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

Ricardo G. Sanfelice and Laurent Praly

Abstract

We study how convergence of an observer whose state lives in a copy of the given system's space can be established using a Riemannian metric. We show that the existence of an observer guaranteeing the property that a Riemannian distance between system and observer solutions is nonincreasing implies that the Lie derivative of the Riemannian metric along the system vector field is conditionally negative. Moreover, we establish that the existence of this metric is related to the observability of the system's linearization along its solutions. Moreover, if the observer has an infinite gain margin then the level sets of the output function are geodesically convex. Conversely, we establish that, if a complete Riemannian metric has a Lie derivative along the system vector field that is conditionally negative and is such that the output function has a monotonicity property, then there exists an observer with an infinite gain margin.

I. Introduction

For a nonlinear system of the form

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

with $x \in \mathbb{R}^n$ being the system's state and $y \in \mathbb{R}^m$ the measured system's output, we study the problem of obtaining an estimate $\hat{x}$ of the state x by means of the dynamical system, called *observer*,

$$\dot{\chi} \;=\; F(\chi, y)\ , \quad \hat{x} \;=\; H(\chi, y) \tag{2}$$

with $\chi \in \mathbb{R}^p$ being the observer's state and $\hat{x} \in \mathbb{R}^n$ the observer's output, used as the system's state estimate. We focus on the case where the state χ of the observer evolves in a copy of the space of the system's state x, i.e., they both belong to $\mathbb{R}^n$, with, moreover, an output function H such that $\hat{x} = \chi$. We consider the following observer design problem:

($\star$) Given functions f and h, design a function F such that for the system

$$\dot{x} \;=\; f(x)\ , \quad \dot{\hat{x}} \;=\; F(\hat{x}, h(x)), \tag{3}$$

R. G. Sanfelice is with the Department of Aerospace and Mechanical Engineering, University of Arizona 1130 N. Mountain Ave, AZ 85721, Email: sricardo@u.arizona.edu

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

the zero estimation error set

$$\mathcal{A} \;=\; \{(x,\hat{x}) \in \mathbb{R}^n \times \mathbb{R}^n \;:\; x = \hat{x}\} \tag{4}$$

is globally asymptotically stable (see the text below (8)).

Many contributions from different viewpoints have been made to address problem ($\star$). While a summary of the very rich literature on the topic is out of the scope of this paper, it is important to point out the interest of exploiting a possible contraction property of the flow generated by the observer. Study of contracting flows has a very long history and has been proposed independently by several authors; see, e.g., [18], [10], [7], [20], [19] (see [14] for a historical discussion). In the context of observers, Riemannian metrics have been used in [1], [3], [4], for instance, with the objective of guaranteeing that the Riemannian distance between the system and observer solutions decreases to zero. In these papers, the authors consider systems whose dynamics follow from a principle of least action involving a Riemannian metric, such as Lagrangian systems with a Lagrangian that is quadratic in the generalized velocities. The observer design therein exploits some properties of this metric and local convergence is established via some ad-hoc modification of this metric or choice of coordinates.

This paper advocates that, since the observability of the system linearized along each of its solutions may vary significantly from one solution to another, the native Euclidean geometry of the state space may not be appropriate to study convergence properties of an observer. Instead of insisting in using a Riemannian metric associated to the system's dynamics, we propose to study Riemannian metrics incorporating information on the system's dynamics and observability. In Section II-B, we show that if for a given Riemannian metric an observer whose state χ lives in a copy of the given system's state space and makes the Riemannian distance along system and observer solutions nonincreasing then, necessarily, the Lie derivative of the metric along the system solutions satisfies an inequality involving the output function. Section II-C shows that if the same conditions hold and the observer has an infinite gain margin then, necessarily, the level sets of the output function are geodesically convex. In Section II-D we establish that if a Riemannian metric with a Lie derivative satisfying the inequality mentioned above is, in some coordinates, uniformly bounded away from zero and upper bounded then the system's linearization along each of its solution must be detectable. With the insight provided by these necessary conditions, Section III proposes a set of sufficient conditions guaranteeing the existence of an observer whose flow leads to a decreasing Riemannian distance between system's state and estimated state.

For the sake of simplicity, we assume throughout the paper that the functions are differentiable sufficiently many times. Moreover, we work under restrictions that can be further relaxed, such as time independence of the right-hand sides and forward completeness of the systems[1].

This paper is devoted to analysis. In a companion paper, we focus on observer design, namely, on the construction of a Riemannian metric satisfying the desired inequality on its Lie derivative and making the level sets of the output function possibly totally geodesic.

[1] A system is said to be *forward complete* if each of its solutions exists on $[0,+\infty)$.

Example 1.1 (Motivational example): We illustrate our results in the following academic system

$$\dot{x}_1 = x_2\sqrt{1+x_1^2}, \quad \dot{x}_2 \;=\; -\frac{x_1}{\sqrt{1+x_1^2}}\,x_2^2, \quad y \;=\; x_1 \;. \tag{5}$$

For this system (5), by following [16], we get the observer

$$\begin{aligned} \dot{\bar{\hat{x}}}_1 \;&=\; \bar{\hat{x}}_2 - (\bar{\hat{x}}_1 - y), \quad \dot{\bar{\hat{x}}}_2 \;=\; -(\bar{\hat{x}}_1 - y), \\ \hat{x}_1 \;&=\; \bar{\hat{x}}_1, \quad \hat{x}_2 \;=\; \frac{\bar{\hat{x}}_2}{\sqrt{1+y^2}} \;. \end{aligned} \tag{6}$$

This observer is in the form (2), but cannot be written in the form of (3) with the $(\hat{x}_1, \hat{x}_2)$ coordinates since this would involve x_2. Nevertheless, with the Lyapunov function

$$\begin{aligned} V(\hat{x}, x) \;=\;\; & (\hat{x}_1 - x_1)^2 \;-\; (\hat{x}_1 - x_1)\,(\hat{x}_2 - x_2)\,\sqrt{1+x_1^2} \\ & + (\hat{x}_2 - x_2)^2\,(1+x_1^2) \end{aligned} \tag{7}$$

we obtain for the system-observer interconnection (5)-(6)

$$\dot{\overbrace{V(\hat{x}, x)}} \;=\; -V(\hat{x}, x) \;.$$

Since V satisfies, for all $(x, \hat{x}) \in \mathbb{R}^2 \times \mathbb{R}^2$,

$$\begin{aligned} \frac{(\hat{x}_1 - x_1)^2 + (\hat{x}_2 - x_2)^2}{2} &\leq V(\hat{x}, x) \\ &\leq \frac{3}{2}\left[(\hat{x}_1 - x_1)^2 + (\hat{x}_2 - x_2)^2\right]\left(1 + x_1^2\right), \end{aligned}$$

this implies that, for all $t \geq 0$ and all $(x, \hat{x}) \in \mathbb{R}^2 \times \mathbb{R}^2$,

$$|X(x,t) - \hat{X}((\hat{x}, x), t)|^2 \leq 3\exp(-t)(1+x_1^2)|x - \hat{x}|^2 \;, \tag{8}$$

where $(X(x,t), \hat{X}((\hat{x}, x), t))$ is the solution issued from points $(x, \hat{x})$ for the system-observer interconnection (5)-(6). This establishes that the set $\mathcal{A}$ is globally asymptotically stable (nonuniformly in x but uniformly in $x - \hat{x}$).

As it will be shown in Section II-A, the key point here is that V is the square of a Riemannian distance between $\hat{x}$ and x that is associated to an x-dependent Riemannian metric. Moreover, as justified in Section II-B, no matter what the observer is, it is impossible to find a standard quadratic form expressed in the given coordinates (i.e., a Riemannian distance associated with a constant Riemannian metric) that is nonincreasing along solutions. This is a motivation for the analysis of observers using x-dependent Riemannian metrics. □

II. Necessary conditions for having a Riemannian distance between system and observer solutions to decrease.

A. Riemannian Distance

As discussed in Section I, the notions of nonexpanding/contracting flow and geodesically monotone vector fields are suitable for studying asymptotic stability of the zero error set $\mathcal{A}$ in (4). We start by recalling some basic facts on Riemannian distance.

Let $P : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be a C^3 symmetric covariant two-tensor (see, e.g., [24, Page 17]). If x and $\bar{x}$ are two sets of coordinates related by $\bar{x} = \phi(x)$ with ϕ being a diffeomorphism, then P expressed in x coordinates as $P(x)$ and in $\bar{x}$ coordinates as $\bar{P}(\bar{x})$ are related by (see, e.g., [24, Example II.2])

$$P(x) \;=\; \frac{\partial\phi}{\partial x}(x)^\top \bar{P}(\bar{x})\, \frac{\partial\phi}{\partial x}(x)\ . \tag{9}$$

If P takes positive definite values then the length of a C^1 path γ between points x_1 and x_2 is defined as

$$L(\gamma)\Big|_{s_1}^{s_2} \;=\; \int_{s_1}^{s_2} \sqrt{\frac{d\gamma}{ds}(s)^\top P(\gamma(s))\frac{d\gamma}{ds}(s)}\, ds, \tag{10}$$

where

$$\gamma(s_1) \;=\; x_1 \quad , \qquad \gamma(s_2) \;=\; x_2\ .$$

With such a definition, P is also called a Riemannian metric. The Riemannian distance $d(x_1, x_2)$ is the minimum of $L(\gamma)\Big|_{s_1}^{s_2}$ among all possible piecewise C^1 paths γ between x_1 and x_2. To relate the Riemannian distance with geodesics, we invoke the Hopf-Rinow Theorem (see, e.g., [24, Theorem II.1.1]), which asserts the following: if every geodesic can be maximally extended to $\mathbb{R}$ then the minimum of $L(\gamma)\Big|_{s_1}^{s_2}$ is actually given by the length of a (maybe nonunique) geodesic, which is called a *minimal geodesic*; for more details, see, e.g., [5] and [8]. In the appendix we show that, in our context, this maximal extension property holds on $\mathbb{R}^n$ if there exist globally defined coordinates in which P satisfies

$$0 \;<\; P(x) \quad \forall x\in\mathbb{R}^n\ , \quad \lim_{r\to\infty} r^2 \underline{p}(r) \;=\; +\infty, \tag{11}$$

where, for any positive real number r,

$$\underline{p}(r) \;=\; \min_{x:|x|\leq r} \lambda_{\min}\left(P(x)\right)\ ,$$

with $\lambda_{\min}\left(P(x)\right)$ denoting the minimum eigenvalue of $P(x)$. In this case, the Riemannian metric given by P is said to be complete and, denoting by γ^* a minimal (normalized[2]) geodesic between $x = \gamma^*(0)$ and $\hat{x} = \gamma^*(\hat{s})$, with $\hat{s} \geq 0$, the Riemannian distance $d(\hat{x}, x)$ is

$$d(\hat{x}, x) \;=\; L(\gamma^*)\Big|_{0}^{\hat{s}} \;=\; \hat{s}\ . \tag{12}$$

Example 2.1: As an illustration, consider the symmetric covariant two-tensor expressed in x coordinates as $P(x) = \begin{bmatrix} 1 - \frac{x_1x_2}{\sqrt{1+x_1^2}} + \frac{x_1^2x_2^2}{1+x_1^2} & -\frac{\sqrt{1+x_1^2}}{2} + x_1x_2 \\ -\frac{\sqrt{1+x_1^2}}{2} + x_1x_2 & 1 + x_1^2 \end{bmatrix}$. Since condition (11) holds with $\underline{p}(r) = \frac{1}{2}$ for all $r > 0$, it is a complete Riemannian metric. Moreover, using (9), it is easy to check that in the coordinates $\bar{x} = \phi(x) =$

[2]A normalized geodesic γ^* satisfies $\frac{d\gamma^*}{ds}(s)^\top P(\gamma^*(s))\, \frac{d\gamma^*}{ds}(s) \; = 1$ for all s in its domain of definition. In the following, the adjective "normalized" is omitted.

$\begin{bmatrix} x_1 \\ x_2\sqrt{1+x_1^2} \end{bmatrix}$, its expression is $\bar{P}(\bar{x}) = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$. Since $\bar{P}(\bar{x})$ is constant, any minimal geodesic $\bar{\gamma}^*$ takes the form $\bar{\gamma}^*(s) = \bar{x} + s\bar{v}$ with $\bar{v} \in \mathbb{R}^2$ satisfying $\bar{v}^\top \bar{P}(\bar{x})\bar{v} = 1$. Then, a minimal geodesic in x coordinates is given by $\gamma^*(s) = \phi^{-1}(\bar{x} + s\bar{v})$. Accordingly, the Riemannian distance between $\hat{x}$ and x is

$$\begin{aligned}
\int_0^{\hat{s}} \sqrt{\tfrac{d\gamma^*}{ds}(s)^\top P(\gamma^*(s))\tfrac{d\gamma^*}{ds}(s)}\, ds \quad = \quad d(\hat{x}, x) \;&=\; d(\bar{\hat{x}}, \bar{x}) \\
&= \int_0^{\hat{s}} \sqrt{\tfrac{d\bar{\gamma}^*}{ds}(s)^\top \bar{P}(\bar{\gamma}^*(s))\tfrac{d\bar{\gamma}^*}{ds}(s)}\, ds \\
&= \sqrt{(\bar{\hat{x}} - \bar{x})^\top \bar{P}(\bar{x})(\bar{\hat{x}} - \bar{x})} \\
&= \sqrt{(\phi(\hat{x}) - \phi(x))^\top \bar{P}(\bar{x})(\phi(\hat{x}) - \phi(x))} \\
&= \sqrt{V(\hat{x}, x)},
\end{aligned}$$

where V is given in (7) and $\bar{\hat{x}} = \phi(\hat{x})$. □

Having a Riemannian distance, we say that a system $\dot{x} = f(x)$, with solutions $X(x,t)$, generates a nonexpanding (respectively, contracting) flow if, for any pair (x_1, x_2) in $\mathbb{R}^n \times \mathbb{R}^n$, the function $t \mapsto d(X(x_1,t), X(x_2,t))$ is nonincreasing (respectively, strictly decreasing); see, e.g., [13]. Also, the vector field f is said to be geodesically monotonic (respectively, strictly monotonic) if we have

$$\mathcal{L}_f P(x) \;\leq\; 0 \qquad (\text{respectively} \;\; < \; 0) \qquad \forall x \in \mathbb{R}^n \;, \tag{13}$$

where $\mathcal{L}_f P$ is the Lie derivative of the symmetric covariant two-tensor P, whose expression in x coordinates is

$$\begin{aligned}
&v^\top \mathcal{L}_f P(x)\, v \\
&\quad = \lim_{r\to 0} \left[\frac{[(I + r\frac{\partial f}{\partial x}(x))v]^\top P(x + rf(x))[(I + r\frac{\partial f}{\partial x}(x))v]}{r} \right. \\
&\qquad\qquad \left. - \frac{v^\top P(x)v}{r} \right] \\
&\quad = \frac{\partial}{\partial x}\left(v^\top P(x)\, v\right) f(x) \;+\; 2\, v^\top P(x) \left(\frac{\partial f}{\partial x}(x)\, v\right)
\end{aligned} \tag{14}$$

for all $v \in \mathbb{R}^n$; see [5, Exercise V.2.8], [24, Page 17], or [17]. We have the following result (see, for instance, [13] or [1] for a proof).

Lemma 2.2: A geodesically monotonic (respectively, strictly monotonic) vector field generates a nonexpanding (respectively, contracting) flow.

If inequality (13) holds for the observer vector field F then $t \mapsto d(\hat{X}((\hat{x}_1, x), t), \hat{X}((\hat{x}_2, x), t))$ is (respectively, strictly) decreasing; however, this property is more than what is needed for the zero estimation error set $\mathcal{A}$ to be (respectively, asymptotically) stable. Actually, it is sufficient to have an observer giving rise to a (respectively, strictly) decreasing function $t \mapsto d(\hat{X}((\hat{x}, x), t), X(x,t))$ for all pairs $(\hat{x}, x)$ in $\mathbb{R}^n \times \mathbb{R}^n$. That is, we do not insist on having a Riemannian distance between any two arbitrary observer solutions to decrease, but only to have a decreasing Riemannian distance between any observer solution and its corresponding system solution (which is a particular observer solution).

B. Necessity of geodesic monotonicity in the directions tangent to the level sets of the output function

Since the Riemannian distance between $\hat{x}$ and x is locally Lipschitz, its upper right-hand Dini derivative is given by

$$\mathfrak{D}^{+}d(\hat{x},x) \;=\; \limsup_{t\to 0_{+}} \frac{d(\hat{X}((\hat{x},x),t),X(x,t))-d(\hat{x},x)}{t} \tag{15}$$

for each $(\hat{x},x)\in\mathbb{R}^n\times\mathbb{R}^n$. It is nonpositive when the function $t\mapsto d(\hat{X}((\hat{x},x),t),X(x,t))$ is nonincreasing.

Theorem 2.3: *Assume there exists a complete C^3 Riemannian metric P such that, for each $(\hat{x},x)\in\mathbb{R}^n\times\mathbb{R}^n$,*

$$\mathfrak{D}^{+}d(\hat{x},x)\leq 0 \tag{16}$$

holds along any solution of (3)*, then*

$$v^{\top}\mathcal{L}_{f}P(x)v \;\leq\; 0 \qquad \forall (x,v)\in\mathbb{R}^n\times\mathbb{R}^n \;\; \text{such that} \;\; \frac{\partial h}{\partial x}(x)v=0\ . \tag{17}$$

Furthermore, if there exists a function $\omega:\mathbb{R}^n\times\mathbb{R}^n\to[0,+\infty)$ such that $(\hat{x},x)\mapsto d(\hat{x},x)\omega(\hat{x},x)$ is a C^2 function on a neighborhood $\mathcal{N}_{\mathcal{A}}$ of $\mathcal{A}$ with the property that, for some $\varepsilon>0$,

$$\frac{\partial^2(d\,\omega)}{\partial\hat{x}^2}(x,x)\geq\varepsilon P(x) \qquad \forall x\in\mathbb{R}^n \tag{18}$$

and, for each $(\hat{x},x)\in\mathcal{N}_{\mathcal{A}}$,

$$\mathfrak{D}^{+}d(\hat{x},x)\leq-\omega(\hat{x},x) \tag{19}$$

holds along any solution of (3)*, then there exists a continuous function $\rho:\mathbb{R}^n\to\mathbb{R}$ satisfying*

$$\mathcal{L}_{f}P(x) \;\leq\; \rho(x)\,\frac{\partial h}{\partial x}(x)^{\top}\frac{\partial h}{\partial x}(x)-\frac{\varepsilon}{2}P(x) \qquad \forall x\in\mathbb{R}^n. \tag{20}$$

Proof: To simplify the notation, let $V:\mathbb{R}^n\times\mathbb{R}^n\to[0,+\infty)$ be the function defined as the square of the Riemannian distance, i.e., $V(\hat{x},x)=d(\hat{x},x)^2$, and notice that[3]

$$\mathfrak{D}^{+}V(\hat{x},x)=\mathfrak{D}^{+}d^2(\hat{x},x)\leq 2\,d(\hat{x},x)\,\mathfrak{D}^{+}d(\hat{x},x). \tag{21}$$

Pick an arbitrary point x in $\mathbb{R}^n$. From [15, Theorem 3.6], there exists a (normal coordinate) neighborhood $\mathcal{N}_x$ such that V is C^2 on $\mathcal{N}_x\times\mathcal{N}_x$. From (21) and (16) (respectively, from (21) and (19), on $(\mathcal{N}_x\times\mathcal{N}_x)\cap\mathcal{N}_{\mathcal{A}}$), we have

$$\mathfrak{D}^{+}V(\hat{x},x)\leq 0 \qquad (\,\text{respectively }\;\leq -2\,d(\hat{x},x)\omega(\hat{x},x)\,).$$

[3]Since $\limsup(a\,b)\leq\limsup a\cdot\limsup b$.

Let r_* be a strictly real number such that, for any v in $\mathbb{S}^n$, the unit sphere, and for all $r \in [0, r_*)$, $(\hat{x} + rv, x)$ are the coordinates of a point in $(\mathcal{N}_x \times \mathcal{N}_x) \cap \mathcal{N}_\mathcal{A}$. We have [4]

$$\frac{\partial^2 V}{\partial \hat{x}^2}(x,x) = \frac{\partial^2 V}{\partial x^2}(x,x) \ = \ 2P(x) \tag{22}$$

and[5]

$$\begin{gathered} V(x,x) \ = \ 0 \ , \frac{\partial V}{\partial x}(x,x) \ = \ \frac{\partial V}{\partial \hat{x}}(x,x) \ = \ 0 \\ \frac{\partial^2 V}{\partial x^2}(x,x) + \frac{\partial^2 V}{\partial \hat{x} \partial x}(x,x) = \frac{\partial^2 V}{\partial \hat{x}^2}(x,x) \ + \ \frac{\partial^2 V}{\partial x \partial \hat{x}}(x,x) \ = \ 0 \end{gathered} \tag{23}$$

and, for all $r \in [0, r^*)$ and $v \in \mathbb{S}^n$,

$$\begin{aligned} \mathfrak{D}^+ V(x+rv, x) \quad &= \quad \frac{\partial V}{\partial x}(x+rv,x)\, f(x) \\ &\qquad + \frac{\partial V}{\partial \hat{x}}(x+rv,x)\, F(x+rv, h(x)) \\ &\leq \quad 0 \\ (\text{ respectively} \quad &\leq \quad -2\, d(x+rv,x)\, \omega(x+rv,x) \ ). \end{aligned}$$

With the definition of d, this implies that $\mathcal{A}$ is forward invariant, i.e., the solutions to (3) with $x = \hat{x}$ as initial condition remain in $\mathcal{A}$ for all $t \geq 0$. This implies

$$F(x, h(x)) = f(x) \ . \tag{24}$$

By differentiating this identity with respect to x, we get

$$\frac{\partial F}{\partial x}(x, h(x)) \ + \ \frac{\partial F}{\partial y}(x, h(x)) \frac{\partial h}{\partial x}(x) \ = \ \frac{\partial f}{\partial x}(x). \tag{25}$$

For r in $(0, r_*)$, we obtain

$$\begin{aligned} &\frac{1}{r^2}\left[\frac{\partial V}{\partial x}(x+rv,x) + \frac{\partial V}{\partial \hat{x}}(x+rv,x)\right] f(x) + \\ &\qquad\qquad \frac{\frac{\partial V}{\partial \hat{x}}(x+rv,x)}{r}\, \frac{F(x+rv, h(x)) - f(x)}{r} \leq 0 \\ &\qquad\qquad\qquad\qquad (\text{respectively} \leq \ -\frac{2}{r^2} d(x+rv,x)\, \omega(x+rv,x)). \end{aligned} \tag{26}$$

To compute the limit for r approaching 0 note that we have the following Taylor expansion around (x,x)

$$\begin{aligned} V(x+rv,x) \quad &= \quad V(x,x) + r\, \frac{\partial V}{\partial \hat{x}}(x,x)\, v \\ &\qquad + \frac{r^2}{2}\, v^\top \frac{\partial^2 V}{\partial \hat{x}^2}(x,x)\, v + O_{x,v}(r^3) \ , \end{aligned}$$

[4]This follows from the fact that a first order approximation of the geodesic is $\gamma(s) = x + s\, v + O_{x,v}(s^2)$ with $v^\top P(x)\, v \ = \ 1$, which yields $V(\hat{x}, x) = d(\hat{x}, x)^2 \ = \ \hat{s}^2 \ = \ (\hat{x} - x)^\top P(x)\, (\hat{x} - x) \ + \ O_{x,v}(\hat{s}^3)$, where the subindex in $O_{x,v}$ indicates dependence on (x, v).

[5]This follows from $x = \hat{x}$ being a minimizer of V for all x.

$$
\begin{aligned}
\frac{\partial V}{\partial \hat{x}}(x+rv,x) \quad &= \quad \frac{\partial V}{\partial \hat{x}}(x,x) + r\,\frac{\partial^2 V}{\partial \hat{x}^2}(x,x)\,v + O_{x,v}(r^2)\ , \\
\frac{F(x+rv,h(x))-f(x)}{r} \quad &= \quad \frac{F(x,h(x))-f(x)}{r} \\
&\qquad + \frac{\partial F}{\partial \hat{x}}(x,h(x))\,v\ +\ O_{x,v}(r).
\end{aligned}
$$

Define $W(x) = V(x + r\,v, x)$ and note that

$$
\frac{\partial W}{\partial x}(x) = \frac{\partial V}{\partial \hat{x}}(x+rv,x) + \frac{\partial V}{\partial x}(x+rv,x).
$$

With (22) and (23), we get

$$
\begin{aligned}
W(x) \quad &= \quad r^2\, v^\top P(x)\, v + O_{x,v}(r^3)\ , \\
\frac{1}{r}\frac{\partial V}{\partial \hat{x}}(x+rv,x) \quad &= \quad 2\, v^\top P(x) + O_{x,v}(r),
\end{aligned}
$$

and with (24)

$$
\frac{F(x+rv,h(x))-f(x)}{r} \quad = \quad \frac{\partial F}{\partial \hat{x}}(x,h(x))\,v\ +\ O_{x,v}(r).
$$

This yields

$$
\begin{aligned}
&\lim_{r\to 0}\frac{1}{r^2}\left[\frac{\partial V}{\partial \hat{x}}(x+rv,x) + \frac{\partial V}{\partial x}(x+rv,x)\right] f(x) \\
&\qquad = \lim_{r\to 0}\frac{1}{r^2}\frac{\partial W}{\partial x}(x) f(x) = \frac{\partial\left(v^\top P\, v\right)}{\partial x}(x) f(x).
\end{aligned}
\tag{27}
$$

Also, with (24), we get

$$
\begin{aligned}
&\lim_{r\to 0}\frac{\frac{\partial V}{\partial \hat{x}}(x+rv,x)}{r}\,\frac{F(x+rv,h(x))-f(x)}{r} \\
&\qquad = 2v^\top P(x)\,\frac{\partial F}{\partial x}(x,h(x))\,v.
\end{aligned}
\tag{28}
$$

Similarly, we can obtain

$$
\lim_{r\to 0}\frac{2}{r^2} d(x+rv,x)\,\omega(x+rv,x)\ =\ v^\top \frac{\partial^2(d\,\omega)}{\partial \hat{x}^2}(x,x)\,v\ . \tag{29}
$$

Then, combining (27), (28), and (29), we have that inequality (26) gives

$$
\begin{aligned}
&\frac{\partial\left(v^\top P v\right)}{\partial x}(x) f(x) + 2v^\top P(x)\frac{\partial F}{\partial x}(x,h(x))v \quad \leq \quad 0 \\
&\qquad \text{(respectively } \leq -v^\top \frac{\partial^2(d\,\omega)}{\partial \hat{x}^2}(x,x)v \qquad \forall v\in \mathbb{S}^n\ ),
\end{aligned}
$$

or, equivalently, using (25) and (14),

$$
v^\top \mathcal{L}_f P(x) v - 2v^\top P(x)\frac{\partial F}{\partial y}(x,h(x))\frac{\partial h}{\partial x}(x)v\ \leq\ 0 \tag{30}
$$

$$
\text{(respectively } \leq -v^\top \frac{\partial^2(d\,\omega)}{\partial \hat{x}^2}(x,x)v \quad \forall v\in \mathbb{S}^n). \tag{31}
$$

It follows that (30) already implies (17). Also, when (19) holds, by completing squares and using Cauchy-Schwarz inequality, we get successively, for any function $\rho : \mathbb{R}^n \to (0,+\infty)$ and all (x,v) in $\mathbb{R}^n \times \mathbb{S}^n$,

$$\begin{aligned} 2v^\top P(x)\frac{\partial F}{\partial y}(x,h(x))\frac{\partial h}{\partial x}(x)v \quad &\leq \quad \rho(x)\left|\frac{\partial h}{\partial x}(x)v\right|^2 \\ &\quad + \frac{1}{\rho(x)}\left|v^\top P(x)\frac{\partial F}{\partial y}(x,h(x))\right|^2 \\ &\leq \quad \rho(x)v^\top \frac{\partial h}{\partial x}(x)^\top \frac{\partial h}{\partial x}(x)v \\ &\quad + \frac{\left|\frac{\partial F}{\partial y}(x,h(x))^\top P(x)\frac{\partial F}{\partial y}(x,h(x))\right|}{\rho(x)}\, v^\top P(x)v \ . \end{aligned}$$

Equation (20) follows from (18) by picking ρ as any continuous function satisfying

$$\frac{2}{\varepsilon}\left|\frac{\partial F}{\partial y}(x,h(x))^\top P(x)\frac{\partial F}{\partial y}(x,h(x))\right| \ \leq\ \rho(x)$$

for all $x \in \mathbb{R}^n$. ■

When compared with (13), which says f is (respectively, strictly) geodesically monotonic, the necessary condition (17) (respectively, (20)) says only that the vector field f is geodesically (respectively, strictly) monotonic in the directions v satisfying $\frac{\partial h}{\partial x}(x)v = 0$, i.e., in the directions tangent to the level sets of the output function h.

Remark 2.4: Theorem 2.3 can be interpreted as an extension of [21, Proposition 3]. In this reference, a C^∞ function V depending only on $\hat{x} - x$, called a *state-independent error Lyapunov function*, is obtained from stability properties of $\mathcal{A}$. In such a case, the conditions in (23) yield a constant matrix P. Then, Theorem 2.3 implies that, for all $x \in \mathbb{R}^n$, P is a semidefinite positive matrix that satisfies, for all $x \in \mathbb{R}^n$,

$$P\,\frac{\partial f}{\partial x}(x) \ +\ \frac{\partial f}{\partial x}(x)^\top P \ \leq\ \rho(x)\,\frac{\partial h}{\partial x}(x)^\top \frac{\partial h}{\partial x}(x) \ -\ \frac{\varepsilon}{2}\,P \ .$$

It follows that, for all $x \in \mathbb{R}^n$ and $c \in [0, \frac{\varepsilon}{4}]$, we have the implication

$$\frac{\partial h}{\partial x}(x)v = 0 \quad \Rightarrow \quad v^\top P\,\frac{\partial f}{\partial x}(x)v \ \leq\ -\,c\,v^\top P v \ . \tag{32}$$

When $c = 0$, this property corresponds to the one established in [21, Proposition 3]. It is worth pointing out that a limitation of the work in [21] is that the results are extrinsic, i.e., they depend on the coordinates since a quadratic form may not be quadratic after a nonlinear change of coordinates. On the other hand, the necessary conditions in Theorem 2.3 are intrinsic. In fact, let ϕ be a diffeomorphism on $\mathbb{R}^n$ leading to the new coordinates

$$\bar{x} \ =\ \phi(x) \quad , \qquad \bar{\hat{x}} \ =\ \phi(\hat{x}) \ . \tag{33}$$

Let $\bar{h}$, $\overline{d}, \overline{\omega}$, $\bar{\rho}$, $\bar{f}$, and $\bar{P}$ be h, d, ω, ρ, f, and P, respectively, in the new coordinates. We have (9) and

$$\bar{h}(\bar{x}) = h(x)\ ,\quad \frac{\partial h}{\partial x}(x) \ = \ \frac{\partial \bar{h}}{\partial \bar{x}}(\bar{x})\frac{\partial \phi}{\partial x}(x),$$

$$\bar{f}(\bar{x}) = \frac{\partial \phi}{\partial x}(x)\, f(x)\ ,$$

$$\bar{d}(\bar{\hat{x}}, \bar{x}) = d(\hat{x}, x), \qquad \bar{\omega}(\bar{\hat{x}}, \bar{x}) = \omega(\hat{x}, x)$$

$$\frac{\partial^2 (d\,\omega)}{\partial \hat{x}^2}(x,x) = \frac{\partial \phi}{\partial x}(x)^\top \frac{\partial^2 (\overline{d}\,\overline{\omega})}{\partial \bar{\hat{x}}^2}(\bar{x},\bar{x})\frac{\partial \phi}{\partial x}(x)\,,$$

$$\bar{\rho}(\bar{x}) = \rho(x)\ ,\quad \mathcal{L}_f P(x) \ = \ \frac{\partial \phi}{\partial x}(x)^\top \mathcal{L}_{\bar{f}} \bar{P}(\bar{x})\frac{\partial \phi}{\partial x}(x)\ .$$

Substituting these expressions in (20), we get

$$\begin{aligned}\frac{\partial \phi}{\partial x}(x)^\top \mathcal{L}_{\bar{f}} \bar{P}(\bar{x})\frac{\partial \phi}{\partial x}(x) \ \leq \ \bar{\rho}(\bar{x}) \ &\left[\frac{\partial \bar{h}}{\partial \bar{x}}(\bar{x})\frac{\partial \phi}{\partial x}(x)\right]^\top \times \\ &\left[\frac{\partial \bar{h}}{\partial \bar{x}}(\bar{x})\frac{\partial \phi}{\partial x}(x)\right] - \frac{1}{2}\frac{\partial \phi}{\partial x}(x)^\top \frac{\partial^2 (\overline{d}\,\overline{\omega})}{\partial \bar{\hat{x}}^2}(\bar{x},\bar{x})\frac{\partial \phi}{\partial x}(x)\end{aligned}$$

and since $\frac{\partial \phi}{\partial x}(x)$ is invertible it gives

$$\mathcal{L}_{\bar{f}} \bar{P}(\bar{x}) \quad \leq \quad \bar{\rho}(\bar{x})\,\frac{\partial \bar{h}}{\partial \bar{x}}(\bar{x})^\top \frac{\partial \bar{h}}{\partial \bar{x}}(\bar{x}) - \frac{1}{2}\frac{\partial^2 (\overline{d}\,\overline{\omega})}{\partial \bar{\hat{x}}^2}(\bar{x},\bar{x}),$$

which is inequality (20) in $\bar{x}$ coordinates.

Furthermore, from the definition of $\mathcal{L}_f P$ and with completion of squares as in the proof of Theorem 2.3, it can be checked that condition (20) is preserved, but with a modified function ρ, after an output-dependent time scaling of the system, i.e., when f is replaced by $\bar{f}(x) = \theta(h(x))f(x)$ with θ taking strictly positive values. □

The necessary conditions in Theorem 2.3 can be used to characterize the family of Riemannian metrics possibly leading to a Riemannian distance that is nonincreasing (via (17)) or strictly decreasing (via (20)) along solutions. For instance, condition (17) can be used to justify that, for system (5), there is no such a Riemannian metric that is constant.

Example 2.5 (Motivational example – continued): For the family of constant Riemannian metrics of the form $P = \begin{bmatrix} p & q \\ q & r \end{bmatrix}$, $, p, r > 0\ , p\, r > q^2$ for (5), for each $v \in \mathbb{R}^2$ such that

$$\frac{\partial h}{\partial x}(x)v = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0\ ,$$

we obtain

$$
\begin{aligned}
&v^\top P \, \frac{\partial f}{\partial x}(x) v \;+\; v^\top \frac{\partial f}{\partial x}(x)^\top P v \\
&\qquad = \frac{2}{\sqrt{1+x_1^2}} \, v^\top P \begin{bmatrix} x_1 x_2 & 1+x_1^2 \\ -\dfrac{x_2^2}{1+x_1^2} & -2x_1x_2 \end{bmatrix} v \\
&\qquad = \frac{v_2^2 (2\, q \,(1+x_1^2) - 4\, r\, x_1\, x_2)}{\sqrt{1+x_1^2}} \ ,
\end{aligned}
$$

which cannot be nonpositive for each x. On the other hand, it can be shown that the family of Riemannian metrics satisfying (17) can be described as

$$
P(x) = \begin{bmatrix} 1 & \frac{x_1x_2}{\sqrt{1+x_1^2}} \\ 0 & \sqrt{1+x_1^2} \end{bmatrix} \begin{bmatrix} \bar{p}(\bar{x}) & \bar{q}(\bar{x}) \\ \bar{q}(\bar{x}) & \bar{r}(\bar{x}) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \frac{x_1x_2}{\sqrt{1+x_1^2}} & \sqrt{1+x_1^2} \end{bmatrix} \tag{34}
$$

with $(\bar{x}_1, \bar{x}_2) = (x_1, x_2\sqrt{1+x_1^2})$ and $\bar{r}(\bar{x}) = a(\bar{x})^2$, $\bar{q}(\bar{x}) = -b(\bar{x})^2 - \frac{1}{2}\frac{\partial \bar{r}}{\partial \bar{x}_1}(\bar{x})\bar{x}_2$, $\bar{p}(\bar{x}) = c(\bar{x})^2 + \frac{\bar{q}(\bar{x})^2}{\bar{r}(\bar{x})}$, where $a, b, c : \mathbb{R}^2 \to \mathbb{R}$ are sufficiently smooth functions with a and c not vanishing. A particular choice is $a(\bar{x}) = 1$, $b(\bar{x}) = \frac{1}{(1+\bar{x}_1^2)^{\frac{1}{4}}}$, and $c(\bar{x})^2 = 1 + \left(\frac{\bar{x}_2}{1+\bar{x}_1^2} + \frac{\bar{x}_1}{\sqrt{1+\bar{x}_1^2}}\right)^2$, which leads to

$$
P(x) \;=\; \begin{bmatrix} 2+x_2^2 & x_1x_2-1 \\ x_1x_2-1 & 1+x_1^2 \end{bmatrix} . \tag{35}
$$

□

C. *Necessity of geodesic convexity of the level sets of the output function*

In Theorem 2.3, we studied the implications of the existence of an observer making $t \mapsto d(\hat{X}((\hat{x},x),t), X(x,t))$ nonincreasing, in particular, when $\hat{x}$ converges to x (in the proof, $(x+rv,x)$ approaches (x,x)). Now we study the implications of the existence of such an observer for the case when $\hat{x}$ is far away from x. To this end, for each s in $[0,\hat{s}]$, let $t \mapsto \Gamma(s,t)$ be a C^1 function satisfying

$$
\begin{aligned}
\frac{\partial X}{\partial t}(x,t) \;&=\; f(X(x,t)) \ , \quad X(x,0) \;=\; x \ , \\
\frac{\partial \hat{X}}{\partial t}(\hat{x},t) \;&=\; F(\hat{X}(\hat{x},t), h(X(x,t))) \ , \quad \hat{X}(\hat{x},0) \;=\; \hat{x} \ , \\
\frac{\partial \Gamma}{\partial t}(s,t) \;&=\; F(\Gamma(s,t), h(X(x,t))) \ , \quad \Gamma(s,0) \;=\; \gamma^*(s) \ ,
\end{aligned}
$$

with γ^* a minimal geodesic between x and $\hat{x}$. Then, we have $\hat{X}((\hat{x},x),t) \;=\; \Gamma(\hat{s},t)$ and hence, at time t, $s \mapsto \Gamma(s,t)$ is a path between $X(x,t)$ and $\hat{X}((\hat{x},x),t)$. Also, we have

$$
d(\hat{x},x) \;=\; d(\Gamma(\hat{s},0), \Gamma(0,0)) \;=\; L(\Gamma(\,\bullet\,,0))\Big|_0^{\hat{s}} \ .
$$

Also, we know from the first order variation formula (see, for instance, [25, Theorem 6.14] or [13, Theorem 5.7]) that we have

$$\begin{aligned}
&\frac{d}{dt}\, L(\Gamma(\,\bullet\,,t))\Big|_0^{\hat{s}}\Big|_{t=0} \\
&\qquad = \frac{d}{dt}\int_0^{\hat{s}} \sqrt{\frac{\partial\Gamma}{\partial s}(s,t)^\top P(\Gamma(s,t))\frac{\partial\Gamma}{\partial s}(s,t)}\; ds\Big|_{t=0} \\
&\qquad = \frac{d\gamma^*}{ds}(\hat{s})^\top P(\gamma^*(\hat{s}))\, F(\gamma^*(\hat{s}),y) \\
&\qquad\quad -\frac{d\gamma^*}{ds}(0)^\top P(\gamma^*(0))\, F(\gamma^*(0),y).
\end{aligned}$$

On the other hand, in general, for each t in the domain of definition, we have only

$$d(\hat{X}(\hat{x},t),X(x,t)) \;=\; d(\Gamma(\hat{s},t),\Gamma(0,t)) \;\leq\; L(\Gamma(\,\bullet\,,t))\Big|_0^{\hat{s}} \;.$$

Then, the upper right-hand Dini derivative of the distance between $\hat{x}$ and x in (15) satisfies

$$\begin{aligned}
\mathfrak{D}^+ d(\hat{x},x) \;&\leq\; \frac{d}{dt}\, L(\Gamma(\,\bullet\,,t))\Big|_0^{\hat{s}}\Big|_{t=0} \\
&\leq\; \frac{d\gamma^*}{ds}(\hat{s})^\top P(\gamma^*(\hat{s}))\, F(\gamma^*(\hat{s}),y) \\
&\qquad -\frac{d\gamma^*}{ds}(0)^\top P(\gamma^*(0))\, f(\gamma^*(0)) \;.
\end{aligned} \tag{36}$$

Even though (36) is an inequality condition, we proceed as if it were an equality. In such a case, if the observer makes the distance $d(\hat{x},x)$ nonincreasing along solutions then necessarily the right-hand side of (36) has to be nonpositive. To get a better understanding of what this means, consider the case when[6]

$$-\frac{d\gamma^*}{ds}(0)^\top P(\gamma^*(0))\, f(\gamma^*(0)) \;\geq\; 0 \;. \tag{37}$$

Then, for the right-hand side of (36) to be nonpositive, with $\hat{x}=\gamma^*(\hat{s})$, we must have

$$\frac{d\gamma^*}{ds}(\hat{s})^\top P(\hat{x})F(\hat{x},y) \;\leq\; 0 \;. \tag{38}$$

At this point, it is important to note that $\frac{d\gamma^*}{ds}(\hat{s})$ is the direction in which the state estimate $\hat{x}$ "sees" the system state x along a minimal geodesic. Such a direction is unknown to the observer. The only known information is that, for given y, x belongs to the following y-level set[7] of the output function:

$$\mathfrak{H}(y) \;=\; \{x : h(x)=y\} \;.$$

[6] For a given $x\in\mathbb{R}^n$, this condition holds for every minimal geodesic γ^* such that $\frac{d\gamma^*}{ds}(0)$ belongs to the closed half space $\{w\in\mathbb{R}^n : w^\top P(x)f(x)\leq 0\}$.

[7] By y-level set of h we mean the intersection, for each $i=1,2,\ldots,m$, of the sets $\{x\in\mathbb{R}^n : h_i(x)=y_i\}$.

Hence, (38) implies the following property: given $\hat{x}$ and y, the level set of the output function $\mathfrak{H}(y)$ is "seen" from $\hat{x}$ along a minimal geodesic, within a cone whose aperture is less than π. As stated in Lemma 2.7 below, this property implies that $\mathfrak{H}(y)$ is geodesically convex; see [23, Definition 6.1.1] and [11, Section 9.4].

Definition 2.6 (geodesic convexity): A subset S of $\mathbb{R}^n$ is said to be geodesically convex if, for any pair of points $(x_1, x_2) \in S \times S$, there exists a minimal geodesic γ^* between $x_1 = \gamma^*(s_1)$ and $x_2 = \gamma^*(s_2)$ satisfying

$$\gamma^*(s) \in S \qquad \forall s \in [s_1, s_2].$$

Lemma 2.7: Let $P : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be a complete Riemannian metric. Assume S is a subset of $\mathbb{R}^n$ such that, for any $\hat{x}$ in $\mathbb{R}^n \setminus S$, there exists a unit vector $v_{\hat{x}}$ such that, for any x in S and any minimal geodesic γ^* between $x = \gamma^*(0)$ and $\hat{x} = \gamma^*(\hat{s})$, with $\hat{s} > 0$, we have

$$\frac{d\gamma^*}{ds}(\hat{s})^\top P(\hat{x})\, v_{\hat{x}} \;<\; 0 \;.$$

Then, S is geodesically convex.

Proof: Assume that S is not geodesically convex. Then, there is a pair $(x_1, x_2) \in S$ such that, for any minimal geodesic γ_1^* between $x_1 = \gamma_1^*(0)$ and $x_2 = \gamma_1^*(s_2)$, there exists $\hat{s}_1$ in $(0, s_2)$ for which $\gamma_1^*(\hat{s}_1)$ is not in S. Let $\hat{x} = \gamma_1^*(\hat{s}_1) \notin S$. Note that $\gamma_2^*(s) = \gamma_1^*(s_2 - s)$ defines a minimal geodesic between $x_2 = \gamma_2^*(0) \in S$ and $\hat{x} = \gamma_2^*(\hat{s}_2) \notin S$, with $\hat{s}_2 = s_2 - \hat{s}_1 > 0$. With our assumption, since x_1 and x_2 are in S, there exists a unit vector $v_{\hat{x}}$ satisfying

$$\frac{d\gamma_1^*}{ds}(\hat{s}_1)^\top P(\hat{x})\, v_{\hat{x}} \;<\; 0 \quad , \qquad \frac{d\gamma_2^*}{ds}(\hat{s}_2)^\top P(\hat{x})\, v_{\hat{x}} \;<\; 0 \;.$$

But this impossible since we have $\frac{d\gamma_1^*}{ds}(\hat{s}_1) \;=\; -\frac{d\gamma_2^*}{ds}(\hat{s}_2)$. ■

For Example 1.1, we shall see in the following section that, with the help of item 2a of Proposition A.3, for any y, the level set $\mathfrak{H}(y) = \{(x_1, x_2) \;:\; x_1 = y\}$ is geodesically convex for the Riemannian metric given in (35).

As announced above, we conclude from Lemma 2.7 that geodesic convexity of the levels sets of the output function is a necessary property in the "general situation" where (37) holds (and when (36) is an equality). Actually, it is necessary, without any extra condition, when the observer has an infinite gain margin.

Definition 2.8 (infinite gain margin): The observer $\dot{\hat{x}} = F(\hat{x}, y)$ for $\dot{x} = f(x)$ is said to have an infinite gain margin with respect to P if (24) holds for every $x \in \mathbb{R}^n$ and, for any geodesic γ^* minimal on $[0, \hat{s})$, we have

$$\frac{d\gamma^*}{ds}(s)P(\gamma^*(s))\,[F(\gamma^*(s), h(\gamma^*(0)) - f(\gamma^*(s))] \;<\; 0 \tag{39}$$

for all $s \in (0, \hat{s})$.

The term *infinite gain margin* follows from the fact that, if the observer $\dot{\hat{x}} = F(\hat{x}, y)$ makes $t \mapsto d(\hat{X}((\hat{x}, x), t), X(x, t))$ nonincreasing (for each solution) and (39) holds, then the same holds for the observer $\dot{\hat{x}} = f(\hat{x}) + \ell\,[F(\hat{x}, y) - f(\hat{x})]$ for any real number $\ell > 1$.

D. Necessity of Uniform Detectability

The necessary condition in (20) is linked to an observability property of the family of linear time-varying systems obtained from linearizing (1) along its solutions. Assuming the system (1) is forward complete, for each x, the corresponding solution to (1) $t \mapsto X(x,t)$ is defined on $[0,+\infty)$. For each x, the linearization of f and h evaluated along a solution $X(x,t)$ gives the following functions defined on $[0,+\infty)$

$$A_x(t) \;=\; \frac{\partial f}{\partial x}(X(x,t)) \;, \quad C_x(t) \;=\; \frac{\partial h}{\partial x}(X(x,t)).$$

These functions define the following family of linear time-varying systems with state $\xi \in \mathbb{R}^n$ and output $\eta \in \mathbb{R}^m$:

$$\dot{\xi} \;=\; A_x(t)\,\xi \;, \quad \eta \;=\; C_x(t)\,\xi. \tag{40}$$

Systems (40) are parameterized by the initial condition x of the chosen solution $X(x,t)$.

The following theorem establishes a relationship between a detectability property of (40) and the existence of a bounded away from zero, upper bounded symmetric covariant two-tensor whose Lie derivative satisfies (20).

Theorem 2.9: Assume system (1) *is forward complete and that there exist a* C^1 *symmetric covariant two-tensor* $P:\mathbb{R}^n \to \mathbb{R}^{n\times n}$ *and strictly positive real numbers* $\underline{p}$ *and* $\overline{p}$ *satisfying* (20) *and*

$$0 \;<\; \underline{p}\,I \;\leq\; P(x) \leq\; \overline{p}\,I, \quad \forall x \in \mathbb{R}^n. \tag{41}$$

Then, for each $x \in \mathbb{R}^n$*, there exists a continuous*[8] *function* $t \in [0,+\infty) \to K_x(t)$ *such that the origin of the linear time-varying system*

$$\dot{\xi} \;=\; (A_x(t) - K_x(t) C_x(t))\;\xi \tag{42}$$

is uniformly exponentially stable.

Proof: To any $x \in \mathbb{R}^n$, we associate the functions $\Pi_x : [0,+\infty) \to \mathbb{R}^{n\times n}$, $K_x : [0,+\infty) \to \mathbb{R}^n$, and $\mathcal{V}_x : \mathbb{R}^n \times [0,+\infty) \to \mathbb{R}$ defined as

$$\begin{aligned} \Pi_x(t) \;&=\; P(X(x,t)), \quad \mathcal{V}_x(\xi,t) \;=\; \xi^\top \Pi_x(t)\xi \;, \\ K_x(t) \;&=\; \frac{\rho(X(x,t))}{2}\,\Pi_x(t)^{-1}\,C_x(t)^\top \;. \end{aligned} \tag{43}$$

We have

$$\underline{p}\,|\xi|^2 \;\leq\; \mathcal{V}_x(\xi,t) \;\leq\; \overline{p}\,|\xi|^2 \qquad \forall (x,t,\xi) \tag{44}$$

[8]We do not ask the function K_x to be bounded.

and, with (20), (18), (14), and the definitions in (43), we get

$$\begin{aligned}
\frac{d}{dt}\left(v^\top \Pi_x(t)\, v\right) &= \frac{\partial}{\partial \chi}\left(v^\top P(\chi)\, v\right)\, f(\chi)\Big|_{\chi = X(x,t)} , \\
&\leq -\frac{\varepsilon}{2} v^\top \Pi_x(t)\, v \\
&\quad - 2\, v^\top \Pi_x(t)\left(A_x(t) - K_x(t) C_x(t)\right) v \ .
\end{aligned}$$

Then, with (42), we have $\frac{d}{dt}\mathcal{V}_x(\xi,t) \ \leq\ -\frac{\varepsilon}{2}\mathcal{V}_x(\xi,t)$. The conclusion follows with (44). ■

It follows from this proof that, if we do not have the upper bound $\overline{p}$ in (41), we still have exponential stability, but we loose the uniformity property. This would be the case, for instance, for the system (5) of Example 1.1 with P given by (35) whose eigenvalues satisfy

$$\begin{aligned}
\lambda_{\min}(P(x)) \quad &\geq \quad \frac{(2+x_2^2)(1+x_1^2)-(x_1x_2-1)^2}{3+x_2^2+x_1^2} = \\
&\qquad \frac{1+x_1^2+(x_1+x_2)^2}{3+x_2^2+x_1^2}\ \geq\ \frac{1}{3},
\end{aligned} \tag{45}$$

$$\lambda_{\max}(P(x)) \quad \leq \quad 3 + x_2^2 + x_1^2. \tag{46}$$

Exponential stability of the origin of (42) is a detectability property for (40). The necessity of this property for the existence of P can be exploited to actually construct it, as it will be shown in the companion paper.

III. A Sufficient Condition

In the previous section, we assumed the existence of an observer making the function $t \mapsto d(\hat{X}((\hat{x},x),t), X(x,t))$ nonincreasing (respectively, strictly decreasing) with d being the distance associated with a Riemannian metric P. We showed that P has to satisfy a (respectively, strict) inequality involving the output function. In this section, we start from the data of such a metric and investigate the possibility of designing an observer making the corresponding Riemannian distance $d(\hat{x},x)$ strictly decreasing along solutions.

In view of Theorem 2.3, we assume that P satisfies

$$\mathcal{L}_f P(x) \ \leq\ \rho(x)\, \frac{\partial h}{\partial x}(x)^\top \frac{\partial h}{\partial x}(x) - q\, P(x) \qquad \forall x \in \mathbb{R}^n$$

with q a strictly positive real number. But, also, willing to be in a "general situation" in which (37) holds and motivated by Lemma 2.7, we restrict our attention to the case where the level set of the output function $\mathfrak{H}(y)$ is geodesically convex for any y in $\mathbb{R}^m$. Actually, we ask for the stronger (see Proposition A.3) property that the sets $\mathfrak{H}(y)$ are totally geodesic (see [6, Section V.II]).

Definition 3.1 (totally geodesic set): Given a C^1 function $\varphi : \mathbb{R}^n \mapsto \mathbb{R}^m$ and a closed subset $\mathcal{C}$ of $\mathbb{R}^n$, the set

$$S \ =\ \{x \in \mathbb{R}^n : \varphi(x) = 0\} \cap \mathcal{C}$$

is said to be totally geodesic if, for any pair (x,v) in $S\times\mathbb{R}^n$ such that

$$\frac{\partial\varphi}{\partial x}(x)\,v \;=\; 0 \quad , \qquad v^\top P(x)\,v \;=\; 1\ ,$$

any geodesic γ with

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

satisfies

$$\varphi(\gamma(s)) \;=\; 0 \qquad \forall s\in J_\gamma\ ,$$

where J_γ is the maximal interval containing 0 so that $\gamma(J_\gamma)$ is contained in $\mathcal{C}$.

In the appendix, we establish a necessary and sufficient checkable condition for the sets $\mathfrak{H}(y)$ to be totally geodesic.

Example 3.2 (Motivational example – continued): For the system in Example 1.1, it is sufficient to check that the Christoffel symbol Γ^1_{22} (see (67)) associated with the particular choice of P in (35) for the family (34) is zero. In fact, we have $\Gamma^1_{22}=\frac{1}{1+x_1^2+(x_1+x_2)^2}\begin{pmatrix}1+x_1^2 & 1-x_1x_2\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix}=0.$ □

The following theorem gives a sufficient condition for the existence of an observer for the single output case.

Theorem 3.3: Assume there exist a complete C^2 Riemannian metric P and a set $\mathcal{C}\subset\mathbb{R}^n$ such that

H1 : *$\mathcal{C}$ is geodesically convex, closed, and with nonempty interior;*

H2 : *there exist a C^1 function $\rho:\mathbb{R}^n\to[0,+\infty)$ and a strictly positive real number q such that*

$$\mathcal{L}_f P(x) \;\le\; \rho(x)\,\frac{\partial h}{\partial x}(x)^\top\frac{\partial h}{\partial x}(x) - q\,P(x) \qquad \forall x\in\mathcal{C}, \tag{47}$$

H3 : *The number of outputs is $m=1$ and, for each y in $h(\mathcal{C})$, the set $\mathfrak{H}(y)\cap\mathcal{C}$ is totally geodesic.*

Then, for any positive real number E there exists a continuous function $k_E:\mathbb{R}^n\to\mathbb{R}$ such that, with the observer given by

$$F(\hat{x},y) \;=\; f(\hat{x}) \;-\; k_E(\hat{x})\,P(\hat{x})^{-1}\frac{\partial h}{\partial x}(\hat{x})^\top\frac{\partial\delta}{\partial y_1}(h(\hat{x}),y)\ , \tag{48}$$

where

$$\delta(y_1,y_2) \;=\; |y_1-y_2|^2\ , \tag{49}$$

the following holds (see (15)):

$$\begin{array}{l}\mathfrak{D}^+ d(\hat{x},x) \quad\le\quad -\dfrac{q}{4}\,d(\hat{x},x)\\ \quad\forall(x,\hat{x})\in\{(x,\hat{x})\;:\;d(\hat{x},x)<E\}\bigcap\;(\mathrm{int}(\mathcal{C})\times\mathrm{int}(\mathcal{C}))\ \ .\end{array} \tag{50}$$

Moreover, expression (48) is intrinsic (i.e., coordinate independent) and gives an observer with infinite gain margin.

Example 3.4 (Motivational example – continued): We have already checked that, for the system (5) and with P given in (35) all the conditions of Theorem 3.3 hold globally, i.e., with $\mathcal{C}=\mathbb{R}^2$. Hence, the observer given by (48) becomes

$$\begin{pmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{pmatrix} = \begin{pmatrix} \hat{x}_2\sqrt{1+\hat{x}_1^2} \\ -\dfrac{\hat{x}_1\hat{x}_2^2}{\sqrt{1+\hat{x}_1^2}} \end{pmatrix} - \frac{2k_E(\hat{x})}{1+\hat{x}_1^2+(\hat{x}_1+\hat{x}_2)^2}\begin{pmatrix} 1+\hat{x}_1^2 \\ 1-\hat{x}_1\hat{x}_2 \end{pmatrix}(\hat{x}_1-y) \ .$$

□

Remark 3.5:

- Theorem 3.3 gives a (nonglobal) solution to problem ($\star$). When the assumptions of Theorem 3.3 hold globally, i.e., they hold for $\mathcal{C} = \mathbb{R}^n$, the observer given by (48) guarantees convergence of the estimated state to the system state, semiglobally with respect to the zero estimation error set $\mathcal{A}$.
  The fact that we do not get global asymptotic stability is likely due to the elementary form of the observer (48) and its infinite gain margin. We expect that other choices for this observer are possible to obtain a global asymptotic stability result.
- As discussed in II-B, we do not claim in Theorem 3.3 that the flow generated by the observer has a contraction property but simply that the Riemannian distance between estimated state and system state decays along the solutions. In other words, this result establishes that the function $(\hat{x}, x) \mapsto d(\hat{x}, x)$ can be used as a Lyapunov function for the zero error set $\mathcal{A}$ and guarantees this function has an exponential decay along the solutions. But it does no say that $d(\hat{x}_1, \hat{x}_2)$ decays along two arbitrary solutions of the flow generated by the observer.□

Theorem 3.3 is a direct consequence of the following lemma (for which there is no restriction on the number of outputs) and the fact that, when the number of outputs is $m = 1$, assumption H3 implies the assumption H3' of the lemma; see Proposition A.3.

Lemma 3.6: Assume there exist a complete C^2 Riemannian metric P, a set $\mathcal{C} \subset \mathbb{R}^n$, a C^1 function $\rho : \mathbb{R}^n \to [0, +\infty)$, and a strictly positive real number q satisfying H1 and H2 of Theorem 3.3. Assume also there exists a C^3 function $\delta : \mathbb{R}^m \times \mathbb{R}^m \to [0, +\infty)$ satisfying

$$\delta(h(x), h(x)) = 0 \quad , \qquad \frac{\partial^2 \delta}{\partial y_1^2}(y_1, y_2)\Big|_{y_1=y_2=h(x)} > 0 \tag{51}$$

for all $x \in \mathcal{C}$, and, such that

H3': *for any pair (x_1, x_2) in $\mathcal{C} \times \mathcal{C}$ satisfying*

$$h(x_1) \neq h(x_2)$$

and for any minimal geodesic γ^* between $x_1 = \gamma^*(s_1)$ and $x_2 = \gamma^*(s_2)$ satisfying $\gamma^*(s) \in \mathcal{C}$ for all

$s \in [s_1, s_2]$, *with* $s_1 \leq s_2$, *we have*

$$\frac{d}{ds}\delta(h(\gamma^*(s)), h(\gamma^*(s_1))) \ > \ 0 \qquad \forall s \in (s_1, s_2] \ . \tag{52}$$

Then, the claim of Theorem 3.3 holds true with a function δ *satisfying H3' (instead of* δ *as in (49)).*

Remark 3.7:

- Property H3' says that we can find a "distance-like" function δ in the output space allowing us to express that the output function h preserves some kind of monotonicity. Namely, as the distance increases along a geodesic in the state space, the same holds in the output space measured by δ. This property has some relationship with the notions of metric-monotone function introduced in [22] and of geodesically monotone function defined in [23, Definition 6.2.3]. In the appendix, we establish a connection with totally geodesic sets and geodesic convexity.

  With such a property, by following a descent direction for the "distance" in the output space, we are guaranteed to decrease the distance in the state space. This feature is exploited in the observer given by (48) via a high-gain term which enforces that such a descent direction is dominating.
- Property H3' with $\delta(y_1, y_2) = |y_1 - y_2|^2$ has been invoked already in [26] but for the case when P is constant. $\square$

Proof: Note that since we have

$$\hat{x} \ = \ x \qquad \Rightarrow \qquad F(\hat{x}, y) \ = \ f(\hat{x}) \ = \ f(x),$$

the result already holds when $d(x, \hat{x})$ is zero. Therefore, the remainder of the proof only considers pairs $(\hat{x}, x)$ that are in $(\mathcal{C} \times \mathcal{C}) \setminus \mathcal{A}$.

The Riemannian metric P being complete, any geodesic is defined on $(-\infty, +\infty)$ and the Riemannian distance $d(x_1, x_2)$ is given by the length of a minimal geodesic γ^* between x_1 and x_2. Since $\mathcal{C}$ is geodesically convex by H1, for any pair (x_1, x_2) in $(\mathcal{C} \times \mathcal{C}) \setminus \mathcal{A}$, there exists a minimal geodesic γ^* between $x_1 = \gamma^*(s_1)$ and $x_2 = \gamma^*(s_2)$ satisfying $\gamma^*(s) \in \mathcal{C}$ for all $s \in [s_1, s_2]$.

Let $(\hat{x}, x)$ be any pair in $(\mathcal{C} \times \mathcal{C}) \setminus \mathcal{A}$ and γ^* denote a minimal geodesic between $x = \gamma^*(0)$ and $\hat{x} = \gamma^*(\hat{s})$ satisfying $\gamma^*(s) \in \mathcal{C}$ for all $s \in [0, \hat{s}]$. With $y = h(x)$, take F as in (48). It gives

$$\begin{aligned}
&\frac{d\gamma^*}{ds}(\hat{s})^\top P(\gamma^*(\hat{s}))\,[F(\gamma^*(\hat{s}), y) - f(\gamma^*(\hat{s}))] \\
&\qquad\qquad - \frac{d\gamma^*}{ds}(0)^\top P(\gamma^*(0))\,[F(\gamma^*(0), y) - f(\gamma^*(0))] \\
&\qquad\qquad\qquad\qquad = \ -k_E(\hat{x})\,\frac{d\,h\circ\gamma^*}{ds}(\hat{s})^\top \frac{\partial\delta}{\partial y_1}(h(\gamma^*(\hat{s})), y).
\end{aligned} \tag{53}$$

On the other hand, we have

$$\frac{d\gamma^*}{ds}(\hat{s})^\top P(\hat{x})\,f(\hat{x}) - \frac{d\gamma^*}{ds}(0)^\top P(x)\,f(x) \tag{54}$$

$$= \int_0^{\hat{s}} \frac{d}{ds}\left(\frac{d\gamma^*}{ds}(s)^\top P(\gamma^*(s))\, f(\gamma^*(s))\right) ds \ .$$

Also the Euler-Lagrange form of the geodesic equation reads, for the i-th coordinate,

$$2\,\frac{d}{ds}\left(\sum_k P_{ik}(\gamma^*(s))\frac{d\gamma_k^*}{ds}(s)\right) \;=\; \sum_{k,l}\frac{d\gamma_k^*}{ds}(s)^\top\,\frac{\partial P_{kl}}{\partial x_i}(\gamma^*(s))\,\frac{d\gamma_l^*}{ds}(s)\ .$$

Then, with the definition of the Lie derivative $\mathcal{L}_f P$ and (47), we get

$$\begin{aligned}
&\frac{d}{ds}\left(\frac{d\gamma^*}{ds}(s)^\top P(\gamma^*(s))\, f(\gamma^*(s))\right)\\
&\qquad = \frac{1}{2}\,\frac{d\gamma^*}{ds}(s)^\top \mathcal{L}_f P(\gamma^*(s))\frac{d\gamma^*}{ds}(s)\ ,\\
&\qquad \leq \frac{\rho(\gamma^*(s))}{2}\left|\frac{\partial h}{\partial x}(\gamma^*(s))\frac{d\gamma^*}{ds}(s)\right|^2\\
&\qquad\qquad - \frac{q}{2}\,\frac{d\gamma^*}{ds}(s)^\top P(\gamma^*(s))\frac{d\gamma^*}{ds}(s)\\
&\qquad \leq \frac{\rho(\gamma^*(s))}{2}\left|\frac{d\,h\circ\gamma^*}{ds}(s)\right|^2 - \frac{q}{2}\ ,
\end{aligned} \tag{55}$$

where, in the last inequality, we have used

$$\frac{d\gamma^*}{ds}(s)^\top P(\gamma^*(s))\frac{d\gamma^*}{ds}(s) = 1$$

since γ^* is normalized. With $d(\hat{x},x) = \hat{s}$ as given in (12), replacing (55) into (54) yields

$$\begin{aligned}
&\frac{d\gamma^*}{ds}(\hat{s})^\top P(\gamma^*(\hat{s}))\, f(\gamma^*(\hat{s})) - \frac{d\gamma^*}{ds}(0)^\top P(\gamma^*(0))\, f(\gamma^*(0))\\
&\qquad \leq \int_0^{\hat{s}} \frac{\rho(\gamma^*(s))}{2}\left|\frac{d\,h\circ\gamma^*}{ds}(s)\right|^2 ds \;-\; \frac{q}{2}\, d(\hat{x},x)\ .
\end{aligned} \tag{56}$$

Then, from (36), using (53) and (56), we obtain

$$\begin{aligned}
&\mathfrak{D}^+ d(\hat{x},x)\\
&\qquad \leq \left[\frac{d\gamma^*}{ds}(\hat{s})^\top P(\gamma^*(\hat{s}))\,(F(\gamma^*(\hat{s}),y) - f(\gamma^*(\hat{s})))\right.\\
&\qquad\qquad \left. -\frac{d\gamma^*}{ds}(0)^\top P(\gamma^*(0))\,(F(\gamma^*(0),y) - f(\gamma^*(0)))\right]\\
&\qquad\quad + \left[\frac{d\gamma^*}{ds}(\hat{s})^\top P(\gamma^*(\hat{s})) f(\gamma^*(\hat{s})) - \frac{d\gamma^*}{ds}(0)^\top P(\gamma^*(0)) f(\gamma^*(0))\right]\\
&\leq -\,k_E(\hat{x})\,\frac{d\,h\circ\gamma^*}{ds}(\hat{s})^\top\,\frac{\partial\delta}{\partial y_1}(h(\gamma^*(\hat{s})),y)^\top\\
&\qquad + \int_0^{\hat{s}} \frac{\rho(\gamma^*(s))}{2}\left|\frac{d\,h\circ\gamma^*}{ds}(s)\right|^2 ds - \frac{q}{2}\, d(\hat{x},x)\ .
\end{aligned} \tag{57}$$

To proceed it is appropriate to associate two functions a and b to any triple $(\hat{x},x,\gamma^*)$ with $(\hat{x},x)$ in $(\mathcal{C}\times\mathcal{C})\setminus\mathcal{A}$

and γ^*, a minimal geodesic between $x = \gamma^*(0)$ and $\hat{x} = \gamma^*(\hat{s})$ satisfying $\gamma^*(s) \in \mathcal{C}$ for all $s \in [0, \hat{s}]$. These functions are defined on $[0, \hat{s}]$ as follows:[9]

$$a_{(\hat{x},x,\gamma^*)}(r) \quad = \quad \frac{1}{r}\, \frac{d\, h\circ\gamma^*}{ds}(r)^\top \frac{\partial \delta}{\partial y_1}(h(\gamma^*(r)), h(\gamma^*(0)))^\top$$

if $0 < r \leq \hat{s}$, and

$$a_{(\hat{x},x,\gamma^*)}(0) \quad = \quad \frac{d\, h\circ\gamma^*}{ds}(0)^\top \frac{\partial^2 \delta}{\partial y_1^2}(h(\gamma^*(0)), h(\gamma^*(0)))^\top \frac{d\, h\circ\gamma^*}{ds}(0);$$

and

$$b_{(\hat{x},x,\gamma^*)}(r) \quad = \quad \frac{1}{r}\int_0^r \frac{\rho(\gamma^*(s))}{2}\left|\frac{d\, h\circ\gamma^*}{ds}(s)\right|^2 ds$$

if $0 < r \leq \hat{s}$, and

$$b_{(\hat{x},x,\gamma^*)}(0) \quad = \quad \frac{\rho(\gamma^*(0))}{2}\left|\frac{d\, h\circ\gamma^*}{ds}(0)\right|^2 .$$

We remark with (51) that δ reaches its global minimum at $y_1 = y_2 = h(x)$. This implies

$$\frac{\partial \delta}{\partial y_1}(h(\gamma^*(r)), h(\gamma^*(0))) \quad = \quad \left[\int_0^1 \left(\frac{\partial^2 \delta}{\partial y_1^2}(h(\gamma^*(\sigma r)), \gamma^*(0)) \frac{d\, h\circ\gamma^*}{ds}(\sigma r)\right) d\sigma\right] r$$

for all $r \in [0, \hat{s}]$. As a consequence, the functions a an b are continuous on $[0, \hat{s}]$. Moreover the property H3' gives readily the implication

$$h(x) \;\neq\; h(\hat{x}) \qquad \Longrightarrow \qquad a_{(\hat{x},x,\gamma^*)}(r) \;>\; 0 \quad \forall r \in (0, \hat{s}] \ .$$

In the case when $h(x) = h(\hat{x})$, we are only left with the following two possibilities:

1) $h\circ\gamma^*$ is constant on $[0, \hat{s}]$. Then we have $\frac{dh\circ\gamma^*}{ds}(s) = 0$ for all $s \in [0, \hat{s}]$ and therefore $a_{(\hat{x},x,\gamma^*)}(r) = b_{(\hat{x},x,\gamma^*)}(r) = 0$ for all $r \in [0, \hat{s}]$.
2) $h\circ\gamma^*$ is not constant on $[0, \hat{s}]$. Then, there exists some s_1 in $(0, \hat{s}]$ such that $h(\gamma(s_1)) \neq h(\gamma^*(0)) = h(x)$. With H3', this implies that the function $s \mapsto \delta(h(\gamma^*(s)), h(\gamma^*(0)))$ is not constant on $[0, \hat{s}]$. But since we have $\delta(h(\gamma^*(\hat{s})), h(\gamma^*(0))) = \delta(h(\gamma^*(0)), h(\gamma^*(0))) = 0$, this function must reach a maximum at some point s_m in $(0, \hat{s})$ where we have
$$\delta(h(\gamma^*(s_m)), h(\gamma^*(0))) > 0,$$
$$\frac{d}{ds}\delta(h(\gamma^*(s_m)), h(\gamma^*(0))) = 0,$$

[9]When $\hat{s} = 0$ the functions $a_{(\hat{x},x,\gamma^*)}$ and $b_{(\hat{x},x,\gamma^*)}$ are only defined at zero.

and therefore $h(\gamma^*(s_m)) \neq h(\gamma^*(0))$. But this contradicts H3'. So this case is impossible.

In any case, we have established that $a_{(\hat{x},x,\gamma^*)}(\hat{s})$ is non negative and if it is zero then $b_{(\hat{x},x,\gamma^*)}(r) = 0$ for all $r \in [0,\hat{s}]$.

Now, let $\check{x}$ be an arbitrary point in $\mathcal{C}$. Call it origin. For each integer i, we introduce the set

$$\mathcal{K}_i = \{(x,\hat{x}) \in \mathcal{C}\times\mathcal{C} : d(\hat{x},x) \leq E\ ,\ i \leq d(\check{x},\hat{x}) \leq i+1\}\ .$$

From the Hopf-Rinow Theorem [24, Theorem II.1.1] $\mathcal{K}_i$ is compact.

To conclude it is sufficient to prove the existence of a real number k_i such that, for any pair $(\hat{x},x)$ in $\mathcal{K}_i \setminus \mathcal{A}$ and any minimal geodesic γ^* between $x=\gamma^*(0)$ and $\hat{x}=\gamma^*(\hat{s})$ satisfying $\gamma^*(s) \in \mathcal{C}$ for all $s \in [0,\hat{s}]$, we have

$$\frac{q}{4} + k_i\, a_{(\hat{x},x,\gamma^*)}(\hat{s}) > b_{(\hat{x},x,\gamma^*)}(\hat{s}).$$

Indeed, with this inequality, the definitions of a and b and (57) where $d(\hat{x},x)=\hat{s}$, we obtain (50) provided the function k_E satisfies

$$k_E(\hat{x}) \geq k_i \qquad \forall \hat{x} \in \mathcal{C} : i \leq d(\hat{x},x) \leq i+1\ .$$

Proceeding by contradiction, suppose that such k_i does not exist. Then, there exists a sequence $(\hat{s}_n, x_n, \hat{x}_n, \gamma_n^*)$, with $\hat{s}_n \geq 0$, $(x_n,\hat{x}_n)$ in $\mathcal{K}_i \setminus \mathcal{A}$, and γ_n^* a minimal geodesic between $x_n = \gamma_n^*(0)$ and $\hat{x}_n = \gamma_n^*(\hat{s}_n)$ satisfying $\gamma_n^*(s) \in \mathcal{C}$ for all $s \in [0,\hat{s}_n]$ and

$$\frac{q}{4} + n\, a_{(\hat{x}_n,x_n,\gamma_n^*)}(\hat{s}_n) \leq b_{(\hat{x}_n,x_n,\gamma_n^*)}(\hat{s}_n)\ . \tag{58}$$

Moreover, the functions $a_{(\hat{x},x,\gamma^*)}$ and $b_{(\hat{x},x,\gamma^*)}$ are C^1 on $[0,\hat{s}]$. Indeed, they can be written as

$$a_{(\hat{x},x,\gamma^*)} = \frac{f_a(r)}{r}\ ,\quad a_{(\hat{x},x,\gamma^*)} = \frac{f_b(r)}{r} \forall r \in]0,\hat{s}]$$

where the function f_a, respectively f_b, is C^2 since h, γ^* and δ are C^3, respectively, ρ is C^1 and h and γ^* are C^2. We have the following technical property.

Claim 1: Let f be a C^2 function defined on a neighborhood of 0 in $\mathbb{R}$, where it is 0. The function φ defined as $\varphi(r) = \frac{f(r)}{r}$ if $r \neq 0$ and $\varphi(0) = f'(0)$ is C^1.

Proof : Clearly, φ is C^2 everywhere except may be at 0. Its first derivative is $\varphi'(r) = \frac{f(r)-rf'(0)}{r^2}$. It is also continuous at 0 since $\lim_{r\to 0}\varphi(r) = f'(0) = \varphi(0)$. Its first derivative at 0 exists if $\lim_{r\to 0}\frac{\varphi(r)-\varphi(0)}{r} = \lim_{r\to 0}\frac{f(r)-rf'(0)}{r^2}$ exists, which is the case since, due to f being C^2, we have

$$\begin{aligned}\frac{f(r)-rf'(0)}{r^2} &= \frac{1}{r^2}\int_0^r [f'(s)-f'(0)]ds \\ &= \frac{1}{r^2}\int_0^r\int_0^s f''(t)dtds \\ &= \frac{1}{r^2}\int_0^r f''(t)[r-t]dt\end{aligned}$$

which leads to $\varphi'(0) = \frac{1}{2} f''(0)$. We have also

$$\frac{f(r) - rf'(r)}{r^2} = -\frac{1}{r^2} \int_0^r s f''(s) ds$$

This implies

$$\lim_{r \to 0} \varphi'(r) = \varphi'(0)$$

and therefore φ' is continuous.

We also have the following claim.

Claim 2: There exists a subsequence $(\hat{s}_{n_1}, x_{n_1}, \hat{x}_{n_1}, \gamma^*_{n_1})$ *of* $(\hat{s}_n, x_n, \hat{x}_n, \gamma^*_n)$ *such that*

$$\lim_{n_1 \to \infty} (\hat{s}_{n_1}, x_{n_1}, \hat{x}_{n_1}) \; = \; (\hat{s}_\omega, x_\omega, \hat{x}_\omega), \tag{59}$$

$$\lim_{n_1 \to \infty} \gamma^*_{n_1}(s) \; = \; \gamma_\omega(s) \text{ uniformly in } s \in [0, E] \; , \tag{60}$$

where $\gamma_\omega : [0, \hat{s}_\omega] \to \mathcal{C}$ *is a minimal geodesic between* x_ω *and* $\hat{x}_\omega$*.*

To prove the claim, not that since $(x_n, \hat{x}_n)$ is in the compact set $\mathcal{K}_i$ and γ^*_n is a minimal geodesic taking values in $\mathcal{C}$ when restricted to $[0, \hat{s}_n]$, from

$$\sqrt{\underline{p}}\, |x_1 - x_2| \; \leq \; d(x_1, x_2) \; \leq \; \sqrt{\overline{p}}\, |x_1 - x_2| \qquad \forall (x_1, x_2) \in \mathcal{C} \times \mathcal{C} \; ,$$

we get

$$\sqrt{\underline{p}}\, |\gamma^*_n(s) - x_n| \; \leq \; d(\gamma^*_n(s), x_n) \; \leq \; \hat{s}_n \; \leq \; E \quad \forall s \in [0, \hat{s}_n]$$

and

$$\begin{aligned} |x_n| \; \leq \; |\hat{x}_n - x_n| \; + \; |\hat{x}_n| \quad &\leq \quad \frac{d(\hat{x}_n, x_n)}{\sqrt{\underline{p}}} \; + \; (i+1) \; , \\ &\leq \quad \frac{E}{\sqrt{\underline{p}}} \; + \; (i+1) \; . \end{aligned}$$

This implies that $\gamma^*_n : [0, E] \to \mathcal{C}$ takes its values in a compact set independent of the index n. Moreover, γ^*_n being a solution of the geodesic equation, there exists a subsequence with index n_1 and a quadruple $(\hat{s}_\omega, x_\omega, \hat{x}_\omega, \gamma_\omega)$ such that (59)-(60) hold (see, for instance, [9, Theorem 5, §1]), where γ_ω is a solution of the geodesic equation and, since $\mathcal{C}$ is closed, it satisfies

$$\gamma_\omega(0) \; = \; x_\omega \; , \quad \gamma_\omega(\hat{s}_\omega) \; = \; \hat{x}_\omega \; , \quad \gamma_\omega(s) \in \mathcal{C} \quad \forall s \in [0, \hat{s}_\omega] \; .$$

Finally, according to [24, Lemma II—.4.2], it is minimizing between x_ω and $\hat{x}_\omega$.

Now, the functions h, ρ and $\frac{\partial h}{\partial x}$ restricted to the compact set where the functions γ^*_n take their values, are continuous and bounded. Also, from the geodesic equation and completeness, the same holds for γ^*_n, $\frac{d\gamma^*_n}{ds}$ and $\frac{d^2\gamma^*_n}{ds^2}$ restricted to $[0, \hat{s}_n]$. With the definition of $b_{(\hat{x}_n, x_n, \gamma^*_n)}$, this implies that the right-hand side of (58) is upper bounded, say by B. Consequently, we have

$$\frac{q}{4} \;+\; n\, a_{(\hat{x}_n, x_n, \gamma_n^*)}(\hat{s}_n) \;\leq\; B \qquad \forall n.$$

Since $a_{(\hat{x}_n, x_n, \gamma_n^*)}(\hat{s}_n)$ is nonnegative, this implies that $a_{(\hat{x}_\omega, x_\omega, \gamma_\omega)}(\hat{s}_\omega) = 0$.

If $\hat{x}_\omega \neq x_\omega$, since $a_{(\hat{x}_\omega, x_\omega, \gamma_\omega)}(\hat{s}_\omega)$ is zero, we have seen that the same holds for $b_{(\hat{x}_\omega, x_\omega, \gamma_\omega)}(r)$, for all for all $r \in [0, \hat{s}_\omega]$. On the other hand, (58) yields

$$\frac{q}{4} \;\leq\; b_{(\hat{x}_\omega, x_\omega, \gamma_\omega^*)}(\hat{s}_\omega)$$

where q is strictly positive. So we have a contradiction.

If $\hat{x}_\omega = x_\omega$, also by compactness, there exists a subsequence with index n_2 of the subsequence with index n_1 in Claim 1 such that we have

$$v_\omega \;=\; \lim_{n_2 \to \infty} \frac{\hat{x}_{n_2} - x_{n_2}}{d(\hat{x}_{n_2}, x_{n_2})} \;=\; \lim_{n_2 \to \infty} \frac{\hat{x}_{n_2} - x_{n_2}}{\hat{s}_{n_2}}$$

Note that since $\hat{x}_\omega = x_\omega$, we have $\hat{s}_{n_1}$ (and also $\hat{s}_{n_2}$) converging to zero. But, with the identity

$$\hat{x}_{n_2} \;=\; x_{n_2} \;+\; \int_0^{\hat{s}_{n_2}} \frac{d\gamma_{n_2}^*}{ds}(s)ds\ ,$$

this gives also

$$v_\omega \;=\; \lim_{n_2 \to \infty} \frac{\gamma_{n_2}^*(\hat{s}_{n_2}) - \gamma_{n_2}^*(0)}{\hat{s}_{n_2}} \;=\; \frac{d\gamma_\omega^*}{ds}(0)$$

On the other hand, since the functions $a_{(\hat{x}_n, x_n, \gamma_n^*)}$ and $b_{(\hat{x}_n, x_n, \gamma_n^*)}$ are C^1 on $[0, E]$, and the way they depend on n is only via γ_n^* (which takes its values in a compact set independent of n), there exist real numbers A_1 and B_1 such that we have

$$\begin{aligned} a_{(\hat{x}_n, x_n, \gamma_n^*)}(\hat{s}_n) \quad &\geq \quad a_{(\hat{x}_n, x_n, \gamma_n^*)}(0) \;-\; A_1\, \hat{s}_n\ , \\ b_{(\hat{x}_n, x_n, \gamma_n^*)}(\hat{s}_n) \quad &\leq \quad b_{(\hat{x}_n, x_n, \gamma_n^*)}(0) \;+\; B_1\, \hat{s}_n\ . \end{aligned}$$

Since we are in the case where $\hat{s}_{n_1}$ goes to 0, this implies

$$\begin{aligned} 0 \quad &= \quad \lim_{n_2 \to \infty} a_{(\hat{x}_{n_2}, x_{n_2}, \gamma_{n_2}^*)}(0)\ , \\ &= \quad v_\omega^\top \frac{\partial h}{\partial x}(x_\omega)^\top \frac{\partial^2 \delta}{\partial y_1^2}(x_\omega, x_\omega)^\top \frac{\partial h}{\partial x}(x_\omega) v_\omega\ . \end{aligned}$$

With (51), we obtain

$$\frac{\partial h}{\partial x}(x_\omega) v_\omega \;=\; 0$$

and therefore :

$$\lim_{n_2 \to \infty} b_{(\hat{x}_{n_2}, x_{n_2}, \gamma_{n_2}^*)}(\hat{s}_{n_2})$$

$$\leq \lim_{n_2\to\infty} b_{(\hat{x}_{n_2},x_{n_2},\gamma^*_{n_2})}(0) \;=\; \frac{\rho(\gamma^*_\omega(0))}{2}\left|\frac{\partial h}{\partial x}(x_\omega)v_\omega\right|^2 \; , \\ = \; 0 \; .$$

This contradicts (58).

So we have stablished the existence of k_i.

Finally, in (53), we have, with (52),

$$\frac{d\,h\circ\gamma^*}{ds}(\hat{s})^\top \frac{\partial \delta}{\partial y_1}(h(\gamma^*(\hat{s})),y) \\ = \frac{d}{d\hat{s}}\delta(h(\gamma^*(\hat{s})),h(\gamma^*(s))) \; > \; 0$$

and $F(\gamma^*(0),y) \;=\; f(\gamma^*(0))$. So (39) holds and the observer has an infinite gain margin.

To prove the last point of Theorem 3.3, let ϕ define a diffeomorphism as in (33). Let $\bar{h}$, $\bar{k}_E$, $\bar{f}$, $\bar{F}$ and $\bar{P}$ be the expressions of h, k_E, f, F and P respectively in the new coordinates. We have (9), (34), and $\bar{k}_E(\bar{x}) = k_E(x)$, $\bar{F}(\bar{x},y) \;=\; \frac{\partial \phi}{\partial x}(x)\,F(x,y)$. This implies

$$\begin{aligned}
\bar{F}(\hat{\bar{x}},y) \;&=\; \frac{\partial \phi}{\partial x}(\hat{x})\left[f(\hat{x}) \;-\; k_E(\hat{x})\,P(\hat{x})^{-1}\frac{\partial h}{\partial x}(\hat{x})^\top \times \right. \\
&\qquad \left. \frac{\partial \delta}{\partial y_1}(h(\hat{x}),y)\right] \; , \\
&=\; \bar{f}(\hat{\bar{x}}) \;-\; \bar{k}_E(\bar{x})\left(\frac{\partial \phi}{\partial x}(\hat{x})P(\hat{x})^{-1}\frac{\partial \phi}{\partial x}(\hat{x})^\top\right) \times \\
&\qquad \left[\frac{\partial \phi}{\partial x}(\hat{x})^\top\right]^{-1}\frac{\partial h}{\partial x}(\hat{x})^\top \frac{\partial \delta}{\partial y_1}(h(\hat{x}),y) \; , \\
&=\; \bar{f}(\hat{\bar{x}}) \;-\; \bar{k}_E(\bar{x})\bar{P}(\hat{\bar{x}})^{-1}\frac{\partial \bar{h}}{\partial \bar{x}}(\hat{\bar{x}})^\top \frac{\partial \delta}{\partial y_1}(\bar{h}(\hat{\bar{x}},y)) \; .
\end{aligned}$$

Therefore, the expression of the observer remains the same after the change of coordinates. ■

IV. Conclusion

If for a Riemannian metric P and an observer such that the distance between estimated state and system state decreases along the solutions, then the Lie derivative of P along the systems solutions satisfies the inequality in Theorem 2.3 involving the output function. Also, the satisfaction of such an inequality together with the existence of upper and lower bounds for P (see (41)) imply detectability of the linear time-varying systems obtained from linearizing the given system (1) along its solutions. Moreover, we have seen how the geodesic convexity of the output function level sets is necessary if the observer has an infinite gain margin and, in a general situation, when the Riemannian distance between estimated state and system state decreases along the solutions of (3).

Conversely, from the data of a Riemannian metric satisfying the necessary conditions in Theorem 2.3 and (41), and when the level sets of the output function are totally geodesic, we showed how to construct, for the single output case, an observer guaranteeing convergence of the estimated state to the system state, semiglobally with respect to zero estimation error set $\mathcal{A}$.

Also, although in Section II we have given an expression of an observer, at this time, we consider this only as an existence result and not as an observer design interesting for application. Actually we have investigated mainly only the possibility and interest of studying observer convergence via a Riemannian metric, crystallizing the idea of using a contraction property. In a companion paper, we focus on observer design, where we study several scenarios in which it is possible to construct a Riemannian metric satisfying the desired inequality on its Lie derivative and making the level sets of the output function possibly totally geodesic.

As a final remark, we observe that extensions of the results to nonautonomous systems, in particular those with inputs, seem possible using the proof techniques proposed here. Also time scaling exploiting the concept of unbounded observability, as in [2], is expected to be useful in relaxing the system completeness assumption.

APPENDIX

A. A necessary condition for completeness

The following lemma provides conditions on P that guarantee that geodesics can be maximally extended to $\mathbb{R}$.

Lemma A.1: Suppose that a symmetric covariant two-tensor $P : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ *satisfies*

$$0 < P(x) \quad \forall x \in \mathbb{R}^n \ , \quad \lim_{r\to\infty} r^2 \underline{p}(r) = +\infty, \tag{61}$$

where, for any positive real number r, $\underline{p}(r) = \min_{x:|x|\leq r} \lambda_{\min}(P(x))$. *Then, with* P *as Riemannian metric on* $\mathbb{R}^n$, *any geodesic can be maximally extended to* $\mathbb{R}$.

Proof: Let x_1 and x_2 be any point in the ball B_r in $\mathbb{R}^n$ centered at the origin and with radius r. The Euclidean distance $|x_1 - x_2|$ satisfies $\int_{s_1}^{s_2} \left|\frac{d\gamma}{ds}(s)\right| ds \geq |x_1 - x_2|$, where γ is any piecewise C^1 path between x_1 and x_2. Using (10), this implies that, for any positive number r,

$$L(\gamma)\Big|_{s_1}^{s_2} \geq \sqrt{\underline{p}(r)} \int_{s_1}^{s_2} \left|\frac{d\gamma}{ds}(s)\right| ds \geq \sqrt{\underline{p}(r)}\, |x_1 - x_2| \, . \tag{62}$$

Let γ be any normalized geodesic maximally defined on (σ_-, σ_+). By definition, it satisfies

$$\frac{d\gamma}{ds}(s)^\top P(\gamma(s)) \frac{d\gamma}{ds}(s) = 1 \qquad \forall s \in (\sigma_-, \sigma_+) \ . \tag{63}$$

Let $[s_1, s_2]$ be any closed interval contained in (σ_-, σ_+). The function $\gamma : [s_1, s_2] \to \mathbb{R}^n$ is bounded (with the Euclidean norm). We denote $r_{[s_1,s_2]} = \max_{s\in[s_1,s_2]} |\gamma(s)|$. By continuity, there exists s_{12} in $[s_1, s_2]$ satisfying $r_{[s_1,s_2]} = |\gamma(s_{12})|$. Then, from (62) and (63), we obtain

$$\sqrt{\underline{p}(|\gamma(s_{12})|)}\, |\gamma(s_{12}) - \gamma(s_2)| \leq L(\gamma)\Big|_{s_{12}}^{s_2} = |s_{12} - s_2| \ . \tag{64}$$

Because (σ_-, σ_+) is the maximal interval of definition of γ, if σ_- is finite, we must have $\lim_{s_1\to\sigma_-} \left|\left(\gamma(s_1), \frac{d\gamma}{ds}(s_1)\right)\right| = +\infty$. Now in the case where we have $\lim_{s_1\to\sigma_-} |\gamma(s_1)| = +\infty$ the definition of s_{12} implies $\lim_{s_1\to\sigma_-} \max_{s\in[s_1,s_2]} |\gamma(s)| = \lim_{s_1\to\sigma_-} |\gamma(s_{12})| = +\infty$. Then, with assumption (61) and (64), we get

$$|\sigma_- - s_2| \quad \geq \quad \lim_{s_1 \to \sigma_-} \sqrt{\underline{p}(|\gamma(s_{12})|)}\, |\gamma(s_{12}) - \gamma(s_2)| \ \geq$$
$$\lim_{s_1 \to \sigma_-} \sqrt{\underline{p}(|\gamma(s_{12})|)}(|\gamma(s_{12})| - |\gamma(s_2)|) \ \geq \ +\infty \ .$$

This is a contradiction. Then, we are left with the case $\lim_{s_1 \to \sigma_-} \left|\frac{d\gamma}{ds}(s_1)\right| = +\infty$. But this contradicts (63) since we just established that γ is bounded on (σ_-, s_2), which, with (61), implies that $P \circ \gamma$ is bounded away from 0.

The same arguments apply to show that $\sigma_+ = +\infty$. ■

B. *On totally geodesic sets and property H3'*

Proposition A.2: Let P be a complete Riemannian metric on $\mathbb{R}^n$ and $\mathcal{C}$ be a geodesically convex subset of $\mathbb{R}^n$.

1) *If there exists x_0 in $\mathcal{C}$ satisfying $\frac{\partial h}{\partial x}(x_0) = 0$ and all the sets $\mathfrak{H}(y) \cap \mathcal{C}$ for y in $h(\mathcal{C})$ are totally geodesic then h is constant on $\mathcal{C}$.*
2) *Let $\mathcal{O}$ be the following open subset of $\mathbb{R}^n$:*

$$\mathcal{O} \ = \ \left\{ x \in \mathrm{int}(\mathcal{C}) \ : \ \mathrm{Rank}\left(\frac{\partial h}{\partial x}(x)\right) = m \right\} \ . \tag{65}$$

If all the sets $\mathfrak{H}(y) \cap \mathcal{C}$ for y in $h(\mathcal{C})$ are totally geodesic then we have, for all (j,k,l) and all $x \in \mathcal{O}$,

$$\begin{aligned} \frac{\partial^2 h_j}{\partial x_k \partial x_l}(x) - \sum_{i=1}^n \frac{\partial h_j}{\partial x_i}(x)\, \Gamma_{kl}^i(x) \quad &= \\ \sum_{i=1}^n \left(g_{jik}(x) \frac{\partial h_i}{\partial x_l}(x) + g_{jil}(x) \frac{\partial h_i}{\partial x_k}(x) \right) \ &, \end{aligned} \tag{66}$$

where $g_{jik} : \mathcal{O} \to \mathbb{R}$ are continuous arbitrary functions and Γ_{kl}^i are the Christoffel symbols

$$\begin{aligned} \Gamma_{kl}^i(x) \quad = \quad \frac{1}{2} \sum_{m=1}^n \left(P(x)^{-1}\right)_{im} &\left(\frac{\partial P_{mk}}{\partial x_l}(x) \right. \\ &\left. + \frac{\partial P_{ml}}{\partial x_k}(x) - \frac{\partial P_{kl}}{\partial x_m}(x) \right) . \end{aligned} \tag{67}$$

Conversely, if (66) holds for any x in $\mathcal{C}$, then all the sets $\mathfrak{H}(y) \cap \mathcal{C}$ for y in $h(\mathcal{C})$ are totally geodesic.

Proof of item 1: The set $\mathcal{C}$ being geodesically convex, for any x there exists a minimal geodesic γ^* between $x_0 = \gamma^*(0)$ and $x = \gamma^*(s)$ satisfying $\gamma^*(\sigma) \in \mathcal{C}\ \forall \sigma \in [0, s]$. Since we have $\frac{\partial h}{\partial x}(x_0) \frac{d\gamma^*}{ds}(0) = 0$ and the set $\mathfrak{H}(h(x_0)) \cap \mathcal{C}$ is totally geodesic, we get $h(x) = h(x_0)$, x being arbitrary in $\mathcal{C}$, h must be constant on $\mathcal{C}$.

Proof of item 2 (necessity): If $\mathcal{O}$ is empty, the statement holds vacuously. If $\mathcal{O}$ is nonempty, let x be in $\mathcal{O}$. It is in the totally geodesic set $\mathfrak{H}(h(x)) \cap \mathcal{C}$. Then, for any v in $\mathbb{R}^n$ satisfying

$$\frac{\partial h}{\partial x}(x)\, v \ = \ 0 \quad , \qquad v^\top P(x)\, v \ = \ 1 \ , \tag{68}$$

consider a geodesic γ satisfying

$$\gamma(0) \ = \ x \quad , \qquad \frac{d\gamma}{ds}(0) \ = \ v \tag{69}$$

with values in $\mathcal{C}$ on an interval (σ_-, σ_+). We have $h(\gamma(s)) = 0$ for all $s \in (\sigma_-, \sigma_+)$. This implies that we have

$$\frac{dh \circ \gamma}{ds}(0) \;=\; \frac{d^2 h \circ \gamma}{ds^2}(0) \;=\; 0\;. \tag{70}$$

But, with the geodesic equation, if we let $Q_{jkl}(x) = \frac{\partial^2 h_j}{\partial x_k \partial x_l}(x) - \sum_{i=1}^n \frac{\partial h_j}{\partial x_i}(x)\,\Gamma^i_{kl}(x)$, we have

$$\frac{d^2 h_j \circ \gamma}{ds^2}(s) \;=\; \sum_{k=1}^n \sum_{l=1}^n Q_{jkl}(\gamma(s)) \frac{d\gamma_k}{ds}(s) \frac{d\gamma_l}{ds}(s)\;. \tag{71}$$

Then, using (69) and (70), we have

$$\sum_{k=1}^n \sum_{l=1}^n Q_{jkl}(x) v_k v_l \;=\; 0 \qquad \forall j \in \{1, 2, \ldots, m\}, \tag{72}$$

where v_k is the kth component of v. Hence, we have established $\sum_{k=1}^n \sum_{l=1}^n Q_{jkl}(x) v_k v_l = 0$ for all $(j, v = (v_k), x) \,:\, j \in \{1, 2, \ldots, m\}\,,\ \frac{\partial h}{\partial x}(x) v = 0\,,\ x \in \mathcal{O}$. The result follows from the S-Lemma (see [12] for instance). In particular, we can pick the functions $g_{jik}(x)$ satisfying (66) as, for each j, the entries of the matrix

$$\begin{array}{l}\left[\dfrac{\partial h}{\partial x}(x) \dfrac{\partial h}{\partial x}(x)^\top\right]^{-1} \dfrac{\partial h}{\partial x}(x) Q_{j\bullet\bullet}(x) \times \\ \qquad \left(I - \dfrac{\frac{\partial h}{\partial x}(x)^\top [\frac{\partial h}{\partial x}(x) \frac{\partial h}{\partial x}(x)^\top]^{-1} \frac{\partial h}{\partial x}(x)}{2}\right)\;.\end{array}$$

Proof of item 2 (sufficiency): For any y in $h(\mathcal{C})$, let (x, v) be any pair in $(\mathfrak{H}(y) \cap \mathcal{C}) \times \mathbb{R}^n$ satisfying $h(x) = y$, $\frac{\partial h}{\partial x}(x)\, v = 0$, $v^\top P(x)\, v = 1$ and let γ be any geodesic satisfying $\gamma(0) = x$, $\frac{d\gamma}{ds}(0) = v$. Let J_γ be the maximal interval containing 0 so that $\gamma(J_\gamma)$ is contained in $\mathcal{C}$. If J_γ is reduced to a point, there is nothing to prove. If not J_γ is an interval with a non empty interior. Then, with (71) and (66), for any interior point s of J_γ, we have, for each j in $\{1, \ldots, m\}$,

$$\begin{aligned}\frac{d}{ds} \frac{dh_j \circ \gamma}{ds}(s) \;&=\; \sum_{k=1}^n \sum_{l=1}^n Q_{jkl}(\gamma(s)) \frac{d\gamma_k}{ds}(s) \frac{d\gamma_l}{ds}(s) \\ &=\; 2 \sum_{i=1}^n \left[\sum_{k=1}^n g_{jik}(\gamma(s)) \frac{d\gamma_k}{ds}(s)\right] \frac{dh_i \circ \gamma}{ds}(s)\;.\end{aligned}$$

Let M be the matrix with entries M_{ji} defined as, $M_{ji}(s) = 2\left[\sum_{k=1}^n g_{jik}(\gamma(s)) \frac{d\gamma_k}{ds}(s)\right]$, for each $s \in \text{int}(J_\gamma)$. The linear time varying system $\frac{dz}{ds} = M(s) z$ has unique solutions. The only one satisfying $z(0) = 0$ is identically 0. So with the uniqueness of the solution of the geodesic equation we must also have $\frac{dh_j \circ \gamma}{ds}(s) = 0\ \forall s \in \text{int}(J_\gamma)$ and therefore $h_j(\gamma(s)) = y_j$ for each $s \in \text{int}(J_\gamma)$ and each j. Also, by continuity, if the upper bound σ_+ (respectively lower bound σ_-) of J_γ is in J_γ, then we have also $h_j(\sigma_+) = y_j$ (respectively $h_j(\sigma_+) = y_j$).

Proposition A.3: Let P be a complete Riemannian metric on $\mathbb{R}^n$ and $\mathcal{C}$ be a geodesically convex subset of $\mathbb{R}^n$.

1) *If property H3' holds then all the sets $\mathfrak{H}(y) \cap \mathcal{C}$ for y in $h(\mathcal{C})$ are*

   a) *totally geodesic,*

   b) *and geodesically convex.*

2) *If $m = 1$ and all the sets $\mathfrak{H}(y) \cap \mathcal{C}$ for y in $h(\mathcal{C})$ are totally geodesic then*

   a) *they are all geodesically convex,*

   b) *and property H3' holds with*

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

Proof of item 1a: Let (x, v) be an arbitrary pair in $\mathcal{C} \times \mathbb{R}^n$ satisfying

$$\frac{\partial h}{\partial x}(x)\, v \;=\; 0 \quad , \qquad v^\top P(x)\, v \;=\; 1 \; . \tag{73}$$

Consider the geodesic γ_v satisfying

$$\gamma_v(0) \;=\; x \quad , \qquad \frac{d\gamma_v}{ds}(0) \;=\; v \; . \tag{74}$$

Since P is complete, γ_v is defined on $(-\infty, +\infty)$. Let J_{γ_v} be the maximal interval containing 0 so that $\gamma_v(J_{\gamma_v})$ is contained in $\mathcal{C}$.

If J_{γ_v} is reduced to a point, there is nothing to prove. In the other case, for the sake of getting a contradiction, assume that h is not constant along this geodesic on J_{γ_v}, i.e., there exists s_0 in J_{γ_v}, say positive, satisfying $h(\gamma_v(s_0)) \neq h(x)$, $\gamma_v(\sigma) \in \mathcal{C}$ for all $\sigma \in [0, s_0]$. Let s_1 be the infimum of the real numbers s in $[0, s_0]$ satisfying $h(\gamma_v(s)) \;\neq\; h(x)$. By continuity s_1 is in $[0, s_0)$ and we have $h(\gamma_v(s_1)) = h(x)$. Also, the definition of s_1 implies that, for any ε in $(0, s_0 - s_1]$, there exits s_ε in $[s_1, s_1 + \varepsilon]$ such that $h(\gamma_v(s_\varepsilon)) \neq h(\gamma_v(s_1))$. Also, when $s_1 \neq 0$, the function $s \mapsto h(\gamma_v(s))$ being constant on $[0, s_1]$, we have

$$\frac{\partial h}{\partial x}(\gamma_v(s_1))\, \frac{d\gamma_v}{ds}(s_1) \;=\; 0 \; . \tag{75}$$

Note that, with (73) and (74), the same holds when $s_1 = 0$.

Now let $\mathcal{B}_\varepsilon(\gamma_v(s_1))$ be a geodesic ball centered at $\gamma_v(s_1)$ with geodesic radius ε sufficiently small to ensure that each geodesic between $\gamma_v(s_1)$ and any point in this ball is minimal. See [5, Theorem VI.7.2]. With s_ε associated with ε as shown above, we define a function γ^* as $\gamma^*(s) \;=\; \gamma_v(s_\varepsilon - s)$ for all $s \in [0, s_\varepsilon - s_1]$. It is a minimal geodesic between $\gamma^*(0) = \gamma_v(s_\varepsilon)$ and $\gamma^*(s_\varepsilon - s_1) = \gamma_v(s_1)$ satisfying $\gamma^*(s) \in \mathcal{C} \cap \mathcal{B}_\varepsilon(\gamma_v(s_1))$ for all $s \in [0, s_\varepsilon - s_1]$ and $h(\gamma^*(0)) \;\neq\; h(\gamma^*(s_\varepsilon - s_1))$. So, according to H3', we have

$$\frac{d}{ds}\delta(h(\gamma^*(s)), h(\gamma^*(0))) \;>\; 0$$

for all $s \in (0, s_\varepsilon - s_1]$. In particular, we have

$$\frac{\partial \delta}{\partial y_1}(h(\gamma^*(s_\varepsilon - s_1)), h(\gamma^*(0))) \times \\ \frac{\partial h}{\partial x}(\gamma^*(s_\varepsilon - s_1))\, \frac{d\gamma^*}{ds}(s_\varepsilon - s_1) \;>\; 0.$$

But (75) leads to a contradiction since

$$\begin{aligned}\frac{\partial h}{\partial x}(\gamma^*(s_\varepsilon - s_1))\,\frac{d\gamma^*}{ds}(s_\varepsilon - s_1) &= -\frac{\partial h}{\partial x}(\gamma_v(s_1))\,\frac{d\gamma_v}{ds}(s_1)\\ &= 0.\end{aligned}$$

Proof of item 1b: Let $(x_1, x_2) \in \mathcal{C} \times \mathcal{C}$ be any arbitrary pair of points satisfying $h(x_1) = h(x_2) = y$. Since $\mathcal{C}$ is geodesically convex, there exists a minimal geodesic γ^* between $x_1 = \gamma^*(s_1)$ and $x_2 = \gamma^*(s_2)$ satisfying $\gamma^*(s) \in \mathcal{C}$ for all $s \in [s_1, s_2]$. We have $\delta(h(\gamma^*(s_2)), h(\gamma^*(s_1))) = \int_{s_1}^{s_2} \frac{\partial \delta}{\partial y_1}(h(\gamma^*(s)), h(\gamma^*(s_1)))\ \frac{d\,h\circ\gamma^*}{ds}(s)\,ds$. But (52) implies the left-hand side of this equation is zero if and only if we have $h(\gamma^*(s)) = h(\gamma^*(s_1))$ for all $s \in [s_1, s_2]$, that is, the geodesic γ^* remains in the set $\mathfrak{H}(h(x_1)) \cap \mathcal{C}$ for all s in $[s_1, s_2]$.

Proof of item 2a: Let $(x_1, x_2) \in \mathcal{C} \times \mathcal{C}$ be any arbitrary pair of points satisfying $h(x_1) = h(x_2) = y$. Since $\mathcal{C}$ is geodesically convex, there exists a minimal geodesic γ^* between $x_1 = \gamma^*(s_1)$ and $x_2 = \gamma^*(s_2)$ satisfying $\gamma^*(s) \in \mathcal{C}$ for all $s \in [s_1, s_2]$. For the sake of getting a contradiction, assume that $\mathfrak{H}(y) \cap \mathcal{C}$ is not geodesically convex. Then, there exists $\hat{s} \in [s_1, s_2]$ such that $\gamma^*(s) \notin \mathfrak{H}(y) \cap \mathcal{C}$. But $\gamma^*(\hat{s})$ being in $\mathcal{C}$, this implies $|h(\gamma^*(\hat{s})) - h(x_1)|^2 \neq 0$. By continuity and compactness, the function $s \in [s_1, s_2] \mapsto |h(\gamma^*(s)) - h(x_1)|^2$ admits a maximum at some $s_{\max}$ in (s_1, s_2) and, hence

$$h(\gamma^*(s_{\max})) \neq h(x_1), \tag{76}$$

$$(h(\gamma^*(s_{\max})) - h(x_1))^\top\,\frac{dh\circ\gamma^*}{ds}(s_{\max}) \;=$$

$$(h(\gamma^*(s_{\max})) - h(x_1))^\top\,\frac{\partial h}{\partial x}(\gamma^*(s_{\max}))\frac{d\gamma^*}{ds}(s_{\max}) = 0\ .$$

When the dimension m of outputs is one, this implies $\frac{\partial h}{\partial x}(\gamma^*(s_{\max}))\frac{d\gamma^*}{ds}(s_{\max}) = 0$. Since the set $\mathfrak{H}(h(\gamma^*(s_{max}))) \cap \mathcal{C}$ is totally geodesic and γ^* takes its values in $\mathcal{C}$ on the interval $[s_1, s_2]$ containing s_{max}, we conclude that γ^* takes actually its values in $\mathfrak{H}(h(\gamma^*(s_{\max}))) \cap \mathcal{C}$ on $[s_1, s_2]$. This contradicts (76), and so $\mathfrak{H}(y) \cap \mathcal{C}$ must be geodesically convex.

Proof of item 2b: Let $(\hat{x}, x)$ be an arbitrary pair of points in $\mathcal{C}\times\mathcal{C}$ satisfying $h(\hat{x}) \neq h(x)$. Since $\mathcal{C}$ is geodesically convex, there exists a minimal geodesic γ^* between $x = \gamma^*(0)$ and $\hat{x} = \gamma^*(\hat{s})$ satisfying $\gamma^*(s) \in \mathcal{C}$ for all $s \in [0, \hat{s}]$. Assume there exists s in $[0, \hat{s}]$ satisfying $\frac{d\,h\circ\gamma^*}{ds}(s) = \frac{\partial h}{\partial x}(\gamma^*(s))\,\frac{d\gamma^*}{ds}(s) = 0$, $\frac{d\gamma^*}{ds}(s)^\top P(\gamma^*(s))\,\frac{d\gamma^*}{ds}(s) = 1$. Then, since $\mathfrak{H}(h(\gamma^*(s))\cap\mathcal{C}$ is totally geodesic, and γ^* takes its values in $\mathcal{C}$ on $[0, \hat{s}]$, we have $h\circ\gamma^*(s) = h\circ\gamma^*(0) = h(x)$ for all $s \in [0, \hat{s}]$ which contradicts $h\circ\gamma^*(\hat{s}) = h(\hat{x}) \neq h(x)$. Then, $\dfrac{d\,h\circ\gamma^*}{ds}$ has a constant sign. But, since we have $h(\hat{x}) - h(x) = \int_0^{\hat{s}} \frac{d\,h\circ\gamma^*}{ds}(s)\,ds$, this sign must be the same as the one of $h(\hat{x}) - h(x)$. We conclude that we have

$$\frac{d}{ds}|h(\gamma(s)) - h(\gamma(0))|^2 = [h(\gamma(s)) - h(\gamma(0))]\,\frac{d\,h\circ\gamma^*}{ds}(s) > 0$$

for all $s \in (0, \hat{s}]$.

REFERENCES

[1] N. Aghannan and P. Rouchon. An intrinsic observer for a class of Lagrangian systems. *IEEE Trans. Automatic Control*, 48(6):936–945, 2003.

[2] A. Astolfi and L. Praly. Global complete observability and output-to-state stability imply the existence of a globally convergent observer. *Mathematics of Control, Signals, and Systems (MCSS)*, 18(1):32–65, 2006.

[3] S. Bonnabel. *Observateurs Asymptotiques Invariants*. PhD thesis, Spécialité Mathématiques et Automatique. École des Mines de Paris, 2007.

[4] S. Bonnabel. A simple intrinsic reduced-observer for geodesic flow. *IEEE Transactions on Automatic Control*, 55(9):2186–2191, 2010.

[5] W. M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, 1975.

[6] E. Cartan. *Geometry of Riemannian spaces*. Maths Sci Press, second edition, 1951.

[7] B. P. Demidovich. Dissipativity of a nonlinear system of differential equations. *Ser. Mat; Mekh. Part I.6 (1961); Part II.1, 3-8(1962) (in Russian).*, pages 19–27.

[8] M.P. DoCarmo. *Riemannian Geometry*. Birkhauser, Boston, 1992.

[9] F. Filippov. *Differential equations with discontinuous right hand sides*. Kluwer Academic Publishers. Mathematics and Its Applications, 1988.

[10] P. Hartmann. *Ordinary differential equations*. Wiley, 1964.

[11] N. J. Hicks. *Notes on Differential Geometry*. Van Nostrand Publishing Company, Princeton, New Jersey, 1965.

[12] T. Terlaky I. Pólik. A survey of the s-lemma. *SIAM Review*, 49(3):371–418, september 2007.

[13] G. Isac and S. Z. Németh. *Scalar and asymptotic scalar derivatives : theory and applications*. Springer, 4th edition, 2008.

[14] J. Jouffroy. Some ancestors of contraction analysis. *Proc. IEEE Conf. Dec. Control*, pages 5450–5455, 2005.

[15] S. Kaboyashi and K. Nomizu. *Foundations of Differential Geometry*, volume 1. Wiley Classics Library Edition, 1996.

[16] A. Krener and A. Isidori. Linearization by output injection and nonlinear observers. *Systems & Control Letters*, 3:47–52, 1983.

[17] L. P. Lebedev and M. J. Cloud. *Tensor analysis*. Norwood Mass., 2005.

[18] D. C. Lewis. Metric properties of differential equations. *American Journal of Mathematics*, 71:294–312, 1949.

[19] W. Lohmiller and J.-J. Slotine. On contraction analysis for nonlinear systems. *Automatica*, 34(6):671–682, 1998.

[20] S. Z. Németh. *Geometric aspects of Minty-Browder monotonicity*. PhD thesis, E ötvös Loránd University. Budapest., 1998.

[21] L Praly. On observers with state independent error Lyapunov function. In *Proceedings of the 5th IFAC Symposium "Nonlinear Control Systems" (NOLCOS'01)*, 2001.

[22] G. B. Price. Definitions and properties of monotone functions. *Bull. Amer. Math. Soc.*, 46:77–80, 1940.

[23] T. Rapcsák. *Smooth Nonlinear Optimization in* $\mathbb{R}^n$. Kluwer Academic, 1997.

[24] T. Sakai. *Riemannian geometry*, volume 149. Translation of Mathematical monographs, American Mathematical Soc., 1996.

[25] M. Spivak. *(A Comprehensive Introduction to) Differential geometry*, volume 2. Publish or Perish, Inc., 1979. 2nd Edition.

[26] J. Tsinias. Further results on the observer design problem. *Syst. Contr. Lett.*, 14:411–418, 1990.