




Machine learning, alignment of covariant Lyapunov vectors, and predictability in Rikitake's geomagnetic dynamo model

Cite as: Chaos 30, 083106 (2020); <https://doi.org/10.1063/5.0009765>

Submitted: 02 April 2020 . Accepted: 13 July 2020 . Published Online: 03 August 2020

Eduardo L. Brugnago , Jason A. C. Gallas , and Marcus W. Beims 



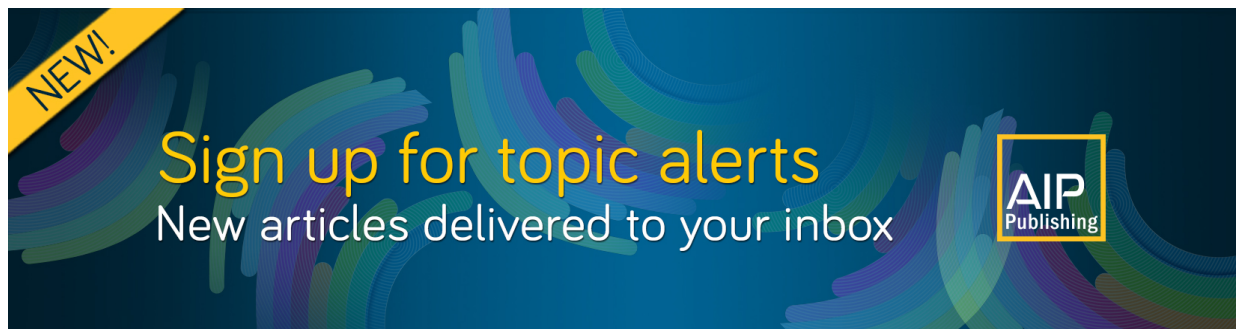
View Online



Export Citation



CrossMark



NEW!

Sign up for topic alerts

New articles delivered to your inbox

AIP Publishing



Machine learning, alignment of covariant Lyapunov vectors, and predictability in Rikitake's geomagnetic dynamo model

Cite as: Chaos 30, 083106 (2020); doi: 10.1063/5.0009765

Submitted: 2 April 2020 · Accepted: 13 July 2020 ·

Published Online: 3 August 2020



View Online



Export Citation



CrossMark

Eduardo L. Brugnago,^{1,2} Jason A. C. Gallas,^{2,3,4} and Marcus W. Beims^{1,2,a)}

AFFILIATIONS

¹Departamento de Física, Universidade Federal do Paraná, 81531-990 Curitiba, Brazil

²Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Str. 38, 01187 Dresden, Germany

³Complexity Sciences Center, 9225 Collins Avenue Suite 1208, Surfside, Florida 33154-3001, USA

⁴Instituto de Altos Estudos da Paraíba, Rua Silvino Lopes 419-2502, 58039-190 João Pessoa, Brazil

^{a)}Author to whom correspondence should be addressed: mbeims@fisica.ufpr.br

ABSTRACT

In this paper, the alignment of covariant Lyapunov vectors is used to train multi-layer perceptron ensembles in order to predict the duration of regimes in chaotic time series of Rikitake's geomagnetic dynamo model. The machine learning procedure reveals the relevance of the alignment of *distinct* covariant Lyapunov vectors for the predictions. To train multi-layer perceptron, we use a classification procedure that associates the number of maxima (or minima) inside regimes of motion with the duration of the corresponding regime. Remarkably accurate predictions are obtained, even for the longest regimes whose duration times are around 17.5 Lyapunov times. We also found long duration regimes with a distinctive statistical behavior, namely, the longest regimes are more likely to occur, a quite unusual behavior. In fact, we observed a largest regime above which no regimes were observed.

Published under license by AIP Publishing. <https://doi.org/10.1063/5.0009765>

It was recently observed^{1,2} that in some chaotic dynamical systems, the alignment of covariant Lyapunov vectors is associated with changes in regimes of motion or the appearance of large peaks along the temporal evolution of physical quantities. In such systems, either long duration regimes or extreme peaks appear after the strong alignment between covariant Lyapunov vectors. Here, such alignments are used to train multi-layer perceptron for the prediction of regime durations in chaotic time series of Rikitake's geomagnetic dynamo model. The strategy is to transform the problem of the determination of time intervals into a classification problem by associating regime duration with the number of maxima (or minima) observed inside each regime. The input data for the training are, therefore, the extreme values of the alignment of covariant Lyapunov vectors, and the number of maxima (or minima) inside the predicted regime. High accuracies are obtained for the predictions, even for the longest regimes that include times up to 17.5 Lyapunov times. Machine learning allows us to identify the most relevant covariant Lyapunov vectors needed for a successful prediction. In addition, we found that the probability to observe the longest duration regimes increases

with the regime duration. This contrasts sharply with what is usually observed in related systems, namely, that extreme events are rare.

I. INTRODUCTION

Forecasting is an extremely challenging open problem in realistic systems like stock markets, laser pulses, climate changes, extreme weathers, critical transitions, and giant ocean waves, among others. Due to its relevance, forecasting has become one of the most active research problems in the present days. Recent advances based on machine learning (ML) techniques open new and interesting ways to attack the issue of forecasting. In the context of prediction in chaotic time series, we mention the use of recurrent neural networks (RNNs)³⁻⁵ and multi-layer perceptron (MLP)⁶⁻⁹ and comparison between several popular methods in ML.¹⁰

It was found recently that in some dynamical systems, there is a relation between the alignment of a specific covariant Lyapunov vector (CLV) along the flow direction and the occurrence of peaks,

or abrupt changes, in the time evolution of physical quantities.^{1,2} The CLVs point along the stable and unstable manifolds in the tangent space of the specific system. For example, it was shown that peaks along the chaotic time series in Chua's and Rössler's systems occur *after* the alignment of the CLV, related to the stable manifold, along the flow direction.¹ The same was observed in the Lorenz system¹¹ for the alignment of the CLV related to the unstable manifold along the flow direction and in complex chemical reactions,² for the alignment of CLVs related to stable and unstable manifolds along the flow direction. Critical transitions can be also predicted in this way.¹² Therefore, the alignment of one or all CLVs along the flow direction may contain intrinsic information about the occurrence of abrupt changes in the dynamics. Such intrinsic information is used here to train MLP ensembles to predict regimes of motion. More specifically, we use the alignment of all CLVs that occur along the chaotic time evolution of variables in Rikitake's model as an input to train MLPs.

The geomagnetic kinematic dynamo model of Rikitake presents two distinct regimes of motion, corresponding to the polarity reversals observed in the Earth magnetic field.^{13–15} Thus, we predict regime duration in Rikitake's model using a classification strategy, based on the number of maxima (or minima) inside a given regime, as explained below. The accuracy in the prediction of the number of maxima (or minima) is astonishing. Using the alignment of CLVs to train the MLPs allows us to elucidate the relevant physical properties behind the success of the prediction. Furthermore, in distinction to the Lorenz system, Rikitake's model presents a very peculiar behavior for the parameters used here. Usually, the probability to obtain regimes with long duration decays exponentially with the duration of the regime. In Rikitake's model, it decays as a power-law and, after that, the probability to obtain the longest duration regimes starts to increase again. This means that the longest regimes become more probable to occur, and there is a bounded interval of regime lengths, above which no regimes were observed. Such extreme-like but not rare events are explained using the alignment of CLVs.

Machine learning allows us to discuss a problematic issue in this context, namely, to know the relevance of the alignment of *distinct* CLVs to describe with success the prediction of peaks, regime changes, and their duration. As far as we know, there is no mathematical argument allowing one to say why in some systems the alignment of just one CLV along the flow direction is enough to make predictions, while in other systems, alignments of more than one, or all CLVs, are needed to do the task. The present work attacks this problematic issue and significantly outperforms recent work combining one CLV and bred vectors expansions to improve predictions.¹¹ Apart from using the ML, the present work analyzes the relevance of the alignment of *all* CLVs to the prediction, and it is thus a kind of higher-dimensional analysis of the CLVs alignment with respect to the aforementioned Lorenz case.¹¹ Furthermore, we mention that the classification procedure was first proposed analyzing the *phase space* of the Lorenz system.⁹ In this case, it relies on the *visual* information that regime changes occur when the variable x passes through zero, something specific of the systems. The present work checks another classification technique, applied to the *tangent space* (namely, to the CLVs) of Rikitake's attractor, being independent of particular properties of the attractor.

Before proceeding, we mention that Rikitake's 1958 ingenious idea of electrically coupling together two disks resulted in the first kinematic dynamo model to display polarity reversals of the Earth magnetic field. A remarkable landmark is that Rikitake observed that *polarity reversals occur at random intervals*, i.e., chaotically, five years before the 1963 celebrated paper of Lorenz. Rikitake studied his model without the benefit of computers.

The paper is divided as follows. Section II briefly presents Rikitake's model together with some dynamical properties that are relevant here. Section III presents details and results using ML to predict regime changes. Section IV summarizes our main results.

II. THE RIKITAKE MODEL

A simple and inspiring mechanical model used to study the reversals of the Earth's magnetic field is a two-disk dynamo system proposed by the geophysicist Rikitake (see Ref. 16). The corresponding non-dimensional equations of motion are

$$\dot{x} = zy - \mu x, \quad (1)$$

$$\dot{y} = (z - \alpha)x - \mu y, \quad (2)$$

$$\dot{z} = 1 - xy, \quad (3)$$

where x and y are the currents of the dynamo coils, z is the angular velocity of the dynamo disks, and $\alpha = \mu(K^2 - 1/K^2)$. The parameters μ and α must be positive and represent the resistive dissipation and the difference of the angular velocities of the two dynamo disks, respectively. We use the parameters combination $\mu = 1.1$ and $K = 2$, which leads to $\alpha = 4.125$. The Lyapunov exponent of the chaotic attractor from this model was analyzed¹⁷ in the parameter space interval $0 < \mu < 5$ and $0 < \alpha < 30$.

A. Covariant Lyapunov vectors

For the parameters considered here, the dynamics of Rikitake's model is chaotic and the Lyapunov spectrum of this chaotic attractor is $(\lambda_1, \lambda_2, \lambda_3) = (0.23, 0, -3.41)$. To each exponent, there is an associated invariant manifold defined in the tangent space: the unstable manifold related to λ_1 , the stable manifold related to λ_3 , and the manifold related to λ_2 , which is tangent to the flow direction since $\lambda_2 = 0$. The correct directions of the invariant manifolds along the attractor can be obtained by calculating the corresponding normalized CLVs: \vec{v}_1 , \vec{v}_2 , and \vec{v}_3 . These vectors point along the unstable, flow, and stable directions, respectively. From the CLVs, it is possible to determine the angles between invariant manifolds from^{1,2}

$$\theta_{ij} = \arccos(\vec{v}_i \cdot \vec{v}_j), \quad \text{for } i < j = 1, 2, 3, \quad i \neq j. \quad (4)$$

Here, CLVs are computed using the methodology explained in Refs. 1, 2, 9, 18–21, and 22. Below, our results will be given in terms of the Lyapunov time $t_L = 1/\lambda_1 \approx 4.35$, related to the positive Lyapunov exponent.

B. The dynamics and regimes

Figure 1 shows a typical chaotic attractor. To obtain it, a numerical integration was started from $(x_0, y_0, z_0) = (0.1, 0.01, 0.001)$,

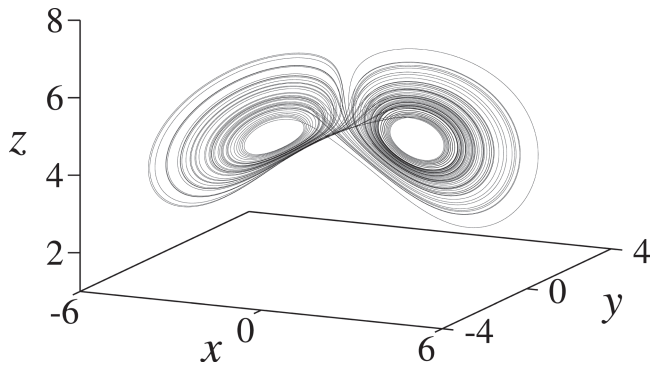


FIG. 1. The two wings of an illustrative chaotic attractor of the Rikitake model, corresponding to the two regimes of oscillation observed for $\mu = 1.1, K = 2$, and $\alpha = 4.125$.

a transient $10^7 h$ ($h = 0.01$) time steps was discarded, the plotting of a total integration time was $4 \times 10^4 h$. The attractor has two wings called *regimes* of motion. An alternative way to visualize the regimes is shown in Fig. 2, which displays the time evolution of the x variable (black curve) from a representative trajectory. The two regimes are separated by the sign of x and represent the reversal of the magnetic field. Inside each regime, one sees oscillations that are directly related to the duration of the individual regimes.

The duration of regimes can vary significantly along a given trajectory and reveals interesting properties of the underlying dynamics. In order to apply the classification strategy adopted in Sec. III, it is appropriate to use the number of maxima (or minima) inside each regime. For $x > 0$, we calculate the maxima inside a regime and for $x < 0$ the minima and use k_n to quantify them. Thus, k_n is the number of local maxima (or minima) inside the regime n . Along the whole simulations, we found 2×10^5 regimes containing $1 \leq k_n \leq 24$ extrema. The duration of the n th regime is denoted by $D_n = \tau_{n+1} - \tau_n$, where τ_n is the time at which the regime starts. So, regime n occurs during the time interval $t \in (\tau_n, \tau_{n+1}]$. In the

example from Fig. 2, we see that, at time τ_{n-1} , a regime with $k_{n-1} = 4$ minima starts and, at time τ_n , a regime with $k_n = 14$ maxima starts.

In Fig. 2, we also plotted in blue the quantity θ_{23} . Red dots indicate maximum and minimum values of θ_{23} before regime changes. As can be seen, at the red dots, the values of θ_{23} are very close to 0 or π , indicating the alignment of the stable manifold along the flow direction. The alignments occur *before* regime changes are observed. These are the same alignments of manifolds along the flow direction that have been used to predict regime changes and durations in other systems.^{1,2,11}

C. Distribution of k_n maxima (or minima)

Figure 3 presents the normalized histogram of the number of regimes having a given k_n inside each regime. As may be seen, until about 20 the number of events decreases when k_n increases. These events obey a power-law decay with exponent -1.55 (see the fitted red curve in Fig. 3). This is close to the exponent -1.5 found nearly three decades ago using a discrete version of Rikitake’s model.²³ For longer duration regimes, k_n lies between 21 and 24, and the probability of their occurrence *increases* with k_n . This means that such longest events are more likely to occur.

The physical reason for such statistically different long events can be explained with the alignment of CLVs. Figure 4 shows an example of a regime with $k_n = 24$ local maxima, together with the angle θ_{12} . In this case, the $k_n = 24$ maxima occur during the time $\sigma_n - \sigma_{n+1} = 12\,946.40 - 12\,871.07 = 75.33$. It can be clearly seen that, in contrast to θ_{23} from Fig. 2, the angle θ_{12} is very close to π at the *beginning* of the long regime (see black square). Thus, there is a strong alignment between the unstable manifold and the flow direction, which occurs at the beginning of the regimes. This strong alignment remains constant for a long time inside the regime and seems to be responsible for the long regime duration.

During all simulations, we consistently found that the alignment of CLVs along the flow direction is responsible for two kinds of predictions: (i) the alignment that predicts regimes changes, marked as red dots in Figs. 2 and 4 and (ii) the alignment leading to very long duration regimes, indicated by the black square in Fig. 4.

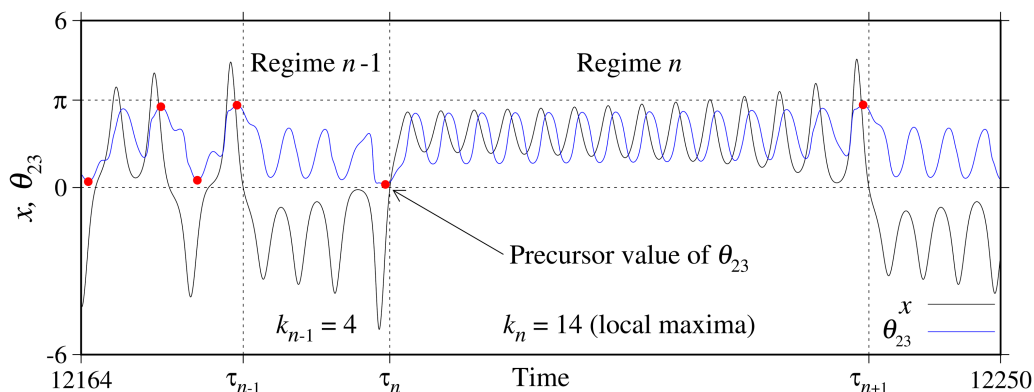


FIG. 2. Temporal evolution of x , black curve, displaying regime changes and duration, plotted together with the angle θ_{23} shown in blue. Red dots indicate positions of the extreme values of θ_{23} preceding regime changes. For reference, horizontal dashed lines indicate values of 0 and π .

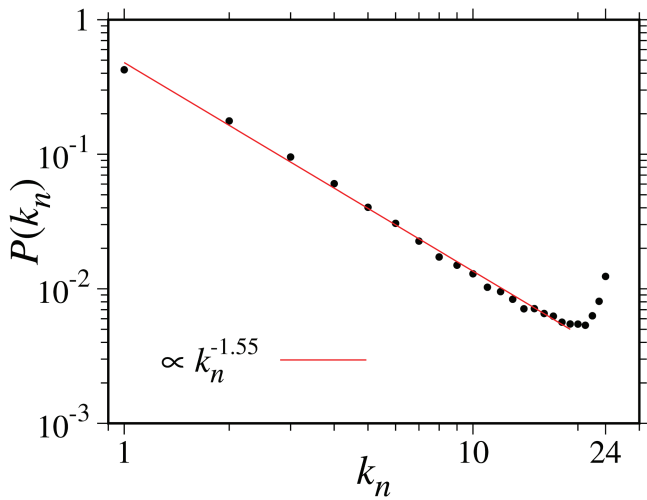


FIG. 3. Normalized distribution of the number of k_n inside each regime.

For clarity, we name the former as *regime change alignment* and the latter as *regime duration alignment*. It has been shown before that the regime change alignment is also related to the duration of the predicted regime,^{1,2,9,11} but the property observed in Fig. 4, the regime duration alignment, has never been observed and has different behavior than the regime change alignment, at least from the statistical point of view. The longest regime duration found has 76.39, which corresponds to 17.56 Lyapunov times.

We observed (not shown here) that results for angles close to π and close 0 are similar so that we superposed them using

$$\theta'_{ij} = \begin{cases} \pi - \theta_{ij}, & \text{if } \theta_{ij} > \frac{\pi}{2}, \\ \theta_{ij}, & \text{in other cases,} \end{cases} \quad (5)$$

and θ'_{ij} remains inside the interval $[0, \pi/2]$, with $i < j = 1, 2, 3$. This simplifies simulations and alignments occur when $\theta'_{ij} \rightarrow 0$. Figure 4

illustrates the behavior of θ_{12} while Fig. 2 shows the behavior of θ_{23} . In fact, the *combined* properties from θ_{12} and θ_{23} are found to be of relevance for the prediction. This will become evident below in Sec. III, when machine learning techniques are applied.

Next, we define the following *extreme* quantities:

$$\begin{aligned} A_n &= \min\{\theta'_{12}(t) : t \in (\tau_n, \tau_{n+1})\}, \\ B_n &= \min\{\theta'_{13}(t) : t \in (\tau_n, \tau_{n+1})\}, \\ C_n &= \min\{\theta'_{23}(t) : t \in (\tau_n, \tau_{n+1})\}. \end{aligned} \quad (6)$$

These quantities are the minima of θ'_{ij} inside regime n . In Fig. 5, we plot C_{n-1} against A_{n-1} with colors indicating the number k_n inside the regimes. As can be observed, when $C_{n-1} \rightarrow 0$, k_n inside the regime increases. Simultaneously, the number of available A_{n-1} decreases since the number of points in Fig. 5 decreases. For values close to $(A_{n-1}, C_{n-1}) \approx (0.08, 0.06)$, we observe a turning point that occurs around a turning line magnified in the inset of Fig. 5. Distinct values of k_n are observed in this magnification. These events have low probability to occur and will not be discussed further.

More important is the gray stripe on the top left of Fig. 5, where $A_{n-1} \rightarrow 0$ and $C_{n-1} \rightarrow 0.25$ (see black arrow). Since $A_{n-1} \rightarrow 0$, these events are related to very strong alignments of the unstable direction along the flow and lead to the longest regimes. In such cases, the angle C_{n-1} can have only one specific angle close to 0.25. This fact provides a physical interpretation for the origin of such long regimes: the strong alignment between the unstable manifold and the flow direction, together with a specific constant direction of the stable manifold along the flow. This occurs for a reasonably long time (roughly 72 time units) in the example shown in Fig. 4, where the whole season duration is 75.3 time units. Additional simulations (not shown) revealed that the largest finite-time Lyapunov exponent (not the asymptotic value 0.23), calculated inside the time window of the regimes, approaches zero for larger values of k_n . This means that the distinct behavior observed in Fig. 4 is related to a transient quasi-stable motion, which leads to such long regimes. In fact, the resulting transient quasi-stable motion looks like a transient limit

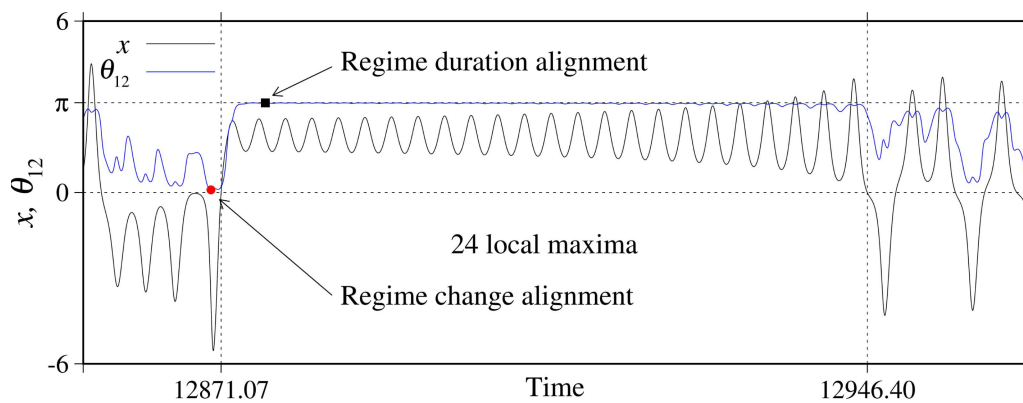


FIG. 4. Temporal evolution of x (black curve) displaying a regime with 24 maxima plotted together with the angle θ_{12} , shown in blue. Red dot indicates the minimum alignment of θ_{12} , which predicts the regime change. The black square indicates the strong alignment at the beginning of the regime.

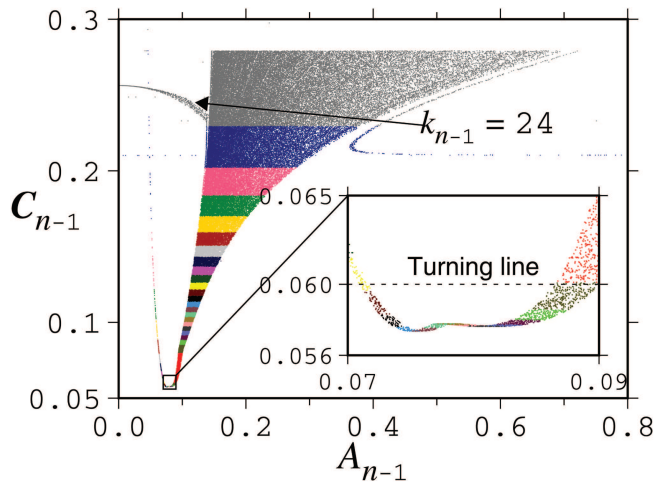


FIG. 5. C_{n-1} vs A_{n-1} with colors indicating $k_n = 1 \rightarrow 24$ inside the regimes.

cycle. To explain this recall that, besides the strong alignment of the unstable manifold along the flow direction, the angle of the stable manifold along the flow direction is also small ($\theta_{23} = 0.25$). Thus, when the trajectory evolves along the flow and the unstable direction, and is simultaneously affected by the stable manifold, it is *trapped* to a kind of transient limit cycle, which leads to the long regimes. The point to mention is that such trapping behavior looks similar to trajectories being trapped around regular tori or almost invariant structures in typical weakly chaotic Hamiltonian systems.^{24,25} In fact, the relation between the transversality of CLVs and dynamical trapping in Hamiltonian systems has been demonstrated.²⁶ The trapping motion, also known as stickiness, increases the probability of a chaotic trajectory staying close to a quasi-regular motion. Something analogous occurs here for a dissipative system.

Furthermore, we realize from Fig. 5 that for one value of C_{n-1} , there is only one corresponding value for k_n (apart from some small exceptional points). For example, for $C_{n-1} = 0.2$, we have only one possible value of $k_n = 3$. However, for one value of A_{n-1} , for example, $A_{n-1} = 0.2$, there can be five possible correspondent values of k_n , namely, 1, 2, 3, 4, and 5. This means that C_{n-1} is a better quantity than A_{n-1} to identify the correct value of k_n . We come to this point again later on, since the machine learning technique is able to recognize the superiority of C_{n-1} to find the correct k_n when compared to the other quantities.

III. MACHINE LEARNING PREDICTION OF k_n

Machine learning is nowadays a popular technique to perform specific tasks without using explicit instructions, relying on inference based on intrinsic patterns. It is a well-documented technique. Here, we follow the main steps described in, e.g., Refs. 6 and 27. MLPs are feedforward artificial neural networks. Initial data are used as input of the network. The propagation of the information goes from the input layer to the output layer. All neurons from one

layer are connected to all neurons from the next layer. The connections have weights, which are adjusted during the training. Once the MLP is trained, decision frontiers are established from which the points are classified. In this way, points are classified according to the regions they are found. In what follows, we briefly review the main steps of the procedure adopted.

A. Artificial neural network

In Fig. 6, we represent one multi-layer perceptron used in this work. Its input layer has neurons whose activation function is the identity function, represented as black rectangles. These neurons establish the connection to the first hidden layer. We use three hidden layers with 50, 100, and 50 neurons, in this order. All neurons from one hidden layer are connected to all neurons from the next hidden layer and just with them. All connections have some weights. For the activation functions in the hidden layers, we use the Rectified Linear Unit (ReLU).²⁸ Furthermore, the propagation of the information is just forward, i.e., from left to right.

In the output layer, an output vector returns one probability for each class. One class of the 24 classes observed is then used according to the largest probability given by the function *softmax*.²⁹ In this way, we relate one point, for example, $(A_{n-1}, B_{n-1}, C_{n-1})$, to the predicted class $k_n^{(pred)}$ (for simplicity, along the text we just use k_n).

The training procedure is nothing more than a mathematical method that minimizes errors by updating the weights in the connections between neurons. The correction of the weights is performed using the backpropagation method with *gradient descent* in weight space.²⁷ Once the MLP is trained, the weights of the connections between neurons do not change anymore, and we proceed with the tests. The accuracy of the method is determined by comparing $k_n^{(pred)}$ with the real class k_n along all regimes.

For the simulation, we use an ensemble of up to 40 distinct MLPs, each one with distinct initial conditions. The MLPs can give different answers, and the pointing of the majority is considered to be the correct one. In the case of a tie, we selected the lowest k_n .

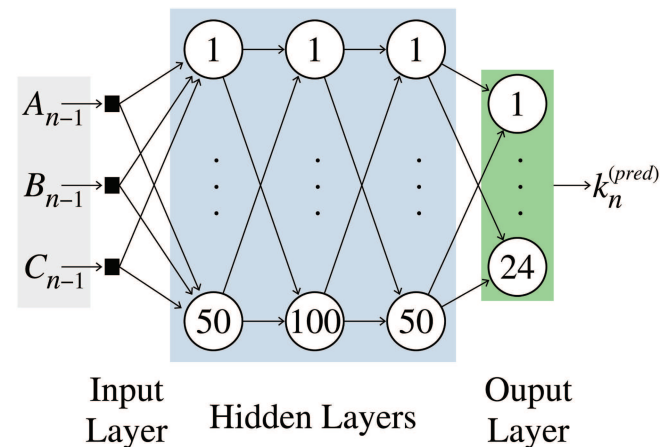


FIG. 6. Schematic representation of one multi-layer perceptron used in the simulations here. See the text.

B. Classification procedure and strategies

From the dynamical point of view, the dynamics of Rikitake's model has some similarities with the dynamics of the Lorenz model. The most relevant similarity is the appearance of two distinct regimes of motion, as illustrated by the pair of wings in Fig. 1. This suggests that a kind of classification procedure as proposed to train the MLPs in the Lorenz systems¹ could be implemented to predict polarity reversals. Here, instead, as inputs for the ML, we use the extreme alignment of CLVs to predict a regime n and the number of k_n . The strategy is to transform the time interval determination problem into a classification problem, namely, to associate the duration of the regime with k_n , the number of maxima (or minima) inside each regime.

More specifically, we use an ensemble of MLPs to classify points in the *space of extreme angles* according to the number k_n found in the regime (which is the predicted regime). Simulations to train the MLPs used a distinct combination of data obtained from the definitions given in Eq. (6). Therefore, we define the following strategies:

$$ABC : [A_{n-1}, B_{n-1}, C_{n-1}] \rightarrow k_n,$$

$$AB : [A_{n-1}, B_{n-1}] \rightarrow k_n,$$

$$AC : [A_{n-1}, C_{n-1}] \rightarrow k_n,$$

$$BC : [B_{n-1}, C_{n-1}] \rightarrow k_n.$$

To obtain strategy AC , for example, we use data like $(A_{n-1}, k_n) = (0.91, 24)$ or $(C_{n-1}, k_n) = (0.113, 14)$, where A_{n-1} (C_{n-1}) is the red dot right before the regime with $k_n = 24$ ($k_n = 14$), as indicated by an arrow in Fig. 4 (Fig. 2). In other words, for strategy AC , we establish a relationship between the pair (A_{n-1}, C_{n-1}) and the label, which identifies the class k_n . All strategies adopted consider the three CLVs. They differ by the angle between the invariant manifolds. Thus, each of the combinations above represents an ensemble of data that leads to a given class k_n .

C. Results

For the training, we generate, from distinct initial conditions, 40 sets with 2.5×10^3 elements containing the information about the alignment of CLVs along the time series. In other words, we have used the alignment of CLVs from 40 distinct time series with 2.5×10^3 values of k_n . Each set is used in one MLP. Thus, the MLPs were trained individually and their combined work is called an ensemble.

In order to test each individual classification strategy, 40 sets with 2.5×10^3 elements were generated from different random initial conditions. This gives a total of 10^5 elements, which are concatenated in one block of data. Since random processes may lead to differences in the learning process, tests and training were realized 10 times for each ensemble. Results are, therefore, presented as averages over 10 repetitions. This ensures the robustness of the method. Input data were multiplied by 10 to facilitate the determination of the classes. Training data were not chosen randomly but according to the way they were generated. In other words, in the numerical integration of Eqs. (1)–(3), data were obtained as time

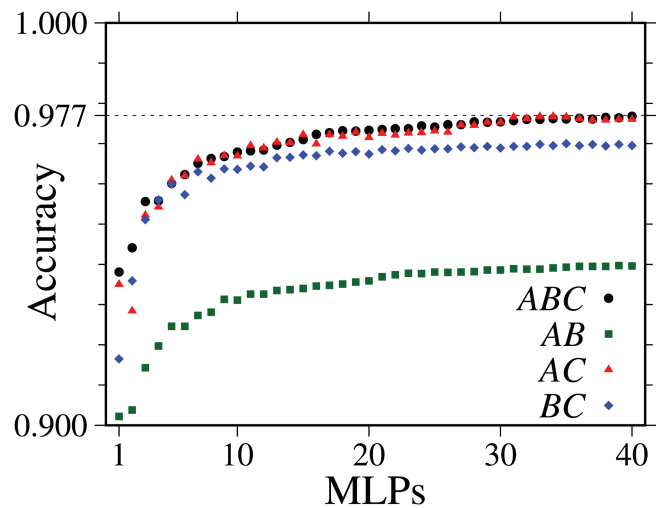


FIG. 7. Accuracies of distinct classes ABC , AB , AC , and BC as a function of the number of MLPs used for the training.

increases and used in this time order as input for the training procedure. Other methods generate the data in the same way, but for the input of the training procedure, the time is chosen randomly. Results are apparently independent of this choice.

As mentioned, as classifiers we use ensembles up to 40 MLPs with three hidden layers and 50, 100, and 50 neurons. For the activation functions, we use the rectified linear unit (ReLU) and choose the optimizer *Adam*³⁰ and loss measure *categorical cross-entropy*.³¹ Here, 500 epochs are used for the training and the connection weights are recalculated in each epoch in accumulations of 40 samples. For the validation, we reserved 20% of training datasets.

Figure 7 presents results for the accuracy of predictions of k_n for the distinct strategies as a function of the number of MLPs. The accuracies increase with the number of MLPs and reach a threshold. Strategies ABC and AC reach a threshold of accuracy 0.977, BC reaches a threshold of accuracy 0.970, and AB a threshold of accuracy 0.940. The optimal number of MLPs used to obtain these accuracies are, respectively, 40, 35, 33, and 39.

We emphasize that the performance of strategies ABC and AC is very similar. This indicates that the information coming from B , namely, the angle between stable and unstable manifolds, is unnecessary for the prediction. Furthermore, besides the strategy ABC , the best accuracies were obtained when information from C was present. Therefore, the angle between the stable manifold and the flow direction is the most relevant for the prediction. This is a confirmation of what can be observed in Fig. 5, where the relation between C_{n-1} and k_n is unique.

Figure 8 presents the accuracies in each class (k_n) for the ensemble with the best result in each strategy (ABC , AB , AC , and BC). It can be seen that all strategies are almost equally accurate to predict values $k_n \leq 12$. At $k_n = 13$, the efficiency of the strategies start to separate from each other and slowly decrease. In the interval $13 \leq k_n \leq 21$, the efficiency of the strategy AB diminishes significantly. Thus, the less accurate predictions from strategies AB and

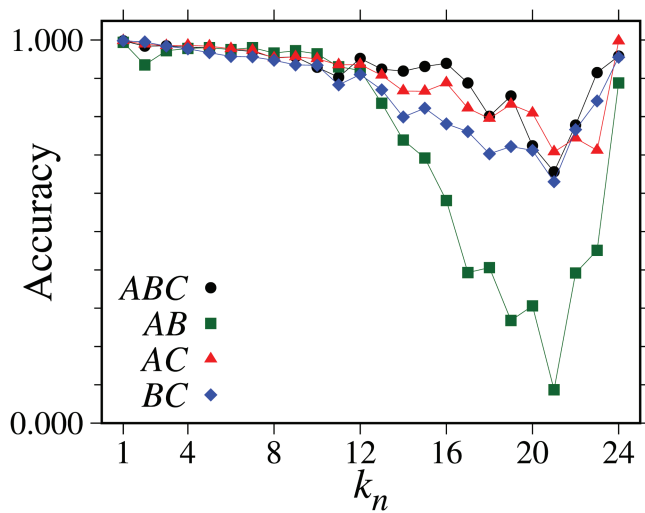


FIG. 8. Accuracies to predict specific values of k_n for the distinct strategies ABC, AB, AC, and BC.

BC, observed in Fig. 7, are related to the not very precise prediction of classes $13 \leq k_n \leq 21$. After that, for $22 \leq k_n \leq 24$, the accuracy of AB and BC increases again. All strategies are equally efficient to predict $k_n = 24$. There is a clear qualitative change in the performance of the distinct strategies for the prediction of $k_n \leq 21$ or $k_n > 21$. When compared to Fig. 3, we see that the first interval is related to a power-law decay, while the second interval is associated with the increase of the probability to obtain regimes with $k_n > 21$. Results from Fig. 8 show that the accuracy of all strategies decreases inside the interval $1 \leq k_n \leq 21$, for which a power-law decay is observed in Fig. 3.

IV. CONCLUSIONS

This paper investigated the ability of machine learning to predict reversals of the Earth's magnetic field according to the coupled dynamo model of Rikitake. Machine learning is found to produce exceptionally reliable predictions, as illustrated in Figs. 7 and 8. Our analysis corroborates previous results that the alignment of CLVs is intrinsically related to abrupt changes along the dynamics of chaotic time series and can be used for prediction tasks. More specifically, this paper shows that the alignment of CLVs can be successfully used to train MLPs ensembles in order to predict the duration of regimes in the chaotic time series of Rikitake's model.

For the classification problem, instead of using the duration of regimes, we use k_n , which is the number of maxima (or minima) inside each regime. Highly accurate predictions are obtained for most k_n , also for the longest ones related to 17.5 Lyapunov times. To train the MLPs, we use input data that combine the information about the extreme alignments of CLVs, following (6), and the corresponding k_n . The combination of aligned CLVs data (the strategies) that provide the best and relatively similar results was ABC and AC. Therefore, information coming from B, namely, from

the angle between stable and unstable manifolds, is unnecessary for the prediction. Furthermore, the best accuracies were obtained when information from C was present, *i.e.*, the angle between the stable manifold and the flow direction.

In general, the accuracy for the prediction of classes ABC, AB, AC, and BC decreases with increasing k_n , until $k_n = 21$ (see Fig. 8). In this interval, the probability to obtain a regime with k_n maxima (or minima) decreases with k_n and obeys a power-law. For $20 < k_n \leq 24$, the performance for prediction of all strategies tested increases again. This is related to a peculiar behavior found in Rikitake's model for the specific parameter combination used here: the probability to obtain an event with given k_n increases with k_n (see Fig. 3). These extreme-like but not rare events become significantly more predictable using the CLVs.

The success in predicting the behavior of regimes of motion in chaotic time series using the alignment of CLVs and machine learning techniques opens new and interesting lines of research in geophysical fluid flows, for example. Using a recent technique, called the objective eulerian coherent structures,³² it is possible to identify, from a single snapshot of the velocity field, regions where particles transported by a flow will converge to and diverge from in a time interval. From the single snapshot of the velocity field, it is possible to determine approximately the alignment of CLVs and, using machine learning, to predict abrupt changes in the fluid dynamics.

ACKNOWLEDGMENTS

This work was supported by the Max Planck Institute for the Physics of Complex Systems, Dresden, in the framework of the Advanced Study Group on Forecasting with Lyapunov vectors. The authors acknowledge support from CNPq (Grant Nos. 304719/2015-3 and 310792/2018-5).

DATA AVAILABILITY

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

REFERENCES

- ¹M. W. Beims and J. A. C. Gallas, "Alignment of Lyapunov vectors: A quantitative criterion to predict catastrophes?," *Sci. Rep.* **6**, 37102 (2016).
- ²M. W. Beims and J. A. C. Gallas, "Predictability of the onset of spiking and bursting in complex chemical reactions," *Phys. Chem. Chem. Phys.* **20**, 18539–18546 (2018).
- ³D. P. Mandic and J. Chambers, *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability* (John Wiley & Sons, Inc., 2001).
- ⁴J. Pathak, B. Hunt, M. Girvan, Z. Lu, and E. Ott, "Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach," *Phys. Rev. Lett.* **120**, 024102 (2018).
- ⁵J. Pathak, Z. Lu, B. R. Hunt, M. Girvan, and E. Ott, "Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data," *Chaos* **27**, 121102 (2017).
- ⁶M. W. Gardner and S. R. Dorling, "Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences," *Atmos. Environ.* **32**, 2627–2636 (1998).

- ⁷C. Ikuta, Y. Uwate, and Y. Nishio, "Chaos glial network connected to multi-layer perceptron for solving two-spiral problem," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems* (IEEE, 2010), pp. 1360–1363.
- ⁸M. Lellep, J. Prexl, M. Linkmann, and B. Eckhardt, "Using machine learning to predict extreme events in the Hénon map," *Chaos* **30**, 013113 (2020).
- ⁹E. L. Brugnago, T. A. Hild, D. Weingaertner, and M. W. Beims, "Classification strategies in machine learning techniques predicting regime changes and durations in the Lorenz system," *Chaos* **30**, 053101 (2020).
- ¹⁰P. Amil, M. C. Soriano, and C. Masoller, "Machine learning algorithms for predicting the amplitude of chaotic laser pulses," *Chaos* **29**, 113111 (2019).
- ¹¹E. L. Brugnago, J. A. C. Gallas, and M. W. Beims, "Predicting regime changes and durations in Lorenz's atmospheric convection model" (submitted) (2020).
- ¹²N. Sharafi, M. Timme, and S. Hallerberg, "Critical transitions and perturbation growth directions," *Phys. Rev. E* **96**, 032220 (2017).
- ¹³T. Rikitake, "Oscillations of a system of disk dynamos," *Proc. Camb. Philos. Soc.* **54**, 89–105 (1958).
- ¹⁴T. Rikitake, *Electromagnetism and the Earth's Interior* (Elsevier, 2012), Vol. 2.
- ¹⁵M. W. McElhinny and P. L. McFadden, *Paleomagnetism: Continents and Oceans* (Elsevier, 1999).
- ¹⁶A. E. Cook and P. H. Roberts, "Rikitake two-disc dynamo system," *Proc. Camb. Philos. Soc.* **68**, 547 (1970).
- ¹⁷R. A. da Silva and P. C. Rech, "A parameter-space analysis of the Rikitake system," *Chin. Phys. Lett.* **30**, 120501 (2013).
- ¹⁸Y. Pomeau, A. Pumir, and P. Pelce, "Intrinsic stochasticity with many degrees of freedom," *J. Stat. Phys.* **37**, 39–49 (1984).
- ¹⁹C. L. Wolfe and R. M. Samelson, "An efficient method for recovering Lyapunov vectors from singular vectors," *Tellus Ser. A* **59**, 355–366 (2007).
- ²⁰A. Norwood, E. Kalnay, K. Ide, S. Yang, and C. Wolfe, "Lyapunov, singular and bred vectors in a multi-scale system an empirical exploration of vectors related to instabilities," *J. Phys. A* **46**, 254021 (2013).
- ²¹F. Ginelli, H. Chaté, R. Livi, and A. Politi, "Covariant Lyapunov vectors," *J. Phys. A* **46**, 254005 (2013).
- ²²P. V. Kuptsov and U. Parlitz, "Theory and computation of covariant Lyapunov vectors," *J. Nonlinear Sci.* **22**, 727–762 (2012).
- ²³M. Seki and K. Ito, "A phase-transition model for geomagnetic polarity reversals," *J. Geomag. Geoelec.* **45**, 79–88 (1993).
- ²⁴G. M. Zaslavsky, "Chaos, fractional kinetics, and anomalous transport," *Phys. Rep.* **371**, 461 (2002).
- ²⁵R. M. da Silva, C. Manchein, M. W. Beims, and E. G. Altmann, "Characterizing weak chaos using time series of Lyapunov exponents," *Phys. Rev. E* **91**, 062907 (2015).
- ²⁶T. S. Kruger, P. P. Galuzio, T. L. Prado, S. R. Lopes, J. D. Szezech, and R. L. Viana, "A mechanism for stickiness, dealing with extreme events," *Phys. Rev. E* **91**, 062903 (2015).
- ²⁷S. S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. (Prentice Hall PTR, 1999).
- ²⁸D. Yarotsky, "Error bounds for approximations with deep ReLU networks," *Neural Netw.* **94**, 103–114 (2017).
- ²⁹J. S. Bridle, "Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition," in *Neurocomputing* (Springer, 1990), pp. 227–236.
- ³⁰I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016).
- ³¹R. Y. Rubinstein and D. P. Kroese, *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation and Machine Learning* (Springer Science & Business Media, 2013).
- ³²G. Haller, "Lagrangian coherent structures," *Annu. Rev. Fluid Mech.* **47**, 137 (2006).