

Computationally efficient Bayesian quantum state tomography

Joseph M. Lukens,^{1,*} Kody J. H. Law,^{2,†} Ajay Jasra,³ and Pavel Lougovski¹

¹Quantum Information Science Group, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

²School of Mathematics, University of Manchester, Manchester, M13 9PL, UK

³Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal, 23955-6900, KSA

*lukensjm@ornl.gov; †kodylaw@gmail.com

Abstract: We describe a method for Bayesian quantum state estimation combining efficient parameterization, a pseudo-likelihood, and advanced numerical sampling techniques. Examples reveal significant computational speedup, indicating the approach’s promise in practical quantum state tomography. © 2020 The Author(s)

Bayesian inference is a powerful paradigm for quantum state tomography (QST), producing a complete probability distribution on the state ρ , quantifying uncertainty in a natural way, utilizing available information optimally, and avoiding unjustifiably optimistic estimates [1]. Yet the numerical challenges associated with sampling from complex probability distributions hampers Bayesian tomography in practical settings. Here we summarize a recently developed method for Bayesian QST [2] that is straightforward to implement and numerically efficient. Our stand-alone approach leverages developments in multiple fields, including density matrix parameterization [3], PAC-Bayesian machine learning [4], and Markov chain Monte Carlo (MCMC) algorithms [5]. Compatible with any number of observations, this method enjoys all the standard advantages of Bayesian QST but with significantly improved computational efficiency, highlighting the power of advanced statistical techniques in practical quantum tomography.

Consider a quantum state in a D -dimensional Hilbert space described by the parameterized density matrix $\rho(\mathbf{x})$. Bayes’ theorem states that the posterior probability distribution of \mathbf{x} , given results \mathcal{D} of some experiment, follows via $\pi(\mathbf{x}) = \frac{1}{Z} L_{\mathcal{D}}(\mathbf{x}) \pi_0(\mathbf{x})$, where the likelihood $L_{\mathcal{D}}(\mathbf{x}) \propto \mathcal{P}(\mathcal{D}|\rho)$, the probability of receiving the particular set of outcomes given state ρ , $\pi_0(\mathbf{x})$ is the prior distribution, and Z is a normalizing constant such that $\int d\mathbf{x} \pi(\mathbf{x}) = 1$. With access to $\pi(\mathbf{x})$, the expectation of any function of ρ can be obtained according to $\langle \phi(\rho) \rangle = \int d\mathbf{x} \pi(\mathbf{x}) \phi(\rho(\mathbf{x}))$. Evaluating integrals of this form is numerically challenging due to their generally complicated features and high dimensionality, even for moderate-size systems (e.g., two qubits). Accordingly, a variety of MCMC sampling methods have been invoked in the literature; however, the computing time required for convergence can easily make these methods intractable for systems of interest. The current work is motivated by the realization that a more tailored sampling procedure can attain significant improvements.

Our proposed Bayesian QST workflow is summarized in Fig. 1. Key aspects include: (i) over-parameterization [3] of the density matrix ($2D^2 + D$ independent real numbers, i.e., D complex D -dimensional vectors \mathbf{z}_k with D weights y_k); (ii) use of a pseudo-likelihood centered on a simple least-squares estimate of the data, which in the case of tomographically complete measurements reduces the computational cost of likelihood evaluation from $\mathcal{O}(D^4)$ to $\mathcal{O}(D^2)$ operations (for tomographically incomplete measurements, this estimate is a projection $P_{\mathcal{M}}(\cdot)$ onto the measured subspace); and (iii) application of a preconditioned Crank–Nicolson (pCN) sampling algorithm, which eliminates the slowdown in convergence from increasing dimensionality intrinsic to the standard random-walk Metropolis algorithm [5]. (See [2] for details on the sampling procedure.)

To explore the effectiveness of the new method, we perform QST on results from the frequency-bin quantum

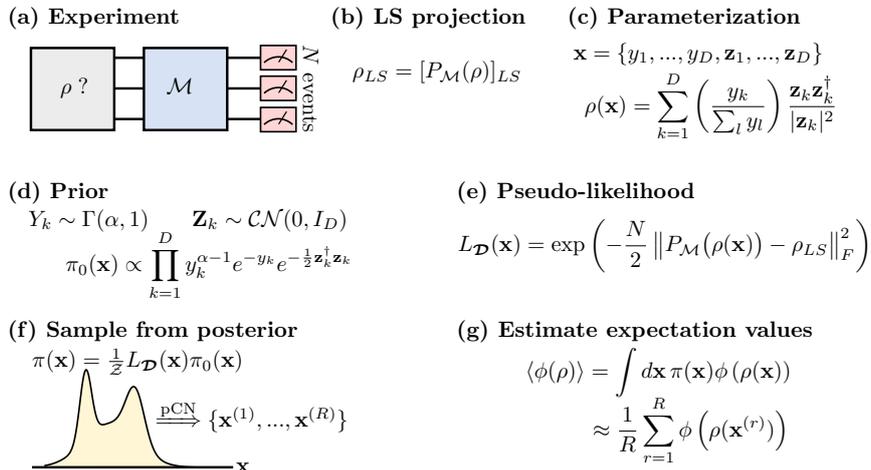


Fig. 1. Overview of proposed Bayesian QST method.

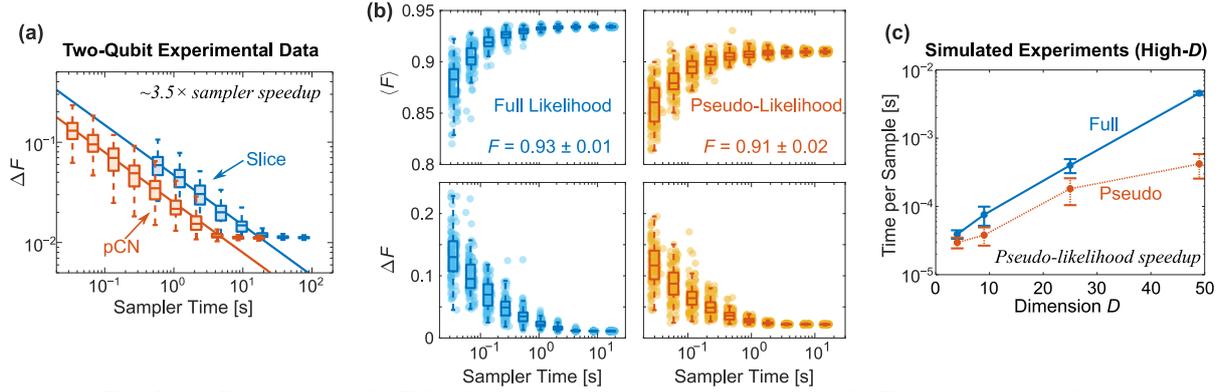


Fig. 2. (a) Comparison of pCN and slice sampling convergence rates. (b) Full and pseudo-likelihood convergence with increasing Markov chain lengths. (c) Computational cost of our algorithm for simulated data.

optics experiment of [6]. For the first test, we compare the speed of pCN to slice sampling, a common and flexible MCMC approach. In this example, we take $\alpha = 1$ for a uniform prior and consider a full likelihood (for consistency with [6]). We increase the number of points in the Markov chain and monitor the standard deviation ΔF of the fidelity $F(\mathbf{x}) = \langle \Psi | \rho(\mathbf{x}) | \Psi \rangle$, where $|\Psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ is the ideal entangled state. $R = 2^{10}$ samples are kept from a total of RT , with T the thinning parameter used to reduce serial correlation in the chain. For each value of T , we run 100 independent samplers, returning 100 separate estimates of ΔF , which we present in Fig. 2(a), using a box plot to summarize the statistics: the center mark denotes the median, upper and lower lines enclose the 25th–75th percentiles, and the whiskers extend to the smaller of the farthest point or $1.5\times$ the length of the box. Initially, both algorithms see ΔF reduce with a log-log slope of $-1/2$, until converging to final values. Linear least-squares fits to the first five and seven points of the slice and pCN curves, respectively, give a $\sim 3.5\times$ temporal speedup for pCN at a given convergence level. Such an improvement—even for this comparatively small system of two qubits—is significant for practical QST, where computational time represents a precious commodity.

In the second test, we compare the behavior of a *full* (multinomial) likelihood [6] to that of the proposed *pseudo*-likelihood. Figure 2(b) plots $\langle F \rangle$ and ΔF for 100 samplers utilizing the pCN algorithm, as the number of points increases, taking $\alpha = 1$ in the prior. Both likelihoods converge to similar values, though the pseudo case returns slightly lower mean and higher uncertainty. The general congruity between the two cases offers evidence in favor of both our choice for the loss function variance ($\sigma^2 = 1/N$) and the projector-based approach to limited measurements [encapsulated by $P_{\mathcal{M}}(\cdot)$]. Even the slightly lower fidelity for the pseudo-likelihood is a positive feature, in that it does not overestimate the state’s fidelity beyond that predicted by a complete model.

Finally, in order to explore dimensionalities beyond that of the experimental results available to us, we then generate simulated tomographic data for a two-qudit state of the form $\rho = \lambda |\Psi\rangle \langle \Psi| + \frac{1-\lambda}{D} I_D$, where $|\Psi\rangle = \sum_{k=1}^d |k\rangle_A |k\rangle_B$ is a high-dimensional Bell state, $D = d^2$, and I_D is the $D \times D$ identity matrix. Count data are obtained by cycling through all pairwise combinations of $(d+1)$ mutually unbiased bases on qudits A and B and drawing from a multinomial to emulate an experimental coincidence distribution. Running the pCN algorithm on these observations and recording the time per sample, we find the trends in Fig. 2(c). While comparable at low D , the evaluation times for the full and pseudo-likelihood approaches become increasingly disparate as D grows, reaching $\sim 10\times$ for $D = 49$. Due to limits on the size of the datasets we could generate, we were not able to reach the predicted $\mathcal{O}(D^2)$ asymptotic scaling improvement, yet these time tests confirm that the pseudo-likelihood does offer computational speedups under appropriate conditions.

We thank R. S. Bennink and B. P. Williams for discussions. This work was performed in part at Oak Ridge National Laboratory, operated by UT-Battelle for the U.S. Department of Energy under contract no. DE-AC05-00OR22725. Funding was provided by the U.S. Department of Energy, Office of Advanced Scientific Computing Research, through the Quantum Algorithm Teams and Early Career Research Programs.

References

1. R. Blume-Kohout, New J. Phys. **12**, 043034 (2010).
2. J. M. Lukens, K. J. H. Law, A. Jasra, and P. Lougovski, arXiv:2002.10354 (2020).
3. T. T. Mai and P. Alquier, J. Stat. Plan. Inference **184**, 62 (2017)
4. B. Guedj, arXiv:1901.05353 (2019).
5. S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White, Statist. Sci. **28**, 424 (2013).
6. H.-H. Lu, J. M. Lukens, N. A. Peters, B. P. Williams, A. M. Weiner, and P. Lougovski, Optica **5**, 1455 (2018).