Learning Temporal Quantum Tomography

Quoc Hoan Tran[1,*] and Kohei Nakajima[1,2,†]
[1]Graduate School of Information Science and Technology, The University of Tokyo, Tokyo 113-8656, Japan
[2]Next Generation Artificial Intelligence Research Center, The University of Tokyo, Tokyo 113-8656, Japan

(Received 5 April 2021; revised 11 October 2021; accepted 30 November 2021; published 22 December 2021)

Quantifying and verifying the control level in preparing a quantum state are central challenges in building quantum devices. The quantum state is characterized from experimental measurements, using a procedure known as tomography, which requires a vast number of resources. However, tomography for a quantum device with temporal processing, which is fundamentally different from standard tomography, has not been formulated. We develop a practical and approximate tomography method using a recurrent machine learning framework for this intriguing situation. The method is based on repeated quantum interactions between a system called quantum reservoir with a stream of quantum states. Measurement data from the reservoir are connected to a linear readout to train a recurrent relation between quantum channels applied to the input stream. We demonstrate our algorithms for representative quantum learning tasks, followed by the proposal of a quantum memory capacity to evaluate the temporal processing ability of near-term quantum devices.

DOI: 10.1103/PhysRevLett.127.260401

Introduction.—The impressive progress in realizing quantum-enhanced technologies places a demand on the characterization and validation of quantum hardware. One of the most quintessential parts of building quantum devices is quantum process tomography (QPT), which is used in verifying quantum devices via the reconstruction of an unknown quantum channel from measurement data [1,2]. Standard QPT approaches, which have been focused recently on small system size [3–7], assume the quantum device processes input states separately in a time-independent manner. In the envisioned picture of quantum time-series processing, the quantum device may output the states in a sequence where the current output depends on the past inputs and outputs. For example, the quantum device may generate temporal and input-dependent noise or fluctuations, which may have effects on the output states [8,9]. Moreover, optical quantum states defined in temporal modes can be manipulated in a time-dependent and input-dependent manner using nonlinear optical processes [10,11]. Performing tomography for such devices differs from standard QPT, because the memory effects need to be taken into account.

Given a sequence of quantum states $\beta_1, \beta_2, \ldots$ in a D_A-dimensional Hilbert space, a quantum device processes this sequence via a temporal map $\mathcal{F}$ to output quantum states $\mathcal{F}(\beta_1), \mathcal{F}(\beta_2), \ldots$ in a D_B-dimensional Hilbert space with a temporal dependency behavior: $\mathcal{F}(\beta_n)$ only depends on a finite input history. An intriguing example is the temporal depolarizing channel $\mathcal{F}(\beta_n) = p_n(I/D) + (1 - p_n)\beta_n$, which replaces β_n with a completely mixed state I/D with probability p_n and leaves the state untouched otherwise ($D_A = D_B = D$ for notational simplicity). The temporal dependency can be established if p_n depends on the recent inputs. As similar to the model of quantum channels with memory [12,13], we model $\mathcal{F}$ via a stream of quantum channels $\{\Omega_n\}$ that applies to $\{\beta_n\}$ and introduces the correlations between outputs. $\mathcal{F}$ can be considered as a coherent superposition or a convex mixture of channels at different times [Fig. 1(a)]. The output of Ω_n is independent of future inputs; it is determined by finite input history and past channels to make the temporal dependency in $\mathcal{F}$. For example, Ω_n corresponds to the experimental studies on quantum processors where the dynamics of decoherence effects act as time-varying quantum channels [9,14–18]. The effect of previous quantum inputs to the output has also been demonstrated experimentally for optical fiber channels [19,20]. Here, $\mathcal{F}$ envisions a typical case in the future realization of quantum communication and quantum internet [21–23] where quantum data can be transmitted via time-dependent and delayed channels.

In this Letter, we propose a supervised learning framework to perform the approximate tomography of $\mathcal{F}$. A naive approach is to perform state tomography of $\mathcal{F}(\beta_n)$ for every n. This requires many repetition experiments on copies of $\mathcal{F}(\beta_n)$ and the inversion of a huge linear system for every n [24]. Our idea is to simplify the experimental protocol and reduce the implementation cost under the assumption of the temporal dependency in $\mathcal{F}$. We assume that it is possible to perform state tomography at some time steps as $\mathcal{F}(\beta_1), \ldots, \mathcal{F}(\beta_L)$. We consider a quantum system $\mathcal{S}$, called a *quantum reservoir* (QR), interacting with the input stream belonging to an auxiliary system $\mathcal{E}$. Each β_n interacts for a certain time with $\mathcal{S}$ before being replaced by another one. Between two consecutive interactions by β_n

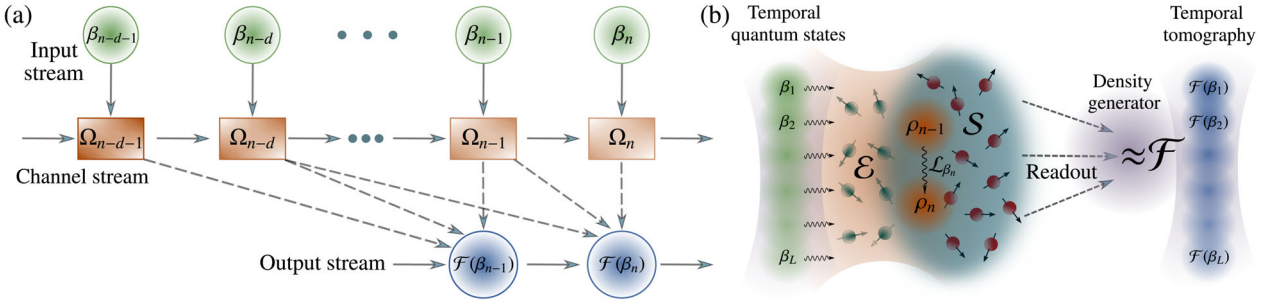

FIG. 1. Our framework can learn the tomography of a device that is supposed to implement an unknown temporal quantum map $\mathcal{F}$; or emulate a predefined $\mathcal{F}$. (a) A stream of quantum channels $\{\Omega_n\}$ applies to the input stream $\{\beta_n\}$ where each channel's output is determined by finite input history and past channels. The device's output is a function of input history such as a coherent superposition or a convex mixture of channels. (b) Our framework consists of a quantum reservoir $\mathcal{S}$ interacting with $\{\beta_n\}$ with memory effects. The reservoir's internal state is evolved via CPTP maps $\mathcal{L}_{\beta_n}$, which transfer the information from the input states to the reservoir. Measurement results in $\mathcal{S}$ are used in a readout layer to reconstruct $\mathcal{F}$.

and β_{n+1}, the QR's internal evolution is described by a completely positive and trace-preserving (CPTP) map $\mathcal{L}_{\beta_n}$ whose role is to effectively transfer the information of β_n from $\mathcal{E}$ into $\mathcal{S}$. With an effective setting of these CPTP maps, the QR's state after applying $\mathcal{L}_{\beta_n}$ will depend more on the recent past inputs than distant past inputs. This scheme ensures the fading memory property [50], which is the ability to retain information about recent inputs for proper learning of functions of the past inputs. Therefore, the results of measurements in $\mathcal{S}$ can be used as high-dimensional quantum features to train a regression model to output density matrices that approximate $\mathcal{F}(\beta_1), \ldots, \mathcal{F}(\beta_L)$ [Fig. 1(b)]. After this procedure, we are able to reconstruct $\mathcal{F}(\beta_n)$ for $n > L$ from the trained parameters by performing a single measurement protocol in $\mathcal{S}$.

Our framework is a quantum extension of classical reservoir computing (RC) to perform quantum tasks (Sec. I A-B in the Supplemental Material [24]). The crucial principle of RC is to represent the input sequence by feeding the input into a dynamical system, called the reservoir, to encode all relevant nonlinear dynamics in high-dimensional trajectories [51–54]. Our proposal exploits quantum dynamics as a reservoir in the time-series processing of quantum data. This idea develops the initial proposal of harnessing disordered quantum dynamics for machine learning with classical time-series data [55–57]. While the RC approaches in tomography tasks focused on a static quantum state [58–60], our approach can process time-dependent quantum states. We further propose the concept of quantum memory capacity to uncover the temporal processing ability of near-term quantum devices.

Model.—Assume that the initial state of the coupled system $(\mathcal{S}, \mathcal{E})$ is a product state $\varrho = \rho \otimes \beta$, where ρ and β are the state of $\mathcal{S}$ and $\mathcal{E}$, respectively. The coupled system is evolved under a unitary evolution U and the state ρ of $\mathcal{S}$ is transformed via the CPTP *reduced dynamics map* $\mathcal{L}_\beta$, where $\mathcal{L}_\beta(\rho) = \mathrm{Tr}_{\mathcal{E}}[U(\rho \otimes \beta)U^\dagger]$. The successive interactions are described as

$$\rho_n = \mathcal{L}_{\beta_n}(\rho_{n-1}) = \mathrm{Tr}_{\mathcal{E}}[U(\rho_{n-1} \otimes \beta_n)U^\dagger], \tag{1}$$

where ρ_n is the state of $\mathcal{S}$ for the nth interaction. We measure local observables $O_1, \ldots, O_K$ on ρ_n to obtain a high-dimensional feature vector called *reservoir state* $\boldsymbol{x}_n$. The kth element in $\boldsymbol{x}_n$ can be calculated as $x_{nk} = \mathrm{Tr}[O_k\rho_n] = \langle O_k \rangle_{\rho_n}$, which is the expectation of the measurement result via O_k. Between two inputs, M cycles of the unitary evolution are processed and each of them is followed by measurements. M is called the measurement multiplexity, thus we obtain MK elements in $\boldsymbol{x}_n$.

In the training stage, we are given an input sequence $\{\beta_1, \ldots, \beta_L\}$ and the target sequence $\{\hat{\boldsymbol{y}}_1, \ldots, \hat{\boldsymbol{y}}_L\}$, where $\hat{\boldsymbol{y}}_k$ is the real vector form to stack the real and imaginary elements of $\mathcal{F}(\beta_k)$. Our framework includes a readout map h, which is simply taken as a linear combination of the reservoir states as $\boldsymbol{y}_n = h(\boldsymbol{x}_n) = \boldsymbol{w}^\top \boldsymbol{x}_n$. Here, $\boldsymbol{w}$ is the parameter to be optimized by minimizing the mean-square error between $\boldsymbol{y}_n$ and $\hat{\boldsymbol{y}_n}$ over $n = 1, \ldots, L$. In the evaluation stage, we are given an input sequence $\{\beta_{L+1}, \ldots, \beta_{L+T}\}$ with the target $\{\hat{\sigma}_{L+1}, \ldots, \hat{\sigma}_{L+T}\}$, where $\hat{\sigma}_i = \mathcal{F}(\beta_i)$. The reconstructed output sequence is $\{\boldsymbol{y}_{L+1}, \ldots, \boldsymbol{y}_{L+T}\}$, which is rearranged in the matrix form $\{\sigma_{L+1}, \ldots, \sigma_{L+T}\}$ [61]. Since targets are density matrices, we use the fidelity $F(\rho, \sigma) = \mathrm{Tr}[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}]$ to estimate the reconstruction error. In error-free tomography, $F = 1$, and $F < 1$ otherwise. We calculate the root mean square of fidelities in the evaluation stage as $\mathrm{RMSF} = \sqrt{(1/T)\sum_{i=L+1}^{i=L+T} F^2(\hat{\sigma}_i, \sigma_i)}$.

The reservoir's response to the same input sequence may differ with different initial states of the reservoir and may result in the loss of reproducibility in the temporal processing. To prevent this effect, the time evolution of the reservoir must mostly depend on the input sequence after enough transient time. This property is known as the echo state property [51] in classical RC or the quantum echo state property (QESP) in quantum RC [63,64] to ensure the fading memory. Based on the spectrum of

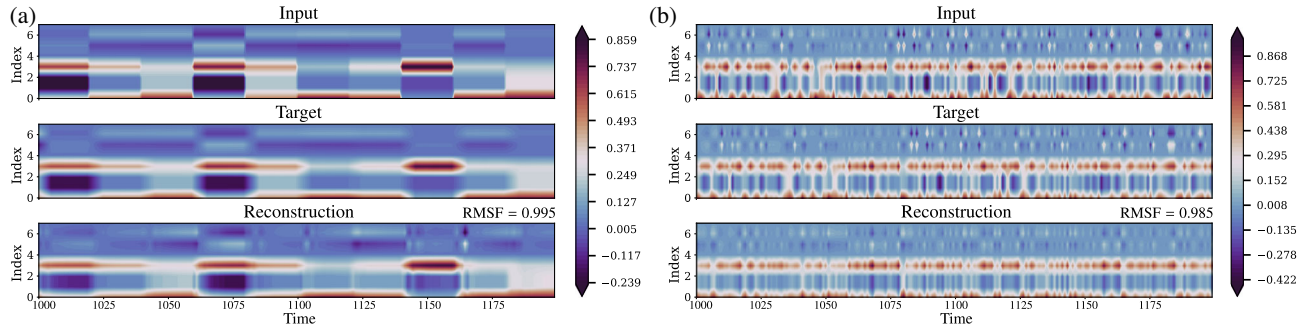

FIG. 2. Temporal tomography for (a) the quantum simple moving average filter and (b) the delayed depolarizing map with $d = 5$, $N_m = 5$, $N_e = 1$, $\alpha = 1.0$, $J/B = 1.0$, and $\tau B = 2.6$ [for (a)] and $\tau B = 2.0$ [for (b)]. At each time point, the density matrix is vectorized by stacking the real and imaginary parts, where the range of values is indicated in the color bars.

reduced dynamics maps, we can evaluate the QESP timescale, which indicates the transient time to forget the QR's initial state before learning (Sec. II in the Supplemental Material [24]).

Results.—We present concrete applications of learning temporal tomography. We consider $(\mathcal{S}, \mathcal{E})$ as a closed system of the transverse field Ising model with the unitary $U = \exp(i\tau H)$, where $H = \sum_{i>j=1}^{N} J_{i,j}\hat{s}_i^x\hat{s}_j^x + B\sum_{j=1}^{N}\hat{s}_j^z$ is unchanged during interaction time τ. Here, B is the natural frequency and $\hat{s}_j^\gamma$ ($\gamma \in \{x, y, z\}$) are the Pauli operators measuring the qubit j along the γ direction. We consider the *power-law decaying* for $J_{ij} = J|i-j|^{-\alpha}/N(\alpha)$ with an interaction strength J, power coefficient α ($0 < \alpha < 3$), and $N(\alpha) = \sum_{i>j}|i-j|^{-\alpha}/(N-1)$ [65–67]. $\mathcal{E}$ includes the first N_e qubits where the remaining $N_m = N - N_e$ qubits form the reservoir $\mathcal{S}$.

In our demonstrations, the number of observables is set to $K = N_m$ if we select observables as spin projections $\hat{s}_j^z$ over the z axis for all j, and to $K = N_m(N_m+1)/2$ if we further select observables as two-spin correlations $\hat{s}_i^z\hat{s}_j^z$ for all $i < j$. We consider the time-dependent depolarizing quantum channel $\Omega_n(\beta) = p_n(I/D) + (1-p_n)\beta$. We introduce a temporal dependency in Ω_n by formulating p_n as the rth-order nonlinear sequence: $p_n = \kappa p_{n-1} + \eta p_{n-1}(\sum_{j=0}^{r-1} p_{n-j-1}) + \gamma u_{n-r+1}u_n + \delta$, where $r = 10$, $\kappa = 0.3$, $\eta = 0.04$, $\gamma = 1.5$, and $\delta = 0.1$. Here, $\{u_n\}$ is considered depending on $\{\beta_n\}$. It is randomly generated with the same random seed used to generate $\{\beta_n\}$ [68]. The QR can encode $\{\beta_n\}$ to reconstruct the output states, including the reconstruction of the nonlinear sequence $\{p_n\}$. This reconstruction is possible since observables in the QR become nonlinear functions of the input history due to the mixing of higher-order correlations in the quantum-chaotic dynamics.

We first consider $\mathcal{F}$ as a quantum simple moving average filter $\mathcal{F}(\beta_n) = [1/(d+1)]\sum_{i=0}^{d}\Omega_{n-i}(\beta_{n-i})$ [Fig. 2(a)] or a delayed depolarizing map $\mathcal{F}(\beta_n) = \Omega_{n-d}(\beta_{n-d})$ [Fig. 2(b)] for a delay $d \geq 0$. We refer to Sec. V B-C in the Supplemental Material [24] for tomography of coherent superposition of channels in times and temporal entanglers. The QR must memorize the previous inputs and learn the properties of previous channels. We set $d = 5$, $K = N_m = 5$, $N_e = 1$, and $M = 5$. Other model parameters are $\alpha = 1.0$, $J/B = 1.0$, and the normalized time $\tau B = 2.6$ [Fig. 2(a)] and $\tau B = 2.0$ [Fig. 2(b)]. The number of time steps used in the initial transients, training, and evaluation stages are 500, 500, and 200, respectively. At each time point, the density matrix is represented as a $2D^2$-dimensional vector ($D = 2^{N_e}$) by stacking the real and imaginary parts. In Fig. 2(a), the inputs jump to a new random quantum state every 20 time steps, thus introducing temporal dependencies between inputs. Alternately, we consider a sequence of i.i.d. random inputs in Fig. 2(b). The target sequences (middle panels in Fig. 2) in the evaluation stage can be almost perfectly reconstructed since the RMSF values are above 98% (bottom panels).

Next, we investigate the dependency of the task's performance on the QR's parameters. Figure 3 illustrates the tomography errors (1.0—RMSF) according to τB in the reconstruction of $\mathcal{F}(\beta_n) = \Omega_{n-1}(\beta_{n-1})$ at $N_e = 2$ and $N_m = 4$, 5 qubits. The transients, training, and evaluation time steps are 1000, 3000, and 1000, respectively. The errors are averaged over 10 different runs with random trials of the input sequence and initial state. The errors reduce quickly at low values of τB and then settle to the stable lower values. Table I presents the average errors along with their standard deviations at $\tau B = 10.0$ and $M = 5$. Particularly, with $N_m = 5$ and $K = 15$, the errors are lower than 1%, 4%, 6%, and 8% for $N_e = 1$, 2, 3, and 4

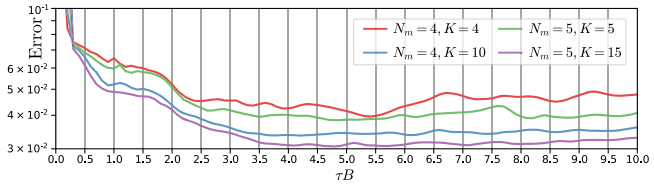

FIG. 3. Tomography errors according to τB for $\mathcal{F}(\beta_n) = \Omega_{n-1}(\beta_{n-1})$ with $\alpha = 1.0$ and $J/B = 1.0$.

TABLE I. Average and standard deviation (mean ± sd) of the tomography errors (%) for $\mathcal{F}(\beta_n) = \Omega_{n-1}(\beta_{n-1})$ of the baseline and our method at $\tau B = 10.0$ and $M = 5$.

(N_m, K)	$N_e = 1$	$N_e = 2$	$N_e = 3$	$N_e = 4$
(4,4)	0.9 ± 0.0	4.8 ± 0.1	8.0 ± 0.0	8.9 ± 0.1
(5,5)	0.9 ± 0.0	4.0 ± 0.1	7.6 ± 0.0	8.6 ± 0.1
(4,10)	0.9 ± 0.0	3.6 ± 0.1	6.4 ± 0.0	8.3 ± 0.1
(5,15)	0.9 ± 0.0	3.3 ± 0.0	5.6 ± 0.0	7.7 ± 0.0
Baseline	3.2 ± 0.1	9.2 ± 0.1	10.0 ± 0.1	10.2 ± 0.1

qubits, respectively. We further compare our method with a classical baseline method, in which we assume that a full tomography of input states can be obtained. Instead of using measurements, the reservoir state $\boldsymbol{x}_n$ is constructed directly from β_n by stacking the real and imaginary parts in the corresponding density matrix to construct the vector form. Table I shows that our method outperforms the classical baseline, which does not have memory effects.

We further investigate the *short-term memory* (STM) of the QR via the delay-reconstruction task $\mathcal{F}(\beta_n) = \beta_{n-d}$. The STM in the classical context is defined via the coefficient of determination to measure how much variance of the delay inputs can be recovered from outputs [70]. Since the input and output of our framework are density matrices, we define the d-delay STM of the QR by the squared distance correlation [71] between the output $\{\sigma_n\}$ and the target $\{\hat{\sigma}_n\} = \{\beta_{n-d}\}$:

$$\mathcal{R}^2(d) = \frac{\mathcal{V}^2(\{\sigma_n\}, \{\hat{\sigma}_n\})}{\sqrt{\mathcal{V}^2(\{\sigma_n\}, \{\sigma_n\})\mathcal{V}^2(\{\hat{\sigma}_n\}, \{\hat{\sigma}_n\})}}. \quad (2)$$

Here, $\mathcal{V}^2$ represents the squared distance covariance of two random sequences of density matrices [24]. $\mathcal{R}^2(d)$ is between 0 and 1, and it represents the QR's ability at the input β_n to reconstruct the previous input β_{n-d}. The behavior of $\mathcal{R}^2(d)$ is indicated by the forgetting curve, which approaches to a small value for large values of d, thus realizing the STM. We then define the quantum memory capacity, QMC $= \sum_{d=0}^{\infty} \mathcal{R}^2(d)$, to measure how much information of the delay input states can be recovered from output states, summed over all delays. If the value of QMC increases, so does the duration of quantum inputs that can be memorized via the QR. Quantifying QMC provides insights into the ability of the QR to reconstruct a temporal function of quantum inputs.

Figure 4(a) shows QMC as a function of τB broken down in values of d ($0 \leq d \leq 7$) with the model parameters $\alpha = 1$, $J/B = 1.0$, $K = N_m = 4$, $N_e = 2$, and $M = 1$ (other results in Sec. IV in the Supplemental Material [24]). $\mathcal{R}^2(d)$ is averaged over different runs with 10 random trials of the initial state and input sequence. The total $\sum_{d=0}^{d=7} \mathcal{R}^2(d)$ increases and obtains the peak value at the onset of the dynamical transition region ($1.5 < \tau B < 2.5$) of the reduced

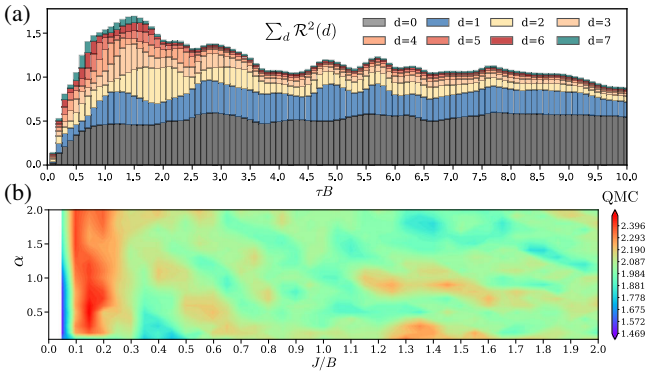

FIG. 4. (a) QMC broken down in delays d according to τB with $N_e = 2$, $N_m = 4$ with $\alpha = 1.0$ and $J/B = 1.0$. (b) The color map of QMC as the function of α and J/B at $\tau B = 10.0$.

dynamics map $\mathcal{L}_\beta$ [Fig. S2(a) in the Supplemental Material [24]]. We further examine the relation between QMC and other model parameters in Fig. 4(b), which displays the average QMC (calculated until $d_{\max} = 10$) as the function of α and J/B at $\tau B = 10.0$. Interestingly, QMC achieves highest values in the region $0.1 < J/B < 0.2$, which is referred to the dynamical transition in Figs. S3 and S6 in the Supplemental Material [24]. These observations can be explained by the difference in the eigenvalues' distribution of $\mathcal{L}_\beta$ [Figs. S2(a) and S3 in the Supplemental Material [24]]. The regime with eigenvalues concentrated near the border of the unit disk leads toward a unitary behavior. Therefore, only a little information of the input state β is remained after applying $\mathcal{L}_\beta$, which is unfavorable on temporal learning tasks. In contrast, the regime with eigenvalues concentrated near the center of the unit disk guarantees enough information in the input states to be entangled with the QR. Here, the dynamics becomes ergodic and the local observables become functions of a finite number of past inputs. We anticipate that QMC builds up first as the dynamics moves from more unitary to more ergodic regime and obtains the peak at the transition between these regimes [72].

Conclusion and discussion.—We formulate and propose the general framework for learning tomography of temporal quantum maps acting on quantum data. We establish the concept of quantum memory capacity, which opens opportunities in developing the theoretical magnitude on the quantum time-series processing [77].

The measurement protocol on the reservoir may lead to the effect of backaction, which is the problem of changes in quantum states due to measurement. Each physical implementation can have measures that may successfully work around this problem [78,79]. One can consider weak measurements on multiple copies of the same systems, such as a huge ensemble of identical molecules in a solid [55,78]. We can explore the implementation based on ion traps since it is possible to experimentally exploit nontrivial

degrees of freedom with the measurements of the spin projections and correlations [80,81] [82].

To perform tomography of the temporal quantum map depending on long distant past inputs, we need to increase the quantum memory capacity of the reservoir, which can be a bottleneck with more scale of resources such as qubits, observables, and measurements. Therefore, it can be helpful in the resource design if we can quantify in advance the required information processing ability of a temporal quantum map, such as how far and what combinations of the past inputs are processed in this map. This directly relates to the information processing framework in input-driven dynamical systems [84,85] but presents further challenges in the quantum context.

The authors acknowledge Shumpei Kobayashi for fruitful discussions. This work is supported by MEXT Quantum Leap Flagship Program (MEXT Q-LEAP) Grants No. JPMXS0118067394 and No. JPMXS0120319794.

*tran_qh@ai.u-tokyo.ac.jp
†k_nakajima@mech.t.u-tokyo.ac.jp

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* 10th ed (Cambridge University Press, USA, 2011).
[2] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, Quantum-process tomography: Resource analysis of different strategies, Phys. Rev. A **77**, 032322 (2008).
[3] J. L. O'Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph, and A. G. White, Quantum Process Tomography of a Controlled-NOT Gate, Phys. Rev. Lett. **93**, 080502 (2004).
[4] M. Riebe, K. Kim, P. Schindler, T. Monz, P. O. Schmidt, T. K. Körber, W. Hänsel, H. Häffner, C. F. Roos, and R. Blatt, Process Tomography of Ion Trap Quantum Gates, Phys. Rev. Lett. **97**, 220407 (2006).
[5] R. C. Bialczak, M. Ansmann, M. Hofheinz, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, H. Wang, J. Wenner, M. Steffen, A. N. Cleland, and J. M. Martinis, Quantum process tomography of a universal entangling gate implemented with Josephson phase qubits, Nat. Phys. **6**, 409 (2010).
[6] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White, Efficient Measurement of Quantum Dynamics via Compressive Sensing, Phys. Rev. Lett. **106**, 100401 (2011).
[7] L. C. G. Govia, G. J. Ribeill, D. Ristè, M. Ware, and H. Krovi, Bootstrapping quantum process tomography via a perturbative ansatz, Nat. Commun. **11**, 1084 (2020).
[8] E. Gehrig and O. G. Hess, Spatio-temporal dynamics and fluctuations in quantum dot lasers: Mesoscopic theory and modeling, in *Quantum Dot Devices and Computing*, edited by J. A. Lott, N. N. Ledentsov, K. J. Malloy, B. E. Kane, and T. W. Sigmon, International Society for Optics and Photonics Vol. 4656 (SPIE, San Jose, 2002), pp. 69–78, 10.1117/12.460803.
[9] J. E. Martinez, P. Fuentes, P. Crespo, and J. Garcia-Frias, Time-varying quantum channel models for superconducting qubits, npj Quantum Inf. **7**, 115 (2021).
[10] B. Brecht, D. V. Reddy, C. Silberhorn, and M. G. Raymer, Photon Temporal Modes: A Complete Framework for Quantum Information Science, Phys. Rev. X **5**, 041017 (2015).
[11] M. G. Raymer and I. A. Walmsley, Temporal modes in quantum optics: Then and now, Phys. Scr. **95**, 064002 (2020).
[12] D. Kretschmann and R. F. Werner, Quantum channels with memory, Phys. Rev. A **72**, 062323 (2005).
[13] F. Caruso, V. Giovannetti, C. Lupo, and S. Mancini, Quantum channels and memory effects, Rev. Mod. Phys. **86**, 1203 (2014).
[14] J. J. Burnett, A. Bengtsson, M. Scigliuzzo, D. Niepce, M. Kudra, P. Delsing, and J. Bylander, Decoherence benchmarking of superconducting qubits, npj Quantum Inf. **5**, 54 (2019).
[15] P. V. Klimov *et al.*, Fluctuations of Energy-Relaxation Times in Superconducting Qubits, Phys. Rev. Lett. **121**, 090502 (2018).
[16] S. Schlör, J. Lisenfeld, C. Müller, A. Bilmes, A. Schneider, D. P. Pappas, A. V. Ustinov, and M. Weides, Correlating Decoherence in Transmon Qubits: Low Frequency Noise by Single Fluctuators, Phys. Rev. Lett. **123**, 190502 (2019).
[17] Z. Wang, S. Shankar, Z. K. Minev, P. Campagne-Ibarcq, A. Narla, and M. H. Devoret, Cavity Attenuators for Superconducting Qubits, Phys. Rev. Applied **11**, 014031 (2019).
[18] A. Stehli, J. D. Brehm, T. Wolz, P. Baity, S. Danilin, V. Seferai, H. Rotzinger, A. V. Ustinov, and M. Weides, Coherent superconducting qubits from a subtractive junction fabrication process, Appl. Phys. Lett. **117**, 124005 (2020).
[19] J. Ball, A. Dragan, and K. Banaszek, Exploiting entanglement in communication channels with correlated noise, Phys. Rev. A **69**, 042324 (2004).
[20] K. Banaszek, A. Dragan, W. Wasilewski, and C. Radzewicz, Experimental Demonstration of Entanglement-Enhanced Classical Communication over a Quantum Channel with Correlated Noise, Phys. Rev. Lett. **92**, 257901 (2004).
[21] H. J. Kimble, The quantum internet, Nature (London) **453**, 1023 (2008).
[22] C. Simon, Towards a global quantum network, Nat. Photonics **11**, 678 (2017).
[23] S. Wehner, D. Elkouss, and R. Hanson, Quantum internet: A vision for the road ahead, Science **362**, eaam9288 (2018).
[24] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.127.260401 for detailed explanations of the reservoir computing, the quantum reservoir computing, the standard quantum process tomography, the temporal tomography, the convergence and metastability analysis of the quantum reservoir, the quantum memory capacity, and other results on the temporal tomography tasks, which include Refs. [25–49].
[25] J. Nokkala, R. Martínez-Peña, G. L. Giorgi, V. Parigi, M. C. Soriano, and R. Zambrini, Gaussian states of continuous-variable quantum systems provide universal and versatile reservoir computing, Commun. Phys. **4**, 53 (2021).
[26] L. C. G. Govia, G. J. Ribeill, G. E. Rowlands, H. K. Krovi, and T. A. Ohki, Quantum reservoir computing with a single nonlinear oscillator, Phys. Rev. Research **3**, 013077 (2021).

[27] S. Dasgupta, K. E. Hamilton, and A. Banerjee, Designing a NISQ reservoir with maximal memory capacity for volatility forecasting, arXiv:2004.08240.
[28] D. S. Abrams and S. Lloyd, Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors, Phys. Rev. Lett. **83**, 5162 (1999).
[29] B. I. Bantysh, A. Y. Chernyavskiy, and Y. I. Bogdanov, Quantum tomography benchmarking, Quantum Inf. Process. **20**, 339 (2021).
[30] Y. Quek, S. Fort, and H. K. Ng, Adaptive quantum state tomography with neural networks, npj Quantum Inf. **7**, 105 (2021).
[31] L. Bruneau, A. Joye, and M. Merkli, Infinite products of random matrices and repeated interaction dynamics, Ann. Inst. H. Poincaré Probab. Stat. **46**, 442 (2010).
[32] I. Nechita and C. Pellegrini, Random repeated quantum interactions and random invariant states, Probab. Theory Relat. Fields **152**, 299 (2012).
[33] R. Movassagh and J. Schenker, An ergodic theorem for homogeneously distributed quantum channels with applications to matrix product states, arXiv:1909.11769.
[34] R. Movassagh and J. Schenker, Theory of Ergodic Quantum Processes, Phys. Rev. X **11**, 041001 (2021).
[35] J. M. Steele, Kingman's subadditive ergodic theorem, Ann. Inst. H. Poincaré Probab. Stat. **25**, 93 (1989), http://www.numdam.org/item/AIHPB_1989__25_1_93_0/.
[36] H. Hennion, Limit theorems for products of positive random matrices, Ann. Probab. **25**, 1545 (1997).
[37] K. Macieszczak, M. Guţă, I. Lesanovsky, and J. P. Garrahan, Towards a Theory of Metastability in Open Quantum Dynamics, Phys. Rev. Lett. **116**, 240404 (2016).
[38] J. B. Lasserre, A trace inequality for matrix product, IEEE Trans. Autom. Control **40**, 1500 (1995).
[39] B. Misra and E. C. G. Sudarshan, The Zeno's paradox in quantum theory, J. Math. Phys. (N.Y.) **18**, 756 (1977).
[40] G. Chiribella, Perfect discrimination of no-signalling channels via quantum superposition of causal structures, Phys. Rev. A **86**, 040301(R) (2012).
[41] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. A. Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, Experimental superposition of orders of quantum gates, Nat. Commun. **6**, 7913 (2015).
[42] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, Experimental verification of an indefinite causal order, Sci. Adv. **3**, e1602589 (2017).
[43] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, Indefinite Causal Order in a Quantum Switch, Phys. Rev. Lett. **121**, 090503 (2018).
[44] K. Wei, N. Tischler, S.-R. Zhao, Y.-H. Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, Y.-A. Chen, B. C. Sanders, Q. Zhang, G. J. Pryde, F. Xu, and J.-W. Pan, Experimental Quantum Switching for Exponentially Superior Quantum Communication Complexity, Phys. Rev. Lett. **122**, 120504 (2019).
[45] Y. Guo, X.-M. Hu, Z.-B. Hou, H. Cao, J.-M. Cui, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and G. Chiribella, Experimental Transmission of Quantum Information using a Superposition of Causal Orders, Phys. Rev. Lett. **124**, 030502 (2020).
[46] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Phys. Rev. A **88**, 022318 (2013).
[47] D. Ebler, S. Salek, and G. Chiribella, Enhanced Communication with the Assistance of Indefinite Causal Order, Phys. Rev. Lett. **120**, 120502 (2018).
[48] D. Meschede, H. Walther, and G. Müller, One-Atom Maser, Phys. Rev. Lett. **54**, 551 (1985).
[49] B. T. H. Varcoe, S. Brattke, M. Weidinger, and H. Walther, Preparing pure photon number states of the radiation field, Nature (London) **403**, 743 (2000).
[50] S. Boyd and L. Chua, Fading memory and the problem of approximating nonlinear operators with volterra series, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **32**, 1150 (1985).
[51] H. Jaeger, The "echo state" approach to analysing and training recurrent neural networks with an erratum note, Tech. Rep. 34, German National Research Center for Information Technology, 2001, http://www.faculty.jacobs-university.de/hjaeger/pubs/EchoStatesTechRep.pdf.
[52] W. Maass, T. Natschläger, and H. Markram, Real-time computing without stable states: A new framework for neural computation based on perturbations, Neural Comput. **14**, 2531 (2002).
[53] M. Lukoševičius and H. Jaeger, Reservoir computing approaches to recurrent neural network training, Comput. Sci. Rev. **3**, 127 (2009).
[54] *Reservoir Computing: Theory, Physical Implementations, and Applications*, edited by K. Nakajima and I. Fischer (Springer Singapore, Singapore, 2021).
[55] K. Fujii and K. Nakajima, Harnessing Disordered-Ensemble Quantum Dynamics for Machine Learning, Phys. Rev. Applied **8**, 024030 (2017).
[56] K. Nakajima, K. Fujii, M. Negoro, K. Mitarai, and M. Kitagawa, Boosting Computational Power through Spatial Multiplexing in Quantum Reservoir Computing, Phys. Rev. Applied **11**, 034021 (2019).
[57] K. Fujii and K. Nakajima, Quantum reservoir computing: A reservoir approach toward quantum machine learning on near-term quantum devices, in *Reservoir Computing: Theory, Physical Implementations, and Applications*, edited by K. Nakajima and I. Fischer (Springer Singapore, Singapore, 2021), pp. 423–450.
[58] S. Ghosh, A. Opala, M. Matuszewski, T. Paterek, and T. C. Liew, Quantum reservoir processing, npj Quantum Inf. **5**, 35 (2019).
[59] S. Ghosh, A. Opala, M. Matuszewski, T. Paterek, and T. C. H. Liew, Reconstructing quantum states with quantum reservoir networks, IEEE Trans. Neural Netw. Learn. Syst. **32**, 3148 (2020).
[60] S. Ghosh, K. Nakajima, T. Krisnanda, K. Fujii, and T. C. H. Liew, Quantum Neuromorphic Computing with Reservoir Computing Networks, Adv. Quantum Technol. **4**, 2100053 (2021).
[61] For sufficient training samples, our framework can reconstruct the density matrices, which are positive semidefinite. However, due to statistical fluctuations, there are some cases

in which the reconstructed matrix A is not positive semidefinite. We project A onto the spectrahedron to obtain a positive semidefinite matrix $\hat{A}$ such that the trace of $\hat{A}$ is equal to 1 and the Frobenius norm between A and $\hat{A}$ is minimized [62].

[62] Y. Chen and X. Ye, Projection onto a simplex, arXiv:1101.6081.

[63] J. Chen and H. I. Nurdin, Learning nonlinear input–output maps with dissipative quantum systems, Quantum Inf. Process. **18**, 198 (2019).

[64] Q. H. Tran and K. Nakajima, Higher-order quantum reservoir computing, arXiv:2006.08999.

[65] D. Porras and J. I. Cirac, Effective Quantum Spin Systems with Trapped Ions, Phys. Rev. Lett. **92**, 207901 (2004).

[66] K. Kim, M.-S. Chang, R. Islam, S. Korenblit, L.-M. Duan, and C. Monroe, Entanglement and Tunable Spin-Spin Couplings between Trapped Ions using Multiple Transverse Modes, Phys. Rev. Lett. **103**, 120502 (2009).

[67] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt, and C. F. Roos, Quasiparticle engineering and entanglement propagation in a quantum many-body system, Nature (London) **511**, 202 (2014).

[68] The channels Ω_n can be considered quantum noises applying to the input states where there is a temporal correlation in these noises. The sequence $\{p_n\}$ resembles the NARMA benchmark [69], which is commonly used for evaluating the computational capability of temporal processing with long time dependence. Here, $\{u_n\}$ is a random sequence of scalar values in [0, 1] but is rescaled into [0, 0.2] before creating p_n to set p_n into the stable range in [0, 1].

[69] A. Atiya and A. Parlos, New results on recurrent network training: Unifying the algorithms and accelerating convergence, IEEE Trans. Neural Netw. Learn. Syst. **11**, 697 (2000).

[70] H. Jaeger, Short term memory in echo state networks, Tech. Rep. 152, German National Research Center for Information Technology, 2002, http://publica.fraunhofer.de/eprints/urn_nbn_de_0011-b-731310.pdf.

[71] G. J. Székely, M. L. Rizzo, and N. K. Bakirov, Measuring and testing dependence by correlation of distances, Ann. Stat. **35**, 2769 (2007).

[72] For temporal classical tasks in quantum spin networks, similar observations [73] have also been investigated recently to address the impact of the transition between localization and thermalization manifest. Our theoretical investigation is not limited to quantum spin networks, but is more generally associated with the spectrum analysis of reduced dynamics maps. Furthermore, the intriguing results in QMC under varying model parameters remind us of the well-known phenomenon in classical RC called the edge of chaos, where in some situations, a memory capacity achieves the maximum values at the edge of stability between different dynamics regimes [74–76].

[73] R. Martínez-Peña, G. L. Giorgi, J. Nokkala, M. C. Soriano, and R. Zambrini, Dynamical Phase Transitions in Quantum Reservoir Computing, Phys. Rev. Lett. **127**, 100502 (2021).

[74] N. Bertschinger and T. Natschläger, Real-time computation at the edge of chaos in recurrent neural networks, Neural Comput. **16**, 1413 (2004).

[75] T. Toyoizumi and L. F. Abbott, Beyond the edge of chaos: Amplification and temporal integration by recurrent networks in the chaotic regime, Phys. Rev. E **84**, 051908 (2011).

[76] T. Haruna and K. Nakajima, Optimal short-term memory before the edge of chaos in driven random recurrent networks, Phys. Rev. E **100**, 062312 (2019).

[77] P. Mujal, R. Martínez-Peña, J. Nokkala, J. García-Beni, G. L. Giorgi, M. C. Soriano, and R. Zambrini, Opportunities in quantum reservoir computing and extreme learning machines, Adv. Quantum Technol. **4**, 2100027 (2021).

[78] M. Negoro, K. Mitarai, K. Nakajima, and K. Fujii, Toward NMR quantum reservoir computing, in *Reservoir Computing: Theory, Physical Implementations, and Applications*, edited by K. Nakajima and I. Fischer (Springer Singapore, Singapore, 2021), pp. 451–458.

[79] J. Chen, H. I. Nurdin, and N. Yamamoto, Temporal Information Processing on Noisy Quantum Computers, Phys. Rev. Applied **14**, 024065 (2020).

[80] P. Richerme, Z.-X. Gong, A. Lee, C. Senko, J. Smith, M. Foss-Feig, S. Michalakis, A. V. Gorshkov, and C. Monroe, Non-local propagation of correlations in quantum systems with long-range interactions, Nature (London) **511**, 198 (2014), 1401.5088.

[81] J. Zhang, G. Pagano, P. W. Hess, A. Kyprianidis, P. Becker, H. Kaplan, A. V. Gorshkov, Z.-X. Gong, and C. Monroe, Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator, Nature (London) **551**, 601 (2017).

[82] In this platform, one can surpass the overhead on the large number of measurements repeated at each time step by utilizing the shadow tomography protocol to estimate many observables from a reasonable number of projective measurements [83].

[83] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).

[84] J. Dambre, D. Verstraeten, B. Schrauwen, and S. Massar, Information processing capacity of dynamical systems, Sci. Rep. **2**, 514 (2012).

[85] T. Kubota, H. Takahashi, and K. Nakajima, Unifying framework for information processing in stochastically driven dynamical systems, Phys. Rev. Research **3**, 043135 (2021).