Unsupervised learning of interacting topological and symmetry-breaking phase transitions

En-Jui Kuo [1,2,3,*] and Hossein Dehghani [1,2,*]

[1]*Department of Physics, University of Maryland, College Park, Maryland 20742, USA*

[2]*Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742, USA*

[3]*Joint Center for Quantum Information and Computer Science, University of Maryland, College Park, Maryland 20742, USA*

(Received 29 November 2021; revised 27 April 2022; accepted 8 June 2022; published 27 June 2022)

Symmetry-protected topological (SPT) phases are short-range entangled phases of matter with a nonlocal order parameter which are preserved under a local symmetry group. Here, by using an unsupervised learning algorithm, namely, diffusion maps, we demonstrate that we can detect phase transitions between symmetry-broken and topologically ordered phases, and between nontrivial topological phases in different classes. In particular, we show that the phase transitions associated with these phases can be detected in different bosonic and fermionic models in one dimension. This includes the interacting Su-Schrieffer-Heeger model, the Affleck-Kennedy-Lieb-Tasaki model and its variants, and weakly interacting fermionic models. Our approach serves as an inexpensive computational method for detecting topological phase transitions associated with SPT systems which can also be applied to experimental data obtained from quantum simulators.

DOI: 10.1103/PhysRevB.105.235136

I. INTRODUCTION

In the last two decades, theoretical prediction and experimental observation of topological phases of matter [1] have been one of the most remarkable advancements in the field of condensed matter physics. In topological quantum phases of matter, although the energy spectrum is gapped, a Landau-Ginzburg framework based on the existence of local order parameters cannot explain some of the most important features of these systems such as the existence of symmetry-protected edge states [2]. While initially most efforts were mostly focused on noninteracting topological insulators [3–9], in recent years, attention has turned to interacting topological phases of matter.

Classifying topological insulators and superconductors which are described by noninteracting fermionic Hamiltonians is less challenging than characterizing interacting topological phases of matter. In particular, since the spectrum of noninteracting models are often exactly solvable, classifying them can be done rigorously via several methods including the classification of random matrices [10], of Dirac operators [11], or, more abstractly, via K theory [12]. Nevertheless, for interacting topological systems, more refined methods such as group cohomology methods have been proposed [13,14]. Correspondingly, for the former systems, using the wave functions of the Hamiltonian or the Green's functions in the momentum space, one can define their corresponding topological invariants, while for the latter, one needs to define nonlocal order parameters in the real space [15–18], which is both computationally and experimentally costly to probe [19,20]. Therefore identification of topological properties in a given set of interacting topological systems is computationally more involved, and designing efficient computational methods to detect possible topological phase transitions is an indispensable task for studying these phases.

On the other hand, over the past few years, machine learning (ML) has emerged as a powerful tool to assist physicists to study a plethora of different problems in condensed matter and quantum sciences. A nonexhaustive list of notable examples includes classifying phases of matter [21–24], studying nonequilibrium dynamics of physical systems [25–27], simulating dynamics of quantum systems [28–30], and augmenting capabilities of quantum devices [31,32]. In particular, in classifying applications, most such techniques rely on supervised ML techniques where the ML algorithm after being trained with labeled systems learns to classify systems with new parameters [21,27,33–38]. However, more interestingly, in unsupervised machine learning no prior knowledge of the phase of the systems is provided, and the algorithm by detecting the hidden structure of the input data learns to cluster wave functions or Hamiltonians in different phases [22,39,40].

II. GOAL

In this paper, we use an unsupervised learning model (ULM), namely, the diffusion maps algorithm (DMA), to detect different topological phase transitions associated with symmetry-protected interacting topological phases with short-range entanglement also known as symmetry-protected topological (SPT) systems [13]. However, we note that our method can only detect the presence of topological phase transitions and cannot distinguish between different topological phases. The diffusion maps algorithm, which can capture the hidden geometrical and topological structure of data sets [41,42], has recently been used to classify phases of matter in the local order of the Ising model [43], thermal topological vortex structures and temperature-driven phase

*These authors contributed equally to this work.

transitions [44,45], and noninteracting topological insulators [46–48].

Here, by applying this algorithm to different bosonic and fermionic systems in one dimension which host a SPT phase, we demonstrate that this technique can approximately identify topological phase transitions and reproduce the phase diagram of models which in addition to symmetry-broken phases can host one or several interacting SPT phases. Hence, in systems where, based on physical symmetry arguments, it is known that there are finite possibilities for the formation of symmetry-broken phases, detection of additional phases via our method is evidence that the system can host nontrivial SPT phases. We should highlight that since we only need snapshots of the ground-state wave functions of the system, unlike other methods which, to detect topological phases, rely on nonlocal (stringlike) operators [15,49–52] or nonlinear functions of the wave functions such as entanglement entropy and the entanglement spectrum of the wave function [53–55], our approach is computationally more convenient. We also note that since we use ground states directly, this approach is also suitable for experimental data obtained from quantum simulators [56] and noisy intermediate-scale quantum (NISQ) devices [57].

III. DIFFUSION MAPS

The DMA is based on the classical idea of integration that the global structure of a manifold can be determined by traversing it provided that local transition rules to move from one point to another are at hand. To do so, we create a Markov matrix of transitions for the points in the input data set, which guides a random walker to traverse the data set through a diffusion process, which results in gaining information about the global geometrical properties of the input data. However, in order to probe the structure of the input data at different scales, a family of transition matrices are employed, and hence the name diffusion "maps."

More concretely, suppose we have n input data, each represented by an N-dimensional vector $\vec{x}_\alpha = (x_\alpha^1, \ldots, x_\alpha^N)$ in $\mathbb{R}^N$, where Greek letters denote the sample indices. For every two sampled vectors $\vec{x}_\alpha, \vec{x}_\beta$, we use the Euclidean metric to define their distance. To set the local rules of transition between two points, we start by defining a symmetric positivity-preserving function which for most applications could be a Gaussian kernel:

$$k(\vec{x}_\alpha, \vec{x}_\beta) = \exp\left(-\frac{|\vec{x}_\alpha - \vec{x}_\beta|^2}{\epsilon}\right), \tag{1}$$

where ϵ is the scaling hyperparameter of the DMA which determines the speed of the diffusion process of the corresponding random walker and correspondingly the number of clusters. We also define the diffusion matrix K, which is a version of the graph Laplacian matrix with components $K_{\alpha\beta} = k(\vec{x}_\alpha, \vec{x}_\beta)$. To interpret this matrix as a probability distribution, we need to normalize it by the diagonal matrix D whose components are given by $D_{\alpha\alpha} = d_\alpha$, where $d_\alpha = \sum_\beta K_{\alpha\beta}$ and plays the role of the local degree of the graph. Then, the normalized Laplacian matrix is defined by

$$M = D^{-1}K. \tag{2}$$

Physically, the matrix M denotes the transition probabilities between different samples such that the probability of transition from sample α to sample β in t time steps is given by the tth power of this matrix, i.e., $M^t_{\alpha\beta}$. We also notice that M is a stochastic matrix [58] with positive entries where each column sums to 1 and hence the largest eigenvalue is 1. We should also notice that there are other choices for the distance function. One can change the L_2 norm to an L_p norm for $p \geqslant 3$. As $p \to \infty$, we get the so-called Chebyshev's distance $d(\vec{x}_\alpha, \vec{x}_\beta) = \max_i |x_{\alpha i} - x_{\beta i}|$. One finds that in some cases, Chebyshev's distance may give sharper boundary conditions [47]. We have tried other distance functions, and our results for these distance functions are presented in the Supplemental Material [59].

Next, we compute the diagonal representation of M. As in principal component analysis (PCA), where we only keep the largest eigenvalues to determine the number of clusters [60], here we also keep the largest eigenvalues in the diagonal representation of M. We note that due to the spectral decay of the eigenvalues only a finite number of terms are necessary to achieve a given relative accuracy in reproducing M via its diagonal representation. Therefore, to reach an accuracy labeled by δ, we only need to keep the eigenvalues larger than $1 - \delta$ in the spectrum. Since δ is arbitrary, we can fix it and then study how the number of relevant clusters varies with the scale parameter ϵ to probe the underlying geometric structure at different scales. We note that the main advantage of this method over more conventional ULMs such as PCA is that while maintaining the local structures of the data sets, it can also recognize their underlying nonlinear manifolds due to its nonlinear kernel, which makes it advantageous compared with other unsupervised learning methods [59].

IV. MODEL INTRODUCTION AND RESULTS

Let us now apply the DMA explained above to our quantum many-body problem. We imagine we are given a Hamiltonian with a set of unknown parameters $\boldsymbol{\theta} \in \mathbb{R}^s$. We assume that we have direct access to the ground state of an N-particle many-body Hamiltonian and we can sample from its many-body wave function in the configuration space. In particular, in this paper, for spin-σ chains our sampled wave function determines the configuration of the spins in the real space, and therefore the components of the αth input vector are $x_\alpha^i = \{\pm\sigma\}$. For our fermionic and hard-core bosonic Hamiltonians, the measured wave function is represented in the Fock space, and therefore $x_\alpha^i = \{0, 1\}$. In this setting, the schematic flowchart of our classification approach is illustrated in Fig. 1, where $\psi_{\text{ground}}(\boldsymbol{\theta})$ denotes the ground-state wave functions. In general, this state is a superposition of different product states. In the next step, the sampling is performed in the space of these product states resulting in the vectors $\vec{x}_\alpha(\boldsymbol{\theta})$ by which we can produce the corresponding Laplacian matrix $M(\boldsymbol{\theta})$. Since across quantum phase transitions usually a change in the relevant degrees of freedom that describe the system occurs, we expect that for a given accuracy δ, once the scale ϵ is chosen properly, the number of clusters obtained from $M(\boldsymbol{\theta})$ should reveal the underlying phase diagram of the system.

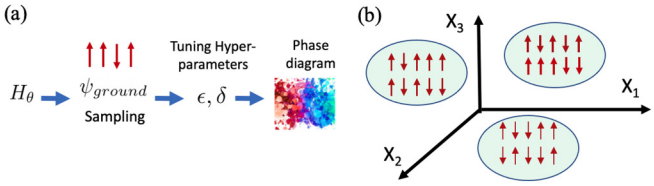

FIG. 1. Schematic illustration of the machine-learning algorithms to identify different phases. (a) Flowchart of the diffusion maps algorithm. From a given Hamiltonian, one (1) obtains the ground state and projectively samples from the ground state, (2) forms the diffusion matrix and tunes the hyperparameters, and (3) obtains the phase diagram by finding the conspicuous transitions in the number of eigenvalues. (b) Different ground states in a cluster can transform into each other via a diffusion process in the data set basis.

To determine the proper regime for ϵ as the parameter that describes the dynamics of the diffusion process [42], we use the approximate analysis of the spectrum of M proposed in Refs. [43,44]. Assuming clusters of samples with r particles in different configuration states, the diffusion matrix K will acquire a block-diagonal form with each block corresponding to a single such cluster. In this representation, diagonal elements are close to the identity, while off-diagonal components are bounded by $O(e^{-r/\epsilon})$. Denoting the number of such clusters with N_c, after diagonalizing K, we will approximately have N_c eigenvalues larger than $1 - e^{-r/\epsilon}$, and therefore the accuracy can be approximated by $\delta \sim e^{-r/\epsilon}$. Thus by choosing $\epsilon \sim -r/\ln\delta$, we can find the number of clusters of a particular size r. Consequently, since the number of particles in the clusters can change between 1 and N, we need to span the corresponding range of ϵ, to obtain the best value of ϵ empirically where changes in N_c show a conspicuous change for different values of $\boldsymbol{\theta}$. Therefore sharp changes in N_c as a function of $\boldsymbol{\theta}$ can be presumed to trace the topological phase transitions.

In what follows we employ the procedure above and consider some of the most well-known models which describe SPT phases, and we demonstrate that we can reproduce their phase diagram with an acceptable precision. We summarize our models in Table I, where for all models we have $\delta = 0.000\,01$), and we use periodic boundary conditions. For all of these models we generate the many-body ground state of the system in the Fock space and then sample from the states whose superposition creates the ground state of the system. We collect 500 samples for each set of fixed parameters in the Hamiltonian.

Model 1: Nontrivial SPT vs symmetry-protected trivial phases. In this section, we ask whether we can distinguish between ground states with a trivial symmetry-protected phase and those with a nontrivial SPT phase [56]. As one of the simplest models, we start with the interacting version of the Su-Schrieffer-Heeger (SSH) model, which has been recently realized via Rydberg atoms and describes a topological phase transition between a SPT nontrivial phase and a symmetry-protected trivial phase. This model can be described by hard-core bosons $b_i^2 = 0 = b_i^{\dagger 2}$. We consider two alternating coupling constants denoted by J and J' between adjacent sites $(n, n+1)$, where the former couples sites with odd n and the latter couples sites with even n,

$$H_{J,J'} = \sum_{ij} J_{ij}(b_i b_j^\dagger + b_j b_i^\dagger). \tag{3}$$

As shown in Ref. [56] this model is mapped to the fermionic SSH model, whose energy spectrum consists of two bands separated by a spectral gap $2(|J| - |J'|)$. Most importantly, this model for $|J'| > |J|$ is a topologically trivial phase, while for $|J'| < |J|$ it becomes a SPT phase which with open boundary conditions hosts delocalized zero-energy edge modes. In this case, the symmetry group $G = \mathbb{Z}_2 \times \mathbb{Z}_2$ based on group cohomology classification only has one nontrivial topological phase since $H^2[\mathbb{Z}_2 \times \mathbb{Z}_2; U(1)] = \mathbb{Z}_2$. Using Jordan-Wigner transformation, the original representation of this model transforms into a spin chain Hamiltonian [56]:

$$H = -J' \sum_i \sigma_{2i-1}^- \sigma_{2i}^+ - J \sum_i \sigma_{2i}^- \sigma_{2i+1}^+ + \text{H.c.} \tag{4}$$

We use the natural spin basis as our sampling result. We can see in Fig. 2(a) that our method reproduces the phase diagram obtained in Ref. [56].

Model 2: Affleck-Kennedy-Lieb-Tasaki model. In this section, compared with model 1, as a more challenging situation, we ask whether our approach can detect SPT phases with more involved symmetries where symmetry-broken phases may exist, too. The Haldane phase of $S = 1$ antiferromagnetic spin chains is a well-known example of such SPT phases with a $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry group whose parent Hamiltonian can host several symmetry-broken phases. An example of a state in the Haldane phase which can be written in a closed form is the Affleck-Kennedy-Lieb-Tasaki (AKLT) state [61]. This state is a special case of the ground state of the following bilinear-biquadratic spin-1 Hamiltonian ($\vec{S}_i$ are spin-1 operators) when $\theta = 1/3$:

$$H_{\text{bb}} = \sum_i \cos(\theta_{\text{bb}})(\vec{S}_i \cdot \vec{S}_{i+1}) + \sin(\theta_{\text{bb}})(\vec{S}_i \cdot \vec{S}_{i+1})^2. \tag{5}$$

TABLE I. Summary of one-dimensional Hamiltonians. Phase types: Haldane phase (H), dimerized phase, dimer phase, ferromagnetic phase (FM), antiferromagnetic phase, XY phase, valence bound state, $\mathbb{Z}_2$ symmetry-breaking phase ($\mathbb{Z}_2$ SB), and critical phase.

Hamiltonian H	Parameters	Type of phase	ϵ	N
$-J'\sum_i \sigma_{2i-1}^-\sigma_{2i}^+ - J\sum_i \sigma_{2i}^-\sigma_{2i+1}^+ + \text{H.c.}$	J, J'	Topological, trivial	0.088	9
$\sum_i \cos(\theta)(\vec{S}_i \cdot \vec{S}_{i+1}) + \sin(\theta)(\vec{S}_i \cdot \vec{S}_{i+1})^2$	θ	H, dimer, FM	$\frac{1}{32}$	9
$\sum_i(\vec{S}_i \cdot \vec{S}_{i+1}) + B\sum_i S_x^i + \sum_i D(S_z^i)^2$	B, D	Spin polarized, $\mathbb{Z}_2$ SB, H	$\frac{1}{32}$	10
$H = -\sum_{<ij>\sigma} c_{i\sigma}^\dagger c_{j\sigma} - 2\Delta_s\sum_j c_{j\uparrow}^\dagger c_{j\downarrow}^\dagger \pm i\Delta_p/2\sum_j(c_{j+1\uparrow}^\dagger c_{j\downarrow}^\dagger + c_{j+1\downarrow}c_{j\uparrow}) + \text{H.c.}$	Δ_s, Δ_p	$N = 0, \pm 1$	$\frac{1}{32}$	6

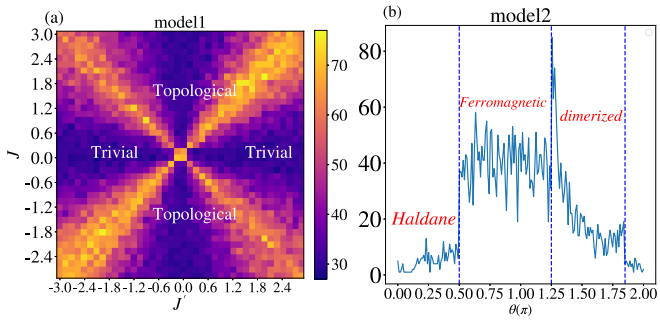

FIG. 2. Phase diagrams of models 1 and 2. The color bar in (a) and vertical axis in (b) represent the number of eigenvalues larger than $1-\delta$ (see text). (a) Model 1 is an interacting bosonic SPT model with a topological phase and a trivial phase with phase boundaries at $J=\pm J'$. (b) Model 2 is biquadratic spin chain model whose ground state at $\theta=1/3$ is the AKLT state. The dashed lines illustrate the phase boundaries obtained via density matrix renormalization group (DMRG) methods.

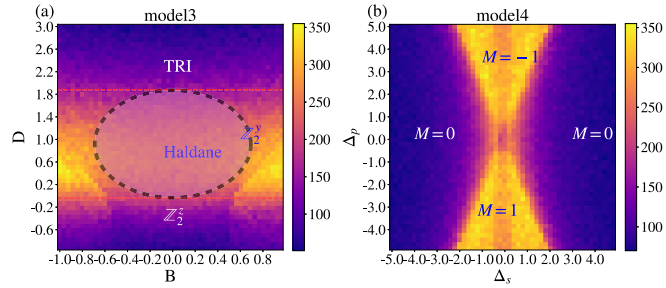

FIG. 3. Phase diagrams of models 3 and 4. (a) Model 3 is a variant of the AKLT model which hosts one trivial phase (TRI), two symmetry-breaking phases ($\mathbb{Z}_2^y$ and $\mathbb{Z}_2^x$), and one topologically ordered phase (Haldane). The topological classification of the system is described by the $\mathbb{Z}_2$ group. The dashed line shows the approximate boundaries of the Haldane phase obtained via exact methods. (b) Model 4 is a weakly interacting fermionic model with unconventional pairing. C labels the states via their distinct zero modes in different phases and is obtained via exact diagonalization. Color bars represent the number of eigenvalues larger than $1-\delta$ (see text).

However, by varying θ this model describes two topological and nontopological phase transitions.

We can apply the DMA to this model with $N=9$ and plot the resulting phase diagram [Fig. 2(b)], which is akin to the results in Ref. [62]. Here, we again see another successful detection of an SPT phase where two additional symmetry-broken phases may compete with each other.

Model 3: Variant of the AKLT model. As another more complicated model, we study a variant of the AKLT model where in addition to symmetry-broken and topological phases we also have a time-reversal-invariant (TRI) phase with the same symmetries as the Haldane phase [63],

$$H=\sum_i(\vec{S}_i\cdot\vec{S}_{i+1})+B\sum_i S_x^i+\sum_i D\big(S_z^i\big)^2. \tag{6}$$

This Hamiltonian is invariant under translational symmetry $\mathbb{Z}_2$, time-reversal symmetry $S_y\to -S_y$, and $\mathbb{Z}_2$ parity symmetry. The complete description of the phase diagram is found in Ref. [63]; it contains a TRI phase with $S^z=0$, two symmetry-broken phases $\mathbb{Z}_2^{x,y}$, and the Haldane phase. Our results for this model are depicted in Fig. 3(a). In this figure, by comparing our phase diagram with those obtained in Ref. [63], we can see a noticeable transition between the $\mathbb{Z}_2^x$ phase, the TRI phase, and the Haldane phase. However, the distinction between the $\mathbb{Z}_2^y$ phase and the Haldane phase is less pronounced, and we only see a partial change of color between these two regions. *Model 4.* In models 1–3, based on group cohomology arguments the topological classification of the phases is described by the $\mathbb{Z}_2$ group, and there is only one topologically nontrivial phase. In model 4, in addition to spin-1 chains, we consider a *weakly* interacting fermionic spin-1/2 Hamiltonian which can host more than one nontrivial SPT phase whose topological classification is described by the $\mathbb{Z}_2\times\mathbb{Z}_2$ group. This model respects both time-reversal symmetry and S_z spin-rotation symmetries. Hence, in the absence of interactions, based on the tenfold classification, its topological phases are described by the $\mathbb{Z}$ group. Nevertheless, by adding interactions to this Hamiltonian, the resulting topological classes of this system will reduce to the $\mathbb{Z}_2\times\mathbb{Z}_2$ group. The Bogoliubov–de Gennes (BdG) form of this Hamiltonian is given by [64]

$$H=-\sum_{<ij>\sigma}c_{i\sigma}^\dagger c_{j\sigma}-2\Delta_s\sum_j c_{j\uparrow}^\dagger c_{j\downarrow}^\dagger \pm i\Delta_p/2\sum_j(c_{j+1\uparrow}^\dagger c_{j\downarrow}^\dagger+c_{j+1\downarrow}c_{j\uparrow})+\text{H.c.}, \tag{7}$$

where the last term is an unconventional pairing between electrons on adjacent sites. This model can host three phases labeled by $M=0,+1,-1$, which represents the states with distinct zero modes [1]. We simulate this model with $N=6$ and plot the resulting phase diagram in Fig. 3(b).

As we can see in this figure, our approach can regenerate the phase diagram of this model as obtained in Ref. [64] with a high accuracy. To be more specific, we see sharp transitions between the topologically nontrivial phases with $M=\pm1$ and trivial phases labeled by $M=0$. While in principle, from this plot we cannot distinguish the $M=\pm1$ phases from each other, from the fact that we see a conspicuous change in the number of eigenvalues at their intersection at $\Delta_{s,p}\simeq 0$ we may speculate that the two regions correspond to distinct phases. This indicates that DMA not only can differentiate between topologically trivial phases and topologically nontrivial phases, but also can provide evidence for differentiating between different topological phases with similar properties.

V. DISCUSSION AND OUTLOOK

In this paper, we have used an unsupervised learning algorithm called diffusion maps to detect topological phase transitions between different symmetry-broken phases and

[1] As demonstrated in [64] by stacking two chains with $C=+1$, we get another nontrivial phase with two zero modes. In the presence of strong interactions the phase with $C=-1$ and $C=3$ are identified and hence the classification is $\mathbb{Z}_4$.

different SPT phases. Our results demonstrate that using diffusion maps as a computationally inexpensive method can provide helpful evidence for the presence of SPT phase transitions and SPT phases. Since previously this algorithm has been used for detecting other symmetry-broken phase transitions in one dimension [43], and due to the lack of intrinsic topological order in one dimension, our results demonstrate that the DMA can detect all different types of equilibrium phase transitions in one dimension.

From a conceptual viewpoint, the reason behind the efficacy of the DMA in detecting SPT phases is thought-provoking. One possible explanation is that the DMA in the continuum limit approximates the differential heat equation. On the other hand, the spectrum of elliptic differential operators, such as the heat kernel appearing in the heat equation, can include important information about topological invariants [65,66]. Based on our results, we speculate that these conceptual relations between the spectrum of differential operators and SPT invariants have practical application for computational and experimental purposes.

An immediate question to be pursued in future is the detection of SPT phases in higher dimensions via the DMA. Another intriguing question to investigate is the detection of long-range topological order in symmetry-enriched topological phases [67] which can be realized in two and higher dimensions without using thermal sampling [46]. Also, detection of topological phases in nonequilibrium phases of matter, especially Floquet systems, is left for future studies.

ACKNOWLEDGMENTS

We are grateful to Mohammad Hafezi, Victor Albert, and Alireza Seif for helpful discussions. We acknowledge funding from Grants No. MURI-AFOSR FA95501610323 and No. MURI-ARO W911NF-15-1-0397.

[1] X.-G. Wen, Colloquium: Zoo of quantum-topological phases of matter, Rev. Mod. Phys. **89**, 041004 (2017).
[2] X.-G. Wen and Q. Niu, Ground-state degeneracy of the fractional quantum Hall states in the presence of a random potential and on high-genus Riemann surfaces, Phys. Rev. B **41**, 9377 (1990).
[3] C. L. Kane and E. J. Mele, Z_2 Topological Order and the Quantum Spin Hall Effect, Phys. Rev. Lett. **95**, 146802 (2005).
[4] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Quantum spin Hall effect and topological phase transition in HgTe quantum wells, Science **314**, 1757 (2006).
[5] J. E. Moore and L. Balents, Topological invariants of time-reversal-invariant band structures, Phys. Rev. B **75**, 121306(R) (2007).
[6] L. Fu, C. L. Kane, and E. J. Mele, Topological Insulators in Three Dimensions, Phys. Rev. Lett. **98**, 106803 (2007).
[7] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Quantum spin Hall insulator state in HgTe quantum wells, Science **318**, 766 (2007).
[8] D. Hsieh, D. Qian, L. Wray, Y. Xia, Y. S. Hor, R. J. Cava, and M. Z. Hasan, A topological Dirac insulator in a quantum spin Hall phase, Nature (London) **452**, 970 (2008).
[9] M. Z. Hasan and J. E. Moore, Three-dimensional topological insulators, Annu. Rev. Condens. Matter Phys. **2**, 55 (2011).
[10] A. Altland and M. R. Zirnbauer, Nonstandard symmetry classes in mesoscopic normal-superconducting hybrid structures, Phys. Rev. B **55**, 1142 (1997).
[11] A. P. Schnyder, S. Ryu, A. Furusaki, and A. W. W. Ludwig, Classification of topological insulators and superconductors in three spatial dimensions, Phys. Rev. B **78**, 195125 (2008).
[12] A. Kitaev, Periodic table for topological insulators and superconductors, in *Advances in Theoretical Physics: Landau Memorial Conference, AIP Conference Proceedings* Vol. 1134 (American Institute of Physics, Melville, NY, 2009), pp. 22–30.
[13] X. Chen, Z.-C. Gu, Z.-X. Liu, and X.-G. Wen, Symmetry-protected topological orders in interacting bosonic systems, Science **338**, 1604 (2012).
[14] M. Barkeshli, P. Bonderson, M. Cheng, and Z. Wang, Symmetry fractionalization, defects, and gauging of topological phases, Phys. Rev. B **100**, 115147 (2019).
[15] F. Pollmann and A. M. Turner, Detection of symmetry-protected topological phases in one dimension, Phys. Rev. B **86**, 125441 (2012).
[16] J. Haegeman, D. Pérez-García, I. Cirac, and N. Schuch, Order Parameter for Symmetry-Protected Phases in One Dimension, Phys. Rev. Lett. **109**, 050402 (2012).
[17] K. Shiozaki and S. Ryu, Matrix product states and equivariant topological field theories for bosonic symmetry-protected topological phases in (1+1) dimensions, J. High Energy Phys. 04 (2017) 100.
[18] H. Dehghani, Z.-P. Cian, M. Hafezi, and M. Barkeshli, Extraction of the many-body Chern number from a single wave function, Phys. Rev. B **103**, 075102 (2021).
[19] A. Elben, J. Yu, G. Zhu, M. Hafezi, F. Pollmann, P. Zoller, and B. Vermersch, Many-body topological invariants from randomized measurements in synthetic quantum matter, Sci. Adv. **6**, eaaz3666 (2020).
[20] Z.-P. Cian, H. Dehghani, A. Elben, B. Vermersch, G. Zhu, M. Barkeshli, P. Zoller, and M. Hafezi, Many-Body Chern Number from Statistical Correlations of Randomized Measurements, Phys. Rev. Lett. **126**, 050501 (2021).
[21] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, Nat. Phys. **13**, 431 (2017).
[22] S. J. Wetzel, Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders, Phys. Rev. E **96**, 022140 (2017).
[23] E. P. L. Van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Learning phase transitions by confusion, Nat. Phys. **13**, 435 (2017).
[24] E.-J. Kuo, A. Seif, R. Lundgren, S. Whitsitt, and M. Hafezi, Decoding conformal field theories: From supervised to unsupervised learning, arXiv:2106.13485.
[25] E. van Nieuwenburg, E. Bairey, and G. Refael, Learning phase transitions from dynamics, Phys. Rev. B **98**, 060301(R) (2018).
[26] F. Schindler, N. Regnault, and T. Neupert, Probing many-body localization with neural networks, Phys. Rev. B **95**, 245134 (2017).
[27] A. Seif, M. Hafezi, and C. Jarzynski, Machine learning the thermodynamic arrow of time, Nat. Phys. **17**, 105 (2021).

[28] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).

[29] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Neural-network quantum state tomography, Nat. Phys. **14**, 447 (2018).

[30] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, Reconstructing quantum states with generative models, Nat. Mach. Intell. **1**, 155 (2019).

[31] A. Seif, K. A. Landsman, N. M. Linke, C. Figgatt, C Monroe, and M. Hafezi, Machine learning assisted readout of trapped-ion qubits, J. Phys. B: At., Mol. Opt. Phys. **51**, 174006 (2018).

[32] G. Torlai, B. Timar, E. P. L. van Nieuwenburg, H. Levine, A. Omran, A. Keesling, H. Bernien, M. Greiner, V. Vuletić, M. D. Lukin, R. G. Melko, and M. Endres, Integrating Neural Networks with a Quantum Simulator for State Reconstruction, Phys. Rev. Lett. **123**, 230504 (2019).

[33] D.-L. Deng, X. Li, and S. Das Sarma, Machine learning topological states, Phys. Rev. B **96**, 195145 (2017).

[34] J. Venderley, V. Khemani, and E.-A. Kim, Machine Learning Out-of-Equilibrium Phases of Matter, Phys. Rev. Lett. **120**, 257204 (2018).

[35] P. Zhang, H. Shen, and H. Zhai, Machine Learning Topological Invariants with Neural Networks, Phys. Rev. Lett. **120**, 066401 (2018).

[36] Q. Ni, M. Tang, Y. Liu, and Y.-C. Lai, Machine learning dynamical phase transitions in complex networks, Phys. Rev. E **100**, 052312 (2019).

[37] B. S. Rem, N. Käming, M. Tarnowski, L. Asteria, N. Fläschner, C. Becker, K. Sengstock, and C. Weitenberg, Identifying quantum phase transitions using artificial neural networks on experimental data, Nat. Phys. **15**, 917 (2019).

[38] H.-Y. Huang, R. Kueng, and J. Preskill, Information-Theoretic Bounds on Quantum Advantage in Machine Learning, Phys. Rev. Lett. **126**, 190505 (2021).

[39] L. Wang, Discovering phase transitions with unsupervised learning, Phys. Rev. B **94**, 195105 (2016).

[40] W. Hu, R. R. P. Singh, and R. T. Scalettar, Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination, Phys. Rev. E **95**, 062122 (2017).

[41] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker, Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps, Proc. Natl. Acad. Sci. USA **102**, 7426 (2005).

[42] R. R. Coifman and S. Lafon, Diffusion maps, Appl. Comput. Harmonic Anal. **21**, 5 (2006).

[43] A. Lidiak and Z. Gong, Unsupervised Machine Learning of Quantum Phase Transitions Using Diffusion Maps, Phys. Rev. Lett. **125**, 225701 (2020).

[44] J. F. Rodriguez-Nieva and M. S. Scheurer, Identifying topological order through unsupervised machine learning, Nat. Phys. **15**, 790 (2019).

[45] J. Wang, W. Zhang, T. Hua, and T.-C. Wei, Unsupervised learning of topological phase transitions using the Calinski-Harabaz index, Phys. Rev. Research **3**, 013074 (2021).

[46] M. S. Scheurer and R.-J. Slager, Unsupervised Machine Learning and Band Topology, Phys. Rev. Lett. **124**, 226401 (2020).

[47] Y. Che, C. Gneiting, T. Liu, and F. Nori, Topological quantum phase transitions retrieved through unsupervised machine learning, Phys. Rev. B **102**, 134213 (2020).

[48] N. Käming, A. Dawid, K. Kottmann, M. Lewenstein, K. Sengstock, A. Dauphin, and C. Weitenberg, Unsupervised machine learning of topological phase transitions from experimental data, Mach. Learn.: Sci. Technol. **2**, 035037 (2021).

[49] M. den Nijs and K. Rommelse, Preroughening transitions in crystal surfaces and valence-bond phases in quantum spin chains, Phys. Rev. B **40**, 4709 (1989).

[50] M. Oshikawa, Hidden $Z_2 \times Z_2$ symmetry in quantum spin chains with arbitrary integer spin, J. Phys.: Condens. Matter **4**, 7469 (1992).

[51] T. Kennedy and H. Tasaki, Hidden Z_2Z_2 symmetry breaking in Haldane-gap antiferromagnets, Phys. Rev. B **45**, 304 (1992).

[52] M. Levin and X.-G. Wen, Detecting Topological Order in a Ground State Wave Function, Phys. Rev. Lett. **96**, 110405 (2006).

[53] A. Kitaev and J. Preskill, Topological Entanglement Entropy, Phys. Rev. Lett. **96**, 110404 (2006).

[54] H. Li and F. D. M. Haldane, Entanglement Spectrum as a Generalization of Entanglement Entropy: Identification of Topological Order in Non-Abelian Fractional Quantum Hall Effect States, Phys. Rev. Lett. **101**, 010504 (2008).

[55] P. Calabrese and A. Lefevre, Entanglement spectrum in one-dimensional systems, Phys. Rev. A **78**, 032329 (2008).

[56] S. de Léséleuc, V. Lienhard, P. Scholl, D. Barredo, S. Weber, N. Lang, H. P. Büchler, T. Lahaye, and A. Browaeys, Observation of a symmetry-protected topological phase of interacting bosons with Rydberg atoms, Science **365**, 775 (2019).

[57] J. Preskill, Quantum computing in the NISQ era and beyond, Quantum **2**, 79 (2018).

[58] J. De la Porte, B. M. Herbst, W. Hereman, and S. J. Van Der Walt, An introduction to diffusion maps, in *Proceedings of the 19th Symposium of the Pattern Recognition Association of South Africa (PRASA 2008), Cape Town, South Africa* (Department of Electrical Engineering, University of Cape Town, Cape Town, 2008), pp. 15–25.

[59] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.105.235136, where we consider two additional spin chain models which can host SPT phases. We consider the following Hamiltonian with periodic boundary conditions as considered for other models in the main text. This includes Ref. [68].

[60] I. T. Jolliffe and J. Cadima, Principal component analysis: A review and recent developments, Philos. Trans. R. Soc. A **374**, 20150202 (2016).

[61] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Rigorous results on valence-bond ground states in antiferromagnets, in *Condensed Matter Physics and Exactly Soluble Models* (Springer, New York, 2004), pp. 249–252.

[62] C. Gils, E. Ardonne, S. Trebst, D. A. Huse, A. W. W. Ludwig, M. Troyer, and Z. Wang, Anyonic quantum spin chains: Spin-1 generalizations and topological stability, Phys. Rev. B **87**, 235120 (2013).

[63] Z.-C. Gu and X.-G. Wen, Tensor-entanglement-filtering renormalization approach and symmetry-protected topological order, Phys. Rev. B **80**, 155131 (2009).

[64] E. Tang and X.-G. Wen, Interacting One-Dimensional Fermionic Symmetry-Protected Topological Phases, Phys. Rev. Lett. **109**, 096403 (2012).

[65] M. F. Atiyah, V. K. Patodi, and I. M. Singer, Spectral asymmetry and Riemannian geometry. I, Math. Proc. Cambridge Philos. Soc. **77**, 43 (1975).

[66] M. F. Atiyah, V. K. Patodi, and I. M. Singer, Spectral asymmetry and Riemannian geometry. II, Math. Proc. Cambridge Philos. Soc. **78**, 405 (1975).

[67] A. Mesaros and Y. Ran, Classification of symmetry enriched topological phases with exactly solvable models, Phys. Rev. B **87**, 155115 (2013).

[68] T. Kennedy and Y. H. Tasaki, Hidden symmetry breaking and the Haldane phase in $S = 1$ quantum spin chains, Commun. Math. Phys. **147**, 431 (1992).

SUPPLEMENTAL MATERIAL

En-Jui Kuo[1,2] and Hossein Dehghani[1,2]

[1]*Department of Physics, University of Maryland, College Park, MD 20742, USA*
[2]*Joint Quantum Institute, NIST/University of Maryland, College Park, MD 20742, USA*

In this supplemental, we consider two additional spin chain models which can host SPT phases. We consider the following Hamiltonian with periodic boundary condition as considered for other models in the main text, [1]:

$$H = \sum_i S_x^i S_x^{i+1} + S_y^i S_y^{i+1} + \lambda S_z^i S_z^{i+1} + D(S_z^i)^2. \tag{S.1}$$

We can use the same techniques and with $L = 9$ and 2500 samples to implement the diffusion map algorithm (DMA) by using Euclidean distance and Chebyshev's distance. We can see sharper phase boundary condition by using Chebyshev's distance. We use L_2 and L_∞ to denote them, respectively. We plot the phase diagram via the DMA. The phase diagram Fig. 1 matches with the predictions in [1] to a great extent. In particular the distinction between the antiferromagnetic phase (AFM), Haldane and the ferromagnetic phases are completely distinguishable, however, we observe a less evident transition between the Haldane and the dimerized phases. We also notice that we have not detected the XY phase whose presence is still undetermined via exact methods [1].

The second model that we consider here is the following spin Hamiltonian with periodic boundary condition [1]:

$$H = \sum_i J_i (S_i \cdot S_{i+1} - \beta (S_i \cdot S_{i+1})^2). \tag{S.2}$$

$J_i = 1$ for i even, $J_i = \omega$ for i is odd. We can see our method Fig. 2 matches qualitatively with [1] and we can see a sharp transition between the topological Haldane phase and two symmetry broken phases i.e. valence bond state (VBS) and dimerized phases. We can also predict $\beta = -\frac{1}{3}$ [1] for the transition from VBS to Haldane phase.

I. PCA AND K-MEANS ALGORITHM

To compare the performance of the diffusion maps algorithm and other unsupervised methods, here, we try one of the most standard unsupervised algorithm called PCA (principle component analysis) on our sampling and compute the mean of 2 components. After doing this, we choose the number of clusters and use the standard k means algorithm. It should be highlighted that we do not have to pick the number of clusters in the diffusion map algorithm. We choose model 3 [2] in our main text.

$$H = \sum_i (\vec{S}_i \cdot \vec{S}_{i+1}) + B \sum_i S_x^i + \sum_i D(S_z^i)^2. \tag{S.3}$$

Here, we find that this method fails to identify the phase transition in fig 3 completely. This superior performance of the diffusion maps in comparison to other unsupervised learning methods, is due to the fact that in this approach we use non-linear kernel functions which allow us to detect complicated non-linear structures in data sets.

[1] T. Kennedyand H. Tasaki, Hidden symmetry breaking and the haldane phase in s= 1 quantum spin chains, Communications in mathematical physics **147**, 431 (1992).

[2] Z.-C. Guand X.-G. Wen, Tensor-entanglement-filtering renormalization approach and symmetry-protected topological order, Physical Review B **80**, 155131 (2009).

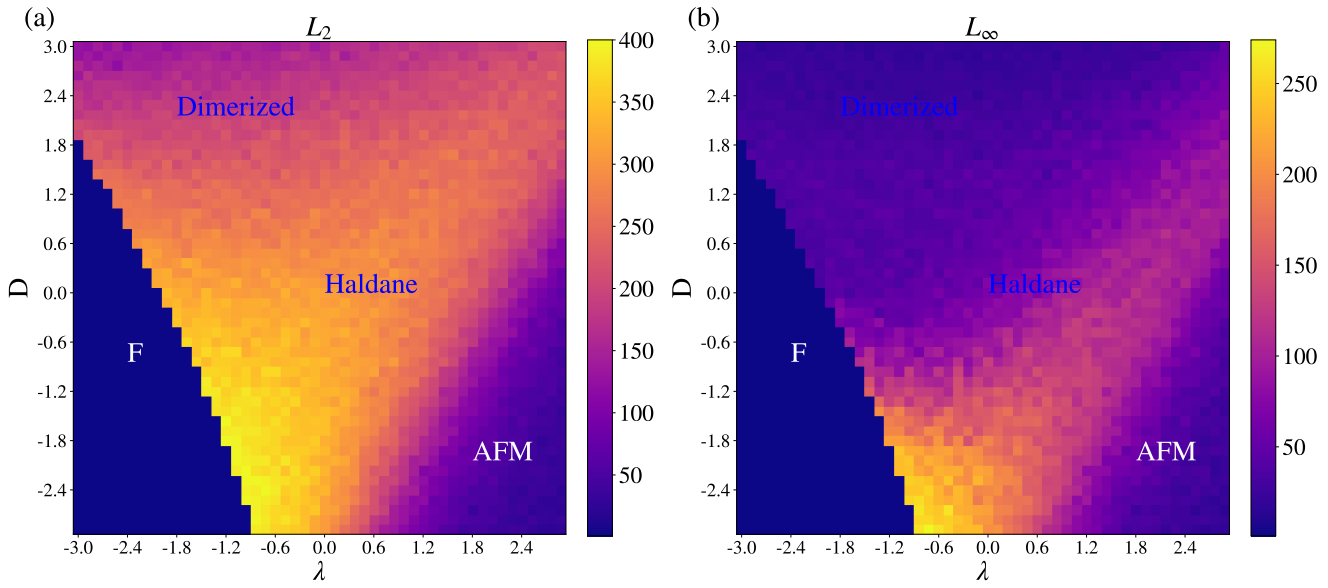

FIG. 1. Phase diagram of Hamiltonian in Eq.S.1 in terms of λ, D obtained via the diffusion maps algorithm. Dashed lines represent the exact phase boundaries obtained by DMRG methods.

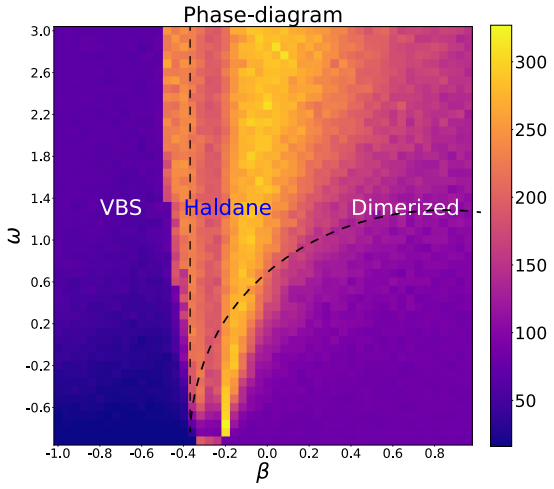

FIG. 2. Phase diagram of Hamiltonian in Eq.S.2 in terms of β, ω obtained via the diffusion maps algorithm. Dashed lines represent the exact phase boundaries obtained by DMRG methods.

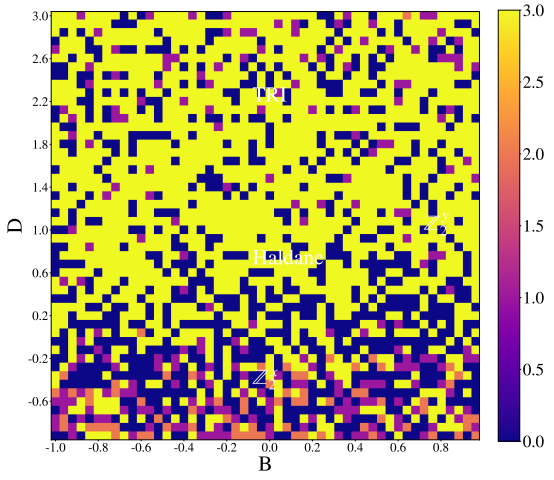

FIG. 3. PCA and k means on model 3