Observability measures for nonlinear systems

Shen Zeng

Abstract— **In this paper, the nonlinear observability problem is revisited and studied in a novel framework that is particularly well-suited for the consideration of different practical aspects of the corresponding state estimation problem. In establishing this framework, we highlight connections between fundamental theoretical aspects of the general observability problem with more quantitative considerations similar to those encountered in numerical analysis. A key result of our novel analysis of the observability problem is the introduction of quantitative observability measures for nonlinear systems as a generalization of the notion of observability gramians for linear systems.**

I. Introduction

The observability problem, along with its counterpart, the controllability problem, can be considered as one of the key forces in the establishment of systems theory as an independent field in the 1960s, see e.g. [1]. Indeed, those two problems carry a highly distinctive flavor that distinguishes it from the otherwise quite similar studies of differential equations and, more generally, dynamical systems.

In the 1970s, these problems were started being considered in a nonlinear setting [2], [3], [4], [5]. From a purely theoretical point of view, our understanding of the matter was settled through the work of Hermann and Krener [5], which showed how the description of observability and controllability in a (differential) geometric framework are naturally generalized by means of a Lie algebraic framework. Since then a vast literature has accumulated on the general topic of observability of nonlinear systems, with a continuing stream of works on topics related to both qualitative and quantitative aspects [6], [7], [8], [9].

Nevertheless, the topic has never been really settled in a satisfactory manner from a practical perspective, leaving the search for fundamental and practically applicable results on the observability problem ongoing [10], [11], [7], [12], [13]. In particular, a major point of criticism is that most theoretical frameworks invoke a "differential viewpoint", in which theoretical results on the observability of nonlinear systems are, in some form or another, based on the consideration of higher order derivatives of the output signal. This, however, should in fact strike us as odd, as from a practical perspective, the idea of utilizing only the derivatives of an output signal at one single time point is arguably the least favorable strategy that one could pursue. In this paper, we attempt to address this apparent gap between previous theoretical considerations and practically relevant scenarios for the state estimation problem. In doing so, we hope to gain a deeper understanding of the mechanisms governing the property of observability on a quantitative level, which is also inherently linked to the question of how strongly the actual reconstruction of a specific state is affected by noise. A main contribution of this paper is the introduction of novel numerical considerations into the study of observability of dynamical systems that are deeply rooted in fundamental theoretical considerations. This leads to a framework that reflects the practical side of the observability problem more adequately and in which practical aspects such as the relative degree of observability can be analyzed in detail. More specifically, in light of this framework, it is natural to derive nonlinear analogs for important observability properties of linear systems, namely local observability gramians and global (un)observability measures. Throughout the paper, we make an attempt to bridge the relevant, quite abstract theoretical foundations with very descriptive simple examples and illustrations.

This paper is organized as follows. In Section II we review some fundamental aspects of the general observability problem and introduce some notation used in the remainder of the paper. In Section III we attempt to bridge the fundamental aspects of the linear observability problem with more quantitative considerations similar to those encountered in numerical analysis. This leads to some novel insights about the definition of quantitative measures for the degree of observability of linear systems. In Section IV, the insights obtained from the linear case are systematically generalized to the nonlinear case, leading to the introduction of local observability gramians and global (un)observability measures. For each of the newly defined quantities, the connections to the linear case are highlighted in detail. It is also shown how these newly introduced measures can be readily computed, allowing us to visualize the observability properties of nonlinear systems.

II. Review of the general observability problem

We consider a nonlinear system (without control inputs)

$$\dot{x}(t) = f(x(t)),$$
$$y(t) = h(x(t)),$$

where $x(t) \in \mathbb{R}^n$ describes the state and $y(t) \in \mathbb{R}^m$ describes the output of the system, at time t, respectively. We are particularly interested in the case $m < n$, in which the role of the output is to take into account the practical circumstance that one is typically unable to measure the full state and in some cases has to resort to cases in which only some subset of states (or a function thereof) can be directly acquired.

Shen Zeng is with the Department of Electrical and Systems Engineering, Washington University, St. Louis, MO 63130, USA.
Correspondence to s.zeng@wustl.edu

Given a set of measurement times $\mathcal{T}$, we define the equivalence relation

$$x' \sim_{\mathcal{T}} x'' :\Leftrightarrow \ \forall t \in \mathcal{T} \ \ (h \circ \Phi_t)(x') = (h \circ \Phi_t)(x''),$$

where $\{\Phi_t\}_{t\in\mathbb{R}}$ denotes the flow of the system $\dot{x} = f(x)$. Points that are equivalent in this so-called indistinguishability relation are those that cannot be distinguished on the basis of their corresponding output trajectories (restricted to the set $\mathcal{T}$). We denote the associated equivalence classes by

$$I_{\mathcal{T}}(x_0) := \{x \in \mathbb{R}^n : x \sim_{\mathcal{T}} x_0\}.$$

Given this notation, it is not difficult to see that a system is observable if $I_{\mathcal{T}}(x_0) = \{x_0\}$, i.e. the partition of the state space into indistinguishable sets is maximally fine. It is useful to also define the information/indistinguishability set with respect to one specific time point $t \in \mathcal{T}$, denoted by $I_{\{t\}}$, which describes the content of information about the initial state x_0 that we can extract from a single output measurement $y(t)$ gathered at a time point $t \in \mathcal{T}$. This set can be written as

$$I_{\{t\}}(x_0) = (h \circ \Phi_t)^{-1}(\{y(t)\}) = \Phi_{-t}(h^{-1}(\{y(t)\})),$$

which is constructive as it implicitly encodes a way to compute these sets, namely by evolving the level set $h^{-1}(\{y(t)\})$ with the flow backward in time.

The process of combining all the different pieces of information about x_0 contained in the sets $I_{\{t\}}$ is equivalent to computing the intersection

$$I_{\mathcal{T}}(x_0) = \bigcap_{t\in\mathcal{T}} I_{\{t\}}(x_0),$$

which offers a more favorable way of thinking about the observability and state estimation problem as opposed to the less intuitive and also less practical differential perspective.

III. The linear observability problem

In the case of linear systems $\dot{x} = Ax, y = Cx$, the sets $I_{\{t\}}$ can be written more explicitly as $I_{\{t\}} = x_0 + \ker Ce^{At}$. Each single information set is an affine subspace anchored in the initial state x_0 and spanned by the vectors in the non-trivial null space $\ker Ce^{At}$. This graphical description contains the (incomplete) information about x_0 that we can infer from a single output measurement $y(t)$. By intersecting these information sets for different $t \in \mathcal{T}$, we hope to obtain a trivial intersection at the actual initial state x_0.

Though extremely elementary, these geometric descriptions were only recently being considered and employed in the study of ensemble observability [14], [15], where the richness of the directions generated by $\ker Ce^{At}$ with varying $t \in \mathcal{T}$ emerged rather prominently as a key property. In the end, however, these considerations are relevant not only to the ensemble observability problem but can also shed new light on the classical observability problem. For instance, the way the sets $I_{\{t\}}$ intersect provides clues as to how stable the reconstruction is in the presence of measurement noise.

Generally speaking, a narrow range of directions $\ker Ce^{At}$ is typically associated with less favorable observability properties, and one would ideally wish for the family $\ker Ce^{At}$ to cover as many different "directions" as possible. The exact situation, however, is often more sophisticated and requires the consideration of further properties beyond the purely geometric ones, such as the magnitude of Ce^{At} that directly relates to the "signal-to-noise ratio". In our geometric picture, we may think of this as the "binding strengths" of the information sets $I_{\{t\}}$ to the actual initial state x_0 in the presence of measurement noise. In the remainder of this paper, we bridge the very elementary and intuitive geometric viewpoint with more analytic and quantitative elements. This synthesis turns out to be quite fruitful and in particular results in the introduction of a novel yet most natural generalization of observability gramians and global (un)observability measures to nonlinear systems. We first highlight these novel considerations by means of a simple linear example.

Example 1: Consider the double integrator

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x,$$
$$y = \begin{pmatrix} 1 & 0 \end{pmatrix} x,$$

which corresponds to the situation of observing the position of an object traveling with constant speed along a single dimension. Solving the observability problem here amounts to estimating the velocity of the object from position measurements. Suppose we are only allowed to measure the system twice, at time points t_1 and t_2, where $t_1 < t_2$. It is a straightforward computation to see that

$$\begin{pmatrix} 1 & t_1 \\ 1 & t_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y(t_1) \\ y(t_2) \end{pmatrix}.$$

In the geometric framework, we may associate to this equation a picture in which two lines, parameterized by $x_1 + t_k x_2 = y(t_k)$, are intersecting in the solution to the problem. A basic trigonometric consideration yields the corresponding angles $\alpha_k = \pi/2 + \arctan(t_k)$. From a purely geometric perspective ignoring the magnitude of the row vectors $(1, t_1)$ and $(1, t_2)$, the single objective would be to maximize the spread of the two lines, yielding the most stable geometric configuration. In general, however, spread and magnitude have to be considered jointly in quantifying the observability properties. The actual interplay between these two components may in general be quite sophisticated. Our current lack of understanding of this interplay highlights the need for general quantitative tools to better analyze and, in a second step, enhance the observability properties of a dynamical system. The following analysis for the simple example provides some first hints that we then take as the basis for further considerations.

A natural candidate for a quantitative measure of observability is given by the "condition number" of the matrix $\mathcal{O}_{\mathcal{T}}$ relating the initial state with the output measurements, which captures how much the solution changes upon perturbation of the right-hand side from y to some other value $\tilde{y}$. It is straightforward to see that

$$\|x-\tilde{x}\| \leq \|\mathcal{O}_{\mathcal{T}}^{-1}\|\|y-\tilde{y}\|,$$

where $\tilde{x}$ is the solution to $\mathcal{O}_{\mathcal{T}}\tilde{x}=\tilde{y}$. Moving back to our specific observability problem, we have

$$\mathcal{O}_{\mathcal{T}}^{-1}=\begin{pmatrix}1 & t_1\\ 1 & t_2\end{pmatrix}^{-1}=\frac{1}{t_2-t_1}\begin{pmatrix}t_2 & -t_1\\ -1 & 1\end{pmatrix},$$

and thus

$$\|\mathcal{O}_{\mathcal{T}}^{-1}\|_{\mathrm{F}}=\frac{1}{t_2-t_1}\sqrt{2+t_1^2+t_2^2}. \tag{1}$$

Here we make use of the bound $\|\mathcal{O}_{\mathcal{T}}^{-1}\|_2 \leq \|\mathcal{O}_{\mathcal{T}}^{-1}\|_{\mathrm{F}}$ and the fact that $\|\mathcal{O}_{\mathcal{T}}^{-1}\|_{\mathrm{F}}$ is slightly simpler to work with. It can be seen that for the "condition number" (1) to be as small as possible, the strategy is to maximize the difference t_2-t_1 rendering the first term as small as possible while trying to keep $t_1^2+t_2^2$ as small as possible. Fulfilling these two requirements can be achieved by the choice $t_1=0$ and some $t_2 \gg 1$. In the limit $t_2 \to \infty$, one would thus have $\|\mathcal{O}_{\mathcal{T}}^{-1}\|_{\mathrm{F}} \to 1$. This measurement strategy corresponds to measuring the initial position and then to wait as long as possible to record the position a second time. Associated to this measurement strategy is the very intuitive reconstruction strategy in which the optimal estimate of the second coordinate, the velocity, would then be given by computing

$$\hat{x}_2=\frac{y(t_2)-y(t_1)}{t_2-t_1}.$$

This is in accordance with computing the least squares solution to $\mathcal{O}_{\mathcal{T}}x=y$, given by $\hat{x}=\mathcal{O}_{\mathcal{T}}^{\dagger}y$, which also yields the (less obvious) optimal position estimate as

$$\hat{x}_1=\frac{t_2y(t_1)-t_1y(t_2)}{t_2-t_1},$$

which in the case that $t_1=0$ is equal to $\hat{x}_1=y(t_1)$. Of course, the measurement times would not matter much in the case of ideal, noise-free, measurements. However, in the case of (additive) white noise in the measurements, the above strategy is expected to yield more reliable results for the velocity estimation as the signal-to-noise ratio for the second measurement improves when $t_2 \gg 1$.

Now the situation becomes much more complicated even for the above considered simple example when taking more than two measurement times into account. Then, as we will see in the following, a reasonable measure for quantitative observability would be given by the *smallest* singular value of $\mathcal{O}_{\mathcal{T}}$, and the strategy for enhancing the observability properties of a system would be to maximize the smallest singular value. We note that in the context of optimal designs, this "max-min objective" is also referred to as *E-optimality*. In the general linear case, we have

$$\mathcal{O}_{\mathcal{T}}=\begin{pmatrix}Ce^{At_1}\\ \vdots\\ Ce^{At_M}\end{pmatrix}.$$

It is reasonable to assume that $\mathcal{O}_{\mathcal{T}}$ has full column rank, i.e. to assume that the (linear) system is fully observable, so that the left inverse $\mathcal{O}_{\mathcal{T}}^{\dagger}$ is well-defined. In view of the general case of measurement times $t_1,\ldots,t_M$, we would then consider the observability measure

$$\|\mathcal{O}_{\mathcal{T}}^{\dagger}\|_2=\max_{\|y\|=1}\|\mathcal{O}_{\mathcal{T}}^{\dagger}y\|=\sigma_{\max}(\mathcal{O}_{\mathcal{T}}^{\dagger}).$$

A description of $\|O_{\mathcal{T}}^{\dagger}\|_2$ purely in terms of $O_{\mathcal{T}}$ is given by considering the relation

$$\|\mathcal{O}_{\mathcal{T}}^{\dagger}\|_2=\sigma_{\max}(\mathcal{O}_{\mathcal{T}}^{\dagger})=\frac{1}{\sigma_{\min}(\mathcal{O}_{\mathcal{T}})},$$

which follows from the fact that if $\mathcal{O}_{\mathcal{T}}$ admits the singular value decomposition $\mathcal{O}_{\mathcal{T}}=U\Sigma V^{\top}$, then $\mathcal{O}_{\mathcal{T}}^{\dagger}$ admits the singular value decomposition $\mathcal{O}_{\mathcal{T}}^{\dagger}=V\Sigma^{\dagger}U^{\top}$. Now instead of demanding that $\|\mathcal{O}_{\mathcal{T}}^{\dagger}\|_2$ be as small as possible, we naturally could instead demand that

$$\sigma_{\min}(\mathcal{O}_{\mathcal{T}})=\min_{\|x\|=1}\|O_{\mathcal{T}}x\|$$

be as large as possible, yielding the claimed condition.

By reconsidering the same simple double-integrator example but now allowing more measurement times $t_1,\ldots,t_M$, it quickly becomes clear that the general consideration by means of the above condition is quite tedious. This illustrates that a purely analytical approach such as that pursued in the simple case of two measurement times is not even feasible in the general linear case. This leads us to treat the general problem in a completely computational, more encompassing, nonlinear framework in the next section.

IV. PRACTICAL OBSERVABILITY ANALYSIS FOR NONLINEAR SYSTEMS

Consider the nonlinear system (without control inputs)

$$\begin{aligned}\dot{x}(t)&=f(x(t)),\\ y(t)&=h(x(t)),\end{aligned}$$

and let $\mathcal{T}=\{t_1,t_2,\ldots,t_M\}$ denote the set of measurement times, with $0 \leq t_1 < t_2 < \cdots < t_M$, i.e. we only consider the case of finitely many measurement times here for the sake of simplicity (at the same time, this also rules out any discussion of a "differential approach"). In the spirit of [16], we define the function

$$\mathrm{Rec}_{\mathcal{T}}:\mathbb{R}^n \to \mathbb{R}^{mM}, \quad x_0 \mapsto \begin{pmatrix}(h\circ\Phi_{t_1})(x_0)\\ \vdots\\ (h\circ\Phi_{t_M})(x_0)\end{pmatrix},$$

mapping some initial state $x_0 \in \mathbb{R}^n$ to the corresponding *recorded* (discrete-time) signal that one would acquire when measuring the output at the time points specified in the M-tuple $\mathcal{T}$. The framework associated with this viewpoint quite naturally leads to the definition of local observability gramians for nonlinear systems in the following.

A. Local observability gramians of nonlinear systems

We define the family of functionals $\{J_{x_0}\}_{x_0 \in \mathbb{R}^n}$, parameterized by the reference initial state $x_0 \in \mathbb{R}^n$, where each functional is of the form

$$J_{x_0}(x) = \| \operatorname{Rec}_{\mathcal{T}}(x) - \operatorname{Rec}_{\mathcal{T}}(x_0)\|.$$

For a fixed $x_0 \in \mathbb{R}^n$, the functional $J_{x_0} : \mathbb{R}^n \to \mathbb{R}$ captures all the relevant observability properties of the nonlinear system at the reference point $x_0 \in \mathbb{R}^n$. We first note that

$$J_{x_0}(x) \geq 0 \text{ for all } x \in \mathbb{R}^n, \text{ and } J_{x_0}(x_0) = 0.$$

Moreover, the existence of some state $x \neq x_0$ for which $J_{x_0}(x) = 0$ is equivalent to $x \sim_{\mathcal{T}} x_0$. Thus, we have

$$I_{\mathcal{T}}(x_0) := \{x \in \mathbb{R}^n : x \sim_{\mathcal{T}} x_0\} = \{x \in \mathbb{R}^n : J_{x_0}(x) = 0\}.$$

But the mapping J_{x_0} also provides some more detailed information beyond the qualitative one through its 0-level set. Indeed, when plotting the other level sets of J_{x_0} in terms of a "heat map", we obtain detailed information about the composition of the state space into different regions with a gradual transition of different levels of indistinguishability from the reference initial state x_0 by means of the corresponding output signals.

In Figure 1 we illustrate a local observability gramian for the Van der Pol system

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -x_1 + \frac{1}{2}(1 - x_1^2)x_2, \\ y &= x_1, \end{aligned} \tag{2}$$

which also highlights the connection to the information sets.

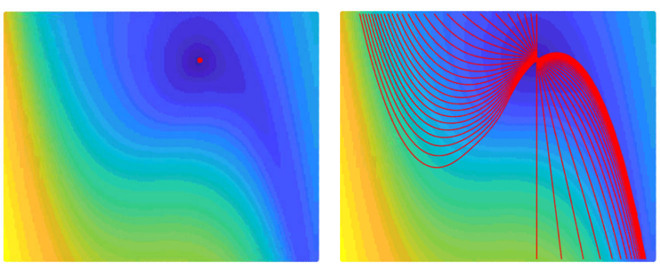

Fig. 1. Left: Level sets of J_{x_0} for the point $x_0 = (1, 2.5)$ highlighted in red for a Van der Pol oscillator considered on the region $[-4, 4]^2$ and $\mathcal{T} = \{0, 0.1, \ldots, 2.5\}$. Right: Information sets $I_{\{t_k\}}$ for $t_k \in \mathcal{T}$. In this illustration, one can see how the elongation of the level sets towards the lower right is related to the red level sets bunching up in that area. The exact connection between the red level sets and the resulting heat map J_{x_0} can, in general, not be inferred, which, again, is a justification for the fully computational approach advocated in this paper.

This very detailed information contained in J_{x_0} becomes particularly important when taking the effects of measurement noise into account, i.e. when $y_{\text{measured}} = \operatorname{Rec}_{\mathcal{T}}(x_0) + v$, with i.i.d. $v_k \sim N(0, \Sigma)$. By inspecting the functional J_{x_0} one can get a good idea of how strongly the minimizer of

$$V(x) = \| \operatorname{Rec}_{\mathcal{T}}(x) - y_{\text{measured}}\|, \tag{3}$$

may deviate from x_0. For instance, the more rapidly J_{x_0} is increasing away from x_0, the lesser the influence of the measurement noise on the reconstruction of x_0.

Remark 1 (Discussion of the linear case): In the case of linear systems, we clearly have

$$J_{x_0}(x) = \|\mathcal{O}_{\mathcal{T}}(x - x_0)\| = (x - x_0)^\top \mathcal{O}_{\mathcal{T}}^\top \mathcal{O}_{\mathcal{T}}(x - x_0),$$

where we recognize the occurring product

$$\mathcal{O}_{\mathcal{T}}^\top \mathcal{O}_{\mathcal{T}} = \sum_{k=1}^{M} e^{A^\top t_k} C^\top C e^{A t_k}$$

as the observability gramian. Our earlier introduced observability measure $\sigma_{\min}(\mathcal{O}_{\mathcal{T}})$ is therefore equal to the smallest eigenvalue of the observability gramian $\lambda_{\min}(\mathcal{O}_{\mathcal{T}}^\top \mathcal{O}_{\mathcal{T}})$. Because in the linear case, the local observability gramian J_{x_0} can be written as a function of the difference $x - x_0$, i.e.

$$J_{x_0}(x) = \psi(x - x_0)$$

where $\psi(\xi) = \xi^\top \mathcal{O}_{\mathcal{T}}^\top \mathcal{O}_{\mathcal{T}} \xi$, the local observability properties of a linear system are exactly the same for all initial states $x_0 \in \mathbb{R}^n$ under consideration. In particular, in the linear case, the heat maps J_{x_0} for different initial states $x_0 \in \mathbb{R}^n$ are given by the level sets of the observability gramian centered at the initial state $x_0 \in \mathbb{R}^n$ under consideration. This discussion shows how J_{x_0} can be considered *the* natural generalization of observability gramians.

Remark 2 (On the practical state estimation problem): As noted earlier, the map J_{x_0} introduced in this section is also very insightful with regard to the practical state estimation problem for nonlinear systems. More specifically, given an M-tuple of (noisy) measurements y_{measured} gathered at the time points specified in $\mathcal{T}$, the solution to the state estimation problem is the minimizer of the cost functional (3). A practical reconstruction algorithm which is deeply rooted in this particular formulation is a simple gradient descent scheme in the state space, as suggested in Figure 2.

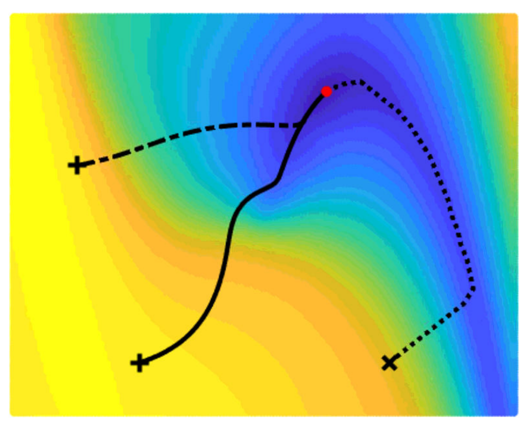

Fig. 2. Illustration of a state estimation algorithm for a Van der Pol oscillator based on the "heat map viewpoint" using an (approximate) gradient descent. The plot shows the region $[-4, 4]^2$ and $x_0 = (1, 2.5)$ in red. The measurement times associated with the particular heat map are given by $\mathcal{T} = \{0, 0.1, \ldots, 10\}$. The trajectories starting from the states marked with a plus are obtained from an approximate gradient descent. The trajectory starting from the state marked with a cross is obtained from an algorithm in which in each local window, the grid point with the smallest value is chosen for the next step. In fact, a gradient descent approach would yield a very poor result for the cross as the initial state as the trajectory would feature a zig-zag behavior about the "ridge".

As we can see from the trajectory starting from the lower right, the *shape* of the level sets of J_{x_0} is highly relevant from a practical perspective as it dictates how well the actual initial state $x_0 \in \mathbb{R}^n$ can be *practically* found by gradient descent techniques. In this regard, a heat map which is very convex and rapidly increasing away from the reference point $x_0 \in \mathbb{R}^n$ is more favorable. The observability measure for nonlinear systems that we introduce in the following does not take these numerical considerations into account, as the purely theoretical observability property is sought to be independent of any specific practical implementation. Even though the discussion of these more practical aspects are beyond of this paper, this is not to say that these practical issues are any less relevant. Rather, when it comes to solving the least squares state estimation problem for a nonlinear system, this is a route one would need to consider.

B. A global observability measure for nonlinear systems

The family of functionals $\{J_{x_0}\}_{x_0 \in \mathbb{R}^n}$ introduced in the previous section provide a complete insight into both the qualitative and, perhaps more importantly, also the quantitative observability properties of a given system. This detailed observability measure attaches to each point $x_0 \in \mathbb{R}^n$ a whole functional containing the mismatch in the resulting output signals. In this sense, this assignment of functionals J_{x_0} to initial states x_0 may be viewed as *the* natural generalization of observability gramians for nonlinear systems. One is, however, often only interested in a less detailed measure attaching a single value rather than a field to a point. This would then allow one to compare the degrees of observability of different points in the state space more directly.

Such a measure can be obtained by considering the compression of the field J_{x_0} attached to x_0 to the scalar

$$O^\delta(x_0) = \inf_{x \in \mathbb{R}^n \setminus B_\delta(x_0)} J_{x_0}, \tag{4}$$

i.e. we consider for each $x_0 \in \mathbb{R}^n$ the infimum of J_{x_0} taken over the "punctured" plane $\mathbb{R}^n \setminus B_\delta(x_0)$. Here the parameter δ is to have the same fixed value for all considered states $x_0 \in \mathbb{R}^n$ and its purpose is to take points "too close" to the reference initial state x_0 out of the competition for the smallest mismatch between the output trajectories. In other words, we ignore the points whose small mismatch is the result of their mere closeness to the reference initial state. We stress that it would not be sufficient to only consider $\inf_{x \in \mathbb{R}^n \setminus \{x_0\}} J_{x_0}$ for the value of observability as in that case each point $x_0 \in \mathbb{R}^n$ would be assigned the value 0 since $J_{x_0}(x_0) = 0$ and that J_{x_0} is a continuous function, which highlights the importance of the *uniform* "separation" from the reference point.

While the global observability measure $O^\delta : \mathbb{R}^n \to \mathbb{R}$ clearly involves a compression/reduction of the spatially extended observability gramian $J : \mathbb{R}^n \to C(\mathbb{R}^n, \mathbb{R})$, we note that it does contain *global* information. In particular, if for a fixed reference $x_0 \in \mathbb{R}^n$, there exists a state $x \neq x_0$ *not necessarily in the vicinity of* x_0 with the same output trajectory, we will have the value $O^\delta(x_0) = 0$, which is indeed the degree of observability that we expect x_0 to have.

Remark 3 (Generalization of linear observability measure): Recall that in our study of the linear case, we found the useful relation

$$\|\mathcal{O}_\mathcal{T}^\dagger\|_2 = \max_{\|y\|=1} \|\mathcal{O}_\mathcal{T}^\dagger y\| = \sigma_{\max}(\mathcal{O}_\mathcal{T}^\dagger) = \frac{1}{\sigma_{\min}(\mathcal{O}_\mathcal{T})},$$

which highlighted $\sigma_{\min}(\mathcal{O}_\mathcal{T}) = \min_{\|x\|=1} \|\mathcal{O}_\mathcal{T}\|$ as a natural quantitative measure for observability. Given our discussion in Remark 1, the definition of O^δ can in fact be recognized as a generalization of this linear observability measure.

Another reasonable (un)observability measure, which, as we will show in the following, also is deeply rooted in the foregoing linear observability analysis, is the definition of

$$U^\delta(x_0) := \sup_{\{x \in \mathbb{R}^n : J_{x_0}(x) \le \delta\}} \|x - x_0\|.$$

Intuitively speaking, it serves as a description of the expansiveness of the δ-level set of the map J_{x_0}. Thus, smaller values $U^\delta(x_0)$ correspond to better observability properties, i.e. U^δ is strictly speaking a global unobservability measure.

Remark 4 (Generalization of linear unobservability measure): This unobservability measure is in fact a direct generalization of the observability measure $\|\mathcal{O}_\mathcal{T}^\dagger\|_2$ from the linear case. Indeed, note first that we have the lower bound

$$\max_{\|y\|=1} \|\mathcal{O}_\mathcal{T}^\dagger y\| \ge \max_{\|\mathcal{O}_\mathcal{T} x\|=1} \|x\|,$$

which follows from the consideration of $y \in R(\mathcal{O}_\mathcal{T})$. Secondly, it can be shown that

$$\max_{\|\mathcal{O}_\mathcal{T} x\|=1} \|x\| = \frac{1}{\sigma_{\min}(\mathcal{O}_\mathcal{T})}, \tag{5}$$

which is also the value of $\max_{\|\mathcal{O}_\mathcal{T} x\|=1} \|x\|$, as we saw earlier, so that we can then conclude the equality

$$\|\mathcal{O}_\mathcal{T}^\dagger\|_2 = \max_{\|y\|=1} \|\mathcal{O}_\mathcal{T}^\dagger y\| = \max_{\|\mathcal{O}_\mathcal{T} x\|=1} \|x\|.$$

In order to show (5), we again consider the singular value decomposition $\mathcal{O}_\mathcal{T} = U\Sigma V^\top$. Modulo unitary transformations, the set $\|\mathcal{O}_\mathcal{T} x\| = 1$ is an ellipsoid $\sigma_1^2 \tilde{x}_1^2 + \cdots + \sigma_n^2 \tilde{x}_n^2 = 1$. By traveling along the longest principle axis, i.e. the axis associated with the minimal singular value σ_n, we can attain the maximal value of $\|\tilde{x}^\star\| = \frac{1}{\sigma_n}$.

While in the linear case, the observability measure $\sigma_{\min}(\mathcal{O}_\mathcal{T})$ and the unobservability measure $1/\sigma_{\min}(\mathcal{O}_\mathcal{T})$ are related through a mere inversion, there is no such simple relation between the measures O^δ and U^δ in the nonlinear case. In fact, in the nonlinear case, the measures O^δ and U^δ emphasize quite different (un)observability properties of a nonlinear system. For instance, in contrast to the observability measure O^δ, the unobservability measure U^δ does not possess the "qualitative exactness" property. In more detail, if there exist states $x \neq x_0$ with the same exact output trajectory as the one generated by x_0, we will have

$$U^\delta(x_0) = \sup_{\{x \in \mathbb{R}^n : J_{x_0}(x) = 0\}} \|x - x_0\|,$$

i.e. in the assessment of the unobservability, the (minimal) distance to the indistinguishable state $x \sim_\mathcal{T} x_0$ matters, whereas for the observability measure O^δ, we have $O^\delta(x_0) = 0$ regardless of the actual distances.

C. Computing the introduced observability measure

In this section, we show that the novel observability measures introduced in this paper are in fact feasible for computations. While the definition of e.g. O^δ in (4) appears quite nested and possibly quite involved, we will see that it is in fact quite naturally tied with its computation by numerical means. First of all, the local observability gramians J_{x_0} occurring in (4) are computed exactly as they are defined, i.e. through the mismatch in output trajectories $\mathrm{Rec}_{\mathcal{T}}(x)$ and $\mathrm{Rec}_{\mathcal{T}}(x_0)$. For computational purposes, we introduce a grid $\{x^{(i)}\}$ on the state space and consider

$$J_{x^{(i)}}(x^{(j)}) = \| \mathrm{Rec}_{\mathcal{T}}(x^{(j)}) - \mathrm{Rec}_{\mathcal{T}}(x^{(i)})\|.$$

Because this computation of J_{x_0} is somewhat exhaustive, it is crucial to have a tailored integration technique for forward simulating the dynamics for a vast number of different initial states in a specified area. The discussion of the particular integration technique best suited for these computations will be discussed in future work.

Regarding the computation of O^δ, the first step is again to introduce a grid on the state space. The parameter δ introduced in (4) for theoretical purposes herein quite remarkably admits a natural interpretation as the mesh size of the (regular) grid. As in the computation of J_{x_0} for a fixed initial state $x_0 \in \mathbb{R}^n$, we compute and store output trajectories with each grid point as an initial state. To compute the value of O^δ defined on the grid points $x_0 = x^{(i)}$, we would go through all the grid points $x^{(j)}$ and compare their distances of the corresponding output trajectories in the search for the minimal one. If the system under scrutiny is observable, such as in the case of the Van der Pol oscillator, we can actually restrict our attention to grid points $x^{(j)}$ in a fixed (small) window about the considered initial state $x_0 = x^{(i)}$ to accelerate the computation. We stress that even though the described computation appears to be quite expensive, with the aforementioned specific integration technique, as well as the restriction to a small window about $x_0 = x^{(i)}$, it is in fact feasible to compute O^δ with a high resolution very quickly. In Figure 3 we illustrate the observability measure O^δ for measurement times $\mathcal{T} = \{0, 0.1, \ldots, 10\}$ and in Figure 4, we illustrate two dozen observability gramians for the Van der Pol system (2) computed with increasing number of measurement times, highlighting the time evolution of the observability measure. Since the observability measure O^δ is defined in quite computational terms, it is perhaps somewhat surprising to see that it admits a very smooth and continuous dynamics, which is rather interesting in its own right.

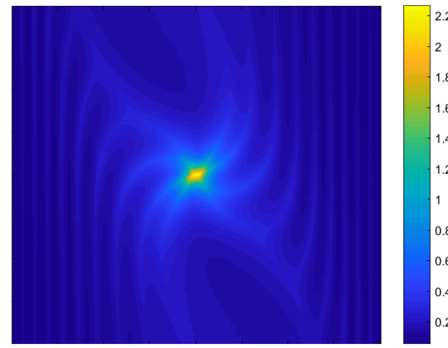

Fig. 3. This figure illustrates the global observability measure O^δ for the Van der Pol oscillator (2) on the region $[-4, 4]^2$. The parameters are $\delta = 5 \times 10^{-3}$ and $\mathcal{T} = \{0, 0.1, \ldots, 10\}$.

V. Summary and outlook

We pointed out novel ways to think about the general observability problem in systems theory, which led to the introduction of a framework that better reflects the actual circumstances in practical state estimation problems. In this framework, one can quite naturally discuss the degree of observability and also connect these quantitative aspects to the fundamental geometric perspective of the observability problem. This was first described for linear systems and then systematically generalized to the nonlinear case in an expository manner. In particular, we elucidated the "right" generalizations of quantitative observability measures for nonlinear systems. In future work, we will investigate how the framework established in this paper may be leveraged in order to determine output functions, measurement times, as well as input signals that best distinguish between different states in state space, both theoretically and practically.

References

[1] R. Kalman, "On the general theory of control systems," *IRE Transactions on Automatic Control*, vol. 4, no. 3, pp. 110–110, 1959.

[2] E. W. Griffith and K. Kumar, "On the observability of nonlinear systems: I," *Journal of Mathematical Analysis and Applications*, vol. 35, no. 1, pp. 135–147, 1971.

[3] S. R. Kou, D. L. Elliott, and T. J. Tarn, "Observability of nonlinear systems," *Information and Control*, vol. 22, no. 1, pp. 89–99, 1973.

[4] D. Aeyels and D. Elliott, "Global observability for nonlinear autonomous differential equations," in *Proc. 1978 IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes*, vol. 17. IEEE, 1979, pp. 1057–1061.

[5] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Trans. Autom. Control*, vol. 22, no. 5, pp. 728–740, 1977.

[6] B. Tibken, "Observability of nonlinear systems-an algebraic approach," in *Proc. 43rd IEEE Conference on Decision and Control*, vol. 5. IEEE, 2004, pp. 4824–4825.

[7] A. J. Krener and K. Ide, "Measures of unobservability," in *Proc. 48th IEEE Conference on Decision and Control*, 2009, pp. 6401–6406.

[8] M.-A. Belabbas, "Geometric methods for optimal sensor design," in *Proc. R. Soc. Lond. A*, vol. 472, no. 2185, 2016, p. 20150312.

[9] S. Zeng and F. Allgöwer, "A general sampled observability result and its applications," in *Proc. 55th IEEE Conference on Decision and Control*, 2016, pp. 3997–4002.

[10] C. Letellier, L. A. Aguirre, and J. Maquet, "Relation between observability and differential embeddings for nonlinear dynamics," *Physical Review E*, vol. 71, no. 6, p. 066213, 2005.

[11] U. Vaidya, "Observability gramian for nonlinear systems," in *Proc. 46th IEEE Conference on Decision and Control*, 2007, pp. 3357–3362.

[12] T. Paradowski, B. Tibken, and R. Swiatlak, "An approach to determine observability of nonlinear systems using interval analysis," in *Proc. 2017 American Control Conference*, 2017, pp. 3932–3937.

[13] M. Guay and K. T. Atta, "An extremum-seeking control observer design technique for nonlinear systems," in *In Proc. 56th IEEE Conference on Decision and Control*, 2017, pp. 2120–2125.

[14] S. Zeng and F. Allgöwer, "On the ensemble observability problem for nonlinear systems," in *Proc. 54th IEEE Conference on Decision and Control*, 2015, pp. 6318–6323.

[15] S. Zeng, S. Waldherr, C. Ebenbauer, and F. Allgöwer, "Ensemble observability of linear systems," *IEEE Transactions on Automatic Control*, vol. 61, no. 6, pp. 1452–1465, 2016.

[16] F. Takens, "Reconstruction and observability, a survey," *IFAC Proceedings Volumes*, vol. 31, no. 17, pp. 427–429, 1998.

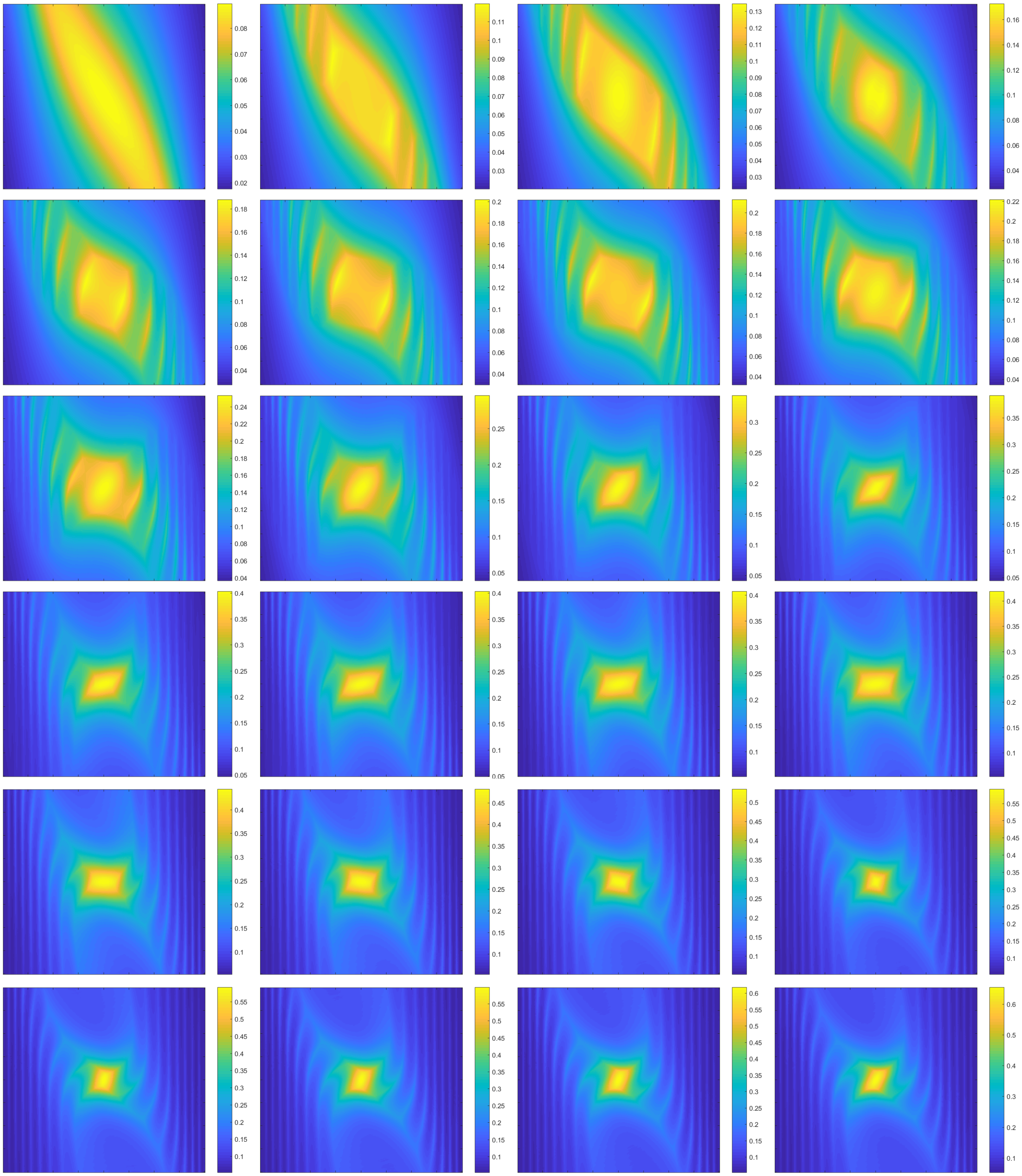

Fig. 4. This figure illustrates several global observability measures computed with increasing lengths of recording times for the considered output trajectories (from left to right, top to bottom). By inspecting the color bar, we see that the values of $O^{\delta}(x_0)$ are steadily increasing as more measurement times are taken into consideration. Regarding the detailed "shape" of the observability measures O^{δ}, in the beginning, we can see a motion that could be described as a pumping motion in the diagonal direction emerging from the origin. By following the evolution, we can see that this pumping motion leaves behind traces that are later set into what could be described as arms of the structure that emerges in the end, cf. Figure 3. We refer to `https://systemstheorylab.wustl.edu/files/2018/03/vdp_og4-2m4k2sp.gif` for a high resolution animation