Paper Review of Quantum State Tomography and Randomized Benchmarking

Kailiang Nan

1 Introduction

Quantum State Tomography (QST) and Randomized Benchmarking (RB) are pivotal techniques in the field of quantum computing and quantum information science. Quantum State Tomography is an essential method for reconstructing the quantum state of a system. In the quantum realm, where uncertainty and superposition reign, QST provides a window into the state of quantum bits (qubits). This technique is crucial for the verification and validation of quantum states, an essential step in quantum computing and quantum information processing.[1]
Randomized Benchmarking, on the other hand, is a statistical method used to assess the fidelity of quantum gates, the fundamental building blocks of quantum computers. RB stands as a cornerstone for quantifying the performance of quantum processors. It is believed to provide a way to measure the average error rates of quantum gates that are independent of state preparation and measurement errors. This is vital in a field where error rates are inherently low yet critically impactful, and where the pursuit of fault-tolerant quantum computing is paramount. This review paper aims to delve into some advancements in the theory of both randomized benchmarking and quantum tomography.[2]
We will start with introducing what are the techniques usually used in the field of Quantum State Tomography and some major challenges it has. Then, this paper will also describe what Randomized Benchmarking does. The first three reviewed papers[3][4][5] tried to address some challenges in QST while the last reviewed paper [6] clarified a common fallacy in the field of RB. This paper review summarizes the key findings in the reviewed papers and describe how they mitigate the problems and correct the misconception respectively. This review will conclude with comparing the techniques proposed by four papers and discussing the challenges that are not solved by these new techniques.

Kailiang Nan: kn171@duke.edu

2 Summary of Quantum Tomography

Acknowledgment: the next few sections refer to the paper from G. M. D'Ariano, etc[1] and the paper from the group of P. G. Kwiat[7].

2.1 History

Quantum state tomography is basically the process of quantum state determination through repeated measurements on identically prepared systems. Quantum state tomography emerged with Vogel and Risken's proposal. They introduced the idea of adapting algorithms used in medical imaging for the quantum realm. This allowed for recovering two-dimensional distributions from one-dimensional projections in various directions. The first experiments, conducted by Michael Raymer's group at the University of Oregon, demonstrated the reconstruction of coherent and squeezed states. The first exact unbiased tomographic method was proposed by G. M. D'Ariano[8] which has then been practically implemented in many experiments. Following the development of the original exact method, quantum tomography has undergone significant advancements. It has been expanded to facilitate the estimation of any observable within the field, accommodating an arbitrary number of modes. Additionally, it now applies to various quantum systems through the application of group theory. This evolution has been enriched by key improvements such as noise deconvolution techniques, the introduction of adaptive tomographic methods, and the adoption of maximum likelihood estimation (MLE) strategies. The reviewed paper "Reliable Quantum State Tomography" used exactly

the adaptations of techniques from MLE to construct the confidence region.

2.2 Steps

There are two main steps involved in the process of quantum state tomography: data acquisition and algorithmic state reconstruction. Data acquisition is basically the process we obtain data from measurement. During the data collection phase, various observables, denoted as O_i, are recorded from multiple, ideally identical, replicates of the quantum state that needs to be reconstructed. This process aims to establish the likelihood of different results for these observables. Algorithmic state reconstruction regards how we can best approximate the real state from the measurement data. Often, the precise count of the state copies being tested is unclear, and the data captured includes stochastic occurrences of all kinds of errors for each potential result. We therefore need a protocol to maximize the precision of our predicted states, and that's how state reconstruction algorithm come into play. When a complete basis is measured, these counts can be translated into probabilities. Subsequently, a tomographic algorithm utilizes these probabilities to formulate a quantum state that corresponds to the observed measurements.

2.3 Theories and Challenges

2.3.1 Maximum Likelihood Estimation

A common approach to identifying a quantum state that corresponds to the collected measurements is to pick the state that is the most probable originator of those measurements. This method is widely adopted in the practice of tomography. By formulating a function $\mathcal{L}(\rho)$ that represents the probability of a specific state ρ, the process of performing maximum likelihood tomography boils down to an optimization challenge: the objective is to determine the parameters t that configure the density matrix $\rho(t)$ in such a way that the likelihood function $\mathcal{L}(\rho(t), M)$ reaches its highest value. Here, $\rho(t)$ represents a parameterized version of the density matrix, with t indicating the set of parameters to be optimized. For example, in quantum optics, measured photon counts obey Poissonian statistics. Thus, a Poissonian distribution can be used as the basis for our likelihood. The problem is then to find a parameter t that maximize Poissonian distribution function based on the measurement data[9].

2.3.2 Bayesian Mean Estimation

The Bayesian Mean Estimation (BME) technique in quantum state tomography is based on Bayesian inference[9]. The likelihood function we have discussed in the MLE section is proportional to a related quantity in BME: the probability of a given measurement outcome occurring, conditioned on ρ being the state measured: $\Pr(M \mid \rho)$. Bayes's rule gives:

$$\Pr(\rho \mid M) = \Pr(M \mid \rho)\frac{\Pr(\rho)}{\Pr(M)}$$

where the term $\Pr(\rho \mid M)$ is known as the posterior probability, which represents the probability of the state ρ after taking the measurements M into account. On the other hand, $Pr(\rho)$ is referred to as the prior probability, indicating the assumed probability of the state ρ before any measurements are considered. The likelihood function is $\Pr(M \mid \rho)$. Lastly, $Pr(M)$ serves as a normalization constant that ensures the posterior probability forms a valid probability distribution. Then the mean state (the estimated state) is calculated by computing the integral:

$$\rho_{BME} = \int_{\rho} \rho \Pr(\rho \mid M) d\rho$$

With these concepts in mind, we can perform an iterative algorithm to improve the accuracy of the estimate as follows

1. To begin with, sample states are selected from a unbiased prior distribution, and the likelihood of these states is then determined.
2. After we sample sufficient states from this prior distribution, posterior distribution is calculated with the sampled states and associated likelihoods.
3. We denote the estimate of ρ_{BME} before sampling the current estimate of the posterior as $\rho_{previous}$. We sample states from the estimate of the posterior distribution, and calculate likelihoods to update the value of ρ_{BME}. After a pre-defined number of states are sampled from this posterior estimate, We calcu-

late the stability of ρ_{BME} by comparing its current value to the previous value $\rho_{previous}$.

4. If the stability is larger than our defined threshold, the algorithm should return ρ_{BME} as the final result. Otherwise, we repeat this process by making the new state-likelihood pairs sampled from the previous posterior estimate the new $\rho_{previous}$ and new likelihood function respectively. We repeat step 1-4 to update the posterior estimate (final result), evaluate the stability of ρ_{BME}. The algorithm is terminated only if the stability is larger than our defined threshold

2.3.3 Challenges

Two major problems exist in the previous two prevalent state reconstruction algorithm

1. Maximum likelihood estimation (MLE) can unjustifiably predict zero probabilities and often results in rank-deficient solutions on the boundary of the valid space, implying overconfidence in certain outcomes.[4] [10]
2. The quantity of measurements required to completely perform quantum state tomography on a system with multiple particles grows exponentially as the particle count increases. This rapid scaling renders the process impractical for even moderately sized systems.[5]

The reviewed papers[3][4][5] basically tried to response to these two challenges.

2.4 Paper Reviews

2.4.1 Reliable Quantum State Tomography[4]

In this article, the authors demonstrate that through quantum state tomography, combined with a suitable data analysis approach, it's possible to achieve dependable and precise error limits. These limits are defined in terms of confidence regions, a notion borrowed from classical statistics. The paper first proposes a general setup for tomography where $n + k$ particles prepared with the same Hilbert space $\mathcal{H}$ that corresponds to the degree of freedom of each particle. From these particles, a random sample of n systems is chosen and subjected to measurement using an arbitrary but complete positive operator valued measure (POVM), represented as $POVM\{B^n\}$ and $\sum B^n = \mathbb{I}^n$. The objective in quantum state tomography is to deduce the state of the remaining k systems based on the results of these measurements.

Based on this setup, the paper proposed a theory (reliable predictions from probability distribution) and a corollary (confidence regions from probability distribution). We will briefly describe them in this section.

The paper assume the measurement outcomes B^n can be processed by a data analysis routine that outputs a probability distribution μ_{B^n} on the set of mixed states, given as

$$\mu_{B^n}(\sigma)d\sigma = \frac{1}{c_{B^n}} \operatorname{tr}\left[\sigma^{\otimes n} B^n\right] d\sigma$$

Where $d\sigma$ denotes the Hilbert-Schmidt measure with $\int d\sigma = 1$ and c_{B^n} is a normalization constant. Note that $\mu_{B^n}(\sigma)d\sigma$ corresponds to the posterior probability distribution when using Bayesian Mean Estimation and $\sigma \mapsto \operatorname{tr}\left[\sigma^{\otimes n} B^n\right]$ is just the likelihood function. If we treat the process of verifying the fidelity of predictions as hypothetical tests which can be viewed as measurements yielding binary results, either "success" or "failure", this can be defined by a joint Positive Operator-Valued Measure (POVM) $\mathcal{T}_{\mu_{B^n}} = \left\{T^{\text{fail}}_{\mu_{B^n}}, 1^{\otimes k}_{\mathcal{H}} - T^{\text{fail}}_{\mu_{B^n}}\right\}$ on $\mathcal{H}^{\otimes k}$ which depends only on μ_{B^n}. Denote ρ^{n+k} as the (unknown) joint state of the n+k systems (particles) and the probability that the test $\mathcal{T}_{\mu_{B^n}}$ fails as $\operatorname{tr}(T^{\text{fail}}_{\mu_{B^n}} \rho^k_{\mu_{B^n}})$ ($\rho^k_{\mu_{B^n}}$ is the postmeasurement state after measuring n systems). We can formalize the theory of reliable predictions as follows:

Theorem For all B^n satisfying

$$\int \mu_{B^n}(\sigma) \operatorname{tr}\left[T^{\text{fail}}_{\mu_{B^n}} \sigma^{\otimes k}\right] d\sigma \leq \epsilon c^{-1}_{n+k,d},$$

For any ρ^{n+k} and $\epsilon > 0$, we have

$$\left\langle \operatorname{tr}\left[T^{\text{fail}}_{\mu_{B^n}} \rho^k_{B^n}\right] \right\rangle_{B^n} \leq \epsilon,$$

Here $c_{N,d} = \begin{pmatrix} N + d^2 - 1 \\ d^2 - 1 \end{pmatrix}$. $\langle\cdot\rangle_{B^n}$ denotes the expectation taken over all possible measurement outcomes B^n when measuring ρ^n. The theorem basically tells us, if we choose good tomography data analysis strategies which determine our probability distribution μ_{B^n} and test function $T^{\text{fail}}_{\mu_{B^n}}$, we can always get good test pass rate

that can give the good estimation of the state we are interested in. The proof of this theory is not included in the paper and is not discussed in my review.
Lastly, the paper proves a corollary that provides a sufficient criterion under which the tests are passed. The idea is based on confidence region from μ_{B^n}. A confidence region is a subset of the single-particle state space with a high probability of encompassing the 'true' state. We can define these confidence regions as follows:

$$\Gamma^{\delta}_{\mu_{B^n}} = \left\{\sigma : \exists \sigma' \in \Gamma_{\mu_{B^n}} \text{ with } F\left(\sigma, \sigma'\right)^2 \geq 1 - \delta^2\right\}$$

where $\delta^2 = \frac{2}{n}\left(\ln \frac{2}{\epsilon} + 2 \ln c_{2n,d}\right)$ and $F\left(\sigma, \sigma'\right) = \left\|\sqrt{\sigma}\sqrt{\sigma'}\right\|_1$ the fidelity. The corollary goes as follows:
Corollary For all B^n, we assume $\Gamma_{\mu_{B^n}}$ to the set of states satisfying

$$\int_{\Gamma_{\mu_{B^n}}} \mu_{B^n}(\sigma) d\sigma \geq 1 - \frac{\epsilon}{2} c^{-1}_{2n,d}.$$

Then, for any σ,

$$\text{Prob}_{B^n}\left[\sigma \in \Gamma^{\delta}_{\mu_{B^n}}\right] \geq 1 - \epsilon,$$

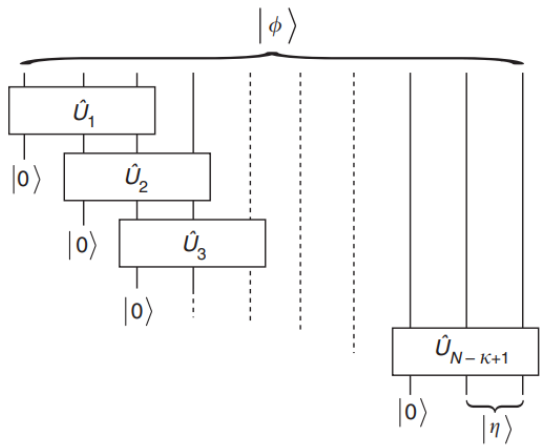
where Prob_{B^n} refers to the distribution of the measurement outcomes B^n when measuring $\sigma^{\otimes n}$. The corollary basically tells us, with reasonable choices of the measurement $\{B^n\}$, we can get good approximations (estimates) of the true states with high probability. Note that the claim is valid for all initial (true) states σ and, according to this paper, the set of confidence regions therefore does not depend on any extra assumptions about state σ and its preparation procedure. In conclusion, the paper basically proves the reliability of the process of quantum state tomography and theoretically ascertains that we can always get good enough estimations of the original states if we properly choose the observation set and prior probability distribution.

2.4.2 Efficient quantum state tomography[3]

The paper tackles the challenge of scaling quantum state tomography efficiently with system size. Traditional quantum state tomography becomes unfeasible for larger systems due to the exponential growth in the number of measurements and computational processing required. The paper aims to present tomography schemes that scale more favorably. The paper argues that the traditional way of representing quantum states is often seen as overly broad. In practical scenarios, the states encountered are usually defined by a limited set of parameters. Experts in fields like many-body physics and quantum information theory have identified several categories of states. These states are characterized by parameters that increase polynomially with system size N and closely resemble the states observed in real-world physical systems. Nevertheless, the potential application of these specific state categories in the field of quantum tomography has not been extensively explored. The paper therefore proposes two schemes that can be applied to systems that are well approximated by a matrix product states (MPS). The schemes are restricted to the case where system consists of a linear chain of N particles (or qudits) or multi-dimensional linear structures, like those found in optical lattices. In such configurations, it's highly likely that interactions between adjacent particles are significantly stronger (due to direct interaction) compared to those between distant particles (which might occur through global fluctuations in control fields). The paper focuses on a well-studied category of these states, known as finitely correlated states (FCS) or matrix product states (MPS). The two schemes are described as follows.
Scheme based on unitary transformations This work introduces a method based on unitary transformations to analyze quantum states. The process involves sequentially disentangling the left and right sides of a chain using unitary operations with a small interaction length, independent of the system size N. The result is a product state alongside a series of local unitary operations to reconstruct the original state.
The method begins with a hypothetical ideal state in the lab $\hat{\rho} = |\phi\rangle\langle\phi|$ assumed to be a Matrix Product State (MPS) with a specific bond dimension. The bond dimension here refers to the number of values that the indices connecting the matrices in the MPS representation can take. Essentially, it measures the size of the matrices that are used in the MPS representation of a quantum state. This implies a constant bound R on the rank of reductions to one part of a bipartite split (left vs. right) of the chain. The protocol starts

with an estimation of the reduced density matrix for the first $\kappa = \lceil \log_d(R) \rceil + 1$ sites , denoted as $\hat{\sigma}$. This matrix's eigendecomposition is given by $\hat{\sigma} = \Sigma_{r=1}^{R} \sigma_r \left|\phi_r\right\rangle \left\langle\phi_r\right|$.
A unitary operation $\hat{U}$ is then applied to the first κ-sites, which disentangles the first site from the rest κ - 1 sites, leading to the state $\hat{U}|\phi\rangle = \langle 0\rangle_1 \otimes |v\rangle_{2,\ldots,N}$, where $|v\rangle$ is some pure state on sites $2, \ldots, N$. This process is repeated, resulting in a sequence of unitaries $\hat{U}_1, \ldots, \hat{U}_{N-\kappa+1}$, where each $\hat{U}_i$ disentangles the ith site. This sequence transforms $|\phi\rangle$ into $|0\rangle^{\otimes N-\kappa+1} \otimes |\eta\rangle$, where $|\eta\rangle$ is some pure state on the last $\kappa - 1$ sites. The above method leads to an MPS approximation of $\hat{\rho}$. The accuracy of this approximation can be verified by monitoring truncation effects. The paper also provides a formalized method to prove the errors of MPS approximation accumulate linearly with the number of particles which is not covered in this review due to word limitation.

Figure 1: Quantum circuit to obtain matrix product state



Scheme Based on Local Measurements The previous scheme is challenging due to the requirement of unitary control of κ neighboring qudits. The paper proposes another scheme that use local measurements to construct a MPS estimate $|\Psi\rangle$ to the original state $\hat{\rho}$ and and a lower bound to the fidelity $\langle \Psi \mid \hat{\rho} \mid \Psi \rangle$ that does not require assumptions on the nature of the state $\hat{\rho}$.
The process begins with conducting a complete set of local measurements on groups of k neighbouring qudits. This results in estimates $\hat{\sigma}$ of the local reductions of $\hat{\rho} = \mathrm{tr}_{1,\cdots,i;i+k+1,\cdots.N(\hat{\rho})}$ with the condition $\|\hat{\rho}_i - \hat{\sigma}_i\|_{\mathrm{tr}} \leq \varepsilon_i$, where ε_i is a small error term.
The paper shows the inequality for fidelity bound calculation. Suppose $|\psi\rangle$ is the unique ground state of a local Hamiltonian $\hat{H} = \Sigma_i \hat{h}_i$, where the $\hat{h}_i$ is a projection operator acting only on sites $i+1, \ldots, i+k$. The fidelity between $|\psi\rangle$ and $\hat{\rho}$ is bounded by $1 - \frac{1}{\Delta E}\left(\sum_i \mathrm{tr}\left[\hat{h}_i \hat{\sigma}_i\right] + \varepsilon_i\right)$, where ΔE is the energy gap above the ground state. This bound gets tighter if the experimental estimates $\hat{\rho}$ align closely with the reductions of the state $\hat{\sigma}_i$, i.e. ε_i is small. The proof goes as follows:
Expanding the Hamiltonian $\hat{H}$ in the eigenbasis $\hat{H} = \Sigma_{n=0}^{2^N-1} E_n \left|E_n\right\rangle \left\langle E_n\right|$, we have

$$\mathrm{tr}[\hat{H}\hat{\rho}] = E_n \sum_{n>0} \langle E_n|\hat{\rho}|E_n\rangle \geq$$
$$\Delta E \sum_{n>0} \langle E_n|\hat{\rho}|E_n\rangle = \Delta E(1 - \langle\psi|\hat{\rho}|\psi\rangle)$$
$$\implies \langle\psi|\hat{\rho}|\psi\rangle \geq 1 - \frac{\sum_i \mathrm{tr}\left[h_i\hat{\rho}\right]}{\Delta E} \geq$$
$$1 - \frac{1}{\Delta E}\left(\sum_i \mathrm{tr}\left[\hat{h}_i\hat{\sigma}_i\right] + \varepsilon_i\right)$$

The above proof tells us the next step is to identify a local gapped Hamiltonian $\hat{H}$ whose ground state reductions are close to $\hat{\sigma}_i$. While it may not be immediately clear whether such a Hamiltonian exists or how to find it efficiently, formal methods suggest that a suitable witness Hamiltonian always exists for a sufficiently large k ($k \geq 2\lceil \log_2 R \rceil + 1$, where R is the bound of the rank on the reduced states). This Hamiltonian can be constructed using an algorithm, and its properties, including the energy gap, can be efficiently computed. The paper also devises an algorithm that can efficiently compute the quantities appeared in the scheme. This review will not include the description and explanation of the algorithm due to word limitation.
In summary, in order to address the problem of exponentially growth of parameters with system size N,the paper Efficient quantum state tomography[3] proposes two schemes to reduce the parameters needed to characterize a less general state $\hat{\rho}$. The first scheme is about finding a sequence of disentangling unitary operators such that, after applying them, the result state is an MPS state which is a good approximation (estimation) of the original state. The second scheme involves experimental measurements to estimate the local reductions of a quantum state and clas-

sical post-processing to find a lower bound for the fidelity of this estimation. The key lies in constructing a local Hamiltonian whose ground state closely aligns with the estimated reductions of the quantum state in question.

2.4.3 Quantum State Tomography via Compressed Sensing[5]

Compressed sensing is a technique commonly used in underdetermined scenario of signal processing. It can reconstruct the information with less measurements (samplings) if the original signal is sparse in some basis. Similarly, the key idea of using compressed sensing in quantum state tomography is to reconstruct the original state from less measurements than normally required, therefore reducing the exponentially growing complexity of QST to a great extent.

The paper first proposes a method named "Matrix recovery using Pauli measurements". It then proves this method can work well on low-rank density matrix and is resilient against noise. The details of the method and proofs are as follows:

Matrix recovery using Pauli measurements In the system comprising n spin- $1/2$ sites in an uncharacterized state ρ. This state is investigated using a specific kind of n-qubit Pauli matrix, denoted by $w = \bigotimes_{i=1}^{n} w_i$, where $w_i \in \{\mathbb{1}, \sigma^x, \sigma^y, \sigma^z\}$. Given that there are d^2 distinct matrices of this form, labeled as $w(a), a \in [1, d^2]$, the method involves randomly selecting m integers from range $[1, d^2]$ and measuring the expectation values $\operatorname{tr} \rho w(A_i)$ for these matrices. The core of this technique is a convex optimization problem aimed at minimizing the trace norm $\|\sigma\|_{\mathrm{tr}}$ subject to

$$\operatorname{tr} \sigma = 1, \quad \operatorname{tr} w(A_i)\, \sigma = \operatorname{tr} w(A_i)\, \rho.$$

Low-rank tomography The theorem: if one knows m randomly chosen Pauli expectation values, where m is determined by the formula $m = cdr \log^2 d$ with c being a constant, d the dimension of the Hilbert space, and r the rank of ρ. Then it is possible to uniquely reconstruct the state ρ by solving the previously mentioned convex optimization problem with probability of failure exponentially small in c.

To prove the above theorem, the paper first defines the "sampling operator" $\mathcal{R} : \rho \mapsto \frac{d}{m} \sum_{i=1}^{m} w(A_i) \operatorname{tr} \rho w(A_i)$ and rewrites the optimization problem above as $\min_\sigma \|\sigma\|_{\mathrm{tr}}$ s.t. $\mathcal{R}\sigma = \mathcal{R}\rho$. If the deviations $\Delta := \sigma - \rho$ satisfying $\mathcal{R}\Delta \neq 0$ or $\|\rho + \Delta\|_{\mathrm{tr}} > \|\rho\|_{\mathrm{tr}}$, then the uniqueness of solution is proved (**uniqueness criteria**). The key point of the following proof steps revolves around constructing a strict subgradient Y for the norm (A matrix Y is a strict subgradient if $\|\rho + \Delta\|_{\mathrm{tr}} > \|\rho\|_{\mathrm{tr}} + \operatorname{tr} Y\Delta$ for all $\Delta \neq 0$). The proof decomposes the deviation $\Delta = \Delta_T + \Delta_T^{\perp}$ into components within and orthogonal to a certain subspace T related to ρ leading to two distinct cases for analysis. The first case deals with a scenario where Δ is largely within the subspace T. Here, the focus is on showing the invertibility of a restricted version of the sampling operator $\mathcal{A}$, with probabilities involved in the analysis derived from noncommutative large-deviation bounds. The second case is a more complex scenario where the proof must establish the existence of an almost subgradient. This involves detailed calculations to show that $\mathcal{R}\Delta = 0$ implies the uniqueness condition. A recursive method, referred to as the "golfing scheme", is also employed to converge towards a subgradient. This process involves multiple iterations of drawing batches of Pauli observables and adjusting the operators and projections accordingly. The proof concludes by calculating the total probability of failure p_f, ensuring it is acceptably low.

Robustness to noise Real-world situations might differ from ideal ones in two main ways: the true state ρ_t may not be low rank but only closely approximated by a low-rank state ρ, and due to noise, the estimated Pauli expectations may not be precise. The true state ρ_t is approximated within error bars ε_1 by a state ρ and the Pauli expectations errors are given by $\|\mathcal{R}\omega - \mathcal{R}\rho_t\|_2 \leq \varepsilon_2$. To estimate ρ_t, a modified convex program is solved to minimize the trace norm of σ,

$$\min \|\sigma\|_{\mathrm{tr}}, \quad \text{subject to } \|\mathcal{R}\sigma - \mathcal{R}\omega\|_2 \leq \varepsilon.$$

The paper then gives an observation without proofs. The first observation is, if Pauli expectations are known up to an error of ε, the difference between the solution $\sigma^\star$ and the true state ρ_t is smaller than $O(\varepsilon\sqrt{rd})$ with a specific probability of failure.

Certified Tomography of Almost Pure States One challenge of the aforementioned scheme is the problem of unknown parameters. When conducting tomography on an unknown state, neither the

rank r nor the closeness δ_1 to a low-rank state are known in advance. The paper gives an estimation approach. The parameters r and σ_1 can be estimated from physical parameters or from the same data used for reconstruction. This is particularly effective for almost pure states (where $r = 1$) as δ_1 relates to the state's purity. The paper also provides another observation which states, for states close to being pure, it's possible to certify this assumption and reconstruct the state with explicit error bounds using a polynomial number of Pauli expectation values.
The paper concludes with presenting numerically simulation results on both the random Pauli and hybrid approaches. The results show the methods provided in the paper can reconstruct the original state with high fidelity by exploiting the sparsity of the original state, which also significantly reduce the complexity of quantum state tomography.

3 Summary of Randomized Benchmarking

3.1 Introduction

Randomized benchmarking is a technique used in quantum computing to measure the average error rate of quantum gates, which are the basic operations used to manipulate qubits[11]. A significant obstacle in developing quantum information processors is the impracticality of fully characterizing the noise in a quantum system through process tomography, especially as the system size increases. Thoroughly understanding the noise is essential as it enables the identification of effective error-correction strategies, which are crucial for the dependable transmission of quantum information. However, given that complete process tomography becomes unmanageable for larger quantum systems, there is an increasing focus on finding scalable techniques that can partially characterize the noise impacting a quantum system[12]. Randomized benchmarking serve as one such method that is widely used in both industry and academia.
For example, in quantum computing experiments, a crucial task is to implement gates that consistently exhibit low error rates, regardless of their application context[12]. Randomization serves as a method for analyzing specific aspects of quantum gate errors in this context. This technique generally involves executing a series of random unitary operators U, followed by their corresponding inverse operations U^{-1}. Under the assumption that the noise model remains unaffected by the quantum operations, the role of this randomization is to depolarize the noise. Consequently, the mean fidelity of the process, when applied to a pure initial state, is believed to be equivalent to the average fidelity of the noise operation across all pure states[12]. In the reviewed paper "What Randomized Benchmarking Actually Measures"[6], the authors of the paper refute the theory that randomized benchmarking is a measure of average gate fidelity (or infidelity).

3.2 Common Steps

There exists more than one variants of how to do randomized benchmarking. The steps listed here are described by the paper being reviewed[6].

1. Execute various sequences of Clifford gates, each designed to revert the Quantum Information Processor (QIP) back to its original state. (U followed by U^{-1})
2. Conduct measurements after each sequence to determine if the QIP has successfully returned to its initial state, a status referred to as "surviving".
3. Chart the observed probabilities of survival against the length of each gate sequence and modeling these observations with an exponential decay function. The rate at which the survival probability diminishes, adjusted for a dimensionality constant and disregarding any errors due to finite sampling, is called the "RB number" r. The purpose of RB (Randomized Benchmarking) experiments is to calculate this r value, which then serves as an indicator of the quantum processor's performance.

3.3 Paper Reviews

3.3.1 What Randomized Benchmarking Actually Measures[6]

The paper challenges the conventional belief that the RB number (r) approximates the average gate set infidelity (AGsI), indicating significant discrepancies between them. It also introduces a

new, more accurate theory for RB decay that is representation-independent. This new approach shows that the RB decay curve is always exponential for small errors describable by process matrices, offering a more reliable measure of physical gate sets' error rates.

In the basic Randomized Benchmarking (RB) protocol for n-qubit Clifford gates, where $n \geq 1$, the method estimates the parameter r from RB data by fitting the average of the sampled survival probabilities (P_m) for various sequence lengths m. This fitting uses the model $P_m = A + (B + Cm)p^m$, where A, B, C, and p are fit parameters determined through fitting. The estimated value of p, denoted as $\hat{p}$, is used to calculate r as $\hat{r} = (d-1)(1-\hat{p})/d$, where $d = 2^n$. The paper points out,as long as the P_m decay exponentially with m, r and the average gate set infidelity (AGsI) ϵ are equally defined. Based on this setting, the paper tries to answer the following two questions

1. Under what circumstances does r equals to the average gate set infidelity (AGsI) ε?
2. To what property of the imperfect gates does r correspond?

To answer the above two questions, the paper emphasizes that the theory of RB is built upon process matrix G_i, state preparations ρ_j, and POVM operators $\mathcal{M}_k = \{E_{k,l}\}$ and imperfect physical gate set can be represented by some $\tilde{\mathcal{G}} = \left\{\tilde{G}_i, \tilde{\rho}_j, \tilde{E}_{k,l}\right\}$, and an idealized perfect device by some $\mathcal{G} = \{G_i, \rho_j, E_{k,l}\}$. The paper then contrasts gate-independent errors, where a single process matrix Λ can represent the imperfection in all gates and $\tilde{C}_i = \Lambda C_i$, with gate-dependent errors, which have individual error maps for each gate Λ_i and $\tilde{C}_i = \Lambda_i C_i$. It points out r is exactly equal to the AGI ε for the gate-independent errors, but the same thing is not true for gate-dependent errors even if it is the "weakly gate-dependent" errors which is believed to have an r equals to ε in lots of literature. The paper then shows that r and ε can differ significantly for realistic weakly gate-dependent noise models through a 1-qubit simulation example. It attributes this discrepancy to the fact that while r is a physical property of the gates and independent of their representation, ε is not. Two representations of a physical gate are called equivalent if they cannot be distinguished by any experiment. The discrepancy arises from the fact that the AGsI ε is not gauge invariant, the value of which can vary from representations to representations. In order to fix this problem, AGsI should be defined as a gauge-invariant property. The paper states this can be achieved by comparing $\tilde{\mathcal{C}}$ not to the usual fixed representation of the Clifford gates $\mathcal{C}$, but to a $\mathcal{C}$-dependent representation of them,$\mathcal{C}_{\tilde{\mathcal{C}}}$. This answers the first question we mentioned above. r can be a representative of AGsI ε only if they are both gate/gauge-independent.

In order to answer the second question, the paper brings up a new interpretive theory for the RB decay curve that associates r with a calculable, representation-independent property of the physical gate set. The theory begins with the formula for calculating the average survival probability over all RB sequences of length m. This probability P_m is a function of the sum of traces of a sequence of operations involving the imperfect Clifford gates $\left\{\tilde{C}_i\right\}$, the state preparation ρ, and the measurement E. The formula encompasses all possible sequences s of length m in the Clifford group $\mathcal{C}$.

$$P_m = \frac{1}{|\mathcal{C}|^m} \sum_s \mathrm{Tr}\left[E\tilde{C}_{s^{-1}}\tilde{C}_{s_m}\ldots\tilde{C}_{s_1}(\rho)\right],$$

It then defines map $\mathcal{S}_m$ which is the average over all sequences s of the combined operation of the imperfect Clifford gates.

$$\mathcal{S}_m = \mathrm{avg}_s\left[\tilde{C}_{s^{-1}}\tilde{C}_{s_m}\ldots\tilde{C}_{s_1}\right] = |\mathcal{C}|\vec{v}^T\mathcal{R}^{m+1}\vec{v}$$

The matrix $\mathcal{R}$ is central to this theory. It is formed by the elements $\tilde{C}_{j\to k} = \tilde{C}_j^{-1}\tilde{C}_k$, and

$$\mathcal{R} = \frac{1}{|\mathcal{C}|}\begin{pmatrix} \tilde{C}_{1\to 1} & \tilde{C}_{2\to 1} & \cdots & \tilde{C}_{|\mathcal{C}|\to 1} \\ \tilde{C}_{1\to 2} & \tilde{C}_{2\to 2} & \cdots & \tilde{C}_{|\mathcal{C}|\to 2} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{C}_{1\to|\mathcal{C}|} & \tilde{C}_{2\to|\mathcal{C}|} & \cdots & \tilde{C}_{|\mathcal{C}|\to|\mathcal{C}|} \end{pmatrix},$$

The theory then relates P_m to the trace of a function involving $\mathcal{R}^{m+1}$: $P_m = |\mathcal{C}|\,\mathrm{Tr}\left[E\left(\vec{v}^T\mathcal{R}^{m+1}\vec{v}\right)(\rho)\right]$, which can be simplified with eigenvalues of $\mathcal{R}$

$$P_m = \sum_i \alpha_i \lambda_i^{m+1}$$

where $\{\lambda_i\}$ are the $4^n|\mathcal{C}|$ eigenvalues of $\mathcal{R}, n$ is the number of qubits, and $\{\alpha_i\}$ are constants de-

pending on ρ, E, and the eigenvectors of $\mathcal{R}$. This means the actual physical quantity the RB number r represents is a joint function of ρ, E, eigenvalues and eigenvectors of $\mathcal{R}$.
The paper also verifies the new theory is consistent with the fact that decays with a functional form of $A + Bp_m$ are normally observed in practice. It shows the new theory overlaps with the old one when errors are gate-independent. The paper also refers to Wallman's work[13], and claims there exists at least one such a representation of the imperfect gates for which $r = \varepsilon$ despite the gates in this representation are not generally completely positive.
In summary, this new theory of RB provides a more nuanced and accurate method for calculating the RB decay, which is crucial for assessing the fidelity and reliability of quantum gates in quantum computing and information processing. It emphasizes a representation-independent approach, focusing on the physical properties of the gate set rather than specific models or representations.

4 Discussion and Conclusion

The first paper "Reliable Quantum State Tomography" essentially validates the efficacy of quantum state tomography, demonstrating that accurate estimations of original quantum states can be consistently achieved by carefully selecting the observation set and prior probability distribution. The second and the third paper attempt to mitigate the problem of exponentially growing complexity of quantum tomography. The paper "Efficient quantum state tomography" develops two schemes. The first one can approximate states and convert them into the form of Matrix Product State (MPS) which can be characterized with far less parameters. The second one requires us to find a gapped local Hamiltonian whose ground state closely aligns with the estimated reductions of the quantum state we desired. The paper "Quantum State Tomography via Compressed Sensing" applies the technique of compressed sensing to quantum state tomography such that we can reconstruct the original state with less parameters than normally needed.
The last paper "What Randomized Benchmarking Actually Measures" points out a common misunderstanding on what randomized benchmarking number represents. It proposes a new theory to better characterize the benchmark number r and also clarifies that r will equal to the average gate set infidelity ε only if we choose the right representation of gate errors.
In general, They all try to optimize or solve some outstanding problems in the original techniques, and their experimental (simulation) results show that their attempts have alleviated the original problems to some extent.

4.1 Challenges

Despite the effort, the challenges remain in the field of quantum state tomography and randomized benchmarking.

1. In terms of quantum state tomography, the complexity reduction techniques proposed by these papers only apply to some specific states. These states usually have some distinct structures in their parameters such as sparsity or one-dimensional configuration. These states only account for a small proportion of the states we are interested in. A more general protocol is needed to further optimize the technique of quantum state tomography.
2. In terms of randomized benchmarking, the paper "What Randomized Benchmarking Actually Measures" proposes a new theory to explain what quantity RB actually represents and does dispel our misconceptions about it. However, current theories for various Randomized Benchmarking (RB) protocols, like interleaved, dihedral, and unitarity RB, are based on methods that depend on representation-dependent techniques. We want to ask, similar to the approach adopted in this study, whether these methods can be converted into a format that is representation-independent. If we can do this, we can improve the stability of these existing RB protocols.

References

[1] G Mauro D'Ariano, Matteo GA Paris, and Massimiliano F Sacchi. "Quantum tomography". Advances in imaging and electron physics **128**, 206–309 (2003).

arXiv:https://arxiv.org/abs/quant-ph/0302028.

[2] Emanuel Knill, Dietrich Leibfried, Rolf Reichle, Joe Britton, R Brad Blakestad, John D Jost, Chris Langer, Roee Ozeri, Signe Seidelin, and David J Wineland. "Randomized benchmarking of quantum gates". Physical Review A **77**, 012307 (2008).

[3] Marcus Cramer, Martin B Plenio, Steven T Flammia, Rolando Somma, David Gross, Stephen D Bartlett, Olivier Landon-Cardinal, David Poulin, and Yi-Kai Liu. "Efficient quantum state tomography". Nature communications **1**, 149 (2010). arXiv:https://arxiv.org/abs/1101.4366.

[4] Matthias Christandl and Renato Renner. "Reliable quantum state tomography". Physical Review Letters **109**, 120403 (2012). arXiv:https://arxiv.org/abs/1108.5329.

[5] David Gross, Yi-Kai Liu, Steven T Flammia, Stephen Becker, and Jens Eisert. "Quantum state tomography via compressed sensing". Physical review letters **105**, 150401 (2010). arXiv:https://arxiv.org/abs/0909.3304.

[6] Timothy Proctor, Kenneth Rudinger, Kevin Young, Mohan Sarovar, and Robin Blume-Kohout. "What randomized benchmarking actually measures". Physical review letters **119**, 130502 (2017). arXiv:https://arxiv.org/abs/1702.01853.

[7] Joseph B. Altepeter, Daniel F.V. James, and Paul G. Kwiat. "4 qubit quantum state tomography". Pages 113–145. Springer Berlin Heidelberg. Berlin, Heidelberg (2004).

[8] G. M. D'Ariano, C. Macchiavello, and M. G. A. Paris. "Detection of the density matrix through optical homodyne tomography without filtered back projection". Phys. Rev. A **50**, 4298–4302 (1994).

[9] Chris Zeitler. "Characterization of quantum states: advances in quantum tomography and tests of nonlocality". PhD thesis. Illinois U., Urbana. (2021). url: `https://inspirehep.net/literature/2101348`.

[10] Robin Blume-Kohout. "Optimal, reliable estimation of quantum states". New Journal of Physics **12**, 043034 (2010).

[11] Easwar Magesan, Jay M. Gambetta, and Joseph Emerson. "Characterizing quantum gates via randomized benchmarking". Phys. Rev. A **85**, 042311 (2012).

[12] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. "Randomized benchmarking of quantum gates". Phys. Rev. A **77**, 012307 (2008).

[13] Joel J. Wallman. "Randomized benchmarking with gate-dependent noise". Quantum **2**, 47 (2018).