



Neural networks for quantum state tomography with constrained measurements

Hailan Ma^{1,2} · Daoyi Dong¹ · Ian R. Petersen¹ · Chang-Jiang Huang³ · Guo-Yong Xiang³

Received: 5 April 2024 / Accepted: 20 August 2024 / Published online: 9 September 2024
© The Author(s) 2024

Abstract

Quantum state tomography (QST) aiming at reconstructing the density matrix of a quantum state plays an important role in various emerging quantum technologies. Recognizing the challenges posed by imperfect measurement data, we develop a unified neural network (NN)-based approach for QST under constrained measurement scenarios, including limited measurement copies, incomplete measurements, and noisy measurements. Through comprehensive comparison with other estimation methods, we demonstrate that our method improves the estimation accuracy in scenarios with limited measurement resources, showcasing notable robustness in noisy measurement settings. These findings highlight the capability of NNs to enhance QST with constrained measurements.

Keywords Quantum state tomography · Neural networks · Errors · Robustness

✉ Daoyi Dong
daoyi.dong@anu.edu.au

Hailan Ma
hailanma0413@gmail.com

Ian R. Petersen
i.r.petersen@gmail.com

Chang-Jiang Huang
hcj00@mail.ustc.edu.cn

Guo-Yong Xiang
gyxiang@ustc.edu.cn

¹ CIICADA Lab, School of Engineering, The Australian National University, Canberra, ACT 2601, Australia

² School of Engineering and Technology, University of New South Wales, Canberra, ACT 2600, Australia

³ Key Laboratory of Quantum Information, University of Science and Technology of China, Chinese Academy of Sciences, Hefei 230026, China

1 Introduction

Quantum state tomography (QST) plays a significant role in verifying and benchmarking quantum tasks, including quantum computation [1], quantum control [2] and quantum communication [3]. It involves reconstructing the state of a quantum system via quantum measurements [4]. To achieve this, least-squares inversion was proposed to solve 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 many 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].

The exponential scaling of parameters in a quantum state requires an exponentially increasing number of complete measurements [11], each of which requires a sufficient number of identical copies, posing a great challenge in practical applications. For example, measuring quantum states can be time-consuming and requires complex and costly experimental setups, which might limit the available measurements [12]. Within coherence time, one may obtain a complete but inaccurate data vector (a few copies assigned to each measurement operator) or an accurate but incomplete data vector (sufficient copies assigned to partial measurements that are experimentally easy to generate). Another persistent challenge is systematic errors in the inherent susceptibility of quantum measurements to noise, which increases the difficulty of fully characterizing quantum states [13]. These imperfect factors are collectively referred to as constrained measurements in this paper, which hinder the precise reconstruction of quantum states in practical applications.

Recently, machine learning (ML) has been utilized to address various quantum tasks [14–17]. ML's capability to handle complex data makes it suitable for extracting useful patterns from imperfect measured data for QST issues [18]. For example, neural networks (NNs) act as a useful tool to denoise the state-preparation-and-measurement errors [19]. Conditional generative adversarial networks have been introduced for the reconstruction of quantum optical states [20, 21], and convolutional neural networks (CNNs) have been applied to estimate quantum states from incomplete measurements [22, 23, 25]. Despite these advancements, the exploration of how NNs enhance QST with various constrained measurements, especially the advantage over traditional methods, has been largely unexplored.

In this context, we develop an NN-based QST method designed for the effective reconstruction of quantum states under measurement constraints. In particular, we design a straightforward yet expressive NN architecture composed of multiple layers of fully connected networks. These networks map measured statistics to intermediate real vectors which are then used to construct lower-triangular matrices, ultimately leading to the generation of density matrices. In contrast to other methods that rely on prior assumptions (e.g., the low-rank assumption in compressed sensing [26]), our approach offers a unified solution applicable to a variety of constrained measurement scenarios common in quantum applications. To benchmark the performance, we compare our

method with MLE, LRE, and a CNN-based QST method [22]. An improvement in average fidelity demonstrates the advantage of our approach in reconstructing quantum states with limited measurement resources. Moreover, our method has demonstrated favorable robustness in estimating quantum states when tomographic measurements suffer from noise. These findings highlight the potential of NNs to enhance QST beyond the traditional dependence on extensive measurements.

The rest of this paper is organized as follows. In Sect. 2, 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 Sect. 3. Section 4 investigates the performance of QST with noise. Concluding remarks are given in Sect. 5.

2 Deep neural networks-based quantum state tomography

In this section, a machine learning-aided quantum state tomography method (called DNN-QST) that takes advantage of deep neural networks to reconstruct quantum states is proposed. Quantum state generation and its representation are first presented, and then, quantum measurement settings are introduced. Finally, the DNN-QST method is provided in detail.

2.1 Quantum state generation and its representation

In this work, we first consider pure states. Let \mathbb{U}^d be the set of all d -dimensional unitary operators, i.e., any $U \in \mathbb{U}^d$ satisfies $UU^\dagger = U^\dagger U = \mathbb{I}_d$. Owing to its invariance under group multiplication (i.e., any region of \mathbb{U}^d carries the same weight in a group average) [27], the Haar metric is utilized as a probability measure on a compact group for generating random unitary matrix. As such, random pure states can be generated by performing random unitary transformations following the Haar measure [23, 28], which can be formulated as $|\psi\rangle_{\text{haar}} = U_{\text{haar}}|\psi_0\rangle$, where $|\psi_0\rangle$ is a fixed pure state. Then, we consider mixed states in the following form

$$\rho_p = p|\psi\rangle\langle\psi|_{\text{haar}} + (1 - p)\frac{\mathbb{I}_d}{d}, \tag{1}$$

where $p \in (0, 1)$ represents the ratio of pure element. The purity of the quantum state in Eq. (1) is closely related to the value of p , namely, $\text{Tr}(\rho_p^2) = p^2(1 - 1/d) + 1/d$. Let $\mathcal{N}(0, 1)_d$ represent random normal distributions of size $d \times d$ with zero mean and unity variance. Then, we consider random mixed states generated from the Ginibre ensembles using the Hilbert–Schmidt metric [24], given by

$$\rho_{\text{mixed}} = \frac{\rho_G \rho_G^\dagger}{\text{Tr}(\rho_G \rho_G^\dagger)}, \quad \rho_G = \mathcal{N}(0, 1)_d + i\mathcal{N}(0, 1)_d. \tag{2}$$

when only considering pure states, splitting $|\psi\rangle$ into real and imaginary parts is natural solution for parameterizing quantum states. Here, we try to propose a general QST method that can be applied to mixed states. To achieve a unified notation for pure states and mixed states, a density matrix ρ , 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 ρ_L , a physical density matrix can be obtained as

$$\rho = \frac{\rho_L \rho_L^\dagger}{\text{Tr}(\rho_L \rho_L^\dagger)}. \quad (3)$$

ρ in Eq. (3) satisfies the three conditions of Hermitian, positive semidefinite, and unit trace. In addition, according to the Cholesky decomposition [29], for any density matrix ρ , there exists a lower triangular matrix ρ_L that achieves $\rho_L \rho_L^\dagger = \rho$. Note that a tiny perturbation term (e.g., $\epsilon = 10^{-7}$) is usually added to the simulated pure states to avoid convergence issues using the Cholesky decomposition [29]

$$\rho = (1 - \epsilon)|\psi\rangle\langle\psi| + \frac{\epsilon}{d}\mathbb{I}_d. \quad (4)$$

As such, any density matrix ρ can be parameterized as a corresponding lower triangular matrix ρ_L , which has been widely used in previous work [20, 21, 30, 31]. By comparison, the Bloch representation requires additional efforts to pull a non-physical state into a physical state [32]. 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 splitting the real and imaginary parts.

2.2 Quantum measurement settings

Theoretically, the minimum number of measurements to fully characterize a quantum state with dimension d is d^2 [11]. For example, a linearly independent set of projectors from pairwise combinations of eigenstates of the Pauli operators can be constructed [11]. An improved estimate of 2-qubit quantum states can be obtained when tomography is performed using projections onto 36 tensor products of Pauli eigenstates [33]. In addition, mutually unbiased bases where all inner products between projectors of different bases are equal to $\frac{1}{d}$ have the potential to maximize information extraction per measurement [34]. Hence, the measurement operators play an important role in the estimation problem.

Let the eigenvectors of σ_x be the diagonal state $|D\rangle$ and the anti-diagonal state $|A\rangle$. Define the eigenvectors of σ_y as the left circular polarization state $|L\rangle$ and the right circular polarization state $|R\rangle$ and take the eigenvectors of σ_z as the horizontal state $|H\rangle$ and the vertical state $|V\rangle$. These six basis states form the 1-qubit Pauli measurement [33], $P_{\text{Pauli}} = \{|H\rangle, |V\rangle, |D\rangle, |A\rangle, |R\rangle, |L\rangle\}$. They are informationally complete for reconstructing 1-qubit quantum states.

In this work, we consider two types of measurement settings: (i) tensor products of Pauli matrices, which is also called the cube measurement [33], (ii) the mutually

Table 1 The measurement bases used in cube and MUB [34]

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, HL\rangle, VR\rangle, VL\rangle$	$ DR\rangle, DL\rangle, AR\rangle, AL\rangle$
$ DH\rangle, DV\rangle, AH\rangle, AV\rangle$	$\frac{1}{\sqrt{2}}(RL\rangle + i LR\rangle), \frac{1}{\sqrt{2}}(RL\rangle - i LR\rangle)$
$ DD\rangle, DA\rangle, AD\rangle, AA\rangle$	$\frac{1}{\sqrt{2}}(RR\rangle + i LL\rangle), \frac{1}{\sqrt{2}}(RR\rangle - i LL\rangle)$
$ DR\rangle, DL\rangle, AR\rangle, AL\rangle$	$\frac{1}{\sqrt{2}}(RV\rangle + i LH\rangle), \frac{1}{\sqrt{2}}(RV\rangle - i LH\rangle)$
$ RH\rangle, RV\rangle, LH\rangle, LV\rangle$	$\frac{1}{\sqrt{2}}(RH\rangle + i LV\rangle), \frac{1}{\sqrt{2}}(RH\rangle - i LV\rangle)$
$ RD\rangle, RA\rangle, LD\rangle, LA\rangle$	
$ RR\rangle, RL\rangle, LR\rangle, LL\rangle$	

unbiased bases (MUB) measurement [34]. For n -qubits, there are 6^n measurement basis operators involved in cube measurement and $(4^n + 2^n)$ involved in MUB measurement. The measurement basis states for cube and MUB are usually grouped into sets, each containing $d = 2^n$ orthogonal projectors. Typically, the 6^n measurement operators in the cube measurement are arranged into 3^n sets. Similarly, the MUB measurements are grouped into $(2^n + 1)$ sets. For the 2-qubit case, their measurement basis states are summarized in Table 1. Denoting each ideal measurement operator as M_i , the presence of noise may lead to the utilization of perturbed operators \tilde{M}_i . This disparity between M_i and \tilde{M}_i consequently impedes the reconstruction process.

2.3 DNN-QST

According to the universal approximation theorem [35], any continuous function on a compact subset of \mathbb{R}^n can be theoretically approximated by deep NNs with multiple layers of fully connected neural networks. Hence, multi-layer NNs can be constructed to approximate a function mapping from measured frequencies to physical density matrices. As such, QST is transformed into a regression problem, which can be solved in a supervised learning procedure. The input can be the observed frequencies, which depend on the choice of measurement operators. Recall the measurements are usually grouped into different sets (see Sect. 2.2 for detailed information), the dimension of the input vector is determined by the number of sets among the measurements. In particular, the dimension of the input feature can be calculated as dK , with K the number of measurement sets. In that case, the input vector can be denoted as $\mathbf{f} = [f_1, f_2, \dots, f_{dK}] \in \mathbb{R}^{dK}$. For example, for the complete measurement case, its input would be 6^n -dimensional feature vectors for the cube measurement and $2^n(2^n + 1)$ -dimensional feature vectors for the MUB measurement, respectively. For the incomplete case, the input dimension would decay. In general, any measurement setting that can be grouped according to certain rules can be applied to our method.

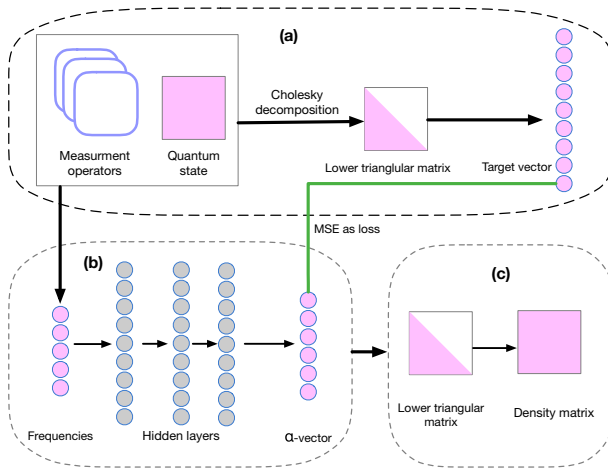


Fig. 1 Schematic of the DNN-QST approach. **a** Obtain the measured data and compute the Cholesky decomposition of the density matrices as the target vectors. **b** A multi-layer NN maps the frequencies to the α -vector. **c** Obtain the quantum states from the networks' output

For example, one can establish a measurement setting using random square root measurements [36].

Recall a lower triangular matrix ρ_L is correlated with a positive Hermitian matrix ρ in Eq. (3). When designing deep NNs to reconstruct a quantum state, a lower triangular matrix ρ_L can be used as an intermediate state. ρ_L is a complex matrix with real and positive diagonal entries and complex off-diagonal elements and thus can be further represented as a real vector with length d^2 (defined as the α -vector in this work). In particular, the positive diagonal entries are first ordered, and then the real and imaginary parts of off-diagonal elements are considered. As such, the output layer can be designed with d^2 neurons to represent 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 created through the generation of different density matrices, whose target vectors are calculated by the Cholesky decomposition. Then, measured frequencies are gathered through simulations of the measurement process using density matrices and measurement operators. Denote the training sample as $\{\mathbf{f}, \alpha\}$, where $\mathbf{f} \in \mathbb{R}^{dK}$ represents the measured frequencies on quantum state ρ and $\alpha \in \mathbb{R}^{d^2}$ represents the real vector obtained by performing Cholesky decomposition on ρ (equivalent to the ground truth of ρ). In (b), an architecture that includes input-hidden-output layers is utilized as a parameterized function to map a feature vector \mathbf{f} comprising of the measurement outcomes to the α -vector. The loss function for training the NNs adopts the mean squared error (MSE) between the reconstructed α -vector (from the output of the NNs) and the expected target vector (from the decomposition of the density matrix). Then, the networks are trained using a gradient descent algorithm with AdamOptimizer to minimize the MSE [37].

The detailed training procedures of DNN-QST are summarized in Algorithm 1. Practically, the entire training dataset is divided into mini-batches to be fed into the networks one by one. An epoch elapses when an entire dataset is passed forward and backward through the NNs exactly one time (line 10-17 in Algorithm 1). After the weights of the NNs have been learned, the α -vector is transformed into a physical density matrix in (c). Unlike the training process in (b), a procedure of $\alpha \rightarrow \rho_L \rightarrow \rho$ is 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 perspective, DNN-QST is efficient in reconstructing quantum states.

Algorithm 1 Procedures for DNN-QST.

```

1: Input: Training samples with density matrices  $\{\rho\}$ , batch size  $B$ , epoch size  $T$ 
2: Output: The approximated function  $\mathcal{G}(\xi) : \mathbf{f} \rightarrow \alpha$  with weights  $\xi$ 
3: for each sample in the training dataset  $j = 1, 2, \dots$  do
4:   Perform measurements on the quantum state  $\rho_j$  and collect the measurement outcomes into measured frequencies  $\mathbf{f}_j = [f_1, f_2, \dots, f_{dK}]$ 
5:   Calculate the Cholesky decomposition on  $\rho_j$  to obtain a lower-triangular matrix and a real vector  $\alpha_j$ 
6: end for
7: Randomly initialize weights of NN, i.e.,  $\xi$ 
8: for  $epoch = 1, 2, \dots, T$  do
9:   Randomly shuffle all the training samples
10:  while not all training samples have been visited do
11:    Sequentially select a mini-batch of  $B$  samples from all the training set
12:    for each sample in the mini-batch  $j = 1, 2, \dots, B$  do
13:      Forward the neural networks as  $\hat{\alpha}_j = \mathcal{G}(\mathbf{f}_j; \xi)$ 
14:    end for
15:    Backward by minimizing MSE, i.e.,  $(\sum_j |\hat{\alpha}_j - \alpha_j|^2)/B$ 
16:    Update  $\xi$  with gradient descent
17:  end while
18: end for

```

Remark 1 For the case of a few copies, we utilize ideal measurement operators $\{M_i\}$ to simulate the measurement process. Here, the number of copies are limited, and the obtained frequencies have large errors. For the case of incomplete measurements, we use sufficient copies to simulate the measurement process, but the total measurements are not informational complete. For the case of noisy measurements, we utilize the noisy operator $\{\tilde{M}_i\}$ to simulate the measurement process.

3 Results on QST with limited measurement resources

In this section, parameter settings are first introduced, and then QST using few measurement copies and incomplete measurements is investigated and analyzed to demonstrate the effectiveness of our approach.

Table 2 Comparison of computation efforts for four methods when reconstructing pure states using cube measurements

Methods	2 qubit			3 qubit		
	Params	Train	Infer (ms)	Params	Train	Infer (ms)
LRE	N/A	N/A	18.21	N/A	N/A	141.52
MLE	N/A	N/A	21.96	N/A	N/A	145.28
CNN	765 K	1 h	1.25	2.8 M	5 h	1.49
DNN	56 K	38 min	1.26	85 K	48 min	1.40

3.1 Parameter settings

For each case, a large number of samples are generated by randomly sampling the parameters and split for different purposes. 98,800 samples are used for training the parameters of DNN-QST. 1000 samples are used for testing the model. Hence, training data and testing data are generated from the same distributions in our work. In addition, those 1000 samples are also tested on LRE and MLE. The measured frequencies are obtained by simulating the measurement process with a finite number of copies. In this work, we introduce S to represent the average number of copies for each measurement, with the total number of copies being SK , where K represents the number of measurements. For the case of few copies, we set S to be different values, such as $\{10, 20, 20, 40, 60, 80, 100\}$ to investigate the performance under varying numbers of copies. For the cases of incomplete and noisy measurements, S is fixed as 100,000 to reduce the impact of limited measurement copies.

For simulations, the Pytorch framework is utilized to construct the NNs 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.

We also implement a CNN-based QST method [22], where the measured frequency vectors are first encoded as an image (e.g., 6×6 pixel for 2-qubit states with the cube measurements) and are then processed via conventional layers of kernel 2×2 , stride length of 1, 25 feature mapping and max-pooling layers of kernel 2×2 , followed by a fully connected layer. 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 operator [13]. The training of 98,800 samples is conducted on GPU with 2000 epochs and 256 batch sizes for CNN and our method. For inference time, we conduct 1000 times of reconstructing the density matrix on CPU and report the average time per inference in Table 2. MLE and LRE do not involve a training process and their computations are both accomplished on CPU. To demonstrate the efficiency of the DNN-QST method, the infidelity of two states is defined as $\bar{F} = 1 - F(\rho, \sigma)$, with the fidelity between the reconstructed state and the expected state defined as $F(\rho, \sigma) = |\text{Tr}(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}})|$.

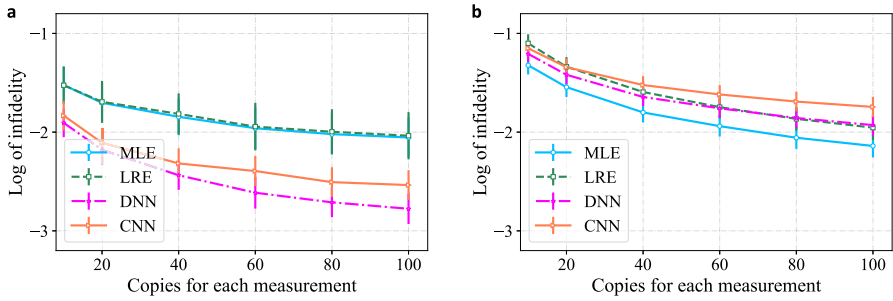


Fig. 2 Comparison of several methods on 2-qubit states using cube measurements. **a** Infidelity versus copies of each measurement for pure states. **b** Infidelity versus copies of each measurement for mixed states

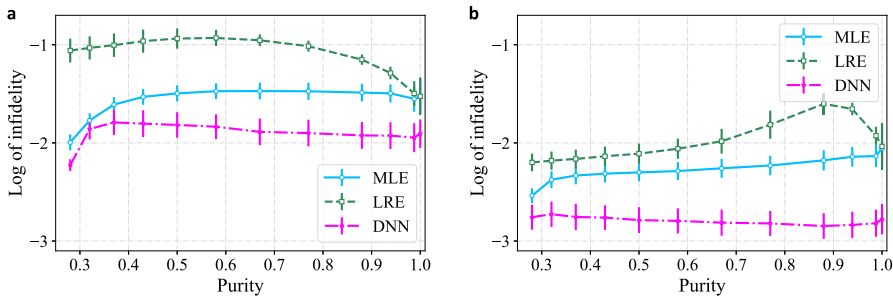


Fig. 3 The performance for 2-qubit states with different purities using few measurement copies under the cube measurement. **a** Infidelity versus purity for cube basis with $S = 10$. **b** Infidelity versus purity for cube basis with $S = 100$

3.2 Few copies

In real applications, taking one measurement on quantum systems may return a statistical value of probability, i.e., frequency. Considering that an individual outcome only provides limited information on the estimated state, an infinite number of measurements are required to determine a quantum state precisely. However, only finite measurements are available and performed on identically prepared states, with observed outcomes gathered as frequencies. A low number of measurement copies will cause the measured frequencies to deviate from the true probabilities.

We first compare the performance of different methods under different measurement copies, with the results of cube measurements presented in Fig. 2. The error bars in all of the figures represent half of the standard deviation. It is clear that DNN exhibit advantage over CNN for both random pure states and random mixed states, and the gaps increase with S . For pure states, our method achieves a superior performance to MLE and LRE. For mixed states, our method achieves a comparative performance to MLE. Then, we investigate the performance of mixed states with different purities. Considering the high computation of the CNN method and the similarity of pure states and mixed states in the form of Eq. (1), we focus on the comparison of our method against the two traditional methods. Figures 3 and 4 summarize the results of $S = 10$ and $S = 100$ for cube measurement and MUB measurement, respectively.

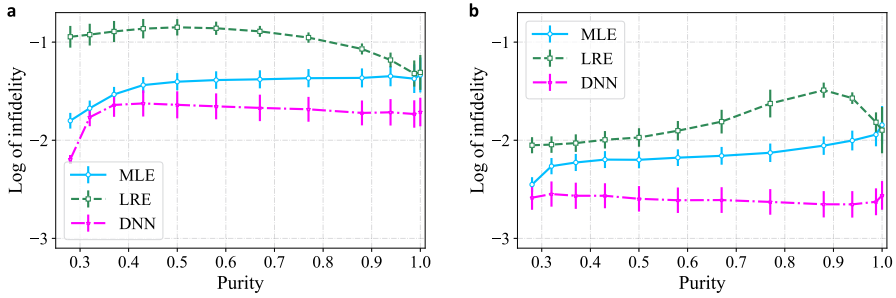
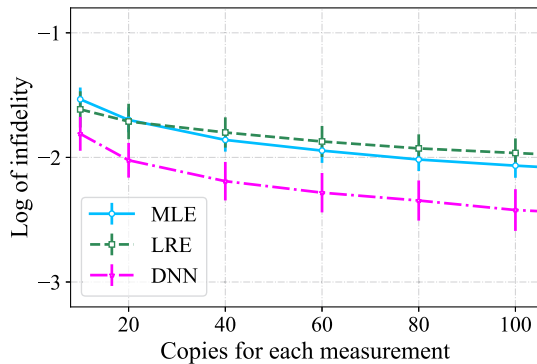


Fig. 4 The performance for 2-qubit states with different purities using few measurement copies under the MUB measurement. **a** Infidelity versus purity for MUB basis with $S = 10$. **b** Infidelity versus purity for MUB basis with $S = 100$

Fig. 5 Results for 3-qubit random pure states when the cube measurement has few copies



DNN achieves the best performance, 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 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 a tiny fluctuation for $S = 100$. One should note that the discrepancy of results for mixed states in Fig. 2 and Fig. 3, which arises from the different classes of mixed states, i.e., the total mixed states in Eq. (2) and mixed states with different purities in Eq. (1). Figure 5 summarizes the results for QST with few copies for 3-qubit pure states following the Haar metric, where DNN achieves the lowest log of infidelity, while MLE is better than LRE.

3.3 Incomplete measurements

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 cuben (mubn) includes the first n sets of projectors from the complete cube (MUB) measurement operators following the order in Table 1. Since the measurement

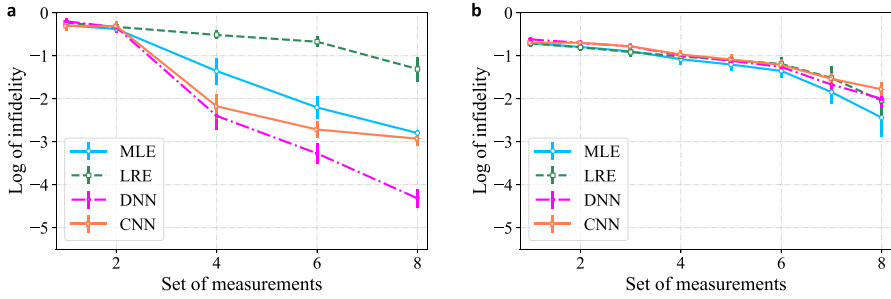


Fig. 6 The performance of 2-qubit pure states and mixed states using an incomplete cube measurement. **a** Infidelity versus sets of the cube measurement for pure states. **b** Infidelity versus sets of the cube measurement for mixed states

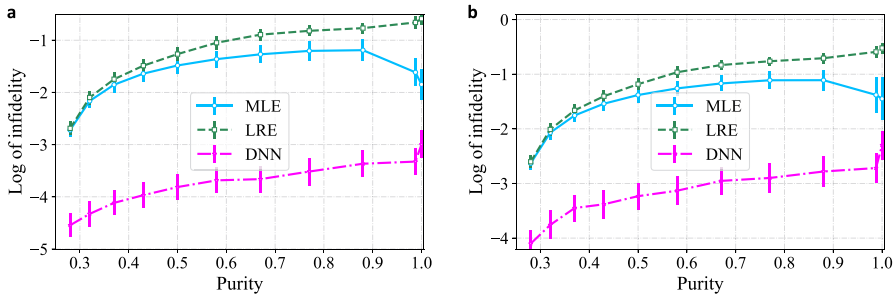
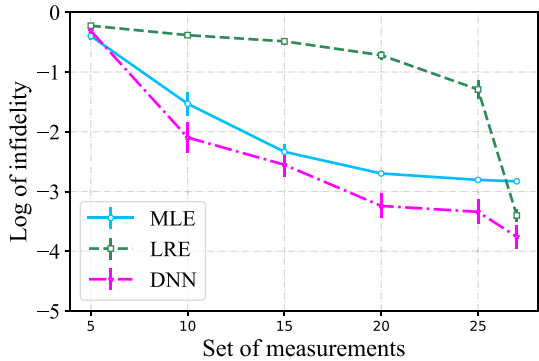


Fig. 7 The performance of 2-qubit states under incomplete measurements. **a** Infidelity versus purity for cube5. **b** Infidelity versus purity for mub3

projectors are usually arranged into several sets, the incompleteness of measurement operators is defined concerning the number of measurement sets.

We first investigate the comparison results of several methods based on the cube measurement. Figure 6 demonstrates that our method exhibits better performance than CNN in addressing incomplete measurements for pure states and achieve comparable performance to traditional methods for random mixed states. 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. Figure 7 demonstrates that 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. The simulation results for 3-qubit pure states are summarized in Fig. 8, where LRE fails to achieve good infidelity when the measurement set is close to the complete case, suggesting that LRE is sensitive to the completeness of measurements.

Fig. 8 Results for 3-qubit random pure states using different numbers of measurement sets for the cube measurement



4 Results on QST with noisy measurements

In real applications, measurements usually suffer from noise. For each measurement operator M_i , the existence of noise in measurements may cause it to be \tilde{M}_i . For one qubit, we consider the unitary rotation operator, given by [38–40]

$$U(\theta_1, \theta_2, \theta_3) = \begin{bmatrix} e^{i\theta_1} \cos(\theta_2) & -ie^{i\theta_3} \sin(\theta_2) \\ -ie^{-i\theta_3} \sin(\theta_2) & e^{-i\theta_1} \cos(\theta_2) \end{bmatrix}. \tag{5}$$

For multiple qubits, we consider the local errors of unitary rotation, defined as $U_e = U_{\Theta_1} \otimes U_{\Theta_2} \cdots \otimes U_{\Theta_n}$, where $\Theta_k = (\theta_1^k, \theta_2^k, \theta_3^k)^T$ denotes the noise vector for the k -th qubit [22]. Hence, we have the noisy measurement operator $\tilde{M}_i = U_e M_i U_e^\dagger$. In this case, the measured frequencies are observed using the noisy operator \tilde{M}_i rather than the ideal operator M_i . However, the reconstruction algorithm utilizes the measurement operator M_i for reconstructing the density matrices, since the noise parameters are not available to the reconstruction algorithms.

In particular, we consider two types of unitary noise: (i) noise parameters are drawn from a Gaussian distribution (denoted as $\mathcal{N}(\mu, \sigma)$, with μ and σ representing the mean value and the standard deviation value, respectively); (ii) noise parameters are drawn from a uniform distribution (denoted as $\mathcal{U}(\mu_1, \mu_2)$, with μ_1 and μ_2 representing the lower bound and the upper bound, respectively). For simplicity, a noise ratio ξ is introduced to control the sampling of three rotation angles. The noise parameters for the first case are sampled as $\theta_1 \sim \mathcal{N}(0, \pi\xi)$, $\theta_2 \sim \mathcal{N}(0, 2\pi\xi)$, and $\theta_3 \sim \mathcal{N}(0, 2\pi\xi)$ [22]. For the uniform case, the noise parameters are sampled as $\theta_1 \sim \mathcal{U}(0, 2\pi\xi)$, $\theta_2 \sim \mathcal{U}(0, 0.5\pi\xi)$, and $\theta_3 \sim \mathcal{U}(0, 2\pi\xi)$. Here, we set $\theta_2 \in [0, 0.5\pi\xi]$ owing to the symmetry of θ_1 and θ_3 in Eq. (5).

The comparison results when the cube measurement suffers from Gaussian noise and uniform noise are summarized in Figs. 9 and 10, respectively. Our method outperforms CNN for both pure and mixed states. The advantage of our method over the two traditional methods are obvious for pure states or mixed states when measurements suffer from uniform noise. Then, we focus on uniform noise and investigate the performance of states with different purities. Figure 11 reveals that DNN ranks first, followed by MLE. The infidelity of DNN and MLE increases with purity, while the

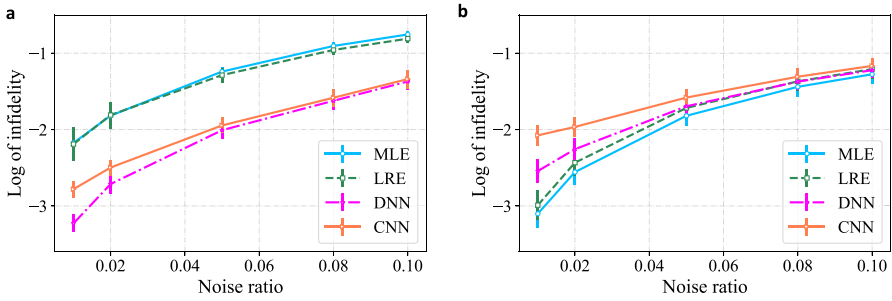


Fig. 9 Comparison of several methods when the cube measurement suffers from Gaussian noise. **a** Infidelity versus noise ratio for pure states. **b** Infidelity versus noise ratio for mixed states

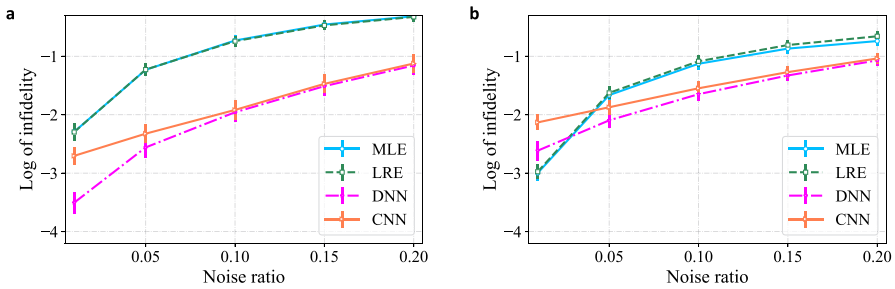


Fig. 10 Comparison of several methods when the cube measurement suffers from uniform noise. **a** Infidelity versus noise ratio for pure states. **b** Infidelity versus noise ratio for mixed states

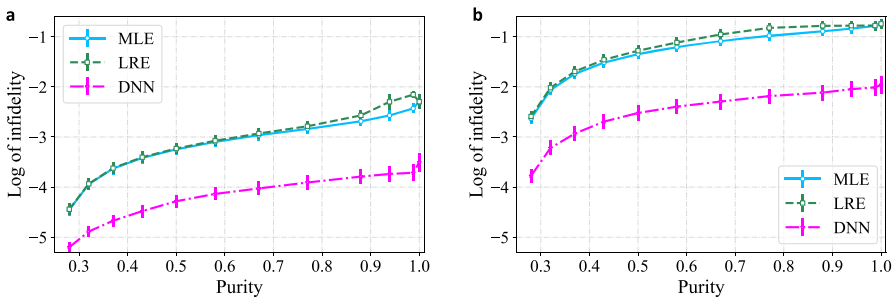
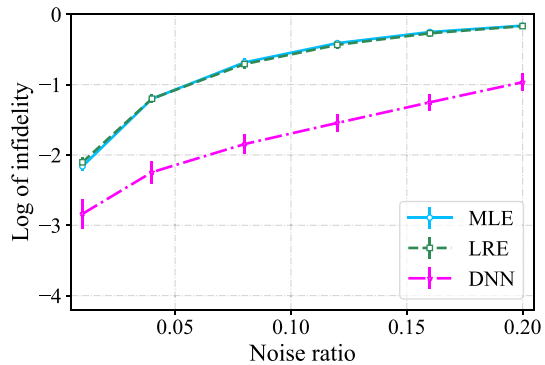


Fig. 11 The performance of 2-qubit states with different purities using noisy cube measurement. **a** Infidelity versus purity for uniform noise with ratio 0.01. **b** Infidelity versus purity for uniform noise with ratio 0.1

infidelity of LRE first increases slowly and then drops greatly with purity. The results of QST with noise for 3-qubit pure states are summarized in Fig. reffig:3qubitnoise, where LRE and MLE achieve similar performance when the cube measurement suffers from uniform noise. Compared with the two methods, DNN exhibits a superior performance under different noise ratios.

Fig. 12 Results for 3-qubit random pure states when the cube measurement suffers from uniform noise



5 Discussion and conclusion

QST is a significant task that has implications for many other quantum information processing tasks. Recognizing the challenges posed by imperfect measurement settings, we developed a general framework that leverages NNs to estimate quantum states with constrained measurements. To demonstrate its efficiency, we applied it to three cases, including a few measurement copies and incomplete measurements as well as noisy measurement operators. Numerical results demonstrate that our approach has a great potential to achieve higher efficiency than when estimating states with limited resources. Besides, our method exhibits favorable robustness when reconstructing quantum optical states with noisy measurements. If the states being tested differ significantly from those used to train the DNN, it is anticipated that the accuracy of these testing samples will degrade. To mitigate this, transfer learning techniques can be employed to endow the NNs with improved generalization capabilities. In our future work, we will incorporate the noise distribution of the state preparation and measurement into the design of NN architecture to achieve improved estimation of quantum states under different noises. In addition, we may explore the use of ML to search for optimal measurement settings, aiming to achieve improved estimation of quantum states.

Acknowledgements We thank Zhi-Bo Hou for useful discussions. This work was supported by the Australian Research Council's Future Fellowship funding scheme under Project FT220100656, the Australian Research Council's Discovery Projects funding scheme DP210101938, and the U.S. Office of Naval Research Global under Grant N62909-19-1-2129.

Author contributions H. M., D. D., and I.R. P. developed the method. C.-J. H. and G.-Y. X contributed to the part of noisy measurement. H. M. performed the simulation with the help from D. D. and C.-J. H. H. M. and D. D. drafted the manuscript with revisions from all the other authors.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Data availability The data generated in this study have been deposited in the Figshare database, which can be accessed at <https://doi.org/10.6084/m9.figshare.24311704>.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Nielsen, M.A., Chuang, I.L.: Quantum computation and quantum information. Cambridge University Press, Cambridge (2010)
2. Dong, D., Petersen, I.R.: Quantum control theory and applications: a survey. *IET Control Theory Appl.* **4**, 2651–2671 (2010)
3. Gisin, N., Thew, R.: Quantum communication. *Nat. Photonics* **1**, 165–171 (2007)
4. Dong, D., Petersen, I.R.: Quantum estimation, control and learning: opportunities and challenges. *Annu. Rev. Control.* **54**, 243–251 (2022)
5. Opatrný, T., Welsch, D.-G., Vogel, W.: Least-squares inversion for density-matrix reconstruction. *Phys. Rev. A* **56**, 1788 (1997)
6. Huszár, F., Houlby, N.M.: Adaptive Bayesian quantum tomography. *Phys. Rev. A* **85**, 052120 (2012)
7. Blume-Kohout, R.: Optimal, reliable estimation of quantum states. *New J. Phys.* **12**, 043034 (2010)
8. Řeháček, J., Hradil, Z., Ježek, M.: Iterative algorithm for reconstruction of entangled states. *Phys. Rev. A* **63**, 040303 (2001)
9. Ježek, M., Fiurášek, J., Hradil, Z.: Quantum inference of states and processes. *Phys. Rev. A* **68**, 012305 (2003)
10. Qi, B., et al.: Quantum state tomography via linear regression estimation. *Sci. Rep.* **3**, 3496 (2013)
11. James, D.F.V., Kwiat, P.G., Munro, W.J., White, A.G.: Phys. Rev. A. Measurement of qubits **64**, 052312 (2001)
12. Chantasri, A., Pang, S., Chalermphisitarak, T., Jordan, A.N.: Quantum state tomography with time-continuous measurements: reconstruction with resource limitations. *Quantum Stud.: Math. Found.* **7**, 23–47 (2019)
13. Smolin, J.A., Gambetta, J.M., Smith, G.: Efficient method for computing the maximum-likelihood quantum state from measurements with additive Gaussian noise. *Phys. Rev. Lett.* **108**, 070502 (2012)
14. Carleo, G., Troyer, M.: Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017)
15. Lennon, D., et al.: Efficiently measuring a quantum device using machine learning. *npj Quantum Inf.* **5**, 79 (2019)
16. Fösel, T., Tighineanu, P., Weiss, T., Marquardt, F.: Reinforcement learning with neural networks for quantum feedback. *Phys. Rev. X* **8**, 031084 (2018)
17. Huang, C.-J., et al.: Realization of a quantum autoencoder for lossless compression of quantum data. *Phys. Rev. A* **102**, 032412 (2020)
18. Biamonte, J., et al.: Quantum machine learning. *Nature* **549**, 195 (2017)
19. Palmieri, A.M., et al.: Experimental neural network enhanced quantum tomography. *npj Quantum Inf.* **6**, 20 (2020)
20. Ahmed, S., Muñoz, C.S., Nori, F., Kockum, A.F.: Quantum state tomography with conditional generative adversarial networks. *Phys. Rev. Lett.* **127**, 140502 (2021)
21. Ahmed, S., Muñoz, C.S., Nori, F., Kockum, A.F.: Classification and reconstruction of optical quantum states with deep neural networks. *Phys. Rev. Res.* **3**, 033278 (2021)
22. Lohani, S., Kirby, B., Brodsky, M., Danaci, O., Glasser, R.T.: Machine learning assisted quantum state estimation. *Mach. Learn.: Sci. Technol.* **1**, 035007 (2020)

23. Danaci, O., Lohani, S., Kirby, B., Glasser, R.T.: Machine learning pipeline for quantum state estimation with incomplete measurements. *Mach. Learn.: Sci. Technol.* **2**, 035014 (2021)
24. Al Osipov, V., Sommers, H.-J., Życzkowski, K.: Random Bures mixed states and the distribution of their purity. *J. Phys. A: Math. Theor.* **43**(5), 055302 (2010)
25. Lohani, S., Searles, T.A., Kirby, B.T., Glasser, R.T.: On the experimental feasibility of quantum state reconstruction via machine learning. *IEEE Trans. Quantum Eng.* **2**, 2103410 (2021)
26. Gross, D., Liu, Y.-K., Flammia, S.T., Becker, S., Eisert, J.: Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **105**, 150401 (2010)
27. Mezzadri, F.: How to generate random matrices from the classical compact groups. arXiv preprint [arXiv:math-ph/0609050](https://arxiv.org/abs/math-ph/0609050) (2006)
28. Qi, B., et al.: Adaptive quantum state tomography via linear regression estimation: theory and two-qubit experiment. *npj Quantum Inf* **3**, 19 (2017)
29. Higham, N.J.: Analysis of the Cholesky decomposition of a semi-definite matrix. Oxford University Press, Oxford (1990)
30. Ma, H., Dong, D., Petersen, I.R.: On how neural networks enhance quantum state tomography with limited resources (2021)
31. Wang, Y., Cheng, S., Li, L., Chen, J.: Ultrafast quantum state tomography with feed-forward neural networks. arXiv preprint [arXiv:2207.05341](https://arxiv.org/abs/2207.05341) (2022)
32. Koutný, D., Motka, L., Hradil, Z., Řeháček, J., Sánchez-Soto, L.L.: Neural-network quantum state tomography. *Phys. Rev. A* **106**, 012409 (2022)
33. De Burgh, M.D., Langford, N.K., Doherty, A.C., Gilchrist, A.: Choice of measurement sets in qubit tomography. *Phys. Rev. A* **78**, 052122 (2008)
34. Adamson, R., Steinberg, A.M.: Improving quantum state estimation with mutually unbiased bases. *Phys. Rev. Lett.* **105**, 030406 (2010)
35. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**, 436–444 (2015)
36. Motka, L., Pař, M., Řeháček, J., Hradil, Z., Sánchez-Soto, L.: Efficient tomography with unknown detectors. *Quantum Sci. Technol.* **2**, 035003 (2017)
37. Kingma, D. P., & Adam, B. J.: A method for stochastic optimization (2015)
38. Mahler, D.H., et al.: Adaptive quantum state tomography improves accuracy quadratically. *Phys. Rev. Lett.* **111**, 183601 (2013)
39. Chapman, R.J., Ferrie, C., Peruzzo, A.: Experimental demonstration of self-guided quantum tomography. *Phys. Rev. Lett.* **117**, 040402 (2016)
40. Hou, Z., et al.: Experimental realization of self-guided quantum process tomography. *Phys. Rev. A* **101**, 022317 (2020)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.