Quantum Loop Topography for Machine Learning

Yi Zhang and Eun-Ah Kim

Quantum Loop Topography for Machine Learning

Yi Zhang* and Eun-Ah Kim†

Department of Physics, Cornell University, Ithaca, New York 14853, USA and
Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA

(Dated: April 3, 2017)

Despite rapidly growing interest in harnessing machine learning in the study of quantum many-body systems, training neural networks to identify quantum phases is a nontrivial challenge. The key challenge is in efficiently extracting essential information from the many-body Hamiltonian or wave function and turning the information into an image that can be fed into a neural network. When targeting topological phases, this task becomes particularly challenging as topological phases are defined in terms of non-local properties. Here we introduce quantum loop topography (QLT): a procedure of constructing a multi-dimensional image from the "sample" Hamiltonian or wave function by evaluating two-point operators that form loops at independent Monte Carlo steps. The loop configuration is guided by characteristic response for defining the phase, which is Hall conductivity for the cases at hand. Feeding QLT to a fully-connected neural network with a single hidden layer, we demonstrate that the architecture can be effectively trained to distinguish Chern insulator and fractional Chern insulator from trivial insulators with high fidelity. In addition to establishing the first case of obtaining a phase diagram with topological quantum phase transition with machine learning, the perspective of bridging traditional condensed matter theory with machine learning will be broadly valuable.

Introduction– Machine learning techniques have been enabling neural networks to recognize and interpret big data sets of images and speeches[1]. Through supervised training with a large set of data, neural networks 'learn' to recognize key features of a universal class. Very recently, rapid and promising development has been made from this perspective on numerical studies of condensed matter systems, including dynamical systems[2–6] as well as classical and quantum systems undergoing phase transitions[7–13]. Also established is the theory connection to renormalization group[14, 15]. Exciting successes in application of machine learning to symmetry broken phases[7–10] may be attributed to the locality of the defining property of the target phases: the order parameter field. The snap-shots of order parameter configuration form images readily fed into neural networks well-developed to recognize patterns in images.

Unfortunately many novel states cannot be detected through a local order parameter. All topological phases are intrinsically defined in terms of non-local topological properties. Not only many-body localized states of growing interest[16] fit into this category, even a superconducting state fits in here since the superconducting order parameter explicitly breaks particle number conservation[17]. In order for neural networks to learn and recognize such phases, we need to supply them with "images" that contain relevant non-local information. Clearly information based on single site is insufficient. One approach to topological phase was to add complexity to the neural network architecture and use layers of convolutional filters to detect local constraints in the presence of translational symmetry, targeting a single topological phase at a time[7, 10]. Another approach was to detect the topological edge states[13]. In addition, ensemble of the Green's function was used to detect charge-ordered phases[9].

Here we introduce quantum loop topography (QLT): a procedure that designs and selects the input data based on the target phases of interest guided by relevant response functions. We focus on the fermionic topological phases but the procedure can be generalized to other situations that are not captured by purely local information, as all physically meaningful states are characterized by their response functions. The subject of topological phases of matter has grown with the appeal that topological properties are non-local and hence more robust[18–20]. Ironically this attractive feature makes it difficult to detect and identify topological phases even in numerics. Importantly, detection of strongly-correlated topological phases as fractional quantum Hall states[21, 22], fractional Chern insulators[23, 24], quantum spin liquids[25–27] requires arduous calculations of topological entanglements entropies[28, 29]. On the other hand, quantization[21–24, 30–34] is a natural theme of topological states and one may wonder whether there is an intelligent way to detect them due to the discreteness in defining properties. In this letter we demonstrate that QLT enables even a rather simple architecture consisting of a fully-connected neural network with a single hidden layer to recognize Chern insulator and fractional Chern insulator states and rapidly produce a phase diagram containing topological quantum phase transition (TQPT). We then discuss insights into the effectiveness of QLT and future directions based on its versatility.

Quantum Loop Topography and our algorithm– The procedure we dubbed QLT constructs an input image from a given Hamiltonian or many-body wave function that contains minimal but sufficient amount of non-local information. The response function that characterizes the phase of interest determines the geometry of the loop
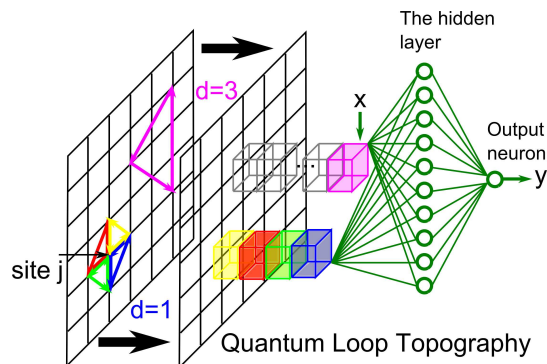
FIG. 1. Schematic illustration of our machine learning algorithm consisting of QLT and a neural network architecture. QLT for each site j consists of 4 loops of linear dimension $d = 1$. One loop of $d = 3$ is also shown for illustration. QLT of $d \leq d_c$ form a $D(d_c)$-dimensional vector for each site j, e.g., $D(1) = 4$ on a square lattice.

objects that enter QLT. But instead of brute-force evaluation of the response functions, we use QLT obtained from instances of Monte Carlo to train and utilize a network.

For Chern insulators, the relevant response function is the Hall conductivity. Interestingly Kitaev [35] pointed out that

$$\sigma_{xy} = \frac{e^2}{h} \cdot \frac{1}{N} \sum 4\pi i P_{jk} P_{kl} P_{lj} S_{\triangle jkl} \tag{1}$$

for free fermion systems[36], where $P_{ij} \equiv \langle c_i^\dagger c_j \rangle$ is the equal-time two-point correlation function between sites i and j, $S_{\triangle jkl}$ is the signed area of triangle jkl, and N is the total number of sites. Taking hints from Eq. (1) we use triangular loops to define QLT for Chern insulators. But instead of the full expectation value for two-point correlation functions in Eq. (1) which are costly to evaluate (requiring many instances of Monte Carlo walking down the Markov chain), we evaluate the bilinear operator with a single Monte Carlo sample α defining $\tilde{P}_{jk}|_\alpha \equiv \langle c_j^\dagger c_k \rangle_\alpha$. Further we note that smaller triangles dominate in a gapped system and keep the loops of linear dimension $d \leq d_c$, where d_c is the cut-off.

Now we define QLT to be a quasi-two-dimensional "image" of $D(d_c)$-dimensional vector of complex numbers assigned to each lattice site j. Each entry of this vector is associated with a distinct triangle cornered at site j which defines a chained product

$$\tilde{P}_{jk}|_\alpha \tilde{P}_{kl}|_\beta \tilde{P}_{lj}|_\gamma \tag{2}$$

where k and l are two other sites of the particular triangle, and $\tilde{P}$'s are evaluated at three independent Monte Carlo steps without averaging over Markov chain. $D(d_c)$ is the total number of triangles with $d \leq d_c$ assigned to each site(see Fig. 1). This way, QLT can be systemat-

ically expanded to include longer-ranged correlations by increasing cut-off d_c until convergence.

By construction QLT is quite versatile. Firstly, QLT can be obtained for and mapped between different lattice geometry, which only enters through $D(d_c)$. Secondly, the entire procedure takes place in real space without need for diagonalization or flux insertion and does not depend on translational invariance. Hence QLT should naturally accommodate heterogeneity, disorder and interaction. Finally, it is clear that the strategy underlying QLT construction we have laid out can be generalized to other novel phases such as $\mathbb{Z}_2$ topological order, quantum spin Hall insulator, or superconductivity[37]. In the rest of this paper we use Variational Monte Carlo(VMC), without loss of generality, to build QLT by sampling the many-body ground states of interest at randomly selected Monte Carlo steps (see Supplemental Material).

Once QLT is obtained for a given model, we feed it to a neural network(Fig. 1). For this, we design a feed-forward fully-connected neural network with only one hidden layer consisting of $n = 10$ sigmoid neurons. The network takes QLT as an input x and each neurons processes the input through independent weights and biases $w \cdot x + b$. After the sigmoid function, the outcome is fed forward to be processed by the output neuron. The final output y corresponds to the neural network's judgement whether the input QLT is topological. We use cross entropy as the cost function with L2 regularization to avoid over-training and a mini-batch size of 10[1]. For the rest of this paper, we use randomly-mixed 20000 data samples within the VMC Metropolis of the topological and trivial phases as the training group. We reserve a separate group of 4000 data samples (also half trivial and half topological) for validation purposes including learning speed control and termination[1]. Once the machine learning is successful, the trained neural network can rapidly process QLT's from different parts of the phase space to yield a phase diagram. In order to establish level of confidence on the trained network's assessment of whether the system is topological or not, we process 2000 QLT's at each point and take the ratio p of 'topological' output, i.e., $y > 0.5$. When p is close to 1 for topological phase and 0 for trivial phase, it indicates even a single QLT can reliably land a trustworthy detection.

Topological quantum phase transition in a free fermion model– We first apply the QLT-based machine learning to the TQPT between trivial insulator and Chern insulator. Consider the following tight-binding model on a square lattice:

$$H(\kappa) = \sum_{\vec{r}} (-1)^y c_{\vec{r}+\hat{x}}^\dagger c_{\vec{r}} + [1 + (-1)^y (1 - \kappa)] c_{\vec{r}+\hat{y}}^\dagger c_{\vec{r}}$$

$$+ (-1)^y \frac{i\kappa}{2} \left[c_{\vec{r}+\hat{x}+\hat{y}}^\dagger c_{\vec{r}} + c_{\vec{r}+\hat{x}-\hat{y}}^\dagger c_{\vec{r}} \right] + \text{h.c.} \tag{3}$$

where $\vec{r} = (x, y)$ (see Fig. 2) and $0 \leq \kappa \leq 1$ is a tuning parameter. The $\kappa = 1$ limit is the π-flux square lattice
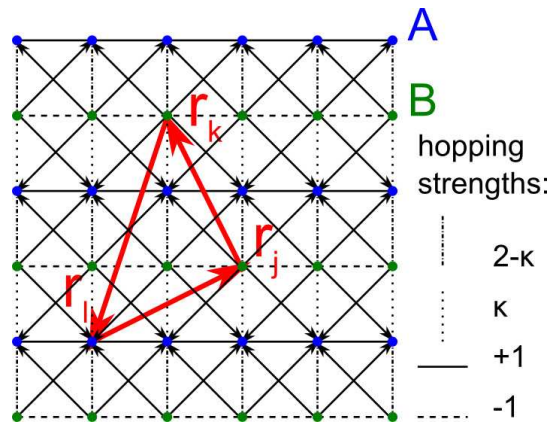
FIG. 2. Model illustration of Eq. 3. The unit cell consists of two sublattice sites A and B. Hopping strengths are different for horizontal and vertical bonds and staggered. The diagonal hopping is $i\kappa$ ($-i\kappa$) along (against) the arrow. The red arrows denotes a triangle that defines the operators of our **QLT**.

FIG. 3. The ratio p of 'topological' response from the neural network on the model in Eq. 3 over the parameter region $\kappa \in [0.1, 1, 0]$. The neural network is trained with $\kappa = 0.1$ for $y = 0$ and $\kappa = 1$ for $y = 1$. The green square symbols represent the results using fermion occupation configurations as an input data. Red dashed line marks the expected topological phase transition at $\kappa = 0.5$. The inset: an enlarged view over the critical region $0.4 \leq \kappa \leq 0.6$. $d_c = 2$ for all.

model for a Chern insulator with Chern number $C = 1$ [31], while the $\kappa = 0$ limit amounts to decoupled two-leg ladders. $H(\kappa)$ interpolates between Chern insulator and trivial insulator with a TQPT at $\kappa = 0.5$. To see this, Fourier transform the Hamiltonian to momentum space

$$H\left(\kappa\right) = \sum_k \left[2\cos k_y + 2i\sin k_y \left(1 - \kappa + \kappa \sin k_x\right)\right] c_{k,A}^\dagger c_{k,B}$$
$$+ 2\cos k_x (c_{k,A}^\dagger c_{k,A} - c_{k,B}^\dagger c_{k,B}) + \text{h.c.} \quad (4)$$

where A and B label the two sublattices, and note that in this two-band model with two Dirac points at $(\pm\pi/2, \pi/2)$, one of the Dirac masses changes sign across $\kappa = 0.5$.

Our complete knowledge of the phase diagram makes Eq. 3 an ideal testing ground for our algorithm. Hence we implement supervised machine learning using two deep-in-phase models of $\kappa = 1.0$ (Chern insulator) and $\kappa = 0.1$ (trivial insulator) for training[38]. The system size is 12×12 lattice spacings unless noted otherwise. First we establish that indeed single-site-based inputs of the fermion occupation configurations $n(\vec{r}) = c_{\vec{r}}^\dagger c_{\vec{r}}$ fail to transmit the topological information to the neural network, as we expected. With $n(\vec{r})$ as inputs, the learning is inefficient, signaled by high yields in the cost function[1]. Moreover, as shown in Fig. 3, even after a long period of training the neural network keeps incorrectly judging the system to be a trivial insulator for all values of κ, except for $\kappa = 1.0$ where the result returns $> 80\%$ 'nontrivial'. This indicates that the neural network unfortunately does not pick up the universal features about the topological phase, but rather memorizes the details of the specific model at $\kappa = 1.0$.

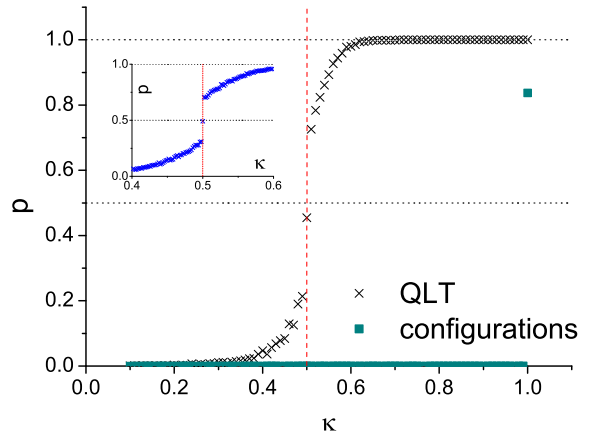The contrast in the results based on QLT input is striking. Fig. 3 shows that the trained network's assessment

achieves $> 99.9\%$ accuracy deep in either the topological phase or trivial phase even with $d_c = 2$. Moreover, even though we have provided the training group with only large-gap models in both topological and trivial phases focusing on identifying phases[39], we find a non-analytical behavior in p as a function of κ at the critical point [see Fig. 3 inset].

Generalizations– Next we consider a fractional Chern insulator (FCI) as an example of strongly-correlated topological phase. Here the $\nu = 1/3$ FCI is represented by a VMC wave function that is the free fermion wave function of the model in Eq. 3 raised to the third power[40]. Surprisingly the neural network trained on non-interacting parent Chern insulator already serves as a 'poor man's network' (see the inset of Fig. 4). This network recognizes that FCI phase is distinct from the parent Chern insulator and hence it only gives $p \sim 0.01$ 'nontrivial' response for the FCI phase. Nevertheless it also notices that FCI is a topologically distinct state from the trivial insulator since $p \sim 0.01$ is large enough to exclude statistical error. Once trained with the FCI wave function at two reference points $\kappa = 0.1$ for trivial and $\kappa = 1.0$ for FCI, the network once again detects FCI phase with high accuracy [see Fig. 4].

Remarkably the network automatically recognizes topological degeneracy. Even we train the network with only one of the degenerate ground states in the topological phase (GS#1 in Fig. 4), it correctly assess topological nature of the two other ground states that are related to GS#1 by flux threading. The network also
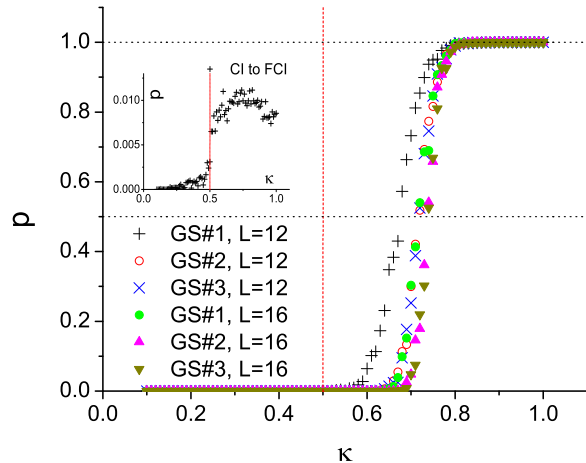
FIG. 4. Application to a $\nu = 1/3$ FCI. The topological phase transition in the *parent* Chern insulator at $\kappa = 0.5$ is marked by a vertical red dashed line. The inset shows the results using neural network trained with the parent free fermion model, where p is calculated over 20000 samples for each κ to reduce statistical error. The main panel shows the results using FCI wave functions for both training ($\kappa = 0.1$ for trivial and $\kappa = 1.0$ for the FCI, first ground state only) and testing (all three degenerate ground states, see Supplemental Material). $L = 16$ data is shown in addition to $L = 12$ to help attribute the differences between κ_c of the topological phase transitions to the finite-size effect. $d_c = 2$ for all.

detects a TQPT at $0.67 \leq \kappa_c \leq 0.77$. The uncertainty in the critical value κ_c is a finite-size effect on the quasi-degeneracy as is clear from the convergence upon increasing the system size. The fact that $\kappa_c > 0.5$ when the single-particle gap closes at $\kappa = 0.5$ could raise concern in light of the findings on single-particle Green's function based approaches[41, 42]. We note that single-particle gap is a pre-requisite for the VMC wave function to represent a topological phase since only then partons may be integrated out, hence the shift of $\kappa_c > 0.5$ is expected. Nevertheless, the result is the first report of the TQPT providing the target, and calls for further study on the critical point using independent measures such as many-body gap or the more established yet costly entanglement entropy based approaches[28, 29, 40, 43, 44].

Finally, we demonstrate that we can train the network to learn the topological protection of the topological phases, and show indifference to the microscopic details such as lattice structure or impurities. The key to successful machine learning this celebrated feature is the diversity of the training inputs. Without diverse inputs, the network looks for features that are specific to its training set. For instance, the network trained only with square lattice cannot recognize the topological phase in the honeycomb lattice. But if we provide diverse input taken from both the square lattice and the honeycomb
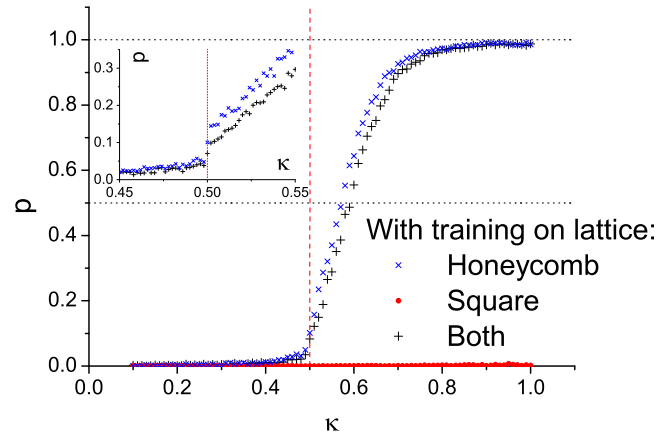

FIG. 5. The ratio p of 'topologically nontrivial' response from the neural networks for the honeycomb lattice model (Supplemental Material) over the parameter region $\kappa \in [0.1, 1, 0]$. The topological phase transition is at $\kappa = 0.5$ (vertical red dashed line). The neural networks are trained using the Chern insulators and trivial insulators only on the honeycomb lattice, only on the square lattice, and on both. The inset zooms into the critical region $0.4 \leq \kappa \leq 0.6$. $d_c = 2$ for all.

lattice systems, the network can be trained to recognize topological phases on both lattices with little penalty on accuracy (see Fig. 5). We also note that the network recognizes the difference between different Chern numbers (e.g., $C = -1$ v.s. $C = 1$) as well as higher Chern numbers (e.g., $C = 2$).

Conclusion– In summary, we have successfully implemented supervised machine learning for topological phases by introducing QLT as an interface between traditional concept of response theory and a simple neural network.

Three major strengths of our QLT-based machine learning approaches are 1) efficiency, 2) accuracy, and 3) versatility. Firstly, the network can be trained with quasi-two-dimensional QLT in gapped phases. Furthermore since QLT bypasses the time-consuming process of averaging over Markov chains, one can quickly scan the phase space once the network is trained. Although our focus was on the phases, we demonstrated that non-analyticity in the ratio of non-trivial response allows us to pinpoint the phase transition. Finally, as a real-space based formalism that does not requires translational symmetry, or diagonalization or flux insertion, QLT is quite versatile.

Our approach can also be applied to systems with other fillings as well as higher dimensions. The fact that QLT readily handles degenerate ground states adds to its versatility. Moreover there is nothing restricting QLT to VMC data. It can be applied to Hamiltonian-based quantum Monte Carlo samples[41, 42] as well as other rep-

resentations of many-body wave functions such as matrix product states and PEPS. Most importantly, the procedure of defining appropriate QLT guided by relevant response function we established here is readily expanded to other state of interest such as superconducting state, quantum spin Hall insulator, and $\mathbb{Z}_2$ topological order[37]. Hence our construction opens door to applications of the machine learning to novel states of broad interest.

* frankzhangyi@gmail.com
† eun-ah.kim@cornell.edu
[1] Michael Nielsen, *Neural Networks and Deep Learning* (Free Online Book, 2013).
[2] S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, and A. J. Liu, Nat Phys **12**, 469 (2016).
[3] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy, and R. Melko, arXiv:1601.02036.
[4] G. Carleo and M. Troyer, arXiv:1606.02318.
[5] G. Torlai and R. G. Melko, (), arXiv:1606.02718.
[6] G. Torlai and R. G. Melko, (), arXiv:1610.04238.
[7] J. Carrasquilla and R. G. Melko, arXiv:1605.01735.
[8] L. Wang, Phys. Rev. B **94**, 195105 (2016).
[9] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, arXiv:1608.07848.
[10] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, arXiv:1609.02552.
[11] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, arXiv:1610.03137.
[12] L. Huang and L. Wang, arXiv:1610.02746.
[13] T. Ohtsuki and T. Ohtsuki, arXiv:1610.00462.
[14] C. Bény, arXiv:1301.3124.
[15] P. Mehta and D. J. Schwab, arXiv:1410.3831.
[16] R. Nandkishore and D. A. Huse, Annual Review of Condensed Matter Physics **6**, 15 (2015).
[17] D. J. Scalapino, S. R. White, and S. Zhang, Phys. Rev. B **47**, 7995 (1993).
[18] A. Kitaev, Annals of Physics **303**, 2 (2003).
[19] S. Das Sarma, M. Freedman, and C. Nayak, Phys. Rev. Lett. **94**, 166802 (2005).
[20] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, Rev. Mod. Phys. **80**, 1083 (2008).
[21] R. B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983).
[22] D. C. Tsui, H. L. Stormer, and A. C. Gossard, Phys. Rev. Lett. **48**, 1559 (1982).
[23] T. Neupert, L. Santos, C. Chamon, and C. Mudry, Phys. Rev. Lett. **106**, 236804 (2011).
[24] N. Regnault and B. A. Bernevig, Phys. Rev. X **1**, 021014 (2011).
[25] V. Kalmeyer and R. B. Laughlin, Phys. Rev. Lett. **59**, 2095 (1987).
[26] X.-G. Wen, International Journal of Modern Physics B **04**, 239 (1990).
[27] L. Balents, Nature **464**, 199 (2010).
[28] A. Kitaev and J. Preskill, Phys. Rev. Lett. **96**, 110404 (2006).
[29] M. Levin and X.-G. Wen, Phys. Rev. Lett. **96**, 110405 (2006).
[30] K. v. Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).
[31] A. W. W. Ludwig, M. P. A. Fisher, R. Shankar, and G. Grinstein, Phys. Rev. B **50**, 7526 (1994).
[32] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988).
[33] M. Z. Hasan and C. L. Kane, Rev. Mod. Phys. **82**, 3045 (2010).
[34] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
[35] A. Kitaev, Annals of Physics **321**, 2 (2006).
[36] Our alternative proof of Eq. (1) (see Supplemental Material) builds on adiabatic continuity protected by the gap in the spectra without requiring the system be non-interacting.
[37] Y. Zhang, R. G. Melko, and E.-A. Kim, in preparation.
[38] The system becomes decoupled two-leg ladders at $\kappa = 0.0$, which is non-generic for a two-dimensional insulator and we choose to avoid such specialty.
[39] See Supplemental Material for further discussion and details on the impact of training models and QLT cut-off d_c on machine learning phases as well as phase transitions.
[40] Y. Zhang, T. Grover, and A. Vishwanath, Phys. Rev. B **84**, 075128 (2011).
[41] T. C. Lang, A. M. Essin, V. Gurarie, and S. Wessel, Phys. Rev. B **87**, 205101 (2013).
[42] Y.-Y. He, H.-Q. Wu, Z. Y. Meng, and Z.-Y. Lu, Phys. Rev. B **93**, 195164 (2016).
[43] Y. Zhang, T. Grover, A. Turner, M. Oshikawa, and A. Vishwanath, Phys. Rev. B **85**, 235151 (2012).
[44] Y. Zhang, T. Grover, and A. Vishwanath, Phys. Rev. B **91**, 035127 (2015).