Neural Network Tomography for Noisy Quantum States

Syed Muhammad Kazim*, Junaid ur Rehman*, Youngmin Jeong†, Kyesan Lee*, and Hyundong Shin*
*Department of Electronics and Information Convergence Engineering, Kyung Hee University, Yongin-si, 17104 Korea
†Networks Business, Samsung Electronics Co., Ltd, Yeongtong-gu, Suwon-si, Gyeonggi-do 16677, Korea
Email: hshin@khu.ac.kr

***Abstract*—We present convolutional neural network (CNN) as a solution for simultaneous quantum state tomography and noise mitigation of high-dimensional qudits. Other approaches including maximum likelihood estimation and particle filtering have a propensity of converging to incorrect quantum states due to state preparation errors and decoherence. We provide details of the model architecture, and report the mean squared error and fidelity of our estimates. In this manner, we demonstrate the usefulness of CNNs for real quantum devices that are noisy.**

I. Introduction

Characterization of quantum states is an essential task required to benchmark the performance of quantum devices. It has important implications in research areas such as quantum information, computing and communication [1]–[3]. Given measurements on an ensemble of identically prepared copies of a quantum system, quantum state tomography (QST) estimates the density matrix approximating the quantum system [4], [5]. Other popular branches of tomography include characterization of quantum logic gates through Gate Set tomography and and estimation of quantum channels through quantum process tomography [6], [7].

A common approach of QST methods is to first perform a series of measurements in different measurement settings on a state ρ, and then use any post processing scheme that can estimate ρ. 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 [8], [9]. So the goal of QST methods is to introduce a method that improves the accuracy of the estimate for a finite amount of resources.

A valid qubit is described by a density matrix which is positive semi-definite, Hermitian with unit trace. Given $N \gg 1$ measurements, multiple techniques are used in QST for estimation such as linear inversion, maximum likelihood estimation, Bayesian inference, and more recently, neural networks (NN). The first three methods converge on the state in the optimal setting. However, in the presence of state preparation errors and decoherence, the desired quantum state ρ changes, and our heuristics converge to another state η. However, neural networks can be trained to be perform both state and channel tomography by using one model, that is, given η, we can estimate ρ.

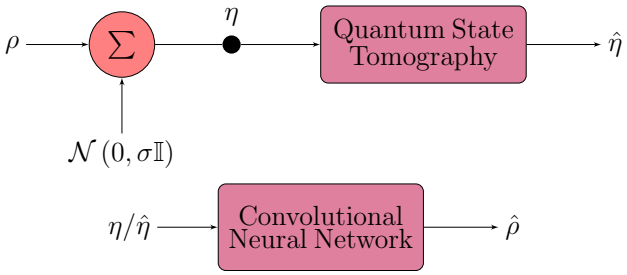

Fig. 1. Process used to generate $\hat{\eta}$, where $\hat{\eta}$ is an estimate of η which in turn is a noisy version of ρ. We implement CNN based tomography by taking η and $\hat{\eta}$ as our input to simultaneously mitigate noise and perform tomography to estimate ρ by reconstructing $\hat{\rho}$.

In existing literature, neural networks have been used for quantum state tomography and noise mitigation. In this paper, we utilize convolutional nueral network (CNN) to perform combined state and channel tomography. Moreover, we report our model's effectiveness with respect to state preparation errors and statistical noise.

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 quantum tomography of 8-dimensional qudits, provide results, and discuss their significance. We conclude in section IV.

II. Methodology

A d-dimensional quantum state ρ can be represented by its Bloch vector $\boldsymbol{r}$ and vector of Pauli words $\boldsymbol{\Omega}$ [5]

$$\rho = \frac{1}{d}\left(\mathbb{I} + \boldsymbol{r}\cdot\boldsymbol{\Omega}\right), \tag{1}$$

where $\mathbb{I}$ is a $d \times d$ identity matrix, $\boldsymbol{r} = (r_1, r_2, \cdots, r_{d^2-1}) \in \mathbb{R}^3$, and $\boldsymbol{\Omega} = (\Omega_1, \Omega_2, \cdots, \Omega_{d^2-1})$ is a vector of Pauli words. For ρ to be valid, it must be a positive semi-definite Hermitian with unit trace.

We generate mixed states ρ from the Ginibre ensemble. We assume that state preparation errors to be Gaussian in nature. Therefore, for each ρ, we calculate

$$\eta^* = \rho + \mathcal{N}(0, \sigma\mathbb{I}) + \iota\mathcal{N}(0, \sigma\mathbb{I}),$$

where $\mathcal{N}(\mu, \Sigma)$ is a mulitvariate Gaussian distribution with mean μ and covariance Σ, and σ is a real valued scalar. Note

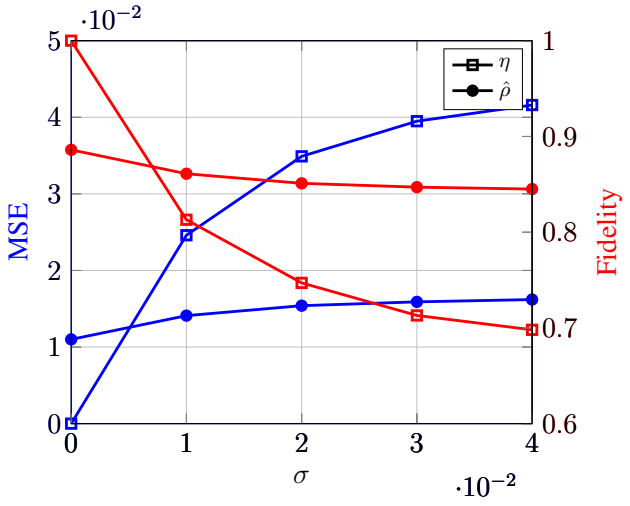

Fig. 2. Fidelity $\mathcal{F}$ and mean squared error (MSE) for the proposed CNN model averaged over 10^4 eight-dimensional mixed states. The squared curves represent $\mathcal{F}$ and MSE of η and ρ, where as the circled curves represent $\mathcal{F}$ and MSE of the output of CNN and ρ, where the input to the CNN is η.

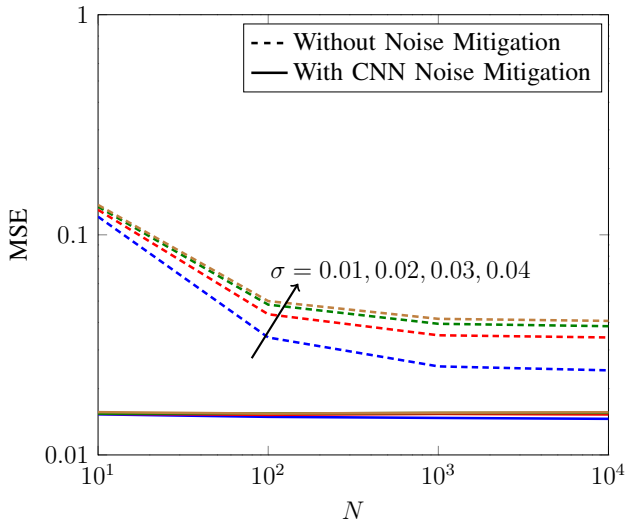

Fig. 3. Mean squared error (MSE) for the proposed CNN model averaged over 10^4 eight-dimensional mixed states. The dashed curves represent MSE of the Bloch vectors of $\hat{\eta}$ and ρ, where $\hat{\eta}$ has been reconstructed using standard tomography on $N \times \left(d^2-1\right)$ copies. The solid lines represent MSE of the Bloch vectors of $\hat{\rho}$ and ρ, where $\hat{\rho}$ are the Bloch vectors at the output of our CNN model when the input is $\hat{\eta}$.

that if η^* is an invalid state, we truncate negative eigenalues of η^*, normalize the remaining and reconstruct state η. Next we perform N measurements of η in each configuration setting Ω_i for $i \in \left\{1, 2, \cdots, d^2-1\right\}$ and observe outcome $|\ell\rangle$ for $\ell \in \{+1_i, -1_i\}$. Let n_i^ℓ represent the number of times we observe ℓ. Then the relative frequency defined as $f = \frac{n_i^\ell}{N}$ can be used to calculate the Bloch vectors of $\hat{\eta}$, which is an estimate of η. This process is illustrated in Figure 1. In this paper, we perform the noise mitigation and QST for $d = 8$.

First we will use η to estimate ρ. This will help us quantify the effectiveness of our CNN model in reducing state preparation errors for different values of σ. Next we will estimate ρ, given $\hat{\eta}$ for different values of N. This quantifies the robustness of our model to statistical noise. We convert our Bloch vectors into 8×8 matrices. The first two-dimensional convolutional layer has a kernel size 3×3, stride length of 1, 24 feature maps, zero padding, and an exponential linear unit (ELU) activation function. In the second layer, we increase feature maps to 128 while keeping the remaining settings invariant. Moreover, we attach two fully connected layers (FCL) with 2048 and 64, respectively, using ELU activation function for each. We perform 2-d dropout in our feature maps and 1-d dropout in FCLs with probability 0.25 and 0.1, respectively. We also perform batch normalization for all the hidden layers.

III. Numerical Results

In this section, we demonstrate the performance of our method detailed previously. We simulate eight-dimensional mixed states randomly sampled from the Ginibre ensemble. We report fidelity $\mathcal{F}$ between the estimated states $\hat{\rho}$ and true states ρ defined as

$$\mathcal{F}\left(\rho, \hat{\rho}\right) = \sqrt{\mathrm{tr}\left(\sqrt{\rho}\hat{\rho}\sqrt{\rho}\right)}, \tag{2}$$

and mean squared error (MSE) of their Bloch vectors defined as

$$\mathrm{MSE} = \frac{1}{d^2}\sum_{i=1}^{d^2-1}\left(r_\rho - r_{\hat{\rho}}\right)^2, \tag{3}$$

where r_ρ and $r_{\hat{\rho}}$ are the Bloch vectors of ρ and $\hat{\rho}$, respectively. Fidelity captures the idea of closeness between ρ and $\hat{\rho}$ such that $\mathcal{F}\left(\rho, \hat{\rho}\right) = 1$ if and only if $\rho = \hat{\rho}$, while MSE is a general measure of distance between corresponding Bloch vectors.

The squared curves in Figure 2 demonstrate the scaling of fidelity and MSE of quantum states with respect to the noise parameter σ. The circled curves demonstrate the scaling of fidelity and MSE of estimated quantum states $\hat{\rho}$ when η is the input to our CNN. Figure 2 shows that our model becomes more useful as σ increases.

Figure 3 compares the MSE of Bloch vectors of density matrix $\hat{\eta}$ reconstructed using standard tomography, and the matrix estimation $\hat{\rho}$ after post-processing, using our CNN model. The dashed curves represent standard tomography without noise mitigation. Therefore as the variance σ of our noise model increases, the average MSE for any value of N increases, and similarly, for any increasing N, MSE decreases for each σ curve. The solid curves demonstrate the MSE of $\hat{\rho}$ and ρ, where $\hat{\rho}$ is the estimate with CNN noise mitigation. The curves offer improvement over standard tomography for all values of σ and N. Moreover, the curves remain mostly invariant. An implication of this is that we can achieve high fidelities and low MSE, and offer improvement in estimation over contemporary methods, with small values of N for noisy quantum channels.

IV. Conclusions

In this work, we formulated QST and noise mitigation as a CNN problem. We have provided details of the architecture required to reproduce results. Moreover, we report the MSE and fidelity of reconstructed states generated with different variance. These results demonstrate that the utility of CNN model increases with as our system becomes more noisy. This formulation, therefore, is suitable for QST and noise mitigation of high dimensional quantum states prepared by real quantum devices.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1A2C2007037).

References

[1] U. Khalid, J. ur Rehman, and H. Shin, "Measurement-based quantum correlations for quantum information processing," *Sci. Rep.*, vol. 10, no. 1, p. 2443, 2020.

[2] U. Khalid, Y. Jeong, and H. Shin, "Measurement-based quantum correlation in mixed-state quantum metrology," *Quantum Inf. Process*, vol. 17, no. 12, p. 343, Nov. 2018.

[3] M. A. Ullah, J. ur Rehman, and H. Shin, "Quantum frequency synchronization of distant clock oscillators," *Quantum Inf. Process*, vol. 19, no. 5, p. 144, Mar. 2020.

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

[5] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, 10th ed. New York: Cambridge University Press, 2010.

[6] R. Blume-Kohout, J. K. Gamble, E. Nielsen, K. Rudinger, J. Mizrahi, K. Fortier, and P. Maunz, "Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography," *Nat. Commun.*, vol. 8, no. 1, pp. 1–13, 2017.

[7] C. E. Granade, C. Ferrie, N. Wiebe, and D. G. Cory, "Robust online Hamiltonian learning," *New J. Phys*, vol. 14, no. 10, p. 103013, 2012.

[8] Blume-Kohout and Robin, "Optimal, reliable estimation of quantum states," *New J. Phys*, vol. 12, no. 4, p. 043034, 2010.

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