



# Scalable algorithms for calculating power functions of random quantum states in the noisy intermediate-scale quantum era

Ruyu Yang <sup>1</sup>, Wencheng Zhao <sup>2,\*</sup> and Tingting Chen <sup>2,†</sup>

<sup>1</sup>*Graduate School of China Academy of Engineering Physics, Beijing 100193, China*

<sup>2</sup>*China University of Mining and Technology, College of Sciences, Beijing 100083, China*



(Received 27 December 2023; accepted 3 May 2024; published 14 June 2024)

This article focuses on the development of scalable and qubit-efficient algorithms for computing power functions of random quantum states. Two algorithms, based on the Hadamard Test and Gate Set Tomography, are proposed. We provide a comparative analysis of their computational outcomes, accompanied by a meticulous evaluation of inherent errors in the Gate Set Tomography approach. The second algorithm exhibits a significant reduction in the utilization of two-qubit gates compared to the first. Consequently, the second algorithm exhibits reduced susceptibility to noise. As an illustration, we apply both methods to compute the von Neumann entropy of randomly generated quantum states. We evaluate the performance of two algorithms by applying noise obtained from actual superconducting systems. The numerical simulation indicates that, in the majority of scenarios, the second algorithm outperforms the first.

DOI: [10.1103/PhysRevA.109.062421](https://doi.org/10.1103/PhysRevA.109.062421)

## I. INTRODUCTION

Random quantum states form the foundational basis for our understanding of quantum information [1,2], black holes [3,4], and related fields. Numerous important functions of random states, such as Renyi entropy, von Neumann entropy, quantum Fisher information, the fidelity of random states, virtual distillation, and separation of density matrices [5–11], play crucial roles in quantum information, condensed matter physics, quantum chemistry, and beyond [12–18]. In the context of quantum computing applications, nonlinear subroutines are also expected to be of significant importance. The exploration of quantum neural networks and kernel methods highlights an ongoing interest in utilizing quantum computational models for advanced applications [19,20]. Kernel methods depend on nonlinear encodings, which require the capability to execute nonlinear quantum operations. Additionally, the capacity to solve nonlinear equations on quantum platforms through these subroutines [21] could enhance computational approaches in various fields, including finance [22] and fluid dynamics [23,24]. Moreover, leveraging nonlinear subroutines for improved signal detection in the presence of noise may also advance error mitigation techniques, showcasing the versatility of these operations in quantum computing [10,25]. Quantum computing is considered to possess the capability to address numerous problems more swiftly than classical computing systems [18,26–28]. The current development of quantum devices is situated in the noisy intermediate-scale quantum (NISQ) era, characterized by the handling of qubits in the tens or hundreds,

accompanied by inevitable quantum noise [29–31]. Exploiting the advantages and addressing the challenges of NISQ quantum computers, we tackle the fundamental yet challenging task of developing algorithms for computing nonlinear functions of random quantum state.

Prior methodologies for nonlinear transformations relied on simultaneously preparing multiple copies of a quantum state [5,32–35] and collective measurements [33–35]. These approaches necessitated a large number of qubits. For instance, when computing  $\text{Tr}\{\rho^m\}$ , with  $\rho$  representing the density matrix defined over  $n$  qubits, these methods required  $nm$  qubits. However, in the NISQ era, the number of qubits is still insufficient, rendering it inadequate to achieve quantum advantage within these algorithms [29]. Conversely, researchers have advocated for constructing the classical shadow of  $\rho$  and subsequently employing it to compute the purity  $\text{Tr}\{\rho^2\}$  [36–42]. While this approach still entails exponential resources relative to the number of qubits, it is perceived as an improvement over traditional state tomography [43]. Nevertheless, ongoing exploration of such methods is delimited to purity, which corresponds to quadratic functions of the density matrix. For higher-order functions like  $\text{Tr}\{\rho^m\}$ , there is no verified indication that these methods sustain an advantage over classical approaches.

To more efficiently exploit quantum computers in the NISQ era, we aim to design algorithms that employ the same number of qubits as  $\rho$ , and the circuit depth exhibits polynomial growth with the order of nonlinear function. A technique for generating random states involves initiating from an initial state and applying quantum gates randomly based on a specific probability distribution. The resultant final states post the application of diverse quantum gates to the initial state might not be orthogonal. We ascertain the presence of the algorithm we want, assuming the knowledge about how to construct the intended random state by utilizing random circuits.

\*Corresponding author: [wczwuli@163.com](mailto:wczwuli@163.com)

†Corresponding author: [wulicct@163.com](mailto:wulicct@163.com)

In this study, we introduce two distinct algorithms, both characterized by their shared utilization of the Grover gate  $G = I - 2|0\rangle\langle 0|$ . The primary aim of both algorithms is to compute the power series expansion  $\text{Tr}\{\rho^m\}$  for a nonlinear function in the context of a multiqubit quantum random state  $\rho$ . The first algorithm is based on the Hadamard Test (HT). It involves transforming an auxiliary qubit (usually  $|0\rangle\langle 0|$ ) into a superposition state using the Hadamard gate. After a controlled gate operation, another Hadamard gate extracts essential data, finalizing the calculation. Our algorithm is Hadamard Test-based but introduces an innovative approach: we deploy a quantum pure state circuit to simulate  $\text{Tr}\{G^{m-1}\rho\}$  computation for a quantum random state, by employing weighted averages across multiple measurements. The second algorithm begins by mathematically converting the calculation of  $\text{Tr}\{\rho^m\}$  for the desired quantum state into  $\text{Tr}\{G^m\}$ . A comprehensive understanding of  $G^m$  is acquired through Gate Set Tomography (GST) [44,45] in a subspace, facilitating the calculation and estimation of  $\text{Tr}\{G^m\}$  through mathematical processing. Compared to the Hadamard Test-based algorithm, this approach entails fewer qubits and two-qubit gates within the circuit, making it less affected by noise. Moreover, this method introduces a result-processing technique rooted in reconstruction. Both of these algorithms are scalable, with their time complexity growing polynomially with the number of qubits.

The structure of this paper is outlined as follows. We begin by introducing the Hadamard test-based algorithm in Sec. II. Subsequently, we explain the application of the GST method in Sec. III for extracting relevant information from the subspace. Section IV is devoted to error analysis and complexities. In Sec. V, we present computational results garnered from the preparation of the random quantum state  $\rho$  and the subsequent application of both algorithms. Notably, we compare the variations in calculations for the identical quantum state when utilizing the two distinct algorithms. Furthermore, we evaluate the performance of both algorithms by employing noise models derived from actual devices. A concise summary of this article is offered in Sec. VI.

## II. TRACE ESTIMATION OF HADAMARD TEST

### A. Theoretical part of Hadamard test

In this section, we show how to calculate  $\text{Tr}\{\rho^m\}$  using HT [46–49]. Firstly, we show how the quantum state is encoded in a quantum channel.

Suppose the  $n$ -qubits random quantum state is given by  $\rho = \sum_{i=1}^{\alpha} p_i U_i |0\rangle\langle 0| U_i^\dagger = \sum_{i=1}^{\alpha} p_i |\psi_i\rangle\langle \psi_i|$ , where  $\alpha$  random unitary gates  $U_i$  and probabilities  $p_i$  are known. Define the  $G$  gate as  $G = \sum_{j=1}^{\alpha} p_j U_j G_0 U_j^\dagger$ , where  $G_0 = I_{2^n} - 2|0\rangle\langle 0|$  is the Grover operator. Then we can express the quantum channel as  $G = I_{2^n} - 2\rho$ , and  $G^k = (I_{2^n} - 2\rho)^k$ . In this way, We encode the state  $\rho$  into a nonunitary quantum channel  $G$ .

Next, we show how HT works. In general, one ancilla qubit is required for HT. The quantum circuit is shown in Fig. 1. The computation process is as follows.

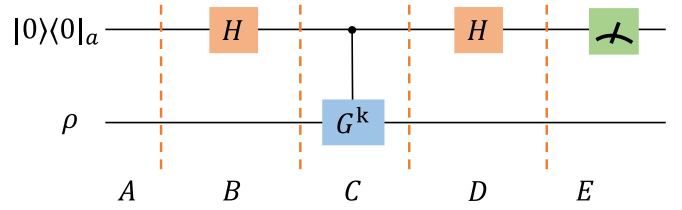


FIG. 1. Illustrative diagram of the Hadamard Test circuit, which is divided into five regions.  $H$  denotes the Hadamard gate. Within region C,  $k$  Controlled- $G$  gates are applied. Finally, in region E, the auxiliary qubit is measured on the computational basis.

In the region denoted as A, the quantum state of the circuit at this point is given by

$$\rho_A = |0\rangle\langle 0|_a \otimes \rho. \quad (1)$$

At the end of the B region, after applying the Hadamard gate, the state of the ancillary qubit is transformed as Eq. (2)

$$H|0\rangle\langle 0|_a H^\dagger = \frac{1}{2}(|0\rangle\langle 0|_a + |0\rangle\langle 1|_a + |1\rangle\langle 0|_a + |1\rangle\langle 1|_a). \quad (2)$$

The overall state of the circuit is given by

$$\rho_B = \frac{1}{2}(|0\rangle\langle 0|_a + |0\rangle\langle 1|_a + |1\rangle\langle 0|_a + |1\rangle\langle 1|_a) \otimes \rho. \quad (3)$$

Then, the state undergoes the action of the controlled  $G$  gate  $CG^k$ :

$$CG^k = |0\rangle\langle 0|_a \otimes I + |1\rangle\langle 1|_a \otimes G^k. \quad (4)$$

Thus, the state at the end of region C is

$$\rho_C = \frac{1}{2}(|0\rangle\langle 0|_a \otimes \rho) + \frac{1}{2}|0\rangle\langle 1|_a \otimes \rho(G^k)^\dagger + \frac{1}{2}|1\rangle\langle 0|_a \otimes G^k \rho + \frac{1}{2}|1\rangle\langle 1|_a \otimes G^k \rho(G^k)^\dagger. \quad (5)$$

Subsequently, the auxiliary qubit undergoes another Hadamard gate operation. Consequently, the state at the end of region D is then given by

$$\rho_D = \frac{1}{2}[(H|0\rangle\langle 0|_a H^\dagger) \otimes \rho + (H|0\rangle\langle 1|_a H^\dagger) \otimes \rho(G^k)^\dagger + (H|1\rangle\langle 0|_a H^\dagger) \otimes G^k \rho + (H|1\rangle\langle 1|_a H^\dagger) \otimes G^k \rho(G^k)^\dagger]. \quad (6)$$

Expanding each term results as

$$\begin{aligned} \rho_D = & \frac{1}{4}|0\rangle\langle 0|_a \otimes [\rho + \rho(G^k)^\dagger + (G^k)\rho + (G^k)\rho(G^k)^\dagger] \\ & + \frac{1}{4}|0\rangle\langle 1|_a \otimes [\rho - \rho(G^k)^\dagger + (G^k)\rho - (G^k)\rho(G^k)^\dagger] \\ & + \frac{1}{4}|1\rangle\langle 0|_a \otimes [\rho + \rho(G^k)^\dagger - (G^k)\rho - (G^k)\rho(G^k)^\dagger] \\ & + \frac{1}{4}|1\rangle\langle 1|_a \otimes [\rho - \rho(G^k)^\dagger - (G^k)\rho + (G^k)\rho(G^k)^\dagger]. \end{aligned} \quad (7)$$

The expression above can be denoted as follows:

$$\begin{aligned} \rho_D = & |0\rangle\langle 0|_a \otimes a_{11} + |0\rangle\langle 1|_a \\ & \otimes a_{12} + |1\rangle\langle 0|_a \otimes a_{21} + |1\rangle\langle 1|_a \otimes a_{22}. \end{aligned} \quad (8)$$

In region E, the measurement operators are defined as  $M_0 = |0\rangle\langle 0|_a \otimes I$  and  $M_1 = |1\rangle\langle 1|_a \otimes I$ . The probability distribution

ALGORITHM 1.

**Input:**  $n, p_i, U_i, N$

**Output:**  $\text{Tr}\{\rho^{m+1}\}$

- 1: Set the initial state to  $|0\rangle^{\otimes n+1}$ .
- 2: Randomly select  $U_i$  with probability  $p_i$  to act on the target qubits.
- 3: Apply a Hadamard gate to the ancilla qubit.
- 4: Sample  $m$  times, with a  $1/2$  probability of adding a  $CG$  gate to the circuit and a  $1/2$  probability of doing nothing.
- 5: Apply another Hadamard gate to the ancilla qubit.
- 6: Perform a computational basis measurement on the auxiliary qubit circuit.
- 7: Repeat the above steps  $N$  times and take the average.

over the outcomes of the measurement are

$$\begin{aligned} P(0) &= \text{Tr}\{M_0 \rho' M_0^\dagger\} = |a_{11}|^2 \\ &= \frac{1}{2} + \frac{1}{2} \text{Tr}\{G^k \rho\}, \end{aligned} \quad (9)$$

$$\begin{aligned} P(1) &= \text{Tr}\{M_1 \rho' M_1^\dagger\} = |a_{22}|^2 \\ &= \frac{1}{2} - \frac{1}{2} \text{Tr}\{G^k \rho\}. \end{aligned} \quad (10)$$

The expression for  $\text{Tr}\{G^k \rho\}$  can be derived, with the ultimate goal of estimating

$$\begin{aligned} \text{Tr}\{\rho^{m+1}\} &= \text{Tr}\left\{\left(\frac{1}{2}I - \frac{1}{2}G\right)^m \rho\right\} \\ &= \frac{1}{2^m} \sum_{k=0}^m C_m^k (-1)^k \text{Tr}\{G^k \rho\}. \end{aligned} \quad (11)$$

Let  $p_k = \frac{1}{2^m} C_m^k$  and  $x_k = (-1)^k \text{Tr}\{G^k \rho\}$ , the above equation can be transformed into an expectation calculation

$$\text{Tr}\{\rho^{m+1}\} = \sum_{k=0}^m p_k x_k = E_k(x_k). \quad (12)$$

To estimate the expectation, we need to generate quantum circuits using a random sampling method. For each circuit, we sample  $m$  times, with a  $1/2$  probability of adding a  $CG$  gate to the circuit and a  $1/2$  probability of doing nothing. After generating multiple circuits, we take the average of the results.

**B. Algorithm for calculating  $\text{Tr}\{\rho^{m+1}\}$ .**

The pseudocode in Algorithm 1 is for calculating  $\text{Tr}\{\rho^{m+1}\}$ .

**III. TRACE ESTIMATION OF QUANTUM TOMOGRAPHY**

**A. Theoretical part of quantum tomography**

Quantum tomography denotes a suite of techniques aiming to reconstruct an unknown quantum channel or state through experimental measurements. This process is pivotal for the comprehensive understanding and authentication of quantum apparatus [50–54]. Nevertheless, the scalability of quantum tomography poses a challenge, as the indispensable measurements and computational resources experience exponential growth in tandem with qubit numbers. Note that, although the resources required for tomography can be reduced by repre-

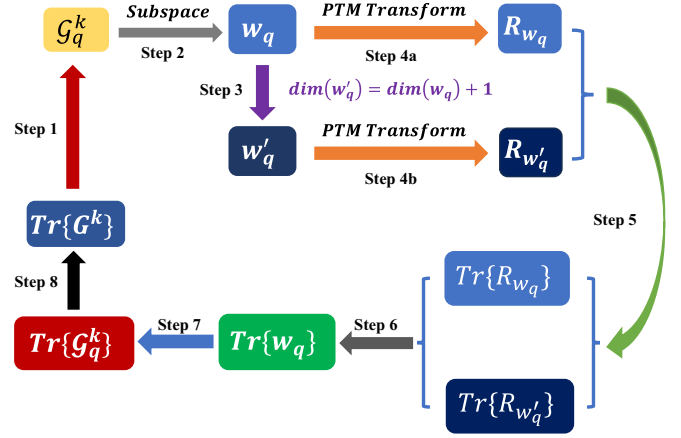


FIG. 2. The calculation of  $\text{Tr}\{G^k\}$  via Gate Set Tomography involves several steps. Initially, step 1 entails breaking down  $\text{Tr}\{G^k\}$  into individual calculations of  $\text{Tr}\{G_q^k\}$ . Following this, step 2 focuses on  $w_q$ , which is the representation of  $\text{Tr}\{G_q^k\}$  in the subspace. The subsequent step 3 involves expanding  $w_q$  by an additional dimension. In step 4a, the matrix  $w_q$  is converted into its corresponding Pauli transfer matrix,  $R_{w_q}$ . Similarly, step 4b involves the transformation of  $w'_q$  into its respective Pauli transfer matrix,  $R_{w'_q}$ . Step 5 then entails calculating the traces of these matrices individually. Utilizing specific mathematical approaches in step 6 allows for the computation of  $\text{Tr}w_q$ , leading to the determination of  $\text{Tr}\{G_q^k\}$  in Step 7. Ultimately, step 8 completes the process, yielding  $\text{Tr}\{G^k\}$ , effectively reversing the initial step.

senting quantum states as matrix product states, in the worst case the resources required for tomography still increase exponentially with the number of bits [52,55]. In the context of our current research problem, there is a silver lining: the subspace we are investigating maintains a dimension that remains unaffected by the number of qubits. This distinctive feature becomes particularly advantageous. In the subsequent section, we expound on our utilization of the GST method to extract the pertinent information from this designated subspace.

In comparison with the preceding context, the process of preparing random quantum states adheres to the same approach as the Hadamard Test method. This methodology necessitates the application of an assortment of stochastically selected gates  $\{(U_i)_{2^n \times 2^n}, i = 1, 2, \dots, \alpha\}$  onto the initial quantum state (typically  $|0\rangle|0\rangle^{\otimes n}$ ), resulting in the emergence of a random quantum state. The primary objective revolves around the computation of  $\text{Tr}\{\rho^m\}$ , which is achieved through the intermediary of  $\text{Tr}\{G^m\}$ , where  $G = I - 2\rho$ . We can express  $\text{Tr}\{\rho^m\}$  as

$$\text{Tr}\{\rho^m\} = \frac{1}{2^m} \sum_{k=0}^m C_m^k (-1)^k \text{Tr}\{G^k\}. \quad (13)$$

Similar to the approach employed in the HT-based algorithm, we can estimate this summation by leveraging the Monte Carlo method. This involves conducting multiple circuit samplings in accordance with their respective probabilities and subsequently calculating the average. Through this process, we can attain the sought-after value of  $\text{Tr}\{\rho^m\}$ . The figure shown in Fig. 2 presents our computation process through a simplified flowchart, starting from the decomposition of  $\text{Tr}\{G^k\}$ , transitioning to the calculation of

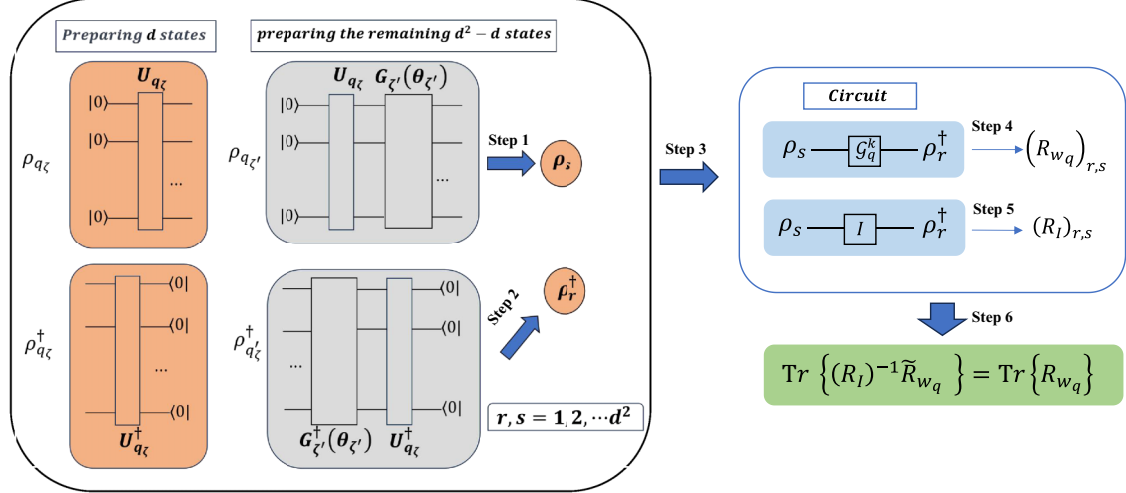


FIG. 3. Schematic diagram of a subprocess for computing  $\text{Tr}\{R_{w_q}\}$  within the tomography workflow. The diagram runs from left to right, with the boxes on the left representing the preparation of the initial state (step 1) and the measurement (step 2). The blue box on the upper right depicts the placement of the prepared states into the circuit (step 3), where the  $\mathcal{G}_q^k$  gate and the identity gate are inserted. Steps 4 and 5 allow us to obtain the matrix elements of the PTMs corresponding to each of these gates. Finally, in step 6,  $\text{Tr}\{R_{w_q}\}$  is indirectly obtained through a similarity transformation.

each individual  $\mathcal{G}_q^k$ , and then proceeding with a series of computations within their respective subspaces. Finally, the computed results  $\text{Tr}\{\mathcal{G}_q^k\}$  are weighted and summed to obtain the desired calculation  $\text{Tr}\{G^k\}$ . We illustrate the sub-process of computing  $\text{Tr}\{R_{w_q}\}$  in Fig. 3, which, together with Fig. 2, forms the complete tomography computation process.

### 1. Mathematical treatment

We start with the unitary quantum gate  $G_0$ :

$$(G_0)_{2^n \times 2^n} = I_{2^n} - 2|0\rangle\langle 0|^{\otimes n} = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \quad (14)$$

Through the application of random unitary gates  $U_i$  to the initial gate  $G_0$ , a set of  $\alpha$  distinct gates is generated. These gates are denoted as  $G_i = U_i G_0 U_i^\dagger$ . The composite gate  $G$  is then defined as the weighted summation of these transformed gates:  $G = \sum_{i=1}^{\alpha} p_i G_i$ .

An insightful observation can be made that in the presence of  $k$  occurrences of  $G$  gates within the circuit, the total count of possible arrangements aggregates to  $\alpha^k$ . These arrangements are uniquely labeled by the index  $q = \{1, 2, \dots, \alpha^k\}$ . As a result, this algorithm effectively dissects the trace  $\text{Tr}\{G^k\}$  of the higher-order powers of  $G$  into computations encompassing  $\alpha^k$  arrangements denoted as  $\mathcal{G}_q^k$ . The calculation process is thereby executed on higher-order random quantum states via the application of a Monte Carlo methodology.

The corresponding probability combination  $\prod_{t=1}^k p_{q_t}$  is represented as  $\mathcal{P}_q$ . Therefore,  $\text{Tr}\{G^k\}$  can be expressed as

$$\text{Tr}\{G^k\} = \sum_{q=1}^{\alpha^k} \mathcal{P}_q \text{Tr}\{\mathcal{G}_q^k\}, \quad (15)$$

where we use  $\mathcal{G}_q^k$  to denote  $\prod_{t=1}^k G_{q_t}$  for convenience.

### 2. Matrix representation of $\mathcal{G}_q^k$

For each of the  $\alpha^k$  instances of  $\mathcal{G}_q^k$ , a specific  $\mathcal{G}_q^k$  is chosen for computation where  $q = 1, 2, \dots, \alpha^k$ . In this scenario, the calculation method is provided for arbitrary combinations, while the computation process remains similar for other combinations.

Upon choosing a specific combination  $\mathcal{G}_q^k$ , a set of  $k$  corresponding  $G_{q_t}$  gates is determined, thereby giving rise to  $k$  specific  $|\psi_{q_t}\rangle$  states, where  $t = 1, 2, \dots, k$ .

For example, consider the cases

$$\begin{aligned} q = 1, & \{ \mathcal{G}_1^k : \underbrace{G_1 G_1 \cdots G_1 G_1}_{k \text{ layers}} \}, \\ q = 2, & \{ \mathcal{G}_2^k : \underbrace{G_1 G_1 \cdots G_1 G_2}_{k \text{ layers}} \}, \\ & \vdots \\ q = \alpha^k, & \{ \mathcal{G}_{\alpha^k}^k : \underbrace{G_\alpha G_\alpha \cdots G_\alpha G_\alpha}_{k \text{ layers}} \}. \end{aligned}$$

Therefore, under the matrix background denoted as  $\mathcal{G}_q^k$ , the subspace dimension  $d$  is not constant. Hence, the determination of the subspace dimension relies entirely on the count of unique gate types present in the given order  $q$ . Let us define  $d$  as the dimension of the nontrivial subspace corresponding to the simplified merge of  $\mathcal{G}_q^k := \prod_{t=1}^k G_{q_t}$ , while representing the random gate sets used to prepare  $G_{q_t}$  as  $U_{q_t}$ , i.e.,  $G_{q_t} = U_{q_t} G_0 U_{q_t}^\dagger$ . Quantum states prepared by different random gates are represented as  $|\psi_{q_\zeta}\rangle$ ,  $\zeta = 1, 2, \dots, d$ .

To ensure completeness, a set of state vectors  $\{|\phi_\eta\rangle, \eta = 1, 2, \dots, 2^n - d\}$  is introduced, which are orthogonal to all the state vectors  $|\psi_{q_\zeta}\rangle$ . Due to the condition  $\langle \psi_{q_\zeta} | \phi_\eta \rangle = 0$ , it

can be inferred that

$$\mathcal{G}_q^k(|\phi_\eta\rangle) = \left[ \prod_{\zeta=1}^d (I - 2(|\psi_{q_\zeta}\rangle\langle\psi_{q_\zeta}|)) \right] |\phi_\eta\rangle = |\phi_\eta\rangle.$$

It is evident that  $\{|\phi_\eta\rangle, \eta = 1, 2, \dots, 2^n - d\}$  forms a set of eigenstates of  $\mathcal{G}_q^k$  with eigenvalue 1. In the representation with  $2^n$  state vectors

$$\left\{ \underbrace{|\psi_{q_1}\rangle, |\psi_{q_2}\rangle, \dots, |\psi_{q_d}\rangle}_{d \text{ terms}}, \underbrace{|\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_\eta\rangle}_{(2^n-d) \text{ terms}} \right\}, \quad (16)$$

as the basis in the  $V_q$  space, the matrix  $\mathcal{G}_q^k$  can be expressed as

$$\begin{pmatrix} w_{11} & w_{12} & \dots & w_{1d} & 0 & 0 & \dots \\ w_{21} & w_{22} & \dots & w_{2d} & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots & 0 & 0 & \dots \\ w_{d1} & w_{d2} & \dots & w_{dd} & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & 1 \end{pmatrix}. \quad (17)$$

The top-left  $d \times d$  matrix  $w_q$  can be considered as a representation of  $\mathcal{G}_q^k$  in the  $d$ -dimensional invariant subspace  $V_{q_1}$  spanned by  $d$  states  $|\psi_{q_\zeta}\rangle$ . Through this expression, the calculation of  $\text{Tr}\{\mathcal{G}_q^k\}$  can be performed

$$\begin{aligned} \text{Tr}\{\mathcal{G}_q^k\} &= \text{Tr}\{w_q\} + \text{Tr}\{I_{2^n-d}\} \\ &= \text{Tr}\{w_q\} + 2^n - d. \end{aligned} \quad (18)$$

### 3. PTM representation of $w_q$

After characterizing  $\mathcal{G}_q^k$ , we need to obtain the trace  $\text{Tr}\{w_q\}$ . However, the matrix  $w_q$  is unknown.  $w_q$  as a mapping, finding the trace  $\text{Tr}\{w_q\}$  requires the use of the Pauli transfer matrix (PTM). We denote the PTM corresponding to  $w_q$  as  $R_{w_q}$ .

Based on the previous discussion, an important relationship can be used

$$\begin{aligned} R_{w_q}|\rho_s\rangle\rangle &= |w_q(\rho_s)\rangle\rangle = |w_q\rho_s w_q^\dagger\rangle\rangle \\ r, s &\in \{1, 2, \dots, d^2\}. \end{aligned} \quad (19)$$

For simplicity, the subscript  $q$  indicating that  $\rho$  belongs to the ordering  $\mathcal{G}_q^k$  will be omitted below.

Although  $R_{w_q}$  is an  $d^2 \times d^2$  matrix, the vector  $|\rho_s\rangle\rangle$  has dimensions of  $(2^n)^2 \times 1$ . Since  $|\rho_s\rangle\rangle$  only has nonzero elements in the  $d^2$ -dimensional subspace, the remaining part of the vector is trivial.

By left-multiplying Eq. (19) by  $\langle\langle\rho_r|$ , the matrix elements of the PTM matrix  $(R_{w_q})_{d^2 \times d^2}$  are given as follows:

$$(R_{w_q})_{rs} = \langle\langle\rho_r|R_{w_q}|\rho_s\rangle\rangle = \langle\langle\rho_r|w_q\rho_s w_q^\dagger\rangle\rangle. \quad (20)$$

To calculate the matrix elements  $(R_{w_q})_{rs}$  of  $R_{w_q}$  using this method,  $d^2$  quantum states in the Hilbert-Schmidt space are required, represented by vectors  $|\rho_r\rangle\rangle$ . The proof of the completeness of these states can be found in Appendix A.

We need to calculate the trace of PTM. Although we can obtain each matrix element of the PTM sequentially through

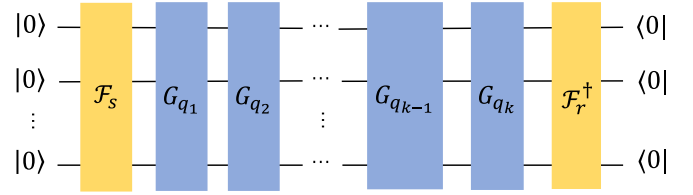


FIG. 4. This is a quantum circuit with  $n$ -qubits used to compute  $\text{Tr}\{\rho_r^\dagger \mathcal{G}_q^k \rho_s \mathcal{G}_q^k\}$ . First, we add the prepared state  $\rho_r^\dagger$  to the circuit. Next, we apply the operator  $\mathcal{G}_q^k$ . Then, we add  $\rho_s$  to the circuit.

the circuit, we cannot directly compute its trace because the basis vectors in the subspace we generate are not orthogonal. There are two feasible approaches: Schmidt decomposition to orthogonalize its basis vectors and then directly calculate the trace by summing the main diagonal elements; or calculate its trace indirectly through a matrix similarity transformation. This paper adopts the GST method, which is the latter approach of indirectly calculating the trace of PTM through similarity transformations.

### 4. Gate Set Tomography

In quantum process tomography (QPT), the information required to reconstruct each gate  $R_{w_q}$  is contained in the measurements of  $\langle\langle\rho_r|R_{w_q}|\rho_s\rangle\rangle$ , and  $R_{w_q}$  is the PTM of  $w_q$  in the Hilbert-Schmidt space. QPT assumes that the initial state and final measurements are known. In practice, these states and measurements must be prepared using quantum gates, and these gates  $\{F_r, F_s'\}$  themselves may have imperfections [50]

$$\begin{aligned} \langle\langle\rho_r| &= \langle\langle 0|F_r, \\ |\rho_s\rangle\rangle &= F_s'|0\rangle\rangle. \end{aligned}$$

Indeed, the initial states and final measurements that were prepared using gates are not directly known and can introduce errors in the estimation process. GST aims to characterize the fully unknown set of gates and states [56]

$$\mathcal{R} = \{|\rho\rangle\rangle, \langle\langle E|R_{w_1}, \dots, R_{w_q}, \dots\rangle\rangle, q = 1, 2, \dots, \alpha^k. \quad (21)$$

GST has similar requirements to QPT, i.e., the ability to measure the set of gates  $\mathcal{R} = \{|\rho\rangle\rangle, \langle\langle E|R_{w_1}, \dots, R_{w_{\alpha^k}}\}$  in the form of expectation values

$$p = \text{Tr}\{\rho_r^\dagger w_q(\rho_s)\} = \langle\langle\rho_r|R_{w_q}|\rho_s\rangle\rangle = \text{Tr}\{\rho_r^\dagger w_q \rho_s w_q^\dagger\}.$$

To simplify the expression, we use  $F_r, F_s'$ , where  $(r, s = 1, 2, \dots, d^2)$ , to denote the quantum gates used for preparing quantum states and measurements (as shown in Fig. 4), which are  $G_{q_r}(\theta_r)U_{q_r}$  and  $U_{q_s}^\dagger G_{q_s}^\dagger(\theta_s)$ , respectively. The density matrices of these prepared quantum states are linearly independent. Please refer to Appendix A for details. By constructing a quantum circuit, it is possible to compute the  $d^2$  matrix elements of the PTM matrix  $R_{w_q}$ . We define  $(R_{w_q})_{rs}$  as follows:

$$(R_{w_q})_{rs} = \langle\langle\rho_r|R_{w_q}|\rho_s\rangle\rangle = \langle\langle 0|F_r R_{w_q} F_s'|0\rangle\rangle. \quad (22)$$

Inserting the completeness state into it yields

$$\begin{aligned}
 p_{rs} &= (R_{w_q})_{rs} = \langle\langle \rho_r | R_{w_q} | \rho_s \rangle\rangle \\
 &= \langle\langle 0 | F_r R_{w_q} F_s | 0 \rangle\rangle \\
 &= \sum_{a,b} \langle\langle 0 | F_r | a \rangle\rangle \langle\langle a | R_{w_q} | b \rangle\rangle \langle\langle b | F_s | 0 \rangle\rangle \\
 &= \sum_{a,b} A_{ra} (R_{w_q})_{bs},
 \end{aligned}$$

where we define

$$\begin{aligned}
 p_{rs} &= (AR_{w_q}B)_{rs}, \\
 A &= \sum_r |r\rangle\langle\langle 0 | F_r, \\
 B &= \sum_s F_s |0\rangle\langle\langle s|. \quad (23)
 \end{aligned}$$

Let  $\tilde{R}_{w_q} = AR_{w_q}B$ , the identity matrix  $I$  also serves as a mapping, and its PTM has the following properties:

$$R_I |\rho_s\rangle\rangle = |I(\rho_s)\rangle\rangle = |I\rho_s I^\dagger\rangle\rangle = |\rho_s\rangle\rangle.$$

It can be observed that the action of  $R_I$  is similar to the identity matrix and its matrix elements can be expressed as follows:

$$(R_I)_{rs} = \langle\langle \rho_r | R_I | \rho_s \rangle\rangle = \text{Tr}\{\rho_r^\dagger I \rho_s I^\dagger\} = \text{Tr}\{\rho_r^\dagger \rho_s\}.$$

Denoting  $g_{rs} = (R_I)_{rs} = \langle\langle \rho_r | \rho_s \rangle\rangle = \langle\langle 0 | F_r F_s | 0 \rangle\rangle$ , we can insert the completeness state and obtain

$$g_{rs} = \sum_{a,b} \langle\langle 0 | F_r | a \rangle\rangle \langle\langle a | b \rangle\rangle \langle\langle b | F_s | 0 \rangle\rangle = (AB)_{rs}. \quad (24)$$

For a given combination  $\mathcal{G}_q^k$ , let  $R_I = AB$ , where  $A$  and  $B$  are matrices. The experimental measurement value  $p_{rs}$  corresponds to the  $rs$  component of the matrix  $A(R_{w_q})B$ , while  $g_{rs}$  corresponds to the  $rs$  component of the matrix  $AB$ . The quantum channel we reconstruct will differ from the real quantum channel by a similarity transformation

$$(R_I)^{-1} \tilde{R}_{w_q} = B^{-1} A^{-1} A R_{w_q} B = B^{-1} R_{w_q} B.$$

Therefore, we can estimate the trace of  $R_{w_q}$ :

$$\text{Tr}\{(R_I)^{-1} \tilde{R}_{w_q}\} = \text{Tr}\{B^{-1} R_{w_q} B\} = \text{Tr}\{R_{w_q}\}. \quad (25)$$

Based on the calculations mentioned earlier, the value of  $\text{Tr}\{R_{w_q}\} = |\text{Tr}\{w_q\}|^2$  can be determined. However, this is not the final result for  $\text{Tr}\{w_q\}$ .

Note that to ensure that  $R_I$  is not an ill-conditioned matrix, it is necessary to perform subspace selection. For a detailed analysis, please refer to Sec. IV A.

### 5. Mathematical processing of results

To calculate  $\text{Tr}\{w_q\}$ , an operation involving taking the square root is required because of  $|\text{Tr}\{w_q\}|^2 = \text{Tr}\{R_{w_q}\}$ . In general,  $\text{Tr}\{w_q\}$  can be decomposed into real and imaginary parts

$$\text{Tr}\{w_q\} = \text{Re}[\text{Tr}\{w_q\}] + i \cdot \text{Im}[\text{Tr}\{w_q\}].$$

Then the trace of  $R_{w_q}$  can be expressed as

$$\text{Tr}\{R_{w_q}\} = (\text{Re}[\text{Tr}\{w_q\}])^2 + (\text{Im}[\text{Tr}\{w_q\}])^2.$$

In fact, only the real part needs to be estimated because the sum of all the imaginary parts of  $\text{Tr}\{w_q\}$  vanishes after summation

$$\sum_{q=1}^{\alpha^k} \text{Tr}\{w_q\} = \sum_{q=1}^{\alpha^k} \text{Re}[\text{Tr}\{w_q\}]. \quad (26)$$

Next, we show how to estimate the real part of  $\text{Tr}\{w_q\}$ . Consider  $(w'_q)_{d+1} = (w_q)_d \oplus 1$ , where  $w'_q$  is the representation of  $\mathcal{G}_q^k$  on the subspace  $V'_q : \text{span}\{|\psi_{q_1}\rangle, |\psi_{q_2}\rangle, \dots, |\psi_{q_d}\rangle; |\phi\rangle\}$ .

$$w'_q = \begin{pmatrix} w_q & 0 \\ 0 & 1 \end{pmatrix},$$

where  $|\phi\rangle \notin V_q$  and  $|\phi\rangle$  can be prepared through the quantum circuit. It can be seen that  $V'_q$  is also a noninvariant subspace of  $\mathcal{G}_q^k$ . It's obvious that

$$\text{Tr}\{w'_q\} = \text{Tr}\{w_q\} + 1, \quad (27)$$

and

$$\begin{aligned}
 \text{Tr}\{R_{w'_q}\} &= |\text{Tr}\{w'_q\}|^2 = |\text{Tr}\{w_q\} + 1|^2 \\
 &= |\text{Re}[\text{Tr}\{w_q\}] + i \cdot \text{Im}[\text{Tr}\{w_q\}] + 1|^2 \\
 &= (\text{Re}[\text{Tr}\{w_q\}])^2 + 2(\text{Re}[\text{Tr}\{w_q\}]) \\
 &\quad + (\text{Im}[\text{Tr}\{w_q\}])^2 + 1. \quad (28)
 \end{aligned}$$

Recall the relation that

$$\begin{aligned}
 \text{Tr}\{R_w\} &= |\text{Tr}\{w_q\}|^2 \\
 &= (\text{Re}[\text{Tr}\{w_q\}])^2 + (\text{Im}[\text{Tr}\{w_q\}])^2,
 \end{aligned}$$

we can find that

$$\text{Re}[\text{Tr}\{w_q\}] = \frac{1}{2} [\text{Tr}\{R_{w'_q}\} - \text{Tr}\{R_{w_q}\} - 1]. \quad (29)$$

The procedure to calculate  $\text{Tr}\{R_{w'_q}\}$  follows a similar algorithm as for computing  $\text{Tr}\{R_{w_q}\}$ , with the distinction that in this case,  $(d+1)^2$  quantum states are required.

### B. Algorithm process of tomography

Algorithms 2 and 3 are the procedure for calculating  $\text{Tr}\{\rho^m\}$  for the quantum state  $\rho = \sum_{i=1}^{\alpha} p_i U_i |0\rangle\langle 0| U_i^\dagger$ . Algorithm 2 serves as the main program for GST, while algorithm 3 functions as a subroutine called multiple times within GST, responsible for iteratively computing  $\text{Tr}\{\mathcal{G}_q^k\}$ .

## ALGORITHM 2.

---

**Input:**  $N, m, p_i, U_i, \alpha, n$

**Output:**  $\text{Tr}\{\rho^m\}$

- 1: **for** iteration = 1 to  $N$  **do**
- 2: Set the initial state as  $|0\rangle^{\otimes n}$ .
- 3: Initialize  $d \leftarrow 0$
- 4: Initialize the subspace  $V_q \leftarrow []$ .
- 5: // Dimension of the subspace is determined.
- 6: **for**  $i = 1$  to  $m$  **do**
- 7: Randomly choose gate with probability 1/2 (for  $G$ ) or 1/2 (for  $I$ ).
- 8: If it is a  $G$  gate, choose  $G_i$  with probability  $p_i$ .
- 9: **end for**
- 10: // Select  $G_q^k$  with probability  $\mathcal{P}_q$
- 11: **for**  $i = 1$  to  $k$  **do**
- 12: // Here  $k$  is the number of  $G$  in  $G_q^k$ .
- 13: Apply gate  $U[i]$  to prepare quantum state  $|\psi_i\rangle = U[i]|0\rangle^{\otimes n}$ .  $U[i]$  is the unitary corresponding to  $i$ th  $G$  gate,  $G[i] = U[i]G_0U[i]^\dagger$ .
- 14: Compute the Gram matrix  $g_i$  in the subspace  $V \oplus |\psi_i\rangle$ .
- 15: Compare eigenvalues of Gram matrix  $\epsilon_g$  of  $g_i$  with threshold  $\epsilon$ .
- 16: // Perform linear correlation analysis on  $|\psi_i\rangle$  states.
- 17: **if**  $\epsilon_g \geq \epsilon$  **then**
- 18: Update the dimension  $d \leftarrow d + 1$ .
- 19: Update the subspace  $V_q \leftarrow V_q \oplus |\psi_i\rangle$ .
- 20: **end if**
- 21: **end for**
- 22:  $t_1 \leftarrow \text{Subroutine}(G_q^k, d, V_q)$
- 23:  $t_2 \leftarrow \text{Subroutine}(G_q^k, d + 1, V_q \oplus |\phi\rangle)$
- 24:  $\text{Tr}\{G_q^k\} \leftarrow 2^n - d + (t_1 - t_2 - 1)/2$ .
- 25: // Invoke the subroutine to calculate  $\text{Tr}\{G_q^k\}$ .
- 26: **end for**
- 27: Calculate  $\text{Tr}\{\rho^m\} = \text{Tr}\{(\frac{1}{2}I - \frac{1}{2}G)^m\} = \sum_{k=0}^m p_k x_k$
- 28: //  $p_k = (\frac{1}{2})^m C_m^k$  and  $x_k = (-1)^k \text{Tr}\{G^k\}$

---

## IV. ERROR ANALYSIS

## A. Eigenvalue truncation of the simplest case

To prevent the introduction of significant statistical errors during subsequent numerical computations, it becomes imperative to truncate the eigenvalues of the Gram matrix  $R_I$ . To ensure that the eigenvalues of the Gram matrix are all greater than threshold  $\epsilon$ , we start from  $|\psi_2\rangle$  to select the quantum states that constitute the subspace. Here we default to  $|\psi_1\rangle$  in the subspace,

Commencing from the iteration  $k = 2$ , the computation of the Gram matrix  $R_I^k$  is carried out. This matrix is associated with four distinct eigenvectors, which can be identified as follows:

$$|\Psi_1\rangle\langle\Psi_1|, |\Psi_1\rangle\langle\Psi_2|, |\Psi_2\rangle\langle\Psi_1|, |\Psi_2\rangle\langle\Psi_2|, \quad (30)$$

where  $|\Psi_1\rangle = |\psi_1\rangle$  and  $|\Psi_2\rangle$  is defined from the Schmidt orthogonalization

$$|\Psi_2\rangle = \frac{1}{\Delta_2}|\psi_2\rangle - \frac{x_{1,1}}{\Delta_2}|\psi_1\rangle. \quad (31)$$

## ALGORITHM 3.

---

**Input:**  $\mathcal{G}_q^k, d, V$

**Output:**  $\text{Tr}\{R_{w_q}\}$

- 1: **for**  $\zeta = 1$  to  $d$  **do**
- 2:  $\rho_{q_\zeta} = |\psi_{q_\zeta}\rangle\langle\psi_{q_\zeta}| = U_{q_\zeta}|0\rangle\langle 0|U_{q_\zeta}^\dagger$
- 3: **end for**
- 4: **for**  $\zeta' = 1$  to  $d$  **do**
- 5: **if**  $\zeta \neq \zeta'$  **then**
- 6:  $\rho_{q_{\zeta'}} = |\psi_{q_{\zeta'}}\rangle\langle\psi_{q_{\zeta'}}|$
- 7:  $= G_{\zeta'}(\theta_{\zeta'})U_{q_{\zeta'}}|0\rangle\langle 0|U_{q_{\zeta'}}^\dagger G_{\zeta'}^\dagger(\theta_{\zeta'})$
- 8: // Prepare  $(d^2 - d)$  additional quantum states
- 9: **else**
- 10: Break // Exit the loop
- 11: **end if**
- 12: **end for**
- 13: **for**  $r = 1$  to  $d^2$  **do**
- 14: **for**  $s = 1$  to  $d^2$  **do**
- 15:  $|\psi_s\rangle \xrightarrow{\mathcal{G}_q^k} \langle\psi_r|$
- 16: //  $(\tilde{R}_{w_q})_{r,s} = (AR_{w_q}B)_{r,s}$
- 17:  $|\psi_s\rangle \xrightarrow{I} \langle\psi_r|$
- 18: //  $(\tilde{R}_I)_{r,s} = (AR_I B)_{r,s} = (AB)_{r,s}$
- 19: **end for**
- 20: **end for**
- 21:  $\text{Tr}\{R_{w_q}\} \leftarrow \text{Tr}\{(\tilde{R}_I^{-1} \cdot \tilde{R}_{w_q})\}$

---

Here  $x_{1,1}$  is the overlap  $\langle\psi_1|\psi_2\rangle$  and  $|\Delta_2|^2 = 1 - |x_{1,1}|^2$  is a normalization factor. The eigenvalues are respectively given by

$$1, \frac{|\Delta_2|^2}{1 + |x_{1,1}|^2}, \frac{|\Delta_2|^2}{1 - |x_{1,1}|^2}, \left| \frac{|\Delta_2|^2}{1 + |x_{1,1}|^2} \right|^2. \quad (32)$$

If the smallest eigenvalue  $|\frac{|\Delta_2|^2}{1 + |x_{1,1}|^2}|^2$  is smaller than the threshold  $\epsilon$  we preset, we add  $|\psi_2\rangle$  into the subspace. If not, we discard  $|\psi_2\rangle$  and then consider whether  $|\psi_3\rangle$  can be added into the subspace.

## B. Eigenvalue truncation for general cases

We generalize the discussion in the previous subsection to the case of a higher-dimensional subspace. Assume that the current subspace already contains  $|\psi_1\rangle, \dots, |\psi_{k-1}\rangle$ , and now we need to decide whether we can add  $|\psi_k\rangle$  to the subspace.  $|\psi_k\rangle$  can be written as

$$|\psi_k\rangle = \sum_i^{k-1} x_{k,i}|\psi_i\rangle + \Delta_k|\Psi_k\rangle, \quad (33)$$

where

$$|\Psi_k\rangle = \frac{1}{\Delta_k}|\psi_k\rangle - \frac{1}{\Delta_k} \sum_i^{k-1} x_{k,i}|\psi_i\rangle. \quad (34)$$

In this case, the smallest singular value is  $|\frac{|\Delta_k|^2}{1 + |x_k|^2}|^2$ , where  $|x_k|^2 = \sum_i^{k-1} |x_{k,i}|^2$ .

We must assess the relationship between the smallest singular value and  $\epsilon$ . If the smallest eigenvalue surpasses  $\epsilon$ , we should proceed with enlarging the subspace. However, if the smallest singular value is less than  $\epsilon$ , it is advisable to disregard the state  $|\psi_k\rangle$ .

### C. Eigenvalue estimation error

In the actual process, the measurement of the Gram matrix has statistical errors  $\Delta(R_I)$ . This will lead to errors in the calculation of the eigenvalues of the Gram matrix. Given that each element of  $\Delta(R_I)$  is subjected to measurement  $N$  times, it follows that every element holds an order of magnitude around  $O(\frac{1}{\sqrt{N}})$ . By referencing the disk theorem, no eigenvalue surpasses  $O(\frac{d^2}{\sqrt{N}})$ . Applying the Weyl inequality [57] allows us to deduce that the disparity between the computed minimum eigenvalue of  $R_I$  and the actual minimum eigenvalue does not exceed  $O(\frac{d^2}{\sqrt{N}})$ . Employing Hoeffding's inequality, the likelihood of each element within  $\Delta(R_I)$  is less than  $\tilde{\epsilon}_{R_I}$  amounts to  $1 - \delta_{R_I} < 1 - 2e^{-2N\tilde{\epsilon}_{R_I}^2}$ . Consequently, the probability of every element being less than  $\tilde{\epsilon}_{R_I}$ , i.e., the error of the eigenvalue will not exceed  $d^2\tilde{\epsilon}_{R_I}$ , can be expressed as  $(1 - \delta_{R_I})^{d^2} > 1 - d^2\delta_{R_I} = 1 - \tilde{\delta}_{R_I}$ , where

$$\tilde{\delta}_{R_I} = d^2\delta_{R_I} \geq 2d^2e^{-2N\tilde{\epsilon}_{R_I}^2}. \quad (35)$$

This means that the estimation for eigenvalues, to achieve precision  $d^2\tilde{\epsilon}_{R_I}$  with a probability of  $1 - d^2\tilde{\delta}_{R_I}$ , requires no more than

$$N = O\left(\frac{\log \frac{d^2}{\tilde{\delta}_{R_I}}}{\tilde{\epsilon}_{R_I}^2}\right) \quad (36)$$

number of measurements for each matrix element.

### D. Error analysis of $R_I^{-1}$

In the GST process, we need to invert the Gram matrix. Next, we consider the effect of the error  $\Delta(R_I)$  on the inverse of  $R_I$ . Using Taylor expansion,

$$\begin{aligned}^{-1} &= R_I^{-1} - R_I^{-1}\Delta(R_I)R_I^{-1} \\ &+ g^{-1}\Delta(R_I)R_I^{-1}\Delta(R_I)R_I^{-1} + \dots, \quad (37)\end{aligned}$$

each matrix term can be bounded as

$$\begin{aligned}(R_I^{-1})_{i,j} &< \epsilon, \\ [R_I^{-1}\Delta(R_I)R_I^{-1}]_{i,j} &< \frac{d^2\tilde{\epsilon}_{R_I}}{\epsilon^2}, \\ [(R_I^{-1}\Delta(R_I))^k R_I^{-1}]_{i,j} &< \frac{d^{2k}\tilde{\epsilon}_{R_I}^k}{\epsilon^{k+1}},\end{aligned}$$

where  $i, j = 0, 1, \dots, d-1$ . Let  $\tilde{\epsilon}_{R_I} = \epsilon_1\epsilon^2$ , then

$$[R_I^{-1} - (R_I + \Delta(R_I))^{-1}]_{i,j} < d^2\epsilon_1 \sum_{l=0}^{+\infty} (d^2\epsilon_1\epsilon)^l,$$

which can be bounded as

$$\{R_I^{-1} - [R_I + \Delta(R_I)]^{-1}\}_{i,j} < \frac{d^2\epsilon_1}{1 - d^2\epsilon_1\epsilon}. \quad (38)$$

### E. Sampling error

In this section, we consider the statistical error when estimating the trace of a quantum gate. Let  $(R_{w_y})_{i,j} = \langle \langle \rho_i | G | \rho_j \rangle \rangle$ . The error in computing  $\text{Tr}\{R_I^{-1}R_{w_y}\}$  is given by

$$\begin{aligned}\mathcal{E} &= \text{Tr}\{\Delta(R_I^{-1})R_{w_y}\} + \text{Tr}\{R_I^{-1}\Delta(R_{w_y})\} \\ &+ \text{Tr}\{\Delta(R_I^{-1})\Delta(R_{w_y})\}.\end{aligned} \quad (39)$$

The first term can be bounded as

$$\begin{aligned}\text{Tr}\{\Delta(R_I^{-1})R_{w_y}\} &< d^2\text{Tr}\{R_I^{-1} - [R_I + \Delta(R_I)]^{-1}\} \\ &< \frac{d^4\epsilon_1}{1 - d^2\epsilon_1\epsilon}.\end{aligned}$$

The second term is

$$\text{Tr}\{R_I^{-1}\Delta(R_{w_y})\} < \frac{d^2\epsilon_2}{\epsilon},$$

where we use

$$[\Delta(R_{w_y})]_{i,j} < \epsilon_2,$$

and  $i, j = 0, 1, \dots, d-1$ . The last term is

$$\text{Tr}\{\Delta(R_I^{-1})\Delta(R_{w_y})\} < \frac{d^4\epsilon_1\epsilon_2}{1 - d^2\epsilon_1\epsilon}.$$

So all the terms must add up to less than

$$\mathcal{E} < \frac{d^4\epsilon_1}{1 - d^2\epsilon_1\epsilon} + \frac{d^2\epsilon_2}{\epsilon} + \frac{d^4\epsilon_1\epsilon_2}{1 - d^2\epsilon_1\epsilon}. \quad (40)$$

According to Hoeffding's inequality, the probability that one of the terms in  $\Delta(R_{w_y})$  is less than  $\epsilon_2$  is  $1 - \delta_2 > 1 - 2e^{-2N\epsilon_2^2}$ . Therefore, the probability that each term is less than  $\epsilon_2$  is  $(1 - \delta_2)^{d^2} > 1 - d^2\delta_2 = 1 - \tilde{\delta}_2$ .

Therefore, The matrix elements of a quantum gate do not require more measurements than

$$N_2 = O\left(\frac{\log \frac{d^2}{\tilde{\delta}_2}}{\epsilon_2^2}\right). \quad (41)$$

### F. Truncation error

Unlike the error caused by statistical fluctuations, truncation error emerges from the omission of specific quantum states during the subspace construction. We analyze disparities between two quantum circuits: one denoted as  $G$ , representing the implemented circuit within our setup, and the other denoted as  $G'$ , derived by substituting a gate layer in  $G$ . Specifically,  $G = G_1G_2G_3 \dots G_n$  constitutes an  $n$ -layer gate circuit corresponding to states  $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle$ . During subspace construction, certain states like  $|\psi_k\rangle$  might be excluded. In the case of  $G'$ ,  $|\psi_k\rangle$  is replaced with  $|\psi'_k\rangle$ , effectively eliminating the constituent  $|\phi_k\rangle$ . To discard  $|\phi_k\rangle$ , a certain condition must be satisfied

$$\left|\frac{|\Delta_k|^2}{1 + |x_k|^2}\right| < \epsilon. \quad (42)$$

In other words,  $|\Delta_k|$  can be at most  $\epsilon^{1/4}$ . Since there is an error in estimating the eigenvalues of  $g$ , the maximum value of  $|\Delta_k|$  is  $\epsilon_3 = (\epsilon + O(\frac{d^2}{\sqrt{N}}))^{1/4}$ . Therefore,  $|\text{Tr}\{G'\} -$



$\text{Tr}\{G\} \sim O(n\epsilon_3)$  since we discard at most  $n$  states  $|\phi_j\rangle$ . Then  $||\text{Tr}\{G'\} - \text{Tr}\{G\}|| \sim O(nd\epsilon_3)$ . Let  $G$  and  $G'$  correspond to  $R_{w_y}$  and  $R'_{w_y}$ , respectively, then we have  $|\text{Tr}\{R'_{w_y}\} - \text{Tr}\{R_{w_y}\}| \sim O(dn\epsilon_3)$ .

**V. NUMERICAL SIMULATION**

Initially, we conducted numerical simulations to compute the von Neumann entropy for a randomly generated collection of quantum states. Subsequently, we simulated the real-time evolution of the Ising model, incorporating stochastic noise. In this second simulation, the noise characteristic of an actual quantum device was taken into account.

**A. Quantum state preparation**

In the preceding text, we postulated that for the algorithm to be effective, knowledge of the random quantum state’s preparation method is imperative. To ensure universality, we express any single-qubit gate through a combination of three fundamental rotation gates. Thus, in this paper, we opt for the  $U(\theta, \phi, \lambda)$  gate

$$U(\theta, \phi, \lambda) = \begin{pmatrix} \cos(\theta/2) & -e^{i\lambda} \sin(\theta/2) \\ e^{i\phi} \sin(\theta/2) & e^{i\lambda+i\phi} \cos(\theta/2) \end{pmatrix} \quad (43)$$

as the random gate to prepare the random quantum state, allowing us to perform computations using both methods and subsequently compare the outcomes, where  $\theta, \phi,$  and  $\lambda$  are real numbers and  $i$  is the imaginary unit. By encoding this general parameterized gate into a single-qubit quantum gate in the quantum circuit, an  $n$ -qubit gate can be obtained through tensor product operation  $U^{\otimes n}$ . Different quantum states are prepared by using various quantum gates and then a random quantum state is simulated by employing a probability-weighted method. However, our random gate can naturally extend to the direct product of  $n$  distinct single-bit random gates.

**B. Processing of the calculation results of the Hadamard Test**

The random gates we need are denoted as  $U_{y_s}$ , where  $s = 1, 2, \dots, k + 1, y_s \in \{1, 2, 3, 4\}$ . The probability of the random gate  $U_{y_s}$  is denoted by  $p_{y_s}$ . Let  $U_{y_s} = U(\theta_{y_s}, \phi_{y_s}, \lambda_{y_s})$ . To simulate the computation of  $\text{Tr}\{\rho^{k+1}\}$ , the circuit we construct is shown in Fig. 5.

We iterate through  $y_s$  to obtain the corresponding  $P_y(0)$  for each set of  $U_{y_s}$ . Summing up all these cases yields the overall  $\mathcal{P}(0)$ :

$$\mathcal{P}(0) = \sum_{y=1}^{\alpha^{k+1}} \left[ \prod_{s=1}^{k+1} p_{y_s} \right] P_y(0). \quad (44)$$

Similarly, by following the same procedure we can obtain

$$\mathcal{P}(1) = \sum_{y=1}^{\alpha^{k+1}} \left[ \prod_{s=1}^{k+1} p_{y_s} \right] P_y(1). \quad (45)$$

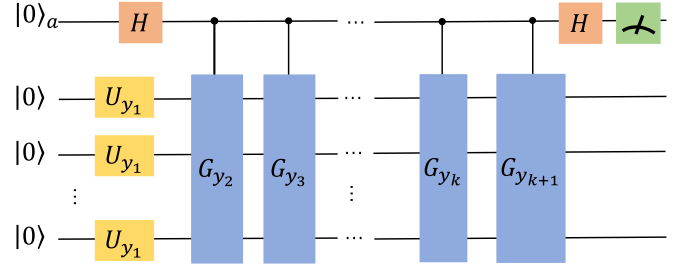


FIG. 5.  $n$ -qubits quantum circuit: The first line represents an ancillary circuit, with the initial state set to  $|0\rangle_a$ . It then undergoes a Hadamard gate, resulting in a superposition state. The three lines below represent the preparation and measurement of the target quantum states, with initial states all set to  $|0\rangle$ . After applying the first random gate  $U_{y_1}$ , the target quantum state  $\rho_{y_1} = U_{y_1}|0\rangle\langle 0|U_{y_1}^\dagger$  is obtained through tensor product. Next, controlled quantum gates  $G_{y_2}, G_{y_3}, \dots, G_{y_{k+1}}$  are applied, followed by another Hadamard gate on the ancillary circuit. Finally, measurement in the computational basis is performed on the ancillary circuit.

With Eqs. (9) and (10), we can deduce  $\text{Tr}\{G^k \rho\}$ , subsequently leading us to the derivation of  $\text{Tr}\{\rho^{m+1}\}$ .

**C. Obtaining the data and exploring potential applications**

In this study, the random gates in the circuit are simulated by selecting different parameters for  $U(\theta, \phi, \lambda)$ . For generality, the parameter selection is based on random numbers. The parameters for the random gates used in this study are shown in Table I. The probabilities for the four random gates are 0.1, 0.2, 0.3, and 0.4. Based on the above content, the values of  $\text{Tr}\{\rho^m\}$  obtained using two different methods are shown in Table II.

From the analysis of the data above, we can see that the numerical values of  $\text{Tr}\{\rho^m\}$  obtained using the HT and GST are very close. More data considering shot noise can be found in Appendix B.

However, compared to HT, the GST method does not require an additional ancillary qubit, reducing the implementation cost of the quantum circuit. This implies that practical applications could benefit from using fewer physical qubits and controlled gates, which is a critical factor to consider. Reducing the number of controlled gates could lower the error rate of the quantum circuit, thereby enhancing the system’s reliability.

Based on the obtained  $\text{Tr}\{\rho^m\}$ , it is possible to calculate  $\text{Tr}\{\rho \ln \rho\}$ ,  $\text{Tr}\{e^{\rho t}\}$ ,  $\text{Tr}\{e^{i\rho t}\}$ , and so on. For example,  $\text{Tr}\{\rho \ln \rho\}$ : the expansion of  $\rho \ln \rho$  with respect to  $G$  yields

TABLE I. The values of the parameters.

	$\theta_{y_s}$	$\phi_{y_s}$	$\lambda_{y_s}$
$i = 1$	$0.29\pi$	$0.07\pi$	$0.11\pi$
$i = 2$	$0.46\pi$	$0.62\pi$	$0.82\pi$
$i = 3$	$0.41\pi$	$0.59\pi$	$0.53\pi$
$i = 4$	$0.55\pi$	$0.31\pi$	$0.60\pi$

\*The parameters of the random gates used to generate the above data.

TABLE II. The numerical value of the parameter.

	Hadamard Test	Tomography
$\text{Tr}\{\rho^2\}$	0.650	0.650
$\text{Tr}\{\rho^3\}$	0.486	0.486
$\text{Tr}\{\rho^4\}$	0.375	0.375

\*The computational results obtained using the two algorithms in this article are presented with three decimal places.

the following expression:

$$\rho \ln \rho \approx -\frac{1}{2} \ln 2(I - G) - \frac{1}{2}G + \frac{1}{2} \left\{ \left(1 - \frac{1}{2}\right)G^2 + \left(\frac{1}{2} - \frac{1}{3}\right)G^3 + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right)G^n + \frac{1}{n}G^{n+1} \right\}. \quad (46)$$

For  $\text{Tr}\{\rho \ln \rho\}$ , the theoretical value can be obtained using the method of matrix multiplication. After expansion, it can be calculated using the two methods described in this paper.

As shown in Table III, it's evident that as the power  $m$  increases, the theoretical and experimental values converge, yielding a diminishing relative error. Upon juxtaposing Table II, one can discern the remarkable similarity between the outcomes derived from both algorithms. The code can be found in [HT](#) and [GST](#).

#### D. Simulation with real noise

To further evaluate our algorithm, we use real noise measured from the superconducting quantum computer 'ibm\_brisbane'. What we consider is a three-qubit transversal field Ising model

$$H = -J \sum_{(i,j)} \sigma_i^z \sigma_j^z - h \sum_i \sigma_i^x, \quad (47)$$

where  $J$  is the coupling constant between neighboring spins,  $\sigma_i^z$  and  $\sigma_j^z$  are the Pauli-Z operators for spins  $i$  and  $j$ , respectively, indicating the interaction between spins;  $h$  is the strength of the transverse magnetic field,  $\sigma_i^x$  is the Pauli-X operator for spin  $i$ , representing the effect of the transverse

 TABLE III. Comparison between the calculated  $\text{Tr}\{\rho \ln \rho\}$  using the method described in this paper and its theoretical value.

	$\text{Tr}\{G^m\}$	$T$ value	$M$ value	RE
$m = 2$	6.600	-0.600	-0.569	5.17%
$m = 3$	5.914	-0.600	-0.543	9.50%
$m = 4$	6.066	-0.600	-0.574	4.30%
$m = 5$	5.814	-0.600	-0.573	4.50%
$m = 6$	5.830	-0.600	-0.582	3.00%
$m = 7$	5.726	-0.600	-0.583	2.83%
$m = 8$	5.710	-0.600	-0.586	2.33%

\*Note: **T-value** represents the theoretical value of  $\text{Tr}\{\rho \ln \rho\}$ , and **M-value** represents the measured value of  $\text{Tr}\{\rho \ln \rho\}$ . **RE** represents the relative error. The value of  $\text{Tr}\{G^m\}$  is measured using the method described in this article. The measured value of  $\text{Tr}\{\rho \ln \rho\}$  can be calculated using  $\text{Tr}\{G^m\}$ .

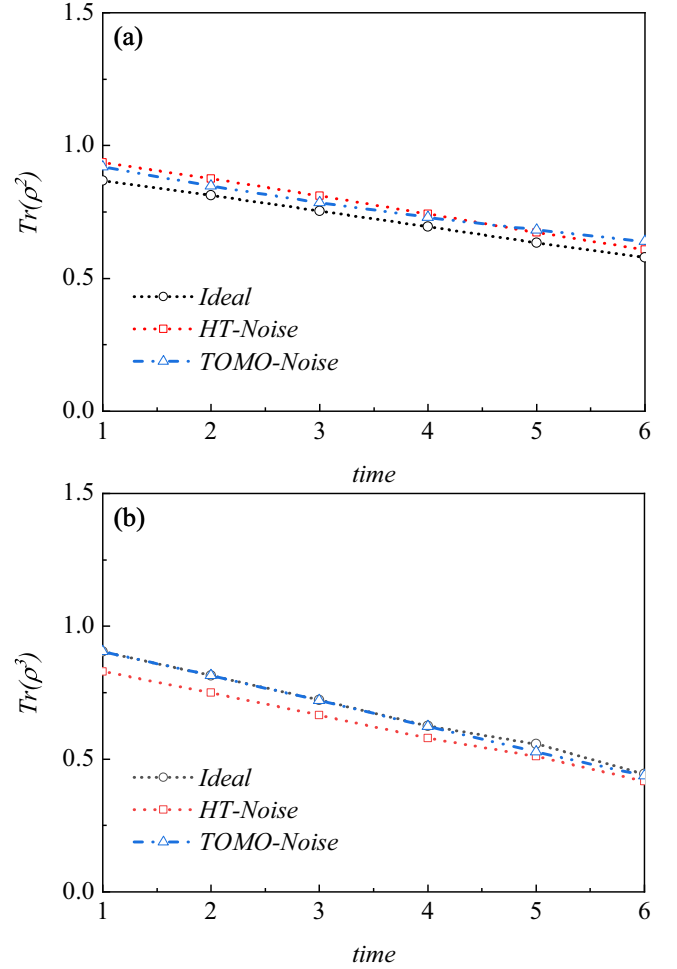


FIG. 6. (a) The function  $\text{Tr}(\rho^2)$  over evolution time. (b) The function  $\text{Tr}(\rho^3)$  over evolution time. *Ideal* denotes the theoretical values of  $\text{Tr}(\rho^2)$  and  $\text{Tr}(\rho^3)$  during the real-time evolution of the Ising model under stochastic noise  $\Lambda$ . *HT-Noise* corresponds to the values obtained for  $\text{Tr}(\rho^2)$  and  $\text{Tr}(\rho^3)$  via an algorithm utilizing the Hadamard test. *TOMO-Noise* indicates the values derived for  $\text{Tr}(\rho^2)$  and  $\text{Tr}(\rho^3)$  through a tomography-based algorithm. *HT-Noise* and *Tomo-Noise* calculations are performed within quantum circuit simulations incorporating noise from real quantum devices.

magnetic field on the spin. The notation  $\sum_{(i,j)}$  denotes summation over all nearest-neighbor spin pairs  $(i, j)$ . In our numerical simulations, we set both  $J$  and  $h$  equal to 1. We use Trotterization to achieve real-time evolution of the Ising model. The real-time evolution of the experimental Ising model will be affected by noise. We use stochastic noise to simulate this effect, that is, a channel with a fidelity of 0.95 is applied after each small step of Trotterization

$$\Lambda(\cdot) = \sum_i U_i \cdot U_i^\dagger, \quad (48)$$

where  $U_i$  is the three-qubit unitary. The dynamical state during the real-time evolution is denoted as  $\rho(t)$ . To evaluate  $\text{Tr}[\rho(t)^2]$  and  $\text{Tr}[\rho(t)^3]$  at various instances, we employed algorithms based on the Hadamard test and tomography.

As depicted in Fig. 6, the Hadamard test-based algorithm outperforms the tomography-based one in calculating

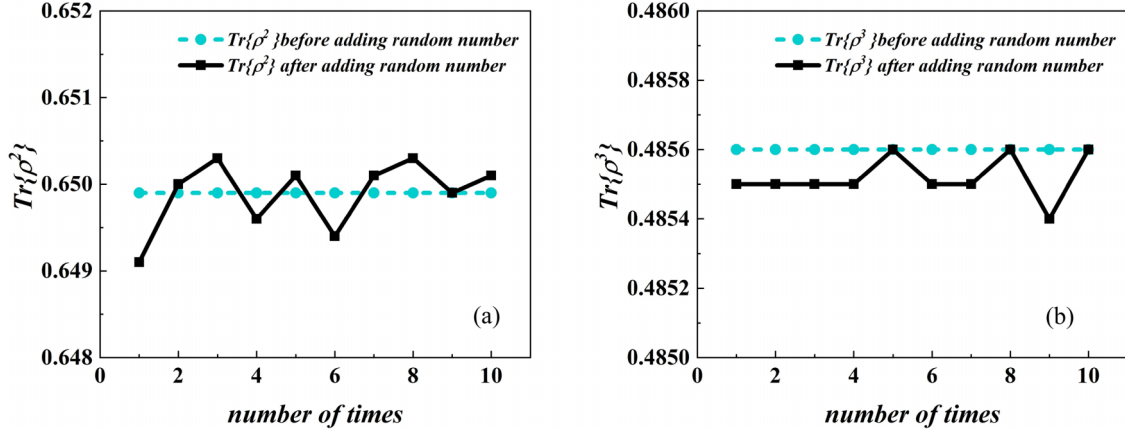


FIG. 7. This figure presents numerical results for  $\text{Tr}\{\rho^2\}$  and  $\text{Tr}\{\rho^3\}$ , taking into account noise, using the Hadamard Test method. Panel (a) displays the numerical results for  $\text{Tr}\{\rho^2\}$ , and panel (b) shows the numerical results for  $\text{Tr}\{\rho^3\}$ . The black lines represent the results after adding random numbers, while the red lines represent the results before adding random gates.

$\text{Tr}[\rho(t)^2]$ , except at the time  $t = 5$  and 6. Conversely, for most other instances, the tomography-based algorithm exhibits superior performance. This discrepancy arises because the tomography approach only necessitates the execution of a three-qubit  $G_0$  gate, whereas the Hadamard test strategy requires implementing a controlled three-qubit  $G_0$  gate, which demands additional two-qubit gates and consequently introduces more noise.

### VI. CONCLUSION

In this article, we present two algorithms for computing the power function of the density matrix by encoding the quantum state into a quantum channel. The first algorithm is based on the HT. In the absence of noise, this algorithm provides an unbiased estimate of the power function. However, it requires a significant number of control- $G_0$  gates, which is not favorable for current hardware limitations. Therefore, we propose an alternative algorithm based on GST. The original GST is not scalable, but for our specific problem, we can perform GST only within the nontrivial subspace and extract

the necessary information. The advantage of this algorithm lies in significantly reducing the utilization of two-qubit gates. Nevertheless, due to the need to mitigate the impact of shot noise on results, subspace selection introduces minor biases to the outcomes. As an example, we apply both methods to compute the von Neumann entropy of a randomly generated quantum state. Additionally, we simulated and evaluated the performance comparison of the two algorithms under noisy conditions. It is worth noting that these algorithms can be used not only for computing the power function of the density matrix but also for evaluating nonlinear functions of physical quantities such as  $\text{Tr}\{O\rho^m\}$ .

### ACKNOWLEDGMENTS

This work is supported by National Natural Science Foundation of China (Grants No. 12225507, and No.12088101) and NSAF (Grant No. U1930403). The algorithm in this article utilizes MINDQUANTUM [58] for numerical simulation.

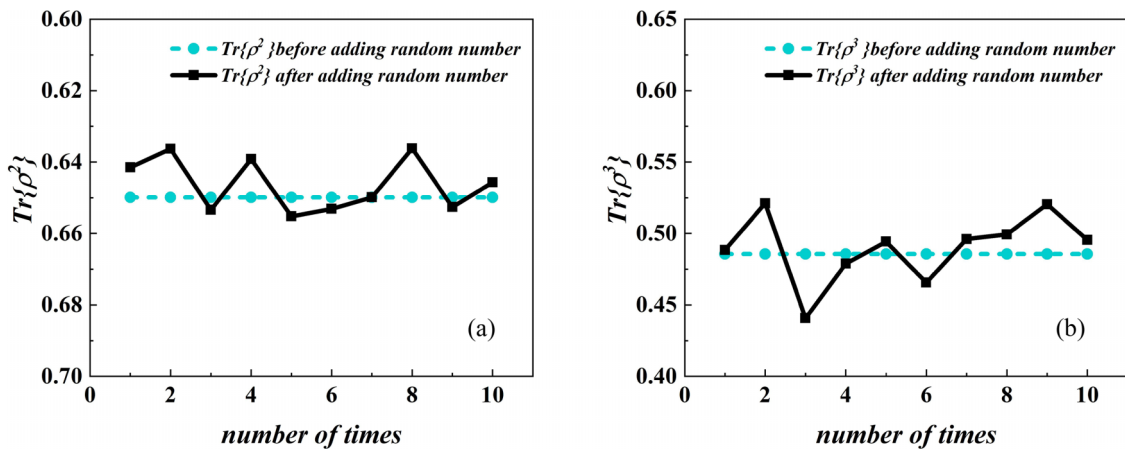


FIG. 8. This figure showcases the numerical outcomes of  $\text{Tr}\{\rho^2\}$  and  $\text{Tr}\{\rho^3\}$  calculated using the GST method while accounting for noise. Panel (a) presents the numerical results for  $\text{Tr}\{\rho^2\}$ , and panel (b) illustrates the numerical results for  $\text{Tr}\{\rho^3\}$ . The black lines denote the results obtained after the addition of random numbers, while the red lines represent the results before the introduction of random gates.

### APPENDIX A: COMPLETENESS PROOF

In the main text, the calculation of  $\text{Tr}\{\rho^k\}$  was mathematically transformed into investigating  $\text{Tr}\{G^k\}$ , where  $\text{Tr}\{G^k\}$  represented the trace of  $G^k$ , which was obtained by decomposing the random state into pure states and then taking a weighted average. Therefore, the primary focus is on calculating  $\text{Tr}\{G_q^k\}$ .

After selecting a combination  $G_q^k : G_{q_1} G_{q_2} \cdots G_{q_i} \cdots G_{q_k}$ , it is necessary to determine the dimension of the subspace  $d$ . For the subspace determined by the combination  $G_q^k$ , if the dimension of the subspace is  $d$ , then the quantum states  $\rho_s$  prepared by different random gates of dimension  $d$ , where  $s \in \{1, 2, \dots, d\}$ , are guaranteed to be linearly independent. Preparation of  $d$  linearly independent quantum states using  $d$   $n$ -bit random gates:  $\rho_s = U_s |0\rangle\langle 0|^{\otimes n} U_s^\dagger$ . The random gate  $U_s$  belongs to different random gates within a certain combination  $G_q^k$ .

Preparation of the remaining  $d^2 - d$  quantum states: Assuming we already prepared  $d$  linearly independent quantum states  $\rho_1, \rho_2, \dots, \rho_d$ , we need to prepare an additional  $d^2 - d$  quantum states to form a complete  $d^2$ -dimensional inner product space. In a finite-dimensional linear space, it is sufficient for all  $d^2$  quantum states to be linearly independent from each other.

For the quantum state  $\rho_s = U_s |0\rangle\langle 0|^{\otimes n} U_s^\dagger$ , where  $s, s' \in \{1, 2, \dots, d\}$ ,  $s \neq s'$ , we apply the quantum gate  $G_{s'} = U_{s'} G_0 U_{s'}^\dagger = I - 2\rho_{s'}$  to obtain

$$\begin{aligned} \rho_{ss'} &= G_{s'} \rho_s G_{s'}^\dagger = (I - 2\rho_{s'}) \rho_s (I - 2\rho_{s'}) \\ &= \rho_s - 2\rho_{s'} \rho_s - 2\rho_s \rho_{s'} + 4\rho_{s'} \rho_s \rho_{s'}. \end{aligned} \quad (\text{A1})$$

For the corresponding  $\rho_{s's}$ , we have

$$\rho_{s's} = \rho_{s'} - 2\rho_{s'} \rho_s - 2\rho_s \rho_{s'} + 4\rho_{s'} \rho_s \rho_{s'}. \quad (\text{A2})$$

It can be shown that the quantum states  $\rho_s, \rho_{s'}, \rho_{ss'}, \rho_{s's}$  are linearly dependent. If we construct all quantum states using this method, then these quantum states will be linearly dependent and cannot form a complete  $d^2$ -dimensional inner product space.

To ensure linear independence, we modify the gate  $G_0$  to  $G_\theta$ , where  $\theta \neq k\pi$ ,  $k \in \mathbb{Z}$ :

$$\begin{aligned} G_0 &= \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \\ \rightarrow G_\theta &= \begin{pmatrix} e^{i\theta} & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} - \begin{pmatrix} 1 - e^{i\theta} & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}. \end{aligned} \quad (\text{A3})$$

After this change, the transformed  $G_s(\theta) = U_s G_\theta U_s^\dagger$ . We represent  $G_s(\theta) = U_s G(\theta) U_s^\dagger$  as  $G_s(\theta) = I - \rho_s(\theta)$ , and  $\rho_s(\theta) \neq$

$\rho_s$ . Following the same approach as before, we prepare the quantum states

$$\begin{aligned} 1\rho_s &= U_s |0\rangle\langle 0|^{\otimes n} U_s^\dagger, \\ 2\rho_{s'} &= U_{s'} |0\rangle\langle 0|^{\otimes n} U_{s'}^\dagger, \\ 3\rho_{ss'} &= (I - 2\rho_{s'}) \rho_s (I - 2\rho_{s'}) \\ &= \rho_s - 2\rho_{s'} \rho_s - 2\rho_s \rho_{s'} + 4\rho_{s'} \rho_s \rho_{s'}, \\ 4\rho_{s's}(\theta) &= G_{s'}(\theta) \rho_s G_{s'}(\theta)^\dagger \\ &= [I - \rho_{s'}(\theta)] \rho_s [I - \rho_{s'}(\theta)] \\ &= \rho_s - \rho_{s'} \rho_s(\theta) - \rho_s(\theta) \rho_{s'} + \rho_{s'}(\theta) \rho_s \rho_{s'}(\theta). \end{aligned} \quad (\text{A4})$$

From the above equation, we observe that rotating the quantum gate  $G_0$  by an arbitrary angle  $\theta \neq k\pi$ ,  $k = 0, 1, 2, \dots$ , changes the originally linearly dependent quantum states, prepared using  $G_0$ , into linearly independent ones.

By using similar methods, it is possible to prepare the remaining  $d(d-1)$  quantum states and ensure that they are linearly independent. This allows us to construct a complete  $d^2$ -dimensional linear space for performing a full tomography of the PTM under study.

### APPENDIX B: MORE NUMERICAL SIMULATION

For the Hadamard Test procedure, we incorporate noise simulation into its measurement process by injecting random numbers drawn from a Gaussian distribution with a standard deviation of 0.01. Figure 7 illustrates the outcomes prior to and subsequent to the integration of these random numbers. For tomography, we introduce noise simulation by adding random numbers to its PTM elements as well as  $g$ -matrix elements. Since the singular values of the  $g$ -matrix are small, we choose random numbers sampled from a Gaussian distribution with a standard deviation of 0.0001 to simulate the noise. The results before and after adding the random numbers are shown in Fig. 8.

Derived from the data depicted in the provided graph, we are able to distinctly discern the repercussions stemming from the introduction or absence of noise. When noise is absent, the output data from the quantum circuit manifests as consistently stable and precise. Nonetheless, the scenario takes a discernible turn upon the infusion of noise. The introduction of noise precipitates volatility in the quantum circuit's outcomes, potentially culminating in errors. Noise can engender an unreliable exchange of information amidst quantum bits, thereby injecting substantial ambiguity into the computational process. These sources of noise might encompass phenomena such as dephasing, errors in quantum gate operations, imprecisions in measurements, or other external disturbances.

In real-world scenarios, noise presents a significant hurdle for the advancement of quantum computing and quantum information processing. To counteract the disruptive influence of noise, researchers are tirelessly engaged in refining quantum error correction and noise suppression techniques. The primary goal of these approaches is to bolster the resilience of quantum circuits and heighten the precision of outcomes, ultimately striving for a more dependable realm of quantum computing. Furthermore, the ramifications of noise

can exhibit variability across diverse quantum algorithms and tasks. As a result, practical applications necessitate a thorough assessment of algorithm robustness alongside the prevailing noise levels. Achieving equilibrium between these considerations is crucial in determining the most optimal course of action.

### APPENDIX C: SUBSPACE DIMENSION

In this Appendix, we study the dimensions of the invariant subspaces corresponding to nonlinear functions of different types of density matrices. Let's start with the simplest case,  $\text{Tr}\{\rho^m\}$ . In our algorithm, we decompose the computation of  $\text{Tr}\{\rho^m\}$  into the computation of  $\text{Tr}\{G_{q_1}G_{q_2}\dots G_{q_k}\}$ , for  $k = 0, 1, \dots, m$ . It is obvious that  $\text{span}(|\psi_{q_1}\rangle, |\psi_{q_2}\rangle, \dots, |\psi_{q_k}\rangle)$

forms a nontrivial invariant subspace of  $G_{q_1}G_{q_2}\dots G_{q_k}$ . Therefore, the largest nontrivial invariant subspace involved in the calculation of  $\text{Tr}\{\rho^m\}$  is  $m$ -dimensional.

A more general nonlinear function is  $\text{Tr}\{P_1\rho_1P_2\rho_2\dots P_m\rho_m\}$ . We can rewrite this function as  $\text{Tr}\{P'\rho'_1\rho'_2\dots\rho'_m\}$ , where  $\rho'_i = (\prod_{l=1}^i P_l)\rho_n(\prod_{l=1}^i P_l)^\dagger$  and  $P' = \prod_{l=1}^m P_l$ . Remark that although  $\rho'_1 \neq \rho'_2 \neq \dots \neq \rho'_m$ , the algorithms we proposed in this paper can also be used. During the calculation, we need to calculate  $\text{Tr}\{P'G_{q_1}G_{q_2}\dots G_{q_k}\}$ , for  $k = 0, 1, \dots, m$ , where  $P'$  is a Pauli operator. Note that  $G_{q_i}$  acts on  $|\psi\rangle$  to get a superposition of  $|\psi_{q_i}\rangle$  and  $|\psi\rangle$ . Thus the nontrivial invariant subspace corresponding to  $PG_{q_1}G_{q_2}\dots G_{q_k}$  is  $\text{span}(|\psi_{q_1}\rangle, \dots, |\psi_{q_m}\rangle, P|\psi_{q_1}\rangle, \dots, P|\psi_{q_m}\rangle)$ , of which the largest possible dimension is  $2m$  dimensions.

- 
- [1] B. Swingle, G. Bentsen, M. Schleier-Smith, and P. Hayden, *Phys. Rev. A* **94**, 040302(R) (2016).
- [2] F. G. S. L. Brandão, W. Chemissany, N. Hunter-Jones, R. Kueng, and J. Preskill, *PRX Quantum* **2**, 030316 (2021).
- [3] P. Hayden and J. Preskill, *J. High Energy Phys.* **09** (2007) 120.
- [4] J. Kudler-Flam, *Phys. Rev. Lett.* **126**, 171603 (2021).
- [5] Z. Holmes, N. J. Coble, A. T. Sornborger, and Y. Subaşı, *Phys. Rev. Res.* **5**, 013105 (2023).
- [6] S. Subramanian and M.-H. Hsieh, *Phys. Rev. A* **104**, 022428 (2021).
- [7] K. Zyczkowski and I. Bengtsson, An introduction to quantum entanglement: A geometric approach, [arXiv:quant-ph/0606228](https://arxiv.org/abs/quant-ph/0606228).
- [8] A. Rath, C. Branciard, A. Minguzzi, and B. Vermersch, *Phys. Rev. Lett.* **127**, 260501 (2021).
- [9] R. Jozsa, *J. Mod. Opt.* **41**, 2315 (1994).
- [10] B. Koczor, *Phys. Rev. X* **11**, 031057 (2021).
- [11] Q. Wang, J. Guan, J. Liu, Z. Zhang, and M. Ying, New quantum algorithms for computing quantum entropies and distances, *IEEE Trans. Inf. Theory* (2024), doi:10.1109/TIT.2024.3399014.
- [12] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, *Nature (London)* **549**, 242 (2017).
- [13] H. A. Carteret, *Phys. Rev. Lett.* **94**, 040502 (2005).
- [14] M. Lubasch, J. Joo, P. Moinier, M. Kiffner, and D. Jaksch, *Phys. Rev. A* **101**, 010301(R) (2020).
- [15] B. Georgeot and D. L. Shepelyansky, *Phys. Rev. Lett.* **86**, 2890 (2001).
- [16] A. Elben, J. Yu, G. Zhu, M. Hafezi, F. Pollmann, P. Zoller, and B. Vermersch, *Sci. Adv.* **6**, eaaz3666 (2020).
- [17] J. Braumüller, A. H. Karamlou, Y. Yanay, B. Kannan, D. Kim, M. Kjaergaard, A. Melville, B. M. Niedzielski, Y. Sung, A. Vepsäläinen *et al.*, *Nat. Phys.* **18**, 172 (2022).
- [18] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. Sawaya *et al.*, *Chem. Rev.* **119**, 10856 (2019).
- [19] M. Schuld, [arXiv:2101.11020](https://arxiv.org/abs/2101.11020).
- [20] M. Schuld and N. Killoran, *Phys. Rev. Lett.* **122**, 040504 (2019).
- [21] S. Lloyd, G. De Palma, C. Gokler, B. Kiani, Z.-W. Liu, M. Marvian, F. Tennie, and T. Palmer, [arXiv:2011.06571](https://arxiv.org/abs/2011.06571).
- [22] D. Herman, C. Googin, X. Liu, A. Galda, I. Safro, Y. Sun, M. Pistoia, and Y. Alexeev, [arXiv:2201.02773](https://arxiv.org/abs/2201.02773).
- [23] F. Gaitan, *npj Quantum Inf.* **6**, 61 (2020).
- [24] F. Oz, R. K. Vuppala, K. Kara, and F. Gaitan, *Quantum Info. Proc.* **21**, 30 (2022).
- [25] W. J. Huggins, S. McArdle, T. E. O'Brien, J. Lee, N. C. Rubin, S. Boixo, K. B. Whaley, R. Babbush, and J. R. McClean, *Phys. Rev. X* **11**, 041036 (2021).
- [26] National Academies of Sciences and Medicine, *Quantum Computing Progress and Prospects*, edited by E. Grumbling and M. Horowitz (The National Academies Press, Washington, DC, 2019).
- [27] A. J. Daley, I. Bloch, C. Kokail, S. Flannigan, N. Pearson, M. Troyer, and P. Zoller, *Nature (London)* **607**, 667 (2022).
- [28] D. Hangleiter and J. Eisert, *Rev. Mod. Phys.* **95**, 035001 (2023).
- [29] J. Preskill, *Quantum* **2**, 79 (2018).
- [30] J. W. Z. Lau, K. H. Lim, H. Shrotriya, and L. C. Kwek, *AAPPS Bulletin* **32**, 27 (2022).
- [31] Y. Ding and F. T. Chong, *Quantum Computer Systems: Research for Noisy Intermediate-Scale Quantum Computers* (Springer Nature, New York, 2022).
- [32] Y. Zhou and Z. Liu, [arXiv:2208.08416](https://arxiv.org/abs/2208.08416).
- [33] F. A. Bovino, G. Castagnoli, A. Ekert, P. Horodecki, C. M. Alves, and A. V. Sergienko, *Phys. Rev. Lett.* **95**, 240407 (2005).
- [34] P. Horodecki, *Phys. Rev. Lett.* **90**, 167901 (2003).
- [35] A. K. Ekert, C. M. Alves, D. K. L. Oi, M. Horodecki, P. Horodecki, and L. C. Kwek, *Phys. Rev. Lett.* **88**, 217901 (2002).
- [36] S. H. Sack, R. A. Medina, A. A. Michailidis, R. Kueng, and M. Serbyn, *PRX Quantum* **3**, 020365 (2022).
- [37] T. Zhang, J. Sun, X.-X. Fang, X.-M. Zhang, X. Yuan, and H. Lu, *Phys. Rev. Lett.* **127**, 200501 (2021).
- [38] A. Seif, Z.-P. Cian, S. Zhou, S. Chen, and L. Jiang, *PRX Quantum* **4**, 010303 (2023).
- [39] A. Elben, S. T. Flammia, H.-Y. Huang, R. Kueng, J. Preskill, B. Vermersch, and P. Zoller, *Nat. Rev. Phys.* **5**, 9 (2023).
- [40] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, *Science* **364**, 260 (2019).

- [41] A. Elben, B. Vermersch, C. F. Roos, and P. Zoller, *Phys. Rev. A* **99**, 052323 (2019).
- [42] A. Elben, R. Kueng, H.-Y. R. Huang, R. van Bijnen, C. Kokail, M. Dalmonte, P. Calabrese, B. Kraus, J. Preskill, P. Zoller *et al.*, *Phys. Rev. Lett.* **125**, 200501 (2020).
- [43] R. O’Donnell and J. Wright, in *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing* (ACM, New York, 2016), pp. 899–912.
- [44] G. M. D’Ariano, M. G. Paris, and M. F. Sacchi, *Adv. Imaging Electron Phys.* **128**, 206 (2003).
- [45] R. Yang and Y. Li, *Phys. Rev. A* **103**, 032421 (2021).
- [46] B. Wu, M. Ray, L. Zhao, X. Sun, and P. Rebentrost, *Phys. Rev. A* **103**, 042422 (2021).
- [47] L. Xu, X.-Y. Zhang, J.-M. Liang, J. Wang, M. Li, L. Jian, and S.-q. Shen, *Commun. Theor. Phys.* **74**, 055106 (2022).
- [48] T. L. Patti, J. Kossai, A. Anandkumar, and S. F. Yelin, *Quantum* **7**, 1057 (2023).
- [49] D. Aharonov, V. Jones, and Z. Landau, in *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing* (ACM, New York, 2006), pp. 427–436.
- [50] D. Greenbaum, [arXiv:1509.02921](https://arxiv.org/abs/1509.02921).
- [51] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, *Nat. Phys.* **14**, 447 (2018).
- [52] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, *Nat. Commun.* **1**, 149 (2010).
- [53] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, *Phys. Rev. A* **77**, 032322 (2008).
- [54] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O’Brien, M. A. Nielsen, and A. G. White, *Phys. Rev. Lett.* **90**, 193601 (2003).
- [55] B. Lanyon, C. Maier, M. Holzäpfel, T. Baumgratz, C. Hempel, P. Jurcevic, I. Dhand, A. Buyskikh, A. Daley, M. Cramer *et al.*, *Nat. Phys.* **13**, 1158 (2017).
- [56] R. Blume-Kohout, J. K. Gamble, E. Nielsen, J. Mizrahi, J. D. Sterk, and P. Maunz, [arXiv:1310.4492](https://arxiv.org/abs/1310.4492).
- [57] M. G. Cowling and J. F. Price, *SIAM J. Math. Anal.* **15**, 151 (1984).
- [58] M. Developer, “MINDQUANTUM, version 0.9.11,” (2021).