

Transition role of entangled data in quantum machine learning

Received: 21 June 2023

Accepted: 17 April 2024

Published online: 02 May 2024

Xinbiao Wang^{1,2,3}, Yuxuan Du^{3,4}✉, Zhuozhuo Tu⁵, Yong Luo^{1,2}✉, Xiao Yuan^{6,7} & Dacheng Tao⁴✉

Entanglement serves as the resource to empower quantum computing. Recent progress has highlighted its positive impact on learning quantum dynamics, wherein the integration of entanglement into quantum operations or measurements of quantum machine learning (QML) models leads to substantial reductions in training data size, surpassing a specified prediction error threshold. However, an analytical understanding of how the entanglement degree in data affects model performance remains elusive. In this study, we address this knowledge gap by establishing a quantum no-free-lunch (NFL) theorem for learning quantum dynamics using entangled data. Contrary to previous findings, we prove that the impact of entangled data on prediction error exhibits a dual effect, depending on the number of permitted measurements. With a sufficient number of measurements, increasing the entanglement of training data consistently reduces the prediction error or decreases the required size of the training data to achieve the same prediction error. Conversely, when few measurements are allowed, employing highly entangled data could lead to an increased prediction error. The achieved results provide critical guidance for designing advanced QML protocols, especially for those tailored for execution on early-stage quantum computers with limited access to quantum resources.

Quantum entanglement, an extraordinary characteristic of the quantum realm, drives the superiority of quantum computers beyond classical computers¹. Over the past decade, diverse quantum algorithms leveraging entanglement have been designed to advance cryptography^{2,3} and optimization^{4–8}, delivering runtime speedups over classical approaches. Motivated by the exceptional abilities of quantum computers and the astonishing success in machine learning, a nascent frontier known as quantum machine learning (QML) has emerged^{9–15}, seeking to outperform classical models in specific learning tasks^{16–25}. Substantial progress has been made in this field, exemplified by the introduction of QML

protocols that offer provable advantages in terms of query or sample complexity for learning quantum dynamics^{26–31}, as a fundamental problem toward understanding the laws of nature. Most of these protocols share a common strategy to gain advantages: the incorporation of entanglement into quantum operations and measurements, leading to reduced complexity. Nevertheless, an overlooked aspect in prior works is the impact of incorporating entanglement in quantum input states, or entangled data, on the advancement of QML in learning quantum dynamics. Due to the paramount role of data in learning^{32–37} as well as entanglement in quantum computing, addressing this question will significantly

¹Institute of Artificial Intelligence, School of Computer Science, Wuhan University, Hubei 430072, China. ²National Engineering Research Center for Multimedia Software, Wuhan University, Hubei 430072, China. ³JD Explore Academy, Beijing 101111, China. ⁴School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798, Singapore. ⁵School of Computer Science, Faculty of Engineering, University of Sydney, Sydney, NSW 2008, Australia. ⁶Center on Frontiers of Computing Studies, Peking University, Beijing 100871, China. ⁷School of Computer Science, Peking University, Beijing 100871, China. ✉e-mail: duyuxuan123@gmail.com; yluo180@gmail.com; dacheng.tao@ntu.edu.sg

enhance our comprehension of the capabilities and limitations of QML models.

A fundamental concept in machine learning that characterizes the capabilities of learning models in relation to datasets is the no-free-lunch (NFL) theorem^{38–41}. The NFL theorem yields a key insight: regardless of the optimization strategy employed, the ultimate performance of models is contingent upon the size and types of training data. This observation has spurred recent breakthroughs in large language models, as extensive and meticulously curated training data consistently yield superior results^{42–46}. In this regard, establishing the quantum NFL theorem enables us to elucidate the specific impact of entangled data on the efficacy of QML models in learning quantum dynamics. Concretely, the achieved theorem can shed light on whether the utilization of entangled data empowers QML models to achieve comparable or even superior performance compared to low-entangled or unentangled data, while simultaneously reducing the sample complexity required. Although initial attempts^{47,48,49} have been made to establish quantum NFL theorems, they have relied on infinite query complexity, thus failing to address our concerns adequately (see Supplementary Note 1 and Supplementary Note 2 for details). Building upon prior findings on the role of entanglement and the classical NFL theorem, a reasonable speculation is that high-entangled data contributes to the improved performance of QML models associated with the reduced sample complexity, albeit at the cost of using extensive quantum resources to prepare such data that may be unaffordable in the early stages of quantum computing⁵⁰.

In this study, we *negate* the above speculation and exhibit the *transition role* of entangled data when QML models incoherently learn quantum dynamics, as shown in Fig. 1. In the incoherent learning scenario, the quantum learner is restricted to utilizing datasets with varying degrees of entanglement to operate on an unknown unitary and inferring its dynamics using the finite measurement outcomes collected under the projective measurement, differing from ref. 48 in learning problems and training data. The entangled data refers to quantum states that are entangled with a reference system, with the degree of entanglement quantitatively characterized by the Schmidt rank r . We rigorously show that within the context of NFL, the entangled data has a *dual effect* on the prediction error according to the number of measurements m allowed. Particularly, with sufficiently

large m , increasing r can consistently reduce the required size of training data for achieving the same prediction error. On the other hand, when m is small, the train data with large r not only requires a significant volume of quantum resources for states preparation, but also amplifies the prediction error. As a byproduct, we prove that the lower bound of the query complexity for achieving a sufficiently small prediction error matches the optimal lower bound for quantum state tomography with nonadaptive measurements. To cover a more generic learning setting, we consider the problem of dynamic learning under arbitrary observable by using ℓ -outcome positive-operator valued measure (POVM) to collect the measurement output. This setting covers the shadow-based learning models^{26,51,52}. We show that the transition role still holds for arbitrary POVM and increasing the possible outcomes ℓ could significantly reduce the query complexity. Numerical simulations are conducted to support our theoretical findings. In contrast to the previous understanding that entanglement mostly confers benefits to QML in terms of sample complexity, the transition role of entanglement identified in this work deepens our comprehension of the relation between quantum information and QML, which facilitates the design of QML models with provable advantages.

Results

We first recap the task of learning quantum dynamics. Let $U \in \text{SU}(2^n)$ be the target unitary and $O \in \mathbb{C}^{2^n \times 2^n}$ be the observable which is a Hermitian matrix acting on an n -qubit quantum system. Here we specify the observable as the projective measurement $O = |\mathbf{o}\rangle\langle\mathbf{o}|$ since any observable reads out the classical information from the quantum system via their eigenvectors. The goal of the quantum dynamics learning is to predict the functions of the form

$$f_U(\psi) = \text{Tr}(OU|\psi\rangle\langle\psi|U^\dagger), \tag{1}$$

where $|\psi\rangle$ is an n -qubit quantum state living in a 2^n -dimensional Hilbert space \mathcal{H}_X . This task can be done by employing the training data S to construct a unitary V_S , i.e., the learned hypothesis has the form of $h_S(\psi) = \text{Tr}(OV_S|\psi\rangle\langle\psi|V_S^\dagger)$, which is expected to accurately approximate $f_U(\psi)$ for the unseen data. While the learned unitary acts on an n -qubit system \mathcal{H}_X , the input state could be entangled with a reference

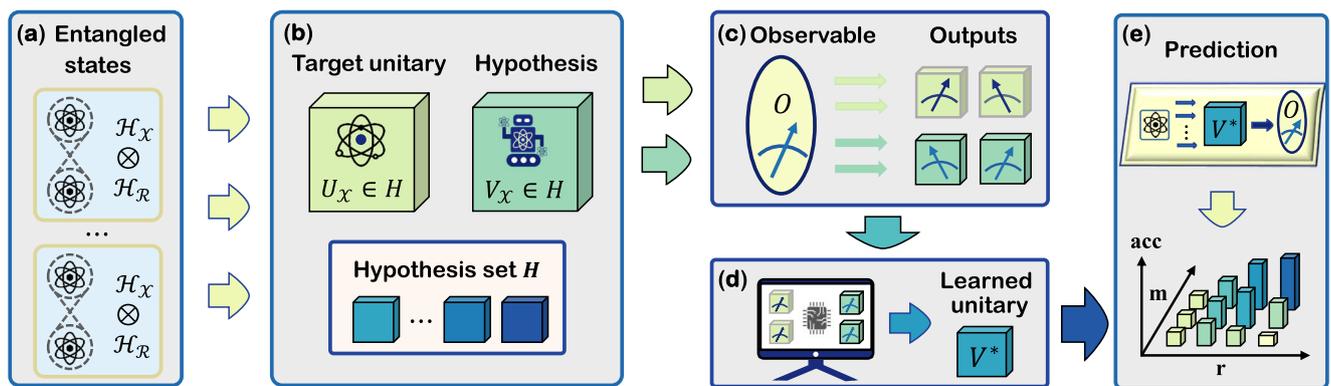


Fig. 1 | Illustration of quantum NFL setting with the entangled data. The goal of the quantum learner is to learn a unitary V_X that can accurately predict the output of the target unitary U_X under a fixed observable O , where the subscript X refers to the quantum system in which the operator O act on. The learning process is as follows. **a** A total number of N entangled bipartite quantum states living in Hilbert space $\mathcal{H}_X \otimes \mathcal{H}_R$ (R denotes the reference system) are taken as inputs, dubbed *entangled data*. **b** Quantum learner proceeds incoherent learning. The entangled data *separately* interacts with the target unitary U_X (agnostic) and the candidate hypothesis V_X extracted from the same Hypothesis set H . **c** The quantum learner is restricted to leverage the finite measured outcomes of the observable O on the

output states of U_X and V_X to conduct learning. **d** A classical computer is exploited to infer V that best estimates U_X according to the measurement outcomes. For example, in the case of variational quantum algorithms, the classical computer serves as an optimizer to update the tunable parameters of the ansatz V_X . **e** The learned unitary V is used to predict the output of unseen quantum states in Hilbert space \mathcal{H}_X under the evolution of the target unitary U_X and the measurement of O . A large Schmidt rank r can enhance the prediction accuracy when combined with a large number of measurements m , but may lead to a decrease in accuracy when m is small.

system \mathcal{H}_R , i.e., $|\psi\rangle \in \mathcal{H}_X \otimes \mathcal{H}_R$. We suppose that all input states have the same Schmidt rank $r \in \{1, \dots, 2^n\}$. Then the response of the state $|\psi_j\rangle$ is given by the measurement output $\mathbf{o}_j = \sum_{k=1}^m \mathbf{o}_{jk}/m$, where m is the number of measurements and \mathbf{o}_{jk} is the output of the k -th measurement of the observable \mathbf{O} on the output quantum state $(\mathbf{U} \otimes \mathbb{I}_R)|\psi_j\rangle$. In this manner, the training data with N examples takes the form $S = \{(|\psi_j\rangle, \mathbf{o}_j) : |\psi_j\rangle \in \mathcal{H}_X \otimes \mathcal{H}_R, \mathbb{E}[\mathbf{o}_j] = \mathbf{u}_j\}_{j=1}^N$ with $\mathbf{u}_j = \text{Tr}((\mathbf{U}^\dagger \mathbf{O} \mathbf{U} \otimes \mathbb{I}_R)|\psi_j\rangle\langle\psi_j|)$ being the expectation value of the observable \mathbf{O} on the state $(\mathbf{U} \otimes \mathbb{I}_R)|\psi_j\rangle$ and N being the size of the training data. Notably, in quantum dynamics learning, sample complexity refers to the size of training data N , or equivalently, the number of quantum states in the training data; query complexity refers to the total number of queries of the explored quantum system, i.e., the production of sample complexity and the number of measurements Nm .

The risk function is a crucial measure in statistical learning theory to quantify how well the hypothesis function h_S performs in predicting f_U , defined as

$$R_U(\mathbf{V}_S) = \int d\psi (f_U(\psi) - h_S(\psi))^2, \tag{2}$$

where the integral is over the uniform Haar measure $d\psi$ on the state space. Intuitively, $R_U(\mathbf{V}_S)$ amounts to the average square error distance between the true output $f(\psi)$ and the hypothesis output $h_S(\psi)$. Moreover, we follow the treatments in ref. 48 choosing the Haar unitary as the target unitary. Additionally, we construct a sampling rule of the training input states which approximates the uniform distribution of all entangled states with Schmidt rank r (refer to Supplementary Note 2).

Under the above setting, we prove the following quantum NFL theorem in learning quantum dynamics, where the formal statement and proof are deferred to Supplementary Note 3.

Theorem 1. (Quantum NFL theorem in learning quantum dynamics, informal). Following the settings in Eq. (1), suppose that the training error of the learned hypothesis on the training data S is less than $\varepsilon = \mathcal{O}(1/2^n)$. Then the lower bound of the averaged prediction error in Eq. (2) yields

$$\mathbb{E}_{U,S} R_U(\mathbf{V}_S) \geq \Omega\left(\frac{\tilde{\varepsilon}^2}{4^n} \left(1 - \frac{N \cdot \min\{m/(2^n r c_1), rn\}}{2^n c_2}\right)\right),$$

where $c_1 = 128/\tilde{\varepsilon}^2$, $c_2 = \min\{(1 - 2\tilde{\varepsilon})^2, (64\tilde{\varepsilon}^2 - 1)^2\}$, $\tilde{\varepsilon} = \Theta(2^n \varepsilon)$, and the expectation is taken over all target unitary \mathbf{U} , entangled states $|\psi_j\rangle$ and measurement outputs \mathbf{o}_j .

The achieved results indicate the transition role of the entangled data in determining the prediction error. Particularly, when a sufficient number of measurements m is allowed such that the Schmidt rank r obeys $r < \sqrt{m/(c_1 2^n n)}$, the minimum term in the achieved lower bound refers to Nrn and hence increasing r can constantly decrease the prediction error. Accordingly, in the two extreme cases of $r=1$ and $r=2^n$, achieving zero averaged risk requires $N = 2^n c_2/n$ and $N=1$ training input states, where the latter achieves an exponential reduction in the number of training data compared with the former. This observation implies that the entangled data empower QML with provable quantum advantage, which accords with the achieved results of ref. 48 in the ideal coherent learning protocol with infinite measurements.

By contrast, in the scenario with $r \geq \sqrt{m/(c_1 2^n n)}$, increasing r could enlarge the prediction error. This result indicates that the entangled data can be harmful to achieving quantum advantages, which contrasts with previous results where the entanglement (e.g., entangled operations or measurements) is believed to contribute to the quantum advantage^{48,53–55}. This counterintuitive phenomenon stems from the fact that when incoherently learning quantum

dynamics, information obtained from each measurement decreases with the increased r and hence a small m is incapable of extracting all information of the target unitary carried by the entangled state.

Another implication of Theorem 1 is that although the number of measurements m contributes to a small prediction error, it is *not decisive to the ultimate performance* of the prediction error. Specifically, when $m \geq r^2 c_2 2^n n$, further increasing m could not help decrease the prediction error which is determined by the entanglement and the size of the training data, i.e., r and N . Meanwhile, at least $r^2 c_2 2^n n$ measurements are required to fully utilize the power of entangled data. These results suggest that the value of m should be adaptive to r to pursue a low prediction error.

We next comprehend the scenario in which the lower bound of averaged risk in Theorem 1 reaches zero and correlate with the results in quantum state learning and quantum dynamics learning^{26,27,29,30,56,57}. In particular, the main focus of those studies is proving the minimum query complexity of the target unitary to warrant zero risk. The results in Theorem 1 indicate that the minimum query complexity is $Nm = \Omega(4^n r c_1 c_2)$, implying the proportional relation between the entanglement degree r and the query complexity. Notably, *this lower bound is tighter than that achieved in ref. 26* in the same setting. The achieved results in terms of query complexity are also non-trivial, as previous works show that query complexity can benefit from using entanglement in quantum data^{58,59} and quantum measurements^{26,30}. The advance of our results stems from the fact that ref. 26 simply employs Holevo’s theorem to give an upper bound on the extracted information in a single measurement, while our bound integrates more refined analysis such as the consideration of Schmidt rank r , the direct use of a connection between the mutual information of the target unitary \mathbf{U} and the measurement outputs \mathbf{o}_j , and the KL-divergence of related distributions (refer to Supplementary Note 3 for more details). Moreover, the adopted projective measurement \mathbf{O} in Eqn. (1) hints that the learning task explored in our study amounts to learning a pure state $\mathbf{U}\mathbf{O}\mathbf{U}$. From the perspective of state learning, *the derived lower bound in Theorem 1 is optimal for the nonadaptive measurement with a constant number of outcomes*⁶⁰. Taken together, while the entangled data hold the promise of gaining advantages in terms of the sample complexity for achieving the same level of prediction error, they may be inferior to the training data without entanglement in terms of query complexity.

The transition role of entanglement explained above leads to the following construction rule of quantum learning models. First, when a large number of measurements is allowed, the entangled data is encouraged to be used for improving the prediction performance. To this end, initial research efforts^{61–66}, which develop effective methods for preparing and storing entangled states, may contribute to QML. Second, when the total number of measurements is limited, it is advised to refrain from using entangled data for learning quantum dynamics.

Remark. (i) The training error scaling $\varepsilon = \mathcal{O}(1/2^n)$ in Theorem 1 and the factor of the achieved lower bound $\tilde{\varepsilon}^2/4^n$ comes from the consideration of average performance over Haar unitaries where the expectation value of observable \mathbf{O} scales as $\text{Tr}(\mathbf{O})/2^n$ (Refer to Supplementary Note 2). (ii) The results of the transition role for entangled data achieved in Theorem 1 can be generalized to the mixed states because the mixed state can be produced by taking the partial trace of a pure entangled state.

In a more generic learning setting, the observable used in the target function defined in Eqn. (1) and the measurement used for collecting the response of training data \mathbf{o} could be arbitrary and varied. In particular, we consider that the observable \mathbf{O} defined in Eqn. (1) could be arbitrary Hermitian operator satisfying $\|\mathbf{O}\|_1 \leq \infty$. The response \mathbf{a}_j for given input states $|\psi_j\rangle$ could be obtained from measuring the output states on system \mathcal{X} with ℓ -outcome POVM. The training dataset in this case refers to $S_\ell = \{(|\psi_j\rangle, \mathbf{a}_j) : |\psi_j\rangle \in \mathcal{H}_{\mathcal{X}\mathcal{R}}, \mathbf{a}_j = (\mathbf{a}_{j1}, \dots, \mathbf{a}_{jm}), \mathbf{a}_{jk} \in \{z_1, \dots, z_\ell\}\}_{j=1}^N$,

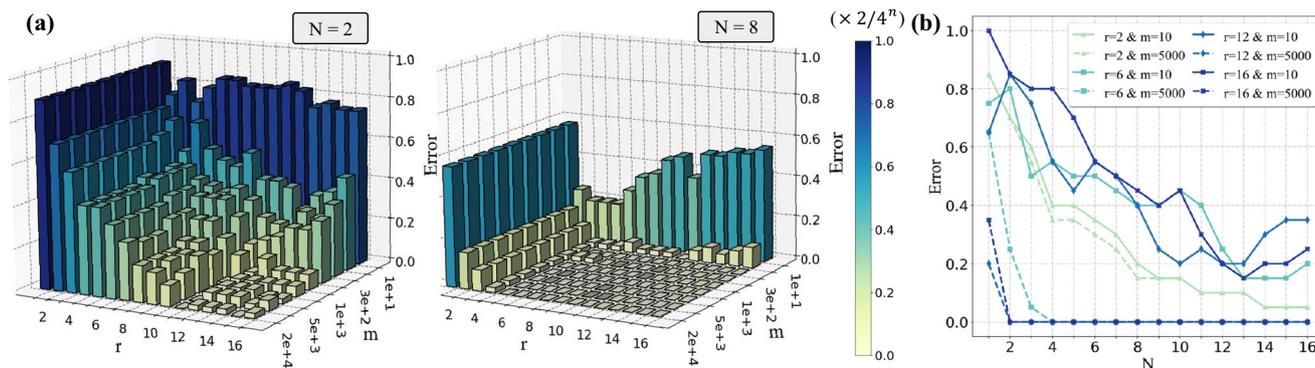


Fig. 2 | Simulation results of quantum NFL theorem when incoherently learning quantum dynamics. **a** The averaged prediction error with a varied number of measurements m and Schmidt rank r when $N=2$ and $N=8$. The z-axis refers to the averaged prediction error defined in Eq. (1). **b** The averaged prediction error with

the varied sizes of training data. The label “ $r = a \& m = b$ ” refers that the Schmidt rank is a and the number of measurements is b . The label “ $(\times 2/4^n)$ ” refers that the plotted prediction error is normalized by a multiplier factor $2/4^n$.

where $|\psi_j\rangle$ refers to the entangled states with Schmidt rank r , \mathbf{a}_j is the m -measurement outputs with ℓ -outcome POVM, and $\{z_i\}_{i=1}^\ell$ is the ℓ possible outcomes of the employed POVM. In this case, denoting the learned unitary as \mathbf{V}_{S_ℓ} , we get the following quantum NFL theorem in learning quantum dynamics for generic measurements, where the formal statement and proof are deferred to Supplementary Note 4.

Theorem 2. (Quantum NFL theorem in learning quantum dynamics for generic measurements, informal) Following the settings in Eq. (1) with arbitrary \mathbf{O} satisfying $\|\mathbf{O}\|_1 \leq \infty$, suppose the learned hypothesis is learned from training data S_ℓ . Then the lower bound of the averaged prediction error in Eqn. (2) yields

$$\mathbb{E}_{U, S_\ell} R_U(\mathbf{V}_{S_\ell}) \geq \varepsilon^2 \left(1 - \frac{N \cdot \min\{4m/r, 6m\ell/2^n r, rn\}}{\log(|\mathcal{X}_{2\varepsilon}(\mathbf{O})|)} \right)$$

where $|\mathcal{X}_{2\varepsilon}(\mathbf{O})|$ refers to the model complexity and only depends on ε and the employed observable \mathbf{O} . For projective measurement $\mathbf{O} = |\mathbf{o}\rangle\langle\mathbf{o}|$, $\log(|\mathcal{X}_{2\varepsilon}(\mathbf{O})|) = 2^n c_2$ is given in the denominator of the achieve lower bound in Theorem 1.

The achieved results in Theorem 2 deliver three implications. First, the transition role of entangled data still holds for arbitrary observable and POVM. In particular, no matter how large the number of possible outcomes of POVM ℓ is, increasing the Schmidt rank will decrease the prediction error as long as the number of measurements m satisfies $\min\{4m/r, 6m\ell/2^n r\} \leq rn$, and increase the prediction error otherwise. Second, when the observable is projective measurement and the number of possible outcomes ℓ is of constant order, the achieved result in Theorem 2 reduces to the results achieved in Theorem 1 for the case of employing projective measurement up to a constant factor. Third, increasing the number of possible outcomes of POVM ℓ can exponentially reduce the number of measurements required to achieve the same level of prediction error. Particularly, considering two extreme cases of the possible outcomes of POVM ℓ being constant scaling $\Theta(1)$ and exponential scaling $\Theta(2^n)$, achieving the same level of prediction error requires the query complexity scaling with the order of $2^n r \log(|\mathcal{X}_{2\varepsilon}(\mathbf{O})|)$ and $r \log(|\mathcal{X}_{2\varepsilon}(\mathbf{O})|)$, where the latter case achieves an exponential reduction in terms of the query complexity.

Numerical results

We conduct numerical simulations to exhibit the transition role of entangled data, the effect of the number of measurements, and the training data size in determining the prediction error. The omitted construction details and results are deferred to Supplementary Note 5.

We focus on the task of learning an n -qubit unitary under a fixed projective measurement $\mathbf{O} = (|\mathbf{o}\rangle\langle\mathbf{o}|)^{\otimes n}$. The number of qubits is $n=4$. The target unitary \mathbf{U}_X is chosen uniformly from a discrete set $\{\mathbf{U}_i\}_{i=1}^M$, where $M=2^n$ refers to the set size and the operators $\mathbf{U}_j^\dagger \mathbf{O} \mathbf{U}_j$ with \mathbf{U}_j in this set are orthogonal such that the operators $\mathbf{U}_j^\dagger \mathbf{O} \mathbf{U}_j$ are well distinguished. The entangled states in S is uniformly sampled from the set $\{(\sum_{j=1}^r \sqrt{c_j} \mathbf{U}_j |\mathbf{o}\rangle) \otimes |\xi_j\rangle | (\sqrt{c_1}, \dots, \sqrt{c_r})^T \in \text{SU}(r), |\xi_j\rangle \in \text{SU}(2^n)\}$. The size of training data is $N \in \{1, 2, \dots, 16\}$ and the Schmidt rank takes $r = \{2^0, \dots, 2^4\}$. The number of measurements takes $m \in \{10, 100, 300, \dots, 5000, 20000\}$. We record the averaged prediction error by learning four different 4-qubit unitaries for 10 training data.

The simulation results are displayed in Fig. 2. Particularly, Fig. 2a shows that for both the cases of $N=2$ and $N=8$, the prediction error constantly decreases with respect to an increased number of measurements m and increased Schmidt rank r when the number of measurements is large enough, namely $m > 1000$. On the other hand, for a small number of measurements with $m \leq 100$ in the case of $N=8$, as the Schmidt rank is continually increased, the averaged prediction error initially decreases and then increases after the Schmidt rank surpasses a critical point which is $r=3$ for $m=10$ and $r=4$ for $m=100$. This phenomenon accords with the theoretical results in Theorem 1 in the sense that the entangled data play a transition role in determining the prediction error for a limited number of measurements. This observation is also verified in Fig. 2b for the varied sizes of training data, where for the small measurement times $m=10$, increasing the Schmidt rank could be not helpful for decreasing the prediction error. By contrast, a large training data size consistently contributes to a small prediction error, which echoes with Theorem 1.

Discussion

In this study, we exploited the effect of the Schmidt rank of entangled data on the performance of learning quantum dynamics with a fixed observable. Within the framework of the quantum NFL theorem, our theoretical findings reveal the transition role of entanglement in determining the ultimate model performance. Specifically, increasing the Schmidt rank below a threshold controlled by the number of measurements can enhance model performance, whereas surpassing this threshold can lead to a deterioration in model performance. Our analysis suggests that a large number of measurements is the precondition to use entangled data to gain potential quantum advantages. In addition, our results demystify the negative role of entangled data in the measure of query complexity. Last, as with the classical NFL theorem, we prove that increasing the size of the training data always contributes to a better performance in QML.

Our results motivate several important issues and questions needed to be further investigated. The first research direction is exploring whether the transition role of entangled data exists for other QML tasks such as learning quantum unitaries or learning quantum channels with the response being measurement output^{26,30,55,67–76}. These questions can be considered in both the coherent and incoherent learning protocols, which are determined by whether the target and model system can coherently interact and whether quantum information can be shared between them. Obtaining such results would have important implications for using QML models to solve practical tasks with provable advantages.

Another research direction is inquiring whether there exists a similar transition role when exploiting entanglement in quantum dynamics and measurements through the use of an ancillary quantum system. The answer for the case of entangled measurement has been given under many specific learning tasks^{26,29,30,77} where the learning protocols with entangled measurements are shown to achieve an exponential advantage over those without in terms of the access times to the target unitary. This quantum advantage arises from the powerful information-extraction capabilities of entangled measurements. In this regard, it is intriguing to investigate the effect of quantum entanglement on model performance when entanglement is introduced in both the training states and measurements, as entangled measurements offer a potential solution to the negative impact of entangled data resulting from insufficient information extraction via weak projective measurements. A positive result could further enhance the quantum advantage gained through entanglement exploitation.

Methods

Here we first outline the proof strategy that establishes the lower bound of the averaged prediction error in Theorem 1. Then we present an intuitive explanation of the transition role of entangled data according to the achieved numerical results.

Proof sketch

The backbone of the proof refers to Fano’s method, which is widely used to derive the lower bound of prediction error in classical learning theory⁷⁸. This method involves the following three parts. Part (I): The space of the target dynamics $\mathcal{U} = \{U \in \text{SU}(d)\}$ is discretized into a 2ε -packing $\mathcal{M}_{2\varepsilon} = \{U_{x'}\}_{x'=1}^{|\mathcal{M}_{2\varepsilon}|}$ such that the dynamics within $\mathcal{M}_{2\varepsilon}$ are sufficiently distinguishable under a distance metric related to the target function in Eq. (1). Part (II): The dynamics learning problem in Eq. (1) is translated to the hypothesis testing problem related to the 2ε -packing $\mathcal{M}_{2\varepsilon}$. Such a hypothesis testing problem amounts to a communication protocol between two parties, namely Alice and Bob. Particularly, Alice chooses an element X of $\{1, \dots, |\mathcal{M}_{2\varepsilon}|\}$ uniformly at random and employs the corresponding unitary U_X to construct the training data $\mathcal{S} = \{|\psi_j\rangle, \mathbf{o}_j^{(x)}\}_{j=1}^N$ with $|\psi_j\rangle$ being the randomly sampled entangled input state and \mathbf{o}_j being the associated measurement output of the state $(U_X \otimes \mathbb{I})|\psi_j\rangle$ under the projective measurement \mathbf{O} . Bob’s goal is to retrieve the information of X from the discrete set $\{1, \dots, |\mathcal{M}_{2\varepsilon}|\}$ based on \mathcal{S} . The inferred index by Bob is denoted by \hat{X} . This leads to the hypothesis testing problem with the null hypothesis $\hat{X} \neq X$. In this regard, we demonstrate that the averaged prediction error is greater than a quantity related to the error probability $\mathbb{P}(\hat{X} \neq X)$ of the hypothesis testing problem, namely $\mathbb{E}_{U, \mathcal{S}} R_U(\mathbf{V}_\mathcal{S}) \geq \mathbb{E}_{X, \mathcal{S}} \varepsilon^2 \mathbb{P}(\hat{X} \neq X)$. This provides the theoretical guarantee of reducing the learning problem to the hypothesis testing problem. Part (III): Fano’s inequality is utilized to establish an upper bound on the error probability $\mathbb{P}(\hat{X} \neq X)$ of the hypothesis testing problem, i.e.,

$$\mathbb{P}(\hat{X} \neq X) \geq 1 - \frac{I(X; \hat{X}) + \log 2}{\log(|\mathcal{M}_{2\varepsilon}|)}, \quad (3)$$

which is dependent on two factors: the cardinality of the 2ε -packing $\mathcal{M}_{2\varepsilon}$, and the mutual information $I(X; \hat{X})$ between the target index X and the estimated index \hat{X} .

To summarize, Fano’s method reduces the challenging problem of lower bounding the prediction error to separately lower bounding the packing cardinality and upper bounding the mutual information, which we could develop techniques for tackling. In particular, we obtain the lower bound of the 2ε -packing cardinality $|\mathcal{M}_{2\varepsilon}|$ by employing the probabilistic argument to show the existence of a large but well-separated collection of quantum dynamics (i.e., the 2ε -packing $\mathcal{M}_{2\varepsilon}$) under a metric dependent on the observable \mathbf{O} .

We establish an upper bound for the mutual information $I(X; \hat{X})$ by considering two cases: one with a small number of measurements and another with a sufficiently large number of measurements. For the former case, the mutual information $I(X; \hat{X})$ is upper bounded by a quantity involving the KL-divergence between the probability distributions of the measurement output $\mathbf{o}_j^{(x)}$ related to various index x , which has the order of $\mathcal{O}(Nmd\varepsilon^2/r)$ in the average case. On the other hand, while the mutual information $I(X; \hat{X})$ cannot grow infinitely with the number of measurements, we derive another upper bound of $I(X; \hat{X})$ with the mutual information $I(X; \{(\mathbf{U}_X \otimes \mathbb{I})|\psi_j\rangle\}_{j=1}^N)$ for the case of a sufficiently large number of measurements which could extract the maximal amount of information from each output state. In this regard, the mutual information is upper bounded by an m -independent quantity with the order of $\mathcal{O}(rn)$. This leads to the final upper bound of the mutual information $\min\{\mathcal{O}(Nmd\varepsilon^2/r), \mathcal{O}(rn)\}$ where the Schmidt rank r plays the opposite role in the two scenarios with the various number of measurements allowed, resulting in the transition role of entangled data.

Taken all together, we can obtain the lower bound of the averaged prediction error in Theorem 1.

Remark. While many similar pipelines involving the utilization of Fano’s inequality have been used for obtaining the lower bound of sample complexity in quantum state learning tasks, we give a detailed explanation about how our results differ from previous studies in Supplementary Note 1E.

Intuitive explanations based on numerical results

We now give an intuitive explanation about the transition role of the entangled data based on the numerical simulations.

Before elucidating, we first detail the construction rule of the target unitaries set and the entangled input states. Particularly, to construct the set consisting of well-distinguished target unitaries under the distance metric related to the observable $\mathbf{O} = (|\mathbf{O}\rangle\langle\mathbf{O}|)^{\otimes n}$, one way is to choose the unitaries U_j in $\text{SU}(2^n)$ at uniformly random such that the target operators $U_j^\dagger \mathbf{O} U_j = |e_j\rangle\langle e_j|$ are mutually orthogonal. In this regard, learning the target unitary under the observable \mathbf{O} is equivalent to identifying the unknown index k corresponding to the target operator $U_k^\dagger \mathbf{O} U_k = |e_k\rangle\langle e_k|$. The entangled input states have the form of $|\psi_j\rangle = \sum_{k=1}^r \sqrt{c_{jk}} |\xi_{jk}\rangle_{\mathcal{X}} |\zeta_{jk}\rangle_{\mathcal{R}}$ where the Schmidt coefficients $\{c_{jk}\}$ satisfy $\sum_{k=1}^r c_{jk} = 1$. As the target unitary U_k acts on the quantum system \mathcal{X} , the identification of the corresponding index k solely depends on the partial trace of the entangled states, i.e., $\sigma_j := \text{Tr}_{\mathcal{R}}(|\psi_j\rangle\langle\psi_j|) = \sum_{k=1}^r c_{jk} |\xi_{jk}\rangle\langle\xi_{jk}|_{\mathcal{X}}$. To this end, we consider that the states $|\xi_{jk}\rangle$ in the reduced states are sampled from the computational basis $\{|e_i\rangle\}_{i=1}^{2^n}$ and the coefficient vector $\mathbf{c}_j = (\sqrt{c_{j1}}, \dots, \sqrt{c_{jr}})$ is sampled from the Haar distribution in the r -dimensional Hilbert space \mathcal{H}_r . In this manner, we construct $\mathcal{S} = \{|\psi_j\rangle, \mathbf{o}_j\}_{j=1}^N$ and use it to learn the unknown index k by solving the following minimization

problem

$$\hat{k} = \arg \min_{k \in [2^n]} \sum_{j=1}^N (\mathbf{o}_j^{(k)} - \mathbf{o}_j)^2, \quad (4)$$

where $[2^n]$ refers to the set $\{1, \dots, 2^n\}$ and $\mathbf{o}_j^{(k)}$ is the collected measurement outputs by applying the observable $\mathbf{U}_k^\dagger \mathbf{O} \mathbf{U}_k$ with $\mathbf{U}_k \in \{\mathbf{U}_{k'}\}_{k' \in [2^n]}$ to input states $|\psi_j\rangle$.

The successful identification of the target index k^* relies on the satisfaction of two key conditions:

1. The states set $\{|\xi_{ji}\rangle\}_{j,i=1}^{N,r}$ contains the target operator $\mathbf{U}_{k^*}^\dagger \mathbf{O} \mathbf{U}_{k^*} = |\mathbf{e}_{k^*}\rangle\langle\mathbf{e}_{k^*}|$.
2. The measurement outputs $\{\mathbf{o}_j\}_{j=1}^N$ closely approximate the corresponding Schmidt coefficient c_k of the operator $\mathbf{U}_k^\dagger \mathbf{O} \mathbf{U}_k = |\mathbf{e}_k\rangle\langle\mathbf{e}_k| \in \{|\xi_{ji}\rangle\}_{j,i=1}^{N,r}$.

The first condition ensures that the measurement outputs $\{\mathbf{o}_j\}_{j=1}^N$ are non-zero, enabling the identification of the target index k^* . Moreover, if the target operator $\mathbf{U}_{k^*}^\dagger \mathbf{O} \mathbf{U}_{k^*}$ is not included in $\{|\xi_{ji}\rangle\}_{j,i=1}^{N,r}$, measuring the output states with observable \mathbf{O} will always yield zero for all $k \in [2^n]$ such that any $k \in [2^n]$ is the solution of Eqn. (4).

The second condition ensures that the non-zero measurement outcomes can closely approximate the ground truth to facilitate the identification of the target index k^* . In this regard, the states set $\{|\xi_{ji}\rangle\}_{j,i=1}^{N,r}$, which associates with highly entangled input states (a large Schmidt rank r), is more likely to contain the unknown target operator $\mathbf{U}_{k^*}^\dagger \mathbf{O} \mathbf{U}_{k^*}$. Consequently, they exhibit a higher identifiability with a greater probability of producing non-zero measurement outcomes. On the other hand, the average magnitude of the Schmidt coefficients $\{c_{ji}\}_{i=1}^r$ of a highly entangled state $|\psi_j\rangle$ decreases with the Schmidt rank r . This leads to a small probability of the target operator $\mathbf{U}_{k^*}^\dagger \mathbf{O} \mathbf{U}_{k^*}$ being measured according to the Born rule. As a result, achieving a close approximation necessitates a larger number of measurements. For instance, the entangled states with $r=2^n$ could always yield non-zero measurement outputs with a large number of measurements. Conversely, if the unentangled state $|\psi_j\rangle\langle\psi_j|$ with $r=1$ is exactly identical to the target operator $\mathbf{U}_{k^*}^\dagger \mathbf{O} \mathbf{U}_{k^*}$, one measurement is sufficient to identify the target index k^* . These observations provide an intuitive indication about the transition role of entangled data from the lens of quantum information that the entangled data could contain more information than unentangled data, but at the same time increase the difficulty of information extraction using projective measurement.

Data availability

The entangled data generated in this study are available at the [Github](#) repository.

Code availability

The code used in this study are available at the [Github](#) repository.

References

1. Feynman, R. P. Simulating physics with computers. In *Feynman and computation*. p. 133–153. (CRC Press, 2018).
2. Shor, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Rev.* **41**, 303–332 (1999).
3. Lanyon, B. P. et al. Experimental demonstration of a compiled version of shor's algorithm with quantum entanglement. *Phys. Rev. Lett.* **99**, 250505 (2007).
4. Deutsch, D. & Jozsa, R. Rapid solution of problems by quantum computation. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* **439**, 553–558 (1992).
5. Grover, L. K. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, 212–219 (1996).
6. Harrow, A. W., Hassidim, A. & Lloyd, S. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.* **103**, 150502 (2009).
7. Lloyd, S., Mohseni, M. & Rebentrost, P. Quantum principal component analysis. *Nat. Phys.* **10**, 631 (2014).
8. Du, Y., Hsieh, Min-Hsiu, Liu, T., You, S. & Tao, D. Quantum differentially private sparse regression learning. *IEEE Trans. Inf. Theory* **68**, 5217–5233 (2022).
9. Schuld, M., Sinayskiy, I. & Petruccione, F. An introduction to quantum machine learning. *Contemp. Phys.* **56**, 172–185 (2015).
10. Biamonte, J. et al. Quantum machine learning. *Nature* **549**, 195 (2017).
11. Ciliberto, C. et al. Quantum machine learning: a classical perspective. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **474**, 20170551 (2018).
12. Dunjko, V. & Briegel, H. J. Machine learning & artificial intelligence in the quantum domain: a review of recent progress. *Rep. Prog. Phys.* **81**, 074001 (2018).
13. Li, W. & Deng, Dong-Ling Recent advances for quantum classifiers. *Sci. China Phys. Mech. Astron.* **65**, 220301 (2022).
14. Tian, J. et al. Recent advances for quantum neural networks in generative learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* **45**, 12321–12340 (2023).
15. Cerezo, M., Verdon, G., Huang, Hsin-Yuan, Cincio, L. & Coles, P. J. Challenges and opportunities in quantum machine learning. *Nat. Comput. Sci.* **2**, 567–576 (2022).
16. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
17. Moll, N. et al. Quantum optimization using variational algorithms on near-term quantum devices. *Quantum Sci. Technol.* **3**, 030503 (2018).
18. Havlíček, Vojtěch et al. Supervised learning with quantum-enhanced feature spaces. *Nature* **567**, 209 (2019).
19. Abbas, A. et al. The power of quantum neural networks. *Nat. Comput. Sci.* **1**, 403–409 (2021).
20. Huang, Hsin-Yuan et al. Power of data in quantum machine learning. *Nat. Commun.* **12**, 1–9 (2021).
21. Liu, Y., Arunachalam, S. & Temme, K. A rigorous and robust quantum speed-up in supervised machine learning. *Nat. Phys.* **17**, 1013–1017 (2021).
22. Wang, X., Du, Y., Luo, Y. & Tao, D. Towards understanding the power of quantum kernels in the NISQ era. *Quantum* **5**, 531 (2021).
23. Du, Y. & Tao, D. On exploring practical potentials of quantum auto-encoder with advantages. Preprint at [arXiv https://doi.org/10.48550/arXiv.2106.15432](https://doi.org/10.48550/arXiv.2106.15432) (2021).
24. Du, Y., Tu, Z., Wu, B., Yuan, X., & Tao, D. Power of quantum generative learning. Preprint at [arXiv https://doi.org/10.48550/arXiv.2205.04730](https://doi.org/10.48550/arXiv.2205.04730) (2022).
25. Du, Y., Yang, Y., Tao, D. & Hsieh, Min-Hsiu Problem-dependent power of quantum neural networks on multiclass classification. *Phys. Rev. Lett.* **131**, 140601 (2023).
26. Huang, Hsin-Yuan, Kueng, R. & Preskill, J. Information-theoretic bounds on quantum advantage in machine learning. *Phys. Rev. Lett.* **126**, 190505 (2021).
27. Bădescu, C. & O'Donnell, R. Improved quantum data analysis. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pp. 1398–1411 (2021).

28. Aharonov, D., Cotler, J. & Qi, Xiao-Liang Quantum algorithmic measurement. *Nat. Commun.* **13**, 887 (2022).
29. Chen, S., Cotler, J., Huang, Hsin-Yuan, & Li, J. Exponential separations between learning with and without quantum memory. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 574–585. (IEEE, 2022).
30. Huang, Hsin-Yuan et al. Quantum advantage in learning from experiments. *Science* **376**, 1182–1186 (2022).
31. Fanizza, M., Quek, Y., & Rosati, M. Learning quantum processes without input control. Preprint at arXiv <https://doi.org/10.48550/arXiv.2211.05005> (2022).
32. Polyzotis, N. & Zaharia, M. What can data-centric ai learn from data and ml engineering? Preprint at arXiv <https://doi.org/10.48550/arXiv.2112.06439> (2021).
33. Jakubik, J., Vössing, M., Kühn, N., Walk, J., & Satzger, G. Data-centric artificial intelligence. *Business & Information Systems Engineering*. pp. 1–9 (2024).
34. Jarrahi, MohammadHossein, Memariani, A. & Guha, S. The principles of data-centric ai. *Commun. ACM* **66**, 84–92 (2023).
35. Whang, S. E., Roh, Y., Song, H. & Lee, J.-G. Data collection and quality challenges in deep learning: a data-centric ai perspective. *VLDB J.* **32**, 1–23 (2023).
36. Zha, D., Bhat, Zaid Pervaiz, Lai, Kwei-Herng, Yang, F., & Hu, X. Data-centric ai: perspectives and challenges. In *Proceedings of the 2023 SIAM International Conference on Data Mining (SDM)*, pp. 945–948. (SIAM, 2023).
37. Zha, D. et al. Data-centric artificial intelligence: a survey. Preprint at arXiv <https://doi.org/10.48550/arXiv.2303.10158> (2023).
38. Wolpert, D. H. & Macready, W. G. No free lunch theorems for optimization. *IEEE Trans. Evolut. Comput.* **1**, 67–82 (1997).
39. Ho, Yu-Chi & Pepyne, D. L. Simple explanation of the no-free-lunch theorem and its implications. *J. Optim. Theory Appl.* **115**, 549–570 (2002).
40. Wolf, M. M. Mathematical foundations of supervised learning. *Lecture Notes from Technical University of Munich*, 1–168 (2018).
41. Adam, S. P., Alexandropoulos, Stamatios-Aggelos N, Pardalos, P. M., & Vrahatis, M. N. No free lunch theorem: a review. *Approximation and Optimization: Algorithms, Complexity and Applications*, pp. 57–82 (2019).
42. Brown, T. et al. Language models are few-shot learners. *Adv. Neural Inf. Process. Syst.* **33**, 1877–1901 (2020).
43. Ouyang, L. et al. Training language models to follow instructions with human feedback. *Adv. Neural Inf. Process. Syst.* **35**, 27730–27744 (2022).
44. Bai, Y. et al. Training a helpful and harmless assistant with reinforcement learning from human feedback. Preprint at arXiv <https://doi.org/10.48550/arXiv.2204.05862> (2022).
45. Touvron, H. et al. Llama: Open and efficient foundation language models. Preprint at arXiv <https://doi.org/10.48550/arXiv.2302.13971> (2023).
46. Zhao, Wayne Xin et al. A survey of large language models. Preprint at arXiv <https://doi.org/10.48550/arXiv.2303.18223> (2023).
47. Poland, K., Beer, K., & Osborne, T. J. No free lunch for quantum machine learning. Preprint at arXiv <https://doi.org/10.48550/arXiv.2003.14103> (2020).
48. Sharma, K. et al. Reformulation of the no-free-lunch theorem for entangled datasets. *Phys. Rev. Lett.* **128**, 070501 (2022).
49. Zhao, H. et al. Learning quantum states and unitaries of bounded gate complexity. Preprint at arXiv <https://doi.org/10.48550/arXiv.2310.19882> (2023).
50. Preskill, J. Quantum computing in the nisq era and beyond. *Quantum* **2**, 79 (2018).
51. Huang, Hsin-Yuan, Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16**, 1050–1057 (2020).
52. Elben, A. et al. The randomized measurement toolbox. *Nat. Rev. Phys.* **5**, 9–24 (2023).
53. Jozsa, R. & Linden, N. On the role of entanglement in quantum-computational speed-up. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **459**, 2011–2032 (2003).
54. Yoganathan, M. & Cade, C. The one clean qubit model without entanglement is classically simulable. Preprint at arXiv <https://doi.org/10.48550/arXiv.1907.08224> (2019).
55. Khatri, S. et al. Quantum-assisted quantum compiling. *Quantum* **3**, 140 (2019).
56. Yuen, H. An improved sample complexity lower bound for (fidelity) quantum state tomography. *Quantum* **7**, 890 (2023).
57. Anshu, A. & Arunachalam, S. A survey on the complexity of learning quantum states. *Nat. Rev. Phys.* **6**, 59–69 (2024).
58. Piani, M. & Watrous, J. All entangled states are useful for channel discrimination. *Phys. Rev. Lett.* <https://doi.org/10.1103/PhysRevLett.102.250501> (2009).
59. Bae, J., Chruściński, D. & Piani, M. More entanglement implies higher performance in channel discrimination tasks. *Phys. Rev. Lett.* **122**, 140404 (2019).
60. Lowe, A. & Nayak, A. Lower bounds for learning quantum states with single-copy measurements. Preprint at arXiv <https://doi.org/10.48550/arXiv.2207.14438> (2022).
61. Wu, Y., Payne, M. G., Hagle, E. W. & Deng, L. Preparation of multiparty entangled states using pairwise perfectly efficient single-probe photon four-wave mixing. *Phys. Rev. A* **69**, 063803 (2004).
62. Basharov, A. M., Gorbachev, V. N. & Rodichkina, A. A. Decay and storage of multiparticle entangled states of atoms in collective thermostat. *Phys. Rev. A* **74**, 042313 (2006).
63. Lemr, K. & Fiurášek, Jaromír Preparation of entangled states of two photons in several spatial modes. *Phys. Rev. A* **77**, 023802 (2008).
64. Lin, Y. et al. Preparation of entangled states through hilbert space engineering. *Phys. Rev. Lett.* **117**, 140502 (2016).
65. Klco, N. & Savage, M. J. Minimally entangled state preparation of localized wave functions on quantum computers. *Phys. Rev. A* **102**, 012612 (2020).
66. Schatzki, L., Arrasmith, A., Coles, P. J., & Cerezo, M. Entangled datasets for quantum machine learning. Preprint at arXiv <https://doi.org/10.48550/arXiv.2109.03400> (2021).
67. Caro, M. C. et al. Out-of-distribution generalization for learning quantum dynamics. *Nat. Commun.* **14**, 3751 (2023).
68. Jerbi, S. et al. The power and limitations of learning quantum dynamics incoherently. Preprint at arXiv <https://doi.org/10.48550/arXiv.2303.12834> (2023).
69. Caro, M. C. et al. Generalization in quantum machine learning from few training data. *Nat. Commun.* **13**, 4919 (2022).
70. Bisio, A., Chiribella, G., D’Ariano, GiacomoMauro, Facchini, S. & Perinotti, P. Optimal quantum learning of a unitary transformation. *Phys. Rev. A* **81**, 032324 (2010).
71. Jones, T. & Benjamin, S. C. Robust quantum compilation and circuit optimisation via energy minimisation. *Quantum* **6**, 628 (2022).
72. Heya, K., Suzuki, Y., Nakamura, Y., & Fujii, K. Variational quantum gate optimization. Preprint at arXiv <https://doi.org/10.48550/arXiv.1810.12745> (2018).
73. Cirstoiu, C. et al. Variational fast forwarding for quantum simulation beyond the coherence time. *npj Quantum Inf.* **6**, 82 (2020).
74. Gibbs, J. et al. Dynamical simulation via quantum machine learning with provable generalization. *Phys. Rev. Res.* **6**, 013241 (2024).
75. Huang, Hsin-Yuan, Chen, S. & Preskill, J. Learning to predict arbitrary quantum processes. *PRX Quantum* **4**, 040337 (2023).

76. Caro, M. C. Learning quantum processes and hamiltonians via the pauli transfer matrix. Preprint at *arXiv* <https://doi.org/10.48550/arXiv.2212.04471> (2022).
77. Bubeck, S., Chen, S., & Li, J. Entanglement is necessary for optimal quantum property testing. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 692–703. (IEEE, 2020).
78. Duchi, J. Lecture notes for statistics 311/electrical engineering 377. *notes pdf*. Last visited **2**, 23 (2016).

Acknowledgements

Y.L. acknowledges support from National Natural Science Foundation of China (Grant Nos. U23A20318 and 62276195). X.Y. acknowledges support from the National Natural Science Foundation of China (Grant Nos. 12175003 and 12361161602), NSAF (Grant No. U2330201).

Author contributions

The project was conceived by Y.D. and D.T. Theoretical results were proved by X.W., Y.D. and Z.T. Numerical simulations and analysis were performed by X.W., Y.D., Y.L. and X.Y. All authors contributed to the write-up.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-47983-1>.

Correspondence and requests for materials should be addressed to Yuxuan Du, Yong Luo or Dacheng Tao.

Peer review information *Nature Communications* thanks Nai-Hui Chia and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024