On the Linearization up to Multi-Output Injection for a Class of Systems With Implicitly Defined Outputs

Sérgio S. Rodrigues and A. 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 an observer with linear error dynamics. We propose a methodology to find the coordinate transformation. Several examples illustrate the proposed procedure.**

***Index Terms*—Nonlinear multi-output systems, time-varying linearization, observer design.**

I. 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 [11], [13]–[16], [18]–[20]. More precisely, given (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 in [18], and Krener and Isidori in [19]. 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 to a linear system up to an output injection: $\dot{z} = Az + \Phi(y),\ y = Cz$, where A and C 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 $\dot{\hat{z}} = A\hat{z} + \Phi(y) + L(y - C\hat{x})$ with L selected so that $A - LC$ is Hurwitz achieves a linear error dynamics $\dot{\tilde{z}} = (A-LC)\tilde{z}$, $\tilde{z} := \hat{z}-z$, where $\tilde{z}(t) \to 0$ as $t \to +\infty$. More recently, using tools from Differential Geometry, Hammouri and Gauthier in [13], [14] and Hammouri and Kinnaert in [16] extended the linearization problem to systems in form (1) to obtain time-varying linear systems up to output injection of the form

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

where $P^q z = [z^1\, z^2\, \ldots\, z^q]^\top \in \mathbb{R}^q$ collects the first q coordinates[1] of $z = [z^1\, z^2\, \ldots\, z^n]^\top \in \mathbb{R}^n$, $q \le n$. In their work, necessary and sufficient conditions to the existence of the desired change of coordinates $z = \Theta(x)$ are provided. Other approaches consist in linearizing a nonlinear system up to output injection by means of input/output transformation, see for example [10], [21].

In this paper, we depart from a different point of view 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 [13], [14], [16], although it is still possible to construct an observer with linear error dynamics, see [3]. A simple example is the following system with a perspective output y and state $x = [x^1\, x^2]^\top$, $x^1 \neq 0$: $\dot{x} = \begin{bmatrix} x^2 \\ -x^1 + y + u \end{bmatrix}$, $\quad y = {}^{x^2}\!/_{x^1}$. It turns out, as we will see later on in Section IV-A, that it is not possible to write it in the target form considered in [13], [14], [16], which is $\dot{z} = A_u z + \Phi_u(y), \quad y = z^1$, for $z = [z^1\, z^2]^\top$, or even other recent suitable target forms that were considered in [5], [6], [12]. However, the system takes the simple form $\dot{x} = Ax + \Phi_u(y), \quad x^1 y = x^2$, for which there exists an observer with linear error dynamics. This last fact can be concluded from the results in [2], [3], where actually this observation holds for more general systems in the form

$$\dot{z} = A_u z + \Phi_u(y) \tag{3a}$$
$$0 = C_u P^s z + (P^s z)^\top D_u y + E_u y + F_u \tag{3b}$$

where A_u, C_u, E_u, and F_u are matrices with real entries, D_u is a column matrix whose entries are matrices D_u^i; $i = 1, \ldots, p$, and $s \le n$. In (3), both the input and output are assumed to be known. Following [3], an optimal observer for (3) is given by

Sérgio S. Rodrigues is with the Johann Radon Institute for Computational and Applied Mathematics (RICAM), ÖAW, Linz, Austria. sergio.rodrigues@oeaw.ac.at

A. Pedro Aguiar is with the Faculty of Engineering, University of Porto (FEUP), Portugal. pedro.aguiar@fe.up.pt

This work was supported by projects CONAV/FCT-PT (PTDC/EEACRO/113820/2009) and MORPH (EU FP7 ICT 288704).

[1] For a given matrix M, $M^\top$ denotes its transpose. Notice that we use superscripts 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.

$\dot{\hat{z}} = A_u\hat{x} + \Phi_u(y) - Q(M_1\hat{z} + M_2)$ where Q satisfies a dynamic Riccati-like equation, and M_1 and M_2 are matrices depending on (y, u). It turns out that the resulting error dynamics $\tilde{z} = z - \hat{z}$ is linear and its convergence depends on some suitable observability conditions, see [3, section 3.1] for the details.

Problem statement: Consider a general nonlinear system of the form

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

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

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. Our goal is to find the necessary and sufficient conditions under which system (4) can be rewritten, up to a change of coordinates, in a system like (3). Note that once in (3), it is possible to construct an observer with linear error dynamics. In (4), we consider that $f_u \colon \mathbb{R}^n \to \mathbb{R}^n$ and $g \colon \mathbb{R}^n \to \mathbb{R}^s$, with $s \leq n$, are smooth (C^∞) functions in Ω, $C_u \in \mathcal{M}_{p\times s}(\mathbb{R})$, $E_u \in \mathcal{M}_{p\times q}(\mathbb{R})$ and $F_u \in \mathcal{M}_{p\times 1}(\mathbb{R})$ are matrices with real entries and $D_u \in \mathcal{M}_{p\times 1}(\mathcal{M}_{s\times q}(\mathbb{R}))$ 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 this 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 $v_1^\top D_u v_2$ is to be understood as $v_1^\top D_u v_2 := [v_1^\top D_u^1 v_2 \quad v_1^\top D_u^2 v_2 \quad \dots \quad v_1^\top D_u^p v_2]^\top$ where $D_u^1, D_u^2, \dots, D_u^p$ are the entries of the column matrix D_u. Thus, the output equation (4b) is an identity in $\mathcal{M}_{1\times p}(\mathbb{R})$ (i.e., in $\mathbb{R}^p$); we consider that it completely defines y in a neighborhood of an interesting point x_0, $y(u, x) = \Psi_u(g(x))$.

The rest of the paper is organized as follows: Section II describes the necessary and sufficient conditions to be able to rewrite 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 III we present an algorithm to find the suitable s-tuple of vector fields. Section IV illustrates the contribution of the paper with two examples. Brief conclusions are discussed in Section V.

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. For a more complete discussion on what follows we suggest the works [1], [7], [17], see also [22].

Given a system of coordinates $(x^1, x^2, \dots, x^n)$, we consider the vector fields in $\mathbb{R}^n$, $\partial/\partial x^k$ 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. We denote by $\mathcal{V}(\Omega)$ the $C^\infty(\Omega)$-module of smooth vector fields in Ω and, by $\Lambda^k(\Omega)$ the $C^\infty(\Omega)$-module of (differential) k-forms, $k \in \mathbb{N}$, defined in the Cartesian product $\mathcal{V}(\Omega)^k$. We denote by $\alpha\wedge\beta$ the wedge product between the forms α and β, and by $\iota_X w$ the interior product $\iota_X w(V_1, V_2, \dots, V_{k-1}) := w(X, V_1, V_2, \dots, V_{k-1})$ between a vector field X and a k-form w; for a r-tuple of vector fields, with $r \leq k$, we define recursively $\iota_{(X_1, X_2, \dots, X_r)} w := \iota_{X_r}\iota_{(X_1, X_2, \dots, X_{r-1})} w$. The exterior derivative of a k-form w will be denoted by $\mathrm{d}w$. The tuple $(h^1, h^2, \dots, h^n)$ of smooth functions is a system of coordinates in $\Omega \subseteq \mathbb{R}^n$ if $\mathrm{d}h^1 \wedge \mathrm{d}h^2 \wedge \dots \wedge \mathrm{d}h^n|_x \neq 0$, for all the points $x \in \Omega$. Given two systems of local coordinates $(x^1, x^2, \dots, x^n)$ and $(h^1, h^2, \dots, h^n)$, respectively, we have

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

We denote by $\mathcal{L}$ the Lie derivative operator. For a vector field X, and a k-form w we have: $\mathcal{L}_X w := \iota_X \mathrm{d}w = \mathrm{d}w(X)$ if $k = 0$, and $\mathcal{L}_X w := (\iota_X \mathrm{d} + \mathrm{d}\iota_X)w$ if $k \geq 1$.

We denote $D_u = \mathrm{col}_{\mathcal{M}}(D_u^1, D_u^2, \dots, D_u^p)$, where the subscript $\mathcal{M}$ means that the entries of the matrix D_u, appearing in (3b), are matrices. We may also define the matrix $\widehat{D}_u = \mathrm{col}_{\mathbb{R}}(D_u^1, D_u^2, \dots, D_u^p)$ with real entries. As an illustration, $\mathrm{col}_{\mathbb{R}}(D_u^1, D_u^2)$ and $\mathrm{col}_{\mathcal{M}}(D_u^1, D_u^2)$ stand, respectively, for $\begin{bmatrix} a_{11} & a_{12} \\ b_{11} & b_{12} \end{bmatrix}$ and $\begin{bmatrix} \begin{bmatrix} a_{11} & a_{12} \end{bmatrix} \\ \begin{bmatrix} b_{11} & b_{12} \end{bmatrix} \end{bmatrix}$, for given matrices $D_u^1 = \begin{bmatrix} a_{11} & a_{12} \end{bmatrix}$ and $D_u^2 = \begin{bmatrix} b_{11} & b_{12} \end{bmatrix}$.

II. Main theorem

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 introduce the following auxiliary system

$$\dot{x} = f_u(x), \quad \bar{y} = g(x) \tag{6}$$

where f_u and g are obtained from (4). The auxiliary output $\bar{y}$ will provide the set of candidates to the first new s coordinate functions. We assume that this auxiliary system is observable in the rank sense (see, e.g., [4]) in a neighborhood of a given point x_0, which implies that $\{\mathrm{d}w|_{x_0} \mid w \in \mathcal{O}\}$ is n-dimensional, where $\mathrm{d}w|_{x_0}$ is the evaluation at x_0 of the form $\mathrm{d}w$, and $\mathcal{O}$ the smallest set containing $\{g^1, g^2, \dots, g^s\}$ and closed under all the Lie derivatives $\{\mathcal{L}_{f_u} \mid u \text{ is a constant in } \mathbb{R}^m\}$. Let $X = (X_1, X_2, \dots, X_s)$ be a s-tuple of vector fields. Following a similar idea in [13], [14], define a sequence of vector spaces as follows: set $\mathrm{d}\Gamma := \mathrm{d}g^1 \wedge \mathrm{d}g^2 \wedge \dots \wedge \mathrm{d}g^s$ and denote by Ω_1^X the real vector space

$$\Omega_1^X := \mathrm{span}_{\mathbb{R}}\{\mathrm{d}\mathcal{L}_{f_u} g^j \wedge \mathrm{d}\Gamma \mid 1 \leq j \leq s \text{ and } u \in \mathbb{R}^m\}. \tag{7}$$

Recursively, for $k \geq 2$, define the real vector space $\Omega_k^X := \mathrm{span}_{\mathbb{R}}\{\mathcal{L}_{f_u}(\iota_X w) \wedge \mathrm{d}\Gamma \mid w \in \Omega_{k-1}^X \text{ and } u \in \mathbb{R}^m\}$. Define also the smallest real vector space 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 $\{\mathrm{d}g\} := \{\mathrm{d}g^j \mid j = 1, 2, \dots, s\}$. Finally, denoting the q-form $\mathrm{d}\Upsilon := \mathrm{d}y^1 \wedge \mathrm{d}y^2 \wedge \dots \wedge \mathrm{d}y^q$, where $y(u, x) = [y^1\, y^2 \, \dots \, y^q]^\top$ is the (known) output signal, we can now state the following result.

Theorem 2.1: Suppose that system (6) is observable in the rank sense in a neighborhood U of a given point x_0, $\mathrm{d}\Upsilon \neq 0$, and $\mathrm{d}y^j \wedge \mathrm{d}\Gamma = 0$ for $j = 1, 2, \dots, q$. 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 if, and only if,

a. $\mathrm{d}\Gamma|_{x_0} \neq 0$;
b. there exists a s-tuple of vector fields $X = (X_1, X_2, \dots, X_s)$ 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 $\mathcal{N}$, for all $u\in\mathbb{R}^m$.

Remark 2.1: To have the desired output equation (3b) we select $z^i = g^i$ for $i \in \{1, 2, \ldots, s\}$. For the rest of the coordinate functions, we may choose them such that $\{\mathrm{d}z^j \mid j = 1, 2, \ldots, n\}$ form a basis for $\Omega[X, g]$ and $\mathrm{d}\Gamma \wedge \mathrm{d}z^{s+1} \wedge \mathrm{d}z^{s+2} \wedge \cdots \wedge \mathrm{d}z^n \neq 0$.

Remark 2.2: Comparing the required conditions with those in [13], [14], [16] we notice that we have mainly two new statements: $dy^j \wedge d\Gamma = 0$ and b.v, which do not introduce any restriction for the case that the output equation (4b) is in explicit form. In that case, $\mathrm{d}\Upsilon = \mathrm{d}\Gamma$, and b.v follows from the preceding conditions and from the definitions.

Next, we prove the necessity and sufficiency of the conditions in Theorem 2.1. To simplify the writing, we present the following auxiliary lemma. The proof is straightforward and omitted; it can be found in [22].

Lemma 2.2: 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 x_0.

Proof of Theorem 2.1. Necessity: 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 now claim that the conditions in Theorem 2.1 hold for a system in the target form (3) with the s-tuple of vector fields $\tilde{X} = \left({\partial}/{\partial z^1}, {\partial}/{\partial z^2}, \ldots, {\partial}/{\partial z^s}\right)$. To show this, first observe that conditions from a. to b.iv. can be proved similarly as in [13], [14]. For b.v. we proceed as follows: first we denote the right hand side of (3a) by $\mathcal{F}_u(z) := A_u z + \Phi_u(y)$, then we compute $\mathrm{d}(\iota_{\mathcal{F}_u}\mathrm{d}z^j)\wedge\mathrm{d}\Upsilon = \mathrm{d}(\sum_{i=1}^n A_u^{ji}z^i + \Phi_u^j(y)) \wedge \mathrm{d}\Upsilon$. Since $\mathrm{d}\Phi_u^j \wedge \mathrm{d}\Upsilon = 0$, it follows that $\mathrm{d}(\iota_{\mathcal{F}_u}\mathrm{d}z^j) \wedge \mathrm{d}\Upsilon = \sum_{i=1}^n A_u^{ji}\mathrm{d}z^i \wedge \mathrm{d}\Upsilon$ is in $\mathrm{span}_{\mathbb{R}}\{\mathrm{d}z^j \mid j = 1, 2, \ldots, n\} \wedge \mathrm{d}\Upsilon = \Omega[\tilde{X}, g] \wedge \mathrm{d}\Upsilon$, and this concludes the necessity proof.

Sufficiency: It is clear that the condition a. cannot hold for $s > n$. Let us first consider the case $s < n$. The proof of sufficiency uses some results described in [13], [14]; 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 (6) into a system in the form $\dot{z} = \bar{A}_u Q^s z + \bar{\Phi}_u(P^s z)$, $\bar{y} = P^s z$, where $\bar{A}_u = [\bar{A}_u^{ji}] \in \mathcal{M}_{n\times(n-s)}(\mathbb{R})$ is a matrix; $j = 1, 2, \ldots, n$; $i = s+1, s+2, \ldots, n$. Now, it is clear that the same change of coordinates transforms system (4) into the system

$$\dot{z} = \bar{A}_u Q^s z + \bar{\Phi}_u(P^s z) \tag{8a}$$
$$0 = C_u P^s z + (P^s z)^\top D_u y + E_u y + F_u \tag{8b}$$

with $Q^s[z^1\, z^2\, \ldots\, z^n]^\top := [z^{s+1}\, z^{s+2}\, \ldots\, z^n]^\top$. Notice that, although the output function $y = \Psi_u(g)$ is a function of g, and $P^s z = \bar{y} = g$, system (8) is not yet necessarily in the form of system (3) because Ψ_u is not necessarily invertible. From [14], 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. From (8a) we have that the vector field f_u in the coordinates $(z^1, z^2, \ldots, z^n)$ reads $f_u = \sum_{j=1}^n \mathcal{F}_u^j {\partial}/{\partial z^j}$ with $\mathcal{F}_u^j = \sum_{i=s+1}^n \bar{A}_u^{ji}z^i + \bar{\Phi}_u^j(P^s z)$ for each $j = 1, 2, \ldots, n$. Since $\mathcal{F}_u^j = \mathcal{L}_{f_u}z^j = \iota_{f_u}\mathrm{d}z^j$ and $\mathrm{d}z^j \in \Omega[X, g]$, we have, from b.v., that there exist $\alpha_u^j = \sum_{i=1}^n \alpha_u^{ji}\mathrm{d}z^i$, with $\alpha_u^{ji} \in \mathbb{R}$ for $j, i = 1, 2, \ldots, n$, such that $\mathrm{d}\mathcal{F}_u^j \wedge \mathrm{d}\Upsilon = \sum_{i=1}^n \alpha_u^{ji}\mathrm{d}z^i \wedge \mathrm{d}\Upsilon$, that is,

$$0 = \mathrm{d}\left(\sum_{i=s+1}^n \bar{A}_u^{ji}z^i + \bar{\Phi}_u^j(P^s z) - \sum_{i=1}^n \alpha_u^{ji}z^i\right) \wedge \mathrm{d}\Upsilon. \tag{9}$$

By Lemma 2.2 it follows that $\sum_{i=s+1}^n \bar{A}_u^{ji}z^i + \bar{\Phi}_u^j(P^s z) - \sum_{i=1}^n \alpha_u^{ji}z^i$ is a function that can be defined by means of $y = [y^1\, y^2\, \ldots\, y^q]^\top$ and, from the conditions $\mathrm{d}y^j \wedge \mathrm{d}\Gamma = 0$, each y^j is a function of $g = P^s z$, that is, $y^j(u, x) = \Psi_u^j(P^s z)$. In particular, since $\mathrm{d}\Upsilon \neq 0$, we have $s \geq q$. Thus, from $\mathrm{d}y^j = \sum_{i=1}^s {\partial y^j}/{\partial z^i}\mathrm{d}z^i$ we can conclude that $\mathrm{d}\Upsilon = \sum_{\sigma\in\mathcal{S}^q_{(1,2,\ldots,s)}} \Theta_u^\sigma(P^s z)\mathrm{d}\Gamma^\sigma$, 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_u^σ are smooth functions. Since $\mathrm{d}\Upsilon \neq 0$, there exists at least one $\sigma_0 \in \mathcal{S}^q_{(1,2,\ldots,s)}$ with non-vanishing $\Theta_u^{\sigma_0}(P^s z)$. Then, from (9) we have that

$$0 = \sum_{\sigma\in\mathcal{S}^q_{(1,2,\ldots,s)}} \Theta_u^\sigma \mathrm{d}\left(\sum_{i=s+1}^n (\bar{A}_u^{ji} - \alpha_u^{ji})z^i + \bar{\Phi}_u^j - \sum_{i=1}^s \alpha_u^{ji}z^i\right) \wedge \mathrm{d}\Gamma^\sigma. \tag{10}$$

If $s > q$, after (wedge) 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 $0 = \overline{\Theta}_u(P^s z)\sum_{i=s+1}^n (\bar{A}_u^{ji} - \alpha_u^{ji})\mathrm{d}z^i \wedge \mathrm{d}\Gamma$, with $\overline{\Theta}_u = \Theta_u^{\sigma_0}$. Also, if $s = q$, (10) reduces to $0 = \overline{\Theta}_u(P^s z)\sum_{i=s+1}^n (\bar{A}_u^{ji} - \alpha_u^{ji})\mathrm{d}z^i \wedge \mathrm{d}\Gamma$, for a suitable $\overline{\Theta}_u$. Since $\{z^j \mid j = 1, 2, \ldots, n\}$ is a system of coordinates, we can conclude that the family $\{\mathrm{d}z^i \wedge \mathrm{d}\Gamma \mid s+1 \leq i \leq n\}$ is linearly independent. Then, since $\overline{\Theta}_u(P^s z) \neq 0$, we have that $\bar{A}_u^{ji} - \alpha_u^{ji} = 0$ for $s+1 \leq i \leq n$ and, returning to (9), we derive that $\mathrm{d}\left(\bar{\Phi}_u^j(P^s z) - \sum_{i=1}^s \alpha_u^{ji}z^i\right) \wedge \mathrm{d}\Upsilon = 0$, that is, for some function Φ_u, we have $\bar{\Phi}_u^j(P^s z) - \sum_{i=1}^s \alpha_u^{ji}z^i = \Phi_u^j(y)$, with $y = [y^1\, y^2\, \ldots\, y^q]^\top$. Therefore, we can rewrite $\mathcal{F}_u^j(z) = \sum_{i=s+1}^n \bar{A}_u^{ji}z^i + \bar{\Phi}_u^j(P^s z) = \sum_{i=s+1}^n \bar{A}_u^{ji}z^i + \sum_{i=1}^s \alpha_u^{ji}z^i + \Phi_u^j(y)$ and so, setting $A_u^{ji} = \begin{cases} \alpha_u^{ji} & \text{if } i = 1, 2, \ldots, s \\ \bar{A}_u^{ji} & \text{if } i = s+1, s+2, \ldots, n \end{cases}$ we arrive to a system in the form of (3).

Finally, we consider the case $s = n$. In this case, due to condition a., the elements of g form a local coordinate system. Then, in coordinates $z = g(x)$, equation (4a) becomes $\dot{z} = D_{g^{-1}z}g f_u(g^{-1}z) =: \bar{\Phi}_u(z)$. Proceeding as above, from b.v., there exist $\alpha_u^j = \sum_{i=1}^n \alpha_u^{ji}\mathrm{d}z^i$, with $\alpha_u^{ji} \in \mathbb{R}$ for $j, i = 1, 2, \ldots, n$, such that $0 = \mathrm{d}\left(\bar{\Phi}_u^j(P^s z) - \sum_{i=1}^n \alpha_u^{ji}z^i\right) \wedge \mathrm{d}\Upsilon$, that is, $\bar{\Phi}_u^j(P^s z) = \sum_{i=1}^n \alpha_u^{ji}z^i + \Phi_u^j(y)$, for a suitable function Φ_u. Setting $A_u^{ji} = \alpha_u^{ji}$ we conclude that system (4) takes the target form in coordinates $z = g$.

III. An algorithm to find the s-tuple X

In the previous section 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 X of vector fields. This section addresses the problem of finding these vector fields. It is important to note that, as it will be seen, the proposed algorithm will end up with a (possible nonlinear) PDE system. The s-tuple will be any solution of that PDE (we may have more variables than equations), which depending on the system under consideration may not be an easy task. To find an algorithm that provides an unique s-tuple X (avoiding solving the system of PDEs), we may need to restrict the class of systems (4) under consideration (cf. [23]). Before we present the algorithm, we present the following result, which generalizes for the multi-output case the results in [16].

Lemma 3.1: Let system (6) 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 $1 \le k_1 \le k_0$, $\mathcal{B}_{k_1} := \{\mathrm{d}g\} \cup \left\{ \mathrm{d}\mathcal{L}_{u_k}\mathcal{L}_{u_{k-1}} \ldots \mathcal{L}_{u_1} g^i \;\middle|\; \begin{array}{l} k = 1, 2, \ldots, k_1 \text{ and} \\ (i, u_1, u_2, \ldots, u_k) \in \mathcal{S}_k \end{array} \right\}$ is, in a suitable neighborhood U of x_0, a basis for the $C^\infty(U)$-module spanned by $\{\mathrm{d}g\} \cup \left\{ \mathrm{d}\mathcal{L}_{u_k}\mathcal{L}_{u_{k-1}} \ldots \mathcal{L}_{u_1} g^i \;\middle|\; \begin{array}{l} k = 1, 2, \ldots, k_1 \\ i = 1, 2, \ldots, s \\ u_1, u_2, \ldots, u_k \in \mathbb{R}^m \end{array} \right\}$;
- $\mathcal{B}_{k_0}$ is a basis for $\mathrm{d}\mathcal{O}$, where $\mathcal{O}$ is the observable space.

Due to the space restrictions the proof is omitted. Basically, it consists of "generalizing" the ideas from [16] to the multi-output case; details can be found in [22].

THE ALGORITHM: *The case $s < n$.* Consider system (4) and suppose that the auxiliary system (6) 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 $X = (X_1, X_2, \ldots, X_s)$, 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 $r \in \{1, 2, \ldots, s\}$ and $u_1 \in \mathbb{R}^m$, and $\mathcal{I}^X_{(r,\,u_1,\,u_2,\,\ldots,\,u_{k-1},\,u_k)} := (\mathcal{L}_{u_k} \iota_X \mathcal{I}^X_{(r,\,u_1,\,u_2,\,\ldots,\,u_{k-1})}) \wedge \mathrm{d}\Gamma$, for all $r \in \{1, 2, \ldots, s\}$, $k \ge 2$ and $u_1, u_2, \ldots, u_k \in \mathbb{R}^m$. Then, fix a sequence $\mathcal{S}$ as in Lemma 3.1. We look for a s-tuple $X = (X_1, X_2, \ldots, X_s)$ of vector fields, with $X_i = \sum_{j=1}^n X_i^j \partial/\partial x^j$ solving, step by step, the following:

1. $\mathrm{d}g^j(X_i) = \delta_i^j$ for all $i, j \in \{1, 2, \ldots, s\}$;
2. successively for $1 \le k \le k_0$:
   a. for all $j \in \{1, 2, \ldots, s\}$ and $v_1, v_2, \ldots, v_k \in \mathbb{R}^m$, $\mathcal{I}^X_{(j,\,v_1,\,v_2,\,\ldots,\,v_k)} \in \mathrm{span}_{\mathbb{R}}\{\mathcal{I}^X_{(r,\,u_1,\,u_2,\,\ldots,\,u_l)} \mid l \in \{1, 2, \ldots, k\}$ and $(r, u_1, u_2, \ldots, u_l) \in \mathcal{S}_l\}$;
   b. for all $(r, u_1, u_2, \ldots, u_k) \in \mathcal{S}_k$, $\mathrm{d}\iota_X \mathcal{I}^X_{(r,\,u_1,\,u_2,\,\ldots,\,u_k)} = 0$;
3. for all $(j, u_1, u_2, \ldots, u_{k_0}) \in \mathcal{S}_{k_0}$ and $v \in \mathbb{R}^m$, $\mathcal{I}^X_{(j,\,u_1,\,u_2,\,\ldots,\,u_{k_0},\,v)} \in \mathrm{span}_{\mathbb{R}}\{\mathcal{I}^X_{(r,\,u_1,\,u_2,\,\ldots,\,u_l)} \mid l = 1, 2, \ldots, k_0$ and $(r, u_1, u_2, \ldots, u_l) \in \mathcal{S}_l\}$;
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 turns out, 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. This can be verified by noticing that the first steps consist in deriving the properties that the vector fields X_i must satisfy. The last step is related with the implicit nature of the output equation. Note that the solution for X is not necessarily 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 IV-B describes an example that illustrates this algorithm.

The case $s = n$. Let system (6) be observable in the rank sense. From Theorem 2.1, the system (4) can be rewritten in target form iff a. holds together with $\mathrm{d}\mathcal{L}_{f_u} \mathrm{d}g^i \wedge \mathrm{d}\Upsilon \in \mathrm{span}_{\mathbb{R}}\{\mathrm{d}g\} \wedge \mathrm{d}\Upsilon$, for all $i = 1, 2, \ldots, n$.

IV. EXAMPLES

We start to show a simple example that is not covered by the frameworks proposed previously in [5], [6], [8], [12]–[16]. Then, with a second example we elucidate the details of the proposed algorithm.

A. A perspective output system

Consider the following perspective output system with state $x = [x^1\ x^2]^\top \in \mathbb{R}^2$, single input $u \in \mathbb{R}$, and with output $y \in \mathbb{R}$ given by the ratio x^2/x^1 (with $x^1 \neq 0$) that also acts as an external input on the system,

$$\dot{x} = f_u(x) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ y + u \end{bmatrix}; \qquad y = x^2/x^1. \quad (11)$$

Notice that the output equation can be written as $x^1 y = x^2$ or as $\begin{bmatrix} 0 & -1 \end{bmatrix} x + x^\top \left[\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right] y = 0$, which is in the form of equation (3b), because $x = P^2 x$. Thus, system (11) can be rewritten in the form of system (3). This implies that it is possible to design an observer with linear error dynamics, for example, from [3, section 3], we can construct the following robust H_∞ type optimal observer with dynamics $\dot{\hat{x}} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \hat{x} + \begin{bmatrix} 0 \\ y + u \end{bmatrix} - \gamma^2 Q \begin{bmatrix} y^2 & -y \\ -y & 1 \end{bmatrix} \hat{x}$, where $\gamma > 0$ is a given gain level. The matrix Q is a matrix solving $\dot{Q} = \begin{bmatrix} \lambda & 1 \\ -1 & \lambda \end{bmatrix} Q + Q \begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix} - \gamma^2 Q \begin{bmatrix} y^2 - \gamma^{-2} & -y \\ -y & 1 - \gamma^{-2} \end{bmatrix} Q$, $Q(0) = Q_0$, with $Q_0^{-1} > 0$, where $\lambda \ge 0$ denotes a forgetting factor. It follows that the estimation error $\tilde{x} = \hat{x} - x$ satisfies the linear dynamics $\dot{\tilde{x}} = \left(\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} - \gamma^2 Q \begin{bmatrix} y^2 & -y \\ -y & 1 \end{bmatrix} \right) \tilde{x}$.

Now, we show that (11) cannot be written in the form $\dot{z} = A_u z + \Phi_u(y)$; $y = z^1$. To this end, we use the results from [13] (see also [14]–[16]) or from Theorem 2.1, that say that a necessary condition is that the statements a. and b.i.-b.iv. in Theorem 2.1 must hold with $\mathrm{d}\Gamma = \mathrm{d}y$ and $g = y$. From (11), it follows that $\mathrm{d}y = \frac{1}{(x^1)^2}(x^1 \mathrm{d}x^2 - x^2 \mathrm{d}x^1)$, $\mathrm{d}y(f_u) = -1 - y^2 + \frac{y+u}{x^1}$, and $\mathrm{d}\mathcal{L}_{f_u} y \wedge \mathrm{d}y = -\frac{y+u}{(x^1)^3} \mathrm{d}x^1 \wedge \mathrm{d}x^2$. Then, from (7), we can conclude that $\Omega_1^X = \mathrm{span}_{\mathbb{R}} \left\{ \frac{y+u}{(x^1)^3} \mathrm{d}x^1 \wedge \mathrm{d}x^2 \mid u \in \mathbb{R} \right\} =$

$\mathrm{span}_{\mathbb{R}}\left\{\frac{y}{(x^1)^3}\mathrm{d}x^1\wedge\mathrm{d}x^2,\ \frac{y+1}{(x^1)^3}\mathrm{d}x^1\wedge\mathrm{d}x^2\right\}$ has dimension 2. Since $\Omega_1^X\subseteq\Omega^X$, we conclude that b.ii. does not hold.

Also, system (11) cannot be, in a neighborhood of an arbitrary point $x_0=[x_0^1\,x_0^2]^\top$ (with $x_0^1\neq 0$), rewritten in the target forms of [12] and [6]. In [12, eq. (6)] the desired form is: $\dot z=\bar F(u,z)=[\bar F^1(u,z)\ \bar F^2(u,z)]^\top;\quad y=C_1z=z^1$, where $|{\partial\bar F^1}/{\partial z^2}(u,z)|$ must remain away from zero. Notice that by rewriting $f_u=x^2\partial/\partial x^1+(-x^1+y+u)\partial/\partial x^2$ as $f_u=\bar F^1\partial/\partial z^1+\bar F^2\partial/\partial z^2$, we find that $\bar F^1=x^2\partial z^1/\partial x^1+(-x^1+y+u)\partial z^1/\partial x^2=-(z^1)^2-1+{z^1+u}/{x^1}$, and $\partial\bar F^1/\partial z^2(u,z)=(z^1+u)\partial(1/x^1)/\partial z^2$ that vanishes for $z^1+u=0$, that is, the condition does not hold in neighborhoods of points x_0 with $x_0^2=-ux_0^1$. On the other hand, in [6, syst. (3)–(4)] (see also [5]) the desired form is: $\dot z=Az+\beta(\tilde y);\ \tilde y=\Psi(y)=z^2,\ (x^1\neq 0)$; where $A=\begin{bmatrix}0&0\\1&0\end{bmatrix}$, β is a suitable function, and Ψ is a local diffeomorphism. Rewriting $f_0=x^2\partial/\partial x^1+(-x^1+y)\partial/\partial x^2$ as $f_0=\bar{\mathcal F}^1\partial/\partial z^1+\bar{\mathcal F}^2\partial/\partial z^2$ (we fix the input because in [6, syst. (3)–(4)] it is not considered), we find that $\bar{\mathcal F}^2=x^2\partial z^2/\partial x^1+(-x^1+y)\partial z^2/\partial x^2=-x^2D_y\Psi x^2/(x^1)^2+(-x^1+y)D_y\Psi 1/x^1=(-y^2-1+y/x^1)D_y\Psi$. Since $\tilde y=z^2$, the system will be in the target form of [6] only if $\partial\bar{\mathcal F}^2/\partial z^1=1$ and, from $y=\Psi^{-1}(z^2)$, we find that $1=\partial\bar{\mathcal F}^2/\partial z^1=yD_y\Psi\partial(1/x^1)/\partial z^1$ and we may write $\mathrm{d}z^1\wedge\mathrm{d}z^2=yD_y\Psi\partial(1/x^1)/\partial z^1\mathrm{d}z^1\wedge\mathrm{d}z^2=yD_y\Psi\mathrm{d}(1/x^1)\wedge\mathrm{d}z^2$, from which we obtain $\mathrm{d}z^1\wedge\mathrm{d}z^2=yD_y\Psi(-1/(x^1)^2)\mathrm{d}x^1\wedge D_y\Psi 1/x^1\mathrm{d}x^2=-y/(x^1)^3(D_y\Psi)^2\mathrm{d}x^1\wedge\mathrm{d}x^2$. Therefore $\mathrm{d}z^1\wedge\mathrm{d}z^2$ vanishes if $y=0$, that is, (z^1,z^2) is not a system of coordinates in the neighborhoods of points x_0 with $x_0^2=0$.

The example also shows that the possibility of finding a transformation that yields a system in the target form (3) depends on the choice of g in the output equation (4b). If, for example, we rewrite $(x^1)^2y=x^1x^2$, in a small neighborhood of $[1\ 0]^\top$ we have that $\hat g=x^1x$ remains close to $g=x$ but, we loose the target form in the variables $z=\hat g=[(x^1)^2\ x^1x^2]^\top$. That is, we would have $\Omega[X,\hat g]=\mathrm{span}_{\mathbb{R}}\{\mathrm{d}z^1,\mathrm{d}z^2\}$ and $\mathrm{d}\iota_{f_0}\mathrm{d}z^2\wedge\mathrm{d}y=(2y^2-2+y/x^1)\mathrm{d}x^1\wedge\mathrm{d}x^2$, $\mathrm{d}z^1\wedge\mathrm{d}y=2\mathrm{d}x^1\wedge\mathrm{d}x^2$, and $\mathrm{d}z^2\wedge\mathrm{d}y=2y\mathrm{d}x^1\wedge\mathrm{d}x^2$. Since $2y^2-2+y/x^1\notin\mathrm{span}_{\mathbb{R}}\{1,y\}$, in any neighborhood of $[1\ 0]^\top$, condition b.v will not hold.

Notice also that we do not require the auxiliary output to be flat (see definition in [8], [9]) . Indeed system (11) can be written in target form for $g=x$ and, because $\dot g^1-g^2=0$, g is not differentially algebraically independent, which means that g is not flat.

B. Illustration of the algorithm

Consider the following system

$$\dot x=f_u(x)=\left[\begin{array}{l}\mathrm{e}^{-x^1}(x^3+u)\\ x^3+(x^1+x^2-\mathrm{e}^{x^1})^2-\mathrm{e}^{-x^1}(x^3+u)\\ x^3+\sin(x^1+x^2-\mathrm{e}^{x^1})\end{array}\right],$$

$$y=\left[\begin{array}{l}x^1+x^2-\mathrm{e}^{x^1}\\ \mathrm{e}^{-x^1}(x^1+x^2)\end{array}\right],$$

with state $x=[x^1\ x^2\ x^3]^\top$, output $y=[y^1\,y^2]^\top$ and input $u\in\mathbb{R}$, where the goal is to write it in the form (3) in a neighborhood $\mathcal N$ of the point $x_0=[0\ 0\ 0]^\top$. From the output equation select $g^1(x)=\mathrm{e}^{x^1}$ 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

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

with $D_u=D=\mathrm{col}_{\mathcal M}\left(\begin{bmatrix}0&0\\0&0\end{bmatrix},\begin{bmatrix}0&1\\0&0\end{bmatrix}\right)$ and therefore it can be written in the form (4b). In this case, $\mathrm{d}\Gamma:=\mathrm{d}g^1\wedge\mathrm{d}g^2=\mathrm{e}^{x^1}\mathrm{d}x^1\wedge\mathrm{d}x^2\neq 0$, $\mathrm{d}y^1=(1-\mathrm{e}^{x^1})\mathrm{d}x^1+\mathrm{d}x^2$, $\mathrm{d}y^2=\mathrm{e}^{-x^1}(1-x^1-x^2)\mathrm{d}x^1+\mathrm{e}^{-x^1}\mathrm{d}x^2$ and $\mathrm{d}\Upsilon=\mathrm{e}^{-x^1}y^1\mathrm{d}x^1\wedge\mathrm{d}x^2$. Moreover, $\mathrm{d}y^1\wedge\mathrm{d}\Gamma=0=\mathrm{d}y^2\wedge\mathrm{d}\Gamma$. The next step consists in checking the observability of the auxiliary system $\dot x=f_u(x)$, $\left\{\begin{array}{l}\bar y^1=g^1=\mathrm{e}^{x^1}\\ \bar y^2=g^2=x^1+x^2\end{array}\right.$, in the rank sense. To this end, we see that $S:=\{g^1,g^2,\mathcal L_0g^1\}$ is a subset of the observable space $\mathcal O$, 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 $\mathcal S_1=\{(1,0)\}\subset\{1,2\}\times\mathbb{R}$ satisfies the conditions of Lemma 3.1. Let $X=(X_1,X_2)$ be a pair of vector fields with $X_1:=X_1^1\partial/\partial x^1+X_1^2\partial/\partial x^2+X_1^3\partial/\partial x^3$ and $X_2:=X_2^1\partial/\partial x^1+X_2^2\partial/\partial x^2+X_2^3\partial/\partial x^3$. We can now go through the algorithm proposed in Section III: From the conditions on Step 1, $\mathrm{d}g^i(X_j)=\delta^i_j$, we derive the identities $X_1^1=\mathrm{e}^{-x^1}$, $X_1^2=-\mathrm{e}^{-x^1}$, $X_2^1=0$, and $X_2^2=1$. Therefore,

$$\begin{array}{rcl}X_1&=&\mathrm{e}^{-x^1}\partial/\partial x^1-\mathrm{e}^{-x^1}\partial/\partial x^2+X_1^3\partial/\partial x^3;\\ X_2&=&\partial/\partial x^2+X_2^3\partial/\partial x^3.\end{array}\tag{12}$$

Now, we move to Step 2.a. From $\mathcal L_ug^1=x^3+u$ and $\mathcal L_ug^2=x^3+(y^1)^2$ 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 does 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 $\mathrm{d}x^1\wedge\mathrm{d}x^2\wedge\mathrm{d}x^3(X_1,X_2,V)=(-X_1^3-X_2^3\mathrm{e}^{-x^1})V^1-X_2^3\mathrm{e}^{-x^1}V^2+\mathrm{e}^{-x^1}V^3$ we obtain

$$\begin{aligned}\iota_X\mathcal I^X_{(1,0)}&=\iota_X(\mathrm{e}^{x^1}\mathrm{d}x^1\wedge\mathrm{d}x^2\wedge\mathrm{d}x^3)\\&=(-X_1^3\mathrm{e}^{x^1}-X_2^3)\mathrm{d}x^1-X_2^3\mathrm{d}x^2+\mathrm{d}x^3\end{aligned}\tag{13}$$

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, rewriting 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 $\mathrm{e}^{x^1}\frac{\partial X_1^3}{\partial x^2}+\frac{\partial X_2^3}{\partial x^2}-\frac{\partial X_2^3}{\partial x^1}=0$, that is,

$$\frac{\partial X_2^3}{\partial x^3}=\frac{\partial X_1^3}{\partial x^3}=0\quad\text{and}\quad\mathrm{e}^{x^1}\frac{\partial X_1^3}{\partial x^2}+\frac{\partial X_2^3}{\partial x^2}=\frac{\partial X_2^3}{\partial x^1}.\tag{14}$$

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, after some direct computations, we find

$$\begin{aligned}&\mathcal L_v\iota_X\mathcal I^X_{(1,0)}\\&=\mathrm{d}(x^3(1-X_1^3-X_2^3)-vX_1^3-(y^1)^2X_2^3+\sin(y^1)).\end{aligned}\tag{15}$$

Then, using the first identity in (14), 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 it follows that $1 - X_1^3 - X_2^3$ must be a constant, that is,

$$\mathrm{d}(X_1^3 + X_2^3) = 0. \tag{16}$$

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 (13), $\iota_X\mathcal{I}^X_{(1,\,0)} \wedge \mathrm{d}\Upsilon = ((-X_1^3\mathrm{e}^{x^1} - X_2^3)\mathrm{d}x^1 - X_2^3\mathrm{d}x^2 + \mathrm{d}x^3) \wedge \mathrm{d}\Upsilon = \mathrm{d}x^3 \wedge \mathrm{d}\Upsilon$. Hence, from the Step 4 we must have that for all $v \in \mathbb{R}$ the 3-forms $(\mathrm{d}\iota_{f_v}\mathrm{d}g^1) \wedge \mathrm{d}\Upsilon$, $(\mathrm{d}\iota_{f_v}\mathrm{d}g^2) \wedge \mathrm{d}\Upsilon$, 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 $(\mathrm{d}\iota_{f_v}\mathrm{d}g^1) \wedge \mathrm{d}\Upsilon = \mathrm{d}(x^3 + u) \wedge \mathrm{d}\Upsilon = \mathrm{d}x^3 \wedge \mathrm{d}\Upsilon \in \mathbb{E}$, $(\mathrm{d}\iota_{f_v}\mathrm{d}g^2) \wedge \mathrm{d}\Upsilon = \mathrm{d}(x^3 + (y^1)^2) \wedge \mathrm{d}\Upsilon = \mathrm{d}x^3 \wedge \mathrm{d}\Upsilon \in \mathbb{E}$, and $(\mathrm{d}\iota_{f_v}\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Upsilon = (\mathcal{L}_v\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Upsilon$. From (15) and (16) we have $(\mathcal{L}_v\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Upsilon = (1 - X_1^3 - X_2^3)\mathrm{d}x^3 \wedge \mathrm{d}\Upsilon - v\mathrm{d}X_1^3 \wedge \mathrm{d}\Upsilon - (y^1)^2\mathrm{d}X_2^3 \wedge \mathrm{d}\Upsilon$. Then, from the first identity in (14), we have that $(\mathcal{L}_v\iota_X\mathcal{I}^X_{(1,\,0)}) \wedge \mathrm{d}\Upsilon = (1 - X_1^3 - X_2^3)\mathrm{d}x^3 \wedge \mathrm{d}\Upsilon$, and from (16), 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 (12), (16) and (14). We can now select a pair of vector fields satisfying those conditions; 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 (13), 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 $\partial/\partial x^1 = \mathrm{e}^{x^1}\partial/\partial z^1 + \partial/\partial z^2$, $\partial/\partial x^2 = \partial/\partial z^2$, and $\partial/\partial x^3 = \partial/\partial z^3$. After substitution, we conclude that $f_u = \mathrm{e}^{-x^1}(x^3+u)\partial/\partial x^1 + (x^3 + (x^1+x^2-\mathrm{e}^{x^1})^2 - \mathrm{e}^{-x^1}(x^3+u))\partial/\partial x^2 + (x^3 + \sin(x^1 + x^2 - \mathrm{e}^{x^1}))\partial/\partial x^3$ reads $f_u = (z^3+u)\partial/\partial z^1 + (z^3 + (y^1)^2)\partial/\partial z^2 + (z^3 + \sin(y^1))\partial/\partial z^3$ in the new coordinates. In other words we arrive to the system in the desired target form, with $P^2z = [z^1\; z^2]^\top$

$$\dot{z} = \begin{bmatrix} z^3 + u \\ z^3 + (y^1)^2 \\ z^3 + \sin(y^1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} z + \begin{bmatrix} u \\ (y^1)^2 \\ \sin(y^1) \end{bmatrix}$$
$$0 = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} P^2z + (P^2z)^\top D_u y + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} y.$$

Again, note that the previous change of variables works globally and not only in a neighborhood of the origin x_0, while the "explicit choice": $g^1 = y^1$, $g^2 = y^2$, cannot be even part of a system of coordinates around a point $\bar{x}$ where $y^1(\bar{x}) = \bar{x}^1 + \bar{x}^2 - \mathrm{e}^{\bar{x}^1}$ vanishes, because $\mathrm{d}y^1 \wedge \mathrm{d}y^2|_{\bar{x}} = \mathrm{e}^{-x^1}y^1\mathrm{d}x^1 \wedge \mathrm{d}x^2|_{\bar{x}} = 0$.

V. Concluding remarks

In this paper, for a specific class of nonlinear systems, 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 an algorithm. The procedure has been illustrated with examples.

References

[1] A. A. Agrachev and Y. L. Sachkov, *Control Theory from the Geometric Viewpoint*, ser. Encyclopaedia of Mathematical Sciences, no. 87, Springer, 2004.

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

[3] ——, "Robust filtering for deterministic systems with implicit outputs," *Systems Control Lett.*, vol. 58, no. 4, pp. 263–270, 2009.

[4] G. Besançon, "An overview on observer tools for nonlinear systems," in *Nonlinear Observers and Applications*. Springer, 2007, ch. 1.

[5] D. Boutat and K. Busawon, "Extended nonlinear observable canonical form for multi-output dynamical systems," in *Proceedings Joint 48th IEEE CDC and 28th Chinese Control Conference*, December 2009, pp. 6520–6525.

[6] ——, "On the transformation of nonlinear dynamical systems into the extended nonlinear observable canonical form," *Internat. J. Control*, vol. 84, no. 1, pp. 94–106, 2011.

[7] M. P. do Carmo, *Differential Forms and Applications*, ser. Universitext. Springer, 1994.

[8] M. Fliess, J. Lévine, P. Martin, and P. Rouchon, "Flatness and defect of nonlinear systems: Introductory theory and examples," *Int. J. Contr.*, vol. 61, no. 6, pp. 1327–1361, 1995.

[9] ——, "A Lie–Bäcklund approach to equivalence and flatness of nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 44, no. 5, pp. 922–937, 1999.

[10] A. Glumineau, C. H. Moog, and F. Plestan, "New algebro-geometric conditions for the linearization by input-output injection," *IEEE Trans. Automat. Contr.*, vol. 41, no. 4, pp. 598–603, 1996.

[11] O. Grasselli and A. Isidori, "An existence theorem for observers of bilinear systems," *IEEE Trans. Automat. Control*, vol. 26, pp. 1299–1300, 1981.

[12] H. Hammouri, G. Bornard, and K. Busawon, "High gain observer for structured multi-output nonlinear systems," *IEEE Trans. Automat. Control*, vol. 55, no. 4, pp. 987–992, 2010.

[13] H. Hammouri and J.-P. Gauthier, "Bilinearization up to output injection," *Systems Control Lett.*, vol. 11, pp. 139–149, 1988.

[14] ——, "Time-varying linearization up to output injection," in *Proceedings of the 28th CDC, Tampa, Florida*, December 1989.

[15] ——, "Global time-varying linearization up to output injection," *SIAM J. Control Optim.*, vol. 30, pp. 1295–1310, 1992.

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

[17] V. Jurdjević, *Geometric Control Theory*, ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1997, no. 51.

[18] A. J. Krener, "On the equivalence of control systems and the linearization of nonlinear systems," *SIAM J. Control*, vol. 11, pp. 670–676, 1973.

[19] A. J. Krener and A. Isidori, "Linearization by output injection and nonlinear observers," *Systems Control Lett.*, vol. 3, pp. 47–52, 1983.

[20] A. J. Krener and W. Respondek, "Nonlinear observers with linearizable error dynamics," *SIAM J. Control Optim*, vol. 23, pp. 1295–1310, 1985.

[21] V. López Morales, F. Plestan, and A. Glumineau, "Linearization by completely generalized input-output injection," *Kibernetika*, vol. 35, no. 6, pp. 793–802, 1999.

[22] S. S. Rodrigues and A. P. Aguiar, "On the linearization up to multi-output injection for a class of nonlinear systems with implicitly defined outputs," 2012, RICAM-Report No. 2012-16. [Online]. Available: http://www.ricam.oeaw.ac.at/publications/reports/

[23] ——, "A new algorithm for linearization up to multi-output and multi-input injection for a class of systems with implicitly defined outputs," in *Proceedings of the 2013 European Control Conference (ECC), Zurich, Switzerland*, July 2013, pp. 1734–1739.