www.oeaw.ac.at

Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

On the linearization up to multi-output injection for a class of nonlinear systems with implicitly defined outputs

S. Rodrigues, S. Pedro Aguiar

RICAM-Report 2012-16

www.ricam.oeaw.ac.at

ON THE LINEARIZATION UP TO MULTI-OUTPUT INJECTION FOR A CLASS OF NONLINEAR SYSTEMS WITH IMPLICITLY DEFINED OUTPUTS

SÉRGIO S. RODRIGUES AND ANTÓNIO PEDRO AGUIAR

Abstract. For a class of nonlinear systems we investigate the problem of under what conditions there exists a coordinate transformation that yields a state affine linear system up to output injection with implicit outputs. In particular, we provide necessary and sufficient conditions for time-varying linearization up to multi-output injection. We highlight that if the conditions hold, as a consequence, it is possible in the new coordinates to construct for a larger class of systems an observer with linear error dynamics. We propose a methodology to find the coordinate transformation. Several examples illustrate the proposed procedure.

Keywords: Nonlinear multi-output systems, time-varying linearization, observer design.

Contents

Date: June 20, 2012.

1. Introduction

The problem of identifying sub-classes in the general class of continuous-time nonlinear systems described by

$$\dot{x} = f_u(x) := f(u,\, x) \tag{1a}$$

$$y = h_u(x) := h(u,\, x) \tag{1b}$$

for which there exists, at least locally, an observer with linear error dynamics has been an active topic along the years [6, 7, 9, 10, 12–14]. More precisely, given system (1) where f and h are sufficiently smooth functions, $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is an input signal and $y \in \mathbb{R}^q$ is the measured output, a classical question is whether does exist a smooth change of coordinates $z = \Theta(x)$ (a local diffeomorphism) such that for the system written in the new coordinates it is possible to construct a state observer $\hat{z}$ so that the estimation error $\tilde{z} = \hat{z} - z$ is governed by an asymptotically stable linear (possibly time varying) dynamical system. If this is the case, then $\hat{x} = \Theta^{-1}(\hat{z})$ is an observer for the original system and the estimation error converges exponentially fast to zero as time goes to infinity.

Motivated by the above question, pioneering work in this area includes the results by Krener [12], Krener and Isidori [13]. In the former, the linearization of a nonlinear system is addressed with no reference to any output; in the latter the linearization is studied up to output injection, that is, the aim is to find $z = \Theta(x)$ for a nonlinear system $\dot{x} = f(x),\ y = h(x)$ that leads into a linear system up to an output injection

$$\dot{z} = Lz + \Phi(y), \quad y = Az$$

where L and A are linear maps and the vector field $\Phi(y)$ only depends on the known output signal y. In this case a Luenberger type observer described by

$$\dot{\hat{z}} = L\hat{z} + \Phi(y) + K(y - A\hat{x})$$

with K selected so that $L - KA$ is Hurwitz achieves a linear error dynamics $\dot{\tilde{z}} = (L - KA)\tilde{z}$, $\tilde{z} := \hat{z} - z$, where $\tilde{z}(t) \to 0$ as $t \to +\infty$.

More recently, using tools from Differential Geometry, Hammouri and Gauthier in [7,9] and Hammouri and Kinnaert in [10] extended the linearization problem to systems in form (1) to obtain time-varying linear systems up to output injection of the form

$$\dot{z} = L_u z + \Phi_u(y),\ y = P^q z \tag{2}$$

where $P^q z = [z^1\, z^2\, \ldots\, z^q]^\top \in \mathbb{R}^q$ collects the first q coordinates[1] of $z = [z^1\, z^2\, \ldots\, z^n]^\top \in \mathbb{R}^n$, $q \leq n$. In their work, necessary and sufficient conditions to the existence of the desired change of coordinates $z = \Theta(x)$ are provided.

In this paper, we depart from a different point by not restricting the *target system* to be in the form of (2). In fact, the motivation of this work emerged from the following observation: there exist relevant classes of systems that do not satisfy the conditions in [7,9,10], although it is still possible to construct an observer with linear error dynamics, see [3]. As an example, consider the following system with a perspective output and state $x = [x^1\, x^2]^\top$, $x^1 \neq 0$,

$$\dot{x} = \begin{bmatrix} x^2 \\ -x^1 + y + u \end{bmatrix}, \quad y = {}^{x^2}\!/_{x^1}.$$

[1]For a given matrix M, $M^\top$ denotes its transpose. Notice that we use super-scripts to denote the coordinates of a vector $v = [v^1\, v^2\, \ldots\, v^k]^\top \in \mathbb{R}^k$. The reason of this is because we will use some tools from Differential Geometry where often that notation is convenient.

It turns out, as we will see later on in Section 5, that it is not possible to write it in the form

$$\dot{z} = L_u z + \Phi_u(y), \quad y = z^1$$

for $z = [z^1\, z^2]^\top$. However it takes the simple form

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ y+u \end{bmatrix}, \quad x^1 y = x^2$$

for which there exists an observer with linear error dynamics. This last fact can be concluded from the results in [2,3]. Actually, we can generalize this to target systems in the form

$$\begin{aligned} &\dot{z} = L_u z + \Phi_u(y) \\ &0 = p_u^i(P^s z,\, y) := A_u^i P^s z + b_u^i(P^s z,\, y) + C_u^i y + D_u^i,\ i = 1,\, 2,\, \dots,\, p \end{aligned}$$

where $P^s z = [z^1\, z^2\, \dots\, z^s]^\top$, $A_u^i(\cdot)$ and $C_u^i(\cdot)$ are linear (row matrices with suitable dimensions), $b_u^i(\cdot,\, \cdot)$ is bilinear and $D_u^i \in \mathbb{R}$, i.e., $p_u^i(P^s z,\, y)$ $i = 1,\, 2,\, \dots,\, p$, are in particular, polynomials of degree not greater than 2 in the coordinates $z^1,\, z^2,\, \dots,\, z^s$ and $y^1,\, y^2,\, \dots,\, y^q$ of $P^s z = [z^1\, z^2\, \dots\, z^s]^\top$ and $y = [y^1\, y^2\, \dots\, y^q]^\top$. In fact, $p_u^i(P^s z,\, y)$ may be written as $p_u^i(P^s z,\, y) = A_u^i P^s z + (P^s z)^\top B_u^i y + C_u^i y + D_u^i$, where B_u^i is the matrix associated with the bilinear form b_u^i. To simplify the writing we collect all the p output equations in a single equation in $\mathbb{R}^p$ as

$$0 = A_u P^s z + (P^s z)^\top B_u y + C_u y + D_u$$

where $A_u = \mathrm{col}_{\mathbb{R}}(A_u^1,\, A_u^2,\, \dots,\, A_u^p)$ is the matrix, with entries in $\mathbb{R}$, whose rows are the A_u^i; $i = 1, \dots, p$. Similarly $C_u = \mathrm{col}_{\mathbb{R}}(C_u^1,\, C_u^2,\, \dots,\, C_u^p)$, $D_u = [D_u^1\, D_u^2\, \dots\, D_u^p]^\top \in \mathbb{R}^p$ and finally $B_u = \mathrm{col}_{\mathcal{M}}(B_u^1,\, B_u^2,\, \dots,\, B_u^p)$ is a column matrix whose entries are the matrices B_u^i; $i = 1, \dots, p$. In this case we define

$$(P^s z)^\top B_u y := [(P^s z)^\top B_u^1 y \quad (P^s z)^\top B_u^2 y \quad \dots \quad (P^s z)^\top B_u^p y]^\top \in \mathbb{R}^p.$$

As a simple illustration of these col operators, we have

$$\mathrm{col}_{\mathbb{R}}(M_a,\, M_b) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \quad \text{and} \quad \mathrm{col}_{\mathcal{M}}(M_a,\, M_b) = \begin{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \\ \\ \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \end{bmatrix}$$

for given matrices $M_a = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$ and $M_b = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}$.

In summary, the reason to consider target systems in the form

$$\dot{z} = L_u z + \Phi_u(y) \tag{3a}$$

$$0 = A_u P^s z + (P^s z)^\top B_u y + C_u y + D_u \tag{3b}$$

where $L_u \in \mathcal{M}_{n\times n}(\mathbb{R})$ is a matrix, and P^s is the "orthogonal projection" onto the first s coordinates, is that it is possible, using e.g. the results from [3] to construct an observer with linear error dynamics.

Inspired by the above arguments and the results in [3, 7, 9, 10], in the first part of the paper we address the problem of finding necessary and sufficient conditions to obtain a coordinate transformation that transform a meaningful class of systems, that will be precisely defined later on, into the target system (3). To illustrate the scope of the results derived, the second part of the paper presents a general procedure to obtain the desired coordinate transformation. The complexity of the algorithm is highly dependent on the original nonlinear system.

We now formulate precisely the problem statement.

Problem statement. We consider nonlinear systems of the form

$$\dot{x} = f_u(x) \tag{4a}$$

$$0 = A_u g(x) + g(x)^\top B_u y + C_u y + D_u \tag{4b}$$

where $x \in \Omega \subseteq \mathbb{R}^n$ is the state of the system, Ω an open subset, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^q$ are respectively the input and output signals, both assumed to be known. We consider that $f_u\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ and $g\colon \mathbb{R}^n \to \mathbb{R}^s$, with $s \le n$, are smooth (C^∞) functions in Ω; $A_u \in \mathcal{M}_{p\times s}(\mathbb{R})$, $C_u \in \mathcal{M}_{p\times q}(\mathbb{R})$ and $D_u \in \mathcal{M}_{p\times 1}(\mathbb{R})$ are matrices with real entries and $B_u \in \mathcal{M}_{p\times 1}(\mathcal{M}_{s\times q}(\mathbb{R}))$ is a column matrix with entries in $\mathcal{M}_{s\times q}(\mathbb{R})$. Notice that, as the notation suggests, these matrices and f_u may depend on the input u; we suppose that that dependence is smooth. For vectors $v_1 = [v_1^1\, v_1^2\, \dots\, v_1^s]^\top \in \mathbb{R}^s$ and $v_2 = [v_2^1\, v_2^2\, \dots\, v_2^q]^\top \in \mathbb{R}^q$, the operation $v_1^\top B_u v_2$ is to be understood as

$$v_1^\top B_u v_2 := [v_1^\top B_u^1 v_2 \quad v_1^\top B_u^2 v_2 \quad \dots \quad v_1^\top B_u^p v_2]^\top$$

for $B_u = \mathrm{col}_{\mathcal{M}}(B_u^1,\, B_u^2,\, \dots,\, B_u^p)$. Thus, the output equation (4b) is an identity in $\mathcal{M}_{1\times p}(\mathbb{R})$, i.e., in $\mathbb{R}^p$. We also consider that it completely defines y in a neighborhood of an interesting point $x_0 \in \Omega$, which means that $\mathrm{col}_{\mathbb{R}}(g(x)B_u^1,\, g(x)B_u^2,\, \dots,\, g(x)B_u^p) + C_u$ has rank q in that neighborhood. In particular, we must have $p \ge q$. For y to be q-dimensional, it also follows that $s \ge q$.

Our aim is to find conditions under which system (4) may be rewritten, up to a change of coordinates, as a system like (3).

The rest of the paper is organized as follows: Section 2 describes the necessary and sufficient conditions to be able to re-write the original system (4) in the desired target form (3). These conditions consist mainly in the existence of a suitable s-tuple of vector fields. In Section 3 we prove that those conditions are indeed necessary and sufficient. In Section 4 we present a general algorithm to find the suitable s-tuple of vector fields. A second more elegant algorithm is also described, but it works only for a particular class of systems. Section 5 illustrates the contribution of the paper with two examples that elucidate some of the particularities of the given algorithms. Brief conclusions are discussed in Section 6.

2. Linearization up to output injection

2.1. **Notation and definitions.** We assume that the reader has some familiarity with basic concepts of Differential Geometry and Control Theory. We briefly recall some terminology and provide some definitions. For a more complete discussion on what follows we suggest the works [1, 5, 11].

Given a system of coordinates $(x^1,\, x^2,\, \dots,\, x^n)$, particular vector fields in $\mathbb{R}^n$ are the ones ${}^{\partial}\!/\!_{\partial x^k}$ defined by ${}^{\partial}\!/\!_{\partial x^k}(x) = [\delta_k^1\ \delta_k^2\ \dots\ \delta_k^n]^\top \in T_x\mathbb{R}^n \sim \mathbb{R}^n$, where δ_i^j, $i,\, j \in \mathbb{N}$, is the Kronecker delta function $\delta_i^j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \ne j \end{cases}$. Related to these are the 1-forms $\mathrm{d}x^k$, defined by $\mathrm{d}x^k({}^{\partial}\!/\!_{\partial x^j}) = \delta_j^k$.

We denote by $\mathcal{V}(\Omega)$ the $C^\infty(\Omega)$-module of smooth vector fields in Ω. Any smooth vector field in an open set $\Omega \subseteq \mathbb{R}^n$ can be written, in coordinates, as a sum $V = \sum_{j=1}^n V^j(x)\partial/\partial x^j$, where $V^j \in C^\infty(\Omega)$.

We consider a differential 0-form in an open set $\Omega \subseteq \mathbb{R}^n$ as being a smooth function in $C^\infty(\Omega)$ and a differential k-form w ($k \in \mathbb{N}_0 := \mathbb{N} \setminus \{0\}$ a positive natural number) in Ω as a multi-linear and anti-symmetric map $w\colon \mathcal{V}(\Omega) \times \mathcal{V}(\Omega) \times \cdots \times \mathcal{V}(\Omega) \to \mathbb{R}$, with k factors of $\mathcal{V}(\Omega)$ in the product. For $k \geq 1$, by multi-linear we mean multi-linear in the $C^\infty(\Omega)$-module, i.e., for $V_i = \sum_{j=1}^2 \alpha_j W_j$,

$$w(V_1, \ldots, V_{i-1}, \sum_{j=1}^{2} \alpha_j W_j, V_{i+1}, \ldots, V_k) = \sum_{j=1}^{2} \alpha_j w(V_1, \ldots, V_{i-1}, W_j, V_{i+1}, \ldots, V_k)$$

for all $i \in \{1, 2, \ldots, k\}$, $\alpha_1, \alpha_2 \in C^\infty(\Omega)$ and $W_1, W_2 \in \mathcal{V}(\Omega)$. By anti-symmetric, for $k \geq 2$, we mean that if we change two vector fields the sign must be changed, i.e., for $i, j \in \{1, 2, \ldots, k\}$ with $i \neq j$ we have

$$\begin{aligned} &w(V_1, \ldots, V_{i-1}, V_i, V_{i+1}, \ldots, V_{j-1}, V_j, V_{j+1}, \ldots, V_k) \\ &= -w(V_1, \ldots, V_{i-1}, V_j, V_{i+1}, \ldots, V_{j-1}, V_i, V_{j+1}, \ldots, V_k). \end{aligned}$$

We denote by $\Lambda^k(\Omega)$ the $C^\infty(\Omega)$-module of (differential) k-forms defined in a given open set $\Omega \subseteq \mathbb{R}^n$. Notice that, by definition, $\Lambda^0(\Omega)$ coincides with the set of scalars $C^\infty(\Omega)$. In coordinates, all the 1-forms $w \in \Lambda^1(\Omega)$ may be written as $w = \sum_{i=1}^n w_i(x)\mathrm{d}x^i$ where $w_i \in C^\infty(\Omega)$ and $\mathrm{d}x^i$ was defined above. To see how the k-forms, for $k > 1$, read in coordinates we need to introduce the so-called wedge product, also called exterior product.

As usual we denote by $\alpha \wedge \beta$ the wedge product between a k-form α and a m-form β; the result is a $(k+m)$-form. If $k = 0$ we have $\alpha \wedge \beta = \alpha\beta$, the product between the scalar α and the element β in the $C^\infty(\Omega)$-module $\Lambda^m(\Omega)$. The wedge product is associative and the following property holds: $\alpha \wedge \beta = (-1)^{km}\beta \wedge \alpha$.

Given the 1-forms $w^1, w^2, \ldots, w^k$, the wedge product $w^1 \wedge w^2 \wedge \cdots \wedge w^k$ is defined by

$$w^1 \wedge w^2 \wedge \cdots \wedge w^k(V_1, V_2, \ldots, V_k) := \det[w^i(V_j)] = \det \begin{bmatrix} w^1(V_1) & w^1(V_2) & \ldots & w^1(V_k) \\ w^2(V_1) & w^2(V_2) & \ldots & w^2(V_k) \\ \vdots & \vdots & \ddots & \vdots \\ w^k(V_1) & w^k(V_2) & \ldots & w^k(V_k) \end{bmatrix}$$

for every k-tuple of vector fields $(V_1, V_2, \ldots, V_k)$. A corollary is that if for some $i \neq j$, with $i, j \in \{1, 2, \ldots, k\}$, the 1-forms w^i and w^j are equal, then $w^1 \wedge w^2 \wedge \cdots \wedge w^k = 0$.

From the above terminology, we can now conclude that every k-form $w \in \Lambda^k(\Omega)$, in coordinates with $k \geq 1$, can be written as

$$w = \sum_{\sigma \in \mathcal{S}^k_{(1,\,2,\,\ldots,\,n)}} w_\sigma(x)\mathrm{d}x^{\sigma(1)} \wedge \mathrm{d}x^{\sigma(2)} \wedge \cdots \wedge \mathrm{d}x^{\sigma(k)}$$

where $\mathcal{S}^k_{(1,\,2,\,\ldots,\,n)}$ denotes the set of all the strictly increasing length-k sub-sequences of $(1, 2, \ldots, n)$ and $w_\sigma \in C^\infty(\Omega)$. Moreover the k-forms in $\{\mathrm{d}x^{\sigma(1)} \wedge \mathrm{d}x^{\sigma(2)} \wedge \cdots \wedge \mathrm{d}x^{\sigma(k)} \mid \sigma \in \mathcal{S}_{(1,\,2,\,\ldots,\,n)}\}$ are linearly independent in $\Lambda^k(\Omega)$, i.e., $\sum_{\sigma \in \mathcal{S}^k_{(1,\,2,\,\ldots,\,n)}} w_\sigma(x)\mathrm{d}x^{\sigma(1)} \wedge \mathrm{d}x^{\sigma(2)} \wedge \cdots \wedge \mathrm{d}x^{\sigma(k)}$ vanishes iff the functions $w_\sigma(x)$ are identically zero in Ω.

The following definitions will also be needed. For $k \geq 1$, given a k-form w and a vector field X the interior product $\iota_X w$ is a $(k-1)$-form defined as

$$\iota_X w(V_1, V_2, \ldots, V_{k-1}) := w(X, V_1, V_2, \ldots, V_{k-1})$$

for every vector fields $V_1, V_2, \ldots, V_{k-1}$. Analogously for $k \geq r$, and r vector fields $X_1, X_2, \ldots, X_r$ we may define $\iota_{(X_1, X_2, \ldots, X_r)}w$ recursively as $\iota_{(X_1, X_2, \ldots, X_r)}w := \iota_{X_r}\iota_{(X_1, X_2, \ldots, X_{r-1})}w$, which gives

$$\iota_{(X_1, X_2, \ldots, X_r)}w(V_1, V_2, \ldots, V_{k-r}) := w(X_1, X_2, \ldots, X_r, V_1, V_2, \ldots, V_{k-r})$$

for every vector fields $V_1, V_2, \ldots, V_{k-r}$.

Recall also that the exterior derivative d in a local coordinate system $(x^1, x^2, \ldots, x^n)$ is given by $\mathrm{d}f = \sum_{i=1}^n {}^{\partial f}/_{\partial x^i}\mathrm{d}x^i$ for any function $f \in \Lambda^0(\Omega)$ and is the unique $\mathbb{R}$-linear mapping sending k-forms into $(k+1)$-forms and satisfying $\mathrm{d}(\mathrm{d}\alpha) = 0$ and $\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^k \alpha \wedge \mathrm{d}\beta$ for every k-form α and r-form β; $k, r \in \mathbb{N}$. It turns out that the form $\mathrm{d}x^k$ is the exterior derivative of the coordinate function $[x^1\ x^2\ \ldots\ x^n]^\top \mapsto x^k$.

We say that a system $(h^1, h^2, \ldots, h^n)$ of smooth functions is a system of coordinates in $\Omega \subseteq \mathbb{R}^n$ if $\mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^n \neq 0$.

Remark 2.1. Notice that from

$$\mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^n({}^{\partial}/_{\partial x^1}, {}^{\partial}/_{\partial x^2}, \ldots, {}^{\partial}/_{\partial x^n}) = \det \begin{bmatrix} \frac{\partial h^1}{\partial x^1} & \frac{\partial h^1}{\partial x^2} & \cdots & \frac{\partial h^1}{\partial x^k} \\ \frac{\partial h^2}{\partial x^1} & \frac{\partial h^2}{\partial x^2} & \cdots & \frac{\partial h^2}{\partial x^k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h^k}{\partial x^1} & \frac{\partial h^k}{\partial x^2} & \cdots & \frac{\partial h^k}{\partial x^k} \end{bmatrix} =: \det \frac{\partial h}{\partial x}$$

it follows that $(h^1, h^2, \ldots, h^n)$ is a system of coordinates if $\det \frac{\partial h}{\partial x} \neq 0$. In particular, $\mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^n \neq 0$. On the other hand if $\det \frac{\partial h}{\partial x} = 0$, by the multi-linearity and anti-symmetry of $\mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^n$, re-writing $V_i = \sum_{j=1}^n V_i^j {}^{\partial}/_{\partial x^j}$, $i = 1, 2, \ldots, n$, we can conclude that $\mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^n(V_1, V_2, \ldots, V_n) = 0$, for all n-tuple $(V_1, V_2, \ldots, V_n)$, i.e., $\mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^n = 0$.

Remark 2.2. From $\mathrm{d}f = \sum_{i=1}^n {}^{\partial f}/_{\partial x^i}\mathrm{d}x^i$ we can conclude that $\mathrm{d}f({}^{\partial}/_{\partial x^k}) = {}^{\partial f}/_{\partial x^k}$ for all $k \in \{1, 2, \ldots, n\}$. On the other side, given another local system of coordinates $(h^1, h^2, \ldots, h^n)$ we define, for each $i \in \{1, 2, \ldots, n\}$ the vector field ${}^{\partial}/_{\partial h^i}$ by the equations: $\mathrm{d}h^j({}^{\partial}/_{\partial h^i}) = \delta_i^j$, $\quad j = 1, 2, \ldots, n$. For every smooth function f we find

$$\mathrm{d}f = \sum_{i=1}^n {}^{\partial f}/_{\partial x^i}\mathrm{d}x^i = \sum_{i=1}^n \left(\sum_{j=1}^n {}^{\partial f}/_{\partial h^j}{}^{\partial h^j}/_{\partial x^i} \right) \mathrm{d}x^i = \sum_{j=1}^n {}^{\partial f}/_{\partial h^j} \sum_{i=1}^n {}^{\partial h^j}/_{\partial x^i}\mathrm{d}x^i = \sum_{j=1}^n {}^{\partial f}/_{\partial h^j}\mathrm{d}h^j$$

and $\mathrm{d}f({}^{\partial}/_{\partial h^k}) = {}^{\partial f}/_{\partial h^k}$, for all $k \in \{1, 2, \ldots, n\}$. Thus

$$\mathrm{d}f({}^{\partial}/_{\partial x^k}) = \sum_{j=1}^n {}^{\partial f}/_{\partial h^j}{}^{\partial h^j}/_{\partial x^k} = \mathrm{d}f\left(\sum_{j=1}^n {}^{\partial h^j}/_{\partial x^k}{}^{\partial}/_{\partial h^j} \right),$$

from which we can conclude that

$$ {}^{\partial}/_{\partial x^k} = \sum_{j=1}^n {}^{\partial h^j}/_{\partial x^k}{}^{\partial}/_{\partial h^j} \tag{5}$$

because f is an arbitrary function. This identity gives us the relation between the "canonical" bases of vector fields associated to the two system of coordinates $(x^1, x^2, \ldots, x^n)$ and $(h^1, h^2, \ldots, h^n)$.

Given a vector field X, the Lie derivative $\mathcal{L}_X$ can be defined on k-forms as follows: $\mathcal{L}_X(w) := \mathcal{L}_X w := \mathrm{d}w(X) = \iota_X \mathrm{d}w$ for a function w and as $\mathcal{L}_X w := (\iota_X \mathrm{d} + \mathrm{d}\iota_X)w$ for a

k-form w with $k \geq 1$. In particular, given a vector field $V = \sum_{j=1}^{n} V^j \partial/\partial x^j$ we have that $V^j = \mathrm{d}x^j(V) = \mathcal{L}_V x^j$.

For the same vector field $X = \sum_{i=1}^{n} X^i \partial/\partial x^i$, we may also define the Lie derivative of a vector field $Y = \sum_{i=1}^{n} Y^i \partial/\partial x^i$ by setting, in coordinates,

$$\mathcal{L}_X Y = [X,\, Y] := \sum_{i=1}^{n} \sum_{j=1}^{n} \left(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \right) \partial/\partial x^i \tag{6}$$

where the vector field $[X,\, Y]$ is called the Lie bracket between the vector fields X and Y. The following identity that we will use later on holds:

$$\mathcal{L}_{[X,\, Y]} f = \mathcal{L}_X \mathcal{L}_Y f - \mathcal{L}_Y \mathcal{L}_X f \tag{7}$$

for every function $f \in \Lambda^0(\Omega)$.

2.2. **Main theorem: necessary and sufficient conditions.** We recall that our task is to find the conditions to the existence of a change of coordinates that carries (4) into the target form (3). To this effect, we need to introduce the following auxiliary system

$$\dot{x} = f_u(x) \tag{8a}$$

$$\bar{y} = g(x) \tag{8b}$$

where f_u and g are obtained from (4). As we will see later the auxiliary output $\bar{y}$ will be the set of candidates to the first new s coordinate functions.

We assume that this auxiliary system is observable in the rank sense (see definition e.g., [4]) in a neighborhood of a given point x_0, i.e., denoting by $\mathcal{O}$ the smallest set containing $\{g^1,\, g^2,\, \ldots,\, g^s\}$ and closed under all the Lie derivatives $\{\mathcal{L}_{f_u} \mid u \text{ is a constant in } \mathbb{R}^m\}$ we have that $\{\mathrm{d}w|_{x_0} \mid w \in \mathcal{O}\}$ is n-dimensional. Here $\mathrm{d}w|_{x_0}$ is the evaluation at x_0 of the so-called exterior derivative $\mathrm{d}w$ of w.

Let $X = (X_1,\, X_2,\, \ldots,\, X_s)$ be a s-tuple of vector fields. Following a similar idea in [7,9], define a sequence of vector spaces as follows: set $\mathrm{d}\Gamma := \mathrm{d}g^1 \wedge \mathrm{d}g^2 \wedge \cdots \wedge \mathrm{d}g^s$ and denote by $\Omega_1^X = (\Omega_1^X)_{[\text{system } (8)]}$ the real vector space generated by the set of $(1+s)$-forms $\{\mathrm{d}\mathcal{L}_{f_u} g^j \wedge \mathrm{d}\Gamma \mid 1 \leq j \leq s \text{ and } u \in \mathbb{R}^m\}$, that is

$$\Omega_1^X := \mathrm{span}_{\mathbb{R}}\{\mathrm{d}\mathcal{L}_{f_u} g^j \wedge \mathrm{d}\Gamma \mid 1 \leq j \leq s \text{ and } u \in \mathbb{R}^m\}. \tag{9}$$

Recursively, for $k \geq 2$, define the real vector space $\Omega_k^X = (\Omega_k^X)_{[\text{system } (8)]}$ as

$$\Omega_k^X := \mathrm{span}_{\mathbb{R}}\{\mathcal{L}_{f_u}(\iota_X w) \wedge \mathrm{d}\Gamma \mid w \in \Omega_{k-1}^X \text{ and } u \in \mathbb{R}^m\}.$$

Define also the smallest real vector space $\Omega^X = (\Omega^X)_{[\text{system } (8)]}$ containing all these previous ones by

$$\Omega^X := \mathrm{span}_{\mathbb{R}}\{w \mid w \in \Omega_k^X \text{ and } k \in \mathbb{N}_0\},$$

where $\mathbb{N}_0 := \mathbb{N} \setminus \{0\}$ denotes the set of positive natural numbers, and consider the vector space

$$\Omega[X,\, g] := \mathrm{span}_{\mathbb{R}}(\, \iota_X \Omega_X \cup \{\mathrm{d}g\}\,)$$

with $\{\mathrm{d}g\} := \{\mathrm{d}g^j \mid j = 1,\, 2,\, \ldots,\, s\}$.

Finally denoting the q-form $\mathrm{d}\Upsilon := \mathrm{d}y^1 \wedge \mathrm{d}y^2 \wedge \cdots \wedge \mathrm{d}y^q$, we are ready to present the necessary and sufficient conditions to be able to transform the system (4) into the target form (3).

Theorem 2.1. *Consider a given point x_0 and suppose that in a neighborhood U of x_0, system (8) is observable in the rank sense, $\mathrm{d}\Upsilon \neq 0$ and $\mathrm{d}y^j \wedge \mathrm{d}\Gamma = 0$ for every $j = 1,\, 2,\, \ldots,\, q$. Then up to a change of coordinates, system (4) can be written in the form of (3) in a sub-neighborhood $\mathcal{N} \subseteq U$ of x_0 iff*

a. $\mathrm{d}\Gamma|_{x_0} \neq 0$;
b. *there exists a s-tuple of vector fields* $X = (X_1, X_2, \ldots, X_s)$ *such that:*
   i. $\mathcal{L}_{X_i} g^j = \delta_i^j$ *in* U;
   ii. *the real dimension of* Ω^X *is equal to* $n - s$ *in* $\mathcal{N}$;
   iii. $\mathrm{d}\iota_X \Omega^X = \{0\}$ *in* $\mathcal{N}$;
   iv. $\bigwedge^{n-s} \iota_X \Omega^X \wedge \mathrm{d}\Gamma|_{x_0} \neq \{0\}$ *and*
   v. $\mathrm{d}\iota_{f_u} \Omega[X, g] \wedge \mathrm{d}\Upsilon \subseteq \Omega[X, g] \wedge \mathrm{d}\Upsilon$ *in* $\mathcal{N}$.

At this point, a number of remarks on the obtained conditions are discussed.

(1) The first observation is that comparing the required conditions with the results in [7,9,10] we notice that we have mainly two new statements: $\mathrm{d}y^j \wedge \mathrm{d}\Gamma = 0$ and b.v. The former means that it is possible to write the output equation (4b) in explicit form as a function of g. The latter is due to the fact that, although that possibility exists, we would like to preserve the implicit notation. It is important to stress that if the output is written in explicit form we recover the results in [7,9,10] because in that case we have $\mathrm{d}\Upsilon = \mathrm{d}\Gamma$ and the last condition b.v. will follow from the preceding ones and from the definitions.
(2) A second observation refers to condition a. Since the goal is to find a suitable set of functions forming a new coordinate system $(z^1, z^2, \ldots, z^n)$, where the functions $g^1, g^2, \ldots, g^s$ will be part of this new coordinate system, that is, $z^i = g^i$ for $i \in \{1, 2, \ldots, s\}$, then as the condition essentially says, the functions $g^1, g^2, \ldots, g^s$ must be independent.
(3) Regarding condition b.i., we may view the vector fields X_i as the vector fields $\partial/\partial g^i$ in the new system of coordinates that we are looking for. In this case, that condition is reasonable. The reason to think about X_i as $\partial/\partial g^i \sim \partial/\partial z^i$ will be clear in the proof of the theorem in Section 3.
(4) The set Ω^X, in Theorem 2.1, is constructed by means of iterated application of Lie derivatives and wedge product with $\mathrm{d}\Gamma$. The latter is used to kill at each step any dependence on g, which roughly speaking means that up to some dependence on g we have that Ω^X is composed by elements of the form $w \wedge \mathrm{d}\Gamma$, where w "is not far from" an iterated Lie derivative of some g^j. For a system in target form these $w \wedge \mathrm{d}\Gamma$ are essentially products of matrices (times $[\mathrm{d}z^1\, \mathrm{d}z^2\, \ldots\, \mathrm{d}z^n]^\top$); so the dimension of Ω^X is at most $n - s$. Thus, using the observability rank condition it follows that the dimension has to be $n - s$. Note that in b.ii. we consider the dimension as a vector space over $\mathbb{R}$.
(5) Condition b.iii. essentially says that for any $w \in \iota_X \Omega^X$ we can find z such that $\mathrm{d}z = w$ in a suitable neighborhood. Together with condition b.iv. it follows that we can find z^j, $j = s+1, s+2, \ldots, n$ such that $\mathrm{d}z^{s+1} \wedge \mathrm{d}z^{s+2} \wedge \cdots \wedge \mathrm{d}z^n \wedge \mathrm{d}\Gamma|_{x_0} \neq \{0\}$, i.e., the set $(z^1, z^2, \ldots, z^n)$ is a system of coordinates. Thus, at this point we now have a candidate for the system of coordinates we are looking for. We stress that, as indicated, the first s candidates $z^1, z^2, \ldots, z^s$ are completed by functions $z^{s+1}, z^{s+2}, \ldots, z^n$ derived from a basis $\{\mathrm{d}z^j \mid j = s+1, s+2, \ldots, n\}$ for $\iota_X \Omega^X$.
(6) If all the conditions in Theorem 2.1 until b.iv. holds we can conclude that there exists a new system of coordinates that transforms the vector field f_u of (8a) into a linear one up to a function $\bar{\Phi}_u(P^s z)$ that depends on the first s coordinates $g = P^s z = [z^1\, z^2\, \ldots\, z^s]^\top$. It is at this point that we need the last condition b.v.. It essentially guarantees that we can re-write the function $\bar{\Phi}_u(P^s z)$ as a sum of a

linear function Π_u depending on $P^s z$ with a function Φ_u depending only on the output y, that is, $\bar{\Phi}_u(P^s z) = \Pi_u P^s z + \Phi_u(y)$.

Remark 2.3. To have the desired output equation (3b) we set $z^i = g^i$ for $i \in \{1, 2, \ldots, s\}$. For the rest of the coordinate functions we may choose them as indicated, from a basis $\{\mathrm{d}z^j \mid j = s+1, s+2, \ldots, n\}$ for $\iota_X \Omega^X$. Notice, however that these last coordinates can be replaced by a family $\{\tilde{z}^j \mid j = s+1, s+2, \ldots, n\}$, where $\tilde{z}^j \in \mathrm{span}_{\mathbb{R}}\{z^i \mid i = 1, 2, \ldots, n\}$ and $\mathrm{d}\Gamma \wedge \mathrm{d}\tilde{z}^{s+1} \wedge \mathrm{d}\tilde{z}^{s+2} \wedge \cdots \wedge \mathrm{d}\tilde{z}^n \neq 0$. Applying this linear change preserves the form of the target system, which means that both $(z^1, z^2, \ldots, z^n)$ and $(z^1, z^2, \ldots, z^s, \tilde{z}^{s+1}, \tilde{z}^{s+2}, \ldots, \tilde{z}^n)$ lead us to the desired target system form.

3. Proof of the main Theorem

In this section we provide the proof of Theorem 2.1. To simplify the reading, the proof is organized into several parts.

3.1. **Proof of necessity.** We start to denote by $\mathcal{S}_{(u_1, u_2, \ldots, u_k)}$ the set of all sub-sequences α of $(u_1, u_2, \ldots, u_k)$, $\alpha = \alpha(\sigma) = (u_{\sigma(1)}, u_{\sigma(2)}, \ldots, u_{\sigma(k_1)})$, where $\sigma \in \mathcal{S}_{(1,2,\ldots,k)}$ is a strictly increasing sub-sequences of $(1, 2, \ldots, k)$, i.e., σ is represented by a strictly increasing function from $\{1, 2, \ldots, k\}$ into itself; in particular $k_1 \leq k$. If σ_1 and σ_2 are two of these strictly increasing functions with disjoint images, i.e., if $\sigma_1(\{1, 2, \ldots, k\}) \cap \sigma_2(\{1, 2, \ldots, k\}) = \emptyset$, then we write $\alpha(\sigma_1) \cap \alpha(\sigma_2) = \emptyset$.

Now, we need the following auxiliary lemmas. The proofs can be found in the Appendix.

Lemma 3.1. *Consider a system of coordinates* $(h^1, h^2, \ldots, h^n)$ *and a natural number* $r < n$. *The* $(r+1)$*-forms* $\{\mathrm{d}h^k \wedge \mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^r \mid k = r+1, r+2, \ldots, n\}$ *are linearly independent in* $\bigwedge^{r+1}(\Omega)$.

Lemma 3.2. *Let* $\{h^j \mid j = 0, 1, \ldots, r\}$ *be a set of smooth functions such that* $\mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^r \neq 0$ *in a neighborhood* $U \subseteq \Omega$ *of a given point* x_0. *Then* $\mathrm{d}h^0 \wedge \mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^r = 0$ *in* U *only if there exists a function* $\phi : \mathbb{R}^r \to \mathbb{R}$ *such that* $h^0 = \phi(h^1, h^2, \ldots, h^r)$ *in a neighborhood* $U_1 \subseteq U$ *of the point* x_0.

Lemma 3.3. *If* $h^0 = \phi(h^1, h^2, \ldots, h^r)$ *in a neighborhood* U *of the point* x_0, *then* $\mathrm{d}h^0 \wedge \mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^r = 0$.

A similar result to the following lemma may be found in [7] for the case of a single (one-dimensional) output.

Lemma 3.4. *Let the target system* (3) *with output* $y \in \mathbb{R}^q$ *satisfy* $\mathrm{d}y^j \wedge \mathrm{d}\Gamma = 0$ *for every* $j \in \{1, 2, \ldots, q\}$ *with* $\mathrm{d}\Gamma = \mathrm{d}z^1 \wedge \mathrm{d}z^2 \wedge \cdots \wedge \mathrm{d}z^s$. *Rewrite* $F_u(z) := L_u z + \Phi_u(y)$ *as* $\bar{L}_u Q^s z + \tilde{L}_u P^s z + \Phi_u(y)$ *where* $P^s[z^1\, z^2\, \ldots\, z^n]^\top := [z^1\, z^2\, \ldots\, z^s]^\top$, $Q^s[z^1\, z^2\, \ldots\, z^n]^\top := [z^{s+1}\, z^{s+2}\, \ldots\, z^n]^\top$, $\bar{L}_u \in \mathcal{M}_{n\times(n-s)}(\mathbb{R})$ *is the matrix whose columns are the last* $(n-s)$ *columns of* L_u *and* $\tilde{L}_u \in \mathcal{M}_{n\times s}(\mathbb{R})$ *is the matrix whose columns are the first* s *columns of* L_u *and* $L_u = [\tilde{L}_u\ \bar{L}_u]$. *Then denote* $M_{u_1 u_2 \ldots u_k} := \bar{L}_{u_1} Q^s \bar{L}_{u_2} Q^s \ldots \bar{L}_{u_k} Q^s$. *Set also* $\mathrm{d}z := [\mathrm{d}z^1\, \mathrm{d}z^2\, \ldots\, \mathrm{d}z^n]^\top$. *With this notation, setting* $\tilde{X} = ({\partial}/{\partial z^1}, {\partial}/{\partial z^2}, \ldots, {\partial}/{\partial z^s})$, *we have that for every* $k \in \mathbb{N}_0$

$$\Omega_k^{\tilde{X}} = \mathrm{span}_{\mathbb{R}}\{\pi^j M_{u_1 u_2 \ldots u_k} \mathrm{d}z \wedge \mathrm{d}\Gamma \mid j = 1, 2, \ldots, s \text{ and } u_1, u_2, \ldots, u_k \in \mathbb{R}^m\} \tag{10}$$

and, denoting $\mathcal{L}_u := \mathcal{L}_{F_u}$ *for* $u \in \mathbb{R}^m$,

$$\begin{aligned}\mathcal{L}_{u_k} \ldots \mathcal{L}_{u_2}\mathcal{L}_{u_1} z^j &= \pi^j M_{u_1 u_2 \ldots u_k} z + a^k \\ &+ \sum_{\substack{r\in\{1,2,\ldots,k-1\}\\ \alpha_i\in\mathcal{S}_{(u_2,\,u_3,\,\ldots,\,u_k)}\\ \alpha_i\cap\alpha_j=\emptyset\\ \{i_1,i_2,\ldots,i_r\}\subseteq\{1,2,\ldots,s\}.}} (a^k_{\alpha_1\alpha_2\ldots\alpha_r})(\pi^{i_1} M_{\alpha_1} z)(\pi^{i_2} M_{\alpha_2} z)\ldots(\pi^{i_r} M_{\alpha_r} z)\end{aligned} \tag{11}$$

where a^k *and* $a^k_{\alpha_1\alpha_2\ldots\alpha_r}$ *are real functions depending only in* $P^s z$ *and* $u_1,\, u_2,\, \ldots,\, u_k$.

With the previous lemmas we have now the tools to prove the necessity of the conditions in Theorem 2.1. First note that conditions a. and b.ii.-b.v. are intrinsic ones, that is, they do not depend on the system of coordinates. Thus it is sufficient to check them for the target system (3) for a given point z_0. Notice that in this case $\mathrm{d}\Gamma = \mathrm{d}z^1 \wedge \mathrm{d}z^2 \wedge \cdots \wedge \mathrm{d}z^s$ and $\{\mathrm{d}g\} = \{\mathrm{d}z^1,\, \mathrm{d}z^2,\, \ldots,\, \mathrm{d}z^s\}$.

We claim that the conditions in Theorem 2.1 hold for a system in the target form (3) with the s-tuple of vector fields

$$\tilde{X} = \left({}^{\partial}\!/\!_{\partial z^1},\, {}^{\partial}\!/\!_{\partial z^2},\, \ldots,\, {}^{\partial}\!/\!_{\partial z^s}\right).$$

Condition a. is satisfied because $(z^1,\, z^2,\, \ldots,\, z^n)$ is a system of coordinates.

By hypothesis the auxiliary system

$$\begin{aligned}\dot{z} &= L_u z + \Phi_u(y)\\ \bar{y} &= P^s z\end{aligned}$$

is observable in the rank sense at z_0. Then by the identity (11) we have that the real vector space $\mathcal{V}$ spanned by the 1-forms of the form

$$\begin{aligned}&\pi^j M_{u_1u_2\ldots u_k}\mathrm{d}z + \sum_{m=1}^{s} {}^{\partial a^k}\!/\!_{\partial z^m}|_{z_0}\mathrm{d}z^m\\ &+\sum_{\substack{r\in\{1,2,\ldots,k-1\}\\ \alpha_i\in\mathcal{S}_{(u_2,\,u_3,\,\ldots,\,u_k)}\\ \alpha_i\cap\alpha_j=\emptyset\\ \{i_1,i_2,\ldots,i_r\}\subseteq\{1,2,\ldots,s\}.}} \left(\sum_{m=1}^{s} {}^{\partial a^k_{\alpha_1\alpha_2\ldots\alpha_r}}\!/\!_{\partial z^m}|_{z_0}\mathrm{d}z^m\right)(\pi^{i_1}M_{\alpha_1}z_0)(\pi^{i_2}M_{\alpha_2}z_0)\ldots(\pi^{i_r}M_{\alpha_r}z_0)\\ &+\sum_{\substack{r\in\{1,2,\ldots,k-1\}\\ \alpha_i\in\mathcal{S}_{(u_2,\,u_3,\,\ldots,\,u_k)}\\ \alpha_i\cap\alpha_j=\emptyset\\ \{i_1,i_2,\ldots,i_r\}\subseteq\{1,2,\ldots,s\}\\ l\in\{1,2,\ldots,r\}}} (a^k_{\alpha_1\alpha_2\ldots\alpha_r}(z_0))(\pi^{i_1}M_{\alpha_1}z_0)\ldots(\pi^{i_{l-1}}M_{\alpha_{l-1}}z_0)(\pi^{i_l}M_{\alpha_l}\mathrm{d}z)(\pi^{i_{l+1}}M_{\alpha_{l+1}}z_0)\ldots(\pi^{i_r}M_{\alpha_r}z_0),\end{aligned}$$

where $u_1,\, u_2,\, \ldots,\, u_k \in \mathbb{R}^m$, $j \in \{1,\, 2,\, \ldots,\, s\}$ and $k \in \mathbb{N}_0$, together with the 1-forms in $\{\mathrm{d}g\}$, is n-dimensional so, necessarily the space $\mathcal{V} \wedge \mathrm{d}\Gamma$ is $(n-s)$-dimensional. This is the space spanned by the vectors of the form

$$\begin{aligned}&\pi^j M_{u_1u_2\ldots u_k}\mathrm{d}z\wedge\mathrm{d}\Gamma \quad +\\ &\sum_{\substack{r\in\{1,2,\ldots,k-1\}\\ \alpha_i\in\mathcal{S}_{(u_2,\,u_3,\,\ldots,\,u_k)}\\ \alpha_i\cap\alpha_j=\emptyset\\ \{i_1,i_2,\ldots,i_r\}\subseteq\{1,2,\ldots,s\}\\ l\in\{1,2,\ldots,r\}}}(a^k_{\alpha_1\alpha_2\ldots\alpha_r}(z_0))(\pi^{i_1}M_{\alpha_1}z_0)\ldots(\pi^{i_{l-1}}M_{\alpha_{l-1}}z_0)(\pi^{i_l}M_{\alpha_l}\mathrm{d}z\wedge\mathrm{d}\Gamma)(\pi^{i_{l+1}}M_{\alpha_{l+1}}z_0)\ldots(\pi^{i_r}M_{\alpha_r}z_0)\end{aligned}$$

that, from (10), is a subspace of $\cup_{k\in\mathbb{N}_0}\Omega_k^{\tilde{X}} \subseteq \Omega^{\tilde{X}}$. Necessarily $\Omega^{\tilde{X}}$ is $(n-s)$-dimensional and so b.ii. holds.

Since $\Omega^{\tilde{X}}$ is $(n-s)$-dimensional, from (10) it follows that necessarily $\Omega^{\tilde{X}} = \mathrm{span}_{\mathbb{R}}\{\mathrm{d}z^j \wedge \mathrm{d}\Gamma \mid j = s+1,\, s+2,\, \ldots,\, n\}$ and, from $\iota_{\tilde{X}}(\mathrm{d}z^j \wedge \mathrm{d}\Gamma) = (-1)^s \mathrm{d}z^j$ for every $j = s+1,\, s+2,\, \ldots,\, n$, we find $\iota_{\tilde{X}}\Omega^{\tilde{X}} = \mathrm{span}_{\mathbb{R}}\{\mathrm{d}z^j \mid j = s+1,\, s+2,\, \ldots,\, n\}$ so b.iii. holds. Since $\mathrm{d}z^{s+1} \wedge \mathrm{d}z^{s+2} \wedge \cdots \wedge \mathrm{d}z^n \wedge \mathrm{d}\Gamma \neq 0$ we can also conclude that b.iv. holds. Finally for b.v. just notice that $\mathrm{d}(\iota_{F_u}\mathrm{d}z^j) \wedge \mathrm{d}\Upsilon = \mathrm{d}(\sum_{i=1}^n L_u^{ji} z^i + \Phi_u^j(y)) \wedge \mathrm{d}\Upsilon$ and since $\mathrm{d}\Phi_u^j \wedge \mathrm{d}\Upsilon = 0$ we have that $\mathrm{d}(\iota_{F_u}\mathrm{d}z^j) \wedge \mathrm{d}\Upsilon = \sum_{i=1}^n L_u^{ji}\mathrm{d}z^i \wedge \mathrm{d}\Upsilon$ is in $\mathrm{span}_{\mathbb{R}}\{\mathrm{d}z^j \mid j = 1,\, 2,\, \ldots,\, n\} \wedge \mathrm{d}\Upsilon = \Omega[\tilde{X},\, g] \wedge \mathrm{d}\Upsilon$.

The necessity of the conditions is proven.

3.2. **Proof of sufficiency.** The proof of sufficiency resorts to some results described in [7,9]. Proceeding as it is presented in there, from the above conditions a. and b.i.-b.iv. we conclude that there exists a change of coordinates which transforms (8) into a system in the form

$$\dot{z} = \bar{L}_u Q^s z + \bar{\Phi}_u(P^s z), \quad \bar{y} = P^s z$$

where $\bar{L}_u = [\bar{L}_u^{ji}] \in \mathcal{M}_{n\times(n-s)}(\mathbb{R})$ is a matrix; $j = 1,\, 2,\, \ldots,\, n$; $i = 1,\, 2,\, \ldots,\, n-s$. Now it is clear that the same change of coordinates transforms the system (4) into the system

$$\dot{z} = \bar{L}_u Q^s z + \bar{\Phi}_u(P^s z) \tag{12a}$$

$$0 = A_u P^s z + (P^s z)^\top B_u y + C_u y + D_u. \tag{12b}$$

Notice that, although the function output y is defined implicitly by means of g and $P^s z = \bar{y} = g$, system (12) is not yet necessarily in the form of system (3) because $\bar{\Phi}$ does not depend necessarily only on y. It is at this point that we need the new condition b.v. in the statement of the theorem. From [9], we know that $\mathrm{d}z = [\mathrm{d}z^1\ \mathrm{d}z^2\ \ldots\ \mathrm{d}z^n]^\top$ is given by $\begin{cases} \mathrm{d}z^j = \mathrm{d}g^j, & \text{if } j = 1,\, \ldots,\, s; \\ \mathrm{d}z^j = \iota_X w^j, & \text{if } j = s+1,\, \ldots,\, n; \end{cases}$ where $\{w^j \mid j = s+1,\, \ldots,\, n\}$ is a basis for the real vector space Ω^X. Actually a new system of coordinates may be obtained by setting $z^j = g^j$ for $j = 1,\, \ldots,\, s$ and by a local solution z^j for $\mathrm{d}z^j = \zeta^j$ for $j = s+1,\, \ldots,\, n$, where $\zeta^j \in \Omega[X,\, g] = \mathrm{span}_{\mathbb{R}}(\{\mathrm{d}g^i \mid i = 1,\, \ldots,\, s\} \cup \{w^j \mid j = s+1,\, \ldots,\, n\})$ and $\mathrm{d}g^1 \wedge \mathrm{d}g^2 \wedge \cdots \wedge \mathrm{d}g^s \wedge \zeta^{s+1} \wedge \zeta^{s+2} \wedge \cdots \wedge \zeta^n \neq 0$. The existence of these solutions is due to the condition b.iii., saying that all the forms in $\iota_X\Omega^X$ are closed, and to condition b.iv..

From (12a) we have that the vector field f_u in the coordinates $(z^1,\, z^2,\, \ldots,\, z^n)$ reads

$$f_u = \sum_{j=1}^n F_u^j \partial/\partial z^j$$

with $F_u^j = \sum_{i=s+1}^n \bar{L}_u^{ji} z^i + \bar{\Phi}_u^j(P^s z)$ for each $j = 1,\, 2,\, \ldots,\, n$. Since $F_u^j = \mathcal{L}_{f_u} z^j = \iota_{f_u}\mathrm{d}z^j$ and $\mathrm{d}z^j \in \Omega[X,\, g]$ we have, from b.v., that there exist $\alpha^j = \sum_{i=1}^n \alpha^{ji}\mathrm{d}z^i$, with $\alpha^{ji} \in \mathbb{R}$ for $j,\, i = 1,\, 2,\, \ldots,\, n$, such that

$$\begin{aligned} &\mathrm{d}F_u^j \wedge \mathrm{d}\Upsilon - \sum_{i=1}^n \alpha^{ji}\mathrm{d}z^i \wedge \mathrm{d}\Upsilon = 0 \\ \Leftrightarrow &\sum_{i=s+1}^n \bar{L}_u^{ji}\mathrm{d}z^i \wedge \mathrm{d}\Upsilon + \mathrm{d}\bar{\Phi}_u^j \wedge \mathrm{d}\Upsilon - \sum_{i=1}^n \alpha^{ji}\mathrm{d}z^i \wedge \mathrm{d}\Upsilon = 0 \\ \Leftrightarrow &\mathrm{d}\left(\sum_{i=s+1}^n \bar{L}_u^{ji} z^i + \bar{\Phi}_u^j(P^s z) - \sum_{i=1}^n \alpha^{ji} z^i\right) \wedge \mathrm{d}\Upsilon = 0. \end{aligned} \tag{13}$$

By Lemma 3.2 it follows that $\sum_{i=s+1}^{n} \bar{L}_u^{ji} z^i + \bar{\Phi}_u^j(P^s z) - \sum_{i=1}^{n} \alpha^{ji} z^i$ is a function that may be defined by means of $(y^1, y^2, \ldots, y^q)$ and, from the conditions $\mathrm{d}y^j \wedge \mathrm{d}\Gamma = 0$ each y^j is a function of $g = [g^1\ g^2\ \ldots\ g^s]^\top$, i.e., $y^j = y^j(g) = y^j(P^s z)$ and $P^s z = [z^1\ z^2\ \ldots\ z^s]^\top$. Thus, from $\mathrm{d}y^j = \sum_{i=1}^{s} {}^{\partial y^j}\!/\!_{\partial z^i} \mathrm{d}z^i$ we can conclude that $\mathrm{d}\Upsilon = \sum_{\sigma \in \mathcal{S}^q_{(1,\,2,\,\ldots,\,s)}} \Theta^\sigma(P^s z)\mathrm{d}\Gamma^\sigma$ where $\mathrm{d}\Gamma^\sigma := \mathrm{d}z^{\sigma(1)} \wedge \mathrm{d}z^{\sigma(2)} \wedge \cdots \wedge \mathrm{d}z^{\sigma(q)}$, $\mathcal{S}^q_{(1,\,2,\,\ldots,\,s)}$ denotes the set of length-q strictly increasing sub-sequences of $(1, 2, \ldots, s)$ and Θ^σ are smooth functions. Since by hypothesis $\mathrm{d}\Upsilon \neq 0$, there exists at least one $\sigma_0 \in \mathcal{S}^q_{(1,\,2,\,\ldots,\,s)}$ with non-vanishing $\Theta^{\sigma_0}(P^s z)$. Then, from (13) we have that

$$\begin{aligned} 0 =& \mathrm{d}\left(\sum_{i=s+1}^{n} \bar{L}_u^{ji} z^i + \bar{\Phi}_u^j(P^s z) - \sum_{i=1}^{n} \alpha^{ji} z^i\right) \wedge \sum_{\sigma \in \mathcal{S}^q_{(1,\,2,\,\ldots,\,s)}} \Theta^\sigma(P^s z)\mathrm{d}\Gamma^\sigma \\ =& \sum_{\sigma \in \mathcal{S}^q_{(1,\,2,\,\ldots,\,s)}} \Theta^\sigma(P^s z)\mathrm{d}\left(\sum_{i=s+1}^{n} (\bar{L}_u^{ji} - \alpha^{ji}) z^i + \bar{\Phi}_u^j(P^s z) - \sum_{i=1}^{s} \alpha^{ji} z^i\right) \wedge \mathrm{d}\Gamma^\sigma \end{aligned}$$

and after multiplication by $\mathrm{d}z^{\hat{\sigma}_0(1)} \wedge \mathrm{d}z^{\hat{\sigma}_0(2)} \wedge \cdots \wedge \mathrm{d}z^{\hat{\sigma}_0(q)}$, where $\hat{\sigma}_0$ is the length-$(s-q)$ sub-sequence of $(1, 2, \ldots, s)$ whose elements are the ones in $\{1, 2, \ldots, s\}$ and not in the range $\{\sigma_0(1), \sigma_0(2), \ldots, \sigma_0(q)\}$ of σ_0, we obtain

$$\begin{aligned} 0 &= \Theta^{\sigma_0}(P^s z)\mathrm{d}\left(\sum_{i=s+1}^{n} (\bar{L}_u^{ji} - \alpha^{ji}) z^i + \bar{\Phi}_u^j(P^s z) - \sum_{i=1}^{s} \alpha^{ji} z^i\right) \wedge \mathrm{d}\Gamma \\ &= \Theta^{\sigma_0}(P^s z)\mathrm{d}\left(\sum_{i=s+1}^{n} (\bar{L}_u^{ji} - \alpha^{ji}) z^i\right) \wedge \mathrm{d}\Gamma \\ &= \Theta^{\sigma_0}(P^s z) \sum_{i=s+1}^{n} (\bar{L}_u^{ji} - \alpha^{ji}) \mathrm{d}z^i \wedge \mathrm{d}\Gamma. \end{aligned}$$

Since $\Theta^{\sigma_0}(P^s z) \neq 0$ and $\{z^j \mid j = 1, 2, \ldots, n\}$ is a system of coordinates, by Lemma 3.1, $\bar{L}_u^{ji} - \alpha^{ji} = 0$. Hence, coming back to (13), we may conclude that

$$\mathrm{d}\left(\bar{\Phi}_u^j(P^s z) - \sum_{i=1}^{s} \alpha^{ji} z^i\right) \wedge \mathrm{d}\Upsilon = 0,$$

i.e., for some function Φ_u, we have

$$\bar{\Phi}_u^j(P^s z) - \sum_{i=1}^{s} \alpha^{ji} z^i = \Phi_u^j(y), \qquad \text{with } y = [y^1\, y^2\, \ldots\, y^q]^\top.$$

Therefore we may rewrite $F_u^j(z) = \sum_{i=s+1}^{n} \bar{L}_u^{ji} z^i + \bar{\Phi}_u^j(P^s z)$ as

$$\begin{aligned} F_u^j(z) &= \sum_{i=s+1}^{n} \bar{L}_u^{ji} z^i + \sum_{i=1}^{s} \alpha^{ji} z^i - \sum_{i=1}^{s} \alpha^{ji} z^i + \bar{\Phi}_u^j(P^s z) \\ &= \sum_{i=s+1}^{n} \bar{L}_u^{ji} z^i + \sum_{i=1}^{s} \alpha^{ji} z^i + \Phi_u^j(y) \end{aligned}$$

and so, setting $L_u^{ji} = \begin{cases} \alpha^{ji} & \text{if } i = 1,\, 2,\, \ldots,\, s \\ \bar{L}_u^{ji} & \text{if } i = s+1,\, s+2,\, \ldots,\, n \end{cases}$ we arrive to the system

$$\begin{aligned} \dot{z} &= L_u z + \Phi_u(y); \\ 0 &= A_u P^s z + (P^s z)^\top B_u y + C_u y + D_u \end{aligned}$$

in the form of system (3).

4. Algorithms for finding the s-tuple vector fields

In the previous sections we saw that the necessary and sufficient conditions to cast the original system (4) in the desired target form (3) boil down to find a suitable s-tuple of vector fields. This section addresses the problem of finding these vector fields. We start to propose an algorithm that uses similar ideas described in [10], but extended for the multiple outputs case and the implicit form. The algorithm however only applies to a particular class of systems. We end the section with an algorithm for the general case, which is not so elegant as the first one, but this does not necessarily mean more computations.

Before we present the algorithms we first have to derive the following result, which generalizes for the multi-output case the results in [10].

Lemma 4.1. *Let system* (8) *be observable in the rank sense at x_0 with $s < n$. Denote $\{\mathrm{d}g\} := \{\mathrm{d}g^j \mid j = 1,\, 2,\, \ldots,\, s\}$ and $\mathcal{L}_u := \mathcal{L}_{f_u}$ for every $u \in \mathbb{R}^m$. Then, it is possible to construct a length-k_0 sequence of subsets $\mathcal{S}_1,\, \mathcal{S}_2,\, \ldots,\, \mathcal{S}_{k_0}$ such that*

- *$\mathcal{S}_1 \subset \{1,\, 2,\, \ldots,\, s\} \times \mathbb{R}^m$ and $\mathcal{S}_k \subset \mathcal{S}_{k-1} \times \mathbb{R}^m$, for $k = 2,\, 3,\, \ldots,\, k_0$;*
- *for $1 \le k_1 \le k_0$ the family*

$$\mathcal{B}_{k_1} := \{\mathrm{d}g\} \cup \{\mathrm{d}\mathcal{L}_{u_k}\mathcal{L}_{u_{k-1}} \ldots \mathcal{L}_{u_1} g^i \mid k = 1,\, 2,\, \ldots,\, k_1 \text{ and } (i,\, u_1,\, u_2,\, \ldots,\, u_k) \in \mathcal{S}_k\}$$

  is a basis for the $C^\infty(U)$-module

$$\operatorname{span}_{C^\infty(U)} \left(\{\mathrm{d}g\} \cup \left\{ \mathrm{d}\mathcal{L}_{u_k}\mathcal{L}_{u_{k-1}} \ldots \mathcal{L}_{u_1} g^i \;\middle|\; \begin{array}{l} k = 1,\, 2,\, \ldots,\, k_1 \\ i = 1,\, 2,\, \ldots,\, s \\ u_1,\, u_2,\, \ldots,\, u_k \in \mathbb{R}^m \end{array} \right\} \right)$$

  in a suitable neighborhood U of x_0;
- *$\mathcal{B}_{k_0}$ is a basis for $\mathrm{d}\mathcal{O}$, where $\mathcal{O}$ is the observable space.*

Proof. From the family $\mathcal{F}_1 := \{\mathrm{d}g\} \cup \{\mathrm{d}\mathcal{L}_{u_1} g^i \mid i = 1,\, 2,\, \ldots,\, s \text{ and } u_1 \in \mathbb{R}^m\}$ we can set a subset $\mathcal{S}_1 \subset \{1,\, 2,\, \ldots,\, s\} \times \mathbb{R}^m$ such that $\{\mathrm{d}g\} \cup \{\mathrm{d}\mathcal{L}_u g^i \mid (i,\, u) \in \mathcal{S}_1\}$ is a basis for $\operatorname{span}_{C^\infty(U)} \mathcal{F}_1$ in a neighborhood U of x_0.

If $s + \#\mathcal{S}_1 = n$, where $\#\mathcal{S}_1$ is the number of elements of $\mathcal{S}_1$, the lemma holds with $k_0 = 1$. Otherwise, we construct one more set $\mathcal{S}_2$ as follows. First we notice that for any vector field Y and any $j \in \{1,\, 2,\, \ldots,\, s\}$ we can write $\mathrm{d}\mathcal{L}_Y g^j = \sum_{(i,\, u) \in \mathcal{S}_1} \alpha_{(i,\, u)}(Y,\, j) \mathrm{d}\mathcal{L}_u g^i + \sum_{k=1}^{s} a_k(Y,\, j) \mathrm{d}g^k$ in the neighborhood U for some $\alpha_{(i,\, u)}(Y,\, j)(\cdot),\, a_k(Y,\, j)(\cdot) \in C^\infty(U)$.

Now, given two vector fields V and W, we find the following:

$$
\begin{aligned}
&\mathrm{d}\mathcal{L}_V\mathcal{L}_W g^j\\
&= \mathrm{d}\iota_V\mathrm{d}\mathcal{L}_W g^j = \mathrm{d}\iota_V\left(\sum_{(i,\,u)\in\mathcal{S}_1}\alpha_{(i,\,u)}(W,\,j)\mathrm{d}\mathcal{L}_u g^i + \sum_{k=1}^{s} a_k(W,\,j)\mathrm{d}g^k\right)\\
&= \mathrm{d}\left(\sum_{(i,\,u)\in\mathcal{S}_1}\alpha_{(i,\,u)}(W,\,j)\iota_V\mathrm{d}\mathcal{L}_u g^i + \sum_{k=1}^{s} a_k(W,\,j)\iota_V\mathrm{d}g^k\right)\\
&= \sum_{(i,\,u)\in\mathcal{S}_1}(\iota_V\mathrm{d}\mathcal{L}_u g^i)\mathrm{d}\alpha_{(i,\,u)}(W,\,j) + \alpha_{(i,\,u)}(W,\,j)\mathrm{d}\iota_V\mathrm{d}\mathcal{L}_u g^i\\
&\quad + \sum_{k=1}^{s}(\iota_V\mathrm{d}g^k)\mathrm{d}a_k(W,\,j) + \sum_{k=1}^{s} a_k(W,\,j)\mathrm{d}\iota_V\mathrm{d}g^k\\
&= \sum_{(i,\,u)\in\mathcal{S}_1}\iota_V(\mathrm{d}\mathcal{L}_u g^i\wedge\mathrm{d}\alpha_{(i,\,u)}(W,\,j)) + \iota_V\mathrm{d}\alpha_{(i,\,u)}(W,\,j)\mathrm{d}\mathcal{L}_u g^i + \alpha_{(i,\,u)}(W,\,j)\mathrm{d}\mathcal{L}_V\mathcal{L}_u g^i\\
&\quad + \sum_{k=1}^{s}\iota_V(\mathrm{d}g^k\wedge\mathrm{d}a_k(W,\,j)) + \iota_V\mathrm{d}a_k(W,\,j)\mathrm{d}g^k + \sum_{k=1}^{s} a_k(W,\,j)\mathrm{d}\mathcal{L}_V g^k
\end{aligned}
$$

and since $\sum_{(i,\,u)\in\mathcal{S}_1}\iota_V(\mathrm{d}\mathcal{L}_u g^i\wedge\mathrm{d}\alpha_{(i,\,u)}(W,\,j))+\sum_{k=1}^{s}\iota_V(\mathrm{d}g^k\wedge\mathrm{d}a_k(W,\,j)) = -\iota_V\mathrm{d}\mathrm{d}\mathcal{L}_W g^j = 0$ and $\mathrm{d}\mathcal{L}_V g^k = \sum_{(r,\,v)\in\mathcal{S}_1}\alpha_{(r,\,v)}(V,\,k)\mathrm{d}\mathcal{L}_v g^r + \sum_{l=1}^{s} a_l(V,\,k)\mathrm{d}g^l$ we have

$$
\begin{aligned}
\mathrm{d}\mathcal{L}_V\mathcal{L}_W g^j = &\sum_{(i,\,u)\in\mathcal{S}_1}\mathcal{L}_V\alpha_{(i,\,u)}(W,\,j)\mathrm{d}\mathcal{L}_u g^i + \alpha_{(i,\,u)}(W,\,j)\mathrm{d}\mathcal{L}_V\mathcal{L}_u g^i + \sum_{k=1}^{s}\mathcal{L}_V a_k(W,\,j)\mathrm{d}g^k\\
&+ \sum_{k=1}^{s} a_k(W,\,j)\left(\sum_{(r,\,v)\in\mathcal{S}_1}\alpha_{(r,\,v)}(V,\,k)\mathrm{d}\mathcal{L}_v g^r + \sum_{l=1}^{s} a_l(V,\,k)\mathrm{d}g^l\right)
\end{aligned}
$$

and so we may conclude that

$$
\mathrm{d}\mathcal{L}_V\mathcal{L}_W(g^j) \in \operatorname{span}_{C^\infty(U)}\left(\{\mathrm{d}g\}\cup\{\mathrm{d}\mathcal{L}_u g^i \mid (i,\,u)\in\mathcal{S}_1\}\cup\{\mathrm{d}\mathcal{L}_V\mathcal{L}_u g^i \mid (i,\,u)\in\mathcal{S}_1\}\right).
$$

Hence, we can set a subset $\mathcal{S}_2\subset\mathcal{S}_1\times\mathbb{R}^m$ such that

$$
\{\mathrm{d}g\}\cup\{\mathrm{d}\mathcal{L}_u g^i \mid (i,\,u)\in\mathcal{S}_1\}\cup\{\mathrm{d}\mathcal{L}_v\mathcal{L}_u g^i \mid (i,\,u,\,v)\in\mathcal{S}_2\}
$$

is a basis for

$$
\operatorname{span}_{C^\infty(U)}\left(\{\mathrm{d}g\}\cup\{\mathrm{d}\mathcal{L}_u g^i,\,\mathrm{d}\mathcal{L}_v\mathcal{L}_u g^i \mid i=1,\,2,\,\ldots,\,s \text{ and } u,\,v\in\mathbb{R}^m\}\right).
$$

Again if $s+\#\mathcal{S}_1+\#\mathcal{S}_2=n$, the theorem holds with $k_0=2$; otherwise we construct recursively, proceeding analogously as above, a sequence of sets $\mathcal{S}_3,\,\mathcal{S}_4,\,\ldots,\,\mathcal{S}_{\bar{k}}$, satisfying the first two conditions of the lemma until we reach the condition $s+\sum_{i=1}^{\bar{k}}\#\mathcal{S}_i=n$. At this point the third condition is also satisfied and the lemma holds with $k_0=\bar{k}$.

Finally, notice that the sequence $\mathcal{S}_1,\,\mathcal{S}_2,\,\ldots,\,\mathcal{S}_{k_0}$ is not unique and that the recursive procedure is finite because necessarily $k_0\le n-s$. □

At this point we are ready to describe the algorithms to find the coordinate transformation. We first consider that the number of s-tuple vector fields required is strictly less than the state-space dimension, i.e., $s<n$.

4.1. **Algorithm I (a particular case).** Resorting to the ideas presented in [10] that only considers the one-dimensional output, we now present for the multi-output case an algorithm to find the s-tuple of vector fields for a class of systems for which there exists a special sequence among those described in Lemma 4.1.

Proposition 4.2. *Let system* (8) *be observable in the rank sense at* x_0 *with* $s < n$. *Denote* $\{\mathrm{d}g\} := \{\mathrm{d}g^j \mid j = 1, 2, \ldots, s\}$ *and* $\mathcal{L}_u := \mathcal{L}_{f_u}$ *for every* $u \in \mathbb{R}^m$. *Suppose that we can construct a length-*k_0 *sequence* $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{k_0}$ *satisfying the properties in Lemma 4.1 and such that there exist* $v_1, v_2, \ldots, v_{k_0} \in \mathbb{R}^m$ *with*

$$(i, v_1, v_2, \ldots, v_{k_0}) \in \mathcal{S}_{k_0} \text{ for all } i \in \{1, 2, \ldots, s\}.$$

Then, for each $i \in \{1, 2, \ldots, s\}$, *define the vector field* Y_i *by*

$$\begin{aligned}
\mathcal{L}_{Y_i} g^j &= 0 \text{ for all } j \in \{1, 2, \ldots, s\} \\
\mathcal{L}_{Y_i}\mathcal{L}_{u_k}\mathcal{L}_{u_{k-1}}\ldots\mathcal{L}_{u_1} g^j &= 0 \text{ for all } k = 1, 2, \ldots, k_0 - 1 \text{ and } (j, u_1, \ldots, u_{k-1}, u_k) \in \mathcal{S}_k \\
\mathcal{L}_{Y_i}\mathcal{L}_{u_{k_0}}\mathcal{L}_{u_{k_0-1}}\ldots\mathcal{L}_{u_1} g^j &= 0 \text{ for all } (j, u_1, \ldots, u_{k_0-1}, u_{k_0}) \in \mathcal{S}_{k_0} \setminus \{(i, v_1, v_2, \ldots, v_{k_0})\} \\
\mathcal{L}_{Y_i}\mathcal{L}_{v_{k_0}}\mathcal{L}_{v_{k_0-1}}\ldots\mathcal{L}_{v_1} g^i &= 1.
\end{aligned}$$

Finally, for X_i *set the iterated Lie bracket*

$$X_i := (-1)^{k_0}[f_{v_1}, [f_{v_2}, [\ldots [f_{v_{k_0-1}}, [f_{v_{k_0}}, Y_i]] \ldots]]].$$

Then, system (4) *can be rewritten in the target form* (3) *iff the* s*-tuple of vector fields* $X = (X_1, X_2, \ldots, X_s)$ *satisfy the conditions in Theorem 2.1.*

Proof. Following a similar reasoning as in [10] we can prove that for a system in target form in coordinates z the proposed algorithm yields constant vector fields $X_i = \partial/\partial z^i + \sum_{j=s+1}^{n} C_i^j \partial/\partial z^j$ for each $i = 1, 2, \ldots, s$. Then, we can easily find the change of coordinates $z^i = w^i$ for $i = 1, 2, \ldots, s$ and $z^j = w^j + \sum_{i=1}^{s} C_i^j w^i$ for $j = s+1, s+2, \ldots, n$. Note that, by (5), we find $\partial/\partial z^j = \partial/\partial w^j$ for $j = s+1, s+2, \ldots, n$ and $\partial/\partial z^i = \partial/\partial w^i + \sum_{j=s+1}^{n} C_i^j \partial/\partial w^j$ for $i = 1, 2, \ldots, s$. We can conclude that (3a), $F_u(z) = \sum_{j=1}^{n} (\sum_{i=1}^{n} L_u^{ji} z^i + \Phi_u^j(y))\, \partial/\partial z^j$, reads $F_u(w) = \sum_{j=1}^{n} (\sum_{i=1}^{n} L_u^{ji} M w + \Phi_u^j(y))\, N[\partial/\partial w^1\ \partial/\partial w^2\ \ldots\ \partial/\partial w^n]^\top$, in new coordinates, for suitable matrices M and N, i.e., $F_u(w) = A_u w + \tilde{\Phi}(y)$ for a suitable matrix A_u and a suitable function $\tilde{\Phi}$. In other words, the system is still in target form when re-written in the new coordinates $(w^1, w^2, \ldots, w^n)$. Moreover we have the identity $X_i = \partial/\partial w^i$ and, from the proof of necessity in Theorem 2.1, we have that the s-tuple X satisfies the conditions in that theorem in the coordinates $(w^1, w^2, \ldots, w^n)$. Since the conditions are intrinsic they are also satisfied by the same s-tuple in the coordinates $(z^1, z^2, \ldots, z^n)$. This means that for a system in target form, the s-tuple defined in the proposition satisfies the conditions in Theorem 2.1. Since the s-tuple is defined in an intrinsic way, it must also satisfy that conditions in any coordinates. □

Remark 4.1. This idea to construct the sequence $\mathcal{S}$ and the vector fields Y_i; $i = 1, \ldots, s$ is borrowed from [10]. We only have to perform some small adjustments to the multi-output case. In fact those adjustments imply that for a system in target form we have

$$\mathcal{L}_{(-1)^{k_0} X_i} g^j = (-1)^{k_0} \mathcal{L}_{Y_i}\mathcal{L}_{v_{k_0}}\mathcal{L}_{v_{k_0-1}}\ldots\mathcal{L}_{v_2}\mathcal{L}_{v_1} g^j,$$

from which we can conclude that $\mathcal{L}_{X_i} g^j = (-1)^{2k_0}\delta_i^j = \delta_i^j$. To check this we may make use of the identity (7) and the definitions of Y_i and $\mathcal{S}$.

Remark 4.2. The difficulty or perhaps impossibility to construct an analogous algorithm to the one in [10] for the multi-output case may be related to some important differences between this case and the single-output case referred in [8, 9]. One of these differences is that in the multi-output case there are, in general, nonlinear changes of coordinates that preserve the target form.

4.2. **Algorithm II (general case).** Unfortunately the sequence stated in Proposition 4.2 will exist only for a very particular class of systems. For the more general case we suggest that, if no such sequence exists, to follow directly the conditions presented in the Theorem 2.1. We may compare this with the algorithm in [7].

Even if the following algorithm does not look as elegant as the one in Proposition 4.2, in fact it may lead to easier computations. The real advantage of the particular algorithm in Proposition 4.2 is that it gives exactly one s-tuple of vector fields while, the general algorithm that follows will give the entire family of the suitable s-tuples.

We start by noticing that, for the system (3), in target form, we can see from (10) and (11) that, for each $k \in \mathbb{N}_0$:

$$\begin{aligned}\Omega_k^{\tilde{X}} &= \operatorname{span}_{\mathbb{R}}\{\pi^j M_{u_1 u_2 \dots u_k} \mathrm{d}z \wedge \mathrm{d}\Gamma \mid j = 1, 2, \dots, s \text{ and } u_1, u_2, \dots, u_k \in \mathbb{R}^m\} \\ &= \operatorname{span}_{\mathbb{R}} \left\{ (\mathrm{d}\mathcal{L}_{u_l} \dots \mathcal{L}_{u_2}\mathcal{L}_{u_1} z^j \wedge \mathrm{d}\Gamma)|_{z_0} \;\middle|\; \begin{array}{l} 1 \le l \le k,\ j = 1, 2, \dots, s, \\ u_1, u_2, \dots, u_l \in \mathbb{R}^m \end{array} \right\}\end{aligned}$$

(recall that we suppose the system is smooth in a neighborhood of z_0, in particular $z \to \Phi_u(y)$ is supposed to be smooth). Then given a sequence $\mathcal{S}$ as in Lemma 4.1 we find, for each $k \in \{1, 2, \dots, k_0\}$:

$$\begin{aligned} &\operatorname{span}_{\mathbb{R}}\{\pi^j M_{u_1 u_2 \dots u_k} \mathrm{d}z \wedge \mathrm{d}\Gamma \mid j = 1, 2, \dots, s \text{ and } u_1, u_2, \dots, u_k \in \mathbb{R}^m\} \\ &= \operatorname{span}_{\mathbb{R}} \left\{ (\mathrm{d}\mathcal{L}_{u_l} \dots \mathcal{L}_{u_2}\mathcal{L}_{u_1} z^j \wedge \mathrm{d}\Gamma)|_{z_0} \;\middle|\; \begin{array}{l} 1 \le l \le k,\ j = 1, 2, \dots, s, \\ (j, u_1, u_2, \dots, u_l) \in \mathcal{S}_l \end{array} \right\} \\ &= \operatorname{span}_{\mathbb{R}} \left\{ \pi^j M_{u_1 u_2 \dots u_l} \mathrm{d}z \wedge \mathrm{d}\Gamma \;\middle|\; \begin{array}{l} 1 \le l \le k,\ j = 1, 2, \dots, s, \\ (j, u_1, u_2, \dots, u_l) \in \mathcal{S}_l \end{array} \right\}. \end{aligned} \tag{14}$$

Consider system (4) and suppose that the auxiliary system (8) is observable in the rank sense. To find a s-tuple of vector fields satisfying the conditions of Theorem 2.1 we may proceed as follows: first of all, for a s-tuple of vector fields $X = (X_1, X_2, \dots, X_s)$, define recursively the following $(s+1)$-forms

$$\mathcal{I}^X_{(r,\, u_1)} := \mathrm{d}\mathcal{L}_{u_1} g^r \wedge \mathrm{d}\Gamma$$

for all $r \in \{1, 2, \dots, s\}$ and $u_1 \in \mathbb{R}^m$,

$$\mathcal{I}^X_{(r,\, u_1,\, u_2,\, \dots,\, u_{k-1},\, u_k)} := (\mathcal{L}_{u_k} \iota_X \mathcal{I}^X_{(r,\, u_1,\, u_2,\, \dots,\, u_{k-1})}) \wedge \mathrm{d}\Gamma$$

for all $r \in \{1, 2, \dots, s\}$, $k \ge 2$ and $u_1, u_2, \dots, u_k \in \mathbb{R}^m$.

Fix a sequence $\mathcal{S}$ as in Lemma 4.1; we look for a s-tuple $X = (X_1, X_2, \dots, X_s)$ of vector fields, with $X_i = \sum_{j=1}^n X_i^j \partial/\partial x^j$ solving, step by step, the following equations

1. $\mathrm{d}g^j(X_i) = \delta_i^j$ for all $i, j \in \{1, 2, \dots, s\}$;
2. successively for $1 \le k \le k_0$:
   a. for all $j \in \{1, 2, \dots, s\}$ and $v_1, v_2, \dots, v_k \in \mathbb{R}^m$,
   $$\mathcal{I}^X_{(j,\, v_1,\, v_2,\, \dots,\, v_k)} \in \operatorname{span}_{\mathbb{R}} \left\{ \mathcal{I}^X_{(r,\, u_1,\, u_2,\, \dots,\, u_l)} \;\middle|\; \begin{array}{l} l \in \{1, 2, \dots, k\}, \\ (r, u_1, u_2, \dots, u_l) \in \mathcal{S}_l \end{array} \right\};$$
   b. for all $(r, u_1, u_2, \dots, u_k) \in \mathcal{S}_k$, $\mathrm{d}\iota_X \mathcal{I}^X_{(r,\, u_1,\, u_2,\, \dots,\, u_k)} = 0$;

3. for all $(j,\, u_1,\, u_2,\, \ldots,\, u_{k_0}) \in \mathcal{S}_{k_0}$ and $v \in \mathbb{R}^m$,
$\mathcal{I}^X_{(j,\, u_1,\, u_2,\, \ldots,\, u_{k_0},\, v)} \in \operatorname{span}_{\mathbb{R}} \left\{ \mathcal{I}^X_{(r,\, u_1,\, u_2,\, \ldots,\, u_l)} \;\middle|\; \begin{array}{l} l = 1,\, 2,\, \ldots,\, k_0, \\ (r,\, u_1,\, u_2,\, \ldots,\, u_l) \in \mathcal{S}_l \end{array} \right\}$;
4. for all $j \in \{1,\, 2,\, \ldots,\, s\}$, $k \in \{1,\, 2,\, \ldots,\, k_0\}$ and $(r,\, u_1,\, u_2,\, \ldots,\, u_k) \in \mathcal{S}_k$, both $(\mathrm{d}\iota_{f_u} \mathrm{d}g^j) \wedge \mathrm{d}\Upsilon$ and $(\mathrm{d}\iota_{f_u} \iota_X \mathcal{I}^X_{(r,\, u_1,\, u_2,\, \ldots,\, u_k)}) \wedge \mathrm{d}\Upsilon$ are elements of the real vector space spanned by

$$\{\mathrm{d}g^i \wedge \mathrm{d}\Upsilon \mid i = 1,\, 2,\, \ldots,\, s\} \cup \left\{ (\iota_X \mathcal{I}^X_{(r,\, u_1,\, u_2,\, \ldots,\, u_k)}) \wedge \mathrm{d}\Upsilon \;\middle|\; \begin{array}{l} k \in \{1,\, 2,\, \ldots,\, k_0\}, \\ (r,\, u_1,\, u_2,\, \ldots,\, u_k) \in \mathcal{S}_k \end{array} \right\}.$$

It now follows from the way the s-tuple X is defined that it will satisfy the conditions of Theorem 2.1 iff it results from the algorithm. Note that the first steps consist in deriving the properties that the vector fields X_i, or their coordinates X_i^j, must satisfy. The last step is perhaps the more cumbersome one but, it may be omitted in the case of explicit output notation because it follows from the preceding ones.

As we see the solution for X is not unique but, as soon as we have derived the properties of each X_i we only have to chose a s-tuple satisfying them. Section 5.3 describes an example that illustrates this algorithm.

4.3. **The case $s = n$.** In this case, from the fact that the n-tuple must satisfy the equations $\mathrm{d}g^j(X_i) = \delta_i^j$, for $i,\, j = 1,\, 2,\, \ldots,\, n$, with $\mathrm{d}\Gamma \neq 0$, it follows that there is only one candidate to satisfy the conditions of Theorem 2.1. Let system (8) be observable in the rank sense. Then, the system (4) can be rewritten in target form iff the n-tuple of vector fields $X = (X_1,\, X_2,\, \ldots,\, X_n)$ satisfying the equations $\mathrm{d}g^j(X_i) = \delta_i^j$, for $i,\, j = 1,\, 2,\, \ldots,\, n$, satisfy also the condition $\mathsf{b.v}$.

Remark 4.3. Notice that the uniqueness of the n-tuple of vector fields $X = X(g)$ follows from the fact that we suppose that the candidates for new suitable coordinate functions g^i for $i \in \{1,\, 2,\, \ldots,\, n\}$ had already been chosen. A different choice of suitable candidates $\tilde{g}^i$ for $i \in \{1,\, 2,\, \ldots,\, n\}$ may lead to a different n-tuple $X(\tilde{g})$.

5. Examples

This section illustrates the main results of the paper. Three examples are proposed. The first one shows a simple dynamical system that do not fit in the framework proposed in [7–10], but it is still possible to construct an observer with linear error dynamics. The other two examples elucidate the details of the algorithm proposed for a particular class of systems and for the general case, respectively.

5.1. **A perspective output system.** Consider the following academic example with state $x = [x^1\ x^2]^\top$, single input $u = u(t) \in \mathbb{R}$, and where the measured output y is the ratio ${x^2}/{x^1}$ that also acts as an external input on the system

$$\dot{x} = F_u(x) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ y+u \end{bmatrix}; \qquad y = {x^2}/{x^1}, \quad x^1 \neq 0. \tag{15}$$

We notice that, the output equation $y = {x^2}/{x^1}$ can be written as $x^1 y = x^2$ or, as

$$\begin{bmatrix} 0 & -1 \end{bmatrix} x + x^\top \begin{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix} y = 0. \tag{16}$$

Since $x = P^2 x$ we conclude that (16) is in the form of equation (3b), which means that system (15) can be re-written in the form of system (3). Now, we show that it cannot be written in the form $\dot{z} = L_u z + \Phi_u(y)$ with the explicit form output $y = z^1$. It is known, from [7] or from Theorem 2.1, that a necessary condition for that is that the conditions a. and b.i.-b.iv. in Theorem 2.1 hold with $\mathrm{d}\Gamma = \mathrm{d}y$ and $g = y$. From $\mathrm{d}(x^1 y) = \mathrm{d}x^2$ we find that $x^1\mathrm{d}y = \mathrm{d}x^2 - y\mathrm{d}x^1$, i.e.,

$$\mathrm{d}y = \frac{1}{(x^1)^2}(x^1\mathrm{d}x^2 - x^2\mathrm{d}x^1).$$

Thus a. holds. Now we prove that b.ii. does not hold. We have that $\Omega_1^X \subseteq \Omega^X$ and, from its definition, see (9), we may find Ω_1^X: To this end, we have that $\mathrm{d}y(F_u) = \frac{1}{(x^1)^2}(x^1(-x_1 + y + u) - x^2x^2) = -1 - y^2 + \frac{y+u}{x^1}$, so $\mathrm{d}\mathcal{L}_u y \wedge \mathrm{d}y = -\frac{y+u}{(x^1)^2}\mathrm{d}x^1 \wedge \mathrm{d}y$, and from $\mathrm{d}x^1 \wedge \mathrm{d}y = \frac{1}{x^1}\mathrm{d}x^1 \wedge \mathrm{d}x^2$, we have $\mathrm{d}\mathcal{L}_u y \wedge \mathrm{d}y = -\frac{y+u}{(x^1)^3}\mathrm{d}x^1 \wedge \mathrm{d}x^2$ and

$$\Omega_1^X = \operatorname{span}_{\mathbb{R}}\left\{\frac{y+u}{(x^1)^3}\mathrm{d}x^1 \wedge \mathrm{d}x^2 \mid u \in \mathbb{R}\right\} = \operatorname{span}_{\mathbb{R}}\left\{\frac{y}{(x^1)^3}\mathrm{d}x^1 \wedge \mathrm{d}x^2,\ \frac{y+1}{(x^1)^3}\mathrm{d}x^1 \wedge \mathrm{d}x^2\right\}.$$

Therefore Ω_1^X has real dimension equal to 2 and for b.ii. to hold the dimension should be 1.

5.2. **Illustration of the algorithm for a specific class of systems.** In this section we illustrate the algorithm described in Proposition 4.2.

Consider the system

$$\dot{x} = \begin{bmatrix}\dot{x}^1\\ \dot{x}^2\\ \dot{x}^3\\ \dot{x}^4\end{bmatrix} = f_u(x) = \begin{bmatrix} x^1(u + x^3)\\ -x^1(u + x^3) + x^1x^4 + 1 + {}^{x^2}\!/_{x^1}\\ -x^3(u + x^3) + 1 + {}^{(x^2+u)}\!/_{x^1}\\ -x^4(u + x^3) + x^3 \end{bmatrix}, \quad x^1 y = x^1 + x^2.$$

where $x = [x^1\ x^2\ x^3\ x^4]^\top$ is the state space, $u \in \mathbb{R}$ is the input and $y = y^1$ is the output. Let $x_0 = [1\ 0\ 0\ 0]^\top$. Our aim is to write this system in the form (3) in a neighborhood of this point.

The first step is to choose the candidates for the first two new coordinates. From the output equation, consider $g^1(x) = x^1$ and $g^2 = x^1 + x^2$. With this choice we see that for $g := [g^1\ g^2]^\top$ the output equation can be written as

$$\begin{bmatrix}0 & -1\end{bmatrix} g(x) + g(x)^\top \left[\begin{bmatrix}1\\0\end{bmatrix}\right] y = 0,$$

which is in the form of equation (4b). In this case, $\mathrm{d}\Gamma := \mathrm{d}g^1 \wedge \mathrm{d}g^2 = \mathrm{d}x^1 \wedge \mathrm{d}x^2 \neq 0$ and $\mathrm{d}\Upsilon = \mathrm{d}y = \frac{1}{x^1}(\mathrm{d}x^2 + (1 - y)\mathrm{d}x^1)$. We also have $\mathrm{d}y \wedge \mathrm{d}\Gamma = 0$. The next step is to check if the auxiliary system

$$\dot{x} = \begin{bmatrix} x^1(u + x^3)\\ -x^1(u + x^3) + x^1x^4 + 1 + {}^{x^2}\!/_{x^1}\\ -x^3(u + x^3) + 1 + {}^{(x^2+u)}\!/_{x^1}\\ -x^4(u + x^3) + x^3 \end{bmatrix}, \qquad \begin{aligned}\bar{y}^1 &= g^1 = x^1,\\ \bar{y}^2 &= g^2 = x^1 + x^2\end{aligned}$$

is observable in the rank sense. For that we notice that

$$S := \{g^1,\ g^2,\ \mathcal{L}_0 g^1,\ \mathcal{L}_0 g^2\} = \{x^1,\ x^1 + x^2,\ x^1x^3,\ x^1x^4 + 1 + {}^{x^2}\!/_{x^1}\} \tag{17}$$

is a subset of the observable space $\mathcal{O}$ and after some straightforward computations we find that $\mathrm{d}\mathcal{L}_0 g^1 = x^3\mathrm{d}x^1 + x^1\mathrm{d}x^3$, $\mathrm{d}\mathcal{L}_0 g^2 = (x^4 - {}^{x^2}\!/_{(x^1)^2})\mathrm{d}x^1 + {}^{1}\!/_{x^1}\mathrm{d}x^2 + x^1\mathrm{d}x^4$ and $\mathrm{d}S|_{x_0} := \{\mathrm{d}x^1,\ \mathrm{d}x^1 + \mathrm{d}x^2,\ \mathrm{d}x^3,\ \mathrm{d}x^2 + \mathrm{d}x^4\}$, which has rank 4. Therefore, the auxiliary system is observable in the rank sense.

Notice also that the length-1 sequence $\mathcal{S}_1 = \{(1, 0),\ (2,\ 0)\} \subset \{1,\ 2\} \times \mathbb{R}$ satisfies the condition in Proposition 4.2. Thus, we are now ready to follow the algorithm described in that Proposition. Let us now compute the pair of vector fields $Y = (Y_1,\ Y_2)$ with

$Y_1 = \sum_{j=1}^4 Y_1^j \partial/\partial x^j$ and $Y_2 = \sum_{j=1}^4 Y_2^j \partial/\partial x^j$. Following the algorithm, for the vector field Y_1 we find the equations

$$
\begin{aligned}
&Y_1^1 = 0; && Y_1^1 + Y_1^2 = 0\\
&x^3Y_1^1 + x^1Y_1^3 = 1; && (x^4 - x^2/(x^1)^2)Y_1^1 + 1/x^1 Y_1^2 + x^1Y_1^4 = 0
\end{aligned}
$$

from which we derive $Y_1 = 1/x_1 \partial/\partial x^3$. For Y_2 we obtain

$$
\begin{aligned}
&Y_2^1 = 0; && Y_2^1 + Y_2^2 = 0\\
&x^3Y_2^1 + x^1Y_2^3 = 0; && (x^4 - x^2/(x^1)^2)Y_2^1 + 1/x^1 Y_2^2 + x^1Y_2^4 = 1,
\end{aligned}
$$

that is, $Y_2 = 1/x_1 \partial/\partial x^4$.

Next we compute the pair $X = (X_1,\, X_2)$ with $X_i = -[f_0,\, Y_i] = [Y_i,\, f_0]$. In coordinates this may be computed by the formula (6): From $f_0(x) = x^1x^3\partial/\partial x^1 + (x^1x^4 - x^1x^3 + 1 + x^2/x^1)\partial/\partial x^2 + (1 - (x^3)^2 + x^2/x^1)\partial/\partial x^3 + (x^3 - x^3x^4)\partial/\partial x^4$ we obtain

$$
\begin{aligned}
&X_1^1 = 1/x^1(x^1) = 1, && X_1^2 = 1/x^1(-x^1) = -1\\
&X_1^3 = 1/x^1(-2x^3) - [x^1x^3(-1/(x^1)^2)] = -x^3/x^1, && X_1^4 = 1/x^1(1 - x^4) = (1-x^4)/x^1
\end{aligned}
$$

and

$$
\begin{aligned}
&X_2^1 = 0, && X_2^2 = 1/x^1(x^1) = 1\\
&X_2^3 = 0, && X_1^4 = 1/x^1(-x^3) - [x^1x^3(-1/(x^1)^2)] = 0.
\end{aligned}
$$

Thus, $X_1 = \partial/\partial x^1 - \partial/\partial x^2 - \frac{x^3}{x^1}\partial/\partial x^3 + \frac{1-x^4}{x^1}\partial/\partial x^4$ and $X_2 = \partial/\partial x^2$.

To obtain the change of coordinates we have to compute $\iota_X \mathcal{I}^X_{(1,\,0)}$ and $\iota_X \mathcal{I}^X_{(2,\,0)}$. From $\mathcal{I}^X_{(1,\,0)} = x^1 \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3$, $\mathcal{I}^X_{(2,\,0)} = x^1 \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^4$, $\det \begin{bmatrix} 1 & 0 & V^1 \\ -1 & 1 & V^2 \\ -x^3/x^1 & 0 & V^3 \end{bmatrix} = x^3/x^1 V^1 + V^3$ and $\det \begin{bmatrix} 1 & 0 & V^1 \\ -1 & 1 & V^2 \\ (1-x^4)/x^1 & 0 & V^4 \end{bmatrix} = (x^4-1)/x^1 V^1 + V^4$ we obtain $\iota_X(\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3) = x^3/x^1 \mathrm{d}x^1 + \mathrm{d}x^3$ and $\iota_X(\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^4) = (x^4-1)/x^1 \mathrm{d}x^1 + \mathrm{d}x^4$. Therefore we obtain the following

$$\iota_X \mathcal{I}^X_{(1,\,0)} = x^3\mathrm{d}x^1 + x^1\mathrm{d}x^3 = \mathrm{d}(x^1x^3) \text{ and } \iota_X \mathcal{I}^X_{(2,\,0)} = (x^4 - 1)\mathrm{d}x^1 + x^1\mathrm{d}x^4 = \mathrm{d}(x^1x^4) - \mathrm{d}x^1.$$

We can now construct a coordinate transformation $(z^1,\, z^2,\, z^3,\, z^4)$ with $z^1 = g^1 = x^1$, $z^2 = g^2 = x^1 + x^2$ and z^3 and z^4 such that $\mathrm{d}z^3$ and $\mathrm{d}z^4$ are in the real vector space spanned by the 1-forms in $\{\mathrm{d}z^1,\, \mathrm{d}z^2,\, \iota_X \mathcal{I}^X_{(1,0)},\, \iota_X \mathcal{I}^X_{(2,0)}\}$ and $\mathrm{d}z^1 \wedge \mathrm{d}z^2 \wedge \mathrm{d}z^3 \wedge \mathrm{d}z^4 \neq 0$. We may set for example $z^3 = x^1x^3$ and $z^4 = x^1x^4$. We can now confirm that applying this coordinate transformation, the original system can be rewritten in the target form. A quick check of this fact can be done, using (5), by noticing that

$$
\begin{aligned}
&\partial/\partial x^1 = \partial/\partial z^1 + \partial/\partial z^2 + \frac{z^3}{z^1}\partial/\partial z^3 + \frac{z^4}{z^1}\partial/\partial z^4, && \partial/\partial x^2 = \partial/\partial z^2\\
&\partial/\partial x^3 = z^1\partial/\partial z^3, && \partial/\partial x^4 = z^1\partial/\partial z^4.
\end{aligned}
$$

Thus, the vector field

$$
\begin{aligned}
f_u = {}& x^1(u + x^3)\partial/\partial x^1 + (-x^1(u + x^3) + x^1x^4 + 1 + x^2/x^1)\partial/\partial x^2\\
&+ (-x^3(u + x^3) + 1 + (x^2+u)/x^1)\partial/\partial x^3 + (-x^4(u + x^3) + x^3)\partial/\partial x^4
\end{aligned}
$$

can be re-written as

$$\begin{aligned} f_u &= x^1(u+x^3)\partial/\partial z^1 + (x^1(u+x^3) - x^1(u+x^3) + x^1x^4 + 1 + x^2/x^1)\partial/\partial z^2 \\ &\quad + \left(x^1(u+x^3)\frac{z^3}{z^1} + (-x^3(u+x^3) + 1 + (x^2+u)/x^1)z^1\right)\partial/\partial z^3 \\ &\quad + \left(x^1(u+x^3)\frac{z^4}{z^1} + (-x^4(u+x^3) + x^3)z^1\right)\partial/\partial z^4 \\ &= (z^3 + uz^1)\partial/\partial z^1 + (z^4 + (x^2+x^1)/x^1)\partial/\partial z^2 \\ &\quad + ((u+x^3)z^3 - (u+x^3)x^3z^1 + z^1 + z^1(x^2+u)/x^1)\partial/\partial z^3 \\ &\quad + ((u+x^3)z^4 - x^4(u+x^3)z^1 + x^3z^1)\partial/\partial z^4 \\ &= (z^3 + uz^1)\partial/\partial z^1 + (z^4 + y)\partial/\partial z^2 + (z^2 + u)\partial/\partial z^3 + z^3\partial/\partial z^4 \end{aligned}$$

and therefore

$$\dot{z} = \begin{bmatrix} z^3 + uz^1 \\ z^4 \\ z^2 \\ z^3 \end{bmatrix} + \begin{bmatrix} 0 \\ y \\ u \\ 0 \end{bmatrix} = \begin{bmatrix} u & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} z + \begin{bmatrix} 0 \\ y \\ u \\ 0 \end{bmatrix}, \qquad z^1 y = z^2,$$

which is a system in the target form (3).

5.3. **Illustration of the general algorithm.** Consider the following system with state $x = [x^1\ x^2\ x^3]^\top$, output $y = [y^1\ y^2]^\top$ and input $u \in \mathbb{R}$,

$$\dot{x} = \begin{bmatrix} \dot{x}^1 \\ \dot{x}^2 \\ \dot{x}^3 \end{bmatrix} = f_u(x) = \begin{bmatrix} e^{-x^1}(x^3+u) \\ x^3 + (x^1 + x^2 - e^{x^1})^2 - e^{-x^1}(x^3+u) \\ x^3 + \sin(x^1 + x^2 - e^{x^1}) \end{bmatrix}, \qquad \begin{aligned} y^1 &= x^1 + x^2 - e^{x^1}, \\ y^2 &= e^{-x^1}(x^1 + x^2) \end{aligned}$$

where the goal is to write this system in the form (3) in a neighborhood of the point $x_0 = [0\ 0\ 0]^\top$. From the output equation select $g^1(x) = e^{x^1}$ and $g^2(x) = x^1 + x^2$ as candidates for the first two coordinates. With this choice we see that for $g := [g^1\ g^2]^\top$ we can write the output equation as

$$\begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} g(x) + g(x)^\top \mathrm{col}_{\mathcal{M}}\left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) y + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} y = 0,$$

and therefore the output can be written in the form (4b). In this case, $\mathrm{d}\Gamma := \mathrm{d}g^1 \wedge \mathrm{d}g^2 = e^{x^1}\mathrm{d}x^1 \wedge \mathrm{d}x^2 \neq 0$, $\mathrm{d}y^1 = (1 - e^{x^1})\mathrm{d}x^1 + \mathrm{d}x^2$, $\mathrm{d}y^2 = e^{-x^1}(1 - x^1 - x^2)\mathrm{d}x^1 + e^{-x^1}\mathrm{d}x^2$ and $\mathrm{d}\Upsilon = e^{-x^1}y^1\mathrm{d}x^1 \wedge \mathrm{d}x^2$. Moreover $\mathrm{d}y^1 \wedge \mathrm{d}\Gamma = 0 = \mathrm{d}y^2 \wedge \mathrm{d}\Gamma$. The next step consist in checking the observability of the auxiliary system

$$\dot{x} = \begin{bmatrix} e^{-x^1}(x^3+u) \\ x^3 + (x^1 + x^2 - e^{x^1})^2 - e^{-x^1}(x^3+u) \\ x^3 + \sin(x^1 + x^2 - e^{x^1}) \end{bmatrix}, \qquad \begin{aligned} \bar{y}^1 &= g^1 = e^{x^1}, \\ \bar{y}^2 &= g^2 = x^1 + x^2 \end{aligned}$$

in the rank sense. To this end, we see that

$$S := \{g^1,\, g^2,\, \mathcal{L}_0 g^1\} \tag{18}$$

is a subset of the observable space $\mathcal{O}$ and after some straightforward computations we find that $\mathcal{L}_u g^1 = x^3 + u$ and $\mathrm{d}S|_{x_0} := \{\mathrm{d}x^1,\, \mathrm{d}x^1 + \mathrm{d}x^2,\, \mathrm{d}x^3\}$, which has rank 3. Therefore, the observability holds.

The length-1 sequence $\mathcal{S}_1 = \{(1,0)\} \subset \{1,2\} \times \mathbb{R}$ satisfy the conditions of Lemma 4.1. Let $X = (X_1,\, X_2)$ be a pair of vector fields with

$$X_1 := X_1^1\partial/\partial x^1 + X_1^2\partial/\partial x^2 + X_1^3\partial/\partial x^3 \text{ and } X_2 := X_1^2\partial/\partial x^1 + X_2^2\partial/\partial x^2 + X_2^3\partial/\partial x^3.$$

We can now go through the algorithm proposed in Section 4.2: From the conditions $\mathrm{d}g^i(X_j) = \delta^i_j$ on Step 1 of the algorithm we derive the identities

$$X_1^1 = \mathrm{e}^{-x^1}, \quad X_1^2 = -\mathrm{e}^{-x^1}, \quad X_2^1 = 0 \quad \text{and} \quad X_2^2 = 1.$$

Therefore the vector fields are of the form

$$X_1 = \mathrm{e}^{-x^1}\partial/\partial x^1 - \mathrm{e}^{-x^1}\partial/\partial x^2 + X_1^3\partial/\partial x^3 \quad \text{and} \quad X_2 = \partial/\partial x^2 + X_2^3\partial/\partial x^3. \tag{19}$$

Now, we move to Step 2.a. From $\mathcal{L}_u g^1 = x^3 + u$ and $\mathcal{L}_u g^2 = x^3 + (y^1)^2$ we have that

$$\mathcal{I}^X_{(1,\,u)} = \mathcal{I}^X_{(2,\,u)} = \mathcal{I}^X_{(1,\,0)} = \mathrm{d}x^3 \wedge \mathrm{d}\Gamma.$$

Thus, the inclusion $\{\mathcal{I}^X_{(1,\,u)}, \mathcal{I}^X_{(2,\,u)}\} \subseteq \mathrm{span}_{\mathbb{R}}\{\mathcal{I}^X_{(1,\,0)}\}$ is trivial and we see that this step of the algorithm do not provide additional information about the pair of vector fields X (for the present example). From this step we must keep the expression found for $\mathcal{I}^X_{(1,\,0)}$ since it will be needed in the following steps.

In Step 2.b we use the identity $\mathrm{d}\iota_X\mathcal{I}^X_{(1,\,0)} = 0$. From

$$\begin{aligned}\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3(X_1,\, X_2,\, V) &= \det\begin{bmatrix} \mathrm{e}^{-x^1} & 0 & V^1 \\ -\mathrm{e}^{-x^1} & 1 & V^2 \\ X_1^3 & X_2^3 & V^3 \end{bmatrix} \\ &= (-X_1^3 - X_2^3\mathrm{e}^{-x^1})V^1 - X_2^3\mathrm{e}^{-x^1}V^2 + \mathrm{e}^{-x^1}V^3\end{aligned}$$

we obtain

$$\begin{aligned}\iota_X\mathcal{I}^X_{(1,\,0)} &= \iota_X(\mathrm{e}^{x^1}\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3) = \mathrm{e}^{x^1}((-X_1^3 - X_2^3\mathrm{e}^{-x^1})\mathrm{d}x^1 - X_2^3\mathrm{e}^{-x^1}\mathrm{d}x^2 + \mathrm{e}^{-x^1}\mathrm{d}x^3) \\ &= (-X_1^3\mathrm{e}^{x^1} - X_2^3)\mathrm{d}x^1 - X_2^3\mathrm{d}x^2 + \mathrm{d}x^3 \end{aligned} \tag{20}$$

and

$$\mathrm{d}\iota_X\mathcal{I}^X_{(1,\,0)} = -\mathrm{e}^{x^1}\mathrm{d}X_1^3 \wedge \mathrm{d}x^1 - \mathrm{d}X_2^3 \wedge \mathrm{d}x^1 - \mathrm{d}X_2^3 \wedge \mathrm{d}x^2.$$

Thus, re-writing for $j = 1,\, 2$, $\mathrm{d}X_j^3 = \frac{\partial X_j^3}{\partial x^1}\mathrm{d}x^1 + \frac{\partial X_j^3}{\partial x^2}\mathrm{d}x^2 + \frac{\partial X_j^3}{\partial x^3}\mathrm{d}x^3$ and taking into account that the 2-forms $\mathrm{d}x^1 \wedge \mathrm{d}x^2$, $\mathrm{d}x^1 \wedge \mathrm{d}x^3$ and $\mathrm{d}x^2 \wedge \mathrm{d}x^3$ are linearly independent in $\Lambda^2(\Omega)$, from $\mathrm{d}\iota_X\mathcal{I}^X_{(1,\,0)} = 0$ we obtain $\frac{\partial X_2^3}{\partial x^3} = 0$, $\mathrm{e}^{x^1}\frac{\partial X_1^3}{\partial x^3} + \frac{\partial X_2^3}{\partial x^3} = 0$ and $\mathrm{e}^{x^1}\frac{\partial X_1^3}{\partial x^2} + \frac{\partial X_2^3}{\partial x^2} - \frac{\partial X_2^3}{\partial x^1} = 0$, i.e.,

$$\frac{\partial X_2^3}{\partial x^3} = \frac{\partial X_1^3}{\partial x^3} = 0 \quad \text{and} \quad \mathrm{e}^{x^1}\frac{\partial X_1^3}{\partial x^2} + \frac{\partial X_2^3}{\partial x^2} = \frac{\partial X_2^3}{\partial x^1}. \tag{21}$$

For the next step in the algorithm we need to compute $\mathcal{I}^X_{(1,\,0,\,v)} = (\mathcal{L}_v\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Gamma$ for $v \in \mathbb{R}$. First we find

$$\begin{aligned}\mathcal{L}_v\iota_X\mathcal{I}^X_{(1,\,0)} &= \mathrm{d}\iota_{f_v}((-X_1^3\mathrm{e}^{x^1} - X_2^3)\mathrm{d}x^1 - X_2^3\mathrm{d}x^2 + \mathrm{d}x^3) \\ &= \mathrm{d}((-X_1^3\mathrm{e}^{x^1} - X_2^3)(x^3 + v)\mathrm{e}^{-x^1} - X_2^3(x^3 + (y^1)^2 - (x^3 + v)\mathrm{e}^{-x^1}) + x^3 + \sin(y^1)) \\ &= \mathrm{d}((x^3 + v)(-X_1^3 - X_2^3\mathrm{e}^{-x^1}) - X_2^3(x^3 + (y^1)^2) + X_2^3(x^3 + v)\mathrm{e}^{-x^1} + x^3 + \sin(y^1)) \\ &= \mathrm{d}(x^3(1 - X_1^3 - X_2^3) - vX_1^3 - (y^1)^2X_2^3 + \sin(y^1)). \end{aligned} \tag{22}$$

Then using the first identity in (21) we obtain

$$\mathcal{I}^X_{(1,\,0,\,v)} = (1 - X_1^3 - X_2^3)\mathrm{e}^{x^1}\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3.$$

From Step 3 of the algorithm it follows that $1 - X_1^3 - X_2^3$ must be a constant, in other words

$$\mathrm{d}(X_1^3 + X_2^3) = 0. \tag{23}$$

Finally we move to the last step of the algorithm. At this point we have the identities

$$\mathrm{d}g^1 \wedge \mathrm{d}\Upsilon = 0 = \mathrm{d}g^2 \wedge \mathrm{d}\Upsilon$$

and, from (20),

$$\iota_X \mathcal{I}^X_{(1,\,0)} \wedge \mathrm{d}\Upsilon = ((-X^3_1 \mathrm{e}^{x^1} - X^3_2)\mathrm{d}x^1 - X^3_2 \mathrm{d}x^2 + \mathrm{d}x^3) \wedge \mathrm{d}\Upsilon = \mathrm{d}x^3 \wedge \mathrm{d}\Upsilon.$$

Hence from the Step 4 we must have that for all $v \in \mathbb{R}$ the 3-forms

$$(\mathrm{d}\iota_{f_v}\mathrm{d}g^1) \wedge \mathrm{d}\Upsilon,\; (\mathrm{d}\iota_{f_v}\mathrm{d}g^2) \wedge \mathrm{d}\Upsilon \text{ and } (\mathrm{d}\iota_{f_v}\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Upsilon$$

are all in the real vector space $\mathbb{E}$ spanned by $\{\mathrm{d}x^3 \wedge \mathrm{d}\Upsilon\}$. We compute

$$\begin{aligned}
(\mathrm{d}\iota_{f_v}\mathrm{d}g^1) \wedge \mathrm{d}\Upsilon &= \mathrm{d}(x^3+u) \wedge \mathrm{d}\Upsilon = \mathrm{d}x^3 \wedge \mathrm{d}\Upsilon \in \mathbb{E}\\
(\mathrm{d}\iota_{f_v}\mathrm{d}g^2) \wedge \mathrm{d}\Upsilon &= \mathrm{d}(x^3+(y^1)^2) \wedge \mathrm{d}\Upsilon = \mathrm{d}x^3 \wedge \mathrm{d}\Upsilon \in \mathbb{E}\\
(\mathrm{d}\iota_{f_v}\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Upsilon &= (\mathcal{L}_v\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Upsilon
\end{aligned}$$

From (22) and (23) we have $(\mathcal{L}_v\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Upsilon = (1 - X^3_1 - X^3_2)\mathrm{d}x^3 \wedge \mathrm{d}\Upsilon - v\mathrm{d}X^3_1 \wedge \mathrm{d}\Upsilon - (y^1)^2\mathrm{d}X^3_2 \wedge \mathrm{d}\Upsilon$. Then from the first identity in (21) we have that $(\mathcal{L}_v\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Upsilon = (1-X^3_1-X^3_2)\mathrm{d}x^3 \wedge \mathrm{d}\Upsilon$ and from (23) we derive that $(\mathcal{L}_v\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Upsilon \in \mathbb{E}$. At this point, from the last step we do not obtain any additional information. Therefore, all the pairs of vector fields $X = (X_1,\, X_2)$ resulting from the proposed algorithm are now completely characterized by the conditions (19), (23) and (21), i.e.,

$$\begin{aligned}
X_1 &= \mathrm{e}^{-x^1}\partial/\partial x^1 - \mathrm{e}^{-x^1}\partial/\partial x^2 + X^3_1\partial/\partial x^3;\\
X_2 &= \partial/\partial x^2 + X^3_2\partial/\partial x^3;
\end{aligned}
\qquad
\begin{aligned}
&\tfrac{\partial X^3_2}{\partial x^3} = \tfrac{\partial X^3_1}{\partial x^3} = 0;\\
&(\mathrm{e}^{x^1}-1)\tfrac{\partial X^3_1}{\partial x^2} = \tfrac{\partial X^3_2}{\partial x^1};\\
&\mathrm{d}(X^3_1 + X^3_2) = 0.
\end{aligned}$$

We can now select a pair of vector fields satisfying the conditions above, for example the pair $X = (X_1, X_2)$ with $X_1 = \mathrm{e}^{-x^1}\partial/\partial x^1 - \mathrm{e}^{-x^1}\partial/\partial x^2$ and $X_2 = \partial/\partial x^2$. Notice that for this choice, from (20), we have $\iota_X\mathcal{I}^X_{(1,\,0)} = \mathrm{d}x^3$. This means that the change of coordinates

$$(z^1,\, z^2,\, z^3) := (\mathrm{e}^{x^1},\, x^1 + x^2,\, x^3)$$

should transform the original system to the target form. Indeed, using (5), we obtain

$$\partial/\partial x^1 = \mathrm{e}^{x^1}\partial/\partial z^1 + \partial/\partial z^2; \quad \partial/\partial x^2 = \partial/\partial z^2; \quad \partial/\partial x^3 = \partial/\partial z^3$$

and

$$\begin{aligned}
f_u &= \mathrm{e}^{-x^1}(x^3+u)\partial/\partial x^1 + (x^3 + (x^1+x^2-\mathrm{e}^{x^1})^2 - \mathrm{e}^{-x^1}(x^3+u))\partial/\partial x^2\\
&\quad + (x^3 + \sin(x^1+x^2-\mathrm{e}^{x^1}))\partial/\partial x^3\\
&= (x^3+u)\partial/\partial z^1 + (\mathrm{e}^{-x^1}(x^3+u) + x^3 + (x^1+x^2-\mathrm{e}^{x^1})^2 - \mathrm{e}^{-x^1}(x^3+u))\partial/\partial z^2\\
&\quad + (x^3 + \sin(x^1+x^2-\mathrm{e}^{x^1}))\partial/\partial z^3\\
&= (x^3+u)\partial/\partial z^1 + (x^3 + (x^1+x^2-\mathrm{e}^{x^1})^2)\partial/\partial z^2 + (x^3 + \sin(x^1+x^2-\mathrm{e}^{x^1}))\partial/\partial z^3\\
&= (z^3+u)\partial/\partial z^1 + (z^3 + (y^1)^2)\partial/\partial z^2 + (z^3 + \sin(y^1))\partial/\partial z^3.
\end{aligned}$$

Thus, we arrive to the system

$$\begin{aligned}
\dot z &= \begin{bmatrix} z^3+u \\ z^3+(y^1)^2 \\ z^3+\sin(y^1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} z + \begin{bmatrix} u \\ (y^1)^2 \\ \sin(y^1) \end{bmatrix}\\
0 &= \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} P^2 z + (P^2 z)^\top \mathrm{col}_{\mathcal{M}}\left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) y + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} y,
\end{aligned}$$

that is in the desired target form, with $P^2 z = [z^1\, z^2]^\top$.

6. Concluding remarks

In this paper we have presented the necessary and sufficient conditions for time-varying linearization up to multi-output injection with implicit outputs. It is important to stress that *i)* the conditions obtained encompass the ones for linear systems, *ii)* the class of systems that admit a coordinate transformation is larger since it includes all the systems that can be transformed to a linear system up to output injection, and *iii)* for any system written in the target form (3), there exists an observer (Kalman-like) that exhibits in the new coordinate system linear error dynamics. To obtain the coordinate transformation we have proposed two algorithms: one that is more restrictive in the original class of systems and another for the general case. Several examples illustrated the proposed procedure.

Appendix

Proof of Lemma 3.1. Let the functions $\psi_k \in C^\infty(\Omega)$ solve $\sum_{k=r+1}^{n} \psi_k \mathrm{d}h^k \wedge \mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^r = 0$. For each $j \in \{r+1,\, r+2,\, \ldots,\, n\}$; after multiplication by $\mathrm{d}h^{r+1} \wedge \mathrm{d}h^{r+2} \wedge \cdots \wedge \mathrm{d}h^{j-1} \wedge \mathrm{d}h^{j+1} \wedge \cdots \wedge \mathrm{d}h^n$ we obtain $0 = (-1)^{j-1}\psi_j \mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^n$ and since $(h^1,\, h^2,\, \ldots,\, h^n)$ is a system of coordinates, necessarily $\psi_j = 0$. □

Proof of Lemma 3.2. First, if $r < n$, we may set $\{h^j \mid j = r+1,\, r+2,\, \ldots,\, n\}$ such that $\{h^j \mid j = 1,\, 2,\, \ldots,\, n\}$ forms a smooth system of coordinates in a neighborhood $U_1 \subseteq U$ of x_0. In other words $x = [x^1\; x^2\; \ldots\; x^n]^\top \mapsto h(x) = [h^1(x)\; h^2(x)\; \ldots\; h^n(x)]^\top$ is a diffeomorphism from U_1 onto $h(U_1)$. Then we can write $h^0(x) = h^0(h^{-1}(h(x)))$ and for $\phi := h^0 \circ h^{-1}$ we have

$$\mathrm{d}h^0 = \mathrm{d}(h^0 \circ h^{-1} \circ h) = \sum_{j=1}^{n} {}^{\partial\phi}\!/\!_{\partial h^j} \mathrm{d}h^j; \quad \mathrm{d}h^j = \sum_{i=1}^{n} {}^{\partial h^j}\!/\!_{\partial x^i} \mathrm{d}x^i$$

and, by the hypothesis, we derive that

$$0 = \sum_{j=1}^{n} {}^{\partial\phi}\!/\!_{\partial h^j} \mathrm{d}h^j \wedge \mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^r = \sum_{j=r+1}^{n} {}^{\partial\phi}\!/\!_{\partial h^j} \mathrm{d}h^j \wedge \mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^r.$$

Since $\{h^j \mid j = 1,\, 2,\, \ldots,\, n\}$ forms a smooth system of coordinates in U_1 we have, by the Lemma 3.1, that ${}^{\partial\phi}\!/\!_{\partial h^j} = 0$ for every $j = r+1,\, r+2,\, \ldots,\, n$, i.e., ϕ does not depend on the coordinate functions $\{h^j \mid j = r+1,\, r+2,\, \ldots,\, n\}$. In other words we may write $h^0(x) = \phi(h^1(x),\, h^2(x),\, \ldots,\, h^r(x))$.

Finally if $r = n$ and $\mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^n \neq 0$ in U, $\{h^j \mid j = 1,\, \ldots,\, n\}$ is a system of coordinates in U and for $h(x) = (h^1(x),\, h^2(x),\, \ldots,\, h^n(x))$ it follows that $h^0(x) = h^0 \circ h^{-1}(h(x))$. □

Proof of Lemma 3.3. We can write $\mathrm{d}h^0 = \sum_{i=1}^{r} \frac{\partial\phi}{\partial h^i} \mathrm{d}h^i$, from which we derive that $\mathrm{d}h^0 \wedge \mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \cdots \wedge \mathrm{d}h^r = 0$. □

Proof of Lemma 3.4. We shall prove both identities by induction on k. Denote $F_u(z) := L_u z + \Phi_u(y) = \bar{L}_u Q^s z + \bar{\Phi}_u(P^s z)$ with $\bar{\Phi}_u(P^s z) := \tilde{L}_u P^s z + \Phi_u(y)$. Note that $\bar{\Phi}_u$ can be written in this way because from $\mathrm{d}y^j \wedge \mathrm{d}\Gamma = 0$ we have that y^j is a function of $P^s z$; see Lemma 3.2. Denote also $\mathcal{L}_u := \mathcal{L}_{F_u}$ where $F_u = [F_u^1\; F_u^2\; \ldots\; F_u^n]^\top$ when seen as a vector field must be understood as $F_u = \sum_{j=1}^n F_u^j \partial/\partial x^j$. From $\mathcal{L}_{u_1} z^j = \mathrm{d}z^j(\bar{L}_{u_1} Q^s z + \bar{\Phi}_{u_1}(P^s z))$ we obtain $\mathcal{L}_{u_1} z^j = \pi^j \bar{L}_{u_1} Q^s z + \pi^j \bar{\Phi}_{u_1}(P^s z)$, where π^j is the row matrix $\pi^j = [\delta_j^1\ \delta_j^2\ \ldots\ \delta_j^n]$. Thus, from Lemma 3.3, $\mathrm{d}\mathcal{L}_{u_1} z^j \wedge \mathrm{d}\Gamma = \pi^j \bar{L}_{u_1} Q^s \mathrm{d}z \wedge \mathrm{d}\Gamma = \pi^j M_{u_1} \mathrm{d}z \wedge \mathrm{d}\Gamma$ and

$$\Omega_1^{\tilde{X}} = \mathrm{span}_{\mathbb{R}}\{\pi^j M_{u_1} \mathrm{d}z \wedge \mathrm{d}\Gamma \mid j = 1,\, 2,\, \ldots,\, s \text{ and } u_1 \in \mathbb{R}^m\},$$

i.e., (10) is satisfied for $k = 1$. Suppose now it is satisfied for a given $k \geq 1$. For given $\pi^j M_{u_1 u_2 \ldots u_k} \mathrm{d}z \wedge \mathrm{d}\Gamma \in \Omega_k^{\tilde{X}}$, taking into account that $M_{u_1 u_2 \ldots u_k} \mathrm{d}z$ is independent of $\mathrm{d}z^i$ for all $i = 1,\, 2,\, \ldots,\, s$, and that $\iota_{\tilde{X}} \mathrm{d}\Gamma = 1$ we have

$$\iota_{\tilde{X}}(\pi^j M_{u_1 u_2 \ldots u_k} \mathrm{d}z \wedge \mathrm{d}\Gamma) = (-1)^s \pi^j M_{u_1 u_2 \ldots u_k} \mathrm{d}z$$

and $\mathcal{L}_{u_{k+1}}((-1)^s \pi^j M_{u_1 u_2 \ldots u_k} \mathrm{d}z) = (-1)^s \pi^j M_{u_1 u_2 \ldots u_k}(\bar{L}_{u_{k+1}} Q^s \mathrm{d}z + \mathrm{d}\bar{\Phi}_{u_{k+1}})$, so

$$\mathcal{L}_{u_{k+1}} \iota_{\tilde{X}}(\pi^j M_{u_1 u_2 \ldots u_k} \mathrm{d}z \wedge \mathrm{d}\Gamma) = (-1)^s \pi^j M_{u_1 u_2 \ldots u_k u_{k+1}} \mathrm{d}z \wedge \mathrm{d}\Gamma.$$

By induction, we can conclude that (10) holds for all $k \in \mathbb{N}_0$.

Now, the identity $\mathcal{L}_{u_1} z^j = \pi^j \bar{L}_{u_1} Q^s z + \pi^j \bar{\Phi}_{u_1}(P^s z)$ shows that (11) holds for $k = 1$. If it holds for a given $k \geq 1$ then

$$\begin{aligned}
&\mathcal{L}_{u_{k+1}} \mathcal{L}_{u_k} \ldots \mathcal{L}_{u_2} \mathcal{L}_{u_1} z^j \\
&= \pi^j M_{u_1 u_2 \ldots u_k} F_{u_{k+1}} + \mathrm{d}a^k(F_{u_{k+1}}) \\
&\quad + \sum_{\substack{r \in \{1,2,\ldots,k-1\} \\ \alpha_i \in \mathcal{S}_{(u_2,\, u_3,\, \ldots,\, u_k)} \\ \alpha_i \cap \alpha_j = \emptyset \\ \{i_1, i_2, \ldots, i_r\} \subseteq \{1,2,\ldots,s\}}} (\mathrm{d}a^k_{\alpha_1 \alpha_2 \ldots \alpha_r}(F_{u_{k+1}}))(\pi^{i_1} M_{\alpha_1} z)(\pi^{i_2} M_{\alpha_2} z) \ldots (\pi^{i_r} M_{\alpha_r} z) \\
&\quad + \sum_{\substack{r \in \{1,2,\ldots,k-1\} \\ \alpha_i \in \mathcal{S}_{(u_2,\, u_3,\, \ldots,\, u_k)} \\ \alpha_i \cap \alpha_j = \emptyset \\ \{i_1, i_2, \ldots, i_r\} \subseteq \{1,2,\ldots,s\} \\ l \in \{1,2,\ldots,r\}}} (a^k_{\alpha_1 \alpha_2 \ldots \alpha_r})(\pi^{i_1} M_{\alpha_1} z) \ldots (\pi^{i_{l-1}} M_{\alpha_{l-1}} z)(\pi^{i_l} M_{\alpha_l} F_{u_{k+1}})(\pi^{i_{l+1}} M_{\alpha_{l+1}} z) \ldots (\pi^{i_r} M_{\alpha_r} z).
\end{aligned}$$

Now, from the fact that the functions a^k, $a^k_{\alpha_1 \alpha_2 \ldots \alpha_r}$ depend only on $P^s z$ and $u_1,\, u_2,\, \ldots,\, u_k$, using the identities

$$\begin{aligned}
\pi^j M_{u_1 u_2 \ldots u_k} F_{u_{k+1}} &= \pi^j M_{u_1 u_2 \ldots u_k u_{k+1}} z + \pi^j M_{u_1 u_2 \ldots u_k} \bar{\Phi}_{u_{k+1}}(P^s z) \\
\mathrm{d}a(F_{u_{k+1}}) &= \sum_{i=1}^n \partial a/\partial z^i \pi^i M_{u_{k+1}} z + \mathrm{d}a(\bar{\Phi}_{u_{k+1}}(P^s z)) \\
\pi^{i_l} M_{\alpha_l} F_{u_{k+1}} &= \pi^{i_l} M_{\tilde{\alpha}_l} z + \pi^{i_l} M_{\alpha_l} \bar{\Phi}_{u_{k+1}}(P^s z)
\end{aligned}$$

where $a \in C^\infty(\Omega)$ is any smooth function and $\tilde{\alpha}_l \in \mathcal{S}_{(u_1, u_2, \ldots, u_k, u_{k+1})}$ is the sequence $\tilde{\alpha}_l := (u_{\sigma(1)},\, u_{\sigma(2)},\, \ldots,\, u_{\sigma(l_n)},\, u_{k+1})$ if $\alpha_l \in \mathcal{S}_{(u_1, u_2, \ldots, u_k)}$ is $\alpha_l = (u_{\sigma(1)},\, u_{\sigma(2)},\, \ldots,\, u_{\sigma(l_n)})$, we can write

$$\begin{aligned}
&\mathcal{L}_{u_{k+1}} \mathcal{L}_{u_k} \ldots \mathcal{L}_{u_2} \mathcal{L}_{u_1} z^j \\
&= \pi^j M_{u_1 u_2 \ldots u_k u_{k+1}} z + a^{k+1} + \sum_{\substack{r \in \{1,2,\ldots,k\} \\ \alpha_i \in \mathcal{S}_{(u_2,\, u_3,\, \ldots,\, u_k,\, u_{k+1})} \\ \alpha_i \cap \alpha_j = \emptyset \\ \{i_1, i_2, \ldots, i_r\} \subseteq \{1,2,\ldots,s\}}} (a^{k+1}_{\alpha_1 \alpha_2 \ldots \alpha_r})(\pi^{i_1} M_{\alpha_1} z)(\pi^{i_2} M_{\alpha_2} z) \ldots (\pi^{i_r} M_{\alpha_r} z)
\end{aligned}$$

where a^{k+1} and $a^{k+1}_{\alpha_1 \alpha_2 \ldots \alpha_r}$ depend only on $P^s z$ and $u_1,\, u_2,\, \ldots,\, u_k,\, u_{k+1}$. By induction, (11) holds for all $k \in \mathbb{N}_0$. $\square$

References

[1] A. A. Agrachev and Y. L. Sachkov, *Control theory from the geometric viewpoint*, Encyclopaedia of Mathematical Sciences, no. 87, Springer, 2004.

[2] A. P. Aguiar and J. Hespanha, *Minimum-energy state estimation for systems with perspective outputs*, IEEE Trans. Automat. Control **51** (2006), 226–241.

[3] ______, *Robust filtering for deterministic systems with implicit outputs*, Systems Control Lett. **58** (2009), 263–270.

[4] G. Besançon, *An overview on observer tools for nonlinear systems*, Nonlinear Observers and Applications, Springer, 2007.

[5] M. P. do Carmo, *Differential forms and applications*, Universitext, Springer, 1994.

[6] O.M. Grasselli and A. Isidori, *An existence theorem for observers of bilinear systems*, IEEE Trans. Automat. Control **26** (1981), 1299–1300.

[7] H. Hammouri and J.-P. Gauthier, *Bilinearization up to output injection*, Systems Control Lett. **11** (1988), 139–149.

[8] ______, *Global time-varying linearization up to output injection*, SIAM J. Control Optim. **30** (1992), 1295–1310.

[9] ______, *Time-varying linearization up to output injection*, Proceedings of the 28th CDC, Tampa, Florida, Dec 1989.

[10] H. Hammouri and M. Kinnaert, *A new procedure for time-varying linearization up to output injection*, Systems Control Lett. **28** (1996), 151–157.

[11] V. Jurdjević, *Geometric control theory*, Cambridge Studies in Advanced Mathematics, no. 51, Cambridge University Press, 1997.

[12] A. J. Krener, *On the equivalence of control systems and the linearization of nonlinear systems*, SIAM J. Control **11** (1973), 670–676.

[13] A. J. Krener and A. Isidori, *Linearization by output injection and nonlinear observers*, Systems Control Lett. **3** (1983), 47–52.

[14] A. J. Krener and W. Respondek, *Nonlinear observers with linearizable error dynamics*, SIAM J. Control Optim **23** (1985), 1295–1310.

S. S. Rodrigues is with the Johann Radon Institute for Computational and Applied Mathematics, ÖAW, Linz, Austria. e-mail: sergio.rodrigues@oeaw.ac.at

A. Pedro Aguiar is with the Institute for Systems and Robotics, Instituto Superior Técnico, Lisbon, Portugal. e-mail: pedro@isr.ist.utl.pt.