2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

Observer design for visual-inertial estimation of pose, linear velocity and gravity direction in planar environments

Tarek Bouazza, Tarek Hamel, Claude Samson

***Abstract*— Vision-aided inertial navigation systems combine data from a camera and an IMU to estimate the position, orientation, and linear velocity of a moving vehicle. In planar environments, existing methods assume knowledge of the vertical direction and ground plane to exploit accelerometer measurements. This paper presents a new solution that extends the estimation to arbitrary planar environments. A deterministic Riccati observer is designed to estimate the direction of gravity along with the vehicle pose, linear velocity, and the normal direction to the plane by fusing bearing correspondences from an image sequence with angular velocity and linear acceleration data. Comprehensive observability and stability analysis establishes an explicit persistent excitation condition under which local exponential stability of the observer is achieved. Simulation and real-world experimental results illustrate the performance and robustness of the proposed approach.**

I. Introduction

Estimating the pose of a moving camera, i.e. determining its position and orientation, from visual data is a fundamental challenge in robotics and computer vision. Over the past 15 years, numerous state-of-the-art methods for pose estimation using vision systems have involved Inertial Measurement Units (IMU). The integration of visual information with inertial measurements has given rise to visual-inertial (VI) systems, which combine data from an IMU with measurements from onboard (monocular or stereo) cameras. These systems find applications in vision-aided inertial navigation [1], visual-inertial odometry (VIO) [2] and Simultaneous Localization and Mapping (SLAM).

When estimating the pose using monocular vision, based on landmark bearing (i.e. direction) measurements, the problem is classically referred to as the Perspective-n-Point (PnP) problem when the coordinates of the landmarks in the inertial frame are known. In the cases where the coordinates are unknown, the relative pose can be obtained using the epipolar constraints. These problems have been widely addressed in the literature [3], [4], [5], and have been revisited in [6] and [7], where nonlinear Riccati observers, exploiting first-order approximations and relying on solutions to the *Continuous Riccati Equation* (CRE), are derived.

When the landmarks belong to a planar surface, the relative pose is generally obtained from the algebraic computation or estimation of the homography matrix [8]. The relative pose up to a scale and the plane's normal direction are typically retrieved using homography decomposition (SVD-based [9], [10], or analytical [11] approaches). The homography estimation problem has been extensively studied in the literature. Classical approaches developed by the computer vision community have tackled this problem frame-by-frame by solving algebraic constraints and/or a minimization problem related to the correspondences of image features [8], [12]. More recently, researchers in the field of Control Systems have addressed the homography estimation differently using nonlinear observers designed on the Special Linear group $\mathbf{SL}(3)$ [13], [14], [15], [16]. By exploiting velocity information to interpolate across a sequence of images, these algorithms provide filtering effects and significantly improve individual homography estimates. However, the common assumption used in these algorithms is that the translational velocity is constant in the reference or body-fixed frame due to the lack of linear velocity measurements. While this assumption has proven useful in many practical situations, it is unreliable when the vehicle experiences large accelerations. Recent solutions [1], [17] have attempted to associate homography and acceleration measurements but require restrictive assumptions of a dominant ground-plane environment and a standard reference (vertical) direction in the inertial frame known *a priori* from inertial data. To the authors' understanding, the present work is the first to jointly estimate the pose and direction of gravity with no assumptions on the vehicle motion or the scene.

This paper addresses the problem of determining the pose, linear velocity, and gravity direction of a camera moving while viewing a planar scene, using bearing measurements extracted from a sequence of images coupled with angular rate and linear acceleration data from an IMU. We propose a deterministic Riccati observer, based on the general framework presented in [6], which does not require prior knowledge of the vertical direction, thus allowing for arbitrary scene orientations and initial camera configurations with respect to the inertial frame. By exploiting the planar homography constraint as in [18], we derive the measurement equation of the Riccati error system. Observability and stability analyses demonstrate that the proposed observer is locally exponentially stable, provided that the set of bearings measured in the reference frame is *uniformly consistent* [13] and that the vehicle's motion is sufficiently exciting.

The remainder of the paper is organized as follows. Section II provides an overview of the Riccati observer design framework and definitions related to the uniform

This work has been supported by the French government, through the EUR DS4H Investments in the Future project managed by the National French Agency (ANR) with the reference number ANR-17-EURE-0004 and the ANR-ASTRID Project ASCAR.

T. Bouazza and T. Hamel are with I3S, CNRS, Université Côte d'Azur, Sophia Antipolis, France, T. Hamel is also with Institut Universitaire de France (IUF), bouazza(thamel)@i3s.unice.fr.

C. Samson is with INRIA Sophia Antipolis and I3S, claude.samson@inria.fr.

observability of linear time-varying systems. The problem addressed is formally stated in Section III. The derivation of the proposed observer with associated observability and stability analysis are provided in Section IV. Simulation and experimental results are presented in Sections V and VI, respectively, followed by concluding remarks in Section VII.

II. Preliminaries

A. Mathematical notation

$\mathcal{I} := \{\mathbf{O}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ denotes a north-east-down (NED) right-handed global inertial frame of reference with a fixed origin $\mathbf{O}$ and the canonical basis of $\mathbb{R}^3$.

The Euclidean norm of the vector $x \in \mathbb{R}^n$ is denoted by $|x|$. The set $\mathcal{B}_r^n := \{x \in \mathbb{R}^n : |x| \leq r\}$ is the closed ball in $\mathbb{R}^n$ of radius r. The sphere embedded in $\mathbb{R}^n$ with radius equal to one is denoted by $\mathcal{S}^{n-1} := \{x \in \mathbb{R}^n : |x| = 1\}$. For any $p \in \mathcal{S}^2$, $\pi_p := I_3 - pp^\top$ denotes the projection operator onto the plane orthogonal to p.

The identity matrix of dimension $n \times n$ is denoted by I_n, the null matrix of dimension $n \times m$ is denoted $0_{n,m}$ and $\mathbb{S}_+(n)$ denotes the set of positive definite $n \times n$ matrices.

The special orthogonal group of 3D rotations is defined by $\mathbf{SO}(3) := \{R \in \mathbb{R}^{3\times 3} : RR^\top = R^\top R = I_3, \det(R) = 1\}$, $\mathfrak{so}(3) := \{a_\times : a \in \mathbb{R}^3\}$ denotes its associated Lie algebra, where $a_\times$ denotes the skew-symmetric matrix associated with the cross product, satisfying $a_\times b = a \times b$ for all $b \in \mathbb{R}^3$.

Let f be a vector-valued function that depends on the two variables x and y, and on the time variable t. We write $f = O(|x|^{k_1}|y|^{k_2})$ with $k_1, k_2 \geq 0$ if $\forall t$: $|f(x,y,t)|/(|x|^{k_1}|y|^{k_2}) \leq \kappa < \infty$ in the neighbourhood of $(x,y) = (0,0)$. If f depends only on x and t then we write $f(x,t) = O(|x|^k)$ if $\forall t$: $|f(x,t)|/|x|^k \leq \kappa < \infty$ in the neighbourhood of $x = 0$.

B. Recalls of the Riccati observer design framework

The observer presented here is a modified version of the Riccati observer design framework as introduced in [6]. While some modifications are made, the proof proceeds analogously to the proof of Theorem 3.1 in [6].

Consider a class of nonlinear systems, whose state $\mathbf{x} := [x_1^\top \; x_2^\top]^\top$, with $x_1 \in \mathcal{B}_r^{n_1}$ and $x_2 \in \mathbb{R}^{n_2}$ (with $n_1 + n_2 = n$), evolves according to the following equations:

$$\begin{cases} \dot{\mathbf{x}} = & A(\mathbf{x},t)\mathbf{x} + \mathbf{u} + O(|\mathbf{x}|^2) + O(|\mathbf{x}||\mathbf{u}|) \\ \mathbf{y} = & C(\mathbf{x},t)\mathbf{x} + O(|\mathbf{x}|^2) \end{cases} \tag{1}$$

A is a continuous matrix-valued function uniformly bounded w.r.t. (with respect to) t and uniformly continuous w.r.t. $\mathbf{x}$, of the form

$$A(\mathbf{x},t) := \begin{bmatrix} A_{1,1}(t) & 0_{n_1\times n_2} \\ A_{2,1}(\mathbf{x},t) & A_{2,2}(\mathbf{x},t) \end{bmatrix} \in \mathbb{R}^{n\times n}$$

and C is a continuous matrix-valued function uniformly bounded w.r.t. t and uniformly continuous w.r.t. $\mathbf{x}$.

$$C(\mathbf{x},t) := \begin{bmatrix} C_1(\mathbf{x},t)^\top & C_2(\mathbf{x},t)^\top \end{bmatrix}^\top \in \mathbb{R}^{m\times n}$$

The following input is applied

$$\mathbf{u} := -PC^\top D(t)\mathbf{y} \tag{2}$$

with $P \in \mathbb{S}_+(n)$ solution to the continuous Riccati equation:

$$\dot{P} = AP + PA^\top - PC^\top D(t)CP + S(t), \; P(0) = P_0 \tag{3}$$

with $S(t)$ a bounded continuous symmetric positive definite matrix-valued function and $D(t)$ a bounded continuous symmetric positive semi-definite matrix-valued function.

Then, from Theorem 3.1 and Corollary 3.2 in [6], the equilibrium $\mathbf{x} = 0$ is locally uniformly exponentially stable provided that the pair $(A^\star(t), C^\star(t))$, with $A^\star(t) := A(0,t)$ and $C^\star(t) := C(0,t)$, is *uniformly observable* and $D(t)$ and $S(t)$ are both larger than some positive matrix.

C. Uniform observability of linear time-varying systems

The following definition outlines the well-known uniform observability condition for linear time-varying systems.

Definition 1 (Uniform observability)**.** *The pair $(A(t), C(t))$ is called uniformly observable if there exists $\delta, \mu > 0$ such that $\forall t \geq 0$*

$$W(t, t+\delta) := \frac{1}{\delta}\int_t^{t+\delta} \Phi^\top(s,t)C^\top(s)C(s)\Phi(s,t)ds \geq \mu I_n \tag{4}$$

where $\Phi(s,t)$ is the transition matrix associated with $A(t)$, such that $\frac{d}{dt}\Phi(s,t) = A(t)\Phi(s,t)$, $\forall s \geq t$, $\Phi(t,t) = I_n$. The nonnegative definite matrix-valued function $W(t,t+\delta)$ is known as the observability Gramian.

Assumption 1. *The k-th order time-derivative of A (resp. C) is well defined and bounded on $[0, +\infty)$ up to $k = K \geq 0$ (resp. up to $k = K+1$).*

Define $N_0 := C, N_{k+1} := N_k A + \dot{N}_k, k = 1, \ldots,$ and the set $\mathcal{M}_K$ of matrix-valued functions $M(\cdot)$ of dimension $(q \times n)$ $(q \geq 1)$ composed of row vectors of $N_0(\cdot), N_1(\cdot), \ldots$

Proposition 1 (necessary condition for non-uniform observability)**.** *The pair $(A(t), C(t))$ is not uniformly observable only if the following statement*

$$\begin{aligned} &\forall \delta > 0, \exists \{t_p\}_{p\in\mathbb{N}}, \exists x \in \mathcal{S}^{n-1} : \\ &\lim_{p\to+\infty} \int_0^\delta |M(t_p+s)\Phi(t_p+s,t_p)x|^2 ds = 0 \end{aligned} \tag{5}$$

holds true for all matrix-valued functions in $\mathcal{M}_K$.

This proposition follows directly from [19], proof of Proposition 1, relation (16).

III. Problem formulation

Consider a vehicle equipped with an attached VI sensor consisting of an onboard monocular camera and an IMU providing angular rate and linear acceleration measurements.

Let $\{\mathcal{C}_t\}$ denote the moving body-fixed frame attached at the focal point of the camera, assumed to be the center of mass of the body, and let $\{\mathring{\mathcal{C}}\}$ represent the constant initial frame aligned with the camera's initial configuration ($t = 0$). The IMU frame is assumed perfectly aligned with $\{\mathcal{C}_t\}$.

Let $R \in \mathbf{SO}(3)$ denote the vehicle attitude w.r.t. to the frame $\{\mathring{\mathcal{C}}\}$ and $\xi \in \mathbb{R}^3$ denote the position of the vehicle expressed in the reference frame $\{\mathring{\mathcal{C}}\}$.

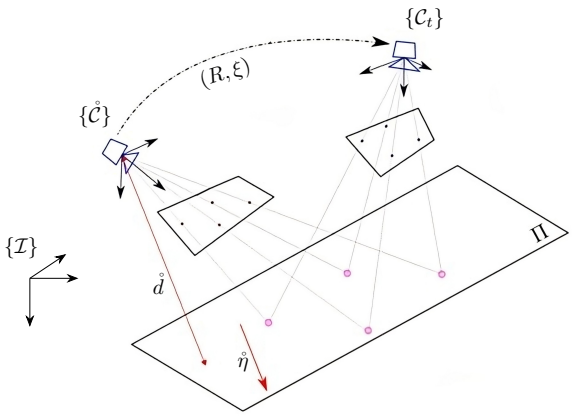

Fig. 1. Euclidean homography depicting the relation among the vehicle's pose, the distance to the plane and the plane's normal vector.

We consider that the attached camera observes a (static) planar scene with arbitrary configuration in the inertial frame. The camera provides the bearing (direction) measurements $p_i \in \mathcal{S}^2$, (resp. $\mathring{p}_i \in \mathcal{S}^2$), $i \in \{1,\dots,l\}$ expressed in frame $\{\mathcal{C}_t\}$ (resp. $\{\mathring{\mathcal{C}}\}$) of l landmarks that belong to the plane. The bearing correspondences $(\mathring{p}_i, p_i)$ of the i-th landmark are related by [16]

$$p_i = \frac{H^{-1}\mathring{p}_i}{|H^{-1}\mathring{p}_i|}, \quad i = \{1,\dots,l\} \tag{6}$$

H is the (Euclidean) homography matrix given by

$$H := R + \frac{\xi\eta^\top}{d} = \left(R^\top - \frac{R^\top\xi\mathring{\eta}^\top}{\mathring{d}}\right)^{-1} \tag{7}$$

where $\mathring{\eta} \in \mathcal{S}^2$ (resp. $\eta \in \mathcal{S}^2$) denotes the unit normal vector pointing towards the target plane expressed in the frame $\{\mathring{\mathcal{C}}\}$ (resp. $\{\mathcal{C}_t\}$), and $\mathring{d} > 0$ (resp. $d > 0$) denotes the distance of the plane to the origin of $\{\mathring{\mathcal{C}}\}$ (resp. $\{\mathcal{C}_t\}$). For a static scene, $\mathring{\eta}$ and $\mathring{d}$ are constant.

The IMU consists of a 3-axis gyrometer that measures the angular velocity $\Omega \in \mathbb{R}^3$, and a 3-axis accelerometer that provides measurements of the *specific acceleration* $a_{\mathcal{S}} \in \mathbb{R}^3$, both expressed in the body-fixed frame $\{\mathcal{C}_t\}$. The kinematics of the linear velocity $V \in \mathbb{R}^3$ expressed in $\{\mathcal{C}_t\}$ are

$$\dot{V} = -\Omega_\times V + a_{\mathcal{S}} + g\gamma \tag{8}$$

where g denotes the gravity constant (typically $g \approx 9.81 m/s^2$) and $\gamma \in \mathcal{S}^2$ denotes the gravitational acceleration direction expressed in $\{\mathcal{C}_t\}$.

IV. Riccati observer design

A. Observer derivation

The kinematic equations expressing the rigid-body motion of the vehicle in the reference frame $\{\mathring{\mathcal{C}}\}$ are given by

$$\dot{R} = R\Omega_\times, \qquad \dot{\xi} = v \tag{9}$$

with $v := RV$ denoting the linear velocity expressed in the reference frame $\{\mathring{\mathcal{C}}\}$, one verifies that

$$\dot{v} = Ra_{\mathcal{S}} + g\mathring{\gamma} \tag{10}$$

where $\mathring{\gamma} := R\gamma$ denotes the (constant) gravity direction expressed in $\{\mathring{\mathcal{C}}\}$. We consider in this work that both $\mathring{\gamma}$ and γ are unknown.

The vehicle position can only be extracted from the homography matrix (7) up to a scale factor. To account for this scaling, we define the inverse of the distance to the scene $\rho := 1/\mathring{d}$. Then, the vehicle position and velocity (up to a scale) are $\bar{\xi} := \xi/\mathring{d} = \rho\xi$ and $\bar{v} := v/\mathring{d} = \rho v$, respectively.

Our aim is to design a nonlinear observer to estimate $(R, \mathring{\gamma}, \mathring{\eta}, \rho, \bar{v}, \bar{\xi}) \in \mathbf{SO}(3) \times (\mathcal{S}^2)^2 \times \mathbb{R}^7$. Given that $\mathring{\gamma}$, $\mathring{\eta}$ and ρ are constant quantities and using (9) and (10), we obtain the following system dynamics

$$\begin{cases} \dot{R} = R\Omega_\times \\ \dot{\mathring{\gamma}} = 0_{3,1} \\ \dot{\mathring{\eta}} = 0_{3,1} \\ \dot{\rho} = 0 \\ \dot{\bar{v}} = \rho R a_{\mathcal{S}} + \rho g \mathring{\gamma} \\ \dot{\bar{\xi}} = \bar{v} \end{cases} \tag{11}$$

The direction vectors $\mathring{\gamma}, \mathring{\eta} \in \mathcal{S}^2$ are over-parametrized by introducing rotation matrices $G, Q \in \mathbf{SO}(3)$, such that

$$\mathring{\gamma} := G^\top \mathbf{e}_3, \qquad \mathring{\eta} := Q^\top \mathbf{e}_3 \tag{12}$$

where the downward direction $\mathbf{e}_3$ of the inertial frame serves as the reference direction.

Remark 1. *This parametrization allows one to lift the kinematics of $\mathcal{S}^2$ on $\mathbf{SO}(3)$ in order to avoid error system complexities in first-order approximations that arise when using minimal parametrizations such as spherical coordinates (for a comprehensive discussion, refer to [20]).*

Let $\hat{G}, \hat{Q} \in \mathbf{SO}(3)$ denotes the estimates of G and Q, respectively. The estimates $\hat{\mathring{\gamma}}, \hat{\mathring{\eta}} \in \mathcal{S}^2$ are obtained by projecting onto $\mathcal{S}^2$, expressed as $\hat{\mathring{\gamma}} := \hat{G}^\top \mathbf{e}_3$ and $\hat{\mathring{\eta}} := \hat{Q}^\top \mathbf{e}_3$. Then, the convergence of $\hat{G}^\top \mathbf{e}_3$ (resp. $\hat{Q}^\top \mathbf{e}_3$) to $G^\top \mathbf{e}_3$ (resp. $Q^\top \mathbf{e}_3$) is equivalent to the one of $\hat{\mathring{\gamma}}$ (resp. $\hat{\mathring{\eta}}$) to $\mathring{\gamma}$ (resp. $\mathring{\eta}$).

Let $\hat{R} \in \mathbf{SO}(3)$, $\hat{\bar{\xi}}, \hat{\bar{v}} \in \mathbb{R}^3$, $\hat{\rho} \in \mathbb{R}$ denote the estimates of $R, \bar{\xi}, \bar{v}, \rho$, respectively. The proposed observer is given by

$$\begin{cases} \dot{\hat{R}} = \hat{R}\Omega_\times + \sigma_{R\times}\hat{R} \\ \dot{\hat{G}} = \sigma_{G\times}\hat{G} \\ \dot{\hat{Q}} = -\sigma_{Q\times}\hat{Q} \\ \dot{\hat{\rho}} = -\sigma_\rho \\ \dot{\hat{\bar{v}}} = \hat{\rho}\hat{R}a_{\mathcal{S}} + \hat{\rho}g\hat{G}^\top\mathbf{e}_3 - \sigma_v \\ \dot{\hat{\bar{\xi}}} = \hat{\bar{v}} - \sigma_\xi \end{cases} \tag{13}$$

where $\sigma_R, \sigma_G, \sigma_Q, \sigma_\xi, \sigma_v \in \mathbb{R}^3$, $\sigma_\rho \in \mathbb{R}$ are the innovation terms to be determined according to the design approach outlined in II-B. Define the following observer errors:

$$\tilde{R} := \hat{R}R^\top,\ \tilde{G} := \hat{G}G^\top,\ \tilde{Q} := Q\hat{Q}^\top \in \mathbf{SO}(3),$$
$$\tilde{\bar{\xi}} := \tilde{R}\bar{\xi} - \hat{\bar{\xi}},\ \tilde{\bar{v}} := \tilde{R}\bar{v} - \hat{\bar{v}} \in \mathbb{R}^3,\ \tilde{\rho} := \rho - \hat{\rho} \in \mathbb{R}$$

Remark 2. *The observer errors of the elements $(R, \bar{\xi}, \bar{v}) \in \mathbf{SO}(3) \times \mathbb{R}^3 \times \mathbb{R}^3$ are defined using the $\mathbf{SE}_2(3)$ Lie group*

error construction [21] which provides a more natural representation of the composition of extended poses.

To derive the Riccati error system, first-order approximations are required. The approximations of the error equations that involve rotation matrices are derived using a (local) minimal parametrization of $\mathbf{SO}(3)$. The chosen parametrization in this work is the vector part of the Rodrigues unit quaternion $\mathbf{q} = [q_0, q_v]$ associated with $\mathbf{R} \in \mathbf{SO}(3)$. Rodrigues' formula that relates $\mathbf{q}$ to $\mathbf{R}$ is given by

$$\mathbf{R}(\mathbf{q}) = I_3 + 2q_{v\times}(q_0 I_3 + q_{v\times})$$

From this relation, one can write

$$\mathbf{R} = I_3 + \lambda_\times + O(|\lambda|^2), \quad \text{with } \lambda := 2\,\mathrm{sign}(q_0)q_v \in \mathcal{B}_2^3$$

One obtains the following first-order approximations

$$\begin{aligned} \tilde{R} &= I_3 + \lambda_{R\times} + O(|\lambda_R|^2) \\ \tilde{G} &= I_3 + \lambda_{G\times} + O(|\lambda_G|^2) \\ \tilde{Q} &= I_3 + \lambda_{Q\times} + O(|\lambda_Q|^2) \end{aligned} \tag{14}$$

where $\lambda_R, \lambda_G, \lambda_Q \in \mathcal{B}_2^3$ denote the (linearized) states associated with the errors of R, G and Q, respectively. Given equation (14) and the first three equations of (11) and (13) one derives the error dynamics of $\tilde{R}$, $\tilde{G}$ and $\tilde{Q}$ as follows

$$\dot{\tilde{R}} = \sigma_{R\times}\tilde{R}, \qquad \dot{\tilde{G}} = \sigma_{G\times}\tilde{G}, \qquad \dot{\tilde{Q}} = \tilde{Q}\sigma_{Q\times}$$

It follows that, in first-order approximations, the dynamics satisfy the following equations

$$\begin{aligned} &\dot{\lambda}_R = \sigma_R + O(|\lambda_R||\sigma_R|), \quad \dot{\lambda}_G = \sigma_G + O(|\lambda_G||\sigma_G|), \\ &\dot{\lambda}_Q = \sigma_Q + O(|\lambda_Q||\sigma_Q|) \end{aligned} \tag{15}$$

For the dynamics of $\tilde{\bar{\xi}}$, from (11) and (13), we have

$$\begin{aligned} \dot{\tilde{\bar{\xi}}} = \dot{\tilde{R}}\bar{\xi} + \tilde{R}\dot{\bar{\xi}} - \dot{\hat{\bar{\xi}}} &= \sigma_{R\times}\tilde{R}\bar{\xi} + \tilde{R}\bar{v} - \hat{\bar{v}} + \sigma_\xi \\ &= \sigma_{R\times}(\tilde{\bar{\xi}} + \hat{\bar{\xi}}) + \tilde{\bar{v}} + \sigma_\xi \\ &= \tilde{\bar{v}} + \bar{\sigma}_\xi + O(|\tilde{\bar{\xi}}||\sigma_R|) \end{aligned} \tag{16}$$

with $\bar{\sigma}_\xi := \sigma_\xi + \sigma_{R\times}\hat{\bar{\xi}} \in \mathbb{R}^3$.
Similarly, the dynamics of $\tilde{\bar{v}}$ can be expressed as

$$\begin{aligned} \dot{\tilde{\bar{v}}} &= \dot{\tilde{R}}\bar{v} + \tilde{R}\dot{\bar{v}} - \dot{\hat{\bar{v}}} \\ &= \sigma_{R\times}\tilde{R}\bar{v} + \rho\tilde{R}\,(Ra_{\mathcal{S}} + g\mathring{\gamma}) - \hat{\rho}\hat{R}a_{\mathcal{S}} - \hat{\rho}g\hat{G}^\top\mathbf{e}_3 + \sigma_v \\ &= \sigma_{R\times}(\tilde{\bar{v}} + \hat{\bar{v}}) + (\rho - \hat{\rho})\,\hat{R}a_{\mathcal{S}} + g(\rho\tilde{R}G^\top - \hat{\rho}\hat{G}^\top)\mathbf{e}_3 + \sigma_v \\ &= g\left(\rho\tilde{R}\tilde{G}\hat{G}^\top - \hat{\rho}\hat{G}^\top\right)\mathbf{e}_3 + \tilde{\rho}\hat{R}a_{\mathcal{S}} + \bar{\sigma}_v + O(|\tilde{\bar{v}}||\sigma_R|) \end{aligned}$$

with $\bar{\sigma}_v := \sigma_v + \sigma_{R\times}\hat{\bar{v}} \in \mathbb{R}^3$. Substituting $\tilde{R}$ and $\tilde{G}$ with the first-order approximations (14), this simplifies to

$$\begin{aligned} \dot{\tilde{\bar{v}}} &= g\left((\tilde{\rho} + \hat{\rho})(I_3 + \lambda_{R\times})\hat{G}^\top(I_3 + \lambda_{G\times}) - \hat{\rho}\hat{G}^\top\right)\mathbf{e}_3 \\ &\quad + \tilde{\rho}\hat{R}a_{\mathcal{S}} + \bar{\sigma}_v + O(|\tilde{\bar{v}}||\sigma_R|) \\ &= \tilde{\rho}\hat{R}a_{\mathcal{S}} + g(\tilde{\rho}\hat{G}^\top + \lambda_{R\times}\hat{\rho}\hat{G}^\top + \hat{\rho}\hat{G}^\top\lambda_{G\times})\mathbf{e}_3 + \bar{\sigma}_v \\ &\quad + O(|\lambda_R||\lambda_G|) + O(|\lambda_R||\tilde{\rho}|) + O(|\lambda_G||\tilde{\rho}|) + O(|\tilde{\bar{v}}||\sigma_R|) \\ &= -\hat{\rho}g(\hat{G}^\top\mathbf{e}_3)_\times\lambda_R - \hat{\rho}g\hat{G}^\top(\mathbf{e}_2\lambda_{G,1} - \mathbf{e}_1\lambda_{G,2}) + \dot{\hat{v}}\tilde{\rho} + \bar{\sigma}_v \\ &\quad + O(|\lambda_R||\lambda_G|) + O(|\lambda_R||\tilde{\rho}|) + O(|\lambda_G||\tilde{\rho}|) + O(|\tilde{\bar{v}}||\sigma_R|) \end{aligned} \tag{17}$$

where $\lambda_{G,1}$ and $\lambda_{G,2}$ are the first two components of λ_G and $\dot{\hat{v}} := (\hat{R}a_{\mathcal{S}} + g\hat{\mathring{\gamma}})$.

To derive the measurement equation, let $\hat{H}$ denote the estimate of the homography, computed from the individual estimates of its components $(\hat{R}, \hat{\bar{\xi}}, \hat{\mathring{\eta}})$. In view of (7), one shows that $\hat{H} = \hat{\Gamma}\hat{R}$, with $\hat{\Gamma} := (I_3 - \hat{\bar{\xi}}\hat{\mathring{\eta}}^\top)^{-1}$. Defining $\hat{\mathring{p}}_i := \frac{\hat{H}p_i}{|\hat{H}p_i|} \in \mathcal{S}^2$ as the estimates of $\mathring{p}_i$, and exploiting the constraint given in (6), yield

$$\hat{\mathring{p}}_i = \frac{\hat{H}p_i}{|\hat{H}p_i|} = \frac{\hat{H}\frac{H^{-1}\mathring{p}_i}{|H^{-1}\mathring{p}_i|}}{|\hat{H}\frac{H^{-1}\mathring{p}_i}{|H^{-1}\mathring{p}_i|}|} = \frac{\tilde{H}\mathring{p}_i}{|\tilde{H}\mathring{p}_i|} \tag{18}$$

where $\tilde{H} := \hat{H}H^{-1}$ denotes the homography error. Using the fact that $\pi_{\hat{\mathring{p}}_i}\hat{\mathring{p}}_i = 0, \forall i \in \{1, \dots, l\}$, one has

$$\begin{aligned} 0 = |\tilde{H}\mathring{p}_i|\pi_{\hat{\mathring{p}}_i}\hat{\mathring{p}}_i = \pi_{\hat{\mathring{p}}_i}\tilde{H}\mathring{p}_i &= \pi_{\hat{\mathring{p}}_i}\hat{H}H^{-1}\mathring{p}_i \\ &= \pi_{\hat{\mathring{p}}_i}\hat{\Gamma}\hat{R}R^\top(I_3 - \bar{\xi}\mathring{\eta}^\top)\mathring{p}_i \\ &= \pi_{\hat{\mathring{p}}_i}\hat{\Gamma}(\tilde{R} - (\tilde{\bar{\xi}} + \hat{\bar{\xi}})\mathbf{e}_3^\top\tilde{Q}\hat{Q})\mathring{p}_i \end{aligned}$$

using the first-order approximations (14), it follows that

$$\begin{aligned} 0 &= \pi_{\hat{\mathring{p}}_i}\hat{\Gamma}\left(I_3 + \lambda_{R\times} - (\tilde{\bar{\xi}} + \hat{\bar{\xi}})\mathbf{e}_3^\top(I_3 + \lambda_{Q\times})\hat{Q}\right)\mathring{p}_i \\ &\quad + O(|\lambda_R|^2) + O(|\lambda_Q|^2) \\ &= \pi_{\hat{\mathring{p}}_i}\mathring{p}_i - \pi_{\hat{\mathring{p}}_i}\hat{\Gamma}\left(-\lambda_{R\times} + \tilde{\bar{\xi}}\mathbf{e}_3^\top\hat{Q} + \hat{\bar{\xi}}\mathbf{e}_3^\top\lambda_{Q\times}\hat{Q}\right)\mathring{p}_i \\ &\quad + O(|\lambda_R|^2) + O(|\lambda_Q|^2) + O(|\tilde{\bar{\xi}}||\lambda_Q|) \end{aligned}$$

Then, by setting $y_i := \pi_{\hat{\mathring{p}}_i}\mathring{p}_i$ for all $i = 1, \dots, l$, one deduces

$$\begin{aligned} y_i = (\hat{Q}\mathring{p}_i)_2\pi_{\hat{\mathring{p}}_i}\hat{\Gamma}\hat{\bar{\xi}}\lambda_{Q,1} - (\hat{Q}\mathring{p}_i)_1\pi_{\hat{\mathring{p}}_i}\hat{\Gamma}\hat{\bar{\xi}}\lambda_{Q,2} + \pi_{\hat{\mathring{p}}_i}\hat{\Gamma}\mathring{p}_{i\times}\lambda_R \\ +(\hat{Q}\mathring{p}_i)_3\pi_{\hat{\mathring{p}}_i}\hat{\Gamma}\tilde{\bar{\xi}} + O(|\lambda_R|^2) + O(|\lambda_Q|^2) + O(|\tilde{\bar{\xi}}||\lambda_Q|) \end{aligned}$$

where $\lambda_{Q,1}, \lambda_{Q,2}$ denote the first two components of λ_Q.

Defining the output vector as $\mathbf{y} := [y_1^\top, \dots, y_l^\top]^\top$, we obtain LTV first-order approximations in the form (1) with

$$\begin{cases} \mathbf{x} := [\lambda_R^\top,\ \lambda_{G,1},\ \lambda_{G,2},\ \lambda_{Q,1},\ \lambda_{Q,2},\ \tilde{\rho},\ \tilde{\bar{v}}^\top,\ \tilde{\bar{\xi}}^\top]^\top \\ \mathbf{u} := [\sigma_R^\top,\ \sigma_{G,1},\ \sigma_{G,2},\ \sigma_{Q,1},\ \sigma_{Q,2},\ \sigma_\rho,\ \bar{\sigma}_v^\top,\ \bar{\sigma}_\xi^\top]^\top \\ A(\mathbf{x},t) := \begin{bmatrix} 0_{8,3} & 0_{8,1} & 0_{8,1} & 0_{8,2} & 0_{8,1} & 0_{8,3} & 0_{8,3} \\ \hat{\alpha}_1 & \hat{\alpha}_2 & \hat{\alpha}_3 & 0_{3,2} & \dot{\hat{v}} & 0_{3,3} & 0_{3,3} \\ 0_{3,3} & 0_{3,1} & 0_{3,1} & 0_{3,2} & 0_{3,1} & I_3 & 0_{3,3} \end{bmatrix} \\ C(\mathbf{x},t) := \begin{bmatrix} \hat{\beta}_1 & 0_{3l,2} & \hat{\beta}_2 & \hat{\beta}_3 & 0_{3l,4} & \hat{\beta}_4 \end{bmatrix} \end{cases} \tag{19}$$

with $\hat{\alpha}_1 = -\hat{\rho}g(\hat{G}^\top\mathbf{e}_3)_\times$, $\hat{\alpha}_2 = -\hat{\rho}g\hat{G}^\top\mathbf{e}_2$, $\hat{\alpha}_3 = \hat{\rho}g\hat{G}^\top\mathbf{e}_1$, $\hat{\beta}_k = [\hat{\beta}_{k,1}^\top \dots \hat{\beta}_{k,l}^\top]^\top, k = 1, \dots, 4$, $\hat{\beta}_{1,i} = \pi_{\hat{\mathring{p}}_i}\hat{\Gamma}\mathring{p}_{i\times}$, $\hat{\beta}_{2,i} = (\hat{Q}\mathring{p}_i)_2\pi_{\hat{\mathring{p}}_i}\hat{\Gamma}\hat{\bar{\xi}}$, $\hat{\beta}_{3,i} = -(\hat{Q}\mathring{p}_i)_1\pi_{\hat{\mathring{p}}_i}\hat{\Gamma}\hat{\bar{\xi}}$, $\hat{\beta}_{4,i} = (\hat{Q}\mathring{p}_i)_3\pi_{\hat{\mathring{p}}_1}\hat{\Gamma}$.

From there, the expressions of the innovation terms are directly computed according to equations (2) and (3), where the matrices S and D (involved in (3)) are chosen larger than some constant positive matrix.

Remark 3. *While $\tilde{G}$ and $\tilde{Q}$ are elements of $\mathbf{SO}(3)$, only the first two components of λ_G and λ_Q appear in system (19), which results in a minimal state representation in the first-order approximations. Accordingly, the last components of the innovation terms σ_G and σ_Q are set to zero for simplicity.*

B. Observability and stability analysis

According to Corollary 3.2 in [6], good conditioning of the solutions $P(t)$ to the CRE (3) and the exponential stability of the error origin of the observer (13) rely on the uniform observability of $(A^\star(t), C^\star(t))$ obtained by setting $\mathbf{x} = 0$ in the expressions of the matrices $A(\mathbf{x}, t)$ and $C(\mathbf{x}, t)$ given in (19). The corresponding matrices are

$$\begin{cases} A^\star(t) := \begin{bmatrix} 0_{8,3} & 0_{8,1} & 0_{8,1} & 0_{8,2} & 0_{8,1} & 0_{8,3} & 0_{8,3} \\ \alpha_1 & \alpha_2 & \alpha_3 & 0_{3,2} & \dot{v}(t) & 0_{3,3} & 0_{3,3} \\ 0_{3,3} & 0_{3,1} & 0_{3,1} & 0_{3,2} & 0_{3,1} & I_3 & 0_{3,3} \end{bmatrix} \\ C^\star(t) := \begin{bmatrix} \beta_1 & 0_{3l,2} & \beta_2 & \beta_3 & 0_{3l,4} & \beta_4 \end{bmatrix} \end{cases} \tag{20}$$

with $\alpha_1 = -\rho g (G^\top \mathbf{e}_3)_\times$, $\alpha_2 = -\rho g G^\top \mathbf{e}_2$, $\alpha_3 = \rho g G^\top \mathbf{e}_1$, $\beta_k = [\beta_{k,1}^\top \dots \beta_{k,l}^\top]^\top, k = 1, \dots, 4$, $\beta_{1,i} = \pi_{\mathring{p}_i} \Gamma \mathring{p}_{i\times}$, $\beta_{2,i} = (Q\mathring{p}_i)_2 \pi_{\mathring{p}_i} \Gamma \bar{\xi}$, $\beta_{3,i} = -(Q\mathring{p}_i)_1 \pi_{\mathring{p}_i} \Gamma \bar{\xi}$, $\beta_{4,i} = (Q\mathring{p}_i)_3 \pi_{\mathring{p}_1} \Gamma$.

Definition 2 (Uniformly consistent set[1])**.** *A set $\mathcal{M}_l$ of $l \geq 4$ constant vector directions $\mathring{p}_i \in \mathcal{S}^2$, $(i = 1 \dots l)$ is termed uniformly consistent if it contains a subset $\mathcal{M}_4 \subset \mathcal{M}_l$ of four vector directions such that all its vector triplets are uniformly linearly independent; that is, there exists $\epsilon > 0$ such that* $\det([\mathring{p}_i\ \mathring{p}_j\ \mathring{p}_k]) = \mathring{p}_i \cdot (\mathring{p}_j \times \mathring{p}_k) \geq \epsilon$ *for all $\mathring{p}_i, \mathring{p}_j, \mathring{p}_k \in \mathcal{M}_4$.*

Assumption 2. *The set $\mathcal{M}_l$ of the measured directions $\mathring{p}_i \in \mathcal{S}^2$ is uniformly consistent, and there exists $\varepsilon > 0$ such that $|\bar{\xi}(t)| \geq \varepsilon$ for all $t \geq 0$.*

Proposition 2. *If Assumption 2 holds and $\dot{v}(t)$ is uniformly continuous and bounded and is "persistently exciting" in the sense that there exists $\delta, \mu > 0$ such that $\forall t \geq 0$,*

$$\frac{1}{\delta} \int_t^{t+\delta} |\dot{v}(s)|^2 ds \geq \mu. \tag{21}$$

Then, the pair $(A^\star(t), C^\star(t))$ is uniformly observable and the equilibrium of the proposed Riccati observer is locally exponentially stable.

The proof is given in the Appendix.

V. Simulation Results

This section provides simulation results to illustrate the performance of the observer (13). The simulation scenario involves a square planar target (with four known corner directions) tilted at a 45-degree angle relative to the inertial frame and a camera that follows a Lissajous reference trajectory with linear acceleration and angular velocity measured in the body-fixed frame (subject to Gaussian additive noise).

We conduct a Monte-Carlo simulation with 100 runs, the initial estimation errors are randomly generated for each run using Gaussian distributions with the following average values: $\tilde{R}(0)$ corresponds to $22.5(\text{deg})$ errors in roll, pitch and yaw, $\tilde{G}(0)$ (resp. $\tilde{Q}(0)$) corresponds to a $15(\text{deg})$ angle between $\mathring{\gamma}(0)$ (resp. $\mathring{\eta}(0)$) and $\hat{\mathring{\gamma}}(0)$ (resp.$\hat{\mathring{\eta}}(0)$), $\bar{\xi}(0) = [0.2, 0.2, 0.2]$, $\tilde{v}(0) = [0.1, 0.1, 0.1]$, $\tilde{\rho}(0) = 0.1$. The following initial Riccati matrix and parameters are chosen: $P(0) = \text{diag}(I_7, 0.4, 2I_6)$, $D = 25 I_{12}$, $S = 0.1 I_{14}$.

[1]This definition involves a stricter *uniformity* condition than that in [13].

The time evolutions of the average estimation errors of the attitude (defined by $\text{trace}(I_3 - \tilde{R})$), the normal to the planar scene (defined by $1 - \hat{\mathring{\eta}}^\top \mathring{\eta}$) and the gravity vector (defined by $1 - \hat{\mathring{\gamma}}^\top \mathring{\gamma}$) are shown in Fig. 2. The average estimation errors of the scaled position and velocity, as well as the inverse of the distance, are shown in Fig. 3. The shaded areas illustrate the 5th to 95th percentile of error. From these figures, one can see that all the estimation errors converge to zero after a short transient period for the generated initial estimates.

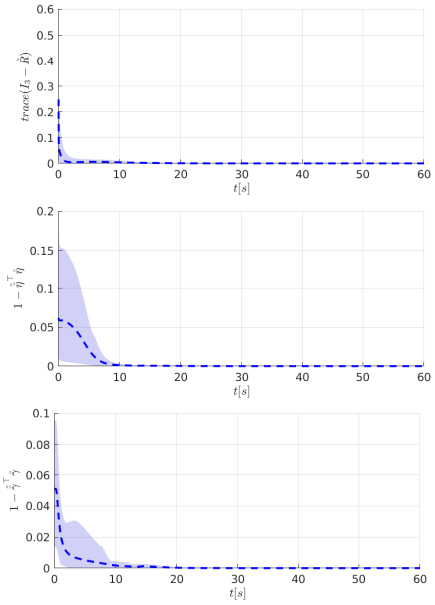

Fig. 2. Average estimation errors of the attitude, normal to the scene vector and gravity vector, respectively.

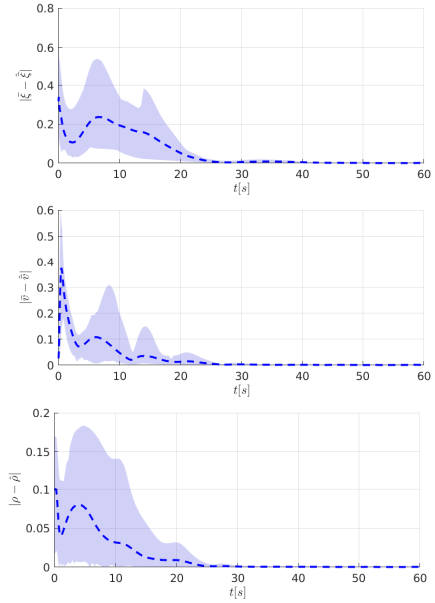

Fig. 3. Average estimation errors of the scaled position, scaled velocity and inverse of the distance, respectively.

VI. Experimental Results

In this section, experimental results are reported to support the simulation results and to demonstrate the robustness and effectiveness of the observer on real-world data.

We perform an online experiment with a mounted Camera-IMU setup viewing a textured planar target on the ground; the system is initially aligned nearly perpendicularly to the target (i.e., $\mathring{\gamma} \simeq \mathring{\eta} \simeq \mathbf{e}_3$). Its motion is then recorded as it is moved in various directions and orientations. The setup includes a Basler camera capturing 25 frames per second at a resolution of 1200×1024 pixels, and an MPU-9250 IMU operating at 100 Hz. The estimation results are compared to ground-truth data collected at 200 Hz using an OptiTrack motion capture system (MOCAP). The observer is implemented in C++, and the collected data are time-synchronized using ROS.

The following initial state estimates and Riccati matrix are considered for the experiment: $\hat{R}_0 = \hat{G}_0 = \hat{Q}_0 = I_3$, $\hat{\rho}_0 = 1$, $\hat{v}_0 = \hat{\xi}_0 = [0,0,0]^\top$ and $P_0 = \mathrm{diag}(0.5I_3; 0.2I_4; 0.2; 0.75I_6)$. The matrices involved were chosen as follows: $D = 10I_{3l}$, $S = 0.01I_{14}$.

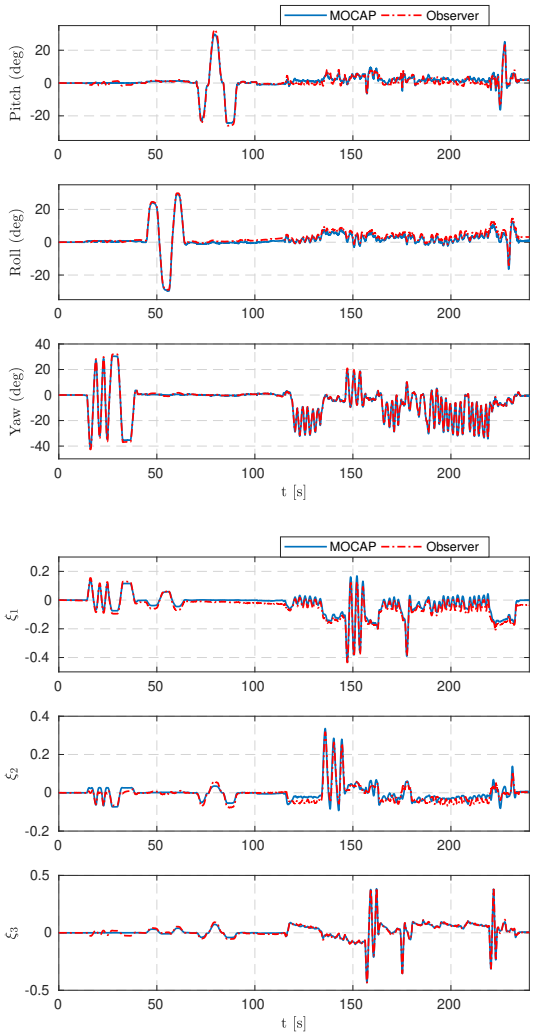

Fig. 4. Attitudes (Euler angles) and scaled positions given by the proposed observer (red) and MOCAP ground truth (blue).

The plots in Fig. 4 show the true and estimated attitude (represented by the Euler angles) and scaled position. It can be observed from these plots that the resulting estimates evolve consistently near the corresponding ground truths and remain particularly stable and noise-free as a result of the filtering property of the proposed observer.

VII. Conclusions

In this paper, the problem of estimating the pose, linear velocity and gravity direction of a vehicle observing a planar target with arbitrary configuration has been addressed. The proposed solution exploits direct bearing correspondences from a sequence of images satisfying the planar homography constraint, along with angular rate and linear acceleration measurements, to derive a nonlinear observer based on a recent Riccati observer design framework. A key contribution lies in establishing an explicit persistent excitation condition under which the local exponential stability of the observer is guaranteed. The theoretical results were supported by extensive simulation and experiments, demonstrating the performance and robustness of the presented approach.

References

[1] S. Zhao, F. Lin, K. Peng, B. Chen, and T. Lee, "Homography-based vision-aided inertial navigation of uavs in unknown environments," in *AIAA Guidance, Navigation, and Control Conference*, 2012, p. 5033.

[2] M. Bloesch, S. Omari, M. Hutter, and R. Siegwart, "Robust visual inertial odometry using a direct ekf-based approach," in *2015 IEEE/RSJ international conference on intelligent robots and systems (IROS)*. IEEE, 2015, pp. 298–304.

[3] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[4] B. M. Haralick, C.-N. Lee, K. Ottenberg, and M. Nölle, "Review and analysis of solutions of the three point perspective pose estimation problem," *International journal of computer vision*, vol. 13, pp. 331–356, 1994.

[5] D. Nistér, "An efficient solution to the five-point relative pose problem," *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 6, pp. 756–770, 2004.

[6] T. Hamel and C. Samson, "Riccati observers for the nonstationary pnp problem," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 726–741, 2017.

[7] 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)*. IEEE, 2020, pp. 4176–4181.

[8] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[9] O. D. Faugeras and F. Lustman, "Motion and structure from motion in a piecewise planar environment," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 2, no. 03, pp. 485–508, 1988.

[10] Z. Zhang and A. R. Hanson, "3d reconstruction based on homography mapping," *Proc. ARPA96*, pp. 1007–1012, 1996.

[11] E. Malis and M. Vargas Villanueva, "Deeper understanding of the homography decomposition for vision-based control," 2007.

[12] J. Y. Kaminski and A. Shashua, "Multiple view geometry of general algebraic curves," *International Journal of Computer Vision*, vol. 56, no. 3, pp. 195–219, 2004.

[13] T. Hamel, R. Mahony, J. Trumpf, P. Morin, and M.-D. Hua, "Homography estimation on the special linear group based on direct point correspondence," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*. IEEE, 2011, pp. 7902–7908.

[14] R. Mahony, T. Hamel, P. Morin, and E. Malis, "Nonlinear complementary filters on the special linear group," *International Journal of Control*, vol. 85, no. 10, pp. 1557–1573, 2012.

[15] M.-D. Hua, J. Trumpf, T. Hamel, R. Mahony, and P. Morin, "Feature-based recursive observer design for homography estimation and its application to image stabilization," *Asian Journal of Control*, vol. 21, no. 4, pp. 1443–1458, 2019.

[16] ——, "Nonlinear observer design on sl (3) for homography estimation by exploiting point and line correspondences with application to image stabilization," *Automatica*, vol. 115, p. 108858, 2020.

[17] V. Grabe, H. H. Bülthoff, and P. R. Giordano, "A comparison of scale estimation schemes for a quadrotor uav based on optical flow and imu measurements," in *2013 IEEE/RSJ international conference on intelligent robots and systems*. IEEE, 2013, pp. 5193–5200.
[18] T. Bouazza, T. Hamel, M. D. Hua, and R. Mahony, "Homography-based riccati observer design for camera pose estimation," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 6862–6868.
[19] P. Morin, A. Eudes, and G. Scandaroli, "Uniform observability of linear time-varying systems and application to robotics problems," in *Geometric Science of Information: Third International Conference, GSI 2017, Paris, France, November 7-9, 2017, Proceedings 3*. Springer, 2017, pp. 336–344.
[20] M.-D. Hua and G. Allibert, "Riccati observer design for pose, linear velocity and gravity direction estimation using landmark position and imu measurements," in *2018 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, 2018, pp. 1313–1318.
[21] 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, 2016.

APPENDIX

Proof of Proposition 2. Using the Peano–Baker series, one verifies that the transition matrix associated with $A^\star$ is

$$\Phi(t+s,t)=\begin{bmatrix} I_8 & & & & & 0_{8,3} & 0_{8,3} \\ \alpha_1 s & \alpha_2 s & \alpha_3 s & 0_{3,2} & a(t+s,t) & I_3 & 0_{3,3} \\ \alpha_1 \frac{s^2}{2} & \alpha_2 \frac{s^2}{2} & \alpha_3 \frac{s^2}{2} & 0_{3,2} & b(t+s,t) & sI_3 & I_3 \end{bmatrix}$$

where $a(t+s,t) := v(t+s)-v(t)$, $b(t+s,t) := \xi(t+s)-\xi(t)-sv(t)$. In view of (20), one can express $C^\star(t)$ as $C^\star(t)=Z(t)\bar{C}$, with $Z(t) := \begin{bmatrix}\beta_1(t) & \beta_2(t) & \beta_3(t) & \beta_4(t)\end{bmatrix}$ and

$$\bar{C} := \begin{bmatrix} I_3 & 0_{3,2} & 0_{3,2} & 0_{3,4} & 0_{3,3} \\ 0_{2,3} & 0_{2,2} & I_2 & 0_{2,4} & 0_{2,3} \\ 0_{3,3} & 0_{3,2} & 0_{3,2} & 0_{3,4} & I_3 \end{bmatrix}$$

Define $\bar{Z}(t) := Z^\top(t)Z(t)$. Then, the observability Gramian associated with the pair $(A^\star, C^\star)$ is of the form

$$W(t,t+\delta)=\frac{1}{\delta}\int_t^{t+\delta}\Phi^\top(s,t)\bar{C}^\top\bar{Z}(s)\bar{C}\Phi(s,t)ds \quad (22)$$

As proven in [18], Assumption 2 establishes a sufficient condition to ensure the full column rank of $Z(t)$. Moreover, this assumption guarantees the existence of $\bar{\epsilon}>0$ such that $\lambda_{\min}(\bar{Z}(t)) \geq \bar{\epsilon}$ for all $t \geq 0$. Then, it follows that $\bar{Z}(t) \geq \bar{\epsilon} I_8$ for all $t \geq 0$. This directly implies from (22) that $W(t,t+\delta) \geq \bar{\epsilon}\bar{W}(t,t+\delta)$, where

$$\bar{W}(t,t+\delta) := \frac{1}{\delta}\int_t^{t+\delta}\Phi^\top(s,t)\bar{C}^\top\bar{C}\Phi(s,t)ds$$

represents the observability Gramian of the pair $(A^\star, \bar{C})$. Consequently, ensuring the uniform observability of the pair $(A^\star, C^\star)$ is equivalent to ensuring that $(A^\star, \bar{C})$ is endowed with this property.

To complete the proof, we introduce the matrix valued-function $M(t) := [N_0^\top\ N_1^\top\ N_2^\top]^\top$, with $N_0=\bar{C}$, $N_1(t)=\bar{C}A^\star(t)$, $N_2(t)=N_1A^\star(t)+\dot{N}_1$. One verifies using the expressions of $A^\star(t)$ and $\bar{C}$ that

$$M(t)=\begin{bmatrix} I_3 & 0_{3,1} & 0_{3,1} & 0_{3,2} & 0_{3,1} & 0_{3,3} & 0_{3,3} \\ 0_{2,3} & 0_{2,1} & 0_{2,1} & I_2 & 0_{2,1} & 0_{2,3} & 0_{2,3} \\ 0_{3,3} & 0_{3,1} & 0_{3,1} & 0_{3,2} & 0_{3,1} & 0_{3,3} & I_3 \\ 0_{3,3} & 0_{3,1} & 0_{3,1} & 0_{3,2} & 0_{3,1} & I_3 & 0_{3,3} \\ \alpha_1 & \alpha_2 & \alpha_3 & 0_{3,2} & \dot{v}(t) & 0_{3,3} & 0_{3,3} \end{bmatrix}$$

For the sake of conciseness, we define $f(t,s,x) := M(t+s)\Phi(t+s,t)x$. This yields

$$f(t,s,x)=\begin{bmatrix} x_1 \\ x_4 \\ x_5 \\ \frac{s^2}{2}(\alpha_1x_1+\alpha_2x_2+\alpha_3x_3)+b(t+s,t)x_6+sx_7+x_8 \\ s(\alpha_1x_1+\alpha_2x_2+\alpha_3x_3)+a(t+s,t)x_6+x_7 \\ \alpha_1x_1+\alpha_2x_2+\alpha_3x_3+\dot{v}(t+s)x_6 \end{bmatrix}$$

with x_i $(i=1,\dots,8)$ denoting the i-th vector-part component of the vector $x \in \mathbb{R}^{14}$ of adequate dimensions, such that $x_1, x_7, x_8 \in \mathbb{R}^3$, $x_j \in \mathbb{R}$, $j=2,\dots,6$ and $f_k(t,s,x)$ $(k=1,\dots,6)$ denoting the k-th vector-part component of $f(t,s,x)$.

We now proceed by contradiction and assume that the pair $(A^\star, \bar{C})$ is not uniformly observable. Then, according to Proposition 1, there exists a sequence $\{t_p\}_{p\in\mathbb{N}}$ and $x \in \mathcal{S}^{13}$ (i.e. $|x|=1$) such that $\lim_{p\to+\infty}\int_0^{\delta+\bar{\delta}}|f(t_p,s,x)|^2ds=0$ with $\bar{\delta}>0$ as large as desired.

In view of the first three components of $f(t_p,s,x)$, the satisfaction of the condition in Proposition 1 requires that $|x_1|=|x_4|=|x_5|=0$. Let us rewrite the fourth component of $f(t_p,s,x)$ as follows

$$f_4(t_p,s,x)=\frac{s^2}{2}\rho g G^\top\begin{bmatrix}-x_3\\ x_2\\ 0\end{bmatrix}+sa(t_p+s,t_p)x_6+sx_7+x_8$$

Given that $\frac{s^2}{2}\rho g G^\top[-x_3\ x_2\ 0]^\top$ significantly dominates all other terms by a factor s, when s is large, the satisfaction of the condition in Proposition 1 implies that $|x_2|$ and $|x_3|$ must become small as fast as $1/\bar{\delta}$ when $\bar{\delta}$ becomes sufficiently large. Then, we may assume from now on that $|x_2|, |x_3| \ll 1$. Now, let us rewrite the last component of $f(t_p,s,x)$ as $f_6(t_p,s,x)=\dot{v}(t_p+s)x_6$. Using the persistent excitation assumption in (21), according to which $\int_{t_p}^{t_p+\delta+\bar{\delta}}|\dot{v}(s)x_6|^2ds \geq \delta\mu|x_6|^2$. This yields

$$\int_0^{\delta+\bar{\delta}}|f_6(t_p,s,x)|^2ds \geq \delta\mu|x_6|^2$$

This relation indicates that the convergence of $\int_{t_p}^{t_p+\delta}|f_6(t_p,s,x)|^2ds$ to zero implies that $|x_6|$ must become small when $\bar{\delta}$ is sufficiently large. Thus, we may also assume that $|x_6| \ll 1$. Similarly, considering the fifth components of $f(t_p,s,x)$ that simplifies to $f_5(t_p,s,x)=x_7$, the satisfaction of the condition in Proposition 1 implies that $|x_7| \ll 1$ when $\bar{\delta}$ is sufficiently large. Then, the fourth component of $f(t_p,s,x)$ reduces to $f_4(t_p,s,x)=x_8$, and likewise, the satisfaction of the condition in Proposition 1 implies that $|x_8| \ll 1$ when $\bar{\delta}$ is sufficiently large.

Given that all components of x must be zero or small as $\bar{\delta}$ becomes sufficiently large, the norm of x must be smaller than one to satisfy the condition of Proposition 1. However, the satisfaction of this condition also requires that $|x|=1$, which results in a contradiction that ultimately proves that the assumption of nonuniform observability of $(A^\star, \bar{C})$ is not valid. As a result, the pair $(A^\star, \bar{C})$, and by extension $(A^\star, C^\star)$, is uniformly observable. The remainder of the proof follows directly from the application of Theorem 3.1 and Corollary 3.2 in [6].

□