

Constructive Observer Design for Visual Simultaneous Localisation and Mapping

Pieter van Goor^a Robert Mahony^a Tarek Hamel^b Jochen Trumpf^a

^a *Australian Centre for Robotic Vision, Research School of Electrical, Energy and Materials Engineering, Australian National University, Australia*

^b *I3S (University Côte d’Azur, CNRS, Sophia Antipolis) and Insitut Universitaire de France*

Abstract

Visual Simultaneous Localisation and Mapping (VSLAM) is a well-known problem in robotics with a large range of applications. This paper presents a novel approach to VSLAM by lifting the observer design to a novel Lie group $\mathbf{VSLAM}_n(3)$ on which the system output is equivariant. The perspective gained from this analysis facilitates the design of a non-linear observer with almost semi-globally asymptotically stable error dynamics. Simulations are provided to illustrate the behaviour of the proposed observer and experiments on data gathered using a fixed-wing UAV flying outdoors demonstrate its performance.

Key words: Equivariant Observer; Visual Odometry (VO); Simultaneous Localisation and Mapping (SLAM).

1 Introduction



Fig. 1. The Disco Parrot UAV used to gather video data to test the proposed observer.

Simultaneous Localisation and Mapping (SLAM) has been an established problem in mobile robotics for at least the last 30 years [10]. Visual SLAM (VSLAM) refers to the special case where the only exteroceptive sensors available are cameras, and is frequently used to refer to

the challenging situation where only a single monocular camera is available. The inherent non-linearity of the VSLAM problem remains challenging [8] and state-of-the-art solutions suffer from high computational complexity and poor scalability [10,21]. Due to the low cost and low weight, as well as the ubiquity of single camera systems, the VSLAM problem remains an active research topic [10,9].

Both the SLAM and VSLAM problems have recently attracted interest in the non-linear observer community, drawing from earlier work on attitude estimation [16,5] and pose estimation [1,25,12]. Barrau and Bonnabel [2] exploited a novel Lie group to design an invariant Extended Kalman Filter for the SLAM problem. Parallel work by Mahony and Hamel [17] proposed the same group structure along with a novel quotient manifold structure for the state-space of the SLAM problem. Work by Zlotnik and Forbes [26] derives a geometrically motivated observer for the SLAM problem that includes estimation of bias in linear and angular velocity inputs. For the VSLAM problem, where only bearing measurements are available, Lourenço *et al.* [14,15] proposed an observer with a globally exponentially stable error system using depths of landmarks as separate components of the observer. Bjerne *et al.* [4] use an attitude heading reference system (AHRS) to determine the orientation of the robot, and then solve the SLAM problem using a linear Kalman filter. A similar approach to VSLAM

* This research was supported by the Australian Research Council through the “Australian Centre of Excellence for Robotic Vision” CE140100016.

**Corresponding author Pieter van Goor.

Email addresses: pieter.vangoor@anu.edu.au (Pieter van Goor), robert.mahony@anu.edu.au (Robert Mahony), thamel@i3s.unice.fr (Tarek Hamel), jochen.trumpf@anu.edu.au (Jochen Trumpf).

is undertaken by Le Bras *et al.* [6]. Hamel and Samson [11] have also introduced a Riccati observer for the case where the orientation of the robot is known. Recent work by the authors [24] introduced a new symmetry structure specifically targeting the VSLAM problem but used an observer design that was a lifted version of that proposed in [11].

In this paper we present a novel non-linear equivariant observer for the VSLAM problem. The approach uses the SLAM manifold state-space proposed in [17] along with a novel symmetry Lie-group, $\mathbf{VSLAM}_n(3)$, introduced by van Goor *et al.* [24] but fully developed for the first time in this paper. We extend the results of [24] by providing equivariant group actions on the state and output spaces leading to the definition of the lifted system, a lifted observer and more importantly an intrinsic error that is globally defined. We propose a Lyapunov function expressed in the intrinsic error coordinates and use this to construct an observer for the visual SLAM problem posed on the symmetry group $\mathbf{VSLAM}_n(3)$. This is in contrast to the majority of state-of-the-art algorithms which depend on local error coordinates and local linearisation [8]. The recent IEKF results [7] exploit a global symmetry of the state-space, however, the symmetry used is not compatible with visual bearings and the resulting algorithm still depends on local linearisation of the measurement function. In our proposed algorithm, separate constant gains for landmark bearing and depth estimates are used, making the design algebraically simple and leading to low computational cost. We show that the error dynamics are *almost semi-globally asymptotically stabilisable* (Def. A.1). The resulting algorithm has low computation and memory requirements, making it ideally suited to embedded systems applications in consumer electronics.

This paper consists of six sections alongside the introduction and conclusion. Section 2 introduces key notation and identities. In Section 3, we formulate the kinematics, state-space and output of the VSLAM system, and in Section 4 we introduce the Lie group $\mathbf{VSLAM}_n(3)$ and its actions on the state and output spaces. In Section 5 we derive a non-linear observer on the Lie group, and in Sections 6 and 7 we provide the results of a simulation and a real-world experiment carried out using a Disco Parrot UAV (Figure 1). The principal contribution of the paper is theoretical and the experimental sections support this by illustrating the properties of the algorithm and demonstrating that it functions on real-world data. We do not aspire to provide a comprehensive benchmark of performance against state-of-the-art SLAM systems in the present paper.

2 Preliminaries

The special orthogonal and special Euclidean matrix Lie groups are denoted $\mathbf{SO}(3)$ and $\mathbf{SE}(3)$, respectively, with Lie algebras $\mathfrak{so}(3)$ and $\mathfrak{se}(3)$. For any column vector $\Omega = (\Omega_1, \Omega_2, \Omega_3) \in \mathbb{R}^3$, the corresponding skew-symmetric matrix is denoted

$$\Omega^\times := \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{pmatrix} \in \mathfrak{so}(3).$$

This matrix has the property that, for any $v \in \mathbb{R}^3$, $\Omega^\times v = \Omega \times v$ where $\Omega \times v$ is the vector (cross) product between Ω and v . For any unit vector $y \in \mathbb{S}^2 \subset \mathbb{R}^3$ and any vector $v \in \mathbb{R}^3$,

$$y^\times y^\times v = yy^\top v - v. \quad (1)$$

Consider a homogeneous matrix $P \in \mathbf{SE}(3)$. The notation $R_P \in \mathbf{SO}(3)$ and $x_P \in \mathbb{R}^3$ is used to represent the rotation and translation components of P , respectively; that is

$$P = \begin{pmatrix} R_P & x_P \\ 0 & 1 \end{pmatrix} \in \mathbf{SE}(3).$$

Likewise, for a matrix $U \in \mathfrak{se}(3)$, the notation $\Omega_U^\times \in \mathfrak{so}(3)$, with $\Omega_U \in \mathbb{R}^3$, and $V_U \in \mathbb{R}^3$ represent the angular and linear velocity components of U , respectively; i.e.

$$U = \begin{pmatrix} \Omega_U^\times & V_U \\ 0 & 0 \end{pmatrix} \in \mathfrak{se}(3).$$

For a background on smooth manifolds, Lie groups and their actions, the authors recommend [13, Chapter 7].

3 Problem Formulation

3.1 VSLAM State Space

Fix an arbitrary reference frame $\{0\}$. Let $P \in \mathbf{SE}(3)$ and $p_i \in \mathbb{R}^3, i = 1, \dots, n$ represent the robot pose and landmark coordinates, respectively, defined with respect to $\{0\}$. The raw coordinates of the SLAM problem are written $(P, p_1, \dots, p_n) \in \mathbf{SE}(3) \times \mathbb{R}^3 \times \dots \times \mathbb{R}^3$. The notation $(P, p_i) \equiv (P, p_1, \dots, p_n)$ is used for simplicity in the sequel.

The physical measurements in a monocular VSLAM system are the bearings (3D directions) of landmarks perceived by the robot. We assume from now on that the

observed landmarks and the robot are not collocated to ensure that the bearing measurements are well defined. Interestingly, this assumption has a substantive impact on the nature of the global symmetries that can be admitted. We make this assumption explicit, defining the *total space* $\mathcal{T}_n^\circ(3)$ of SLAM configurations considered to be

$$\mathcal{T}_n^\circ(3) = \{(P, p_i) \in \mathbf{SE}(3) \times \mathbb{R}^3 \times \cdots \times \mathbb{R}^3 \mid p_i \neq x_P, i = 1, \dots, n\}. \quad (2)$$

We assume that the trajectory of the robot remains in $\mathcal{T}_n^\circ(3)$ for all time.

Given two configurations $(P, p_i), (P', p'_i) \in \mathcal{T}_n^\circ(3)$, then $(P, p_i) \simeq (P', p'_i)$ if there exists $S \in \mathbf{SE}(3)$ such that $(P, p_i) = (S^{-1}P', R_S^\top(p'_i - x_S))$. That is, two sets of coordinates in the total space $\mathcal{T}_n^\circ(3)$ are considered equivalent when they are related by a rigid body transformation of the reference frame $\{0\}$. It is straightforward to show the relation \simeq is an equivalence relation on $\mathcal{T}_n^\circ(3)$ and

$$[P, p_i] := \{(S^{-1}P, R_S^\top(p_i - x_S)) \mid S \in \mathbf{SE}(3)\}.$$

is the associated equivalence class. The VSLAM manifold is the set

$$\mathcal{M}_n^\circ(3) = \{[P, p_i] \mid (P, p_i) \in \mathcal{T}_n^\circ(3)\},$$

with quotient manifold structure. This is the open subset of the SLAM manifold $\mathcal{M}_n(3)$ considered in [17] without those equivalence classes where a landmark is co-located with the robot. Two configurations are equivalent on the SLAM manifold, $(P^1, p_i^1) \simeq (P^2, p_i^2)$, if and only if the ego-centric coordinates of the landmarks are equal, $R_{P^1}^\top(p_i^1 - x_{P^1}) = R_{P^2}^\top(p_i^2 - x_{P^2})$ for all i .

3.2 VSLAM Kinematics

Assume that the robot is moving in a static environment. Define a velocity input vector space $\mathbb{V} = \mathfrak{se}(3)$ to contain the rigid-body velocity $U \in \mathfrak{se}(3)$ of the robot. The kinematics of the VSLAM system are given by the system function $f : \mathcal{T}_n^\circ(3) \times \mathbb{V} \rightarrow \mathbf{T}\mathcal{T}_n^\circ(3)$,

$$\begin{aligned} \frac{d}{dt}(P, p_i) &= f((P, p_i), U), \\ &:= (PU, 0). \end{aligned} \quad (3)$$

3.3 System Output

The camera measurements are modelled as elements of the sphere S^2 . Each individual bearing is given by an

output function $h^i : \mathcal{T}_n^\circ(3) \rightarrow S^2$,

$$h^i(P, p_i) := \frac{R_P^\top(p_i - x_P)}{\|p_i - x_P\|}, \quad (4)$$

The output functions are well defined on $\mathcal{M}_n^\circ(3)$ since

$$\begin{aligned} h^i(S^{-1}P, R_S^\top(p_i - x_S)) &= \frac{(R_S^\top R_P)^\top (R_S^\top(p_i - x_S) - R_S^\top(x_P - x_S))}{\|R_S^\top(p_i - x_S) - R_S^\top(x_P - x_S)\|}, \\ &= \frac{R_P^\top(p_i - x_P)}{\|p_i - x_P\|}. \end{aligned} \quad (5)$$

The full output space of the VSLAM system is defined as a product of n spheres,

$$\mathcal{X}^n(3) := S^2 \times \cdots \times S^2,$$

with a combined output function $h : \mathcal{T}_n^\circ(3) \rightarrow \mathcal{X}^n(3)$

$$h(P, p_i) := \left(\frac{R_P^\top(p_1 - x_P)}{\|p_1 - x_P\|}, \dots, \frac{R_P^\top(p_n - x_P)}{\|p_n - x_P\|} \right). \quad (6)$$

4 Symmetry of the VSLAM Problem

4.1 Symmetry of the Total Space $\mathcal{T}_n^\circ(3)$

Define a Lie group

$$\mathbf{VSLAM}_n(3) = SE(3) \times (\mathbf{SO}(3) \times \mathbf{MR}(1))^n,$$

with product Lie group structure, where $\mathbf{MR}(1)$ is the multiplicative real group of positive real numbers. This Lie group was first proposed in van Goor *et al.* [24]. The associated Lie algebra is denoted $\mathfrak{vslam}_n(3)$. We write elements of $\mathbf{VSLAM}_n(3)$ as $(A, (Q, a)_i) \equiv (A, (Q, a)_1, \dots, (Q, a)_n) \in \mathbf{VSLAM}_n(3)$. The group product, identity and inverse are given by

$$\begin{aligned} (A_1, (Q_1, a_1)_i) \cdot (A_2, (Q_2, a_2)_i) &= (A_1 A_2, (Q_1 Q_2, a_1 a_2)_i), \\ \text{id} &= (I_4, (I_3, 1)_i), \quad (A, (Q, a)_i)^{-1} = (A^{-1}, (Q^\top, a^{-1})_i). \end{aligned}$$

Lemma 4.1 *The mapping $\Upsilon : \mathbf{VSLAM}_n(3) \times \mathcal{T}_n^\circ(3) \rightarrow \mathcal{T}_n^\circ(3)$ defined by*

$$\begin{aligned} \Upsilon((A, (Q, a)_i), (P, p_i)) &:= (PA, a_i^{-1} R_{PA} Q_i^\top R_P^\top (p_i - x_P) + x_{PA}), \end{aligned} \quad (7)$$

is a transitive right group action of $\mathbf{VSLAM}_n(3)$ on $\mathcal{T}_n^\circ(3)$.

PROOF. Trivially, $\Upsilon((I_4, (I_3, 1)_i), (P, p_i)) = (P, p_i)$ for any $(P, p_i) \in \mathcal{T}_n^\circ(3)$. Let $(A_1, (Q_1, a_1)_i), (A_2, (Q_2, a_2)_i) \in$

$\mathbf{VSLAM}_n(3)$ and $(P, p_i) \in \mathcal{T}_n^\circ(3)$ be arbitrary. Then

$$\begin{aligned} & \Upsilon((A_1, (Q_1, a_1)_i), \Upsilon((A_2, (Q_2, a_2)_i), (P, p_i))) \\ &= (PA_2A_1, (a_1^{-1}a_2^{-1}R_{PA_2A_1}Q_1^\top \\ & \quad Q_2^\top R_P^\top(p - x_P) + x_{PA_2A_1})_i), \\ &= \Upsilon((A_2A_1, (Q_2Q_1, a_2a_1)_i), (P, p_i)), \\ &= \Upsilon((A_2, (Q_2, a_2)_i) \cdot (A_1, (Q_1, a_1)_i), (P, p_i)). \end{aligned}$$

Thus Υ is a group action. To see that Υ is transitive, let $(P, p_i), (P', p'_i) \in \mathcal{T}_n^\circ(3)$ be arbitrary. Choose $(A, (Q, a)_i) \in \mathbf{VSLAM}_n(3)$ to satisfy

$$\begin{aligned} A &= P^{-1}P', \quad a_i = \frac{\|p_i - x_P\|}{\|p'_i - x_{P'}\|}, \\ Q_i \frac{R_{P'}^\top(p'_i - x_{P'})}{\|p'_i - x_{P'}\|} &= \frac{R_P^\top(p_i - x_P)}{\|p_i - x_P\|}. \end{aligned}$$

Then $\Upsilon((A, (Q, a)_i), (P, p_i)) = (P', p'_i)$.

The action Υ of $\mathbf{VSLAM}_n(3)$ on $\mathcal{T}_n^\circ(3)$ is shown in Figure 2. Given $(A, (Q_A, a)_i) \in \mathbf{VSLAM}_n(3)$ and $(P, p_i) \in \mathcal{T}_n^\circ(3)$, the action transforms the robot pose P by right-translation by A . The landmark points are transformed by considering their body-fixed coordinates; applying a rotation Q_i^\top and scaling a_i^{-1} ; transforming them along with the robot pose; and finally writing the result in inertial coordinates.

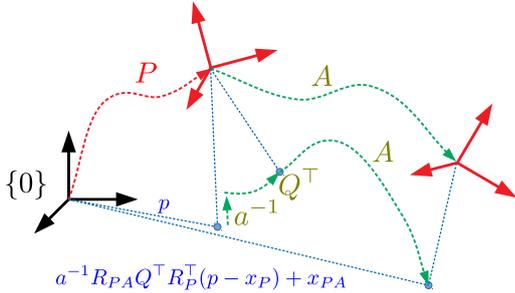


Fig. 2. The action Υ of the VSLAM group on the state space for some given $(A, (Q_A, a)_i) \in \mathbf{VSLAM}_n(3)$ and $(P, p_i) \in \mathcal{T}_n^\circ(3)$. The pose P is mapped to PA . The body fixed frame landmark points $R_P^\top(p_i - x_P)$ are rotated by Q_i^\top and scaled by a_i^{-1} in the body-fixed frame before *transforming with the robot pose* to a new point p'_i which is then rewritten in the inertial frame.

4.2 Symmetry of the Output Space

There is an action ρ of the $\mathbf{VSLAM}_n(3)$ group on the output space such that the measurement function h defined in (6) is equivariant with respect to Υ and ρ . The following lemmas define this action ρ and show the equivariance of h for the proposed $\mathbf{VSLAM}_n(3)$ geometry. The authors know of no output action that makes bearing outputs equivariant with respect to prior geometries

proposed for SLAM [2,17]. The equivariance structure enables the development of a globally-defined intrinsic error.

Proposition 4.2 *The mapping $\rho : \mathbf{VSLAM}_n(3) \times \mathcal{X}^n(3) \rightarrow \mathcal{X}^n(3)$ defined by*

$$\rho((A, (Q, a)_i), y_i) = Q_i^\top y_i, \quad (8)$$

is a right group action of $\mathbf{VSLAM}_n(3)$ on $\mathcal{X}^n(3)$.

PROOF. The proof is straightforward.

Lemma 4.3 *The output $h : \mathcal{T}_n^\circ(3) \rightarrow \mathcal{X}^n(3)$ (6) is equivariant with respect to actions Υ (7) and ρ (8). That is, for any $X \in \mathbf{VSLAM}_n(3)$ and any $\xi \in \mathcal{T}_n^\circ(3)$,*

$$h(\Upsilon(X, \xi)) = \rho(X, h(\xi)).$$

PROOF. Let $X = (A, (Q, a)_i)$ and $\xi = (P, p_i)$ be arbitrary. Then

$$\begin{aligned} & h(\Upsilon(X, \xi)) \\ &= h(PA, a_i^{-1}R_{PA}Q_i^\top R_P^\top(p_i - x_P) + x_{PA}), \\ &= \frac{R_{PA}^\top(a_i^{-1}R_{PA}Q_i^\top R_P^\top(p_i - x_P) + x_{PA} - x_{PA})}{\|a_i^{-1}R_{PA}Q_i^\top R_P^\top(p_i - x_P) + x_{PA} - x_{PA}\|}, \\ &= \frac{Q_i^\top R_P^\top(p_i - x_P)}{\|p_i - x_P\|}, \\ &= \rho(X, h(\xi)). \end{aligned}$$

4.3 Lift of the VSLAM Kinematics

In order to consider the system on the $\mathbf{VSLAM}_n(3)$ group, the kinematics from the total space must be lifted onto the group. A lift is a map $\Lambda : \mathcal{T}_n^\circ(3) \times \mathbb{V} \rightarrow \mathfrak{vslam}_n(3)$ such that

$$D\Upsilon_{(P, p_i)}(\text{id})[\Lambda((P, p_i), U)] = f((P, p_i), U) \quad (9)$$

where $f((P, p_i), U)$ is given by (3).

Lemma 4.4 *The function $\Lambda : \mathcal{T}_n^\circ(3) \times \mathbb{V} \rightarrow \mathfrak{vslam}_n(3)$, defined by*

$$\Lambda((P, p_i), U) := (U, (\Lambda_Q(U, R_P^\top(p_i - x_P)), \Lambda_a(U, R_P^\top(p_i - x_P)))), \quad (10)$$

where $\Lambda_Q : \mathfrak{se}(3) \times (\mathbb{R}^3 \setminus \{0\}) \rightarrow \mathfrak{so}(3)$ is given by

$$\Lambda_Q(U, q) := \left(\Omega_U + \frac{q \times V_U}{|q|^2} \right)^\times,$$

and $\Lambda_a : \mathfrak{se}(3) \times (\mathbb{R}^3 \setminus \{0\}) \rightarrow \mathfrak{mr}(1)$ is given by

$$\Lambda_a(U, q) := \frac{q^\top V_U}{|q|^2},$$

is a lift in the sense of (9) of the kinematics (3) onto $\mathfrak{vslam}_n(3)$ with respect to the group action (7).

PROOF. Given $(P, p_i) \in \mathcal{T}_n^\circ(3)$, let $q_i := R_P^\top(p_i - x_P)$, $W_i^\times := \Lambda_Q(U, q_i)$, and $w_i := \Lambda_a(U, q_i)$. Then

$$\begin{aligned} \text{DY}_{(P, p_i)}(\text{id}) [\Lambda((P, p_i), U)] \\ &= \text{DY}_{(P, p_i)}(\text{id}) [(U, (W_i^\times, w_i))] \\ &= (PU, v_i) \end{aligned} \quad (11)$$

where v_i is

$$v_i = \left. \frac{d}{ds} \right|_{s=0} \left[(1 + sw_i)^{-1} R_{P(I_4 + sU)} (I_3 + sW_i^\times)^\top q_i + x_{P(I_4 + sU)} \right].$$

Computing this derivative and evaluating at $s = 0$ one obtains

$$\begin{aligned} v_i &= -w_i R_P q_i + R_P \Omega_U^\times q_i + R_P (W_i^\times)^\top q_i + R_P V_U, \\ &= -\frac{q_i^\top V_U}{|q_i|^2} R_P q_i + R_P \Omega_U^\times q_i + R_P V_U \\ &\quad - R_P \left(\Omega_U + \frac{q_i \times V_U}{|q_i|^2} \right)^\times q_i, \quad (12) \\ &= -R_P \frac{q_i q_i^\top}{|q_i|^2} V_U + R_P V_U + R_P \frac{q_i q_i^\top}{|q_i|^2} V_U - R_P V_U, \\ &= 0, \end{aligned} \quad (13)$$

where (12) follows from substituting for w_i and W_i with the full expressions for Λ_a and Λ_Q , and (13) follows from cancelling the Ω_U term and using the relationship (1) as well as rearranging the first term. The result follows from substituting directly into (11) and recalling (3).

5 Observer Design

Figure 3 shows a schematic overview of the proposed observer. The key features of equivariant observer design are the distinction between the observer state $\hat{X} \in \mathfrak{VSLAM}_n(3)$ and the estimated state $(\hat{P}, \hat{p}_i) \in \mathcal{T}_n^\circ(3)$, and the design of the correction term around the output error $(d_i) = \rho(\hat{X}^{-1}, y_i)$ rather than the raw measurements $(y_i) \in \mathcal{X}^n(3)$.

5.1 Lifted System Kinematics

Let $\xi = (P, p_i) \in \mathcal{T}_n^\circ(3)$ denote the true state of the VSLAM system. The kinematics of ξ are given by (3). Choose an arbitrary origin configuration $\xi^\circ = (P^\circ, p_i^\circ) \in \mathcal{T}_n^\circ(3)$. The lifted system is (Lemma 4.4)

$$\begin{aligned} \frac{d}{dt} X &= X \Lambda(\Upsilon(X, \xi^\circ), U), \\ &= (AU, (Q\Lambda_Q, a\Lambda_a)_i), \end{aligned} \quad (14)$$

for $X = (A, (Q, a)_i) \in \mathfrak{VSLAM}_n(3)$. If $\Upsilon(X(0), \xi^\circ) = \xi(0)$ then trajectories of the lifted system kinematics project to trajectories of the VSLAM kinematics (3) [18]. That is, recalling (7)

$$\begin{aligned} (P(t), p_i(t)) &= \Upsilon(X(t), \xi^\circ) \\ &= (P^\circ A, a_i^{-1} R_{P^\circ A} Q_i^\top R_{P^\circ}^\top (p_i^\circ - x_{P^\circ}) + x_{P^\circ A}) \end{aligned}$$

for all $t \geq 0$, where we have dropped the time dependence of A , Q_i and a_i from the notation to improve readability.

5.2 Landmark Observer

Let $X = (A, (Q, a)_i)$ be a trajectory of the lifted system (14) associated with a trajectory $\xi = (P, p_i)$ of the true system satisfying (3) for measured input signal $U = (\Omega_U^\times, V_U)$. Let $y_i := h^i(\xi)$ (4) denote the output.

Fix an arbitrary origin configuration $\xi^\circ = (P^\circ, p_i^\circ) \in \mathcal{T}_n^\circ(3)$. The observer state is $\hat{X} = (\hat{A}, (\hat{Q}, \hat{a})_i) \in \mathfrak{VSLAM}_n(3)$, with kinematics given by

$$\begin{aligned} \frac{d}{dt} \hat{X} &:= \hat{X} \Lambda(\Upsilon(\hat{X}, \xi^\circ), U) - \Delta_{\hat{X}} \hat{X}, \\ \hat{X}(0) &= \text{id}, \end{aligned} \quad (15)$$

where Λ is the lift defined by (10) and $\Delta_{\hat{X}} = (\Delta, (\Gamma, \gamma)_i) \in \mathfrak{vslam}_n(3)$ is a correction term that is chosen later. The state estimate is given by $\hat{\xi} = (\hat{P}, \hat{p}_i) = \Upsilon(\hat{X}, \xi^\circ)$ (7). Let $y_i^\circ := h^i(\xi^\circ)$ denote the origin output, and let d_i denote the output error [18], defined as

$$d_i = \rho(\hat{X}^{-1}, y_i) = \rho(X \hat{X}^{-1}, y_i^\circ). \quad (16)$$

Define the true range $r_i = \|p_i - x_P\|$ and estimated range $\hat{r}_i = \|\hat{p}_i - x_{\hat{P}}\|$ for each i . The range error is

$$\tilde{r}_i = \frac{\hat{r}_i}{r_i}, \quad (17)$$

for each i .

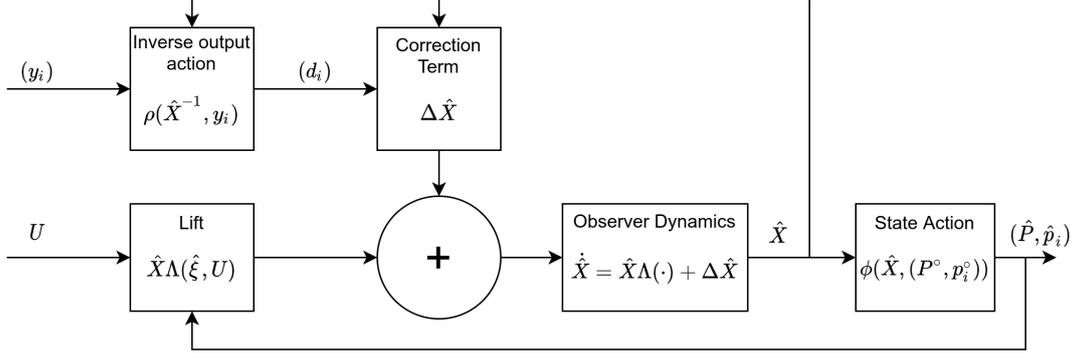


Fig. 3. An overview of the observer system. Note the distinction between the observer state \hat{X} and the estimated state (\hat{P}, \hat{p}_i) , and that the correction term Δ is based on the output error (d_i) rather than the raw measurement (y_i) .

Define the twice differentiable barrier function $\beta_{\underline{c}}^\epsilon : (\epsilon, \infty) \rightarrow [0, \infty)$ to be

$$\beta_{\underline{c}}^\epsilon(c) := \begin{cases} \frac{(c-\epsilon)^2}{(c-\epsilon)^2(c-\underline{c})}, & \epsilon < c < \underline{c}, \\ 0, & c \geq \underline{c} \end{cases}, \quad (18)$$

for parameters $0 < \epsilon < \underline{c}$.

The landmark correction terms Γ_i and γ_i for $\Delta_{\hat{X}}$ are defined to be

$$\begin{aligned} \Gamma_i &:= \left(\frac{d_i^\top \hat{Q}_i V_U}{\hat{r}_i (1 + d_i^\top y_i^\circ)} - \frac{k_i}{(1 + d_i^\top y_i^\circ)^2} \right) (d_i^\times y_i^\circ)^\times \\ &\quad + \frac{1}{\hat{r}_i} \left((y_i^\circ - d_i)^\times \hat{Q}_i V_U \right)^\times, \\ \gamma_i &:= \frac{\alpha_i}{\hat{r}_i^2} \left((1 - d_i^\top y_i^\circ) d_i^\top \hat{Q}_i V_U - y_i^{\circ\top} (d_i \times \hat{Q}_i V_U)^\times d_i \right) \\ &\quad + \frac{1}{\hat{r}_i} (y_i^\circ - d_i)^\top \hat{Q}_i V_U + \frac{\alpha_i}{\hat{r}_i} \beta_{\underline{c}}^\epsilon(\hat{r}_i), \end{aligned} \quad (19)$$

where k_i and α_i are constant positive scalars. The robot pose correction term $\Delta \in \mathbf{SE}(3)$ can be chosen to be any continuous function of the observer state and measurements (cf. §5.3).

Theorem 1 Consider the observer $\hat{X} \in \mathbf{VSLAM}_n(3)$ with kinematics (15). Assume that U is bounded, and that $y_i^\times y_i^\times V_U$ is persistently exciting, in the sense that there exist $T > 0$ and $\mu > 0$ such that

$$\frac{1}{T} \int_t^{t+T} \|y_i^\times y_i^\times V_U\| d\tau \geq \mu, \quad (20)$$

for each i and all $t > 0$. Assume that there exist bounds $0 < \underline{r} < \bar{r} \in \mathbb{R}$ such that

$$r \leq \|p_i - x_P\| \leq \bar{r}$$

for all time.

Then the landmark correction terms (19) define an almost semi-globally asymptotically stabilising (Def. A.1) correction term for the error dynamics of d_i, \tilde{r}_i (16,17) around the equilibrium $(y_i^\circ, 1)$ with exception set

$$\chi = \{(d, \tilde{r})_i \in (\mathbb{S}^2 \times \mathbb{R}_+)^n \mid d_i = -y_i^\circ \text{ for some } i \in [1, \dots, n]\}. \quad (21)$$

Moreover, as $(d_i, \tilde{r}_i) \rightarrow (y_i^\circ, 1)$, the estimated state $\hat{\xi} \rightarrow \xi$ converges to the true state up to the SLAM manifold equivalence.

PROOF. The outline of the proof is as follows. We begin by choosing the parameters of the correction terms to depend on the initial conditions. Next, it is shown that the observer equations are well-defined for all time. We introduce storage functions in (25) and proceed to show that they are non-increasing over time in (26). Then we apply Barbalat's lemma along with persistence of excitation to show that the storage functions converge to zero over time, leading to (27) and (28). Finally, we show that convergence of the error dynamics is equivalent to convergence of the true and estimated states on the SLAM manifold.

Let K be a compact set in the complement of χ (21). Then there exists $\tilde{r}^m > 0$ such that $(d_i, \tilde{r}_i) \in K$ implies that $\tilde{r}_i \geq \tilde{r}^m$. Choose $k_0 = \frac{1}{2}\tilde{r}^m$ and assume that $(d_i(0), \tilde{r}_i(0)) \in K$. Then

$$\tilde{r}_i(0) \geq k_0, \quad \hat{r}_i(0) \geq k_0 r_i(0) \geq k_0 \underline{r} > 0,$$

for each i . Choose $\underline{c} = \underline{r}$ and $\epsilon = \min(k_0, \frac{1}{2})\underline{r}$. Then $\hat{r}_i(0) > \epsilon > 0$ for every i . By continuity of the solutions there exist $T_i > 0$ such that $\hat{r}_i(t) > \epsilon$ and $1 + d_i^\top y_i^\circ > 0$, and hence $\beta_{\underline{c}}^\epsilon(\hat{r}_i)$ and Γ_i are well-defined, for $t \in [0, T_i)$. (We will show later that T_i can be chosen arbitrarily large, and that $\hat{r}_i(t) > \epsilon$ and $1 + d_i^\top y_i^\circ > 0$ hold for all time.)

By definition, $r_i = a_i^{-1}r_i^\circ$, and $\hat{r}_i = \hat{a}_i^{-1}r_i^\circ$. Differentiating these with respect to time yields

$$\begin{aligned}\dot{r}_i &= -V_U^\top y_i, \\ &= -d_i^\top \hat{Q}_i V_U, \\ \dot{\hat{r}}_i &= -V_U^\top \hat{y}_i + \hat{r}_i \gamma_i, \\ &= -V_U^\top y_i + \alpha_i \beta_r^\epsilon(\hat{r}_i) \\ &\quad + \frac{\alpha_i}{\hat{r}_i} \left((1 - d_i^\top y_i^\circ) d_i^\top \hat{Q}_i V_U - y_i^{\circ\top} (d_i \times \hat{Q}_i V_U)^\times d_i \right),\end{aligned}\quad (22)$$

where $\hat{y}_i = h^i(\hat{\xi})$ is the estimated measurement. Since β_r^ϵ (18) is a barrier function ensuring that $\hat{r}_i(t) > \epsilon$ ($\beta_r^\epsilon(\hat{r}_i) \rightarrow \infty$ as $\hat{r}_i \searrow \epsilon$) and the remaining terms in (22) are bounded, it follows that \hat{r}_i is well defined $\forall t \in [0, T_i)$.

Differentiating the output error d_i yields

$$\begin{aligned}\dot{d}_i &= \frac{d}{dt} \rho(X \hat{X}^{-1}, y^\circ)_i = \frac{d}{dt} \hat{Q}_i Q_i^\top y_i^\circ, \\ &= \left(\hat{Q}_i \Lambda_Q(U, R_{\hat{P}}^\top(\hat{p}_i - x_{\hat{P}})) - \Gamma_i \hat{Q}_i \right) Q_i^\top y_i^\circ \\ &\quad - \hat{Q}_i \Lambda_Q(U, R_P^\top(p_i - x_P)) Q_i^\top y_i^\circ.\end{aligned}$$

Observe that $R_{\hat{P}}^\top(\hat{p}_i - x_{\hat{P}}) = \hat{r}_i \hat{y}_i$ and $R_P^\top(p_i - x_P) = r_i y_i$. Using this, the derivative of d_i is simplified to

$$\begin{aligned}\dot{d}_i &= \text{Ad}_{\hat{Q}_i} (\Lambda_Q(U, \hat{r}_i \hat{y}_i) - \Lambda_Q(U, r_i y_i)) d_i - \Gamma_i d_i, \\ &= \text{Ad}_{\hat{Q}_i} \left(\frac{\hat{y}_i \times V_U}{\hat{r}_i} - \frac{y_i \times V_U}{r_i} \right)^\times d_i - \Gamma_i d_i, \\ &= \left(\frac{y_i^{\circ\top} \hat{Q}_i V_U}{\hat{r}_i} - \frac{d_i^\times \hat{Q}_i V_U}{r_i} \right)^\times d_i - \Gamma_i d_i,\end{aligned}\quad (23)$$

where the last step is obtained by using various identities involving the skew symmetric operator.

Recalling (23) and using identities involving the skew symmetric operator and projector, the derivative of 1 -

$y_i^{\circ\top} d_i$ is

$$\begin{aligned}\frac{d}{dt} (1 - y_i^{\circ\top} d_i) &= -y_i^{\circ\top} \dot{d}_i, \\ &= -y_i^{\circ\top} \left(\frac{y_i^{\circ\top} \hat{Q}_i V_U}{\hat{r}_i} - \frac{d_i^\times \hat{Q}_i V_U}{r_i} \right)^\times d_i + y_i^{\circ\top} \Gamma_i d_i, \\ &= -y_i^{\circ\top} \left(\frac{d_i^\times \hat{Q}_i V_U}{\hat{r}_i} - \frac{d_i^\times \hat{Q}_i V_U}{r_i} \right)^\times d_i \\ &\quad + \left(\frac{d_i^\top \hat{Q}_i V_U}{\hat{r}_i (1 + d_i^\top y_i^\circ)} - \frac{k_i}{(1 + d_i^\top y_i^\circ)^2} \right) y_i^{\circ\top} (d_i^\times y_i^\circ)^\times d_i, \\ &= (r_i^{-1} - \hat{r}_i^{-1}) y_i^{\circ\top} (d_i^\times \hat{Q}_i V_U)^\times d_i \\ &\quad + \left(\frac{d_i^\top \hat{Q}_i V_U}{\hat{r}_i (1 + d_i^\top y_i^\circ)} - \frac{k_i}{(1 + d_i^\top y_i^\circ)^2} \right) y_i^{\circ\top} \Pi_{d_i} y_i^\circ, \\ &= (r_i^{-1} - \hat{r}_i^{-1}) y_i^{\circ\top} (d_i^\times \hat{Q}_i V_U)^\times d_i \\ &\quad + \left(\frac{d_i^\top \hat{Q}_i V_U}{\hat{r}_i} - \frac{k_i}{1 + d_i^\top y_i^\circ} \right) (1 - d_i^\top y_i^\circ).\end{aligned}\quad (24)$$

Observe that $k_i \frac{1 - d_i^\top y_i^\circ}{1 + d_i^\top y_i^\circ} \rightarrow \infty$ as $1 - d_i^\top y_i^\circ \nearrow 2$. Since all other terms in (24) are bounded, it follows that $1 - d_i^\top y_i^\circ < 2 - \nu$ for some small $\nu > 0$, and hence $1 + d_i^\top y_i^\circ > \nu > 0$ and Γ_i is well-defined and bounded for all $t \in [0, T_i)$.

We show by contradiction that the domain of definition $[0, T_i)$ can be extended arbitrarily. Suppose, for some i , that T_i' is the largest value such that Γ_i and $\beta_r^\epsilon(\hat{r}_i)$ are well-defined. Then Γ_i and $\beta_r^\epsilon(\hat{r}_i)$ are both bounded on $[0, T_i')$ by the arguments above, and continuous. It follows that their limits as $t \rightarrow T_i'$ exist and are finite. But then, by continuity of solutions, Γ_i and $\beta_r^\epsilon(\hat{r}_i)$ can be extended to $[0, T_i' + t_\Delta)$ for some sufficiently small $t_\Delta > 0$. This contradicts the assumption that T_i' is the maximum value for which Γ_i and $\beta_r^\epsilon(\hat{r}_i)$ are well-defined on $[0, T_i')$, and therefore no such maximum value can exist. Hence Γ_i and $\beta_r^\epsilon(\hat{r}_i)$ are well-defined on $[0, \infty)$, and so are the observer dynamics.

For each i , define the storage function

$$\begin{aligned}l_i(d_i, \tilde{r}_i; r_i) &:= \frac{r_i}{2} \|y_i^\circ - d_i\|^2 + \frac{r_i^2}{2\alpha_i} (1 - \tilde{r}_i)^2 \\ &= r_i (1 - y_i^{\circ\top} d_i) + \frac{1}{2\alpha_i} (r_i - \tilde{r}_i)^2,\end{aligned}\quad (25)$$

in the variables (d_i, \tilde{r}_i) where the second line follows from $\|d_i\| = \|y_i^\circ\| = 1$ and substituting for \tilde{r}_i . Note that l_i depends on the time varying range r_i as a parameter. By assumption, $r_i \geq \underline{r} > 0$ and the storage functions l_i are positive definite.

The derivative of l_i may now be computed as follows:

$$\begin{aligned}
\dot{l}_i &= \dot{r}_i(1 - y_i^{\circ\top} d_i) + r_i \frac{d}{dt}(1 - y_i^{\circ\top} d_i) \\
&\quad + \frac{1}{\alpha_i}(r_i - \hat{r}_i)(\dot{r}_i - \dot{\hat{r}}_i), \\
&= -(1 - y_i^{\circ\top} d_i) d_i^\top \hat{Q}_i V_U \\
&\quad + r_i(r_i^{-1} - \hat{r}_i^{-1}) y_i^{\circ\top} \left(d_i^\times \hat{Q}_i V_U \right)^\times d_i \\
&\quad + r_i \left(\frac{d_i^\top \hat{Q}_i V_U}{\hat{r}_i} - \frac{k_i}{1 + d_i^\top y_i^\circ} \right) (1 - d_i^\top y_i^\circ) \\
&\quad + (\hat{r}_i - r_i) \beta_r^\epsilon(\hat{r}_i) \\
&\quad + (\hat{r}_i - r_i) \hat{r}_i^{-1} (1 - d_i^\top y_i^\circ) d_i^\top \hat{Q}_i V_U \\
&\quad - (\hat{r}_i - r_i) \hat{r}_i^{-1} y_i^{\circ\top} (d_i \times \hat{Q}_i V_U)^\times d_i, \\
&= r_i(r_i^{-1} - \hat{r}_i^{-1}) y_i^{\circ\top} \left(d_i^\times \hat{Q}_i V_U \right)^\times d_i \\
&\quad - k_i r_i \frac{1 - d_i^\top y_i^\circ}{1 + d_i^\top y_i^\circ} + (\hat{r}_i - r_i) \beta_r^\epsilon(\hat{r}_i) \\
&\quad - (\hat{r}_i - r_i) \hat{r}_i^{-1} y_i^{\circ\top} (d_i \times \hat{Q}_i V_U)^\times d_i, \\
&= -k_i r_i \frac{1 - d_i^\top y_i^\circ}{1 + d_i^\top y_i^\circ} + (\hat{r}_i - r_i) \beta_r^\epsilon(\hat{r}_i). \tag{26}
\end{aligned}$$

The second term is negative semi-definite, since it is zero when $\hat{r}_i \geq \underline{r}$ (18), and otherwise $\hat{r}_i < \underline{r} \leq r_i$ and $\hat{r}_i - r_i \leq 0$. It follows that $\dot{l}_i(t) \leq 0$, and $l_i(t) \leq l_i(0)$ for all t .

Barbalat's Lemma [20, Lemma 4.2] is used to prove $\dot{l}_i \rightarrow 0$. To show that \dot{l}_i is uniformly continuous, it is sufficient to show that \ddot{l}_i is bounded. Recall that r_i is bounded above and below by assumption, and that $\hat{r}_i(t) > \epsilon$ for all time. Computing the second derivative of l_i , one has

$$\ddot{l}_i = -k_i \dot{r}_i \frac{1 - y_i^{\circ\top} d_i}{1 + y_i^{\circ\top} d_i} - \frac{2r_i y_i^{\circ\top} \dot{d}_i}{(1 + y_i^{\circ\top} d_i)^2} + (\hat{r}_i - r_i) \frac{\partial \beta_r^\epsilon}{\partial \hat{r}_i} \dot{\hat{r}}_i.$$

It suffices to show that the component terms are all bounded. In the first and second terms, $1/(1 + y_i^{\circ\top} d_i)$ is bounded from the above discussion. The components $r_i y_i^{\circ\top} \dot{d}_i$ and $\dot{r}_i(1 - y_i^{\circ\top} d_i)$ are also bounded due to the boundedness of \hat{r}_i , r_i , and the assumption that U is bounded. As for the third term, the component $\dot{\hat{r}}_i$ is bounded since the velocity input V_U is bounded and \hat{r}_i is lower bounded. This means that both $\beta_r^\epsilon(\hat{r}_i)$ and its derivative with respect to \hat{r}_i are bounded. Therefore, \dot{l}_i is uniformly continuous, and by Barbalat's lemma, $\dot{l}_i \rightarrow 0$. This implies that $d_i \rightarrow y_i^\circ$, since both terms that appear in \dot{l}_i (26) are non-positive.

It remains to show $\tilde{r}_i \rightarrow 1$ (or equivalently $\hat{r}_i \rightarrow r_i$). Applying Barbalat's Lemma to $d_i(t)$ and exploiting the same bounding arguments as before, it is straightforward

to verify that $\dot{d}_i \rightarrow 0$. It is also easily verified that $\Gamma_i \rightarrow 0$ as $d_i \rightarrow y_i^\circ$.

From (23), \dot{d}_i may be written as

$$\begin{aligned}
\dot{d}_i &= \left(\frac{y_i^{\circ\top} \hat{Q}_i V_U}{\hat{r}_i} - \frac{d_i^\times \hat{Q}_i V_U}{r_i} \right)^\times d_i - \Gamma_i d_i, \\
&= \frac{d_i^\times d_i^\times \hat{Q}_i V_U}{r_i} - \frac{d_i^\times y_i^{\circ\top} \hat{Q}_i V_U}{\hat{r}_i} - \Gamma_i d_i, \\
&= (r_i^{-1} - \hat{r}_i^{-1}) d_i^\times d_i^\times \hat{Q}_i V_U + \frac{d_i^\times (d_i - y_i^\circ)^\times \hat{Q}_i V_U}{\hat{r}_i} - \Gamma_i d_i, \\
&= (r_i^{-1} - \hat{r}_i^{-1}) \hat{Q}_i y_i^\times y_i^\times V_U + \frac{d_i^\times (d_i - y_i^\circ)^\times \hat{Q}_i V_U}{\hat{r}_i} - \Gamma_i d_i.
\end{aligned}$$

Clearly, $\hat{r}_i^{-1} d_i^\times (d_i - y_i^\circ)^\times \hat{Q}_i V_U \rightarrow 0$ as $d_i \rightarrow y_i^\circ$. Hence

$$\dot{d}_i \rightarrow (r_i^{-1} - \hat{r}_i^{-1}) \left(\hat{Q}_i y_i^\times y_i^\times V_U \right), \tag{27}$$

as $d_i \rightarrow y_i^\circ$. Therefore, $(\hat{r}_i - r_i) \frac{1}{\hat{r}_i r_i} \|y_i^\times y_i^\times V_U\| \rightarrow 0$. Recall that $r_i \leq \bar{r}$. The estimated range \hat{r}_i is also bounded above by a constant \bar{r}_i depending on the initial value of $l_i(0)$. Since $|r_i - \hat{r}_i| < (\bar{r} \bar{r}_i) |r_i^{-1} - \hat{r}_i^{-1}|$, it follows from (27) that

$$(r_i - \hat{r}_i) \|y_i^\times y_i^\times V_U\| \rightarrow 0. \tag{28}$$

Each l_i must converge to a positive constant $c_i^0 \leq l_i(0)$ as $\dot{l}_i \rightarrow 0$. Therefore, (25) ensures that $(\hat{r}_i - r_i) \rightarrow \pm \sqrt{2\alpha_i c_i^0}$. Integrating (28) over a period of time T , and using the fact that $(r_i - \hat{r}_i)$ is converging to a constant, it follows that

$$(r_i - \hat{r}_i) \int_t^{t+T} \|y_i^\times y_i^\times V_U\| d\tau \rightarrow 0.$$

Using the persistence of excitation assumption (20), it must be that $\hat{r}_i \rightarrow r_i$.

Recall the exception set χ (21). It is straightforward to verify that this set has measure zero. Since the initial choice of compact set K in the complement of χ was arbitrary, then the equilibrium $(y_i^\circ, 1)$ of (d_i, \tilde{r}_i) is almost semi-globally asymptotically stabilisable by the proposed correction terms (Def. A.1).

Observe that, at the equilibrium,

$$\begin{aligned}
\hat{y}_i &= \rho(\hat{X}, y_i^\circ) = \rho(\hat{X}, d_i) = y_i, \\
\hat{r}_i &= \tilde{r}_i r_i = r_i.
\end{aligned}$$

Using this, the ego-centric coordinates of the SLAM configuration and their estimates satisfy

$$R_P^\top (p_i - x_P) = r_i y_i = \hat{r}_i \hat{y}_i = R_{\hat{P}}^\top (\hat{p}_i - x_{\hat{P}}),$$

and thus

$$p_i = (R_{\hat{P}P^{-1}})^\top (\hat{p}_i - x_{\hat{P}P^{-1}}).$$

Therefore, at the equilibrium point, each p_i is related to each \hat{p}_i by the same rigid body transformation $S = \hat{P}P^{-1}$. Moreover, it is clear that $P = S^{-1}\hat{P}$. Hence, the two configurations on $\mathcal{T}_n^\circ(3)$ are equivalent on the SLAM manifold, $\xi \simeq \hat{\xi}$. This completes the proof.

5.3 Total Space Representative

In Theorem 1 the convergence result is independent of the choice of correction term Δ . This is due to the ego-centric nature of the group action considered and the invariance properties of the SLAM manifold, which cause the inertial frame of a SLAM system to be unobservable. In essence, choosing Δ will influence the element $S \in \mathbf{SE}(3)$ that relates the reference and estimated states, but does not influence the convergence of the SLAM error. Nevertheless, it is clear that as the error converges, it is desirable that the $S(t)$ that relates the reference and established states converges to a constant, essentially capturing the ‘‘inertial map’’ property that is desired in visual odometry. It is a key contribution of this paper to observe that imposing this constraint is a separate requirement from the underlying SLAM solution, that is, we must introduce an additional criterion that captures this property and then use this to design the correction term Δ .

The criterion that we propose to minimize is the weighted mean velocity of the landmark points

$$\sum_{i=1}^n \kappa_i \|\dot{\hat{p}}_i\|^2$$

For a static environment, the true landmark points are not moving. For the observer estimate these points may be moving, due to residue velocity associated with the landmark error correction, but also importantly, due to the correction term Δ that is moving the entire SLAM configuration. The motion due to Δ will be strongly correlated, while it is expected that the residue velocity due to the correction terms will be uncorrelated and for large constellations of points average to zero. Choosing Δ to minimize this additional criteria can be thought of as minimizing instantaneous map drift.

Proposition 5.1 *Let the origin configuration $\xi^\circ = (P^\circ, p^\circ) \in \mathcal{T}_n^\circ(3)$, the observer state $\hat{X} = (\hat{A}, (\hat{Q}, \hat{a})_i) \in \mathbf{VSLAM}_n(3)$, and the correction term $\Delta_{\hat{X}} = (\Delta, (\Gamma, \gamma)_i)$ be defined as in the statement of Theorem 1. Let $\hat{\xi} = (\hat{P}, \hat{p}_i) = \Upsilon(\hat{X}, \xi^\circ) \in \mathcal{T}_n^\circ(3)$ be the estimated state on the total space, and let $\hat{q}_i = R_{\hat{P}}^\top(\hat{p}_i - x_{\hat{P}})$ for each i .*

Then the solution to

$$\Delta = \underset{\Delta \in \mathbf{se}(3)}{\operatorname{argmin}} \left\{ \sum_{i=1}^n \kappa_i \|\dot{\hat{p}}_i\|^2 \right\}, \quad (29)$$

where κ_i are positive scalars, is given by

$$\Delta = \operatorname{Ad}_{\hat{A}}(\Omega_\Delta^\times, V_\Delta), \quad (30)$$

where Ω_Δ and V_Δ are determined by

$$\begin{pmatrix} \Omega_\Delta \\ V_\Delta \end{pmatrix} = - \left(\sum_{i=1}^n \kappa_i \begin{pmatrix} \hat{q}_i^\times \hat{q}_i^\times & \hat{q}_i^\times \\ -\hat{q}_i^\times & I_3 \end{pmatrix} \right)^{-1} \left(\sum_{i=1}^n \kappa_i \begin{pmatrix} \hat{q}_i^\times \operatorname{Ad}_{\hat{Q}_i^\top}(\Gamma_i) \hat{q}_i \\ \gamma_i \hat{q}_i + \operatorname{Ad}_{\hat{Q}_i^\top}(\Gamma_i) \hat{q}_i \end{pmatrix} \right), \quad (31)$$

so long as the inverse in (31) remains well-defined.

PROOF. First, observe that

$$\begin{aligned} \hat{q}_i &= R_{\hat{P}}^\top(\hat{p}_i - x_{\hat{P}}), \\ &= R_{\hat{P}}^\top(\hat{a}_i^{-1} R_{\hat{P}} \hat{Q}_i^\top R_{P^\circ}^\top(p_i^\circ - x_{P^\circ}) + x_{\hat{P}} - x_{P^\circ}), \\ &= \hat{a}_i^{-1} \hat{Q}_i^\top R_{P^\circ}^\top(p_i^\circ - x_{P^\circ}). \end{aligned}$$

Equation (29) presents a weighted least squares problem, and to solve it we analyse the component expressions $\kappa_i \dot{\hat{p}}_i$. The time derivative of each \hat{p}_i needs to be computed. Recall that the velocity lift Λ is defined precisely so that $\dot{\hat{p}}_i = 0$ when the correction terms are set to zero. Since differentiation is a linear operation, this means that

$$\begin{aligned} \dot{\hat{p}}_i &= \frac{d}{dt} \left(R_{P^\circ \hat{A}} \hat{a}_i^{-1} \hat{Q}_i^\top R_{P^\circ}^\top(p_i^\circ - x_{P^\circ}) + x_{P^\circ} - x_{\hat{A}} \right) \\ &= -R_{P^\circ \hat{A}} \Omega_\Delta^\times \hat{a}_i^{-1} \hat{Q}_i^\top R_{P^\circ}^\top(p_i^\circ - x_{P^\circ}) - R_{P^\circ \hat{A}} V_\Delta \\ &\quad + \gamma_i R_{P^\circ \hat{A}} \hat{a}_i^{-1} \hat{Q}_i^\top R_{P^\circ}^\top(p_i^\circ - x_{P^\circ}) \\ &\quad + R_{P^\circ \hat{A}} \hat{a}_i^{-1} \hat{Q}_i^\top \Gamma_i R_{P^\circ}^\top(p_i^\circ - x_{P^\circ}), \\ &= -R_{P^\circ \hat{A}} \Omega_\Delta^\times \hat{q}_i - R_{P^\circ \hat{A}} V_\Delta + \gamma_i R_{P^\circ \hat{A}} \hat{q}_i \\ &\quad + R_{P^\circ \hat{A}} \operatorname{Ad}_{\hat{Q}_i^\top}(\Gamma_i) \hat{q}_i, \\ \|\dot{\hat{p}}_i\| &= \left\| -\Omega_\Delta^\times \hat{q}_i - V_\Delta + \gamma_i \hat{q}_i + \operatorname{Ad}_{\hat{Q}_i^\top}(\Gamma_i) \hat{q}_i \right\|, \\ &= \left\| \begin{pmatrix} -\hat{q}_i^\times & I_3 \end{pmatrix} \begin{pmatrix} \Omega_\Delta \\ V_\Delta \end{pmatrix} - \left(\gamma_i \hat{q}_i + \operatorname{Ad}_{\hat{Q}_i^\top}(\Gamma_i) \hat{q}_i \right) \right\|. \end{aligned}$$

Therefore, by the theory of Weighted Least Squares, (31) is exactly the solution to (29), as required.

Proposition 5.1 provides a clear way to choose a correction term based on the static landmark assumption,

and allows for scalars κ_i to be chosen to weight the optimisation. The computational and memory costs of the correction terms scale linearly with the number of landmarks as opposed to alternative observer designs [10,21] which scale quadratically.

6 Simulation Results

To verify the landmark observer design in Theorem 1 and the robot correction term in Proposition 5.1, we conducted a simulation of a flying vehicle equipped with a monocular camera, observing 5 stationary landmarks as it moves in a circular trajectory with a constant body-fixed velocity $U = (\Omega_U^x, V_U)$, where $\Omega_U = (0, 0, 0.5)$ rad/s and $V_U = (1.5, 0, 0)$ m/s. For simplicity, it is assumed that the camera frame coincides with the body-fixed frame of the vehicle. The initial position of the vehicle was set to $(3, 3, 5)$ m with its rotational axes aligned with the inertial frame $\{0\}$. The positions of the landmarks were initialised to random positions $(p_i^1, p_i^2, 0)$ on the ground plane with $p_i^1, p_i^2 \sim N(0, 5^2)$ m.

The origin position P^o of the robot is set to the identity I_4 , and the origin landmark positions p_i^o are set to $10y_i(0)$, where $y_i(0)$ are the measured bearings to the true landmark positions at time 0. That is, the estimated points are initialised with correct bearings and an arbitrary depth of 10 m. The observer is defined on **VSLAM**₅(3) with kinematics given by (15), landmark correction terms (Γ_i, γ_i) given by (19), where $k_i = 5$ and $\alpha_i = 500$, and robot correction term Δ given as in Proposition 5.1 with $\kappa_i = 1$. At the end of the simulation, the estimated system state is aligned with the true system state by matching the true and estimated robot poses. Figure 4 shows the trajectories of estimated landmark positions and robot position over time, as well as the true landmark positions and the true robot trajectory, and figure 5 shows the evolution of each of the landmarks' associated storage functions, as defined in (25).

This simulation provides a simple demonstration of performance of the proposed observer and illustrates typical trajectories of the landmark estimates during a repeating motion such as the circle. The estimated landmark positions can be seen to converge to the true landmark positions in a natural manner. The choice to initialise landmarks as having a bearing matching the initial measurement is a natural one for practical implementation of the algorithm, although Theorem 1 provides that almost any initial conditions will converge. This almost semi-global convergence is a key property of the observer presented here that is not available in many of the state-of-the-art solutions.

7 Experimental Results

To demonstrate the observer described in Theorem 1 in a real-world scenario, we gathered video, GPS, and IMU

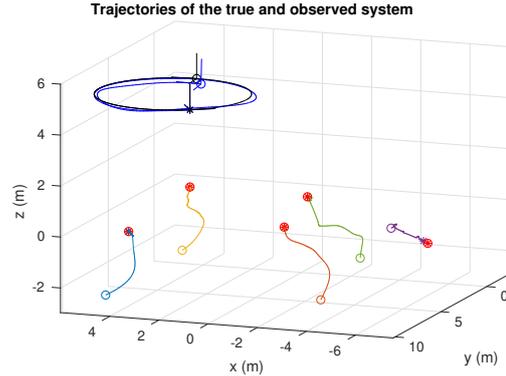


Fig. 4. The simulated trajectories of observer landmark estimates from initial positions with correct bearings but fixed depth of 10m. The observer landmark trajectories are shown in a range of colours matching those used in Figure 5, and the true landmark positions are shown in red. The observer robot trajectory is shown in blue, and the true robot trajectory is shown in black. The (\circ) and (\star) markers, respectively, denote the start and end of the trajectories of all the objects shown.

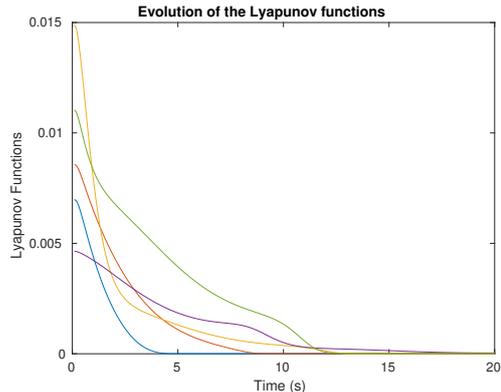


Fig. 5. The evolution of the storage functions of each of the landmarks in the system shown in Figure 4. The colours match the colours of the landmark trajectories in Figure 4. The initial convergence of the landmarks is quick as the bearings converge, and then slows as the depths gradually converge.

data from a Disco Parrot fixed-wing UAV flying outdoors. Image features were identified using OpenCV's goodFeaturesToTrack, and subsequently tracked using OpenCV's calcOpticalFlowPyrLK. These image features were then corrected for camera intrinsics and converted to spherical bearing coordinates before being used as landmark inputs to the observer. Landmarks are added to the system state after being observed for two frames so that their depths can be initialised from optical flow. When a landmark is no longer visible, it is removed from the observer state.

Initially, the input velocities $U = (\Omega_U, V_U)$ to the sys-

tem were estimated by combining the GPS signal (to obtain scale information) with egomotion estimated using the IMU and optical flow from the video stream, as outlined in [19]. Once sufficiently many landmarks are initialised, the optical flow vectors of each of the landmarks were combined with the existing landmark estimates to compute the input velocity $U = (\Omega_U, V_U)$. The observer was implemented using Euler integration with gain parameters set to $k_i = 5.0$ and $\alpha_i = 0.5$ for each i . The video recorded had a frame rate of 30 fps, leading to the Euler integration step being set to $dt = 0.033$ s. GPS data was recorded at 25 Hz in order to compare with the observer’s estimated trajectory. The observer trajectory was aligned to the GPS trajectory using the Umeyama method [23]. Figure 6 shows the aligned trajectories according to the observer and according to the GPS in the x and y directions, where the z direction refers to the plane’s altitude. Figure 7 shows the final positions of all landmark points in addition to the observer- and GPS-estimated trajectories. Figure 8 shows a frame taken from the video stream used in the experiment, with lines to represent the optical flow tracking overlaid. A video showcasing the feature tracking system is available online¹. The quality of the trajectory and map provided in Figure 7 show the robustness of the observer to noisy bearing measurements in practice.

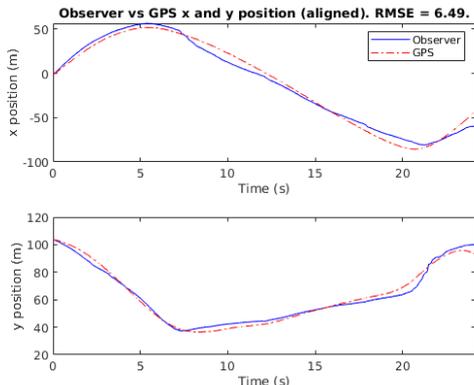


Fig. 6. The x and y positions of the UAV according to the aligned Observer (blue) and GPS (red).

8 Conclusion

This paper presents an observer design posed on a symmetry group for the VSLAM problem. The SLAM manifold introduced in [17] and the symmetry group discussed in [24] are reintroduced and exploited in the observer design. The observer is formulated on output errors, and provides a clear way to change the gains for bearing and depth of landmarks separately. The almost semi-global convergence of the proposed observer improves on the properties of state-of-the-art Extended Kalman Filter

¹ https://www.youtube.com/watch?v=QzIxx2eM1_s

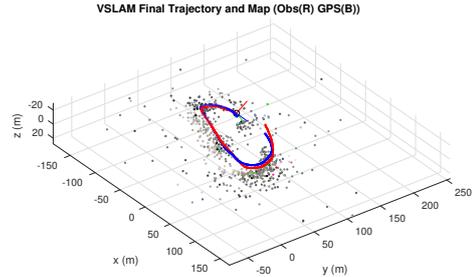


Fig. 7. The full trajectory of the UAV according to the aligned Observer (blue) and GPS (red), and the final positions of all of the landmarks, coloured with the colour of the pixel where they were first observed.

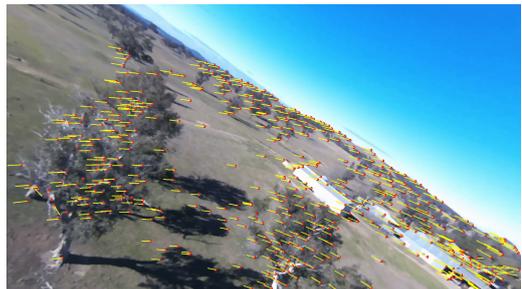


Fig. 8. A single frame of the video stream used in the experiment. Red circles represent the image features being tracked, and yellow lines represent the vector of motion of the image features between the current frame and the previous frame.

systems, which suffer from linearisation errors. While research into the development of non-linear observers for the SLAM problem is only recent, the observer for VSLAM presented in this paper demonstrates some of the key advantages the approach can offer.

A Almost Semi-Globally Asymptotically Stabilising Controls and Corrections

The concept of semi-global asymptotic stabilisability was introduced in [22] to model the dependence of gain on the basin of attraction in feedback stabilisation of a dynamical system. In the context of an observer analysis, this definition can be transferred to stabilisability of the error dynamics by correction. However, the classical concept introduced by Teel and Praly does not capture topological constraints associated with stability analysis on manifolds. For a large class of manifolds, including Lie-groups with $\mathbf{SO}(3)$ as a subgroup, these topological constraints prevent the existence of globally smooth asymptotically stable error dynamics [3]. On such spaces, smooth error dynamics will always admit an exception set χ of unstable or hyperbolic critical points that cannot be part of the basin of attraction of the desired equilibrium.

We consider a system-observer pair, where the observer

has the internal model principle, coupled through a correction function $\Delta(\hat{x}, y)$ depending on the observer state and system output. We assume that there is a well defined error function $e : \mathbf{G} \times \mathcal{M} \rightarrow \mathcal{M}$ where \mathbf{G} is the observer state space and \mathcal{M} is the system state space. The error dynamics evolve on the manifold \mathcal{M} depending on the system and observer state evolution as well as any exogenous inputs such as velocities.

Definition A.1 *An equilibrium e_* of the error dynamics of a system-observer pair, on a manifold \mathcal{M} , is almost globally stable if its basin of attraction is the complement of an exception set $\chi \subset \mathcal{M}$ of measure zero.*

An equilibrium e_ of the error dynamics of a system-observer pair, on a manifold \mathcal{M} , is said to be almost semi-globally stabilizable if, for each compact set $K \subset \mathcal{M}$ in the complement of an exception set $\chi \subset \mathcal{M}$ of measure zero, there exists a choice of correction $\Delta(\hat{x}, y)$ such that e_* is an asymptotically stable equilibrium of the error dynamics with basin of attraction containing K .*

References

- [1] G. Baldwin, R. Mahony, and J. Trumf. 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)*, pages 2237–2242, 2009.
- [2] Axel Barrau and Silvere Bonnabel. An EKF-SLAM algorithm with consistency properties, 2016. arXiv:1510.06263.
- [3] S.P. Bhata and D.S. Bernstein. Atopological obstruction to continuous global stabilization of rotational motion and the unwinding phenomenon. *Systems & Control Letters*, 39:63–70, 2000.
- [4] E. Bjerne, T. A. Johansen, and E. F. Brekke. Redesign and analysis of globally asymptotically stable bearing only SLAM. In *2017 20th International Conference on Information Fusion (Fusion)*, pages 1–8, July 2017.
- [5] S. Bonnabel, P. Martin, and P. Rouchon. Symmetry-preserving observers. *IEEE Transactions on Automatic Control*, 53(11):2514–2526, 2008.
- [6] F. Le Bras, T. Hamel, R. Mahony, and C. Samson. Observers for position and velocity bias estimation from single or multiple direction outputs. In T.I. Fossen, K.Y. Pettersen, and H. Nijmeijer, editors, *Sensing and Control for Autonomous Vehicles*, chapter 1. Lecture Notes in Control and Information Sciences 474, Springer, 2017.
- [7] Martin Brossard, Silvere Bonnabel, and Axel Barrau. Unscented kalman filter on lie groups for visual inertial odometry. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 649–655. IEEE, 2018.
- [8] Cesar Cadena, Luca Carlone, Henry Carrillo, Yasir Latif, Davide Scaramuzza, José Neira, Ian Reid, and John J Leonard. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on robotics*, 32(6):1309–1332, 2016.
- [9] J. Delmerico and D. Scaramuzza. A benchmark comparison of monocular visual-inertial odometry algorithms for flying robots. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2018.
- [10] Jorge Fuentes-Pacheco, José Ruiz-Ascencio, and Juan Manuel Rendón-Mancha. Visual simultaneous localization and mapping: a survey. *Artificial Intelligence Review*, 43(1):55–81, 2015.
- [11] T. Hamel and C. Samson. Riccati observers for the nonstationary pnp problem. *IEEE Transactions on Automatic Control*, 63(3):726–741, March 2018.
- [12] Minh-Duc Hua, Mohammad Zamani, Jochen Trumf, Robert Mahony, and Tarek Hamel. 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, December 2011.
- [13] John M Lee. Smooth manifolds. In *Introduction to Smooth Manifolds*. Springer, 2013.
- [14] Pedro Lourenço, Bruno Guerreiro, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. Simultaneous localization and mapping for aerial vehicles: a 3-d sensor-based gas filter. *Autonomous Robots*, 40(5):881–902, 2016.
- [15] Pedro Lourenço, Pedro Batista, Paulo Oliveira, and Carlos Silvestre. A globally exponentially stable filter for bearing-only simultaneous localization and mapping with monocular vision. *Robotics and Autonomous Systems*, 100:61 – 77, 2018.
- [16] R. Mahony, T. Hamel, and J. Pflimlin. Nonlinear complementary filters on the special orthogonal group. *IEEE Transactions on Automatic Control*, 53(5):1203–1218, June 2008.
- [17] Robert Mahony and Tarek Hamel. A geometric nonlinear observer for simultaneous localisation and mapping. In *Conference on Decision and Control*, page 6 pages, Melbourne, December 2017.
- [18] Robert Mahony, Jochen Trumf, and Tarek Hamel. Observers for kinematic systems with symmetry. In *Proceedings of 9th IFAC Symposium on Nonlinear Control Systems (NOLCOS)*, page 17 pages, 2013. Plenary paper.
- [19] Felix Schill, Robert Mahony, and Peter Corke. Estimating ego-motion in panoramic image sequences with inertial measurements. In *Robotics Research*, pages 87–101. Springer, 2011.
- [20] Jean-Jacques E Slotine, Weiping Li, et al. *Applied nonlinear control*, volume 199. Prentice hall Englewood Cliffs, NJ, 1991.
- [21] H. Strasdat, J.M.M. Montiel, and A.J. Davison. Visual SLAM: Why filter? *Computer Vision and Image Understanding (CVIU)*, 30(2):65–77, 2012.
- [22] A.R. Teel and L. Praly. Global stabilizability and observability imply semi-global stabilizability by output feedback. *Systems and Control Letters*, (22):313–325, 1994.
- [23] S Umeyama. Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(4):376–380, 1991.
- [24] Pieter van Goor, Robert Mahony, Tarek Hamel, and Jochen Trumf. A geometric observer design for visual localisation and mapping. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 2543–2549. IEEE, 2019.
- [25] J.F. Vasconcelos, R. Cunha, C. Silvestre, and P. Oliveira. A nonlinear position and attitude observer on SE(3) using landmark measurements. *Systems & Control Letters*, 59(3-4):155–166, 2010.
- [26] David Evan Zlotnik and James Richard Forbes. Gradient-based observer for simultaneous localization and mapping. *IEEE Transactions on Automatic Control*, 63(12):4338–4344, 2018.



Pieter van Goor is a PhD student at the Australian National University. He received his BEng(R&D) / BSc (Mechatronics and Advanced Mathematics) with first class Honours in 2018 from the Australian National University. His research interests are in applications of Lie group symmetries and geometric methods to problems in visual spatial aware-

ness. He is an IEEE student member.



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).