

A Quantum Hamiltonian Identification Algorithm: Computational Complexity and Error Analysis

Yuanlong Wang, Daoyi Dong, Bo Qi, Jun Zhang, Ian R. Petersen, Hidehiro Yonezawa

Abstract—Quantum Hamiltonian identification is important for characterizing the dynamics of quantum systems, calibrating quantum devices and achieving precise quantum control. In this paper, an effective two-step optimization (TSO) quantum Hamiltonian identification algorithm is developed within the framework of quantum process tomography. In the identification method, different probe states are inputted into quantum systems and the output states are estimated using the quantum state tomography protocol via linear regression estimation. The time-independent system Hamiltonian is reconstructed based on the experimental data for the output states. The Hamiltonian identification method has computational complexity $O(d^6)$ where d is the dimension of the system Hamiltonian. An error upper bound $O(\frac{d^3}{\sqrt{N}})$ is also established, where N is the resource number for the tomography of each output state, and several numerical examples demonstrate the effectiveness of the proposed TSO Hamiltonian identification method.

Index Terms—Quantum system, Hamiltonian identification, process tomography, computational complexity.

I. INTRODUCTION

AS quantum technology develops, there is an increasing demand for characterizing an unknown quantum process since it is vital to verify and benchmark quantum devices for quantum computation, communication and metrology [1]. The standard solution to characterizing a quantum process is Quantum Process Tomography (QPT), wherein usually known input quantum states (probe states) are applied to the process and the output states are measured to reconstruct the quantum process [2]-[4]. Hamiltonian identification for closed quantum systems is a special class of QPT that corresponds to a unitary quantum process and is an essential component to characterize the dynamics of a quantum system.

System identification has been widely investigated in classical (non-quantum) systems and control theory, and many

identification algorithms have been developed to estimate unknown dynamical parameters of linear or nonlinear input-output systems [5]-[7]. In recent years, the problem of quantum system identification has attracted more and more attention due to the rapid development of emerging quantum technology [8], [9] and increasing demand of characterizing quantum devices. For example, a framework for quantum system identification has been established in [10] to classify how much knowledge about a quantum system is attainable from a given experimental setup. Guřa and Yamamoto [11] considered a class of passive linear quantum input-output systems, and investigated the problem of identifiability and how to optimize the identification precision by preparing good input states and performing appropriate measurements on the output states.

In this paper, we focus on the problem of quantum Hamiltonian identification (QHI), which is a key task in characterizing the dynamics of quantum systems and achieving high-precision quantum control. There exist some results on QHI and various aspects of QHI have been investigated [12]-[15]. For example, a symmetry-preserving observer has been developed for the Hamiltonian identification of a two-level quantum system [16]. The identifiability problem for a Hamiltonian corresponding to a dipole moment has been investigated [17] and the question of how to utilize quantum control to identify such Hamiltonian has been addressed [18]. Closed-loop learning control has been presented to optimally identifying Hamiltonian information [19] and compressed sensing has been proposed to enhance the efficiency of identification algorithms for Hamiltonian with special structures [20], [21]. Several Hamiltonian identification algorithms have been developed using only measurement in a single fixed basis [22]-[24]. Wang *et al.* [25] utilized dynamical decoupling to identify Hamiltonians for quantum many-body systems with arbitrary couplings. Cole *et al.* [26] discussed the estimation error in identifying a two-state Hamiltonian and Zhang *et al.* [27] presented a QHI protocol using measurement time traces. Most of these existing results have limitations for practical applications (e.g., estimating a single parameter [24], [28], identifying special Hamiltonian [22], [26]), and there are few theoretical results on the analysis of computational complexity and upper bounds on estimation errors. This paper presents an identification algorithm for general time-independent Hamiltonians and analyzes its computational complexity and upper bounds on estimation errors.

Our quantum Hamiltonian identification is presented within the framework of quantum process tomography. Some different input states are prepared for quantum systems and the

This work was supported by the Australian Research Council's Discovery Projects funding scheme under Project DP130101658, Laureate Fellowship FL110100020, Centres of Excellence CE110001027 and the National Natural Science Foundation of China under Grant No. 61374092.

Y. Wang and H. Yonezawa are with the School of Engineering and Information Technology, University of New South Wales, Canberra, ACT 2600, Australia, and also with the Centre for Quantum Computation and Communication Technology, Australian Research Council, Canberra, ACT 2600, Australia (e-mail: yuanlong.wang.qc@gmail.com; h.yonezawa@adfa.edu.au).

D. Dong and I. R. Petersen are with the School of Engineering and Information Technology, University of New South Wales, Canberra, ACT 2600, Australia (e-mail: daoyidong@gmail.com; i.r.petersen@gmail.com).

B. Qi is with the Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China and with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: qibo@amss.ac.cn).

J. Zhang is with the Joint Institute of UM-SJTU, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: zhangjun12@sjtu.edu.cn).

corresponding output states are measured after a fixed time evolution under the Hamiltonian to be identified. These output states are reconstructed using the quantum state tomography technique via linear regression estimation (LRE) [29]. Using the information of estimated output states, the Hamiltonian is reconstructed via an identification algorithm. The main contributions of this paper are summarized as follows.

- The quantum Hamiltonian identification (QHI) problem is formulated within the framework of quantum process tomography (QPT) and several relevant points in QPT are clarified in order to present an efficient QHI algorithm.
- A Two-Step Optimization (TSO) identification algorithm is presented and its computational complexity is analyzed. Our identification algorithm has the computational complexity $O(d^6)$ where d is the dimension of the quantum system.
- Analytical results of estimation error are presented and an error upper bound is established as $O(\frac{d^3}{\sqrt{N}})$, where N is the resource number in the tomography of each output state.
- Numerical examples are presented to demonstrate the performance of our QHI algorithm. It is then compared with the QHI method using measurement time traces in [27], and our identification algorithm shows an efficiency advantage over the method in [27] in terms of the computational time.

The structure of this paper is as follows. In Section II we present some preliminaries and briefly introduce QPT. Section III formulates the QHI problem within the framework of QPT. Section IV presents a TSO Hamiltonian identification algorithm and analyzes the computational complexity. Section V analyzes the estimation error theoretically and establishes an upper bound. In Section VI, we present two numerical examples to demonstrate performance and also compare our identification algorithm with the QHI method using time traces in [27]. Section VII concludes this paper.

Notation: a^* denotes the conjugate of a ; $A_{m \times n}$ denotes an m -row and n -column matrix; A^T is the transpose of A ; A^\dagger is the conjugate and transpose of A ; \mathbb{H} denotes a Hilbert space; \mathbb{R} and \mathbb{C} are the sets of all real and complex numbers, respectively; I is the identity matrix (dimension omitted if without ambiguity); $\|A\|$ denotes the Frobenius norm of A ; $\text{Tr}(A)$ is the trace of A ; $|\psi\rangle$ is a unit complex vector representing a quantum (pure) state; ρ is a density matrix representing a quantum state; \hat{a} is the estimate of a ; $\langle A, B \rangle$ represents the inner product of A and B defined as $\langle A, B \rangle = \text{Tr}(A^\dagger B)$; $\langle a, b \rangle$ denotes the inner product of a and b with $\langle a, b \rangle = a^\dagger b$; \mathbb{C}_d is the set of all d -dimension complex vectors; $\mathbb{C}_{d \times d}$ is the set of all $d \times d$ complex matrices; $\text{vec}(\cdot)$ denotes the vectorization function; $\text{vec}^{-1}(\cdot)$ is the inverse function of vectorization from \mathbb{C}_{d^2} to $\mathbb{C}_{d \times d}$; $A \otimes B$ denotes the tensor product of A and B ; $\text{Tr}_1(X)$ means partial trace on space \mathbb{H}_1 where $X \in \mathbb{H}_1 \otimes \mathbb{H}_2$; δ is the Dirac Delta function; i as a subscript means an integer index, otherwise i means imaginary unit, i.e., $i = \sqrt{-1}$.

II. PRELIMINARIES AND QUANTUM PROCESS TOMOGRAPHY

A. Matrix and Vectorization Fundamentals

For a matrix $A_{m \times n} = [a_{ij}]$, its Frobenius norm is defined as

$$\|A\| = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2} = \sqrt{\text{Tr}(A^\dagger A)}.$$

Two important properties of the Frobenius norm are:

$$\|A\| = \|UA\|, \quad (1)$$

$$\|AB\| \leq \|A\| \cdot \|B\|, \quad (2)$$

where U is any $m \times m$ unitary matrix.

The tensor product of matrices $A_{m \times n} = [a_{ij}]$ and $B_{p \times q} = [b_{kl}]$ is defined as follows:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \dots & \dots & \dots & \dots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}_{mp \times nq}. \quad (3)$$

We introduce the vectorization function $\text{vec} : \mathbb{C}_{m \times n} \mapsto \mathbb{C}_{mn}$. For a matrix $A_{m \times n} = [a_{ij}]$,

$$\text{vec}(A_{m \times n}) = [a_{11}, a_{21}, \dots, a_{m1}, a_{12}, \dots, a_{m2}, \dots, a_{1n}, \dots, a_{mn}]^T.$$

The function $\text{vec}(\cdot)$ (also denoted as $|\cdot\rangle$ or $|\cdot\rangle$ in the physics community) is linear. Its common properties are listed as follows [30], [31]:

$$\text{vec}(|a\rangle\langle b|) = |b\rangle^* \otimes |a\rangle, \quad (4)$$

$$\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X), \quad (5)$$

$$\langle A, B \rangle = \langle \text{vec}(A), \text{vec}(B) \rangle, \quad (6)$$

$$\text{Tr}_1(\text{vec}(A)\text{vec}(B)^\dagger) = AB^\dagger. \quad (7)$$

$$\text{Tr}_2(\text{vec}(A)\text{vec}(B)^\dagger) = (B^\dagger A)^T. \quad (8)$$

In this paper, we also define that $\text{vec}^{-1}(\cdot)$ maps a $d^2 \times 1$ vector into a $d \times d$ square matrix rather than matrices with other sizes. In (7) and (8), $\text{Tr}_1(X)$ means partial trace on the space \mathbb{H}_1 where X belongs to the space $\mathbb{H}_1 \otimes \mathbb{H}_2$. Similarly $\text{Tr}_2(X)$ means partial trace on the space \mathbb{H}_2 .

B. Quantum System and Evolution

The state of a closed quantum system can be described by a unit complex vector $|\psi\rangle$ in the underlying Hilbert space and its dynamics is governed by the Schrödinger equation

$$i \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle, \quad (9)$$

where H is the system Hamiltonian and we set $\hbar = 1$ using atomic units in this paper. When the quantum system under consideration is an open quantum system or the quantum state is a mixed state, we need to use a Hermitian positive semidefinite matrix ρ satisfying $\text{Tr}(\rho) = 1$ to describe the quantum state. For a closed quantum system with state $|\psi\rangle$, we have $\rho = |\psi\rangle\langle\psi|$. Its evolution from the initial state $\rho(0)$

to $\rho(t)$ at time t can be determined by a unitary propagator U :

$$\rho(t) = U(t)\rho(0)U^\dagger(t), \quad (10)$$

where $U(t) = e^{-iHt}$ if H is independent of t .

For an open quantum system, the dynamics of its state can be described by a master equation. Alternatively, the transformation from an input state ρ_{in} to an output state ρ_{out} is given by Kraus operator-sum representation [1]

$$\rho_{out} = \mathcal{E}(\rho_{in}) = \sum_i A_i \rho_{in} A_i^\dagger, \quad (11)$$

where the quantum operation \mathcal{E} maps ρ_{in} to ρ_{out} and $\{A_i\}$ is a set of mappings from the input Hilbert space to the output Hilbert space with $\sum_i A_i^\dagger A_i \leq I$. In this paper, we only consider trace-preserving operations which means that the completeness relation

$$\sum_i A_i^\dagger A_i = I \quad (12)$$

is satisfied. In particular, we consider d -dimensional quantum systems and have $A_i \in \mathbb{C}_{d \times d}$.

C. Quantum Measurement and Quantum State Tomography

We aim to identify the system Hamiltonian H from the input states (usually known) and the output states. To extract information from the output quantum states, a positive-operator valued measurement (POVM) is usually performed on these states. A POVM is a set $\{M_i\}$, where all the elements are Hermitian positive semidefinite and $\sum_i M_i = I$. When a set of POVM is performed, the probability of outcome i occurring is determined by the Born Rule $p_i = \text{Tr}(\rho M_i)$. A special class of POVM are the projective measurement operators $\{P_i\}$, which are projectors satisfying $P_i P_j = \delta_{ij} P_j$.

In real experiments, it is impossible to implement infinitely many measurements. Hence, p_i can only be approximated within a limited accuracy. The methodology for designing the measurement operators $\{M_i\}$ and estimating ρ from experimental data is called quantum state tomography, where usually a large number of identified independent copies of ρ are used. Common quantum state tomography methods include Maximum likelihood estimation [32]-[34], Bayesian mean estimation [32], [35] and linear regression estimation (LRE) [29]. In this paper, the LRE method will be used in numerical simulation and error analysis although our QHI method is also applicable to other quantum state tomography methods. In the LRE method, the quantum state reconstruction problem is converted into a parameter estimation problem for a linear regression model and the least-squares method can be used to obtain estimates of the unknown parameters. The LRE method of quantum state tomography was first presented in [29] and it has also been used to experimentally reconstruct quantum states for various tasks [36], [37]. Its advantages of high efficiency and an analytical error upper bound make it especially beneficial in presenting numerical results and error analysis for our TSO QHI method.

D. Standard Quantum Process Tomography

We rephrase the framework of general quantum process tomography in [1] in the matrix form and later we will consider QHI problem under this framework.

By expanding $\{A_i\}$ in (11) in a fixed family of basis matrices $\{E_i\}$, we obtain

$$A_i = \sum_j c_{ij} E_j, \quad (13)$$

and then

$$\mathcal{E}(\rho) = \sum_{jk} E_j \rho E_k^\dagger x_{jk},$$

with $x_{jk} = \sum_i c_{ij} c_{ik}^*$. If we define the matrix $C = [c_{ij}]$ and the matrix $X = [x_{ij}]$, then

$$X = C^T C^*, \quad (14)$$

which indicates that X must be Hermitian and positive semidefinite. X is called the *process matrix* [38]. The completeness constraint equation (12) becomes

$$\sum_{j,k} x_{jk} E_k^\dagger E_j = I. \quad (15)$$

It is difficult to further simplify this relationship before the structure of $\{E_i\}$ is determined. Note that the matrix X and the process \mathcal{E} are in a one-to-one correspondence. Hence, we can obtain a full characterization of \mathcal{E} by reconstructing X [1].

Let $\{\rho_m\}$ be a complete basis set of $\mathbb{C}_{d \times d}$. For example, all Pauli matrices $\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ and $\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, together with $I_{2 \times 2}$, form a complete basis set of $\mathbb{C}_{2 \times 2}$. If we let $\{\rho_m\}$ be linearly independent matrices (with respect to addition between matrices, and multiplication between a scalar and a matrix) and we input ρ_m to the process, then each process output can be expanded uniquely in the basis set $\{\rho_n\}$; i.e.,

$$\rho_{out} = \mathcal{E}(\rho_{in}) = \mathcal{E}(\rho_m) = \sum_n \lambda_{mn} \rho_n. \quad (16)$$

For simplicity, we choose $\{\rho_n\}$ to be the same set as $\{\rho_m\}$ although they could be different. We then need to find the relationship between X and λ , which is independent of the bases $\{E_i\}$. Considering the effects of the bases $\{\rho_n\}$ on $\{\rho_m\}$, we have

$$E_j \rho_m E_k^\dagger = \sum_n \beta_{mn}^{jk} \rho_n. \quad (17)$$

Hence,

$$\sum_n \sum_{jk} \beta_{mn}^{jk} \rho_n x_{jk} = \sum_n \lambda_{mn} \rho_n.$$

From the linear independence of $\{\rho_n\}$, one can obtain

$$\sum_{jk} \beta_{mn}^{jk} x_{jk} = \lambda_{mn}. \quad (18)$$

To rewrite this equation into a compact form, define the matrix $\Lambda = [\lambda_{mn}]$ and arrange the elements β_{mn}^{jk} into a matrix B :

$$B = \begin{pmatrix} \beta_{11}^{11} & \beta_{11}^{21} & \cdots & \beta_{11}^{12} & \beta_{11}^{22} & \cdots & \beta_{11}^{d^2 d^2} \\ \beta_{21}^{11} & \beta_{21}^{21} & \cdots & \beta_{21}^{12} & \beta_{21}^{22} & \cdots & \beta_{21}^{d^2 d^2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \beta_{12}^{11} & \beta_{12}^{21} & \cdots & \beta_{12}^{12} & \beta_{12}^{22} & \cdots & \beta_{12}^{d^2 d^2} \\ \beta_{22}^{11} & \beta_{22}^{21} & \cdots & \beta_{22}^{12} & \beta_{22}^{22} & \cdots & \beta_{22}^{d^2 d^2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \beta_{d^2 d^2}^{11} & \beta_{d^2 d^2}^{21} & \cdots & \beta_{d^2 d^2}^{12} & \beta_{d^2 d^2}^{22} & \cdots & \beta_{d^2 d^2}^{d^2 d^2} \end{pmatrix} \quad (19)$$

so that we have

$$B\text{vec}(X) = \text{vec}(\Lambda). \quad (20)$$

Here, B is determined once the bases $\{E_i\}$ and $\{\rho_m\}$ are chosen, and Λ is obtained from experimental data. B , X and Λ are in general complex matrices. Note that X should be Hermitian and positive semidefinite and satisfy the constraint (15). Hence, direct inversion or pseudo-inversion of B may fail to generate a physical solution. We try to find a physical estimate \hat{X} which will generate an output $\hat{\rho}'_{ge}$ as close as possible to the estimated results $\hat{\rho}'$ from quantum state tomography. Because $\hat{\rho}'_{ge}$ and $\hat{\rho}'$ are characterized by $\hat{\Lambda}_{ge}$ and $\hat{\Lambda}$ separately, we should minimize $\|\hat{\Lambda}_{ge} - \hat{\Lambda}\|$. Since

$$\|\hat{\Lambda}_{ge} - \hat{\Lambda}\| = \|\text{vec}(\hat{\Lambda}_{ge}) - \text{vec}(\hat{\Lambda})\| = \|B\text{vec}(\hat{X}) - \text{vec}(\hat{\Lambda})\|,$$

we will take $\|B\text{vec}(\hat{X}) - \text{vec}(\hat{\Lambda})\|$ as a performance index.

The problem is now the following optimization problem:

Problem 1: Given the matrix B and experimental data $\hat{\Lambda}$, find a Hermitian and positive semidefinite estimate \hat{X} minimizing $\|B\text{vec}(\hat{X}) - \text{vec}(\hat{\Lambda})\|$, such that (15) is satisfied.

It is difficult to obtain an analytical solution to Problem 1. In this paper, we do not directly solve Problem 1 since the problem of QHI can be further specified based on Problem 1. After one obtains an estimate \hat{X} , it is straightforward to obtain Kraus operators $\{\hat{A}_i\}$. Since \hat{X} is Hermitian, it has spectral decomposition

$$\hat{X} = \sum_{i=1}^{d^2} u_i |v_i\rangle\langle v_i|,$$

where u_i are real eigenvalues. Then

$$\hat{C}^T = \sum_{i=1}^{d^2} \sqrt{u_i} |v_i\rangle\langle v_i|,$$

and

$$\hat{A}_i = \sum_j c_{ij} E_j.$$

Though X and \mathcal{E} are in one-to-one correspondence, the notable property of the Kraus operator-sum representation is its non-uniqueness; i.e., there may be more than one different sets of Kraus operators that give rise to the same process \mathcal{E} . This comes from the procedure of decomposing X into $C^T C^*$, which is in fact non-unique because

$$X = C^T C^* = (C^T U^T)(U^* C^*)$$

holds for any unitary U . Hence, the deduction of C from X is non-unique.

III. PROBLEM FORMULATION OF HAMILTONIAN IDENTIFICATION

The objective of this paper is to develop a new algorithm to identify a time-independent Hamiltonian H . If we compare (10) with the Kraus representation (11), it is clear that the unitary propagator $U(t)$ is the only Kraus operator. Then from (13) we know that the matrix C is a row vector. Hence, from (14) we know X is of rank one. It is worth mentioning that, for any given process \mathcal{E} , although the Kraus operator-sum representation is not unique, the process matrix X is in fact uniquely determined. Although there might be other Kraus operator-sum representations where the number of operators is more than 1, the conclusion that X is of rank one is always true. When X is of rank one, the semidefinite requirement is naturally satisfied. Let $X = gg^\dagger$ and $g = \text{vec}(G)$.

Now we need to determine basis sets $\{E_i\}$ and $\{\rho_m\}$. Proper choice of these basis sets can greatly simplify the QHI problem, and we thus choose both of them as the natural basis $\{|j\rangle\langle k|\}_{1 \leq j, k \leq d}$, because the natural basis can simplify the completeness requirement (15) and Problem 1. These advantages can be demonstrated as follows.

Proposition 1: If $\{E_i\}$ is chosen as the natural basis and the relationship between i , j and k is $i = (j-1)d + k$, then the completeness constraint reads $\text{Tr}_1 X = I_d$.

The proof of Proposition 1 is presented in Appendix A.

The natural basis is also useful in transforming Problem 1 into an optimization problem in a more convenient form:

Problem 2: Given the matrix B and experimental data $\hat{\Lambda}$, find a Hermitian and positive semidefinite estimate \hat{X} minimizing $\|\hat{X} - \text{vec}^{-1}(B^{-1}\text{vec}(\hat{\Lambda}))\|$, such that constraint (15) is satisfied.

Problem 2 is not necessarily equivalent to Problem 1. We need to determine when B is invertible and when these two problems are equivalent. To answer these two questions, we give the following conditions to characterize B .

Theorem 1: Let $\{E_i\}_{i=1}^{d^2}$ be a set of matrices in the space $\mathbb{C}_{d \times d}$ and let $\{\rho_m\}_{m=1}^{d^2}$ be a set of linearly independent bases of $\mathbb{C}_{d \times d}$. Define B through (17) and (19). Then $\{E_i\}$ is a set of linearly independent bases of $\mathbb{C}_{d \times d}$ if and only if B is invertible.

Theorem 2: Let $\{E_i\}_{i=1}^{d^2}$ be a set of matrices in $\mathbb{C}_{d \times d}$ and let $\{\rho_m\}_{m=1}^{d^2}$ be a set of normal orthogonal bases of $\mathbb{C}_{d \times d}$. Define B through (17) and (19). Then $\{E_i\}$ forms normal orthogonal basis of $\mathbb{C}_{d \times d}$ if and only if B is unitary.

The detailed proofs of Theorem 1 and Theorem 2 are presented in Appendix B and Appendix C, respectively. Under the conditions in Theorem 2, B is unitary, and we have

$$\begin{aligned} \|B\text{vec}(\hat{X}) - \text{vec}(\hat{\Lambda})\| &= \|\text{vec}(\hat{X}) - B^{-1}\text{vec}(\hat{\Lambda})\| \\ &= \|\hat{X} - \text{vec}^{-1}(B^{-1}\text{vec}(\hat{\Lambda}))\|, \end{aligned}$$

which means Problem 1 is equivalent to Problem 2 in this case. The natural basis set satisfies the requirements in Theorem 1 and Theorem 2.

With the natural basis $\{|j\rangle\langle k|\}_{1 \leq j, k \leq d}$ for $\{E_i\}$ and $\{\rho_m\}$, we have

$$\text{Tr}_1(\text{vec}(G)\text{vec}(G)^\dagger) = I_d = GG^\dagger,$$

which means the completeness constraint (12) is equivalent to the requirement that G is unitary. Hence, we can transform Problem 2 into the following problem which is critical for QHI.

Problem 3: Assume that $\{\rho_m\}_{m=1}^{d^2}$ is a set of normal orthogonal bases of the space $\mathbb{C}_{d \times d}$, $\{E_i\}$ is chosen as $\{|j\rangle\langle k|\}_{1 \leq j, k \leq d}$, and the relationship between i, j and k is $i = (j-1)d + k$. Given the unitary matrix B and experimental data $\hat{\Lambda}$, find a unitary matrix \hat{G} minimizing $\|\text{vec}(\hat{G})\text{vec}(\hat{G})^\dagger - \text{vec}^{-1}(B^\dagger \text{vec}(\hat{\Lambda}))\|$.

Remark 1: Note that we can experimentally measure only Hermitian physical variables. Hence, we cannot directly use $|j\rangle\langle k|$ ($j \neq k$) as probe states. According to [1], when $j \neq k$, one can take $|j\rangle\langle j|$, $|k\rangle\langle k|$, $|+\rangle\langle +|$ and $|-\rangle\langle -|$ as inputs where $|+\rangle = (|j\rangle + |k\rangle)/\sqrt{2}$ and $|-\rangle = (|j\rangle + i|k\rangle)/\sqrt{2}$. Then $\mathcal{E}(|j\rangle\langle k|)$ can be obtained from

$$\begin{aligned} \mathcal{E}(|j\rangle\langle k|) = & \mathcal{E}(|+\rangle\langle +|) + i\mathcal{E}(|-\rangle\langle -|) \\ & - \frac{1+i}{2}\mathcal{E}(|j\rangle\langle j|) - \frac{1-i}{2}\mathcal{E}(|k\rangle\langle k|). \end{aligned} \quad (21)$$

IV. HAMILTONIAN IDENTIFICATION ALGORITHM AND COMPUTATIONAL COMPLEXITY

A. Solution to Problem 3: Two-step Optimization (TSO)

The direct solution to Problem 3 is difficult [39] and we split it into two sub-problems (which is the reason we name our method Two-Step Optimization):

Problem 3.1: Let $\hat{D} = \text{vec}^{-1}(B^\dagger \text{vec}(\hat{\Lambda}))$ be a given matrix. Find a $d \times d$ matrix \hat{S} minimizing $\|\text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger - \hat{D}\|$.

Problem 3.2: Let \hat{S} be given. Find a $d \times d$ unitary matrix \hat{G} minimizing $\|\text{vec}(\hat{G})\text{vec}(\hat{G})^\dagger - \text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger\|$.

1) : For Problem 3.1, let

$$\begin{aligned} L_1 = & \|\text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger - \hat{D}\|^2 \\ = & \text{Tr}\{[\text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger - \hat{D}][\text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger - \hat{D}]\} \\ = & [\text{vec}(\hat{S})^\dagger \text{vec}(\hat{S})]^2 - \text{vec}(\hat{S})^\dagger (\hat{D} + \hat{D}^\dagger) \text{vec}(\hat{S}) + \text{Tr}(\hat{D}\hat{D}^\dagger). \end{aligned}$$

Then by partial differentiation we obtain the conjugate gradient matrix

$$\frac{\partial L_1}{\partial \text{vec}(\hat{S})^*} = 2\text{vec}(\hat{S})^\dagger \text{vec}(\hat{S}) \text{vec}(\hat{S}) - (\hat{D}^\dagger + \hat{D}) \text{vec}(\hat{S}), \quad (22)$$

which leads to

$$(\hat{D}^\dagger + \hat{D}) \text{vec}(\hat{S}) = 2\text{vec}(\hat{S})^\dagger \text{vec}(\hat{S}) \text{vec}(\hat{S}). \quad (23)$$

Therefore the optimal $\text{vec}(\hat{S})$ must be an eigenvector of $(\hat{D}^\dagger + \hat{D})$ corresponding to the positive eigenvalue $2\text{vec}(\hat{S})^\dagger \text{vec}(\hat{S})$. Then

$$\begin{aligned} L_1 = & [\text{vec}(\hat{S})^\dagger \text{vec}(\hat{S})]^2 - \text{vec}(\hat{S})^\dagger (\hat{D} + \hat{D}^\dagger) \text{vec}(\hat{S}) + \text{Tr}(\hat{D}\hat{D}^\dagger) \\ = & \text{Tr}(\hat{D}\hat{D}^\dagger) - [2\text{vec}(\hat{S})^\dagger \text{vec}(\hat{S})]^2/4. \end{aligned}$$

Since $\hat{D}^\dagger + \hat{D}$ is Hermitian, we have the spectral decomposition

$$\hat{D}^\dagger + \hat{D} = \sum_{i=1}^{d^2} \hat{\alpha}_i \text{vec}(\hat{P}_i) \text{vec}(\hat{P}_i)^\dagger, \quad (24)$$

where $\hat{P}_i \in \mathbb{C}_{d \times d}$ and $\hat{\alpha}_1 \geq \dots \geq \hat{\alpha}_{d^2}$. To minimize L_1 , we should choose $2\text{vec}(\hat{S})^\dagger \text{vec}(\hat{S}) = \hat{\alpha}_1$ and $\hat{S} = \sqrt{\frac{\hat{\alpha}_1}{2}} \hat{P}_1$.

2) : For Problem 3.2, note that

$$\begin{aligned} & \|\text{vec}(\hat{G})\text{vec}(\hat{G})^\dagger - \text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger\|^2 \\ = & \text{Tr}\{[\text{vec}(\hat{G})\text{vec}(\hat{G})^\dagger - \text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger]^2\} \\ = & [\text{vec}(\hat{G})^\dagger \text{vec}(\hat{G})]^2 + [\text{vec}(\hat{S})^\dagger \text{vec}(\hat{S})]^2 \\ & - 2\text{vec}(\hat{G})^\dagger \text{vec}(\hat{S}) \text{vec}(\hat{S})^\dagger \text{vec}(\hat{G}) \\ = & d^2 + [\text{Tr}(\hat{S}^\dagger \hat{S})]^2 - 2|\text{Tr}(\hat{G}^\dagger \hat{S})|^2. \end{aligned}$$

Hence, Problem 3.2 is equivalent to maximizing $L_2 = |\text{Tr}(\hat{G}^\dagger \hat{S})|^2$ among all unitary \hat{G} . We make a polar decomposition [44] of \hat{S} to obtain $\hat{S} = \hat{V} \hat{Q}$, where $\hat{V} = \hat{S}(\hat{S}^\dagger \hat{S})^{-\frac{1}{2}}$ is unitary and $\hat{Q} = (\hat{S}^\dagger \hat{S})^{\frac{1}{2}}$ is positive semidefinite. We make a spectral decomposition on \hat{Q} to obtain $\hat{Q} = \hat{Z} \hat{R} \hat{Z}^\dagger$, where \hat{Z} is unitary and $\hat{R} = \text{diag}(\hat{R}_{11}, \hat{R}_{22}, \dots, \hat{R}_{dd})$ with $\hat{R}_{jj} \geq 0$. Without loss of generality, we assume $\hat{R}_{jj} > 0$ for all $1 \leq j \leq d$. Let $\hat{F} = \hat{Z}^\dagger \hat{G}^\dagger \hat{V} \hat{Z}$, and assume that $\hat{F}_{jj} = \hat{r}_j e^{i\hat{\psi}_j}$ with $\hat{r}_j \geq 0$ and $0 \leq \hat{\psi}_j < 2\pi$. Because \hat{F} is unitary, we must have $\hat{r}_j \leq 1$. Hence, we have

$$\begin{aligned} L_2 = & |\text{Tr}(\hat{G}^\dagger \hat{V} \hat{Q})|^2 \\ = & |\text{Tr}(\hat{F} \hat{R})|^2 \\ = & |\sum_j \hat{R}_{jj} \hat{r}_j e^{i\hat{\psi}_j}|^2 \\ = & (\sum_j \hat{r}_j \hat{R}_{jj} \cos \hat{\psi}_j)^2 + (\sum_j \hat{r}_j \hat{R}_{jj} \sin \hat{\psi}_j)^2. \end{aligned} \quad (25)$$

Then we let $\frac{\partial L_2}{\partial \hat{\psi}_j} = 0$ for all j and we obtain

$$\frac{\sum_j \hat{r}_j \hat{R}_{jj} \sin \hat{\psi}_j}{\sum_j \hat{r}_j \hat{R}_{jj} \cos \hat{\psi}_j} = \tan \hat{\psi}_1 = \tan \hat{\psi}_2 = \dots = \tan \hat{\psi}_d.$$

Note that $L_2(\hat{F}) = L_2(e^{i\hat{\psi}_0} \hat{F})$ for any $\hat{\psi}_0 \in \mathbb{R}$. Hence, we can choose $\hat{\psi}_1 = 0$, which means $\hat{\psi}_j = 0$ or π for $2 \leq j \leq d$. To maximize L_2 , we should let all $\hat{\psi}_j$ equal to 0. Therefore, $L_2 = (\sum_j \hat{r}_j \hat{R}_{jj})^2$, which indicates $\hat{r}_j = 1$ for all j . If all the diagonal elements of a unitary matrix are equal to one, then it must be the identity matrix. Hence, for the optimal value we have $\hat{F} = I$. Considering an extra global phase, we finally have the optimal solution

$$\hat{G} = e^{i\hat{\psi}} \hat{V} = e^{i\hat{\psi}} \hat{S} (\hat{S}^\dagger \hat{S})^{-\frac{1}{2}},$$

where $\hat{\psi} \in \mathbb{R}$. Combining the results of Problem 3.1 and Problem 3.2, we obtain the final solution.

After we solve Problem 3, we should calculate the Kraus operator \hat{A} (which is also the unitary propagator $\hat{U}(t)$) from \hat{G} , and finally we calculate \hat{H} from \hat{A} . Note that $\hat{U}(t)$ must be a unitary matrix. Then the questions arise of how to calculate \hat{A} from \hat{G} , and whether the matrix \hat{A} calculated from \hat{G} is always unitary? We answer these questions as follows.

Proposition 2: Under the assumptions of Problem 3, suppose we have obtained a solution

$$\hat{X} = \text{vec}(\hat{G}) \text{vec}(\hat{G})^\dagger.$$

Then there is essentially only one Kraus operator \hat{A} calculated from \hat{G} . \hat{A} must be unitary and in fact \hat{A} is equal to $e^{i\phi} \hat{G}^T$, where $\phi \in \mathbb{R}$.

Proof: Denote $\text{vec}(\hat{G})_j$ as the j -th element of $\text{vec}(\hat{G})$. Since

$$\hat{X} = \text{vec}(\hat{G}) \text{vec}(\hat{G})^\dagger,$$

then

$$\begin{aligned}
\hat{\mathcal{E}}(\rho) &= \sum_{j,k=1}^{d^2} E_j \rho E_k^\dagger \hat{x}_{jk} \\
&= \sum_{j,k=1}^{d^2} E_j \rho E_k^\dagger [\text{vec}(\hat{G})_j \text{vec}(\hat{G})_k^*] \\
&= \sum_{j=1}^{d^2} E_j \text{vec}(\hat{G})_j \rho \sum_{k=1}^{d^2} E_k^\dagger \text{vec}(\hat{G})_k^* \\
&= \sum_{m,n=1}^d |m\rangle\langle n| \text{vec}(\hat{G})_{(m-1)d+n} \rho \\
&\quad \times \sum_{s,t=1}^d |t\rangle\langle s| \text{vec}(\hat{G})_{(s-1)d+t}^* \\
&= \sum_{m,n=1}^d |m\rangle\langle n| \hat{G}_{nm} \rho \sum_{s,t=1}^d |t\rangle\langle s| \hat{G}_{ts}^* \\
&= \hat{G}^T \rho \hat{G}^* \\
&= e^{i\phi} \hat{G}^T \rho e^{-i\phi} \hat{G}^*.
\end{aligned} \tag{26}$$

Therefore, there is essentially only one Kraus operator, which is $e^{i\phi} \hat{G}^T$ with $\phi \in \mathbb{R}$ undetermined, and $\hat{A} = e^{i\phi} \hat{G}^T$ is unitary. \blacksquare

Remark 2: If \hat{S} and \hat{G} are the solutions to Problem 3.1 and Problem 3.2, respectively, then for any $\phi_1, \phi_2 \in \mathbb{R}$, $e^{i\phi_1} \hat{S}$ and $e^{i\phi_2} \hat{G}$ are also optimal solutions, respectively. Hence, there is in fact an undetermined global phase in \hat{G} , which can also be seen from Proposition 2. This stems from the global phase in the Hamiltonian, which is physically unobservable. Through proper prior knowledge, this global phase can be eliminated. For example, in [27] the prior knowledge of $\text{Tr}H = 0$ is assumed. In our simulations of Section VI, we use the assumption that the smallest eigenvalue of H is set to a determined value.

After obtaining \hat{A} , we need to solve $\hat{A} = e^{-i\hat{H}t}$ to obtain \hat{H} . Note that in real physical systems we always require \hat{H} to be Hermitian. Another question which naturally arises is whether every solution \hat{H} of the equation $\hat{A} = e^{-i\hat{H}t}$ is Hermitian? We introduce Theorem 1.43 from [40] as well as its proof, since the proof provides a method to obtain \hat{H} .

Lemma 1 ([40]): $A \in \mathbb{C}_{n \times n}$ is unitary if and only if $A = e^{iH}$ for some Hermitian H . In this representation H can be taken to be Hermitian positive definite.

Proof: The Schur decomposition of A has the form $A = QDQ^\dagger$ with Q unitary and

$$D = \text{diag}(\exp(i\theta_j)) = \exp(i\Theta),$$

where $\Theta = \text{diag}(\theta_j) \in \mathbb{R}_{n \times n}$. Hence,

$$A = Q \exp(i\Theta) Q^\dagger = \exp(iQ\Theta Q^\dagger) = \exp(iH),$$

where $H = H^\dagger$. Without loss of generality we can take $\theta_j > 0$, which implies that H is positive definite. \blacksquare

Lemma 1 satisfies our needs perfectly. Instead of using the general matrix logarithm function, we can just use the Schur decomposition to obtain the logarithm of unitary matrix \hat{A} . Furthermore, from the proof of Lemma 1 we notice that all θ_j should lie in a region no larger than π , otherwise they can not be uniquely determined. This indicates that the sampling period should be small enough. This can also be viewed as a result of Nyquist sampling theorem, as stated in [27]. Hence, in this paper we employ the following assumption.

Assumption 1: The evolution time t satisfies

$$0 < t < \frac{\pi}{h_d - h_1}, \tag{27}$$

where h_d and h_1 are the largest and smallest eigenvalues of Hamiltonian H , respectively.

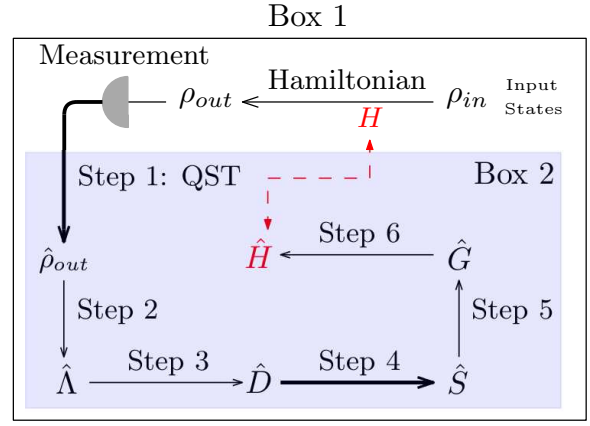


Fig. 1. General procedure of the quantum Hamiltonian identification method, where QST indicates quantum state tomography and this paper focuses on Box 2.

In Appendix D we give an example of a sufficient condition for Assumption 1, which might be more convenient to determine t in practice. Now with Assumption 1 satisfied and h_1 set, we design an algorithm to recover the Hamiltonian from a unitary \hat{G} as the following.

Algorithm 1: (i) Perform a Schur decomposition of \hat{G}^T to get $\hat{G}^T = \hat{Q}\hat{J}\hat{Q}^\dagger$ with \hat{Q} unitary, and $\hat{J} = \exp(i\hat{\Theta})$, where $\hat{\Theta} = \text{diag}(\hat{\theta}_j)$, $0 \leq \hat{\theta}_1 \leq \hat{\theta}_2 \leq \dots \leq \hat{\theta}_d < 2\pi$.

(ii) If $\hat{\theta}_d - \hat{\theta}_1 < \pi$, go to step (iii); otherwise, find the smallest k so that $\hat{\theta}_k - \hat{\theta}_1 \geq \pi$. Then for $j = k, k+1, \dots, d$, replace $\hat{\theta}_j$ with $\hat{\theta}_j - 2\pi$. This step aims to ensure the reconstructed Hamiltonian has spectral region no larger than $h_d - h_1$.

(iii) Let $\hat{\theta}_0 = \max_j \hat{\theta}_j$. For all $1 \leq j \leq d$, take $\hat{\theta}_j = \hat{\theta}_j - h_1 t - \hat{\theta}_0$. If we denote $\hat{\Theta} = \text{diag}(\hat{\theta}_j)$, then $\hat{H} = -\hat{Q}\hat{\Theta}\hat{Q}^\dagger/t$ is the final estimated Hamiltonian.

B. General Procedure and Computational Complexity

In Fig. 1, we summarize the general procedure of the QHI framework. All steps in Box 2 are data processing steps performed on a computer. Step 1 is quantum state tomography, which includes the acquisition of experimental data and post-processing of the experimental data. In this paper we do not consider the time spent on experiments, since it depends on the experimental realization. In the following, we briefly summarize each step and illustrate their corresponding computational complexity.

Step 1. Choose basis sets $\{E_i\}$ and $\{\rho_m\}$ and calculate B . Then use quantum state tomography to reconstruct experimental output states of the system. The number of resource copies N in state tomography determines the estimation error, but does not affect the computational complexity of the estimation algorithm. Generally the calculation of B according to (17) has $O(d^{11})$ computational complexity. However, under the natural basis, this complexity can be reduced to only $O(d^4)$. For state reconstruction, we employ the method of quantum state tomography using LRE for our numerical simulations. The computational complexity of LRE state tomography is $O(d^6)$ offline and $O(d^4)$ online [29]. Considering there are d^2 output states to be reconstructed, the total computational complexity of our LRE method for QHI is $O(d^6)$.

Step 2. Use (16) to determine $\hat{\Lambda}$. Generally the computational complexity to solve (16) is $O(d^{12})$. But it is only $O(d^2)$ using the orthogonal property under the natural basis.

Step 3. Calculate $\hat{D} = \text{vec}^{-1}(B^\dagger \text{vec}(\hat{\Lambda}))$. Generally the complexity is $O(d^8)$. But under the natural basis, we already know the specific structure and value of B (see (29)). Thus, the complexity now is only $O(d^4)$.

Step 4. Calculate \hat{S} according to the spectral decomposition of $\hat{D} + \hat{D}^\dagger$. The computational complexity is determined by spectral decomposition, which is $O(d^6)$ (the computational complexity of spectral decomposition is cubic in a Hermitian matrix's dimension, see [41]).

Step 5. Use matrix polar decomposition to obtain $\hat{G} = \hat{S}(\hat{S}^\dagger \hat{S})^{-\frac{1}{2}}$. The computational complexity is $O(d^3)$ [41].

Step 6. Use the Schur decomposition to obtain the final estimated Hamiltonian \hat{H} from \hat{G} . The computational complexity of Schur decomposition is $O(d^3)$ [41], [42].

Our Hamiltonian identification procedure has the following advantages. Firstly, the framework is general, since we formulate it within the QPT framework. We do not impose any restriction (such as sparseness) on the Hamiltonian. Secondly, Step 1 has the potential for parallel processing. One can deal with data on hand to reconstruct existing output states while at the same time inputting new probe states to the process and making measurements on them. Thirdly, the computational complexity can be analyzed. Regardless of the time spent in experiments, all steps in our QHI framework have clear computational complexity (at most $O(d^6)$). Finally, it is possible to analytically investigate an error upper bound and a detailed error analysis is presented in Section V.

C. Practical Consideration of Storage Requirements

One issue in the calculations is that the dimension of B may increase rapidly. When there are 4 qubits, B has 2^{32} elements. If it takes one byte to store one element of B , then we need 4GB of storage space, which is already a very heavy task for a common PC. We notice that B generated from the natural basis is a permutation matrix. This is vital to computation efficiency. A permutation matrix is a (square in this paper) matrix such that all elements are 0 except exactly one 1 in each column and each row.

Notice that after B is determined from equations (17) and (19), its real usage is in Problem 3, where we need to multiply B^\dagger to a vector. This multiplication task can be done in an alternative way where B 's full storage is avoided. To be specific, we aim to make B sparse. Hence, we only need to store the information of its very small number of nonzero elements and thus ignore a large number of zero elements, while still being able to perform the multiplication. This idea is realized by the following theorem:

Theorem 3: Let $\{E_i\}_{i=1}^{d^2}$ be a set of matrices in $\mathbb{C}_{d \times d}$. Choose $\{\rho_m\}_{m=1}^{d^2} = \{|j\rangle\langle k|\}_{1 \leq j, k \leq d}$. Define B through (17) and (19). Then $\{E_i\} = e^{i\theta} \{\rho_m\}$ if and only if B is a permutation matrix. Here, $\theta \in \mathbb{R}$ is any fixed global phase.

Proof: Using Theorem 2 we know that equation (52) holds.

Sufficiency: Define $W(j, k)$ as a $d^2 \times d^2$ matrix where $W(j, k)$'s element in position (m, n) is the number β_{nm}^{jk} , and denote $(x, y) = (x-1)K + y$ for $1 \leq x, y \leq K$. Using equation (4), we consider each element of $W(j, k)$,

$$\begin{aligned} W(j, k)_{(p, q)(s, t)} &= \text{vec}(\rho_{(p, q)})^\dagger (e^{-i\theta} |g\rangle^* \langle h|^* \otimes e^{i\theta} |m\rangle \langle n|) \text{vec}(\rho_{(s, t)}) \\ &= (\langle q|^* \otimes \langle p|) (|g\rangle^* \otimes |m\rangle) (\langle h|^* \otimes \langle n|) (|t\rangle^* \otimes |s\rangle) \\ &= (\langle q|g\rangle^* \otimes \langle p|m\rangle) (\langle h|t\rangle^* \otimes \langle n|s\rangle) \\ &= \delta_{qg} \delta_{pm} \delta_{th} \delta_{sn}. \end{aligned} \quad (28)$$

Hence, each matrix $W(j, k)$ has exactly one 1 and all other elements are 0. From equation (53) we know each row of B has exactly one 1 and all other elements are 0. When indices j and k run from 1 to d^2 , the index combination (g, h, m, n) never repeats, therefore, $W(j_1, k_1)$ and $W(j_2, k_2)$ have different positions of 1 as long as index pair $(j_1, k_1) \neq (j_2, k_2)$. This means each row of B has no more than one 1. Since B is square, we know that each row of B has exactly one 1. Hence, B is a permutation matrix.

Necessity: When B is a permutation matrix, from equation (53) we know that each matrix $W(j, k)$ has exactly one 1 and all other elements are 0. According to V 's permutation property and equation (54) we know this property for each $W(j, k)$ also holds for each matrix $E_k^* \otimes E_j$. This means that each matrix E_j has exactly one nonzero element, denoted as x_j . Then we have $x_k^* x_j = 1$ holds for every $k, j = 1, 2, \dots, d^2$. Let $j = k$, and we find $x_j = e^{i\theta_j}$. Then we know $\theta_1 = \theta_2 = \dots = \theta_{d^2} = \theta$, where θ is any fixed real number. Since B is invertible, from Theorem 1 we know $\{E_j\}$ is a linearly independent set. Thus each pair of matrices in $\{E_j\}$ have different positions of $e^{i\theta}$. Hence, we can write $\{E_j\} = e^{i\theta} \{\rho_m\}$. ■

From the proof of Theorem 3, one can also deduce an equation to directly calculate B . Substituting this into equation (28), we obtain

$$\beta_{(s, t)(p, q)}^{jk} = \delta_{qg} \delta_{pm} \delta_{th} \delta_{sn} = \beta_{(s, t)(p, q)}^{(m, n)(g, h)}. \quad (29)$$

Therefore, one can easily write down B when the size d is given.

A special case of the sufficiency of Theorem 3; i.e., when $\{E_i\}$ and $\{\rho_m\}$ are the same natural basis sets with the same order of elements, also appeared in [43]. Our theorem and proof here is more general. Using this theorem, we only need to store all 1's positions in B , which only requires d^4 storage space. This is a great reduction compared with d^8 , and the cost is only some more coding in calculating multiplication by B . Furthermore, the computational complexity in writing down B is also reduced to only $O(d^4)$.

V. ERROR ANALYSIS

The error in the Hamiltonian identification method under consideration has only three possible sources. The first one occurs in state estimation, where measurement frequency in practical simulations or experiments is used to approximate the measurement probability. The second one is that state reconstruction algorithm might produce errors. The third one is that our TSO QHI algorithm may also produce errors. In this

section, we give an error upper bound. We first fix the given evolution time t and analyze the error of our QHI method. Then we utilize the similar method to analyze the relationship between the error and the time t .

A. Upper Error Bound for Fixed Evolution Time

Theorem 4: If $\{E_i\}$ and $\{\rho_m\}$ are chosen as natural basis of $\mathbb{C}_{d \times d}$ and the evolution time t is fixed and satisfies Assumption 1, then the estimation error of the TSO QHI method $E\|\hat{H} - H\|$ scales as $O(\frac{d^3}{\sqrt{N}})$, where N is the number of resources in state tomography for each output state.

Proof: The proof of this theorem is divided into the following seven parts.

1) *Error in step 1:* The quantum state tomography algorithm used in this paper is from [29], and the upper bound on the state estimation error is given by

$$\sup_{\rho} E \text{Tr}(\hat{\rho} - \rho)^2 = \frac{M}{4N} \text{Tr}(X^T X)^{-1}, \quad (30)$$

where ρ is the true state and $\hat{\rho}$ its estimator, M is the number of measurement bases, N is the number of experiments (i.e., number of copies of ρ) in state tomography, X is a matrix determined by the measurement basis set (for details, see [29]). Henceforth, we denote this error upper bound (i.e., the RHS of (30)) as Δ_{st} . Following the deduction in the Methods section of [29], one can prove $\Delta_{st} \sim O(\frac{d^4}{N})$. In the following, we will label other errors in the form of Δ with a subscript.

When ρ_m is Hermitian, $E\|\hat{\mathcal{E}}(\rho_m) - \mathcal{E}(\rho_m)\|^2 \leq \Delta_{st}$. When ρ_m is not Hermitian, its process output is in fact calculated according to equation (21) rather than directly probed. Hence, we must analyze this situation specifically. Under the choice of $\{\rho_m\}$ as the natural basis, for $j \neq k$,

$$\begin{aligned} & E\|\hat{\mathcal{E}}(|j\rangle\langle k|) - \mathcal{E}(|j\rangle\langle k|)\|^2 \\ &= E\|\left[\hat{\mathcal{E}}(|+\rangle\langle +|) - \mathcal{E}(|+\rangle\langle +|)\right] + i\left[\hat{\mathcal{E}}(|-\rangle\langle -|) - \mathcal{E}(|-\rangle\langle -|)\right] \\ &\quad - \frac{1+i}{2}\left[\hat{\mathcal{E}}(|j\rangle\langle j|) - \mathcal{E}(|j\rangle\langle j|)\right] - \frac{1+i}{2}\left[\hat{\mathcal{E}}(|k\rangle\langle k|) - \mathcal{E}(|k\rangle\langle k|)\right]\|^2 \\ &\leq (1 + |i| + |\frac{1+i}{2}| + |\frac{1+i}{2}|)^2 \Delta_{st} \\ &= (6 + 4\sqrt{2})\Delta_{st}. \end{aligned} \quad (31)$$

2) *Error in step 2:* Now we calculate the error in the experimental data:

$$\begin{aligned} & E\|\hat{\Lambda} - \Lambda\|^2 \\ &= E \sum_m \sum_{n,k} (\hat{\lambda}_{mn}^* - \lambda_{mn}^*)(\hat{\lambda}_{mk} - \lambda_{mk}) \delta_{nk} \\ &= E \sum_m \sum_{n,k} (\hat{\lambda}_{mn}^* - \lambda_{mn}^*)(\hat{\lambda}_{mk} - \lambda_{mk}) \text{Tr}(\rho_n^\dagger \rho_k) \\ &= E \sum_m \text{Tr}[\sum_n (\hat{\lambda}_{mn}^* - \lambda_{mn}^*) \rho_n^\dagger \sum_k (\hat{\lambda}_{mk} - \lambda_{mk}) \rho_k] \\ &= E \sum_m \text{Tr}(\hat{\mathcal{E}}(\rho_m) - \mathcal{E}(\rho_m))^2 \\ &= E[\sum_{j=1}^d \sum_{k=1, k \neq j}^d \|\hat{\mathcal{E}}(|j\rangle\langle k|) - \mathcal{E}(|j\rangle\langle k|)\|^2 \\ &\quad + \sum_{l=1}^d \|\hat{\mathcal{E}}(|l\rangle\langle l|) - \mathcal{E}(|l\rangle\langle l|)\|^2] \\ &\leq (6 + 4\sqrt{2})d(d-1)\Delta_{st} + d\Delta_{st}. \end{aligned} \quad (32)$$

Also, we denote $\Delta_\Lambda = \|\hat{\Lambda} - \Lambda\|$.

3) *Error in step 3:* From Theorem 3, we know B^\dagger is a permutation matrix. Hence, its effect on $\text{vec}(\Lambda)$ is merely a series of interchanging two elements of $\text{vec}(\Lambda)$, and thus $D = \text{vec}^{-1}(B^\dagger \text{vec}(\Lambda))$ is just a reordering of Λ 's elements. For the same reason, $\hat{D} - D$ is just reordering of $\hat{\Lambda} - \Lambda$. Therefore

$$\|\hat{D} - D\| = \left(\sum_{j,k} |\hat{D}_{jk} - D_{jk}|^2\right)^{\frac{1}{2}} = \|\hat{\Lambda} - \Lambda\| = \Delta_\Lambda. \quad (33)$$

4) *Error in step 4:* We present a lemma to be used in this part.

Lemma 2: Let b and c be two complex vectors with the same finite dimension and assume that they are not both zero simultaneously. Then we have

$$\frac{\|bb^\dagger - cc^\dagger\|}{\|b\| + \|c\|} \leq \min_{\theta \in \mathbb{R}} \|e^{i\theta}b - c\| \leq \frac{\sqrt{2}\|bb^\dagger - cc^\dagger\|}{\sqrt{\|b\|^2 + \|c\|^2}}. \quad (34)$$

The detailed proof of Lemma 2 can be found in Appendix E. We first estimate $\|\hat{S}\|$.

$$\|\hat{S}\|^2 = \text{Tr}(\hat{S}^\dagger \hat{S}) = \frac{\hat{\alpha}_1}{2} \text{Tr}(\hat{P}_1^\dagger \hat{P}_1) = \frac{\hat{\alpha}_1}{2}.$$

Remember that $\hat{\alpha}_1$ is the largest eigenvalue of $\hat{D} + \hat{D}^\dagger$. Using Theorem 3,

$$|\hat{\alpha}_1 - 2d| \leq \|(\hat{D} + \hat{D}^\dagger) - 2D\| \leq 2\|\hat{D} - D\| = 2\Delta_\Lambda.$$

We thus have

$$2d - 2\Delta_\Lambda \leq \hat{\alpha}_1 \leq 2d + 2\Delta_\Lambda.$$

Therefore,

$$\sqrt{d - \Delta_\Lambda} \leq \|\hat{S}\| = \sqrt{\hat{\alpha}_1/2} \leq \sqrt{d + \Delta_\Lambda}. \quad (35)$$

We also need to estimate $\|\hat{S} - S\|$. Using Lemma 2, we have

$$\begin{aligned} & \|\hat{S} - S\| = \|\text{vec}(\hat{S}) - \text{vec}(S)\| \\ &\leq \frac{\sqrt{2}\|\text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger - \text{vec}(S)\text{vec}(S)^\dagger\|}{\sqrt{\|\text{vec}(\hat{S})\|^2 + \|\text{vec}(S)\|^2}} \\ &\leq \frac{\sqrt{2}\|\text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger - \text{vec}(S)\text{vec}(S)^\dagger\|}{\sqrt{2d - \Delta_\Lambda}} \\ &= \left[\frac{1}{\sqrt{d}} + o(1)\right] \|\text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger - D\| \\ &\leq \left[\frac{1}{\sqrt{d}} + o(1)\right] [\|\text{vec}(\hat{S})\text{vec}(\hat{S})^\dagger - \hat{D}\| + \|\hat{D} - D\|] \\ &= \left[\frac{1}{\sqrt{d}} + o(1)\right] [\min_{\tilde{S} \in \mathbb{C}_{d \times d}} \|\text{vec}(\tilde{S})\text{vec}(\tilde{S})^\dagger - \hat{D}\| + \|\hat{D} - D\|] \\ &\leq \left[\frac{1}{\sqrt{d}} + o(1)\right] [\|\text{vec}(G)\text{vec}(G)^\dagger - \hat{D}\| + \|\hat{D} - D\|] \\ &= \left[\frac{1}{\sqrt{d}} + o(1)\right] \cdot 2\|\hat{D} - D\| \sim \frac{2}{\sqrt{d}}\Delta_\Lambda. \end{aligned} \quad (36)$$

5) *Error in step 5:* We introduce *Weyl's Perturbation Theorem*, which can be found in [44].

Lemma 3 ([44]): Let A, B be Hermitian matrices with eigenvalues $\lambda_1(A) \geq \dots \geq \lambda_n(A)$ and $\lambda_1(B) \geq \dots \geq \lambda_n(B)$, respectively. Then

$$\max_j |\lambda_j(A) - \lambda_j(B)| \leq \|A - B\|. \quad (37)$$

Remark 3: The original version of Lemma 3 was for the operator norm. However, from [44] we know for any finite-dimension square matrix, its operator norm is not larger than its Frobenius norm. Therefore this theorem also holds for the Frobenius norm, which is our main focus throughout this paper.

For the true value we have $S^\dagger S = G^\dagger G = I$ and $\|S\| = \sqrt{d}$. Denote $\|\hat{S}^\dagger \hat{S} - S^\dagger S\| = \Delta_{S^\dagger S}$. From the spectral decomposition $\hat{S}^\dagger \hat{S} = \hat{U} \hat{E} \hat{U}^\dagger$, where $\hat{E} = \text{diag}(1 + t_1, 1 + t_2, \dots, 1 + t_d)$. Hence, $t_j \in \mathbb{R}$. Then

$$\|\hat{S}^\dagger \hat{S} - S^\dagger S\|^2 = \|\hat{U} \hat{E} \hat{U}^\dagger - I\|^2 = \|\hat{E} - I\|^2 = \sum_j t_j^2 = \Delta_{S^\dagger S}^2.$$

Thus we know

$$\begin{aligned}
\|\hat{G} - \hat{S}\|^2 &= \text{Tr}[(\hat{G}^\dagger - \hat{S}^\dagger)(\hat{G} - \hat{S})] \\
&= d - 2\text{Tr}\sqrt{\hat{S}^\dagger \hat{S}} + \text{Tr}(\hat{S}^\dagger \hat{S}) \\
&= d - 2\sum_j \sqrt{1+t_j} + \sum_j (1+t_j) \\
&= \sum_j (\sqrt{1+t_j} - 1)^2 = \sum_j \frac{t_j^2}{2+2t_j+\sqrt{1+t_j}} \\
&= \sum_j t_j^2 [\frac{1}{4} - \frac{1}{8}t_j + o(t_j)] = \frac{1}{4}\Delta_{S^\dagger S}^2 + o(\Delta_{S^\dagger S}^2). \tag{38}
\end{aligned}$$

For $\Delta_{S^\dagger S}$, using property (2), we have

$$\begin{aligned}
\Delta_{S^\dagger S} &= \|\hat{S}^\dagger \hat{S} - S^\dagger S\| \\
&\leq \|\hat{S}^\dagger \hat{S} - \hat{S}^\dagger S\| + \|\hat{S}^\dagger S - S^\dagger S\| \\
&\leq \|\hat{S}^\dagger\| \cdot \|\hat{S} - S\| + \|S\| \cdot \|\hat{S}^\dagger - S^\dagger\| \\
&= (\|\hat{S}\| + \sqrt{d})\|\hat{S} - S\| \\
&\leq (\sqrt{d+\Delta_\Lambda} + \sqrt{d})\|\hat{S} - S\| \\
&\sim (\sqrt{d+\Delta_\Lambda} + \sqrt{d})\frac{2}{\sqrt{d}}\Delta_\Lambda \sim 4\Delta_\Lambda. \tag{39}
\end{aligned}$$

Combining (38) and (39), we obtain

$$\|\hat{G} - \hat{S}\| \sim \frac{1}{2}\Delta_{S^\dagger S} \leq 2\Delta_\Lambda. \tag{40}$$

From subsection IV-A, we know there is in fact an extra degree of freedom ϕ in the estimated $e^{i\phi}\hat{G}^T$ and it can be eliminated using prior knowledge. Here, we take

$$\|\hat{G} - G\| = \min_\phi \|e^{i\phi}\hat{G} - G\|.$$

Then we have

$$\begin{aligned}
\|\hat{G} - G\| &\leq \|\hat{G} - \hat{S}\| + \|\hat{S} - S\| + \|S - G\| \\
&= \|\hat{G} - \hat{S}\| + \|\hat{S} - S\|. \tag{41}
\end{aligned}$$

Now by substituting equations (36) and (40) into (41), we have

$$\begin{aligned}
\|\hat{G} - G\| &\leq \|\hat{G} - \hat{S}\| + \|\hat{S} - S\| \\
&\leq 2\Delta_\Lambda + \frac{2}{\sqrt{d}}\Delta_\Lambda \sim O(\Delta_\Lambda). \tag{42}
\end{aligned}$$

6) *Error in step 6:* In this part we need the following lemma:

Lemma 4: For $\theta \in [-\pi, \pi]$, $\frac{2}{\pi^2}\theta^2 \leq 1 - \cos\theta$.

Based on differential analysis up to the second-order derivative, the proof of Lemma 4 is straightforward and hence we omit the details.

Suppose the system Hamiltonian has a spectral decomposition $tH = -Q\Theta Q^\dagger$, where $\Theta = \text{diag}(\theta_j)$. Since t satisfies Assumption 1, we have $0 \leq \theta_j \leq \pi$ for every $j = 1, 2, \dots, d$. Let $M = \hat{Q}^\dagger Q$, which is also unitary. Then, we have

$$\begin{aligned}
&t^2\|\hat{H} - H\|^2 \\
&= \|\hat{Q}\hat{\Theta}\hat{Q}^\dagger - Q\Theta Q^\dagger\|^2 = \|\hat{\Theta} - M\Theta M^\dagger\|^2 \\
&= \text{Tr}(\hat{\Theta}^2 + \Theta^2) - 2\text{Tr}(\hat{\Theta}M\Theta M^\dagger) \\
&= \sum_j (\hat{\theta}_j^2 + \theta_j^2) - 2\sum_{j,k} \hat{\theta}_j \theta_k |M_{jk}|^2 \\
&= \sum_{j,k} (\hat{\theta}_j^2 + \theta_k^2) |M_{jk}|^2 - 2\sum_{j,k} \hat{\theta}_j \theta_k |M_{jk}|^2 \\
&= \sum_{j,k} (\hat{\theta}_j - \theta_k)^2 |M_{jk}|^2. \tag{43}
\end{aligned}$$

Now using Lemma 4, we have

$$\begin{aligned}
&\frac{4t^2}{\pi^2}\|\hat{H} - H\|^2 \\
&\leq 2\sum_{j,k} [1 - \cos(\hat{\theta}_j - \theta_k)] |M_{jk}|^2 \\
&= 2\sum_{j,k} |M_{jk}|^2 - 2\sum_{j,k} \cos(\theta_k - \hat{\theta}_j) |M_{jk}|^2 \\
&= 2d - 2\text{Re}(\sum_{j,k} e^{i(\theta_k - \hat{\theta}_j)} |M_{jk}|^2) \\
&= \text{Tr}(I + I) - 2\text{Re}(\sum_j e^{-i\hat{\theta}_j} \sum_k e^{i\theta_k} |M_{jk}|^2) \\
&= \text{Tr}(e^{i\hat{\Theta}} e^{-i\hat{\Theta}} + e^{i\Theta} e^{-i\Theta}) - 2\text{Re}[\text{Tr}(e^{-i\hat{\Theta}} M e^{i\Theta} M^\dagger)] \\
&= \|e^{i\hat{\Theta}} - M e^{i\Theta} M^\dagger\|^2 = \|\hat{Q} e^{i\hat{\Theta}} \hat{Q}^\dagger - Q e^{i\Theta} Q^\dagger\|^2 \\
&= \|\hat{G}^T - G^T\|^2 = \|\hat{G} - G\|^2. \tag{44}
\end{aligned}$$

Hence

$$\|\hat{H} - H\| \sim O(\|\hat{G} - G\|) \tag{45}$$

7) *Total Error:* We combine equations (45), (42), (32) and (30) to obtain

$$\begin{aligned}
E\|\hat{H} - H\|^2 &\sim E[O(\|\hat{G} - G\|^2)] \sim E[O(\Delta_\Lambda^2)] \\
&\sim O(d^2 \Delta_{st}) \sim O(\frac{d^6}{N}), \tag{46}
\end{aligned}$$

which concludes the proof of Theorem 4. \blacksquare

From Theorem 4, we can also obtain the following corollary.

Corollary 1: If $\{E_i\}$ and $\{\rho_m\}$ are chosen as natural basis of $\mathbb{C}_{d \times d}$, and the evolution time t is fixed and satisfies Assumption 1, then the TSO Hamiltonian identification method is asymptotically unbiased.

B. Upper Error Bound vs Evolution Time

Using a similar idea to the above, we can characterize the estimation error for different evolution times t :

Theorem 5: If $\{E_i\}$ and $\{\rho_m\}$ are chosen as a natural basis of $\mathbb{C}_{d \times d}$ and N is fixed, then the estimation error of the TSO Hamiltonian identification method scales as $E\|\hat{H} - H\| \sim O(\frac{1}{t})$ where t satisfies Assumption 1.

The proof of this theorem is similar to the proof of Theorem 4. Note from (44), we have

$$\frac{2t}{\pi}\|\hat{H} - H\| \leq \|\hat{G} - G\|, \tag{47}$$

which combined with (42), (32) and (30) leads to the conclusion in the theorem. It is worth pointing out that in this theorem, the evolution time cannot be arbitrarily large, rather it must be upper bounded according to Assumption 1. Hence, this scaling only holds in a certain region.

VI. NUMERICAL RESULTS

We perform numerical simulations using MATLAB on PC. It is worth mentioning that the selection of natural bases is only a mathematical representation tool in the identification algorithm. When performing measurements on the output states, our framework is applicable to many general measurement bases, such as cube bases [45], MUB bases [46]-[48], SIC-POVMs [49], etc. In our simulations for the TSO method, we choose cube measurement bases. The single-qubit cube measurement set consists of six measurement operators: $\{\frac{I \pm \sigma_x}{2}, \frac{I \pm \sigma_y}{2}, \frac{I \pm \sigma_z}{2}\}$, and the multi-qubit cube measurement set is the tensor product of the single-qubit cube set. After the measurements, we then use the LRE method to reconstruct the output states.

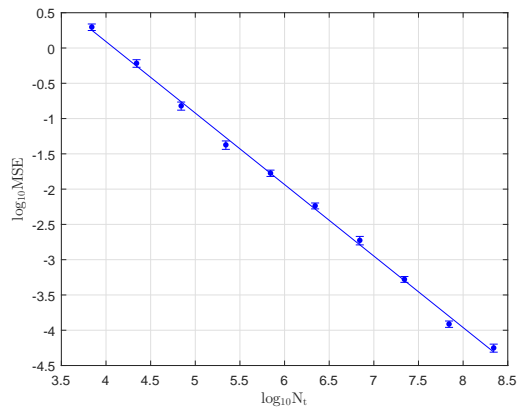


Fig. 2. MSE versus the logarithm of the total resource number N_t .

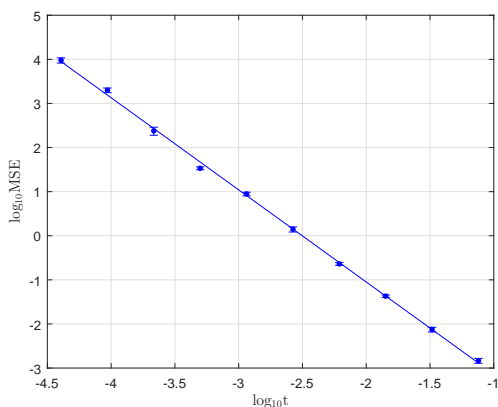


Fig. 3. MSE versus the logarithm of the different evolution times t .

A. Performance Illustration

First we illustrate the relationship between the Mean Squared Error (MSE) and the resource number. Let N_t be the total number of resources; i.e., the total number of copies of different quantum states used as probes. Considering (21), we have $N_t = \frac{3d^2-d}{2}N$. H is the real Hamiltonian, and \hat{H} its estimation through IV-A. In Fig. 2, the vertical axis is $\log_{10} E\text{Tr}(\hat{H} - H)^2$ and the horizontal axis is $\log_{10} N_t$. The real Hamiltonian is taken as

$$H = \begin{pmatrix} 5 & 0.1 & 3i & 4i \\ 0.1 & -1 & 1.8 & 0.9 \\ -3i & 1.8 & 2 & 0.7i \\ -4i & 0.9 & -0.7i & 3 \end{pmatrix}. \quad (48)$$

The distance between its largest and smallest eigenvalues is 11.95. The evolution time $t = 0.1$ and each point is repeated for 10 times. The fitting slope is -1.0131 ± 0.0154 , which matches the theoretical result in Theorem 4.

Now we demonstrate the relationship between the MSE and the evolution time. For the same 2-qubit Hamiltonian in (48), we fix the number of copies in state tomography for each output state as $3^6 \times 1000$ and perform simulations for different evolution times t . The result is in Fig. 3 and each point is repeated 10 times. The fitting slope is -2.0891 ± 0.0215 , which matches the theoretical result in Theorem 5.

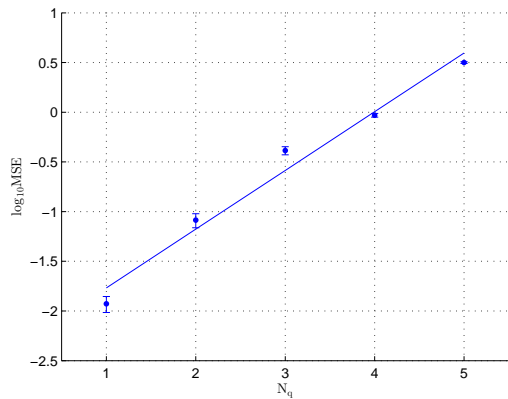


Fig. 4. MSE versus number of qubits N_q .

Moreover, we present an example to illustrate the relationship between the MSE and the qubit number. Let N_q denote the number of qubits; i.e., $d = 2^{N_q}$. We perform simulations when N_q increases from 1 to 5. We set

$$H = \begin{pmatrix} 1 & 0.9 + 0.9i \\ 0.9 - 0.9i & 2 \end{pmatrix}^{\otimes N_q}$$

and $t = 0.01$. For $N_q = 5$, the distance between the largest and smallest eigenvalues of H is 193.87. The number of copies for state tomography is $3^6 \times 1000$ for each output state. The result is in Fig. 4 and each point is repeated 10 times. We observe that as N_q increases, the errorbar decreases. This is because as N_q increases, the error is also increasing. Therefore, the fluctuations gradually become relatively small. We examine various Hamiltonians and obtain similar results. Furthermore, we observe that the upper error bound in Theorem 4 indicates a slope larger than that of the fitted line in Fig. 4. The observation may come from the fact that the bound in Theorem 4 is an upper bound and the experimental result in Fig. 4 does not necessarily reach the bound.

B. Performance Comparison

Now we compare the performance of the TSO QHI method with the QHI approach developed by Zhang and Sarovar in [27], which is based on the eigenstate realization algorithm in classical identification (abbreviated as the ERA method hereafter).

The ERA method can be used to give a general solution to QHI although it was originally presented for the identification of partial parameters in the system Hamiltonian. The ERA method first converts QHI into a system identification problem in the real domain, where the transfer function of the equivalent linear system can be obtained. From temporal records of system observables, it can reconstruct the transfer function. Then equating the coefficients of the transfer functions with unknown parameters to those from the experimental data, the ERA method leads to a set of multivariate polynomial equations, whose solution yields the estimates of the Hamiltonian parameters.

This approach is only efficient if the number of parameters to be identified in the Hamiltonian is small. This is because

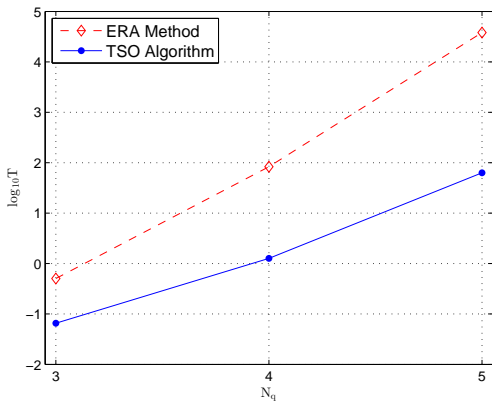


Fig. 5. Running time T versus qubit number N_q for the ERA method in [27] and our TSO method.

solving multivariate polynomial equation takes a considerable amount of time, especially for high dimensional systems or for full Hamiltonian identification with complex quantum systems. In fact, common algorithms solving multivariate polynomial equations can be super-exponential when the number of variables scales up [50].

To illustrate the efficiency of the TSO Hamiltonian identification method, we compare it with the ERA method by numerical simulations, which we performed on a single thread, computer cluster with 2 Intel Xeon E5-2680v3 CPUs and 256 GB memory. We consider the following Hamiltonian for a 1D chain of N_q qubits, which is the example investigated in [27]:

$$H = \sum_{k=1}^{N_q} \frac{\omega_k}{2} \sigma_z^k + \sum_{k=1}^{N_q-1} \delta_k (\sigma_+^k \sigma_-^{k+1} + \sigma_-^k \sigma_+^{k+1}). \quad (49)$$

Here ω_k and δ_k are unknown parameters to be identified. δ_k are the coupling strength between k -th and $(k+1)$ -th spins, $\sigma_+ = \frac{1}{2}(\sigma_x + i\sigma_y)$ and $\sigma_- = \frac{1}{2}(\sigma_x - i\sigma_y)$. Running on the same computer cluster, we compare the consumed time of our TSO QHI method versus the ERA method for the cases of $N_q = 3, 4$, and 5. For the TSO method, we do not utilize the prior structural knowledge (1D-chain) of the targeted Hamiltonian, whereas this information is used in the ERA method. Fig. 5 shows the numerical result, where the vertical axis is the running time T (in units of seconds) in a logarithmic scale, and the horizontal axis is the number of qubits N_q . The red diamonds are the times from the ERA method, whereas the blue dots are for the TSO identification method. The numerical results show that the TSO method is much faster (e.g., around 100 times faster for $N_q = 4$) than the ERA method even if we do not use the prior knowledge of Hamiltonian's structure. It is worth mentioning that the efficiency of the TSO algorithm usually depends on the system size but not the number of parameters for a given system size, while the performance of the ERA method significantly depends on the system size as well as the number of parameters to be identified. The efficiency advantage of TSO algorithm becomes remarkable as the system size and the number of parameters to be identified increase.

VII. CONCLUSION

We have presented a new TSO Hamiltonian identification method and analyzed its computational complexity. This identification method is applicable to general time-independent Hamiltonians for closed quantum systems. We have also provided a theoretical upper bound for the identification error and demonstrate the performance of the identification algorithm using numerical examples. Future work includes the extension of the TSO algorithm to quantum process tomography for open quantum systems and the investigation of whether quantum entanglement can enhance the performance of Hamiltonian identification.

APPENDIX A

PROOF OF PROPOSITION 1

Proof: In this paper, whenever we need to endow orders to number pairs (x, y) ($1 \leq x, y \leq K$) we identify (x, y) with $(x-1)K + y$ unless declared otherwise.

When $\{E_i\}$ is chosen as $\{|j\rangle\langle k|\}_{1 \leq j, k \leq d}$, expand (15) as

$$\begin{aligned} \sum_{s,t,u,v=1}^d x_{(s,t)(u,v)} |v\rangle\langle u|s\rangle\langle t| &= \sum_{s,t,u,v=1}^d \delta_{us} x_{(s,t)(u,v)} |v\rangle\langle t| \\ &= \sum_{s,t,v=1}^d x_{(s,t)(s,v)} |v\rangle\langle t| \\ &= I_d = \sum_{t,v=1}^d \delta_{tv} |v\rangle\langle t|. \end{aligned}$$

Therefore, we must have

$$\sum_{s=1}^d x_{(s,t)(s,v)} = \delta_{tv} = \sum_{s=1}^d x_{((s-1)d+t)((s-1)d+v)}$$

for $t, v = 1, 2, \dots, d$, which is just $\text{Tr}_1 X = I_d$. ■

APPENDIX B

PROOF OF THEOREM 1

Proof: Using equation (5), we vectorize equation (17) to obtain

$$(E_k^* \otimes E_j) \text{vec}(\rho_m) = \sum_n \beta_{nm}^{jk} \text{vec}(\rho_n). \quad (50)$$

Let $\{W(j, k)\}_{j,k=1}^{d^2}$ be a family of d^4 matrices. The matrix $\{W(j, k)\}$ is $d^2 \times d^2$ and its element in position (m, n) is the number β_{nm}^{jk} . Let $V = (\text{vec}(\rho_1), \text{vec}(\rho_2), \dots, \text{vec}(\rho_{d^2}))$. From equation (50) we have

$$(E_k^* \otimes E_j) V = V W(j, k). \quad (51)$$

Since $\{\rho_m\}_{m=1}^{d^2}$ is a set of linearly independent matrices forming a basis of the space $\mathbb{C}_{d \times d}$, V must be invertible. Therefore we know

$$W(j, k) = V^{-1} (E_k^* \otimes E_j) V. \quad (52)$$

Sufficiency: Since $\{E_i\}_{i=1}^{d^2}$ is a set of linearly independent matrices forming a basis of the space $\mathbb{C}_{d \times d}$, we know $\{E_k^* \otimes E_j\}_{j,k=1}^{d^2}$ is a set of linearly independent matrices forming a basis of the space $\mathbb{C}_{d^2 \times d^2}$. Therefore, $\{W(j, k)\}_{j,k=1}^{d^2}$ is also a set of linearly independent matrices forming a basis

of $\mathbb{C}_{d^2 \times d^2}$. We then know $\{\text{vec}(W(j,k)^T)\}_{j,k=1}^{d^2}$ is a set of linearly independent column vectors forming a basis of the space $\mathbb{C}_{d^4 \times 1}$, which leads to the conclusion that

$$B = \begin{bmatrix} \text{vec}(W(1,1)^T), \text{vec}(W(1,2)^T), \dots, \text{vec}(W(2,1)^T), \\ \text{vec}(W(2,2)^T), \dots, \text{vec}(W(d^2, d^2)^T) \end{bmatrix} \quad (53)$$

must be invertible.

Necessity: When B is invertible, from equation (53) we know that $\{\text{vec}(W(j,k)^T)\}_{j,k=1}^{d^2}$ is a set of linearly independent column vectors forming a basis of the space $\mathbb{C}_{d^4 \times 1}$. Therefore $\{W(j,k)\}_{j,k=1}^{d^2}$ is a set of linearly independent basis of $\mathbb{C}_{d^2 \times d^2}$, and from equation (52) $\{E_k^* \otimes E_j\}_{j,k=1}^{d^2}$ is also a set of linearly independent basis of $\mathbb{C}_{d^2 \times d^2}$.

Now suppose that $\{E_i\}_{i=1}^{d^2}$ is not linearly independent. Then from equation (52), one can easily prove $\{E_k^* \otimes E_j\}_{j,k=1}^{d^2}$ is not linearly independent, which leads to a contradiction. Hence, we have proved necessity. ■

APPENDIX C PROOF OF THEOREM 2

Proof: We follow the notations in the Proof of Theorem 1. Since $\{\rho_m\}_{m=1}^{d^2}$ is a set of normal orthogonal basis of space $\mathbb{C}_{d \times d}$, we know V is unitary. Therefore, we know

$$W(j,k) = V^\dagger (E_k^* \otimes E_j) V. \quad (54)$$

Sufficiency: Since $\{E_i\}_{i=1}^{d^2}$ is a set of normal orthogonal basis of the space $\mathbb{C}_{d \times d}$, we have

$$\delta_{(p,q)(k,j)} = \delta_{pk} \delta_{qj} = \langle E_p^*, E_k^* \rangle \langle E_q, E_j \rangle = \langle E_p^* \otimes E_q, E_k^* \otimes E_j \rangle,$$

which means $\{E_k^* \otimes E_j\}_{j,k=1}^{d^2}$ is a set of normal orthogonal basis of the space $\mathbb{C}_{d^2 \times d^2}$. Therefore from (54), we know that $\{W(j,k)\}_{j,k=1}^{d^2}$ is also a set of normal orthogonal basis of the space $\mathbb{C}_{d^2 \times d^2}$. Hence B must be unitary.

Necessity: Since B is unitary, from (53) we know $\{W(j,k)\}_{j,k=1}^{d^2}$ is a set of normal orthogonal basis of $\mathbb{C}_{d^2 \times d^2}$. According to (54), we know $\{E_k^* \otimes E_j\}_{j,k=1}^{d^2}$ is also a set of normal orthogonal basis of $\mathbb{C}_{d^2 \times d^2}$. Hence, we have

$$\langle E_p^* \otimes E_q, E_k^* \otimes E_j \rangle = \delta_{(p,q)(k,j)} = \delta_{pk} \delta_{qj} = \langle E_p^*, E_k^* \rangle \langle E_q, E_j \rangle. \quad (55)$$

Now we concentrate on the third equality in (55). Setting $p = k = q = j$, we obtain $1 = |\langle E_j, E_j \rangle|^2$. Since $\langle E_j, E_j \rangle = \text{Tr}(E_j^\dagger E_j)$ is a positive real number, we must have $\langle E_j, E_j \rangle = 1$ for every $j = 1, 2, \dots, d^2$. Setting $p = k$, we obtain $\delta_{qj} = \langle E_q, E_j \rangle$, which means that $\{E_i\}_{i=1}^{d^2}$ is a set of normal orthogonal basis of the space $\mathbb{C}_{d \times d}$. ■

APPENDIX D EXAMPLES OF ELIMINATING MULTIVALUED SOLUTIONS

Example 1: Let $\|\cdot\|_x$ be any submultiplicative matrix norm (i.e., $\|\cdot\|_x$ satisfies (2)). Suppose we know *a priori* that $\|H\|_x$ is upper bounded by a known value h_m . Then we can set the evolution time $t < \frac{\pi}{2h_m}$.

The proof is straightforward. *Theorem 1* in Chapter 10.3 of [51] states that the absolute value of the eigenvalue of

any matrix is no larger than the submultiplicative norm of the matrix. Hence, the prior knowledge in Example 1 is a sufficient condition for Assumption 1 to be satisfied.

APPENDIX E PROOF OF LEMMA 4

Proof:

$$\begin{aligned} \|e^{i\theta} b - c\|^2 &= (e^{-i\theta} b^\dagger - c^\dagger)(e^{i\theta} b - c) \\ &= b^\dagger b + c^\dagger c - (e^{i\theta} b c^\dagger + e^{-i\theta} b^\dagger c). \end{aligned} \quad (56)$$

Let $b^\dagger c = r e^{i\phi}$, where $r > 0, \phi \in \mathbb{R}$. Then

$$\begin{aligned} \|e^{i\theta} b - c\|^2 &= b^\dagger b + c^\dagger c - (r e^{i\theta} e^{-i\phi} + r e^{-i\theta} e^{i\phi}) \\ &= b^\dagger b + c^\dagger c - 2r \cos(\theta - \phi). \end{aligned} \quad (57)$$

Therefore, we should take $\theta = \phi$ to obtain

$$\min_{\theta} \|e^{i\theta} b - c\|^2 = b^\dagger b + c^\dagger c - 2r. \quad (58)$$

We have

$$\begin{aligned} \|bb^\dagger - cc^\dagger\|^2 &= \text{Tr}(bb^\dagger bb^\dagger + cc^\dagger cc^\dagger - 2bb^\dagger cc^\dagger) \\ &= (b^\dagger b)^2 + (c^\dagger c)^2 - 2r^2. \end{aligned} \quad (59)$$

From the Cauchy-Schwartz inequality,

$$r = |\langle b, c \rangle| \leq \|b\| \cdot \|c\| = \sqrt{b^\dagger b} \sqrt{c^\dagger c}.$$

We thus have

$$\begin{aligned} &\frac{b^\dagger b + c^\dagger c}{2} \min_{\theta} \|e^{i\theta} b - c\|^2 \\ &= \frac{1}{2} (b^\dagger b + c^\dagger c)^2 - r (b^\dagger b + c^\dagger c) \\ &\leq (b^\dagger b)^2 + (c^\dagger c)^2 - r (b^\dagger b + c^\dagger c) \\ &\leq (b^\dagger b)^2 + (c^\dagger c)^2 - 2r \sqrt{b^\dagger b} \sqrt{c^\dagger c} \\ &\leq (b^\dagger b)^2 + (c^\dagger c)^2 - 2r^2 \\ &= \|bb^\dagger - cc^\dagger\|^2. \end{aligned} \quad (60)$$

On the other hand,

$$\begin{aligned} &(\sqrt{b^\dagger b} + \sqrt{c^\dagger c})^2 \min_{\theta} \|e^{i\theta} b - c\|^2 \\ &= (b^\dagger b + c^\dagger c + 2\sqrt{b^\dagger b} \sqrt{c^\dagger c})(b^\dagger b + c^\dagger c - 2r) \\ &= (b^\dagger b + c^\dagger c)^2 - 4r \sqrt{b^\dagger b} \sqrt{c^\dagger c} \\ &\quad + 2(b^\dagger b + c^\dagger c)(\sqrt{b^\dagger b} \sqrt{c^\dagger c} - r) \\ &\geq (b^\dagger b)^2 + (c^\dagger c)^2 + 2b^\dagger b c^\dagger c - 4r \sqrt{b^\dagger b} \sqrt{c^\dagger c} \\ &= (b^\dagger b)^2 + (c^\dagger c)^2 + 2(\sqrt{b^\dagger b} \sqrt{c^\dagger c} - r)^2 - 2r^2 \\ &\geq (b^\dagger b)^2 + (c^\dagger c)^2 - 2r^2 \\ &= \|bb^\dagger - cc^\dagger\|^2. \end{aligned} \quad (61)$$

■

ACKNOWLEDGEMENT

The authors would like to thank Zhibo Hou and Guo-Yong Xiang for helpful discussions.

REFERENCES

- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2000.
- [2] J. Fiurásek and Z. Hradil, "Maximum-likelihood estimation of quantum processes," *Phys. Rev. A*, vol. 63, no. 2, 2001, Art. ID 020101.
- [3] M. F. Sacchi, "Maximum-likelihood reconstruction of completely positive maps," *Phys. Rev. A*, vol. 63, no. 5, 2001, Art. ID 054104.
- [4] 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, "Ancilla-assisted quantum process tomography," *Phys. Rev. Lett.*, vol. 90, no. 19, 2003, Art. ID 193601.
- [5] H. F. Chen and W. Zhao, *Recursive Identification and Parameter Estimation*. Singapore: CRC Press, Taylor & Francis, 2014.
- [6] L. Ljung, *System Identification - Theory for the User*. Upper Saddle River, N.J.: Prentice Hall, 1999.
- [7] E. D. Sontag, Y. Wang, A. Megretski, "Input classes for identifiability of bilinear systems," *IEEE Trans. Autom. Control*, vol. 54, no. 2, pp. 195-207, 2009.
- [8] H. Yonezawa, D. Nakane, T. A. Wheatley, K. Iwasawa, S. Takeda, H. Arao, K. Ohki, K. Tsumura, D. W. Berry, T. C. Ralph, H. M. Wiseman, E. H. Huntington and A. Furusawa, "Quantum-enhanced optical-phase tracking," *Science*, vol. 337, no. 6101, pp. 1514-1517, 2012.
- [9] G.-Y. Xiang, B. L. Higgins, D. W. Berry, H. M. Wiseman and G. J. Pryde, "Entanglement-enhanced measurement of a completely unknown optical phase," *Nat. Photonics*, vol. 5, no. 1, pp. 43-47, 2011.
- [10] D. Burgarth and K. Yuasa, "Quantum system identification," *Phys. Rev. Lett.*, vol. 108, no. 8, 2012, Art. ID 080502.
- [11] M. Guță and N. Yamamoto, "System identification for passive linear quantum systems," *IEEE Trans. Autom. Control*, vol. 61, no. 4, pp.921-936, 2016.
- [12] D. Burgarth, K. Maruyama and F. Nori, "Indirect quantum tomography of quadratic Hamiltonians," *New J. Phys.*, vol. 13, no. 1, 2011, Art. ID 013019.
- [13] C. Di Franco, M. Paternostro and M. S. Kim, "Hamiltonian tomography in an access-limited setting without state initialization," *Phys. Rev. Lett.*, vol. 102, no. 18, 2009, Art. ID 187203.
- [14] C. Ferrie, C. E. Granade and D. G. Cory, "Adaptive Hamiltonian estimation using Bayesian experimental design," *arXiv preprint*, quant-ph, arXiv: 1111.0935, 2011.
- [15] A. Sone and P. Cappellaro, "Hamiltonian identifiability assisted by single-probe measurement," *arXiv preprint*, quant-ph, arXiv: 1609.09446, 2016.
- [16] S. Bonnabel, M. Mirrahimi and P. Rouchon, "Observer-based Hamiltonian identification for quantum systems," *Automatica*, vol. 45, no. 5, pp. 1144-1155, 2009.
- [17] C. Le Bris, M. Mirrahimi, H. Rabitz and G. Turinici, "Hamiltonian identification for quantum systems: well-posedness and numerical approaches," *ESAIM Control Optim. Calc. Var.*, vol. 13, no. 2, pp. 378-395, 2007.
- [18] Z. Leghtas, G. Turinici, H. Rabitz and P. Rouchon, "Hamiltonian identification through enhanced observability utilizing quantum control," *IEEE Trans. Autom. Control*, vol. 57, no. 10, pp. 2679-2683, 2012.
- [19] J. M. Geremia and H. Rabitz, "Optimal identification of Hamiltonian information by closed-loop laser control of quantum systems," *Phys. Rev. Lett.*, vol. 89, no. 26, 2002, Art. ID 263902.
- [20] A. Shabani, M. Mohseni, S. Lloyd, R. L. Kosut and H. Rabitz, "Estimation of many-body quantum Hamiltonians via compressive sensing," *Phys. Rev. A*, vol. 84, no. 1, 2011, Art. ID 012107.
- [21] K. Rudinger and R. Joynet, "Compressed sensing for Hamiltonian reconstruction," *Phys. Rev. A*, vol. 92, no. 5, 2015, Art. ID 052322.
- [22] S. G. Schirmer and D. K. L. Oi, "Two-qubit Hamiltonian tomography by Bayesian analysis of noisy data," *Phys. Rev. A*, vol. 80, no. 2, 2009, Art. ID 022333.
- [23] S. G. Schirmer, A. Kolli and D. K. L. Oi, "Experimental Hamiltonian identification for controlled two-level systems," *Phys. Rev. A*, vol. 69, no. 5, 2004, Art. ID 050306.
- [24] A. Sergeevich, A. Chandran, J. Combes, S. D. Bartlett and H. M. Wiseman, "Characterization of a qubit Hamiltonian using adaptive measurements in a fixed basis," *Phys. Rev. A*, vol. 84, no. 5, 2011, Art. ID 052315.
- [25] S. T. Wang, D. L. Deng and L. M. Duan, "Hamiltonian tomography for quantum many-body systems with arbitrary couplings," *New J. Phys.*, vol. 17, no. 9, 2015, Art. ID 093017.
- [26] J. H. Cole, S. G. Schirmer, A. D. Greentree, C. J. Wellard, D. K. L. Oi and L. C. L. Hollenberg, "Identifying an experimental two-state Hamiltonian to arbitrary accuracy," *Phys. Rev. A*, vol. 71, no. 6, 2005, Art. ID 062312.
- [27] J. Zhang and M. Sarovar, "Quantum Hamiltonian identification from measurement time traces," *Phys. Rev. Lett.*, vol. 113, no. 8, 2014, Art. ID 080401.
- [28] H. Yuan and C. F. Fung, "Optimal feedback scheme and universal time scaling for Hamiltonian parameter estimation," *Phys. Rev. Lett.*, vol. 115, no. 11, 2015, Art. ID 110401.
- [29] B. Qi, Z. Hou, L. Li, D. Dong, G.-Y. Xiang, G.-C. Guo, "Quantum state tomography via linear regression estimation," *Sci. Rep.*, vol. 3, no. 3496, 2013.
- [30] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
- [31] J. Watrous. (2011). *Theory of quantum information* [Online]. Available: <https://cs.uwaterloo.ca/~watrous/LectureNotes.html>
- [32] M. Paris and J. Řeháček, *Quantum State Estimation*, vol. 649 of Lecture Notes in Physics, Springer, Berlin, 2004.
- [33] Z. Hradil, "Quantum-state estimation," *Phys. Rev. A*, vol. 55, no. 3, 1997, Art. ID R1561.
- [34] Y. S. Teo, H. Zhu, B.-G. Englert, J. Řeháček and Z. Hradil, "Quantum-state reconstruction by maximizing likelihood and entropy," *Phys. Rev. Lett.*, vol. 107, no. 2, 2011, Art. ID 020404.
- [35] R. Blume-Kohout, "Optimal, reliable estimation of quantum states," *New J. Phys.*, vol. 12, no. 4, 2010, Art. ID 043034.
- [36] B. Qi, Z. Hou, Y. Wang, D. Dong, H.-S. Zhong, L. Li, G.-Y. Xiang, H. M. Wiseman, C.-F. Li and G.-C. Guo, "Recursively adaptive quantum state tomography: theory and two-qubit experiment," *npj Quantum Information*, 2016, in press, online: *arXiv preprint*, quant-ph, arXiv: 1512.01634, 2015.
- [37] Z. Hou, H.-S. Zhong, Y. Tian, D. Dong, B. Qi, L. Li, Y. Wang, F. Nori, G.-Y. Xiang, C.-F. Li and G.-C. Guo, "Full reconstruction of a 14-qubit state within four hours," *New J. Phys.*, vol. 18, 2016, Art. ID 083036.
- [38] J. L. O'Brien, G. J. Pryde, A. Gilchrist, D. F. V. James, N. K. Langford, T. C. Ralph and A. G. White, "Quantum process tomography of a controlled-NOT gate," *Phys. Rev. Lett.*, vol. 93, no. 8, 2004, Art. ID 080502.
- [39] Y. Wang, B. Qi, D. Dong and I. R. Petersen, "An iterative algorithm for Hamiltonian identification of quantum systems," in *Proc. 55th IEEE Conf. Decision Control*, Las Vegas, NV, USA, Dec. 12-14, 2016.
- [40] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008.
- [41] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed. Baltimore, MD: JHU Press, 2013.
- [42] Y. M. Nechepurenko, "New spectral analysis technology based on the Schur decomposition," *Russ. J. Numer. Anal. Math. Modelling*, vol. 14, no. 3, pp. 265-274, 1999.
- [43] X. Wu and K. Xu, "Partial standard quantum process tomography," *Quantum Inf. Process.*, vol. 12, no.2, pp. 1379-1393, 2013.
- [44] R. Bhatia, *Matrix Analysis*, Grad. Texts in Maths. 169, Springer, Berlin, 1997.
- [45] M. D. de Burgh, N. K. Langford, A. C. Doherty and A. Gilchrist, "Choice of measurement sets in qubit tomography," *Phys. Rev. A*, vol. 78, no. 5, 2008, Art. ID 052122.
- [46] W. K. Wootters and B. D. Fields, "Optimal state-determination by mutually unbiased measurements," *Ann. Phys.*, vol. 191, no. 2, pp. 363-381, 1989.
- [47] Z. Hou, G.-Y. Xiang, D. Dong, C.-F. Li and G.-C. Guo, "Realization of mutually unbiased bases for a qubit with only one wave plate: theory and experiment," *Optics Express*, vol. 23, no. 8, pp. 10018-10031, 2015.
- [48] A. Miranowicz, K. Bartkiewicz, J. Peřina Jr., M. Koashi, N. Imoto and F. Nori, "Optimal two-qubit tomography based on local and global measurements: Maximal robustness against errors as described by condition numbers," *Phys. Rev. A*, vol. 90, no. 6, 2014, Art. ID 062123.
- [49] J. M. Renes, R. Blume-Kohout, A. J. Scott and C. M. Caves, "Symmetric informationally complete quantum measurements," *J. Math. Phys.*, vol. 45, no. 6, pp. 2171-2180, 2004.
- [50] D. Bleichenbacher and P. Q. Nguyen, "Noisy polynomial interpolation and noisy Chinese remaindering," in *EUROCRYPT*, Bruges, Belgium, 2000, pp. 53-69.
- [51] P. Lancaster and M. Tismenetsky, *The Theory of Matrices: with Applications*. New York, USA: Academic, 1985.