

# Nonlinear Observers Design for Vision-Aided Inertial Navigation Systems

Miaomiao Wang, Soulaïmane Berkane, and Abdelhamid Tayebi

**Abstract**—This paper deals with the simultaneous estimation of the attitude, position and linear velocity for vision-aided inertial navigation systems. We propose a nonlinear observer on  $SO(3) \times \mathbb{R}^{15}$  relying on body-frame acceleration, angular velocity and (stereo or monocular) bearing measurements of some landmarks that are constant and known in the inertial frame. Unlike the existing local Kalman-type observers, our proposed nonlinear observer guarantees almost global asymptotic stability and local exponential stability. A detailed uniform observability analysis has been conducted and sufficient conditions are derived. Moreover, a hybrid version of the proposed observer is provided to handle the intermittent nature of the measurements in practical applications. Simulation and experimental results are provided to illustrate the effectiveness of the proposed state observer.

**Index Terms**—Nonlinear observer, Vision-aided inertial navigation system (Vision-aided INS), Bearing measurements, Uniform observability.

## I. INTRODUCTION

The design of reliable state observers for the simultaneous estimation of the attitude (orientation), position, and linear velocity for inertial navigation systems (INSs) is crucial in many robotics and aerospace applications. Visual sensors, which provide a rich information about the environment, are becoming ubiquitous in a wide range of applications from mobile augmented reality to autonomous vehicles' navigation. A vision system together with an inertial measurement unit (IMU) forms a vision-aided INS (or visual-inertial navigation system). Depending on the camera type, vision-aided INSs can be categorized as monocular vision-aided INS (single camera) or stereo vision-aided INS (two cameras in a stereo setup). Vision-aided INS is widely used in robotics, for instance, visual-inertial odometry (VIO) and Simultaneous Localization and Mapping (SLAM) [2]–[5]. Unlike the VIO and SLAM, where the landmarks/features are unknown, the estimation problem at hand, referred to as vision-aided INS, assumes known landmarks in the inertial frame. Visual measurements

of these landmarks together with IMU measurements are used to simultaneously estimate the attitude, position and linear velocity. The proposed estimation scheme, endowed with almost global asymptotic stability<sup>1</sup> (AGAS) guarantees, is able to handle different types of measurements, including monocular-bearing and stereo-bearing measurements.

### A. Motivation and Prior Literature

The position and linear velocity of a rigid body can be obtained, for instance, from a Global Positioning System (GPS), while its attitude can be estimated from a set of body-frame measurements of some known inertial vectors [6]–[8]. Typically, low-cost IMU-based estimation techniques assume that the accelerometer provides body-frame measurements of the gravity vector, which is not true in applications involving non-negligible accelerations. One solution to this problem consists in using the so-called velocity-aided attitude observers [9]–[11], which make use of the linear velocity and IMU measurements to estimate the attitude. Instead of assuming the linear velocity available (as in the velocity-aided observers), GPS-aided navigation observers use IMU and GPS information to simultaneously estimate the attitude, position and linear velocity; see, for instance, [12] and references therein. These estimation schemes, however, are not suitable for implementations in GPS-denied environments (*e.g.*, indoor applications).

One alternative approach, that is widely adopted in GPS-denied environments, is the vision-aided INS. Most of the existing estimation schemes for vision-aided INSs in the literature are of Kalman-type such as the Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF) [2], [3]. Different versions of these algorithms have been proposed in the literature depending on the type of vision-based measurements used. Due to the nonlinearity of the vision-aided INS kinematics, these Kalman-type filters, relying on local linearizations, do not provide strong stability guarantees. Recently, several nonlinear observers using three-dimensional (3D) landmark position measurements, have been proposed in the literature [13]–[15]. In contrast with the standard EKF and its variants, an invariant EKF (IEKF), with provable local stability guarantees, has been proposed in [13]. A local Riccati-based nonlinear observer, inspired from [16], has been proposed in [14]. Motivated by the work in [17], [18], hybrid nonlinear observers, with global exponential stability guarantees, have been proposed in [15].

<sup>1</sup>The equilibrium point is stable and asymptotically attractive from almost all initial conditions except from a set of Lebesgue measure zero.

This work was supported by the National Sciences and Engineering Research Council of Canada (NSERC), under the grants NSERC-DG RGPIN 2020-06270 and NSERC-DG RGPIN-2020-04759. A preliminary and partial version of this work was presented in [1].

M. Wang is with the Department of Electrical and Computer Engineering, Western University, London, ON N6A 3K7, Canada (e-mail: mwang448@uwo.ca).

S. Berkane is with the Department of Department of Computer Science and Engineering, University of Quebec in Outaouais, QC J8X 3X7, Canada (e-mail: soulaïmane.berkane@uqo.ca).

A. Tayebi is with the Department of Electrical and Computer Engineering, Western University, London, ON N6A 3K7, Canada, and also with the Department of Electrical Engineering, Lakehead University, Thunder Bay, ON P7B 5E1, Canada (e-mail: atayebi@lakeheadu.ca).

It is clear that vision systems do not directly provide 3D landmark position measurements. In fact, they can be obtained from the images of a stereo-vision system using additional algorithms [19]. In [1], we developed nonlinear observers for attitude, position, linear velocity and gravity vector estimation, using direct stereo-bearing measurements. A local Riccati observer for attitude, position, linear velocity and accelerometer-bias estimation, with monocular-bearing measurements, has been proposed in [20]. On the other hand, in practical applications, the used sensors may have different sampling rates. For instance, the sampling rates of the vision sensors are much lower than those of the IMU, which is due to the hardware of the vision sensors and the heavy image processing computations. In this situation, one should be careful as the stability results derived for continuous-time observers are not preserved when the measurements are intermittent. To address this problem, hybrid nonlinear observers have been considered in [21]–[23].

### B. Contributions and Organization of the Paper

In the present paper, we propose an AGAS nonlinear observer for the simultaneous estimation of the attitude, position and linear velocity using body-frame accelerometer and gyro measurements as well as monocular or stereo bearing measurements. Note that, due to the motion space topology, AGAS is the strongest result one can achieve with smooth time-invariant observers. We also provide a hybrid version of the proposed observer that takes into account the sampled and intermittent nature of the visual measurements for practical implementation purposes. Some highlights of the contributions of this paper are as follows:

- 1) To the best of our knowledge, this work is the first to achieve AGAS results for vision-aided INS with bearing measurements. Note that the Riccati observers in [16], [20] provide only local stability guarantees. The key idea that allowed to achieve this strong stability result is the introduction of some auxiliary state variables allowing to appropriately design some output-driven signals leading to linear time-varying dynamics for the translational estimation errors.
- 2) The proposed observer uses generic vision-based measurements, including stereo-bearing and monocular-bearing measurements. This is a distinct feature from the existing nonlinear observers which are tailored to a specific type of vision-based measurements [1], [13]–[15], [20], [23]. This is achieved by introducing some auxiliary basis vectors in the estimation procedure, allowing to derive a generic (possibly time-varying) innovation term for the translational state estimation that captures the different types of vision-based measurements used in this work.
- 3) A detailed uniform observability analysis has been carried out for the two types of visual landmark information (i.e., stereo-bearing measurements and monocular bearing measurements). Sufficient conditions on the number and location of the landmarks as well as the motion of the vehicle, have been derived. The most challenging analysis

was the one related to the monocular vision system which required a tedious proof.

- 4) In practice, the sampling rates of the visual measurements are much lower than those of the IMU measurements (which can be assumed continuous). In this context, we propose a hybrid version of our nonlinear observer to handle the discrete and intermittent nature of the visual measurements, which has been experimentally validated using the EuRoc dataset [24].

The rest of this paper is organized as follows. After some preliminaries in Section II, we formulate our estimation problem in Section III. Section IV is devoted to the design of a generic nonlinear observer for vision-aided INS using different types of vision-based measurements with stability and observability analysis. A hybrid version of the proposed observer, taking into account the discrete and intermittent nature of the visual measurements, is presented in Section V. Simulation and experimental results are presented in Sections VI and VII, respectively.

## II. PRELIMINARY MATERIAL

### A. Notations and Definitions

The sets of real, non-negative real, natural numbers and nonzero natural numbers are denoted by  $\mathbb{R}$ ,  $\mathbb{R}_{\geq 0}$ ,  $\mathbb{N}$  and  $\mathbb{N}_{>0}$ , respectively. We denote by  $\mathbb{R}^n$  the  $n$ -dimensional Euclidean space, and by  $\mathbb{S}^n$  the set of unit vectors in  $\mathbb{R}^{n+1}$ . The Euclidean norm of a vector  $x \in \mathbb{R}^n$  is denoted by  $\|x\|$ , and the Frobenius norm of a matrix  $X \in \mathbb{R}^{n \times m}$  is denoted by  $\|X\|_F = \sqrt{\text{tr}(X^\top X)}$ . The  $n$ -by- $n$  identity and zeros matrices are denoted by  $I_n$  and  $0_n$ , respectively. For a given matrix  $A \in \mathbb{R}^{n \times n}$ , we define  $\lambda(A)$  as the set of all eigenvalues of  $A$ , and  $\mathcal{E}(A)$  as the set of all unit-eigenvectors of  $A$ . The minimum and maximum eigenvalues of  $A$  are, respectively, denoted by  $\lambda_{\min}^A$  and  $\lambda_{\max}^A$ . By  $\text{blkdiag}(\cdot)$ , we denote the block diagonal matrix. Let  $e_i$  denote the  $i$ -th basis vector of  $\mathbb{R}^n$  which represents the  $i$ -th column of the identity matrix  $I_n$ .

The Special Orthogonal group of order three, denoted by  $SO(3)$ , is defined as  $SO(3) := \{R \in \mathbb{R}^3, RR^\top = R^\top R = I_3, \det(R) = 1\}$ . The Lie algebra of  $SO(3)$  is given by  $\mathfrak{so}(3) := \{\Omega \in \mathbb{R}^3 : \Omega = -\Omega^\top\}$ . Let  $\times$  be the vector cross-product on  $\mathbb{R}^3$  and define the map  $(\cdot)^\times : \mathbb{R}^3 \rightarrow \mathfrak{so}(3)$  such that  $x \times y = x^\times y$ , for any  $x, y \in \mathbb{R}^3$ . Let  $\text{vec} : \mathfrak{so}(3) \rightarrow \mathbb{R}^3$  be the inverse isomorphism of the map  $(\cdot)^\times$ , such that  $\text{vec}(\omega^\times) = \omega$  for all  $\omega \in \mathbb{R}^3$ . For a matrix  $A \in \mathbb{R}^{3 \times 3}$ , we denote by  $\mathbb{P}_a : \mathbb{R}^{3 \times 3} \rightarrow \mathfrak{so}(3)$  the anti-symmetric projection of  $A$  such that  $\mathbb{P}_a(A) := (A - A^\top)/2$ . Define the composition map  $\psi_a := \text{vec} \circ \mathbb{P}_a$  such that, for a matrix  $A = [a_{ij}] \in \mathbb{R}^{3 \times 3}$ , one has  $\psi_a(A) = \frac{1}{2}[a_{32} - a_{23}, a_{13} - a_{31}, a_{21} - a_{12}]^\top$ . For any  $R \in SO(3)$ , we define  $|R|_I \in [0, 1]$  as the normalized Euclidean distance on  $SO(3)$  with respect to the identity  $I_3$ , which is given by  $|R|_I^2 = \text{tr}(I_3 - R)/4$ . We introduce the following important orthogonal projection operator:  $\pi : \mathbb{S}^2 \rightarrow \mathbb{R}^{3 \times 3}$  that will be used throughout this paper:

$$\pi(x) = I_3 - xx^\top, \quad x \in \mathbb{S}^2. \quad (1)$$

Note that  $\pi(x)$  is an orthogonal projection matrix which geometrically projects any vector in  $\mathbb{R}^3$  onto the plane or-

thogonal to  $x$ . Moreover, one verifies that  $\pi(x)$  is bounded and positive semi-definite,  $\pi(x)y = 0_{3 \times 1}$  if  $x, y$  are collinear, and  $R\pi(x)R^\top = \pi(Rx)$  for any  $R \in SO(3)$ . For the sake of simplicity, the argument of the time-dependent signals is omitted unless otherwise required for the sake of clarity.

Consider  $A(t) \in \mathbb{R}^{n \times n}$  and  $C(t) \in \mathbb{R}^{m \times n}$  as matrix-valued functions of time  $t$ , and suppose that  $A(t)$  and  $C(t)$  are continuous and bounded on  $\mathbb{R}_{\geq 0}$ . The following definition formulates the well-known uniform observability condition in terms of the Gramian matrix

**Definition 1.** The pair  $(A(t), C(t))$  is uniformly observable if there exist constants  $\delta, \mu$  such that

$$W_o(t, t + \delta) := \frac{1}{\delta} \int_t^{t+\delta} \Phi^\top(\tau, t) C^\top(\tau) C(\tau) \Phi(\tau, t) d\tau \geq \mu I_n, \quad \forall t \geq 0, \quad (2)$$

where  $\Phi(\tau, t)$  is the transition matrix associated to  $A(t)$  such that  $\frac{d}{dt}\Phi(t, \tau) = A(t)\Phi(t, \tau)$  and  $\Phi(t, t) = I_n$ .

### III. PROBLEM FORMULATION

#### A. Kinematic Model

Let  $\{\mathcal{I}\}$  be an inertial frame and  $\{\mathcal{B}\}$  be a body-fixed frame attached to the center of mass of a rigid body. Let the rotation matrix  $R \in SO(3)$  be the attitude of the frame  $\{\mathcal{B}\}$  with respect to the frame  $\{\mathcal{I}\}$ . Let the vectors  $p \in \mathbb{R}^3$  and  $v \in \mathbb{R}^3$  denote the position and linear velocity of the rigid body expressed in frame  $\{\mathcal{I}\}$ , respectively. The kinematic equations of a rigid body navigating in 3D space is given by:

$$\dot{R} = R\omega^\times \quad (3a)$$

$$\dot{p} = v \quad (3b)$$

$$\dot{v} = g + Ra \quad (3c)$$

where  $g \in \mathbb{R}^3$  denotes the gravity vector in frame  $\{\mathcal{I}\}$ ,  $\omega \in \mathbb{R}^3$  denotes the body-frame angular velocity expressed in frame  $\{\mathcal{B}\}$ , and  $a \in \mathbb{R}^3$  denotes the body-frame ‘‘apparent acceleration’’ capturing all non-gravitational forces applied to the rigid body expressed in frame  $\{\mathcal{B}\}$ .

**Assumption 1.** The measurements of the angular velocity  $\omega$  and acceleration  $a$  are continuous and bounded.

#### B. Estimation Problem for Vision-Aided INS

This work focuses on the problem of attitude, position and linear velocity estimation for INS using the body-frame acceleration and angular velocity measurements, as well as vision-based body-frame position information of a family of  $N \in \mathbb{N}_{>0}$  landmarks that are constant and known in the inertial frame. Let  $p_i$  denote the (constant and known) position of the  $i$ -th landmark in frame  $\{\mathcal{I}\}$ , and  $p_i^{\mathcal{B}} := R^\top(p_i - p)$  denote the position of the  $i$ -th landmark in frame  $\{\mathcal{B}\}$ . In the following, we detail the two measurement models considered in this work (see Fig. 1).

- 1) *Stereo-bearing measurements:* Let  $(R_{c1}, p_{c1})$  and  $(R_{c2}, p_{c2})$  denote the homogeneous transformation from the body-fixed frame  $\{\mathcal{B}\}$  to the right camera frame  $\{\mathcal{C}_1\}$  and left camera frame  $\{\mathcal{C}_2\}$ , respectively. Then, the

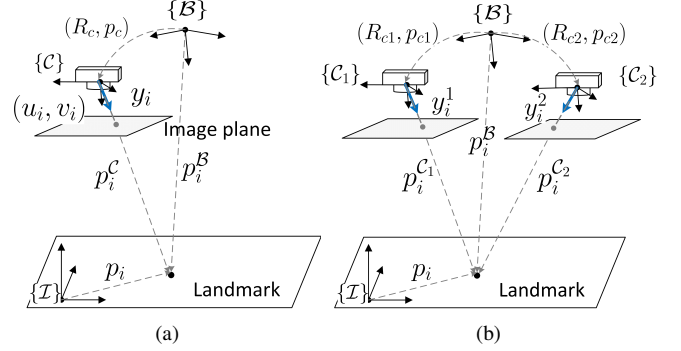


Fig. 1: The geometry models of vision systems: (a) monocular vision system (b) stereo vision system.

model of the stereo-bearing vectors of the  $i$ -th landmark in frame  $\{\mathcal{C}_s\}$ ,  $s \in \{1, 2\}$  are given as

$$y_i^s := \frac{p_i^{C_s}}{\|p_i^{C_s}\|} = \frac{R_{cs}^\top(p_i^{\mathcal{B}} - p_{cs})}{\|p_i^{\mathcal{B}} - p_{cs}\|}, \quad i = 1, 2, \dots, N, \quad (4)$$

where  $p_i^{C_s} = R_{cs}^\top(p_i^{\mathcal{B}} - p_{cs})$  denotes the coordinates of the  $i$ -th landmark expressed in frame  $\{\mathcal{C}_s\}$  for  $s = 1, 2$ .

- 2) *Monocular-bearing measurements:* Let  $(R_c, p_c)$  denote the homogeneous transformation from the body-fixed frame  $\{\mathcal{B}\}$  to the camera frame  $\{\mathcal{C}\}$ . Then, the model of the monocular-bearing vector of the  $i$ -th landmark expressed in frame  $\{\mathcal{C}\}$  is given as

$$y_i := \frac{p_i^{\mathcal{C}}}{\|p_i^{\mathcal{C}}\|} = \frac{R_c^\top(p_i^{\mathcal{B}} - p_c)}{\|p_i^{\mathcal{B}} - p_c\|}, \quad i = 1, 2, \dots, N, \quad (5)$$

where  $p_i^{\mathcal{C}} = R_c^\top(p_i^{\mathcal{B}} - p_c)$  denotes the coordinates of the  $i$ -th landmark expressed in frame  $\{\mathcal{C}\}$ .

**Remark 1.** The bearing measurements of the  $i$ -th landmark can be easily obtained from the direct pixel measurements  $(u_i, v_i)$  (in the image plane) as  $y_i = \mathcal{K}^{-1}z_i / \|\mathcal{K}^{-1}z_i\| \in \mathbb{S}^2$  with  $z_i = [u_i, v_i, 1]^\top$  and  $\mathcal{K}$  denoting the intrinsic matrix of the camera. As we can see in (4) and (5), only partial information of the 3D landmark positions in frame  $\{\mathcal{B}\}$  are available in the monocular-bearing and stereo-bearing measurements.

#### C. Objectives

Our main objectives in this work are as follows:

- 1) Design an almost globally asymptotically stable non-linear observer for the simultaneous estimation of the attitude  $R(t)$ , position  $p(t)$  and linear velocity  $v(t)$ , using the IMU measurements  $(\omega(t), a(t))$  and the visual measurements from either (4) or (5). The observer should be generic in the sense that it does not require any modification when using any of the above mentioned measurements.
- 2) Carry out a detailed uniform observability analysis for the above mentioned visual measurements scenarios and provide sufficient feasibility conditions depending on the

number and location of the landmarks, as well as the motion of the vehicle.

- 3) Provide a version of the observer with continuous IMU measurements, and sampled and intermittent visual measurements for practical implementation purposes. This is motivated by the fact that vision systems in general provide measurements at much lower sampling rates compared to the IMU sampling rates.

#### IV. A NONLINEAR OBSERVER USING CONTINUOUS VISION-BASED MEASUREMENTS

##### A. Nonlinear Observer Design

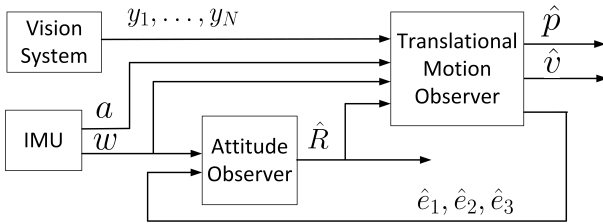


Fig. 2: Structure of the proposed nonlinear observer on  $SO(3) \times \mathbb{R}^{15}$  for vision-aided INSs.

To solve the vision-aided state estimation problem described in Section III-B, we propose the following nonlinear observer on  $SO(3) \times \mathbb{R}^{15}$ :

$$\dot{\hat{R}} = \hat{R}(\omega + \hat{R}^\top \sigma_R)^\times, \quad (6a)$$

$$\dot{\hat{p}} = \hat{v} + \sigma_R^\times \hat{p} + \hat{R} K_p \sigma_y, \quad (6b)$$

$$\dot{\hat{v}} = \hat{g} + \hat{R} a + \sigma_R^\times \hat{v} + \hat{R} K_v \sigma_y, \quad (6c)$$

$$\dot{\hat{e}}_i = \sigma_R^\times \hat{e}_i + \hat{R} K_i \sigma_y, \quad i = 1, 2, 3, \quad (6d)$$

where  $\hat{g} = \sum_{i=1}^3 g_i \hat{e}_i$  with  $g = [g_1, g_2, g_3]^\top$ . The matrix  $\hat{R} \in SO(3)$  denotes the estimate of the attitude  $R$ , and the vectors  $\hat{p} \in \mathbb{R}^3$  and  $\hat{v} \in \mathbb{R}^3$  denote the estimates of the position  $p$  and linear velocity  $v$ , respectively. The structure of the overall observer is shown in Fig. 2. The attitude innovation term  $\sigma_R$  is given as follows:

$$\sigma_R := \frac{k_R}{2} \hat{R} \sum_{i=1}^3 \rho_i (\hat{R}^\top \hat{e}_i)^\times (\hat{R}^\top e_i) = \frac{k_R}{2} \sum_{i=1}^3 \rho_i \hat{e}_i^\times e_i \quad (7)$$

with constant scalars  $k_R > 0, \rho_i \geq 0, i = 1, 2, 3$ . The observer gain matrices are designed as follows:

$$K = PC^\top(t)Q(t), \quad (8)$$

with  $K := [K_p^\top, K_1^\top, K_2^\top, K_3^\top, K_v^\top]^\top$ . The matrix  $C(t)$  will be defined later depending on type of visual measurements used. The matrix  $P$  is the solution to the following CRE:

$$\dot{P} = A(t)P + PA^\top(t) - PC^\top(t)Q(t)C(t)P + V(t) \quad (9)$$

where  $P(0) \in \mathbb{R}^{15 \times 15}$  is a symmetric positive definite matrix, matrices  $V(t) \in \mathbb{R}^{15 \times 15}$  and  $Q(t) \in \mathbb{R}^{15 \times 3N}$  are continuous,

bounded and uniformly positive definite, and matrix  $A(t) \in \mathbb{R}^{15 \times 15}$  is given as

$$A(t) = \begin{bmatrix} -\omega^\times & 0_3 & 0_3 & 0_3 & I_3 \\ 0_3 & -\omega^\times & 0_3 & 0_3 & 0_3 \\ 0_3 & 0_3 & -\omega^\times & 0_3 & 0_3 \\ 0_3 & 0_3 & 0_3 & -\omega^\times & 0_3 \\ 0_3 & g_1 I_3 & g_2 I_3 & g_3 I_3 & -\omega^\times \end{bmatrix}. \quad (10)$$

The matrix  $A(t)$  is obtained from the closed-loop translational error dynamics as it will be shown in the following subsection. In the traditional Kalman filter, matrices  $V(t)$  and  $Q^{-1}(t)$  are the covariance matrices of the additive noise on the system state and output, respectively. The explicit design of  $\sigma_y = [\sigma_{y1}^\top, \sigma_{y2}^\top, \dots, \sigma_{yN}^\top]^\top \in \mathbb{R}^{3N}$  and the matrix  $C(t)$  in the CRE (9) for each type of visual measurement described in Section III-B are given as follows:

- 1) *Stereo-bearing measurements*: From the stereo-bearing measurements defined in (4), define the vector  $\sigma_{yi}$  as

$$\sigma_{yi} = \sum_{s=1}^2 \pi(R_{cs} y_i^s) (\hat{R}^\top (\hat{p}_i - \hat{p}) - p_{cs}) \quad (11)$$

for all  $i = 1, 2, \dots, N$ , with  $p_i := [p_{i1}, p_{i2}, p_{i3}]^\top = \sum_{j=1}^3 p_{ij} e_j$ ,  $\hat{p}_i = \sum_{j=1}^3 \hat{p}_{ij} \hat{e}_j$  for all  $i = 1, 2, \dots, N$  and the projection map  $\pi$  defined in (1). The matrix  $C(t)$  is given by:

$$C(t) = \begin{bmatrix} \Pi_1 & -p_{11}\Pi_1 & -p_{12}\Pi_1 & -p_{13}\Pi_1 & 0_3 \\ \Pi_2 & -p_{21}\Pi_2 & -p_{22}\Pi_2 & -p_{23}\Pi_2 & 0_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Pi_N & -p_{N1}\Pi_N & -p_{N2}\Pi_N & -p_{N3}\Pi_N & 0_3 \end{bmatrix} \quad (12)$$

with  $\Pi_i := \sum_{s=1}^2 \pi(R_{cs} y_i^s), i = 1, \dots, N$ .

- 2) *Monocular-bearing measurements*: From the monocular-bearing measurements in (5), vector  $\sigma_{yi}$  is designed as

$$\sigma_{yi} = \pi(R_{ci} y_i) (\hat{R}^\top (\hat{p}_i - \hat{p}) - p_c) \quad (13)$$

for all  $i = 1, 2, \dots, N$ , with the map  $\pi$  defined in (1). The matrix  $C(t)$  is given by:

$$C(t) = \begin{bmatrix} \Pi_1 & -p_{11}\Pi_1 & -p_{12}\Pi_1 & -p_{13}\Pi_1 & 0_3 \\ \Pi_2 & -p_{21}\Pi_2 & -p_{22}\Pi_2 & -p_{23}\Pi_2 & 0_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Pi_N & -p_{N1}\Pi_N & -p_{N2}\Pi_N & -p_{N3}\Pi_N & 0_3 \end{bmatrix} \quad (14)$$

with  $\Pi_i := \pi(R_{ci} y_i), i = 1, \dots, N$ .

**Remark 2.** Note that the non-standard innovation term  $\sigma_R$  in (7) relies on the inertial frame axes  $e_i, i \in \{1, 2, 3\}$  and the auxiliary dynamical signals  $\hat{e}_i, i \in \{1, 2, 3\}$ . The motivation behind this construction is as follows. Typically, the attitude can be estimated using body-frame measurements of at least two non-collinear inertial frame vectors [6]. These body-frame vector measurements can be easily constructed from full landmark position measurements, see for instance [23]. However, in the case of body-frame bearing measurements, the problem is quite challenging since we do not have the corresponding inertial vectors that will allow the construction of an appropriate innovation term  $\sigma_R$ . To overcome



this challenge, we consider the inertial basis vectors  $e_i$  and their corresponding body-frame vectors  $R^\top e_i$ . Since  $R^\top e_i$  is unknown, we design the adaptive auxiliary vectors  $\hat{e}_i$  such that  $\hat{R}^\top \hat{e}_i$  tends exponentially to  $R^\top e_i$ . This idea is somewhat similar to the idea of the velocity-aided attitude estimation schemes where we introduce an auxiliary variable to overcome the lack of the acceleration in the inertial frame [9]–[11].

**Remark 3.** The projection operator  $\pi(R_{cy_i})$  projects vectors onto the plane orthogonal to the body-frame bearing  $R_{cy_i}$ . This operator allows to eliminate, from the projected vector, the component which is collinear to  $R_{cy_i}$ . For instance, in (13) the projection of  $\hat{R}^\top(\hat{p}_i - \hat{p}) - p_c$  onto the plane orthogonal to the body-frame unit vector  $R_{cy_i}$  will boil down to the projection of  $\hat{R}^\top(\hat{p}_i - \hat{p}) - p_i^B$ , since the component  $p_i^B - p_c$ , which is parallel to  $R_{cy_i}$ , will be eliminated. This mechanism, will allow us to generate the needed terms to put the error injection vector  $\sigma_y$  in the sought-after form  $\sigma_y = C(t)\tilde{x}$  as it will be shown later. The projection idea, however, is not new; it has been used in different ways in many references, for instance, [25]–[28].

**Remark 4.** Note that 3D landmark positions can be algebraically reconstructed from stereo-bearing measurements. The main motivation for the direct use of the stereo-bearing measurements in our observer is related to the robustness of the resulting estimation algorithm. In fact, as shown in the experimental results (Section VII), the observers relying on 3D landmark position measurements may fail in situations where one of the cameras of the stereo-vision system loses sight of the landmarks for some period of time. However, our observer using direct stereo-bearing measurements handles this situation very well by switching to a monocular bearing configuration.

## B. Error Dynamics and Stability Analysis

Define the attitude estimation error  $\tilde{R} := R\hat{R}^\top$ , and the translational estimation error  $\tilde{x} := [\tilde{p}^\top, \tilde{e}_1^\top, \tilde{e}_2^\top, \tilde{e}_3^\top, \tilde{v}^\top]^\top \in \mathbb{R}^{15}$  with  $\tilde{p} = R^\top p - \hat{R}^\top \hat{p}$ ,  $\tilde{v} = R^\top v - \hat{R}^\top \hat{v}$  and  $\tilde{e}_i = R^\top e_i - \hat{R}^\top \hat{e}_i$ ,  $i \in \{1, 2, 3\}$ . These geometric estimation errors are motivated from [1], [15], [23], and are different from the standard errors used in classical EKF-based filters [2], [3]. The innovation term  $\sigma_R$  in (7) can be rewritten as

$$\sigma_R = k_R \psi_a(M\tilde{R}) + \Gamma(t)\tilde{x} \quad (15)$$

with  $\Gamma(t) := \frac{1}{2}k_R[0_3, \rho_1 e_1^\times \hat{R}(t), \rho_2 e_2^\times \hat{R}(t), \rho_3 e_3^\times \hat{R}(t), 0_3] \in \mathbb{R}^{3 \times 15}$ , where we made use of the facts  $\hat{e}_i = \hat{R}^\top e_i - \hat{R}^\top \hat{e}_i$  and  $\frac{1}{2} \sum_{i=1}^3 \rho_i e_i^\times \hat{R}^\top e_i = -\psi_a(M\tilde{R})$  with  $M = \sum_{i=1}^3 \rho_i e_i e_i^\top = \text{diag}(\rho_1, \rho_2, \rho_3)$ . It is not difficult to show that  $\Gamma(t)$  is continuous and bounded, i.e.,  $\|\Gamma(t)\|_F \leq \frac{\sqrt{2}}{2}k_R \sum_{i=1}^3 \rho_i := c_\Gamma$ . Moreover, for any distinct non-negative scalars  $\rho_i$ ,  $i = 1, 2, 3$ , the matrix  $M$  is positive semi-definite with three distinct eigenvalues. From (15), one can notice that  $\sigma_R$  has two terms: the first term  $k_R \psi_a(M\tilde{R})$  is commonly used for the establishment of the stability proofs of the attitude estimation subsystem; see for instance [6], [17]. The second term depending on the estimation error  $\tilde{x}$  is an asymptotically vanishing

perturbation term as it will be shown later. In view of (3), (6) and (15), one obtains the following closed-loop system:

$$\dot{\tilde{R}} = \tilde{R}(-k_R \psi_a(M\tilde{R}) - \Gamma(t)\tilde{x})^\times \quad (16a)$$

$$\dot{\tilde{x}} = A(t)\tilde{x} - K\sigma_y \quad (16b)$$

where  $A(t)$  is defined in (10), and  $K$  is designed in (8) relying on the solution  $P(t)$  to the CRE (9). Note that the overall closed-loop system (16) is nonlinear and it can be seen as a cascade interconnection of a time-varying linear (LTV) system on  $\mathbb{R}^{15}$  and a nonlinear system evolving on  $SO(3)$ . Given continuous and bounded matrices  $A(t)$ ,  $C(t)$ ,  $Q(t)$ ,  $V(t)$ , with  $Q(t)$ ,  $V(t)$  being uniformly positive definite and the pair  $(A(t), C(t))$  uniformly observable, it follows that the solution  $P(t)$  to the CRE (9) is well defined on  $\mathbb{R}_{\geq 0}$  and there exist positive constants  $0 < p_m \leq p_M < \infty$  such that  $p_m I_n \leq P(t) \leq p_M I_n$  for all  $t \geq 0$  [29], [30].

**Theorem 1.** Consider the nonlinear system (16) with  $\sigma_y = C(t)\tilde{x}$  and  $C(t)$  being continuous and bounded. Let Assumption 1 hold, and suppose that the pair  $(A(t), C(t))$  is uniformly observable. Pick  $k_R > 0$  and three distinct scalars  $\rho_i \geq 0$ ,  $i = 1, 2, 3$ . Let  $K$  be given in (8) with matrices  $Q(t)$  and  $V(t)$  in (9) being continuous, bounded and uniformly positive definite. Then, the following statements hold:

- i) All solutions of the closed-loop system (16) converge to the set of equilibria given by  $(I_3, 0_{15 \times 1}) \cup \Psi_M$  where

$$\Psi_M := \{(\tilde{R}, \tilde{x}) \in SO(3) \times \mathbb{R}^{15} | \tilde{R} = \mathcal{R}_\alpha(\pi, v), v \in \mathcal{E}(M), \tilde{x} = 0_{15 \times 1}\}. \quad (17)$$

- ii) The desired equilibrium  $(I_3, 0_{15 \times 1})$  is locally exponentially stable.
- iii) All the undesired equilibria in  $\Psi_M$  are unstable, and the desired equilibrium  $(I_3, 0_{15 \times 1})$  is almost globally asymptotically stable.

*Proof.* See Appendix A.  $\square$

Theorem 1 provides AGAS and local exponential stability results for the proposed nonlinear observer. Among the interesting features of our observer is the fact that  $\tilde{x}$  is guaranteed to converge globally exponentially to zero independently from the dynamics of  $\tilde{R}$ , as long as  $\sigma_y = C(t)\tilde{x}$  and the pair  $(A(t), C(t))$  is uniformly observable. Note that  $A(t)$  in (10) is continuous and bounded as long as  $\omega(t)$  is continuous and bounded. It is worth pointing out that AGAS is the strongest result one can aim at with a smooth vector field on  $SO(3) \times \mathbb{R}^{15}$ . This is due to the topological obstruction on  $SO(3)$  which consists in the fact that no continuous time-invariant vector field on  $SO(3)$  leads to a globally asymptotically stable equilibrium [31].

## C. Observability Analysis

In this subsection, we derive sufficient conditions for the uniform observability of the pair  $(A(t), C(t))$  for the previously mentioned two types of vision-based measurements. An important technical result that will be used to carry out our uniform observability proofs is given in the following lemma:

**Lemma 1.** Consider a constant matrix  $A \in \mathbb{R}^{n \times n}$  and a (possibly) time-varying matrix  $C(t) \in \mathbb{R}^{m \times n}$  such that:

- 1) All eigenvalues of  $A$  are real.
- 2)  $C(t)$  is continuous and bounded.

Let  $N = A - S$  be a nilpotent matrix with index  $s \leq n$ , where  $S$  is a diagonalizable matrix. Let  $\mathcal{O}(t) \in \mathbb{R}^{r \times n}$  be a matrix composed of ( $r > 0$ ) row vectors of  $C(t)$ ,  $C(t)N, \dots, C(t)N^{s-1}$ . Suppose that there exist constant scalars  $\delta, \mu > 0$  such that

$$\int_t^{t+\delta} \mathcal{O}^\top(\tau) \mathcal{O}(\tau) d\tau > \mu I_n, \quad \forall t \geq 0. \quad (18)$$

Then, the pair  $(A, C(t))$  is uniformly observable.

*Proof.* See Appendix B.  $\square$

Note that the decomposition  $A = S + N$  with a nilpotent matrix  $N$  and a diagonalizable matrix  $S$  is known as the Jordan-Chevalley decomposition. It is important to mention that matrices  $N$  and  $S$  are uniquely determined and commute (i.e.,  $SN = NS$ ), see [32, Theorem 1]. The main advantage of Lemma 1 is to provide a relaxed condition for the uniform observability of the pair  $(A, C(t))$  when the nilpotent part of  $A$  is non-zero (i.e.,  $N \neq 0_n$ ). Condition (18) is equivalent to the Kalman observability if  $A$  is a nilpotent matrix and  $C$  is constant. Moreover, if  $A$  is a diagonalizable matrix (i.e.,  $N = 0_n$ ), condition (18) reduces to the persistency of excitation (PE) requirement on  $C(t)$ . Note that other uniform observability conditions were proposed in the literature such as [33, Lemma 3.1], which involves high-order derivatives of  $C(t)$ , and [27, Lemma 2.7], which is more suitable when  $C(t)$  is the product of a PE matrix and a constant matrix guaranteeing Kalman observability.

1) *Stereo-bearing measurements:* From (4), one can rewrite  $\sigma_{yi}$  in (11) in terms of the estimation errors as

$$\begin{aligned} \sigma_{yi} &= \sum_{s=1}^2 R_{cs} \pi(y_i^s) R_{cs}^\top (\hat{R}^\top (\hat{p}_i - \hat{p}) - R^\top (p_i - p)) \\ &= \Pi_i \tilde{p} - p_{i1} \Pi_i \tilde{e}_1 - p_{i2} \Pi_i \tilde{e}_2 - p_{i3} \Pi_i \tilde{e}_3 \end{aligned} \quad (19)$$

for all  $i = 1, \dots, N$ , where we made use of the facts  $\Pi_i = \sum_{s=1}^2 \pi(R_{cs} y_i^s)$ ,  $p_i = \sum_{j=1}^3 p_{ij} e_j$  and  $\pi(R_{cs} y_i^s) (R^\top (p_i - p) - p_{cs}) = 0_{3 \times 1}$ . From the definition of  $\tilde{x}$ , one obtains  $\sigma_y = C(t) \tilde{x}$  with  $C(t)$  defined in (12). For each  $i = 1, \dots, N$ , the matrix  $\Pi_i$  is positive definite if the vectors  $R_{c1} y_i^1$  and  $R_{c2} y_i^2$  are non-collinear. Note that for stereo vision systems,  $R_{c1} y_i^1$  and  $R_{c2} y_i^2$  are naturally non-collinear since all the visible landmarks are within the limited sensing distance in practice. Hence, we assume that matrices  $\Pi_i, i = 1, \dots, N$  are uniformly positive definite. Note that the matrix  $C(t)$  is time-varying due to the time-varying matrices  $\Pi_i(t), i = 1, 2, \dots, N$ . Defining the  $3N$ -by- $3N$  block diagonal matrix  $\Theta(t) := \text{blkdiag}(\Pi_1(t), \Pi_2(t), \dots, \Pi_N(t))$ , one has  $C(t) = \Theta(t) \bar{C}$  with

$$\bar{C} = \begin{bmatrix} I_3 & -p_{11} I_3 & -p_{12} I_3 & -p_{13} I_3 & 0_3 \\ I_3 & -p_{21} I_3 & -p_{22} I_3 & -p_{23} I_3 & 0_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ I_3 & -p_{N1} I_3 & -p_{N2} I_3 & -p_{N3} I_3 & 0_3 \end{bmatrix}. \quad (20)$$

It is easy to show that  $C(t)$  is continuous and bounded since  $\|\Theta(t)\|_F \leq \sum_{i=1}^N \|\Pi_i\|_F$  and  $\|\Pi_i\|_F, i = 1, \dots, N$  are bounded.

**Lemma 2.** Consider the matrices  $A(t)$  defined in (10) and  $C(t)$  defined in (12). Suppose that there exist three non-aligned landmarks among the  $N \geq 3$  measurable landmarks, whose plane is not parallel to the gravity vector. Then, the pair  $(A(t), C(t))$  is uniformly observable.

*Proof.* See Appendix C.  $\square$

**Remark 5.** The sufficient observability conditions in Lemma 2 using stereo bearing measurements are mildly stronger than those needed for the local observers in the literature [13], [14], [23] using 3D landmark position measurements. This is mainly due to the over-parameterization of our observer with the additional auxiliary signals  $\tilde{e}_i$ , which is the paid price for the almost global asymptotic stability results achieved in Theorem 1.

2) *Monocular-bearing measurements:* From (5), one can rewrite  $\sigma_{yi}$  in (13) in terms of the estimation errors as

$$\begin{aligned} \sigma_{yi} &= R_c \pi(y_i) R_c^\top (\hat{R}^\top (\hat{p}_i - \hat{p}) - p_c - (R^\top (p_i - p) - p_c)) \\ &= \Pi_i \tilde{p} - p_{i1} \Pi_i \tilde{e}_1 - p_{i2} \Pi_i \tilde{e}_2 - p_{i3} \Pi_i \tilde{e}_3 \end{aligned} \quad (21)$$

for all  $i = 1, \dots, N$  with  $\Pi_i = \pi(R_c y_i)$ . From the definition of  $\tilde{x}$ , one obtains  $\sigma_y = C(t) \tilde{x}$  with  $C(t)$  defined in (14). Note that the matrix  $C(t)$  defined in (14) is similar to the matrix  $C(t)$  defined in (12), and the main difference is that the matrix  $\Pi_i$  in (14) is only guaranteed to be positive semi-definite. One can also show that  $C(t) = \Theta(t) \bar{C}$  with  $\bar{C}$  defined in (20) and  $\Theta(t) = \text{blkdiag}(\Pi_1, \Pi_2, \dots, \Pi_N)$  being continuous and bounded.

**Lemma 3.** Consider the matrices  $A(t)$  defined in (10) and  $C(t)$  defined in (14). Suppose that there exist three non-aligned landmarks, indexed by  $\ell_1, \ell_2, \ell_3$ , among the  $N \geq 3$  measurable landmarks, whose plane is not parallel to the gravity vector, and one of following statements holds:

- i) The camera is in motion with bounded velocity and there exists a constant  $\epsilon > 0$  such that for any time  $t^* \geq 0$  and landmark  $i \in \{\ell_1, \ell_2, \ell_3\}$ , there exists some time  $t > t^*$  such that  $\|(R(t) R_c y_i(t)) \times (R(t^*) R_c y_i(t^*))\| \geq \epsilon$ .
- ii) The camera is motionless and the following matrix has full rank of  $15 + N$

$$\mathcal{C}' = \begin{bmatrix} \bar{C} & M \\ N_1 & 0_{3 \times N} \\ N_2 & 0_{3 \times N} \end{bmatrix} \in \mathbb{R}^{(3N+6) \times (15+N)} \quad (22)$$

where  $\bar{C}$  defined in (20),  $M = \text{blkdiag}(p_1 - p', \dots, p_N - p')$ ,  $N_1 = [0_3, 0_3, 0_3, 0_3, I_3, 0_3]$ ,  $N_2 = [0_3, g_1 I_3, g_2 I_3, g_3 I_3, 0_3]$ , and  $p' = p + R p_c$  denoting the position of the camera in the inertial frame.

Then, the pair  $(A(t), C(t))$  is uniformly observable.

*Proof.* See Appendix D.  $\square$

**Remark 6.** The proof of this lemma relies on the application of the technical Lemma 1. For the condition i), using the fact

$R(t)R_c y_i(t) = (p_i - p'(t))/\|p_i - p'(t)\|$ , it follows that the camera is not indefinitely moving in a straight line passing through one of landmarks  $\ell_1, \ell_2, \ell_3$ . Note that this condition involves an extra condition on the motion of the vehicle, with respect to Lemma 2, to generate sufficient information for uniform observability. Condition ii) also holds when the camera is in motion and the minimum number of required landmarks for the matrix  $\mathcal{O}'$  to have a full rank of  $N + 15$  is  $N = 5$ .

**Proposition 1.** Consider the case where the camera is motionless with  $N \geq 5$  measurable non-aligned landmarks. Suppose that none of the following scenarios hold:

- (a) All the landmarks are located in the same plane.
- (b) There exist three non-aligned landmarks and the rest of landmarks are located in the plane parallel to the gravity vector and contains two of these three landmarks.
- (c) There exist three non-aligned landmarks and the rest of the landmarks are aligned with one of these three landmarks and the position of the camera.
- (d) There exist three non-aligned landmarks and the rest of the landmarks are either located in the plane parallel to the gravity vector and contains two of these three landmarks, or aligned with the third of these three landmarks and the position of the camera.

Then, the matrix  $\mathcal{O}'$  defined in (22) has full rank.

*Proof.* See Appendix E  $\square$

All the four cases stated in Proposition 1 are summarized in Fig 3. Note that the cases (a) and (b) are independent from the location of the camera. According to the Lemma 3, if the matrix  $\mathcal{O}'$  has full rank, one can conclude that the pair  $(A(t), C')$  is uniformly observable. Note that the state estimation in this static case is also known as the static Perspective-n-Point (PnP) problem [16], [34]. In this scenario, one aims at determining the pose (position and orientation) of a camera given its intrinsic parameters and a set of  $N$  correspondences between 3D points and their 2D projections.

## V. HYBRID OBSERVER USING INTERMITTENT VISION-BASED MEASUREMENTS

In practical applications, the IMU measurements can be obtained at a high rate, while the vision-based measurements are often obtained at a much lower rate due to the hardware design of the vision sensors and the heavy image processing computations. Hence, the IMU measurements can be assumed as continuous and the measurements from the vision systems are sampled intermittently. This motivates us to redesign the proposed continuous nonlinear observer in terms of continuous IMU and intermittent vision-based measurements.

**Assumption 2.** We assume that the vision-based measurements are available at strictly increasing time instants  $\{t_k\}_{k \in \mathbb{N}_{>0}}$ , and there exist constants  $0 < T_m \leq T_M < \infty$  such that  $t_1 < T_M$  and  $T_m \leq t_{k+1} - t_k \leq T_M$  for all  $k \in \mathbb{N}_{>0}$ .

This assumption implies that the time difference between two consecutive vision-based measurements are lower and

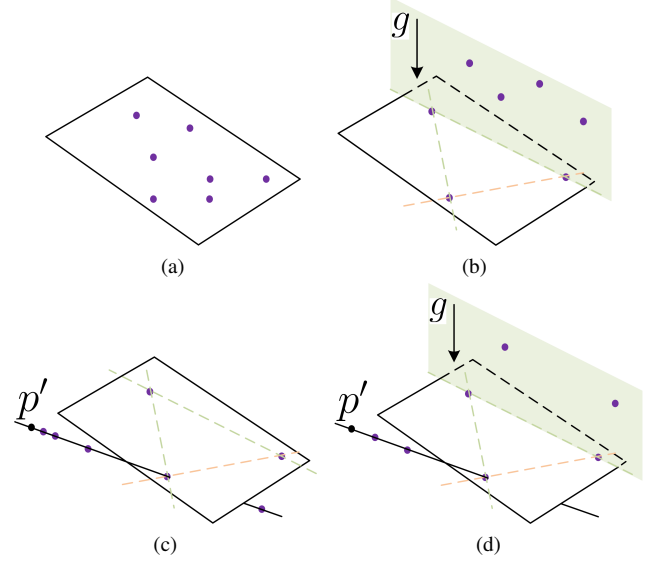


Fig. 3: The four possible cases of Proposition 1 where the matrix  $\mathcal{O}'$  in (22) does not have full rank. The locations of  $N \geq 5$  non-aligned landmarks are depicted by purple dots, and the location of the motionless monocular camera is depicted by a black dot.

upper bounded. The positive lower bound  $T_m$  is required to avoid Zeno behaviors. Note that, if  $T_m = T_M$ , the vision-based measurements are sampled periodically.

Motivated by the work in [23] and making use of the framework of hybrid dynamical systems presented in [35], [36], we propose the following hybrid nonlinear observer modified from (6) as

$$\begin{aligned} \dot{\hat{R}} &= \hat{R}(\omega + \hat{R}^\top \sigma_R)^\times & \hat{R}^+ &= \hat{R} \\ \dot{\hat{p}} &= \sigma_R^\times \hat{p} + \hat{v} & \hat{p}^+ &= \hat{p} + \hat{R} K_p \sigma_y \\ \dot{\hat{v}} &= \sigma_R^\times \hat{v} + \hat{g} + \hat{R} a & \hat{v}^+ &= \hat{v} + \hat{R} K_v \sigma_y \\ \dot{\hat{e}}_i &= \sigma_R^\times \hat{e}_i, \quad i \in 1, 2, 3 & \hat{e}_i^+ &= \hat{e}_i + \hat{R} K_i \sigma_y \end{aligned} \quad (23)$$

$t \in [t_{k-1}, t_k], k \in \mathbb{N}_{>0}$        $t \in \{t_k\}, k \in \mathbb{N}_{>0}$

where  $k_R > 0$ ,  $\hat{g} = \sum_{i=1}^3 g_i \hat{e}_i$ ,  $\sigma_R$  is given in (7), and the vector  $\sigma_y$  is given in (11) for stereo-bearing measurements and in (13) for monocular-bearing measurements. Let  $K := [K_p^\top, K_1^\top, K_2^\top, K_3^\top, K_v^\top]^\top$ , which is designed as

$$K = PC^\top(t)(C(t)PC^\top(t) + Q^{-1}(t))^{-1} \quad (24)$$

where  $P$  is the solution to the following continuous-discrete Riccati equations (CDRE)

$$\dot{P} = A(t)P + PA(t)^\top + V(t), \quad t \in [t_k, t_{k+1}] \quad (25a)$$

$$P^+ = (I - KC(t))P, \quad t \in \{t_k\} \quad (25b)$$

where  $A(t)$  is given by (10),  $P(0)$  is symmetric positive definite, and  $Q(t) \in \mathbb{R}^{3N \times 3N}$ ,  $V(t) \in \mathbb{R}^{15 \times 15}$  are continuous, bounded and uniformly positive definite. Then, as per [13], [37] and [23, Lemma 7], the solution  $P$  to (25) exists, and there exist constants  $0 < p_m \leq p_M < \infty$  such that  $p_m I_{15} \leq P \leq p_M I_{15}$  for all  $t \geq 0$  if there exist positive

constants  $\mu \in \mathbb{R}$  and  $\delta \in \mathbb{N}_{>0}$  such that  $W_o^h(t_j, t_{j+\delta}) = \sum_{i=j}^{j+\delta} \Phi^\top(t_i, t_j) C^\top(t_i) C(t_i) \Phi(t_i, t_j) > \mu I_{15}, \forall j \in \mathbb{N}_{>0}$ , where  $\Phi(t, \tau)$  denotes the state transition matrix associated to  $A(t)$ . Note that  $W_o^h$  is the discrete version of  $W_o$  in (2). Hence, the same conditions as in Lemma 2 and 3 can be derived for the existence of the positive constants  $\mu \in \mathbb{R}$  and  $\delta \in \mathbb{N}_{>0}$  such that  $W_o^h(t_j, t_{j+\delta}) > \mu I_{15}$  for all  $j \in \mathbb{N}_{>0}$ .

In view of (3), (7) and (23), one obtains the following hybrid closed-loop system  $\mathcal{H}$ :

$$\left. \begin{aligned} \dot{\tilde{R}} &= \tilde{R}(-k_R \psi_a(M\tilde{R}) - \Gamma(t)\tilde{x})^\times \\ \dot{\tilde{x}} &= A(t)\tilde{x} \end{aligned} \right\} \quad t \in [t_k, t_{k+1}] \quad (26a)$$

$$\left. \begin{aligned} \tilde{R}^+ &= \tilde{R} \\ \tilde{x}^+ &= (I_{15} - KC(t))\tilde{x} \end{aligned} \right\} \quad t \in \{t_k\} \quad (26b)$$

where the AGAS proof for the equilibrium  $(I_3, 0_{15 \times 1})$  of  $\mathcal{H}$  can be easily conducted by combing the proof of Theorem 1 and [23, Theorem 9], and is therefore omitted here. The proposed hybrid nonlinear observer for vision-aided INSs with intermittent vision-based measurements has been summarized in Algorithm 1.

---

**Algorithm 1** Nonlinear observer for vision-aided INSs

---

**Input:** Continuous IMU measurements  $\omega$  and  $a$ , and intermittent visual measurements at a set of time instants  $\{t_k\}_{k \in \mathbb{N}_{>0}}$ .

**Output:**  $\hat{R}(t), \hat{p}(t)$  and  $\hat{v}(t)$  for all  $t \geq 0$

```

1: for  $k \geq 1$  do
2:   while  $t \in [t_{k-1}, t_k]$  do
3:      $\dot{\hat{R}} = \hat{R}(\omega + \hat{R}^\top \sigma_R)^\times$  /*  $\sigma_R$  defined in (7) */
4:      $\dot{\hat{p}} = \sigma_R^\times \hat{p} + \hat{v}$ 
5:      $\dot{\hat{v}} = \sigma_R^\times \hat{v} + \sum_{i=1}^3 g_i \hat{e}_i + \hat{R}a$ 
6:      $\dot{\hat{e}}_i = \sigma_R^\times \hat{e}_i$  /* for all  $i = 1, 2, 3$  */
7:      $\dot{P} = A(t)P + PA^\top(t) + V(t)$  /*  $A$  defined in (10)
        and  $V(t)$  being uniformly positive definite */
8:   end while
9:   Obtain the vector  $\sigma_y$  and matrix  $C(t)$  from the visual
        measurements at time  $t_k$  /*Using (11) and
        (12) for stereo-bearing measurements, or (13) and (14)
        for monocular-bearing measurements */
10:   $K = PC^\top(t_k)(C(t_k)PC^\top(t_k) + Q^{-1}(t_k))^{-1}$  /*  $Q(t)$ 
        being uniformly positive definite */
11:  Compute matrices  $K_p, K_1, K_2, K_3$  and  $K_v$  from  $K$ 
        /*using  $K = [K_p^\top, K_1^\top, K_2^\top, K_3^\top, K_v^\top]^\top$  */
12:   $\hat{R}^+ = \hat{R}$ 
13:   $\hat{p}^+ = \hat{p} + \hat{R}K_p\sigma_y$ 
14:   $\hat{v}^+ = \hat{v} + \hat{R}K_v\sigma_y$ 
15:   $\hat{e}_i^+ = \hat{e}_i + \hat{R}K_i\sigma_y$  /* for all  $i = 1, 2, 3$  */
16:   $P^+ = (I_{15} - KC)P$ 
17: end for

```

---

**Remark 7.** Note that our proposed observer (23) is deterministic and the gain matrix  $K$  is designed based on the CDRE (25) with any uniformly positive definite matrices  $V(t)$  and  $Q(t)$ . However, it is possible to relate (locally) the matrices  $V(t)$  and  $Q(t)$  to the measurements noise properties leading to a local sub-optimal design of the observer gains in the

spirit of the Kalman filter. As in [23], let  $\omega$  and  $a$  be the noisy measurements and replace  $\omega$  by  $\omega + n_\omega$  and  $a$  by  $a + n_a$  in (3) with  $n_\omega$  and  $n_a$  denoting the noise signals associated to  $\omega$  and  $a$ , respectively. In view of (3), (23) and (26b), the time derivative of  $\tilde{x}$  can be approximated around  $\tilde{x} = 0$  by  $\dot{\tilde{x}} \approx A(t)\tilde{x} + G_t n_x$  with  $n_x = [n_\omega^\top, n_a^\top]^\top$ ,  $A(t)$  defined in (10), and

$$G_t = - \begin{bmatrix} (\hat{R}^\top \hat{p})^\times & (\hat{R}^\top \hat{e}_1)^\times & (\hat{R}^\top \hat{e}_2)^\times & (\hat{R}^\top \hat{e}_3)^\times & (\hat{R}^\top \hat{v})^\times \\ 0_3 & 0_3 & 0_3 & 0_3 & -I_3 \end{bmatrix}^\top$$

Moreover, replacing  $y_i$  in (5) by  $y_i + n_{y_i}$  for each bearing measurement with  $n_{y_i}$  denoting the noise signals. From (13),  $\sigma_{y_i}$  can be rewritten in the form of  $\sigma_{y_i} = \Pi_i \tilde{p} - \sum_{j=1}^3 \Pi_i \tilde{e}_j + \|p - p_i\| \Pi_i n_{y_i}$ , and  $\sigma_y$  can be approximated around  $\tilde{x} = 0$  by  $\sigma_y \approx C(t)\tilde{x} + M_t n_y$  with  $n_y = [n_{y_1}^\top, \dots, n_{y_N}^\top]^\top$  and  $M_t = \text{blkdiag}(\|p - \hat{p}_1\| \Pi_1, \dots, \|p - \hat{p}_N\| \Pi_N)$ . Then, matrices  $V(t)$  and  $Q(t)$  can be related to the covariance matrices of the measurements noise as follows:

$$V(t) = G_t \text{Cov}(n_x) G_t^\top, \quad Q^{-1}(t) = M_t \text{Cov}(n_y) M_t^\top. \quad (27)$$

In practice, a small positive definite matrix can be added to  $V(t)$  and  $Q^{-1}(t)$  to guarantee that  $V(t)$  and  $Q^{-1}(t)$  are uniformly positive definite.

## VI. SIMULATION RESULTS

In this simulation, we consider an autonomous vehicle moving on the ‘8’-shape trajectory described by  $p(t) = 2[\sin(t), \sin(t) \cos(t), 1]^\top$  (m) with the initial attitude  $R(0) = I_3$  and angular velocity  $\omega(t) = [-\cos(2t), 1, \sin(2t)]^\top$  (rad/s). Let  $g = [0, 0, -9.81]^\top$  (m/s<sup>2</sup>) be the gravity vector expressed in the inertial frame, and assume that there are  $N = 5$  landmarks randomly selected. Different types of continuous vision-based measurements are generated using (4) and (5). For comparison purposes, 3D landmark position measurements  $y_i = R^\top(p_i - p)$ ,  $i = 1, 2, \dots, N$  are considered. In this case, the vector  $\sigma_y = [\sigma_{y_1}^\top, \sigma_{y_2}^\top, \dots, \sigma_{y_N}^\top]^\top \in \mathbb{R}^{3N}$  used in (6) is designed as

$$\sigma_{y_i} = \hat{R}^\top(\hat{p}_i - \hat{p}) - y_i, \quad i = 1, 2, \dots, N \quad (28)$$

with  $\hat{p}_i = \sum_{j=1}^3 p_{ij} \hat{e}_j$ . It follows that  $\sigma_y = C\tilde{x}$  with a constant matrix  $C = \bar{C}$  and  $\bar{C}$  defined in (20)<sup>2</sup>.

The same initial conditions are considered for each case as:  $\hat{R}(0) = \exp(0.5\pi u^\times)$  with  $u \in \mathbb{S}^2$ ,  $\hat{v}(0) = \hat{p}(0) = 0_{3 \times 1}$ ,  $\hat{e}_i(0) = e_i$ ,  $i = 1, 2, 3$ , and  $P(0) = I_{15}$ . The gain parameters for each case are chosen as  $\rho_1 = 0.5, \rho_2 = 0.3, \rho_3 = 0.2$  and  $k_R = 1$  for  $\sigma_R$  in (7), and  $Q = 10^3 I_{3N}, V = 10^{-4} I_{15}$  for the CRE (9). Simulation results are shown in Fig. 4. As one can see, the estimated states from all the cases converge, after a few seconds, to the vicinity of the real states. Roughly speaking, with the same tuning parameters, the performances of the proposed continuous observer with three types of vision-based measurements is quite similar.

<sup>2</sup>In practice, 3D landmark positions can be obtained, for instance, using stereo cameras, and the sufficient condition for uniform observability of the pair  $(A(t), C)$  with  $A(t)$  in (10) is the same as the one in Lemma 2.



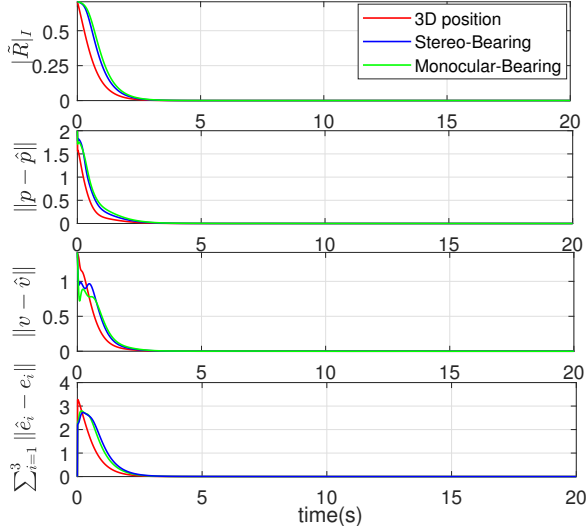


Fig. 4: Simulation results of the nonlinear observer (6) using 3D landmark position, stereo-bearing and monocular-bearing measurements.

## VII. EXPERIMENTAL RESULTS

To further validate the performance of the proposed observer, two sets of experiments have been presented using the data from the EuRoC dataset [24], where the trajectories are generated by a real flight of a quadrotor. This dataset includes stereo images, IMU measurements, and ground truth obtained from Vicon motion capture system. More details about the EuRoC dataset can be found in [24]. Since the sampling rate of the stereo camera (20Hz) is much lower than that of the IMU (200Hz), the proposed hybrid observer (23) was implemented.

The accelerometer and gyro measurements are compensated using the biases provided in the dataset. The features are tracked via the Kanade-Lucas-Tomasi (KLT) tracker [38], with the RANSAC outliers removal algorithm. Due to the lack of physical landmarks in this dataset, a set of ‘virtual’ landmarks (in the inertial frame) are generated as [23]. The IMU measurements are not continuous although obtained at a high rate. The same numerical integration methods in [23] for the estimated states are considered. The monocular-bearing measurements are obtained from the right camera.

For comparison purposes, the IEKF developed in [13] using 3D landmark position measurements has been considered. All the observers are concurrently executed using the same set of landmarks and visual measurements. The initial conditions for the estimated states are given as  $\hat{R}(0) = \exp(0.1\pi u^\times)$  with  $u \in \mathbb{S}^2$ ,  $\hat{p}(0) = \hat{v}(0) = 0_{3 \times 1}$  and  $\hat{e}_i(0) = e_i, i \in \{1, 2, 3\}$  and  $P(0) = I_{15}$ . The scalar gain parameters are chosen as  $k_R = 20$ ,  $\rho_1 = 0.5$ ,  $\rho_2 = 0.3$  and  $\rho_3 = 0.2$ . For the CDRE (25), matrices  $V(t)$  and  $Q(t)$  are tuned using (27) (with an additional small identity matrix  $0.002I$ ) with  $\text{Cov}(n_x) = \text{blkdiag}(0.0024I_3, 0.028I_3)$ ,  $\text{Cov}(n_y) = 0.0005I_{3N}$  for bearing measurements and  $\text{Cov}(n_y) = 0.06I_{3N}$  for 3D position measurements. For the IEKF, the gain matrices are tuned as per Section V.B in [13] using the same above mentioned covariance of the measurements noise. The results of the

first experiment are shown in Fig. 5. The estimates, provided by the proposed observer and the IEKF converge, after a few seconds, to the vicinity of the ground truth with a nice performance in terms of noise attenuation. The averaged position estimation errors after 10sec (i.e., at steady state) are as follows: 3.26cm for the proposed observer with 3D position measurements, 3.29cm for the proposed observer with stereo-bearing measurements, 10.99cm for the proposed observer with monocular-bearing measurements and 2.89cm for the IEKF with 3D position measurements. The proposed observer using monocular-bearing measurements comes with the largest position estimation error, which is mainly due to the fact that it takes less measurement information and requires stronger conditions (on the motion of the camera and the location of the landmarks) for uniform observability than the other settings.

In the second experiment, we consider the scenario where the measurements of the left camera are not available after 120sec. In this situation, after 120sec, the 3D landmark positions can not be constructed from a single camera, while the stereo-bearing measurements become the monocular-bearing measurements. The results of the second experiment are shown in Fig. 6. As one can see, the estimates, provided by the proposed observer using 3D landmark position measurements and IEKF, diverge after 120sec; while the estimates, provided by the proposed observer using stereo-bearing measurements, stay in the vicinity of the ground truth.

## VIII. CONCLUSION

We addressed the problem of simultaneous estimation of the attitude, position and linear velocity for vision-aided INSs. An AGAS nonlinear observer on  $SO(3) \times \mathbb{R}^{15}$  has been proposed using body-frame acceleration and angular velocity measurements, as well as body-frame stereo (or monocular) bearing measurements of some known landmarks in the inertial frame. A detailed uniform observability analysis has been carried out for the monocular and stereo bearing measurements cases. In the stereo bearing measurements case, uniform observability is guaranteed as long as there exist three non-aligned landmarks whose plane is not parallel to the gravity vector. In the monocular bearing case, on top of the condition of the stereo bearing case, it is required that none of the body-frame bearings (of the three non-aligned landmarks whose plane is not orthogonal to the gravity vector) maintains the same direction indefinitely. In the case of a monocular bearing, with a motionless camera, which is known as the static PnP problem, our observer provides a viable solution as long as we have five landmarks that are not in one of the four configurations shown in Fig. 3.

For practical implementation purposes, we proposed a hybrid version of our nonlinear observer to handle the case where the IMU measurements are continuous and the bearing measurements are intermittently sampled. This observer has been validated using the EuRoC dataset experimental data of a real quadrotor flight. To illustrate the benefit of using bearing measurements over landmark position measurements, we implemented the bearing-based and landmark-position-based observers in a scenario where one of the cameras loses sight

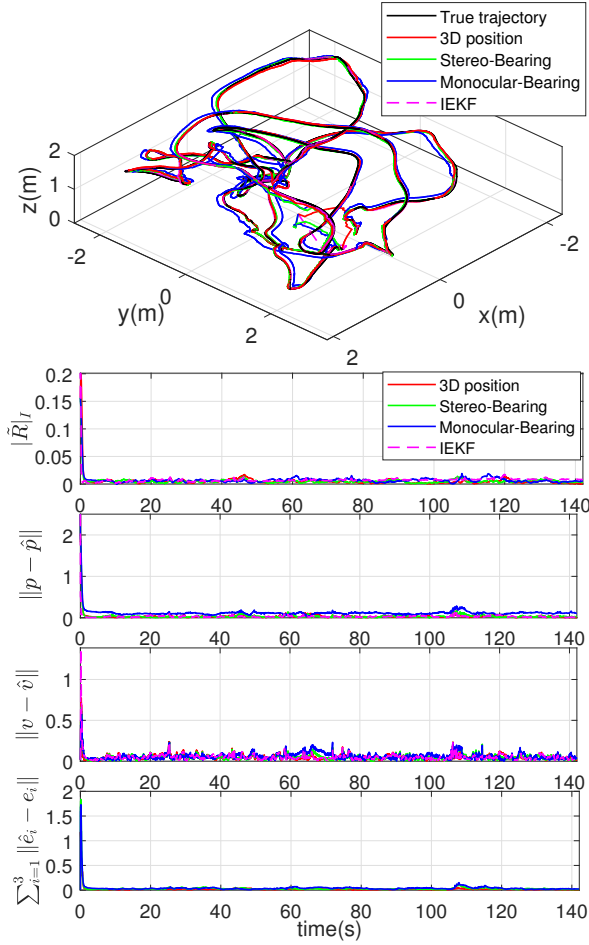


Fig. 5: Experimental results using Vicon Room 1 01 of the EuRoC dataset [24].

of the landmarks after some time, and the results are shown in Fig. 6. As a future work, we intend to enhance our proposed observer with the estimation of the accelerometer and angular velocity biases while preserving the AGAS property for the overall closed-loop system.

## APPENDIX

### A. Proof of Theorem 1

Before proceeding with the proof of Theorem 1, some useful properties on  $SO(3)$  are given in the following lemma, whose proof can be found in [17], [39]

**Lemma 4.** Consider the trajectory  $\dot{R} = R\omega^\times$  with  $R(0) \in SO(3)$  and  $\omega \in \mathbb{R}^3$ . Let  $\mathcal{L}_M(R) = \text{tr}((I_3 - R)M)$  be the potential function on  $SO(3)$  with  $M = M^\top$  a positive semi-definite matrix. Then, for all  $x, y \in \mathbb{R}^3$  the following properties hold:

$$4\lambda_{\min}^{\bar{M}}|R|_I^2 \leq \mathcal{L}_M(R) \leq 4\lambda_{\max}^{\bar{M}}|R|_I^2, \quad (29)$$

$$\|\psi_a(MR)\|^2 = \alpha(M, R) \text{tr}((I_3 - R)\underline{M}), \quad (30)$$

$$\dot{\psi}_a(MR) = E(MR)\omega, \quad (31)$$

$$\|E(MR)\|_F \leq \|\bar{M}\|_F, \quad (32)$$

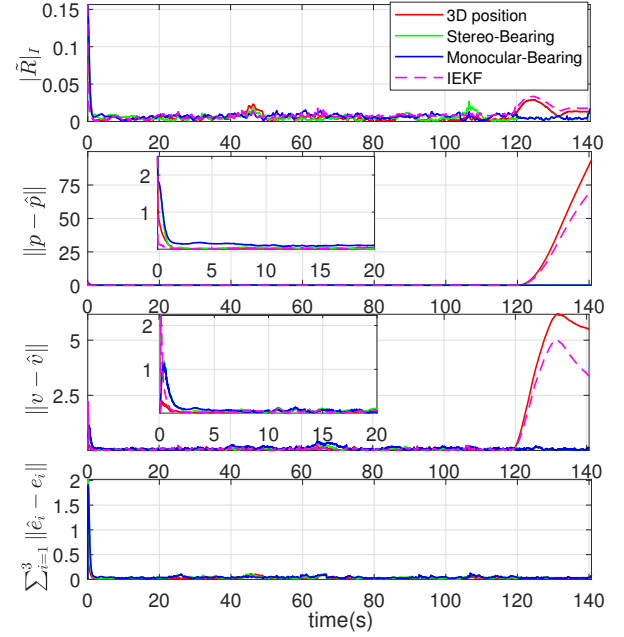


Fig. 6: Experimental results using Vicon Room 1 01 of the EuRoC dataset [24]. The measurements of the left camera are not available after 120sec.

where  $\bar{M} := \frac{1}{2}(\text{tr}(M)I - M)$ ,  $\underline{M} := \text{tr}(\bar{M}^2)I_3 - 2\bar{M}^2$ ,  $E(MR) = \frac{1}{2}(\text{tr}(MR)I_3 - R^\top M)$ , and the map  $\alpha(M, R) := (1 - |R|_I^2) \cos\langle u, \bar{M}u \rangle$  with  $u \in \mathbb{S}^2$  denoting the axis of the rotation  $R$  and  $\langle \cdot, \cdot \rangle$  denoting the angle between two vectors.

Consider the following real-valued function on  $\mathbb{R}^{15}$ :

$$\mathcal{L}_P(\tilde{x}) = \tilde{x}^\top P^{-1} \tilde{x}, \quad (33)$$

where  $P(t)$  is solution to the CRE (9). Note that  $A(t)$  in (10) is continuous and bounded since  $\omega(t)$  is continuous and bounded. By assumptions, it easy to conclude from [29], [30] that there exist  $0 < p_m \leq p_M < \infty$  such that  $p_m I_{15} \leq P(t) \leq p_M I_{15}$  for all  $t \geq 0$ . It follows that  $\frac{1}{p_M} \|\tilde{x}\|^2 \leq \mathcal{L}_P(\tilde{x}) \leq \frac{1}{p_m} \|\tilde{x}\|^2$ . In view of (16b) and (9), the time-derivative of  $\mathcal{L}_P$  is given by

$$\begin{aligned} \dot{\mathcal{L}}_P &= \tilde{x}^\top (P^{-1}A + A^\top P^{-1} - 2C^\top QC + \dot{P}^{-1})\tilde{x} \\ &= -\tilde{x}^\top P^{-1}VP^{-1}\tilde{x} - \tilde{x}^\top C^\top QC\tilde{x} \\ &\leq -\frac{v_m}{p_M^2} \|\tilde{x}\|^2 \leq -\lambda \mathcal{L}_P \end{aligned} \quad (34)$$

with  $\lambda = v_m p_m / p_M^2$ ,  $v_m := \inf_{t \geq 0} \lambda_{\min}^{V(t)}$ , where we made use of the facts  $C^\top QC \geq 0$  and  $\dot{P}^{-1} = -P^{-1}\dot{P}P^{-1} = -P^{-1}A - A^\top P^{-1} + C^\top QC - P^{-1}VP^{-1}$ . Hence, one has  $\|\tilde{x}(t)\| \leq \sqrt{p_M/p_m} \exp(-\frac{\lambda}{2}t) \|\tilde{x}(0)\|$ , which implies that  $\tilde{x}$  converges to zero exponentially, and  $\tilde{x}, \dot{\tilde{x}}$  are bounded. Note that the convergence of  $\tilde{x}$  is independent from the dynamics of the rotation. From (16), the equilibrium points of the system are given as  $(R^*, 0_{15 \times 1})$  with  $\|\psi_a(MR^*)\| = 0$ . Using the facts  $\psi_a(MR) = \text{vec} \circ \mathbb{P}_a(MR)$  and  $\mathbb{P}_a(MR) = (MR - R^\top M)/2$ , it follows that  $\|\psi_a(M\tilde{R})\| = 0$  implies, as shown in [6], that  $\tilde{R} \in \{R \in SO(3) : R = \mathcal{R}_\alpha(\pi, v), v \in \mathcal{E}(M)\}$ .

On the other hand, consider the following real-valued func-

tion on  $SO(3)$ :

$$\mathcal{L}_M(\tilde{R}) = \text{tr}((I - \tilde{R})M), \quad (35)$$

whose time-derivative is given by

$$\begin{aligned} \dot{\mathcal{L}}_M &= \text{tr}(-M\tilde{R}(-k_R\psi_a(M\tilde{R}) - \Gamma(t)\tilde{x})^\times) \\ &\leq -2k_R\|\psi_R\|^2 + 2c_\Gamma\|\tilde{x}\|\|\psi_R\| \end{aligned} \quad (36)$$

where  $\psi_R := \psi_a(M\tilde{R})$ , and we made use of the facts  $\text{tr}(-Au^\times) = \langle \langle A, u^\times \rangle \rangle = 2u^\top \psi_a(A)$  for any  $A \in \mathbb{R}^{3 \times 3}$ ,  $u \in \mathbb{R}^3$ , and  $\|\Gamma(t)\| \leq c_\Gamma$  for all  $t \geq 0$ .

Consider the following Lyapunov function candidate:

$$\mathcal{L}(\tilde{R}, \tilde{x}) = \mathcal{L}_M(\tilde{R}) + \kappa \mathcal{L}_P(\tilde{x}), \quad (37)$$

with some  $\kappa > 0$ . Let  $\zeta := [\|\psi_R\|, \|\tilde{x}\|]^\top$ . From (34) and (36), the time-derivative of  $\mathcal{L}$  is given by

$$\begin{aligned} \dot{\mathcal{L}} &\leq -2k_R\|\psi_R\|^2 + 2c_\Gamma\|\tilde{x}\|\|\psi_R\| - \kappa \frac{v_m}{p_M^2} \|\tilde{x}\|^2 \\ &= -\zeta^\top H \zeta, \quad H := \begin{bmatrix} 2k_R & -c_\Gamma \\ -c_\Gamma & \frac{\kappa v_m}{p_M^2} \end{bmatrix}. \end{aligned} \quad (38)$$

Choosing  $\kappa > c_\Gamma^2 p_M^2 / (2k_R v_m)$  such that matrix  $H$  is positive definite, it follows that  $\dot{\mathcal{L}} \leq 0$  and  $\mathcal{L}$  is non-increasing. Then, making use of the facts  $\mathcal{L}(\tilde{R}(t), \tilde{x}(t)) - \mathcal{L}(\tilde{R}(0), \tilde{x}(0)) = \int_0^t \dot{\mathcal{L}}(\tilde{R}(\tau), \tilde{x}(\tau)) d\tau \leq -\int_0^t \zeta^\top(\tau) H \zeta(\tau) d\tau$ , one verifies that  $\lim_{t \rightarrow \infty} \int_0^t \zeta^\top(\tau) H \zeta(\tau) d\tau$  exists and is finite. Since  $\tilde{x}$  is bounded and matrices  $A(t)$ ,  $C(t)$ ,  $Q(t)$  and  $P(t)$  are bounded, it is clear that  $\tilde{x}$  is bounded. From (29) and (30) in Lemma 4, one shows that  $\psi_R$  is bounded. In view of (16) and (31)-(32), one can easily verify that  $\dot{\psi}_R$  is bounded. Thus, it follows that the time-derivative of  $\zeta^\top H \zeta$  is bounded, which implies the uniform continuity of  $\zeta^\top H \zeta$ . Therefore, by virtue of Barbalat's lemma, one has  $\lim_{t \rightarrow \infty} \zeta^\top(t) H \zeta(t) = 0$ , i.e.,  $(\|\psi_R\|, \|\tilde{x}\|) \rightarrow (0, 0)$  as  $t \rightarrow \infty$ . This implies that, for any initial condition  $(\tilde{R}(0), \tilde{x}(0)) \in SO(3) \times \mathbb{R}^{15}$ , the solution  $(\tilde{R}, \tilde{x})$  to (16) converges to the set  $(I_3, 0_{15 \times 1}) \cup \Psi_M$ , which proves item (i).

Next, let us prove the local exponential stability of the equilibrium  $(I_3, 0_{15 \times 1})$  in item (ii). Let  $0 < \varepsilon_R < 4\lambda_{\min}^M$  and define the set  $U_{\varepsilon_R} = \{(R, x) \in SO(3) \times \mathbb{R}^{15} : \mathcal{L}(R, x) \leq \varepsilon_R\}$ . From (37)-(38) with  $\kappa > c_\Gamma^2 p_M^2 / (2k_R v_m)$ , for any initial condition  $(\tilde{R}(0), \tilde{x}(0)) \in U_{\varepsilon_R}$ , one has  $\mathcal{L}_M(\tilde{R}(t)) \leq \mathcal{L}(\tilde{R}(t), \tilde{x}(t)) \leq \mathcal{L}(\tilde{R}(0), \tilde{x}(0)) \leq \varepsilon_R$  for all  $t \geq 0$ . It follows from (29)-(30) that

$$|\tilde{R}|_I^2 \leq \varepsilon_R / 4\lambda_{\min}^M < 1 \quad (39)$$

$$\varrho |\tilde{R}|_I^2 \leq \|\psi_R\|^2 \leq 4\lambda_{\max}^W |\tilde{R}|_I^2 \quad (40)$$

with  $W := \frac{1}{2}(\text{tr}(\underline{M})I_3 - \underline{M}) = \bar{M}^2$  and  $\varrho := \min_{(R, x) \in U_{\varepsilon_R}} 4\alpha(M, R)\lambda_{\min}^W \geq 4(1 - \varepsilon_R / 4\lambda_{\min}^M)\lambda_{\min}^W > 0$ . Let  $\bar{\zeta} := [|\tilde{R}|_I, \|\tilde{x}\|]^\top$ . In view of (29), (33) and (35), one obtains

$$\underline{\alpha} \|\bar{\zeta}\|^2 \leq \mathcal{L} \leq \bar{\alpha} \|\bar{\zeta}\|^2 \quad (41)$$

where  $\underline{\alpha} := \min\{4\lambda_{\min}^M, \frac{\kappa}{p_M}\}$  and  $\bar{\alpha} := \max\{4\lambda_{\max}^M, \frac{\kappa}{p_M}\}$ . Substituting (40) into (38), one has

$$\dot{\mathcal{L}} \leq -2k_R \varrho |\tilde{R}|_I^2 + 4\sqrt{\lambda_{\max}^W c_\Gamma} \|\tilde{x}\| |\tilde{R}|_I - \frac{\kappa v_m}{p_M^2} \|\tilde{x}\|^2$$

$$\leq -\bar{\zeta}^\top \bar{H} \bar{\zeta}, \quad \bar{H} := \begin{bmatrix} 2k_R \varrho & -2\sqrt{\lambda_{\max}^W c_\Gamma} \\ -2\sqrt{\lambda_{\max}^W c_\Gamma} & \frac{\kappa v_m}{p_M^2} \end{bmatrix}. \quad (42)$$

Choosing  $\kappa > 2\lambda_{\max}^W c_\Gamma^2 p_M^2 / (k_R \varrho v_m)$ , one can show that both matrices  $H$  and  $\bar{H}$  are positive definite since  $\varrho \leq 4\lambda_{\min}^W \leq 4\lambda_{\max}^W$ . In view of (41) and (42), one concludes

$$\|\bar{\zeta}(t)\| \leq \sqrt{\bar{\alpha}/\underline{\alpha}} \exp(-\frac{1}{2}\lambda_{\min}^{\bar{H}} t) \|\bar{\zeta}(0)\| \quad (43)$$

for all  $t \geq 0$ , which implies that  $(\tilde{R}, \tilde{x})$  converges to  $(I_3, 0_{15 \times 1})$  exponentially for any initial condition  $(\tilde{R}(0), \tilde{x}(0)) \in U_{\varepsilon_R}$ . This completes the proof of item (ii).

Now, we need to show that the equilibria in  $\Psi_M$  are unstable. From (34), one shows that  $\tilde{x}$  converges to zero exponentially, and is independent from the dynamics of  $\tilde{R}$ . Then, we focus on the dynamics of (16) at  $\|\tilde{x}\| = 0$ . For each  $v \in \mathcal{E}(M)$ , let us define  $R_v^* = \mathcal{R}_\alpha(\pi, v)$  and the open set  $U_v^\delta := \{(R, x) \in SO(3) \times \mathbb{R}^{15} : R = R_v^* \exp(\epsilon^\times), \|\epsilon\| \leq \delta, \|x\| = 0\}$  with  $\delta$  sufficiently small. For any  $(\tilde{R}, \tilde{x}) \in U_v^\delta$ , pick a sufficiently small  $\epsilon$  such that  $(R_v^*)^\top \tilde{R} := \exp(\epsilon^\times) \approx I_3 + \epsilon^\times$ . Consequently, from (16a) one obtains the linearized dynamics of  $\epsilon$  around  $R_v^*$  as follows:

$$\dot{\epsilon} = -k_R W_v \epsilon \quad (44)$$

where  $\psi_a(k_R M R_v^*(I_3 + \epsilon^\times)) = k_R W_v \epsilon$  with  $W_v = W_v^\top = \frac{1}{2}(\text{tr}(M R_v^*) I_3 - (M R_v^*)^\top) = \frac{1}{2}(2v^\top M v - \text{tr}(M)) I_3 - \frac{1}{2}(2\lambda_v^M v v^\top - M)$ , and we made use of the facts  $M v = \lambda_v^M v$ ,  $\psi_a(M R_v^*) = 0$ ,  $\psi_a(M R_v^* \epsilon^\times) = \text{vec} \circ \mathbb{P}_a(M R_v^* \epsilon^\times)$  and  $\mathbb{P}_a(M R_v^* \epsilon^\times) = \frac{1}{2}(M R_v^* \epsilon^\times + \epsilon^\times (M R_v^*)^\top) = (W_v \epsilon)^\times$ . Since  $W_v = W_v^\top$  and  $M$  is positive semi-definite with three distinct eigenvalues, one verifies that  $-2v^\top W_v v = -v^\top M v + \text{tr}(M) > 0$ , which implies that  $-W_v, \forall v \in \mathcal{E}(M)$  has at least one positive eigenvalue. Then, one can conclude that all the equilibrium points in  $\Psi_M$  are unstable. Consider the subsystem  $\dot{\tilde{R}} = \tilde{R}(-k_R \psi_a(M \tilde{R}))^\times$  and its linearized dynamics around the undesired equilibrium points  $R_v^* = \mathcal{R}_\alpha(\pi, v)$ ,  $v \in \mathcal{E}(M)$  in (44). For each undesired equilibrium point  $R_v^*$ , there exist (local) stable and unstable manifolds, and the union of the stable manifolds and the undesired equilibria has dimension less than three [32, The Stable Manifold Theorem]. Then, the set of the union of the stable manifolds and the undesired equilibria has Lebesgue measure zero. It follows that the solution  $\tilde{R}(t)$  converges to  $I_3$  starting from all initial conditions except from a set of Lebesgue measure zero. Hence, one concludes that the equilibrium  $(I_3, 0_{15 \times 1})$  of the overall system (16) is almost globally asymptotically stable. This completes the proof.

## B. Proof of Lemma 1

The proof of this Lemma is motivated from [33, Lemma 3.1] and [27, Lemma 2.7]. In order to show that the pair  $(A, C(t))$  is uniformly observable, we are going to verify the existence of constants  $\delta, \mu > 0$  such that  $z^\top W_o(t, t + \delta) z = \frac{1}{\delta} \int_t^{t+\delta} \|C(\tau) \Phi(\tau, t) z\|^2 d\tau \geq \mu$  for all  $t \geq 0$  and  $z \in \mathbb{S}^{n-1}$ . Let us proceed by contradiction and assume that the pair

$(A, C(t))$  is not uniformly observable, i.e.,

$$\forall \bar{\delta}, \bar{\mu} > 0, \exists t \geq 0, \min_{z \in \mathbb{S}^{n-1}} \frac{1}{\bar{\delta}} \int_t^{t+\bar{\delta}} \|C(\tau)\Phi(\tau, t)z\|^2 d\tau < \bar{\mu}. \quad (45)$$

Consider a sequence  $\{\mu_q\}_{q \in \mathbb{N}}$  of positive numbers converging to zero, and an arbitrary positive scalar  $\bar{\delta}$ . Then, there must exist a sequence of time instants  $\{t_q\}_{q \in \mathbb{N}}$  and a sequence of vectors  $\{z_q\}_{q \in \mathbb{N}}$  with  $z_q \in \mathbb{S}^{n-1}$  such that  $\frac{1}{\bar{\delta}} \int_{t_q}^{t_q+\bar{\delta}} \|C(\tau)\Phi(\tau, t_q)z_q\|^2 d\tau < \bar{\mu}_q$  for any  $q \in \mathbb{N}$ . Since the set  $\mathbb{S}^{n-1}$  is compact, there exists a sub-sequence of  $\{z_q\}_{q \in \mathbb{N}}$  which converges to a limit  $\bar{z} \in \mathbb{S}^{n-1}$ . Moreover, since  $C(t)$  is bounded and the interval of integration in (45) is fixed, it follows from (45) that

$$\lim_{q \rightarrow \infty} \int_0^{\bar{\delta}} \|C(t_q + \tau)\Phi(t_q + \tau, t_q)\bar{z}\|^2 d\tau = 0 \quad (46)$$

by a change of integration variable. Since  $A = S + N$  and  $SN = NS$ , the state transition matrix  $\Phi(t_q + \tau, t_q)$  can be explicitly expressed as  $\Phi(t_q + \tau, t_q) = \exp(A\tau) = \exp(S\tau)\exp(N\tau)$ . Then, (46) is equivalent to  $\lim_{q \rightarrow \infty} \int_0^{\bar{\delta}} \|C(t_q + \tau)\exp(S\tau)\exp(N\tau)\bar{z}\|^2 d\tau = 0$ , which implies

$$\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)\exp(S\tau)\exp(N\tau)\bar{z}\|^2 d\tau = 0 \quad (47)$$

with some  $0 < \delta < \bar{\delta}$ . Consider now the following technical results whose proofs are given after this proof.

**Lemma 5.** *From (47) with  $\bar{\delta}$  large enough, one has*

$$\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)\exp(N\tau)\bar{z}\| d\tau = 0. \quad (48)$$

**Lemma 6.** *From (48) with  $\bar{\delta}$  large enough, one has*

$$\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)N^k\bar{z}\|^2 d\tau = 0 \quad (49)$$

for all  $k = 0, 1, \dots, s-1$ .

Using the fact that the matrix  $\mathcal{O}(t)$  is composed of row vectors of  $C(t)$ ,  $C(t)N$ ,  $\dots$ ,  $C(t)N^{s-1}$ , one has  $\sum_{k=0}^{s-1} \|C(t_q + \tau)N^k\bar{z}\|^2 \geq \|\mathcal{O}(t_q + \tau)\bar{z}\|^2$  for every  $q \in \mathbb{N}_{>0}$ . It follows from (48) and (49) that

$$\begin{aligned} \lim_{q \rightarrow \infty} \int_{t_q+\bar{\delta}-\delta}^{t_q+\bar{\delta}} \|\mathcal{O}(\tau)\bar{z}\|^2 d\tau &= \lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|\mathcal{O}(t_q + \tau)\bar{z}\|^2 d\tau \\ &\leq \lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \sum_{k=0}^{s-1} \|C(t_q + \tau)N^k\bar{z}\|^2 d\tau = 0. \end{aligned} \quad (50)$$

From (18), one can show that  $\int_{t_q+\bar{\delta}-\delta}^{t_q+\bar{\delta}} \|\mathcal{O}(\tau)\bar{z}\|^2 d\tau = \bar{z}^\top (\int_{t_q+\bar{\delta}-\delta}^{t_q+\bar{\delta}} \mathcal{O}^\top(\tau)\mathcal{O}(\tau) d\tau) \bar{z} > \mu$ , which contradicts (50). It implies that (49) is not true for all  $k = 0, 1, \dots, s-1$ , and in turns (46) and (47) are not true when  $\bar{\delta}$  is large enough. Therefore, one can always find  $\bar{\delta}$  large enough, such that (45) does not hold. Consequently, one concludes that the pair  $(A, C(t))$  is uniformly observable. It remains to prove Lemma 5 and 6.

*1) Proof of Lemma 5:* We are going to show that (47) implies (48) provided that  $\bar{\delta}$  is large enough. Since  $S$  is diagonalizable, from [32, Theorem 1] there exists an invertible matrix  $P$  such that  $D = \text{diag}(\lambda_1, \dots, \lambda_n) = P^{-1}SP$  with  $\lambda_1, \dots, \lambda_n$  denoting the eigenvalues of  $A$  repeated according to their multiplicity. Then, one obtains  $\exp(S\tau) = P \exp(D\tau)P^{-1}$ . Let  $\bar{z}'(\tau) = P^{-1} \exp(N\tau)\bar{z}$  and rewrite the eigenvalues of  $D$  as  $\lambda'_1 < \dots < \lambda'_d$  with  $d$  denoting the number of distinct eigenvalues. It follows that  $\exp(D\tau)P^{-1} \exp(N\tau)\bar{z} = \exp(D\tau)\bar{z}'(\tau) = \sum_{i=1}^d \exp(\lambda'_i \tau) \bar{z}'_i(\tau)$  with  $\bar{z}'_i(\tau) \in \mathbb{R}^{15}$  for all  $i = 1, \dots, d$  and  $\sum_{i=1}^d \bar{z}'_i(\tau) = \bar{z}'(\tau)$ . We assume that there exist a constant  $\epsilon' > 0$  and a sub-sequence of  $\{i\}_{1 \leq i \leq d} \subset \mathbb{N}_{>0}$  such that

$$\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)P\bar{z}'_i(\tau)\| d\tau > \epsilon'. \quad (51)$$

Then, one can always pick the largest  $j$  from the sub-sequence. For the sake of simplicity, define  $\nu_j(\tau) = \sum_{i=1}^j \exp(\lambda'_i \tau) \bar{z}'_i(\tau)$ . Using the facts  $\|C(t_q + \tau)P\nu_j(\tau)\| \leq \|C(t_q + \tau)P\nu_d(\tau)\| + \sum_{i=j+1}^d \exp(\lambda'_i \tau) \|C(t_q + \tau)P\bar{z}'_i(\tau)\|$ ,  $P\nu_d(\tau) = P \exp(D\tau)P^{-1} \exp(N\tau)\bar{z} = \exp(S\tau)\exp(N\tau)\bar{z}$  and  $\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)P\bar{z}'_i(\tau)\| = 0$  for all  $j+1 \leq i \leq d$ , from (47) one can show that

$$\begin{aligned} &\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)P\nu_j(\tau)\| d\tau \\ &\leq \sum_{i=j+1}^d \exp(|\lambda'_i|\bar{\delta}) \lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)P\bar{z}'_i(\tau)\| d\tau \\ &\quad + \lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)\exp(S\tau)\exp(N\tau)\bar{z}\| d\tau = 0. \end{aligned} \quad (52)$$

Let  $\eta'(\tau) = \sum_{i=1}^{j-1} \exp((\lambda'_i - \lambda'_j)\tau) \bar{z}'_i(\tau)$  such that  $\exp(-\lambda'_j \tau) \nu_j(\tau) = \bar{z}'_j(\tau) + \eta'(\tau)$ . Using the facts  $\sum_{i=1}^d \bar{z}'_i(\tau) = \bar{z}'(\tau) = P^{-1} \exp(N\tau)\bar{z} = P^{-1} \sum_{k=0}^{s-1} N^k \bar{z}$  and  $\lim_{t \rightarrow \infty} \exp(-at)t^b = 0$  for any  $a, b > 0$ , one has  $\eta'(\tau) \rightarrow 0$  as  $\tau \rightarrow \infty$ . Since  $C(t)$  is continuous and bounded by assumption, there exist a positive constant  $\gamma'$  such that  $\gamma' = \sup_{t \geq 0} \|C(t)P\|$ . Then, from (51) one can show that

$$\begin{aligned} &\exp(|\lambda'_j|\bar{\delta}) \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)P\nu_j(\tau)\| d\tau \\ &\geq \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)P \exp(-\lambda'_j \tau) \nu_j(\tau)\| d\tau \\ &\geq \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)P(\bar{z}'_j(\tau) + \eta'(\tau))\| d\tau \\ &\geq \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)P\bar{z}'_j(\tau)\| d\tau - \gamma' \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|\eta'(\tau)\| d\tau \end{aligned}$$

Choosing  $\bar{\delta}$  large enough such that  $\sup_{\tau \in [\bar{\delta}-\delta, \bar{\delta}]} \|\eta'(\tau)\| \leq \frac{\epsilon'}{2\delta\gamma'}$ , it follows that  $\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)P\nu_j(\tau)\| d\tau \geq \frac{\epsilon'}{2} \exp(-|\lambda'_j|\bar{\delta}) > 0$ , which contradicts (52). It means that the sub-sequence satisfying (51) does not exist, i.e., the assumption according to (51) does not hold. Therefore, one obtains  $\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau)P\bar{z}'_i(\tau)\| d\tau = 0$  for all



$i = 1, 2, \dots, d$ . Using the fact  $\exp(N\tau)\bar{z} = P \sum_{i=1}^d \bar{z}'_i(\tau)$ , one can show that  $\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) \exp(N\tau)\bar{z}\| d\tau \leq \sum_{i=1}^d \lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) P \bar{z}'_i(\tau)\| d\tau = 0$ , which gives (48).

2) *Proof of Lemma 6:* We are going to show (49) for all  $k = 0, 1, \dots, s-1$  from (48) with some  $\bar{\delta}$  large enough. Let us proceed by contradiction and assume that (49) does not hold for any  $0 \leq k \leq s-1$ , i.e., there exist a constant  $\epsilon > 0$  and a sub-sequence of  $\{k\}_{0 \leq k \leq s-1} \subset \mathbb{N}$  such that

$$\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) N^k \bar{z}\| d\tau > \epsilon. \quad (53)$$

One can always pick the largest  $\bar{k}$  such that (53) holds. Since  $N$  is a nilpotent matrix of order  $s \leq n$ , one has  $\exp(N\tau) = \sum_{k=0}^{s-1} \frac{\tau^k}{k!} N^k$ . For the sake of simplicity, define  $\Sigma_k(\tau) = \sum_{i=0}^k \frac{\tau^i}{i!} N^i$  with  $k \in \mathbb{N}$ . Using the fact  $\|C(t_q + \tau) \Sigma_{\bar{k}}(\tau) \bar{z}\| \leq \|C(t_q + \tau) \exp(N\tau) \bar{z}\| + \sum_{i=\bar{k}+1}^{s-1} \frac{\tau^i}{i!} \|C(t_q + \tau) N^i \bar{z}\|$  and  $\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) N^i \bar{z}\| d\tau = 0$  for all  $\bar{k} + 1 \leq i \leq s-1$ , from (48) one can show that

$$\begin{aligned} & \lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) \Sigma_{\bar{k}}(\tau) \bar{z}\| d\tau \\ & \leq \lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) \exp(N\tau) \bar{z}\| d\tau \\ & + \lim_{q \rightarrow \infty} \sum_{i=\bar{k}+1}^{s-1} \frac{\bar{\delta}^i}{i!} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) N^i \bar{z}\| d\tau = 0. \end{aligned} \quad (54)$$

Let  $\eta(\tau) = \frac{\bar{k}!}{\tau^{\bar{k}}} \Sigma_{\bar{k}-1}(\tau) \bar{z}$  such that  $\frac{\bar{k}!}{\tau^{\bar{k}}} \Sigma_{\bar{k}}(\tau) \bar{z} = N^{\bar{k}} \bar{z} + \eta(\tau)$ . It is easy to show that  $\eta(\tau)$  is bounded and  $\lim_{t \rightarrow \infty} \eta(\tau) = 0$ . Since  $C(t)$  is bounded by assumption, there exist a positive constant  $\gamma$  such that  $\gamma = \sup_{t \geq 0} \|C(t)\|$ . Then, one obtains

$$\begin{aligned} & \frac{\bar{k}!}{(\bar{\delta}-\delta)^{\bar{k}}} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) \Sigma_{\bar{k}}(\tau) \bar{z}\| d\tau \\ & \geq \int_{\bar{\delta}-\delta}^{\bar{\delta}} \left\| C(t_q + \tau) \frac{\bar{k}!}{\tau^{\bar{k}}} \Sigma_{\bar{k}}(\tau) \bar{z} \right\| d\tau \\ & = \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) (N^{\bar{k}} \bar{z} + \eta(\tau))\| d\tau \\ & \geq \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) N^{\bar{k}} \bar{z}\| d\tau - \gamma \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|\eta(\tau)\| d\tau. \end{aligned}$$

Choosing  $\bar{\delta}$  large enough such that  $\sup_{\tau \in [\bar{\delta}-\delta, \bar{\delta}]} \|\eta(\tau)\| \leq \frac{\epsilon}{2\delta\gamma}$ , from (53) one can show that  $\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \tau) \Sigma_{\bar{k}}(\tau) \bar{z}\| d\tau \geq \epsilon(\bar{\delta} - \delta)^{\bar{k}} / (2\bar{k}!) > 0$ , which contradicts (54), which means that the sub-sequence satisfying (53) does not exist, i.e., assumption (53) does not hold. Thus, one can show that  $\lim_{q \rightarrow \infty} \int_{\bar{\delta}-\delta}^{\bar{\delta}} \|C(t_q + \bar{\delta} - \delta + \tau) N^k \bar{z}\| d\tau = 0$  for all  $k = 0, 1, \dots, s-1$ . Therefore, one can conclude that (49) holds for all  $k = 0, 1, \dots, s-1$  with some  $\bar{\delta}$  large enough. This completes the proof.

### C. Proof of Lemma 2

From the definition of  $A(t)$  in (10), one can rewrite  $A(t) = \bar{A} + S(t)$  with a block diagonal skew symmetric matrix  $S(t) = \text{blkdiag}(-\omega^\times, -\omega^\times, \dots, -\omega^\times) \in \mathbb{R}^{15 \times 15}$

and a constant matrix  $\bar{A}$  such that  $\bar{A}S(t) = S(t)\bar{A}$ . Let us introduce the following block diagonal matrix:

$$T(t) = \text{blkdiag}(R^\top, R^\top, \dots, R^\top) \in \mathbb{R}^{15 \times 15} \quad (55)$$

whose dynamics are given as  $\dot{T}(t) = S(t)T(t)$ . One can verify that  $T(t)T^\top(t) = I_{15}$  and  $T(t)\bar{A} = \bar{A}T(t)$ . Let  $\bar{\Phi}(t, \tau) = \exp(\bar{A}(t - \tau))$  be the state transition matrix associated to  $\bar{A}$ . Using similar steps as in the proof of [15, Lemma 3], the state transition matrix associated to  $A(t)$  can be written as

$$\Phi(t, \tau) = T(t)\bar{\Phi}(t, \tau)T^{-1}(\tau), \quad (56)$$

with the properties:  $\frac{d}{dt}\Phi(t, \tau) = A(t)\Phi(t, \tau)$ ,  $\Phi(t, t) = I$ ,  $\Phi^{-1}(t, \tau) = \Phi(\tau, t)$  for all  $t, \tau \geq 0$ , and  $\Phi(t_3, t_2)\Phi(t_2, t_1) = \Phi(t_3, t_1)$  for every  $t_1, t_2, t_3 \geq 0$ . Moreover, one can rewrite  $C(t)$  as  $C(t) = \Theta(t)\bar{C}$  with  $\Theta(t) = \text{blkdiag}(\Pi_1(t), \dots, \Pi_N(t)) \in \mathbb{R}^{3N \times 3N}$  and  $\bar{C}$  defined in (20). Define a new block matrix

$$\bar{T}(t) = \text{blkdiag}(R^\top, R^\top, \dots, R^\top) \in \mathbb{R}^{3N \times 3N}. \quad (57)$$

One can show that  $\bar{C}T(t) = \bar{T}(t)\bar{C}$  and  $\bar{T}^\top(t)\bar{T}(t) = I_{3N}$ . From (2) and (55)-(57), one can show that

$$\begin{aligned} W_o(t, t + \delta) &= \frac{1}{\delta} \int_t^{t+\delta} \Phi^\top(\tau, t) \bar{C}^\top \Theta^\top(\tau) \Theta(\tau) \bar{C} \Phi(\tau, t) d\tau \\ &\geq T(t) W'_o(t, t + \delta) T^{-1}(t), \end{aligned} \quad (58)$$

where  $W'_o(t, t + \delta) = \frac{1}{\delta} \int_t^{t+\delta} \bar{\Phi}^\top(\tau, t) \bar{C}^\top \bar{\Theta}^\top(\tau) \bar{\Theta}(\tau) \bar{C} \bar{\Phi}(\tau, t) d\tau$  with  $\bar{\Theta}(t) = \bar{T}^\top(t) \Theta(t) \bar{T}(t) = \text{blkdiag}(\bar{\Pi}_1, \dots, \bar{\Pi}_N)$  and  $\bar{\Pi}_i = R \Pi_i R^\top$  for each  $i = 1, 2, \dots, N$ .

Next, we are going to show that the pair  $(\bar{A}, \bar{\Theta}(t)\bar{C})$  is uniformly observable. Since  $\bar{A}$  is nilpotent with  $\bar{A}^3 = 0$ , from Lemma 1, the pair  $(\bar{A}, \bar{\Theta}(t)\bar{C})$  is uniformly observable if there exist scalars  $\delta, \mu > 0$  such that condition (18) holds with  $\mathcal{O}(t) = [(\bar{\Theta}(t)\bar{C})^\top, (\bar{\Theta}(t)\bar{C}\bar{A})^\top, (\bar{\Theta}(t)\bar{C}\bar{A}^2)^\top]^\top$ . Since  $\Pi_i$  is uniformly positive definite for each  $i = 1, 2, \dots, N$ , one obtains  $\bar{\Pi}_i, i = 1, 2, \dots, N$  and  $\bar{\Theta}(t)$  uniformly positive definite. Then, one obtains that  $\text{rank}(\mathcal{O}) = \text{rank}([\bar{C}^\top, (\bar{C}\bar{A})^\top, (\bar{C}\bar{A}^2)^\top]^\top)$ . Let  $N_1, N_2 \in \mathbb{R}^{3 \times 15}$  be the first three rows of  $\bar{C}\bar{A}$  and  $\bar{C}\bar{A}^2$ , respectively. One can show that  $\text{rank}(\mathcal{O}) = \text{rank}(\bar{\mathcal{O}})$  with

$$\bar{\mathcal{O}} := [\bar{C}^\top, N_1^\top, N_2^\top]^\top. \quad (59)$$

Since there exist at least three non-aligned landmarks, for the sake of simplicity, we assume that the first three landmarks are not aligned. Define  $u_i = [u_{i1}, u_{i2}, u_{i3}]^\top := p_1 - p_{i+1}$  for  $i = 1, 2$ . One can easily show that  $u_1, u_2$  and  $g$  are linearly independent, since the plane of these three landmarks is not parallel to the gravity vector. Let  $N_0$  be the first nine rows of  $\bar{C}$ . Applying the matrix row and column operations on matrix  $[N_0^\top, N_1^\top, N_2^\top]^\top$ , we obtain the following matrix:

$$\bar{\mathcal{O}}' = \begin{bmatrix} I_3 & -p_{11}I_3 & -p_{12}I_3 & -p_{13}I_3 & 0_3 \\ 0_3 & u_{11}I_3 & u_{12}I_3 & u_{13}I_3 & 0_3 \\ 0_3 & u_{21}I_3 & u_{22}I_3 & u_{23}I_3 & 0_3 \\ 0_3 & 0_3 & 0_3 & 0_3 & I_3 \\ 0_3 & g_1I_3 & g_2I_3 & g_3I_3 & 0_3 \end{bmatrix}. \quad (60)$$

One can show that  $\bar{\mathcal{O}}'$  has full rank of 15 since  $u_1, u_2$  and  $g$

are linearly independent. Then, it is straightforward to verify that  $\text{rank}(\mathcal{O}) = \text{rank}(\bar{\mathcal{O}}) \geq \text{rank}(\bar{\mathcal{O}}') = 15$ , which implies that the pair  $(\bar{A}, \bar{C})$  is uniformly observable, and there exist constants  $\delta, \mu' > 0$  such that  $W_o'(t, t + \delta) \geq \mu' I_{15}$  for all  $t \geq 0$ . Choosing  $0 < \mu < \bar{\epsilon}\mu'$ , from (58), it follows that  $W_o(t, t + \delta) \geq T(t)W_o'(t, t + \delta)T^{-1}(t) \geq \bar{\epsilon}\mu'T(t)T^{-1}(t) \geq \mu I_{15}$ , i.e., the pair  $(A(t), C(t))$  is uniformly observable. This completes the proof.

#### D. Proof of Lemma 3

Following similar steps as in the proof of Lemma 2, from (2) and (55)-(57), one can show that

$$W_o(t, t + \delta) = \frac{1}{\delta} \int_t^{t+\delta} \Phi(\tau, t)^\top C^\top(\tau) C(\tau) \Phi(\tau, t) d\tau \\ = T(t)W_o'(t, t + \delta)T^\top(t), \quad (61)$$

where  $W_o'(t, t + \delta) = \frac{1}{\delta} \int_t^{t+\delta} \bar{\Phi}^\top(\tau, t) \bar{C}^\top \bar{\Theta}^\top(\tau) \bar{\Theta}(\tau) \bar{C} \bar{\Phi}(\tau, t) d\tau$  with  $\bar{\Theta}(t) = \bar{T}^\top(t) \bar{\Theta}(t) \bar{T}(t) = \text{blkdiag}(\bar{\Pi}_1, \dots, \bar{\Pi}_N)$  and  $\bar{\Pi}_i = R \Pi_i R^\top = I_3 - \frac{p_i - p'}{\|p_i - p'\|}$  with  $p' = p + R p_c$  for each  $i = 1, 2, \dots, N$ .

Next, we are going to show that the pair  $(\bar{A}, \bar{\Theta}(t)\bar{C})$  is uniformly observable. Since  $\bar{A}$  is nilpotent with  $\bar{A}^3 = 0$ , from Lemma 1, the pair  $(\bar{A}, \bar{\Theta}(t)\bar{C})$  is uniformly observable if there exist scalars  $\delta, \mu > 0$  such that inequality (18) holds with  $\mathcal{O}(t) = [(\bar{\Theta}(t)\bar{C})^\top, (\bar{\Theta}(t)\bar{C}\bar{A})^\top, (\bar{\Theta}(t)\bar{C}\bar{A}^2)^\top]^\top$ . To verify the above condition, we proceed by contradiction. Assume that

$$\forall \bar{\delta}, \bar{\mu} > 0, \exists t \geq 0, \min_{z \in \mathbb{S}^{n-1}} \frac{1}{\bar{\delta}} \int_t^{t+\bar{\delta}} \|\mathcal{O}(\tau)z\|^2 d\tau < \bar{\mu}. \quad (62)$$

Similar to the arguments in the proof of Lemma (1), consider a sequence  $\{\mu_q\}_{q \in \mathbb{N}}$  of positive numbers converging to zero, and an arbitrary positive scalar  $\bar{\delta}$ . Then, there must exist a sequence of time instants  $\{t_q\}_{q \in \mathbb{N}}$  and a sequence of vectors  $\{z_q\}_{q \in \mathbb{N}}$  with  $z_q \in \mathbb{S}^{n-1}$  such that for any  $q \in \mathbb{N}$  one has  $\int_{t_q}^{t_q+\bar{\delta}} \|\mathcal{O}(\tau)z_q\|^2 d\tau < \bar{\mu}_q$ . Since  $\mathbb{S}^{n-1}$  is compact, there exists a sub-sequence of  $\{z_q\}_{q \in \mathbb{N}}$  which converges to a limit  $\bar{z} \in \mathbb{S}^{n-1}$ . Therefore, since  $\bar{\Theta}(t)$  is bounded, one has

$$\lim_{q \rightarrow \infty} \int_0^{\bar{\delta}} \|\mathcal{O}(t_q + \tau)\bar{z}\|^2 d\tau = 0 \quad (63)$$

by a change of integration variable. Consider the following technical result whose proof is given at the end of this proof.

**Lemma 7.** Let  $\bar{\mathcal{O}}(t) = [(\bar{\Theta}(t)\bar{C})^\top, N_1^\top, N_2^\top]^\top$ . From (63), one has

$$\lim_{q \rightarrow \infty} \|\bar{\mathcal{O}}(t_q + s)\bar{z}\|^2 = 0, \forall s \in [0, \bar{\delta}]. \quad (64)$$

Next, we consider the following cases

- i) *Camera in motion:* Let  $\bar{C}_i = [I_3, p_{i1}I_3, p_{i2}I_3, p_{i3}I_3, 0_3]$  associated to  $i$ -th landmark. Note that  $\ell_1, \ell_2, \ell_3$  are not-aligned and their plane is not parallel to the gravity vector. Inequality  $\bar{C}_i \bar{z} \neq 0_{3N \times 1}$  implies that there exists at least one of  $\bar{C}_i \bar{z}, i = \ell_1, \ell_2, \ell_3$ , which is different from zero. Without loss of generality, let  $\bar{z}'' := \bar{C}_i \bar{z} \neq 0$  with  $i \in \{\ell_1, \ell_2, \ell_3\}$ . From (64), one has

$$\lim_{q \rightarrow \infty} \|\bar{\Pi}_i(t_q + s)\bar{C}_i \bar{z}\|^2 = 0, \quad s \in [0, \bar{\delta}]. \quad (65)$$

For the sake of simplicity, let  $u(t) := R(t)R_c y_i(t) \in \mathbb{S}^2$ . Using the facts  $\bar{\Pi}_i = \pi(u)$  and  $\|\pi(u)y\|^2 = y^\top \pi(u)y = -y^\top (u^\times)^2 y = \|x \times y\|^2$ , for any  $u \in \mathbb{S}^2, y \in \mathbb{R}^3$ , it follows from (65) that  $\lim_{q \rightarrow \infty} \|u(t_q + s) \times \bar{z}''\|^2 = 0, \forall s \in [0, \bar{\delta}]$ . This implies that for any  $\bar{\mu}'$ , there exists  $q^*$  such that for all  $q \geq q^*$

$$\|u(t_q + s) \times \bar{z}''\|^2 < \bar{\mu}', \quad \forall s \in [0, \bar{\delta}]. \quad (66)$$

Motivated by the proof of [28, Lemma 4], let  $u_1 = u(t_q)$  and  $u_2 = u(t_q + \bar{\delta})$ , and choose  $\bar{\mu}' = (\epsilon \|\bar{z}''\|)^2 / (4 + 2\epsilon^2)$  such that  $\|u_1 \times \bar{z}''\|^2 + \|u_2 \times \bar{z}''\|^2 < (\epsilon \|\bar{z}''\|)^2 / (2 + \epsilon^2)$ . The case where  $\|u_1 \times u_2\| = 0$  is trivial. Let  $\|u_1 \times u_2\| \neq 0$  and  $\bar{z}'' = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_1 \times u_2$  with constants  $\alpha_i \in \mathbb{R}, i = 1, 2, 3$ . Then, one has  $\|u_1 \times \bar{z}''\|^2 + \|u_2 \times \bar{z}''\|^2 = (\alpha_1^2 + \alpha_2^2 + 2\alpha_3^2) \|u_1 \times u_2\|^2$ , where we made use of the fact  $\|u_i \times \bar{z}''\|^2 = \alpha_i^2 \|u_1 \times u_2\|^2 + \alpha_3^2 \|u_1 \times u_2\|^2$  for  $i = 1, 2$ . One can also show  $\|\bar{z}''\|^2 \leq 2\alpha_1^2 + 2\alpha_2^2 + \alpha_3^2 \|u_1 \times u_2\|^2 \leq (2 + \|u_1 \times u_2\|^2)(\alpha_1^2 + \alpha_2^2 + \alpha_3^2)$ . Hence, one obtains

$$\frac{\|u_1 \times u_2\|^2}{2 + \|u_1 \times u_2\|^2} \leq \frac{\|u_1 \times \bar{z}''\|^2 + \|u_2 \times \bar{z}''\|^2}{\|\bar{z}''\|^2} < \frac{\epsilon^2}{2 + \epsilon^2}.$$

Since the function  $f(x) = x^2 / (2 + x^2)$  is monotonically increasing (i.e.,  $\partial f / \partial x > 0$ ) for all  $x > 0$ , one obtains  $\|u(t_q) \times u(t_q + \bar{\delta})\| = \|u_1 \times u_2\| < \epsilon$ . From the definition of  $u(t)$  and the fact that  $\bar{\delta}$  can be arbitrary large, this contradicts item (i) of the Lemma that for any  $t^* \geq 0$ , there exists  $t > t^*$  such that  $\|u(t) \times u(t^*)\| > \epsilon$ .

- ii) *Motionless Camera:* In this case the matrix  $\bar{\mathcal{O}}$  in (64) is constant. For any  $u \in \mathbb{R}^3$ , one shows  $\bar{\Pi}_i u = u + (p_i - p') \bar{z}'_i$  with  $\|p_i - p'\| \neq 0$  and  $\bar{z}'_i = -(p_i - p')^\top u / \|p_i - p'\|^2 \in \mathbb{R}$  for all  $i = 1, 2, \dots, N$ . Note that identity  $p' = p_i$  implies that the camera is located at location the  $i$ -th landmark, which does not hold since all the landmarks are measurable by assumption. Let  $\bar{z}' = [\bar{z}'_1, \dots, \bar{z}'_N]^\top$  and  $\bar{z}'' := [\bar{z}^\top, (\bar{z}')^\top]^\top$ , which is non-zero since  $\bar{z}$  is non-zero. Hence, identity (64) can be rewritten as

$$\lim_{q \rightarrow \infty} \|\mathcal{O}' \bar{z}''\|^2 = 0. \quad (67)$$

with  $\mathcal{O}'$  defined in (22), which contradicts item (ii) of the Lemma that  $\mathcal{O}'$  has full rank of  $15 + N$ , i.e.,  $\|\mathcal{O}' \bar{z}''\|^2 > 0, \forall t \geq 0$ .

Therefore, the assumption (62) does not hold. Hence, for both cases, the pair  $(\bar{A}, C(t))$  is uniformly observable. It follows from (61) that the pair  $(A(t), C(t))$  is also uniformly observable.

It remains to prove Lemma 7. Since the velocity of the camera  $v' = \frac{d}{dt} p'(t)$  is bounded, the derivative of  $\mathcal{O}(t_p + \tau) \bar{z}$  is bounded for all  $q \in \mathbb{N}, \tau \in [0, \bar{\delta}]$ , and then it follows that  $\|\mathcal{O}(t_p + \tau) \bar{z}\|^2$  is uniformly continuous. Since every uniformly continuous function is also Cauchy-continuous, it implies that  $\|\mathcal{O}(t_p + \tau) \bar{z}\|^2$  is a Cauchy sequence of continuous functions and  $\lim_{q \rightarrow \infty} \|\mathcal{O}(t_p + \tau) \bar{z}\|^2$  exists. Applying Lebesgue theorem, one has  $\lim_{q \rightarrow \infty} \int_0^{\bar{\delta}} \|\mathcal{O}(t_q + \tau) \bar{z}\|^2 d\tau = \int_0^{\bar{\delta}} \lim_{q \rightarrow \infty} \|\mathcal{O}(t_q + \tau) \bar{z}\|^2 d\tau = 0$ . Again, making use of the fact that  $\lim_{q \rightarrow \infty} \|\mathcal{O}(t_q + \tau) \bar{z}\|^2$  is uniformly continuous and

non-negative, it follows that

$$\lim_{q \rightarrow \infty} \|\mathcal{O}(t_q + s)\bar{z}\|^2 = 0, \quad \forall s \in [0, \bar{\delta}]. \quad (68)$$

From the definition of  $\mathcal{O}$  and (68), it follows that  $\sum_{i=0}^2 \lim_{q \rightarrow \infty} \|\bar{\Theta}(t_q + s)\bar{C}\bar{A}^i\bar{z}\|^2 = 0, \forall s \in [0, \bar{\delta}]$ . Now, we are going to show that  $\bar{C}\bar{A}^2\bar{z} = \bar{C}\bar{A}\bar{z} = 0_{3N \times 1}$ . Let  $\bar{z} = [\bar{z}_1^\top, \bar{z}_2^\top, \dots, \bar{z}_5^\top]^\top \in \mathbb{R}^{15}$  with  $\bar{z}_i \in \mathbb{R}^3$  for all  $i = 1, \dots, 5$ . Making use of the fact  $\bar{\Theta}(t)\bar{C}\bar{A}^2\bar{z} = \bar{\Theta}(t)[(g_1\bar{z}_2 + g_2\bar{z}_3 + g_3\bar{z}_4)^\top, \dots, (g_1\bar{z}_2 + g_2\bar{z}_3 + g_3\bar{z}_4)^\top]^\top$ , one can show that, if  $g_1\bar{z}_2 + g_2\bar{z}_3 + g_3\bar{z}_4 \neq 0$ , the only solution of  $\lim_{q \rightarrow \infty} \|\bar{\Theta}(t_q + s)\bar{C}\bar{A}^2\bar{z}\| = 0$  is when the constant vector  $g_1\bar{z}_2 + g_2\bar{z}_3 + g_3\bar{z}_4$  is collinear with  $p_i - p'(t_q + s)$  for all  $i = 1, \dots, N$ . This does not hold since there are at least three non-aligned landmarks by assumption. Hence, identity  $\lim_{q \rightarrow \infty} \|\bar{\Theta}(t_q + s)\bar{C}\bar{A}^2\bar{z}\|^2 = 0$  implies  $\bar{C}\bar{A}^2\bar{z} = 0_{3N \times 1}$ . Moreover, using the fact  $\bar{\Theta}(t)\bar{C}\bar{A}\bar{z} = [(\bar{\Pi}_1\bar{z}_5)^\top, \dots, (\bar{\Pi}_N\bar{z}_5)^\top]^\top$ , one can show that, if  $\bar{z}_5 \neq 0_3$ , the only solution of  $\lim_{q \rightarrow \infty} \|\bar{\Theta}(t_q + s)\bar{C}\bar{A}\bar{z}\| = 0$  is when the constant vector  $\bar{z}_5$  is collinear with  $p_i - p'(t_q + s)$  for all  $i = 1, \dots, N$ . This does not hold since there are at least three non-aligned landmarks by assumption. These identities are not satisfied since there are at least three non-aligned landmarks by assumption. Hence,  $\lim_{q \rightarrow \infty} \|\bar{\Theta}(t_q + s)\bar{C}\bar{A}\bar{z}\|^2 = 0$  implies  $\bar{C}\bar{A}\bar{z} = 0_{3N \times 1}$ . From (59) and the facts  $\bar{C}\bar{A}^2\bar{z} = \bar{C}\bar{A}\bar{z} = 0_{3N \times 1}$ , one has  $\bar{C}\bar{z} \neq 0_{3N \times 1}$ . Then, from the definition of  $\bar{\mathcal{O}}$ , identity (68) can be reduced to (64). This completes the proof.

### E. Proof of Proposition 1

Consider the matrix  $\mathcal{O}'$  defined in (22) with  $N \geq 5$ . Since all the landmarks are not aligned, for the sake of simplicity, we assume that the first three landmarks are not aligned. Let  $u_i = [u_{i1}, u_{i2}, u_{i3}]^\top := p_1 - p_{i+1}$  for  $i = 1, 2$ . Hence, for each landmark  $p_i$ , there exist scalars  $\alpha_{ij}, j = 1, 2, 3$  such that  $p_i - p_1 = \sum_{j=1}^3 \alpha_{ij}u_j + \alpha_{i3}g$ . Note that if all the landmarks are located in the same plane, one has  $\alpha_{i3} = 0$  for all  $i = 4, \dots, N$ . Otherwise, one can always find three non-aligned landmarks, denoted by  $p_1, p_2$  and  $p_3$ , whose plane is not parallel to the gravity vector, i.e.,  $u_1, u_2$  and  $g$  are linearly independent. Applying similar column and row operations as in (60) on the matrix  $\mathcal{O}'$ , one obtains  $\mathcal{O}''$  in (69) (i.e.,  $\text{rank}(\mathcal{O}') = \text{rank}(\mathcal{O}'')$ ), which can be rewritten in the form of a block upper triangular matrix as

$$\mathcal{O}'' = \begin{bmatrix} \bar{\mathcal{O}}' & \Delta \\ 0_{(3N-9) \times 15} & M' \end{bmatrix},$$

where  $\bar{\mathcal{O}}' \in \mathbb{R}^{15 \times 15}$  is equivalent to the one defined in (60). It is obvious to show that matrix  $\mathcal{O}''$  does not have full rank if all the landmarks are located in the plane parallel to the gravity vector (i.e.,  $\text{rank}(\mathcal{O}') < 15$ ).

Now, we assume that all the landmarks are not located in the same plane, and there exist three non-aligned landmarks, denoted by  $p_1, p_2$  and  $p_3$ , whose plane is not parallel to the gravity. It follows that matrix  $\mathcal{O}'$  has full rank and is invertible. Then, one can show that any non-zero column of  $\Delta$  is linearly dependent on the columns of  $\bar{\mathcal{O}}'$ . Hence,  $\mathcal{O}''$  has full rank of  $15 + N$  if the matrix  $M'$  has rank of  $N$ . Since all the landmarks

are measurable by assumption, one has  $p' - p_i \neq 0$  for all  $i = 1, 2, \dots, N$  and the last  $N - 3$  columns of  $M'$  are linearly independent. Applying the column and row operations, one can show that matrix  $M'$  has the same rank as matrix  $M''$  defined in

$$M'' = \begin{bmatrix} \alpha'_{43}\check{p}_{14} & \alpha_{41}\check{p}_{24} & \alpha_{42}\check{p}_{34} & p_4 - p' & \dots & 0_{3 \times 1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha'_{N3}\check{p}_{1N} & \alpha_{N1}\check{p}_{2N} & \alpha_{N3}\check{p}_{3N} & 0_{3 \times 1} & \dots & p_N - p' \end{bmatrix} \quad (70)$$

where  $\check{p}_{ij} := p_i - p_j, \forall i, j = 1, \dots, N$ , and we made use of the facts  $\alpha_{i3}g - \alpha_{i1}\check{p}_{2i} - \alpha_{i2}\check{p}_{3i} = \sum_{j=1}^2 \alpha_{ij}u_j + \alpha_{i3}g - (\alpha_{i1} + \alpha_{i2})\check{p}_{1i} = \alpha'_{i3}\check{p}_{1i}$  with  $\alpha'_{i3} = -(1 + \alpha_{i1} + \alpha_{i2})$  for all  $i = 4, \dots, N$ .

- Consider the case where all the landmarks are located in the same plane. It implies that  $\alpha_{i3} = 0$  for all  $i = 4, \dots, N$ . From (70), one can show that the first three columns of  $M''$  are linearly dependent, i.e.,  $\alpha'_{i3}\check{p}_{1i} + \alpha_{i1}\check{p}_{2i} + \alpha_{i2}\check{p}_{3i} = \alpha_{i3}g = 0_{3 \times 1}$  for each  $i = 4, \dots, N$ . Hence, the rank of  $M''$  is less than  $N$ .
- Consider the case where one of the first three columns is zero.
  - If  $\alpha_{i1} = 0$  for all  $i = 4, \dots, N$ , one has  $\check{p}_{i1} = -\alpha_{i2}(p_3 - p_1) + \alpha_{i3}g$ . It follows that all the landmarks are located in the plane that contains  $p_1, p_3$  and is parallel to the gravity vector.
  - If  $\alpha_{i2} = 0$  for all  $i = 4, \dots, N$ , one has  $\check{p}_{i1} = -\alpha_{i1}(p_2 - p_1) + \alpha_{i3}g$ . It follows that all the landmarks are located in the plane that contains  $p_1, p_2$  and is parallel to the gravity vector.
  - If  $\alpha'_{i3} = 0$  (i.e.,  $\alpha_{i1} + \alpha_{i2} = -1$ ) for all  $i = 4, \dots, N$ , one has  $\check{p}_{i1} = -\alpha_{i1}(p_2 - p_1) - \alpha_{i2}(p_3 - p_1) + \alpha_{i3}g$ , which can be rewritten as  $\check{p}_{i3} = -\alpha_{i1}(p_2 - p_3) + \alpha_{i3}g$ . It follows that all the landmarks are located in the plane that contains  $p_2, p_3$  and is parallel to the gravity vector.

Hence, the rank of  $M'$  is less than  $N$  if landmarks  $p_4, \dots, p_N$  are located in the same plane that contains two of landmarks  $p_1, p_2, p_3$  and is parallel to the gravity vector.

- Consider the case where landmarks  $p_4, \dots, p_N$  are aligned with one of the landmarks  $p_1, p_2, p_3$  and the camera position  $p'$ . Without loss of generality, let  $p_1, p_4, \dots, p_N$  and  $p'$  be aligned. Then, one can show that the first column of the matrix  $M''$  is linearly dependent on its last  $N - 3$  columns, which implies that the rank of  $M'$  is less than  $N$ .
- Consider the case where there exists an index  $4 < \ell < N$  such that landmarks  $p_4, \dots, p_\ell$  are located in the same plane that contains two of the landmarks  $p_1, p_2, p_3$  (for example,  $p_1$  and  $p_2$ ) and is parallel to the gravity vector, and landmarks  $p_{\ell+1}, \dots, p_N$  are aligned with the third landmark (i.e.,  $p_3$ ) and the camera position  $p'$ . It follows that  $\alpha_{i2} = 0$  for all  $i = 4, \dots, \ell$ ,  $p_3 - p_i$  and  $p_i - p'$  are collinear for all  $i = \ell + 1, \dots, N$ , which implies that the rank of  $M'$  is less than  $N$ .

If none of these cases hold, one can conclude that matrix  $\mathcal{O}'$

$$\mathcal{O}'' = \begin{bmatrix} I_3 & -p_{11}I_3 & -p_{12}I_3 & -p_{13}I_3 & 0_3 & p_1 - p' & 0_{3 \times 1} & 0_{3 \times 1} & 0_{3 \times 1} & \dots & 0_{3 \times 1} \\ 0_3 & u_{11}I_3 & u_{12}I_3 & u_{13}I_3 & 0_3 & -u_1 & p_2 - p' & 0_{3 \times 1} & 0_{3 \times 1} & \dots & 0_{3 \times 1} \\ 0_3 & u_{21}I_3 & u_{22}I_3 & u_{23}I_3 & 0_3 & -u_2 & 0_{3 \times 1} & p_3 - p' & 0_{3 \times 1} & \dots & 0_{3 \times 1} \\ 0_3 & 0_3 & 0_3 & 0_3 & I_3 & 0_{3 \times 1} & 0_{3 \times 1} & 0_{3 \times 1} & 0_{3 \times 1} & \dots & 0_{3 \times 1} \\ 0_3 & g_1I_3 & g_2I_3 & g_3I_3 & 0_3 & 0_{3 \times 1} & 0_{3 \times 1} & 0_{3 \times 1} & 0_{3 \times 1} & \dots & 0_{3 \times 1} \\ \hline 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \alpha_{43}g & \alpha_{41}(p_2 - p') & \alpha_{42}(p_3 - p') & p_4 - p' & \dots & 0_{3 \times 1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_3 & 0_3 & 0_3 & 0_3 & 0_3 & \alpha_{N3}g & \alpha_{N1}(p_2 - p') & \alpha_{N2}(p_3 - p') & 0_{3 \times 1} & \dots & p_N - p' \end{bmatrix} \quad (69)$$

in (22) has full rank. This completes the proof.

## REFERENCES

- [1] M. Wang and A. Tayebi, "Nonlinear observers for stereo-vision-aided inertial navigation," in *Proc. 58th IEEE Conference on Decision and Control*, 2019, pp. 2516–2521.
- [2] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint Kalman filter for vision-aided inertial navigation," in *Proc. IEEE International Conference on Robotics and Automation*. IEEE, 2007, pp. 3565–3572.
- [3] M. Li and A. I. Mourikis, "High-precision, consistent EKF-based visual-inertial odometry," *The International Journal of Robotics Research*, vol. 32, no. 6, pp. 690–711, 2013.
- [4] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM: a versatile and accurate monocular slam system," *IEEE transactions on robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.
- [5] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.
- [6] R. Mahony, T. Hamel, and J.-M. Pfimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Transactions on automatic control*, vol. 53, no. 5, pp. 1203–1218, 2008.
- [7] S. Bonnabel, P. Martin, and P. Rouchon, "Non-linear symmetry-preserving observers on lie groups," *IEEE Transactions on Automatic Control*, vol. 54, no. 7, pp. 1709–1713, 2009.
- [8] M.-D. Hua, G. Ducard, T. Hamel, R. Mahony, and K. Rudin, "Implementation of a nonlinear attitude estimator for aerial robotic vehicles," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 1, pp. 201–213, 2013.
- [9] M.-D. Hua, "Attitude estimation for accelerated vehicles using GPS/INS measurements," *Control Engineering Practice*, vol. 18, no. 7, pp. 723–732, 2010.
- [10] A. Roberts and A. Tayebi, "On the attitude estimation of accelerating rigid-bodies using GPS and IMU measurements," in *Proc. 50th IEEE conference on decision and control and European Control conference*, 2011, pp. 8088–8093.
- [11] S. Berkane and A. Tayebi, "Attitude and gyro bias estimation using GPS and IMU measurements," in *Proc. 56th IEEE Conference on Decision and Control*, 2017, pp. 2402–2407.
- [12] T. H. Bryne, J. M. Hansen, R. H. Rogne, N. Sokolova, T. I. Fossen, and T. A. Johansen, "Nonlinear observers for integrated ins/gnss navigation: implementation aspects," *IEEE Control Systems Magazine*, vol. 37, no. 3, pp. 59–86, 2017.
- [13] 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.
- [14] M.-D. Hua and G. Allibert, "Riccati observer design for pose, linear velocity and gravity direction estimation using landmark position and IMU measurements," in *Proc. IEEE Conference on Control Technology and Applications*, 2018, pp. 1313–1318.
- [15] 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.
- [16] 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.
- [17] S. Berkane, A. Abdessameud, and A. Tayebi, "Hybrid attitude and gyro-bias observer design on SO(3)," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 6044–6050, 2017.
- [18] M. Wang and A. Tayebi, "Hybrid pose and velocity-bias estimation on SE(3) using inertial and landmark measurements," *IEEE Transactions on Automatic Control*, vol. 64, no. 8, pp. 3399–3406, 2019.
- [19] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge University Press, 2003.
- [20] S. De Marco, M.-D. Hua, T. Hamel, and C. Samson, "Position, velocity, attitude and accelerometer-bias estimation from IMU and bearing measurements," in *Proc. European Control Conference*. IEEE, 2020, pp. 1003–1008.
- [21] F. Ferrante, F. Gouaisbaut, R. G. Sanfelice, and S. Tarbouriech, "State estimation of linear systems in the presence of sporadic measurements," *Automatica*, vol. 73, pp. 101–109, 2016.
- [22] S. Berkane and A. Tayebi, "Attitude estimation with intermittent measurements," *Automatica*, vol. 105, pp. 415–421, 2019.
- [23] M. Wang and A. Tayebi, "Nonlinear state estimation for inertial navigation systems with intermittent measurements," *Preprint submitted to Automatica*, vol. 122, p. 109244, 2020.
- [24] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, "The EuRoC micro aerial vehicle datasets," *The International Journal of Robotics Research*, vol. 35, no. 10, pp. 1157–1163, 2016.
- [25] G. Baldwin, R. Mahony, and J. Trumpf, "A nonlinear observer for 6 dof pose estimation from inertial and bearing measurements," in *Proc. IEEE International Conference on Robotics and Automation*. IEEE, 2009, pp. 2237–2242.
- [26] P. Batista, C. Silvestre, and P. Oliveira, "Navigation systems based on multiple bearing measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 4, pp. 2887–2899, 2015.
- [27] T. Hamel and C. Samson, "Position estimation from direction or range measurements," *Automatica*, vol. 82, pp. 137–144, 2017.
- [28] S. Berkane, A. Tayebi, and S. de Marco, "A nonlinear navigation observer using IMU and generic position information," *Preprint submitted to Automatica*, 2020, [arXiv:2006.14056v1](https://arxiv.org/abs/2006.14056v1).
- [29] R. S. Bucy, "Global theory of the riccati equation," *Journal of computer and system sciences*, vol. 1, no. 4, pp. 349–361, 1967.
- [30] —, "The riccati equation and its bounds," *Journal of computer and system sciences*, vol. 6, no. 4, pp. 343–353, 1972.
- [31] D. E. Koditschek, "Application of a new lyapunov function to global adaptive attitude tracking," in *Proc. 27th IEEE Conference on Decision and Control*. IEEE, 1988, pp. 63–68.
- [32] L. Perko, *Differential equations and dynamical systems*, 3rd ed. Texts in Applied Mathematics 7, Springer-Verlag, New York. Inc. 2001, 2013, vol. 7.
- [33] G. G. Scandaroli, "Visuo-inertial data fusion for pose estimation and self-calibration," Ph.D. dissertation, Université Nice Sophia Antipolis, 2013.
- [34] V. Lepetit, F. Moreno-Noguer, and P. Fua, "Epnnp: An accurate O(n) solution to the PnP problem," *International journal of computer vision*, vol. 81, no. 2, p. 155, 2009.
- [35] R. Goebel, R. G. Sanfelice, and A. R. Teel, "Hybrid dynamical systems," *IEEE control systems magazine*, vol. 29, no. 2, pp. 28–93, 2009.
- [36] —, *Hybrid Dynamical Systems: modeling, stability, and robustness*. Princeton University Press, 2012.
- [37] J. Deyst and C. Price, "Conditions for asymptotic stability of the discrete minimum-variance linear estimator," *IEEE Transactions on Automatic Control*, vol. 13, pp. 702–705, 1968.
- [38] J. Shi and C. Tomasi, "Good features to track," in *Proc. IEEE conference on Computer Vision and Pattern Recognition*. IEEE, 1994, pp. 593–600.
- [39] S. Berkane, "Hybrid attitude control and estimation on SO(3)," Ph.D. dissertation, The University of Western Ontario, 2017.