Article

Machine learning on quantum experimental data toward solving quantum many-body problems

Gyungmin Cho [1] & Dohun Kim [1]

Advancements in the implementation of quantum hardware have enabled the acquisition of data that are intractable for emulation with classical computers. The integration of classical machine learning (ML) algorithms with these data holds potential for unveiling obscure patterns. Although this hybrid approach extends the class of efficiently solvable problems compared to using only classical computers, this approach has been only realized for solving restricted problems because of the prevalence of noise in current quantum computers. Here, we extend the applicability of the hybrid approach to problems of interest in many-body physics, such as predicting the properties of the ground state of a given Hamiltonian and classifying quantum phases. By performing experiments with various error-reducing procedures on superconducting quantum hardware with 127 qubits, we managed to acquire refined data from the quantum computer. This enabled us to demonstrate the successful implementation of theoretically suggested classical ML algorithms for systems with up to 44 qubits. Our results verify the scalability and effectiveness of the classical ML algorithms for processing quantum experimental data.

Progress in information storage and processing techniques[1–5] has given rise to the generation of large amounts of data, and the use of machine learning (ML) to process these data is being actively explored in biology[6], chemistry[7], and physics[8]. Areas of potential application of ML in physics include the study of many-body physics. One of the interesting problems is the prediction of the ground-state properties of a given Hamiltonian, for example the electronic structure Hamiltonian. The other intriguing problem entails the exploration of the boundaries between different quantum phases, which may enable the identification of exotic quantum phases such as high-temperature superconductivity. Widely used classical algorithms[9,10], however, in spite of their many successful applications, have fundamental limitations in approximating strongly interacting systems.

Although quantum computers are expected to excel at solving quantum chemistry and many-body physics problems, current devices are still prone to errors, which compromise the accuracy of results. While quantum error correction is believed as a solution[11,12], the large-scale operation is not immediately realizable[13]. Hybrid approaches in which classical computers are combined with quantum computers as to circumvent this issue have been introduced[14].

Research aimed at broadening the utility of the hybrid approach leads to quantum state learning methods such as classical shadow[15–17] which is a succinct classical representation of a quantum state. The ability to convert quantum states into classical forms naturally allows for their use as data in classical ML, and recent research has focused on theoretical analyses of the performance in many-body physics applications[18–21]. One advantage in this direction is that it allows well-developed classical devices and ML techniques to be leveraged. Additionally, under widely believed complexity conjectures, the combined use of data from quantum computers and learning on classical computers can solve, in a computationally efficient manner, some problems that are challenging for non-ML algorithms relying solely on classical devices[18,22]. Despite the aforementioned advantages, prevailing errors in data from quantum computers limit the range of

[1]Department of Physics and Astronomy, and Institute of Applied Physics, Seoul National University, Seoul 08826, South Korea.
✉e-mail: km950501@snu.ac.kr; dohunkim@snu.ac.kr

problems that are addressable and challenge the scalability as the system size increases[23–25]. Thus, appropriate quantum error mitigations (QEMs) are necessary during and after data acquisition.

Here, we experimentally verify the applicability of the hybrid approach to problems of interest in many-body physics. Previous research[18,19] obtained training data for ML from tensor network algorithms on classical computers. These approaches face limitations as system sizes increase, particularly beyond two dimensions, highlighting the utility of quantum computers for generating training data. Our experiments demonstrate that classical ML on data from quantum computers is effective not only in one-dimensional but also in two-dimensional many-body physics problems. Experimentally, prior works[17] have used classical shadows primarily for estimating observables of a given quantum state. However, building on the rigorous algorithm with various error-reducing procedures applied to the training data, we have successfully used classical shadows obtained from quantum computers for ML tasks, extending their applications beyond predicting physical observables for a given state.

We implement classical ML algorithms[18] to solve problems related to the prediction of ground state properties and the classification of quantum phases, as illustrated in Fig. 1. These problems can be considered as regression and classification in traditional ML, respectively. For regression, because the prediction of accurate values is required, various QEMs were used, which resulted in the application of accurate ML models to a 12-qubit system. Regarding the classification task, we expanded on the previous work[26,27] of distinguishing the Symmetry Protected Topological (SPT) phase[28] by performing the classification task in a more general setting and increasing the system size up to 44 qubits. Also, with the help of a measurement-assisted state preparation method[29], which enabled the generation of suitable training data, we demonstrated successful phase classification between topologically ordered and trivial phases of a system comprising as many as 25 qubits, thereby confirming the applicability of the scalable ML algorithms. We conducted our experiments on a device consisting of superconducting qubits provided by IBM. Detailed error statistics on the hardware are presented in the Supplementary Note 1.

Results

Classical shadow

Despite many efforts to reduce the sample complexity for quantum state tomography (QST), exponential scaling of the sample complexity in terms of the system size (n) is unavoidable[30]. To circumvent the exponentially increasing sample complexity, research focusing on predicting physical quantities instead of the full description of a quantum state was introduced[31]. Subsequent studies presented the more experimentally friendly protocol of classical shadows[15]. In addition to linear functions $f_{\text{linear}}(\rho) = \text{Tr}(O\rho)$, nonlinear functions such as $f_{\text{nonlinear}}(\rho) = \text{Tr}(O\rho^{\otimes k})$ can also be estimated by the classical shadow[27]. Utilization of the classical shadow as the data for ML would be expected to enable the ML model to learn the nonlinear properties of the state[18,22].

We obtained the classical shadows of the state by applying a unitary transformation sampled from a random unitary ensemble, followed by the measurement. By repeating this process T times, we can obtain $S_T(\rho) = \{(b^{(t)}, U^{(t)})\}_{t=1}^{T}$ as the experimental results, where $b^{(t)} \in \{0, 1\}^n$ is the measurement outcome, $U^{(t)} = \otimes_{i=1}^{n} U_i^{(t)}$ where each $U_i^{(t)}$ is sampled from the Haar measure over the unitary group $\mathbb{U}(2)$[32] instead of widely used Cl(2) (Clifford group on single qubit). As a result, the unbiased estimator $\hat{\sigma}_T(\rho)$ can be written as

$$\hat{\sigma}_T(\rho) = 1/T \sum_{t=1}^{T} \left\{ \otimes_{i=1}^{n} \left(3U_i^{(t)\dagger} \left| b_i^{(t)} \right\rangle \left\langle b_i^{(t)} \right| U_i^{(t)} - I_2 \right) \right\} \quad (1)$$

where $\mathbb{E}_{U,b}(\hat{\sigma}_T(\rho)) = \rho$ and I_2 is the 2×2 identity matrix

In practice, classical shadow estimations sometimes lead to inferior results compared to direct measurement due to the restriction on available QEMs resulting from the randomized measurements (RM)[16] for the classical shadow. Here, direct
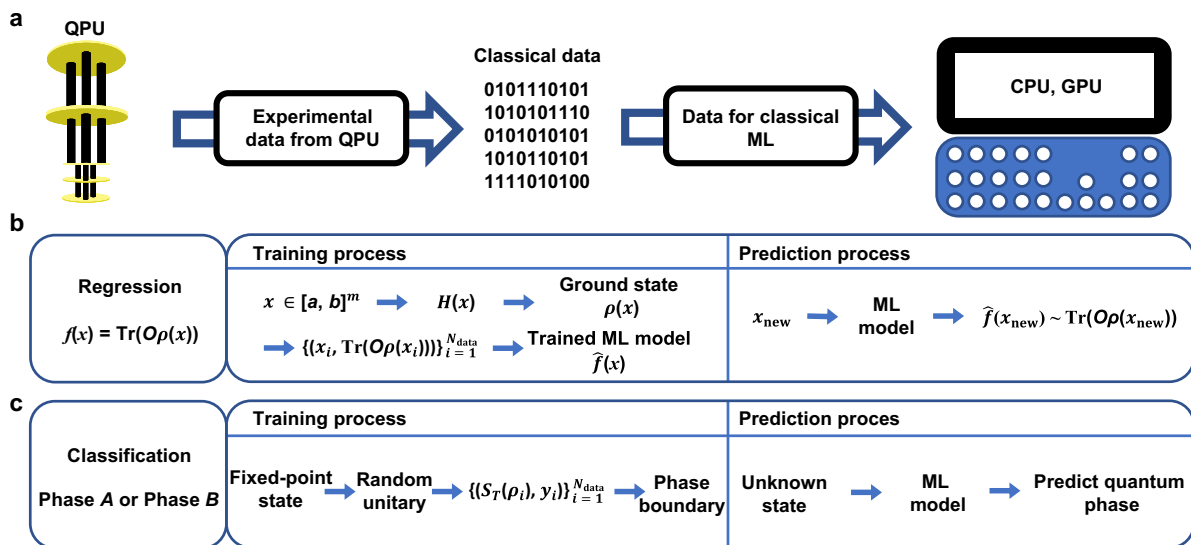
Fig. 1 | Classical machine learning (ML) on quantum experimental data.
a Conversion of the information represented in a quantum state into a classical form for the application of classical ML. The training data were the expectation values measured via the quantum computer or the classical shadows of the quantum state. **b** Predicting the properties of the ground state of a given Hamiltonian casts to regression in classical ML. The objective of the regression is to approximate the target function $f(x) = \text{Tr}(O\rho(x))$ as closely as possible by trained ML model using training data $\{(x_i, \text{Tr}(O\rho(x_i)))\}_{i=1}^{N_{\text{data}}}$. After training, when a new input x_{new} is given, the trained ML model outputs $\hat{f}(x_{\text{new}})$ as a guess for $f(x_{\text{new}})$. **c** In classification, the objective is to distinguish a particular quantum phase from others by training a ML model that gives rise to the phase boundary on a relevant space. After training, when a new state is given, the trained ML model assigns a phase of the state by considering its position relative to the phase boundary on the space.
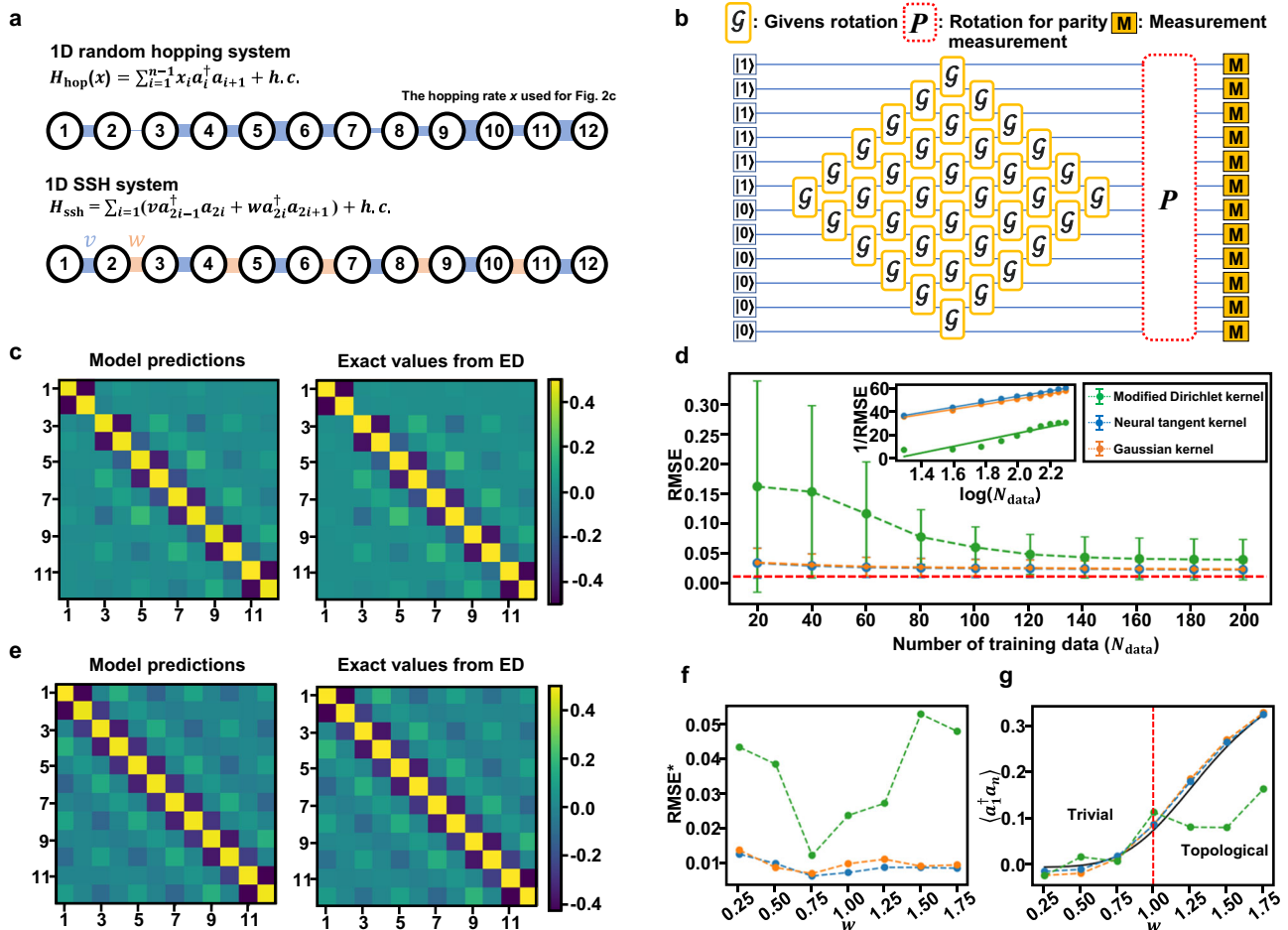
Fig. 2 | Predicting ground state properties. a Hamiltonian for random hopping system and Su–Schrieffer–Heeger (SSH) system. x_i in $H_{\text{hop}}(x)$ is sampled uniformly from [0, 2]. **b** Ground state preparation circuit. Gates labeled $\mathcal{G}$ and P represent a Givens rotation and a basis rotation for parity measurement, respectively. **c** Correlation matrix for random hopping system. Results from ML prediction using the Gaussian kernel and exact values from exact diagonalization (ED) are shown on the left and right, respectively. **d** Prediction error. Model performance is measured by the root-mean-square error (RMSE). The error bars represent the standard deviation of the RMSEs of each observable. The dotted red line means the average RMSE for the training data. Inset shows the relationship between the $\log(N_{\text{data}})$ and 1/RMSE. The solid lines indicate a linear fit to the data points. **e** Correlation matrix

for 1D SSH system ($v = 1$, $w = 1$). Results from ML predictions using Gaussian kernel and exact values from ED are shown on the left and right, respectively. **f** Prediction error. RMSE* $\left(= \sqrt{2/n(n-1)\sum_{i \geq j}^{n}(\langle a_i^\dagger a_j\rangle - \langle a_i^\dagger a_j\rangle_{\text{ML}})^2}\right)$ from different kernels while keeping $v = 1$ and varying w in 0.25 intervals from 0.25 to 1.75. **g** Edge correlation. The edge correlation $\langle a_1^\dagger a_n\rangle$ is measured between sites located at either end of a chain. The dotted red line is the known phase boundary between the trivial and topological phases. The black curve indicates exact values obtained by ED. In **d**, **f**, **g** the green, blue, and orange points each indicate the results predicted by the ML model trained by a (modified) Dirichlet kernel, neural tangent kernel, and Gaussian kernel, respectively.

measurement was used for the regression, whereas classical shadow representation obtained by randomized measurement was employed for the classification in which errors were tolerable to a certain extent, but nonlinear properties were necessary.

Case 1: predicting the properties of the ground state
In our first experiment, we aimed to predict the properties of the ground state, specifically by focusing on learning that $f(x) = \text{Tr}(O\rho(x))$ ($\rho(x)$ is the ground state of the Hamiltonian $H(x)$). By mapping an input vector x to a high-dimensional space through a feature map $\varphi : x \in R^m \to R^{m_\varphi}$, functions $f(x)$ can be approximated by $w^{\mathsf{T}}\varphi(x)$ where w is a model parameter. However, if the dimension of the feature space (m_φ) is too large, it is impractical to conduct calculations directly using feature vector $\varphi(x)$. Instead of having to process high-dimensional vectors, a relatively simple relation known as the kernel trick[33] $k(x, x') = \varphi(x)^{\mathsf{T}}\varphi(x')$ can be used. Among the many available algorithms, we used the previously studied kernel ridge regression (KRR)[18,19] with minor modifications and a closed-form expression for

predicting $f(x_{\text{new}})$ based on N_{data} samples is given by

$$\hat{f}(x_{\text{new}}) = \sum_{i=1}^{N_{\text{data}}}\sum_{j=1}^{N_{\text{data}}} k(x_{\text{new}}, x_i)(K + \lambda I)^{-1}_{ij} f(x_j) \qquad (2)$$

where λ is the hyperparameter, $K_{ij}(= k(x_i, x_j))$ is a kernel matrix, and I is the $N_{\text{data}} \times N_{\text{data}}$ identity matrix.

We selected the 1D nearest-neighbor (NN) random hopping system $H_{\text{hop}}(x)$ with 12 sites ($n = 12$) (illustrated in Fig. 2a) as a benchmark for the learning task. The number of parameters of the system increases linearly as n grows. Although it may be challenging to train the system with many parameters, after training, the ML model would have the extended interpolation regions for inferences. The ground state of $H_{\text{hop}}(x)$ was prepared on a quantum computer by using the Givens rotations[34,35] in Fig. 2b. To reduce errors, apart from diverse error mitigation methods (Dynamical Decoupling, Pauli twirling, McWeeny purification)[36,37], we implemented a parity measurement by

recompiling the circuit. Further information is provided in the Supplementary Note 3.

We obtained the expectation values corresponding to the site correlations $\langle a_i^\dagger a_j \rangle$ $(=1/4(X_i\text{-}iY_i)Z_{i+1}.. Z_{j-1}(X_j+iY_j)$ by employing a Jordan-Wigner transformation, where X_i, Y_i, and Z_i are the Pauli operators at site i), and constructed a correlation matrix of which each (i, j)-element corresponds to $\langle a_i^\dagger a_j \rangle$. We uniformly sampled the hopping rate $x \in [0, 2]^{n-1}$ and performed 20,000 measurements to obtain each $\langle a_i^\dagger a_j \rangle$. Then, we used ML with the aim of predicting the correlation matrix for a new ground state at x_{new}. We collected 200 data points from the quantum computer to train the ML model. The performance of the trained model was evaluated using 10,000 (N_{test}) test data obtained by Exact Diagonalization (ED). The correlation matrix predicted by the ML model has reasonable similarity to the exact values (Fig. 2c). As shown in Fig. 2d, the model achieved an average root-mean-square error (RMSE) $(=\sqrt{1/N_{test}\sum_{i=1}^{N_{test}}(f_O(x_i)-\hat{f}_O(x_i))^2}$ where $f_O(x) = \text{Tr}(O\rho(x)))$ of 0.0168 on the test data, and the decrease in the RMSE as the number of training data increases confirms the importance of the data in ML for predicting the ground state properties. The red dotted line represents the average RMSE of the error-mitigated training data, which becomes 0.0244 in the case of raw data. We observed the scaling relationship $\log(N_{data}) \sim \mathcal{O}(1/RMSE)$ of some trained models (inset in Fig. 2d), even in the multi-phase case, which outperforms theoretical predictions[18]. This can be attributed to the fact that the physical systems often encountered are structurally simpler, allowing ML models to predict the properties of interest more effectively[19–21]. For a detailed comparison of the performance of the ML model across different values of m and between with and without error mitigation, refer to Supplementary Note 5.

The trained ML model was subsequently used to predict the ground state properties of the Su–Schrieffer–Heeger (SSH) Hamiltonian (Fig. 2a), and as seen in Fig. 2e–g, the results confirmed the ability of the ML model to predict not only the correlation matrix but also edge correlations originating from the topological properties. We therefore succeeded in experimentally confirming one of the promising applications of the ML approach, namely that the properties of other systems of interest can be estimated without performing additional experiments. Details of the noisy circuit simulations can be found in the Supplementary Note 5.

Case 2: classifying quantum phases
In the second experiment, we classified quantum phases in many-body physics by using principal component analysis (PCA) and a support vector machine (SVM) as classical ML algorithms[18,33]. To generate data for ML, we prepared the fixed-point state of a given phase on the quantum computer and applied a local random unitary to generate different states within the same phase as the training data[38]. The advantage of this approach is that it allows for model-independent data acquisition, thereby reducing the biases in the training data[39]. With a classical shadow that contains sufficient information to compute the nonlinear properties of the state, we employed a shadow kernel[18] defined by

$$k_{shadow}(S_T(\rho),S_T(\tilde{\rho})) = \exp\{\tau/T^2 \sum_{t,t'=1}^{T} \exp[\gamma/n \sum_{i=1}^{n} \text{Tr}(\sigma_i^{(t)}\tilde{\sigma}_i^{(t')})]\} \quad (3)$$

where $\sigma_i^{(t)} = 3U_i^{(t)\dagger}|b_i^{(t)}\rangle\langle b_i^{(t)}|U_i^{(t)} - I_2$ and τ, $\gamma > 0$ are hyperparameters. Each k_{shadow} can be efficiently evaluated using $\mathcal{O}(nT^2)$ computation time. Further explanations on the shadow kernel can be found in the Supplementary Note 2.

Case 2-1: distinguishing a short-range entangled state from a trivial one
Without symmetry, both the SPT phase, having short-range entanglement, and the trivial phase exhibit trivial order, and the states from these two phases can be connected via constant-depth local unitary (LU) transformations[38] (Fig. 3a). However, in cases the applied unitary preserves a specific symmetry, the constant-depth circuit protecting the symmetry is known to be non-existent[40]. As a fixed-point state in the SPT and trivial phases, we utilized the ground state of $H_{ZXZ} = -\sum_i Z_{i-1}X_iZ_{i+1}$ and $H_X = -\sum_i X_i$, respectively, with a 44-site periodic boundary condition, and examined whether the ML model can distinguish the SPT phase protected by $\mathbb{Z}_2 \otimes \mathbb{Z}_2$ symmetry generated by $X_{even(odd)} = \prod_{i=even(odd)} X_i$ or time-reversal symmetry (TRS) $\mathcal{T} = (\prod_i X_i)K$, where $\mathbb{Z}_2$ is a second-order group and K denotes complex conjugation. Specifically, we do not assume that the system is translationally invariant when applying a symmetric random unitary[39] (Fig. 3b). This means that attempts to measure the string order parameters (SOP) to distinguish the SPT phase are highly likely to fail, as shown in Fig. 3c[27,28]. However, ML using classical shadows to generate data could increase the probability of classifying the SPT phase because it contains non-linear information about the state[41].

We obtained 20 data points for each phase in the form of classical shadows using $T = 100$, and used half of these points as training data to identify the phase boundary and the other half as test data. For classical ML, PCA with a shadow kernel was used to reduce the dimensionality of the data, and the phase boundary was obtained using SVM with a Gaussian kernel. Figure 3d shows that the ML model can distinguish between the SPT phase protected by $\mathbb{Z}_2 \otimes \mathbb{Z}_2$ symmetry or TRS with high accuracy. However, if the symmetry is not respected, the two phases can be connected by a constant depth local unitary, which would complicate classification. Figure 3e, f shows the distribution of the test data and phase boundary of the trained ML model for the $\mathbb{Z}_2 \otimes \mathbb{Z}_2$ symmetry and TRS. In the case of TRS, the symmetric local unitary gates are less restrictive than in the $\mathbb{Z}_2 \otimes \mathbb{Z}_2$ symmetry, which permits the generation of a wider variety of states from the fixed-point state to ultimately give rise to a more dispersed data distribution.

The dependency of the algorithm's computational time on n appears when calculating the shadow kernel $k^{(shadow)}(S_T(\rho_1), S_T(\rho_2))$ and is proportional to $\mathcal{O}(n)$. Therefore, even as the system size increases, the computational time only increases linearly. In addition, we found that when the number of layers of symmetric random unitary (d_{LU}) and T are the same, the larger the system size, the more clearly the quantum phase separates. Further details on this are elaborated in Supplementary Note 5.

Lastly, we attempted to distinguish the quantum phases of the $H_{Cl} = -J\sum_i Z_i X_{i+1} Z_{i+2} - h_1 \sum_i X_i - h_2 \sum_i X_i X_{i+1}$, for which multiple quantum phases exist depending on the values of h_1 and h_2 (with $J = 1$)[26,42,43]. Utilizing the same experimental data under $\mathbb{Z}_2 \otimes \mathbb{Z}_2$ symmetry, we trained a ML model and attempted to distinguish between the SPT and trivial phases among a total of 40 test data points from H_{Cl}−20 from each phase−obtained by density matrix renormalization group (DMRG) simulation[44,45], as shown in Fig. 3g. These results showed that all test data points, except for one, were correctly classified into appropriate phase, demonstrating the applicability of ML with the classical shadow.

Case 2-2: distinguishing a long-range entangled state from a trivial one
Topologically ordered states, having long-range entanglement, directly prepared on quantum devices have been confirmed by measuring the non-zero expectation value of topological string operators[46] or topological entanglement entropy (TEE)[47–49]. Here, we classified topologically ordered phases using the ML model trained by data. We utilized a surface code (a planar version of the
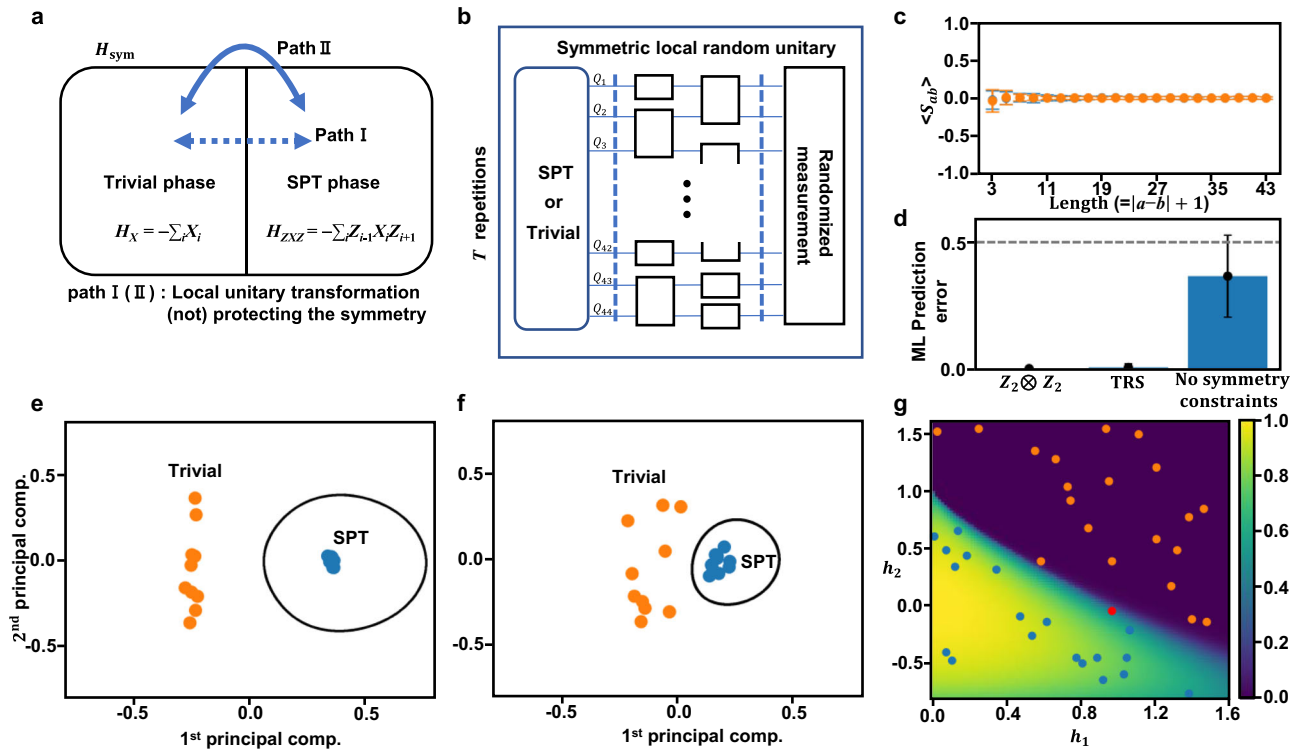
Fig. 3 | Experimental results for distinguishing between symmetry protected topological (SPT) and trivial phases. a Trivial and SPT phases. When considering a set of gapped symmetric Hamiltonians only, the two phases cannot be connected by a constant depth symmetric local unitary (Path I). However, without consideration of the symmetry, there exists a constant depth local unitary circuit (Path II) connecting them. **b** Generation of data. The process involves applying two layers of symmetric local random unitary to a fixed-point state associated with a particular phase, followed by randomized measurement for classical shadows. Random gates used can be found in the Supplementary Note 5. **c** String order parameter (SOP) ($S_{ab} = Z_a X_{a+1} X_{a+3} ... X_{b-3} X_{b-1} Z_b$) of various lengths measured by direct measurements on the quantum computer. The blue and orange dots correspond to the SPT and trivial phases, respectively. The error bars represent the standard deviation of 300 SOPs. **d** Classification errors. The chart shows the classification error of the ML model for cases with $\mathbb{Z}_2 \otimes \mathbb{Z}_2$, TRS, and without symmetry constraints. The error bars represent the standard deviation from 10 instances. **e, f** 2D representation of the test data. Each figure corresponds to $\mathbb{Z}_2 \otimes \mathbb{Z}_2$ (**e**) and time-reversal symmetry (TRS) (**f**). **g** Classification results. The background represents the expectation values of a SOP measured from the ground states of the H_{CI}, and these values are used to determine the ground truth of the quantum phase. For the 40 randomly sampled test data (marked by circles), each phase predicted by the trained ML model is colored in blue for the SPT phase and orange for the trivial phase. The red dot represents a point where the model incorrectly predicted the phase.

toric code), which is well known to have a $\mathbb{Z}_2$ topological order, as a fixed-point state for a topologically ordered phase and a random product state for a trivial phase. The method that was previously used for preparing the topologically ordered state[49] requires hardware-specific qubit connectivity conditions, which would otherwise incur additional swap gates and lower the state preparation fidelity. In addition, the required circuit depth for these methods increases as $\mathcal{O}(d_{code})$ where d_{code} is the code distance of the surface code[49]. We avoided these problems by utilizing the measurement-assisted state preparation[29] method instead. Specifically, after applying a sequence of unitary transformations in Fig. 4a, we performed a projection by measuring some ancilla qubits and applied Pauli Z or an identity gate adaptively to certain data qubits conditioned on the measurement results. Despite the non-deterministic nature of the projection, it is always possible to prepare $|0_L\rangle$ of surface code as follows at $\mathcal{O}(1)$ circuit depths.

$$|0_L\rangle = 1/\sqrt{2^{N_p}} \prod_p (I + B_P)|0\rangle^{\otimes n} \qquad (4)$$

Here, $B_P = \Pi_{i \in p} X_i$ is a X-plaquette operator where each X-plaquette is shaded in blue in Fig. 4a, N_p is the number of X-plaquettes. In addition, we applied the adaptive virtual gate to classical shadows to eliminate possible errors arising from idle data qubits during the measurement of the ancilla qubits. This flexibility is another advantage of using the

classical shadow as data for classical ML. Further details on this method can be found in the Supplementary Note 3.

We prepared a $d_{code} = 5$ surface code and applied tensor products of the random single qubit gates to generate data for the topologically ordered phase. For the trivial phase, we prepared a random product state and applied local random unitary at depth (d_{LU}) from 0 to 5 (Fig. 4b). We obtained 20 data points from each phase using $T = 300$. The ML model was trained by following the same procedure as in Case 2-1. As shown in Fig. 4c, it was possible to distinguish the phases by unsupervised learning when $d_{LU} = 0$. The distribution of the trivial phases spread out as the d_{LU} increased, and the phase boundary faded beyond $d_{LU} \sim (d_{code} - 1)/2 = 2$. To verify that the observed phenomenon originated from a robust topological property, we carried out the same experiments for a randomly sampled state with the same gate complexity for preparing $|0_L\rangle$ but without any topological order. The results indicated that a phase boundary does not exist, even when d_{LU} is small (Fig. 4d).

In comparison to previously studied direct measurements of the topological order parameter, measuring topological string operators[46] within error ε only required $\mathcal{O}(1/\varepsilon^2)$ measurements; however, deviations of the state from the fixed-point state degraded its classification ability[18]. This could be avoided by evaluating the nonlinear properties such as TEE[47,48] but it would require more measurements than were made in our case.
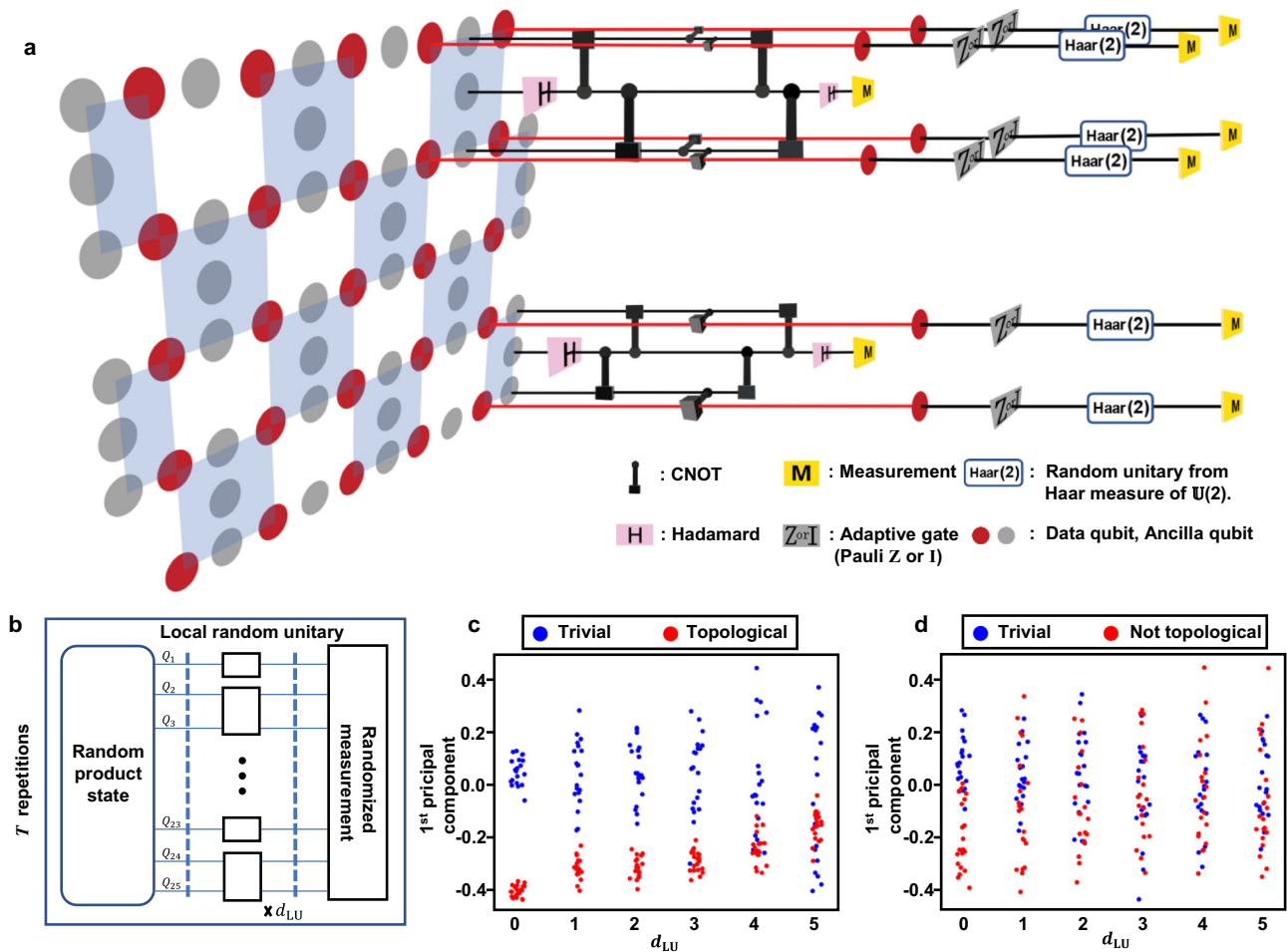
Fig. 4 | Experimental results for distinguishing between topologically ordered and trivial phases. a Schematic diagram for measurement-assisted state preparation method on the heavy-hexagonal lattice. The application of a sequence of unitary transformations and adaptive gates depending on the measurement results of some ancilla qubits was followed by randomized measurements for classical shadows. The blue shaded areas represent the X-plaquettes. **b** Generation of data in the trivial phase. States in this phase are generated by applying local random unitary to the product state at varying circuit depth (d_{LU}). Detailed information about the local random unitary that was used is provided in the Supplementary Note 5. **c** 1D projection of data. Using kernel principal component analysis with the shadow kernel in Eq. (3), each data point was projected onto the first principal axis. **d** 1D projection of data without topological order.

Discussion

Unlocking the black-box of the ML model

A drawback of ML-based analyses is the black-box nature of the ML model, which makes it challenging to understand which aspects of the data are utilized in these analyses. To address this issue, we conducted experiments to extract relations among the data that were used by the ML model. In experiment, we collected 10 data points each from both the topologically ordered and trivial phases consisting of 9 qubits with local random unitary applied to a fixed-point state. Then, we used SVM on the feature space formed by the feature map $\varphi(\rho)$, which transformed the input quantum state into a vector consisting of the subsystem Renyi-2 entanglement entropy (Fig. 5a). ML in this case corresponds to finding an optimal linear combination of subsystem Renyi-2 entanglement entropy. Details of the mapping procedure and a specific form of the ML classifier can be found in the Supplementary Note 5. The evaluation of the trained ML model presented in Fig. 5b was conducted using 100 datasets (each set consisted of 6000 states, generated in the same way as the training data, but through classical simulation). We observed that the phase classifier derived from the ML model distinguishes quantum phases more effectively than the previously used nonlinear order parameter of Renyi-2 TEE[49]. This result is attributed to the intentional introduction of errors in training data, a technique named as data augmentation[50]. Furthermore, a comparison of the prediction error of the model with and without measurement error mitigation (MEM) in the training data enabled us to confirm that appropriate error mitigation techniques can help improve the ML performance.

Our results highlight an interesting aspect, namely the use of classical ML to process quantum experimental data with problem-specific error-reducing procedures for studying quantum many-body physics. Extensions of our work and interesting future directions can be outlined as follows. Rather than conducting the measurements immediately after preparing the quantum state, compressive quantum transformations preserving an important feature of the state before measurements could lower the dimensionality of the quantum state[43], which could reduce the computational time for classical ML. Moreover, investigating the non-equilibrium properties using ML based on data obtained from dynamical simulations of quantum systems[51] would be a promising generalization of our work. Additionally, training the ML model with experimental data obtained by applying various error mitigation techniques or other methodology suitable for classical shadows represents a promising avenue for further research[52–63]. Consequently, we anticipate that future work based on our results would continue to extend the useful applications of quantum devices before fault-tolerant quantum computers.

a

b
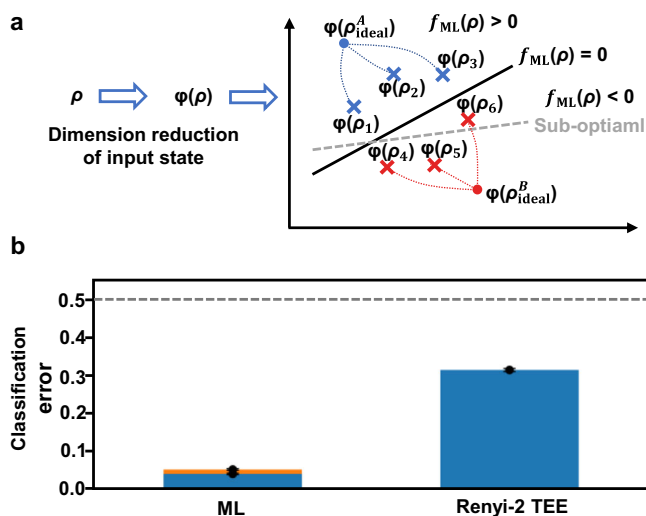
Fig. 5 | Extracting a phase classifier from a trained ML model. a Illustrative diagram of ML procedure. In the first step, given a quantum state ρ, we used feature mapping $\varphi(\rho)$ to reduce the dimensionality of the data. To achieve an error-resilient phase classifier $f_{ML}(\rho)$, training data are generated by applying random unitary transformations to fixed-point states. **b** Classification error. The orange bars represent the results from the ML model trained on raw data without measurement error mitigation (MEM). The blue bars represent the classification error of the ML model trained by the data with MEM and the Renyi-2 Topological entanglement entropy (TEE). The error bar denotes the standard deviation of 100 instances of the test data set.

Methods

Kernel ridge regression
To predict the properties of the ground state from the measured expectation values, we utilized the classical ML algorithm, Kernel Ridge Regression (KRR). KRR aims to estimate the value of a real-valued function $f(x)$ for a new input vector x_{new} using the given data $\{(x_i, f(x_i))\}_i$. Here, $f(x_{new})$ was approximated as $f_{ML}(x_{new}; w) = w^T\varphi(x_{new})$ by employing w, which minimizes the L_2 loss with the regularization $C(w) = \frac{1}{2}\sum_{i=1}^{N_{data}}|w^T\varphi(x_i) - f(x_i)|^2 + \frac{1}{2}\|w\|_2^2$. As the feature vector $\varphi(x)$ typically resides in high-dimensional space, we used the kernel trick instead of directly calculating the high dimensional feature vector $\varphi(x)$.

Kernel principal component analysis
For reducing the dimensionality of data, we applied kernel principal component analysis (PCA) to quantum states represented as classical shadows, using the shadow kernel. PCA was implemented by defining the covariance matrix $C = \frac{1}{N_{data}}\sum_{i=1}^{N_{data}}\varphi(x_i)\varphi(x_i)^T$ and by diagonalizing C, which led to principal axes that could describe the data with the minimum possible number of variables. But, mapping an input vector x into a large dimensional vector through the feature map $\varphi : x \in R^m \to R^{m_\varphi}$, as in our case, would cause the direct diagonalization of C to become computationally challenging. Using the kernel trick as a workaround, by diagonalizing the kernel matrix K, defined by $K_{ij} = k(x_i, x_j) = \varphi(x_i)^T\varphi(x_j)$ for indices i, j ranging from 1 to N_{data}, we can derive the expression for principal axes instead of diagonalizing covariance matrix C. Using kernel PCA, we effectively reduced the dimensionality of the data, which led to the clustering of data points within the same quantum phase.

Error-reducing procedures
We utilized various methods to reduce the errors arising from the data acquisition process in quantum computing. Both Dynamical Decoupling (DD) and Pauli twirling (PT) were employed in all experiments. Additional techniques such as particle number conservation, McWeeny purification, and parity measurement by recompiling the circuit were used to train ML model for the 1D nearest-neighbor random hopping system. In the task of distinguishing quantum phases, we were able to eliminate the swap gate overhead induced by the hardware qubit connectivity through the measurement-assisted state preparation method and reduce the errors from idle data qubits by using the virtual (adaptive) gate on the classical shadow. Furthermore, in the experiment that involved extracting the ML classifier from the trained ML model, we utilized measurement error mitigation, which improved the performance of the trained ML model. Detailed explanations for each method can be found in the Supplementary Note 3.

Classical simulation for ground states
The test data in Fig. 3g were obtained by using matrix product state (MPS) representation and density matrix renormalization group (DMRG) to compute the ground state of $H_{CI} = -J\sum_i Z_i X_{i+1}Z_{i+2} - h_1 \sum_i X_i - h_2\sum_i X_i X_{i+1}$. In our simulation, we utilized a bond dimension of $\mathcal{X} = 100$ for a system of 44 qubits and employed a perfect sampling method to enable the simulation of randomized measurements to obtain the classical shadows of the prepared quantum state represented by MPS.

Hardware characteristics
We used the quantum hardware (*ibm_sherbrooke*) with 127 qubits available through IBM Cloud. The backend is composed of fixed-frequency transmon qubits where each qubit, embedded in a heavy-hexagonal lattice, is directly connected with two or three other qubits. Detailed device characteristics are summarized in the Supplementary Note 1.

Data availability
The data that support the findings of this study are available from the corresponding author upon request.

References
1. Philips, S. G. J. et al. Universal control of a six-qubit quantum processor in silicon. *Nature* **609**, 919–924 (2022).
2. Moses, S. A. et al. A race-track trapped-ion quantum processor. *Phys. Rev. X* **13**, 041052 (2023).
3. Bravyi, S., Dial, O., Gambetta, J. M., Gil, D. & Nazario, Z. The future of quantum computing with superconducting qubits. *J. Appl. Phys.* **132**, 160902 (2022).
4. Abobeih, M. H. et al. Fault-tolerant operation of a logical qubit in a diamond quantum processor. *Nature* **606**, 884–889 (2022).
5. Bluvstein, D. et al. A quantum processor based on coherent transport of entangled atom arrays. *Nature* **604**, 451–456 (2022).
6. Alipanahi, B., Delong, A., Weirauch, M. T. & Frey, B. J. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nat. Biotechnol.* **33**, 831–838 (2015).
7. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
8. Carleo, G. & Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017).
9. Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **136**, B864–B871 (1964).
10. White, S. R. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.* **69**, 2863–2866 (1992).
11. Aharonov, D. & Ben-Or, M. Fault-Tolerant Quantum Computation with Constant Error Rate. *SIAM J. Comput.* **38**, 1207–1282 (2008).
12. Knill, E., Laflamme, R. & Zurek, W. H. Resilient Quantum Computation. *Science* **279**, 342–345 (1998).
13. Google Quantum, A. I. et al. Suppressing quantum errors by scaling a surface code logical qubit. *Nature* **614**, 676–681 (2023).

14. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
15. Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16**, 1050–1057 (2020).
16. Elben, A. et al. The randomized measurement toolbox. *Nat. Rev. Phys.* **5**, 9–24 (2022).
17. Zhang, T. et al. Experimental Quantum State Measurement with Classical Shadows. *Phys. Rev. Lett.* **127**, 200501 (2021).
18. Huang, H.-Y., Kueng, R., Torlai, G., Albert, V. V. & Preskill, J. Provably efficient machine learning for quantum many-body problems. *Science* **377**, eabk3333 (2022).
19. Lewis, L. et al. Improved machine learning algorithm for predicting ground state properties. *Nat. Commun.* **15**, 895 (2024).
20. Che, Y., Gneiting, C., & Nori, F. Exponentially improved efficient machine learning for quantum many-body states with provable guarantees. *Phys. Rev. Res.* **6**, 033035 (2024).
21. Onorati, E., Rouzé, C., França, D. S. & Watson, J. D. Efficient learning of ground & thermal states within phases of matter. Preprint at http://arxiv.org/abs/2301.12946 (2023).
22. Huang, H.-Y. et al. Power of data in quantum machine learning. *Nat. Commun.* **12**, 2631 (2021).
23. Jung, K. et al. Deep learning enhanced individual nuclear-spin detection. *npj Quantum Inf.* **7**, 41 (2021).
24. Miles, C. et al. Machine learning discovery of new phases in programmable quantum simulator snapshots. *Phys. Rev. Res.* **5**, 013026 (2023).
25. Huang, H.-Y. et al. Quantum advantage in learning from experiments. *Science* **376**, 1182–1186 (2022).
26. Herrmann, J. et al. Realizing quantum convolutional neural networks on a superconducting quantum processor to recognize quantum phases. *Nat. Commun.* **13**, 4144 (2022).
27. Smith, A., Jobst, B., Green, A. G. & Pollmann, F. Crossing a topological phase transition with a quantum computer. *Phys. Rev. Res.* **4**, L022020 (2022).
28. Pollmann, F. & Turner, A. M. Detection of symmetry-protected topological phases in one dimension. *Phys. Rev. B* **86**, 125441 (2012).
29. Lu, T.-C., Lessa, L. A., Kim, I. H. & Hsieh, T. H. Measurement as a Shortcut to Long-Range Entangled Quantum Matter. *PRX Quantum* **3**, 040337 (2022).
30. Haah, J., Harrow, A. W., Ji, Z., Wu, X. & Yu, N. Sample-optimal tomography of quantum states. *IEEE Trans. Inform. Theory* 1–1 https://doi.org/10.1109/TIT.2017.2719044 (2017).
31. Aaronson, S. Shadow tomography of quantum states. *SIAM J. Comput.* **49**, STOC18-368–STOC18-394 (2020).
32. Paini, M., Kalev, A., Padilha, D., & Ruck, B. Estimating expectation values using approximate quantum states. *Quantum* **5**, 413 (2021).
33. Bishop, C. M. *Pattern recognition and machine learning* (Springer, 2006).
34. Google AI Quantum and Collaborators et al. Hartree-Fock on a superconducting qubit quantum computer. *Science* **369**, 1084–1089 (2020).
35. Jiang, Z., Sung, K. J., Kechedzhi, K., Smelyanskiy, V. N. & Boixo, S. Quantum Algorithms to Simulate Many-Body Physics of Correlated Fermions. *Phys. Rev. Appl.* **9**, 044036 (2018).
36. Kim, Y. et al. Scalable error mitigation for noisy quantum circuits produces competitive expectation values. *Nat. Phys.* **19**, 752–759 (2023).
37. Sung, K. J., Rančić, M. J., Lanes, O. T. & Bronn, N. T. Simulating Majorana zero modes on a noisy quantum processor. *Quantum Sci. Technol.* **8**, 025010 (2023).
38. Zeng, B. & Wen, X.-G. Gapped quantum liquids and topological order, stochastic local transformations and emergence of unitarity. *Phys. Rev. B* **91**, 125121 (2015).
39. Liu, Y.-J., Smith, A., Knap, M. & Pollmann, F. Model-Independent Learning of Quantum Phases of Matter with Quantum Convolutional Neural Networks. *Phys. Rev. Lett.* **130**, 220603 (2023).
40. Chen, X., Zeng, B., Wen, X.-G., & Zhou, D.-L. (eds) *Quantum Information Meets Quantum Matter: From Quantum Entanglement to Topological Phases of Many-Body Systems* (Springer, 2019).
41. Elben, A. et al. Many-body topological invariants from randomized measurements. *Sci. Adv.* **6**, eaaz3666 (2020).
42. Caro, M. C. et al. Generalization in quantum machine learning from few training data. *Nat. Commun.* **13**, 4919 (2022).
43. Cong, I., Choi, S. & Lukin, M. D. Quantum convolutional neural networks. *Nat. Phys.* **15**, 1273–1278 (2019).
44. Ferris, A. J. & Vidal, G. Perfect sampling with unitary tensor networks. *Phys. Rev. B* **85**, 165146 (2012).
45. Schollwöck, U. The density-matrix renormalization group in the age of matrix product states. *Ann. Phys.* **326**, 96–192 (2011).
46. Semeghini, G. et al. Probing topological spin liquids on a programmable quantum simulator. *Science* **374**, 1242–1247 (2021).
47. Levin, M. & Wen, X.-G. Detecting Topological Order in a Ground State Wave Function. *Phys. Rev. Lett.* **96**, 110405 (2006).
48. Kitaev, A. & Preskill, J. Topological Entanglement Entropy. *Phys. Rev. Lett.* **96**, 110404 (2006).
49. Satzinger, K. J. et al. Realizing topologically ordered states on a quantum processor. *Science* **374**, 1237–1241 (2021).
50. Perez, L. & Wang, J. The Effectiveness of Data Augmentation in Image Classification using Deep Learning. Preprint at http://arxiv.org/abs/1712.04621 (2017).
51. Kim, Y. et al. Evidence for the utility of quantum computing before fault tolerance. *Nature* **618**, 500–505 (2023).
52. Chen, S., Yu, W., Zeng, P. & Flammia, S. T. Robust Shadow Estimation. *PRX Quantum* **2**, 030348 (2021).
53. Koh, D. E. & Grewal, S. Classical Shadows With Noise. *Quantum* **6**, 776 (2022).
54. Jnane, H., Steinberg, J., Cai, Z., Nguyen, H. C., & Koczor, B. Quantum error mitigated classical shadows. *PRX Quantum* **5**, 010324 (2024).
55. Brieger, R., Heinrich, M., Roth, I. & Kliesch, M. Stability of classical shadows under gate-dependent noise. Preprint at http://arxiv.org/abs/2310.19947 (2023).
56. Zhao, A. & Miyake, A. Group-theoretic error mitigation enabled by classical shadows and symmetries. *npj Quantum Inf.* **10**, 57 (2024).
57. Wu, B. & Koh, D. E. Error-mitigated fermionic classical shadows on noisy quantum devices. *npj Quantum Inf.* **10**, 39 (2024).
58. Huang, H.-Y., Kueng, R. & Preskill, J. Efficient estimation of Pauli observables by derandomization. *Phys. Rev. Lett.* **127**, 030503 (2021).
59. Hu, H.-Y. & You, Y.-Z. Hamiltonian-driven shadow tomography of quantum states. *Phys. Rev. Res.* **4**, 013054 (2022).
60. Bu, K., Koh, D. E., Garcia, R. J. & Jaffe, A. Classical shadows with Pauli-invariant unitary ensembles. *npj Quantum Inf.* **10**, 6 (2024).
61. Akhtar, A. A., Hu, H.-Y. & You, Y.-Z. Scalable and Flexible Classical Shadow Tomography with Tensor Networks. *Quantum* **7**, 1026 (2023).
62. Zhou, Y. & Liu, Q. Performance analysis of multi-shot shadow estimation. *Quantum* **7**, 1044 (2023).
63. Zhou, Y. & Liu, Z. A hybrid framework for estimating nonlinear functions of quantum states. *npj Quantum Inf.* **10**, 62 (2024).

Acknowledgements

Author contributions

G.C. conceived the project, performed cloud-based experiments, conducted classical error simulations, and analyzed the data with D.K. G.C. and D.K. prepared the manuscript. D.K. supervised the project.

Competing interests

The author declares no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-51932-3.

Correspondence and requests for materials should be addressed to Gyungmin Cho or Dohun Kim.

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