

A Geometric Observer Design for Visual Localisation and Mapping

Pieter van Goor, *Student Member, IEEE*, Robert Mahony, *Fellow, IEEE*, Tarek Hamel, *Senior Member, IEEE*,
and Jochen Trumpf *Senior Member, IEEE*

Abstract—This paper builds on recent work on Simultaneous Localisation and Mapping (SLAM) in the non-linear observer community, by framing the visual localisation and mapping problem as a continuous-time equivariant observer design problem on the symmetry group of a kinematic system. The state-space is a quotient of the robot pose expressed on $SE(3)$ and multiple copies of real projective space, used to represent both points in space and bearings in a single unified framework. An observer with decoupled Riccati-gains for each landmark is derived and we show that its error system is almost globally asymptotically stable and exponentially stable in-the-large.

I. INTRODUCTION

Simultaneous Localisation and Mapping (SLAM) is a well-known problem in mobile robotics and has been an active area of research for the last 30 years [1]. Visual localisation and mapping refers to the particular case of the SLAM problem where the only exteroceptive sensors available are cameras. The visual localisation and mapping problem, and particularly the case where only a single monocular camera is available, continues to be of substantial interest due to the low cost and low weight, as well as the ubiquity of single camera systems [1]. While visual localisation and mapping is an established research topic with a rich history [2], it remains an active research topic, especially in the area of low-cost light-weight embedded systems [3]. State-of-the-art filters and observers approach the SLAM problem through linearisation, and do not deal well with poor initial estimation or choice of linearisation point [2]. Additionally, these methods suffer from high computational complexity and poor scalability [1], [4].

Both the SLAM and visual localisation and mapping problems have attracted interest recently in the non-linear observer community. Approaches to these problems have emerged from earlier work on attitude estimation [5], [6] and pose estimation [7], [8], [9]. Bonnabel *et al.* [10] exploited a novel Lie group to design an invariant Kalman Filter for the SLAM problem. Parallel work by Mahony *et al.* [11] developed the same Lie group and proposed a quotient manifold structure for the state-space of the SLAM problem. Work by Zlotnik *et al.* [12] derives a geometrically motivated observer for the SLAM problem that includes estimation of bias in linear and angular velocity inputs. For the visual localisation and mapping problem, where only

bearing measurements are available, Lourenco *et al.* [13], [14] proposed an observer with a globally exponentially stable error system using depths of landmarks as separate components of the observer. Grabe *et al.* [15] derived a non-linear observer for the case where a significant number of the bearings measured are of coplanar landmarks by using the instantaneous homography constraint. Bjorne *et al.* [16] uses an attitude heading reference system (AHRS) to determine the orientation of the robot, and then solves the SLAM problem using a linear Kalman filter. A similar approach to the visual localisation and mapping case is undertaken in [17]. Hamel *et al.* have also introduced a Riccati observer [18] for the case where the orientation of the robot is known.

In this paper we present a novel non-linear geometric observer for the visual localisation and mapping problem. The approach extends the SLAM manifold presented in [11] to include bearings (such as magnetometer or gravity measurements) and landmark points in the same formulation by exploiting the structure of the real-projective space \mathbb{RP}^3 and homogeneous coordinates for bound and free vectors. The proposed \mathbb{RP}^3 state-space also allows modelling of visual features as a simple linear projection of \mathbb{RP}^3 onto \mathbb{RP}^2 . A novel Lie group termed the $VSLAM_n(3)$ group is introduced and shown to be a symmetry on the measurement function of the visual localisation and mapping problem. The proposed observer uses decoupled gain matrices for each landmark point that satisfy a simple Riccati equation. As a consequence of decoupling the Riccati observer for each landmark, the computational complexity of our approach is only $\mathcal{O}(n)$. Finally, the innovation on the pose of the robot is determined through finding the minimum of a novel cost function on the tangent space of \mathbb{RP}^3 , and is based on the static environment assumption common in SLAM algorithms. The resulting observer is shown to have an error system that is almost globally asymptotically stable (the basin of attraction excludes a set of measure zero) and exponentially stable in-the-large (exponentially stable on any compact set contained in the basin of attraction).

This paper consists of five sections alongside the introduction and conclusion. Section II introduces key notation and identities, and provides an in-depth explanation of the application of \mathbb{RP}^3 to representing points and bearings in 3D space. In Section III, we formulate the kinematics, state-space and output of the visual localisation and mapping system, and in Section IV we introduce the new Lie group $VSLAM_n(3)$ that acts on the state-space. In Section V we derive a non-linear observer on the Lie group, and in Section VI we provide the results of a simulation. The experimental

R. Mahony, J. Trumpf and P. van Goor are with the Australian Centre for Robotic Vision, Research School of Electrical, Energy and Materials Engineering, Australian National University, Australia {first name}. {last name}@anu.edu.au

T. Hamel is with I3S (University Côte d'Azur, CNRS, Sophia Antipolis) and Institut Universitaire de France, thamel@i3s.unice.fr

results are designed to verify the theory developed throughout the paper, not to provide a comprehensive evaluation of performance.

II. PRELIMINARIES

A. Notation

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

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

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 .

Consider a matrix $P \in \mathbf{SE}(3)$. The notations $R_P \in \mathbf{SO}(3)$ and $x_P \in \mathbb{R}^3$ are used to represent the rotation and translation components of P respectively, and P may be written as

$$P = \begin{pmatrix} R_P & x_P \\ 0 & 1 \end{pmatrix}.$$

Likewise, for a matrix $U \in \mathfrak{se}(3)$, the notations $\Omega_U \in \mathfrak{so}(3)$ and $V_U \in \mathbb{R}^3$ represent the rotational and translational velocity components of U respectively, and U may be written as

$$U = \begin{pmatrix} \Omega_U & V_U \\ 0 & 1 \end{pmatrix}.$$

For any $y \in \mathbb{R}^3 \setminus \{0\}$ the *projector* Π_y is given by

$$\Pi_y := I_3 - \frac{yy^\top}{|y|^2}.$$

The operator Π_y projects vectors onto the subspace of \mathbb{R}^3 orthogonal to y . The projector and the skew-symmetric matrix are related by

$$\Pi_y = -\frac{y^\times y^\times}{|y|^2}, \quad (1)$$

for any $y \in \mathbb{R}^3 \setminus \{0\}$. For any $\bar{y} \in \mathbb{R}^4 \setminus \{0\}$ the projector is similarly defined as

$$\bar{\Pi}_{\bar{y}} := I_4 - \frac{\bar{y}\bar{y}^\top}{|\bar{y}|^2}.$$

B. Real Projective Space

For $x \in \mathbb{R}^4 \setminus \{0\}$, define the set of equivalence classes

$$[x] := \{ax \mid a \in \mathbb{R} \setminus \{0\}\}.$$

Given two elements $x, y \in \mathbb{R}^4 \setminus \{0\}$, the notation $x \simeq y$ indicates $x = ay$ for some $a \in \mathbb{R} \setminus \{0\}$. The 3-dimensional real-projective space $\mathbb{RP}^3 = \{[x] \mid x \in \mathbb{R}^4 \setminus \{0\}\}$ is a smooth quotient manifold [19]. For any full rank matrix $A \in \mathbb{R}^{4 \times 4}$, the operation

$$A[x] := [Ax] \quad (2)$$

is well-defined.

Let $x \in \mathbb{R}^4 \setminus \{0\}$, and define an horizontal space $H_x = \{v \in \mathbb{R}^4 \mid v^\top x = 0\}$. Define an equivalence relationship $(x, v) \equiv (ax, av)$ for $a \in \mathbb{R} \setminus \{0\}$ between elements of H_x and H_{ax} . A tangent vector $v_{[x]} \in T_{[x]}\mathbb{RP}^3$ is the equivalence class $[x, v] = \{(ax, av) \mid v \in H_x\}$ and the tangent space $T_{[x]}\mathbb{RP}^3$ is the set of $av \in H_{ax}$.

For any $[x] \in \mathbb{RP}^3$, define the projector

$$\bar{\Pi}_{[x]} := \bar{\Pi}_x.$$

To see this is well-defined, let $a \in \mathbb{R}$ be a non-zero scalar, and check

$$\bar{\Pi}_{ax} = I_4 - \frac{(ax)(ax)^\top}{|(ax)|^2} = I_4 - \frac{a^2 xx^\top}{a^2 |x|^2} = \bar{\Pi}_x.$$

Analogously, the projector $\Pi_{[y]} := \Pi_y$ is well-defined for any $y \in \mathbb{RP}^2$.

Let $p \in \mathbb{R}^3$ be a vector representing the position of a point in space. Define the homogeneous coordinates

$$\bar{p} := \begin{pmatrix} p \\ 1 \end{pmatrix}$$

as an embedding $\mathbb{R}^3 \hookrightarrow \mathbb{R}^4$ and refer to such points \bar{p} as *bound vectors* with foot at the origin of the reference frame and tip at the \mathbb{R}^3 point it represents. Let $b \in S^2 = \{b \in \mathbb{R}^3 \mid |b| = 1\}$ be a vector representing a bearing or direction and define homogeneous coordinates

$$\bar{b} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

as an embedding $S^2 \hookrightarrow \mathbb{R}^4$. We term \bar{b} a *free vector*. Using these embeddings it is possible to define a map $\alpha : \mathbb{R}^3 \sqcup S^2 \rightarrow \mathbb{RP}^3$

$$\begin{aligned} \alpha(p) &:= [\bar{p}], & p \in \mathbb{R}^3, \\ \alpha(b) &:= [\bar{b}], & b \in S^2. \end{aligned}$$

A *point-type* element of \mathbb{RP}^3 is any element in the subset $\{[x] \mid x_4 \neq 0\}$. A *bearing-type* element of \mathbb{RP}^3 is any element in the subset $\{[x] \mid x_4 = 0\}$. A full inverse of α is not uniquely defined due to the sign ambiguity of elements of \mathbb{RP}^3 . However, it is possible to define a unique map $\gamma : \mathbb{RP}^3 \rightarrow \mathbb{R}^3 \sqcup \mathbb{RP}^2$ by

$$\gamma([x]) := \begin{cases} x_{1:3}/x_4 \in \mathbb{R}^3, & \text{if } x_4 \neq 0 \\ [x_{1:3}] \in \mathbb{RP}^2, & \text{if } x_4 = 0 \end{cases}, \quad (3)$$

where $x_{1:3} \in \mathbb{R}^3$ denotes the first three elements of x and $[x_{1:3}] = \{ax_{1:3} \mid a \in \mathbb{R} \setminus \{0\}\}$, analogous to the \mathbb{R}^4 definition. Define a projection $\beta : \mathbb{R}^3 \sqcup S^2 \rightarrow \mathbb{R}^3 \sqcup \mathbb{RP}^2$ by

$$\beta(x) := \begin{cases} x \in \mathbb{R}^3, & \text{if } x \in \mathbb{R}^3 \\ [x] \in \mathbb{RP}^2 & \text{if } x \in S^2 \end{cases}.$$

The following commutative diagram holds

$$\begin{array}{ccc} \mathbb{R}^3 \sqcup S^2 & & \\ \beta \downarrow & \searrow \alpha & \\ \mathbb{R}^3 \sqcup \mathbb{RP}^2 & \xleftarrow{\gamma} & \mathbb{RP}^3 \end{array}$$

The map γ is smooth under restriction to either point-type elements or bearing-type elements of \mathbb{RP}^3 . Although γ is unable to reconstruct the full direction vector b from a bearing-type \mathbb{RP}^3 element, the unsigned direction $[b]$ is sufficient for the observer construction that we undertake in the sequel.

III. PROBLEM FORMULATION

A. VSLAM Total Space

The formulation of the total space for the VSLAM problem is an extension of the formulation in [11] to include not only points in 3D space but also bearings through their \mathbb{RP}^3 representations.

Raw coordinates for the VSLAM problem can be defined by fixing an arbitrary reference frame $\{0\}$. Let $P \in \mathbf{SE}(3)$ and $\eta_i \in \mathbb{RP}^3$ represent the robot pose and landmark coordinates respectively, defined with respect to $\{0\}$. Note that each $\eta_i \in \mathbb{RP}^3$ is either point-type or bearing-type depending on whether its last entry is zero. The total space of the VSLAM problem is the product space

$$\mathcal{T}_n(3) = \mathbf{SE}(3) \times \mathbb{RP}^3 \times \dots \times \mathbb{RP}^3,$$

with elements

$$(P, \eta_1, \dots, \eta_n).$$

The notation $(P, \eta_i) \equiv (P, \eta_1, \dots, \eta_n)$ is used to simplify notation in the sequel.

Given $(P, \eta_i) \in \mathcal{T}_n(3)$, recalling (2) define

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

Given two elements $(P, \eta_i), (Q, \theta_i) \in \mathcal{T}_n(3)$, the notation $(P, \eta_i) \simeq (Q, \theta_i)$ means that $(P, \eta_i) = (S^{-1}Q, S^{-1}\theta_i)$ for some $S \in \mathbf{SE}(3)$. The SLAM manifold is the set

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

with quotient manifold structure [11].

An expression is well-defined on the SLAM manifold $\mathcal{M}_n(3)$ if it is invariant to the action of a rigid-body transformation of the reference frame. An important example is $(P, \eta_i) \mapsto P^{-1}\eta_i$. Given any $S \in \mathbf{SE}(3)$, one has

$$(S^{-1}P, S^{-1}\eta_i) \mapsto (S^{-1}P)^{-1}S^{-1}\eta_i = P^{-1}SS^{-1}\eta_i = P^{-1}\eta_i. \quad (4)$$

B. VSLAM Kinematics

The assumption will be made that the robot is moving through a static environment. Consider the velocity input space $\mathbb{V} = \mathfrak{se}(3)$. The kinematics of the VSLAM system are given by the function

$$f : \mathcal{T}_n(3) \times \mathbb{V} \rightarrow T\mathcal{T}_n(3), \quad ((P, \eta_i), U) \mapsto (PU, 0). \quad (5)$$

C. System Output

The physical measurements taken by our robot in the VSLAM system are the bearings of landmarks. Let $\eta'_i = P^{-1}\eta_i$ be the body-fixed frame coordinates of a landmark $\eta_i \neq \alpha(x_P)$. Using the basic pinhole camera model as described in [20] with invertible 3×3 camera matrix K , the measurement of η'_i taken by the camera is $(K \ \mathbf{0}_{3 \times 1})\eta'_i$. Assuming the camera is calibrated matrix K , it is easy to recover the element

$$K^{-1}(K \ \mathbf{0}_{3 \times 1})\eta'_i = (I_3 \ \mathbf{0}_{3 \times 1})\eta'_i$$

although the scale of this element is arbitrary and cannot be known. If η_i is a bearing-type element, then $\theta_{i,4} = 0$ and no information is lost through the camera projection. However, if η_i is a point-type element, then the scale of the vector is not recoverable. In this formulation the sign of the landmark measurement (representing whether the landmark is in front of or behind the camera) is ambiguous, but this is sufficient for the observer design undertaken in Section V. The choice of bearing-type or point-type for a particular landmark η_i is a modelling choice based on the requirements for the resulting map of the environment.

The output space of the VSLAM system is defined as

$$\mathcal{N}_n(3) := \mathbb{RP}^2 \times \dots \times \mathbb{RP}^2.$$

The output function of the VSLAM system is defined as

$$h : \mathcal{T}_n(3) \rightarrow \mathcal{N}_n(3), \quad (P, \eta_i) \mapsto (I_3 \ \mathbf{0})P^{-1}\eta_i. \quad (6)$$

The output function transforms each η_i into body-fixed frame coordinates, and projects the result into \mathbb{RP}^2 , representing bearing-type or point-type landmark measurements with a calibrated pinhole camera.

IV. SYMMETRY OF THE VSLAM PROBLEM

A. Symmetry of the Total Space

We introduce a group we term Scaled Orthogonal Transformations $\mathbf{SOT}(n)$, a subgroup of the group of similarity transforms on \mathbb{R}^n .

Lemma 4.1: For any $n \in \mathbb{N}$, the set

$$\mathbf{SOT}(n) = \left\{ \begin{pmatrix} R & 0 \\ 0 & a \end{pmatrix} \mid R \in \mathbf{SO}(n), a \in \mathbb{R} \setminus \{0\} \right\},$$

with matrix multiplication is a subgroup of $\mathbf{SIM}(n)$.

Proof: Assigning matrix multiplication as the group action it is clear that $\mathbf{SOT}(n)$ is the direct product of $\mathbf{SO}(3) \times \mathbb{R}_*$, where \mathbb{R}_* is the Lie group formed by assigning multiplication as the operation on $\mathbb{R} \setminus \{0\}$. It is straightforward to verify that $\mathbf{SOT}(n)$ is a subgroup of $\mathbf{SIM}(n)$ by considering the action $x \mapsto \frac{1}{a}Rx$ for $x \in \mathbb{R}^n$. ■

The action of $\mathbf{SOT}(3)$ on landmarks is a rotation combined with a scaling for point-type landmarks. Recalling (2) and taking advantage of the equivalence class structure of \mathbb{RP}^3 ,

$$\begin{pmatrix} R & 0 \\ 0 & a \end{pmatrix} \begin{bmatrix} p \\ 1 \end{bmatrix} = \begin{bmatrix} Rp \\ a \end{bmatrix} = \begin{bmatrix} \frac{1}{a}Rp \\ 1 \end{bmatrix}, \quad \begin{pmatrix} R & 0 \\ 0 & a \end{pmatrix} \begin{bmatrix} b \\ 0 \end{bmatrix} = \begin{bmatrix} Rb \\ 0 \end{bmatrix}.$$

There are exactly three orbits of **SOT**(3) acting on \mathbb{RP}^3 , defined by

$$\begin{aligned}\mathbb{RP}_p^3 &:= \{[x] \in \mathbb{RP}^3 \mid x_4 \neq 0, [x] \neq \alpha(\mathbf{0})\}, \\ \mathbb{RP}_b^3 &:= \{[x] \in \mathbb{RP}^3 \mid x_4 = 0\}, \\ \mathbb{RP}_0^3 &:= \{\alpha(\mathbf{0}) = \mathbf{e}_4\},\end{aligned}\quad (7)$$

where x_4 refers to the fourth coordinate of x .

The symmetry group $\mathbf{VSLAM}_n(3)$ for the VSLAM problem with n landmarks in 3 dimensions is defined as a Lie group

$$\mathbf{VSLAM}_n(3) = SE(3) \times \mathbf{SOT}(3) \times \dots \times \mathbf{SOT}(3),$$

with product Lie group structure. The associated Lie algebra is denoted $\mathfrak{vslam}_n(3)$.

Lemma 4.2: The mapping $\Upsilon : \mathbf{VSLAM}_n(3) \times \mathcal{T}_n(3) \rightarrow \mathcal{T}_n(3)$ defined by

$$\Upsilon((A, Q_i), (P, \eta_i)) = (PA, PAQ_i^{-1}P^{-1}\eta_i), \quad (8)$$

where the right-hand expression depends on definition (2), is a right group action of $\mathbf{VSLAM}_n(3)$ on $\mathcal{T}_n(3)$.

Proof: Trivially, $\Upsilon((I_4, I_4), (P, \eta_i)) = (P, \eta_i)$ for any $(P, \eta_i) \in \mathcal{T}_n(3)$. Let $(A_1, Q_{i,1}), (A_2, Q_{i,2}) \in \mathbf{VSLAM}_n(3)$ and (P, η_i) be arbitrary. Then

$$\begin{aligned}\Upsilon((A_1, Q_{i,1}), \Upsilon((A_2, Q_{i,2}), (P, \eta_i))) &= \Upsilon((A_1, Q_{i,1}), (PA_2, PA_2Q_{i,2}^{-1}P^{-1}\eta_i)), \\ &= (PA_2A_1, PA_2A_1Q_{i,1}^{-1}(PA_2)^{-1}PA_2Q_{i,2}^{-1}P^{-1}\eta_i), \\ &= (P(A_2A_1), P(A_2A_1)(Q_{i,2}Q_{i,1})^{-1}P^{-1}\eta_i), \\ &= \Upsilon((A_2, Q_{i,2}) \cdot (A_1, Q_{i,1}), (P, \eta_i)).\end{aligned}$$

This demonstrates that Υ is a right action as required. \blacksquare

Recall the orbits of **SOT**(3) described in (7). Given a configuration $(P^\circ, \eta_i^\circ) \in \mathcal{T}_n(3)$, let $(P, \eta_i) = \Upsilon((A, Q_i), (P^\circ, \eta_i^\circ))$ for some $(A, Q_i) \in \mathbf{VSLAM}_n(3)$. Observe that if $P^{\circ-1}\eta_j^\circ \in \mathbb{RP}_0^3$ for some j , then $P^{-1}\eta_j \in \mathbb{RP}_0^3$ also, independent of the particular element (A, Q_i) . To overcome this, in the remainder of the paper it is assumed that there is never a j such that $P^{\circ-1}\eta_j^\circ \in \mathbb{RP}_0^3$. This assumption is reasonable, in that it is equivalent to assuming there are no landmarks coinciding precisely with the origin of the robot. Additionally, it is assumed that the type of each landmark (point or bearing) is known, and the landmarks are enumerated such that $i = 1, \dots, n_p$ and $i = n_p+1, \dots, n_p+n_b = n$ represent of point- and bearing-type landmarks respectively. The reduced total space is defined as

$$\begin{aligned}\mathcal{T}_{n_p, n_j}^\circ(3) &:= \{(P, \eta_i) \in \mathcal{T}_{n_p+n_j}(3) \mid 1 \leq i \leq n_p \Leftrightarrow \eta_i \in \mathbb{RP}_p^3, \\ &\quad 1 \leq i - n_p \leq n_b \Leftrightarrow \eta_i \in \mathbb{RP}_p^3\},\end{aligned}$$

and only elements $(P, \eta_i) \in \mathcal{T}_{n_p, n_j}^\circ(3)$ are considered from here going forward.

B. Lift of the VSLAM Kinematics

In order to consider the system on the $\mathbf{VSLAM}_n(3)$ group, the kinematics from the state space must be lifted onto the group. The following lemma provides the lift function.

Lemma 4.3: The function $\lambda : \mathcal{T}_{n_p, n_j}^\circ(3) \times \mathbb{V} \rightarrow \mathfrak{vslam}_n(3)$, defined by

$$\lambda((P, \eta_i), U) = (U, W(U, P^{-1}\eta_i)),$$

where $W : \mathfrak{se}(3) \times (\mathbb{RP}^3 \cup \mathbb{RP}_p^3) \rightarrow \mathfrak{so}(3)$ is given by

$$W\left((\Omega_U, V_U), \begin{bmatrix} q \\ r \end{bmatrix}\right) = \begin{pmatrix} \left(\Omega_U - r \frac{V_U \times q}{|q|^2}\right)^\times & 0 \\ 0 & -r \frac{V_U^\top q}{|q|^2} \end{pmatrix},$$

is a velocity lift of the kinematics (5) onto $\mathbf{VSLAM}_n(3)$ with respect to the group action (8).

Proof: To show that λ is a velocity lift, it is required that

$$D\Upsilon_{(P, \eta_i)}(\text{id})[\lambda((P, \eta_i), U)] = f((P, \eta_i), U).$$

Equivalently, it is required to show that

$$(PU, \bar{\Pi}_{\eta_i}(PUP^{-1} - PW_iP^{-1})\eta_i) = (PU, 0), \quad (9)$$

where $W_i := W(U, P^{-1}\eta_i)$.

First, it is necessary to show that W is well-defined whenever $q \neq \mathbf{0}$. To see this, let $a \in \mathbb{R}$ be any non-zero scalar, and observe that

$$\begin{aligned}W\left(U, \begin{bmatrix} aq \\ ar \end{bmatrix}\right) &= \begin{pmatrix} \left(\Omega_U - ar \frac{V_U \times (aq)}{|aq|^2}\right)^\times & 0 \\ 0 & -ar \frac{V_U^\top (aq)}{|aq|^2} \end{pmatrix}, \\ &= \begin{pmatrix} \left(\Omega_U - r \frac{V_U \times q}{|q|^2}\right)^\times & 0 \\ 0 & -r \frac{V_U^\top q}{|q|^2} \end{pmatrix} = W\left(U, \begin{bmatrix} q \\ r \end{bmatrix}\right).\end{aligned}\quad (10)$$

Recalling the expression for f provided in (5), it is clear that the first terms on both sides of (9) are equal. Let

$$\begin{bmatrix} q_i \\ r_i \end{bmatrix} := P^{-1}\eta_i. \quad (11)$$

In order to aid in the readability of the following equations, q_i and r_i in (11) are chosen such that $|q_i| = 1$. However, it is important to note this choice is arbitrary as shown in (10). To show (9), consider that

$$\begin{aligned}\bar{\Pi}_{\eta_i}PUP^{-1}\eta_i &= \bar{\Pi}_{\eta_i}PU \begin{bmatrix} q_i \\ r_i \end{bmatrix}, \\ &= \bar{\Pi}_{\eta_i}P \begin{bmatrix} \Omega_U^\times q_i + r_i V_U \\ 0 \end{bmatrix}, \\ &= \bar{\Pi}_{\eta_i} \left[P \begin{pmatrix} \Omega_U^\times q_i + r_i V_U \\ 0 \end{pmatrix} - r_i V_U^\top q_i P \begin{pmatrix} q_i \\ r_i \end{pmatrix} \right], \\ &= \bar{\Pi}_{\eta_i}P \begin{bmatrix} \Omega_U^\times q_i + r_i (I_3 - q_i q_i^\top) V_U \\ -r_i V_U^\top q_i r_i \end{bmatrix}, \\ &= \bar{\Pi}_{\eta_i}P \begin{bmatrix} \Omega_U^\times q_i - r_i q_i^\times q_i^\times V_U \\ -r_i V_U^\top q_i r_i \end{bmatrix},\end{aligned}$$

using the identity (1). This further reduces to

$$\begin{aligned}\bar{\Pi}_{\eta_i} P U P^{-1} \eta_i &= \bar{\Pi}_{\eta_i} P \begin{bmatrix} \Omega_U^\times q_i - r_i (V_U^\times q_i)^\times q_i \\ -r_i V_U^\top q_i r_i \end{bmatrix}, \\ &= \bar{\Pi}_{\eta_i} P \begin{pmatrix} (\Omega_U - r_i V_U \times q_i)^\times & 0 \\ 0 & -r_i V_U^\top q_i \end{pmatrix} \begin{bmatrix} q_i \\ r_i \end{bmatrix}, \\ &= \bar{\Pi}_{\eta_i} P W_i P^{-1} \eta_i,\end{aligned}$$

where the last step follows from (11) and the choice of $|q_i| = 1$. From here, (9) clearly resolves to

$$\begin{aligned}D\Upsilon_{(P, \eta_i)}(\text{id})[(U, W_i)] &= (PU, 0) \\ &= f((P, \eta_i), U),\end{aligned}$$

as required. This completes that proof that λ is a velocity lift. \blacksquare

The kinematics of the true state $\xi = (P, \eta_i) \in \mathcal{T}_{n_p, n_j}^\circ(3)$ of the VSLAM system are given by

$$\dot{\xi} = f(\xi, U). \quad (12)$$

Choose a reference configuration $\xi^\circ = (P^\circ, \eta_i^\circ) \in \mathcal{T}_{n_p, n_j}^\circ(3)$. By construction, the trajectories of the lifted system kinematics

$$\dot{X} = X\lambda(\Upsilon(X, \xi^\circ), U)$$

project to trajectories of the VSLAM kinematics (12) via $\xi(t) = \Upsilon(X(t), \xi^\circ)$.

V. OBSERVER DESIGN

A. Observer Kinematics

Define the observer state to lie on the VSLAM group, $\hat{X} = (\hat{A}, \hat{Q}_i) \in \mathbf{VSLAM}_n(3)$, with kinematics given by

$$\begin{aligned}\frac{d}{dt} \hat{X} &= \hat{X} \lambda(\Upsilon(\hat{X}, \xi^\circ), U) + \hat{X} \Delta_{\hat{X}}, \\ \hat{X}(0) &= \text{id},\end{aligned} \quad (13)$$

where $\Delta_{\hat{X}} = (\Delta_{\hat{A}}, \Delta_{\hat{Q}_i}) \in \mathbf{vslam}_n(3)$ is an innovation term. The estimated state $\hat{\xi} = (\hat{P}, \hat{\eta}_i) \in \mathcal{T}_{n_p, n_j}^\circ(3)$ is given by

$$\hat{\xi} = \Upsilon(\hat{X}, \xi^\circ). \quad (14)$$

Additional notation is helpful in simplifying the expressions that follow in the observer design. Define

$$\hat{y}_i := h((\hat{P}, \hat{\eta}_i)), \quad y_i := h((P, \eta_i)). \quad (15)$$

All expressions above are well-defined for equivalence classes in the SLAM manifold.

B. Landmark Observer

Theorem 5.1: Let $\xi = (P, \eta_i) \in \mathcal{T}_{n_p, n_j}^\circ(3)$ be the true state of the system, evolving with the kinematics (12). Let $\xi^\circ \in \mathcal{T}_{n_p, n_j}^\circ(3)$ be arbitrary up to the requirement that, for all i , η_i° and η_i are members of the same orbit of \mathbb{RP}^3 under the action of $\mathbf{SOT}(3)$. Define $\hat{X} = (\hat{A}, \hat{Q}_i) \in \mathbf{VSLAM}_n(3)$ to be the observer state with kinematics defined by (13), and define $\hat{\xi} = (\hat{P}, \hat{\eta}_i)$ as in (14).

Now, for $i = 1, \dots, n_p$, define $\Sigma_i \in \mathbb{R}^{3 \times 3}$ by

$$\begin{aligned}\dot{\Sigma}_i &= \Sigma_i \Omega_U^\times - \Omega_U^\times \Sigma_i + H_i - \Sigma_i \Pi_{y_i} G_i \Pi_{y_i} \Sigma_i, \\ \Sigma(0)_i &= \Sigma_{i,0} > 0, \quad G_i = k_G I_3, \quad H_i = k_H I_3,\end{aligned} \quad (16)$$

where $k_G, k_H > 0$ are constants, and assume that there exist $\delta > 0$ and $\mu > 0$ such that

$$\frac{1}{\delta} \int_t^{t+\delta} \Pi_{R_P(s) y_i(s)} ds \geq \mu I_3. \quad (17)$$

for any time $t > 0$ and for any $i = 1, \dots, n_p$. For $i = n_p + 1, \dots, n_p + n_b$, define

$$\Sigma_i \equiv I_3, \quad G_i = I_3, \quad H_i = I_3. \quad (18)$$

Then, for every landmark $i = 1, \dots, n_p + n_b$, define $\Delta_{\hat{Q}_i}$ as

$$\begin{aligned}\Delta_{\hat{Q}_i} &= \begin{pmatrix} (\hat{y}_i^\times K_i \Pi_{y_i} \hat{y}_i)^\times & 0 \\ 0 & -\hat{y}_i^\top K_i \Pi_{y_i} \hat{y}_i \end{pmatrix}, \\ K_i &= k \Sigma_i \Pi_{y_i} G_i, \quad k > 0.5,\end{aligned} \quad (19)$$

where y_i and \hat{y}_i are given by (15). Let the innovation term $\Delta_{\hat{A}}$ be given by the least-squares solution to

$$\begin{aligned}\min_{(\Delta_{\hat{R}}, \Delta_{\hat{x}})} \sum_{i=1}^n \left| \frac{1}{|\hat{\theta}_i|} \bar{\Pi}_{\hat{\theta}_i} \left(\begin{pmatrix} -(\hat{\theta}_i^{1:3})^\times & \hat{\theta}_i^4 I_3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta_{\hat{R}} \\ \Delta_{\hat{x}} \end{pmatrix} + \Delta_{\hat{Q}_i} \hat{\theta}_i \right) \right|, \\ \Delta_{\hat{A}} &= \begin{pmatrix} \Delta_{\hat{R}}^\times & \Delta_{\hat{x}} \\ 0 & 0 \end{pmatrix}, \quad \hat{\theta}_i := \hat{P}^{-1} \hat{\eta}_i.\end{aligned} \quad (20)$$

Then the estimated state coordinates $\hat{\xi}$ converge to the true coordinates ξ almost-globally asymptotically and exponentially in the large¹ up to equivalence on the SLAM manifold $\mathcal{M}_n(3)$.

Proof: To verify that $\Delta_{\hat{A}}$ is well-defined note that the cost in (20) is invariant to scale in the data $\hat{\theta}_i \mapsto a_i \hat{\theta}_i$ for $a_i \in \mathbb{R} \setminus \{0\}$. A Lyapunov analysis proves the desired result.

For $i = 1, \dots, n_p$, recalling (3), define the error coordinates and candidate storage function as

$$\begin{aligned}e_i &:= \gamma(\hat{P}^{-1} \hat{\eta}_i) - \gamma(P^{-1} \eta_i), \\ l_i &:= \frac{1}{2} e_i^\top \Sigma_i^{-1} e_i,\end{aligned}$$

respectively. The condition (17) ensures that Σ_i is well-conditioned, and remains bounded and positive-definite for all time $t \geq 0$ [21]. Therefore the candidate storage function l_i is positive definite. It remains to show that l_i is monotonically decreasing. The kinematics of e_i are

$$\dot{e}_i = -\Omega_U^\times e_i - K_i \Pi_{y_i} e_i.$$

¹For any compact set in the basin of attraction of the equilibrium, the value of the Lyapunov function converges exponentially to zero.

Differentiating the candidate storage function, one has

$$\begin{aligned}
\dot{l}_i &= e_i^\top \Sigma_i^{-1} \dot{e}_i - \frac{1}{2} e_i^\top \Sigma_i^{-1} \dot{\Sigma}_i \Sigma_i^{-1} e_i, \\
&= e_i^\top \Sigma_i^{-1} (-\Omega_U^\times e_i - k \Sigma_i \Pi_{y_i} G_i \Pi_{y_i} e_i) - \frac{1}{2} e_i^\top \Sigma_i^{-1} (\Sigma_i \Omega_U^\times \\
&\quad - \Omega_U^\times \Sigma_i + H_i - \Sigma_i \Pi_{y_i} G_i \Pi_{y_i} \Sigma_i) \Sigma_i^{-1} e_i, \\
&= -\frac{1}{2} e_i^\top \Sigma_i^{-1} \Omega_U^\times e_i - \frac{1}{2} e_i^\top \Omega_U^\times \Sigma_i^{-1} e_i \\
&\quad + \left(\frac{1}{2} - k\right) e_i^\top \Pi_{y_i} G_i \Pi_{y_i} e_i - \frac{1}{2} e_i^\top \Sigma_i^{-1} H_i \Sigma_i^{-1} e_i, \\
&\leq -\frac{1}{2} e_i^\top \Sigma_i^{-1} H_i \Sigma_i^{-1} e_i, \\
&\leq -\frac{1}{2} \frac{\sigma_{i,m}^2}{\sigma_{i,M}} k_H l_i,
\end{aligned}$$

where $\sigma_{m,i}$ and $\sigma_{M,i}$ denote the infimum of the smallest and the supremum of the largest eigenvalues of Σ_i over time, respectively. Since $k_H > 0$ is chosen as a constant, and Σ_i remains well-conditioned and bounded, the equilibrium $e_i = 0$ is exponentially stable. Equivalently, this provides that $\hat{P}^{-1} \hat{\eta}_i \rightarrow P^{-1} \eta_i$ globally exponentially.

For $i = n_p + 1, \dots, n_p + n_b$, define the candidate storage function

$$l_i := \frac{1}{2} \left(1 - \left(\frac{y_i^\top \hat{y}_i}{|y_i| |\hat{y}_i|} \right)^2 \right). \quad (21)$$

Observe that l_i is well-defined as a function of \mathbb{RP}^2 elements, since the expression is invariant to multiplication of y_i or \hat{y}_i by any non-zero scalar. Clearly l_i is positive definite. The kinematics of the bearing $y_i \in \mathbb{RP}^2$ are given by

$$\begin{aligned}
\dot{y}_i &= \frac{d}{dt} (I_3 \quad \mathbf{0}) P^{-1} \eta_i, \\
&= - (I_3 \quad \mathbf{0}) \bar{\Pi}_{P^{-1} \eta_i} W_i P^{-1} \eta_i, \\
&= -\Pi_{y_i} \Omega_U^\times (I_3 \quad \mathbf{0}) P^{-1} \eta_i, \\
&= -\Omega_U^\times y_i.
\end{aligned}$$

This is well-defined as an element of the tangent space $T_{y_i} \mathbb{RP}^2$ since any scaling of y_i results in the same scaling of the expression for \dot{y}_i . Since $\hat{y}_i^\top y_i = 0$, the dynamics of the norm of any chosen representative of y_i are given by $\frac{d}{dt} |y_i| = 0$. Analogously, recalling (18) and (19), the kinematics of $\hat{y}_i \in \mathbb{RP}^2$ are given by

$$\begin{aligned}
\dot{\hat{y}}_i &= (-\Omega_U^\times - (\hat{y}_i^\times K_i \Pi_{y_i} \hat{y}_i)^\times) \hat{y}_i, \\
&= -\Omega_U^\times \hat{y}_i + \hat{y}_i^\times \hat{y}_i^\times (k \Sigma_i \Pi_{y_i} G_i) \Pi_{y_i} \hat{y}_i, \\
&= -\Omega_U^\times \hat{y}_i - k \Pi_{\hat{y}_i} \Pi_{y_i} \hat{y}_i,
\end{aligned}$$

and hence the dynamics of the norm of any representative of \hat{y}_i are given by $\frac{d}{dt} |\hat{y}_i| = 0$. As a consequence of this and the scale invariance of (21), we may choose $|y_i| = |\hat{y}_i| = 1$ for readability without loss of generality. Differentiating the candidate storage function leads to

$$\begin{aligned}
\dot{l}_i &= -(y_i^\top \hat{y}_i) (\dot{y}_i^\top \hat{y}_i + y_i^\top \dot{\hat{y}}_i), \\
&= k (y_i^\top \hat{y}_i) y_i^\top \Pi_{\hat{y}_i} \Pi_{y_i} \hat{y}_i, \\
&= k (y_i^\top \hat{y}_i)^2 ((y_i^\top \hat{y}_i)^2 - 1) \\
&= -k (y_i^\top \hat{y}_i)^2 l_i
\end{aligned}$$

which is negative definite as long as the initial directions $y_i(0)$ and $\hat{y}_i(0)$ are not orthogonal. There are two situations in which $l_i = 0$. The first one corresponds to the stable case where $l_i = 0$ (\hat{y}_i and y_i are parallel) while the second one corresponds to the unstable case for which $l_i = 1$ (\hat{y}_i and y_i are orthogonal). To prove the exponential stability in the large, suppose that $0 < l_i \leq \epsilon < 1$ for some fixed ϵ . Then,

$$\begin{aligned}
\dot{l}_i &= -k (y_i^\top \hat{y}_i)^2 l_i, \\
&= -k (1 - l_i) l_i, \\
&\leq -k (1 - \epsilon) l_i.
\end{aligned}$$

Observe that, unless $l_i = 1$, such an ϵ can always be found. Therefore, $l_i \rightarrow 0$ almost-globally asymptotically, and exponentially in the large. Since the measurement function h is invertible on bearing-type elements, this provides the desired result that $\hat{P}^{-1} \hat{\eta}_i \rightarrow P^{-1} \eta_i$ almost-globally asymptotically and exponentially in the large.

Define the whole-of-system Lyapunov function

$$\mathcal{L} := \sum_{i=1}^n l_i. \quad (22)$$

From the analysis of each individual l_i , it is clear that $\mathcal{L} \rightarrow 0$ almost-globally asymptotically and exponentially in the large. The convergence of each \mathcal{L} provides that

$$\begin{aligned}
(\hat{P}, \hat{\eta}_i) &\simeq ((P \hat{P}^{-1}) \hat{P}, (P \hat{P}^{-1}) \hat{\eta}_i), \\
&= (P, P(\hat{P}^{-1} \hat{\eta}_i)), \\
&\rightarrow (P, P(P^{-1} \eta_i)), \\
&= (P, \eta_i),
\end{aligned}$$

almost-globally asymptotically and exponentially in the large as well. This completes the proof. \blacksquare

VI. SIMULATION RESULTS

To verify the observer derived in Theorem 5.1, we conducted a simulation of a vehicle equipped with a single monocular camera, observing 4 point-type landmarks and 2 bearing-type landmarks as it moves through space. The vehicle moves in a circular trajectory at a fixed height of 3 m. The body-fixed velocity U is fixed to be constant, with $\Omega_U = (0, 0, -0.5)^\top$ rad/s and $V_U = (1.5, 0, 0)$ m/s. For simplicity, the camera frame is assumed to coincide with the body-fixed frame of the vehicle, which avoids the need for a separate computation to transform the body-fixed velocity into the camera frame. Let the true state be $(P, \eta_i) \in \mathcal{T}_{n_p, n_b}^\circ(3)$. The reference configuration is chosen as $\xi^\circ = (I_4, \eta_i^\circ)$, where

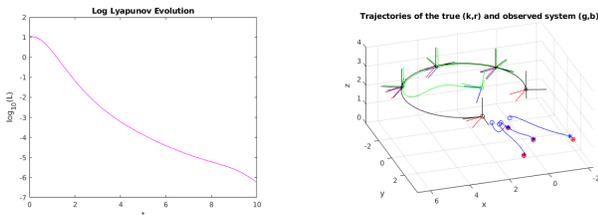
$$\eta_i^\circ = \alpha \left(2 \left(\frac{h(\eta_i)}{|h(\eta_i)|} + \epsilon_i \right) \right)$$

where the ϵ_i terms represent errors in the initial measurements. The observer is defined on $\mathbf{VSLAM}_n(3)$, with kinematics given by (13) and innovation terms given by Theorem 5.1. The initial conditions and gains for the observer are chosen as

$$\Sigma_i(0) = 25I_3, \quad k_H = 0.5, \quad k_G = 2.0, \quad k = 1.0.$$

The simulation was carried out by implementing the continuous time system with Euler integration using a time step of $dt = 0.02$ s.

Figure 1a shows the evolution of $\log_{10}(\mathcal{L})$, where \mathcal{L} is the Lyapunov function of the simulated system as defined in (22). This clearly shows exponential convergence of the observer error dynamics. Figure 1b shows the evolution of the trajectory of the simulated system. Since the estimated state only converges to the true state up to equivalence on the SLAM manifold $\mathcal{M}_n(3)$, it is necessary to assign total space coordinates to the estimate to aid the comparison. In Figure 1b the choice of total space coordinates for the estimated state is made so that the final robot pose is aligned with that of the true state. This shows that the landmarks have correctly converged to the true landmarks up to the SLAM manifold equivalence.



(a) The evolution of \log_{10} of the Lyapunov function \mathcal{L} (22) with respect to time.

(b) The trajectory of the simulated system (green, blue) compared with the true system evolution (black, red).

VII. CONCLUSION

This paper presents an observer design posed on a novel symmetry group for the visual SLAM problem. The total space and SLAM manifold conceptualised in [11] have been extended to include free vectors. The development of the symmetry group $\mathbf{VSLAM}_n(3)$ has allowed both point-type and bearing-type landmarks to be treated in a unified framework. Riccati observers were incorporated for each of the point-type landmarks, and grant the user refined control over their convergence. The almost-global convergence of the proposed observer on both point-type and bearing-type landmarks is a contrast to many state-of-the-art Extended Kalman Filter 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 visual SLAM presented in this paper demonstrates some of the key advantages the approach can offer.

REFERENCES

- [1] J. Fuentes-Pacheco, J. Ruiz-Ascencio, and J. M. Rendón-Mancha, "Visual simultaneous localization and mapping: a survey," *Artificial Intelligence Review*, vol. 43, no. 1, pp. 55–81, 2015.
- [2] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. Leonard, "Past, present, and future of simultaneous localization and mapping: Towards the robust-perception age," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1309–1332, December 2016.

- [3] 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.
- [4] H. Strasdat, J. Montiel, and A. Davison, "Visual SLAM: Why filter?," *Computer Vision and Image Understanding (CVIU)*, vol. 30, no. 2, pp. 65–77, 2012.
- [5] R. Mahony, T. Hamel, and J. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1203–1218, June 2008.
- [6] S. Bonnabel, P. Martin, and P. Rouchon, "Symmetry-preserving observers," *IEEE Transactions on Automatic Control*, vol. 53, no. 11, pp. 2514–2526, 2008.
- [7] G. Baldwin, R. Mahony, and J. Trunpf, "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)*, 2009, pp. 2237–2242.
- [8] J. Vasconcelos, R. Cunha, C. Silvestre, and P. Oliveira, "A nonlinear position and attitude observer on SE(3) using landmark measurements," *Systems & Control Letters*, vol. 59, no. 3–4, pp. 155–166, 2010.
- [9] M.-D. Hua, M. Zamani, J. Trunpf, R. Mahony, and T. 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.
- [10] A. Barrau and S. Bonnabel. (2016) An EKF-SLAM algorithm with consistency properties. ArXiv:1510.06263. [Online]. Available: <https://arxiv.org/abs/1510.06263v3>
- [11] R. Mahony and T. Hamel, "A geometric nonlinear observer for simultaneous localisation and mapping," in *Conference on Decision and Control*, Melbourne, December 2017, p. 6 pages.
- [12] D. E. Zlotnik and J. R. Forbes, "Gradient-based observer for simultaneous localization and mapping," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4338–4344, 2018.
- [13] P. Lourenço, B. Guerreiro, P. Batista, P. Oliveira, and C. Silvestre, "Simultaneous localization and mapping for aerial vehicles: a 3-d sensor-based gas filter," *Autonomous Robots*, vol. 40, no. 5, pp. 881–902, 2016.
- [14] P. Lourenço, P. Batista, P. Oliveira, and C. Silvestre, "A globally exponentially stable filter for bearing-only simultaneous localization and mapping with monocular vision," *Robotics and Autonomous Systems*, vol. 100, pp. 61 – 77, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0921889017300234>
- [15] V. Grabe, H. H. Bühlhoff, D. Scaramuzza, and P. R. Giordano, "Nonlinear ego-motion estimation from optical flow for online control of a quadrotor uav," *The International Journal of Robotics Research*, vol. 34, no. 8, pp. 1114–1135, 2015.
- [16] E. Bjorne, 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)*, July 2017, pp. 1–8.
- [17] F. L. Bras, T. Hamel, R. Mahony, and C. Samson, "Observers for position and velocity bias estimation from single or multiple direction outputs," in *Sensing and Control for Autonomous Vehicles*, T. Fossen, K. Pettersen, and H. Nijmeijer, Eds. Lecture Notes in Control and Information Sciences 474, Springer, 2017, ch. 1.
- [18] T. Hamel and C. Samson, "Riccati observers for the nonstationary pnp problem," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 726–741, March 2018.
- [19] P. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ, USA: Princeton University Press, January 2008.
- [20] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [21] T. Hamel and C. Samson, "Riccati observers for position and velocity bias estimation from direction measurements," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, Dec 2016, pp. 2047–2053.