

Optical neural network quantum state tomography

Ying Zuo,^{a,†} Chenfeng Cao^{ⓧ,a,†} Ningping Cao,^{b,c} Xuanying Lai^{ⓧ,d} Bei Zeng,^{a,*} and Shengwang Du^{d,*}

^aThe Hong Kong University of Science and Technology, Department of Physics, Hong Kong, China

^bUniversity of Guelph, Department of Mathematics and Statistics, Guelph, Ontario, Canada

^cUniversity of Waterloo, Institute for Quantum Computing, Waterloo, Ontario, Canada

^dThe University of Texas at Dallas, Department of Physics, Richardson, Texas, United States

Abstract. Quantum state tomography (QST) is a crucial ingredient for almost all aspects of experimental quantum information processing. As an analog of the “imaging” technique in quantum settings, QST is born to be a data science problem, where machine learning techniques, noticeably neural networks, have been applied extensively. We build and demonstrate an optical neural network (ONN) for photonic polarization qubit QST. The ONN is equipped with built-in optical nonlinear activation functions based on electromagnetically induced transparency. The experimental results show that our ONN can determine the phase parameter of the qubit state accurately. As optics are highly desired for quantum interconnections, our ONN-QST may contribute to the realization of optical quantum networks and inspire the ideas combining artificial optical intelligence with quantum information studies.

Keywords: optical neural network; quantum state tomography; quantum optics.

Received Nov. 23, 2021; revised manuscript received Jan. 30, 2022; accepted for publication Feb. 21, 2022; published online Mar. 24, 2022.

© The Authors. Published by SPIE and CLP under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.

[DOI: [10.1117/1.AP.4.2.026004](https://doi.org/10.1117/1.AP.4.2.026004)]

1 Introduction

Quantum state tomography (QST) is a standard process of reconstructing quantum information of an unknown quantum state through measurements of its copies. QST is used to verify state preparation, examine state properties such as correlations, and calibrate experimental systems. It is a crucial part of almost all aspects of experimental quantum information processing, including quantum computing, quantum metrology, and quantum communication.^{1–6}

As an analog of the “imaging” technique in quantum settings, QST is born to be a data science problem. Given limited copies of an unknown state ρ , we can extract its information via QST. QST is essentially an inverse problem, and such information recovering tasks are well suited to machine learning. Quantum learning theory indicates that $\Theta(2^{2n}/\epsilon^2)$ copies of ρ are necessary and sufficient to learn ρ up to trace distance ϵ .⁷ Although the tremendous resource requirement makes full-state QST

impractical for large-scale systems, several weaker quantum learning models (e.g., probably approximately correct learning,⁸ online learning,⁹ and shadow tomography^{10,11}) can exponentially reduce the computational resource for learning some 2-outcome measurement expectation values or “shadows.”

An artificial neural network (NN), a powerful algorithm in machine learning to fit a specific function, has been widely used for solving quantum information problems, such as quantum optimal control,^{12,13} quantum maximum entropy estimation,¹⁴ and Hamiltonian reconstruction.¹⁵ NNs have also been widely applied for QST applications, such as efficiently recovering the information of local-Hamiltonian ground states from local measurements,¹⁶ performing tomography on highly entangled state with large system size,¹⁷ mitigating the state preparation and measurement (SPAM) errors in experiments,¹⁸ and improving the state fidelity.^{19,20} Generative models with NNs can also perform QST with dramatically lower costs.^{21,22}

In this work, we demonstrate QST with an optical NN (ONN). Several optical implementations for realizing fully connected NN hardware have been proposed and demonstrated recently.^{23–28} Optical computing takes advantage of the bosonic wave nature of light: superposition and interference give rise to

*Address all correspondence to Bei Zeng, zengb@ust.hk; Shengwang Du, dusw@utdallas.edu

[†]These authors contributed equally to this work.

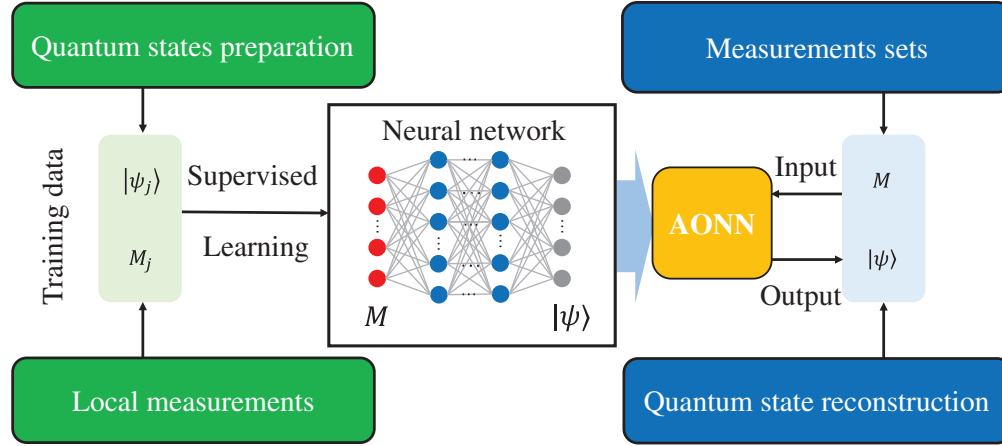


Fig. 1 Schematic of ONN-based QST.

its intrinsic parallel computing ability. Meanwhile, light is the fastest information carrier in nature. ONN is promising for next-generation artificial intelligence hardware, which provides high energy efficiency, low cross-talk, light-speed processing, and massive parallelism. As compared with the electronic version, ONNs are ideal for dealing with visual signals and information that are naturally generated and coded in light, such as image recognition and vehicular automation. However, most ONN demonstrations are still restricted to linear computation only due to the lack of suitable nonlinearity at a low light level for a large amount of optical neurons.^{25–27} Without nonlinear activation functions, ONN is always equivalent to a single-layer structure that cannot be applied for “real” deep machine learning. This problem had not been solved until most recently optical nonlinearity based on electromagnetically induced transparency (EIT),^{28,29} phase-change materials,³⁰ and saturated absorption^{31,32} was implemented to realize artificial optical neurons for ONNs.

Figure 1 illustrates a general scheme of ONN-QST. First, we collect the training data set from a known quantum state $\{|\psi_j\rangle\}$ and the corresponding local measurements $\{M_j\}$. Second, we train NNs under supervised learning with some nonlinear activation functions in their hidden neurons to obtain the optimal network parameters. Third, we take the trained network parameters to configure the ONN and perform some fine adjustments to optimize the hardware performance. Last, we feed measurement data sets to the trained ONN to reconstruct unknown quantum states. To validate this scheme, in the following sections, we start with a general discussion of QST with the computer-simulated NN and then describe our ONN experimental approach.

2 NN for QST

We consider a general n -qubit space with Pauli operators (removed the all identity terms) defined as

$$P = \{\sigma_{i_1}^{(1)} \otimes \cdots \otimes \sigma_{i_n}^{(n)} | \sigma_{i_k}^{(k)} \in \mathcal{P}, \sum_{k=1}^n i_k \neq 0\}, \quad (1)$$

where $\mathcal{P} = \{\sigma_0 = I, \sigma_1 = X, \sigma_2 = Y, \sigma_3 = Z\}$. Every term in P is specified by its index (i_1, i_2, \dots, i_n) . Measuring every element in P performs a QST for any n -qubit quantum state

ρ . For instance, when $n = 1$, we need to measure all three Paulis X, Y, Z for QST. Clearly, the cardinality of P grows exponentially with n . When ρ is a pure state, one may use techniques to reduce the number of measurements for $n > 1$. Compressed sensing is an efficient technique for recovering low-rank quantum states from randomly sampled Pauli operators.^{33,34}

When ρ is a pure state, it can be written as

$$|\psi\rangle = \sum_{k=1}^{2^n} a_k |\phi_k\rangle, \quad (2)$$

where $\{|\phi_k\rangle\}$ are the computational basis, and the amplitudes $a_k \in \mathbb{C}$ are normalized [i.e., $\sum_{k=1}^{2^n} (a_{k,r}^2 + a_{k,im}^2) = 1$, where $a_{k,r} \in \mathbb{R}$ and $a_{k,im} \in \mathbb{R}$ are the real and imaginary parts of a_k , respectively].

The measurement expectation values P are $\vec{c} = \text{tr}(\rho \cdot P) = (\text{tr}(\rho P_1), \text{tr}(\rho P_2), \dots, \text{tr}(\rho P_{4^n-1}))$. For a single-qubit pure state $\alpha|0\rangle + \beta|1\rangle$, its density matrix can be expressed as

$$\rho = \frac{1}{2} (1 + \vec{c} \cdot \vec{\sigma}), \quad (3)$$

where $\vec{\sigma} = \{X, Y, Z\}$ and $\vec{c} = (\text{tr}(\rho X), \text{tr}(\rho Y), \text{tr}(\rho Z)) \equiv (\langle X \rangle, \langle Y \rangle, \langle Z \rangle)$.

In compressed sensing, one needs to randomly sample a set $P^m = \{P_1, \dots, P_m\}$ of m Pauli operators from P , then use $\vec{c} = \text{tr}(\rho \cdot P^m) = (\text{tr}(\rho P_1), \text{tr}(\rho P_2), \dots, \text{tr}(\rho P_m))$ to recover the unknown state ρ , more precisely, the parameters of ρ . This can be regarded as a regression problem to estimate the function between \vec{c} and the parameters of ρ (e.g., $a_{k,r}$ and $a_{k,im}$).

NNs are excellent tools for solving regression problems. When using NNs for QST, the expectation values \vec{c} from random-sampled P^m are inputs to the network, and state parameters $(a_{k,r}, a_{k,im})$ are the outputs. Compared with compressed sensing, the NN for QST can be significantly faster when processing many data points. Once the NN is well-trained, it can produce reliable unseen results within an instance, while one needs to solve a convex optimization problem for each data point when applying compressed sensing. Note that both NN-QST and compressed sensing use much fewer measurement settings than the standard method. Without loss of generality, we use the simplest

type of NNs in this paper—fully connected feed-forward NNs. The neurons between the nearest layers are fully connected, and the information only passes forward while training. The supervised training process is to compare the ideal outputs $(a_{k,r}, a_{k,im})$ with current NN outputs and update parameters embedded in the NN to minimize their difference.

We numerically trained computer-based NNs nonlinear activation functions for 1-qubit, 2-qubit, and 3-qubit QST. For the 1-qubit system, the number of sampled operators $m \in [1, 2, 3]$; for the 2-qubit system, the number of sampled operators $m \in [6, 8, 10, 12]$; and for the 3-qubit system, $m \in [20, 25, 30, 35, 40]$. Plainly, m equals the number of input neurons, and n decides the number of output neurons. For each m , three sets of Pauli operators have been sampled and tested. Figure 2 plots the average fidelities (green bars) of both cases as functions of the number of randomly sampled Paulis. For the single-qubit system, the fidelity reaches 99.99% with three Paulis [Fig. 2(a)]. For the 2-qubit system, the fidelity reaches 99.9% with 10 randomly sampled Paulis [Fig. 2(b)]. For the 3-qubit system,

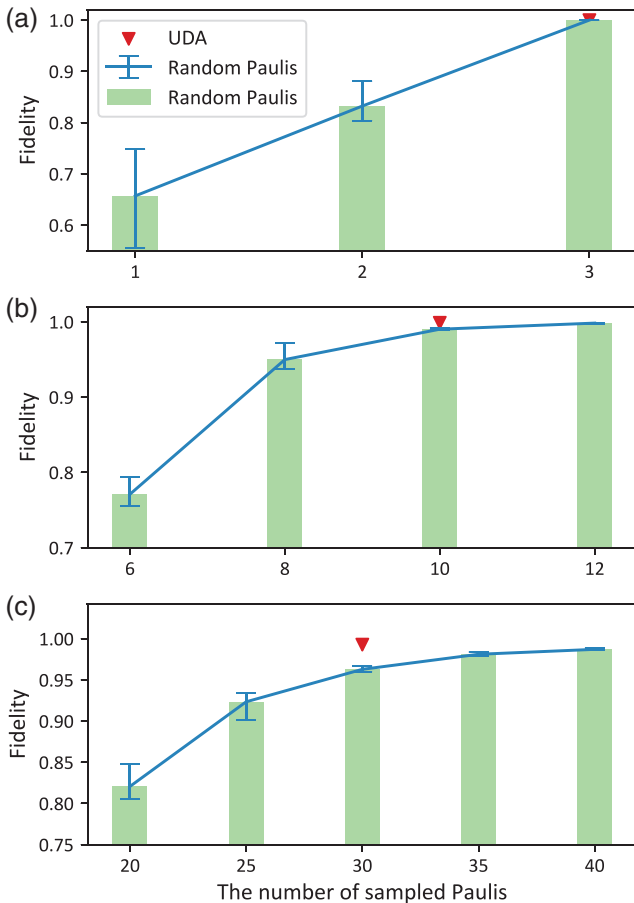


Fig. 2 The fidelities of NN predictions for different samples of Pauli operators: the red triangles are the average fidelities for UDA Pauli operator sets, which are very close to 1. A Pauli operator set is said to be “UDA” if measuring these operators can uniquely determine a pure state among all states. The green bars are the average fidelities for random sampled Pauli operator sets. The blue lines are the error bars for different samples. We train NN to predict state wavefunctions from measurements for (a) 1 qubit, (b) 2 qubits, and (c) 3 qubits.

a fidelity of higher than 99.9% requires more than 35 randomly sampled Paulis [Fig. 2(c)]. Details of training can be found in Sec. S1 in the [Supplemental Materials](#).

Theoretically, a pure state ρ uniquely determined among all states (UDA) for measuring a set of operators F means that there is no other state, pure or mixed, that has the same expectation values while measuring F .³⁵ In Ref. 36, the authors discovered two sets of Pauli operators, P_{2-UDA} and P_{3-UDA} , that are UDA for all 2-qubit and 3-qubit pure states, respectively (see Sec. S2 in the [Supplemental Materials](#) for the particular sets P_{2-UDA} and P_{3-UDA}). Namely, they are special cases of Pauli operator sets that the map between expectation values and the measured state ρ is bijective. Similarly, we apply NNs for these two sets of UDA operators and obtain the prediction fidelities of 99.9% for the 2-qubit case and 99.3% for the 3-qubit case (red triangles in Fig. 2).

We remark that our UDA scheme is not readily scalable for larger systems. However, there exist protocols with better scalability, e.g., compressed sensing,³³ shadow tomography,^{10,11} where NNs can also be naturally used. In addition, our NN-based scheme can be adapted to quantum tomography in the optical system by taking physical constraints into account, which we will discuss in the next section.

3 ONN-QST Experiment

In this first proof-of-principle experimental demonstration, we implement the single-qubit space with light polarizations, i.e., horizontal polarization $|H\rangle = |0\rangle$ and vertical polarization $|V\rangle = |1\rangle$. Instead of making a full QST, here we focus our task to determine the phase parameter of a pure state $|\psi\rangle = \frac{1}{\sqrt{2}}(|H\rangle + e^{i\theta}|V\rangle)$. The experimental ONN-QST setup is displayed in Fig. 3. In conventional QST, an arbitrary polarization state can be reconstructed by measuring the expectation values of the three Pauli operators. Figure 3(a) illustrates such an optical measurement setup. A laser beam passes through a polarization beam splitter (PBS_1) and becomes horizontally polarized ($|H\rangle$). The target state $|\psi\rangle = \frac{1}{\sqrt{2}}(|H\rangle + e^{i\theta}|V\rangle)$ is prepared by letting this horizontally polarized light pass through a half-wave plate (HWP_1) and a quarter-wave plate (QWP_1). The expectation values $\langle X \rangle$, $\langle Y \rangle$, and $\langle Z \rangle$ are obtained by sending the light polarization qubit state to the measurement units II, III, and IV shown in Fig. 3(a). To determine $\langle Z \rangle$, we send the polarization qubit directly to PBS_2 which projects $|H\rangle$ and $|V\rangle$ into two photodetectors in the measurement unit III. The normalized differential output from these two photodetectors gives the value $\langle Z \rangle$. The same setup can also be used to determine $\langle X \rangle$ or $\langle Y \rangle$ by placing HWP_2 or QWP_2 before PBS_2 as shown in II or IV, respectively (see Sec. S3 in the [Supplemental Materials](#) for details).

We obtain a data set $\{M_i\} = \{|\phi_i\rangle : 1 - \langle X \rangle_i, 1 - \langle Y \rangle_i, 1 - \langle Z \rangle_i\}$ by varying the phase $\theta \in [0, \pi/2]$ in the qubit state $|\psi\rangle = \frac{1}{\sqrt{2}}(|H\rangle + e^{i\theta}|V\rangle)$ and use them to train our ONN in Fig. 3(b). The ONN comprises an input layer of three neurons, a hidden layer of 20 neurons, and a single-neuron output layer.^{28,29} Figure 3(b) shows the optical layout of the ONN, and its network structure diagram is displayed in Fig. 3(c). The three coupling laser beams in the optical input layer are generated by a spatial light modulator (SLM_1) in Fig. 3(b), lenses L_2 and L_3 , and an aperture, as shown in unit I of Fig. 3(b). The SLM_1 is divided into three parts and each part is encoded with the sine phase pattern $m\pi \sin[(2\pi/T_{mj})j + (2\pi/T_{mk})k]$,

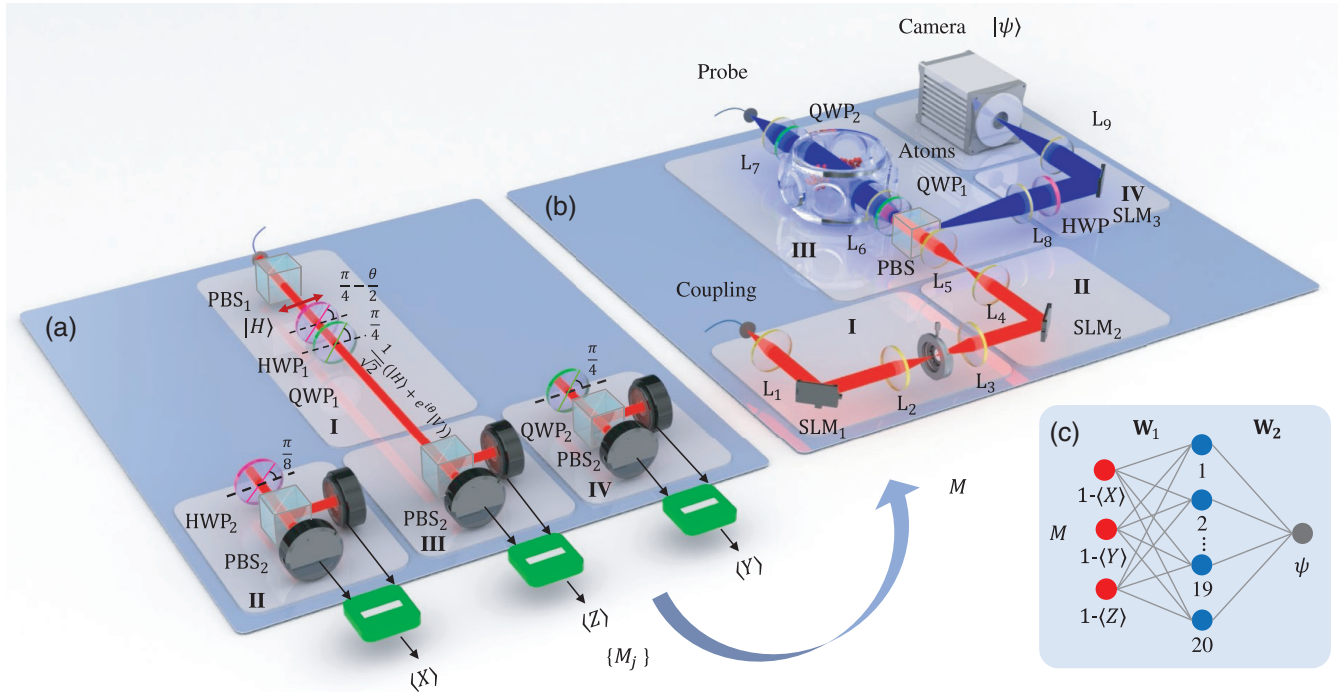


Fig. 3 Schematics of optical implementation of QST. (a) Optical layout of qubit QST, including generation of polarization state (top panel), measurement of $\langle Z \rangle$, $\langle X \rangle$, and $\langle Y \rangle$ (bottom panel). The fast axis of the HWP₁ is aligned with an angle $\pi/4 - \theta/2$ to the horizontal direction. The fast axis of the QWP₁ is aligned with an angle $\pi/4$ to the horizontal direction. (b) Schematic of ONN. (I) Input generation, (II) linear operation of the first layer, (III) nonlinear operation, and (IV) linear operation of the second layer. Spatial light modulators: SLM₁ (HOLOEYE LETO), SLM₂ (HOLOEYE PLUTO-2), and SLM₃ (HOLOEYE GEAE-2). Camera: Hamamatsu C11440-22CU. Lenses: L₁–L₉. Atoms are trapped in a MOT. (c) The NN structure employed.

where m is the modulation depth, T_{mj} and T_{mk} are the period of modulation along x and y directions, and j and k are the pixel number along the x and y directions. The sine phase encoded on SLM₁ modulates the beams into separated beams at the focal plane of lens L₂. The aperture behaves as a filter to keep the zero-order beam, whose intensity is determined by the modulation depth m . Thus, the intensity of the three beams is changed according to the input. The focal beams pass through lens L₃ and are collimated and incident to the SLM₂. These weighted beams, as the input vector, are incident on SLM₂, which diffracts each beam into 20 directions with designed weights (see Sec. S4 in the [Supplemental Materials](#) for the algorithm to calculate the pattern encoded on SLM₂). A Fourier lens L₄ performs linear summation for the beams diffracted into the same direction and forms 20 spots on its front focal plane. Thus, the combination of SLM₂ and L₄ completes the first linear operation W_1 and generates the input to the hidden layer. We then image these 20 spots with lenses L₅ and L₆ to laser-cooled ⁸⁵Rb atoms in a two-dimensional magneto-optical trap (MOT),^{37,38} where these 20-spot coupling beam patterns spatially modulate the transparency of the atomic medium through EIT.^{39,40} Another relatively weak collimated probe beam counterpropagates through the MOT, and its spatial transmission is nonlinearly controlled by the 20-spot coupling beam pattern. Here the nonlinear optical activation functions are realized with EIT in cold atoms. The equation of nonlinear activation functions is as follows:

$$I_p^{\text{out}} = f(I_c) = I_p^{\text{in}} e^{-\text{OD} \frac{4\gamma_{12}\gamma_{13}}{\Omega_c^2 + 4\gamma_{12}\gamma_{13}}}, \quad (4)$$

where I_p^{in} is the power of the input probe beam. Ω_c is the Rabi frequency of the coupling beam, and Ω_c^2 is proportional to coupling beam intensity I_c . Here, $\gamma_{13} = 2\pi \times 3$ MHz is fixed and determined by the spontaneous emission of the excited state $|3\rangle$. The ground-state dephasing rate γ_{12} can be engineered by applying an external magnetic field. OD is the atomic optical depth on the probe transition.

The image of the probe beam transmission pattern by lenses L₆ and L₈ becomes the output of the 20 hidden neurons. SLM₃ and Fourier lens L₉ perform the second linear matrix operation W_2 , and a camera records the output. The technical details of our ONN are described in Refs. 28 and 29.

In this work, because we encode trained NN model and input data into the power of beams, the ONN can only handle positive values: input, output, linear matrix elements, and input/output of nonlinear activation functions are all positive values.^{28,29} Meanwhile, the EIT optical nonlinear activation functions are increasing and convex. The lack of negative values in the NN limits its ability. Therefore the ONN is only able to perform regression tasks on increasing and convex functions. To match the ONN constraints, we perform a transform to the input variable, e.g., $\langle X \rangle$ to $1 - \langle X \rangle$, so that all input values to the ONN nodes are positive. We add these conditions to NN to simulate the ONN performance. The optimizer we use is Adam.⁴¹ We find

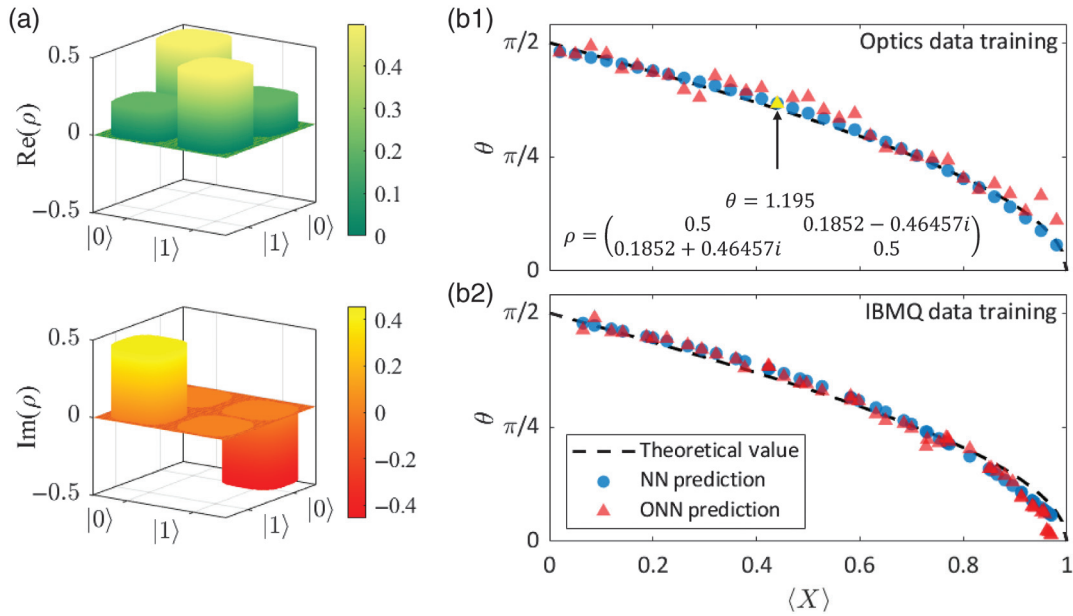


Fig. 4 (a) Optical tomography of the qubit and (b) experimental ONN tomography result. The ONN is training by (b1) optical tomography data and (b2) IBMQ tomography data. The black dashed line is the theoretical value of the phase θ according to the $\langle X \rangle$. The blue circles are the phase θ numerically predicted by the trained NN, and the red triangles are the experimentally measured predictions of θ according to $\langle X \rangle$. The yellow triangle is an example of the experimental ONN predicted state.

that this specific ONN fails to describe the whole range of non-monotonic functions. For the first proof-of-principle experimental demonstration, we will only apply the ONN for single-qubit QST with phase θ within $[0, \pi/2]$. It is surprising that such a positive-valued ONN is still able to perform some types of QST.

To train the ONN, we prepared the training data set $\{M_i\}$ from 23 phase values from a uniform distribution $\theta_j \sim U(0, \pi/2)$, corresponding to the optical polarization states $\{\rho_j = \mathcal{N}(|\phi_j\rangle\langle\phi_j|)\}$. Here, \mathcal{N} is the noise channel in experiments, and measures the Pauli expectation values $\langle X \rangle$, $\langle Y \rangle$, $\langle Z \rangle$. In a similar way, we prepare a test set with 32 independent data samples.

In addition to optical quantum states, we sample data from the IBM quantum (IBMQ) computer `ibmq_ourense`,⁴² and implement the same ONN training for comparison. The quantum circuit to prepare $|\psi\rangle = (|H\rangle + e^{i\theta}|V\rangle)/\sqrt{2}$ is the initial state $|H\rangle$ going through a Hadamard gate and then going through an RZ rotation gate. On `ibmq_ourense`, we uniformly sample 158 data points as the training set; 50 data points as the test set. Experimental optical quantum state and IBMQ tomography data are used to train two NNs. Details of training ONN can be found in Sec. S5 in the [Supplemental Materials](#).

Figure 4 shows the ONN state construction results using NN models trained by the ONN-QST training set and the IBMQ computer training set separately. The theoretical value is calculated from $\langle X \rangle$ directly. With the ONN system set up for the training results, we sent a set of the input vectors to the system. The examples of the real and imaginary parts of the density matrix are shown in Fig. 4(a). The experimentally measured state example is predicted by the ONN QST training model. The example input vector for the ONN model is $(\langle X \rangle, \langle Y \rangle, \langle Z \rangle) = (0.440, 0.898, 0)$ and the experimental ONN predicted state is

$$\theta = 1.195 \text{ and } \rho = \begin{pmatrix} 0.5 & 0.1852 - 0.46457i \\ 0.1852 + 0.46457i & 0.5 \end{pmatrix}$$

which is close to the theoretical value $\theta = 1.1152$ and NN predicted value $\theta = 1.1532$. The state is also marked with a yellow triangle in Fig. 4(b1). The experimental results are shown in Fig. 4(b). The theoretical value, NN predicted value, and experimental ONN predicted value agree with both optics data training [Fig. 4(b1) shows] and IBMQ data training [Fig. 4(b2) shows]. The theoretical value, NN prediction value, and ONN predicted value are consistent in both cases. The results suggest that our positive-valued ONN with EIT nonlinear activation functions is capable of qubit QST.

4 Discussion and Conclusion

While most demonstrations of ONNs took classification tasks to verify their feasibility,^{26,27,30} we performed the first regression task, i.e., ONN-QST. To accomplish regression tasks, the nonlinear function is essential as long as the relation between the input vector and output vector cannot be expressed linearly. The tunable EIT nonlinear optical activation functions in our ONN offer opportunities for performing regression tasks with convex and increasing/decreasing functions. Although our ONN has some certain limitations that the linear operation matrix elements are all positive valued, it has the potential to do large-size QST with restrictions.

Further, ONN can play a positive role in the noisy intermediate-scale quantum (NISQ) era. In NISQ algorithms, one usually only needs to reconstruct some reduced density matrix and extract the required local information instead of characterizing the whole system through a full-state tomography. ONN-QST can serve as an efficient subroutine to speed up this process. For example, within each Trotter step of the quantum imaginary

time evolution,^{43,44} we can train an ONN to reconstruct the reduced density matrix of some neighboring qubits, then use this information to determine the direction of the next step.

To perform QST for a higher dimensional space requires more active neurons. Our theoretical simulation shows 10 and 30 inputs are needed for the 2-qubit and 3-qubit cases, respectively. However, while the number of optical neurons is not a limiting factor in our current experimental setup, the ONN input/output and matrix weights are all positive-valued. Meanwhile, the nonlinear activation functions we implemented are increasing and convex, and it is impossible to conduct the regression task of nonmonotonic functions experimentally. These physical limitations limit us to performing more complicated QST. We believe the next generation of complex-valued ONNs with data encoded in both light amplitude and phase will be more powerful. The future development of complex-valued ONNs may enable large-size QST and more applications.

Optical quantum networks⁴⁵ have been brought to the fore by the reduced decoherence and high speed of photons. Recently, apart from generating optical quantum states⁴⁶ and optical quantum communication over a long distance,⁴⁷ multiple state-of-the-art experiments on optical quantum interfaces to store⁴⁸ and distribute entanglements^{49,50} have been exhibited. Among all of these, QST is essential for characterizing the generation and preservation of quantum states and has the potential to verify the entanglement distributed across the whole network. We believe that our optical setup of integrated ONN-QST will shed light on replenishing the optical quantum network with one more brick.

Acknowledgments

Authors acknowledge the use of IBMQ services for this work. The authors declare no competing interests. C.C. and B.Z. are supported by General Research Fund (No. GRF/16300220).

Code, Data, and Materials Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. F. Bouchard et al., “Quantum process tomography of a high-dimensional quantum communication channel,” *Quantum* **3**, 138 (2019).
2. G. M. D’Ariano, M. G. Paris, and M. F. Sacchi, “Quantum tomography,” in *Adv. Imaging Electron Phys.*, P. W. Hawkes, Ed., Vol. **128**, 206–309, Elsevier (2003).
3. U. Leonhardt, “Quantum-state tomography and discrete wigner function,” *Phys. Rev. Lett.* **74**(21), 4101–4105 (1995).
4. R. Thew et al., “Qudit quantum-state tomography,” *Phys. Rev. A* **66**(1), 012303 (2002).
5. A. I. Lvovsky and M. G. Raymer, “Continuous-variable optical quantum-state tomography,” *Rev. Mod. Phys.* **81**, 299–332 (2009).
6. M. Rambach et al., “Robust and efficient high-dimensional quantum state tomography,” *Phys. Rev. Lett.* **126**(10), 100402 (2021).
7. R. O’Donnell and J. Wright, “Efficient quantum tomography,” in *Proc. Forty-Eighth Annu. ACM Symp. Theory of Comput.*, pp. 899–912 (2016).
8. S. Aaronson, “The learnability of quantum states,” *Proc. R. Soc. A* **463**(2088), 3089–3114 (2007).
9. S. Aaronson et al., “Online learning of quantum states,” *J. Stat. Mech.* **2019**(12), 124019 (2019).
10. S. Aaronson, “Shadow tomography of quantum states,” *SIAM J. Comput.* **49**(5), STOC18-368 (2019).
11. H.-Y. Huang, R. Kueng, and J. Preskill, “Predicting many properties of a quantum system from very few measurements,” *Nat. Phys.* **16**(10), 1050–1057 (2020).
12. M. Y. Niu et al., “Universal quantum control through deep reinforcement learning,” *NPJ Quantum Inf.* **5**, 33 (2019).
13. Z. An et al., “Quantum optimal control of multilevel dissipative quantum systems with reinforcement learning,” *Phys. Rev. A* **103**(1), 012404 (2021).
14. N. Cao et al., “Neural networks for quantum inverse problems,” arXiv:2005.01540 (2021).
15. C. Cao et al., “Supervised learning in Hamiltonian reconstruction from local measurements on eigenstates,” *J. Phys.: Condens. Matter* **33**(6), 064002 (2020).
16. T. Xin et al., “Local-measurement-based quantum state tomography via neural networks,” *NPJ Quantum Inf.* **5**, 109 (2019).
17. G. Torlai et al., “Neural-network quantum state tomography,” *Nat. Phys.* **14**(5), 447–450 (2018).
18. A. M. Palmieri et al., “Experimental neural network enhanced quantum tomography,” *NPJ Quantum Inf.* **6**, 20 (2020).
19. Y. Quek, S. Fort, and H. K. Ng, “Adaptive quantum state tomography with neural networks,” *NPJ Quantum Inf.* **7**, 105 (2021).
20. S. Ahmed et al., “Classification and reconstruction of optical quantum states with deep neural networks,” *Phys. Rev. Res.* **3**(3), 033278 (2021).
21. J. Carrasquilla et al., “Reconstructing quantum states with generative models,” *Nat. Mach. Intell.* **1**(3), 155–161 (2019).
22. S. Ahmed et al., “Quantum state tomography with conditional generative adversarial networks,” *Phys. Rev. Lett.* **127**(14), 140502 (2021).
23. G. Wetzstein et al., “Inference in artificial intelligence with deep optics and photonics,” *Nature* **588**(7836), 39–47 (2020).
24. B. J. Shastri et al., “Photonics for artificial intelligence and neuro-morphic computing,” *Nat. Photonics* **15**(2), 102–114 (2021).
25. D. Woods and T. J. Naughton, “Photonic neural networks,” *Nat. Phys.* **8**(4), 257–259 (2012).
26. Y. Shen et al., “Deep learning with coherent nanophotonic circuits,” *Nat. Photonics* **11**(7), 441–446 (2017).
27. X. Lin et al., “All-optical machine learning using diffractive deep neural networks,” *Science* **361**(6406), 1004–1008 (2018).
28. Y. Zuo et al., “All-optical neural network with nonlinear activation functions,” *Optica* **6**(9), 1132–1137 (2019).
29. Y. Zuo et al., “Scalability of all-optical neural networks based on spatial light modulators,” *Phys. Rev. Appl.* **15**(5), 054034 (2021).
30. J. Feldmann et al., “All-optical spiking neurosynaptic networks with self-learning capabilities,” *Nature* **569**(7755), 208–214 (2019).
31. X. Guo et al., “Backpropagation through nonlinear units for the all-optical training of neural networks,” *Photonics Res.* **9**(3), B71–B80 (2021).
32. A. Ryou et al., “Free-space optical neural network based on thermal atomic nonlinearity,” *Photonics Res.* **9**(4), B128–B134 (2021).
33. D. Gross et al., “Quantum state tomography via compressed sensing,” *Phys. Rev. Lett.* **105**(15), 150401 (2010).
34. S. T. Flammia et al., “Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators,” *New J. Phys.* **14**(9), 095022 (2012).
35. J. Chen et al., “Uniqueness of quantum states compatible with given measurement results,” *Phys. Rev. A* **88**(1), 012109 (2013).
36. X. Ma et al., “Pure-state tomography with the expectation value of pauli operators,” *Phys. Rev. A* **93**(3), 032140 (2016).
37. H. J. Metcalf and P. van der Straten, “Laser cooling and trapping of atoms,” *J. Opt. Soc. Am. B* **20**(5), 887–908 (2003).
38. S. Zhang et al., “A dark-line two-dimensional magneto-optical trap of ⁸⁵Rb atoms with high optical depth,” *Rev. Sci. Instrum.* **83**(7), 073102 (2012).
39. S. E. Harris, “Electromagnetically induced transparency,” *Phys. Today* **50**(7), 36–42 (1997).

40. M. Fleischhauer, A. Imamoglu, and J. P. Marangos, “Electromagnetically induced transparency: optics in coherent media,” *Rev. Mod. Phys.* **77**(2), 633–673 (2005).
41. D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization,” arXiv:1412.6980 (2017).
42. 5-qubit backend: IBM Q Team, “IBM Q 5 Ourense backend specification V1.3.5,” 2020, <https://quantum-computing.ibm.com>
43. M. Motta et al., “Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution,” *Nat. Phys.* **16**(2), 205–210 (2020).
44. C. Cao et al., “Quantum imaginary time evolution steered by reinforcement learning,” arXiv: 2105.08696 (2021).
45. H. J. Kimble, “The quantum internet,” *Nature* **453**(7198), 1023–1030 (2008).
46. Z. Gu, C. Yang, and J. F. Chen, “Characterization of the photon-number state of a narrow-band single photon generated from a cold atomic cloud,” *Opt. Commun.* **439**, 206–209 (2019).
47. Y. Yu et al., “Entanglement of two quantum memories via fibres over dozens of kilometres,” *Nature* **578**(7794), 240–245 (2020).
48. C. Li et al., “Quantum communication between multiplexed atomic quantum memories,” *Phys. Rev. Lett.* **124**(24), 240504 (2020).
49. K. S. Choi et al., “Entanglement of spin waves among four quantum memories,” *Nature* **468**(7322), 412–416 (2010).
50. Y. Pu et al., “Experimental entanglement of 25 individually accessible atomic quantum interfaces,” *Sci. Adv.* **4**(4), eaar3931 (2018).
51. B. Xu et al., “Empirical evaluation of rectified activations in convolutional network,” arXiv:1505.00853 (2015).
52. R. Di Leonardo, F. Ianni, and G. Ruocco, “Computer generation of optimal holograms for optical trap arrays,” *Opt. Express* **15**(4), 1913–1922 (2007).
53. F. Nogrette et al., “Single-atom trapping in holographic 2D arrays of microtraps with arbitrary geometries,” *Phys. Rev. X* **4**(2), 021034 (2014).

Ying Zuo received her BSc degree from the University of Science and Technology of China, Hefei, Anhui, China, in 2017. Upon completion

of this work, she is a PhD candidate at the Hong Kong University of Science and Technology, supervised by Prof. Shengwang Du and Prof. Bei Zeng. Her research interests focus on optical computing and quantum optics.

Chenfeng Cao received his BSc degree from the University of Chinese Academy of Sciences, Beijing, China, in 2019. Currently, he is a PhD candidate supervised by Prof. Bei Zeng at the Hong Kong University of Science and Technology. His research interests lie in quantum information and quantum computing.

Ningping Cao received her PhD from the University of Guelph, Ontario, Canada, in 2021. Currently, she is a postdoc fellow at the Institute of Quantum Computing (IQC), University of Waterloo, Ontario, Canada. She is interested in a wide range of topics in quantum information and quantum computation.

Xuanying Lai received her BS degree from East China Normal University in 2020. Currently, she is pursuing her PhD under the supervision of Prof. Shengwang Du in the Department of Physics at the University of Texas at Dallas. Her research focus is quantum optics.

Bei Zeng is a professor at the Hong Kong University of Science and Technology. Her research focus is on the design of quantum error-correcting codes with nice properties that are suitable for high rate quantum information transmission through practical physical channels, and reliable quantum computation with high noise tolerance and low resource requirement. She is a fellow of APS.

Shengwang Du, currently a professor of physics at the University of Texas at Dallas since 2021, worked at the Hong Kong University of Science and Technology from 2008 to 2020. His group is exploring fundamentals in the field of atomic, molecular, and optical (AMO) physics, and their applications. His current research activities include quantum networks, all-ONNs, and applied optical microscopy. He is a fellow of both APS and Optica (formerly OSA).