Robust Ground State Tomography with Neural Networks

Syed Muhammad Kazim, Junaid ur Rehman, and Hyundong Shin
Department of Electronics and Information Convergence Engineering, Kyung Hee University, Yongin-si, 17104 Korea
Email: hshin@khu.ac.kr

***Abstract*—In this paper, we present a neural network quantum state tomography scheme, which offers improvement in generalization to unseen data over conventional methods. Quantum state tomography (QST) is a resource intensive task and requires prohibitively large processing for even moderately sized quantum systems. Here, we perform the tomography of a 4-qubit 2-local Hamiltonian with true and estimated expectations with respect to a small set of observables, which is sufficient to achieve high fidelity. This method is scalable to larger states and Hamiltonians with arbitrary structures.**

I. INTRODUCTION

The ability to characterize quantum states allows us to benchmark the efficacy of quantum devices. Therefore it forms one of the fundamentals of quantum research with various application including, but not limited to, quantum information, communication and computing. This field of research is formally known as quantum state tomography (QST) [1], [2]. At the heart of QST, we perform measurements on the copies of an unknown and arbitrary quantum state and use the information gained from the outcomes to estimate the state. Given an infinite copies of a unique quantum state, QST is trivial. However in the real world, there is a limitation on the number of resources. As the number of qubits increase, the dimension of the quantum state grows exponentially. Even for modestly high dimensions, post-processing becomes intractable by most methods.

More concretely, QST methods generally either perform a fixed number of measurements on a state ρ given a static measurement configuration and post-process the information gained for estimation [3], [4] or apply some adaptive algorithm which does the same but changes the measurement setting periodically. The latter method is said to be adaptive [5]–[7]. Depending on the statistical scheme, the estimate $\hat{\rho}$ can be a good or bad approximate of ρ. However, as the number of measurements increase, generally the accuracy of the estimate increase as well because the relative frequency approaches true frequency in probability. So one aim of QST is to find a method that improves the accuracy of the estimate for a finite amount of resources.

One method that reduces the required resources is QST via reduced density matrices (RDMs). RDMs are accurate and powerful since they utilize local measurements. RDMs have been used for QST of ground state of local Hamiltonians. A many-body Hamiltonian is said to be k-local if $H = \sum_i Hi^{(k)}$, where each term H_i^k acts non-trivially on at most k particles. Only a polynomial number of parameters are required to completely characterize such a system. Since, generally, the ground states can encode the information of the entire system, only k-local measurements are required for QST [8], [9].

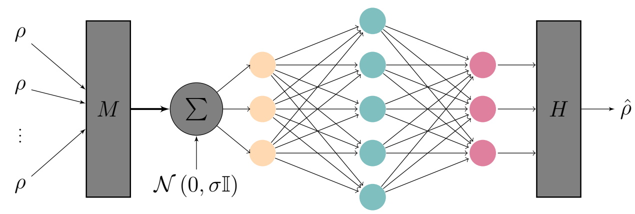


Fig. 1. Complete process of neural network quantum state tomography. An esnemble of ground states ρ are measured using a set of defined observables. Noise is added to each element in this set. Neural network is trained to predict the coefficients of the Hamiltonian H given noisy measurements as input. $\hat{\rho}$ is an estimate of the true ground state ρ.

QST via machine learning has burgeoned in the past few years. Currently extensive research on exploring popular deep learning architectures for QST is ongoing. This includes simple fully connected (FC) neural networks, recurrent neural networks (RBM and LSTMs), and more recently, generative adversarial networks [10], [11]. Neural networks (NNs) have been used for regression of data with high-dimensional features. The task of state reconstruction can be reformulated as a regression problem to make use of NNs to reconstruct ground states of local Hamiltonians.

In existing literature, FC NNs have been used for local-measurement-based QST of multiple qubits. In this paper, we propose improvements which help the model generalize better to noisy data for 4-qubit ground states of 2-local Hamiltonians [12]. However, our method is general and can be extended for any Hamiltonian. We further investigate how this method scales with respect to the main resource, i.e., copies of the quantum state.

This paper is structured as follows. We provide details of state preparation, model architecture, and overall scheme in section II. In section III, we apply our understanding to simulate QST via NN of 4-qubit ground states of 2-local Hamiltonians, provide results, and discuss their significance. We conclude in section IV.

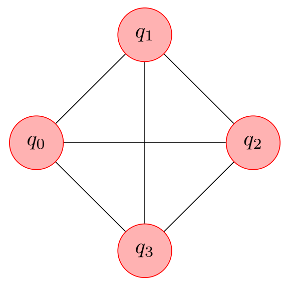


Fig. 2. configuration of the 4-qubits states. Each q_i represents a qubit and each qubit interacts with the other.

II. Methodology

A. Problem Description

The 4-qubit 2-local Hamiltonian H is defined as

$$H = \sum_{i=1}^{4} \sum_{1 \leq k \leq 4} w_k^i \sigma_k^i + \sum_{1 \leq i \leq j \leq 4} \sum_{1 \leq n,m \leq 3} J_{nm}^{ij} \sigma_n^i \otimes \sigma_m^j \quad (1)$$

where $\sigma_1, \sigma_2, \sigma_3$ and σ_4 are the Pauli matrices X, Y and Z, and the identity matrix respectively. We denote the set of Hamiltonian coefficients as $h = \{w_k^i, J_{nm}^{ij}\}$. The ground state configuration can be visualized in Fig. 2.

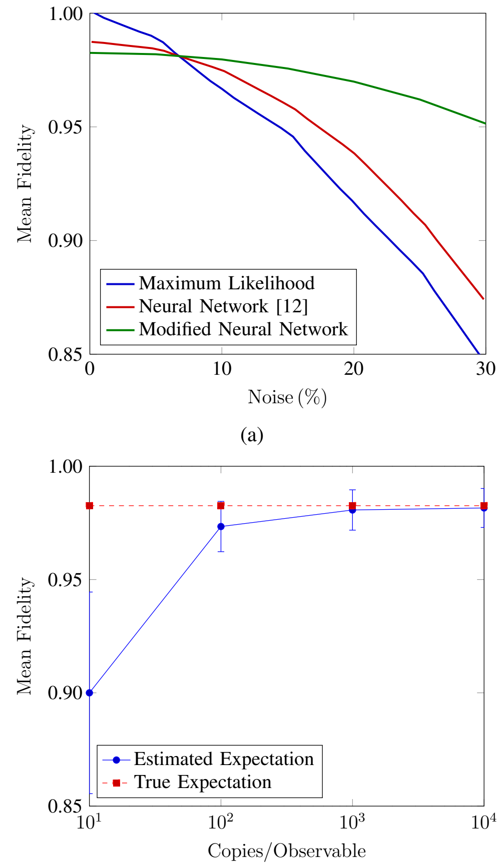
The basis set for H is $\boldsymbol{B} = \{\sigma_m \otimes \sigma_n : n + m \neq 8\}$. We further denote ρ to be the ground state of H, i.e., the density matrix of the eigenvector corresponding to the smallest eigenvalue of H. We define the set of true expectations of the ground state based on the local observables in the set B as $\boldsymbol{M} = \{s_{m,n}^{ij} \colon s_{m,n}^{ij} = \text{tr}\,(\text{tr}_{ij}\,\rho B_{m,n}), B_{m,n} \in \boldsymbol{B}, 1 \leq i < j \leq 4, 1 \leq n,m \leq 4\}$, where $\text{tr}_{ij}\,\rho$ is the partial trace of ρ with respect to the remaining indices.

Given the expectation values (true values or estimations) of the observable on the ground state, we want to map $\boldsymbol{M} \rightarrow \hat{h}$, where $\hat{h}$ is an estimate of h. Note that since our goal here is to reconstruct ρ and not H, our estimate $\hat{h}$ need not be very accurate. However, the ground state $\hat{\rho}$ of $\hat{H}$ with respect to $\hat{h}$ must be a good estimate of ρ.

B. State Preparation

In the section, we describe the method for generating data required to train the FC NN. For our NN, our input is the expectation with respect to the set $\boldsymbol{B}$. The number of input neurons equals $|\boldsymbol{M}|$. Similarly, since we estimate h, the number of output neurons equals $|h|$.

To generate a single point of the training data, we sample the multivariate normal distribution once to get coefficients h of the Hamiltonian. For each experiment, we generate 5×10^4 random Hamiltonian coefficients for training and 5×10^3 for testing. We further randomly choose 20% of the training data, and allocate it for validation. When we train our NN to find the true expectations, our target is $\boldsymbol{M}$ corresponding to each h. Moreover, since we want to analyze the performance of our NN with respect to the number of copies of ρ, N, we create further target sets where each element of the target set is the relative frequency when N copies are measured in the observable $B_{m,n}$.



(b)

Fig. 3. Neural network (NN) quantum state tomography with true and observed expectations averaged over 5×10^3 quantum states. (a) The mean fidelity of the maximum likeihood estimation and NN [12] is compared to our modified network with better generalization. As the percentage of noise is increased in the measured data, the advantage of our method becomes more pronounced. (b) The mean fidelity of our method as the number of copies of ρ used to calculate the expectations in the dataset is increased. The fidelity quickly saturates to the maximum achieved fidelity with true expectations.

C. Fully Connected Neural Network

Now that we know what our inputs and targets are, in this section we will discuss the details of our NN architecture. For the 4-qubit 2-local Hamiltonian given by 1, $|\boldsymbol{M}| = |h| = 66$ which is the number of input and output neurons of our network. We use a fully connected feedforwad network i.e., all neurons in one layer are connected to each neuron is layer before and after it. We use three hidden layers of sizes 300, 1500 and 300 in that order. To introduce non-linearity, we use a rectified linear unit (ReLU) at the out put of each layer except for the last. Moreover, we use mean square error as our loss function and Adam as our optimizer. The learning rate is set to 10^{-3}. To improve the generalization capability of our model, we found that weight decay of 10^{-4}, a dropout of 5×10^{-3} after each layer and a batch size of 500 to be optimal after hyperparameter optimization.

III. Numerical Results

In this section, we demonstrate the performance of our NN detailed previously. We report the fidelity, F, of our estimations $\hat{\rho}$ of the ground states ρ of the numerically generated 4-qubit 2-local Hamiltonians. We define the fidelity as

$$F(\rho, \hat{\rho}) = \sqrt{\mathrm{tr}\left(\sqrt{\rho}\hat{\rho}\sqrt{\rho}\right)}.$$

Fidelity captures the idea of closeness of one state to another. $F(\rho, \hat{\rho}) = 1$ only when $\rho = \hat{\rho}$.

Fig. 3(a) compares the performance of maximum likelihood estimation (MLE), neural network [] and our method when noise in injected into the dataset $\boldsymbol{M}$. For each measurement in the dataset, a scaled random sample from $\mathcal{N}(0, 1)$ is added. The percentage scaling termed noise forms the abscissa of the figure. This figure demonstrate the robustness of our more generalized NN compared to regular NN and conventional MLE estimates. The difference is more pronounced as the noise increases.

In Fig. 3(b), we measure multiple copies of ρ to estimate relative frequencies. These relative frequencies are then used to calculate expectations with respect to each observable. When we train our NN on this new dataset, the fidelity increases as the number of copies measured increase (blue). The vertical bars on the blue curve represent the variance of the fidelity the samples. The plot is averaged over 5×10^3 samples of the test set. For comparison, the red dashed line representing the performance of the network with the idea dataset $\boldsymbol{M}$ is plotted.

IV. Conclusions

In this work, we formulated QST as a regression problem so that we can use neural network to reconstrcut quantum states. We provide details of the dataset, architecture and optimized hyperparameters required to reproduce results. Moreover, we have demonstrated that our model is more robust against noise. We further show that our model can estimate states to high accuracy with meager quantum resources. This scheme can be used for tomography of any Hamiltonian.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1A2C2007037) and by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP$-2021-0-02046$) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

References

[1] R. Schmied, “Quantum state tomography of a single qubit: comparison of methods,” *J. Mod. Opt.*, vol. 63, no. 18, pp. 1744–1758, Feb. 2016.

[2] W. Munro, D. James, A. White, and P. Kwiat, “Measurement of qubits,” *Phys. Rev. A*, vol. 64, no. 030302, Oct. 2001.

[3] K. Banaszek, G. D'ariano, M. Paris, and M. Sacchi, “Maximum-likelihood estimation of the density matrix,” *Phys. Rev. A*, vol. 61, no. 1, p. 010304, Dec. 1999.

[4] R. Adamson and A. M. Steinberg, “Improving quantum state estimation with mutually unbiased bases,” *Phys. Rev. Lett.*, vol. 105, no. 3, p. 030406, Jul. 2010.

[5] S. S. Straupe, “Adaptive quantum tomography,” *JETP Lett.*, vol. 104, no. 7, pp. 510–522, Sep. 2016.

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

[7] S. M. Kazim, A. Farooq, J. ur Rehman, and H. Shin, “Adaptive quantum state tomography with iterative particle filtering,” *Quantum Inf. Process*, vol. 20, no. 10, pp. 1–19, Oct. 2021.

[8] J. Chen, Z. Ji, B. Zeng, and D. Zhou, “From ground states to local Hamiltonians,” *Phys. Rev. A*, vol. 86, no. 2, p. 022339, Aug. 2012.

[9] J. Chen, H. Dawkins, Z. Ji, N. Johnston, D. Kribs, F. Shultz, and B. Zeng, “Uniqueness of quantum states compatible with given measurement results,” *Phys. Rev. A*, vol. 88, no. 1, p. 012109, Jul. 2013.

[10] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, “Reconstructing quantum states with generative models,” *Nat. Mach. Intell.*, vol. 1, no. 3, pp. 155–161, Mar. 2019.

[11] S. Ahmed, C. S. Muñoz, F. Nori, and A. F. Kockum, “Quantum state tomography with conditional generative adversarial networks,” *Phys. Rev. Lett.*, vol. 127, no. 14, p. 140502, Sep. 2021.

[12] T. Xin, S. Lu, N. Cao, G. Anikeeva, D. Lu, J. Li, G. Long, and B. Zeng, “Local-measurement-based quantum state tomography via neural networks,” *npj Quantum Inform.*, vol. 5, no. 1, pp. 1–8, Nov. 2019.