On exploring the potential of quantum auto-encoder for learning quantum systems

Yuxuan Du and Dacheng Tao, Fellow, IEEE

Abstract—The frequent interactions between quantum computing and machine learning revolutionize both fields. One prototypical achievement is the quantum auto-encoder (QAE), as the leading strategy to relieve the curse of dimensionality ubiquitous in the quantum world. Despite its attractive capabilities, practical applications of QAE have yet largely unexplored. To narrow this knowledge gap, here we devise three effective QAE-based learning protocols to address three classically computational hard learning problems when learning quantum systems, which are low-rank state fidelity estimation, quantum Fisher information estimation, and Gibbs state preparation. Attributed to the versatility of QAE, our proposals can be readily executed on near-term quantum machines. Besides, we analyze the error bounds of the trained protocols and showcase the necessary conditions to provide practical utility from the perspective of complexity theory. We conduct numerical simulations to confirm the effectiveness of the proposed three protocols. Our work sheds new light on developing advanced quantum learning algorithms to accomplish hard quantum physics and quantum information processing tasks.

Index Terms—Auto-encoders, Quantum machine learning, Quantum computing, Quantum information processing.

I. INTRODUCTION

MANIPULATING quantum systems and predicting their properties are of fundamental importance for developing quantum technologies [1]–[3]. Despite the importance, exactly capturing all information of a given quantum system is computationally hard, since the space of a quantum state exponentially scales with the number of particles. For instance, quantum state tomography requests $\Theta(4^N)$ copies to determine an unknown $N$-qubit quantum state [4], the recognition of entanglement of quantum states has proven to be NP-hard [5].

With the aim of understanding quantum world, two leading approaches, i.e., shadow-tomography-based methods [6]–[8] and data compression methods [9]–[11], are developed to extract useful information from investigated quantum systems from different angles. Namely, the former concentrates on predicting certain properties of an unknown quantum state, e.g., estimating the expectation values of quantum observables. In other words, this approach aims to extract partial information from the explored quantum system. The latter delves into fully reconstructing a special class of quantum systems satisfying the low-rank property. Concisely, this approach first compresses the quantum state into a low-dimensional subspace, followed by quantum tomography with a polynomial runtime complexity.

Among various quantum data compression techniques, the quantum auto-encoder (QAE) [12] and its variants [13], [14], as the quantum extension of classical auto-encoders [15], have recently attracted great attention. Conceptually, QAE harnesses a variational quantum circuit $U(\theta)$ to encode an initial input state into a low-dimensional latent space so that the input state can be recovered from this compressed state by a decoding operation $U(\theta)^\dagger$, as shown in Figure 1. Analogous to their classical counterparts, QAE is flexible, which allows us to implement it on noisy intermediate-scale quantum (NISQ) machines [16] to solve different learning problems. In particular, QAE carried out on linear optical systems has been applied to reduce qudits to qubits with low error levels [17] and to compress two-qubit states into one-qubit states [18]. Additionally, QAE has been realized on superconducting quantum processors to compress three-qubit quantum states [14]. Besides quantum data compression, initial studies have exhibited that QAEs can be exploited to denoise GHZ states which subject to spin-flip errors and random unitary noise [19] and to conduct efficient error-mitigation [20] and data encoding [21]. Despite the above achievements, practical applications of QAE in the field of quantum science remain largely unknown.

To conquer aforementioned issues, here we revisit QAE from its spectral properties and explore its potentials in learning quantum systems [22], [23], as the prerequisite to use quantum techniques to benefit various fields such as spectroscopy, chemistry, and material sciences. This exploration also aligns with the recent perspective in the quantum computing community, suggesting that quantum machine learning has the potential to significantly accelerate quantum data analysis compared to classical methods [24], [25]. To do so, we exhibit that the compressed state of QAE living in the latent space conveys sufficient spectral information of the input $N$-qubit state $\rho$. This means that when $\rho$ is low-rank, the spectral information of $\rho$ can be efficiently acquired via applying quantum tomography [26] to the compressed state in polynomial runtime with $N$. This result hints a number of applications of QAE in learning quantum systems, because numerous tasks in this field can be recast to capture the spectral information of the given state. Based on this understanding, we devise three QAE-based learning protocols, i.e., QAE-based fidelity estimator (Section IV), QAE-based quantum Fisher information estimator (Section V), and QAE-based Gibbs state solver (Section VI), to address three classically computational hard problems in learning quantum systems, i.e., low-rank state fidelity estimation [27], quantum Fisher information estimation [28], and quantum Gibbs state preparation [29], [30], respectively. Our study
Then, we analyze the spectral property of QAE in Section 0, and 
\( P \) is the unit vector whose \( a | \) refers to a two-dimensional vector. Under Dirac notation, a \( A \). Basic concepts of quantum computing

conclude this study and discuss future research directions.

II. BACKGROUND

A. Basic concepts of quantum computing

The fundamental unit in quantum computing is qubit, which refers to a two-dimensional vector. Under Dirac notation, a qubit state is denoted by \( |a_0\rangle = a_0 |0\rangle + a_1 |1\rangle \in \mathbb{C}^2 \), where \( |0\rangle = [1, 0]^\top \) and \( |1\rangle = [0, 1]^\top \) specify two unit bases, and the coefficients \( a_0, a_1 \in \mathbb{C} \) satisfy \( |a_0|^2 + |a_1|^2 = 1 \). An \( N \)-qubit state is denoted by \( |\Psi\rangle = \sum_{i=1}^{2^N} a_i |i\rangle \in \mathbb{C}^{2^N} \), where \( |i\rangle \in \mathbb{R}^{2^N} \) is the unit vector whose \( i \)-th entry being 1 and other entries are 0, and \( \sum_{i=0}^{2^N-1} |a_i|^2 = 1 \) with \( a_i \in \mathbb{C} \). Following conventions, an \( N \)-qubit separable state can be written as \( |a_1, \ldots, a_N\rangle \) or \( |a_1, \ldots, a_N\rangle \). We sometimes denote \( |a\rangle |b\rangle \) as \( |a\rangle_A |b\rangle_B \), which means that the qubits \( |a\rangle_A |b\rangle_B \) are assigned in the quantum register \( A (B) \). Besides Dirac notation, the density matrix can be used to describe more general qubit states. For example, the density matrix of the state \( |\Psi\rangle \) is \( \rho = |\Psi\rangle \langle \Psi| \in \mathbb{C}^{2^N \times 2^N} \). For a set of qubit states \( \{ p_i |\psi_i\rangle \}_{i=1}^{m} \) with \( p_i > 0 \), \( \sum_{i=1}^{m} p_i = 1 \), and \( |\psi_i\rangle \in \mathbb{C}^{2^N} \) for all \( \in [m] \), its density matrix is \( \rho = \sum_{i=1}^{m} p_i |\psi_i\rangle \langle \psi_i| \) with \( p_i = |\psi_i\rangle \langle \psi_i| \) and \( \text{Tr} (\rho) = 1 \).

Quantum gates refer to unitary transformations and serve as the computational toolkit for quantum circuit models, i.e., an \( N \)-qubit gate \( U \in U(2^N) \) obeys \( U U^\dagger = U^\dagger U = \mathbb{I}_{2^N} \), where \( \mathbb{I} \) stands for the unitary group. Throughout the whole study, we focus on the single-qubit and two-qubit quantum gate set, i.e., \( \text{RZ}(\theta) = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix} \), \( \text{RY}(\theta) = \begin{pmatrix} \cos \theta/2 & -\sin \theta/2 \\ \sin \theta/2 & \cos \theta/2 \end{pmatrix} \), and \( \text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \). Note that this gate set is universal such that these quantum gates can be used to reproduce the functions of all the other quantum gates [26].

Quantum measurements are employed to extract quantum information of the evolved quantum state into the classical form. In this study, we concentrate on the positive operator-valued measures (POVM), which is described by a collection of positive operators \( 0 \leq \Pi_i \) satisfying \( \sum_i \Pi_i = \mathbb{I} \). Specifically, applying the measurement \( \{ \Pi_i \} \) to the state \( \rho \), the probability of outcome \( i \) is given by \( \text{Pr}(i) = \text{Tr}(\rho \Pi_i) \).

B. Quantum machine learning

Quantum machine learning (QML) employs quantum computers to gain computational advantages in learning both classical data (e.g., images, videos, and texts) and quantum data (e.g., dynamics and correlation properties of quantum systems) [1]. Learning quantum systems is significant as quantum technology development relies on creating and manipulating quantum systems [32], [33], which become exponentially complex with size, making their full description intractable [22], [23]. Thus, a key research goal is to develop QML models that surpass classical models in solving crucial problems in quantum science such as quantum computing, communication, and metrology [34], [35].

A leading candidate of QML models in NISQ era is variational quantum algorithms (VQAs), also known as quantum neural networks (QNNs) [36]–[43]. Conceptually, a VQA contains four principal blocks: the input state \( \rho_0 \), a variational quantum circuit (or equivalently ansatz) \( U(\theta) \), the measurement \( M \), and a loss function \( \mathcal{L}(\theta) \). In the training stage, VQA follows an iterative manner to proceed optimization, where the classical optimizer continuously leverages the output of the quantum circuit, i.e., \( \text{Tr}(MU(\theta)\rho_0U(\theta)^\dagger) \), to update trainable parameters \( \theta \) to minimize \( \mathcal{L}(\theta) \). The updating rule at the \( t \)-th iteration is \( \theta^{(t+1)} = \theta^{(t)} - \eta \nabla_\theta \mathcal{L}(\theta) \), where \( \eta \) is the learning rate and \( \nabla_\theta \mathcal{L}(\theta) \) is the gradients of \( \mathcal{L}(\theta) \) that can be acquired via parameter shift rule [44]. The updating is terminated if \( \mathcal{L}(\theta) \) is converged or the iteration \( t \) exceeds the threshold.
C. Classical auto-encoder

Auto-encoder (AE) is a technique in the context of machine learning to learn efficient codings of unlabeled data and has been broadly applied to clustering [45], spatio-temporal data mining [46], information retrieval [47], anomaly detection [48], and image processing [49]. Intuitively, AE completes the data compression by mapping the given input from a high-dimensional space to a low-dimensional space in the sense that the input can be recovered from this dimension-reduced point. As shown in the upper left panel of Figure 1, two basic elements of AE are the encoder \( f(\cdot) \) and the decoder \( g(\cdot) \), which are realized by neural networks [15]. Suppose that the given dataset is \( X \in \mathbb{R}^{m \times n} \) whose \( i \)-th column \( x_i \in \mathbb{R}^m \) represents the \( i \)-th data point. The encoder \( f : \mathbb{R}^m \to \mathbb{R}^k \) with \( k \ll m \) compresses each data point into a \( k \)-dimensional space \( f(x_i) \), while the decoder \( g : \mathbb{R}^k \to \mathbb{R}^m \) maps the latent representation \( f(x_i) \) to the reconstructed data \( \hat{x}_i = g(f(x_i)) \) [15]. The optimization of AE amounts to minimizing the reconstruction error between \( X \) and its reconstruction \( \hat{X} \) whose \( i \)-th column is \( \hat{x}_i \).

D. Quantum auto-encoder

Quantum auto-encoder (QAE), as the quantum analog of AE, is composed of the quantum encoder and quantum decoder as exhibited in the upper right panel of Figure 1. The construction of the quantum encoder (or decoder) is completed by variational quantum circuits. Mathematically, define an \( N \)-qubit input state as \( \rho \in \mathbb{C}^{2^N \times 2^N} \) and the number of qubits that represents the latent space as \( K \) with \( K < N \). Let the quantum encoder be a variational quantum circuit \( U(\theta) \), where \( U(\theta) = \prod_{i=1}^{N} U_i(\theta) \) and \( U_i(\theta) \) is a sequence of parameterized single-qubit and two-qubits gates with \( \theta \) being tunable parameters. Under the above setting, QAE aims to find optimal parameters \( \theta^\star \) that minimize the loss function

\[
\mathcal{L}(U(\theta), \rho) = \text{Tr} \left( M_G U(\theta) \rho U(\theta)^\dagger \right),
\]

where \( M_G = (\mathbb{1}_N - (|0\rangle \langle 0|)^\otimes (N-K) \otimes \mathbb{1}_K) \) refers to the measurement operator and \( \mathbb{1}_N \in \mathbb{R}^{2^N \times 2^N} \) is the identity. While both AEs and QAEs use encoders and decoders to process data, QAEs are specifically designed for quantum computers and are therefore better suited for learning quantum data, as illustrated in the lower panel of Figure 1 in contrast to AEs which are typically used for classical data.

E. Related work

Previous literature related to our work can be divided into two categories. The first category is QAEs and the second group is other quantum learning methods to accomplish the tasks explored in this study. Here we separately compare them with our work.

The variants of QAEs have been proposed to compensate certain deficiencies existed in the original QAE proposed by [12]. In particular, the proposal developed in [14], [50] combines QAE with quantum adders, which can approximately add two unknown quantum states supported in different quantum systems. Ref. [51] proposed a hybrid QAE to tackle clustering and classification task. On par with the model design, initial studies have been conducted to explore applications of QAEs in state compression [17], [18], denoising [19] and error-mitigation [20]. Different from prior studies that mainly focus on the feasibility of QAEs, our work explores their practical applications in learning quantum systems.

Several studies have designed other variational quantum algorithms to address the learning problems exploited in this study. For the task of fidelity estimation, the protocol proposed by Ref. [52] can be applied to full-rank quantum states; the protocol proposed by Ref. [53] intended to the low-rank fidelity estimation. Compared with these two method, our proposal is more resource-efficient. For the task of Gibbs state preparation, a relevant study is Ref. [53]. In that study, the protocol leverages the technique of Taylor truncation to estimate Von-Neumann entropy, which introduces the estimation error. Conversely, our approach does not require such an estimation.

III. Spectral properties of QAE

We now revisit QAE from its spectral properties. Denote the rank of the state \( \rho \) as \( r^\star \) and its spectral decomposition as

\[
\rho = \sum_{i=1}^{r^\star} \lambda_i |\psi_i\rangle \langle \psi_i| \quad \text{with} \quad \sum_{i=1}^{r^\star} \lambda_i = 1,
\]

where \( \lambda_i \) and \( |\psi_i\rangle \) are the \( i \)-th eigenvalues and eigenvectors, respectively. Without loss of generality, we suppose \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{r^\star} > 0 \).

![Figure 1: The paradigm of AE and QAE.](image-url)
\[ \lambda_2 \geq \ldots \geq \lambda_r. \] The optimal \( U(\theta^*) \) in Eqn. (1) is shown in the following theorem whose proof is given in Appendix A.

**Theorem 1.** QAE with the loss function \( L(U(\theta), \rho) \) in Eqn. (2) has multiple critical (global minimum) points of \( U(\theta^*) \). Let \( r^* = 2^K \). The generic form of \( U(\theta^*) \) is

\[ U(\theta^*) := \sum_{i=1}^{r^*} 0^{\odot N-K} |\varphi_i⟩ynchronormal vectors with \( r \)

\[ \quad + \sum_{j=r+1}^{2N} |\phi_j⟩\langle ψ_j|, \quad (3) \]

where \( \{0\}^{\odot N-K} |\varphi_i⟩ \}_{i=1}^{r^*+1} \cup \{|\phi_j⟩\}_{j=r+1}^{2N} \) and \( \{|\psi_i⟩\}_{i=1}^{r^*} \) are two sets of orthonormal vectors and \( \{\varphi_i⟩\}_{i=r^*+1}^{2N} \) is defined in Eqn. (2).

Given access to the optimal quantum encoder \( U(\theta^*) \) in Eqn. (3), the compressed state of QAE yields

\[ \sigma^* = Tr_E \left( \frac{M_G U(\theta^*) \rho U(\theta^*)^\dag M_G^\dag}{Tr(M_G U(\theta^*) \rho U(\theta^*)^\dag)} \right), \quad (4) \]

where the term `'Tr_E(·)′' represents the partial trace of all qubits except for the K latent qubits [26]. The spectral information between \( \sigma^* \) and \( \rho \) is closely related, as indicated by the following lemma, where the corresponding proof is provided in Appendix B.

**Lemma 1.** The compressed state has the same eigenvalues with \( \rho \), i.e., \( \sigma^* = \sum_{i=1}^{r^*} \lambda_i |\varphi_i⟩\langle \varphi_i| \). Moreover, the subspace spanned by the eigenvectors \( \{|\varphi_i⟩\}_{i=1}^{r^*} \) of \( \rho \) can be recovered by \( U(\theta^*) \) and eigenvectors \( \{|\varphi_i⟩\}_{i=1}^{r^*} \).

The result of Lemma 1 indicates that when the input state \( \rho \) is low-rank, the spectral information of \( \sigma^* \) can be efficiently obtained by quantum tomography associated with classical post-processing. In particular, the size of the \( K \)-qubit compressed state \( \sigma^* \) scales with \( r^* \times r^* \), which implies that the classical description of \( \sigma^* \) can be obtained in \( O(\text{poly}(r^*)) \) runtime. Moreover, the collected classical information suggests that the spectral decomposition of \( \sigma^* \), i.e., \( \{\lambda_i\} \) and \( \{|\varphi_i⟩\} \), can be computed in \( O(\text{poly}(r^*)) \) runtime. Taken together, the spectral information of a low-rank state \( \rho \) can be efficiently acquired via QAE in the optimal case.

In the subsequent sections, we elucidate how to employ QAE, especially the acquired spectral information, to facilitate the low-rank state fidelity estimation, quantum Fisher information estimation, and Gibbs state preparation tasks. Due to the intrinsic hardness, any such improvement would not only be immensely valuable for various downstream tasks in quantum science such as quantum computing, quantum sensing, and quantum metrology but also indicates the practical utility of near-term quantum computers. For clarity, we unify some notations throughout the rest study. Denote the spectral decomposition of the variational quantum circuits \( U(\hat{\theta}) \) as

\[ U(\hat{\theta}) := \sum_{i=1}^{r} 0^{\odot N-K} |\varphi_i⟩\langle \varphi_i| + \sum_{j=r+1}^{2N} |\phi_j⟩\langle \psi_j|, \quad (5) \]

where \( \hat{\theta} \) refers to the trained parameters, and \( \{0\}^{\odot N-K} |\varphi_i⟩ \}_{i=1}^{r+1} \cup \{|\phi_j⟩\}_{j=r+1}^{2N} \) and \( \{|\varphi_i⟩\}_{i=1}^{r^*} \) are two sets of orthonormal vectors with \( r = 2^K \). Note that \( r \) may not be equal to \( r^* \) when the training procedure is not optimal. Besides, since the Eigen-decomposition of \( U(\theta) \) is not unique and it operates with the projector \( M_G \), in Eqn. (1), we set \( U(\theta) \) a similar form as shown in Eqn. (3) for ease of analysis. Based on Eqn. (3), the explicit form of the compressed state of QAE in the non-optimal case is

\[ \hat{\sigma} = Tr_E \left( \frac{M_G U(\theta) \rho U(\theta)^\dag M_G^\dag}{Tr(M_G U(\theta) \rho U(\theta)^\dag)} \right) = \sum_{i=1}^{r^*} \hat{\lambda}_i |\hat{\varphi}_i⟩\langle \hat{\varphi}_i|, \quad (6) \]

where \( \hat{\lambda}_i = \langle \hat{\varphi}_i|\rho|\hat{\varphi}_i⟩ / \left( \sum_{j=1}^{r^*} \langle \hat{\varphi}_j|\rho|\hat{\varphi}_j⟩ \right) \).

**IV. QAE-BASED FIDELITY ESTIMATOR**

The quantification of the quantum state fidelity plays a central role in verifying and characterizing the states prepared by a quantum processor [26] and studying quantum phase transitions [54]. In other words, devising an effective fidelity estimator is the precondition of using quantum computer to benefit chemistry, quantum physics, and material sciences. Despite its importance, quantum state fidelity calculation is computationally hard for both classical and quantum devices [31]. Mathematically, the fidelity of two states \( \rho \) and \( \kappa \) yields

\[ F(\rho, \kappa) = Tr \left( \sqrt{\sqrt{\rho} \kappa \sqrt{\rho}} \right) = \|\sqrt{\rho} \kappa \|_1. \quad (7) \]

However, when we slightly relax the problem and target to the low-rank state fidelity estimation (i.e., the rank of \( \rho \) is no larger than \( O(\text{poly}(N)) \) and \( \kappa \) is unconstrained), quantum computers may provide potential advantages over classical devices, supported by the following lemma.

**Lemma 2** (Proposition 5, [31]). The problem of the \( N \)-qubit low-rank state fidelity estimation within the precision \( \pm 1/\text{poly}(N) \) is DQC1-hard.

Recall that classical computers can not efficiently simulate DQC1 models unless the polynomial-time hierarchy collapses to the second level [55]. The result of Lemma 2 shows that if a quantum algorithm can achieve low-rank state fidelity estimation in polynomial runtime with an error below \( \text{poly}(1/N) \), it has a runtime advantage compared to the optimal classical method.

Here we devise a QAE-based fidelity estimator to complete the low-rank state fidelity estimation task and discuss its potentials. The underlying concept of our proposal is as follows. Recall an equivalent formula of fidelity is

\[ F(\rho, \kappa) = Tr \left( \sum_{i,j} W_{i,j} |\psi_i⟩\langle \psi_j| \right), \quad (8) \]

where \( W_{i,j} = \sqrt{\lambda_i} \langle \psi_i|\kappa|\psi_j⟩ \) for \( \forall i \in [r^*] \). In other words, calculating \( F(\rho, \kappa) \) amounts to acquiring \( W \in \mathbb{C}^{r^* \times r^*} \), which can be effectively achieved once the state \( \kappa \) and the eigenvalues and eigenvectors of \( \rho \) are accessible. The spectral information of \( \rho \) provided by QAE allows us to estimate the fidelity \( F(\rho, \kappa) \).
estimator. As shown in Figure 2, the proposed method consists of four steps. (i) QAE in Eqn. (1) is applied to optimize \( \theta \) with \( T \) iterations. (ii) After optimization, the compressed state \( \sigma^{(T)} \) in Eqn. (10) is extracted into the classical register via quantum state tomography. Once the classical form of \( \sigma^{(T)} \) is accessible, the classical Eigen-solver is employed to acquire its spectral information, i.e., \( \{ \lambda_i, |\varphi_i\rangle \} \). (iii) When the spectral information of \( \rho \) is available and the state \( \kappa \) is accessible, we can calculate \( W \) in Eqn. (3) based on Ref. [31]. (iv) Given access to \( W \), the fidelity \( F(\rho, \kappa) \) can be effectively obtained.

![Fig. 2: The paradigm of the QAE-based fidelity estimator.](image)

### A. Implementation of QAE-based fidelity estimator

We now elaborate on the protocol of the QAE-based fidelity estimator. As shown in Figure 2, the proposed method consists of four steps.

1) We apply QAE formulated in Eqn. (1) to compress the given state \( \rho \). Let the total number of training iterations be \( T \). The updating rule of the parameters \( \theta \) at the \( t \)-th iteration yields

\[
\theta^{(t+1)} = \theta^{(t)} - \eta \frac{\partial L(U(\theta^{(t)}), \rho)}{\partial \theta^{(t)}},
\]

where \( \eta \) refers to the learning rate. The gradients \( \frac{\partial L(U(\theta^{(t)}), \rho)}{\partial \theta^{(t)}} \) can be acquired by using the parameter shift rule [41], [42]. After the optimization is completed, the trained parameters \( \theta^{(T)} \) are exploited to prepare the compressed state in Eqn. (6), i.e.,

\[
\sigma^{(T)} = \text{Tr} (M_G U(\theta^{(T)})) \text{Tr}(U(\theta^{(T)})^\dagger M_G^\dagger) / \text{Tr}(M_G U(\theta^{(T)})) \text{Tr}(U(\theta^{(T)})^\dagger M_G^\dagger)).
\]

2) The quantum state tomography techniques [26] are applied to \( \sigma^{(T)} \) with size \( 2^K \times 2^K \) to obtain its classical description. For simplicity, here we only focus on the ideal case with the infinite number of measurements so that the state \( \sigma^{(T)} \) can be exactly recorded in the classical register. When the measurement error is considered, the likelihood recovering strategy can be used to efficiently find the closest probability distribution of \( \sigma^{(T)} \) [56]. Once the classical form of \( \sigma^{(T)} \) is in hand, its spectral decomposition is \( \{ \lambda_i \} \) and \( \{ |\varphi_i\rangle \} \). Given the spectral information of \( \sigma^{(T)} \), the fidelity \( F(\rho, \kappa) \) is estimated by

\[
\hat{F}(\rho, \kappa) = \text{Tr}(\sum_{i,j} \sqrt{\lambda_i \lambda_j} |\hat{\varphi}_i| \kappa |\hat{\varphi}_j\rangle \langle \hat{\varphi}_i| \langle \hat{\varphi}_j|)^{-1/2}),
\]

where \( |\hat{\varphi}_i\rangle = U(\theta^{(T)}) |0\rangle \otimes N-K, |\varphi_i\rangle \). An important question is quantifying the estimation error, i.e., \( |\hat{F}(\rho, \kappa) - F(\rho, \kappa)| \). The following theorem indicates that the estimation error can be bounded by the quantity

\[
\zeta = \sqrt{1 - F^2(\rho, \tilde{\rho})},
\]

where \( F(\rho, \tilde{\rho}) \equiv \hat{F}(\rho, \rho) \) refers to the fidelity between the reconstructed state of the trained QAE \( \tilde{\rho} \) and the target state \( \rho \). The evaluation of \( \zeta \) is computationally efficient, since as the calculation of \( \hat{F}(\rho, \kappa) \), \( F(\rho, \kappa) \) can be efficiently attained by using Eqn. (14).

**Theorem 2.** Following notations in Eqns. (14) and (15), the QAE-based fidelity estimator after \( T \) iterations returns the estimated fidelity \( \hat{F}(\rho, \kappa) \) with

\[
\hat{F}(\rho, \kappa) - \sqrt{\zeta} \leq F(\rho, \kappa) \leq \hat{F}(\rho, \kappa) + \sqrt{\zeta}.
\]

The proof of Theorem 2 is provided in Appendix C. The achieved results indicate how the computationally-efficient quantity \( \zeta \) reflects the upper bound of the estimation error \( F(\rho, \kappa) - \hat{F}(\rho, \kappa) \) when the optimization of QAE is not optimal. Moreover, the estimation error in Eqn. (16) hints conditional advantages of our proposal. In particular, after a polynomial number of iterations with \( T \sim O(\text{poly}(N)) \), our proposal could reach runtime advantages when \( \sqrt{\zeta} \). This is because such a restriction on \( T \) leads that the total runtime complexity of our proposal polynomially scales with \( N \) to attain a low estimation error, which is unachievable by classical methods, according to Lemma 2.
Another intriguing property of the quantity $\zeta$ is that it is lower bounded by the training loss of QAE $\delta$, as elaborated in Appendix D. In this regard, to suppress the estimation error of QAE-based fidelity estimator, the key is suppressing $\delta$. Possible strategies towards this goal encompass the employment of problem-oriented ansatz to implement the quantum encoder $U(\theta)$ [58] and the adoption of advanced optimization techniques [39], [60] to improve the trainability of QAE. Moreover, it is noteworthy that according to the computational complexity theory, there may exist a class of computationally-hard quantum states that cannot be efficiently compressed by QAE in which $\zeta > 1/\text{poly}(N)$.

**Remark.** We emphasize two key points regarding the QAE-based fidelity estimator. First, our analysis aligns with convention of NISQ algorithms to highlight the conditional advantages of our approach, contingent upon the fulfillment of specific requirements [24], [61]. This contrasts from the study of fault-tolerant quantum algorithms, where the deterministic advantage should be demonstrated by analyzing an explicit runtime complexity bound and showcasing provable speedups over the best classical counterparts [62]. Second, our proposal distinguishes itself from the classical shadow [8], as it cannot calculate fidelity between two mixed states. Namely, the classical shadow is specifically designed to estimate the expectation values of observables $\{O_i\}$ with $\{\text{Tr}(\rho O_i)\}$. However, the fidelity defined in Eqn. (7) can only be reduced to this form when one of the states is a pure state and its classical description is known.

### B. Numerical simulations

We conduct numerical simulations to evaluate performance of QAE-based fidelity estimator.

**The construction of the employed quantum states.** We synthesize two pairs of states and employ QAE-based Fidelity estimator to calculate their fidelity. The first pair of states contains two eight-qubit state $\rho$ and $\sigma$ with $N = 8$. The construction of $\rho$ and $\sigma$ can be decomposed into two steps. In the first step, we prepare two pure states, i.e., $|\Psi(0)\rangle = |0\rangle^{\otimes 8}$ and $|\Psi(1)\rangle = \alpha_1 |10\ldots0\rangle + \alpha_2 |01\ldots0\rangle + \ldots + \alpha_8 |00\ldots0\rangle$, where the coefficients are fixed to be $\alpha_1 = \alpha_2 = \ldots = \alpha_8 = 1/8$. Once the pure state $|\Psi\rangle$ (e.g., $|\Psi(0)\rangle$ or $|\Psi(1)\rangle$) is prepared, the noisy channel $N(\cdot)$ is applied to this state, i.e.,

$$N(|\Psi\rangle, p, r, a) = \begin{cases} |\Psi\rangle, & \text{with probability } p; \\ V(r), & \text{with probability } 1-p, \end{cases}$$

where $a \in \mathbb{R}$ is a hyper-parameter, $V(r) \in \mathbb{R}^{2^8 \times 2^8}$ refers to a density operator with $V_{ii} = 1.5^{-ai}$ when $i \leq r$ and $V_{ii} = 0$ when $i > r$, and $V_{ij} = 0$ when $i \neq j$. To construct $\rho$ and $\kappa$, we set $\rho = N(|\Psi(0)\rangle, p = 0.1, r = 8, a = 2)$ and $\kappa = N(|\Psi(1)\rangle, p = 0.5, r = 17, a = 5)$, respectively. In this way, $\rho$ and $\kappa$ correspond to two eight-qubit states with rank $r^* = 8$ and $r^* = 17$. The exact fidelity between $\rho$ and $\sigma$ is $F(\rho, \sigma) = 0.765$.

We follow the same construction rule to produce the second pair of states. The number of qubits is $N = 8$ and we set the two states to be the same with the rank $2^8$, i.e., $\rho = \sigma$. The parameter setting of the noisy channel is $p = 0.5, r = 2^8, a = 1/10$. The explicit form of the two states is $\rho = \sigma = N(|\Psi(0)\rangle, 0.5, 2^8, 1/10)$. The exact fidelity is $F(\rho, \sigma) = 1$.

**Hyper-parameters settings.** For both of the two pair states, we vary the size of latent space with $K = 1, \ldots, 7$. The initialized parameters are sampled from Gaussian distribution with the same mean and the differed variance. Namely, for all settings, the mean is set as zero while the variance is set as $0.4 \times (N-K)$. The Adam optimizer is employed to complete all simulations with the learning rate being $0.08$. The implementation of the quantum encoder $U(\theta)$ employs the hardware-efficient ansatz

$$U(\theta) = \prod_{l=1}^{L} U_l(\theta).$$

An intuition is shown in the lower panel of Figure 3(a). The number of iterations is $T = 200$. Considering that a higher expressive ansatz is necessary to suppress the input state into a lower dimensional space, the layer number of ansatz depends on the number of latent qubits $K$. That is, for the first pair of states, we set $L = 11 - K$; for the second pair of states, we set $L = 13 - K$. To analyze the robustness of our proposal, each setting described above is repeated with five times.

**Code and device information.** The proposed three QAE-based learning protocols are implemented by Python, with PennyLane as the backbone. We conduct all numerical simulations on TESLA P40 GPUs using the simulation backends provided by PennyLane. The source code is available at https://github.com/yuxuan-du/Quantum-auto-encoders-based-learning-protocols.

**The sub-fidelity and super-fidelity bounds.** We employ the sub-fidelity and super-fidelity bounds (SFB) [27] as reference to benchmark performance of the QAE-based fidelity estimator. Specifically, the sub-fidelity between $\rho$ and $\kappa$ yields

$$F_L(\rho, \kappa) = \text{Tr}(\rho \kappa) = \sqrt{2 ((\text{Tr}(\rho \kappa))^2 - \text{Tr}(\rho \kappa \rho \kappa))}.$$ (19)

Moreover, the super-fidelity between $\rho$ and $\kappa$ yields

$$F_U(\rho, \kappa) = \text{Tr}(\rho \kappa) = \sqrt{(1 - \text{Tr}(\rho^2))(1 - \text{Tr}(\kappa^2))}.$$ (20)

**Simulation results.** The simulation results for the first pair of states are illustrated in Figure 3(b)-(e). In particular, the sub-fidelity between $\rho$ and $\kappa$ yields

$$F_L(\rho, \kappa) = \text{Tr}(\rho \kappa) = \sqrt{2 ((\text{Tr}(\rho \kappa))^2 - \text{Tr}(\rho \kappa \rho \kappa))}.$$ (19)

Moreover, the super-fidelity between $\rho$ and $\kappa$ yields

$$F_U(\rho, \kappa) = \text{Tr}(\rho \kappa) = \sqrt{(1 - \text{Tr}(\rho^2))(1 - \text{Tr}(\kappa^2))}.$$ (20)
true fidelity $F(\rho, \kappa)$, as indicated by the green solid and dashed line with cross markers. In the measure of top-1 performance, QAE-based Fidelity estimator achieves better performance than SSFB methods (the solid and dashed grey lines) when $K \geq 3$. The above analysis exhibits the potential of our proposal when the number of latent qubits satisfies $2^K \geq r^*$. The simulation results for the second pair of states are shown in Figure 4. The subplot (a) shows the training loss of QAE with respect to the different $K$. An observation is that QAE-based fidelity estimator is robust, since its variance can be ignorable. Moreover, under both the average and top-1 measures, the training loss $\delta$ is lower than 0.02 only when $K = 7$. The degraded performance compared to the first task is caused by the fact that $\rho$ is full rank, which is incompressible by QAE. As such, some information in $\rho$ must be lost in the compressed state $\sigma$ when $2^K < r^*$. As a result, increasing $K$ enables $\sigma$ to preserve more information in $\rho$ and reduce the training error $\delta$. Subplot (b) depicts the estimated fidelity bounds of the QAE-based fidelity estimator, the SSFB, and the exact result. On average, QAE-based fidelity estimator outperforms SSFB when $K = 7$. In the measure of top-1 performance, QAE-based fidelity estimator outperforms SSFB when $K \geq 6$.

V. QAE-BASED QFI ESTIMATOR

The high precision of sensing is of great importance to many scientific and real-world applications, which include but not...
limit to chemical structure identification [63] and gravitational wave detection [64]. However, the central limit theorem indicates that the statistical error in the classical scenario must be lower bounded by $\nu^{-1/2}$, where $\nu$ is the number of measurements. Driven by the significance of improving the precision of sensing, it highly desired to utilize quantum effects to further reduce the statistical error.

Quantum metrology provides a positive response towards the above questions [65]–[67]. Through leveraging coherence or entangled quantum states, the statistical error can be reduced by an amount proportional to $\nu^{-1}$. A fundamental quantity in quantum metrology is quantum Fisher information (QFI) [28], which determines the best possible estimation error for a certain quantum state. Mathematically, let $\rho_0(\gamma)$ be the evolved state, which is prepared by interacting a tunable $N$-qubit state $\rho(\gamma)$ with a source that encodes the environmental information in a single parameter $\gamma$. The QFI $\mathcal{I}(\theta; \rho_0(\gamma))$ measures the ultimate precision $\Delta\theta$, i.e., $(\Delta\theta)^2 \geq 1/(\nu \mathcal{I}(\theta; \rho_0(\gamma)))$, where $\nu$ is the number of measurements to estimate $\theta$. At the expense of computational simplicity, $\mathcal{I}(\theta; \rho_0(\gamma))$ is estimated by

$$\mathcal{I}_\tau(\theta; \rho_0) = \frac{1 - F(\rho_0, \rho_{0+\tau})}{\tau^2}. \quad (21)$$

When $\tau \to 0$, the approximation error suppresses to 0, i.e., $\mathcal{I}_\tau(\theta; \rho_0(\gamma)) = \mathcal{I}_{\tau=0}(\theta; \rho_0(\gamma))$. Intuitively, a true state $\rho_0(\gamma)$ with a high QFI $\mathcal{I}_\tau(\theta; \rho_0(\gamma))$ is sharply different from the error state $\rho_{0+\tau}(\gamma)$, which implies that the parameter $\gamma$ can be effectively estimated via measurement.

As indicated in Lemma 2, the calculation of the fidelity $F(\rho_0(\gamma), \rho_{0+\tau}(\gamma))$ in Eqn. (21) is classically computational hard for general density matrices. Such an inefficiency prohibits the applicability of quantum metrology. To address this issue, Ref. [68] generalized the fidelity estimation solver [31] to estimate $\mathcal{I}_\tau(\theta; \rho_0(\gamma))$ with potential runtime advantages. Their proposal employed variational quantum circuits to prepare states $\rho_0(\gamma)$ and $\rho_{0+\tau}(\gamma)$ by adjusting $\gamma$. Once these two quantum states are prepared, the variational fidelity estimator [31] is applied to estimate $\mathcal{I}_\tau(\theta; \rho_0(\gamma))$. The optimization is conducted to find optimal $\gamma^*$ that maximizes $\mathcal{I}_\tau(\theta; \rho_0(\gamma))$.

Here we combine variational quantum circuits with the QAE-based fidelity estimator to construct a more advanced QFI estimator. Recall that to balance the computational efficiency and the learning performance, the variational QFI estimator in [68] exploits the variational fidelity estimator to estimate $F(\rho_0(\gamma), \rho_{0+\tau}(\gamma))$. In this regard, we can employ QAE-based fidelity estimator to devise the QAE-based QFI estimator, which can be integrated into advanced quantum metrology protocols to obtain precise estimation using few measurements.

### A. Implementation of the QAE-based QFI estimator

The implementation of the QAE-based QFI estimator is illustrated in Figure 5. In particular, our proposal repeats the following procedure with in total $T$ times. At the $t$-th iteration, the variational quantum circuit $U(\gamma(t))$ is utilized to generate the probe state $\rho_0(\gamma(t))$, as highlighted by the red box. The calculated fidelity is used to compute $\mathcal{I}_\tau(\theta, \rho_0(\gamma))$ in Eqn. (21) and then update parameters $\gamma$ to maximize this quantity. Repeating the above procedures with $T$ times, the classical processor outputs $\gamma^{(T)}$, which can be used to prepare the estimated probe state $\rho(\gamma^{(T)})$.

![Diagram](image.png)
B. Numerical simulations

Problem setup. We apply the QAE-based QFI estimator to prepare the probe state $\rho(\gamma)$ that maximizes the QFI for a magnetic field described by the parameter $\gamma$. Mathematically, the probe state interacting with the magnetic field transforms to $\rho_d(\gamma) = W(\gamma)\rho(\gamma)W(\gamma)^\dagger$, where $W(\gamma) = e^{-i\theta G}$ is a unitary transformation, $G = \sum_{i=1}^J Z_i$ is the problem Hamiltonian, and $Z_i$ is the Pauli-Z matrix on the $i$-th qubit. We set $J = 4$ and $\theta = 0.1$ to proceed the following simulations.

Fig. 6: The simulation results of the QAE-based QFI estimator. The outer and inner plots depict the average and top-1 performance of the proposed QAE-based QFI estimator. The shaded region represents the standard deviation. The labels ‘Exact’ and ‘Estimator’ separately represent the maximum QFI that can be achieved and the estimated QFI returned by the proposed estimator.

Hyper-parameters settings. We adopt the following hyper-parameter settings to conduct all simulations. Specifically, the variational quantum circuits $U(\gamma)$ and $U(\theta)$, which are used to prepare the probe state $\rho_d(\gamma)$ and to build the quantum encoder respectively, follow the hardware-efficient ansatz in Eqn. [18]. The layer number $L$ for $U(\gamma)$ and $U(\theta)$ is set as 6 and 5, respectively. A slight modification in building $U(\gamma)$ is removing all $R_Z$ gates in the single-qubit layer to accelerate training. For QAE, the number of latent qubits is set as 2, the number of iterations is 300, and in each iteration, the Adam optimizer is used to update $\theta$ with the learning rate being 0.02. For the probe state, the total number of iterations is 75 and in each iteration, the parameters $\gamma$ are updated by parameter shift rule with the learning rate being 0.08. Each setting is repeated with five times to obtain the statistical results.

Simulation results. The simulation results are exhibited in Figure 6. Recall that a well-known conclusion in quantum metrology is that the optimal probe state is an $N$-qubit GHZ state in the noiseless scenario, where the QFI reaches the Heisenberg limit $4N^2$. Compared with the optimal result 64, the average QFI of the optimized probe state oscillates around 60 after 30 iterations with the standard deviation up to 3. The inner plot illustrates the top-1 performance of our proposal. In the 46-th step, its QFI reaches 63.317. The aforementioned observation highlights the feasibility of utilizing QAE-based QFI estimator for preparing the probe state, and suggests an intriguing avenue for future research to enhance the robustness and stability of the proposed QAE-based QFI estimator.

VI. QAE-BASED QUANTUM GIBBS STATE SOLVER

Gibbs state describes the thermal equilibrium properties of quantum systems. Efficient preparation of quantum Gibbs states is a central problem in quantum simulation, quantum optimization, and quantum machine learning [30]. As such, an efficient quantum Gibbs state preparation solver can benefit many applications. For example, the quantum semi-definite programs exploit quantum Gibbs states to earn runtime speedups [29]. The mathematical expression of quantum Gibbs state is

$$\rho_G = \frac{e^{-\beta H}}{\text{Tr}(e^{-\beta H})},$$

where $\beta$ refers to the inverse temperature and $H \in \mathbb{C}^{2N\times2N}$ is a specified $N$-qubit Hamiltonian. However, theoretical results based on computational complexity theory [69] indicate that the preparation of quantum Gibbs states at low-temperature is difficult, i.e., computing the partition function $\text{Tr}(e^{-\beta H})$ in two-dimensions is NP-hard [70]. Driven by the significance and the intrinsic hardness of preparation, Ref. [53] explored how to use variational quantum circuits implemented on NISQ devices to generate a state $\rho(\gamma^*) \in \mathbb{C}^{2N\times2N}$ approximating $\rho_G$. In particular, $\rho(\gamma^*)$ minimizes the free energy $F$ such that

$$\rho(\gamma^*) = \arg\min_{\rho(\gamma)} F(\rho(\gamma)),$$

where $F(\rho(\gamma)) = \text{Tr}(H\rho(\gamma)) - \beta^{-1}S(\rho(\gamma))$ and $S(\rho(\gamma)) = -\sum_i \lambda_i \log \lambda_i$ refers to the Von-Neumann entropy [26]. To further improve the computational efficiency, Ref. [53] replaces $S(\rho(\gamma))$ with its truncated version, i.e., $S_t(\rho(\gamma)) = \sum_{j<t} C_j \text{Tr}(\rho(\gamma)^{j+1})$ with $C_j$ being certain constants. When $S(\rho(\gamma))$ is estimated by $S_t(\rho(\gamma))$, the quantity used to optimize $\rho(\gamma)$ yields $\arg\min_{\rho(\gamma)} F(\rho(\gamma)) := \text{Tr}(H\rho(\gamma)) - \frac{2\text{Tr}(\rho(\gamma)^2)}{\beta} - \frac{\text{Tr}(\rho(\gamma)^3)}{2\beta^3} + 3$. Enlightened by the fact that $S(\rho(\gamma))$ amounts to the calculation for the eigenvalues of $\rho(\gamma)$, we apply QAE to estimate $S(\rho(\gamma))$. Mathematically, the approximation of the free energy takes the form as

$$\hat{F}(\rho(\gamma)) = \text{Tr}(H\rho(\gamma)) - \beta^{-1}\text{S}(\hat{\rho}(\gamma)),$$

where $S(\hat{\rho}(\gamma))$ refers to the Von-Neumann entropy of the reconstructed state $\hat{\rho}\gamma$ returned by QAE with $S(\hat{\rho}(\gamma)) = -\sum_i \lambda_i \log \lambda_i$. This strategy allows a dramatically reduction of the required quantum resources but the estimation accuracy can be improved. In the optimal case, the spectral information of the compressed state $\sigma^*$, i.e., a set of eigenvalues $\{\lambda_i\}_{i=1}^r$, is sufficient to recover $S(\rho(\gamma))$.

A. Implementation of the QAE-based Gibbs state solver

We now elaborate on the implementation of the proposed QAE-based quantum Gibbs state solver shown in Figure 8. Suppose that the total number of iterations to optimize the free energy $\hat{F}(\rho(\gamma))$ is $T$. The mechanism of our proposal at the $t$-th iteration with $t \in [T]$ is as follows.

1) The variational quantum Gibbs state $\rho(\gamma^{(t)})$ prepared by the variational quantum circuit $U(\gamma^{(t)})$, i.e., $\rho(\gamma^{(t)}) = \text{Tr}_A(U(\gamma^{(t)})(|0\rangle\langle 0|)U(\gamma^{(t)})^\dagger)$, is interacted with $H$ to compute $\text{Tr}(H\rho(\gamma^{(t)}))$, as highlighted in the brown box.
2) QAE is employed to compress the state $\rho(\gamma(t))$ prepared by $U(\gamma(t))$. Once the optimization of QAE is completed, the compressed state $\sigma^T(t)$ is extracted into the classical form followed by the spectral decomposition $\{\tilde{\lambda}_i\}$. This step is highlighted in the blue box.

3) Given access to $S(\rho(\gamma(t)))$ and $\text{Tr}(H\rho(\gamma(t)))$, the classical optimizer can minimize $\mathcal{F}(\rho(\gamma(t)))$. The minimization is achieved by the gradient-based methods, i.e., the update rule at the $t$-th iteration is

$$
\gamma(t+1) = \gamma(t) - \eta \frac{\partial \mathcal{F}(\rho(\gamma(t)))}{\partial \gamma(t)}.
$$

After repeating the above procedures with in total $T$ iterations, the QAE-based quantum Gibbs state solver outputs the optimized free energy $\mathcal{F}(\rho(\gamma(T)))$ and the trained parameters $\gamma(T)$.

We quantify the lower bound of the fidelity between the target state $\rho_G$ and the state $\rho(\gamma)$ prepared by the QAE-based quantum Gibbs state solver in the following lemma. The corresponding proof is provided in Appendix F.

**Lemma 4.** Following notation in Eqn. (24), let $\hat{\gamma}$ be the optimized parameters of the QAE-based quantum Gibbs state solver. Suppose $|\mathcal{F}(\rho(\hat{\gamma})) - \mathcal{F}(\rho_G)| \leq \epsilon_1$ as the discrepancy between the optimized and the original free energy is less than the threshold $\epsilon_1$. Let $\delta$ be the training loss of QAE when compressing the state $\rho(\gamma)$ and $\xi$ be the quantity defined in Eqn. (15). The fidelity between the approximated state $\rho(\gamma)$ returned by QAE-based quantum Gibbs state solver and the target state $\rho_G$ is lower bounded by

$$
\mathcal{F}(\rho_G, \rho(\gamma)) \geq 1 - \sqrt{2\xi \left(1 + \log \frac{2^N - 1}{1 - \sqrt{1 - \delta}}\right) + 2\beta \epsilon_1}.
$$

The achieved results demonstrate that even though the optimization of QAE is not optimal, the collected spectral information of the compressed state can also be used to quantify the Von-Von-Neumann entropy with the estimation guarantee, while the price to pay is decreasing the fidelity between $\rho(\gamma)$ and $\rho_G$. Furthermore, when the inverse temperature $\beta$, the gap $\epsilon_1$, and the training loss of QAE $\delta$ are small, the prepared state $\rho(\gamma)$ can well approximate the target state $\rho_G$ in Eqn. (24).

In this respect, as with prior two protocols, the key to gain computation advantages is exploiting advanced optimization techniques to improve the trainability of QAE and hence the performance of the QAE-based quantum Gibbs state solver.

**B. Numerical simulations**

We apply our proposal to prepare the Gibbs state corresponding to the Ising chain model and evaluate its performance.

**Construction rule of the target state.** The problem Hamiltonian in Eqn. (23) is

$$
H_B = -\sum_{i=1}^{J} Z_i Z_{i+1} + Z_i \text{ is the Pauli-Z matrix on the } i\text{-th qubit.}
$$

The quantum Gibbs state takes the form as $\rho_G = \frac{1}{|U(\theta)|^2} \in \mathbb{C}^{2^J \times 2^J}$. We adopt this construction rule to generate three target Gibbs states with $J = 3$ and $\beta \in \{1, 2, 3, 4\}$.

**Hyper-parameters setting.** We apply the same hyper-parameters setting to learn three Gibbs states. In particular, the variational Gibbs state $\rho(\gamma)$ is prepared by interacting $U(\gamma(t))$ with a 4-qubit state $|0\rangle_A |00\rangle_B$ (the subscripts $A$ and $B$ refer to two subsystems), followed by tracing out the subsystem $A$, i.e., $\rho(\gamma(t)) = \text{Tr}_A(U(\gamma(t))|0\rangle_A |00\rangle_B \langle 0|_A \langle 00|_B) U(\gamma(t))$. The trainable unitary $U(\gamma(t))$ is implemented by the hardware-efficient ansatz in Eqn. (18) with $L = 5$. The total number of iterations to optimize $\mathcal{F}(\rho(\gamma(t)))$ in Eqn. (24) is set as $T = 200$. The parameter shift rule is applied to update $\gamma$ with the learning rate $\eta = 0.2$. In each iteration $t$, QAE is used to acquire the spectral information of $\rho(\gamma(t))$. The hardware-efficient ansatz in Eqn. (18) is exploited to implement the quantum encoder $U(\theta)$ with $L = 4$. The number of latent qubits is set as $K = 2$. The Adam optimizer is used to update $\theta$ with the learning rate $\eta = 0.02$. The number of iterations to optimize $\theta$ is set as 100. The initialized parameters for both $\gamma$ and $\theta$ are uniformly sampled from the interval $[0, 1]$. We repeat each setting with five times to collect the statistical results.

**Simulation results.** The simulation results are illustrated in Figure 8. The outer plot exhibits the average fidelity between the prepared quantum Gibbs state $\rho(\gamma)$ and the target state $\rho_G$, i.e., for all settings $\beta \in \{1, 2, 3, 4\}$, the average fidelity is above 0.97 after 200 iterations. More concisely, the average
fidelity $F(\rho(\gamma), \rho_G)$ converges to 0.972, 0.987, and 0.994 for the setting $\beta = 1.2, 1.5$, and 4, respectively. The ignorable shaded region indicates the robustness of QAE-based Gibbs state solver. The inner plot depicts the top-1 performance of QAE-based Gibbs state solver with respect to $\beta = 1.2, 1.5$, and 4. The corresponding fidelity is 0.983, 0.994, and 0.997, respectively. The achieved high fidelity for all settings reflects the effectiveness of the proposed QAE-based quantum Gibbs state solver. Moreover, the degraded performance with respect to the decreased $\beta$ accords with Lemma [4] i.e., a smaller $\beta$ leads to a worse fidelity bound.

VII. IMPLEMENTATION ON REAL QUANTUM HARDWARE

We now discuss the feasibility of implementing the proposed QAE-based learning protocols on NISQ machines. QAE-based fidelity estimator requires access to quantum resources in Step (i) and Step (iii). In Step (i), QAE is employed to compress the input state, which can be accomplished by NISQ machines due to the flexibility of QAEs. Experimental studies have demonstrated the feasibility of executing QAEs on NISQ machines [14], [17]. In Step (iii) (or the evaluation of $\gamma$ in Eqn. (15)), a quantum computer executes the SWAP test on the explored state $\gamma$ (or $\rho$) and the reconstructed state $|\tilde{\omega}_i\rangle |0\rangle^{\otimes N-K}$ [57]. Various platforms such as photons and superconducting quantum chips have experimentally demonstrated the implementation of SWAP test circuit [71], [72]. To conduct SWAP test, the eigenvector $|\tilde{\omega}_i\rangle$ should be encoded into a $K$-qubit quantum state, which can be efficiently achieved using quantum state preparation methods since the number of latent qubits $K$ is small [26], [73].

We next discuss how to implement QAE-based QFI estimator and QAE-based Gibbs state solver on NISQ machines. The quantum computer plays a similar role in constructing these two learning protocols. First, quantum computer is employed to carry out on a variational quantum circuit, i.e., $U(\gamma)$, to prepare the probe state and the variational Gibbs state, respectively. As with other variational quantum algorithms, we can employ hardware-efficient ansatz to implement $U(\gamma)$, since many experimental studies have demonstrated the effectiveness of hardware-efficient ansatz on NISQ machines [74]. Second, the prepared variational quantum states is fed into QAE to obtain the corresponding spectral information. As explained in the construction of QAE-based fidelity estimator, this procedure can also be efficiently accomplished by NISQ machines. Taken together, QAE-based QFI estimator and Gibbs state solver can be realized on NISQ machines.

VIII. DISCUSSION AND CONCLUSION

In this study, we revisit QAE and demonstrate how to exploit its spectral properties to benefit quantum computation, quantum information, and quantum sensing tasks. We provide rigorous error bounds for the QAE-based learning protocols, which are crucial for utilizing QAE-based methods to solve practical tasks in quantum system learning. Additionally, we conduct extensive simulations to validate the effectiveness of our proposals. Theoretical analysis implies that the performance of the proposed QAE-based learning protocols can be further improved by adopting advanced optimization methods and ansatz to suppress the training loss. Empirical results validate our theoretical results.

There are two promising future research directions to further advance our protocols. The first future research direction is exploring advanced learning strategies that can effectively suppress the training loss [75]–[78]. The second future research direction is exploring effective initialization methods for the QAE-based quantum Fisher estimator and QAE-based quantum Gibbs solver [79]–[81]. Without a suitable initialization, the prepared variational quantum state may become high rank, leading to a significant estimation error in the training process and resulting in inferior performance.

In the end, we would like to point out the potential of applying advanced dimension reduction techniques [82]–[85] to facilitate many other complicated learning problems in the quantum world such as Hamiltonian identification [86] and quantum state reconstruction [87]. Besides, in quantum machine learning [88]–[94], theoretical studies [95], [96] proved that quantum kernels [61], [97]–[99] may achieve better generalization ability than that of classical kernels. Considering that the core of quantum kernels is calculating the similarity of two quantum inputs, which amounts to accomplishing the fidelity estimation, efficient dimension reduction techniques can boost the performance of quantum kernels.

REFERENCES

[1] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, “Quantum machine learning,” Nature, vol. 549, no. 7671, p. 195, 2017. [Online]. Available: https://www.nature.com/articles/nature23474

[2] J. Carrasquilla and R. G. Melko, “Machine learning phases of matter,” Nature Physics, vol. 13, no. 5, p. 431, 2017.

[3] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, “Machine learning for molecular and materials science,” Nature, vol. 559, no. 7715, pp. 547–555, 2018.

[4] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, “Sample-optimal tomography of quantum states,” IEEE Transactions on Information Theory, vol. 63, no. 9, pp. 5628–5641, 2017.

[5] L. Gurvits, “Classical deterministic complexity of edmonds’ problem and quantum entanglement,” in Proceedings of the thirty-fifth annual ACM symposium on Theory of computing, 2003, pp. 10–19.
[6] S. Aaronson and G. N. Rothblum, “Gentle measurement of quantum states and differential privacy,” pp. 322–333, 2019.
[7] S. Aaronson, “Shadow tomography of quantum states,” SIAM Journal on Computing, vol. 49, no. 5, pp. STOC18–368, 2019.
[8] H.-Y. Huang, R. Kueng, and J. Preskill, “Predicting many properties of a quantum system from very few measurements,” Nature Physics, vol. 16, no. 10, pp. 1050–1057, 2020.
[9] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, “Quantum state tomography via compressed sensing,” Physical review letters, vol. 105, no. 15, p. 150401, 2010.
[10] L. A. Rozema, D. H. Mahler, A. Hayat, P. S. Turner, and A. M. Steinberg, “Quantum data compression of a qubit ensemble,” Physical review letters, vol. 113, no. 16, p. 160504, 2014.
[11] A. Shabani, M. Mohseni, S. Lloyd, R. L. Kosut, and H. Rabitz, “Estimation of many-body quantum Hamiltonians via compressed sensing,” Phys. Rev. A, vol. 84, p. 012107, Jul 2011. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.84.012107
[12] J. Romero, J. F. Olson, and A. Aspuru-Guzik, “Quantum autoencoders for efficient compression of quantum data,” Quantum Science and Technology, vol. 2, no. 4, p. 045001, 2017.
[13] C. Cao and X. Wang, “Noise-assisted quantum autoencoder,” Physical Review Applied, vol. 15, no. 5, p. 054012, 2021.
[14] L. Lamata, U. Alvarez-Rodriguez, J. D. Martín-Guerrero, M. Sanz, and E. Solano, “Quantum autoencoders via quantum adders with genetic algorithms,” Quantum Science and Technology, vol. 4, no. 1, p. 014007, 2018.
[15] I. Goodfellow, Y. Bengio, and A. Courville, Deep learning. MIT press, 2016.
[16] J. Preskill, “Quantum computing in the nisq era and beyond.” Quantum, vol. 2, p. 79, 2018.
[17] A. Pepper, N. Tischler, and G. J. Pryde, “Experimental realization of a quantum autoencoder: The compression of qutrits via machine learning,” Physical review letters, vol. 122, no. 6, p. 060501, 2019.
[18] C.-J. Huang, H. Ma, Q. Yin, J.-F. Tang, D. Dong, C. Chen, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, “Realization of a quantum autoencoder for lossless compression of quantum data,” Physical Review A, vol. 102, no. 3, p. 032421, 2020.
[19] D. Bondarenko and P. Feldmann, “Quantum autoencoders to denoise quantum data,” Physical Review Letters, vol. 124, no. 13, p. 130502, 2020.
[20] X.-M. Zhang, W. Kong, M. U. Farooq, M.-H. Yung, G. Guo, and X. Wang, “Generic detection-based error mitigation using quantum autoencoders,” Physical Review A, vol. 103, no. 4, p. L040403, 2021.
[21] C. Bravo-Friero, “Quantum autoencoders with enhanced data encoding,” Machine Learning: Science and Technology, 2021.
[22] V. Gebhart, R. Santagati, A. A. Gentile, E. M. Gauger, D. Craig, N. Ares, L. Banchi, F. Marquardt, L. Pezzè, and C. Bonato, “Learning quantum systems,” Nature Reviews Physics, pp. 1–16, 2023.
[23] V. Dunjko, “Quantum learning unravels quantum system,” Science, vol. 365, p. 6598, pp. 1154–1155, 2022.
[24] M. Cerezo, G. Verdon, H.-Y. Huang, L. Cincio, and P. J. Coles, “Challenges and opportunities in quantum machine learning,” Nature Machine Learning, vol. 2, no. 9, p. 567–576, 2022.
[25] H.-Y. Huang, M. Broughton, J. Cotler, S. Chen, J. Li, M. Mohseni, H. Neven, R. Babbush, R. Kueng, J. Preskill et al., “Quantum advantage in learning from experiments,” Science, vol. 367, no. 6598, pp. 1182–1186, 2022.
[26] M. A. Nielsen and I. L. Chuang, Quantum computation and quantum information. Cambridge University Press, 2010.
[27] J. A. Miszczak, Z. Puchała, P. Horodecki, A. Uhlmann, and K. Życzkowski, “Sub-and super- fidelity as bounds for quantum fidelity,” Quantum Information and Computation, vol. 9, no. 1&2, pp. 0103–0130, 2009.
[28] D. Petz and C. Ghinea, “Introduction to quantum fisher information,” in Quantum probability and related topics. World Scientific, 2011, pp. 261–281.
[29] F. G. Brandao and K. M. Svore, “Quantum speed-ups for solving semidefinite programs,” in 2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS). IEEE, 2017, pp. 415–426.
[30] D. Poulin and P. Wojcik, “Sampling from the thermal quantum Gibbs state and evaluating partition functions with a quantum computer,” Physical review letters, vol. 103, no. 22, p. 220502, 2009.
[31] M. Cerezo, A. Paremba, L. Cincio, and P. J. Coles, “Variational quantum feature estimation,” Quantum, vol. 4, p. 248, 2020.
[32] H.-Y. Huang, R. Kueng, and J. Preskill, “Information-theoretic bounds on quantum advantage in machine learning,” Physical Review Letters, vol. 126, no. 19, p. 190505, 2021.
APPENDIX A
PROOF OF THEOREM 1

Proof of Theorem 1. Note that the loss \( \mathcal{L} \) defined in Eqn. (1) is lower bounded by 0, i.e.,
\[
\mathcal{L}(U(\theta), \rho) = \text{Tr} \left( M_G U(\theta) \rho \right) = \text{Tr} \left( \left( I_N - \langle 0 | 0 \rangle \otimes I_K \right) U(\theta) \rho U(\theta)^\dagger \right) \\
\geq 0,
\]
(28)
since both the measurement operator \( M_G \) and the state \( U(\theta) \rho U(\theta)^\dagger \) are positive semi-definite matrices. Such a result implies that the optimal parameters \( \theta^* \), or equivalently \( U(\theta^*) \), satisfy
\[
\mathcal{L}(U(\theta^*), \rho) = 0.
\]
(29)

We next derive the explicit form of \( U(\theta^*) \) that enables \( \mathcal{L}(U(\theta^*), \rho) = 0 \). Recall that the loss function \( \mathcal{L}(U(\theta), \rho) \) can be rewritten as
\[
\mathcal{L}(U(\theta), \rho) = -\text{Tr} \left( \left( \langle 0 | 0 \rangle \otimes I_K \right) U(\theta) \rho U(\theta)^\dagger \right).
\]
(30)

To obtain \( \mathcal{L}(U(\theta^*), \rho) = 0 \), it necessitates to ensure
\[
\text{Tr} \left( \left( \langle 0 | 0 \rangle \otimes I_K \right) U(\theta^*) \rho U(\theta^*)^\dagger \right) = 1.
\]
(31)

Considering that any unitary operator is full rank and all its singular values equal to 1, without loss of generality, we denote the singular value decomposition of the optimal quantum encoder
\[
U(\theta^*) = \sum_{i=1}^{r^*} |\varphi_i\rangle \langle \varphi_i| + \sum_{i=r^*+1}^{2N} |\varphi_i\rangle \langle \varphi_i|,
\]
(32)
where \( r^* = 2^K \), and \( \{|0\rangle \otimes \varphi_i\} \cup \{|\varphi_i\rangle \} \) is a set of orthonormal basis to form the left (or right) singular vectors of \( U(\theta^*) \).

In conjunction with the above two equations, we obtain
\[
\text{Tr} \left( \left( \langle 0 | 0 \rangle \otimes I_K \right) U(\theta^*) \rho U(\theta^*)^\dagger \right)
\]
\[
= \text{Tr} \left( \sum_{i=1}^{r^*} |\varphi_i\rangle \langle \varphi_i| \rho U(\theta^*)^\dagger \right)
\]
\[
= \text{Tr} \left( \sum_{i=1}^{r^*} |\varphi_i\rangle \langle \varphi_i| \sum_{i=1}^{r^*} |\varphi_i\rangle \langle \varphi_i| \rho \right)
\]
\[
= \text{Tr} \left( \sum_{i=1}^{r^*} |\varphi_i\rangle \langle \varphi_i| \rho \right),
\]
(33)

where the first equality uses the explicit form of \( U(\theta^*) \).

Recall \( \rho = \sum_{i=1}^{2N} \lambda_i |\psi_i\rangle \langle \psi_i| \), where \( \sum_{i=1}^{2N} \lambda_i = 1 \), \( 0 \leq \lambda_i \leq 1 \) for \( \forall i \in [2N] \). Combining the explicit form of \( \rho \) with the last term in Eqn. (33), we obtain
\[
\text{Tr} \left( \sum_{i=1}^{r^*} |\varphi_i\rangle \langle \varphi_i| \rho \right) = \sum_{i=1}^{r^*} \lambda_i \sum_{j=1}^{r^*} |\langle \varphi_j| \psi_i| \rangle|^2.
\]
(34)

Since \( \sum_{i=1}^{r^*} \lambda_i = 1 \) and \( \{|\varphi_i\rangle\} = \{|\psi_i\rangle\} \) is a set of orthonormal basis, the term \( \text{Tr} \left( \sum_{i=1}^{r^*} |\varphi_i\rangle \langle \varphi_i| \rho \right) \) equals to one if and only if \( \sum_{i=1}^{r^*} |\langle \varphi_j| \psi_i| \rangle|^2 = 1 \). Due to the orthonormality of \( \{|\varphi_j\rangle\} \) and \( \{|\psi_i\rangle\} \), such an equality can be achieved when \( |\psi_i\rangle \) for \( \forall i \in [r^*] \) can be linearly spanned by \( \{|\varphi_j\rangle\}_{j=1}^{r^*} \), i.e., for \( \forall i \in [r^*] \),
\[
|\psi_i\rangle = |\varphi_j\rangle, \quad \exists j \in [r^*]
\]
(35)

which ensures \( |\langle \varphi_j| \psi_i| \rangle|^2 = 1 \).

Combining Eqn. (32) and Eqn. (35), the optimal quantum encoder yields the form \( \sum_{i=1}^{r^*} |0\rangle \otimes \varphi_i \langle \varphi_i| \chi^i + \sum_{i=r^*+1}^{2N} |\varphi_i\rangle \langle \varphi_i| \psi_i | \). Since the set of orthonormal basis \( \{|\varphi_i\rangle\} \) is non-unique, there are multiple critical points.

We end the proof by indicating that any other form of \( U(\theta) \) differing with Eqn. (32) does not promise the optimal result, i.e., \( \mathcal{L}(U(\theta), \rho) > 0 \). Specifically, we define such a quantum encoder as
\[
U(\theta) = \sum_{i=1}^{2N} |\varphi_i^\prime\rangle \langle \varphi_i^\prime|,
\]
(36)
where \( \{|\varphi_i^\prime\rangle\} \) is a set of orthonormal basis to form the left (or right) singular vectors of \( U(\theta) \). Notably, to guarantee that a subset orthonormal basis in \( \{|\varphi_i^\prime\rangle\} \) can not be spanned by \( \{|0\rangle \otimes \varphi_i^\prime\} \), or equivalently, \( U(\theta) \neq U(\theta^*) \), we impose a restriction on \( U(\theta) \) such that
\[
\sum_{j=1}^{2N} |\varphi_j^\prime\rangle \langle \varphi_j^\prime| \leq \sum_{j=1}^{2N} \langle 0 \otimes |\varphi_j^\prime\rangle | \chi^j + \sum_{k=1}^{2N} |\varphi_k^\prime\rangle \langle \varphi_k^\prime| = \sum_{j=1}^{r^*} |\varphi_j^\prime\rangle \langle \varphi_j^\prime|,
\]
(37)

The restriction in Eqn. (37) ensures the following result, i.e.,
\[
\begin{align*}
U(\theta)^\dagger \langle 0 | 0 \rangle \otimes I_K U(\theta)
\end{align*}
\]
\[
= U(\theta)^\dagger \langle 0 | 0 \rangle \otimes I_K U(\theta)
\]
\[
= U(\theta)^\dagger \langle 0 | 0 \rangle \otimes I_K U(\theta)
\]
\[
= \sum_{j=1}^{2N} |\varphi_j^\prime\rangle \langle \varphi_j^\prime| \leq \sum_{j=1}^{2N} \langle 0 \otimes |\varphi_j^\prime\rangle | \chi^j + \sum_{k=1}^{2N} |\varphi_k^\prime\rangle \langle \varphi_k^\prime| = \sum_{j=1}^{r^*} |\varphi_j^\prime\rangle \langle \varphi_j^\prime|.
\]
(38)

where the first equality uses the property of the projector with \( \Pi^2 = \Pi \), the first inequality employs Eqn. (37), and the last equality uses the orthonormality of \( \{|\varphi_i^\prime\rangle\} \).

Connecting Eqn. (33) with Eqn. (38), we conclude that when \( U(\theta) \) satisfies the form in Eqn. (36), the following relation holds
\[
\text{Tr} \left( \left( \langle 0 | 0 \rangle \otimes I_K \right) U(\theta) \rho U(\theta)^\dagger \right)
\]
\[
= \text{Tr} \left( \sum_{i=1}^{r^*} |\varphi_i^\prime\rangle \langle \varphi_i^\prime| \rho \right)
\]
\[
\leq 1.
\]
(39)

Such a relation implies \( \mathcal{L}(U(\theta), \rho) > 0 \), according to Eqn. (30). \( \square \)

APPENDIX B
PROOF OF LEMMA 1

Proof of Lemma 1. Let us first derive the explicit form of the compressed state \( \sigma^* \) based on the spectral decomposition of
where the first equality employs the spectral form of $U$ with the optimal unitary $U$ we can prepare the input state $|\varphi_i\rangle$, and the second equality uses Eqn. (2) and Eqn. (3); the difference of the fidelity operators $A$ and $B$ yields $F(\rho, \kappa) = Tr(\rho A \kappa B)$, where $\kappa$ is the reconstructed state of QAE $\hat{\rho}$ (i.e., $\|\rho - \hat{\rho}\|_1/2$), and the error quantity $\varsigma$ in Eqn. (45) yields the following relation

$$1 - \sqrt{1 - \delta} \leq 1 - F(\rho, \hat{\rho}) \leq 1 - F(\rho, \kappa) = \varsigma.$$  

where the first inequality uses the result of Lemma 5 [26] and the second inequality exploits $1 - Tr(\sqrt{\rho} \sqrt{\hat{\rho}}) \leq 1/2 Tr(\|\rho - \hat{\rho}\|_1)$ supported by Lemma 2.

Then, according to the Fuchs-van de Graaf’s inequality [26], we have

$$\frac{1}{2} \text{Tr}(\|\rho - \hat{\rho}\|_1) \leq \sqrt{1 - F^2(\rho, \hat{\rho})} = \varsigma. \quad (44)$$

In conjunction with the above results, we obtain that the fidelity $F(\rho, \kappa)$ is bounded by the estimated fidelity with an additive error term, i.e.,

$$\hat{F}(\rho, \kappa) - \sqrt{\varsigma} \leq F(\rho, \kappa) \leq \hat{F}(\rho, \kappa) + \sqrt{\varsigma}. \quad (45)$$

\[\square\]

\section*{Appendix D}

\section*{On the relation of the training loss $\delta$ and the quantity $\varsigma$}

In this section, we interpret the training loss of QAE $\mathcal{L}$ in Eqn. (1), the trace distance between the target state $\rho$ and the reconstructed state of QAE $\hat{\rho}$ (i.e., $\|\rho - \hat{\rho}\|_1/2$), and the error quantity $\varsigma$ in Eqn. (15) yields the following relation

$$1 - \sqrt{1 - \delta} \leq 1 - F(\rho, \hat{\rho}) \leq 1 - F(\rho, \kappa) = \varsigma. \quad (46)$$

This relation can be attained by separately showing (i) $\|\rho - \hat{\rho}\|_1/2 \leq \varsigma$; (ii) $1 - \sqrt{1 - \delta} \leq \|\rho - \hat{\rho}\|_1/2$. (i) $\|\rho - \hat{\rho}\|_1/2 \leq \varsigma$. This relation has been proved in Eqn. (44). (ii) $\delta \leq \|\rho - \hat{\rho}\|_1$. To achieve this relation, we only need to illustrate that the training loss of QAE is upper bounded by the infidelity of the target state and the reconstructed state, i.e.,

$$1 - \sqrt{1 - \delta} \leq 1 - F(\rho, \hat{\rho}) \leq 1 - F(\rho, \kappa) \quad (47)$$

Then, according to the Fuchs-van de Graaf’s inequality [26] with

$$1 - F(\rho, \hat{\rho}) \leq \frac{1}{2} \|\rho - \hat{\rho}\|_1, \quad (48)$$

it immediately obtains $\delta \leq \|\rho - \hat{\rho}\|_1/2$.

The inequality in Eqn. (47) is hold due to the monotonicity of fidelity. Note that the training loss of QAE can be reformulated as

$$\mathcal{L}(U(\theta), \rho) = \text{Tr} \left( (I_N - (|0\rangle \langle 0|) \otimes \mathbb{1}_K) U(\theta) \rho U(\theta)^\dagger \right)$$

$$= 1 - \text{Tr} \left( |0\rangle \langle 0| \otimes \mathbb{1}_K \right) \text{Tr}_E(U(\theta) \rho U(\theta)^\dagger)$$

$$= 1 - F^2(\|0\rangle \langle 0\| \otimes \mathbb{1}_K, U(\theta) \rho U(\theta)^\dagger) \quad (49)$$

where the second equality uses the property of partial trace and the operation $\text{Tr}_E(C)$ follows Eqn. (10), which refers to the partial trace on $K$ latent qubits, and the third inequality

\[\square\]

\section*{Appendix C}

\section*{Proof of Theorem 2}

Two results used in the proof of Theorem 2 are as follows.

\begin{lemma}[Lemma 1, [31]] For any positive semi-definite operators $A$, $B$, and $C$, where $C$ is normalized, i.e., $\text{Tr}(C) = 1$, the difference of the fidelity $F(A, C)$ and the fidelity $F(B, C)$ yields

$$|F(A, C) - F(B, C)| = |\sqrt{A} \sqrt{C} - \sqrt{B} \sqrt{C}|_1 \leq \sqrt{\text{Tr} \left( (\sqrt{A} - \sqrt{B})^2 \right)} = \sqrt{2(1 - \text{Tr} \left( \sqrt{A} \sqrt{B} \right))}. \quad (42)$$

We are now ready to leverage Lemma 5 to prove Theorem 2.

\begin{proof}[Proof of Theorem 2] Here we evaluate the discrepancy between $\hat{F}(\rho, \kappa)$ and $F(\rho, \kappa)$. In particular, the difference $|\hat{F}(\rho, \kappa) - F(\rho, \kappa)|$ yields

$$\left| \hat{F}(\rho, \kappa) - F(\rho, \kappa) \right| \leq \sqrt{2 \left( 1 - \text{Tr} \left( \sqrt{\rho} \sqrt{\hat{\rho}} \right) \right)} \leq \sqrt{\text{Tr}(\|\rho - \hat{\rho}\|_1)}, \quad (43)$$

where the first inequality uses the result of Lemma 5 and the second inequality exploits $1 - \text{Tr}(\sqrt{\rho} \sqrt{\hat{\rho}}) \leq 1/2 \text{Tr}(\|\rho - \hat{\rho}\|_1)$ supported by Lemma 2.

Then, according to the Fuchs-van de Graaf’s inequality [26], we have

$$\frac{1}{2} \text{Tr}(\|\rho - \hat{\rho}\|_1) \leq \sqrt{1 - F^2(\rho, \hat{\rho})} = \varsigma. \quad (44)$$

In conjunction with the above results, we obtain that the fidelity $F(\rho, \kappa)$ is bounded by the estimated fidelity with an additive error term, i.e.,

$$\hat{F}(\rho, \kappa) - \sqrt{\varsigma} \leq F(\rho, \kappa) \leq \hat{F}(\rho, \kappa) + \sqrt{\varsigma}. \quad (45)$$

\[\square\]
employs the definition of fidelity when one state is the pure state. In other words, we have

$$F((|0\rangle \langle 0|) \otimes (N-K), \text{Tr}_E(U(\theta)\rho U(\theta)^\dagger)) = \sqrt{1 - \delta}.$$ \hspace{1cm} (50)

Moreover, the fidelity between the reconstructed state and the target state satisfies the following relation, i.e.,

$$1 - F(\rho, \hat{\rho}) = 1 - F\left(\rho, \left(U(\theta)\rho U(\theta)^\dagger \right) \otimes \text{Tr}_E(U(\theta)\rho U(\theta)^\dagger)\right)$$

$$= 1 - F\left(\left(|0\rangle \langle 0| \otimes N-K\right) \otimes \text{Tr}_E(U(\theta)\rho U(\theta)^\dagger), U(\theta)\rho U(\theta)^\dagger\right)$$

$$\geq F\left(|0\rangle \langle 0| \otimes (N-K), \text{Tr}_E(\rho U(\theta)^\dagger)\right),$$

where the first equality follows the explicit form of the reconstructed state, the second equation uses the invariant property of fidelity with respect to an isometry, and the inequality comes from the monotonicity of fidelity as $F(\rho, \sigma) \leq F(N(\rho), N(\sigma))$ with $N$ being a feasible quantum channel.

In conjunction with the above two equations, it is evident that

$$1 - F(\rho, \hat{\rho}) \geq 1 - \sqrt{1 - \delta}.$$ \hspace{1cm} (51)

**APPENDIX E**

**PROOF OF LEMMA 3**

**Proof of Lemma 3** The results shown in Lemma 3 can be easily obtained by leveraging Theorem 2. Recall the explicit form of the estimated QFI, i.e.,

$$\hat{I}_\tau(\vartheta, \rho_0(\gamma)) = \frac{1}{2} - F(\rho_0(\gamma), \rho_{0+\tau}(\gamma)).$$ \hspace{1cm} (52)

Combining the above equation with Eqn. (51), we have

$$\hat{I}_\tau(\vartheta, \rho_0(\gamma)) \leq I_\tau(\vartheta, \rho_0(\gamma)) + \frac{8\sqrt{\frac{\delta}{\tau^2}}}{\tau^2}$$

$$\hat{I}_\tau(\vartheta, \rho_0(\gamma)) \geq I_\tau(\vartheta, \rho_0(\gamma)) - \frac{8\sqrt{\frac{\delta}{\tau^2}}}{\tau^2}.$$ \hspace{1cm} (53)

After simplification, we have

$$\hat{I}_\tau(\vartheta, \rho_0(\gamma)) - \frac{8\sqrt{\frac{\delta}{\tau^2}}}{\tau^2} \leq I_\tau(\vartheta, \rho_0(\gamma)) \leq \hat{I}_\tau(\vartheta, \rho_0(\gamma)) + \frac{8\sqrt{\frac{\delta}{\tau^2}}}{\tau^2}.$$

**APPENDIX F**

**PROOF OF LEMMA 4**

Before presenting the proof of Lemma 4, let us first present a necessary lemma, which connects the relative entropy with the fidelity.

**Lemma 6** (Lemma 2, [53]). Given quantum states $\rho$ and $\kappa$ and a constant $\epsilon_2$, suppose that the relative entropy $S(\rho||\kappa)$ is less than $\epsilon_2$. Then the fidelity between $\rho$ and $\kappa$ satisfies

$$F(\rho, \kappa) \geq 1 - \sqrt{2S(\rho||\kappa)} \geq 1 - \sqrt{2\epsilon_2}.$$ \hspace{1cm} (54)

We next present the proof of Lemma 4.

**Proof of Lemma 4** Supported by Lemma 6, we have

$$F(\rho_G, \rho(\gamma)) \geq 1 - \sqrt{2S(\rho(\gamma)||\rho_G)}.$$ \hspace{1cm} (55)