



This is the accepted manuscript made available via CHORUS. The article has been published as:

Predicting amplitude death with machine learning

Rui Xiao, Ling-Wei Kong, Zhong-Kui Sun, and Ying-Cheng Lai

Phys. Rev. E **104**, 014205 — Published 12 July 2021

DOI: [10.1103/PhysRevE.104.014205](https://doi.org/10.1103/PhysRevE.104.014205)

Predicting amplitude death with machine learning

Rui Xiao,^{1,2} Ling-Wei Kong,¹ Zhong-Kui Sun,² and Ying-Cheng Lai^{1,3,*}

¹*School of Electrical, Computer and Energy Engineering,
Arizona State University, Tempe, AZ 85287, USA*

²*Department of applied mathematics, Northwestern Polytechnical University, Xi'an 710129, China*

³*Department of Physics, Arizona State University, Tempe, Arizona 85287, USA*

In nonlinear dynamics, a parameter drift can lead to a sudden and complete cessation of the oscillations of the state variables - the phenomenon of amplitude death. The underlying bifurcation is one at which the system settles into a steady state from chaotic or regular oscillations. As the normal functioning of many physical, biological, and physiological systems hinges on oscillations, amplitude death is undesired. To predict amplitude death in advance of its occurrence based solely on oscillatory time series collected while the system still functions normally is a challenge problem. We exploit machine learning to meet this challenge. In particular, we develop the scheme of “parameter-aware” reservoir computing, where training is conducted for a small number of bifurcation parameter values in the oscillatory regime to enable prediction upon a parameter drift into the regime of amplitude death. We demonstrate successful prediction of amplitude death for three prototypical dynamical systems in which the transition to death is preceded by either chaotic or regular oscillations. Because of the completely data driven nature of the prediction framework, potential applications to real world systems can be anticipated.

I. INTRODUCTION

In nonlinear dynamical systems, amplitude death is a phenomenon in which the oscillatory behaviors of the state variables halt suddenly and completely [1, 2]. A typical route to amplitude death is the drift of a system parameter through a critical point at which a bifurcation from oscillations to a steady state occurs. From the point of view of system functioning, amplitude death is often associated with a catastrophic type of behaviors. For example, in biomedicine, normal physiological conditions are associated with oscillations, while the system’s settling into a steady state is often viewed as the onset of pathological conditions or is associated with death. Because of the relevance of the phenomenon of amplitude death to physical, chemical, biological, and physiological systems [3–7], it has been studied extensively for three decades. Earlier it was found that a parameter mismatch among a network of coupled oscillators can lead to amplitude death [8, 9] and the study was extended to various network settings [10–13]. It was also found that the phenomenon can arise in dynamical systems with time delayed coupling [14–24]. Alternative mechanisms leading to amplitude death include conjugate coupling [25–28], dynamical (time-varying) coupling [29–31], mean-field interaction [32–34], and nonlinear coupling [35].

As amplitude death is undesired in real-world systems, it is of interest and importance to be able to predict its occurrence while the underlying system is in the regime of normal functioning. In the real world, the system equations are most likely unknown, so it is imperative to formulate the prediction problem as one that is based solely on data or time series. These considerations have

motivated the present work. In particular, we ask the following question: suppose the system is in a “normal” regime with oscillations and there is a slow parameter drift due, e.g., to environmental changes, can amplitude death be predicted in advance of its possible occurrence? The problem is extremely challenging because of the requirement to predict the catastrophic behavior based on the presently accessible information which indicates that the system should and would be completely normal by all measures. In fact, if one measures the dynamical variables of the system, the resulting time series are “healthy” in the sense that they all exhibit oscillations, giving no traceable sign that a catastrophic event such as amplitude death would occur upon some amount of parameter drift or a perturbation. To our knowledge, in the literature the only available method is one based on sparse optimization such as compressive sensing [36, 37] where, if the mathematical structure of the system equations is such that they contain only a small number of terms belonging to a power or a Fourier series, the coefficients of these terms can be determined from time series data through sparse optimization. Here, we assume that the governing equations of the system do not have such a simple mathematical structure, rendering inapplicable any of the existing optimization methods. Our solution is to exploit machine learning to develop a model-free, fully data based paradigm to predict amplitude death.

For applications of machine learning in model-free prediction of nonlinear dynamical systems, reservoir computing, a class of recurrent neural networks [38–41], has stood out as an effective paradigm and has received growing attention in recent years [42–58]. Briefly, a reservoir computing machine is a nonlinear dynamical network which, when properly trained with time series data from the target system, becomes effectively a high-dimensional representation of the original system. Starting from the same initial condition, temporal synchronization can be

* Ying-Cheng.Lai@asu.edu

maintained between the reservoir machine and the target system [57], enabling prediction. It is important to note that, in the current literature on model free prediction of chaotic systems with reservoir computing [42–58], training and prediction occur at the same set of parameter values of the target system. However, to predict amplitude death caused by a parameter drift, training and prediction will need to be done at different parameter values. As an essential requirement, it is necessary for the machine to keep track of the variations in the bifurcation parameter, which can be accommodated by designating a specific input channel to the machine. A viable design is to connect this input parameter channel to every node in the reservoir network, which has recently been investigated for predicting transient chaos [59]. Here, we exploit this idea to predict amplitude death.

Figure 1(a) shows our modified reservoir computing structure with an additional input parameter channel, and Fig. 1(b) explains the basic working of this machine learning scheme to predict amplitude death. In particular, let p be the bifurcation parameter of the target system. As p varies, a critical point arises: p_c , where the system exhibits oscillations for $p < p_c$ and the system settles into a steady state (the regime of amplitude death) for $p > p_c$. Training of the reservoir machine is done based on time series taken from a small number of parameter values in the oscillation regime, as indicated by the three vertical dashed blue lines at $p_1 < p_2 < p_3$ ($p_3 < p_c$) in Fig. 1(b). For each of the three parameter values, sufficient training is required in the sense that the machine is able to predict correctly and accurately the oscillatory behavior at the same parameter value for a reasonable amount of time. Suppose that, currently, the system functioning is normal and it operates at the parameter value $p_0 < p_c$, and suppose a parameter drift $\Delta p > 0$ occurs. In the prediction phase, we input the new parameter value $p_0 + \Delta p$ into the reservoir machine through the parameter channel. The prediction is deemed successful if the machine generates oscillations for $p_0 + \Delta p < p_c$ but exhibits amplitude death for $p_0 + \Delta p > p_c$. We demonstrate successful prediction using three prototypical systems of coupled nonlinear oscillators. The broad implication is that the so-articulated reservoir machine can be used to predict the characteristic changes in the system behavior as a result of a parameter drift.

II. BASICS OF RESERVOIR COMPUTING WITH AN INPUT PARAMETER CHANNEL

As shown in Fig. 1 (a), a reservoir computing machine has three components: (1) an input layer that converts an M -dimensional input signal into an N -dimensional signal through an $N \times M$ input weighted matrix W_{in} , (2) a reservoir network (recurrent hidden layer) with N nodes characterized by W_r , an $N \times N$ weighted matrix, and (3) an output layer that maps the N -dimensional vector characterizing the dynamical state of the reser-

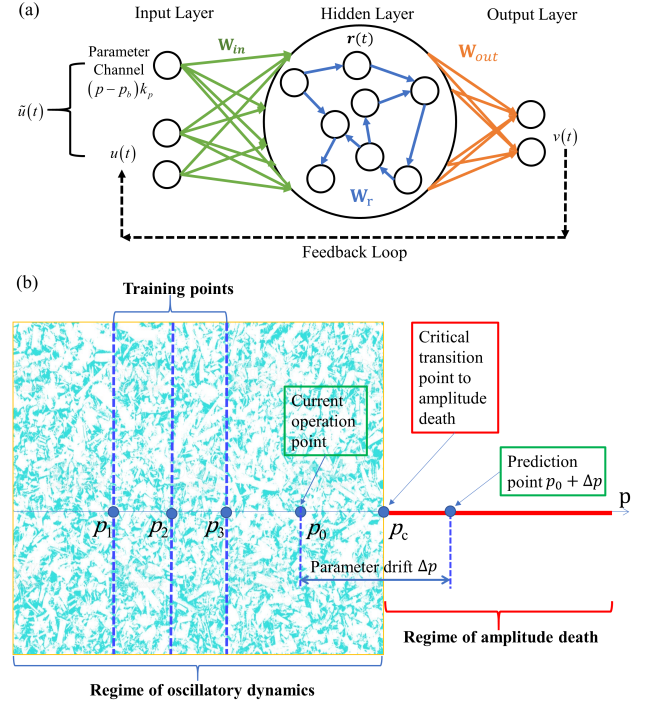


FIG. 1. Modified reservoir-computing scheme and illustration of model-free prediction of amplitude death. (a) Structure of the modified reservoir computing scheme with an additional parameter channel. The reservoir machine has three layers: the input, hidden, and output layers. Input data vector $\tilde{\mathbf{u}}(t)$ consists of the value of the bifurcation parameter and the associated measured time series from the dynamical variables of the target system, $\mathbf{r}(t)$ denotes the dynamical state of the complex network in the hidden layer, and $\mathbf{v}(t)$ denotes the output data vector representing the prediction result. (b) Training of the reservoir computing machine is done in the pre-transition regime for a small number of bifurcation parameter values (as indicated by the three vertical dashed blue lines), where the system generates oscillatory time series. The critical transition to amplitude death occurs at p_c . The target system currently operates at p_0 . Prediction is done for $p = p_0 + \Delta p$, where $\Delta p > 0$ is a parameter drift. Depending on whether the value of $p_0 + \Delta p$ is before or after the transition, a properly trained machine shall be able to predict either an oscillatory behavior or amplitude death, respectively.

voir network into an L -dimensional signal through the $L \times N$ output matrix W_{out} . In a typical application in nonlinear dynamics, the input and output signals are low-dimensional, while the reservoir state is high-dimensional, i.e., $L \sim M \ll N$. The output matrix W_{out} is determined by the training process, while the input weighted matrix W_{in} and the reservoir matrix W_r are initially randomly drawn and then fixed.

In the training phase, the M -dimensional vector $\tilde{\mathbf{u}}(t)$ consisting of the $(M - 1)$ -dimensional time series data $\mathbf{u}(t)$ and the one-dimensional parameter channel is fed into the reservoir as $W_{in} \cdot \tilde{\mathbf{u}}(t)$. The time series data are mapped into the reservoir such that the $(M - 1)$

dimensions are evenly divided among the nodes of the neural network: data from each dimension are input to $N/(M-1)$ nodes. However, the parameter channel is connected to each and every node in the network. The input data through the parameter channel is $(p-p_b)k_p$, where p_b and k_p are two hyperparameters. Under this coupling configuration, the elements of the input matrix $\mathcal{W}_{in}$ are chosen randomly from a uniform distribution in the interval $[-\sigma, \sigma]$. The reservoir is a large, sparse and random network with average degree $\langle k \rangle$, which is described by the matrix $\mathcal{W}_r$, whose values are randomly chosen from the uniform distribution $(0, 1)$. For a given value of the spectral radius, the element values of $\mathcal{W}_r$ are rescaled such that its largest eigenvalue is ρ . The state of each reservoir node at time t is $\mathbf{r}(t) = [r_1(t), r_2(t), \dots, r_N(t)]^T$, and the initial state is set to be $\mathbf{r}(0) = \mathbf{0}$. The reservoir dynamics are described by the following mapping function:

$$\mathbf{r}(t+dt) = (1-\alpha)\mathbf{r}(t) + \alpha \tanh[\mathcal{W}_r \cdot \mathbf{r}(t) + \mathcal{W}_{in} \cdot \tilde{\mathbf{u}}(t)], \quad (1)$$

where $\alpha \in (0, 1]$ is the leakage rate - a hyperparameter. To map the reservoir state to the output layer, it is conventional [49] to define a new vector $\tilde{\mathbf{r}}$ with the same odd row elements as those of $\mathbf{r}$ but with the even row elements as the squared values of the corresponding even row elements of $\mathbf{r}$.

The reservoir computing machine is trained independently for multiple values of the bifurcation parameter. After training with data from one bifurcation parameter value, we reset the time and the initial states to zero, and repeat the training with data from another parameter value. After training is done with the available data from all bifurcation parameter values, we have a vector of multiple input data and multiple recordings of the reservoir state vector $\mathbf{r}(t)$ as well. To find the output matrix through optimization and to express it explicitly, we introduce an input data matrix. In particular, let N_t be the number of training time steps for each training value of the bifurcation parameter. Since we train the reservoir machine using time series obtained from m values of the bifurcation parameter, the total number of training time steps is mN_t . The available data points are stacked in the sequence of these parameter values into a single, $M \times mN_t$ matrix. In this paper, we use $m = 3$. For each input time series, we evolve the state of the reservoir network for $N_\tau \ll N_t$ time steps and disregard these “transient” states from the calculation of the output matrix, so the dimension of the effective input data matrix, denoted as $\tilde{\mathcal{U}}$, is $M \times 3(N_t - N_\tau)$. Note that, the first row of $\tilde{\mathcal{U}}$ consists of three segments: one for each value of the bifurcation parameter, and each of the remaining $(M-1)$ rows consists of three respective segments of the time series of the specific dynamical variables of the target system, which is denoted as $\mathcal{U}$. Since the input bifurcation parameter values are fixed, the effective input data used to calculate the output matrix is $\mathcal{U}$. Likewise, with all the effective input data, the normalized reservoir state matrix, denoted as $\tilde{\mathcal{R}}$, has the dimension

$N \times 3(N_t - N_\tau)$. The output matrix $\mathcal{W}_{out}$ can be calculated using the regression scheme that minimizes the following loss function:

$$\mathcal{L} = \sum_t \left\| \mathcal{U}(t) - \mathcal{W}_{out} \tilde{\mathcal{R}}(t) \right\|^2 + \beta \|\mathcal{W}_{out}\|^2, \quad (2)$$

where β is a small positive regulation constant to prevent over-fitting by imposing penalty on large values of the fitting parameters.

The regularized regression can be described as

$$\mathcal{W}_{out} = \mathcal{U} \cdot \tilde{\mathcal{R}}^T \cdot \left(\tilde{\mathcal{R}} \cdot \tilde{\mathcal{R}}^T + \beta \mathcal{I} \right)^{-1}, \quad (3)$$

where $\mathcal{I}$ is an $N \times N$ identity matrix.

In the prediction phase, the input data vector $\mathbf{u}(t)$ is replaced by the output vector $\mathbf{v}(t)$, so the whole reservoir computing machine becomes a closed loop, self-evolving dynamical system that maps $\mathbf{v}(t)$ to $\mathbf{v}(t+dt)$ according to the following rules:

$$\begin{aligned} \tilde{\mathbf{v}}(t) &= [\mathbf{v}(t); (p-p_b)k_p \cdot \mathcal{B}], \\ \mathbf{r}(t+dt) &= (1-\alpha)\mathbf{r}(t) + \alpha \tanh[\mathcal{W}_r \cdot \mathbf{r}(t) + \mathcal{W}_{in} \cdot \tilde{\mathbf{v}}(t)], \\ \mathbf{v}(t+dt) &= \mathcal{W}_{out} \cdot \tilde{\mathbf{r}}(t+dt), \end{aligned} \quad (4)$$

where $\mathcal{B}$ is an $(N_t - N_\tau)$ -dimensional row vector with all elements equal to one. We use the Bayesian optimization method [54] to determine the optimal values of the hyperparameters of the reservoir computing machine, which are $\langle k \rangle$, σ , ρ , α , β , p_b , and k_p . In particular, for each value of the bifurcation parameter, we ensure that training is completed so that the machine is capable of predicting the state evolution of the target dynamical system for several Lyapunov times, as characterized by below-threshold root-mean squared errors (RMSE) during the time period. Among the three sets of errors from the three values of the bifurcation parameter, we choose the largest RMSE as the criterion for determining the optimal hyperparameter values. After the values of the hyperparameters have been fixed, we train the reservoir machine once again using all the input data to finalize the output matrix $\mathcal{W}_{out}$. The machine so trained is now ready for predicting the system behavior for values of the bifurcation parameter that are different from the three values used for training. Especially, we input the bifurcation parameter value of interest into the parameter channel, and use any of the three available data set to restart or “warm-up” the reservoir network.

Note that, as described, a basic requirement is that the reservoir computing machine be well trained at each of the three selected bifurcation parameter values so that it can possibly be used to predict the system behavior at other parameter values that the machine has not been exposed to. The training data thus consists of the time series from the three parameter values only. When predicting the system dynamics for other parameter values, a small segment of any of the three time series is used as the initial condition to restart or “warm-up” the machine. Since no time series from other parameter values

are assumed to be available, the performance of short-term prediction at these parameter values cannot be assessed.

It is also noteworthy that the training ε parameter values are chosen based on the empirical criterion that they are not too far away from the transition point and are reasonably spaced, and the corresponding time series should be oscillatory. Simulations reveal that the reservoir computing machine is tolerant to relatively small variations in the training parameter values.

III. RESULTS

We demonstrate the ability of our “parameter-aware” reservoir-computing scheme to predict amplitude death in three representative systems: coupled Rössler and Lorenz, coupled Lorenz and coupled Stuart-Landau oscillators. Prior to bifurcating to amplitude death, the system exhibits either periodic or chaotic oscillations. Training is done in the oscillatory parameter regime preceding the bifurcation.

A. Coupled Rössler and Lorenz oscillators

The six-dimensional system of coupled Rössler and Lorenz oscillators is described by [16]

$$\begin{aligned}\dot{x}_1 &= -x_2 - x_3, \\ \dot{x}_2 &= x_1 + ax_2 + \varepsilon(y_2 - x_2), \\ \dot{x}_3 &= b + x_3(x_1 - c), \\ \dot{y}_1 &= \mu(y_2 - y_1), \\ \dot{y}_2 &= -y_1y_3 - y_2 + ry_1 + \varepsilon(x_2 - y_2), \\ \dot{y}_3 &= y_1y_2 - dy_3,\end{aligned}\tag{5}$$

where the parameters are $a = b = 0.1$, $c = 18$, $\mu = 10$, $r = 28$, and $d = 8/3$, and ε is the coupling strength. The Rössler or the Lorenz oscillator, when isolated, is in an oscillatory state. When they are coupled, amplitude death can occur, as exemplified by the bifurcation diagram in Fig. 2(a). The diagram seems to indicate that, as ε increases through a critical point $\varepsilon_c \approx 0.426$, a sudden transition from chaotic oscillations to amplitude death occurs. Figure 2(b) shows a time series of sustained chaotic oscillations for ε below but not too close to “ ε_c ” ($\varepsilon = 0.34$). For ε values above but not too close to “ ε_c ”, the asymptotic state of the system is a stable equilibrium point corresponding to amplitude death, whose occurrence is preceded by transient chaos, as shown in Fig. 2(c) for $\varepsilon = 0.5$.

The fixed points associated with amplitude death in the reservoir prediction agree well with the ones in the real system. For the coupled Rössler-Lorenz system, the fixed points change only slightly with the coupling parameter. Take $\varepsilon = 0.50$ as an example, the original sys-

tem has two fixed points:

$$\begin{aligned}(x_1, x_2, x_3, y_1, y_2, y_3) &= \\ (-4.205, -0.0045, 0.0045, 8.4063, 8.4063, 26.4997)\end{aligned}$$

and

$$\begin{aligned}(x_1, x_2, x_3, y_1, y_2, y_3) &= \\ (4.2003, -0.0073, 0.0073, -8.4064, -8.4064, 26.5004).\end{aligned}$$

The corresponding machine-predicted ones are

$$\begin{aligned}(x_1, x_2, x_3, y_1, y_2, y_3) &= \\ (-4.2041, -0.0065, 0.0045, 8.4038, 8.4037, 26.4977)\end{aligned}$$

and

$$\begin{aligned}(x_1, x_2, x_3, y_1, y_2, y_3) &= \\ (4.1974, -0.0055, 0.0072, -8.4032, -8.4033, 26.5021),\end{aligned}$$

which agree well with the ground truth.

We use a number of initial conditions to generate the bifurcation diagram in Fig. 2(a). For some values of ε , the system has bistability in that there are two coexisting attractors associated with oscillation and amplitude death, respectively. The reason that the transition exemplified in Fig. 2(a) appears rather abrupt is that the system moves into the basin of the amplitude-death attractor as ε changes. In fact, the transition is a gradual process in the sense of multistability. In particular, there exists an interval of ε values about “ ε_c ”, where the coupled system possesses two coexisting attractors: a chaotic attractor and a stable steady-state attractor. Denote this interval as $[\varepsilon_{c1}, \varepsilon_{c2}]$, where $\varepsilon_{c1} < \varepsilon_c < \varepsilon_{c2}$. For $\varepsilon < \varepsilon_{c1}$, chaotic oscillations are the only possible state of the system, i.e., almost every initial condition will generate a trajectory that lands on the chaotic attractor. Likewise, for $\varepsilon > \varepsilon_{c2}$, the only attractor in the system is the stable steady state resulting in amplitude death. For $\varepsilon_{c1} < \varepsilon < \varepsilon_{c2}$, if we choose a large number of random initial conditions, a fraction of them will lead to the chaotic attractor while the remaining to amplitude death. As the value of ε increases from ε_{c1} , the fraction of initial conditions leading to amplitude death, denoted as R_{AD} , gradually increases towards a value (denoted as R_{AD}^* , where $R_{AD}^* < 1$), when ε reaches ε_{c2} . As ε increases through ε_{c2} , R_{AD} increases abruptly from R_{AD}^* to one. At the same time, as ε increases from ε_{c1} to ε_{c2} , the fraction of initial conditions approaching the chaotic attractor will gradually decrease from one to the value $(1 - R_{AD}^*)$ and then suddenly to zero at ε_{c2} .

The dynamical mechanism responsible for the scenario of emergence of amplitude death, as described in the preceding paragraph, has been fairly well understood in nonlinear dynamics [60], which can be attributed to a saddle-node bifurcation at ε_{c1} and a crisis [61] at ε_{c2} . In particular, the saddle-node bifurcation creates a stable steady-state attractor at ε_{c1} . At ε_{c1} , the basin volume of the new steady-state attractor (relative to that of the chaotic attractor) is zero. As ε increases through ε_{c1} ,

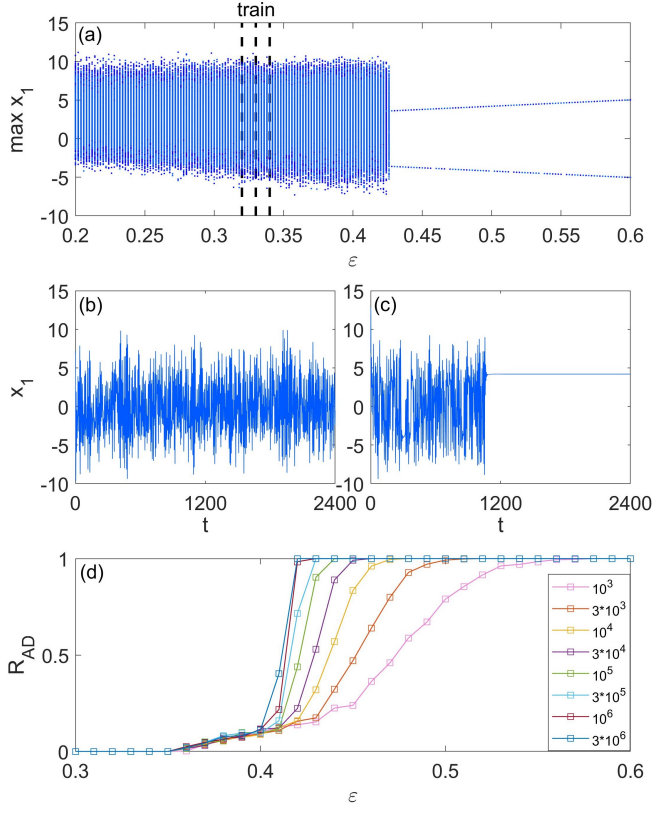


FIG. 2. Scenario of transition to amplitude death. For the coupled Rössler-Lorenz system, (a) a bifurcation diagram with the coupling parameter ϵ from a number of initial conditions, with simulation time $t = 10^5$ and integration time step $dt = 0.01$, (b) chaotic oscillations for $\epsilon = 0.34$, and (c) transient chaos and amplitude death for $\epsilon = 0.50$. (d) Transition to amplitude death revealed by taking an ensemble of random initial conditions, as characterized by R_{AD} , the ratio of the fraction of initial conditions that result in amplitude death. As explained in the text, for an infinite trajectory, as ϵ increases from ϵ_{c1} , R_{AD} increases from zero to a small value and then suddenly to one at ϵ_{c2} . For numerical trajectories of a finite length, the transition would appear gradual but it becomes increasingly sharp as the trajectory length is increased. Here, 1000 randomly initial conditions are chosen from the phase-space region $\{(-10, 10), (-10, -10), (0, 1), (-20, 20), (-20, -20), (0, 50)\}$.

its basin grows, together with the basin boundary. For $\epsilon_{c1} < \epsilon < \epsilon_{c2}$, there are thus two coexisting basins. At ϵ_{c2} , the chaotic attractor collides with the basin boundary and is destroyed - a crisis [61]. For $\epsilon > \epsilon_{c2}$, the only basin of attraction in the system is the one of the stable steady-state attractor, so almost every initial condition leads to amplitude death. The sudden increase of R_{AD} from R_{AD}^* to one holds for infinitely long trajectories. For finite numerical trajectories, the increase would be gradual. As the length of the trajectories increases, the transition would appear increasingly sharp, as exemplified in Fig. 2(d).

Will the reservoir computing machine be able to learn

the essential dynamics of the coupled Rössler-Lorenz system through training in the regime of chaotic oscillations and then to predict the occurrence of amplitude death? To obtain an answer, we train the machine at three values of the coupling parameter: $\epsilon = 0.32, 0.33$, and 0.34 - all in the regime where the system exhibits chaotic oscillations, as indicated by the three vertical dashed lines in Fig. 2(a). For each chosen ϵ value, we use time series of length $t = 800$ (which corresponds to approximately 700 Lyapunov times) to train the machine. A typical prediction result is demonstrated by the time evolution of two dynamical variables, as shown in Figs. 3(a) and 3(b) for $\epsilon = 0.34$, where the machine is capable of replicating the evolution of the dynamical variables for about 6 Lyapunov times ($t \approx 7.5$), after which the prediction still remains meaningful in the sense that the machine generates time series that are statistically indistinguishable from the real time series.

We can now tune the parameter to values beyond the range of training and make predictions. Figures 3(c) and 3(d) show, for $\epsilon = 0.36$, the predicted time series of chaotic oscillations, which indicate correctly that the system is still in the oscillatory regime. Figures 3(e) and 3(f) show, for $\epsilon = 0.50$, that the machine predicts correctly a transient chaotic behavior and the subsequent amplitude death.

The reservoir computing system contains two types of parameters: hyperparameters and free parameters such as the weighted elements of the input matrix $\mathcal{W}_{in}$, the topology and elements of the reservoir network matrix $\mathcal{W}_{res}$. Once the hyperparameter values have been determined and fixed, the prediction performance is robust against variations in the free parameters. Figure 3(g) shows the predicted R_{AD} versus the coupling parameter ϵ for ten random realizations of the reservoir computing system, for $t = 10^3$. The predicted R_{AD} versus ϵ curves are clustered together and agree with the ground truth. For a fixed realization, as the trajectory length increases, the transition becomes sharper, as illustrated in Fig. 3(h) for $t = 10^3$, $t = 3 \times 10^3$, and $t = 10^4$.

For the coupled Rössler-Lorenz system, the transition from oscillatory dynamics to amplitude death is mediated by bistability: there exists a parameter interval $[\epsilon_{c1}, \epsilon_{c2}]$ in which the system possesses two coexisting attractors: a chaotic attractor (oscillations) and a stable steady-state attractor (amplitude death). Numerically, the manifestation is that, in the bistability interval, the curve of R_{AD} versus ϵ is independent of the trajectory length. The estimated values of the end points of the interval are $\epsilon_{c1} \approx 0.36$ and $\epsilon_{c2} \approx 0.40$.

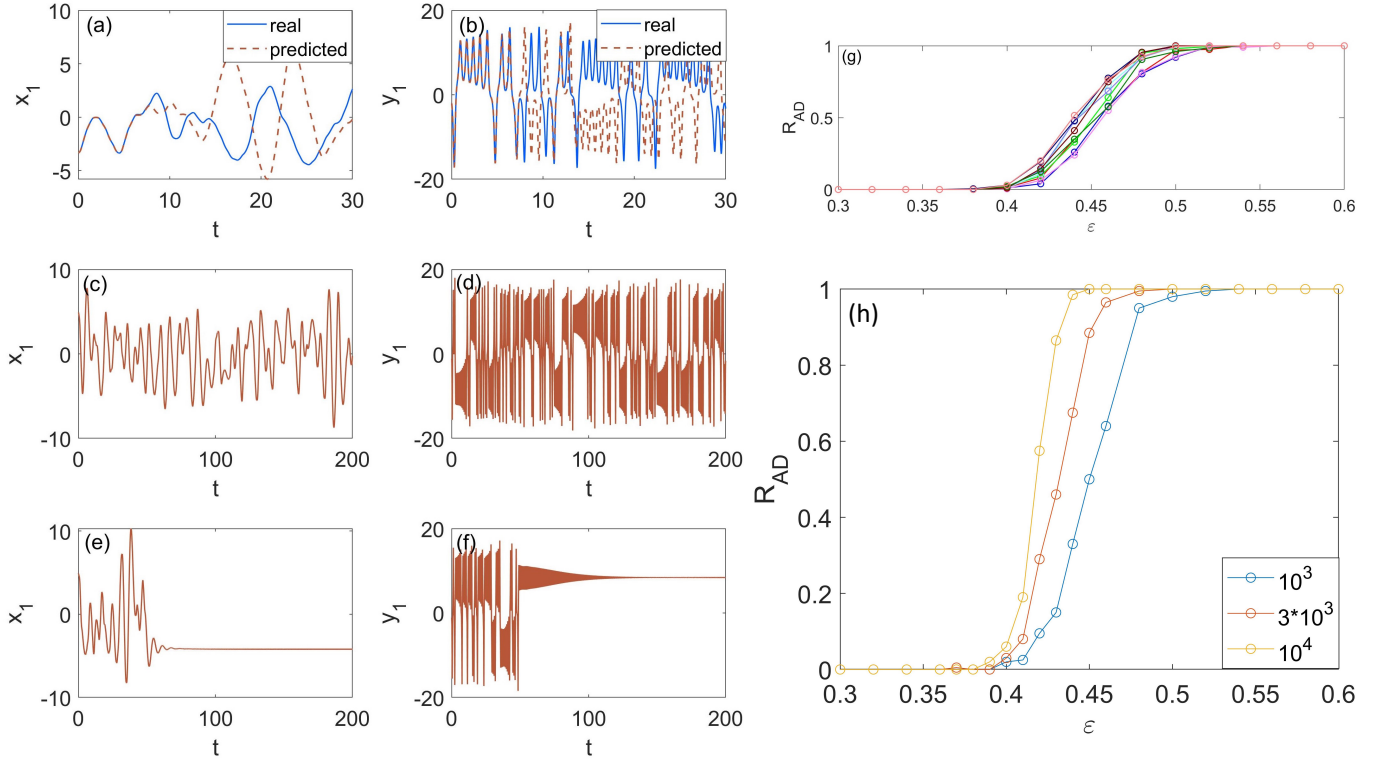


FIG. 3. Prediction of amplitude death in the coupled Rössler-Lorenz system. The training of the reservoir computing machine is done for three values of the coupling parameter: $\varepsilon = 0.32, 0.33$, and 0.34 - all in the regime of chaotic oscillations. Prediction can be made beyond the parameter interval for training through tuning of the input parameter channel to the reservoir network. (a,b) True and predicted chaotic time series $x_1(t)$ and $y_1(t)$ from Eq. (5), for validating the correctness and quality of training. The machine is able to predict the time series exactly for about 6 Lyapunov times beyond which the prediction is still meaningful in the sense that the predicted and true time series are statistically indistinguishable. (c,d) Predicted sustained chaotic time series $x_1(t)$ and $y_1(t)$ for $\varepsilon = 0.36$. (e,f) Predicted transient chaos and amplitude death for $\varepsilon = 0.5$. (g) The predicted ratio R_{AD} versus ε in the regime of transition to amplitude death from ten random realizations of the free parameters of the machine. The predicted results agree with those from direct simulations of the original system Eq. (5). (h) Transition profile for three different values of the trajectory length, demonstrating that the transition becomes sharper with longer trajectory length. Other parameter values of the reservoir system are $N = 1200$, $\langle k \rangle = 20$, $\sigma = 0.1$, $\rho = 0.8833$, $\alpha = 0.312$, $\beta = 1.4303 \times 10^{-5}$, $p_b = 0.5899$, $k_p = 0.1925$, $M = 7$, $L = 6$, $N_t = 80000$, $N_\tau = 50$, and $dt = 0.01$.

B. Coupled Lorenz oscillators with a time delay

We consider the following system of time-delay coupled chaotic Lorenz oscillators:

$$\begin{aligned}
 \dot{x}_1 &= -\mu(x_1 - x_2), \\
 \dot{x}_2 &= -x_1x_3 - x_2 + rx_1 + \varepsilon[y_2(t - \tau) - x_2(t)], \\
 \dot{x}_3 &= x_1x_2 - dx_3, \\
 \dot{y}_1 &= -\mu(y_1 - y_2), \\
 \dot{y}_2 &= -y_1y_3 - y_2 + ry_1 + \varepsilon[x_2(t - \tau) - y_2(t)], \\
 \dot{y}_3 &= y_1y_2 - dy_3,
 \end{aligned} \tag{6}$$

where $\mu = 10$, $r = 28$, and $d = 8/3$ are the parameter values of the classical Lorenz oscillator, ε is the coupling parameter, and τ is the time delay. When uncoupled, each Lorenz oscillator exhibits sustained chaotic oscillations. When coupled, amplitude death can occur in a range of ε values. Figure 4(a) shows a typical bifurcation diagram of the coupled system: the local maxi-

mum values of the dynamical variable $x_1(t)$ versus ε for $\tau = 0.15$, which indicates that amplitude death can occur for $\varepsilon \gtrsim 0.3$. Figures 4(b) and 4(c) show, respectively, the time series from the sustained chaotic oscillating state for $\varepsilon = 0.23$ and transient chaos followed by amplitude death for $\varepsilon = 0.35$. The ratio R_{AD} versus ε obtained directly from the original system is shown in Fig. 4(d). Because of the transient dynamics, the probability for a short trajectory to land on the stable steady-state attractor is small, leading to a slow increase of R_{AD} with ε . As the length of the trajectories increases, there is a higher probability for a trajectory to approach the steady-state attractor, making R_{AD} to switch from zero to one in a more abrupt fashion.

We train the reservoir computing machine at three values of the coupling parameter: $\varepsilon = 0.21, 0.22$, and 0.23 - all in the regime of chaotic oscillations, as indicated by the three vertical dashed lines in Fig. 4(a). Figure 5(a) shows a predicted time series $x_1(t)$ in comparison with

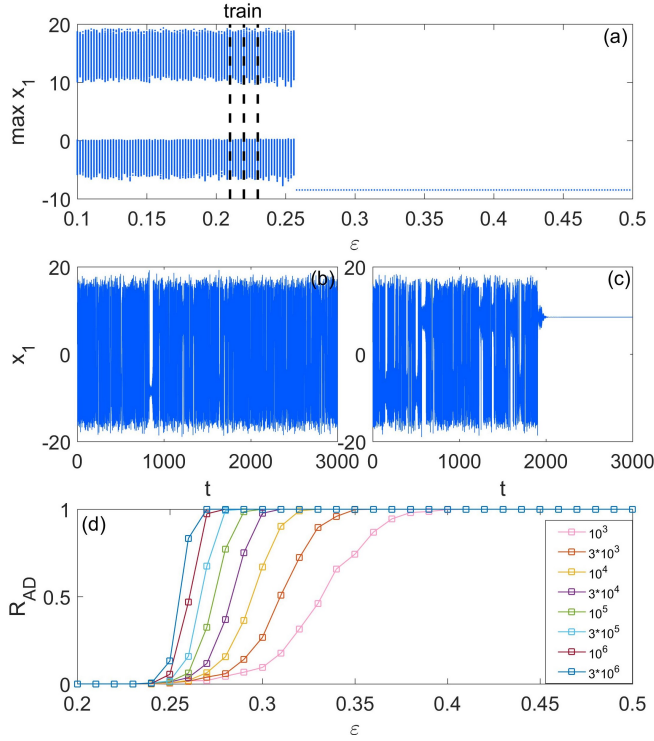


FIG. 4. Transition to amplitude death in the system of a pair of time-delay coupled Lorenz chaotic oscillators. (a) Bifurcation diagram for simulation time $t = 5 \times 10^4$ with $dt = 0.01$. (b) Time series of sustained chaotic oscillation for $\varepsilon = 0.23$. (c) Transient chaos and amplitude death for $\varepsilon = 0.35$. (d) The ratio of the amplitude-death state, R_{AD} , versus ε , for a number of different trajectory lengths. For each value of ε , the ratio is calculated from 1000 trajectories with initial conditions randomly chosen from the phase-space region $-20 \leq (x_1, y_1) \leq 20$, $-20 \leq (x_2, y_2) \leq 20$, and $0 \leq (x_3, y_3) \leq 50$. For all the cases, the time delay is fixed at $\tau = 0.15$.

the true time series for $\varepsilon = 0.23$. The machine is able to predict the state evolution approximately exactly for about 5 Lyapunov times ($t \approx 5$). While the prediction deviates from the true state evolution for $t > 5$, the statistical characteristics of the predicted and true trajectories are indistinguishable. Similar behaviors are obtained with the other two values of the training parameter. These results thus serve to validate the training process. Figure 5(b) shows a predicted time series for $\varepsilon = 0.25$, which indicates that the system is still in the regime of chaotic oscillations. When the value of the input parameter channel to the reservoir system is tuned to $\varepsilon = 0.50$, the machine correctly predicts that the system will settle into the amplitude-death state after a chaotic transient, as shown in Fig. 5(c). Figure 5(d) shows the predicted ratio R_{AD} versus ε for ten random realizations of the reservoir computing system for $t = 10^3$. The predicted transition curves are clustered and consistent with the true curve for this trajectory length. For the coupled Lorenz-Lorenz system, the transition from oscillatory dy-

namics to amplitude death is “abrupt” without any bistability, where the dependence of R_{AD} on ε approaches a step function as the length of the trajectories increases, as shown in Fig. 4(d).

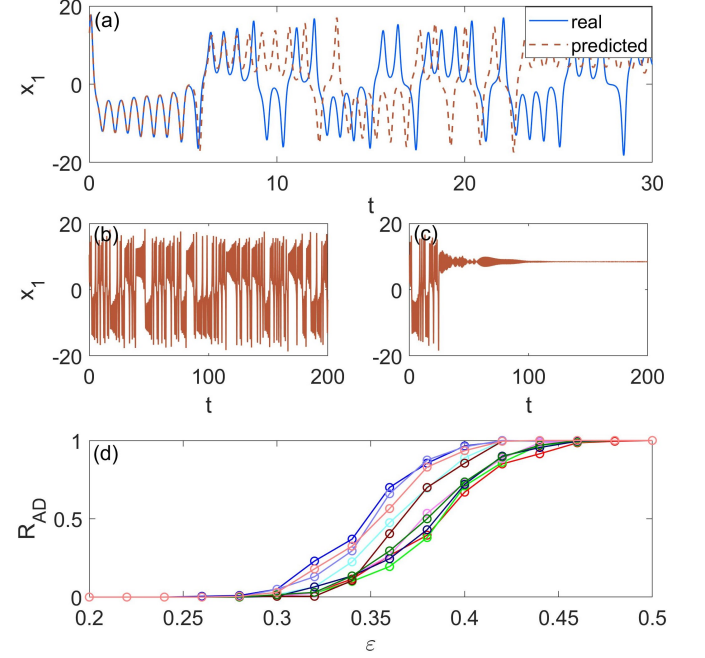


FIG. 5. Predicting amplitude death in the system of a pair of time-delay coupled chaotic Lorenz oscillators. (a) True and predicted chaotic time series $x_1(t)$ from Eq. (6), for validating the training. The machine is able to predict the time series exactly for about 5 Lyapunov times, beyond which the prediction remains meaningful in the statistical sense. (b) Predicted sustained chaotic time series for $\varepsilon = 0.25$. (c) Predicted transient chaos and amplitude death for $\varepsilon = 0.50$. (d) The predicted transition curve (R_{AD} versus ε) from ten random realizations of the machine. The results agree with those from direct simulations of the original system Eq. (6). The time delay is $\tau = 0.15$. The values of other parameters of the reservoir computing system are $N = 1200$, $\langle k \rangle = 20$, $\sigma = 0.1639$, $\rho = 0.5057$, $\alpha = 0.6057$, $\beta = 4.7487 \times 10^{-5}$, $p_b = 0.0061$, $k_p = 0.4956$, $M = 7$, $L = 6$, $N_t = 80000$, $N_\tau = 50$, and $dt = 0.01$.

C. System of coupled Stuart-Landau oscillators

The system is described by:

$$\begin{aligned} \dot{z}_1 &= (1 + i\omega_1 - |z_1|^2) z_1 + \varepsilon (z_2 - z_1), \\ \dot{z}_2 &= (1 + i\omega_2 - |z_2|^2) z_2 + \varepsilon (z_1 - z_2), \end{aligned} \quad (7)$$

where $z_i = x_i + iy_i$ ($i = 1, 2$) are complex variables, ω_1 and ω_2 are parameters. Without coupling, i.e., $\varepsilon = 0$, both oscillators have an unstable fixed point at $z_{1,2}^* = 0$. A previous study [9] revealed that amplitude death occurs for $\varepsilon > 1$ and $\Delta\omega > 2\sqrt{2\varepsilon - 1}$, where $\Delta\omega \equiv$

$|\omega_1 - \omega_2|$ represents the mismatch between the two oscillators. Depending on the amount of mismatch, the system dynamics can be quite different. Figure 6(a) shows a bifurcation diagram of the dynamical variable x_1 . It can be seen that the system is in an oscillatory state for $\varepsilon \in (0, 1)$, and amplitude death occurs for $\varepsilon > 1$. Figures 6(b) and 6(c) show an oscillatory time series for $\varepsilon = 0.99$ and amplitude death (with a short transient) for $\varepsilon = 1.02$, respectively.

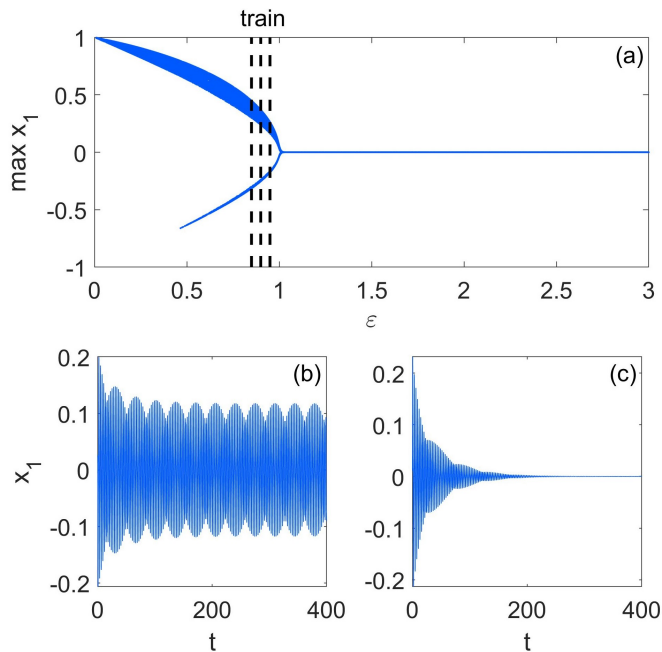


FIG. 6. Oscillations and amplitude death in the system of coupled Stuart-Landau oscillators. (a) A bifurcation diagram with the coupling parameter ε . The system is oscillatory for $0 < \varepsilon < 1$ and amplitude death occurs for $\varepsilon > 1$. (b) An oscillatory time series for $\varepsilon = 0.99$. (c) Amplitude death preceded by a short transient for $\varepsilon = 1.02$. Other parameter values are $\omega_1 = 2.0$ and $\omega_2 = 7.0$.

We train the reservoir computing machine at parameter values $\varepsilon = 0.85, 0.9$ and 0.95 , as indicated by the three vertical dashed lines in Fig. 6(a)). Figure 7(a) and 7(b) show an example of the predicted time series $x_1(t)$ and the difference between the predicted and real time series, respectively. The machine predicts correctly that the system is in an oscillatory state for $\varepsilon < 1$, as exemplified in Fig. 7(c) for $\varepsilon = 0.99$. For $\varepsilon > 1$, the machine predicts successfully amplitude death, as demonstrated in Fig. 7(d) for $\varepsilon = 1.02$. Figure 7(e) shows a machine predicted bifurcation diagram, which agrees with the real diagram in Fig. 6(a). Figure 7(f) shows a distribution of the predicted transition point from 1000 random realizations of the reservoir machine, where all predictions are close to the true transition point $\varepsilon^* = 1$.

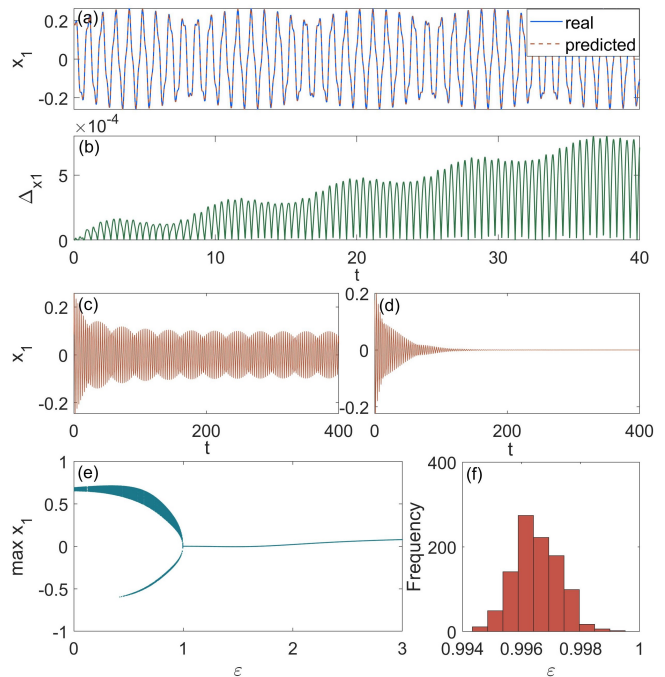


FIG. 7. Predicting amplitude death in the system of coupled Stuart-Landau oscillators. (a) Predicted and real time series for $\varepsilon = 0.95$. (b) The difference between the predicted and real time series for $\varepsilon = 0.95$. (c) Predicted state of oscillation for $\varepsilon = 0.99$. (d) Predicted amplitude death preceded by a transient for $\varepsilon = 1.02$. (e) Predicted bifurcation diagram. (f) The distribution of the predicted transition point to amplitude death from 1000 random realizations of reservoir machine and initial conditions. The parameter values of the machine are $N = 600$, $\langle k \rangle = 8$, $\sigma = 2.0$, $\rho = 0.1893$, $\alpha = 1.0$, $\beta = 10^{-4}$, $p_b = 0.0$, $k_p = 0.5223$, $M = 5$, $L = 4$, $N_t = 16000$, $N_\tau = 50$, and $dt = 0.05$.

IV. DISCUSSION

In biological and physiological contexts, oscillations are essential to maintaining the normal functions of the system, whereas complete cessation of oscillations is associated with severe pathological conditions or even death. In terms of nonlinear dynamics, the occurrence of this amplitude-death phenomenon is the result of a bifurcation to a stable steady state as induced by parameter drifting. It is of interest to forecast amplitude death prior to its actual occurrence when the system is still in a normal oscillation state. Specifically, the prediction problem can be formulated as follows. Suppose a control or bifurcation parameter has been specified and the system is currently in the parameter regime in which the dynamical variables exhibit normal and “healthy” oscillations, where oscillatory time series from a number of parameter values in this regime have been measured. Suppose the bifurcation parameter begins to drift towards a regime that the system has never been in, i.e., no information is available about the system dynamics in the new parameter territory. For a given amount of parameter change,

how do we predict with confidence that amplitude death would occur? If the system equations are known, this prediction problem is trivial, as it can be solved by a simple computational bifurcation analysis. The problem becomes challenging, in fact extremely challenging, when there is no knowledge about the system equations and the only available information is the oscillatory time series collected when the system is in the healthy regime.

We have developed a machine-learning approach to solving the problem of data-based prediction of amplitude death. In particular, exploiting the conventional reservoir computing machine for predicting the state evolution of chaotic systems with a modification to make the machine to be “aware” of the values of the bifurcation parameter of the target system, we have established that amplitude death can be reliably predicted. The structural modification is the addition of a parameter channel through which the machine gains the ability that is essential for predicting the future state of the system: the ability to distinguish the dynamical states associated with different values of bifurcation parameter. We have demonstrated that a proper leaning scheme enables one to fully instill this ability into the machine: training the machine with multiple sets of oscillatory time series data, each from a different value of the bifurcation parameter. We have illustrated our general principle of machine-learning based prediction through three prototypical dynamical systems that represent the paradigmatic systems to study amplitude death. The three systems studied range from the “dynamically” simple coupled Stuart-Landau system to the relatively more complex coupled Lorenz-Lorenz system with a time delay.

In all the systems, prior to the occurrence of amplitude death, there are oscillations: either chaotic or regular. Insofar as the reservoir computing machine has been adequately trained through our learning scheme, any possible future occurrence of amplitude death can be successfully forecasted. In fact, in all three cases, our study reveals that even when the training parameter values are not close to the transition point to amplitude death, the reservoir computing machine is able to predict the critical transition.

Finally, we note that there is a recent work on explaining the success of reservoir computing forecaster of chaotic systems from a mathematical point of view [62], where *linear* activation functions for the artificial neurons in the reservoir network are assumed to make the analysis based on vector auto-regressive averages possible. For general nonlinear activation functions as in Eq. (1), a mathematical understanding of the inner working of reservoir computing is not available at the present.

ACKNOWLEDGEMENT

We thank Dr. Junjie Jiang for discussion. The codes associated with this work are available at <https://github.com/Sherry-ljk/Predict-AD-with-ML.git>. This work was supported by ONR under Grand No. N00014-21-1-2323. RX was supported by the Innovation Foundation for Doctor Dissertation of Northwestern Polytechnical University under Grant No. CX201923.

-
- [1] G. Saxena, A. Prasad, and R. Ramaswamy, “Amplitude death: The emergence of stationarity in coupled nonlinear systems,” *Phys. Rep.* **521**, 205 (2012).
 - [2] A. Koseska, E. Volkov, and J. Kurths, “Oscillation quenching mechanisms: Amplitude vs. oscillation death,” *Phys. Rep.* **531**, 173 (2013).
 - [3] B. Kuntsevich and A. Pisarchik, “Synchronization effects in a dual-wavelength class-B laser with modulated losses,” *Phys. Rev. E* **64**, 046221 (2001).
 - [4] M. Heinrich, T. Dahms, V. Flunkert, S. W. Teitworth, and E. Schöll, “Symmetry-breaking transitions in networks of nonlinear circuit elements,” *New J. Phys.* **12**, 113030 (2010).
 - [5] M. Dolnik and M. Marek, “Extinction of oscillations in forced and coupled reaction cells,” *J. Phys. Chem.* **92**, 2452 (1988).
 - [6] A. Kuznetsov, M. Kærn, and N. Kopell, “Synchrony in a population of hysteresis-based genetic oscillators,” *SIAM J. Appl. Math.* **65**, 392 (2004).
 - [7] A. Koseska, E. Volkov, A. Zaikin, and J. Kurths, “Inherent multistability in arrays of autoinducer coupled genetic oscillators,” *Phys. Rev. E* **75**, 031916 (2007).
 - [8] R. E. Mirollo and S. H. Strogatz, “Amplitude death in an array of limit-cycle oscillators,” *J. Stat. Phys.* **60**, 245 (1990).
 - [9] D. G. Aronson, G. B. Ermentrout, and N. Kopell, “Amplitude response of coupled oscillators,” *Physica D* **41**, 403 (1990).
 - [10] Z. Hou and H. Xin, “Oscillator death on small-world networks,” *Phys. Rev. E* **68**, 055103 (2003).
 - [11] J. Yang, “Transitions to amplitude death in a regular array of nonlinear oscillators,” *Phys. Rev. E* **76**, 016204 (2007).
 - [12] A. Koseska, E. Volkov, and J. Kurths, “Transition from amplitude to oscillation death via turing bifurcation,” *Phys. Rev. Lett.* **111**, 024103 (2013).
 - [13] W. Zou, M. Zhan, and J. Kurths, “Amplitude death in globally coupled oscillators with time-scale diversity,” *Phys. Rev. E* **98**, 062209 (2018).
 - [14] D. R. Reddy, A. Sen, and G. L. Johnston, “Time delay induced death in coupled limit cycle oscillators,” *Phys. Rev. Lett.* **80**, 5109 (1998).
 - [15] F. M. Atay, “Distributed delays facilitate amplitude death of coupled oscillators,” *Phys. Rev. Lett.* **91**, 094101 (2003).
 - [16] A. Prasad, “Amplitude death in coupled chaotic oscillators,” *Phys. Rev. E* **72**, 056204 (2005).

- [17] R. Karnatak, R. Ramaswamy, and A. Prasad, “Amplitude death in the absence of time delays in identical coupled oscillators,” *Phys. Rev. E* **76**, 035201 (2007).
- [18] W. Zou and M. Zhan, “Partial time-delay coupling enlarges death island of coupled oscillators,” *Phys. Rev. E* **80**, 065204 (2009).
- [19] G. Saxena, A. Prasad, and R. Ramaswamy, “Dynamical effects of integrative time-delay coupling,” *Phys. Rev. E* **82**, 017201 (2010).
- [20] J. Zhang and X. Gu, “Stability and bifurcation analysis in the delay-coupled van der pol oscillators,” *Appl. Math. Model.* **34**, 2291 (2010).
- [21] W. Zou, X. Zheng, and M. Zhan, “Insensitive dependence of delay-induced oscillation death on complex networks,” *Chaos* **21**, 023130 (2011).
- [22] Y. Sugitani, K. Konishi, and N. Hara, “Delay-and topology-independent design for inducing amplitude death on networks with time-varying delay connections,” *Phys. Rev. E* **92**, 042928 (2015).
- [23] H. Teki, K. Konishi, and N. Hara, “Amplitude death in a pair of one-dimensional complex Ginzburg-Landau systems coupled by diffusive connections,” *Phys. Rev. E* **95**, 062220 (2017).
- [24] Y. Sugitani and K. Konishi, “Design of coupling parameters for inducing amplitude death in Cartesian product networks of delayed coupled oscillators,” *Phys. Rev. E* **96**, 042216 (2017).
- [25] T. Singla, N. Pawar, and P. Parmananda, “Exploring the dynamics of conjugate coupled Chua circuits: Simulations and experiments,” *Phys. Rev. E* **83**, 026210 (2011).
- [26] W. Zou, D. Senthikumar, A. Koseska, and J. Kurths, “Generalizing the transition from amplitude to oscillation death in coupled oscillators,” *Phys. Rev. E* **88**, 050901 (2013).
- [27] R. Karnatak, R. Ramaswamy, and U. Feudel, “Conjugate coupling in ecosystems: Cross-predation stabilizes food webs,” *Chaos Soli. Frac.* **68**, 48 (2014).
- [28] K. Ponrasu, K. Sathiyadevi, V. Chandrasekar, and M. Lakshmanan, “Conjugate coupling-induced symmetry breaking and quenched oscillations,” *EPL* **124**, 20007 (2018).
- [29] K. Konishi, “Amplitude death induced by dynamic coupling,” *Phys. Rev. E* **68**, 067202 (2003).
- [30] K. Konishi and N. Hara, “Topology-free stability of a steady state in network systems with dynamic connections,” *Phys. Rev. E* **83**, 036204 (2011).
- [31] V. Varshney, G. Saxena, B. Biswal, and A. Prasad, “Oscillation death and revival by coupling with damped harmonic oscillator,” *Chaos* **27**, 093104 (2017).
- [32] T. Banerjee and D. Ghosh, “Transition from amplitude to oscillation death under mean-field diffusive coupling,” *Phys. Rev. E* **89**, 052912 (2014).
- [33] T. Banerjee, P. S. Dutta, and A. Gupta, “Mean-field dispersion-induced spatial synchrony, oscillation and amplitude death, and temporal stability in an ecological model,” *Phys. Rev. E* **91**, 052919 (2015).
- [34] K. Ishibashi and R. Kanamoto, “Oscillation collapse in coupled quantum van der Pol oscillators,” *Phys. Rev. E* **96**, 052210 (2017).
- [35] A. Prasad, M. Dhamala, B. M. Adhikari, and R. Ramaswamy, “Amplitude death in nonlinear oscillators with nonlinear coupling,” *Phys. Rev. E* **81**, 027201 (2010).
- [36] W.-X. Wang, R. Yang, Y.-C. Lai, V. Kovanis, and C. Grebogi, “Predicting catastrophes in nonlinear dynamical systems by compressive sensing,” *Phys. Rev. Lett.* **106**, 154101 (2011).
- [37] W.-X. Wang, Y.-C. Lai, and C. Grebogi, “Data based identification and prediction of nonlinear and complex dynamical systems,” *Phys. Rep.* **644**, 1 (2016).
- [38] H. Jaeger, “The echo state approach to analysing and training recurrent neural networks-with an erratum note,” *German National Research Center for Information Technology GMD Technical Report* **148**, 13 (2001).
- [39] W. Mass, T. Nachtschlaeger, and H. Markram, “Real-time computing without stable states: A new framework for neural computation based on perturbations,” *Neur. Comp.* **14**, 2531 (2002).
- [40] H. Jaeger and H. Haas, “Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication,” *Science* **304**, 78 (2004).
- [41] G. Manjunath and H. Jaeger, “Echo state property linked to an input: Exploring a fundamental characteristic of recurrent neural networks,” *Neur. Comp.* **25**, 671 (2013).
- [42] N. D. Haynes, M. C. Soriano, D. P. Rosin, I. Fischer, and D. J. Gauthier, “Reservoir computing with a single time-delay autonomous Boolean node,” *Phys. Rev. E* **91**, 020801 (2015).
- [43] L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udaltsov, Y. K. Chembo, and M. Jacquot, “High-speed photonic reservoir computing using a time-delay-based architecture: Million words per second classification,” *Phys. Rev. X* **7**, 011015 (2017).
- [44] J. Pathak, Z. Lu, B. Hunt, M. Girvan, and E. Ott, “Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data,” *Chaos* **27**, 121102 (2017).
- [45] Z. Lu, J. Pathak, B. Hunt, M. Girvan, R. Brockett, and E. Ott, “Reservoir observers: Model-free inference of unmeasured variables in chaotic systems,” *Chaos* **27**, 041102 (2017).
- [46] T. Duriez, S. L. Brunton, and B. R. Noack, *Machine Learning Control-Taming Nonlinear Dynamics and Turbulence* (Springer, 2017).
- [47] Z. Lu, B. R. Hunt, and E. Ott, “Attractor reconstruction by machine learning,” *Chaos* **28**, 061104 (2018).
- [48] J. Pathak, A. Wilner, R. Fussell, S. Chandra, B. Hunt, M. Girvan, Z. Lu, and E. Ott, “Hybrid forecasting of chaotic processes: Using machine learning in conjunction with a knowledge-based model,” *Chaos* **28**, 041101 (2018).
- [49] J. 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).
- [50] T. L. Carroll, “Using reservoir computers to distinguish chaotic signals,” *Phys. Rev. E* **98**, 052209 (2018).
- [51] K. Nakai and Y. Saiki, “Machine-learning inference of fluid variables from data using reservoir computing,” *Phys. Rev. E* **98**, 023111 (2018).
- [52] Z. S. Roland and U. Parlitz, “Observing spatio-temporal dynamics of excitable media using reservoir computing,” *Chaos* **28**, 043118 (2018).
- [53] T. Weng, H. Yang, C. Gu, J. Zhang, and M. Small, “Synchronization of chaotic systems and their machine-learning models,” *Phys. Rev. E* **99**, 042203 (2019).
- [54] A. Griffith, A. Pomerance, and D. J. Gauthier, “Forecasting chaotic systems with very low connectivity reservoir computers,” *Chaos* **29**, 123108 (2019).

- [55] J. Jiang and Y.-C. Lai, “Model-free prediction of spatiotemporal dynamical systems with recurrent neural networks: Role of network spectral radius,” *Phys. Rev. Research* **1**, 033056 (2019).
- [56] P. R. Vlachas, J. Pathak, B. R. Hunt, T. P. Sapsis, M. Girvan, E. Ott, and P. Koumoutsakos, “Forecasting of spatio-temporal chaotic dynamics with recurrent neural networks: A comparative study of reservoir computing and backpropagation algorithms,” *arXiv preprint arXiv:1910.05266* (2019).
- [57] H. Fan, J. Jiang, C. Zhang, X. Wang, and Y.-C. Lai, “Long-term prediction of chaotic systems with machine learning,” *Phys. Rev. Research* **2**, 012080 (2020).
- [58] C. Zhang, J. Jiang, S.-X. Qu, and Y.-C. Lai, “Predicting phase and sensing phase coherence in chaotic systems with machine learning,” *Chaos* **30**, 083114 (2020).
- [59] L.-W. Kong, H.-W. Fan, C. Grebogi, and Y.-C. Lai, “Machine learning prediction of critical transition and system collapse,” *Phys. Rev. Research* **3**, 013090 (2021).
- [60] E. Ott, *Chaos in Dynamical Systems*, 2nd ed. (Cambridge University Press, Cambridge, UK, 2002).
- [61] C. Grebogi, E. Ott, and J. A. Yorke, “Crises, sudden changes in chaotic attractors and chaotic transients,” *Physica D* **7**, 181 (1983).
- [62] E. M. Bollt, “On explaining the surprising success of reservoir computing forecaster of chaos? The universal machine learning dynamical system with contrast to VAR and DMD,” *Chaos* **31**, 013108 (2021).