Non linear Unknown Input Observability: Basic Properties and the Case of a Single Unknown Input

Agostino Martinelli

Non linear Unknown Input Observability: Basic Properties and the Case of a Single Unknown Input

Agostino Martinelli

ISRN INRIA/RR--8608--FR+ENG
ISSN 0249-6399

RESEARCH REPORT
N° 8608
October 2014
Project-Team Emotion

Non linear Unknown Input Observability: Basic Properties and the Case of a Single Unknown Input

Agostino Martinelli

Project-Team Emotion

Research Report n° 8608 — October 2014 — 23 pages

Abstract: This paper investigates the problem of non linear observability when part (or even all) of the system inputs is unknown. A new definition of indistinguishable states is first provided. Then, in order to separate the effect of the known inputs from the effect of the unknown inputs on the system outputs, the state is augmented. This allows us to obtain the extension of basic properties, which hold in the case of known inputs. Starting from these properties, the paper analyzes the state observability and provides sufficient conditions that allow us to establish if a given system is weakly locally observable. On the other hand, the proposed method computes a codistribution defined in the augmented space. This makes the computational cost dependent on the dimension of the augmented state. In the case of a single unknown input, the paper proposes a method to operate a separation on the previous codistribution. Thanks to this separation, it is possible to derive the observability properties by computing a codistribution that is defined in the original space. As it will be seen, the analytic derivations required to perform this separation are complex and we are currently extending them to the multiple unknown inputs case. The proposed approach is used to derive the observability properties of many systems, starting from very simple ones. The last application is a very complex unknown input observability problem. Specifically, we derive the observability properties for the visual-inertial structure from problem in the case when part of the inertial inputs are missing.

Key-words: Non linear 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: propriétés de base et le cas d'une seule entrée inconnue

Résumé : Cet article étudie le problème de l'observabilité non linéaire quand une partie (voire la totalité) des entrées du système est inconnu. Une nouvelle définition des états indiscernables est d'abord donné. Ensuite, afin de séparer l'effet des entrées connues de l'effet des entrées inconnues sur les sorties du système, l'état est augmenté. Cela nous permet d'obtenir l'extension de propriétés de base, qui détiennent dans le cas des entrées toutes connues. A partir de ces propriétés, le document analyse l'observabilité de l'état et fournit des conditions suffisantes qui nous permettent d'établir si un système est observable. D'autre part, la méthode proposée calcule une codistribution définie dans l'espace augmenté. Cela rend le coût du calcul dépendent de la dimension de l'état. Dans le cas d'une seule entrée inconnue, le document propose une méthode pour opérer une séparation sur la codistribution précédent. Grace a cette séparation, il est possible de calculer les propriétés d'observabilité par une codistribution qui est définie dans l'espace original. Comme on le verra, les dérivations analytiques nécessaires pour effectuer cette séparation sont complexes et nous sommes en train de les étendre au cas de plusieurs entrés inconnues. L'approche proposé est utilisé pour calculer les propriétés d'observabilité de nombreux systèmes, en partant de très simples. La dernière application est un problème très complexe. Plus précisément, nous dérivons les propriétés d'observabilité pour la le problème de la fusion de données visuels et inertielles dans le cas où une partie des entrées est manquante.

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 [1, 4, 8, 25]. A huge effort has then been devoted to design observers for both linear and nonlinear systems in presence of UI. On the other hand, to the best of our knowledge, a theory of non linear observability that extends the results obtained in [10] to the case UI still lacks. In [10] the observability properties of a non linear 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 basically means that, the set where all the Lie derivatives take the same values taken at x_0, includes the set I_{x_0}.

This paper provides a first contribution towards a general theory of unknown inputs observability (UIO) in the non linear case, i.e., a general theory that extends the results derived in the case when part (or even all) of the inputs are unknown. Our results hold for systems whose dynamics are non linear in the state and affine in both the known and unknown inputs. We start by reminding the reader few basic definitions for the case with only known inputs (KI). Then, in section 2.2 we introduce the new definition of *indistinguishable states* for the case UI. Section 3 provides a discussion which motivates the key idea, that consists in augmenting the original state in order to make null the Lie derivatives of the output (up to a given order) along all the directions that correspond to the unknown inputs. Thanks to this fact, we obtain the extension of the property of the theory in [10] mentioned above, i.e., that the Lie derivatives of any output (up to a given order) computed along any direction allowed by the system dynamics take the same values at the states which are indistinguishable (see property 4).

Section 4 and 5 introduce the extended state together with its dynamics and some basic properties. The extended 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 [2] 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 [2] a sufficient condition for the state observability has been provided.

Based on this augmented state the paper proposes a method to analyze the state observability (section 6). This method is based on the computation of a codistribution defined in the augmented space. This makes the computational cost dependent on the dimension of the augmented state.

In sections 7 and 8 we analyze the case of a single unknown input. In this case, the paper proposes a method to operate a separation on the previous codistribution. Specifically, it provides an automatic method able to derive a basis for this codistribution. This basis consists of two set of covectors. The first set consists of gradients of scalar functions that only depend on the original state. The second set consists of gradients of scalar functions that depend on the entire augmented state. However, the functions of this second set do not contain additional information on the original state. As it will be seen, the analytic derivations required to perform this separation are complex and we are currently extending them to the multiple unknown inputs case.

The proposed approach is illustrated by deriving the observability properties of many systems, starting from very simple ones. The last application is a very complex unknown input observability problem. Specifically, we derive the observability properties for the visual-inertial structure from problem in the case when part of the inertial inputs are missing.

2 Basic Definitions

In the following we will refer to a non-linear control system with m_u known inputs ($u \equiv [u_1, \cdots, u_{m_u}]^T$) and m_w unknown inputs or disturbances ($w \equiv [w_1, \cdots, w_{m_w}]^T$). The state is the vector $x \in U$, with U an open set of $\mathbb{R}^n$. We assume that the dynamics are non-linear 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:

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

where the functions $f_i(x)$, $i = 0, 1, \cdots, m_u$, and $g_j(x)$, $j = 1, \cdots, m_w$, are $\mathbb{R}^n$-valued functions defined on the open set U and the function $h(x)$ is a scalar function defined on the open set U. For the sake of simplicity, we will assume that all these functions are analytic functions in U.

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 following, 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, $\forall t \in \mathcal{I}$.

2.1 The case of known inputs

Let us start by reminding the reader some basic concepts and definitions for the case when all the inputs are known in the time interval $\mathcal{I}$, i.e., when $m_w = 0$. The definition of indistinguishable states is:

Definition 1 (Indistinguishable states) *Two states x_a and x_b are indistinguishable if for any $u(t)$ (the vector containing the inputs), $h(x(t; \ x_a; \ u)) = h(x(t; \ x_b; \ u))$ $\forall t \in \mathcal{I}$.*

Additionally, 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. According to the theory of observability, a system is observable in x_0 if $I_{x_0} = x_0$. In the sequel, we will refer to the concept of *weak local observability* as defined in [10]. Starting from this definition, in [10] the observability rank criterion has been introduced and proved that it is a sufficient condition for a system to be locally weakly observable (theorem 3.1 in [10]). Additionally, the derivations in [10] proves that the viceversa of this theorem holds generically (theorem 3.11). We address the reader to [10] for further details. Our goal is to extend the observability rank condition to the case of UI.

2.2 The case of known and unknown inputs

Let us now consider the general case, i.e., when $m_w \neq 0$. The only definition that differs with respect to the previous case ($m_w = 0$) is the one of indistinguishable states. We introduce the following definition:

Definition 2 (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 definition 1, 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 The key idea

This section provides an intuitive procedure in order to motivate our approach, which will be introduced in the next sections. We believe useful to show our basic idea through this intuitive procedure before introducing the approach and deriving for it all the analytical properties starting from the definitions given in the previous sections.

Roughly speaking, our objective is to get information on the initial state x_0 starting from the knowledge of the input $u(t)$ and the output $y(t)$, $\forall t \in \mathcal{I}$, and by knowing the analytical expression of the functions in (1).

Let us start this discussion by considering the case $m_w = 0$. We remark that, because of the Taylor theorem, the knowledge of the input $u(t)$ and the output $y(t)$ is equivalent to the knowledge of their time derivatives at $t = 0$, up to any order. Additionally, it is possible to analytically derive the expression of the $m-$order time derivative of the output function at $t = 0$ in terms of all the Lie derivatives of the function h along all the vector fields $f_i(x)$, $i = 0, 1, \cdots, m_u$ computed at x_0 up to the m^{th} order and all the time derivatives of the functions u_i $(i = 1, \cdots, m_u)$ computed at $t = 0$. We have:

$$\left.\frac{d^m h(x(t;\ x_0;\ u))}{dt^m}\right|_{t=0} = \sum_{p=1}^{m} \sum_{i_1 i_2 \cdots i_p = 0}^{m_u} \mathcal{L}^p_{i_1 i_2 \cdots i_p} h(x_0) \tag{2}$$

$$\sum_{k_1, k_2, \cdots, k_p = 0,\ |\sum_{j=1}^p k_j = m-p}^{m-p} C^{m,\ p}_{k_1, k_2, \cdots, k_p} u^{(k_1)}_{i_1} \cdots u^{(k_p)}_{i_p}$$

where:

- $u_i^{(k)} \equiv \left.\frac{d^k u_i}{dt^k}\right|_{t=0}$, $k = 0, 1, \cdots, m$; $i = 1, \cdots, m_u$;
- $u_0 \equiv 1$ and $u_0^{(k)} = 0$, $k > 1$;
- $C^{m,\ p}_{k_1, k_2, \cdots, k_p}$, are real numbers satisfying a recursive equation which can be obtained by directly differentiating the expression in (2) with respect to time.
- $\mathcal{L}^p_{i_1 i_2 \cdots i_p} h(x_0)$ is the $p-$order Lie derivative of the function h along the vector fields $f_{i_1}, \cdots, f_{i_p}$, computed at x_0.

Equation (2) provides the analytical expression for the link between the input and the output of our system. This link depends on the initial state x_0 through the Lie derivatives of the function h along all the vector fields f_i, $i = 0, 1, \cdots, m_u$. By exciting the system with different inputs, we can get information on x_0 by analyzing the output. Since the dependence on the initial state is fully encompassed in the Lie derivatives, equation (2) allows us to conclude that, an upper bound on the knowledge of the initial state, is provided by the knowledge of the values that all the Lie derivatives take at the initial state. On the other hand, we cannot a priori exclude that the system contains enough information to determine all these values.

Let us now consider the general case, i.e., when $m_w \neq 0$. The expression in (2) must be changed to include the disturbances with their time derivatives and the Lie derivatives along the vector fields g_j, $j = 1, \cdots, m_w$. In other words, the time derivatives of the output depend on the time derivatives of the disturbances w_j, $j = 1, \cdots, m_w$. Hence, by exciting our system with different known inputs, we can not even get information on the vector constituted by all the Lie derivatives computed at the initial state. Indeed, the output is simultaneously excited by the known inputs and the unknown disturbances. Let us set $\tilde{u}_i \equiv u_i$ if $0 \leq i \leq m_u$ and $\tilde{u}_i \equiv w_{i-m_u}$ if $m_u + 1 \leq i \leq m_u + m_w$. We write the new analytical expression that provides the link between the input (both known and unknown) and the output as the sum of two parts:

$$\frac{d^m h(x(t;\ x_0;\ u;\ w)}{dt^m}\bigg|_{t=0} = \sum_{p=1}^{m} \Bigg\{ \sum_{i_1 i_2 \cdots i_p = 0}^{m_u} \mathcal{L}^p_{i_1 i_2 \cdots i_p} h(x_0) \tag{3}$$

$$\sum_{k_1, k_2, \cdots, k_p = 0,\ |\sum_{j=1}^p k_j = m-p}^{m-p} C^{m,\ p}_{k_1, k_2, \cdots, k_p} u^{(k_1)}_{i_1} \cdots u^{(k_p)}_{i_p}$$

$$+ \sum_{i_1 i_2 \cdots i_p = remaining} \mathcal{L}^p_{i_1 i_2 \cdots i_p} h(x_0)$$

$$\sum_{k_1, k_2, \cdots, k_p = 0,\ |\sum_{j=1}^p k_j = m-p}^{m-p} C^{m,\ p}_{k_1, k_2, \cdots, k_p} \tilde{u}^{(k_1)}_{i_1} \cdots \tilde{u}^{(k_p)}_{i_p} \Bigg\}$$

The first sum only contains the know inputs (i.e., $i_1, i_2, \cdots, i_p = 0, 1, \cdots m_u$) while, for each addend in the second sum, i.e., the sum where the indexes $i_1, i_2, \cdots, i_p$ take the *remaining* values, at least one input is unknown (in this second sum, the Lie derivatives $\mathcal{L}^p_{i_1 i_2 \cdots i_p} h(x_0)$ are computed along the vector fields $g_1, \cdots, g_{m_w}$ in correspondence of the indexes larger than m_u). In the very special case when all the Lie derivatives $\mathcal{L}^p_{i_1 i_2 \cdots i_p} h(x_0)$ vanish when at least one index $i_1, i_2, \cdots, i_p$ is larger than m_u, the second sum, which also contains unknown inputs, vanishes as well. Hence, we could do the same considerations provided for the case $m_w = 0$. In other words, in this very special case, we could conclude that, an upper bound on the knowledge of the initial state, is provided by the values that all the Lie derivatives along the first $m_u + 1$ vector fields $(f_0, f_1, \cdots, f_{m_u})$ take at the initial state. Additionally, we could optimistically hope that the system contains enough information to determine all these values. Obviously, this is a very special case. Our idea is to extend the original state in order to artificially reproduce such a situation. In particular, we include the unknown inputs in the state together with their time derivatives. By including the time derivatives up to the $(k-1)^{th}$ order, we will obtain $\mathcal{L}^p_{i_1 i_2 \cdots i_p} h(x_0) = 0$ when at least one index $i_1, i_2, \cdots, i_p$ is larger than m_u and for all $p = 0, 1, \cdots, k$ (see property 1 in the next section). This property will be fundamental and in particular it allows us to obtain the result stated by property 4, which is the extension of the well known result in the theory of non-linear observability (with known inputs), stating that if two states are indistinguishable, the

Lie derivatives of the output along any vector fields appearing in the dynamics take the same values on these two states. In the next sections we introduce our approach.

4 Extended system

In accordance with the previous discussion, in order to separate the effect of the disturbances on the output from the effect of the known inputs on the output, we augment the original state by including the unknown inputs together with their time-derivatives. For each unknown input w_j $(j = 1, \cdots, m_w)$, let us denote by $w_j^{(k)}$ its k^{th} time derivative, i.e., $w_j^{(k)} \equiv \frac{d^k w_j}{dt^k}$.

We define a new system, which 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 $x^{(k)}$ the extended state that includes the time derivatives up to the $(k-1)-$order:

$$x^{(k)} \equiv [x^T, \ w^T, \ w^{(1)\ T}, \ \cdots, \ w^{(k-1)\ T}]^T \tag{4}$$

The dimension of the extended state is $n+km_w$. From (1) it is immediate to obtain the dynamics for the extended state:

$$\dot{x}^{(k)} = f_0^{(k)}(x^{(k)}) + \sum_{i=1}^{m_u} f_i^{(k)}(x)u_i + \sum_{j=1}^{m_w} 1_{n+km_w}^{n+(k-1)m_w+j} w_j^{(k)} \tag{5}$$

where:

$$f_0^{(k)}(x^{(k)}) \equiv \left[\begin{array}{c} f_0(x) + \sum_{i=1}^{m_w} g_i(x)w_i \\ w^{(1)} \\ w^{(2)} \\ \cdots \\ w^{(k-1)} \\ 0_{m_w} \end{array} \right] \tag{6}$$

$$f_i^{(k)}(x) \equiv \left[\begin{array}{c} f_i(x) \\ 0_{km_w} \end{array} \right] \tag{7}$$

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 $f_i^{(k)}$, $(i = 1, \cdots, m_u)$ and it depends on the disturbances via the unit vectors $1_{n+km_w}^{n+(k-1)m_w+j}$, $(j = 1, \cdots, m_w)$. Finally, we remark that only the vector field $f_0^{(k)}$ depends on the new state elements.
We derive simple 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 $w_j^{(f)}$, $j = 1, \cdots, m_w$, $\forall f \geq m$.*

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, $\nabla g = [\frac{\partial g}{\partial x}, \frac{\partial g}{\partial w}, \frac{\partial g}{\partial w^{(1)}}, \cdots, \frac{\partial g}{\partial w^{(m-1)}}, 0_{(k-m)m_w}]$. It is immediate to realize that the product of this gradient by any vector filed in (5) 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:

Property 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 $1_{n+km_w}^{n+(k-1)m_w+j}$ $(j = 1, \cdots, m_w)$ 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 (4). Then, the proof follows from the fact that any vector among $1_{n+km_w}^{n+(k-1)m_w+j}$ $(j = 1, \cdots, m_w)$ has the first $n + (k-1)m_w$ components equal to zero ■

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

5 Properties of the extended system

In the following we will use the notation: $\xi \equiv [w^T, \ w^{(1)\ T}, \ \cdots, \ w^{(k-1)\ T}]^T$. In this notation we have $x^{(k)} = [x^T, \ \xi^T]^T$. We also denote by $\Sigma^{(0)}$the original system, i.e., the one characterized by the state x and the equations in (1). The definition 2, given for $\Sigma^{(0)}$, can be applied to $\Sigma^{(k)}$. Specifically, in $\Sigma^{(k)}$, two states $[x_a, \xi_a]$ and $[x_b, \xi_b]$ are indistinguishable if and only if, for any $u(t)$ (the known inputs), there exist two vector functions $w_a^{(k)}(t)$ and $w_b^{(k)}(t)$ (the k^{th} time derivative of two disturbance vectors) such that, $h(x(t; \ [x_a, \xi_a]; \ u; \ w_a^{(k)})) = h(x(t; \ [x_b, \xi_b]; \ u; \ w_b^{(k)}))$ $\forall t \in \mathcal{I}$. It is immediate to prove the following property:

Property 3 *x_a and x_b are indistinguishable in $\Sigma^{(0)}$if and only if there exist ξ_a and ξ_b such that $[x_a, \xi_a]$ and $[x_b, \xi_b]$ are indistinguishable in $\Sigma^{(k)}$.*

It also holds the following fundamental property:

Property 4 *If $[x_a, \xi_a]$ and $[x_b, \xi_b]$ are indistinguishable in $\Sigma^{(k)}$then the Lie derivatives of the output up to the k^{th}-order computed on these points take the same values.*

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

$$\tilde{u}_i(t) = \tag{8}$$

$$\begin{cases} u_i^1 & t \in [0,\ t_1) \\ u_i^2 & t \in [t_1,\ t_1+t_2) \\ \cdots & \\ u_i^g & t \in [t_1+t_2+\cdots+t_{g-1},\ t_1+t_2+\cdots+t_{g-1}+t_g) \end{cases}$$

Since $[x_a, \xi_a]$ and $[x_b, \xi_b]$ are indistinguishable in $\Sigma^{(k)}$, there exist two disturbance functions $w_a^{(k)}(t)$ and $w_b^{(k)}(t)$ such that:

$$h(x(t;\ [x_a, \xi_a];\ \tilde{u};\ w_a^{(k)})) = h(x(t;\ [x_b, \xi_b];\ \tilde{u};\ w_b^{(k)})) \tag{9}$$

$\forall t \in [0,\ t_1+t_2+\cdots+t_{g-1}+t_g) \subset \mathcal{I}$. On the other hand, by taking the two quantities in (9) 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:

$$\frac{\partial^g h(x(t_1+\cdots+t_g;\ [x_a, \xi_a];\ \tilde{u};\ w_a^{(k)}))}{\partial t_1 \partial t_2 \cdots \partial t_g} = \tag{10}$$

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

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

$$\mathcal{L}^g_{\theta_1\theta_2\cdots\theta_g} h(x_a) = \mathcal{L}^g_{\theta_1\theta_2\cdots\theta_g} h(x_b) \tag{11}$$

where $\theta_h = f_0^{(k)} + \sum_{i=1}^{m_u} f_i^{(k)} u_i^h$, $h = 1, \cdots, g$. The equality in (11) 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:

$$\mathcal{L}^g_{v_1 v_2\cdots v_g} h(x_a) = \mathcal{L}^g_{v_1 v_2\cdots v_g} h(x_b) \tag{12}$$

where $v_1 v_2 \cdots v_g$ are vector fields belonging to the set $\{f_0^{(k)}, f_1^{(k)}, \cdots, f_{m_u}^{(k)}\}$ $\blacksquare$
The indistinguishable set in $\Sigma^{(k)}$of a given $[x_0, \xi_0] \in U \times W$ (where U and W are open sets in $\mathbb{R}^n$ and $\mathbb{R}^{km_w}$, respectively), is the set of all the states $[x, \xi]$ such that $[x, \xi]$ and $[x_0, \xi_0]$ are indistinguishable in $\Sigma^{(k)}$. This set will be denoted by $I_{[x_0, \xi_0]}$.

Finally, for a given $[x_0, \xi_0] \in U \times W$ we introduce the set $K_{[x_0, \xi_0]}$, which is the set of points $[x, \xi] \in U \times W$ such that all the Lie derivatives of the output up to the order k take the same values as in $[x_0, \xi_0]$.
An immediate consequence of property 4 is the following property:

Property 5 $\forall [x_0, \xi_0] \in U \times W$, $K_{[x_0, \xi_0]} \supseteq I_{[x_0, \xi_0]}$.

6 Observability of the original state

We discuss the observability of the original state (x) in a given point x_0. In particular, we will refer to the weak local observability and we start by considering the case without disturbances, i.e., $m_w = 0$. According to the observability rank condition, the weak local observability of the system in (1) with $m_w = 0$ in a given point x_0 can be investigated by analyzing the codistribution generated by the gradients of its Lie derivatives. Specifically, if the dimension of this codistribution is equal to the dimension of the state in a given neighbourhood of x_0, we conclude that the state is weakly locally observable in x_0. As we mentioned in section 2.1, in [10] the observability rank criterion has been introduced and proved that it is a sufficient condition for a system to be locally weakly observable (theorem 3.1). The proof of this theorem is based on the fact that the Lie derivatives of the outputs are constant on the indistinguishable sets.

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 satisfies the same property (see property 4). Therefore, we can adopt the observability rank criterion to investigate the weak local observability, provided that we suitably augment the state. As we mentioned at the beginning of this section, we are interested in deriving the observability properties of the original state x and not in the observability of the entire extended state. We can check the weak local observability of the original state component by component, i.e., by proceeding as follows. Let us consider the i^{th} component of the original state, i.e., x_i ($i = 1, \cdots, n$) and let us denote by $\bar{\Omega}_k$ the codistribution that is the span of the gradients of the Lie derivatives of the system $\Sigma^{(k)}$up to the $k-$order. We check if the gradient of x_i with respect to the extended state belongs to $\bar{\Omega}_k$. In other words, we check if the row vector of dimension $n + km_w$, with all the entries equal to zero with the exception of the i^{th} entry equal to 1, is in the span of the gradients of the Lie derivatives up to the $k-$order. If there exists a k such that this is true, we conclude that x_i is weakly locally observable. If this holds for $i = 1, \cdots, n$, we conclude that the original state is weakly locally observable.

On the other hand, the computation demanded to check if a given covector belongs to $\bar{\Omega}_k$ can be very complex because by increasing k we also increase the dimension of the extended state. In the sequel, we will focus our attention on this fundamental issue. Specifically, we want to solve the following problem: separate the information on the original state from the information on its extension.

We fully address this issue in the case $m_w = 1$. In section 7 we introduce the basic equations that characterize this system. In section 8 we operate a separation on the codistribution generated by all the Lie derivatives up to the $k-$order. Specifically, we provide an automatic method able to derive a basis for this codistribution. This basis consists of two set of covectors. The first set consists of gradients of scalar functions that only depend on the original state. The second set consists of gradients of scalar functions that depend on the entire augmented state. However, the functions of this second set do not contain additional information on the original state. We are currently extending these results to the general case (i.e., for any m_w).

7 Single Input and single disturbance

We will refer to the following system:

$$\begin{cases} \dot{x} = f(x)u + g(x)w \\ y = h(x) \end{cases} \tag{13}$$

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

$$x^{(k)} \equiv [x^T, \ w^T, \ w^{(1) \ T}, \ \cdots, \ w^{(k-1) \ T}]^T \tag{14}$$

The dimension of the extended state is $n + k$. From (13) it is immediate to obtain the dynamics for the extended state:

$$\dot{x}^{(k)} = G(x^{(k)}) + F(x)u + 1_{n+k}^{n+k} w^{(k)} \tag{15}$$

where:

$$F \equiv \begin{bmatrix} f(x) \\ 0 \\ 0 \\ \cdots \\ 0 \\ 0 \end{bmatrix} \qquad G \equiv \begin{bmatrix} g(x)w \\ w^{(1)} \\ w^{(2)} \\ \cdots \\ w^{(k-1)} \\ 0 \end{bmatrix} \tag{16}$$

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., $L_g^1 \equiv \mathcal{L}_g h$. We denote with the symbol d the gradient with respect to the extended state in (14). When we want to denote the gradient only respect to x, we adopt the symbol d_x. Additionally, for a given codistribution Ω, we will denote by $\mathcal{L}_F\Omega$ the codistribution whose covectors are the Lie derivatives along F of the covectors in Ω. Finally, given two vector spaces V_1 and V_2, we denote with $V_1 + V_2$ their sum, i.e., the span of all the generators of both V_1 and V_2.

The derivations provided in the next section are based on the assumption that $L_g^1 \neq 0$ in 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 $L_g^1 = 0$. We introduce the following system associated with the system in (15):

$$\begin{cases} \dot{x} = f(x) + g(x)u \\ y = h(x) \end{cases} \tag{17}$$

This is a system without disturbances and with a single known input u. Let us denote by r the relative degree of this system in x_0. Since $L_g^1 = 0$, we have $r > 1$. Additionally, we can introduce new local coordinates $x' = \mathcal{Q}(x) = \begin{bmatrix} \mathcal{Q}_1(x) \\ \cdots \\ \mathcal{Q}_n(x) \end{bmatrix}$ such that the first new r coordinates are: $\mathcal{Q}_1(x) = h(x)$, $\mathcal{Q}_2(x) = \mathcal{L}_f^1 h(x)$, $\cdots$, $\mathcal{Q}_r(x) = \mathcal{L}_f^{r-1} h(x)$ (see proposition 4.1.3 in [12]). Now let us derive the equations of the original system (i.e., the one in (15)) in these new coordinates. We have:

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

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

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

We remark that the first r components of x' are weakly locally observable since they are the output and its Lie derivatives along f. In order to investigate the observability properties of the remaining components, we augment the state as in (14) and we can consider the new output $\tilde{h}(x') = x'_r$. We set $L_g^1 = \tilde{g}_r = \mathcal{L}_g \mathcal{L}_f^{r-1} h \neq 0$.

8 Separating the information on the original state from the information on its extension

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 provide an automatic method able to derive a basis for this codistribution. This basis consists of two set of covectors. The first set consists of gradients of scalar functions that only depend on the original state. The second set consists of gradients of scalar functions that depend on the entire augmented state. However, the functions of this second set do not contain additional information on the original state.

For $\Sigma^{(k)}$the observable codistribution is the span of the gradients of all the Lie derivatives along F and G up to the k-order. The observable codistribution that includes a given order $m \leq k$ will be denoted by $\bar{\Omega}_m$. It is defined recursively by the following algorithm:

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

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

Definition 1 (Ω codistribution) *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 involutive, since it 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., $L^m \equiv span\{\mathcal{L}_G^1 dh, \mathcal{L}_G^2 dh, \cdots, \mathcal{L}_G^m dh\}$.

We finally introduce the following codistribution:

Definition 2 ($\tilde{\Omega}$ 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 $\mathcal{L}^1_G dh, \mathcal{L}^2_G dh, \cdots, \mathcal{L}^m_G dh$. The second set of generators, i.e., the gradients $\mathcal{L}^1_G dh, \mathcal{L}^2_G dh, \cdots, \mathcal{L}^m_G dh$, are m and, with respect to the first set, they are gradients of functions that also depend on the state extension $\xi = [w,\ w^{(1)}, \cdots,\ w^{(m-1)}]^T$. Let us stuck all the m scalar functions $\mathcal{L}^1_G h, \mathcal{L}^2_G h, \cdots, \mathcal{L}^m_G h$ in a column vector. We consider the map:

$$\mathcal{K} : \xi \to \left[\begin{array}{c} \mathcal{L}^1_G h \\ \mathcal{L}^2_G h \\ \cdots \\ \mathcal{L}^m_G h \end{array} \right]$$

In other words, we regard the previous vector as a function of only the state extension ξ. The fundamental point is that this map is one to one in a given open set. This can be easily realized by a direct computation of the Lie derivatives. In particular, it is possible to see that each component $\mathcal{L}^i_G h$, $(i = 1, \cdots, m)$ linearly depends on $w^{(i-1)}$ by the product $L^1_g w^{(i-1)}$ (we remind the reader that $L^1_g \neq 0$). As a result, 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 other words, we can adopt the procedure described in section 6 by considering the codistribution Ω_m instead of $\bar{\Omega}_m$. Specifically, we check if the row vector of dimension n, with all the entries equal to zero with the exception of the i^{th} entry equal to 1, belongs to Ω_m. If there exists an integer m such that this is true, we conclude that x_i is weakly locally observable. If this holds for $i = 1, \cdots, n$, we conclude that the original state is weakly locally observable.

In the rest of this section we will prove this fundamental theorem, stating that $\tilde{\Omega}_m = \bar{\Omega}_m$.

Theorem 1 (Separation of the information) $\bar{\Omega}_m = \tilde{\Omega}_m$

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 2 $\mathcal{L}_G \bar{\Omega}_m + \mathcal{L}_{\Phi_m} dh = \mathcal{L}_G \bar{\Omega}_m + \mathcal{L}_F \mathcal{L}^m_G dh$

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

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

Lemma 3 $\breve{\Phi}_m = \sum_{j=1}^m c^n_j (\mathcal{L}_G h, \mathcal{L}^2_G h, \cdots, \mathcal{L}^m_G h) \phi_j$, *i.e., the vector* $\breve{\Phi}_m$ *is a linear combination of the vectors* ϕ_j $(j = 1, \cdots, m)$*, where the coefficients* (c^n_j) *depend on the state only through the functions that generate the codistribution* L^m

Proof: We proceed by induction. By definition, $\breve{\Phi}_0 = \phi_0$.

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

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

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

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

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

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 $c_j^n (\equiv \tilde{c}_j + \bar{c}_j)$ 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 4 $\hat{\phi}_m = \sum_{j=1}^{m} b_j^n (\mathcal{L}_G h, \mathcal{L}_G^2 h, \cdots, \mathcal{L}_G^m h)\Phi_j$, *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 definition, $\Phi_0 = \hat{\phi}_0$.
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:

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

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

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

Property 6 *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 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: $\bar{\Omega}_m = \bar{\Omega}_{m-1} + \mathcal{L}_F \bar{\Omega}_{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1} = \bar{\Omega}_{m-1} + \mathcal{L}_F \tilde{\Omega}_{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1} = \bar{\Omega}_{m-1} + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_F L^{m-1} + \mathcal{L}_G \bar{\Omega}_{m-1}$. 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

2 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 property 6 we obtain: $\bar{\Omega}_m = [\Omega_{m-1}, 0_k] + L^m + [\mathcal{L}_f \Omega_{m-1}, 0_k] + \mathcal{L}_{\hat{\phi}_{m-1}} dh + [\mathcal{L}_{\frac{g}{L_g^1}} \Omega_{m-1}, 0_k] = \tilde{\Omega}_m$ ■

9 Applications

We apply the theory developed in the previous sections in order to investigate the observability properties of several systems with unknown inputs. In 9.1 we consider systems described by eq. (13). In 9.2 we consider the case of multiple unknown inputs.

9.1 Systems with a single disturbance

We consider a vehicle moving in 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:

$$\left[\begin{array}{l} \dot{x}_v = v\cos\theta \\ \dot{y}_v = v\sin\theta \\ \dot{\theta} = \omega \end{array}\right. \tag{20}$$

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:

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);
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: $\rho \equiv \sqrt{x_v^2 + y_v^2}$, $\phi = atan\frac{y_v}{x_v}$. We have, for the three cases, $y = \rho$, $y = \pi - (\theta - \phi)$ 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. The dynamics in these new coordinates become:

$$\left[\begin{array}{l} \dot{\rho} = v\cos(\theta - \phi) \\ \dot{\phi} = \frac{v}{\rho}\sin(\theta - \phi) \\ \dot{\theta} = \omega \end{array}\right. \tag{21}$$

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

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

By a direct computation we obtain $dim(\Omega_2) = 2$. We remark that in the case when both v and ω are known the system is invariant under rotations around the vertical axis. In other words, for a given initial state $x_0 = [\rho_0, \theta_0, \phi_0]$ the transformation defined by the continuous parameter τ: $\rho_0 \to \rho_0$, $\theta_0 \to \theta_0 + \tau$, $\phi_0 \to \phi_0 + \tau$ defines an indistinguishable set. Obviously, this set is still indistinguishable in the case when we miss one of the input. We conclude that the dimension of the observable space is equal to 2 and two independent physical quantities which are weakly locally observable are ρ and $\theta - \phi$.

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

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

We easily obtain $L^1_g = 0$ and $\mathcal{L}^1_f h = \cos(\theta - \phi)$. The relative degree of the associated system is $r = 2$. We can build the codistribution Ω_m by setting: $h = \cos(\theta - \phi)$, $L^1_g = -\sin(\theta - \phi)$, $\Omega_1 = d_x\rho + d_x h$ and by using $\Omega_m = \Omega_{m-1} + \mathcal{L}_f\Omega_{m-1} + \mathcal{L}_{\frac{g}{L^1_g}}\Omega_{m-1} + \mathcal{L}_{\phi_{m-1}} d_x h$. We have $dim(\Omega_1) = 2$. Additionally, by following the procedure adopted in the previous case we also known that the dimension of the observable space cannot exceed 2. Therefore, also in this case the dimension of the observable space is equal to 2 and two independent physical quantities which are weakly locally observable are ρ and $\theta - \phi$.

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

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

The dimension of the observable space is 1 since, in addition to the rotation invariance, the system misses information about the absolute scale (the output and the known input are angular measurements). Therefore, also ρ is unobservable. We can verify this same results by computing the dimension of Ω_m for any m: we obtain always 1.

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

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

By a direct computation we obtain $dim(\Omega_1) = 2$. Because of the rotation invariance we also known that the dimension of the observable space cannot exceed 2. Therefore, also in this case the dimension of the observable space is equal to 2 and two independent physical quantities which are weakly locally observable are ρ and $\theta - \phi$.

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

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

By a direct computation we obtain $dim(\Omega_2) = 2$. Since the system misses information about the absolute scale (the output and the known input are angular measurements). Therefore, ρ is unobservable. We conclude that the dimension of the observable space is equal to 2 and two independent physical quantities which are weakly locally observable are ϕ and θ.

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

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

We easily obtain $L^1_g = 0$ and $\mathcal{L}^1_f h = \frac{\sin(\theta-\phi)}{\rho}$. The relative degree of the associated system is $r = 2$. We can build the codistribution Ω_m by setting: $h = \frac{\sin(\theta-\phi)}{\rho}$, $L^1_g = \frac{\cos(\theta-\phi)}{\rho}$, $\Omega_1 = d_x\phi + d_x h$ and by using $\Omega_m = \Omega_{m-1} + \mathcal{L}_f\Omega_{m-1} + \mathcal{L}_{\frac{g}{L^1_g}}\Omega_{m-1} + \mathcal{L}_{\phi_{m-1}} d_x h$. We obtain $dim(\Omega_2) = 3$. 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 6. Specifically, 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 [5, 13, 14, 18, 24] and in the framework of neuroscience (e.g., [3, 6, 7]). 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 [11, 13, 14, 15, 16, 18, 23] and [26] 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 [9] and [19]. This result can be summarized as follows. In the standard Vi-SfM problem all the independent observable modes 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 angle, 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.

9.2.1 Simple 2D localization problem

We consider the system characterized by the same dynamics given in (20). 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:

$$y = [x_v,\ y_v]^T \tag{22}$$

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 (20) 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 crondition introduced in [10].

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: $\mathcal{L}^1_1 h_x = \cos\theta_v$, $\mathcal{L}^1_1 h_y = \sin\theta_v$, $\mathcal{L}^1_2 h_x = \mathcal{L}^1_2 h_y = 0$. 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 observable (actually, weakly locally observable).

We now consider the case when both the system inputs are unknown. In this case, by comparing (1) with (20) 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 in a given time interval, we also know their time derivatives. In particular, we know $\dot{x}_v(0)$ and $\dot{y}_v(0)$. From (21) we easily obtain: $\theta_v(0) = atan\left(\frac{\dot{y}_v(0)}{\dot{x}_v(0)}\right)$. Hence, also the initial orientation is observable, by only using the GPS measurements.

Let us proceed by applying the results of our theory. We start by computing the codistribution Ω^0. As in the case of only known inputs we have:

$$\Omega^0 = span\{[1,\ 0,\ 0],\ [0,\ 1,\ 0]\}$$

From this we know that any function of x_v and y_v is an observable mode. We want to know if also a function of θ_v is an observable mode (in which case the entire state would be observable). We have to compute Ω^1. For, we build the system $\Sigma^{(1)}$. We have: $x^{(1)} = [x_v,\ y_v,\ \theta_v,\ v,\ \omega]^T$. We can easily obtain the analytical expression for the quantities appearing in (5). 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 filed. We have: $\mathcal{L}^1_0 h_x = \nabla h_x \cdot f_0^{(1)}(x) = [1,\ 0,\ 0,\ 0,\ 0] \cdot [\cos\theta_v v,\ \sin\theta_v v,\ \omega,\ 0,\ 0] = \cos\theta_v v$ (similarly, we obtain $\mathcal{L}^1_0 h_y = \sin\theta_v v$). We obtain:

$$\Omega^1 = span\{[1,0,0,0,0],[0,1,0,0,0],$$
$$[0,0,-\sin\theta_v v,\cos\theta_v,0],[0,0,\cos\theta_v v,\sin\theta_v,0]\}$$

from which we obtain that Δ^1 is the span of the vector $[0,0,0,0,1]^T$. Therefore, any function of θ is an observable mode and the entire original state is observable.

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 [21, 22] (see also the work in [17] 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 non linear 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.

In [21, 22] 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 Δ^k, $\forall k \geq 2$, contains a single vector. This vector describes a system symmetry, which is an internal 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 symmetry 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 [21, 22] we compute all the Lie derivatives up to the 7^{th} 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 function 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 [21, 22]) [1].

10 Conclusion

In this paper we considered the problem of non linear observability in the case when part (or even all) of the system inputs is unknown. A new definition of indistinguishable states was provided. Then, in order to separate the effect of the known inputs from the effect of the unknown inputs on the system outputs, the original state was augmented by including the unknown inputs together with their time derivatives up to a given order. This allowed us to obtain the extension of basic properties, which hold in the standard case (i.e., when all the inputs are known). Thanks to this properties, it was possible to extend the observability rank condition introduced in the seventies to the case of unknown inputs. On the other hand, the use of this method, requires to compute a codistribution defined in the extended space. As a result, the computation demanded to apply the proposed extended observability rank condition dramatically depends on the dimension of the extended state. For this reason, we approached the fundamental issue of separating the information on the original state from the information on its extension. We fully solved

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

this problem only in the case characterized by a single disturbance and a single known input. Specifically, we provided an automatic method able to derive a basis for the aforementioned codistribution. This basis consists of two set of covectors. The first set consists of gradients of scalar functions that only depend on the original state. The second set consists of gradients of scalar functions that depend on the entire augmented state. However, the functions of this second set do not contain additional information on the original state. The analytic derivations required to perform this separation were complex and we are currently extending them to the multiple unknown inputs case.

The proposed approach was used to derive the observability properties of several systems in the framework of wheeled robotics and of the visual inertial sensor fusion.

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, The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems October 18-22, 2010, Taipei, Taiwan

[3] 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).

[4] S.P. Bhattacharyya Observer design for linear systems with unknown inputs IEEE Transactions on Automatic Control, 23 (1978)

[5] M. Bryson and S. Sukkarieh, Observability Analysis and Active Control for Airbone SLAM, IEEE Transaction on Aerospace and Electronic Systems, vol. 44, no. 1, 261–280, 2008

[6] 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

[7] 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

[8] 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.

[9] 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

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

[11] 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.

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

[13] 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.

[14] 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.

[15] 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.

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

[17] 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

[18] 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.

[19] A. Martinelli, Visual-inertial structure from motion: observability and resolvability, iros 2013, Tokyo, Japan

[20] A. Martinelli, Closed-form solution of visual-inertial structure from motion, Int. J. of Computer Vision, Vol. 106, N. 2, 138–152, 2014.

[21] A. Martinelli, Visual-inertial structure from motion: observability vs minimum number of sensors, ICRA 2014, Hong Kong.

[22] A. Martinelli, Observability Properties and Deterministic Algorithms in Visual-Inertial Structure from Motion, Foundations and Trends in Robotics, 139–209, http://dx.doi.org/10.1561/2300000030, 2014

[23] 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

[24] D. Strelow and S. Singh, Motion estimation from image and inertial measurements, International Journal of Robotics Research, 23(12), 2004

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

[26] Weiss., S., Vision Based Navigation for Micro Helicopters, PhD thesis, Diss. ETH No. 20305

Contents

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