Hierarchy of the echo state property in quantum reservoir computing

Shumpei Kobayashi[1*] **Quoc Hoan Tran**[2,3] **Kohei Nakajima**[1,2,3]
[1]Department of Creative Informatics, The University of Tokyo, Japan
[2]Next Generation Artificial Intelligence Research Center (AI Center),
Graduate School of Information Science and Technology, The University of Tokyo, Japan
[3]Department of Mechano-Informatics, The University of Tokyo, Japan
*s-kobayashi@isi.imi.i.u-tokyo.ac.jp

Abstract

The echo state property (ESP) represents a fundamental concept in the reservoir computing framework that ensures stable output-only training of reservoir networks. However, the conventional definition of ESP does not aptly describe possibly non-stationary systems, where statistical properties evolve. To address this issue, we introduce two new categories of ESP: *non-stationary ESP* designed for possibly non-stationary systems, and *subspace/subset ESP* designed for systems whose subsystems have ESP. Following the definitions, we numerically demonstrate the correspondence between non-stationary ESP in the quantum reservoir computer (QRC) framework with typical Hamiltonian dynamics and input encoding methods using nonlinear autoregressive moving-average (NARMA) tasks. These newly defined properties present a new understanding toward the practical design of QRC and other possibly non-stationary RC systems.

1 Introduction

Physical reservoir computing Tanaka *et al.* (2019); Nakajima (2020), which utilizes non-linear natural dynamics of physical substrate for temporal information processing, has garnered much attention because it can mitigate the massive need for computational resources for sophisticated machine learning methods, such as deep learning. However, not all physical systems can be effectively used as reservoir substrates because of possible initial-state sensitivity in natural dynamics, such as chaotic systems. One precondition to exclude such a system beforehand is the echo state property (ESP), which requires the initial state dependency to vanish over time.

Quantum systems have been attracting attention as one of the promising substrates for physical reservoir computing. However, we here argue that the quantum system, in general, is not always stationary and that, in some cases, the traditional definition of ESP is not helpful for ensuring its capability of temporal information processing. In this manuscript, we define and analyze new conditions that secure such a possibly non-stationary system to behave as a practical reservoir.

1.1 Quantum reservoir computing

In the NISQ Preskill (2018) era, non-universal quantum computation schemes gained much attention because of their near-term feasibility on physical devices. Such computational procedure includes, for instance, variational quantum computation (VQC) McClean *et al.* (2016); Mitarai *et al.* (2018) and quantum reservoir computing (QRC) Fujii and Nakajima (2017); Ghosh *et al.* (2019a). VQC and QRC apply to one-shot and autoregressive quantum machine learning algorithms, which also became a general prospective application of quantum computation. Recent works on QRC in-

37th First Workshop on Machine Learning with New Compute Paradigms at NeurIPS 2023(MLNPCP 2023).

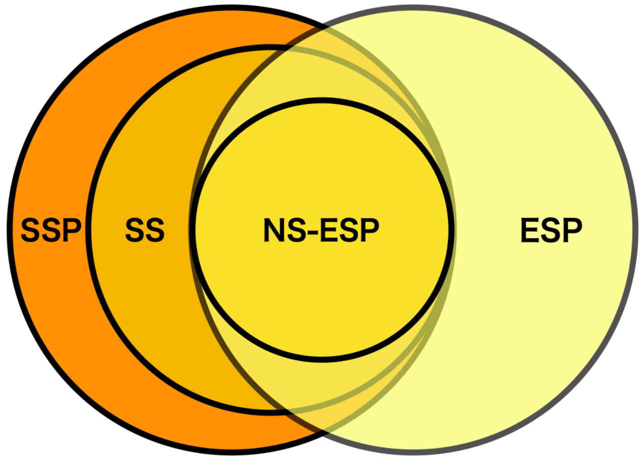


Figure 1: Inclusion relationship of conventional ESP and the one defined in this paper. NS stands for *non-stationary*, SS stands for *subset* and SSP stands for *subspace* ESP, respectively.

clude proposals of QRC in various physical apparatus Negoro *et al.* (2018); Ghosh *et al.* (2019b); Chen *et al.* (2020); Nokkala *et al.* (2021); Govia *et al.* (2021); Spagnolo *et al.* (2022), with some of them performing actual physical experiments, and theoretical analyses Martínez-Peña and Ortega (2023). Specifically, some works Chen *et al.* (2020); Suzuki *et al.* (2022); Sannia *et al.* (2022); Kubota *et al.* (2023); Fry *et al.* (2023) focus on the dissipative nature of the natural quantum system to find a relationship between the existence of dissipation and the trainability of QRC. Kubota *et al.* (2023) analyzed the behavior of QRC driven by natural noise in quantum processing unit (QPU). The work of Martínez-Peña and Ortega (2023) further examined this research direction to describe ESP from the standpoint of a time-independent filter and dynamical systems that the authors call state affine systems (SAS).

1.2 Echo state property

Reservoir computing (RC) Nakajima and Fischer (2021) is a temporal information processing method that incorporates a dynamical system as a feature map generator. The ESP Gallicchio (2019); Yildiz *et al.* (2012); Manjunath and Jaeger (2013); Jaeger (2002) is known to be one of the necessary conditions for RC to perform information processing tasks through diminishing initial state sensitivity. The ESP is a condition of an input-driven dynamical system to have a fading memory. It can be experimentally checked by the following conditions.

Definition 1.1. Echo state property

Let a system state space be $\mathcal{S}$, an input space be $\mathcal{X}$ and a set of time index be $\mathcal{T}$. For an input-driven dynamical system with dynamical map $s_t = f(\{\mathbf{u}_\tau\}_{\tau<t}; s_0)$ such that $f : \mathcal{X}^{\mathcal{T}} \times \mathcal{S} \to \mathcal{S}$, where s_0 is the initial state and $\{\mathbf{u}_\tau\}$ is a sequence of inputs indexed by time τ, the ESP holds if and only if

$$\forall\{\mathbf{u}_\tau\}, \forall(s_0, s_0'),\ \|f(\{\mathbf{u}_\tau\}_{\tau\le t}; s_0) - f(\{\mathbf{u}_\tau\}_{\tau\le t}; s_0')\| \underset{t\to\infty}{\to} 0. \tag{1}$$

1.3 Contributions

In this paper, we define a two-way extension of conventional ESP. One direction is *non-stationary* ESP, which requires finite variance output signals relative to initial-state difference decay. Another direction is *subset/subspace* non-stationary ESP, which focuses on a situation in which a part of the system has a non-stationary ESP. However, the entire system is possibly initial state dependent. Our analysis includes numerical analysis of non-stationary ESP in a typical QRC setup with a specific type of the system Hamiltonian.

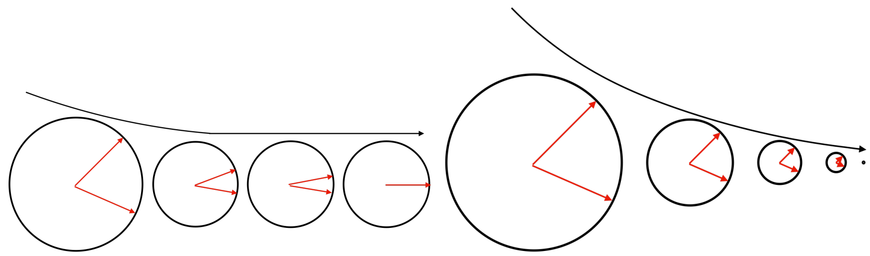
Our contributions are as follows.

- Defined non-stationary and subspace/subset versions of ESP, which could be practical for QRC and other non-conventional systems.

- Numerically showed a relationship between non-stationary ESP and the information processing capability of QRC.

2 Main results

2.1 Non-stationary ESP



(a) Non-stationary ESP (b) ESP that is not non-stationary ESP

Figure 2: Schematics of two types of ESPs. State space is illustrated as a circle, and different states are illustrated as red arrows. (a) Non-stationary ESP. State difference decays, but the state space remains finite. (b) With ESP but not with non-stationary ESP. The state difference decays as fast as the state decay.

Although ESP is supposed to work on stationary systems, a quantum system, for instance, is not always stationary, even if the system dynamics map does not depend on time. A trivial example is the case in which a Pauli noise exists. Let us depict an example where uniform depolarization of rate ϵ exists. When the depolarization is the only noise that exists in the system, the norm of the system state ρ measured using trace norm; $D(\rho, I)$ vanishes when $t \to \infty$, namely, $D(\rho, I) \propto (1-\epsilon)^t$. In this case, ESP does not mean forgetting memory, because the state difference $\|f(\{\mathbf{u}_\tau\}_{\tau\le t}; s_0) - f(\{\mathbf{u}_\tau\}_{\tau\le t}; s_0')\|$ relative to $\|f(\{\mathbf{u}_\tau\}_{\tau\le t}; s_0')\|$ does not change. The following modified definition of ESP was made to handle such a non-stationary system in our analysis.

Definition 2.1. Non-stationary ESP

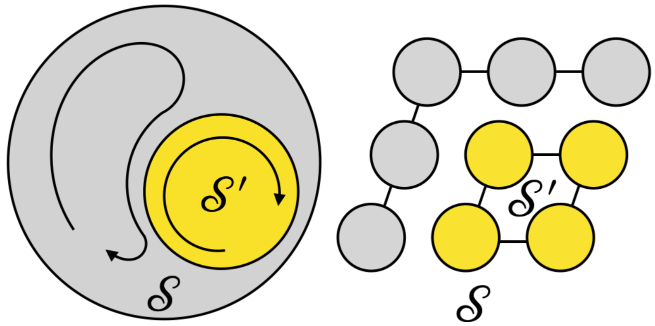
Given a system dynamics $f : \mathcal{X}^{\mathcal{T}} \times \mathcal{S} \to \mathcal{S}$, where $\mathcal{S}$ is bounded, f has *non-stationary* ESP if the following condition holds:

$$\begin{aligned}&\forall\{\mathbf{u}_\tau\},\ \forall(s_0, s_0'),\ \exists w \in \mathbb{N} < +\infty \text{ s.t.}\\ &\liminf_{t\to\infty}\left[\mathrm{Var}_w^t(\{\mathbf{u}_\tau\})\right] > 0 \Rightarrow \frac{\|f(\{\mathbf{u}_\tau\}_{\tau\le t}; s_0) - f(\{\mathbf{u}_\tau\}_{\tau\le t}; s_0')\|}{\sqrt{\min\left[\overline{\mathrm{Var}}_w^t(f; s_0), \overline{\mathrm{Var}}_w^t(f; s_0')\right]}} \underset{t\to\infty}{\to} 0,\end{aligned} \tag{2}$$

where

$$\begin{aligned}\mathrm{E}_w^t(\{\mathbf{u}_\tau\}) &\equiv \frac{1}{w}\sum_{k=0}^{w-1}\mathbf{u}_{t-k}\\ \mathrm{Var}_w^t(\{\mathbf{u}_\tau\}) &\equiv \mathrm{E}_w^t\left(\{\|\mathbf{u}_\tau - \overline{\mathrm{E}}_w^t(\{\mathbf{u}_\tau\})\|^2\}\right)\\ \overline{\mathrm{E}}_w^t(f; s_0) &\equiv \frac{1}{w}\sum_{k=0}^{w-1} f(\{\mathbf{u}_\tau\}_{\tau\le t-k}; s_0)\\ \overline{\mathrm{Var}}_w^t(f; s_0) &\equiv \overline{\mathrm{E}}_w^t\left(\|f - \overline{\mathrm{E}}_w^t(f, s_0)\|^2; s_0\right).\end{aligned} \tag{3}$$

The normalizing part on the denominator makes a non-stationary system, such as one that has a strong depolarizing channel, not satisfy the condition. Therefore, checking the system property for machine learning tasks is helpful. Additionally, it follows that if non-stationary ESP holds, then ESP holds.



(a) Subspace ESP (b) Subset ESP

Figure 3: (a) In the subspace ESP, the full state space $\mathcal{S}$ has an invariant subspace $\mathcal{S}' \subseteq \mathcal{S}$ under the input-driven system dynamics f. Furthermore, $\mathcal{S}'$ satisfies the non-stationary ESP under $f|_{\mathcal{S}'}$. (b) In the subset ESP, the subspace is limited to a subset of the $\mathcal{S}$ element.

2.2 Subspace and subset non-stationary ESP

In the RC setup, we can post-process output signals from the reservoir. Simple post-processing methods include linear transformation and *subset selection*. Here, *subset selection* denotes the selection of $m \leq n$ elements from the system output $x \in \mathbb{R}^n$.

For instance, if an invariant subspace of the input-driven dynamics exists, that subsystem holds non-stationary ESP. Then, we have non-stationary ESP-compatible output signals by post-projecting the outputs to that invariant subspace. On the other hand, if the system dynamics has a disjoint structure among its elements and there exists a non-stationary ESP-compatible subset of the outputs, then we can post-restrict them to that subset to obtain non-stationary ESP-compatible output signals.

By following the observation above, we now formally define a weak version of non-stationary ESP to preclude a situation in which only a part of the system is initial-state sensitive.

The *subspace* non-stationary ESP holds if only a linear system subspace holds non-stationary ESP.

Definition 2.2. Subspace non-stationary ESP

Given a system dynamics $f : \mathcal{X}^{\mathcal{T}} \times \mathcal{S} \to \mathcal{S}$ where $\mathcal{S}$ is bounded, f has *subspace non-stationary ESP* if there exists a linear transformation $\mathbb{P} : \mathcal{S} \to \mathcal{S}'$ such that $\mathcal{S}' \subseteq \mathcal{S}$ and $\mathbb{P} \circ f$ holds the non-stationary ESP.

Specifically, to treat cases in which a *subset* $\mathcal{S}'$ of $\mathcal{S}$ holds the non-stationary ESP, we define the following version.

Definition 2.3. Subset non-stationary ESP

Given a system dynamics $f : \mathcal{X}^{\mathcal{T}} \times \mathcal{S} \to \mathcal{S}$ where $\mathcal{S}$ is bounded. f has *subset non-stationary ESP* if there exists a subset selection procedure $\mathcal{P} : \mathcal{S} \to \mathcal{S}'$ such that $\mathcal{S}' \leq \mathcal{S}$ and $\mathcal{P} \circ f$ holds the non-stationary ESP.

The expression $A \leq B$ denotes that A is a non-void element-wise subset of B. If $A \equiv \mathbb{R}^n$, then $B = \mathbb{R}^m$ $(m \leq n)$ for instance.

It follows that if subset non-stationary ESP holds, then subspace non-stationary ESP holds because we can define $\mathbb{P}$ as a diagonal matrix such that it has 1 in the dimension included in $\mathcal{S}'$ and 0 otherwise. In addition, if non-stationary ESP holds, then every element of the system state has non-stationary ESP. Therefore, the subset non-stationary ESP holds. These relationships can be written as

$$\text{NS-ESP} \subsetneq \text{Subset NS-ESP} \subsetneq \text{Subspace NS-ESP}. \tag{4}$$

More generally, we have the inclusion relationship of the ESP variants, as seen in Fig. 1.

The definition of the subset non-stationary ESP is natural for practical QRC because we can select any observable for our system output. That is usually a subset of all Pauli strings or a linear combination of them, which can be reconstructed by measuring some of the observables.

If the subset non-stationary ESP holds, linear regression with sparsity regularization likely yields a stationary result for its task because including an ESP-incompatible subset of the signal into the regression should result in performance degradation.

Using the definitions above, we target systems in which some parts of the system have ESP while the remaining portion does not. An example of such a system in a quantum case is when the system dynamics is a tensor product of unitary and dissipative evolution. Ensuring a subset non-stationary ESP guarantees that such a system can be used as a reservoir with a simple transformation of output signals.

2.3 Non-stationary ESP of QRC

We have numerically examined whether the defined non-stationary ESP corresponds to temporal information processing capability using nonlinear autoregressive moving-average (NARMA) tasks Atiya and Parlos (2000). A NARMA sequence $\{\mathbf{y}_t\}$ of order k, given an input sequence $\{\mathbf{u}_t\}$, is defined as a nonlinear combination of $\{\mathbf{y}_t\}$ and $\{\mathbf{u}_t\}$ in the past:

$$\mathbf{y}_t = 0.3\mathbf{y}_{t-1} + 0.05\mathbf{y}_{t-1}\sum_{i=t-k}^{t-1}\mathbf{y}_i + 1.5\mathbf{u}_{t-1}\mathbf{u}_{t-k} + 0.1. \tag{5}$$

We use the QRC of the following Sherrington–Kirkpatrick (SK) Hamiltonian Sherrington and Kirkpatrick (1975) with external field:

$$H = \sum_{i>j=1}^{N} J_{ij}\sigma_i^x\sigma_j^x + \frac{1}{2}\sum_{i=1}^{N} h_i\sigma_i^z, \tag{6}$$

where N is the number of qubits in the system, σ_i^x and σ_i^z are Pauli X and Y operator of i-th qubit, respectively, J_{ij} and h_i are sampled from some distribution. The input sequence $\mathcal{U} \equiv \{\mathbf{u}_t\}_{t\in\mathcal{T}} \in \mathbb{R}^{|\mathcal{T}|}$ are fed into the reservoir using the following input encoding method named *reset-input encoding*:

$$\begin{aligned}\rho' &= \mathcal{E}(\rho, \mathbf{u}; \theta)\\ &= \mathrm{tr}_{\mathrm{A}}(\rho) \otimes \sigma_{\mathrm{A}}(\mathbf{u}; \theta),\end{aligned} \tag{7}$$

where A is the subsystem that we use for qubit state replacement by subsystem state $\sigma_{\mathrm{A}}(\mathbf{u})$ of form

$$\sigma_A(\mathbf{u}; \theta) \equiv U(\mathbf{u}; \theta)\left(|0\rangle\langle 0|^{\otimes |A|}\right)U^\dagger(\mathbf{u}; \theta). \tag{8}$$

Therefore, the overall state update is

$$\rho_{t+1} = e^{-iH}\mathcal{E}(\rho_t, \mathbf{u}_t; \theta)e^{iH}. \tag{9}$$

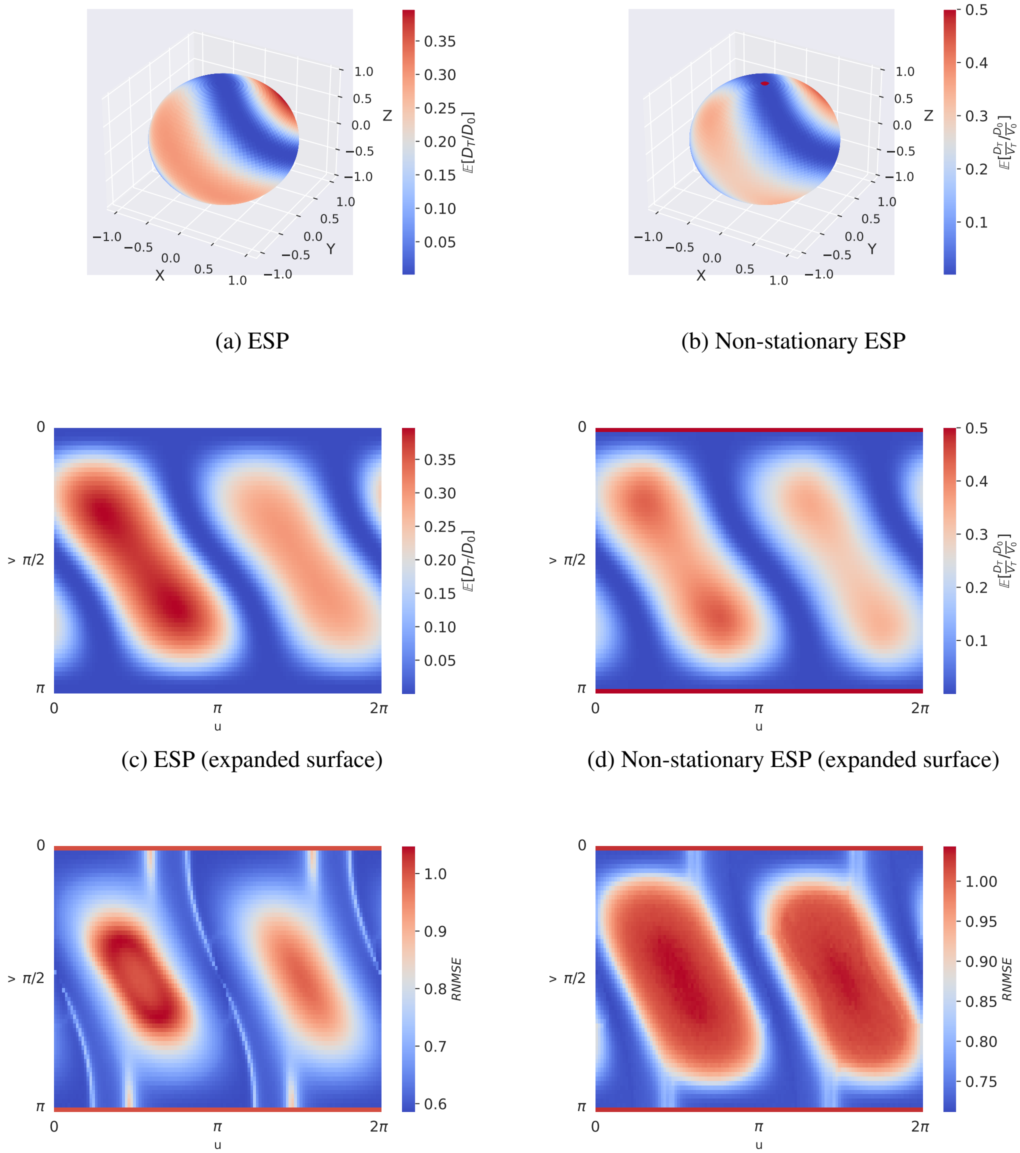
Here, we parametrize the input encoding unitary $U(\cdot\,; \theta)$ by a parameter θ to explore different configurations to ensure that the QRC has different non-stationary ESP, while system Hamiltonian H is fixed. The actual experiment was done on a 2-qubit setup, with $|A|$ = 1, while θ stands for the axis of single-qubit rotation in the Bloch sphere.

Reservoir output signals are $\left\{\mathrm{tr}(P\rho_t) \mid P \in \{I, X, Y, Z\}^{\otimes N}\right\}$ for each t. The first 80% of $\{\mathbf{u}_t\}$ are used to fit a linear regression to optimize MSE against the first 80% of $\{\mathbf{y}_t\}$. The test results are computed using remaining 20% of the sequences. We have 5 NARMA sequences for each order k and the results in Fig.4 are averaged over the sequences. We can observe that the RNMSE results of both NARMA2 and NARMA10 tasks in Fig. 4e and Fig. 4f, respectively, almost perfectly corresponds to the non-stationary ESP results in Fig. 4d. It should be noted that the red line at the top and bottom of Fig. 4d, which corresponds to the large RNMSE in Fig. 4e and Fig. 4f, cannot be reproduced by conventional ESP value Eq. (1), as shown in Fig. 4c.

3 Conclusion

This paper proposes a non-stationary and subspace/subset ESPs that are considered helpful in real-world RC scenarios. As a concrete application, we have numerically analyzed a QRC with a well-known SK Hamiltonian and a reset-input encoding method. We found a good correspondence between non-stationary ESP and information processing capability using NARMA tasks.

In a follow-up study, we conducted an in-depth theoretical analysis of non-stationary ESP to enhance our understanding of the QRC dynamics Kobayashi *et al.*. Our theoretical framework provides novel perspectives for the practical design of QRC and other non-stationary RC systems.



(a) ESP

(b) Non-stationary ESP

(c) ESP (expanded surface)

(d) Non-stationary ESP (expanded surface)

(e) NARMA2 RNMSE (expanded surface)

(f) NARMA10 RNMSE (expanded surface)

Figure 4: Grid search result of input rotation axis θ for a QRC with SK Hamiltonian with external field and qubit-reset for input encoding. Each spherical coordinate corresponds to the input rotation axis θ as described in the main text. Expanded surface plot is the plainer surface of the Bloch sphere, in which top line and bottom line corresponds to north and south pole, respectively. a,c) ESP. b, d) Non-stationary ESP, where the value is cut-off to be upper bounded by 0.5. $D_t \equiv D(\rho_t^{(i)}, \rho_t^{(j)})$ stands for the trace distance of two different states starting from different initial states $\rho_i^{(0)}$ and $\rho_j^{(0)}$ but fed the same input sequence. V_t stands for the denominator of Eq. (2) at time t. At $t = 0$, variance calculation is done looking forward instead of backward for simplicity. e-f) Root normalized mean squared error (RNMSE) of e) NARMA2. f) NARMA10 tasks with the same setup. Our experiments are done with $T = 200$ and $w = 10$ where T and w are variables defined in Eq. (2).

References

G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose, Neural Networks **115**, 100 (2019).

K. Nakajima, Japanese Journal of Applied Physics **59**, 060501 (2020).

J. Preskill, Quantum **2**, 79 (2018).

J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New Journal of Physics **18**, 023023 (2016).

K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Physical Review A **98**, 10.1103/physreva.98.032309 (2018).

K. Fujii and K. Nakajima, Physical Review Applied **8**, 10.1103/physrevapplied.8.024030 (2017).

S. Ghosh, A. Opala, M. Matuszewski, T. Paterek, and T. C. H. Liew, npj Quantum Information **5**, 35 (2019a).

M. Negoro, K. Mitarai, K. Fujii, K. Nakajima, and M. Kitagawa, Machine learning with controllable quantum dynamics of a nuclear spin ensemble in a solid (2018), arXiv:1806.10910 [quant-ph] .

S. Ghosh, T. Paterek, and T. C. H. Liew, Phys. Rev. Lett. **123**, 260404 (2019b).

J. Chen, H. I. Nurdin, and N. Yamamoto, Physical Review Applied **14**, 10.1103/physrevapplied.14.024065 (2020).

J. Nokkala, R. Martínez-Peña, G. L. Giorgi, V. Parigi, M. C. Soriano, and R. Zambrini, Communications Physics **4**, 10.1038/s42005-021-00556-w (2021).

L. C. G. Govia, G. J. Ribeill, G. E. Rowlands, H. K. Krovi, and T. A. Ohki, Phys. Rev. Res. **3**, 013077 (2021).

M. Spagnolo, J. Morris, S. Piacentini, M. Antesberger, F. Massa, A. Crespi, F. Ceccarelli, R. Osellame, and P. Walther, Nature Photonics **16**, 318 (2022).

R. Martínez-Peña and J.-P. Ortega, Physical Review E **107**, 10.1103/physreve.107.035306 (2023).

Y. Suzuki, Q. Gao, K. C. Pradel, K. Yasuoka, and N. Yamamoto, Scientific Reports **12**, 1353 (2022).

A. Sannia, R. Martínez-Peña, M. C. Soriano, G. L. Giorgi, and R. Zambrini, Dissipation as a resource for quantum reservoir computing (2022), arXiv:2212.12078 [quant-ph] .

T. Kubota, Y. Suzuki, S. Kobayashi, Q. H. Tran, N. Yamamoto, and K. Nakajima, Phys. Rev. Res. **5**, 023057 (2023).

D. Fry, A. Deshmukh, S. Y.-C. Chen, V. Rastunkov, and V. Markov, Optimizing quantum noise-induced reservoir computing for nonlinear and chaotic time series prediction (2023), arXiv:2303.05488 [quant-ph] .

K. Nakajima and I. Fischer, eds., *Reservoir Computing: Theory, Physical Implementations, and Applications* (Springer Singapore, Singapore, 2021).

C. Gallicchio, in *27th European Symposium on Artificial Neural Networks* (2019).

I. B. Yildiz, H. Jaeger, and S. J. Kiebel, Neural networks **35**, 1 (2012).

G. Manjunath and H. Jaeger, Neural computation **25**, 671 (2013).

H. Jaeger, *Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the "echo state network" approach*, Vol. 5 (GMD-Forschungszentrum Informationstechnik Bonn, 2002).

A. Atiya and A. Parlos, IEEE Transactions on Neural Networks **11**, 697 (2000).

D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).

S. Kobayashi, Q. H. Tran, and K. Nakajima, In preparation .