Convergence of Nonlinear Observers on $\mathbb{R}^n$ with a Riemannian Metric (Part III)

Ricardo G. Sanfelice and Laurent Praly

***Abstract*—This paper is the third and final component of a three-part effort on observers contracting a Riemannian distance between the state of the system and its estimate. In Part I, we showed that such a contraction property holds if the system dynamics and the Riemannian metric satisfy two key conditions: a differential detectability property and a geodesic monotonicity property. With the former condition being the focus of Part II, in this Part III, we study the latter condition in relationship to the nullity of the second fundamental form of the output function. We formulate sufficient and necessary conditions for it to hold. We establish a link between it and the infinite gain margin property, and we provide a systematic way for constructing a metric satisfying this condition. Finally, we illustrate cases where both conditions hold.**

I. Introduction

A. Background

We consider nonlinear systems on $\mathbb{R}^n$ of the form

$$\dot{\boldsymbol{x}} \;=\; \boldsymbol{f}(\boldsymbol{x}) \;, \quad \boldsymbol{y} \;=\; \boldsymbol{h}(\boldsymbol{x}), \tag{1}$$

where $\boldsymbol{x}$ represents the state living in $\mathbb{R}^n$, $\boldsymbol{y} : \mathbb{R}^n \to \mathbb{R}^n$ represents the measured output living in $\mathbb{R}^p$ and $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ and $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^p$ are functions.

For this class of systems, we continue the study, started in [1] and [2], of designing

1) a state observer, namely, a dynamical system

$$\dot{\hat{\boldsymbol{x}}} \;=\; \boldsymbol{F}(\hat{\boldsymbol{x}}, \boldsymbol{h}(\boldsymbol{x})) \;, \tag{2}$$

with a state $\hat{\boldsymbol{x}}$ living in the same manifold as the system state $\boldsymbol{x}$ to be estimated, such that the zero estimation error set

$$\mathcal{A} = \{(\boldsymbol{x}, \hat{\boldsymbol{x}}) \in \mathbb{R}^n \times \mathbb{R}^n : \boldsymbol{x} = \hat{\boldsymbol{x}}\} \tag{3}$$

2) is forward invariant,
3) solutions to (1)-(2) converge to it – a property that is guaranteed when a Riemannian distance between true and estimated state strictly decreases,
4) and has an infinite gain margin (see Definition 3.1).

There is a large corpus of contributions dedicated to this problem in the literature. The case when the distance is Euclidean, in appropriate coordinates, has been deeply investigated, giving rise to the well-known Luenberger observer [3], Kalman filter [4], and high-gain observer [5]. The case where the distance is derived from a Riemannian metric given by the dynamics or by the manifold the state belongs to was studied in [6], [7], [8]. The design procedure proposed there exploits properties of the given metric to establish local convergence of the distance to zero, specifically, via an appropriate choice of coordinates or modification of the metric. In this paper, which continues from [1] and [2], the Riemannian metric is not given but properly chosen as part of the design of the observer.

R. G. Sanfelice is with the Department of Electrical and Computer Engineering, University of California, Santa Cruz, CA 95064, USA. Email: ricardo@ucsc.edu

L. Praly is with CAS, ParisTech, Ecole des Mines, 35 rue Saint Honoré, 77305, Fontainebleau, France. Email: Laurent.Praly@ensmp.fr

B. Motivation

The choice of the Riemannian metric mentioned above is dictated by the following result reported in Theorem 3.3 and Lemma 3.6 in [1] (see also [9]). We state it slightly differently but keep the original numbering of the conditions. The symbols and notions – e.g., *complete*, *Riemannian metric*, $\boldsymbol{d}_1^2\wp$, *geodesic*, and *Riemannian distance* – are defined in Appendix A1.

Theorem 1.1: Given C^3 functions $\boldsymbol{f}$ and $\boldsymbol{h}$, suppose there exists a complete C^3 Riemannian metric $\boldsymbol{P}$ and an open subset Ω of $\mathbb{R}^n$ such that

A2 : There exist a continuous function $\rho : \Omega \to [0, +\infty)$ and a strictly positive real number q satisfying[1]

$$\mathcal{L}_{\boldsymbol{f}}\boldsymbol{P}(\boldsymbol{x}) \leq \; \rho(\boldsymbol{x})\, \boldsymbol{dh}(\boldsymbol{x}) \otimes \boldsymbol{dh}(\boldsymbol{x}) - q\, \boldsymbol{P}(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Omega\,. \tag{4}$$

A3 : There exists a C^3 function $\wp : (\boldsymbol{y}_1, \boldsymbol{y}_2) \in \mathbb{R}^p \times \mathbb{R}^p \mapsto \wp(\boldsymbol{y}_1, \boldsymbol{y}_2) \in [0, +\infty)$ satisfying

$$\wp(\boldsymbol{y}, \boldsymbol{y}) = 0 \;, \quad \boldsymbol{d}_1^2\wp(\boldsymbol{y}, \boldsymbol{y}) > 0 \qquad \forall \boldsymbol{y} \in \boldsymbol{h}(\Omega) \tag{5}$$

and, for any geodesic γ^*, taking values in Ω and minimizing on the maximal interval (s_1, s_2), we have

$$\begin{aligned} \frac{d}{ds}\left\{\wp(\boldsymbol{h}(\gamma^*(s)), \boldsymbol{h}(\gamma^*(s_3)))\right\} \;>\; 0 \qquad & \forall s \in (s_3, s_4), \\ \forall s_3, s_4 \in (s_1, s_2) : \qquad & \qquad (6) \\ s_3 < s_4 \quad \& \quad \boldsymbol{h}(\gamma^*(s_3)) \neq \boldsymbol{h}(\gamma^*(s_4)) \;. \end{aligned}$$

Under these conditions, for any strictly positive real number E and any closed subset $\mathcal{C}$ of Ω with a nonempty interior, there exists a continuous function $k_E^* : \mathcal{C} \to \mathbb{R}_{>0}$ such that

– for any continuous function $k_E : \mathcal{C} \to \mathbb{R}$ satisfying

$$k_E(\hat{\boldsymbol{x}}) \geq k_E^*(\hat{\boldsymbol{x}}) \qquad \forall \hat{\boldsymbol{x}} \in \mathcal{C} \;,$$

– for the observer given by

$$\dot{\hat{\boldsymbol{x}}} \;=\; \boldsymbol{F}(\hat{\boldsymbol{x}}, y) \;:=\; \boldsymbol{f}(\hat{\boldsymbol{x}}) \;-\; k_E(\hat{\boldsymbol{x}})\, \boldsymbol{g}_{\boldsymbol{P}}[\wp \circ \boldsymbol{h}](\hat{\boldsymbol{x}}, \boldsymbol{y}) \;, \tag{7}$$

[1] Component-wise the inequality (4) is

$$\sum_c \left[\frac{\partial P_{ab}}{\partial x_c}(x) f_c(x) + P_{ac}(x)\frac{\partial f_c}{\partial x_b}(x) + P_{bc}(x)\frac{\partial f_c}{\partial x_a}(x)\right] \leq \rho(x) \sum_i \frac{\partial h_i}{\partial x_a}(x)\frac{\partial h_i}{\partial x_b}(x) - qP_{ab}(x) \;,$$

and the observer equation (7) is

$$\dot{\hat{x}}_a = f_a(\hat{x}) - k_E(\hat{x}) \sum_b [P(\hat{x})^{-1}]_{ab} \sum_i \frac{\partial h_i}{\partial x_b}(\hat{x}) \frac{\partial \wp}{\partial y_{1i}}(h(\hat{x}), y) \;.$$

where, for each $\boldsymbol{y}$, $\hat{\boldsymbol{x}} \mapsto \boldsymbol{g}_{\boldsymbol{P}}[\wp \circ \boldsymbol{h}](\hat{\boldsymbol{x}}, \boldsymbol{y})$ is the Riemannian gradient (with respect to $\hat{\boldsymbol{x}}$) of the function $\hat{\boldsymbol{x}} \mapsto \wp(\boldsymbol{h}(\hat{\boldsymbol{x}}), \boldsymbol{y})$,

– and, for all x and $\hat{x}$ in $\mathcal{C}$ satisfying

$$d(\hat{\boldsymbol{x}}, \boldsymbol{x}) < E \ , \tag{8}$$

where d denotes the Riemannian distance induced by $\boldsymbol{P}$, and linked by a minimizing normalized geodesic γ^* satisfying

$$\boldsymbol{x} = \boldsymbol{\gamma}^*(0) \ , \quad \hat{\boldsymbol{x}} = \boldsymbol{\gamma}^*(\hat{s}) \ , \quad \gamma^*(s) \in \mathcal{C} \quad \forall s \in [0, \hat{s}] \ ,$$

we have[2]

$$\mathfrak{D}^+ d(\hat{\boldsymbol{x}}, \boldsymbol{x}) \ \leq \ -\frac{q}{4}\, d(\hat{\boldsymbol{x}}, \boldsymbol{x}) \ . \tag{9}$$

The consequence of (9) in Theorem 1.1 is that, as long as the assumptions are satisfied, the distance between the true state $\boldsymbol{x}$ and the estimated state $\hat{\boldsymbol{x}}$ is exponentially decreasing. Among the assumptions is the fact that the Riemannian metric $\boldsymbol{P}$ must satisfy two key conditions that are of complete different nature.

The first condition, referred to as Condition A2, named *strong differential detectability with respect to the metric* $\boldsymbol{P}$, is related to detectability of (1), and, as such pertains to control theory. It involves the right-hand side $\boldsymbol{f}$, the output map $\boldsymbol{h}$, and the Riemannian metric $\boldsymbol{P}$ to be chosen. Geometrically, it says that the flow generated by (1) contracts along directions that are tangent to the level sets of h. In [1], we show that a weak form of this differential detectability property is necessary for the existence of an observer with state $\hat{\boldsymbol{x}}$ in the same space as the system state $\boldsymbol{x}$ and making the set $\mathcal{A}$ in (3) invariant and a Riemannian distance between true state and its estimate to decrease exponentially when evaluated along solutions. We show also that uniform detectability of the linearization of (1) along each of its solutions is necessary for Condition A2 to hold. In [2], we present techniques for the design of the Riemannian metric $\boldsymbol{P}$ for given functions $\boldsymbol{f}$ and $\boldsymbol{h}$ so that Condition A2 holds. We show that such a design is possible when (1) satisfies any of the following properties:

i) Strongly infinitesimally reconstructible (see [2, Definition 3.1]) in the sense that each time-varying linear system resulting from the linearization along a solution to the system (1) satisfies a uniform reconstructibility property;
ii) Strongly differentially observable, in the sense that the state to output derivatives mapping is an injective immersion (see [2, Proposition 4.4]);

The second condition, referred to as Condition A3, says roughly that, if, along a geodesic, the distance between the true state $\boldsymbol{x}$ and its estimate $\hat{\boldsymbol{x}}$ reduces then the same holds between the corresponding true (measured) output $\boldsymbol{y}$ and its estimate $\boldsymbol{h}(\hat{\boldsymbol{x}})$. Following [10, Definition 6.2.3], a function $\boldsymbol{h}$ satisfying Condition A3 is known as being *geodesically monotone*. This property involves the output map $\boldsymbol{h}$ and the Riemannian metric $\boldsymbol{P}$, but not $\boldsymbol{f}$. In [1, Proposition A3], we established that this property implies that the level sets of $\boldsymbol{h}$ are strongly convex, which is a property that is typically exploited in optimization theory; see, e.g., [11]. Actually it is needed only to allow E in (8) to be arbitrary – in this way, making the result semiglobal. Indeed, in [2] we show that, without it, Condition A2 alone guarantees the existence of a locally (i.e., E is imposed and small enough) convergent observer and a locally convergent reduced order observer. See [2, Propositions 2.4 and 2.8].

[2] $\mathfrak{D}^+ d(\hat{\boldsymbol{x}}, \boldsymbol{x}) \ = \ \limsup_{t \searrow 0} \dfrac{d(\hat{\boldsymbol{X}}((\hat{\boldsymbol{x}}, \boldsymbol{x}), t), \boldsymbol{X}(\boldsymbol{x}, t)) - d(\hat{\boldsymbol{x}}, \boldsymbol{x})}{t}$

C. Contributions

Parallel to [1], dedicated to the study and design of a metric satisfying the strong differential detectability property of Condition A2, we devote this paper to the geodesic monotonicity property in Condition A3. Our contributions are as follows:

1) In Section III-A, we show that Condition A3 is equivalent to the infinite gain margin property (see Definition 3.1) when the correction (or innovation) term in the observer is of gradient type as in (7).
2) In Section III-B and Proposition 3.4, we show that Condition A3 holds if $\boldsymbol{h}$ is a (geodesically) affine function.
3) In Section III-C, we give necessary conditions for Condition A3 to hold. In particular, we reveal the key role played by the second fundamental form of the function $\boldsymbol{h}$ (see Definition 3.7), and the fact that $\boldsymbol{h}$ is a Riemannian submersion (see Definition 3.17), with totally geodesic (see Definition 3.13) level sets and an integrable orthogonal distribution (see Definition 3.10).
4) In Corollary 3.9, we propose a test to check if Condition A3 holds. The conditions to check depend on symbolic computations involving $\boldsymbol{h}$, $\boldsymbol{P}$, and their differentials.
5) In Theorem 3.24, we present a systematic way to construct a metric $\boldsymbol{P}$ satisfying Condition A3.
6) In Section IV, we illustrate, via examples, situations in which both Conditions A2 and A3 hold.

Because of space limitations, details behind routine (but sometimes lengthy) computations involved in the proofs are omitted. They can be found in [12], along with additional material not referred to in this paper.

The reading of this paper requires the knowledge of well established concepts and results from Riemannian geometry, in particular, on Riemannian submersions and on optimization on Riemannian manifolds. [13], [14], [15], [11] are relevant references on such topics.

As a difference to our previous work, coordinates play a significant role in the solution we have found to design a Riemannian metric $\boldsymbol{P}$ satisfying Condition A3. Our first step is to introduce our notation involving coordinates and related basic assumptions.

II. Preliminaries and Notation

Symbols in bold style represent coordinate-free objects. In particular, $\boldsymbol{x}$ and $\boldsymbol{y}$ are points in a manifold, $\boldsymbol{f}$ is a vector field on a tangent bundle, $\boldsymbol{h}$ is a function between manifolds, $\boldsymbol{P}$ is a symmetric 2-covariant tensor, etc.

Once coordinates, defined below, x and y, with letters in normal style type, have been chosen for $\boldsymbol{x}$ and $\boldsymbol{y}$, we can express the corresponding objects: $f(x)$ for the value of $\boldsymbol{f}$ at the point $\boldsymbol{x}$, $h(x)$ for $\boldsymbol{h}(\boldsymbol{x})$, $P(x)$ for $\boldsymbol{P}(\boldsymbol{x})$, etc.

The writing of the system in (1) and of the observer in (7) with the state, the output, and the functions in bold style means that the coordinates used therein play no role, namely, the expression of both the plant and the observer dynamics are coordinate free. However, the use of normal style type for k_E and $\wp$ in the observer (7), and for d in (8) and in (9) is to indicate that a change of coordinates in $[0, +\infty)$ is not allowed, i.e., these are scalar invariant functions taking values in $[0,+\infty)$.

As a general rule, when not used as indices, the symbols x, P, d, γ, ϕ, $\mathcal{C}$ ... are used for the $\boldsymbol{x}$-manifold $\mathbb{R}^n$, while the symbols y, Q, e, δ, χ, $\mathcal{D}$, ..., following in the alphabetical order, are used for the $\boldsymbol{y}$-manifold $\mathbb{R}^p$. For example, the expression of the Riemannian norm of the velocity of a path and of the distance between two points are denoted $\frac{d\gamma}{ds}(s)^\top P(\gamma(s))\frac{d\gamma}{ds}(s)$ and $d(x_a,x_b)$ in the $\boldsymbol{x}$-manifold, while, in the $\boldsymbol{y}$-manifold, they are denoted $\frac{d\delta}{ds}(s)^\top Q(\delta(s))\frac{d\delta}{ds}(s)$ and $e(y_a,y_b)$, respectively.

As usual (see [16, p. 40]) we call *coordinate chart* a triple, respectively, $(x,\mathcal{M},\phi)$ for the $\boldsymbol{x}$-manifold $\mathbb{R}^n$ and $(y,\mathcal{N},\chi)$ for the $\boldsymbol{y}$-manifold $\mathbb{R}^p$, such that $\mathcal{M}$ and $\mathcal{N}$, called *coordinate domains*, are open subsets of, respectively, $\mathbb{R}^n$ and $\mathbb{R}^p$ and $\phi:\mathcal{M}\to\mathbb{R}^n$ and $\chi:\mathcal{N}\to\mathbb{R}^n$ are homeomorphisms, called *coordinate maps*, satisfying

$$x \;=\; \phi(\boldsymbol{x}) \quad \forall \boldsymbol{x}\in\mathcal{M} \quad , \qquad y \;=\; \chi(\boldsymbol{y}) \quad \forall \boldsymbol{y}\in\mathcal{N} \; .$$

where $x=(x_a,x_b,\dots)$ in $\mathbb{R}^n$ and $y=(y_i,y_j,\dots)$ in $\mathbb{R}^p$ are called *local coordinates*. Roman letters a, b, ..., are used as indices for x and run over the range $\{1,2,\dots,n\}$ and roman letters i, j, ..., are used as indices for y and run over the range $\{1,2,\dots,p\}$.

Denoting the family of (as many times as necessary) continuously differentiable functions C^s, the coordinate charts $(x,\mathcal{M},\phi)$ and $(y,\mathcal{N},\chi)$ are assumed to assure a C^s structure, in the sense that, for any two coordinate charts $(x_1,\mathcal{M}_1,\phi_1)$ and $(x_2,\mathcal{M}_2,\phi_2)$, $\phi_1\circ\phi_2^{-1}$ is a C^s diffeomorphism. A coordinate chart $(x,\mathcal{M},\phi)$ is said to be *a coordinate chart around* $\boldsymbol{x}_0$ if $\boldsymbol{x}_0$ belongs to $\mathcal{M}$.

As an illustration of these definitions, given coordinate charts $(x,\mathcal{M},\phi)$ around $\boldsymbol{x}_0$ and $(y,\mathcal{N},\chi)$ around $\boldsymbol{h}(\boldsymbol{x}_0)$, with $\boldsymbol{h}(\mathcal{M})$ contained in $\mathcal{N}$, the expression h of the function $\boldsymbol{h}$ is

$$h(x) \;=\; \chi(\boldsymbol{h}(\phi^{-1}(x))) \qquad \forall x\in\phi(\mathcal{M}).$$

Given coordinate charts $(x,\mathcal{M},\phi)$ and $(y,\mathcal{N},\chi)$, and C^s diffeomorphisms $\mathcal{C}:\phi(\mathcal{M})\to\mathbb{R}^n$ and $\mathcal{D}:\chi(\mathcal{N})\to\mathbb{R}^p$, we obtain new coordinates charts $(\bar{x},\bar{\mathcal{M}},\bar{\phi})$ and $(\bar{y},\bar{\mathcal{N}},\bar{\chi})$, where

$$\begin{array}{rclcrclcrcl} \bar{x} &=& \mathcal{C}(x) \;, & & \bar{\mathcal{M}} &=& \mathcal{M} \;, & & \bar{\phi} &=& \mathcal{C}\circ\phi \;, \\ \bar{y} &=& \mathcal{D}(y) \;, & & \bar{\mathcal{N}} &=& \mathcal{N} \;, & & \bar{\chi} &=& \mathcal{D}\circ\chi \;. \end{array}$$

Then we have, for example, the following relationships between the expressions of the vector field $\boldsymbol{f}$, the function $\boldsymbol{h}$, and the symmetric 2-covariant tensor $\boldsymbol{P}$:

$$\begin{array}{rcl} \bar{f}(\mathcal{C}(x)) &=& \frac{\partial \mathcal{C}}{\partial x}(x)\, f(x) \;, \\ \bar{h}(\mathcal{C}(x)) &=& \mathcal{D}(h(x)) \;, \\ \frac{\partial \mathcal{C}}{\partial x}(x)^\top \bar{P}(\mathcal{C}(x))\, \frac{\partial \mathcal{C}}{\partial x}(x) &=& P(x) \;, \end{array} \tag{10}$$

where (f,h,P) and $(\bar{f},\bar{h},\bar{P})$ are the expressions of $(\boldsymbol{f},\boldsymbol{h},\boldsymbol{P})$ in the corresponding coordinates (x,y) and $(\bar{x},\bar{y})$, respectively.

More insight can be gained in the context of observers when y, coordinates for $\boldsymbol{h}(\boldsymbol{x})$, can be used as part of coordinates for $\boldsymbol{x}$. This motivates the following assumption.

Assumption I: *The function $\boldsymbol{h}:\mathbb{R}^n\to\mathbb{R}^p$ is a submersion*[3] *on a set* Ω.

[3]This means, the set $\Omega = \{\boldsymbol{x}\in\mathbb{R}^n : \mathrm{Rank}\,(\boldsymbol{dh}(\boldsymbol{x}))=p\}$, is open, where $\boldsymbol{dh}$ is the differential of $\boldsymbol{h}$.

When this assumption holds, we have the following result. See [17, Theorem I.2.1(2)].

Theorem 2.1 (Local Submersion Theorem): If $\boldsymbol{h}:\mathbb{R}^n\to\mathbb{R}^p$ is a submersion on a set Ω then $\boldsymbol{h}(\Omega)$ is an open set and, for any point $\boldsymbol{x}_0$ of Ω, there exist a coordinate chart $(x,\mathcal{M},\phi)$ around $\boldsymbol{x}_0$, a coordinate chart $(y,\mathcal{N},\chi)$ around $\boldsymbol{h}(\boldsymbol{x}_0)$, with $\mathcal{N}$ containing $\boldsymbol{h}(\mathcal{M})$, and a submersion $h^{\mathsf{com}}:\phi(\mathcal{M})\to\mathbb{R}^{n-p}$ on $\phi(\mathcal{M})$ such that $\mathcal{C}=(h,h^{\mathsf{com}}):\phi(\mathcal{M})\to h(\phi(\mathcal{M}))\times\mathbb{R}^{n-p}$ is a C^s diffeomorphism. Consequently, $((y,z),\mathcal{M},\phi_{\mathcal{N}})$, with $\phi_{\mathcal{N}}=\mathcal{C}\circ\phi$, is a coordinate chart[4] around $\boldsymbol{x}_0$.

In this statement, $z=(z_\alpha,z_\beta,\dots)$ is in the open set $h^{\mathsf{com}}(\phi(\mathcal{M}))$. The greek letters α, β, ..., used as indices, run over the range $\{1,2,\dots,n-p\}$. A coordinate chart $((y,z),\mathcal{M},\phi_{\mathcal{N}})$ is, by nature, paired with the coordinate chart $(y,\mathcal{N},\chi)$, with the same y and, without loss of generality,

$$\boldsymbol{h}(\mathcal{M}) \;=\; \mathcal{N} \;, \quad y \;=\; \chi(\boldsymbol{h}(\phi_{\mathcal{N}}^{-1}(y,z))) \quad \forall (y,z)\in\phi_{\mathcal{N}}(\mathcal{M}) \;.$$

When (y,z) are used as coordinates for $\boldsymbol{x}$, we decompose the expression P of $\boldsymbol{P}$ as

$$P(y,z) \;=\; \left(\begin{array}{cc} P_{yy}(y,z) & P_{yz}(y,z) \\ P_{zy}(y,z) & P_{zz}(y,z) \end{array}\right) \;. \tag{11}$$

In such a case, changes of coordinates take the particular form

$$(\bar{y},\bar{z}) \;=\; (\mathcal{D}(y),\mathcal{E}(y,z))$$

where $\mathcal{D}:\mathbb{R}^p\to\mathbb{R}^p$ is a C^s diffeomorphism and $\mathcal{E}:\mathbb{R}^p\times\mathbb{R}^{n-p}\to\mathbb{R}^{n-p}$ is such that $(\mathcal{D},\mathcal{E})$ is a C^s diffeomorphism.

Throughout the paper, we use objects from differential geometry. Some of these objects are explained in the glossary in Appendix A1.

III. On Condition A3

A. Is Condition A3 necessary?

To answer this question, we invoke the infinite gain margin property, which, in the context of our proposed observer, is defined as follows.

Definition 3.1 (Infinite gain margin [1, Definition 2.8]): Let $\boldsymbol{P}$ be a Riemannian metric on $\mathbb{R}^n$ and Ω an open subset of $\mathbb{R}^n$. An observer

$$\dot{\hat{\boldsymbol{x}}} \;=\; \boldsymbol{f}(\hat{\boldsymbol{x}}) - \mathfrak{C}(\hat{\boldsymbol{x}},\boldsymbol{y}) \;, \tag{12}$$

where $\mathfrak{C}$ is a correction term, is said to have an infinite gain margin on Ω with respect to $\boldsymbol{P}$ if, for any geodesic $\boldsymbol{\gamma}^*$ taking values in Ω and minimizing on the maximal interval (s_1,s_2), we have

$$\begin{array}{ll} \text{either} & \dfrac{d\boldsymbol{\gamma}^*}{ds}(s_4)^\top \boldsymbol{P}(\boldsymbol{\gamma}^*(s_4))\,\mathfrak{C}(\boldsymbol{\gamma}^*(s_4),\boldsymbol{h}(\boldsymbol{\gamma}^*(s_3))) \;>\; 0 \\ \text{or} & \mathfrak{C}(\boldsymbol{\gamma}^*(s_4),\boldsymbol{h}(\boldsymbol{\gamma}^*(s_3))) \;=\; 0 \end{array} \tag{13}$$

$$\forall s_3,s_4\in(s_1,s_2) : \; s_3<s_4 \;,$$

From the first order variation formula (see, for instance, [18, Theorem 6.14] or [19, Theorem 5.7]), and properties of Riemannian distances and geodesics, the upper right-hand Dini derivative of the distance between $\boldsymbol{x}=\boldsymbol{\gamma}^*(s_3)$ and $\hat{\boldsymbol{x}}=\boldsymbol{\gamma}^*(s_4)$ satisfies

[4]The subscript $\mathcal{N}$ in $\phi_{\mathcal{N}}$ is introduced to emphasize that this particular coordinate chart involves, in its construction, a coordinate chart for the $\boldsymbol{y}$-manifold $\mathbb{R}^p$.

$$\mathfrak{D}^+ d(\hat{\boldsymbol{x}},\boldsymbol{x}) \;\leq\; -\frac{d\boldsymbol{\gamma}^*}{ds}(s_4)^\top \boldsymbol{P}(\boldsymbol{\gamma}^*(s_4))\,\boldsymbol{\mathfrak{C}}(\boldsymbol{\gamma}^*(s_4),\boldsymbol{h}(\boldsymbol{\gamma}^*(s_3))))$$
$$+\left[\frac{d\boldsymbol{\gamma}^*}{ds}(s_4)^\top \boldsymbol{P}(\boldsymbol{\gamma}^*(s_4))\boldsymbol{f}(\boldsymbol{\gamma}^*(s_4)) \right. \tag{14}$$
$$\left. -\frac{d\boldsymbol{\gamma}^*}{ds}(s_3)^\top \boldsymbol{P}(\boldsymbol{\gamma}^*(s_3))\boldsymbol{f}(\boldsymbol{\gamma}^*(s_3))\right] \; .$$

Hence, when (13) holds, the correction term $\mathfrak{C}$ always contributes to the decrease of the distance between $\hat{\boldsymbol{x}}$ and $\boldsymbol{x}$. If the contribution of $\mathfrak{C}$ were to be negative, then the desired decrease of the distance would have to be provided by the dynamics of the system dictated by $\boldsymbol{f}$ – namely, by the term around brackets in (14).

Remark 3.2: Although observers without infinite gain margin do exist, those with infinite gain margin are quite common. They are guaranteed to exist for any system as in (1) belonging to the following family:

"Euclidean family": There exists a coordinate chart $(x,\mathcal{M},\phi)$ for which the expression h of the output is

$$h(x) \;=\; Hx$$

and the expression P of a Riemannian metric $\boldsymbol{P}$ satisfying Condition A2 is a constant matrix.

Indeed, when h satisfies this property, (13) simplifies to[5]

$$(\hat{x}-x)^\top P\,\boldsymbol{\mathfrak{C}}(\hat{x},Hx) \;>\; 0 \qquad \forall(\hat{x},x)\in\phi(\mathcal{M})^2 : \hat{x}\neq x$$

and it suffices to pick

$$\boldsymbol{\mathfrak{C}}(\hat{x},Hx) \;=\; k_E(\hat{x})P^{-1}H^\top(H\hat{x}-Hx) \; ,$$

where $k_E : \mathbb{R}^n \to \mathbb{R}$ is a continuous function.

In the particular case where the correction term in the observer is of gradient type as in (7), the infinite gain margin property is equivalent to Condition A3. This equivalence results from the following consequence of (7):

$$\begin{aligned}
&\frac{d}{ds}\wp(h(\gamma^*(s)),h(\gamma^*(s_3)))\\
&\quad=\; \frac{\partial\wp}{\partial y_1}(h(\gamma^*(s)),h(\gamma^*(s_3)))\frac{\partial h}{\partial x}(\gamma^*(s))\frac{d\gamma^*}{ds}(s) \; ,\\
&\quad=\; \tfrac{1}{k_E(x)}\,\frac{d\gamma^*}{ds}(s)^\top P(\gamma^*(s))\mathfrak{C}(\gamma^*(s),h(\gamma^*(s_3)))^\top \; ,
\end{aligned}$$

written with the coordinate chart $(x,\mathcal{M},\phi)$ where

$$x \;=\; \gamma^*(s_3) \quad , \qquad \hat{x} \;=\; \gamma^*(s) \; .$$

We conclude that, if we want an observer of gradient type with a correction term contributing to the decrease of the distance between the state and its estimate, Condition A3 must hold.

B. A Sufficient Condition for the Satisfaction of Condition A3

When the output function $\boldsymbol{h}$ is given, the function $\wp$ and the metric $\boldsymbol{P}$ are the only objects remaining at our disposal to satisfy Condition A3.

1) Choice of the Function $\wp$: We need the function $\wp$ to quantify the "gap" between $\boldsymbol{h}(\hat{\boldsymbol{x}})$ and $\boldsymbol{h}(\boldsymbol{x})$. This motivates us to equip the $\boldsymbol{y}$-manifold $\mathbb{R}^p$ with a C^s complete Riemannian metric $\boldsymbol{Q}$ and obtain a distance, which we denote e. Then, we choose the function $\wp$ as

$$\wp(\boldsymbol{y}_1,\boldsymbol{y}_2) \;=\; e(\boldsymbol{y}_1,\boldsymbol{y}_2)^2 \; . \tag{15}$$

[5]The fact that P is constant allows to express any minimal geodesic γ^* between x and $\hat{x}$ as a straight line connecting $x=\gamma^*(s_3)$ to $\hat{x}=\gamma^*(s_4)$.

To have the required smoothness property on $\wp$, we need an extra property for $\boldsymbol{Q}$, as the following lemma states.

Lemma 3.3: Assume the complete metric $\boldsymbol{Q}$ is such that any piece of geodesic $\boldsymbol{\delta}$ is minimizing. Then the function $\wp$ defined in (15) is C^s and, for any $\boldsymbol{y}_1$, any coordinate chart $(y,\mathcal{N},\chi)$ around $\boldsymbol{y}_1$, and any $\boldsymbol{y}_2$ in $\mathcal{N}$ that is linked to $\boldsymbol{y}_1$ by a (minimizing) geodesic $\boldsymbol{\delta}^*$ with values in $\mathcal{N}$ on (s_1,s_2), i.e. we have

$$\delta^*(s_1)=y_1 \; , \quad \delta^*(s_2)=y_2 \; , \quad \delta^*(s)\in\chi(\mathcal{N}) \;\forall s\in(s_1,s_2)\; ,$$

we obtain

$$\begin{aligned}
\left.\frac{\partial^2\wp}{\partial y_1^2}(y_1,y_2)\right|_{y_2=y_1} \;&=\; 2\,Q(y_1)\; ,\\
\frac{\partial e^2}{\partial y_1}(y_1,y_2)^\top \;&=\; 2\,\frac{Q(y_1)\frac{d\delta^*}{ds}(s_1)e(y_1,y_2)}{\sqrt{\frac{d\delta^*}{ds}(s_1)^\top Q(\delta^*(s_1))\frac{d\delta^*}{ds}(s_1)}}\; .
\end{aligned} \tag{16}$$

Proof: The claim follows from [17, Proposition III.4.8] since, whatever $\boldsymbol{y}_1$ is in the $\boldsymbol{y}$-manifold $\mathbb{R}^p$, no geodesic emanating from this point has a cut point. ■

With what precedes, we consider the following assumption.

Assumption II: *The $\boldsymbol{y}$-manifold $\mathbb{R}^p$ is equipped with a complete metric $\boldsymbol{Q}$ such that any corresponding geodesic $\boldsymbol{\delta}$ is minimizing on $\mathbb{R}$ and the function $\wp$ is the square of the corresponding distance.*

The simplest way to satisfy the condition in Lemma 3.3 is to choose the metric $\boldsymbol{Q}$ flat. Precisely, with

– $\mathbb{R}^p$ equipped with a global coordinate chart $(\bar{y},\mathbb{R}^p,i_d)$ with the identity matrix I_p as expression of a metric;
– $\boldsymbol{\mathcal{D}}$ an homeomorphism from $\mathbb{R}^p$ onto $\mathbb{R}^p$,

for any $\boldsymbol{y}_0$ in $\mathbb{R}^p$ and any coordinate chart $(y,\mathcal{N},\chi)$ around $\boldsymbol{y}_0$, the expressions $\mathcal{D}$ and Q of $\boldsymbol{\mathcal{D}}$ and $\boldsymbol{Q}$ satisfy (see (10)):

$$\bar{y}=\mathcal{D}(y)\;,\quad Q(y)=\frac{\partial\mathcal{D}}{\partial y}(y)^\top I_p\frac{\partial\mathcal{D}}{\partial y}(y) \qquad \forall y\in\chi(\mathcal{N})\;.$$

This implies that the distance e is Euclidean when $\boldsymbol{y}$ is expressed with the specific coordinates $\bar{y}$ which may not necessarily be the physical quantities provided by the sensors.

2) Choice of the Metric $\boldsymbol{P}$: With the function $\wp$ being the square of the distance e in the $\boldsymbol{y}$-manifold $\mathbb{R}^p$ as in (15), in view of the interpretation of Condition A3 in terms of geodesic monotonicity, the simplest case for this condition to hold is when the image under $\boldsymbol{h}$ of a geodesic in the $\boldsymbol{x}$-manifold is a geodesic in the $\boldsymbol{y}$-manifold.

Proposition 3.4: Suppose Assumptions I and II hold. Then, Condition A3 is satisfied if any geodesic $\boldsymbol{\gamma}^*$, in the $\boldsymbol{x}$-manifold $\mathbb{R}^n$, that takes values in the open set $\Omega\subset\mathbb{R}^n$ on a maximal interval (s_1,s_2), is such that $s\in(s_1,s_2)\mapsto\boldsymbol{h}(\boldsymbol{\gamma}^*(s))$ is either constant or a geodesic in the $\boldsymbol{y}$-manifold $\mathbb{R}^p$.

Remark 3.5: The "Euclidean family" introduced in Remark 3.2 gives the simplest case we can think of for this property to hold. Indeed with choosing a constant matrix for the expression of $\boldsymbol{Q}$, we take advantage of the property that the image by a linear function of a straight line is a straight line.

Proof: By assumption, for any geodesic $\boldsymbol{\gamma}^*$ taking values in Ω and minimizing on the maximal interval (s_1,s_2), the function $s\in(s_1,s_2)\mapsto\boldsymbol{h}(\boldsymbol{\gamma}^*(s))$ is either constant or a geodesic $s\in(s_1,s_2)\mapsto\delta^*(s)$ in the $\boldsymbol{y}$-manifold $\mathbb{R}^p$. In the former case, we have

$$\boldsymbol{h}(\boldsymbol{\gamma}^*(s_3)) = \boldsymbol{h}(\boldsymbol{\gamma}^*(s_4)) \qquad \forall s_3, s_4 \in (s_1, s_2) : \ s_3 < s_4 \ .$$

In the latter case, we have

$$\frac{d\boldsymbol{h}\circ\boldsymbol{\gamma}^*}{ds}(s) = \frac{d\delta^*}{ds}(s) \neq 0 \quad \forall s \in (s_3, s_4)$$

and

$$\boldsymbol{h}(\boldsymbol{\gamma}^*(s_3)) \neq \boldsymbol{h}(\boldsymbol{\gamma}^*(s_4)) \qquad \forall s_3, s_4 \in (s_1, s_2) : \ s_3 < s_4 \ .$$

On another hand, with Assumption II, any geodesic in the $\boldsymbol{y}$-manifold is minimizing, so (16) gives,

$$\frac{d}{ds}\left\{\wp(\boldsymbol{h}(\boldsymbol{\gamma}^*(s), \boldsymbol{h}(\boldsymbol{\gamma}^*(s_3))\right\} \ = \ 2\, e(\boldsymbol{h}(\boldsymbol{\gamma}^*(s), \boldsymbol{h}(\boldsymbol{\gamma}^*(s_3)) \times \times\sqrt{\frac{d\boldsymbol{h}\circ\boldsymbol{\gamma}^*}{ds}(s)^\top \boldsymbol{Q}(\boldsymbol{\gamma}^*(s))\frac{d\boldsymbol{h}\circ\boldsymbol{\gamma}^*}{ds}(s)} \qquad \forall s \in (s_3, s_2) \ , \quad \forall s_3 \in (s_1, s_2) \ .$$

where, e being a distance for $\mathbb{R}^p$, the right-hand side is strictly positive. Hence, (6) holds. ■

To make the condition in Proposition 3.4 more explicit, we note that $\boldsymbol{h}(\boldsymbol{\gamma})$ is a geodesic in the $\boldsymbol{y}$-manifold $\mathbb{R}^p$ if and only if $\boldsymbol{h}(\boldsymbol{\gamma})$ satisfies the geodesic equation and therefore, thanks to [16, Theorem 9.12] for example, if and only if, the following holds:

– for any $\boldsymbol{x}_0$ in Ω and any pair of coordinate charts $(x, \mathcal{M}, \phi)$ around $\boldsymbol{x}_0$ and $(y, \mathcal{N}, \chi)$ around $\boldsymbol{h}(\boldsymbol{x}_0)$, with "objects" expressed in these coordinates and with Γ^a and Δ^i denoting the respective Christoffel symbol matrices, namely
$$\Gamma^a \ = \ (\Gamma^a_{bc}) \quad , \qquad \Delta^i \ = \ (\Delta^i_{jk}) \ ,$$
– for any geodesic γ^*, with values in $\phi(\mathcal{M})$ and their image by $\boldsymbol{h}$ in $\chi(\mathcal{N})$, and minimizing on the maximal interval (s_1, s_2),

the geodesic equation in $\chi(\mathcal{N})$ and in $\phi(\mathcal{M})$ are, respectively, for all s in (s_1, s_2),

$$\begin{aligned} 0 \ = \ & \frac{d^2 h_i\circ\gamma^*}{ds^2}(s) \\ & + \frac{d\gamma^*}{ds}(s)^\top \frac{\partial h}{\partial x}(\gamma^*(s))^\top \Delta^i(h(\gamma^*(s)))\frac{\partial h}{\partial x}(\gamma^*(s))\frac{d\gamma^*}{ds}(s) \\ & \qquad\qquad\qquad\qquad \forall i \in \{1,2,\dots,p\} \ , \\ 0 \ = \ & \frac{d^2\gamma^*_c}{ds^2}(s) \ + \ \frac{d\gamma^*}{ds}(s)^\top \Gamma^c(\gamma^*(s))\frac{d\gamma^*}{ds}(s) \\ & \qquad\qquad\qquad\qquad \forall c \in \{1,2,\dots,n\} \ . \end{aligned}$$

Then, with the uniqueness of the solution to the geodesic equation, we have the following result. Its proof can be found in [13] or [11].

Lemma 3.6 ([13, Proposition 1.5] or [11, Theorem 6.4.1]): Let Ω be an open subset of $\mathbb{R}^n$. Suppose Assumption I holds. The following two properties are equivalent[6]

1) Any geodesic $\boldsymbol{\gamma}^*$, in the $\boldsymbol{x}$-manifold $\mathbb{R}^n$ that takes values in Ω on a maximal interval (s_1, s_2), is such that $s \in (s_1, s_2) \mapsto \boldsymbol{h}(\boldsymbol{\gamma}^*(s))$ is either constant or a geodesic in the $\boldsymbol{y}$-manifold $\mathbb{R}^p$.
2) For any point $\boldsymbol{x}_0$ in Ω, any coordinate chart $(x, \mathcal{M}, \phi)$ around $\boldsymbol{x}_0$ and any coordinate chart $(y, \mathcal{N}, \chi)$ around $\boldsymbol{h}(\boldsymbol{x}_0)$, with $\boldsymbol{h}(\mathcal{M})$ contained in $\mathcal{N}$, we have, for all x in $\phi(\mathcal{M}\cap\Omega)$,

$$\begin{aligned} 0 \ = \ & \frac{\partial^2 h_i}{\partial x_a \partial x_b}(x) \\ & -\sum_c \Gamma^c_{ab}(x)\frac{\partial h_i}{\partial x_c}(x) + \sum_{j,k}\Delta^i_{jk}(h(x))\frac{\partial h_j}{\partial x_a}(x)\frac{\partial h_k}{\partial x_b}(x) \end{aligned} \tag{17}$$

Lemma 3.6 motivates the following definition.

Definition 3.7 ([21, p. 123]): We call second fundamental form $\mathit{I\!I}_{\boldsymbol{P}}\boldsymbol{h}$ of $\boldsymbol{h}$ the object defined as follows: For any point $\boldsymbol{x}_0$, any coordinate chart $(x, \mathcal{M}, \phi)$ around $\boldsymbol{x}_0$ and any coordinate chart $(y, \mathcal{N}, \chi)$ around $\boldsymbol{h}(\boldsymbol{x}_0)$, with $\boldsymbol{h}(\mathcal{M})$ contained in $\mathcal{N}$, the expression of $\mathit{I\!I}_{\boldsymbol{P}}\boldsymbol{h}$ is

$$\begin{aligned} \mathit{I\!I}_P h^i_{ab}(x) \ = \ & \frac{\partial^2 h_i}{\partial x_a \partial x_b}(x) \\ & -\sum_c \Gamma^c_{ab}(x)\frac{\partial h_i}{\partial x_c}(x) + \sum_{j,k}\Delta^i_{jk}(h(x))\frac{\partial h_j}{\partial x_a}(x)\frac{\partial h_k}{\partial x_b}(x) \ . \end{aligned} \tag{18}$$

A coordinate-free version of this definition can be found in [14, Definition I.1.4.1] or [15, Definition 3.1.1]. See also [22, Definition 8.1].

Lemma 3.8: The second fundamental form $\mathit{I\!I}_{\boldsymbol{P}}\boldsymbol{h}$ is a bilinear map of a pair of vector fields on the $\boldsymbol{x}$-manifold $\mathbb{R}^n$ into a vector field on the $\boldsymbol{y}$-manifold $\mathbb{R}^p$. It is a 2-covariant to 1-contravariant tensor, i.e. by changing to coordinates

$$\bar{x} = \mathcal{C}(x) \quad , \qquad \bar{y} \ = \ \mathcal{D}(y) \ ,$$

where $\mathcal{C}$ and $\mathcal{D}$ are C^s diffeomorphisms, the expression $\mathit{I\!I}_P\overline{h}$ of the second fundamental form in the new coordinates satisfies

$$\sum_{c,d}\frac{\partial\mathcal{C}_c}{\partial x_a}(x)\frac{\partial\mathcal{C}_d}{\partial x_b}(x)\mathit{I\!I}_P\overline{h}^k_{cd}(\bar{x}) \ = \ \sum_i \frac{\partial\mathcal{D}_k}{\partial y_i}(h(x))\mathit{I\!I}_P h^i_{ab}(x) \ . \tag{19}$$

A proof of Lemma 3.8 can be found in Appendix A2.

We can rephrase the way to guarantee Condition A3 stated in Proposition 3.4 as follows.

Corollary 3.9: Suppose Assumptions I and II hold. Then, Condition A3 is satisfied if the second fundamental form of $\boldsymbol{h}$ is zero on the open subset Ω.

Consequently, (17) provides a test for Condition A3 that involves h, P, and their first derivatives.

C. Necessity of the nullity of $\mathit{I\!I}_{\boldsymbol{P}}\boldsymbol{h}$

With Corollary 3.9, it is tempting to forget about Condition A3 and base our design of the Riemannian metric $\boldsymbol{P}$ on guaranteeing that the second fundamental form of $\boldsymbol{h}$ is zero.

Next, we investigate such an approach. Namely, given $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^p$, a submersion on a set Ω, as in Assumption I and $\boldsymbol{Q}$, a metric equipping the $\boldsymbol{y}$-manifold $\mathbb{R}^p$ and satisfying Assumption II, let the function $\wp$ be the square of the distance e in the $\boldsymbol{y}$-manifold $\mathbb{R}^p$, as written in (15). We are interested in the following question:

(Q1) *Is the second fundamental form of $\boldsymbol{h}$ being null, i.e. the property in (17), a necessary for Condition A3 to hold?*

We address this question, by grouping, as in [13, §3], the equations in (17) into three blocks. We employ the following definition; see [21, & C p.127], [13, p. 77], or [23, p. 205].

[6] According to [20] (but not [13]) a function $\boldsymbol{h}$ satisfying the property in item 1 is said to be affine on Ω. In the case where $p = 1$ and the metric $\boldsymbol{Q}$ is flat, $\boldsymbol{h}$ is said linear affine in [11, p. 88 and following pages] where its necessary and sufficient conditions are presented.

Definition 3.10: The tangent space of the level sets of $\boldsymbol{h}$, denoted $\mathfrak{D}^{\mathsf{tan}}$, is called the tangent distribution. It satisfies

$$\mathfrak{D}^{\mathsf{tan}}(\boldsymbol{x}) \;=\; \left\{ v^{\mathsf{tan}} : \boldsymbol{dh}(\boldsymbol{x}) v^{\mathsf{tan}} = 0 \right\} \ ,$$

and does not depend on $\boldsymbol{P}$. Its elements v^{tan} are called tangent vectors.

The $\boldsymbol{P}$-orthogonal complement to the tangent distribution, denoted $\mathfrak{D}^{\mathsf{ort}}_{\boldsymbol{P}}$, is called the orthogonal distribution. It satisfies

$$\mathfrak{D}^{\mathsf{ort}}_{\boldsymbol{P}}(\boldsymbol{x}) \;=\; \left\{ v^{\mathsf{ort}} : v^{\mathsf{tan}\top} \boldsymbol{P}(\boldsymbol{x}) v^{\mathsf{ort}} \;=\; 0 \quad \forall v^{\mathsf{tan}} \in \mathfrak{D}^{\mathsf{tan}}(\boldsymbol{x}) \right\}$$

and does depend on $\boldsymbol{P}$. Its elements v^{ort} are called orthogonal vectors.

What we refer to as *orthogonal distribution* is usually called *horizontal distribution*, and variations of the letter h are used to denote it. But since we use the letter h to denote the output function in this paper, we employ the term *orthogonal* and use the symbol ort. For consistency, we use the symbol tan and call *tangent distribution* what is usually called *vertical distribution*.

Ignoring the $\boldsymbol{x}$-dependence, basic linear algebra leads to the following result.

Lemma 3.11: For any pair of coordinate charts $(x, \mathcal{M}, \phi)$ and $(y, \mathcal{N}, \chi)$, by letting h, P and $\mathfrak{D}^{\mathsf{ort}}_P(x)$ be the corresponding expressions of $\boldsymbol{h}$, $\boldsymbol{P}$ and $\mathfrak{D}^{\mathsf{ort}}_{\boldsymbol{P}}(\boldsymbol{x})$, the orthogonal distribution $\mathfrak{D}^{\mathsf{ort}}_P(x)$ is spanned by the columns of the gradient of h, i.e. by the columns of $P(x)^{-1} \frac{\partial h}{\partial x}(x)^\top$.

Given a pair of coordinate charts $(x, \mathcal{M}, \phi)$ and $(y, \mathcal{N}, \chi)$, $\mathfrak{D}^{\mathsf{ort}}_P(x)$ and $\mathfrak{D}^{\mathsf{tan}}(x)$ are complementary linear subspaces of the tangent space at x of the $\boldsymbol{x}$-manifold $\mathbb{R}^n$. As a consequence, any vector v in this tangent space can be decomposed as

$$v \;=\; v^{\mathsf{ort}} + v^{\mathsf{tan}}$$

with v^{ort} in $\mathfrak{D}^{\mathsf{ort}}_P(x)$ and v^{tan} in $\mathfrak{D}^{\mathsf{tan}}(x)$. This property allows us to decompose (17) in the following three blocks of equations, for each i in $\{1, 2, \dots, p\}$ and each x in $\phi(\mathcal{M} \cap \Omega)$:

$$v^{\mathsf{tan}\top} I\!I_P h^i(x) v^{\mathsf{tan}} \;=\; 0 \quad \forall v^{\mathsf{tan}} \in \mathfrak{D}^{\mathsf{tan}}(x) \ , \tag{20}$$

$$v^{\mathsf{ort}\top} I\!I_P h^i(x) v^{\mathsf{ort}} \;=\; 0 \quad \forall v^{\mathsf{ort}} \in \mathfrak{D}^{\mathsf{ort}}_P(x) \, , \tag{21}$$

$$v^{\mathsf{ort}\top} I\!I_P h^i(x) v^{\mathsf{tan}} \;=\; 0 \quad \forall (v^{\mathsf{tan}}, v^{\mathsf{ort}}) \in \mathfrak{D}^{\mathsf{tan}}(x) \times \mathfrak{D}^{\mathsf{ort}}_P(x) \, . \tag{22}$$

With the above, we can rephrase question (Q1) as follows:

(Q1') *Does Condition A3 imply that equations (20), (21), and (22) are satisfied?*

Our answer builds from the study in [13], [14], [15]. For the sake of completeness, we rewrite in our setting some of the results therein.

1) About Necessity of (20):

Definition 3.12: Given $\boldsymbol{y}_0$ in the $\boldsymbol{y}$-manifold $\mathbb{R}^p$, the set

$$\mathfrak{H}(\boldsymbol{y}_0) \;=\; \{ \boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{y}_0 \} \ .$$

is called the $\boldsymbol{y}_0$-level set of $\boldsymbol{h}$.

Definition 3.13: Given an open subset Ω of $\mathbb{R}^n$ and a point $\boldsymbol{x}_0$ in Ω, the $\boldsymbol{h}(\boldsymbol{x}_0)$-level set $\mathfrak{H}(\boldsymbol{h}(\boldsymbol{x}_0))$ of $\boldsymbol{h}$ is said to be totally geodesic on Ω if any geodesic $\boldsymbol{\gamma}$ taking values in Ω on the maximal interval (s_1, s_2) and satisfying

$$\boldsymbol{dh}(\boldsymbol{\gamma}(s_3)) \frac{d\boldsymbol{\gamma}}{ds}(s_3) = 0 \ ,$$

for some s_3 in (s_1, s_2), satisfies

$$\boldsymbol{h}(\boldsymbol{\gamma}(s)) = \boldsymbol{h}(\boldsymbol{\gamma}(s_3)) \qquad \forall s \in (s_1, s_2) \ .$$

An equivalent definition is that, for any $\boldsymbol{x}_0$ in Ω, there exists a coordinate chart $(x, \mathcal{M}, \phi)$ around $\boldsymbol{x}_0$ such that we have

$$v^{\mathsf{tan}\top} H_P h(x) \, v^{\mathsf{tan}} \;=\; 0 \qquad \forall v^{\mathsf{tan}} \in \mathfrak{D}^{\mathsf{tan}}(x) \ , \ \forall x \in \phi(\mathcal{M} \cap \Omega) \, , \tag{23}$$

with $H_P h$ the expression of the Riemannian Hessian of h.

We have the following result. Its proof can be found in [13, Lemma 3.2(i)] or [1, Proposition A.2.2 and A.3.1.a].

Lemma 3.14: Let Ω be an open set coming from Assumption I and suppose Condition A3 holds with such choice of Ω. Then, for any $\boldsymbol{x}_0$ in Ω and any coordinate chart $(x, \mathcal{M}, \phi)$ around $\boldsymbol{x}_0$, we have (20) or, equivalently, for any $\boldsymbol{x}_0$ in Ω, the $\boldsymbol{h}(\boldsymbol{x}_0)$-level set $\mathfrak{H}(\boldsymbol{h}(\boldsymbol{x}_0))$ is totally geodesic on Ω.

Hence, the answer to Question (Q1') is "yes" as far as (20) is concerned.

Example 3.15: Consider the harmonic oscillator with unknown frequency. Its dynamics are given as

$$\dot{y} = z_\alpha \quad , \qquad \dot{z}_\alpha \;=\; -y z_\beta \quad , \qquad \dot{z}_\beta \;=\; 0 \tag{24}$$

Given $\varepsilon > 0$, we consider the invariant open set[7]

$$\Omega_\varepsilon \;=\; \left\{ (y, z_\alpha, z_\beta) \in \mathbb{R}^3 : \varepsilon < z_\beta y^2 + z_\alpha^2 < \frac{1}{\varepsilon} \ , \ \varepsilon < z_\beta < \frac{1}{\varepsilon} \right\} \ . \tag{25}$$

In [2, Example 4.5], we have obtained the following metric satisfying Condition A2:

$$P(y, z_\alpha, z_\beta) \;=\; \begin{pmatrix} 1 & 0 & -z_\beta & 0 \\ 0 & 1 & 0 & -z_\beta \\ 0 & 0 & -y & -z_\alpha \end{pmatrix} \mathcal{P} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -z_\beta & 0 & -y \\ 0 & -z_\beta & -z_\alpha \end{pmatrix}, \tag{26}$$

where $\mathcal{P}$ is a positive definite symmetric $(4,4)$ matrix. Symbolic computations give $\Gamma^y_{\alpha\alpha} \;=\; \Gamma^y_{\alpha\beta} \;=\; 0$ and, up to some nonzero factor k,

$$\begin{aligned} k\Gamma^y_{\beta\beta} = \\ -y \left[z_\beta^2 \det\left(E_2^\top \mathcal{P} E_2\right) + 2 z_\beta \det\left(E_4^\top \mathcal{P} E_2\right) + \det\left(E_4^\top \mathcal{P} E_4\right) \right] \\ -z_\alpha \Big[z_\beta^2 \det\left(E_1^\top \mathcal{P} E_2\right) + z_\beta \left[\det\left(E_2^\top \mathcal{P} E_3\right) + \det\left(E_1^\top \mathcal{P} E_4\right) \right] \\ + \det\left(E_3^\top \mathcal{P} E_4\right) \Big] \end{aligned}$$

where E_i is the 4×3 matrix made of the 3×3 identity matrix with an additional raw of zeros inserted as its i-th row. Hence (20) and therefore Condition A3 do not hold.

Similarly, it can be shown that Condition A3 does not hold for the other metric, satisfying Condition A2, obtained in [2, Example 3.7].

Fortunately, with the techniques presented below, we shall be able to obtain, in Example 4.5, a metric satisfying Conditions A2 and A3.

2) About (21): We shall not try to establish that (21) is necessary for Condition A3 to hold. Instead, we show below that, maybe after modifying appropriately $\boldsymbol{P}$, we can always

[7]The system is not observable if $y = z_\alpha = 0$ or $z_\beta = 0$.

guarantee that (21) holds. To do so, we start by providing an expression of (21).

Let $(x, \mathcal{M}, \phi)$ be an arbitrary coordinate chart around some $\boldsymbol{x}_0$ in Ω. With Lemma 3.11, an expression, in these coordinates, of the restriction, denoted $\boldsymbol{I\!I}_{\boldsymbol{P}}\boldsymbol{h}^{\mathsf{ort,ort}}$, to the orthogonal distribution $\mathfrak{D}_{\boldsymbol{P}}^{\mathsf{ort}}$ of the second fundamental form of $\boldsymbol{h}$ is, for each i in $\{1, 2, \ldots, p\}$ and each x in $\phi(\mathcal{M} \cap \Omega)$,

$$[I\!I_P h^{\mathsf{ort,ort}}]^i(x) \;=\; \frac{\partial h}{\partial x}(x)P(x)^{-1} I\!I_P h^i(x) P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top \; . \tag{27}$$

By definition, equation (21) is equivalent to the nullity of $[I\!I_P h^{\mathsf{ort,ort}}]^i(x)$, for each i in $\{1, 2, \ldots, p\}$ and each x in $\phi(\mathcal{M} \cap \Omega)$. Via computations using the components of the second fundamental form above, and by expansion using the identity

$$[P^{-1}]_{i\alpha} \;=\; -\sum_l [P^{-1}]_{il} \sum_\eta P_{l\eta}[P_{zz}^{-1}]_{\eta\alpha} \; ,$$

we can establish the following result.

Lemma 3.16: For any $\boldsymbol{x}_0$ in Ω and any pair of coordinate charts $((y, z), \mathcal{M}, \phi_{\mathcal{N}})$ around $\boldsymbol{x}_0$ and $(y, \mathcal{N}, \chi)$ around $\boldsymbol{h}(\boldsymbol{x}_0)$, the expression of $\boldsymbol{I\!I}_{\boldsymbol{P}}\boldsymbol{h}^{\mathsf{ort,ort}}$ in (27) is

$$\begin{aligned} 2[I\!I_P h^{\mathsf{ort,ort}}]^i_{jk} \;=&\\ &\sum_l Q_{il}^{-1}\left(\frac{\partial Q_{lk}}{\partial y_j} + \frac{\partial Q_{lj}}{\partial y_k} - \frac{\partial Q_{jk}}{\partial y_l}\right)\\ &\qquad -\sum_l [P_y^{-1}]_{il}\left(\frac{\partial [P_y]_{lk}}{\partial y_j} + \frac{\partial [P_y]_{lj}}{\partial y_k} - \frac{\partial [P_y]_{jk}}{\partial y_l}\right)\\ &+\sum_l [P_y^{-1}]_{il}\sum_{\alpha,\beta}[P_{zz}^{-1}]_{\alpha\beta}\times\\ &\qquad \times\left(\frac{\partial [P_y]_{lk}}{\partial z_\alpha}P_{\beta j} + \frac{\partial [P_y]_{lj}}{\partial z_\alpha}P_{\beta k} - \frac{\partial [P_y]_{jk}}{\partial z_\alpha}P_{\beta l}\right) \; , \end{aligned}$$

where we have denoted

$$P_y(y,z) = P_{yy}(y,z) - P_{yz}(y,z)P_{zz}(y,z)^{-1}P_{zy}(y,z) \; . \tag{28}$$

A proof of Lemma 3.16 can be found in [12].

It follows from this expression that, if P_y in (28) does not depend on z and is chosen such that it has the same Christoffel symbols as those of Q, then condition (21) holds. These two conditions are trivially satisfied if we simply have for all (y, z) in $\phi_{\mathcal{N}}(\mathcal{M})$,

$$Q(y) = P_y(y) = P_{yy}(y,z) - P_{yz}(y,z)P_{zz}(y,z)^{-1}P_{zy}(y,z) \; .$$

If, instead of the coordinates (y, z), we use coordinates x, this equation is

$$\left(\frac{\partial h}{\partial x}(x)P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top\right)^{-1} \;=\; Q(h(x)) \qquad \forall x \in \phi(\mathcal{M}) \tag{29}$$

where h and Q are expressed with the same coordinates.

Definition 3.17 ([24, axiom S2]): A submersion $\boldsymbol{h}$ satisfying (29) is called a Riemannian submersion. An equivalent definition is that $\boldsymbol{h}$ is a submersion preserving length of orthogonal vectors, i.e., we have for all v^{ort} in $\mathfrak{D}_P^{\mathsf{ort}}(x)$,

$$v^{\mathsf{ort}\top}\frac{\partial h}{\partial x}(x)^\top Q(h(x))\frac{\partial h}{\partial x}(x)v^{\mathsf{ort}} = v^{\mathsf{ort}\top}P(x)v^{\mathsf{ort}} \, .$$

Now, as indicated above, instead of showing the necessity of (21) for Condition A3 to hold, we answer the following question:

(Q2) *If we are given a metric $\boldsymbol{P}$ satisfying (20) and (22) but neither (21) nor (29), can we modify it to satisfy the three conditions?*

To answer this question, we propose the modification $\boldsymbol{P}_{mod}$ of the metric $\boldsymbol{P}$, the expression of which is, for the coordinate chart $(x, \mathcal{M}, \phi)$, given by

$$\begin{aligned} P_{mod}(x) \;=&\; P(x) + \\ &\frac{\partial h}{\partial x}(x)^\top\left[Q(h(x)) - \left(\frac{\partial h}{\partial x}(x)P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top\right)^{-1}\right]\frac{\partial h}{\partial x}(x) \; , \end{aligned} \tag{30}$$

for all x in $\phi(\mathcal{M})$. This is a positive definite matrix[8]. This definition of $\boldsymbol{P}_{mod}$ via its expression with coordinates gives a covariant 2-tensor that is invariant under a change of coordinates for $\boldsymbol{y}$ (and $\boldsymbol{h}$). Namely, with $\bar{P}_{mod}$ defined as

$$\begin{aligned} \bar{P}_{mod}(\bar{x}) \;=&\; \bar{P}(\bar{x}) + \\ &\frac{\partial \bar{h}}{\partial \bar{x}}(\bar{x})^\top\left[\bar{Q}(\bar{h}(\bar{x})) - \left(\frac{\partial \bar{h}}{\partial \bar{x}}(\bar{x})\bar{P}(\bar{x})^{-1}\frac{\partial \bar{h}}{\partial \bar{x}}(\bar{x})^\top\right)^{-1}\right]\frac{\partial \bar{h}}{\partial \bar{x}}(\bar{x}) \; , \end{aligned}$$

where we have (see (10))

$$\begin{aligned} \bar{x} = \mathcal{C}(x) \quad , &\qquad \bar{h}(\mathcal{C}(x)) = \mathcal{D}(h(x)) \; , \\ \frac{\partial \mathcal{C}}{\partial x}(x)^\top \bar{P}(\mathcal{C}(x))\frac{\partial \mathcal{C}}{\partial x}(x) &\;=\; P(x) \; , \\ \frac{\partial \mathcal{D}}{\partial y}(y)^\top \bar{Q}(\mathcal{D}(y))\frac{\partial \mathcal{D}}{\partial y}(y) &\;=\; Q(y) \; , \end{aligned}$$

we obtain

$$\frac{\partial \mathcal{C}}{\partial x}(x)^\top \bar{P}_{mod}(\bar{x})\frac{\partial \mathcal{C}}{\partial x}(x) \;=\; P_{mod}(x) \; . \tag{31}$$

The metric $\boldsymbol{P}_{mod}$ has the following properties.

Lemma 3.18: Given the metric $\boldsymbol{P}$, let $\boldsymbol{P}_{mod}$ be a metric, the expression of which, with the coordinate chart $(x, \mathcal{M}, \phi)$, is as in (30). The following holds:

1) $\boldsymbol{P}_{mod}$ satisfies condition (29); i.e., $\boldsymbol{h}$ is a Riemannian submersion with $\boldsymbol{P}_{mod}$.
2) Condition (20) holds for $\boldsymbol{P}$ if and only if it holds for $\boldsymbol{P}_{mod}$.
3) $$\mathfrak{D}_{\boldsymbol{P}}^{\mathsf{ort}}(\boldsymbol{x}) = \mathfrak{D}_{\boldsymbol{P}_{mod}}^{\mathsf{ort}}(\boldsymbol{x}) \qquad \forall \boldsymbol{x} \in \Omega \; . \tag{32}$$

Proof: Item 1 follows from the expression of $\boldsymbol{P}_{mod}$ with coordinates (y, z) which is

$$P_{mod}(y,z) \;= \tag{33}$$
$$\begin{pmatrix} Q(y) + P_{yz}(y,z)P_{zz}(y,z)^{-1}P_{zy}(y,z) & P_{yz}(y,z) \\ P_{zy}(y,z) & P_{zz}(y,z) \end{pmatrix} .$$

For item 2, we note that condition (20) is equivalent to (23) where the Hessian of $\boldsymbol{h}$ is related to its gradient by (see (56))

$$H_P h(x) \;=\; \frac{1}{2}\,\mathcal{L}_{g_P h}P(x) \qquad \forall x \in \phi(\mathcal{M}) \; .$$

So the claim follows from the fact that the product rule for Lie differentiation is formally identical with the product rule of ordinary differentiation. Indeed the Lie differentiation of

[8] For any full row rank matrix H and symmetric positive definite matrices P and Q, the matrix $P + H^\top(Q - (HP^{-1}H^\top)^{-1})H$ is positive definite.

(30) gives a matrix $M(x)$ satisfying

$$\mathcal{L}_{g_{P_{mod}}h}P_{mod}(x) = \mathcal{L}_{g_P h}P(x) \\ \qquad + M(x)\frac{\partial h}{\partial x}(x) + \frac{\partial h}{\partial x}(x)^\top M(x)^\top \qquad \forall x \in \phi(\mathcal{M}) \ . \tag{34}$$

For item 3, from (30) we obtain that, for any coordinate chart $(x, \mathcal{M}, \phi)$,

$$P_{mod}(x)P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top \\ = \frac{\partial h}{\partial x}(x)^\top Q(h(x))\left(\frac{\partial h}{\partial x}(x)P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top\right) \ .$$

Since $Q(h(x))$ and $\left(\frac{\partial h}{\partial x}(x)P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top\right)$ are invertible matrices, this establishes that the columns of $P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top$ span the same vector space as the columns of $P_{mod}(x)^{-1}\frac{\partial h}{\partial x}(x)^\top$. With Lemma 3.11, this establishes that the orthogonal distributions $\mathfrak{D}^{\mathrm{ort}}_P(x)$ and $\mathfrak{D}^{\mathrm{ort}}_{P_{mod}}(x)$ are identical. Since the coordinate chart $(x, \mathcal{M}, \phi)$ is arbitrary, we have (32). ∎

With Lemma 3.18 we have answered positively Question (Q2) but only partially. Indeed, we have not established that (22) holds for $\boldsymbol{P}_{mod}$, when (22) holds for $\boldsymbol{P}$. As we show next, working directly with $\boldsymbol{P}_{mod}$ allows to establish this property.

3) About (22)*:* Postponing the study of the necessity of (22) to the next paragraph, here we study what it implies.

Lemma 3.19 ([13, Lemma 3.2(ii)], [14, Proposition I.5.4]): Assume $\boldsymbol{h}$ is a Riemannian submersion on Ω. If, for any point $\boldsymbol{x}_0$ in Ω, there exists a coordinate chart $(x, \mathcal{M}, \phi)$ around $\boldsymbol{x}_0$ such that (22) holds then the distribution $\mathfrak{D}^{\mathrm{ort}}_{\boldsymbol{P}}$ is integrable everywhere locally on Ω, i.e., for any $\boldsymbol{x}_0$ in Ω, and for any pair of coordinate charts $(x, \mathcal{M}, \phi)$ around $\boldsymbol{x}_0$ and $(y, \mathcal{N}, \chi)$ around $\boldsymbol{h}(\boldsymbol{x}_0)$ there exists a C^s function $h^{\mathrm{ort}} : \phi(\mathcal{M}) \to \mathbb{R}^{n-p}$ satisfying

$$\frac{\partial h^{\mathrm{ort}}}{\partial x}(x)P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top \ = \ 0 \qquad \forall x \in \phi(\mathcal{M}) \ , \tag{35}$$

with h being the expression of $\boldsymbol{h}$ with the coordinates y, and such that the function

$$x \mapsto \theta(x) \ = \ (h(x), h^{\mathrm{ort}}(x))$$

is a diffeomorphism.

A proof of Lemma 3.19 can also be found in [12].

In this statement, thanks to Lemma 3.18, we can omit the assumption that $\boldsymbol{h}$ is a Riemannian submersion if we replace $\boldsymbol{P}$ by $\boldsymbol{P}_{mod}$ and (22) holds for $\boldsymbol{P}_{mod}$.

Compared with the claim in the Local Submersion Theorem 2.1, the novelty here is in the fact that the function h^{ort} satisfies (35). This is very useful. Indeed we have the following result.

Lemma 3.20: Assume $\boldsymbol{h}$ is a Riemannian submersion on Ω and, for any $\boldsymbol{x}_0$ in Ω, and for any pair of coordinate charts $(x, \mathcal{M}, \phi)$ around $\boldsymbol{x}_0$ and $(y, \mathcal{N}, \chi)$ around $\boldsymbol{h}(\boldsymbol{x}_0)$, there exists a C^s function $h^{\mathrm{ort}} : \phi(\mathcal{M}) \to \mathbb{R}^{n-p}$ satisfying the properties listed in Lemma 3.19. Under these conditions, the expression $\overline{P}$ of $\boldsymbol{P}$ with the coordinates

$$(y, z) \ = \ (h(x), h^{\mathrm{ort}}(x)) \tag{36}$$

is in the following block diagonal form:

$$\overline{P}(y, z) \ = \ \begin{pmatrix} \overline{P}_y(y,z) & 0 \\ 0 & \overline{P}_z(y,z) \end{pmatrix} \tag{37}$$

Proof: Let P be the expression of $\boldsymbol{P}$ in the coordinates x. From (10), its expression $\overline{P}$ in the coordinates (y, z) satisfies for all x in $\phi(\mathcal{M})$,

$$P(x) \ = \\ \begin{pmatrix} \frac{\partial h}{\partial x}(x)^\top & \frac{\partial h^{\mathrm{ort}}}{\partial x}(x)^\top \end{pmatrix} \overline{P}(h(x), h^{\mathrm{ort}}(x)) \begin{pmatrix} \frac{\partial h}{\partial x}(x) \\ \frac{\partial h^{\mathrm{ort}}}{\partial x}(x) \end{pmatrix} .$$

Post-multiplying by $P(x)^{-1}\begin{pmatrix} \frac{\partial h}{\partial x}(x)^\top & \frac{\partial h^{\mathrm{ort}}}{\partial x}(x)^\top \end{pmatrix}$ and exploiting the invertibility of $\begin{pmatrix} \frac{\partial h}{\partial x}(x)^\top & \frac{\partial h^{\mathrm{ort}}}{\partial x}(x)^\top \end{pmatrix}$, this gives

$$\begin{aligned} I_n \ &= \ \overline{P}(h(x), h^{\mathrm{ort}}(x)) \times \\ &\qquad \times \begin{pmatrix} \frac{\partial h}{\partial x}(x) \\ \frac{\partial h^{\mathrm{ort}}}{\partial x}(x) \end{pmatrix} P(x)^{-1} \begin{pmatrix} \frac{\partial h}{\partial x}(x)^\top & \frac{\partial h^{\mathrm{ort}}}{\partial x}(x)^\top \end{pmatrix} , \\ &= \ \overline{P}(h(x), h^{\mathrm{ort}}(x)) \times \end{aligned} \tag{38}$$

$$\times \begin{pmatrix} \frac{\partial h}{\partial x}(x)P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top & \frac{\partial h}{\partial x}(x)P(x)^{-1}\frac{\partial h^{\mathrm{ort}}}{\partial x}(x)^\top \\ \frac{\partial h^{\mathrm{ort}}}{\partial x}(x)P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top & \frac{\partial h^{\mathrm{ort}}}{\partial x}(x)P(x)^{-1}\frac{\partial h^{\mathrm{ort}}}{\partial x}(x)^\top \end{pmatrix} .$$

But, when (35) holds, the last matrix on the right-hand side is block diagonal. Since this matrix is the inverse of $\overline{P}$, $\overline{P}$ is also block diagonal. ∎

4) About Question (Q1'): Up to now, we have established that, if Condition A3 holds, then (20) and (21) hold, perhaps after changing $\boldsymbol{P}$ into $\boldsymbol{P}_{mod}$. On the other hand, we know with Lemmas 3.18 and 3.19 that, if (22) holds for $\boldsymbol{P}_{mod}$ then the orthogonal distributions $\mathfrak{D}^{\mathrm{ort}}_{\boldsymbol{P}}$ and $\mathfrak{D}^{\mathrm{ort}}_{\boldsymbol{P}_{mod}}$ are integrable. It turns out that, conversely, if, in addition to Condition A3 we have this integrability property, then (22) holds.

Proposition 3.21: If Condition A3 holds with a metric $\boldsymbol{P}$ such that the orthogonal distribution $\mathfrak{D}^{\mathrm{ort}}_{\boldsymbol{P}}$ is integrable, then the second fundamental form of $\boldsymbol{h}$ for the metric $\boldsymbol{P}_{mod}$ is zero on Ω.

Proof: It follows from Lemmas 3.19 and 3.20 that, for any $\boldsymbol{x}_0$ in Ω, there exists a coordinate chart $((y, z), \mathcal{M}, \phi_{\mathcal{N}})$ (see (36)) such that the expressions $\overline{P}$ and $\overline{P}_{mod}$ of $\boldsymbol{P}$ and $\boldsymbol{P}_{mod}$, respectively, are (see (33) and (37))

$$\begin{aligned} \overline{P}(y,z) \ &= \ \begin{pmatrix} \overline{P}_y(y,z) & 0 \\ 0 & \overline{P}_z(y,z) \end{pmatrix} , \\ \overline{P}_{mod}(y,z) \ &= \ \begin{pmatrix} Q(y) & 0 \\ 0 & \overline{P}_z(y,z) \end{pmatrix} . \end{aligned}$$

On the other hand, in these specific coordinates, (20), implied by Condition A3 (see Lemma 3.14), is equivalent to

$$\sum_j [\overline{P}_y^{-1}]_{ij} \frac{\partial [\overline{P}_z]_{\alpha\beta}}{\partial y_j} = \overline{\Gamma}^i_{\alpha\beta} = 0 \ .$$

This implies $\frac{\partial [\overline{P}_z]_{\alpha\beta}}{\partial y_k}$ is zero. Hence $\overline{P}_z$ does not depend on y and its associated Christoffel symbols are

$[\Gamma_{mod}]^i_{jk}(y,z)$

$$= \frac{1}{2}\sum_l [Q(y)^{-1}]_{il}\left(\frac{\partial Q_{lj}}{\partial y_k}(y) + \frac{\partial Q_{lk}}{\partial y_j}(y) - \frac{\partial Q_{jk}}{\partial y_l}(y)\right) \ , \\ = \Delta^i_{jk}(y) \ ,$$

$$[\overline{\Gamma}_{mod}]^i_{j\alpha}(y,z) \ = \ 0 \quad , \qquad [\overline{\Gamma}_{mod}]^i_{\alpha\beta}(y,z) \ = \ 0 \ .$$

The result follows from the fact that the nullity of the second fundamental form does not depend on the coordinates and that we have

$$\begin{aligned} \overline{I\!I}_{P_{mod}} h^i_{jk}(y,z) \ &= \ -[\overline{\Gamma}_{mod}]^i_{jk}(y,z) + \Delta^i_{jk}(y) \ , \\ \overline{I\!I}_{P_{mod}} h^i_{j\alpha}(y,z) \ &= \ -[\overline{\Gamma}_{mod}]^i_{j\alpha}(y,z) \ , \qquad (39) \\ \overline{I\!I}_{P_{mod}} h^i_{\alpha\beta}(y,z) \ &= \ -[\overline{\Gamma}_{mod}]^i_{\alpha\beta}(y,z) \ . \end{aligned}$$

■

This statement is not satisfactory because it requires the extra condition of integrability of the orthogonal distribution $\mathfrak{D}^{\text{ort}}_{\boldsymbol{P}}$. Whether this integrability is implied by Condition A3 is, for us, an open problem. Fortunately, as shown below, when the dimension p of the $\boldsymbol{y}$-manifold is 1, the integrability condition is not needed in the statement of Proposition 3.21. So, in this case, if Condition A3 holds, the second fundamental form of $\boldsymbol{h}$ is zero for the metric $\boldsymbol{P}_{mod}$. Namely, our sufficient condition is necessary but, perhaps after modifying $\boldsymbol{P}$ into $\boldsymbol{P}_{mod}$. Here we recover in some way [1, Proposition A.3.2.b].

What we wrote above about the peculiarity of the case $p = 1$ is a consequence of the fact that the assumption of Lemma 3.20 is always satisfied. Indeed, we have the following result.

Lemma 3.22: Suppose Assumption I is satisfied. If the dimension p of the $\boldsymbol{y}$-manifold $\mathbb{R}^p$ is 1, and (29) and Condition A3 hold, then $\boldsymbol{h}$ is a Riemannian submersion, the second fundamental form of which is zero on Ω.

Proof: Let $\boldsymbol{x}_0$ be any point in Ω and $(x, \mathcal{M}, \phi)$ around $\boldsymbol{x}_0$ and $(y, \mathcal{N}, \chi)$ around $\boldsymbol{h}(\boldsymbol{x}_0)$ be any coordinate charts. With all "objects" expressed with these coordinates, let also $\mathfrak{X}(x,t)$ denote the solution at time t of

$$\dot{x} = P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top \ ,$$

passing through x in $\mathcal{M}$ at time 0. Because the function $t \mapsto h(\mathfrak{X}(x,t))$ is strictly increasing, there exists an open neighborhood $\mathcal{M}$ of $\phi(\boldsymbol{x}_0)$ such that, for any x in $\mathcal{M}$, there exists a (unique) $\tau(x)$ satisfying $h(\mathfrak{X}(x,\tau(x))) = h(\phi(\boldsymbol{x}_0))$ Let h^{com} be the submersion given by the Local Submersion Theorem 2.1 in Section II. We define a function $h^{\text{ort}} : \mathcal{M} \to \mathbb{R}^{n-p}$ as

$$h^{\text{ort}}(x) \ = \ h^{\text{com}}(\mathfrak{X}(x,\tau(x))) \ . \qquad (40)$$

It satisfies

$$\frac{\partial h^{\text{ort}}}{\partial x}(x)P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top \ = \ 0 \qquad \forall x \in \mathcal{M}.$$

Then, with Lemma 3.20, the coordinates

$$(y,z) = (h(x), h^{\text{ort}}(x))$$

defined on $\mathcal{M}$, are such that the expression $\bar{P}$ of $\boldsymbol{P}$ has the following block diagonal form (see [25, p. 57 §19] or [2, Theorem 2.6])

$$P(y,z) \ = \ \begin{pmatrix} P_y(y,z) & 0 \\ 0 & P_z(y,z) \end{pmatrix} \ .$$

From here we proceed as in the proof of Proposition 3.21. ■

Because Condition A3 and the nullity of the second fundamental form of $\boldsymbol{h}$ for $\boldsymbol{P}_{mod}$ are equivalent when $p = 1$, it would be interesting to know if we can always "massage" the measurements to go to this case. In other words, if we have a p-dimensional output function for which the observer problem can be solved, does there exists a 1-dimensional function of this output function with which we can still solve the observer problem. In the linear case, this result is known as Heymann's Lemma[9] [26], [27].

D. Construction of the metric $\boldsymbol{P}$

A design of an appropriate metric $\boldsymbol{P}$ will be provided by another necessary condition for having the second fundamental form zero when $\boldsymbol{h}$ is a Riemannian submersion.

Lemma 3.23: If $\boldsymbol{P}$ and $\boldsymbol{Q}$ are complete and $\boldsymbol{h}$ is a Riemannian submersion on $\mathbb{R}^n$, the second fundamental form of which is zero on $\mathbb{R}^n$, then $\boldsymbol{h}$ is surjective and there exist an $n-p$ dimensional C^s manifold $\boldsymbol{\mathcal{Z}}_h$, a surjective submersion $\boldsymbol{h}^{\text{ort}} : \mathbb{R}^n \to \boldsymbol{\mathcal{Z}}_h$ and a complete metric $\boldsymbol{R}$ on $\boldsymbol{\mathcal{Z}}_h$ such that $\boldsymbol{\theta} = (\boldsymbol{h}, \boldsymbol{h}^{\text{ort}}) : \mathbb{R}^n \to \mathbb{R}^p \times \boldsymbol{\mathcal{Z}}_h$ is a C^s diffeomorphism. Moreover, if $(x, \mathbb{R}^n, \phi)$ and $(y, \mathbb{R}^p, \chi)$ are globally defined coordinate charts, then, for any coordinate chart $(z, \mathcal{O}, \psi)$ for $\boldsymbol{\mathcal{Z}}_h$ used to express $\boldsymbol{h}^{\text{ort}}$ and $\boldsymbol{R}$, we have for all x in $[h^{\text{ort}}]^{-1}(\mathcal{O})$,

$$P(x) \ = \ \frac{\partial h}{\partial x}(x)^\top Q(h(x))\,\frac{\partial h}{\partial x}(x) \qquad (41) \\ + \frac{\partial h^{\text{ort}}}{\partial x}(x)^\top R(h^{\text{ort}}(x))\,\frac{\partial h^{\text{ort}}}{\partial x}(x) \ ,$$

where P, h, h^{ort}, θ, Q, and R denote expressions in the above mentioned coordinates.

Lemma 3.23 is a specific version of the more general [13, Corollary 3.7]. A proof can be found in [12] under more restrictive assumptions of boundedness of the function $\boldsymbol{y} \mapsto \boldsymbol{Q}(\boldsymbol{y})$ and, instead of involutivity, commutation of particular vector fields spanning $\mathfrak{D}^{\text{ort}}_{\boldsymbol{P}}$.

Actually (41) provides a procedure to construct a metric $\boldsymbol{P}$ making the second fundamental form of $\boldsymbol{h}$ zero and, consequently, satisfying Condition A3. The following theorem presents this construction and the forthcoming Example 3.26 illustrates it.

Theorem 3.24: Suppose Assumption I holds. Assume there exist

i) a metric $\boldsymbol{Q}$ for $\mathbb{R}^p$ satisfying Assumption II;
ii) with $q \geq 0$, an $n-p+q$-dimensional C^s manifold $\boldsymbol{\Xi}$ equipped with a metric $\boldsymbol{R}$;
iii) a C^s function $\boldsymbol{h}^{\text{ort}} : \Omega \to \boldsymbol{\Xi}$ is , with rank $n-p$ on Ω such that $\boldsymbol{\theta} = (\boldsymbol{h}, \boldsymbol{h}^{\text{ort}})$ has rank n on Ω.

Then, the metric $\boldsymbol{P}$ defined on Ω as the pull back via $\boldsymbol{\theta}$ of the product metric $\boldsymbol{Q} \oplus \boldsymbol{R}$ (see its expression with coordinates in (41)) is such that $\boldsymbol{h}$ is a Riemannian submersion with a second fundamental form that is zero on Ω. Furthermore, Condition A3 holds when $\wp$ is the square of the distance given by $\boldsymbol{Q}$.

Proof: It follows from our assumptions and the Rank Theorem that the restriction of $\boldsymbol{h}^{\text{ort}}$ to Ω is a subimmersion, i.e., for each $\boldsymbol{x}_0$ in Ω, there exist an open neighborhood $\mathcal{M}_0$ of $\boldsymbol{x}_0$, a C^s manifold $\boldsymbol{\mathcal{Z}}_0$ of dimension $n-p$, a submersion $\boldsymbol{\mathfrak{s}}_0 : \mathcal{M}_0 \to \boldsymbol{\mathcal{Z}}_0$ and an immersion $\boldsymbol{i}_0 : \boldsymbol{\mathcal{Z}}_0 \to \boldsymbol{\Xi}$ satisfying

$$\boldsymbol{h}^{\text{ort}}(\boldsymbol{x}) \ = \ \boldsymbol{i}_0(\boldsymbol{\mathfrak{s}}_0(\boldsymbol{x})) \qquad \forall \boldsymbol{x} \in \mathcal{M}_0$$

The index 0 is used here to insist on the fact that all the corresponding objects are $\boldsymbol{x}_0$ dependent.

[9] We are very grateful to Alessandro Astolfi, from Imperial College in London, for pointing out this lemma.

Let $(\xi, \mathcal{P}, \omega)$ be a coordinate chart around $\boldsymbol{h}^{\mathsf{ort}}(\boldsymbol{x}_0)$ in Ξ, $(z, \mathcal{O}, \psi)$ be a coordinate chart around $\boldsymbol{\mathscr{s}}_0(\boldsymbol{x}_0)$ in $\boldsymbol{\mathcal{Z}}_0$ and $(x, \mathcal{M}, \phi)$ be a coordinate chart around $\boldsymbol{x}_0$ in Ω with

$$\mathcal{M} \subset \mathscr{M}_0 \quad , \qquad \boldsymbol{\mathscr{s}}_0(\mathcal{M}) \subset \mathcal{O} \quad , \qquad \boldsymbol{i}_0(\mathcal{O}) \subset \mathcal{P} \ .$$

We have

$$\xi \ = \ h^{\mathsf{ort}}(x) \ = \ i_0(z) \ , \qquad z \ = \ \mathscr{s}_0(x) \qquad \forall x \in \phi(\mathcal{M}) \ .$$

With the function $(\boldsymbol{h}, \boldsymbol{h}^{\mathsf{ort}})$ having rank n, we get

$$n \ = \ \mathsf{Rank} \begin{pmatrix} I_p & 0 \\ 0 & \dfrac{\partial i_0}{\partial z}(\mathscr{s}_0(x)) \end{pmatrix} \begin{pmatrix} \dfrac{\partial h}{\partial x}(x) \\ \dfrac{\partial \mathscr{s}_0}{\partial x}(x) \end{pmatrix}$$

where $\frac{\partial i_0}{\partial z}(\mathscr{s}_0(x))$ is an $((n-p+q),(n-p))$ matrix of rank $n-p$. This implies the square matrix

$$\begin{pmatrix} \dfrac{\partial h}{\partial x}(x)^\top & \dfrac{\partial \mathscr{s}_0}{\partial x}(x)^\top \end{pmatrix}$$

is invertible. It follows that $x \mapsto (y,z) = (h(x), \mathscr{s}_0(x))$ is a diffeomorphism and (y,z) can be used as coordinates for $\boldsymbol{x}$ in a neighborhood of $\boldsymbol{x}_0$. According to (41), the expression P in these coordinates of the metric $\boldsymbol{P}$ is

$$\begin{aligned} &P(y,z) \ = \\ &\begin{pmatrix} I_p & 0 \\ 0 & \dfrac{\partial i_0}{\partial z}(z)^\top \end{pmatrix} \begin{pmatrix} Q(y) & 0 \\ 0 & R(i_0(z)) \end{pmatrix} \begin{pmatrix} I_p & 0 \\ 0 & \dfrac{\partial i_0}{\partial z}(z) \end{pmatrix} \\ &\qquad\qquad = \begin{pmatrix} Q(y) & 0 \\ 0 & P_z(z) \end{pmatrix} \end{aligned}$$

where

$$P_z(z) \ = \ \frac{\partial i_0}{\partial z}(z)^\top R(i_0(z)) \frac{\partial i_0}{\partial z}(z) \ .$$

From here, the proof can be concluded as in the proof of Proposition 3.21. ■

Remark 3.25:

1) The restriction that $\boldsymbol{h}^{\mathsf{ort}}$ has rank $n-p$ is crucial. A counterexample is given by the metric (26), which does not make the level sets of the output function totally geodesic. Indeed, in (26) we have that $n = 3$ and $p = 1$, but $\boldsymbol{h}^{\mathsf{ort}}$ has (generic) rank 3.
2) Formula (41) is remarkable because of the decomposition of P as a sum. In the upcoming Section IV-B, we observe that it simplifies the verification of Condition A2.
3) The family of metrics given by (41) would exactly correspond to the one of those making the second fundamental form $\boldsymbol{h}$ zero if we were not imposing the extra condition that $\boldsymbol{h}$ is a Riemannian submersion.
4) The metric $\boldsymbol{P}$ given by Theorem 3.24 is defined only on Ω and we do not claim it is complete.
5) Once the manifold Ξ is chosen, the existence of the function $\boldsymbol{h}^{\mathsf{ort}}$ satisfying the conditions is not guaranteed. Indeed, as a consequence of Lemma 3.23, $\boldsymbol{h}$ and Ξ cannot be arbitrary. For example $\boldsymbol{h}$ must be surjective and its level sets must be diffeomorphic to each other. Also, Ξ may not be minimal in terms of dimension and there should exist an immersion between $\boldsymbol{\mathcal{Z}}_h$ and Ξ. We illustrate this point in the following example.

Example 3.26: Let the $\boldsymbol{x}$-manifold be $\mathbb{R}^2$ equipped with globally defined coordinates (x_1, x_2). Let $\mathbb{R}$ be the $\boldsymbol{y}$-manifold equipped with a globally defined coordinate y. This yields $n = 2$ and $p = 1$. The output function $\boldsymbol{h}$ we consider is, when expressed in these coordinates,

$$y \ = \ h(x) := x_1^2 + x_2^2 \ .$$

It is a submersion on $\Omega := \mathbb{R}^2 \setminus \{0\}$. The level sets of this output function are diffeomorphic with the unit circle $\mathbb{S}^1$.

To design a metric $\boldsymbol{P}$ satisfying Condition A3, we follow the lines of Theorem 3.24. We could select Ξ as a connected 1-dimensional manifold, this implying $q = 0$ in Theorem 3.24. Instead, we select Ξ as $\mathbb{R}^2 \setminus \{0\}$, in which $\mathbb{S}^1$ is embedded. Then, q in Theorem 3.24, is equal to 1 and $\boldsymbol{h}^{\mathsf{ort}} : \Omega \to \mathbb{R}^2 \setminus \{0\}$ is to be chosen with rank 1. Then, we choose coordinates for $\boldsymbol{x}$ and $\boldsymbol{y}$. For $\boldsymbol{x}$, we keep those defined above, namely, (x_1, x_2). To get an extra degree of freedom, for $\boldsymbol{y}$ we change, via a C^s function $\mathscr{D} : \mathbb{R} \to \mathbb{R}$ with nonvanishing derivative, the original coordinate y in

$$\bar{y} \ = \ \mathscr{D}(y) \ .$$

Let also $h^{\mathsf{ort}} = (h_\alpha^{\mathsf{ort}}, h_\beta^{\mathsf{ort}})$ be the expression of the function $\boldsymbol{h}^{\mathsf{ort}}$ we are interested in. For the needs of Theorem 3.24, by letting

$$\bar{h}(x) \ = \ \mathscr{D}(h(x)) \ ,$$

since $n = 2$ and $p = 1$, we want, for each $x \in \mathbb{R}^2 \setminus \{0\}$,

$$\mathsf{Rank} \begin{pmatrix} \frac{\partial \bar{h}}{\partial x}(x) \\ \frac{\partial h^{\mathsf{ort}}}{\partial x}(x) \end{pmatrix} \ = \ 2 \ , \quad \mathsf{Rank} \left(\frac{\partial h^{\mathsf{ort}}}{\partial x}(x) \right) \ = \ 1$$

A solution to these equations is

$$\begin{aligned} h_\alpha^{\mathsf{ort}}(x_1, x_2) \quad &= \quad \frac{k_\alpha(x_1, x_2)}{\sqrt{k_\alpha(x_1, x_2)^2 + k_\beta(x_1, x_2)^2}} \ , \\ h_\beta^{\mathsf{ort}}(x_1, x_2) \quad &= \quad \frac{k_\beta(x_1, x_2)}{\sqrt{k_\alpha(x_1, x_2)^2 + k_\beta(x_1, x_2)^2}} \end{aligned}$$

where $(k_1, k_2) : \mathbb{R}^2 \setminus \{0\} \to \mathbb{R}^2 \setminus \{0\}$ are C^s. The rank 1 condition is satisfied because of the normalization. To meet the rank 2 condition, we require, for all (x_1, x_2) in $\mathbb{R}^2 \setminus \{0\}$,

$$x_1 \left(k_\beta \frac{\partial k_\alpha}{\partial x_2} - k_\alpha \frac{\partial k_\beta}{\partial x_2} \right) - x_2 \left(k_\beta \frac{\partial k_\alpha}{\partial x_1} - k_\alpha \frac{\partial k_\beta}{\partial x_1} \right) \ \neq \ 0 \tag{42}$$

In this way the image of $\boldsymbol{h}^{\mathsf{ort}}$ is indeed the unit circle, not as an "abstract" manifold, but as an immersed submanifold of $\mathbb{R}^2$. Then, following Theorem 3.24 and according to (41), a metric satisfying Condition A3 on $\mathbb{R}^2 \setminus \{0\}$ is, with Q a C^s function with strictly positive values and $\begin{pmatrix} R_{\alpha\alpha} & R_{\alpha\beta} \\ R_{\alpha\beta} & R_{\beta\beta} \end{pmatrix}$ a C^s function with positive definite values,

$$\begin{aligned} P(x_1, x_2) &= \mathscr{D}'(x_1^2 + x_2^2)^2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} Q(\mathscr{D}(x_1^2 + x_2^2)) \begin{pmatrix} x_1 & x_2 \end{pmatrix} \\ &\qquad + \frac{\partial h^{\mathsf{ort}}}{\partial x}(x)^\top R(h^{\mathsf{ort}}(x)) \frac{\partial h^{\mathsf{ort}}}{\partial x}(x) \\ &= \mathscr{D}'(x_1^2 + x_2^2)^2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} Q(\mathscr{D}(x_1^2 + x_2^2)) \begin{pmatrix} x_1 & x_2 \end{pmatrix} \\ &\qquad + \mathscr{K}(x) \widetilde{R}(k_\alpha(x), k_\beta(x)) \mathscr{K}(x)^\top \end{aligned}$$

where $\mathscr{D}'$ is the derivative of $\mathscr{D}$ and

$$\mathcal{K}(x) = \begin{pmatrix} k_\beta(x)\dfrac{\partial k_\alpha}{\partial x_1}(x) - k_\alpha(x)\dfrac{\partial k_\beta}{\partial x_1}(x) \\ k_\beta(x)\dfrac{\partial k_\alpha}{\partial x_2}(x) - k_\alpha(x)\dfrac{\partial k_\beta}{\partial x_2}(x) \end{pmatrix}$$

$$\widetilde{R}(k_\alpha, k_\beta) = \frac{k_\alpha^2 R_{\beta\beta}(k_\alpha, k_\beta) - 2k_\alpha k_\beta R_{\alpha\beta}(k_\alpha, k_\beta)}{(k_\alpha^2 + k_\beta^2)^3} + \frac{k_\beta^2 R_{\alpha\alpha}(k_\alpha, k_\beta)}{(k_\alpha^2 + k_\beta^2)^3} .$$

For example, the particular choice

$$\mathcal{D}(s) = s \ , \ Q = 1 \ , \ k_\alpha = x_1 \ , \ k_\beta = x_2 \ , \ R = I$$

gives

$$P(x_1, x_2) = \begin{pmatrix} x_1^2 + \frac{x_2^2}{(x_1^2+x_2^2)^2} & x_1x_2 - \frac{x_1x_2}{(x_1^2+x_2^2)^2} \\ x_1x_2 - \frac{x_1x_2}{(x_1^2+x_2^2)^2} & x_2^2 + \frac{x_1^2}{(x_1^2+x_2^2)^2} \end{pmatrix}$$

IV. On Simultaneous Satisfaction of Conditions A2 and A3

We have observed that Conditions A2 and A3 are of completely different nature.

The next example shows both of these conditions may not always hold simultaneously.

In this section, we investigate ways, from a design standpoint, to guarantee that Condition A3 holds when Condition A2 is already satisfied, and vice versa.

Example 4.1: Consider the system

$$\dot{x}_1 = 2x_2 \quad , \qquad \dot{x}_2 = 1 - x_1 \quad , \qquad y = x_1^2 + x_2^2 \tag{43}$$

It is differentially observable of order five. Furthermore, there exists a globally convergent observer with linear dynamics estimating the output and its four derivatives. We have seen in Example 3.26 how to construct a metric satisfying Condition A3. Also there exists an expression P of the metric $\boldsymbol{P}$, which is polynomial of degree 2 in (x_1, x_2), satisfying Condition A2.

For this system (43), the observer (7) takes the form

$$\begin{pmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{pmatrix} = \begin{pmatrix} 2\hat{x}_2 \\ 1 - \hat{x}_1 \end{pmatrix} + kP(\hat{x}_1, \hat{x}_2)^{-1} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} (y - \hat{y}) . \tag{44}$$

With the first order variation formula, this observer leads to a strict decrease of the Riemannian distance induced by P, if, for any normalized geodesic γ^*, minimizing on $[s, \hat{s}]$,

$$\begin{aligned} &\frac{d\gamma^*}{ds}(\hat{s})^\top P(\gamma^*(\hat{s})) \times \\ &\times \left[\begin{pmatrix} 2\gamma_2^*(\hat{s}) \\ 1 - \gamma_1^*(\hat{s}) \end{pmatrix} - kP(\gamma^*(\hat{s}))^{-1}\gamma^*(\hat{s})(|\gamma^*(\hat{s})|^2 - |\gamma^*(s)|^2)\right] \\ &\qquad - \frac{d\gamma^*}{ds}(s)^\top P(\gamma^*(s)) \begin{pmatrix} 2\gamma_2^*(s) \\ 1 - \gamma_1^*(s) \end{pmatrix} < 0 \end{aligned} \tag{45}$$

The correction term contributes strictly to this decrease if Condition A3 holds, i.e.

$$|\gamma^*(\hat{s})| \neq |\gamma^*(s)| \ \Rightarrow \ \frac{d\gamma^*}{ds}(\hat{s})^\top \gamma^*(\hat{s})(|\gamma^*(\hat{s})|^2 - |\gamma^*(s)|^2) > 0 . \tag{46}$$

Unfortunately, there is no complete metric $\boldsymbol{P}$ such that the observer (44) satisfies (45) and (46) together. Indeed we know with [1, A.3.1.a] that (46) implies that, for any x different from the origin and any unit vector v tangent at x to the circle with radius $|x|$ and centered at the origin, the geodesic γ^* satisfying

$$\gamma^*(0) = x \ , \quad \frac{d\gamma^*}{ds}(0) = v \ ,$$

remains in that circle. Actually, there are two normalized geodesics issued from x, say γ_+^* and γ_-^*, satisfying

$$\frac{d\gamma_+^*}{ds}(0) = +v \quad , \qquad \frac{d\gamma_-^*}{ds}(0) = -v$$

which remain in the circle. The metric being complete by assumption, the orbits of these geodesics are the complete circle and there exist s_+ and s_- such that[10]

$$\gamma_+^*(s_+) = \gamma_-^*(s_-) \quad , \qquad \frac{d\gamma_+^*}{ds}(s_+) = -\frac{d\gamma_-^*}{ds}(s_-) \tag{47}$$

and γ_+^*, respectively γ_-^*, is minimizing on $[0, s_+]$, respectively $[0, s_-]$. But if (45) holds, we obtain

$$\frac{d\gamma_+^*}{ds}(s_+)^\top P(\gamma_+^*(s_+)) \begin{pmatrix} 2\gamma_{+2}^*(s_+) \\ 1 - \gamma_{+1}^*(s_+) \end{pmatrix} < v^\top P(x) \begin{pmatrix} 2x_2 \\ 1 - x_1 \end{pmatrix}$$

and

$$\frac{d\gamma_-^*}{ds}(s_-)^\top P(\gamma_-^*(s_-)) \begin{pmatrix} 2\gamma_{-2}^*(s_-) \\ 1 - \gamma_{-1}^*(s_-) \end{pmatrix} < -v^\top P(x) \begin{pmatrix} 2x_2 \\ 1 - x_1 \end{pmatrix}$$

With (47), these inequalities cannot hold simultaneously. On the other hand, it is possible to satisfy either Condition A2 or Condition A3 by properly choosing the metric.

A. Satisfying Condition A2 first

We know with [2, Proposition 2.4] that a Riemannian metric satisfying Condition A2 gives a locally convergent observer. This motivates starting with Condition A2.

In [2, (47) and Propositions 3.2 and 3.5] we have given procedures for obtaining metrics satisfying Condition A2. Then, with a metric constructed via such procedures, it remains to check if Condition A3 holds. Such a check consists of testing whether or not the second fundamental form of $\boldsymbol{h}$ is zero. For this test to be positive, we must have (21) which is satisfied if (29) holds. We know the latter condition can always be satisfied by modifying the given metric $\boldsymbol{P}$ into $\boldsymbol{P}_{mod}$ as given in (30). Fortunately, the satisfaction of Condition A2 is not affected by this modification, as the following result shows.

Proposition 4.2: Condition A2 holds for $\boldsymbol{P}$ if and only if Condition A2 holds for $\boldsymbol{P}_{mod}$.

Proof: The claim is a direct consequence of the identity

$$\frac{\partial}{\partial x}\left\{\left[\frac{\partial h}{\partial x}(x)v^{\mathsf{tan}}\right]^\top \left[Q(h(x)) - \left(\frac{\partial h}{\partial x}(x)P(x)^{-1}\frac{\partial h}{\partial x}(x)^\top\right)^{-1}\right] \times \left[\frac{\partial h}{\partial x}(x)v^{\mathsf{tan}}\right]\right\} = 0 \quad \forall v^{\mathsf{tan}} \in \mathfrak{D}^{\mathsf{tan}}(x) \ , \ \forall x \in \phi(\mathcal{M}) \tag{48}$$

being valid for any coordinate chart $(x, \mathcal{M}, \phi)$. ■

Example 4.3: (Systems that are strongly differentially observable of order n) In [2, §IV], we have seen that, when

[10] $\gamma_+^*(s_+) = \gamma_-^*(s_-)$ is a cut point of x. See [28, Cut Points ch. 10].

$p=1$ and

$$\boldsymbol{i}_n(\boldsymbol{x}) \;=\; \begin{pmatrix} \boldsymbol{h}(\boldsymbol{x}) \\ L_f\boldsymbol{h}(\boldsymbol{x}) \\ \vdots \\ L_{\boldsymbol{f}}^{n-1}\boldsymbol{h}(\boldsymbol{x}) \end{pmatrix}$$

is a diffeomorphism from some open set Ω to $\mathbb{R}^n$, the expression in some coordinate chart $(x,\mathcal{M},\phi)$ of a metric satisfying Condition A2 on Ω is

$$P(x)=\frac{\partial i_n}{\partial x}(x)^\top \bar{P}\frac{\partial i_n}{\partial x}(x) \ ,$$

where $\bar{P}$ is a symmetric positive definite matrix to be chosen (see [2, Lemma 4.2]). Actually, $\bar{x} \;=\; i_n(x)$ are other coordinates for $\boldsymbol{x}$ in which the expression of the metric $\boldsymbol{P}$ is simply the constant matrix $\bar{P}$. Moreover, the expression of $\boldsymbol{h}$ in the same particular coordinates is linear, i.e.

$$y \;=\; C\bar{x} \ ,$$

with the notation

$$C \;=\; \begin{pmatrix} 1 & 0 & \dots & 0 \end{pmatrix} \ .$$

Hence, the system belongs to the "Euclidean family" and the observer (7) takes the form (see [29])

$$\dot{\hat{x}} \;=\; f(\hat{x})-k_E(\hat{x})\frac{\partial i_n}{\partial x}(\hat{x})^{-1}\bar{P}^{-1}C^T(C\hat{x}-y) \ .$$

B. Satisfying Condition A3 First

For Condition A2 to hold, the Riemannian metric $\boldsymbol{P}$ must satisfy the inequality (4). Instead, for A3 to hold, according to Lemma 3.14, we must have at least the equalities (20). It may be easier to satisfy first the equalities and then the inequalities. Namely, instead of starting with a metric that satisfies Condition A2, we start with a metric given by Theorem 3.24, which is guaranteed to satisfy Condition A3. Then, it remains to define the degrees of freedom $\boldsymbol{h}^{\mathsf{ort}}$ and $\boldsymbol{R}$ involved in its construction so as to satisfy Condition A2.

In this context, the fact that the formula (41) for P is a sum implies that Condition A2 takes a particular form. Indeed, for any $\boldsymbol{x}_0$ in $\mathbb{R}^n$, for any coordinate charts $(x,\mathbb{R}^n,\phi)$ around $\boldsymbol{x}_0$ and $(y,\mathbb{R}^p,\chi)$ around $\boldsymbol{h}(\boldsymbol{x}_0)$ in $\mathbb{R}^p$, for any $(\xi,\mathcal{P},\omega)$ around $\boldsymbol{h}^{\mathsf{ort}}(\boldsymbol{x}_0)$ in $\boldsymbol{\Xi}$, for all vectors v satisfying

$$\sum_a \frac{\partial h_i}{\partial x_a}(x)v_a \;=\; 0 \ , \tag{49}$$

and with the definitions

$$g_\gamma(x)=\sum_c \frac{\partial h^{\mathsf{ort}}_\gamma}{\partial x_c}(x)f_c(x) \ , \ w_\alpha(x)=\sum_a \frac{\partial h^{\mathsf{ort}}_\alpha}{\partial x_a}(x)v_a \ , \tag{50}$$

the expression in (41) of the metric $\boldsymbol{P}$ gives

$$\begin{aligned} v^\top\mathcal{L}_fP(x)v \;=\; & \sum_{\alpha,\beta,\gamma} w_\alpha\frac{\partial R_{\alpha\beta}}{\partial \xi_\gamma}(h^{\mathsf{ort}}(x))\, g_\gamma(x)w_\beta \\ & +2\sum_{a,\beta,\gamma} v_a\frac{\partial g_\gamma}{\partial x_a}(x)\, R_{\gamma\beta}(h^{\mathsf{ort}}(x))w_\beta \ . \end{aligned}$$

By invoking the S-Lemma (see [30]), we obtain that Condition A2 is satisfied if, when (49) holds, we have

$$\begin{aligned} \sum_{\alpha,\beta}\Bigg(& w_\alpha(x)\left[\sum_\gamma\frac{\partial R_{\alpha\beta}}{\partial\xi_\gamma}(h^{\mathsf{ort}}(x))g_\gamma(x)\right]w_\beta(x) \\ & +2\left[\sum_a v_a\frac{\partial g_\alpha}{\partial x_a}(x)\right]R_{\alpha\beta}(h^{\mathsf{ort}}(x))w_\beta(x)\Bigg) \\ & \leq\; -q\sum_{\alpha,\beta}w_\alpha(x)R_{\alpha\beta}(h^{\mathsf{ort}}(x))w_\beta(x) \end{aligned} \tag{51}$$

for some strictly positive q.

With the above, we have reduced the design of the observer (7) to the problem of finding functions h^{ort}, of rank $n-p$ and such that (h,h^{ort}) has rank n, and R with positive definite values, satisfying the inequality above.

Example 4.4 (Systems of dimension two): We consider a general system written as

$$\dot{y} \;=\; f_y(y,z) \quad , \qquad \dot{z} \;=\; f_z(y,z) \tag{52}$$

with $n=2$ and $p=1$. It follows from Theorem 3.24, that Condition A3 is satisfied if P is in the form (see (41))

$$P(y,z)=\begin{pmatrix} 1+\frac{\partial h^{\mathsf{ort}}}{\partial y}(y,z)^2 & \frac{\partial h^{\mathsf{ort}}}{\partial y}(y,z)\frac{\partial h^{\mathsf{ort}}}{\partial z}(y,z) \\ \frac{\partial h^{\mathsf{ort}}}{\partial y}(y,z)\frac{\partial h^{\mathsf{ort}}}{\partial z}(y,z) & \frac{\partial h^{\mathsf{ort}}}{\partial z}(y,z)^2 \end{pmatrix} \ ,$$

where h^{ort} is any C^3 function with $\frac{\partial h^{\mathsf{ort}}}{\partial z}(y,z)$ strictly positive for all (y,z). In this case, we choose

$$\wp(y_1,y_2) \;=\; |y_1-y_2|^2$$

and Condition A2 holds if we have

$$\frac{\partial}{\partial z}\left\{\frac{\partial h^{\mathsf{ort}}}{\partial y}(y,z)f_y(y,z)+\frac{\partial h^{\mathsf{ort}}}{\partial z}(y,z)f_z(y,z)\right\}\leq -\frac{\partial h^{\mathsf{ort}}}{\partial z}(y,z)^2$$

In this case, the observer (7) is

$$\begin{aligned} \dot{\hat{y}} \;&=\; f_y(\hat{y},\hat{z})-k_E(\hat{y},\hat{z})(\hat{y}-y) \ , \\ \dot{\hat{z}} \;&=\; f_z(\hat{y},\hat{z})+k_E(\hat{y},\hat{z})\frac{\frac{\partial h^{\mathsf{ort}}}{\partial y}(y,z)}{\frac{\partial h^{\mathsf{ort}}}{\partial z}(y,z)}(\hat{y}-y) \ . \end{aligned}$$

Example 4.5: We consider again the harmonic oscillator with unknown frequency in (24), the state of which evolves in the invariant set Ω_ε defined in (25). We have seen in Example 3.15 that the metrics considered thus far satisfying Condition A2 do not satisfy Condition A3. Following the observations at the beginning of this section, we proceed by constructing P so that Condition A3 holds, and then assess the satisfaction of Condition A2.

Following Theorem 3.24, the level sets of the output function being diffeomorphic to $\mathbb{R}^2$, we choose $\boldsymbol{\Xi}$ as $\mathbb{R}^2$. Then, following Theorem 3.24, a metric satisfying Condition A3 is, with the notation $\frac{\partial}{\partial x}=\begin{pmatrix}\frac{\partial}{\partial y} & \frac{\partial}{\partial z_\alpha} & \frac{\partial}{\partial z_\beta}\end{pmatrix}$,

$$\begin{aligned} P(y,z_\alpha,z_\beta) \;=\; & \mathscr{D}'(y)^2\begin{pmatrix}1\\0\\0\end{pmatrix}Q(\mathscr{D}(y))\begin{pmatrix}1 & 0 & 0\end{pmatrix} \\ & +\frac{\partial\begin{pmatrix}h^{\mathsf{ort}}_\alpha\\ h^{\mathsf{ort}}_\beta\end{pmatrix}}{\partial x}(y,z_\alpha,z_\beta)^\top R(h^{\mathsf{ort}}(y,z_\alpha,z_\beta))\frac{\partial\begin{pmatrix}h^{\mathsf{ort}}_\alpha\\ h^{\mathsf{ort}}_\beta\end{pmatrix}}{\partial x}(y,z_\alpha,z_\beta) \end{aligned} \tag{53}$$

where it remains to choose

- Q as a C^s function with strictly positive values,
- $\mathscr{D}$ and h^{ort} such that $(y,z)\mapsto(\mathscr{D}(y),h^{\mathsf{ort}}(y,z))$ is a C^s diffeomorphism on Ω_ε,

– and R as a C^s function with positive definite values

to satisfy Condition A2 or its sufficient condition (51). To help in this task, we remind the reader of the findings in [2, Example 2.2]. In that example we show that, with $(\xi,\mathcal{P},\omega)$ as a global coordinate chart for $\Xi=\mathbb{R}^2$, the arrival set of $\boldsymbol{h}^{\mathsf{ort}}$, and the choices

$$\begin{gathered}\mathcal{D}(y)=y\ ,\quad Q(y)=c\ ,\quad h^{\mathsf{ort}}_\alpha(y,z)=z_\alpha-y\ ,\\ h^{\mathsf{ort}}_\beta(y,z)=z_\beta+\tfrac{1}{2}y^2\ ,\quad R(\xi)\ =\ I_2\ ,\end{gathered}$$

where c is a strictly positive real number, the expression P of the metric obtained in (53), namely,

$$P(y,z_\alpha,z_\beta)\ =\ \begin{pmatrix}1\\0\\0\end{pmatrix}c\begin{pmatrix}1&0&0\end{pmatrix}+\begin{pmatrix}-1&y\\1&0\\0&1\end{pmatrix}\begin{pmatrix}-1&1&0\\y&0&1\end{pmatrix}$$

is such that Condition A2 is satisfied but not strictly – namely, it only certifies weak differential detectability.

From this point we proceed with a "deformation" of the metric above to meet both conditions. We choose

$$\begin{gathered}h^{\mathsf{ort}}_\alpha(y,z)=\xi_\alpha=z_\alpha-y\ ,\quad h^{\mathsf{ort}}_\beta(y,z)=\xi_\beta=z_\beta+\tfrac{1}{2}y^2+abyz_\alpha\ ,\\ R(\xi)=\begin{pmatrix}1&0\\0&1+a\xi_\alpha^2\end{pmatrix}\ ,\end{gathered}$$

where a and b are strictly positive real numbers to be chosen, with, a priori, a being small. We express the inequality (51) with the coordinates (y,z), restricted to the set Ω_ε. Since (50) reads

$$\begin{aligned}g_\alpha(y,z_\alpha,z_\beta)\ &=\ -yz_\beta-z_\alpha\ =\ -yz_\beta-\xi_\alpha-y\ ,\\ g_\beta(y,z_\alpha,z_\beta)\ &=\ yz_\alpha+abz_\alpha^2-aby^2z_\beta\ ,\end{aligned}$$

$$\begin{gathered}w_\alpha=v_\alpha-v_y\ ,\quad w_\beta=v_\beta+yv_y+abz_\alpha v_y+abyv_\alpha\ ,\\ \begin{pmatrix}v_\alpha&v_\beta\end{pmatrix}\ =\ \begin{pmatrix}w_\alpha&w_\beta\end{pmatrix}\begin{pmatrix}1&-aby\\0&1\end{pmatrix}\ ,\end{gathered}$$

inequality (51) is

$$\begin{aligned}&w_\beta^2 2a\xi_\alpha g_\alpha+2\begin{pmatrix}w_\alpha&w_\beta\end{pmatrix}\begin{pmatrix}1&-aby\\0&1\end{pmatrix}\begin{pmatrix}-1&y+2abz_\alpha\\-y&-aby^2\end{pmatrix}\times\\ &\qquad\times\begin{pmatrix}1&0\\0&1+a\xi_\alpha^2\end{pmatrix}\begin{pmatrix}w_\alpha\\w_\beta\end{pmatrix}\\ &\qquad\qquad\leq\ -q\begin{pmatrix}w_\alpha&w_\beta\end{pmatrix}\begin{pmatrix}1&0\\0&1+a\xi_\alpha^2\end{pmatrix}\begin{pmatrix}w_\alpha\\w_\beta\end{pmatrix}\end{aligned}$$

for some strictly positive q. This inequality can be rewritten

$$\begin{aligned}&-w_\beta^2\left[2a\xi_\alpha(yz_\beta+\xi_\alpha+y)+2aby^2(1+a\xi_\alpha^2)-q(1+a\xi_\alpha^2)\right]\\ &+\quad 2w_\alpha w_\beta a\left[2b(\xi_\alpha+y)(1+a\xi_\alpha^2)+y\xi_\alpha^2+ab^2y^3(1+a\xi_\alpha^2)\right]\\ &\qquad\qquad-\ w_\alpha^2\left[2(1-aby^2)-q\right]\ \leq\ 0\end{aligned}$$

It is satisfied if we have $2(1-aby^2)-q\ >\ 0$, and

$$\begin{aligned}&4a^2\left[2b(\xi_\alpha+y)(1+a\xi_\alpha^2)+y\xi_\alpha^2+ab^2y^3(1+a\xi_\alpha^2)\right]^2\ <\\ &4\left[2a\xi_\alpha(yz_\beta+\xi_\alpha+y)+2aby^2(1+a\xi_\alpha^2)-q(1+a\xi_\alpha^2)\right]\times\\ &\qquad\qquad\times\left[2(1-aby^2)-q\right]\ ,\end{aligned}$$

for all (y,z) in Ω_ε, and therefore satisfying

$$\frac{4}{\varepsilon^2}>y^2+\xi_\alpha^2>\frac{\varepsilon^2}{4}\quad,\qquad\varepsilon<z_\beta<\frac{1}{\varepsilon}\ .\tag{54}$$

We have

$$\begin{aligned}&2a\xi_\alpha(yz_\beta+\xi_\alpha+y)+2aby^2(1+a\xi_\alpha^2)-q(1+a\xi_\alpha^2)\\ &\quad=\ a\left[2\xi_\alpha y\,(z_\beta+1)+(2-q)\xi_\alpha^2+2by^2\right]+2a^2by^2\xi_\alpha^2-q\ .\end{aligned}$$

So, by choosing b large enough to satisfy $2\left(\frac{1}{\varepsilon}+1\right)^2\leq(2-q)b$, we obtain successively

$$\begin{aligned}&2\xi_\alpha y\,(z_\beta+1)+(2-q)\xi_\alpha^2+2by^2\ \geq\ \frac{1}{2}\left[(2-q)\xi_\alpha^2+2by^2\right]\\ &2a\xi_\alpha(yz_\beta+\xi_\alpha+y)+2aby^2(1+a\xi_\alpha^2)-q(1+a\xi_\alpha^2)\\ &\qquad\qquad\geq\ \tfrac{a}{2}\min\{(2-q),2b\}\tfrac{\varepsilon^2}{4}+2a^2by^2\xi_\alpha^2-q\ .\end{aligned}$$

Also, using (54), we have $2(1-aby^2)-q\ \geq\ 2-ab\frac{8}{\varepsilon^2}-q$, and

$$\begin{aligned}&2b(\xi_\alpha+y)(1+a\xi_\alpha^2)+y\xi_\alpha^2+ab^2y^3(1+a\xi_\alpha^2)\\ &\qquad\leq 2b\frac{4}{\varepsilon}\left(1+a\frac{4}{\varepsilon^2}\right)+\frac{8}{\varepsilon^3}+ab^2\frac{8}{\varepsilon^3}\left(1+a\frac{4}{\varepsilon^2}\right)\\ &\qquad\leq\tfrac{8}{\varepsilon^3}\left(b+1+ab^2\right)\left(1+a\tfrac{4}{\varepsilon^2}\right)\ .\end{aligned}$$

Then, a sufficient condition for Condition A2 to hold is

$$\left.\begin{aligned}&2\left(\frac{1}{\varepsilon}+1\right)^2\leq(2-q)b\quad,\qquad 2-ab\frac{8}{\varepsilon^2}-q\ >\ 0\\ &\frac{64a^2}{\varepsilon^6}\left(b+1+ab^2\right)^2\left(1+a\frac{4}{\varepsilon^2}\right)^2\ <\\ &\quad\left(\frac{a}{2}\min\{(2-q),b\}\frac{\varepsilon^2}{4}-q\right)\left(2-ab\frac{8}{\varepsilon^2}-q\right)\end{aligned}\right\}\tag{55}$$

With b fixed as $b=4\left(\frac{1}{\varepsilon}+1\right)^2$, since the following inequality is satisfied when $a=0$, there exists a strictly positive real number $\bar{a}$ such that, for all a in $[0,\bar{a})$, we have

$$\frac{64a}{\varepsilon^6}\left(b+1+ab^2\right)^2\left(1+a\frac{4}{\varepsilon^2}\right)^2\ <\ \frac{\varepsilon^2}{8}\left(2-ab\frac{8}{\varepsilon^2}\right)$$

We fix a in $(0,\bar{a})$. By continuity, there exists q satisfying (55).

We have established the existence of a triplet (a,b,q) such that Conditions A2 and A3 are satisfied on Ω_ε by the metric $\boldsymbol{P}$, the expression of which, with the coordinate (y,z), is

$$\begin{aligned}&P(y,z_\alpha,z_\beta)\ =\ \begin{pmatrix}1\\0\\0\end{pmatrix}c\begin{pmatrix}1&0&0\end{pmatrix}\\ &+\begin{pmatrix}-1&y+abz_\alpha\\1&aby\\0&1\end{pmatrix}\begin{pmatrix}1&0\\0&1+a(z_\alpha-y)^2\end{pmatrix}\begin{pmatrix}-1&1&0\\y+abz_\alpha&aby&1\end{pmatrix}\ ,\end{aligned}$$

The observer (7) for the harmonic oscillator with unknown frequency is

$$\begin{aligned}\overbrace{\begin{pmatrix}\hat{y}\\\hat{z}_\alpha\\\hat{z}_\beta\end{pmatrix}}^{\cdot}\ &=\ \begin{pmatrix}\hat{z}_\alpha\\-\hat{y}\hat{z}_\beta\\0\end{pmatrix}\\ &\qquad-\frac{k_E(\hat{y},\hat{z}_\alpha,\hat{z}_\beta)}{c}\begin{pmatrix}1\\1\\-\hat{y}-ab(\hat{z}_\alpha+\hat{y})\end{pmatrix}(\hat{y}-y)\ .\end{aligned}$$

As a final remark, we note that the expression of the metric with the coordinates (y, ξ) is (by definition)

$$\bar{P}(y, \xi_\alpha, \xi_\beta) \;=\; \begin{pmatrix} c & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 + a\xi_\alpha^2 \end{pmatrix} \;.$$

All the corresponding Christoffel symbols are zero, except $\bar{\Gamma}^\alpha_{\beta\beta} = -\frac{\xi_\alpha}{2}$, $\bar{\Gamma}^\beta_{\alpha\beta} = -\frac{\xi_\alpha}{2(a\xi_\alpha^2+1)}$. It follows that the component $\mathfrak{R}^\alpha_{\alpha\beta\beta} = \frac{\partial \bar{\Gamma}^\alpha_{\beta\beta}}{\partial \xi_\alpha} - \bar{\Gamma}^\alpha_{\beta\beta}\bar{\Gamma}^\beta_{\alpha\beta} = -2a - \frac{\xi_\alpha^2}{4(a\xi_\alpha^2+1)}$ of the Riemann curvature tensor is not zero. So there is no coordinates for which the expression of the metric is Euclidean.

V. Conclusions

A. Conclusions of this paper

In [1], we have established that an observer, the correction term of which is based on a gradient of a "gap" function between measured output and estimated output, converges when a strong differential detectability condition (Condition A2) holds and when the output function is geodesic monotone (Condition A3) holds.

In [2], we have shown how, for a given system for which all the variational systems are reconstructible, we can design a metric satisfying Condition A2.

In this paper, we have shown that Condition A3 is strongly linked to the nullity of the second fundamental form of the output function. Actually these two properties are equivalent when the dimension p of the $\boldsymbol{y}$-manifold $\mathbb{R}^p$ is 1. When p is larger than 1, the latter implies always the former but, for the converse, we need the extra assumption that the orthogonal distribution is involutive. With this study we have been able to propose a design tool for obtaining a metric satisfying Condition A3. This tool, described in Theorem 3.24, is systematic in the sense that it does not rely on some equation or inequality to solve. It is a formula (see (41)) giving the expression of the metric in any given coordinate chart.

In our study, we have left open many problems. We point out only two of them:

1) *Is the involutivity of the orthogonal distribution necessary for Condition A3 to hold?*
   In case of a positive answer, the nullity of the second fundamental form of the output function would be equivalent to Condition A3 and, more interestingly, our design procedure would construct all the metrics satisfying Condition A3.
2) *Under which conditions does a nonlinear counterpart of Heymann's Lemma [26], [27] holds?*
   More precisely, the problem to solve is the following: *Assume system (1) satisfies Condition A2 and the level sets of the given output function $\boldsymbol{h}$ are totally geodesic. Do there exist functions $\mathfrak{f} : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^n$ and $\mathfrak{h} : \mathbb{R}^p \to \mathbb{R}$ satisfying $\mathfrak{f}(\boldsymbol{x}, \boldsymbol{h}(\boldsymbol{x}), \boldsymbol{h}(\boldsymbol{x})) = \boldsymbol{f}(\boldsymbol{x})$ and such that Conditions A2 and A3 hold for the modified system $\dot{\boldsymbol{x}} = \mathfrak{f}(\boldsymbol{x}, \boldsymbol{h}(\boldsymbol{x}), y(t))$, $\boldsymbol{y}_{mod} = \mathfrak{h}(\boldsymbol{h}(\boldsymbol{x}))$ where $t \mapsto \boldsymbol{y}(t)$ is considered as an (known) input function.*
   This modified system is obtained with output injection in the function $\mathfrak{f}$ and the reduction to 1 of the number of outputs via the function $\mathfrak{h}$. In case of a positive answer, the extra conditions of involutivity of the orthogonal distribution would be unnecessary.

B. Conclusions on our study of convergence of observers with a Riemannian Metric

As written in the introduction, the three papers ([1], [2], and this one) formulate sufficient conditions that are as close as possible to necessary conditions for the design of

1) an observer; namely, a dynamical system with a state evolving in the same space as the (true) state of the given system,
2) with convergence established by the decrease of a Riemannian distance between the estimated and the true state,
3) with the set (3) of points where the estimated state is equal to the true state being forward invariant,
4) and with an infinite gain margin.

A key motivation for this effort is assessing if contraction theory is a fundamental tool for analyzing observer convergence. We have put this into practice by studying the effect of the flow of the system-observer pair on a Riemannian distance between the estimate generated by the observer and the system state, knowing that the Euclidean case, with therefore appropriately chosen coordinates, had been dealt with already (see, e.g., [31], [32]).

From our study, we conclude that the expected condition of differential detectability (Condition A2), related to a contraction property in the tangent space to the level sets of the output function, is not by itself sufficient to obtain a (at most local) convergent observer. It is also required for these level sets to be at least totally geodesic. Our results indicate that this latter condition, related to convexity, is likely a consequence of the fact that an observer is, to some extent, searching for the global minimum of a cost function that depends on the output error. To the best of our knowledge, this condition has not been proposed before, the reason being perhaps that, when the coordinates are such that the output function is linear and the metric is Euclidean; i.e., when the system is in the "Euclidean family" of Remark 3.2, the said condition is automatically satisfied.

Our study remained at a theoretical level and has not addressed real-world applications, mainly due to the difficulty of satisfying Conditions A2 and A3 simultaneously. These conditions are of completely different nature and, for the time being, we do not know of a systematic way for having both satisfied for general systems. In one way or the other, other methods assume the knowledge of a family of metrics satisfying the two properties, e.g., there are coordinates for which the output function is linear and the pair (f, h) is differentially detectable with respect to a constant metric. Nevertheless, we have reduced the problem of simultaneously satisfying Conditions A2 and A3 to finding functions h^{ort}, of rank $n - p$ and such that (h, h^{ort}) has rank n, and R with positive definite values, satisfying (51) for some $q > 0$. Also, fortunately, as shown in [2], Condition A2 only is already sufficient to obtain a locally convergent observer.

Ultimately, rather than advocating for a new observer design technique, our work presents substantiated arguments about the advantages and disadvantages/limitations of designing observers based on contracting a Riemannian distance. In these regards we appropriate the words of D.C. Lewis who, after proposing in [33] contraction by flows of a Finsler distance (which includes as a special case a Riemannian distance) to study the dependence of solutions of dynamical systems on the initial conditions, wrote in [34]

> *It thus was felt that the method [= contraction] as so far developed [in [33]] was intrinsically too crude to yield the desired results in applications.*

REFERENCES

[1] R. G. Sanfelice and L. Praly. Convergence of nonlinear observers on $\mathbb{R}^n$ with a Riemannian metric (Part I). IEEE Transactions on Automatic Control, 57(7):1709–1722, July 2012.
[2] R. G. Sanfelice and L. Praly. Convergence of nonlinear observers on $\mathbb{R}^n$ with a riemannian metric (part ii). IEEE Transactions on Automatic Control, 61(10):2848–2860, October 2016.
[3] D. Luenberger. Observing the state of a linear system. IEEE Transactions on Military Electronics, 8:74–80, 1964.
[4] R.E Kalman and R.S Bucy. New results in linear filtering and prediction theory. Journal of Basic Engineering, 108:83–95, 1961.
[5] H. Khalil and L. Praly. High-gain observers in nonlinear feedback. Int. J. Robust. Nonlinear Control, 24:991–1157, 2013.
[6] N. Aghannan and P. Rouchon. An intrinsic observer for a class of Lagrangian systems. IEEE Trans. Automatic Control, 48(6):936–945, 2003.
[7] S. Bonnabel. Observateurs Asymptotiques Invariants. PhD thesis, Spécialité Mathématiques et Automatique. École des Mines de Paris, 2007.
[8] S. Bonnabel. A simple intrinsic reduced-observer for geodesic flow. IEEE Transactions on Automatic Control, 55(9):2186–2191, 2010.
[9] R. G. Sanfelice and L. Praly. Convergence of nonlinear observers on $\mathbb{R}^n$ with a Riemannian metric (Part I) (revised). In arXiv: http://arxiv.org/abs/1412.6730, 2014.
[10] T. Rapcsák. Smooth Nonlinear Optimization in $\mathbb{R}^n$. Kluwer Academic, 1997.
[11] C. Udriste. Convex Functions and Optimization Methods on Riemannian Manifolds. Kluwer Academic Publishers, 1994.
[12] R. G. Sanfelice and L. Praly. Convergence of nonlinear observers on $\mathbb{R}^n$ with a Riemannian metric (Part III) (report). In arXiv: https://arxiv.org/abs/2102.08340, 2021.
[13] J. Vilms. Totally geodesic maps. Journal of differential geometry, 4(1):73–79, 1970.
[14] T. Nore. Second fundamental form of a map. Annali di Matematica pura ed Applicata, 146(1):281–310, 1986.
[15] D. Kupeli E. García-Río. Semi-Riemannian Maps and Their Applications. 1999.
[16] J. M. Lee. Introduction to Smooth Manifolds. Springer, 2013.
[17] T. Sakai. Riemannian geometry, volume 149. Translation of Mathematical monographs, American Mathematical Soc., 1996.
[18] M. Spivak. (A Comprehensive Introduction to) Differential geometry, volume 2. Publish or Perish, Inc., 1979. 2nd Edition.
[19] G. Isac and S. Z. Németh. Scalar and asymptotic scalar derivatives : theory and applications. Springer, 4th edition, 2008.
[20] N. Innami. Splitting theorems of Riemannian manifolds. Compositio Mathematica, 47:237–247, 1982.
[21] J. Eells and J.H. Sampson. Harmonic mappings of Riemannian manifolds. Amer. J. Math., 86(1):109–160, 1964.
[22] A. P. Aguiar A. Saccon, J. Hauser. Optimal control on non-compact lie groups: A projection operator approach. In Proceedings of the 49th IEEE Conference on Decision and Control, pages 7111–7116, 2010.
[23] B. O'Neill. Semi-Riemannian geometry. Academic Press Inc., 1983.
[24] B. O'Neill. The fundamental equations of a submersion. The Michigan Mathematical Journal, 13(4):459–469, 1966.
[25] L.P. Eisenhart. Riemannian Geometry. Princeton University Press, 1925.
[26] M. Heymann and W. Wonham. Comments" on pole assignment in multi-input controllable linear systems". IEEE Transactions on Automatic Control, 13(6):748–749, 1968.
[27] M. L. J. Hautus. A simple proof of Heymann's lemma. IEEE Transactions on Automatic Control, 22(5):885–886, 1977.
[28] J. M. Lee. Introduction to Riemannian Manifolds. Springer, second edition edition, 2018.
[29] F. Deza, E. Busvelle, J. P. Gauthier, and D. Rakotopara. High gain estimation for nonlinear systems. Systems & Control Letters, 18:295–299, April 1992.
[30] T. Terlaky I. Pólik. A survey of the s-lemma. SIAM Review, 49(3):371–418, september 2007.
[31] J. Tsinias. Further results on the observer design problem. Syst. Contr. Lett., 14:411–418, 1990.
[32] L. Praly. On observers with state independent error Lyapunov function. In Proceedings of the 5th IFAC Symposium "Nonlinear Control Systems" (NOLCOS'01), 2001.
[33] D. C. Lewis. Metric properties of differential equations. American Journal of Mathematics, 71:294–312, 1949.
[34] D. C. Lewis. Differential equations referred to a variable metric. American Journal of Mathematics, 73(1):48–58, 1951.
[35] A. Figalli and C. Villani. Optimal transport and curvature. In Nonlinear PDEs and Applications, pages 171–217. Springer, 2011.
[36] H. Rund D. Lovelock. Tensors, differential forms, and variational principles. Dover, 1989.

APPENDIX

A1. GLOSSARY

We give here a complement to the glossary in [2]. Also, we recommend reading [35, Section 1].

1) A Riemannian metric $\boldsymbol{P}$ is a symmetric covariant 2-tensor with positive definite values.
2) The length of a C^1 path $\boldsymbol{\gamma}$ between points $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$ is defined as
$$L(\boldsymbol{\gamma})\Big|_{s_a}^{s_b} = \int_{s_a}^{s_b} \sqrt{\frac{d\boldsymbol{\gamma}}{ds}(s)^\top \boldsymbol{P}(\boldsymbol{\gamma}(s))\frac{d\boldsymbol{\gamma}}{ds}(s)}\, ds,$$
where $\boldsymbol{\gamma}(s_a) = \boldsymbol{x}_a$ and $\boldsymbol{\gamma}(s_b) = \boldsymbol{x}_b$.
3) The Riemannian distance $d(\boldsymbol{x}_a, \boldsymbol{x}_b)$ is the minimum of $L(\boldsymbol{\gamma})\Big|_{s_a}^{s_b}$ among all possible piecewise C^1 paths $\boldsymbol{\gamma}$ between $\boldsymbol{x}_a$ and $\boldsymbol{x}_b$. A minimizer giving the distance is called a minimizing geodesic and is denoted $\boldsymbol{\gamma}^*$.
4) A Riemannian metric $\boldsymbol{P}$ is said complete when every geodesic can be maximally extended to $\mathbb{R}$.
5) $\boldsymbol{d^2h}$ denotes the second differential the expression of which, with the coordinates x, is $\frac{\partial^2 h_i}{\partial x_a \partial x_b}(x)$.
6) The second differential with respect to the first argument of the function $(\boldsymbol{y}_1, \boldsymbol{y}_2) \mapsto \wp(\boldsymbol{y}_1, \boldsymbol{y}_2)$ is denoted $\boldsymbol{d}_1^2\wp$.
7) $\boldsymbol{H_P h}$ denotes the (Riemannian) Hessian of $\boldsymbol{h}$. It is a 2-covariant to 1-contravariant tensor defined as
$$\boldsymbol{H_P h}(\boldsymbol{x}) = \frac{1}{2}\mathcal{L}_{\boldsymbol{g_P h}}\boldsymbol{P}(\boldsymbol{x}) . \tag{56}$$
Its expression with coordinates x is
$$(H_P h(x))_{ab} = \frac{\partial^2 h}{\partial x_a \partial x_b}(x) - \sum_c \Gamma_{ab}^c(x)\frac{\partial h}{\partial x_c}(x) .$$
Given a geodesic $\boldsymbol{\gamma}$, we have (see [23, Exercise 3.16])
$$\frac{d^2}{ds^2}\Big\{\boldsymbol{h}(\boldsymbol{\gamma}(s))\Big\} = \frac{d\boldsymbol{\gamma}}{ds}(s)^\top \boldsymbol{H_P h}(\boldsymbol{\gamma}(s))\frac{d\boldsymbol{\gamma}}{ds}(s) \tag{57}$$

A2. PROOF OF LEMMA 3.8

Let $\bar\Gamma$ and $\bar\Delta$ be the expressions of the Christoffel symbols in the new coordinates. From [36, (3.5.22)], we get

$$\sum_{d,e}\frac{\partial \mathcal{C}_d}{\partial x_a}\bar\Gamma_{de}^c\frac{\partial \mathcal{C}_e}{\partial x_b} = \sum_d \frac{\partial \mathcal{C}_c}{\partial x_d}\Gamma_{ab}^d - \frac{\partial^2 \mathcal{C}_c}{\partial x_a \partial x_b} \tag{58}$$
$$\sum_{l,m}\frac{\partial \mathcal{D}_l}{\partial y_i}\bar\Delta_{lm}^k\frac{\partial \mathcal{D}_m}{\partial y_j} = \sum_l \frac{\partial \mathcal{D}_k}{\partial y_l}\Delta_{ij}^l - \frac{\partial^2 \mathcal{D}_k}{\partial y_i \partial y_j}$$

We also have $\bar h(\mathcal{C}(x)) = \mathcal{D}(h(x))$,

$$\sum_c \frac{\partial \bar h}{\partial \bar x_c}(\mathcal{C}(x))\frac{\partial \mathcal{C}_c}{\partial x_a}(x) = \sum_m \frac{\partial \mathcal{D}}{\partial y_m}(h(x))\frac{\partial h_m}{\partial x_a}(x) ,$$
$$\frac{\partial \bar h}{\partial \bar x_c}(\mathcal{C}(x)) = \sum_{e,m}\frac{\partial \mathcal{D}}{\partial y_m}(h(x))\frac{\partial h_m}{\partial x_e}(x)\left[\frac{\partial \mathcal{C}}{\partial x}(x)^{-1}\right]_{ec} ,$$

and

$$\sum_{c,d} \frac{\partial^2 \bar{h}}{\partial \bar{x}_c \partial \bar{x}_d}(\mathcal{C}(x)) \frac{\partial \mathcal{C}_c}{\partial x_a}(x) \frac{\partial \mathcal{C}_d}{\partial x_b}(x) + \sum_c \frac{\partial \bar{h}}{\partial \bar{x}_c}(\mathcal{C}(x)) \frac{\partial^2 \mathcal{C}_c}{\partial x_a \partial x_b}(x)$$
$$= \sum_{i,j} \frac{\partial^2 \mathcal{D}}{\partial y_i \partial y_j}(h(x)) \frac{\partial h_i}{\partial x_a}(x) \frac{\partial h_j}{\partial x_b}(x) + \sum_i \frac{\partial \mathcal{D}}{\partial y_i}(h(x)) \frac{\partial^2 h_i}{\partial x_a \partial x_b}(x) \ .$$

Using these expressions, we obtain

$$I\!I_P \overline{h}^k_{cd}(\bar{x}) = \frac{\partial^2 \bar{h}_k}{\partial \bar{x}_c \partial \bar{x}_d}(\bar{x}) - \sum_e \bar{\Gamma}^e_{cd}(x) \frac{\partial \bar{h}_k}{\partial \bar{x}_e}(\bar{x}) + \sum_{l,m} \bar{\Delta}^k_{lm}(\bar{h}(\bar{x})) \frac{\partial \bar{h}_l}{\partial \bar{x}_c}(\bar{x}) \frac{\partial \bar{h}_m}{\partial \bar{x}_d}(\bar{x}) \ ,$$

Substituting it in (58), we obtain

$$\begin{aligned}
&\sum_{c,d} \frac{\partial \mathcal{C}_c}{\partial x_a}(x) \frac{\partial \mathcal{C}_d}{\partial x_b}(x) I\!I_P \overline{h}^k_{cd}(\bar{x}) \\
&= \sum_{c,d} \frac{\partial \mathcal{C}_c}{\partial x_a}(x) \frac{\partial \mathcal{C}_d}{\partial x_b}(x) \frac{\partial^2 \bar{h}_k}{\partial \bar{x}_c \partial \bar{x}_d}(\bar{x}) \\
&\quad - \sum_{c,d} \frac{\partial \mathcal{C}_c}{\partial x_a}(x) \frac{\partial \mathcal{C}_d}{\partial x_b}(x) \sum_e \bar{\Gamma}^e_{cd}(x) \frac{\partial \bar{h}_k}{\partial \bar{x}_e}(\bar{x}) \\
&\quad + \sum_{c,d} \frac{\partial \mathcal{C}_c}{\partial x_a}(x) \frac{\partial \mathcal{C}_d}{\partial x_b}(x) \sum_{l,m} \bar{\Delta}^k_{lm}(\bar{h}(\bar{x})) \frac{\partial \bar{h}_l}{\partial \bar{x}_c}(\bar{x}) \frac{\partial \bar{h}_m}{\partial \bar{x}_d}(\bar{x}) \\
&= \sum_{i,j} \frac{\partial^2 \mathcal{D}_k}{\partial y_i \partial y_j}(h(x)) \frac{\partial h_i}{\partial x_a}(x) \frac{\partial h_j}{\partial x_b}(x) \\
&\quad + \sum_i \frac{\partial \mathcal{D}_k}{\partial y_i}(h(x)) \frac{\partial^2 h_i}{\partial x_a \partial x_b}(x) \\
&\quad - \sum_{c,e,i} \frac{\partial \mathcal{D}_k}{\partial y_i}(h(x)) \frac{\partial h_i}{\partial x_e}(x) \left[\frac{\partial \mathcal{C}}{\partial x}(x)^{-1}\right]_{ec} \frac{\partial^2 \mathcal{C}_c}{\partial x_a \partial x_b}(x) \\
&\quad - \sum_c \left[\sum_d \frac{\partial \mathcal{C}_c}{\partial x_d} \Gamma^d_{ab} - \frac{\partial^2 \mathcal{C}_c}{\partial x_a \partial x_b}\right] \sum_{e,i} \frac{\partial \mathcal{D}_k}{\partial y_i}(h(x)) \frac{\partial h_i}{\partial x_e}(x) \times \left[\frac{\partial \mathcal{C}}{\partial x}(x)^{-1}\right]_{ec} \\
&\quad + \sum_{l,m} \bar{\Delta}^k_{lm}(\bar{h}(\bar{x})) \left[\sum_i \frac{\partial \mathcal{D}_l}{\partial y_i}(h(x)) \frac{\partial h_i}{\partial x_a}(x)\right] \times \left[\sum_j \frac{\partial \mathcal{D}_m}{\partial y_j}(h(x)) \frac{\partial h_j}{\partial x_b}(x)\right] \\
&= \sum_{i,j} \frac{\partial^2 \mathcal{D}_k}{\partial y_i \partial y_j}(h(x)) \frac{\partial h_i}{\partial x_a}(x) \frac{\partial h_j}{\partial x_b}(x) \\
&\quad + \sum_i \frac{\partial \mathcal{D}_k}{\partial y_i}(h(x)) \left[\frac{\partial^2 h_i}{\partial x_a \partial x_b}(x) - \sum_d \Gamma^d_{ab} \frac{\partial h_i}{\partial x_d}(x)\right] \\
&\quad + \sum_{i,j} \frac{\partial h_i}{\partial x_a}(x) \frac{\partial h_j}{\partial x_b}(x) \left[\sum_l \frac{\partial \mathcal{D}_k}{\partial y_l} \Delta^l_{ij} - \frac{\partial^2 \mathcal{D}_k}{\partial y_i \partial y_j}\right]
\end{aligned}$$

which reduces to

$$\sum_i \frac{\partial \mathcal{D}_k}{\partial y_i}(h(x)) \times \left[\frac{\partial^2 h_i}{\partial x_a \partial x_b}(x) - \sum_d \Gamma^d_{ab} \frac{\partial h_i}{\partial x_d}(x) + \sum_{l,m} \frac{\partial h_l}{\partial x_a}(x) \frac{\partial h_m}{\partial x_b}(x) \Delta^i_{lm}\right]$$

and establishes (19).

PLACE PHOTO HERE

Ricardo G. Sanfelice received the B.S. degree in Electronics Engineering from the Universidad de Mar del Plata, Buenos Aires, Argentina, in 2001, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of California, Santa Barbara, CA, USA, in 2004 and 2007, respectively. In 2007 and 2008, he held postdoctoral positions at the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology and at the Centre Automatique et Systèmes at the École des Mines de Paris. In 2009, he joined the faculty of the Department of Aerospace and Mechanical Engineering at the University of Arizona, Tucson, AZ, USA, where he was an Assistant Professor. In 2014, he joined the University of California, Santa Cruz, CA, USA, where he is currently Professor in the Department of Electrical and Computer Engineering. Prof. Sanfelice is the recipient of the 2013 SIAM Control and Systems Theory Prize, the National Science Foundation CAREER award, the Air Force Young Investigator Research Award, the 2010 IEEE Control Systems Magazine Outstanding Paper Award, and the 2020 Test-of-Time Award from the Hybrid Systems: Computation and Control Conference. His research interests are in modeling, stability, robust control, observer design, and simulation of nonlinear and hybrid systems with applications to power systems, aerospace, and biology.

PLACE PHOTO HERE

Laurent Praly graduated as an engineer from ÉcoleNationale Supérieure des Mines de Paris in 1976 and got his PhD in Automatic Control and Mathematics in 1988 from Université Paris IX Dauphine. After working in industry for three years, in 1980 he joined the Centre Automatique et Systèmes at École des Mines de Paris. From July 1984 to June 1985, he spent a sabbatical year as a visiting assistant professor in the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign. Since 1985 he has continued at the Centre Automatique et Systèmes where he served as director for two years. He has made several long term visits to various institutions (Institute for Mathematics and its Applications at the University of Minnesota, University of Sydney, University of Melbourne, Institut Mittag-Leffler, University of Bologna). His main interest is in feedback stabilization of controlled dynamical systems under various aspects – linear and nonlinear, dynamic, output, under constraints, with parametric or dynamic uncertainty, disturbance attenuation or rejection. On these topics he is contributing both on the theoretical aspect with many academic publications and the practical aspect with applications in power systems, mechanical systems, aerodynamical and space vehicles.