

# Research Article Echo State Property upon Noisy Driving Input

# Junhyuk Woo (),<sup>1</sup> Hyeongmo Kim (),<sup>1,2</sup> Soon Ho Kim (),<sup>1</sup> and Kyungreem Han ()<sup>1</sup>

<sup>1</sup>Laboratory of Computational Neurophysics, Brain Science Institute, Korea Institute of Science and Technology, Seoul 02792, Republic of Korea

<sup>2</sup>Department of Physics and Astronomy, Seoul National University, Seoul 08826, Republic of Korea

Correspondence should be addressed to Kyungreem Han; khan@kist.re.kr

Received 7 October 2023; Revised 1 January 2024; Accepted 9 January 2024; Published 16 February 2024

Academic Editor: Chittaranjan Hens

Copyright © 2024 Junhyuk Woo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The echo state property (ESP) is a key concept for understanding the working principle of the most widely used reservoir computing model, the echo state network (ESN). The ESP is achieved most of the operation time under general conditions, yet the property is lost when a combination of driving input signals and intrinsic reservoir dynamics causes unfavorable conditions for forgetting the initial transient state. A widely used treatment, setting the spectral radius of the weight matrix below the unity, is not sufficient as it may not properly account for the nature of driving inputs. Here, we characterize how noisy driving inputs affect the dynamical properties of an ESN and the empirical evaluation of the ESP. The standard ESN with a hyperbolic tangent activation function is tested using the MNIST handwritten digit datasets at different additive white Gaussian noise levels. The correlations among the neurons, input mapping, and memory capacity of the reservoir nonlinearly decrease with the noise level. These trends agree with the deterioration of the MNIST classification accuracy against noise. In addition, the ESP index for noisy driving input is developed as a tool to help easily assess ESPs in practical applications. Bifurcation analysis explicates how the noise destroys an asymptotical convergence in an ESN and confirms that the proposed index successfully captures the ESP against noise. These results pave the way for developing noise-robust reservoir computing systems, which may promote the validity and utility of reservoir computing for real-world machine learning applications.

## 1. Introduction

Reservoir computing (RC) provides a supervised learning framework for recurrent neural networks (RNNs) that help overcome practical limitations of RNNs such as the convergence issue on training and vanishing/exploding gradient problems [1-7]. Among various types of RC models, the echo state network (ESN) originally developed by Herbert Jaeger in 2001 [5, 8] has been widely used. In an ESN system, a random sparsely connected RNN (acting as a reservoir) maps the time-varying driving inputs into the highdimensional spatiotemporal information in the reservoir, and then a readout layer is used to link the reservoir information and desired output [4, 5, 8-10]. In most cases, the connectivity matrix of the reservoir network is fixed, yet the weights of connections between the reservoir and the readout layer are trainable in a supervised manner. This efficient learning strategy enables high performance at a low

computational cost, compared to conventional RNN settings. This basic working principle is shared with another widely used RC model, the liquid state machine (LSM) which was developed by Maass, Natschlager, and Markram (independently from the ESN) [4, 9, 10]. These RC models have been successfully applied to a wide range of real-world problems including stock market prediction [11, 12], biomedical application [13–17], speech recognition [18, 19], and handwriting recognition [20–23], as well as fundamental problems in physics such as critical transition dynamics [24–27], network link inferring [28, 29], and stochastic/ chaotic time series prediction [7, 30–41].

The echo state property (ESP) was originally defined by Jaeger [5] to grasp the computational capabilities and mechanisms of ESNs, which may be used as a design principle of the reservoirs. Intuitively, a reservoir having this property can be successfully entrained by the driving input to generate high-dimensional, nonlinear, and rich memory properties for computations. In other words, the reservoir dynamics asymptotically washes out its transient initial state induced by the driving input; this is also referred to as "input forgetting" or "state forgetting" [5]. The original ESP [5] claims that an ESN meets the ESP if all state vectors driven by any input sequence from a compact set U asymptotically converge to the same state. While it is a sufficient condition for achieving ESP that the largest singular value of the weight matrix is smaller than unity, it is a necessary condition that the spectral radius of the matrix is smaller than unity. Some following works elaborated on the ESP definition and its sufficient conditions by taking into account the nature of driving input signals. The work by Yildiz et al. [42] considered the distribution of the driving input signals to refine the original ESP concept. This work had important ramifications for the conditions for ESP, a spectral radius greater than unity does not necessarily imply a loss of the ESP, and thus, the commonly used procedure of scaling the spectral radius to under unity to ensure ESPs can be flawed. Meanwhile, the condition on the largest singular value of the weight matrix was found to be too restrictive, leading to poor performance, and alternative conditions were formulated [42, 43]. Subsequently, less restrictive sufficient conditions for the ESP to prevent the fast washing-out problem have been developed [42, 43]. More recently, Manjunath and Jaeger [44] provided an alternative formulation where the ESP is defined concerning a specific input signal rather than a range of possible inputs. For a given input signal, the formulation prescribes the spectral properties of the network weight matrix W that satisfies the ESP. However, work still needs to be done to elucidate when ESP is satisfied given a broad distribution of input signals, e.g., inputs in the presence of noise. Wainrib and Galtier [45] developed a cheap algorithm to establish a local and operational version of the ESP through the computation of the largest Lyapunov exponent. Basterrech [46] presented an empirical analysis of the accuracy and the input mapping of reservoirs. Kubota et al. [47] experimentally demonstrated that cultured neuronal networks can have ESPs to serve as physical reservoir computers.

As stated above, the ESP concept has undergone several refinements since its introduction, yet the validity of widely used literature conditions for the ESP is limited because they do not always properly account for the nature of driving input signals. Of course, the simple technical treatment (i.e., setting the (effective) spectral radius of the weight matrix below the unity) often fails to guarantee the ESP when a combination of driving input and intrinsic reservoir dynamics causes unfavorable conditions for forgetting the transient initial state. For this, the empirical assessment of ESP in practice needs more elaboration in providing sufficient conditions for the ESP, considering that the ESP is a cooperative phenomenon of the intrinsic reservoir dynamics and a set of admissible driving inputs. This study specifically describes numerical simulations and an analytical characterization of the ESP in the presence of noisy driving inputs. The standard ESN model with a tangent hyperbolic activation function is used to examine the dynamical properties and the resulting ESP during the MNIST

handwritten digit classification tasks at different additive white Gaussian noise levels. The effects of the noise on the reservoir dynamics are characterized by various measures including the correlations among the neuronal activities within a reservoir, the mapping of the noisy input to the reservoir, and the memory capacity. These dynamical properties are related to the MNIST classification accuracy and bifurcation dynamics of the reservoir. In addition, the ESP index for noisy driving input is developed based on the work by Gallicchio [48] to help easily assess the property in practical applications. Bifurcation analysis was employed to capture the underlying dynamical properties of the reservoir and to prove the validity of the proposed ESP index.

By way of outline, Section 2 reviews the theoretical frameworks for understanding ESNs and the ESP. Section 3 describes the computational and theoretical methods including the definition of noisy input-driven ESP, bifurcation analysis, and the MNIST classification task. Section 4 begins by describing the changes in the dynamical properties of the reservoir against noise, followed by bifurcation dynamics of the reservoir, and these are related to the deterioration of classification accuracy and the ESP index. Section 5 combines the discussion and conclusions.

# 2. Theoretical Framework

2.1. Standard Echo State Network. The standard echo state network (ESN) model proposed by Herbert Jaeger in 2001 [5] with *N* reservoir units, *K* inputs, and *L* outputs is defined as follows:

$$\begin{aligned} x_{k+1} &= F(x_k, u_{k+1}) \equiv f(Wx_k + W^{\text{in}}u_{k+1} + W^{\text{fb}}y_k), \\ y_k &= g(W^{\text{out}}x_k), \end{aligned}$$
 (1)

where  $x_k \in \mathbb{R}^{N \times 1}$ ,  $u_k \in \mathbb{R}^{K \times 1}$ , and  $y_k \in \mathbb{R}^{L \times 1}$  are the internal, input, and output vectors at time k, respectively,  $W \in \mathbb{R}^{N \times N}$ is the internal weight matrix of the reservoir,  $W^{\text{in}} \in \mathbb{R}^{N \times K}$  is the input matrix,  $W^{\text{fb}} \in \mathbb{R}^{N \times L}$  is the feedback matrix, and  $W^{\text{out}} \in \mathbb{R}^{L \times (N+K)}$  is the output matrix. The state activation function  $f = (f_1, \ldots, f_N)^T$  is a sigmoid function (usually  $f_i$ is a hyperbolic tangent function) applied component-wise with f(0) = 0 and the output activation function is  $g = (g_1, \ldots, g_L)^T$ , where each  $g_i$  is usually the identity or a sigmoid function.

The compactness condition means F is defined on  $X \times U$ , where  $X \subset \mathbb{R}^N$  and  $U \subset \mathbb{R}^K$  are compact sets and F satisfies  $F(x_k, u_{k+1}) \in X$  with  $u_k \in U$ ,  $\forall k \in \mathbb{Z}$ . Let  $U^{-\infty} \coloneqq \{u^{-\infty} = (\dots, u_{-1}, u_0) \mid u_k \in U, \forall k \leq 0\}$  and  $X^{-\infty} \coloneqq \{x^{-\infty} = (\dots, x_{-1}, x_0) \mid x_k \in X, \forall k \leq 0\}$  denote the set of left-infinite input and state vector sequences, respectively. We say  $x^{-\infty}$  is compatible with  $u^{-\infty}$  when  $x_{k+1} = F(u_{k+1}, x_k), \forall k \leq 0$ .

The standard ESN with a hyperbolic tangent activation function without feedback (i.e.,  $f = \tanh$  and  $W^{\text{fb}} = 0$ ) is given by

$$x_{k+1} = F(x_k, u_{k+1}) \equiv \tanh(Wx_k + W^{\text{in}}u_{k+1}).$$
(2)

2.2. Echo State Property for Standard ESNs. The echo state property (ESP) is first defined by Jaeger in his original paper that proposed the ESN [5]: a network  $F: X \times U \longrightarrow X$  with the compactness condition has the ESP with respect to U if for any left infinite input sequence  $u^{-\infty} \in U^{-\infty}$  and any two state vector sequences  $x^{-\infty}$ ,  $y^{-\infty} \in X^{-\infty}$  compatible with  $u^{-\infty}$ , it holds that  $x_0 = y_0$ . It is important to note that the ESP is not a property of a reservoir (or network) per se, but a property of a pair (the reservoir and the set of admissible inputs). The original definition of the ESP takes into account only the range U of admissible inputs, not the probability distribution of the input process. In practice, however, it is this distribution that determines the admissible range of spectral radius for almost all input sequences—which are the practically relevant ones, not the "pathological" ones, which destroy the ESP but occur with zero probability [42]. Therefore, a more useful definition of the ESP is proposed considering the probability of occurrence for admissible inputs [42]: a network F satisfies the echo state property with respect to a stochastic process  $(U_k)_{k\in\mathbb{Z}}$  (where the random variables  $U_k$  take values in a set U) if with probability one,  $\forall u^{-\infty} \in U^{-\infty}$  and  $\forall x^{-\infty}, \forall y^{-\infty} \in X^{-\infty}$  compatible with  $u^{-\infty}$ , it holds that  $x_0 = y_0$ . The definition of ESP for a specific input signal that respects the nature of the expected input signals in more detail is suggested by Manjunath and Jaeger [44]: a network F is said to have the echo state property with respect to an input sequence  $\{u_k\}$  if  $\forall k \in \mathbb{Z}$ ,  $x_{k+1} = F(x_k, u_{k+1})$ , and  $y_{k+1} = F(y_k, u_{k+1})$ , then  $x_k = y_k$  for all  $k \in \mathbb{Z}$ .

The ESP is connected to the spectral properties of the weight matrix W, and some work has been devoted to stating and refining sufficient/necessary conditions for the ESP of the standard ESN [5, 42-44]. A rather restrictive sufficient condition for the ESP was given by Jaeger [5] as  $\sigma^{\max}(W) < 1$ , where  $\sigma^{\max}(W)$  denotes the maximum singular value of W. Since this condition is too restrictive and the input is washed out very fast, it is not commonly used in practice. A less restrictive sufficient condition known to date is that W is diagonally Schur stable [42, 43]. More recently, Manjunath and Jaeger [44] provided an improved formulation of the sufficient condition for the ESP linked to an input:  $\lim_{t\longrightarrow\infty}$  $\sup 1/t \sum_{i=1}^{t} (C(t) - (1 + \ln 2)) I\{C(t) \ge 2\} > \ln(||W||)/2, \text{ where}$  $C(t) = \min(|W^{\text{in}}u(t)|)$ , and *I* is the indicator function that is 1 when its argument is true, and 0 otherwise. In the presence of zero input, a necessary condition for the ESP is  $\rho(W) < 1$ , where  $\rho(W)$  denotes the spectral radius of W [5]. However, for nonzero input signals, the condition is neither sufficient nor necessary for the ESP [42].

Conditions for the ESP used in the literature typically fail to properly account for the effects of driving input signals, often limiting the potentialities of the RC approach. Gallicchio [48] introduced an empirical ESP index that enables analysis of the stability regimes of reservoirs:

$$\text{ESP}_{\text{index}} = \frac{1}{P} \sum_{i=1}^{P} \frac{1}{L - T + 1} \sum_{j=T}^{L} D(\phi(j, 0, 0), \phi(j, 0, z_i)),$$
(3)

where *P* is the number of randomly generated initial states, the first and the last time step used for the calculation of the ESP index are denoted by *L* and *T*, respectively, and D(x, y)represents the Euclidean distance between two vectors *x* and *y*. A process  $\phi$  on a state space *X* is defined as follows:

$$\phi(n,m,x) = \begin{cases} F_{n-1} \circ \dots \circ F_m(x), & \text{if } n > m, \\ x, & \text{if } n = m. \end{cases}$$
(4)

where  $F_k(\cdot) = F(\cdot, u_{k+1})$ .

2.3. Bifurcation Analysis in ESNs. The dynamical properties of ESNs have been investigated using bifurcation theory; it has been considered the standard method to examine the qualitative changes of the dynamical systems such as phase transitions and instabilities [49]. Yildiz et al. [42] used bifurcation analysis to prove that the spectral radius condition  $\rho(W) < 1$  is not a necessary condition for ESP under the zero input environment [42]. With zero input  $u_k = 0$ , the origin becomes a trivial fixed point with a sigmoid activation function. Since stationary origin state  $x_k = 0$  is compatible with zero input, the problem is reduced to the stability and existence problem of additional fixed points, which can be analyzed with bifurcation theory in an autonomous dynamical system. In two-dimensional systems consisting of two nodes, the weight matrix should be in the stable triangle region  $\Delta = \{W \in \mathbb{R}^{2 \times 2} || \operatorname{tr}(W) | -1 < \operatorname{det}(W)$ <1} [50]. For convenience, they assumed the one component in the weight matrix zero  $w_{11} = 0$ . By fixing  $(\det(W), \operatorname{tr}(W))$ = (c, c + 1), where c is positive, new fixed points emerge if  $|w_{12}| < 1/\sqrt{c}$  and disappear if  $|w_{12}| > 1/\sqrt{c}$ . It is called degenerate bifurcation which generates two more fixed points away from the origin in a two-dimensional system. Moreover, increasing determinant det (W) with fixed trace tr (W) induces additional nontrivial fixed points away from the origin. This bifurcation analysis ensures the existence of nontrivial fixed points-at least four-away from the origin which can be either asymptotically stable or unstable. Stabilities of these fixed points in reservoir dynamics should be separated from stabilities in bifurcation analysis.

Using computational simulations, one can easily check the stabilities of these points by observing the basin of points. Yildiz et al. [42] showed the case that there exist three asymptotically stable fixed points including the origin and two fixed points from the degenerate bifurcation and two saddle points from the pitchfork bifurcation. These results can be extended to higher dimensions using appropriate block matrices. However, the bifurcations of the higher dimensional nonautonomous dynamical systems are difficult to solve analytically: in higher dimensions, ESNs can exhibit the Neimar-Sacker bifurcation, indicating fixed points other than the origin can exist at  $\rho(W) < 1$ , and the input noise significantly affects the dynamical properties of reservoirs and in turn the training and inference performance of ESNs. The mathematical theory called stochastic bifurcation theory has dealt with this type of dynamical systems perturbed by additive stochastic noise [51]. However, artificial neural networks including ESNs allow the input signal to be mapped into networks through a nonlinear sigmoid-like activation function. This nonlinearity applied to stochastic terms makes analytic approaches much more challenging due to the difficulty of linearly separating the deterministic and stochastic terms. To circumvent these problems, we resort to numerical simulations to examine the behavior of the stochastic and nonlinear dynamical systems.

#### 3. Methods

We refine the literature definitions of ESP [5, 42–44] and ESP index [48] to adequately reflect the effects of noisy driving input  $u = \tilde{u} + \xi$ , where  $\tilde{u}$  is the input sequence from a compact set *U* and  $\xi$  the additive white Gaussian noise (AWGN). We also provide a sufficient condition of ESP considering the noise.

#### 3.1. Echo State Property upon Noisy Driving Input

Definition 1 (ESP with respect to noisy driving input). A network  $F: X \times U \longrightarrow X$  (with the compactness condition) has the echo state property of tolerance  $\epsilon$  and confidence  $\gamma$  with respect to U and noise process  $\xi$ : if for any left infinite noisy input sequence  $u^{-\infty} = \tilde{u}^{-\infty} + \xi^{-\infty}$ ,  $u^{-\infty} \in U^{-\infty}$ , and any two state vector sequences  $x^{-\infty}$ ,  $y^{-\infty} \in X^{-\infty}$  compatible with different realizations of  $u^{-\infty}$ , it holds that  $P(||x_0 - y_0|| \le \epsilon) \ge \gamma$ .

Here,  $\|\cdot\|$  denotes the norm defined on the state space *X*. If  $\epsilon = 0$ ,  $\gamma = 1$ , and  $\xi^{-\infty} = (..., 0, 0)$ , the definition becomes the original one by Jaeger [5].

A sufficient condition for the ESP with respect to noisy driving input is given by following proposition—sufficient condition for ESP with respect to noisy driving input. For a standard ESN model in equation (2), if the following conditions are satisfied:

- (i)  $\xi$  and  $\eta$  are different realizations of the noise process;
- (ii)  $\sigma^{\max}(W) < 1;$
- (iii)  $P(\sigma^{\max}(W^{in})\sum_{k=1}^{\infty} [\sigma^{\max}(W)]^k \|\xi_{-k} \eta_{-k}\| \le \epsilon) \ge \gamma$ ,

the ESN has ESP of tolerance  $\epsilon$  and confidence  $\gamma$  with respect to noisy driving input. The proof of the proposition is presented in Appendix A. When the probability distribution of the noise is given, the regime of spectral properties of Wand  $W^{\text{in}}$  where ESP holds can be calculated using the given proposition.

Finally, we define the ESP index noisy driving input to empirically assess the property, which leads to:

$$\text{ESP}_{\text{index}} = \frac{1}{P} \sum_{i=1}^{P} \frac{1}{L - T + 1} \sum_{j=T}^{L} D\left(\tilde{\phi}(j, 0, 0), \phi(j, 0, z_i)\right), \quad (5)$$

where  $\tilde{\phi}$  is a process defined using the  $\tilde{F}_k(\cdot) = F(\cdot, \tilde{u}_{k+1})$  instead of  $F_k$  (see equation (4)).

3.2. Bifurcation Analysis. For an ESN that satisfies the ESP, different reservoir states induced by the same driving input from a compact set U should asymptotically converge. If the input is given zero (i.e., u = 0), the reservoir starting from

any initial state should converge to the origin which is a unique global stable fixed point. The stability and uniqueness of the fixed point are analytically provable, yet the analytical approaches seem to be challenging in the case of the nonlinear mapping of input signal containing noise because the problem of separating the stochastic part from the deterministic term is not easy to solve. Instead, we use numerical simulations to examine the nonlinear nonautonomous dynamical properties of the ESN upon noisy driving input.

The ESNs upon a driving input with white Gaussian noise evolve following the relation:

$$x_{k+1} = \tanh\left[Wx_k + W^{\text{in}}\xi_k\right],\tag{6}$$

where  $W \in \mathbb{R}^{N \times N}$  is an internal weight matrix,  $W^{\text{in}} \in \mathbb{R}^{N \times K}$ is an input matrix,  $x_k \in \mathbb{R}^{N \times 1}$  is reservoir state, and  $\xi_k \in \mathbb{R}^{K \times \hat{1}}$  is the K-dimensional vector whose component is a n independent white Gaussian noise. The input matrix is fixed to  $W^{\text{in}} = 0.5\mathbb{I}_{N \times N}$  when noise is given for convenience, and  $W^{in} = 0$  without input. As discussed above, if there is no input (i.e.  $W^{in} = 0$ ), the ESNs with ESP will converge to the origin. Reservoir states were updated at each step following equation (6). After one update, all states are bounded within (-1, 1) since the activation function is bounded. Hence, initial states within [-1, 1] are enough to examine the full dynamics of the reservoirs. In onedimensional ESNs, 50 initial states within [-1, 1] with equal intervals randomly chosen from the uniform distribution U(-1, 1) and the reservoirs were updated for 300 steps. In two-dimensional ESNs, 100 initial states are chosen randomly in  $[-1, 1] \times [-1, 1]$ —each component is extracted from a uniform distribution U(-1, 1)—and updated for 300 steps. Increasing the spectral radius of W (by multiplying constants), we observe the stability in ESNs to track the ESP. In the simulations, the input noise was given in the same way as the MNIST classification task (Section 3.3).

3.3. Handwritten Digit Classification Tasks. The modified National Institute of Standards and Technology (MNIST) image database for handwritten digits ("0," "1," ..., and "9") is used to test the performances of the standard ESN. The training and test accuracies are obtained using different numbers of examples: the sizes of the training sets vary from 120 to 240, ..., 30,000, and the test sets are onesixth of the corresponding training set size, i.e., from 20 to 40, ..., 5,000 examples. Each training or testing set contains an equal number of samples from each digit. The pixel dimension of a digit number image was  $28 \times 28 = 784$  pixels. The AWGN was added to each image with different signal-to-noise ratios (SNRs). The SNR was defined as  $\mathbb{E}(\tilde{P}_{n,(i,j)}^2)/\operatorname{Var}(\sigma)$ , where  $\tilde{P}_{n,(i,j)}$  is the pixel intensity of the (i, j) pixel (i, j = 1, 2, ..., 28) in the  $n^{\text{th}}$ image without noise, and  $\sigma$  denotes a white Gaussian noise with a mean of zero. The variance of  $\sigma$  is determined by the given SNR value. Pixel intensities of the noisy image are given as follows:

$$P_{n,(i,j)} = \begin{cases} 0, & \text{if } \tilde{P}_{n,(i,j)} + \sigma < 0, \\ 255, & \text{if } \tilde{P}_{n,(i,j)} + \sigma > 255, \\ \tilde{P}_{n,(i,j)} + \sigma, & \text{otherwise.} \end{cases}$$
(7)

Each image is transformed into 28 temporal driving input signals by presenting the normalized intensities of 28 pixels in a column at each time step, i.e.,  $j^{\text{th}}$  element (j = 1, 2, ..., 28) of the input signal  $u_k$  at the time step k = 28n + i is given by  $P_{n,(i,j)}/255$ . The readout layer consists of 10 simple neurons with a linear activation function, where the output signal  $\hat{y}_D$  of  $D^{\text{th}}$  readout neuron represents the likelihood that the input image at time t belongs to class D. We train the output  $\hat{y}_D$  to copy a target signal  $y_D$ , where  $y_D(t) = 1$  if the current image at time t belongs to class D,  $y_D(t) = 0$  otherwise. For the test, class D with the highest probability is chosen.

### 4. Results

Figure 1 exhibits the effect of noise on the handwritten digit classification accuracy using the standard ESN model (see equation (2)). For each SNR level, the training and test accuracies converged as the training/test dataset size increased (left side of Figure 1(a); training and test accuracies colored cyan and blue, respectively), where the accuracies declined as the SNR level decreased (i.e., increasing AWGN levels). The neuronal activities in the reservoir fluctuated more frequently as the SNR decreased (Figure 1(a)) and these caused the decrease of the mean absolute value of cross-correlations with a lag of zero ( $\overline{C}_{abs}$ ) between all pairs of neuronal activities in the reservoir (Figure 1(b)). The dependence of the training and test accuracies on the SNR was nonlinear, as displayed in Figure 1(c): both the training and test accuracies drastically dropped at SNR <~1.0.

The short-term memory capacity (MC) of the reservoir, which is the primary measure of the network's ability to store the past input information [52], nonlinearly decreased upon reducing the SNR level (i.e., promoting AWGN) (Figure 2(a)). The MC is the sum of  $\delta$ -delay short-term memory capacity  $MC_{\delta}$  ( $\delta = 1, ..., \delta_{max}$ ) against SNR levels (i.e., MC =  $\sum_{\delta=1}^{\delta_{max}} MC_{\delta}$ ), where the MC<sub> $\delta$ </sub> means the coefficient of determination  $(R^2)$  of the linear regression with the reservoir state vector  $x_k$  to predict driving input signal  $u_{k-\delta}$ with a delay of  $\delta$ ;  $\delta_{max}$  was set to 20 in Figure 2(a) and each of  $MC_{\delta}$  values against SNR levels are displayed in Appendix B. The mapping score  $(S_M)$  of the driving input signal also decreased by promoting the AWGN levels (Figure 2(b)). The score quantifies how much information of the driving input signal is mapped to the reservoir against noise by comparing the reservoir state induced by the driving inputs with and without the noise. It is defined as  $S_M = 1/(1 + D(\tilde{x}, x))$ , where *D* denotes the Euclidean distance and  $\tilde{x}$  and *x* mean the state vector by the input sequence without noise  $(\tilde{u}_k)$  and that by the noisy input sequence  $(u_k = \tilde{u}_k + \xi_k)$ , respectively. The trends of these two information processing indices of the reservoir are in good agreement with the deterioration

pattern of the classification accuracy due to the noise (Figure 1(c)). This indicates that the addition of noise strongly influences the information processing of the reservoir via changing the neural activities, which in turn impacts the computational performance.

As a way to assess the computational capability of the reservoir, the echo state property (ESP) index for noisy driving input which reflects the dynamical changes of the reservoir is devised based on the work by Gallicchio [48]. The ESP index measures the average deviation of noisedriven trajectories induced by random initial states from a noiseless trajectory starting from the zero state (see Section 3.1 for a formal definition). It thus captures the existence of the ESP: intuitively, an ESP index value close to zero strongly suggests that the ESN possesses the ESP, while a larger value means the reservoir is far from having the ESP. Figure 3 displays the ESP index and test accuracy for noisy driving input as functions of spectral radius and input scaling at different noise conditions-no addition of noise and adding the AWGN with SNR from 4.0 to 0.1. The reservoirs were generated by varying spectral radii and input scaling values as follows: the elements of the input matrix  $W^{in}$  were randomly sampled from a uniform distribution on the interval [-a, a], where a represents the input scaling. The internal weight matrix W was randomly generated so that 1% of the elements are nonzero, and these nonzero values are uniformly sampled from the range [-1, 1]. Then, W was rescaled to achieve the desired spectral radius. For the case without noise (Figure 3(a)), both the ESP index for noisy driving input and the original index by Gallicchio [48] generally agree with the distribution of the test accuracies on the spectral radius-input scaling plane: the index value becomes higher (i.e., less tendency to have ESP) as the spectral radius increases and the input scaling decreases (i.e., towards the top-left corner on the plane). However, when increasing the noise level (Figures 3(b)-3(d)), the ESP indices for noisy input well capture the collapse of computational properties of the reservoirs (i.e., deterioration of test accuracies), while the original index did not capture the effect of noise on the computation.

The detailed dynamical properties of ESNs upon noisy driving input are investigated using the one-dimensional (N=1) and two-dimensional (N=2) ESN models in agreement with Yildiz's method for zero input cases [42]. In N=1 case,  $W = \rho(W)$  is a constant scalar and it is the spectral radius itself. x = 0 is a trivial fixed point in this system for arbitrary W. The stability of a trivial fixed point is determined by the spectral radius; x = 0 is stable if  $|\rho(W)| < 1$  and unstable if  $|\rho(W)| > 1$ ; additional nonzero stable fixed points emerge for  $|\rho(W)| > 1$ . The evolution of the reservoir states obtained by numerical simulations confirmed these results. Figure 4 exhibits the dynamics of one-dimensional ESNs with zero input and with only the white Gaussian noise. For the zero input case, onedimensional ESNs converge to trivial fixed point x = 0 for  $\rho(W) = 0.9$  (first column in Figure 4(a)) or two nonzero fixed points for  $\rho(W) = 1.1$  (second column in Figure 4(a)). The reservoir states manifested a pitchfork bifurcation (last column in Figure 4(a): the states converged to a stable fixed



FIGURE 1: Influence of signal-to-noise ratio (SNR) on the handwritten digit classification: (a) (left) training (cyan) and test (blue) accuracies against training/test dataset size and (right) the activity of a neuron in the reservoir given input digits without noise and in cases of additive white Gaussian noise (AWGN) with SNR = 4.0, SNR = 1.0, and SNR = 0.1. The size of the training dataset varied from 120 to 240, . . . , 30,000, and the test data sizes were one-sixth of the training size (i.e., from 20 to 40, . . . , 5,000). The accuracies using the largest dataset are displayed in each panel. (b) The mean absolute value of cross-correlations with a lag of zero ( $\overline{C}_{abs}$ ) among the neuronal activities in the reservoir at different SNR levels (SNR = 4.0, 3.0, . . . , 0.2, 0.1): cross-correlations with a lag of zero between *i*<sup>th</sup> and *j*<sup>th</sup> neurons are defined as  $C_{ij} = \mathbb{E}_k[(x_k^i - \overline{x_k^i})(x_k^j - \overline{x_k^j})]/\sqrt{\mathbb{E}_k[(x_k^i - \overline{x_k^i})^2]}\mathbb{E}_k[(x_k^j - \overline{x_k^j})^2]}$  and  $\overline{C}_{abs} = \mathbb{E}_{ij}[|C_{ij}|]$ . (c) Training and test accuracies over SNR levels for 30,000 training examples and 5,000 test examples. The horizontal dashed line indicates the test accuracy without noise, the same as indicated by the blue color on the top left panel in (a). Error bars indicate the standard error over 20 independent simulations, which are negligible in all cases.

point x = 0 for  $\rho(W) < 1$ . As spectral radius increases, an unstable fixed point emerges around  $\rho(W) = 1$  and two nonzero stable fixed points emerge for  $\rho(W) > 1$ (Figure 4(a)). Upon white Gaussian noise, while the overall tendencies were similar to the zero input case, the reservoir states fluctuated by the noise: different initial states, in turn, converged to an asymptotically same trajectory for all noise levels (SNR = 4.0, 1.0, and 0.1), and the degree of fluctuation increased with the noise level (Figures 4(b)-4(d)). In the cases of  $\rho(W) = 0.9$  (first column in Figures 4(b)-4(d)), reservoir states fluctuate around x = 0 and the degree of fluctuation increases as the noise level increases. In the cases of  $\rho(W) = 1.1$  (second column in Figures 4(b)-4(d)), states flow into the vicinity of one of the two original fixed points. As the noise level increases (or SNR decreases), noise induces more blurring of fixed points, while we can still see the footprints of fixed points observed in the dynamics of ESNs without input. The last 100 steps are used to test the convergence of the reservoir states upon 10 different white Gaussian noise levels (third column in Figures 4(b)–4(d)). For given spectral radius  $\rho(W)$ , states are colored with dark green if states are bounded within tolerance  $\epsilon = 0.05$ , i.e.,  $|x_k| < \epsilon = 0.05$  for  $200 < k \le 300$ . Otherwise, states are colored light green if they are not bounded. In a noisy environment, states can deviate from the origin (i.e.,  $|x_k| \ge \epsilon = 0.05$ ) even at  $\rho(W) < 1$ .

We extend our analysis into two-dimensional ESNs where two neighboring nodes (N = 2) are connected. In N = 2 case, W is a 2 × 2 matrix and the reservoir state in each step k can be represented by a two-dimensional vector  $x_k =$ 



FIGURE 2: Memory capacity and mapping score of the input digit against SNR levels. (a) Short-term memory capacity (MC).  $MC = \sum_{\delta=1}^{20} MC_{\delta}$ , where  $MC_{\delta}$  denotes the coefficient of determination  $(R^2)$  of the linear regression with the reservoir state vector  $x_k$  to predict driving input signal  $u_{k-\delta}$  with a delay of  $\delta$ . (b) Mapping score  $S_M$  of driving input signals. The score is defined as  $S_M = 1/(1 + D(\tilde{x}, x))$ , where D denotes the Euclidean distance and  $\tilde{x}$  and x are the state vectors compatible with the input sequence without noise  $(\tilde{u}_k)$  and the noisy input sequence  $(u_k = \tilde{u}_k + \xi_k)$ , respectively. Error bars indicate the standard error over 20 independent simulations; the values were negligible in most cases.



FIGURE 3: Continued.



FIGURE 3: The echo state property (ESP) index for noisy driving input, original ESP index [48], and test accuracies of the handwritten digit classification tasks against spectral radius and input scaling: (a) MNIST examples without noise; (b) those with additive white Gaussian noise (AWGN) with a signal-to-noise ratio (SNR) = 4.0; (c) AWGN with SNR = 1.0; (d) AWGN with SNR = 0.1. The mean ESP index and test accuracies were obtained over 5 independent simulations with different random initial states, each using 30,000 training examples and 5,000 test examples.



FIGURE 4: Continued.



FIGURE 4: Reservoir states of one-dimensional ESNs with zero input and with only the white Gaussian noise. Evolution of reservoir states with different initial states for one white Gaussian noise input against the spectral radius of the internal weight matrix  $\rho(W) = 0.9$  (left), 1.1 (center), and reservoir states against  $\rho(W)$  (right) in the last 100 steps ( $201 \le k \le 300$ ; green shaded area for the first and second columns): (a) reservoir states without noise; (b) reservoir states driven by white Gaussian input noise with SNR = 4.0; (c) with SNR = 1.0; (d) SNR = 1.0. Each panel of the first and second columns displays 100 reservoir states starting from random initial states, while that of the third column includes 1000 different states (100 randomized reservoir states  $\times$  10 randomized noise generations at each SNR). The noise levels are the same as MNIST tasks in Section 3.3. The last 100 states are colored dark green if they are bounded with tolerance  $\epsilon = 0.05$ , i.e.,  $|x_k| < \epsilon = 0.05$ , and otherwise light green.



FIGURE 5: Continued.



FIGURE 5: Reservoir states  $x_k = [x_k^1, x_k^2]^T$  of two-dimensional ESNs with zero input and with only the white Gaussian noise. Evolution of reservoir states with different initial states for one white Gaussian noise input against the spectral radius of the internal weight matrix  $\rho(W) = 0.9$  (left), 1.1 (center), and reservoirs states against  $\rho(W)$  (right) in the last 100 steps ( $201 \le k \le 300$ ; green shaded area for the first and second columns): (a) reservoir states without noise; (b) reservoir states driven by white Gaussian input noise with SNR = 4.0; (c) with SNR = 1.0; (d) SNR = 1.0. Each panel of the first and second columns displays 100 reservoir states starting from random initial states, while that of the third column includes 1000 different states (100 randomized reservoir states × 10 randomized noise generations at each SNR). The internal weight matrix was chosen to  $W = c \begin{bmatrix} 0.274 & -1.09 \\ 0.83 & 0.35 \end{bmatrix}$ , for all simulations, where *c* is a positive constant to rescale the spectral radius  $\rho(W)$ . The noise levels are the same as MNIST tasks in Section 3.3. The last 100 states are colored dark green if they are bounded with tolerance  $\epsilon = 0.05$ , i.e.,  $|x_k| < \epsilon = 0.05$ , and otherwise light green.

 $[x_k^1, x_k^2]^T$ . Reflecting that the stabilities of N = 2 cases are not solely determined by  $\rho(W)$ , the models exhibit various types of bifurcations compared to the N = 1 systems. In addition to constant rescaling, there are more bifurcation parameters such as tr(W) [42]. Here, we analyze the system which exhibits a Hopf bifurcation under a noise-free environment. In a Hopf bifurcation, a stable fixed point becomes unstable, and a limit cycle arises around the fixed point as the bifurcation parameter crosses the critical value. Internal weight matrix  $W = c \begin{bmatrix} 0.274 & -1.09 \\ 0.83 & 0.35 \end{bmatrix}$  was used and spectral

radius  $\rho(W)$  is rescaled by multiplying the constant c > 0. The reservoir states in two-dimensional ESNs without noise converge to the trivial fixed point  $x = [0, 0]^T$  for  $\rho(W) = 0.9$ as in the one-dimensional case (first column in Figure 5(a)). However, the reservoir states in two-dimensional ESNs with  $\rho(W) = 1.1$  oscillate (second column in Figure 5(a)). The last 100 steps from different initial states converge (third column in Figure 5(a)). They exhibit a Hopf bifurcation; a stable fixed point becomes unstable, and a limit cycle arises around the fixed point as  $\rho(W)$  crosses 1. For spectral radii  $\rho(W) > 1$ , the states are distributed around the origin, implying the



FIGURE 6: ESP index for noisy driving input against spectral radius in one-dimensional (N = 1) and two-dimensional ESNs (N = 2): (a) without noise; (b) driven by white Gaussian input noise with SNR = 4.0; (c) with SNR = 1.0; (d) SNR = 0.1. The noise levels are the same as MNIST tasks. All reservoir parameters are the same as in Figure 4 for N = 1 and in Figure 5 for N = 2; 100 independent simulations with different initial states for each spectral radius and each noise level were done. The original ESP index [48] values were compared for the last 250 time steps.

existence of a limit cycle. This result indicates that ESP does not hold if  $\rho(W) > 1$ . In the case of white Gaussian noise with different levels (i.e., SNR = 4.0, 1.0, 0.1), the reservoir states in two-dimensional ESNs start to fluctuate (Figures 5(b)-5(d)). For  $\rho(W) = 0.9$  (first column in Figures 5(b)-5(d)), reservoir states fluctuate around the fixed point  $x = [0, 0]^T$  and the degree of fluctuation increases as the noise level increases. In the cases of  $\rho(W) =$ 1.1 (second column in Figures 5(b)-5(d)), the states quickly oscillate within the bounded range. The last 100 steps upon 10 different white Gaussian noise levels are examined to confirm the convergence of the reservoir states (third column in Figures 5(b)-5(d)). For given spectral radius  $\rho(W)$ , states are colored with dark green if states are bounded within tolerance  $\epsilon = 0.05$ ;  $|x_k| < \epsilon = 0.05$  for  $200 < k \le 300$  (third column in Figures 5(b)–5(d)). Otherwise, states are colored light green if they are not bounded. As in the case of the one-dimensional model, ESNs for some initial states can deviate from the origin (i.e.,  $|x_k| \ge \epsilon = 0.05$ ) even for  $\rho(W) < 1$  as the noise level increases.

Results from the bifurcation analysis are related to the ESP index for noisy driving input (equation (5)) and the original index by Gallicchio [48]. The reservoir states in bifurcation analysis (i.e., last 250 time steps (T = 51, L = 300)) in Figures 4 and 5 were used to compute the ESP indices. As expected, both ESP indices well agreed with the case of zero input (i.e., no addition of white Gaussian noise) (Figure 6(a)). The values become drastically increased as the spectral radius is promoted beyond the unity (i.e.,  $\rho(W) > 1$ ), whereby ESP is easily destroyed.



FIGURE 7:  $\delta$ -delay short-term memory capacity MC<sub> $\delta$ </sub> against SNR levels. MC<sub> $\delta$ </sub> is defined as the coefficient of determination ( $R^2$ ) of the linear regression with the reservoir state vector  $x_k$  to predict the driving input signal  $u_{k-\delta}$  with a delay of  $\delta$ . Error bars indicate the standard error over 20 independent simulations; the values were negligible in most cases.

In the presence of noise, only the ESP index for noisy driving input could capture the deterioration of the ESP with the noise level for  $\rho(W) < 1$ ; the original ESP index remained almost zero for this regime (Figures 6(b)–6(d)). For  $\rho(W) > 1$ , both indices increase with the spectral radius.

## 5. Discussion and Conclusions

Despite the importance of the empirical assessment of the ESP for the logical design and optimal operation of reservoir computers, the commonly used conditions for the ESP do not explicitly account for the interference from the input noise which can significantly affect their performance. To provide useful information about the empirical and analytical assessment of the computational capabilities of ESNs, a series of extensive numerical simulations and analytic characterization of the ESP were performed.

The analysis began with the comparison of the primary dynamical properties of the standard ESN model with different input noise levels (i.e., no addition of noise and adding the AWGN with SNR from 4.0 to 0.1). The significant and

distinct relationship between the dynamical measures and the MNIST classification accuracy indicated that the noiseinduced dynamical changes and the computational capability of the reservoir are fundamentally intertwined. We then provided the ESP index for noisy driving input, reflecting these dynamical changes based on the work by Gallicchio [48], to help easily assess the computational capability of ESNs in practical applications. We have extended the bifurcation analysis of Yildiz et al. [42] using the onedimensional and two-dimensional ESN models, by taking into account the effects of AWGN on reservoir dynamics to explicate the underlying physics of the noise effects and to confirm the validity of the proposed ESP index. For both cases in one-dimensional ESN (Figure 4) and twodimensional ESN (Figure 5), the convergence of ESN bifurcates with an increasing spectral radius of the internal weight matrix  $\rho(W)$ : pitchfork bifurcations and Hopf bifurcations were observed and the origin was a unique fixed point for small  $\rho(W)$  without noise. However, the fixed point was distracted by the AWGN (Figures 4 and 5); this means that all state vectors driven by any input sequence from a compact set U would not asymptotically converge to the same state due to the interference of the noise. For both one- and two-dimensional ESNs, when they have a fixed point at  $\rho(W) = 0.9$ , the reservoir states were bounded within the tolerance of  $\epsilon = 0.05$  without AWGN (Figures 4(a) and 5(a)), yet some reservoir states deviated from the fixed point under SNR = 4, 1, 0.1 (Figures 4(b)-4(d) and 5(b)-5(d)). This loss of ESP by the AWGN significantly changed the dynamical properties and information processing of the reservoir as captured by the neural correlations (Figure 1(b)), the memory capacity (Figure 2(a)), and the mapping of the noisy input to the reservoir (Figure 2(b)), and these, in turn, led to the collapse of computational capability as indicated by the deterioration of MNIST classification accuracy against noise (Figure 1(c)). The proposed ESP index was defined as the average deviation of noise-driven trajectories from a noiseless trajectory (see Section 3.1 for a formal definition), and the index well characterized the collapse of the computational properties of the reservoirs (Figures 3 and 6). While the original ESP index [48] was designed for noise-free conditions, the proposed ESP index considered the changes in the dynamical properties due to the input noise (Figures 1, 2, 4, and 5) to better characterize the computational capabilities of ESNs, especially in dealing with real-world problems where the interference from the input noise significantly affect the performance of ESNs. As exhibited in Figures 4 and 5, the reservoir states are strongly

entrained by the noise and the strongly entrained fluctuating patterns of the reservoir states would wash out their transient initial states (Figures 4(b)–4(d) and 5(b)–5(d)). While the original index does not differentiate these effects of noise from the convergence of relevant information of the driving input, the newly proposed ESP index can circumvent these problems by incorporating both trajectories of the reservoir states induced by driving inputs with and without noise in defining the index (equation (5)).

Our work provides a framework to understand ESP in the context of a noisy driving input—the proposed definitions of the ESP and ESP index may enable the empirical assessments of the computational capabilities of the reservoirs for noisy input conditions. This may promote the validity, reliability, and utility of reservoir computers for real-world machine-learning applications.

# Appendix

# A. The Proof of Proposition for Defining the Echo State Property upon Noisy Driving Input

*Proof.* For any two state vector sequences  $x^{-\infty}$  and  $y^{-\infty} \in X^{-\infty}$  compatible with  $u^{-\infty} = \tilde{u}^{-\infty} + \xi^{-\infty}$  and  $v^{-\infty} = \tilde{u}^{-\infty} + \eta^{-\infty}$ , respectively.

$$\begin{aligned} \|x_{k+1} - y_{k+1}\| &= \|F(x_k, u_k) - F(y_k, v_k)\| \\ &= \|\tanh(Wx_k + W^{\text{in}}u_k) - \tanh(Wy_k + W^{\text{in}}v_k)\| \\ &\leq \|W(x_k - y_k) + W^{\text{in}}(u_k - v_k)\| (\because \tanh(x) \text{ is a 1-Lipschitz function}) \\ &\leq \|W(x_k - y_k)\| + \|W^{\text{in}}(\xi_k - \eta_k)\| (\text{by triangular inequality}) \\ &= \sigma^{\max}(W)\|x_k - y_k\| + \sigma^{\max}(W^{\text{in}})\|\xi_k - \eta_k\|. \end{aligned}$$
(A.1)

Applying the inequality above, we obtain

$$\begin{aligned} \|x_{0} - y_{0}\| &\leq \sigma^{\max}(W) \|x_{-1} - y_{-1}\| + \sigma^{\max}(W^{\text{in}}) \|\xi_{-1} - \eta_{-1}\| \\ &\leq \lim_{k \to \infty} \left[ \sigma^{\max}(W) \right]^{k} \|x_{-k} - y_{-k}\| + \sigma^{\max}(W^{\text{in}}) \sum_{k=1}^{\infty} \left[ \sigma^{\max}(W) \right]^{k} \|\xi_{-k} - \eta_{-k}\| \\ &= \sigma^{\max}(W^{\text{in}}) \sum_{k=1}^{\infty} \left[ \sigma^{\max}(W) \right]^{k} \|\xi_{-k} - \eta_{-k}\|. \end{aligned}$$
(A.2)

Therefore, if  $P(\sigma^{\max}(W^{\text{in}})\sum_{k=1}^{\infty} [\sigma^{\max}(W)]^k \|\xi_{-k} - \eta_{-k}\| \le \epsilon) \ge \gamma$ , then  $P(\|x_0 - y_0\| \le \epsilon) \ge P(\sigma^{\max}(W^{\text{in}})\sum_{k=1}^{\infty} [\sigma^{\max}(W)]^k \|\xi_{-k} - \eta_{-k}\| \le \epsilon) \ge \gamma$ .

## **B. Short-Term Memory Capacity for Each Delay**

Short-term memory capacity for each delay is provided in Figure 7.

## **Data Availability**

The codes used to support the findings of this study have been deposited in the GitHub repository (https://github. com/LCNP-KIST/Echo-state-property-upon-noisy-drivinginput).

## Disclosure

The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

# **Conflicts of Interest**

The authors declare that they have no conflicts of interest regarding the publication of this paper.

## **Authors' Contributions**

K.H. conceptualized the study. J.W., H.K., and K.H. contributed to modeling and simulations. J.W., H.K., S.H.K., and K.H. performed analysis. K.H., H.K., J.W., and S.H.K. wrote and prepared the original draft. K.H. wrote, reviewed, and edited the article. K.H. performed supervision. All authors have read and agreed to the published version of the manuscript. Junhyuk Woo and Hyeongmo Kim contributed equally to this work.

## Acknowledgments

This research was funded by the Korea Institute of Science and Technology (KIST) Institutional Program (Project nos. 2E32921 and 2E32163), National R&D Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2021M3F3A2A01037808), and Institute of Information and Communications Technology Planning and Evaluation (IITP) grant funded by the Korea government (MSIT) (no. 2022-0-00159, Development of AI edge devices for the prediction of senior cardiocerebrovascular disease and dementia).

## References

- M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review*, vol. 3, no. 3, pp. 127–149, 2009.
- [2] B. Schrauwen, D. Verstraeten, and J. Van Campenhout, "An overview of reservoir computing: theory, applications and implementations," in *The 15th European Symposium on Artificial Neural Networks*, ESANN, Bruges, Belgium, 2007.

- [3] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Processings of the 30th International Conference on Machine Learning*, PMLR, Atlanta, GA, USA, April 2013.
- [4] W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: a new framework for neural computation based on perturbations," *Neural Computation*, vol. 14, no. 11, pp. 2531–2560, 2002.
- [5] H. Jaeger, "The "echo state" approach to analysing and training recurrent neural networks," in *GMD Report*, German National Research Center for Information Technology, Bonn, Germany, 2001.
- [6] J. J. Steil, "Backpropagation-decorrelation: online recurrent learning with O(N) complexity," in *Processings of the 2004 IEEE International Joint Conference on Neural Networks* (*IEEE Cat. No.04CH37541*), Budapest, Hungary, July 2004.
- [7] D. Verstraeten, B. Schrauwen, M. D'Haene, and D. Stroobandt, "An experimental unification of reservoir computing methods," *Neural Networks*, vol. 20, no. 3, pp. 391–403, 2007.
- [8] H. Jaeger and H. Haas, "Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 5667, pp. 78–80, 2004.
- [9] W. Maass and H. Markram, "On the computational power of circuits of spiking neurons," *Journal of Computer and System Sciences*, vol. 69, no. 4, pp. 593–616, 2004.
- [10] W. Maass, "Liquid state machines: motivation, theory, and applications," in *Computability in Context*, pp. 275–296, IMPERIAL COLLEGE PRESS, London, UK, 2011.
- [11] X. Lin, Z. Yang, and Y. Song, "Short-term stock price prediction based on echo state networks," *Expert Systems with Applications*, vol. 36, no. 3, pp. 7313–7317, 2009.
- [12] I. Ilies, H. Jaeger, O. Kosuchinas, M. Rincon, V. Sakenas, and N. Vaskevicius, "Stepping forward through echoes of the past: forecasting with echo state networks. Short report on the winning entry to the NN3 financial forecasting competition," 2007, http:// www.neural-forecasting-competition.com/downloads/NN3/ methods/27-NN3\_Herbert\_Jaeger\_report.pdf.
- [13] M. A. Escalona-Morán, M. C. Soriano, I. Fischer, and C. R. Mirasso, "Electrocardiogram classification using reservoir computing with logistic regression," *Institute of Electrical and Electronics Engineers Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 892–898, 2015.
- [14] O. Al Zoubi, M. Awad, and N. K. Kasabov, "Anytime multipurpose emotion recognition from EEG data using a Liquid State Machine based framework," *Artificial Intelligence in Medicine*, vol. 86, pp. 1–8, 2018.
- [15] A. Das, P. Pradhapan, W. Groenendaal et al., "Unsupervised heart-rate estimation in wearables with Liquid states and a probabilistic readout," *Neural Networks*, vol. 99, pp. 134– 147, 2018.
- [16] P. Buteneers, D. Verstraeten, P. van Mierlo et al., "Automatic detection of epileptic seizures on the intra-cranial electroencephalogram of rats using reservoir computing," *Artificial Intelligence in Medicine*, vol. 53, no. 3, pp. 215–223, 2011.
- [17] S. Ghosh, A. Senapati, A. Mishra et al., "Reservoir computing on epidemic spreading: a case study on COVID-19 cases," *Physical Review A*, vol. 104, no. 1, Article ID 14308, 2021.
- [18] D. Verstraeten, B. Schrauwen, D. Stroobandt, and J. Van Campenhout, "Isolated word recognition with the Liquid State Machine: a case study," *Information Processing Letters*, vol. 95, no. 6, pp. 521–528, 2005.
- [19] M. D. Skowronski and J. G. Harris, "Minimum mean squared error time series classification using an echo state network

prediction model," in *Processings of the 2006 IEEE International Symposium on Circuits and Systems*, Kos, Greece, May 2006.

- [20] S. Krishnagopal, Y. Aloimonos, and M. Girvan, "Similarity learning and generalization with limited data: a reservoir computing approach," *Complexity*, vol. 2018, Article ID 6953836, 15 pages, 2018.
- [21] N. Schaetti, M. Salomon, and R. Couturier, "Echo state networks-based reservoir computing for MNIST handwritten digits recognition," in *Processings of the 2016 IEEE Intl Conference on Computational Science and Engineering (CSE)* and IEEE Intl Conference on Embedded and Ubiquitous Computing (EUC) and 15th Intl Symposium on Distributed Computing and Applications for Business Engineering (DCABES), IEEE, Paris, France, August 2016.
- [22] P. Wijesinghe, G. Srinivasan, P. Panda, and K. Roy, "Analysis of liquid ensembles for enhancing the performance and accuracy of liquid state machines," *Frontiers in Neuroscience*, vol. 13, p. 504, 2019.
- [23] J. Woo, S. H. Kim, H. Kim, and K. Han, "Characterization of the neuronal and network dynamics of liquid state machines," *Physica A: Statistical Mechanics and Its Applications*, vol. 633, Article ID 129334, 2024.
- [24] L.-W. Kong, H.-W. Fan, C. Grebogi, and Y.-C. Lai, "Machine learning prediction of critical transition and system collapse," *Physical Review Research*, vol. 3, no. 1, Article ID 13090, 2021.
- [25] R. Xiao, L.-W. Kong, Z.-K. Sun, and Y.-C. Lai, "Predicting amplitude death with machine learning," *Physical Review A*, vol. 104, no. 1, Article ID 14205, 2021.
- [26] S. H. Lim, L. Theo Giorgini, W. Moon, and J. S. Wettlaufer, "Predicting critical transitions in multiscale dynamical systems using reservoir computing," *Chaos*, vol. 30, no. 12, Article ID 123126, 2020.
- [27] M. Roy, S. Mandal, C. Hens, A. Prasad, N. V. Kuznetsov, and M. Dev Shrimali, "Model-free prediction of multistability using echo state network," *Chaos*, vol. 32, no. 10, Article ID 101104, 2022.
- [28] A. Banerjee, J. D. Hart, R. Roy, and E. Ott, "Machine learning link inference of noisy delay-coupled networks with optoelectronic experimental tests," *Physical Review X*, vol. 11, no. 3, Article ID 31014, 2021.
- [29] A. Banerjee, S. Chandra, and E. Ott, "Network inference from short, noisy, low time-resolution, partial measurements: application to *C. elegans* neuronal calcium dynamics," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 120, no. 12, Article ID 2216030120, 2023.
- [30] H. Jaeger, "Adaptive nonlinear system identification with echo state networks," Advances in Neural Information Processing Systems, vol. 15, 2002.
- [31] A. Lazar, G. Pipa, and J. Triesch, "Fading memory and time series prediction in recurrent networks with different forms of plasticity," *Neural Networks*, vol. 20, no. 3, pp. 312–322, 2007.
- [32] J. L. Rosselló, M. L. Alomar, A. Morro, A. Oliver, and V. Canals, "High-density liquid-state machine circuitry for time-series forecasting," *International Journal of Neural Systems*, vol. 26, no. 5, Article ID 1550036, 2016.
- [33] 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," *Physical Review Letters*, vol. 120, no. 2, Article ID 24102, 2018.
- [34] T. Weng, H. Yang, C. Gu, J. Zhang, and M. Small, "Synchronization of chaotic systems and their machine-learning models," *Physical Review A*, vol. 99, no. 4, Article ID 42203, 2019.

- [35] J. Jiang and Y.-C. Lai, "Model-free prediction of spatiotemporal dynamical systems with recurrent neural networks: role of network spectral radius," *Physical Review Research*, vol. 1, no. 3, Article ID 33056, 2019.
- [36] F. Borra, A. Vulpiani, and M. Cencini, "Effective models and predictability of chaotic multiscale systems via machine learning," *Physical Review A*, vol. 102, no. 5, Article ID 52203, 2020.
- [37] C. Zhang, J. Jiang, S.-X. Qu, and Y.-C. Lai, "Predicting phase and sensing phase coherence in chaotic systems with machine learning," *Chaos*, vol. 30, no. 8, Article ID 83114, 2020.
- [38] V. Pyragas and K. Pyragas, "Using reservoir computer to predict and prevent extreme events," *Physics Letters A*, vol. 384, no. 24, Article ID 126591, 2020.
- [39] S. Nag Chowdhury, A. Ray, A. Mishra, and D. Ghosh, "Extreme events in globally coupled chaotic maps," *Journal of Physics: Complexity*, vol. 2, no. 3, Article ID 35021, 2021.
- [40] N. A. K. Doan, W. Polifke, and L. Magri, "Short-and longterm predictions of chaotic flows and extreme events: a physics-constrained reservoir computing approach," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 477, no. 2253, Article ID 20210135, 2021.
- [41] M. Roy, A. Senapati, S. Poria, A. Mishra, and C. Hens, "Role of assortativity in predicting burst synchronization using echo state network," *Physical Review A*, vol. 105, no. 6, Article ID 64205, 2022.
- [42] I. B. Yildiz, H. Jaeger, and S. J. Kiebel, "Re-visiting the echo state property," *Neural Networks*, vol. 35, pp. 1–9, 2012.
- [43] M. Buehner and P. Young, "A tighter bound for the echo state property," *Institute of Electrical and Electronics Engineers Transactions on Neural Networks*, vol. 17, no. 3, pp. 820–824, 2006.
- [44] G. Manjunath and H. Jaeger, "Echo state property linked to an input: exploring a fundamental characteristic of recurrent neural networks," *Neural Computation*, vol. 25, no. 3, pp. 671–696, 2013.
- [45] G. Wainrib and M. N. Galtier, "A local echo state property through the largest Lyapunov exponent," *Neural Networks*, vol. 76, pp. 39–45, 2016.
- [46] S. Basterrech, "Empirical analysis of the necessary and sufficient conditions of the echo state property," in *Processings of the 2017 International Joint Conference on Neural Networks* (IJCNN), Anchorage, AK, USA, May 2017.
- [47] T. Kubota, K. Nakajima, and H. Takahashi, "Echo state property of neuronal cell cultures," in *Processings of the International Conference on Artificial Neural Networks* (*ICANN*), Springer International Publishing, Munich, Germany, September 2019.
- [48] C. Gallicchio, "Chasing the echo state property," in *Processings of the European Symposium on Artificial Neural Networks* (ESANN), Bruges, Belgium, April 2019.
- [49] S. H. Strogatz, Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering, CRC Press, Boca Raton, FL, USA, 2018.
- [50] J. M. T. Thompson and H. B. Stewart, Nonlinear Dynamics and Chaos, Wiley, Hoboken, NJ, USA, 2002.
- [51] C. Meunier and A. D. Verga, "Noise and bifurcations," *Journal of Statistical Physics*, vol. 50, no. 1-2, pp. 345–375, 1988.
- [52] H. Jaeger, "Short term memory in echo state networks," in *GMD Report*, GMD Forschungszentrum Informationstechnik, Bonn Germany, 2002.