Self-Guided Quantum Tomography

Christopher Ferrie

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

We introduce a self-learning tomographic technique in which the experiment guides itself to an estimate of its own state. Self-guided quantum tomography uses measurements to directly test hypotheses in an iterative algorithm which converges to the true state. We demonstrate through simulation on many qubits that Self-guided quantum tomography 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 postprocessed 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 Refs. [4,5], but converges to the 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 ρ 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

$$\Pr(D|\rho) = \prod_k \mathrm{Tr}(\rho E_k). \tag{1}$$

The prevailing method is to solve the inverse problem of identifying an accurate estimate, σ, of ρ 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 σ such that we converge to ρ 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 $O(2^{2n})$ 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 Δ_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

$$\sigma_{k+1} = \sigma_k + \alpha_k g_k. \qquad (4)$$

The functions α_k and β_k control the convergence and are user defined, although they are usually specified in the forms

$$\alpha_k = \frac{a}{(k+1+A)^s}, \qquad \beta_k = \frac{b}{(k+1)^t}, \qquad (5)$$

where $a, A, b, s,$ 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. It is useful to note that much of the former task has been done and generally good choices are [7] $s = 0.602$ and $t = 0.101$. These values indeed work for our problem but the asymptotically optimal [6] values $s = 1$ and $t = 1/6$ seem to perform well even early on. The random direction Δ_k is arbitrary, up to some constraints. Given some vector representation of the operators, we take the common choice $\Delta_k = \pm 1$ for each element of the vector and the sign randomly assigned by a fair coin toss. For $a, A,$ and b, the optimal parameters are more problem dependent and here we have two different sets depending on whether we are demonstrating asymptotic performance or not—these values will be noted when the results are discussed later on. Convergence results on SPSA [6,7] imply that the infidelity, for example, decreases at rate $O(1/k^\gamma)$, where the exponent actually achieved is highly problem dependent. Asymptotic results, however, give a rate $\gamma \approx 1$ to first order.

We illustrate the iterations of the algorithm in Fig. 1 for the metric $m(\rho, \sigma) = 1 - F(\rho, \sigma)$, the infidelity between the two states. For pure states this is equivalent to $m(\psi, \phi) = 1 - |\langle \psi | \phi \rangle|^2$ 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

$$m(\psi, \phi) \approx 1 - \frac{n(\phi_+)}{n(\phi_+) + n(\phi_-)}, \qquad (6)$$
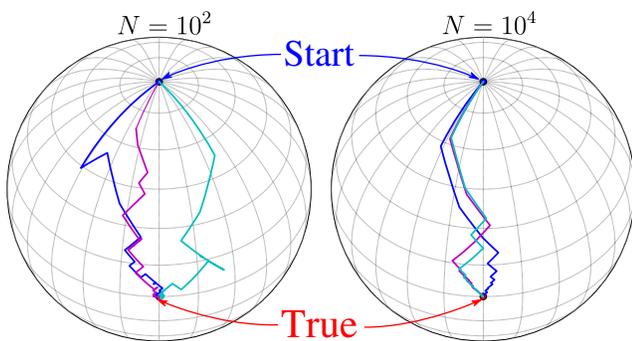
FIG. 1 (color online).   Estimating the state of a qubit via SGQT and the fidelity metric. On the left, each fidelity estimate was generated with $N = 10^2$ single shot experiments—on the right, $N = 10^4$. Each line represents one of three runs with $k = 10^3$ iterations.

which fluctuates due to statistical noise. In Fig. 1, we see this manifest through the volatility of the path taken by the algorithm when $n(\phi_+) + n(\phi_-) \equiv N = 10^2$ 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. We will see, however, that this intuition fails us. 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 following 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 $N_{\text{tot}} = NMk$. For the 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. (6). In Fig. 2, we plot the infidelity as a function k. We find, independent of N, the asymptotic scaling of infidelity is
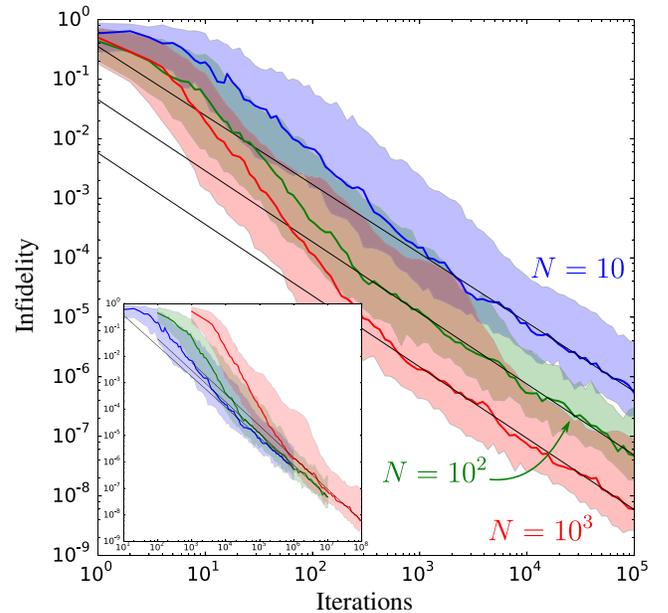
FIG. 2 (color online).   The infidelity vs the number of iterations k achieved by SGQT. Each line is the median performance of SGQT over 100 randomly (according to Haar measure) chosen pure states. The shaded regions represent the interquartile range of infidelities. The inset shows the performance as function Nk by simply shifting the original lines by their corresponding N. The SGQT parameters chosen for these simulations where $a = 3$, $A = 0$, and $b = 0.1$.

$O(1/k^\gamma)$ with $\gamma \in (1.16, 1.20)$. This is slightly better than what we would expect from the asymptotic rate of $O(1/k)$.

The inset in Fig. 2 shows that in the asymptotic regime, the performance is roughly independent N, but also shows that, in terms of the total number of experiments, fewer repetitions per measurement setting is better initially. That is, contrary to what we might expect, 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. 3, we generalize to multiple qubits. The fits to the $O(1/k^\gamma)$ scaling gave values $\gamma \in (0.80, 1.05)$, although it is difficult to trust this as an asymptotic fit for the data on 10 qubits (which gave $\gamma = 0.80$). Separately fixing k and fitting the infidelity to $O(d^\eta)$ (shown in the inset of Fig. 3) gave values $\eta \in (1.02, 1.35)$. As expected, since even estimating the fidelity to an arbitrary single pure target state is not efficient

in the number of qubits, n, the convergence of SGQT is not efficient in the number of qubits.

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

$$|\psi\rangle = \alpha_1 |10\ldots0\rangle + \alpha_2 |01\ldots0\rangle + \cdots + \alpha_n |00\ldots1\rangle. \quad (7)$$

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 a 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 of finding the "closest" state within a desirable subclass. The performance of SGQT for this problem is demonstrated in Fig. 4, where SGQT is shown to find the highest fidelity W-class state to a randomly generated mixed state. The mixed state is generated by first Haar randomly choosing a W state, then subjecting it to 5% depolarizing noise. Fits to $O(1/k^\gamma)$ asymptotic scaling for each qubit number lie in $\gamma \in (0.95, 1.03)$, in line with the optimal performance. Separately fixing k and fitting the infidelity to $O(d^\eta)$ (shown in the inset of Fig. 4) gave values $\eta \in (1.41, 1.53)$. Since $d = 2(n-1)$, this is efficient in n.

FIG. 3 (color online). The infidelity vs the number of iterations k achieved by SGQT for states of increasing qubit number (shown on the right). The number of experiments per fidelity estimate for all cases was $N = 10^4$. Each line is the median performance of SGQT over 100 randomly (according to Haar measure) chosen pure states, where the initial guess was a random perturbation of 0.01 standard deviation in each dimension of the parametrized space (this was done to more efficiently extract the asymptotic scaling). The shaded regions represent the interquartile range of infidelities. The inset shows the performance as a function of the number of qubits for fixed values of k as marked. The SGQT parameters used for these simulations (and all further simulations) were $a = 0.3$, $A = 1000$, and $b = 0.1$.

FIG. 4 (color online). The rescaled infidelity (difference between the minimum achievable infidelity and actual infidelity) vs the number of iterations k achieved by SGQT for W states of increasing qubit number (shown on the right). Parameters are as in Fig. 3.

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. 5, 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 slower than noiseless measurements with fits giving $\gamma \in (0.48, 0.59)$. However, the convergence demonstrates that SGQT is robust to both statistical and technical measurement noise.

So SGQT works, but in what sense is it more efficient than standard quantum tomography (SQT)? There are four things to consider: (1) how the infidelity scales with the total number of copies (here N_{tot}) of the system and the dimension (d); (2) the total number of different measurement directions; (3) the space complexity (number of real numbers to store); and (4) the computational time required to arrive at the final estimate.

Asymptotic arguments relate fidelity to Euclidean distance on some parameterization of pure states which transform the problem into one with a known solution from classical statistics [9]. The infidelity of estimating pure states is at least $O(d/N_{\text{tot}})$ and can be achieved by maximum likelihood estimation SQT. The extracted scaling of SGQT is slightly worse giving $O(d^\eta/N_{\text{tot}})$ with $\eta > 1$. Noting that the implementation of the SGQT algorithm was written with a single line of code and a static set of algorithmic parameters, we think it is quite good for a first
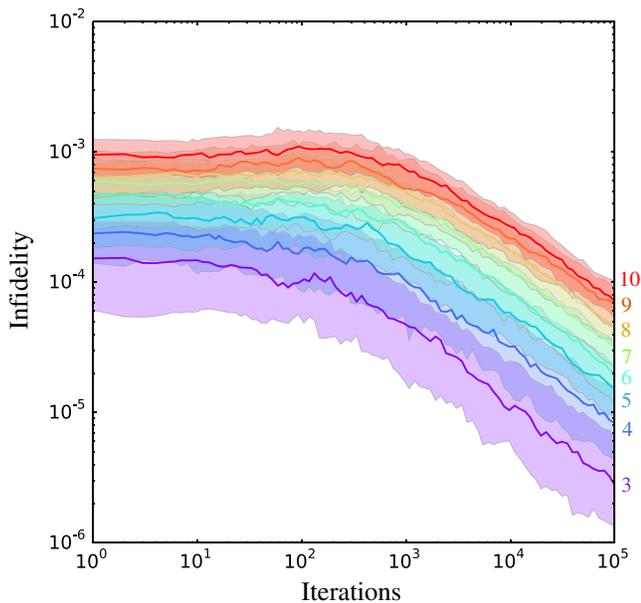
FIG. 5 (color online). The infidelity vs the number of iterations k achieved by SGQT for W states of increasing qubit number. For each fidelity estimation measurement, an unknown zero-mean, 0.1 standard deviation, Gaussian random perturbation was applied to the target state. All other parameters are as in Fig. 3.

investigation in self-learning algorithms. With more sophisticated numerical algorithms, we conjecture that self-learning approaches can achieve $O(d/N_{\text{tot}})$ and with hopefully less effort than has been put into the theoretical analysis of tomography.

The total number of measurement directions for SGQT is $O(N_{\text{tot}})$, whereas for tomography it is a free parameter m_{tot}. For informationally complete tomography $m_{\text{tot}} = O(d)$, but overcomplete measurement sets are often considered. Finally, where SGQT can provide substantial improvements in efficiency is in the computational complexity of calculating the estimator. The amount of storage required for tomography is at least $O(m_{\text{tot}}d)$, the number of elements in a measurement consisting of rank-1 elements times the space required to store each element. Since SGQT is online, it forgets the past measurements and only requires $O(d)$ storage space. To actually compute an estimator from a data set requires at least $O(m_{\text{tot}}d^2)$ time (the complexity of linear least squares regression), with more sophisticated techniques (such as maximum likelihood) taking much longer. Since SGQT ends with an estimate of the state, there is no computation needed—an enormous advantage. These considerations are summarized in Table I.

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 that 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 [10,11] 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 [12,13], such as the swap test [14].

In summary, we have provided an experimental protocol to learn quantum states without the need for classical

TABLE I. Summary of the trade-off in complexities for SGQT and SQT (standard quantum tomography). Numerical fits here give $\eta > 1$, but it is conjectured that $\eta = 1$ is achievable with optimized choices on gain sequences $\{\alpha_k\}$ and $\{\beta_k\}$.

	Infidelity	Measurements	Space	Time
SQT	$O(d/N_{\text{tot}})$	m_{tot}	$O(m_{\text{tot}}d)$	$O(m_{\text{tot}}d^2)$
SGQT	$O(d^\eta/N_{\text{tot}})$	$O(N_{\text{tot}})$	$O(d)$	$O(d)$

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.

———————————

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 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, Nat. Commun. **1**, 149 (2010).

[3] G. Tóth, W. Wieczorek, D. Gross, R. Krischek, C. Schwemmer, and H. Weinfurter, Phys. Rev. Lett. **105**, 250403 (2010).

[4] S. T. Flammia and Y.-K. Liu, Phys. Rev. Lett. **106**, 230501 (2011).

[5] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, Phys. Rev. Lett. **107**, 210404 (2011).

[6] J. C. Spall, IEEE Trans. Autom. Control **37**, 332 (1992).

[7] P. Sadegh and J. C. Spall, IEEE Trans. Autom. Control **43**, 1480 (1996).

[8] R. D. Somma, J. Chiaverini, and D. J. Berkeland, Phys. Rev. A **74**, 052302 (2006).

[9] I. Olkin and M. Sobel, Ann. Stat. **7**, 284 (1979).

[10] O. Moussa, M. P. da Silva, C. A. Ryan, and R. Laflamme, Phys. Rev. Lett. **109**, 070504 (2012).

[11] E. Magesan *et al.*, Phys. Rev. Lett. **109**, 080505 (2012).

[12] N. Wiebe, C. Granade, C. Ferrie, and D. G. Cory, Phys. Rev. Lett. **112**, 190501 (2014).

[13] N. Wiebe, C. Granade, C. Ferrie, and D. G. Cory, Phys. Rev. A **89**, 042314 (2014).

[14] H. Buhrman, R. Cleve, J. Watrous, and R. de Wolf, Phys. Rev. Lett. **87**, 167902 (2001).