A comparative study on how neural networks enhance quantum state tomography
Hailan Ma, Daoyi Dong, Ian R. Petersen, Chang-Jiang Huang, Guo-Yong Xiang

Abstract—Quantum state tomography aiming at reconstructing the density matrix of a quantum state plays an important role in various emerging quantum technologies. Inspired by the intuition that machine learning has favorable robustness and generalization, we propose a deep neural networks based quantum state tomography (DNN-QST) approach, that can be applied to three cases, including few measurement copies and incomplete measurements as well as noisy measurements. Numerical results demonstrate that DNN-QST exhibits a great potential to achieve high fidelity for quantum state tomography with limited measurement resources and can achieve improved estimation when tomographic measurements suffer from noise. In addition, the results for 2-qubit states from quantum optical devices demonstrate the generalization of DNN-QST and its robustness against possible error in the experimental devices.

Index Terms—Quantum state tomography, Deep neural networks, Few copies, Incomplete measurement, Noise.

I. INTRODUCTION

Quantum state tomography (QST), which aims at reconstructing the state of a quantum system via quantum measurements [1] has drawn considerable efforts, owing to its significant role in verifying and benchmarking quantum tasks, including quantum computation [2], quantum control [3] and quantum communication [4].

To estimate a quantum state, a large number of measurements on an ensemble of identical quantum systems are performed to deduce its density matrix based on the measured outcomes. Several effective methods for QST have been proposed. Among them, least-squares inversion aims at solving the inverse of linear equations that relate the measured quantities to the density-matrix elements of a quantum system [5]. Bayesian tomography constructs a state using an integral averaging over all possible quantum states with proper weights [6], [7]. Maximum likelihood estimation (MLE) chooses the state estimate that maximizes the probability of the observed data, whose solution usually involves a large number of nonlinear equations [8], [9]. Linear regression estimation (LRE) solves the quantum state estimation problem using a linear model, and is usually combined with physical projection techniques to avoid the non-physical quantum states [10], [11].

Owing to its intrinsic capability of exacting information from high-dimensional data [12], [13], machine learning (ML) has been recently utilized to address various quantum tasks [14]. Among them, the representation of many-body quantum states [15], the verification of quantum devices [16], quantum error correction [17], quantum manipulation [18], [19] and quantum data compression [20], [21] have been investigated.

QST is a process of extracting useful information from experimental measurement statistics [9] and can be naturally regarded as a data processing problem. Recently, there have been many efforts to investigate the use of ML in quantum state tomography. For example, deep learning based quantum state tomography has been proposed to reconstruct states [22]. A convolutional neural network has been introduced to reconstruct 2-qubit quantum states from a set of coincidence measurements [23]–[25]. Reconstruction of optical quantum states with reduced data can be achieved via the utilization of conditional generative adversarial networks [26], [27]. ML has also been used to enhance the performance of experimental quantum state estimation [25], [28]. For example, neural networks act as a useful tool to denoise the state-preparation-and-measurement errors [29] and have achieved improved performance when combined with MLE. In [28], QST using neural networks has been benchmarked and compared on experimentally generated 2-qubit entangled states.

As is known, at least \((4^n - 1)\) real parameters are required to specify the density matrix of an \(n\)-qubit quantum state, which means that the required measurement resources scale exponentially with the qubit number \(n\) [30]. For solid-state systems, the measurement process can be time-consuming, which can limit the time available for performing measurements during the data collection procedure. Under limited measurement copies, the approximation of the observed distributions for systems with large sizes tends to be undersampled, resulting in inaccurate measured statistics. From this respective, the number of copies plays an essential role for practical quantum state tomography. Hence, it is meaningful to investigate the case of few measurement copies, which provides insight for QST on large quantum systems.

To determine a unique state, an informationally complete
set of measurement operators is required. For example, 2-qubit state tomography relies on the measured statistics of at least 16 measurements. In many solid-state systems, correlated multi-qubit measurement may not be realizable and experimental readout is often done via single-qubit projective operator measurements or single-spin measurements. In such cases, the measurement operators maybe incomplete. Although incomplete measurements do not characterize a unique state, additional constraints can be introduced to obtain physically valid estimates. For example, the maximum entropy principle finds a state consistent with the measured data with the largest von Neumann entropy. Variational quantum tomography returns a physically valid state that minimizes the expectation value of the missing operators. In this work, we also investigate the performance of machine learning aided QST in the case of incomplete measurements.

In real applications, quantum measurements are prone to errors and noise, which increases the difficulty of characterizing quantum states with high efficiency. For example, it has been found that the computational resources to fully characterize a quantum state scale as \( O(d^4) \) when the observed data suffers from Gaussian noise. Considering the favorable generalization properties of neural networks, we incorporate neural networks into QST to reconstruct quantum states using noisy measurements.

It is known that deep neural networks can approximate different patterns of data efficiently with favorable generalization properties. Considering that neural networks have the potential to characterize the features of quantum states and usually require reduced prior knowledge about quantum states, we propose a deep neural networks based quantum state tomography (DNN-QST) approach, that can be applied in the case of few measurement copies and incomplete measurements. To demonstrate its capability and efficiency, simulations are performed on different cases for comparison with traditional methods. A significant improvement of average fidelity over MLE and LRE demonstrates the advantage of DNN-QST in reconstructing quantum states with limited measurement resources. In addition, DNN-QST is applied to noisy measurement operators and exhibits favorable robustness in reconstructing quantum optical states when tomographic measurements suffer from noise.

The rest of this paper is organized as follows. Section II introduces several basic concepts about quantum state estimation and two traditional QST techniques. In Section III the DNN-QST method is presented in detail. Numerical results of QST with limited resources, i.e., few measurement copies and incomplete measurements are provided in Section IV. Section V investigates the performance of QST with noise. Concluding remarks are given in Section VI.

II. PRELIMINARIES

In this section, several basic concepts about quantum state tomography are introduced, and then two QST techniques, i.e., MLE and LRE are presented.

A. Quantum state tomography

Quantum state tomography is a process of estimating the density matrix (denoted as \( \rho \)) of a quantum state based on a set of measurements (denoted as Hermitian operators \( \mathcal{O} \)). The density matrix \( \rho \) is usually a positive-definite Hermitian operator with unit trace, that is (i) \( \rho = \rho^\dagger \) (where \( \rho^\dagger \) is the conjugate and transpose of \( \rho \)), (ii) \( \text{Tr}(\rho) = 1 \), and (iii) \( \rho \geq 0 \). Measurement operators are usually positive-operator-valued measurements (POVMs), and can be represented as a set of positive semi-definite matrices \( \mathcal{O}_k \) that sum to identity (\( \sum \mathcal{O}_k = I \)). According to Born’s rule, when measuring the quantum state \( \rho \) with the measurement operator \( \mathcal{O}_k \), the probability of obtaining the outcome \( k \) is calculated as \( p_k = \text{Tr}(\mathcal{O}_k \rho) \).

In principle, if the POVMs are tomographically complete and the probabilities \( p_k \) of different outcomes are available, it would be possible to reconstruct the true state \( \rho \) directly by inverting a linear relation. However, it is impossible to obtain the exact probabilities \( p_k \) since only a finite number \( (N_{\mathcal{O}}) \) of physical systems can be used for measurements. In the case of \( N_{\mathcal{O}} \) occurrences for the outcome \( \mathcal{O}_k \), the relative \( \hat{p}_k = \frac{N_{\mathcal{O}_k}}{N_{\mathcal{O}}} \) is an experimental approximation to the true probability \( p_k \). Hence, a critical goal of QST is to deduce the density matrix \( \rho \) based on the experimental statistics \( \{\hat{p}_k\} \).

B. MLE

The core idea of MLE is to maximize the probability of the observed data. If the probability of observing the outcome \( k \) is \( p_k \), then the probability of observing the outcome \( k \) with occurrence \( n_k \) is calculated as \( \prod_k p_k^{n_k} \). Performing a logarithm transformation and a division by \( N_{\mathcal{O}} \), a log-likelihood functional is obtained as \( \mathcal{L}(\rho) = \sum \hat{p}_k \ln \text{Tr}(\rho \mathcal{O}_k) \). As such, the solution of MLE can be reduced to searching for a density operator that generates probabilities \( \{\hat{p}_k\} \) that are as close to the observed frequencies \( \{\hat{p}_k\} \) as possible, which is formulated as

\[
\hat{\rho}_{\text{MLE}} = \arg \max_{\rho} \sum \hat{p}_k \ln \text{Tr}(\rho \mathcal{O}_k).
\]  

By analyzing the corresponding extremal equation, one can obtain a solution in a recursive way. (i) Initialize the state using a physical quantum state, e.g., \( \rho = 1/d \), where \( d \) is the dimension of the quantum system; (ii) Compute the evolving operator as \( R = \sum \frac{\hat{p}_k \mathcal{O}_k}{\text{Tr}(\mathcal{O}_k \rho)} \); (iii) Update the estimated density matrix as \( \rho' = \frac{\text{Tr}(\mathcal{O} \rho)}{\text{Tr}(\mathcal{O} \rho')} \); (iv) Terminate the iteration if the distance between \( \rho' \) and \( \rho \) falls below a given threshold. It is straightforward to check that the above solution satisfies the requirement of positive semidefiniteness and trace 1. From this respective, MLE always returns a physical estimate of a quantum state.

C. LRE

Quantum state tomography can be converted into a parameter estimation problem for a linear regression model.
For a system with dimension $d$, there exists a complete basis set of Hermitian and orthonormal operators $\{ \Omega_j \}_{j=0}^{d^2-1}$ on the corresponding Liouville space, that satisfies $\text{Tr}(\Omega^j \Omega) = \delta_{ij}$, where $\delta_{ij}$ denotes the Kronecker function. Without loss of generality, let $\Omega_0 = 1/d$, and the other bases be traceless, i.e., $\text{Tr}(\Omega_j) = 0$ for $j = 1, 2, \ldots, d^2-1$. Under such a basis set, the quantum state $\rho$ to be reconstructed may be parameterized as

$$\rho = \frac{1}{d} + \sum_{j=1}^{d^2-1} \gamma_j \Omega_j,$$

where $\gamma_j = \text{Tr}(\rho \Omega_j)$ is a real number. In fact, $\gamma_0 = 1$ remains unchanged, and it is preferable to take the effective parametrization vector as $\Gamma = (\gamma_1, \gamma_2, \ldots, \gamma_{d^2-1})^T$. Similarly, each measurement operator $\Omega_k$ can be parameterized as

$$\Omega_k = \frac{1}{d} + \sum_{i=1}^{d^2-1} \lambda_{ij}^{(k)} \Omega_j,$$

with $\lambda_{ij}^{(k)} = \text{Tr}(\Omega_k \Omega_j)$. When performing $\Omega_k$ on the quantum state $\rho$, the probability of observing the outcome $k$ is calculated as

$$p_k = \text{Tr}(\Omega_k \rho) = \frac{1}{d} + \sum_{j=1}^{d^2-1} \gamma_j \lambda_{ij}^{(k)}.$$

Denote $e_k = p_k - p_k$ and $\Lambda^{(k)} = (\lambda_{11}^{(k)}, \lambda_{12}^{(k)}, \ldots, \lambda_{d^2-1}^{(k)})^T$, a linear regression equation is obtained as $p_k = \frac{1}{d} + \Gamma^T \Lambda^{(k)} + e_k$. As such, the problem of QST is reduced to the estimation of the unknown vector $\Gamma$. Denote $Y = (\rho_1 - \frac{1}{d}, \ldots, \rho_M - \frac{1}{d})$, $X = (\Lambda^{(1)}, \ldots, \Lambda^{(M)})$, $E = (e_1, \ldots, e_M)^T$. A set of linear regression equations can be represented in a compact form $Y = X\Gamma + E$ [10], [11]. Finally, a least-square solution can be obtained as

$$\Gamma_{LS} = (X^T WX)^{-1}X^T W Y,$$

where $W$ represents a diagonal weighting matrix. In fact, the solution to $\Gamma$ can also be calculated in an iterative way [10]. Finally, an estimated density matrix $\hat{\rho}$ is obtained based on (2). However, it is clear that $\hat{\rho}$ satisfies (i) $\hat{\rho} = \hat{\rho}^\dagger$ and (ii) $\text{Tr}(\hat{\rho}) = 1$, but might have negative eigenvalues due to inaccuracies. Hence, it is essential to pull it back to a physical state using a projection technique [37].

### III. Deep Neural Networks Based Quantum State Tomography

In this section, a machine learning aided quantum state tomography method (called as DNN-QST) that takes advantage of deep neural networks to reconstruct quantum states is proposed. Since QST involves measuring quantum states using a set of measurement projectors, the quantum states and the measurement operators are important ingredients in DNN-QST. Here, quantum state generation and its representation are firstly presented, and then different measurement settings are introduced. Finally, the DNN-QST method is provided in detail.

#### A. Quantum state generation and its representation

The core idea of deep neural networks lies in learning patterns from big data. The first key step is to generate samples for learning. Let $\mathcal{U}^d$ be the set of all $d$-dimensional unitary operators, i.e., any $U \in \mathcal{U}^d$ satisfies $UU^\dagger = U^\dagger U = 1$. Owing to its invariance under group multiplication (i.e., any region of $\mathcal{U}^d$ carries the same weight in a group average) [39], the Haar metric is utilized as a probability measure on a compact group. As such, random pure states can be generated by performing random unitary transformations following the Haar measure [11], [40], which can be formulated as $|\psi\rangle_{Haar} = U_{Haar}|\psi_0\rangle$, where $|\psi_0\rangle$ is a fixed pure state.

In addition, we also consider mixed states in the following form

$$\rho_p = p|\psi\rangle\langle\psi|_{Haar} + (1-p)\frac{1}{d},$$

where $p \in (0,1)$ represents the ratio of pure element. The purity of the quantum state in (3) is closely related to the value of $p$. That is

$$\text{Tr}(\rho_p^2) = (1 - \frac{2}{d})p^2 + \frac{p}{d} + \frac{1}{d}.$$ 

To achieve a unified notation for pure states and mixed states, a density matrix $\rho_d$, i.e., a $d \times d$ matrix is utilized to describe a quantum state.

Although there are many ways to generate a Hermitian operator, the condition of positivity is difficult to guarantee. Given any lower triangular matrix $\rho_L$, a physical density matrix can be obtained as

$$\rho = \frac{\rho_L \rho^\dagger_L}{\text{Tr}(\rho_L \rho^\dagger_L)}.$$ 

It is easy to check that the density matrix in (8) satisfies the three conditions as described in Section II-A. In addition, according to the Cholesky decomposition [41], for any physical density matrix, there exists a lower triangular matrix $\rho_L$ that achieves

$$\rho_L \rho^\dagger_L = \rho.$$ 

Note that a tiny perturbation term (e.g., $\varepsilon = 10^{-7}$) is usually added to the simulated pure states to avoid convergence issues using the Cholesky decomposition [41]

$$\rho = (1 - \varepsilon)|\psi\rangle\langle\psi| + \frac{\varepsilon}{d} \frac{1}{d}.$$ 

As such, for both pure states and mixed states, there exists a corresponding lower triangular matrix $\rho_L$ associated with a density matrix $\rho$. Hence, the search for a physical quantum state can be converted to the search for a lower triangular matrix, which can be further transformed into a real vector by considering the real and the imaginary parts.

#### B. Quantum measurement settings

Quantum state tomography involves measurements to obtain a set of expectation values. Theoretically, the minimum number of measurements to fully characterize a quantum state with dimension $d$ is $d^2$ [31]. For example, a linearly
An improved estimate of 2-qubit quantum states can be obtained when tomography is performed using projections onto 36 tensor products of Pauli eigenstates \[42\]. In addition, mutually unbiased bases where all inner products between projectors of different bases are equal to \(\frac{1}{2}\) have the potential to maximize information extraction per measurement \[43\]. Hence, the measurement operators play an important role in the estimation problem.

Before introducing different quantum measurement operators, several important concepts are introduced. Denote the Pauli matrices as \(\sigma = (\sigma_x, \sigma_y, \sigma_z)\), with

\[
\sigma_x = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.
\]

Let the eigen-vectors of \(\sigma_x\) be the diagonal state \(|D\rangle\) and the anti-diagonal state \(|A\rangle\). Define the eigen-vectors of \(\sigma_y\) as the left circular polarization state \(|L\rangle\) and the right circular polarization state \(|R\rangle\) and take the eigen-vectors of \(\sigma_z\) as the horizontal state \(|H\rangle\) and the vertical state \(|V\rangle\). These six basis states form the 1-qubit Pauli measurement

\[
P_{\text{pauli}} = \{|H\rangle, |V\rangle, |D\rangle, |A\rangle, |R\rangle, |L\rangle\}.
\]

It is clear that they are informationally complete for reconstructing 1-qubit quantum states.

In this work, we consider two types of measurement settings. The first one involves tensor products of Pauli matrices, which is also called the cube measurement \[42\]. The second one is the mutually unbiased bases (MUB) measurement \[43\]. Generalizing the 1-qubit Pauli measurement in \[12\] to \(n\)-qubits using the tensor product, we obtain the following Pauli measurement basis

\[
P = P_1 \otimes P_2 \cdots \otimes P_n, \quad \text{with} \quad P_i \in P_{\text{pauli}}.
\]

For \(n\)-qubits, there are \(6^n\) measurement basis operators involved in Pauli measurement or cube measurement. Typically, the \(6^n\) measurement operators are arranged into several sets, with each set containing \(d = 2^n\) orthogonal projectors. These \(6^n\) basis states for the cube measurement can be grouped into \(3^n\) sets of orthogonal projectors \[43\].

Consider that the set of \((4^n - 1)\) Pauli operators that are required to determine an arbitrary mixed state can be partitioned into \((2^n + 1)\) distinct subsets, each consisting of \((2^n - 1)\) internally commuting observables \[44\]. The MUB measurement usually includes \((2^n + 1)\) set of measurement basis states \[43\].

In real applications, measurement usually suffers from noise. In this work, we consider unitary rotations on the measurement operators to simulate the noise. For 1-qubit, an arbitrary rotation operator can be defined as \[45\]

\[
U(\theta_1, \theta_2, \theta_3) = \begin{bmatrix} e^{i\theta_1 \cos(\theta_2)} & -ie^{i\theta_1 \sin(\theta_2)} \\ -ie^{-i\theta_1 \sin(\theta_2)} & e^{-i\theta_1 \cos(\theta_2)} \end{bmatrix}.
\]

For \(n\)-qubit systems, we consider the local noise, with the unitary rotation transformation defined as

\[
U_e = U_{\theta_1} \otimes U_{\theta_2} \cdots \otimes U_{\theta_n},
\]
represented as a real vector with length $d^2$, and is called the\textbf{ α-vector}. In particular, the positive diagonal entries are first ordered, and then the real and imaginary parts of off-diagonal elements are considered. Finally, a layer with $d^2$ neurons is designed as the output layer to generate the α-vector, which can be utilized to produce physical density matrices with additional transformations.

The schematic of DNN-QST is summarized in Fig. 1. In (a), samples are generated from the density matrices and measurement operators, and the useful data includes the measured frequencies and the target vectors calculated by the Cholesky decomposition. In (b), an architecture that includes input-hidden-output layers is utilized as a parameterized function to map a feature vector comprising of the measurement outcomes to the α-vector, which is then transformed into a physical density matrix (see (c)). The loss function for training the network adopts the mean square error (MSE) between the reconstructed density matrix of a quantum state). Also, the networks are trained using a gradient descent algorithm with AdamOptimizer to minimize the MSE.

The process of reconstructing quantum states using DNN-QST is carried in two stages. In the first stage, the networks are iteratively optimized using the training samples and the parameters are updated with the purpose of minimizing the MSE. After the weights of the neural networks have been learnt, we come to the second stage, i.e., to apply the neural networks to reconstruct the density matrices of quantum states from new measurement outcomes. Unlike the training process, the procedures of $α \rightarrow \rho_L \rightarrow \rho$ are required to obtain physically valid quantum states. Although the first stage requires many iterations to optimize, the second stage only requires a single feed-forward calculation without iterations. From this respective, DNN-QST is efficient in reconstructing quantum states.

To demonstrate the efficiency of the DNN-QST method, the fidelity between the reconstructed state and the real state is evaluated, which is formulated as

$$ F(\rho, \sigma) = |\text{Tr}(\sqrt{\sqrt{\rho} \sigma \sqrt{\rho}})|. $$

(16)

Furthermore, the infidelity of two states is defined as $F = 1 - F(\rho, \sigma)$, and can also be used for evaluation.

IV. NUMERICAL RESULTS ON QST WITH LIMITED RESOURCES

To test the performance of the DNN-QST method, several groups of numerical simulations are carried out on different states under different scenarios. In this subsection, parameter settings are first introduced, and then QST using few measurement copies and incomplete measurements is investigated and analyzed to demonstrate the effectiveness of DNN-QST compared with LRE and MLE.

### A. Parameter settings

For each case, a large number of samples are generated by randomly sampling the parameters. 98800 samples are used for training the parameters of DNN-QST. Then the DNN-QST model is tested on 1000 samples. In addition, those 1000 samples are also tested on LRE and MLE. For simulations, the Pytorch framework is utilized to construct the deep neural networks to run the model. In particular, three hidden layers are utilized, with each hidden layer using the same neurons. The number of neurons for the hidden layers is set as 128 for 2-qubit pure states, and is set as 256 for 2-qubit mixed states and 3-qubit states. All neurons are activated by the leaky relu function except the one before the α-vector.

For MLE, the iteration process is terminated once the gap of infidelity between two successive runs falls below $10^{-8}$. For LRE, a fast version of projection is performed to project a Hermitian operator with negative eigenvalues to a physical density matrix. For 2-qubit quantum systems, their measurement basis states are summarized in Table I.

### B. Few copies

In real applications, a measurement on quantum systems may return a statistical value of probability, i.e., frequency. Consider that an individual outcome only provides limited information on the estimated state, an infinite number of measurements is required to determine a quantum state precisely. However, only finite measurements are available in practical quantum state tomography. In that case, the accuracy of measured frequencies can be influenced by the number of measurement copies used. In this subsection, we focus on the case of QST when the measurement copies are limited. In particular, we assume that the copies for each measurement are equal, denoted as $S$, and the total number of copies is $N = Sk$ with $k$ being the number of measurement operators.

The results for 2-qubit states with different purities are shown in Fig. 2. For $S = 10$ and $S = 100$, DNN achieves the best performance under both the cube measurement and the MUB measurement, followed by MLE. Also, graphs of infidelity vs purity for the three methods exhibit different trends. In particular, the infidelity of MLE increases with purity, while the infidelity of LRE first increases slowly and then drops greatly with purity. By comparison, the infidelity.

### Table I

| cube | MUB |
|------|-----|
| $|HH\rangle,|HV\rangle,|VH\rangle,|VV\rangle$ | $|HH\rangle,|HV\rangle,|VH\rangle,|VV\rangle$ |
| $|HD\rangle,|HA\rangle,|VD\rangle,|VA\rangle$ | $|RD\rangle,|RA\rangle,|LD\rangle,|LA\rangle$ |
| $|HR\rangle,|HE\rangle,|VR\rangle,|VL\rangle$ | $|DR\rangle,|DL\rangle,|AR\rangle,|AL\rangle$ |
| $|DH\rangle,|DV\rangle,|AH\rangle,|AV\rangle$ | $\frac{1}{2}(|RL\rangle + i|LR\rangle), \frac{1}{2}(|RL\rangle - i|LR\rangle),|RR\rangle + i|LL\rangle),$ $|RR\rangle - i|LL\rangle)$ |
| $|DD\rangle,|DA\rangle,|AD\rangle,|AA\rangle$ | $\frac{1}{2}(|RV\rangle + i|LV\rangle), \frac{1}{2}(|RV\rangle - i|LV\rangle),$ $|RH\rangle + i|HL\rangle), |RH\rangle - i|HL\rangle)$ |
| $|DR\rangle,|DL\rangle,|AR\rangle,|AL\rangle$ | $\frac{1}{2}(|RV\rangle + i|LV\rangle), \frac{1}{2}(|RV\rangle - i|LV\rangle),|RH\rangle + i|HL\rangle), |RH\rangle - i|HL\rangle)$ |
| $|RH\rangle,|RV\rangle,|LV\rangle,|VL\rangle$ | $\frac{1}{2}(|RV\rangle + i|LV\rangle), \frac{1}{2}(|RV\rangle - i|LV\rangle),|RH\rangle + i|HL\rangle), |RH\rangle - i|HL\rangle)$ |
| $|RD\rangle,|RA\rangle,|LD\rangle,|LA\rangle$ | $\frac{1}{2}(|RV\rangle + i|LV\rangle), \frac{1}{2}(|RV\rangle - i|LV\rangle),|RH\rangle + i|HL\rangle), |RH\rangle - i|HL\rangle)$ |
| $|RR\rangle,|RL\rangle,|LR\rangle,|LL\rangle$ | $\frac{1}{2}(|RV\rangle + i|LV\rangle), \frac{1}{2}(|RV\rangle - i|LV\rangle),|RH\rangle + i|HL\rangle), |RH\rangle - i|HL\rangle)$ |
Fig. 2. The performance for 2-qubit states with different purities using few measurement copies. (a1) Infidelity vs purity for cube basis with $S = 10$; (a2) Infidelity vs purity for cube basis with $S = 100$; (b1) Infidelity vs purity for MUB with $S = 10$; (b2) Infidelity vs purity for MUB with $S = 100$.

Fig. 3. The performance of 2-qubit states for pure states and mixed states using few measurement copies. (a1) Infidelity vs copies for cube basis for pure states; (a2) Infidelity vs copies for MUB for pure states; (b1) Infidelity vs copies for cube basis for mixed states with $p = 0.9$; (b2) Infidelity vs copies for MUB for mixed states with $p = 0.9$. 
of DNN first increases until around purity of 0.35 and drops slowly with purity for $S = 10$, and the infidelity of DNN nearly keeps the same with tiny fluctuation for $S = 100$.

Fig 3 demonstrates that the infidelity of the three methods decreases with the number of copies for each measurement on both the cube and the MUB measurement settings. For both pure states and mixed states with $p = 0.9$ (purity $\text{Tr}(\rho^2) = 0.88$), the performance of DNN is much better than the other two methods, and the advantage of DNN over LRE and MLE becomes obvious with an increasing number of copies. Unlike the case of pure states where MLE and LRE achieve similar performance in (a1)-(a2), the two methods achieve different estimation fidelities when reconstructing mixed states in (b1)-(b2).

To demonstrate the generalization of DNN-QST, the performance for 3-qubits using cube measurement is also investigated. Fig. 4 summarizes the results for 3-qubit pure states following the Haar metric. Judging from the decreasing curves, the estimation accuracy of the three methods improves with increasing number of copies for each measurement. In particular, DNN achieves the lowest infidelity, while MLE is better than LRE with an increasing number of measurement copies. The numerical results demonstrate that the proposed method has the potential to achieve efficient quantum state tomography using incomplete measurements.

C. Incomplete measurements

In this subsection, quantum state tomography using incomplete measurements is investigated. Recall the measurement projectors are usually arranged into several sets. In this work, the incompleteness of measurement operators is defined with respect of the number of measurement sets, rather than the number of measurement projectors. For 2-qubit systems, we consider 9 cases for the cube setting, including cube1, cube2, ..., cube9 and 5 cases for the MUB setting, including mub1, mub2, ..., mub5. Note that cube1 (mub1) includes the first $n$ sets of projectors from the complete cube (MUB) measurement operators following the order in Table 1.

The results of 2-qubit states with different purities are shown in Fig. 5. As we can see, DNN achieves the best fidelity for both the cube and the MUB measurements. Besides, LRE is superior to MLE for the complete measurements in (a1) and (b1), but is worse than MLE for incomplete measurements including cube5 and mub3. Besides, the curves of infidelity versus purity for the three methods exhibit different trends. In particular, the infidelity of DNN always increases with purity for the four cases, which is similar to that of LRE. By comparison, the infidelity of MLE increases with purity for complete measurements, but exhibits a trend of first increasing and then dropping with purity for incomplete measurements. In Fig. 6, DNN achieves better estimation accuracy than the other two methods for both pure states and mixed states with $p = 0.5$ (purity $\text{Tr}(\rho^2) = 0.5$). Also, LRE is sensitive to the completeness of measurement operators, since it achieves the worst performance when the number of sets is 4-8 for the cube measurement and 2-4 for the MUB measurement, but can achieve excellent infidelity for the complete cube and MUB measurements.

To verify the generalization of DNN-QST, we also implement the simulation for 3-qubit pure states following the Haar metric. From Fig. 7, LRE fails to achieve good estimation when the measurement is close to the complete case, suggesting that LRE is sensitive to the completeness of measurements. By comparison, DNN achieves better estimation than MLE with few measurement operators. Based on the above results, DNN has the advantage of reconstructing quantum states using incomplete measurements.

A. Numerical results for 2-qubit and 3-qubit states

In this subsection, we focus on the case of measurement projectors suffering from unitary noise as in (15). In particular, we consider two types of unitary noise: (i) noise parameters follow Gaussian distributions (denoted as $\mathcal{N}$); (ii) noise parameters follow uniform distributions (denoted as $\mathcal{U}$). For simplicity, three ratios $\xi_1, \xi_2, \xi_3$ are introduced. The noise parameters for the first case are sampled as $\theta_1 \sim \mathcal{N}(0, \pi\xi_1)$, $\theta_2 \sim \mathcal{N}(0, 2\pi\xi_2)$, $\theta_3 \sim \mathcal{N}(0, 2\pi\xi_3)$. For the uniform cases, the noise parameters are sampled as $\theta_1 \sim \mathcal{U}(0, 2\pi\xi_1)$, $\theta_2 \sim \mathcal{U}(0, 5\pi\xi_2)$, $\theta_3 \sim \mathcal{U}(0, 2\pi\xi_3)$. For these cases, during the data collecting process, the measured frequencies are observed using the noisy operator $U_M U\dagger$ rather than the ideal operator $M$. However, during
Fig. 5. The performance of 2-qubit states with different purities using incomplete measurements. (a1) Infidelity vs purity for cube9; (a2) Infidelity vs purity for cube5; (b1) Infidelity vs purity for mub5; (b2) Infidelity vs purity for mub3.

Fig. 6. The performance of 2-qubit pure states and mixed states using incomplete measurements. (a1) Infidelity vs sets of the cube measurement for pure states; (a2) Infidelity vs sets of the MUB measurement for pure states; (b1) Infidelity vs sets of the cube measurement for mixed states with $p = 0.5$; (b2) Infidelity vs sets of the MUB measurement for mixed states with $p = 0.5$. 
controls the path qubit of the photon. The polarization of the photon in each path is manipulated by an HWP and a quarter-wave plate (QWP). 2) **Unitary gate** (Fig. 11 (iii)): 2-qubit unitary gate is produced from another interferometer. A special beam splitter cube that is half PBS coated and half coated by a non-polarizer beam splitter (NBS) is applied in the junction of two Sagnac interferometers. Then four unitary polarization operators \( V_1, V_2, V_R \) and \( V_L \) are generated from the combination of two QWPs, an HWP, and a phase shifter (PS) composed of a pair of wedge-shaped plates. 3) **Measurements** (Fig. 11 (iv)-(v)): Measurements can be realized with the combination of a QWP, an HWP and a PBS.

Recall 2-qubit unitary gates can be realized by combining a path unitary gate with a polarization gate \([47]\).

Denote a 2-qubit unitary gate as \( U = \begin{bmatrix} U_{RR} & U_{RL} \\ U_{LR} & U_{LL} \end{bmatrix} \), where \( U_{RR}, U_{RL}, U_{LR}, U_{LL} \) are 2 \( \times \) 2 matrices following the form of \( U_{RR} = \frac{1}{2} V_2 (V_R + V_L) \), \( U_{LL} = \frac{1}{2} (V_R + V_L) \), \( U_{RL} = -\frac{i}{2} V_2 (V_R - V_L) \), and \( U_{LR} = \frac{i}{2} (V_R - V_L) \). Thus, a set of QWPs, HWPs, and phase shifters may generate the unitary operators \( V_1, V_2, V_R \) and \( V_L \).

By performing 2-qubit unitary gate on the initial states, the output states are used for tomography. In this work, mixed states with purity larger than 0.99 are selected as basis states. To include more quantum optical states with high purity, 1000 unitary gates are simulated on each basis state by sampling the parameters that are associated with \( V_1, V_2, V_R \) and \( V_L \). Finally, 5005 quantum optical states are obtained for the following comparison.

To test the generalization of DNN-QST, we directly adopt the model trained from the quantum states sampled according to \([7]\) with \( p = 0.99 \) in Section V-A and test it on the generated quantum optical states. The results of reconstructing those quantum optical states using noisy measurements are revealed in Fig. 12 As is shown, LRE and MLE achieve similar performance and DNN achieves better performance with increasing noise ratios for both the cube and the MUB measurements. Based on these results, DNN-QST has strong generalization in estimating quantum optical states and exhibits robustness against errors in the measurement operators.

### VI. Conclusion

Quantum state tomography is a significant task that has implications for many other quantum information processing tasks. Owing to the potential of ML to capture complex patterns from data, we proposed a general framework that utilizes neural networks to estimate quantum states. To demonstrate its efficiency, we applied it to three cases, including few measurement copies and incomplete measurements as well as noisy measurement operators. Numerical results demonstrate that DNN-QST has a great potential to achieve higher efficiency than LRE and MLE when estimating states with limited resources. Besides, DNN-QST is efficient in dealing with the noise in measurement operators and exhibits...
Fig. 8. The performance of 2-qubit states with different purities using noisy cube measurement. (a1) Infidelity vs purity for Gaussian noise with ratio 0.01; (a2) Infidelity vs purity for Gaussian noise with ratio 0.05; (b1) Infidelity vs purity for uniform noise with ratio 0.01; (b2) Infidelity vs purity for uniform noise with ratio 0.1.

Fig. 9. The performance of pure states and mixed state using noisy MUB measurement. (a1) Infidelity vs noise ratio (Gaussian noise) for pure states; (a2) Infidelity vs noise ratio (Gaussian noise) for mixed states with $p = 0.7$; (b1) Infidelity vs noise ratio (uniform noise) for pure states; (b2) Infidelity vs noise ratio (uniform noise) for mixed states with $p = 0.7$. 
favorable robustness when reconstructing quantum optical states with noisy measurements. Our future work will focus on adaptative quantum state tomography using ML and taking the advantages of ML to speed up tomography-based quantum experiments.

REFERENCES

[1] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control*. Cambridge University Press, 2009.

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[3] D. Dong and I. R. Petersen, “Quantum control theory and applications: a survey,” *IET Control Theory & Applications*, vol. 4, no. 12, pp. 2651–2671, 2010.

[4] N. Gisin and R. Thew, “Quantum communication,” *Nature Photonics*, vol. 1, no. 3, pp. 165–171, 2007.

[5] T. Opatrný, D.-G. Welsch, and W. Vogel, “Least-squares inversion for density-matrix reconstruction,” *Physical Review A*, vol. 56, no. 3, p. 1788, 1997.

[6] F. Huszár and N. M. Houlsby, “Adaptive Bayesian quantum tomography,” *Physical Review A*, vol. 85, no. 5, p. 052120, 2012.

[7] R. Blume-Kohout, “Optimal, reliable estimation of quantum states,” *New Journal of Physics*, vol. 12, no. 4, p. 043034, 2010.

[8] J. Řeháček, Z. Hadil, and M. Ježek, “Iterative algorithm for reconstruction of entangled states,” *Physical Review A*, vol. 63, no. 4, p. 040303, 2001.

[9] M. Ježek, J. Fiurášek, and Z. Hadil, “Quantum inference of states and processes,” *Physical Review A*, vol. 68, no. 1, p. 012305, 2003.

[10] B. Qi, Z. Hou, L. Li, D. Dong, G. Xiang, and G. Guo, “Quantum state tomography via linear regression estimation,” *Scientific Reports*, vol. 3, p. 3496, 2013.

[11] B. Qi, Z. Hou, Y. Wang, D. Dong, H.-S. Zhong, L. Li, G.-Y. Xiang, H. M. Wiseman, C.-F. Li, and G.-C. Guo, “Adaptive quantum state tomography via linear regression estimation: Theory and two-qubit experiment,” *npj Quantum Information*, vol. 3, no. 1, pp. 1–7, 2017.

[12] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[13] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[14] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, “Quantum machine learning,” *Nature*, vol. 549, no. 7671, p. 195, 2017.

[15] G. Carleo and M. Troyer, “Solving the quantum many-body problem with artificial neural networks,” *Science*, vol. 355, no. 6325, pp. 602–606, 2017.

[16] D. Lennon, H. Moon, L. Camenzind, L. Yu, D. Zumbühl, G. Briggs, M. Osborne, E. Laird, and N. Ares, “Efficiently measuring a quantum device using machine learning,” *npj Quantum Information*, vol. 5, no. 1, pp. 1–8, 2019.

[17] T. Fösel, P. Tighineanu, T. Weiss, and F. Marquardt, “Reinforcement learning with neural networks for quantum feedback,” *Physical Review X*, vol. 8, no. 3, p. 031084, 2018.

[18] C. Chen, D. Dong, H.-X. Li, J. Chu, and T.-J. Tarn, “Fidelity-based probabilistic Q-learning for control of quantum systems,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 920–933, 2013.
H. Ma, D. Dong, S. X. Ding, and C. Chen, “Curriculum-based deep reinforcement learning for quantum control,” arXiv preprint arXiv:2012.15427, 2020.

C.-J. Huang, H. Ma, Q. Yin, J.-F. Tang, D. Dong, C. Chen, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, “Realization of a quantum autoencoder for lossless compression of quantum data,” Physical Review A, vol. 102, no. 3, p. 032412, 2020.

H. Ma, C.-J. Huang, C. Chen, D. Dong, Y. Wang, R.-B. Wu, and G.-Y. Xiang, “On compression rate of quantum autoencoders: Control design, numerical and experimental realization,” arXiv preprint arXiv:2005.11149, 2020.

Q. Xu and S. Xu, “Neural network state estimation for full quantum state tomography,” arXiv preprint arXiv:1811.06654, 2018.

S. Lohani, B. Kirby, M. Brodsky, O. Danaci, and R. T. Glasser, “Machine learning assisted quantum state estimation,” Machine Learning: Science and Technology, vol. 1, no. 3, p. 035007, 2020.

O. Danaci, S. Lohani, B. T. Kirby, and R. T. Glasser, “Machine learning pipeline for quantum state estimation with incomplete measurements,” arXiv preprint arXiv:2012.03104, 2020.

S. Lohani, T. A. Searles, B. T. Kirby, and R. T. Glasser, “On the experimental feasibility of quantum state reconstruction via machine learning,” arXiv preprint arXiv:2012.09432, 2020.

S. Ahmed, C. S. Muñoz, F. Nori, and A. F. Kockum, “Quantum state tomography with conditional generative adversarial networks,” arXiv preprint arXiv:2008.03240, 2020.

S. Ahmed, C. S. Muñoz, F. Nori, and A. F. Kockum, “Classification and reconstruction of optical quantum states with deep neural networks,” arXiv preprint arXiv:2012.02185, 2020.

M. Neugebauer, L. Fischer, A. Jäger, S. Czischek, S. Jochim, M. Weitemüller, and M. Gätterer, “Neural-network quantum state tomography in a two-qubit experiment,” Physical Review A, vol. 102, no. 4, p. 042604, 2020.

A. M. Palmieri, E. Kovlakov, F. Bianchi, D. Yudin, S. Straupe, J. D. Biamonte, and S. Kulik, “Experimental neural network enhanced quantum tomography,” npj Quantum Information, vol. 6, no. 1, pp. 1–5, 2020.

Z. Hou, H.-S. Zhong, Y. Tian, D. Dong, B. Qi, L. Li, Y. Wang, F. Nori, G.-Y. Xiang, C.-F. Li, et al., “Full reconstruction of a 14-qubit state within four hours,” New Journal of Physics, vol. 18, no. 8, p. 083036, 2016.

D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, “Measurement of qubits,” Physical Review A, vol. 64, p. 052312, 2001.

Y.-x. Liu, L. Wei, and F. Nori, “Quantum tomography for solid-state qubits,” EPL (Europhysics Letters), vol. 67, no. 6, p. 874, 2004.

A. Chantasri, S. Pang, T. Chalermpusitarak, and A. N. Jordan, “Quantum state tomography with time-continuous measurements: reconstructions with resource limitations,” Quantum Studies: Mathematics and Foundations, pp. 1–25, 2019.

D. Gonçalves, C. Lavor, M. Gomes-Ruggiero, A. Cesário, R. Vianna, and T. Maciel, “Quantum state tomography with incomplete data: Maximum entropy and variational quantum tomography,” Physical Review A, vol. 87, no. 5, p. 052140, 2013.

Y. S. Teo, B. Stoklasa, B.-G. Englert, J. Řeháček, and Z. Hradil, “Incomplete quantum state estimation: A comprehensive study,” Physical Review A, vol. 85, no. 4, p. 042317, 2012.

T. O. Maciel, A. T. Cesário, and R. O. Vianna, “Variational quantum tomography with incomplete information by means of semidefinite programs,” International Journal of Modern Physics C, vol. 22, no. 12, pp. 1361–1372, 2011.

J. A. Smolin, J. M. Gambetta, and G. Smith, “Efficient method for computing the maximum-likelihood quantum state from measurements with additive Gaussian noise,” Physical Review Letters, vol. 108, no. 7, p. 070502, 2012.

G. Torlai and R. G. Melko, “Machine-learning quantum states in the nisq era,” Annual Review of Condensed Matter Physics, vol. 11, pp. 325–344, 2020.

F. Mezzadri, “How to generate random matrices from the classical compact groups,” arXiv preprint math-ph/0609050, 2006.

O. Danaci, S. Lohani, B. Kirby, and R. T. Glasser, “Machine learning pipeline for quantum state estimation with incomplete measurements,” Machine Learning: Science and Technology, 2021.

N. J. Higham, Analysis of the Cholesky decomposition of a semi-definite matrix. Oxford University Press, 1990.