State Estimation for Invariant Systems on Lie Groups with Delayed Output Measurements☆

Alireza Khosravian[a], Jochen Trumpf[a], Robert Mahony[a], Tarek Hamel[b]

[a]*Research School of Engineering, Australian National University, Canberra ACT 2601, Australia. (e-mails: alireza.khosravian@anu.edu.au; jochen.trumpf@anu.edu.au; robert.mahony@anu.edu.au)*
[b]*University of Nice-Sophia Antipolis I3S UNS-CNRS, France (e-mail: thamel@i3s.unice.fr)*

Abstract

This paper proposes a state estimation methodology for invariant systems on Lie groups where outputs of the system are measured with delay. The proposed method is based on a cascade of an observer and a predictor. The observer uses delayed measurements and provides estimates of delayed states. The predictor uses those estimates together with the current inputs of the system to compensate for the delay and to provide a prediction of the current state of the system. We consider three classes of left-invariant, right-invariant, and mixed-invariant systems and propose predictors tailored to each class. The key contribution of the paper is to exploit the underlying symmetries of systems to design novel predictors that are computationally simple and generic, in the sense that they can be combined with any stable observer or filter. We provide rigorous stability analysis demonstrating that the prediction of the current state converges to the current system trajectory if the observer state converges to the delayed system trajectory. Performance of the proposed approach is demonstrated using a sophisticated Software-In-The-Loop simulator indicating the robustness of the observer-predictor methodology even when large measurement delays are present.

1. Introduction

Recent work on applications in navigation and control of mechanical systems has demonstrated advantages of modeling such systems as kinematics on appropriate Lie groups [1–3]. In particular, state estimation for systems on Lie groups has been motivated by real world applications such as; attitude estimation on $\mathsf{SO}(3)$ [3–11], pose estimation on $\mathsf{SE}(3)$ [12–15], homography estimation on $\mathsf{SL}(3)$ [16], and motion estimation of chained systems on nilpotent Lie groups [17] (e.g. front-wheel drive cars or kinematic cars with k trailers).

State observers rely on fusing the measurements of inputs and outputs of systems to estimate their internal states. For mechanical systems, output measurements usually provide low frequency information about the state of the system while input measurements provide high frequency information about the rate of change of the state. For invariant systems on Lie groups, systematic observer design methodologies have been proposed that lead to strong stability and robustness properties of the resulting observers. Specifically, Bonnabel *et al.* [7] propose a symmetry-preserving observer that utilizes the invariance of the system and the moving frame method, leading to local convergence properties of the observers. The authors propose methods in [18–20] to achieve almost globally convergent observers using globally well defined errors and error dynamics. Other work in this area is the consideration of multiple output maps formulated by actions of the Lie group on homogeneous output spaces and also adaptive estimation of an unknown input bias [7, 19–21]. All of the observer design methodologies in this literature require delay free measurement of the current outputs and inputs of the system. In many practical scenarios, however, measurement of the outputs of the system is inherently delayed (comparing to the ideal outputs of the system) while the inputs are measured without significant delays. For example, in the velocity aided attitude estimation scenario, measurements of linear velocity (output) provided by commercial GPS units are usually delayed with respect to the actual velocity of the vehicle. The delay can be several hundred milliseconds (and even up to half a second) due to various environmental effects and in-sensor processing delays [22]. In contrast, measurements of the vehicle's angular velocity and linear acceleration provided by an onboard IMU are almost instantaneous. Another example is attitude estimation for satellites using star trackers and gyros. The image processing inside a star-tracker sensor can cause significant delays in the order of tens of milliseconds while gyroscope measurements on the satellite are obtained without significant delays. Similar delay problems occur in attitude estimation for aerial robots when vision based sensors compute landmarks. Even in instrumented indoor flight environments, the attitude data from devices such as VICON or OptiTrack are delayed by the communication channel from these sensors to the onboard attitude estimation system of the vehicle. Finally, heavy filtering of noisy sensor measurements before data fusion can introduce significant delays in the filtered data. This is particularly important, for example, for obtaining air velocity

☆This work was partially supported by the Australian Research Council through the ARC Discovery Project DP120100316 "Geometric Observer Theory for Mechanical Control Systems" and by the project "Sensory Control of Aerial Robot" from French Agence Nationale de la Recherche through the ANR ASTRID SCAR.

measurements from noisy air pressure readings [23]. It is well understood that measurement delays can negatively affect the stability and robustness of observers or filters and degrade their performance [24–31].

The classical approach to tackle sensor delay in estimation problem is to take an estimator that has the desired performance for delay free measurements, and modify its innovation term such that it compares each delayed measurement with its corresponding backward time-shifted estimate. If the delay-free estimator has a Lyapunov stability proof, the stability analysis for the modified estimator is approached by using Lyapunov-Krasovskii or Lyapunov-Razumikhin functions [26–28, 32, 33]. Although commonly used in practice, such modified estimators require complicated stability analyses and careful and conservative gain tuning which may lead to poor transient responses. Combined observer-predictor design is an alternative method that has received recent interests, see e.g. [24, 25, 29–31, 34] and the references therein. These methods take observers that use delayed measurements to provide estimates of the delayed state and combine them with appropriate predictors to compensate for the delay. The authors have recently proposed a cascade observer-predictor methodology to handle sensor delay in the particular case of attitude estimation on the Lie group $\mathsf{SO}(3)$ [35]. Exponential convergence of the attitude estimate to its current value was shown using the specific algebraic structure of the Lie group $\mathsf{SO}(3)$. To the authors' knowledge, there is currently no state estimation methodology (with stability proof) for systems on general Lie groups when output measurements are affected by delay.

In this paper, we tackle the problem of state estimation for systems on general matrix Lie groups when measurements of system outputs are delayed. We propose an observer-predictor methodology for three classes of systems on Lie groups; left-invariant, right-invariant, and mixed-invariant. Our proposed methodology employs observers that use the delayed output measurements and estimate the delayed state of the system. We propose dynamic predictors that use the delayed estimates from the observers together with the current estimates of the inputs in order to predict the current state of the system (see Fig. 1). The key contribution of the paper is effective use of the underlying symmetry of systems in order to design the predictors such that the observer-predictor pairs are co-stable meaning that the observer-predictor combination provides (asymptotically/exponentially) stable estimates of the current state if the observer itself provides (asymptotically/exponentially) stable estimates of the delayed state. We ensure separation of the observer design problem from the predictor design problem which gives us the freedom of employing any stable observer or filter without affecting the proposed co-stability analysis. An advantage of this approach is that the gain tuning process of the observer is independent of the magnitude of the delay, yielding robust performance of the observer-predictor pairs even with large delays. Rather than assuming the availability of current input measurements, we consider a more general case where only estimates of the current inputs of the system are available. This generalization is particularly useful when the input measurements have unknown biases and scaling factors that are adaptively estimated at the observer stage. The proposed predictors are recursive and hence computationally cheap making them ideal for embedded implementation in real-world applications. As an example, we consider the velocity aided attitude estimation problem and we provide realistic numerical simulations using a sophisticated Software-In-The-Loop (SITL) system designed for ArduPilot/APM, an open-source autopilot system that is widely used among the UAV enthusiasts community [36–40]. Using the SITL simulator, we demonstrate robustness of the observer-predictor approach in practical situations even when large GPS delays are present.

The paper is organized as follows. After briefly clarifying our notation in Section 2, we formulate the problem in Section 3. The observer-predictor approach is discussed in Section 4 where the proposed predictors are introduced and their co-stability properties are investigated. The relation of the proposed recursive predictors to the non-recursive predictors of [41] is discussed in Section 5. Realistic simulation studies in Section 6 and brief conclusions in Section 7 complete the paper.

2. Notations and Definitions

Let G be a matrix Lie group with associated Lie algebra $\mathfrak{g}$ which is identified with the tangent space at the identity element of the Lie group. Choose an inner product $\langle\!\langle .,. \rangle\!\rangle$ on $\mathfrak{g}$ and denote the corresponding induced norm on $\mathfrak{g}$ by $\|.\|$. Denote by $\langle .,. \rangle^r_X$ (resp. $\langle .,. \rangle^l_X$) a right-invariant (resp. left-invariant) Riemannian metric at a point $X \in G$ induced by $\langle\!\langle .,. \rangle\!\rangle$ using right (resp. left) translation of the Lie algebra $\mathfrak{g}$. Denote the corresponding induced right-invariant (resp. left-invariant) norm on T_XG by $|.|^r_X$ (resp. $|.|^l_X$). Denote the geodesic distance with respect to (wrt.) $\langle .,. \rangle^r$ (resp. $\langle .,. \rangle^l$) between the points $Y \in G$ and $Z \in G$ by $d_r(Y,Z)$ (resp. $d_l(Y,Z)$). We denote the condition number of a matrix $X \in G$ by $\mathrm{cond}(X)$. For the adjoint map $\mathrm{Ad}_X : \mathfrak{g} \to \mathfrak{g}$, the notation $\overline{\sigma}(\mathrm{Ad}_X)$ denotes the largest singular value of the adjoint map wrt. the norm $\|.\|$, defined by $\overline{\sigma}(\mathrm{Ad}_X) := \max_{\|v\|=1} \|\mathrm{Ad}_X v\|$. We say the trajectory $a(t) \in \mathbb{R}^+$ converges to zero and denote $a(t) \to 0$ if $\lim_{t\to+\infty} a(t) = 0$. We write $a(t) \xrightarrow{\exp} 0$ and we say $a(t)$ converges exponentially to zero if there exist positive constants c and α such that $a(t) \leq c \exp(-\alpha t)$ for all $t \geq 0$.

3. Problem formulation

We consider three classes of invariant systems on the Lie group G given by

$$\dot{X}_l(t) = X_l(t)u_l(t), \qquad X_l(0) = X_{l_0} \tag{1}$$

$$\dot{X}_r(t) = u_r(t)X_r(t), \qquad X_r(0) = X_{r_0} \tag{2}$$

$$\dot{X}_m(t) = X_m(t)u_l(t) + u_r(t)X_m(t), \quad X_m(0) = X_{m_0}, \tag{3}$$

where X_l, X_r and $X_m \in G$ are internal states of the systems (1), (2), and (3), respectively, and $u_l, u_r \in \mathfrak{g}$ are input signals. We drop the subscripts l, r and m wherever this is possible without causing confusion. Similar to [18, 21], we call the systems

(1) and (2) left-invariant and right-invariant, respectively. We call the system (3) *mixed-invariant* since its vector field is composed of a left-invariant term $X_m(t)u_l(t)$ and a right-invariant term $u_r(t)X_m(t)$, but the combined vector field is in general neither right nor left invariant. We assume that the inputs $u_l(t)$ and $u_r(t)$ are admissible in the sense that corresponding solutions for the relevant systems exist for all initial conditions and that these solutions are unique and continuously differentiable for all time. In the context of mechanical systems, equation (1) is typically used to model the kinematics when the system input u_l is measured in the body-fixed reference frame of the system. Similarly, the system model (2) is typically employed when the system input u_r is measured in the inertial frame of reference. In the cases where some of the system inputs are measured in the body-fixed frame while other inputs are measured in the inertial frame, the system model (3) might be employed.

Assume that either direct measurements of system inputs or an estimate of those inputs are available but the measurement of the outputs of the system encounter a delay of τ seconds. Such a situation occurs in many applications involving state estimation of mobile robots where inputs of the system are typically angular velocities and accelerations obtained almost delay-free using high rate sensors such as gyros and accelerometers while partial measurements of states (i.e system outputs which are position and orientation in this example) are obtained using magnetometers, cameras, GPS, etc. which are usually subject to considerable amounts of delay. In some practical scenarios direct measurements of inputs are not available but an (asymptotically stable) estimate of those inputs can be obtained using adaptive estimation techniques. The problem we discuss in this paper is to use the estimates of current inputs together with delayed measurements of outputs in order to estimate the current state of systems of the form (1), (2), or (3), assuming that the amount of the output delay is known.

Example 1. *The orientation of a vehicle wrt. a known frame of reference is called its attitude. Attitude estimation is a classical problem which is still a popular research topic [3–11, 13–15]. In mobile robotics applications, a 3-axis accelerometer typically employed as an inclinometer to measure the Earth's gravitational direction for use in attitude estimation algorithms. When a vehicle performs high acceleration maneuvers, accelerometers usually do not provide accurate measurements of the direction of the gravitational field. In this situation, it is common to enhance the attitude estimation algorithm by employing measurements of the linear velocity of the vehicle. Such algorithms are called velocity aided attitude estimators [5, 8, 42–45]. The motion of a flying rigid body in the Earth's gravitational field is then described by the following equations [8, 35, 42, 43].*

$$\dot{R}(t) = R(t)\Omega(t)_\times, \qquad R(0) = R_0 \tag{4}$$

$$\dot{v}(t) = ge_3 + R(t)a(t), \quad v(0) = v_0 \tag{5}$$

where the attitude matrix $R \in SO(3)$ represents the rotation of the body-fixed frame $\mathcal{B}$ wrt. the inertial frame $\mathcal{I}$, $\Omega \in \mathbb{R}^3$ is the angular velocity vector of $\mathcal{B}$ wrt. $\mathcal{I}$ expressed in $\mathcal{B}$, $v \in \mathbb{R}^3$ is the linear velocity of $\mathcal{B}$ wrt. $\mathcal{I}$ expressed in $\mathcal{I}$, $a \in \mathbb{R}^3$ is the so-called specific acceleration of the rigid body which represents the sum of all non-gravitational forces applied to the body divided by its mass and is expressed in $\mathcal{B}$, and ge_3 is the (constant) gravitational acceleration vector expressed in $\mathcal{I}$. The linear operator $(.)_\times$ maps any vector in $\mathbb{R}^3$ to its corresponding skew-symmetric matrix in $\mathfrak{so}(3)$ such that $(x)_\times y$ is equal to the cross product $x \times y$ for all $x, y \in \mathbb{R}^3$. The internal states of the dynamical system (4)-(5) are the attitude matrix $R(t)$ and the velocity vector $v(t)$ and its inputs are the angular velocity vector $\Omega(t)$ and the specific acceleration $a(t)$.

Similar to [21], we rewrite the dynamics (4)-(5) as a mixed-invariant system on the Lie group $SE(3)$ with inputs living on its Lie algebra $\mathfrak{se}(3)$. Defining $X = \begin{bmatrix} R & v \\ 0 & 1 \end{bmatrix} \in SE(3)$, $u_l = \begin{bmatrix} \Omega_\times & a \\ 0 & 0 \end{bmatrix} \in \mathfrak{se}(3)$, and $u_r = \begin{bmatrix} 0 & ge_3 \\ 0 & 0 \end{bmatrix} \in \mathfrak{se}(3)$, it is straight-forward to verify that $\dot{X}(t) = X(t)u_l(t) + u_r(t)X(t)$ where the group product is simply given by matrix multiplication. If the angular velocity $\Omega(t)$ and the acceleration $a(t)$ are directly measured (by a 3-axis gyro and accelerometers), then the inputs $u_l(t)$ and $u_r(t)$ would be available and can be used to estimate the state $X(t)$. However, in practice, measurements of angular velocity and linear acceleration are sometimes corrupted by unknown biases and scaling factors. It is common practice to assume that these biases and scaling factors are constant and that they are adaptively estimated with observers [4, 5, 11, 20, 33, 42, 46]. These estimates then can be used to obtain estimates of $\Omega(t)$ and $a(t)$.

In outdoor environments, a measurement of linear velocity is usually obtained using GPS units attached to the rigid body. Due to the internal processing time of GPS chips, the velocity measurement is usually delayed wrt. the actual linear velocity $v(t)$ of the vehicle. The amount of this delay can be up to hundreds of milliseconds and hence is not negligible. Assume that the GPS velocity measurement is given by $v_m(t) = v(t-\tau_v)$ where τ_v is a known constant delay. A magnetometer is another type of sensor usually employed for attitude estimation. A 3-axis magnetometer measures the magnetic field of the earth in the body-fixed frame. An ideal magnetometer output $m(t)$ is related to the attitude matrix via $b(t) = R(t)^\top \mathring{b}$ where $\mathring{b}$ is the vector of the Earth's magnetic field at the position of the rigid body expressed in $\mathcal{I}$. Assume that the magnetometer measurement is given by $b_m(t) = R(t-\tau_b)^\top \mathring{b}$ where τ_b is a known constant delay. In practice, the magnetometer delay is much shorter compared to the GPS delay. Defining $\tau := max(\tau_v, \tau_b)$ both outputs (i.e. GPS velocity and magnetometer measurements) at time $t-\tau$ are available at time t. The velocity aided attitude estimation problem is to use the estimates (or measurements) of current inputs $\Omega(t)$ and $a(t)$ together with the delayed GPS velocity and magnetometer measurements in order to estimate the current states $R(t)$ and $v(t)$. □

4. Observer-predictor approach

The approach that we take in this paper to tackle the problem defined in Section 3 is to employ an observer that uses the de-

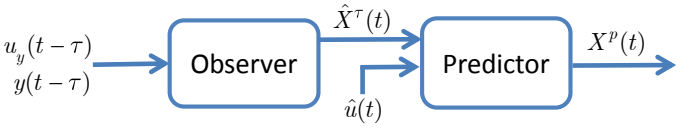

Figure 1: Proposed observer-predictor methodology.

layed output measurements and artificially delayed input measurements (denoted by $u_y(t-\tau)$) and provides an estimate of the delayed state $X(t-\tau)$, denoted by $\hat{X}^\tau(t)$. We then design a predictor that uses the delayed estimate $\hat{X}^\tau(t)$ together with measurements or estimates of the current input $u(t)$ (denoted by $\hat{u}(t)$) and provides "predictions" of the current state $X(t)$ denoted by $X^p(t)$. This observer-predictor scheme is illustrated in Fig. 1. In the special case where measurements of the true input $u(t)$ are available, one can replace $u_y(t-\tau)$ and $\hat{u}(t)$, respectively, with $u(t-\tau)$ and $u(t)$ in Fig 1. Otherwise, $\hat{u}(t)$ can be provided by an adaptive observer which estimates bias, scaling, and other unknown parameters of the input. This adaptive observer can sometimes be a part of the observer block of Fig. 1, in which case $\hat{u}(t)$ is indeed fed to the predictor from the observer block.

When outputs of the system are measured delay-free, various observer design methodologies are available in the literature that are capable of asymptotically/exponentially estimating the system state [4–9, 11–21, 42–45]. If a given observer is time-invariant[1], one can simply feed that observer with delayed outputs and delayed inputs of the system to obtain an estimate of the delayed state $X(t-\tau)$. Most observers employed in control design are formulated by a set of ODEs that does not explicitly depend on time. These observers are naturally time-invariant systems and hence the method described above is applicable to obtain an estimate of $X(t-\tau)$ for the systems (1)-(3). Hence, in this paper, we focus only on designing the predictor block of our proposed observer-predictor approach for the case of systems in the form (1), (2), or (3). Roughly speaking, the predictor's job is to use the information provided by the input signal in order to compute how much the state has changed in the period $t-\tau$ to t and use the result of this computation together with the estimate of $X(t-\tau)$ to provide a prediction of the current state $X(t)$.

We propose the following predictor for the left-invariant system (1)

$$\dot{\Delta}_l(t) = \Delta_l(t)\hat{u}_l(t), \qquad \Delta_l(0) = \Delta_{l_0}, \tag{6a}$$

$$X_l^p(t) = \hat{X}_l^\tau(t)\Delta_l(t-\tau)^{-1}\Delta_l(t), \quad t \geq \tau \tag{6b}$$

where $\Delta_l(t) \in G$ is the internal state of the predictor, $\hat{u}_l(t) \in \mathfrak{g}$ is an estimate of the input $u_l(t)$, and $X_l^p(t) \in$ is the prediction of $X_l(t)$. If direct measurements of $u_l(t)$ are available, those measurements can replace $\hat{u}_l(t)$. The predictor dynamics (6a), which is a copy of the system dynamics (1), generates the trajectory of $\Delta_l(t)$. This trajectory is stored in a buffer for the period $[t-\tau, t]$ in order to compute the prediction $X_l^p(t)$ using the static equation (6b). The term $\Delta_l(t-\tau)^{-1}\Delta_l(t)$ in (6b) is a prediction of the "increment" of the state $X(t)$ from the time $t-\tau$ to the time t. The predictor recursively computes this increment using the information provided by the input via (6a).

Similar to (6a)-(6b), we propose the following predictor for the system (2)

$$\dot{\Delta}_r(t) = \hat{u}_r(t)\Delta_r(t), \qquad \Delta_r(0) = \Delta_{r_0}, \tag{7a}$$

$$X_r^p(t) = \Delta_r(t)\Delta_r(t-\tau)^{-1}\hat{X}_r^\tau(t), \quad t \geq \tau \tag{7b}$$

where $\Delta_r(t) \in G$ and $X_r^p(t) \in G$ are the internal state of the predictor and the prediction of $X_r(t)$, respectively, and $\hat{u}_r(t) \in \mathfrak{g}$ is an estimate of $u_r(t)$. Again, we need to buffer $\Delta_r(t)$ for the period $[t-\tau, t]$. Note that this time, the increment of the state is predicted by the term $\Delta_r(t)\Delta_r(t-\tau)^{-1}$ such that it is compatible with the right-invariant nature of system (2).

The internal dynamics of the predictor that we propose for the mixed-invariant system (3) consists of both (6a) and (7a). The trajectories $\Delta_l(t)$ and $\Delta_r(t)$ are then buffered for $[t-\tau, t]$ and are used in the following equation to provide the prediction $X_m^p(t) \in G$ of $X_m(t)$.

$$X_m^p(t) = \Delta_r(t)\Delta_r(t-\tau)^{-1}\hat{X}_m^\tau(t)\Delta_l(t-\tau)^{-1}\Delta_l(t), \quad t \geq \tau \tag{8}$$

The term $\Delta_l(t-\tau)^{-1}\Delta_l(t)$ in (8) takes into account the increment of the state $X_m(t)$ (from $t-\tau$ to t) due to the input $u_l(t)$ while the term $\Delta_r(t)\Delta_r(t-\tau)^{-1}$ takes into account the increment due to the input $u_r(t)$. An interesting property of this predictor is that it separates the effect of $u_r(t)$ and $u_l(t)$ on the prediction $X_m^p(t)$. We will describe later why this separation is possible for the mixed-invariant system (3) and how it helps in predictor design.

Note that all of the above predictors can be implemented *recursively* together with *any* observer algorithm that provides the estimate $\hat{X}^\tau(t)$.

In the observer-predictor arrangement, the trajectory of the predicted state $X^p(t)$ inherently depends on the trajectory of the observer $X^\tau(t)$. Hence, rather than discussing the stability of the predictor, we need a notion of stability that relates the stability properties of the predictor to the stability properties of its corresponding observer. The following definition formalizes this notion.

Definition 1. *Consider a distance $d(.,.)$ on G. With respect to this distance, we say an observer-predictor pair is*

- *co-stable if $d(\hat{X}^\tau(t), X(t-\tau))$ bounded for all $t \geq \tau$ yields that $d(X^p(t), X(t))$ is bounded for all $t \geq \tau$.*
- *asymptotically co-stable if it is co-stable and $d(\hat{X}^\tau(t), X(t-\tau)) \to 0$ yields $d(X^p(t), X(t)) \to 0$.*
- *exponentially co-stable if it is co-stable and $d(\hat{X}^\tau(t), X(t-\tau)) \xrightarrow{exp} 0$ yields $d(X^p(t), X(t)) \xrightarrow{exp} 0$.*

We say a predictor is universally (asymptotically/exponentially) co-stable wrt. $d(.,.)$ if the combination of that predictor with any (asymptotically/exponentially) stable observer yields a (asymptotically/exponentially) co-stable observer-predictor pair wrt. $d(.,.)$. □

[1] Time-invariant is in the sense of considering the system inputs and outputs altogether as the input of the observer and considering the state estimate as the output of the observer.

The following Theorem summarizes the co-stability properties of the predictor (6) (resp. (7)) for the left-invariant system (1) (resp. right-invariant system (2)).

Theorem 1. *Consider system (1) with the predictor (6) (resp. system (2) with the predictor (7)). Assume that cond$(X_l(t))$ and $\|\hat{u}_l(t)-u_l(t)\|$ (resp. cond$(X_r(t))$ and $\|\hat{u}_r(t)-u_r(t)\|$) are bounded for all $t \geq 0$. For all $\tau \geq 0$ and all choices of $\Delta_{l_0} \in G$ (resp. $\Delta_{r_0} \in G$) we have;*

(a) The predictor is universally co-stable wrt. the distance $d_r(.,.)$ (resp. $d_l(.,.)$).

(b) If $\|\hat{u}_l(t) - u_l(t)\| \to 0$ (resp. $\|\hat{u}_r(t) - u_r(t)\| \to 0$) then the predictor is universally asymptotically co-stable wrt. $d_r(.,.)$ (resp. $d_l(.,.)$).

(c) If $\|\hat{u}_l(t) - u_l(t)\| \xrightarrow{exp} 0$ (resp. $\|\hat{u}_r(t) - u_r(t)\| \xrightarrow{exp} 0$), then the predictor is universally exponentially co-stable wrt. $d_r(.,.)$ (resp. $d_l(.,.)$).

□

Proof: We prove parts (a)-(c) of Theorem 1 for the left-invariant system (1) with the predictor (6a)-(6b). The proof for the right-invariant case can be obtained similarly.

Consider the observer error

$$E_l^\tau(t) = \hat{X}_l^\tau(t) X_l(t-\tau)^{-1} \tag{9}$$

and the total prediction error

$$E_l^p(t) = X_l^p(t) X_l(t)^{-1}. \tag{10}$$

Using (6b) and (10) we have $E_l^p(t) = \hat{X}_l^\tau(t)\Delta_l(t-\tau)^{-1}\Delta_l(t)X_l(t)^{-1}$ $= \hat{X}_l^\tau(t)X_l(t-\tau)^{-1}X_l(t-\tau)\Delta_l(t-\tau)^{-1}\Delta_l(t)X_l(t)^{-1}$. Defining $\Xi_l(t) := X_l(t)\Delta_l(t)^{-1}$ and using (9) we have

$$E_l^p(t) = E_l^\tau(t)\Xi_l(t-\tau)\Xi_l(t)^{-1} \tag{11}$$

It is easy to verify that the distance $d_r(.,.)$ is right-invariant such that $d_r(XZ, YZ) = d_r(X, Y)$ for all $X, Y, Z \in G$. Hence, using (11) and (9) we have

$$\begin{aligned} d_r(X_l^p(t), X_l(t)) &= d_r(E_l^p(t), I) = d_r(E_l^\tau(t)\Xi_l(t-\tau)\Xi_l(t)^{-1}, I) \\ &= d_r(E_l^\tau(t), \Xi_l(t)\Xi_l(t-\tau)^{-1}) \\ &\leq d_r(E_l^\tau(t), I) + d_r(\Xi_l(t)\Xi_l(t-\tau)^{-1}, I) \\ &= d_r(X_l^\tau(t), X_l(t-\tau)) + d_r(\Xi_l(t), \Xi_l(t-\tau)) \end{aligned} \tag{12}$$

The above inequality relates the scalar prediction error $d_r(X_l^p(t), X_l(t))$ to the scalar observer error $d_r(X_l^\tau(t), X_l(t-\tau))$ and the virtual error $d_r(\Xi_l(t), \Xi_l(t-\tau))$ which indirectly depends on the input estimation error $\|\hat{u}_l(t) - u_l(t)\|$. Using (1) and (6a), the dynamics of Ξ_l is given by $\dot{\Xi}_l(t) = \dot{X}_l(t)\Delta_l(t)^{-1} - X_l(t)\Delta_l(t)^{-1}\dot{\Delta}_l(t)\Delta_l(t)^{-1} = -[\mathrm{Ad}_{X_l(t)}(\hat{u}_l(t) - u_l(t))]\Xi_l(t)$. So,

$$\dot{\Xi}_l(t) = u_l^b(t)\Xi_l(t), \quad u_l^b(t) := -\mathrm{Ad}_{X_l(t)}(\hat{u}_l(t) - u_l(t)) \in \mathfrak{g} \tag{13}$$

Lemma 4 in the appendix summarizes the convergence properties of the error $d_r(\Xi_l(t), \Xi_l(t-\tau))$ based on eq (13). Also, according to Lemma 2, $\overline{\sigma}(\mathrm{Ad}_{X_l})$ is bounded if cond(X_l) is bounded. Hence, invoking both Lemma 2 and Lemma 4, the proof of parts (a)-(c) for the left invariant system directly follows from inequality (12).

For the right-invariant system (2) with the predictor (7), the proof of parts (a)-(c) can be obtained by direct adaptation of the proof of the left-invariant part using the following alternative error definitions.

$$E_r^\tau(t) = X_r(t-\tau)^{-1}\hat{X}_r^\tau(t), \tag{14}$$

$$E_r^p(t) = X_r(t)^{-1}X_r^p(t), \tag{15}$$

$$E_r^p(t) = \Xi_r(t)^{-1}\Xi_r(t-\tau)E_r^\tau(t), \tag{16}$$

where $\Xi_r(t) := \Delta_r(t)^{-1}X_r(t)$. ■

In order to discuss the co-stability properties of the predictor (6a), (7a), (8) for the mixed-invariant system (3), we rewrite this system as a combination of a right-invariant and a left-invariant system. Consider the systems

$$\dot{Z}_l(t) = Z_l(t)u_l(t), \quad Z_l(0) = Z_{l_0} \in G, \tag{17}$$

$$\dot{Z}_r(t) = u_r(t)Z_r(t), \quad Z_r(0) = Z_{r_0} \in G \tag{18}$$

where the initial conditions Z_{l_0} and Z_{r_0} are chosen such that $X_{m_0} = Z_{r_0}Z_{l_0}$. We have $\frac{d}{dt}(Z_r(t)Z_l(t)) = \dot{Z}_r(t)Z_l(t) + Z_r(t)\dot{Z}_l(t) = u_r(t)Z_r(t)Z_l(t) + Z_r(t)Z_l(t)u_l(t) = X_m(t)u_l(t) + u_r(t)X_m(t)$. This proves that $X_m(t) = Z_r(t)Z_l(t)$ for all $t \geq 0$. Using this decomposition, the following Theorem summarizes the co-stability properties of the proposed predictor for the mixed-invariant system.

Theorem 2. *Consider system (3) with the predictor (6a), (7a), (8). Assume that $\|\hat{u}_l(t) - u_l(t)\|$, $\|\hat{u}_r(t) - u_r(t)\|$, cond$(Z_r(t))$, cond$(Z_l(t))$, and cond$(\Delta_r(t)\Delta_r(t-\tau)^{-1})$ (resp. cond$(\Delta_l(t-\tau)^{-1}\Delta_l(t))$) are bounded for all $t \geq 0$. For all $\tau \geq 0$ and all choices of $\Delta_{l_0} \in G$ and $\Delta_{r_0} \in G$ we have;*

(a) The predictor is universally co-stable wrt. the distance $d_r(.,.)$ (resp. $d_l(.,.)$).

(b) If $\|\hat{u}_l(t) - u_l(t)\| \to 0$ and $\|\hat{u}_r(t) - u_r(t)\| \to 0$ then the predictor is universally asymptotically co-stable wrt. $d_r(.,.)$ (resp. $d_l(.,.)$).

(c) If $\|\hat{u}_l(t) - u_l(t)\| \xrightarrow{exp} 0$ and $\|\hat{u}_r(t) - u_r(t)\| \xrightarrow{exp} 0$, then the predictor is universally exponentially co-stable wrt. $d_r(.,.)$ (resp. $d_l(.,.)$). □

Proof: We prove the Theorem for the right-invariant distance $d_r(.,.)$. The proof for the left-invariant distance follows similarly. Define the following observer and prediction errors for the mixed-invariant system (3).

$$E_m^\tau(t) = Z_r(t-\tau)^{-1}\hat{X}_m^\tau(t)Z_l(t-\tau)^{-1}, \tag{19}$$

$$E_m^p(t) = Z_r(t)^{-1}X_m^p(t)Z_l(t)^{-1}. \tag{20}$$

Using (8) and (20), we have $E_m^p(t) = Z_r(t)^{-1}\Delta_r(t)\Delta_r(t-\tau)^{-1}(Z_r(t-\tau)Z_r(t-\tau)^{-1})\hat{X}_m^\tau(t)(Z_l(t-\tau)^{-1}Z_l(t-\tau))\Delta_l(t-\tau)^{-1}\Delta_l(t)Z_l(t)^{-1}$. Defining $\Xi_l^z(t) := Z_l(t)\Delta_l(t)^{-1}$ and $\Xi_r^z(t) := Z_r(t)^{-1}\Delta_r(t)$ and using (19) yields

$$E_m^p(t) = \Xi_r^z(t)\Xi_r^z(t-\tau)^{-1}E_m^\tau(t)\Xi_l^z(t-\tau)\Xi_l^z(t)^{-1} \tag{21}$$

Using (19), (20), (21), and Lemma 3 in the appendix and dropping the argument t, we have

$$
\begin{aligned}
&d_r(X_m^p, X_m) = d_r(X_m^p, Z_r Z_l)\\
&\le \overline{\sigma}(\mathrm{Ad}_{Z_r}) d_r(Z_r^{-1} X_m^p, Z_r^{-1}(Z_r Z_l))\\
&= \overline{\sigma}(\mathrm{Ad}_{Z_r}) d_r(Z_r^{-1} X_m^p Z_l^{-1}, I) = \overline{\sigma}(\mathrm{Ad}_{Z_r}) d_r(E_m^p, I)\\
&= \overline{\sigma}(\mathrm{Ad}_{Z_r}) d_r(\Xi_r^z \Xi_r^z(t-\tau)^{-1} E_m^\tau \Xi_l^z(t-\tau) \Xi_l^{z-1}, I)\\
&= \overline{\sigma}(\mathrm{Ad}_{Z_r}) d_r(\Xi_r^z \Xi_r^z(t-\tau)^{-1} E_m^\tau, \Xi_l^z \Xi_l^z(t-\tau)^{-1})\\
&\le \overline{\sigma}(\mathrm{Ad}_{Z_r}) \Big(d_r(\Xi_r^z \Xi_r^z(t-\tau)^{-1} E_m^\tau, I) + d_r(\Xi_l^z \Xi_l^z(t-\tau)^{-1}, I) \Big)\\
&= \overline{\sigma}(\mathrm{Ad}_{Z_r}) \Big(d_r(\Xi_r^z \Xi_r^z(t-\tau)^{-1} E_m^\tau \Xi_r^z(t-\tau) \Xi_r^{z-1}, \Xi_r^z(t-\tau) \Xi_r^{z-1})\\
&\qquad + d_r(\Xi_l^z \Xi_l^z(t-\tau)^{-1}, I) \Big)\\
&\le \overline{\sigma}(\mathrm{Ad}_{Z_r}) \cdot\\
&\Big(d_r((\Delta_r(t-\tau)\Delta_r^{-1} Z_r)^{-1} \hat{X}_m^\tau (Z_r(t-\tau) Z_l(t-\tau))^{-1} (\Delta_r(t-\tau)\Delta_r^{-1} Z_r), I)\\
&\qquad + d_r(\Xi_r^z(t-\tau) \Xi_r^{z-1}, I) + d_r(\Xi_l^z, \Xi_l^z(t-\tau)) \Big)\\
&= \overline{\sigma}(\mathrm{Ad}_{Z_r}) \cdot\\
&\Big(\overline{\sigma}(\mathrm{Ad}_{Z_r^{-1}\Delta_r \Delta_r(t-\tau)^{-1}}) d_r(\hat{X}_m^\tau X_m(t-\tau)^{-1} \Delta_r(t-\tau)\Delta_r^{-1} Z_r, \Delta_r(t-\tau)\Delta_r^{-1} Z_r)\\
&\qquad + d_r(\Xi_r^z(t-\tau) \Xi_r^{z-1}, I) + d_r(\Xi_l^z, \Xi_l^z(t-\tau)) \Big)\\
&\le \overline{\sigma}(\mathrm{Ad}_{Z_r}) \Big(\overline{\sigma}(\mathrm{Ad}_{Z_r^{-1}\Delta_r \Delta_r(t-\tau)^{-1}}) d_r(\hat{X}_m^\tau, X_m(t-\tau))\\
&\qquad + d_r(\Xi_r^z(t-\tau), \Xi_r^z) + d_r(\Xi_l^z, \Xi_l^z(t-\tau)) \Big) \qquad (22)
\end{aligned}
$$

Similar to (13), one can verify that

$$\dot{\Xi}_l^z(t) = u_l^z(t) \Xi_l^z(t), \quad u_l^z(t) := -\mathrm{Ad}_{Z_l(t)}(\hat{u}_l(t) - u_l(t)) \qquad (23a)$$

$$\dot{\Xi}_r^z(t) = u_r^z(t) \Xi_r^z(t), \quad u_r^z(t) := \mathrm{Ad}_{Z_r(t)^{-1}}(\hat{u}_r(t) - u_r(t)). \qquad (23b)$$

Lemma 4 summarizes the stability properties of the scalar errors $d_r(\Xi_r^z(t-\tau), \Xi_r^z(t))$ and $d_r(\Xi_l^z(t), \Xi_l^z(t-\tau))$ based on eq. (23). Also, using Lemma 2 in the Appendix we have that if $\mathrm{cond}(Z_r(t))$, $\mathrm{cond}(Z_l(t))$, and $\mathrm{cond}(\Delta_r \Delta_r(t-\tau)^{-1})$ are bounded, then $\overline{\sigma}(\mathrm{Ad}_{Z_r(t)^{-1}})$, $\overline{\sigma}(\mathrm{Ad}_{Z_l(t)})$, and $\overline{\sigma}(\mathrm{Ad}_{Z_r^{-1}(t)\Delta_r(t)\Delta_r(t-\tau)^{-1}})$ are bounded. Hence, invoking both Lemma 2 and Lemma 4, proof of parts (a)-(c) for the right-invariant distance $d_r(.,.)$ follows directly from inequality (22). ■

In the following, we consider a particular case where the measurements of the actual inputs $u_l(t)$ and $u_r(t)$ (instead of their estimates $\hat{u}_l(t)$ and $\hat{u}_r(t)$) are used in the predictors. The following Theorem shows that in this case the total prediction error of the current state exactly equals the observer error of estimating the delayed state. This shows that the proposed predictors do not change the estimation error behavior of the corresponding observer that feeds the predictor, even though the predictor fully compensates for the effect of sensor delays.

Theorem 3. *Consider system (1) with the predictor (6), or system (2) with the predictor (7), or system (3) with the predictor (6a), (7a), (8). Assume that $\hat{u}_l(t) = u_l(t)$ and $\hat{u}_r(t) = u_r(t)$ for all $t \ge 0$. Consider the observer errors (9), (14) and (19) and their corresponding total prediction errors (10), (15), and (20), respectively. We have $E_l^\tau(t) = E_l^p(t)$, $E_r^\tau(t) = E_r^p(t)$, and $E_m^\tau(t) = E_m^p(t)$ for all $t \ge \tau$, all $\tau \ge 0$, and all choices of $\Delta_{l_0}, \Delta_{r_0} \in G$.* □

Proof: The proof follows from the derivations made in the proof of Theorem 1. According to (13), if $\hat{u}_l(t) = u_l(t)$ then $\dot{\Xi}_l(t) = 0$ and hence $\Xi_l(t)$ is constant which in particular implies $\Xi_l(t) = \Xi_l(t-\tau)$ for all $t \ge \tau$. This together with (11) yields $E_l^p(t) = E_l^\tau(t)$ for all $t \ge \tau$. A similar argument for the right-invariant case shows that $\dot{\Xi}_r(t) = 0$ if $\hat{u}_r(t) = u_r(t)$ and proves that $E_r^p(t) = E_r^\tau(t)$ for all $t \ge \tau$. Also, the same argument is applicable to (23) and (21) to conclude $E_m^p(t) = E_m^\tau(t)$ if both $\hat{u}_l(t) = u_l(t)$ and $\hat{u}_r(t) = u_r(t)$. ■

The internal dynamics of the predictors (6a) and (7a) are simple forward integration. One possible drawback of using these pure forward integrators is that their internal states $\Delta_l(t)$ and $\Delta_r(t)$ can become larger and larger (due to the input measurement noise or numerical integration inaccuracies) as we continue the integration procedure. One possible way to overcome the above issue is to periodically reset the initial condition of dynamics (6a) and (7a) to a constant value so that the trajectories of $\Delta_l(t)$ $\Delta_r(t)$ remain bounded for all times. An arbitrary resetting of the initial conditions of (6a) and (7a) can potentially destroy the co-stability properties of the predictors. However, owing to the invariance of the systems (6a) and (7a), the following Lemma proposes a resetting methodology that keeps the trajectories of the internal states of the predictors bounded while maintaining their co-stability properties.

Lemma 1. *Consider the system*

$$\dot{\Delta}_l^s(t) = \Delta_l^s(t) \hat{u}_l(t), \qquad (24)$$

and assume that the trajectory $\Delta_l^s(t)$ is stored in a buffer for the period $[t-\tau, t]$ (i.e. the last τ seconds). Assume that the initial condition of the system (24) resets to a fixed value Δ_{l_0} every T_s seconds where $T_s \ge \tau$ (i.e. $\Delta_l^s(nT_s) = \Delta_{l_0}$ for $n = 0, 1, \ldots$). For $t \in [nT_s, (n+1)T_s)$, compute the following variable that depends only on the buffered values of Δ_l^s.

$$\Delta_l^b(t) := \begin{cases} \Delta_{l_0} \Delta_l^s(nT_s^-)^{-1} \Delta_l^s(t-\tau), & \text{for all } t \in [nT_s, nT_s + \tau), \\ \Delta_l^s(t-\tau), & \text{for all } t \in [nT_s + \tau, (n+1)T_s) \end{cases} \qquad (25)$$

Then, we have $\Delta_l^b(t)^{-1} \Delta_l^s(t) = \Delta_l(t-\tau)^{-1} \Delta_l(t)$ for all $t \ge \tau$ (i.e. we can replace $\Delta_l(t-\tau)^{-1} \Delta_l(t)$ with the bounded signal $\Delta_l^b(t)^{-1} \Delta_l^s(t)$ in (6b)). Similarly, consider the system

$$\dot{\Delta}_r^s(t) = \hat{u}_r(t) \Delta_r^s(t), \qquad (26)$$

and buffer the trajectory $\Delta_r^s(t)$ for $[t-\tau, t]$. Reset the initial condition of the system (26) to a fixed value Δ_{r_0} every $T_s \ge \tau$ seconds. For $t \in [nT_s, (n+1)T_s)$, compute the following variable.

$$\Delta_r^b(t) := \begin{cases} \Delta_r^s(t-\tau) \Delta_r^s(nT_s^-)^{-1} \Delta_{r_0}, & \text{for all } t \in [nT_s, nT_s + \tau), \\ \Delta_r^s(t-\tau), & \text{for all } t \in [nT_s + \tau, (n+1)T_s) \end{cases} \qquad (27)$$

Then, we have $\Delta_r^s(t) \Delta_r^b(t)^{-1} = \Delta_r(t) \Delta_r(t-\tau)^{-1}$ for all $t \ge \tau$ (i.e. we can replace $\Delta_r(t) \Delta_r(t-\tau)^{-1}$ with the bounded signal $\Delta_r^s(t) \Delta_r^b(t)^{-1}$ in (7b)). □

Proof: We prove the Lemma for the left-invariant case. The right-invariant case can be proved similarly. Consider the systems (6a) and (24), both with the same initial condition Δ_{l_0}. We have $\Delta_l^s(t) = \Delta_l(t)$ for all $t \in [0, T_s)$ and clearly $\Delta_l^b(t)^{-1}\Delta_l^s(t) = \Delta_l^s(t{-}\tau)^{-1}\Delta_l^s(t) = \Delta_l(t{-}\tau)^{-1}\Delta_l(t)$ holds for all $t \in [\tau, T_s)$. In the following, we show that the statement of the Lemma also holds for all $t \in [nT_s, (n+1)T_s)$, $n = 1, 2, \ldots$. Using (6a) and (24), we have $\frac{d}{dt}(\Delta_l^s(t)\Delta_l(t)^{-1}) = \Delta_l^s(t)\hat{u}_l(t)\Delta_l(t)^{-1} - \Delta_l^s(t)\hat{u}_l(t)\Delta_l(t)^{-1} = 0$ for all $t \in [nT_s, (n{+}1)T_s)$. Hence, $\Delta_l^s(t)\Delta_l(t)^{-1}$ is constant for all $t \in [nT_s, (n+1)T_s)$ and we have

$$\Delta_l^s(t_1)\Delta_l(t_1)^{-1} = \Delta_l^s(t_2)\Delta_l(t_2)^{-1} \quad \text{for all} \quad t_1, t_2 \in [nT_s, (n{+}1)T_s). \tag{28}$$

Choosing $t_1 = t \in [nT_s, (n{+}1)T_s)$ and $t_2 = nT_s$ we obtain

$$\Delta_l^s(t) = \Delta_l^s(nT_s)\Delta_l(nT_s)^{-1}\Delta_l(t) = \Delta_{l_0}\Delta_l(nT_s^-)^{-1}\Delta_l(t) \tag{29}$$

where we used the continuity of $\Delta_l(t)$ at $t = nT_s$ to replace $\Delta_l(nT_s)$ by $\Delta_l(nT_s^-)$. One can use the same method as was done to derive (28) to conclude $\Delta_l^s(t_1)\Delta_l(t_1)^{-1} = \Delta_l^s(t_2)\Delta_l(t_2)^{-1}$ for all $t_1, t_2 \in [(n{-}1)T_s, nT_s)$. Choosing $t_1 = t - \tau \in [nT_s{-}\tau, nT_s) \subset [(n{-}1)T_s, nT_s)$ and $t_2 = nT_s^-$ we obtain

$$\Delta_l^s(t-\tau)^{-1}\Delta_l^s(nT_s^-) = \Delta_l(t{-}\tau)^{-1}\Delta_l(nT_s^-). \tag{30}$$

Using (25) and (29), for all $t \in [nT_s, nT_s + \tau)$ we have

$$\Lambda(t)^{-1}\Delta_l^s(t) = \Delta_l^s(t{-}\tau)^{-1}\Delta_l^s(nT_s^-)\Delta_l(nT_s^-)^{-1}\Delta_l(t) \tag{31}$$

Substituting (30) into (31) yields $\Delta_l^b(t)^{-1}\Delta_l^s(t) = \Delta_l(t{-}\tau)^{-1}\Delta_l(t)$ for all $t \in [nT_s, nT_s + \tau)$. For $t \in [nT_s + \tau, (n+1)T_s)$, we chose $t_1 = t{-}\tau \in [nT_s, (n{+}1)T_s{-}\tau)$ and $t_2 = t$ and recall (28) to obtain $\Delta_l^s(t{-}\tau)^{-1}\Delta_l^s(t) = \Delta_l(t{-}\tau)^{-1}\Delta_l(t)$. Using (25) for $t \in [nT_s{+}\tau, (n{+}1)T_s)$ we have $\Delta_l^b(t)^{-1}\Delta_l^s(t) = \Delta_l^s(t{-}\tau)^{-1}\Delta_l^s(t) = \Delta_l(t{-}\tau)^{-1}\Delta_l(t)$. Consequently, the Lemma holds for all $t \in [nT_s, (n{+}1)T_s)$. This completes the proof. ■

Lemma 1 proposes a resetting technique that keeps the trajectories of the internal states of the predictor bounded while it does not affect the trajectory of the prediction $X^p(t)$ and hence preserves the co-stability properties of those predictors. An alternative resetting technique has been proposed in [35, Lemma 1] to bound the internal states of predictors while maintaining their stability. That method involves employing two copies of the internal dynamics of the predictor and appropriately resetting the initial conditions of those copies and also switching between the trajectories of those copies such that the final predictions (6a) and (7a) do not change. The method proposed here requires less computational power compared to [35, Lemma 1] since it employs only one copy of the predictor's internal dynamics.

Remark 1. *Theorem 2 assumes that cond($Z_r(t)$), cond($Z_l(t)$), cond($\Delta_r(t)\Delta_r(t-\tau)^{-1}$), and cond($\Delta_l(t-\tau)^{-1}\Delta_l(t)$) are bounded for all $t \geq 0$. If G is a compact group, then these conditions are automatically satisfied. Nevertheless, for general Lie groups these conditions may not hold in general. One way to address this issue is to resort to Lemma 1 and periodically switch the initial conditions of (32), (34), (17) and (18) such that the trajectories of $\Delta_l(t)$, $\Delta_r(t)$, $Z_l(t)$, and $Z_r(t)$ remain bounded for all $t \geq 0$. It is possible to show that switching can be done in a way that does not change the trajectory of $X_m^p(t)$ while the trajectories of the switched $Z_l(t)$ and $Z_r(t)$ still satisfy $X_m(t) = Z_r(t)Z_l(t)$ for all $t \geq 0$, and $\Xi_l^z(t)$ and $\Xi_r^z(t)$ remain continuous. That is to say, with the resetting scheme of Lemma 1, we can replace the boundedness conditions of cond($Z_r(t)$), cond($Z_l(t)$), cond($\Delta_r(t)\Delta_r(t{-}\tau)^{-1}$), and cond($\Delta_l(t{-}\tau)^{-1}\Delta_l(t)$) with cond($X_m(t)$) being bounded and the results of Theorem 2 still remain valid for general Lie groups.* □

Example 2. *As was shown in Example 1, the velocity-aided attitude estimation problem with GPS delay can be formulated as a special case of predictor design for a mixed-invariant system on the Lie group $SE(3)$. Hence, one can directly employ the predictor (6a), (7a), (8). Nevertheless, here we show that the specific structure of dynamics (4)-(5) allows reducing the dimension of the mixed-invariant predictor. We assume that an estimate of $X(t{-}\tau) \in SE(3)$ is available and we decompose it into $\hat{X}^\tau(t) = \begin{bmatrix} \hat{R}^\tau(t) & \hat{v}^\tau(t) \\ 0 & 1 \end{bmatrix}$ where $\hat{R}^\tau \in SO(3)$ and $\hat{v}^\tau \in \mathbb{R}^3$. Decompose Δ_l and Δ_r into $\Delta_l = \begin{bmatrix} \bar{\Delta}_l & \delta_l \\ 0 & 1 \end{bmatrix}$ and $\Delta_r = \begin{bmatrix} \bar{\Delta}_r & \delta_r \\ 0 & 1 \end{bmatrix}$ where $\Delta_l, \Delta_r \in SO(3)$ and $\delta_l, \delta_r \in \mathbb{R}^3$. We need estimates of the inputs $u_l(t)$ and $u_r(t)$ to formulate the predictor. As was shown in Example 1, $u_r(t)$ is a known constant in this scenario. So, we simply choose $\hat{u}_r(t) = u_r = \begin{bmatrix} 0 & ge_3 \\ 0 & 0 \end{bmatrix} \in \mathfrak{se}(3)$. We assume that estimates of $\hat{a}(t)$ and $\hat{\Omega}(t)$ of the current specific acceleration and angular velocity are available and we choose $\hat{u}_l(t) = \begin{bmatrix} \hat{\Omega}(t)_\times & \hat{a}(t) \\ 0 & 0 \end{bmatrix} \in \mathfrak{se}(3)$. If accelerometers and gyros directly measure $a(t)$ and $\Omega(t)$, we use those direct measurements as $\hat{a}(t)$ and $\hat{\Omega}(t)$. Otherwise, if the accelerometers and gyros have constant unknown biases and scaling factors, the observer part of Fig 1 estimates those values and provides estimates of $\hat{a}(t)$ and $\hat{\Omega}(t)$. Substituting for Δ_r and u_r in (6a) we have*

$$\begin{aligned} \begin{bmatrix} \dot{\bar{\Delta}}_l(t) & \dot{\delta}_l(t) \\ 0 & 0 \end{bmatrix} &= \begin{bmatrix} \bar{\Delta}_l(t) & \delta_l(t) \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \hat{\Omega}(t)_\times & \hat{a}(t) \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \bar{\Delta}_l(t)\hat{\Omega}(t)_\times & \bar{\Delta}_l(t)\hat{a}(t) \\ 0 & 0 \end{bmatrix}. \end{aligned} \tag{32}$$

Also, we have

$$\begin{aligned} \Delta_l(t{-}\tau)^{-1}\Delta_l(t) &= \begin{bmatrix} \bar{\Delta}_l(t)^\top & -\bar{\Delta}_l(t)^\top\delta_l(t) \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \bar{\Delta}_l(t) & \delta_l(t) \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \bar{\Delta}_l(t{-}\tau)^\top\bar{\Delta}_l(t) & \Delta_l(t{-}\tau)^\top(\delta_l(t) - \delta_l(t{-}\tau)) \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{33}$$

Similarly, substituting for Δ_r and u_r in (7a) we have

$$\begin{bmatrix} \dot{\bar{\Delta}}_r(t) & \dot{\delta}_r(t) \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & ge_3 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \bar{\Delta}_r(t) & \delta_r(t) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & ge_3 \\ 0 & 0 \end{bmatrix}. \tag{34}$$

This yields $\dot{\bar{\Delta}}_r(t) = 0$ *and* $\dot{\delta}_r(t) = ge_3$*, implying* $\bar{\Delta}_r(t-\tau) = \bar{\Delta}_r(t)$ *and* $\delta_r(t) - \delta_r(t-\tau) = \tau ge_3$*. These together yield*

$$\Delta_r(t)\Delta_r(t-\tau)^{-1} = \begin{bmatrix} \bar{\Delta}_r(t) & \delta_r(t) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{\Delta}_r(t-\tau)^\top & -\bar{\Delta}_r(t-\tau)^\top \delta_r(t) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} I & \tau g e_3 \\ 0 & 1 \end{bmatrix}. \tag{35}$$

Since the right hand side of (35) is a known constant, we do not actually need to implement the right-invariant part of the predictor (i.e. the dynamics (34)). We can use the known value given by the right hand side of (35) to obtain the prediction (8). This reduces the dimension of the predictor to the same dimension as the actual system. Decomposing X^p *into* $X^p = \begin{bmatrix} R^p & v^p \\ 0 & 1 \end{bmatrix}$ *and replacing (33) and (35) into (8) and dropping the argument t, we have*

$$\begin{aligned} &\begin{bmatrix} R^p & v^p \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} I & \tau g e_3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{R}^\tau & \hat{v}^\tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \bar{\Delta}_l(t-\tau)^\top \bar{\Delta}_l & \Delta_l(t-\tau)^\top(\delta_l - \delta_l(t-\tau)) \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \hat{R}^\tau \bar{\Delta}_l(t-\tau)^\top \bar{\Delta}_l & \hat{v}^\tau + \tau g e_3 + \hat{R}^\tau \Delta_l(t-\tau)^\top(\delta_l - \delta_l(t-\tau)) \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{36}$$

Finally, using (32) and (36), we simplify the predictor (6a), (7a), and (8) to

$$\dot{\bar{\Delta}}_l(t) = \bar{\Delta}_l(t)\hat{\Omega}(t)_\times, \quad \bar{\Delta}_l(0) = \bar{\Delta}_{l_0} \in SO(3), \quad t \geq 0 \tag{37}$$

$$\dot{\delta}_l(t) = \bar{\Delta}_l(t)\hat{a}(t), \quad \delta_l(0) = \delta_{l_0} \in \mathbb{R}^3, \quad t \geq 0 \tag{38}$$

$$R^p(t) = \hat{R}^\tau(t)\bar{\Delta}_l(t-\tau)^\top \bar{\Delta}_l(t), \quad t \geq \tau \tag{39}$$

$$v^p(t) = \hat{v}^\tau(t) + \tau g e_3 + \hat{R}^\tau(t)\bar{\Delta}_l(t-\tau)^\top(\delta_l(t) - \delta_l(t-\tau)), \quad t \geq \tau. \tag{40}$$

Since $SO(3)$ *is a compact manifold, there is no concern regarding the boundedness of the internal state* $\bar{\Delta}_l(t)$*. However, the state* $\delta_l(t)$ *might grow larger and larger as it lives in* $\mathbb{R}^3$*. Hence, we need to employ the resetting technique proposed in Lemma 1 to bound the trajectory of* $\delta_l(t)$*. If we periodically reset the initial condition of (38) to* δ_{l_0} *every* T_s *seconds (and do not reset (37)), (25) simplifies to adding* $\delta_{l_0} - \delta_l(nT_s^-)$ *to the term* $\delta_l(t-\tau)$ *in (40) for all* $t \in [nT_s, nT_s + \tau)$*,* $n = 1, 2, \ldots$*. This resetting scheme ensures the co-stability of the predictor as long as* $v(t)$ *is bounded (see Remark 1). The attitude predictor (37) and (39) has exactly the same form as the attitude predictor proposed by the authors in [35, Eq. (6)-(7)], however, the velocity predictor (38) and (40) has an essentially different form compared to [35, Equ. (8)-(9)]. It is easy to verify that the predictor proposed in [35] is an example of predictor (6a)-(6b) when the underlying Lie group is* $SO(3) \times \mathbb{R}^3$ *and the group multiplication is simply given by* $(R_1, v_1)(R_2, v_2) = (R_1R_2, v_1 + v_2)$*.* □

5. Recursive implementation of non-recursive predictors

The predictor design methodologies presented in the previous sections only suit invariant systems on Lie groups. For a general system of the form

$$\dot{X}(t) = F(X(t), u(t)), \tag{41}$$

where $X \in G$ and u belongs to an input manifold $\mathcal{M}_u$ and $F : G \times \mathcal{M}_u \to TG$, $F(X, u) \in T_X G$, a simple predictor with desirable co-stability properties was presented in [41]. At time t, this predictor uses the estimate $\hat{X}^\tau(t)$ as the initial condition of a copy of the system dynamics at time $t-\tau$. Then it forward integrates the system dynamics from $t-\tau$ to t using the information of the input $u(t)$ to obtain a prediction of $X(t)$. This procedure is mathematically formulated as follows

$$\frac{d}{ds}X_t^p(s) = F(X_t^p(s), u(s)), \qquad s \in [t-\tau, t], \tag{42a}$$

$$X_t^p(t-\tau) = \hat{X}^\tau(t), \tag{42b}$$

$$X^p(t) = X_t^p(t). \tag{42c}$$

where X_t^p is the internal state of the predictor and X^p is the prediction of $X(t)$. The subscript t of X_t^p emphasizes that the predictor (42) is *non-recursive* in the sense that at each time t, the forward integration of dynamics (42a) from $s = t-\tau$ to $s = t$ should be completely performed in order to compute the final value $X_t^p(t)$ and obtain the prediction $X^p(t)$. This non-recursive nature of predictor (42) is in fact its major drawback which prevents its application in practical scenarios involving real-time implementations in on-board computers of robots. Despite this drawback, universally asymptotic and exponential co-stability of this predictor has been shown in [41] for systems on $\mathbb{R}^n$ and this result can be generalized for systems whose states evolve on differentiable manifolds, including Lie groups.

There is an interesting link between the non-recursive predictor (42) and the recursive predictors we proposed in Section 4. When the underlying system (41) is of the form (1), (2), or (3), the following lemma shows that the resulting predictor (42) produces exactly the same prediction trajectory as is produced by the recursive predictors we proposed in this paper. In other words, for the case where the vector field $F(X, u)$ is left, right, or mixed-invariant, the non-recursive predictor (42) can be *recursively* realized using the method we presented in this paper.

Proposition 1. *Consider the predictor (42) for the system (41).*

(a) (left resp. right-invariant case) Assume that the system (41) is left-invariant (resp. right-invariant) in the sense that $F(X, u) = XF_l^{\mathfrak{g}}(u)$ *(resp.* $F(X, u) = F_r^{\mathfrak{g}}(u)X$*) where* $F_l^{\mathfrak{g}} : \mathcal{M}_u \to \mathfrak{g}$ *(resp.* $F_r^{\mathfrak{g}} : \mathcal{M}_u \to \mathfrak{g}$*) is a known map. Define* $u_l := F_l^{\mathfrak{g}}(u)$ *(resp.* $u_r := F_r^{\mathfrak{g}}(u)$*), rewrite the system dynamics as* $\dot{X}(t) = X(t)u_l(t)$ *(resp.* $\dot{X}(t) = u_r(t)X(t)$*), and obtain the corresponding predictor (6) (resp. (7)) for this system assuming that* $\hat{u}_l = u_l$ *(resp.* $\hat{u}_r = u_r$*). If both predictors (42) and (6) (resp. (7)) are fed with the same observer trajectories i.e.* $\hat{X}^\tau(t) = \hat{X}_l^\tau(t)$ *(resp.* $\hat{X}^\tau(t) = \hat{X}_r^\tau(t)$*), then the prediction* $X^p(t)$ *of (42c) equals* $X_l^p(t)$ *of (6b) (resp.* $X_r^p(t)$ *of (7b)) for all* $t \geq \tau$*.*

(b) (mixed-invariant case) Assume that the system (41) is mixed-invariant in the sense that $F(X, u) = XF_l^{\mathfrak{g}}(u) + F_r^{\mathfrak{g}}(u)X$ *where* $F_l^{\mathfrak{g}}, F_r^{\mathfrak{g}} : \mathcal{M}_u \to \mathfrak{g}$ *are known maps. Define*

$u_l := F_l^{\mathfrak{g}}(u)$ and $u_r := F_l^{\mathfrak{g}}(u)$, rewrite the system dynamics as $\dot{X}(t) = X(t)u_l(t) + u_r(t)X(t)$, and obtain the corresponding predictor (6a), (7a), and (8) for this system (assuming $\hat{u}_l = u_l$ and $\hat{u}_r = u_r$). If both predictors are fed with the same observer trajectories (i.e. $\hat{X}^\tau(t) = \hat{X}^\tau_m(t)$), then the prediction $X^p(t)$ of (42c) equals $X^p_r(t)$ of (8) for all $t \geq \tau$.

□

Proof: We prove part (a) for the left-invariant system. Proof for the right-invariant case can be obtained similarly. We suggest that the solution $X^p_t(s)$, $s \geq t-\tau$ of (42a)-(42b) is given by $X^p_t(s) = \hat{X}^\tau(t)\Delta_l(t-\tau)^{-1}\Delta_l(s)$ where $\frac{d}{ds}\Delta_l(s) = \Delta_l(s)u_l(s)$. To show this, we note that this solution satisfies the initial condition requirement $X^p_t(s)|_{s=t-\tau} = \hat{X}^\tau(t)\Delta_l(t-\tau)^{-1}\Delta_l(t-\tau) = \hat{X}^\tau(t)$ and the derivative requirement $\frac{d}{ds}X^p_t(s) = \hat{X}^\tau(t)\Delta_l(t-\tau)^{-1}\frac{d}{ds}\Delta_l(s) = \hat{X}^\tau(t)\Delta_l(t-\tau)^{-1}\Delta_l(s)u_l(s) = X^p_t(s)u_l(s) = F(X^p_t(s), u_l(s))$. By (42c) we have $X^p(t) = X^p_t(s)|_{s=t} = \hat{X}^\tau(t)\Delta_l(t-\tau)^{-1}\Delta_l(t)$ which is equal to (6b). Proof of part (b) is obtained similarly by verifying that the solution of (42a)-(42b) is given by $X^p_t(s) = \Delta_r(s)\Delta_r(t-\tau)^{-1}\hat{X}^\tau(t)\Delta_l(t-\tau)^{-1}\Delta_l(s)$ with $\frac{d}{ds}\Delta_l(s) = \Delta_l(s)u_l(s)$ and $\frac{d}{ds}\Delta_r(s) = u_r(s)\Delta_r(s)$ and by evaluating this solution at $s = t$. ■

The results of [41] suggest that recursive implementability of the predictor (42) is essentially (up to a change of variable) restricted to the systems of the form (1), (2), or (3).

6. Simulations

A predictor for the velocity-aided attitude estimation problem with GPS delay is proposed by (37)-(40) in Example 2. In this section, we provide extensive simulation studies to evaluate the performance of the predictor (37)-(40) in practical situations where there are inaccuracies such as sensor noises, biases, sampling, etc. According to Fig. 1, the total observer-predictor estimation error (i.e. the difference between $X^p(t)$ and $X(t)$) is a combination of the pure error due to the observer (i.e. the difference between $\hat{X}^\tau(t)$ and $X(t-\tau)$) and the pure error due to the predictor (i.e. the difference between $X^p(t)$ and $X(t)$ if the predictor is fed with $X(t-\tau)$ instead of $\hat{X}^\tau(t)$). We first provide a set of simulations to evaluate the pure predictor error. Then we present realistic simulations using a sophisticated Software-In-The-Loop (SITL) system to demonstrate the total observer-predictor error.

6.1. Pure prediction error:

Using MATLAB®, we feed the predictor (37)-(40) with the true $R(t-\tau)$ and $v(t-\tau)$ instead of $\hat{R}^\tau(t)$ and $\hat{v}^\tau(t)$, respectively. We add Gaussian noises with the high standard deviation of 2.5 (deg/s) and 1.5 (m/s^2) to each axis of the gyro and accelerometer, respectively. We consider various amounts of delay ranging from a small delay of 0.2 (s) to a large delay of 1 (s). We consider a plane that is flying in a circular trajectory with a linear velocity of 22.5 (m/s). In order to demonstrate the effects of both low acceleration and high acceleration maneuvers, we perform separate simulations with different values for the radius of the plane's flight trajectory (smaller radius implies higher acceleration maneuvers). For each value of GPS delay and each flight radius, we perform 100 simulations and we average the attitude prediction error[2] and the velocity prediction error[3] during each simulation. Fig. 2 summarizes the results by showing an error bar for the results of each set of 100 simulations. According to Fig. 2, the average error due to the predictor is small even with large delays and even though we considered high gyro and accelerometer noises. The error increases as the amount of delay becomes larger. This is because the predictor relies on the forward integration of noisy gyro and accelerometer data from $t-\tau$ to t and hence the larger this period is, the more noise aggregates in the final prediction. Also, the errors increase when the flight radius decreases (i.e. when the plane performs higher acceleration maneuvers). The attitude prediction error is much less sensitive to this effect than the velocity prediction error.

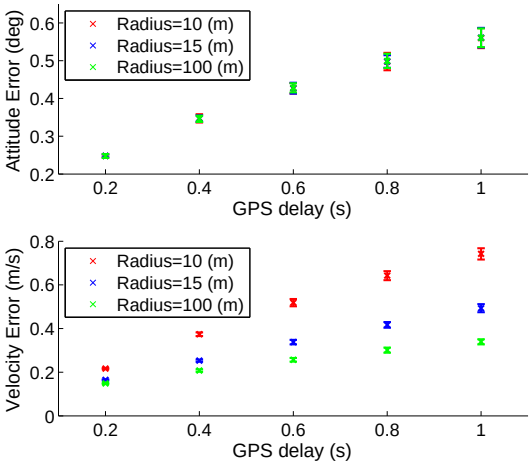

Figure 2: Pure attitude and velocity prediction errors. The error bars show two standard deviations. The linear velocity of the plane is 22.5 (m/s).

6.2. Total observer-predictor error:

Here we aim to provide realistic simulations to demonstrate that the observer-predictor approach is indeed capable of providing good estimates in practical situations. To this end, we use a comprehensive open source SITL system designed mostly by a group of ArduPilot/APM developers and CanberraUAV team [36–39]. This simulator allows building the ArduPilot/APM[4] autopilot code using an ordinary C++ compiler, making a native executable that allows testing the autopilot code without implementing on an actual hardware. The native executable emulates the hardware of the APM board at the register level, so the key low level hardware drivers (such as gyros, accelerometers, GPS, ADC, etc.) all run in the same way that they would run in a real flight. The SITL consists of three main

[2] Attitude error between two rotation matrices R_1 and R_2 is computed using the angle of rotation in the angle-axis representation of the error matrix $R_1R_2^\top$ given by $\frac{180}{\pi}\arccos(1 - 0.5\mathrm{tr}(I - R_1R_2^\top))$.

[3] Error between two velocity vectors is simply computed using the Euclidean norm of the difference of two velocity vectors.

[4] An open-source autopilot system that is widely used among the UAV enthusiasts community.

modules that interact with each other to simulate the whole system. The first module is JSBSim, an open source flight dynamics model [47] that simulates the trajectory of the vehicle. The second module is the ArduPilot/APM code that emulates the onboard autopilot software of the robot (including the estimation, control, navigation algorithms, etc.). The third module is MAVProxy which is a MAVLink ground station written in python [37].

For navigation purposes, SITL is capable of emulating various sensor measurements including GPS (longitude/latitude/linear velocity vector), accelerometers, gyros, etc. We setup the SITL parameters such that the GPS measurements are delayed. Here, we aim to demonstrate the performance of the observer-predictor approach when ArduPilot's native EKF [40] is used as an observer that takes the delayed measurements and provides the estimates $\hat{R}^\tau(t)$ and $\hat{v}^\tau(t)$ to the predictor (37)-(40)[5]. We take the estimates of gyro and accelerometer bias from the observer and use them along with the current gyro and accelerometer measurements to obtain estimates of current inputs $\hat{\Omega}(t)$ and $\hat{a}(t)$ fed into the predictor. We setup SITL to use ArduPlane (a fixed wing plane simulator) which tries to follow a desired square path of about 700 (m) by 200 (m) with the desired linear velocity of about 22.5 (m/s). Except for the GPS delay, all other simulation parameters (including the sensor sampling rates, noises, biases, EKF gain matrices, etc.) are set to the SITL's default values which correspond to typical low cost sensor suites. We choose various amounts of GPS delays and for each amount, we perform the simulation 20 times where in each simulation the plane's path is slightly different from other simulations[6]. We average the error trajectory during each flight simulation and then we compute the mean and standard deviation of all 20 flights. Repeating this procedure for various amounts of GPS delay, Fig. 3 summarizes the results by showing an error bar for the results of each set of 20 simulations. This figure shows that the observer-predictor remains stable even for large amounts of delay, which demonstrates the robustness of the proposed approach. A large portion of the total observer-predictor error is due to the observer itself. The average pure error due to the predictor is roughly the difference between the means of the red and the green plots of Fig. 3. This difference increases as the amount of delay increases, which is compatible to the results of Fig. 2[7]. Nevertheless, even when large GPS delay is present, the total error is within a very reasonable range for navigation and control using low cost sensor suites. It is worth noting that alternative estimation methods that either do not compensate for the delay or compensate with Lyapunov-Krasovskii terms perform poorly (or even become unstable) with large delays (for instance see [35, Fig. 2] and [48, Fig. 7]) while the proposed observer-predictor approach remains stable for all amounts of delay, demonstrating the robustness of the proposed approach.

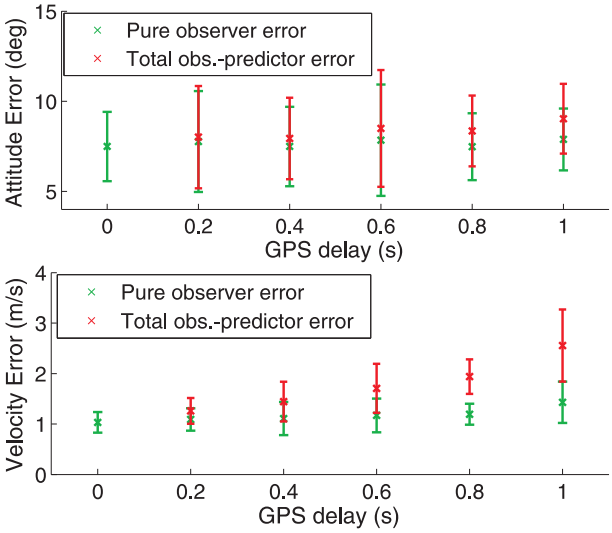

Figure 3: Pure observer error and total observer-predictor error. The eror bars show two standard deviations.

7. Conclusion

We consider a state estimation problem for invariant systems on Lie groups where measurements of the outputs of the system are delayed. Given an observer or filter that has desired stability properties when the system outputs are delay-free, we propose an estimation methodology that preserves those stability properties when the system outputs are delayed. The proposed approach relies on combining the observer with a predictor that compensates for the delay. The delayed measurements are fused in the observer to obtain estimates of the delayed state. Those delayed estimates are then fed into the predictor to obtain the prediction of the current state. We employ invariance of the underlying systems together with the Lie group structure of the state space in order to design recursive predictors that are computationally simple and demonstrate strong costability properties ensuring that the observer-predictor combination preserves the stability properties of the observer. Using a sophisticated Software-In-The-Loop simulator, we demonstrate the robustness of the proposed predictors in practical situations even when large sensor delay is present.

8. Acknowledgment

The Authors would like to thank Sean O'Brien, Andrew Tridgell, Paul Riseborough, Evan Slatyer, Benjamin Nizette, Salim Masoumi, Arash Khodaparastsichani, and Juan Adarve for their contributions in developing the simulations with the ArduPilot/APM SITL system.

[5] For simulation studies of this paper, we omit the delay compensation parts of the ArduPilot's native EKF and turn it into an standard EKF before we combine it with the proposed predictor.

[6] This is because the ArduPilot system implements closed-loop control based on noisy data

[7] The pure predictor error of Fig. 3 is larger than Fig. 2 since the observer's estimation error of gyro and accelerometer biases (in addition to the input noise and plane's acceleration) contribute to the prediction error.

APPENDIX

Lemma 2. *The largest singular value of the adjoint map $Ad_X : \mathfrak{g} \to \mathfrak{g}$ wrt. the induced Euclidean norm on $\mathfrak{g}$ is bounded by cond(X).* □

Proof: Embed $\mathfrak{g}$ into $\mathbb{R}^{m\times m}$. Since $\dim(\mathfrak{g}) \leq m^2$, the largest singular value of the adjoint map $\mathrm{Ad}_X : \mathfrak{g} \to \mathfrak{g}$ is less than or equal to the largest singular value of the map $\mathrm{Ad}_X : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$. We show that the later equals to cond(X). Invoking the property $\mathrm{vec}(XwX^{-1}) = X^{-\top} \otimes X\mathrm{vec}(w)$ where $\mathrm{vec}(w) \in \mathbb{R}^{m^2}$ is the vectorization of the matrix $w \in \mathbb{R}^{m\times m}$ and $\otimes$ denotes the Kronecker product, one can conclude that the matrix representation of $\mathrm{Ad}_X : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$ wrt. the standard basis for its domain and co-domain is given by $[\![\mathrm{Ad}_X]\!] = X^{-\top} \otimes X$. This implies that $\overline{\sigma}([\![\mathrm{Ad}_X]\!]) = \overline{\sigma}(X^{-\top})\overline{\sigma}(X) = \underline{\sigma}(X)^{-1}\overline{\sigma}(X) = \mathrm{cond}(X)$. This concludes the proof. ■

Lemma 3. *We have $d_r(X, Y) \leq \overline{\sigma}(Ad_Z)d_r(Z^{-1}X, Z^{-1}Y)$ for all $X, Y, Z \in G$.*

Proof: Define a right-invariant distance between $X, Y \in G$ as the infimum of the lengths of the curves connecting X and Y wrt. the metric $\langle ., . \rangle^r$, such that $d_r(X, Y) = \inf_{\gamma(0)=X,\ \gamma(1)=Y} \int_0^1 |\dot{\gamma}(s)|_r^{\gamma(s)} ds$. Using a change of variable $\bar{\gamma}(s) := Z^{-1}\gamma(s)$ we have $d_r(X, Y) = \inf_{\bar{\gamma}(0)=Z^{-1}X,\ \bar{\gamma}(1)=Z^{-1}Y} \int_0^1 |Z\dot{\bar{\gamma}}(s)|_r^{Z\bar{\gamma}(s)} ds = \inf_{\bar{\gamma}(0)=Z^{-1}X,\ \bar{\gamma}(1)=Z^{-1}Y} \int_0^1 \|Z\dot{\bar{\gamma}}(s)\bar{\gamma}(s)^{-1}Z^{-1}\| ds \leq \overline{\sigma}(\mathrm{Ad}_Z) \inf_{\bar{\gamma}(0)=Z^{-1}X,\ \bar{\gamma}(1)=Z^{-1}Y} \int_0^1 \|\dot{\bar{\gamma}}(s)\bar{\gamma}(s)^{-1}\| ds = \overline{\sigma}(\mathrm{Ad}_Z) \inf_{\bar{\gamma}(0)=Z^{-1}X,\ \bar{\gamma}(1)=Z^{-1}Y} \int_0^1 |\dot{\bar{\gamma}}(s)|_r^{\bar{\gamma}(s)} ds = \overline{\sigma}(\mathrm{Ad}_Z) d_r(Z^{-1}X, Z^{-1}Y)$. ■

Lemma 4. *Consider the system $\dot{\Xi}_l(t) = u_l^b(t)\Xi_l(t)$ (resp. $\dot{\Xi}_r(t) = \Xi_r(t)u_r^b(t)$) with $\Xi_l(0) = \Xi_0 \in G$ (resp. $\Xi_r(0) = \Xi_0 \in G$) where $u_l^b(t), u_r^b(t) \in \mathfrak{g}$.*

(I) If $u_l^b(t)$ (resp. $u_r^b(t)$) is bounded for all $t \geq 0$, then $d_l(\Xi_r(t-\tau), \Xi_r(t))$ (resp. $d_l(\Xi_r(t-\tau), \Xi_r(t))$) is bounded for all $\tau \geq 0$ and all $t \geq \tau$.

(II) If $\|u_l^b(t)\| \to 0$ (resp. $\|u_r^b(t)\| \to 0$), then $d_r(\Xi_l(t-\tau), \Xi_l(t)) \to 0$ (resp. $d_l(\Xi_r(t-\tau), \Xi_r(t)) \to 0$) for all $\tau \geq 0$.

(III) If $\|u_l^b(t)\| \xrightarrow{exp} 0$ (resp. $\|u_r^b(t)\| \xrightarrow{exp} 0$), then $d_r(\Xi_l(t-\tau), \Xi_l(t)) \xrightarrow{exp} 0$ (resp. $d_l(\Xi_r(t-\tau), \Xi_r(t)) \xrightarrow{exp} 0$) for all $\tau \geq 0$.

□

Proof: We prove the Lemma for the right-invariant system $\dot{\Xi}_l(t) = u_l^b(t)\Xi_l(t)$ and the right-invariant distance $d_r(.,.)$. The proof for the left-invariant case can be obtained similarly. The length of the curve $s \mapsto \Xi_l(s),\ s \in [t-\tau, t]$ connecting $\Xi_l(t-\tau)$ to $\Xi_l(t)$ wrt. the metric $\langle ., . \rangle_r$ is given by $L_r(\Xi_l(t-\tau), \Xi_l(t)) = \int_{t-\tau}^t |\dot{\Xi}_l(s)|_r^{\Xi_l(s)} ds = \int_{t-\tau}^t |u_l^b(s)\Xi_l(s)|_r^{\Xi_l(s)} ds = \int_{t-\tau}^t \|u_l^b(s)\| ds$.

If $u_l^b(t)$ is bounded, there exists a constant c_u such that $\|u_l^b(t)\| \leq c_u$ for all $t \geq 0$. Hence $L_l(\Xi_l(t-\tau), \Xi_l(t)) \leq \tau c_u$ and consequently $d_r(\Xi_l(t-\tau), \Xi_l(t))$ is bounded by $d_r(\Xi_l(t-\tau), \Xi_l(t)) \leq L_r(\Xi_l(t-\tau), \Xi_l(t)) \leq \tau c_u$. This proves part (I).

If $\|u_l^b(t)\| \to 0$, then for all $\epsilon_u > 0$ there exist a T_u such that for all $t \geq T_u$ we have $\|u_l^b(t)\| < \epsilon_u$ and hence, $L_r(\Xi_l(t-\tau), \Xi_l(t)) < \tau\epsilon_u$. Hence, we have $d_r(\Xi_l(t-\tau), \Xi_l(t)) \leq L_r(\Xi_l(t-\tau), \Xi_l(t)) < \tau\epsilon_u$ for all $t \geq T_u$. Consequently, $d_r(\Xi_l(t-\tau), \Xi_l(t)) \to 0$. This proves part (II).

If $\|u_l^b(t)\| \xrightarrow{\exp} 0$, then there exist positive constants c and $\alpha > 0$ such that $\|u_l^b(t)\| \leq c\exp(-\alpha t)$ for all $t > 0$. We have $L_r(\Xi_l(t-\tau), \Xi_l(t)) = \int_{t-\tau}^t \|u_l^b(s)\| ds \leq \int_{t-\tau}^t c\exp(-\alpha s) ds = \frac{c}{\alpha}(\exp(\alpha\tau) - 1)\exp(-\alpha t)$. Hence $d_r(X(t-\tau), X(t)) \leq L_r(\Xi_l(t-\tau), \Xi_l(t)) \xrightarrow{\exp} 0$. This proves part (III). ■

References

[1] V. Jurdjevic, Geometric control theory, Cambridge university press, 1997.
[2] F. Bullo, Geometric control of mechanical systems, Vol. 49, Springer, 2005.
[3] F. L. Markley, J. L. Crassidis, Fundamentals of Spacecraft Attitude Determination and Control, Springer, 2014.
[4] R. Mahony, T. Hamel, J.M. Pflimlin, Nonlinear complementary filters on the special orthogonal group, IEEE Trans. Autom. Control 53 (5) (2008) 1203–1218.
[5] H. F. Grip, T. I. Fossen, T. A. Johansen, A. Saberi, Globally exponentially stable attitude and gyro bias estimation with application to GNSS/INS integration, Automatica 51 (2015) 158–166.
[6] J. Vasconcelos, C. Silvestre, P. Oliveira, A nonlinear observer for rigid body attitude estimation using vector observations, in: Proc. IFAC World Congr., Korea, 2008.
[7] S. Bonnabel, P. Martin, P. Rouchon, Symmetry-preserving observers, IEEE Trans. Autom. Control 53 (11) (2008) 2514–2526.
[8] A. Roberts, A. Tayebi, On the attitude estimation of accelerating rigid-bodies using GPS and IMU measurements, in: IEEE Conf. Decision and Control and European Control Conf. (CDC-ECC), 2011, pp. 8088–8093.
[9] M. Izadi, A. K. Sanyal, Rigid body attitude estimation based on the Lagrange–d'Alembert principle, Automatica 50 (10) (2014) 2570–2577.
[10] A. K. Sanyal, T. Lee, M. Leok, N. H. McClamroch, Global optimal attitude estimation using uncertainty ellipsoids, Systems & Control Letters 57 (3) (2008) 236–245.
[11] M. Zamani, J. Trumpf, R. Mahony, Minimum-energy filtering for attitude estimation, IEEE Trans. Autom. Control 58 (2013) 2917–2921.
[12] J. Vasconcelos, R. Cunha, C. Silvestre, P. Oliveira, A nonlinear position and attitude observer on SE(3) using landmark measurements, Systems & Control Letters 59 (3) (2010) 155–166.
[13] M.-D. Hua, M. Zamani, J. Trumpf, R. Mahony, 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.
[14] H. Rehbinder, B. K. Ghosh, Pose estimation using line-based dynamic vision and inertial sensors, IEEE Trans. Automatic Control 48 (2) (2003) 186–199.
[15] S. S. Rodrigues, N. Crasta, A. P. Aguiar, F. S. Leite, State estimation for systems on SE(3) with implicit outputs: An application to visual servoing, in: Proceedings of NOLCOS, 2010.
[16] T. Hamel, R. Mahony, J. Trumpf, P. Morin, M.-D. Hua, Homography estimation on the special linear group based on direct point correspondence, in: IEEE Conf. Decision and Control and European Control Conf. (CDC-ECC), 2011, 2011, pp. 7902–7908.
[17] N. E. Leonard, P. S. Krishnaprasad, Motion control of drift-free, left-invariant systems on Lie groups, IEEE Trans. Autom. Control 40 (9) (1995) 1539–1554.
[18] C. Lageman, J. Trumpf, R. Mahony, Gradient-like observers for invariant dynamics on a Lie group, IEEE Trans. Autom. Control 55 (2) (2010) 367–377.

[19] R. Mahony, J. Trumpf, T. Hamel, Observers for kinematic systems with symmetry, in: Proc. IFAC Symposium on Nonlinear Control Systems, 2013, pp. 617–633.
[20] A. Khosravian, J. Trumpf, R. Mahony, C. Lageman, Observers for invariant systems on lie groups with biased input measurements and homogeneous outputs, Automatica 55 (2015) 19–26, (to appear).
[21] A. Barrau, S. Bonnabel, The invariant extended kalman filter as a stable observer, arXiv preprint arXiv:1410.1465.
[22] D. B. Kingston, R. W. Beard, Real-time attitude and position estimation for small UAVs using low-cost sensors, in: AIAA 3rd Unmanned Unlimited Technical Conference, Workshop and Exhibit, 2004.
[23] P. Riseborough, 3DRobotics, private communication (2015).
[24] T. Ahmed-Ali, I. Karafyllis, F. Lamnabhi-Lagarrigue, Global exponential sampled-data observers for nonlinear systems with delayed measurements, Systems & Control Letters 62 (7) (2013) 539–549.
[25] I. Karafyllis, C. Kravaris, From continuous-time design to sampled-data design of observers, IEEE Trans. Automatic Control 54 (9) (2009) 2169–2174.
[26] V. Van Assche, T. Ahmed-Ali, C. Hann, F. Lamnabhi-Lagarrigue, High gain observer design for nonlinear systems with time varying delayed measurements, in: IFAC World congress, 2011, pp. 692–696.
[27] W. Aggoune, M. Boutayeb, M. Darouach, Observers design for a class of nonlinear systems with time-varying delay, in: Proc. IEEE Conf. on Decision and Control, Vol. 3, 1999, pp. 2912–2913.
[28] F. Cacace, A. Germani, C. Manes, An observer for a class of nonlinear systems with time varying observation delay, Systems & Control Letters 59 (5) (2010) 305–312.
[29] A. Germani, C. Manes, P. Pepe, A new approach to state observation of nonlinear systems with delayed output, IEEE Trans. Automatic Control 47 (1) (2002) 96–101.
[30] S. Battilotti, Nonlinear predictors for systems with bounded trajectories and delayed measurements¡, in: Proc. IFAC World Congr., 2014, pp. 6812–9817.
[31] N. Kazantzis, R. A. Wright, Nonlinear observer design in the presence of delayed output measurements, Systems & Control letters 54 (9) (2005) 877–886.
[32] T. Ahmed-Ali, E. Cherrier, F. Lamnabhi-Lagarrigue, Cascade high gain predictors for a class of nonlinear systems, IEEE Trans. Automatic Control 57 (1) (2012) 221–226.
[33] S. Bahrami, M. Namvar, Delay compensation in global estimation of rigid-body attitude under biased velocity measurement, in: Proc. IEEE Conf. on Decision and Control, 2014.
[34] A. Khosravian, J. Trumpf, R. Mahony, State estimation for nonlinear systems with delayed output measurements, in: IEEE Conf. Decision and Control, 2015, (under review).
[35] A. Khosravian, J. Trumpf, R. Mahony, T. Hamel, Velocity aided attitude estimation on SO(3) with sensor delay, in: Proc. IEEE Conf. on Decision and Control, 2014.
[36] Setting up SITL on Linux, `http://dev.ardupilot.com/wiki/simulation-2/setting-up-sitl-on-linux/` (visited on January 2015).
[37] MAVProxy, `http://tridge.github.io/MAVProxy/` (visited on January 2015).
[38] ArduPilot Project, `https://github.com/diydrones/ardupilot` (visited on January 2015).
[39] CanberraUAV, `http://www.canberrauav.com/` (visited on January 2015).
[40] ArduPilot/APM, APM: Navigation Extended Kalman Filter Overview, `http://plane.ardupilot.com/wiki/common-apm-navigation-extended-kalman-filter-overview/` (visited on March 2014).
[41] S. O'Brien, Predictors for systems on manifolds with time-delayed measurements, Australian National University, Bachelor's Thesis, 2014.
[42] P. Martin, E. Salaün, Design and implementation of a low-cost observer-based attitude and heading reference system, Control Engineering Practice 18 (7) (2010) 712–722.
[43] M.-D. Hua, Attitude estimation for accelerated vehicles using GPS/INS measurements, Control Engineering Practice 18 (7) (2010) 723–732.
[44] M.-D. Hua, P. Martin, T. Hamel, Velocity-aided attitude estimation for accelerated rigid bodies, arXiv preprint arXiv:1411.3953.
[45] G. Allibert, D. Abeywardena, M. Bangura, R. Mahony, Estimating body-fixed frame velocity and attitude from inertial measurements for a quadrotor vehicle, in: IEEE Conference on Control Applications, 2014, pp. 978–983.
[46] J. F. Vasconcelos, C. Silvestre, P. Oliveira, A nonlinear GPS/IMU based observer for rigid body attitude and position estimation, in: IEEE Conf. Decision and Control, 2008, pp. 1255–1260.
[47] JSBSim, `http://jsbsim.sourceforge.net/` (visited on January 2015).
[48] A. Khosravian, J. Trumpf, R. Mahony, T. Hamel, Recursive attitude estimation in the presence of multi-rate and multi-delay vector measurements, in: Proc. American Control Conference, 2015, (to appear).