Optimal single-qubit tomography: Realization of locally optimal measurements on a quantum computer

Bacui Li [1,*] Lorcán O. Conlon [1,†] Ping Koy Lam[1,2,‡] and Syed M. Assad[1,§]

[1]*Centre for Quantum Computation and Communication Technology, Department of Quantum Science,*
Australian National University, Canberra, ACT 2601, Australia
[2]*Institute of Materials Research and Engineering, Agency for Science,*
Technology and Research (A*STAR), 2 Fusionopolis Way, 08-03 Innovis 138634, Singapore

Quantum bits, or qubits, are the fundamental building blocks of present quantum computers. Hence, it is important to be able to characterize the state of a qubit as accurately as possible. By evaluating the qubit characterization problem from the viewpoint of quantum metrology, we are able to find optimal measurements under the assumption of good prior knowledge. We implement these measurements on a superconducting quantum computer. Our demonstration produces sufficiently low error to allow the saturation of the theoretical limits, given by the Nagaoka-Hayashi bound. We also present simulations of adaptive measurement schemes utilizing the proposed method. The results of the simulations show the robustness of the method in characterizing arbitrary qubit states with different amounts of prior knowledge.

I. INTRODUCTION

The qubit is the basic unit of information in most quantum computers. Qubit-based quantum computers have shown great promise in sensing [1] and computation technologies [2,3]. Thus, it is essential for the development of quantum computers to study qubit states in detail. The field of research which focuses on characterizing quantum states is called state tomography [4–6]. More specifically, state tomography concerns the problem of finding the most accurate method of reconstructing the state with a fixed amount of quantum resources, or equivalently, finding the most efficient method that reconstructs the state up to some fixed threshold uncertainty with the minimal amount of quantum resources.

Any reconstruction method will involve measurements of the quantum state. All physically possible measurements in the context of quantum systems are described by a positive operator-valued measure (POVM). Among all possible POVMs, projective measurements form a special subset. Projective measurements can be described by a single observable operator and are orthogonal to one another. To reconstruct a quantum state without ambiguity, we require that the measurement used is informationally complete [7]. With certain measurements, the information obtained is insufficient to reconstruct the state. For example, a projective measurement along the z axis of a single-qubit state cannot provide any information about the x or y components of the Bloch vector,

and this measurement alone will fail to reconstruct an arbitrary state.

Two reconstruction methods that are widely regarded as optimal are projective measurements onto mutually unbiased bases (MUB) [6] and the symmetric informationally complete POVM (SIC-POVM) [8,9]. Both of these measurements have been realized experimentally [10–12], and the SIC-POVM has already been demonstrated for quantum entanglement detection [13,14] and tomography protocols [11,12,15,16]. However, for reconstructing an unknown state, the repeated measurement of an ensemble of identical unknown states with either the SIC-POVM or MUB will not be the optimal method. This is because neither of these measurements exploits the knowledge gained from previous measurement results. Indeed, the optimality of the SIC-POVM for state reconstruction has been shown only under the condition of measuring a completely unknown state uniformly distributed following the Haar measure of the unitary operators [17,18]. Thus, when we have partial knowledge of the state in question, other measurements may become optimal.

The optimality of the SIC-POVM and MUB is questioned in a localized setting, where the quantum state is well known, and we attempt to sense small changes in the state. This bears great similarity to the setting used in quantum metrology, where physical parameters are encoded onto quantum states, and we attempt to sense small changes in those parameters [19]. It is natural, therefore, to utilize tools from quantum metrology to analyze quantum state tomography in the localized setting. In quantum metrology, it is common to use Cramér-Rao bounds to place limits on how accurately physical quantities can be measured [19–21]. One bound which is of particular practical relevance is the Nagaoka-Hayashi (N-H) bound, as it applies when probe states are measured individually [22,23]. In this work, we map the qubit tomography

*patrick.li.bj@qq.com
†lorcanconlon@gmail.com
‡Ping.Lam@anu.edu.au
§cqtsma@gmail.com

032605-1

problem to quantum metrology, allowing us to use the N-H bound as an experimentally accessible bound for tomography. The N-H bound, thus, provides a reference against which the measurement proposed in this paper can be compared.

Under the local setting, a measurement inspired by MUB has been proposed. This measurement used a probabilistic mixture of three projective measurements that are orthogonal to each other, such as projective measurements on the x, y, and z axes of the Bloch sphere and was shown to be optimal for the single-qubit state reconstruction problem [24]. In this work we investigate an optimal reconstruction using a POVM inspired by the SIC-POVM. The SIC-POVM can be represented by an equiangular constellation of complex vectors in Hilbert space. For qubit states, the SIC-POVM, therefore, forms a tetrahedron. The POVM we propose is similar, being a tetrahedron squashed along one axis. Like the SIC-POVM, our measurement requires only four POVM elements, the minimal number for an informationally complete measurement on a single-qubit state.

From a theoretical perspective, the connection between our proposed POVM and the SIC-POVM suggests the existence of some unexplored optimal POVMs in higher dimensions[1] given any specific amount of prior knowledge. From a practical viewpoint, we anticipate that the proposal and implementation of an optimal measurement for localized state tomography will stimulate the development of optimal adaptive protocols for quantum tomography that converge to the locally optimal measurement we proposed.

The structure of this paper is as follows: in Sec. II we give a detailed description of the state tomography problem in the localized setting. We demonstrate that the SIC-POVM, which is optimal for uniformly distributed states, does not saturate the N-H bound. Hence, it is not optimal in this context. In Sec. III we introduce optimal POVMs that saturate the N-H bound. We present the analytical form of the POVMs, which depend on the density matrix of the mixed state ρ.

In Sec. IV we show the experimental procedure that implements the optimal POVM on a quantum computer. The method proposed is robust for estimating all mixed states except for marginal cases when the purity is exactly 1. Here we utilize an efficient circuit decomposition of a two-qubit unitary [26], which uses fewer CNOT gates and single-qubit gates than the previous approach of implementing qubit POVMs [27]. The processing of the raw quantum circuit output integrates measurement gate error mitigation [28] and a live calibration scheme which ensures the unbiasedness of the measurement and improves the overall performance. These error mitigation approaches have been shown to improve the quality of results in other areas of quantum information research [29–31].

In Sec. V we show the results generated by methods discussed in Sec. IV. The results validate the feasibility and demonstrate the optimality of the proposed POVMs. Nonetheless, due to various noises in the current generation of quantum computers, the resulting mean squared error (MSE)

value has room for improvement, as the resulting MSE is at worst 15% above the predicted theoretical MSE.

Finally, in Sec. VI we will discuss the application of the proposed POVMs in a broader context and address the limitation of the method when purity is close to or is exactly 1. We address one of the major drawbacks of the localized approach, namely, that a good prior knowledge of the state is required. We present a two-step adaptive method incorporating both the SIC-POVM and our proposed POVMs. This adaptive method outperforms either the SIC-POVM or the proposed POVM when either of them is used independently. We also consider a generalization of our optimal POVM for arbitrary amounts of prior knowledge. In the limit of a completely unknown state, this reduces to the SIC-POVM as expected. The result from simulations of these two adaptive measurement approaches demonstrates the robustness of the method without the assumption of good prior knowledge.

II. BACKGROUND

To find the optimal measurement in the local setting discussed above, we first define the problem and the figure of merit, which will be used throughout the paper. The tomographical problem we will investigate is formulated as a linear quantum state tomography problem [32]: We would like to reconstruct an arbitrary mixed single-qubit state linearly parameterized as

$$\rho = (I + \vec{\sigma} \cdot \boldsymbol{\theta})/2, \tag{1}$$

where $\vec{\sigma}$ is the vector of Pauli matrices and $\boldsymbol{\theta}$ is the vector of parameters we wish to learn about. We define r as the length of the vector $\boldsymbol{\theta}$, i.e., $r = |\boldsymbol{\theta}|$. r also determines the purity by $(r^2 + 1)/2$ and is a crucial parameter for the qubit state as we shall see in the later sections. For the estimated state, $\hat{\rho}$, the metric of deviation we use is the trace norm $D_{\rho,\hat{\rho}} = \mathrm{tr}[\sqrt{(\hat{\rho} - \rho)^{\dagger}(\hat{\rho} - \rho)}]$ (twice the trace distance). In this formulation, the trace norm equals the Euclidean distance between the estimated parameter vector $\hat{\boldsymbol{\theta}}$ and the true parameter vector $\boldsymbol{\theta}$, i.e., $D_{\rho,\hat{\rho}} = |\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}|$ where $\hat{\boldsymbol{\theta}}$ is the estimated vector of parameters $\boldsymbol{\theta}$ [19].

This formulation enables us to minimize the expectation value of the square of the trace norm between the true state and the estimated state, $E[D^2_{\rho,\hat{\rho}}]$ by minimizing the total MSE of the three parameters. In addition to the formulation presented above, some studies also consider the Hilbert-Schmidt norm or fidelity as their figure of merit [16,17]. In the case of qubit state tomography, the Hilbert-Schmidt norm is proportional to the trace norm between two arbitrary mixed states; thus, it is equivalent to minimizing either of the two norms. Under the formulation discussed above, the MSE produced by the SIC-POVM for a mixed state is [5]

$$\mathrm{MSE}_{\mathrm{SIC}} = 9 - r^2. \tag{2}$$

While the MSE that the SIC-POVM can attain is optimal in the nonlocal setting, the development of quantum metrology theory suggests that better measurements may be possible in the local setting we are considering. One such development is the precision bound for multiparameter estimation, the N-H bound [22,23]. This bound generalizes information trade-off relations such as the Heisenberg uncertainty principle and

---

[1]Although the existence of such an object for arbitrary dimension N is still an open problem [25].

gives the precision limit of any multiparameter estimation problem when the probe states are measured individually. The MSE described above can be lower bounded by the N-H bound as we have essentially transformed the problem of linear quantum state tomography into a multiparameter estimation problem [22], for which the N-H bound is applicable. Facilitated by the method of semidefinite programming for efficient calculation of the N-H bound [23], we verified that the N-H bound for the single-qubit tomography problem is

$$C_{\mathrm{NH}} = (2 + \sqrt{1 - r^2})^2. \tag{3}$$

This shows that the MSE for estimating $\boldsymbol{\theta}$ obtained from identical measurements on N probe states, $\mathrm{MSE}(N)$ when measuring them one by one is lower bounded by C_{NH}/N for all unbiased measurement methods.[2] A measurement that reaches the value of C_{NH}/N will be deemed optimal. It is easy to show that $C_{\mathrm{NH}} < \mathrm{MSE}_{\mathrm{SIC}}$ except at $r = 0$. Thus, the SIC-POVM is not optimal, and there may exist an optimal measurement that performs better than the SIC-POVM and saturates the N-H bound.

III. OPTIMAL POVM

We now present the analytical form of the optimal POVMs which saturate the N-H bound. Depending on the mixed state, ρ [as parameterized in Eq. (1)], we have different solutions for the optimal POVM, $\{\Pi_i(\boldsymbol{\theta})\}$. Each POVM consists of four rank-1 operators. There is a simple decomposition for the elements, $\Pi_i = r_i |\psi_i\rangle\langle\psi_i|$ where r_i is the real amplitude and $|\psi_i\rangle$ is a pure-state state vector. Notice that the amplitude r_i denotes the trace of the corresponding POVM element, i.e., $\mathrm{tr}(\Pi_i) = r_i$. Hence, we can visualize the POVM element Π_i by visualizing a vector in the Bloch sphere whose direction is defined by the direction of $|\psi_i\rangle$ and its length by r_i. The four vectors are oriented such that the representation of the POVM element, Π_z, points along the opposite direction of the probe state ρ while the other three point to vertices of an equilateral triangle in the opposite direction. A POVM in this symmetrical shape is inspired by the SIC-POVM which forms a tetrahedron; thus, we call this POVM the *squashed tetrahedron* POVM (ST-POVM).

The ST-POVM is parameterized by two parameters, $0 \leqslant r_p < 1$ which determine the "squashedness" of the tetrahedron and the angle ϕ which orientates the legs forming the equilateral triangle. Explicitly, the four POVM outcomes are given by $r_z = (1 + \sqrt{\frac{1+r_p}{1-r_p}})^{-1}$, $r_1 = r_2 = r_3 = \frac{2-r_z}{3}$ and

$$
\begin{aligned}
|\psi_z\rangle &= |1\rangle, \\
|\psi_1\rangle &= |0\rangle A_0 + |1\rangle A_1 e^{i\phi}, \\
|\psi_2\rangle &= |0\rangle A_0 + |1\rangle A_1 e^{i(\phi+2\pi/3)}, \\
|\psi_3\rangle &= |0\rangle A_0 + |1\rangle A_1 e^{i(\phi+4\pi/3)},
\end{aligned}
\tag{4}
$$

where $A_0 = 1/\sqrt{3r_1}$ and $A_1 = \sqrt{1 - 1/3r_1}$. These four POVM elements with $\phi = 0$ are shown in Fig. 1.
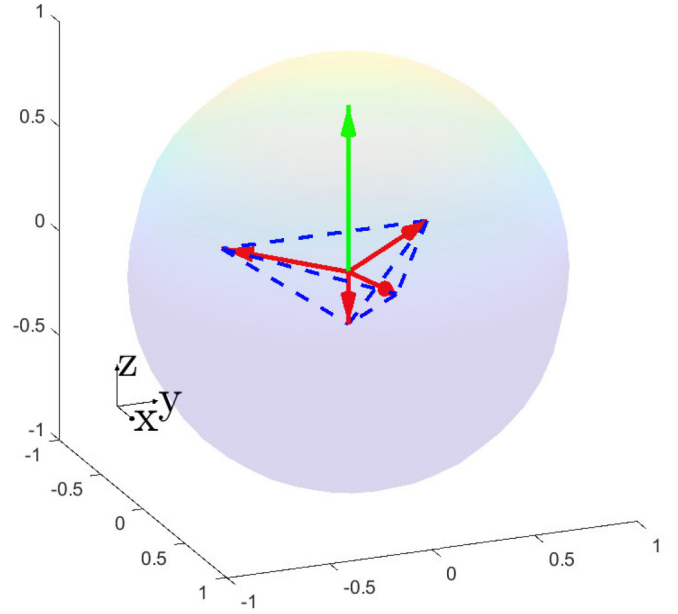

FIG. 1. A Bloch sphere representation of the four elements of the ST-POVM (with $r_p = 0.8$ and $\phi = 0$) shown by the red arrows. The green arrow shows the state being estimated. The shape formed by the four vertices is a triangular pyramid. When the state being estimated lies closer to the surface of the Bloch sphere (higher purity), the shape formed by the four elements is squashed further. When estimating the maximally mixed state, r_p goes to zero, and the ST-POVM converges to the SIC-POVM.

With this POVM and setting $\phi = 0$, we obtain an estimate of $\boldsymbol{\theta}$ from $\hat{\theta}_j = \sum_k \mathcal{E}_{jk} f_k$, where

$$
\mathcal{E} = \begin{pmatrix} 0 & 2a & -a & -a \\ 0 & 0 & \sqrt{3}a & -\sqrt{3}a \\ b & 1 & 1 & 1 \end{pmatrix}, \tag{5}
$$

$a = \sqrt{1 + \sqrt{\frac{1-r_p}{1+r_p}}}$ and $b = -1 - 2\sqrt{\frac{1+r_p}{1-r_p}}$. $f_k = n_k/N$ ($k = z, 1, 2, 3$) are the frequencies with which each POVM outcome occurs given N probes are measured and the outcome Π_k was observed n_k times. One can verify the unbiasedness of $\hat{\boldsymbol{\theta}}$ for all r_p and ϕ by calculating $\mathbb{E}[\hat{\theta}_j]$. This is given by $\mathbb{E}[\hat{\theta}_j] = \sum_k \mathcal{E}_{jk} p_k = \theta_j$ where $p_k = \mathrm{tr}[\Pi_k \rho]$ are the probabilities for outcome k. The expected MSE for this estimator is $\mathbb{E}[\mathrm{MSE}] = \sum_{j,k} p_k (\mathcal{E}_{jk} - \theta_j)^2$. For the state $\rho = |0\rangle \frac{1+r}{2}\langle 0| + |1\rangle \frac{1-r}{2}\langle 1|$, the expected MSE is minimized when $r_p = r$, and this saturates the N-H bound.

Without loss of generality, we have provided the analytical form of the POVM when the probe state is oriented along the z axis. For any other state oriented differently, it is always possible to use a coordinate system that aligns its positive z axis with the state and use the formulation given above before transforming the estimated state back to the original coordinate system. This change of basis doesn't affect the optimality of the measurement as the MSE based on trace norm is invariant under unitary transformations.

In principle, it is always possible that the estimated state will not lie within the Bloch sphere, i.e., $|\hat{\boldsymbol{\theta}}| > 1$. However, in practical scenarios, as pointed out in Ref. [18], the probability

---

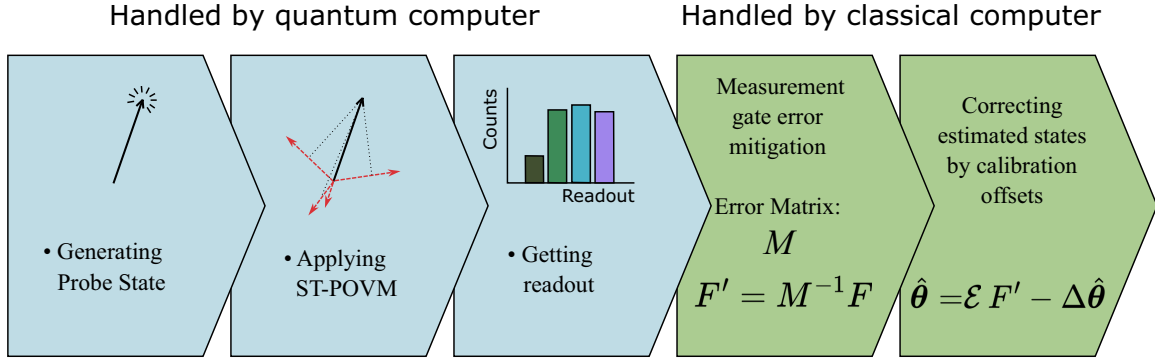[2]This bound agrees with the bound derived by Gill and Massar [24,33].

FIG. 2. An illustration of the procedures involved in the demonstration. The first three stages are handled by the quantum computer, and the last two stages are handled by the classical computer. The classical operations take the frequencies of outcomes, F, and apply the measurement error mitigation. At the last step, by applying the preconstructed estimator $\mathcal{E}$ and the calibration offset $\Delta\hat{\theta}$, we obtain the final estimates of the parameters.

of obtaining a physically impossible state will decay to zero with increasing repetitions of estimations due to the large deviation theory [34] as long as the estimator is unbiased. However, if the purity of the state is close to 1, the number of experiments required to guarantee the absence of a physically impossible estimate can be too big to realize experimentally. In this paper we do not consider this issue because neither states with r close to 1 nor an insufficient number of experiments occur.

IV. METHODOLOGY

We implemented the ST-POVM on a superconducting quantum computer to demonstrate its feasibility and utility. An illustration of the general procedure of the demonstration is shown in Fig. 2. The demonstration involves both operations on quantum computers and classical computers. For the quantum operations, we generated the state and applied the ST-POVM before measuring it at the end. For the classical statistical operations, we utilize the measurement gate error mitigation technique [28] and a calibration procedure to improve our results. The procedure shown in Fig. 2 performs estimation with a fixed number of probe states N. However, through this paper, we will subsample these data to obtain estimates of the quantum state based on different numbers of trials

In this section we will first discuss the method of preparing probe states and then the method of implementing the POVM on the quantum computer. Finally, we will discuss the detailed implementation of the demonstration with error-mitigating techniques and calibration.

A. State preparation

Quantum circuits initialize only pure states, and the unitary quantum gates available do not change the purity of the initialized state in principle. Thus, to prepare a mixed state, we use the result that a mixed-qubit state can be generated by a probabilistic combination of two pure states with probability P_1 and P_2. For a single-qubit mixed state ρ, it is always possible to decompose the mixed state as follows:

$$\rho = P_1 |\phi_1\rangle\langle\phi_1| + P_2 |\phi_2\rangle\langle\phi_2|. \tag{6}$$

Thus, implementing an estimation scheme on a certain mixed state is equivalent to performing estimations with pure states randomly chosen between $|\phi_1\rangle$ and $|\phi_2\rangle$ with probabilities P_1 and P_2. The U_1 block in Fig. 3(a) generates the desired pure state $|\phi_i\rangle$ for the simulation of a mixed state.

U_1 is designed to bring the initial $|0\rangle$ state of the probe qubit to the pure state of either $|\phi_1\rangle$ or $|\phi_2\rangle$. The decomposition shown in Eq. (6) is not unique. In our demonstration, we use pure states $\rho(\boldsymbol{\theta})$ ($|\boldsymbol{\theta}| = 1$) that have the same θ_y and θ_z as the target mixed state ρ. The probabilities assigned to the two states will have a ratio of $P_1 : P_2 = |\theta_x - \theta_x^{(2)}| : |\theta_x - \theta_x^{(1)}|$. θ_x is the x axis projection of the mixed state. $\theta_x^{(1)}$ and $\theta_x^{(2)}$ are the x-axis projections of the two pure states $\rho_1(\boldsymbol{\theta}) = |\phi_1\rangle\langle\phi_1|$ and $\rho_2(\boldsymbol{\theta}) = |\phi_2\rangle\langle\phi_2|$. During the initialization of the mixed state ρ, a pure state will be randomly selected from $\{\rho_1, \rho_2\}$ with probabilities P_1 and P_2, respectively.

B. Naimark extension

The quantum circuit uses two qubits for each instance of the chosen pure state. One is the probe qubit, and the other is the auxiliary qubit which helps the implementation of the POVM. The U_2 block in Fig. 3(a) represents the combined unitary operator of a series of two-qubit and single-qubit gates that entangle the probe qubit and the auxiliary qubit. This is to use the expanded Hilbert space for implementing the ST-POVM on the probe qubit.

We followed the principles of Naimark's dilation theorem [19,35,36] to construct U_2, which requires an auxiliary qubit for our experiment. The U_2 constructed should entangle the probe qubit and the auxiliary qubit initialized in the $|0\rangle$ state in a particular way such that a chosen projective value measurement on the extended two-qubit state E_i^+ produces the same probability distribution as the original POVM:

$$\text{tr}(\rho \otimes |0\rangle\langle 0| U_2^\dagger E_i^+ U_2) = \text{tr}(\rho \Pi_i) \quad \text{for } \forall \rho. \tag{7}$$

For more detail on Naimark's dilation theorem see Appendix A.

In our demonstration, the quantum computer hardware supports projective measurements on each qubit; thus, the chosen measurement on the extended Hilbert space will be $E_i^+ = |e_i\rangle\langle e_i|$ where $|e_i\rangle$ denotes each of the four computational basis
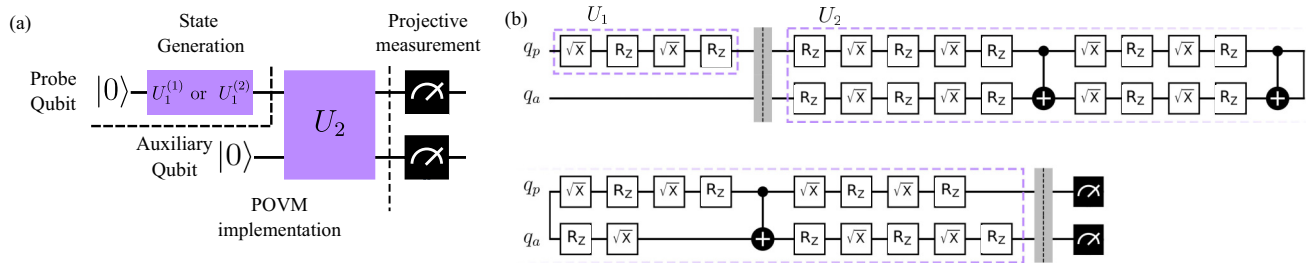
FIG. 3. (a) A schematic diagram of the quantum circuit for the demonstration. $U_1^{(1)}$ or $U_1^{(2)}$ is a single-qubit unitary used to prepare the appropriate probe state. U_2 implements the optimal ST-POVM. (b) An example of the circuit; all the circuits have the same structure. The dotted black lines separate the circuit into three steps, which correspond to U_1, U_2 and measurement gates in (a). R_Z denotes the parametric R_z gates with the corresponding input angles that realize U_1 and U_2.

states: $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$. By Naimark's dilation theory and an numerical approach of implementing the POVM, we obtained the required U_2 that fulfils Eq. (A3). By applying U_2 on the two-qubit state and making projective measurements along the four computational basis states of the quantum computer available, we performed the ST-POVM on the probe qubit.

C. Implementing on IBM quantum computer with autonomous package

The decomposition of these unitary U_1 and U_2 is handled by the autonomous algorithms provided by IBM's quantum computing package, Qiskit [37]. An example of the native circuit running on one of the IBM quantum systems, `ibmq_jakarta`, is shown in Fig. 3(b). We use this quantum system to generate all experimental results in this paper. For a single-qubit unitary, Refs. [38,39] provide a scheme for the device with restricted sets of single-qubit operations to construct arbitrary single-qubit unitary gate. For a two-qubit unitary, Ref. [26] shows that with three CNOT gates acting on the two qubits and some other single-qubit unitary operation, one can construct arbitrary two-qubit unitary operators.

1. Readout noise and measurement gate error mitigation

Readout noise refers to the noise of the measurement gate. With up to 10% error for assigning the $|0\rangle$ state as $|1\rangle$ or vice versa for some qubits on the IBM cloud computing devices, the readout noise posed a significant challenge to the accuracy of our repeated measurements. Thus, it is crucial to mitigate the readout noise produced by the measurement gates. In our demonstration, we followed the error mitigation scheme proposed in Ref. [28] for minimizing readout noise. This error model proposes that the output frequencies vector F' of a set of measurements have a linear correlation to the true frequencies F, which would be obtained on an ideal quantum computer. The linear correlation is as follows:

$$F' = MF, \tag{8}$$

where M is a probability matrix that maps the ideal output F to the experimental output F'. The method we implement first generates the error matrix M, a nonunitary operator describing the error. A set of four calibration circuits are run measuring only trivially generated states, $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$, which are assumed to be error-free. Then we use the reverse correlation

M^{-1} to calculate what the true frequencies should be for the experimental runs:

$$F = M^{-1}F'. \tag{9}$$

Whenever we perform an offset calibration, we also generate the matrix M so that the matrix M is always up to date. This accounts for any possible drifting of hardware parameters.

2. Time-dependent calibration

For any localized measurement that demands great precision, calibration before the measurement is recommended as it reduces systematic errors which cannot be reduced by repeated estimation, unlike random error due to statistical fluctuations. The execution of circuits on the quantum computer lasted only 12 h. However, the entire execution is split into four time segments due to the availability of IBM's quantum computers. Thus, the possible time-dependent variables, such as the drift of hardware parameters and updated hardware calibrations, may affect the raw output of the circuits. To account for this time-dependent issue, we performed calibrations regularly and compensated for the systematic offset for each corresponding main run.

The calibration runs use the same U_2 (ST-POVM) but slightly different U_1 (state preparation) compared to the main runs. For precise calibration, we need to remove any possibility of circuit degeneration. The degeneracy of the circuit refers to the reduction of circuit complexity when initializing different states. For example, initializing a general state such as $1/\sqrt{2}(|0\rangle + i|1\rangle)$ requires a more complex circuit than initializing state $|0\rangle$. To avoid degeneracy while keeping the control protocols on the quantum computers autonomous and robust, we adopted a semirandomized approach to the initialization step of the main runs.

For every two of the main runs, we generate one calibration run consisting of 10 randomized mixed states in the unit sphere of radius 0.1 centered around the true state. We also randomized the exact mixed state in a small sphere of radius 0.01 to eliminate any degeneracy of state generation. The exact randomization scheme of the main run is as follows: We choose a mixed state for a set of two main runs randomly following a uniform distribution inside a small sphere of radius 0.01 centered around the mixed state we intended to simulate. Then we initialize and measure the state with the exact procedure also implemented on the calibration circuits.
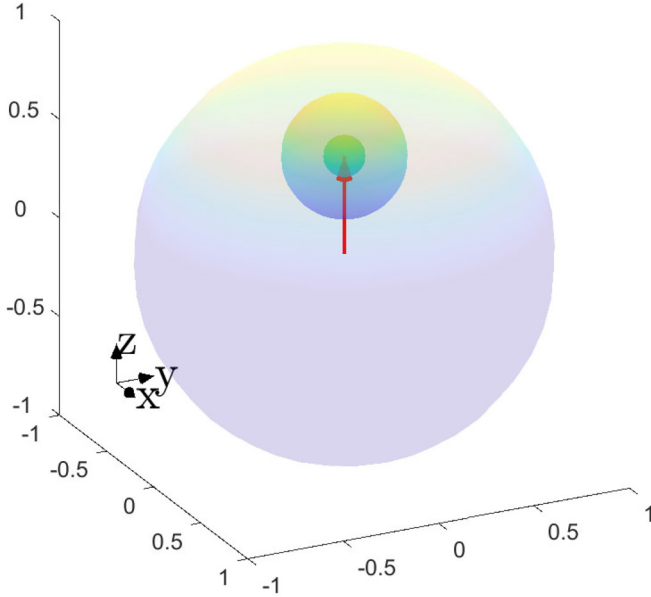
FIG. 4. The ranges of the semirandomized initialization scheme. The outermost sphere is the Bloch sphere, and the medium size sphere shows the range of randomly chosen calibration states. These calibration initial states will be used to find the offsets on parameter θ. The smallest sphere shows the range of randomly chosen for the main runs. The medium sphere and the small sphere are centered around the intended true state shown by the vector arrow. The sizes of spheres in this graph are exaggerated for improved visual clarity. The exact radii of these spheres are given in the main text.

FIG. 5. The effective MSE using 100 probe states characterized by the length of the Bloch vector, r. The result is obtained from a total number of 180 000 identical probe states for each data point via repeated subsampling. The MSE is scaled to measure the effectiveness of the method per probe state, i.e., MSE $= N$ MSE$_N$ where N is the number of probes. The data points are the combined scaled MSE of the estimated state $\hat{\rho}(\hat{\theta}_x, \hat{\theta}_y, \hat{\theta}_z)$ when the group's size is 100. The blue (upper) line shows the theoretical MSE for a perfect SIC-POVM. The red dotted line indicates the N-H bound that lower bounds the MSE of any possible separable measurement. Error bars are one standard deviation obtained via bootstrapping.

The ranges of the semirandomized initialization are shown in Fig. 4. The possible error generated from the semirandomized approach will be discussed in the next section.

To correct the systematic error, we use the calibration runs to find the offsets for estimating each component of the Bloch vector. This is done by fitting the known input Bloch vector components to the estimated components via linear regression, with a fixed gradient of 1, to ensure that our estimates are locally unbiased. Next, we subtract the systematic offsets, $\Delta\hat{\theta}$ in Fig. 2, from the corresponding estimate of parameters obtained from the main runs. The method described above provided significant improvement to the overall accuracy, especially for estimates from large ensembles of identical copies of the qubit state.

V. RESULTS

In this section, we will present the results of the demonstration. We will also address our statistical method to analysed the raw results. For the very last stage of the demonstration shown as "Correcting estimated states by calibration offsets" in Fig. 2, the input is the measurement results after being corrected by measurement gate error mitigation. This result is a series of random POVM outcomes chosen from the four possible outcomes X_1, X_2, X_3, and X_4 (such as $\{X_1, X_2, X_1, \ldots, X_4, X_1\}$). The four possible outcomes correspond to the four POVM elements, and the length of the series is equal to the total number of probe states involved in the demonstration. I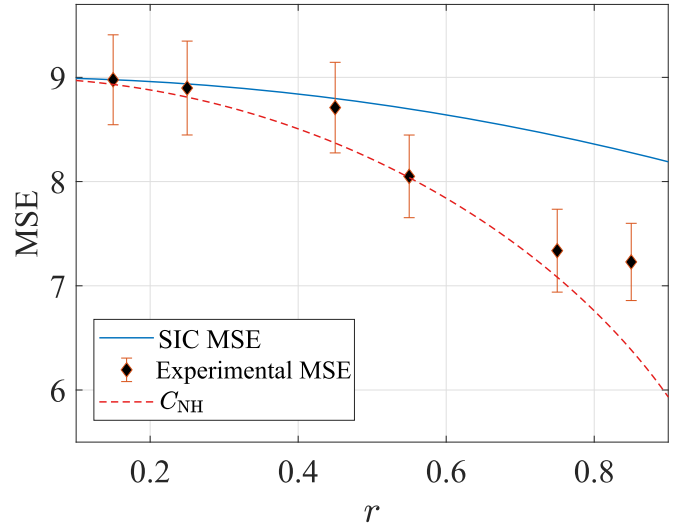n a practical setting, the number of probes may range from a few tens to millions. To comprehensively analyze the performance and the potential of the ST-POVM, we analyze the series of outcomes for the probe states by dividing them into groups of the same sizes, N.

From each group with N outcomes, we calculate the frequency with which each POVM element is observed $F = (f_z, f_1, f_2, f_3)^{\mathsf{T}}$. Using Eq. (5), we calculate the estimated value of θ, $\hat{\theta}$, for each group. Then we calculate the MSE for each group and report the average MSE as the result. In this section, as we are working with multiple probes and the precision limit given by the N-H bound is MSE $\geqslant C_{\text{NH}}/N$, we expect the MSE to scale down with an increased number of probes (in each group) MSE$_N \rightarrow$ MSE$_1/N$. In a practical setting, a large number of probes will be needed to obtain a satisfactory estimate of the state. Here we first present results using 100 probes to demonstrate the effectiveness of the ST-POVM for a lower number of probes. We then proceed to establish a model for the performance, in terms of MSE, when varying numbers of probes are used to estimate θ.

A. MSE of 100-probe estimation

We implemented the ST-POVM scheme on IBM's quantum computers. We obtained a set of results that shows the advantages and limitations of the measurement method we proposed in an experimental setting. The MSE in our demonstration conducted using 100 probe states is shown in Fig. 5. This result demonstrates that the POVM we suggest can saturate the N-H bound for some instances. Note, however, that in
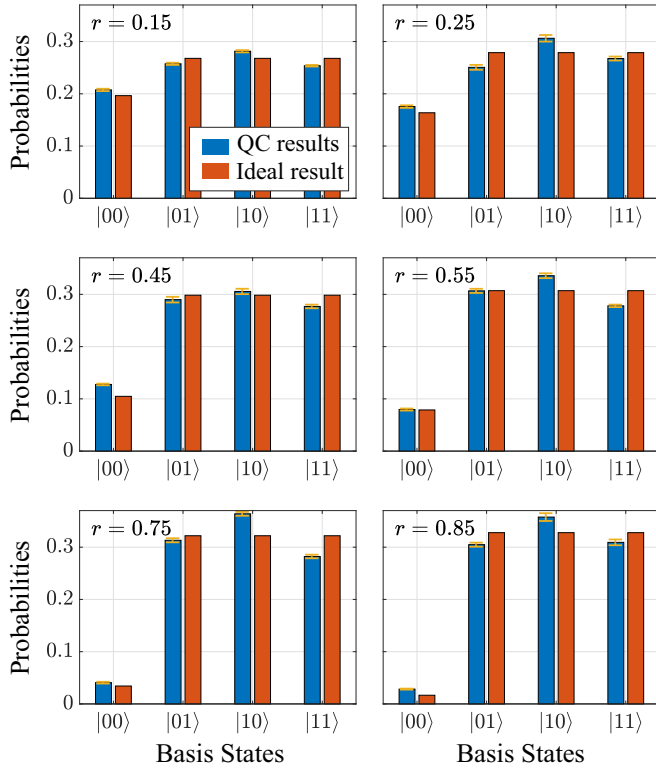
FIG. 6. The deviation of the output state from the ideal state is demonstrated by the differences in probability for each measurement outcome. The probabilities of the outcomes marked by the four basis states are labeled by "Quantum Computer (QC) result." The probabilities of ideal outcomes are produced on simulated noiseless circuits and shown as the "Ideal result." The deviation of the amplitude from the ideal amplitude accounts for the time drift of the experimental setting across six sets of measurement. The error bar in the diagrams indicates the estimated statistical error of sampling from a multinomial distribution. Each set of measurements is performed within 30 min on IBM's quantum devices.

TABLE I. The result of the model fitting according to Eq. (10).

r	C_{NH}	C	δ
0.15	8.932	8.939 ± 0.096	$2.72 \pm 0.92 \times 10^{-5}$
0.25	8.810	8.937 ± 0.066	$33.41 \pm 0.63 \times 10^{-5}$
0.45	8.370	8.785 ± 0.065	$7.45 \pm 0.62 \times 10^{-5}$
0.55	8.038	8.108 ± 0.115	$14.06 \pm 11.04 \times 10^{-5}$
0.75	7.083	7.358 ± 0.072	$2.10 \pm 1.69 \times 10^{-5}$
0.85	6.385	7.225 ± 0.068	$13.07 \pm 0.65 \times 10^{-5}$

We did not conduct any further investigation on the sources of errors with the POVM implementation as our focus is on the realization of the ST-POVM. To confirm and investigate the error contribution from the POVM implementation, we suggest a simpler measurement scheme such as the one shown in [40], which can potentially reconstruct any middle state in different stages of the circuit with a much lower error rate compared to the existing errors in the circuit.

B. MSE with varying number of probe states

In the previous section, we presented the MSE analyzed in groups of 100 probe states. For practical use, we would like to consider various different numbers of probe states. Accounting for the possible error generated during the demonstration, we propose the following model for how the MSE scales with the number of probes used:

$$\mathrm{MSE}(N) = \frac{C}{N} + \delta, \tag{10}$$

where C is the experimental bound which depends on the specific implementation, and δ is a constant that accounts for the systematic error and drift of the experimental settings. As the N-H bound poses a fundamental limit to the state estimation problem, the C value in the model above should not be less than the N-H bound, i.e., $C \geqslant C_{\mathrm{NH}}$. In an ideal experiment, $C = C_{\mathrm{NH}}$ and $\delta = 0$.

Given a finite number of measurement outcomes ($N_{\mathrm{total}} = 180\,000$ in our case), to accurately fit the model, we utilized repeated subsampling to get many more instances of the MSE(N) than just N_{total}/N sets. To conduct subsampling, we randomly divided the entire sample into N_{total}/N sets and used the N outcomes in each set to generate an instance of MSE(N). Then we repeated the division of the entire data set multiple times to generate a greater number of separate instances of MSE(N) values, from which an estimate of the average MSE(N) will be calculated by averaging over all instances of MSE(N). We observed no further convergence after reaching 10\,000 instances of MSE(N) for all probe states. Thus, we use 10\,000 instances for all the results presented in this section.

In Fig. 7 we fit a modified version of Eq. 10 shown below by linear regression

$$N\,\mathrm{MSE}(N) = C + N\,\delta. \tag{11}$$

With an increasing number of estimations, we can see that the specific MSE, $N\,\mathrm{MSE}(N)$, increases almost linearly. This show that our model is a valid description of the results. In Table I, for different probe states characterized by r, the values

some cases, the average experimental MSE deviates from the theoretical prediction. We provide a brief discussion of this experimental error in the MSE obtained from sets of 100 probe states in this section, and a further elaborated analysis regarding the error in the experimental MSE using larger numbers of probes will be presented in Sec. V B.

By analyzing the entire series of POVM outcomes of 180\,000 measurement for each r, Fig. 6 hints at the kind of error we encounter when operating the quantum computer. For all six different quantum states, the probability amplitudes of the $|10\rangle$ state are consistently higher than the amplitudes of the ideal state, and the amplitudes of the $|11\rangle$ state are consistently lower than the ideal amplitudes. This could be the consequence of a systematic gate error or a gate alignment problem, but a definite conclusion about the origin of the error could not be made based on the data. As we have already performed measurement gate error mitigation before getting the result in Fig. 6, we suspect that the influences of the measurement gate error on the data as a whole are relatively small compared to the observed deviation from the ideal result of up to 10%.
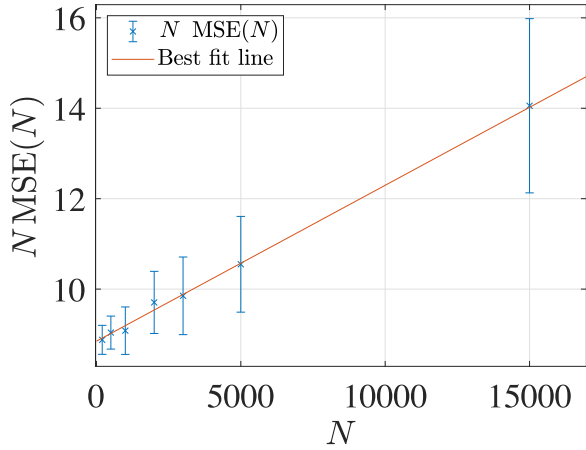
FIG. 7. A demonstration of the model fitting according to Eq. (11) for $r = 0.25$. The error bar shows the standard deviation of MSE for each number of probes obtained by bootstrapping.

of the N-H bound, C_{NH}, and the values of the fitted parameters, C and δ, are shown. We attribute the deviation of C from the N-H bound to the imperfect implementation of the POVM, and we attribute the value δ to imperfect calibration and drift of the experimental setup. In the following paragraphs, we will try to explain and validate the above claim.

Due to the experimental imperfections, the POVM that is physically implemented by the imperfect circuit differs from the theoretical optimal POVM we proposed. As long as we treat all the quantum gates as unitary gates, the imperfect circuit always corresponds to some POVM acting on the first qubit. We can see in Fig. 6 that the probabilities of arriving at the four different POVM outcomes differ from the ideal probabilities calculated from the optimal POVM. Assuming that the initialization step gives negligible errors, this indicates the implementation of the POVM can be slightly different from the proposed POVM. This imperfect POVM is less optimal, and the variance resulting from the suboptimal POVM will be greater than the N-H bound by an amount depending on the degree of error in the circuit.

We propose that the major contributor to the values of δ is the statistical error of the calibration runs. As we use a finite number of qubits for the calibration data, the calibration procedure will not be perfectly accurate. Thus, there is always a finite gap between the correction proposed by the calibration runs and the actual deviation between the means of the main measurement and the true values. Repeating the main runs will not eliminate this systematic error, so the effect of the slight inaccuracy of calibration will contribute to the constant term in Eq. (10), δ. We can verify this by calculating the expected standard deviation based on the finite amount of data used in calibration. For every 30 000 data points for the main measurement, we run 150 000 calibration measurements. The estimated standard deviation from the $n_{calib} = 150\,000$ calibration runs will be approximately $C_{NH}/n_{calib} = 3 \times 10^{-5}$ to 7×10^{-5}, which is on the same scale as the δ listed in Table. I.

In addition to the statistical error of calibration, we propose other possible contributors to δ that are comparatively harder to estimate. One of them is random fluctuations of the experimental circuit (including the initialization and the

POVM implementation) in terms of its average effect on an input state, i.e., fluctuation of the physical parameters when the circuit is implemented on the quantum computer. The fluctuation of measurement gate performance contributes to δ as well because the independent and identical performance of the measurement gate is one of the premises of the measurement error mitigation scheme discussed in Sec. IV C 1. Error reports based on the method of randomized benchmarking [41] can be fetched from the database of IBMQ. Still, the reports available based on the randomized method are insufficient for analyzing the contribution to δ from the circuit. Other possible contributors can be the drift of experimental settings across different times and the statistical error in the result on the original 180 000 probes before subsampling.

VI. COMPLEMENTARY SIMULATIONS

The method above and the N-H bound are designed to solve a localized problem, i.e., when the state is known relatively accurately, and we are trying to infer small changes in the state. However, in an alternative scenario, the qubit state can be completely unknown, or we know only approximately what the state might be. For such purposes, it is neither possible nor efficient to apply the ST-POVM measurement directly, as the orientation of the ST-POVM depends on the state. Thus, we present simulations of two adaptive methods in this section to demonstrate the potential usefulness of the ST-POVM in the scenario where we have some limited amount of prior knowledge. Adaptive approaches to tomography have been considered before both theoretically [40] and experimentally [42–44]. In this section, we will first illustrate the simulations of an adaptive measurement scheme on states assumed to be completely unknown. Then we will consider optimal measurements for differing degrees of prior knowledge. We show that in the limit of a large number of probes, with an adaptive method utilizing the ST-POVM, the MSE approaches the N-H bound. We also show that as the amount of prior knowledge increases, the optimal measurement tends to the ST-POVM.

A. Adaptive POVM using the SIC and ST-POVM

The simulation discussed in this section will show an adaptive scheme which gives an effective MSE that saturates the N-H bound asymptotically with the total probe number. We present a general guideline for implementing a two-step adaptive method, which uses the SIC-POVM initially before switching to the ST-POVM. We show that the optimal number of preliminary steps, given a fixed amount of total probe states, N, is proportional to $\sqrt{N}$ when N is large. This section demonstrates the effectiveness of applying the ST-POVM on a more general problem of an unknown qubit state.

To measure a completely unknown state, one approach is an adaptive measurement scheme where the measurement of the next step will depend on the measurement result of the steps before it. In this way, the measurement applied for later steps will have reduced variance compared to the measurements applied before. In this example, we consider a two-step adaptive measurement where the SIC-POVM will serve as the preliminary measurement, and ST-POVM will be the secondary fine measurement. As the optimal ST-POVM

is state-dependent, in the second stage, the ST-POVM chosen will be the most suitable measurement from the family of ST-POVMs based on the estimated state given by the preliminary step. After the two stages are completed, the final estimated state will be a weighted average of the result of the preliminary stage and the secondary stage. The final estimation of the parameters will be

$$\hat{\boldsymbol{\theta}} = W\hat{\boldsymbol{\theta}}^{(1)} + (1-W)\hat{\boldsymbol{\theta}}^{(2)}, \qquad (12)$$

where $0 \leqslant W \leqslant 1$ is the weight of the preliminary measurement. The superscripts (1) and (2) denote the first step and the second step of the adaptive measurement. The MSE that we aim to minimize for optimizing this adaptive method is

$$\text{MSE} = E[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^2], \qquad (13)$$

where $\boldsymbol{\theta}$ is the probe state. With a given number of probes N to complete the two-step adaptive measurement, we will need to allocate an appropriate amount of probes to the preliminary measurement and assign an appropriate value to the weight W to minimize the MSE.

To implement simulations with up to 10^{10} total available probes, we choose to use an approximated analytical approach instead of Monte Carlo sampling. We will justify the approximation used in this approach below and, in Fig. 9(a) below, a referential set of data is given to further demonstrate the asymptotic accuracy of the approximated analytical approach. For this analytical approach, the central limit theorem that applies to a large number of repeated sampling facilitates our analysis. For a large number of available probe states, by the multidimensional central limit theorem [45], the distribution of the estimated state for each stage can be well approximated by a multivariate normal distribution centered around the true state with covariance matrix $\Sigma_N = \Sigma/N$. However, as shown in Fig. 9(a), the approximation tends to introduce more error in the low probe number regime and the results obtained converge to the Monte Carlo simulation in higher probe number.

The central limit theorem also leads to a simplification of finding the optimal weight W from Eq. (12). Assuming that the results obtained from both stages are unbiased, the estimated states for the two stages are uncorrelated on all three parameters θ_x, θ_y, and θ_z. For θ_j, this translates to

$$\text{cov}\big(\theta_j^{(1)}, \theta_j^{(2)}\big) = 0 \quad (j = x, y, z), \qquad (14)$$

where $\text{cov}(\,\cdot\,, \,\cdot\,)$ is the covariance between the two random variables. Note that the two outcomes from the two stages remain statistically dependent on each other as the parameter of the second stage depends on the result of the first stage. Let the MSE of the preliminary stage be MSE_1, and the MSE of the secondary stage be MSE_2. By Eq. (14), we simplify the total MSE based on Eq. (12) to be

$$\text{MSE} = W^2\text{MSE}_1 + (1-W)^2\text{MSE}_2. \qquad (15)$$

The MSE is minimized by setting

$$W = \frac{\text{MSE}_2}{\text{MSE}_1 + \text{MSE}_2}. \qquad (16)$$

Using the optimal value of the weight W, we can find the average MSE for different numbers of allocated preliminary measurements by Eq. (12). For a given preliminary outcome
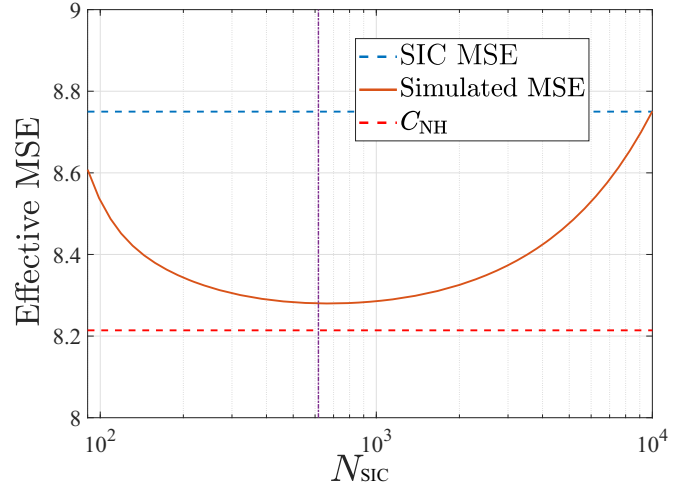

FIG. 8. The effective MSE given 10 000 probe states, $N \text{ MSE}(N)$, for different number of allocated preliminary steps, N_{SIC}. There is an optimal number of preliminary step N'_{SIC} that gives a minimum MSE. The MSE obtained when only SIC-POVM is used is indicated by the upper flat line. The lower flat line indicates the N-H bound. The optimal N_{SIC} is indicated by the vertical line.

$\hat{\boldsymbol{\theta}}^{(1)}$, we have $\text{MSE}_1 = (\hat{\theta}^{(1)} - \theta)^2$. If the function $\text{MSE}_2(\hat{\theta}^{(1)})$ is known, given initial parameter, θ, we are able to evaluate both MSE_1 and $\text{MSE}_2(\hat{\theta}^{(1)})$. By Eq. (15) and Eq. (16), a MSE value that depends on the preliminary estimate $\hat{\theta}^{(1)}$ can be constructed. The method that we use to find the function, $\text{MSE}_2(\hat{\theta}^{(1)})$, is a change of the coordinate system. Based on the preliminary estimate, $\hat{\theta}^{(1)}$, we switch to a coordinate system such that the predicted state lies along the negative z axis. This means that the optimal POVM, used in the second stage, is oriented along the positive z axis. The MSE obtained in this setting is equal to the MSE in the original coordinate system because the MSE derived from the trace norm between states is invariant under a unitary rotation. Then, after transforming $\hat{\theta}^{(1)}$ back to the original coordinate system, we have the function $\text{MSE}_2(\hat{\theta}^{(1)})$.

Last, we will need to iterate through different numbers of probes allocated for the two stages to find the optimal number of preliminary measurements together with the minimum MSE achieved by the two-step adaptive method. Thus, we perform numerical integration with a fixed number of preliminary measurements using the SIC-POVM, N_{SIC}, as shown below:

$$\overline{\text{MSE}} = \iiint_S \text{MSE}(\hat{\boldsymbol{\theta}}^{(1)}) \, \text{pdf}(\hat{\boldsymbol{\theta}}^{(1)}) \, d\hat{\boldsymbol{\theta}}^{(1)}, \qquad (17)$$

where $\text{pdf}(\hat{\boldsymbol{\theta}}^{(1)})$ is the probability density function of the multivariate normal distribution that we use to approximate the discrete distribution of the preliminary outcomes. S is the entire parameter space with $|\hat{\boldsymbol{\theta}}^{(1)}| < 1$.

We provide the detailed graphed result for one sample qubit state in Fig. 8 and Fig. 9, and the results of five other states are presented in Table II. For the detailed examples, we use the state $\rho = (I + \vec{\sigma} \cdot \boldsymbol{\theta})/2$ where $\boldsymbol{\theta} = (0, 0, 0.5)^\top$ is the vector of parameters. For this specific state, Fig. 8 shows the change in MSE when allocating different numbers of qubits to the SIC-POVM preliminary experiments when the total
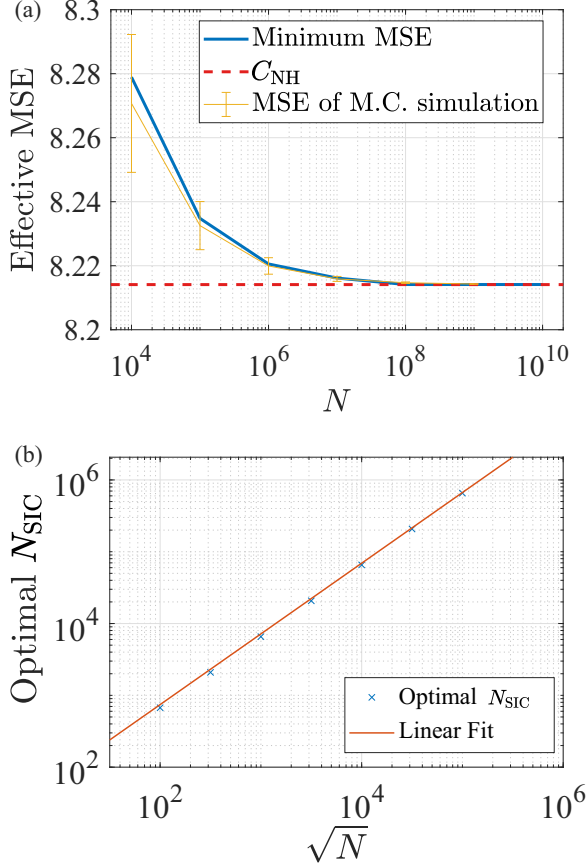
FIG. 9. (a) Asymptotic behavior of the effective MSE when the number of total probes approaches infinity. A set of data generated from a Monte Carlo simulation is shown for reference. The error bar indicates the statistical error of the Monte Carlo simulation. The dotted red line shows the N-H bound. (b) The change of the optimal preliminary step number, N_{SIC} with different numbers of total probes, N. The linear fit verified Eq. (18).

number of qubits is 10 000. It shows that there is one and only one minimum MSE. Either more preliminary steps or fewer preliminary steps will increase the MSE. For 10 000 qubits, the optimal number of preliminary SIC-POVM experiments is 673 (without accounting for modeling error). In this case, the effective MSE is only 0.8% above the N-H bound at 8.28. In contrast, for a pure SIC-POVM approach, the effective MSE will be more than 6.5% above the N-H bound at 8.75. For a larger number of total probes, the effective MSE is closer to the N-H bound and will eventually converge to the N-H

TABLE II. The coefficient B of various states on the z axis. The value of B grows to infinity when the θ_z value approaches 0, 1, and -1. Thus, we did not give a value of B for $\theta_z = 0$. The algorithms we have designed can solve the problem of finding the optimal number of preliminary steps with arbitrary parameters. This table serves as a general guideline of how many preliminary numbers are preferred and as a demonstration of our algorithms.

θ_z	-0.9	-0.5	-0.3	0	0.3	0.5	0.9
B	8.39	6.76	10.35	NA	10.28	6.55	5.24

bound when $N \to \infty$. This asymptotic behavior is shown in Fig. 9(a). We also found that for this particular two-step adaptive measurement on the state specified above, the optimal number of preliminary experiments, which we denote N'_{SIC}, can be approximated by

$$N'_{SIC} = B\sqrt{N}, \qquad (18)$$

where B is a constant, which depends on the state's purity and orientation. Figure 9(b) shows the change of N'_{SIC}. By linear regression, we found that the $B = 6.55$ is the coefficient.

We have included the suggested coefficient B for other states when the ST-POVM is aligned along the z axis in Table II. The initial SIC-POVM in our simulated experiment always points downward, similar to the "squashed" tetrahedron shown in Fig. 1. We also found that the value of B grows when the θ_z value approaches 0, 1, and -1. For 1 and -1, as the ST-POVM becomes more specialized, it requires more preliminary steps to ensure a correct input estimated state until at $r = 1$, the formulation introduced in Sec. III breaks as the bottom left entry, $b + r$, in the estimator matrix $\mathcal{E}$ in Eq. (5) becomes infinity. This incompatibility of the ST-POVM with pure states is expected because setting the preliminary expectation of $r = 1$ eliminates the uncertainty of estimating $\hat{\theta}$ in the radial direction. For example, the state $|0\rangle$ has a vanishing z direction variance:

$$\mathbb{E}[(\hat{\theta}_z - \theta_z)^2] = \sum_{i=0,1} p_{|i\rangle}(\theta_{z,|i\rangle} - 1)^2 = 0. \qquad (19)$$

This changes the three-parameter estimation problem to a two-parameter problem. Thus, setting $r = 1$ violates the assumption of estimating three parameters simultaneously for the ST-POVM estimation scheme. For $\theta_z = 0$ the SIC-POVM is already the optimal POVM. Thus, any further adaptive measurement will be less optimal.

Note that Table II is computed assuming that the state is known. In an experiment with a completely unknown probe state, it will not be possible to use the optimal B value or the optimal weight in Eq. (16). For an experiment with a flexible amount of probe states, one could use approximately $20\sqrt{N}$ of the probes for the preliminary SIC-POVM without inducing significant overhead for the ST-POVM while maintaining high suitability for the follow-up measurement. In this simulation, the relative orientation between the initial state and the SIC-POVM will change the coefficient B. The reason for this is that although estimations given by the SIC-POVM give a uniform MSE $= 9 - r_p^2$ [Eq. (2)] regardless of the orientation of the POVM, the distribution of SIC-POVM estimates is not invariant with a unitary rotation. This can be explained by the discrete structure of the SIC-POVM. We should bear in mind that an arbitrary rotation of the tetrahedron representing the SIC-POVM does not guarantee the orientation invariance like a rotation acting on the sphere does. This is what gives rise to the asymmetry in Table II.

B. Evolving ST-POVM for general adaptive measurement

In the previous simulation, we considered a simple two-step adaptive measurement. This MSE approached the quantum limit, the N-H bound, asymptotically when the number of available probe states tends to infinity. Thus, there is

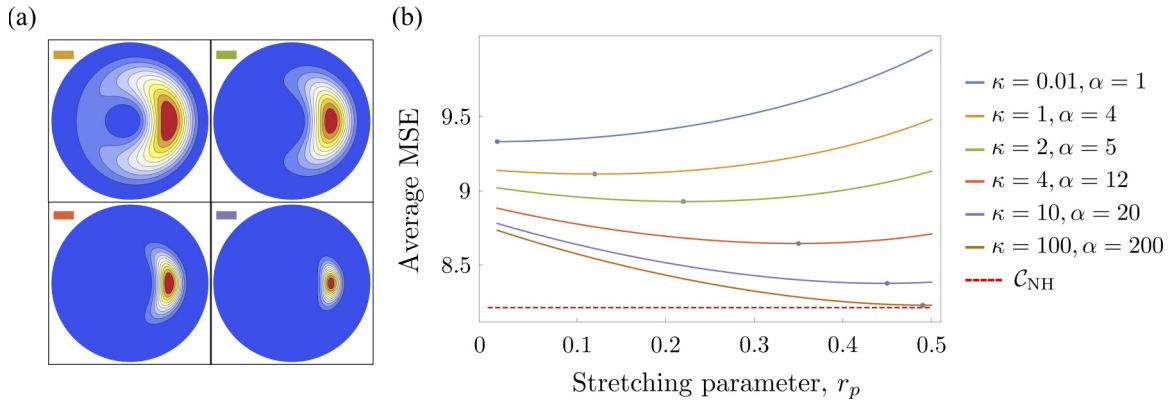(a)   (b)

FIG. 10. Minimizing the average MSE given a certain prior distribution of states. (a) Examples of the prior distribution of states for a cross section of the Bloch sphere in the *x-y* plane. Legend is the same as in (b). (b) Average MSE as a function of the stretching parameter, r_p, for different prior distributions. The blue dots indicate the minimum MSE for each prior distribution. As the distribution gets more sharply peaked, the optimal r_p tends to the r, which saturates the N-H bound, and the MSE tends to the limit set by the N-H bound.

little space for other more sophisticated adaptive schemes to improve upon. However, when the total number of probes is low, the evolving adaptive measurement utilizing the Bayesian estimation [46] procedure is more advantageous. During an evolving adaptive measurement, the POVM implemented is adjusted more frequently according to the results obtained from the preceding steps. In such an adaptive measurement scheme, the stretching parameter[3] r_p (first shown in Sec. III) is updated over time so that, at each stage, the optimal POVM is being used.

In each iteration of an evolving adaptive measurement, a fixed algorithm will decide the most suitable POVM based on the prior distribution of states given from the previous measurement results. The objective of the entire adaptive procedure will be to improve the sharpness of the posterior distribution so that the uncertainty of the estimated state is minimized. After a POVM is decided, it will be applied to one of the probes. Using the results of this POVM and following the formulation of Bayesian estimation, one could obtain the posterior distribution, which will serve as the prior for the next iteration. In this section, we will give a paradigmatic example of finding the most suitable POVM for a given Bayesian prior distribution of state based on the minimization of Bayesian risk [47]. We emphasize that this is different to simulating the full adaptive measurement scheme. This section merely shows what is the optimal POVM given a certain prior distribution.

For the purpose of this demonstration, we assume that the prior distribution is invariant under a rotation around its radial axis. Thus, the state distribution under the spherical coordinate system, $f(r, \lambda, \phi)$, can be decomposed to a separable form,

$$f(r, \lambda, \phi) = f_r(r) f_{\lambda\phi}(\lambda, \phi). \tag{20}$$

For simplicity, we use the 3D Von Mises-Fisher distribution [48] for the probability density function $f_{\lambda\phi}$ with a unified

concentration parameter κ for all values of r. For the distribution along the radial direction, f_r, we use the Beta distribution with two positive shape parameters α and β. We set $\beta = \alpha(1 - r)/r$ where r is the distance from the center of the distribution to the origin of the Bloch sphere.

For a given prior distribution, we will assume that we always point our ST-POVM along the direction of the distribution's radial axis. The Bayesian risk, $\bar{\epsilon}(r_p)$, when using a POVM with parameter r_p can be calculated as follows [47]:

$$\bar{\epsilon}(r_p) = \int dr \, d\lambda \, d\phi \, \text{MSE}(\boldsymbol{\theta}_p, r, \lambda, \phi) f(r, \lambda, \phi), \tag{21}$$

where $\text{MSE}(\boldsymbol{\theta}_p, r, \lambda, \phi)$ is the MSE of a measurement on the probe state $\rho(r, \lambda, \phi)$, where (r, λ, ϕ) are the spherical coordinates of the probe state in the Bloch sphere, when the ST-POVM is orientated along the direction of $\boldsymbol{\theta}_p$ with parameter $r_p = |\boldsymbol{\theta}_p|$. Due to the symmetry of the prior distribution, we choose the orientation of the POVM to be the same as the orientation of the distribution. Thus, the chosen ST-POVM and the distribution are invariant when rotating around the axis of $\boldsymbol{\theta}_p$. Consequentially, without loss of generality, we write the Bayesian risk on the left-hand side as a function of r_p. The equation above indicates that the Bayesian risk, $\bar{\epsilon}(r_p)$, is the average of the MSE we first defined in the introduction over the distribution of quantum states using the same POVM with parameter r_p.

We compute the Bayesian risk for a range of distributions in the form of Eq. (20), centered on a state with $r = 0.5$. Examples of this prior distribution are shown in Fig. 10(a). For each distribution, to find the best ST-POVM with parameter r_p, the Bayesian risk is calculated for a range of different stretching parameters.

The results of the calculation are shown in Fig. 10(b). It shows that with an increase of the parameter κ and α, the minimum Bayesian risk approaches the N-H bound. As an increase in κ and α implies a more localized prior distribution, the results also suggest that the ST-POVM is applicable and asymptotically optimal in an evolving adaptive measurement scheme that utilizes Bayesian estimation. It is also worth noting that the ST-POVM should be close to the optimal measurement in terms of MSE for the later stages (when the

---

[3]The stretching parameter, r_p, may be seen as a representation of the degree to which the squashed tetrahedron POVM is squashed. When $r_p = 0$, there is zero squashing and the ST-POVM converges to the SIC-POVM.

uncertainties of the priors are low) of any evolving adaptive tomography scheme on the single-qubit state. As the optimal measurement with zero prior knowledge was found to be the ST-POVM with a stretching parameter $r_p = 0$, i.e. a SIC-POVM, this framework offers a nice way to visualize the evolution of the optimal measurement.

VII. CONCLUSION

In this paper we have mapped qubit tomography to a quantum multiparameter estimation problem. Based on the multiparameter estimation bound for separable measurements (N-H bound), we found the theoretically optimal measurement for the local tomography problem, which we call the ST-POVM. As such local measurements that characterize the qubit state have direct implications for quantum computing, we performed a demonstration on an IBM superconducting quantum computer, and the result shows the saturation of the N-H bound. In Sec. VI we argued that the ST-POVM could also be effective in the nonlocalized setting by showing the results of two complementary simulations of two different types of adaptive measurement.

In this paper, single-qubit state estimation is the focus. However, to better evaluate the performances of a more general quantum circuit, one needs to consider more than one qubit. We will briefly discuss two possible ways of extending the scope of the demonstration. Previous studies have shown that entangling measurements help reduce the experimental errors [16,49,50] and can be useful in state discrimination [51]. Thus, the first extension can be to evaluate more than one copy of the qubit state at once, as the entanglement between multiple probes state will be able to reduce the experimental errors further. While the exact method that attains the N-H bound for multiple copies of a state $\rho^{\otimes n}$ remains unknown, the ST-POVM with its corresponding estimator introduced in this paper may have implications in finding a series of POVMs exploiting the symmetry of a multicopy quantum system that saturates the N-H bound. For the SIC-POVM, a two-copy extension has been tested on the photonic system [16], which suggests a possible generalization of the ST-POVM to multiple copies of a qubit system.

On the other hand, a less complicated extension of the results in this research would be to reduce the complexity by considering two of the three parameters. The search for the two-parameter optimal POVM is promising because it has been shown that there is an equivalent of the SIC-POVM in this problem, the conditional SIC-POVM [32,52]. It is optimal when the figure of merit is the average Hilbert-Schmidt norm deviation. Alternatively, for a general Hilbert space of dimension d with $d(d+1)/2$ free parameters, there might be a conditional optimal POVM with a similar structure as the ST-POVM when three of the $d(d+1)/2$ parameters are unknown, similar to the three parameters for the qubit tomography problem, θ_x, θ_y, and θ_z.

In summary, the result of our demonstration validates the theory and illustrates the power of a general quantum com-

puter in implementing POVM measurements. The scheme we developed may also provide an opportunity for understanding the characteristics of the qubits and the quantum gates by facilitating a more well-rounded assessment of any single-qubit output state.

APPENDIX: NAIMARK'S DILATION THEOREM

Given any POVM $\{\Pi_i\}$, Naimark's dilation theorem [19,35,36] guarantees the existence of an isometry $\tilde{V} : \mathcal{H} \to \mathcal{H}^+$, and a projective measurement $\{E_i^+\}$ on an extended Hilbert space $\mathcal{H}^+$ (with the same number of elements), which satisfy

$$\tilde{V}^\dagger E_i^+ \tilde{V} \rho = \Pi_i \rho \quad \text{for } \forall \rho. \tag{A1}$$

As we shall see later, the above equation is solvable for $\mathcal{H}^+$ being a two-qubit Hilbert space. In the following discussion, we will use a single auxiliary qubit to realize the POVM. Because the auxiliary qubit on a quantum computer is initialized at the $|0_a\rangle$ state, we consider an isometry $\tilde{V}$ with the following effect on a mixed state ρ:

$$\tilde{V}\rho = U_2(\rho \otimes |0_a\rangle\langle 0_a|). \tag{A2}$$

This simplifies Eq. (A1) to

$$\langle 0_a|U_2^\dagger E_i^+ U_2|0_a\rangle \rho = \Pi_i \rho \quad \text{for } \forall \rho. \tag{A3}$$

For projectors E_i^+, the unitary U_2 effectively rotates the hyperplane that the operator projects onto. Thus, we can write the operators on the extended Hilbert space $U_2^\dagger E_i^+ U_2$ as another projector, $|\psi_i^+\rangle\langle\psi_i^+|$. Now we force the operators acting on ρ on each side to be the same, and we have

$$\langle 0_a||\psi_i^+\rangle = \sqrt{r_i}|\psi_i\rangle, \tag{A4}$$

where r_i and $|\psi_i\rangle$ defines the ST-POVM in Sec. III. A numerical calculation can then be performed to find all four elements of orthonormal $\{|\psi_i^+\rangle\}$ given Eq. (A4). It is also worth noting that Eq. (A3) also guarantees that the measurement outcome distribution is unchanged by this process, which is crucial for the success of the estimation:

$$\text{tr}(\rho \otimes |0\rangle\langle 0|U_2^\dagger E_i^+ U_2) = \text{tr}(\rho \Pi_i) \quad \text{for} \quad \forall \rho. \tag{A5}$$

[1] C. D. Marciniak, T. Feldker, I. Pogorelov, R. Kaubruegger, D. V. Vasilyev, R. van Bijnen, P. Schindler, P. Zoller, R. Blatt, and T. Monz, Optimal metrology with programmable quantum sensors, Nature (London) **603**, 604 (2022).

[2] Y. Wu *et al.*, Strong Quantum Computational Advantage Using a Superconducting Quantum Processor, Phys. Rev. Lett. **127**, 180501 (2021).

[3] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[4] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, 2010).

[5] J. Řeháček, B. G. Englert, and D. Kaszlikowski, Minimal qubit tomography, Phys. Rev. A **70**, 052321, (2004).

[6] W. K. Wootters and B. D. Fields, Optimal state-determination by mutually unbiased measurements, Ann. Phys. **191**, 363 (1989).

[7] A. Czerwinski, Quantum state tomography with information-ally complete POVMs generated in the time domain, Quantum Info. Proc. **20**, 105 (2021).

[8] H. Zhu, SIC POVMs and Clifford groups in prime dimensions, J. Phys. A: Math. Theor. **43**, 305305 (2010).

[9] J. M. Renes, R. Blume-Kohout, and A. J. Scott, Symmetric in-formationally complete quantum measurements, J. Math. Phys **45**, 2171 (2004).

[10] C.-F. Li, D. Dong, G.-C. Guo, G. Xiang, and Z. Hou, Real-ization of mutually unbiased bases for a qubit with only one wave plate: Theory and experiment, Optics Express **23**, 10018 (2015).

[11] Y. Y. Zhao, N. K. Yu, P. Kurzyński, G. Y. Xiang, C. F. Li, and G. C. Guo, Experimental realization of generalized qubit mea-surements based on quantum walks, Phys. Rev. A **91**, 042101 (2015).

[12] Z. Bian, J. Li, H. Qin, X. Zhan, R. Zhang, B. C. Sanders, and P. Xue, Realization of Single-Qubit Positive-Operator-Valued Measurement via a One-Dimensional Photonic Quantum Walk, Phys. Rev. Lett. **114**, 203602 (2015).

[13] J. Bae, B. C. Hiesmayr, and D. McNulty, Linking entanglement detection and state tomography via quantum 2-designs, New J. Phys. **21**, 013012 (2019).

[14] J. Li and L. Chen, Entanglement criterion via general symmet-ric informationally complete measurement, J. Phys. A: Math. Theor. **55**, 015302 (2022).

[15] Z. E. D Medendorp, F. A. Torres-Ruiz, L. K. Shalm, G. N. Tabia, C. A. Fuchs, and A. M. Steinberg, Experimental charac-terization of qutrits using symmetric informationally complete positive operator-valued measurements, Phys. Rev. A **83**, 051801(R) (2011).

[16] Z. Hou, J.-F. Tang, J. Shang, H. Zhu, J. Li, Y. Yuan, K.-D. Wu, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Deterministic realization of collective measurements via photonic quantum walks, Nat. Commun. **9**, 1414 (2018).

[17] A. J. Scott, Tight informationally complete quantum measure-ments, J. Phys. A: Math. Gen. **39**, 13507 (2006).

[18] D. Petz and L. Ruppert, Efficient quantum tomography needs complementary and symmetric measurements, Rep. Math. Phys. **69**, 161 (2012).

[19] A. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (Springer, Basel, 2011).

[20] C. W. Helstrom, Minimum mean-squared error of estimates in quantum statistics, Phys. Lett. A **25**, 101 (1967).

[21] C. W. Helstrom and R. S. Kennedy, Noncommuting observables in quantum detection and estimation theory, IEEE Trans. Inf. Theory **20**, 16 (1974).

[22] H. Nagaoka, A new approach to Cramér-Rao bounds for quan-tum state estimation, *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*, edited by H. Masahito (World Sci-entific, Singapore, 2005), Chap. 8.

[23] L. O. Conlon, J. Suzuki, P. K. Lam, and S. M. Assad, Efficient computation of the Nagaoka-Hayashi bound for multiparameter estimation with separable measurements, npj Quantum Inf. **7**, 110 (2021).

[24] Z. Hou, H. Zhu, G. Y. Xiang, C. F. Li, and G. C. Guo, Achieving quantum precision limit in adaptive qubit state tomography, npj Quantum Inf. **2**, 16001 (2016).

[25] P. Horodecki, Ł. Rudnicki, and K. Zyczkowski, Five open prob-lems in quantum information theory, PRX Quantum **3**, 010101 (2022).

[26] F. Vatan and C. Williams, Optimal quantum circuits for general two-qubit gates, Phys. Rev. A **69**, 032315 (2004).

[27] Y. S. Yordanov and C. H. W. Barnes, Implementation of a general single-qubit positive operator-valued measure on a circuit-based quantum computer, Phys. Rev. A **100**, 062317 (2019).

[28] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, Mitigating measurement errors in multi-qubit exper-iments, Phys. Rev. A **103**, 042605 (2021).

[29] K. Yamamoto, S. Endo, H. Hakoshima, Y. Matsuzaki, and Y. Tokunaga, Error-Mitigated Quantum Metrology via Virtual Pu-rification, Phys. Rev. Lett. **129**, 250503 (2022).

[30] G. J. Mooney, G. A. White, C. D. Hill, and L. C. Hollenberg, Generation and verification of 27-qubit Greenberger-Horne-Zeilinger states in a superconducting quantum computer, J. Phys. Commun. **5**, 095004 (2021).

[31] F. B. Maciejewski, Z. Zimborás, and M. Oszmaniec, Mitigation of readout noise in near-term quantum devices by classical post-processing based on detector tomography, Quantum **4**, 257 (2019).

[32] D. Petz, L. Ruppert, and A. Szanto, Conditional SIC-POVMs, IEEE Trans. Inf. Theory **60**, 351 (2014).

[33] R. D. Gill and S. Massar, State estimation for large ensembles, Phys. Rev. A **61**, 042312 (2000).

[34] S. R. Varadhan, Large deviations, Ann. Probab. **36**, 397 (2008).

[35] M. Neumark, On spectral functions of a symmetric operator, Izv. Ross. Akad. Nauk. **7**, 285 (1943).

[36] I. Gelfand and M. Neumark, On the imbedding of normed rings into the ring of operators in Hilbert space, Rec. Math. [Mat. Sbornik] N.S. **54**, 197 (1943).

[37] Qiskit contributors, Qiskit: An open-source framework for quantum computing (2021), doi:10.5281/zenodo.2573505.

[38] A. Y. Kitaev, Quantum computations: Algorithms and error correction, Russian Math. Surveys **52**, 1191 (1997).

[39] C. M. Dawson and M. A. Nielsen, The Solovay-Kitaev algo-rithm, Quantum Inf. Comput. **6**, 81 (2006).

[40] F. Huszár and N. M. T. Houlsby, Adaptive Bayesian quantum tomography, Phys. Rev. A **85**, 052120 (2012).

[41] E. Magesan, J. M. Gambetta, and J. Emerson, Characterizing quantum gates via randomized benchmarking, Phys. Rev. A **85**, 042311 (2012).

[42] R. Okamoto, S. Oyama, K. Yamagata, A. Fujiwara, and S. Takeuchi, Experimental demonstration of adaptive quantum state estimation for single photonic qubits, Phys. Rev. A **96**, 022124 (2017).

[43] K. S. Kravtsov, S. S. Straupe, I. V. Radchenko, N. M. T. Houlsby, F. Huszár, and S. P. Kulik, Experimental adaptive Bayesian tomography, Phys. Rev. A **87**, 062122 (2013).

[44] D. H. Mahler, L. A. Rozema, A. Darabi, C. Ferrie, R. Blume-Kohout, and A. M. Steinberg, Adaptive Quantum State Tomography Improves Accuracy Quadratically, Phys. Rev. Lett. **111**, 183601 (2013).

[45] A. W. van der Vaart, *Asymptotic Statistics* (Cambridge University Press, Cambridge, 1998).

[46] C. P. Robert, *The Bayesian Choice* (Springer, New York, 2007).

[47] J. Rubio and J. Dunningham, Quantum metrology in the presence of limited data, New J. Phys. **21**, 043037 (2019).

[48] R. A. Fisher, Dispersion on a sphere, Proc. R. Soc. A **217**, 295 (1953).

[49] L. O. Conlon, T. Vogl, C. D. Marciniak, I. Pogorelov, S. K. Yung, F. Eilenberger, D. W. Berry, F. S. Santana, R. Blatt, T. Monz *et al.*, Approaching optimal entangling collective measurements on quantum computing platforms, Nat. Phys. **19**, 351 (2023).

[50] E. Roccia, I. Gianani, L. Mancino, M. Sbroscia, F. Somma, M. G. Genoni, and M. Barbieri, Entangling measurements for multiparameter estimation with two qubits, Quantum Sci. Technol. **3**, 01LT01 (2018).

[51] L. O. Conlon, F. Eilenberger, P. K. Lam, and S. M. Assad, Discriminating qubit states with entangling collective measurements (2023), arXiv:2302.08882.

[52] H. Ohno and D. Petz, Examples of conditional SIC-POVMs, Quantum Info. Proc. **14**, 3983 (2015).