

# Cooperative Visual-Inertial Odometry: Analysis of Singularities, Degeneracies and Minimal Cases

Agostino Martinelli

► **To cite this version:**

Agostino Martinelli. Cooperative Visual-Inertial Odometry: Analysis of Singularities, Degeneracies and Minimal Cases. IEEE Robotics and Automation Letters, IEEE 2020, 5 (2), pp.668 - 675. 10.1109/LRA.2020.2965063 . hal-02427991

**HAL Id: hal-02427991**

**<https://hal.inria.fr/hal-02427991>**

Submitted on 14 Jan 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Cooperative Visual-Inertial Odometry: Analysis of Singularities, Degeneracies and Minimal Cases

Agostino Martinelli

## Abstract

This paper provides an exhaustive analysis of all the singularities and minimal cases in cooperative visual-inertial odometry. Specifically, the case of two agents is analysed. As in the case of a single agent (addressed in [1]) and in the case of other computer vision problems, the key of the analysis is the establishment of an equivalence between the cooperative visual-inertial odometry problem and a Polynomial Equation System (PES). In the case of a single agent, the PES consists of linear equations and a single polynomial of second degree. In the case of two agents, the number of second degree equations becomes three and, also in this case, a complete analytic solution can be obtained [2]. The power of the analytic solution is twofold. From one side, it allows us to determine the state without the need of an initialization. From another side, it provides fundamental insights into all the structural properties of the problem. This paper focuses on this latter issue. Specifically, we obtain all the minimal cases and singularities depending on the number of camera images and the relative trajectory between the agents. The problem, when non singular, can have up to eight distinct solutions. The usefulness of this analysis is illustrated with simulations. In particular, we show quantitatively how the performance of the state estimation worsens near a singularity.

## I. INTRODUCTION

Vision and inertial sensing have received great attention by the mobile robotics community. These sensors require no external infrastructure and this is a key advantage for robots operating in unknown environments where GPS signals are shadowed. Additionally, these sensors have very interesting complementarities and together provide rich information to build a system capable of vision-aided inertial navigation and mapping. Recently, this sensor fusion problem has been successfully addressed by using optimization-based approaches [3], [4], [5], [6], [7], [8], [9], [10]. These optimization methods outperform filter-based algorithms in terms of accuracy due to their capability of re-linearizing past states. On the other hand, the optimization process can be affected by the presence of local minima. Deterministic solutions able to automatically determine the state without initialization have also been introduced [1], [11], and they can overcome this obstacle.

In the case of only vision, the problem of determining the three-dimensional structure of the scene from a two-dimensional image sequence (the structure from motion problem), has been investigated since more than half century and analytic results have been obtained [12], [13], [14]. In particular, these results translate the problem into polynomial equations and, the analysis of these polynomial equations, allows us to detect all the structural properties of the problem (singularities, degeneracies, etc). In the case of the problem of visual and inertial sensor fusion, such type of analysis has only been carried out very recently. Specifically, [1] established that this sensor fusion problem is equivalent to a very simple Polynomial Equation System (PES). In particular, this PES consists of a single polynomial equation of second degree and several linear equations. This PES can be easily solved in closed-form and, this solution, has the advantage to provide the state without initialization. Even more importantly, this PES contains all the structural properties of the problem. In particular, by studying this PES, the author of [1] carried out a detailed analysis of the problem by providing all the system singularities and minimal cases depending on the trajectory, on the number of camera images and on the features layout. The problem can have up to two solutions in its minimal cases.

Visual and inertial sensors have also been used in a cooperative scenario (e.g., for cooperative mapping in [15]). In [16] we investigated its observability properties in order to obtain the observable state in several conditions. In [2], we established that, the cooperative visual-inertial odometry in the case of two agents (from now on *COVIO*), is equivalent to a PES that consists of several linear equations and three polynomial equations of second degree. This PES was solved by using the method based on the Macaulay resultant matrices [17]. This is the analytic solution of *COVIO*. In [2] we exploited the power of this solution to obtain an automatic method for state initialization. In particular, we analyzed its performance in terms of precision and robustness both by using synthetic and real data. On the other hand, the equivalence between *COVIO* and the aforementioned PES allows us to obtain all the structural properties of *COVIO* (e.g., how the minimal cases, singularities and degeneracies depend on the trajectories and on the number of camera images). This is fundamental in many applications and, so far, it has been discussed in the only-vision case [18].

The goal of this paper is precisely to perform this analysis of singularities and minimal cases<sup>1</sup>. The PES that is equivalent to *COVIO*, together with its analytic solution is summarized in section II (the reader is addressed to [2] for its analytic derivation). The PES has the following feature. The unknowns are the components of the relative state (relative position,

The author is with Univ. Grenoble Alpes, INRIA, 38000 Grenoble, France [agostino.martinelli@inria.fr](mailto:agostino.martinelli@inria.fr)

<sup>1</sup>Note that, by using the PES in [2], we will obtain the singularities of the most extreme scenario (i.e., when no external point feature is available). These singularities include all the singularities of any less critical scenario.

velocity and orientation between the two agents) while the coefficients of the system and the vector with the constant terms only depend on the measurements from the two cameras and the inertial sensors delivered during a short time interval. Starting from the analytic solution of this PES it is possible to obtain the number of solutions in terms of the trajectories and the number of camera images. This analysis is carried out in section III. The importance and usefulness of this analysis is illustrated by showing how the performance of the state estimation deteriorates near a singularity. In particular, this study is carried out in section IV via simulations. Finally, section V provides our conclusion.

## II. THE ANALYTIC SOLUTION OF *COVIO*

This section provides a description of the PES that fully characterizes *COVIO* together with its analytic solution. A complete derivation is available in [2]. To maintain consistency with the notation adopted in [2], we will denote the two agents by  $\mathcal{B}_1$  and  $\mathcal{B}_2$ . The PES consists of a linear system that we called the *linear system associated to COVIO* and was denoted by  $\Sigma_{Lin}$ . The unknowns of this system are not independent. To account for this,  $\Sigma_{Lin}$  must be completed with three polynomial equations of second degree.

### A. Linear system ( $\Sigma_{Lin}$ )

We distinguish the case when only  $\mathcal{B}_1$  is equipped with a camera, from the case when both  $\mathcal{B}_1$  and  $\mathcal{B}_2$  are equipped with a camera. In this latter case we assume that the two cameras are synchronized (the case when the cameras are not synchronized provides two distinct problems that can be solved separately). We assume that the camera performs  $n$  observations during the time interval  $(t_A, t_B)$  at the times  $t_1 = t_A, t_2, \dots, t_{n-1}, t_n = t_B$ . The linear system is:

$$\lambda_j \mu_j = \xi_A + \eta_A(t_j - t_A) + R_A \beta_j^2 - \beta_j^1 \quad (1)$$

where:

- $\lambda_j$  ( $j = 1, \dots, n$ ) is the distance between the two bodies<sup>2</sup> at the time  $t_j$ .
- $\mu_j$  ( $j = 1, \dots, n$ ) is the unit vector with the same direction of the vector that connects the two bodies at the time  $t_j$ , expressed in the local frame of  $\mathcal{B}_1$  at the time  $t_1 = t_A$ .
- $\xi_A$  is the position of  $\mathcal{B}_2$  in the local frame of  $\mathcal{B}_1$  at the time  $t_1 = t_A$ .
- $\eta_A$  is the velocity of  $\mathcal{B}_2$  in the local frame of  $\mathcal{B}_1$  at the time  $t_1 = t_A$ .
- $R_A$  is the  $3 \times 3$  orthogonal matrix that characterizes the relative rotation between  $\mathcal{B}_1$  and  $\mathcal{B}_2$ , at the time  $t_1 = t_A$ .
- The two vectors  $\beta_j^1$  and  $\beta_j^2$  are defined by equation (5) in [2] for  $t = t_j$ . They are the doubled integrated acceleration, from  $t_1$  up to  $t_j$ , of the two bodies ( $\mathcal{B}_1$  and  $\mathcal{B}_2$ ), computed in their local frames at time  $t_1 = t_A$ .

The system in (1) is a linear equations' system in  $15 + n$  unknowns, which are:

- The distances  $\lambda_1, \dots, \lambda_n$ .
- The three components of  $\xi_A$ .
- The three components of  $\eta_A$ .
- The nine entries of the matrix  $R_A$ .

The quantities  $\mu_j, \beta_j^1, \beta_j^2$  ( $j = 1, \dots, n$ ), are directly provided by the IMU and the camera measurements delivered in the time interval  $(t_A, t_j)$ . This is a fundamental aspect, since makes the solution initialization-free (the experiments carried out in [2] showed that, as in the case of a single agent, the time interval never exceeds 4 seconds).

Note that equation (1) is a vector equation, providing 3 scalar equations. Since this holds for each  $j = 1, \dots, n$ , we obtain a linear system of  $3n$  equations in  $15 + n$  unknowns. On the other hand, we also know that the matrix  $R_A$  is orthonormal and consequently it is fully characterized by three parameters (instead of nine). This will be accounted later (section II-B).

Let us consider now the case when also  $\mathcal{B}_2$  is equipped with a camera and, the measurements made by this camera, occur at the same times  $t_1, \dots, t_n$  (synchronized cameras). In this case,  $\Sigma_{Lin}$  also includes the following  $3n$  equations:

$$R_A \nu_j = -\mu_j \quad j = 1, \dots, n \quad (2)$$

where  $\nu_j$  ( $j = 1, \dots, n$ ) is the unit vector with the same direction of the vector that connects the two bodies at the time  $t_j$ , expressed in the local frame of  $\mathcal{B}_2$  at the time  $t_1 = t_A$ . Note that, the unknowns in this second set of equations, are only the entries of  $R_A$ .

### B. Polynomial equation system ( $\mathcal{P}$ )

So far, we have obtained a system of linear equations in  $15 + n$  unknowns. In the case of a single camera, the equations are the  $3n$  scalar equations given in (1). In the case of two synchronized cameras, we also have the  $3n$  equations given in (2). On the other hand, the unknowns are not independent since the matrix  $R_A$  is orthonormal. To account for this, in [2] we proceeded with the following three steps: (i) Elimination of  $\xi_A, \eta_A$  and  $\lambda_1, \dots, \lambda_n$  from  $\Sigma_{Lin}$  by using part of its equations; (ii) quaternion parametrization of  $R_A$ ; (iii) reduction to a quadratic system in three unknowns. The reader is addressed to [2] for all the details about the above steps.

<sup>2</sup>The distance is between the origins of the two local frames.

### III. ANALYSIS OF SINGULARITIES AND MINIMAL CASES

The goal of the analysis provided in this section is to find out how the number of the solutions of  $\mathcal{COVIO}$  depends on the relative motion and on the number of camera images. For this analysis, we do not need to optimally solve  $\mathcal{P}$ . In particular, this analysis assumes that all the measurements are noiseless. Also note that, for our analysis, there are many cases for which it is unnecessary to use the complete  $\mathcal{P}$ . Indeed, in many cases, the structure of  $\Sigma_{Lin}$  allows us to separate  $R_A$  from the rest (see the properties 2 and 3). As we will see, for these cases we only need to use  $\Sigma_{Lin}$ . The complete  $\mathcal{P}$  will be only used in few cases.

We start by eliminating the unknown  $\xi_A$  by using the first equation of  $\Sigma_{Lin}$  (i.e., the one for  $j = 1$  in (1)). Since  $t_1 = t_A$  and both  $\beta_1^1$  and  $\beta_1^2$  vanish, we have:  $\xi_A = \lambda_1 \mu_1$ . Substituting in the remaining equations, we obtain:

$$\eta_A \Delta_j + R_A \beta_j^2 + \lambda_1 \mu_1 - \lambda_j \mu_j = \beta_j^1, \quad j = 2, \dots, n \quad (3)$$

with  $\Delta_j \triangleq t_j - t_A$ . We denote this linear system by  $\tilde{\Sigma}_{Lin}$ . It is characterized by the vector of unknowns:  $[\eta_A \ E^1 \ E^2 \ E^3 \ \lambda_1 \ \dots \ \lambda_n]$  (where  $E^1, E^2, E^3$  are the three columns of  $R_A$ ) and the matrix:

$$\begin{bmatrix} \Delta_2 I_3 & \Gamma_2 & \mu_1 & -\mu_2 & 0_3 & \dots & \dots & \dots & 0_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \Delta_j I_3 & \Gamma_j & \mu_1 & \dots & 0_3 & -\mu_j & 0_3 & \dots & 0_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \Delta_n I_3 & \Gamma_n & \mu_1 & \dots & \dots & \dots & \dots & 0_3 & -\mu_n \end{bmatrix} \quad (4)$$

where  $I_3$  is the  $3 \times 3$  identity matrix,  $0_3$  is the zero  $3 \times 1$  vector and  $\Gamma_j$  is the  $3 \times 9$  matrix block

$$\Gamma_j \triangleq [ \beta_{xj}^2 I_3 \mid \beta_{yj}^2 I_3 \mid \beta_{zj}^2 I_3 ]$$

and the indices  $x, y, z$  on  $\beta_j^2$  denote its three components.

We provide a first characterization of the relative motion by introducing the concept of collinearity and coplanarity. We denote by  $\mathcal{V}$  the vector space generated by the vectors  $\mu_1, \mu_2, \dots, \mu_n$ . We introduce the following definition:

**Definition 1 (Collinearity and coplanarity)** *The relative motion is collinear when the dimension of  $\mathcal{V}$  is 1. It is coplanar when the dimension of  $\mathcal{V}$  is  $\leq 2$ .*

We also define the concept of singularity as follows:

**Definition 2 (Singularity)** *The  $\mathcal{COVIO}$  is fully singular when neither any component of the relative orientation nor the relative position and velocity (even up to a scale) can be determined, even with multiple (but finite) solutions. It is singular, when at least one of the above quantities cannot be determined, even with multiple (but finite) solutions.*

In other words, the  $\mathcal{COVIO}$  is singular when we have infinite solutions at least for one unknown among the three components of the relative position, speed and orientation. It is fully singular when we have simultaneously infinite solutions for all the above unknowns and even for the relative position and the relative speed up to a scale.

We denote by  $a_r$  the relative inertial acceleration between the two agents (in the global frame). For the sake of brevity, with  $a_r = 0_3$  we mean that the relative acceleration is null during the considered time interval. Finally, we define  $\hat{\beta}_j^2 \triangleq \frac{\beta_j^2}{|\beta_j^2|}$ , i.e.,  $\hat{\beta}_j^2$  is the unit vector along  $\beta_j^2$ .

We have the following fundamental property:

**Property 1 (Scale invariance)** *If  $a_r = 0_3$ , the  $\mathcal{COVIO}$  is scale invariant (i.e., the absolute scale is unobservable).*

**Proof :** When  $a_r = 0_3$  we have:

$$\xi_j = \xi_1 + \Delta_j \eta_A$$

Hence, the following vector belongs to the null space of the matrix in (4):

$$[\eta_A, 0_3, 0_3, 0_3, \lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n]^T \quad (5)$$

Additionally, since (2) only depends on  $R_A$ , the above vector belongs to the null space of  $\tilde{\Sigma}_{Lin}$  also in the case of two cameras. The killing vector in (5) expresses the system invariance with respect to the absolute scale ■

When this property holds, in order to further investigate which conditions provide the observability of the remaining unknowns, we fix arbitrarily one of the  $\lambda$  value and we remove it from the unknowns.

For the rest of our analysis, we need to distinguish the case of single camera from the case of two cameras.

### A. Single camera

In this case  $\tilde{\Sigma}_{Lin}$  only consists of equations (3). We have the following important property:

**Property 2 (Separation in the case of a single camera)** *When  $|\beta_j^2| = 0 \forall j$ , no component of  $R_A$  is observable.*

**Proof :** This is a trivial consequence of the fact that, when  $|\beta_j^2| = 0$ , equation (3) is independent of  $R_A$  ■  
When this property holds, in order to further investigate which conditions provide the observability of the remaining unknowns, we simply remove from the state all the components of  $R_A$ . In other words,  $\tilde{\Sigma}_{Lin}$  becomes:

$$\eta_A \Delta_j + \lambda_1 \mu_1 - \lambda_j \mu_j = \beta_j^1, \quad j = 2, \dots, n \quad (6)$$

Note that, in this case,  $\mathcal{P}$  reduces to  $\tilde{\Sigma}_{Lin}$  since  $\tilde{\Sigma}_{Lin}$  is independent of  $R_A$ .

Because of the property above, the vector  $\beta_j^2$  (together with the vector  $a_r$ ) plays a fundamental role in the analysis of the singularities of COVIO. Note that  $\beta_j^2$  is independent of  $\mathcal{B}_2$ 's rotations because, by definition, it is computed in a fixed frame. It is the doubled integrated acceleration from  $t_1$  up to  $t_j$  perceived by  $\mathcal{B}_2$  and rotated of the rotation done in  $(t_A, t_j)$ .

In the rest of this section we provide a set of results that provide an exhaustive analysis by analyzing all the possible scenarios, case by case. The analytic proofs of these results will be given separately, in appendix A .

**Case 1 ( $\beta_j^2 = a_r = 0_3$ )** *Let us suppose that  $|\beta_j^2| = 0$  for any  $j = 1, \dots, n$  and that  $a_r = 0_3$ . No component of  $R_A$  is observable. In addition:*

- *If  $n \leq 2$  and/or the relative motion is collinear the COVIO is fully singular.*
- *If  $n \geq 3$  and the relative motion is not collinear, then  $\xi_1, \xi_2, \xi_3$  and  $\eta_A$  can be uniquely determined up to a scale.*

In presence of gravity, this case occurs when both  $\mathcal{B}_1$  and  $\mathcal{B}_2$  are in free fall. In absence of gravity, this case occurs when both  $\mathcal{B}_1$  and  $\mathcal{B}_2$  accomplish uniform rectilinear motions.

**Case 2 ( $\beta_j^2 = 0_3$  and  $a_r \neq 0_3$ )** *Let us suppose that  $|\beta_j^2| = 0$  for any  $j = 1, \dots, n$ . No component of  $R_A$  is observable. In addition:*

- *If  $n \leq 2$  and/or the relative motion is collinear the COVIO is fully singular.*
- *If  $n = 3$  and the relative motion is coplanar the COVIO is fully singular.*
- *If  $n = 3$  and the relative motion is not coplanar, then  $\xi_1, \xi_2, \xi_3$  and  $\eta_A$  can be uniquely determined (**minimal case**).*
- *If  $n \geq 4$  and the relative motion is not collinear, then  $\xi_1, \dots, \xi_n$  and  $\eta_A$  can be uniquely determined.*

In presence of gravity, this case occurs when  $\mathcal{B}_2$  is in free fall and  $\mathcal{B}_1$  is not. In absence of gravity, this case occurs when  $\mathcal{B}_2$  accomplishes a uniform rectilinear motion and  $\mathcal{B}_1$  does not.

We now consider the cases when  $\beta_j^2$  does not vanish but  $\hat{\beta}_j^2$  is independent of  $j$ , i.e., when the vector  $\beta_j^2$  is along a fix axis. We denote by  $\hat{\beta}^2$  this axis and by  $\hat{\tau} \triangleq R_A \hat{\beta}^2$ . For these cases, we denote by  $\alpha_r, \alpha_p$  and  $\alpha_y$ , the roll, pitch and yaw angles that characterize the relative orientation between  $\mathcal{B}_1$  and  $\mathcal{B}_2$ , and computed in a frame with the vertical axis aligned with  $\hat{\tau}$ . In presence of gravity, this condition occurs in two situations. The first situation is realized when  $\mathcal{B}_2$  accelerates along the same direction of gravity, with the exception of the free fall. The second situation is realized when the inertial acceleration is constant and its direction does not coincide with the gravity.

**Case 3 ( $\beta_j^2$  along a fix axis and  $a_r = 0_3$ )** *Let us suppose that  $\hat{\beta}_j^2 = \hat{\beta}^2 \forall j$  and that the special condition  $|\beta_j^2| \propto \Delta_j$  does not hold. Let us suppose also that  $a_r = 0_3$ . We have:*

- *If  $n \leq 3$  and/or the relative motion is collinear the COVIO is fully singular.*
- *If  $n = 4$  and the relative motion is not coplanar, then  $\xi_1, \dots, \xi_n$  and  $\eta_A$  up to scale, together with  $\alpha_r$  and  $\alpha_p$ , can be uniquely determined.*
- *If  $n = 4$  and the relative motion is coplanar (but not collinear), the same above quantities can be determined but we have two distinct solutions.*
- *If  $n \geq 5$  and the relative motion is not collinear, then the same above quantities can be uniquely determined.*

If the special condition  $|\beta_j^2| \propto \Delta_j$  holds, no component of  $R_A$  is observable. Additionally, it can be proved that also  $\eta_A$  cannot be determined and, for the remaining unknowns ( $\xi_1, \dots, \xi_n$ ), the same results described in case 1 hold.

**Case 4 ( $\beta_j^2$  along a fix axis and  $a_r \neq 0_3$ )** *Let us suppose that  $\hat{\beta}_j^2 = \hat{\beta}^2 \forall j$  and that the special condition  $|\beta_j^2| \propto \Delta_j$  does not hold. We have:*

- *If  $n \leq 3$  and/or the relative motion is collinear the COVIO is fully singular.*
- *If  $n = 4$  and the relative motion is coplanar the COVIO is fully singular.*
- *If  $n = 4$  and the relative motion is not coplanar, then  $\xi_1, \xi_2, \xi_3, \eta_A$  together with  $\alpha_r$  and  $\alpha_p$ , can be determined. In particular, there are two distinct solutions. This is a **minimal case***

- If  $n = 5$  and the relative motion is coplanar (but not collinear), then the same above quantities can be determined. In particular, there are two distinct solutions.
- If  $n = 5$  and the relative motion is not coplanar, then the same above quantities can be uniquely determined.
- If  $n \geq 6$  and the relative motion is not collinear, then the same above quantities can be uniquely determined.

If the special condition  $|\beta_j^2| \propto \Delta_j$  holds, no component of  $R_A$  is observable. Additionally, it can be proved that also  $\eta_A$  cannot be determined and, for the remaining unknowns ( $\xi_1, \dots, \xi_n$ ), the same results described in case 2 hold.

The last two cases contemplate any motion of  $\mathcal{B}_2$  that does not satisfy the special conditions considered so far.

**Case 5 ( $\beta_j^2$  along at least two axes and  $a_r = 0_3$ )** Let us suppose that the span of  $\hat{\beta}_1^2, \dots, \hat{\beta}_n^2$  has dimension at least equal to 2 and that  $a_r = 0_3$ . We have:

- If  $n \leq 3$  and/or the relative motion is collinear the COVIO is fully singular.
- If  $n = 4$  and the relative motion is not collinear, then  $\xi_1, \xi_2, \xi_3$  and  $\eta_A$  up to a scale together with the three Euler angles that characterize the relative orientation, can be determined. In particular, this is a **minimal case** with up to eight distinct solutions.
- If  $n \geq 5$  and the relative motion is not collinear, then the same above quantities can be uniquely determined.

**Case 6 ( $\beta_j^2$  along at least two axes and  $a_r \neq 0_3$ )** Let us suppose that the span of  $\hat{\beta}_1^2, \dots, \hat{\beta}_n^2$  has dimension at least equal to 2 and that  $a_r \neq 0_3$ . We have:

- If  $n \leq 4$  and/or the relative motion is collinear the COVIO is fully singular.
- If  $n \geq 5$  and the relative motion is not collinear, then  $\xi_1, \xi_2, \xi_3, \eta_A$  together with the three Euler angles that characterize the relative orientation can be uniquely determined.

## B. Two cameras

In this case  $\tilde{\Sigma}_{Lin}$  consists of equations (2) and (3). The analysis is much easier thanks to the following property:

**Property 3 (Separation in the case of 2 cameras)** When the relative motion is not collinear,  $R_A$  is fully observable and can be separated from the remaining unknowns.

**Proof:** : This is a trivial consequence of the fact that, when the dimension of  $\mathcal{V}$  is at least 2, equation (2) allows us to uniquely determine  $R_A$  ■

All the results can be provided by distinguishing two cases (the proofs of these results will be given in appendix B).

**Case 7 (Non-collinear motion)** Let us suppose that the relative motion is non-collinear.  $R_A$  can be uniquely determined. In addition:

- If  $n \leq 2$  the remaining unknowns cannot be determined even up to a scale.
- If  $n = 3$ ,  $a_r \neq 0_3$  and the relative motion is coplanar the remaining unknowns cannot be determined even up to a scale.
- If  $n = 3$ ,  $a_r \neq 0_3$  and the relative motion is not coplanar, then  $\xi_1, \xi_2, \xi_3$  and  $\eta_A$  can be uniquely determined (**minimal case**).
- If  $n \geq 4$  and  $a_r \neq 0_3$ , then  $\xi_1, \dots, \xi_n$  and  $\eta_A$  can be uniquely determined.
- If  $n \geq 3$  and  $a_r = 0_3$ , then  $\xi_1, \xi_2, \xi_3$  and  $\eta_A$  can be uniquely determined up to a scale.

Let us consider the case of a collinear relative motion. By definition,  $\mu_1 = \dots = \mu_n \triangleq \mu$  and, consequently, also  $\nu_1 = \dots = \nu_n \triangleq \nu$ . Additionally, the condition  $\hat{\beta}_j^2 = \nu, \forall j = 1, \dots, n$  is equivalent to  $\hat{\beta}_j^1 = \mu$ . This condition cannot be satisfied by maintaining the relative motion collinear (e.g., consider the case when  $R_A \beta_j^2 - \beta_j^1 = 0_3$  in equation (1)). We have the following result:

**Case 8 (Collinear motion)** Let us suppose that the relative motion is collinear. We have:

- If  $n \leq 2$  and/or  $\hat{\beta}_j^2 = \nu, \forall j = 1, \dots, n$ , then  $R_A$  can be uniquely determined up to a rotation around  $\mu$ .
- If  $n \geq 3$  and  $\hat{\beta}_j^2 \neq \nu$  for at least one  $j$ , then  $R_A$  can be uniquely determined.

Additionally, none of the remaining unknowns can be determined.

All the above results are summarized in tables I and II.

$\beta^2$	$a_r$	$n$	What can be determined
$\beta_j^2 = 0_3, \forall j$	$a_r = 0_3$	$n \leq 2$ and/or collinear motion $n \geq 3$ and non collinear motion	Nothing Relative Position and Speed up to a scale - Unique Solution
	$a_r \neq 0_3$	$n \leq 2$ and/or collinear motion $n = 3$ and coplanar motion $n = 3$ and non coplanar motion $n \geq 4$ and non collinear motion	Nothing Nothing Minimal Case - Relative Position and Speed - Unique Solution Relative Position and Speed - Unique Solution
$\beta_j^2$ along fix axis	$a_r = 0_3$	$n \leq 3$ and/or collinear motion $n = 4$ and non coplanar motion $n = 4$ and coplanar motion $n \geq 5$ and non collinear motion	Nothing Relative Position and Speed up to a scale and roll and pitch - Unique Solution Relative Position and Speed up to a scale and roll and pitch - Two Solutions Relative Position and Speed up to a scale and roll and pitch - Unique Solution
	$a_r \neq 0_3$	$n \leq 3$ and/or collinear motion $n = 4$ and coplanar motion $n = 4$ and non coplanar motion $n = 5$ and coplanar motion $n = 5$ and non coplanar motion $n \geq 6$ and non collinear motion	Nothing Nothing Minimal Case - Relative Position and Speed and roll and pitch - Two Solutions Relative Position and Speed and roll and pitch - Two Solutions Relative Position and Speed and roll and pitch - Unique Solution Relative Position and Speed and roll and pitch - Unique Solution
$\beta_j^2$ along 2 axes	$a_r = 0_3$	$n \leq 3$ and/or collinear motion $n = 4$ and non collinear motion $n \geq 5$ and non collinear motion	Nothing Minimal Case - Relative Orientation and Position and Speed up to a scale - Eight Solutions Relative Orientation and Position and Speed up to a scale - Unique Solution
	$a_r \neq 0_3$	$n \leq 4$ and/or collinear motion $n \geq 5$ and non collinear motion	Nothing Relative Position Speed and Orientation - Unique Solution

TABLE I: Singularities, number of solutions and minimal cases in the case of a single camera

Motion	$n, a_r$	What can be determined
Non Collinear	$n \leq 2$ $n = 3, a_r \neq 0_3$ and coplanar motion $n = 3, a_r \neq 0_3$ and non coplanar motion $n \geq 3, a_r = 0_3$ $n \geq 4, a_r \neq 0_3$	Relative Orientation - Unique Solution Relative Orientation - Unique Solution Minimal Case - Relative Orientation, Position and Speed - Unique Solution Relative Orientation, Position and Speed up to a scale - Unique Solution Relative Orientation, Position and Speed - Unique Solution
Collinear	$n \leq 2$ and/or $\beta_j^2 = \nu$ $n \geq 3$ and $\beta_j^2 \neq \nu$	Relative roll and pitch - Unique Solution Relative Orientation - Unique Solution

TABLE II: Singularities, number of solutions and minimal cases in the case of two synchronized cameras

#### IV. DISCUSSION

In this section, we provide results obtained via synthetic measurements to quantitatively figure out how the performance of the estimation process worsens as we approach a singularity. In particular, we study the following three singularities:

- $\mathcal{S}_1$ , which arises when  $|a_r| = 0$  (cases 1, 3, 5, and the last item in the case 7).
- $\mathcal{S}_2$ , which arises when  $|\beta_j^2| = 0, \forall j$ , and only  $\mathcal{B}_2$  is equipped by a camera (cases 1, 2).
- $\mathcal{S}_3$ : As above but  $\beta_j^2$  is along a fix axis  $\forall j$  (cases 3, 4).

When we estimate the relative state, we expect a large error on the absolute scale near  $\mathcal{S}_1$ , a large error on the entire relative orientation near  $\mathcal{S}_2$ , a large error on the yaw ( $\alpha_y$ ) near  $\mathcal{S}_3$ .

We adopt the same simulation set-up adopted in [2]. In this set-up, we perform Monte Carlo simulations by generating many trials characterized by the following features. The initial position of  $\mathcal{B}_1$  is set at the origin while the position of  $\mathcal{B}_2$  is randomly generated, with a normal distribution, centered at the origin, and with covariance matrix  $1 m^2 I_3$  ( $I_3$  is the  $3 \times 3$  identity matrix). The initial velocities of both  $\mathcal{B}_1$  and  $\mathcal{B}_2$  are randomly generated with normal distribution, zero mean, and covariance matrix  $1 (m/s)^2 I_3$ . The initial orientations are characterized by the roll, pitch and yaw angles. These are also randomly generated, with zero mean and variance  $(50 deg)^2$ . The trajectories of both  $\mathcal{B}_1$  and  $\mathcal{B}_2$  are also randomly generated. Their time duration is 30 seconds. The angular speeds are Gaussian, their values at each step follow a zero-mean Gaussian distribution with covariance matrix equal to  $(1 deg)^2 I_3$ . The relative acceleration ( $a_r$ ) is Gaussian, zero-mean and covariance matrix equal to  $\sigma_{a_r}^2 I_3$ . Finally, the inertial acceleration of  $\mathcal{B}_2$  (denoted by  $a_2$ ) is also randomly generated. The value of  $\sigma_{a_r}$  and the generation model for  $a_2$  are set to reproduce the conditions of the three singularities above. They will be provided in the discussion below. All the experiments are carried out in presence of gravity. Both  $\mathcal{B}_1$  and  $\mathcal{B}_2$  are equipped with inertial sensors able to measure at each time step the acceleration (the sum of the gravity and the inertial acceleration) and the angular speed. These measurements are affected by errors. Specifically, each measurement is generated at every time step of  $0.002 s$  by adding to

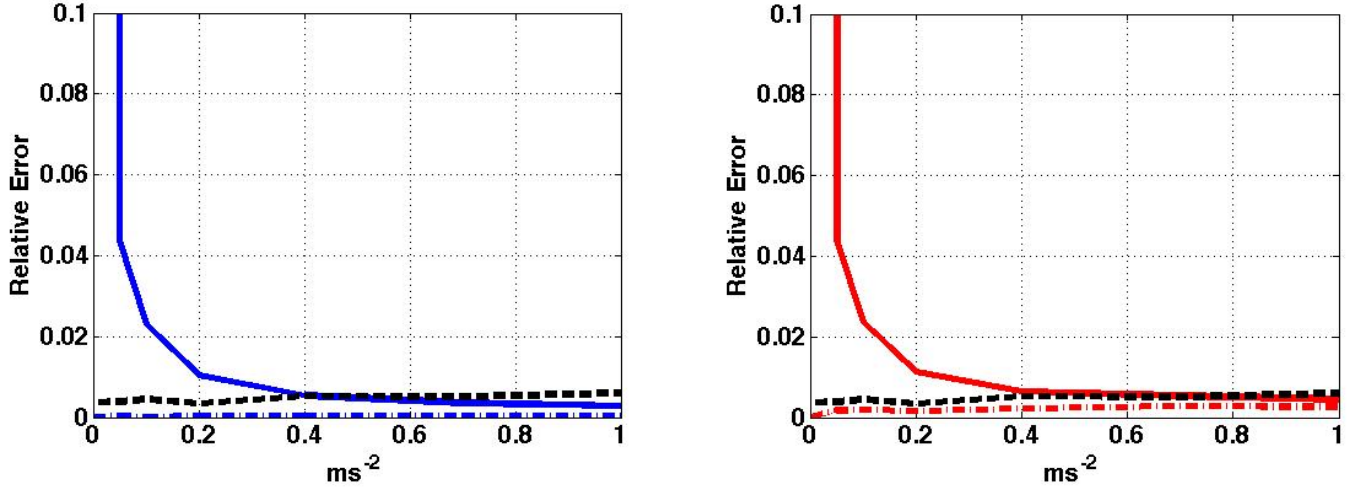


Fig. 1: Results on the  $\mathcal{S}_1$  singularity. The relative error vs  $\sigma_{a_r}$ . On the left: the error on the relative position (solid blue), on the relative position up to a scale (dashed blue) and on the relative orientation (dashed black). On the right: the error on the relative speed (solid red), on the relative speed up to a scale (dashed red) and on the relative orientation (dashed black).

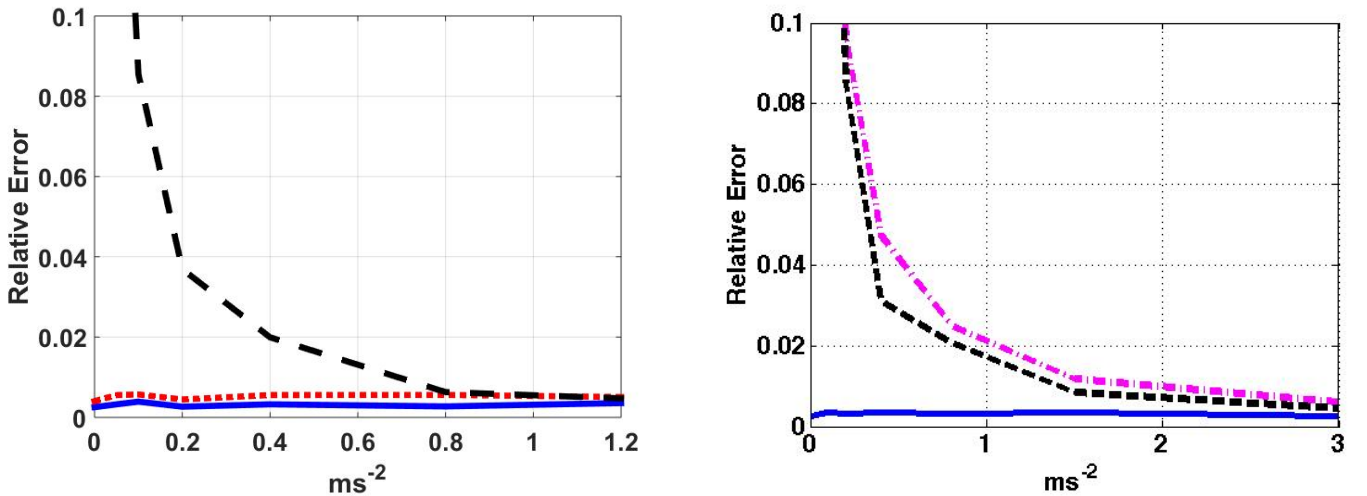


Fig. 2: On the left, results on the  $\mathcal{S}_2$  singularity. The relative error vs  $\sigma_{a_2}$ . Relative position (solid blue), relative speed (dotted red), and relative orientation (dashed black). On the right, results on the  $\mathcal{S}_3$  singularity. The relative error vs  $\sigma_{|a_2|}$ . Relative position (solid blue),  $\alpha_r$  (dashed black),  $\alpha_p$  (dashed magenta).

the true value a random error that follows a Gaussian distribution. The mean value of this error is zero. The standard deviation is  $0.03 \text{ ms}^{-2}$  for the accelerometers and  $0.1 \text{ deg s}^{-1}$  for the gyroscopes. The camera measurements are generated each  $0.2 \text{ s}$ . They are randomly generated by adding to the true value a random error that follows a zero-mean Gaussian distribution, with variance  $1 \text{ deg}^2$ .

For the estimation, we adopt an Extended Kalman Filter (EKF) and we compute the error on the final state. The filter is initialized by setting the initial state randomly, with a normal distribution centered at the true state and with standard deviation equal to 0.2 times the true value (20% initialization error). The results obtained with this EKF are comparable with the results obtained with our analytic solution<sup>3</sup>.

We provide the precision of the EKF in estimating the relative position, the relative speed and the relative orientation. For these three quantities, we provide the relative error, which is obtained by performing 1000 trials and by computing the mean values. Furthermore, as regards the relative position and the relative speed, we compute the average of the errors on their three components. For the relative orientation, to study the first two singularities, we compute the average of the errors on the three

<sup>3</sup>Note that for the EKF we compute the error on the final state while, for the analytic solution, we compute the error on the initial state.



angles that characterize the relative orientation. To study the last singularity, we provide separately the errors on  $\alpha_r$ ,  $\alpha_p$  and  $\alpha_y$  (the roll, pitch and yaw with respect to a frame whose vertical axis coincides with  $\hat{\tau}$ ).

Figure 1 examines the first singularity ( $\mathcal{S}_1$ ).  $a_2$  follows a Gaussian, zero-mean distribution. Its covariance matrix is set equal to  $\sigma_{a_2}^2 I_3$ , with  $\sigma_{a_2} = 1 \text{ ms}^{-2}$ . We plot the relative error vs  $\sigma_{a_r}$ . As expected, when  $\sigma_{a_r} \rightarrow 0$ , the errors on the position and speed diverge (solid blue on the left and dashed red on the right). On the other hand, the relative errors on the same quantities up to a scale are unaffected by  $\sigma_{a_r}$  (dashed blue on the left and dashed red on the right). Similarly, also the error on the relative orientation is independent of  $\sigma_{a_r}$  (dashed black in the two plots).

Figure 2 (left side) examines the singularity  $\mathcal{S}_2$ . We set  $\sigma_{a_r} = 1 \text{ ms}^{-2}$ .  $a_2$  follows a Gaussian distribution. Its mean value coincides with the gravity. Its covariance matrix is set equal to  $\sigma_{a_2}^2 I_3$ . We plot the relative error vs  $\sigma_{a_2}$ . As expected, when  $\sigma_{a_2} \rightarrow 0$ , the error on the relative orientation diverges (dashed black). On the other hand, the errors on the relative position (solid blue) and on the relative speed (dotted red) are independent of  $\sigma_{a_2}$ .

Figure 2 (right side) examines the singularity  $\mathcal{S}_3$ . We set  $\sigma_{a_r} = 1 \text{ ms}^{-2}$ .  $a_2$  is set along the gravity axis and its magnitude follows a zero-mean Gaussian distribution, with variance  $\sigma_{|a_2|}^2$ . We plot the relative error vs  $\sigma_{|a_2|}$ . As expected, when  $\sigma_{|a_2|} \rightarrow 0$ , the error on  $\alpha_r$  (dashed black) and  $\alpha_p$  (dashed magenta) diverge. The error on  $\alpha_y$  is too large to appear on the plot, independently of  $\sigma_{|a_2|}$ . The error on the relative position (solid blue) is independent of  $\sigma_{|a_2|}$ .

## V. CONCLUSION

This paper provided a detailed analysis of all the singularities and minimal cases for the cooperative visual-inertial sensor fusion problem. The analysis was done in the case of two agents and extended the results provided in [1] (for the case of a single agent). As in that case, the key of the analysis was the establishment of the equivalence between the cooperative visual-inertial odometry problem and a Polynomial Equation System (PES). In the case of a single agent, the PES consists of several linear equations and a single polynomial equation of second degree. In the case of two agents, the polynomial equations of second degree become three. The analytic solution of this PES was first provided. The PES provides fundamental insights into all the structural properties of the problem. By analyzing this PES, we obtained all the minimal cases and singularities depending on the number of camera images and the relative trajectory between the agents. The problem, when non singular, can have up to eight distinct solutions. The usefulness of this analysis was illustrated by performing simulations where we adopted an Extended Kalman Filter for the state estimation. We studied the performance of this filter near three singularities, obtained with our analysis. The results fully agree with our expectation: the deterioration of the estimate concerns the quantities that become unobservable in the corresponding singularity.

Future works will be focused on the extension of these results to more complex scenarios. These include the case when the camera is not extrinsically calibrated and/or the case when the inertial measurements are biased. Finally, a further important extension is the case of any number of agents. We believe that, even if the analytic solution cannot be obtained and, consequently, an exhaustive analysis cannot be provided (the PES becomes too complex since it will include any number of polynomial equations of second degree), some important general properties can be derived. To this regard, the study of the case of two agents will play a fundamental role. A special case of this extension is the case when all agents except 2 do not move and have no camera. This coincides with the problem analyzed in this paper when also external point features are available.

## APPENDIX A

### PROOF OF THE RESULTS PROVIDED IN SECTION III-A

*Case 1:* Both properties 1 and 2 hold. Because of property 2, we can ignore the matrix  $R_A$ , which is fully unobservable. Because of property 1, we can fix the value of one  $\lambda$ . Hence, the number of unknowns reduces to  $2 + n$  and the number of equations of  $\tilde{\Sigma}_{Lin}$  is  $3(n - 1)$ . In order to solve the linear system, we must meet the condition  $2n \geq 5$ , i.e., for  $n \leq 2$  the problem is fully singular. Let us assume that the relative motion is not collinear. This means that  $\mu_1, \dots, \mu_n$  do not coincide. We distinguish two cases: (i)  $\mu_2, \dots, \mu_n$  do not coincide (ii)  $\mu_2, \dots, \mu_n$  coincide. In the case (i), the easiest is to fix  $\lambda_1$ . Hence, the matrix that characterizes the remaining linear system in (6), namely  $\tilde{\Sigma}_{Lin}$ , is:

$$\left[ \begin{array}{c|c|c|c|c|c|c|c} \Delta_2 I_3 & -\mu_2 & 0_3 & \cdots & \cdots & \cdots & 0_3 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ \Delta_j I_3 & \cdots & 0_3 & -\mu_j & 0_3 & \cdots & 0_3 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ \Delta_n I_3 & \cdots & \cdots & \cdots & \cdots & 0_3 & -\mu_n & \end{array} \right]$$

It is immediate to verify that this matrix is full rank if at least two unit vectors among  $\mu_2, \dots, \mu_n$  do not coincide (which is the case). Let us consider the case (ii). We have  $\mu_2 = \dots = \mu_n \neq \mu_1$ . By fixing  $\lambda_2$ , the matrix that characterizes  $\tilde{\Sigma}_{Lin}$  becomes:

$$\left[ \begin{array}{c|c|c|c|c|c|c|c} \Delta_2 I_3 & \mu_1 & 0_3 & \cdots & \cdots & \cdots & 0_3 & \\ \Delta_3 I_3 & \mu_1 & -\mu_3 & 0_3 & \cdots & \cdots & 0_3 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ \Delta_n I_3 & \mu_1 & \cdots & \cdots & \cdots & 0_3 & -\mu_n & \end{array} \right]$$

By performing on this matrix simple operations that preserve the rank, it is possible to achieve the same structure of the matrix in the case (i). Specifically, with  $\mu_1$  instead of  $\mu_2$  and with different quantities instead of  $\Delta_2, \dots, \Delta_n$  (but these new quantities still differ one each other). Since  $\mu_1$  is different from the other  $\mu$ , this matrix is full rank.

*Case 2:* Because of property 2 we can ignore  $R_A$ . The number of equations  $(3n - 3)$  equals the number of unknowns  $(3 + n)$  precisely when  $n = 3$ . We consider this case. The matrix is:  $\left[ \begin{array}{c|c|c|c} \Delta_2 I_3 & \mu_1 & -\mu_2 & 0_3 \\ \Delta_3 I_3 & \mu_1 & 0_3 & -\mu_3 \end{array} \right]$ , which is full rank iff  $\mu_1 \cdot (\mu_2 \wedge \mu_3) \neq 0$  (where the symbol " $\wedge$ " denotes the cross product). Finally, if  $n \geq 4$  it suffices that the relative motion is not collinear to make the matrix of  $\tilde{\Sigma}_{Lin}$  full rank.

*Case 3:*  $\tilde{\Sigma}_{Lin}$  becomes (see eq. (3) with  $R_A \hat{\beta}_j^2 = |\beta_j^2| \hat{\tau}$ ):

$$\Delta_j \eta_A + |\beta_j^2| \hat{\tau} + \lambda_1 \mu_1 - \lambda_j \mu_j = \beta_j^1 \quad (7)$$

Because of property 1, the scale is unobservable. Hence, we fix arbitrarily  $\lambda_1$  and we remove it from the unknowns. The remaining unknowns in the linear system in (7) are:  $\eta_A, \hat{\tau}, \lambda_2, \dots, \lambda_n$ . On the other hand, they are not independent because  $\hat{\tau}$  is a unit vector. Hence, the number of independent unknowns is  $4 + n$ . To have  $3n - 3 \geq 4 + n$  we must meet the condition  $2n \geq 7$ . Hence,  $n \geq 4$ . Let us study the case  $n = 4$ . The matrix of the linear system in (7) is:

$$\left[ \begin{array}{c|c|c|c|c} \Delta_2 I_3 & |\beta_2^2| I_3 & -\mu_2 & 0_3 & 0_3 \\ \Delta_3 I_3 & |\beta_3^2| I_3 & 0_3 & -\mu_3 & 0_3 \\ \Delta_4 I_3 & |\beta_4^2| I_3 & 0_3 & 0_3 & -\mu_4 \end{array} \right],$$

whose dimension is  $9 \times 9$ . Its determinant, when the condition  $|\beta_j^2| \propto \Delta_j$  does not hold, is proportional to  $\mu_2 \cdot (\mu_3 \wedge \mu_4)$ . Hence, a sufficient condition for the resolvability is that the relative motion is not coplanar (note that we could have chosen to remove another  $\lambda$  and the co planarity regards also  $\mu_1$ ). This condition guarantees a unique solution. Let us consider the case of a coplanar relative motion. The determinant of the previous matrix vanishes. We compute its null space. It consists of a single vector. We consider the part of this vector that corresponds to  $\hat{\tau}$ . When it does not vanish, we can still solve the problem and we have two solutions. This part is equal to a combination of the following 3 vectors:  $|\mu_2| |\mu_3 \wedge \mu_4|$ ,  $|\mu_3| |\mu_2 \wedge \mu_4|$  and  $|\mu_4| |\mu_2 \wedge \mu_3|$ . Hence, it vanishes iff all the positions are collinear. This means that, with the exception of a collinear relative motion, when  $n = 4$  and the relative motion is coplanar, the *COVIO* has 2 distinct solutions (which are obtained by requiring that  $\hat{\tau}$  is a unit vector). Let us refer now to the case  $n \geq 5$ . By performing on the corresponding matrix simple operations that preserve the rank, it is possible to obtain the following matrix:

$$\left[ \begin{array}{c|c|c|c|c|c|c|c} \Delta_2 I_3 & |\beta_2^2| I_3 & \mu_2 & 0_3 & 0_3 & \dots & 0_3 \\ 0_{33} & \tilde{\beta}_3 I_3 & -\eta_3 \mu_2 & \mu_3 & 0_3 & \dots & 0_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0_{33} & 0_{33} & (\eta_3 \tilde{\eta}_j - \eta_j) \mu_2 & -\tilde{\eta}_j \mu_3 & \dots & \mu_j & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0_{33} & 0_{33} & (\eta_3 \tilde{\eta}_n - \eta_n) \mu_2 & -\tilde{\eta}_n \mu_3 & 0_3 & \dots & \mu_n \end{array} \right]$$

where  $0_{33}$  is the zero  $3 \times 3$  matrix,  $\tilde{\eta}_j = \frac{\tilde{\beta}_j}{\beta_3}$ ,  $\eta_j = \frac{\Delta_j}{\Delta_2}$  and  $\tilde{\beta}_j = |\beta_j^2| - \eta_j |\beta_2^2|$  ( $\neq 0$ , by assumption, since the special condition  $|\beta_j^2| \propto \Delta_j$  does not hold). This matrix is not full rank only when the unit vectors  $\mu_2, \dots, \mu_n$  coincide. On the other hand, we could have chosen to remove another  $\lambda$ , e.g.  $\lambda_2$ . By proceeding as in the second part of the proof of the case 1, we can prove that the matrix that characterizes the resulting linear system is not full rank only when the unit vectors  $\mu_1, \mu_3, \dots, \mu_n$  coincide. Hence, if the relative motion is not collinear,  $\xi_1, \dots, \xi_n$  and  $\eta_A$  up to scale, together with the roll and pitch angles with respect to the axis  $\hat{\beta}^2$ , can be uniquely determined.

*Case 4:* As in the previous case,  $\tilde{\Sigma}_{Lin}$  is given by equation (7). The number of independent unknowns is  $5 + n$ . To have  $3n - 3 \geq 5 + n$  we must meet the condition  $2n \geq 8$ . Hence,  $n \geq 4$  and  $n = 4$  is a minimal case. Let us study this case. The matrix of  $\tilde{\Sigma}_{Lin}$  is:

$$\left[ \begin{array}{c|c|c|c|c|c} \Delta_2 I_3 & |\beta_2^2| I_3 & \mu_1 & -\mu_2 & 0_3 & 0_3 \\ \Delta_3 I_3 & |\beta_3^2| I_3 & \mu_1 & 0_3 & -\mu_3 & 0_3 \\ \Delta_4 I_3 & |\beta_4^2| I_3 & \mu_1 & 0_3 & 0_3 & -\mu_4 \end{array} \right],$$

whose dimension is  $9 \times 10$ . It is full rank (i.e., rank 9) iff the relative motion is not coplanar. Additionally, the part of the single vector in the null space that corresponds to  $\hat{\tau}$ , does not vanish. Hence we have two solutions<sup>4</sup>. Let us refer now to the case  $n = 5$ . The matrix of this linear system has dimension is  $12 \times 11$ . By performing on this matrix simple operations that preserve the rank, it is possible to obtain the following matrix:

<sup>4</sup>Note that, if the nullspace of a matrix has dimension 1, the solution of the corresponding linear system is  $\vec{x} = \vec{x}_p + \lambda \hat{n}$ , where:  $\vec{x}_p$  is a particular solution of the linear system,  $\hat{n}$  is the unit vector that generates the nullspace and  $\lambda$  is a free real parameter (to be determined). Since in this case 3 entries of  $\vec{x}$  correspond to the unit vector  $\hat{\tau}$ ,  $\lambda$  is determined by the constraint that these three entries make a unit vector. Enforcing this constraint results in a second order polynomial equation in  $\lambda$  that provides 2 solutions for  $\lambda$  and, as a result, 2 solutions for  $\vec{x}$ .

$$\left[ \begin{array}{c|c|c|c|c|c|c|c|c} \Delta_2 I_3 & \beta_2^2 I_3 & \mu_1 & \mu_2 & 0_3 & 0_3 & 0_3 & & \\ \hline 0_{33} & \tilde{\beta}_3 I_3 & (1 - \eta_3) \mu_1 & -\eta_3 \mu_2 & \mu_3 & 0_3 & 0_3 & & \\ \hline 0_{33} & 0_{33} & \gamma_4 \mu_1 & (\eta_3 \tilde{\eta}_4 - \eta_4) \mu_2 & -\tilde{\eta}_4 \mu_3 & \mu_4 & 0_3 & & \\ \hline 0_{33} & 0_{33} & \gamma_5 \mu_1 & (\eta_3 \tilde{\eta}_5 - \eta_5) \mu_2 & -\tilde{\eta}_5 \mu_3 & 0_3 & \mu_5 & & \end{array} \right]$$

where  $\gamma_j = 1 - \eta_j - \frac{\tilde{\beta}_j}{\beta_3} (1 - \eta_3)$ , for  $j = 4, 5$ . It is possible to check that, when the span of  $\mu_1, \dots, \mu_5$  has dimension 3, this matrix is full rank. When is 2, the null space of the matrix has dimension 1, in which case, 2 distinct solutions can be determined by exploiting the fact that  $\hat{r}$  is a unit vector. Let us consider the case  $n \geq 6$ . By performing the same operations on the matrix that characterizes the linear system, it is possible to verify that, even in the coplanar case (but not collinear), the matrix is full rank.

*Case 5:* In this case, no component of  $R_A$  can be removed from  $\tilde{\Sigma}_{Lin}$ . On the other hand, because of property 1, the scale is unobservable and we can remove  $\lambda_1$ . Hence, the unknowns are  $\eta_A, R_A, \lambda_2, \dots, \lambda_n$  and the number of independent unknowns is  $3 + 3 + n - 1 = 5 + n$ . To have  $3n - 3 \geq 5 + n$  we must meet the condition  $n \geq 4$ .  $n = 4$  is a minimal case. Let us study this case. The matrix is:

$$\left[ \begin{array}{c|c|c|c|c|c|c|c|c} \Delta_2 I_3 & \beta_{x2}^2 I_3 & \beta_{y2}^2 I_3 & \beta_{z2}^2 I_3 & -\mu_2 & 0_3 & 0_3 & & \\ \hline \Delta_3 I_3 & \beta_{x3}^2 I_3 & \beta_{y3}^2 I_3 & \beta_{z3}^2 I_3 & 0_3 & -\mu_3 & 0_3 & & \\ \hline \Delta_4 I_3 & \beta_{x4}^2 I_3 & \beta_{y4}^2 I_3 & \beta_{z4}^2 I_3 & 0_3 & 0_3 & -\mu_4 & & \end{array} \right]$$

whose dimension is  $9 \times 15$ . We use 6 equations to eliminate  $\eta_A, \lambda_2, \lambda_3, \lambda_4$ . This elimination is possible iff the relative motion is not collinear. Note that, this condition also guarantees that the remaining 3 equations are independent. By solving  $\mathcal{P}$  we obtain up to eight real solutions. If  $n \geq 5$ , by proceeding as above, we obtain that  $\mathcal{P}$  includes two further quadratic equations. This allows us to obtain a unique solution.

*Case 6:* With respect to the previous case, property 1 does not hold and no  $\lambda$  variable can be fixed and removed. As a result, we have a further unknown and we must meet the condition  $3n - 3 \geq 6 + n$ . Hence,  $2n \geq 9$  and, consequently,  $n \geq 5$ . In the case  $n = 5$ , we use 8 equations to eliminate  $\eta_A, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ . This elimination is possible iff the relative motion is not collinear. Note that, this condition also guarantees that the remaining 4 equations are independent. As a result, we obtain four quadratic equations in three unknowns and the solution is unique<sup>5</sup>.

## APPENDIX B

### PROOF OF THE RESULTS PROVIDED IN SECTION III-B

*Case 7:* From property 3 we know that  $R_A$  can be uniquely determined. We substitute its value in (3), which is now a system of equations in the unknowns  $\lambda_1, \dots, \lambda_n$  and  $\eta_A$ . This is precisely the same linear system that characterizes the cases 1 and 2. Hence, regarding the remaining unknowns, the same results hold.

*Case 8:* Equation (2) becomes:  $R_A \nu = -\mu$ . This equation allows us to determine  $R_A$  up to a rotation around  $\mu$ . Specifically, we obtain  $R_A = R_\mu(\theta) R_0$ , where:

- $R_0$  is a rotation matrix that transforms the vector  $\nu$  in  $-\mu$ . It can be determined from  $\mu$  and  $\nu$  and so it is known.
- $R_\mu(\theta)$  is the rotation matrix that performs a rotation of  $\theta$  around  $\mu$ . It depends linearly on  $\cos \theta$  and  $\sin \theta$ .

We substitute  $R_A = R_\mu(\theta) R_0$  in equation (3). We obtain a linear system in the unknowns:  $[\eta_A \cos \theta \sin \theta \lambda_1 \dots \lambda_n]$ . If  $\hat{\beta}_j^2 = \nu, \forall j = 1, \dots, n$ , this linear system is independent of  $\theta$ . Hence  $\theta$  is unobservable. Regarding the remaining unknowns, we remark that the linear system is precisely the same of the cases 1 and 2. As a result, the remaining unknowns cannot be determined even up to a scale (we remind the reader that the relative motion is collinear). Let us consider the case when  $\hat{\beta}_j^2$  is along an axis different from  $\nu$  for at least one  $j$ . The resulting linear system is characterized by the following matrix:

$$\left[ \begin{array}{c|c|c|c|c|c|c|c|c|c|c|c|c|c} \Delta_2 I_3 & \beta_2^{\cos} & \beta_2^{\sin} & \mu & -\mu & 0_3 & \dots & \dots & \dots & \dots & 0_3 & & & \\ \hline \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & & & \\ \hline \Delta_j I_3 & \beta_j^{\cos} & \beta_j^{\sin} & \mu & \dots & 0_3 & -\mu & 0_3 & \dots & \dots & 0_3 & & & \\ \hline \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & & & \end{array} \right]$$

where  $\beta_j^{\cos} = \tilde{\beta}_j - \mu(\tilde{\beta}_j \cdot \mu)$ ,  $\beta_j^{\sin} = \mu \wedge \tilde{\beta}_j$  and  $\tilde{\beta}_j = R_0 \hat{\beta}_j^2$ . The following two vectors belong to its null space:  $[\mu^T, 0_2, 0, \Delta_2, \Delta_3, \dots, \Delta_n]$  and  $[0_3, 0_2, 1, 1, \dots, 1]$ . By fixing two  $\lambda$  (e.g.,  $\lambda_1$  and  $\lambda_2$ ), we obtain a linear system that is full rank, meaning that  $\theta$  can be determined.

<sup>5</sup>However, if the 5 vectors  $\hat{\beta}_j^2$  are coplanar, the rank of the remaining linear system is 3 and consequently we have up to eight solutions (this result does not hold when  $n \geq 6$ , for which we always have a unique solution, even if the 6 vectors  $\hat{\beta}_j^2$  are coplanar).

## REFERENCES

- [1] A. Martinelli, Closed-form solution of visual-inertial structure from motion, *International Journal of Computer Vision (IJCV)*, vol. 106, no. 2, pp. 138–152, 2014.
- [2] A. Martinelli, A. Oliva and B. Mourain, Cooperative Visual-Inertial Sensor Fusion: the Analytic Solution, *IEEE Robotics and Automation Letters*, Volume: 4 , Issue: 2 , April 2019, pages: 453–460
- [3] A. Mourikis and S. Roumeliotis, A multi-state constraint kalman filter for vision-aided inertial navigation, *IEEE International Conference on Robotics and Automation*, 2007, Rome, Italy.
- [4] A. Mourikis and S. Roumeliotis, A dual-layer estimator architecture for long-term localization, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008, Anchorage, USA.
- [5] G. Huang, A. Mourikis, S. Roumeliotis, An observability-constrained sliding window filter for slam, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011, S. Francisco, USA.
- [6] T. Lupton and S. Sukkariéh, Visual-inertial-aided navigation for high-dynamic motion in built environments without initial conditions, *IEEE Transactions on Robotics*, vol. 28, pp. 61–76, 2012.
- [7] V. Indelman, S. Williams, M. Kaess, and F. Dellaert, Information fusion in navigation systems via factor graph based incremental smoothing, *Robotics and Autonomous Systems*, pp. 721–738, 2013.
- [8] S. Leutenegger, P. Furgale, V. Rabaud, M. Chli, K. Konolige, and R. Siegwart, Keyframe-based visual-inertial odometry using nonlinear optimization, *The International Journal of Robotics Research*, vol 34, pp. 314–334, 2015.
- [9] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, Imu preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation, *Robotics: Science and Systems*, 2015, Rome, Italy.
- [10] C. Forster, L. Carlone, F. Dellaert, and D. Scaramuzza, On-manifold preintegration for real-time visual-inertial odometry, *IEEE Trans. Robot.*, vol. 33, no. 1, pp. 1–21, Feb. 2017
- [11] J. Kaiser, A. Martinelli, F. Fontana and D. Scaramuzza, Simultaneous State Initialization and Gyroscope Bias Calibration in Visual Inertial Aided Navigation, *IEEE Robotics and Automation Letters*, vol 2, pp. 18–25, 2017.
- [12] H. Christopher Longuet-Higgins (September 1981). "A computer algorithm for reconstructing a scene from two projections". *Nature* 293: 133–135.
- [13] Richard I. Hartley (June 1997). "In Defense of the Eight-Point Algorithm". *IEEE Transaction on Pattern Recognition and Machine Intelligence* 19 (6): 580–593.
- [14] D. Nistér, An efficient solution to the five-point relative pose problem, *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 26(6):756-770, June 2004
- [15] C. X. Guo, K. Sartipi, R. C. DuToit, G. A. Georgiou, R. Li, J. O. Leary, E. Nerurkar, J. Hesck, S. Roumeliotis, Large-Scale Cooperative 3D Visual-Inertial Mapping in a Manhattan World, *IEEE International Conference on Robotics and Automation*, 2016, Stockholm, Sweden.
- [16] A. Martinelli, A. Renzaglia and A. Oliva, Cooperative Visual-Inertial Sensor Fusion: Fundamental Equations and State determination in Closed Form, *Autonomous Robots* (2019). <https://doi.org/10.1007/s10514-019-09841-8>,
- [17] S. Telen, B. Mourrain and M. Van Barel, Solving Polynomial Systems via a Stabilized Representation of Quotient Algebras, arXiv:1711.04543 (to appear on the *SIAM Journal on Matrix Analysis and Applications*)
- [18] G.L. Mariottini, S. Martini, and M. Egerstedt, A Switching Active Sensing Strategy to Maintain Observability for Vision-Based Formation Control, *IEEE International Conference on Robotics and Automation*, 2009, Kobe, Japan.