Applied Bayesian Qubit State Tomography

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

***Abstract*—In this paper we present a simple Bayesian inference based single-stage quantum state tomography. Previous approaches such as maximum likelihood estimation are compared with Bayesian analysis and its advantages especially concerning error bars are highlighted. We describe the underpinnings of Monte Carlo based methods and show how they can be applied to reconstruct a valid density matrix representing a single qubit. We also demonstrate the formulation of qubit tomography as a problem such that statistical packages such as Stan can be accessed as a black-box and their results can be validated.**

I. Introduction

The ability to characterize the state of a quantum system is becoming essential with developments in research areas such as quantum computing, communications, and metrology [1] [2] [3]. Quantum state tomography (QST) employs a series of measurements on unknown, identically prepared quantum states to statistically reconstruct a density matrix that represents the state of the quantum system. Recent advancements in QST have also precipitated the characterization of more complex quantum systems such as those involving entanglement and higher dimensions [4] [5]. QST, therefore, serves as the cornerstone of any upcoming quantum technology [6].

QST techniques utilize sets of measurement observables, statistical models, and quantum resources. One of the earliest methods used simple linear transformations of measured data to uniquely determine a quantum state. Although, this method is straightforward, it can also reconstruct density matrices which are Hermitian with unit trace, but violate the condition of positive semi-definiteness due to experimental noise, and therefore do not correspond to any physical quantum state [7]. To ensure the validity of the reconstructed quantum state, maximum likelihood estimation (MLE), a statistical method, is used for tomography. A parametric function of MLE is created which always produces a valid density matrix. MLE is an efficient tool that converts QST to a convex optimization problem [8]. Just like the previous method, MLE constructs a density matrix that must completely agree with the frequencies of POVM measurements. Only when such a state is invalid (eigenvalue < 0), constraints of MLE force it to produce a valid but rank deficient state (i.e. some eigenvalues are 0). Therefore, despite the widespread usage of MLE in QST, it suffers from similar problems which arise due to the frequentists description of the statistical model. Moreover, there are no inherent error bars associated with the estimated state in MLE which are required to ascertain the quality of our estimate [9].

A viable solution to the problem of fitting states to agree with probabilities generated from noisy measurements is to move towards Bayesian statistics based QST. Bayesian mean estimation (BME) complements the likelihood function with a prior, which explicitly delineates our knowledge (or lack thereof) before any experiment is performed, to produce a valid distribution. Major benefits inherent to BME include the presence of error bars associated with the estimated state, and that it produces the most accurate estimation for a finite number of copies N of the state [9] [10]. As the number of experiments increase, the posterior distribution produced by BME tends to become independent of the prior, and relies more on the likelihood function. That is, BME based QST relies more on prior knowledge in the presence of insufficient data, and this reliance reduces with increasing data [11].

In this paper, we examine the key elements of BME and carry out a single stage QST using only a fraction of identically prepared copies required for full scale quantum state tomography for a single qubit. We also quantify the error of our posterior distribution. This is in fact the first step of more complex tomography techniques involving sequential importance sampling (SIS) algorithms. The tomography problem is coded in Stan, a probabilistic programming language that implements an advance version of the Monte Carlo algorithm to efficiently traverse the posterior space.

This paper is structured as follows. We expound the basics of Bayesian inference qualitatively, required to understand its application to state tomography in section II. Section III briefly explains the fundamentals of Stan. In section IV, we apply our understanding of BME and Stan to simulate a single stage quantum state tomography of a qubit, provide results, and analyze our posterior. We conclude in section V.

II. Applied Bayesian Inference

In Bayesian analysis, there are three key terms which interact to produce a posterior distribution, $p\left(\theta|\mathrm{data}\right)$, which is used to make estimates about the parameter θ. The prior, $p\left(\theta\right)$, is a valid probability distribution that explicitly represents our pre-data (measurement) knowledge of the experiment. This could be driven by an existing study, understanding of the process or any other bias [12]. Depending on the amount of reliable a priori knowledge, we can choose from weakly to strongly informative priors.

$$p\left(\theta|\mathrm{data}\right) = \frac{\mathcal{L}\left(\theta|\mathrm{data}\right) \times p\left(\theta\right)}{p\left(\mathrm{data}\right)}, \tag{1}$$

where $\mathcal{L}\left(\theta|\mathrm{data}\right) = p\left(\mathrm{data}|\theta\right)$.

The likelihood function, $\mathcal{L}\left(\theta|\mathrm{data}\right)$, is a data driven element which reduces the reliance of the posterior on the prior as data increases. In $\mathcal{L}(\theta|\mathrm{data})$, the data is kept fixed and parameter θ is varied. The denominator in (1) is often intractable for complex multidimensional models, and for practical purposes can be thought of as a term that normalizes the product of the prior and the likelihood to produce a valid posterior distribution.

Since the multidimensional nature of Bayesian integrals make analytical and computational solutions impossible (or at least impractical), we must circumvent this process. As it turns out, we do not need to calculate the exact probability density function (pdf) of our posterior distribution. As long as we can effectively sample from the product of our likelihood and prior at sufficient points in parameter space, the relative sampling frequency provides us with an un-normalized posterior distribution. (1) can be rewritten as

$$p(\theta|data) \propto p(data|\theta) \times p(\theta)\,. \tag{2}$$

This is the key insight underlying Monte Carlo Markov Chain (MCMC) algorithms. MCMC solvers utilize dependent sampling to traverse the posterior. Generally, this involves iterations of generating random samples, and at each point in parameter space, performing an accept-reject rule which governs whether to move to the next sample or not. Thereby creating a Markov chain where the next step is dependent on the current position. The frequency of samples traversed during this process forms our approximate posterior. This equals the true posterior distribution in the limiting case when the number of iterations approach infinity.

Since infinite iterations are not possible, it begs the question when does our solver approach convergence? For the purpose, multiple Markov chains are initialized. The average of the variance of each chain, the within chain variance, and the variance of the mean of each chain, the between-chain variance, are calculated. When the ratio, $\hat{\mathcal{R}}$, of these variances approaches 1, our chains are said to have converged [11].

$$\hat{\mathcal{R}} = \sqrt{\frac{W + \frac{1}{n}(B - W)}{W}}, \tag{3}$$

where n is the total number of samples per chain, and W and B are the within and between-chain variances, respectively.

III. Overview of Stan

Stan implements the No-U-Turn Sampler (NUTS) version of HMC. Stan's MCMC solver works as a black-box that conveniently applies Bayesian inference on our problem. However the problem setup is coded in Stan. This involves providing at least 3 blocks of information for meaningful Bayesian inference namely data, parameters and the model.

Since Stan is statically typed language, the type of data or parameters used must be declared. In the data block, variables are created which are used to pass data to Stan. In the parameter block, we declare all the parameters that we want to infer from our model. The model block is used to specify the likelihood function and priors. Stan essentially works in the negative log posterior (NLP) space, and hence for a new

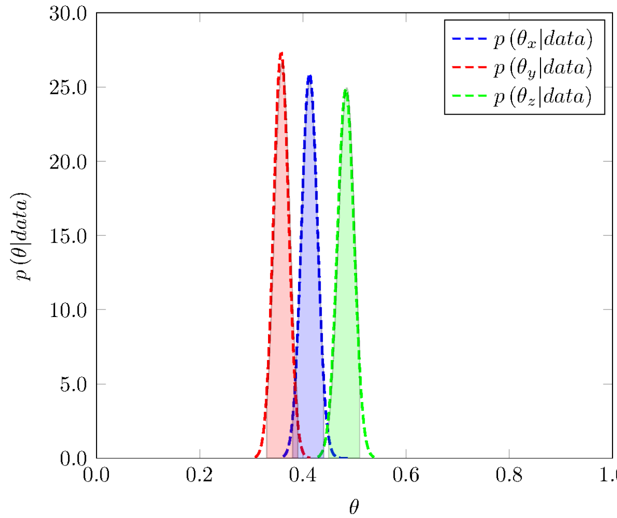

Fig. 1. PDF of posterior distributions for θ_x, θ_y and θ_z. Each θ is centered at BME, and the width of the shaded areas on the horizontal axis correspond to 95% credible regions for $N = 1000$.

TABLE I
Statistical Summary of QST.

θ	Mean	2.5%	97.5%	$\hat{\mathcal{R}}$
θ_x	0.41	0.38	0.44	1.0
θ_y	0.36	0.33	0.39	1.0
θ_z	0.48	0.45	0.51	1.0

sample at each step of HMC, it evaluates the following instead of (2) [11]

$$\log p(\theta|data) \propto \log p(data|\theta) + \log p(\theta)\,. \tag{4}$$

IV. Method and Results

Any valid two dimensional density matrix, ρ, can be represented by

$$\rho = \frac{1}{2}\left(\mathbb{I} + \vec{r}\cdot\vec{\sigma}\right), \tag{5}$$

where $\vec{r} = (r_x, r_y, r_z)$ is called a Bloch vector such that $\|\vec{r}\|_2 \leq 1$ and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ [7] is a vector of Pauli operators. Therefore to successfully perform QST, we must determine $\vec{r}$.

The first step is to perform measurements on an ensemble of identically prepared quantum states, ρ. Measurements of qubits involve a set of configurations $a \in \mathcal{A}$, where measurements in each a correspond to observing one of the finite outcomes, $|\psi\rangle$, from a positive operator-valued measure (POVM) set, $\mathcal{M}_a$ [13]. Here we perform the standard six-state tomography on a qubit where $\mathcal{A} = \{\sigma_x, \sigma_y, \sigma_z\}$ and outcomes for a are their eigenstates such that $\mathcal{M}_a = \{|\psi_1\rangle, |\psi_2\rangle\}$ where $|\psi_1\rangle$ corresponds to the +1 eigenvalue and $|\psi_2\rangle$ corresponds to the -1 eigenvalue. Therefore, the probability of outcome of $|\psi\rangle \in \mathcal{M}_a$ is $p(|\psi_i\rangle) = \mathrm{Tr}(|\psi_i\rangle\langle\psi_i|\rho)$ for $i = 1, 2$ according to Born's rule of quantum theory [14]. Then for each configuration a we can determine the corresponding element in $\vec{r}$ using the following relation

$$r = p(|\psi_1\rangle) - p(|\psi_2\rangle). \tag{6}$$

As $\mathrm{Tr}(|\psi_1\rangle\langle\psi_1|\rho) = 1 - \mathrm{Tr}(|\psi_2\rangle\langle\psi_2|\rho)$, (6) reduces to

$$r = 2 \times p(|\psi_1\rangle) - 1, \tag{7}$$

which means that $\vec{r}$ can be calculated from the projections of ρ on the $+1$ eigenstates of $\vec{\sigma}$.

Since we cannot determine $p(|\psi_1\rangle)$ without knowing ρ, we prepare $3N$ copies of state ρ and measure σ_x, σ_y and σ_z on N qubits each, and note the frequencies of outcome n_i of observing the +1 eigenstate of σ_i for $i = x, y, z$. This is the data that we provide to Stan which treats it as three separate studies, one for each configuration a. To estimate $p(|\psi_1\rangle)$ via Bayesian inference, we declare parameters θ_x, θ_y and θ_z in Stan's parameter block to represent the probability of outcome $|\psi_1\rangle$ for σ_x, σ_y and σ_z respectively. For each θ, we use a weakly informative prior centered on the positive axis of the Bloch sphere with high variance.

We use a Bernoulli distribution where θ is used to represent the probability of observing the +1 eigenstate. Each measurement is assumed to be independent. Hence, our likelihood function for N measurements takes the following form for each study

$$\mathcal{L}(\theta|data) \propto \theta^{n_i}(1-\theta)^{(N-n_i)}. \tag{8}$$

We performed QST on a random qubit defined by the density matrix

$$\rho = \begin{pmatrix} 0.4891 & -0.1315 + 0.1453j \\ -0.1315 - 0.1453j & 0.5109 \end{pmatrix}$$

using 4 chains and 1000 HMC iterations for N = 1000. Using the BMEs tabulated in Table 1 of the posterior PDFs illustrated in Fig. 1 and (7), we can approximate the Bloch vector so that $\vec{r} \approx (-0.18, -0.28, -0.04)$. The log infidelity of the reconstructed state $\hat{\rho}$ is calculated to be less than -3.5. Moreover, using the quantile values provided in Table 1, we can see that θ_x, θ_y and θ_z have a 95% credible interval of $0.38 < \theta_x < 0.44$, $0.33 < \theta_y < 0.39$ and $0.45 < \theta_z < 0.51$ respectively. Values of $\vec{\mathcal{R}}$ show that the chains converged successfully.

V. Conclusions

In this work, we have formulated QST as a Bayesian problem. We have highlighted fundamentals of MCMC required to apply Bayesian Inference. Moreover, through an example, we have shown that it does not suffer from the drawbacks of MLE and reconstructs a valid quantum states with easily quantifiable error bars. Moreover, this formulation can also be used with different tomographically complete measurement bases to make QST adaptive.

Acknowledgment

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

References

[1] J. L. O'Brien, G. Pryde, A. Gilchrist, D. James, N. K. Langford, T. Ralph, and A. White, "Quantum process tomography of a controlled-NOT gate," *Phys. Rev. Lett.*, vol. 93, no. 8, p. 080502, 2004.

[2] 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.

[3] 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.

[4] A. Khan, J. ur Rehman, K. Wang, and H. Shin, "Unified monogamy relations of multipartite entanglement," *Sci. Rep.*, vol. 9, no. 1, p. 16419, Nov. 2019.

[5] J. Bae, B. C. Hiesmayr, and D. McNulty, "Linking entanglement detection and state tomography via quantum 2-designs," *New J. Phys*, vol. 21, no. 1, p. 013012, 2019.

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

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

[8] 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, 1999.

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

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

[11] B. Lambert, *A Student's Guide to Bayesian Statistics*. London: SAGE, 2018.

[12] C. Granadeand, J. Combes, and D. Cory, "Practical Bayesian tomography," *New J. Phys*, vol. 18, no. 3, p. 033024, 2016.

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

[14] J. Řeháček, B.-G. Englert, and D. Kaszlikowski, "Minimal qubit tomography," *Phys. Rev. A*, vol. 70, no. 5, p. 052321, 2004.