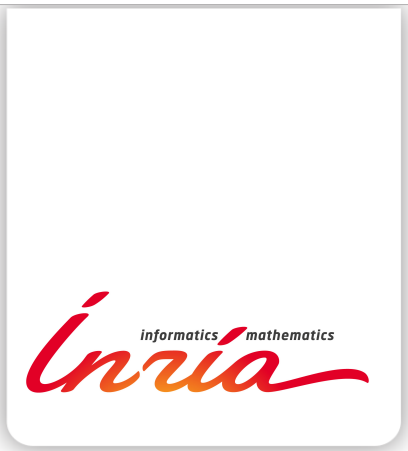


Nonlinear Unknown Input Observability: The Analytic Solution in the case of a Single Unknown Input

Agostino Martinelli

RESEARCH REPORT
N° 9027
Avril 2017
Project-Team CHROMA

ISSN 0249-6399 ISRN INRIA/RR--9027--FR+ENG



Nonlinear Unknown Input Observability: The Analytic Solution in the case of a Single Unknown Input

Agostino Martinelli

Project-Team CHROMA

Research Report n° 9027 — Avril 2017 — 55 pages

Abstract: This work presents the analytic solution of a fundamental open problem in the framework of nonlinear observability, which is the unknown input observability problem (UIO problem). The solution here provided holds in the case of a single unknown input. The first part of the work presents this analytic solution, namely the analytic criterion that allows us to obtain the observability of a nonlinear system in presence of a single unknown input and multiple known inputs. As for the observability rank condition, the proposed analytic criterion is based on the computation of the observable codistribution. Similarly to the case of only known inputs, the observable codistribution is obtained by recursively computing the Lie derivatives of the outputs along the vector fields that characterize the dynamics. However, in correspondence of the unknown input, the corresponding vector field must be suitably rescaled. Additionally, the Lie derivatives of the outputs must also be computed along a new set of vector fields that are obtained by recursively performing suitable Lie bracketing of the vector fields that define the dynamics. In practice, the entire observable codistribution is obtained by a very simple recursive algorithm. Finally, it is shown that the recursive algorithm converges in a finite number of steps and the criterion to establish that the convergence has been reached is provided. The proposed analytic extension of the observability rank condition is illustrated by checking the weak local observability of several nonlinear systems driven by multiple known inputs and a single unknown input. For these systems, extensive simulations are also provided. It is shown that, a simple estimator based on an Extended Kalman Filter, provides results that agree with what we could expect from the observability analysis. Finally, it is shown that the analytic criterion introduced in this work is not only a key analytical tool to automatically obtain the state observability but it can be also a fundamental tool to design an appropriate estimator.

Key-words: Nonlinear observability; Unknown Input Observability; Observability Rank Condition; Observable Codistribution; Robotics Applications

RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

Inovallée
655 avenue de l'Europe Montbonnot
38334 Saint Ismier Cedex

Observabilité non linéaire avec entrées inconnues: Solution Analytique dans le cas d'une seule entrée inconnue

Résumé : Ce travail présente la solution analytique d 'un problème ouvert fondamental dans le cadre de l' observabilité non linéaire, qui est le problème de l' observabilité non linéaire quand une partie (voire la totalité) des entrées du système est inconnu. La solution ici fournie est valable dans le cas d'une seule entrée inconnue. La première partie du travail présente cette solution analytique, c'est à dire le critère analytique qui permet d'obtenir l'observabilité d'un système non linéaire en présence d'une seule entrée inconnue et de multiples entrées connues. Pareil que pour la solution dans le cas sans entrées inconnues, le critère analytique proposé est basé sur le calcul de la codistribution observable. De même pour le cas des entrées connues seulement, la codistribution observable est obtenue en calculant de façon récursive les dérivées de Lie des sorties du système le long des champs vectoriels qui caractérisent la dynamique. Cependant, en correspondance de l'entrée inconnue, le champ vectoriel correspondant doit être réétalonné de manière appropriée. De plus, les dérivées de Lie des sorties du système doivent également être calculées le long d'un nouvel ensemble de champs vectoriels qui sont obtenus en effectuant récursivement un bracketing de Lie approprié des champs vectoriels qui définissent la dynamique. En pratique, la totalité de la codistribution observable est obtenue par un algorithme récursif très simple. Enfin, il est montré que l'algorithme récursif converge en un nombre fini d'étapes et que le critère pour établir que la convergence a été atteinte est fourni. La solution analytique proposée est illustrée par l'étude de l'observabilité de plusieurs systèmes non linéaires pilotés par plusieurs entrées connues et une seule entrée inconnue. Pour ces systèmes, des simulations sont également fournies. Il est montré qu'un estimateur simple basé sur un filtre de Kalman étendu fournit des résultats qui concordent avec ce que l'on pourrait attendre de l'analyse d'observabilité. Enfin, il est montré que le critère analytique introduit dans ce travail n'est pas seulement un outil analytique clé pour obtenir automatiquement l'observabilité de l'état mais il peut aussi être un outil fondamental pour concevoir un estimateur approprié.

Mots-clés : observabilité non linéaire; Observabilité avec entrées inconnues; Applications dans la robotique

1 Introduction

State observability is a necessary condition that a state must satisfy to be estimated. Checking the system observability is a fundamental step in state estimation. It has become praxis in many application fields to provide an observability analysis prior to solving an estimation problem (e.g., in robotics [4, 12, 13, 14, 20, 33, 34, 35, 37], in visual-inertial sensor fusion [9, 10, 11, 17, 18, 19, 21, 25, 26, 30, 32], in sensor calibration [5, 7, 22, 23, 31]). Investigating the observability properties is very simple in the linear case. Unfortunately, real systems are very rarely characterized by linearity.

The control theory community has developed the analytic tool necessary to check the state observability for nonlinear systems provided that all the system inputs are known. This is the observability rank condition introduced by Herman and Krener in 1977 [8]. In accordance with this criterion, it is possible to derive all the observability properties of a nonlinear system by performing automatic computation. On the other hand, in many real scenarios, one or more disturbances can dramatically impact the system dynamics. A disturbance can be considered as an unknown input. Its presence can dramatically affect the observability properties of the state. This is for instance the case of a drone that operates in presence of wind. The wind is in general unknown, time-variant and acts on the system dynamics as an unknown input.

The problem of deriving an analytic tool able to determine the observability properties in presence of unknown inputs is known as the *Unknown Input Observability* (UIO) problem. This problem was introduced and firstly investigated by the automatic control community in the seventies [1, 3, 6, 36]. In particular, Basile and Marro provided the solution of this problem in the linear case [1]. In many application fields, most of the systems are characterized by nonlinear dynamics, even in very simple cases (e.g. in planar robotics). Additionally, the presence of disturbances cannot be ignored in many cases and can significantly affect the observability properties. The UIO problem in the nonlinear case has recently been investigated by the robotics community and partial solutions have recently been proposed [2, 26, 28, 29]. These solutions only provide sufficient conditions for the state observability. They are based on a suitable state extension.

This work provides the analytic solution of the UIO problem in the nonlinear case in the case of a single unknown input. This solution has been presented at the SIAM symposium on control and applications 2015 [27]. In section 2 we define the class of systems for which we provide the solution. This class includes any nonlinear system linear in the inputs (both known and unknown).

In [8, 16] the observability properties of a nonlinear system are obtained by computing the observable codistribution. The computation of this codistribution is the core of the observability rank condition introduced in [8]. In order to deal with the case of unknown inputs, we need to derive a new algorithm able to generate the observable codistribution. In section 3 we remind the reader the algorithm to compute the observable codistribution in the case without disturbances, together with some basic properties that characterize its convergence (section 3.1). Then, in section 3.2, we introduce the new algorithm that generates the entire observable codistribution in presence of a single unknown input, together with some basic properties that characterize its convergence. The solution of the UIO problem in the case of a single unknown input is summarized in section 4 and it is illustrated in section 5 by checking the observability of several robotics systems driven by multiple known inputs and a single unknown input, ranging from planar robotics up to aerial navigation systems. This section also contains extensive simulations and introduces a simple estimator. The results obtained by estimating the state with this estimator fully agree with the observability analysis carried out by using the proposed analytic criterion. In section 6 we provide all the analytical derivations necessary to prove the validity of the analytic

results presented in section 3. In section 7 we provide our conclusion by remarking that, the analytic criterion here introduced, is not only a powerful and simple criterion to automatically detect the observable states for a general nonlinear system affected by a disturbance (i.e., the analytic solution of the UIO problem in the case of a single unknown input) but it can be also a key tool for designing an appropriate estimator.

2 System definition

We will refer to a nonlinear control system with m_u known inputs ($u \triangleq [u_1, \cdots, u_{m_u}]^T$) and a single unknown input or disturbance (w). The state is the vector $x \in M$, with M an open set of $\mathbb{R}^n$. We assume that the dynamics are nonlinear with respect to the state and affine with respect to the inputs (both known and unknown). Finally, for the sake of simplicity, we will refer to the case of a single output y and we provide the extension to multiple outputs, which is trivial, separately. This will allow us to avoid the introduction of a further index. Our system is characterized by the following equations:

$$\begin{cases} \dot{x} = \sum_{i=1}^{m_u} f^i(x)u_i + g(x)w \\ y = h(x) \end{cases} \tag{1}$$

where $f^i(x)$, $i = 1, \cdots, m_u$ and $g(x)$ are vector fields in M and the function $h(x)$ is a scalar function defined on the open set M. Finally, we assume that the unknown input w is an analytic function of time.
Throughout this document, in the case when $m_u = 1$, we denote by $f(x)$ the vector field $f^1(x)$.

3 Observable Codistribution

As for the observability rank condition, the analytic criterion to obtain the observability properties of the system in (1), is obtained by computing the observable codistribution[1]. In this section we provide the algorithm to compute this codistribution. For the clarity sake, we start by reminding the reader the standard algorithm that generates the observable codisribution in absence of unknown inputs (section 3.1). Then, in section 3.2, we provide the algorithm for the system in (1).

Note that this section directly provides the analytic results. All the analytic proofs will be provided later, in section 6.

3.1 Observable codistribution in the case without unknown inputs

We consider the system in (1) when all the inputs are known. We will denote with the symbol $\mathcal{D}$ the differential with respect to the state x. For instance, if $x = [x_1,\ x_2]^T$ and $h = x_1 + x_2^2$, we have[2]: $\mathcal{D}h = \mathcal{D}x_1 + 2x_2\mathcal{D}x_2$.

For a given codistribution Ω and a given vector field $\theta = \theta(x)$ (both defined on the open set M), we denote by $\mathcal{L}_\theta\Omega$ the codistribution whose covectors are the Lie derivatives along θ of the covectors in Ω. We remind the reader that the Lie derivative of a scalar function $h(x)$ along the vector field $f(x)$ is defined as follows:

$$\mathcal{L}_f h \triangleq \frac{\partial h}{\partial x} f$$

which is the product of the row vector $\frac{\partial h}{\partial x}$ with the column vector f. Hence, it is a scalar function. Additionally, by definition of Lie derivative of covectors, we have: $\mathcal{L}_f\mathcal{D}h = \mathcal{D}\mathcal{L}_f h$.

Finally, given two vector spaces V_1 and V_2, we denote by $V_1 + V_2$ their sum, i.e., the span of all the generators of both V_1 and V_2.

The observable codistribution is generated by the following recursive algorithm (see [8] and [16]):

Algorithm 1 Observable codistribution when w is known

1. $\Omega_0 = span\{\mathcal{D}h\}$;
2. $\Omega_m = \Omega_{m-1} + \sum_{i=1}^{m_u} \mathcal{L}_{f^i}\Omega_{m-1} + \mathcal{L}_g\Omega_{m-1}$

In presence of multiple outputs, we only need to add to the codistribution Ω_0, the span of the differentials of the remaining outputs. In [16] it is proven that this algorithm converges. In particular, it is proven that it has converged when $\Omega_m = \Omega_{m-1}$. From this, it is easy to realize that the convergence is achieved in at most $n-1$ steps[3].

[1] The reader non-familiar with the concept of *distribution*, as it is used in [16], should not be afraid by the term *distribution* and the term *codistribution*. Very simply speaking, a distribution is a vector space defined on M (our set in $\mathbb{R}^n$ where the system is defined). In particular, this vector space changes by moving on M. This vector space is in fact the span of a set of vector functions (vector fields) defined on M. A codistribution is the dual of a distribution. Very simply speaking (and this is enough to understand this manuscript) a distribution is generated by a set of column vectors. A codistribution is generated by a set of line vectors. All these vectors are vector functions (i.e., they depends on the point $x \in M$) and they have the same dimension of x.

[2] The span of the differentials of a set of scalar functions is a codistribution. The reader non familiar with the theory of distributions can simply consider the differential as the gradient operator. The gradient of a scalar function is a line vector. For instance, if $x = [x_1,\ x_2]^T$ and $h = x_1 + x_2^2$, we obtain for its gradient the line vector function $[1,\ 2x_2]$. Later, in section 5, we adopt this representation. According to this, a codistribuion will be the span of a set of line vectors and a covector (i.e., an element of a codistribution) will be a line vector.

[3] This is a consequence of lemmas 1.9.1, 1.9.2 and 1.9.6 in [16].

3.2 Observable codistribution in presence of a single unknown input

We now provide the new algorithm that generates the observable codistribution when w is an unknown input. We will denote by $L_g^1 = L_g^1(x)$ the first order Lie derivative of the function $h(x)$ along the vector field $g(x)$, i.e.,

$$L_g^1 \triangleq \mathcal{L}_g h \tag{2}$$

The analytic computation of the observable codistribution is based on the assumption that $L_g^1 \neq 0$ on a given neighbourhood of x_0. In appendix A, we deal with the case when this condition is not satisfied. Specifically, we show that it is either possible to redefine the output, without altering the system observability properties, such that the Lie derivative of the new output along g does not vanish or the unknown input is spurious (i.e., it does not affect the observability properties). For these reasons, we can assume that $L_g^1 \neq 0$.

Before introducing the new algorithm that generates the entire observable codistribution, we introduce a new set of vector fields ${}^i\phi_m \in \mathbb{R}^n$ ($i = 1, \cdots, m_u$ and for any integer m). They are obtained recursively by the following algorithm:

Algorithm 2

1. ${}^i\phi_0 = f^i$;
2. ${}^i\phi_m = \frac{[{}^i\phi_{m-1},\ g]}{L_g^1}$

where the parenthesis $[\cdot, \cdot]$ denote the Lie bracket of vector fields, defined as follows:

$$[a,\ b] \triangleq \frac{\partial b}{\partial x} a(x) - \frac{\partial a}{\partial x} b(x)$$

In other words, for each $i = 1, \cdots, m_u$, we have one new vector field at every step of the algorithm. Throughout this document, in the case when $m_u = 1$, we denote by ϕ_m the vector field ${}^1\phi_m$.

We are now ready to provide the algorithm that generates the entire observable codistribution. It is the following:

Algorithm 3 Observable codistribution when w is unknown

1. $\Omega_0 = span\{\mathcal{D}h\}$;
2. $\Omega_m = \Omega_{m-1} + \sum_{i=1}^{m_u} \mathcal{L}_{f^i} \Omega_{m-1} + \mathcal{L}_{\frac{g}{L_g^1}} \Omega_{m-1} + \sum_{i=1}^{m_u} span\left\{\mathcal{L}_{{}^i\phi_{m-1}} \mathcal{D}h\right\}$

In presence of multiple outputs, we only need to add to the codistribution Ω_0, the span of the differentials of the remaining outputs. Note that, in presence of multiple outputs, the function L_g^1 is still a scalar function since it is still defined by using a single output. The result is independent of the chosen output (provided that L_g^1 does not vanish[4]).

In section 6.2.2 we investigate the convergence properties of algorithm 3. We consider first the case of a single known input (i.e., $m_u = 1$) and then the results are easily extended to the case of multiple known inputs ($m_u > 1$) in section 6.2.3. We prove that algorithm 3 converges and we also provide the analytic criterion to check that the convergence has been attained. This proof and the convergence criterion cannot be the same that hold for algorithm 1, because of the last term

[4]The case when $L_g^1 = 0$ for all the outputs is dealt in appendix A.

that appears at the recursive step[5], i.e., the term $\sum_{i=1}^{m_u} span\left\{\mathcal{L}_{^i\phi_{m-1}}\mathcal{D}h\right\}$ (the special case when, the contribution due to this last term is included in the other terms, is considered separately by lemma 6). In general, the criterion to establish that the convergence has been attained is not simply obtained by checking the condition $\Omega_{m+1} = \Omega_m$. Deriving the new analytic criterion is not immediate. It requires to derive the analytic expression that describes the behaviour of the last term in the recursive step. This fundamental equation is provided in section 6 and it is the equation (32). The analytic derivation of this equation allows us to detect the key quantity that governs the convergence of algorithm 3, in particular regarding the contribution due to the last term in the recursive step. This key quantity is the following scalar:

$$\tau = \frac{\mathcal{L}_g^2 h}{(L_g^1)^2} \tag{3}$$

We prove (see lemma 10 in section 6) that, in general, it exists m' such that $\mathcal{D}\tau \in \Omega_{m'}$ (and therefore $\mathcal{D}\tau \in \Omega_m \ \forall m \geq m'$). Additionally, we prove that the convergence of the algorithm has been reached when $\Omega_{m+1} = \Omega_m$, $m \geq m'$ and $m \geq 2$ (theorem 2). We also prove that the required number of steps is at most $n + 2$.

In section 6.2.1 it is also shown that the computed codistribution is the entire observable codistribution. Also in this case, the proof is given by first considering the case of a single known input (see theorem 1) and then, its validity is extended to the case of multiple inputs in section 6.2.3. Note that this proof is based on the assumption that the unknown input (w) is a differentiable function of time, up to a given order (the order depends on the specific case).

Algorithm 3 differs from the standard algorithm 1 because of the following reasons:

- In the recursive step, the vector field that corresponds to the unknown input (i.e., g) must be rescaled by dividing by L_g^1.
- The recursive step also contains the sum of the contributions $\sum_{i=1}^{m_u} \mathcal{L}_{^i\phi_{m-1}}\mathcal{D}h$. In other words, we need to compute the Lie derivatives of the differential of the output along the vector fields obtained through the recursive algorithm 2.
- The convergence of algorithm 3 is achieved in at most $n + 2$ steps, instead of $n - 1$ steps (in the special case dealt by lemma 6, this upper bound is $n - 1$ for both cases).
- When $\Omega_m = \Omega_{m-1}$ algorithm 1 has converged. For algorithm 3, we also need to check that $\mathcal{D}\tau \in \Omega_m$ and $m \geq 2$ (with the exception of the special case dealt by lemma 6).

[5]The convergence criterion of algorithm 1 is a consequence of the fact that, all the terms that appear at the recursive step of algorithm 1, are the Lie derivative of the codistribution at the previous step, along fixed vector fields (i.e., vector fields that remain the same at each step of the algorithm). This is not the case for the last term at the recursive step of algorithm 3.

4 The Analytic Criterion

In this section, we outline all the steps to investigate the weak local observability at a given point x_0 of a nonlinear system characterized by (1). Basically, these steps are the steps necessary to compute the observable codistribution (i.e., the steps of algorithm 3) and to prove that the differential of a given state component belongs to this codistribution.

Note that, in the trivial case analyzed by lemma 6, the criterion provided below simplifies, since we do not need to compute the quantity $\tau \left(= \frac{\mathcal{L}_g^2 h}{(\mathcal{L}_g^1 h)^2}\right)$, and we do not need to check that its differential belongs to the codistribution computed at every step of algorithm 3. In practice, we skip the steps 4 and 5 in the procedure below.

1. For the chosen x_0, compute $L_g^1 (= \mathcal{L}_g^1 h)$. In the case when $L_g^1 = 0$, choose another function in the space of functions $\mathcal{F}$ (defined as the space that contains h and its Lie derivative up to any order along the vector fields $f^1, \cdots, f^{m_u}$) such that its Lie derivative along g does not vanish[6].

2. Compute the codistribution Ω_0 and Ω_1 (at x_0) by using algorithm 3.

3. Compute the vector fields ${}^i\phi_m$ $(i = 1, \cdots, m_u)$ by using algorithm 2, starting from $m = 0$, to check if the considered system is in the special case dealt by lemma 6[7]. In this trivial case, set $m' = 0$, use the recursive step of algorithm 3 to build the codistribution Ω_m for $m \geq 2$, and skip to step 6.

4. Compute $\tau \left(= \frac{\mathcal{L}_g^2 h}{(L_g^1)^2}\right)$ and $\mathcal{D}\tau$.

5. Use the recursive step of algorithm 3 to build the codistribution Ω_m for $m \geq 2$, and, for each m, check if $\mathcal{D}\tau \in \Omega_m$. Denote by m' the smallest m such that $\mathcal{D}\tau \in \Omega_m$.

6. For each $m \geq m'$, check if $\Omega_{m+1} = \Omega_m$ and denote by $\Omega^* = \Omega_{m^*}$, where m^* is the smallest integer such that $m^* \geq m'$ and $\Omega_{m^*+1} = \Omega_{m^*}$ (note that $m^* \leq n + 2$).

7. If the differential of a given state component $(x_j,\ j = 1, \cdots, n)$ belongs to Ω^* (namely if $\mathcal{D}x_j \in \Omega^*$) on a given neighbourhood of x_0, then x_j is weakly locally observable at x_0. If this holds for all the state components, the state x is weakly locally observable at x_0. Finally, if the dimension of Ω^* is smaller than n on a given neighbourhood of x_0, then the state is not weakly locally observable at x_0.

[6] If the Lie derivative of any function in $\mathcal{F}$ vanishes, it means that the unknown input can be ignored to obtain the observability properties, as it is shown in appendix A.

[7] Note that this check requires at most $n - 1$ (simple) operations.

5 Applications

We apply the criterion described in section 4 in order to investigate the observability properties of several nonlinear systems characterized by the equations given in (1). These examples will be discussed in sections 5.1, 5.2 and 5.3. The dynamics of the first two examples are the ones of the unicycle. In accordance with the unicycle dynamics, the motion is powered by two independent controls, which are the linear and the angular speed, respectively. In section 5.1 we consider the case when one of these two inputs is unknown and acts as an unknown input. In section 5.2 we consider the case when both these two inputs are known. However, the dynamics are also affected by an external unknown input. Finally, the last example regards a vehicle that moves in $3D$ (section 5.3).

These examples are deliberately uncomplicated in order to allow us to compare the analytic results provided by the proposed criterion with what we can expect by following intuitive reasoning.

Note that the analytic criterion is very powerful and can be used to automatically obtain the observability properties of any system that satisfies (1), i.e., independently of the state dimension (intuitive reasoning becomes often prohibitive for systems characterized by high-dimensional states).

This section also contains extensive simulations and introduces a simple estimator. The results obtained by estimating the state with this estimator fully agree with the observability analysis carried out by using the proposed analytic criterion. These results also show that the analytic criterion is not only a criterion to check the state observability but also a fundamental tool to design an appropriate estimator for a given nonlinear system in presence of a disturbance (see section 7 for a discussion about this fundamental aspect).

5.1 Unicycle with one input unknown

5.1.1 The system

We consider a vehicle that moves on a $2D$-environment. The configuration of the vehicle in a global reference frame, can be characterized through the vector $[x_v,\ y_v,\ \theta]^T$ where x_v and y_v are the Cartesian vehicle coordinates, and θ is the vehicle orientation. We assume that the dynamics of this vector satisfy the unicycle differential equations:

$$\begin{cases} \dot{x}_v &= v\cos\theta \\ \dot{y}_v &= v\sin\theta \\ \dot{\theta} &= \omega \end{cases} \tag{4}$$

where v and ω are the linear and the rotational vehicle speed, respectively, and they are the system inputs. We consider the following three cases of output (see also figure 1 for an illustration):

1. the distance (r) from the origin (e.g., a landmark is at the origin and its distance is measured by a range sensor);

2. the bearing angle (β) of the origin in the local frame (e.g., a landmark is at the origin and its bearing angle is measured by an on-board camera);

3. the bearing angle (ϕ) of the vehicle in the global frame (e.g., a camera is placed at the origin).

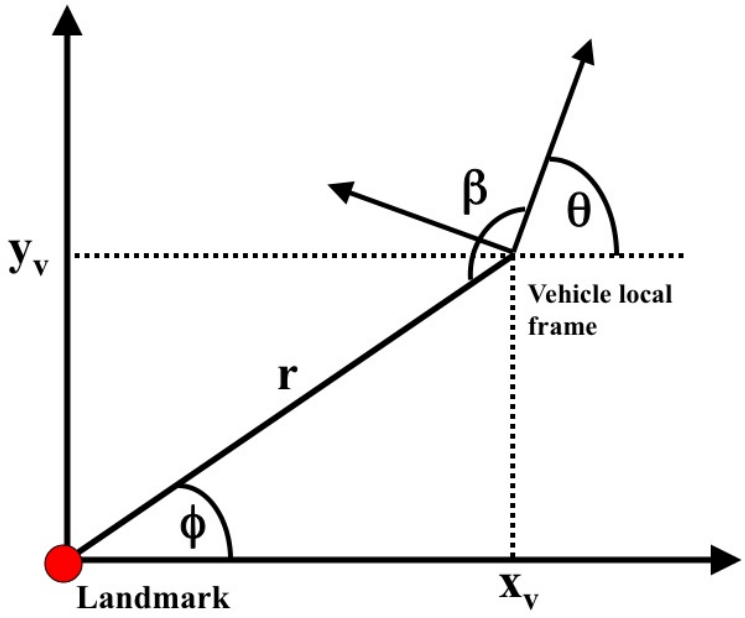


Figure 1: The vehicle state in cartesian and polar coordinates ($[x_v, y_v, \theta]^T$ and $[r, \phi, \theta]^T$, respectively) together with the three considered outputs (r, β and ϕ).

We can analytically express the output in terms of the state. We remark that the expressions become very simple if we adopt polar coordinates: $r \triangleq \sqrt{x_v^2 + y_v^2}$, $\phi = atan\frac{y_v}{x_v}$. We have, for the three cases, $y = r$, $y = \pi - (\theta - \phi)$ and $y = \phi$, respectively. For each of these three cases, we consider the following two cases: v is known, ω is unknown; v is unknown, ω is known. Hence, we have six cases. The dynamics of the unicycle in polar coordinates become:

$$\begin{cases} \dot{r} = v\cos(\theta - \phi) \\ \dot{\phi} = \dfrac{v}{r}\sin(\theta - \phi) \\ \dot{\theta} = \omega \end{cases} \tag{5}$$

5.1.2 Intuitive procedure to obtain the observability properties

By using the observability rank condition in [8], we easily obtain that, when both the inputs are known, the dimension of the observable codistribution is 2 for the first two observations ($y = r$ and $y = \theta - \phi$) and 3 for the last one ($y = \phi$). In particular, for the first two observations all the initial states rotated around the vertical axis are indistinguishable. When one of the inputs misses, this unobservable degree of freedom obviously remains. On the other hand, when the linear speed is unknown (i.e., it acts as an unknown input ($w = v$)) and the observation is an angle (second and third observation, i.e., $y = \theta - \phi$ and $y = \phi$, respectively), we lose a further degree of freedom, which corresponds to the absolute scale. In table 1 we provide the dimension of the observable codistribuion obtained by following this intuitive reasoning for the six considered cases.

Case	Dimension of Ω
1^{st}: $y = r$, $u = \omega$, $w = v$	2
2^{nd}: $y = r$, $w = \omega$, $u = v$	2
3^{d}: $y = \theta - \phi$, $u = \omega$, $w = v$	1
4^{th}: $y = \theta - \phi$, $w = \omega$, $u = v$	2
5^{th}: $y = \phi$, $u = \omega$, $w = v$	2
6^{th}: $y = \phi$, $w = \omega$, $u = v$	3

Table 1: Dimension of the observable codistribution (Ω) obtained by following intuitive reasoning

5.1.3 Analytic results

We now derive the observability properties by applying the analytic criterion described in section 4. For all the cases we have $m_u = 1$. Hence, we adopt the following notation: $f \triangleq f^1$ (for the vector field in (1)) and $\phi_m \triangleq {}^1\phi_m$ (for the vectors defined by algorithm 2). We consider the six cases defined in section 5.1, separately.

First Case: $y = r$**,** $u = \omega$**,** $w = v$

We have:

$$f = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad g = \begin{bmatrix} \cos(\theta - \phi) \\ \frac{\sin(\theta-\phi)}{r} \\ 0 \end{bmatrix}$$

We apply the analytic criterion in section 4. We obtain:

Step 1

We have $L_g^1 = \cos(\theta - \phi)$, which does not vanish, in general.

Step 2

We have: $\Omega_0 = span\{[1, 0, 0]\}$. Additionally, $\Omega_1 = \Omega_0$.

Step 3

We have $\mathcal{L}_{\phi_0} L_g^1 = \mathcal{L}_f L_g^1 = -\sin(\theta - \phi)$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau \triangleq \frac{L_g^2}{(L_g^1)^2} = \frac{\tan^2(\theta-\phi)}{r}$ and

$$\mathcal{D}\tau = \frac{\tan(\theta - \phi)}{r} \left[-\frac{\tan(\theta - \phi)}{r}, -\frac{2}{\cos^2(\theta - \phi)}, \frac{2}{\cos^2(\theta - \phi)} \right]$$

Step 5

We need to compute Ω_2 and, in order to do this, we need to compute ϕ_1. We obtain:

$$\phi_1 = \begin{bmatrix} -\tan(\theta - \phi) \\ \frac{1}{r} \\ 0 \end{bmatrix}$$

and

$$\Omega_2 = span\left\{[1,0,0],\ \left[0, \frac{1}{\cos^2(\theta-\phi)}, -\frac{1}{\cos^2(\theta-\phi)}\right]\right\}$$

It is immediate to check that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$

Step 7

The dimension of the observable codistribution is 2. We conclude that the state is not weakly locally observable. This result agrees with the one in table 1 (second line).

Second Case: $y = r$**,** $u = v$**,** $w = \omega$

We have:

$$f = \begin{bmatrix} \cos(\theta-\phi) \\ \frac{\sin(\theta-\phi)}{r} \\ 0 \end{bmatrix} \quad g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We apply the analytic criterion in section 4. We obtain:

Step 1

We easily obtain $\mathcal{L}_g h = 0$. We consider the function $\mathcal{L}_f h \in \mathcal{F}$. We have: $\mathcal{L}_f h = \cos(\theta-\phi)$ and $\mathcal{L}_g\mathcal{L}_f h = -\sin(\theta-\phi)$, which does not vanish, in general. Hence, we can proceed with the steps in section 4 by setting:

$$h = \cos(\theta-\phi)$$

We obtain: $L_g^1 = -\sin(\theta-\phi)$

Step 2

We have:

$$\Omega_0 = span\{[1,0,0],\ [0,\sin(\theta-\phi),\ -\sin(\theta-\phi)]\}$$

as long as the function r is also a system output. Additionally, $\Omega_1 = \Omega_0$.

Step 3

We have $\mathcal{L}_{\phi_0} L_g^1 = \mathcal{L}_f L_g^1 = \frac{\sin(\theta-\phi)\cos(\theta-\phi)}{r}$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have:

$$\tau = -\frac{\cos(\theta-\phi)}{\sin^2(\theta-\phi)}$$

Step 5

By a direct computation we obtain $\Omega_2 = \Omega_1$. Additionally, it is immediate to check that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$.

Step 7

The dimension of the observable codistribution is 2. We conclude that the state is not weakly locally observable. This result agrees with the one in table 1 (third line).

Third Case: $y = \theta - \phi$**,** $u = \omega$**,** $w = v$ We have:

$$f = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad g = \begin{bmatrix} \cos(\theta - \phi) \\ \frac{\sin(\theta - \phi)}{r} \\ 0 \end{bmatrix}$$

We apply the analytic criterion in section 4. We obtain:

Step 1

We have $L_g^1 = -\frac{\sin(\theta - \phi)}{r}$, which does not vanish, in general.

Step 2

We have $\Omega_0 = span\{[0, -1, 1]\}$ and $\Omega_1 = \Omega_0$.

Step 3

We have $\mathcal{L}_{\phi_0} L_g^1 = \mathcal{L}_f L_g^1 = -\frac{\cos(\theta - \phi)}{r}$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = 2\cot(\theta - \phi)$ and

$$\mathcal{D}\tau = \frac{2}{\sin^2(\theta - \phi)}\,[0, 1, -1]$$

Step 5

By a direct computation we obtain $\Omega_2 = \Omega_1$. Additionally, it is immediate to check that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$.

Step 7

The dimension of the observable codistribution is 1. We conclude that the state is not weakly locally observable. This result agrees with the one in table 1 (fourth line). Note that, the new unobservable direction with respect to the case when both inputs are known, is precisely the absolute scale, since the vector $\mathcal{D}r = [1, 0, 0] \notin \Omega^*$.

Fourth Case: $y = \theta - \phi$**,** $u = v$**,** $w = \omega$ We have:

$$f = \begin{bmatrix} \cos(\theta - \phi) \\ \frac{\sin(\theta - \phi)}{r} \\ 0 \end{bmatrix} \quad g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We apply the analytic criterion in section 4. We obtain:

Step 1

We have $L_g^1 = 1 \neq 0$.

Step 2

By a direct computation we obtain: $\Omega_0 = span\{[0, -1, 1]\}$ and

$$\Omega_1 = span\left\{[0,-1,1], \left[-\frac{\sin(\theta-\phi)}{r^2}, -\frac{\cos(\theta-\phi)}{r}, \frac{\cos(\theta-\phi)}{r}\right]\right\}$$

Step 3

Since $L_g^1 = 1$, it is immediate to realize that $\mathcal{L}_{\phi_j} L_g^1 = 0$, for any integer $j \geq 0$. Hence, the considered system is in the special case considered by lemma 6 and we skip to step 6 by setting $m' = 0$.

Step 6

By a direct computation, it is possible to check that $\Omega_2 = \Omega_1$ meaning that $m^* = 1$ and $\Omega^* = \Omega_1$.

Step 7

The dimension of the observable codistribution is 2. We conclude that the state is not weakly locally observable. This result agrees with the one in table 1 (fifth line).

Fifth Case: $y = \phi$**,** $u = \omega$**,** $w = v$ We have:

$$f = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad g = \begin{bmatrix} \cos(\theta-\phi) \\ \frac{\sin(\theta-\phi)}{r} \\ 0 \end{bmatrix}$$

We apply the analytic criterion in section 4. We obtain:

Step 1

We have $L_g^1 = \frac{\sin(\theta-\phi)}{r}$, which does not vanish, in general.

Step 2

We easily obtain $\Omega_0 = span\{[0,1,0]\}$ and $\Omega_1 = \Omega_0$.

Step 3

We have $\mathcal{L}_{\phi_0} L_g^1 = \mathcal{L}_f L_g^1 = \frac{\cos(\theta-\phi)}{r}$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = -2\cot(\theta-\phi)$ and

$$\mathcal{D}\tau = \frac{2}{\sin^2(\theta-\phi)}[0,-1,1]$$

Step 5

To compute Ω_2 we need to compute ϕ_1. We obtain:

$$\phi_1 = \begin{bmatrix} -r \\ \cot(\theta-\phi) \\ 0 \end{bmatrix}$$

and

$$\Omega_2 = span\left\{[0,1,0], \ \frac{1}{\sin^2(\theta-\phi)}[0,1,-1]\right\}$$

It is immediate to check that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation we obtain $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$, whose dimension is 2.

Step 7

The dimension of the observable codistribution is 2. We conclude that the state is not weakly locally observable. This result agrees with the one in table 1 (sixth line). Note that, the new unobservable direction with respect to the case when both inputs are known, is precisely the absolute scale, since the vector $\mathcal{D}r = [1, 0, 0] \notin \Omega^*$.

Sixth Case: $y = \phi$**,** $u = v$**,** $w = \omega$ We have:

$$f = \begin{bmatrix} \cos(\theta - \phi) \\ \frac{\sin(\theta-\phi)}{r} \\ 0 \end{bmatrix} \quad g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We apply the analytic criterion in section 4. We obtain:

Step 1

We easily obtain $\mathcal{L}_g h = 0$. We consider the function $\mathcal{L}_f h \in \mathcal{F}$. We have: $\mathcal{L}_f h = \frac{\sin(\theta-\phi)}{r}$ and $\mathcal{L}_g \mathcal{L}_f h = \frac{\cos(\theta-\phi)}{r}$, which does not vanish, in general. Hence, we can proceed with the steps in section 4 by setting:

$$h = \frac{\sin(\theta - \phi)}{r} \qquad L_g^1 = \frac{\cos(\theta - \phi)}{r}$$

Step 2

We have:

$$\Omega_0 = span \left\{ [0, 1, 0], \ \left[-\frac{\sin(\theta - \phi)}{r^2}, -\frac{\cos(\theta - \phi)}{r}, \frac{\cos(\theta - \phi)}{r} \right] \right\}$$

as long as the function ϕ is also a system output. We compute Ω_1. By a direct computation, we obtain that its dimension is 3. Hence, we do not need to proceed with the remaining steps since we can directly conclude that the entire state is weakly locally observable. This result agrees with the one in table 1 (seventh line).

5.2 Unicycle in presence of an external unknown input

5.2.1 The system

We consider the same vehicle considered in section 5.1 and we adopt the same state $[x_v,\ y_v,\ \theta]^T$ to characterize its position and orientation. We assume that the vehicle motion is also affected by an unknown input that produces an additional (and unknown) robot speed (denoted by w) along a fixed direction (denoted by γ). Hence, the dynamics are characterized by the following differential equations:

$$\begin{bmatrix} \dot{x}_v &= v \cos\theta + w \cos\gamma \\ \dot{y}_v &= v \sin\theta + w \sin\gamma \\ \dot{\theta} &= \omega \end{bmatrix} \tag{6}$$

where v and ω are the linear and the rotational speed, respectively, in absence of the unknown input. We assume that these two speeds are known (we refer to them as to the known inputs), w is unknown (we refer to it as to the unknown input or disturbance) and γ is constant in time. See also figure 2 for an illustration.

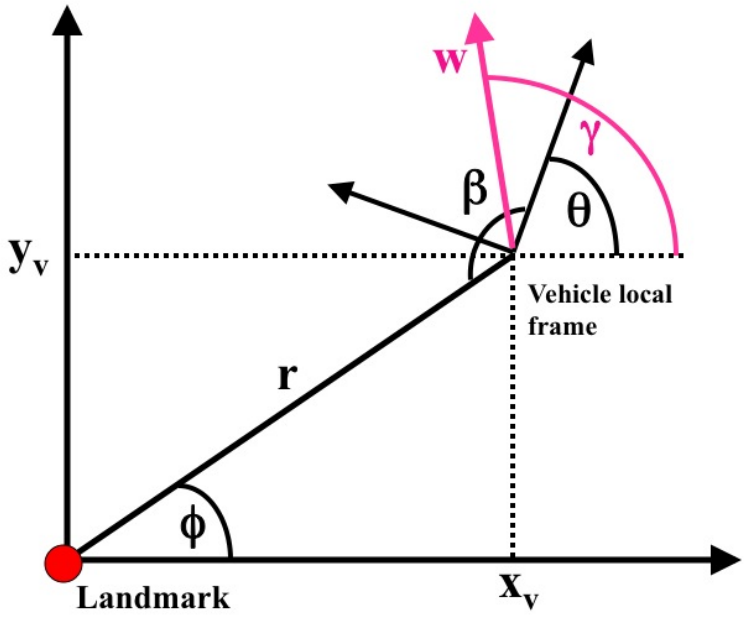


Figure 2: The vehicle state together with the three considered outputs.

We consider the same three cases of output considered in section 5.1. Additionally, we deal with both the cases when γ is known and unknown (in section 5.2.2 and 5.2.3, respectively).

5.2.2 Observability properties when γ is known

The state is $[x_v,\ y_v,\ \theta]^T$ and its dynamics are provided by the three equations in (6), where γ is a known parameter. These equations are a special case of (1). From (6) and (1) we easily obtain: $m_u = 2$, $u_1 = v$, $u_2 = \omega$,

$$f^1 = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix}, \quad f^2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad g = \begin{bmatrix} \cos\gamma \\ \sin\gamma \\ 0 \end{bmatrix}$$

We consider the three outputs separately. For the simplicity sake, we actually consider the following three outputs: $y = r^2 = x_v^2 + y_v^2$ instead of $y = r$, $y = \tan\beta = \frac{y_v - x_v \tan\theta}{x_v + y_v \tan\theta}$ instead of $y = \beta$ and $y = \tan\phi = \frac{y_v}{x_v}$ instead of $y = \phi$. Obviously, the result of the observability analysis does not change.

First Case: $y = r^2$ We apply the analytic criterion in section 4. We obtain:

Step 1

We have $L_g^1 = 2(x_v \cos\gamma + y_v \sin\gamma)$, which does not vanish, in general.

Step 2

We have $\Omega_0 = span\{[x_v, y_v, 0]\}$ and $\Omega_1 = span\{[x_v, y_v, 0], [\cos\theta, \sin\theta, y_v\cos\theta - x_v\sin\theta]\ \}$.

Step 3

We have $\mathcal{L}_{^1\phi_0} L_g^1 = \mathcal{L}_{f^1} L_g^1 = 2\cos(\gamma - \theta)$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = \frac{1}{2(x_v \cos\gamma + y_v \sin\gamma)^2}$ and

$$\mathcal{D}\tau = -\frac{1}{(x_v \cos\gamma + y_v \sin\gamma)^3} [\cos\gamma, \sin\gamma, 0]$$

Step 5

We need to compute Ω_2 and, in order to do this, we need to compute $^1\phi_1$ and $^2\phi_1$ through algorithm 2. We obtain: $^1\phi_1 = ^2\phi_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$. On the other hand, we obtain that $\mathcal{L}_{\frac{g}{L_g^1}} \mathcal{D}\mathcal{L}_{f^1} h \notin \Omega_1$. Hence, by using algorithm 3 we obtain that Ω_2 has dimension equal to 3. As a result, we do not need to proceed with the remaining steps, since we can directly conclude that the entire state is weakly locally observable.

Second Case: $y = \tan\beta$ We apply the analytic criterion in section 4. We obtain:

Step 1

We have $L_g^1 = -\frac{y_v \cos\gamma - x_v \sin\gamma}{x_v^2 \cos^2\theta + 2\sin\theta\cos\theta x_v y_v - y_v^2 \cos^2\theta + y_v^2}$, which does not vanish, in general.

Step 2

By an explicit computation (by using algorithm 3) we obtain that the dimension of Ω_0 is 1 and the dimension of Ω_1 is 2.

Step 3

We have $\mathcal{L}_{^2\phi_0} L_g^1 = \mathcal{L}_{f^2} L_g^1 \neq 0$, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = \frac{x_v \cos\gamma + y_v \sin\gamma + x_v \cos(\gamma - 2\theta) - y_v \sin(\gamma - 2\theta)}{y_v \cos\gamma - x_v \sin\gamma}$

Step 5

We need to compute Ω_2. Also in this case, we obtain that $\mathcal{L}_{\frac{g}{L_g^1}} \mathcal{D}\mathcal{L}_{f^1} h \notin \Omega_1$. Hence, by using algorithm 3 we obtain that Ω_2 has dimension equal to 3. As a result, we do not need to proceed with the remaining steps, since we can directly conclude that the entire state is weakly locally observable.

Third Case: $y = \tan\phi$ We apply the analytic criterion in section 4. We obtain:

Step 1

We have $L_g^1 = -\frac{y_v \cos\gamma - x_v \sin\gamma}{x_v^2}$, which does not vanish, in general.

Step 2

We have $\Omega_0 = span\{[-y_v, x_v, 0]\}$. In addition, by an explicit computation (by using algorithm 3) we obtain that the dimension of Ω_1 is 2.

Step 3

We have $\mathcal{L}_{^2\phi_0} L_g^1 = \mathcal{L}_{f^2} L_g^1 \neq 0$, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = -\frac{2x_v^4 \cos\gamma}{x_v^4 \sin\gamma - x_v^3 y_v \cos\gamma}$

Step 5

We need to compute Ω_2. Also in this case, we obtain that $\mathcal{L}_{\frac{g}{L_g^1}}\mathcal{D}\mathcal{L}_{f^1}h \notin \Omega_1$. Hence, by using algorithm 3 we obtain that Ω_2 has dimension equal to 3. As a result, we do not need to proceed with the remaining steps, since we can directly conclude that the entire state is weakly locally observable.

5.2.3 Observability properties when γ is unknown

The state is $[x_v,\ y_v,\ \theta,\ \gamma]^T$ and its dynamics are provided by the following four equations:

$$\begin{bmatrix} \dot{x}_v &= v\cos\theta + w\cos\gamma \\ \dot{y}_v &= v\sin\theta + w\sin\gamma \\ \dot{\theta} &= \omega \\ \dot{\gamma} &= 0 \end{bmatrix} \tag{7}$$

From (7) and (1) we easily obtain: $m_u = 2$, $u_1 = v$, $u_2 = \omega$,

$$f^1 = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \\ 0 \end{bmatrix}, \quad f^2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad g = \begin{bmatrix} \cos\gamma \\ \sin\gamma \\ 0 \\ 0 \end{bmatrix}$$

We consider the three outputs separately. As in section 5.2.2, we actually consider the following three outputs: $y = r^2 = x_v^2 + y_v^2$ instead of $y = r$, $y = \tan\beta = \frac{y_v - x_v\tan\theta}{x_v + y_v\tan\theta}$ instead of $y = \beta$ and $y = \tan\phi = \frac{y_v}{x_v}$ instead of $y = \phi$. Obviously, the result of the observability analysis does not change.

First Case: $y = r^2$ We apply the analytic criterion in section 4. We obtain:

Step 1

We obviously obtain the same expression as in the case $y = r^2$ of section 5.2.2, i.e., $L_g^1 = 2(x_v\cos\gamma + y_v\sin\gamma)$, which does not vanish, in general.

Step 2

We have $\Omega_0 = span\{[x_v, y_v, 0, 0]\}$ and $\Omega_1 = span\{[x_v, y_v, 0, 0], [\cos\theta, \sin\theta, y_v\cos\theta - x_v\sin\theta, 0]\ \}$.

Step 3

We have $\mathcal{L}_{^1\phi_0}L_g^1 = \mathcal{L}_{f^1}L_g^1 = 2\cos(\gamma - \theta)$, which does not vanish, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = \frac{1}{2(x_v\cos\gamma + y_v\sin\gamma)^2}$, as in the case $y = r^2$ of section 5.2.2. On the other hand, the differential of τ also includes the derivative with respect to γ, namely:

$$\mathcal{D}\tau = \frac{-1}{(x_v\cos\gamma + y_v\sin\gamma)^3}\left[\cos\gamma,\ \sin\gamma,\ 0,\ y_v\cos\gamma - x_v\sin\gamma\right]$$

Step 5

We need to compute Ω_2 and, in order to do this, we need to compute $^1\phi_1$ and $^2\phi_1$ through algorithm 2. We obtain: $^1\phi_1 = {}^2\phi_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$. By using algorithm 3 we compute Ω_2 and we

obtain that its dimension is 3. Additionally, it is possible to verify that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$.

Step 7

We conclude that the dimension of the observable codistribution is equal to $3(< 4)$ and the state is not weakly locally observable. In particular, since the differential of every state component does not belong to Ω_2, we conclude that no state component is observable.

Second Case: $y = \tan\beta$ We apply the analytic criterion in section 4. We obtain:

Step 1

We obviously obtain the same expression as in the case $y = \tan\beta$ of section 5.2.2, i.e., $L_g^1 = -\frac{y_v \cos\gamma - x_v \sin\gamma}{x_v^2 \cos^2\theta + 2\sin\theta\cos\theta x_v y_v - y_v^2 \cos^2\theta + y_v^2}$, which does not vanish, in general.

Step 2

We compute Ω_0 and Ω_1: their dimension are 1 and 2, respectively.

Step 3

We have $\mathcal{L}_{^2\phi_0} L_g^1 = \mathcal{L}_{f^2} L_g^1 \neq 0$, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = \frac{x_v \cos\gamma + y_v \sin\gamma + x_v \cos(\gamma - 2\theta) - y_v \sin(\gamma - 2\theta)}{y_v \cos\gamma - x_v \sin\gamma}$, as in the case $y = \tan\beta$ of section 5.2.2. On the other hand, the differential of τ also includes the derivative with respect to γ.

Step 5

By using algorithm 3 we compute Ω_2 and we obtain that its dimension is 3. Additionally, it is possible to verify that $\mathcal{D}\tau \in \Omega_2$, meaning that $m' = 2$.

Step 6

By a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$.

Step 7

We conclude that the dimension of the observable codistribution is equal to $3(< 4)$ and the state is not weakly locally observable. In particular, since the differential of every state component does not belong to Ω_2, we conclude that no state component is observable.

Third Case: $y = \tan\phi$ We apply the analytic criterion in section 4. We obtain:

Step 1

We obviously obtain the same expression as in the case $y = \tan\phi$ of section 5.2.2, i.e., $L_g^1 = -\frac{y_v \cos\gamma - x_v \sin\gamma}{x_v^2}$, which does not vanish, in general.

Step 2

We compute Ω_0 and Ω_1: their dimension are 1 and 2, respectively.

Step 3

We have $\mathcal{L}_{^2\phi_0} L_g^1 = \mathcal{L}_{f^2} L_g^1 \neq 0$, in general. This suffices to conclude that the considered system is not in the special case considered by lemma 6 and we need to continue with step 4.

Step 4

We have $\tau = -\frac{2x_v^4 \cos\gamma}{x_v^4 \sin\gamma - x_v^3 y_v \cos\gamma}$, as in the case $y = \tan\phi$ of section 5.2.2. On the other hand, the differential of τ also includes the derivative with respect to γ.

Step 5

γ	Output	State observability
known	$y = r$	yes
known	$y = \beta$	yes
known	$y = \phi$	yes
unknown	$y = r$	no
unknown	$y = \beta$	no
unknown	$y = \phi$	yes

Table 2: Weak local observability of the state in all the considered scenarios

By using algorithm 3 we compute Ω_2 and we obtain that its dimension is 4. As a result, we do not need to proceed with the remaining steps, since we can directly conclude that the entire state is weakly locally observable.

Table 2 summarizes the results of the observability analysis carried out in this section. We conclude this section by remarking that these results agree with our expectation. By using the observability rank condition in [8], we easily obtain that, in absence of the unknown input, the dimension of the observable codistribution is 2 for the first two observations ($y = r$ and $y = \beta$) and 3 for the last one ($y = \phi$). In particular, for the first two observations, all the initial states rotated around the vertical axis are indistinguishable. In other words, in these two cases, the system exhibits a continuous symmetry [24]. In presence of the unknown input, when γ is known, the aforementioned system invariance is broken and the entire state becomes observable. When γ is unknown, the symmetry still remains (and obviously also concerns the new state component γ). Note also a very important aspect. The presence of an unknown input improves the observability properties of a system (this regards the case when γ is known). In particular, if $w = 0$ (absence of unknown input), the state becomes unobservable despite the knowledge of the unknown input (we know that it is zero), while, when $w \neq 0$, the state is observable even if w is unknown. Note that having an unknown input equal to zero is an event that occurs with zero probability and our theory accounts this fact since it is based on definition 1. To this regard, note also that the validity of theorem 1, which allows us to introduce the algorithms 2 and 3, holds when the unknown input is different from 0.

5.2.4 Simulations

The scope of this section is to show that, by using a very simple estimator based on an Extended Kalman Filter, we obtain results which agree with our observability analysis.

Simulated trajectories and robot sensors

The trajectories are simulated as follows. The equations in (6) are discretized with a time step of $5 * 10^{-4}$ s and they last for 10 s. The linear speed, i.e. v, is constant and equal to 0.1 $m\ s^{-1}$. The angular speed, i.e. ω, is generated randomly. Specifically, its value is settled at each step and follows a Gaussian distribution with mean value $\frac{2\pi}{10}\ rad\ s^{-1}$ and variance $\left(\frac{6\pi}{10}\right)^2\ rad^2\ s^{-2}$. A typical example of trajectory, obtained with this setting, is displayed in fig. 3, left side. The disturbance is generated as follows. The parameter γ is set equal to 0.7 π. The unknown input w is generated randomly. Specifically, its value is settled at each step and follows a Gaussian distribution with mean value 0.05 $m\ s^{-1}$ and variance $(0.005)^2\ m^2\ s^{-2}$. A typical example of trajectory is displayed in fig. 3, right side.

The robot is equipped with proprioceptive sensors able to measure at each time step the linear and

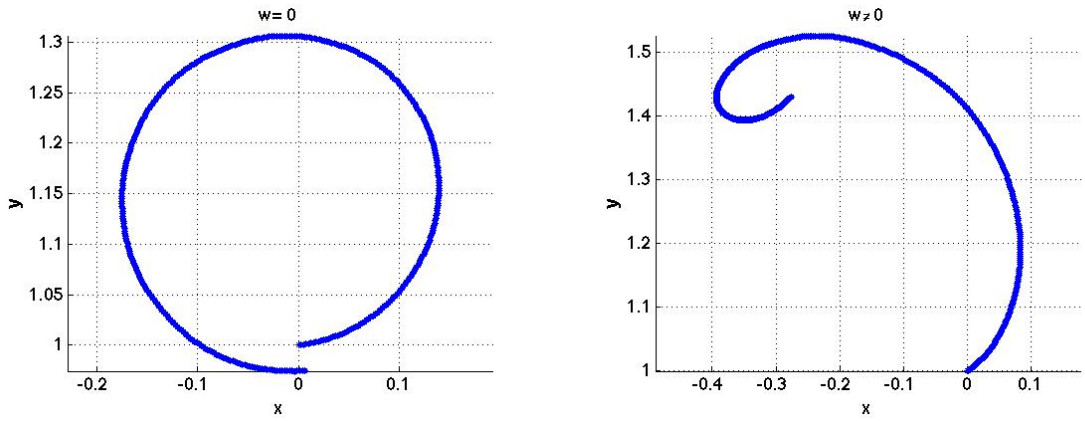


Figure 3: Typical simulated trajectories. Left side without disturbance, right side with disturbance.

the angular speed. These measurements are affected by errors. Specifically, each measurement is generated at every time step of $5 * 10^{-4}$ s by adding to the true value a random error that follows a Gaussian distribution. The mean value of this error is zero and the standard deviation is 0.01 times the true value for both the linear and the angular speed.
Regarding the exteroceptive measurements, they are generated at a lower frequency. Specifically, the measurements are generated each $2.5*10^{-2}$ s. Also these measurements are affected by errors. Specifically, the measurement is generated by adding to the true value a random error that follows a Gaussian distribution. The mean value of this error is zero and the standard deviation is 0.01 m for the range measurements and 1 deg for the two angular measurements (bearing of the origin in the robot frame and bearing of the robot in the global frame).

Estimation results

We adopt an Extended Kalman Filter that estimates an extended state that includes the unknown input together with its first order time derivative. In other words, in the case when γ is known, the estimated state is: $[x_v,\ y_v,\ \theta,\ w,\ w^{(1)}]$. Its dynamics are obtained from (6) and are:

$$\begin{bmatrix} \dot{x}_v & = v\cos\theta + w\cos\gamma \\ \dot{y}_v & = v\sin\theta + w\sin\gamma \\ \dot{\theta} & = \omega \\ \dot{w} & = w^{(1)} \\ \dot{w}^{(1)} & = w^{(2)} \end{bmatrix} \tag{8}$$

When γ is unknown, the estimated state also includes γ and the dynamics are given by (8) and the additional equation $\dot{\gamma} = 0$.

In order to implement the prediction phase of our Extended Kalman Filter we have to provide the value of $w^{(2)}$, which is unknown. We set this quantity to zero. Note that the simulated trajectory does not satisfy this hypothesis since the disturbance is randomly generated. However, the estimator is able to provide good performance as it is shown in figures 4-6.

Figure 4 displays the estimated trajectory in the case when the output is provided by the range sensor ($h = r$). Dots blue are the ground truth, red circles the estimated trajectory by

only using the knowledge of the proprioceptive measurements (i.e., the measurements of v and ω), black stars the trajectory estimated by our estimator. Left side is the case when γ is known and right side when it is unknown. In accordance with our observability analysis, the estimator follows the true trajectory only in the case when γ is known. Figure 5 displays the estimated trajectory in the case when the output is provided by the on-board bearing sensor ($h = \beta$). Also in this case, the estimator follows the true trajectory only in the case when γ is known, in accordance with our observability analysis. Finally, figure 6 displays the estimated trajectory in the case when the output is provided by the bearing sensor at the origin ($h = \phi$). In this case, the estimator follows the true trajectory both in the case when γ is known and when it is unknown. Also this result agrees with our observability analysis.

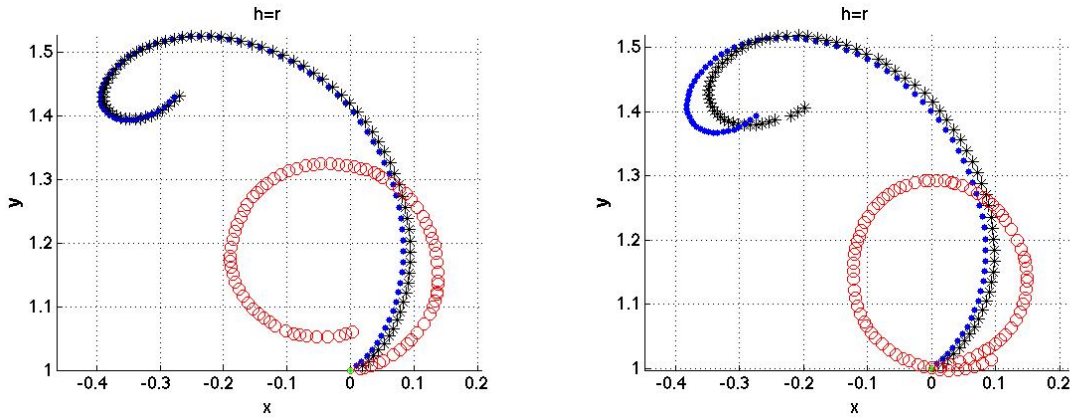


Figure 4: Estimated trajectory in the case when the output is $h = r$. Dots blue are the ground truth, red circles the estimated trajectory by only using the knowledge of the proprioceptive measurements, black stars the trajectory estimated by our estimator. On the left side is the case when γ is known and on the right side when it is unknown.

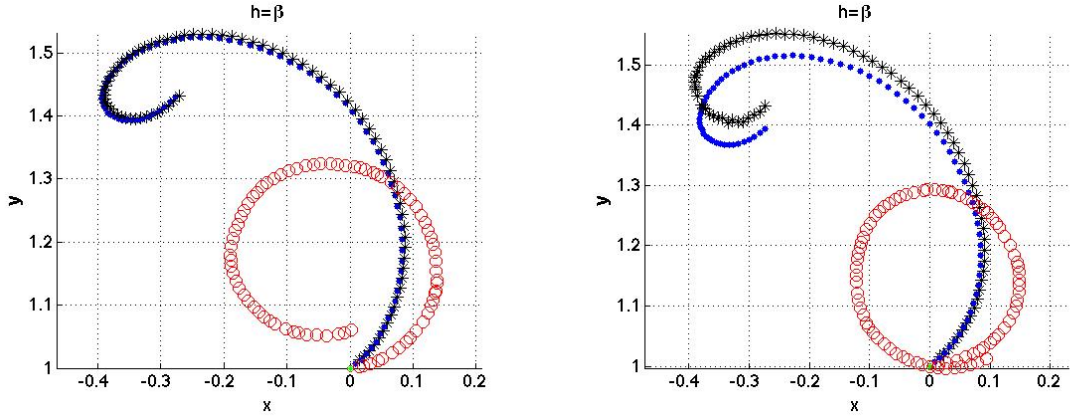


Figure 5: The same as in figure 4 in the case when the output is $h = \beta$.

We conclude this section by remarking that the results of our simulations fully agree with the observability analysis provided in sections 5.2.2 and 5.2.3. Note that we showed the result of a

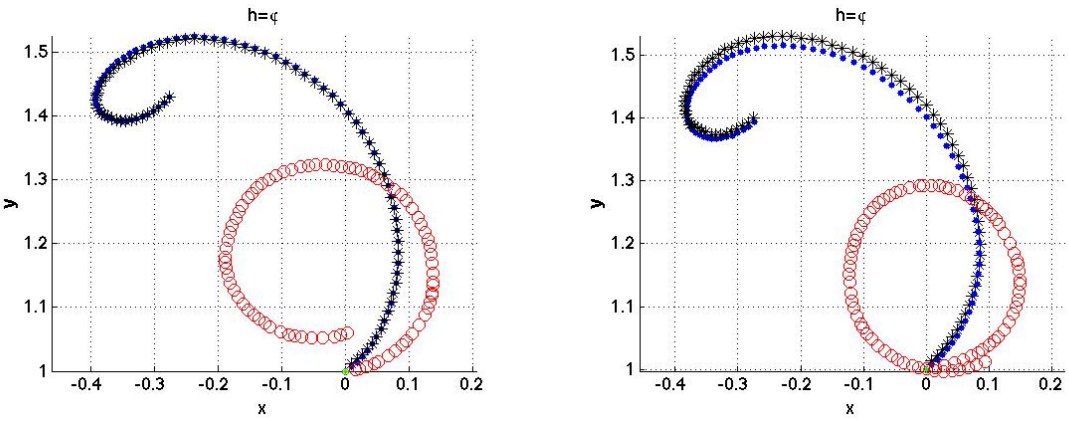


Figure 6: The same as in figure 4 in the case when the output is $h = \phi$.

single trial. However, we found always similar results by running many trials.

5.3 Vehicle moving in 3D in presence of a disturbance

5.3.1 The system

We consider a vehicle that moves in a $3D-$environment. We assume that the dynamics of the vehicle are affected by the presence of a disturbance (e.g., this could be an aerial vehicle in presence of wind). We assume that the direction of the disturbance is constant in time and a priori known. Conversely, the disturbance magnitude is unknown and time dependent. The vehicle is equipped with speed sensors (e.g., airspeed sensors in the case of an aerial vehicle), gyroscopes and a bearing sensor (e.g., monocular camera). We assume that all the sensors share the same frame (in other words, they are extrinsically calibrated). Without loss of generality, we define the vehicle local frame as this common frame. The airspeed sensors measure the vehicle speed with respect to the air in the local frame. The gyroscopes provide the angular speed in the local frame. Finally, the bearing sensor provides the bearing angles of the features in the environment expressed in its own local frame. We consider the extreme case of a single point feature and, without loss of generality, we set the origin of the global frame at this point feature (see figure 7 for an illustration). Additionally, we assume that the $z-$axis of the global frame is aligned with the direction of the disturbance.
Our system can be characterized by the following state:

$$X \triangleq [x,\ y,\ z,\ q_t,\ q_x,\ q_y,\ q_z]^T \tag{9}$$

where $r = [x,\ y,\ z]$ is the position of the vehicle in the global frame and $q = q_t + q_x i + q_y j + q_z k$ is the unit quaternion that describes the transformation change between the global and the local frame. The dynamics are affected by the presence of the disturbance. The disturbance is characterized by the following vector (in the global frame):

$$\bar{w} = w \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \tag{10}$$

where w is its unknown magnitude.

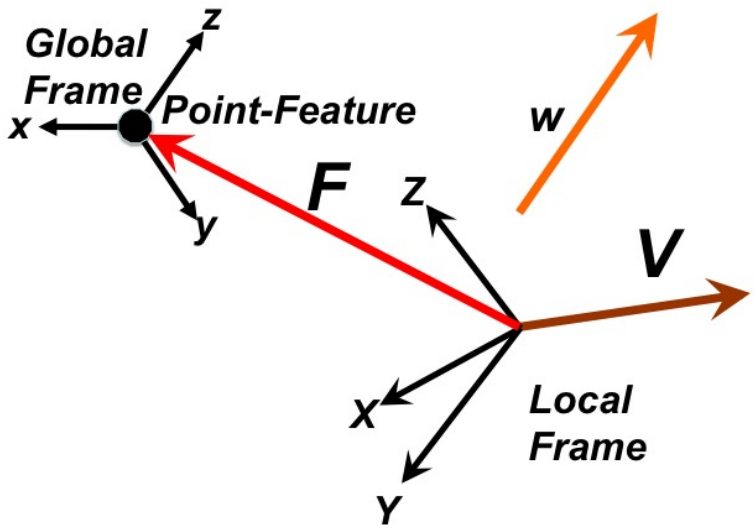


Figure 7: Local and global frame for the considered problem. The $z-$axis of the latter is aligned with the direction of the disturbance (assumed to be known and constant in time). The speed V is the vehicle speed with respect to the air, which differs from the ground speed because of the disturbance (w).

In the sequel, for each vector defined in the $3D$ space, the pedix q will be adopted to denote the corresponding imaginary quaternion. For instance, regarding the vehicle position, we have: $r_q = 0 + x\ i + y\ j + z\ k$. Additionally, we denote by V and Ω the following physical quantities:

- $V = [V_x,\ V_y,\ V_z]$ is the vehicle speed with respect to the air expressed in the local frame (hence, $w\ k + qV_qq^*$ is the vehicle speed with respect to the ground expressed in the global frame).
- $\Omega \triangleq [\Omega_x\ \Omega_y\ \Omega_z]$ is the angular speed (and $\Omega_q = 0 + \Omega_x\ i + \Omega_y\ j + \Omega_z\ k$).

The dynamics of the state are:

$$\begin{bmatrix} \dot{r}_q & = w\ k + qV_qq^* \\ \dot{q} & = \frac{1}{2}q\Omega_q \end{bmatrix} \tag{11}$$

The monocular camera provides the position of the feature in the local frame ($F_q = -q^*r_qq$) up to a scale. Hence, it provides the ratios of the components of F:

$$h_{cam}(X) \triangleq [h_u,\ h_v]^T = \left[\frac{(q^*r_qq)_x}{(q^*r_qq)_z},\ \frac{(q^*r_qq)_y}{(q^*r_qq)_z}\right]^T \tag{12}$$

where the pedices x, y and z indicate respectively the i, j and k component of the corresponding quaternion. We have also to consider the constraint $q^*q = 1$. This provides the further observation:

$$h_{const}(X) \triangleq q^*q \tag{13}$$

Our system is characterized by the state in (9), the dynamics in (11) and the three outputs h_u, h_v and h_{const} in (12) and (13).

5.3.2 Observability in absence of disturbance

Our system is characterized by the state in (9), the dynamics in (11) with $w = 0$ and the three outputs h_u, h_v and h_{const} in (12) and (13).

By comparing (11) with (1) we obtain that our system is characterized by six known inputs ($m_u = 6$) that are: $u_1 = \Omega_x$, $u_2 = \Omega_y$, $u_3 = \Omega_z$, $u_4 = V_x$, $u_5 = V_y$ and $u_6 = V_z$. Additionally, we obtain:

$$f^1 = \frac{1}{2}\begin{bmatrix} 0 \\ 0 \\ 0 \\ -q_x \\ q_t \\ q_z \\ -q_y \end{bmatrix},\ f^2 = \frac{1}{2}\begin{bmatrix} 0 \\ 0 \\ 0 \\ -q_y \\ -q_z \\ q_t \\ q_x \end{bmatrix},\ f^3 = \frac{1}{2}\begin{bmatrix} 0 \\ 0 \\ 0 \\ -q_z \\ q_y \\ -q_x \\ q_t \end{bmatrix},$$

$$f^4 = \begin{bmatrix} q_t^2 + q_x^2 - q_y^2 - q_z^2 \\ 2q_tq_z + 2q_xq_y \\ 2q_xq_z - 2q_tq_y \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},\ f^5 = \begin{bmatrix} 2q_xq_y - 2q_tq_z \\ q_t^2 - q_x^2 + q_y^2 - q_z^2 \\ 2q_tq_x + 2q_yq_z \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

$$f^6 = \begin{bmatrix} 2q_tq_y + 2q_xq_z \\ 2q_yq_z - 2q_tq_x \\ q_t^2 - q_x^2 - q_y^2 + q_z^2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Finally, in absence of disturbance we have:

$$g = [0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0]^T$$

In this case we can apply the observability rank condition, i.e., algorithm 1, to obtain the observable codistribution. We compute the codistribution Ω_0 by computing the differentials of the three functions h_u, h_v and h_{const}. We obtain that this codistribution has dimension equal to 3. We use algorithm 1 to compute Ω_1. We obtain that its dimension is 4. In particular, the additional covector is obtained by the differential of the following Lie derivative:

$$\mathcal{L}_{f^4} h_u$$

In other words:

$$\Omega_1 = span\left\{\mathcal{D}h_u,\ \mathcal{D}h_v,\ \mathcal{D}h_{const},\ \mathcal{D}\mathcal{L}_{f^4}h_u,\right\}$$

All the remaining first order Lie derivatives have differential that is in the above codistribution. Additionally, by an explicit computation, it is easy to realize that $\Omega_2 = \Omega_1$. This means that algorithm 1 has converged and the observable codistribution is Ω_1.
By an explicit computation, it is possible to check that the differentials of the components of the vector F belong to Ω_1. This means that all the observable modes are the components of F, i.e., the position of the feature in the local frame (obviously, the fourth observable mode is the norm of the quaternion). In particular, no component of the vehicle orientation is observable.

5.3.3 Observability in presence of the disturbance

We now consider the case when the dynamics are affected by the presence of the disturbance. By comparing (11) with (1) we obtain that the vector fields that characterize the dynamics are the same that characterize the dynamics in absence of disturbance with the exception of the last one, which becomes:

$$g = [0,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0]^T$$

To derive the observability properties we apply the proposed analytic tool, by following the seven steps provided in section 4.

First Step

We start by computing the Lie derivatives of the outputs h_u, h_v and h_{const} along the vector field g. We find that the result differs from zero for the first two outputs. Hence, we use the first output (h_u) to define L_g^1 (we could choose also the second output h_v). In particular, we obtain: $L_g^1 \triangleq \mathcal{L}_g h_u =$

$$\frac{-y(2q_tq_z - 2q_xq_y) - x(q_t^2 - q_x^2 + q_y^2 - q_z^2)}{[z(q_t^2 - q_x^2 - q_y^2 + q_z^2) + 2x(q_tq_y + q_xq_z) + 2y(q_yq_z - q_tq_x)]^2}$$

Second Step

We compute the codistribution Ω_0 by computing the differentials of the three functions h_u, h_v and h_{const}. This coincides with the case without disturbance, and we obtain that this codistribution has dimension equal to 3.

We use algorithm 3 to compute Ω_1. We obtain that its dimension is 5. In particular, the additional two independent covectors are obtained by the differentials of the following two Lie derivatives:

$$\mathcal{L}_{f^4} h_u, \qquad \mathcal{L}_{\frac{g}{L_g^1}} h_v$$

In other words:

$$\Omega_1 = span\left\{ \mathcal{D}h_u,\ \mathcal{D}h_v,\ \mathcal{D}h_{const},\ \mathcal{D}\mathcal{L}_{f^4} h_u,\ \mathcal{D}\mathcal{L}_{\frac{g}{L_g^1}} h_v \right\}$$

All the remaining first order Lie derivatives have differential that is in the above codistribution.

Third Step

We compute ${}^1\phi_1$, ${}^2\phi_1$, ${}^3\phi_1$, ${}^4\phi_1$, ${}^5\phi_1$ and ${}^6\phi_1$ by using algorithm 2. We obtain that all these vectors vanish. As a result, all the subsequent steps of algorithm 2 provide null vectors. Therefore, the assumptions of lemma 6 are trivially met. We set $m' = 0$ and we skip to the sixth step.

Sixth Step

We use algorithm 3 to compute Ω_2 and we obtain:

$$\Omega_2 = \Omega_1 + span\left\{ \mathcal{D}\mathcal{L}_{f^4}\mathcal{L}_{\frac{g}{L_g^1}} h_v \right\}$$

Hence, its dimension is 6. Finally, by using again algorithm 3 it is possible to compute Ω_3 and to check that $\Omega_3 = \Omega_2$. This means that the algorithm has converged and the observable codistribution is $\Omega^* = \Omega_2$.

Seventh Step

By computing the distribution orthogonal to the codistribution Ω^* we can find the continuous symmetry that characterizes the unobservable space [24]. By an explicit computation we obtain the following vector:

$$\left[-y,\ x,\ 0,\ -\frac{q_z}{2},\ -\frac{q_y}{2},\ \frac{q_x}{2}, \frac{q_t}{2}\right]^T$$

This symmetry corresponds to an invariance with respect to a rotation around the $z-$axis of the global frame. This means that we have a single unobservable mode that is the yaw in the global frame. We conclude by remarking that the presence of the disturbance, even if its magnitude is unknown and is not constant, makes observable the roll and the pitch angles. This result is similar to the result that we obtain in the case of visual and inertial sensor fusion in presence of gravity. The presence of gravity makes observable the roll and the pitch angles, even if its magnitude is unknown [25, 26]. What it is non intuitive in the case now investigated, is that, not only the magnitude of the disturbance is unknown, but it is also time dependent.

5.3.4 Simulations

As in section 5.2.4, the scope of this section is to show that, by using a simple estimator based on an Extended Kalman Filter, we obtain results that agree with our observability analysis.

Simulated trajectories and robot sensors

The trajectories are simulated as follows. The equations in (11) are discretized with a time step of 0.01 s. Each trial lasts 200 s. The initial vehicle speed is set to zero. The initial position is set equal to $[0.5\ 0.5\ 0.5]m$. The vehicle motion is randomly generated. The angular speed, i.e. Ω, is Gaussian. Specifically, its value at each step follows a zero mean Gaussian distribution with covariance matrix equal to $(1\ deg)^2 I_3$, where I_3 is the identity 3×3 matrix. At each time step, the vehicle speed is incremented by adding a random vector with zero mean Gaussian distribution. In particular, the covariance matrix of this distribution is set equal to $\sigma^2 I_3$, with $\sigma = 10^{-4}m$. Finally, we simulate the disturbance as a vector along the $z-$axis, whose magnitude is generated as a random Gaussian variable, with mean value 0.2 ms^{-1} and variance 0.05^2 m^2s^{-2}.

Typical trajectories, obtained with this setting, are displayed in figures 8 and 9.

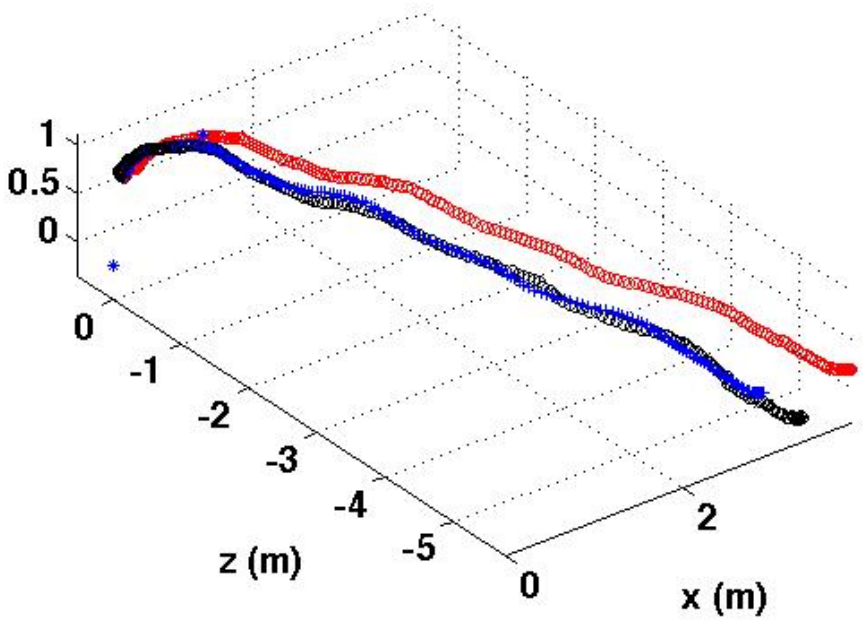
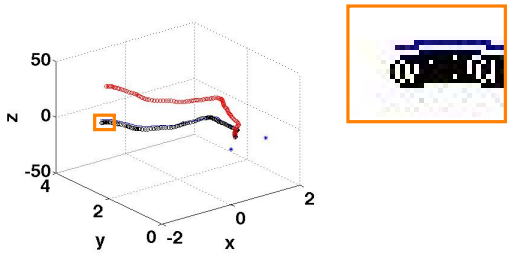


Figure 8: A typical simulated trajectory in absence of disturbance. In blue the true trajectory, in red the trajectory estimated by only integrating the gyroscope and the airspeed measurements and in black the trajectory estimated by the proposed EKF.

The vehicle is equipped with proprioceptive sensors able to measure at each time step the speed with respect to the air and the angular speed. These measurements are affected by errors. Specifically, each measurement is generated at every time step of 0.01 s by adding to the true value a random error that follows a Gaussian distribution. The mean value of this error is zero and the standard deviation is 0.01 ms^{-1} for the airspeed and 1 $deg\ s^{-1}$ for the gyroscope.



(a) (b)

Figure 9: A typical simulated trajectory in presence of disturbance along the $z-$axis. In blue the true trajectory, in red the trajectory estimated by only integrating the gyroscope and the airspeed measurements and in black the trajectory estimated by the proposed EKF.

Regarding the exteroceptive measurements, they are generated at a lower frequency. Specifically, the measurements are generated each 0.1 s. Also these measurements are affected by errors. Specifically, each measurement is generated by adding to the true value a random error that follows a zero mean Gaussian distribution, with variance 1 deg^2. We simulate two point features, one at the origin and the other one at the position $[2\ 0\ 0]m$ (note that the result of the observability analysis is independent of the number of point features).

Estimation results

We adopt an EKF that estimates an extended state that includes the state in (9) and the unknown input (w) together with its first order time derivative ($w^{(1)} \triangleq \frac{dw}{dt}$). In other words, the estimated state is: $\left[x,\ y,\ z,\ q_t,\ q_x,\ q_y,\ q_z,\ w,\ w^{(1)}\right]^T$. Its dynamics are given in (11) with the two further equations:

$$\begin{bmatrix} \dot{w} & = w^{(1)} \\ \dot{w}^{(1)} & = w^{(2)} \end{bmatrix} \tag{14}$$

In order to implement the prediction phase of our EKF we have to provide the value of $w^{(2)}$ ($\triangleq \frac{d^2w}{dt^2}$), which is unknown. We set this quantity to zero. Note that the simulated trajectory does not satisfy this hypothesis since the disturbance is randomly generated. However, the estimator is able to provide good performance as it is shown in figures 10-14.

Figure 8 displays the estimated trajectory in absence of the disturbance. The red line is the trajectory obtained by only using the proprioceptive measurements, while the black line is the trajectory obtained by the proposed EKF. The true trajectory is in blue. Both the estimated trajectories diverge. However, the divergence is slower for the trajectory estimated by the EKF. This divergence is consistent with our observability analysis. The vehicle position is not observable (both with and without disturbance). Figure 9 displays the same of figure 8 for a trial in presence of the disturbance. Note that, in this case, the $z-$axis has now a different scale (much larger). This is to show the trajectory obtained by only using the proprioceptive measurements, whose divergence is much faster due to the presence of the disturbance along the $z-$axis.

In the remaining figures, we show the performance of the proposed EKF in estimating the distances of the two features and the roll, the pitch and the yaw angles. In accordance with our observability analysis, the distances of the features are observable both with and without the disturbance. In contrast, the roll and the pitch angles are observable only in presence of

the disturbance. Finally, the yaw angle is always unobservable. All these results are clearly confirmed by our simulation results.

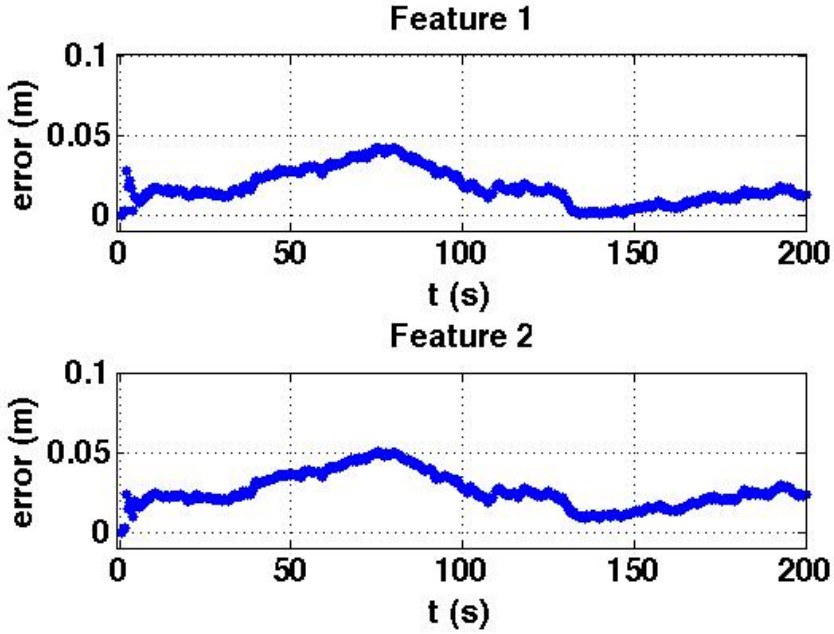


Figure 10: Estimation error on the distance of the two observed point features in absence of the disturbance.

Figures 10 and 11 show the error on the two distances, in absence and in presence of the disturbance, respectively. We note that, the error does not increase with time in both cases.

Figures 12, 13 and 14 show the error on the roll, pitch and yaw angles, respectively. The left side is in absence of the disturbance while the right side is in presence of the disturbance. In absence of the disturbance, the error increases with time for all the three angles. In presence of the disturbance, it does not increase for the roll and the pitch angle.
We conclude this section by remarking that the results of our simulations fully agree with the observability analysis provided in section 5.3.3. Also note that we showed the result of a single trial. However, we found always similar results by running many trials.

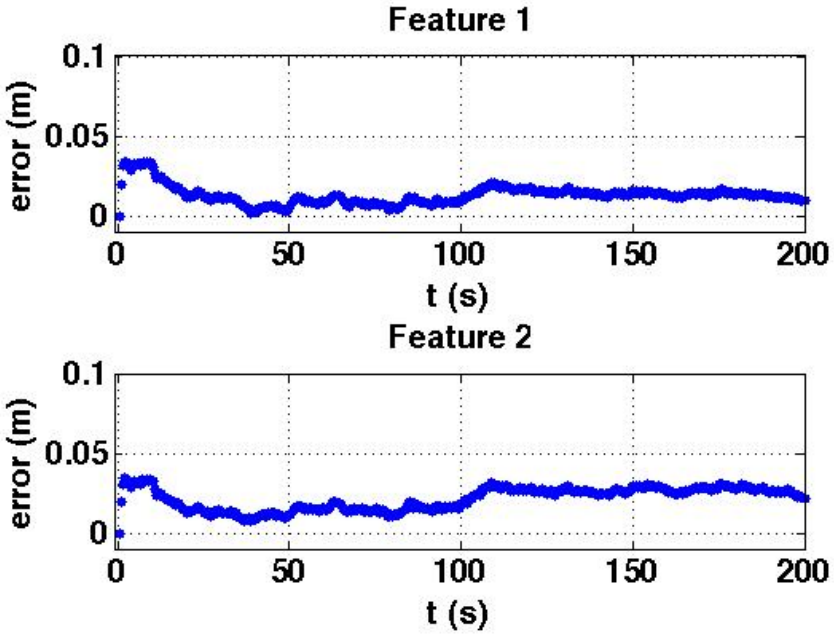


Figure 11: As in fig. 10 but in presence of the disturbance.

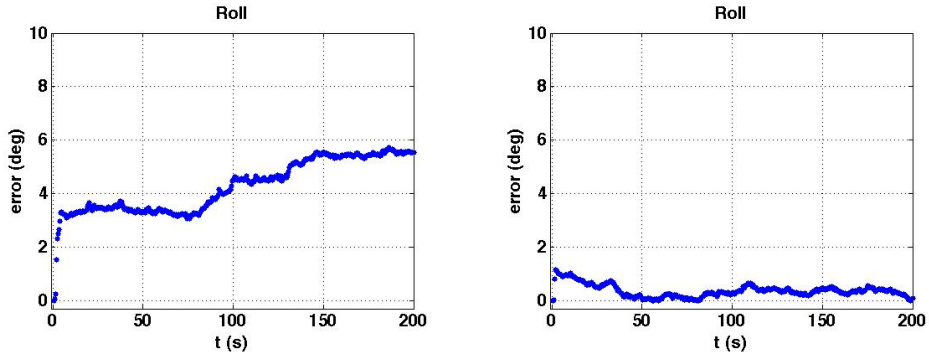


Figure 12: Estimation error on the vehicle roll angle. Left side in absence of the disturbance, right side in presence of the disturbance.

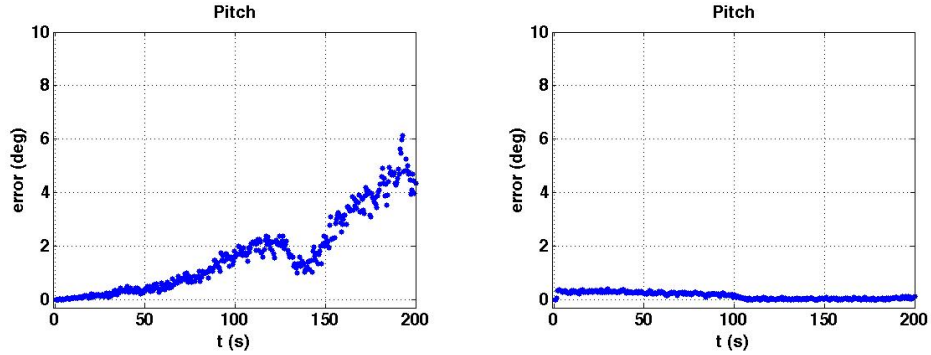


Figure 13: As in fig. 12 but for the pitch angle.

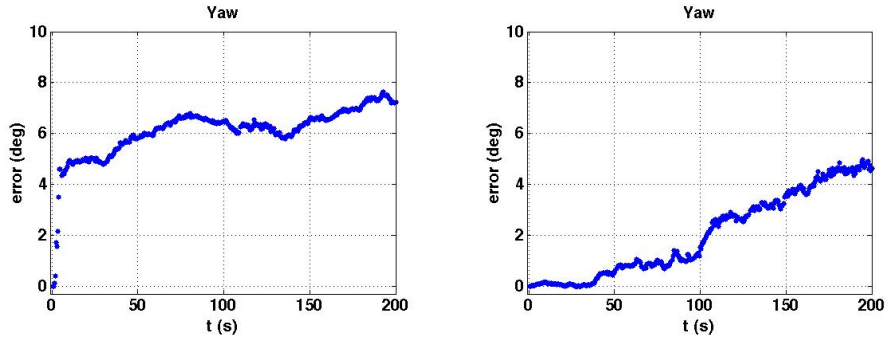


Figure 14: As in fig. 12 but for the yaw angle.

6 Analytic Derivations

In this section, we prove the validity of the analytic results presented in section 3. We start by introducing several important concepts starting by providing the definition of state indistinguishability in presence of unknown inputs (section 6.1). These concepts hold for any number of unknown inputs and when the dynamics are affine in the inputs (and not simply linear). Then, section 6.2 is devoted to prove the results that hold in the case of a single unknown input and therefore to prove the validity of the results provided in sections 3 and 4.

6.1 State Indistinguishability, State augmentation and Observable Codistribution

As mentioned in the introduction of this section, the concepts introduced in this section, hold for any number of unknown inputs (m_w) and when the dynamics are affine in the inputs. Specifically, we will refer to the following dynamics:

$$\dot{x} = g^0(x) + \sum_{i=1}^{m_u} f^i(x)u_i + \sum_{j=1}^{m_w} g^j(x)w_j \tag{15}$$

where $g^0(x)$, $f^i(x)$, $i = 1, \cdots, m_u$, and $g^j(x)$, $j = 1, \cdots, m_w$, are vector fields in $M \subset \mathbb{R}^n$.

In [8, 16] the observability properties of a nonlinear system driven by only known inputs are derived starting from the definition of indistinguishable states. According to this definition, it is proven the following fundamental property:

Proposition 1 *The Lie derivatives of any output computed along any direction allowed by the system dynamics take the same values at the states which are indistinguishable.*

Starting from this fundamental property it is possible to prove that algorithm 1 generates the observable codistribution [16]. In presence of unknown inputs, we first need to introduce a new definition of indistinguishability. To be conservative, the new definition must consider two states indistinguishable if there exists at least one pair of unknown inputs, such that, the outputs obtained starting from the first state under the effect of the first unknown input on a given time interval ($\mathcal{I}$), coincide with the outputs obtained starting from the second state under the effect of the second unknown input on the same time interval, and this holds for any choice of the known inputs. On the other hand, if this condition is achieved only for a unique pair of unknown inputs, the probability that this event occurs is zero. For this reason, we require that this condition is met for infinite pairs of unknown inputs. The concept of measure in spaces with infinite dimensions is not trivial. In particular, there is no analogue of Lebesgue measure on an infinite-dimensional Banach space. One possibility, which is frequently adopted, is to use the concept of prevalent and shy sets [15]. Let us denote by $\mathcal{W}$ the functional space of all the possible unknown input functions. The probability that a given unknown input belongs to a shy subset of $\mathcal{W}$, is 0. The probability that a given unknown input belongs to a prevalent subset of $\mathcal{W}$, is 1. Finally, the probability that a given unknown input belongs to a non-shy subset of $\mathcal{W}$, is strictly larger than 0.

We introduce the following definition:

Definition 1 (Indistinguishable states) *Two states x_a and x_b are indistinguishable if, for every $u(t)$ (the known input vector function), there exists a non-shy subset $\mathcal{W}_a$ in the functional space of all the possible unknown input functions, such that, for any unknown input function $w_a \in$*

$\mathcal{W}_a$ it exists an unknown input function w_b such that $h(x(t;\ x_a;\ u;\ w_a)) = h(x(t;\ x_b;\ u;\ w_b))$ $\forall t \in \mathcal{I}$.

The property stated by proposition 1 does not hold in presence of unknown inputs. Our first objective is to extend the original state in order to obtain a similar property for the resulting extended system. This new property will be the one stated by proposition 4.

To obtain such a result, we proceed as follows. We extend the original state by including the unknown inputs together with their time derivatives. Specifically, we denote by ${}^k x$ the extended state that includes the unknown inputs and their time derivatives up to the $(k-1)-$order:

$$ {}^k x \triangleq [x^T,\ w^T,\ w^{(1)\ T},\ \cdots,\ w^{(k-1)\ T}]^T \tag{16}$$

where $w^{(k)} \triangleq \frac{d^k w}{dt^k}$ and ${}^k x \in M^{(k)}$, with $M^{(k)}$ an open set of $\mathbb{R}^{n+km_w}$. From (15) it is immediate to obtain the dynamics for the extended state:

$$ {}^k\dot{x} = G + \sum_{i=1}^{m_u} F^i u_i + \sum_{j=1}^{m_w} W^j w_j^{(k)} \tag{17}$$

where:

$$ G \triangleq \begin{bmatrix} g^0 + \sum_{j=1}^{m_w} g^j w_j \\ w^{(1)} \\ w^{(2)} \\ \cdots \\ w^{(k-1)} \\ 0_{m_w} \end{bmatrix} \tag{18}$$

$$ F^i \triangleq \begin{bmatrix} f^i \\ 0_{km_w} \end{bmatrix}, \quad W^j \triangleq \begin{bmatrix} 0_{n+(k-1)m_w+j-1} \\ 1 \\ 0_{m_w-j} \end{bmatrix} \tag{19}$$

and we denoted by 0_m the $m-$dimensional zero column vector. We remark that the resulting system has still m_u known inputs and m_w unknown inputs. However, while the m_u known inputs coincide with the original ones, the m_w unknown inputs are now the $k-$order time derivatives of the original unknown inputs. The state evolution depends on the known inputs via the vector fields F^i, $(i = 1, \cdots, m_u)$ and it depends on the unknown inputs via the unit vectors W^j, $(j = 1, \cdots, m_w)$. Finally, we remark that only the vector field G depends on the new state elements.

In the sequel, we will denote the extended system by $\Sigma^{(k)}$. Additionally, we use the notation: $\xi \triangleq [w^T,\ w^{(1)\ T},\ \cdots,\ w^{(k-1)\ T}]^T$. In this notation we have ${}^k x = [x^T,\ \xi^T]^T$. We also denote by $\Sigma^{(0)}$ the original system, i.e., the one characterized by the state x and the equations in (15).

We start by providing a simple result for $\Sigma^{(k)}$:

Lemma 1 *In $\Sigma^{(k)}$, the Lie derivatives of the output up to the m^{th} order ($m \leq k$) are independent of $w_j^{(f)}$, $j = 1, \cdots, m_w$, $\forall f \geq m$.*

Proof: We proceed by induction on m for any k. When $m = 0$ we only have one zero-order Lie derivative (i.e., $h(x)$), which only depends on x, namely it is independent of $w^{(f)}$, $\forall f \geq 0$. Let us assume that the previous assert is true for m and let us prove that it holds for $m+1$. If it is true for m, any Lie derivative up to the m^{th} order is independent of $w^{(f)}$, for any $f \geq m$. In other words, the analytical expression of any Lie derivative up to the $m-$order is represented

by a function $g(x, w, w^{(1)}, \cdots, w^{(m-1)})$. Hence, $dg = [\frac{\partial g}{\partial x}, \frac{\partial g}{\partial w}, \frac{\partial g}{\partial w^{(1)}}, \cdots, \frac{\partial g}{\partial w^{(m-1)}}, 0_{(k-m)m_w}]$. It is immediate to realize that the product of this differential by any vector field in (17) depends at most on $w^{(m)}$, i.e., it is independent of $w^{(f)}$, $\forall f \geq m+1$ ■

A simple consequence of this lemma is the following property:

Proposition 2 *Let us consider the system $\Sigma^{(k)}$. The Lie derivatives of the output up to the k^{th} order along at least one vector among W^j $(j = 1, \cdots, m_w)$ are identically zero.*

Proof: From the previous lemma it follows that all the Lie derivatives, up to the $(k-1)-$order are independent of $w^{(k-1)}$, which are the last m_w components of the extended state in (16). Then, the proof follows from the fact that any vector among W^j $(j = 1, \cdots, m_w)$ has the first $n + (k-1)m_w$ components equal to zero ■

We have the following property:

Proposition 3 *The Lie derivatives of the output up to the k^{th} order along any vector field G, $F^1, \cdots, F^{m_u}$ for the system $\Sigma^{(k)}$ coincide with the same Lie derivatives for the system $\Sigma^{(k+1)}$*

Proof: We proceed by induction on m for any k. When $m = 0$ we only have one zero-order Lie derivative (i.e., $h(x)$), which is obviously the same for the two systems, $\Sigma^{(k)}$and $\Sigma^{(k+1)}$. Let us assume that the previous assert is true for m and let us prove that it holds for $m+1 \leq k$. If it is true for m, any Lie derivative up to the m^{th} order is the same for the two systems. Additionally, from lemma 1, we know that these Lie derivatives are independent of $w^{(f)}$, $\forall f \geq m$. The proof follows from the fact that the first $n + mm_w$ components of the vector fields G, $F^1, \cdots, F^{m_u}$ for $\Sigma^{(k)}$, coincide with the first $n + mm_w$ components of the same vector fields for $\Sigma^{(k+1)}$, when $m < k$ ■

For $\Sigma^{(k)}$ we have a fundamental property that is the extension of the one stated by proposition 1:

Proposition 4 *If x_a and x_b are indistinguishable, there exist ξ_a and ξ_b such that, in $\Sigma^{(k)}$, the Lie derivatives of the output up to the k^{th}-order, along all the vector fields that characterize the dynamics of $\Sigma^{(k)}$, take the same values at $[x_a, \xi_a]$ and $[x_b, \xi_b]$.*

Proof: We consider a piecewise-constant input $\tilde{u}$ as follows $(i = 1, \cdots, m_u)$:

$$\tilde{u}_i(t) = \tag{20}$$
$$\begin{cases} u_i^1 & t \in [0,\ t_1) \\ u_i^2 & t \in [t_1,\ t_1 + t_2) \\ \cdots & \\ u_i^g & t \in [t_1 + t_2 + \cdots + t_{g-1},\ t_1 + t_2 + \cdots + t_{g-1} + t_g) \end{cases}$$

Since x_a and x_b are indistinguishable, there exist two unknown input functions $w_a(t)$ and $w_b(t)$ such that the output coincide on x_a and x_b. In particular, we can write:

$$h(x(t;\ [x_a, \xi_a];\ \tilde{u};\ w_a^{(k)})) = h(x(t;\ [x_b, \xi_b];\ \tilde{u};\ w_b^{(k)})) \tag{21}$$

$\forall t \in [0,\ t_1 + t_2 + \cdots + t_{g-1} + t_g) \subset \mathcal{I}$. On the other hand, by taking the two quantities in (21) at $t = t_1 + t_2 + \cdots + t_{g-1} + t_g$, we can consider them as functions of the g arguments $t_1, t_2, \cdots, t_g$. Hence, by differentiating with respect to all these variables, we also have:

$$\frac{\partial^g h(x(t_1 + \cdots + t_g;\ [x_a, \xi_a];\ \tilde{u};\ w_a^{(k)}))}{\partial t_1 \partial t_2 \cdots \partial t_g} = \tag{22}$$

$$= \frac{\partial^g h(x(t_1 + \cdots + t_g;\ [x_b, \xi_b];\ \tilde{u};\ w_b^{(k)}))}{\partial t_1 \partial t_2 \cdots \partial t_g}$$

By computing the previous derivatives at $t_1 = t_2 = \cdots = t_g = 0$ and by using proposition 2 we obtain, if $g \leq k$:

$$\mathcal{L}^g_{\theta_1\theta_2\cdots\theta_g} h \Big|_{\substack{x = x_a \\ \xi = \xi_a}} = \mathcal{L}^g_{\theta_1\theta_2\cdots\theta_g} h \Big|_{\substack{x = x_b \\ \xi = \xi_b}} \tag{23}$$

where $\theta_h = G + \sum_{i=1}^{m_u} F^i u_i^h$, $h = 1, \cdots, g$. The equality in (23) must hold for all possible choices of $u_1^h, \cdots, u_{m_u}^h$. By appropriately selecting these $u_1^h, \cdots, u_{m_u}^h$, we finally obtain:

$$\mathcal{L}^g_{v_1 v_2 \cdots v_g} h \Big|_{\substack{x = x_a \\ \xi = \xi_a}} = \mathcal{L}^g_{v_1 v_2 \cdots v_g} h \Big|_{\substack{x = x_b \\ \xi = \xi_b}} \tag{24}$$

where $v_1 v_2 \cdots v_g$ are vector fields belonging to the set $\{G, F^1, \cdots, F^{m_u}\}$ ■

The main difference between propositions 1 and 4 is that, in the latter, we cannot consider any order Lie derivative, since the order cannot exceed k. This will have important consequences, as we will see. Before discussing this point, we remark that in [8] it was also defined the concept of $V-$indistinguishable states, with V a subset of the definition set that includes the two considered states. From this definition and the previous proof we can alleviate the assumptions in the previous proposition. Specifically, we have the following:

Remark 1 *The statement of proposition 4 also holds if x_a and x_b are $V-$indistinguishable.*

Now let us discuss how we can use the result stated by the proposition 4 to investigate the observability properties. Thanks to the results stated by propositions 3 and 4 we can easily build a codistribution that is observable. In the sequel, we will denote by d the differential with respect to the entire extended state. We will call *original state* the vector x and, we remind the reader that we denote by $\mathcal{D}$ the differential with respect to the original state. The observable codistribution for $\Sigma^{(k)}$ is the span of the differentials of all the Lie derivatives of the output along G, $F^1, \cdots, F^{m_u}$ up to the k-order. Hence, for any $m \leq k$, it is obtained recursively by the following algorithm:

Algorithm 4 *Observable codistribution for the extended state ($m \leq k$)*

1. $\bar{\Omega}_0 = span\{dh\}$;
2. $\bar{\Omega}_m = \bar{\Omega}_{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1} + \sum_{i=1}^{m_u} \mathcal{L}_{F^i} \bar{\Omega}_{m-1}$

It is possible to obtain all the observability properties of the original state starting from the codistributions generated by the previous algorithm. In order to show this, we introduce the concept of *Observable Mode*. This concept can be also useful in the case without unknown inputs.
Given x_0, we denote by I_{x_0} the set of all the states x such that x and x_0 are indistinguishable. According to the theory of observability, a system is observable in x_0 if $I_{x_0} = x_0$. We introduce here the following new definition:

Definition 2 (Observable Mode) *A scalar function is observable in x_0 if it is constant on I_{x_0}. Additionally, it is weakly observable in x_0 if it exists an open neighbourhood B_{x_0} such that it is constant on $B_{x_0} \cap I_{x_0}$.*

The basic idea behind definition 2 is the following. Let us suppose that the true initial state is x_0. According to the definition of indistinguishable set (which is based on definition 1), the system only contains the information to establish whether the initial state belongs to the indistinguishable set I_{x_0} or not. Hence, if a scalar function takes the same value on this set, we conclude that the system has enough information to know the value of this scalar function at the initial time. We also remark that definition 2 generalizes the definition of observability. Specifically, a system is observable in x_0 when all the components of the state x are observable modes in x_0. When the scalar function is only weakly observable at x_0, we conclude that the system has enough information to know the value of this scalar function at the initial time, provided that we a priori know that the initial state is *sufficiently* close to x_0.

The following two propositions generalize theorem 3.1 and 3.11 in [8].

Proposition 5 *Given a scalar function $\theta(x)$, if it exists an integer k such that for the extended system $\Sigma^{(k)}$, $d\theta \in \bar{\Omega}_k$ in a given x_0 and for a given extension ξ_0, then θ is weakly observable in x_0.*

Proof: $\theta(x)$ can be expressed in terms of the Lie derivatives of the output along the fields that characterize the dynamics of $\Sigma^{(k)}$, up to the $k-$order. We can write $\theta(x) = \mathcal{G}(\phi_1(x,\xi), \cdots, \phi_L(x,\xi))$, $\forall x \in B_{x_0}$ where $\phi_1(x,\xi), \cdots, \phi_L(x,\xi)$ are L Lie derivatives among the ones of above and $\mathcal{G}$ is a given function.

Let us consider a given x indistinguishable from x_0 and that belongs to B_{x_0}. In other words, $x \in B_{x_0} \cap I_{x_0}$. From proposition 4 there are two extensions ξ_a and ξ_b, such that, the Lie derivatives $\phi_1(x,\xi), \cdots, \phi_L(x,\xi)$ take the same values on the two extended states $[x, \xi_a]$ and $[x_0, \xi_b]$. Therefore, $\forall x \in B_{x_0} \cap I_{x_0}$ we have: $\theta(x) = \mathcal{G}(\phi_1(x,\xi_a), \cdots, \phi_L(x,\xi_a)) = \mathcal{G}(\phi_1(x_0,\xi_b), \cdots, \phi_L(x_0,\xi_b)) = \theta(x_0)$ and θ is weakly observable in x_0 ■

Proposition 6 *If the scalar function $\theta(x)$ is weakly observable in x_0, then $\exists k$ and ξ_0 such that $d\theta \in \bar{\Omega}_k$ a.e. on an open neighbourhood B_{x_0}.*

Proof: We proceed by contradiction. $\forall k$ and ξ and for any open ball centered on x_0 with radius r ($B^r_{x_0}$), it exists a set $C^r_{x_0} \subseteq B^r_{x_0}$ with measure strictly larger than zero for which $d\theta \notin \bar{\Omega}_k$. This means that $\forall k$ and ξ it exists a distribution with dimension larger than zero that is orthogonal to the codistribution $\bar{\Omega}_k$ in $C^r_{x_0}$. This means that it exists at least one vector field s, of the same dimension of the original state, such that, $\forall k$ and ξ, it exists at least one vector field s_k of dimension km_w, such that the vector field $[s^T, s_k^T]^T$ is orthogonal to the codistribution $\bar{\Omega}_k$ in $C^r_{x_0}$. From this, we obtain that it exists $\epsilon > 0$ such that the states x_0 and $x_0 + \epsilon s$ are indistinguishable. Indeed, it is possible to express the $m-$time derivative of the output at the initial time, for a generic integer m, in terms of the Lie derivatives of the output along the fields of $\Sigma^{(k)}$, up to the $m-$order, and the known inputs and their time derivatives (see [26], sect 2.5.1). Since this holds for any k, the $m-$order time derivative of the output at the initial time coincides in x_0 and $x_0 + \epsilon s$. Since this holds for any order m, from the Taylor theorem the output coincides on a given time interval and consequently x_0 and $x_0 + \epsilon s$ are indistinguishable. But this means that also $\theta(x_0) = \theta(x_0 + \epsilon s)$. Hence, $d\theta \cdot s = 0$ a.e. on an open neighbourhood B_{x_0} and $d\theta \in \bar{\Omega}_k$ a.e. on an open neighbourhood B_{x_0} ■

Propositions 5 and 6 state that all the observability properties of the original state are contained in the codistributions generated by algorithm 4. From this, we can easily obtain a sufficient

condition for the observability of the original state. Indeed, if on a given x_0, the differential of a given component of x (the original state) belongs to $\bar{\Omega}_m$ for a given integer $m \leq k$, we can conclude that this state component is weakly locally observable (in x_0). If this holds for all the state components, we can conclude that the entire original state is weakly locally observable. More in general, we can conclude that, a given scalar function of the original state, is weakly observable in a given point x_0 if its differential (computed in x_0) belongs to $\bar{\Omega}_m$ (computed in x_0) for a given integer $m \leq k$. On the other hand, we remark the following two fundamental differences between algorithms 1 and 4:

1. In the latter, since the state augmentation can be continued indefinitely, we do not have convergence;
2. The latter provides a codistribution that describes simultaneously the observability properties of the original state and its extension.

The goal of the next section is to address these fundamental issues in the case of a single unknown input. In particular, we show that it is possible to directly compute the entire observable codistribution of the original system, namely, without the need of extending the state.

6.2 Proof of the validity of the analytic criterion

This section is devoted to the case of a single unknown input (i.e., $m_w = 1$) and dynamics linear in the inputs (the vector g^0 is the null vector). In other words, we are considering the system characterized by (1). When $m_w = 1$, the extended state that includes the time derivatives of w up to the $(k-1)-$order is:

$$ {}^k x \triangleq [x^T,\ w,\ w^{(1)},\ \cdots,\ w^{(k-1)}]^T \tag{25}$$

where $w^{(j)} \triangleq \frac{d^j w}{dt^j}$. The dimension of the extended state is in this case $n + k$. For the clarity sake, let us consider the case of a single known input (i.e., we consider the system in (1) with $m_u = 1$). We provide all the analytic results in this simpler case and then we extend them to the case of $m_u > 1$. From (1) it is immediate to obtain the dynamics for the extended state:

$$ {}^k\dot{x} = G({}^k x) + F(x)u + W w^{(k)} \tag{26}$$

where:

$$ F \triangleq \begin{bmatrix} f(x) \\ 0 \\ 0 \\ \cdots \\ 0 \\ 0 \end{bmatrix} \quad G \triangleq \begin{bmatrix} g(x)w \\ w^{(1)} \\ w^{(2)} \\ \cdots \\ w^{(k-1)} \\ 0 \end{bmatrix} \quad W \triangleq \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cdots \\ 0 \\ 1 \end{bmatrix} \tag{27}$$

and we set $f(x) \triangleq f^1(x)$.

Algorithm 4 becomes:

Algorithm 5 *Observable codistribution for the extended state in the case $m_u = 1$ ($m \leq k$)*

1. $\bar{\Omega}_0 = span\{dh\}$;
2. $\bar{\Omega}_m = \bar{\Omega}_{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1} + \mathcal{L}_F \bar{\Omega}_{m-1}$

where G is now the simpler vector field given in (27) and d denotes the differential respect to the state in (25).

We will address the two fundamental issues mentioned at the end of the previous section. This is obtained by two separates steps. In the first step (sect. 6.2.1) we perform a separation on the observable codistribution defined by algorithm 5. This codistribution can be split into two codistributions: the former is the codistribution generated by algorithm 3, once embedded in the extended space, and the latter is the codistribution L^m defined in section 6.2.1 (see theorem 1). We prove (Lemma 2) that the second codistribution (L^m) can be ignored when deriving the observability properties of the original state. In the second step (sect. 6.2.2) we prove that algorithm 3 converges in at most $n+2$ steps and we provide the convergence criterion. Finally, in section 6.2.3 we extend the results of the previous two theorems to the case of multiple known inputs ($m_u > 1$).

6.2.1 Separation

For each integer m, we generate the codistribution Ω_m by using algorithm 3 (note that here we are considering $m_u = 1$). By construction, the generators of Ω_m are the differentials of scalar functions that only depend on the original state (x) and not on its extension. In the sequel, we need to embed this codistribution in $\mathbb{R}^{n+k}$. We will denote by $[\Omega_m, 0_k]$ the codistribution made by covectors whose first n components are covectors in Ω_m and the last components are all zero. Additionally, we will denote by L^m the codistribution that is the span of the Lie derivatives of dh up to the order m along the vector G, i.e., $L^m \triangleq span\{\mathcal{L}_G^1 dh, \mathcal{L}_G^2 dh, \cdots, \mathcal{L}_G^m dh\}$. We finally introduce the following codistribution:

Definition 3 ($\tilde{\Omega}$ codistribution) *This codistribution is defined as follows:* $\tilde{\Omega}_m \triangleq [\Omega_m, 0_k] + L^m$

The codistribution $\tilde{\Omega}_m$ consists of two parts. Specifically, we can select a basis that consists of exact differentials that are the differentials of functions that only depend on the original state (x) and not on its extension (these are the generators of $[\Omega_m, 0_k]$) and the differentials $\mathcal{L}_G^1 dh, \mathcal{L}_G^2 dh, \cdots, \mathcal{L}_G^m dh$. The second set of generators, i.e., the differentials $\mathcal{L}_G^1 dh, \mathcal{L}_G^2 dh, \cdots, \mathcal{L}_G^m dh$, are m and, with respect to the first set, they are differentials of functions that also depend on the state extension $\xi = [w,\ w^{(1)}, \cdots,\ w^{(m-1)}]^T$. We have the following result:

Lemma 2 *Let us denote with x_j the j^{th} component of the state ($j = 1, \cdots, n$). We have: $\mathcal{D}x_j \in \Omega_m$ if and only if $dx_j \in \tilde{\Omega}_m$*

Proof: The fact that $\mathcal{D}x_j \in \Omega_m$ implies that $dx_j \in \tilde{\Omega}_m$ is obvious since $[\Omega_m, 0_k] \subseteq \tilde{\Omega}_m$ by definition. Let us prove that also the contrary holds, i.e., that if $dx_j \in \tilde{\Omega}_m$ then $\mathcal{D}x_j \in \Omega_m$. Since $dx_j \in \tilde{\Omega}_m$ we have $dx_j = \sum_{i=1}^{N_1} c_i^1 \omega_i^1 + \sum_{i=1}^{N_2} c_i^2 \omega_i^2$, where $\omega_1^1, \omega_2^1, \cdots, \omega_{N_1}^1$ are N_1 generators of $[\Omega_m, 0_k]$, $\omega_1^2, \omega_2^2, \cdots, \omega_{N_2}^2$ are N_2 generators of L^m and $c_1^1, \cdots, c_{N_1}^1, c_1^2, \cdots, c_{N_2}^2$ are suitable coefficients. We want to prove that $N_2 = 0$.

We proceed by contradiction. Let us suppose that $N_2 \geq 1$. We remark that the first set of generators have the last k entries equal to zero, as for dx_j. The second set of generators consists of the Lie derivatives of dh along G up to the m order. Let us select the one that is the highest order Lie derivative and let us denote by j' this highest order. We have $1 \leq N_2 \leq j' \leq m$. By a direct computation, it is immediate to realize that this is the only generator that depends on $w^{(j'-1)}$. Specifically, the dependence is linear by the product $L_g^1 w^{(j'-1)}$ (we remind the reader that $L_g^1 \neq 0$). But this means that dx_j has the $(n+j')^{th}$ entry equal to $L_g^1 \neq 0$ and this is not possible since $dx_j = [\mathcal{D}x_j, 0_k]$ ■

A fundamental consequence of this lemma is that, if we are able to prove that $\tilde{\Omega}_m = \bar{\Omega}_m$, the weak local observability of the original state x, can be investigated by only considering the codistribution Ω_m. In the rest of this section we prove this fundamental theorem, stating that $\tilde{\Omega}_m = \bar{\Omega}_m$.

For a given $m \leq k$ we define the vector $\Phi_m \in \mathbb{R}^{n+k}$ by the following algorithm:

1. $\Phi_0 = F$;
2. $\Phi_m = [\Phi_{m-1},\ G]$

where now the Lie brackets $[\cdot, \cdot]$ are computed with respect to the extended state, i.e.:

$$[F,\ G] \triangleq \frac{\partial G}{\partial^k x} F(^k x) - \frac{\partial F}{\partial^k x} G(^k x)$$

By a direct computation it is easy to realize that Φ_m has the last k components identically null. In the sequel, we will denote by $\breve{\Phi}_m$ the vector in $\mathbb{R}^n$ that contains the first n components of Φ_m. In other words, $\Phi_m \triangleq [\breve{\Phi}_m^T, 0_k^T]^T$. Additionally, we set $\hat{\phi}_m \triangleq \begin{bmatrix} \phi_m \\ 0_k \end{bmatrix}$ (ϕ_m is defined by algorithm 2).

We have the following result:

Lemma 3 $\mathcal{L}_G \bar{\Omega}_m + \mathcal{L}_{\Phi_m} dh = \mathcal{L}_G \bar{\Omega}_m + \mathcal{L}_F \mathcal{L}_G^m dh$

Proof: We have $\mathcal{L}_F \mathcal{L}_G^m dh = \mathcal{L}_G \mathcal{L}_F \mathcal{L}_G^{m-1} dh + \mathcal{L}_{\Phi_1} \mathcal{L}_G^{m-1} dh$.

The first term $\mathcal{L}_G \mathcal{L}_F \mathcal{L}_G^{m-1} dh \in \mathcal{L}_G \bar{\Omega}_m$. Hence, we need to prove that $\mathcal{L}_G \bar{\Omega}_m + \mathcal{L}_{\Phi_m} dh = \mathcal{L}_G \bar{\Omega}_m + \mathcal{L}_{\Phi_1} \mathcal{L}_G^{m-1} dh$. We repeat the previous procedure m times. Specifically, we use the equality $\mathcal{L}_{\Phi_j} \mathcal{L}_G^{m-j} dh = \mathcal{L}_G \mathcal{L}_{\Phi_j} \mathcal{L}_G^{m-j-1} dh + \mathcal{L}_{\Phi_{j+1}} \mathcal{L}_G^{m-j-1} dh$, for $j = 1, \cdots, m$, and we remove the first term since $\mathcal{L}_G \mathcal{L}_{\Phi_j} \mathcal{L}_G^{m-j-1} dh \in \mathcal{L}_G \bar{\Omega}_m$ ■

Lemma 4 $\breve{\Phi}_m = \sum_{j=1}^m c_j^n (\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h) \phi_j$, *i.e., the vector* $\breve{\Phi}_m$ *is a linear combination of the vectors* ϕ_j $(j = 1, \cdots, m)$, *where the coefficients* (c_j^n) *depend on the state only through the functions that generate the codistribution* L^m

Proof: We proceed by induction. By a direct computation it is immediate to obtain: $\breve{\Phi}_1 = \phi_1 \mathcal{L}_G h$.

Inductive step: Let us assume that $\breve{\Phi}_{m-1} = \sum_{j=1}^{m-1} c_j (\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^{m-1} h) \phi_j$. We have:

$$\Phi_m = [\Phi_{m-1},\ G] = \sum_{j=1}^{m-1} \left[c_j \begin{bmatrix} \phi_j \\ 0_k \end{bmatrix},\ G \right] =$$

$$\sum_{j=1}^{m-1} c_j \left[\begin{bmatrix} \phi_j \\ 0_k \end{bmatrix},\ G \right] - \sum_{j=1}^{m-1} \mathcal{L}_G c_j \begin{bmatrix} \phi_j \\ 0_k \end{bmatrix}$$

We directly compute the Lie bracket in the sum (note that ϕ_j is independent of the unknown input w and its time derivatives):

$$\left[\begin{bmatrix} \phi_j \\ 0_k \end{bmatrix},\ G \right] = \begin{bmatrix} [\phi_j,\ g]w \\ 0_k \end{bmatrix} = \begin{bmatrix} \phi_{j+1} \mathcal{L}_G^1 h \\ 0_k \end{bmatrix}$$

Regarding the second term, we remark that $\mathcal{L}_G c_j = \sum_{i=1}^{m-1} \frac{\partial c_j}{\partial (\mathcal{L}_G^i h)} \mathcal{L}_G^{i+1} h$. By setting $\tilde{c}_j = c_{j-1} \mathcal{L}_G^1 h$ for $j = 2, \cdots, m$ and $\tilde{c}_1 = 0$, and by setting $\bar{c}_j = -\sum_{i=1}^{m-1} \frac{\partial c_j}{\partial (\mathcal{L}_G^i h)} \mathcal{L}_G^{i+1} h$ for $j =$

$1, \cdots, m-1$ and $\bar{c}_m = 0$, we obtain $\breve{\Phi}_m = \sum_{j=1}^{m}(\tilde{c}_j + \bar{c}_j)\phi_j$, which proves our assert since $c_j^n (\triangleq \tilde{c}_j + \bar{c}_j)$ is a function of $\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h$ ■

It also holds the following result:

Lemma 5 *If $w \neq 0$, $\hat{\phi}_m = \sum_{j=1}^{m} b_j^n(\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h)\Phi_j$, i.e., the vector $\hat{\phi}_m$ is a linear combination of the vectors Φ_j $(j = 1, \cdots, m)$, where the coefficients (b_j^n) depend on the state only through the functions that generate the codistribution L^m*

Proof: We proceed by induction. By a direct computation it is immediate to obtain: $\hat{\phi}_1 = \Phi_1 \frac{1}{\mathcal{L}_G h}$ (note that $\mathcal{L}_G h = L_g^1 w \neq 0$).
Inductive step: Let us assume that $\hat{\phi}_{m-1} = \sum_{j=1}^{m-1} b_j(\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^{m-1} h)\Phi_j$. We need to prove that $\hat{\phi}_m = \sum_{j=1}^{m} b_j^n(\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h)\Phi_j$. We start by applying on both members of the equality $\hat{\phi}_{m-1} = \sum_{j=1}^{m-1} b_j(\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^{m-1} h)\Phi_j$ the Lie bracket with respect to G. We obtain for the first member: $[\hat{\phi}_{m-1},\ G] = \hat{\phi}_m \mathcal{L}_G^1 h$. For the second member we have:

$$\sum_{j=1}^{m-1}[b_j\Phi_j,\ G] = \sum_{j=1}^{m-1} b_j[\Phi_j,\ G] - \sum_{j=1}^{m-1}\mathcal{L}_G b_j \Phi_j =$$

$$= \sum_{j=1}^{m-1} b_j \Phi_{j+1} - \sum_{j=1}^{m-1}\sum_{i=1}^{m-1} \frac{\partial b_j}{\partial(\mathcal{L}_G^i h)} \mathcal{L}_G^{i+1} h \Phi_j$$

Since $\mathcal{L}_G h = L_g^1 w \neq 0$, by setting $\tilde{b}_j = \frac{b_{j-1}}{\mathcal{L}_G^1 h}$ for $j = 2, \cdots, m$ and $\tilde{b}_1 = 0$, and by setting $\bar{b}_j = -\sum_{i=1}^{m-1} \frac{\partial b_j}{\partial(\mathcal{L}_G^i h)} \frac{\mathcal{L}_G^{i+1} h}{\mathcal{L}_G^1 h}$ for $j = 1, \cdots, m-1$ and $\bar{b}_m = 0$, we obtain $\hat{\phi}_m = \sum_{j=1}^{m}(\tilde{b}_j + \bar{b}_j)\Phi_j$, which proves our assert since $b_j^n (\triangleq \tilde{b}_j + \bar{b}_j)$ is a function of $\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h$ ■

An important consequence of the previous two lemmas is the following result:

Proposition 7 *If $w \neq 0$, the following two codistributions coincide:*

1. *$span\{\mathcal{L}_{\Phi_0} dh, \mathcal{L}_{\Phi_1} dh, \cdots, \mathcal{L}_{\Phi_m} dh, \mathcal{L}_G^1 dh, \cdots \mathcal{L}_G^m dh\}$;*
2. *$span\{\mathcal{L}_{\hat{\phi}_0} dh, \mathcal{L}_{\hat{\phi}_1} dh, \cdots, \mathcal{L}_{\hat{\phi}_m} dh, \mathcal{L}_G^1 dh, \cdots \mathcal{L}_G^m dh\}$;*

We are now ready to prove the following fundamental result:

Theorem 1 (Separation) *If $w \neq 0$, $\bar{\Omega}_m = \tilde{\Omega}_m \triangleq [\Omega_m, 0_k] + L^m$*

Proof: We proceed by induction. By definition, $\bar{\Omega}_0 = \tilde{\Omega}_0$ since they are both the span of dh.
Inductive step: Let us assume that $\bar{\Omega}_{m-1} = \tilde{\Omega}_{m-1}$. We have: $\bar{\Omega}_m = \bar{\Omega}_{m-1} + \mathcal{L}_F \bar{\Omega}_{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1} = \bar{\Omega}_{m-1} + \mathcal{L}_F \tilde{\Omega}_{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1} = \bar{\Omega}_{m-1} + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_F L^{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1}$. On the other hand, $\mathcal{L}_F L^{m-1} = \mathcal{L}_F \mathcal{L}_G^1 dh + \cdots + \mathcal{L}_F \mathcal{L}_G^{m-2} dh + \mathcal{L}_F \mathcal{L}_G^{m-1} dh$. The first $m-2$ terms are in $\bar{\Omega}_{m-1}$. Hence we have: $\bar{\Omega}_m = \bar{\Omega}_{m-1} + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_F \mathcal{L}_G^{m-1} dh + \mathcal{L}_G \bar{\Omega}_{m-1}$. By using lemma 3 we obtain: $\bar{\Omega}_m = \bar{\Omega}_{m-1} + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_{\Phi_{m-1}} dh + \mathcal{L}_G \bar{\Omega}_{m-1}$. By using again the induction assumption we obtain: $\bar{\Omega}_m = [\Omega_{m-1}, 0_k] + L^{m-1} + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_{\Phi_{m-1}} dh + \mathcal{L}_G[\Omega_{m-1}, 0_k] + \mathcal{L}_G L^{m-1} = [\Omega_{m-1}, 0_k] + L^m + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_{\Phi_{m-1}} dh + [\mathcal{L}_{\frac{g}{L_g^1}} \Omega_{m-1}, 0_k]$, where we used $L^m + \mathcal{L}_G[\Omega_{m-1}, 0_k] = L^m + [\mathcal{L}_{\frac{g}{L_g^1}} \Omega_{m-1}, 0_k]$, which holds because $\mathcal{L}_G h = L_g^1 w \neq 0$. By using proposition 7, we obtain: $\bar{\Omega}_m = [\Omega_{m-1}, 0_k] + L^m + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_{\hat{\phi}_{m-1}} dh + [\mathcal{L}_{\frac{g}{L_g^1}} \Omega_{m-1}, 0_k] = \tilde{\Omega}_m$
■

Theorem 1 is fundamental. It allows us to obtain all the observability properties of the original state by restricting the computation to the codistribution defined by algorithm 3, namely a codistribution whose covectors have the same dimension of the original space. In other words, the dimension of these covectors is independent of the state augmentation.

6.2.2 Convergence

Algorithm 3 is recursive and $\Omega_m \subseteq \Omega_{m+1}$. This means that, if for a given m the differentials of the components of the original state belong to Ω_m, we can conclude that the original state is weakly locally observable. On the other hand, if this is not true, we cannot exclude that it is true for a larger m. The goal of this section is precisely to address this issue. We will show that the algorithm converges in a finite number of steps and we will also provide the criterion to establish that the algorithm has converged (theorem 2). This theorem will be proved at the end of this section since we need to introduce several important new quantities and properties.
When investigating the convergence properties of algorithm 3, we remark that, the main difference between algorithm 1 and 3, is the presence of the last term in the recursive step of the latter. Without this term, the convergence criterion would simply consist of the inspection of the equality $\Omega_{m+1} = \Omega_m$, as for algorithm 1. Indeed, the convergence criterion of algorithm 1 is a consequence of the fact that, all the terms that appear at the recursive step of algorithm 1, are the Lie derivative of the codistribution at the previous step, along fixed vector fields (i.e., vector fields that remain the same at each step of the algorithm). This is not the case for the last term at the recursive step of algorithm 3.

The following result provides the convergence criterion in a very special case that basically occurs when the contribution due to the last term in the recursive step of algorithm 3 is included in the other terms. In this case, we obviously obtain that the convergence criterion consists of the inspection of the equality $\Omega_{m+1} = \Omega_m$, as for algorithm 1. For any integer $j \geq 0$ we define:

$$\chi_j \triangleq \frac{\mathcal{L}_{\phi_j} L_g^1}{L_g^1} \tag{28}$$

We have the following result:

Lemma 6 *Let us denote by Λ_j the distribution generated by $\phi_0, \phi_1, \cdots, \phi_j$ and by $m(\leq n-1)$ the smallest integer for which $\Lambda_{m+1} = \Lambda_m$ (n is the dimension of the state x). In the very special case when $\chi_j = 0$, $\forall j = 0, \cdots, m$, algorithm 3 converges at the integer j such that $\Omega_{j+1} = \Omega_j$ and this occurs in at most $n-1$ steps*[8].

Proof: First of all, we remind the reader that the existence of an integer $m(\leq n-1)$ such that $\Lambda_{m+1} = \Lambda_m$ is proved in [16]. In particular, the first chapter in [16] analyzes the convergence of Λ_j with respect to j. It is proved that the distribution converges to Λ^* and that the convergence is achieved at the smallest integer for which we have $\Lambda_{m+1} = \Lambda_m$. Additionally, we have $\Lambda_{m+1} = \Lambda_m = \Lambda^*$ and m cannot exceed $n-1$.

In the very special case when $\chi_j = 0$, $\forall j = 0, \cdots, m$, thanks to the aforementioned convergence of the distribution Λ_j, we easily obtain that $\mathcal{L}_{\phi_{j-1}} \mathcal{L}_g h = 0$ $\forall j \geq 1$. Now, let us consider

[8]Note that, in the case of multiple known inputs (i.e., $m_u > 1$), the previous result does not change. In particular, Λ_j becomes the distribution generated by ${}^1\phi_0, {}^1\phi_1, \cdots, {}^1\phi_j, \cdots, {}^{m_u}\phi_0, {}^{m_u}\phi_1, \cdots, {}^{m_u}\phi_j$ and the very special case is now characterized by $\frac{\mathcal{L}_{i_{\phi_j}} L_g^1}{L_g^1} = 0$, $\forall j = 0, \cdots, m$ and $\forall i = 1, \cdots, m_u$.

the following equation:

$$\mathcal{L}_{\phi_j} h = \frac{1}{L_g^1} \left(\mathcal{L}_{\phi_{j-1}} \mathcal{L}_g h - \mathcal{L}_g \mathcal{L}_{\phi_{j-1}} h \right) \tag{29}$$

Since $\mathcal{L}_{\phi_{j-1}} \mathcal{L}_g h = 0 \; \forall j \geq 1$, we have $\mathcal{L}_{\phi_j} h = -\mathcal{L}_{\frac{g}{L_g^1}} \mathcal{L}_{\phi_{j-1}} h$, for any $j \geq 1$. Therefore, we conclude that, the last term in the recursive step of algorithm 3, is included in the second last term and, in this special case, algorithm 3 has converged when $\Omega_{m+1} = \Omega_m$. This occurs in at most $n-1$ steps, as for algorithm 1. ■

Let us consider now the general case. To proceed we need to introduce several important new quantities and properties.
For a given positive integer j we define the vector $\psi_j \in \mathbb{R}^n$ by the following algorithm:

1. $\psi_0 = f$;
2. $\psi_j = [\psi_{j-1}, \; \frac{g}{L_g^1}]$

It is possible to find the expression that relates these vectors to the vectors ϕ_j, previously defined. Specifically we have:

Lemma 7 *It holds the following equation:*

$$\psi_j = \phi_j + \left\{ \sum_{i=0}^{j-1} (-)^{j-i} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-1} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1} \tag{30}$$

Proof: We proceed by induction. By definition $\psi_0 = \phi_0 = f$ and equation (30) holds for $j = 0$.
Inductive step: Let us assume that it holds for a given $j - 1 \geq 0$ and let us prove its validity for j. We have:

$$\psi_j = \left[\psi_{j-1}, \; \frac{g}{L_g^1} \right] = \left[\phi_{j-1}, \; \frac{g}{L_g^1} \right]$$
$$+ \left[\left\{ \sum_{i=0}^{j-2} (-)^{j-i-1} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-2} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1}, \; \frac{g}{L_g^1} \right]$$

On the other hand:

$$\left[\phi_{j-1}, \; \frac{g}{L_g^1} \right] = \phi_j - \frac{\mathcal{L}_{\phi_{j-1}} L_g^1}{L_g^1} \frac{g}{L_g^1}$$

and

$$\left[\left\{ \sum_{i=0}^{j-2} (-)^{j-i-1} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-2} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1}, \; \frac{g}{L_g^1} \right] =$$
$$-\mathcal{L}_{\frac{g}{L_g^1}} \left\{ \sum_{i=0}^{j-2} (-)^{j-i-1} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-2} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1} =$$
$$\left\{ \sum_{i=0}^{j-2} (-)^{j-i} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-1} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1}$$

Hence:

$$\psi_j = \phi_j - \frac{\mathcal{L}_{\phi_{j-1}} L_g^1}{L_g^1} \frac{g}{L_g^1} + \left\{ \sum_{i=0}^{j-2} (-)^{j-i} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-1} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1},$$

which coincides with (30) ■

We have the following result:

Lemma 8 *For $i = 0, 1, \cdots, m-2$, we have:*

$$\mathcal{D} \frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \in \Omega_m \tag{31}$$

Proof: By construction, $\mathcal{D}\mathcal{L}_{\phi_i} h \in \Omega_m$, for any $i = 1, \cdots, m-1$. On the other hand, we have:

$$\mathcal{L}_{\phi_i} h = \frac{1}{L_g^1} [\mathcal{L}_{\phi_{i-1}} \mathcal{L}_g h - \mathcal{L}_g \mathcal{L}_{\phi_{i-1}} h] = \frac{\mathcal{L}_{\phi_{i-1}} L_g^1}{L_g^1} - \mathcal{L}_{\frac{g}{L_g^1}} \mathcal{L}_{\phi_{i-1}} h$$

We compute the differential of both members of this equation. Since $\mathcal{D}\mathcal{L}_{\frac{g}{L_g^1}} \mathcal{L}_{\phi_{i-1}} h \in \Omega_m$, for any $i = 1, \cdots, m-1$, also $\mathcal{D} \frac{\mathcal{L}_{\phi_{i-1}} L_g^1}{L_g^1} \in \Omega_m$ ■

From lemma 7 with $j = 1, \cdots, m-1$ and lemma 8 it is immediate to obtain the following result:

Proposition 8 *If Ω_m is invariant with respect to $\mathcal{L}_f$ and $\mathcal{L}_{\frac{g}{L_g^1}}$ then it is also invariant with respect to $\mathcal{L}_{\phi_j}$, $j = 1, \cdots, m-1$.*

In order to obtain the convergence criterion for algorithm 3 we need to find the analytic expression that relates three consecutive Lie derivatives, $\mathcal{L}_{\phi_j} h$, $\mathcal{L}_{\phi_{j-1}} h$ and $\mathcal{L}_{\phi_{j-2}} h$. This will allow us to detect the key quantity that governs the convergence of algorithm 3, in particular regarding the contribution due to the last term in the recursive step. This quantity is a scalar and it is the one provided in (3). For the sake of clarity, we provide equation (3) below:

$$\tau \triangleq \frac{L_g^2}{(L_g^1)^2}$$

where $L_g^2 \triangleq \mathcal{L}_g^2 h$. The behaviour of the last term in the recursive step of algorithm 3 is given by the following lemma:

Lemma 9 *We have the following key equality:*

$$\mathcal{L}_{\phi_j} h = \mathcal{L}_{\phi_{j-2}} \tau + \tau \frac{\mathcal{L}_{\phi_{j-2}} L_g^1}{L_g^1} - \mathcal{L}_{\frac{g}{L_g^1}} \left(\frac{\mathcal{L}_{\phi_{j-2}} L_g^1}{L_g^1} + \mathcal{L}_{\phi_{j-1}} h \right) \tag{32}$$

$j \geq 2$.

Proof: We will prove this equality by an explicit computation. We have:

$$\mathcal{L}_{\phi_j} h = \frac{1}{L_g^1} \left(\mathcal{L}_{\phi_{j-1}} \mathcal{L}_g h - \mathcal{L}_g \mathcal{L}_{\phi_{j-1}} h \right)$$

The second term on the right hand side simplifies with the last term in (32). Hence we have to prove:

$$\frac{1}{L_g^1}\mathcal{L}_{\phi_{j-1}}L_g^1 = \mathcal{L}_{\phi_{j-2}}\tau + \tau\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1} - \mathcal{L}_{\frac{g}{L_g^1}}\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1} \tag{33}$$

We have:

$$\frac{1}{L_g^1}\mathcal{L}_{\phi_{j-1}}L_g^1 = \frac{1}{(L_g^1)^2}\left(\mathcal{L}_{\phi_{j-2}}L_g^2 - \mathcal{L}_g\mathcal{L}_{\phi_{j-2}}L_g^1\right) \tag{34}$$

We remark that:

$$\frac{1}{(L_g^1)^2}\mathcal{L}_{\phi_{j-2}}L_g^2 = \mathcal{L}_{\phi_{j-2}}\tau + 2\tau\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1}$$

and

$$\frac{1}{(L_g^1)^2}\mathcal{L}_g\mathcal{L}_{\phi_{j-2}}L_g^1 = \tau\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1} + \mathcal{L}_{\frac{g}{L_g^1}}\frac{\mathcal{L}_{\phi_{j-2}}L_g^1}{L_g^1}$$

By substituting these two last equalities in (34) we immediately obtain (33) ■

Lemma 10 *In general, it exists an integer $m \leq n+2$ (being n the dimension of x) such that $\mathcal{D}\tau \in \Omega_m$.*

Proof: Let us introduce the following notation, for a given integer j:

- $\mathcal{Z}_j \triangleq \mathcal{L}_{\phi_{j+2}}h$;
- $\mathcal{B}_j \triangleq \mathcal{L}_{\phi_j}\tau$;
- $\chi_j \triangleq \frac{\mathcal{L}_{\phi_j}L_g^1}{L_g^1}$.

By construction, $\mathcal{D}\mathcal{Z}_j \in \Omega_{j+3}$. On the other hand, from equation (32), we immediately obtain:

$$\mathcal{D}\mathcal{Z}_j = \mathcal{D}\mathcal{B}_j + \chi_j\mathcal{D}\tau + \tau\mathcal{D}\chi_j - \mathcal{L}_{\frac{g}{L_g^1}}\left(\mathcal{D}\chi_j + \mathcal{D}\mathcal{L}_{\phi_{j+1}}h\right) \tag{35}$$

By using lemma 8 we obtain the following results:

- $\tau\mathcal{D}\chi_j \in \Omega_{j+2}$;
- $\mathcal{L}_{\frac{g}{L_g^1}}\mathcal{D}\chi_j \in \Omega_{j+3}$.

Additionally, $\mathcal{L}_{\frac{g}{L_g^1}}\mathcal{D}\mathcal{L}_{\phi_{j+1}}h \in \Omega_{j+3}$. Hence, from (35), we obtain that the following covector:

$$\mathcal{Z}_j' \triangleq \mathcal{D}\mathcal{B}_j + \chi_j\mathcal{D}\tau \tag{36}$$

belongs to Ω_{j+3}. Let us denote by j^* the smallest integer such that:

$$\mathcal{D}\mathcal{B}_{j^*} = \sum_{j=0}^{j^*-1} c_j\mathcal{D}\mathcal{B}_j + c_{-1}\mathcal{D}h \tag{37}$$

Note that j^* is a finite integer and in particular $j^* \leq n-1$. Indeed, if this would not be the case, the dimension of the codistribution generated by $\mathcal{D}h, \mathcal{D}\mathcal{B}_0, \mathcal{D}\mathcal{B}_1, \cdots, \mathcal{D}\mathcal{B}_{n-1}$ would be $n+1$, i.e., larger than n. From (37) and (36) we obtain:

$$\mathcal{Z}'_{j*} = \sum_{j=0}^{j^*-1} c_j \mathcal{D}\mathcal{B}_j + c_{-1}\mathcal{D}h + \chi_{j*}\mathcal{D}\tau \tag{38}$$

From equation (36), for $j = 0, \cdots, j^*-1$, we obtain: $\mathcal{D}\mathcal{B}_j = \mathcal{Z}'_j - \chi_j \mathcal{D}\tau$. By substituting in (38) we obtain:

$$\mathcal{Z}'_{j*} - \sum_{j=0}^{j^*-1} c_j \mathcal{Z}'_j - c_{-1}\mathcal{D}h = \left(- \sum_{j=0}^{j^*-1} c_j \chi_j + \chi_{j*} \right) \mathcal{D}\tau \tag{39}$$

We remark that the left hand side consists of the sum of covectors that belong to Ω_{j^*+3}. Since in general $\chi_{j*} \neq \sum_{j=0}^{j^*-1} c_j \chi_j$, we have $\mathcal{D}\tau \in \Omega_{j^*+3}$. By setting $m \triangleq j^* + 3$, we have $m \leq n+2$ and $\mathcal{D}\tau \in \Omega_m$ ■

The previous lemma ensures that, in general, it exists a finite $m \leq n+2$ such that $\mathcal{D}\tau \in \Omega_m$. Note that the previous proof holds if the quantity $\chi_{j*} - \sum_{j=0}^{j^*-1} c_j \chi_j$ does not vanish. This holds in general, with the exception of the trivial case considered in lemma 6, in which case $\chi_j = 0, \ \forall j$.

The following theorem allows us to obtain the criterion to stop algorithm 3:

Theorem 2 *If $\mathcal{D}\tau \in \Omega_m$ and Ω_m is invariant under $\mathcal{L}_f$ and $\mathcal{L}_{\frac{g}{L_g^1}}$, then $\Omega_{m+p} = \Omega_m \ \forall p \geq 0$*

Proof: We proceed by induction. Obviously, the equality holds for $p = 0$.
Inductive step: let us assume that $\Omega_{m+p} = \Omega_m$ and let us prove that $\Omega_{m+p+1} = \Omega_m$. We have to prove that $\mathcal{D}\mathcal{L}_{\phi_{m+p}} h \in \Omega_m$. Indeed, from the inductive assumption, we know that $\Omega_{m+p} (= \Omega_m)$ is invariant under $\mathcal{L}_f$ and $\mathcal{L}_{\frac{g}{L_g^1}}$. Additionally, because of this invariance, by using proposition 8, we obtain that Ω_m is also invariant under $\mathcal{L}_{\phi_j}$, for $j = 1, 2, \cdots, m+p-1$. Since $\mathcal{D}\tau \in \Omega_m$ we have $\mathcal{D}\mathcal{L}_{\phi_{m+p-2}}\tau \in \Omega_m$. Additionally, $\mathcal{D}\mathcal{L}_{\phi_{m+p-1}} h \in \Omega_m$ and, because of lemma 8, we also have $\mathcal{D}\frac{\mathcal{L}_{\phi_{m+p-2}} L_g^1}{L_g^1} \in \Omega_m$. Finally, because of the invariance under $\mathcal{L}_{\frac{g}{L_g^1}}$, also the Lie derivatives along $\frac{g}{L_g^1}$ of $\mathcal{D}\mathcal{L}_{\phi_{m+p-1}} h$ and $\mathcal{D}\frac{\mathcal{L}_{\phi_{m+p-2}} L_g^1}{L_g^1}$ belong to Ω_m. Now, we use equation (32) for $j = m+p$. By computing the differential of this equation it is immediate to obtain that $\mathcal{D}\mathcal{L}_{\phi_{m+p}} h \in \Omega_m$ ■

We conclude this section by providing an upper bound for the number of steps that are in general necessary to achieve the convergence. The dimension of Ω_{j^*+2} is at least the dimension of the span of the covectors: $\mathcal{D}h, \ \mathcal{Z}'_0, \ \mathcal{Z}'_1, \ \cdots, \ \mathcal{Z}'_{j^*-1}$. From the definition of j^*, we know that the vectors $\mathcal{D}h, \ \mathcal{D}\mathcal{B}_0, \ \mathcal{D}\mathcal{B}_1, \ \cdots, \ \mathcal{D}\mathcal{B}_{j^*-1}$ are independent meaning that the dimension of their span is j^*+1. Hence, from (36), it easily follows that the dimension of the span of the vectors $\mathcal{D}h, \ \mathcal{Z}'_0, \ \mathcal{Z}'_1, \ \cdots, \ \mathcal{Z}'_{j^*-1}, \ \mathcal{D}\tau$ is at least j^*+1. Since Ω_{j^*+3} contains this span, its dimension is at least j^*+1. Therefore, the condition $\Omega_{m+1} = \Omega_m$, for $m \geq j^*+3$ is achieved for $m \leq n+2$.

6.2.3 Extension to the case of multiple known inputs

It is immediate to repeat all the steps carried out in the previous two subsections and extend the validity of theorem 1 to the case of multiple known inputs ($m_u > 1$). Additionally, also theorem 2 can be easily extended to cope with the case of multiple known inputs. In this case,

requiring that $\Omega_{m+1} = \Omega_m$ means that Ω_m must be invariant with respect to $\mathcal{L}_{\frac{g}{L_g^1}}$ and all $\mathcal{L}_{f^i}$ simultaneously.

7 Conclusion

The goal of this work was to present the analytic solution of a fundamental open problem in control theory. The problem is known in the literature as the Unknown Input Observability (UIO) problem. It consists in deriving the analytic criterion that allows us to check the system observability in presence of unknown inputs (throughout this work we also used the term *disturbance* to mean an unknown input). The solution here provided holds in the case of a single unknown input. In other words, it is provided the analytic criterion that allows us to obtain the observability of a nonlinear system in presence of a single unknown input and multiple known inputs (see section 2 for a definition of the considered systems).

As for the observability rank condition, the proposed analytic criterion is based on the computation of the observable codistribution. Similarly to the case of only known inputs, the observable codistribution is obtained by recursively computing the Lie derivatives of the outputs along the vector fields that characterize the dynamics. However, in correspondence of the unknown input, the corresponding vector field must be suitably rescaled. Additionally, the Lie derivatives of the outputs must also be computed along a new set of vector fields that are obtained by recursively performing suitable Lie bracketing of the vector fields that define the dynamics. In practice, the entire observable codistribution is obtained by a very simple recursive algorithm. This algorithm was provided in section 3. Finally, it was shown that the recursive algorithm converges in a finite number of steps and the criterion to establish that the convergence has been reached was provided.

The analytic criterion here introduced is the extension of the observability rank condition to the case when the dynamics are driven by an unknown input (and multiple known inputs). The analytic criterion was quickly outlined in section 4 to make easy its automatic implementation. Its validity was proven in section 6. Finally, it was illustrated in section 5 by checking the weak local observability of several nonlinear systems driven by multiple known inputs and a single unknown input.

We are currently extending the proposed analytic criterion to deal with the general case of multiple unknown inputs and when the dynamics are affine (and not simply linear) in the inputs (both known and unknown). In few words, to deal with the general case described by equation (15).

We conclude this work with the following remark, which has a fundamental practical importance for designing an estimator in presence of a single unknown input. We remark that, the state estimated by the two estimators adopted in section 5 (sections 5.2.4 and 5.3.4), includes the disturbance together with its first order time derivative. We found a much worse performance by using an estimator that only estimates the original state or by including in the state only the disturbance (without its time derivative).

We conjecture that this is a consequence of the fact that the observable codistibution is Ω_2, i.e., for the considered system, algorithm 3 converges at its second step. Indeed, we proved that, for a system for which algorithm 3 converges to Ω_m (i.e., converges in m steps) the observable modes[9] satisfy a fundamental property. These functions are constant on the indistinguishable sets only in the extended state space that includes the disturbance and its time derivatives up to the $m-1$ order (see proposition 4 together with theorem 1 and lemma 2). Note that, in the space of the original state (or in a less extended space, namely, that does not include all the first $m-1$ time derivatives), the observable modes are not constant on the indistinguishable sets. Proving this conjecture will require further theoretical investigations. The validity of this conjecture would make the analytic tool not only a tool to check the state observability but also a fundamental tool to design an appropriate estimator for a given nonlinear system in presence

[9]See definition 2 for the definition of *observable mode*.

of a disturbance.

A The case when $L_g^1 = 0$

We consider the case of multiple outputs. In other words, our system is defined by the following equations:

$$\begin{cases} \dot{x} = \sum_{i=1}^{m_u} f^i(x)u_i + g(x)w \\ y_k = h_k(x), \quad k = 1, \cdots, p \end{cases} \tag{40}$$

We introduce the following notation:

- $\mathcal{F}$ is the space of functions that contains all the outputs $h_1, \cdots, h_p$ and their Lie derivatives along the vector fields $f^1, \cdots, f^{m_u}$, up to any order;
- $\mathcal{DF}$ is the codistribution generated by the differentials, with respect to the state x, of the functions in $\mathcal{F}$;
- $\mathcal{L}_G\mathcal{F}$ is the space of functions that consists of the Lie derivatives along G of the functions in $\mathcal{F}$ (we remind the reader that G is the vector defined in (27));

It is immediate to prove the following properties:

- From $\mathcal{F}$ we can select a set of functions such that their differentials generate $\mathcal{DF}$.
- It exists an integer m such that, by running m recursive steps of algorithm 4, we obtain a codistribution $\bar{\Omega}_m$ that contains $[\mathcal{DF},\ 0_m]$ (i.e., the codistribution $\mathcal{DF}$ once embedded in $\mathbb{R}^{n+m}$)[10].
- In general, the functions belonging to $\mathcal{L}_G\mathcal{F}$, are functions of x and w.

Because of the second property, we are allowed to choose any function in $\mathcal{F}$ as a system output (note that the functions in $\mathcal{F}$ only depend on the state x and not on the unknown input w). Hence, if for a given system the condition $\mathcal{L}_g h_k = 0$, $\forall k = 1, \cdots p$, we can redefine the output by choosing any function in $\mathcal{F}$.

To conclude, let us consider the case when no function in $\mathcal{L}_G\mathcal{F}$ depends on w. This means that $\mathcal{L}_G\mathcal{F}$ only contains the zero function (the Lie derivative along G of any function in $\mathcal{F}$ vanishes). As a result, any order Lie derivative of any outputs, computed along $F^1, \cdots, F^{m_u}$ (which are the vector fields defined in (27)) and at least once along G, vanishes. This means that, in algorithm 4, we can ignore the contribution due to the Lie derivative along G. But this trivially means that the observable codistribution is precisely $\mathcal{DF}$ and the unknown input is spurious.

[10] Note that, from proposition 5, the functions in $\mathcal{F}$ are weakly observable.

References

[1] G. Basile and G. Marro. On the observability of linear, time invariant systems with unknown inputs. J. Optimization Theory Appl., 3:410–415, 1969.

[2] F. A. W. Belo, P. Salaris, and A. Bicchi, 3 Known Landmarks are Enough for Solving Planar Bearing SLAM and Fully Reconstruct Unknown Inputs, IROS 2010, Taipei, Taiwan

[3] S.P. Bhattacharyya Observer design for linear systems with unknown inputs IEEE Trans. on Aut. Control, 23 (1978)

[4] Bryson, M. and Sukkarieh, S., Observability analysis and active control for airborne SLAM, IEEE Transactions on Aerospace and Electronic Systems, Vol 44 , Issue 1, pp 261–280, January 2008

[5] A. Censi, Antonio Franchi, Luca Marchionni, and Giuseppe Oriolo. Simultaneous calibration of odometry and sensor parameters for mobile robots. IEEE Transactions on Robotics, April 2013.

[6] R. Guidorzi and G. Marro. On Wonham stabilizability condition in the synthesis of observers for unknown-input systems. Automatic Control, IEEE Transactions on, 16(5):499–500, oct 1971.

[7] Chao X. Guo and Stergios I. Roumeliotis, IMU-RGBD Camera 3D Pose Estimation and Extrinsic Calibration: Observability Analysis and Consistency Improvement, ICRA 2013, Karlsruhe, Germany

[8] Hermann R. and Krener A.J., 1977, Nonlinear Controllability and Observability, Transaction On Automatic Control, AC-22(5): 728–740

[9] J. A. Hesch, D. G. Kottas, S. L. Bowman, and S. I. Roumeliotis, Towards consistent vision-aided inertial navigation, Int. Workshop on the Algorithmic Foundations of Robotics, Cambridge, MA, Jun, 2012.

[10] J.A. Hesch, D.G. Kottas, S.L. Bowman, and S.I. Roumeliotis, "Camera-IMU-based Localization: Observability Analysis and Consistency Improvement," International Journal of Robotics Research, 33(1), Jan. 2014, pp. 182-201.

[11] J.A. Hesch, D.G. Kottas, S.L. Bowman, and S.I. Roumeliotis, "Consistency Analysis and Improvement of Vision-aided Inertial Navigation," IEEE Transactions on Robotics, 30(1), Feb. 2014, pp. 158-176.

[12] Huang, Guoquan P., Mourikis, A.I. and Roumeliotis, S.I., Analysis and improvement of the consistency of extended Kalman filter based SLAM, IEEE International Conference on Robotics and Automation, 2008, 19-23 May 2008, Pasadena, CA.

[13] G. Huang, A.I. Mourikis, S.I. Roumeliotis: "Observability-based Rules for Designing Consistent EKF SLAM Estimators," International Journal of Robotics Research, 29(5), pp. 502-528, 2010.

[14] G. Huang, A.I. Mourikis, S.I. Roumeliotis: "A Quadratic-Complexity Observability-Constrained Unscented Kalman Filter for SLAM," IEEE Transactions on Robotics, 29(5), pp. 1226-1243, 2013.

[15] B. R. Hunt, T. Sauer and J. A. Yorke, Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces, BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY Volume 27, Number 2, October 1992

[16] Isidori A., Nonlinear Control Systems, 3rd ed., London, Springer Verlag, 1995.

[17] E. Jones and S. Soatto, "Visual-inertial navigation, mapping and localization: A scalable real-time causal approach", The International Journal of Robotics Research, vol. 30, no. 4, pp. 407–430, Apr. 2011.

[18] J. Kelly and G. Sukhatme, Visual-inertial simultaneous localization, mapping and sensor-to-sensor self-calibration, Int. Journal of Robotics Research, vol. 30, no. 1, pp. 56–79, 2011.

[19] D. G. Kottas, J. A. Hesch, S. L. Bowman, and S. I. Roumeliotis, On the consistency of vision-aided inertial navigation, in Proc. of the Int. Symposium on Experimental Robotics, Canada, Jun 2012.

[20] Kwang Wee Lee, Wijesoma, W.S. and Javier, I.G., On the Observability and Observability Analysis of SLAM, IEEE/RSJ International Conference on Intelligent Robots and Systems, 9-15 Oct. 2006, Beijing, China.

[21] M. Li and A. I. Mourikis, Improving the accuracy of EKF-based visualinertial odometry, ICRA 2012, St. Paul, MN.

[22] M. Li, A.I. Mourikis: "Online Temporal Calibration for Camera-IMU Systems: Theory and Algorithms," International Journal of Robotics Research, 33(7), pp. 947-964, 2014.

[23] Martinelli, A., Scaramuzza, D., Siegwart, R. Automatic Self-Calibration of a Vision System during Robot Motion. IEEE International Conference on Robotics and Automation (ICRA 2006), Orlando, USA, May 2006.

[24] A. Martinelli, State Estimation Based on the Concept of Continuous Symmetry and Observability Analysis: the Case of Calibration, IEEE Transactions on Robotics, vol. 27, no. 2, pp. 239–255, 2011

[25] A. Martinelli, Vision and IMU data fusion: closed-form solutions for attitude, speed, absolute scale and bias determination, IEEE Transactions on Robotics, Volume 28 (2012), Issue 1 (February), pp 44–60.

[26] A. Martinelli, Observability Properties and Deterministic Algorithms in Visual-Inertial Structure from Motion, Found. and Trends in Robotics, Vol 3, 139–209.

[27] A. Martinelli, Nonlinear Unknown Input Observability: Analytical expression of the observable codistribution in the case of a single unknown input, SIAM - CT15 Symposium, 2015, Paris, France.

[28] A. Martinelli, Minimalistic sensor design in visual-inertial structure from motion, 2015 IEEE/RSJ International Conference on Robotics and Automation, Seattle (USA).

[29] A. Martinelli, Visual-inertial structure from motion: observability vs minimum number of sensors, 2014 IEEE/RSJ International Conference on Robotics and Automation, Hong Kong.

[30] Mirzaei F.M. and Roumeliotis S.I., A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation, Transactions on Robotics, 2008, Vol. 24, No. 5, 1143–1156

[31] F.M. Mirzaei, D.G. Kottas, and S.I. Roumeliotis, 3D Lidar-Camera Intrinsic and Extrinsic Calibration: Observability Analysis and Analytical Least Squares-based Initialization, International Journal of Robotics Research, Special Issue on Robot Vision, 31(4), Apr. 2012, pp. 452-467.

[32] G. Panahandeh, C. X. Guo, M. Jansson, and S. I. Roumeliotis, "Observability Analysis of a Vision-Aided Inertial Navigation System using Planar Features on the Ground", in Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4187-4194, Tokyo, Japan, Nov. 3-7, 2013.

[33] Perera, L.D.L., Melkumyan, A. and Nettleton, E., On the linear and nonlinear observability analysis of the SLAM problem, IEEE International Conference on Mechatronics, 14-17 April 2009, Malaga.

[34] N. Trawny, G. Huang, A.I. Mourikis, S.I. Roumeliotis: "Observability-based Consistent EKF Estimators for Multi-robot Cooperative Localization," Autonomous Robots, 30(1), pp. 99-122, 2011.

[35] Vidal-Calleja, T., Bryson, M., Sukkarieh, S., Sanfeliu, Alberto, Andrade-Cetto, J., On the Observability of Bearing-only SLAM, EEE International Conference on Robotics and Automation, 10-14 April 2007, Roma.

[36] S.H. Wang, E.J. Davison, P. Dorato Observing the states of systems with unmeasurable disturbance IEEE Transactions on Automatic Control, 20 (1975)

[37] Wang Z. and Dissanayake, G., Observability analysis of SLAM using fisher information matrix, International Conference on Control, Automation, Robotics and Vision, 2008, Hanoi.

Contents

1 Introduction **3**

2 System definition **5**

3 Observable Codistribution **6**
3.1 Observable codistribution in the case without unknown inputs 6
3.2 Observable codistribution in presence of a single unknown input 7

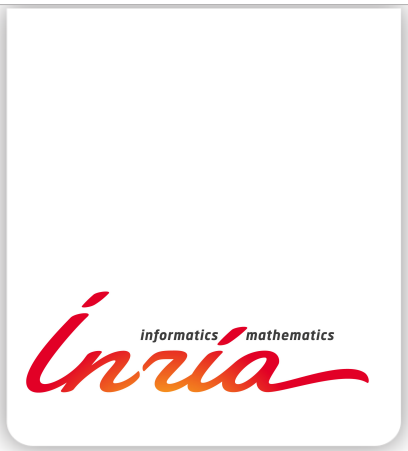
4 The Analytic Criterion **9**

5 Applications **10**
5.1 Unicycle with one input unknown 10
5.1.1 The system 10
5.1.2 Intuitive procedure to obtain the observability properties 11
5.1.3 Analytic results 12
5.2 Unicycle in presence of an external unknown input 16
5.2.1 The system 16
5.2.2 Observability properties when γ is known 17
5.2.3 Observability properties when γ is unknown 19
5.2.4 Simulations 21
5.3 Vehicle moving in 3D in presence of a disturbance 24
5.3.1 The system 24
5.3.2 Observability in absence of disturbance 26
5.3.3 Observability in presence of the disturbance 27
5.3.4 Simulations 29

6 Analytic Derivations **34**
6.1 State Indistinguishability, State augmentation and Observable Codistribution . . 34
6.2 Proof of the validity of the analytic criterion 39
6.2.1 Separation 40
6.2.2 Convergence 43
6.2.3 Extension to the case of multiple known inputs 47

7 Conclusion **49**

A The case when $L_g^1 = 0$ **51**



RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

Inovallée
655 avenue de l'Europe Montbonnot
38334 Saint Ismier Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399