PAPER

Adaptive tomography of qubits: purity versus statistical fluctuations

View the article online for updates and enhancements.

IOP ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection−download the first chapter of every title for free.

Adaptive tomography of qubits: purity versus statistical fluctuations

Aonan Zhang[1,2] , **Yujie Zhang**[3], **Feixiang Xu**[1,2], **Long Li**[2,4] **and Lijian Zhang**[1,2]

[1] National Laboratory of Solid State Microstructures and College of Engineering and Applied Sciences, Nanjing University, Nanjing 210093, People's Republic of China
[2] Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, People's Republic of China
[3] Kuang Yaming Honors School, Nanjing University, Nanjing 210093, People's Republic of China
[4] School of Physics, Nanjing University, Nanjing 210093, People's Republic of China

E-mail: aonan.zhang@smail.nju.edu.cn and lijian.zhang@nju.edu.cn

Abstract

The success of quantum information processing applications relies on accurate and efficient characterization of quantum states, especially nearly-pure states. In this work, we investigate a procedure for adaptive qubit state tomography which achieves $O(1/N)$ scaling in accuracy for large N. We analyze the performance of the adaptive protocol on the characterization of pure, mixed and nearly-pure states, and clarify the interplay between the purity of the state and the statistical fluctuations of the measurement results. Our results highlight the subtle difference between the characterization of nearly-pure states and those of pure or highly mixed states.

(Some figures may appear in colour only in the online journal)

1. Introduction

Implementation of quantum information processing (QIP) relies on accurate characterization and manipulation of quantum states. Therefore the estimation of an unknown quantum state is of fundamental importance in quantum information theory [1, 2]. Quantum state tomography (QST) is a process of identifying the density matrix ρ of an unknown quantum state via measurements on N identical copies of the state. Due to the statistical fluctuations of the measurement results, there would be a statistical distance between the estimate $\hat{\rho}$ and the true state ρ.

Tomographic protocols aim at minimizing the statistical distance with limited resources [3–5], while a variety of metrics have been used to quantify the statistical distance [6]. One of the well-motivated metrics is the *infidelity*, which is defined as

$$1 - F(\rho, \hat{\rho}) = 1 - \left[\text{Tr} \left(\sqrt{\sqrt{\rho}\hat{\rho}\sqrt{\rho}} \right) \right]^2.$$

(1)

Strictly speaking, estimating the ' *unknown state*' means that there is no priori information, i.e. the probability distribution for all possible states within the Hilbert space is homogeneous before the estimation procedure. In practice, priori knowledge about the quantum state can be given in various ways, such as that the state is pure, or that the priori probability distribution of the state is not uniform [7, 8]. Such priori information may help to design accurate and efficient QST protocols.

How to gain priori knowledge about an unknown state? One simple idea is to perform pre-measurements of the state and acquire a partial characterization. Such procedure is known as the adaptive quantum state tomography, which has been investigated in a host of theoretical and experimental works [9–17]. Adaptive tomography has the advantages of improving the characterization accuracy and being less sensitive to systematic errors than standard tomography. Recently, a simple protocol for adaptive tomography based on mutually unbiased bases (MUB) has been proposed [10], and improves the estimation accuracy, which is quantified by the infidelity, from $O(1/\sqrt{N})$ to $O(1/N)$ for pure states. Furthermore, adaptive tomography using Bayesian estimation has been demonstrated and achieves $O(1/N)$ scaling in the estimation accuracy for pure states [11].

However, in practical setups the noise is unavoidable in the preparation of a quantum state and measurement apparatus, resulting in nearly-pure states rather than pure states. Therefore nearly-pure states are more practical cases in physical implementations of QIP applications. To date, the performance of adaptive tomography on the nearly-pure states and the reason for such performance remain to be elucidated even for the single-qubit scenario. In particular, how the purity of the state affects the performance has not been clarified entirely. In this work, we address the problem of estimating the performance of adaptive tomography where the state with moderate mixture is typical. We study the adaptive qubit tomography with two quintessential measurements: MUB and the symmetric informationally complete positive operator-valued measure (SIC-POVM), and analyze their performance on pure, highly mixed and nearly-pure states. It is shown that the interplay between the purity of the state and the statistical fluctuations affects the scaling of the estimation accuracy, and defines the transition region from pure to mixed states, i.e. the nearly-pure states.

The rest of the paper is organized as follows. In section 2 adaptive quantum state tomography using MUB and SIC-POVM are introduced, including the optimal configuration, the maximum likelihood estimation (MLE), and the adaptive protocols. Then we analyze the performance of QST for pure and highly mixed states. Special attention is paid to nearly-pure states in section 3, in which the purity of the state and the statistical fluctuations in the measurement results are comparable and affect the performance of tomography. Finally, we discuss the criterion of nearly-pure states in QST.

2. Adaptive quantum state tomography

2.1. Tomography using MUB and SIC-POVM

A quantum state reveals its information through measurements [5]. Optimal measurement sets for QST based on various figure of merits have been investigated [18–23]. MUB are optimal

fixed measurements for von Neumann measurements [24] and widely used in standard state tomography. Two orthogonal bases $\{|u_1\rangle, ..., |u_d\rangle\}$ and $\{|v_1\rangle, ..., |v_d\rangle\}$ are mutually unbiased if

$$|\langle u_i|v_j\rangle|^2 = 1/d \quad \forall i,j. \tag{2}$$

We should indicate that the MUB measurements are projective measurements. For a d-dimensional system, a quantum state is specified by $d^2 - 1$ real parameters, therefore MUB measurements need at least $(d^2 - 1)/(d - 1) = d + 1$ orthogonal measurements. If $d = 2$, MUB measurements require three orthogonal measurements to identify the density matrix:

$$\rho = \frac{1}{2}\left(\mathbb{1} + \vec{s}\cdot\vec{\sigma}\right), \tag{3}$$

where $\vec{s} = (s_x, s_y, s_z)$ is the Pauli vector and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is the tensor of Pauli operators. Complete sets of MUB have been constructed in Hilbert spaces whose dimensions are any power of prime, but not scalable to arbitrary high dimensions [25, 26], even as low as $d = 6$ [27].

On the other side, positive-operator valued measures (POVM) allow one to gain more information from a single measurement than the standard projective measurements. In quantum measurement theory, a POVM can be defined as a set of non-negative operators $\{E_i\}$ which fulfills the completeness condition $\sum_i E_i = \mathbb{1}$. The POVM is informationally complete when every state is determined completely by the measurement statistics, and more useful when the POVM is symmetric when all pairwise inner products between the POVM elements are equal [19, 25]. In mathematical viewpoint, MUB and SIC-POVM for a qubit are both measurement sets as platonic solids in Bloch sphere [22, 28], as shown in figure 1.

SIC measurements are of vital importance in quantum information theory, especially in quantum tomography [29–32] and quantum communication [33–36]. State tomography with SIC-POVM has been experimentally demonstrated for qubits [30], qutrits [31] and qudits [32]. In contrast to the MUB, SIC-POVM minimizes the amount of measurements, and has been conjectured to exist in arbitrary dimensions [19, 32, 37, 38]. SIC-POVM can be implemented with a single measurement setting [39–42], while MUB usually changes the configuration to realize multiple measurement sets and divides the ensemble of states into subgroups. In the following, we will focus on the SIC-POVM because previous works have not given a detailed analysis of adaptive tomography using SIC-POVM, while we also give the results for MUB. For a d-dimensional quantum system, SIC-POVM is described by a set of projectors $P_i = |\psi_i\rangle\langle\psi_i|$ $(i = 0, 1, 2...d^2 - 1)$, which satisfy

$$\text{Tr}\left(P_iP_j\right) = \frac{1}{d + 1} \text{ (for } i \neq j), \quad E_i = \frac{1}{d}P_i. \tag{4}$$

The projectors P_i at Bloch sphere are sketched in figure 1(b). Attributed to the constraint of completeness, all possible measurement results are limited to the interior of the tangential tetrahedron of the Bloch sphere. Geometrically, the four tangent points are opposite to the four projectors. In accordance with Born's rule, the probability of observing outcome i is $p_i = \text{Tr}(E_i\rho)$, which can be represented by the Pauli parameters
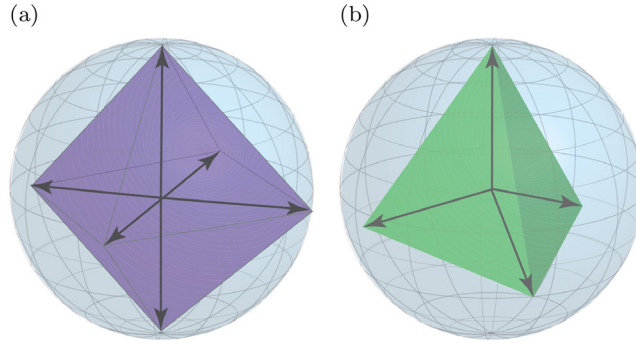
Figure 1. Projections of MUB and SIC-POVM in Bloch sphere ($d = 2$). (a) The six measurement outcomes of MUB constitute the vertices of an octahedron at Bloch sphere. (b) The four measurement outcomes of SIC-POVM constitute the vertices of a tetrahedron.

$$p_0 = \frac{1}{4}\left(1 + s_z\right),$$

$$p_1 = \frac{1}{4}\left(1 + \frac{2\sqrt{2}}{3}s_x - \frac{1}{3}s_z\right),$$

$$p_2 = \frac{1}{4}\left(1 - \frac{\sqrt{2}}{3}s_x + \frac{\sqrt{6}}{3}s_y - \frac{1}{3}s_z\right),$$

$$p_3 = \frac{1}{4}\left(1 - \frac{\sqrt{2}}{3}s_x - \frac{\sqrt{6}}{3}s_y - \frac{1}{3}s_z\right). \tag{5}$$

The density matrix ρ can be parameterized through probabilities $\vec{p} = (p_0, p_1, p_2, p_3)$ as $\rho = 6\vec{p} \cdot \vec{E} - 1$, which yields

$$\rho = \begin{pmatrix} 2p_0 & \sqrt{2}\left(p_1 + p_2 e^{-i2\pi/3} + p_3 e^{i2\pi/3}\right) \\ \sqrt{2}\left(p_1 + p_2 e^{i2\pi/3} + p_3 e^{-i2\pi/3}\right) & 1 - 2p_0 \end{pmatrix}. \tag{6}$$

2.2. Tomography accuracy and optimal strategy

In the single-qubit tomography scenario, the outcomes of SIC-POVM subject to a multinomial distribution $\mathcal{M}(N, (p_0, p_1, p_2, p_3))$, of which the probability mass function is

$$f(n_i; N; p_i) = \frac{N}{\prod_i n_i!} \prod_i p_i^{n_i}. \tag{7}$$

If we focus on the estimation of a specific parameter p_i, n_i becomes a Bernoulli trial $\sim B(N, p_i)$ with a binomial distribution. The expectation and standard deviation of estimator $\hat{p}_i$ are

$$\langle \hat{p}_i \rangle = p_i, \quad \Delta \hat{p}_i = \sqrt{\frac{1}{N}} \sqrt{p_i(1 - p_i)}. \tag{8}$$

The estimated state is subject to such statistical fluctuations and the physical constraints. As an example, for the parameter p_0 in equation (5), we can substitute s_z (equation (5)) into equation (8), and yield

$$\langle \hat{s}_z \rangle = s_z, \quad \Delta \hat{s}_z = \sqrt{\frac{4}{N}} \sqrt{1 - \left(\frac{1 - s_z}{2} \right)^2}. \tag{9}$$

In general cases, $\Delta \hat{s}_x, \Delta \hat{s}_y$ are $O(1/\sqrt{N})$ as well. Meanwhile, the infidelity defined in equation (1) can also be given in terms of Pauli vector representation [43]

$$1 - F(\vec{s}, \vec{s}_{\text{est}}) = \frac{1}{2} \left(1 - \vec{s} \cdot \vec{s}_{\text{est}} - \sqrt{1 - s^2} \sqrt{1 - s_{\text{est}}^2} \right), \tag{10}$$

where $\vec{s}_{\text{est}}$ is the Pauli vector of the estimate state and $s = |\vec{s}|$ ($s_{\text{est}} = |\vec{s}_{\text{est}}|$) denotes the length of $\vec{s}$ ($\vec{s}_{\text{est}}$). The discrepancy between the true state $\vec{s}$ and the estimate state $\vec{s}_{\text{est}}$ is denoted as $\delta \vec{s} = \vec{s} - \vec{s}_{\text{est}}$, and $|\delta \vec{s}| \propto O(1/\sqrt{N})$ since $\Delta \hat{s}_x, \Delta \hat{s}_y, \Delta \hat{s}_z \propto O(1/\sqrt{N})$. For a typical mixed state, the average infidelity can be estimated under the approximation $|\delta \vec{s}| \ll 1 - s$ [44]

$$1 - F \approx \frac{1}{4} |\delta \vec{s}|^2 \propto O(1/N). \tag{11}$$

However, if the true state is (or approaches) a general pure state, $|\vec{s}|$ equals (or approaches) 1. The third term on the right side of equation (10) vanishes, and the infidelity can be inferred by

$$\begin{aligned} 1 - F &= \frac{1}{2} \left(1 - \vec{s} \cdot \vec{s}_{\text{est}} \right) \\ &= \frac{1}{2} \left(1 - \vec{s} \cdot \vec{s} \right) + \frac{1}{2} \vec{s} \cdot \delta \vec{s} \propto O(1/\sqrt{N}). \end{aligned} \tag{12}$$

The $O(1/\sqrt{N})$ term dominates in the average of equation (12) over the distribution of the measurement results, therefore the average infidelity scales as $O(1/\sqrt{N})$, as shown in figure 2(a). However, not all of the pure states behave as $O(1/\sqrt{N})$ scaling. When the measurement is aligned in several specific configurations, the tomography accuracy would be significantly improved. Such configurations are called the optimal strategy of QST. The optimal strategy for MUB has been investigated in previous works [10]. If one basis diagonalizes the true state ρ, the infidelity would be minimized. For SIC-POVM, the optimal strategy is the antiparallel strategy from both the tomographic purpose [45] and the viewpoint of informational power [33, 46]. In equation (9), when $s_z \approx -1$ (i.e. $p_0 \approx 0$), the uncertainty in the estimation $\hat{s}_z$ (i.e. $\hat{p}_0$) is reduced to 0. The true state is aligned antiparallel to $|\psi_0\rangle$ in Bloch sphere. Similarly, if the state is antiparallel to any $|\psi_i\rangle$, we will have similar results.

Figure 2(b) illustrates the scenario of estimating a pure state in the optimal configuration. In this case, one of the outcomes is not registered at all; the distribution of the measurement results is compressed into a plane without statistical error in s_z direction. The statistical fluctuations on s_x and s_y are $O(1/\sqrt{N})$. Meanwhile, because most of trials violate the physical constraint, the estimation $\vec{s}_{\text{est}}$, as well as the true state $\vec{s}$, is a pure state. In consequence, we can estimate the infidelity

$$\begin{aligned} 1 - F &= \frac{1}{2} \left(1 - \vec{s} \cdot \vec{s}_{\text{est}} \right) \\ &\approx \frac{1}{2} \left(1 - \cos \theta \right) \propto O(1/N). \end{aligned} \tag{13}$$

The angle $\theta \approx |\delta \vec{s}|/s \propto O(1/\sqrt{N})$ indicates the discrepancy between $\vec{s}_{\text{est}}$ and $\vec{s}$. Back to equation (12), the projection between $\vec{s}$ and the direction vector $\delta \vec{s}/|\delta \vec{s}|$, as well as $|\delta \vec{s}|$ itself, is
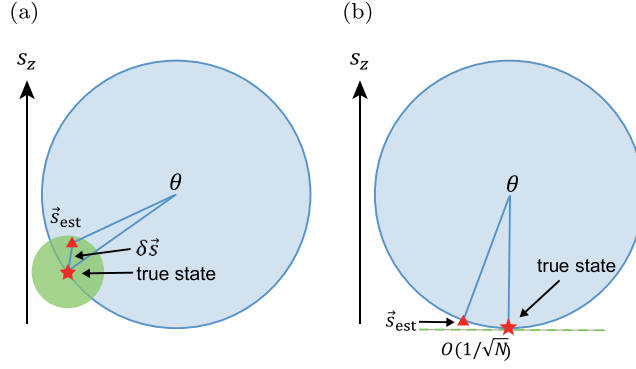
Figure 2. The scenario of estimating pure states. Each figure represents the cross-section of Bloch sphere. The red star denotes the true state $\vec{s}$, and the triangle denotes the estimation $\vec{s}_{est}$. (a) Estimation of a general pure state. The average infidelity is obtained by averaging the infidelity over the distribution of the measurement results (shaded area, green). (b) Estimation of a pure state in the optimal configuration. The probability of observing one of the outcomes becomes zero, therefore the distribution of the measurement results is a planar region (dashed line, green).

$O(1/\sqrt{N})$, therefore the inner product $\vec{s} \cdot \delta\vec{s}$ in equation (12) leads to an infidelity scaling as $O(1/N)$.

2.3. Maximum likelihood estimation

In actual experiments, we collect the count n_i of the measurement outcome i rather than the well-defined probability p_i. Accordingly we substitute empirical frequencies $\hat{p}_i$ that is statistically determined by p_i and reconstruct the density matrix. Due to the statistical fluctuations of the measurement values, it is possible to violate the positivity constraint if we directly apply a linear inversion with the raw data (equation (6)), especially for pure states. In practical tomography, MLE is used extensively to reconstruct a physical state [4, 47]. In compliance with the Bayesian principle of statistical inference, MLE maximizes the likelihood function [45, 48]

$$\mathcal{L}\left(n_i|\rho\right) = \prod_i \left[\mathrm{Tr}\left(E_i\rho\right)\right]^{n_i} \tag{14}$$

subjects to the conditions $\rho \geqslant 0$ and $\mathrm{Tr}(\rho) = 1$. Generally, this problem can be conducted using standard tools for convex optimization. Nevertheless, the extremum algorithm proposed in [45] is instructive to give an analytical derivation. Here we take the scenario of SIC-POVM as an example to explain the mechanism of MLE. By incorporating the likelihood function with Lagrange multipliers, the extremum equations can be written of the form

$$\mu + 2 - \frac{1}{2}\sum_i \sqrt{1 - \mu^2 + 12\mu\hat{p}_i} = 0, \tag{15}$$

$$\frac{\hat{p}_i}{\tilde{p}_i} = 1 - \mu + 3\mu\tilde{p}_i, \tag{16}$$

where μ is a Lagrange multiplier, $\tilde{p}_i$ denotes the probability distribution which optimizing the logarithm of the maximum likelihood function $\log \mathcal{L}\left(n_i|\rho\right)$. For a pure state that is antiparallel

to one of the projections in SIC-POVM, for example $|\psi_0\rangle$, equations (15) and (16) are solved near $\vec{p} = (0, 1/3, 1/3, 1/3)$ and give the solution

$$\tilde{p}_i = \sqrt{\frac{1}{3}\hat{p}_i}. \tag{17}$$

According to the multinomial distribution, the measurement results of the outcomes are

$$\hat{p}_0 = 0;$$
$$\langle \hat{p}_i \rangle = \frac{1}{3}, \quad \Delta\hat{p}_i = \sqrt{\frac{2}{9N}} \quad (i = 1, 2, 3). \tag{18}$$

By means of equation (17), we get the probability distribution $\tilde{p}_i$ which optimizing $\log \mathcal{L}(n_i|\rho)$ under large N approximation

$$\tilde{p}_0 = 1 - \tilde{p}_1 - \tilde{p}_2 - \tilde{p}_3;$$
$$\langle \tilde{p}_i \rangle = \frac{1}{3}, \quad \Delta\tilde{p}_i \approx \frac{1}{2}\sqrt{\frac{2}{9N}} \quad (i = 1, 2, 3). \tag{19}$$

Consequently, MLE not only compresses the estimation $\vec{s}_{\text{est}}$ to a physical state, but also reduces the statistical error of $\hat{p}_i$ by a factor of 2. Then we define the effective length of the state

$$s_{\text{eff}} = \sqrt{\langle \hat{s} \rangle^2 + \langle |\delta\vec{s}|^2 \rangle} = \sqrt{12 \sum_i \langle \hat{p}_i^2 \rangle - 3} \tag{20}$$

to represent the typical length of the reconstructed Pauli vector. This quantity infers a typical reconstructed state given by the linear-inversion of measurement results, which we would use to define the transition region of tomographic performance in the following analyses. Following the multinomial distribution equation (7), one can derive the sum of $\langle \hat{p}_i^2 \rangle$:

$$\sum_i \langle \hat{p}_i^2 \rangle = \frac{1}{N} + \frac{N-1}{N} \sum_i p_i^2$$
$$= \frac{3+s^2}{12} + \frac{9-s^2}{12N}, \tag{21}$$

where $\sum_i p_i^2 = (3+s^2)/12$. Therefore, equation (20) can be represented as $s_{\text{eff}} = \sqrt{s^2 + (9-s^2)/N}$ in general cases. In the estimation of a pure state in the optimal configuration, $\hat{s}$ is always greater or equal to one, and MLE is scarcely avoidable, as shown in figure 2(b). In virtue of the compression effect of MLE, the discrepancy angle $\theta \approx |\delta\vec{s}| = \sqrt{(9-s^2)/4N}$. To see the MLE of the measurement result is denoted by $\vec{s}_{\text{mle}}$, it suffices to note that

$$1 - F = \frac{1}{2}(1 - \vec{s} \cdot \vec{s}_{\text{mle}}) \approx \frac{1}{4}\theta^2 \approx \frac{1}{2N}. \tag{22}$$

For MUB, we can utilize similar analysis and obtain $\theta \approx |\delta\vec{s}| = \sqrt{(9-3s^2)/N}$ in general cases. The infidelity for a pure state in the optimal configuration can be estimated as

$$1 - F = \frac{1}{2}(1 - \vec{s} \cdot \vec{s}_{\text{mle}}) \approx \frac{2}{3N}. \tag{23}$$

J. Phys. A: Math. Theor. **51** (2018) 395304

A Zhang *et al*

2.4. Adaptive protocol and simulation results

From the above analysis, it is shown that the behavior of QST is basis dependent. However, to approach the optimal configuration one requires some priori knowledge about the state. To achieve this goal, a two-step adaptive protocol can be employed: firstly, perform pre-estimation on half of copies $N_1 = N/2$ of the state through static tomography and acquire an estimation $\hat{\rho}_0$; secondly, transform the measurement to the optimal configuration with respect to $\hat{\rho}_0$, and obtain a more accurate estimation of the state through the remaining $N/2$ copies. The two-step adaptive protocol allows us to approach the optimal configuration in the second step with the information acquired on the state within the first step. This procedure is expected to increase the accuracy prominently, resulting in $O(1/N)$ scaling for pure states. The adaptive protocol using MUB has been demonstrated for single qubits [10]. For single qubit, arbitrary set of rank 1 and rank 2 POVM elements including SIC-POVM can be implemented by a one-dimensional quantum walk [41], which suggests that the adaptive protocol using SIC-POVM is also feasible in experimental implementation.

2.4.1. Simulation results. We perform Monte Carlo simulations of the adaptive tomography for the pure state

$$\rho_1^{\text{pure}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{24}$$

with initial outcome probabilities $\vec{p} = (1/2, 1/6, 1/6, 1/6)$ for SIC-POVM; and

$$\rho_2^{\text{pure}} = \begin{pmatrix} \frac{1}{2} + \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{3}}\mathrm{i} \\ \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}}\mathrm{i} & \frac{1}{2} - \frac{1}{\sqrt{3}} \end{pmatrix}, \tag{25}$$

i.e. $\vec{s} = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ for MUB. The results of static tomography and *known-basis tomography* (knowing the true state and resorting to optimal measurement sets) are performed for comparison, as shown in figure 3. Every marker represents an infidelity averaged over 200 repeated trials; the dashed line represents a linear fitting of $\log_{10}(1 - F)$ versus $\log_{10}(N)$, which gives a relation $1 - F = cN^{\alpha}$. The fitting results of α are summarized in table 1. Notably, the adaptive protocol outperforms static tomography and approaches the scaling of known-basis tomography, for pure states. Figure 3 also shows the enormous advantages of the optimal strategy (known-basis tomography), departing from the typical static tomography. We obtain $c = 0.4810$ and $c = 0.5987$ in known-basis power-law fitting for SIC-POVM and MUB respectively, which is consistent with our predictions in equations (22) and (23).

In the same way, we perform simulations for the maximum mixed state $\rho^{\text{mix}} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}$, as shown in figure 4. In this case, the static tomography and known-basis tomography are identical. All the three procedures scale as $O(1/N)$; the adaptive tomography shows no superiority compared with the non-adaptive one.

In summary, due to the physical constraint on the reconstructed state, different configurations for pure states will result in distinct performance of tomography, as we derived in section 2.2. Therefore, the adaptive protocol can improve the accuracy significantly for pure states. Yet for highly mixed states, the measurement results hardly violate the physical constraint. As a result, the accuracy of tomography is completely determined by the statistical fluctuations.

Figure 5 compares the fitting curves of tomography using SIC-POVM and MUB. MUB reveals better accuracy in static tomography, while the SIC-POVM performs better in optimal configuration. In the two-step adaptive procedure, how close the configuration in the second
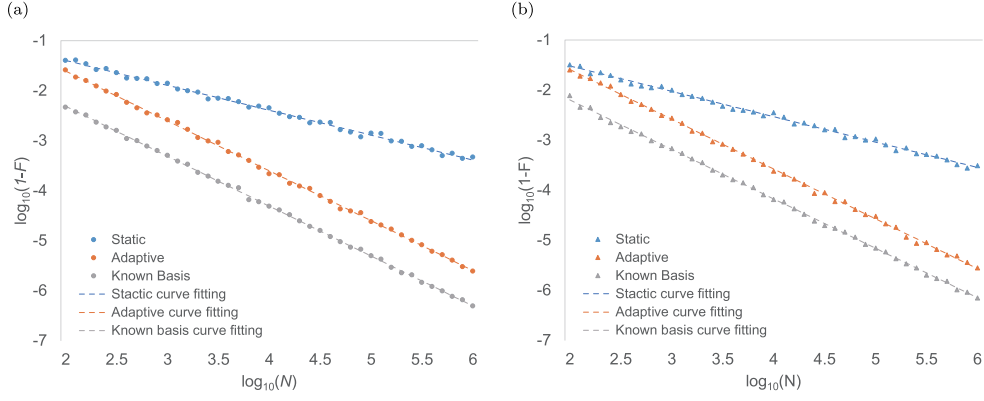
Figure 3. Monte Carlo simulations of different tomographic procedures for pure states: static tomography (blue), adaptive tomography (red) and known-basis tomography (grey). Every data points to the infidelity averaged over 200 repetitions, and the dashed line represents corresponding power-law fitting to illustrate the scaling of infidelity versus *N*. (a) The results for tomography using SIC-POVM (for the state ρ_1^{pure} in equation (24)); (b) The results for tomography using MUB (for the state ρ_2^{pure} in equation (25)).

Table 1. Curve fitting results of α for pure states.

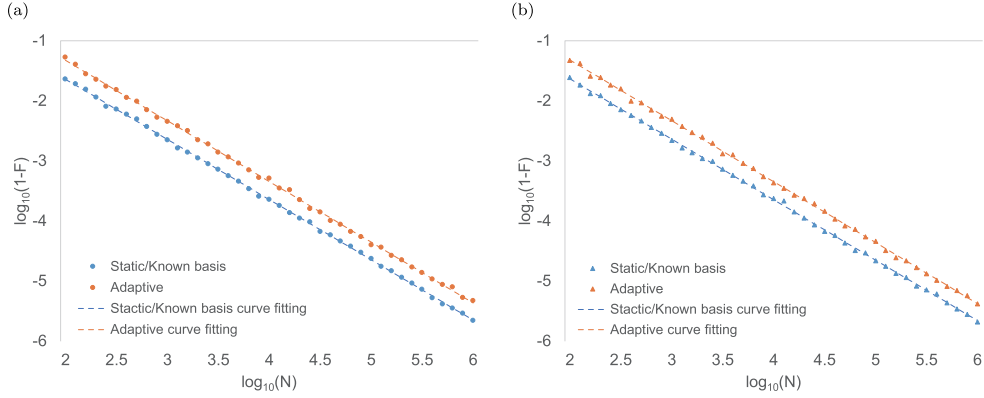|	SIC-POVM	MUB
Static	-0.4977 ± 0.0054	-0.5069 ± 0.0051
Adaptive	-0.9976 ± 0.0042	-0.9951 ± 0.0048
Known-basis	-0.9979 ± 0.0036	-0.9873 ± 0.0043

Figure 4. Monte Carlo simulations of different tomographic procedures for highly mixed states (the maximum mixed state $\rho^{\text{mix}} = \frac{1}{2}\mathbb{1}$): static tomography/adaptive tomography (red) and known basis tomography (grey). (a) The results for tomography using SIC-POVM; (b) The results for tomography using MUB.

step is to the optimal configuration depends on the accuracy of the first step, which will in turn affect the accuracy of the second step. Consequently, the accuracy of the final estimation depends on both the statistical fluctuations in the first step and the accuracy of the optimal configuration. The better accuracy of MUB in pre-estimation results in a smaller discrepancy for
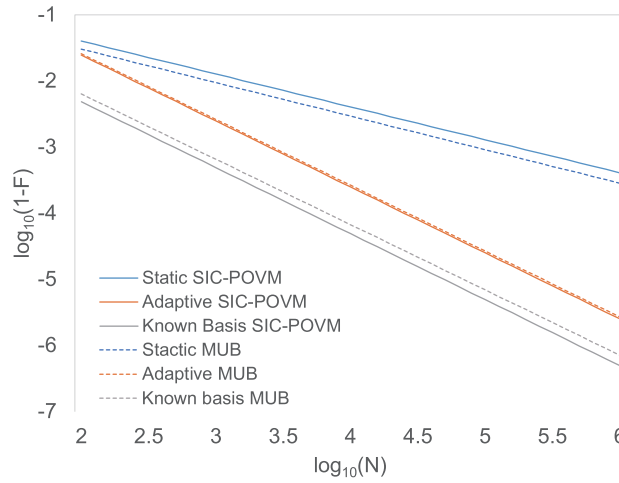
Figure 5. The comparison of the performance of MUB (dashed lines) and SIC-POVM (solid lines) in different tomographic procedures for pure states: static tomography (blue), adaptive tomography (red) and known-basis tomography (grey).

the optimal configuration in the second step, where SIC-POVM obtains more accurate estimation. The estimation accuracy in the two stages for MUB ($3/2N$ for typical static tomography and $2/3N$ for known-basis tomography) and SIC-POVM ($2/N$ for typical static tomography and $1/2N$ for known-basis tomography) are exactly canceled out. Finally, MUB and SIC-POVM reveal similar accuracy in adaptive tomography.

3. Nearly-pure states in quantum tomography

The adaptive protocol manifests its superiority for pure states, but not for highly mixed states. As far as we know, nearly-pure states with a moderate mixture are more popular cases in physical implementations of QIP, therefore understanding the performance of adaptive tomography for nearly-pure states is a key for its applications in practical setups. In the first instance, we perform simulations for states of the form $\rho_1^{\text{near}} = (1-\lambda)\rho_1^{\text{pure}} + \lambda\mathbb{1}$ [$\rho_2^{\text{near}} = (1-\lambda)\rho_2^{\text{pure}} + \lambda\mathbb{1}$ for MUB], where λ is the smaller eigenvalue of ρ_1^{near} (ρ_2^{near}). As an example, the simulation result for $\lambda = 0.0002$ using SIC-POVM is shown in figure 6, which reveals remarkable distinction from pure states. It is shown in section 2.3 that MLE demonstrates its superiority for completely pure states; conversely, statistical fluctuations are mainly responsible for the estimation of highly mixed states. In the transition region of the two kind of states, the effects of purity and statistical fluctuations are comparable and the mechanism of tomography still remains unclear, even for single qubit.

In practice, we simulate a range of cases and observe that the curve contains a turning point, which divides the curve into two segments. The adaptive tomography achieves $O(1/N)$ scaling at large N, but becomes '*invalid*' at small N ($N < O(1/\lambda)$ in the simulation). Notably, the 'invalidation' for nearly-pure states also exists in tomography based on iterative algorithms [12]. To date, there has not been a theory to rigorously interpret this phenomenon.
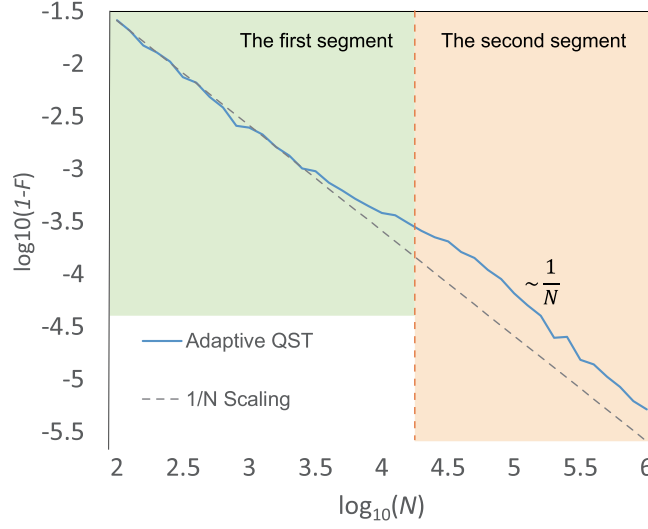
Figure 6. Monte Carlo simulations of adaptive tomography using SIC-POVM. The true state is a nearly pure state $\rho_1^{\text{near}} = \begin{pmatrix} 0.9998 & 0 \\ 0 & 0.0002 \end{pmatrix}$. The curve contains two segments and a turning point. The scaling of the infidelity becomes ' *invalid*' around the demarcation of two segments (vertical dashed line).

3.1. Nearly-pure states in the optimal configuration

We start from the analysis of the first segment in figure 6. As we show in appendix, the 'invalidation' at small N is attributed to the statistical fluctuations of the outcomes instead of the impurity of the state. When the probability p_i of observing one of outcomes i is exceedingly small, a small N is insufficient to estimate the parameter p_i. Envisage a scenario which determines a parameter $p_i = 1/10\,000$ utilizing samples $N = 100$ within one trial. For each trial, we may obtain $\hat{p}_i = 0.01$ with a probability 0.01 and $\hat{p}_i = 0.01$ wih a probability 0.99. If we repeat the experiment with many trials, the expectation of $\hat{p}_i$ should be $1/10\,000$. This situation is exactly what we encounter in QST for nearly-pure states, resulting in a decrease of the average fidelity.

Figure 7(a) illustrates the situation of estimating nearly-pure states with the optimal configuration. For $N \ll 1/\lambda$, it is highly probable to obtain a measurement result in the tangent plane of Bloch sphere, and MLE will obtain a pure state inferred as $\vec{s}_{\text{mle}}$. This situation is very close to the estimation of a pure state in the optimal configuration, therefore we expect that the infidelity can be estimated through $\vec{s}_{\text{mle}}$ in the first segment in figure 6. With the increase of N, the statistical fluctuations decrease and $s_{\text{eff}} \leqslant 1$. The measurement results are not likely to violate the positivity constraint. This situation is very close to the estimation of a highly mixed state, therefore the infidelity can be estimated by equations (10) and (11) directly in the second segment. Recall that in equation (20) we introduce s_{eff} to denote the typical length of the reconstructed Pauli vector. s_{eff} can be considered as kind of figure of merit which represents how close the measurement result is from a pure state. Based on the above analysis, we can set $s_{\text{eff}} = 1$ as the turning point of the two segments, then the average infidelity can be evaluated by
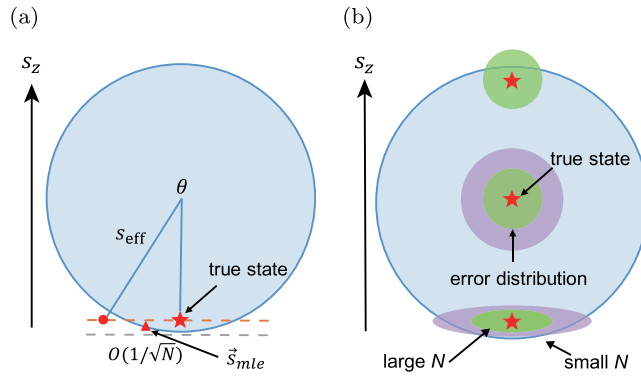
Figure 7. The cross-section of Bloch sphere to illustrate the adaptive quantum state tomography. (a) State estimation for a nearly-pure state (red star) in the optimal configuration. The tangent plane of Bloch sphere (dashed line, grey) denotes the ensemble of measurement results with a high probability to obtain. The representative measurement results (red dot) are inferred by $\vec{s}_{\mathrm{eff}}$ with $\theta \sim O(1/\sqrt{N})$, and $\vec{s}_{\mathrm{mle}}$ denotes the MLE of $\vec{s}_{\mathrm{eff}}$. (b) The distribution of the estimation $\hat{\rho}$ (shaded area) for tomography using SIC-POVM. The distribution of $\hat{\rho}$ (shaded area) depends on both the purity of the true state (red star) and the number of samples N.

$$1 - F = \begin{cases} \frac{1}{2}\left(1 - \vec{s} \cdot \vec{s}_{\mathrm{est}}\right) & s_{\mathrm{eff}} > 1 \\ \frac{1}{2}\left(1 - \vec{s} \cdot \vec{s}_{\mathrm{est}} - \sqrt{1 - s^2}\sqrt{1 - s_{\mathrm{est}}^2}\right) & s_{\mathrm{eff}} \leqslant 1. \end{cases} \tag{26}$$

For nearly-pure states in optimal configuration, the criterion $s_{\mathrm{eff}} \leqslant 1$ could be simplified approximately as $N \geqslant 2/\lambda$, where $\lambda = (1 - s)/2$. With equation (26), one can conclude

(i) For SIC-POVM:

$$\begin{aligned} 1 - F &= \frac{1}{2}\left(1 - ss_{\mathrm{est}}\cos\theta\right) \\ &\approx \frac{1}{2}\left(1 - s\right) + \frac{1}{4}\theta^2 s \approx \lambda + \frac{1}{2N} \quad N \ll \frac{2}{\lambda}; \end{aligned} \tag{27}$$

$$1 - F \approx \frac{1}{4}|\delta\vec{s}|^2 \approx \frac{2}{N} \qquad\qquad N \gg \frac{2}{\lambda}. \tag{28}$$

(ii) For MUB:

$$1 - F \approx \lambda + \frac{2}{3N} \qquad\qquad N \ll \frac{3}{2\lambda}, \tag{29}$$

$$1 - F \approx \frac{3}{2N} \qquad\qquad N \gg \frac{3}{2\lambda}. \tag{30}$$

The theoretical predictions from equations (27) and (29), together with the $1/N$ scaling for the second segment, are shown in figure 8 (dashed line, grey). Simulation results (grey data) are in well agreement with our theoretical predictions. Note the deviation from theoretical prediction is observed near the turning point, because the contribution of MLE attenuates gradually with the increase of N. Nonetheless, our theoretical model is

sufficient to give clear description of the tomography accuracy and is in well agreement with the turning point.

3.2. Adaptive tomography

Let us proceed to deduce the infidelity for adaptive protocol. The pre-estimation step gives a preliminary knowledge of the state with typical error as $O(1/\sqrt{N_1})$. This statistical fluctuation introduces a discrepancy from the optimal configuration in the second step. On average, this discrepancy is equivalent to an extra $O(1/N_1)$ impurity in the final estimation, and can be encapsulated in a coefficient β for SIC-POVM (γ for MUB). Recall that the pre-estimation step consume $N_1 = N/2$ samples of resources. Therefore, two modifications should be applied on equations (27) and (28) in adaptive tomography: (i) λ should be replaced by $\lambda + \beta/N_1$ for SIC-POVM($\lambda + \gamma/N_1$ for MUB); (ii) N should be replaced by $N/2$. Then the formulas of infidelity can be recast into

(i) For SIC-POVM:

$$1 - F \approx \lambda + \frac{2\beta}{N} + \frac{1}{N} \qquad N \ll \frac{4}{\lambda}, \tag{31}$$

$$1 - F \approx \frac{4}{N} \qquad\qquad N \gg \frac{4}{\lambda}. \tag{32}$$

(ii) For MUB:

$$1 - F \approx \lambda + \frac{2\gamma}{N} + \frac{4}{3N} \qquad N \ll \frac{3}{\lambda}, \tag{33}$$

$$1 - F \approx \frac{3}{N} \qquad\qquad N \gg \frac{3}{\lambda}. \tag{34}$$

Figure 8 shows the results for the nearly-pure state ρ using static tomography, adaptive tomography and known basis tomography respectively. The adaptive scheme significantly improves the average fidelity, and achieves $1/N$ scaling for large N. For modest number of samples, adaptive tomography of nearly-pure states performs more prominent improvements on accuracy in comparison with pure states, and approaches the fidelity of known-basis tomography, especially at the number N around the turning point. The dependence of the infidelity on N reveals how the purity of the state and the statistical fluctuations affect the accuracy, as we discussed in section 3.1.

By the comparison of SIC-POVM and MUB shown in figure 9, we conclude that MUB still performs better in static tomography for nearly-pure states. In known-basis tomography, SIC-POVM shows better fidelity at small N, and behaves similarly as MUB at large N. The performance of adaptive tomography is similar for the two measurements. The similar performance is also testified by the fitting results $\beta = 0.6662 \pm 0.0292$ and $\gamma = 0.5296 \pm 0.0362$, which lead to very close values for equations (31) and (33).

3.3. Criterion of nearly-pure states

Back from infidelity to Bloch sphere, we now explain what 'nearly-pure states' is in terms of QST. In previous works, the criteria of nearly-pure states was considered as $\lambda < O(1/\sqrt{N})$. It has been discussed in [10] that why adaptive tomography achieves a better infidelity. Standard
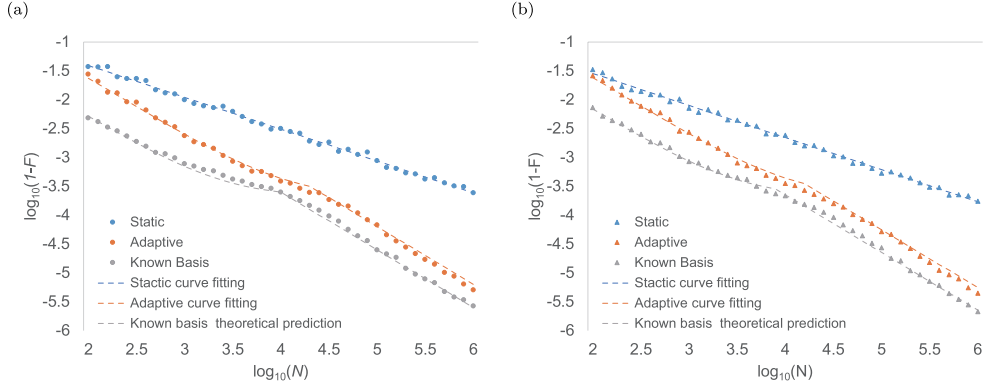
Figure 8. Tomography results for the nearly-pure state with a small eigenvalue $\lambda = 0.0002$ using two measurements: (a) SIC-POVM (for the state ρ_1^{near}); (b) MUB (for the state ρ_2^{near}). Every marker represents an infidelity averaged over 200 repeated trails. Blue: static tomography. Red: adaptive tomography. Grey: known-basis tomography.
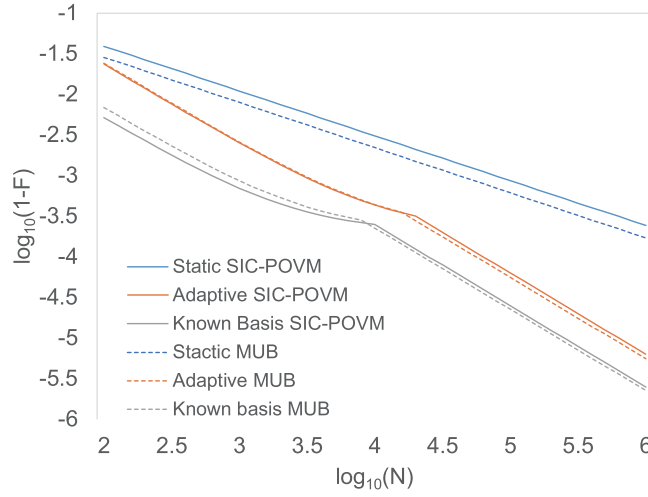

Figure 9. The comparison of the performance of MUB (dashed lines) and SIC-POVM (solid lines) in different tomographic procedures for nearly-pure states: static tomography (blue), adaptive tomography (red) and known-basis tomography (grey).

tomography does not minimize the expected infidelity for two reasons: first, the variance of the estimate $\hat{\rho}$ depends also on ρ itself; second, the dependence of infidelity on the error, $\Delta\rho$, also varies with ρ. Ulteriorly, we assert that criterion of nearly-pure states is also *anisotropic* and *inhomogeneous*, varies with ρ, not invariably $\lambda < O(1/\sqrt{N})$.

It is reasonable that a state is 'nearly-pure' when the reconstructed state $\hat{\rho}$, the linear inversion of the measurement results, can be a pure state. In other words, if the error distribution of measurement results is all within the Bloch sphere, the state behaves as a complete mixed state, therefore the scaling of infidelity turn to $O(1/N)$. For instance, in the optimal configuration for SIC-POVM, the distribution of $\hat{p}_0$ is defined as equation (8). The state becomes nearly-pure when the standard deviation $\Delta\hat{p}_0$ exceeds p_0 itself:

$$\sqrt{\frac{1}{N}}\sqrt{p_0(1-p_0)} > p_0. \tag{35}$$

The result $\lambda < 2/(N+1)$ is in accordance with our simulation results and theoretical derivation (figure 6). One may wonder the question: why the turning point exists? The turning point is existent for all mixed states, typically at $\lambda \sim 1/\sqrt{N}$, but not always at $1/\sqrt{N}$, especially when the optimal configuration is applied. As shown in figure 7(b), the error is compressed in some directions in the Bloch sphere. The purity of the state and the statistical fluctuations of the measurement outcomes jointly define the transition region in the performance of the adaptive tomography. The optimal configuration actually shifts the turning point, rather than transfers the scaling from $O(1/\sqrt{N})$ to $O(1/N)$. Our results are generally valid for all mixed states, and become visible for nearly-pure states, rather than highly mixed states. Pure states are special cases that the smaller eigenvalue is zero.

In adaptive tomography, diagonalizing the density matrix ρ is not the recipe of achieving a better fidelity, but the probability p_i of observing one of the POVM outcomes equals (or approaches) zero is crucial. Both parallel and antiparallel strategies in SIC-POVM diagonalize the density matrix but provide totally different performance on nearly pure states. In substance, the recipe is to eliminate (or reduce) the *randomness* of outcomes. In optimal configuration for pure states, the error in one of the outcomes is eliminated completely, rendering the response of this outcome from a *probabilistic* process to a *deterministic* one. Therefore, absolute $1/N$ scaling would be achieved if the state is completely pure and the measurement is aligned in the truly optimal configuration, when the full information about the state is known.

4. Conclusion

In conclusion, we investigated adaptive quantum state tomography using MUB and SIC-POVM both of which give rise to $O(1/N)$ infidelity for large N. In comparison with previous works, we gave a clear picture for scaling improvement of adaptive strategy compared to the conventional tomography, and further interpreted why adaptation achieves a better fidelity by considering the effect of stochastic fluctuations. In particular, we discussed how the interplay between the purity of the state and the stochastic fluctuations of measurement results affect the accuracy of reconstruction. Our analyses can be generalized to the tomography of quantum states with more than a single qubit. Specifically, the effects of the purity of the state on the accuracy of reconstruction would be equally effective. Our results highlight the unique behavior of nearly-pure states in quantum state tomography, and fill the gap between the tomography in fully optimal configuration for a pure state and adaptive tomography for practical cases.

Acknowledgments

J. Phys. A: Math. Theor. **51** (2018) 395304

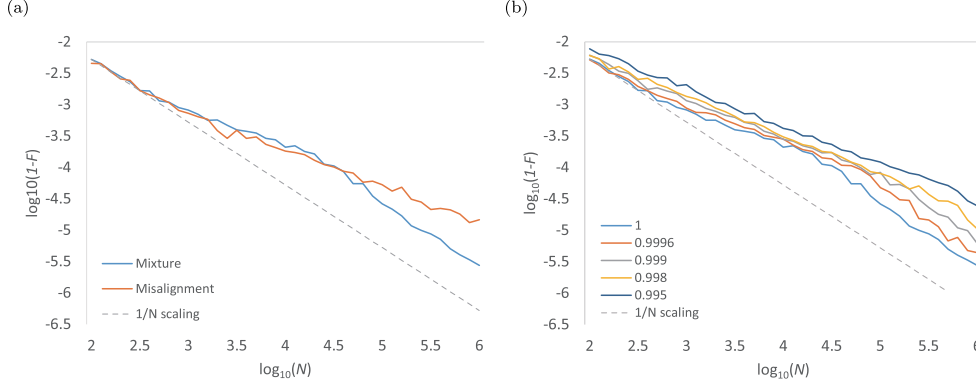A Zhang *et al*

(a)

(b)

Figure A1. Simulation results for static tomography using SIC-POVM: (a) the misaligned pure state ρ_1 (red) and the nearly pure state ρ_2 (blue). The two states perform similarly in the first segment, but diverges at higher *N*. (b) The states with same purity but various misalignments, which are defined by the discrepancy angle θ from the optimal configuration. The solid lines in different colors represent different misalignments with respect to the optimal configuration.

Appendix. Statistical fluctuations and impurity for nearly-pure states

In the scenario of QST, there are two differences for the estimated state between a pure state and a nearly-pure state: (i) the probabilities of the outcomes p_i, which is directly relevant to the statistical fluctuations; (ii) the purity of the true state, which is quantified by the eigenvalues of the states. To identify why the adaptive protocol becomes *invalid* for nearly-pure states, we perform static tomography using SIC-POVM for two states: the misaligned pure state ρ_1 (by which we refer to the misalignment from the optimal configuration) and the nearly pure state ρ_2:

$$\rho_1 = \begin{pmatrix} 0.0002 & 0.0141\mathrm{i} \\ 0.0141\mathrm{i} & 0.9998 \end{pmatrix}, \rho_2 = \begin{pmatrix} 0.0002 & 0 \\ 0 & 0.9998 \end{pmatrix}.$$

Both of the two states lead to the same probability $p_0 = 0.0001$ for the outcome zero, but only the nearly pure state ρ_2 has some impurity ($\lambda = 0.0002$). The simulation results of static tomography for the two states are shown in figure A1(a). We conclude from this comparison that the infidelities for both states behave similarly at small *N*, which should be attributed to the probability $p_0 = 0.0001$; conversely, the scaling of two states diverge at higher number of samples ($\log_{10} N > 4.0$). The infidelity of the nearly-pure state ρ_2 transforms into $O(1/N)$ scaling while the infidelity of the misaligned pure state ρ_1 maintains $O(1/\sqrt{N})$ scaling.

What if the states have the same purity but different misalignments? The results for static tomography are illustrated in figure A1(b). The misalignment is quantified by its angle θ deviated from the optimal configuration. We take $\cos\theta = 1, 0.9996, 0.999, 0.998, 0.995$ in the numerical simulations respectively. The turning point in each situation is obvious compared with the $1/N$ scaling. It infers that even the states have identical degree of impurity, these states reveal different manifestations for a range of misalignments, due to the fact that the statistical fluctuations are not the same, and result in different conditions of nearly-pure states.

ORCID iDs

Aonan Zhang ⓘ https://orcid.org/0000-0002-6310-4769

References

[1] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge: Cambridge University Press)
[2] Kok P, Munro W J, Nemoto K, Ralph T C, Dowling J P and Milburn G J 2007 *Rev. Mod. Phys.* **79** 135–74
[3] Hradil Z 1997 *Phys. Rev.* A **55** R1561
[4] Paris M and Řeháček J (ed) 2004 *Quantum State Estimation* vol 649 (New York: Springer)
[5] James D F V, Kwiat P G, Munro W J and White A G 2001 *Phys. Rev.* A **64** 052312
[6] Braunstein S L and Caves C M 1994 *Phys. Rev. Lett.* **72** 3439–43
[7] Demkowicz-Dobrzański R 2011 *Phys. Rev.* A **83** 061802
[8] Petz D and Ruppert L 2012 *J. Phys. A: Math. Theor.* **45** 085306
[9] Hou Z, Zhu H, Xiang G Y, Li C F and Guo G C 2016 *NPJ Quantum Inf.* **2** 16001
[10] Mahler D H, Rozema L A, Darabi A, Ferrie C, Blume-Kohout R and Steinberg A M 2013 *Phys. Rev. Lett.* **111** 183601
[11] Kravtsov K S, Straupe S S, Radchenko I V, Houlsby N M T, Huszár F and Kulik S P 2013 *Phys. Rev.* A **87** 062122
[12] Qi B, Hou Z, Wang Y, Dong D, Zhong H S, Li L, Xiang G Y, Wiseman H M, Li C F and Guo G C 2017 *NPJ Quantum Inf.* **3** 19
[13] Chapman R J, Ferrie C and Peruzzo A 2016 *Phys. Rev. Lett.* **117** 040402
[14] Okamoto R, Iefuji M, Oyama S, Yamagata K, Imai H, Fujiwara A and Takeuchi S 2012 *Phys. Rev. Lett.* **109** 130404
[15] Yin Q, Li L, Xiang X, Xiang G Y, Li C F and Guo G C 2017 *Phys. Rev.* A **95** 012129
[16] Wang H *et al* 2016 *Sci. China Phys. Mech. Astron.* **59** 100313
[17] Sugiyama T, Turner P S and Murao M 2012 *Phys. Rev.* A **85** 052107
[18] Nunn J, Smith B J, Puentes G, Walmsley I A and Lundeen J S 2010 *Phys. Rev.* A **81** 042109
[19] Zhu H and Englert B G 2011 *Phys. Rev.* A **84** 022327
[20] Ferrie C and Blume-Kohout R 2016 *Phys. Rev. Lett.* **116** 090407
[21] Zhu H 2014 *Phys. Rev.* A **90** 012115
[22] Foreman M R, Favaro A and Aiello A 2015 *Phys. Rev. Lett.* **115** 263901
[23] Zhang L, Chan K W C and Verma P K 2016 *Phys. Rev.* A **93** 032137
[24] Wootters W K and Fields B D 1989 *Ann. Phys.* **191** 363–81
[25] Durt T, Englert B G, Bengtsson I and Życzkowski K 2010 *Int. J. Quantum Inf.* **08** 535–640
[26] Wootters W K 2006 *Found. Phys.* **36** 112–26
[27] McNulty D and Weigert S 2012 *Int. J. Quantum Inf.* **10** 1250056
[28] de Burgh M D, Langford N K, Doherty A C and Gilchrist A 2008 *Phys. Rev.* A **78** 052122
[29] Scott A J 2006 *J. Phys. A: Math. Gen.* **39** 13507–30
[30] Durt T, Kurtsiefer C, Lamas-Linares A and Ling A 2008 *Phys. Rev.* A **78** 042338
[31] Medendorp Z E D, Torres-Ruiz F A, Shalm L K, Tabia G N M, Fuchs C A and Steinberg A M 2011 *Phys. Rev.* A **83** 051801
[32] Bent N, Qassim H, Tahir A A, Sych D, Leuchs G, Sánchez-Soto L L, Karimi E and Boyd R W 2015 *Phys. Rev.* X **5** 041006
[33] Dall'Arno M, D'Ariano G M and Sacchi M F 2011 *Phys. Rev.* A **83** 062304
[34] Szymusiak A 2014 *J. Phys. A: Math. Theor.* **47** 445301
[35] Dall'Arno M, Buscemi F and Ozawa M 2014 *J. Phys. A: Math. Theor.* **47** 235302
[36] Brandsen S, Dall'Arno M and Szymusiak A 2016 *Phys. Rev.* A **94** 022335
[37] Zauner G 2011 *Int. J. Quantum Inf.* **09** 445–507
[38] Renes J M, Blume-Kohout R, Scott A J and Caves C M 2004 *J. Math. Phys.* **45** 2171
[39] Bian Z, Li J, Qin H, Zhan X, Zhang R, Sanders B C and Xue P 2015 *Phys. Rev. Lett.* **114** 203602
[40] Zhao Y Y, Yu N K, Kurzyński P, Xiang G Y, Li C F and Guo G C 2015 *Phys. Rev.* A **91** 042101
[41] Kurzynski P and Wojcik A 2013 *Phys. Rev. Lett.* **110** 200404

[42] Clarke R B M, Kendon V M, Chefles A, Barnett S M, Riis E and Sasaki M 2001 *Phys. Rev.* A **64** 012303
[43] Jozsa R 1994 *J. Mod. Opt.* **41** 2315–23
[44] Bagan E, Baig M, Muñoz Tapia R and Rodriguez A 2004 *Phys. Rev.* A **69** 010304
[45] Řeháč J, Englert B G and Kaszlikowski D 2004 *Phys. Rev.* A **70** 052321
[46] Słomczyński W and Szymusiak A 2016 *Quantum Inf. Process.* **15** 565–606
[47] Banaszek K, D'Ariano G M, Paris M G A and Sacchi M F 1999 *Phys. Rev.* A **61** 010304
[48] Blume-Kohout R 2010 *New J. Phys.* **12** 043034