



Nonlinear Unknown Input Observability: Extension of the Observability Rank Condition and the Case of a Single Unknown Input

Agostino Martinelli

► To cite this version:

Agostino Martinelli. Nonlinear Unknown Input Observability: Extension of the Observability Rank Condition and the Case of a Single Unknown Input. [Research Report] RR-8608, INRIA Grenoble. 2014. hal-01071314v5

HAL Id: hal-01071314

<https://inria.hal.science/hal-01071314v5>

Submitted on 24 Mar 2016

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

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



Nonlinear Unknown Input Observability: Extension of the Observability Rank Condition and the Case of a Single Unknown Input

Agostino Martinelli

**RESEARCH
REPORT**

N° 8608

October 2014

Project-Team Emotion



Nonlinear Unknown Input Observability: Extension of the Observability Rank Condition and the Case of a Single Unknown Input

Agostino Martinelli

Project-Team Emotion

Research Report n° 8608 — version 4 — initial version October 2014 —
revised version 2 February 2015 — 30 pages

Abstract: This paper investigates the unknown input observability problem in the nonlinear case under the assumption that the unknown inputs are differentiable functions of time (up to a given order). The goal is not to design new observers but to provide simple analytic conditions in order to check the weak local observability of the state. The analysis starts by extending the observability rank condition. This is obtained by a state augmentation and is called the extended observability rank condition (first contribution). The proposed extension of the observability rank condition only provides sufficient conditions for the state observability. On the other hand, in the case of a single unknown input, the paper provides a simple algorithm to directly obtain the entire observable codistribution (second and main contribution). As in the standard case of only known inputs, the observable codistribution is obtained by recursively computing the Lie derivatives 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 must be computed also 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. However, the analytic derivations required to prove that this codistribution fully characterizes the weak local observability of the state are complex. 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. Also this proof is based on several tricky analytical steps. Several applications illustrate the derived theoretical results, both in the case of a single unknown input and in the case of multiple unknown inputs.

Key-words: nonlinear observability; Unknown Input Observability; vision aided inertial navigation

**RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES**

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

Observabilité non linéaire avec entrées inconnues: Extended Observability Rank Criterion et le cas d'une seule entrée inconnue

Résumé : Pas de résumé

Mots-clés : observabilité non linéaire; Observabilité avec entrées inconnues; vision par ordinateur; navigation inertielle

1 Introduction

The problem of state observability for systems driven by unknown inputs (UI) is a fundamental problem in control theory. This problem was introduced and firstly investigated in the seventies [3, 6, 15, 39]. A huge effort has then been devoted to design observers for both linear and nonlinear systems in presence of UI, e.g., [1, 2, 8, 9, 12, 13, 14, 17, 18, 21, 25, 26, 41].

The goal of this paper is not to design new observers for systems driven by UI but to provide simple analytic conditions in order to check the weak local observability of the state. The obtained results hold for systems whose dynamics are nonlinear in the state and affine in both the known and the unknown inputs. Additionally, the unknown inputs are supposed to be smooth functions of time (specifically, they are supposed to be \mathcal{C}^k , for a suitable integer k).

In [19] the observability properties of a nonlinear system are derived starting from the definition of indistinguishable states. According to this definition, 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. Hence, if a given state x belongs to the indistinguishable set of a state x_0 (i.e., to I_{x_0}) all the Lie derivatives computed at x and at x_0 take the same values. This is a fundamental property. In particular, based on this property, the observability rank condition was introduced in [19].

Our first objective is to extend the observability rank condition. For, we introduce a new definition of *indistinguishable states* for the case UI (section 2). Then, in section 3 we introduce a new system by a suitable state extension. For this extended system, we show that, the Lie derivatives of the outputs up to a given order, take the same values at the states which are indistinguishable. In other words, the new system satisfies the same property derived in [19] mentioned above and this allows us to extend the observability rank condition (section 4). We will refer to this extension as to the Extended Observability Rank Condition (*EORC*).

The new system is obtained by a state augmentation. In particular, the augmented state is obtained by including the unknown inputs together with their time-derivatives up to given order. This augmented state has already been considered in the past. Specifically, in [4] the authors adopted this augmented state to investigate the observability properties of a fundamental problem in the framework of mobile robotics (the bearing SLAM). In particular, starting from the idea of including the time-derivatives of the unknown input in the state, in [4] a sufficient condition for the state observability has been provided.

The *EORC* is based on the computation of a codistribution defined in the augmented space. In other words, the *EORC* allows us to check the weak local observability of the original state together with its extension and not directly of the original state. This makes the computational cost dependent on the dimension of the augmented state. Additionally, the *EORC* only provides sufficient conditions for the weak local observability of the original state since the state augmentation can be continued indefinitely. For these reasons, the paper focuses on the following two fundamental issues:

- Understanding if it is possible to derive the weak local observability of the original state by computing a codistribution defined in the original space, namely a codistribution consisting of covectors with the same dimension of the original state.
- Understanding if there exists a given augmented state such that, by further augmenting the state, the observability properties of the original state provided by *EORC* remain unvaried.

Both these issues have been fully addressed in the case of a single unknown input (see theorems 1 and 2).

Thanks to the result stated by theorem 1 (section 6), the algorithm in definition 3 in section 6 (for the case of a single known input) and in definition 5 in section 8 (for the case of multiple known inputs) can be used to obtain the entire observable codistribution. In other words, the observability properties of the original state are obtained by a very simple algorithm. As it will be seen, the analytic derivations required to prove theorem 1 are complex and we are currently extending them to the multiple unknown inputs case. Theorem 2 (section 7) ensures the convergence of the algorithm in a finite number of steps and it also provides the criterion to establish that this convergence has been reached. Also this proof is based on several tricky and complex analytical steps.

Both theorems 1 and 2 are first proved in the case of a single known input (sections 6 and 7) but in section 8 their validity is extended to the case of multiple known inputs.

All the theoretical results are illustrated in section 9 by deriving the observability properties of several nonlinear systems driven by unknown inputs.

2 Basic Definitions

In the sequel we will refer to a nonlinear control system with m_u known inputs ($u \equiv [u_1, \dots, u_{m_u}]^T$) and m_w unknown inputs or disturbances ($w \equiv [w_1, \dots, w_{m_w}]^T$). 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 (the extension to multiple outputs is straightforward). Our system is characterized by the following equations:

$$\begin{cases} \dot{x} = f_0(x) + \sum_{i=1}^{m_u} f_i(x)u_i + \sum_{j=1}^{m_w} g_j(x)w_j \\ y = h(x) \end{cases} \quad (1)$$

where $f_i(x)$, $i = 0, 1, \dots, m_u$, and $g_j(x)$, $j = 1, \dots, m_w$, are vector fields in M and the function $h(x)$ is a scalar function defined on the open set M . For the sake of simplicity, we will assume that all these functions are analytic functions in M .

Let us consider the time interval $\mathcal{I} \equiv [0, T]$. Note that, since the equations in (1) do not depend explicitly on time, this can be considered as a general time interval of length T . In the sequel, we will assume that the solution of (1) exists in \mathcal{I} and we will denote by $x(t; x_0; u; w)$ the state at a given time $t \in \mathcal{I}$, when $x(0) = x_0$ and the known input and the disturbance are $u(t)$ and $w(t)$, respectively, $\forall t \in \mathcal{I}$.

We introduce the following definition:

Definition 1 (Indistinguishable states in presence of UI) *Two states x_a and x_b are indistinguishable if, for any $u(t)$ (the known input vector function), there exist $w_a(t)$ and $w_b(t)$ (i.e., two unknown input vector functions in general, but not necessarily, different from each other) such that $h(x(t; x_a; u; w_a)) = h(x(t; x_b; u; w_b)) \forall t \in \mathcal{I}$.*

This definition states that, if x_a and x_b are indistinguishable, then, for any known input, by looking at the output during the time interval \mathcal{I} , we cannot conclude if the initial state was x_a and the disturbance w_a or if the initial state was x_b and the disturbance w_b . We remark that, contrary to the definition of indistinguishable states in the case without disturbances, the new definition does not establish an equivalence relation. Indeed, we can have x_a and x_b indistinguishable, x_b and x_c indistinguishable but x_a and x_c are not indistinguishable. As in the

case of known inputs, given x_0 , the indistinguishable set I_{x_0} is the set of all the states x such that x and x_0 are indistinguishable. Starting from this definition, we can use exactly the same definitions of observability and weak local observability adopted in the case without disturbances.

3 Extended system and basic properties

The results contained in this and in the next section are also available in [35].

In order to extend the observability rank condition to the case of unknown inputs we introduce a new system (the extended system) such that its Lie derivatives are constant on the indistinguishable sets. The new system will be denoted by $\Sigma^{(k)}$. It is simply obtained by extending 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 time derivatives up to the $(k-1)$ -order:

$${}^k x \equiv [x^T, w^T, w^{(1)T}, \dots, w^{(k-1)T}]^T \quad (2)$$

where $w^{(k)} \equiv \frac{d^k w}{dt^k}$ and ${}^k x \in M^{(k)}$, with $M^{(k)}$ an open set of \mathbb{R}^{n+km_w} . From (1) it is immediate to obtain the dynamics for the extended state:

$${}^k \dot{x} = f_0^{(k)}({}^k x) + \sum_{i=1}^{m_u} f_i^{(k)}(x) u_i + \sum_{j=1}^{m_w} 1_{n+km_w}^{n+(k-1)m_w+j} w_j^{(k)} \quad (3)$$

where:

$$f_0^{(k)}({}^k x) \equiv \begin{bmatrix} f_0(x) + \sum_{i=1}^{m_w} g_i(x) w_i \\ w^{(1)} \\ w^{(2)} \\ \vdots \\ w^{(k-1)} \\ 0_{m_w} \end{bmatrix} \quad (4)$$

$$f_i^{(k)}(x) \equiv \begin{bmatrix} f_i(x) \\ 0_{km_w} \end{bmatrix} \quad (5)$$

and we denoted by 0_m the m -dimensional zero column vector and by 1_m^l the m -dimensional unit column vector, with 1 in the l^{th} position and 0 elsewhere. We remark that the resulting system has still m_u known inputs and m_w disturbances. 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 disturbances. The state evolution depends on the known inputs via the vector fields $f_i^{(k)}$, ($i = 1, \dots, m_u$) and it depends on the disturbances via the unit vectors $1_{n+km_w}^{n+(k-1)m_w+j}$, ($j = 1, \dots, m_w$). Finally, we remark that only the vector field $f_0^{(k)}$ depends on the new state elements.

In the rest of this section we derive several properties satisfied by $\Sigma^{(k)}$.

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

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 $g(x, w, w^{(1)}, \dots, w^{(m-1)})$. Hence, $\nabla g = [\frac{\partial g}{\partial x}, \frac{\partial g}{\partial w}, \frac{\partial g}{\partial w^{(1)}}, \dots, \frac{\partial g}{\partial w^{(m-1)}}]$. It is immediate to realize that the product of this gradient by any vector filed in (3) 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 are the following two properties:

Proposition 1 *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 $1_{n+(k-1)m_w+j}^{n+(k-1)m_w+j}$ ($j = 1, \dots, m_w$) 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 (2). Then, the proof follows from the fact that any vector among $1_{n+(k-1)m_w+j}^{n+(k-1)m_w+j}$ ($j = 1, \dots, m_w$) has the first $n + (k-1)m_w$ components equal to zero ■

Proposition 2 *The Lie derivatives of the output up to the k^{th} order along any vector field $f_0^{(k)}, f_1^{(k)}, \dots, f_{m_u}^{(k)}$ 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 $f_0^{(k)}, f_1^{(k)}, \dots, f_{m_u}^{(k)}$ coincide with the first $n + mm_w$ components of $f_0^{(k+1)}, f_1^{(k+1)}, \dots, f_{m_u}^{(k+1)}$ when $m < k$ ■

In the sequel we will use the notation: $\xi \equiv [w^T, w^{(1)T}, \dots, w^{(k-1)T}]^T$. 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 (1).

It holds the following fundamental result:

Proposition 3 *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 ($i = 1, \dots, m_u$):

$$\tilde{u}_i(t) = \begin{cases} u_i^1 & t \in [0, t_1) \\ u_i^2 & t \in [t_1, t_1 + t_2) \\ \dots & \\ u_i^g & t \in [t_1 + t_2 + \dots + t_{g-1}, t_1 + t_2 + \dots + t_{g-1} + t_g) \end{cases} \quad (6)$$

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

$$h(x(t; [x_a, \xi_a]; \tilde{u}; w_a^{(k)})) = h(x(t; [x_b, \xi_b]; \tilde{u}; w_b^{(k)})) \quad (7)$$

$\forall t \in [0, t_1 + t_2 + \dots + t_{g-1} + t_g) \subset \mathcal{I}$. On the other hand, by taking the two quantities in (7) at $t = t_1 + t_2 + \dots + t_{g-1} + t_g$, we can consider them as functions of the g arguments t_1, t_2, \dots, t_g . Hence, by differentiating with respect to all these variables, we also have:

$$\begin{aligned} \frac{\partial^g h(x(t_1 + \dots + t_g; [x_a, \xi_a]; \tilde{u}; w_a^{(k)}))}{\partial t_1 \partial t_2 \dots \partial t_g} &= \\ &= \frac{\partial^g h(x(t_1 + \dots + t_g; [x_b, \xi_b]; \tilde{u}; w_b^{(k)}))}{\partial t_1 \partial t_2 \dots \partial t_g} \end{aligned} \quad (8)$$

By computing the previous derivatives at $t_1 = t_2 = \dots = t_g = 0$ and by using proposition 1 we obtain, if $g \leq k$:

$$\mathcal{L}_{\theta_1 \theta_2 \dots \theta_g}^g h \Big|_{\substack{x = x_a \\ \xi = \xi_a}} = \mathcal{L}_{\theta_1 \theta_2 \dots \theta_g}^g h \Big|_{\substack{x = x_b \\ \xi = \xi_b}} \quad (9)$$

where $\theta_h = f_0^{(k)} + \sum_{i=1}^{m_u} f_i^{(k)} u_i^h$, $h = 1, \dots, g$. The equality in (9) must hold for all possible choices of $u_1^h, \dots, u_{m_u}^h$. By appropriately selecting these $u_1^h, \dots, u_{m_u}^h$, we finally obtain:

$$\mathcal{L}_{v_1 v_2 \dots v_g}^g h \Big|_{\substack{x = x_a \\ \xi = \xi_a}} = \mathcal{L}_{v_1 v_2 \dots v_g}^g h \Big|_{\substack{x = x_b \\ \xi = \xi_b}} \quad (10)$$

where $v_1 v_2 \dots v_g$ are vector fields belonging to the set $\{f_0^{(k)}, f_1^{(k)}, \dots, f_{m_u}^{(k)}\}$ ■

In [19] 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 3 also holds if x_a and x_b are V -indistinguishable.*

Thanks to the results stated by propositions 2 and 3 we will introduce the extension of the observability rank condition in the next section.

4 Extension of the Observability Rank condition

According to the observability rank condition, the weak local observability of the system in (1) with $m_w = 0$ at a given point x_0 can be investigated by analyzing the codistribution generated by the gradients of the Lie derivatives of its output. Specifically, if the dimension of this codistribution is equal to the dimension of the state on a given neighbourhood of x_0 , we conclude that the state is weakly locally observable at x_0 (theorem 3.1 in [19]). We can also check the weak local observability of a subset of the state components. Specifically, a given component of the state is weakly locally observable at x_0 , if its gradient belongs to the aforementioned codistribution¹. The proof of theorem 3.1 in [19] is based on the fact that all the Lie derivatives (up to any order) of the output computed along any direction allowed by the system dynamics take the same values at the states which are indistinguishable.

¹A component of the state is observable at x_0 if it is constant on the indistinguishable set I_{x_0} .

Let us consider now the general case, i.e., when $m_w \neq 0$. In the extended system $(\Sigma^{(k)})$ we know that the Lie derivatives up to the k -order satisfy the same property (see proposition 3). Therefore, we can extend the validity of theorem 3.1 in [19] to our case, provided that we suitably augment the state and that we only include the Lie derivatives up to the k -order to build the observable codistribution.

In the sequel, we will introduce the following notation:

- $\bar{\Omega}_m$ will denote the observable codistribution for $\Sigma^{(k)}$ that includes all the Lie derivatives of the output along $f_0^{(k)}, f_1^{(k)}, \dots, f_{m_u}^{(k)}$ up to the order $m \leq k$;
- The symbol d will denote the gradient with respect to the extended state in (2) and the symbol d_x will denote the gradient only respect to x ;
- For a given codistribution Λ and a given vector field η , we will denote by $\mathcal{L}_\eta \Lambda$ the codistribution whose covectors are the Lie derivatives along η of the covectors in Λ (we are obviously assuming that the dimension of these covectors coincides with the dimension of η).
- Given two vector spaces V_1 and V_2 , we will denote with $V_1 + V_2$ their sum, i.e., the span of all the generators of both V_1 and V_2 .
- For a given $V \subseteq M^{(k)}$ and a given $[x_0, \xi_0] \in V$, we will denote with $I_{[x_0, \xi_0]}^V$ the set of all the states V -indistinguishable from $[x_0, \xi_0]$.

The codistribution $\bar{\Omega}_m$ can be computed recursively by the following algorithm:

Algorithm 1 (Computation of $\bar{\Omega}_m$, $m \leq k$)

```

Set  $\bar{\Omega}_0 = \text{span}\{dh\}$ ;
Set  $i = 0$ 
while  $i < m$  do
  Set  $i = i + 1$ 
  Set  $\bar{\Omega}_i = \bar{\Omega}_{i-1} + \sum_{i'=0}^{m_u} \mathcal{L}_{f_{i'}^{(k)}} \bar{\Omega}_{i-1}$ 
end while

```

Let us denote by x_j the j^{th} component of the state ($j = 1, \dots, n$). We introduce the following definition:

Definition 2 (EORC) For the system $\Sigma^{(k)}$, the j^{th} component of the state (i.e., x_j , $j = 1, \dots, n$) satisfies the extended observability rank condition at $[x_0, \xi_0]$ if $dx_j \in \bar{\Omega}_k$ at $[x_0, \xi_0]$. If this holds $\forall j = 1, \dots, n$, we say that the state x satisfies the extended observability rank condition at $[x_0, \xi_0]$ in $\Sigma^{(k)}$.

We have the following result, which is the extension of the result stated by theorem 3.1 in [19]:

Proposition 4 For $\Sigma^{(k)}$, if x_j ($j = 1, \dots, n$) satisfies the observability rank condition at $[x_0, \xi_0]$, then x_j is weakly locally observable at $[x_0, \xi_0]$. Additionally, x_j remains weakly locally observable by further extending the state (i.e., in every system $\Sigma^{(f)}$ ($f > k$)).

Proof: We prove that it exists an open neighbourhood U of $[x_0, \xi_0]$ such that, for every open neighbourhood $V \subseteq U$ of $[x_0, \xi_0]$, x_j is constant on the set $I_{[x_0, \xi_0]}^V$. Since $dx_j \in \bar{\Omega}_k$ at $[x_0, \xi_0]$, it exists some open neighborhood U of $[x_0, \xi_0]$, such that x_j can be expressed in terms of the Lie derivatives of h along the directions $f_{i'}^{(k)}$ ($i' = 0, 1, \dots, m_u$) up to the k order. If $V \subseteq U$ is an

open neighborhood of $[x_0, \xi_0]$, then proposition 3 and remark 1 imply that all the Lie derivatives up to the k order are constant on the set $I_{[x_0, \xi_0]}^V$ and, consequently, also x_j is constant on this set. Finally, the fact that x_j is weakly locally observable in every system $\Sigma^{(f)}$ ($f > k$) directly follows from proposition 2 ■

In accordance with the previous result, the *EORC* is a tool to analyze the observability properties of a nonlinear system driven by known and unknown inputs. However, we remark two important limitations of the *EORC*. The former consists in the fact that the state augmentation can be continued indefinitely. As a result, the *EORC* only provides sufficient conditions for the weak local observability of the state components. The latter regards the computational cost demanded to check if it is satisfied. Specifically, the computation demanded to check if dx_j belongs to $\bar{\Omega}_k$ can be very complex because by increasing k we also increase the dimension of the extended state.

In the rest of this paper, we will focus our attention on these fundamental issues and we will provide the main paper contributions:

- obtaining a new codistribution (Ω_k) that is the span of covectors whose dimension is n (i.e., independent of the state extension) such that $dx_j \in \Omega_k$ if and only if $dx_j \in \bar{\Omega}_k$;
- understanding if there exists a given \hat{k} such that, if $dx_j \notin \bar{\Omega}_{\hat{k}}$, then $dx_j \notin \bar{\Omega}_k \forall k > \hat{k}$.

We fully address both these issues in the case $m_w = 1$. In section 5 we introduce the basic equations that characterize this case. In section 6 we provide a complete answer to the first issue by operating a separation on the codistribution generated by all the Lie derivatives up to the k -order. Specifically, we prove that the observable codistribution can be splitted into two codistributions. The former is generated by the gradients of scalar functions that only depend on the original state. The latter is generated by the gradients of scalar functions that depend on the entire augmented state. However, this latter codistribution can be ignored when deriving the observability properties of the original state. The former codistribution, namely the one generated by the gradients of scalar functions that only depend on the original state, is defined by a simple recursive algorithm. In section 7 we provide a complete answer to the second issue by proving that this algorithm converges in a finite number of steps and by also providing the criterion to establish that the convergence of the algorithm has been reached (theorem 2). Also this proof is based on several tricky analytic steps. For the sake of clarity, we start this discussion by considering the case when the system is characterized by a single known input, i.e., when $m_u = 1$ (sections 5, 6 and 7). In particular, both theorems 1 and 2 are proved in this simplified case. However, in section 8, their validity is extended to the case of multiple known inputs (i.e., $\forall m_u > 1$).

5 Single known Input and single disturbance

We will refer to the following system:

$$\begin{cases} \dot{x} = f(x)u + g(x)w \\ y = h(x) \end{cases} \quad (11)$$

In other words, we consider the case when f_0 is the null vector and $m_u = m_w = 1$. In this case, the extended state that includes the time derivatives up to the $(k-1)$ -order is:

$${}^k x \equiv [x^T, w, w^{(1)}, \dots, w^{(k-1)}]^T \quad (12)$$

The dimension of the extended state is $n + k$. From (11) it is immediate to obtain the dynamics for the extended state:

$${}^k\dot{x} = G({}^kx) + F(x)u + 1_{n+k}^{n+k}w^{(k)} \quad (13)$$

where:

$$F \equiv \begin{bmatrix} f(x) \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \end{bmatrix} \quad G \equiv \begin{bmatrix} g(x)w \\ w^{(1)} \\ w^{(2)} \\ \dots \\ w^{(k-1)} \\ 0 \end{bmatrix} \quad (14)$$

In the sequel, we will denote by L_g^1 the first order Lie derivative of the function $h(x)$ along the vector field $g(x)$, i.e., $L_g^1 \equiv \mathcal{L}_g h$. The derivations provided in the next sections are based on the assumption that $L_g^1 \neq 0$ on a given neighbourhood of x_0 . We conclude this section by showing that, when this assumption does not hold, it is possible to introduce new coordinates and to show that the observability properties can be investigated starting from a new output that satisfies the assumption.

Let us suppose that $L_g^1 = 0$ on a given neighbourhood of x_0 . We introduce the following system associated with the system in (11):

$$\begin{cases} \dot{x} = f(x) + g(x)u \\ y = h(x) \end{cases} \quad (15)$$

This is a system without disturbances and with a single known input u . Let us denote by r the relative degree of this system at x_0 . Since $L_g^1 = 0$ on a given neighbourhood of x_0 , we have $r > 1$. Additionally, we can introduce the following new local coordinates (see proposition 4.1.3 in [22]):

$$x' = \mathcal{Q}(x) = \begin{bmatrix} \mathcal{Q}_1(x) \\ \dots \\ \mathcal{Q}_n(x) \end{bmatrix} \quad (16)$$

such that the first new r coordinates are:

$$\mathcal{Q}_1(x) = h(x), \quad \mathcal{Q}_2(x) = \mathcal{L}_f^1 h(x), \quad \dots, \quad \mathcal{Q}_r(x) = \mathcal{L}_f^{r-1} h(x) \quad (17)$$

Now let us derive the equations of the original system (i.e., the one in (11)) in these new coordinates. We have:

$$\begin{cases} \dot{x}' = \tilde{f}(x')u + \tilde{g}(x')w \\ y = x'_1 \end{cases} \quad (18)$$

where \tilde{f} and \tilde{g} have the following structure:

$$\tilde{f} \equiv \begin{bmatrix} x'_2 \\ x'_3 \\ \dots \\ x'_r \\ \tilde{f}_r(x') \\ \dots \\ \tilde{f}_n(x') \end{bmatrix} \quad \tilde{g} \equiv \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \\ \tilde{g}_r(x') \\ \dots \\ \tilde{g}_n(x') \end{bmatrix} \quad (19)$$

We remark that the first r components of x' are weakly locally observable since they are the output and its Lie derivatives along f up to the $(r-1)$ -order (note that we do not need to augment the state to use the first $(r-1)$ Lie derivatives because all the Lie derivatives up to the $(r-1)$ -order that includes at least one direction along g vanish automatically). In order to investigate the observability properties of the remaining components, we augment the state as in (12) and we can consider the new output $\tilde{h}(x') = x'_r$. We set $L_g^1 = \tilde{g}_r = \mathcal{L}_g \mathcal{L}_f^{r-1} h \neq 0$.

6 The observable codistribution (Ω)

In this section we operate a separation on the codistribution generated by all the Lie derivatives up to the m -order ($m \leq k$). Specifically, we prove that this codistribution can be splitted into two codistributions. The former is generated by the gradients of scalar functions that only depend on the original state. The latter is generated by the gradients of scalar functions that depend on the entire augmented state. However, this latter codistribution can be ignored when deriving the observability properties of the original state.

The observable codistribution is given by the algorithm 1 and, in this case $m_u = m_w = 1$, it reduces to the span of the gradients of all the Lie derivatives along F and G up to the k -order. Hence, for any $m \leq k$, it is obtained recursively by the following algorithm:

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

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

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

where the parenthesis $[\cdot, \cdot]$ denote the Lie brackets of vector fields. Similarly, we define $\Phi_m \in \mathbb{R}^{n+k}$ by the following algorithm:

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

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 $\check{\Phi}_m$ the vector in \mathbb{R}^n that contains the first n components of Φ_m . In other words, $\Phi_m \equiv [\check{\Phi}_m^T, 0_k^T]^T$. Additionally, we set $\hat{\phi}_m \equiv \begin{bmatrix} \phi_m \\ 0_k \end{bmatrix}$.

We define the Ω codistribution as follows (see definition 5 in section 8 for the case when $m_u > 1$):

Definition 3 (Ω codistribution, $m_u = m_w = 1$) *This codistribution is defined recursively by the following algorithm:*

1. $\Omega_0 = d_x h;$
2. $\Omega_m = \Omega_{m-1} + \mathcal{L}_f \Omega_{m-1} + \mathcal{L}_{\frac{g}{L_g^1}} \Omega_{m-1} + \mathcal{L}_{\phi_{m-1}} d_x h$

Note that this codistribution is completely integrable by construction. More importantly, its generators are the gradients of 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., $L^m \equiv \text{span}\{\mathcal{L}_G^1 dh, \mathcal{L}_G^2 dh, \dots, \mathcal{L}_G^m dh\}$.

We finally introduce the following codistribution:

Definition 4 ($\tilde{\Omega}$ codistribution) *This codistribution is defined as follows: $\tilde{\Omega}_m \equiv [\Omega_m, 0_k] + L^m$*

The codistribution $\tilde{\Omega}_m$ consists of two parts. Specifically, we can select a basis that consists of exact differentials that are the gradients 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 gradients $\mathcal{L}_G^1 dh, \mathcal{L}_G^2 dh, \dots, \mathcal{L}_G^m dh$. The second set of generators, i.e., the gradients $\mathcal{L}_G^1 dh, \mathcal{L}_G^2 dh, \dots, \mathcal{L}_G^m dh$, are m and, with respect to the first set, they are gradients of functions that also depend on the state extension $\xi = [w, w^{(1)}, \dots, w^{(m-1)}]^T$.

We have the following result:

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

Proof: The fact that $d_x x_j \in \Omega_m$ 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 $d_x x_j \in \Omega_m$. 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 $\omega_1^1, \omega_2^1, \dots, \omega_{N_1}^1$ are N_1 generators of $[\Omega_m, 0_k]$ and $\omega_1^2, \omega_2^2, \dots, \omega_{N_2}^2$ are N_2 generators of L^m . 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 $(n + j')^{\text{th}}$ entry equal to $L_g^1 \neq 0$ and this is not possible since $dx_j = [d_x x_j, 0_k]$ ■

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 will prove this fundamental theorem, stating that $\tilde{\Omega}_m = \bar{\Omega}_m$.

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

The proof of this theorem is complex and is based on several results that we prove before. Based on them, we provide the proof of the theorem at the end of this section.

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 $j = 1, \dots, m$, 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 $\check{\Phi}_m = \sum_{j=1}^m c_j^n (\mathcal{L}_G h, \mathcal{L}_G^2 h, \dots, \mathcal{L}_G^m h) \phi_j$, i.e., the vector $\check{\Phi}_m$ is a linear combination of the vectors ϕ_j ($j = 1, \dots, m$), 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: $\check{\Phi}_1 = \phi_1 \mathcal{L}_G h$.

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

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

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 $j = 2, \dots, m$ 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 $j = 1, \dots, m-1$ and $\bar{c}_m = 0$, we obtain $\check{\Phi}_m = \sum_{j=1}^m (\tilde{c}_j + \bar{c}_j) \phi_j$, which proves our assert since $c_j^n (\equiv \tilde{c}_j + \bar{c}_j)$ is a function of $\mathcal{L}_G h, \mathcal{L}_G^2 h, \dots, \mathcal{L}_G^m h$ ■
It also holds the following result:

Lemma 5 $\hat{\phi}_m = \sum_{j=1}^m b_j^n (\mathcal{L}_G h, \mathcal{L}_G^2 h, \dots, \mathcal{L}_G^m h) \Phi_j$, i.e., the vector $\hat{\phi}_m$ is a linear combination of the vectors Φ_j ($j = 1, \dots, m$), 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: $\hat{\phi}_1 = \Phi_1 \frac{1}{\mathcal{L}_G h}$.

Inductive step: Let us assume that $\hat{\phi}_{m-1} = \sum_{j=1}^{m-1} b_j (\mathcal{L}_G h, \mathcal{L}_G^2 h, \dots, \mathcal{L}_G^{m-1} h) \Phi_j$. We need to prove that $\hat{\phi}_m = \sum_{j=1}^m b_j^n (\mathcal{L}_G h, \mathcal{L}_G^2 h, \dots, \mathcal{L}_G^m h) \Phi_j$. We start by applying on both members of the equality $\hat{\phi}_{m-1} = \sum_{j=1}^{m-1} b_j (\mathcal{L}_G h, \mathcal{L}_G^2 h, \dots, \mathcal{L}_G^{m-1} h) \Phi_j$ 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:

$$\begin{aligned} \sum_{j=1}^{m-1} [b_j \Phi_j, G] &= \sum_{j=1}^{m-1} b_j [\Phi_j, G] - \sum_{j=1}^{m-1} \mathcal{L}_G b_j \Phi_j = \\ &= \sum_{j=1}^{m-1} b_j \Phi_{j+1} - \sum_{j=1}^{m-1} \sum_{i=1}^{m-1} \frac{\partial b_j}{\partial (\mathcal{L}_G^i h)} \mathcal{L}_G^{i+1} h \Phi_j \end{aligned}$$

By setting $\tilde{b}_j = \frac{b_{j-1}}{\mathcal{L}_G^1}$ for $j = 2, \dots, m$ and $\tilde{b}_1 = 0$, and by setting $\bar{b}_j = -\sum_{i=1}^{m-1} \frac{\partial b_j}{\partial (\mathcal{L}_G^i h)} \frac{\mathcal{L}_G^{i+1} h}{\mathcal{L}_G^1}$ for $j = 1, \dots, m-1$ 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 $b_j^n (\equiv \tilde{b}_j + \bar{b}_j)$ is a function of $\mathcal{L}_G h, \mathcal{L}_G^2 h, \dots, \mathcal{L}_G^m h$ ■

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

Proposition 5 *The following two codistributions coincide:*

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

We are now ready to prove theorem 1.

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: $\bar{\Omega}_m = \bar{\Omega}_{m-1} + \mathcal{L}_F \bar{\Omega}_{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1} = \bar{\Omega}_{m-1} + \mathcal{L}_F \bar{\Omega}_{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1} = \bar{\Omega}_{m-1} + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_F L^{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1}$. On the other hand, $\mathcal{L}_F L^{m-1} = \mathcal{L}_F \mathcal{L}_G^1 dh + \dots + \mathcal{L}_F \mathcal{L}_G^{m-2} dh + \mathcal{L}_F \mathcal{L}_G^{m-1} dh$. 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: $\bar{\Omega}_m = [\Omega_{m-1}, 0_k] + L^{m-1} + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_{\Phi_{m-1}} dh + \mathcal{L}_G [\Omega_{m-1}, 0_k] + \mathcal{L}_G L^{m-1} = [\Omega_{m-1}, 0_k] + L^m + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_{\Phi_{m-1}} dh + [\mathcal{L}_{\frac{g}{L_g^1}} \Omega_{m-1}, 0_k]$ and by using proposition 5 we obtain: $\bar{\Omega}_m = [\Omega_{m-1}, 0_k] + L^m + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_{\hat{\phi}_{m-1}} dh + [\mathcal{L}_{\frac{g}{L_g^1}} \Omega_{m-1}, 0_k] = \tilde{\Omega}_m$ ■

7 Convergence of the algorithm that defines Ω

Theorem 1 is fundamental. It allows us to obtain all the observability properties of the original state by restricting the computation to the Ω codistribution, 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. The Ω codistribution is defined recursively and $\Omega_m \subseteq \Omega_{m+1}$ (see definition 3 in section 6). This means that, if for a given m the gradients 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.

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 a useful expression that relates these vectors to the vectors ϕ_j , previously defined. Specifically we have:

Lemma 6 *It holds the following equation:*

$$\psi_j = \phi_j + \left\{ \sum_{i=0}^{j-1} (-)^{j-i} \mathcal{L}_{\frac{g}{L_g^1}}^{j-i-1} \left(\frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \right) \right\} \frac{g}{L_g^1} \quad (20)$$

Proof: We proceed by induction. By definition $\psi_0 = \phi_0 = f$ and equation (20) 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

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

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 (20) ■

We have the following result:

Lemma 7 For $i = 0, 1, \dots, m-2$, we have:

$$d_x \frac{\mathcal{L}_{\phi_i} L_g^1}{L_g^1} \in \Omega_m \quad (21)$$

Proof: By construction, $d_x \mathcal{L}_{\phi_i} h \in \Omega_m$, for any $i = 1, \dots, m-1$. 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 gradient of both members of this equation. Since $d_x \mathcal{L}_{\frac{g}{L_g^1}} \mathcal{L}_{\phi_{i-1}} h \in \Omega_m$, for any $i = 1, \dots, m-1$, also $d_x \frac{\mathcal{L}_{\phi_{i-1}} L_g^1}{L_g^1} \in \Omega_m$ ■

From lemma 6 with $j = 1, \dots, m-1$ and lemma 7 it is immediate to obtain the following result:

Proposition 6 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}_{ϕ_j} , $j = 1, \dots, m-1$.

Let us denote by $L_g^2 \equiv \mathcal{L}_g^2 h$ and by $\rho \equiv \frac{L_g^2}{(L_g^1)^2}$.

Lemma 8 *We have the following key equality:*

$$\mathcal{L}_{\phi_j} h = \mathcal{L}_{\phi_{j-2}} \rho + \rho \frac{\mathcal{L}_{\phi_{j-2}} L_g^1}{L_g^1} - \mathcal{L}_{\frac{g}{L_g^1}} \left(\frac{\mathcal{L}_{\phi_{j-2}} L_g^1}{L_g^1} + \mathcal{L}_{\phi_{j-1}} h \right) \quad (22)$$

$j \geq 2$.

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

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

The second term on the right hand side simplifies with the last term in (22). Hence we have to prove:

$$\frac{1}{L_g^1} \mathcal{L}_{\phi_{j-1}} L_g^1 = \mathcal{L}_{\phi_{j-2}} \rho + \rho \frac{\mathcal{L}_{\phi_{j-2}} L_g^1}{L_g^1} - \mathcal{L}_{\frac{g}{L_g^1}} \frac{\mathcal{L}_{\phi_{j-2}} L_g^1}{L_g^1} \quad (23)$$

We have:

$$\frac{1}{L_g^1} \mathcal{L}_{\phi_{j-1}} L_g^1 = \frac{1}{(L_g^1)^2} (\mathcal{L}_{\phi_{j-2}} L_g^2 - \mathcal{L}_g \mathcal{L}_{\phi_{j-2}} L_g^1) \quad (24)$$

We remark that:

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

and

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

By substituting these two last equalities in (24) we immediately obtain (23) ■

Lemma 9 *In general, it exists a finite m such that $d_x \rho \in \Omega_m$.*

Proof: For a given m , Ω_m contains all the covectors $d_x \mathcal{L}_{\phi_j} h$ ($j = 0, \dots, m-1$). From equation (22), we immediately obtain that, for a given $m \geq 3$, Ω_m contains the covectors ($j = 0, \dots, m-3$):

$$\mu_j \equiv d_x \mathcal{L}_{\phi_{j+2}} h = d_x \rho_j + \chi_j d_x \rho + \rho d_x \chi_j - \mathcal{L}_{\frac{g}{L_g^1}} (d_x \chi_j + d_x \mathcal{L}_{\phi_{j+1}} h) \quad (25)$$

with $d_x \rho_j \equiv d_x \mathcal{L}_{\phi_j} \rho$ and $\chi_j \equiv \frac{\mathcal{L}_{\phi_j} L_g^1}{L_g^1}$.

By using lemma 7 we obtain the following results:

- $\rho d_x \chi_j \in \Omega_m$, $j = 0, \dots, m-2$;
- $\mathcal{L}_{\frac{g}{L_g^1}} d_x \chi_j \in \Omega_m$, $j = 0, \dots, m-3$;

Additionally, also $\mathcal{L}_{\frac{g}{L_g^1}} d_x \mathcal{L}_{\phi_{j+1}} h \in \Omega_m$, $j = 0, \dots, m-3$. Hence, from (25), we obtain that Ω_m also contains the covectors ($j = 0, \dots, m-3$):

$$\mu'_j \equiv d_x \rho_j + \chi_j d_x \rho \quad (26)$$

Let us denote by j^* the smallest integer such that:

$$d_x \rho_{j^*} = \sum_{j=0}^{j^*-1} c_j d_x \rho_j + c_{-1} d_x h \quad (27)$$

Note that j^* is a finite integer and in particular $j^* \leq n-1$, where n is the dimension of the state. Indeed, if this would be not the case, the dimension of the codistribution generated by $d_x h, d_x \rho_0, d_x \rho_1, \dots, d_x \rho_{n-1}$ would be $n+1$, i.e., larger than n .

From (27) and (26) we obtain:

$$\mu'_{j^*} = \sum_{j=0}^{j^*-1} c_j d_x \rho_j + c_{-1} d_x h + \chi_{j^*} d_x \rho \quad (28)$$

From equation (26), for $j = 0, \dots, j^*-1$, we obtain: $d_x \rho_j = \mu'_j - \chi_j d_x \rho$. By substituting in (28) we obtain:

$$\mu'_{j^*} - \sum_{j=0}^{j^*-1} c_j \mu'_j - c_{-1} d_x h = \left(- \sum_{j=0}^{j^*-1} c_j \chi_j + \chi_{j^*} \right) d_x \rho \quad (29)$$

We remark that the left hand side consists of the sum of covectors that belong to Ω_m . Since in general $\chi_{j^*} \neq \sum_{j=0}^{j^*-1} c_j \chi_j$, we conclude that $d_x \rho \in \Omega_m$ ■

The previous lemma ensures that it exists a finite m such that $d_x \rho \in \Omega_m$. In particular, from the previous proof, it is possible to check that this value of m cannot exceed $n+2$. The following theorem allows us to obtain the criterion to stop the algorithm in definition 3:

Theorem 2 *If $d_x \rho \in \Omega_m$ 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 $d_x \mathcal{L}_{\phi_{m+p}} h \in \Omega_m$. 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 6, we obtain that Ω_m is also invariant under \mathcal{L}_{ϕ_j} , for $j = 1, 2, \dots, m+p-1$. Since $d_x \rho \in \Omega_m$ we have $d_x \mathcal{L}_{\phi_{m+p-2}} \rho \in \Omega_m$. Additionally, $d_x \mathcal{L}_{\phi_{m+p-1}} h \in \Omega_m$ and, because of lemma 7, we also have $d_x \frac{\mathcal{L}_{\phi_{m+p-2}} L_g^1}{L_g^1} \in \Omega_m$. Finally, because of the invariance under $\mathcal{L}_{\frac{g}{L_g^1}}$, also the

Lie derivatives along $\frac{g}{L_g^1}$ of $d_x \mathcal{L}_{\phi_{m+p-1}} h$ and $d_x \frac{\mathcal{L}_{\phi_{m+p-2}} L_g^1}{L_g^1}$ belong to Ω_m . Now, we use equation (22) for $j = m+p$. By computing the gradient of this equation it is immediate to obtain that $d_x \mathcal{L}_{\phi_{m+p}} h \in \Omega_m$ ■

We conclude this section by providing an upper bound for the number of steps necessary to achieve the convergence. The dimension of Ω_{j^*+2} is at least the dimension of the span of the covectors: $d_x h, d_x \mu'_0, d_x \mu'_1, \dots, d_x \mu'_{j^*-1}$. From the definition of j^* , we know that the vectors $d_x h, d_x \rho_0, d_x \rho_1, \dots, d_x \rho_{j^*-1}$ are independent meaning that the dimension of their span is j^*+1 . Hence, from (26), it easily follows that the dimension of the span of the vectors

$d_x h, d_x \mu'_0, d_x \mu'_1, \dots, d_x \mu'_{j^*-1}, d_x \rho$ 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 $m \leq n + 2$.

8 Extension to the case of multiple known inputs and method's summary

The previous two sections provide a complete answer to the problem of deriving all the observability properties of a system whose dynamics is driven by a single known input and a single unknown input and that depend non-linearly on the state and linearly on both the inputs. Before providing the steps to be followed in order to obtain the weak local observability properties of such a system, we remark that it is possible to extend our results to the case of multiple known inputs. This extension is simply obtained by re-defining the Ω codistribution.

We are referring to the nonlinear system characterized by the following equations:

$$\begin{cases} \dot{x} = \sum_{i=1}^{m_u} f_i(x)u_i + g(x)w \\ y = h(x) \end{cases} \quad (30)$$

The new Ω codistribution is defined as follows:

Definition 5 (Ω codistribution, $m_w = 1, \forall m_u$) *This codistribution is defined recursively by the following algorithm:*

1. $\Omega_0 = d_x h$;
2. $\Omega_m = \Omega_{m-1} + \sum_{i=1}^{m_u} \mathcal{L}_{f_i} \Omega_{m-1} + \mathcal{L}_{\frac{g}{L_g^1}} \Omega_{m-1} + \sum_{i=1}^{m_u} \mathcal{L}_{\phi_{m-1}^i} d_x h$

where the vectors $\phi_m^i \in \mathbb{R}^n$ ($i = 1, \dots, m_u$) are defined by the following algorithm:

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

It is immediate to repeat all the steps carried out in section 6 and extend the validity of theorem 1 to the system characterized by (30). This extension states that all the observability properties of the state that satisfies the nonlinear dynamics in (30) can be derived by analyzing the codistribution defined by definition 5. Finally, 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.

We conclude this section by outlining the steps to investigate the weak local observability at a given point x_0 of a nonlinear system driven by a single disturbance and several known inputs. In other words, to investigate the weak local observability of a system defined by a state that satisfies the dynamics in (30). The validity of the following procedure is a consequence of the theoretical results previously derived (in particular theorem 1 and theorem 2). This procedure is also available in [36].

1. For the chosen x_0 , compute $L_g^1(= \mathcal{L}_g^1 h)$ and $\rho \left(= \frac{\mathcal{L}_g^2 h}{(L_g^1)^2}\right)$. In the case when $L_g^1 = 0$, introduce new local coordinates, as explained at the end of section 5 and re-define the output².
2. Build the codistribution Ω_m (at x_0) by using the algorithm provided in definition 5, starting from $m = 0$ and, for each m , check if $d_x \rho \in \Omega_m$.
3. Denote by m' the smallest m such that $d_x \rho \in \Omega_m$.
4. 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 $m^* \leq n + 2$).
5. If the gradient of a given state component (x_j , $j = 1, \dots, n$) belongs to Ω^* (namely if $d_x x_j \in \Omega^*$) 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 .

9 Applications

We apply the theory developed in the previous sections in order to investigate the observability properties of several nonlinear systems driven by unknown inputs. In 9.1 we consider systems with a single disturbance, namely characterized by the equations given in (30). In this case we will use the results obtained in sections 5, 6, 7 and 8. In particular, we will follow the steps outlined at the end of section 8. In 9.2 we consider the case of multiple disturbances, i.e., when the state dynamics satisfy the first equation in (1). In this section, we also consider the case of multiple outputs and we use directly the *EORC*, as discussed in section 4.

9.1 Systems with a single disturbance

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 $[x_v, y_v, \theta_v]^T$ where x_v and y_v are the cartesian vehicle coordinates, and θ_v is the vehicle orientation. We assume that the dynamics of this vector satisfy the unicycle differential equations:

$$\begin{cases} \dot{x}_v = v \cos \theta_v \\ \dot{y}_v = v \sin \theta_v \\ \dot{\theta}_v = \omega \end{cases} \quad (31)$$

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

²Note that in the case of multiple known inputs, for the local coordinates we have the possibility to choose among the m_u functions f_i . The most convenient choice is the one that corresponds to the highest relative degree (if this degree coincides with n it means that the state is weakly locally observable and we do not need to pursue the observability analysis).

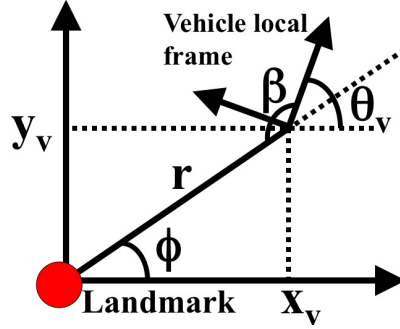


Figure 1: The vehicle state in cartesian and polar coordinates together with the three considered outputs.

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

We can analytically express the output in terms of the state. We remark that the expressions become very simple if we adopt polar coordinates: $r \equiv \sqrt{x_v^2 + y_v^2}$, $\phi = \text{atan}\frac{y_v}{x_v}$. We have, for the three cases, $y = r$, $y = \pi - (\theta_v - \phi)$ 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. The dynamics in these new coordinates become:

$$\begin{cases} \dot{r} = v \cos(\theta_v - \phi) \\ \dot{\phi} = \frac{v}{r} \sin(\theta_v - \phi) \\ \dot{\theta}_v = \omega \end{cases} \quad (32)$$

9.1.1 $y = r$, $u = \omega$, $w = v$

In this case we have $f = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and $g = \begin{bmatrix} \cos(\theta_v - \phi) \\ \frac{\sin(\theta_v - \phi)}{r} \\ 0 \end{bmatrix}$.

We follow the five steps mentioned at the end of section 8. We have $L_g^1 = \cos(\theta_v - \phi)$ and $\rho \equiv \frac{L_g^2}{(L_g^1)^2} = \frac{\tan^2(\theta_v - \phi)}{r}$. Additionally:

$$d_x \rho = \frac{\tan(\theta_v - \phi)}{r} \left[-\frac{\tan(\theta_v - \phi)}{r}, -\frac{2}{\cos^2(\theta_v - \phi)}, \frac{2}{\cos^2(\theta_v - \phi)} \right]$$

We also have $\Omega_0 = \text{span}\{[1, 0, 0]\}$. Hence, $d_x \rho \notin \Omega_0$. Additionally, $\Omega_1 = \Omega_0$. We need to compute

Ω_2 and, in order to do this, we need to compute ϕ_1 . We obtain: $\phi_1 = \begin{bmatrix} -\tan(\theta_v - \phi) \\ \frac{1}{r} \\ 0 \end{bmatrix}$ and

$\Omega_2 = \text{span} \left\{ [1, 0, 0], \left[0, \frac{1}{\cos^2(\theta_v - \phi)}, -\frac{1}{\cos^2(\theta_v - \phi)} \right] \right\}$. It is immediate to check that $d_x \rho \in \Omega_2$, meaning that $m' = 2$. Additionally, by a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$, whose dimension is 2. We conclude that the dimension of the observable space is equal to 2 and the state is not weakly locally observable.

9.1.2 $y = r, u = v, w = \omega$

In this case we have $f = \begin{bmatrix} \cos(\theta_v - \phi) \\ \frac{\sin(\theta_v - \phi)}{r} \\ 0 \end{bmatrix}$ and $g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

We follow the five steps mentioned at the end of section 8. We easily obtain $L_g^1 = 0$. Hence, we have to introduce new local coordinates, as explained at the end of section 5. We obtain $\mathcal{L}_f^1 h = \cos(\theta_v - \phi)$ and we obtain that the relative degree of the associated system in (15) is $r = 2$. Let us denote the new coordinates by x'_1, x'_2, x'_3 . In accordance with (16) and (17) we should set $x'_1 = r$ and $x'_2 = \cos(\theta_v - \phi)$. On the other hand, to simplify the computation, we set $x'_2 = \theta_v - \phi$. Finally, we set $x'_3 = \theta_v$. We compute the new vector fields that characterize the dynamics in the new coordinates. We have:

$$\tilde{f} \equiv \begin{bmatrix} \cos(x'_2) \\ -\frac{\sin(x'_2)}{x'_1} \\ 0 \end{bmatrix} \quad \tilde{g} \equiv \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad (33)$$

Additionally, we set $\tilde{h} = \cos(x'_2)$ and $\Omega_1 = \text{span}\{[1, 0, 0], [0, -\sin(x'_2), 0]\}$. In the new coordinates we obtain: $L_g^1 = -\sin(x'_2)$ and $\rho = -\frac{\cos(x'_2)}{\sin^2(x'_2)}$. It is immediate to check that $d_x \rho \in \Omega_1$, meaning that $m' = 1$. Additionally, by a direct computation, it is possible to check that $\Omega_2 = \Omega_1$ meaning that $m^* = 1$ and $\Omega^* = \Omega_1$, whose dimension is 2. We conclude that the dimension of the observable space is equal to 2 and the state is not weakly locally observable.

9.1.3 $y = \theta_v - \phi, u = \omega, w = v$

In this case we have $f = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and $g = \begin{bmatrix} \cos(\theta_v - \phi) \\ \frac{\sin(\theta_v - \phi)}{r} \\ 0 \end{bmatrix}$.

We follow the five steps mentioned at the end of section 8. We have $L_g^1 = -\frac{\sin(\theta_v - \phi)}{r}$ and $\rho = 2 \cot(\theta_v - \phi)$. Additionally:

$$d_x \rho = \frac{2}{\sin^2(\theta_v - \phi)} [0, 1, -1]$$

We also have $\Omega_0 = \text{span}\{[0, -1, 1]\}$. Hence, $d_x \rho \in \Omega_0$, meaning that $m' = 0$. Additionally, by a direct computation, it is possible to check that $\Omega_1 = \Omega_0$ meaning that $m^* = 0$ and $\Omega^* = \Omega_0$, whose dimension is 1. We conclude that the dimension of the observable space is equal to 1 and the state is not weakly locally observable.

9.1.4 $y = \theta_v - \phi, u = v, w = \omega$

In this case we have $f = \begin{bmatrix} \cos(\theta_v - \phi) \\ \frac{\sin(\theta_v - \phi)}{r} \\ 0 \end{bmatrix}$ and $g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

We follow the five steps mentioned at the end of section 8. We have $L_g^1 = 1$ and $\rho = 0$. Hence, $d_x \rho = [0, 0, 0]$ and we do not need to check if $d_x \rho \in \Omega_m$. In other words, we can set $m' = 0$. By a direct computation we obtain: $\Omega_0 = \text{span}\{[0, -1, 1]\}$, $\Omega_1 = \text{span}\{[0, -1, 1], [-\frac{\sin(\theta_v - \phi)}{r^2}, -\frac{\cos(\theta_v - \phi)}{r}, \frac{\cos(\theta_v - \phi)}{r}]\}$. Additionally, we obtain $\Omega_2 = \Omega_1$, meaning that $m^* = 1$ and $\Omega^* = \Omega_1$, whose dimension is 2. We conclude that the dimension of the observable space is equal to 2 and the state is not weakly locally observable.

9.1.5 $y = \phi, u = \omega, w = v$

In this case we have $f = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ and $g = \begin{bmatrix} \cos(\theta_v - \phi) \\ \frac{\sin(\theta_v - \phi)}{r} \\ 0 \end{bmatrix}$.

We follow the five steps mentioned at the end of section 8. We have $L_g^1 = \frac{\sin(\theta_v - \phi)}{r}$ and $\rho = -2 \cot(\theta_v - \phi)$. Additionally:

$$d_x \rho = \frac{2}{\sin^2(\theta_v - \phi)} [0, -1, 1]$$

We also have $\Omega_0 = \text{span}\{[0, 1, 0]\}$. Hence, $d_x \rho \notin \Omega_0$. Additionally, $\Omega_1 = \Omega_0$. We need to

compute Ω_2 and, in order to do this, we need to compute ϕ_1 . We obtain: $\phi_1 = \begin{bmatrix} -r \\ \cot(\theta_v - \phi) \\ 0 \end{bmatrix}$

and $\Omega_2 = \text{span}\left\{[0, 1, 0], \frac{1}{\sin^2(\theta_v - \phi)} [0, 1, -1]\right\}$. It is immediate to check that $d_x \rho \in \Omega_2$, meaning that $m' = 2$. Additionally, by a direct computation, it is possible to check that $\Omega_3 = \Omega_2$ meaning that $m^* = 2$ and $\Omega^* = \Omega_2$, whose dimension is 2. We conclude that the dimension of the observable space is equal to 2 and the state is not weakly locally observable.

9.1.6 $y = \phi, u = v, w = \omega$

In this case we have $f = \begin{bmatrix} \cos(\theta_v - \phi) \\ \frac{\sin(\theta_v - \phi)}{r} \\ 0 \end{bmatrix}$ and $g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

We follow the five steps mentioned at the end of section 8. We easily obtain $L_g^1 = 0$. Hence, we have to introduce new local coordinates, as explained at the end of section 5. We obtain $\mathcal{L}_f^1 h = \frac{\sin(\theta_v - \phi)}{r}$ and we obtain that the relative degree of the associated system in (15) is $r = 2$. Let us denote the new coordinates by x'_1, x'_2, x'_3 . In accordance with (16) and (17) we set $x'_1 = \phi$ and $x'_2 = \frac{\sin(\theta_v - \phi)}{r}$. Finally, we set $x'_3 = \frac{\cos(\theta_v - \phi)}{r}$.

We compute the new vector fields that characterize the dynamics in the new coordinates. We obtain:

$$\tilde{f} \equiv \begin{bmatrix} x'_2 \\ -2x'_2 x'_3 \\ x'^2_2 - x'^2_3 \end{bmatrix} \quad \tilde{g} \equiv \begin{bmatrix} 0 \\ x'_3 \\ -x'_2 \end{bmatrix} \quad (34)$$

Additionally, we set $\tilde{h} = x'_2$ and $\Omega_1 = \text{span}\{[1, 0, 0], [0, 1, 0]\}$. In the new coordinates we obtain: $L_g^1 = x'_3$ and $\rho = -\frac{x'_2}{x'_3}$. Since ρ depends on x'_3 , $d_x \rho \notin \Omega_1$. Since the dimension of Ω_1 is already 2, because of lemma 9, we know that it exists a given integer m such that the dimension of Ω_m is larger than 2. Hence, we conclude that the entire state is weakly locally observable.

9.2 Systems with multiple disturbances

In this case we refer to the general case, i.e., to systems characterized by the dynamics given in (1). For this general case we do not have the results stated by the theorem of separation (theorem 1) and we have to compute the entire codistribution and to proceed as it has been described in section 4.

We derive the observability properties of two systems with unknown inputs. The first system characterizes a localization problem in the framework of mobile robotics. The state and its

dynamics are the same as in the example discussed in 9.1. However, we consider a different output and also the case when both the inputs are unknown. For this simple example, the use of our theory is not required to derive the observability properties, which can be obtained by using intuitive reasoning.

The second system is much more complex and describes one of the most important sensor fusion problem, which is the problem of fusing visual and inertial measurements. We will refer to this problem as to the visual-inertial structure from motion problem (the Vi-SfM problem). This problem has been investigated by many disciplines, both in the framework of computer science [7, 23, 24, 30, 32, 38] and in the framework of neuroscience (e.g., [5, 10, 11]). Inertial sensors usually consist of three orthogonal accelerometers and three orthogonal gyroscopes. All together, they constitute the Inertial Measurement Unit (IMU). We will refer to the fusion of monocular vision with the measurements from an IMU as to the *standard* Vi-SfM problem. In [20, 23, 24, 27, 28, 30, 37] and [40] the observability properties of the standard Vi-SfM have been investigated in several different scenarios. Very recently, following two independent procedures, the most general result for the standard Vi-SfM problem has been provided in [16] and [31]. This result can be summarized as follows. In the standard Vi-SfM problem all the independent observable states are: the positions in the local frame of all the observed features, the three components of the speed in the local frame, the biases affecting the inertial measurements, the roll and the pitch angles, the magnitude of the gravity and the transformation between the camera and IMU frames. The fact that the yaw angle is not observable is an obvious consequence of the system invariance under rotation about the gravity vector. We want to use here the theory developed in the previous sections in order to investigate the observability properties of the Vi-SfM problem when the number of inertial sensors is reduced, i.e., when the system is driven by unknown inputs.

9.2.1 Simple 2D localization problem

We consider the system characterized by the same dynamics given in (31). Additionally, we assume that the vehicle is equipped with a GPS able to provide its position. Hence, the system output is the following two-components vector:

$$y = [x_v, y_v]^T \quad (35)$$

Let us start by considering the case when both the system inputs, i.e., the two functions $v(t)$ and $\omega(t)$, are available. By comparing (1) with (31) we obtain $x = [x_v, y_v, \theta_v]^T$, $m_u = 2$, $m_w = 0$, $u_1 = v$, $u_2 = \omega$, $f_0(x) = [0, 0, 0]^T$, $f_1(x) = [\cos \theta_v, \sin \theta_v, 0]^T$ and $f_2(x) = [0, 0, 1]^T$.

In order to investigate the observability properties, we apply the observability rank condition introduced in [19].

The system has two outputs: $h_x \equiv x_v$ and $h_y \equiv y_v$. By definition, they coincide with their zero-order Lie derivatives. Their gradients with respect to the state are, respectively: $[1, 0, 0]$ and $[0, 1, 0]$. Hence, the space spanned by the zero-order Lie derivatives has dimension two. Let us compute the first order Lie derivatives. We obtain: $\mathcal{L}_1^1 h_x = \cos \theta_v$, $\mathcal{L}_1^1 h_y = \sin \theta_v$, $\mathcal{L}_2^1 h_x = \mathcal{L}_2^1 h_y = 0$. Hence, the space spanned by the Lie derivatives up to the first order span the entire configuration space and we conclude that the state is weakly locally observable.

We now consider the case when both the system inputs are unknown. In this case, by comparing (1) with (31) we obtain $m_u = 0$, $m_w = 2$, $w_1 = v$, $w_2 = \omega$, $f_0(x) = [0, 0, 0]^T$, $g_1(x) = [\cos \theta_v, \sin \theta_v, 0]^T$ and $g_2(x) = [0, 0, 1]^T$.

Intuitively, we know that the knowledge of both the inputs is unnecessary in order to have the full observability of the entire state. Indeed, the first two state components can be directly obtained from the GPS. By knowing these two components during a given time interval, we also

know their time derivatives. In particular, we know $\dot{x}_v(0)$ and $\dot{y}_v(0)$. From (31) we easily obtain: $\theta_v(0) = \text{atan}\left(\frac{\dot{y}_v(0)}{\dot{x}_v(0)}\right)$. Hence, also the initial orientation is observable, by only using the GPS measurements.

Let us proceed by applying the *EORC*, discussed in section 4. We start by computing the codistribution $\bar{\Omega}_0$ in $\Sigma^{(0)}$. We easily obtain:

$$\bar{\Omega}_0 = \text{span}\{[1, 0, 0], [0, 1, 0]\}$$

From this we know that x_v and y_v are weakly locally observable. We want to know if also θ_v is weakly locally observable (in which case the entire state would be weakly locally observable). We have to compute $\bar{\Omega}_1$ in $\Sigma^{(1)}$. For, we build the system $\Sigma^{(1)}$. We have: ${}^1x = [x_v, y_v, \theta_v, v, \omega]^T$. We can easily obtain the analytical expression for the quantities appearing in (3). We have: $f_0^{(1)}(x) = [\cos \theta_v v, \sin \theta_v v, \omega, 0, 0]^T$. We compute the analytical expression of the first-order Lie derivatives along this vector field. We have: $\mathcal{L}_0^1 h_x = \nabla h_x \cdot f_0^{(1)}(x) = [1, 0, 0, 0, 0] \cdot [\cos \theta_v v, \sin \theta_v v, \omega, 0, 0] = \cos \theta_v v$ (similarly, we obtain $\mathcal{L}_0^1 h_y = \sin \theta_v v$). We obtain:

$$\begin{aligned} \bar{\Omega}_1 = \text{span}\{[1, 0, 0, 0, 0], [0, 1, 0, 0, 0], \\ [0, 0, -\sin \theta_v v, \cos \theta_v v, 0], [0, 0, \cos \theta_v v, \sin \theta_v v, 0]\} \end{aligned}$$

from which we obtain that the gradient of θ_v belongs to $\bar{\Omega}_1$. Therefore, also θ_v is weakly locally observable and so the entire original state.

9.2.2 The Vi-SfM with partial input knowledge

For the brevity sake, we do not provide here the computation necessary to deal with this problem. All the details are available in [33, 34] (see also the work in [29] for the definition of continuous symmetries). Here we provide a summary of these results. First of all, we remark that the Vi-SfM problem can be described by a nonlinear system with six inputs (3 are the accelerations along the three axes, provided by the accelerometers, and 3 are the angular speeds provided by the gyroscopes). The outputs are the ones provided by the vision. In the simplest case of a single point feature, they consist of the two bearing angles of this point feature in the camera frame.

We analyzed the following three cases:

1. camera extrinsically calibrated, only one input known (corresponding to the acceleration along a single axis);
2. camera extrinsically uncalibrated, only one input known (corresponding to the acceleration along a single axis);
3. camera extrinsically uncalibrated, two inputs known (corresponding to the acceleration along two orthogonal axes).

The dimension of the original state is 12 in the first case and 23 in the other two cases. Additionally $m_u = 1$ and $m_w = 5$ in the first two cases while $m_u = 2$ and $m_w = 4$ in the last case.

In [33, 34] we prove that the observability properties of Vi-SfM do not change by removing all the three gyroscopes and one of the accelerometers. In other words, exactly the same properties hold when the sensor system only consists of a monocular camera and two accelerometers. To achieve this result, we computed the Lie derivatives up to the second order for the third case mentioned above. By removing a further accelerometer (i.e., by considering the case of a monocular camera and a single accelerometer) the system loses part of its observability properties. In

particular, the distribution $\Delta^k (\equiv \bar{\Omega}_k^\perp)$, $\forall k \geq 2$, contains a single vector. This vector describes a continuous symmetry that is the invariance under the rotation around the accelerometer axis. This means that some of the internal parameters that define the extrinsic camera calibration, are no longer identifiable. Although this symmetry does not affect the observability of the absolute scale and the magnitude of the velocity, it reflects in an indistinguishability of all the initial speeds that differ for a rotation around the accelerometer axis. On the other hand, if the camera is extrinsically calibrated (i.e., if the relative transformation between the camera frame and the accelerometer frame is known (first case mentioned above)) this invariance disappears and the system still maintains full observability, as in the case of three orthogonal accelerometers and gyroscopes. The analysis of this system (the first case mentioned above) has been done in the extreme case when only a single point feature is available. This required to significantly augment the original state. In particular, in [33, 34] we compute all the Lie derivatives up to the 7th order, i.e., we included in the original state the 5 unknown inputs together with their time-derivatives up to the six order. We prove that the gradient of any component of the original state, with the exception of the yaw angle, is orthogonal to the distribution Δ^k , $\forall k \geq 7$ (see the computational details in [33, 34])³.

³Note that, the yaw angle is not observable even in the case when all the 6 inputs are known. The fact that the yaw is unobservable is a consequence of a symmetry in the considered system, which is the system invariance under rotations about the gravity axis.

10 Conclusion

In this paper we investigated the problem of nonlinear observability when part (or even all) of the system inputs is unknown. We made the assumption that the unknown inputs are differentiable functions of time (up to a given order). The goal was not to design new observers but to provide simple analytic conditions in order to check the weak local observability of the state. An unknown input was also called disturbance. The analysis started by extending the observability rank condition. This was obtained by a state augmentation and was called the extended observability rank condition. In general, by further augmenting the state, the observability properties of the original state also increase. As a result, the extended observability rank condition only provides a sufficient condition for the weak local observability of the original state since the state augmentation can be continued indefinitely. Additionally, the computational cost demanded to obtain the observability properties through the extended observability rank condition, dramatically depends on the dimension of the augmented state. For these reasons, we focused our investigation on the following two fundamental issues. The former consisted in understanding if there exists a given augmented state such that, by further augmenting the state, the observability properties of the original state provided by the extended observability rank condition remain unvaried. The latter consisted in understanding if it is possible to derive the observability properties of the original state by computing a codistribution defined in the original space, namely a codistribution consisting of covectors with the same dimension of the original state. Both these issues have been fully addressed in the case of a single unknown input. In this case, we provided an analytical method to operate a separation on the codistribution computed by the extended observability rank condition, i.e., the codistribution defined in the augmented space. Thanks to this separation, we introduced a method able to obtain the observability properties by simply computing a codistribution that is defined in the original space (theorem 1). The new codistribution is defined recursively by a very simple algorithm. Specifically, the algorithm in definition 3 in section 6 (for the case of a single known input) and in definition 5 in section 8 (for the case of multiple known inputs). Hence, the overall method to obtain all the observability properties is very simple. On the other hand, the analytic derivations required to prove the validity of this separation are complex and we are currently extending them to the multiple unknown inputs case. Finally, we showed that the recursive algorithm converges in a finite number of steps and we also provided the criterion to establish if the algorithm has converged (theorem 2). Also this proof is based on several tricky and complex analytical steps.

Both theorems 1 and 2 have first been proved in the case of a single known input (sections 6 and 7) but in section 8 their validity was extended to the case of multiple known inputs.

All the theoretical results have been illustrated by deriving the observability properties of several nonlinear systems driven by known and unknown inputs.

References

- [1] J.-P. Barbot, M. Fliess, T. Floquet, "An algebraic framework for the design of nonlinear observers with unknown inputs", IEEE Conf. on Decision and Control, New-Orleans, USA, 2007
- [2] J.-P. Barbot, D. Boutat, T. Floquet, "An observation algorithm for nonlinear systems with unknown inputs", Automatica, Volume 45, Issue 8, August 2009, Pages 1970-1974
- [3] G. Basile and G. Marro. On the observability of linear, time invariant systems with unknown inputs. J. Optimization Theory Appl., 3:410–415, 1969.
- [4] F. A. W. Belo, P. Salaris, and A. Bicchi, 3 Known Landmarks are Enough for Solving Planar Bearing SLAM and Fully Reconstruct Unknown Inputs, The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems October 18-22, 2010, Taipei, Taiwan
- [5] A. Berthoz, B. Pavard and L.R. Young, Perception of Linear Horizontal Self-Motion Induced by Peripheral Vision Basic Characteristics and Visual-Vestibular Interactions, Exp. Brain Res. 23, 471–489 (1975).
- [6] S.P. Bhattacharyya Observer design for linear systems with unknown inputs IEEE Transactions on Automatic Control, 23 (1978)
- [7] M. Bryson and S. Sukkarieh, Observability Analysis and Active Control for Airborne SLAM, IEEE Transaction on Aerospace and Electronic Systems, vol. 44, no. 1, 261–280, 2008
- [8] Chen, W. and Saif, M. (2006). Unknown input observer design for a class of nonlinear systems: an LMI approach. In IEEE American Control Conference.
- [9] Darouach, M., Zasadzinski, M., Xu, S. J. (1994). Full-order observers for linear systems with unknown inputs. IEEE Transactions on Automatic Control, 39(3)
- [10] Dokka K., MacNeilage P. R., De Angelis G. C. and Angelaki D. E., Estimating distance during self-motion: a role for visual-vestibular interactions, Journal of Vision (2011) 11(13):2, 1-16
- [11] C. R. Fetsch, G. C. DeAngelis and D. E. Angelaki, Visual-vestibular cue integration for heading perception: Applications of optimal cue integration theory, Eur J Neurosci. 2010 May ; 31(10): 1721-1729
- [12] Floquet, T., Barbot, J. (2004). A sliding mode approach of unknown input observers for linear systems. In 43rd IEEE conference on decision and control. Atlantis, Paradise Island, Bahamas. December.
- [13] Floquet, T., Edwards, C., Spurgeon, S. (2007). On sliding mode observers for systems with unknown inputs. Int. J. Adapt. Control Signal Process Vol 21, no. 89, 638–656, 2007.
- [14] Guan, Y., Saif, M. (1991). A novel approach to the design of unknown input observers. IEEE Transactions on Automatic Control, 36(5).
- [15] 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.

- [16] 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
- [17] Q.P. Ha, H. Trinh State and input simultaneous estimation for a class of nonlinear systems *Automatica*, 40 (2004).
- [18] H. Hammouri and Z. Tmar. Unknown input observer for state affine systems: A necessary and sufficient condition. *Automatica*, 46(2):271–278, 2010.
- [19] Hermann R. and Krener A.J., 1977, Nonlinear Controllability and Observability, *Transaction On Automatic Control*, AC-22(5): 728–740
- [20] 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.
- [21] M. Hou, P.C. MÅller Design of observers for linear systems with unknown inputs *IEEE Transactions on Automatic Control*, 37 (6) (1992).
- [22] Isidori A., *Nonlinear Control Systems*, 3rd ed., Springer Verlag, 1995.
- [23] 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.
- [24] 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.
- [25] Koenig, D. and Mammarr, S. (2001). Design of a class of reduced order unknown inputs nonlinear observer for fault diagnosis. In *IEEE American Control Conference*.
- [26] D. Koenig, B. Marx, D. Jacquet Unknown input observers for switched nonlinear discrete time descriptor system *IEEE Transactions on Automatic Control*, 53 (1) (2008).
- [27] 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.
- [28] M. Li and A. I. Mourikis, Improving the accuracy of EKF-based visualinertial odometry, *ICRA 2012*, St. Paul, MN.
- [29] 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
- [30] 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.
- [31] A. Martinelli, Visual-inertial structure from motion: observability and resolvability, *iros 2013*, Tokyo, Japan
- [32] A. Martinelli, Closed-form solution of visual-inertial structure from motion, *Int. J. of Computer Vision*, Vol. 106, N. 2, 138–152, 2014.

- [33] A. Martinelli, Visual-inertial structure from motion: observability vs minimum number of sensors, ICRA 2014, Hong Kong.
- [34] A. Martinelli, Observability Properties and Deterministic Algorithms in Visual-Inertial Structure from Motion, Foundations and Trends in Robotics, 139–209, <http://dx.doi.org/10.1561/23000000030>, 2014
- [35] A. Martinelli. Extension of the Observability Rank Condition to Nonlinear Systems Driven by Unknown Inputs. 23th Mediterranean Conference on Control and Automation (MED), June, 2015. Torremolinos, Spain
- [36] A. Martinelli. Nonlinear Unknown Input Observability: Analytical Expression of the Observable Codistribution in the Case of a Single Unknown Input. The SIAM Conference on Control and Its Applications (CT15). July, 2015. Paris, France
- [37] 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
- [38] D. Strelow and S. Singh, Motion estimation from image and inertial measurements, International Journal of Robotics Research, 23(12), 2004
- [39] S.H. Wang, E.J. Davison, P. Dorato Observing the states of systems with unmeasurable disturbance IEEE Transactions on Automatic Control, 20 (1975)
- [40] Weiss., S., Vision Based Navigation for Micro Helicopters, PhD thesis, Diss. ETH No. 20305
- [41] F. Yang, R.W. Wilde Observer for linear systems with unknown inputs IEEE Transactions on Automatic Control, 33 (7) (1988).

Contents

1	Introduction	3
2	Basic Definitions	4
3	Extended system and basic properties	5
4	Extension of the Observability Rank condition	7
5	Single known Input and single disturbance	9
6	The observable codistribution (Ω)	11
7	Convergence of the algorithm that defines Ω	14
8	Extension to the case of multiple known inputs and method's summary	18
9	Applications	19
9.1	Systems with a single disturbance	19
9.1.1	$y = r, u = \omega, w = v$	20
9.1.2	$y = r, u = v, w = \omega$	21
9.1.3	$y = \theta_v - \phi, u = \omega, w = v$	21
9.1.4	$y = \theta_v - \phi, u = v, w = \omega$	21
9.1.5	$y = \phi, u = \omega, w = v$	22
9.1.6	$y = \phi, u = v, w = \omega$	22
9.2	Systems with multiple disturbances	22
9.2.1	Simple 2D localization problem	23
9.2.2	The Vi-SfM with partial input knowledge	24
10	Conclusion	26



**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