Variational quantum process tomography of unitaries

Shichuan Xue,[1] Yong Liu [1,2] Yang Wang,[1] Pingyu Zhu,[1] Chu Guo,[3,4,*] and Junjie Wu [1,†]

[1]*Institute for Quantum Information & State Key Laboratory of High Performance Computing, College of Computer Science and Technology, National University of Defense Technology, Changsha 410073, China*

[2]*College of Information and Communication, National University of Defense Technology, Xi'an 710006, China*

[3]*Henan Key Laboratory of Quantum Information and Cryptography, Zhengzhou, Henan 450000, China*

[4]*Key Laboratory of Low-Dimensional Quantum Structures and Quantum Control of Ministry of Education, Department of Physics and Synergetic Innovation Center for Quantum Effects and Applications, Hunan Normal University, Changsha 410081, China*

Quantum process tomography is an experimental technique to fully characterize an unknown quantum process. Standard quantum process tomography suffers from exponentially scaling of the number of measurements with the increasing system size. In this work, we put forward a quantum machine learning algorithm which approximately encodes the unknown unitary quantum process into a relatively shallow depth parametric quantum circuit. We demonstrate our method by reconstructing the unitary quantum processes resulting from the quantum Hamiltonian evolution and random quantum circuits up to eight qubits. Results show that those quantum processes could be reconstructed with high fidelity, while the number of input states required are at least two orders of magnitude less than required by the standard quantum process tomography.

I. INTRODUCTION

Quantum process tomography is an indispensable technique in quantum information processing to fully characterize an unknown quantum process [1]. It is increasingly pivotal in identifying and verifying the performance of a quantum device and its dynamics when the system goes larger.

Standard quantum process tomography (SQPT) works by preparing an informationally complete set of input states and then performing the standard quantum state tomography on the corresponding output quantum states [2–4]. As a result, the total number of quantum measurements scales as 4^{2n} for an n-qubit quantum process. The exponential growth severely limits the problem size on which SQPT can be feasibly applied. Currently, SQPT has only been experimentally implemented up to three qubits [5–10]. In the meantime, with the rapid development of quantum computing hardware [11–13], scalable quantum process tomography schemes are in great demand.

Various schemes have been proposed to alleviate the exponential scaling problem of SQPT. For example, ancilla-assisted process tomography could reduce the exponential number of input states to a single entangled state [3,14,15], and direct characterization of quantum dynamics (DCQD)

could reduce the total number of configurations from 4^{2n} to 4^n [16,17]. In general, the exponential scaling is unavoidable to reconstruct a generic quantum process. However, by assuming certain structures of the unknown quantum process, the number of configurations can be significantly reduced. Such examples include compressed sensing quantum process tomography that assumes the measurement outcomes are sparse [18], and tensor-network-states-based quantum process tomography, which assumes a low entanglement structure of the underlying quantum process [19,20]. Recently, quantum machine learning combined with parametric quantum circuit (PQC) has received extensive focus [21–26]. Specifically, the authors of Ref. [26] showed that it was reasonable to compile an unknown unitary process into a parametric quantum circuit, however, it suffered from double amounts of qubits. The authors of Ref. [27] also demonstrated that one could efficiently encode the information of certain quantum states into a PQC using a gradient-based quantum machine learning algorithm, after which the unknown quantum state can be reconstructed classically with high fidelity using the optimal parameters of the PQC.

In this work, we propose a supervised quantum machine learning algorithm for quantum process tomography, which is also a continuation of Ref. [27]. As shown in Fig. 1, we use a PQC of certain depth d to approximate the unknown quantum process denoted by $\mathcal{U}$, where $\vec{\theta}$ is a list of parameters to be optimized in this PQC. To learn the information of $\mathcal{U}$, we randomly prepare a set of N random quantum states $|\psi_j\rangle$, each of which is separately fed into the unknown quantum process and the PQC. Then, as long as each pair of output quantum states $\mathcal{U}|\psi_j\rangle$ and $\mathcal{C}(\vec{\theta})|\psi_j\rangle$ are equal to each other and N is large enough, the unitary operation represented by $\mathcal{C}(\vec{\theta})$

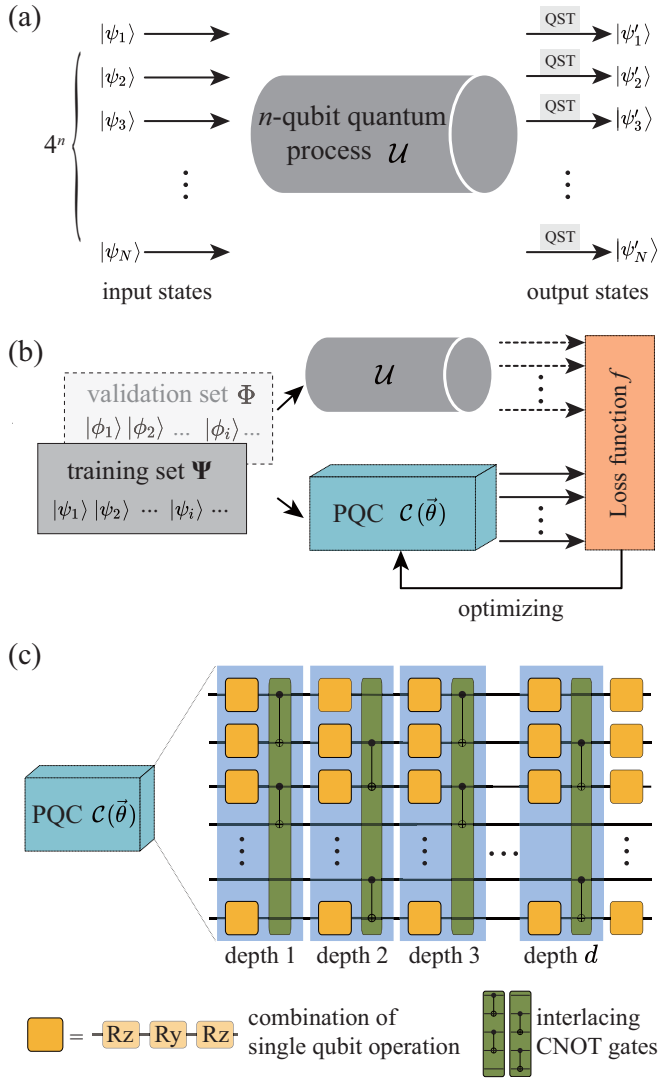*guochu604b@gmail.com

†junjiewu@nudt.edu.cn

FIG. 1. Scheme of variational quantum process tomography. Panel (a) illustrates the procedures of standard quantum process tomography. The information of the unknown quantum process is learned by 4^n pairs of inputs and outputs for an n-qubit quantum process $\mathcal{U}$. Panel (b) gives a general framework of our method. We build a loss function f to evaluate the distance between $\mathcal{U}$ and $\mathcal{C}(\vec{\theta})$. By training the PQC with the quantum states in the training set and validation set, and optimizing the parameters in $\mathcal{C}(\vec{\theta})$ based on a gradient-descending approach, the PQC gradually approximates the physical quantum process $\mathcal{U}$. Panel (c) shows the structure of the parametric quantum circuit which begins and ends with a single-qubit layer. Each two-qubit layer is counted as a depth and the circuit contains d depths of operations and ends with a single-qubit layer.

should be approximate to $\mathcal{U}$. As a result, all the information of $\mathcal{U}$ are stored in the parameters $\vec{\theta}$ and we can systematically reconstruct $\mathcal{U}$ from those parameters using a classical computer. Interestingly, during the training process, we use an additional set of random input states as the validation set, similar to that used in classical machine learning algorithms, which tests the generalization ability of the training outcomes for a particular PQC and set of input states.

Compared to its alternatives, our approach has several advantages. First, near-term quantum computers or simulators may only be able to faithfully run a quantum process with limited depths, therefore, it is reasonable that the same process could be reproduced by a parametric quantum circuit with a relatively low depth. However, the amount of entanglement produced by such quantum computers could be huge [12,13], in which case the tensor-network-states-based methods would be invalid. Second, our approach only requires to measure a single qubit for each configuration, hence it is less prone to errors with relatively shallow circuit and simple measurement [12,13]. Lastly, with a small number of input states, our approach may already be able to reconstruct $\mathcal{U}$ with very high fidelity. We demonstrate our approach on the reconstruction of two unitary processes produced by quantum Hamiltonian evolution and random quantum circuits, respectively. In both examples our numerical results show that we could reconstruct a quantum process up to eight qubits with a similarity value [defined in Eq. (6)] higher than 99%, and the number of required input quantum states is smaller than that required by SQPT by at least two orders of magnitude.

This paper is organized as follows. In Sec. II, we introduce the scheme of our quantum machine learning algorithm for quantum process tomography. In Sec. III, we demonstrate our method with numerical simulations of quantum process tomography for the time evolution of a quantum XXZ spin chain and the randomly generated quantum circuit. We conclude in Sec. IV.

II. APPROXIMATING UNITARY QUANTUM PROCESSES WITH PARAMETRIC QUANTUM CIRCUIT

Our quantum machine algorithm is composed of three parts: design of the PQC, training, and validation, which are shown in Figs. 1(b) and 1(c). Figure 1(a) shows the procedures of standard quantum process tomography as a comparison. In the following, we present the details of each component.

A. Parametric quantum circuit

The design of our PQC is shown in Fig. 1(c), where interlaced layers of single-qubit gates and two-qubits controlled-NOT (CNOT) gates are used. It is designed to quickly generate entanglement between qubits, thus make it possible to approximate complicated quantum processes. In practice, the design of the PQC should also take the underlying quantum hardware into consideration, especially the choice as well as the pattern of the two-qubit gates. Each two-qubit layer is counted as a depth and is varied between odd and even depth. Each single-qubit layer contains three rotational gates (R_z, R_y, and R_z) on each qubit, where R_y and R_z are defined as

$$R_y(\theta) = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}, \tag{1}$$

$$R_z(\theta) = \begin{bmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{bmatrix}. \tag{2}$$

The sequence R_z, R_y, and R_z makes sure that arbitrary single-qubit rotations can be produced with appropriate parameters. Our PQC ends with a single-qubit layer. As a result, for such a circuit with n qubits and d depths, the total number

of parameters is $3n(d+1)$. Actually, our method is a generalized framework with various circuit ansatzes. Based on the numerical results, we utilize such interlacing single-layer parallel CNOT gates in the following simulations. More details are shown in Appendix. B.

B. Training process

We build a loss function which reflects the distance between the unitary represented by $\mathcal{C}(\vec{\theta})$ and the target unitary $\mathcal{U}$. Concretely, we first randomly generate a set of N quantum states, denoted as $\Psi = \{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_N\rangle\}$. Here, a random quantum state is generated by applying an R_y gate with random parameters onto each qubit and controlled-Z (CZ) gates with random control and target qubits. Each state $|\psi_j\rangle$ is fed into the unknown quantum process $\mathcal{U}$ and the PQC $\mathcal{C}(\vec{\theta})$, with the output quantum states denoted as $|\psi_j^{\text{ideal}}\rangle = \mathcal{U}|\psi_j\rangle$ and $|\psi_j'\rangle = \mathcal{C}(\vec{\theta})|\psi_j\rangle$. Then we compute the Euclidean distance between $|\psi_j^{\text{ideal}}\rangle$ and $|\psi_j'\rangle$, which is

$$\big\| |\psi_j'\rangle - |\psi_j^{\text{ideal}}\rangle \big\|^2 = 2 - 2\text{Re}\big(\langle\psi_j^{\text{ideal}}|\psi_j'\rangle\big). \tag{3}$$

The inner product on the right-hand side of Eq. (3) can be efficiently computed with a quantum computer using a generalized SWAP-test algorithm, which is detailed in Appendix. A. The loss function f is defined as the summation of the distance obtained over all input states, which is

$$
\begin{aligned}
f(\vec{\theta}) &= \frac{1}{N}\sum_{j=1}^{N} \big\| |\psi_j'\rangle - |\psi_j^{\text{ideal}}\rangle \big\|^2 \\
&= \frac{2}{N}\sum_{j=1}^{N}\big[1 - \text{Re}\big(\langle\psi_j^{\text{ideal}}|\psi_j'\rangle\big)\big],
\end{aligned} \tag{4}
$$

namely, f is the mean square error between the two set of output quantum states.

The loss function f is a hybrid quantum-classical function and its gradient can be estimated based on the chain rule, where it contains functions to be evaluated with a quantum computer and functions to be evaluated on a classical computer

$$
\begin{aligned}
\frac{\partial f(\vec{\theta})}{\partial \vec{\theta}_j} &= \frac{\partial f(\vec{\theta})}{\partial \mathcal{O}(\vec{\theta})}\frac{\partial \mathcal{O}(\vec{\theta})}{\partial \vec{\theta}_j} \\
&= \frac{\partial f(\vec{\theta})}{\partial \mathcal{O}(\vec{\theta})}\left(\frac{1}{2}\mathcal{O}(\vec{\theta}_j^+) - \frac{1}{2}\mathcal{O}(\vec{\theta}_j^-)\right).
\end{aligned} \tag{5}
$$

$\mathcal{O}$ in our case means the generalized SWAP-test with parametric quantum gates and $\partial\mathcal{O}(\vec{\theta})/\partial\theta_j$ can be computed using parameter-shifting rule [28], where $\vec{\theta}_j$ denotes the jth parameter in the parameter list $\vec{\theta}$ and $\vec{\theta}_j^{\pm} = \vec{\theta}_j \pm \frac{\pi}{2}$. Hence the gradient of the loss function f can be computed following the authors of Ref. [29], who proposed a method to embed Eq. (5) into the classical automatic differentiation framework, such that the gradient of a hybrid quantum-classical loss function can be automatically computed using a hybrid quantum-classical computer. The gradient can then be fed into a gradient-based optimizer to minimize the loss function f.

After the training, we evaluate the distance between $\mathcal{U}$ and $\mathcal{C}(\vec{\theta})$ using the similarity defined as

$$\text{similarity}\big(\mathcal{U}, \mathcal{C}(\vec{\theta})\big) = 1 - \frac{\|\mathcal{C}(\vec{\theta}) - \mathcal{U}\|_F}{2\|\mathcal{U}\|_F}. \tag{6}$$

Here $\|X\|_F$ denotes the Frobenius norm of the matrix X. The similarity$(\mathcal{U}, \mathcal{C}(\vec{\theta})) = 1$ means $\|\mathcal{C}(\vec{\theta}) - \mathcal{U}\|_F = 0$, in which case $\mathcal{U}$ can be perfectly reconstructed from $\mathcal{C}(\vec{\theta})$.

C. Usage of a validation set

As a prior, we do not know whether our PQC is expressive enough or not to represent $\mathcal{U}$, and whether the number of input states is enough or not to ensure convergence to $\mathcal{U}$. Moreover, in practice we may also have the problem of overfitting such that the optimal $\mathcal{C}(\vec{\theta})$ is very distinct from $\mathcal{U}$ but the loss function f has already converged to 0. To overcome these problems, we borrow the idea of the validation set from classical machine learning, which is part of the training data and primarily used to test the generalization ability of the training outcomes without resorting to the testing data.

Concretely, we generate another set of input states denoted as $\Phi = \{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_N\rangle\}$, i.e., the validation set, which is independent of the training set. After the training process, we feed each $|\phi_j\rangle$ into the unknown quantum process and the resulting optimal PQC, obtaining two outputs $|\phi_j^{\text{ideal}}\rangle = \mathcal{U}|\phi_j\rangle$ and $|\phi_j'\rangle = \mathcal{C}(\vec{\theta})|\phi_j\rangle$. Then we compute the quantum fidelity between $|\phi_j^{\text{ideal}}\rangle$ and $|\phi_j'\rangle$ efficiently through the SWAP-test [30] on a quantum computer, and summarize over all the instances of the validation set, which is defined as the accuracy

$$\text{accuracy}\big(\mathcal{U}, \mathcal{C}(\vec{\theta})\big) = \frac{1}{N}\sum_{j\in\Phi}^{N}\text{Re}\big(\langle\phi_j^{\text{ideal}}|\phi_j'\rangle\big). \tag{7}$$

It is noted that accuracy is a faithful tool. If accuracy is close to 1, it means the PQC we obtained can be well generalized to the new input validation states. Actually we show that accuracy and similarity are indeed strongly correlated in our numerical simulation (using the Pearson correlation coefficient r), therefore we can pick out the simulation with larger accuracy value as a more faithful reconstruction of $\mathcal{U}$. Moreover, the accuracy is an efficient evaluation criterion. Since direct characterization of distance between $\mathcal{U}$ and $\mathcal{C}(\vec{\theta})$ as required in Eq. (6) scales exponentially with the number of qubits n, it is possible to determine whether the training is successful or not based on the accuracy, without resorting to the complete characterization of similarity.

III. NUMERICAL RESULTS AND DISCUSSIONS

We demonstrate our quantum machine learning algorithm using numerical simulations based on a classical PQC simulator. Specifically, we concentrate on two cases: (1) the unknown unitary process is produced by a quantum Hamiltonian evolution and (2) by random quantum circuits, respectively.
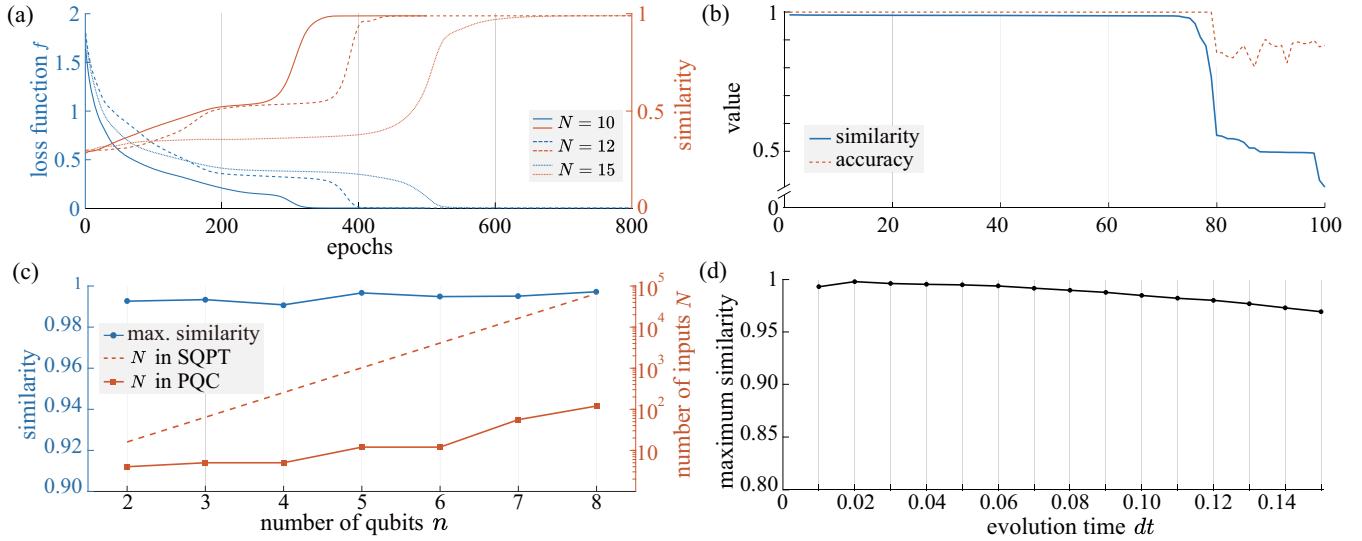
FIG. 2. Numerical results on the Heisenberg XXZ spin-chain time evolution. Panel (a) shows a typical training process on (5,7,10), (5,7,12), and (5,7,15) configurations. The loss function f (blue line, corresponds to the left axis) decreases with the training epochs till converging to a threshold value. Meanwhile, the similarity value (orange line, corresponds to the right axis) approaching 1. Panel (b) shows the correspondence between the accuracy on validation set and the similarity of the final results among 100 trials of (5,6,10) simulations. The data are sorted in a descending similarity order. The blue solid line denotes the similarity between $\mathcal{C}(\vec{\theta})$ and $\mathcal{U}$. The red dotted line shows the accuracy value on the validation set. It is noted that the accuracy and similarity values are indeed strongly correlated with a Pearson correlation coefficient $r = 0.9555$. (c) Scalability tests from 2-qubit to 8-qubit quantum processes. The left-axis (blue circle) shows the maximum similarity reached during repetitive trials and the logarithmic right-axis represents the corresponding number of input states N needed. The orange dotted line denotes the number of inputs in SQPT, and the orange rectangular represents that in PQC. In (a), (b), and (c), the evolution time $dt = 0.01$. (d) Extension tests on quantum processes with longer evolution time dt, ranging from 0.01 to 0.15. The line graph shows the maximum similarity value reached among repetitive trials on a (6,7,18) circuit.

A. *XXZ* spin-chain time evolution

We take the Hamiltonian of the Heisenberg XXZ spin chain [31] in a magnetic field as our example, which is written as

$$\hat{H}_{XXZ} = \sum_{l=1}^{n-1} \left[J\left(\hat{\sigma}_l^x \hat{\sigma}_{l+1}^x + \hat{\sigma}_l^y \hat{\sigma}_{l+1}^y\right) + \Delta \hat{\sigma}_l^z \hat{\sigma}_{l+1}^z \right] + h \sum_{l=1}^{n} \hat{\sigma}_l^z. \tag{8}$$

Here n is the number of spins (qubits), J is the tunneling strength, Δ is the interaction strength, and h is the magnetization strength. The evolutionary operator with time dt is denoted as

$$\mathcal{U}_{XXZ} = e^{-i\hat{H}_{XXZ}dt}. \tag{9}$$

In the simulations, we fix $h = 0.1$, $J = 1$, and set $dt = 0.01$.

As shown in Sec. II, we will encode $\mathcal{U}_{XXZ}$ into the circuit parameters $\vec{\theta}$. We prepare a set of randomly generated states $|\psi_i\rangle$ as inputs, i.e., the training set Ψ, and feed each $|\psi_i\rangle$ into both the quantum process and the PQC. The same training procedure is repeated for 100 times, with each of them initialized independently. After training, another randomly generated validation set is utilized to test the generalizability and pick out the most faithful instance of parametric circuit. Based on Eqs. (6) and (7), we can evaluate the performance of the parametric quantum circuits on the training set and validation set.

Here, we denote (n, d, N) as an n-qubit, d-depth, and N-input PQC configuration in the numerical simulation. Fig-

ure 2(a) illustrates a typical five-qubit training process on different sizes of input states: (5,7,10), (5,7,12), and (5,7,15) cases. The parametric quantum circuit is initialized by randomly generated parameters and evaluated by loss function in Eq. (4) at each epoch. Figure 2(a) shows that the loss function goes down with the training process while the similarity gradually reaches 1 under different N. Moreover, among 100 independent trials with random initialization on the (5,6,10) PQC configuration, we utilize the validation set to distinguish the more faithful PQC. It is plotted in Fig. 2(b) that the accuracy on the validation set shows a strong correspondence with the final similarity value with the Pearson correlation coefficient $r = 0.9555$. Therefore, it is feasible to utilize the accuracy value as a criterion to determine the optimal circuit parameters to reconstruct $\mathcal{U}$ with higher similarity.

In Fig. 2(c), we conduct simulations on Heisenberg XXZ spin chains of different lengths to evaluate the scalability of our method. For an n-qubit Heisenberg XXZ spin chain time evolution process (n ranges from 2 to 8), we utilize different depths of PQC d, choose different sizes of training sets N, repeat such independently identical simulations for 100 times, and calculate the maximum similarity achieved. Compared with standard quantum process tomography, which needs 4^n pairs of input states and output states, the number of required input quantum states in our method is at least two orders of magnitude less (concretely $N = 56$ and $N = 120$ for the seven-qubit and eight-qubit quantum processes, while the corresponding numbers in SQPT are 16 384 and 65 536, respectively). Meanwhile we utilize a relatively shallow depth
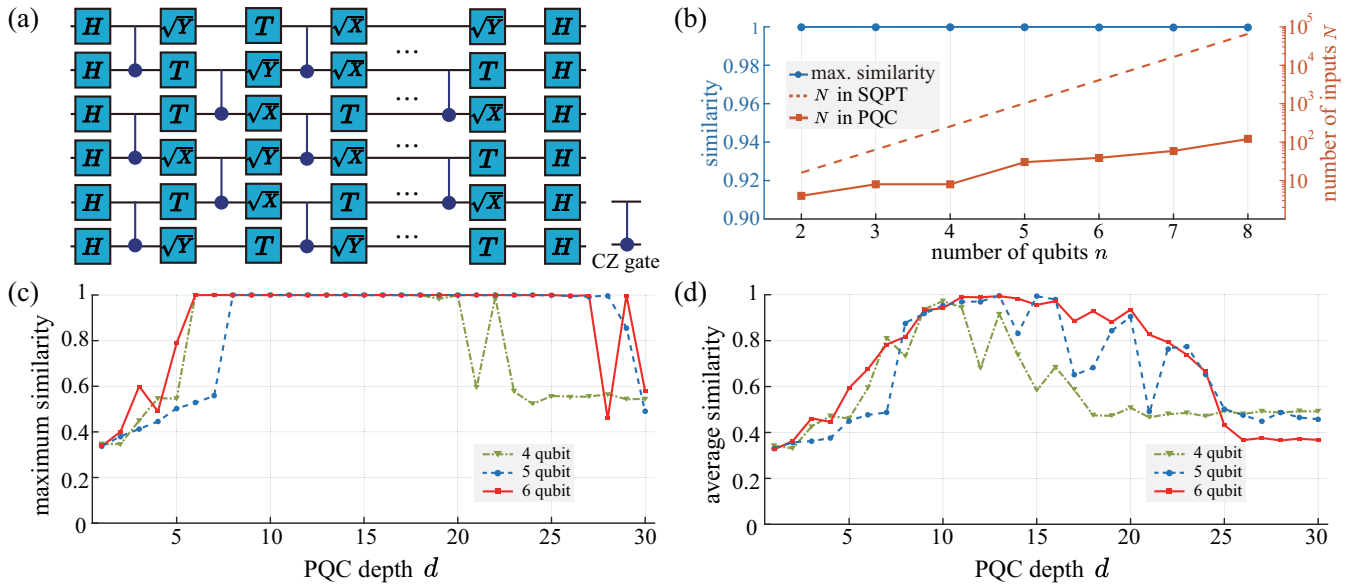
FIG. 3. Numerical results on the randomly generated quantum circuits. (a) A typical six-qubit randomly generated quantum circuit organization. (b) Scalability tests from two-qubit to eight-qubit quantum processes. The left-axis (blue circle) shows the maximum similarity reached during repetitive trials and the logarithmic right-axis represents the corresponding number of input states N needed. The randomly generated circuit are of depth 6, and the evolution time $dt = 0.01$. (c,d) Variational quantum process tomography on randomly generated quantum circuit on four-qubit (green triangle), five-qubit (blue circle), and six-qubit (red rectangle) cases, respectively. The line graph shows the relationship between the similarity [maximum value in (c) and average value in (d)] and the PQC depth d.

circuit ($d \leqslant 8$) and achieve a faithful similarity (higher than 99%) over all cases. More data and configuration details can be seen in Appendix B. In addition, our method only involves single-qubit measurements for each configuration, instead of measuring in the complete set of computational basis.

We further study the influence of the evolution time dt of the Hamiltonian on the results of our method in Fig. 2(d). We take a (6,7,18) PQC configuration as an example, and set various evolution times dt, ranging from 0.01 to 0.15. Under each case, we repeat the numerical simulations and get the maximum similarity value. In Fig. 2(d), we plot the maximum similarity against the evolution times dt. It can be seen that our method still achieves an acceptable similarity when the evolution time goes longer.

B. Randomly generated quantum circuit

In this section, we consider the quantum process of randomly generated circuits. For an n-qubit D-depth randomly generated quantum circuit, we first apply Hadamard gates to initialize the state to a symmetric superposition. Then, the circuit is organized by depth, including controlled-phase (CZ) gates alternating between odd and even configurations to entangle neighboring qubits and the randomly chosen single-qubit gate (T, $\sqrt{X}$, or $\sqrt{Y}$). Finally, Hadamard gates are applied to each qubit. A specific six-qubit randomly generated circuit is organized as shown in Fig. 3(a). It is noted that such randomly generated quantum circuits are hard for efficient simulation on a classical computer [32,33].

Here, we take the six-depth randomly generated quantum circuits as examples. We utilize $(4, d, 8)$, $(5, d, 30)$, and $(6, d, 39)$ PQC configurations to learn the corresponding quantum process, respectively, where PQC depth d ranges

from 1 to 30. Data details can be seen in Appendix B. Similarly, we conduct the scalability tests from two-qubit to eight-qubit six-depth randomly generated quantum circuits in Fig. 3(b). In Figs. 3(c) and 3(d), we plot the maximum similarity and average similarity value against the PQC depth d. It is seen that for such a specific random quantum circuit, our method can give an approximate circuit $\mathcal{C}(\vec{\theta})$ with high similarity among repetitive independent trails, and the results depend on the depth of the parametric circuits d. Taking the four-qubit (green line) randomly generated quantum circuit as an example, the most appropriate PQC depth d is among 6 to 15 (the maximum similarity can reach to 1 and meanwhile the average similarity keeps at a relatively high value). Shallow circuits may not be expressive enough to rebuild $\mathcal{U}$, while deeper depth with more parameters may result in overfitting. Hence, although the maximum similarity can reach to 1 toughly, but the average similarity remains low when the depth $d > 25$.

In addition, there are two points to be noted. First, the loss function in Eq. (4) is a global operation since we need to ensure the absolute phase between all output states. Hence it may result in the barren plateau when the system size grows [34,35]. Based on our simulation results, however, our method is feasible with the qubits' number going to 8 since we tried other approaches including changing circuit ansatzes (sequential or parallel structure), increasing training set states (various N), adding circuit depth (traversing d), and so on to avoid vanishing gratitude during optimization iterations. Moreover, we evaluated the loss value and added disturbance when the loss tends to be flat. Numerical results show that our method is available under the current two models and there indeed exists more possible attempts to solve the barren plateau problem. Second, our approach is initially designed

TABLE I. Numerical simulation details of scalability test on the Heisenberg XXZ spin-chain time evolution.

n	Maximum similarity			d	N_{PQC}	N_{SQPT}
	Seq. setting	2-para setting	1-para setting			
2	99.46%	99.37%	99.26%	2	4	16
3	99.42%	99.14%	99.34%	4	5	64
4	98.82%	98.70%	99.08%	4	5	256
5	99.13%	99.32%	99.66%	7	12	1024
6	98.88%	98.59%	99.48%	7	12	4096
7	98.73%	98.67%	99.51%	8	56	16384
8	97.97%	98.57%	99.72%	8	120	65536

TABLE III. Numerical simulation details of scalability test on the 6-depth randomly generated quantum circuits.

n	Maximum similarity			d	N_{PQC}	N_{SQPT}
	Seq. setting	2-para setting	1-para setting			
2	99.99%	99.99%	99.99%	3	4	16
3	99.99%	99.99%	99.99%	5	8	64
4	99.99%	99.99%	99.99%	6	8	256
5	99.99%	99.99%	99.99%	7	30	1024
6	99.99%	99.99%	99.99%	7	39	4096
7	99.99%	99.99%	99.99%	8	59	16384
8	-	-	99.99%	8	120	65536

only for the unitary quantum process, however, it has the potential to generalize it to a generic quantum process. Possible methods include utilizing the superposition of unitaries or extending the nonunitary to a larger unitary. Specifically, as proved in Ref. [37], we could utilize linear combinations of unitary operators to construct the equivalent nonunitary. It is also worth trying by transforming the nonunitary process into a unitary in a larger Hilbert space by unitary dilation [38] and then learning the extended unitary.

IV. CONCLUSION

In this work, we propose a quantum machine learning algorithm for quantum process tomography, which encodes the unknown unitary quantum process into a parametric quantum circuit of certain depth d. A set of randomly generated product quantum states are used as the training data to minimize the loss function. The training process is repeated with a validation set in the end to filter out the instance with highest similarity, namely the closest to the unknown quantum process.

We demonstrate our method by two numerical examples, including the Hamiltonian evolution with the Heisenberg XXZ spin chain from two-qubit to eight-qubit and random quantum circuits. The results indicate that a faithful reconstruction of $\mathcal{U}$ (similarity higher than 99%) can be reached with a relatively low-depth PQC ($d \leqslant 8$), and a relatively small number of training states (at least two orders of magnitude compared to SQPT). Moreover, only the single-qubit measurement is required in each configuration, instead of measuring in the complete set of the computational basis. The method shows good feasibility on both models under limited

system size and there are indeed more topics that can be considered in future work, including analyzing large-scale open quantum system dynamics and solving the barren plateau. Our work presents a promising application of using the quantum machine learning algorithm to accelerate quantum process tomography.

APPENDIX A: GENERALIZED SWAP-TEST CIRCUIT

The key to calculating the loss function and accuracy is the inner product between the target state $|\psi_j^{\mathrm{ideal}}\rangle = \mathcal{U}|\psi_j\rangle$ and the circuit output state $|\psi_j'\rangle = \mathcal{C}(\vec{\theta})|\psi_j\rangle$. For real state vectors, the standard SWAP-test circuit (shown in Fig. 4 gray shaded box) is enough to evaluate the overlap since it is a real number. However, $|\psi_j^{\mathrm{ideal}}\rangle$ and $|\psi_j'\rangle$ are complex vectors, so we introduce a generalized SWAP-test to evaluate the overlap value.

TABLE II. Numerical simulation details of time extension test on (6,7,18) PQC configuration.

dt	Max. similarity	dt	Max. similarity
0.02	99.80%	0.09	98.78%
0.03	99.62%	0.10	98.48%
0.04	99.54%	0.11	98.22%
0.05	99.49%	0.12	98.02%
0.06	99.38%	0.13	97.69%
0.07	99.18%	0.14	97.30%
0.08	98.98%	0.15	96.94%

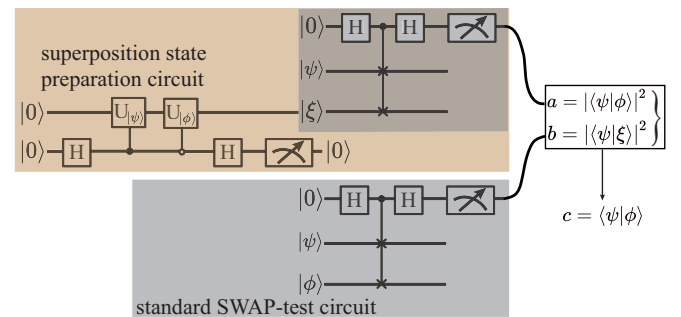FIG. 4. Generalized SWAP-test circuit. The bottom gray shaded circuit is the standard SWAP-test circuit. By measuring the ancillary qubit, we could get access to the fidelity between two input quantum states. The top orange shaded layer involves in a superposition state preparation circuit and a standard SWAP-test circuit. Solving the two equations on the right, we can get two complex state vector's overlap value c.

TABLE IV. Numerical simulation details of random quantum circuit tests.

	Max. similarity			Avg. similarity				Max. similarity			Avg. similarity		
d	4-qubit	5-qubit	6-qubit	4-qubit	5-qubit	6-qubit	d	4-qubit	5-qubit	6-qubit	4-qubit	5-qubit	6-qubit
1	34.73%	34.73%	33.69%	34.26%	32.97%	32.56%	16	99.99%	99.99%	99.99%	68.50%	97.95%	97.20%
2	34.59%	34.59%	40.09%	33.08%	35.74%	36.34%	17	99.98%	99.98%	99.99%	58.64%	64.97%	88.46%
3	44.95%	44.95%	60.00%	42.63%	36.24%	46.17%	18	99.92%	99.99%	99.99%	47.47%	68.22%	92.86%
4	54.79%	54.79%	48.97%	47.13%	37.55%	44.58%	19	98.34%	99.99%	99.99%	47.24%	84.41%	88.08%
5	54.52%	54.52%	78.94%	46.10%	45.11%	59.39%	20	99.85%	99.99%	99.99%	50.81%	90.25%	93.54%
6	99.99%	99.99%	99.99%	59.28%	47.66%	67.63%	21	59.22%	99.91%	99.98%	46.54%	48.39%	82.62%
7	99.99%	99.99%	99.99%	81.24%	48.72%	78.34%	22	99.09%	99.99%	99.97%	48.01%	76.70%	79.27%
8	99.99%	99.99%	99.99%	73.15%	87.78%	81.68%	23	57.71%	99.94%	99.97%	48.52%	77.47%	73.86%
9	99.99%	99.99%	99.99%	93.66%	91.77%	93.21%	24	52.24%	99.97%	99.98%	47.07%	65.41%	66.43%
10	99.99%	99.99%	99.99%	97.26%	95.31%	94.16%	25	55.73%	99.88%	99.99%	49.38%	50.24%	43.23%
11	99.99%	99.99%	99.99%	94.42%	96.93%	99.03%	26	55.48%	99.69%	99.61%	48.04%	47.60%	36.67%
12	99.99%	99.99%	99.99%	67.82%	97.03%	98.89%	27	55.44%	99.21%	99.90%	49.22%	44.83%	37.58%
13	99.99%	99.99%	99.99%	91.75%	99.62%	99.40%	28	56.54%	99.85%	45.77%	48.65%	48.84%	36.54%
14	99.99%	99.99%	99.99%	73.53%	83.19%	98.33%	29	54.25%	85.67%	99.94%	49.26%	46.50%	37.29%
15	99.99%	99.99%	99.99%	58.23%	99.34%	95.60%	30	54.51%	48.65%	57.68%	49.14%	45.78%	36.74%

Given two complex quantum states $|\psi\rangle$ and $|\phi\rangle$, it is already known that the fidelity between the two states can be evaluated using a SWAP-test circuit, namely $a = |\langle\psi|\phi\rangle|^2$ can be efficiently calculated on a quantum device. Thus, as long as we can prepare another superposition state $\xi = \frac{1}{\sqrt{2}}(|\psi\rangle + |\phi\rangle)$, we can also obtain the fidelity $b = |\langle\psi|\xi\rangle|^2$. Based on the two results above, we can arrive at the overlap between the two quantum state vectors

$$c = \langle\psi|\phi\rangle = b - \frac{a+1}{2} + i\sqrt{(a+1)b - b^2 - \frac{(a-1)^2}{4}}. \tag{A1}$$

The superposition state ξ can be easily obtained using a controlled operation with an auxiliary qubit, as $|0\rangle\langle0|\hat{\mathcal{U}}_{|\phi\rangle} + |1\rangle\langle1|\hat{\mathcal{U}}_{|\psi\rangle}$, where $\hat{\mathcal{U}}_{|\psi\rangle}$ denotes the unitary operation to produce the quantum state $|\psi\rangle$. As shown in Ref. [36], we can add control to the arbitrary unitary process. By postselecting the ancillary qubit on $|0\rangle$, we prepare the superposition state.

It is noted that our generalized approach only requires to measure a single qubit for each configuration (while the three-qubit TOFFOLI gate is used as required by the SWAP-test). Since the errors of quantum gate operations are almost one order of magnitude less than that of quantum measurements for current quantum computers [12,13], our method could be less prone to errors, compared to the massive and general measurements involved in SQPT.

APPENDIX B: DETAILS OF NUMERICAL SIMULATIONS

In this work, we put forward a generalized parametric quantum circuit framework to solve the quantum process tomography problem. Actually, the circuit ansatz is not fixed, but problem-specific. There are various circuit organizations and we take the following three kinds of circuit ansatzes as examples in Fig. 5 and conduct numerical simulations. The overall results on two models are showed in Tables I and III. Based on the numerical results, we utilize the single-layer parallel setting in the work.

In Fig. 2(c), we conduct the scalability tests on n-qubit Heisenberg XXZ spin-chain time evolution. Here, we give detailed numerical results listed in Table I where d, N_{PQC},

(a) sequential setting   (b) two-layer parallel setting   (c) single-layer parallel setting

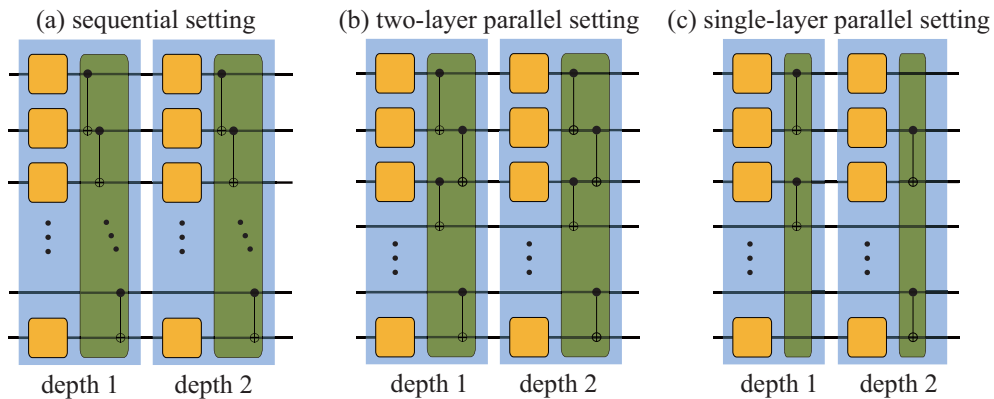depth 1   depth 2   depth 1   depth 2   depth 1   depth 2

FIG. 5. Different circuit ansatzes. (a) Sequential setting, the CNOT gates are one by one. (b) Two-layer parallel setting, the CNOT gates are organized as two staggering layers. (c) Single-layer parallel setting, the CNOT gates are interlacing organized (varied between odd and even depth).

and N_{SQPT} denote PQC depth, number of input quantum states in our method and number of input quantum states in SQPT, respectively, under single-layer parallel setting.

In Fig. 2(d), we conduct extension tests on longer time evolution process with (6,7,18) configuration. Here, we give detailed numerical results listed in Table II, where the maximum similarity is calculated under 100 trials.

In Fig. 3(b), we conduct the scalability tests on n-qubit six-depth randomly generated quantum circuits. Here, we give

detailed numerical results listed in Table III where d, N_{PQC}, and N_{SQPT} denote PQC depth, the number of input quantum states in our method, and the number of input quantum states in SQPT, respectively, under a single-layer parallel setting.

In Figs. 3(c) and 3(d), we utilize (4,6,8), (5,7,30), and (6,7,39) PQC configurations to learn the corresponding four-qubit, five-qubit, and six-qubit six-depth randomly generated quantum circuits, respectively. Here, we give detailed numerical results listed in Table IV.

[1] N. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[2] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).

[3] G. M. D'Ariano and P. LoPresti, Phys. Rev. Lett. **86**, 4195 (2001).

[4] S. Aaronson, Proc. R. Soc. A **463**, 3089 (2007).

[5] R. C. Bialczak, M. Ansmann, M. Hofheinz, E. Lucero, M. Neeley, A. D. O'Connell, D. Sank, H. Wang, J. Wenner, and M. Steffen, Nat. Phys. **6**, 409 (2009).

[6] J. L. O'Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph, and A. G. White, Phys. Rev. Lett. **93**, 080502 (2004).

[7] A. M. Childs, I. L. Chuang, and D. W. Leung, Phys. Rev. A **64**, 012314 (2001).

[8] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White, Phys. Rev. Lett. **106**, 100401 (2011).

[9] M. Riebe, K. Kim, P. Schindler, T. Monz, P. O. Schmidt, T. K. Körber, W. Hänsel, H. Häffner, C. F. Roos, and R. Blatt, Phys. Rev. Lett. **97**, 220407 (2006).

[10] L. Govia, G. Ribeill, D. Ristè, M. Ware, and H. Krovi, Nat. Commun. **11**, 1084 (2020).

[11] C. Guo, Y. Liu, M. Xiong, S. Xue, X. Fu, A. Huang, X. Qiang, P. Xu, J. Liu, S. Zheng, H. L. Huang, M. Deng, D. Poletti, W. S. Bao, and J. Wu, Phys. Rev. Lett. **123**, 190501 (2019).

[12] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell *et al.*, Nature (London) **574**, 505 (2019).

[13] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan *et al.*, Phys. Rev. Lett. **127**, 180501 (2021).

[14] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White, Phys. Rev. Lett. **90**, 193601 (2003).

[15] Z. Hou, J.-F. Tang, C. Ferrie, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Phys. Rev. A **101**, 022317 (2020).

[16] M. Mohseni and D. A. Lidar, Phys. Rev. Lett. **97**, 170501 (2006).

[17] M. Mohseni and D. A. Lidar, Phys. Rev. A **75**, 062331 (2007).

[18] A. V. Rodionov, A. Veitia, R. Barends, J. Kelly, D. Sank, J. Wenner, J. M. Martinis, R. L. Kosut, and A. N. Korotkov, Phys. Rev. B **90**, 144504 (2014).

[19] C. Guo, K. Modi, and D. Poletti, Phys. Rev. A **102**, 062414 (2020).

[20] G. Torlai, C. J. Wood, A. Acharya, G. Carleo, J. Carrasquilla, and L. Aolita, arXiv:2006.02424.

[21] M. C. Caro and I. Datta, Quantum Machine Intelligence **2**, 14 (2020).

[22] C. M. Popescu, Quant. Info. Proc. **20**, 286 (2021).

[23] C.-C. Chen, M. Watabe, K. Shiba, M. Sogabe, K. Sakamoto, and T. Sogabe, ACM Transactions on Quantum Computing **2**, 1 (2021).

[24] A. Abbas, D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner, Nature Computational Science **1**, 403 (2021).

[25] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Nat. Phys. **14**, 447 (2018).

[26] S. Khatri, R. LaRose, A. Poremba, L. Cincio, A. T. Sornborger, and P. J. Coles, Quantum **3**, 140 (2019).

[27] Y. Liu, D. Wang, S. Xue, A. Huang, X. Fu, X. Qiang, P. Xu, H.-L. Huang, M. Deng, C. Guo, X. Yang, and J. Wu, Phys. Rev. A **101**, 052316 (2020).

[28] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, Phys. Rev. A **99**, 032331 (2019).

[29] J. Liu, K. H. Lim, K. L. Wood, W. Huang, C. Guo, and H.-L. Huang, Sci. China Phys. Mech. Astron. **64**, 290311 (2021).

[30] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Phys. Rev. Lett. **87**, 167902 (2001).

[31] M. Žnidarič, T. Prosen, and P. Prelovšek, Phys. Rev. B **77**, 064426 (2008).

[32] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Nat. Phys. **14**, 595 (2018).

[33] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, Nat. Phys. **15**, 159 (2019).

[34] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Nat. Commun. **9**, 4812 (2018).

[35] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Nat. Commun. **12**, 1791 (2021).

[36] X.-Q. Zhou, T. C. Ralph, P. Kalasuwan, M. Zhang, A. Peruzzo, B. P. Lanyon, and J. L. O'Brien, Nat. Commun. **2**, 413 (2011).

[37] A. W. Schlimgen, K. Head-Marsden, LeeAnnM. Sager, P. Narang, and D. A. Mazziotti, Phys. Rev. Lett. **127**, 270503 (2021).

[38] C. Sparrow, E. Martín-López, N. Maraviglia, A. Neville, C. Harrold, J. Carolan, Y. N. Joglekar, T. Hashimoto, N. Matsuda, J. L. O'Brien *et al.*, Nature (London) **557**, 660 (2018).