Self-guided quantum tomography

Christopher Ferrie

Center for Quantum Information and Control, University of New Mexico, Albuquerque, New Mexico, 87131-0001

(Dated: June 17, 2014)

We introduce a self-learning tomographic technique in which the experiment guides itself to an estimate of its own state. Self-guided quantum tomography (SGQT) uses measurements to directly test hypotheses in an iterative algorithm which converges to the true state. We demonstrate through simulation on many qubits that SGQT is a more efficient and robust alternative to the usual paradigm of taking a large amount of informationally complete data and solving the inverse problem of post-processed state estimation.

The act of inferring a quantum mechanical description of a physical system—assigning it a quantum state—is referred to as tomography. Tomography is a required, and now routine, task for designing, testing and tuning qubits—the building blocks of a quantum information processing device [1]. However, in a grand irony, the exact same exponential scaling that gives a quantum information processing device its power also limits our ability to characterize it.

That tomography is a problem exponentially hard in the number of qubits has lead to many proposals for efficient learning within restricted subsets of quantum states [2, 3]. On the other hand, if we expect to have prepared a specific target state, efficient protocols exist to estimate the fidelity to this state [4, 5]. These protocols are direct in the sense that few measurements are required to provide an estimate of the fidelity to the target rather than first reconstructing the state then calculating the fidelity.

The proposal presented here is direct in the same sense as [4, 5], but converges to state itself. The algorithm is iterative—from directly estimating a distance measure to the underlying state, the experiment guides itself to a description of its own state: self-guided quantum tomography (SGQT). The distance measure \( m \) is arbitrary, in the sense that any measure will work. However, the more rapidly the experiment can provide an estimate for \( m \), the more rapidly SGQT will converge, such that if \( m \) can be estimated efficiently, then SGQT will be efficient.

Before describing exactly what SGQT is, we first state what it is not by reviewing the problem of tomography. There is some true state \( \rho \) which generates data, a list of measurement outcomes corresponding to effects \( D = \{ E_0, E_1, \ldots \} \). The probability to observe this data is given by the Born rule

\[
\Pr(D|\rho) = \prod_k \text{Tr}(\rho E_k). \tag{1}
\]

The prevailing method is to solve the inverse problem of identifying an accurate estimate, \( \sigma \), of \( \rho \) given a sample data set drawn from this distribution. Here the approach is quite different. We begin with a distance measure on states \( m(\rho, \sigma) \). The only requirement is that this measure can be estimated from experiment, such that we have access to

\[
f(\sigma) = \langle m(\rho, \sigma) \rangle. \tag{2}
\]

This quantity fluctuates from noise which can come from a variety of sources but is always present due to the fundamental statistical nature of quantum mechanics—also known as shot-noise.

Here we will provide an algorithm to iteratively propose new states \( \sigma \) such that we converge to \( \rho \) only via estimates of \( f(\sigma) \). The core of the algorithm is a stochastic optimization technique known as simultaneous perturbation stochastic approximation (SPSA) [6].

SPSA is an iterative optimization technique which uses only two (noisy) function calls per iteration to estimate the gradient. In the context of state estimation, this means that SGQT requires only two proposal states \( \sigma_\pm \) and experimental estimates \( f(\sigma_\pm) \) to provide an unbiased estimate of the gradient, which in turn provides the direction to the true state. This is the key element which provides SGQT with its efficiency. For one might expect that to estimate the gradient would require \( O(2^n) \) proposal states, where \( n \) is the number of qubits. In the remainder, we detail the algorithm and demonstrate via numerical experiments the claimed efficiency and robustness of SGQT.

The steps of each iteration proceed as follows (suppose we are at iteration \( k \)): (1) Generate a random direction to search in defined by \( \Delta_k \). (2) Calculate the estimated gradient in that direction,

\[
g_k = \frac{f(\sigma_k + \beta_k \Delta_k) - f(\sigma_k - \beta_k \Delta_k)}{2\beta_k} \Delta_k. \tag{3}
\]

(3) Calculate the next iterate via

\[
\sigma_{k+1} = \sigma_k + \alpha_k g_k. \tag{4}
\]

The functions \( \alpha_k \) and \( \beta_k \) control the convergence and are user defined, although they are usually specified in the forms

\[
\alpha_k = \frac{a}{(k + 1)^s}, \quad \beta_k = \frac{b}{(k + 1)^t}, \tag{5}
\]

where \( a, b, s \) and \( t \) are chosen first roughly based on extensive numerical studies for many problems then tweaked based on numerical simulations for the problem at hand. Much of the former task has been done and generally good choices are [7] \( s = 0.602 \) and \( t = 0.101 \). The random direction \( \Delta_k \) is arbitrary, up to some constraints.

Infidelity

\[ \rho \] is the algorithm when \( n = 10^3 \) iterations.

\[ s = \text{ranging in the basis containing } | \psi \rangle \text{ by the algorithm when } n \] which fluctuates due to statistical noise. In Fig. 1, we

\[ m \] to

\[ m(\sigma_k, \rho) - m(\rho, \rho) = \frac{\epsilon}{2} \|H[m](\rho)\| \|\sigma_k - \rho\|_2^2, \] \[ \|H[m](\rho)\| \] is the spectral norm, the largest eigenvalue, of \( H[m](\rho) \). Thus, we are achieving optimal performance when \( m \approx 1/k^{2/3} \).

\[ m(\sigma_k, \rho) = 1 - |\langle \psi | \phi \rangle|^2 \] and can be estimated by measuring in the basis containing \( |\phi\rangle \). That is, by counting the number of outcomes in the direction of \( |\phi\rangle \), say \( n(\phi_+) \), we can estimate

\[ m(\psi, \phi) = 1 - \frac{n(\phi_+)}{n(\phi_+) + n(\phi_-)}, \] which fluctuates due to statistical noise. In Fig. 1, we see this manifest through the volatility of the path taken by the algorithm when \( n(\phi_+) + n(\phi_-) \approx N = 10^2 \) and

\[ N = 10^4 \]. We might expect then that more experiments are needed to mitigate these fluctuations as we converge on the target true state. While this appears to be true for “small” repetition and iteration numbers, we will see, however, that this intuition fails us in the asymptotic regime. That is, for a fixed number of iterations, the performance is roughly independent of the number of experiments. This will demonstrate the superior efficiency of SGQT to converge well beyond what we might expect to be the “noise floor”.

In our discussion, we will refer to the follow three algorithmic and experimental parameters: \( N \), the number of experiments per estimate of \( m \); \( M \), the number of estimates of \( m \) per iteration; and \( k \), the number of iterations. Thus, the total number of experiments is \( N_{\text{tot}} = N \cdot M \cdot k \). For standard finite difference gradient estimation, we have \( M = 2d \), where \( d \) is the real dimension of the state space. For \( n \) pure qubits \( d = 2(2^n - 1) \), which grows exponentially. For SGQT, however, \( M = 2 \) regardless of the dimension, thus we will restrict our attention to \( N \) and \( k \) with the understanding that \( N_{\text{tot}} = 2Nk \).

We continue with the qubit example of Fig. 1 to determine how the performance of SGQT scales with \( N \) and \( k \) retaining the fidelity objective function in Eq. (8). In Fig. 2, we plot the infidelity as a function \( k \). We find, independent of \( N \), the asymptotic scaling of infidelity is \( O(1/k^{\beta/2}) \) with \( \beta \in (0.66, 0.69) \). This is roughly what we would expect from the bound in Eq. (7) if the state converges at the optimal rate of \( O(k^{-1/3}) \).

The lines in Fig. 2 are three slices of the simulation results shown in Fig. 3, where we have shown the median performance as a function of both \( N \) and \( k \). Note that
as we approach the asymptotic regime, the performance depends less on \( N \). In other words, as we convergence, it is not necessary to increase the number of experimental repetitions to increase the accuracy of the estimated fidelity. This false intuition would, however, hold true if we were to use an optimization algorithm (such as a standard gradient descent) which does not take account of the stochasticity in estimating the fidelity.

Above we have explored the efficacy of SGQT for single qubit tomography. In Fig. 4, we generalize to multiple qubits. As expected, since even estimating the fidelity to an arbitrary single pure target state is not efficient, the convergence of SGQT is not efficient in the number of qubits. However, SGQT outperforms standard tomography since the latter still requires the additional exponentially hard problem of reconstructing the state after the measurements are taken.

If, on the other hand, we restrict the class of states to one which can be efficiently specified and the fidelity to which can be efficiently estimated, SGQT becomes efficient. As an example, consider the W-class of states which have found use in the theory of entanglement [8]. An \( n \) qubit W-class state is one of the form
\[
|\psi\rangle = \alpha_1 |01\ldots0\rangle + \alpha_2 |00\ldots1\rangle + \cdots + \alpha_n |00\ldots1\rangle.
\] (9)

Note that the number of parameters grows linearly with the number of qubits. Moreover, the fidelity to a target in this class can be estimated efficiently [4, 5]. But if we actually want to learn the state we are faced with an new problem: the actual true state might not lie in this subclass. Using SGQT with fidelity estimation we can, however, efficiently find the W-class state with highest fidelity to the true state. This points to both the robustness of SGQT and efficacy of solving the problem.
The performance of SGQT for this problem is demonstrated in Fig. 5, where SGQT is shown to find the highest fidelity W-class state to a randomly generate mixed state. The mixed state is generated by first Haar randomly choosing a W-state, then subjecting it to 5% depolarizing noise. Fits to $O(1/k^3)$ asymptotic scaling for each qubit number lie in $\beta \in (0.45, 0.58)$. This is slightly less than optimal ($\beta = 2/3$) and can likely be improved with more extensive numerical exploration of the algorithmic parameters or higher-order methods.

Finally, we show that SGQT is robust to small amounts of measurement errors, which is a part of the ever-present state preparation and measurement (SPAM) error problem. In Fig. 6, we demonstrate the robustness of SGQT to measurement errors for W-state estimation, where the measurement error is simulated by randomly perturbing the measurement target state with zero-mean Gaussian noise with a (quite high) standard deviation of 0.1. The convergence is roughly independent of $n$, the number of qubits, with fits giving $\beta \in (0.57, 0.61)$. In other words, SGQT is robust to both statistical and technical measurement noise.

Here we have considered examples of pure state tomography using infidelity since it has a clear-cut interpretation and is a standard error metric. We reiterate that SGQT will work with any distance metric so long as it is estimable via experiment. The only caveat is that the efficiency of SGQT is directly related to the efficiency with which the distance can be estimated. For example, in direct fidelity estimation [4, 5], only certain classes of states can be validated in an efficient way, such as our W-state example. If only a subclass of states is considered, SGQT will converge to the nearest state within that subclass, as we have demonstrated with W-states. We have also shown the SGQT is robust to certain forms of SPAM errors. In the same way as for states, SGQT can be used to find quantum channels where randomized benchmarking [9, 10] can be used to efficiently estimate the fidelity to certain classes of unitaries. Finally, we note that to further mitigate the issues of complexity, it may become viable in the future to aid the estimation of the distance measure with quantum resources [11, 12], such as the swap test [13].

In summary, we have provided an experimental protocol to learn quantum states without the need for classical reconstruction—that is, the quantum system guides itself to a description of its own state. Using ideas from stochastic optimization theory and the direct estimation of fidelity, we have shown that certain classes of states can be learned efficiently in an iterative experimental protocol which ends with the experiment determining its own state. This result demonstrates that the standard, and prohibitive, paradigm of first collecting massive amounts of data, then solving the inverse problem of state estimation is unnecessary. Perhaps with an eye to the future, the approach considered here is a step toward quantum learning, in which autonomous quantum machines learn and manipulate their environment without the need of a human operator.

Acknowledgments

This work was supported in part by National Science Foundation Grant No. PHY-1212445 and by the Canadian Government through the NSERC PDF program.

[1] M. A. Nielsen and I. L. Chuang, Quantum computation and quantum information. Cambridge University Press (2010).
[2] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, Efficient quantum state tomography, Nature Communications 1, 149 (2010).
[3] G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, Permutationally Invariant Quantum Tomography, Physical Review Letters 105, 250403 (2010).
[4] S. T. Flammia and Y.-K. Liu, Direct fidelity estimation from few Pauli measurements, Physical Review Letters 106, 230501 (2011).
[5] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, Practical characterization of quantum devices without tomography, Physical Review Letters 107, 210404 (2011).

[6] J. C. Spall, Multivariate stochastic approximation using a simultaneous perturbation gradient approximation, IEEE Transactions on Automatic Control 37, 332 (1992).

[7] P. Sadegh and J. C. Spall, Optimal random perturbations for stochastic approximation using a simultaneous perturbation gradient approximation, IEEE Transactions on Automatic Control 43, 1480 (1996).

[8] R. D. Somma, J. Chiaverini and D. J. Berkeland, Lower bounds for the fidelity of entangled-state preparation, Physical Review A 74, 052302 (2006).

[9] O. Moussa, M. P. da Silva, C. A. Ryan, and R. Laflamme, Practical experimental certification of computational quantum gates using a twirling procedure, Physical Review Letters 109, 070504 (2012).

[10] E. Magesan, J. M. Gambetta, B. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, et al., Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking, Physical Review Letters 109, 080505 (2012).

[11] N. Wiebe, C. Granade, C. Ferrie, and D. Cory, Hamiltonian Learning and Certification Using Quantum Resources, arXiv:1309.0876 (2013).

[12] N. Wiebe, C. Granade, C. Ferrie, and D. G. Cory, Quantum Hamiltonian learning using imperfect quantum resources, Physical Review A 89, 042314 (2014).

[13] H. Buhrman, R. Cleve, J. Watrous and R. de Wolf, Quantum Fingerprinting, Physical Review Letters 87, 167902 (2001).