



OPEN

Existence of reservoir with finite-dimensional output for universal reservoir computing

Shuhei Sugiura¹, Ryo Ariizumi^{2✉}, Toru Asai¹ & Shun-ichi Azuma³

In this paper, we prove the existence of a reservoir that has a finite-dimensional output and makes the reservoir computing model universal. Reservoir computing is a method for dynamical system approximation that trains the static part of a model but fixes the dynamical part called the reservoir. Hence, reservoir computing has the advantage of training models with a low computational cost. Moreover, fixed reservoirs can be implemented as physical systems. Such reservoirs have attracted attention in terms of computation speed and energy consumption. The universality of a reservoir computing model is its ability to approximate an arbitrary system with arbitrary accuracy. Two sufficient reservoir conditions to make the model universal have been proposed. The first is the combination of fading memory and the separation property. The second is the neighborhood separation property, which we proposed recently. To date, it has been unknown whether a reservoir with a finite-dimensional output can satisfy these conditions. In this study, we prove that no reservoir with a finite-dimensional output satisfies the former condition. By contrast, we propose a single output reservoir that satisfies the latter condition. This implies that, for any dimension, a reservoir making the model universal exists with the output of that specified dimension. These results clarify the practical importance of our proposed conditions.

Keywords Machine learning, Neural network, Nonlinear dynamical system, Reservoir computing

Reservoir computing (RC) is a machine learning method for dynamical system approximation. An RC model consists of a dynamical system called a reservoir and a static function called a readout. First, the input signal to the RC model is processed in the reservoir. Next, the signal from the reservoir is processed using the readout, and the model output is obtained. The concept of RC is that only the static part of the model is trained, that is, the readout, to make the model behave as desired^{1–5}. RC was proposed initially to simplify the training of recurrent neural networks (RNNs)^{6,7} and is superior in terms of computational cost for training. Generally, a randomly generated RNN is used as the reservoir. However, because the reservoir is fixed, a physical system that is difficult to adjust can also be used as the reservoir. Recently, physical RC, which uses a physical system as the reservoir, has received considerable attention^{8–12}. Physical reservoirs are expected to be superior to an RNN implemented on a general-purpose computer in terms of processing speed and energy consumption^{8,13}.

We say that an RC model is universal if it can approximate an arbitrary system with arbitrary accuracy by training only the readout. Several studies have been performed on the approximation ability of RC models. Grigoryeva et al.¹⁴ studied an echo state network (ESN)¹, which is a typical RC model composed of an RNN reservoir and linear readout. Gonon et al.¹⁵ studied the three classes of RC models for stochastic inputs: one composed of a linear reservoir and polynomial readout, one composed of a state-affine reservoir and linear readout, and an ESN. They showed that each model can approximate any system with any accuracy. However, their results differ from the universality that we deal with in this paper. This is because, for each target system to be approximated, they must train both the readout and the reservoir. Fixing the reservoir regardless of the targets is the concept of RC, and by violating it, advantages such as the hardware implementation of the reservoir and reduction of the computational cost of training are lost.

We say that a reservoir is universal if it makes the RC model universal. To the best of our knowledge, it is unknown whether a universal reservoir with a finite-dimensional output exists. This is an important problem because a reservoir's output is finite in practice. The keys to solving this problem are the study of Maass et al.² and our recent study¹⁶. In each of these studies on continuous-time RC, a sufficient condition for a reservoir to

¹Graduate School of Engineering, Nagoya University, Nagoya 464-8603, Japan. ²Department of Mechanical Systems Engineering, Tokyo University of Agriculture and Technology, Koganei, Tokyo 184-8588, Japan. ³Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan. ✉email: ryoariizumi@go.tuat.ac.jp

be universal was proposed. They use a polynomial readout, evaluate the approximation error using the uniform norm, and assume that the target has fading memory. Fading memory means that if two inputs were close to each other in the recent past, the present outputs are also close. In^{2,16}, an m -output reservoir is represented as a set of m operators, which are maps between input functions and scalar-valued output functions. Therefore, we call a reservoir with a finite-dimensional output a finite reservoir. In², input and output functions of the operators are defined on the bi-infinite-time (BIT) interval \mathbb{R} . Hence, we call these operators BIT operators. Maass et al.² showed that a reservoir is universal if it has the separation property and operators in it have fading memory. In¹⁶, input and output functions of the operators are defined on the right-infinite-time (RIT) interval $\mathbb{R}_+ = [0, \infty)$. Hence, we call these operators RIT operators. In¹⁶, we showed that a reservoir is universal if it has the neighborhood separation property (NSP) and the operators in it are bounded. However, it remains an open question whether there exists any finite reservoir satisfying those conditions^{2,16}.

In this paper, we provide two results. First, we show that no finite reservoir satisfies the condition in². We derive a contradiction from the assumption that an m -output reservoir satisfies the condition in²: the separation property and fading memory. BIT operators have a one-to-one correspondence between functionals, which are maps from input signals to \mathbb{R} . Using the functionals, we construct a map from the compact space of input signals to \mathbb{R}^m . The separation property and fading memory mean that the constructed map is injective and continuous, respectively. This leads to the contradiction that the space of input signals and a subset of \mathbb{R}^m are homeomorphic, although they have different dimensions. As the second result, we show that there is a reservoir that has a single output and the NSP, which is the condition in¹⁶. RIT operators also have a one-to-one correspondence between functionals. We show that a single output reservoir with the NSP exists if a functional with a continuous left inverse exists. To obtain such a functional, we use the Hahn–Mazurkiewicz theorem, which provides continuous surjection from $[0, 1]$ to the space of functional inputs. Assuming the axiom of choice, we can take the right inverse of the surjection, which is the functional that we seek.

Our contribution in this study is to show that there is no finite reservoir satisfying the condition in² but there is one satisfying the condition in¹⁶. Through the discussion, we provide an example of a universal reservoir with a single output. The mathematical meaning of our example is that an operator, which is a map between functions, can be approximated by training only the readout, which is a continuous map from $[0, 1]$ to \mathbb{R} . This is a counter-intuitive and interesting result. The practical meaning of our example is that a reservoir has the possibility of being universal, regardless of the dimension of the output. This is particularly important for using physical reservoirs because they are difficult to adjust.

The structure of this paper is as follows: first, we show that there is no finite reservoir that satisfies the condition in². Second, we show that a universal reservoir with a single output exists that satisfies the condition in¹⁶. Finally, we conclude the paper.

Reservoir computing represented by bi-infinite-time operators

In the study by Maass et al.², which is one of the earliest on RC, they proposed a condition for the reservoir to be universal. We first consider this condition to prove the existence of a universal finite reservoir. Eventually, we conclude that no finite reservoir can satisfy the condition in². In the first subsection, we briefly review², and in the second subsection, we describe one of the two main results of the present study.

Condition for universality

RC is a method to approximate a dynamical system, a map between functions of time. Let $A \subset \mathbb{R}^n$ be a compact set of input values and $K > 0$ be the limit of the speed of input change. We define a set U^B of BIT inputs as follows:

$$U^B = \{u : \mathbb{R} \rightarrow A \mid \forall t_1, t_2 \in \mathbb{R}, \|u(t_1) - u(t_2)\| \leq K|t_1 - t_2|\}. \quad (1)$$

Let Y^B be a set of output functions from \mathbb{R} to \mathbb{R} . We call operators from U^B to Y^B “BIT operators” because they are maps between signals defined on the BIT interval \mathbb{R} . For operator F and input signal u , we write the output signal and its value at time t as Fu and $Fu(t)$, respectively. Maass et al.² defined the reservoir, that is, the dynamical part of a RC model, as a set \mathbb{F} of BIT operators. If the cardinal number $m = |\mathbb{F}|$ is finite, \mathbb{F} represents a dynamical system that returns m output signals $F_1u, \dots, F_mu \in Y^B$ for an input signal $u \in U^B$. Hence, we call m the output dimension of \mathbb{F} and call \mathbb{F} an m -output reservoir. Mathematically, \mathbb{F} can even be an uncountable set: for example, the reservoir representing waves on a liquid surface¹⁷ has the cardinality of continuum.

The RC model $\hat{F} : U^B \rightarrow Y^B$ is defined as follows:

$$\hat{F}u(t) = p(F_1u(t), \dots, F_mu(t)) \quad (u \in U^B, t \in \mathbb{R}), \quad (2)$$

where $i \in \mathbb{N}$ and $F_1, \dots, F_i \in \mathbb{F}$. The function $p : \mathbb{R}^i \rightarrow \mathbb{R}$ is a polynomial called a readout. If m is finite, we can set $i = m$ and $\{F_1, \dots, F_i\} = \mathbb{F}$. In this case, the RC model is trained only by turning the readout p . However, if m is infinite, we must also select a finite number of operators $F_1, \dots, F_i \in \mathbb{F}$. This operation effectively means the training of the reservoir as well as the readout. For instance, in¹⁸, they considered a reservoir composed of all operators defined by a stable linear system. In practice, selecting the operators from such a reservoir means the turning of some parameters of a linear system. Therefore, a finite reservoir is required to achieve the low computational cost of training, which is the advantage of RC. In the remainder of this paper, we assume that m is finite unless otherwise specified.

The RC model and its reservoir are said to be universal if the model can approximate an arbitrary dynamical system with arbitrary accuracy. We formulate the universality of a reservoir represented by BIT operators as follows:

Definition 1 Let reservoir \mathbb{F} be a set of BIT operators F_1, \dots, F_m and let \mathbb{F}^* be another set of BIT operators. Reservoir \mathbb{F} is said to be universal for uniform approximations in \mathbb{F}^* if, for any operator $F^* \in \mathbb{F}^*$ and $\varepsilon > 0$, a polynomial $p : \mathbb{R}^m \rightarrow \mathbb{R}$ exists that satisfies

$$\forall u \in U^B, \forall t \in \mathbb{R}, |F^*u(t) - p(F_1u(t), \dots, F_mu(t))| < \varepsilon. \quad (3)$$

Definition 1 means that, if the reservoir \mathbb{F} is universal, a polynomial of operators F_1, \dots, F_m can approximate any target in \mathbb{F}^* with any accuracy.

To guarantee the universality of RC, Maass et al.² considered BIT operators that have two properties called time invariance and fading memory. An operator $F : U^B \rightarrow Y^B$ is said to be time-invariant if the following holds:

$$\forall t \in \mathbb{R}, \forall u_1, u_2 \in U^B, [(\forall \tau \in \mathbb{R}, u_2(\tau) = u_1(\tau - t)) \Rightarrow (\forall \tau \in \mathbb{R}, Fu_2(\tau) = Fu_1(\tau - t))]. \quad (4)$$

The property time invariance means that a temporal shift of an input also shifts the output.

Fading memory is defined as follows:

Definition 2 A BIT operator $F : U^B \rightarrow Y^B$ is said to have fading memory if the following holds:

$$\forall u_1 \in U^B, \forall \varepsilon > 0, \exists \delta > 0, \exists T > 0, \forall u_2 \in U^B, \left[\max_{\tau \in [-T, 0]} \|u_1(\tau) - u_2(\tau)\| < \delta \Rightarrow |Fu_1(0) - Fu_2(0)| < \varepsilon \right]. \quad (5)$$

Fading memory means that the output value strongly depends on the recent past input but weakly depends on the distant past. Fading memory also means independence from the future, which we define as causality later in the paper.

According to Maass et al.², the following condition is essential for a reservoir to be universal.

Definition 3 Let reservoir \mathbb{F} be a set of BIT operators. Reservoir \mathbb{F} is said to have the separation property if \mathbb{F} satisfies the following:

$$\forall u_1, u_2 \in U^B, \exists F \in \mathbb{F}, [(\exists \tau \leq 0, u_1(\tau) \neq u_2(\tau)) \Rightarrow Fu_1(0) \neq Fu_2(0)]. \quad (6)$$

The separation property means that the reservoir provides different outputs to different inputs. Suppose that the reservoir does not have the separation property, that is, the reservoir returns the same output to two different inputs. Then, the RC model cannot approximate a target that returns different outputs to those inputs. Hence, the separation property is necessary to achieve universality. Note that Eqs. (5) and (6) are the conditions for the output at time 0 because time invariance is assumed. Under time invariance, the same holds at times other than 0.

The result in² is described as follows:

Theorem 1 (Maass and Natschl²) Let \mathbb{F}^* be the set of time-invariant BIT operators with fading memory. Suppose that reservoir $\mathbb{F} \subset \mathbb{F}^*$ has the separation property. Then, reservoir \mathbb{F} is universal for uniform approximations in \mathbb{F}^* .

To summarize the result in², a reservoir with the separation property that contains only time-invariant operators with fading memory is universal. If m can be infinite, a reservoir that satisfies the condition of Theorem 1 exists. For example, the reservoir composed of linear systems¹⁸, which we mentioned before, satisfies the condition.

No finite reservoir satisfies the condition of Theorem 1

We show that any finite reservoir does not satisfy the condition of Theorem 1. To achieve this, we define a functional corresponding to a time-invariant and causal BIT operator, with which fading memory and the separation property are expressed more simply. The causality of a BIT operator is defined as follows:

Definition 4 BIT operator $F : U^B \rightarrow Y^B$ is said to be causal if F satisfies

$$\forall u_1, u_2 \in U^B, \forall t \in \mathbb{R}, [(\forall \tau \leq t, u_1(\tau) = u_2(\tau)) \Rightarrow Fu_1(t) = Fu_2(t)]. \quad (7)$$

A BIT operator with fading memory is causal. With causality, the output value $Fu(t)$ depends only on the input values of u on $(-\infty, t]$. Moreover, with time invariance, this relation does not depend on t . Hence, a causal and time-invariant operator is considered as a relation between input signals on $\mathbb{R}_- = (-\infty, 0]$ and \mathbb{R} . The corresponding functional is defined by such a relation. For the compact set $A \subset \mathbb{R}^n$ and $K > 0$ in Eq. (1), we define the domain V of functionals as follows:

$$V = \{v : \mathbb{R}_- \rightarrow A \mid \forall t_1, t_2 \leq 0, \|v(t_1) - v(t_2)\| \leq K|t_1 - t_2|\}. \quad (8)$$

The correspondence between a BIT operator and a functional is defined as follows:

Definition 5 A causal and time-invariant BIT operator $F : U^B \rightarrow Y^B$ and a functional $f : V \rightarrow \mathbb{R}$ are said to correspond to each other if the following holds:

$$\forall u \in U^B, \forall v \in V, [(\forall \tau \leq 0, u(\tau) = v(\tau)) \Rightarrow Fu(0) = f(v)]. \quad (9)$$

This correspondence is one-to-one¹⁸ and we can express the conditions of operators using their corresponding functionals.

We define a norm on V and a metric derived from it. Let $w : \mathbb{R}_+ \rightarrow (0, 1]$ be a non-increasing function that satisfies $\lim_{t \rightarrow \infty} w(t) = 0$. We define a weighted norm $\|v\|_w$ of $v \in V$ as follows:

$$\|v\|_w = \sup_{\tau \leq 0} \|v(\tau)\| w(-\tau). \tag{10}$$

Using this norm, we define a metric $d : V \times V \rightarrow \mathbb{R}_+$ on V as follows:

$$d(v_1, v_2) = \|v_1 - v_2\|_w. \tag{11}$$

Hereafter, we use d as a metric on V . Fading memory and the separation property are expressed by a functional as follows:

Proposition 1 *A causal and time-invariant operator $F : U^B \rightarrow Y^B$ has fading memory if and only if the functional $f : V \rightarrow \mathbb{R}$ corresponding to F is continuous.*

Proposition 2 *Let $F_1, \dots, F_m : U^B \rightarrow Y^B$ be causal and time-invariant operators and functionals $f_1, \dots, f_m : V \rightarrow \mathbb{R}$ correspond to F_1, \dots, F_m , respectively. Then, reservoir $\mathbb{F} = \{F_1, \dots, F_m\}$ has the separation property if and only if the following map $(f_1, \dots, f_m) : V \rightarrow \mathbb{R}^m$ is injective:*

$$(f_1, \dots, f_m)(v) = (f_1(v), \dots, f_m(v)) \quad (v \in V). \tag{12}$$

We provide the proofs in the supplementary material. Using Propositions 1 and 2, we obtain the following theorem, which provides the conclusion of this section.

Theorem 2 *Suppose that the input range $A \subset \mathbb{R}^n$ in Eq. (1) includes distinct a_1 and a_2 that satisfy*

$$\forall \alpha \in [0, 1], a_1 + \alpha(a_2 - a_1) \in A. \tag{13}$$

Then, no natural number m and time-invariant operators $F_1, \dots, F_m : U^B \rightarrow Y^B$ with fading memory exist such that reservoir $\mathbb{F} = \{F_1, \dots, F_m\}$ has the separation property.

Proof of Theorem 2 We prove the theorem by contradiction. Suppose that natural number m and time-invariant operators $F_1, \dots, F_m : U^B \rightarrow Y^B$ with fading memory exist such that reservoir $\mathbb{F} = \{F_1, \dots, F_m\}$ has the separation property. Let functionals $f_1, \dots, f_m : V \rightarrow \mathbb{R}$ correspond to F_1, \dots, F_m , respectively. Then, from Propositions 1 and 2, the map $(f_1, \dots, f_m) : V \rightarrow \mathbb{R}^m$ is a continuous injection. An injection can be considered as a bijection onto its image, and a continuous bijection from a compact domain to a Hausdorff space is a homeomorphism. Because the set V is compact¹⁶, map (f_1, \dots, f_m) is a topological embedding, that is, a homeomorphism onto its image. We use the following lemma. □

Lemma 1 *A topological embedding $g : [0, 1]^{m+1} \rightarrow V$ exists.*

We provide the proof after the Proof of Theorem 2. From Lemma 1, the composition $(f_1, \dots, f_m) \circ g : [0, 1]^{m+1} \rightarrow \mathbb{R}^m$ is also a topological embedding, that is, $[0, 1]^{m+1}$ and $(f_1, \dots, f_m) \circ g([0, 1]^{m+1}) \subset \mathbb{R}^m$ are homeomorphic. We call the small inductive dimension simply a dimension. Dimensions have the following two properties (see pages 3–4 of reference¹⁹): first, two homeomorphic topological spaces have the same dimension. Second, a topological space has a dimension equal to its subspace or larger. Therefore, we have

$$m + 1 = \text{ind } [0, 1]^{m+1} = \text{ind } (f_1, \dots, f_m) \circ g([0, 1]^{m+1}) \leq \text{ind } \mathbb{R}^m = m, \tag{14}$$

where $\text{ind}(\cdot)$ is the dimension. Eq. (14) is a contradiction, which proves Theorem 2. □

As shown in Theorem 2, any finite reservoir does not satisfy the condition of Theorem 1. Therefore, Theorem 1 cannot support the possibility that a universal finite reservoir exists. We prove Lemma 1 as follows:

Proof of Lemma 1 From the assumption, distinct $a_1, a_2 \in A$ exist that satisfy $a_1 + \alpha(a_2 - a_1) \in A$ for any $\alpha \in [0, 1]$. Let $T = \frac{1}{K} \|a_1 - a_2\|$. For $c \in [0, 1]^{m+1}$, we define a continuous piecewise linear function $\alpha_c : \mathbb{R}_- \rightarrow [0, 1]$ as follows:

$$\alpha_c(t) = c_i + \frac{1}{T} (-iT - t)(c_{i+1} - c_i) \quad (-(i+1)T < t \leq -iT). \tag{15}$$

where $i \in \{0\} \cup \mathbb{N}$, $(c_1, \dots, c_{m+1}) = c$, and $c_0, c_{m+2}, c_{m+3}, \dots$ are zeros. The function α_c satisfies $\alpha_c(-iT) = c_i$. Because the Lipschitz constant of α_c is $\frac{K}{\|a_1 - a_2\|}$ or less, we can define a continuous function $g : [0, 1]^{m+1} \rightarrow V$ as follows:

$$g : c \mapsto v \quad (c \in [0, 1]^{m+1}), \quad v(t) = a_1 + \alpha_c(t)(a_2 - a_1) \quad (t \leq 0). \tag{16}$$

Because $v = g(c)$ satisfies $v(-iT) = a_1 + c_i(a_2 - a_1)$ for any $c = (c_1, \dots, c_{m+1}) \in [0, 1]^{m+1}$ and $i \in \{1, \dots, m+1\}$, g is injective. Therefore, g is a continuous bijection from its compact domain $[0, 1]^{m+1}$ to its image, that is, g is a topological embedding, which proves Lemma 1. \square

Reservoir computing represented by right-infinite-time operators

To obtain the output value of a BIT operator, we have to consider an input signal that continues from the infinite past. In our recent study¹⁶, we avoided this impracticality by defining input and output signals for positive time and proposed another condition for the reservoir to be universal. Using this condition, we prove the existence of a universal finite reservoir. In the first subsection, we briefly explain our previous results¹⁶. In the second subsection, we prove the main result of the present study.

Condition for universality

For compact set $A \subset \mathbb{R}^n$ and $K > 0$ in Eq. (1), we define a set U^R of RIT inputs as follows:

$$U^R = \{u : \mathbb{R}_+ \rightarrow A \mid \forall t_1, t_2 \geq 0, \|u(t_1) - u(t_2)\| \leq K|t_1 - t_2|\}. \quad (17)$$

Let Y^R be the set of output functions from \mathbb{R}_+ to \mathbb{R} . We call operators from U^R to Y^R “RIT operators” because they are maps between signals defined on the RIT interval \mathbb{R}_+ . In¹⁶, the reservoir is defined as a set \mathbb{F} of RIT operators. The RC model $\hat{F} : U^R \rightarrow Y^R$ is defined as follows:

$$\hat{F}u(t) = p(F_1u(t), \dots, F_mu(t)) \quad (u \in U^R, t \geq 0), \quad (18)$$

where $\{F_1, \dots, F_m\} = \mathbb{F}$. The polynomial $p : \mathbb{R}^m \rightarrow \mathbb{R}$ is the readout.

In the RIT operator case, universality is defined as follows:

Definition 6 Let reservoir \mathbb{F} be a set of RIT operators F_1, \dots, F_m and let \mathbb{F}^* be another set of RIT operators. Reservoir \mathbb{F} is said to be universal for uniform approximations in \mathbb{F}^* if, for any operator $F^* \in \mathbb{F}^*$ and $\varepsilon > 0$, a polynomial $p : \mathbb{R}^m \rightarrow \mathbb{R}$ exists that satisfies

$$\forall u \in U^R, \forall t \geq 0, |F^*u(t) - p(F_1u(t), \dots, F_mu(t))| < \varepsilon. \quad (19)$$

RIT operators also correspond to functionals, which are maps to real numbers. We require these corresponding functionals to describe the result in¹⁶. To correspond to a functional, an RIT operator must be causal as follows:

Definition 7 An RIT operator $F : U^R \rightarrow Y^R$ is said to be causal if F satisfies

$$\forall u_1, u_2 \in U^R, \forall t \geq 0, [(\forall \tau \in [0, t], u_1(\tau) = u_2(\tau)) \Rightarrow Fu_1(t) = Fu_2(t)]. \quad (20)$$

With causality, the output value $Fu(t)$ depends only on the input values of u on $[0, t]$. Hence, a causal RIT operator is considered as a relation between input signals of various lengths and \mathbb{R} . The corresponding functional is defined by such a relation. For the set V in Eq. (8), we define the domain V^{res} of functionals as follows:

$$V^{\text{res}} = \left\{ v^{[t]} \mid v \in V, t \geq 0 \right\}, \quad (21)$$

where $v^{[t]}$ is the restriction of v to $[-t, 0]$. Let $\lambda : V^{\text{res}} \rightarrow \mathbb{R}_+$ be a map defined by $v^{[t]} \mapsto t$. For example, $\lambda(v) = 1$ for $v : [-1, 0] \rightarrow A$. The correspondence between an RIT operator and a functional is defined as follows:

Definition 8 A causal RIT operator $F : U^R \rightarrow Y^R$ and a functional $f : V^{\text{res}} \rightarrow \mathbb{R}$ are said to correspond to each other if the following holds:

$$\forall u \in U^R, \forall v \in V^{\text{res}}, [(\forall \tau \in [-t, 0], u(t + \tau) = v(\tau)) \Rightarrow Fu(t) = f(v)], \quad (22)$$

where $t = \lambda(v)$.

Similar to the BIT operator case, this correspondence is one-to-one¹⁶.

We need a metric on V^{res} to explain the condition of reservoirs for universality proposed in¹⁶. Using $w : \mathbb{R}_+ \rightarrow (0, 1]$ in Eq. (10), we define the weighted norm $\|v\|_w$ of $v : [-t, 0] \rightarrow \mathbb{R}^n$ as

$$\|v\|_w = \sup_{\tau \in [-t, 0]} \|v(\tau)\| w(-\tau). \quad (23)$$

Let $\theta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a strictly increasing, bounded, and continuous function. For w and θ , we define the distance $d(v_1, v_2)$ between $v_1, v_2 \in V^{\text{res}}$ as

$$d(v_1, v_2) = \left\| v_1^{[t_{\min}]} - v_2^{[t_{\min}]} \right\|_w + |\theta(t_1) - \theta(t_2)|, \quad (24)$$

where $t_1 = \lambda(v_1)$, $t_2 = \lambda(v_2)$, and $t_{\min} = \min\{t_1, t_2\}$. See Assumption 1 in¹⁶ for detailed conditions for the map $d : V^{\text{res}} \times V^{\text{res}} \rightarrow \mathbb{R}_+$ to be a metric. We extend d onto $V^{\text{res}} \cup V$ by defining $\lambda(v) = \infty$, $v^{[\infty]} = v$, and $\theta(\infty) = \lim_{t \rightarrow \infty} \theta(t)$ for any $v \in V$. With this extension, the metrics in Eqs. (11) and (24) are equivalent on V . Hereafter, we use d as a metric on $V^{\text{res}} \cup V$. Any $v \in V$ is an accumulation point of V^{res} because $v^{[t]} \in V^{\text{res}}$ converges to v as $t \rightarrow \infty$. As shown in Proposition 5 in¹⁶, $V^{\text{res}} \cup V$ is compact. Therefore, $V^{\text{res}} \cup V = \overline{V^{\text{res}}}$ holds, where $\overline{\cdot}$ means closure. In¹⁶, the following condition is proposed for a reservoir to be universal.

Definition 9 Let \mathbb{F} be a set of causal RIT operators F_1, \dots, F_m and let $f_1, \dots, f_m : V^{\text{res}} \rightarrow \mathbb{R}$ correspond to each operator in \mathbb{F} . The set \mathbb{F} is said to have the neighborhood separation property (NSP) if the following holds for any distinct $v_1, v_2 \in \overline{V^{\text{res}}}$:

$$\exists \delta > 0, \overline{(f_1, \dots, f_m)(N_\delta(v_1) \cap V^{\text{res}})} \cap \overline{(f_1, \dots, f_m)(N_\delta(v_2) \cap V^{\text{res}})} = \emptyset, \tag{25}$$

where $N_\delta(v)$ is a δ -neighborhood of $v \in \overline{V^{\text{res}}}$ defined as follows:

$$N_\delta(v) = \{v' \in \overline{V^{\text{res}}} \mid d(v', v) < \delta\}. \tag{26}$$

The NSP guarantees that the images of neighborhoods of distinct points are disjoint. Hence, we can say that the NSP is the “strong injectivity” of the map (f_1, \dots, f_m) .

In¹⁶, target operator $F^* : U^R \rightarrow Y^R$ is assumed to correspond to a uniformly continuous functional. As shown in Proposition 8 in¹⁶, this assumption gives F^* three properties. The first is the following:

$$\forall \varepsilon > 0, \forall u_1 \in U^R, \forall t \geq 0, \exists \delta > 0, \forall u_2 \in U^R, \left[\max_{\tau \in [0, t]} \|u_1(\tau) - u_2(\tau)\| < \delta \Rightarrow |F^*u_1(t) - F^*u_2(t)| < \varepsilon \right]. \tag{27}$$

Roughly speaking, this is a “continuity” of the output value with respect to past inputs. Hence, F^* is also causal. The second property is the equicontinuity of output signals, that is,

$$\forall \varepsilon > 0, \forall t \geq 0, \exists \delta > 0, \forall y \in F^*(U), \forall t' \geq 0, [|t - t'| < \delta \Rightarrow |y(t) - y(t')| < \varepsilon]. \tag{28}$$

The third property is fading memory of an RIT operator¹⁶, which means that older inputs have less influence on the present output. At first glance, the uniform continuity of the target’s corresponding functional is a stricter condition than that in². However, in², continuity implies uniform continuity because the functional domain V is compact. Hence, no difference exists between the two.

The result in¹⁶ is described as follows:

Theorem 3 (Sugiura et al.¹⁶) *Let \mathbb{F}^* be the set of RIT operators corresponding to a uniformly continuous functional and let reservoir \mathbb{F} be a set of bounded RIT operators. Suppose that \mathbb{F} has the NSP. Then, reservoir \mathbb{F} is universal for uniform approximations in \mathbb{F}^* .*

Operator $F : U^R \rightarrow Y^R$ is said to be bounded if $F(U^R)$ is uniformly bounded, that is,

$$\exists c \geq 0, \forall y \in F(U), \forall t \geq 0, |y(t)| \leq c. \tag{29}$$

To summarize the result in¹⁶, a reservoir with the NSP that contains only bounded operators is universal.

Existence of universal finite reservoir

Using Theorem 3, we show that a universal reservoir \mathbb{F} with a single output exists. If the reservoir \mathbb{F} is universal, we can easily construct a universal reservoir with a higher dimensional output by adding arbitrary operators to \mathbb{F} . Hence, the universality of \mathbb{F} means that, for any dimension, a universal reservoir exists with the output of that specified dimension. Let operator $F : U^R \rightarrow Y^R$ and functional $f : V^{\text{res}} \rightarrow \mathbb{R}$ correspond to each other. Then, the NSP of reservoir $\mathbb{F} = \{F\}$ is written as follows:

$$\forall v_1, v_2 \in \overline{V^{\text{res}}}, [v_1 \neq v_2 \Rightarrow (\exists \delta > 0, \overline{f(N_\delta(v_1) \cap V^{\text{res}})} \cap \overline{f(N_\delta(v_2) \cap V^{\text{res}})} = \emptyset)]. \tag{30}$$

The following proposition provides a stronger but simpler condition than Eq. (30):

Proposition 3 *Let $f : V^{\text{res}} \rightarrow \mathbb{R}$. Suppose that there is a continuous left inverse $g : \overline{f(V^{\text{res}})} \rightarrow \overline{V^{\text{res}}}$ of f , that is, a continuous map g that satisfies the following:*

$$g \circ f = \text{id}_{V^{\text{res}}}. \tag{31}$$

Then, f satisfies Eq. (30).

Proof of Proposition 3 We prove the contraposition. Suppose that there are distinct $v_1, v_2 \in \overline{V^{\text{res}}}$ that satisfy the following:

$$\forall \delta > 0, \overline{f(V_1)} \cap \overline{f(V_2)} \neq \emptyset, \tag{32}$$

where $V_1, V_2 \subset V^{\text{res}}$ are defined as follows:

$$V_1 = N_\delta(v_1) \cap V^{\text{res}}, \quad V_2 = N_\delta(v_2) \cap V^{\text{res}}. \tag{33}$$

Because of $v_1 \neq v_2$, we obtain $\overline{V_1} \cap \overline{V_2} = \emptyset$ by setting $\delta = d(v_1, v_2)/3$. Let $(\alpha_{1,i})_{i \in \mathbb{N}} \in f(V_1)$ and $(\alpha_{2,i})_{i \in \mathbb{N}} \in f(V_2)$ be sequences that converge to $\alpha \in \overline{f(V_1)} \cap \overline{f(V_2)}$. Suppose that $g : \overline{f(V^{\text{res}})} \rightarrow V^{\text{res}}$ is a left inverse of f . Then, $g(\alpha_{1,i})$ and $g(\alpha_{2,i})$ are contained in V_1 and V_2 , respectively. Because of $\overline{V_1} \cap \overline{V_2} = \emptyset$, $g(\alpha_{1,i})$ and $g(\alpha_{2,i})$ do not converge to the same point as $i \rightarrow \infty$, even though $\alpha_{1,i}$ and $\alpha_{2,i}$ converge to α . Hence, g is not continuous at $\alpha \in \overline{f(V^{\text{res}})}$. \square

To prove that a functional $f : V^{\text{res}} \rightarrow \mathbb{R}$ with a continuous left inverse exists, we use the following theorem.

Theorem 4 (Hahn–Mazurkiewicz theorem²⁰) *Let E be a connected, locally connected, and compact metric space. Then, a continuous surjection from $[0, 1]$ to E exists.*

Metric space $\overline{V^{\text{res}}} = V^{\text{res}} \cup V$ is compact¹⁶. As we show later in this paper, $\overline{V^{\text{res}}}$ is connected and locally connected. Hence, from Theorem 4, a space-filling curve $g : [0, 1] \rightarrow \overline{V^{\text{res}}}$ exists, which is continuous and surjective. Using the axiom of choice, we obtain a functional $f : V^{\text{res}} \rightarrow \mathbb{R}$ with a continuous left inverse as a right inverse of g . Therefore, a reservoir that contains only one operator can have the NSP. To prove that $\overline{V^{\text{res}}}$ is connected and locally connected, we assume that the set $A \subset \mathbb{R}^n$ in Eq. (8) is convex.

Proposition 4 *The set $\overline{V^{\text{res}}}$ is connected.*

Proof of Proposition 4 We prove that $\overline{V^{\text{res}}}$ is arcwise connected, which is a stronger condition than being connected. Let v_1 and v_2 be arbitrary inputs in $\overline{V^{\text{res}}}$. The set $\overline{V^{\text{res}}}$ is said to be arcwise connected if there is some continuous map $P : [0, 1] \rightarrow \overline{V^{\text{res}}}$ that satisfies $P(0) = v_1$ and $P(1) = v_2$. Such a map P is called a path from v_1 to v_2 . Let $t_1 = \lambda(v_1)$ and $t_2 = \lambda(v_2)$. If a path exists from v_1 to v_2 , one exists from v_2 to v_1 . Hence, we can assume that $t_1 \leq t_2$ without loss of generality. Let $v_3 = v_2^{[t_1]}$, that is, v_3 has the same domain as v_1 and the same value as v_2 . We form a path from v_1 to v_2 through v_3 .

First, we form a path from v_1 to v_3 . Because v_1 and v_3 has the same domain $[-t_1, 0]$, we can define a map $P_1 : [0, 1] \rightarrow \overline{V^{\text{res}}}$ as follows:

$$P_1(\alpha) = (1 - \alpha)v_1 + \alpha v_3 \quad (\alpha \in [0, 1]). \tag{34}$$

Map P_1 satisfies $P_1(0) = v_1$ and $P_1(1) = v_3$. Because the input range $A \subset \mathbb{R}^n$ is convex, the image of P_1 is included in $\overline{V^{\text{res}}}$. Map P_1 is continuous because, for any $\alpha_1, \alpha_2 \in [0, 1]$, the distance between $P_1(\alpha_1)$ and $P_1(\alpha_2)$ is formulated as follows:

$$d(P_1(\alpha_1), P_1(\alpha_2)) = |\alpha_1 - \alpha_2| \|v_1 - v_3\|_w. \tag{35}$$

Therefore, map P_1 is a path from v_1 to v_3 .

Second, we form a path from v_3 to v_2 . The function θ has inverse θ^{-1} because it is strictly increasing and continuous. Using θ^{-1} , we define map $\mu : [0, 1] \rightarrow [t_1, t_2]$ as follows:

$$\mu(\alpha) = \theta^{-1}((1 - \alpha)\theta(t_1) + \alpha\theta(t_2)) \quad (\alpha \in [0, 1]), \tag{36}$$

where $\theta(\infty) = \lim_{t \rightarrow \infty} \theta(t)$ and $\theta^{-1}(\theta(\infty)) = \infty$. We define a map $P_2 : [0, 1] \rightarrow \overline{V^{\text{res}}}$ as follows:

$$P_2(\alpha) = v_2^{[\mu(\alpha)]} \quad (\alpha \in [0, 1]). \tag{37}$$

Map P_2 satisfies $P_2(0) = v_3$ and $P_2(1) = v_2$ because $\mu(0) = t_1$ and $\mu(1) = t_2$ hold. From the definition (21) of V^{res} , the image of P_2 is included in $\overline{V^{\text{res}}}$. Map P_2 is continuous because, for any $\alpha_1, \alpha_2 \in [0, 1]$, the distance between $P_2(\alpha_1)$ and $P_2(\alpha_2)$ is formulated as follows:

$$d(P_2(\alpha_1), P_2(\alpha_2)) = |\theta(\mu(\alpha_1)) - \theta(\mu(\alpha_2))| = |\alpha_1 - \alpha_2| |\theta(t_1) - \theta(t_2)|. \tag{38}$$

Therefore, map P_2 is a path from v_3 to v_2 .

Using P_1 and P_2 , we define path $P : [0, 1] \rightarrow \overline{V^{\text{res}}}$ from v_1 to v_2 as follows:

$$P(\alpha) = \begin{cases} P_1(2\alpha) & (0 \leq \alpha \leq \frac{1}{2}) \\ P_2(2\alpha - 1) & (\frac{1}{2} < \alpha \leq 1) \end{cases}. \tag{39}$$

Therefore, the set $\overline{V^{\text{res}}}$ is arcwise connected. \square

To prove that the set $\overline{V^{\text{res}}}$ is locally connected, we use the following proposition:

Proposition 5 *For any $v_1, v_2 \in \overline{V^{\text{res}}}$, path $P : [0, 1] \rightarrow \overline{V^{\text{res}}}$ from v_1 to v_2 exists that satisfies the following:*

$$\forall \alpha \in [0, 1], d(v_1, v_2) = d(v_1, P(\alpha)) + d(P(\alpha), v_2). \tag{40}$$

Equation (40) means that path P is “the shortest” between v_1 and v_2 , that is, equality holds in the triangle inequality between any point on the path and the two endpoints.

Proof of Proposition 5 Let v_1 and v_2 be arbitrary inputs in $\overline{V}^{\text{res}}$. If path P from v_1 to v_2 exists that satisfies Eq. (40), a similar path from v_2 to v_1 exists. Hence, for $t_1 = \lambda(v_1)$ and $t_2 = \lambda(v_2)$, we can assume that $t_1 \leq t_2$ without loss of generality. We show that path $P : [0, 1] \rightarrow \overline{V}^{\text{res}}$ defined by (39) satisfies Eq. (40) using the following three lemmas: \square

Lemma 2 Path $P_1 : [0, 1] \rightarrow \overline{V}^{\text{res}}$ defined by Eq. (34) satisfies the following:

$$\forall \alpha \in [0, 1], d(v_1, v_3) = d(v_1, P_1(\alpha)) + d(P_1(\alpha), v_3). \quad (41)$$

Proof of Lemma 2 Using Eq. (35), distance $d(v_1, v_3)$ is transformed as follows:

$$\begin{aligned} d(v_1, v_3) &= \|v_1 - v_3\|_w \\ &= |0 - \alpha| \|v_1 - v_3\|_w + |\alpha - 1| \|v_1 - v_3\|_w \\ &= d(P_1(0), P_1(\alpha)) + d(P_1(\alpha), P_1(1)) \\ &= d(v_1, P_1(\alpha)) + d(P_1(\alpha), v_3), \end{aligned} \quad (42)$$

which proves Lemma 2. \square

Lemma 3 Path $P_2 : [0, 1] \rightarrow \overline{V}^{\text{res}}$ defined by Eq. (37) satisfies the following:

$$\forall \alpha \in [0, 1], d(v_3, v_2) = d(v_3, P_2(\alpha)) + d(P_2(\alpha), v_2). \quad (43)$$

Proof of Lemma 3 Using Eq. (38), distance $d(v_3, v_2)$ is transformed as follows:

$$\begin{aligned} d(v_3, v_2) &= |\theta(t_1) - \theta(t_2)| \\ &= |0 - \alpha| |\theta(t_1) - \theta(t_2)| + |\alpha - 1| |\theta(t_1) - \theta(t_2)| \\ &= d(P_2(0), P_2(\alpha)) + d(P_2(\alpha), P_2(1)) \\ &= d(v_3, P_2(\alpha)) + d(P_2(\alpha), v_2), \end{aligned} \quad (44)$$

which proves Lemma 3. \square

Lemma 4 The input $v_3 = v_2^{[t_1]}$ satisfies the following:

$$d(v_1, v_2) = d(v_1, v_3) + d(v_3, v_2). \quad (45)$$

Proof of Lemma 4 Distance $d(v_1, v_2)$ is transformed as follows:

$$\begin{aligned} d(v_1, v_2) &= \left\| v_1 - v_2^{[t_1]} \right\|_w + |\theta(t_1) - \theta(t_2)| \\ &= \|v_1 - v_3\|_w + |\theta(\lambda(v_3)) - \theta(t_2)| \\ &= d(v_1, v_3) + d(v_3, v_2), \end{aligned} \quad (46)$$

which proves Lemma 4. \square

Let α be an arbitrary number in $[0, 1]$. We show that the following holds:

$$d(v_1, P(\alpha)) + d(P(\alpha), v_2) \leq d(v_1, v_2). \quad (47)$$

If $0 \leq \alpha \leq \frac{1}{2}$, we obtain Eq. (47) as follows:

$$d(v_1, P(\alpha)) + d(P(\alpha), v_2) \leq d(v_1, P(\alpha)) + d(P(\alpha), v_3) + d(v_3, v_2) = d(v_1, v_3) + d(v_3, v_2) = d(v_1, v_2). \quad (48)$$

The inequality is a triangle inequality. The first equality is obtained from Lemma 2 because $P(\alpha) = P_1(2\alpha)$ holds. The second equality is obtained from Lemma 4. If $\frac{1}{2} < \alpha \leq 1$, we obtain Eq. (47) as follows:

$$d(v_1, P(\alpha)) + d(P(\alpha), v_2) \leq d(v_1, v_3) + d(v_3, P(\alpha)) + d(P(\alpha), v_2) = d(v_1, v_3) + d(v_3, v_2) = d(v_1, v_2). \quad (49)$$

The inequality is a triangle inequality. The first equality is obtained from Lemma 3 because $P(\alpha) = P_2(2\alpha - 1)$ holds. The second equality is obtained from Lemma 4. From Eq. (47) and the triangle inequality $d(v_1, v_2) \leq d(v_1, P(\alpha)) + d(P(\alpha), v_2)$, we obtain Eq. (40), which proves Proposition 5. \square

Using Proposition 5, we prove the following proposition:

Proposition 6 The set $\overline{V}^{\text{res}}$ is locally connected.

Proof of Proposition 6 We prove that $\overline{V}^{\text{res}}$ is locally arcwise connected, which is a stronger condition than being locally connected. Let v be an arbitrary input in $\overline{V}^{\text{res}}$. The set $\overline{V}^{\text{res}}$ is said to be locally arcwise connected if, for any open set $V_1 \subset \overline{V}^{\text{res}}$ including v , an arcwise connected open set $V_2 \subset V_1$ exists that includes v . Therefore, it is sufficient for the Proof of Proposition 6 to show that a δ -neighborhood $N_\delta(v) \subset \overline{V}^{\text{res}}$ of v is arcwise connected for any $\delta > 0$. Let δ be an arbitrary positive number and v_1, v_2 be arbitrary inputs in $N_\delta(v)$. From Proposition 5 and $d(v, v_1) < \delta$, path $P_1 : [0, 1] \rightarrow \overline{V}^{\text{res}}$ from v_1 to v exists that satisfies the following for any $\alpha \in [0, 1]$:

$$d(v, P_1(\alpha)) = d(v, v_1) - d(P_1(\alpha), v_1) < \delta - d(P_1(\alpha), v_1) \leq \delta. \quad (50)$$

Hence, $P_1([0, 1]) \subset N_\delta(v)$ holds. Similarly, we obtain path $P_2 : [0, 1] \rightarrow N_\delta(v)$ from v to v_2 . Therefore, path $P : [0, 1] \rightarrow N_\delta(v)$ from v_1 to v_2 exists, which proves Proposition 6. \square

Using Theorem 3 and Propositions 3, 4, and 6, we show that a universal reservoir exists that has only one output.

Theorem 5 Let \mathbb{F}^* be the set of RIT operators corresponding to a uniformly continuous functional. Assume the axiom of choice and that the set $A \subset \mathbb{R}^n$ in Eq. (8) is convex. Then, reservoir \mathbb{F} exists that is a set of only one RIT operator and \mathbb{F} is universal for uniform approximations in \mathbb{F}^* .

Proof of Theorem 5 Metric space $\overline{V^{\text{res}}} = V^{\text{res}} \cup V$ is compact¹⁶. From Propositions 4 and 6, $\overline{V^{\text{res}}}$ is connected and locally connected. Hence, from Theorem 4, a continuous surjection $g : [0, 1] \rightarrow \overline{V^{\text{res}}}$ exists. We define a set $S \subset [0, 1]$ as follows:

$$S = \{\alpha \in [0, 1] \mid g(\alpha) \in V^{\text{res}}\}. \quad (51)$$

Let $g|_S : S \rightarrow V^{\text{res}}$ be a restriction of g to S . Because $g|_S$ is a surjection to V^{res} , the axiom of choice provides a right inverse $f : V^{\text{res}} \rightarrow S$ of $g|_S$, that is, f satisfies $g|_S \circ f = \text{id}_{V^{\text{res}}}$. Functional f and another restriction $g|_{\overline{S}} : \overline{S} \rightarrow \overline{V^{\text{res}}}$ of g also satisfy $g|_{\overline{S}} \circ f = \text{id}_{V^{\text{res}}}$. Because the restriction $g|_{\overline{S}}$ is continuous, from Proposition 3, f satisfies Eq. (30). Hence, reservoir $\mathbb{F} = \{f\}$ has the NSP, where $F : U^R \rightarrow Y^R$ is the corresponding operator of f . Because of $S \subset [0, 1]$, functional f and operator F are bounded. Therefore, reservoir \mathbb{F} satisfies the condition of Theorem 3 and is universal for uniform approximations in \mathbb{F}^* . \square

Theorem 5 shows that a universal finite reservoir exists, which achieves the low computational cost of training. Moreover, a universal reservoir exists independent of the dimension of its output. The reservoir obtained from a right inverse of a space-filling curve proposed in the Proof of Theorem 5 is probably chaotic. This result suggests that chaos is key to the reservoir's universality. In practice, chaotic reservoirs have already been considered in research and are known to be useful^{21,22}.

The difference between the results of Sections “Reservoir computing represented by bi-infinite-time operators” and “Reservoir computing represented by right-infinite-time operators” is not caused by the difference between the BIT and RIT operators but the difference between the conditions of functionals f_1, \dots, f_m corresponding to operators in the reservoir. Fading memory and the separation property, which are the conditions of Theorem 1, require the map (f_1, \dots, f_m) to be continuous and injective; however, this is impossible. By contrast, the NSP, which is the condition of Theorem 3, only requires the “strong injectivity” of (f_1, \dots, f_m) , which is possible as described thus far. The NSP is also defined for a reservoir represented by BIT operators as follows:

Definition 10 Let \mathbb{F} be a set of causal BIT operators F_1, \dots, F_m and let $f_1, \dots, f_m : V \rightarrow \mathbb{R}$ correspond to each operator in \mathbb{F} . The set \mathbb{F} is said to have the NSP if the following holds for any distinct $v_1, v_2 \in V$:

$$\exists \delta > 0, \overline{(f_1, \dots, f_m)(N_\delta(v_1))} \cap \overline{(f_1, \dots, f_m)(N_\delta(v_2))} = \emptyset, \quad (52)$$

where $N_\delta(v) \subset V$ is a δ -neighborhood of $v \in V$.

A finite reservoir with the NSP of Definition 10 can be obtained in the same manner as in the RIT operator case. Using Theorem 4 in¹⁶, that is, an extension of the Stone–Weierstrass theorem, we can derive the universality of Definition 1 from the NSP of Definition 10. Therefore, the NSP is the essence of proving the existence of a universal finite reservoir.

Conclusion

We discussed whether a universal finite reservoir exists under the assumption that the readout is a polynomial and the target operator has fading memory. In the discussion, we considered two sufficient conditions for the reservoir to be universal proposed by Maass et al.² and in¹⁶. First, we showed that no finite reservoir satisfies the condition in². Supposing that such a reservoir exists, we derived the contradiction that the input function space and a subset of finite-dimensional vector space are homeomorphic. Next, we proposed an example of a universal reservoir that has a single output and the NSP, which is the condition in¹⁶. The functional corresponding to the operator representing the reservoir has a continuous left inverse. We showed that this is a stronger condition than the NSP. The functional is defined as a right inverse of the continuous surjection from $[0, 1]$ to the input space of the functional. The surjection is given by the Hahn–Mazurkiewicz theorem. Our example means that, for any dimension, a universal reservoir exists that has the output of that specified dimension. This result is particularly important for the use of physical reservoirs, which are difficult to train.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Received: 4 September 2023; Accepted: 11 March 2024

Published online: 11 April 2024

References

1. Jaeger, H. 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, 148.34 (2001).
2. Maass, W. & Natschl, T. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Comput.* **14**(11), 2531–2560 (2002).
3. Steil, J. J. Backpropagation-decorrelation: Online recurrent learning with $O(N)$ complexity. In *2004 IEEE International Joint Conference on Neural Networks* 843–848 (2004).
4. Verstraeten, D., Schrauwen, B., D’Haene, M. & Stroobandt, D. An experimental unification of reservoir computing methods. *Neural Netw.* **20**(3), 391–403 (2007).
5. Lukoševičius, M. & Jaeger, H. Reservoir computing approaches to recurrent neural network training. *Comput. Sci. Rev.* **3**(3), 127–149 (2009).
6. Williams, R. J. & Zipsper, D. A learning algorithm for continually running fully recurrent neural networks. *Neural Comput.* **1**(2), 270–280 (1989).
7. Werbos, P. J. Backpropagation through time: what it does and how to do it. *Proc. IEEE* **78**(10), 1550–1560 (1990).
8. Tanaka, G. *et al.* Recent advances in physical reservoir computing: A review. *Neural Netw.* **115**, 100–123 (2019).
9. Friedman, J. S. Unsupervised learning & reservoir computing leveraging analog spintronic phenomena. *IEEE 16th Nanotechnology Materials and Devices Conference* 1–2 (2021).
10. Stelzer, F., Röhm, A., Lüdge, K. & Yanchuk, S. Performance boost of time-delay reservoir computing by non-resonant clock cycle. *Neural Netw.* **124**, 158–169 (2020).
11. Soriano, M. C. *et al.* Delay-based reservoir computing: Noise effects in a combined analog and digital implementation. *IEEE Trans. Neural Netw. Learn. Syst.* **26**(2), 388–393 (2014).
12. Dong, J., Rafayelyan, M., Krzakala, F. & Gigan, S. Optical reservoir computing using multiple light scattering for chaotic systems prediction. *IEEE J. Sel. Top. Quantum Electron.* **26**(1), 1–12 (2020).
13. Nakajima, K. & Fischer, I. *Reservoir Computing* (Springer, 2021).
14. Grigoryeva, L. & Ortega, J. P. Echo state networks are universal. *Neural Netw.* **108**, 495–508 (2018).
15. Gonon, L. & Ortega, J. P. Reservoir Computing Universality With Stochastic Inputs. *IEEE Trans. Neural Netw. Learn. Syst.* **31**(1), 100–112 (2020).
16. Sugiura, S., Ariizumi, R., Asai, T., & Azuma, S. Nonessentiality of reservoir’s fading memory for universality of reservoir computing. *IEEE Trans. Neural Netw.* <https://doi.org/10.1109/TNNLS.2023.3298013> (2023).
17. Fernando, C. & Sojakka, S. Pattern recognition in a bucket. In *European Conference on Artificial Life* 588–597 (Springer, 2003).
18. Boyd, S. & Chua, L. O. Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Trans. Circuits Syst.* **32**(11), 1150–1161 (1985).
19. Engelking, R. *Dimension Theory* (North-Holland Publishing Company, 1978).
20. Hocking, J. G. & Young, G. S. *Topology* (Addison-Wesley Publishing Company, 1961).
21. Jensen, J. H. & Tufte, G. Reservoir computing with a chaotic circuit. In *Artificial Life Conference Proceedings* 222–229 (MIT Press, 2017).
22. Choi, J. & Kim, P. Reservoir computing based on quenched chaos. *Chaos Solitons Fractals* **140**, 110131 (2020).

Acknowledgements

We thank Edanz (<https://jp.edanz.com/ac>) for editing a draft of this manuscript.

Author contributions

Conceptualization: S.S.; Methodology: S.S.; Writing—original draft preparation: S.S.; Writing—review and editing: All authors; Funding acquisition: R.A., S.A.; Supervision: R.A., T.A., S.A.

Funding

This research was supported by JST FOREST Program under Grant No. JPMJFR2123 and by JSP KAKENHI under Grant No. JP22K04027.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-56742-7>.

Correspondence and requests for materials should be addressed to R.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024