

ARTICLE OPEN



Intelligent certification for quantum simulators via machine learning

Tailong Xiao¹, Jingzheng Huang¹, Hongjing Li¹, Jianping Fan² and Guihua Zeng¹✉

Quantum simulation is a technology of using controllable quantum systems to study new quantum phases of matter. Certification for quantum simulators is a challenging problem whereas identification and properties estimation are two crucial approaches that can be resorted to. In this work, we propose *Ab initio* end-to-end machine learning certification protocol briefly named MLCP. The learning protocol is trained with a million-level size of randomized measurement samples without relying on the assistance of quantum tomography. In the light of MLCP, we can identify different types of quantum simulators to observe their distinguishability hardness. We also predict the physical properties of quantum states evolved in quantum simulators such as entanglement entropy and maximum fidelity. The impact of randomized measurement samples on the identification accuracy is analyzed to showcase the potential capability of classical machine learning on quantum simulation results. The entanglement entropy and maximum fidelity with varied subsystem partitions are also estimated with satisfactory precision. This work paves the way for large-scale intelligent certification of quantum simulators and can be extended onto an artificial intelligence center to offer easily accessible services for local quantum simulators in the noisy intermediate-size quantum (NISQ) era.

npj Quantum Information (2022)8:138; <https://doi.org/10.1038/s41534-022-00649-6>

INTRODUCTION

Quantum simulation providing access to theoretical investigations that are currently impossible is the most promising application of the NISQ devices^{1–3}. Quantum simulation technologies are beneficial for the study of quantum metrology^{4–6} and quantum computation^{7–11}. In the NISQ era, the random circuit sampling and Boson sampling problems are realized to achieve quantum supremacy (also called quantum advantage) based on superconducting quantum circuit and photonics platform, respectively^{9–11}. These machines are dedicated to solving specific problems and we call them quantum simulation machines (QSMs) or quantum simulators.

Even though the research boom in quantum simulation, the practical challenges of manipulating large-scale quantum matter still prevent progress from constructing a fault-tolerant quantum computer. In addition, due to the absence of universal quantum error correction which is generally hard to be implemented both in digital and analog quantum simulation, the certification for QSMs is necessary and urgent. Currently, there are a large number of studies that concentrate on quantum validation and certification¹², such as quantum Hamiltonian learning¹³, quantum tomography¹⁴, direct fidelity estimation¹⁵, random benchmarking¹⁶, and quantum cross-platform verification¹⁷. As an example, quantum tomography requires exponential quantum measurements ($\sim 4^N$, N is the number of qubits) to reconstruct the full quantum states although it can obtain all the information about the quantum state¹² thereby preventing its application in large-scale QSMs. Without using tomography, previous seminal researches¹⁸ have proposed methods including estimating the overlap of two quantum states from different platforms¹⁹, and the cross-platform verification protocol based on measurement-based quantum computation²⁰. These methods extract statistical information from measurement records of QSMs to estimate the maximum fidelity or cross-correlation such that can intuitively

evaluate the reliability of computational results. Notably, exploring the hidden pattern from a large-scale dataset is highly suited for machine-learning methods. Numerous works demonstrate that machine learning plays a crucial role in quantum physics and simulation, such as state discrimination²¹, tomography^{22–26} and parameter estimation⁵, whereas machine learning-based certification for QSMs that does not resort to state tomography is rarely investigated. In ref. ²⁷, a neural network classifier is constructed to directly estimate the fidelity using positive operator value measurements (POVM). Machine-learning approach for certification via randomized measurement has a potential advantage and is required to be studied.

In this regard, we devise a generic machine-learning certification protocol (MLCP) by leveraging the advanced LSTM-Transformer hybrid framework from natural language processing. Through investigating two critical approaches of certification, i.e., identification and quantum properties estimation for QSMs, the MLCP model achieves impressive performance, particularly in terms of identification accuracy. More specifically, the MLCP can automatically classify different types of QSMs with the randomized measurement results gathered from different simulated or experimental implementations by mapping the sparse measurement results into a high-dimensional embedding. The learned mapping can further be applied to quantum properties estimation such as entanglement entropy²⁸ and maximum fidelity prediction by using the regression method. The identification accuracy showcases the distinguishability hardness of different types of QSMs through training a million-level number of samples. Interestingly, the impact of repeated measurement settings on the ultimate accuracy provides a straightforward witness for the potential capability of classical machine learning in discriminating quantum channels/states, which also provides essential guidance for practical measurements in QSMs. The estimation of physical properties in small-scale quantum systems achieves satisfactory

¹State Key Laboratory of Advanced Optical Communication Systems and Networks, and Center of Quantum Sensing and Information Processing, Shanghai Jiao Tong University, 200240 Shanghai, China. ²Department of Computer Science, University of North Carolina-Charlotte, Charlotte, NC 28223, USA. ✉email: ghzeng@sjtu.edu.cn

precision, especially when estimating maximum fidelity. As for relatively larger sizes, the entanglement entropy can also be predicted without using strictly exponentially scaled measurement samples based on the MLCP model. The entanglement entropy and maximum fidelity can be predicted with polynomial number of measurements at the inference stage to diagnose the evolution of the QSMs. The model can be readily extended and transferred into an artificial intelligence cloud center such that local QSMs can upload their measurement data to obtain certification via the machine-learning model.

RESULTS

Machine-learning certification protocol

We present two functions of the MLCP based on the classification and regression methods in machine learning. The procedures of the protocol can be seen in Fig. 1 where four main steps are presented to accomplish the certification. The first step is to generate quantum simulation data, which can collect the experiments or classical simulations based on matrix product state or density matrix simulator. The quantum Hamiltonians can be different types such as disorder/ordered. Besides, quantum evolution can consist of different evolution times that generate the measurement records for different quantum states.

In this work, we investigate the long-range XY model²⁹ in the presence of a transverse field as a case study of our protocol shown in Fig. 1a. We remark that the analog system can also be a two-dimensional many-body systems and the protocol can be feasible in digital quantum evolution with a variational quantum circuit in different physical platforms such as superconducting and cold ion systems as depicted in Fig. 1a. The quantum Hamiltonian of the XY model is given by

$$H_{XY} = \hbar \sum_{i < j} J_{ij} (\sigma_i^+ \sigma_j^- + \sigma_i^- \sigma_j^+) + \hbar B \sum_j \sigma_j^z, \quad (1)$$

where \hbar is Planck's constant divided by 2π , σ_i^β ($\beta = x, y, z$) denotes the spin- $\frac{1}{2}$ Pauli operators, σ_i^+ (σ_i^-) are the spin-raising (lowering) operators acting on site i , and $J_{ij} \approx J_0/|i - j|^\alpha$ are the coupling coefficients with an approximate power-law decay and $0 < \alpha < 3$ ³⁰. Alternatively, a locally disordered potential can be added to realize the Hamiltonian $H = H_{XY} + H_D$ where $H_D = \hbar \sum_j \Gamma_j \sigma_j^z$ and Γ_j is the magnitude of disorder acting on site j . To investigate larger sizes of quantum systems, we also conduct classical simulations for the one-dimensional Heisenberg model. The Hamiltonian is given by

$$H_{hg} = J_z \sum_{j=1}^{N-1} \sigma_j^z \sigma_{j+1}^z + \frac{1}{2} \sigma_j^+ \sigma_{j+1}^- + \frac{1}{2} \sigma_j^- \sigma_{j+1}^+, \quad (2)$$

where the model only consists of nearest-neighbor interactions, which can be efficiently simulated based on matrix product state time evolution. Two typical couplings are chosen: $J = 1$ or uniformly sampled $J \in (0, 1]$. Note that the Planck constant \hbar is normalized into the coupling coefficients J_j . A common quantum computation process is composed of three steps: (1) the quantum probe state preparation denoted as $|\psi_0\rangle$, (2) the quantum evolution or a series of gate operations which can be represented as a quantum operation or channel $\mathcal{U} = \exp(-iHt)$ as shown in Fig. 1. The analog quantum evolution can be quench dynamics or thermalization. The digital quantum evolution can be represented by a variational circuit composed of interleaved multi-rotation and entangled layers. (3) Randomized measurement unitaries are written as $U = U_1 \otimes \dots \otimes U_N$ on a computational basis for each local qubit. There are two typical randomized unitary measurements: (i) circular unitary ensemble (CUE)³¹, (ii) random Pauli measurement (also referred as classical shadow³²). The final measurement operation generates bit string $\mathbf{s} \in \{0, 1\}^N$ with probability $P(\mathbf{s}) = \langle \mathbf{s} | U \rho U^\dagger | \mathbf{s} \rangle$, where ρ is the quantum state to

be measured. The repeated measurement collects the statistical properties of the evolved quantum states.

The physical probe state in both models is prepared as the Néel ordered state denoted by $\rho_0 = |\psi_0\rangle\langle\psi_0|$ with $|\psi_0\rangle = |0101 \dots 01\rangle$. This state was subsequently time-evolved under the specified Hamiltonian into the final state $\rho(t)$. The initial product state will be entangled with various types under the coherent interactions. Subsequently, randomized measurements are performed on $\rho(t)$ through local rotations of each qubit i.e., U_i uniformly sampled from the CUE or Pauli group. In reality, each U_i can be decomposed into three rotations $R_z(\theta_3)R_y(\theta_2)R_z(\theta_1)$, and this sandwich structure can ensure that drawing of the U_i is stable against small drifts of physical parameters controlling the rotation angles θ_i ²⁹. Random Pauli measurement that performs x, y, z operations before the final z -measurement is easier to be implemented compared to sampling unitaries from CUE. Finally, computational basis measurements are performed on each local qubit to generate the measurement bit strings. For each set of applied unitaries U , the measurement is repeated N_M times and we implement N_U unitaries, thus generating $N_U N_M$ measurement samples. In reality, repeating N_M is much easier to be implemented compared to switching the measurement settings N_U since the latter requires constantly producing external control pulse to manipulate the qubits, which will introduce quantum noise and measurement errors and slow down the speed of data acquisition. The former, however, only requires repeating the same quantum evolution and measuring the quantum state. More details about the randomized measurement scheme can be found in Supplementary Note 2.

The one-dimensional XY Hamiltonian of Eq. (1) is a representative model for studying many-body localization (MBL). The disorder Hamiltonian terms in this model have an impact on the entropy growth of different subsystems. In addition, the Hamiltonian with different field couplings J_{ij} has distinct evolution characteristics. We consider two different field couplings: $J = 370 \text{ s}^{-1}$, $\alpha = 1.01$ and $J = 420 \text{ s}^{-1}$, $\alpha = 1.24$ with and without disorder effect, thus giving rise to four categories of QSMs. In addition, for different total evolution times, the quantum system undergoes a different evolution with a distinguishable variation pattern of entanglement entropy. MLCP learns from the quantum simulation data with different evolution times by given supervised signals such as the true categories of the data or the theoretical entanglement and maximum fidelity. Besides, the noisy time evolution of quantum simulators can be regarded as a quantum channel. Therefore, identifying different types of quantum simulators can be viewed as quantum channel discrimination which is also a basic problem in quantum information.

To study the performance of MLCP in properties estimation in large system sizes, we conduct additional quantum simulations based on the one-dimensional Heisenberg model to generate randomized measurement data with different time evolutions. We note Hamiltonian (2) only has short-range interactions so that the matrix product state can effectively simulate the evolutions even if the system size is large. The non-Hermitian evolution is relatively hard to simulate in a matrix product state such that we only consider the unitary evolutions.

Through large amounts of numerical simulation for the XY model and Heisenberg model, we collect the randomized measurement samples to construct the training dataset. It turns out that generating the quantum simulation data is highly expensive, particularly in those long-range systems where tensor network states are hard to simulate as the bond dimension increases dramatically. The generated quantum simulation dataset can be used to assist the real quantum simulation experiments. Compared to estimating the fidelity or entropy of the states directly, the information gained from distinguishing the states from different physical conditions is not much stronger. However, the following quantum properties estimation can fulfill this

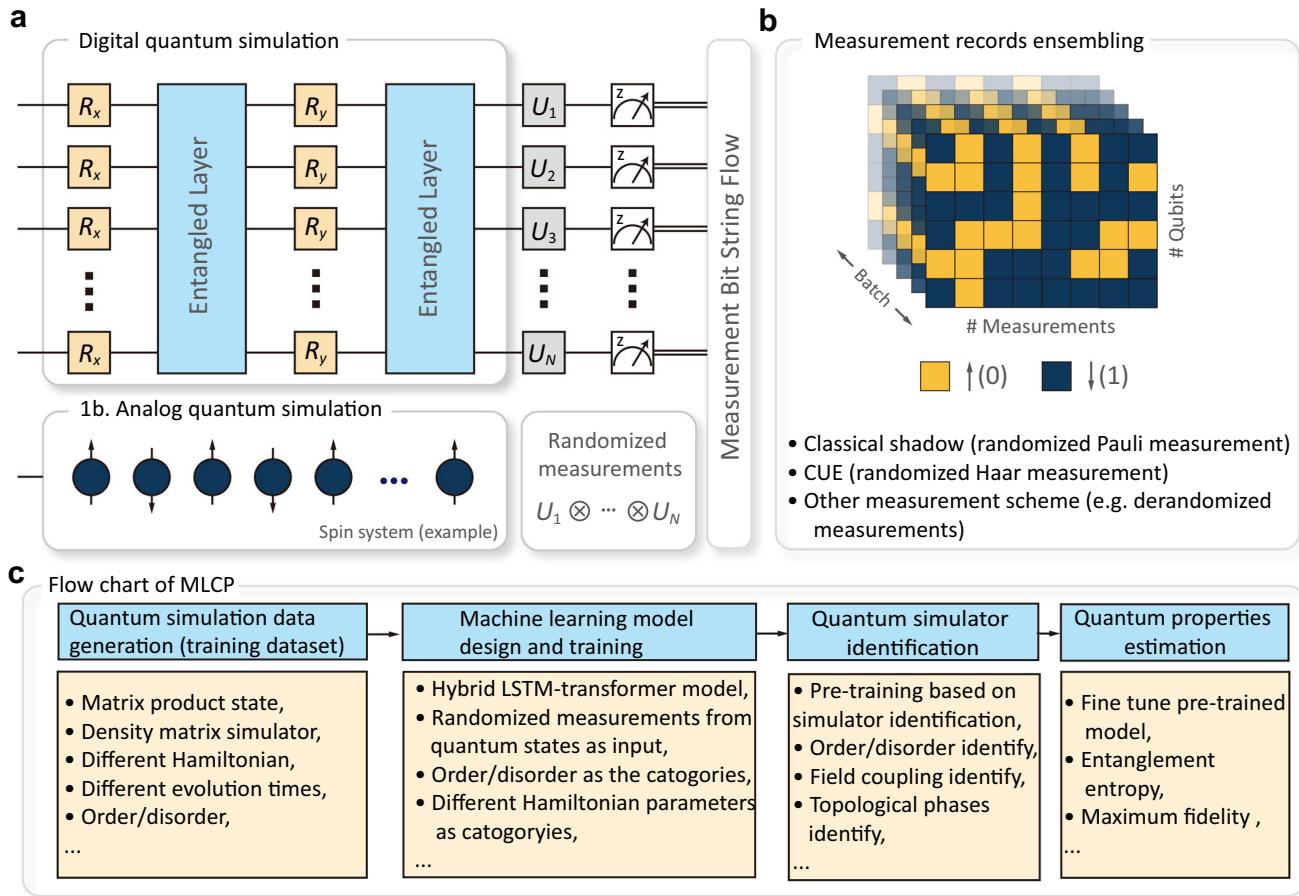


Fig. 1 Ab initio end-to-end machine-learning certification protocol (MLCP) for quantum simulation machines (QSMs). **a** The digital quantum evolution represented by variational quantum circuit, or the analog quantum evolution such as quantum Ising spin model or Bose–Hubbard model is prepared. Both two evolutions adopt randomized unitaries to collect the measurement records. **b** Reshaping the measurement records and concatenating them into one batch computational tensor. Each row denotes records of the total number of measurements repeated for one qubit. Each column represents the full-size qubit records with single-shot measurement. **c** The flowchart of MLCP works. Our protocol consists of four procedures: quantum simulation data generation, machine-learning design and training, quantum simulator identification, and properties estimation. For the large-scale quantum simulation in classical computers, matrix product states are highly suitable for simulating states with entanglement entropy growth limited to $\log D$ with D denoting the maximum bond dimension.

deficiency. The details of the physical system, the randomized measurement, and the training dataset generation can be found in Supplementary Notes 1 and 2.

Identification phase

The first function of the protocol is to identify different types of QSMs based on different Hamiltonians e.g., distinguishing the final state belonging to different field couplings and whether the system is clean or disordered, as described in Fig. 1c. The identification process can be implemented by using an LSTM-Transformer-LSTM (LTL) encoder with a sandwich architecture (see “Methods”). First, we conduct quantum simulations for the noisy XY model to collect the randomized measurement records. Digital quantum simulation such as the variational quantum circuit is also feasible. Secondly, we construct a batch-style computational tensor by concatenating the measurement bit strings where the row dimension denotes the number of qubits N and the column denotes the measurement dimension with N_M and the batch dimension denotes shuffled measurement settings. The encoder learns the high-dimensional representation of the measurement records ensemble (as Fig. 1b shows) via supervised learning in which the labels are determined by the varied evolution times and the type of the QSMs. The details of the label settings can be found in Supplementary Note 1. The qubit site matters in our

protocol as the physical properties are closely related to the partition of the many-body system. Remarkably, the number of recurrent neurons scales linearly with N since each local site is regarded as a time slot in time-series modeling implying that MLCP has flexible scalability over system sizes. Note that the number of measurements required to accomplish the identification process in the classical computer still scale exponentially with N where the demonstration can be found in Supplementary Note 3. This inherent sample inefficiency is due to the quantum and classical separation gap³³ demonstrating only using classical processing is hard to complete specific quantum tasks such as identifying topological phases and estimating non-linear properties. Even though, MLCP is more concentrated on the practical applicability in the quantum certification when given a large number of training samples (which may be exponentially sufficient) combining pre-training techniques. In this function, we aim to apply the MLCP model to study the distinguishability of different types of QSMs. In addition, we also study the impact of different measurements N_U, N_M on the ultimate performance of the machine-learning model to analyze the capability of the classical learning model. As a result, we can leverage the LTL encoder to map the measurement records into the high-dimensional embedding and then identify the different physical conditions. The identification phase can pre-train the LTL model

according to specified criteria such as order/disorder. The learned model contains the prior information about the quantum system, which is likely to be beneficial for the downstream estimation phase. In addition, the protocol can uncover the unknown physical evolution conditions of the final quantum state to provide quantum certification.

Property estimation phase

The protocol can also be used to estimate the quantum properties of the quantum state under various QSMs such as the purity p and the second-order Rényi entropy $S^{(2)}$. We regard the estimation process as the quantum downstream task, a concept motivated by the classical machine-learning research field^{34–36}. The downstream estimation task can be referred to as transfer learning which uses the prior knowledge of the pre-trained model to improve the performance of the specific task. Physical properties such as entropy and mutual information can be utilized to provide an insightful diagnostic tool for QSMs. We make use of single-state estimation to predict entanglement entropy. The cross-correlation of quantum states from different platforms (also called maximum fidelity) is also defined to evaluate the reliability of QSMs. The maximum fidelity of two quantum states can be estimated by two-state estimation with the assistance of the LTL model. In the case of single-state estimation, we are required to estimate the purity $p = \text{Tr}[\rho^2]$ which only involves a single final quantum state ρ . Thus, we only collect the measurement records under each random unitary U and map the data into the feature embedding through the LTL encoder, and finally feed the feature embedding into the multi-hidden ANN to estimate the purity sampled from a learned Gaussian distribution $\mathcal{N}(\mu, \sigma)$ (see “Methods”, Fig. 7b). The second-order Rényi entropy can further be calculated analytically by using the formula

$$S^{(2)}(\rho) = -\log_2 \text{Tr}[\rho^2]. \quad (3)$$

The theoretical value such as the entanglement entropy can be calculated during the numerical simulation by directly calculating the trace of the density matrix square. The loss function is selected as the root mean square error (RMSE) between the true entropy and the predicted entropy. Entropy estimation is generally a complex problem and there are numerous unsupervised machine-learning methods³⁷. In our model, we leverage supervised learning to investigate the feasibility of entropy estimation via randomized measurements.

In two-state estimation, we aim to solve the cross-correlation property of two quantum states ρ_1, ρ_2 given by

$$\mathcal{F}_{\max} = \frac{\text{Tr}[\rho_1 \rho_2]}{\max\{\text{Tr}[\rho_1^2], \text{Tr}[\rho_2^2]\}}. \quad (4)$$

The common physical quantities are the fidelity and the overlap normalized over the maximum purity between two quantum states. The two-state estimation can be treated as the cross-platform verification where two quantum states share the same randomized unitary U and measurement shots N_M in case we have no access to the theoretical quantum states. The learning process requires the true label (theoretical properties of the quantum states) of the measurement records. The theoretical overlap between two QSMs should be 1 but in reality, finite measurements cannot guarantee the calculated overlap of two states simulated by two QSMs to be 1 when the quantum noise of different platforms exists. Therefore, we make use of the conventional cross-validation estimator to numerically calculate the estimated overlap as the label of our supervised learning model. A detailed description can be found in Supplementary Note 1B. In addition, we equally split our measurement records into two parts and separately feed the records into the same model to obtain two embeddings. Then we obtain the difference between these two

embeddings. This operation is quite similar to the siamese network architecture^{38,39}. Eventually, we feed the difference into the multi-hidden ANN to obtain the ultimate physical properties estimation sampled from a Gaussian distribution with mean μ and variance σ . The probability sampling from the distribution can enhance the robustness of the model against noisy data samples. We note that the MLCP does not require the quantum state tomography technique to reconstruct the full density matrix to calculate the overlap. In our protocol, we directly input the measurement records into the machine-learning model and obtain the final physical properties estimation. The whole process constitutes a unified end-to-end framework to provide quantum certification for QSMs. A detailed analysis of machine-learning architecture and its realizations can be found in Methods and Supplementary Notes 2 and 3A.

Identification results of QSMs

We simulate two sets of physical parameters for Hamiltonian quench dynamics. Specifically, case (1) specifies: the maximum coupling $J_0 = 420 \text{ s}^{-1}$, $a = 1.24$ with and without disorder term H_D ; case (2) specifies: the maximum coupling $J_0 = 370 \text{ s}^{-1}$, $a = 1.01$ with and without disorder term. The number of qubits $N = 10$ in both two simulations. We can regard these two cases of physical parameters as a category set $\mathcal{C} = \{C_i, i = 1, 2, 3, 4\}$ where each category C_i denotes one specific physical condition. In category C_1 with clean system and case (1) condition, we simulate 11 equal-spaced evolution times $t \in [0, 10] \text{ ms}$. In class C_2 with disorder and case (1) condition, we simulate 21 equal-spaced evolution times $t \in [0, 20] \text{ ms}$. In category C_3 with clean and case (2) condition and C_4 with disorder and case (2) condition, we simulate 11 equal-spaced evolution times $t \in [0, 10] \text{ ms}$. The detailed dataset description can be found in Supplementary Notes 1 and 3B.

We do not adopt four categories identification process. Instead, we merge category C_1, C_3 into one category C_{clean} and C_2, C_4 as another category C_{disorder} . This operation transforms a multi-class classification problem into a binary classification problem. Naturally, we can also merge C_1, C_2 into one category C_{420} and C_3, C_4 into another category C_{370} . The binary classification emphasizes the difference in the physical model between the field coupling strength and disorder or clean. We may obtain some useful and intuitive findings by analyzing the ultimate classification accuracy of these two binary classifications. We use a binary cross entropy loss function to train the model. The simulation results of the identification $C_{\text{clean}}, C_{\text{disorder}}$ and C_{420}, C_{370} are presented in Fig. 2. Figure 2a, b demonstrates the performance of the identification on $C_{\text{clean}}, C_{\text{disorder}}$. Figure 2d, e shows the performance of the identification on C_{420}, C_{370} . Figure 2c, f shows the performance of two identification problems when N_U and N_M given a set of decreased values. The hyperparameters such as the batch size, the learning rate adjustment, and the network parameters can be found in Supplementary Note 5. The Micro-F1 and weighted-F1 scores are calculated to evaluate the accuracy⁴⁰ in the imbalanced dataset as class C_2 owns more data samples than other classes in our simulation dataset. The LTL model is trained without the information of randomized unitary U sampled from CUE and achieves a high identification accuracy implying that U is not the necessary component in our model. Note that, adding the information of U is possible to further enhance the accuracy by designing an appropriate model but it is beyond the scope of our work. Due to the probabilistic feature of the measured bit strings, the LTL model requires burning a few epochs to find a good network initialization in both identifications as Fig. 2a, d shows. The accuracy increases smoothly after the burning period implying the network finds a good direction of parameter optimization. We note that the burning epochs are varied with the batch size and the learning rate when fixing the network parameters. Generally, a large batch size will consume

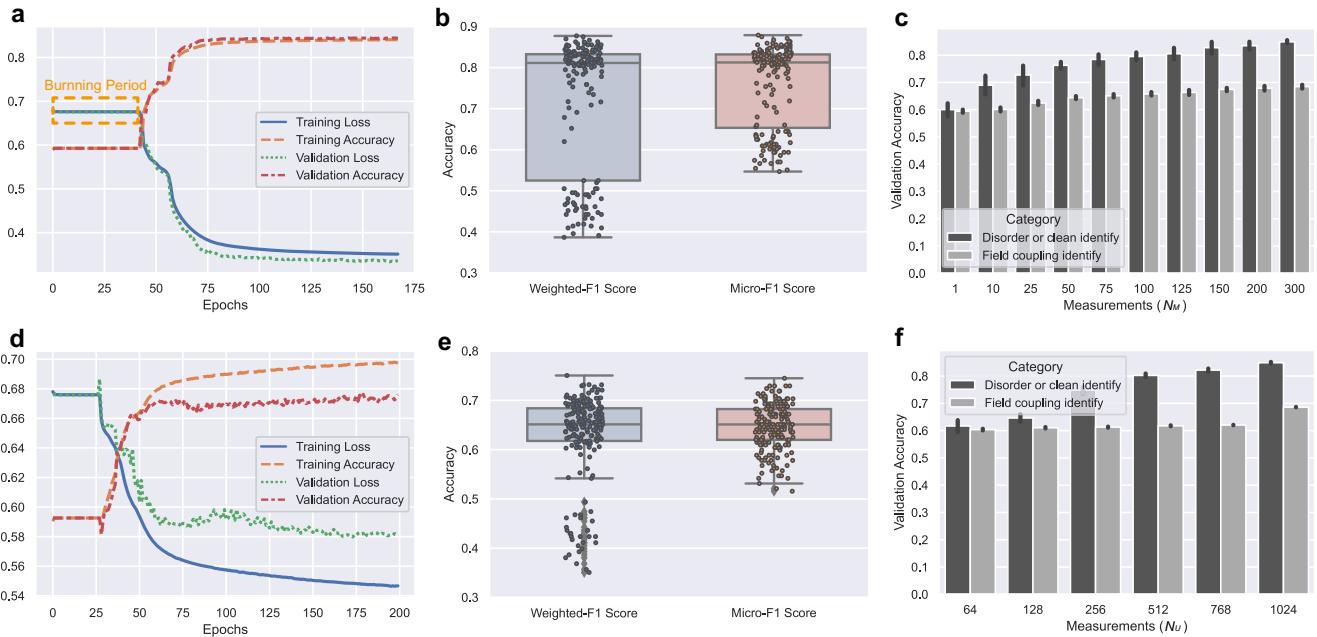


Fig. 2 Machine-learning performance characterized by the loss and accuracy during the training and validation process as well as the Micro-F1 and weighted-F1 scores for identifying C_{disorder} , C_{clean} and C_{420} , C_{370} . **a** The losses continuously decrease as the training epoch increases both in the training and validation dataset. On the contrary, the training and validation accuracy increase correspondingly. The measured bit strings are simulated with $N_U = 1024$, $N_M = 300$. **b** The Micro-F1 and weighted-F1 scores are demonstrated in a box plot with the maximum accuracy approaching 90%. The box plot is a graphical way to depict data through their quartiles. The small circles in (b) and (e) denote the actual validation accuracy during training processing. **c** The performance of two identification tasks is varied with N_M when given $N_U = 1024$. **d** The losses decrease smoothly after the burning period and the ultimate validation accuracy reaches 68% when given $N_U = 1024$, $N_M = 300$. **e** The box plot of Micro-F1 and weighted-F1 score of each epoch during 200 epochs. The maximum weighted-F1 and accuracy can reach approximately 75%. **f** The performance of two identification tasks varies with different N_U when given $N_M = 300$. The error bar in (e, f) denotes the standard variance over repeated simulations.

more epochs to find a good starting point for parameter optimization. We remark that the fact that the validation loss is smaller than the training loss is caused by two possibilities: (1) a relatively large dropout rate (0.5 in our model). In the training phase, the dropout will freeze the neurons but in the validation phase, all neurons are activated. This technique is used to avoid overfitting. (2) sample noise. Since the evolution is simulated with quantum noise, the measurement data is also noisy. The validation set contains less noise than the training set, thus leading to a smaller loss value. In Fig. 2b, e, the box plot shows the distribution of the validation accuracy of each epoch during the whole training process. Particularly in Fig. 2b, only a small portion of the accuracy is less than 2/3 (outliers denoted as circles) which demonstrates that our model can show satisfiable performance even with small training epochs. On the contrary, Fig. 2e, d shows a more balanced distribution of validation accuracy. The ultimate precision can reach 85% when identifying disorder or clean of QSMs in our large-scale dataset.

When identifying the physical system with different couplings, i.e., binary classification on C_{420} , C_{370} , we also adopt the same MLCP architecture and dataset to train the LTL learning model. From Fig. 2d, e, it turns out that the model is harder to be trained and the ultimate identification accuracy that the model can reach is lower compared to the case of identifying disorder or clean. As a consequence, the final performance of the LTL model reveals that the quench dynamics under disorder or clean is easier to be distinguished than the dynamics under different couplings. The finding is well consistent with the experimental result of ref. ²⁹, where the z-magnetization of qubits $\langle \sigma_i^z \rangle$ is calculated to show a more distinct feature between disorder and clean systems. On the contrary, the magnetization of different couplings shares a similar variation pattern which is relatively hard to be distinguished. From this perspective, the LTL model does not rely on any specific

observable to make the classification. It absorbs the observable selection process into a neural network in a black-box fashion. Thus, the LTL model can automatically identify patterns hidden in seemingly random measurement bit strings thereby providing the certification for QSMs although the information gain is not highly large.

In Fig. 2c, f, we present the performance of the LTL model when varying the measurements N_U and N_M to investigate the impact of the dataset size on the ultimate accuracy. The LTL model is a classical machine-learning algorithm and the measured data is not mapped into quantum memory. Therefore, the required number of samples should scale exponentially as the number of qubits^{41,42}. In Fig. 2c, the validation accuracy varied with N_M , and basically, large N_M generates higher accuracy as we expect. For the disorder or clean identification, we can divide the set of N_M into three phases: $N_M \in \{1 - 10\}, \{10 - 75\}, \{75 - 300\}$. In the first phase, the increase of N_M can lead to the largest accuracy enhancement compared to the other two phases. In the second phase, the accuracy enhancement rate decrease compared with the first phase when continuously increasing N_M . In the last phase, the accuracy enhancement rate is continued to slow down when we increase N_M . On the contrary, in the field coupling identification task, the accuracy obeys a linear increase when we increase N_M in a nearly linear schedule although the enhancement rate is lower than the first identification task. The simulation results demonstrate that repeated measurement is necessary when given a local random unitary to enhance the identification accuracy of QSMs. In Fig. 2f, we present the validation accuracy varied with N_U . In disorder or clean identification, the accuracy is linearly increased as we exponentially increase N_U . The results imply the classical machine-learning algorithm is not sample-efficient in handling quantum simulation (computation) measurement data if we do not resort to quantum memory. In contrast, in the field

coupling identification task, the exponential increase of sample still leads to a linear increase in validation accuracy although the linear increase is highly small and is nearly 1%. In Supplementary Note 4, we provide a detailed description of the performance analysis when varying the measurements N_U, N_M and we also present a detailed analysis of model selection and hyperparameters adjustment.

Properties estimation results

To simplify the model training hardness and convince the estimation results, we separately train the samples in categories C_1, C_2 to estimate the Rényi entropy of the probability distribution hidden in measurement bit strings and the maximum fidelity. During training, for different forward partition N_A , the corresponding measurement bit string s_A is fed into the neural network to make the prediction. We note that the unitary information is not necessary when estimating the entanglement entropy and maximum fidelity. Both properties can be estimated by extracting the auto-correlation and cross-correlation of the measured bit strings. The exponentially scaled measurement samples can guarantee the information completeness of randomized measurements⁴³. On the contrary, when reconstructing the quantum states, the local unitary matrix is necessary when using classical shadow. More details can be found in Supplementary Note 2.

As for entanglement entropy estimation, it turns out that the entropy is maximum when the spin chain is equally partitioned (see Supplementary Note 4B). When we consider different evolution times, the value of the entropy distributes in $[0, 4]$ through observing the theoretical value of our data samples. However, the distribution of the measurement data and the target value are discrete since our evolution is discretized leading to a high-resolution distribution of final quantum states. Therefore, our dataset may not be fully sufficient to train the model well to estimate the entropy precisely especially when the value of the entropy is distributed in a wide range. This effect becomes more prominent when the subsystem is closer to equally partitioned. On the contrary, the ground-truth maximum fidelity distributes in a smaller range regardless of the subsystem partition, which means the measurement data can be viewed as a continuous distribution over the target space. The visualization and the detailed analysis can be found in Supplementary Note 4B.

The numerical results of estimating the entropy for two systems are presented as Fig. 3 shows. When estimating the entropy, large N_A has a smaller validation error both for clean and disorder systems, which meets our expectations. It turns out that large N_A has a more compact entropy distribution for different evolution times. Conversely, half-partition with $N_A = 5$ has the largest error up to 0.98 as can be seen in Fig. 3a. The error is almost linearly decreased as N_A increases. The tendency is consistent with the theoretical observation where the entropy is decreased when $N_A > 5$ and continuously increases. From Fig. 3b, it turns out that the error has no linear relation with N_A in the disorder system. The theoretical calculation of the entropy in the disorder system also has no clear relation with N_A among different evolution times. The extra disorder term in XY Hamiltonian may lead to the irregular increase of the entropy. We note that since the number of evolution times is highly limited, the measurement records obtained from different quantum states are not enough to make a precise estimation, especially when the system is half-partitioned. Entropy estimation is a highly important problem in quantum many-body physics. Our work presents the first try based on supervised learning to demonstrate the feasibility of only using randomized measurement records. However, the performance of the model can be further improved by increasing the samples measured from different quantum states. We mention that supervised entropy estimation may be physical system-specific i.e., the different systems may have to train different

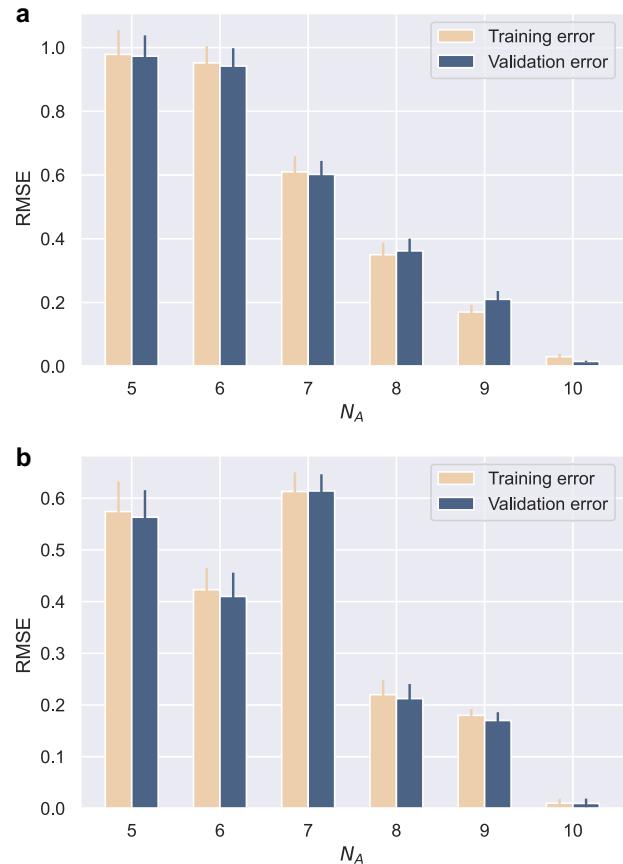


Fig. 3 The RMSE of the MLCP model in estimating the second-order Rényi entropy of varied subsystem sizes. **a** The model is trained with the sub-dataset of category C_1 given $N_U=512, N_M=300$. **b** The model is trained and validated with the sub-dataset of category C_2 given the same measurement settings. The error bar denotes the standard variance of repeated runs with different randomization.

models. But recently, the transfer learning approach⁴⁴ has been demonstrated to alleviate this shortcoming.

As for estimating the maximum fidelity of two quantum states, the error is smaller than estimating the entropy both for clean and disorder systems as Fig. 4a, b shows. The true value of the maximum fidelity is distributed in a relatively compact range according to its definition which naturally leads to small losses. In addition, a larger forward partition has a larger loss in both clean and disorder systems. Moreover, in a disordered system, the loss increase is not as regular as in a clean system. Since the maximum fidelity estimation adopts the Siamese network structure, some inherent noises can be canceled out which may be beneficial for precise estimation. The maximum fidelity estimation aims to present a cross-correlation property of two quantum states rather than characterizing the inherent attribute of the quantum state. Therefore, the maximum fidelity estimation may be empirically easier to be accurately estimated compared to estimating the entanglement entropy based on the MLCP model in a supervised learning fashion. The validation loss in certain cases is smaller than the training loss is highly likely due to the relatively large dropout and sample noise as we speculate in the identification phase.

The RMSE losses evaluate the prediction precision of the model when estimating physical properties. We further conduct simulations on clean systems to generate measurement records for testing with $N_U=512, N_M=300$. The time evolutions are chosen as $t=\{0, 3, 5\}$ ms. When estimating \mathcal{F}_{\max} , the measurements are

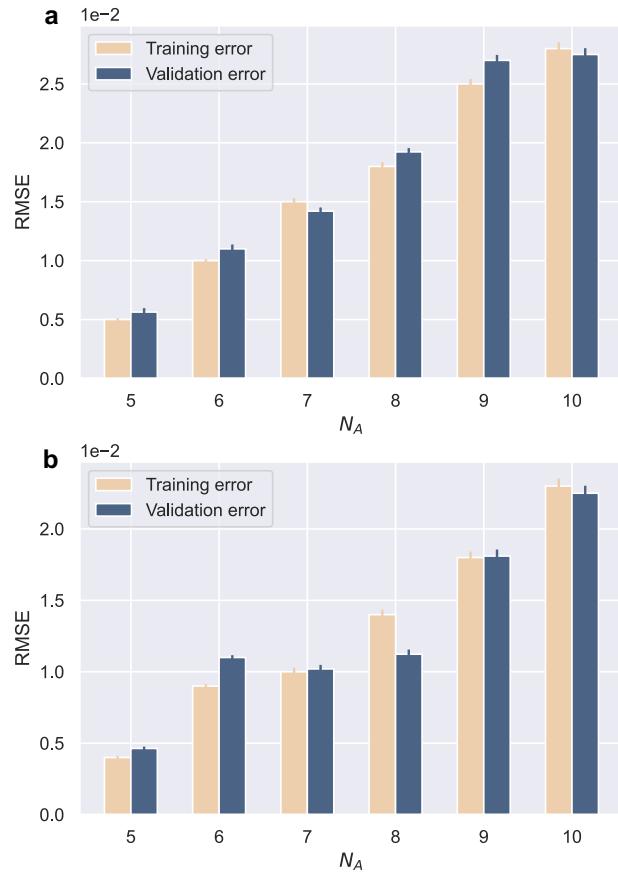


Fig. 4 The numerical performance of the MLCP model in estimating the maximum fidelity of varied subsystem sizes. **a** The model is trained based on the samples in C_1 . **b** The model is trained based on the samples in C_2 . In **(a, b)**, the measurement setting are set to be $N_U = 512, N_M = 150$. The error bar denotes the standard variance of repeated runs with different randomization.

divided into two equal parts to mimic two independent measurements from two QSMs. The estimated entropy holds a similar overall tendency compared with its theoretical value although the error is not negligible when $N_A = 5, 6, 7$ as Fig. 5a shows. Note that $t = 0$ ms, the gap between estimation and true value is relatively smaller than other times. The theoretical entropy of $t = 0$ and $N_A = 10$ should be zero, however, the effect of the imperfections leads to nonzero entropy. The variation tendency of entanglement entropy is still useful in scenarios of qualitative assessment. Estimated \mathcal{F}_{\max} also shows an overall consistent tendency compared with the conventional method as Fig. 5b shows. The merit of the MLCP model for predicting physical properties is more efficient at the inference phase compared to conventional methods. The estimation accuracy is limited by the training samples. Besides, the model is system-specific and the transferability can enhance its applicability.

To investigate the performance of MLCP in large system sizes, we simulate the time evolution $t = 0.5, 1, 3$ for $N = 20, 30$. The time slot is $\Delta t = 0.01$ and for each time slot, we measure the matrix product state and obtain the randomized samples. The system size of $N = 30$ is chosen with uniformly sampled couplings and $N = 20$ is chosen with constant couplings. The randomized measurements are chosen to be $N_U = 200, N_M = 500$ for each quantum state. We train the model for $N = 20, 30$ separately. We only concentrate on estimating the second-order Rényi entropy in the short-range Hamiltonian to explore the performance of MLCP. To study the impact of measurement samples on the estimation performance, we present three measurement settings with

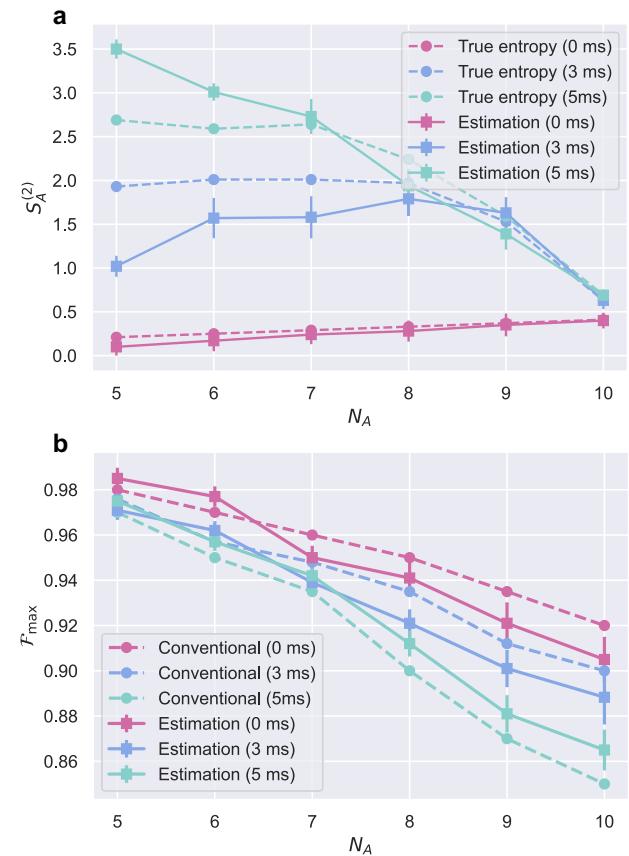


Fig. 5 The estimation performance of the MLCP model when evaluating independent measurements with $t = 0, 3, 5$ ms. **a** The entropy estimation compared with its theoretical value. **b** The maximum fidelity compared with conventional method¹⁹. The error bar denotes the standard variance over repeated simulations.

$N_M = 5, 50, 500$ to study the impact of the number of samples on the final estimation accuracy as Fig. 6 shows. It can be found in Fig. 6a–c that when increasing the evolution time from $0.5 \rightarrow 3$, the estimation accuracy is decreased as the quantum state in longer time evolution has larger entanglement entropy. Although the total number of training samples is increased accordingly, the validation error has a gap compared with smaller evolution times. We require further increasing the training samples when estimating the entanglement entropy in longer evolution times to enhance the estimating accuracy. We also found that $N_U N_M = 3 \times 10^5$ measurement samples also achieve satisfactory mean absolute error (MAE) as Fig. 6c shows. We note that our training samples are still exponentially scaled as $O(2^N)$ which also implies approximate informationally-complete randomized measurements. When we increase N_M regardless of evolution times and system sizes, all the estimation errors of MLCP prediction show decreased tendency demonstrating that more randomized measurement samples can enhance the estimation accuracy. This observation is consistent with the results in ref. ³². The standard method can achieve highly accurate estimations with the number of samples super-exponentially scaled. However, in large-scale systems, collecting sufficiently large samples is resource-intensive. On the other hand, the collected number of samples also can be used to reconstruct the quantum state such that the estimation accuracy is notably high. Therefore, in case one dose does not require high precision estimation, MLCP can provide coarse-grained estimation without using a strictly exponential number of measurement samples such as state tomography.

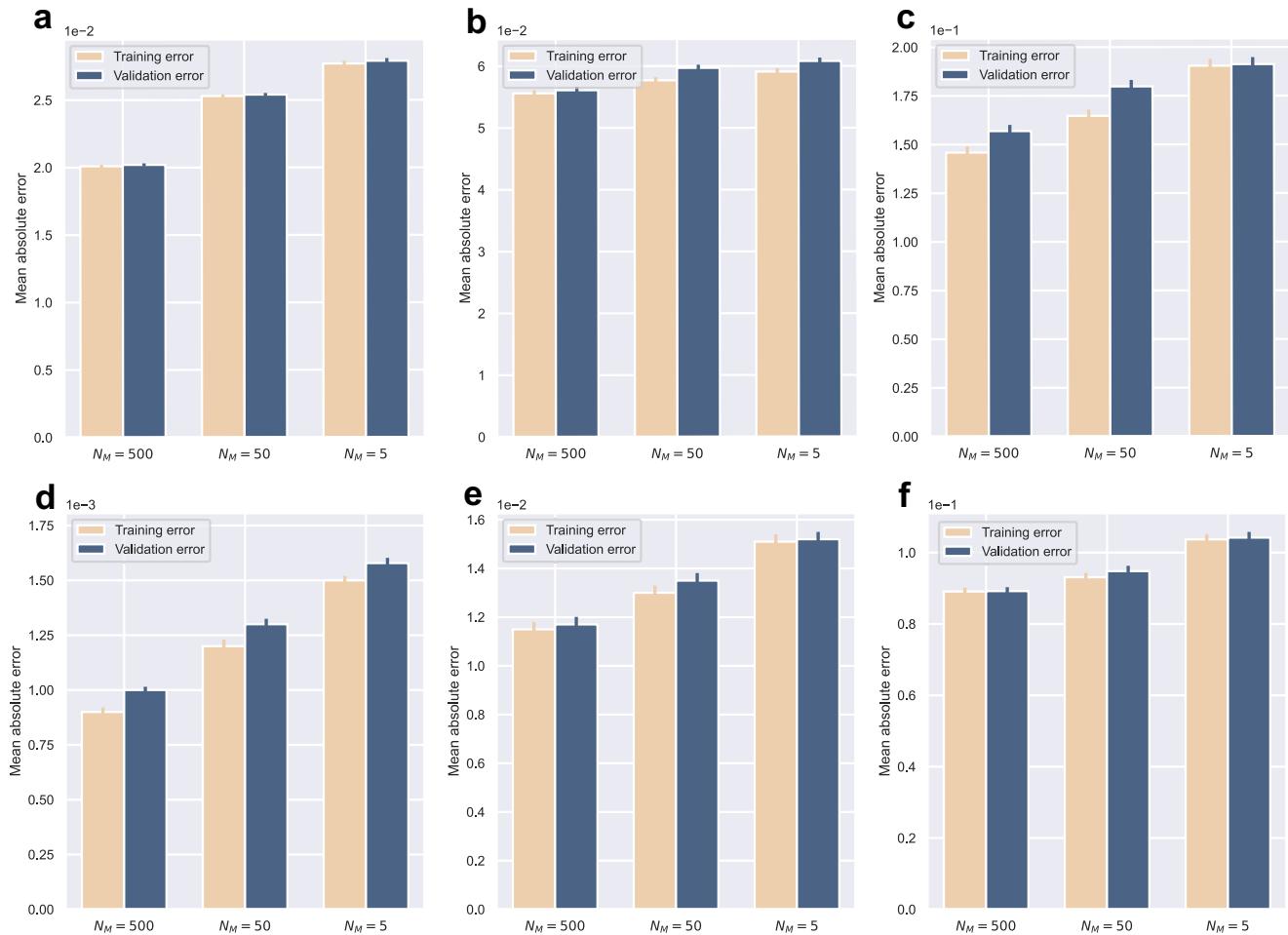


Fig. 6 The mean absolute error (MAE) of different evolution times for estimating the second-order Rényi entropy of half-partition subsystem. a–c are the training and validation MAE with $N = 20, J = 1$, and evolution times $t = \{0.5, 1, 3\}$, respectively. **d–f** are the training and validation MAE with $N = 30, J \sim U(0, 1)$, and evolution times $t = \{0.5, 1, 3\}$, respectively. The error bar denotes the standard variance of repeated runs with different randomization.

As for considering uniformly sampled couplings in the Heisenberg model, the entanglement entropy growth rate is not large as the case in constant couplings. The quantum states have a relatively smaller entanglement entropy compared to the state in constant coupling when we observe at the same evolution time. The training samples in $N = 30$ are still set to be as $N_U N_M = 10^5 \sim 10^7$. The training and validation error still show satisfactory precision as Fig. 6d-f shows. We conceive that although the number of qubits increases, the quantum state has a smaller entanglement entropy such that MLCP is easier to learn the hidden relations in bit strings. Moreover, the consecutive quantum states in neighboring time evolutions have more similar structures. MLCP can make use of this information in the time dimension. These two possibilities lead to the fact that MLCP does not require strictly exponentially scaled number measurement samples to estimate entanglement entropy. We note that in ref. 45, GHZ state is measured to study the law of training samples v.s. the number of qubits. The GHZ state is maximally entangled such that its scaling law can be an upper bound of other entangled states with smaller entanglement entropy. Besides, standard methods do not consider estimating the entanglement entropy with measurement samples by using the time-evolved data cooperatively. MLCP can collect all the evolved training samples to estimate the property which is likely to extract the information of quantum states in the time dimension. Therefore, our method is of more practical interest that aims to identify the quantum phases of matter and

estimate the properties in a unified framework based on the end-to-end MLCP model.

DISCUSSION

In summary, we propose an end-to-end supervised learning protocol called MLCP by integrating the advanced LSTM and Transformer encoder models. The protocol can be readily extended to a different number of qubits by increasing the sequence length of the recurrent model. To train the model, we construct a large-scale dataset by collecting randomized measurement outcomes from different types of QSMs, including the one-dimensional long-range XY model and short-range Heisenberg model.

In the identification phase, we conduct a binary identification to distinguish whether the QSM is clean or disordered and two QSMs from different field couplings. The numerical results demonstrate that identifying the clean or disorder is easier than two different field couplings. The accuracy of the former reaches up to 85% and the latter reaches 68%. More significantly, by increasing repeated measurements for one local randomized unitary setting, the identification accuracy can be polynomially enhanced. On the contrary, increasing the number of local random unitary leads to a linear accuracy enhancement. The results provide a straightforward witness that classical machine learning is not sample-efficient in

handling quantum simulation data which may be viewed as an indirect implication of quantum advantage.

In properties estimation for small system size ($N = 10$) in long-range model, the overall variation tendency of the estimated entanglement entropy and maximum fidelity is coincident with the theoretical calculation. The machine-learning estimation for maximum fidelity achieves competitive performance with the conventional method. The accuracy of the physical properties estimation can be further improved by controlling the smaller evolution time of QSMs to enrich the measurement samples. When estimating the entanglement entropy in short-range large system size ($N = 20, 30$), the estimation accuracy still achieves satisfactory precision without strictly using an exponential number of training samples. We deliberate that the short-range system has a smaller entanglement entropy growth rate, and the quantum states are easier to be learned. Besides, the additional information on quantum states in the time dimension can be further leveraged to reduce the number of training samples. The two methods of certification generate weak-to-strong information gain, which is flexible in the practical certification of QSMs.

Learning from measurement bit strings by extracting their hidden patterns can be used to identify or predict quantum states or properties. Handcrafted features in ref. ⁴⁶ may be incorporated into the MLCP model and further enhance its capability in quantum certification such as reducing measurement samples. We also remark that the MLCP model does not rely on specific measurement schemes such as randomized Haar measurements, randomized Pauli measurements, and different types of informationally-complete POVM (IC-POVM). We notice that recent SIC-POVM measurement is experimentally realized to achieve a single setting POVM by using a four-level quantum state, which can dramatically accelerate the data acquisition speed⁴³. The different scheme has their merits and demerits. (De)randomized measurement is relatively easier to be implemented in actual experiments. IC-POVM however is necessary for reconstructing quantum states but also costs a lot of resources. The described measurement schemes based on post-processing cannot overcome the quantum-classical separation gap⁴⁴. However, machine learning is highly likely to reduce the practical measurement samples by using an important sampling technique⁴⁵. Our method can incorporate the identification and estimation phases into a unified framework to identify the quantum phases of matter and estimate core-grained properties estimation in large-scale system size. Our model can also be extended to the AI cloud center to provide accessible services for local QSMs. In future work, we will continue to investigate the feasibility of our model in two-dimensional many-body QSMs.

METHODS

Quantum simulation dataset generation

Our work mainly makes use of a supervised learning approach to train the numerically simulated quantum simulation dataset. We generate a large-scale dataset and the number of samples is 1,824,768. The total time we cost to simulate the randomized measurement data is 30 days. There are 5 days consumed in profiling the raw data. The profiled data can accelerate the training process of the standard machine-learning library. We then randomly divide the whole dataset into two independent subsets: a training set and a validation set. The ratio of the number of samples in the training set to the validation set is 99:1, which is commonly used in the large-scale dataset. The quantum simulation of the XY Hamiltonian model is accomplished by numerically solving the quantum master equation. We simulate the quantum evolution of 10 qubits. The larger size of qubits is not supported in naive density matrix representation in the evolution by using the quantum master equation. Different evolution time t is simulated

to enrich the dataset. To reflect the practical quantum simulation experiment, we consider the state preparation and measurement (SPAM) error. Specifically, the initial Néel state is not perfectly prepared and has errors. In addition, the measurement is still not perfect. We absorb the local unitary rotation error into the local depolarization error and then regard the randomized measurement as the perfect one. The state evolution error in our implementation consists of the spin-flip error and the spontaneous emission error. These two errors are the main source of error in spin quantum simulation. Finally, there are mainly two types of randomized measurements. In analog quantum simulation, a randomized Haar measurement operator from CUE is usually used. However, we also provide a short introduction to the classical shadow method for randomized measurement, which is commonly used in digital quantum simulation. The detailed dataset construction can be found in Supplementary Note 3B. More description about the physical model, the quantum master equation characterization, the errors modeling and the randomized measurements can be found in Supplementary Note 1A, Note 2A, and 2B. In addition, the numerical data generation process is implemented based on the Qutip quantum package⁴⁷. The reference python code can be found in ref. ²⁹.

Machine-learning model design and training

After the training and validation dataset are constructed, we build the LTL model and then train the model by using the constructed dataset. We divide all the data samples into four main categories. Then we further regard the data samples as the categories of disorder or clean and J370 and J420. We view the certification process as a binary classification problem. The categories can be increased by adding more different physical Hamiltonian models. The backbone of the LTL model is displayed in Fig. 7. We use the LSTM as the trainable embedding layer to map the discrete measurement records into continuous space as shown in Fig. 7a shows. Subsequently, the embedded data is processed by the transformer encoding layer. Finally, the encoder output is further processed by an LSTM layer. The learned encoder can be regarded as the pre-trained model which is trained by the measurement data under various types of Hamiltonians and different evolution times. The quantum simulation certification can be viewed as an identification problem or a physical property estimation problem solved by the regression process based on the pre-trained LTL model as shown in Fig. 7b. We remark that the pre-training process can learn much useful prior information and may render the property estimation process more efficient. The chosen pre-training criteria are not unique and can be engineered more universally for downstream property estimations tasks. A detailed description of the machine-learning model and the realization methods can be found in Supplementary Note 3. The machine-learning hyperparameters and model selection can be found in Supplementary Note 5.

Matrix product states

Matrix product states (MPS) are well summarized in the literature^{48,49}, so we only introduce the important aspects for discussion in this work. MPS methods can approximate the physical state by a linear tensor network with N tensors one at each site. The MPS approximation is controlled by the maximum bond dimension D , which limits the maximum entanglement degree between continuous subsystems. For example, an MPS with maximum dimension D can only represent a quantum state with entanglement entropy with $S = \log(D)$ for the bipartition of two connected subsystems.

For the unitary time evolution, we use time evolution block-decimation (TEBD) to simulate the one-dimensional quantum many-body systems, characterized by at most nearest-neighbor interactions. A detailed clarification of MPS time evolution can be

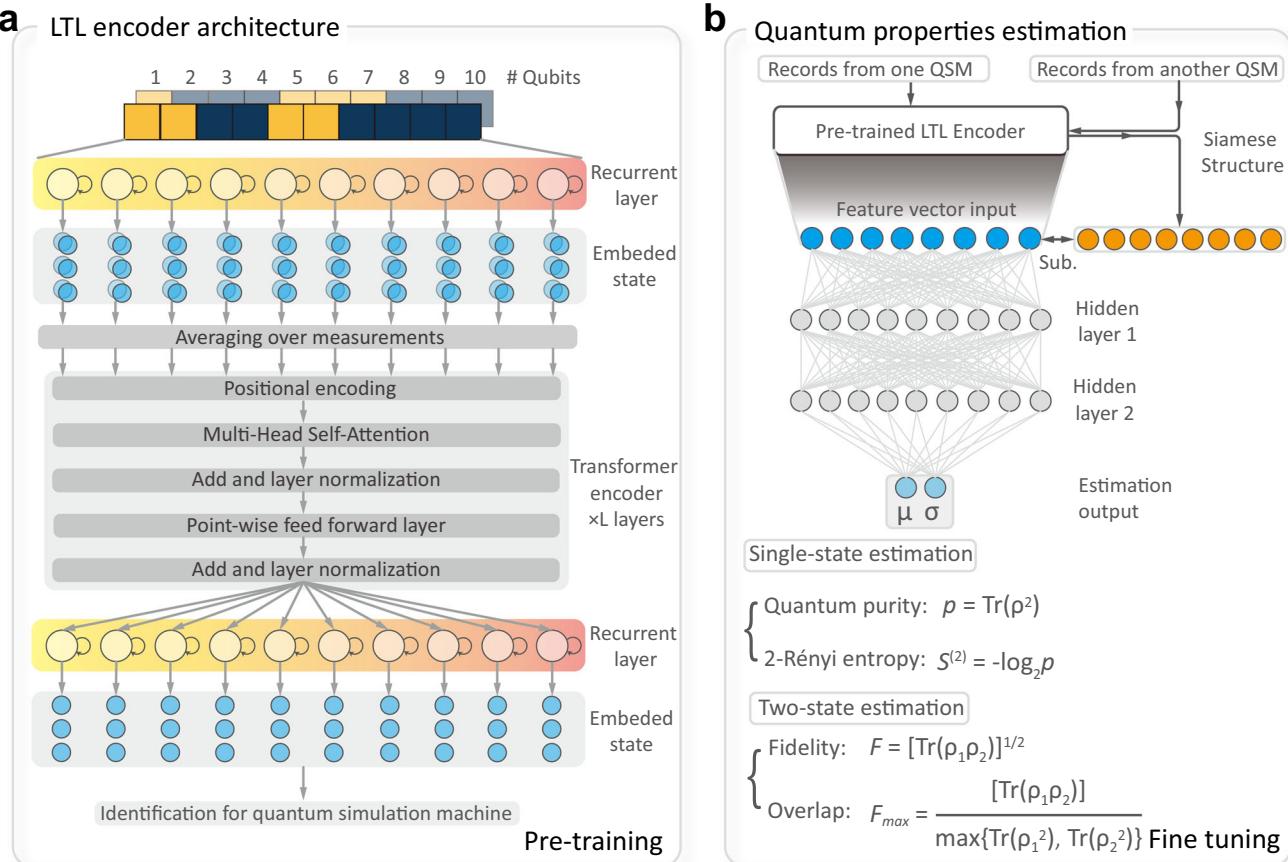


Fig. 7 The backbone of the LSTM-Transformer-LSTM (LTL) model. **a** The LTL encoder works as follows: raw measurement records are split into single-shot records and each single-shot record is embedded through a recurrent layer. Then all embedded states are aggregated and averaged to generate a complete embedding. Subsequently, the embedding is processed via the L transformer encoder layer to further extract the latent pattern. Finally, the output of the transformer layer is fed into a recurrent layer to obtain the final embedding used for quantum downstream tasks. **b** The quantum properties such as purity and fidelity can be estimated by firstly encoding the raw records into the representative high-dimensional embedding and then passing the embedding into an artificial neural network (ANN) to accomplish the estimation. The two-state estimation requires two records of QSMs from different locations and times.

found in Supplementary Note 1. The one-dimensional Heisenberg model is suitable simulated based on the TEBD method. To generate the randomized measurement samples, we require sampling from the MPS wavefunction. Before generating statistically independent samples of a given MPS wavefunction on a computational basis, we require applying local site operations to each site tensor. Then we follow a Monte Carlo process to generate the bit strings. (1) Choose an arbitrary site i among the N unprojected sites of the normalized MPS and obtain the diagonal elements of the single site reduced density matrix $p_0^{(i)}$ and $p_1^{(i)}$. (2) Generate a random number $r \in [0, 1]$ and select $|0\rangle$ if $r \leq p_0^{(i)}$ otherwise $|1\rangle$. Applying the single-site projector for the selected state and normalizing the remaining MPS with $N - 1$ unprojected sites to the value of $p_{0|1}^{(i)}$ depending on the randomly selected state. (3) Repeat the process until all sites are projected. We note the sampling procedure can guarantee each sample is drawn anew and not from a Markov chain seeded by the previous sample. Therefore, the samples have no auto-correlation. The source code in ref. ²³ also provides different measurement schemes such as IC-POVM and randomized measurements for matrix product states. Our sampling procedure is motivated by the provided source code in ref. ⁵⁰. For three typical time evolutions, we cost 6 days to generate the randomized measurement bit strings.

DATA AVAILABILITY

The quantum simulation data for the short-range Heisenberg model are available at <https://github.com/XiaoTailong/short-range-Heisenberg-model>.

CODE AVAILABILITY

The code to generate the long-range XY model is available at https://github.com/TiffBrydes/Renyi_Entanglement_Entropy/tree/v0. The code to preprocess the dataset and analyze the experimental results is available from the corresponding author on reasonable request.

Received: 1 June 2022; Accepted: 7 November 2022;
Published online: 28 November 2022

REFERENCES

1. Bennett, C. H. & DiVincenzo, D. P. Quantum information and computation. *Nature* **404**, 247–255 (2000).
2. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
3. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).
4. Giovannetti, V., Lloyd, S. & Maccone, L. Quantum metrology. *Phys. Rev. Lett.* **96**, 010401 (2006).
5. Xiao, T., Huang, J., Fan, J. & Zeng, G. Continuous-variable quantum phase estimation based on machine learning. *Sci. Rep.* **9**, 1–13 (2019).

6. Xiao, T., Fan, J. & Zeng, G. Parameter estimation in quantum sensing based on deep reinforcement learning. *npj Quantum Inf.* **8**, 2 (2022).

7. Lloyd, S. Universal quantum simulators. *Science* **273**, 1073–1078 (1996).

8. Xiao, T., Bai, D., Fan, J. & Zeng, G. Quantum Boltzmann machine algorithm with dimension-expanded equivalent hamiltonian. *Phys. Rev. A* **101**, 032304 (2020).

9. Zhong, H.-S. et al. Phase-programmable Gaussian boson sampling using stimulated squeezed light. *Phys. Rev. Lett.* **127**, 180502 (2021).

10. Lund, A. P., Bremner, M. J. & Ralph, T. C. Quantum sampling problems, boson-sampling and quantum supremacy. *npj Quantum Inf.* **3**, 15 (2017).

11. Bouland, A., Fefferman, B., Nirkhe, C. & Vazirani, U. On the complexity and verification of quantum random circuit sampling. *Nat. Phys.* **15**, 159–163 (2019).

12. Eisert, J. et al. Quantum certification and benchmarking. *Nat. Rev. Phys.* **2**, 382–390 (2020).

13. Wang, J. et al. Experimental quantum hamiltonian learning. *Nat. Phys.* **13**, 551–555 (2017).

14. D'Ariano, G. M., Paris, M. G. & Sacchi, M. F. Quantum tomography. *Adv. Imaging Electron Phys.* **128**, 206–309 (2003).

15. Brandao, F. G. Quantifying entanglement with witness operators. *Phys. Rev. A* **72**, 022310 (2005).

16. Magesan, E., Gambetta, J. M. & Emerson, J. Scalable and robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.* **106**, 180504 (2011).

17. Zhu, D. et al. Cross-platform comparison of arbitrary quantum states. *Nat. Commun.* **13**, 6620 (2022).

18. Carrasco, J., Elben, A., Kokail, C., Kraus, B. & Zoller, P. Theoretical and experimental perspectives of quantum verification. *PRX Quantum* **2**, 010102 (2021).

19. Elben, A. et al. Cross-platform verification of intermediate scale quantum devices. *Phys. Rev. Lett.* **124**, 010504 (2020).

20. Greganti, C. et al. Cross-verification of independent quantum devices. *Phys. Rev. X* **11**, 031049 (2021).

21. Fanizza, M., Mari, A. & Giovannetti, V. Optimal universal learning machines for quantum state discrimination. *IEEE Trans. Inf. Theory* **65**, 5931–5944 (2019).

22. Torlai, G. et al. Neural-network quantum state tomography. *Nat. Phys.* **14**, 447–450 (2018).

23. Carrasquilla, J., Torlai, G., Melko, R. G. & Aolita, L. Reconstructing quantum states with generative models. *Nat. Mach. Intell.* **1**, 155–161 (2019).

24. Sharir, O., Levine, Y., Wies, N., Carleo, G. & Shashua, A. Deep autoregressive models for the efficient 601 variational simulation of many-body quantum systems. *Phys. Rev. Lett.* **124**, 020503 (2020).

25. Carleo, G. & Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017).

26. Cha, P. et al. Attention-based quantum tomography. *Mach. Learn.: Sci. Technol.* **3**, 01LT01 (2021).

27. Zhang, X. et al. Direct fidelity estimation of quantum states using machine learning. *Phys. Rev. Lett.* **127**, 130503 (2021).

28. Berkovits, R. Extracting many-particle entanglement entropy from observables using supervised machine learning. *Phys. Rev. B* **98**, 241411 (2018).

29. Brydges, T. et al. Probing rényi entanglement entropy via randomized measurements. *Science* **364**, 260–263 (2019).

30. Jurcevic, P. et al. Quasiparticle engineering and entanglement propagation in a quantum many-body system. *Nature* **511**, 202–205 (2014).

31. Zyczkowski, K. & Kus, M. Random unitary matrices. *J Phys. A: Math. Gen.* **27**, 4235 (1994).

32. Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16**, 1050–1057 (2020).

33. Chen, S., Cottler, J., Huang, H.-Y. & Li, J. Exponential separations between learning with and without quantum memory. In *2021 IEEE 62nd Annual Symp. on Foundations of Computer Science (FOCS)*, 574–585 (IEEE, 2022).

34. Jaderberg, B. et al. Quantum self-supervised learning. *Quantum Sci. Technol.* **7**, 035005 (2022).

35. May, A., Zhang, J., Dao, T. & Ré, C. On the downstream performance of compressed word embeddings. In *Advances in Neural Information Processing Systems*, Vol. 32, 11782 (NeurIPS, 2019).

36. Torrey, L. & Shavlik, J. Transfer learning. In: *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques* (eds. Soria, E., Martin, J., Magdalena, R., Martinez, M. & Serrano, A.), 242–264 (IGI Global, 2010).

37. Nir, A., Sela, E., Beck, R. & Bar-Sinai, Y. Machine-learning iterative calculation of entropy for physical systems. *Proc. Natl. Acad. Sci. USA* **117**, 30234–30240 (2020).

38. Koch, G. et al. Siamese neural networks for one-shot image recognition. In *ICML Deep Learning Workshop*, Vol. 2, 0 (Lille, 2015).

39. Liu, W., Liu, Z., Rehg, J. M. & Song, L. Neural similarity learning. In *Advances in Neural Information Processing Systems*, Vol. 32 (NeurIPS, 2019).

40. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).

41. Huang, H.-Y. et al. Quantum advantage in learning from experiments. *Science* **376**, 1182–1186 (2022).

42. Huang, H.-Y., Kueng, R., Torlai, G., Albert, V. V. & Preskill, J. Provably efficient machine learning for quantum many-body problems. *Science* **377**, eabk333 (2022).

43. Stricker, R. et al. Experimental single-setting quantum state tomography. *PRX Quantum* **3**, 040310 (2022).

44. Gresch, A., Bittel, L. & Kliesch, M. Scalable approach to many-body localization via quantum data. Preprint at <https://doi.org/10.48550/arXiv.2202.08853> (2022).

45. Rath, A., van Bijnen, R., Elben, A., Zoller, P. & Vermersch, B. Importance sampling of randomized measurements for probing entanglement. *Phys. Rev. Lett.* **127**, 200503 (2021).

46. Sotnikov, O. et al. Certification of quantum states with hidden structure of their bitstrings. *npj Quantum Inf.* **8**, 1–13 (2022).

47. Johansson, J. R., Nation, P. D. & Nori, F. Qutip: an open-source python framework for the dynamics of open quantum systems. *Comput. Phys. Commun.* **183**, 1760–1772 (2012).

48. da Silva, M. P., Landon-Cardinal, O. & Poulin, D. Practical characterization of quantum devices without tomography. *Phys. Rev. Lett.* **107**, 210404 (2011).

49. Paeckel, S. et al. Time-evolution methods for matrix-product states. *Ann. Phys.* **411**, 167998 (2019).

50. Scholl, P. et al. Quantum simulation of 2d antiferromagnets with hundreds of Rydberg atoms. *Nature* **595**, 233–238 (2021).

ACKNOWLEDGEMENTS

The authors thank the fruitful discussions with Dongyun Bai and Xinliang Zhai. The authors also thank the useful sharing and guidance from Tiff Brydges and Benoit Vermersch on their researches and some software usages. This work is supported by National Natural Science Foundation of China (Grant Nos. 61701302 and 61631014).

AUTHOR CONTRIBUTIONS

T.L.X. developed the simulation and implemented the algorithms, T.L.X., J.Z.H., and H.J.L. elaborated on the Hamiltonian framework of quantum simulation and the certification protocol, T.L.X. and J.P.F. operated the machine-learning analysis, and G.Z. conceived the research. All the authors discussed and contributed to the writing of the manuscript.

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/s41534-022-00649-6>.

Correspondence and requests for materials should be addressed to Guihua Zeng.

Reprints and permission 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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.