static environment then $\overset{\circ}{\overline{v}}_i = 0$.

The velocity of the robot used is the relative motion of the robot frame with respect to the reference frame expressed in the body-fixed frame. This formulation is the most common encountered in the SLAM literature and is associated with robots using onboard velocity sensors. The linear velocity of x_P is denoted $v \in \mathbb{R}^3$ and the angular velocity of R_P is denoted $\Omega \in \mathbb{R}^3$.

The kinematics of the robot pose are then given by

$$\dot{x}_P = R_P v, \qquad \dot{R}_P = R_P \Omega^\times \tag{7}$$

where

$$\Omega^\times = \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{pmatrix}.$$

[2] The notation $\mathcal{F}(\mathrm{T}\mathbb{R}^3)$ denotes the oriented orthonormal frame bundle associated with the tangent bundle $\mathrm{T}\mathbb{R}^3$. That is the set of all possible frames located at any point in $\mathbb{R}^3$ and oriented arbitrarily corresponding to arbitrary pose of a vehicle.

One has that $\Omega^\times w = \Omega \times w$ for any $w \in \mathbb{R}^3$, where $\times$ denotes the vector product. Equation (7) can be written compactly in homogeneous matrix notation

$$\dot{P} = PV, \tag{8}$$

where

$$V = (\Omega, v)^\wedge := \begin{pmatrix} \Omega^\times & v \\ 0 & 0 \end{pmatrix}. \tag{9}$$

The velocity V can be thought of as an element of the matrix Lie algebra $\mathfrak{se}(3)$ associated with the special Euclidean group, although only the vector space structure of $\mathfrak{se}(3)$ is used in modeling the instantaneous velocity of a pose. The tangent space of the frame bundle $\mathcal{F}(\mathrm{T}\mathbb{R}^3)$ at a point P is given by

$$T_P\mathcal{F}(\mathrm{T}\mathbb{R}^3) = \{PV \mid V = (\Omega, v)^\wedge \in \mathfrak{se}(3)\}.$$

Let $\mathbb{P}_{\mathfrak{se}} : \mathbb{R}^{4\times4} \to \mathfrak{se}(3)$ denote the unique orthogonal projection of $\mathbb{R}^{4\times4}$ onto $\mathfrak{se}(3)$ with respect to the Frobenius inner product $\langle\langle A, B\rangle\rangle = \mathrm{tr}\left(A^\top B\right)$ on $\mathbb{R}^{4\times4}$. That is for all $V \in \mathfrak{se}(3)$, $M \in \mathbb{R}^{4\times4}$, one has

$$\langle\langle V, M\rangle\rangle = \langle\langle V, \mathbb{P}_{\mathfrak{se}}(M)\rangle\rangle = \langle\langle \mathbb{P}_{\mathfrak{se}}(M), V\rangle\rangle .$$

One has that for all $M_1 \in \mathbb{R}^{3\times3}$, $m_2, m_3 \in \mathbb{R}^3$, and $m_4 \in \mathbb{R}$,

$$\mathbb{P}_{\mathfrak{se}}\left(\begin{bmatrix} M_1 & m_2 \\ m_3^\top & m_4 \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{2}(M_1 - M_1^\top) & m_2 \\ 0 & 0 \end{bmatrix}. \tag{10}$$

3. SLAM Manifold

A key geometric structure in SLAM problems is a principal (fibre) bundle that we term the SLAM manifold. This manifold provides a geometric model for the configuration state associated with simultaneous estimation of landmarks and robot pose.

3.1. Problem formulation

We begin by defining *raw* or *total space* coordinates for the SLAM problem using an unknown but fixed reference frame $\{0\}$. Let $P \in \mathcal{F}(\mathrm{T}\mathbb{R}^3)$ represent the body-fixed frame coordinates of the robot with respect to this reference frame. Let

$$\overline{p}_i \in \mathbb{R}^3, \qquad i = 1, \dots, n,$$

be the positions of sparse point features in the environment expressed with respect to the reference frame $\{0\}$. The *total space* of the SLAM problem is the product space

$$\mathcal{T}_n(3) = \mathcal{F}(\mathrm{T}\mathbb{R}^3) \times \mathbb{R}^3 \times \cdots \times \mathbb{R}^3 \tag{11}$$

made up of elements with raw homogeneous coordinates $(P, \overline{p}_1, \dots, \overline{p}_n)$. The subscript n denotes the number of landmarks while the number in parentheses denotes the dimension of Euclidean space with which the total space is associated. The total space is the usual state-space used in classical SLAM algorithms. We emphasise that unlike other recent work (Barrau and Bonnabel (2016); Zlotnik and Forbes (2018)) we do not try to model the state of the SLAM problem as a Lie-group.

The tangent space of $\mathcal{T}_n(3)$ at a point $\Xi = (P, \overline{p}_1, \dots, \overline{p}_n)$ can be identified with the matrix subspace

$$T_\Xi\mathcal{T}_n(3) = \{(PV, P\mathring{v}_1, \dots, P\mathring{v}_n) \,\big|\, V \in \mathfrak{se}(3), v_i \in \mathbb{R}^3\}.$$

That is we use the embedded nature of the total coordinates to define $T_\Xi\mathcal{T}_n(3) \subset \mathbb{R}^{4\times4} \times \mathbb{R}^4 \times \cdots \times \mathbb{R}^4$ as a matrix vector subspace.

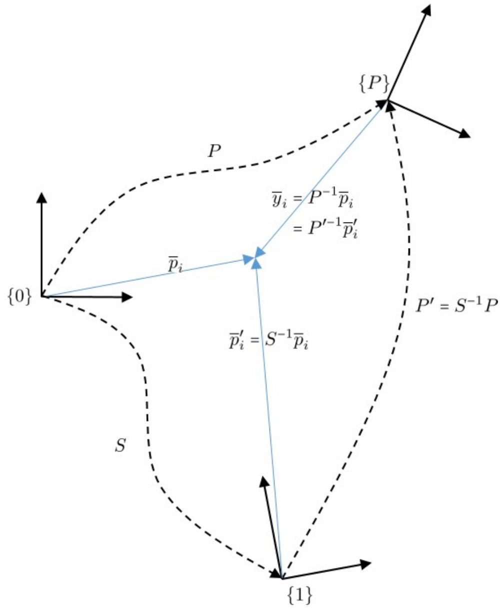

Figure 1: The raw coordinates $P \in \mathcal{F}(\mathrm{T}\mathbb{R}^3)$ and $\overline{p}_i \in \mathbb{R}^3$ are measured relative to a reference frame $\{0\}$. An arbitrary $\mathbf{SE}(3)$ gauge transformation S transforms $\{0\}$ to a new reference $\{1\}$ and new coordinates $P' \in \mathcal{F}(\mathrm{T}\mathbb{R}^3)$ and $\overline{p}'_i \in \mathbb{R}^3$ describe the robot and environment. Note that the output $\overline{y}_i$ does not change and refers to the same physical location as $\overline{p}_i$ and $\overline{p}'_i$.

3.2. Total space kinematics

Recalling (8) and (6) the velocity measurements of the SLAM problem consist of a rigid-body velocity $V \in \mathfrak{se}(3)$ and n landmark velocities $v_i \in \mathbb{R}^3$. These velocities are expressed in coordinates of the body-fixed frame $\{P\}$, and represent the relative motion of the robot and landmarks with respect to a reference frame $\{0\}$. Define

$$\mathbb{V} = \{(V, \vec{v}_1, \dots, \vec{v}_n) \mid V \in \mathfrak{se}(3), v_i \in \mathbb{R}^3\} \tag{12}$$

and note that $\mathbb{V}$ inherits a natural linear vector space structure from the product of the underlying structures on $\mathfrak{se}(3)$ and $\mathbb{R}^3$.

Define a function $f : \mathcal{T}_n(3) \times \mathbb{V} \to T\mathcal{T}_n(3)$:

$$f((P, \overline{p}_1, \dots, \overline{p}_n), (V, \vec{v}_1, \dots, \vec{v}_n)) := (PV, P\vec{v}_1, \dots, P\vec{v}_n), \tag{13}$$

where the matrix multiplication notation $P\vec{v}_i$ transforms $\vec{v}_i$ from body-fixed frame coordinates to reference frame coordinates analogously to (6). The function f encodes the natural kinematics of the SLAM problem in the raw coordinates. That is for time-varying coordinates $\Xi(t) = (P(t), \overline{p}_1(t), \dots, \overline{p}_n(t)) \in \mathcal{T}_n(3)$ then the ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\Xi = f(\Xi, (V, \vec{v}_1, \dots, \vec{v}_n)) = (PV, P\vec{v}_1, \dots, P\vec{v}_n) \tag{14}$$

evolves according to the physical structure described in Section 2. Stationary landmarks will have zero velocity, $\vec{v}_i \equiv 0$.

3.3. Gauge transform invariance

The raw coordinates of the SLAM problem are not intrinsic since they depend on the arbitrary choice of reference frame $\{0\}$. Indeed, one may consider any *gauge* (Kanatani and Morris (2001)) transformation $S \in \mathbf{SE}(3)$ of frame $\{0\}$ to a new reference frame $\{1\}$ (as shown in Figure 1) and generate new 'raw' coordinates $(S^{-1}P, S^{-1}\overline{p}_1, \dots, S^{-1}\overline{p}_n)$ with respect to frame $\{1\}$ that represent the same SLAM configuration.

Lemma 3.1. *Consider the map $\alpha : \mathbf{SE}(3) \times \mathcal{T}_n(3) \to \mathcal{T}_n(3)$ defined by*

$$\alpha(S, (P, \overline{p}_1, \dots, \overline{p}_n)) := (S^{-1}P, S^{-1}\overline{p}_1, \dots, S^{-1}\overline{p}_n). \tag{15}$$

This map defines a smooth, proper and free right group action of $\mathbf{SE}(3)$ *on* $\mathcal{T}_n(3)$. *The quotient (see Definition 3.3)*

$$\mathcal{M}_n(3) = \mathcal{T}_n(3)/\alpha \tag{16}$$

is a smooth manifold of dimension $3n$.

Proof 3.2. *One computes*

$$\begin{aligned}\alpha(S_1, \alpha(S_2, (P, \overline{p}_1, \dots, \overline{p}_n))) &= \alpha(S_1, (S_2^{-1}P, S_2^{-1}\overline{p}_1, \dots, S_2^{-1}\overline{p}_n)) \\ &= (S_1^{-1}S_2^{-1}P, S_1^{-1}S_2^{-1}\overline{p}_1, \dots, S_1^{-1}S_2^{-1}\overline{p}_n) \\ &= \alpha(S_2S_1, (P, \overline{p}_1, \dots, \overline{p}_n))\end{aligned}$$

to demonstrate the right action property. The identity relationship $\alpha(I_4, (P, \overline{p}_1, \dots, \overline{p}_n)) = (P, \overline{p}_1, \dots, \overline{p}_n)$ *is straightforward to verify and it follows that* α *is a group action.*

The action α *is smooth by construction. To see that* α *is free, note that the stabiliser of any point is the trivial group. To see that* α *is proper we note that it is an algebraic map and is both open and continuous. The quotient* $\mathcal{M}_n(3)$ *is a well defined smooth manifold since the group action is both free and proper (Steenrod (1951)). Indeed,* $\mathcal{M}$ *is a principal* $\mathbf{SE}(3)$*-bundle. The dimension of* $\mathcal{M}_n(3)$ *is obtained by a dimension count,* $\dim \mathcal{M}_n(3) = \dim \mathcal{T}_n(3) - \dim \mathbf{SE}(3) = 3n + 6 - 6 = 3n$.

Definition 3.3. *The group action* α (15) *is termed the* gauge *action and an element* $(P, \overline{p}_1, \dots, \overline{p}_n) \in \mathcal{T}_n(3)$ *defines a gauge equivalence class*

$$[P, \overline{p}_1, \dots, \overline{p}_n] = \left\{ (S^{-1}P, S^{-1}\overline{p}_1, \dots, S^{-1}\overline{p}_n) \,\middle|\, S \in \mathbf{SE}(3) \right\} \tag{17}$$

obtained by applying $\alpha(S, \cdot)$ *to* $(P, \overline{p}_1, \dots, \overline{p}_n)$ *for all elements* $S \in \mathbf{SE}(3)$. *An equivalence class* $[P, \overline{p}_1, \dots, \overline{p}_n]$ *is termed a (SLAM) configuration. The quotient manifold* (16) $\mathcal{M}_n(3) = \mathcal{T}_n(3)/\alpha$ *consists of the set of all configurations*

$$\mathcal{M}_n(3) = \left\{ [P, \overline{p}_1, \dots, \overline{p}_n] \,\middle|\, (P, \overline{p}_1, \dots, \overline{p}_n) \in \mathcal{T}_n(3) \right\}.$$

This manifold is termed the SLAM manifold. *We will use* $\xi = [P, \overline{p}_1, \dots, \overline{p}_n] \in \mathcal{M}_n(3)$ *where compressed notation is appropriate. We will also use the notation* $\pi : \mathcal{T}_n(3) \to \mathcal{M}_n(3)$,

$$\pi(P, \overline{p}_1, \dots, \overline{p}_n) := [P, \overline{p}_1, \dots, \overline{p}_n] \tag{18}$$

for the intrinsic projection associated with the quotient structure of $\mathcal{M}_n(3)$. □

The key advantage of this formulation is that the ambiguity associated with specification of the reference frame is not present for configurations in the SLAM manifold. In particular, since a configuration is an equivalence class $\xi = [P, \overline{p}_1, \dots, \overline{p}_n]$, then for any choice of raw coordinates $(P', \overline{p}'_1, \dots, \overline{p}'_n) = (S^{-1}P, S^{-1}\overline{p}_1, \dots, S^{-1}\overline{p}_n)$ related by rigid-body transformation of the reference frame, the associated configuration $\xi = [S^{-1}P, S^{-1}\overline{p}_1, \dots, S^{-1}\overline{p}_n]$ is the *same* element of $\mathcal{M}_n(3)$.

Remark 3.4. *(Robo-centric coordinates) It is possible to globally define a coordinate chart* $h : \mathcal{M}_n(3) \to \mathbb{R}^{3n}$ *by*

$$h([P, \overline{p}_1, \dots, \overline{p}_n]) := (P^{-1}\overline{p}_1, \dots, P^{-1}\overline{p}_n).$$

It is straightforward to verify that this map is well defined on $\mathcal{M}_n(3)$ *since for any* $S \in \mathbf{SE}(3)$

$$\begin{aligned}h([\alpha(S, (P, \overline{p}_1, \dots, \overline{p}_n))] &= (P^{-1}SS^{-1}\overline{p}_1, \dots, P^{-1}SS^{-1}\overline{p}_n) \\ &= (P^{-1}\overline{p}_1, \dots, P^{-1}\overline{p}_n) \\ &= h([P, \overline{p}_1, \dots, \overline{p}_n]).\end{aligned} \tag{19}$$

Furthermore, h *is full rank and defines a coordinate chart into* $\mathbb{R}^{3n}$. *The coordinates defined in this manner correspond to the landmark positions expressed in the body-fixed-frame. These coordinates are the robo-centric coordinates that* *have been considered in prior work (Castellanos et al. (2007); Williams and Reid (2010); Guerreiro et al. (2013); Lourenço et al. (2016)).*

In practice, robocentric local coordinates are quite nonlinear with respect to the underlying SLAM optimisation and although classical EKF solutions have been developed (Castellanos et al. (2007); Williams and Reid (2010)) they were not found to be competitive with an EKF formulated explicitly on the total space coordinates. The quotient structure of the SLAM manifold provides a geometric link between robo-centric coordinates (a nonlinear coordinate chart for $\mathcal{M}_n(3)$) and the raw coordinates (total space coordinates for $\mathcal{M}_n(3)$). The authors believe that algorithms developed in the SLAM manifold geometry but implemented in total space coordinates will inherit the advantages of the robo-centric and relative formulations (Castellanos et al. (2007); Williams and Reid (2010); Mei et al. (2011)) while retaining the advantages of classical SLAM formulation.

3.4. Tangent Space of $\mathcal{M}_n(3)$

The tangent space of $\mathcal{M}_n(3)$ at a point $\xi = \pi(\Xi)$ is formally a quotient of the tangent space of the total space

$$T_{\pi(\Xi)}\mathcal{M}_n(3) := T_\Xi \mathcal{T}_n(3)/\ker \mathrm{d}\pi_\Xi$$

where $\mathrm{d}\pi_\Xi : T_\Xi\mathcal{T}_n(3) \to T_\xi\mathcal{M}_n(3)$ is the differential of π at the point Ξ and $\ker \mathrm{d}\pi_\Xi$ is the kernel of the linear map $\mathrm{d}\pi_\Xi$ (Absil, Mahony and Sepulchre (2008)).

To make this explicit, fix $\Xi_0 = (P, \overline{p}_1, \dots, \overline{p}_n) \in \mathcal{T}_n(3)$ constant and consider a curve in $\mathcal{T}_n(3)$ constructed by only allowing variation in the fibre,

$$\Xi(t) = \alpha(S(t), \Xi_0) = (S^{-1}(t)P, S^{-1}(t)\overline{p}_1, \dots, S^{-1}(t)\overline{p}_n),$$

where $S(t) \in \mathbf{SE}(3)$ is a curve with $S(0) = I_4$ and $\frac{\mathrm{d}}{\mathrm{d}t}S|_{S=I_4} = W \in \mathfrak{se}(3)$. By construction $\Xi(0) = \Xi_0$. Taking the derivative of $\Xi(t)$ at $t = 0$ we get

$$(-WP, -W\overline{p}_1, \dots, -W\overline{p}_n) \in \ker \mathrm{d}\pi_{\Xi_0} \subset T_{\Xi_0}\mathcal{T}_n(3), \tag{20}$$

that is, varying $W \in \mathfrak{se}(3)$ fully characterises the kernel of $\mathrm{d}\pi$ at Ξ_0.

Returning to the generic notation $\Xi = (P, \overline{p}_1, \dots, \overline{p}_n) \in \mathcal{T}_n(3)$, define an equivalence relation on tangent vectors $\dot{\Xi} = (PV, P\mathring{v}_1, \dots, P\mathring{v}_n) \in T_\Xi\mathcal{T}_n(3)$ of the total space by

$$\lfloor \dot{\Xi} \rfloor_\Xi = \{(PV - WP, P\mathring{v}_1 - W\overline{p}_1, \dots, P\mathring{v}_n - W\overline{p}_n) \mid W \in \mathfrak{se}(3)\}.$$

A representation of the tangent space $T_\xi\mathcal{M}_n(3)$ at $\xi = \pi(\Xi)$ is now given by the quotient $T_\Xi\mathcal{T}_n(3)/\ker \mathrm{d}\pi_\Xi$, that is

$$T_\xi\mathcal{M}_n(3) = \{\lfloor \dot{\Xi} \rfloor_\Xi \mid \dot{\Xi} \in T_\Xi\mathcal{T}_n(3)\}.$$

This representation of $T_\xi\mathcal{M}_n(3)$ depends on the particular choice of $\Xi \in \mathcal{T}_n(3)$ (where $\pi(\Xi) = \xi$) used in the construction. It is, however, straightforward to map between different representations of the same tangent space associated with different points in the fibre. In particular, for a given Ξ, consider the point $\alpha(S, \Xi) \in \mathcal{T}_n(3)$ for $S \in \mathbf{SE}(3)$. Then for a time varying trajectory $\Xi(t)$ with $\Xi(0) = \Xi$

$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha(S, \Xi)|_{t=0} = \mathrm{d}\alpha_S(\dot{\Xi}) = (S^{-1}PV, S^{-1}P\mathring{v}_1, \dots, S^{-1}\mathring{v}_n) \in T_{\alpha(S,\Xi)}\mathcal{T}_n(3).$$

By construction $\pi(\alpha(S, \Xi(t))) = \pi(\Xi(t))$ and thus $\lfloor \frac{\mathrm{d}}{\mathrm{d}t}\alpha(S, \Xi)|_{t=0} \rfloor_{\alpha(S,\Xi)}$ and $\lfloor \frac{\mathrm{d}}{\mathrm{d}t}\Xi|_{t=0} \rfloor_\Xi$ are the same element of $T_{\pi(\Xi)}\mathcal{M}_n(3)$ expressed in different representations of the tangent space. Thus, given a tangent vector $\lfloor \dot{\Xi} \rfloor_\Xi$, the equivalent algebraic representation of the same tangent vector at a different point $\alpha(S, \Xi)$ in the fibre is $\lfloor \mathrm{d}\alpha_S \dot{\Xi} \rfloor_{\alpha(S,\Xi)}$. In explicit coordinates

$$\lfloor (PV, P\mathring{v}_1, \dots, P\mathring{v}_n) \rfloor_\Xi = \lfloor (S^{-1}PV, S^{-1}P\mathring{v}_1, \dots, S^{-1}\mathring{v}_n) \rfloor_{\alpha(S,\Xi)}.$$

Note that $\mathrm{d}\pi : T_\Xi\mathcal{T}_n(3) \to T_\xi\mathcal{M}_n(3)$. This map is given by $\mathrm{d}\pi(\dot{\Xi}) = \lfloor \dot{\Xi} \rfloor_\Xi$, which can be written as

$$\mathrm{d}\pi(PV, P\mathring{v}_1, \dots, P\mathring{v}_n) = \lfloor (PV, P\mathring{v}_1, \dots, P\mathring{v}_n) \rfloor_\Xi \tag{21}$$

with respect to the natural representation.

3.5. SLAM manifold kinematics

The natural kinematics of the SLAM problem on the SLAM manifold are given by the projection of the total space kinematics onto the quotient SLAM manifold.

In particular, define a map

$$\begin{aligned} & f : \mathcal{M}_n(3) \times \mathbb{V} \to T\mathcal{M}_n(3), \\ & f(\xi, (V, \vec{v}_1, \dots, \vec{v}_n)) := \lfloor (PV, P\vec{v}_1, \dots, P\vec{v}_n) \rfloor_\Xi \end{aligned} \tag{22}$$

for any $\Xi = (P, \overline{p}_1, \dots, \overline{p}_n) \in \mathcal{T}_n(3)$ such that $\pi(\Xi) = \xi$. To see that this map is well defined compute

$$\begin{aligned} f(\pi(\alpha(S, \Xi)), (V, \vec{v}_1, \dots, \vec{v}_n)) &= \lfloor (S^{-1}PV, S^{-1}P\vec{v}_1, \dots, S^{-1}P\vec{v}_n) \rfloor_{\alpha(S,\Xi)} \\ &= \lfloor \mathrm{d}\alpha_S(PV, P\vec{v}_1, \dots, P\vec{v}_n) \rfloor_{\alpha(S,\Xi)} \\ &= \lfloor (PV, P\vec{v}_1, \dots, P\vec{v}_n) \rfloor_\Xi \\ &= f(\pi(\Xi), (V, \vec{v}_1, \dots, \vec{v}_n)). \end{aligned}$$

That is, the value of f is independent of the choice of element Ξ of the equivalence class $[\Xi]$ used in its definition.

The kinematics

$$\frac{\mathrm{d}}{\mathrm{d}t}\xi = f(\xi, (V, \vec{v}_1, \dots, \vec{v}_n)), \qquad \xi(0) = \xi_0, \tag{23}$$

are the kinematics of the SLAM problem, expressed on the SLAM manifold, for physical velocities $(V, \vec{v}_1, \dots, \vec{v}_n) \in \mathbb{V}$.

3.6. Outputs

In the following development we will not model noise in the sensors. Real measurements are of course corrupted by noise and generative noise models are an important part of stochastic filter formulations. The present development, however, draws its inspiration from the nonlinear observer community and focuses on global analysis of the geometric and non-linear structure of the system. We mark the distinction by using the term outputs rather than measurements in the following development.

The outputs considered are body-fixed frame observations of points in the environment. Associated sensing modalities include RGBD cameras, stereo cameras, LIDAR, etc. The output equation is the map $h : \mathcal{M}_n(3) \to \mathcal{N}_1 \times \cdots \times \mathcal{N}_n$, given by

$$h([P, \overline{p}_1, \dots, \overline{p}_n]) := (P^{-1}\overline{p}_1, \dots, P^{-1}\overline{p}_n) \tag{24}$$

that was shown to be well defined earlier (19). The combined output space is $\mathcal{N}_n(3) = \mathbb{R}^3 \times \cdots \times \mathbb{R}^3$ and we write $y = (\overline{y}_1, \dots, \overline{y}_n)$ where appropriate.

We will assume that the robot velocities $V \in \mathfrak{se}(3)$ are measured using odometry or inertial sensors. This is a standard assumption in SLAM algorithms and there are established sensor suites that provide this information in the body-fixed-frame coordinates. If the landmarks are moving then velocity measurements $(\vec{v}_1, \dots, \vec{v}_n)$ are also required. Such measurements can be provided by doppler radar, high frequency LIDAR, and some vision systems. In practice, many SLAM systems simply avoid choosing moving points as part of the environment map and use a static world assumption $v_i = 0$.

4. Symmetry of the SLAM problem

In this section we show that the SLAM problem is equivariant. The full symmetry consists of four parts, firstly a group, secondly a transitive group action φ on the state space $\mathcal{M}_n(3)$, thirdly a group action on the velocity space $\mathbb{V}$ such that the SLAM kinematics are equivariant, and finally a group action on the output space $\mathcal{N}_n(3)$. The Lie group we consider $\mathbf{SE}_{n+1}(3)$ was first proposed in the work of Barrau and Bonnabel (Barrau and Bonnabel (2016)) and was separately developed in the authors' preliminary work (Mahony and Hamel (2017)) as the $\mathbf{SLAM}_n(3)$ group.

4.1. The Lie group $\mathbf{SE}_{n+1}(3)$

The following development draws from Barrau and Bonnabel (2016) and the parallel work Mahony and Hamel (2017). The group $\mathbf{SE}_{n+1}(3)$ can be parameterised by elements

$$\mathbf{SE}_{n+1}(3) = \{(A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n) \mid A \in \mathbf{SE}(3), \mathring{\vec{a}}_i \in \mathbf{R}^3, i = 1, \dots, n\}.$$

with group multiplication

$$\begin{aligned}(A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n) \cdot (B, \mathring{\vec{b}}_1, \dots, \mathring{\vec{b}}_n) &= (AB, \mathring{\vec{a}}_1 + A\mathring{\vec{b}}_1, \dots, \mathring{\vec{a}}_n + A\mathring{\vec{b}}_n), \\ &= (AB, \mathring{\vec{a}}_1 + \overset{\circ}{\overline{R_A b_1}}, \dots, \mathring{\vec{a}}_n + \overset{\circ}{\overline{R_A b_n}}),\end{aligned} \tag{25}$$

identity $\mathrm{id} = (I, \mathring{\vec{0}}, \dots, \mathring{\vec{0}})$ and inverse

$$(A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n)^{-1} = (A^{-1}, -\overset{\circ}{\overline{R_{A^{-1}} a_1}}, \dots, -\overset{\circ}{\overline{R_{A^{-1}} a_n}}).$$

We will commonly use $X = (A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n) \in \mathbf{SE}_{n+1}(3)$ to denote elements of this group.

4.2. State Symmetry

Lemma 4.1. *The mapping* $\varphi : \mathbf{SE}_{n+1}(3) \times \mathcal{M}_n(3) \to \mathcal{M}_n(3)$ *defined by*

$$\varphi((A, a_1, \dots, a_n), [P, \overline{p}_1, \dots, \overline{p}_n]) = [PA, \overline{p}_1 + P\mathring{\vec{a}}_1, \dots, \overline{p}_n + P\mathring{\vec{a}}_n] \tag{26}$$

is a transitive right group action of $\mathbf{SE}_{n+1}(3)$ *on* $\mathcal{M}_n(3)$*.*

Proof 4.2. *Firstly, it is necessary to verify that the action is well defined on* $\mathcal{M}_n(3)$*. In particular, for any* $S \in \mathbf{SE}(3)$

$$\begin{aligned}\varphi((A, a_1, \dots, a_n), [\alpha(S, (P, \overline{p}_1, \dots, \overline{p}_n))]) &= \varphi((A, a_1, \dots, a_n), [S^{-1}P, S^{-1}\overline{p}_1, \dots, S^{-1}\overline{p}_n]) \\ &= [S^{-1}PA, S^{-1}\overline{p}_1 + S^{-1}P\mathring{\vec{a}}_1, \dots, S^{-1}\overline{p}_n + S^{-1}P\mathring{\vec{a}}_n] \\ &= [\alpha\left(S, (PA, \overline{p}_1 + P\mathring{\vec{a}}_1, \dots, \overline{p}_n + P\mathring{\vec{a}}_n)\right)] \\ &= [PA, \overline{p}_1 + P\mathring{\vec{a}}_1, \dots, \overline{p}_n + P\mathring{\vec{a}}_n] \\ &= \varphi\left((A, a_1, \dots, a_n), [P, \overline{p}_1, \dots, \overline{p}_n]\right).\end{aligned} \tag{27}$$

To verify the group action property:

$$\begin{aligned}\varphi\left((A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n), \varphi((B, \mathring{\vec{b}}_1, \dots, \mathring{\vec{b}}_n), [P, \overline{p}_1, \dots, \overline{p}_n])\right) &= \varphi\left((A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n), [PB, \overline{p}_1 + P\mathring{\vec{b}}_1, \dots, \overline{p}_n + P\mathring{\vec{b}}_n]\right) \\ &= \left[PBA, (\overline{p}_1 + P\mathring{\vec{b}}_1) + PB\overset{\circ}{\overline{a_1}}, \dots, (\overline{p}_n + P\mathring{\vec{b}}_n) + PB\overset{\circ}{\overline{a_n}}\right] \\ &= \left[P(BA), \overline{p}_1 + P(\mathring{\vec{b}}_1 + B\overset{\circ}{\overline{a_1}}), \dots, \overline{p}_n + P(\mathring{\vec{b}}_n + B\overset{\circ}{\overline{a_n}})\right] \\ &= \varphi\left((BA, \mathring{\vec{b}}_1 + B\mathring{\vec{a}}_1, \dots, \mathring{\vec{b}}_n + B\overset{\circ}{\overline{a_n}}), \left[P, \overline{p}_1, \dots, \overline{p}_n\right]\right) \\ &= \varphi\left((B, \mathring{\vec{b}}_1, \dots, \mathring{\vec{b}}_n) \cdot (A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n), \left[P, \overline{p}_1, \dots, \overline{p}_n\right]\right).\end{aligned}$$

It is straightforward to verify the identity property

$$\varphi(\mathrm{id}, [P, \overline{p}_1, \dots, \overline{p}_n]) = [P, \overline{p}_1, \dots, \overline{p}_n],$$

and transitivity follows from the property that

$$\varphi((P, \mathring{\vec{p}}_1, \dots, \mathring{\vec{p}}_n), [I_4, \overline{0}, \dots, \overline{0}]) = [P, \overline{p}_1, \dots, \overline{p}_n],$$

and hence any point in $\mathcal{M}_n(3)$ *can be reached from* $[I_4, \overline{0}, \dots, \overline{0}]$ *by suitable construction of an element of* $\mathbf{SE}_{n+1}(3)$*.*

A visualisation of the action of φ on $\mathcal{M}_n(3)$ is given in Figure 2. In particular, the robot pose is updated by the rigid body transformation $A \in \mathbf{SE}(3)$ while the environment points $\overline{p}_i$ are updated by the translations $\mathring{\vec{a}}_i$ after derotation. We emphasise that the group elements $(A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n)$ are not physical variables and can only be interpreted as transforming SLAM configurations.

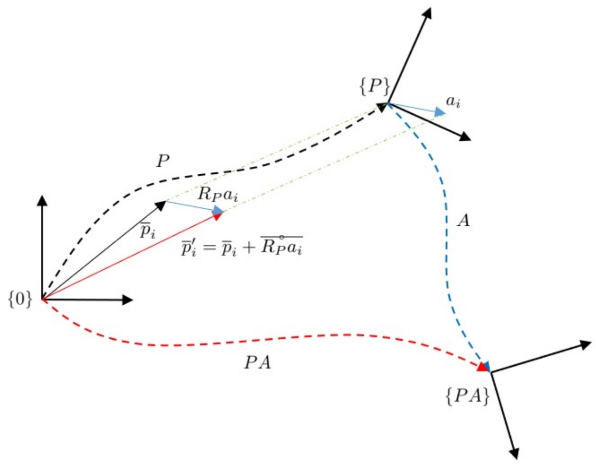

Figure 2: Group action $\varphi((A, \mathring{a}_1, \dots, \mathring{a}_n), (P, \overline{p}_1, \dots, \overline{p}_n))$. The pose $P \mapsto PA$, that is the tip point of the pose is updated by the correction $A \in \mathbf{SE}(3)$. The environment points $\overline{p}_i \mapsto \overline{p}_i + \overline{R_P \mathring{a}_i}$, are updated by the corrections a_i rotated back into the inertial frame.

4.3. Compatibility of the output

A key property of the geometric SLAM formulation is that there is also a compatible group operation on the output $\mathcal{N}_n(3)$ of the system.

Lemma 4.3. *The action* $\rho : \mathbf{SE}_{n+1}(3) \times \mathcal{N}_n(3) \to \mathcal{N}_n(3)$ *defined by*

$$\rho((A, \mathring{a}_1, \dots, \mathring{a}_n), (\overline{y}_1, \dots, \overline{y}_n)) := (A^{-1}(\overline{y}_1 + \mathring{a}_1), \dots, A^{-1}(\overline{y}_n + \mathring{a}_n)) \tag{28}$$

is a transitive right group action on $\mathcal{N}_n(3)$*. Furthermore, one has that the output map h* (24) *is equivariant with respect to* ρ *and* ϕ*. That is*

$$\rho((A, \mathring{a}_1, \dots, \mathring{a}_n), h([\Xi])) = h(\phi((A, \mathring{a}_1, \dots, \mathring{a}_n), [\Xi])).$$

Proof 4.4. *It is straightforward to see that* ρ *is a smooth binary operation. Direct computation shows that*

$$\rho(\mathrm{id}, (\overline{y}_1, \dots, \overline{y}_n)) = (\overline{y}_1, \dots, \overline{y}_n).$$

To verify that ρ *is an right action:*

$$\begin{aligned}
\rho((A, \mathring{a}_1, \dots, \mathring{a}_n), \rho((B, \mathring{b}_1, \dots, \mathring{b}_n), (\overline{y}_1, \dots, \overline{y}_n))) &= \rho((A, \mathring{a}_1, \dots, \mathring{a}_n), (B^{-1}(\overline{y}_1 + \mathring{b}_1), \dots, B^{-1}(\overline{y}_n + \mathring{b}_n))) \\
&= (A^{-1}(B^{-1}(\overline{y}_1 + \mathring{b}_1) + \mathring{a}_1), \dots, A^{-1}(B^{-1}(\overline{y}_n + \mathring{b}_n) + \mathring{a}_n)) \\
&= (A^{-1}B^{-1}((\overline{y}_1 + \mathring{b}_1) + B\mathring{a}_1), \dots, A^{-1}B^{-1}((\overline{y}_n + \mathring{b}_n) + B\mathring{a}_n)) \\
&= ((BA)^{-1}(\overline{y}_1 + (\mathring{b}_1 + B\mathring{a}_1)), \dots, (BA)^{-1}(\overline{y}_n + (\mathring{b}_n + B\mathring{a}_n))) \\
&= \rho((BA, \mathring{b}_1 + B\mathring{a}_1, \dots, \mathring{b}_n + B\mathring{a}_n), (\overline{y}_1, \dots, \overline{y}_n)) \\
&= \rho((B, \mathring{b}_1, \dots, \mathring{b}_n) \cdot (A, \mathring{a}_1, \dots, \mathring{a}_n), (\overline{y}_1, \dots, \overline{y}_n)).
\end{aligned}$$

To see that h is equivariant with respect to ρ *and* φ *compute*

$$\begin{aligned}
h(\varphi((A, \mathring{a}_1, \dots, \mathring{a}_n), [P, \overline{p}_1, \dots, \overline{p}_n])) &= h([PA, \overline{p}_1 + P\mathring{a}_1, \dots, \overline{p}_n + P\mathring{a}_n]) \\
&= \left(A^{-1}P^{-1}(\overline{p}_1 + P\mathring{a}_1), \dots, A^{-1}P^{-1}(\overline{p}_n + P\mathring{a}_n)\right) \\
&= \left(A^{-1}(P^{-1}\overline{p}_1 + \mathring{a}_1), \dots, A^{-1}(P^{-1}\overline{p}_n + \mathring{a}_n)\right) \\
&= \left(A^{-1}(\overline{y}_1 + \mathring{a}_1), \dots, A^{-1}(\overline{y}_n + \mathring{a}_n)\right) \\
&= \rho((A, \mathring{a}_1, \dots, \mathring{a}_n), (\overline{y}_1, \dots, \overline{y}_n)).
\end{aligned}$$

This completes the proof.

4.4. Equivariance of the SLAM kinematics

A system $\dot{\xi} = f(\xi, u)$ for $\xi \in \mathcal{M}$ and $u \in \mathbb{V}$ is termed equivariant if (Mahony et al. (2013, 2020)) there is a group $\mathbf{G}$ and group actions $\varphi : \mathbf{G} \times \mathcal{M} \to \mathcal{M}$ and $\psi : \mathbf{G} \times \mathbb{V} \to \mathbb{V}$ such that

$$\mathrm{d}\varphi_X f(\xi, u) = f(\varphi_X(\xi), \psi_X(u)) \tag{29}$$

for all $X \in \mathbf{G}$, $\xi \in \mathcal{M}$ and $u \in \mathbb{V}$. In this section we show that the SLAM kinematics (23) are equivariant.

We will continue to use the notation introduced in (6), $A\mathring{\vec{v}} = \overline{\mathring{R_A v}}$, for coordinate change of a landmark velocity by a rotation R_A associated with an element $A \in \mathbf{SE}(3)$. Analogous to that notation, one has that for $V = (\Omega, v)^\wedge$ then

$$V\mathring{\vec{a}} = \overline{\Omega \mathring{\times} a}, \qquad a \in \mathbf{R}^3.$$

Lemma 4.5. *The mapping $\psi : \mathbf{SE}_{n+1}(3) \times \mathbb{V} \to \mathbb{V}$, defined by*

$$\psi((A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n), (V, \mathring{\vec{v}}_1, \dots, \mathring{\vec{v}}_n)) := (\mathrm{Ad}_{A^{-1}} V, A^{-1}(\mathring{\vec{v}}_1 + V\mathring{\vec{a}}_1), \dots, A^{-1}(\mathring{\vec{v}}_n + V\mathring{\vec{a}}_n)) \tag{30}$$

is a right group action of $\mathbf{SE}_{n+1}(3)$ on the space of inputs $\mathbb{V}$.

Proof 4.6. *The algebraic map is well defined by construction. Observe that*

$$\begin{aligned}
&\psi((B, \mathring{\vec{b}}_1, \dots, \mathring{\vec{b}}_n), \psi((A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n), (V, \mathring{\vec{v}}_1, \dots, \mathring{\vec{v}}_n))) \\
&\quad = \psi((B, \mathring{\vec{b}}_1, \dots, \mathring{\vec{b}}_n), (\mathrm{Ad}_{A^{-1}} V, A^{-1}(\mathring{\vec{v}}_1 + V\mathring{\vec{a}}_1), \dots, A^{-1}(\mathring{\vec{v}}_n + V\mathring{\vec{a}}_n))) \\
&\quad = (\mathrm{Ad}_{B^{-1}} \mathrm{Ad}_{A^{-1}} V, B^{-1}A^{-1}(\mathring{\vec{v}}_1 + V\mathring{\vec{a}}_1) + B^{-1}\left(\mathrm{Ad}_{A^{-1}} V\right)\mathring{\vec{b}}_1, \dots, B^{-1}A^{-1}(\mathring{\vec{v}}_n + V\mathring{\vec{a}}_n) + B^{-1}\left(\mathrm{Ad}_{A^{-1}} V\right)\mathring{\vec{b}}_n) \\
&\quad = (\mathrm{Ad}_{(AB)^{-1}} V, (AB)^{-1}(\mathring{\vec{v}}_1 + V(\mathring{\vec{a}}_1 + A\mathring{\vec{b}}_1)), \dots, (AB)^{-1}(\mathring{\vec{v}}_n + V(\mathring{\vec{a}}_n + A\mathring{\vec{b}}_n)) \\
&\quad = \psi((AB, \mathring{\vec{a}}_1 + A\mathring{\vec{b}}_1, \dots, \mathring{\vec{a}}_n + A\mathring{\vec{b}}_n), (V, \mathring{\vec{v}}_1, \dots, \mathring{\vec{v}}_n)), \\
&\quad = \psi((A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n) \cdot (B, \mathring{\vec{b}}_1, \dots, \mathring{\vec{b}}_n), (V, \mathring{\vec{v}}_1, \dots, \mathring{\vec{v}}_n)).
\end{aligned}$$

This verifies the right group action property. It is straightforward to verify

$$\psi(\mathrm{id}, (V, \mathring{\vec{v}}_1, \dots, \mathring{\vec{v}}_n)) = (V, \mathring{\vec{v}}_1, \dots, \mathring{\vec{v}}_n).$$

It is now possible to demonstrate equivariance of the SLAM kinematics on $\mathcal{M}_n(3)$.

Lemma 4.7. *The SLAM kinematics (22) are equivariant under the group actions (26) and (30) in the sense of (29).*

Proof 4.8. *Let $X = (A, \mathring{\vec{a}}_1, \dots, \mathring{\vec{a}}_n)$, $\Xi = (P, \bar{p}_1, \dots, \bar{p}_n)$ with $\xi = \pi(\Xi) = [\Xi]$, and $u = (V, \mathring{\vec{v}}_1, \dots, \mathring{\vec{v}}_n)$. For $X \in \mathbf{SE}_{n+1}(3)$ define a map $\Upsilon_X : \mathcal{T}_n(3) \to \mathcal{T}_n(3)$ by*

$$\Upsilon_X(\Xi) := (PA, \bar{p}_1 + P\mathring{\vec{a}}_1, \dots, \bar{p}_n + P\mathring{\vec{a}}_n). \tag{31}$$

It is easily verified that

$$\pi(\Upsilon_X(\Xi)) = \varphi(X, \pi(\Xi)). \tag{32}$$

Consider a trajectory $\Xi(t)$ for a fixed $X \in \mathbf{SE}_{n+1}(3)$ and take the time differential of (32). One has

$$\mathrm{d}\varphi_X \circ \mathrm{d}\pi(\dot{\Xi}) = \mathrm{d}\pi \circ \mathrm{d}\Upsilon_X(\dot{\Xi}).$$

Recalling (21) one can now write

$$\begin{aligned}
\mathrm{d}\varphi_X(\lfloor\dot{\Xi}\rfloor_\Xi) &= \mathrm{d}\varphi_X \circ \mathrm{d}\pi(\dot{\Xi}) \\
&= \mathrm{d}\pi \circ \mathrm{d}\Upsilon_X(\dot{\Xi}) = \lfloor \mathrm{d}\Upsilon_X(\dot{\Xi})\rfloor_{\Upsilon_X(\Xi)}
\end{aligned} \tag{33}$$

for any tangent vector $\dot{\Xi} \in T_\Xi\mathcal{T}_n(3)$. It is easily verified that

$$\mathrm{d}\Upsilon_X(PV, P\mathring{\vec{v}}_1, \dots, P\mathring{\vec{v}}_n) = (PVA, P(\mathring{\vec{v}}_1 + V\mathring{\vec{a}}_1), \dots, P(\mathring{\vec{v}}_n + V\mathring{\vec{a}}_n)). \tag{34}$$

Recalling (22) *one computes*

$$\begin{aligned}
f(\varphi_X(\xi), \psi_X(u)) &= f([PA, \bar{p}_1 + P\vec{a}_1, \dots, \bar{p}_n + P\vec{a}_n], (\mathrm{Ad}_{A^{-1}} V, A^{-1}(\vec{v}_1 + V\vec{a}_1), \dots, A^{-1}(\vec{v}_n + V\vec{a}_n))) \\
&= \lfloor (PVA, P(\vec{v}_1 + V\vec{a}_1), \dots, P(\vec{v}_n + V\vec{a}_n)) \rfloor_{\Upsilon_X(\Xi)} \\
&= \lfloor \mathrm{d}\Upsilon_X(PV, P\vec{v}_1, \dots, P\vec{v}_n) \rfloor_{\Upsilon_X(\Xi)} \\
&= \mathrm{d}\varphi_X \left(\lfloor (PV, P\vec{v}_1, \dots, P\vec{v}_n \rfloor_{\Xi} \right) \\
&= \mathrm{d}\varphi_X f(\xi, u)
\end{aligned}$$

where the third and fourth steps follow from (34) *and* (33)*, respectively. This completes the proof.*

5. Observer design

In the following section we provide a constructive design of a simple constant gain observer for the SLAM problem that is globally asymptotically stable. The contribution of this section is not so much in providing yet another SLAM algorithm, of which there are many, but rather to use the geometric structure and the associated observer design to understand the tradeoff between localisation and mapping that is inherent in the SLAM formulation. The constant gain observer developed can be easily and intuitively tuned to behave either as a localisation algorithm (for known environment points) or a mapping algorithm (for known robot trajectory). The example also provides a good example of the principles of equivariant observer design Mahony et al. (2013, 2020).

We approach the observer design by lifting the system kinematics onto the symmetry group and designing the observer on $\mathbf{SE}_{n+1}(3)$. Let $\xi(t) = [P(t), \bar{p}_1(t), \dots, \bar{p}_n(t)] \in \mathcal{M}_n(3)$ be the 'true' configuration of the SLAM problem. Let $X(t) = (A(t), \vec{a}_1(t), \dots, \vec{a}_n(t)) \in \mathbf{SE}_{n+1}(3)$ and define *lifted kinematics* (Mahony et al. (2013, 2020))

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(A(t), \vec{a}_1(t), \dots, \vec{a}_n(t)) = (AV, A\vec{v}_1, \dots, A\vec{v}_n), \\
A(0) = I_4, a_1(0) = 0, \cdots, a_n(0) = 0.
\end{aligned} \tag{35}$$

Equation (35) evolves on $\mathbf{SE}_{n+1}(3)$ where V is the velocity of the robot and v_i is the velocity of the ith target points. Choose an arbitrary reference configuration

$$\xi^\circ = [P^\circ, \bar{p}_1^\circ, \dots, \bar{p}_n^\circ] \in \mathcal{M}_n(3).$$

If the initial condition $X(0) \in \mathbf{SE}_{n+1}(3)$ of the lifted kinematics satisfies $\varphi(X(0), \xi^\circ) = \xi(0)$ then (35) induces a trajectory that satisfies

$$\xi(t) = \varphi(X(t), \xi^\circ) \in \mathcal{M}_n(3)$$

for all time Mahony et al. (2013, 2020).

We choose the state space of the observer to lie on the SLAM group

$$\hat{X} = (\hat{A}, \hat{\vec{a}}_1, \dots, \hat{\vec{a}}_n) \in \mathbf{SE}_{n+1}(3) \tag{36}$$

and will use the lifted kinematics (35) as the internal model for the observer design. The configuration estimate generated by the observer is given by

$$\hat{\xi} = [\hat{P}, \hat{\bar{p}}_1, \dots, \hat{\bar{p}}_n] = \varphi(\hat{X}, \xi^\circ) \in \mathcal{M}_n(3)$$

given the reference $\xi^\circ \in \mathcal{M}_n(3)$. The goal of the observer design is to estimate both the relative symmetry that takes $\xi^\circ = [P^\circ, \bar{p}_1^\circ, \dots, \bar{p}_n^\circ]$ to $\xi(0) = [P(0), \bar{p}_1(0), \dots, \bar{p}_n(0)]$ as well as to encode the ongoing evolution of $X(t)$ given by (35) while correcting for errors introduced by the noisy measurement of $V \in \mathfrak{se}(3)$.

Theorem 5.1. *Consider the kinematics* (35) *evolving on* $\mathbf{SE}_{n+1}(3)$ *along with outputs* $y = h([\Xi(t)]) \in \mathcal{N}_n(3)$ *given by* (24) *and velocity* $V \in \mathfrak{se}(3)$*. Denote the observer state by* $\hat{X} = (\hat{A}, \hat{\vec{a}}_1, \dots, \hat{\vec{a}}_n) \in \mathbf{SE}_{n+1}(3)$*. Assume that each landmark* $\bar{p}_i$*, for* $i = \{1, \dots, n\}$*, is moving with a* constant *unknown velocity* $\frac{\mathrm{d}}{\mathrm{d}t}\bar{p}_i = \mathrm{const.}$ *and note that* $\vec{v}_i = P^{-1}\dot{\bar{p}}_i$

inherits the ego motion from the changing robot pose. Define $\mathring{\vec{w}}_i = \hat{P}\mathring{\vec{v}}_i = P_0\hat{A}\mathring{\vec{v}}_i$, *for* $i = \{1, \ldots, n\}$ *and note that for* $\hat{A} \to A$ *then* $\mathring{\vec{w}}_i \to P\mathring{\vec{v}}_i = \dot{\bar{p}}_i$.

Fix an arbitrary reference $\xi^\circ = [P^\circ, \bar{p}_1^\circ, \ldots, \bar{p}_n^\circ] \in \mathcal{M}_n(3)$ *and define the* output error *to be*

$$\bar{e} = \rho(\hat{X}^{-1}, (\bar{y}_1, \ldots, \bar{y}_n)). \tag{37}$$

where $(\bar{y}_1^\circ, \ldots, \bar{y}_n^\circ) = h(\xi^\circ)$. *Choose gains* $k_0 > 0$ *and* $k_i, l_i, m_i > 0$ *for* $i = 1, \ldots, n$.

Consider the observer defined by

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{X} = (\hat{A}V + \Delta\hat{A}, \mathring{\vec{\delta}}_1 + \hat{\mathring{\vec{w}}}_1, \ldots, \mathring{\vec{\delta}}_n + \hat{\mathring{\vec{w}}}_n), \quad \hat{X}(0) = \mathrm{id}, \tag{38}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathring{\vec{w}}}_i := \mathrm{Ad}_{P_0}\,\Delta\hat{\mathring{\vec{w}}}_i + m_i k_i(\bar{e}_i - \bar{y}_i^\circ), \quad \hat{\mathring{\vec{w}}}_i(0) = \mathring{\vec{0}}, \tag{39}$$

where $\hat{\mathring{\vec{w}}}_i$ *is an estimate of* $\mathring{\vec{w}}_i$ *and the innovations are given by*

$$\Delta := -k_0\mathbb{P}_{\mathfrak{se}}\left(\sum_{i=1}^{n} k_i(\bar{e}_i - \bar{y}_i^\circ)\bar{e}_i^\top\right) \in \mathfrak{se}(3) \tag{40}$$

$$\mathring{\vec{\delta}}_i := \frac{l_i}{k_i}(\bar{e}_i - \bar{y}_i^\circ) + \Delta\hat{\mathring{\vec{a}}}_i + \hat{\mathring{\vec{w}}}_i. \tag{41}$$

Then, the configuration estimate $\hat{\xi}(t) = \varphi(\hat{X}(t), \xi^\circ) \in \mathcal{M}_n(3)$ *and the estimates* $(\hat{\mathring{\vec{w}}}_1, \ldots, \hat{\mathring{\vec{w}}}_n)$ *converge globally asymptotically and locally exponentially to the true state* $\xi(t) \in \mathcal{M}_n(3)$ *and to the* $(\mathring{\vec{w}}_1, \ldots, \mathring{\vec{w}}_n)$ *respectively.*

Proof 5.2. *Let* $X(t) \in \mathbf{SE}_{n+1}(3)$ *satisfy the lifted kinematics* (35) *and assume* $\phi(X(0), \xi^\circ) = \xi(0)$. *It follows that* $\varphi(X(t), \xi(0)) = \xi(t)$ *Mahony et al. (2013, 2020). Define a group error E by*

$$E = \hat{X}X^{-1} = (\tilde{A}, \tilde{\mathring{\vec{a}}}_1, \ldots, \tilde{\mathring{\vec{a}}}_n) \in \mathbf{SE}_{n+1}(3), \tag{42}$$

where $\tilde{A} := \hat{A}A^{-1}$, *and* $\tilde{\mathring{\vec{a}}}_i := \hat{\mathring{\vec{a}}}_i - \tilde{A}\mathring{\vec{a}}_i$. *Using* (35) *and* (38), *it is straightforward to verify that*

$$\dot{E} = (\Delta\tilde{A}, \mathring{\vec{\delta}}_1 + \hat{\mathring{\vec{w}}}_1 + \Delta(\hat{\mathring{\vec{a}}}_1 - \tilde{\mathring{\vec{a}}}_1) - \mathring{\vec{w}}_1, \ldots, \mathring{\vec{\delta}}_n + \hat{\mathring{\vec{w}}}_n + \Delta(\hat{\mathring{\vec{a}}}_n - \tilde{\mathring{\vec{a}}}_n) - \mathring{\vec{w}}_n). \tag{43}$$

Using the fact that $E^{-1} = (\tilde{A}^{-1}, -\tilde{A}^{-1}\tilde{\mathring{\vec{a}}}_1, \ldots, -\tilde{A}^{-1}\tilde{\mathring{\vec{a}}}_n)$, *it follows*

$$\begin{aligned}\bar{e} &= (\tilde{A}(\bar{y}_1^\circ - \tilde{A}^{-1}\tilde{\mathring{\vec{a}}}_1)), \ldots, \tilde{A}(\bar{y}_n^\circ - \tilde{A}^{-1}\tilde{\mathring{\vec{a}}}_n)) \\ &= (\tilde{A}(\bar{y}_1^\circ - \tilde{A}^{-1}\tilde{\mathring{\vec{a}}}_1), \ldots, \tilde{A}(\bar{y}_n^\circ - \tilde{A}^{-1}\tilde{\mathring{\vec{a}}}_n)) \\ &= (\tilde{A}\bar{y}_1^\circ - \tilde{\mathring{\vec{a}}}_1, \ldots, \tilde{A}\bar{y}_n^\circ - \tilde{\mathring{\vec{a}}}_n).\end{aligned} \tag{44}$$

Based on (43), *it is straightforward to show that the derivative of each element of* (44) *(along with* (41)*) fulfills*

$$\begin{aligned}\dot{\bar{e}}_i &= \Delta\bar{e}_i - \Delta\hat{\mathring{\vec{a}}} - (\mathring{\vec{\delta}}_i + \hat{\mathring{\vec{w}}}_i) + \mathring{\vec{w}}_i, \\ &= \Delta(\bar{e}_i - \bar{y}_i^\circ) + \Delta\bar{y}_i^\circ - \frac{l_i}{k_i}(\bar{e}_i - \bar{y}_i^\circ) + \tilde{\mathring{\vec{w}}}_i,\end{aligned}$$

with $\tilde{\mathring{\vec{w}}}_i = \mathring{\vec{w}}_i - \hat{\mathring{\vec{w}}}_i$. *Differentiating* $\tilde{\mathring{\vec{w}}}_i$ *and recalling* (39), *one has*

$$\frac{d}{dt}\tilde{\mathring{\vec{w}}}_i = \mathrm{Ad}_{P_0}\,\Delta\tilde{\mathring{\vec{w}}}_i - m_i k_i(\bar{e}_i - \bar{y}_i^\circ) \tag{45}$$

We first prove global asymptotic convergence of the observer. Consider the following candidate (positive definite) Lyapunov function,

$$\mathcal{L} = \sum_{i=1}^{n}\left(\frac{k_i}{2}\left|\bar{e}_i - \bar{y}_i^\circ\right|^2 + \frac{1}{2m_i}\tilde{\mathring{\vec{w}}}_i^2\right). \tag{46}$$

Differentiating $\mathcal{L}$ and using the fact that for any $x \in \mathbf{R}^3$ then $\mathring{\bar{x}}^\top \Delta \mathring{\bar{x}} = 0$ (respectively $\mathring{\bar{x}}^\top (\mathrm{Ad}_{P_0} \Delta)\mathring{\bar{x}} = 0$), one has

$$\begin{aligned}
\dot{\mathcal{L}} &= \sum_{i=1}^{n} \left(k_i \left(\bar{e}_i - \bar{y}_i^\circ\right)^\top \dot{\bar{e}}_i + \frac{1}{m_i} \mathring{\tilde{\bar{w}}}_i^\top \frac{d}{dt} \mathring{\tilde{\bar{w}}}_i \right) \\
&= \mathrm{tr}\left(\Delta \sum_{i=1}^{n} k_i \bar{y}_i^\circ \left(\bar{e}_i - \bar{y}_i^\circ\right)^\top \right) - \sum_{i=1}^{n} l_i \left|\bar{e}_i - \bar{y}_i^\circ\right|^2
\end{aligned}$$

Now recalling the expression of Δ (40), it is straightforward to verify that

$$\Delta = k_0 \begin{bmatrix} \frac{1}{2}\sum_{i=1}^{n} k_i (y_i^\circ \times e_i)^\times & -\sum_{i=1}^{n} k_i(\bar{e}_i - \bar{y}_i^\circ) \\ 0 & 0 \end{bmatrix}.$$

and hence

$$\dot{\mathcal{L}} = -\frac{k_0}{2} \left| \sum_{i=1}^{n} k_i y_i^\circ \times e_i \right|^2 - k_0 \left| \sum_{i=1}^{n} k_i (y_i^\circ - e_i) \right|^2 - \sum_{i=1}^{n} l_i \left|\bar{e}_i - \bar{y}_i^\circ\right|^2$$

The derivative of the Lyapunov function is negative semi-definite, equal to zero when $\bar{e}_i = \bar{y}_i^\circ$, and application of LaSalle's principle ensures that the equilibrium $\left((\bar{e}_i - \bar{y}_i^\circ)^\top, \mathring{\tilde{\bar{w}}}_i^\top\right)^\top \to (\mathring{\bar{0}}, \mathring{\bar{0}})$ is globally asymptotically *stable. It follows that $E = (\tilde{A}, -\mathring{\tilde{\bar{a}}}_1, \dots, -\mathring{\tilde{\bar{a}}}_n)$ converges* asymptotically *to a constant. The condition $\bar{e}_i = \bar{y}_i^\circ$ implies (44)*

$$\tilde{A}\bar{y}_i^\circ - \mathring{\tilde{\bar{a}}} = \bar{y}_i^\circ, \quad \forall i = 1, \dots, n.$$

By exploiting the expression of $\mathring{\tilde{\bar{a}}}_i := -\tilde{A}\mathring{\bar{a}}_i + \mathring{\hat{\bar{a}}}_i$ (or equivalently $\tilde{a}_i := -R_{\tilde{A}} a_i + \hat{a}_i$), one gets:

$$\tilde{A}(\bar{y}_i^\circ + \mathring{\bar{a}}_i) = \bar{y}_i^\circ + \mathring{\hat{\bar{a}}}_i$$

This in turn implies that the limit satisfies

$$A^{-1}(\bar{y}_i^\circ + \mathring{\bar{a}}_i) = \hat{A}^{-1}(\bar{y}_i^\circ + \mathring{\hat{\bar{a}}}_i)$$

It follows that

$$\rho(X, y_i^\circ) = h(\varphi(X, \xi^\circ)) = h(\varphi(\hat{X}, \xi^\circ)) = \rho(\hat{X}, y_i^\circ).$$

Regarding just the central equality, and noting that h is full rank on $\mathcal{M}_n(3)$ then

$$\xi = \varphi(X, \xi^\circ) = \varphi(\hat{X}, \xi^\circ) = \hat{\xi}$$

which concludes the proof of global asymptotic stability.

To prove local exponential stability of the observer error we define the following slightly modified Lyapunov function:

$$\mathcal{L}_\epsilon = \sum_{i=1}^{n} \left(\frac{k_i}{2} \left|\bar{e}_i - \bar{y}_i^\circ\right|^2 + \frac{1}{2m_i} \mathring{\tilde{\bar{w}}}_i^2 - \epsilon \mathring{\tilde{\bar{w}}}_i^\top (\bar{e}_i - \bar{y}_i^\circ) \right). \tag{47}$$

with ϵ a positive number such that $\epsilon < k_i^{\min}/m_i^{\max}$.

Differentiating $\mathcal{L}_\epsilon$, it yields

$$\dot{\mathcal{L}}_\epsilon = -\frac{k_0}{2} \left| \sum_{i=1}^{n} k_i y_i^\circ \times e_i \right|^2 - k_0 \left| \sum_{i=1}^{n} k_i (y_i^\circ - e_i) \right|^2 - \sum_{i=1}^{n} l_i^\epsilon \left|\bar{e}_i - \bar{y}_i^\circ\right|^2 - \epsilon \sum_{i=1}^{n} \left|\mathring{\tilde{\bar{w}}}_i\right|^2 + \epsilon \sum_{i=1}^{n} \frac{l_i}{k_i} \mathring{\tilde{\bar{w}}}_i^\top (\bar{e}_i - \bar{y}_i^\circ) - \epsilon \sum_{i=1}^{n} \mathring{\tilde{\bar{w}}}_i^\top \Delta \bar{y}_i^\circ$$

$$- \epsilon \sum_{i=1}^{n} \tilde{\mathring{\overline{w}}}_i^\top \left(\Delta + (\mathrm{Ad}_{P_0} \Delta)^\top \right) (\overline{e}_i - \overline{y}_i^\circ).$$

with $l_i^\epsilon = (l_i - \epsilon k_i m_i)$*, for* $i = \{1, \dots, n\}$*. From the above expression of* Δ (40)*, it is straightforward to verify that around the equilibrium the last term of* $\dot{\mathcal{L}}_\epsilon$ *is a negligible third order term. Hence, in a local neighbourhood of the asymptotic limit*

$$\begin{aligned}
\dot{\mathcal{L}}_\epsilon \approx & -\frac{k_0}{2} \left| \sum_{i=1}^{n} k_i \overline{y}_i^\circ \times e_i \right|^2 - k_0 \left| \sum_{i=1}^{n} k_i (\overline{y}_i^\circ - e_i) \right|^2 - \sum_{i=1}^{n} \left(l_i^\epsilon \left| \overline{e}_i - \overline{y}_i^\circ \right|^2 + \epsilon \left| \tilde{\mathring{\overline{w}}}_i \right|^2 - \epsilon \frac{l_i}{k_i} \tilde{\mathring{\overline{w}}}_i^\top (\overline{e}_i - \overline{y}_i^\circ) \right) \\
& - \epsilon k_0 \sum_{i=1}^{n} \left(\frac{1}{2} \tilde{\mathring{\overline{w}}}_i^\top \left(\sum_{i=1}^{n} k_i (y_i^\circ \times e_i)^\times \right) \overline{y}_i^\circ - \tilde{\mathring{\overline{w}}}_i^\top \left(\sum_{i=1}^{n} k_i (e_i - y_i^\circ) \right) \overline{y}_i^\circ \right) \\
\leq & - \sum_{i=1}^{n} \left(\frac{k_0}{2n} \left| \sum_{i=1}^{n} k_i \overline{y}_i^\circ \times e_i \right|^2 + \frac{k_0}{n} \left| \sum_{i=1}^{n} k_i (\overline{y}_i^\circ - e_i) \right|^2 + l_i^\epsilon \left| \overline{e}_i - \overline{y}_i^\circ \right|^2 + \epsilon \left| \tilde{\mathring{\overline{w}}}_i \right|^2 \right. \\
& \left. - \epsilon \frac{l_i}{k_i} |\tilde{\mathring{\overline{w}}}_i| . |\overline{e}_i - \overline{y}_i^\circ| - \epsilon \frac{k_0}{2} |\tilde{\mathring{\overline{w}}}_i| . \left| \sum_{i=1}^{n} k_i (y_i^\circ \times e_i)^\times \right| - \epsilon k_0 |\tilde{\mathring{\overline{w}}}_i| . \left| \sum_{i=1}^{n} k_i (e_i - y_i^\circ) \right| \right)
\end{aligned}$$

This in turn shows that by choosing $0 < \epsilon < \min \left(\frac{k_i^{\min}}{m_i^{\max}}, \frac{1}{n k_0}, \frac{2 l_i^{\min} {k_i^{\min}}^2}{2 {k_i^{\max}}^3 m_i^{\max} + {l_i^{\max}}^2} \right)$*, then there exists a local neighbourhood of* $(\mathring{0}, \mathring{0})$ *such that* $\dot{\mathcal{L}}_\epsilon < -2 m_i^{\min} \epsilon \mathcal{L}_\epsilon$ *and local exponential stability is proved.*

Although the proof provided is somewhat complicated, the algebraic structure of the observer, (38), (39), (40), (41) is remarkably simple compared to comparable SLAM algorithms in the literature. The innovation terms (40) and (41) are simply lifted versions of the gradients of the cost function with respect to the variables. There is no covariance or information matrix to propagate in this algorithm, there are no linearisations, the variables in the algorithm are directly related to the physical variables of the problem.

Furthermore, although the observer is derived explicitly using the geometric structure of the SLAM manifold it is posed on the symmetry group $\mathbf{SE}_{n+1}(3)$. This symmetry also acts on the total space through the group action Υ (31). As such, one can define a trajectory

$$(\hat{P}, \hat{\overline{p}}_1, \dots, \hat{\overline{p}}_n) := \Upsilon_{\hat{X}}(P^\circ, \overline{p}_1^\circ, \dots, \overline{p}_n^\circ) = (P^\circ \hat{A}, \overline{p}_1^\circ + P^\circ \hat{\mathring{a}}_1, \dots, \overline{p}_n^\circ + P^\circ \hat{\mathring{a}}_n). \tag{48}$$

and consider the evolution of the robot pose $\hat{P}$ and environment map points $(\hat{\overline{p}}_1, \dots, \hat{\overline{p}}_n)$ separately. That is, the proposed observer induces an observer estimate on the total space coordinates analogous to the sort of trajectory estimates provided by classical SLAM algorithms. A key point that we emphasise is that the total space trajectory is not observable from the data. The invariance associated with the gauge transform is always present in the data, and the evolution of the total space trajectory along the fibre can actually be arbitrarily assigned. The geometric invariance associated with the SLAM manifold is present in all SLAM algorithms, however it is usually swept-under-the-carpet by introducing priors on the initial condition or other tricks. The structure of the proposed observer makes this relationship explicit. The observer is posed on the symmetry group and acts on the SLAM manifold to define an globally asymptotically stable system. The action of the observer trajectory on the total space provides insight into how the trajectory on the SLAM manifold can be lifted to relate to the separate localisation and mapping problems.

To see this, consider how choosing the gains can provide focus on estimating the environment (mapping), estimating the pose with respect to known environment (pose estimation), or solve the complete SLAM problem.

- **Mapping:** By choosing $k_0 = 0$ the robot pose estimate is no longer corrected by landmark measurements. This formulation solves the mapping problem independent of the robot pose and is closely related to the work of Guerreiro Guerreiro et al. (2013) and Lourenço *et al.* Lourenço et al. (2016) but posed in the inertial-frame. The robot pose estimate is still present in the total space oberver (48), however, the estimate becomes a forward integration of the measured velocity.

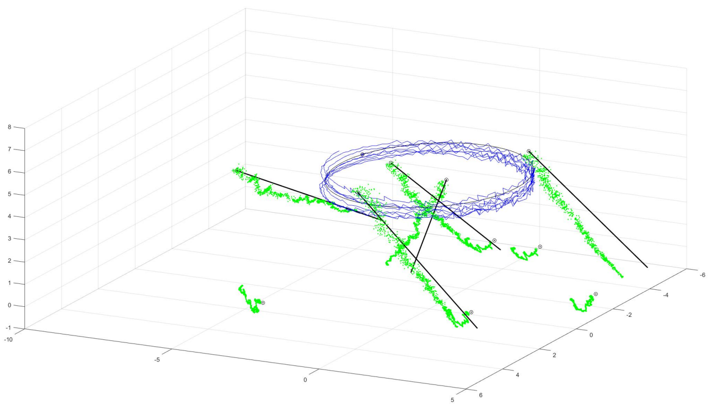

Figure 3: Configuration trace for Simulation 6.1 in the presence of noise. The blue trace is the estimate of the position of the robot frame with its final estimated position plotted as a blue star, while the underlying black trace is the true position of the robot frame with its final location plotted as a black circle. The true landmark feature positions are shown in black, with their final positions shown as black circles. The landmark estimates are plotted in green with their final position shown as green stars Note that the true robot trajectory (in black) is not subject to the velocity noise, that is, the noise process in velocity was assumed to be in the measurement device and not in the robot motion. The high levels of noise are evident in the spread of the measurements as well as the trace of the estimated trajectory.

- **Pose Estimation:** By choosing $l_i = m_i = 0$ the map estimates are no longer updated by the measurements. (Note that (39) has solution $\hat{\vec{w}}_i(t) = \mathring{\vec{0}}$ for all time.) Assuming that the landmark priors are correct (that is $\{\overline{p}_i^\circ\}$ are the true positions of the global landmarks in some frame of reference) then the pose estimate $\hat{P}$ (48) converges to the true pose (in the given frame of reference) analogous to previously published pose estimation algorithms Baldwin et al. (2009); Vasconcelos et al. (2010).

In practice, if this algorithm is employed, a tradeoff between these two extremes must be chosen by relative gain scaling of the two innovations. Choosing gains is an important aspect of any practical implementation of an observer and is discussed in the Simulation section.

6. Simulation results

Details of two simulations are provided to demonstrate the behaviour and capability of the proposed algorithm. The first simulation demonstrates the functionality of the algorithm (Fig. 3-4) with noise in landmark measurements and robot velocity estimates as well as moving landmarks. Although this is already a challenging scenario, the real power of the proposed algorithm is demonstrated in the second simulation (Fig's 5, 6, 7) where we use a simple 2D example to study the global stability, robustness and computational cost of the proposed algorithm.

6.1. Demonstration of functionality

We consider the case of a vehicle equipped with a 3D-sensor, such as a stereo camera, observing an unknown constellation of point feature landmarks. The vehicle moves along a roughly circular trajectory at a fixed altitude ($z = 5m$) above the ground with forward velocity 1.5m.s^{-1} while turning with angular velocity 0.5 rad.s^{-1}, tracing out a circular trajectory of radius 3m (Fig. 3). A Gaussian noise of standard deviation 5-10% relative error is added to the system velocity and the corresponding trajectory is not exactly circular.

There are ten landmark features modeled of which five points are stationary while five points are initialised with a random positive velocity of magnitude between 0 and 0.1m.s^{-1}. The vehicle does not know which landmarks are stationary in advance. All points are initialised randomly on the ground plane ($z = 0$) with the vehicle viewing them from above (Fig. 3). Perfect data association was assumed and no landmark labelling errors are simulated.

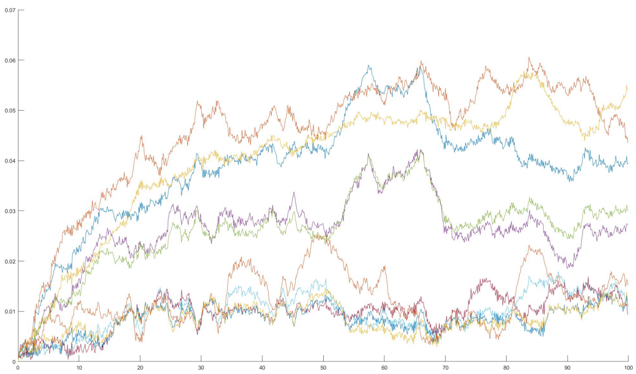

Figure 4: Evolution of the magnitude of velocity estimates $|\hat{\hat{w}}_i|$ for Experiment 2 (with noise). The five velocity estimates associated with stationary points are clearly identified with small velocity estimates. The mean of the vector velocity estimates for these estimates converge to zero.

The observer gains $k_i = 0.05$ and $l_i = 0.05$ for $i = 1 \dots 10$ are used. The velocity observer gain $m_i = 1$ for $i = 1, \dots, 10$, was used in order to impose time-scale separation between the convergence of the configuration error and the velocity estimation. A higher gain in velocity estimation could be used, however, this would couple error in the velocity estimation into the configuration estimation and impact the asymptotic performance of the filter (Fig. 4). The performance of the integral velocity observer is clear in the convergence of the stationary landmark velocity estimates to a noise floor near to zero, and the convergence of the moving landmark velocity estimates to constant magnitudes in Figure 4. The stationary landmark velocity estimates have zero mean over time as can be seen from Figure 3.

The observer estimate is $\hat{\xi}(t) = \varphi(\hat{X}, \xi^{\circ})$. This is an element of the SLAM manifold $\mathcal{M}_n(3)$, and as such does not have separate robot pose and landmark estimates. The configuration plot, Figure 3, uses total space coordinates $(x_{\hat{P}}, \hat{\hat{p}}_1, \dots, \hat{\hat{p}}_n)$ with respect to a reference frame $\{0\}$. In order to visualise the error in the configuration estimate effectively, we have used an invariance transformation $\alpha_{S_{\text{final}}}$ where $S_{\text{final}} \in \mathbf{SE}(3)$ is chosen to map $\hat{P}_{\text{final}}$ to P_{final}. That is, for the purposes of displaying the results, plot the total space coordinates of the observer trajectory in a frame of reference such that the final pose of the observer matches the final pose of the true robot trajectory. As a consequence, there will be zero error in the robot pose in the final instant, all the errors in the observer will be visible in the landmark errors.

6.2. Robustness and Computational Comparison

In this section we consider the robustness of the proposed algorithm, and in particular its robustness to non-Gaussian noise introduced through data association errors (mislabeling of landmarks). These errors are common in real world systems Tombari et al. (2013) where feature detection and data association algorithms that identify and remove outliers Tombari et al. (2013); Chin et al. (2016); Bustos and Chin (2017) consume a major part of the computational resources required for modern SLAM algorithms. To provide strong comparative evidence we keep the case considered very simple, a square 2D trajectory (Fig. 5), and simulate the vanilla Extended Kalman Filter (EKF) SLAM algorithm for comparison. This scenario minimizes the linearisation errors inherent in the EKF formulation and the algorithm is known to perform well for 2D problems Stachniss et al. (2016); Durrant-Whyte and Bailey (2006); Bailey and Durrant-Whyte (2006); Aulinas et al. (2008); Thrun (2002).

Figure 5 shows the comparison of the reconstructed trajectories generated by the proposed equivariant observer and a classical extended Kalman filter algorithm for a typical simulation. The sensor range of the vehicle is modelled as a hemi-disk of radius 1m and only map points within this range are observed at any given time. There were a total of 200 points in the environment models, although not all were observed by the vehicle during the simulation. In addition, when each point is observed and labelled we introduce a 5% probability of mislabelling, that is of associating the observed point with the wrong point from previous measurements. To avoid unrealistic mislabeling we only allow points to be mislabelled within the visible points at the time of sensing, a typical matching error for data association al-

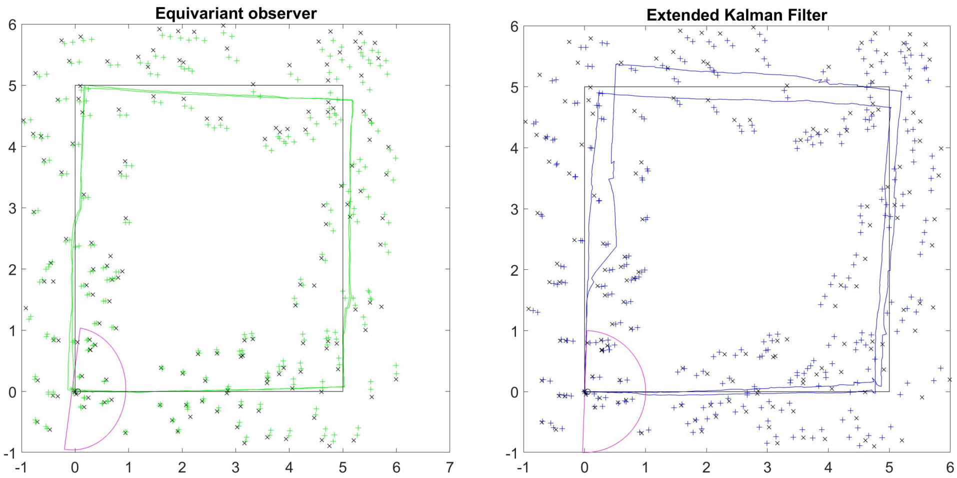

Figure 5: Reconstructed trajectory and map for Equivariant observer (in green) and Extended Kalman Filter (in blue). The black trace in both plots indicates the true trajectory and the black crosses the true features. The robot is located at the radial centre of the red hemi-disk, which denotes the sensor range. Sensor measurements and vehicle odometry are corrupted by noise of relative magnitude 5-10%. In addition, the data association fails with probably 5%, however, we only allow mislabeling within the locally observed subset of points. The global robustness of the proposed equivariant observer deals well with the high noise levels and provides a better solution than the extended Kalman filter.

gorithms. In addition, measurement noise in the order of 5-10% (5cm standard deviation) is added to the measurements and additional noise is added to the robot odometry. The noise parameters are deliberately highly challenging and the algorithms are run on all data with no removal of outliers to emphasise the base level performance of the algorithms. We use the true noise covariances (ignoring data association error) in both measurements and state processes to tune the EKF gains. The equivariant observer uses observer gains $k_i = 1$ and $l_i = 0.2$ for $i = 1 \dots 10$. The gain $m_i = 0$ was set to zero as only static environment points were considered.

Figure 6 graphs the results of the statistical study undertaken. The Root Mean Square Error (RMSE) of the map quality is computed for a sample size of 500 simulations for each data association error probability setting. That is 500 separate simulations similar to Figure 5, for independent noise and map locations, were run and the results compiled to generate Fig. 6. The resulting sample statistics are shown in box plot format. The graph clearly shows the advantage of the Extended Kalman Filter for very low data association error (less than 3% labelling error). Once the data association error exceeds 3 mislabelled points per hundred matches, then the global stability and robustness of the proposed equivariant observer comes into play outperforming the EKF formulation that relies on Gaussian noise assumptions and local linearisation of the system equations.

The fragility of stochastic SLAM algorithms to data association errors is well known (Stachniss et al. (2016); Cadena et al. (2016)). Modern SLAM pipelines include a range of heuristics and sophisticated algorithms to identify and remove outliers (Tombari et al. (2013); Chin et al. (2016); Bustos and Chin (2017)). Although such algorithms are effective, they also consume significant computational resources. The cost of outlier removal is in addition to the cost of the feature extraction (common to all feature based algorithms) and the cost of computing the filter updates. It is not possible to evaluate our algorithm against the plethora of robust stochastic algorithms used to reduce data association errors, however, it is possible to evaluate the computational cost of computing the state estimate using the equivariant algorithm and the extended Kalman filter as a lower bound to the expected computational cost of any modern SLAM algorithm. Figure 7 plots statistics for a series of experiments in which the maps size was increased from five points to two hundred points. At five points the underlying SLAM problem is not well conditioned, and even

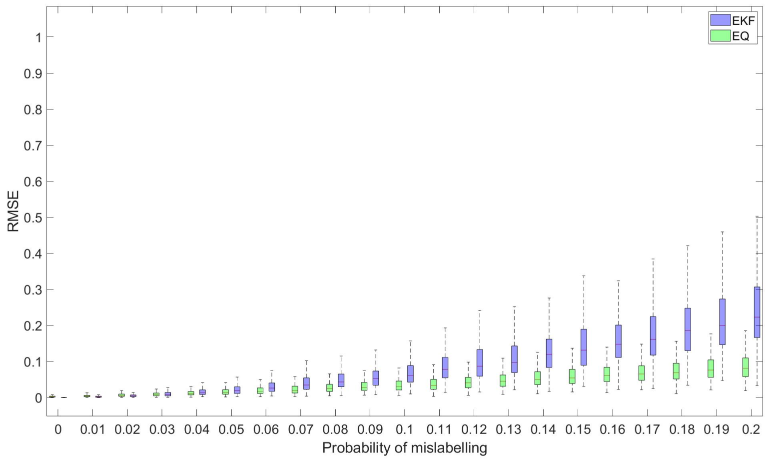

Figure 6: Errors in data association are introduced for each landmark "seen" by the robot sensors with a probability of mislabling each point shown on the x-axis. The Root Mean Square Error (RMSE) of the map quality is computed for a sample size of 500 simulations for each data probability of mislabeling. The mean and standard deviation of the RMSE is shown in the box, and the full range of the RMSE is shown by the wiskers.

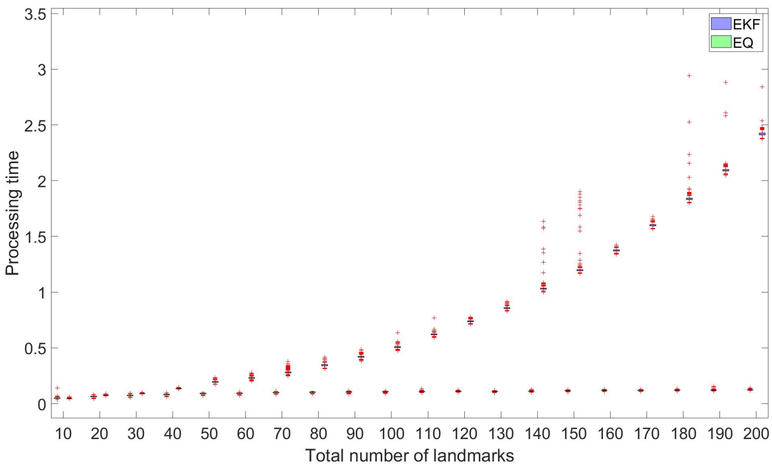

Figure 7: Statistics of the computational cost of updating one step of the SLAM algorithm is shown for a range of map sizes. Neither algorithm was optimized and both algorithm were coded as variations of the same data structure, feature extraction, etc. Satistics from a run of 500 indepedent simulations for each size of map are displayed as box plot.

highly optimized real-time SLAM and VIO algorithms tend to require at least 20-30 feature points active at any one time. In this simulation that would require around 100 map points since most of the points are not visible to the robot most of the time.

Most of the computational cost in the proposed algorithm is associated with the feature extraction process and the cost of processing many points is negligible. The EKF algorithm, that maintains a covariance matrix has computational cost that grows quadratically with map size in both computation and memory. In classical algorithms, this issue is addressed by limiting the map size and working with sub-maps, however, the down side of such an approach is a loss of performance in loop closure over longer trajectories.

The simulations demonstrate the potential of the proposed algorithm as a low computational cost, highly robust

SLAM algorithm that can handle large map size and high levels of data association errors. A classical EKF outperforms the proposed equivariant observer in map RMS error for data association error rates lower than two in a hundred matches, is on par for mislabeling in the order of 3%, and the equivariant observer outperforms the EKF for higher levels of mislabeling error. Moreover, a typical map size to demonstrate this performance is of the order of 100-200 points and the error correction benefit from correctly estimating the error covariance comes at significant computation cost. Outlier detection is a computational cost that would be an additional load on the processor.

The example simulation Fig. 5 was done for 5% probability of mislabeling and shows that the equivariant observer map and trajectory are clearly of higher quality than the corresponding EKF trajectory. Having said this, neither trajectory demonstrates the high precision mapping quality that state-of-the-art SLAM algorithms with sophisticated outlier removal can obtain. The map and pose estimate, however, are quite sufficient to provide spatial awareness to IOT devices and to use as a navigational aid for consumer mobile robotic vehicles. In such applications there is a need for a low complexity SLAM algorithm that is highly robust to non-Gaussian noise such as data association errors and provides good enough map and pose estimates reliably.

7. Conclusion

This paper provides a detailed development of the equivariant structure of the classical landmark point SLAM problem. The key contributions of the paper include; the development of the SLAM manifold $\mathcal{M}_n(3)$, demonstrating that the SLAM kinematics are equivariant with respect to a group action of $\mathbf{SE}_{n+1}(3)$ on $\mathcal{M}_n(3)$, and proposing an observer with globally asymptotically stable error dynamics for landmark SLAM in dynamic environments.

Acknowledgment

Thanks to Pieter van Goor for his invaluable help in running the comparison experiments. This research was supported by the Australian Research Council through the “Australian Centre of Excellence for Robotic Vision” CE140100016 and the Franco-Australian International Research Project “Advancing Autonomy for Unmanned Robotic Systems” (IRP ARS).

References

Absil, P.A., Mahony, R., Sepulchre, R., 2008. Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton, NJ, USA.

Aulinas, J., Petillot, Y., Salvi, J., Lladó., X., 2008. The SLAM problem: A survey, in: Proceedings of the International Conference of the Catalan Association for Artificial Intelligence, IOS Press. pp. 363–371.

Bailey, T., Durrant-Whyte, H.F., 2006. Simultaneous localisation and mapping (SLAM): Part II. IEEE Robotics and Automation Magazine (RAM) 13, 108–117.

Baldwin, G., Mahony, R., Trumpf, J., 2009. A nonlinear observer for 6 DOF pose estimation from inertial and bearing measurements, in: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), pp. 2237–2242.

Barrau, A., Bonnabel, S., 2016. An EKF-SLAM algorithm with consistency properties. arXiv:1510.06263.

Barrau, C., Bonnabel, S., 2017. The invariant extended Kalman filter as a stable observer. IEEE Transactions on Automatic Control 62, 1797–1812. doi:DOI:10.1109/TAC.2016.2594085.

Bjorne, E., Johansen, T.A., Brekke, E.F., 2017. Redesign and analysis of globally asymptotically stable bearing only SLAM, in: 2017 20th International Conference on Information Fusion (Fusion), pp. 1–8.

Bonnabel, S., 2007. Left-invariant extended Kalman filter and attitude estimation, in: Procedings of the IEEE Conference on Decision and Control (CDC), New Orleans, LA, USA. p. 6 pages. doi:DOI:10.1109/CDC.2007.4434662.

Bonnabel, S., Martin, P., Rouchon, P., 2008. Symmetry-preserving observers. IEEE Transactions on Automatic Control 53, 2514–2526.

Bonnabel, S., Martin, P., Salaun, E., 2009. Invariant Extended Kalman Filter: Theory and application to a velocity-aided attitude estimation problem, in: IEEE Conference on Decision and Control, pp. 1297–1304.

Bustos, Á.P., Chin, T.J., 2017. Guaranteed outlier removal for point cloud registration with correspondences. IEEE Transactions on Pattern Analysis and Machine Intelligence 40, 2868–2882. doi:10.1109/TPAMI.2017.2773482.

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I.D., Leonard, J.J., 2016. Past, present, and future of simultaneous localization and mapping: Towards the robust-perception age. IEEE Transactions on Robotics 32, 1309–1332.

Castellanos, J., Martinez-Cantin, R., Tardós, J., Neira, J., 2007. Robocentric map joining: Improving the consistency of EKF-SLAM. Robotics and Autonomous Systems 55, 21–29.

Chin, T.J., Kee, Y.H., Eriksson, A., Neumann, F., 2016. Guaranteed outlier removal with mixed integer linear programs, in: Procedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5858–5866.

Dayoub, F., Cielniak, G., Duckett, T., 2011. Long-term experiments with an adaptive spherical view representation for navigation in changing environments. Robotics and Autonomous Systems (RAS), 59, 285–295.

Dellaert, F., 2012. Factor Graphs and GTSAM: A Hands-on Introduction - Technical Report. Technical Report GT-RIM-CP&R-2012-002. Georgia Institute of Technology.

Dissanayake, G., Huang, S., Wang, Z., Ranasinghe., R., 2011. A review of recent developments in Simultaneous Localization and Mapping, in: International Conference on Industrial and Information Systems, pp. 477–482.

Durrant-Whyte, H.F., Bailey, T., 2006. Simultaneous localisation and mapping (SLAM): Part I. IEEE Robotics and Automation Magazine (RAM) 13, 99–110.

Engel J., Schöps T., C.D., 2014. LSD-SLAM: Large-scale direct monocular SLAM, in: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), Computer Vision – ECCV 2014. Springer. volume 8690.

Grisetti, G., Kummerle, R., Stachniss, C., Burgard, W., 2010. A tutorial on graph-based SLAM. IEEE Intelligent Transportation Systems Magazine 2, 31–43.

Guerreiro, B.J.N., Batista, P., Silvestre, C., Oliveira, P., 2013. Globally asymptotically stable sensor-based simultaneous localization and mapping. IEEE Transactions on Robotics 29, 1380–1395. doi:DOI:10.1109/TRO.2013.2273838.

Hamel, T., Samson, C., 2016. Riccati observers for position and velocity bias estimation from direction measurements, in: 2016 IEEE 55th Conference on Decision and Control (CDC), pp. 2047–2053. doi:10.1109/CDC.2016.7798565.

Hamel, T., Samson, C., 2018. Riccati observers for the nonstationary PnP problem. IEEE Transactions on Automatic Control 63, 726–741. doi:10.1109/TAC.2017.2726179.

andJ.A. Hesch, D.K., Bowman, S., Roumeliotis., S., 2012. On the consistency of vision-aided Inertial Navigation, in: Proceedings of the International Symposium on Experimental Robotics (ISER), pp. 303–317.

Hesch, J., Kottas, D., Bowman, S., Roumeliotis, S., 2014. Camera-IMU-based localization: Observability analysis and consis-tency improvement. The International Journal of Robotics Research (IJRR) 33, 182–201.

Hua, M.D., Zamani, M., Trumpf, J., Mahony, R., Hamel, T., 2011. Observer design on the special euclidean group SE(3), in: Proceedings of the IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, USA.

Huang, G., Mourikis, A., Roumeliotis, S., 2011. An observability-constrained sliding window filter for SLAM. In Proceedings of theIEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) , 65–72.

Huang, G., Mourikis, A., Roumeliotis, S.I., 2010. Observability-based rules for designing consistent EKF SLAM estimators. The International Journal of Robotics Research 29, 502–528.

Ila, V., Porta, J.M., Andrade-Cetto, J., 2010. Information-based compact pose SLAM. IEEE Transactions on Robotics 26, 78–93. doi:10.1109/TRO.2009.2034435.

Johansen, T., Brekke, E., 2016. Globally exponentially stable kalman filtering for SLAM with AHRS, in: 19th International Conference on Information Fusion (FUSION).

Judd, K.M., Gammell, J.D., 2019. The Oxford Multimotion Dataset: Multiple SE(3) Motions With Ground Truth. IEEE Robotics and Automation Letters 4, 800–807. doi:10.1109/LRA.2019.2892656.

Kaess, M., Johannsson, H., Roberts, R., Ila, V., Leonard, J., Dellaert, F., 2012. iSAM2: Incremental smoothing and mapping using the Bayes tree. The International Journal of Robotics Research 31, 216–235. doi:doi.org/10.1177/0278364911430419.

Kanatani, K., Morris, D.D., 2001. Gauges and gauge transformations for uncertainty description of geometric structure with indetermincy. IEEE Transactions on Information Theory 47, 12 pages.

Krajn/'ıık, T., Fentanes, J., Mozos, O., Duckett, T., Ekekrantz, J., Hanheide, M., 2014. Long-term topological localisation for service robots in dynamic environments using spectral maps, in: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4537–4542.

Le Bras, F., Hamel, T., Mahony, R., Samson, C., 2017. Observers for position and velocity bias estimation from single or multiple direction outputs, in: Fossen, T., Pettersen, K., Nijmeijer, H. (Eds.), Sensing and Control for Autonomous Vehicles. Lecture Notes in Control and Information Sciences 474, Springer. chapter 1.

Lourenço, P., Guerreiro, B., Batista, P., Oliveira, P., Silvestre, C., 2016. Simultaneous localization and mapping for aerial vehicles: A 3-D sensor-based GAS filter. Autonomous Robots 40, 881–902.

Mahony, R., Hamel, T., 2017. A geometric nonlinear observer for simultaneous localisation and mapping, in: Conference on Decision and Control, Melbourne. p. 6 pages.

Mahony, R., Hamel, T., Pflimlin, J.M., 2008. Non-linear complementary filters on the special orthogonal group. IEEE Transactions on Automatic Control 53, 1203–1218. doi:10.1109/TAC.2008.923738.

Mahony, R., Hamel, T., Trumpf, J., 2020. Equivariant Systems Theory and Observer Design. arXiv:2006.08276 [cs, eess] arXiv:2006.08276.

Mahony, R., Trumpf, J., Hamel, T., 2013. Observers for kinematic systems with symmetry, in: IFAC Proceedings Volumes, Elsevier. pp. 617–633.

Mei, C., Sibley, G., Cummins, M., Newman, P., Reid, I., 2011. RSLAM: A system for large-scale mapping in constant-time using stereo. International journal of computer vision 94, 198–214.

Mur-Artal, R., Montiel, J., Tardós, J., 2015. ORB-SLAM: A versatile and accurate monocular SLAM system. IEEE Transactions on Robotics 31, 1147–1163. doi:DOI:10.1109/TRO.2015.2463671.

Scaramuzza, D., Fraundorfer, F., 2011. Part I: The first 30 years and fundamentals. IEEE Robotics and Automation Magazine 18, 80–92.

Stachniss, C., Thrun, S., Leonard, J., 2016. Simultaneous localization and mapping. In b. Siciliano, in: Siciliano, B., Khatib, O. (Eds.), Springer Handbook of Robotics. second ed.. Springer. chapter 46, pp. 1153—1176.

Steenrod, N., 1951. The Topology of Fibre Bundles. Princeton: Princeton University Press.

Strasdat, H., Montiel, J., Davison, A., 2012. Visual SLAM: Why filter? Computer Vision and Image Understanding (CVIU) 30, 65–77.

Thrun, S., 2002. Exploring artificial intelligence in the new millenium, The Morgan Kaufmann Series in Artificial Inteligence. chapter Robotic mapping: A survey. doi:ISBNISBN-10:1558608117.

Tombari, F., Salti, S., Di Stefano, L., 2013. Performance evaluation of 3D keypoint detectors. International Jouranl Computer Vision 102, 198–220.

Triggs, B., McLauchlan, P., Hartley, R., Fitzgibbon, A., 2000. Bundle adjustment — A modern synthesis, in: Triggs, B., Zisserman, A., Szeliski,

R. (Eds.), Vision Algorithms: Theory and Practice. Springer, Berlin, Heidelberg. volume 1883.

Vasconcelos, J., Cunha, R., Silvestre, C., Oliveira, P., 2010. A nonlinear position and attitude observer on SE(3) using landmark measurements. Systems & Control Letters 59, 155–166. doi:DOI:10.1016/j.sysconle.2009.11.008.

Vasconcelos, J.F., Silvestre, C., Oliveira, P., 2008. A nonlinear observer for rigid body attitude estimation using vector observations, in: Proceedings of the 17th IFAC World Congress, pp. 8599–8604.

Vik, B., Fossen, T., 2001. A nonlinear observer for GPS and INS integration, in: Proceedings of the 40th IEEE Conference on Decision and Control, pp. 2956–2961.

Walcott-Bryant, A., Kaess, M., Johannsson, H., Leonard, J.J., 2012. Dynamic pose graph SLAM: Longterm mapping in low dynamic environments. Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) , 1871–1878.

Williams, B., Reid, I., 2010. On combining visual SLAM and visual odometry, in: International Conference on Robotics and Automation.

Yang, S., Scherer, S., 2019. CubeSLAM: Monocular 3-D Object SLAM. IEEE Transactions on Robotics 35, 925–938. doi:10.1109/TRO.2019.2909168.

Zhang, J., Henein, M., Mahony, R., Ila, V., 2020. VDO-SLAM: A Visual Dynamic Object-Aware SLAM System. arXiv:2005.11052 [cs] arXiv:2005.11052.

Zhang, T., Wu, K., Song, J., Huang, S., Dissanayake, G., 2017. Convergence and consistency analysis for a 3D invariant-EKF SLAM. IEEE Robotics and Automation Letters 2, 733–740.

Zlotnik, D., Forbes, J., 2018. Gradient-based observer for simultaneous localization and mapping. IEEE Transactions on Automatic Control , 8 pagesdoi:10.1109/TAC.2018.2829467.

Robert Mahony is a Professor in the School of Engineering at the Australian National University. He received his BSc in 1989 (applied mathematics and geology) and his PhD in 1995 (systems engineering) both from the Australian National University. He is a fellow of the IEEE and was president of the Australian Robotics Association from 2008-2011. He was Director of the Research School of Engineering at the Australian National University 2014-2016. His research interests are in non-linear systems theory with applications in robotics and computer vision. He is known for his work in aerial robotics, equivariant observer design, matrix subspace optimisation and image based visual servo control.

Tarek Hamel is Professor at the University of Nice Sophia Antipolis since 2003. He received his Ph.D. degree in Robotics from the University of Technology of Compiègne (UTC), France, in 1996. After two years as a research assistant at the UTC, he joined the Centre d'Etudes de Mécanique d'Ile de France in 1997 as Associate Professor. His research interests include nonlinear control theory, estimation and vision-based control with applications to Unmanned Robotic Systems. He is IEEE Fellow and senior member of the Institut Universitaire de France. He served as Associate Editor for IEEE Transactions on Robotics, IEEE Transaction on Control System Technology and for Control Engineering Practice.

Jochen Trumpf received the Dipl.-Math. and Dr. rer. nat. degrees in mathematics from the University of Würzburg, Germany, in 1997 and 2002, respectively. He is currently a Professor in the School of Engineering at the Australian National University. His research interests include observer theory and design, linear systems theory and optimization on manifolds with applications in robotics, computer vision and wireless communication. He currently serves as Associate Editor for Mathematics of Control, Signals, and Systems (MCSS).