Machine learning Z_{2} quantum spin liquids with quasiparticle statistics

Yi Zhang, Roger G. Melko, and Eun-Ah Kim

Machine Learning $\mathbb{Z}_2$ Quantum Spin Liquids with Quasi-particle Statistics

Yi Zhang[1],[*] Roger G. Melko[2,3], and Eun-Ah Kim[1][†]

[1] *Department of Physics, Cornell University, Ithaca, New York 14853, USA*
[2] *Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada and*
[3] *Department of Physics and Astronomy, University of Waterloo, Ontario, N2L 3G1, Canada*
(Dated: November 30, 2017)

After decades of progress and effort, obtaining a phase diagram for a strongly-correlated topological system still remains a challenge. Although in principle one could turn to Wilson loops and long-range entanglement, evaluating these non-local observables at many points in phase space can be prohibitively costly. With growing excitement over topological quantum computation comes the need for an efficient approach for obtaining topological phase diagrams. Here we turn to machine learning using quantum loop topography (QLT), a notion we have recently introduced. Specifically, we propose a construction of QLT that is sensitive to quasi-particle statistics. We then use mutual statistics between the spinons and visons to detect a $\mathbb{Z}_2$ quantum spin liquid in a multi-parameter phase space. We successfully obtain the quantum phase boundary between the topological and trivial phases using a simple feed forward neural network. Furthermore we demonstrate advantages of our approach for the evaluation of phase diagrams relating to speed and storage. Such statistics-based machine learning of topological phases opens new efficient routes to studying topological phase diagrams in strongly correlated systems.

INTRODUCTION

Despite much interest in topological phases of matter, the search for and detection of the finite regions of phase space that support topological order has been a long standing challenge. This is a nontrivial challenge because microscopic models of strongly-correlated topological order are usually established in exactly solvable models at first [1–7]. Nevertheless, universal properties of topological phases from low-energy effective theories, i.e. topological quantum field theories, provide a broader support for the understanding of topological phases. Naturally, much effort has gone into perturbing exactly solvable models both theoretically [8–13] and numerically [14–20] to investigate the stability of the associated topological phases beyond the fine-tuned solvable points. Moreover, growing enthusiasm over the idea of using exotic statistics of excitations for topological quantum computation [1, 3, 21] and the experimental quests driven by related proposals [22, 23] have raised the need to understand the stability of various topological phases and establish the corresponding phase diagrams.

It is perhaps the key interesting feature of topological phases that simultaneously underlies challenges in their numerical diagnosis: the absence of local order parameter. Nevertheless various measures of non-local correlations have enabled progress in evaluating phase diagrams and detecting phase transitions. Among the most successful approaches are expectation value of Wilson loop[24, 25], and entanglement entropy[19, 26–28]. Unfortunately, the non-local nature of the respective estimators can make the algorithms for measuring these costly in general. In addition, in some cases one can use a thermodynamic signature such as specific heat to detect the phase transition [15, 16, 26, 29]. Although a singularity in specific heat is an effective indicator of a phase transition, it has the drawback that it does not reveal any information regarding the topological aspects of the associated

phases. Hence, in addition to these standard techniques, developing a cost-effective approach that can map out a phase diagram with topological quantum phase transitions using key features of the topological phase, such as non-trivial statistics, is highly desirable.

A new strategy for a dramatic speed-up in the approximate evaluation of phase diagrams is to use neural-network based machine learning [30]. Efforts in this direction fall into one of two broad categories: unsupervised learning and supervised learning. In the unsupervised context, the task of classifying raw state configurations with phase labels (e.g. "clustering" to find hidden patterns or grouping in data) is one actively-pursued goal. Several different approaches have been used, including principal component analysis and neural networks [31–33], resulting in a rapidly-developing sub-field. Within the supervised learning approach, the algorithmic strategy is perhaps more well established. There, neural networks can be trained with data in the form of raw state configurations, each labelled by its respective phase. Once the neural network is trained, a new ("test") data set is given to it, and it is tasked with labelling each configuration with one of the phases it has been trained to recognize. This approach has been successful in obtaining phase diagrams with conventional ordered phases[34–36], many-body localizations[37], or chiral topological phases[38–40]. Nevertheless, identifying non-chiral topological order remains surprisingly challenging for such supervised machine learning approaches[34].

Here we propose a learning strategy based on the non-trivial statistics between fundamental excitations – the key defining property of correlated topological phases. This approach is inspired by earlier efforts to calculate quasi-particle statistics that rely on evaluating quantities reflecting long-range entanglement of correlated topological phases [41–48]. Since such calculations are typically computationally costly, attempting to sweep a large phase space can become prohibitively difficult. Hence we propose using a technique called quantum loop topography (QLT) [39], designed around quasi-particle

statistics in conjunction with neural-network based machine learning. The notion of the QLT was first introduced in Ref. [39] as a preprocessing interface between a quantum many-body state and a feed forward neural network. The key idea was to select and organize the input data based on the physical response defining the target phases to detect. This method was successfully implemented for chiral topological phases in Ref. [39]. Here we build a QLT protocol for non-chiral topological phases with non-trivial quasi-particle statistics. We then demonstrate the effectiveness of our QLT-based machine learning strategy by obtaining the phase diagram of the toric code in the presence of magnetic fields, mapping out the $\mathbb{Z}_2$ quantum spin liquid in the parameter space of the external field strengths.

The rest of this paper is organized as follows. In section II, we present the general strategy for evaluating phase diagrams involving topological phases using neural network based supervised machine learning with quantum loop topography. In Sec. III, we discuss the model of field-perturbed toric code and the QLT guided by the quasi-particle mutual statistics and specified for the detection of the $\mathbb{Z}_2$ quantum spin liquid phase. In Sec. IV, we present the architecture of our neural network and the algorithm for performing supervised machine learning. In Sec. V, we present the results bench-marked against conventional specific heat collapse method. We then close with a discussion and outlook.

QUANTUM LOOP TOPOGRAPHY FOR MACHINE LEARNING

The evaluation of a phase diagram using neural network based supervised machine learning is a process made of three key stages [34]. The first stage is to assemble diverse groups of input data representing each of the phases of interest. This collection of data sets forms the "training set". Each element of the training set is labeled with the respective phase. The second stage is to construct the neural network and train it to categorize the training set correctly by providing feedback with the known label. During the training process the neural network adjusts the parameters that define the network (the weight matrix w and bias vector b) to minimize the error in its output y (see Fig. 1). The final stage is to sweep through the phase space and a "test" data set for each phase space point is analyzed by the neural network. Such efficient and auto-mated study of the entire phase space is one important merit of the neural-network based approach and is fueling rapidly increasing efforts in this direction[34–36, 38, 39].

Quantum loop topography was introduced in Ref. [39] af-ter it became clear through previous efforts [35] that for effi-ciency or even success in training, it is critically advantageous to use input data containing information relevant to the target phases. QLT is a data preparation (or preprocessing) stage that builds on the characteristics of a phase [39]. The necessity for QLT arises when local information is insufficient for the phase identification as in topological phases (and in superconduct-
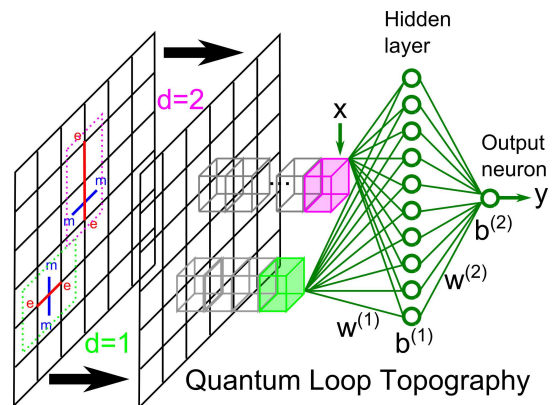

FIG. 1. Schematic illustration of our machine learning architec-ture. Quantum loop topography uses an ensemble of 'minimally non-local' operators to extract information from the many-body systems and creates 'image' with relevant information for the input into the artificial neural network.

ing phases). Traditionally, one would either use maximally non-local information such as entanglement entropy or explic-itly evaluate the relevant response function. Using QLT com-bined with machine learning, one avoids both of these time consuming approaches while keeping their key elements. A QLT is a "topographic" map made of un-averaged and there-fore fluctuating values of products of loop-forming operators that are relevant for the phase of interest. By retaining only the loops of sizes below certain cutoff, i.e. minimally non-local loops, a QLT can readily take advantage of the short cor-relation lengths in gapped phases, nevertheless incorporating non-local correlations.

The key to success for any QLT is to involve operators rel-evant for the defining properties of the target phase and this approach is particularly useful when usual equilibrium order parameter is not computationally accessible. Earlier, two of us introduced the notion of QLT and successfully implemented it targeting integer and fractional Chern insulator [39]. For such chiral topological phases, the natural response function to guide the construction of QLT was the Hall conductivity. On the other hand, of increasing importance are non-chiral topological phases with multiple types of fractional excita-tions. Here we focus on such non-chiral topological phases when one cannot employ the Chern number inspired QLT from Ref. [39]. The new defining propoerty we will focus on is the quasi-particle statistics.

One of the defining properties of intrinsic topological phases is the non-trivial exchange and braiding statistics that their quasi-particle excitations must obey. Such statistical in-formation is encoded into the topological quantum field the-ory associated with a given topological phase through the ex-pectation values of Wilson loops forming non-trivial knots. When the world-line of a quasi-particle forms a knot with that of another quasi-particle, such a knot can only be resolved through an appropriate unitary transformation in the quasi-particle Hilbert space [1, 49, 50]. To be specific, let us con-
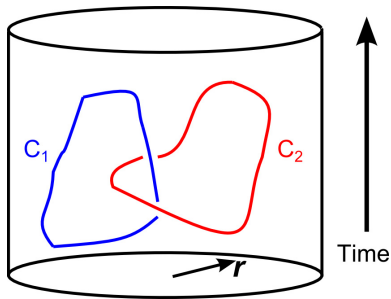
FIG. 2. A nontrivial knot with linking number $L = 1$ between two Wilson loops C_1 and C_2 in 2+1-dimension space-time.

sider a gapped Abelian topological phase. The long-distance (IR) effective theories associated with such phases are Chern Simons theories with K-matrices [51, 52]:

$$\mathcal{L}_{CS} = \frac{K_{IJ}}{4\pi} \epsilon^{\mu\nu\lambda} a_\mu^I \partial_\nu a_\lambda^J - a_\mu^I j_I^\mu \tag{1}$$

where a_μ^I are $U(1)$ gauge fields coupled to quasi-particle currents j_I^μ, $I = 1, 2, \cdots$ labels the fundamental types of quasi-particle excitations. For instance, the K matrix associated with Laughlin states at filling $\nu = 1/m$ has a single entry $K_{11} = m$ while that associated with $\mathbb{Z}_2$ quantum spin liquid is

$$K = \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}. \tag{2}$$

From the field theory perspective, the Wilson loops formed by the quasi-particle (quasi-hole) trajectory C defined by

$$W_C^I \equiv \mathcal{P} \exp\left(i \oint_C a_l^I dl\right), \tag{3}$$

where $\mathcal{P}$ is the path-ordering symbol, are the only gauge-invariant observables. The quasi-particle statistics are encoded in Wilson loops through [1, 49, 50]:

$$\mathcal{P} W_{C_1}^I W_{C_2}^J = e^{2\pi i K_{IJ}^{-1} L(C_1, C_2)}$$

where $L(C_1, C_2)$ is the space-time linking number between the C_1 and C_2 loops. For example, we illustrate a two-Wilson-loop configuration forming a nontrivial knot ($L = 1$) in Fig. 2. In a topological quantum field theory, the expectation values of Wilson loop configurations only depend on the topology irrespective of the size of Wilson loops involved, since a topological quantum field theory is essentially scale-free.

Nevertheless when topological phases are investigated through lattice models, microscopic details and short-distance (UV) physics enter the problem and long-distance (IR) physics will only appear at length scales much longer than the short-distance cutoff. As a result, previous numerical studies focused on the largest loops the system can support, e.g. the two cycles of a torus [41, 42, 46, 47]. Unfortunately, trying to numerically evaluate such global objects with precision is highly costly. Clearly, the statistical information would be
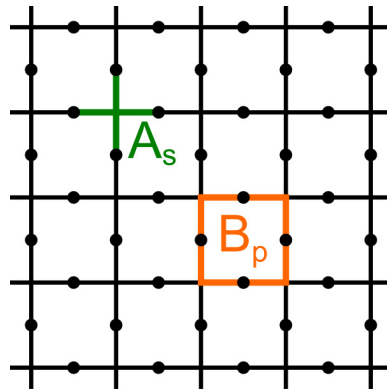

FIG. 3. Illustration of the lattice spin model in Eq. 4. The spin-1/2's reside on the bonds of the two-dimensional square lattice. A_s and B_p are the products of σ^x and σ^z operators around a site s and on the boundaries of a plaquette p, respectively.

present in the knots of smaller size as well [49, 50, 53], such as the one in Fig. 4b. However, the short-distance details and perturbations prevent us from reaching comparative theoretical predictions and making use of the information from such smaller knots, even though they can be much easier to study numerically.

Instead, we now propose using QLT based upon these semi-local knots to select and build the input data, which is analyzed from a novel perspective using machine learning with an artificial neural network (see Fig. 1). When supervised machine learning is successfully accomplished, the resulting architecture can scan the phase space and obtain the phase regions of the topological phases efficiently and automatically.

$\mathbb{Z}_2$ QUANTUM SPIN LIQUID AND QUANTUM LOOP TOPOGRAPHY

The $\mathbb{Z}_2$ quantum spin liquid is the prototypical example of a state with a non-chiral topological order defined on a lattice [1, 54]. It is a strongly-correlated quantum spin liquid, with four-fold ground-state degeneracy separated from the excited states with a full gap when defined on a torus. Its non-chiral nature and the lack of topological edge states make the detection of $\mathbb{Z}_2$ quantum spin liquid even more elusive. Importantly, its fundamental types of quasi-particle excitations, the spinon and the vison, both have trivial self statistics yet semionic mutual statistics: the system picks up an overall phase factor $e^{i\theta}$ with statistical angle $\theta = \pi$ upon braiding a spinon around a vison, or vice versa.

A model known to support a $\mathbb{Z}_2$ quantum spin liquid phase is [12, 15, 20]:

$$H_{2D} = -J_x \sum_s A_s - J_z \sum_p B_p - h_x \sum_j \sigma_j^x - h_z \sum_j \sigma_j^z \tag{4}$$

where the spin-1/2 lives on the bonds of a square lattice and $A_s = \prod_{j \in s} \sigma_j^x$ and $B_p = \prod_{j \in p} \sigma_j^z$ are the products of spin operators around a site s and on the boundaries of a plaquette
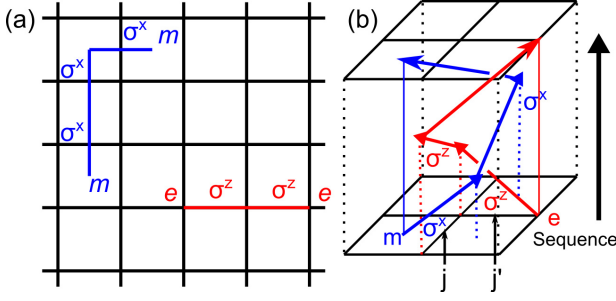
FIG. 4. (a) A string of σ^z (σ^x) operators creates at its two ends a pair of spinons (visons) denoted as 'e' ('m'), or equivalently, moves one spinon (vison) from one end to the other. Note that the spinon string C_e and vison string $\tilde{C}_m$ live on the bond of the original and dual square lattices, respectively. (b) Illustration of a nontrivial knot between the world-lines of a spinon and a vison as a discrete lattice version of Fig. 2.

p, respectively [see Fig. 3]. The h_z and h_x terms are external magnetic fields in the $\hat{x}$ and $\hat{z}$ directions, respectively. For the rest of the paper, we consider system size 12×12 unless noted otherwise.

At the special point $h_x = h_z = 0$, H_{2D} amounts to Kitaev's toric code [1]. The exact solvability of the toric code has allowed for much progress in explicit understanding of $\mathbb{Z}_2$ quantum spin liquid. Since all A_s and B_p commute in the Hamiltonian, the ground states are simply given by allowing $A_s = B_p = 1$, while the quasi-particle excitations are associated with violations that cost finite energy penalties: a spinon at site s with $A_s = -1$ and a vison in plaquette p with $B_p = -1$. Note that a string of σ^z (σ^x) operators $\prod_{j\in C}\sigma^z_j$ ($\prod_{j\in\tilde{C}}\sigma^x_j$) moves one spinon (vison) from one end of C ($\tilde{C}$) to the other end, see Fig. 4a. To illustrate the mutual statistics, consider the process $\prod_{j\in p}\sigma^z_j$ that circles a spinon around a plaquette p: it gives rise to a phase factor B_p, which is -1 if there exists a vison inside the plaquette.

The toric code is, however, non-generic - its correlation length is zero and quasi-particles are strictly not allowed in the ground states. We consider presence of magnetic fields: the h_z (h_x) magnetic field makes it more preferential to generate spinon (vison) pairs, which will eventually condense and give rise to a spin-polarized phase beyond a critical field strength. The two topologically-trivial phases, usually called magnetically ordered phase and disordered phase at either large h_z or large h_x are dual to each other under $\sigma_x \leftrightarrow \sigma_z$ transformation.

In practice, Eq. 4 can be mapped to a three-dimensional classical system - the (anisotropic) $\mathbb{Z}_2$ gauge Higgs model [11] via imaginary time evolution and Trotter decomposition [15]. After separating the imaginary time into a large number of small intervals $\beta = n\Delta\tau$ the operator $\exp(-\beta H)$ for the quantum partition function can be approximated by $[\exp(-\Delta\tau H_x)\exp(-\Delta\tau H_z)]^n$. $\beta = 1/k_B T$ is the inverse temperature, and H_z (H_x) are the terms with σ_z operators (σ_x operators). This maps the partition function to a three-dimensional system with n lattice spacings along the imagi-

nary time direction, and Ising spin-1/2 denoted as S_j lives on the cubic lattice bonds. The statistical weight of a given $\{S_j\}$ configuration on each imaginary time slice and between adjacent slices is determined by H_z and H_x, respectively. After careful treatment of the gauge redundancy, the J_x, h_z, h_x and J_z terms in Eq. 4 are interpreted as the vertical bond, horizontal bond, vertical plaquette and horizontal plaquette terms in the anisotropic $\mathbb{Z}_2$ gauge Higgs model. For simplicity, we consider the isotropic case where

$$\beta H_{3D} = -\lambda_b \sum_j S_j - \lambda_p \sum_p \prod_{j\in p} S_j \qquad (5)$$

where dimensionless parameters λ_b and λ_p are related to the parameters ot the model Eq. (4) through

$$\lambda_b = h_z\Delta\tau = -\frac{1}{2}\ln\tanh J_x\Delta\tau$$

$$\lambda_p = J_z\Delta\tau = -\frac{1}{2}\ln\tanh h_x\Delta\tau \qquad (6)$$

in the limit of small $\Delta\tau$. We note that the method we will present also offers an alternative and convenient signature of de-confinement [55] in an equivalent lattice gauge theory, even though our focus here is mainly on the $\mathbb{Z}_2$ quantum spin liquids. For the rest of the paper, we use λ_b and λ_p as parameters, sticking to the notation of Ref. 15 for benchmark purposes.

The three-dimensional classical model in Eq. 5 offers a convenient way to measure operators within QLT, which are sampled within classical Monte Carlo Metropolis with both on-site and cluster updates:

$$\langle\hat{O}\rangle = \mathrm{tr}\left[\rho_{2D}\hat{O}\right] = \sum_{\alpha\gamma}\rho^\gamma_{\alpha\alpha}\langle\hat{O}\rangle_\alpha \qquad (7)$$

where $\rho_{2D} = \sum_{\alpha\beta\gamma}\rho^\gamma_{\alpha\beta}|\alpha\rangle\langle\beta|$ is the quantum density matrix of the original two-dimensional system in Eq. 4. α and β are two-dimensional spin-1/2 configurations. $\rho^\gamma_{\alpha\beta} = \exp\left(-E_{\alpha\beta\gamma}\right)/Z$ is the thermal statistical weight of the three-dimensional classical configuration with open boundary conditions α and β along the imaginary time direction and bulk configuration γ, whose energy $E_{\alpha\beta\gamma}$ is given by Eq. 5. Similarly, $\rho^\gamma_{\alpha\alpha}$ is the normalized and positive-definite weight with periodic boundary condition α, serving as the sample probability in our Monte Carlo metropolis. Then the quantity that contributes to the operator expectation value is

$$\langle\hat{O}\rangle_\alpha = \sum_\beta\langle\alpha|\hat{O}|\beta\rangle\cdot\rho^\gamma_{\alpha\beta}/\rho^\gamma_{\alpha\alpha} = \sum_\beta\langle\alpha|\hat{O}|\beta\rangle\cdot\exp\left(-\Delta E_{\alpha\to\beta}\right)$$

where $E_{\alpha\to\beta}$ is the energy difference.

Since the $\mathbb{Z}_2$ topological order has trivial quasi-particle self statistics, we focus on the mutual statistics and consider for QLT only operators with nontrivial knots between the spinon and vison world-lines. Namely, these are strings of σ^z and σ^x operators with double intersections. The example illustrated in Fig. 4b corresponds to the operator $\langle\sigma^x_j\sigma^z_j\sigma^x_j\sigma^z_j\rangle =$

tr$\left[\rho\sigma_j^x\sigma_j^z\sigma_{j'}^z\sigma_{j'}^x\right]$, where the vison and spinon trajectories intersect twice at r and r', before the density matrix ρ weighs the quasi-particle correlations between the initial and final positions and effectively closes the trajectories.

In the spirit of machine learning, instead of calculating their full expectation values, which are relatively expensive due to averaging over the Markov chain, we settle with an individual Monte Carlo sample α for QLT. To further simplify the QLT, we consider instead the operators $\prod_{j\in C_e}\sigma_j^z\prod_{k\in\tilde{C}_m}\sigma_k^x$ where the strings C_e and $\tilde{C}_m$ intersect. When valued at a particular Monte Carlo step, the original operators within QLT can be straightforwardly derived from the ensemble of such intersecting $\prod_{j\in C_e}\sigma_j^z\prod_{k\in\tilde{C}_m}\sigma_k^x$, making the latter an equally informative candidate for machine learning with QLT. Finally, string tension and the rapidly decaying spinon-spinon and vison-vison correlations allow us to focus on strings no longer than a cut-off, which we choose to set to $d_c = 2$ unless noted otherwise.

MACHINE LEARNING TOPOLOGICAL PHASE DIAGRAM

Now we describe the specific procedure for studying the phase diagram of the model in Eq. (5). Since the large λ_p and small λ_b limit and the small λ_p and λ_b limit are well established for the $\mathbb{Z}_2$ topological phase and the trivial phase, respectively, they offer the labeled samples needed for supervised training with target neural output $y = 1$ and $y = 0$, respectively. We illustrate two parallel machine learning approaches with merits of their own. In the efficiency-focused approach, we train a *universal* neural network with a diverse training set of multiple points within the known phase space, and then apply the single trained network to the entire (λ_p, λ_b) phase space. In the precision-focused approach, we train a separate neural network for each λ_b-slice with two sample points at large and small λ_p, and then apply the network to the intermediate phase space on the given λ_b-slice.

For each value of λ_p and λ_b chosen for training, we construct 10000 QLT data sets sampled through a classical Monte Carlo Metropolis procedure. We also reserve a separate test set of 20% of the size of the training set for validation purposes including learning speed control and termination [30]. As illustrated in Fig. 1, once obtained, each data set serves as an "image" for a fully-connected feed-forward neural network with one hidden layer consisting of $n = 20$ sigmoid neurons. Each neuron processes the input through independent weights and biases $w \cdot x + b$. After the hidden layer, the outcome is fed forward to be processed by the output neuron. The final output $0 \leq y \leq 1$ corresponds to the neural network's judgment of whether the QLT input state is topological. The neural network is trained via back-propagation and gradient descent to optimize the weights and biases. We use the cross entropy as the cost function, with L2 regularization to avoid over-fitting, and a mini-batch size of 10 [30].

Once the neural network is trained successfully, it can rapidly process QLT data obtained from different parts of the
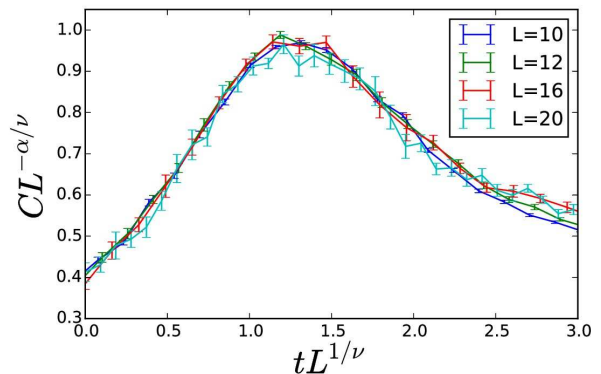

FIG. 5. Specific heat data of the three-dimensional system H_{3D} from Eq. 5 for a range of system sizes $L = 10, 12, 16, 20$ is collapsed according to the finite-size scaling using critical exponents of the three-dimensional Ising universality class. The best collapse is achieved at critical temperature $k_B T_c = 1.324(1)$ for $\lambda_p/\beta = 1$ and $\lambda_b/\beta = 0.29$.

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 inputs at each point to obtain statistics. Specifically, for each phase space point (λ_p, λ_b), we find the fraction $p(\lambda_p, \lambda_b)$ of test inputs with 'topological' output, i.e., $y > 0.5$. To further suppress the uncertainty due to sample fluctuations, we base our judgment upon the average neural output $\bar{y}$ over 5 QLT inputs from uncorrelated Monte Carlo samples: a 'topological' output corresponds to $\bar{y} > 0.5$; otherwise, the $\bar{y} < 0.5$ output is considered 'trivial'. See Supplemental Materials for details.

The extreme corners of the phase diagram in the space of (λ_p, λ_b) have been schematically understood for a long time[11]. First of all, the topological phase is expected in the region with large λ_p and small λ_b, with the $\lambda_p \to \infty$ and $\lambda_b \to 0$ limit equivalent to the zero-field toric code. Moreover, magnetically ordered phase is expected in the large λ_b limit, and the disordered phase is expected at the high temperature limit of small λ_p and λ_b. The two topological trivial phases are dual to each other with a self-dual line $\lambda_b = -0.5\ln\tanh\lambda_p$ as their phase boundary. Numerical pursuits of the phase diagram with increased system sizes and computational efforts[15–18] established the phase boundary between a disordered (magnetically ordered) phase and topological phase at around $\lambda_p \lesssim 0.76$ ($\lambda_b \gtrsim 0.223$) with the critical λ_p value from the topological phase toward the disordered phase displaying a very slight negative dependence upon increasing λ_b. The goal of our QLT-based machine learning approach is to reproduce these known results in a way that can be extended to other models whose phase diagrams remain an open question in the future.

In order to precisely benchmark the outcome, we also determine one point of the phase boundary with a traditional techniques of collapsing specific heat data[56]. This approach requires calculating specific heat at over a mesh of phase space points repeatedly for a range of system sizes and then finding
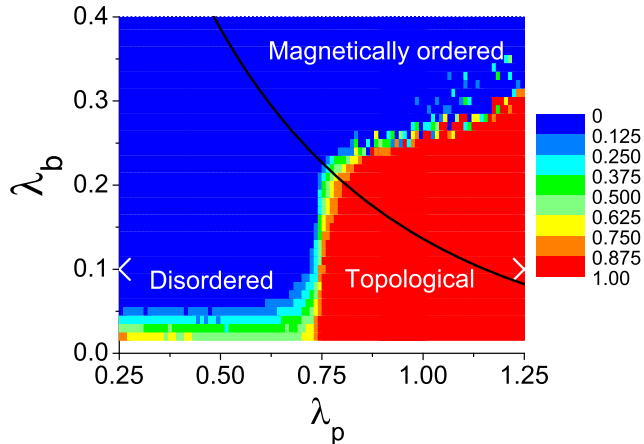
FIG. 6. Phase diagram in both λ_b and λ_p obtained with a single, universal neural network with supervised machine learning and quantum loop topography. The color scale indicates the fraction $p(\lambda_p, \lambda_b)$ of topological response: red (blue) color is the topological (trivial) phase. The white crosses denote the training sets. The solid black line labels the self-dual parameters $\lambda_b = -0.5 \ln \tanh \lambda_p$.

the phase space point that offers the best critical scaling collapse. Specifically, in order to locate the critical point intersected by the line $\lambda_b/\lambda_p = 0.29$ in the (λ_p, λ_b) phase diagram, we obtain the specific heat data of the classical system in Eq. 5 with $\lambda_p/\beta = 1$ and $\lambda_b/\beta = 0.29$ and a variable temperature $k_B T = \beta^{-1}$ for system sizes $L = 10, 12, 16, 20$ via conventional Monte Carlo calculations. We then attempt to collapse the data to the finite-size scaling function using critical exponents of the expected three-dimensional Ising universality class for each of different temperatures, looking for the best data collapse. Since we find the best data collapse at temperature $k_B T = 1.324$ (see Fig. 5) with a noticeable deviation nearby temperatures as close as $k_B T = 1.323$ and $k_B T = 1.325$, we conclude $k_B T_c = 1.324 \pm 0.001$. This locates one critical point in the phase diagram at $(\lambda_p, \lambda_b) = (0.7553 \pm 0.006, 0.2192)$, see Fig. 7. Accuracy of this magnitude comes with a cost, and the Monte Carlo computation involved in producing the large number of samples needed to reduce statistical errors involved about two core-years of CPU time in this example. Unfortunately, this level of accuracy is necessary when the varying quality of scaling collapse at different points in the phase space is used to single out the critical point. Moreover, a priori knowledge of the critical exponents is required in this approach. Although viable in principle, mapping out the entire two-parameter phase diagram to this accuracy, where this Monte Carlo process needs to be repeated over many other cuts, would quickly incur insurmountable costs[57].

RESULTS

First we attempt to obtain the full two-dimensional phase diagram with a single *universal* neural network trained with data sets from only two phase-space points: $(\lambda_p, \lambda_b) = (0.25, 0.1)$ for the non-topological phase and $(1.25, 0.1)$ for the topological phase. The resulting phase diagram shown in Fig. 6 reproduces the known phase diagram reasonably well, especially taking into account the rapid speed of the neural network scanning for the $\mathbb{Z}_2$ topological phase. Specifically, the entire process of training set preparation, neural network machine learning, and eventually scanning a phase space of 101×39 pairs of (λ_p, λ_b) parameters takes just about 4 hours of CPU time. The lack of diversity in the training set as a result of including data from only two points in phase space [58] affects the reliability of the neural network for small $\lambda_b < 0.05$, and also near the tri-critical point around $(\lambda_p, \lambda_b) \sim (0.75, 0.225)$. Nevertheless one remarkable outcome of this result (obtained from a single neural network trained with only one disordered phase-space point and one topological phase-space point) is the apparent understanding of the duality. Specifically, although the network was trained deep in the disordered phase as the only topologically trivial example, it correctly recognized the magnetically ordered phase also as a topologically trivial phase. Indeed, this suggests that the decision-vital information is not specific to spinon or vison alone, but between spinon and vison and invariant under the duality transformation, such as their mutual statistics.

To improve the critical region, we consider a slice-by-slice construction of the phase diagram where a series of neural networks are trained. The resulting phase diagram in Fig. 7(a) is obtained with 34 different neural networks. We train one neural network with quantum loop topography inputs from $\lambda_p = 1.25$ for the topological phase ($y = 1$) and $\lambda_p = 0.25$ for the disordered phase ($y = 0$), at a fixed value of $\lambda_b \leq 0.22$ below the magnetic ordering threshold. Then the phase space interpolating between $0.25 \leq \lambda_p \leq 1.25$ is scanned by that particular neural network. To analyze the extent of accuracy we achieve, 2000 QLT test inputs are assessed at each (λ_p, λ_b) for the fraction p of 'topological' response. Indeed, the results that $p \to 0$ deep in the trivial phase and $p \to 1$ deep in the topological phase, indicating even a single measurement with as few as five Monte Carlo samples can reliably provide a trustworthy detection. Notice that not only is the critical region relatively sharp, but the phase boundary between the disordered phase and the topological phase now reproduces the slightly negative slope seen in Ref. [15]. Here we only focused on refining the phase boundary between the disordered phase and the topological phase relying on duality to give us the other side of the phase boundary. However, to apply the QLT to unknown phase diagrams, one may consider multiple-neuron output layer to differentiate all phases under equal footing.

We now compare the neural network based location of

(a)

(b)

$\lambda_b = 0.2192$
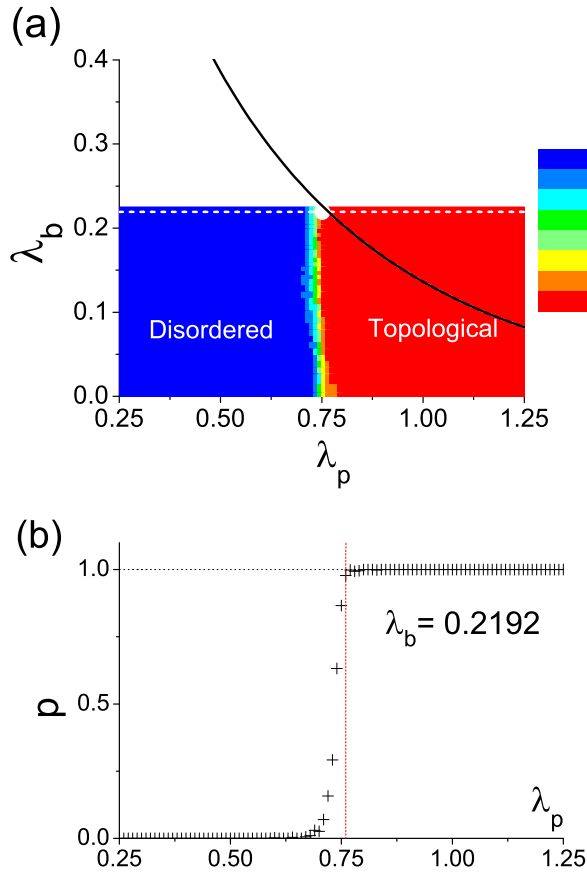
λ_p

(a)

(b)

$\lambda_b = 0.2192$

FIG. 7. (a): For each $\lambda_b \leq 0.22$, we implement supervised training with quantum loop topography inputs on the neural network with $\lambda_p = 1.25$ for the topological phase ($y = 1$) and $\lambda_p = 0.25$ for the trivial phase ($y = 0$). Then the resulting neural network is applied to the phase space in between and builds up the phase diagram slice by slice. The color scales indicates the fraction $p(\lambda_p, \lambda_b)$ of 'topological' output, from red for the most likely to blue for the least likely. The solid black line labels self-dual conditions. (b): one slice at $\lambda_b = 0.2192$ corresponding to the white dashed line in (a). The vertical red dashed line in (b) and white dot in (a) label $\lambda_p = 0.7553$, the critical value pinpointed using specific heat collapse method, see Fig. 5.

FIG. 8. (a) The neural output $\bar{y}(\lambda_p)$ over the range $0.25 \leq \lambda_p \leq 1.25$. (b) The second derivative $-d^2\bar{y}/d\lambda_p^2$ in the critical region $0.68 \leq \lambda_p \leq 0.78$ from higher-resolution data. $\bar{y}$ is averaged over 20000 Monte Carlo samples at $\lambda_b = 0.2192$ and different system sizes L. Inset: The ratio of topological response 'p' for different QLT cut-off $d_c = 1, 2, 3$ at $L = 12$. The red dashed lines mark the location of the transition $\lambda_p = 0.7553$ obtained using specific heat collapse, see Fig. 5.

phase boundary to the critical point on the line $\lambda_b/\lambda_p = 0.29$ obtained using specific heat collapse: $(\lambda_p, \lambda_b) = (0.7553 \pm 0.006, 0.2192)$ by considering a cut of the two-dimensional phase diagram along $\lambda_b = 0.2192$ shown in Fig. 7(b) [see the white dashed line in Fig. 7(a)]. The neural network shows a statistical ratio of 'topological' response $p < 0.005$ for $\lambda_p < 0.7$, deep in the trivial phase, and $p > 0.995$ for $\lambda_p > 0.8$, deep in the topological phase putting a bound on the critical point $0.7 < \lambda_p^c < 0.8$. This is clearly consistent with the specific heat collapse guided determination of the critical point, albeit with lower precision. However, we note that the total computational time for the cut shown in Fig. 7(b) was mere 18 minutes of CPU time. Hence, this comparison establishes how the QLT based machine learning can facilitate and speed-up the evaluation of phase diagram dramatically by quickly
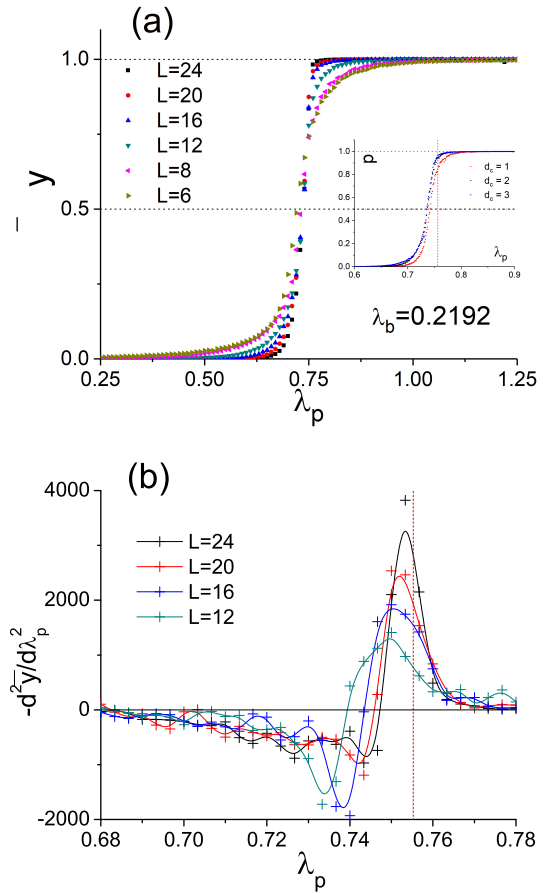
zooming into the narrow region containing the phase boundary.

Before closing, we now turn to the finite-size effects on the neural network output $\bar{y}(\lambda_p)$ averaged over 20000 Monte Carlo samples [see Fig. 8(a)]. Again for fixed $\lambda_b = 0.2192$ for comparison with the specific heat collapse, the training sets are taken from $\lambda_p = 0.25, 0.60$ for the trivial phase and $\lambda_p = 0.90, 1.25$ for the topological phase, respectively. The resulting neural network then tests the Monte Carlo samples generated between $0.25 \leq \lambda_p \leq 1.25$. We find the outputs to be reliably definitive ($y \to 0$ in the trivial phase and $y \to 1$ in the topological phase) away from the critical region even for the smallest system size $L = 6$. In addition, no obvious hysteresis is observed. These behaviors are in sharp contrast with the more global probes based upon large Wilson loops [18]. The fact that even rather small systems can be used to detect the topological phase away from the critical region is clearly advantageous. Noticeably, the transition gets sharper

with increasing system size in Fig. 8(a). To capture this trend in a more revealing manner we looked into the curvature of $\bar{y}$ as a function of λ_p. Specifically, we fit the higher-resolution data in the critical region to an analytic curve and take derivatives [59]. The data points of Fig. 8(b) are the results and the solid curves are guide to the eyes. This curvature plot clearly highlights how the peak in curvature is sharpening and moving towards the reference point obtained from the specific heat collapse, demonstrating that our machine learning based approach can do a respectable job near criticality, upon increasing system size. We also find a fast convergence in the QLT cut-off d_c, see Fig. 8 inset, and $d_c = 2$ is indeed adequate for our exemplary purposes.

SUMMARY AND DISCUSSIONS

We have proposed a quantum-loop-topography-based supervised machine learning strategy for detecting strongly-correlated topological phases, using quasi-particle statistics as defining features. The quantum loop topography (QLT) we developed here is a feature selection process [60] that effectively picks out Wilson loop operators on a semi-local scale – the key feature of a topological phase that is nevertheless minimally extended. Relying on the mutual statistics between the vison and spinon excitations, QLT enabled us to successfully train and test a simple (shallow, fully-connected feed forward) neural network to recognize a $\mathbb{Z}_2$ quantum spin liquid, map out its parameter region, and locate the topological quantum phase transitions of a microscopic interacting Hamiltonian.

Compared to existing approaches of studying the phase diagram of $\mathbb{Z}_2$ quantum spin liquids, the QLT-based machine learning we have implemented here has two clear advantages: (1) significant savings on computational resources as we have summarized in the last section, and (2) targeting small-scale Wilson loops. Indeed, the combination of our QLT and machine learning makes effective use of small-scale Wilson loops reflecting quasi-particle statistics. Traditionally, diagnosis of topological phases and de-confined gauge theories [55] has largely relied on long-range entanglement properties in order to avoid intractable short-distance fluctuations. However, we have demonstrated that one can still manage to analyse the information within semi-local operators with neural networks, which excel at being trained to recognize key features even in noisy data. Further, since QLT is semi-local, the approach does not depend on boundary conditions, and even small systems can yield reliable results on labelling the phase. Moreover, this selective nature of QLT preempts memory required for storing and processing data becoming a bottle-neck, which can be an issue with machine learning based approaches.

Going beyond the $\mathbb{Z}_2$ quantum spin liquids, the model of QLT featuring quasi-particle statistics that we have introduced here can be generalized to other non-trivial topological phases, including ones with non-Abelian excitations. Our development will be particularly beneficial for the study of higher-dimensional systems, such as the three-dimensional $\mathbb{Z}_2$

topological phases [6] and Fracton topological order [8, 9]. The fact that these phases also feature multiple types of quasi-particles with nontrivial statistics and fusion makes them an exciting future direction to study with our algorithm. Finally, we would like to point out that the QLT constructed here and the QLT we constructed in Ref. [39] are guided by distinct defining properties of the phases of interest: Hall conductivity for chiral topological phases and mutual statistics for the $\mathbb{Z}_2$ quantum spin liquid. Nevertheless, both QLT layers enabled successful evaluation of the phase diagrams using a shallow network with orders of magnitude speed up. From the broader perspective of machine learning, the idea of a QLT layer using particular operators to select relevant features as inputs to the neural network is an example of a feature selection layer [60]. Our success in Ref. [39] and again in this paper advocates QLT – and the idea of selecting defining features and relevant operators via responses, statistics, or beyond – as a general strategy for machine learning applied to quantum condensed matter problems. We also note important progress on representing quantum states using neural networks [61–63].

[*] frankzhangyi@gmail.com
[†] eun-ah.kim@cornell.edu

[1] A. Kitaev, Annals of Physics **303**, 2 (2003).
[2] A. Kitaev, Annals of Physics **321**, 2 (2006).
[3] J. Haah, Phys. Rev. A **83**, 042330 (2011).
[4] X. Chen, Z.-C. Gu, Z.-X. Liu, and X.-G. Wen, Science **338**, 1604 (2012).
[5] X.-G. Wen, Phys. Rev. B **95**, 205142 (2017).
[6] C. Chamon, Phys. Rev. Lett. **94**, 040402 (2005).
[7] H. Yao and S. A. Kivelson, Phys. Rev. Lett. **99**, 247203 (2007).
[8] S. Vijay, arXiv:1701.00762.
[9] S. Vijay, J. Haah, and L. Fu, Phys. Rev. B **94**, 235157 (2016).
[10] V. Karimipour, L. Memarzadeh, and P. Zarkeshian, Phys. Rev. A **87**, 032322 (2013).
[11] E. Fradkin and S. H. Shenker, Phys. Rev. D **19**, 3682 (1979).
[12] J. Vidal, S. Dusuel, and K. P. Schmidt, Phys. Rev. B **79**, 033109 (2009).
[13] R. Mohseninia, S. S. Jahromi, L. Memarzadeh, and V. Karim-

ipour, Phys. Rev. B **91**, 245110 (2015).

[14] S. Dusuel, M. Kamfor, R. Orús, K. P. Schmidt, and J. Vidal, Phys. Rev. Lett. **106**, 107203 (2011).

[15] I. S. Tupitsyn, A. Kitaev, N. V. Prokof'ev, and P. C. E. Stamp, Phys. Rev. B **82**, 085114 (2010).

[16] S. Trebst, P. Werner, M. Troyer, K. Shtengel, and C. Nayak, Phys. Rev. Lett. **98**, 070602 (2007).

[17] L. Genovese, F. Gliozzi, A. Rago, and C. Torrero, Nuclear Physics B - Proceedings Supplements **119**, 894 (2003).

[18] G. A. Jongeward, J. D. Stack, and C. Jayaprakash, Phys. Rev. D **21**, 3360 (1980).

[19] C. Castelnovo and C. Chamon, Phys. Rev. B **77**, 054433 (2008).

[20] F. Wu, Y. Deng, and N. Prokof'ev, Phys. Rev. B **85**, 195104 (2012).

[21] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, Rev. Mod. Phys. **80**, 1083 (2008).

[22] R. M. Lutchyn, J. D. Sau, and S. Das Sarma, Phys. Rev. Lett. **105**, 077001 (2010).

[23] Y. Oreg, G. Refael, and F. von Oppen, Phys. Rev. Lett. **105**, 177002 (2010).

[24] R. D. Sedgewick, D. J. Scalapino, and R. L. Sugar, Phys. Rev. B **65**, 054508 (2002).

[25] J. Yu, S.-P. Kou, and X.-G. Wen, EPL (Europhysics Letters) **84**, 17004 (2008).

[26] S. C. Morampudi, C. von Keyserlingk, and F. Pollmann, Phys. Rev. B **90**, 035117 (2014).

[27] A. Kitaev and J. Preskill, Phys. Rev. Lett. **96**, 110404 (2006).

[28] M. Levin and X.-G. Wen, Phys. Rev. Lett. **96**, 110405 (2006).

[29] K. Park and S. Sachdev, Phys. Rev. B **65**, 220405 (2002).

[30] Michael Nielsen, *Neural Networks and Deep Learning* (Free Online Book, 2013).

[31] L. Wang, Phys. Rev. B **94**, 195105 (2016).

[32] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nature Physics **13** (2017).

[33] W. Hu, R. R. P. Singh, and R. T. Scalettar, Phys. Rev. E **95**, 062122 (2017).

[34] J. Carrasquilla and R. G. Melko, Nature Physics **13**, 431 (2017).

[35] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Scientific Reports **7**, 8823 (2017).

[36] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Phys. Rev. X **7**, 031038 (2017).

[37] F. Schindler, N. Regnault, and T. Neupert, Phys. Rev. B **95**, 245134 (2017).

[38] T. Ohtsuki and T. Ohtsuki, Journal of the Physical Society of Japan **85**, 123706 (2016).

[39] Y. Zhang and E.-A. Kim, Phys. Rev. Lett. **118**, 216401 (2017).

[40] T. Ohtsuki and T. Ohtsuki, Journal of the Physical Society of Japan **86**, 044708 (2017), http://dx.doi.org/10.7566/JPSJ.86.044708.

[41] Y. Zhang, T. Grover, A. Turner, M. Oshikawa, and A. Vishwanath, Phys. Rev. B **85**, 235151 (2012).

[42] Y. Zhang, T. Grover, and A. Vishwanath, Phys. Rev. B **91**, 035127 (2015).

[43] H.-H. Tu, Y. Zhang, and X.-L. Qi, Phys. Rev. B **88**, 195412 (2013).

[44] Y. Zhang and X.-L. Qi, Phys. Rev. B **89**, 195144 (2014).

[45] J.-W. Mei and X.-G. Wen, Phys. Rev. B **91**, 125123 (2015).

[46] M. P. Zaletel, R. S. K. Mong, and F. Pollmann, Phys. Rev. Lett. **110**, 236801 (2013).

[47] L. Cincio and G. Vidal, Phys. Rev. Lett. **110**, 067208 (2013).

[48] J. Wildeboer, A. Seidel, and R. G. Melko, Phys. Rev. B **95**, 100402 (2017).

[49] E. Witten, Communications in Mathematical Physics **121**, 351 (1989).

[50] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, Rev. Mod. Phys. **80**, 1083 (2008).

[51] X. G. Wen and A. Zee, Phys. Rev. B **46**, 2290 (1992).

[52] N. Read, Phys. Rev. Lett. **65**, 1502 (1990).

[53] M. B. Hastings and X.-G. Wen, Phys. Rev. B **72**, 045141 (2005).

[54] S. Yan, D. A. Huse, and S. R. White, Science **332**, 1173 (2011).

[55] K. Gregor, D. A. Huse, R. Moessner, and S. L. Sondhi, New Journal of Physics **13**, 025009 (2011).

[56] M. E. Fisher and M. N. Barber, Phys. Rev. Lett. **28**, 1516 (1972).

[57] Note the use of Binder cumulant, which is often preferred over specific heat collapse for finding critical points for its efficiency, requires an order parameter and hence it is not an option for any topological order.

[58] One way to improve the neural network's performance is to increase the diversity of the training set and include more phase space.

[59] We do not find a discontinuity in the neural output which offered a sharp location of the critical point in Ref. [39]. This is possibly due to the fact that the topologically trivial phases are gapless in the present case.

[60] I. Guyon and A. Elisseeff, J. Mach. Learn. Res. **3**, 1157 (2003).

[61] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[62] D.-L. Deng, X. Li, and S. Das Sarma, arXiv:1609.09060.

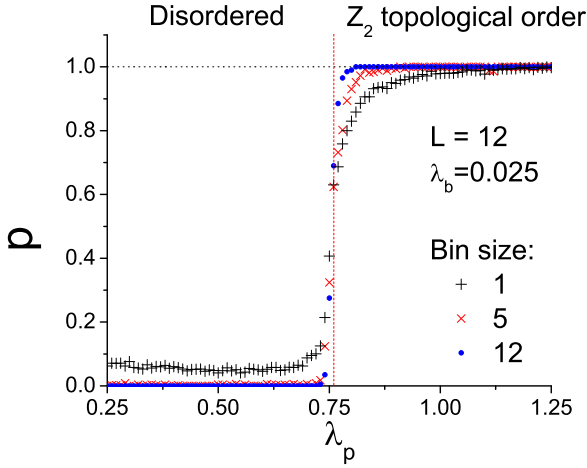[63] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X **7**, 021021 (2017).

FIG. 9. The ratio of 'topological' response as a function of λ_p at fixed $\lambda_b = 0.025$. The neural network is prepared with training set using $\lambda_p = 1.25$ for the topological phase ($y = 1$) and $\lambda_p = 0.25$ for the trivial (disordered) phase ($y = 0$), and then applied to the phase space in between. The decision is based upon the criteria that average neural output $\bar{y} > 0.5$ over a number of uncorrelated samples. The vertical red dashed line at $\lambda_p = 0.76$ is the approximate location of the transition.

APPENDIX: FLUCTUATION SUPPRESSION VIA SAMPLE BINNING

Since our quantum loop topography is built upon individual Monte Carlo samples and forgoes the Markov chain, the input into and thus the output y from the artificial neural network naturally have fluctuations irrespective of the implementation of the supervised machine learning, and may sometimes influence the decision's accuracy. For example, using the simple criteria $y > 0.5$ to determine whether a single output is 'topological', one can achieve accuracy of $> 99.5\%$ deep in the topological phase ($\lambda_p > 0.8$) as well as the trivial phase ($\lambda_p < 0.y$) at relatively large $\lambda_b \geq 0.1$; however, at small $\lambda_b \leq 0.05$, such accuracy decreases to $\sim 96\%$, see Fig. 9.

To suppress the impacts of the fluctuation, we instead judge 'topological' or 'trivial' based upon whether $\bar{y} > 0.5$, where the neural output is average over a small bin that consists of multiple Monte Carlo samples. Indeed, with bins as small as 5 samples, accuracy away from the critical region $> 99.5\%$ is obtained once again even at small λ_b, see Fig. 9.

Note that we have kept intact the algorithm of supervised machine learning and the architecture of the neural network and quantum loop topography. Instead, our binning only kicks in after we have obtained our original neural network's outputs, thus retaining full legitimacy of the discussion and analysis on the supervised machine learning and quantum loop topography.