# Quantum entanglement recognition

Jun Yong Khoo [1,2] and Markus Heyl [1]

[1]*Max-Planck Institute for the Physics of Complex Systems, D-01187 Dresden, Germany*
[2]*Institute of High Performance Computing, Agency for Science, Technology, and Research, Singapore 138632*

Entanglement constitutes a key characteristic feature of quantum matter. Its detection, however, still faces major challenges. In this paper, we formulate a framework for probing entanglement based on machine learning techniques. The central element is a protocol for the generation of statistical images from quantum many-body states, with which we perform image classification by means of convolutional neural networks. We show that the resulting quantum entanglement recognition task is accurate and can be assigned a well-controlled error across a wide range of quantum states. We discuss the potential use of our scheme to quantify quantum entanglement in experiments. Our developed scheme provides a generally applicable strategy for quantum entanglement recognition in both equilibrium and nonequilibrium quantum matter.

## I. INTRODUCTION

Entanglement has turned into a central concept across various branches in physics ranging from quantum technological applications [1] to the characterization of quantum matter [2,3]. It has remained, however, a key challenge to quantify the entanglement content of a given quantum state especially under realistic experimental conditions beyond the ground and pure state paradigm. This challenge is rooted in the fundamental property that entanglement measures are nonlinear functions of the density matrix, but quantum measurements only yield direct information linear in it according to the axiomatic foundations of quantum mechanics. For quantum systems involving a limited number of degrees of freedom, entanglement can still be quantified also experimentally upon reconstructing the full density matrix via tomography [4–8], by means of measurements on identical copies of quantum states [9–13], or through the statistics of randomized measurements [14–16].

By applying machine learning techniques we show in this paper that entanglement measures can be accurately extracted merely from limited information obtainable directly from quantum measurements, thereby reducing significantly the necessary measurement resources. The key element of the proposed scheme is a protocol to generate a two-dimensional statistical image from a given quantum many-body state. Utilizing conventional image recognition based on convolutional neural networks, we then perform a *quantitative* classification of the entanglement entropy and logarithmic negativity, rang-

ing over a broad class of quantum many-body wave functions from ground and excited states to quantum nonequilibrium evolution for pure and mixed states. We apply these techniques to various one-dimensional quantum spin-1/2 chains and find that this classification is remarkably accurate and can be assigned a well-defined error. Importantly, we find that the artificial neural networks (ANNs) for quantum entanglement recognition are robust to weak perturbations. Unlike previous machine-learning-based schemes that exploited ANNs to classify states belonging to different entanglement classes [17,18], our focus on quantitative entanglement prediction broadens significantly the applicability and provides a direct means of performance characterization through the prediction error. In particular, the obtained ANNs can, at a controlled error, even predict nonequilibrium dynamics, generated, e.g., by a Lindblad master equation. We further discuss how this quantum entanglement recognition might be utilized experimentally as an entanglement estimator upon training with simulated data.

## II. IMAGE GENERATION FROM QUANTUM STATES

The key element of our quantum entanglement recognition scheme is the statistical image generating protocol, which we now introduce. Let $\rho_0$ denote the density matrix of a quantum (many-body) system. In the following, we study models of $L$ spin-1/2 degrees of freedom for simplicity, although the protocol can be extended straightforwardly to any lattice model with finite local Hilbert spaces. We perform measurements on the quantum state with the string $O = \sigma_1^z \otimes \cdots \otimes \sigma_L^z$ of Pauli matrices $\sigma_l^z$ yielding as outcomes spin configurations $\mathbf{s} = (s_1, \ldots, s_L)$ with $s_l = \pm 1$. Such measurements can be performed on quantum computing platforms such as trapped ions [19] and superconducting qubits [20–22] or ultracold atomic systems via quantum gas microscopes [23]. In general, a single measurement basis is not sufficient for entanglement detection. We therefore generate a more
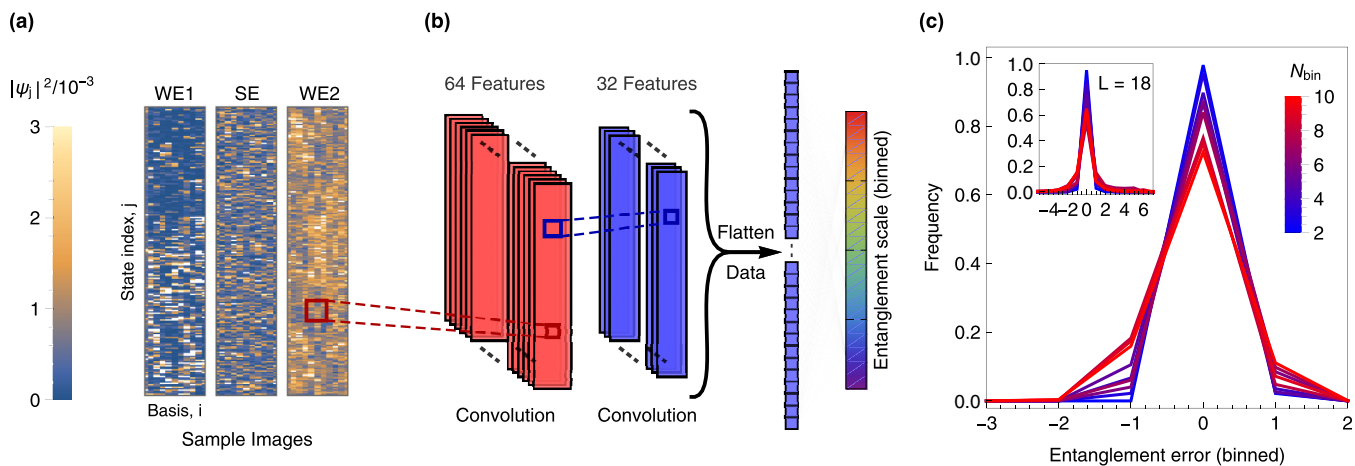
FIG. 1. Schematic illustration of the quantum entanglement recognition scheme. (a) Representative statistical images of weakly entangled pure states (WE1), strongly entangled pure states (SE), and weakly entangled mixed states (WE2) of the transverse-field ferromagnetic Ising model. (b) Adapting an image recognition neural network for entanglement quantification. The network reads statistical images such as those in (a). After processing the image information the network classifies the entanglement by assigning the images to different labels corresponding to binned intervals of the considered entanglement measure. Shown in the figure is the case of $N_{\mathrm{bin}} = 4$ total number of bins. (c) Distribution of the error in entanglement quantification for weakly entangled ground states (WE1) of the ferromagnetic transverse-field Ising model with field strengths $h > 1$ over a range of $2 \leqslant N_{\mathrm{bin}} \leqslant 10$ for a spin chain with $L = 10$ sites. Error distances $\delta$ are measured in units of bins. A similar performance is obtained for larger system sizes, as is shown in the inset for the case of $L = 18$.

detailed picture of $\rho_0$ by applying a set of $\mathcal{W}$ fixed but random local unitary transformations $U_{i>1} = u_{i,1} \otimes \cdots \otimes u_{i,L}$ with $i = 2, \ldots, \mathcal{W}$, where each local unitary $u_{i>1,l}$ on site $l$ is drawn independently from the circular unitary ensemble (CUE) [24]. The full information about $\rho_0$ can be obtained by measuring along $\mathcal{W} = 2^L$ orthogonal directions; this would lead to full state tomography. Here, we explore whether a limited number of measurement axes is sufficient for entanglement quantification. For that purpose, we start with a simple experimentally accessible way of generating independent measurement directions via local unitary rotations. Measuring the rotated $\rho \mapsto \rho_i = U_i \rho_0 U_i^\dagger$ as before, we obtain the probabilities

$$p_{ij} = \langle j | \rho_i | j \rangle = \mathrm{Tr}[U_i \rho_0 U_i^\dagger |j\rangle\langle j|], \qquad (1)$$

with $j = 1, \ldots, \mathcal{D}$ labeling the $\mathcal{D} = 2^L$ spin configurations. To begin, we consider the case with $\mathcal{W} = 11$. In this way we obtain a two-dimensional representation of $\rho_0$ in terms of the probabilities $p_{ij}$, shown for three exemplary states of different entanglement classes in Fig. 1(a).

While we have chosen the simplest form of local unitary transformations $U_i$ for the image generating protocol, note that it is not unique and can just as well comprise many-site unitary transformations respecting the symmetry of the entanglement measure (see Appendix F for details). For any protocol, a natural choice, though not absolutely essential, is to have $U_1$ be the $\mathcal{D} \times \mathcal{D}$ identity operator, i.e., $\rho_1 = \rho_0$, such that $p_{1j} = \langle j | \rho_0 | j \rangle$. We found that while this choice does not affect the performance of the network trained with the simplest protocol proposed here, it turns out to be crucial for the network to achieve a comparable performance when trained with some of these alternative protocols.

## III. SUPERVISED LEARNING OF ENTANGLEMENT

In the following we now outline how the introduced statistical image generation can be used to perform a quantitative entanglement classification task by means of a supervised learning procedure. For pure states a natural entanglement measure is the half-chain entanglement entropy

$$S(\rho) = -\mathrm{Tr}_<[\rho_< \ln \rho_<], \qquad (2)$$

with $\rho_< = \mathrm{Tr}_> \rho$ being the reduced density for the first half of the chain, obtained by tracing out the degrees of freedom of the remainder $>$ from the full density matrix $\rho$. For mixed states we use instead the logarithmic negativity [25],

$$E_{\mathcal{N}}(\rho) = \log_2 ||\rho^{T_<}||, \qquad (3)$$

where $||\rho^{T_<}||$ denotes the trace norm of $\rho^{T_<}$ and $\rho^{T_<}$ is the partial transpose of $\rho$ on the degrees of freedom of the first half of the chain.

For the desired quantification task we bin the range of values for the respective entanglement measure into $N_{\mathrm{bin}}$ equally spaced intervals. We fix an interval $I_S = [0, S_{\mathrm{max}}]$ for the entanglement entropy $S$, say, with a suitably chosen $S_{\mathrm{max}}$ and decompose $I_S = \bigcup_{n=1}^{N_{\mathrm{bin}}} I_{S_n}$ into $I_{S_n} = [(n-1)\Delta S, n\Delta S)$ with $n = 1, \ldots, N_{\mathrm{bin}} = S_{\mathrm{max}}/\Delta S$. Each of these bins, labeled by $n$, corresponds to the category with which we aim to associate the images. This binning process is necessary when there are no training images in some range of entanglement values, and consequently, the largest value of $N_{\mathrm{bin}}$ permissible is such that there are no empty intervals $I_{S_n}$. With this we can now perform an entanglement classification problem as in conventional image recognition. We adapt a convolutional neural network to process the two-dimensional image of the quantum state as shown in Fig. 1(b). Two layers of feature maps, comprising 64 and 32 features, respectively, are extracted from the images.

These feature data are then flattened and fitted across the images from the training library to their respective $N_{bin}$ labels. See Appendix A for more details on the network. We quantify the classification accuracy on test images by the signed binning error distance $\delta = n - n_{ANN}$ measuring the difference between the bin label $n$ from the exact entanglement content and $n_{ANN}$ predicted by the ANN. A typical distribution of $\delta$ for various $2 \leqslant N_{bin} \leqslant 10$ is shown in Fig. 1(c) for a specific benchmark problem. Importantly, the entanglement classification exhibits a well-defined error, whose distribution is sharply peaked around $\delta = 0$, implying no error, with some further appreciable weight only for $\delta = \pm 1$.

## IV. RESULTS

It is the key goal of the following analysis to explore the performance of the quantum entanglement recognition scheme for a large variety of quantum states.

### A. Models

As benchmark systems we choose a set of paradigmatic one-dimensional quantum many-body models. This includes the transverse-field Ising chain with either ferromagnetic (TFI+) or antiferromagnetic (TFI−) couplings as well as the XX model:

$$H_{TFI\pm} = \mp \sum_{\langle i,j \rangle} \sigma_i^x \sigma_j^x - h \sum_{i=1}^{L} \sigma_i^z, \qquad (4)$$

$$H_{XX} = -\sum_{\langle i,j \rangle} \left( \sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y \right) - h \sum_{i=1}^{L} (-1)^i \sigma_i^z. \qquad (5)$$

Here, $\sigma_i^{x,y,z}$ denote the Pauli matrices at site $i = 1, \ldots, L$, and $h$ denotes a magnetic field strength. While throughout this paper we show numerical data for $L = 10$, we emphasize that the entanglement recognition does not depend significantly on $L$. This scalability with system size is exemplified in Fig. 1(c), where we also include data for $L = 18$. Although we focus on a particular set of models, we find that our results do not depend crucially on the model details as discussed below, suggesting that our observations are generic and therefore applicable in a broad context.

### B. Ground and excited states

We start by studying weakly entangled ground states of the considered spin chains. First, we explore the performance by testing on the same class of states, e.g., with ground states of TFI+ after training the ANN with ground states of the same model (see Appendix B for details). The network performs remarkably well in this case, as can be seen in Fig. 1(c), showing a typical distribution of the binning error $\delta$. Almost independent of $N_{bin}$, one can recognize a strongly peaked distribution with an appreciable weight beyond $\delta = 0$ only at $\delta = \pm 1$. As the mean error is practically vanishing, the performance is effectively captured by the standard deviation, which can therefore be used as a well-defined error quantifier. The fact that the network typically fails at most by assigning a state to the nearest-neighboring bin suggests that the dominant error originates from those instances where the actual
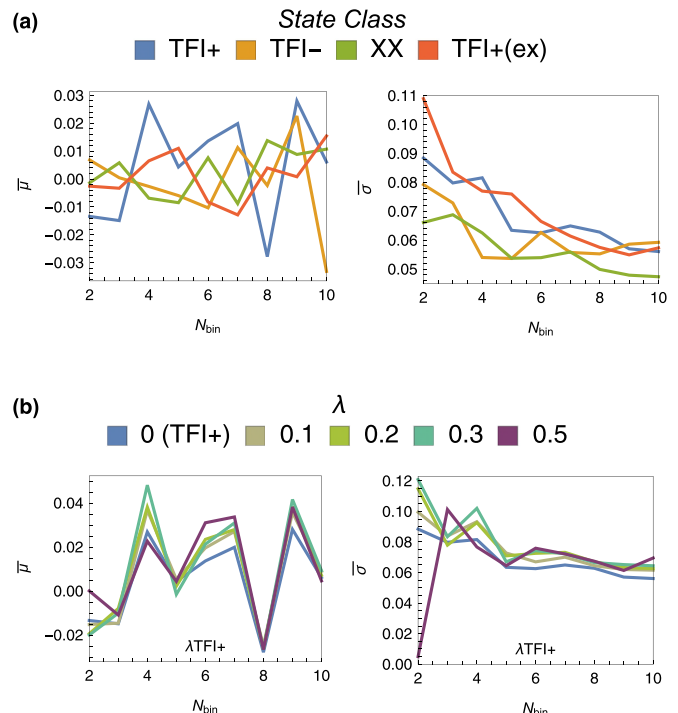


FIG. 2. Characterization of network performance on weakly entangled states through its binning error statistics. (a) The rescaled mean $\bar{\mu}$ (left) fluctuates negligibly about zero while the rescaled standard deviation $\bar{\sigma}$ (right) decreases with increasing total number of bins $N_{bin}$ used for the training. The binning error is rescaled by $N_{bin}$ to reflect the reduction in the maximum entanglement entropy error as $N_{bin}$ increases. The plots are shown for the three classes of weakly entangled ground states (i.e., obtained for fields $h > 1$); considered are those from the transverse-field ferromagnetic (TFI+) and antiferromagnetic (TFI−) Ising models as well as the XX model. Shown also is the corresponding plot for the full spectrum of TFI+ excited [TFI+(ex)] states (red), indicating the network's applicability beyond area-law-type entanglement. (b) Analogous plots showing the robustness of the TFI+ trained network to perturbations when tested on $\lambda$TFI+ ground states for various strengths of perturbation $\lambda$.

entanglement entropy resides close to the border between two bins. A particularly important consequence of Fig. 1(c) is that the performance of the network improves upon enlarging $N_{bin}$. This can be directly quantified by the statistics of the error $\delta S = S_n - S_n^{ANN} = \delta \times \Delta S$, where $S_n$ denotes the binned value of the computed entanglement entropy and $S_n^{ANN}$ denotes the one predicted by the ANN. For comparing between the performances for different values of $\Delta S$, we considered instead the mean $\bar{\mu}$ and standard deviation $\bar{\sigma}$ of the rescaled error, $\bar{\delta} = \delta S / S_{max} = \delta / N_{bin}$, the results of which are shown in Fig. 2(a) for all the different models considered. In each case, $|\bar{\mu}| < 0.03$ is negligible while a decreasing $\bar{\sigma}$ trend indicates a clear improvement in network prediction accuracy for larger $N_{bin}$.

As a next step, we show that our scheme also applies to excited states with volume-law entanglement, where we again find strongly peaked distributions for $\delta$ similar to that in Fig. 1(c). We have included for one representative case, $H_{TFI+}$, the corresponding $\bar{\mu}$ and $\bar{\sigma}$ in Fig. 2(a). We have also checked

that the network performs equally well when making the Ising chain nonintegrable and across different entanglement measures (see Appendix D for details). Specifically, we added an integrability-breaking perturbation,

$$H_{\lambda\text{TFI+}} = H_{\text{TFI+}} + \lambda \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z, \qquad (6)$$

where $\lambda$ characterizes the strength of the perturbation. Remarkably, as shown in Fig. 2(b), the ANN trained with ground states of $H_{\text{TFI+}}$ is robust to such a perturbation, whose performance when tested on ground states of $H_{\lambda\text{TFI+}}$ does not suffer significantly for weak perturbations ($\lambda = 0.1$), or even for strong perturbations ($\lambda = 0.5$) when $N_{\text{bin}}$ is large.

### C. States obtained from unitary dynamics

As a next step we aim to explore quantum entanglement recognition in nonequilibrium dynamics. This is of particular importance for many quantum simulator platforms, where it is much more natural to realize time evolution than, for instance, ground state preparation. Here, we will be exclusively interested in whether the ANN can be trained to quantify the entanglement dynamics associated with time evolution under the same Hamiltonian $H$. This effectively probes the ability of the network to differentiate between the entanglements present in different superpositions of the eigenstates of $H$. For the case of unitary dynamics, the training library consists of images of states evenly sampled across the statistical image time series of 1000 different initial randomly polarized states evolving in time under $H$. In this case, there are essentially no gaps in the entanglement distribution of the training images, so that unlike for ground and excited states, binning into intervals is not necessary. Consequently, we are no longer constrained as before to having the network perform a "discrete classification" task. Instead, the network can be modified to perform a "continuous classification" task by appending to the network in Fig. 1(b) a single output neuron that directly predicts the value of the entanglement measure $S$ or $E_{\mathcal{N}}$. See Appendix A for details on network structure and Appendices B and C for details on the training image library.

As a benchmark, we consider the evolution of randomly polarized states under $H_{\text{TFI+}}$. The resulting prediction for the entanglement entropy $S^{\text{ANN}}(t)$, based on its statistical image time series, in comparison to the exact real-time evolution $S(t)$, is shown in Fig. 3(a) for a specific instance (thin green solid curve and thin black dashed curve respectively) and averaged over 100 different initial randomly polarized states (thick green solid curve and thick black dashed curve respectively) with a shaded region indicating the associated standard deviation of the error in the network prediction. As one can see, the network is able to almost precisely predict the entanglement dynamics, achieving essentially vanishing error for $S(t)$ under unitary evolution across the timescales probed.

### D. States obtained from dissipative dynamics

In an experimentally realistic context the major challenge is to quantify entanglement for mixed states. For this purpose we study exemplarily the logarithmic negativity $E_{\mathcal{N}}$ for nonunitary time evolution as described by a Lindblad master
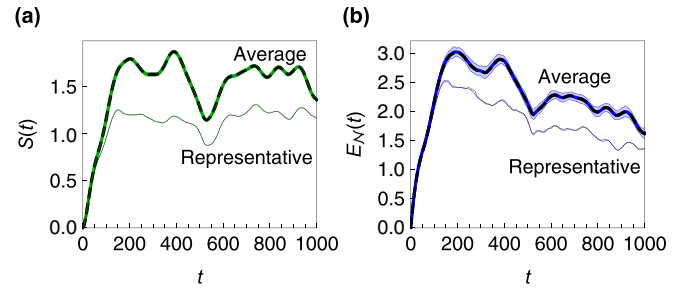


FIG. 3. Excellent quantification of entanglement dynamics by a network trained under the same time evolution. (a) Unitary dynamics under $H_{\text{TFI+}}$ with $h = 1$ of the entanglement entropy $S(t)$ and (b) dissipative dynamics with small dissipation parameter $\gamma = 0.01$ of the logarithmic negativity $E_{\mathcal{N}}(t)$, for a representative initial random polarized state (Representative) and averaged over 100 different initial randomly polarized states (Average). In contrast to the actual values plotted with black dashed curves, the network predictions are plotted in (a) green and (b) blue, with a corresponding shaded region indicating the associated standard deviation of the prediction error over the initial random states.

equation of the form $\partial_t \rho = -i[H_{\text{TFI+}}, \rho] + \gamma \sum_{i=1}^{L} (\sigma_i^x \rho \sigma_i^x - \rho)$ with $\gamma$ characterizing the dissipation strength. Analogous to the case of unitary dynamics, the training images are obtained from a statistical image time series of 1000 different initial randomly polarized states evolving in time under the above Lindblad master equation. We find that the ANN is capable of accurately tracking the evolution of $E_{\mathcal{N}}$ for the studied dissipative dynamics. Analogous to the purely unitary case, we compare in Fig. 3(b) the network-predicted $E_{\mathcal{N}}^{\text{ANN}}(t)$ (blue curves) with the exact result $E_{\mathcal{N}}(t)$ (black curves) for the case of weak dissipative dynamics with $\gamma = 0.01$. For early times $t < 100 \sim \gamma^{-1}$, the evolution is approximately unitary, and $E_{\mathcal{N}}(t)$ increases linearly with time. At later times $t > 100$, however, $E_{\mathcal{N}}(t)$ gradually decays as dissipation kicks in leading to a reduction in entanglement. The network performs very well, although slightly worse than the unitary case as can be seen by its slightly broader range of fluctuations in prediction error (shaded region) in Fig. 3(b) compared with that in Fig. 3(a). We characterize the average performance of the network by the time-dependent mean $\mu(t)$ [Fig. 4(a)] and standard deviation $\sigma(t)$ [Fig. 4(b)] of the error in entanglement prediction by the network, $\delta S^{\text{ANN}}(t)$ for the unitary case and $\delta E_{\mathcal{N}}^{\text{ANN}}(t)$ for the nonunitary case, taken over 100 different initial randomly polarized states.

### E. Generalizing to time evolution under a perturbed Hamiltonian

An additional challenge in experiments is the presence of perturbations to a model Hamiltonian that is being emulated. For applications in experiments, therefore, it is desirable for an ANN that is trained for a specific Hamiltonian $H$ to still be relatively accurate in quantifying the entanglement of states evolved by a weakly perturbed Hamiltonian $\tilde{H}$. To this end, for our case with $H = H_{\text{TFI+}}$, we consider again the perturbed Hamiltonian $\tilde{H} = H_{\lambda\text{TFI+}}$ [Eq. (6)]. The effects of perturbation on the network's performance are shown in Fig. 5. We find that when the perturbation is weak, $\lambda = 0.01$, the ANN trained by
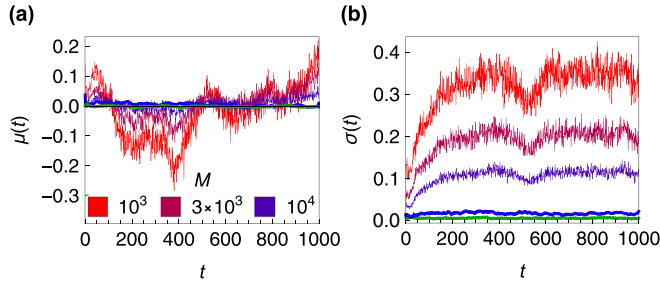
FIG. 4. Effects of image noise on the network performance on dynamical entanglement quantification. (a) The mean $\mu(t)$ and (b) standard deviation $\sigma(t)$ of the prediction error in the respective entanglement measures taken over the 100 initial states used to produce the plots in Figs. 3(a) and 3(b). For the dissipative case, color variation from blue (noiseless) to red indicates the level of noise introduced to the statistical image time series via sampling the wave function a finite $M$ number of times to generate each row of the image.

images of states evolving under $H_{\mathrm{TFI+}}$ with weak dissipation $\gamma = 0.01$ is still able to accurately quantify the entanglement dynamics on average but does so with a larger error at times $t > 100 \sim \gamma^{-1}$, when dissipation begins to dominate. When the perturbation is strong, $\lambda = 0.2$, the error begins to grow from $t > 0$, while it starts to underpredict the entanglement on average after $t \gtrsim 2\gamma^{-1}$. In the worst-case scenario, when in addition the actual dissipation is twice as large as what was used for training, the ANN begins to overpredict on average after $t \gtrsim 4\gamma^{-1}$ by an amount that increases with time. Interestingly, a stronger dissipation does not further increase the standard deviation of the prediction error. The above suggests that dissipation modifies the statistical images in a systematic way which on the one hand can be recognized by the ANN but on the other hand lacks a simple extraction of the dissipation strength. Consequently, the ANN overpredicts (underpredicts) on average the late time entanglement when the actual dissipation is weaker (stronger) than what was used for training. The presence of a strong perturbation to the Hamiltonian, however, does strongly deform the statistical images and cripples the performance of the ANN.

## V. SUMMARY AND DISCUSSION

In this paper we have introduced quantum entanglement recognition based on machine learning techniques. Our protocol provides a controlled and unbiased way to extract quantum state information by essentially applying projective measurements in $\mathcal{W}$ predetermined but randomly selected bases. By organizing this information into a statistical image, the entanglement quantification task is mapped into the conventional image recognition task, for which convolutional neural networks have been optimized to achieve excellent performance. In the large $\mathcal{W} \sim 2^L$ limit, these statistical images essentially capture all the information present in the density matrix, such that it would not be surprising if a trained convolutional neural network is able to successfully reconstruct the entanglement associated with a statistical image.

The central result of this work is to show that indeed the above expectation is correct and, more remarkably, only a
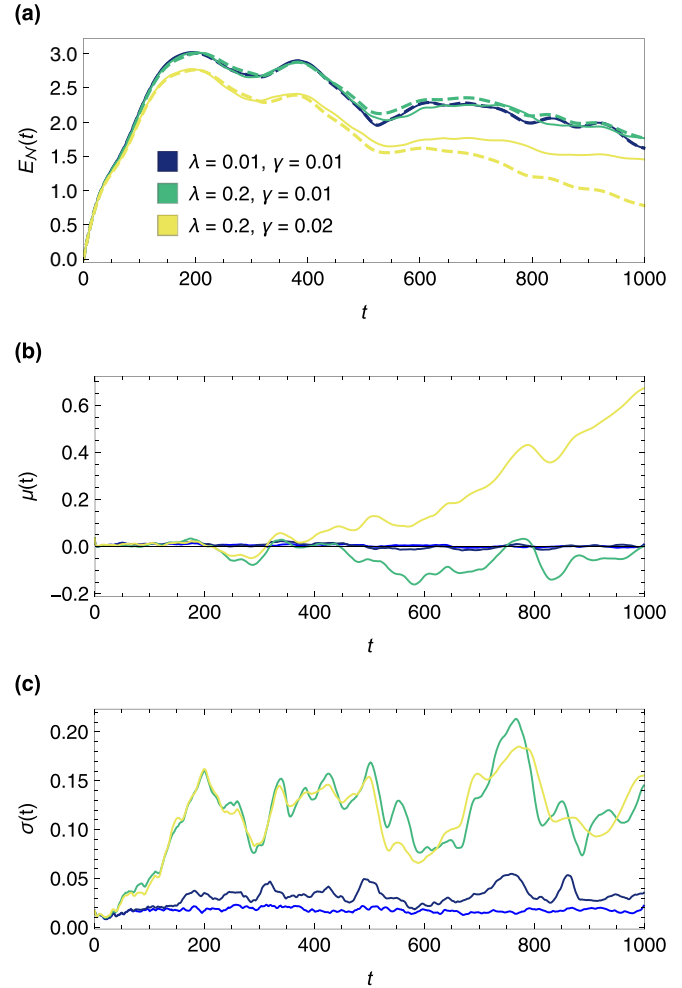


FIG. 5. Effects of perturbations on the network performance of dynamical entanglement quantification. Testing the network that was trained on statistical images of states evolving under $H_{\mathrm{TFI+}}$ with dissipation parameter $\gamma = 0.01$ on those evolving under a perturbed version $H_{\lambda\mathrm{TFI+}} = H_{\mathrm{TFI+}} + \lambda \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z$ with weak (dark blue) and strong (green) perturbations $\lambda$ for the same dissipation, as well as with stronger dissipation for the strongly perturbed case (yellow). Plots of the (a) logarithmic negativity $E_{\mathcal{N}}(t)$ (dashed curves) and those predicted by the network $E_{\mathcal{N}}^{\mathrm{ANN}}(t)$ (solid curves) averaged over 100 initial states and the corresponding (b) mean $\mu(t)$ and (c) standard deviation $\sigma(t)$ of the prediction error. (b) and (c) For comparison, we show the blue reference plots from Figs. 4(a), and 4(b) respectively, i.e., for the case of $\lambda = 0$ and $\gamma = 0.01$.

small set of measurement bases $\mathcal{W} \ll 2^L$ is required. This is the key feature going beyond previous works on applying machine learning techniques to characterize quantum matter [26–30], which operate either in a single measurement basis [26,27,29,30] or by considering low-order correlation functions [28]. By applying to various different classes of quantum many-body states we show that the resulting quantum entanglement recognition can be assigned a well-defined error making the scheme accurate and reliable. While the networks primarily learn model-specific features of the entanglement, which was demonstrated by the inability of a network trained for one class of ground states to always generalize to another class of ground states (see Fig. 7 in Appendix E), we show that

the networks in our scheme are at least capable of generalizing to weak perturbations, e.g., in the context of nonequilibrium dynamics, implying that the networks can learn universal features of quantum entanglement, and from a theoretical standpoint, that such features are already well encoded within the state information obtained from $\mathcal{W} \ll 2^L$ measurement bases. A generalizable network is particularly important for its potential use in experiments, where the microscopic details of the dynamics might not be known in full detail.

In the experimental context, we showed that indeed the network is able to perform accurately in the presence of weak perturbations. In addition, the probabilities $p_{ij}$ from Eq. (1) can only be estimated from a finite number of measurements $M$ of spin configurations, introducing noise into the statistical images. In Figs. 4(a) and 4(b), we have included results for such noisy images, where one can see that even with $M = 3000$ per rotation the dynamics can be reproduced well, with a mean $|\mu(t)| \lesssim 0.1$ and standard deviation $\sigma(t) \lesssim 0.2$ of the prediction error $\delta E_{\mathcal{N}}^{\text{ANN}}(t)$, i.e., within 5–10% of the actual values of $E_{\mathcal{N}}(t)$. For the considered $\mathcal{W} = 5$ rotations this implies a total number of 15 000 measurements, which is much less compared with a recent experiment on probing Rényi entropies [16]. Let us emphasize, however, that we have not attempted to optimize the ANN for the entanglement recognition task at finite $M$, so that further improvements on this front will likely appear in the future. While the perfectly optimized ANN is not expected to overcome the intrinsic exponential-in-system-size scaling barrier associated with quantum measurements, our scheme provides the means to push the experimental limits today substantially closer towards the intrinsic barrier. This opens the door to the development of machine learning tools that directly enable experimental studies on quantum entanglement in systems beyond the few-body context. This is all the more important as many phases of strongly correlated quantum matter are characterized by their entanglement content such as in, e.g., quantum spin liquids as they have been recently experimentally realized on noisy intermediate-scale quantum devices [31] and programmable quantum simulators [32].

### ACKNOWLEDGMENTS

### APPENDIX A: CONVOLUTIONAL NEURAL NETWORKS FOR ENTANGLEMENT RECOGNITION

Based on the statistical images generated out of quantum many-body states, we perform a conventional image recognition task using a convolutional neural network. In the case of ground and excited states for which images are labeled by their half-chain entanglement entropy binned into $N_{\text{bin}}$ intervals, we used a network consisting of two consecutive convolution layers followed by a hidden dense layer with $N_{\text{bin}}$ nodes established by the different image labels. The first convolution layer scans a statistical image and constructs 64 different feature maps, which are in turn scanned by the second layer to construct 32 new feature maps. This feature information is then flattened and fitted across the images from the training library to their respective $N_{\text{bin}}$ labels of the dense layer via the "categorical cross-entropy" loss function.

In the case of states obtained from dynamics for which images are labeled directly by their (continuous) half-chain entanglement entropy or logarithmic negativity, an additional dense layer with a single node is appended to the above network. The data are then fitted to this last node via the "mean-square error" loss function. To increase the quantification precision, we increase the number of nodes in the preceding dense layer to $N_{\text{bin}} = 50$.

### APPENDIX B: TRAINING LIBRARIES

For unbiased training of the ANN, the training library is constructed by generating an approximately equal number of widely varied reference images in each of the $N_{\text{bin}}$ categories, i.e., images from states with entanglement spread across $I_{S_n}$ for each bin $n$.

#### 1. Ground states

We solve for the ground states for different magnetic field strengths $h > 1$ for each of the model classes to obtain the corresponding libraries of labeled images with a uniformly distributed entanglement entropy $S \in I_S = [0, \ln 2]$. To ensure that there are no directional biases in the images concerning a specific orientation on the Bloch sphere, a uniform random rotation $U = \otimes_{i=1}^{L} u$ is applied to each state before generating its corresponding statistical image. We separately train the ANN with the library of a particular model class and then test the entanglement classification with new states it has not seen before from either the same or a different model class. Each training library contains 50 000 images, while an additional 10 000 images are generated for each test library.

#### 2. Excited states

We focus on the TFI+ model (as well as a nonintegrable modification $\lambda$TFI+) for the study of excited states. We solve for the eigenstates and their corresponding entanglement entropies (also logarithmic negativities) for a range of near-critical field strengths $h \in (1, 1.07)$. For each $h$ value, the eigenstates are grouped into $N_{\text{bin}}$ bins ranging between their maximum and minimum entanglement entropies. One state is randomly selected from each bin for image generation, while the rest of the eigenstates are discarded.

#### 3. Unitary and dissipative dynamics

We time-evolve 1000 different initial randomly polarized states for 1000 time steps under $H_{\text{TFI+}}$ for the unitary case, and under the Lindblad master equation of the form $\partial_t \rho = -i[H_{\text{TFI+}}, \rho] + \gamma \sum_{i=1}^{L} (\sigma_i^x \rho \sigma_i^x - \rho)$ with $\gamma = 0.01$ for the dissipative case. For each initial state, its time-evolved states at 250 random but evenly spaced time steps are selected for image generation, giving a total of 250 000 images on which the ANN is trained.
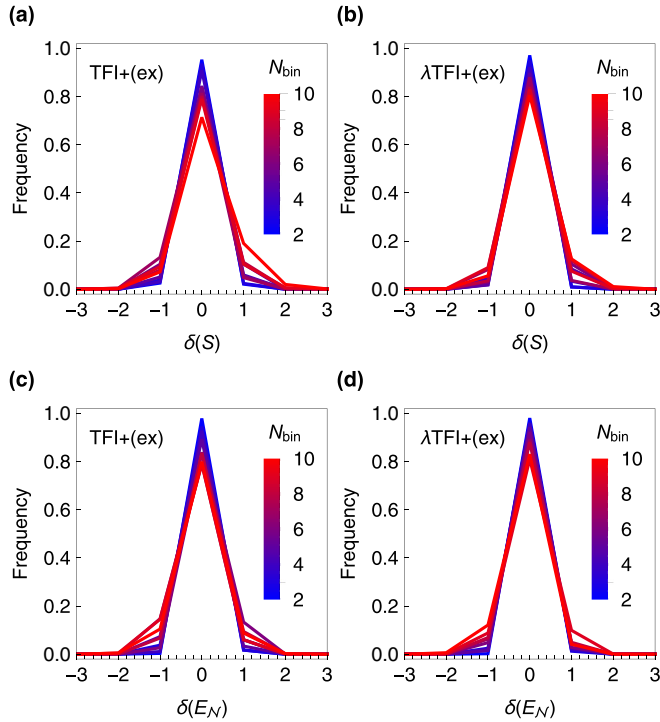
FIG. 6. Distribution of the signed binning error distance $\delta$ in entanglement quantification of the ANN when trained and tested on excited states of (a) and (c) the ferromagnetic transverse-field Ising model $H_{\mathrm{TFI+}}$ and (b) and (d) a nonintegrable version of it, $H_{\lambda\mathrm{TFI+}} = H_{\mathrm{TFI+}} + \lambda \sum_{<i,j>} \sigma_i^z \sigma_j^z$ with $\lambda = 0.3$, over a range of $2 \leqslant N_{\mathrm{bin}} \leqslant 10$. The network accurately quantifies entanglement independent of the entanglement measure, where in (a) and (b) the network was trained on images labeled by their half-chain entanglement entropy $S$ (binned), while in (c) and (d) the network was trained on images labeled by their half-chain logarithmic negativity $E_{\mathcal{N}}$ (binned). The respective error distances $\delta$ are measured in units of bins.

## APPENDIX C: NUMERICAL TOOLS

The ground and excited states are solved via exact diagonalization for spin chains of $L = 10$ sites. To speed up the process of image generation, ITensor [33] was employed to solve for the ground state of $H_{\mathrm{TFI+}}$ on a spin chain with $L = 18$ sites as well as perform the localized unitary rotations [inset of Fig. 1(c)]. In this case, smaller images with $\mathcal{W} = 5$ had to be used due to limited computational memory. The time evolution of a given initial state is numerically solved by employing the `mesolve` function of Quantum Toolbox in PYTHON (QuTiP) [34,35].

## APPENDIX D: NETWORK PERFORMANCE ON EXCITED STATES OF THE NONINTEGRABLE ISING MODEL BASED ON DIFFERENT ENTANGLEMENT MEASURES

We explore the network performance on excited states of the nonintegrable Ising model based on different entanglement measures (Fig. 6). For both measures investigated – the entanglement entropy $S$ and the logarithmic negativity $E_{\mathcal{N}}$ – the network performs equally well as compared to the integrable cases shown in Fig. 1(c) of the main text.
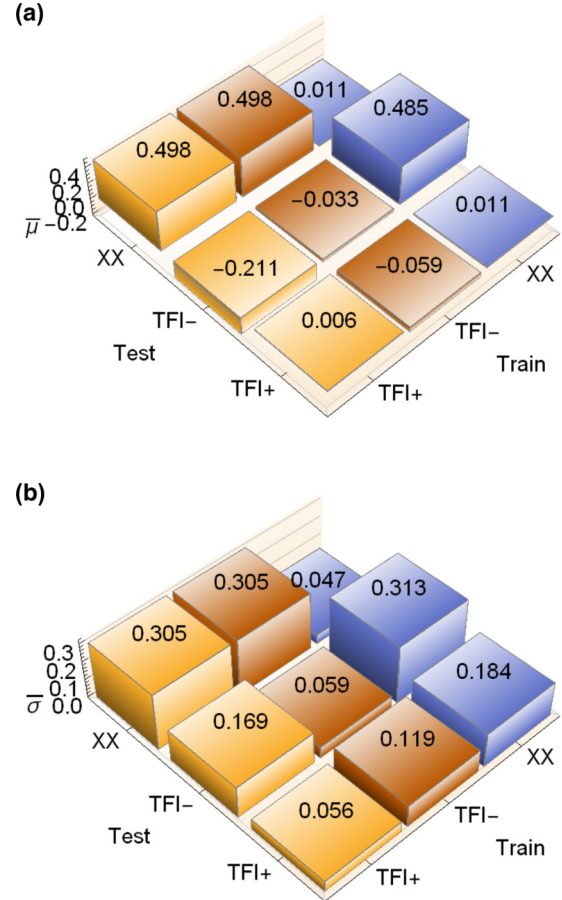


FIG. 7. Characterization of network performance on weakly entangled states. (a) The mean $\mu$ and (b) the standard deviation $\sigma$ of the binning error of the network for all possible training-test combinations from these three classes. The results shown are for the case of $N_{\mathrm{bin}} = 10$.

## APPENDIX E: GENERALIZABILITY OF ANN TO GROUND STATES OF DIFFERENT CLASSES

We explore the capabilities of the ANN to generalize to unfamiliar data by testing the network with states from model classes different from those it was trained on. A respective summary of all training-test combinations is shown in Figs. 7(a) and 7(b) for $N_{\mathrm{bin}} = 10$ containing both the mean $\mu$ and the standard deviation $\sigma$ of $\delta S$. As expected, the network performs best when tested on the same class of states that it was trained on. However, it can also generalize to different states in some instances, e.g., when testing on TFI+ states, albeit with a slightly poorer performance and in a nonreciprocal fashion. These observations suggest that while the ANN primarily learns model-specific features of the entanglement, it does in fact learn also about some universal features of quantum entanglement, the extent of which appears to depend on the type of states used for training.

## APPENDIX F: ENTANGLEMENT QUANTIFICATION WITH ALTERNATE PROTOCOLS

In addition to the (simplest) protocol $U_i = u_{i,1} \otimes \cdots \otimes u_{i,L}$ presented in the main text, we have considered statistical
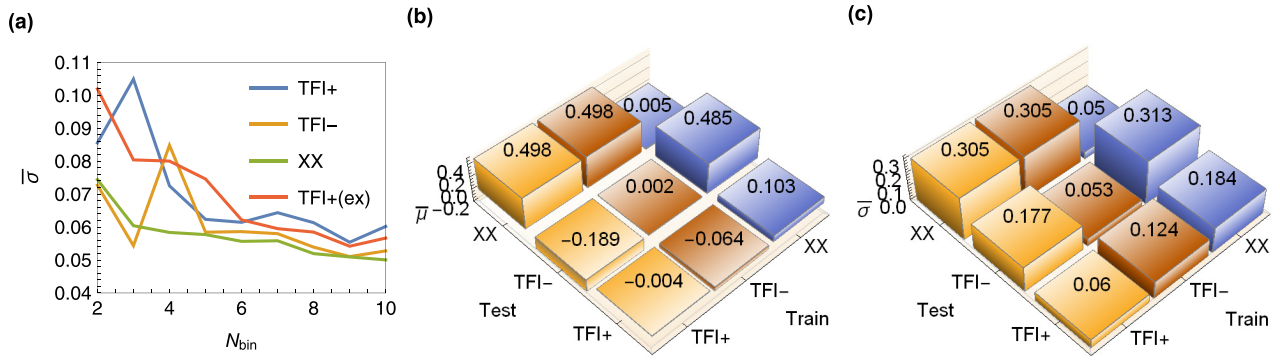
FIG. 8. Analogs of (a) the right panel of Fig. 2(a) in the main text and (b) and (c) Figs. 7(a) and 7(b), respectively, for protocol $U_{i>1}^A$.
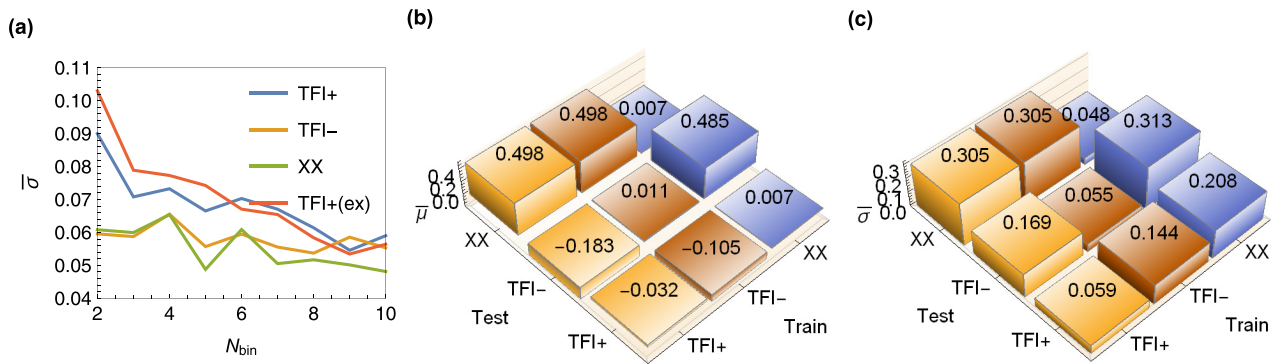


FIG. 9. Analogs of (a) the right panel of Fig. 2(a) in the main text and (b) and (c) Figs. 7(a) and 7(b), respectively, for protocol $U_{i>1}^B$.
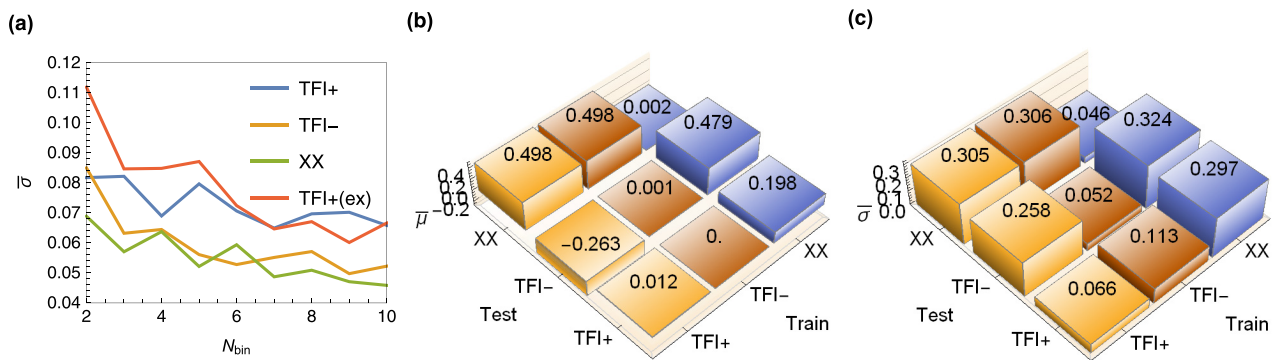


FIG. 10. Analogs of (a) the right panel of Fig. 2(a) in the main text and (b) and (c) Figs. 7(a) and 7(b), respectively, for protocol $U_{i>1}^C$.

image generation via alternate protocols. To illustrate the independence of protocols respecting the symmetry of the (half-chain) entanglement measures, we show in Figs. 8–10 the network performance based on the following multiqubit unitary transformations:

$$U_{i>1}^A = u_{i,(1,5)} \otimes u_{i,2} \otimes u_{i,3} \otimes u_{i,4} \otimes u_{i,6} \otimes \cdots \otimes u_{i,L}, \quad \text{(F1)}$$

$$U_{i>1}^B = u_{i,(1,4,5)} \otimes u_{i,2} \otimes u_{i,3} \otimes u_{i,6} \otimes \cdots \otimes u_{i,L}, \quad \text{(F2)}$$

$$U_{i>1}^C = u_{i,(1,2,3,4,5)} \otimes u_{i,(6,7,8,9,10)}, \quad \text{(F3)}$$

with $i = 2, \ldots, \mathcal{W}$, where $u_{i,(j_1,j_2,\ldots,j_n)}$ denotes an SU($2^n$) unitary operator drawn independently from the circular unitary ensemble (CUE) [24] that acts on sites $j_1, j_2, \ldots, j_n$. The performances based on the different protocols are essentially identical when testing on the same class of ground states that the ANN was trained on but vary in the extent to which they are able to generalize to the other classes.

### APPENDIX G: ABILITY OF THE NETWORK TO INTERPOLATE

We show here an interesting finding that the network is able to interpolate for the case of dissipative dynamics. Specifically, we trained a network on images drawn from states evolved under $\partial_t \rho = -i[H_{\text{TFI}+}, \rho] + \gamma \sum_{i=1}^{L} (\sigma_i^x \rho \sigma_i^x - \rho)$ with $\gamma = 0.01$ and $\gamma = 0.03$ and tested it on states evolved
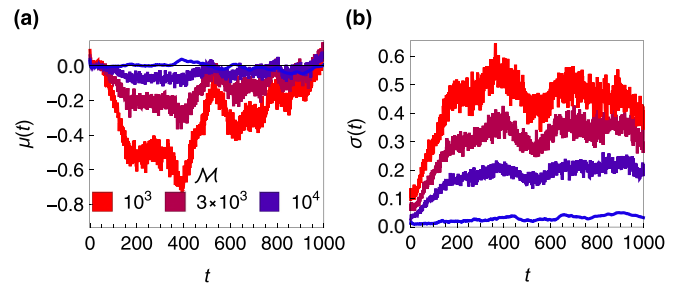


FIG. 11. Ability of network to interpolate in the dissipation parameter space. (a) The mean $\mu(t)$ and (b) standard deviation $\sigma(t)$ of the prediction error in $E_{\mathcal{N}}(t)$ analogous to Figs. 4(a) and 4(b) for the network trained on dissipative evolution with $\gamma = 0.01$ and $\gamma = 0.03$ but tested on $\gamma = 0.02$.

under the same dissipative dynamics but with $\gamma = 0.02$. The network performance is shown in Fig. 11. In the absence of image noise, remarkably, the network is able to accurately predict the entanglement dynamics with a performance comparable to the case in which it was tested with the same dissipation parameter as it was trained on, shown for the case of $\gamma = 0.01$ in Figs. 4(a) and 4(b) of the main text. In the presence of sampling noise, the quantitative deviation is two to three times worse compared with those of Figs. 4(a) and 4(b).

[1] J. P. Dowling and G. J. Milburn, Quantum technology: the second quantum revolution, Philos. Trans. R. Soc., A **361**, 1655 (2003).

[2] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, Entanglement in many-body systems, Rev. Mod. Phys. **80**, 517 (2008).

[3] N. Laflorencie, Quantum entanglement in condensed matter systems, Phys. Rep. **646**, 1 (2016).

[4] K. Kim, M.-S. Chang, S. Korenblit, R. Islam, E. E. Edwards, J. K. Freericks, G.-D. Lin, L.-M. Duan, and C. Monroe, Quantum simulation of frustrated Ising spins with trapped ions, Nature (London) **465**, 590 (2010).

[5] P. Jurcevic, B. P. Lanyon, P. Hauke, C. Hempel, P. Zoller, R. Blatt, and C. F. Roos, Quasiparticle engineering and entanglement propagation in a quantum many-body system, Nature (London) **511**, 202 (2014).

[6] T. Lanting, A. J. Przybysz, A. Yu. Smirnov, F. M. Spedalieri, M. H. Amin, A. J. Berkley, R. Harris, F. Altomare, S. Boixo, P. Bunyk, N. Dickson, C. Enderud, J. P. Hilton, E. Hoskinson, M. W. Johnson, E. Ladizinsky, N. Ladizinsky, R. Neufeld, T. Oh, I. Perminov *et al.*, Entanglement in a Quantum Annealing Processor, Phys. Rev. X **4**, 021041 (2014).

[7] T. Fukuhara, S. Hild, J. Zeiher, P. Schauß, I. Bloch, M. Endres, and C. Gross, Spatially Resolved Detection of a Spin-Entanglement Wave in a Bose-Hubbard Chain, Phys. Rev. Lett. **115**, 035302 (2015).

[8] B. P. Lanyon, C. Maier, M. Holzäpfel, T. Baumgratz, C. Hempel, P. Jurcevic, I. Dhand, A. S. Buyskikh, A. J. Daley, M. Cramer, M. B. Plenio, R. Blatt, and C. F. Roos, Efficient

[9] tomography of a quantum many-body system, Nat. Phys. **13**, 1158 (2017).

[9] A. J. Daley, H. Pichler, J. Schachenmayer, and P. Zoller, Measuring Entanglement Growth in Quench Dynamics of Bosons in an Optical Lattice, Phys. Rev. Lett. **109**, 020505 (2012).

[10] D. A. Abanin and E. Demler, Measuring Entanglement Entropy of a Generic Many-Body System with a Quantum Switch, Phys. Rev. Lett. **109**, 020504 (2012).

[11] R. Islam, R. Ma, P. M. Preiss, M. Eric Tai, A. Lukin, M. Rispoli, and M. Greiner, Measuring entanglement entropy in a quantum many-body system, Nature (London) **528**, 77 (2015).

[12] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, Quantum thermalization through entanglement in an isolated many-body system, Science **353**, 794 (2016).

[13] A. Lukin, M. Rispoli, R. Schittko, M. E. Tai, A. M. Kaufman, S. Choi, V. Khemani, J. Léonard, and M. Greiner, Probing entanglement in a many-body–localized system, Science **364**, 256 (2019).

[14] S. J. van Enk and C. W. J. Beenakker, Measuring Tr$\rho^n$ on Single Copies of $\rho$ Using Random Measurements, Phys. Rev. Lett. **108**, 110503 (2012).

[15] A. Elben, B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller, Rényi Entropies from Random Quenches in Atomic Hubbard and Spin Models, Phys. Rev. Lett. **120**, 050406 (2018).

[16] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Probing Rényi entanglement entropy via randomized measurements, Science **364**, 260 (2019).

[17] S. Lu, S. Huang, K. Li, J. Li, J. Chen, D. Lu, Z. Ji, Y. Shen, D. Zhou, and B. Zeng, Separability-entanglement classifier via machine learning, Phys. Rev. A **98**, 012315 (2018).

[18] Y.-C. Ma and M.-H. Yung, Transforming Bell's inequalities into state classifiers with machine learning, npj Quantum Inf. **4**, 34 (2018).

[19] R. Blatt and C. F. Roos, Quantum simulations with trapped ions, Nat. Phys. **8**, 277 (2012).

[20] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, and A. Wallraff, Rapid High-Fidelity Single-Shot Dispersive Readout of Superconducting Qubits, Phys. Rev. Appl. **7**, 054020 (2017).

[21] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, Rapid High-Fidelity Multiplexed Readout of Superconducting Qubits, Phys. Rev. Appl. **10**, 034040 (2018).

[22] S. Touzard, A. Kou, N. E. Frattini, V. V. Sivak, S. Puri, A. Grimm, L. Frunzio, S. Shankar, and M. H. Devoret, Gated Conditional Displacement Readout of Superconducting Qubits, Phys. Rev. Lett. **122**, 080502 (2019).

[23] S. Kuhr, Quantum-gas microscopes: a new tool for cold-atom quantum simulators, Natl. Sci. Rev. **3**, 170 (2016).

[24] F. Mezzadri, How to generate random matrices from the classical compact groups, Not. Am. Math. Soc. **54**, 592 (2007).

[25] G. Vidal and R. F. Werner, Computable measure of entanglement, Phys. Rev. A **65**, 032314 (2002).

[26] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, Nat. Phys. **13**, 431 (2017).

[27] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Machine Learning Phases of Strongly Correlated Fermions, Phys. Rev. X **7**, 031038 (2017).

[28] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Machine learning quantum phases of matter beyond the fermion sign problem, Sci. Rep. **7**, 8823 (2017).

[29] I. Cong, S. Choi, and M. D. Lukin, Quantum convolutional neural networks, Nat. Phys. **15**, 1273 (2019).

[30] A. Bohrdt, C. S. Chiu, G. Ji, M. Xu, D. Greif, M. Greiner, E. Demler, F. Grusdt, and M. Knap, Classifying snapshots of the doped Hubbard model with machine learning, Nat. Phys. **15**, 921 (2019).

[31] K. J. Satzinger, Y. Liu, A. Smith, C. Knapp, M. Newman, C. Jones, Z. Chen, C. Quintana, X. Mi, A. Dunsworth, C. Gidney, I. Aleiner, F. Arute, K. Arya, J. Atalaya, R. Babbush, J. C. Bardin, R. Barends, J. Basso, A. Bengtsson *et al.*, Realizing topologically ordered states on a quantum processor, arXiv:2104.01180 [quant-ph].

[32] G. Semeghini, H. Levine, A. Keesling, S. Ebadi, T. T. Wang, D. Bluvstein, R. Verresen, H. Pichler, M. Kalinowski, R. Samajdar, A. Omran, S. Sachdev, A. Vishwanath, M. Greiner, V. Vuletic, and M. D. Lukin, Probing topological spin liquids on a programmable quantum simulator, arXiv:2104.04119 [quant-ph].

[33] M. Fishman, S. R. White, and E. M. Stoudenmire, The ITensor software library for tensor network calculations, arXiv:2007.14822 [cs.MS].

[34] J. R. Johansson, P. D. Nation, and F. Nori, Qutip: An open-source Python framework for the dynamics of open quantum systems, Comput. Phys. Commun. **183**, 1760 (2012).

[35] J. R. Johansson, P. D. Nation, and F. Nori, Qutip 2: A Python framework for the dynamics of open quantum systems, Comput. Phys. Commun. **184**, 1234 (2013).