

RESEARCH ARTICLE | DECEMBER 17 2024

## Generalized synchronization in the presence of dynamical noise and its detection via recurrent neural networks

José M. Amigó  ; Roberto Dale ; Juan C. King ; Klaus Lehnertz 



*Chaos* 34, 123156 (2024)

<https://doi.org/10.1063/5.0235802>



### Articles You May Be Interested In

“Metric” complexity for weakly chaotic systems

*Chaos* (March 2007)

The two-parametric scaling and new temporal asymptotic of survival probability of diffusing particle in the medium with traps

*Chaos* (March 2017)

An analysis of the stability and bifurcation of a discrete-time predator–prey model with the slow–fast effect on the predator

*Chaos* (March 2024)



**Chaos**

**Special Topics Open  
for Submissions**

[Learn More](#)

# Generalized synchronization in the presence of dynamical noise and its detection via recurrent neural networks

Cite as: Chaos 34, 123156 (2024); doi: 10.1063/5.0235802

Submitted: 29 August 2024 · Accepted: 1 December 2024 ·

Published Online: 17 December 2024







View Online



Export Citation



CrossMark

José M. Amigó,<sup>1,a)</sup>  Roberto Dale,<sup>1,b)</sup>  Juan C. King,<sup>1,c)</sup>  and Klaus Lehnertz<sup>2,d)</sup> 

## AFFILIATIONS

<sup>1</sup>Centro de Investigación Operativa, Universidad Miguel Hernández, 03202 Elche, Spain

<sup>2</sup>Department of Epileptology, University of Bonn Medical Centre, Venusberg Campus 1, 53127 Bonn, Germany; Helmholtz Institute for Radiation and Nuclear Physics, University of Bonn, Nussallee 14–16, 53115 Bonn, Germany; and Interdisciplinary Center for Complex Systems, University of Bonn, Brühler Straße 7, 53175 Bonn, Germany

<sup>a)</sup> Author to whom correspondence should be addressed: [jm.amigo@umh.es](mailto:jm.amigo@umh.es)

<sup>b)</sup> Electronic mail: [rdale@umh.es](mailto:rdale@umh.es)

<sup>c)</sup> Electronic mail: [juan.king@goumh.umh.es](mailto:juan.king@goumh.umh.es)

<sup>d)</sup> Electronic mail: [klaus.lehnertz@ukbonn.de](mailto:klaus.lehnertz@ukbonn.de)

## ABSTRACT

Given two unidirectionally coupled nonlinear systems, we speak of generalized synchronization when the responder “follows” the driver. Mathematically, this situation is implemented by a map from the driver state space to the responder state space termed the synchronization map. In nonlinear times series analysis, the framework of the present work, the existence of the synchronization map amounts to the invertibility of the so-called cross map, which is a continuous map that exists in the reconstructed state spaces for typical time-delay embeddings. The cross map plays a central role in some techniques to detect functional dependencies between time series. In this paper, we study the changes in the “noiseless scenario” just described when noise is present in the driver, a more realistic situation that we call the “noisy scenario.” Noise will be modeled using a family of driving dynamics indexed by a finite number of parameters, which is sufficiently general for practical purposes. In this approach, it turns out that the cross and synchronization maps can be extended to the noisy scenario as families of maps that depend on the noise parameters, and only for “generic” driver states in the case of the cross map. To reveal generalized synchronization in both the noiseless and noisy scenarios, we check the existence of synchronization maps of higher periods (introduced in this paper) using recurrent neural networks and predictability. The results obtained with synthetic and real-world data demonstrate the capability of our method.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International (CC BY-NC-ND) license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). <https://doi.org/10.1063/5.0235802>

The first description of synchronization of two coupled dynamical systems (two pendulum clocks hanging from a beam) is attributed to Christiaan Huygens in 1665. In 1990, Pecora and Carroll demonstrated that chaotic systems that are unidirectionally coupled (i.e., drive-response systems) can also synchronize. By synchronization in both of the previous cases, we mean that the two systems evolve in finite time to a dynamic with a constant relationship between their states. In turn, chaotic synchronization gave rise to generalized synchronization, where now the relationship between states may be arbitrary. Precisely,

our work deals with a mathematical formulation of generalized synchronization in the more realistic case of drivers perturbed by dynamical noise. In addition, we do not assume knowledge of the states but only scalar observables of them in the form of time series. We also discuss other practical issues, most importantly, a method to detect generalized synchronization for both noiseless and noisy drivers, based on recurrent neural networks. The capability of this method is successfully tested with numerical simulations and real-world data.

## I. INTRODUCTION

The framework of this paper is nonlinear time series analysis, and the topic is synchronization of two unidirectionally coupled nonlinear systems and its generalization when noise is present in the driving system. By synchronization we mean *generalized* (or *general*) synchronization in the sense of Afraimovich *et al.*<sup>1</sup> and Rulkov *et al.*,<sup>2</sup> i.e., there is a map, called the *synchronization map* that transforms the states of the driving system (*driver*) into states of the driven system (*responder*), possibly with a time delay or after an initial transient time. Identical (complete, full, etc.) synchronization corresponds then to the synchronization map being the identity between two structurally equal systems; other, more interesting examples include lag, intermittent-lag, and phase synchronizations.<sup>3</sup> Synchronization, whether identical or generalized, plays an important role in many fields of science and engineering, particularly in nonlinear dynamics,<sup>4,5</sup> telecommunications,<sup>4,6,7</sup> neuroscience,<sup>8–10</sup> and cryptography;<sup>11–14</sup> see, e.g., Pikovsky *et al.*,<sup>15</sup> and Pecora and Carroll<sup>16</sup> for overviews and historical notes.

In the noiseless or fully deterministic scenario, synchronization has been extensively studied using a number of techniques, including cross (or mutual) prediction,<sup>17,18</sup> conditional Lyapunov exponents,<sup>4,19,20</sup> replica synchronization,<sup>21</sup> asymptotic stability of the responder,<sup>22</sup> nonlinear interdependence measures,<sup>23–25</sup> cellular nonlinear networks,<sup>26,27</sup> complexity measures extracted from symbolic representations,<sup>28–30</sup> reservoir computing,<sup>31,32</sup> and more. We will use prediction because predictability is a fingerprint of determinism, i.e., functional dependence.<sup>33</sup>

Furthermore, it is known<sup>18</sup> that, in the case of two unidirectionally coupled nonlinear systems with a noiseless driver, there exists typically a continuous map defined from the reconstructed state space of the responder to the reconstructed state space of the driver, which was called the closeness mapping in Amigó and Hirata<sup>34</sup> and will be called the *cross map* here. As it turns out, the definition of synchronization amounts to the invertibility (i.e., bijectivity) of the cross map; in fact, the inverse of the cross map is the “translation” of the synchronization map (if any) from the original domains (driver and responder state spaces) to the reconstructed ones. The existence of the cross map has been used to study interdependence and causal relationships in nonlinear time series analysis.<sup>34–37</sup> In a nutshell, these methods harness some actual or hypothetical property of the cross map (continuity, smoothness, or local expansiveness) to reveal, given bivariate time series of a coupled dynamics, what the driving system is. We will generalize the cross and synchronization maps to multi-time versions that are well suited to the application of recurrent neural networks in time series analysis.

The main objective of this paper is the extension of the cross and synchronization maps from noiseless to noisy drivers. To model noise in the driver, we replace the dynamic of a noiseless driver with a family of driving dynamics indexed by a finite number of parameters whose values are randomly chosen, an idea called finitely parameterized stochasticity.<sup>38</sup> To implement this idea in our setting, we will use the *stochastic forcing* approach of Stark *et al.*<sup>39</sup> First, noise is formulated as an autonomous dynamical system called a *shift system*, whose states comprise all possible noise realizations in form of parametric sequences; the  $n$ th component of a given sequence indicates which is the chosen driving dynamic at time  $n$ .

Second, the noisy driver is then formulated as a non-autonomous system, namely, a system forced by that shift system. As a result, our approach to synchronization in the presence of dynamical noise is based on state space reconstruction for unidirectionally coupled systems<sup>40</sup> and stochastic forcing<sup>39</sup> and is sufficiently general for practical purposes. We will show that the cross and synchronization maps can be extended from the noiseless to the noisy scenario by incorporating an additional dependence on the noise parameters and only for typical driver states in the case of the cross map.

In addition to discussing theoretical results in the noiseless and noisy scenarios, we also explore synchronization with synthetic and real world data. Prompted by multi-time expressions for the synchronization map, we not only use *perceptrons* but also *long short-term memory* (LSTM) *nets*.<sup>41</sup> In this approach, synchronization is detected via estimation of a responder state by a contemporaneous driver state or, in case of LSTM nets, by a segment of contemporaneous and past driver states. As the benchmark in numerical simulations, we chose *nearest neighbor cross prediction* because it is based precisely on the existence of the cross map, so it fits very well in our approach. In fact, the continuity of the cross map entails that near neighbors of a point in the responder state space map to near neighbors of its image in the driver state space (and vice versa when synchronization sets in). Therefore, one expects that this correspondence stays if low-amplitude noise affects the driving dynamic. We also apply LSTM nets to detect coupling directionality and synchronization in electroencephalograms (EEGs) from a subject with epilepsy. The optimization of the parameters and metaparameters of our numerical tools is beyond the scope of the present work.

To address the points described above, the rest of this paper is divided into a first, theoretical, and a second, numerical, part. Thus, in Sec. II, we first review the basics of our approach to make this paper self-contained. For didactic reasons, we start with the Takens and Stark (or forced Takens) embedding theorems (Sec. II A), along with the concept of cross map (Sec. II B); then we introduce the concept of generalized synchronization (Sec. II C) and discuss its relationship with the cross map (Sec. II D). Novel concepts such as the cross and synchronization maps of higher periods are introduced for further applications in Secs. VI and VII. The presentation is rigorous from a mathematical point of view but unnecessary technicalities are avoided. Along the way, practical issues are also considered with a view to the second part of the paper. Once the traditional, noiseless scenario has been presented, the noisy scenario is set in two steps: in Sec. III, we revisit stochastic forcing and an embedding theorem that is used in the second step, Sec. IV, where a unidirectional coupling with a noisy driver is modeled as stochastic forcing. The generalization of the cross and synchronization maps to the noisy scenario is the subject of Sec. V. The theoretical concepts of Secs. II–V are illustrated and put into practice in the second part of the paper. For this purpose, in Sec. VI, we resort to two unidirectionally coupled Hénon maps, synchronization being detected with recurrent neural networks and compare the results with those obtained with nearest-neighbor cross predictability. In Sec. VII, we tackle the applicability of our tools to the analysis of real data in the form of EEGs, where noise and bidirectional coupling are the rules. In this case, the “driver” is identified by the direction with the strongest coupling. Our findings are discussed

in light of results published in the specialized literature. Finally, the main contributions and conclusions of this paper are summarized in Sec. VIII.

## II. THE NOISELESS SCENARIO

This section is a compact, mathematically oriented account of the cross map, synchronization, and their interplay in the absence of noise.

### A. Embedding theorems

Following Stark,<sup>40</sup> let  $Y$  be a non-autonomous dynamical system (the responder) evolving under the influence of an autonomous dynamical system  $X$  (the driver).  $X$  is also called the driving or forcing system, and  $Y$  is the driven or forced system. In the case of discrete-time deterministic dynamical systems or flows observed at discrete times, this situation is described by the difference equations

$$\begin{cases} x_{t+1} = f(x_t), \\ y_{t+1} = g(x_t, y_t), \end{cases} \quad (1)$$

where (i)  $x_t \in \mathcal{M}_X$  is the state of system  $X$  at time  $t$ ; (ii)  $y_t \in \mathcal{M}_Y$  is the state of system  $Y$  at time  $t$ ; (iii)  $\mathcal{M}_X$  and  $\mathcal{M}_Y$  are compact manifolds of dimensions  $\dim_X \geq 0$  and  $\dim_Y \geq 1$ , respectively; (iv)  $f: \mathcal{M}_X \rightarrow \mathcal{M}_X$  is a  $C^1$  diffeomorphism (i.e., a  $C^1$  invertible map such that  $f^{-1}$  is also  $C^1$ , where  $C^1$  is shorthand for continuously differentiable); and (v)  $g: \mathcal{M}_X \times \mathcal{M}_Y \rightarrow \mathcal{M}_Y$  is a  $C^1$  map such that  $g(x, \cdot)$  is a diffeomorphism of  $\mathcal{M}_Y$  for every  $x \in \mathcal{M}_X$ . Alternatively, we say that there is a (unidirectional) coupling from  $X$  to  $Y$  and use the shorthand  $X \rightsquigarrow Y$ . Since we assume that  $f$  and  $g(x, \cdot)$  are invertible, we may take  $t \in \mathbb{Z}$ , although in applications time series have a beginning that we will set at  $t = 0$ .

**Remark 1.** By defining the map

$$[f, g](x, y) = (f(x), g(x, y)), \quad (2)$$

the forced system (1) becomes an autonomous dynamical system on the full state space  $\mathcal{M}_X \times \mathcal{M}_Y$ , called the *skew product* of  $f$  and  $g$ . Due to the properties (ii)–(v) above,  $[f, g]$  is a  $C^1$  diffeomorphism of the compact manifold  $\mathcal{M}_X \times \mathcal{M}_Y$ .

As stated in the Introduction, we are mainly interested in nonlinear time series analysis. So, suppose further that the only information available about the systems  $X$  and  $Y$  are scalar observations  $\varphi_X(x_t)$  of the states  $x_t$  and  $\varphi_Y(y_t)$  of the states  $y_t$ , where the observation functions  $\varphi_X: \mathcal{M}_X \rightarrow \mathbb{R}$  and  $\varphi_Y: \mathcal{M}_Y \rightarrow \mathbb{R}$  are assumed to be  $C^1$ . To reconstruct the state spaces of the driver  $X$  and the responder  $Y$  from the corresponding observations  $\varphi_X(x_t)$  and  $\varphi_Y(y_t)$ , we use the Takens and Stark theorems, respectively, which we remind below for further reference.

**Theorem 2** [Takens theorem<sup>42</sup>]. *If  $d \geq 2 \dim_X + 1$ , then the map  $E_{f, \varphi_X}: \mathcal{M}_X \rightarrow \mathbb{R}^d$  defined as*

$$E_{f, \varphi_X}(x) = (\varphi_X(f^0(x)), \varphi_X(f^1(x)), \dots, \varphi_X(f^{d-1}(x))) \quad (3)$$

*is an embedding for generic  $f$  and  $\varphi_X$ .*

As usual,  $f^0$  is the identity,  $f^1 = f$ , and  $f^n$  is the  $n$ th iterate of  $f$ . Here, “generic  $f$  and  $\varphi_X$ ” formally means that the set  $\{f, \varphi_X\}$  for which  $E_{f, \varphi_X}$  is an embedding (i.e., a  $C^1$  diffeomorphism onto its image) is open and dense in the  $C^1$ -topology (uniform convergence of a map

and its derivative) of the respective function spaces, namely, diffeomorphisms of  $\mathcal{M}_X$ , and  $C^1$  maps from  $\mathcal{M}_X$  to  $\mathbb{R}$ . In general, a property is *generic* in a topological space  $\mathcal{T}$  if it holds on a *residual subset*  $\mathcal{S} \subset \mathcal{T}$ , i.e., on a subset that contains a countable intersection of open sets. It turns out that an open and dense set of maps  $f$  for which  $E_{f, \varphi_X}$  is an embedding for generic  $\varphi_X$  is built by those  $C^1$  diffeomorphisms of  $\mathcal{M}_X$  that have only a finite number of periodic orbits of period less than  $d$ , and the eigenvalues of each such periodic orbits are distinct (Stark,<sup>40</sup> Theorem 2.2).

**Remark 3.** Theorem 2 was generalized by Sauer *et al.*<sup>43</sup> in two ways. First, by replacing “generic” with “probability-one” (in the sense of prevalence). Second, by replacing the manifold  $\mathcal{M}_X$  by a compact invariant set  $A$  that may have fractal box-counting dimension, and the restriction  $d \geq 2 \dim_X + 1$  (which comes from Whitney’s embedding theorem<sup>44</sup>) by  $d \geq 2 \text{boxdim}(A) + 1$ , where  $\text{boxdim}(A)$  is the box-counting dimension of  $A$ .

For our purposes, we need to generalize Theorem 2 to the forced dynamic

$$(x_{t+1}, y_{t+1}) = [f, g](x_t, y_t)$$

defined by the diffeomorphism (2) in the full state space  $\mathcal{M}_X \times \mathcal{M}_Y$ . As before, set  $[f, g]^0 = \text{identity}$ ,  $[f, g]^1 = [f, g]$  and  $[f, g]^{t+1} = [f, g] \circ [f, g]^t$  for the iterates of  $[f, g]$ , so that

$$[f, g]^t(x, y) = (f^t(x), g^{(t)}(x, y)), \quad (4)$$

where  $g^{(t)}(x, y): \mathcal{M}_X \times \mathcal{M}_Y \rightarrow \mathcal{M}_Y$  is recursively defined by  $g^{(0)}(x, y) = y$  and

$$g^{(t)}(x, y) = g(f^{t-1}(x), g^{(t-1)}(x, y)) \quad (5)$$

for  $t \geq 1$ . Application of the Takens theorem to the skew product  $[f, g]$  would provide a map  $E_{[f, g], \varphi_{X, Y}}: \mathcal{M}_X \times \mathcal{M}_Y \rightarrow \mathbb{R}^D$ , with  $D \geq 2(\dim_X + \dim_Y) + 1$ , which would be an embedding for open dense sets of diffeomorphisms of  $\mathcal{M}_X \times \mathcal{M}_Y$  and observation  $C^1$  maps  $\varphi_{X, Y}(x, y): \mathcal{M}_X \times \mathcal{M}_Y \rightarrow \mathbb{R}$ , in the  $C^1$  topology of the respective function spaces. However, what we need for applications to nonlinear time series analysis is an embedding for generic maps  $f$ ,  $g$ , and observation maps  $\varphi_Y$  on  $\mathcal{M}_Y$ , and this is not guaranteed by this approach.

The generalization of the Takens theorem to forced dynamics that we need is the following due to Stark.

**Theorem 4** [forced Takens theorem<sup>40</sup>]. *If  $D \geq 2(\dim_X + \dim_Y) + 1$ , then the map  $E_{f, g, \varphi_Y}: \mathcal{M}_X \times \mathcal{M}_Y \rightarrow \mathbb{R}^D$  defined as*

$$E_{f, g, \varphi_Y}(x, y) = (\varphi_Y(g^{(0)}(x, y)), \varphi_Y(g^{(1)}(x, y)), \dots, \varphi_Y(g^{(D-1)}(x, y))) \quad (6)$$

*is an embedding for generic  $f$ ,  $g$ , and  $\varphi_Y$ .*

Specifically, generic  $g$  means that  $E_{f, g, \varphi_Y}$  is an embedding for an open and dense set of diffeomorphisms  $g(x, y)$  (such that  $g(x, \cdot)$  is a diffeomorphism of  $\mathcal{M}_Y$  for every  $x \in \mathcal{M}_X$ ) in the  $C^1$ -topology of  $\mathcal{M}_X \times \mathcal{M}_Y$ . In this case, an open and dense set of maps  $f$  for which  $E_{f, g, \varphi_Y}$  is an embedding for generic  $g$  and  $\varphi_Y$  is built by those  $C^1$  diffeomorphisms of  $\mathcal{M}_X$  whose periodic orbits of period less than  $2d$  are isolated and have distinct eigenvalues (Stark,<sup>40</sup> Theorem 3.1).

**B. The cross map**

Hereinafter, we tacitly assume that  $f, g, \varphi_X$ , and  $\varphi_Y$  are generic in the sense of Theorems 2 and 4. Also, “smooth” stands for  $C^1$  smoothness in the following.

Given the scalar observations  $(\varphi_X(x_t))_{t \in \mathbb{Z}}$  and  $(\varphi_Y(y_t))_{t \in \mathbb{Z}}$ , Theorems 2 and 4 allow to “reconstruct” the (possibly unknown) dynamics of the underlying systems  $X$  and  $Y$  in the manifolds

$$\mathcal{N}_X = E_{f, \varphi_X}(\mathcal{M}_X) \subset \mathbb{R}^d$$

and

$$\mathcal{N}_Y = E_{f, g, \varphi_Y}(\mathcal{M}_X \times \mathcal{M}_Y) \subset \mathbb{R}^D$$

called the *reconstructed driver* and *responder state spaces*, respectively, by means of the *time-delay vectors*

$$\mathbf{x}_t = E_{f, \varphi_X}(x_t) = (\varphi_X(x_t), \varphi_X(x_{t+1}), \dots, \varphi_X(x_{t+d-1})) \in \mathbb{R}^d \quad (7)$$

and

$$\mathbf{y}_t = E_{f, g, \varphi_Y}(x_t, y_t) = (\varphi_Y(y_t), \varphi_Y(y_{t+1}), \dots, \varphi_Y(y_{t+D-1})) \in \mathbb{R}^D. \quad (8)$$

In turn, the dynamics  $x_{t+1} = f(x_t)$  in  $\mathcal{M}_X$  translates into the *reconstructed driving dynamics*

$$\mathbf{x}_{t+1} = (E_{f, \varphi_X} \circ f \circ E_{f, \varphi_X}^{-1})(\mathbf{x}_t) =: \tilde{f}(\mathbf{x}_t) \quad (9)$$

in  $\mathcal{N}_X$ , while the dynamics  $(x_{t+1}, y_{t+1}) = [f, g](x_t, y_t)$  in  $\mathcal{M}_X \times \mathcal{M}_Y$  translates into the *reconstructed coupled dynamics*

$$\mathbf{y}_{t+1} = (E_{f, g, \varphi_Y} \circ [f, g] \circ E_{f, g, \varphi_Y}^{-1})(\mathbf{y}_t) =: \tilde{[f, g]}(\mathbf{y}_t) \quad (10)$$

in  $\mathcal{N}_Y$ , the manifolds  $\mathcal{N}_X$  and  $\mathcal{N}_Y$  being diffeomorphic copies of  $\mathcal{M}_X$  and  $\mathcal{M}_X \times \mathcal{M}_Y$ , respectively. Therefore, all coordinate-independent properties of  $f$  and  $[f, g]$  can be determined in  $\mathcal{N}_X$  and  $\mathcal{N}_Y$ .

**Remark 5.** Without loss of generality, it can be assumed that  $d = D$ . In nonlinear time series analysis, where the underlying dynamical system is unknown, the embedding dimension of a time series is usually chosen by the method of false nearest neighbors.<sup>45</sup>

Let  $\Pi_X : \mathcal{M}_X \times \mathcal{M}_Y \rightarrow \mathcal{M}_X$  be the projection onto  $\mathcal{M}_X$ , i.e.,  $\Pi_X(x, y) = x$ . From the diagram

$$\begin{array}{ccc} \mathcal{M}_X \times \mathcal{M}_Y \ni (x_t, y_t) & \xrightarrow{\Pi_X} & x_t \in \mathcal{M}_X \\ E_{f, g, \varphi_Y}^{-1} \uparrow & & \downarrow E_{f, \varphi_X} \\ \mathcal{N}_Y \ni \mathbf{y}_t & & \mathbf{x}_t \in \mathcal{N}_X \end{array} \quad (11)$$

along with the smoothness of the embeddings  $E_{f, g, \varphi_Y}^{-1}, E_{f, \varphi_X}$ , and the projection  $\Pi_X$ , we conclude the following proposition.

**Proposition 6.** A unidirectional coupling  $X \rightsquigarrow Y$  necessarily implies the existence of a smooth map

$$\Phi := E_{f, \varphi_X} \circ \Pi_X \circ E_{f, g, \varphi_Y}^{-1} : \mathcal{N}_Y \rightarrow \mathcal{N}_X \quad (12)$$

called the *cross map of the coupling*  $X \rightsquigarrow Y$ , which sends  $\mathbf{y}_t$  to  $\mathbf{x}_t$ , i.e.,

$$\mathbf{x}_t = \Phi(\mathbf{y}_t). \quad (13)$$

Intuitively, Eq. (13) spells out that the responder signal carries information about the dynamics of the driver because of the time evolution law  $y_{t+1} = g(x_t, y_t)$ .

**Remark 7.** Equation (13) is equivalent to the existence of a map

$$\mathbf{x}_t = \Phi^{(k)}(\mathbf{y}_{t-k}), \quad (14)$$

for any  $k \in \mathbb{Z}$ , where  $\Phi^{(k)} : \mathcal{N}_Y \rightarrow \mathcal{N}_X$ . Indeed, from

$$\mathbf{x}_t = \tilde{f}^k(\mathbf{x}_{t-k}) \quad (15)$$

[see Eq. (9)] and  $\mathbf{x}_{t-k} = \Phi(\mathbf{y}_{t-k})$ , it follows

$$\Phi^{(k)} = \tilde{f}^k \circ \Phi = E_{f, \varphi_X} \circ \tilde{f}^k \circ E_{f, \varphi_X}^{-1} \circ \Phi \quad (16)$$

and, hence,

$$\mathbf{x}_t = \frac{1}{K} \sum_{k=0}^{K-1} \Phi^{(k)}(\mathbf{y}_{t-k}) =: \Phi_K(\mathbf{y}_t, \mathbf{y}_{t-1}, \dots, \mathbf{y}_{t-K+1}) \quad (17)$$

for all  $K \geq 1$ . Note that  $\Phi^{(k)}$  and  $\Phi_K$  are continuous, and  $\Phi^{(0)} = \Phi_1 = \Phi$ .

By changing the summation limits in (17), one can construct other similar multi-time expressions. For definiteness, we will use only definition (17).

**Definition 8.** We call the continuous map  $\mathbf{x}_t = \Phi^{(k)}(\mathbf{y}_{t-k})$  the *cross map of order*  $k \in \mathbb{Z}$ , and the continuous map  $\mathbf{x}_t = \Phi_K(\mathbf{y}_t, \mathbf{y}_{t-1}, \dots, \mathbf{y}_{t-K+1})$  the *cross map of period*  $K \geq 1$ .

The continuity of the cross map  $\Phi$  has been used in nonlinear time series analysis to discriminate functional (deterministic, causal, etc.) relationships between observations due to coupled dynamics from statistical correlation. In its simplest version, the continuity of the cross map  $\mathbf{x}_t = \Phi(\mathbf{y}_t)$  belonging to the coupling  $X \rightsquigarrow Y$  implies that, given an open ball  $B_\varepsilon(\mathbf{x}_t) \subset \mathbb{R}^d$  with center  $\mathbf{x}_t$  and arbitrary radius  $\varepsilon > 0$ , there exists an open ball  $B_\delta(\mathbf{y}_t) \subset \mathbb{R}^D$  with center  $\mathbf{y}_t$  and radius  $\delta = \delta(\varepsilon) > 0$  such that  $\Phi(B_\delta(\mathbf{y}_t)) \subset B_\varepsilon(\mathbf{x}_t)$ . Therefore, the  $k$  nearest neighbors  $\mathbf{y}_{t_1}, \dots, \mathbf{y}_{t_k}$  of a time-delay vector  $\mathbf{y}_t \in \mathcal{N}_Y$  in a time series  $(\mathbf{y}_t)_{0 \leq t \leq T}$  of the responder are mapped by  $\Phi$  to close neighbors  $\mathbf{x}_{t_1}, \dots, \mathbf{x}_{t_k}$  of the contemporaneous vector  $\mathbf{x}_t \in \mathcal{N}_X$  in the corresponding time series  $(\mathbf{x}_t)_{0 \leq t \leq T}$  of the driver. Methods that take advantage of the continuity of  $\Phi$  in this way to unveil the coupling  $X \rightsquigarrow Y$  include *cross prediction*,<sup>18</sup> *convergent cross mapping*,<sup>35</sup> and *continuity scaling*.<sup>37</sup>

**C. Generalized synchronization**

According to Rulkov *et al.*<sup>2</sup> and Pikovsky *et al.*,<sup>15</sup> the systems  $X \rightsquigarrow Y$  are in *generalized* (or *general*) *synchronization* if there exists a continuous map  $h : \mathcal{M}_X \rightarrow \mathcal{M}_Y$  such that

$$y_t = h(x_t), \quad (18)$$

for all  $t \in \mathbb{Z}$ . That is, the responder follows the driver but in a weaker form than in identical synchronization, which corresponds to  $h$  being the identity (i.e.,  $X$  and  $Y$  are structurally the same and  $y_t = x_t$ ). We will also say that  $Y$  is synchronized to  $X$  if Eq. (18) holds and call  $h(x)$  the *synchronization map*.

Therefore, in case of synchronization the full state space  $\mathcal{M}_X \times \mathcal{M}_Y$  shrinks into the subspace  $\{(x, y) \in \mathcal{M}_X \times \mathcal{M}_Y : y = h(x)\}$ , which is the graph of the synchronization map  $x \mapsto h(x)$ . This subspace is usually called the *synchronization manifold*, even

when  $h$  is not smooth. It follows then that the projection map from  $\mathcal{M}_X \times \mathcal{M}_Y$  onto  $\mathcal{M}_X$ ,  $\Pi_X(x, y) = x$ , is invertible,

$$\Pi_X^{-1}(x) = (x, h(x)), \tag{19}$$

and the range of  $\Pi_X^{-1} : \mathcal{M}_X \rightarrow \mathcal{M}_X \times \mathcal{M}_Y$  is the synchronization manifold.

**Remark 9.** Plug the driver dynamic  $x_t = f(x_{t-1})$  into Eq. (18) to derive

$$y_t = h(x_t) = (h \circ f)(x_{t-1}) = \dots = (h \circ f^k)(x_{t-k}) =: h^{(k)}(x_{t-k}), \tag{20}$$

where  $h^{(k)} = h \circ f^k : \mathcal{M}_X \rightarrow \mathcal{M}_Y$  is continuous for all  $k \geq 0$ , and  $h^{(0)} = h$ . Equation (20) with  $k \geq 1$  corresponds to generalized synchronization for responders with an internal delay loop. When  $f$  is invertible (as in our case), the generalized synchronization of the systems  $X$  and  $Y$  can indistinctly be defined by  $y_t = h(x_t)$  or, more generally, by  $y_t = h^{(k)}(x_{t-k})$  for  $k \in \mathbb{Z}$ ; in the latter case,  $h = h^{(k)} \circ f^{-k}$ . The maps  $h^{(k)} : \mathcal{M}_X \rightarrow \mathcal{M}_Y$  are called *synchronization maps of order  $k$* , orders 0 and 1 being the usual choices in applications.

From Eq. (20), it trivially follows that

$$y_t = \frac{1}{K} \sum_{k=0}^{K-1} h^{(k)}(x_{t-k}) =: h_K(x_t, x_{t-1}, \dots, x_{t-K+1}), \tag{21}$$

for all  $K \geq 1$ , in case  $Y$  is synchronized with  $X$ . By changing the summation limits in Eq. (21), one can construct other similar expressions. For definiteness, we will use only Eq. (21) in this paper, so that  $h_1 = h$ .

Therefore, to detect synchronization of a time series  $\{y_t\}_{t \geq 0}$  with another time series  $\{x_t\}_{t \geq 0}$ , we can look for functional dependencies of the form (21) with  $K > 1$  rather than  $y_t = h(x_t)$ . If  $\{x_t\}_{t \geq 0}$  is a deterministic time series (i.e.,  $x_{t+1} = f(x_t)$ ) and  $\{y_t\}_{t \geq 0}$  is synchronized with it, then Eq. (21) holds with a continuous map  $h_K : \mathcal{M}_X^K \rightarrow \mathcal{M}_Y$  such that  $h_K(x_t, f^{-1}(x_t), \dots, f^{-K+1}(x_t)) = h(x_t)$ . The point is that, in time series analysis, multi-time dependencies like (21) can be efficiently detected by recurrent neural nets, as we discuss in Sec. VI.

**Definition 10.** We call the continuous map  $y_t = h_K(x_t, \dots, x_{t-K+1})$  in Eq. (21), the *synchronization map of period  $K \geq 1$* .

The synchronization maps of order  $k$ ,  $h^{(k)} = h \circ f^k$ , satisfy a number of straightforward relations involving also the function  $g(x, y)$ . Indeed, in case of synchronization, the dynamic (1) of  $X \rightsquigarrow Y$  simplifies to

$$\begin{cases} x_{t+1} = f(x_t), \\ y_{t+1} = g(x_t, h(x_t)). \end{cases} \tag{22}$$

Comparing with  $y_{t+1} = h(f(x_t))$  shows that  $h(x)$  fulfills the functional relation

$$h(f(x)) = g(x, h(x)). \tag{23}$$

Replace  $x$  with  $f^k(x)$  in Eq. (23) to obtain

$$h^{(k+1)}(x) = h^{(k)}(f(x)) = g(f^k(x), h^{(k)}(x)). \tag{24}$$

Recursion of Eq. (24) leads to alternative formulas for synchronization maps of arbitrary periods involving function  $g$ .

Contingent upon the structure of  $g(x, y)$ , the synchronization map  $h(x)$  can sometimes be written in closed form; see Pikovsky *et al.*<sup>15</sup> and Parlitz<sup>46</sup> for an example with a baker map. Interestingly,

the parameters of that example can be fine-tuned so that the cross sections  $x^{(2)} = \text{const}$  of  $h(x^{(1)}, x^{(2)})$  are Weierstrass functions, i.e., continuous functions that are nowhere differentiable.

The definition of synchronization can be weakened by requiring condition (18) only asymptotically. In more formal terms, we say that the responder  $Y$  is *asymptotically synchronized* to the driver  $X$  if there exists a continuous map  $h : \mathcal{M}_X \rightarrow \mathcal{M}_Y$  such that

$$\lim_{t \rightarrow \infty} \|y_t - h(x_t)\| = 0, \tag{25}$$

where  $\|\cdot\|$  is a distance in  $\mathcal{M}_Y$ . In this case, the synchronization manifold becomes an attracting set in  $\mathcal{M}_X \times \mathcal{M}_Y$ .

A direct consequence of asymptotic synchronization is the *asymptotic stability* of the responder. We say that the responder  $Y$  is asymptotically stable if all orbits converge to the same orbit regardless of the initial condition, that is, if given two responses  $(y_t)_{t \geq 0}$  and  $(\tilde{y}_t)_{t \geq 0}$  to a signal  $(x_t)_{t \geq 0}$  from the driver with different initial conditions  $y_0 \neq \tilde{y}_0$ , then

$$\lim_{t \rightarrow \infty} \|y_t - \tilde{y}_t\| = 0. \tag{26}$$

Asymptotic stability of the responder is weaker than asymptotic synchronization because the existence of a hypothetical synchronization map does not follow from Eq. (26). This is the case, for example, when a periodic driver has gone through a period-doubling bifurcation<sup>47</sup> or there is a multistability in the responder, i.e., a driver signal  $(x_t)_{t \geq 0}$  can elicit two or more stable responses.<sup>15</sup>

On the other hand, the asymptotic stability of the responder provides a simple method to test synchronization called the *auxiliary system method*.<sup>21</sup> This method boils down to check Eq. (26) for two initial conditions  $y_0 \neq \tilde{y}_0$ ; if (26) does not hold, then  $Y$  is not synchronized to  $X$ .

#### D. Relationship between generalized synchronization and the cross map

According to Proposition 6, a coupling  $X \rightsquigarrow Y$  implies the existence of the cross map  $\mathbf{x}_t = \Phi(\mathbf{y}_t)$ , whereas the synchronization map  $y_t = h(x_t)$  exists in seemingly exceptional cases (unless the coupling is strong enough). Despite this notable difference, both maps are closely related, as we will now see.

Let  $\Pi_Y(x, y) = y$  be the projection map from  $\mathcal{M}_X \times \mathcal{M}_Y$  onto  $\mathcal{M}_Y$ . On the one hand, from the diagram

$$\begin{array}{ccccc} \mathcal{M}_X \ni x_t & \xrightarrow{h} & y_t \in \mathcal{M}_Y & & \\ E_{f, \varphi_X} \downarrow & & \uparrow \Pi_Y \circ E_{f, g, \varphi_Y}^{-1} & & \\ \mathcal{N}_X \ni \mathbf{x}_t & \xleftarrow{\Phi} & \mathbf{y}_t \in \mathcal{N}_Y & & \end{array} \tag{27}$$

we have that if  $\Phi$  is invertible, then  $h$  exists and

$$h = \Pi_Y \circ E_{f, g, \varphi_Y}^{-1} \circ \Phi^{-1} \circ E_{f, \varphi_X}. \tag{28}$$

On the other hand, Eq. (19) spells out that, if  $h$  exists, then  $\Pi_X$  is invertible. From Eq. (12), it follows then that

$$\Phi^{-1} = E_{f, g, \varphi_Y} \circ \Pi_X^{-1} \circ E_{f, \varphi_X}^{-1}. \tag{29}$$

The bottom line of Eqs. (28) and (29) is the following.

**Proposition 11.** *The systems  $X \rightsquigarrow Y$  are synchronized if and only if the cross map  $\mathbf{x}_t = \Phi(\mathbf{y}_t)$  is invertible and bicontinuous*

(i.e.,  $\Phi^{-1}$  is continuous). If this case, the synchronization map  $y_t = h(x_t)$  and  $\mathbf{y}_t = \Phi^{-1}(\mathbf{x}_t)$  are related through the expressions (28) and (29), where  $\Pi_X^{-1}(x) = (x, h(x))$ .

In other words,  $\Phi^{-1} : \mathcal{N}_X \rightarrow \mathcal{N}_Y$  is the synchronization map of the systems  $X \rightsquigarrow Y$  in the reconstructed state spaces (if it exists and is continuous); see Pecora *et al.*<sup>48</sup> for numerical methods to test whether two time series are related by a map with properties such as continuity, invertibility, smoothness, and more. As a rule, the relationship  $x \mapsto (x, y)$  is multivalued owing to folds in the manifold  $\mathcal{M}_X \times \mathcal{M}_Y$ , so generalized synchronization is rather an exception. Multivalued synchronization maps, corresponding to noninvertible cross maps, have been considered, e.g., in Rulkov *et al.*<sup>49</sup> and Parlitz.<sup>46</sup>

To lift  $y_t = h_K(x_t, x_{t-1}, \dots, x_{t-K+1})$ , the synchronization map of period  $K$  (21), to the reconstructed state spaces  $\mathcal{N}_X$  and  $\mathcal{N}_Y$ , use the reconstructed driver dynamic  $\tilde{f} : \mathcal{N}_X \rightarrow \mathcal{N}_X$  defined in Eq. (9) to derive

$$\mathbf{y}_t = \Phi^{-1}(\mathbf{x}_t) = (\Phi^{-1} \circ \tilde{f})(\mathbf{x}_{t-1}) = \dots = (\Phi^{-1} \circ \tilde{f}^k)(\mathbf{x}_{t-k}), \quad (30)$$

where  $k \in \mathbb{Z}$  and

$$\Phi^{-1} \circ \tilde{f}^k = \Phi^{-1} \circ E_{f_{\phi_X}} \circ f^k \circ E_{f_{\phi_X}}^{-1} = (\Phi^{(-k)})^{-1}, \quad (31)$$

by the definition of the cross map of order  $k$ , Eq. (16). Therefore, the synchronization map of period  $K \geq 1$  (21) becomes

$$\mathbf{y}_t = \frac{1}{K} \sum_{k=0}^{K-1} (\Phi^{(-k)})^{-1}(\mathbf{x}_{t-k}) =: H_K(\mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_{t-K+1}) \quad (32)$$

in the reconstructed spaces. Note that  $H_1 = \Phi^{-1}$ .

**Definition 12.** The continuous map  $\mathbf{y}_t = H_K(\mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_{t-K+1})$  defined in Eq. (32) will be called the *reconstructed synchronization map of period  $K \geq 1$* .

We will harness  $H_K$  with  $K > 1$  in the applications with synthetic data (Sec. VI) and real world data (Sec. VII) via recurrent neural networks. We remark already at this point that, unlike the synthetic data of Sec. VI, real world data are generally bidirectionally coupled, as happens with the EEGs of Sec. VII. Following the standard approach, we will measure the coupling strength between pairs of EEGs in both directions, the “driver” being identified by the direction with the strongest coupling. In case of equal strengths, the systems are assumed to be synchronized.

As mentioned in Sec. II B, in case of unidirectional coupling (the framework of this paper) the relationship  $\mathbf{x}_t = \Phi(\mathbf{y}_t)$  due to the coupled dynamic  $X \rightsquigarrow Y$  can be unveiled numerically from the time series  $(\mathbf{x}_t)_{0 \leq t \leq T}$  and  $(\mathbf{y}_t)_{0 \leq t \leq T}$ .<sup>35,37,48</sup> Thus, in the method of nearest-neighbor cross prediction, one estimates  $\mathbf{x}_t$  or  $\mathbf{x}_{t+1}$  based on the nearest neighbors of  $\mathbf{y}_t$  for any  $0 \leq t \leq T$  to test for the existence of  $\Phi$ . Likewise, if  $\Phi^{-1}$  exists and is continuous (i.e.,  $X$  and  $Y$  are synchronized), then one can also discern the inverse relationship  $\mathbf{y}_t = \Phi^{-1}(\mathbf{x}_t)$  by the same techniques. As a matter of fact, in the case of a bijective and bicontinuous  $\Phi$ , there is a one-to-one relation between the neighborhoods of nearest neighbors of  $\mathbf{y}_t$  and  $\mathbf{x}_t$ , so, if  $Q_T(\Phi)$  and  $Q_T(\Phi^{-1})$  are fidelity metrics of the respective estimations, then

$$Q_T(\Phi) - Q_T(\Phi^{-1}) \simeq 0, \quad (33)$$

where it applies that the longer the time series, the better the predictions and, hence, the smaller  $Q_T(\Phi) - Q_T(\Phi^{-1})$ . In other words, the continuity of the cross map and its inverse can be exploited via (33) to test two time series  $(\mathbf{x}_t)_{0 \leq t \leq T}$  and  $(\mathbf{y}_t)_{0 \leq t \leq T}$  for general synchronization.

**Remark 13.** The nonexistence of the cross map can uncover common drivers. Indeed, if  $Z \rightsquigarrow X$  and  $Z \rightsquigarrow Y$ , then  $\mathbf{z}_t = \Phi(\mathbf{x}_t)$  and  $\mathbf{z}_t = \tilde{\Phi}(\mathbf{y}_t)$ , so that  $\Phi(\mathbf{x}_t) = \tilde{\Phi}(\mathbf{y}_t)$ . Here,  $\Phi : \mathcal{N}_X \rightarrow \mathcal{N}_Z$  and  $\tilde{\Phi} : \mathcal{N}_Y \rightarrow \mathcal{N}_Z$  are the cross maps associated to the coupled dynamics  $Z \rightsquigarrow X$  and  $Z \rightsquigarrow Y$ , respectively. However, there is no cross map between  $\mathcal{N}_X$  and  $\mathcal{N}_Y$ , unless  $\Phi$  or  $\tilde{\Phi}$  is invertible and continuous [so that  $\mathbf{x}_t = (\Phi^{-1} \circ \tilde{\Phi})(\mathbf{y}_t)$  or  $\mathbf{y}_t = (\tilde{\Phi}^{-1} \circ \Phi)(\mathbf{x}_t)$ ], in which case  $X$  or  $Y$  is synchronized with  $Z$ .

To wrap up this section, let us point out that the diagram (27) is a particularization of the diagram

$$\begin{array}{ccc} \mathcal{M}_X \ni x_t & \xrightarrow{h^{(k)}} & y_{t+k} \in \mathcal{M}_Y \\ E_{f_{\phi_X}} \downarrow & & \uparrow \Pi_Y \circ E_{f_{\phi_Y}}^{-1} \\ \mathcal{N}_X \ni \mathbf{x}_t & \xleftarrow{\Phi^{(-k)}} & \mathbf{y}_{t+k} \in \mathcal{N}_Y \end{array} \quad (34)$$

to  $k = 0$ , where  $h^{(k)}$  is the synchronization map of order  $k$  [Eq. (20),  $h^{(0)} = h$ ] and  $\Phi^{(-k)}$  is the cross map of order  $-k$  [Eq. (14),  $\Phi^{(0)} = \Phi$ ]. Thus, Eq. (28) is the special case  $k = 0$  of the relationship

$$\begin{aligned} h^{(k)} &= \Pi_Y \circ E_{f_{\phi_Y}}^{-1} \circ (\Phi^{(-k)})^{-1} \circ E_{f_{\phi_X}} \\ &= \Pi_Y \circ E_{f_{\phi_Y}}^{-1} \circ \Phi^{-1} \circ E_{f_{\phi_X}} \circ f^k, \end{aligned} \quad (35)$$

where we used (31) and (9) in the second line. In this regard, note that  $h^{(k)}$  is invertible if and only if  $h$  is invertible [since  $h^{(k)} = h \circ f^k$  by Eq. (20)] and, likewise,  $\Phi^{(-k)}$  is invertible if and only if  $\Phi$  is invertible [since  $\Phi^{(-k)} = E_{f_{\phi_X}} \circ f^{-k} \circ E_{f_{\phi_X}}^{-1} \circ \Phi$  by Eq. (16)].

### III. DYNAMICAL NOISE AS STOCHASTIC FORCING

Random or “noisy” dynamical systems can be modeled in different ways, from the perhaps simplest ones, such as switching systems<sup>50,51</sup> and iterated function systems<sup>52</sup> to nonautonomous dynamical systems<sup>53</sup> and full-fledged random dynamical systems, described by random differential equations.<sup>54</sup> In our setting, a natural way to turn a noiseless dynamic, say,  $x_{t+1} = f(x_t)$  on a compact manifold  $\mathcal{M}_X$ , into a noisy one is to replace the map  $f$  with a family of maps  $\{f_{\omega_t}\}_{t \in \mathbb{Z}}$ , where the index  $\omega_t$  is a (possibly multi-component) parameter belonging to a suitable space that is randomly chosen at each discrete time  $t$ . For example,  $f_{\omega_t}(x_t) = f(x_t) + \omega_t$ ,  $\omega_t \in \mathcal{M}_X$ , models additive dynamical noise, while  $f_{\omega_t}(x_t) = \omega_t f(x_t)$  models multiplicative dynamical noise. This approach, sometimes called “finitely parameterized stochasticity”<sup>38</sup> is sufficient for most practical applications.<sup>55</sup> So, the term *noisy dynamical system* will refer hereafter to such implementation of dynamical noise via stochastic processes in the parameter space; our parameter spaces will be compact topological sets.

At this point, we recall that each *stationary* (discrete-time) stochastic process corresponds in a canonical way to a so-called *shift system*, which is a dynamical system whose states are the realizations of the stochastic process considered.<sup>56</sup> In other words, stationary stochastic processes can be modeled as autonomous dynamical systems. It is, therefore, not surprising that shift systems allow to

formulate noisy dynamical systems as forced systems in a manner formally similar to the noiseless case. However, before getting to that point, we need to introduce the concepts and notation.

Let  $\Omega$  be a compact topological space of parameters and let  $\Omega^{\mathbb{Z}}$  be the set of all two-sided sequences of elements of  $\Omega$ ,

$$\omega = (\dots, \omega_{-t}, \dots, \omega_{-1}, \omega_0, \omega_1, \dots, \omega_t, \dots)$$

endowed with the product topology. As a result,  $\Omega^{\mathbb{Z}}$  is a compact topological space too. Furthermore, let  $\sigma : \Omega^{\mathbb{Z}} \rightarrow \Omega^{\mathbb{Z}}$  be the (left) shift map,

$$\sigma(\omega) = (\dots, \omega_{-t+1}, \dots, \omega_0, \omega_1^*, \omega_2, \dots, \omega_{t+1}, \dots),$$

where the asterisk marks the zeroth component; component-wise,  $[\sigma(\omega)]_t = \omega_{t+1}$  for all  $t \in \mathbb{Z}$ . The shift map is a homeomorphism of  $\Omega^{\mathbb{Z}}$ .

In addition, the continuous (or topological) dynamical system  $(\Omega^{\mathbb{Z}}, \sigma)$  can be further promoted to a measure-preserving dynamical system by introducing a  $\sigma$ -invariant (probability) measure  $\mu$  on the Borel sigma-algebra  $\mathcal{B}^{\times}$  of  $\Omega^{\mathbb{Z}}$  via the finite-dimensional probability distributions of the given or desired  $\Omega$ -valued stochastic process.<sup>56</sup> For instance, the  $n$ -dimensional marginal probabilities  $\mathbb{P}(\omega_{i_1} \in B_{i_1}, \dots, \omega_{i_n} \in B_{i_n})$  on  $(\Omega^{\mathbb{Z}}, \mathcal{B}^{\times})$  are given by

$$\mu_n(B_{i_1} \times \dots \times B_{i_n}) := \mu(B_{i_1} \times \dots \times B_{i_n} \times \prod_{i \neq i_1, \dots, i_n} \Omega), \quad (36)$$

where  $B_{i_1}, \dots, B_{i_n}$  are Borel sets (e.g., open sets) of  $\Omega$ . The resulting dynamical system  $\Sigma = (\Omega^{\mathbb{Z}}, \mathcal{B}^{\times}, \mu, \sigma)$  is the *shift system* mentioned above. When the measure  $\mu$  consists of finitely many atoms, then  $\{f_{\omega_t}\}$  is an iterated function system.<sup>52</sup> Product measures

$$\mu_n(B_{i_1} \times B_{i_2} \times \dots \times B_{i_n}) = \mu_1(B_{i_1})\mu_1(B_{i_2}) \dots \mu_1(B_{i_n})$$

correspond to independent (memoryless) processes such as coin tossing and white noise.

Following Stark *et al.*,<sup>39</sup> a noisy dynamical system  $X$  is then modeled by the skew product

$$\begin{cases} \omega_{t+1} = [\sigma(\omega)]_t, \\ x_{t+1} = f(\omega_t, x_t) = f_{\omega_t}(x_t), \end{cases} \quad (37)$$

where we suppose that  $f_{\omega_t} = f(\omega_t, \cdot) : \mathcal{M}_X \rightarrow \mathcal{M}_X$  is a diffeomorphism for all  $\omega_t \in \Omega$ . Alternatively,

$$\begin{cases} \omega_{t+1} = [\sigma^t(\omega)]_0, \\ x_{t+1} = f([\sigma^t(\omega)]_0, x_t) = f_{[\sigma^t(\omega)]_0}(x_t), \end{cases} \quad (38)$$

i.e., the parameter of the dynamic at time  $t$  is the 0-component of the shifted sequence  $\sigma^t(\omega)$ .

Due to the formal similarity of Eq. (37) with Eq. (1) for the forced dynamic  $X \rightsquigarrow Y$ , the modeling (37) of a noisy dynamical system is called *stochastic forcing*.<sup>39</sup> Indeed, here we have  $\Sigma \rightsquigarrow X$ , where the shift system  $\Sigma = (\Omega^{\mathbb{Z}}, \mathcal{B}^{\times}, \mu, \sigma)$  is also an autonomous dynamical system and  $X$  is *randomly* forced by  $\Sigma$  since  $x_{t+1} = f(\omega_t, x_t)$ . This parallelism also carries over to the embedding  $E_{f, \varphi_X} : \mathcal{M}_X \rightarrow \mathbb{R}^d$ , Eq. (3), as follows.

Let  $\omega = (\omega_t)_{t \in \mathbb{Z}}$  be a two-sided sequence of points in  $\Omega$  and set

$$f_{\omega_t, \dots, \omega_0}(x) = (f_{\omega_t} \circ \dots \circ f_{\omega_0})(x), \quad (39)$$

for all  $x \in \mathcal{M}_X$ , so  $f_{\omega_t, \dots, \omega_0} : \mathcal{M}_X \rightarrow \mathcal{M}_X$  for all  $t \geq 0$ . For every  $\omega$ , define the map  $E_{f, \varphi_X, \omega} : \mathcal{M}_X \rightarrow \mathbb{R}^d$  as

$$E_{f, \varphi_X, \omega}(x) = (\varphi_X(x), \varphi_X(f_{\omega_0}(x)), \dots, \varphi_X(f_{\omega_{d-2}, \dots, \omega_0}(x))). \quad (40)$$

Note that  $E_{f, \varphi_X, \omega}$  actually depends on the  $d-1$  parameters  $\omega_0, \omega_1, \dots, \omega_{d-2}$ .

**Theorem 14** [Stark *et al.*<sup>39</sup>]. *If  $d \geq 2 \dim_X + 1$ , then there exists a residual set of  $(f, \varphi_X)$  such that for any  $(f, \varphi_X)$  in this set there is an open dense set of sequences  $\omega \in \Omega^{\mathbb{Z}}$  such that the map  $E_{f, \varphi_X, \omega}$  is an embedding.*

Finally, let us point out that Theorem 14 generalizes readily to the case of noisy observations. For example, if the observation function  $\varphi_X(x)$  is replaced by the noisy observation function  $\varphi_{X, \eta}(x) = \varphi_X(x, \eta)$ , where  $\eta \in (\Omega')^{\mathbb{Z}}$ ,  $\Omega'$  is a compact set, and  $\Sigma'$  is the corresponding shift space, then the map

$$E_{f, \varphi_{X, \eta}, \omega, \eta}(x) = (\varphi_{X, \eta_0}(x), \varphi_{X, \eta_1}(f_{\omega_0}(x)), \dots, \varphi_{X, \eta_{d-1}}(f_{\omega_{d-2}, \dots, \omega_0}(x))) \quad (41)$$

is an embedding for generic  $(\omega, \eta) \in \Sigma \times \Sigma'$ . See Stark *et al.*<sup>39</sup> for more detail and other possibilities. Therefore, we may assume hereafter that the observations are noiseless for notational simplicity.

#### IV. COUPLED DYNAMICS AND NOISE

Next, we show that the skew product (37) includes the case of two unidirectionally coupled systems  $X \rightsquigarrow Y$ , where the driver is a noisy dynamical system, namely,

$$\begin{cases} x_{t+1} = f_{\omega_t}(x_t), \\ y_{t+1} = g(x_t, y_t). \end{cases} \quad (42)$$

Equivalently,

$$x_{t+1} = f_{\omega_t, \omega_{t-1}, \dots, \omega_0}(x_0), \quad (43)$$

see Eq. (39), and

$$y_{t+1} = (g_{x_t} \circ g_{x_{t-1}} \circ \dots \circ g_{x_0})(y_0) = g_{x_t, x_{t-1}, \dots, x_0}(y_0) \quad (44)$$

for  $t \geq 0$ , where  $g_x(y) := g(x, y)$ .

**Remark 15.** Of course, if  $\omega_t = \omega_0$  for all times  $t$ , then we recover the noiseless case with  $f := f_{\omega_0}$ ,  $y_0 = g^{(0)}(x_0, y_0)$  and

$$g_{x_{t-1}, \dots, x_0}(y_0) = g_{x_{t-1}, \dots, x_1}(g(x_0, y_0)) = g^{(t)}(x_0, y_0) \quad (45)$$

for  $t \geq 1$ ; see Eqs. (4) and (5).

Some basic facts about the noisy driving dynamic  $x_{t+1} = f(\omega_t, x_t)$  follow.

**Fact 1.** Since the parametric sequence  $\omega = (\omega_t)_{t \in \mathbb{Z}}$  is a trajectory of an  $\Omega$ -valued random process modeled by the shift space  $\Sigma = (\Omega^{\mathbb{Z}}, \mathcal{B}^{\times}, \mu, \sigma)$ , the noisy orbit

$$\begin{aligned} \xi(x, \omega) &= (\dots, x, f_{\omega_0}(x), f_{\omega_1, \omega_0}(x), \dots, f_{\omega_{t-1}, \dots, \omega_0}(x), \dots) \\ &= (\dots, x_0, x_1, x_2, \dots, x_t, \dots) \in \mathcal{M}_X^{\mathbb{Z}} \end{aligned} \quad (46)$$

is a trajectory of an  $\mathcal{M}_X$ -valued random process. In general, i.i.d. parametric sequences  $\omega$  (commonly used in applications) do not



generate i.i.d. noisy orbits  $\xi = \xi(x, \omega)$ . Additive noise  $x_{t+1} = f(x_t) + \omega_t$  is a plain example when the invariant measure of  $f: \mathcal{M}_X \rightarrow \mathcal{M}_X$  is not uniform.

**Fact 2.** Since the  $\omega_t$ 's are the outcomes of a stationary process, the  $x_t$ 's are also the outcomes of a stationary process. Indeed, the definition  $x_{t+1} = f(\omega_t, x_t)$  is time-invariant due to the stationarity in the generation of the  $\omega_t$ 's.

**Fact 3.** Under additional assumptions,  $\xi(x, \omega) = (x_t)_{t \in \mathbb{Z}}$  can match any arbitrary stationary sequence in  $\mathcal{M}_X$  (i.e., a trajectory of a stationary  $\mathcal{M}_X$ -valued random process) by fine tuning the sequence  $\omega$ . For example, assume the following mild proviso.

**Condition 16.**  $f(\cdot, x) = f_x: \mathcal{M}_p \rightarrow \mathcal{M}_X$  is an embedding for each  $x \in \mathcal{M}_X$ .

Under this condition, given  $x_t$ , the relationship between  $\omega_t$  and  $x_{t+1}$  is one-to-one for each  $t$ , which implies that the equation  $f_{x_t}(\omega_t) := f(\omega_t, x_t) = x_{t+1}$  can be solved for  $\omega_t$  in a unique way. Therefore, the noisy orbit  $\xi(x, \omega)$  can be recursively transformed into any stationary sequence  $\eta \in \mathcal{M}_X^{\mathbb{Z}}$  by choosing  $x_0 = \eta_0$  and  $\omega_t$  as the unique solution of  $f(\omega_t, x_t) = \eta_{t+1}$  for  $t = 0, 1, 2, \dots$  and  $t = -1, -2, \dots$ . Henceforth, we assume that Condition 16 is met.

**Fact 4.** In particular, by Condition 16, the relationship between  $\xi = \{x_t\}_{t \in \mathbb{Z}}$  and  $\{x_0, \omega\}$  is one-to-one, i.e., the function  $(x, \omega) \mapsto \xi$  is invertible, where  $\xi_0 = x$ . Therefore, we may indistinctly talk of  $x$  and  $\omega$ , or  $\xi = (x_t)_{t \in \mathbb{Z}}$ . In practice, one chooses  $\omega$  so that the noisy orbit  $\xi$  of  $x$  deviates from the noiseless orbit by small perturbations.

By Fact 2, we can view the noisy dynamic (42) as stochastic forcing, the compact manifold  $\mathcal{M}_X$  being the parameter set and the orbits  $\xi = (x_t)_{t \in \mathbb{Z}}$  of the noisy driver playing the role of the parameter sequences  $\omega = (\omega_t)_{t \in \mathbb{Z}}$ . This being the case, replace in Theorem 14(i) the sequence  $\omega \in \Omega^{\mathbb{Z}}$  with the noisy orbit  $\xi = (x_t)_{t \in \mathbb{Z}}$ , (ii) the map  $f: \Omega \times \mathcal{M}_X \rightarrow \mathcal{M}_X$  with  $g: \mathcal{M}_X \times \mathcal{M}_Y \rightarrow \mathcal{M}_Y$ , (iii)  $f_{\omega_t}(x_t) = f(\omega_t, x_t)$  with  $g_{x_t}(y_t) = g(x_t, y_t)$ , (iv)  $f_{\omega_t, \omega_{t-1}, \dots, \omega_0}$  with  $g_{x_t, x_{t-1}, \dots, x_0}$ , and (v)  $\varphi_X$  with  $\varphi_Y$ , to derive the following result.

**Theorem 17.** *If  $\delta \geq 2 \dim_Y + 1$ , then there exists a residual set of  $(g, \varphi_Y)$  such that for any  $(g, \varphi_Y)$  in this set there is an open dense set of sequences  $\xi = (x_t)_{t \in \mathbb{Z}} \in \mathcal{M}_X^{\mathbb{Z}}$  such that the map  $E_{g, \varphi_Y, \xi}: \mathcal{M}_Y \rightarrow \mathbb{R}^\delta$  defined by*

$$E_{g, \varphi_Y, \xi}(y) = (\varphi_Y(y), \varphi_Y(g_{x_0}(y)), \dots, \varphi_Y(g_{x_{\delta-2}, \dots, x_0}(y))) \quad (47)$$

is an embedding.

According to Eq. (47),  $E_{g, \varphi_Y, \xi}: \mathcal{M}_Y \rightarrow \mathbb{R}^\delta$  depends actually on the  $\delta - 1$  parameters  $x_0, x_1, \dots, x_{\delta-2}$ . The points  $(x_0, \dots, x_{\delta-2})$  are dense in the finite-dimensional manifold  $\mathcal{M}_X^{\delta-1}$  if and only if the points  $x_k$  are dense in  $\mathcal{M}_X$  for each  $0 \leq k \leq \delta - 2$ . It follows that  $E_{g, \varphi_Y, \xi}$  is an embedding for a residual set of  $(g, \varphi_Y)$  and dense sets of points  $\{x_0, \dots, x_{\delta-2}\}$  in  $\mathcal{M}_X$ .

**Remark 18.** One can extend the map  $E_{g, \varphi_Y, \xi}$  from  $\mathcal{M}_Y$  to  $\mathcal{M}_X \times \mathcal{M}_Y$  by defining  $E_{g, \varphi_Y, x_1, \dots, x_{\delta-2}}: \{x_0\} \times \mathcal{M}_Y \rightarrow \mathbb{R}^\delta$  as

$$E_{g, \varphi_Y, x_1, \dots, x_{\delta-2}}(x_0, y) := E_{g, \varphi_Y, \xi}(y). \quad (48)$$

Yet,  $E_{g, \varphi_Y, x_1, \dots, x_{\delta-2}}$  does not allow to reconstruct the full state space  $\mathcal{M}_X \times \mathcal{M}_Y$  because, according to Theorem 17, in general, this map is an embedding only for a dense set of points  $(x_0, y) \in \mathcal{M}_X \times \mathcal{M}_Y$ . Nevertheless, this result can be useful in applications where one can assume  $x_0$  to be fixed and generic, like in time series analysis.

## V. THE NOISY SCENARIO

In this section, we discuss some changes and limitations introduced by noise in the conventional framework of Sec. II. Since the driver dynamic now explicitly depends on time through the noise, so do the main concepts like state reconstruction, cross map, and synchronization map.

### A. State reconstruction

Let  $\omega \in \Omega^{\mathbb{Z}}$  be a parametric sequence and suppose  $d \geq 2 \dim_X + 1$ . Then, according to Theorem 14, the map  $E_{f, \varphi_X, \omega}: \mathcal{M}_X \rightarrow \mathbb{R}^d$  defined in Eq. (40) is generically an embedding. Similar to the noiseless case, we define the manifolds

$$\mathcal{N}_{X, \omega} = E_{f, \varphi_X, \omega}(\mathcal{M}_X) \subset \mathbb{R}^d \quad (49)$$

(each one diffeomorphic to  $\mathcal{M}_X$ ), and the noisy time-delay vectors

$$\begin{aligned} \mathbf{x}_t &= E_{f, \varphi_X, \sigma^t(\omega)}(x_t) \\ &= (\varphi_X(x_t), \varphi_X(x_{t+1}), \dots, \varphi_X(x_{t+d-1})) \in \mathcal{N}_{X, \sigma^t(\omega)}, \end{aligned} \quad (50)$$

with  $x_{t+k} = f_{\omega_{t+k-1}, \dots, \omega_t}(x_t)$  for  $k \geq 1$ . Then, the driver dynamics  $x_{t+1} = f_{\omega_t}(x_t)$  translate to

$$\mathbf{x}_{t+1} = F_{\sigma^t(\omega)}(\mathbf{x}_t) \quad (51)$$

in the reconstructed state spaces, where the map  $F_\omega: \mathcal{N}_{X, \omega} \rightarrow \mathcal{N}_{X, \sigma(\omega)}$  defined as

$$F_\omega = E_{f, \varphi_X, \sigma(\omega)} \circ f_{\omega_0} \circ E_{f, \varphi_X, \omega}^{-1} \quad (52)$$

is a diffeomorphism, provided that  $E_{f, \varphi_X, \omega}$  and  $E_{f, \varphi_X, \sigma(\omega)}$  are embeddings. At variance with the noiseless case, the reconstructed dynamic  $\mathbf{x}_t \mapsto \mathbf{x}_{t+1}$  hops from a diffeomorphic copy  $\mathcal{N}_{X, \sigma^t(\omega)}$  of  $\mathcal{M}_X$  to another diffeomorphic copy  $\mathcal{N}_{X, \sigma^{t+1}(\omega)}$ .

Likewise, let  $\xi = \xi(x, \omega) \in \mathcal{M}_X^{\mathbb{Z}}$  be a noisy orbit of  $x = x_0$  [Eq. (46)] and suppose  $\delta \geq 2 \dim_Y + 1$ . Then, according to Theorem 17, the map  $E_{g, \varphi_Y, \xi}: \mathcal{M}_Y \rightarrow \mathbb{R}^\delta$  defined in Eq. (47) is generically an embedding. Define the manifolds

$$\mathcal{N}_{Y, \xi} = E_{g, \varphi_Y, \xi}(\mathcal{M}_Y) \subset \mathbb{R}^\delta \quad (53)$$

(each one diffeomorphic to  $\mathcal{M}_Y$ ) and the noisy time-delay vectors

$$\begin{aligned} \mathbf{y}_t &= E_{g, \varphi_Y, \sigma^t(\xi)}(y_t) \\ &= (\varphi_Y(y_t), \varphi_Y(y_{t+1}), \dots, \varphi_Y(y_{t+\delta-1})) \in \mathcal{N}_{Y, \sigma^t(\xi)}, \end{aligned} \quad (54)$$

with  $y_{t+1} = g(x_t, y_t) =: g_t(y_t)$  and

$$y_{t+k} = g(x_{t+k-1}, g_{x_{t+k-2}, \dots, x_{t+1}, x_t}(y_t)) =: g_{x_{t+k-1}, \dots, x_{t+1}, x_t}(y_t), \quad (55)$$

for  $k \geq 2$ . Then, similar to (52), the map  $G_\xi: \mathcal{N}_{Y, \xi} \rightarrow \mathcal{N}_{Y, \sigma(\xi)}$  defined as

$$G_\xi = E_{g, \varphi_Y, \sigma(\xi)} \circ g_{x_0} \circ E_{g, \varphi_Y, \xi}^{-1} \quad (56)$$

is a diffeomorphism, provided that  $E_{g, \varphi_Y, \xi}$  and  $E_{g, \varphi_Y, \sigma(\xi)}$  are embeddings, and it holds

$$\mathbf{y}_{t+1} = G_{\sigma^t(\xi)}(\mathbf{y}_t). \quad (57)$$

Again, the range of  $G_{\sigma^t(\xi)}$  depends on  $t$  through  $\sigma^t(\xi)$ , but all of them are diffeomorphic copies of  $\mathcal{M}_Y$ .

**B. Cross map**

The definition of the cross map of the systems  $X \rightsquigarrow Y$  in (12) hinges on the reconstruction of both the driver state space  $\mathcal{M}_X$  and the full state space  $\mathcal{M}_X \times \mathcal{M}_Y$ . However, according to Remark 18, the latter reconstruction is generally only possible in the noisy scenario for a dense set of driver states.

This being the case, we are going to define the cross map  $\mathbf{x}_t = \Phi_{\sigma^t(\omega)}(\mathbf{y}_t)$  for two time series  $(\mathbf{x}_t)_{t \geq 0}$  and  $(\mathbf{y}_t)_{t \geq 0}$  of time-delay vectors obtained from a noisy orbit  $\xi = \xi(x_0, \omega)$  of the driver  $X$  [Eq. (46)] and the corresponding response from the system  $Y$ , respectively. This limited approach suffices for the needs of time series analysis, where the focus in practice is on (finite segments of) single orbits rather than on manifolds, and points and parameters can be considered generic. For ease of notation, we will write

$$E_{g,\varphi_Y,\sigma^t(\omega)}(\mathbf{x}_t, \mathbf{y}_t) = E_{g,\varphi_Y,\mathbf{x}_{t+1}, \dots, \mathbf{x}_{t+\delta-2}}(\mathbf{x}_t, \mathbf{y}_t) = \mathbf{y}_t, \tag{58}$$

see Eq. (48), since the relation  $\sigma^t(\xi) \leftrightarrow (\mathbf{x}_t, \sigma^t(\omega))$  is one-to-one by Condition 16 in Sec. IV.

To define the cross map in the presence of dynamical noise,  $\mathbf{x}_t = \Phi_{\sigma^t(\omega)}(\mathbf{y}_t)$ , we mimic the definition of the cross map in Eq. (11) in the form

$$\begin{array}{ccc} \mathcal{M}_X \times \mathcal{M}_Y \ni & (\mathbf{x}_t, \mathbf{y}_t) & \xrightarrow{\Pi_X} \mathbf{x}_t \in \mathcal{M}_X \\ E_{g,\varphi_Y,\sigma^t(\omega)}^{-1} & \uparrow & \downarrow E_{f,\varphi_X,\sigma^t(\omega)} \\ \mathcal{N}_{Y,\sigma^t(\xi)} \ni & \mathbf{y}_t & \mathbf{x}_t \in \mathcal{N}_{X,\sigma^t(\omega)}, \end{array} \tag{59}$$

under the assumption that  $E_{g,\varphi_Y,\sigma^t(\omega)}(\mathbf{x}_t, \mathbf{y}_t)$  is an embedding for the considered states  $\mathbf{x}_t \in \mathcal{M}_X$ . Hence,

$$\mathbf{x}_t = \Phi_{\sigma^t(\omega)}(\mathbf{y}_t) := \left( E_{f,\varphi_X,\sigma^t(\omega)} \circ \Pi_X \circ E_{g,\varphi_Y,\sigma^t(\omega)}^{-1} \right) (\mathbf{y}_t). \tag{60}$$

Let us check that  $\Phi_{\sigma^t(\omega)}(\mathbf{y}_t)$  becomes  $\Phi(\mathbf{y}_t)$ , Eq. (12), when the noise is switched off in Eq. (60), i.e., when  $\omega = \bar{\omega}$  with  $\bar{\omega}_t = \omega_0$  for all  $t \in \mathbb{Z}$ . We suppose that the maps  $E_{f,\varphi_X,\sigma^t(\omega)}$  and  $E_{g,\varphi_Y,\sigma^t(\omega)}$  are embeddings for  $\omega = \bar{\omega}$ .

In that case,  $E_{f,\varphi_X,\sigma^t(\bar{\omega})}(\mathbf{x}_t) = E_{f,\varphi_X}(\mathbf{x}_t)$  with  $f := f_{\omega_0}$ ; see Eqs. (50) and (3). Similarly, by Eqs. (45) and (6) with  $D = \delta - 1$ , and setting  $\bar{\xi} = (x_0, \bar{\omega}) = (f^t(x_0))_{t \in \mathbb{Z}}$ ,

$$\begin{aligned} E_{g,\varphi_Y,\sigma^t(\bar{\omega})}(\mathbf{x}_t, \mathbf{y}_t) &= E_{g,\varphi_Y,\sigma^t(\bar{\xi})}(\mathbf{y}_t) \\ &= (\varphi_Y(\mathbf{y}_t), \varphi_Y(g_{x_t}(\mathbf{y}_t)), \dots, \varphi_Y(g_{x_{t+\delta-2}, \dots, x_t}(\mathbf{y}_t))) \\ &= (\varphi_Y(g^{(0)}(\mathbf{x}_t, \mathbf{y}_t)), \varphi_Y(g^{(1)}(\mathbf{x}_t, \mathbf{y}_t)), \dots, \varphi_Y(g^{(\delta-1)}(\mathbf{x}_t, \mathbf{y}_t))) \\ &= E_{f,g,\varphi_Y}(\mathbf{x}_t, \mathbf{y}_t). \end{aligned} \tag{61}$$

Comparison with Eq. (12) shows that  $\Phi_{\sigma^t(\bar{\omega})}(\mathbf{y}_t) = \Phi(\mathbf{y}_t)$ , as it should.

**C. Synchronization map**

Generalized synchronization (18) can be extended from the noiseless dynamic  $[f, g]$  to the noisy dynamic  $[f_\omega, g]$ ,  $\omega \in \Omega^\mathbb{Z}$ , where the dynamic changes at every time step, by requiring

$$y_t = h_{\sigma^t(\omega)}(\mathbf{x}_t). \tag{62}$$

**Definition 19.** We say that the responder  $Y$  is synchronized to a driver  $X$  perturbed by the noise  $\omega \in \Omega^\mathbb{Z}$ , if there is a sequence of continuous maps  $h_{\sigma^t(\omega)} : \mathcal{M}_X \rightarrow \mathcal{M}_Y$  such that Eq. (62) holds for all  $t \in \mathbb{Z}$ .

More generally, the synchronization map of order  $k \geq 0$ ,  $h^{(k)} = h \circ f^k$  [Eq. (20)], generalizes to  $h_{\sigma^t(\omega)}^{(0)} := h_{\sigma^t(\omega)}$  and

$$y_t = h_{\sigma^t(\omega)}^{(k)}(\mathbf{x}_{t-k}) := h_{\sigma^t(\omega)} \circ f_{\omega_{t-1}, \dots, \omega_{t-k}}(\mathbf{x}_{t-k}), \tag{63}$$

for  $k \geq 1$  in the noisy case, while the synchronization map of period  $K \geq 1$  (21) generalizes to

$$y_t = \frac{1}{K} \sum_{k=0}^{K-1} h_{\sigma^t(\omega)}^{(k)}(\mathbf{x}_{t-k}) =: h_{K,\sigma^t(\omega)}(\mathbf{x}_t, \dots, \mathbf{x}_{t-K+1}). \tag{64}$$

To define the synchronization map in the reconstructed spaces, we replace  $\mathbf{y}_t = \Phi^{-1}(\mathbf{x}_t)$  with  $\mathbf{y}_t = H_{\sigma^t(\omega)}(\mathbf{x}_t)$  in the “noisy” version of diagram (27),

$$\begin{array}{ccc} \mathcal{M}_X \ni & \mathbf{x}_t & \xrightarrow{h_{\sigma^t(\omega)}} \mathbf{y}_t \in \mathcal{M}_Y \\ E_{f,\varphi_X,\sigma^t(\omega)}^{-1} & \uparrow & \uparrow E_{g,\varphi_Y,\sigma^t(\xi)}^{-1} \\ \mathcal{N}_{X,\sigma^t(\omega)} \ni & \mathbf{x}_t & \xrightarrow{H_{\sigma^t(\omega)}} \mathbf{y}_t \in \mathcal{N}_{Y,\sigma^t(\xi)} \end{array} . \tag{65}$$

Here, we used Eqs. (50) and (54) and assume that the maps  $E_{f,\varphi_X,\sigma^t(\omega)}$  and  $E_{g,\varphi_Y,\sigma^t(\xi)}$  are embeddings. Then, it follows from (65)

$$\begin{aligned} \mathbf{y}_t &= (E_{g,\varphi_Y,\sigma^t(\xi)} \circ h_{\sigma^t(\omega)} \circ E_{f,\varphi_X,\sigma^t(\omega)}^{-1})(\mathbf{x}_t) \\ &=: H_{\sigma^t(\omega)}(\mathbf{x}_t). \end{aligned} \tag{66}$$

Furthermore, by Eq. (51),

$$\begin{aligned} \mathbf{x}_t &= F_{\sigma^{t-1}(\omega)}(\mathbf{x}_{t-1}) \\ &= \dots = (F_{\sigma^{t-1}(\omega)} \circ \dots \circ F_{\sigma^{t-k}(\omega)})(\mathbf{x}_{t-k}) = F_{\sigma^{t-1}(\omega)}^{(k)}(\mathbf{x}_{t-k}), \end{aligned} \tag{67}$$

so that

$$\begin{aligned} \mathbf{y}_t &= \frac{1}{K} \sum_{k=0}^{K-1} \left( H_{\sigma^t(\omega)} \circ F_{\sigma^{t-1}(\omega)}^{(k)} \right) (\mathbf{x}_{t-k}) \\ &=: H_{K,\sigma^t(\omega)}(\mathbf{x}_t, \dots, \mathbf{x}_{t-K+1}) \end{aligned} \tag{68}$$

generalizes the reconstructed synchronization map of period  $K$ , Eq. (32), to the noisy case.

To check that  $\mathbf{y}_t = H_{\sigma^t(\omega)}(\mathbf{x}_t)$  becomes  $\mathbf{y}_t = \Phi^{-1}(\mathbf{x}_t)$  when the noise is switched off in Eq. (66), replace

$$E_{g,\varphi_Y,\sigma^t(\xi)}^{-1} : \mathbf{y}_t \longrightarrow \mathbf{y}_t \tag{69}$$

on the right column of diagram (65) with

$$\Pi_Y \circ E_{g,\varphi_Y,\sigma^t(\omega)}^{-1} : \mathbf{y}_t \xrightarrow{E_{g,\varphi_Y,\sigma^t(\omega)}^{-1}} (\mathbf{x}_t, \mathbf{y}_t) \xrightarrow{\Pi_Y} \mathbf{y}_t \tag{70}$$

so that, according to Eq. (61),  $\Pi_Y \circ E_{g,\varphi_Y,\sigma^t(\omega)}^{-1}(\mathbf{y}_t)$  becomes  $\Pi_Y \circ E_{f,g,\varphi_Y}^{-1}(\mathbf{y}_t)$  in the noiseless case  $\omega = \bar{\omega}$ , i.e.,  $\bar{\omega}_t = \omega_0$  for all  $t \in \mathbb{Z}$ . Finally, set  $h(\mathbf{x}_t) = h_{\sigma^t(\bar{\omega})}$  to convert diagram (65) to diagram (27), thus identifying  $H_{\sigma^t(\bar{\omega})}(\mathbf{x}_t)$  with  $\Phi^{-1}(\mathbf{x}_t)$ , as it should be.

The numerical simulations of Sec. VI show that synchronization is robust against dynamical noise for strong enough couplings

and, hence, can occur in the presence of dynamical noise. On the other hand, if synchronization occurs in the presence of dynamical noise but disappears when noise is switched off, then one speaks of noise-induced synchronization.<sup>57</sup>

Finally, we can generalize the concepts of asymptotic synchronization and stability of the responder in the presence of noise in the driver as follows. We say that  $Y$  is asymptotically synchronized to the noisy driver  $X$  if the definition of synchronization, Eq. (66), holds only asymptotically, i.e.,

$$\lim_{t \rightarrow \infty} \|y_t - H_{\sigma^t(\omega)}(\mathbf{x}_t)\| = 0, \tag{71}$$

where  $\|\cdot\|$  is a norm in  $\mathbb{R}^{\dim Y}$ . It follows then that  $Y$  is *asymptotically stable*, i.e., the orbits of  $Y$  converge to  $H_{\sigma^t(\omega)}(\mathbf{x}_t)$ , regardless of their initial conditions. Asymptotic stability can easily be checked in practice. As in the noiseless case, it is a handy method to rule out synchronization.

## VI. NUMERICAL SIMULATIONS

Unlike identical synchronization, which can be easily visualized, generalized synchronization is more difficult to detect. As mentioned in the Introduction, there exists an extensive literature on methods to detect functional dependency (and generalized synchronization for that matter) between two time series. The functional dependency targeted in this section is the synchronization map of a certain period  $K > 1$  given in Eqs. (32) and (68) for noiseless and noisy drivers, respectively. For this reason, we use recurrent neural networks of the type *long short-term memory* (LSTM), which excel at predicting data from time series and are robust to noise. In fact, the LSTM nets outperformed the perceptrons ( $K = 1$ ) in the numerical simulations below, so we will only report the results obtained with the former. As a benchmark we use nearest-neighbor cross prediction (Sec. II B) because it is based on the continuity of the cross map (and its inverse in case of synchronization). In addition, nearest-neighbor cross prediction is robust against noise, particularly if the neighborhoods are well populated.

### A. Models

For the numerical simulations we chose two unidirectionally coupled Hénon maps with several structural parameters and varying coupling strength. This testbed, first proposed by Schiff *et al.*<sup>17</sup> and studied with the normalized mutual error, has been revisited several times in the literature, e.g., in Quian Quiroga *et al.*,<sup>24</sup> where the authors use the conditional Lyapunov exponent and the so-called nonlinear interdependencies.<sup>23</sup>

Thus, the equations of the driver  $X$ , with states  $x = (x^{(1)}, x^{(2)})$  in a trapping region of the attractor, are

$$\begin{cases} x_{t+1}^{(1)} = 1.4 - (x_t^{(1)})^2 + (b_1 + \omega_t)x_t^{(2)}, \\ x_{t+1}^{(2)} = x_t^{(1)}, \end{cases} \tag{72}$$

where  $b_1$  is a constant and  $\omega_t$  are i.i.d. random numbers in the interval  $[-A, A]$ , the noiseless scenario corresponding to  $A = 0$ . The observation function is  $\varphi_X(x_t) = x_t^{(1)}$ , i.e., the projection on the first component.

The equations of the responder  $Y$ , with states  $y = (y^{(1)}, y^{(2)})$ , are

$$\begin{cases} y_{t+1}^{(1)} = 1.4 - [Cx_t^{(1)}y_t^{(1)} + (1 - C)(y_t^{(1)})^2] + b_2y_t^{(2)}, \\ y_{t+1}^{(2)} = y_t^{(1)}, \end{cases} \tag{73}$$

where  $b_2$  is a constant and  $C$  is the *coupling strength*. For  $C = 0$ , systems  $X$  and  $Y$  are uncoupled. The observation function is again the projection on the first component,  $\varphi_Y(y_t) = y_t^{(1)}$ .

The parameter settings are as follows.

- The settings for the constants  $b_1$  and  $b_2$  are the same as in Schiff *et al.*<sup>17</sup> and Quian Quiroga *et al.*<sup>24</sup> So, we first set  $b_1 = b_2 = 0.3$ , the standard values of the Hénon map, to study the coupling of identical systems (*Model Hénon 0.3-0.3*), which allows identical synchronization (i.e.,  $y_t = x_t$ ) for  $C = 1$ . Then, to study the coupling of non-identical systems, we set  $b_1 = 0.3, b_2 = 0.1$  (*Model Hénon 0.3-0.1*) and  $b_1 = 0.1, b_2 = 0.3$  (*Model Hénon 0.1-0.3*).
- For the previous choices of  $b_1$  and  $b_2$ , we found that the driver orbits can diverge for noise amplitudes  $A > 0.013$ , so we restrict them to the interval  $0 \leq A \leq 0.013$ . The amplitudes used in the figures below are  $A = 0$  (noiseless driver), 0.005 and 0.013.
- The range of the *coupling strength*  $C$  is  $0 \leq C \leq 1.2$ ; the increment of  $C$  in the figures below is  $\Delta C = 0.05$ .
- For each case described above (identical/non-identical systems, noiseless/noisy driver), one series  $(x_t)_{0 \leq t \leq T-1}$  and one series  $(y_t)_{0 \leq t \leq T-1}$  were generated with seeds  $x_0 = (0, 0.9)$  and  $y_0 = (0.75, 0)$ , and length  $T = 10^5$  (after discarding the first 1000 points). Since we are only interested in synchronization, one series per case suffices because of asymptotic stability (Sec. II C).
- The embedding dimension in the noiseless and noisy scenarios is  $d = 5$ , i.e.,  $\mathbf{x}_t = (x_t^{(1)}, \dots, x_{t+4}^{(1)})$  and  $\mathbf{y}_t = (y_t^{(1)}, \dots, y_{t+4}^{(1)})$ ,  $0 \leq t \leq T - 5$ . *A posteriori* justification for this choice are the excellent results obtained in the benchmark below.

The methods to test for synchronization in the *noiseless case* ( $A = 0$ ) and *noisy cases* ( $A = 0.005, 0.013$ ) are the following.

**Method 1** Our first method unveils synchronization by detecting functional dependencies, namely, the existence of the synchronization map of period  $K = 10$  for time-delay vectors, i.e.,

$$\mathbf{y}_t = H_{10}(\mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_{t-9}) \tag{74}$$

[see Eq. (32)]. To do this, we used a three-layer neural network to predict  $\mathbf{y}_t$  based on  $\mathbf{x}_t, \dots, \mathbf{x}_{t-9}$ . Specifically, (i) the input layer consisted of an LSTM net with 5 units, hidden states of dimension 10 (corresponding to the inputs  $\mathbf{x}_t, \dots, \mathbf{x}_{t-9}$ ) and the activation function  $\text{ReLU}(x) = \max\{0, x\}$ ; (ii) the intermediate layer had 25 neurons and the activation function  $\text{Sigmoid}(x) = 1/(1 + e^{-x})$ ; and (iii) the output layer had 5 neurons. Hence, the output layer returns 5 states, corresponding to the 5 components of  $\hat{\mathbf{y}}_t$ , the prediction of  $\mathbf{y}_t$ . The network was trained with an 80% of the data (the first 80 000 time-delay vectors) and stochastic gradient descent, while the remaining 20% of the data was used for testing. The accuracy of the predictions  $\hat{\mathbf{y}}_t$  output by the neural network based on the *testing* data  $\mathbf{x}_t, \dots, \mathbf{x}_{t-9}$  (i.e., for each  $t =$

80 000, . . . , 99 990) was measured by their mean squared error (MSE), mean absolute error (MAE), and mean absolute scaled error (MASE). The matching of these three metrics in both the training and testing phases discards overfitting. Furthermore, since predictions based on data patterns are robust against low levels of noise, we expect this method to work well in both the noiseless and noisy cases.

**Method 2** As a benchmark we used nearest-neighbor cross prediction which, in the noiseless case, estimates  $\mathbf{x}_t$  based on the continuity of the cross map  $\mathbf{x}_t = \Phi(\mathbf{y}_t)$  and corresponding nearest neighbors.<sup>18</sup> Following the convergent cross mapping (CCM) method, we measured the accuracy of those estimations by  $r(\mathbf{x}, \hat{\mathbf{x}})$ , the Pearson correlation coefficient of the estimates  $\hat{\mathbf{x}}_t$  obtained with the  $d + 1 = 6$  nearest neighbors of  $\mathbf{y}_t$ . Since  $T = 10^5$ , the time series are sufficiently long to obtain good estimates (actually only the first 10 000 points were used), so  $r(\mathbf{x}, \hat{\mathbf{x}}) \simeq 1$ . On the contrary, if  $r(\mathbf{y}, \hat{\mathbf{y}})$  is the Pearson correlation coefficient of the estimates  $\hat{\mathbf{y}}_t$  obtained via the  $d + 1 = 6$  nearest-neighbors of  $\mathbf{x}_t$ , then we expect  $r(\mathbf{y}, \hat{\mathbf{y}}) \simeq 0$ , unless  $Y$  synchronizes with  $X$ , in which case  $\mathbf{y} = \Phi^{-1}(\mathbf{x})$  and

$r(\mathbf{y}, \hat{\mathbf{y}}) \simeq 1$  (this time due to the continuity of  $\Phi^{-1}$ ). We conclude that if

$$\Delta r = r(\mathbf{x}, \hat{\mathbf{x}}) - r(\mathbf{y}, \hat{\mathbf{y}}) \tag{75}$$

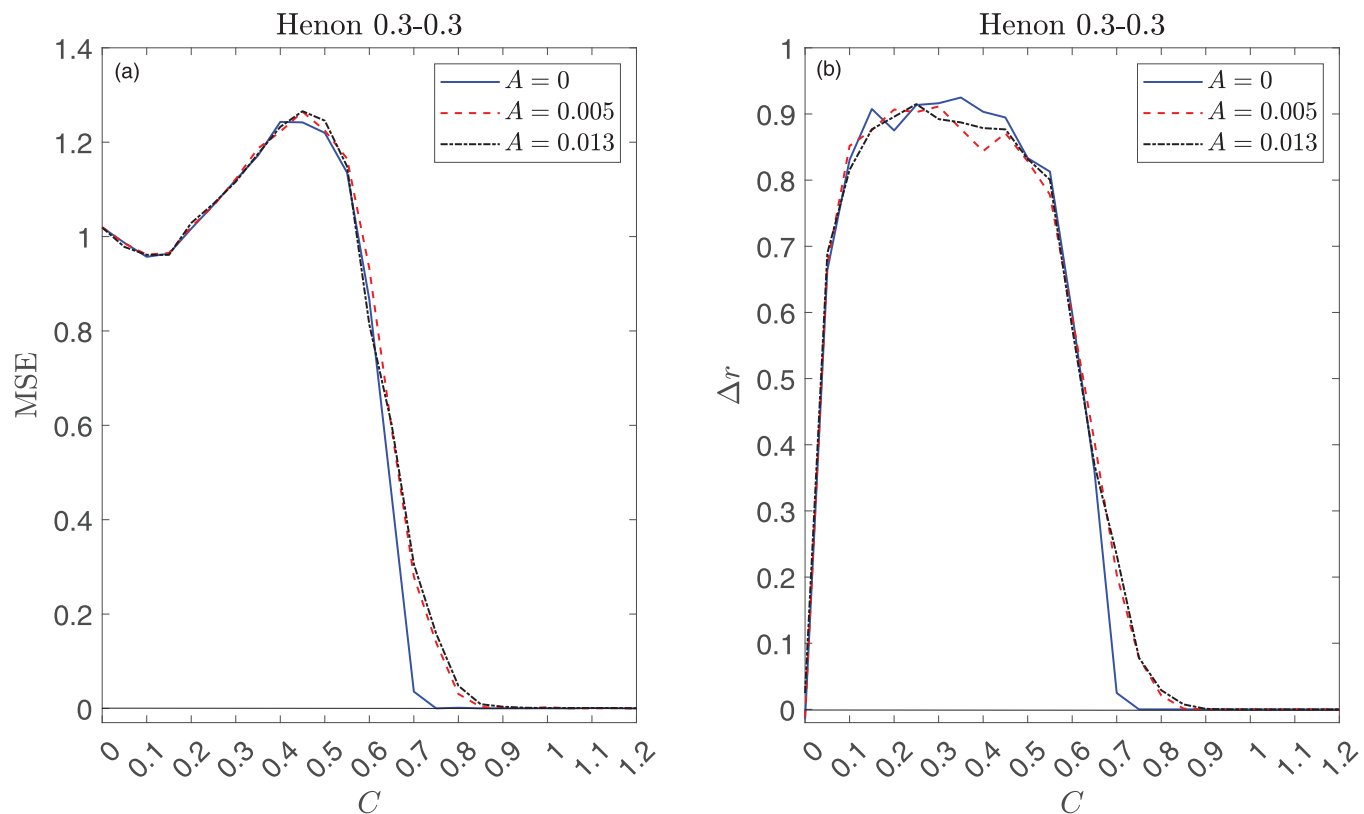
and  $X \rightsquigarrow Y$ , then

- (i)  $0 \leq \Delta r \leq 1$ , and
- (ii)  $\Delta r \simeq 0$  signalizes synchronization, except when  $r(\mathbf{x}, \hat{\mathbf{x}}) = 0 = r(\mathbf{y}, \hat{\mathbf{y}})$ , i.e., when  $X$  and  $Y$  are uncoupled.

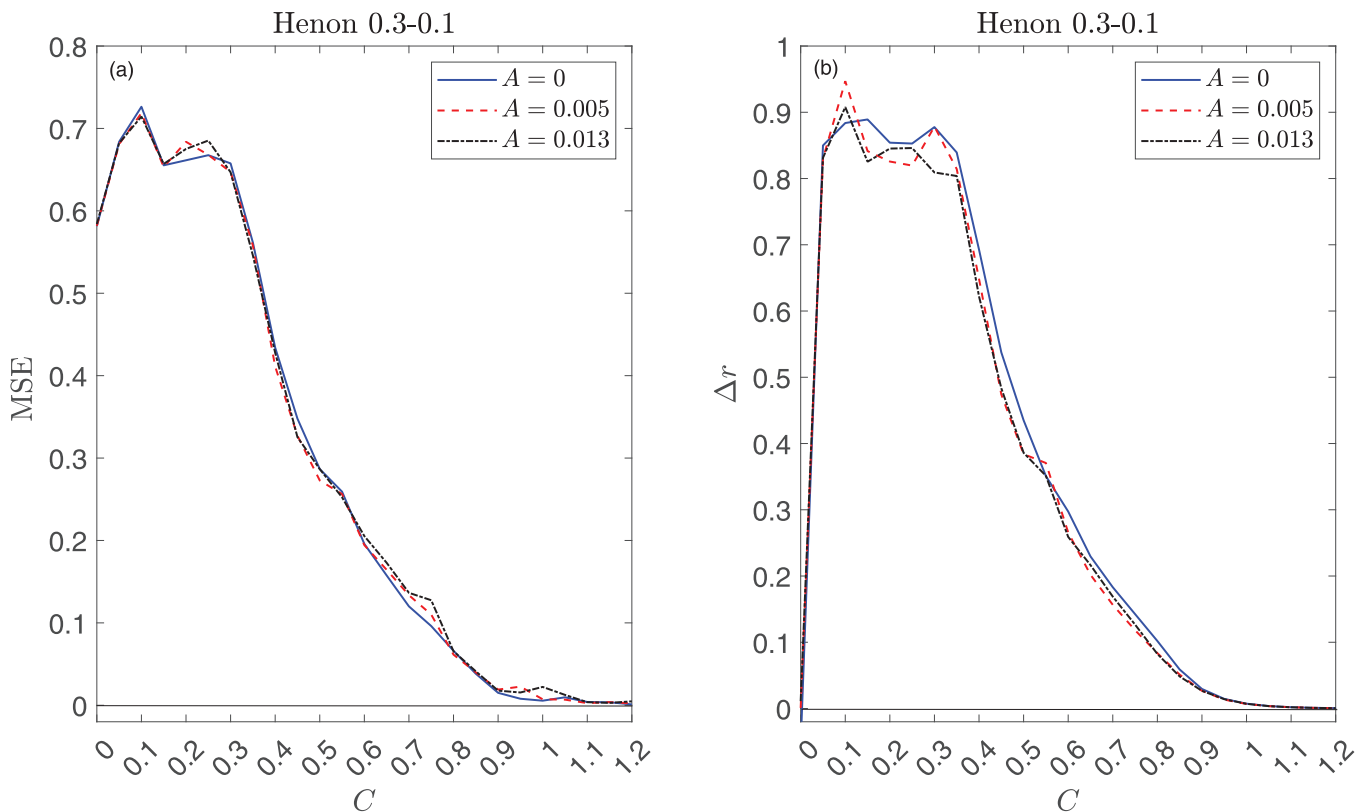
In the noisy cases, the situation is qualitatively the same, thanks to the robustness of nearest-neighbor cross prediction against low levels of noise. See, e.g., Sugihara *et al.*,<sup>35</sup> Mønster *et al.*,<sup>38</sup> and the book by Datsoris and Parlitz<sup>59</sup> for CCM algorithms to compute (75).

### B. Results

Out of the accuracy results obtained with the LSTM network and testing data, we are going to discuss only the MSE vs  $C$  curves since the other two curves, MAE and MASE vs  $C$ , are similar for all models.



**FIG. 1.** Numerical results for the model Hénon 0.3–0.3, i.e.,  $b_1 = 0.3$  in (72) and  $b_2 = 0.3$  in (73). (a) MSE vs the coupling strength  $C$  for a noiseless driver (noise amplitude  $A = 0$ ) and a noisy driver ( $A = 0.005, 0.013$ ) obtained with an LSTM net. (b)  $\Delta r$  vs  $C$  for a noiseless driver ( $A = 0$ ) and a noisy driver ( $A = 0.005, 0.013$ ) obtained via 6-nearest-neighbor cross prediction. See text for more detail.



**FIG. 2.** Numerical results for the model Hénon 0.3–0.1, i.e.,  $b_1 = 0.3$  in (72) and  $b_2 = 0.1$  in (73). The information displayed in the panels (a) and (b) is the same as in Fig. 1.

The results of the numerical simulations are depicted in Figs. 1–3 for Method 1 [panels (a)] and Method 2 [panels (b)] and the three models Hénon 0.3–0.3 (Fig. 1), 0.3–0.1 (Fig. 2), and 0.1–0.3 (Fig. 3). Comparison of both panels for each case and  $C > 0$  shows an excellent agreement of both methods on the synchronization states, i.e.,  $MSE(C) = 0$  in panels (a) and  $\Delta r(C) = 0$  in panels (b). As noted above,  $\Delta r(0) = 0$  in all cases owing to the fact that  $X$  and  $Y$  are uncoupled for  $C = 0$ ; such numerical artifacts can be easily filtered out by checking whether  $r(\mathbf{x}, \hat{\mathbf{x}}) \simeq 0$  and  $r(\mathbf{y}, \hat{\mathbf{y}}) \simeq 0$ .

The main conclusions from the numerical results can be summarized as follows.

- Small-amplitude noise does not destroy all the states of “strong” synchronization (i.e., due to strong enough couplings), it only shifts the synchronization threshold to higher values. So, synchronization can also occur in the presence of dynamical noise.
- Synchronization due to strong enough couplings is robust against small-amplitude dynamical noise, while synchronization states with a weak coupling strength can be unstable whatever the amplitude of the noise. This fact is illustrated in the Model Hénon 0.1–0.3 (Fig. 3), where synchronization is detected in the interval  $0.5 \lesssim C \lesssim 0.6$  for  $A = 0$ .

- Weakly coupled systems can be asymptotically synchronized, which can be detected via the auxiliary systems method both in the noiseless and noisy cases. Indeed, Table I shows the intervals of coupling strengths for which the responder is asymptotically stable, obtained with the auxiliary system method. Therefore, synchronization can occur only for couplings in the corresponding interval (as it does). Note that Table I excludes the spurious synchronization  $\Delta r(0) = 0$ .
- In general, when the noise amplitude increases, the threshold of stable synchronization moves towards stronger couplings. However, the Model Hénon 0.3–0.1 (Fig. 2) shows that there can be parameter settings for which that threshold is virtually the same for the noise amplitudes considered here.

Finally, let us point out that we also performed numerical simulations with perceptrons to detect the possible existence of the conventional synchronization map (period 1). The results were similar but not as sharp regarding weak synchronization states as the results obtained with LSTM nets to detect synchronization maps of period  $K > 1$ . However, no performance analysis of Method 1 with respect to the period  $K$  was carried out and, hence, no attempt was made to optimize the parameter  $K$  (which, anyway, depends on the data at hand).

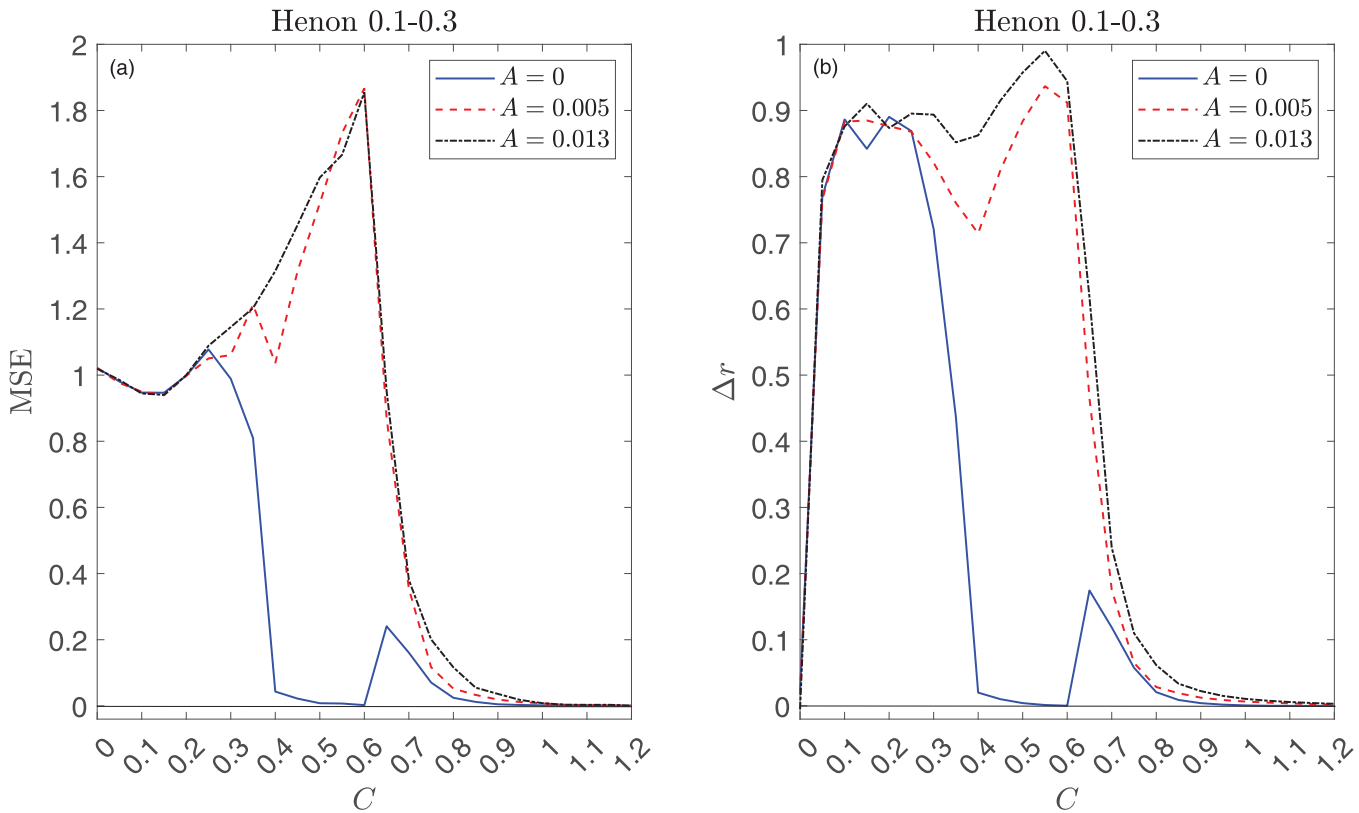


FIG. 3. Numerical results for the model Hénon 0.1–0.3, i.e.,  $b_1 = 0.1$  in (72) and  $b_2 = 0.3$  in (73). The information displayed in the panels (a) and (b) is the same as in Fig. 1.

VII. APPLICATION TO REAL-WORLD DATA: EEGs

The purpose of this section is to illustrate the application of Method 1 to real world data, specifically, intracranial EEG recordings from a subject with epilepsy. Therefore, we will not scrutinize here the complexity of such signals but rather check whether our findings align with results obtained in previous studies.

First of all, we notice that real observations  $(\varphi_X(x_t))_{1 \leq t \leq T}$  and  $(\varphi_Y(y_t))_{1 \leq t \leq T}$  of coupled systems  $X$  and  $Y$ , respectively, can deviate from our assumptions in Secs. II–VI in two important issues: nonstationarity or/and bidirectionality of the coupling, as it actually occurs with the time series in this section. To meet these challenges, this time we will apply Method 1 (Sec. VI) in both directions

TABLE I. Coupling strengths in the range  $0 \leq C \leq 1.2$  for which the responder is asymptotically stable.

	Hénon 0.3–0.3	Hénon 0.3–0.1	Hénon 0.1–0.3
$A = 0$	$0.40 \leq C \leq 1.20$	$0.15 \leq C \leq 1.20$	$0.40 \leq C \leq 1.20$
$A = 0.005$	$0.40 \leq C \leq 1.20$	$0.20 \leq C \leq 1.20$	$0.40 \leq C \leq 1.20$
$A = 0.013$	$0.40 \leq C \leq 1.20$	$0.20 \leq C \leq 1.20$	$0.40 \leq C \leq 1.20$

$X \rightsquigarrow Y$  and  $Y \rightsquigarrow X$ , on sufficiently short data segments to ensure approximate stationarity.

There is a subtlety, though. In the unidirectionally coupling  $X \rightsquigarrow Y$  studied in the previous sections, we detected synchronization by detecting a functional dependency between  $y_t$  and  $x_t, \dots, x_{t-K+1}$ , namely,  $y_t = H_{K, X \rightsquigarrow Y}(x_t, \dots, x_{t-K+1})$ , where  $H_{K, X \rightsquigarrow Y}$  is the reconstructed synchronization map of period  $K$  of the coupling  $X \rightsquigarrow Y$ , defined in Eq. (32). The robustness to noise of  $H_{K, X \rightsquigarrow Y}$  allowed us then to extend our conclusions to signals contaminated with low-amplitude noise. If, for the sake of this argument, we think of a bidirectional coupling  $X \leftrightarrow Y$  as the joint action of two separate unidirectional couplings  $X \rightsquigarrow Y$  and  $Y \rightsquigarrow X$ , then  $y_t$  will depend on  $x_t, \dots, x_{t-K+1}$  (whether  $X$  and  $Y$  are synchronized or not) through the cross map of period  $K$  of the coupling  $Y \rightsquigarrow X$ , i.e.,  $y_t = \Phi_{K, Y \rightsquigarrow X}(x_t, \dots, x_{t-K+1})$ ; see Eq. (17) with  $y$  and  $x$  swapped. Therefore, here we expect  $y_t$  to depend on  $x_t, \dots, x_{t-K+1}$  in general.

The bottom line is that, by using Method 1 in the directions  $X \rightsquigarrow Y$  and  $Y \rightsquigarrow X$ , we will be able to detect the “dominant driver” or the “coupling directionality” of the bidirectional coupling  $X \leftrightarrow Y$ . To this end, we are going to measure the strength of the coupling in both directions via the accuracy of the predictions of  $x_t$  and  $y_t$  made by LSTM nets in short non-overlapping segments over the entire EEGs, the dominant driver being given by the direction

with the strongest coupling. In case of equal strengths, the systems are assumed to be synchronized.

### A. Data description

The data that we are going to analyze is the following; see Lehnertz and Dickten<sup>60</sup> for more detail.

1. The signals are EEGs recorded intracranially from a subject with epilepsy during 86 090 s (23 h, 54 m, 50.4 s) with 48 electrode contacts at a sampling frequency of 200 Hz (sampling time = 5 ms). The subject had signed informed consent that her/his clinical data might be used and published for research purposes, and the study protocol had previously been approved by the ethics committee of the University of Bonn. The recording started at 7:00 am, corresponding to the initial sampling interval  $t = 1$ , and ended at the final sampling time  $t_{\text{final}} = 17.217\,984 \times 10^6$ . The epileptic convulsions occurred at the following sampling times.

- Average time of a first group of subclinical seizures:  $\bar{t}_{C1} = 3.4082 \times 10^6$  (17 041 s). By subclinical seizures we mean localized seizure activity on the EEG with no obvious clinical activity.
- Average time of a second group of subclinical seizures:  $\bar{t}_{C2} = 4.6082 \times 10^6$  (23, 041 s).
- Average time of a third group of subclinical seizures:  $\bar{t}_{C3} = 6.8762 \times 10^6$  (34 381 s).
- Onset time of a clinical seizure (the only one in the whole series):  $t_{C4} = 17.1842 \times 10^6$  (85 921 s).

2. A schematic of the implanted electrodes can be found in Fig. 2 of Lehnertz and Dickten.<sup>60</sup> The electrodes contacts are divided into the following three categories:

- focal (F), which comprises electrode contacts located within the seizure-onset zone;
- neighbor (N), which groups electrode contacts not more than two contacts distant to those of category F;
- other (O), gathering all remaining electrode contacts.

To designate the electrode contacts and their (approximately) 24 h recordings, we use the same labels as in Ref. 60. For example,  $X = \text{TR01}$  means that the system  $X$  is the source of the EEG  $(\varphi_X(x_t))_{1 \leq t \leq t_{\text{final}}}$  recorded at the electrode contact TR01.

3. For the sake of our analysis, we will consider the following five pairs  $(X, Y)$  of electrode contacts.
  - Case 1:  $(X, Y) = (\text{TR05} - \text{TR06})$  in the categories (F-F).
  - Case 2:  $(X, Y) = (\text{TR07} - \text{TBPR1})$  in the categories (F-N).
  - Case 3:  $(X, Y) = (\text{TR05} - \text{TL05})$  in the categories (F-O).
  - Case 4:  $(X, Y) = (\text{TBAR1} - \text{TLL04})$  in the categories (N-O).
  - Case 5:  $(X, Y) = (\text{TLR04} - \text{TLL04})$  in the categories (O-O).
4. As in the previous numerical simulations, the embedding dimension of the systems  $X$  and  $Y$  is 5. Thus,

$$\mathbf{x}_t = (\varphi_X(x_t), \varphi_X(x_{t+1}), \dots, \varphi_X(x_{t+4})) \quad (76)$$

$1 \leq t \leq t_{\text{final}} - 4$ , are the time-delay vectors corresponding to system  $X$ , and analogously with the EEG  $(\varphi_Y(y_t))_{1 \leq t \leq t_{\text{final}}}$  generated by the system  $Y$ .

5. For approximate stationarity,<sup>64,65</sup> we partitioned the time series  $(\mathbf{x}_t)$  and  $(\mathbf{y}_t)$ ,  $1 \leq t \leq t_{\text{final}} - 4$ , into 1434 non-overlapping segments

$$S_{X,n} = (\mathbf{x}_t)_{12\,000(n-1)+1 \leq t \leq 12\,000n-4} \quad (77)$$

and

$$S_{Y,n} = (\mathbf{y}_t)_{12\,000(n-1)+1 \leq t \leq 12\,000n-4} \quad (78)$$

of 11 996 points ( $\simeq 60$  s) each,  $n = 1, 2, \dots, 1434$ , and a last pair of segments

$$S_{X,1435} = (\mathbf{x}_t)_{1\,720\,800 \leq t \leq 1\,721\,798} \quad (79)$$

and

$$S_{Y,1435} = (\mathbf{y}_t)_{1\,720\,800 \leq t \leq 1\,721\,798}, \quad (80)$$

comprising only 9980 points ( $\simeq 50$  s). The segments  $1 \leq n \leq 720$ , correspond to the daylight hours (7 am–7 pm), while the segments  $721 \leq n \leq 1435$  correspond to the night hours. The clinical seizure occurs in the segment  $n = 1433$ , i.e., the third to last segment of the series, and it initiates just one second after the beginning of that segment ( $t_{C4} = 85\,921$  s).

6. As in Sec. VI, we use the first 80% of the data of each  $n$ th segment  $S_{X,n}$  and  $S_{Y,n}$  as training data, and the remaining 20% as testing data. So, this time we obtain two accuracy measures: (i)  $\text{MSE}_{X \rightarrow Y}(n)$ , for the predictions of  $\mathbf{y}_t$  output by the LSTM net, based on  $\mathbf{x}_t, \dots, \mathbf{x}_{t-K+1}$  with testing data of the segments  $S_{Y,n}$  and  $S_{X,n}$ , and (ii)  $\text{MSE}_{Y \rightarrow X}(n)$ , for the predictions of  $\mathbf{x}_t$  output by the LSTM net, based on  $\mathbf{y}_t, \dots, \mathbf{y}_{t-K+1}$  with testing data of the segments  $S_{X,n}$  and  $S_{Y,n}$ . As in Sec. VI, we set  $K = 10$  here. Of course, the parameter  $K$  can be fine-tuned for optimal results, but this is an issue not contemplated in the present work.

### B. Results

Since, at variance with the numerical simulations in Sec. VI, we have here bidirectionally coupled signals and two prediction accuracy measures  $\text{MSE}_{X \rightarrow Y}(n)$  and  $\text{MSE}_{Y \rightarrow X}(n)$ , we are going to use the coupling directionality index

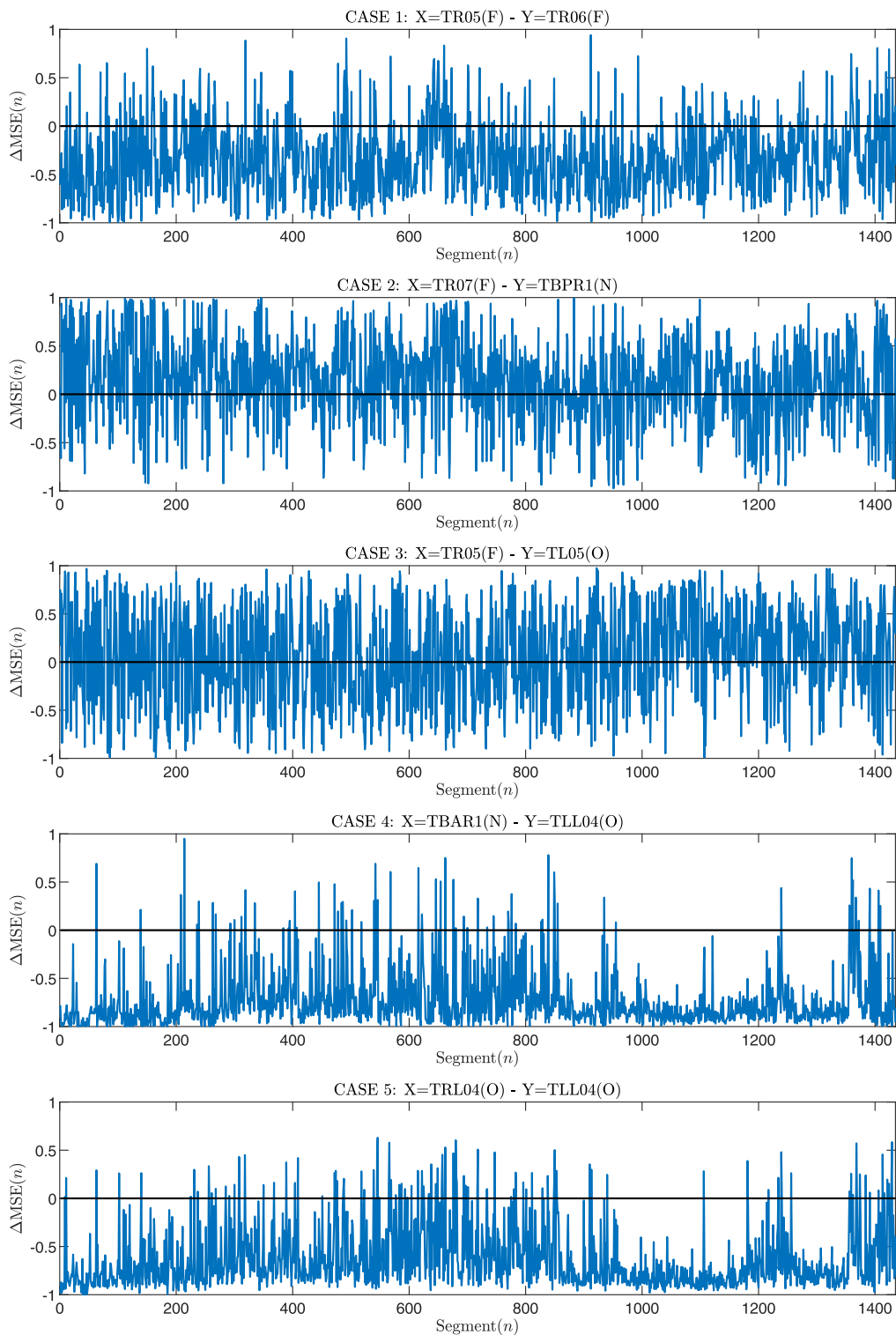
$$\Delta \text{MSE}(n) = \frac{\text{MSE}_{X \rightarrow Y}(n) - \text{MSE}_{Y \rightarrow X}(n)}{\text{MSE}_{X \rightarrow Y}(n) + \text{MSE}_{Y \rightarrow X}(n)}, \quad (81)$$

for each pair of data segments  $S_{X,n}$  and  $S_{Y,n}$ ,  $1 \leq n \leq 1435$ , so that

- (i)  $-1 \leq \Delta \text{MSE}(n) \leq +1$  and
- (ii)  $\Delta \text{MSE}(n) \geq 0$  if and only if  $\text{MSE}_{Y \rightarrow X}(n) \leq \text{MSE}_{X \rightarrow Y}(n)$ , i.e., knowledge of  $\mathbf{x}_t$  in the  $n$ th segment leads to better predictions of  $\mathbf{y}_t$ , than the other way around.

Therefore, if  $\Delta \text{MSE}(n) > 0$  (respectively,  $\Delta \text{MSE}(n) < 0$ ), then we conclude that  $X$  is the dominant driver (respectively,  $Y$  is the dominant driver). This interpretation agrees with other approaches based on the cross map,<sup>34,35</sup> transfer entropy,<sup>66,67</sup> phase dynamics,<sup>68</sup> etc. Otherwise, if  $\Delta \text{MSE}(n) = 0$ , then  $X$  and  $Y$  are assumed to be synchronized in segment  $n$  (although it might be difficult to discern this situation from “no-coupling”).

Figure 4 plots  $\Delta \text{MSE}(n)$  vs the segment number  $n$ ,  $1 \leq n \leq 1435$ . The numerical results are summarized in Table II for the 24 h EEGs, and in Table III for 12 h EEGs corresponding to daylight hours ( $1 \leq n \leq 720$ ) and night hours ( $721 \leq n \leq 1435$ ). It was



**FIG. 4.** Top to bottom: plots of  $\Delta\text{MSE}(n)$ , the directionality indicator (81), obtained using the segments  $S_{x,n}$  and  $S_{y,n}$ ,  $1 \leq n \leq 1435$ , given in Eqs. (77)–(79), for cases 1–5. The clinical seizure occurs in the segment  $n = 1433$ , too close to the right margin to be marked. See Sec. VII A for detail.



TABLE II. Results of cases 1–5 with 24 h EEGs.

Case	$\Delta\text{MSE}(n) > 0$	Dominant electrode
1 (F–F)	16% of segments	$Y = \text{TR06 (F)}$ dominates $X = \text{TR05 (F)}$
2 (F–N)	64% of segments	$X = \text{TR07 (F)}$ dominates $Y = \text{TBPR1 (N)}$
3 (F–O)	56% of segments	$X = \text{TR05 (F)}$ dominates $Y = \text{TL05 (O)}$
4 (N–O)	5% of segments	$Y = \text{TLL04 (O)}$ dominates $X = \text{TBAR1 (N)}$
5 (O–O)	7% of segments	$Y = \text{TLL04 (O)}$ dominates $X = \text{TLR04 (O)}$

not possible to highlight the clinical seizure in Fig. 4 because it occurs in the segment  $n = 1433$ , so any visual marks at that point are indistinguishable from the right margin of the corresponding panel.

In view of Fig. 4 and Tables II and III, we can draw the following general conclusions.

- The coupling directionality, as measured by  $\Delta\text{MSE}(n)$ , depends on the segment  $n$ . The overall dominance of the signals is stable with respect to day and night, although the *dominance degrees*, as measured by the percentages of segments contributing to the dominant coupling direction in the first or second 12 h, respectively, are different in all cases.
- Except in case 2, the sign of  $\Delta\text{MSE}(n)$  during the epileptic convulsions ( $n = 1433$ ) coincides with the overall coupling direction. In fact,

Case	1	2	3	4	5
$\Delta\text{MSE}(1433)$	−0.58	−0.87	0.60	−0.68	−0.40

- According to Osterhage *et al.*,<sup>62,63</sup> an important question in epileptology is whether the pathological interaction between the seizure-onset zone (label F) and other brain areas (labels N and O), a phenomenon called *focal driving*, can also be identified during seizure-free periods. Cases 2 and 3 in Table II answer this question affirmatively, that is, our method detects focal driving in the analyzed EEG.
- In addition, cases 2 and 3 in Table III indicate that focal driving is not diminished during sleep.
- More generally, Table III shows that focal driving does not appear to be influenced by other (possibly “stronger”) synchronization phenomena such as sleep.
- The dominance degree is rather high in cases 1, 4, and 5, with  $\Delta\text{MSE}(n) < 0$  over 80% of the segments both in the 24 h and 12 h

TABLE III. Results of cases 1–5 with 12 h EEGs (day and night).

Case	$\Delta\text{MSE}(n) > 0$ day	$\Delta\text{MSE}(n) > 0$ night
1 (F–F)	18% (Y dominant)	15% (Y dominant)
2 (F–N)	70% (X dominant)	58% (X dominant)
3 (F–O)	51% (X dominant)	60% (X dominant)
4 (N–O)	6% (Y dominant)	4% (Y dominant)
5 (O–O)	9% (Y dominant)	6% (Y dominant)

EEGs. Note that the interaction in those cases is local (cases 1 and 5) or it does not involve the seizure generating area (case 4).

The above findings are in line with the results of more comprehensive studies by Lehnertz and Dickten,<sup>60</sup> Dickten *et al.*,<sup>61</sup> and Osterhage *et al.*,<sup>62,63</sup> which empirically demonstrates the capability of our LSTM net-based method.

## VIII. CONCLUSION

Synchronization of two unidirectionally coupled dynamical systems  $X \rightsquigarrow Y$  is a classical topic in nonlinear dynamics. It is defined by the existence of a continuous function  $y_t = h(x_t)$  between the states  $x_t$  of the driver  $X$  and the states  $y_t$  of the responder, called the synchronization map. While  $h$ , when it exists, points from the state space of the driver (domain) to the state space of the responder (range), the cross map  $\mathbf{x}_t = \Phi(\mathbf{y}_t)$  always exists in that framework, is continuous, and points in the opposite direction between the corresponding reconstructed state spaces. In the standard, noiseless scenario, the existence of the synchronization map (i.e., synchronization between  $X$  and  $Y$ ) amounts to  $\Phi$  being invertible and bicontinuous. These and other fundamentals of generalized synchronization in the absence of dynamical noise were presented in a self-contained and unified way in Sec. II, with emphasis on the relationship between the cross map and the synchronization map.

In this context, the main contributions of the present paper are the following.

- (1) *Introduction of higher-period versions of the cross and synchronization maps* in Eqs. (17) and (21), the period-1 versions corresponding to the conventional concepts. They are based on the corresponding maps of order  $k$ , defined in Eqs. (14) and (20), and, actually, they may be defined in many different ways. A synchronization map of period 10 was used in the numerical simulation of Sec. VI because it gave better results than the conventional map in the detection of synchronization. Higher-period cross maps were invoked in Sec. VII to understand the sign of the directionality index (81) when the coupling is bidirectional. Optimization of the period was not discussed because it is beyond the scope of this paper.
- (2) *Generalizations of the synchronization map and the cross map when the driver is noisy* in Secs. V B and V C, respectively. To this end, the dynamical noise was modeled as stochastic forcing. The generalizations consist of families of maps that depend on noise parameters and coincide with their conventional counterparts when the noise is switched off. As usual, those generalizations have the wished properties under some formal provisos,

e.g., generacy of the maps and parameters involved, as well as the driver states in the case of the cross map. However, this does not mean that they are only useful in theory; they can be also useful in practice, e.g., in laboratory or numerical experiments, where typical properties are taken for granted and even noise parameters may be known.

- (3) *Application of LSTM nets to detect synchronization in synthetic data.* This method harnesses the existence of synchronization maps of higher periods in both the noiseless and noisy scenarios. To be more precise, in the numerical simulations of Sec. VI, synchronization was revealed in the reconstructed state spaces by detecting a period-10 synchronization map [Eq. (32)] using an LSTM net and predictability. The dynamical systems were two coupled Hénon maps with several parameter settings and noise amplitudes. As a benchmark, we used nearest-neighbor cross prediction based on the continuity of the cross map and its inverse (in case of synchronization). The agreement of the results obtained with the two methods was excellent. The results also showed the robustness of both methods against noise.
- (4) *Application of LSTM nets to detect the dominant driver in real-world data* in Sec. VII. The real-world data consisted of a 24 h EEG from a subject with epilepsy. To cope with the nonstationarity of the signal and the bidirectionality of the couplings between different brain areas, we partitioned the EEG in non-overlapping 60 s segments and measured the coupling dominance by the directionality index  $\Delta\text{MSE}(n)$  defined in Eq. (81); this index is based on the mean square error of predictions made by LSTM nets in the  $n$ th segment. The results agreed with results published in the literature, in particular, the existence of focal driving and its robustness to the day/night cycle.

## ACKNOWLEDGMENTS

J. M. Amigó and R. Dale were financially supported by Agencia Estatal de Investigación, Spain, Grant No. PID2019-108654GB-I00/AEI/10.13039/501100011033. J. M. Amigó was also supported by Generalitat Valenciana, Spain, Grant No. PROMETEO/2021/063.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Ethics Approval

The studies involving human participants were reviewed and approved by the ethics committee of the University of Bonn. The subject provided written informed consent to participate in this study.

Ethics approval for experiments reported in the submitted manuscript on animal or human subjects was granted. The studies involving human participants were reviewed and approved by the ethics committee of the University of Bonn. The subject provided written informed consent to participate in this study.

## Author Contributions

**José M. Amigó:** Conceptualization (lead); Formal analysis (lead); Funding acquisition (lead); Investigation (lead); Methodology (equal); Project administration (lead); Resources (lead); Supervision (lead); Validation (equal); Writing – original draft (lead); Writing – review & editing (equal). **Roberto Dale:** Data curation (equal); Software (lead); Validation (equal); Visualization (lead); Writing – review & editing (equal). **Juan C. King:** Data curation (equal); Software (lead); Validation (equal); Visualization (equal); Writing – review & editing (equal). **Klaus Lehnertz:** Data curation (lead); Methodology (equal); Supervision (equal); Validation (equal); Writing – review & editing (equal).

## DATA AVAILABILITY

The EEG dataset presented in this article is not publicly available due [state restrictions such as privacy or ethical restrictions]. Requests to access the dataset should be directed to [klaus.lehnertz@ukbonn.de](mailto:klaus.lehnertz@ukbonn.de).

## REFERENCES

- <sup>1</sup>V. S. Afraimovich, N. N. Verichev, and M. I. Rabinovich, “Stochastic synchronization of oscillation in dissipative systems,” *Radiophys. Quantum Electron.* **29**, 795–803 (1986).
- <sup>2</sup>N. F. Rulkov, M. M. Sushchik, L. S. Tsimring, and H. D. I. Abarbanel, “Generalized synchronization of chaos in directionally coupled chaotic systems,” *Phys. Rev. E* **51**, 980–994 (1995).
- <sup>3</sup>M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, “From phase to lag synchronization in coupled chaotic oscillators,” *Phys. Rev. Lett.* **78**, 4193–4196 (1997).
- <sup>4</sup>L. M. Pecora and T. L. Carroll, “Synchronization in chaotic systems,” *Phys. Rev. Lett.* **64**, 821–824 (1990).
- <sup>5</sup>S. Boccaletti, J. Kurths, G. Osipov, D. L. Valladare, and C. S. Zhou, “The synchronization of chaotic systems,” *Phys. Rep.* **366**, 1–101 (2002).
- <sup>6</sup>K. M. Cuomo, A. V. Oppenheim, and S. H. Strogatz, “Synchronization of Lorenz-based chaotic circuits with applications to communications,” *IEEE Trans. Circuits Syst. II* **40**, 626–633 (1993).
- <sup>7</sup>L. Kocarev and U. Parlitz, “General approach for chaotic synchronization with applications to communication,” *Phys. Rev. Lett.* **74**, 5028–5031 (1995).
- <sup>8</sup>T. Nowotny, R. Huerta, and M. I. Rabinovich, “Neuronal synchrony: Peculiarity and generality,” *Chaos* **18**, 037119 (2008).
- <sup>9</sup>J. M. Amigó, T. S. Mosqueiro, and R. Huerta, “Predicting synchronization of three mutually inhibiting groups of oscillators with strong resetting,” *Appl. Math. Inf. Sci.* **9**, 2245–2256 (2015).
- <sup>10</sup>P. Hövel, A. Viol, P. Loske, L. Merfort, and V. Vuksanović, “Synchronization in functional networks of the human brain,” *J. Nonlinear Sci.* **30**, 2259–2282 (2020).
- <sup>11</sup>L. J. Kocarev, K. S. Halle, K. Eckert, L. O. Chua, and U. Parlitz, “Experimental demonstration of secure communications via chaotic synchronization,” *Int. J. Bifurcation Chaos* **2**, 709–713 (1992).
- <sup>12</sup>J. M. Amigó, L. Kocarev, and J. Szczepanski, “Discrete Lyapunov exponent and resistance to differential cryptanalysis,” *IEEE Trans. Circuits Syst. II* **54**, 882–886 (2007).
- <sup>13</sup>W. Kinzel, A. Englert, and I. Kanter, “On chaos synchronization and secure communication,” *Philos. Trans. R. Soc. A* **368**, 379–389 (2010).
- <sup>14</sup>*Chaos Synchronization and Cryptography for Secure Communications: Applications for Encryption*, edited by S. Banerjee (Information Science Reference, Hershey, PA, 2011).
- <sup>15</sup>A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2001).
- <sup>16</sup>L. M. Pecora and T. L. Carroll, “Synchronization of chaotic systems,” *Chaos* **25**, 097611 (2015).

- <sup>17</sup>S. J. Schiff, P. So, T. Chang, R. E. Burke, and T. Sauer, "Detecting dynamical interdependence and generalized synchrony through mutual prediction in a neural ensemble," *Phys. Rev. E* **54**, 6708–6724 (1996).
- <sup>18</sup>M. Le Van Quyen, J. Martinerie, C. Adam, and F. J. Varela, "Nonlinear analyses of interictal EEG map the brain interdependences in human focal epilepsy," *Physica D* **127**, 250–266 (1999).
- <sup>19</sup>B. R. Hunt, E. Ott, and J. A. Yorke, "Differentiable generalized synchronization of chaos," *Phys. Rev. E* **55**, 4029–4034 (1997).
- <sup>20</sup>K. Pyragas, "Conditional Lyapunov exponents from time series," *Phys. Rev. E* **56**, 5183–5188 (1997).
- <sup>21</sup>H. D. I. Abarbanel, N. F. Rulkov, and M. M. Sushchick, "Generalized synchronization of chaos: The auxiliary system approach," *Phys. Rev. E* **53**, 4528–4535 (1996).
- <sup>22</sup>L. Kocarev and U. Parlitz, "Generalized synchronization, predictability and equivalence of unidirectionally coupled dynamical systems," *Phys. Rev. Lett.* **76**, 1816–1819 (1996).
- <sup>23</sup>J. Arnhold, P. Grassberger, K. Lehnertz, and C. E. Elger, "A robust method for detecting interdependencies: Application to intracranially recorded EEG," *Physica D* **134**, 419–430 (1999).
- <sup>24</sup>R. Quian Quiroga, J. Arnhold, and P. Grassberger, "Learning driver-response relationships from synchronization patterns," *Phys. Rev. E* **61**, 5142–5148 (2000).
- <sup>25</sup>R. Quian Quiroga, A. Kraskov, T. Kreuz, and P. Grassberger, "Performance of different synchronization measures in real data: A case study on electroencephalographic signals," *Phys. Rev. E* **65**, 041903 (2002).
- <sup>26</sup>R. Sowa, A. Chernihovskiy, F. Mormann, and K. Lehnertz, "Estimating phase synchronization in dynamical systems using cellular nonlinear networks," *Phys. Rev. E* **71**, 061926 (2005).
- <sup>27</sup>D. Krug, H. Osterhage, C. E. Elger, and K. Lehnertz, "Estimating nonlinear interdependencies in dynamical systems using cellular nonlinear networks," *Phys. Rev. E* **76**, 041916 (2007).
- <sup>28</sup>R. Monetti, W. Bunk, T. Aschenbrenner, and F. Jamitzky, "Characterizing synchronization in time series using information measures extracted from symbolic representations," *Phys. Rev. E* **79**, 046207 (2009).
- <sup>29</sup>J. M. Amigó, R. Monetti, T. Aschenbrenner, and W. Bunk, "Transcripts: An algebraic approach to coupled time series," *Chaos* **22**, 013105 (2012).
- <sup>30</sup>J. M. Amigó and K. Keller, "Permutation entropy: One concept, two approaches," *Eur. Phys. J. Spec. Top.* **222**, 263–273 (2013).
- <sup>31</sup>D. Ibáñez-Soria, J. García-Ojalvo, A. Soria-Frisch, and G. Ruffini, "Detection of generalized synchronization using echo state networks," *Chaos* **28**, 033118 (2018).
- <sup>32</sup>T. Lyburn, D. M. Walker, M. Small, and T. Jüngling, "The reservoir's perspective on generalized synchronization," *Chaos* **29**, 093133 (2019).
- <sup>33</sup>H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. (Cambridge University Press, Cambridge, 2004).
- <sup>34</sup>J. M. Amigó and Y. Hirata, "Detecting directional couplings from multivariate flows by the joint distance distribution," *Chaos* **28**, 075302 (2018).
- <sup>35</sup>G. Sugihara, R. May, H. Ye, C.-H. Hsieh, E. Deyle, M. Fogarty, and S. Munch, "Detecting causality in complex ecosystems," *Science* **338**, 496–500 (2012).
- <sup>36</sup>D. Harnack, E. Laminski, M. Schünemann, and K. R. Pawelzik, "Topological causality in dynamical systems," *Phys. Rev. Lett.* **119**, 098301 (2017).
- <sup>37</sup>X. Ying, S.-Y. Leng, H.-F. Ma, Q. Nie, Y.-C. Lai, and W. Lin, "Continuity scaling: A rigorous framework for detecting and quantifying causality accurately," *Research* **2022**, 9870149 (2022).
- <sup>38</sup>M. R. Muldoon, D. S. Broomhead, J. P. Huke, and R. Hegger, "Delay embedding in the presence of dynamical noise," *Dyn. Stab. Syst.* **13**, 175–186 (1998).
- <sup>39</sup>J. Stark, D. S. Broomhead, M. E. Davies, and J. Huke, "Delay embeddings for forced systems. II: Stochastic forcing," *J. Nonlinear Sci.* **13**, 519–577 (2003).
- <sup>40</sup>J. Stark, "Delay embeddings for forced systems. I: Deterministic forcing," *J. Nonlinear Sci.* **9**, 255–332 (1999).
- <sup>41</sup>I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (The MIT Press, Cambridge, MA, 2016).
- <sup>42</sup>F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*, Lecture Notes in Mathematics Vol. 898, edited by D. A. Rand and L. S. Young (Springer Verlag, 1981), pp. 366–381.
- <sup>43</sup>T. Sauer, J. A. Yorke, and M. Casdagli, "Embedology," *J. Stat. Phys.* **65**, 579–616 (1991).
- <sup>44</sup>H. Whitney, "Differentiable manifolds," *Ann. Math.* **37**, 645–680 (1936).
- <sup>45</sup>M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Phys. Rev. A* **45**, 3403–3411 (1992).
- <sup>46</sup>U. Parlitz, "Detecting generalized synchronization," *Nonlinear Theory Appl. IEICE* **3**, 113–127 (2012).
- <sup>47</sup>U. Parlitz, L. Junge, and L. Kocarev, "Subharmonic entrainment of unstable period orbits and generalized synchronization," *Phys. Rev. Lett.* **79**, 3158–3161 (1997).
- <sup>48</sup>L. M. Pecora, T. L. Carroll, and J. F. Heagy, "Statistics for mathematical properties of maps between time series embeddings," *Phys. Rev. E* **52**, 3420–3439 (1995).
- <sup>49</sup>N. F. Rulkov, V. S. Afraimovich, C. T. Lewis, J. R. Chazottes, and A. Cordonet, "Multivalued mappings in generalized chaos synchronization," *Phys. Rev. E* **64**, 016217 (2001).
- <sup>50</sup>J. M. Amigó, P. Kloeden, and A. Giménez, "Switching systems and entropy," *J. Differ. Equ. Appl.* **19**, 1872–1888 (2013).
- <sup>51</sup>J. M. Amigó, P. Kloeden, and A. Giménez, "Entropy increase in switching systems," *Entropy* **15**, 2363–2383 (2013).
- <sup>52</sup>M. F. Barnsley and S. Demko, "Iterated function systems and the global construction of fractals," *Proc. R. Soc. London, Ser. A* **339**, 243–375 (1985) See <https://www.jstor.org/stable/2397690>.
- <sup>53</sup>P. E. Kloeden and M. Rasmussen, *Nonautonomous Dynamical Systems* (American Mathematical Society, Providence, RI, 2011).
- <sup>54</sup>L. Arnold, *Random Dynamical Systems* (Springer Verlag, Berlin, 2003).
- <sup>55</sup>Y. Hirata, J. M. Amigó, S. Horai, K. Ogimoto, and K. Aihara, "Forecasting wind power ramps with prediction coordinates," *Chaos* **31**, 103105 (2021).
- <sup>56</sup>K. Petersen, *Ergodic Theory* (Cambridge University Press, Cambridge, 2000).
- <sup>57</sup>O. I. Moskalenko, A. E. Hramov, A. A. Koronovskii, and A. A. Ovchinnikov, "Effect of noise on generalized synchronization of chaos: Theory and experiment," *Eur. Phys. J. B* **82**, 69–82 (2011).
- <sup>58</sup>D. Mønster, R. Fusaroli, K. Tylén, A. Roepstorff, and J. F. Sherson, "Causal inference from noisy time-series data—Testing the convergent cross-mapping algorithm in the presence of noise and external influence," *Future Gener. Comput. Syst.* **73**, 52–62 (2017).
- <sup>59</sup>G. Datsis and U. Parlitz, *Nonlinear Dynamics* (Springer Nature, 2022).
- <sup>60</sup>K. Lehnertz and H. Dickten, "Assessing directionality and strength of coupling through symbolic analysis: An application to epilepsy patients," *Philos. Trans. R. Soc. A* **373**, 20140094 (2015).
- <sup>61</sup>H. Dickten, S. Porz, C. E. Elger, and K. Lehnertz, "Weighted and directed interactions in evolving large-scale epileptic brain networks," *Sci. Rep.* **6**, 34824 (2016).
- <sup>62</sup>H. Osterhage, F. Mormann, T. Wagner, and K. Lehnertz, "Measuring the directionality of coupling: Phase versus state space dynamics and applications to EEG time series," *Int. J. Neural Syst.* **17**, 139–148 (2007).
- <sup>63</sup>H. Osterhage, F. Mormann, T. Wagner, and K. Lehnertz, "Detecting directional coupling in the human epileptic brain: Limitations and potential pitfalls," *Phys. Rev. E* **77**, 011914 (2008).
- <sup>64</sup>C. Rieke, K. Sternickel, R. G. Andrzejak, C. E. Elger, P. David, and K. Lehnertz, "Measuring nonstationarity by analyzing the loss of recurrence in dynamical systems," *Phys. Rev. Lett.* **88**, 244102 (2002).
- <sup>65</sup>C. Rieke, F. Mormann, R. G. Andrzejak, T. Kreuz, P. David, C. E. Elger, and K. Lehnertz, "Discerning nonstationarity from nonlinearity in seizure-free and preseizure EEG recordings from epilepsy patients," *IEEE Trans. Biomed. Eng.* **50**, 634–639 (2003).
- <sup>66</sup>T. Schreiber, "Measuring information transfer," *Phys. Rev. Lett.* **85**, 461–464 (2000).
- <sup>67</sup>M. Staniek and K. Lehnertz, "Symbolic transfer entropy," *Phys. Rev. Lett.* **100**, 158101 (2008).
- <sup>68</sup>M. G. Rosenblum and A. Pikovsky, "Detecting direction of coupling in interacting oscillators," *Phys. Rev. E* **64**, 045202 (2001).