Global Observer Design for a Class of Linear Observed Systems on Groups

Changwu Liu and Yuan Shen

***Abstract*— Linear observed systems on groups encode the geometry of a variety of practical state estimation problems. In this paper, we propose a unified observer framework for a class of linear observed systems by restricting a bi-invariant system on a Lie group to its normal subgroup. This structural property powerfully enables a system immersion of the original system into a linear time-varying system. Leveraging the immersion, an observer is constructed by first designing a Kalman-like observer for the immersed system and then reconstructing the group-valued state via optimization. Under a rank condition, global exponential stability (GES) is achieved provided one global optimum of the reconstruction optimization is found, reflecting the topological difficulties inherent to the non-Euclidean state space. Semi-global stability is guaranteed when input biases are jointly estimated. The theory is applied to the GES observer design for two-frame systems, capable of modeling a family of navigation problems. Two non-trivial examples are provided to illustrate implementation details.**

***Index Terms*— Asymptotic Nonlinear Observers, Geometric Methods, Kalman Filters, Navigation, State Estimation, Systems on Lie Groups.**

I. Introduction

LINEAR observed systems on groups are systems whose flows are compatible with the group automorphisms [1]–[3]. State estimation of those systems is of particular interest to the robotics and automation community as it captures the geometry of many practical problems [4]–[7]. Prototypical examples include the navigation of rigid bodies using multi-sensor information [4], [8]–[14], i.e., estimating the attitudes, positions and linear velocities of some moving objects. Those variables are naturally combined together as transformations between coordinate frames. The set of transformations is a closed subgroup of the general linear group and hence is endowed with a natural Lie group structure. The success of those applications is attributed to respecting geometry, i.e., utilizing the group formulation of the state space as well as system properties mainly linked to linear observed structures.

An observer or state estimator is a dynamical system driven by known inputs and measurements [15]. Asymptotic stability is the central objective in observer design. The main difficulties hindering the guarantee of stability of observers for general linear observed systems on groups are two-fold: the nonlinearities of the system equations and the non-Euclidean state space topology. The first category of observers for linear observed systems is based on linearization respecting the geometry. The invariant filters (IEKF) linearize the system in exponential coordinates to obtain state-independent error dynamics leveraging the linear observed structure. With gains computed from a Riccati equation akin to those in the Kalman filter using error dynamics, IEKF achieves local stability under conditions similar to Kalman filters [5], [6]. Equivariant filters (EqF) consider a wider range of systems on homogeneous spaces, encompassing linear observed systems on groups. EqF lifts the system to its symmetry group acting on the base manifold and designs an observer utilizing the invariant group error [16], [17]. Again, its gain is obtained from a Riccati equation whose coefficients follow from linearization in the coordinates of the base manifold. Though IEKF and EqF are designed for general systems, their stability domain is only local due to linearization. The second category of observers is characterized by constructive approaches that achieve almost-global stability [18], [19]. This is the best one could have if designing the correction by a continuous vector field, originating from topological obstructions [20] when the state space contains a non-contractible component. To the best of our knowledge, there is no out-of-the-box, easy-to-implement observer that achieves global exponential stability (GES) for general linear observed systems on groups.

There are numerous case-by-case studies for global observer design in rigid-body navigation. In addition to classical constructive almost-global observers for attitude estimation [21], [22], recently proposed hybrid observers for IMU-based navigation with landmark or vision-type measurements are discussed in [23]–[25]. Such observers consist of continuous flows and discrete jumps, and achieve GES, thereby overcoming the topological obstructions. The related construction relies heavily on the matrix group representation of IMU dynamics and corresponding innovations, both of which are closely related to the linear observed structure. In contrast to the constructive methods, switching to robocentric coordinates using body-referenced linear velocities and landmark positions simplifies the system model to obtain a linear time-varying (LTV) system [26]–[29]. A Riccati observer [30] is designed for those auxiliary LTVs and the original states are reconstructed later. Strong GES guarantees are obtained under persistent excitation. The above methods focus on specific examples, only a tip of an iceberg of the systems capable of being modeled by linear observed formulation. Moreover, the success of several key techniques implicitly relies on linear observed structures. The feasibility of those techniques for

The authors are with Department of Electronic Engineering, Tsinghua University, 100084 Beijing, China (e-mail: liucw_ee@tsinghua.edu.cn; shenyuan_ee@tsinghua.edu.cn).

general linear observed systems on groups has never been studied before.

In this paper, we delve deep into the linear observed structure and reveal a powerful connection between this structure and the possibility of an LTV immersion of linear observed system on groups. A global observer framework follows naturally. Our contributions are summarized below.

- Group-theoretic conditions allowing LTV immersion of linear observed systems on groups are established. The class of systems arising from the restriction of a bi-invariant system to any normal subgroup of the state space can be immersed into LTVs.
- An out-of-the-box unified observer is proposed for immersible linear observed systems on groups. If a rank condition on the system structure is satisfied, GES is achieved provided one global optimum of the related optimization on the group can be found, reflecting the topological difficulties. Joint estimation of input bias is also tackled.
- The observer framework is applied to two-frame systems [4], a powerful scheme modeling a broad class of practical problems. Implementations of GES observers for immersible two-frame systems are provided with additional extension to bearing and range measurements.
- Two non-trivial navigational examples are given to illustrate the GES observer implementation, never treated before with such strong guarantees in the literature.

For the rest of the paper, Section II reviews key mathematical preliminaries. The central system properties allowing LTV immersion are discussed in Section III. Our GES observer is detailed in Section IV. Joint estimation of input bias is also discussed here. A prototypical application of the theory to immersible two-frame systems is discussed in Section V. Non-trivial navigational examples with selected simulations are presented in Section VI.

II. Preliminaries

A. Notations

Let $\mathbb{R}$, $\mathbb{N}$ denote the sets of real and natural numbers. $\mathbb{R}^n$ denotes an n-dimensional vector space and $\mathbb{S}^n$ is the set of all unit vectors of $\mathbb{R}^{n+1}$, termed the n-sphere. Lowercase and capital letters represent vectors and matrices respectively, unless otherwise noted. We use $\|\cdot\|$ for the Euclidean norm of a vector or the Frobenius norm of a matrix. Other norms are marked explicitly. $\preceq$, $\succeq$, $\prec$, $\succ$ denote the partial order on symmetric matrices. $\mathbb{S}^n_+$ is the cone of symmetric positive definite $\mathbb{R}^{n\times n}$-matrices. $\hat{(\cdot)}$ and $(\cdot)$ denote the estimated and the true state, respectively. $\mathrm{diag}(\cdot)$ denotes the diagonal or block-diagonal matrix.

B. Group Theory Basics

A Lie group G is simultaneously an abstract group and a smooth manifold. The set $L(G)$ or $R(G)$ of left- or right-invariant vector fields is a finite-dimensional vector space closed under commutators, yielding an identification of $L(G)$ or $R(G)$ with its Lie algebra $\mathfrak{g}$, i.e., the tangent space $T_{id}G$ at the identity. This identification also induces a Lie bracket on $\mathfrak{g}$. Let $\Gamma(TG)$ denote the set of all smooth vector fields on G as an infinite-dimensional $\mathbb{R}$-vector space. The automorphism group $\mathrm{Aut}(G)$ is the set of bijective maps $\phi : G \to G$ that preserve multiplication: $\phi(g_1g_2) = \phi(g_1)\phi(g_2), \forall g_1, g_2 \in G$. The conjugations $\{\phi_g : G \to G \mid \phi_g(h) = ghg^{-1}, g \in G\}$ form the inner-automorphism group, denoted $\mathrm{Inn}(G)$. A normal subgroup of G is a subgroup invariant under any conjugation of the original group. $\mathrm{Inn}(G)$ is a normal subgroup of $\mathrm{Aut}(G)$, making the quotient $\mathrm{Out}(G) := \mathrm{Aut}(G)/\mathrm{Inn}(G)$ well-defined, termed the group of outer automorphisms. In the matrix case, $\mathrm{Aut}(G)$ and $\mathrm{Inn}(G)$ are smooth, and their corresponding Lie algebras are denoted $\mathfrak{aut}(G)$ and $\mathfrak{inn}(G)$ respectively. We are interested in matrix Lie groups, namely the closed subgroups of $\mathrm{GL}(n,\mathbb{R})$, which are invertible $\mathbb{R}^{n\times n}$ matrices under multiplication. A typical group representing rotations on $\mathbb{R}^d$ is

$$\mathrm{SO}(d) = \left\{ R \in \mathrm{GL}(d,\mathbb{R}) \,\middle|\, R^\top R = I_{d\times d},\ \det(R) = 1 \right\}. \quad (1)$$

Note that $\mathbb{R}^{n\times n}$ provides a global embedding for matrix groups as well as for their Lie algebras. Let $\mathcal{L}_{\mathfrak{g}} : \mathbb{R}^{\dim \mathfrak{g}} \to \mathbb{R}^{n\times n}$ be the linear isomorphism between some vector space and the matrix embedding of the Lie algebra. The $\mathcal{L}_{\mathfrak{so}(d)}$ embeds elements of $\mathbb{R}^{\frac{d(d-1)}{2}}$ into skew-symmetric $\mathbb{R}^{d\times d}$, e.g.,

$$\mathcal{L}_{\mathfrak{so}(2)}(x) = \begin{bmatrix} 0 & -x \\ x & 0 \end{bmatrix}, \quad (2)$$

$$\mathcal{L}_{\mathfrak{so}(3)}\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}. \quad (3)$$

We also use the compact notation $(\cdot)^\times := \mathcal{L}_{\mathfrak{so}(d)}(\cdot)$.

Let $\varphi : G \times \mathbb{R}^{d_y} \to \mathbb{R}^{d_y}, (\chi, d) \mapsto \varphi(\chi, d)$ be a smooth map termed a left group action of G on a vector space $\mathbb{R}^{d_y}$ satisfying $\varphi(id_G, d) = d$ and $\varphi(\chi_1\chi_2, d) = \varphi(\chi_1, \varphi(\chi_2, d))$ for every $\chi_1, \chi_2 \in G$ and $d \in \mathbb{R}^{d_y}$. Let $\tilde{\varphi} : G \times \mathbb{R}^{d_y} \to \mathbb{R}^{d_y}$ be defined by $\tilde{\varphi}(\chi, d) := \varphi(\chi^{-1}, d)$, where $\varphi(\cdot,\cdot)$ is the previous left action. For every $d \in \mathbb{R}^{d_y}$ and $\chi_1, \chi_2 \in G$, we have $\tilde{\varphi}(id_G, d) = d$, but $\tilde{\varphi}(\chi_1\chi_2, d) = \tilde{\varphi}(\chi_2, \tilde{\varphi}(\chi_1, d))$. Such $\tilde{\varphi}$ is called a right action on $\mathbb{R}^{d_y}$. If fixing either $\chi \in G$ or $d \in \mathbb{R}^{d_y}$, we define two partial maps by $\varphi_\chi : \mathbb{R}^{d_y} \to \mathbb{R}^{d_y}$ via $\varphi_\chi(d) := \varphi(\chi, d)$ and $\varphi_d : G \to \mathbb{R}^{d_y}$ via $\varphi_d(\chi) := \varphi(\chi, d)$. Their push-forwards are then defined. $\varphi_{\chi*} : T_d\mathbb{R}^{d_y} \to T_{\varphi(\chi,d)}\mathbb{R}^{d_y}$ is the push-forward of φ_χ anchored at $d \in \mathbb{R}^{d_y}$. As both $T_d\mathbb{R}^{d_y}$ and $T_{\varphi(\chi,d)}\mathbb{R}^{d_y}$ are isomorphic to $\mathbb{R}^{d_y}$ itself, we identify $\varphi_{\chi*} : \mathbb{R}^{d_y} \to \mathbb{R}^{d_y}$ with a linear operator. $\varphi_{d*} : T_\chi G \to T_{\varphi(\chi,d)}\mathbb{R}^{d_y}$ is the push-forward of the other partial map. We can view this map as $\varphi_{d*} : T_\chi G \to \mathbb{R}^{d_y}$. In this paper, we consider matrix Lie groups and φ being linear actions. Moreover, χ and $\dot{\chi}$ are viewed as invertible square matrices of the same size. This means φ and the push-forwards of the partial maps can all be realized via matrix multiplication. The simplified notation $\chi d := \varphi(\chi, d)$ can be used. Meanwhile, the expressions $\dot{\chi}_1\chi_2 d$ and $\chi_1\dot{\chi}_2 d$ are legal and the terms can be freely associated using parentheses.

Readers are referred to [31], [32] for a thorough mathematical preparation.

C. Observability and Matrix Riccati/Lyapunov Equations

Consider an LTV system $\dot{x} = A_t x + B_t u,\ y = H_t x$ where $x \in \mathbb{R}^{d_x},\ A_t \in \mathbb{R}^{d_x \times d_x},\ B_t \in \mathbb{R}^{d_x \times d_u}$ and $H_t \in \mathbb{R}^{d_y \times d_x}$. $d_x,\ d_u,\ d_y$ are dimensions of the state, input, and output, respectively. The subscripts t indicate the dependence on time. Let $\Phi(t_2, t_1) \in \mathbb{R}^{d_x \times d_x}$ be the state transition matrix satisfying $\frac{\partial \Phi(t_2,t_1)}{\partial t_2} = A_{t_2}\Phi(t_2, t_1)$ with initial condition $\Phi(t_1, t_1) = I$. Define the observability Gramian [30], [33] with respect to $R_\tau \in \mathbb{R}^{d_y \times d_y}$ as

$$\mathcal{O}(t_2, t_1) := \int_{t_1}^{t_2} \Phi^\top(\tau, t_1) H_\tau^\top R_\tau^{-1} H_\tau \Phi(\tau, t_1) d\tau. \quad (4)$$

The LTV system is said to be uniformly observable (persistently excited) if there exist constants $\delta, \alpha > 0$, such that $\mathcal{O}(t + \delta, t) \succeq \alpha I$ holds for every $t \in \mathbb{R}$ [33]. Similarly, the determinability Gramian [34], [35] of such LTV system is given by

$$\mathcal{D}(t_2, t_1) := \int_{t_1}^{t_2} \Phi^\top(\tau, t_2) H_\tau^\top R_\tau^{-1} H_\tau \Phi(\tau, t_2) d\tau. \quad (5)$$

The system is persistently determinable if there exist constants $\delta, \alpha > 0$, such that $\mathcal{D}(t + \delta, t) \succeq \alpha I, \forall t \in \mathbb{R}$. The matrix Riccati equation is

$$\dot{P} = A_t P_t + P_t A_t^\top + Q_t - P_t H_t^\top R_t^{-1} H_t P_t \quad (6)$$

for $P_t, Q_t \in \mathbb{S}_+^{d_x}$. If the LTV system is uniformly controllable and observable, then the eigenvalues of P are uniformly lower and upper bounded [36], i.e., $\exists p_m, p_M > 0, p_m I \preceq P_t \preceq p_M I, \forall t$. Uniform observability is related to the lower bound p_m specifically. (6) is slightly modified [37] to obtain an explicit solution

$$\dot{P} = \lambda P_t + A_t P_t + P_t A_t^\top - P_t H_t^\top R_t^{-1} H_t P_t, \quad (7)$$

where $\lambda > 0$. Persistent determinability uniformly lower bounds P, i.e., $\exists p_m > 0, P \succeq p_m I, \forall t$.

III. Properties Allowing LTV Immersion of Linear Observed Systems on Groups

The flow of a linear observed system evolving on a Lie group G is closely related to the automorphism group $\mathrm{Aut}(G)$ of the state space [1], [3], [6]. These connections with additional group-theoretic conditions enable a powerful construction of a high-dimensional auxiliary LTV system with the same input-output behavior as the original system on the group. This immersion acts as a cornerstone of our global observer design. For simplicity in presentation, we consider a system with a single measurement.

A. Linear Observed Systems on Groups and Their Structural Properties

Let $\chi \in G$ be the group-valued state. Denote by u the input taking values in $\mathbb{R}^{\dim \mathfrak{g}}$. Let $y \in \mathbb{R}^{d_y}$ be the output in some vector space. A linear observed system on G is given by

$$\dot{\chi} = f_u(\chi), \quad y = h(\chi), \quad (8)$$

where $f_u \in \Gamma(TG)$ is a smooth vector field on G depending on u, known as the process model, with group-affine properties [5] satisfying

$$f_u(\chi_1\chi_2) = \chi_1 f_u(\chi_2) + f_u(\chi_1)\chi_2 - \chi_1 f_u(id_G)\chi_2, \quad (9)$$

for every $\chi_1, \chi_2 \in G$, and every $u \in \mathbb{R}^{\dim \mathfrak{g}}$. Moreover, $h : G \to \mathbb{R}^{d_t}$, known as the algebraic measurement, is given by a group action of G on a constant vector $d \in \mathbb{R}^{d_y}$ as

$$h^R(\chi) = \chi^{-1} d \text{ or } h^L(\chi) = \chi d, \quad (10)$$

where we implicitly denote the G-action $\varphi : G \times \mathbb{R}^{d_y} \to \mathbb{R}^{d_y}$ as $\chi d := \varphi(\chi, d)$ or $\chi^{-1} d := \varphi(\chi^{-1}, d)$. The superscript 'L' or 'R' marks the left or right action respectively. Only linear actions with respect to the second slot, d, are considered.

Define a smooth mapping $\phi : \mathbb{R} \times G \times \mathbb{R} \to G$ with $\phi(t; \chi, t) = \chi$ for all $\chi \in G$. We interpret $\phi(s; \chi, t)$ as the state transition from an initial value $\chi \in G$ at timestamp t to timestamp s. If we arbitrarily fix s, t in the first and third slots and further require that ϕ is an automorphism of G, i.e., $\phi(s; \chi_1\chi_2, t) = \phi(s; \chi_1, t)\phi(s; \chi_2, t), \forall \chi_1, \chi_2 \in G$, then, for fixed t, the map $s \mapsto \phi(s; \cdot, t)$ is a smooth curve on $\mathrm{Aut}(G)$. Define the vector field $g : G \to TG$ by

$$g(\chi) := \frac{d}{ds}|_{s=t}\phi(s; \chi, t) \in T_\chi G\ . \quad (11)$$

One can verify that

$$g(\chi_1\chi_2) = g(\chi_1)\chi_2 + \chi_1 g(\chi_2), \forall \chi_1, \chi_2 \in G. \quad (12)$$

Evaluating the derivative at $s = t$ corresponds to differentiating at the identity element of $\mathrm{Aut}(G)$. Consequently, g can be identified with an element of $\mathfrak{aut}(G)$, namely the Lie algebra of the automorphism group $\mathrm{Aut}(G)$. Moreover, $\mathfrak{aut}(G)$ is the collection of all vector fields on G satisfying property (12). Let $L(G) := \{\tilde{f} \in \Gamma(TG) \mid \chi\tilde{f}(\chi') = \tilde{f}(\chi\chi'), \forall \chi, \chi' \in G\}$ and $R(G) := \{\tilde{f} \in \Gamma(TG) \mid \tilde{f}(\chi')\chi = \tilde{f}(\chi'\chi), \forall \chi, \chi' \in G\}$ denote the sets of left- and right-invariant vector fields on G respectively. Note that both $L(G)$ and $R(G)$ are isomorphic to $\mathfrak{g}$ as $\mathbb{R}$-vector spaces [31], [32]. We now restate the characterization of group-affine properties using (9), equivalent to those in [19], [38].

Proposition 1: For any fixed time t, the set of group-affine vector fields f_us is exactly one of the following subsets of $\Gamma(TG)$:

$$\mathfrak{aut}(G) + L(G) := \left\{ g + \tilde{f} \,\middle|\, g \in \mathfrak{aut}(G), \tilde{f} \in L(G) \right\}, \quad (13)$$

$$\mathfrak{aut}(G) + R(G) := \left\{ g + \tilde{f} \,\middle|\, g \in \mathfrak{aut}(G), \tilde{f} \in R(G) \right\}. \quad (14)$$

In other words, f_u is decomposable as one of the following

$$f_u(\chi) = g_u^L(\chi) + \chi f_u(id_G), \quad (15)$$

$$f_u(\chi) = g_u^R(\chi) + f_u(id_G)\chi, \quad (16)$$

with the vector fields $g_u^L(\chi)$ and $g_u^R(\chi)$ satisfying (12).

Proof: See Appendix A. ■

Remark 1: The vector field g_u corresponds to the autonomous error evolution of the dynamics in (8): $\dot{e}^L = g_u^L(e^L)$ with the left-invariant error $e^L := \chi^{-1}\hat{\chi}$ or $\dot{e}^R = g_u^R(e^R)$ with the right-invariant error $e^R := \hat{\chi}\chi^{-1}$. This explains our use of superscripts 'L' and 'R'.

B. Group-theoretic Conditions for LTV Immersion

Unlike invariant [5], [6] or equivariant [16], [17], [39] filters, which linearize the system in exponential (normal) coordinates and thus preclude global convergence, we avoid linearization through a powerful construction allowing immersion of the original system on the Lie group into an auxiliary high-dimensional linear system by delving deeper into the fine structures of the automorphism group.

Definition 1: A system $\dot{\chi} = f_u(\chi),\ y = h(\chi)$ on a Lie group G with output $y \in \mathbb{R}^{d_y}$ is said to be immersible [40], [41] into an LTV system $\dot{z} = F_u z + C_u,\ y = H_t z$ on $\mathbb{R}^{d_z}$ with $F_u \in \mathbb{R}^{d_z \times d_z}, C_u \in \mathbb{R}^{d_z}, H_t \in \mathbb{R}^{d_y \times d_z}$, if there exists a smooth map $\pi : G \to \mathbb{R}^{d_z}$ such that for every input u taking values in $\mathbb{R}^{\dim \mathfrak{g}}$ and every initial value $\chi_0 \in G$ from any initial time t_0,

$$H_t \phi_u^{\mathbb{R}^{d_z}}(t; \pi(\chi_0), t_0) = h \circ \phi_u^G(t; \chi_0, t_0) \tag{17}$$

holds for every t, where $\phi_u^{\mathbb{R}^{d_z}} : \mathbb{R} \times \mathbb{R}^{d_z} \times \mathbb{R} \to \mathbb{R}^{d_z}$ and $\phi_u^G : \mathbb{R} \times G \times \mathbb{R} \to G$ are the flows of the LTV system and the system on the Lie group G emanating from $(\pi(\chi_0), t_0)$ and (χ_0, t_0) under the same input u.

The existence of an immersed system is based on a finite termination criterion when iteratively computing the Lie derivative $\mathcal{L}_f h$ of the output along the system dynamics [40]. As group-affine dynamics is closely related to automorphisms of G as shown in Proposition 1, we must provide explicit characterizations of $\mathfrak{aut}(G)$ before we proceed. We start with a special class of automorphisms, called the inner automorphisms, to give a taste of the mechanism.

The inner-automorphisms form a subgroup of $\mathrm{Aut}(G)$ in the form of conjugates as

$$\mathrm{Inn}(G) := \left\{ \phi_{\chi_0} \,\middle|\, \phi_{\chi_0}(\chi) = \chi_0 \chi \chi_0^{-1}, \forall \chi_0, \chi \in G \right\}. \tag{18}$$

As before, fix s, t and impose $\phi_{\chi_0}(s; \cdot, t) \in \mathrm{Inn}(G)$. Define a vector field $g_{\chi_0} : G \to TG$ by

$$g_{\chi_0}(\chi) := \frac{d}{ds}\,|_{s=t} \phi_{\chi_0}(s; \chi, t)\ . \tag{19}$$

Note that χ_0 depends smoothly on s in (19) and $\chi_0(t) = id_G$. We obtain an explicit formula for $g_{\chi_0}(\chi)$ as

$$g_{\chi_0}(\chi) = \left(\frac{d\chi_0}{ds} \chi - \chi \frac{d\chi_0}{ds} \right)\Bigg|_{s=t}. \tag{20}$$

As $d\chi_0/ds \in \mathfrak{g}$ at $s = t$, the set of vector fields corresponding to $\mathrm{Inn}(G)$ is denoted by

$$\mathfrak{inn}(G) := \{ g | g(\chi) = A\chi - \chi A, \forall \chi \in G, \forall A \in \mathfrak{g} \}, \tag{21}$$

as a subset of $\Gamma(TG)$. The $\mathfrak{inn}(G)$ is the Lie algebra of $\mathrm{Inn}(G)$ and also a Lie subalgebra of $\mathfrak{aut}(G)$. It is immediate that $\mathfrak{inn}(G) \subset L(G) + R(G)$. Following Proposition 1, we see that for any fixed time, $\mathfrak{inn}(G) + L(G)$ or $\mathfrak{inn}(G) + R(G)$ characterizes exactly bi-invariant systems [5]. For group-affine dynamics corresponding to $\mathrm{Inn}(G)$, A in (21) may depend on u, and thus implicitly depends on t. This will create undesirable derivatives preventing the establishment of finite termination when conducting $\mathcal{L}_f h$. We further focus on a special class of group-affine dynamics related to $\mathrm{Inn}(G)$: for any fixed $g_A(\chi) = A\chi - \chi A \in \mathfrak{inn}(G)$ independent of time, let's define

$$g_A + L(G) := \left\{ g_A + \tilde{f} \in \Gamma(TG) \,\middle|\, \tilde{f} \in L(G) \right\}, \tag{22}$$

$$g_A + R(G) := \left\{ g_A + \tilde{f} \in \Gamma(TG) \,\middle|\, \tilde{f} \in R(G) \right\}. \tag{23}$$

Note that for group-affine dynamics corresponding to $g_A + L(G)$ or $g_A + R(G)$, the dependence on input u (and hence on t) can only appear in the $L(G)$ or $R(G)$ component.

Proposition 2: Consider a linear observed system on a Lie group $G \subset \mathbb{R}^{d_y \times d_y}$: $\dot{\chi} = f_u(\chi),\ y = h(\chi)$ with the state $\chi \in G$, the input $u \in \mathbb{R}^{\dim \mathfrak{g}}$ and the output $y \in \mathbb{R}^{d_y}$. Using notations in (22) and (23), if the system structures f_u and h satisfy one of the following conditions:

- Case 1: $f_u \in g_A + L(G),\ h(\chi) = \chi^{-1} d$,
- Case 2: $f_u \in g_A + R(G),\ h(\chi) = \chi d$,

where $d \in \mathbb{R}^{d_y}$ is a constant vector. This linear observed system on G is immersible into an LTV system. The immersion map π and the LTV system are shown below.

- For Case 1, we have $\pi(\chi) = z := [z_0^\top, \ldots, z_{d_y-1}^\top]^\top \in \mathbb{R}^{d_y^2}$, whose row blocks are

$$z_j := \chi^{-1}(A^j d) \in \mathbb{R}^{d_y},\ j \in [0, d_y - 1]. \tag{24}$$

  There exist d_y constants $a_l \in \mathbb{R}, l = 0, \ldots, d_y - 1$ such that the dynamics governing z is

$$\dot{z}_j = -(f_u(id) - A) z_j - z_{j+1},\ j \in [0, d_y - 2], \tag{25}$$

$$\dot{z}_{d_y-1} = -(f_u(id) - A) z_{d_y-1} - \sum_{l=0}^{d_y-1} a_l z_l, \tag{26}$$

  with the measurement equation $y = z_0$.
- For Case 2, we have $\pi(\chi) = z := [z_0^\top, \ldots, z_{d_y-1}^\top]^\top \in \mathbb{R}^{d_y^2}$, whose row blocks are

$$z_j := \chi(A^j d) \in \mathbb{R}^{d_y},\ j \in [0, d_y - 1]. \tag{27}$$

  The same constants a_l and measurement equation $y = z_0$ as in Case 1 are used. The dynamics governing z is

$$\dot{z}_j = (f_u(id) + A) z_j - z_{j+1},\ j \in [0, d_y - 2], \tag{28}$$

$$\dot{z}_{d_y-1} = (f_u(id) + A) z_{d_y-1} - \sum_{l=0}^{d_y-1} a_l z_l. \tag{29}$$

Proof: See Appendix B. ■

The explicit formula of $\mathfrak{inn}(G)$ is crucial for the immersion. Moreover, the time-independence of $g_A \in \mathfrak{inn}(G)$ is essential to establish the finite termination. Unfortunately, Cases 1 and 2 in Proposition 2 cover only a class of bi-invariant systems, far from sufficient for potential applications.

This compels us to consider group-affine f_us corresponding to automorphisms beyond $\mathrm{Inn}(G)$, which are highly non-trivial. $\mathrm{Inn}(G)$ is a normal subgroup of $\mathrm{Aut}(G)$, implying that the quotient $\mathrm{Out}(G) := \mathrm{Aut}(G)/\mathrm{Inn}(G)$ is a well-defined abstract group, not necessarily a continuous group even if G is a Lie group. In general, it's impossible to obtain explicit formulas for $\mathrm{Out}(G)$. Luckily, such obstruction can be overcome through group embedding.

Theorem 1: Let $G \subset \mathrm{GL}(d_y, \mathbb{R})$ be a matrix Lie group. If there exists a bigger matrix Lie group $\tilde{G} \subset \mathrm{GL}(d_y, \mathbb{R})$, such that $G \subset \tilde{G}$ is a non-trivial normal subgroup of $\tilde{G}$. Let the Lie algebra of $\tilde{G}$ be $\tilde{\mathfrak{g}} \subset \mathbb{R}^{d_y \times d_y}$ containing $\mathfrak{g} \subset \mathbb{R}^{d_y \times d_y}$. Define the restriction of the vector field $\mathfrak{inn}(\tilde{G}) \in \Gamma(T\tilde{G})$ to a vector field $\Gamma(TG)$ on G as

$$\mathfrak{inn}(\tilde{G})|_{\Gamma(TG)} := \left\{ g \,\middle|\, g(\chi) = \tilde{A}\chi - \chi\tilde{A}, \forall \chi \in G, \forall \tilde{A} \in \tilde{\mathfrak{g}} \right\}.$$

Then $f_u \in \mathfrak{inn}(\tilde{G})|_{\Gamma(TG)} + L(G) \supset \mathfrak{inn}(G) + L(G)$ or $f_u \in \mathfrak{inn}(\tilde{G})|_{\Gamma(TG)} + R(G) \supset \mathfrak{inn}(G) + R(G)$ defines group-affine dynamics. Moreover, we fix $g_{\tilde{A}} \in \mathfrak{inn}(\tilde{G})|_{\Gamma(TG)}$ with $\tilde{A} \in \tilde{\mathfrak{g}}$ and define

$$g_{\tilde{A}} + L(G) := \left\{ g_{\tilde{A}} + \tilde{f} \in \Gamma(TG) \,\middle|\, \tilde{f} \in L(G) \right\}, \tag{30}$$

$$g_{\tilde{A}} + R(G) := \left\{ g_{\tilde{A}} + \tilde{f} \in \Gamma(TG) \,\middle|\, \tilde{f} \in R(G) \right\}. \tag{31}$$

Let a linear observed system be given by $\dot{\chi} = f_u(\chi),\ y = h(\chi)$ with the state $\chi \in G$, the input $u \in \mathbb{R}^{\dim \mathfrak{g}}$ and the output $y \in \mathbb{R}^{d_y}$. If the system structures f_u and h satisfy one of the following conditions:

- Case 1: $f_u \in g_{\tilde{A}} + L(G),\ h(\chi) = \chi^{-1}d$ corresponding to the linear action of $\tilde{G}$ on $\mathbb{R}^{d_y}$,
- Case 2: $f_u \in g_{\tilde{A}} + R(G),\ h(\chi) = \chi d$ corresponding to the linear action of $\tilde{G}$ on $\mathbb{R}^{d_y}$,

where $d \in \mathbb{R}^{d_y}$ is a constant vector, then the system on group G is immersible into an LTV system. Additionally, the immersion map π and the immersed LTV system match those in (24)–(26) for Case 1 and (27)–(29) for Case 2, if we merely substitute $A \in \mathfrak{g}$ by $\tilde{A} \in \tilde{\mathfrak{g}}$. Furthermore, the constant a_js are adapted accordingly.

Proof: See Appendix C. ■

Remark 2: It's important to emphasize that Theorem 1 covers more than bi-invariant systems on G. Let $\tilde{A} \in \tilde{\mathfrak{g}}$ while $\tilde{A} \notin \mathfrak{g}$. For $\chi \in G$, $g_{\tilde{A}}(\chi) := \tilde{A}\chi - \chi\tilde{A}$ is in $\mathfrak{aut}(G)$ but certainly not in $\mathfrak{inn}(G)$. This means the induced group-affine dynamics $f_u \in g_{\tilde{A}} + L(G)$ or $f_u \in g_{\tilde{A}} + R(G)$ correspond to non-trivial outer automorphisms of G. Theorem 1 is a genuine generalization of Proposition 2 beyond bi-invariant systems on G. Luckily, the class of group-affine systems in Theorem 1 is sufficient for interesting applications, which serves as our system model for global observer design.

Remark 3: To guarantee observability, we have to tackle a system with multiple measurements, e.g., $y^{(i)} = \chi^{-1}d^{(i)} \in \mathbb{R}^{d_y}, i = 1, \ldots, M$. Then we should repeatedly conduct the immersion M times for each measurement, creating M copies of immersed states $z^{(i)} := [z_0^{(i)\top}, \ldots, z_{d_y-1}^{(i)\top}]^\top \in \mathbb{R}^{d_y^2}$. Stacking these states $z^{(i)}$ together, we obtain an immersed LTV on $\mathbb{R}^{Md_y^2}$. More precisely, in Case 1 for $1 \le i \le M$, the immersed system is given by

$$\dot{z}_j^{(i)} = -(f_u(id) - \tilde{A})z_j^{(i)} - z_{j+1}^{(i)},\ j \in [0, d_y - 2], \tag{32}$$

$$\dot{z}_{d_y-1}^{(i)} = -(f_u(id) - \tilde{A})z_{d_y-1}^{(i)} - \sum_{l=0}^{d_y-1} \tilde{a}_l z_l^{(i)}, \tag{33}$$

with M copies $\mathbb{R}^{d_y}$-valued multiple measurements $y^{(i)} = z_0^{(i)}, i = 1, \ldots, M$. In Case 2 for $1 \le i \le M$, the LTV is given by

$$\dot{z}_j^{(i)} = (f_u(id) + \tilde{A})z_j^{(i)} - z_{j+1}^{(i)},\ j \in [0, d_y - 2], \tag{34}$$

$$\dot{z}_{d_y-1}^{(i)} = (f_u(id) + \tilde{A})z_{d_y-1}^{(i)} - \sum_{l=0}^{d_y-1} \tilde{a}_l z_l^{(i)}, \tag{35}$$

with measurements $y^{(i)} = z_0^{(i)}, i = 1, \ldots, M$.

Remark 4: In practice, one often considers the joint estimation of an unknown constant input bias $b \in \mathbb{R}^{\dim \mathfrak{g}}$. Involving b as part of the extended state in $G \times \mathfrak{g}$ destroys the group-affine property and leads to imperfect IEKFs [4]. Despite efforts to modify the filter errors [4], [42], [43], local stability guarantee is not proved with unknown input bias. Nevertheless, our immersion still works with input bias b as we add an additional equation $\dot{b} = 0$ and substitute $f_u(id)$ by $f_{u+b}(id)$ in the immersed LTV system, shedding light on this issue. The equations governing the extended state (z, b) evolving within the vector space $\mathbb{R}^{d_y^2} \times \mathbb{R}^{\dim \mathfrak{g}}$ become nonlinear.

IV. Global Observer Design Using LTV Immersion

We focus on the immersible linear observed systems on Lie groups in Theorem 1. A global observer is constructed by first designing an observer for the immersed LTV, and then reconstructing the state in the Lie group using the estimation of the immersed state. Hereafter, our system model incorporates multiple measurements to ensure observability, similar to the formulation in [5].

A. Observer Structure from the Immersion

Consider a linear observed system on group G as $\dot{\chi} = f_u(\chi),\ y^{(i)} = h^{(i)}(\chi)$ with process dynamics $f_u(\chi)$ and multiple measurements $h^{(i)}(\chi), i = 1, \ldots, M$ verifying one of the cases in Theorem 1. We now define an observer based on this immersion.

Definition 2: An observer for the linear observed system with the above properties is a pair $(\mathcal{F}, \mathcal{T})$, where $\mathcal{F} : \mathbb{R}^{d_z} \times \mathbb{R}^{\dim \mathfrak{g}} \times \mathbb{R}^{Md_y} \to \mathbb{R}^{d_z}$ is a dynamical system on $\mathbb{R}^{d_z}$, which serves as an observer given by $\dot{\hat{z}} = \mathcal{F}(\hat{z}, u, y)$ for the immersed system, and $\mathcal{T} : \mathbb{R}^{d_z} \to G$ is a map that reconstructs the estimate $\hat{\chi} = \mathcal{T}(\hat{z})$ on the group G from the estimate of the immersed state $\hat{z}$. Note that the second and third arguments of $\mathcal{F}$ correspond to the $\mathbb{R}^{\dim \mathfrak{g}}$-valued input and M different $\mathbb{R}^{d_y}$-valued measurements of the original system on G.

We now detail the structure of $\mathcal{F}$ and $\mathcal{T}$. Assume $f_u(\chi) \in g_{\tilde{A}} + L(G)$ and $h^{(i)}(\chi) = \chi^{-1}d^{(i)} \in \mathbb{R}^{d_y}, i = 1, \ldots, M$ as in Case 1 of Theorem 1, hence the immersed LTV system is written in a compact form

$$\dot{z} = F_u z,\ y = Hz, \tag{36}$$

with the immersed LTV state defined by $z := [z^{(1)\top}, \ldots, z^{(M)\top}]^\top \in \mathbb{R}^{d_z}$. Each state component $z^{(i)}$ corresponds to the i-th measurement and is given by $z^{(i)} := [z_0^{(i)\top}, \ldots, z_{d_y-1}^{(i)\top}]^\top \in \mathbb{R}^{d_y^2}$. Hence $d_z = Md_y^2$ and $z_j^{(i)} \in \mathbb{R}^{d_y}, 1 \le i \le M, 0 \le j \le d_y - 1$. The measurement y is $[y^{(1)\top}, \ldots, y^{(M)\top}]^\top \in \mathbb{R}^{Md_y}$. The matrices F_u and H are obtained by stacking the immersed states and measurements as

$$F_u := \operatorname{diag}(F_u^{(1)}, \dots, F_u^{(M)}) \in \mathbb{R}^{d_z \times d_z}, \tag{37}$$

$$H := \operatorname{diag}(H^{(1)}, \dots, H^{(M)}) \in \mathbb{R}^{Md_y \times d_z}, \tag{38}$$

where each block is defined by

$$F_u^{(i)} := \begin{bmatrix} -S_u & -I & 0 & \cdots & 0 \\ 0 & -S_u & -I & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\tilde{a}_0 I & -\tilde{a}_1 I & -\tilde{a}_2 I & \cdots & -S_u - \tilde{a}_{d_y-1} I \end{bmatrix}, \tag{39}$$

$$H^{(i)} := \begin{bmatrix} I & 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{40}$$

with each component being of size $\mathbb{R}^{d_y \times d_y}$ and

$$S_u := f_u(id) - \tilde{A} \in \tilde{\mathfrak{g}} \subset \mathbb{R}^{d_y \times d_y}. \tag{41}$$

A linear observer $\mathcal{F}$ for (36) with an estimated state $\hat{z} \in \mathbb{R}^{d_z}$ and a gain matrix $K \in \mathbb{R}^{d_z \times Md_y}$ is designed as a continuous time Kalman filter

$$\dot{\hat{z}} = F_u \hat{z} + K(y - H\hat{z}), \tag{42}$$

where the variable gain $K := PH^\top R^{-1}$ is computed using a matrix Riccati equation

$$\dot{P} = F_u P + P F_u^\top + Q - PH^\top R^{-1} HP, \tag{43}$$

where $P, Q \in \mathbb{R}^{d_z \times d_z}$ and $R \in \mathbb{R}^{Md_y \times Md_y}$. Note that Q and R are viewed as positive definite tunable gains, often interpreted as the covariances of the process and measurement noises respectively.

Since $d_z > \dim(G)$, directly pulling back the observer $\mathcal{F}$ by $\mathcal{T}$ from $\mathbb{R}^{d_z}$ to a dynamical system on G is troublesome, and thus reconstruction of the estimation $\hat{\chi} \in G$ is formulated as solving an optimization problem [15]. The reconstruction map $\hat{\chi} = \mathcal{T}(\hat{z})$ is the left-inverse of the immersion π as

$$\mathcal{T} : \mathbb{R}^{d_z} \to G, \; \hat{z} \mapsto \operatorname*{arg\,min}_{\hat{\chi} \in G} \|\hat{z} - \pi(\hat{\chi})\|^2. \tag{44}$$

By definition of π, we have $z_j^{(i)} = \chi^{-1}(\tilde{A}^j d^{(i)})$. For simplicity, we define

$$d_j^{(i)} := \tilde{A}^j d^{(i)} \in \mathbb{R}^{d_y}. \tag{45}$$

Hence, the reconstruction process is explicitly reformulated as

$$\hat{\chi} = \mathcal{T}(\hat{z}) = \operatorname*{arg\,min}_{\hat{\chi} \in G} \left\| \hat{Z} - \hat{\chi}^{-1} D \right\|_\Sigma^2, \tag{46}$$

with the matrices $\hat{Z}, D \in \mathbb{R}^{d_y \times Md_y}$ given by

$$\hat{Z} := \left[\hat{z}_0^{(1)}, ..., \hat{z}_{d_y-1}^{(1)}, ..., \hat{z}_j^{(i)}, ..., \hat{z}_0^{(M)}, ..., \hat{z}_{d_y-1}^{(M)}\right], \tag{47}$$

$$D := \left[d_0^{(1)}, ..., d_{d_y-1}^{(1)}, ..., d_j^{(i)}, ..., d_0^{(M)}, ..., d_{d_y-1}^{(M)}\right]. \tag{48}$$

Note that the index i ranges from $[1, M]$ and j ranges from $[0, d_y - 1]$. The optimization problem (46) can be weighted by positive definite matrix $\Sigma \in \mathbb{S}_+^{Md_y}$.

To summarize, the observer for a linear observed system on G with multiple measurements in Case 1 of Theorem 1 is a cascade of the optimization-based state reconstruction (46) after the linear Kalman filter (42)–(43), with matrix blocks defined by (37)–(38), (39)–(40) and (47)–(48).

Similar reasoning applies to the systems in Case 2 of Theorem 1, the observer has the same structure as in case 1, except the S_u in the diagonal blocks in $F_u^{(i)}$ is

$$S_u := -f_u(id) - \tilde{A} \in \tilde{\mathfrak{g}} \subset \mathbb{R}^{d_y \times d_y}, \tag{49}$$

compared to the diagonal block (41) in (39). Moreover, the optimization formulation is slightly different compared with (46):

$$\hat{\chi} = \mathcal{T}(\hat{z}) = \operatorname*{arg\,min}_{\hat{\chi} \in G} \left\| \hat{Z} - \hat{\chi} D \right\|_\Sigma^2. \tag{50}$$

As the constant weighting matrix Σ can be absorbed into the norm by redefining $\hat{Z}$ and D as $\hat{Z}\Sigma^{-\frac{1}{2}}$ and $D\Sigma^{-\frac{1}{2}}$ respectively. We omit Σ for notational simplicity.

Remark 5: Solving the optimization problem defined in (46) or (50) is a non-trivial task. The cost function is generally non-convex and has multiple critical points due to the non-Euclidean topology of the state group G [44]. Luckily, for a large class of applications on two-frame groups, the unique global minimum can be found by explicit formulas through generalizing the Umeyama techniques in [45]. In general, one imposes regularity conditions to simplify the cost function landscape (46) or (50), i.e., a perfect Morse function [44] whose critical points are non-degenerate. If, in addition, the global minimum is unique, any algorithm with the ability to escape a saddle point or local maximum will converge to this global minimum. Hence, we do not specify a particular optimization algorithm for reconstructing $\hat{\chi}$. Instead, we abstractly require $\mathcal{T}$ provides one global optimum.

B. Joint Estimation of Lie-Algebra-Valued Input Biases

If the input of the system on G is corrupted by a constant $\mathfrak{g}$-valued bias, the feasibility of the system immersion into $\mathbb{R}^{d_z}$ is preserved, and such bias may be simultaneously estimated through an observer for the immersed system with an extended bias state. This introduces multiplicative nonlinearities to the immersed system from the coupling of the bias and the original states $z_j^{(i)}$. Let $d_b := \dim \mathfrak{g}$.

To address bias estimation, we need additional structures to characterize the dependence of $f_u(id)$ on the input bias $b \in \mathbb{R}^{d_b}$.

Assumption 1: The $\mathfrak{g}$-valued $f_{u+b}(id)$ is additive with respect to the input bias $b \in \mathbb{R}^{d_b}$, i.e., $f_{u+b}(id) = f_u(id) + \mathcal{L}_\mathfrak{g}(b)$.

For Case 1 in Theorem 1, the immersed system with bias $b \in \mathbb{R}^{d_b}$ writes

$$\dot{z} = F_u z - F_b z, \; \dot{b} = 0, \; y = Hz, \tag{51}$$

where F_u, H are the same as (37) and (38) respectively. F_b is given by

$$F_b = \operatorname{diag}(\mathcal{L}_\mathfrak{g}(b), ..., \mathcal{L}_\mathfrak{g}(b)) \in \mathbb{R}^{d_z \times d_z}, \tag{52}$$

with Md_y-many $\mathcal{L}_\mathfrak{g}(b)$ blocks. An observer for such system is given by

$$\dot{\hat{z}} = F_u \hat{z} - F_{\hat{b}} \hat{z} + K_z(y - H\hat{z}), \tag{53}$$

$$\dot{\hat{b}} = K_b(y - H\hat{z}), \tag{54}$$

with gain matrix $K = [K_z^\top, K_b^\top]^\top \in \mathbb{R}^{(d_z+d_b)\times Md_y}$, whose blocks are $K_z \in \mathbb{R}^{d_z\times Md_y}, K_b \in \mathbb{R}^{d_b\times Md_y}$. Inspired by the approach in [35], we choose to compute the variable gain $K := P\breve{H}^\top R^{-1}$ via a modified Riccati equation as

$$\dot{P} = \lambda P + \breve{F}P + P\breve{F}^\top - P\breve{H}^\top R^{-1}\breve{H}P, \tag{55}$$

to address the nonlinearities arising from the multiplication $F_{\hat{b}}\hat{z}$, aiming to break the coupling loop between gain computation and state estimation, which is a Kalman-like observer. The positive constant λ serves as a forgetting factor. The matrix $\breve{F}$ is derived by linearizing (53)–(54) with respect to the extended state $[\hat{z}^\top, \hat{b}^\top]^\top$. $\breve{F}$ and $\breve{H}$ are given by

$$\breve{F} := \begin{bmatrix} F_u - F_{\hat{b}} & -J_{\hat{z}} \\ 0_{d_b\times d_z} & 0_{d_b\times d_b} \end{bmatrix}, \; \breve{H} := \begin{bmatrix} H & 0_{d_z\times d_b} \end{bmatrix}. \tag{56}$$

$-J_z \in \mathbb{R}^{d_z\times d_b}$ is the Jacobian with respect to $\hat{b}$ resulting from $F_{\hat{b}}\hat{z}$. Define $\mathcal{L}^\dagger(\hat{z}_j^{(i)}) \in \mathbb{R}^{d_y\times d_b}$ to be the linear operator such that $\mathcal{L}^\dagger(\hat{z}_j^{(i)})\hat{b} = \mathcal{L}_{\mathfrak{g}}(\hat{b})\hat{z}_j^{(i)}$. $J_{\hat{z}} \in \mathbb{R}^{d_z\times d_b}$ is given by

$$J_{\hat{z}} = \left[\mathcal{L}^\dagger(\hat{z}_0^{(1)})^\top, ..., \mathcal{L}^\dagger(\hat{z}_j^{(i)})^\top, ..., \mathcal{L}^\dagger(\hat{z}_{d_y-1}^{(M)})^\top\right]^\top. \tag{57}$$

The subsequent reconstruction of the G-valued state is exactly the same as in the non-biased Case 1 shown in (46) using the estimation $\hat{z}$ of the above observer (53)–(55).

For systems in Case 2 of Theorem 1, the same reasoning applies and thus the immersed system with bias $b \in \mathbb{R}^{d_b}$ writes

$$\dot{z} = F_u z + F_b z, \; \dot{b} = 0, \; y = Hz, \tag{58}$$

with F_u composed of blocks S_u as (49). The H, F_b are the same as those in (51). Similarly, a Kalman-like observer for systems in Case 2 reads

$$\dot{\hat{z}} = F_u\hat{z} + F_{\hat{b}}\hat{z} + K_z(y - H\hat{z}), \tag{59}$$

$$\dot{\hat{b}} = K_b(y - H\hat{z}), \tag{60}$$

tuned using (55) with the only difference being $\breve{F}$ as

$$\breve{F} := \begin{bmatrix} F_u + F_{\hat{b}} & J_{\hat{z}} \\ 0_{d_b\times d_z} & 0_{d_b\times d_b} \end{bmatrix}, \tag{61}$$

compared to (56). The subsequent reconstruction of the G-valued state is the same as in the non-biased Case 2 shown in (50) using the estimation $\hat{z}$ from the observer (59)–(60).

C. Global Properties of the Bias-free Observer

The proposed observer framework, as a cascade of an optimization after a Kalman observer for the immersed system on $\mathbb{R}^{d_z}$, will possess global stability if the optimization algorithm can achieve one global minimum under certain regularity conditions while the Kalman observer is globally exponentially stable. We first analyze the uniform observability of the immersed LTV system.

Assumption 2: There exist constants $\alpha, \delta > 0$, such that

$$\Phi^\top(t+\delta, t)\Phi(t+\delta, t) \succeq \alpha I, \; \forall t \in \mathbb{R}, \tag{62}$$

where the transition matrix $\Phi(t_2, t_1) \in \mathbb{R}^{d_y\times d_y}$ is defined by $\frac{\partial\Phi(t_2,t_1)}{\partial t_2} = -S_u\Phi(t_2, t_1)$ and $\Phi(t_2, t_2) = I$ for all t_1, t_2 and S_u given by (41) for Case 1 (respectively, (49) for Case 2).

Lemma 1: Under Assumption 2, the pair (F_u, H) in (37)–(38) with blocks defined by (39)-(40) is uniformly observable.

Proof: See Appendix D. ∎

Remark 6: Many practical systems, e.g., the two-frame systems [4], automatically satisfy Assumption 2 due to state group structure as shown later. Hence, the associated Kalman observer for the immersed LTV is automatically uniformly observable. The assumption is considered very weak.

Assumption 3: The D defined in (48) satisfies rank$(D) \geq d_y$. Note that d_y is both the dimension of a single vector-valued output and the dimension of the square matrix into which G embeds.

Remark 7: Only $d_0^{(i)}, 1 \leq i \leq M$ are related to physical measurements. $d_j^{(i)}, j \geq 1, 1 \leq i \leq M$ are virtual measurements generated by multiplying $\tilde{A}$, i.e., the automorphism structure. For a suitable choice of $\tilde{A}$, fewer than d_y measurements suffice to satisfy the assumption. As G is an invertible square matrix, $Z = \chi^{-1}D$ or $Z = \chi D$ has a unique solution χ determined by (Z, D) if rank$(D) \geq d_y$. This implies that $\min_{\chi\in G}\|Z - \chi^{-1}D\|^2$ or $\min_{\chi\in G}\|Z - \chi D\|^2$ achieves a unique global minimum of 0 under the rank condition. Hence, the true state χ can be uniquely determined by the true value of the immersed system state.

Definition 3: Let $\hat{\chi}$ and χ denote the estimated and true states on G, respectively. As G is identified as a closed subgroup of GL$(d_y, \mathbb{R})$, an extrinsic metric on G is defined as $d(\hat{\chi}, \chi) = \|\chi^{-1} - \hat{\chi}^{-1}\|$ or $d(\hat{\chi}, \chi) = \|\chi - \hat{\chi}\|$.

Remark 8: The former metric is used in the stability analysis for systems satisfying Case 1, while the latter is used for systems satisfying Case 2.

Theorem 2: Under Assumption 2 and 3, provided we can find an optimization algorithm to implement $\mathcal{T}$, the observer (46) cascaded after (42)–(43) with blocks defined in (41) for a linear observed system on G that satisfies Case 1 in Theorem 1, or the observer (50) cascaded after (42)–(43) with blocks defined in (49) for a system that satisfies Case 2 in Theorem 1, is globally exponentially stable with respect to the corresponding metric defined above.

Proof: See Appendix E. ∎

Remark 9: We equate the effort required to design a GES observer for immersible linear observed systems on groups to that of solving an optimization problem on such groups, provided the system structure satisfies the rank condition. Our observer overcomes the topological obstructions [20] to achieve GES through system immersion. The topological difficulties influence the landscape of the cost function (46) or (50), where undesirable critical points, such as saddles, arise due to the compact part of G [44].

Remark 10: The uniqueness of the optimum $\hat{\chi}$ of (46) or (50) is not required in the proof for each $\hat{Z}$. We only require that $\mathcal{T}$ outputs one G-valued estimate that achieves global minimum. The regularity condition in Assumption 3 automatically guarantees that $\mathcal{T}(\hat{Z})$ is the unique optimum when $\hat{Z}$ is sufficiently close to Z.

Remark 11: For the two-frame systems discussed later in Section V, an explicit solution of a global minimum for (46) or (50) can be obtained, thereby enabling global observer realizations as per Theorem 2.

D. Semi-global Properties of the Observer with Unknown Constant Input Bias

With biases, the immersion still works, but the immersed system is no longer linear, and the Jacobians of the extended Kalman filters depend on the estimated state $\hat{z}$. The stability of the extended Kalman filter posed on this system relies on a uniform boundedness hypothesis on the minimum and maximum eigenvalue of P in (43), which in turn depends on the estimated state $\hat{z}$ through the Jacobian, making it unverifiable. [35], [37] breaks the dependency loop by modifying the gain computation as (55), which has an explicit solution through integration. Boundedness of P can then be guaranteed by persistent determinability. Inspired by [35], we present the following non-local results with joint bias estimation.

Assumption 4: The true trajectory on the group evolves in a compact subset $\mathcal{G}_1$ of G.

Assumption 5: The pair $(\breve{F}, \breve{H})$ defined in (56) for system in Case 1 (respectively, (61) for system in Case 2) is persistently determinable.

Remark 12: Verifying Assumption 5 a priori is difficult as it depends on the estimated state. However, monitoring the minimum eigenvalue of determinability Gramian $\mathcal{D}(t, t-\delta)$ in a moving-horizon fashion numerically is possible thanks to its dependence on the estimated state instead of on the true state. Moreover, active choice of trajectories should be considered to maximize $\mathcal{D}(t, t-\delta)$ to improve observer performance [34], [46].

Theorem 3: With unknown constant input bias, let us consider the observer (46) cascaded after (53), (54) and (55) with blocks defined by (56) for systems in Case 1 of Theorem 1 (respectively, the observer (50) cascaded after (59), (60) and (55) with blocks defined by (61) for systems in Case 2). Under Assumption 3, 4 and 5, there exist a compact subset $\mathcal{G}_2$ of G and a compact subset $\hat{\mathcal{B}}$ of $\mathbb{R}^{\dim \mathfrak{g}}$, such that for any initial value $\hat{\chi}(t_0) \in \mathrm{int}(\mathcal{G}_1)$ and $\hat{b}(t_0) \in \mathcal{B} \subset \hat{\mathcal{B}}$, the estimation $\hat{\chi}(t), \hat{b}(t)$ remain in $\mathcal{G}_2, \hat{\mathcal{B}}$ respectively, for $t \in [t_0, \infty)$. Moreover, $d(\hat{\chi}(t), \chi(t))$ and $\|\hat{b}(t) - b\|$ converge to 0 exponentially after some finite time.

Proof: See Appendix F. ∎

V. Application to Two-Frame Systems

We apply our observer design toolbox to two-frame systems, which are linear observed systems on two-frame groups constructed via the semi-direct product of a rotation group and several vectors. Two-frame systems are powerful tools to model a large class of navigation problems. Our theory provides unified GES observer solutions, compared to the local results achieved by the InEKF [4], [5] and case-by-case nonlinear constructive methods [23] for global results.

A. Immersible Two-Frame Systems

Definition 4: The two-frame group, denoted $\mathrm{TFG}(d, n, m)$ [4], where $d = 2$ or 3, $n, m \in \mathbb{N}$, is a matrix Lie group as the closed subgroup of $\mathrm{GL}(d+n+m, \mathbb{R})$ in the form of

$$\mathrm{TFG}(d,n,m) = \left\{ \begin{bmatrix} R & W \\ 0 & I \end{bmatrix} \middle| \begin{array}{l} R \in \mathrm{SO}(d), W = [X \ RY] \\ X \in \mathbb{R}^{d\times n}, Y \in \mathbb{R}^{d\times m} \end{array} \right\}.$$

The size of W is $\mathbb{R}^{d\times(n+m)}$. Each column of X or Y is an $\mathbb{R}^d$-valued vector. The two-frame group describes the rigid geometric transformation between two frames, serving as the state space in single rigid-body kinematics. In navigation problems, suppose R represents the rotation from the body frame to the world frame, the n columns of X are related to $\mathbb{R}^d$-valued states expressed in the world frame, in contrast to the m columns of Y expressed in the body frame. Its Lie algebra $\mathfrak{tfg}(d,n,m) \subset \mathbb{R}^{(d+n+m)\times(d+n+m)}$ is defined as

$$\mathfrak{tfg}(d,n,m) = \left\{ \begin{bmatrix} \omega^\times & \rho \\ 0 & 0 \end{bmatrix} \middle| \omega \in \mathbb{R}^{\frac{d(d-1)}{2}}, \rho \in \mathbb{R}^{d\times(n+m)} \right\}.$$

Definition 5: The extended similarity transformation group, denoted $\mathrm{SIM}_{n+m}(d)$ [18], where $d = 2$ or 3, $n, m \in \mathbb{N}$, is a matrix Lie group as the closed subgroup of $\mathrm{GL}(d+n+m, \mathbb{R})$ in the form of

$$\mathrm{SIM}_{n+m}(d) = \left\{ \begin{bmatrix} R & W \\ 0 & A \end{bmatrix} \middle| \begin{array}{l} R \in \mathrm{SO}(d), W \in \mathbb{R}^{d\times(n+m)} \\ A \in \mathrm{GL}(n+m, \mathbb{R}) \end{array} \right\}.$$

Its Lie algebra $\mathfrak{sim}_{n+m}(d)$ can be checked as

$$\mathfrak{sim}_{n+m}(d) = \left\{ \begin{bmatrix} \Omega^\times & \gamma \\ 0 & L \end{bmatrix} \middle| \begin{array}{l} \Omega \in \mathbb{R}^{\frac{d(d-1)}{2}}, \gamma \in \mathbb{R}^{d\times(n+m)} \\ L \in \mathbb{R}^{(n+m)\times(n+m)} \end{array} \right\}.$$

Direct calculation verifies that $\mathrm{TFG}(d,n,m)$ is a normal subgroup of $\mathrm{SIM}_{n+m}(d)$. This allows the explicit characterization of the structures of two-frame systems in Case 1 or 2 of Theorem 1. Denote the state to be estimated $T \in \mathrm{TFG}(d,n,m)$ with block components as

$$T = \begin{bmatrix} R & W \\ 0 & I \end{bmatrix} \in \mathrm{TFG}(d,n,m). \tag{63}$$

From (30), the dynamics in Case 1 of Theorem 1 is the restriction of $\mathfrak{inn}(\mathrm{SIM}_{n+m}(d))$ on $\mathrm{TFG}(d,n,m)$ plus a time-varying left-invariant vector field $L(\mathrm{TFG})$. Letting Ω, γ, $\tilde{\omega}$, and $\tilde{\rho}$ be matrix blocks of proper sizes, we write

$$\dot{T}_t = \begin{bmatrix} \Omega^\times & \gamma \\ 0 & L \end{bmatrix} T_t - T_t \begin{bmatrix} \Omega^\times & \gamma \\ 0 & L \end{bmatrix} + T_t \begin{bmatrix} \tilde{\omega}_t^\times & \tilde{\rho}_t \\ 0 & 0 \end{bmatrix}, \tag{64}$$

where we explicitly mark the dependence on t. Similarly by (31), the Case-2 dynamics on two-frame groups is given by

$$\dot{T}_t = \begin{bmatrix} \Omega^\times & \gamma \\ 0 & L \end{bmatrix} T_t - T_t \begin{bmatrix} \Omega^\times & \gamma \\ 0 & L \end{bmatrix} + \begin{bmatrix} \tilde{\omega}_t^\times & \tilde{\rho}_t \\ 0 & 0 \end{bmatrix} T_t. \tag{65}$$

For simplicity, we can combine terms and regard $\omega_t := \tilde{\omega}_t - \Omega \in \mathbb{R}^{\frac{d(d-1)}{2}}$ and $\rho_t := \tilde{\rho}_t - \gamma \in \mathbb{R}^{d\times(n+m)}$ as inputs. This leads to the characterization of systems in Case 1 as

$$\begin{cases} \dot{T}_t = \begin{bmatrix} \Omega^\times & \gamma \\ 0 & L \end{bmatrix} T_t + T_t \begin{bmatrix} \omega_t^\times & \rho_t \\ 0 & -L \end{bmatrix}, \\ y^{(i)} = T_t^{-1} d^{(i)}, \ i = 1, 2, ..., M \end{cases} \tag{66}$$

where the M constant vectors $d^{(i)} \in \mathbb{R}^{d+n+m}$ are with i indexing. Note that $y^{(i)} \in \mathbb{R}^{d+n+m}$. Similarly, defining inputs $\omega_t := \tilde{\omega}_t + \Omega$ and $\rho_t := \tilde{\rho}_t + \gamma$, the systems in Case 2 are given by

$$\begin{cases} \dot{T}_t = \begin{bmatrix} \omega_t^\times & \rho_t \\ 0 & L \end{bmatrix} T_t - T_t \begin{bmatrix} \Omega^\times & \gamma \\ 0 & L \end{bmatrix}, \\ y^{(i)} = T_t d^{(i)}, \ i = 1, 2, ..., M \end{cases} \tag{67}$$

where $y^{(i)}$ and $d^{(i)}$ are defined in a similar fashion. Let a constant element be

$$\tilde{A} = \begin{bmatrix} \Omega^\times & \gamma \\ 0 & L \end{bmatrix} \in \mathfrak{sim}_{n+m}(d). \tag{68}$$

Let $N = d + m + n$. Let $\pi : \mathrm{TFG}(d,n,m) \to \mathbb{R}^{d_z}$ be the immersion map for (66), i.e., $\pi(T) = z := [z_0^{(1)\top}, \ldots, z_{N-1}^{(M)\top}]^\top \in \mathbb{R}^{d_z}$ with each row block defined by $z_j^{(i)} = T^{-1}\tilde{A}^j d^{(i)}, 0 \le j \le N-1, 1 \le i \le M$. Hence, $d_z = MN^2$ with each $z_j^{(i)} \in \mathbb{R}^N$. The immersion for (66) is

$$\begin{cases} \dot{z}_j^{(i)} = -\begin{bmatrix} \omega_t^\times & \rho_t \\ 0 & -L \end{bmatrix} z_j^{(i)} - z_{j+1}^{(i)},\ j \in [0, N-2] \\ \dot{z}_{N-1}^{(i)} = -\begin{bmatrix} \omega_t^\times & \rho_t \\ 0 & -L \end{bmatrix} z_{N-1}^{(i)} - \sum_{l=0}^{N-1} \tilde{a}_l z_l^{(i)} \\ y^{(i)} = z_0^{(i)},\ i \in [1, M] \end{cases}, \tag{69}$$

where $\tilde{a}_j$s are derived from the operator equation $\tilde{A}^N = \sum_{l=0}^{N-1} \tilde{a}_l \tilde{A}^l$. The variables involved in (69) are homogeneous coordinates, and thus we are only interested in the first d coordinates of $z_j^{(i)}$ or $y^{(i)}$ in practice. Decompose the state and measurement into $z_j^{(i)} = [\bar{z}_j^{(i)\top}, \underline{z}_j^{(i)\top}]^\top$ and $y^{(i)} = [\bar{y}^{(i)\top}, \underline{y}^{(i)\top}]^\top$, where $\bar{z}_j^{(i)}, \bar{y}^{(i)} \in \mathbb{R}^d$ and $\underline{z}_j^{(i)}, \underline{y}^{(i)} \in \mathbb{R}^{n+m}$. (69) is divided into two sub-systems

$$\begin{cases} \dot{\bar{z}}_j^{(i)} = -\omega_t^\times \bar{z}_j - \rho_t \underline{z}_j^{(i)} - \bar{z}_{j+1}^{(i)},\ j \in [0, N-2] \\ \dot{\bar{z}}_{N-1}^{(i)} = -\omega_t^\times \bar{z}_{N-1}^{(i)} - \rho_t \underline{z}_{N-1}^{(i)} - \sum_{l=0}^{N-1} \tilde{a}_l \bar{z}_l^{(i)} \\ \bar{y}^{(i)} = \bar{z}_0^{(i)},\ i \in [1, M] \end{cases}, \tag{70}$$

$$\begin{cases} \dot{\underline{z}}_j^{(i)} = L\underline{z}_j^{(i)} - \underline{z}_{j+1}^{(i)},\ j \in [0, N-2] \\ \dot{\underline{z}}_{N-1}^{(i)} = L\underline{z}_{N-1}^{(i)} - \sum_{l=0}^{N-1} \tilde{a}_l \underline{z}_l^{(i)} \\ \underline{y}^{(i)} = \underline{z}_0^{(i)},\ i \in [1, M] \end{cases}, \tag{71}$$

where (70) is cascaded after (71).

In view of (45) and (68), we have the notation $d_j^{(i)} = \tilde{A}^j d^{(i)}$. If we further define $d_j^{(i)} = [\bar{d}_j^{(i)\top}, \underline{d}_j^{(i)\top}]^\top$, where $\bar{d}_j^{(i)} \in \mathbb{R}^d$ and $\underline{d}_j^{(i)} \in \mathbb{R}^{n+m}$, these constant components are inductively calculated by

$$\underline{d}_j^{(i)} = L^j \underline{d}^{(i)},\ i \in [1, M],\ j \in [0, N-1], \tag{72}$$

$$\bar{d}_{j+1}^{(i)} = \Omega^\times \bar{d}_j^{(i)} + \gamma \underline{d}_j^{(i)},\ i \in [1, M],\ j \in [0, N-2]. \tag{73}$$

As all underlined variables come from extending the physical coordinates to homogeneous coordinates, one knows a priori that $\underline{y}^{(i)} = \underline{z}_0^{(i)} = \underline{d}_0^{(i)} = \underline{d}^{(i)}$ are constants. Let the initial values of (71) at t_0 be $\underline{z}_j^{(i)}(t_0) = \underline{d}_j^{(i)} = L^j \underline{d}^{(i)}$, the sub-system (71) will remain constant as the right-hand side of (71) is zero for all $t \ge t_0$ by virtue of $L^N = \sum_{l=0}^{N-1} \tilde{a}_l L^l$ from the definition of $\tilde{a}_l$s. Hence, it is only necessary to design an observer for the sub-system (70). Recalling that $\underline{z}_j^{(i)} \equiv \underline{d}_j^{(i)}$ and substituting the underlined variables with the constants $\underline{d}_j^{(i)}$, the immersed system equation for (69) writes

$$\begin{cases} \dot{\bar{z}}_j^{(i)} = -\omega_t^\times \bar{z}_j^{(i)} - \rho_t \underline{d}_j^{(i)} - \bar{z}_{j+1}^{(i)},\ j \in [0, N-2] \\ \dot{\bar{z}}_{N-1}^{(i)} = -\omega_t^\times \bar{z}_{N-1}^{(i)} - \rho_t \underline{d}_{N-1}^{(i)} - \sum_{l=0}^{N-1} \tilde{a}_l \bar{z}_l^{(i)} \\ \bar{y}^{(i)} = \bar{z}_0^{(i)},\ i \in [1, M] \end{cases}. \tag{74}$$

The same rationale applies to systems in Case 2. Let $\pi : \mathrm{TFG}(d,n,m) \to \mathbb{R}^{d_z}$ be the immersion map for (67), i.e., $\pi(T) = z := [z_0^{(1)\top}, ..., z_{N-1}^{(M)\top}]^\top \in \mathbb{R}^{d_z}$ with each row block defined by $z_j^{(i)} = T\tilde{A}^j d^{(i)}$. The ranges of i, j and the dimensions d_z and N are the same as before. The immersion for (67) is

$$\begin{cases} \dot{z}_j^{(i)} = \begin{bmatrix} \omega_t^\times & \rho_t \\ 0 & L \end{bmatrix} z_j^{(i)} - z_{j+1}^{(i)},\ j \in [0, N-2] \\ \dot{z}_{N-1}^{(i)} = \begin{bmatrix} \omega_t^\times & \rho_t \\ 0 & L \end{bmatrix} z_{N-1}^{(i)} - \sum_{l=0}^{N-1} \tilde{a}_l z_l^{(i)} \\ y^{(i)} = z_0^{(i)},\ i \in [1, M] \end{cases}, \tag{75}$$

where the constants $\tilde{a}_j$ are the same as (69). Similarly, the states $z_j^{(i)}$, the measurements $y^{(i)}$ and the constants $d^{(i)}$ can be decomposed into components denoted by bars or underlines. Specifically, the components of $d_j^{(i)}$ are the same as (72)–(73). The immersion for systems in Case 2 is given by

$$\begin{cases} \dot{\bar{z}}_j^{(i)} = \omega_t^\times \bar{z}_j^{(i)} + \rho_t \underline{d}_j^{(i)} - \bar{z}_{j+1}^{(i)},\ j \in [0, N-2] \\ \dot{\bar{z}}_{N-1}^{(i)} = \omega_t^\times \bar{z}_{N-1}^{(i)} + \rho_t \underline{d}_{N-1}^{(i)} - \sum_{l=0}^{N-1} \tilde{a}_l \bar{z}_l^{(i)} \\ \bar{y}^{(i)} = \bar{z}_0^{(i)},\ i \in [1, M] \end{cases}, \tag{76}$$

with the underlined variables $\underline{z}_j^{(i)} \equiv \underline{d}_j^{(i)}$ being the same constants as before.

Kalman Observers in a unified form are implemented for (74) and (76). Let the state of the observer be $\hat{\bar{z}} := [\hat{\bar{z}}_0^{(1)\top}, ..., \hat{\bar{z}}_j^{(i)\top}, ..., \hat{\bar{z}}_{N-1}^{(M)\top}]^\top \in \mathbb{R}^{MNd}$ with $0 \le j \le N-1$ and $1 \le i \le M$. Each $\hat{\bar{z}}_j^{(i)}$ is in $\mathbb{R}^d$. Let $\bar{y} = [\bar{y}^{(1)\top}, ..., \bar{y}^{(M)\top}]^\top \in \mathbb{R}^{Md}$ be the stacked output. The observer equation is given by

$$\dot{\hat{\bar{z}}} = F_u \hat{\bar{z}} + C_u + K(\bar{y} - H\hat{\bar{z}}), \tag{77}$$

with the Kalman gain $K \in \mathbb{R}^{MNd \times Md}$ calculated using the Riccati equations (43) with the pair (F_u, H).

For systems in Case 1 and 2, H is derived from (74) and (76):

$$H = \mathrm{diag}(H^{(1)}, ..., H^{(M)}), \tag{78}$$

where all blocks are $H^{(i)} = [I_{d\times d}, 0_{d\times d(N-1)}], 1 \le i \le M$.

The (F_u, C_u) pair is composed of blocks as $F_u = \mathrm{diag}(F_u^{(1)}, \ldots, F_u^{(M)})$ and $C_u = [C_u^{(1)\top}, \ldots, C_u^{(M)\top}]^\top$ with $F_u^{(i)} \in \mathbb{R}^{Nd\times Nd}$ and $C_u^{(i)} \in \mathbb{R}^{Nd}$ corresponding to the i-th measurement. For systems in Case 1, these blocks are

$$F_u^{(i)} = \begin{bmatrix} -\omega_t^\times & -I & 0 & \cdots & 0 \\ 0 & -\omega_t^\times & -I & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\tilde{a}_0 I & -\tilde{a}_1 I & -\tilde{a}_2 I & \cdots & -\omega_t^\times - \tilde{a}_{N-1} I \end{bmatrix},$$

$$C_u^{(i)} = \left[-\left(\rho_t \underline{d}_0^{(i)}\right)^\top \quad \cdots \quad -\left(\rho_t \underline{d}_{N-1}^{(i)}\right)^\top \right]^\top,$$

by (74). Similarly by (76), the blocks for Case-2 systems are

$$F_u^{(i)} = \begin{bmatrix} \omega_t^\times & -I & 0 & \cdots & 0 \\ 0 & \omega_t^\times & -I & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\tilde{a}_0 I & -\tilde{a}_1 I & -\tilde{a}_2 I & \cdots & \omega_t^\times - \tilde{a}_{N-1} I \end{bmatrix},$$

$$C_u^{(i)} = \left[\left(\rho_t \underline{d}_0^{(i)}\right)^\top \quad \cdots \quad \left(\rho_t \underline{d}_{N-1}^{(i)}\right)^\top \right]^\top.$$

B. Two-frame Group State Reconstruction

We now formulate the state reconstruction procedure (46) or (50) for two-frame systems. Let $\hat{Z} = [\hat{\bar{Z}}^\top, \underline{Z}^\top]^\top$ and $D = [\bar{D}^\top, \underline{D}^\top]^\top$. Their components are from the state estimation of the immersed LTV, as

$$\hat{\bar{Z}} := \left[\hat{\bar{z}}_0^{(1)}, \ldots, \hat{\bar{z}}_{N-1}^{(1)}, \ldots, \hat{\bar{z}}_j^{(i)}, \ldots, \hat{\bar{z}}_0^{(M)}, \ldots, \hat{\bar{z}}_{N-1}^{(M)}\right], \quad (79)$$

$$\bar{D} := \left[\bar{d}_0^{(1)}, \ldots, \bar{d}_{N-1}^{(1)}, \ldots, \bar{d}_j^{(i)}, \ldots, \bar{d}_0^{(M)}, \ldots, \bar{d}_{N-1}^{(M)}\right], \quad (80)$$

$$\underline{D} = \left[\underline{d}_0^{(1)}, \ldots, \underline{d}_{N-1}^{(1)}, \ldots, \underline{d}_j^{(i)}, \ldots, \underline{d}_0^{(M)}, \ldots, \underline{d}_{N-1}^{(M)}\right]. \quad (81)$$

Moreover, $\underline{Z} = \underline{D}$. The sizes of those matrix blocks are easily deduced. Based on the Umeyama algorithm [45], we have the below lemma, which later resolves (46) and (50).

Lemma 2: Let $\hat{\bar{Z}}, \bar{D}$ be of size $\mathbb{R}^{d\times MN}$ and $\underline{Z} = \underline{D}$ be of size $\mathbb{R}^{(n+m)\times MN}$. Suppose $\underline{D}\underline{D}^\top$ is invertible. Let $\bar{U}\Lambda\bar{V}^\top$ be the singular value decomposition of $\hat{\bar{Z}}\left[I_{(n+m)\times(n+m)} - \underline{D}^\top(\underline{D}\underline{D}^\top)\underline{D}\right]\bar{D}^\top$ with singular values $\Lambda := \mathrm{diag}(\sigma_1, ..., \sigma_d)$ in decreasing order. Define $\bar{S}$ by

$$\bar{S} = \begin{cases} I_{d\times d}, & \det(\bar{U}\bar{V}) = 1 \\ \mathrm{diag}(I_{(d-1)\times(d-1)}, -1), & \det(\bar{U}\bar{V}) = -1 \end{cases}. \quad (82)$$

Then, the global minimum of the optimization

$$\operatorname*{arg\,min}_{R\in SO(d), W\in\mathbb{R}^{d\times(n+m)}} \left\| \begin{bmatrix} \hat{\bar{Z}} \\ \underline{Z} \end{bmatrix} - \begin{bmatrix} R & W \\ 0 & I \end{bmatrix} \begin{bmatrix} \bar{D} \\ \underline{D} \end{bmatrix} \right\|^2 \quad (83)$$

can be achieved by

$$R^* = \bar{U}\bar{S}\bar{V}^\top, \quad (84)$$

$$W^* = -(\hat{\bar{Z}} - \bar{V}\bar{S}\bar{U}^\top\bar{D})\underline{D}^\top(\underline{D}\underline{D}^\top)^{-1}. \quad (85)$$

Proof: See Appendix G. ∎

Remark 13: The (R^*, W^*) that achieves the global minimum of (83) may not be unique. Uniqueness is unnecessary in Theorem 2. However, the (R^*, W^*) is automatically the unique optimum of the corresponding optimization when $\mathrm{rank}\left(\hat{\bar{Z}}\left[I_{(n+m)\times(n+m)} - \underline{D}^\top(\underline{D}\underline{D}^\top)\underline{D}\right]\bar{D}^\top\right) \geq d-1$ [45], guaranteed when $\hat{Z}$ is close to Z owing to our assumption that $\mathrm{rank}(D) > d$.

We solve the TFG state reconstruction $\hat{\chi} = \mathcal{T}(\hat{z})$ formulated in (46) or (50) by Lemma 2. The constant weight Σ is absorbed into the formula through substituting $\hat{Z}, D$ by $\hat{Z}\Sigma^{-\frac{1}{2}}, D\Sigma^{-\frac{1}{2}}$. Define the components of the state estimation $\hat{\chi} \in \mathrm{TFG}(d,n,m)$ as

$$\hat{\chi} = \begin{bmatrix} \hat{R} & \hat{W} \\ 0 & I \end{bmatrix} \in \mathrm{TFG}(d,n,m). \quad (86)$$

The state reconstruction $\mathcal{T}$ of systems in Case 1 by (46) is

$$\begin{cases} \hat{R} = \bar{V}\bar{S}\bar{U}^\top \\ \hat{W} = (\bar{V}\bar{S}\bar{U}^\top\hat{\bar{Z}} - \bar{D})\underline{D}^\top(\underline{D}\underline{D}^\top)^{-1} \end{cases}, \quad (87)$$

and the reconstruction $\mathcal{T}$ of systems in Case 2 by (50) is

$$\begin{cases} \hat{R} = \bar{U}\bar{S}\bar{V}^\top \\ \hat{W} = -(\hat{\bar{Z}} - \bar{V}\bar{S}\bar{U}^\top\bar{D})\underline{D}^\top(\underline{D}\underline{D}^\top)^{-1} \end{cases}, \quad (88)$$

where $\bar{U}, \bar{V}, \bar{S}$ are calculated the same as in Lemma 2.

C. Global Properties of the Bias-free Observer for Immersible Two-frame Systems

By virtue of Theorem 2, the global stability of our observer for two-frame systems is determined by a rank condition.

Proposition 3: Consider the two-frame systems in Case 1 on $\mathrm{TFG}(d,n,m)$ in the form of (66). Its observer $(\mathcal{F}, \mathcal{T})$ with $\mathcal{F}$ defined by (77) and $\mathcal{T}$ defined by (87) is globally exponentially stable if the system structure $D = [\bar{D}^\top, \underline{D}^\top]^\top$ defined by (80)–(81) satisfies $\mathrm{rank}(D) \geq d+m+n$. Similarly for the two-frame systems in Case 2 (67), its observer $(\mathcal{F}, \mathcal{T})$ with $\mathcal{F}$ defined by (77) and $\mathcal{T}$ defined by (88) is GES if the system structure satisfies $\mathrm{rank}(D) \geq d+m+n$.

Proof: It suffices to show Assumption 2 holds. The $\Phi(t_1,t_2) \in \mathbb{R}^{d\times d}$ in (62) satisfies $\frac{\partial\Phi(t_1,t_2)}{\partial t_1} = \mp\omega_t^\times\Phi(t_1,t_2)$ for immersible two-frame systems, which implies $\Phi(t_1,t_2) \in SO(d)$. This shows that (77) is uniformly observable, thereby completing the proof by Theorem 2. ∎

D. Joint Estimation of Input Biases

Joint estimation of a constant $\mathfrak{tfg}(d,n,m)$-valued bias b is considered, and it only involves slight modification of the immersed system on $\mathbb{R}^{d_z}$ as well as its corresponding estimator. Let the components of the bias be

$$\mathcal{L}_{\mathrm{TFG}}(b) = \begin{bmatrix} b_\omega^\times & b_\rho \\ 0 & 0 \end{bmatrix} \in \mathfrak{tfg}(d,n,m), \quad (89)$$

where $b_\omega \in \mathbb{R}^{\frac{d(d-1)}{2}}$ and $b_\rho \in \mathbb{R}^{d\times(n+m)}$. By Assumption 1, the immersed system involves bias in an additive fashion, i.e., we can replace (ω_t, ρ_t) in (74) or (76) by $(\omega_t + b_\omega, \rho_t + b_\rho)$ to obtain the biased version. For two-frame systems in Case 1, the biased immersed system writes

$$\begin{cases} \dot{\bar{z}}_j^{(i)} = -(\omega_t + b_\omega)^\times \bar{z}_j^{(i)} - (\rho_t + b_\rho)\underline{d}_j^{(i)} - \bar{z}_{j+1}^{(i)} \\ \dot{\bar{z}}_{N-1}^{(i)} = -(\omega_t + b_\omega)^\times \bar{z}_{N-1}^{(i)} - (\rho_t + b_\rho)\underline{d}_{N-1}^{(i)} - \sum_{l=0}^{N-1}\tilde{a}_l\bar{z}_l^{(i)} \\ \dot{b}_\omega = 0, \ \dot{b}_\rho = 0, \quad j \in [0, N-2] \\ \bar{y}^{(i)} = \bar{z}_0^{(i)}, \ i \in [1, M] \end{cases}. \quad (90)$$

Similarly, for two-frame systems in Case 2, the biased immersed system writes

$$\begin{cases} \dot{\bar{z}}_j^{(i)} = (\omega_t + b_\omega)^\times \bar{z}_j^{(i)} + (\rho_t + b_\rho)\, \underline{d}_j^{(i)} - \bar{z}_{j+1}^{(i)} \\ \dot{\bar{z}}_{N-1}^{(i)} = (\omega_t + b_\omega)^\times \bar{z}_{N-1}^{(i)} + (\rho_t + b_\rho)\, \underline{d}_{N-1}^{(i)} - \sum_{l=0}^{N-1} \tilde{a}_l \bar{z}_l^{(i)} \\ \dot{b}_\omega = 0,\ \dot{b}_\rho = 0, \quad j \in [0, N-2] \\ \bar{y}^{(i)} = \bar{z}_0^{(i)},\ i \in [1, M] \end{cases}. \tag{91}$$

For observer design, it is necessary to vectorize b_ρ by stacking its columns to one vector.

Proposition 4: If the immersible two-frame system is corrupted only by b_ρ related to the vector-part dynamics, the observer composed of (87) cascaded after a Kalman observer for (90) (respectively, (88) cascaded after a Kalman observer for (91)) is globally exponentially stable.

Proof: If $b_\omega = 0$, (90) or (91) is linear. The Kalman observer for (90) or (91) is globally exponentially stable, because the observer with extended state $(\bar{z}, b_\rho)$ is uniformly observable. Such analysis is similar to the accelerometer-bias-only case in [23], [47]. The rest follows from Theorem 2. ■

Proposition 5: If the immersible two-frame system is corrupted by b_ρ and b_ω, the observer composed of (87) cascaded after an extended Kalman-like observer for (90) (respectively, (88) cascaded after an extended Kalman-like observer for (91)) is semi-globally stable, if (1) persistent determinability holds evaluated using $(\hat{\bar{z}}, \hat{b})$; (2) the true trajectory is bounded.

Proof: This directly follows from Theorem 3. ■

E. Extension to Range and Bearing Measurements

LTV immersion makes it possible to handle bearing or range measurements which do not fit into InEKF. In InEKF, the innovation term depends on the invariant error only [5]. Using invariant error to linearize such measurement equations involves additional dependence of the innovation on the estimated TFG state, destroying the local stability guarantee.

Let $\pi_{\mathbb{S}^n} : \mathbb{R}^{n+1} \to \mathbb{S}^n, x \mapsto x/\|x\|_2$ be the projection for bearing. Bearing measurements for system (66) or (67) are

$$y^{(i)} = \pi_{\mathbb{S}^{d-1}}\left(T_t^{-1} d^{(i)}\right) \text{ or } y^{(i)} = \pi_{\mathbb{S}^{d-1}}\left(T_t d^{(i)}\right) \tag{92}$$

respectively. Hence, the original measurement equation $\bar{y}^{(i)} = \bar{z}_0^{(i)}$ in (74) or (91) is replaced by

$$y^{(i)} = \pi_{\mathbb{S}^{d-1}}(\bar{z_0}^{(i)}), \tag{93}$$

while the immersion still works. Although (93) is nonlinear, it could be converted into time-varying linear form by well-known orthogonal projection techniques widely used in [24], [30], [47]–[50] as

$$0 = \left(I_{d\times d} - \bar{y}^{(i)} \bar{y}^{(i)\top}\right) \bar{z}_0^{(i)} := \Pi_{\bar{y}^{(i)}} \bar{z}_0^{(i)}, \tag{94}$$

where the known trajectory of $\bar{y}^{(i)}$ ($\|\bar{y}^{(i)}\| = 1$) is injected into the linear measurement matrix. 0 is viewed as the virtual measurement.

When considering immersible two-frame systems with bearing measurements, our observer is realized in the same manner as in Proposition 3 with the slight difference that the measurement equations in the immersed system should be replaced by (94). In addition to the rank condition on D, global exponential stability is achieved if the Kalman observer for the immersed system with bearing measurements is uniformly observable. This does not hold automatically as in Proposition 3. Similar observability analysis involving bearing has been comprehensively conducted in [24], [30]. Such analyses deviate from the aim of the present paper and thus are omitted.

Another interesting type of measurement is range, writing

$$y^{(i)} = \left\|T_t^{-1} d^{(i)}\right\|_2 \text{ or } y^{(i)} = \left\|T_t d^{(i)}\right\|_2 \tag{95}$$

for system (66) or (67). Hence, the original measurement equation $\bar{y}^{(i)} = \bar{z}_0^{(i)}$ in (74) or (91) is replaced by

$$y^{(i)} = \left\|\bar{z_0}^{(i)}\right\|_2. \tag{96}$$

Proposition 6: For two-frame systems (66) or (67) on TFG(d, n, m) with the range measurements (95), the immersed system (74) or (76) can further be immersed into an LTV system regarding state $(\bar{z}, s)$ with the extended state s

$$\begin{cases} \dot{\bar{z}} = f_1(\bar{z}, u) \\ \dot{s} = f_2(\bar{z}, s, u) \\ \frac{1}{2}\left(y^{(i)}\right)^2 = s_{0,0}^{(i)},\ i \in [1, M] \end{cases}, \tag{97}$$

where f_1 is linear in $\bar{z}$, as in (74) or (76). f_2 is linear with respect to $\bar{z}$ and s. $u \in \mathfrak{tfg}(d, n, m)$ is the input. The extended state is $s := [..., s_{j,k}^{(i)}, ...]^\top$, where each $\mathbb{R}$-scalar is defined by

$$s_{j,k}^{(i)} := \frac{1}{2} \bar{z}_j^{(i)\top} \bar{z}_k^{(i)}, \tag{98}$$

with index ranges $0 \le j \le k$, $0 \le k \le N-1$ and $1 \le i \le M$. Recall that $N = d + n + m$.

Proof: See Appendix H for expressions of f_2. ■

Remark 14: Linear time-invariant system (LTI) with quadratic outputs can be immersed into a higher dimensional LTV system [51], [52]. In general, if the system dynamics becomes dependent on t, taking the Lie derivative of the quadratic output along system dynamics will not terminate after finite steps, preventing the immersion. The above proposition works thanks to the anti-symmetric structure of ω_t. Hence, we provide a valuable example of a successful immersion of an LTV system with quadratic outputs.

When considering non-biased immersible two-frame systems with range measurements, our observer is realized in the same manner as in Proposition 3 with the modification that the Kalman filter for the immersed system should be designed for (97). In addition to the rank condition on D, global exponential stability is achieved if the Kalman observer for (97) is uniformly observable. This does not hold without further assumptions as in Proposition 3. Detailed observability analysis for (97) is beyond the scope of the present paper.

VI. Navigational Examples

Two non-trivial examples modeled by immersible two-frame systems are introduced to illustrate the implementation.

A. Navigation on Rotating Earth

Let the world frame anchored to Earth. We consider estimating the attitude, position and linear velocity of an aircraft using onboard high-precision IMU along with landmark, bearing and range of measurements. The body-frame is a fixed frame moving with the aircraft. The attitude $R_t \in \mathrm{SO}(3)$ is the rotation from the body to the world frame. Let $p_t, v_t \in \mathbb{R}^3$ be the position and linear velocity of the aircraft expressed in the world frame. The Earth angular velocity is $\Omega \in \mathbb{R}^3$ referenced in the world frame. The non-biased IMU dynamics write $\dot{R}_t = -\Omega^\times R_t + R_t\omega_t^\times$, $\dot{p}_t = v_t$ and $\dot{v}_t = R_t a_t + g - 2\Omega^\times v_t - (\Omega^\times)^2 p_t$, where $\omega_t, a_t \in \mathbb{R}^3$ are gyroscope and accelerometer outputs [8]. To fit the system into two-frame systems, define $W_t := [p_t, v_t + \Omega^\times p_t]$ and construct the state space by $T_t = \begin{bmatrix} R_t & W_t \\ 0 & I \end{bmatrix} \in \mathrm{TFG}(3,2,0)$. IMU dynamics satisfies (66) as $\dot{T}_t = \begin{bmatrix} -\Omega^\times & \gamma \\ 0 & L \end{bmatrix} T_t + T_t \begin{bmatrix} \omega_t^\times & \rho_t \\ 0 & -L \end{bmatrix}$, where $\gamma := [0, g] \in \mathbb{R}^{3\times 2}$, $\rho_t := [0, a_t] \in \mathbb{R}^{3\times 2}$. Note that $L := \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}$ and g is the local gravity expressed in the world frame. Consider two landmark-type measurements $\bar{y}^{(i)} = R_t^{-1}(\bar{d}^{(i)} - p_t), (i = 1, 2)$, one bearing measurement $\bar{y}^{(3)} = \pi_{\mathbb{S}^2}(R_t^{-1}(\bar{d}^{(3)} - p_t))$ and one range measurement $\bar{y}^{(4)} = \|R_t^{-1}(\bar{d}^{(4)} - p_t)\|$. The $\bar{d}^{(i)}$ for $i = 1, \dots, 4$ are known $\mathbb{R}^3$-vectors. To fit the measurement equations in the form of $y^{(i)} = T_t^{-1} d^{(i)}$ (66), we introduce homogeneous coordinates $y^{(i)} = \begin{bmatrix} \bar{y}^{(i)} \\ \underline{y}^{(i)} \end{bmatrix}, d^{(i)} = \begin{bmatrix} \bar{d}^{(i)} \\ \underline{d}^{(i)} \end{bmatrix}$ where $\underline{d}^{(i)} = \underline{y}^{(i)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

This example falls within case-1 immersible systems. Let $d_j^{(i)}$ and the state of the immersed LTV $z_j^{(i)}$ be $z_j^{(i)} := T_t^{-1} \begin{bmatrix} -\Omega^\times & \gamma \\ 0 & L \end{bmatrix}^j d^{(i)} := T_t^{-1} d_j^{(i)}$. The index ranges $i \in [1, 4]$, $j \in [0, 4]$. Let $z_j^{(i)} = \begin{bmatrix} \bar{z}_j^{(i)} \\ \underline{z}_j^{(i)} \end{bmatrix}$ and $d_j^{(i)} = \begin{bmatrix} \bar{d}_j^{(i)} \\ \underline{d}_j^{(i)} \end{bmatrix}$. The underlined variables are $\underline{d}_0^{(i)} \equiv \underline{z}_0^{(i)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\underline{d}_1^{(i)} \equiv \underline{z}_1^{(i)} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ and $\underline{d}_j^{(i)} \equiv \underline{z}_j^{(i)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ $(j = 2, 3, 4)$. The barred variables are $\bar{d}_0^{(i)} = \bar{d}^{(i)}$, $\bar{d}_1^{(i)} = -\Omega^\times \bar{d}_0^{(i)}$, $\bar{d}_2^{(i)} = (\Omega^\times)^2 \bar{d}_0^{(i)} - g$, $\bar{d}_3^{(i)} = \|\Omega\|^2 \Omega^\times (\bar{d}_0^{(i)} + g)$, and $\bar{d}_4^{(i)} = -\|\Omega\|^2 (\Omega^\times)^2 (\bar{d}_0^{(i)} + g)$. By Theorem 1, the immersed LTV reads $\dot{\bar{z}}_0^{(i)} = -\omega^\times \bar{z}_0^{(i)} - \bar{z}_1^{(i)}$, $\dot{\bar{z}}_1^{(i)} = -\omega^\times \bar{z}_1^{(i)} + a - \bar{z}_2^{(i)}$, $\dot{\bar{z}}_2^{(i)} = -\omega^\times \bar{z}_2^{(i)} - \bar{z}_3^{(i)}$, $\dot{\bar{z}}_3^{(i)} = -\omega^\times \bar{z}_3^{(i)} - \bar{z}_4^{(i)}$, $\dot{\bar{z}}_4^{(i)} = -\omega^\times \bar{z}_4^{(i)} + \|\Omega\|^2 \bar{z}_3^{(i)}$ for $i \in [1, 4]$. The landmark measurements are $\bar{y}^{(i)} = \bar{z}_0^{(i)} (i = 1, 2)$. The bearing measurement is $0 = (I_{3\times 3} - \bar{y}^{(3)} \bar{y}^{(3)\top}) \bar{z}_0^{(3)}$. For the range measurement, the immersed LTV is extended by $\dot{s}_{0,0} = -2s_{0,1}$, $\dot{s}_{0,1} = -s_{1,1} - \frac{1}{2} a^\top \bar{z}_0^{(4)} - s_{0,2}$, $\dot{s}_{0,2} = -s_{1,2} - s_{0,3}$, $\dot{s}_{0,3} = -s_{1,3} - s_{0,4}$, $\dot{s}_{0,4} = -s_{1,4} + \|\Omega\|^2 s_{0,3}$, $\dot{s}_{1,1} = -a^\top \bar{z}_1^{(4)} - 2s_{1,2}$, $\dot{s}_{1,2} = -\frac{1}{2} a^\top \bar{z}_2^{(4)} - s_{2,2} - s_{1,3}$, $\dot{s}_{1,3} = -\frac{1}{2} a^\top \bar{z}_3^{(4)} - s_{2,3} - s_{1,4}$, $\dot{s}_{1,4} = -\frac{1}{2} a^\top \bar{z}_4^{(4)} - s_{2,4} + \|\Omega\|^2 s_{1,3}$, $\dot{s}_{2,2} = -2s_{2,3}$, $\dot{s}_{2,3} = -s_{3,3} - s_{2,4}$, $\dot{s}_{2,4} = -s_{3,4} + \|\Omega\|^2 s_{2,3}$, $\dot{s}_{3,3} = -2s_{3,4}$, $\dot{s}_{3,4} = -s_{4,4} + \|\Omega\|^2 s_{3,3}$, $\dot{s}_{4,4} = 2\|\Omega\|^2 s_{3,4}$ with each $s_{\cdot,\cdot}$ a scalar. The range measurement is $\frac{1}{2}(y^{(4)})^2 = s_{0,0}$, exactly linear. We use the Kalman observer to obtain estimation for the immersed state. Using this estimate, (87) is used to reconstruct the $\mathrm{TFG}(3, 2, 0)$ state. Related matrices $\hat{\bar{Z}}, \bar{D}, \underline{D}$ are assembled using previously defined constants following (79)–(81).

B. SLAM with Moving Object Tracking

An aircraft with IMU navigates using 3 landmark measurements. Let its attitude, position and velocity be $R_t \in \mathrm{SO}(3)$, $p_t, v_t \in \mathbb{R}^3$. Meanwhile, the aircraft estimates a static landmark with unknown position $l_t \in \mathbb{R}^3$ and tracks a moving object with position $q_t \in \mathbb{R}^3$ and velocity $c_t \in \mathbb{R}^3$, using landmark measurements. The dynamics writes $\dot{R}_t = R_t \omega_t^\times, \dot{p}_t = v_t, \dot{v}_t = R_t a_t + g, \dot{l}_t = 0, \dot{q}_t = c_t, \dot{c}_t = 0$, where $\omega_t, a_t \in \mathbb{R}^3$ are gyroscope and accelerometer readings. g denotes the gravity. The system fits into case-1 immersible two-frame systems as $\dot{T}_t = \begin{bmatrix} 0^\times & \gamma \\ 0 & L \end{bmatrix} T_t + T_t \begin{bmatrix} \omega_t^\times & \rho_t \\ 0 & -L \end{bmatrix}$, where $T_t = \begin{bmatrix} R_t & W_t \\ 0 & I \end{bmatrix} \in \mathrm{TFG}(3, 5, 0)$ with $W_t = [p_t, v_t, l_t, q_t, c_t] \in \mathbb{R}^{3\times 5}$, $\gamma := [0_{3\times 1}, g, 0_{3\times 3}] \in \mathbb{R}^{3\times 5}$, and $\rho_t := [0_{3\times 1}, a_t, 0_{3\times 3}] \in \mathbb{R}^{3\times 5}$. Note that $L = [L_{ij}] \in \mathbb{R}^{5\times 5}$, only $L_{21} = L_{54} = -1$ and other elements are 0. We consider the following measurements $\bar{y}^{(i)} = R_t^{-1}(\bar{d}^{(i)} - p_t), (i = 1, 2, 3)$, $\bar{y}^{(4)} = R_t^{-1}(l_t - p_t)$, $\bar{y}^{(5)} = R_t^{-1}(q_t - p_t)$, and $\bar{y}^{(6)} = R_t^{-1}(c_t - v_t)$. Using the homogeneous coordinates $y^{(i)} = \begin{bmatrix} \bar{y}^{(i)} \\ \underline{y}^{(i)} \end{bmatrix}$, $d^{(i)} = \begin{bmatrix} \bar{d}^{(i)} \\ \underline{d}^{(i)} \end{bmatrix}$, all measurements are in the form of $y^{(i)} = T_t^{-1} d^{(i)}$. One has $\bar{d}^{(i)} (i = 1, 2, 3)$ being known landmark positions and $\bar{d}^{(i)} = 0_{3\times 1} (i = 4, 5, 6)$. Moreover, $\underline{d}^{(i)} = [1, 0_{1\times 4}]^\top (i = 1, 2, 3)$, $\underline{d}^{(4)} = [1, 0, -1, 0_{1\times 2}]^\top$, $\underline{d}^{(5)} = [1, 0_{1\times 2}, -1, 0]^\top$ and $\underline{d}^{(6)} = [0, 1, 0_{1\times 2}, -1]^\top$.

By Theorem 1, let $d_j^{(i)}$ and the state of the immersed LTV $z_j^{(i)}$ be $z_j^{(i)} := T_t^{-1} \begin{bmatrix} 0^\times & \gamma \\ 0 & L \end{bmatrix}^j d^{(i)} := T_t^{-1} d_j^{(i)}$. Let $z_j^{(i)} = \begin{bmatrix} \bar{z}_j^{(i)} \\ \underline{z}_j^{(i)} \end{bmatrix}$ and $d_j^{(i)} = \begin{bmatrix} \bar{d}_j^{(i)} \\ \underline{d}_j^{(i)} \end{bmatrix}$. Hence, the LTV immersion governing $\bar{z}_j^{(i)}$ reads $\dot{\bar{z}}_0^{(i)} = -\omega_t^\times \bar{z}_0^{(i)} - \bar{z}_1 (i = 1, 2, 3, 4)$, $\dot{\bar{z}}_0^{(5)} = -\omega_t^\times \bar{z}_0^{(5)} - \bar{z}_1^{(5)}$, $\dot{\bar{z}}_1 = -\omega_t^\times \bar{z}_1 + a_t - \bar{z}_2$, $\dot{\bar{z}}_1^{(5)} = -\omega_t^\times \bar{z}_1^{(5)} + a_t - \bar{z}_2$, and $\dot{\bar{z}}_2 = -\omega_t^\times \bar{z}_2$. Note the immersed equations from different measurements may be the same, and thus they are represented by one common $\bar{z}_j$ without superscript (i). The original measurements now become those of the immersed LTV, as $\bar{y}^{(i)} = \bar{z}_0^{(i)}, i \in [1, 5]$ and $\bar{y}^{(6)} = \bar{z}_1^{(5)}$. We use the Kalman observer to obtain estimation for the immersed state. Using this estimate, (87) is used to reconstruct the $\mathrm{TFG}(3, 5, 0)$ state. Related matrices $\hat{\bar{Z}}, \bar{D}, \underline{D}$ are assembled according to (79)–(81). Our observer achieves GES as long as $\bar{d}^{(i)} (i = 1, 2, 3)$ are linearly independent. The InEKF for such system achieves only local stability [4], [5]. Simulations comparing our observer to the InEKF are shown in Fig. 1.

VII. Conclusion

The group-theoretic properties that allow LTV immersion of linear observed systems are identified. An observer framework is then proposed as an optimization-based state reconstruction cascaded after a Kalman-like observer for the immersed LTV. GES is achieved provided a suitable rank condition on the

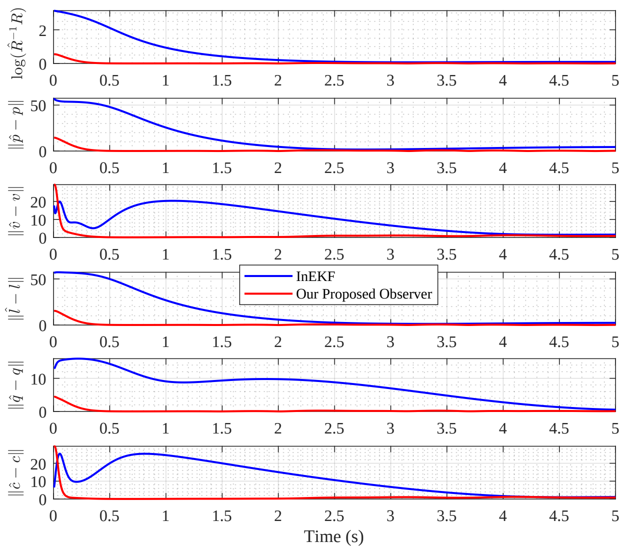

Fig. 1. Comparison of our proposed observer to InEKF for SLAM with moving object tracking. The observers are initialized at the same value very far from the true state. The errors of all components are calculated after the first update. Observers are tuned by associating noises.

system structure is satisfied and one global minimum of the optimization is found. Bias estimation with semi-global stability guarantees is discussed. The theory is applied to two-frame systems and illustrations of the implementation are provided through non-trivial navigation examples. Future work will focus on applying the theory to multi-frame systems involving multiple agents [53].

Appendix

A. Proof of Proposition 1

Define $g_u^L(\chi) := f_u(\chi) - \chi f_u(id)$. A direct calculation yields $g_u^L(\chi_1\chi_2) = f_u(\chi_1\chi_2) - \chi_1\chi_2 f_u(id) = \chi_1(f_u(\chi_2) - \chi_2 f_u(id)) + (f_u(\chi_1) - \chi_1 f_u(id))\chi_2 = \chi_1 g_u^L(\chi_2) + g_u^L(\chi_1)\chi_2$, hence $g_u^L \in \mathfrak{aut}(G)$. The same applies for g_u^R. As $f_u(id)$ is in the tangent space at $id \in G$ and hence lies in $\mathfrak{g}$, $\chi f_u(id)$ or $f_u(id)\chi$ is the left- or right-invariant vector field on G by left- or right-translation of the vector at id to the whole group [31], [32], thus belongs to $L(G)$ or $R(G)$, completing the proof. ■

B. Proof of Proposition 2

In Case 1, the dynamics reads $\dot{\chi} = f_u(\chi) = g_A(\chi) + \chi f_u(id) = A\chi - \chi(A - f_u(id))$ with measurement $y = h(\chi) = \chi^{-1}d \in \mathbb{R}^{d_y}$. Using the push-forward of the group action, define $z_0 := \chi^{-1}d \in \mathbb{R}^{d_y}$ and calculate $\mathcal{L}_{f_u}h$ as

$$
\begin{aligned}
\dot{z}_0 &= -\chi^{-1}\dot{\chi}\chi^{-1}d = -\chi^{-1}(A\chi + \chi(f_u(id) - A))\chi^{-1}d \\
&= -(f_u(id) - A)(\chi^{-1}d) - \chi^{-1}(Ad) \\
&= -(f_u(id) - A)z_0 - \chi^{-1}(Ad).
\end{aligned}
$$

Note that the measurement equation is $y = z_0$. Define $z_j := \chi^{-1}(A^j d)$. Taking the time derivative along f_u, we obtain

$$
\begin{aligned}
\dot{z}_j &= -\chi^{-1}\dot{\chi}\chi^{-1}(A^j d) \\
&= -\chi^{-1}(A\chi + \chi(f_u(id) - A))\chi^{-1}(A^j d) \\
&= -(f_u(id) - A)\chi^{-1}(A^j d) - \chi^{-1}(A^{j+1}d) \\
&= -(f_u(id) - A)z_j - z_{j+1}.
\end{aligned}
$$

Note we have used the fact that A does not depend on u, and thus $\dot{A} = 0$ thanks to our restriction of f_u to just $g_A + L(G)$ instead of $\mathfrak{inn}(G) + L(G)$. It's natural to ask whether the process of inductive definition of z_{j+1} through z_j terminates after finite-many steps. As $A \in \mathfrak{g}$ is now viewed as a time-independent $\mathbb{R}^{d_y \times d_y}$ linear operator on $\mathbb{R}^{d_y}$, by Cayley-Hamilton Theorem, there exist d_y real constants $a_0, a_1, \ldots, a_{d_y-1} \in \mathbb{R}$, such that

$$
A^{d_y} = a_{d_y-1}A^{d_y-1} + a_{d_y-2}A^{d_y-2} + \cdots + a_1 A + a_0 I. \quad (99)
$$

By definition of z_js, this directly yields

$$
z_{d_y} = a_{d_y-1}z_{d_y-1} + a_{d_y-2}z_{d_y-2} + \cdots + a_1 z_1 + a_0 z_0. \quad (100)
$$

To summarize, a closed list of equations on variables $z_0, \ldots, z_{d_y-1}$ are obtained with measurement $y = z_0$:

$$
\begin{aligned}
\dot{z}_j &= -(f_u(id) - A)z_j - z_{j+1}, \ j \in [0, d_y - 2], \\
\dot{z}_{d_y-1} &= -(f_u(id) - A)z_{d_y-1} - z_{d_y} \\
&= -(f_u(id) - A)z_{d_y-1} - \sum_{l=0}^{d_y-1} a_l z_l.
\end{aligned}
$$

The above equation is linear time-varying, completing the immersion for Case 1. Although the mechanics in Case 2 are analogous, we emphasize here that to cancel undesirable χs by multiplication as before, the group-affine dynamics should be adjusted to $f_u \in g_A + R(G)$. To be precise, $\dot{\chi} = f_u(\chi) = A\chi - \chi A + f_u(id)\chi$. Let $z_j := \chi(A^j d)$. Taking derivatives along the new f_u yields

$$
\begin{aligned}
\dot{z}_j &= \dot{\chi}(A^j d) = (A + f_u(id))\chi(A^j d) - \chi(A^{j+1}d) \\
&= (A + f_u(id))z_j - z_{j+1}.
\end{aligned}
$$

As the operator $A \in \mathfrak{g}$ satisfies the same constraint (99), this leads to the same relationship (100) among z_js in Case 2. The immersed LTV system in Case 2 is

$$
\begin{aligned}
\dot{z}_j &= (f_u(id) + A)z_j - z_{j+1}, \ j \in [0, d_y - 2], \\
\dot{z}_{d_y-1} &= (f_u(id) + A)z_{d_y-1} - \sum_{l=0}^{d_y-1} a_l z_l,
\end{aligned}
$$

with the same a_l and measurement equation as Case 1. ■

C. Proof of Theorem 1

First we show $f_u \in \mathfrak{inn}(\tilde{G})|_{\Gamma(TG)} + L(G)$ or $\mathfrak{inn}(\tilde{G})|_{\Gamma(TG)} + R(G)$ defines group-affine dynamics. It suffices to show that $\mathfrak{inn}(\tilde{G})|_{\Gamma(TG)} \subset \mathfrak{aut}G$. Let $\tilde{\phi} : \mathbb{R} \times \tilde{G} \times \mathbb{R} \to \tilde{G}$ be a flow of $\tilde{G}$. Let $\tilde{\phi}(s; \cdot, t) \in \mathrm{Inn}(\tilde{G})$ by fixing s and t, then $\tilde{\phi}(s; \chi, t) = \tilde{\chi}_0(s)\chi\tilde{\chi}_0^{-1}(s)$ is a conjugation. As G is a normal subgroup of $\tilde{G}$, for every $\chi \in G$, $\tilde{\phi}(s; \chi, t) = \tilde{\chi}_0(s)\chi\tilde{\chi}_0^{-1}(s)$ remains in G for all $\chi_0(s) \in \tilde{G}$. This means that $\tilde{\phi}$ can be restricted to a flow $\phi : \mathbb{R} \times G \times \mathbb{R} \to G$ as $\phi(s; \chi, t) := \tilde{\phi}(s; \chi, t)|_G$, $\forall \chi \in G$. Now $\phi(s; \cdot, t)$ is an automorphism of G when s and t are fixed, but not necessarily inner if $\tilde{\chi}(s) \notin G$. Let $g \in \mathfrak{inn}(\tilde{G})|_{\Gamma(TG)}$, then by definition (21)

$$
g(\chi) := \frac{d}{ds}|_{s=t}\phi(s; \chi, t), \ \ \phi(s; \chi, t) \in \mathrm{Aut}(G) \supsetneq \mathrm{Inn}(G),
$$

which implies $g \in \mathfrak{aut}(G)$. Hence, f_u indeed defines group-affine dynamics. The proof of immersion is straightforward, as the linear observed system on G can be viewed as the restriction of a bi-invariant system described in Proposition 2. Similarly to Appendix B, define $z_j := \chi^{-1}(\tilde{A}^j d),\ j = 0, \dots, d_y - 1$ for Case 1, we obtain

$$\dot{z}_j = -(f_u(id) - \tilde{A})z_j - z_{j+1},\ j \in [0, d_y - 2],$$
$$\dot{z}_{d_y-1} = -(f_u(id) - \tilde{A})z_{d_y-1} - \sum_{l=0}^{d_y-1} \tilde{a}_l z_l,$$

where $\tilde{a}_l$ is defined by $\tilde{A}^{d_y} = \sum_{l=0}^{d_y-1} \tilde{a}_l \tilde{A}^l$. For Case 2, define $z_j := \chi(\tilde{A}^j d),\ j \in [0, d_y - 1]$, one obtains

$$\dot{z}_j = (f_u(id) + \tilde{A})z_j - z_{j+1},\ j \in [0, d_y - 2],$$
$$\dot{z}_{d_y-1} = (f_u(id) + \tilde{A})z_{d_y-1} - \sum_{l=0}^{d_y-1} \tilde{a}_l z_l,$$

with the same $\tilde{a}_l$s as in Case 1. The measurement equation in both cases is $y = z_0$. ∎

D. Proof of Lemma 1

Due to the diagonal structure of F_u and H in (37) and (38) exhibiting repeated patterns, it suffices to prove the pair $(F_u^{(i)}, H^{(i)})$ defined by (37) and (40) is uniformly observable.

We first calculate the transition matrix $\Phi^F(t_2, t_1)$ of $F_u^{(i)}$. Let $F_u^{(i)} = \bar{A} + \bar{S}_t$, where $\bar{S}_t = \mathrm{diag}(-S_u, \dots, -S_u)$ and

$$\bar{A} = \begin{bmatrix} 0 & -I & 0 & \cdots & 0 \\ 0 & 0 & -I & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\tilde{a}_0 I & -\tilde{a}_1 I & -\tilde{a}_2 I & \cdots & -\tilde{a}_{d_y-1} I \end{bmatrix} \in \mathbb{R}^{d_y^2 \times d_y^2}.$$

Note that the subscript in $\bar{S}_t$ indicates its dependence on time. One can verify that $\bar{S}_t \bar{A} = \bar{A}\bar{S}_t$. Decomposing $F_u^{(i)}$ into the sum of two commuting matrices is the standard technique in observability analysis [23], [24]. To proceed, denote $\Phi^{\bar{S}}(t_2, t_1)$ to be the state transition matrix of $\bar{S}_t$, then $\Phi^{\bar{S}}(t_2, t_1) = \mathrm{diag}(\Phi(t_2, t_1), \dots, \Phi(t_2, t_1))$ since $\Phi(t_2, t_1)$ is the transition matrix of $-S_u = \mp f_u(id) + \tilde{A}$. Similarly, $\Phi^{\bar{S}}(t_2, t_1)\bar{A} = \bar{A}\Phi^{\bar{S}}(t_2, t_1), \forall t_1, t_2 \in \mathbb{R}$. For any fixed t_0, we claim $\Phi^F(t_2, t_1) = \Phi^{\bar{S}}(t_2, t_0)\Phi^{\bar{A}}(t_2, t_1)[\Phi^{\bar{S}}(t_1, t_0)]^{-1}$, where $\Phi^{\bar{A}}(\cdot, \cdot)$ is the transition matrix of $\bar{A}$. Let $\Psi(t_2, t_1)$ denote the right-hand side. It's clear that $\Psi(t_1, t_1) = I$ and $\Psi(t_2, t_1) = \Psi(t_1, t_2)^{-1}$. Compute the partial derivative $\frac{\partial \Psi}{\partial t_2} = \bar{S}_t \Psi + \Phi^{\bar{S}}(t_2, t_0)\bar{A}\Phi^{\bar{A}}(t_2, t_1)[\Phi^{\bar{S}}(t_1, t_0)]^{-1} = (\bar{S}_t + \bar{A})\Psi$, thereby proving the claim. The constant matrix pair $(\bar{A}, \bar{H})$ is Kalman observable, as $[\bar{H}^\top, (\bar{H}\bar{A})^\top, \dots, (\bar{H}\bar{A}^N)^\top]^\top$ is of full column rank, implying $\exists \delta_1, \alpha_1 > 0$, such that

$$\int_t^{t+\delta_1} \Phi^{\bar{A}}(\tau, t)^\top \bar{H}_\tau^\top R^{-1} \bar{H}_\tau \Phi^{\bar{A}}(\tau, t) d\tau \succeq \alpha_1 I,$$

for every t given $R \in \mathbb{S}_+^{d_y}$. The Lemma condition implies the existence of $\delta, \alpha > 0$, such that $\Phi^{\bar{S}}(t+\delta, t)^\top \Phi^{\bar{S}}(t+\delta, t) \succeq \alpha I$. Note that $\Phi^F(t_2, t_1) = \Phi^{\bar{S}}(t_2, t_1)\Phi^{\bar{A}}(t_2, t_1)$, hence

$$\int_t^{t+\delta_1} \Phi^F(\tau, t)^\top \bar{H}_\tau^\top R^{-1} \bar{H}_\tau \Phi^F(\tau, t) d\tau \succeq \frac{\delta_1}{\delta} \alpha \alpha_1 I,$$

for each $t \in \mathbb{R}$, completing the proof. ∎

E. Proof of Theorem 2

We only prove for a system in Case 1. The reasoning for Case 2 is identical. Let $\chi(t), \hat{\chi}(t)$ be the true and estimated trajectory on G respectively. Let $z = \pi(\chi)$ and $\hat{z} = \pi(\hat{\chi})$ be the true and estimated state of the immersed LTV system, where π is the immersion map defined by (24). Let $N = d_y - 1$. Let $Z = [z_0^{(1)}, \dots, z_N^{(1)}, \dots, z_j^{(i)}, \dots, z_0^{(M)}, \dots, z_N^{(M)}]$ and $\hat{Z} = [z_0^{(1)}, \dots, z_N^{(1)}, \dots, z_j^{(i)}, \dots, z_0^{(M)}, \dots, z_N^{(M)}]$ associated with the true and estimated state of the LTV. Let the system structure be $D = [d_0^{(1)}, \dots, d_N^{(1)}, \dots, d_j^{(i)}, \dots, d_0^{(M)}, \dots, d_N^{(M)}]$. The matrices $Z, \hat{Z}, D$ are of size $\mathbb{R}^{d_y \times M d_y}$.

We first show that $d(\hat{\chi}, \chi)$ can be bounded by $\|\hat{Z} - Z\|$. A well-known fact is $\lambda_{\min}(S)\mathrm{tr}(K) \le \mathrm{tr}(KS) \le \lambda_{\max}(S)\mathrm{tr}(K)$, where $K \in \mathbb{S}_+$, S is symmetric and is of the same size as K. Hence, $\lambda_{\min}(DD^\top)\mathrm{tr}[(\chi^{-1} - \hat{\chi}^{-1})(\chi^{-1} - \hat{\chi}^{-1})^\top] \le \mathrm{tr}[DD^\top(\chi^{-1} - \hat{\chi}^{-1})(\chi^{-1} - \hat{\chi}^{-1})^\top] = \mathrm{tr}[(\chi^{-1} - \hat{\chi}^{-1})DD^\top(\chi^{-1} - \hat{\chi}^{-1})^\top] = \|\chi^{-1}D - \hat{\chi}^{-1}D\|^2 = \|Z - \hat{\chi}^{-1}D\|^2$. On the other hand, as $\hat{\chi} = \arg\min_{\hat{\chi} \in G} \|\hat{Z} - \hat{\chi}^{-1}D\|^2$, we have $\|\hat{Z} - \hat{\chi}^{-1}D\| \le \|\hat{Z} - Z\|$, since there exists $\chi' \in G$ such that $Z = \chi'^{-1}D$. This gives the estimate $\|Z - \hat{\chi}^{-1}D\| \le \|Z - \hat{Z}\| + \|\hat{Z} - \hat{\chi}^{-1}D\| \le 2\|\hat{Z} - Z\|$. By Assumption 3, $\sigma_{\min}(D) > 0$, which bounds the error metric $d(\hat{\chi}, \chi) \le \frac{2}{\sqrt{\sigma_{\min}(D)}}\|\hat{Z} - Z\|$.

We then show that $\hat{Z}$ tends to Z exponentially from any initial value. This is the standard proof of the stability of Kalman observer under uniform observability [30], [54], thus we make it brief. Using Assumption 2, we have (F_u, H) being uniformly observable by Lemma 1. Thus $p_m I \preceq P \preceq p_M I$ is uniformly bounded [36]. Let $\tilde{z} = \hat{z} - z$ be the error state, whose governing ODE is $\dot{\tilde{z}} = (F_u - KH)\tilde{z}$. Denote $V(\tilde{z}) := \tilde{z}^\top P^{-1} \tilde{z}$. One obtains $1/p_M \|\tilde{z}\|^2 \le V \le 1/p_m \|\tilde{z}\|^2$ and $V \to +\infty$ as $\|\tilde{z}\| \to +\infty$. To prove uniform GES, one only needs to confirm $\dot{V} \le -\frac{\lambda_{\min}(Q)}{p_M^2}\|\tilde{z}\|^2$ below a negative quadratic function. There exist constants $c_1, c_2 > 0$ such that $\|\hat{Z}(t) - Z(t)\| \le c_1\|\hat{Z}(t_0) - Z(t_0)\| \exp[-c_2(t - t_0)]$. As $c_3 d(\hat{\chi}(t_0), \chi(t_0)) \ge \|\hat{Z}(t_0) - Z(t_0)\|$ for some $c_3 > 0$. The error metric satisfies $d(\hat{\chi}(t), \chi(t)) \le \frac{2c_1c_3}{\sqrt{\sigma_{\min}(D)}} d(\hat{\chi}(t_0), \chi(t_0)) \exp[-c_2(t - t_0)]$, completing the proof. ∎

F. Proof of Theorem 3

Let the system structure be of Case 1. As the true χ evolves in a compact set $\mathcal{G}_1$, $z := \pi(\chi)$ with $z_j^{(i)} := \chi^{-1}\tilde{A}^j d^{(i)}$ is in a compact set $\mathcal{Z}$ of the immersed space $\mathbb{R}^{M d_y^2}$. By Assumption 3, $d(\hat{\chi}, \chi) \le c_1\|\hat{Z} - Z\|$ for some $c_1 > 0$. It suffices to show that there exists a compact subset $\hat{\mathcal{Z}} \times \hat{\mathcal{B}} \subset \mathbb{R}^{M d_y^2} \times \mathbb{R}^{\dim \mathfrak{g}}$, such that $[\hat{z}^\top, \hat{b}^\top]^\top$ initialized in $\mathrm{int}(\mathcal{Z} \times \mathcal{B})$ with $P(t_0) = P_0$ remains in $\hat{\mathcal{Z}} \times \hat{\mathcal{B}}$ and the convergence of $[\hat{z}^\top, \hat{b}^\top]^\top$ to $[z^\top, b^\top]^\top$ is exponential after finite time. To verify the conditions of [35, Th. 1], we confirm (1) the pair $(\breve{F}, \breve{H})$ is persistently determinable; (2) $[z(t)^\top, b^\top]^\top$ is bounded; (3) the linearization error from the dynamics is given by $\varphi([z^\top, b^\top]^\top, [\hat{z}^\top, \hat{b}^\top]^\top) := F_u(z - \hat{z}) - F_b z + F_{\hat{b}}\hat{z} - \breve{F}[(\hat{z} - z)^\top, (\hat{b} - b)^\top]^\top$, where blocks are defined in (56), thus $\|\varphi\|$ is bounded by a quadratic function of $\|\hat{z} - z\|$ and $\|\hat{b} - b\|$.

With [35, Th. 1], we obtain a semi-global result with joint bias estimation by immersion. ■

G. Proof of Lemma 2

Let the cost of the optimization be $J(R, W)$. We wish to first decouple the vector part from J by completing square, as

$$\begin{aligned} J &= \operatorname{tr}\left[(\hat{\bar{Z}} - R^{-1}(\bar{D} - W\underline{D}))(\hat{\bar{Z}} - R^{-1}(\bar{D} - W\underline{D}))^\top\right] \\ &= (\hat{\bar{Z}} - R^{-1}\bar{D})(\hat{\bar{Z}} - R^{-1}\bar{D})^\top + 2R^{-1}W\underline{D}(\hat{\bar{Z}} \\ &\quad - R^{-1}\bar{D})^\top + R^{-1}W\underline{D}\underline{D}^\top W^\top R \\ &= \operatorname{tr}\bigg[\left(W\underline{D}\underline{D}^\top + (\hat{\bar{Z}} - R^{-1}\bar{D})\underline{D}^\top\right)(\underline{D}\underline{D}^\top)^{-1} \\ &\quad \left(W\underline{D}\underline{D}^\top + (\hat{\bar{Z}} - R^{-1}\bar{D})\underline{D}^\top\right)^\top\bigg] + \operatorname{tr}\bigg[\left(\hat{\bar{Z}} - R^{-1}\bar{D}\right) \\ &\quad \left(I_{(n+m)\times(n+m)} - \underline{D}^\top(\underline{D}\underline{D}^\top)\underline{D}\right)\left(\hat{\bar{Z}} - R^{-1}\bar{D}\right)^\top\bigg]. \end{aligned}$$

As the first $\operatorname{tr}(\cdot)$ term is non-negative, the global minimum (R^*, W^*) of J must be achieved with

$$W^* = -(\hat{\bar{Z}} - R^{*-1}\bar{D})\underline{D}^\top(\underline{D}\underline{D}^\top)^{-1}. \tag{101}$$

This means we could minimize the second term with rotation only, and plug the R which achieves the global minimum of the second $\operatorname{tr}(\cdot)$ term back to (101). Let the Cholesky decomposition be $\bar{L}\bar{L}^\top = I_{(n+m)\times(n+m)} - \underline{D}^\top(\underline{D}\underline{D}^\top)\underline{D}$. It suffices to consider the optimization problem with cost $\tilde{J}(R)$:

$$\underset{R\in \mathrm{SO}(d)}{\arg\min} \left\|\hat{\bar{Z}}\bar{L} - R^{-1}\bar{D}\bar{L}\right\|^2. \tag{102}$$

Expanding $\tilde{J}$ using the properties of trace, we obtain

$$\tilde{J} = \|\hat{\bar{Z}}\bar{L}\|^2 + \|\hat{\bar{D}}\bar{L}\|^2 - 2\operatorname{tr}\left(\hat{\bar{Z}}\bar{L}\bar{L}^\top\bar{D}^\top R\right). \tag{103}$$

By lemma conditions, $\bar{U}\Lambda\bar{V}^\top$ is the singular value decomposition of $\hat{\bar{Z}}\bar{L}\bar{L}^\top\bar{D}^\top$ with singular values $\Lambda := \operatorname{diag}(\sigma_1, ..., \sigma_d)$ in decreasing order. Using the results from [45], the global optimum R^* of (102) and thus (83) is given by

$$R^* = \bar{U}\bar{S}\bar{V}^\top. \tag{104}$$

Substituting (104) back into (101), we get $W^* = -(\hat{\bar{Z}} - \bar{V}\bar{S}\bar{U}^\top\bar{D})\underline{D}^\top(\underline{D}\underline{D}^\top)^{-1}$, which completes the proof. ■

H. Proof of Proposition 6

It suffices to show that taking derivatives of $s^{(i)}_{j,k}$ will not create non-desirable terms which are not included in $\bar{z}$ or s. First, assuming $0 \le j \le k \le N-2$, calculation shows $\dot{s}^{(i)}_{j,k} = \frac{1}{2}\dot{\bar{z}}^{(i)\top}_j\bar{z}^{(i)}_k + \frac{1}{2}\bar{z}^{(i)\top}_j\dot{\bar{z}}^{(i)}_k = \frac{1}{2}(-\omega_t^\times\bar{z}^{(i)}_j - \rho_t\underline{d}^{(i)}_j - \bar{z}^{(i)}_{j+1})^\top\bar{z}^{(i)}_k + \frac{1}{2}\bar{z}^{(i)\top}_j(-\omega_t^\times\bar{z}^{(i)}_k - \rho_t\underline{d}^{(i)}_k - \bar{z}^{(i)}_{k+1}) = -\frac{1}{2}(\rho_t\underline{d}^{(i)}_j)^\top\bar{z}^{(i)}_k - s^{(i)}_{j+1,k} - \frac{1}{2}(\rho_t\underline{d}^{(i)}_k)^\top\bar{z}^{(i)}_j - s^{(i)}_{j,k+1}$. Thanks to the anti-symmetric structure of $\omega_t^\times$, the terms $\frac{1}{2}(-\omega_t^\times\bar{z}^{(i)}_j)^\top\bar{z}^{(i)}_k$ and $\frac{1}{2}\bar{z}^{(i)\top}_j(-\omega_t^\times\bar{z}^{(i)}_k)$ exactly cancel. Otherwise, these two-terms and their subsequent derivatives will create infinitely many new terms preventing the finite termination. The same reasoning applies to the cases $0 \le j \le k = N-1$ or $j = k = N-1$. The $\tilde{a}_\iota$s arise from the dynamics f_1.

Thus, the additional immersion f_2 in (97) for two-frame systems in Case 1 (66) with range measurements is given by $\dot{s}^{(i)}_{j,k} = -\frac{1}{2}(\rho_t\underline{d}^{(i)}_j)^\top\bar{z}^{(i)}_k - s^{(i)}_{j+1,k} - \frac{1}{2}(\rho_t\underline{d}^{(i)}_k)^\top\bar{z}^{(i)}_j - s^{(i)}_{j,k+1}$, for $0 \le j \le k \le N-2$; $\dot{s}^{(i)}_{j,N-1} = -\frac{1}{2}(\rho_t\underline{d}^{(i)}_j)^\top\bar{z}^{(i)}_{N-1} - s^{(i)}_{j+1,N-1} - \frac{1}{2}(\rho_t\underline{d}^{(i)}_{N-1})^\top\bar{z}^{(i)}_j - \sum_{\iota=0}^{N-1}\tilde{a}_\iota s_{j,\iota}$, for $0 \le j \le N-2$; and $\dot{s}^{(i)}_{N-1,N-1} = -(\rho_t\underline{d}^{(i)}_{N-1})^\top\bar{z}^{(i)}_{N-1} - \sum_{\iota=0}^{N-1}2\tilde{a}_\iota s_{\iota,N-1}$.

Similarly, the additional immersion f_2 in (97) for two-frame systems in Case 2 (67) with range measurements is given by $\dot{s}^{(i)}_{j,k} = \frac{1}{2}(\rho_t\underline{d}^{(i)}_j)^\top\bar{z}^{(i)}_k + s^{(i)}_{j+1,k} + \frac{1}{2}(\rho_t\underline{d}^{(i)}_k)^\top\bar{z}^{(i)}_j + s^{(i)}_{j,k+1}$, for $0 \le j \le k \le N-2$; $\dot{s}^{(i)}_{j,N-1} = \frac{1}{2}(\rho_t\underline{d}^{(i)}_j)^\top\bar{z}^{(i)}_{N-1} + s^{(i)}_{j+1,N-1} + \frac{1}{2}(\rho_t\underline{d}^{(i)}_{N-1})^\top\bar{z}^{(i)}_j + \sum_{\iota=0}^{N-1}\tilde{a}_\iota s^{(i)}_{j,\iota}$, for $0 \le j \le N-2$; and $\dot{s}^{(i)}_{N-1,N-1} = (\rho_t\underline{d}^{(i)}_{N-1})^\top\bar{z}^{(i)}_{N-1} + \sum_{\iota=0}^{N-1}2\tilde{a}_\iota s^{(i)}_{\iota,N-1}$. Note the index i ranges from 1 to M, corresponding to the i-th range measurement. ■

References

[1] A. Barrau and S. Bonnabel, "Linear observed systems on groups," *Systems & Control Letters*, vol. 129, pp. 36–42, July 2019.

[2] A. Barrau and S. Bonnabel, "A mathematical framework for imu error propagation with applications to preintegration," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 5732–5738.

[3] C. Liu and Y. Shen, "On the existence of linear observed systems on manifolds with connection," *IEEE Control Systems Letters*, vol. 8, pp. 2607–2612, 2024.

[4] A. Barrau and S. Bonnabel, "The geometry of navigation problems," *IEEE Transactions on Automatic Control*, vol. 68, no. 2, pp. 689–704, 2023.

[5] A. Barrau and S. Bonnabel, "The invariant extended kalman filter as a stable observer," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1797–1812, 2017.

[6] A. Barrau and S. Bonnabel, "Invariant kalman filtering," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 237–257, 2018.

[7] A. Barrau and S. Bonnabel, "Intrinsic filtering on lie groups with applications to attitude estimation," *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 436–449, 2015.

[8] M. Brossard, A. Barrau, P. Chauchat, and S. Bonnabel, "Associating uncertainty to extended poses for on lie group imu preintegration with rotating earth," *IEEE Transactions on Robotics*, vol. 38, no. 2, pp. 998–1015, 2022.

[9] P. van Goor and R. Mahony, "Eqvio: An equivariant filter for visual-inertial odometry," *IEEE Transactions on Robotics*, vol. 39, no. 5, pp. 3567–3585, 2023.

[10] A. Fornasier, P. van Goor, E. Allak, R. Mahony, and S. Weiss, "Msceqf: A multi state constraint equivariant filter for vision-aided inertial navigation," *IEEE Robotics and Automation Letters*, vol. 9, no. 1, pp. 731–738, 2024.

[11] C. Liu, C. Jiang, and H. Wang, "Ingvio: A consistent invariant filter for fast and high-accuracy gnss-visual-inertial odometry," *IEEE Robotics and Automation Letters*, vol. 8, no. 3, pp. 1850–1857, 2023.

[12] M. Brossard, A. Barrau, and S. Bonnabel, "Exploiting symmetries to design ekfs with consistency properties for navigation and slam," *IEEE Sensors Journal*, vol. 19, no. 4, pp. 1572–1579, 2019.

[13] P. Chauchat, A. Barrau, and S. Bonnabel, "Invariant filtering for wheeled vehicle localization with unknown wheel radius and unknown gnss lever arm," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, 2024, pp. 2005–2011.

[14] Z. Ahmed and C. A. Woolsey, "An invariant extended kalman filter for wind estimation using a small, fixed-wing uncrewed aerial vehicle," *IEEE Transactions on Control Systems Technology*, pp. 1–11, 2025.

[15] P. Bernard, *Observer Design for Nonlinear Systems*. Cham, Switzerland: Springer, 2018.

[16] P. van Goor, T. Hamel, and R. Mahony, "Equivariant filter (eqf)," *IEEE Transactions on Automatic Control*, vol. 68, no. 6, pp. 3501–3512, 2023.

[17] R. Mahony, P. van Goor, and T. Hamel, "Observer design for nonlinear systems with equivariance," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 221–252, 2022.

[18] P. van Goor, P. Wickramasinghe, M. Hampsey, and R. Mahony, "Constructive synchronous observer design for inertial navigation with delayed gnss measurements," *European Journal of Control*, vol. 80, p. 101047, 2024, 2024 European Control Conference Special Issue.

[19] P. van Goor and R. Mahony, "Autonomous error and constructive observer design for group affine systems," in *2021 60th IEEE Conference on Decision and Control (CDC)*, 2021, pp. 4730–4737.

[20] S. P. Bhat and D. S. Bernstein, "A topological obstruction to continuous global stabilization of rotational motion and the unwinding phenomenon," *Systems & Control Letters*, vol. 39, no. 1, pp. 63–70, 2000.

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

[22] S. Berkane and A. Tayebi, "On the design of attitude complementary filters on $\mathrm{SO}(3)$," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 880–887, 2018.

[23] M. Wang and A. Tayebi, "Hybrid nonlinear observers for inertial navigation using landmark measurements," *IEEE Transactions on Automatic Control*, vol. 65, no. 12, pp. 5173–5188, 2020.

[24] M. Wang, S. Berkane, and A. Tayebi, "Nonlinear observers design for vision-aided inertial navigation systems," *IEEE Transactions on Automatic Control*, vol. 67, no. 4, pp. 1853–1868, 2022.

[25] M. Wang and A. Tayebi, "Observers design for inertial navigation systems: A brief tutorial," in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 1320–1327.

[26] H. Alnahhal, S. Benahmed, S. Berkane, and T. Hamel, "Attitude estimation using scalar measurements," *IEEE Control Systems Letters*, pp. 1–1, 2025.

[27] S. Benahmed and S. Berkane, "State estimation using single body-frame bearing measurements," in *2024 European Control Conference (ECC)*, 2024, pp. 116–121.

[28] S. Benahmed, S. Berkane, and T. Hamel, "A geometric approach for pose and velocity estimation using imu and inertial/body-frame measurements," 2025. [Online]. Available: https://arxiv.org/abs/2504.03764

[29] S. Benahmed, S. Berkane, and T. Hamel, "A generic observer design for inertial navigation systems using an ltv framework," 2025. [Online]. Available: https://arxiv.org/abs/2410.03846

[30] T. Hamel and C. Samson, "Riccati observers for the nonstationary pnp problem," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 726–741, 2018.

[31] J. M. Lee, *Introduction to Smooth Manifolds*. New York: Springer, 2013.

[32] B. C. Hall, *Lie Groups, Lie Algebras, and Representations An Elementary Introduction*. Cham, Switzerland: Springer, 2015.

[33] C.-T. Chen, *Linear System Theory and Design*, 3rd ed. New York, NY, USA: Oxford University Press, 1999.

[34] P. Salaris, M. Cognetti, R. Spica, and P. R. Giordano, "Online optimal perception-aware trajectory generation," *IEEE Transactions on Robotics*, vol. 35, no. 6, pp. 1307–1322, 2019.

[35] P. Bernard, N. Mimmo, and L. Marconi, "On the semi-global stability of an ek-like filter," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1771–1776, 2021.

[36] M. Pengov, E. Richard, and J.-C. Vivalda, "On the boundedness of the solutions of the continuous riccati equation," *Journal of Inequalities and Applications*, vol. 6, no. 6, pp. 641–649, 2001.

[37] K. Reif, F. Sonnemann, and R. Unbehauen, "An ekf-based nonlinear observer with a prescribed degree of stability," *Automatica*, vol. 34, no. 9, pp. 1119–1123, 1998.

[38] P. van Goor, T. Hamel, and R. Mahony, "Constructive equivariant observer design for inertial navigation," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 2494–2499, 2023, 22nd IFAC World Congress.

[39] P. van Goor, T. Hamel, and R. Mahony, "Equivariant filter (eqf): A general filter design for systems on homogeneous spaces," in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 5401–5408.

[40] J. Levine and R. Marino, "Nonlinear system immersion, observers and finite-dimensional filters," *Systems & Control Letters*, vol. 7, no. 2, pp. 133–142, 1986.

[41] G. Besancon and A. Ticlea, "An immersion-based observer design for rank-observable nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 1, pp. 83–88, 2007.

[42] A. Fornasier, Y. Ng, R. Mahony, and S. Weiss, "Equivariant filter design for inertial navigation systems with input measurement biases," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 4333–4339.

[43] A. Fornasier, "Equivariant symmetries for aided inertial navigation," 2024. [Online]. Available: https://arxiv.org/abs/2407.14297

[44] J. W. Milnor, *Morse Theory*. New Jersey, USA: Princeton University Press, 1969.

[45] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 376–380, 1991.

[46] J. Benhamou, S. Bonnabel, and C. Chapdelaine, "Optimal active sensing control for two-frame systems," in *2023 62nd IEEE Conference on Decision and Control (CDC)*, 2023, pp. 5967–5974.

[47] S. de Marco, M.-D. Hua, T. Hamel, and C. Samson, "Position, velocity, attitude and accelerometer-bias estimation from imu and bearing measurements," in *2020 European Control Conference (ECC)*, 2020, pp. 1003–1008.

[48] M. Wang and A. Tayebi, "Pose, velocity and landmark position estimation using imu and bearing measurements," 2024. [Online]. Available: https://arxiv.org/abs/2407.18099

[49] T. Hamel, M.-D. Hua, and C. Samson, "Deterministic observer design for vision-aided inertial navigation," in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 1306–1313.

[50] M.-D. Hua, S. De Marco, T. Hamel, and R. W. Beard, "Relative pose estimation from bearing measurements of three unknown source points," in *2020 59th IEEE Conference on Decision and Control (CDC)*, 2020, pp. 4176–4181.

[51] S. Berkane, D. Theodosis, T. Hamel, and D. V. Dimarogonas, "State estimation for linear systems with quadratic outputs," *IEEE Control Systems Letters*, vol. 7, pp. 3872–3877, 2023.

[52] L. Rodrigues, "On the relation between observability of a class of linear systems with norm outputs and navigation with range measurements," *IEEE Control Systems Letters*, vol. 9, pp. 1243–1248, 2025.

[53] C. Liu and Y. Shen, "Classification of linear observed systems on multi-frame groups via automorphisms," 2024. [Online]. Available: https://arxiv.org/abs/2412.14673

[54] T. Hamel and C. Samson, "Position estimation from direction or range measurements," *Automatica*, vol. 82, pp. 137–144, 2017.