International Journal of Bifurcation and Chaos, Vol. 32, No. 13 (2022) 2230031 (16 pages)

DOI: 10.1142/S0218127422300312

Forecasting the Duration of Three Connected Wings in a Generalized Lorenz Model

E. L. Brugnago
Instituto de Física, Universidade de São Paulo,
05508-090 São Paulo, SP, Brazil
elb@if.usp.br

C. C. Felicio* and M. W. Beims†
Departamento de Física, Universidade Federal do Paraná,
81531-980 Curitiba, PR, Brazil
**carolini@fisica.ufpr.br*
†*mbeims@fisica.ufpr.br*

Received November 4, 2021; Revised August 8, 2022

This work considers the problem of predicting wing changes, and their duration, in systems able to support three interconnected wings (spirals) in the space of the variables. This is done by exploring the alignment of covariant Lyapunov vectors (CLVs) known to precede the occurrence of peaks and regime changes in some chaotic systems. Here, the alignment of the CLVs is combined with classification procedures, machine learning techniques and nearest-neighbors analysis to predict the duration inside coupled wings. The classification procedure has distinct classes defined as the number of maxima of one variable inside one wing, proportional to the time spent inside each wing. In general, we observe that predictions of significant duration times (around 17 oscillations) inside a wing can be efficiently predicted until four subsequently visited wings. For five to ten subsequently visited wings, the prediction efficiency decreases significantly for more than ten oscillations inside each wing. The numerical comparison shows the superiority of the nearest-neighbors methods over the multilayer perceptrons procedure for the predictions. Remarkably, accuracies larger than 0.9 are obtained for predictions of classes for ≤ 3 subsequently visited wings. Accuracies higher than 0.5 are obtained for predictions of classes in the ≤ 6 subsequently visited wings.

Keywords: Lorenz model; 3-wings chaotic attractor; covariant Lyapunov vectors.

1. Introduction

In general, it is impossible to guarantee accurate predictions of future events, given that they are necessarily uncertain. Thus, any development that provides forecasting insights based on past and present data in specific physical systems is highly desirable. Forecasting is of tremendous relevance in many areas like e.g. extreme and rare events, lasers peaks, extreme weathers, virus propagations, climate changes, critical transitions, stock markets, tipping points in dynamical systems, to mention a few [Lucarini *et al.*, 2016; Nazarimehr *et al.*, 2018; Nazarimehr *et al.*, 2020].

The inability to precisely provide (or control) all initial conditions leads to intrinsic uncertainties to predict the future in chaotic systems. In 1963, Lorenz considered this issue in the context of atmospheric convection by deriving a flow model [Lorenz, 1963]. The model contains the chaotic Lorenz attractor, formed by two connected spiralling motions,

each around its own unstable fixed point. Each spiral (or wing) can be visited many times. When the particle enters one wing, until the moment it leaves the wing, we describe the associated motion as a *regime*, properly defined in Sec. 2.3. Prediction of transitions between wings of the Lorenz attractor, and their duration, was studied using the concept of bred vectors [Evans *et al.*, 2004], the maximum of a peak (for a given variable) right before the regime changes [Yadav *et al.*, 2005], and more recently using machine learning techniques [Brugnago *et al.*, 2020c], together with information extracted from covariant Lyapunov vectors (CLVs) [Brugnago *et al.*, 2020b]. The *alignment* of CLVs along specific directions allows one to anticipate the occurrence of large peaks in chaotic time series [Beims & Gallas, 2016a, 2018], and also critical transitions [Sharafi *et al.*, 2017]. CLVs [Ginelli *et al.*, 2007] are obtained by integrating the equations of motion forward and backward along the same trajectory [Lai *et al.*, 1993; Norwood *et al.*, 2013; Ginelli *et al.*, 2007; Kuptsov & Parlitz, 2012; Ginelli *et al.*, 2013] and have been determined in many realistic systems [Truant & Morriss, 2014; Kuptsov & Kuptsova, 2014; Inubushi *et al.*, 2015; Palatella *et al.*, 2013; Krüger *et al.*, 2015; Jousseph *et al.*, 2016; Hansen & Hannestad, 2014; Beims & Gallas, 2016b; Sala *et al.*, 2012; Trevisan & Pancotti, 1998; Saiki & Kobayashi, 2010].

Broadly speaking, for an improved theoretical understanding of forecasting in chaotic systems, it is smart to methodically increase the complexity of the chaotic dynamics while ensuring the systems are as simple as possible. In other words, create a chaotic system with an extra intricate topological structure such as a multiwing attractor. In this context, the double scroll in Chua's circuit [Chua *et al.*, 1986], for example, was generalized to n-double scrolls [Suykens & Chua, 1997] and to a family of scroll grid attractors [Yalçin *et al.*, 2002]. Machine implementation using cellular neural networks, chaos and synchronization in systems with increasing complexity has important applications [Suykens & Vandewalle, 2005; Yao *et al.*, 2017; Caschera *et al.*, 2011; Monostori, 2003; Maxwell *et al.*, 2018]. Similarly, generalizations of Lorenz's model to n-wings regimes are also feasible and are denoted as *proto-Lorenz* systems [Miranda & Stone, 1993]. The possibility to have such additional wings, immediately raises the question about the predictability of the trajectory permanence on each wing (wings duration) in much more complicated systems.

The present work uses the three-dimensional proto-Lorenz system which presents a chaotic attractor composed of three wings. As time goes on, the three wings are visited many times, and the corresponding duration times can be estimated. The model also possesses three CLVs, each one pointing to one of the directions, the flow, stable or unstable direction. Some of the angles between these directions can *align*, furnishing relevant information about the change between wings. We combine the *alignment* of CLVs, classification procedures and machine learning techniques with supervised learning to predict the duration time inside the three-wings and subsequently visited wings. We treat the problem of predicting the time interval within each wing as a classification problem. Within each wing, there is a number of maxima of one dynamical variable which changes as times goes on. So, the time intervals inside the wings are classified by the number of such maxima. To distinguish the distinct classes we employ Multilayer Perceptrons (MLPs) [Haykin, 1999] and K-nearest neighbors (KNN) algorithm [Cover & Hart, 1967]. While the former is used to establish decision boundaries between the classes, the later counts the largest number of neighbors of the same class. The prediction efficiency using MLPs is compared to the KNN technique. However, the effectiveness of the alignment of the CLVs for the predictions is evident in all methods. We highlight that the technique proposed here, which uses information combined from two angles formed between CLVs, demonstrates an extension of the possibility of predicting the duration of wings beyond just one wing. This fact broadens the view of the employability of CLVs in the context of predictions in chaotic dynamical systems. The technique proposed here is crucial to improve the prediction time for chaotic time series. More specifically, while predictions of around 10 Lyapunov times occur with high accuracy prediction of around 18 Lyapunov times will have smaller accuracies. But still of relevance.

The paper is organized as follows. In Sec. 2, we briefly present the proto-Lorenz model with a 3-wings chaotic attractor, introduce the approach with CLVs and precisely outline the problem addressed in this study. Section 3 presents the results using the alignment of CLVs combined with machine learning techniques and nearest neighbors

methods to predict regime durations in subsequently visited wings. Section 4 summarizes our main results.

2. Model of 3-Wings, CLVs and Problem Delineation

2.1. *Model and numerical integration*

It has been shown that it is possible to construct a three-dimensional proto-Lorenz system, which presents, for certain configurations, a chaotic trajectory that evolves with n-wings around unstable fixed points [Miranda & Stone, 1993]. In the case $n = 3$ used in this study, the model is described as

$$\dot{p} = \frac{1}{3}[-(\sigma+1)p + (\sigma - r + z)q] + \frac{1}{3M}[-(1-\sigma)(p^2 - q^2) + 2(r + \sigma - r - z)pq], \quad (1)$$

$$\dot{q} = \frac{1}{3}[(r - \sigma - z)p - (\sigma+1)q] + \frac{1}{3M}[2(\sigma-1)pq + (r + \sigma - z)(p^2 - q^2)], \quad (2)$$

$$\dot{z} = \frac{1}{2}(3p^2 q - q^3) - bz, \quad (3)$$

with

$$M = \sqrt{p^2 + q^2}, \quad (4)$$

where (p, q, z) are variables, $(\dot{p}, \dot{q}, \dot{z})$ the corresponding derivatives with respect to time and (σ, r, b) the usual Lorenz system parameters. Here, we consider the canonical values $r = 28.0$, $\sigma = 10.0$ and $b = 8/3$.

For the numerical integration, we use the fourth-order Runge–Kutta method [Butcher, 2016] with constant time step $\delta = 10^{-3}$ and discard 10^6 integration steps as transient. The initial conditions (p_0, q_0, z_0) are randomly selected from the equiprobable distribution in the interval $0 < p_0, q_0, z_0 < 1$. Figure 1 displays the $q \times p$ projection of the proto-Lorenz attractor obtained from Eqs. (1)–(3). Regions A, B and C (according to background colors) delimit the three wings of the attractor with angles $2\pi/3$ rad. Each wing surrounds its own instable fixed point (blue) clockwise, as exemplified inside region A's wing with dark green, yellow to red color trajectory (see the dashed line and red arrow). In this example, it becomes evident that

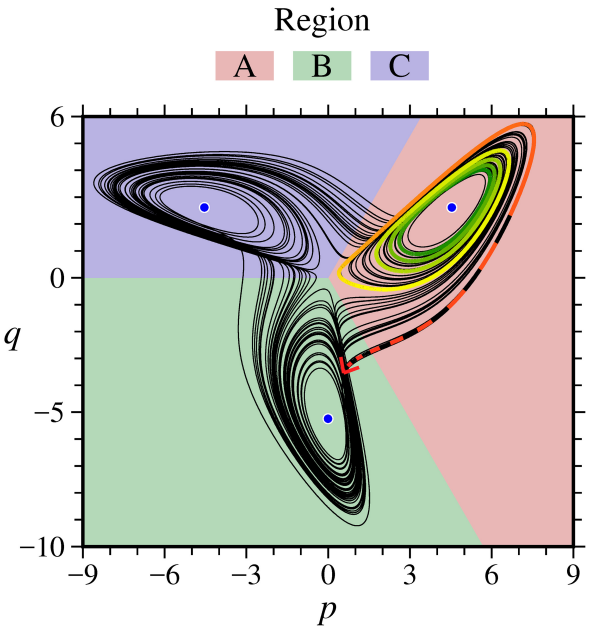

Fig. 1. Illustrative attractor projected in the $q \times p$ plane with 81 600 points. Background colors delimiting the wings located in regions A, B and C. The direction of motion follows the order A → B → C → A . . . . Blue circles are the locations of the unstable fixed points related to each wing. The chaotic trajectory starts in the inner part of the wing, closer to the fixed point (exemplified by the green color line inside region A's ring), moves to the outer part of the wing (yellow to red color line), until it changes the wing (red dashed line and arrow).

the trajectory leaves wing A from its outer part. The trajectory tends to enter the internal part of the wings, evolving outward around the fixed point and leaving the wing from the outer part. The direction of the wing succession is clockwise, so that the alternation A → B → C → A → . . . is observed. A precise definition for regions, wings and regimes is given in Sec. 2.3. For this chaotic attractor, the largest Lyapunov exponent $\lambda \approx 1.306$ was obtained along 10^7 points on the attractor. For that, we use a well-known algorithm described in [Wolf *et al.*, 1985; Wolfe & Samelson, 2007].

2.2. *Covariant Lyapunov vectors*

The CLVs are associated with the Lyapunov exponents (LEs) and are independent of the norm and are invariant under temporal inversion. The knowledge of the CLVs allows us to identify topological information about the tangent space's local structure. Each CLV is associated to a Oseledec subspace, a consequence of Oseledec's theorem

[Oseledets, 1968; Mañé, 1987]. For practical purposes in dynamical systems, each Oseledec subspace is related to stable, unstable or central invariant manifold in the tangent space. For the system and parameters considered in this work, three invariant manifolds in the tangent space are observed, the stable one related to negative LE, the unstable related to positive LE, and the central manifold to zero LE (flow direction). The CLVs furnish the associated contraction/expanding directions in the phase-space. In the present work, we use the most popular method [Ginelli *et al.*, 2013] to determine the CLVs, which is summarized in the Appendix.

Figure 2 displays the same attractor from Fig. 1 with additional colors (see color palette) indicating the angles between pairs of manifolds. Angles between manifolds are denoted by $\theta^{(ij)} = \cos^{-1}(\mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)})$, with $\mathbf{v}^{(i)}$ and $\mathbf{v}^{(j)}$ being normalized CLVs. In Fig. 2(a), we see the angle between unstable and central manifolds, unstable and stable manifolds in Fig. 2(b), and central and stable manifolds in Fig. 2(c). Since we are interested in the alignment of the manifolds, for $\theta^{(ij)} > \pi/2$, we use the transformation $\theta^{(ij)} \mapsto \pi - \theta^{(ij)}$. Each angle behaves similarly along the different wings. For example, in Fig. 2(c) the angle $\theta^{(23)}$ tends to $\pi/2$ close to the origin $(0, 0)$, to $\pi/4$ in intermediate parts and is less than $\pi/8$ at the edges and before the wing changes occurs. This repeats for the three wings. The alignment of unstable and central manifolds along the stable manifold precedes wing changes shown by the light blue color at the edges of the wings in Figs. 2(b) and 2(c).

2.3. *Problem delineation*

In this subsection, we outline a mathematical description and a thoughtful explanation of the regime duration prediction problem as a function of the alignment between pairs of CLVs, according to the treatment given in this work. First, the phase space V of the system given by Eqs. (1)–(3) is conveniently defined as the union of three regions, A, B and C, and a set $\mathrm{Z} = \{(0, 0, z) \in \mathrm{V}\}$: $\mathrm{V} = \mathrm{A} \cup \mathrm{B} \cup \mathrm{C} \cup \mathrm{Z}$. These sets are disjoints of two by two, that is $\mathrm{A} \cap \mathrm{B} = \emptyset$, $\mathrm{A} \cap \mathrm{C} = \emptyset$, $\mathrm{B} \cap \mathrm{C} = \emptyset$ and $(0, 0, z) \notin \mathrm{A} \cup \mathrm{B} \cup \mathrm{C}$, $\forall\ z$. This division of the phase space occurs as a function of the system variables p and q so that each wing of the chaotic attractor $\mathcal{A}$ is in one, and only one, region, as shown in Fig. 1. Observing the symmetry of the attractor, the regions comprise intervals of $2\pi/3$ rad in the

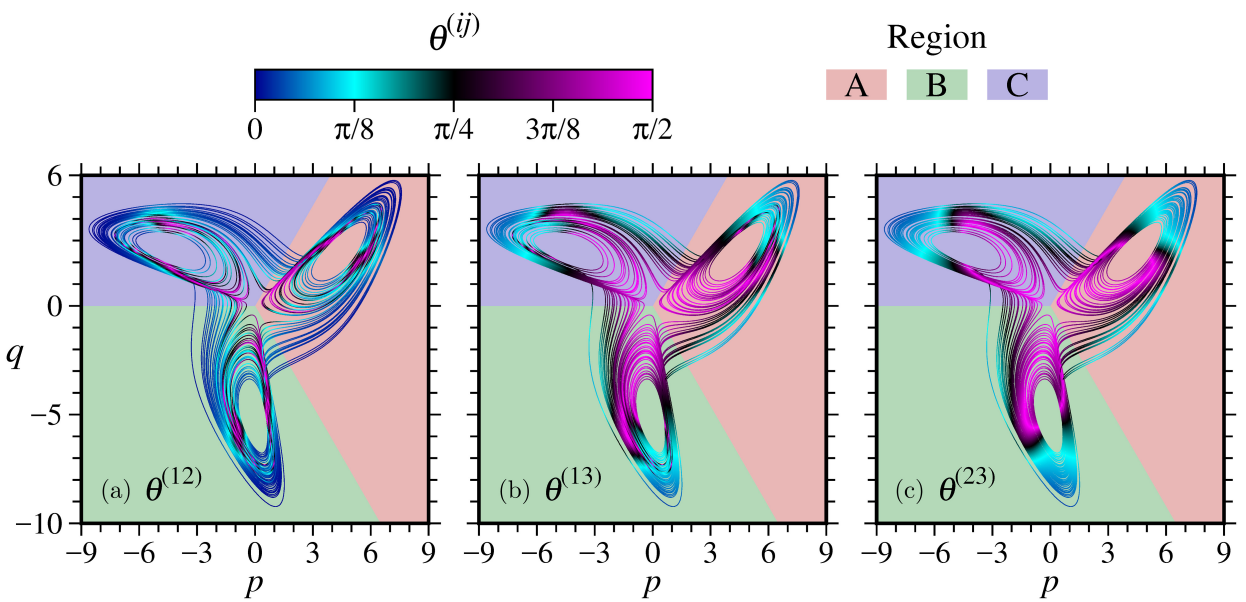

Fig. 2. The same attractor from Fig. 1 but with 10^5 points and additional colors (see color palette) indicating the angles between (a) unstable and central, (b) unstable and stable and (c) central and stable manifolds in the tangent space. The angles $\theta^{(23)}$ in (c) and $\theta^{(13)}$ in (b) tend to $\pi/2$ close to the origin $(0, 0)$, and are less than $\pi/8$ in the outer parts, before changing the wing. This occurs for the three wings. The alignment of unstable and central manifolds along the stable manifold precedes wing changes shown by the light blue color in the outer part of the wings in (b) and (c). Note that such alignment preceding wing changes are stronger in (a), as recognized by the dark blue color.

projection $q \times p$ of V, as

$$\begin{cases} \mathrm{A} = \Big\{(p,q,z) \in \mathrm{V}\backslash\mathrm{Z} : \\ \qquad -\dfrac{\pi}{3} \le \tan^{-1}\left(\dfrac{q}{p}\right) < \dfrac{\pi}{3} \text{ and } 0 < p\Big\}, \\ \mathrm{B} = \Big\{(p,q,z) \in \mathrm{V}\backslash\mathrm{Z} : \\ \qquad \dfrac{\pi}{3} < \cos^{-1}\left(\dfrac{p}{M}\right) \le \pi \text{ and } q \le 0\Big\}, \\ \mathrm{C} = \Big\{(p,q,z) \in \mathrm{V}\backslash\mathrm{Z} : \\ \qquad \dfrac{\pi}{3} \le \cos^{-1}\left(\dfrac{p}{M}\right) < \pi \text{ and } 0 < q\Big\}, \end{cases} \tag{5}$$

where $M = \sqrt{p^2 + q^2}$ as in Eq. (4). The delimitation that define regions provides a precise division of the chaotic attractor into three disjoint subsets, here called wings

$$\mathcal{W}^{\mathrm{A}} \coloneqq \mathcal{A} \cap \mathrm{A}, \quad \mathcal{W}^{\mathrm{B}} \coloneqq \mathcal{A} \cap \mathrm{B}, \quad \mathcal{W}^{\mathrm{C}} \coloneqq \mathcal{A} \cap \mathrm{C}. \tag{6}$$

Denoting by $X(t) = (p(t), q(t), z(t)) \in \mathrm{V}$; from an initial condition $X_0 = X(t = 0)$ belonging to the basin of attraction $\mathcal{B}$ of $\mathcal{A}$, we obtain a trajectory $\mathcal{T}(X_0)$ that transits through the three regions. So we define a regime $\tilde{\mathcal{T}}_n(X_0) \coloneqq \{X(t) : \tau_n \le t < \tau_{n+1} \text{ and } X(t = 0) = X_0\}$ a stretch of $\mathcal{T}(X_0)$, where τ_n is the instant at which the trajectory enters a region and τ_{n+1} corresponds to the start time of the following regime. Precisely, let R be a region such that $\tilde{\mathcal{T}}_n(X_0) \subset \mathrm{R}$, then $\tilde{\mathcal{T}}_{n+1}(X_0) \cap \mathrm{R} = \emptyset$. Taking τ_1 for $\tilde{\mathcal{T}}_1(X_0) \subset \mathrm{A}$, entails that

$$\begin{gathered} \tilde{\mathcal{T}}_{3m-2}(X_0) \subset \mathrm{A}, \\ \tilde{\mathcal{T}}_{3m-1}(X_0) \subset \mathrm{B} \quad \text{and} \\ \tilde{\mathcal{T}}_{3m}(X_0) \subset \mathrm{C}, \end{gathered} \tag{7}$$

with m being a positive integer. Asymptotically,

$$\begin{gathered} \tilde{\mathcal{T}}_{3m-2}(X_0) \subset \mathcal{W}^{\mathrm{A}}, \\ \tilde{\mathcal{T}}_{3m-1}(X_0) \subset \mathcal{W}^{\mathrm{B}} \quad \text{and} \\ \tilde{\mathcal{T}}_{3m}(X_0) \subset \mathcal{W}^{\mathrm{C}}. \end{gathered} \tag{8}$$

For practical purposes, $\tilde{\mathcal{T}}_1(X_0)$ can be contained by any of the three regions, only the cyclic property of the regimes is important.

After delineating the concepts of region, wing and regime in the context of this study, we proceed to the formulation of the prediction problem. The duration of the regime $\tilde{\mathcal{T}}_{n+a}$ is given by

$$\Delta t_{n+a} = \tau_{n+a+1} - \tau_{n+a}, \tag{9}$$

with a as a positive integer and, for most cases, if $m \neq n$ we have $\Delta t_{m+a} \neq \Delta t_{n+a}$. We propose to predict Δt_{n+a} as a function of quantities $\mu_n^{(k)}$ measured, or calculated, along the nth regime. Specifically, we consider extreme alignments $\varphi_n^{(ij)}$ between pairs of CLVs as predictor variables, as

$$\mu_n^{(k)} = \varphi_n^{(ij)} \coloneqq \min\{\theta^{(ij)} : \tau_n \le \tau'_n \le t < \tau_{n+1}\}, \tag{10}$$

where τ'_n is chosen in order to delimit the range of the regime that will be taken to evaluate the alignment between CLVs. For the studied system there are three combinations (ij), it follows that

$$\mu_n^{(1)} = \varphi_n^{(12)}, \quad \mu_n^{(2)} = \varphi_n^{(13)} \quad \text{and} \quad \mu_n^{(3)} = \varphi_n^{(23)}. \tag{11}$$

3. Results

The analyses and results presented in this section were performed on data obtained from 100 different time series of the system Eqs. (1)–(3), each one evolved from a different initial condition randomly selected, as described in Sec. 2.1. For each initial condition, we perform the numerical evolution of the system, discarding the transient of 10^6 integration steps and, later, obtaining 10^7 points along the chaotic attractor. Thus, we compute approximately 5.7×10^3 regimes in each time series, accumulating a total of 573 111 regimes for which we obtain the durations and extreme alignments of the CLVs.

3.1. *Alignment of CLVs, classes and predictions*

First, we analyze the relationship between the alignment of the tangent space manifolds during the current regime with the regime transition. This can be better appreciated in the time window as seen in Fig. 3. The time series of p and q are plotted in Fig. 3(a), with background colors identifying regions, as in Figs. 1 and 2. Variable z is plotted in Fig. 3(b), where the number of local maxima of z inside the nth wing defines the class Ω_n. For example, in Fig. 3(b) we have $\Omega_n = 4$ maxima inside

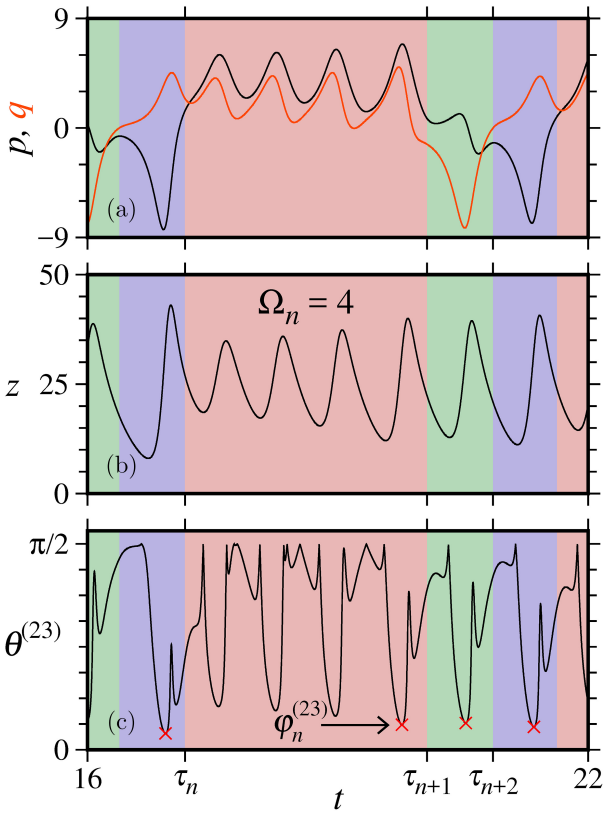

Fig. 3. Time window showing (a) p (black line) and q (orange line) evolution, (b) variable z and (c) the angle $\theta^{(23)}$. Background colors identify the regions A, B and C, like the color scheme in Figs. 1 and 2. As an example, in (b) we show the class $\Omega_n = 4$ (maxima of z) inside region A. The smallest value of $\theta^{(23)}$ ($= \varphi_n^{(23)}$) inside any region occurs before a region change occurs (see red crosses). Consequently, collecting all red crosses along the time series, the data of $\varphi_n^{(23)}$ (the strongest alignment of $\theta^{(23)}$), we get relevant information when regional changes occur.

region A. Figure 3(c) displays the angle $\theta^{(23)}$ inside the same time window. The smallest value of $\theta^{(23)}$ is defined as $\varphi_n^{(23)}$, according to Eq. (10). This alignment, inside any region, occurs before a wing change occurs, as indicated by red crosses in Fig. 3(c). Due to the high alignment of CLVs pairs observed in long-term regimes, for simplicity in the data treatment, we take $\tau'_n = \tau_n$ only when $\Omega_n = 1$, τ'_n being the time of the last minimum of z for $\Omega_n > 1$. We adopt the alignment only after the last minimum of z, in order to consider only the extreme alignment closest to the wing change. Therefore, $\varphi_n^{(23)}$ is the smallest value of $\theta^{(23)}$ inside one region during the last lap in the wing, and is the strongest alignment of $\theta^{(23)}$ just before the change of wings.

In order to discuss the relation between the duration Δt_{n+1} [$a = 1$ in Eq. (9)] and the extreme alignment $\varphi_n^{(ij)}$, we present Fig. 4. In this figure we consider small regions in $\Delta t_{n+1} \times \varphi_n^{(ij)}$ plane, but large enough to support the discussion about the predictability of the duration of regimes. Note that we are considering $\varphi_n^{(ij)}$ during the regime $\tilde{\mathcal{T}}_n$ on any wing and, then, looking at Δt_{n+1}, of the regime $\tilde{\mathcal{T}}_{n+1}$, on the succeeding wing. Colors indicate the values of the classes $\Omega_{n+1} = 1, 2, 3, 4, 5$, corresponding to the number of local maxima in z during $\tilde{\mathcal{T}}_{n+1}$. For clarity, this figure shows a cutout of the whole available space of possibilities. While the duration values and the classes are strongly related, the projections along the angle axis are far from unique. For example, in Fig. 4(a), more values of Δt_{n+1} are related to one $\varphi_n^{(12)}$. This is not desirable when precise predictions of Δt_{n+1} are expected. However, as seen in Figs. 4(b) and 4(c), the projections on the angle axis become better, meaning that the number of allowed angles for a given Δt_{n+1} diminishes. The best case is for $\varphi_n^{(23)}$ in Fig. 4(c).

In this study, we seek to predict the class Ω_n as an alternative to predict the duration Δt_n, given the reciprocity between the number of turns taken in a given regime before switching wings, and the assigned classes are the same. As for the duration averages, using linear regression from the data, we can establish the relationship $\Delta t_n = 0.638\Omega_n + 0.388$ as a good approximation. In the following subsections, we show that the angles between CLVs are related to the number of peaks in z (class) that will be observed while the trajectory is in a given region (A, B or C). Figure 5(a) displays the classes Ω_{n+1} as a function of the angles $\varphi_n^{(13)}$ and $\varphi_n^{(23)}$. Here, we combine two angles, as done by Brugnago *et al.* [2020a]. The colors associates the class. All classes with $\Omega_{n+1} \geq 20$ are marked with black. As can be seen, the classes are well divided in regions, and the class number of each regime increases as both angles approach zero. This can be better observed in the magnification shown in the inset of the same figure. Regarding the prediction of the duration of regimes, we have the following two crucial points:

(i) smaller values of $\varphi_n^{(13)}$ and $\varphi_n^{(23)}$ can predict classes with larger values, and this is related to the strong alignment of the unstable and central manifolds along the stable manifold direction;
(ii) the sharp division between distinct classes can measure the trustworthiness of the prediction.

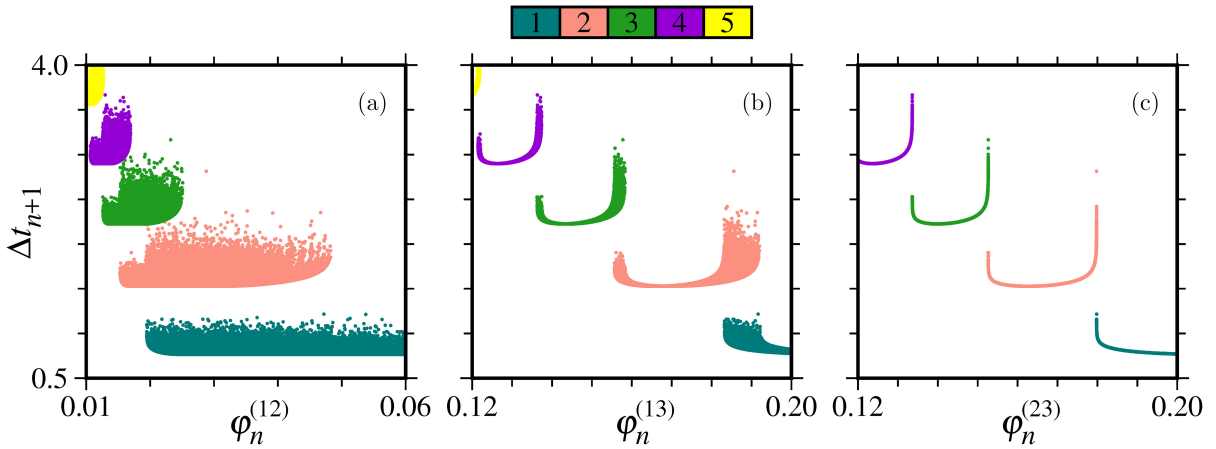

Fig. 4. Plane $\Delta t_{n+1} \times \varphi_n^{(ij)}$ shows the relation between the duration Δt_{n+1} and the extreme alignment (a) $\varphi_n^{(12)}$; (b) $\varphi_n^{(13)}$; and (c) $\varphi_n^{(23)}$ for each nth regime. Colors are the classes of the $(n+1)$th regimes. Results in (a) demonstrate clearly that projections along the angle axes are far from unique, making the prediction of Δt_{n+1}, for a given $\varphi_n^{(12)}$, very imprecise. Such imprecisions tend to significantly diminish for angles $\varphi_n^{(13)}$ in (b), and $\varphi_n^{(23)}$ in (c).

It is measured by the proportion ν of first neighbors containing the same class.

We define $\nu := Q/N$, with N being the number of points $P_n(\mu_n^{(i)}, \mu_n^{(j)})$ taken into account, that is, the total number of elements in the analyzed dataset. In order to determine Q, we consider the two first neighbors at each point, one corresponding to each axis of the plane $\mu_n^{(j)} \times \mu_n^{(i)}$. So, the point $P_n(\mu_n^{(i)}, \mu_n^{(j)})$ has a neighbor $P_{h/v}(\mu_{h/v}^{(i)}, \mu_{h/v}^{(j)})$ in the horizontal/vertical direction. It is found that

$$h = \underset{m}{\text{argmin}}(|\mu_m^{(i)} - \mu_n^{(i)}|, m \neq n),$$
$$v = \underset{m}{\text{argmin}}(|\mu_m^{(j)} - \mu_n^{(j)}|, m \neq n), \tag{12}$$

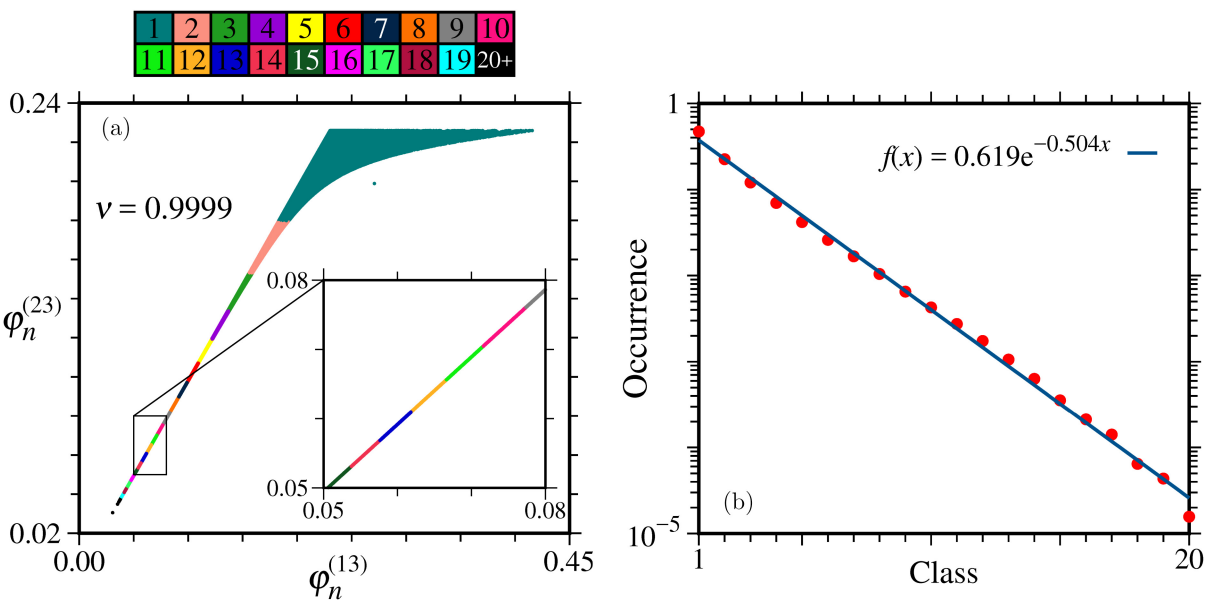

Fig. 5. (a) Classes Ω_{n+1} (see colors legend) as a function of the angles $\varphi_n^{(13)}$ and $\varphi_n^{(23)}$. There are 573 111 points obtained from 100 time series. The classes are well partitioned and the inset shows the growing pattern of classes as the alignment increases. In other words, as the angles $\varphi_n^{(13)}$ and $\varphi_n^{(23)}$ diminish, the classes increase, meaning that larger durations Δt_{n+1} are expected. The quantifier ν close to 1 indicates that points from the same class are closer to each other than to points of another class. (b) Occurrence of each class (points in red). The occurrence decays exponentially, as indicated by the straight line (blue line). The class 20 includes cases with 20 or more turns inside the same wing.

being the correspondence between $\mu_n^{(k)}$ and $\varphi_n^{(ij)}$ as given in Eqs. (11). Thus, Q is the number of points for which at least one of the two neighbors matches in class. In Fig. 5(a) we have $\nu = 0.9999$. For other angle combinations, $\varphi_n^{(12)} \times \varphi_n^{(13)}$ and $\varphi_n^{(12)} \times \varphi_n^{(23)}$, we obtain smaller values for ν, suggesting that grouping-based classification methods will be most effective for the selected pair $(\varphi_n^{(13)}, \varphi_n^{(23)})$. The simultaneous use of the three angles did not show any significant improvement. We see in Fig. 5(a) that the larger value classes occupy smaller regions. The occurrence of each class falls exponentially with its number, as shown in Fig. 5(b). Due to the small number of events with classes larger than 19, the class 20 in Fig. 5(b) concatenates cases with 20 or more turns inside the same wing.

Note that the results above are for predicting regime duration from the next wing (or region). For example, the alignment of $\varphi_n^{(23)}$ in region A can predict the duration of regime on the next visit to B. The same can be said for the alignment of $\varphi_n^{(23)}$ in region B to predict C, and so on. An exciting issue comes to mind when we ask about the possibility to predict the duration of regimes of subsequently visited wings. For example, is the alignment of $\varphi_n^{(23)}$ in region A able to predict the duration of the regimes on sequential visits to C, A, B, ...? In other words, is the alignment of $\varphi_n^{(23)}$ able to predict duration $\Delta t_{n+2}, \Delta t_{n+3}, \ldots$ of the classes $\Omega_{n+2}, \Omega_{n+3}, \ldots$, respectively? Such larger advances correspond to taking $a > 1$ in Eq. (9). Throughout this work we studied advances $1 \leq a \leq 10$.

Figure 6 shows the answer to the above question(s). It displays the classes (a) Ω_{n+2}, (b) Ω_{n+3}, (c) Ω_{n+4}, (d) Ω_{n+5}, (e) Ω_{n+7} and (f) Ω_{n+10} as a function of the angles $\varphi_n^{(13)}$ and $\varphi_n^{(23)}$. The white background is the region of class $\Omega_{n+1} = 1$ from Fig. 5(a), and the khaki background is the region of class $\Omega_{n+1} = 2$ from the same figure. Let us discuss Fig. 6(a) in more detail using an example. Consider to start in region A of Fig. 1, and that the pair $(\varphi_n^{(13)}, \varphi_n^{(23)})$ predict the class Ω_{n+1} in

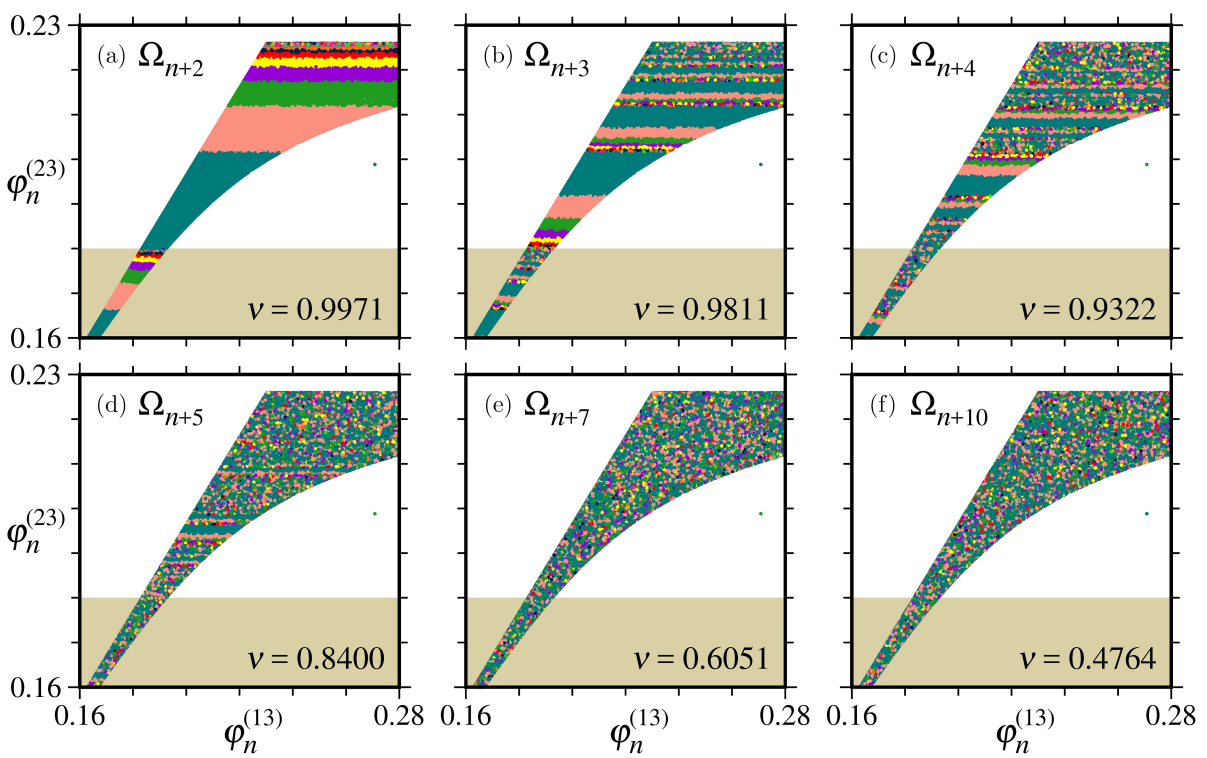

Fig. 6. Classes (a) Ω_{n+2}, (b) Ω_{n+3}, (c) Ω_{n+4}, (d) Ω_{n+5}, (e) Ω_{n+7} and (f) Ω_{n+10} are shown as a function of the angles $\varphi_n^{(13)}$ and $\varphi_n^{(23)}$. The colors of the classes follow the legend of Fig. 5(a). The white background is the region of class $\Omega_{n+1} = 1$ in Fig. 5(a), and the khaki background is the region of class $\Omega_{n+1} = 2$ in the same figure. In (a) we observe that inside the classes $\Omega_{n+1} = 1$ and $\Omega_{n+1} = 2$ in Fig. 5(a), new sequences of higher-order classes emerge. This emergence of new classes occurs more and more, visible in (b) and (c), and colors become more and more mixed in (d)–(f), making the predictions task almost impossible for longer times.

Fig. 5(a), which correspond to the duration of the next regime in region B. Thus, Fig. 6(a) displays the classes obtained in region C predicted by $\varphi_n^{(13)}$ and $\varphi_n^{(23)}$ obtained in region A. While the white background corresponds to $\Omega_{n+1} = 1$ in region B, the khaki background is related to $\Omega_{n+1} = 2$ in region B. In contrast to Fig. 5(a), inside one background color in Fig. 6(a) the classes are larger with increasing values of $\varphi_n^{(13)}$ and $\varphi_n^{(23)}$. This means that for stronger alignments in region A, smaller classes can be predicted in region C. This changes each time a further region is considered. For example, in Fig. 6(b), inside one background color and one class of Fig. 6(a), the predicted classes Ω_{n+3} are larger with decreasing values of $\varphi_n^{(13)}$ and $\varphi_n^{(23)}$. In this case, the predicted classes correspond to one complete loop in the attractor, namely, values of $\varphi_n^{(13)}$ and $\varphi_n^{(23)}$ from A, are used to predict the classes Ω_{n+3} after one loop on the attractor, i.e. after the transitions A $\rightarrow$ B $\rightarrow$ C $\rightarrow$ A. Figures 6(c)–6(f) present the predicted classes for additional transitions among wings. We see that the color become mixed, and the efficiency of the prediction is lost. It is possible to quantify such loss by the quantity ν, presented in Figs. 6(a)–6(f), which decreases when we try to predict the duration of regimes after increasing transitions between regions.

3.2. *MLP and KNN predictions*

Note that the well-separated colors in Fig. 5(a) allows to identify values for the pairs $(\varphi_n^{(13)}, \varphi_n^{(23)})$ to each color. Consequently, the classes have well-defined regions in the plane $\varphi_n^{(23)} \times \varphi_n^{(13)}$ in Fig. 5(a). This suggests to employ a ML-based classifier of the MLP to establish decision boundaries. We use a network with two hidden layers of 100 and 20 neurons, respectively, all with the Rectified Linear Unit (ReLU) [Yarotsky, 2017] type.

A schematic diagram of the MLP employed in this study is presented in Fig. 7. The input layer neurons (black rectangle) distribute the information from the extreme angles to the first hidden layer, which consists of 100 neurons, without making transformations in the input information. As already mentioned, the neurons of the two hidden layers (region highlighted in blue) are of the ReLU type, the last one, consisting of 20 neurons, connects to the output layer (pink rectangle), which returns a sorting from a *softmax* function [Sharma *et al.*, 2020]. Artificial neural network learning takes place by updating the weights of the connections between the successive layers of neurons. In Fig. 7 we represent the matrices of the weights of these connections by $W_{\text{in},1}$ (input layer for first hidden layer), $W_{1,2}$ (first hidden layer for second hidden layer) and $W_{2,\text{out}}$ (second hidden layer for output layer). Weights are updated during training seasons using a *backpropagation* algorithm taking into account *gradient descent* [Haykin, 1999].

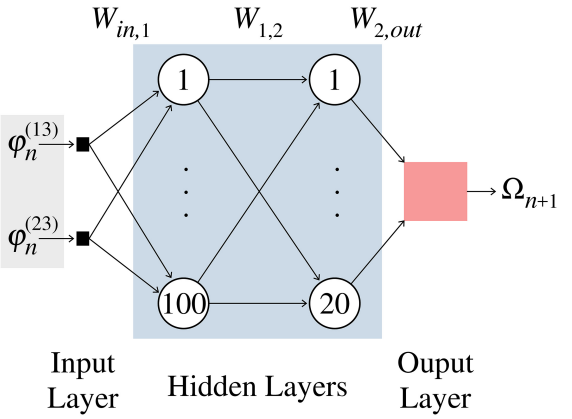

Fig. 7. Schematic diagram of the MLP used. Black rectangles represent the neurons of the input layer, the circles in the blue area represent the neurons of the hidden layers, and the pink square is the output layer. $W_{a,b}$ represents the matrices of the weights of the connections between neurons of different layers. From the implementation point of view, the input layer neurons distribute the information from the extreme angles $\varphi_n^{(ij)}$ to the first hidden layer, without making transformations in the input information. Artificial neural network learning process takes place by updating the weights of the connections between successive layers of neurons. Weights are updated during the training seasons using a backpropagation algorithm.

From the 573 111 points in the data set, we randomly generated two new sets with 250 000 points each, one for the MLP training and the other for testing. Both maintain the same exponential class distribution shown in Fig. 5(b). The training stage consists of 200 epochs, that is, the network updates its weights from the training set for 200 repetitions. It is important to mention some additional MLP configurations: we use the optimizer *Adam* [Goodfellow *et al.*, 2016], for error calculation we use *categorical cross-entropy* [Rubinstein & Kroese, 2013] and the validation in each season on 20% of the sample (50 000 points), with weight updates every 100 points. To predict the class with an advance $a = 1$, we obtained a prediction accuracy of 0.9821. Table 1 shows the prediction accuracy in each class,

Table 1. Prediction accuracy (PA) in each class of the MLP. The columns −1 is the precision deviation for one lower class and +1 is the precision deviation for one higher class.

Class	PA	−1	+1	Class	PA	−1	+1
1	0.9983	—	0.0017	11	0.9642	0.0089	0.0268
2	0.9595	0.0291	0.0114	12	0.9823	0.0000	0.0177
3	0.9526	0.0017	0.0458	13	0.9884	0.0000	0.0116
4	1.0000	0.0000	0.0000	14	0.9620	0.0000	0.0380
5	0.9641	0.0253	0.0106	15	1.0000	0.0000	0.0000
6	1.0000	0.0000	0.0000	16	1.0000	0.0000	0.0000
7	0.9875	0.0125	0.0000	17	0.9318	0.0000	0.0682
8	0.9852	0.0148	0.0000	18	1.0000	0.0000	0.0000
9	0.9969	0.0031	0.0000	19	1.0000	0.0000	0.0000
10	0.9715	0.0285	0.0000	20	1.0000	0.0000	—

emphasizing the highest ones with 100% accuracy. In the −1 columns, we have the forecast deviation for one class below. For example, for class 2 we have accuracy of 0.9595, and in the −1 column we have 0.0291, meaning that 0.0291 should be of class 2 but were recognized as class 1. In the columns +1 the same for the deviation of one class above.

We also employ a KNN algorithm, which consists of a nonparametric technique applicable for both classification and regression. This method uses a reference set of previously labeled sample points and assigns labels to the test points according to the values of the K nearest neighbor. In the problem discussed here, we use this algorithm as a classifier and the reference set formed by labeled points $\Phi_i^{(a)}(\varphi_i^{(13)}, \varphi_i^{(23)}; \Omega_{i+a})$, the same 250 000 used for training the MLP. The class $\tilde{\Omega}_n^{(a)}$ assigned by this algorithm to a test point $P_n(\varphi_n^{(13)}, \varphi_n^{(23)})$ is given by the most frequent class observed in the K reference points closest to the test one. To compute the distances between the reference and test points, different metrics can be adopted, including measures that are not metrics, in the strict sense.

Figure 8 displays the comparison between MLP prediction accuracy and KNN methods. The classification from the first K neighbors assigns classes to the test setpoints according to the class of most neighbors in the training set. In the case of a tie, we assume the lowest class. To determine the neighborhood in the KNN method, we use two different distance calculations, namely, KNN with *Euclidean* metrics in the plane (KNNE) and the *minimum* distance in one of the two directions in the plane (KNNM). The Euclidean metric is well known, and the minimum distance does not consist of a metric and can be written as $d(x, y) = \min(|x_1 - y_1|, |x_2 - y_2|)$.

For the comparison with the MLP, we chose $K = 1$ in the KNNE and KNNM methods. We use the MLP training set as the database to classify the test set points in both cases. The results consider only one advance ($a = 1$), since for $a > 1$ there is a large mixture between the classes' bands, mainly those with $\Omega_{n+a} \geq 10$, which makes it difficult for the MLP to establish decision boundaries. For classes $\Omega_{n+1} \geq 13$ the accuracy is the same for KNNE and KNNM. We employ KNN for advances $a > 1$. As in the previous case, we formed the sets of references (training) and tests, all with 250 000 elements and following the same distribution presented

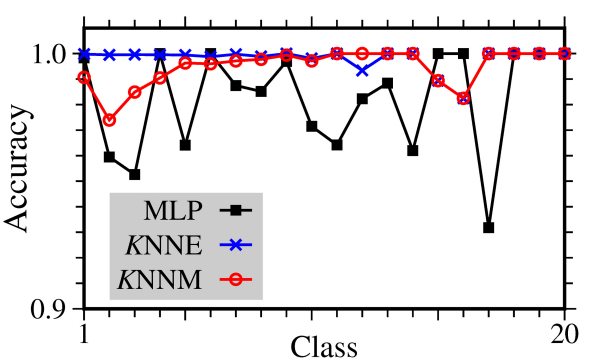

Fig. 8. Comparison between results from MLP and KNN methods for the advance $a = 1$, detailed by the classes. KNN with neighborhood $K = 1$. In general, the MLP method shows (black squares) a smaller accuracy compared to both KNN procedures. Euclidean metric (KNNE in blue) gives better results for classes $\Omega_{n+1} \leq 9$ than the minimum distance (KNNM in red). From $\Omega_{n+1} \geq 13$ both KNN applications present very close accuracies. Classes 15 and 16 are highly predicted by MLP, also the classes $\Omega_{n+1} \geq 18$ have similar results to KNN method.

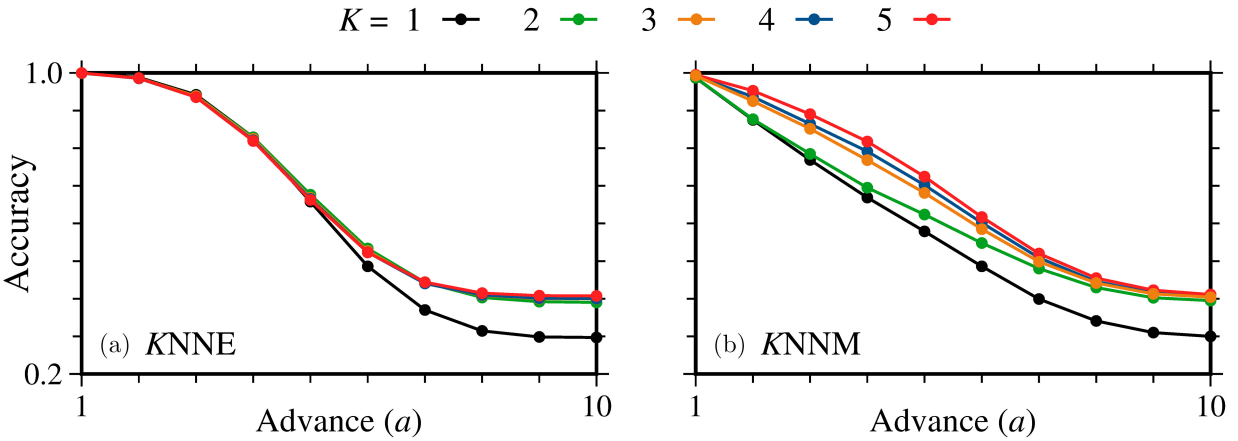

Fig. 9. Accuracy using the KNN methods to predict Ω_{n+a}, where $a = 1, 2, \ldots, 10$. The number K of neighbors is identified by colors, and the distance measure to the neighbors is (a) Euclidean and (b) minimum. While in (a) two neighbors are enough to reach the many neighbors effect, in (b) the accuracy is more dependent on the number of used neighbors.

for the general case in Fig. 5(b). Then we compare the accuracy in the prediction obtained from the two measurements, Euclidean and minimum distance. Figure 8 demonstrates that all prediction quantifiers are very efficient with accuracy larger than 0.9. However, in general, the MLP method shows a smaller efficiency compared to both KNN procedures. The only exceptions are the classes 15 and 16. Furthermore, the KNNM method is less efficient than the KNNE, at least for some specific classes around 2 to 4.

Figure 9 shows the accuracy of the hit rates using KNNE and KNNM as a function of the classes Ω_{n+a} to be predicted. The number of neighbors considered is $K = 1, 2, \ldots, 5$ and is identified by the colors, as shown in Fig. 9. The neighbors are shown using the Euclidean distance in Fig. 9(a) and using the minimum distance in Fig. 9(b). In both cases, hit rates decrease with a. However, while in Fig. 9(a) the curves for $K > 1$ overlap, in Fig. 9(b) they separate for almost every a. In general, we observe that the accuracy becomes better by increasing the number of neighbors and the KNNM refines this behavior. Furthermore, the minimum distance is more appropriate to describe the details of the classes.

Figure 10 displays the accuracy of the hit rates for each class. Figure 10(a) shows the best results for $K = 1$ neighbors with $a \leq 4$ using Euclidean distance and Fig. 10(b) the best results for $K = 5$ neighbors with $5 \leq a \leq 10$ using the minimum distance. We observe that only the hit accuracy for $a = 1$ does not decay for higher classes. In other words, the prediction of higher classes decreases with a. This is certainly a consequence of the increasing mixture of points in Fig. 6 for higher values of a.

In order to improve the prediction correctness index we consider the naive Bayes [Berrar, 2019]

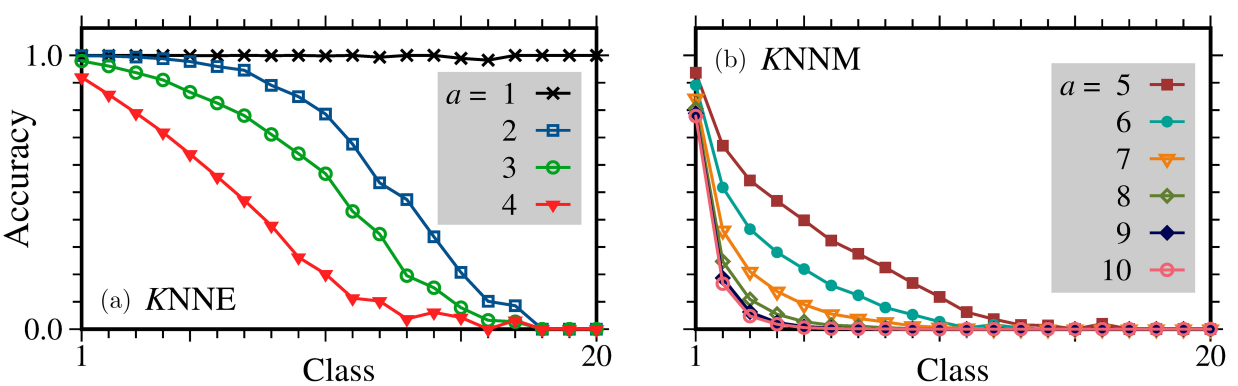

Fig. 10. Prediction accuracy in each class. (a) Best results for $K = 1$ with $a \leq 4$ and (b) best results for $K = 5$ neighbors with $5 \leq a \leq 10$. While only the hit accuracy for $a = 1$ does not decay for higher classes, for all other a values it decays.

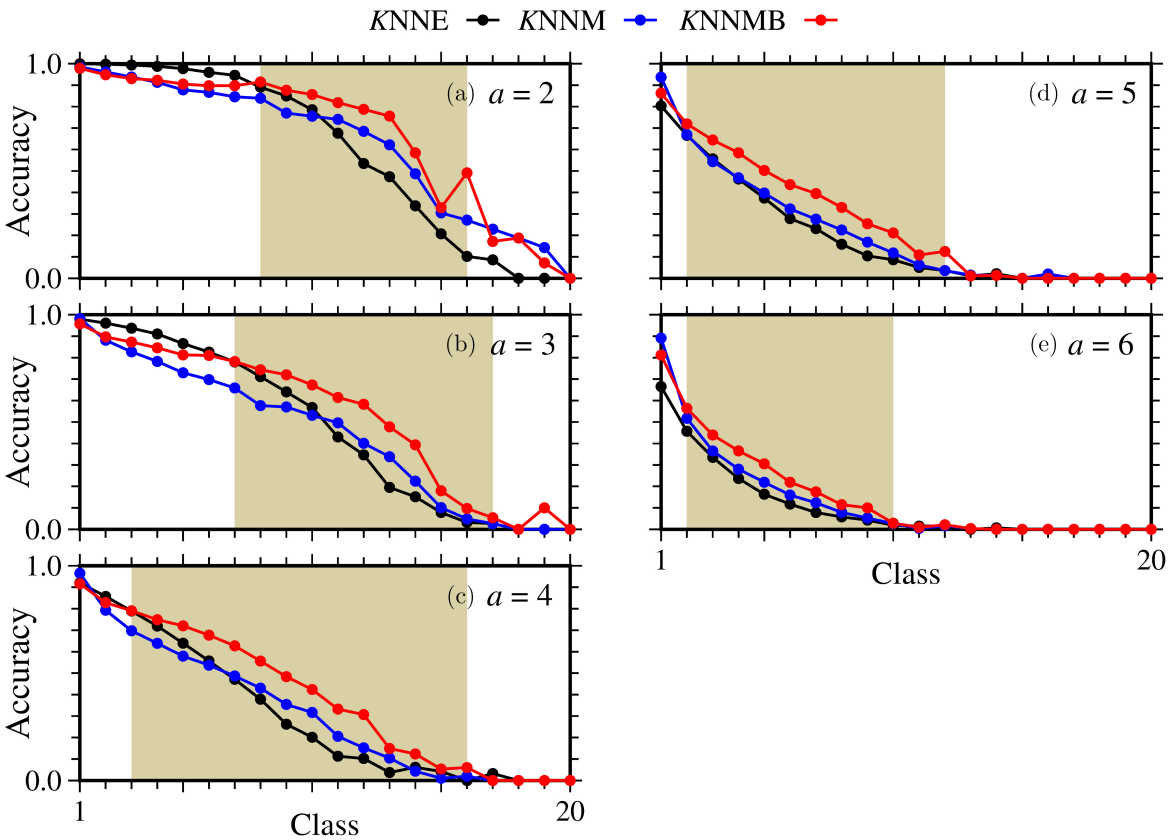

Fig. 11. Prediction accuracy in each class. Comparison between KNN with different metrics and the classification with Bayesian probability. KNNE for $K = 1$, KNNM for $K = 5$ and KNNMB for $K = 5$. The khaki region highlights where KNNMB gives us the best result. Advances (a) $a = 2$, (b) $a = 3$, (c) $a = 4$, (d) $a = 5$ and (e) $a = 6$. The khaki region shifts to the left as the advance increases, as we can see for $a > 4$, with the exception of classes $\Omega = 1$, KNNMB has higher accuracy for lower classes.

probabilistic method, so that we assign a *a posteriori* probability as a function of probabilities *a priori*:

$$P(\Omega\,|\,\mathcal{C}\Omega') = \frac{P(\mathcal{C}\Omega'\,|\,\Omega)P(\Omega)}{P(\mathcal{C}\Omega')}. \tag{13}$$

Let $P(\Omega\,|\,\mathcal{C}\Omega')$ be the conditional probability of the Ω class given that the method KNN classified it as Ω' ($\mathcal{C}\Omega'$ means to classify as Ω'). The above conditional probability is a function of the occurrences of Ω and $\mathcal{C}\Omega'$, as well as the probability $P(\mathcal{C}\Omega'\,|\,\Omega)$ of a class point Ω be classified in Ω'. In this way, the predicted class is assumed to be the one most likely to be *a posteriori* probability. We identify KNNMB as the KNN method with the minimum distance increased by the probability calculated from Eq. (13). Figure 11 shows the accuracy in each class for different versions of the KNN method, and for different advances a. In black, we have KNNE with $K = 1$, in blue KNNM with $K = 5$ and in red the KNN method with Bayesian probability KNNMB and $K = 5$. The khaki colored region identifies the best results for KNNMB, for example, in Fig. 11(c) with $a = 4$, KNNMB is best for $3 \leq \Omega_{n+4} \leq 13$. However, the KNNMB method resulted in accuracies larger that 0.5 in the khaki color range only for the cases $8 \leq \Omega_{n+2} \leq 14$, $7 \leq \Omega_{n+3} \leq 12$, $3 \leq \Omega_{n+4} \leq 8$, $2 \leq \Omega_{n+5} \leq 5$ and $\Omega_{n+6} = 3$. For larger values of a, no significant improvements were observed.

4. Conclusions

In the present work, the alignment of CLVs is used to forecast the duration time inside subsequently visited wings of the proto-Lorenz system. To this end, a classification procedure is used and results using multi-layers perceptrons [Haykin, 1999] are compared to K-nearest neighbors methods [Cover & Hart, 1967], which include Euclidean distances, minimum distances and the naive Bayes probability [Berrar, 2019]. The classification strategy

has distinct classes defined as the number of maxima of the z variable inside one wing. In other words, the classes are related to the times inside a wing.

In all methods, the alignment of the CLVs is shown to be very efficient for the prediction task. Accuracies larger than 0.9 are obtained for predictions of classes in the $a \leq 3$ subsequently visited wings. Accuracies higher than 0.5 are obtained for predictions of classes in the $a \leq 6$ subsequently visited wings. This corresponds to predicting with relatively good precision the time spent inside wings which occur after two full loops in the three-wings attractor. However, larger times (or classes) inside wings become more unpredictable when increasing the value of the advance a. Cases where the accuracies remain below 0.5 are considered unpredictable, for example, the ones obtained for $\Omega_{n+2} \geq 15$, $\Omega_{n+3} \geq 13$, $\Omega_{n+4} \geq 9$, $\Omega_{n+5} \geq 6$ and $\Omega_{n+6} \geq 3$. Nevertheless, accuracies in the predictions are impressive for lower classes. We emphasize that the alignment of CLVs within each regime, when taken in pairs, consists of predictor variables of the time interval that the trajectory will visit the following wings, going beyond just one wing forward in time. Thus, such data provide a better understanding of the employability of CLVs in the context of predictions in chaotic dynamical systems.

Many additional exciting issues emerge when studying predictions in n-wings systems. For example, consider two systems, one with a two-wings attractor and the other with a four-wings attractor, both with the same positive Lyapunov exponent. Is the prediction of two loops in the two-wings system comparable to the prediction of one loop in the four-wings system? Related developments and predictions for a four-wings [Miranda & Stone, 1993] system are in progress. Furthermore, to implement the proposed techniques practically, some comments are needed. The classification procedure and the machine learning technique are easily executed from experimental data. However, a method to determine all unstable and stable directions from experimental data is still unavailable. This is undoubtedly an up-and-coming research issue for future investigations. However, the present technique could be applicable in cases where an alignment occurs between the flow direction and the most unstable direction in a chaotic system. Both quantities can be obtained from the experimental data, and the classification procedure is applied. Nevertheless, forecasting would be less precise in this case.

Acknowledgments

C. C. Felicio thanks CAPES for financial support. E. L. Brugnago is grateful for the financial support from the São Paulo Research Foundation (FAPESP) under Grant Nos. 2021/12232-0 and 2018/03211-6. M. W. Beims was supported by CNPq (Grant No. 310792/2018-5), Brazil. The authors thank Prof. J. A. C. Gallas for suggesting the 3-wings model.

Data Availability

The data that support the findings of this study are available within the article.

References

Beims, M. W. & Gallas, J. A. C. [2016a] "Alignment of Lyapunov vectors: A quantitative criterion to predict catastrophes?" *Sci. Rep.* **6**, 37102.

Beims, M. W. & Gallas, J. A. C. [2016b] "Manifold angles, the concept of self-similarity, and angle-enhanced bifurcation diagrams," *Sci. Rep.* **6**, 18859.

Beims, M. W. & Gallas, J. A. C. [2018] "Predictability of the onset of spiking and bursting in complex chemical reactions," *Phys. Chem. Chem. Phys.* **20**, 18539.

Berrar, D. [2019] "Bayes' theorem and naive Bayes classifier," *Encyclopedia of Bioinformatics and Computational Biology* (Academic Press, Oxford), p. 403.

Brugnago, E. L., Gallas, J. A. C. & Beims, M. W. [2020a] "Machine learning, alignment of covariant Lyapunov vectors, and predictability in Rikitake's geomagnetic dynamo model," *Chaos* **30**, 083106.

Brugnago, E. L., Gallas, J. A. C. & Beims, M. W. [2020b] "Predicting regime changes and durations in Lorenz's atmospheric convection model," *Chaos* **30**, 103109.

Brugnago, E. L., Hild, T. A., Weingärtner, D. & Beims, M. W. [2020c] "Classification strategies in machine learning techniques predicting regime changes and durations in the Lorenz system," *Chaos* **30**, 053101.

Butcher, J. C. [2016] *Runge–Kutta Methods* (John Wiley & Sons, Ltd).

Caschera, F., Bedau, M. A., Buchanan, A., Cawse, J., de Lucrezia, D., Gazzola, G., Hanczyc, M. M. & Packard, N. H. [2011] "Coping with complexity: Machine learning optimization of cell-free protein synthesis," *Biotechnol. Bioengin.* **108**, 2218–2228.

Chua, L. O., Komuro, M. & Matsumoto, T. [1986] "The double scroll family," *IEEE Trans. Circuits Syst.-I* **33**, 1072.

Cover, T. & Hart, P. [1967] "Nearest neighbor pattern classification," *IEEE Trans. Inf. Th.* **13**, 21.

Evans, E., Bhatti, N., Kinney, J., Pann, L., Peña, M., Yang, S. C., Kalnay, E. & Hansen, J. [2004] "Rise undergraduates find that regime changes in Lorenz's model are predictable," *Bull. Am. Meteorol. Soc.* **85**, 520.

Ginelli, F., Poggi, P., Turchi, A., Chaté, H., Livi, R. & Politi, A. [2007] "Characterizing dynamics with covariant Lyapunov vectors," *Phys. Rev. Lett.* **99**, 130601.

Ginelli, F., Chaté, H., Livi, R. & Politi, A. [2013] "Covariant Lyapunov vectors," *J. Phys. A* **46**, 254005.

Goodfellow, I., Bengio, Y. & Courville, A. [2016] *Deep Learning* (MIT Press).

Hansen, R. S. & Hannestad, S. [2014] "Chaotic flavor evolution in an interacting neutrino gas," *Phys. Rev. D* **90**, 025009.

Haykin, S. S. [1999] *Neural Networks*: *A Comprehensive Foundation*, 2nd edition (Prentice Hall PTR).

Inubushi, M., Takehiro, S. & Yamada, M. [2015] "Regeneration cycle and the covariant Lyapunov vectors in a minimal wall turbulence," *Phys. Rev. E* **92**, 023022.

Jousseph, C. A., Kruger, T. S., Manchein, C., Lopes, S. R. & Beims, M. W. [2016] "Weak dissipative effects on trajectories from the edge of basins of attraction," *Physica A* **456**, 68.

Krüger, T. S., Galuzio, P. P., Prado, T. L., Viana, R. L., Szezech, J. D. & Lopes, S. R. [2015] "Mechanism for stickiness suppression during extreme events in Hamiltonian systems," *Phys. Rev. E* **91**, 062903.

Kuptsov, P. V. & Parlitz, U. [2012] "Theory and computation of covariant Lyapunov vectors," *J. Nonlin. Sci.* **22**, 727.

Kuptsov, P. V. & Kuptsova, A. V. [2014] "Predictable nonwandering localization of covariant Lyapunov vectors and cluster synchronization in scale-free networks of chaotic maps," *Phys. Rev. E* **90**, 032901.

Lai, Y. C., Grebogi, C., Yorke, J. A. & Kan, I. [1993] "How often are chaotic saddles nonhyperbolic?" *Nonlinearity* **6**, 779.

Lorenz, E. N. [1963] "Deterministic nonperiodic flow," *J. Atmos. Sci.* **20**, 130.

Lucarini, V., Faranda, D., de Freitas, J. M. M., Holland, M., Kuna, T., Nicol, M., Todd, M., Vaienti, S. *et al.* [2016] *Extremes and Recurrence in Dynamical Systems* (John Wiley & Sons, NY).

Mañé, R. [1987] *Ergodic Theory and Differentiable Dynamics* (Springer-Verlag).

Maxwell, A. E., Warner, T. A. & Fang, F. [2018] "Implementation of machine-learning classification in remote sensing: An applied review," *Int. J. Rem. Sens.* **39**, 2784–2817.

Miranda, R. & Stone, E. [1993] "The proto-Lorenz system," *Phys. Lett. A* **178**, 105.

Monostori, L. [2003] "Ai and machine learning techniques for managing complexity, changes and uncertainties in manufacturing," *Engin. Appl. Artif. Intell.* **16**, 277–291.

Nazarimehr, F., Jafari, S., Golpayegani, S. S. M. R. H., Perc, M. & Sprott, J. C. [2018] "Predicting tipping points of dynamical systems during a period-doubling route to chaos," *Chaos* **28**, 073102.

Nazarimehr, F., Jafari, S., Perc, M. & Sprott, J. C. [2020] "Critical slowing down indicators," *Eur. Phys. Lett.* **132**, 18001.

Norwood, A., Kalnay, E., Ide, K., Yang, S. C. & Wolfe, C. [2013] "Lyapunov, singular and bred vectors in a multi-scale system: An empirical exploration of vectors related to instabilities," *J. Phys. A* **46**, 254021.

Oseledets, V. I. [1968] "A multiplicative ergodic theorem. Characteristic Lyapunov, exponents of dynamical systems," *Trans. Moscow Math. Soc.* **19**, 197–231.

Palatella, L., Trevisan, A. & Rambaldi, S. [2013] "Nonlinear stability of traffic models and the use of Lyapunov vectors for estimating the traffic state," *Phys. Rev. E* **88**, 022901.

Rubinstein, R. Y. & Kroese, D. P. [2013] *The Cross-Entropy Method*: *A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning* (Springer Science & Business Media).

Saiki, Y. & Kobayashi, M. U. [2010] "Numerical identification of nonhyperbolicity of the Lorenz system through Lyapunov vectors," *JSIAM Lett.* **2**, 107.

Sala, M., Manchein, C. & Artuso, R. [2012] "Estimating hyperbolicity of chaotic bidimensional maps," *Int. J. Bifurcation and Chaos* **12**, 1250217-1–11.

Sharafi, N., Timme, M. & Hallerberg, S. [2017] "Critical transitions and perturbation growth directions," *Phys. Rev. E* **96**, 032220.

Sharma, S., Sharma, S. & Athaiya, A. [2020] "Activation functions in neural networks," *Int. J. Eng. Appl. Sci. Technol.* **4**, 310.

Suykens, J. A. K. & Chua, L. O. [1997] "n-double scroll hypercubes in 1-D CNNs," *Int. J. Bifurcation and Chaos* **7**, 1873.

Suykens, J. A. K. & Vandewalle, J. [2005] *Cellular Neural Networks, Multi-Scroll Chaos and Synchronization* (World Scientific, Singapore).

Trevisan, A. & Pancotti, F. [1998] "Periodic orbits, Lyapunov vectors, and singular vectors in the Lorenz system," *J. Atmos. Sci.* **55**, 390.

Truant, D. P. & Morriss, G. P. [2014] "Backward and covariant Lyapunov vectors and exponents for hard-disk systems with a steady heat current," *Phys. Rev. E* **90**, 052907.

Wolf, A., Swift, J. B., Swinney, H. L. & Vastano, J. A. [1985] "Determining Lyapunov exponents from a time series," *Physica D* **16**, 285.

Wolfe, C. L. & Samelson, R. M. [2007] "An efficient method for recovering Lyapunov vectors from singular vectors," *Tellus A* **59**, 355.

Yadav, R. S., Dwivedi, S. & Mittal, A. K. [2005] "Prediction rules for regime changes and length in a new regime for the Lorenz model," *J. Atmos. Sci.* **62**, 2316.

Yalçin, M. E., Suykens, J. A. K., Vandewalle, J. & Özoğuz, S. [2002] "Families of scroll grid attractors," *Int. J. Bifurcation and Chaos* **12**, 23.

Yao, Y., Xiao, Z., Wang, B., Viswanath, B., Zheng, H. & Zhao, B. Y. [2017] *Complexity vs. Performance*: *Empirical Analysis of Machine Learning as a Service* (Association for Computing Machinery).

Yarotsky, D. [2017] "Error bounds for approximations with deep ReLU networks," *Neural Net.* **94**, 103.

Appendix

Determining the Angles Between the Covariant Lyapunov Vectors

The method for obtaining the LEs is well known [Wolf *et al.*, 1985] and involves a Gramm–Schmidt (GS) orthonormalization procedure. Being very efficient to determine the LEs, the GS method destroys the information about the direction of all eigenvectors not related to the largest Lyapunov exponent. To recover the direction of all invariant manifolds, a method was proposed [Ginelli *et al.*, 2013] that involves the forward and backward time integrations.

For the forward time we integrate the equations of motion for variables $(\dot{p}, \dot{q}, \dot{z})$, together with their corresponding linearized equations

$$\begin{pmatrix} \Delta p(t+s) \\ \Delta q(t+s) \\ \Delta z(t+s) \end{pmatrix} = \mathbf{J}(t) \begin{pmatrix} \Delta p(s) \\ \Delta q(s) \\ \Delta z(s) \end{pmatrix}, \qquad \text{(A.1)}$$

where $\mathbf{J}(t)$ is the 3×3 Jacobian matrix of Eqs. (1)–(3). The above linearized equations furnish the time evolution in the tangent space. Choosing an initially orthonormal basis $\mathbf{G}_0 = (\Delta p(0), \Delta q(0), \Delta z(0)) = (\mathbf{a}_0^{(1)}, \mathbf{a}_0^{(2)}, \mathbf{a}_0^{(3)})$, the evolution in tangent space up to a time $t = ns$ $(n = 0, 1, 2, \ldots)$ is achieved from $\tilde{\mathbf{G}}_n = \mathbf{J}_n \mathbf{G}_0$, with $\mathbf{J}_n = \mathbf{J}_{n-1}\mathbf{J}_{n-2}\mathbf{J}_{n-3}\ldots\mathbf{J}_0$. As usual, the matrix $\tilde{\mathbf{G}}_n$ is normalized for each step using the QR decomposition $\tilde{\mathbf{G}}_n = \mathbf{G}_n \mathbf{R}_n$, where $\tilde{\mathbf{R}}_n$ is an upper-triangular matrix with elements $\mathbf{g}_n^{(i)} \cdot \mathbf{a}_n^{(j)}$. Here, $\mathbf{G}_n = (\mathbf{a}_n^{(1)}, \mathbf{a}_n^{(2)}, \mathbf{a}_n^{(3)})$ are the GS vectors before the orthonormalization and $(\mathbf{g}_n^{(1)}, \mathbf{g}_n^{(2)}, \mathbf{g}_n^{(3)})$ the GS vectors after the orthonormalization.

For the backward time, it is crucial to realize that the CLV related to the largest positive LE is now pointing along the stable direction from the forward evolution. Therefore, the CLVs are defined as

$$\mathbf{v}_n^{(1)} = c_n^{(1,1)} \mathbf{g}_n^{(1)}, \qquad \text{(A.2)}$$

$$\mathbf{v}_n^{(2)} = c_n^{(1,2)} \mathbf{g}_n^{(1)} + c_n^{(2,2)} \mathbf{g}_n^{(2)}, \qquad \text{(A.3)}$$

$$\mathbf{v}_n^{(3)} = c_n^{(1,3)} \mathbf{g}_n^{(1)} + c_n^{(2,3)} \mathbf{g}_n^{(2)} + c_n^{(3,3)} \mathbf{g}_n^{(3)}. \qquad \text{(A.4)}$$

The coefficients $c_n^{(i,j)}$ are determined from the dynamics in the tangent space for the backward direction, which in the matrix form is written as $\mathbf{C}_{n-1} = \mathbf{R}_n^{-1} \mathbf{C}_n$, where $\mathbf{R}_n^{-1}$ is the inverse of the $\mathbf{R}_n$ matrix. At last, the angles between the CLVs are computed from $\theta_n^{(i,j)} = \cos^{-1}(\mathbf{v}_n^{(i)} \cdot \mathbf{v}_n^{(j)})$. A schematic representation of the algorithm used to calculate the angles $\theta_n^{(i,j)}$ between the CLVs is displayed in Fig. 12. The circles represent the start (green) and the end (red) of the algorithm. The gray parallelogram indicates data entry and the blue rectangles the processes performed. The yellow diamonds are the decisions made, which have a true (green arrow) or false (red arrow) answer.

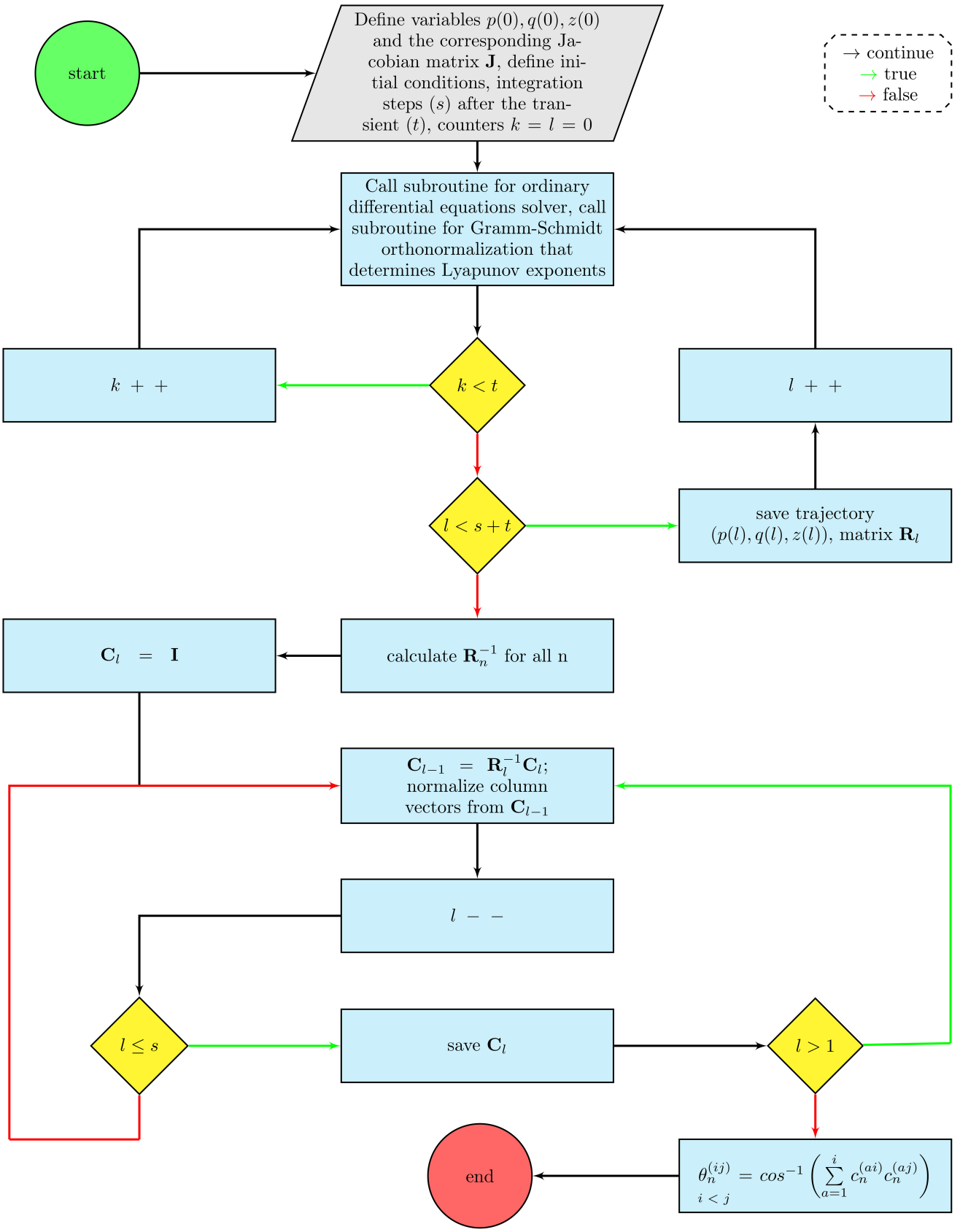

Fig. 12. Block diagram for calculating the angles between the covariant Lyapunov vectors. The algorithm uses the procedure described in [Ginelli *et al.*, 2013].