Optical Neural Network Quantum State Tomography

Ying Zuo[a,+] , Chenfeng Cao[a,+] , Ningping Cao[b,c], Xuanying Lai[d], Bei Zeng[a*], Shengwang Du[d]**

[a]The Hong Kong University of Science and Technology, Department of Physics, Clear Water Bay, Kowloon, Hong Kong, China
[b]University of Guelph, Department of Mathematics & Statistics, Guelph, Canada, N1G 2W1 ON
[c]University of Waterloo, Institute for Quantum Computing, Waterloo, Canada, ON N2L 3G1
[d]The University of Texas at Dallas, Department of Physics, Richardson, Texas, USA, 75080
[+]Those authors contribute equally to this work

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. Here, 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.

*Bei Zeng, zengb@ust.hk
**Shengwang Du dusw@utdallas.edu

1 Introduction

Quantum state tomography (QST) is a standard process of reconstructing quantum information of an unknown quantum state through measurements on its copies. QST is used to verify state preparation, exam 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 $\Theta\left(2^{2n}/\varepsilon^2\right)$ 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] shadow tomography[10,11]) can exponentially reduce the computational resource for learning some 2-outcome measurement expectation values or "shadows".

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] Hamiltonian reconstruction,[15] etc. 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 neural networks can also perform QST with dramatically lower cost.[21,22]

In this work, we demonstrate QST with an optical neural network (ONN). Several optical implementations for realizing fully-connected neural network hardware have been proposed and demonstrated recently.[23–28] Optical computing takes advantages of the bosonic wave nature of light: superposition and interference give rise to 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 to the electronic version, ONNs are ideal for dealing with visual signals and information which 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 low light level for large amount of optical neurons[25–27] . Without nonlinear activation functions, and ONN is always

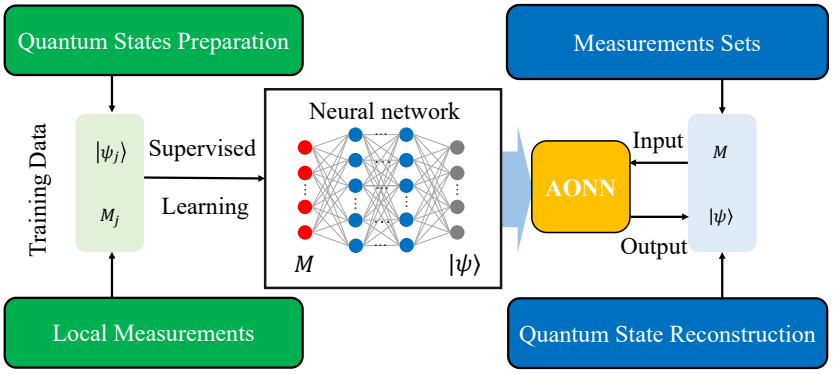

Fig 1 Schematics of optical neural network based quantum state tomography.

equivalent to a single-layer structure that cannot be applied for "real" deep machine learning. This problem has 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. Firstly we collect the training data set from a known quantum state $\{|\psi_j\rangle\}$ and the corresponding local measurements $\{M_j\}$. Secondly, we train neural networks under supervised learning with some nonlinear activation functions in their hidden neurons to obtain the optimal network parameters. Thirdly, we take the trained network parameters to configure the ONN and perform some fine adjustments to optimize the hardware performance. At 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 from 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 term) defined as

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

where $\mathcal{P} = \{\sigma_0 = I, \sigma_1 = X, \sigma_2 = Y, \sigma_3 = Z\}$. Every term in P is specified by its index $(i_1, i_2, \cdots, i_n)$. 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 a ket

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

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

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

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

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 $P^m = \{P_1, \cdots, P_m\}$ of m Pauli

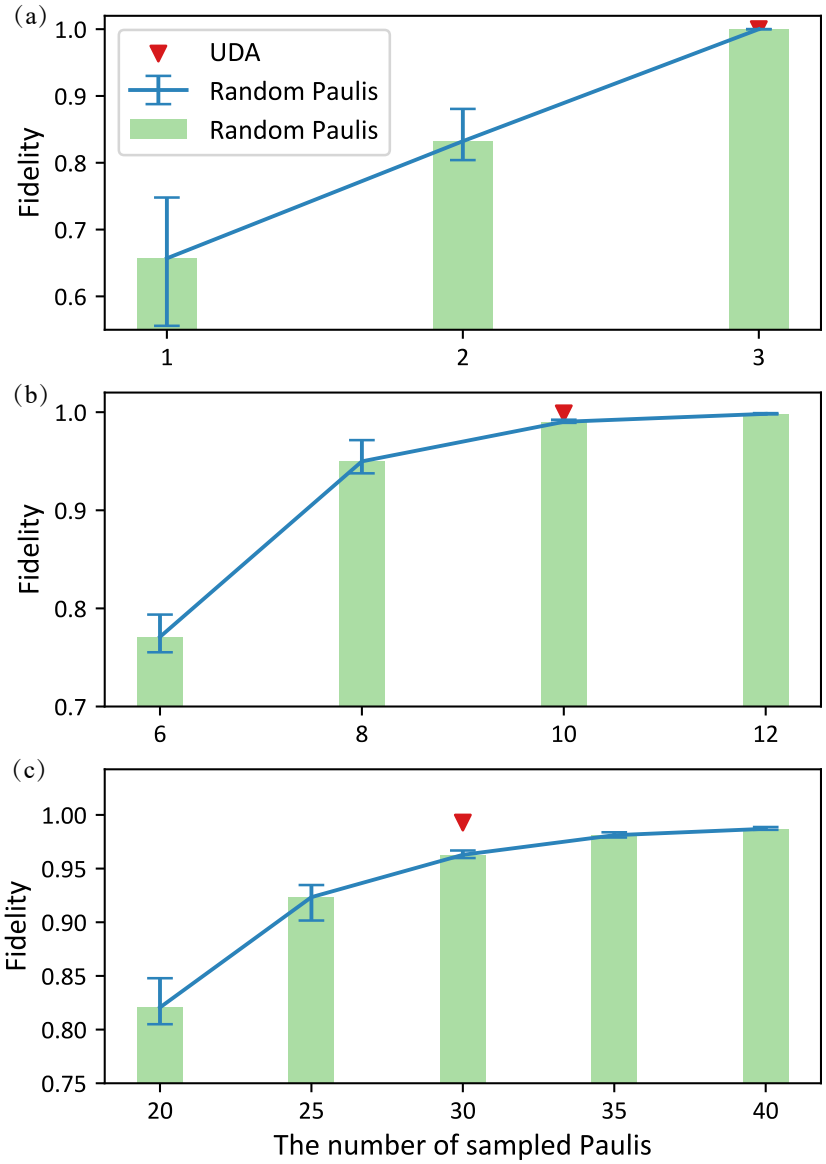

Fig 2 The fidelities of NN predictions for different samples of Pauli operators: The red triangles are the average fidelities for UDA Pauli operator sets, which is very close to 1. A Pauli operator set is said to be "UDA" if measuring these operators can uniquely determine a pure state among all states. 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.

operators from P, then use $\vec{c} = \mathrm{tr}(\rho \cdot P^m) = (\mathrm{tr}(\rho P_1), \mathrm{tr}(\rho P_2), \cdots, \mathrm{tr}(\rho P_m))$ 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 parameters of ρ (e.g. $a_{k,\mathrm{r}}$ and $a_{k,\mathrm{im}}$).

NNs are excellent tools for solving regression problems. When using NNs for QST, the ex-

pectation values $\vec{c}$ from random-sampled P^m are inputs to the network, and state parameters ($a_{k,\mathrm{r}}$, $a_{k,\mathrm{im}}$) are the outputs. Compared to 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 letter – 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 $(a_{k,\mathrm{r}}, a_{k,\mathrm{im}})$ 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 $m \in [6, 8, 10, 12]$; 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 3 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, a fidelity of higher than 99.9% requires more than 35 randomly sampled Paulis [Fig. 2(c)]. Details of training can be found in Supplemental Material S1.

Theoretically, a pure state ρ is Uniquely Determined among all states (UDA) of a set of operators F means that there is no other state, pure or mixed, has the same expectation values while measuring F.[35] In Ref.,[36] authors discovered two sets of Pauli operators, $P_{\text{2-UDA}}$ and $P_{\text{3-UDA}}$, that

are UDA for all 2-qubit and 3-qubit pure states respectively. (See Supplemental Material S2 for the particular sets $P_{\text{2-UDA}}$ and $P_{\text{3-UDA}}$.) 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 Figure 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

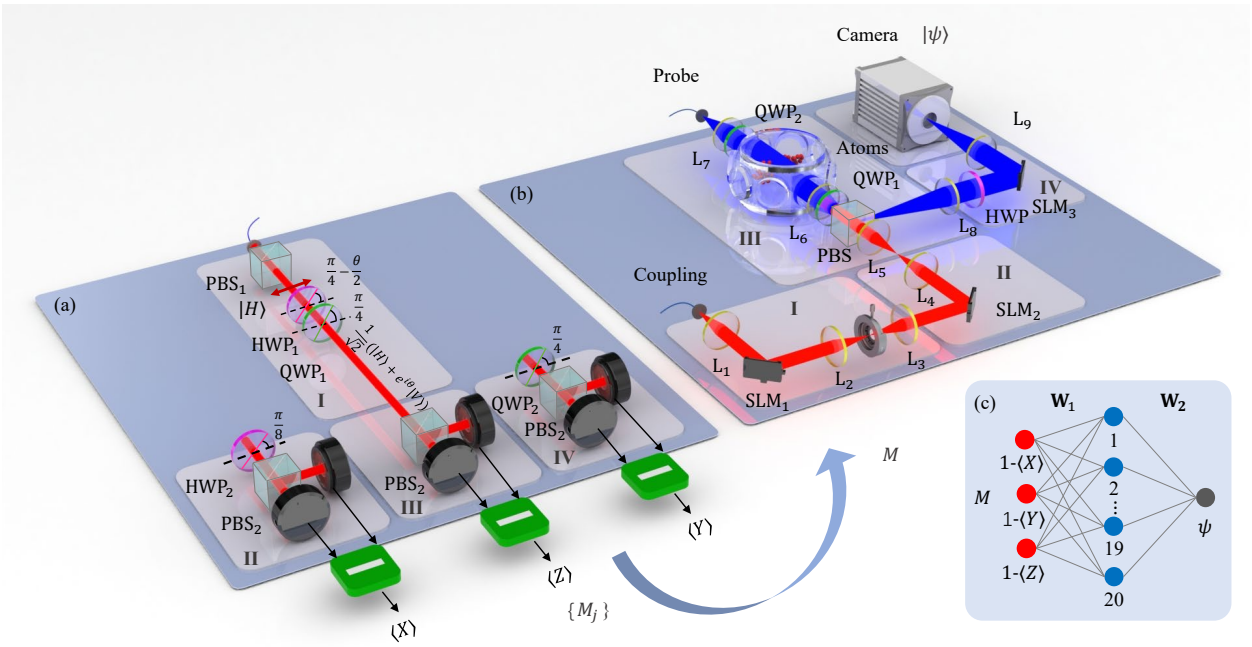

Fig 3 Schematics of optical implementation of quantum state tomography. (a) Optical layout of qubit quantum state tomography, including generation of polarization state (top panel), measurement of $\langle Z\rangle$, $\langle X\rangle$ and $\langle Y\rangle$ (bottom panel). The fast axis of the HWP$_1$ is aligned with an angle $\pi/4-\theta/2$ to the horizontal direction. The fast axis of the QWP$_1$ is aligned with an angle $\pi/4$ to the horizontal direction. (b) Schematics of optical neural network. I. Input generation. II. Linear operation of the first layer. III. Nonlinear operation. IV. Linear operation of the second layer. Spatial light modulators: SLM1 (HOLOEY LETO), SLM2(HOLOEYE PLUTO-2), and SLM3(HOLOEY GEAE-2). Camera: Hamamatsu C11440-22CU. PBS: polarization beam splitter. Lenses: L1-L9 Atoms are trapped in magneto-optical trap. (c) The neural network structure employed.

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 Supplemental Material S3 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 3 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 (SLM1) [Fig. 3(b)], lenses L2 and L3, and an aperture, as shown in unit

I of Fig.3(b). The SLM1 is divided into 3 parts and each part is encoded with sine phase pattern $m\pi \sin(\frac{2\pi}{T_{mj}}j + \frac{2\pi}{T_{mk}}k)$, 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 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 collimated incident to the SLM2. These weighted beams, as the input vector, are incident on SLM2, which diffracts each beam into 20 directions with designed weight (See Supplemental Material S4 for the algorithm 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 ^{85}Rb atoms in a two-dimensional (2D) magneto-optical trap (MOT),[37,38] where these 20-spot coupling beam patterns spatially modulate the transparency of the atomic medium through electromagnetically induced transparency (EIT).[39,40] Another relatively weak collimated probe beam counter-propagates 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 follows Equation 4.

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

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 lens 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 ref.[28,29]

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 neural network limit its ability. Therefore the ONN is only able to perform regression task on increasing and convex functions. To match the ONN constrains, 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 that this specific ONN fails to describe the whole range of nonmonotonic 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 measure 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 computer ibmq_ourense,[42]

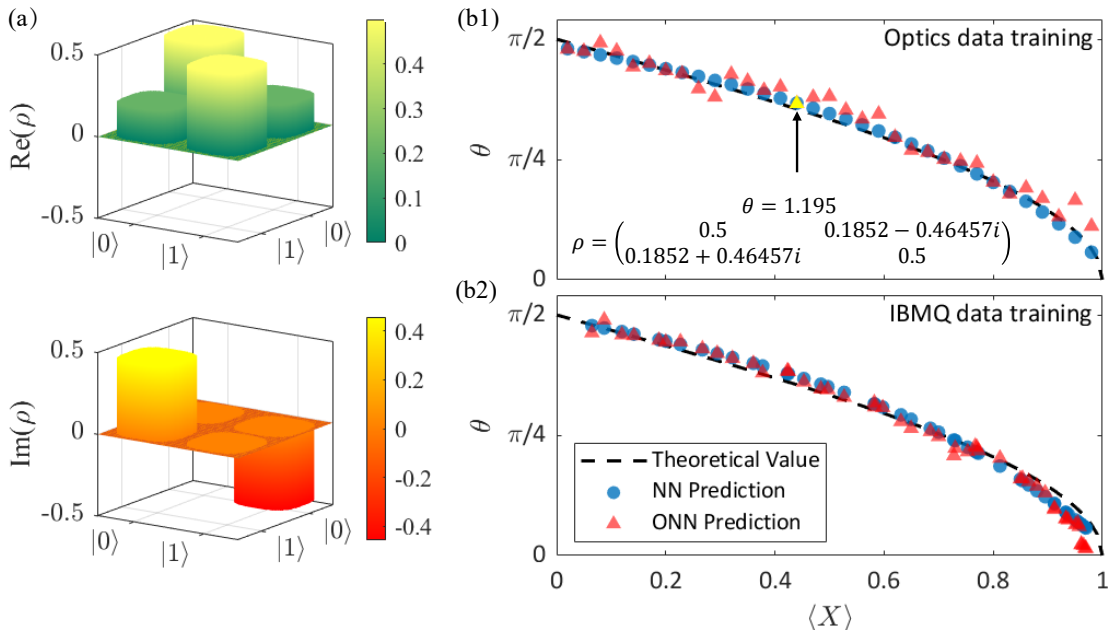

Fig 4 (a) Optical tomography of qubit (b) Experimental ONN tomography result. The ONN is training by optical tomography data (b1) and IBMQ tomography data (b2). The black dashed line is the theoretical value of the phase θ according to the $\langle X\rangle$. The blue circles are the phase θ numerically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

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 a 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 Supplemental Material S5.

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

ONN experimental predicted state is $\theta = 1.195$ 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 neural network predicted value $\theta = 1.1532$. The state is also marked with 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, neural network 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 for qubit QST.

4 Discussion and conclusion

While most demonstrations of optical neural networks took classification tasks to verify their feasibility,[26,27,30] we performed the first regression task, i.e., ONN-QST. To accomplish regression tasks, nonlinear function is essential as long as the relation between 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 limitation 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- 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 its impossible to conduct the regression task of non-monotonic 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 network[45] has been brought to the fore by the reduced decoherence and high speed of photons. Recently, apart from generating optical quantum states[46] and optical quantum communication over long distance,[47] multiple state-of-the-art experiments on optical quantum interfaces to store[48] 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.

Disclosures

The authors declare no competing interests.

Acknowledgments

Authors acknowledge the use of IBM Quantum services for this work.

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, F. Hufnagel, D. Koutnỳ, *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," *Advances in Imaging and Electron Physics* **128**, 206–309 (2003).

3 U. Leonhardt, "Quantum-state tomography and discrete wigner function," *Phys. Rev. Lett.* **74**(21), 4101 (1995).

4 R. Thew, K. Nemoto, A. G. White, *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**(1), 299 (2009).

6 M. Rambach, M. Qaryan, M. Kewming, *et al.*, "Robust and efficient high-dimensional quantum state tomography," *Phys. Rev. Lett.* **126**, 100402 (2021).

7 R. O'Donnell and J. Wright, "Efficient quantum tomography," in *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, 899–912 (2016).

8 S. Aaronson, "The learnability of quantum states," *Proc. R. Soc. A.* **463**(2088), 3089–3114 (2007).

9 S. Aaronson, X. Chen, E. Hazan, *et al.*, "Online learning of quantum states," *J. Stat. Mech.* **2019**(12), 124019 (2019).

10 S. Aaronson, "Shadow tomography of quantum states," *SIAM Journal on Computing* **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," *Nature Physics* **16**(10), 1050–1057 (2020).

12 M. Y. Niu, S. Boixo, V. N. Smelyanskiy, *et al.*, "Universal quantum control through deep reinforcement learning," *npj Quantum Inf.* **5**(1), 1–8 (2019).

13 Z. An, H.-J. Song, Q.-K. He, *et al.*, "Quantum optimal control of multilevel dissipative quantum systems with reinforcement learning," *Phys. Rev. A* **103**, 012404 (2021).

14 N. Cao, J. Xie, A. Zhang, *et al.*, "Neural networks for quantum inverse problems," *arXiv:2005.01540* (2021).

15 C. Cao, S.-Y. Hou, N. Cao, *et al.*, "Supervised learning in hamiltonian reconstruction from local measurements on eigenstates," *J. Phys.: Condens. Matter* **33**(6), 064002 (2020).

16 T. Xin, S. Lu, N. Cao, *et al.*, "Local-measurement-based quantum state tomography via neural networks," *npj Quantum Inf.* **5**(1), 1–8 (2019).

17 G. Torlai, G. Mazzola, J. Carrasquilla, *et al.*, "Neural-network quantum state tomography," *Nat. Phys.* **14**(5), 447 (2018).

18 A. M. Palmieri, E. Kovlakov, F. Bianchi, *et al.*, "Experimental neural network enhanced quantum tomography," *npj Quantum Inf.* **6**(1), 1–5 (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, C. Sánchez Muñoz, F. Nori, *et al.*, "Classification and reconstruction of optical quantum states with deep neural networks," *Phys. Rev. Research* **3**, 033278 (2021).

21 J. Carrasquilla, G. Torlai, R. G. Melko, *et al.*, "Reconstructing quantum states with generative models," *Nat. Mach. Intell* **1**(3), 155–161 (2019).

22 S. Ahmed, C. Sánchez Muñoz, F. Nori, *et al.*, "Quantum state tomography with conditional generative adversarial networks," *Phys. Rev. Lett.* **127**, 140502 (2021).

23 G. Wetzstein, A. Ozcan, S. Gigan, *et al.*, "Inference in artificial intelligence with deep optics and photonics," *Nature* **588**(7836), 39–47 (2020).

24 B. J. Shastri, A. N. Tait, T. F. de Lima, *et al.*, "Photonics for artificial intelligence and neuromorphic 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, N. C. Harris, S. Skirlo, *et al.*, "Deep learning with coherent nanophotonic circuits," *Nat. Photonics* **11**(7), 441 (2017).

27 X. Lin, Y. Rivenson, N. T. Yardimci, *et al.*, "All-optical machine learning using diffractive deep neural networks," *Science* **361**(6406), 1004–1008 (2018).

28 Y. Zuo, B. Li, Y. Zhao, *et al.*, "All-optical neural network with nonlinear activation functions," *Optica* **6**(9), 1132–1137 (2019).

29 Y. Zuo, Y. Zhao, Y.-C. Chen, *et al.*, "Scalability of all-optical neural networks based on spatial light modulators," *Phys. Rev. Applied* **15**, 054034 (2021).

30 J. Feldmann, N. Youngblood, C. D. Wright, *et al.*, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature* **569**(7755), 208–214 (2019).

31 X. Guo, T. D. Barrett, Z. M. Wang, *et al.*, "Backpropagation through nonlinear units for the all-optical training of neural networks," *Photon. Res.* **9**, B71–B80 (2021).

32 A. Ryou, J. Whitehead, M. Zhelyeznyakov, *et al.*, "Free-space optical neural network based on thermal atomic nonlinearity," *Photonics Res.* **9**(4), B128–B134 (2021).

33 D. Gross, Y.-K. Liu, S. T. Flammia, *et al.*, "Quantum state tomography via compressed sensing," *Phys. Rev. Lett.* **105**(15), 150401 (2010).

34 S. T. Flammia, D. Gross, Y.-K. Liu, *et al.*, "Quantum tomography via compressed sensing: error bounds, sample complexity and efficient estimators," *New J. Phys.* **14**(9), 095022 (2012).

35 J. Chen, H. Dawkins, Z. Ji, *et al.*, "Uniqueness of quantum states compatible with given measurement results," *Phys. Rev. A* **88**(1), 012109 (2013).

36 X. Ma, T. Jackson, H. Zhou, *et al.*, "Pure-state tomography with the expectation value of pauli operators," *Phys. Rev. A* **93**, 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, J. F. Chen, C. Liu, *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)**(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 (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,". Retrieved from https://quantum-computing.ibm.com* (2020).

43 M. Motta, C. Sun, A. T. Tan, *et al.*, "Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution," *Nature Physics* **16**(2), 205–210 (2020).

44 C. Cao, Z. An, S.-Y. Hou, *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, F. Ma, X. Y. Luo, *et al.*, "Entanglement of two quantum memories via fibres over dozens of kilometres," *Nature* **578**(7794), 240–245 (2020).

48 C. Li, N. Jiang, Y. K. Wu, *et al.*, "Quantum Communication between Multiplexed Atomic Quantum Memories," *Phys. Rev. Lett.* **124**(24), 1–6 (2020).

49 K. S. Choi, A. Goban, S. B. Papp, *et al.*, "Entanglement of spin waves among four quantum memories," *Nature* **468**(7322), 412–418 (2010).

50 Y. Pu, Y. Wu, N. Jiang, *et al.*, "Experimental entanglement of 25 individually accessible atomic quantum interfaces," *Sci. Adv.* **4**(4), eaar3931 (2018).

51 B. Xu, N. Wang, T. Chen, *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, H. Labuhn, S. Ravets, *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 B.Sc. from University of Science and Technology of China, Hefei, Anhui, China, in 2017. Upon completion of this work, she is a Ph.D. 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 B.Sc. from University of Chinese Academy of Sciences, Beijing, China, in 2019. He is currently a Ph.D. candidate supervised by Professor 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 Ph.D. from the University of Guelph, Ontario, Canada, in 2021. She is currently 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 as a student under the supervision of Prof. Shengwang Du at the Department of Physics, The University of Texas at Dallas. Her research focus is quantum optics.

Bei Zeng is a professor of 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, had 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-optical neural networks, and applied optical microscopy. He is a Fellow of both APS and Optica (formerly OSA).

List of Figures

3 Schematics of optical implementation of quantum state tomography. (a) Optical layout of qubit quantum state tomography, 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 HWP_1 is aligned with an angle $\pi/4 - \theta/2$ to the horizontal direction. The fast axis of the QWP_1 is aligned with an angle $\pi/4$ to the horizontal direction. (b) Schematics of optical neural network. I. Input generation. II. Linear operation of the first layer. III. Nonlinear operation. IV. Linear operation of the second layer. Spatial light modulators: SLM1 (HOLOEY LETO), SLM2(HOLOEYE PLUTO-2), and SLM3(HOLOEY GEAE-2). Camera: Hamamatsu C11440-22CU. PBS: polarization beam splitter. Lenses: L1-L9 Atoms are trapped in magneto-optical trap. (c) The neural network structure employed.

4 (a) Optical tomography of qubit (b) Experimental ONN tomography result. The ONN is training by optical tomography data (b1) and IBMQ tomography data (b2). 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 neural network, and the red triangles are the experimentally measured predictions of θ according to $\langle X \rangle$.The yellow triangle is an example of ONN experimental predicted state.