Bias Estimation for Invariant Systems on Lie Groups with Homogeneous Outputs

A. Khosravian, J. Trumpf, R. Mahony, C. Lageman

Abstract— **In this paper, we provide a general method of state estimation for a class of invariant systems on connected matrix Lie groups where the group velocity measurement is corrupted by an unknown constant bias. The output measurements are given by a collection of actions of a single Lie group on several homogeneous output spaces, a model that applies to a wide range of practical scenarios. The proposed observer consists of a group estimator part, providing an estimate of a bounded state evolving on the Lie group, and a bias estimator part, providing an estimate of the bias in the associated Lie algebra. We employ the gradient of a suitable invariant cost function on the Lie group as an innovation term in the group estimator. We design the bias estimator such that it guarantees uniform local exponential stability of the estimation error dynamics around the zero error state. We propose a systematic methodology for the design of suitable cost functions on Lie groups by lifting invariant cost functions from the homogeneous output spaces. We show that the resulting observer is implementable based on available sensor measurements if the homogeneous output spaces are reductive. As an example, we derive an observer for rigid body attitude using vector and gyro measurements with unknown constant gyro bias.**

I. Introduction

Observer design for invariant systems whose state evolves on a Lie group has recently been subject to intensive research. The interest in this subject is motivated by several applications, in particular attitude and pose estimation for inertial navigation systems. Systematic observer design methodologies have been proposed that lead to strong stability and robustness properties. Specifically, Bonnabel *et al.* [4]–[6] consider observers which consist of a copy of the system and a correction term, along with a constructive method to find suitable symmetry-preserving correction terms. The construction utilizes the invariance of the system and the moving frame method, leading to local convergence properties of the observers. The authors in [14]–[16] propose methods to achieve almost globally convergent observers. A key aspect of the design approach proposed in [14]–[16] is the use of the invariance properties of the system to ensure that the error dynamics are globally defined and are autonomous. This leads to a straight forward stability analysis and excellent performance in practice. A key extension to early works in this area is the consideration of output measurements where a partial state measurement is generated by an action of the Lie group on a homogeneous output space [4]–[6], [14], [15]. Specifically, when the output manifold is a reductive homogeneous space of the Lie group, the authors in [14], [15] propose an observer design on the output manifold and then lift the designed observer to obtain a corresponding observer on the Lie group. The approach can be applied to many interesting real world scenarios such as attitude estimator design on the Lie group $\mathsf{SO}(3)$ or pose estimation on the Lie group $\mathsf{SE}(3)$ [3], [7], [11], [18], [20], [26]. For the specific cases of attitude estimation on $\mathsf{SO}(3)$ and pose estimation on $\mathsf{SE}(3)$, some methods have been proposed for the concurrent estimation of an unknown constant bias corrupting the velocity measurement [18], [26], [27]. These methods strongly depend on particular properties of the specific Lie groups $\mathsf{SO}(3)$ or $\mathsf{SE}(3)$, and hence cannot be easily generalized to a broader class of Lie groups. To the authors' knowledge, there is no existing work on bias estimation for general classes of invariant systems with homogeneous output measurements.

In this paper, we propose a general observer design for invariant systems with biased velocity measurements. The observer consists of a group estimator coupled to a bias estimator. In line with the approach of [14]–[16] for the bias free case, our proposed group estimator employs the gradient of an invariant cost function as an innovation term. We design the bias estimator part using a Lyapunov approach that guarantees uniform local exponential stability of the estimation error dynamics around the zero error state for bounded observed state trajectories. The main contribution of this paper is a bias estimation algorithm for an invariant system on a general connected matrix Lie group together with a stability analysis that is valid for the general case. We also generalize the notion of homogeneous outputs by allowing multiple outputs, where each output belongs to a possibly different homogeneous output space and is modeled by an action of the Lie group. We propose a systematic method for the construction of suitable cost functions on Lie groups by employing and lifting cost functions designed on the homogeneous output spaces. We then use the gradient of the constructed cost function on the Lie group as an innovation term for the estimator. When the homogeneous output spaces are reductive, we prove that this methodology results in an observer that can be implemented using only the available sensor measurements. Hence, the proposed observer can be employed in real world applications.

A. Khosravian, J. Trumpf and R. Mahony are with the Research School of Engineering, Australian National University, Canberra ACT 2601, Australia (e-mail: alireza.khosravian@anu.edu.au; jochen.trumpf@anu.edu.au; robert.mahony@anu.edu.au).

C. Lageman is with the Mathematical Institute, University of Würzburg, 97074 Würzburg, Germany (e-mail: christian.lageman@mathematik.uni-wuerzburg.de).

This work was partially supported by the Australian Research Council through the ARC Discovery Project DP120100316 "Geometric observer theory for mechanical control systems".

The paper is organized as follows. After briefly clarifying our notation in Section II, we formulate the problem in Section III. We introduce the proposed observer in Section IV and investigate its convergence properties. Section V is devoted to the systematic construction of cost functions on Lie groups. The main results of the paper are summarized in Theorem 2 (stability analysis of the observer) and Proposition 4 (constructing cost functions on Lie groups). A detailed example in Section VI and brief conclusions in Section VII complete the paper.

II. Notation and definitions

Let G be a connected matrix Lie group with associated Lie algebra $\mathfrak{g}$. Denote the identity element of G by I. Left (resp. right) multiplication of $X \in G$ by $S \in G$ is denoted by $L_S X = SX$ (resp. $R_S X = XS$). The Lie algebra $\mathfrak{g}$ can be identified with the tangent space at the identity element of the Lie group, i.e. $\mathfrak{g} \cong T_I G$. For any $u \in \mathfrak{g}$, one can obtain a tangent vector at $S \in G$ by left (resp. right) translation of u denoted by $Su := T_I L_S u \in T_S G$ (resp. $uS := T_I R_S u \in T_S G$). The adjoint map at the point $S \in G$ is denoted by $\mathrm{Ad}_S \colon \mathfrak{g} \to \mathfrak{g}$ and is defined by $\mathrm{Ad}_S v := T_S R_{S^{-1}} T_I L_S v = SvS^{-1}$ for all $v \in \mathfrak{g}$. A cost function $\phi \colon G \times G \to \mathbb{R}$ is right-invariant if $\phi(\hat{X}S, XS) = \phi(\hat{X}, X)$ for all $X, \hat{X}, S \in G$. We denote a Riemannian metric at a point $X \in G$ by $\langle ., . \rangle_X$. The Riemannian metric is right-invariant if for all $u, v \in T_X G$ we have $\langle u, v \rangle_X = \langle uX^{-1}, vX^{-1} \rangle_I$. We denote an inner product on $\mathfrak{g}$ by $\langle\langle ., . \rangle\rangle$. Consider a linear map $H \colon \mathfrak{g} \to \mathfrak{g}$. The Hermitian adjoint of H is denoted by H^* and is defined as the linear map $H^* \colon \mathfrak{g} \to \mathfrak{g}$ such that for all $v_1, v_2 \in \mathfrak{g}$ we have $\langle\langle Hv_1, v_2 \rangle\rangle = \langle\langle v_1, H^* v_2 \rangle\rangle$. We use the acronym SPD for symmetric positive definite.

III. Problem formulation

Consider a class of left-invariant systems on G given by

$$\dot{X} = Xu \tag{1}$$

where $u \in \mathfrak{g}$ is the group velocity. Although the ideas presented in this paper are based on the above left-invariant dynamics, they can easily be modified to suite a right-invariant system as was done for instance in [16]. If we interpret $u \colon \mathbb{R} \to \mathfrak{g}$ as an input of system (1), this input is called *admissible* if corresponding solutions for the system exist for all initial values and all initial times, and if these solutions are unique, continuously differentiable and exist for all time. For many systems of practical interest, e.g., affinely bounded, smooth systems, there is always a "rich" set of admissible inputs (every continuous input is admissible in this case). In most mechanical systems, u is interpreted as the velocity of physical objects. Hence, it is reasonable to assume that u is bounded and continuous.

Let $M_i,\ i = 1, \ldots, n$ denote a collection of n homogeneous spaces of G, termed *output spaces*. Denote the outputs of system (1) by $y_i \in M_i$. Suppose each output provides a partial measurement of X via $y_i = h_i(X, \mathring{y}_i)$ where $\mathring{y}_i$ is the reference output associated with y_i and h_i is a right action of G on M_i, i.e. $h_i(XS, y_i) = h_i(S, h_i(X, y_i))$ for all $y_i \in M_i$ and all $X, S \in G$. We will assume that the state X in (1) is observable from measurements of u and $y := (y_1, \ldots, y_n)$. We define the combined reference output $y_0 := (\mathring{y}_1, \ldots, \mathring{y}_n)$ and the combined right action $h(X, y_0) := (h_1(X, \mathring{y}_1), \ldots, h_n(X, \mathring{y}_n))$. The output y belongs to the orbit of G acting on the product space $M_1 \times \ldots \times M_n$ containing y_0, that is

$$\begin{aligned} M :=& \{y \in M_1 \times M_2 \ldots \times M_n \mid y = h(X, y_0),\ X \in G\} \\ &\subset M_1 \times M_2 \ldots \times M_n. \end{aligned}$$

Note that the combined action defined above is transitive on M. Hence, M is a homogeneous space of G while $M_1 \times \ldots \times M_n$ is not necessarily a homogeneous space [13].

We assume that a measurement of group velocity is available, but it is corrupted by a constant unknown additive bias. That is

$$u_y = u + b \tag{2}$$

where $u_y \in \mathfrak{g}$ is the measurement of u and $b \in \mathfrak{g}$ is an unknown constant bias.

Our goal is to design an observer that estimates X and b. Denoting the estimate of X by $\hat{X}$, we consider the following right-invariant group error,

$$E_r = \hat{X} X^{-1} \tag{3}$$

as was proposed in [16]. The above error resembles the usual $\hat{x} - x$ used in classical observer theory when $\hat{x}$ and x belong to a vector space. It is worth recalling that $\hat{X} = X$ if and only if $E_r = I$. Denoting the estimate of b by $\hat{b}$, we consider the following bias estimation error

$$\tilde{b} = \hat{b} - b. \tag{4}$$

Example 1: The attitude estimation problem has been investigated by a range of authors during the past decades [3], [7], [9], [18], [21], [25], [27]. The attitude of a rigid body can be identified with a rotation matrix $R \in \mathrm{SO}(3)$ representing the rotation from a body-fixed frame $\{B\}$ to the inertial frame $\{A\}$. Denoting the set of skew-symmetric matrices by $\mathfrak{g} = \mathfrak{so}(3)$, the Lie algebra of the Lie group $G = \mathrm{SO}(3)$, the natural body-fixed frame kinematics of this system is given by

$$\dot{R} = R\Omega \tag{5}$$

where $\Omega \in \mathfrak{so}(3)$ represents the angular velocity of $\{B\}$ with respect to $\{A\}$ expressed in $\{B\}$. The angular velocity measured by rate gyros is usually corrupted by additive unknown bias such that $\Omega_y = \Omega + b$ where Ω_y is the measured angular velocity and b denotes the unknown constant bias. Assume that partial attitude information is provided by measuring a constant inertial direction in $\{B\}$. Such a measurement can be provided by on-board sensor systems such as a magnetometer or an accelerometer. We recall that in most practical cases, the constant inertial direction $\mathring{y}_1$ is known *a priori*, for example in the case of using an accelerometer, the gravitational direction is known in the inertial frame. The measured direction y_1 is related to R and $\mathring{y}_1$ by

$$y_1 = R^\top \mathring{y}_1. \tag{6}$$

The vectors y_1 and $\mathring{y}_1$ are normalized to have unit norm such that the measured output lives in the sphere S^2. The output map (6) defines a right action of $\mathsf{SO}(3)$ on the homogeneous space $M_1 \cong \mathsf{S}^2$ by $h_1(R, \mathring{y}_1) := R^\top \mathring{y}_1$. The attitude estimation problem is to use the measurement y_1 and the reference $\mathring{y}_1$ in order to estimate the attitude matrix R. It is known that the attitude kinematics are in fact unobservable when only a single constant inertial direction is measured by the sensor system [18], [19], [25]. To resolve this issue, it is common to employ multiple sensors each measuring a different inertial direction. For instance, suppose $y_1, y_2 \in \mathsf{S}^2$ are two vector measurements associated with the inertial directions $\mathring{y}_1, \mathring{y}_2 \in \mathsf{S}^2$. Now we consider $y := (y_1, y_2)$ as the output and $y_0 := (\mathring{y}_1, \mathring{y}_2)$ as the reference output and construct the combined action $h(R, y_0) := (h_1(R, \mathring{y}_1), h_2(R, \mathring{y}_2)) = (R^\top \mathring{y}_1, R^\top \mathring{y}_2)$. Note that h is a transitive right action of $\mathsf{SO}(3)$ on the output space $M := \{y \in \mathsf{S}^2 \times \mathsf{S}^2 |\ y = (R^\top \mathring{y}_1, R^\top \mathring{y}_2)\}$ which is an orbit of $\mathsf{SO}(3)$ acting on $\mathsf{S}^2 \times \mathsf{S}^2$. It is well known that the system is observable if $\mathring{y}_1$ and $\mathring{y}_2$ are non-collinear [18], [19], [25], [27].

IV. Observer design and analysis

We propose the following *group estimator*,

$$\dot{\hat{X}} = \hat{X}(u_y - \hat{b}) - \operatorname{grad}_1 \phi(\hat{X}, X) \tag{7}$$

where $\phi\colon G \times G \to \mathbb{R}^+$ is a *cost function* and $\operatorname{grad}_1 \phi$ denotes the Riemannian gradient of ϕ with respect to its first variable. The above structure for the group estimator was proposed in [16] for the case of a bias free group velocity measurement.

We propose the following *bias estimator*,

$$\dot{\hat{b}} = \gamma \mathrm{Ad}^*_{\hat{X}} \left(\operatorname{grad}_1 \phi(\hat{X}, X) \hat{X}^{-1} \right), \tag{8}$$

where γ is a positive scalar and $\mathrm{Ad}^*_{\hat{X}}\colon \mathfrak{g} \to \mathfrak{g}$ denotes the Hermitian adjoint of $\mathrm{Ad}_{\hat{X}}$, cf. Section II. In Section V below, we propose a systematic method for constructing suitable cost functions $\phi(\hat{X}, X)$ on Lie groups by employing and lifting cost functions designed on the homogeneous output spaces. When the homogeneous output spaces are *reductive* [13], we will show that $\operatorname{grad}_1 \phi(\hat{X}, X)$ (and consequently the observer (7)-(8)) is *implementable* based on sensor measurements, i.e. it is only a function of $\hat{X}, y, y_0$ and it is not a direct function of the unknown variable X. The properties of the above observer are summarized in the following theorem.

Theorem 2: Consider the observer (7)-(8) for the system (1)-(2) and assume that both the admissible group velocity u and the system state X are bounded. Assume that $\phi(\hat{X}, X)\colon G \times G \to \mathbb{R}^+$ is right-invariant and that $\operatorname{grad}_1 \phi(\hat{X}, X)$ is computed with respect to a right-invariant Riemannian metric on G. Suppose in addition that ϕ is chosen such that $\operatorname{grad}_1 \phi(I, I) = 0$ and $\operatorname{Hess}_1 \phi(I, I)$ is SPD which in particular implies that $\phi(., I)\colon G \to \mathbb{R}^+$ has an isolated local minimum at $\phi(I, I)$. Then the error $(E_r(t), \tilde{b}(t))$ is uniformly locally exponentially stable around $(I, 0)$. Moreover, if all *sublevel sets* of $\phi(., I)\colon G \to \mathbb{R}^+$ defined by $\{E \in G |\ \phi(E, I) \leq c\}$, $c \in \mathbb{R}^+$ are compact then the error is bounded for all $t > 0$ and all initial conditions.

Note that Theorem 2 requires the existence of a right-invariant Riemannian metric on G and a suitable right-invariant cost ϕ. A right-invariant Riemannian metric can easily be constructed on any Lie group by transporting an inner product on the Lie algebra to other tangent spaces by right translation. In Section V below, we propose a method for designing a right-invariant cost function ϕ based on invariant cost functions on the homogeneous output spaces. This method is illustrated in Section VI by a detailed example, namely attitude estimator design on $\mathsf{SO}(3)$.

Also note that the requirement for the system state to be bounded is no restriction in practice and is automatically fulfilled for a compact matrix Lie group G such as $\mathsf{SO}(3)$.

A. Proof of Theorem 2

Replacing u_y from (2) into (7) and using (1), the dynamics of the group error (3) are given by

$$\begin{aligned} \dot{E}_r &= \dot{\hat{X}} X^{-1} + \hat{X} (\dot{X^{-1}}) \\ &= \hat{X} u X^{-1} - \hat{X} \tilde{b} X^{-1} - \operatorname{grad}_1 \phi(\hat{X}, X) X^{-1} - \hat{X} u X^{-1}. \end{aligned} \tag{9}$$

In light of the right-invariant property of the considered cost function and the Riemannian metric, [16, Lemma 16] implies

$$\operatorname{grad}_1 \phi(\hat{X}, X) X^{-1} = \operatorname{grad}_1 \phi(E_r, I). \tag{10}$$

Using (10), the group error dynamics (9) is simplified to

$$\dot{E}_r = -\mathrm{Ad}_{\hat{X}} \tilde{b} E_r - \operatorname{grad}_1 \phi(E_r, I). \tag{11}$$

Consider the candidate Lyapunov function,

$$\mathcal{L}(E_r, \tilde{b}) = \phi(E_r, I) + \frac{1}{2\gamma} \langle\langle \tilde{b}, \tilde{b} \rangle\rangle. \tag{12}$$

The time derivative of $\mathcal{L}$ is given by

$$\dot{\mathcal{L}}(E_r, \tilde{b}) = \langle \operatorname{grad}_1 \phi(E_r, I), \dot{E}_r \rangle_{E_r} + \frac{1}{\gamma} \langle\langle \tilde{b}, \dot{\tilde{b}} \rangle\rangle. \tag{13}$$

Recalling $\dot{\tilde{b}} = \dot{\hat{b}}$ and substituting $\dot{E}_r$ from (11) in (13), we have

$$\begin{aligned} \dot{\mathcal{L}}(E_r, \tilde{b}) = &- \|\operatorname{grad}_1 \phi(E_r, I)\|^2 \\ &- \langle \operatorname{grad}_1 \phi(E_r, I), \mathrm{Ad}_{\hat{X}} \tilde{b} E_r \rangle_{E_r} + \frac{1}{\gamma} \langle\langle \tilde{b}, \dot{\hat{b}} \rangle\rangle. \end{aligned}$$

Using the right-invariant property of the Riemannian metric for the second term on the right hand side, it follows that

$$\begin{aligned} \dot{\mathcal{L}}(E_r, \tilde{b}) &= -\|\operatorname{grad}_1 \phi(E_r, I)\|^2 \\ &\quad - \langle\langle \operatorname{grad}_1 \phi(E_r, I) E_r^{-1}, \mathrm{Ad}_{\hat{X}} \tilde{b} \rangle\rangle + \frac{1}{\gamma} \langle\langle \tilde{b}, \dot{\hat{b}} \rangle\rangle \\ &= -\|\operatorname{grad}_1 \phi(E_r, I)\|^2 \\ &\quad + \frac{1}{\gamma} \langle\langle \tilde{b}, -\gamma \mathrm{Ad}^*_{\hat{X}} \left(\operatorname{grad}_1 \phi(E_r, I) E_r^{-1} \right) + \dot{\hat{b}} \rangle\rangle \end{aligned} \tag{14}$$

where the definition of the Hermitian adjoint is used to simplify the second term on the right hand side. Moreover, recalling (8) and using (10) we have

$$\begin{aligned}\dot{\tilde{b}} &= \gamma \mathrm{Ad}^*_{\hat{X}}\left(\ \mathrm{grad}_1\phi(\hat{X}, X)\hat{X}^{-1}\right) \\ &= \gamma \mathrm{Ad}^*_{\hat{X}}\left(\mathrm{grad}_1\phi(E_r, I)E_r^{-1}\right). \end{aligned} \tag{15}$$

Replacing $\dot{\tilde{b}}$ in (14) from (15) yields

$$\dot{\mathcal{L}}(E_r, \tilde{b}) = -\|\mathrm{grad}_1\phi(E_r, I)\|^2 \leq 0 \tag{16}$$

which implies that the Lyapunov function is non-increasing along the system trajectories and we have $\mathcal{L}(t) \leq \mathcal{L}(t_0)$ for all $t > 0$. This proves that the error signals are bounded near $(I, 0)$ if the initial errors belong to a sufficiently small neighborhood of $(I, 0)$. Moreover, if all sublevel sets of $\phi(., I)\colon G \to \mathbb{R}^+$ are compact then the errors are bounded for all initial conditions. Since the system state X is assumed bounded, it follows that the observer states $\hat{X}$ and $\hat{b}$ are also bounded.

In light of (11) and (15), one obtains the following closed loop error dynamics,

$$\dot{E}_r = -(\mathrm{Ad}_{\hat{X}}\tilde{b})E_r - \ \mathrm{grad}_1\phi(E_r, I), \tag{17}$$

$$\dot{\tilde{b}} = \gamma \mathrm{Ad}^*_{\hat{X}}\left(\mathrm{grad}_1\phi(E_r, I)E_r^{-1}\right). \tag{18}$$

Let $\epsilon, \delta \in \mathfrak{g}$ denote the first order approximations of E and $\tilde{b}$, respectively. Linearizing the error dynamics around $(I, 0)$ and neglecting all terms of quadratic or higher order in (ϵ, δ) yields

$$\dot{\epsilon} = -\mathrm{Ad}_{\hat{X}}\delta - \mathrm{Hess}_1\phi(I, I)\epsilon, \tag{19}$$

$$\dot{\delta} = \gamma \mathrm{Ad}^*_{\hat{X}}(\mathrm{Hess}_1\phi(I, I)\epsilon), \tag{20}$$

where $\mathrm{Hess}_1\phi(I, I)$ denotes the Hessian of the function $\phi(., I)\colon G \to \mathbb{R}$ evaluated at the point (I, I), which is assumed to be SPD. To analyze the stability of system (19)-(20), we rewrite that system as

$$\begin{bmatrix} \dot{\epsilon} \\ \dot{\delta} \end{bmatrix} = \begin{bmatrix} -\mathrm{Hess}_1\phi(I, I) & -\mathrm{Ad}_{\hat{X}} \\ \gamma \mathrm{Ad}^*_{\hat{X}}\mathrm{Hess}_1\phi(I, I) & 0 \end{bmatrix} \begin{bmatrix} \epsilon \\ \delta \end{bmatrix}. \tag{21}$$

We observe that the linear map

$$\begin{bmatrix} -\mathrm{Hess}_1\phi(I, I) & -\mathrm{Ad}_{\hat{X}} \\ \gamma \mathrm{Ad}^*_{\hat{X}}\mathrm{Hess}_1\phi(I, I) & 0 \end{bmatrix}$$

on the right hand side of (21) resembles the following matrix structure

$$\Psi := \begin{bmatrix} A & B \\ -B^\top P & 0 \end{bmatrix},$$

where A is a Hurwitz matrix and P is a SPD matrix. Stability analysis of systems with dynamics $\dot{x} = \Psi x$ is of interest in adaptive control design and has been investigated carefully in the literature during the past decades [17], [22], [23]. The following lemma is a coordinate-free version of [17, Theorem 1], adapted to our setting. Applying it to system (21) will complete the proof of Theorem 2.

Lemma 3: Suppose $x, y \in \mathfrak{g}$ and (x, y) is governed by the following linear time-varying parameterized dynamics

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} A(t, \lambda) & B(t, \lambda) \\ -B^*(t, \lambda)P(t, \lambda) & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{22}$$

where $\lambda \in \mathcal{D}$ is a constant parameter and $\mathcal{D}$ is a closed not necessarily compact set. Here, A, B and P are bounded continuous maps of (t, λ), the maps B and P are continuously differentiable with respect to time, the time derivative of B is uniformly bounded (uniform in λ) and $B^*(t, \lambda)\colon \mathfrak{g} \to \mathfrak{g}$ denotes the Hermitian adjoint of the linear map $B(t, \lambda)\colon \mathfrak{g} \to \mathfrak{g}$, cf. Section II. Suppose that $A(t, \lambda)\colon \mathfrak{g} \to \mathfrak{g}$ is a linear map and that $P(t, \lambda)\colon \mathfrak{g} \to \mathfrak{g}$ is a SPD linear map such that the map $Q(t, \lambda) := \dot{P}(t, \lambda) + A^*(t, \lambda)P(t, \lambda) + P^*(t, \lambda)A(t, \lambda)\colon \mathfrak{g} \to \mathfrak{g}$ is symmetric negative definite. Finally, suppose that positive scalars p_m, p_M, q_m, q_M exist such that $p_m \leq \langle\langle P(t, \lambda)v_1, v_2\rangle\rangle \leq p_M$ and $-q_m \leq \langle\langle Q(t, \lambda)w_1, w_2\rangle\rangle \leq -q_M$ for all $v_1, v_2, w_1, w_2 \in \mathfrak{g}$ and all $(t, \lambda) \in \mathbb{R}^+ \times \mathcal{D}$. Then the equilibrium $(0, 0)$ of system (22) is uniformly exponentially stable (uniform with respect to both, initial conditions and λ) if and only if there exists a $T > 0$ so that for all $t \geq 0$ the linear map $\int_t^{t+T} B^*(\tau, \lambda)B(\tau, \lambda)d\tau\colon \mathfrak{g} \to \mathfrak{g}$ is uniformly λ-SPD for all $\lambda \in \mathcal{D}$. That is, there exist $\mu > 0$ and $T > 0$ such that for all $t \geq 0$, all $0 \neq w \in \mathfrak{g}$ and all $\lambda \in \mathcal{D}$

$$\left\langle\left\langle \left(\int_t^{t+T} B^*(\tau, \lambda)B(\tau, \lambda)d\tau\right)w, w\right\rangle\right\rangle \geq \mu\langle\langle w, w\rangle\rangle. \tag{23}$$

To apply Lemma 3 to system (21), consider initial conditions $X(t_0)$ for system (1) and $(\hat{X}(t_0), \hat{b}(t_0))$ for the estimator (7)-(8). Supposing $\lambda = (t_0, X(t_0), \hat{X}(t_0), \hat{b}(t_0)) \in \mathcal{D}$ where $\mathcal{D} := \mathbb{R} \times G \times G \times \mathfrak{g}$, the trajectory $\hat{X}$ can be viewed as a function of t and λ. Now, consider $A := -\mathrm{Hess}_1\phi(I, I)$, $B(t, \lambda) := -\mathrm{Ad}_{\hat{X}(t,\lambda)}$, $B^*(t, \lambda) = -\mathrm{Ad}^*_{\hat{X}(t,\lambda)}$, $P := \gamma\mathrm{Hess}_1\phi(I, I)$. Note that A and P are independent of (t, λ) in this case and both B and the time derivative of B are uniformly bounded since u, X, $\hat{X}$ and $\hat{b}$ are all bounded. Since $\mathrm{Hess}_1\phi(I, I)$ is SPD, we have $\mathrm{Hess}_1\phi(I, I) = \mathrm{Hess}_1^*\phi(I, I)$. It follows that $\dot{P} + A^*P + PA = -2\gamma\mathrm{Hess}_1^*\phi(I, I)\mathrm{Hess}_1\phi(I, I)$ is symmetric negative definite. It only remains to investigate the requirement imposed by (23). There exists $c > 0$ such that for all $0 \neq w \in \mathfrak{g}$ we have $\langle\langle \mathrm{Ad}_{\hat{X}}w, \mathrm{Ad}_{\hat{X}}w\rangle\rangle \geq c\langle\langle w, w\rangle\rangle$ since the adjoint representation of uniformly bounded trajectories is uniformly bounded away from the set of singular operators on $\mathfrak{g}$ for any real finite-dimensional connected Lie group G. Using the definition of Hermitian adjoint, we have $\langle\langle \mathrm{Ad}^*_{\hat{X}}\mathrm{Ad}_{\hat{X}}w, w\rangle\rangle \geq c\langle\langle w, w\rangle\rangle$. Integrating both sides yields $\left\langle\left\langle \left(\int_t^{t+T} \mathrm{Ad}^*_{\hat{X}}\mathrm{Ad}_{\hat{X}}d\tau\right)w, w\right\rangle\right\rangle \geq cT\langle\langle w, w\rangle\rangle$ for any arbitrary trajectory of $\hat{X}$ starting from any arbitrary t_0. That is to say (23) holds for all $\lambda \in \mathcal{D}$, which completes the requirements of Lemma 3. Hence, the equilibrium $(0, 0)$ of the linearized system (21) is uniformly exponentially stable and consequently the equilibrium $(I, 0)$ of the nonlinear closed loop dynamics (17)-(18) is uniformly locally exponentially stable which completes the proof of Theorem 2.

Note that owing to the parameter-dependent analysis, the obtained exponential stability is uniform with respect to the choice of all initial conditions in λ and not only with respect to the choice of $E_r(t_0)$ and $\tilde{b}(t_0)$ for a given $\hat{X}$.

V. Constructing a cost function on the Lie group

In the previous section we proposed the observer (7)-(8) that employs the gradient of a cost function ϕ as its innovation term. In order to guarantee the stability of the observer via Theorem 2, the employed cost function must be right-invariant, must satisfy $\operatorname{grad}_1\phi(I,I) = 0$, and $\operatorname{Hess}_1\phi(I,I)$ must be SPD. In this section, we propose a method for constructing a suitable cost function on the Lie group G by employing cost functions on the homogeneous output spaces $M_i,\ i = 1,\ldots,n$.

Proposition 4: Suppose $f_i\colon M_i \times M_i \to \mathbb{R}^+,\ (\hat{y}_i, y_i) \mapsto f_i(\hat{y}_i, y_i)$ are cost functions on $M_i,\ i = 1,\ldots,n$. Lift each $f_i(\hat{y}_i, y_i)$ from M_i to G via $\phi_i(\hat{X}, X) := f_i(h_i(\hat{X}, \mathring{y}_i), h_i(X, \mathring{y}_i))$ and construct the cost function $\phi(\hat{X}, X) := \sum_{i=1}^n \phi_i(\hat{X}, X)$. The following properties hold for $\phi\colon G \times G \to \mathbb{R}^+$.

(a) Suppose each f_i is invariant under the action h_i of G, that is $f_i(\hat{y}, y) = f_i(h_i(S, \hat{y}_i), h_i(S, y_i))$ for all $S \in G, i = 1,\ldots,n$. Then the cost function $\phi(\hat{X}, X)$ is right-invariant.

(b) Denote the differential of $f_i(.,\mathring{y}_i)\colon M_i \to \mathbb{R}^+$ with respect to the first coordinate by $\mathrm{d}_1 f_i(., y_0)\colon TM \to \mathbb{R}$. If $\mathrm{d}_1 f_i(\mathring{y}_i, \mathring{y}_i) = 0$ for all $i = 1,\ldots,n$ then $\operatorname{grad}_1\phi(I,I) = 0$. If, additionally, $\operatorname{Hess}_1 f_i(\mathring{y}_i, \mathring{y}_i)$ is SPD for all $i = 1,\ldots,n$ and if $\bigcap_{i=1}^n T_I \mathrm{stab}_{h_i}(\mathring{y}_i) = \{0\}$, where $\mathrm{stab}_{h_i}(\mathring{y}_i)$ denotes the *stabilizer* of $\mathring{y}_i$ with respect to the action h_i defined by $\mathrm{stab}_{h_i}(\mathring{y}_i) := \{X \in G | h_i(X, \mathring{y}_i) = \mathring{y}_i\}$, then $\operatorname{Hess}_1\phi(I,I)$ is SPD.

(c) If the functions $f_i(.,\mathring{y}_i)\colon M_i \to \mathbb{R}^+,\ i = 1,\ldots,n$ have compact sublevel sets and $\bigcap_{i=1}^n \mathrm{stab}_{h_i}(\mathring{y}_i)$ is compact, then $\phi(.,I)\colon G \to \mathbb{R}^+$ has compact sublevel sets.

(d) If $M_1,\ldots,M_n$ are *reductive* homogeneous spaces of G, then $\operatorname{grad}_1\phi(\hat{X}, X)$ is implementable based on sensor measurements, i.e. $\operatorname{grad}_1\phi(\hat{X}, X)$ is only a function of $\hat{X}, y_i, \mathring{y}_i$ (which are all available) and it is not a direct function of X (which is not available for measurement).

Proof:

(a) For any $S \in G$ we compute $\phi(\hat{X}S, XS) = \sum_{i=1}^n f_i(h_i(\hat{X}S, \mathring{y}_i), h_i(XS, \mathring{y}_i)) = \sum_{i=1}^n f_i(h_i(S, h_i(\hat{X}, \mathring{y}_i)), h_i(S, h_i(X, \mathring{y}_i))) = \sum_{i=1}^n f_i(h_i(\hat{X}, \mathring{y}_i), h_i(X, \mathring{y}_i)) = \phi(\hat{X}, X)$ which shows that ϕ is right-invariant.

(b) Define the map $h_{\mathring{y}_i}\colon G \to M_i$ by $h_{\mathring{y}_i}X := h_i(X, \mathring{y}_i)$. Differentiating both sides of $\phi(\hat{X}, I) = \sum_{i=1}^n f_i(h_i(\hat{X}, \mathring{y}_i), \mathring{y}_i)$ in an arbitrary direction $v \in T_I G$, using the chain rule and evaluating at $\hat{X} = I$ we have $\mathrm{d}_1\phi(I,I)[v] = \sum_{i=1}^n \mathrm{d}_1 f_i(\mathring{y}_i, \mathring{y}_i)[T_I h_{\mathring{y}_i} v]$. Hence, $\mathrm{d}_1 f_i(\mathring{y}_i, \mathring{y}_i) = 0,\ i = 1,\ldots,n$ implies $\mathrm{d}_1\phi(I,I)[v] = 0$ for all v which in turn yields $\operatorname{grad}_1\phi(I,I) = 0$. One can verify that $\operatorname{Hess}_1\phi_i(I,I) = (T_I h_{\mathring{y}_i})^* \operatorname{Hess}_1 f_i(\mathring{y}_i, \mathring{y}_i) T_I h_{\mathring{y}_i}$ where $(T_I h_{\mathring{y}_i})^*$ denotes the Hermitian adjoint of $T_I h_{\mathring{y}_i}$. Hence we have, $\operatorname{Hess}_1\phi(I,I) = \sum_{i=1}^n \operatorname{Hess}_1\phi_i(I,I) = \sum_{i=1}^n (T_I h_{\mathring{y}_i})^* \operatorname{Hess}_1 f_i(\mathring{y}_i, \mathring{y}_i) T_I h_{\mathring{y}_i}$ which together with the assumption that $\operatorname{Hess}_1 f_i(\mathring{y}_i, \mathring{y}_i),\ i = 1,\ldots,n$ are SPD implies that $\operatorname{Hess}_1\phi(I,I)$ is at least symmetric positive semidefinite and $\ker(\operatorname{Hess}_1\phi(I,I)) = \bigcap_{i=1}^n \ker(T_I h_{\mathring{y}_i})$. Since, $\ker(T_I h_{\mathring{y}_i}) = T_I \mathrm{stab}_{h_i}(\mathring{y}_i)$, we have $\bigcap_{i=1}^n \ker(T_I h_{\mathring{y}_i}) = \bigcap_{i=1}^n T_I \mathrm{stab}_{h_i}(\mathring{y}_i) = \{0\}$. Consequently, $\ker(\operatorname{Hess}_1\phi(I,I)) = \{0\}$ which implies that $\operatorname{Hess}_1\phi(I,I)$ is full rank and hence SPD.

(c) We have to show that $E \mapsto \sum_{i=1}^n f_i(h_i(E, \mathring{y}_i), \mathring{y}_i)$ has compact sublevel sets. Denote by S_c^i the sublevel set of $z_i \mapsto f_i(z_i, \mathring{y}_i)$ for the value c. The sublevel set of $(z_1,\ldots,z_n) \mapsto \sum_{i=1}^n f_i(z_i, \mathring{y}_i)$ for the value c is then a closed set contained in the compact set $S_c^1 \times \ldots \times S_c^n$ (the product is compact by Tychonoff's theorem [24]). It remains to show that the map $G \to M,\ E \mapsto (h_1(E, \mathring{y}_1),\ldots,h_n(E, \mathring{y}_n))$ is proper (i.e. the preimage of compact sets is compact). It is easy to see that this holds if and only if the stabilizer of $(\mathring{y}_1,\ldots,\mathring{y}_n)$ is compact. The latter is given by $\bigcap_{i=1}^n \mathrm{stab}_{h_i}(\mathring{y}_i)$.

(d) For any reductive homogeneous space M_i there is an invariant horizontal distribution H_i and an invariant Riemannian metric on G, denoted by $\langle .,. \rangle^i$, such that the projection of the metric from H_i onto TM_i induces an invariant Riemannian metric on M_i, denoted by $\prec .,. \succ^i$ [8]. Denote the gradient of $\phi_i(\hat{X}, X)$ and $f_i(\hat{y}_i, y_i)$ with respect to $\langle .,. \rangle^i$ and $\prec .,. \succ^i$ by $\operatorname{grad}_1^i\phi_i(\hat{X}, X)$ and $\operatorname{grad}_1^i f_i(\hat{y}_i, y_i)$, respectively. It has been shown in [14, Proposition 22] that $\operatorname{grad}_1^i\phi_i(\hat{X}, X) = (\operatorname{grad}_1^i f_i(\hat{y}_i, y_i))^{H_i(\hat{X})}$ where $(.)^{H_i(\hat{X})}$ denotes the unique horizontal lift associated with the reductive homogeneous output space M_i and $\hat{y}_i = h_i(\hat{X}, \mathring{y}_i)$. Now, consider an arbitrary Riemannian metric $\langle .,. \rangle$ on G. There exist SPD linear maps $Q(\hat{X}), Q_i(\hat{X})\colon T_{\hat{X}}G \to T_{\hat{X}}^*G$ such that $\langle v, w \rangle_{\hat{X}} = Q(\hat{X})[v][w]$ and $\langle v, w \rangle_{\hat{X}}^i = Q_i(\hat{X})[v][w]$ for all $v, w \in \mathfrak{g}$. Denoting the gradient of ϕ_i with respect to $\langle .,. \rangle$ by $\operatorname{grad}_1\phi_i$, it is easy to verify that $\operatorname{grad}_1\phi_i(\hat{X}, X) = Q^{-1}(\hat{X}) \circ Q_i(\hat{X})[\operatorname{grad}_1^i\phi_i(\hat{X}, X)]$ [1]. Hence, $\operatorname{grad}_1\phi(\hat{X}, X) = \sum_{i=1}^n Q^{-1}(\hat{X}) \circ Q_i(\hat{X})[(\operatorname{grad}_1^i f_i(\hat{y}_i, y_i))^{H_i(\hat{X})}]$. Consequently, $\operatorname{grad}_1\phi(\hat{X}, X)$ can be written only as a function of $\hat{X}, y_i, \mathring{y}_i$. ∎

Proposition 4 proposes a systematic method to construct a cost function ϕ which satisfies the requirements of Theorem 2. The gradient of this function can be employed as an innovation term for the observer. We will illustrate this method in the next section by providing a comprehensive example.

Note that part (b) of Proposition 4 imposes the additional condition $\bigcap_{i=1}^n T_I \mathrm{stab}_{h_i}(\mathring{y}_i) = \{0\}$. As will be discussed in the next section, for the attitude estimation problem this condition corresponds to availability of two or more noncollinear vector measurements. When the system has only one homogeneous output space M_1, the condition simplifies to $T_I\text{stab}_{h_1}(\mathring{y}_1) = \{0\}$ which is equivalent to $\text{stab}_{h_1}(\mathring{y}_1) = \{I\}$. This condition corresponds to the observability criterion discussed in [15] for the bias free case. Hence, the condition imposed in part (b) of Proposition 4 can be interpreted as a generalization of that observability criterion when multiple homogeneous output spaces are considered.

Note that in computing the derivative of the Lyapunov function in section IV-A, we implicitly assumed that the cost function ϕ is not an explicit function of time. Hence, the method proposed in Proposition 4 suits those systems whose reference outputs are constant with respect to time. Time-varying reference outputs have been investigated in [2], [10], [12], [25] for the special case of attitude estimation on $\mathsf{SO}(3)$. Nevertheless, in most practical cases, the reference outputs can be assumed constant [7], [18], [26], and Proposition 4 can be applied to design a suitable cost function.

In several robotics applications where the considered Lie group is $\mathsf{SO}(3)$ and where the outputs live in copies of S^2, as well as in the case where the considered Lie group is $\mathsf{SE}(3)$ and the homogeneous output spaces are copies of $\mathbb{R}^3$, it has been shown that the reductivity condition imposed by part (d) of Proposition 4 holds and hence the resulting observer is implementable based on sensor measurements [15], [11].

VI. Example: attitude estimation using biased angular velocity measurements

We return to the attitude estimation problem discussed in Example 1. We aim to employ the observer developed in Section IV and use the guidelines presented in Section V to design a suitable cost function. We first verify the condition imposed by part (b) of Proposition 4. We have $\text{stab}_{h_1}(\mathring{y}_1) = \{R \in \mathsf{SO}(3) |\ R^\top \mathring{y}_1 = \mathring{y}_1\}$ and $\text{stab}_{h_2}(\mathring{y}_2) = \{R \in \mathsf{SO}(3) |\ R^\top \mathring{y}_2 = \mathring{y}_2\}$. Hence we have, $T_I\text{stab}_{h_1}(\mathring{y}_1) = \{\Omega \in \mathfrak{so}(3) |\ \Omega\mathring{y}_1 = 0\}$ and $T_I\text{stab}_{h_2}(\mathring{y}_2) = \{\Omega \in \mathfrak{so}(3) |\ \Omega\mathring{y}_2 = 0\}$. Consider $\Omega = \omega_\times$ where $(.)_\times$ denotes the usual mapping from $\mathbb{R}^3$ to $\mathfrak{so}(3)$. We have $T_I\text{stab}_{h_1}(\mathring{y}_1) \bigcap T_I\text{stab}_{h_2}(\mathring{y}_2) = \{\omega \in \mathbb{R}^3 |\ \omega \times \mathring{y}_1 = 0, \omega \times \mathring{y}_2 = 0\}$ which implies that ω is parallel to both $\mathring{y}_1$ and $\mathring{y}_2$. Hence, $T_I\text{stab}_{h_1}(\mathring{y}_1) \bigcap T_I\text{stab}_{h_2}(\mathring{y}_2) = \{0\}$ if and only if $\mathring{y}_1$ and $\mathring{y}_2$ are non-collinear.

In order to employ the method presented in Section V, we need to construct suitable cost functions $f_1(\hat{y}_1, y_1)$ and $f_2(\hat{y}_2, y_2)$ on M_1 and M_2. Note that both M_1 and M_2 are copies of S^2 in this example. We consider $f_1(\hat{y}_1, y_1) = \frac{k_1}{2}\|\hat{y}_1 - y_1\|^2$ where $\|a\| = a^\top a$ is induced by the standard Euclidian norm on $\mathbb{R}^3$ and $k_1 > 0$ is a positive scalar. It is easy to verify that f_1 is invariant with respect to the action of $\mathsf{SO}(3)$ on S^2 given by (6). One can also verify that $\text{d}_1 f_1(\mathring{y}_1, \mathring{y}_1) = 0$ and $\text{Hess}_1 f_1(\mathring{y}_1, \mathring{y}_1)$ is SPD. Similarly, we propose the cost function $f_2(\hat{y}_2, y_2) = \frac{k_2}{2}\|\hat{y}_2 - y_2\|^2$ on M_2 where $k_2 > 0$. We lift $f_1(\hat{y}_1, y_1)$ and $f_2(\hat{y}_2, y_2)$ to obtain the following right-invariant cost function on $G \times G$.

$$\begin{aligned}\phi(\hat{R}, R) :&= f_1(\hat{R}^\top \mathring{y}_1, R^\top \mathring{y}_1) + f_2(\hat{R}^\top \mathring{y}_2, R^\top \mathring{y}_2) \\ &= \frac{k_1}{2}\|\hat{R}^\top \mathring{y}_1 - R^\top \mathring{y}_1\|^2 + \frac{k_2}{2}\|\hat{R}^\top \mathring{y}_2 - R^\top \mathring{y}_2\|^2.\end{aligned} \tag{24}$$

By Proposition 4 it is guaranteed that $\text{grad}_1\phi(I, I) = 0$ and $\text{Hess}_1\phi(I, I)$ is SPD. We consider the inner product $\langle\langle \Omega_1, \Omega_2\rangle\rangle = \text{tr}(\Omega_1^\top \Omega_2)$, for $\Omega_1, \Omega_2 \in \mathfrak{so}(3)$ induced by the Frobenius inner product in $\mathbb{R}^{3\times 3}$. One can build the following right-invariant Riemannian metric on $\mathsf{SO}(3)$ using right translation of the Lie algebra $\mathfrak{so}(3)$.

$$\langle \Omega_1\hat{R}, \Omega_2\hat{R}\rangle_{\hat{R}} := \langle\langle \Omega_1, \Omega_2\rangle\rangle = \text{tr}(\Omega_1^\top \Omega_2).$$

Recalling that $\text{grad}_1\phi(\hat{X}, X)$ can be computed using the differential of $\phi(\hat{X}, X)$ with respect to the first coordinate in an arbitrary tangential direction $\hat{\Omega}\hat{R} \in T_{\hat{R}}\mathsf{SO}(3)$, we have

$$\begin{aligned}\text{d}_1\phi(\hat{R}, R)[\hat{\Omega}\hat{R}] &= \langle \text{grad}_1\phi(\hat{R}, R), \hat{\Omega}\hat{R}\rangle_{\hat{R}} \\ &= \langle\langle \text{grad}_1\phi(\hat{R}, R)\hat{R}^\top, \hat{\Omega}\rangle\rangle,\end{aligned} \tag{25}$$

where the last equality is obtained using the right-invariant property of the considered Riemannian metric. One can compute $\text{d}_1\phi(\hat{R}, R)[\hat{\Omega}\hat{R}]$ by direct differentiation of (24) as,

$$\begin{aligned}&\text{d}_1\phi(\hat{R}, R)[\hat{\Omega}\hat{R}] \\ &\quad= k_1\mathring{y}_1^\top \hat{\Omega}\hat{R}(\hat{R} - R)^\top \mathring{y}_1 + k_2\mathring{y}_2^\top \hat{\Omega}\hat{R}(\hat{R} - R)^\top \mathring{y}_2 \\ &\quad= -\text{tr}\left(\hat{\Omega}^\top \hat{R}\big(k_1(\hat{y}_1 - y_1)\mathring{y}_1^\top + k_2(\hat{y}_2 - y_2)\mathring{y}_2^\top\big)\right).\end{aligned}$$

Here we have used the shorthand notation $\hat{y}_1 = \hat{R}^\top \mathring{y}_1$ and $\hat{y}_2 = \hat{R}^\top \mathring{y}_2$ corresponding to the lifting process. Using the notation $\mathbb{P}_{\mathfrak{so}(3)}(A) = \frac{1}{2}(A - A^\top)$ for the matrix projection from $\mathbb{R}^{3\times 3}$ onto $\mathfrak{so}(3)$ we have

$$\begin{aligned}&\text{d}_1\phi(\hat{X}, X)[\hat{\Omega}\hat{R}] \\ &\quad= -\text{tr}(\hat{\Omega}^\top \mathbb{P}_{\mathfrak{so}(3)}(\hat{R}\big(k_1(\hat{y}_1 - y_1)\mathring{y}_1^\top + k_2(\hat{y}_2 - y_2)\mathring{y}_2^\top\big)) \\ &\quad= \langle\langle -\mathbb{P}_{\mathfrak{so}(3)}(\hat{R}\big(k_1(\hat{y}_1 - y_1)\mathring{y}_1^\top + k_2(\hat{y}_2 - y_2)\mathring{y}_2^\top\big), \hat{\Omega}\rangle\rangle.\end{aligned} \tag{26}$$

Comparing (25) and (26) and noting that $\hat{\Omega}$ is arbitrary, we infer

$$\begin{aligned}&\text{grad}_1\phi(\hat{R}, R) \\ &\quad= -\mathbb{P}_{\mathfrak{so}(3)}(\hat{R}\big(k_1(\hat{y}_1 - y_1)\mathring{y}_1^\top + k_2(\hat{y}_2 - y_2)\mathring{y}_2^\top\big))\hat{R}.\end{aligned} \tag{27}$$

It is straight forward to show that for all $\hat{R} \in \mathsf{SO}(3)$ and $A \in \mathbb{R}^{3\times 3}$ we have $\mathbb{P}_{\mathfrak{so}(3)}(\hat{R}A) = \hat{R}\mathbb{P}_{\mathfrak{so}(3)}(A\hat{R})\hat{R}^\top$. Hence, (27) can be simplified to

$$\begin{aligned}&\text{grad}_1\phi(\hat{R}, R) \\ &\quad= -\hat{R}\mathbb{P}_{\mathfrak{so}(3)}((k_1(\hat{y}_1 - y_1)\mathring{y}_1^\top + k_2(\hat{y}_2 - y_2)\mathring{y}_2^\top)\hat{R}) \\ &\quad= -\hat{R}\mathbb{P}_{\mathfrak{so}(3)}(k_1(\hat{y}_1 - y_1)\hat{y}_1^\top + k_2(\hat{y}_2 - y_2)\hat{y}_2^\top) \\ &\quad= -\hat{R}(k_1(\hat{y}_1 y_1^\top - y_1\hat{y}_1^\top) + k_2(\hat{y}_2 y_2^\top - y_2\hat{y}_2^\top)).\end{aligned} \tag{28}$$

Note that (28) is implementable as was stated in Proposition 4, since S^2 is a reductive homogeneous space of $\mathsf{SO}(3)$. Using (28) as an innovation term, the attitude estimator is given by

$$\begin{aligned}\dot{\hat{R}} = \hat{R}(\Omega_y - \hat{b}) &+ k_1\hat{R}(\hat{y}_1 y_1^\top - y_1\hat{y}_1^\top) \\ &+ k_2\hat{R}(\hat{y}_2 y_2^\top - y_2\hat{y}_2^\top)\end{aligned} \tag{29}$$

where k_1 and k_2 are used as observer gains. Since $\mathsf{SO}(3) \subset \mathsf{GL}(3)$ is a matrix Lie group and our considered Riemannian metric is induced by the Frobenius inner product on $\mathbb{R}^{3\times 3}$, we can explicitly compute the map $\mathrm{Ad}^*_{\hat{R}} : \mathfrak{so}(3) \to \mathfrak{so}(3)$ as $\mathrm{Ad}^*_{\hat{R}} v = \mathbb{P}_{\mathfrak{so}(3)}\left(\hat{R}^\top v \hat{R}\right) = \hat{R}^\top v \hat{R}$ for all $v \in \mathfrak{so}(3)$. The bias estimator (8) then becomes

$$\begin{aligned}\dot{\hat{b}} &= -\gamma(\hat{R}^\top \hat{R}(k_1(\hat{y}_1 y_1^\top - y_1 \hat{y}_1^\top) + k_2(\hat{y}_2 y_2^\top - y_2 \hat{y}_2^\top))\hat{R}^{-1}\hat{R}) \\ &= -\gamma(k_1(\hat{y}_1 y_1^\top - y_1 \hat{y}_1^\top) + k_2(\hat{y}_2 y_2^\top - y_2 \hat{y}_2^\top) \\ &= -\gamma k_1(\hat{y}_1 y_1^\top - y_1 \hat{y}_1^\top) - \gamma k_2(\hat{y}_2 y_2^\top - y_2 \hat{y}_2^\top). \qquad (30)\end{aligned}$$

Using the property $ab^\top - ba^\top = (b \times a)_\times$, one can conclude that (29)-(30) corresponds to the complementary passive attitude estimator proposed in [18], [27] where almost globally asymptotic and locally exponential convergence of the estimation error to $(I, 0)$ has been proved.

VII. Conclusions

We developed a state observer for left-invariant systems on connected matrix Lie groups. The observer is capable of adaptive estimation of a constant unknown bias in the group velocity measurements. We employed the gradient of a suitable invariant cost function as the innovation term of the group estimator and we proved uniform local exponential convergence of the estimator trajectories to bounded system trajectories. The notion of homogeneous output spaces has been generalized to multiple outputs, each of which is modeled via a right action of the Lie group on a corresponding homogeneous output space. A method for constructing right-invariant cost functions on Lie groups is proposed based on invariant cost functions on the homogeneous output spaces. In this situation, a verifiable condition on the stabilizer of the reference outputs associated with the output spaces was derived that ensures the stability of the observer. This condition is consistent with the observability criterion discussed in [15]. We showed that the resulting observer is implementable based on available sensor measurements when the homogeneous output spaces are reductive. As a case study, attitude estimator design on the Lie group $\mathsf{SO}(3)$ was investigated where the outputs are vector measurements, each providing partial attitude information, and the angular velocity of the rigid body is measured via a biased gyro.

Although we discussed left-invariant systems with a collection of right output actions, the ideas presented in this paper can be modified in a straight forward manner to suite right-invariant systems with left output actions. In this paper, we focused on matrix Lie groups since they are more interesting in practical applications. However, we believe that the ideas presented in this paper can be applied to a wider class of Lie groups without requiring significant modifications.

References

[1] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[2] P. Batista, C. Silvestre, and P. Oliveira, "GES attitude observers-part II: Single vector observations," in *Proc. IFAC World Congr., Milan, Italy*, 2011, pp. 2991–2996.

[3] S. Bonnabel, P. Martin, and P. Rouchon, "A non-linear symmetry-preserving observer for velocity-aided inertial navigation," in *Proc. American Control Conf.*, 2006, pp. 2910–2914.

[4] ——, "Non-linear observer on Lie groups for left-invariant dynamics with right-left equivariant output," in *Proc. IFAC World Congr.*, 2008, pp. 8594–8598.

[5] ——, "Symmetry-preserving observers," *IEEE Trans. Autom. Control*, vol. 53, no. 11, pp. 2514–2526, 2008.

[6] ——, "Non-linear symmetry-preserving observers on Lie groups," *IEEE Trans. Autom. Control*, vol. 54, no. 7, pp. 1709–1713, 2009.

[7] S. Brás, R. Cunha, J. F. Vasconcelos, C. Silvestre, and P. Oliveira, "A nonlinear attitude observer based on active vision and inertial measurements," *IEEE Trans. Robot.*, vol. 27, no. 4, pp. 664–677, 2011.

[8] J. Cheeger and D. G. Ebin, *Comparison theorems in Riemannian geometry*. American Mathematical Society, 1975.

[9] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *Journal of Guidance, Control, and Dynamics*, vol. 30, pp. 12–28, 2007.

[10] H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi, "Attitude estimation using biased gyro and vector measurements with time-varying reference vectors," *IEEE Trans. Autom. Control*, vol. 57, no. 5, pp. 1332–1338, 2012.

[11] M.-D. Hua, M. Zamani, J. Trumpf, R. Mahony, and T. Hamel, "Observer design on the special euclidean group SE (3)," in *Proc. IEEE Conf. on Decision and Control and the European Control Conf.*, 2011, pp. 8169–8175.

[12] A. Khosravian and M. Namvar, "Globally exponential estimation of satellite attitude using a single vector measurement and gyro," in *Proc. 49th IEEE Conf. Decision and Control*, USA, Dec. 2010, pp. 364–369.

[13] S. Kobayashi and K. Nomizu, *Foundations of differential geometry Volume II*. Wiley-Interscience, 1996.

[14] C. Lageman, J. Trumpf, and R. Mahony, "Observer design for invariant systems with homogeneous observations," *arXiv preprint arXiv:0810.0748*, 2008.

[15] ——, "Observers for systems with invariant outputs," in *Proc. the European Control Conf.*, Budapest, Hungary, 2009, pp. 4587–4592.

[16] ——, "Gradient-like observers for invariant dynamics on a Lie group," *IEEE Trans. Autom. Control*, vol. 55, no. 2, pp. 367–377, 2010.

[17] A. Loría and E. Panteley, "Uniform exponential stability of linear time-varying systems: revisited," *Systems and Control Letters*, vol. 47, no. 1, pp. 13–24, 2002.

[18] R. Mahony, T. Hamel, and J.M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1203–1218, 2008.

[19] R. Mahony, T. Hamel, J. Trumpf, and C. Lageman, "Nonlinear attitude observers on SO(3) for complementary and compatible measurements: A theoretical study," in *Proc. 48th IEEE Conf. on Decision and Control*, China, December 2009, pp. 6407–6412.

[20] P. Martin and E. Salaün, "Design and implementation of a low-cost observer-based attitude and heading reference system," *Control Engineering Practice*, vol. 18, no. 7, pp. 712–722, 2010.

[21] N. Metni, J.-M. Pflimlin, T. Hamel, and P. Soueres, "Attitude and gyro bias estimation for a VTOL UAV," *Control Engineering Practice*, vol. 14, no. 12, pp. 1511–1520, 2006.

[22] A. Morgan and K. Narendra, "On the stability of nonautonomous differential equations $\dot{x} = A + B(t)x$, with skew symmetric matrix $B(t)$," *SIAM Journal on Control and Optimization*, vol. 15, no. 1, pp. 163–176, 1977.

[23] ——, "On the uniform asymptotic stability of certain linear nonautonomous differential equations," *SIAM Journal on Control and Optimization*, vol. 15, no. 1, pp. 5–24, 1977.

[24] J. R. Munkres, *Topology*. Prentice Hall Upper Saddle River, 2000.

[25] J. Trumpf, R. Mahony, T. Hamel, and C. Lageman, "Analysis of non-linear attitude observers for time-varying reference measurements," *IEEE Trans. Autom. Control*, vol. 57, no. 11, pp. 2789–2800, 2012.

[26] 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, pp. 155–166, 2010.

[27] J. Vasconcelos, C. Silvestre, and P. Oliveira, "A nonlinear observer for rigid body attitude estimation using vector observations," in *Proc. IFAC World Congr.*, Korea, July 2008, pp. 8599–8604.