Lawrence Berkeley National Laboratory

Title

Application of the variational autoencoder to detect the critical points of the anisotropic Ising model

Permalink

Journal

ISSN

Authors

Baul, Anshumitra

Walker, Nicholas

Moreno, Juana

et al.

Publication Date

DOI

Copyright Information

Peer reviewed

Application of the Variational Autoencoder to Detect the Critical Points of the Anisotropic Ising Model

Anshumitra Baul

Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803, USA

Nicholas Walker

Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, Berkeley, California 94720, USA

Juana Moreno and Ka-Ming Tam

Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803, USA and
Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, USA
(Dated: November 2, 2021)

We generalize the previous study on the application of variational autoencoders to the two-dimensional Ising model to a system with anisotropy. Due to the self-duality property of the system, the critical points can be located exactly for the entire range of anisotropic coupling. This presents an excellent test bed for the validity of using a variational autoencoder to characterize an anisotropic classical model. We reproduce the phase diagram for a wide range of anisotropic couplings and temperatures via a variational autoencoder without the explicit construction of an order parameter. Considering that the partition function of $(d+1)$-dimensional anisotropic models can be mapped to that of the d-dimensional quantum spin models, the present study provides numerical evidence that a variational autoencoder can be applied to analyze quantum systems via Quantum Monte Carlo.

I. INTRODUCTION

Machine learning (ML) has become an indispensable tool to reach the boundaries of scientific understanding in the age of big data. An overflow of information is being analyzed using ML to quantify patterns in a large variety of fields from social networking, object and image recognition, advertising, finance, engineering, medicine, biological physics, and astrophysics among others [1].

Machine learning is a data modeling approach which employs algorithms which favor strategies driven by statistical analysis and based on pattern extraction and inference. ML algorithms, such as deep learning, provide new advances to understand physical data. Opportunities for scientific investigations are being devised particularly in numerical studies, which naturally involve large data sets and complex systems, where explicit algorithms are often challenging. A concerted effort to address large data problems in statistical mechanics and many-body physics using the ML approach is emerging [2–9]. The foundation of ML is deeply connected with statistical physics, and hence is fruitful to combine ML techniques with numerical methods which involve the prediction of phase transition regions. Scaling and renormalization are the core principles to understand macroscopic phenomena from microscopic properties. The way forward for machines to learn from large data sets would incorporate conceptually similar principles [9, 10].

Changes in the macroscopic properties of a physical system occur at phase transitions, which often involve a symmetry breaking process [11]. The theory of phase transitions and symmetry breaking was formulated by Landau as a phenomenological model, and was later devised from microscopic principles using the renormalization group [12]. Phases can be identified by an order parameter, which is zero in the disordered phase and finite in the ordered phase. The order parameter is determined by symmetry considerations of the underlying Hamiltonian. There are states of matter where the order parameter can only be defined in a complicated non-local way. These systems include topological states such as quantum spin liquids [13]. A major goal of the ML approach in complex statistical mechanics models or in strongly correlated systems is to detect the phase transitions from the data itself without explicitly constructing any order parameter [2].

The development artificial neural networks to detect phase transitions is a major advance in the area of ML applications in physics. In earlier works, artificial neural networks were based on supervised learning algorithms [2, 9]. Labeled data is used to train the supervised learning algorithm, from which the algorithm learns to assign labels to the input data points [14, 15]. Apart from supervised learning, another major type of ML is unsupervised learning for which the input data has no label. Conventional unsupervised learning algorithms, such as principal component analysis [16], find structure in unlabeled data without involving any artificial neural network. Here, the data is classified into clusters and labels can then be assigned to the data points accordingly [16].

Autoencoder is a new direction to utilize artificial neural networks in unsupervised machine learning. The very first versions of the autoencoder were being used for dimensional reduction of data before feeding its output into other ML algorithms [17, 18]. It is created by encoding an artificial neural network, which outputs a latent representation of the given data, and is utilized as the input of decoding neural network that tries to accurately reconstruct the input data from the latent representation

[19, 20].

A major shortcoming of autoencoder is the possibility of sharp changes in the latent representation with respect to small differences in the input data. Ideally, the latent representation should be a smooth function of the input data. The variational autoencoder (VAE) represents the latent representation in terms of probability distributions instead of a fixed set of numbers [21, 22]. This probabilistic latent representation allows a smooth latent representation. Since 2013, VAEs have been developed to become one of the most successful unsupervised learning algorithms [21]. Promising results are being shown in both encoding and reconstructing data [21–23].

VAEs are being successfully applied recently to detect phase transitions in classical spin models [23–25]. The input data sets are given by Monte Carlo method and then unsupervised machine learning, such as the VAE is used for deciphering and distinguishing different physics in the input data sets. After successfully applying to classical models, a natural question arises, whether such an approach remains viable for quantum models. In particular, can a VAE be viable to distinguish different quantum phases and transition regions based on the obtained data from quantum Monte Carlo. Recently, various models from statistical mechanics, in particular the Ising model and the Potts model have been investigated [26–28]. In this work, we investigate a rather simple quantum model, one-dimensional transverse field Ising model(TFIM), to address the capability and the limitations of the autoencoder. Given that, the critical line of the model can be calculated analytically due to self-dual property [29–31]. The TFIM is an excellent test bed to address the various aspects of VAEs in quantum models.

This paper is organized as follows. In section II, we briefly describe the transverse field Ising model (TFIM) and the Suzuki-Trotter formulation by mapping it to the anisotropic Ising model. In section III, the Monte Carlo method and the VAE are presented. The results from the VAE are described in the section IV. We conclude and discuss the implication and possible future applications of the method developed in this study in Section V. The self-duality of the anisotropic two-dimensional Ising model and the detail of the VAE are discussed in the appendix.

II. TRANSVERSE FIELD ISING MODEL

A. Model

We consider an Ising model with transverse field [32–34]. The Hamiltonian is given as

$$H = - \sum_{<i,j>} J_{ij} \sigma_i^z \sigma_j^z - \Gamma \sum_i \sigma_i^x, \qquad (1)$$

where $\sigma^\alpha (\alpha = x, y, z)$ are the Pauli matrices, which obey the commutation relation, $[\sigma_i^\alpha, \sigma_j^\beta] = 2\iota \delta_{ij} \epsilon_{\alpha\beta\gamma} \sigma_i^\gamma$. J_{ij} is the coupling between the spins at sites i and j. Only nearest neighbors coupling is considered in this study. Γ is the transverse field applied in the x-direction. ι is an imaginary number, and σ^z has the eigenvalues as ± 1. Their eigenvectors are symbolically denoted by $|\uparrow>.$ and $|\downarrow>$, that is

$$|\uparrow> = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (2)$$

and

$$|\downarrow> = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \qquad (3)$$

The order parameter is given by the average magnetization $m = \sum_i < \sigma_i^z > /N$ (N being the total number of sites) which characterizes the phase transition between paramagnetic and ferromagnetic phases. Without specification, we consider only the one dimensional TFIM with coupling limited to the nearest neighbors.

B. Suzuki-Trotter Formalism

Instead of working with the quantum spins directly, we follow the standard procedure of mapping a d-dimensional quantum Hamiltonian into a $d + 1$-dimensional effective classical Hamiltonian by the Suzuki-Trotter transformation [35–37]. We apply this to the TFIM system. We first define the longitudinal spin coupling terms and the transverse filed terms as follow

$$H_0 \equiv - \sum_{<i,j>} J_{ij} \sigma_i^z \sigma_j^z. \qquad (4)$$

$$V \equiv - \Gamma \sum_{i=1} \sigma_i^x.$$

$$H = V + H_0.$$

The partition function of H reads

$$Z = Tr(e^{-\beta(H_0 + V)}), \qquad (5)$$

β is the inverse temperature. The Trotter formula gives,

$$exp(A_1 + A_2) = \lim_{M \to \infty} [exp(A_1/M) exp(A_2/M)]^M \qquad (6)$$

when $[A_1, A_2] \neq 0$. Using the trotter formula we have

$$Z = \sum_i \lim_{M \to \infty} \langle s_i | [exp(-\beta H_0/M) exp(-\beta V/M)]^M | s_i \rangle, \qquad (7)$$

where s_i is the i-th spin configuration of the whole system, and the above summation runs over all the 2^N possible configurations denoted by i. M identifies the number of identity operators. The identity operator formed from the spin operators is given as,

$$I = \sum_i^{2^N} |s_{i,k}\rangle \langle s_{i,k}| \qquad (8)$$

$$= \sum_{\sigma_{i,k}=-1}^{1} |\sigma_{i,k}, ..., \sigma_{N,k}\rangle \langle \sigma_{i,k}, ..., \sigma_{N,k}|$$

where $k = 1, 2, ..., M$. Hence, it is the product of M exponentials in Z.

$$Z = \lim_{M \to \infty} Tr \prod_{k=1}^{M} \langle \sigma_{i,k}, ..., \sigma_{N,k}| \tag{9}$$

$$exp(-\frac{\beta H_0}{M})exp(-\frac{\beta V}{M})|\sigma_{i,k+1}, ...\sigma_{N,k+1}\rangle \tag{10}$$

Applying the periodic boundary conditions implies $\sigma_{N+1,p} = \sigma_{1,p}$. After evaluating the expression of the partition function,

$$Z = C^{\frac{NM}{2}} Tr_\sigma(-\beta H_{eff}[\{\sigma]\}), \tag{11}$$

where $C = \frac{1}{2} \sinh \frac{2\beta\Gamma}{M}$ and the effective classical Hamiltonian

$$H_{eff}(\{\sigma\}) = \sum_{<i,j>}^{N} \sum_{k=1}^{M} [\frac{-J_{ij}}{M}\sigma_{i,k}\sigma_{j,k} \tag{12}$$
$$-\frac{\delta_{ij}}{(2\beta)} \ln (\coth \frac{\beta\Gamma}{M})\sigma_{i,k}\sigma_{i,k+1}].$$

$\sigma_{i,k}$ involved are the eigenvalues of σ^z and hence there is no non-commuting part in H_{eff}. The effective Hamiltonian shows the system of spins in the $d + 1$-dimensional lattice and one extra label k for each spin variable. Each single quantum spin variable σ_i in the original Hamiltonian, now we have an array of M number of classical spins $\sigma_{i,k}$. Hence, the partition function of the 1D quantum Ising model is mapped to that of the 2D Ising model. This new (time-like) dimension along which these classical spins are spaced is called the Trotter dimension.

In this paper we assume the model only has nearest neighbor coupling on a square lattice. The coupling along the x-direction and y-direction are denoted as J_x and J_y respectively. We also define $K_x = 1/(\beta J_x)$ and $K_y = 1/(\beta J_y)$. We set $N = M$ in this study.

For the two-dimensional classical Ising model, the critical points can be obtained due to the self-dual property. The detail of the dual transformation is shown in the appendix A.

III. METHODOLOGY

A. Monte Carlo Sampling

The spin configurations are generated by using the single spin-flip Metropolis algorithm. In a single spin flip, first we flip the spin of a single lattice site, and then calculate the change in energy ΔE. After that, we use the resulting change in energy to calculate the Metropolis criterion $exp(-\Delta E/T)$. If a randomly generated uniform number is in the range of $[0, 1)$, and is smaller or equal to the Metropolis criterion, the configuration is accepted as the new configuration. The code for the simulation

is written in Python using the NumPy library [38–41]. We note that using generative neural network, instead of Monte Carlo method, for sampling has also been proposed recently for the 1D quantum spin models [42, 43].

B. Variational Autoencoder

Variational autoencoder(VAE) is under the category of generative models. New data can be produced by learning the distribution of the input data [21, 44]. We use an encoder-decoder architecture where the encoder maps the input to a latent representation in term of some chosen distributions. The latent distribution is mapped back to reconstruct the input using the decoder. The latent representation in a well-trained model can be used to generate new samples resembling the original training data [45].

In this study, we only consider a multi-dimensional Gaussian distribution for the latent space representation. We briefly explain the main concepts of the VAE in the following.

Encoder: Encoder is a neural network that takes the higher dimensional input data into a lower dimensional latent space. For example, an image of the spin configuration of size $32 \times 32 = 1024$ lattice points, can be converted into a vector of dimension 8. The encoder is a method for dimensionality reduction. The neural network representing the encoder maps each sample to a distribution and is also called a probabilistic encoder of VAE. The encoder is denoted as $q_\phi(z|x)$, which is a distribution that maps an input sample x (the Ising spin configurations) to provide a latent representation z (the parameters in the multi-dimensional Gaussian distribution). ϕ is a set of learnable parameters in the neural network that is varied to produce outputs from the encoder.

Latent Space: The latent space is the input for the decoder network, and it is outputted from the encoder network. For the VAE, the latent space is represented by multi-dimensional Gaussian distribution. As the Gaussian distribution is completely specified by its mean and standard deviations, the dimension of the latent space is two times as that of the dimension of the multi-dimensional Gaussian distribution. An encoded sample is denoted by z.

The latent space is regularized and then get penalized for deviating from the prior multi-dimensional Gaussian distribution by the Kullback-Leibler (KL) divergence term [46].

Decoder: From the fig. 1 one can see that the decoder in VAE converts the compressed samples in the latent space back to the original input samples [47] . It is represented as $p_\theta(x|z)$, a distribution to produce reconstructed samples x' conditioned on latent representations z. θ is a learnable parameter in the neural network that can be varied to produce different outputs from the decoder. A sampling distribution provides the input to the decoder which is described by the latent representation
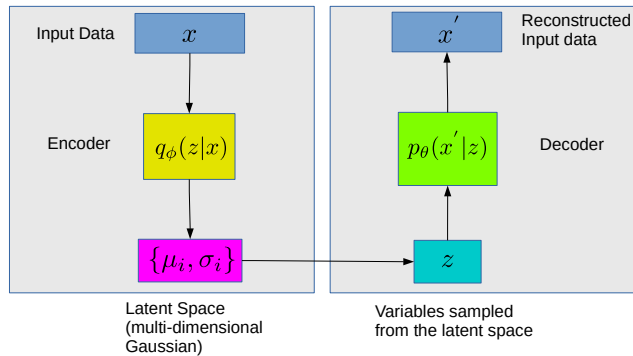
obtained from the encoder.

FIG. 1. Diagram depicting the structure of the VAE. The left hand side is the encoding part and the right hand side is the decoding part. x is the input Ising spin configuration, $q_\phi(z|x)$ is the encoder neural network, μ and σ are the latent means and standard deviations of the latent space distribution, z. $p_\theta(x'|z)$ is the decoder neural network and x' is the reconstructed Ising spin configuration.

Loss Functions in VAE: The VAE contained two sets of trainable parameters, θ and ϕ for the neural networks of the encoder and decoder respectively. They are trained by minimizing the loss function. The loss functions in VAE consist of two terms. The first one measures the "similarity" of the inputs and the reconstructed outputs. The second one measures the difference between the designated priori distribution, chosen to be multi-dimensional Gaussian distribution, and the actual distribution of the compressed inputs.

For the first term of the loss function, the standard reconstruction loss measures the error between the samples generated by the decoder and the original input samples. It is measured by the binary crossentropy between the encoder input and decoder output in our study. This is expressed as $L_{RC} = -E_{z \sim q_\phi(z|x)}[\log p_\theta(x|z)]$ The expectation, E, is over the representations, z with respect to the encoder's distribution.

The second is the Kullback-Leibler divergence (KLD) of the latent representations. KLD measures the divergence between the chosen latent representation $p(z)$ and the approximated distribution from the output of the encoder $q_\phi(z|x)$.

The KLD is defined as

$$L_{KLD} = D_{KL}\left[q_\phi(z|x)||p(z)\right] = -\sum_z q_\phi(z|x) \log\left(\frac{q_\phi(z|x)}{p(z)}\right).$$
(13)

It is minimized to optimize the latent representation from the encoder $q_\phi(z|x)$ to resemble the latent representations from $p(z)$.

The total loss is the sum of the reconstruction loss and the KLD [48],

$$L(\phi,\theta;x,z) = -E_{z \sim q_\phi(z|x)}\left[\log p_\theta(x|z)\right]$$
$$+D_{KL}\left[q_\phi(z|x)||p(z)\right] = L_{RC} + L_{KLD}. \quad (14)$$

The two losses in the VAE together are optimized simultaneously to describe the latent state for an observation with distributions close to the prior approximated distribution but also deviates to describe the salient features of the input.

The linear combination of the reconstruction loss and the KLD is often denoted as a variational lower bound or evidence lower bound loss function as both the reconstruction loss and the KLD are non-negative. Minimizing the loss, $min_{\theta,\phi}L(\theta,\phi;x,z)$, maximizes the lower bound of the probability of generating new samples [45].

C. β-total correlation VAE

A further refinement of the VAE can be achieved by decomposing the KLD into three parts. The three parts describing the index-code mutual information, total correlation, and dimension-wise Kulback-Leibler divergence, which are characterized by parameters often denoted as α, β, and γ respectively. β is the most important one for obtaining good results [48, 49].

A β- total correlation VAE (β-TCVAE) is defined as the VAE with the $\alpha = \gamma = 1$ and tuning β as a parameter. The detail is given in the appendix B. It is well suited for representation learning of patterns [49]. We fix the parameters of the decomposition with $\alpha = \gamma = 1$ and $\beta = 8$ in this study, these values were found to be good for finding the phase diagram of the 2D isotropic Ising model [25, 50].

Our goal is to map the raw Ising spin configurations to a reduced set of descriptors that discriminate between the samples using the criterion inferred by the β-total correlation VAE [25, 50]. The encoder and decoder are implemented as deep convolution neural networks (CNN) to preserve the spatially dependent 2-dimensional structure of the Ising spin configurations [51]. Scaled exponential linear unit (SELU) activation function is used in each convolution layer. The output of the final convolution layer is flattened, and fed into two 2-dimensional dense layers. Then, it is used as the input layer for the decoder CNN and reshaped to match the structure of the output from the final convolution layer in the encoder CNN. Hence, the decoder network is simply the opposite of the encoder network, incorporated with convolution transpose layers in favor of standard convolution layers [50–52]. The final output layer from the decoder network is reproduced from the original input configurations obtained from the encoder network which uses a sigmoid activation function. The loss term consists of the reconstruction loss and the β-total correlation KLD (β-TCKLD) term with $\alpha = \gamma = 1$ and $\beta = 8$ [46, 50]. We employ minibatch stratified sampling on the given data while training.

To optimize the loss, Nesterov-accelerated Adaptive Moment Estimation(Nadam) was used and it efficiently minimizes the loss during the training of the β-TCVAE model [50, 53]. The default parameters provided by the

Keras library and a learning rate of 0.00001 were chosen. The training is performed over 100 epochs with a batch size of $33 \times 33 = 1089$ for both lattice size $N = 64$ and $N = 128$ with a number of sample 1024 per phase point. The reduced descriptors of the 2D Ising spin configuration are given by the latent variables [50]. The β-TCVAE model used in this work was implemented using the Keras ML library with TensorFlow as a supporting backhand [54, 55].

D. Principal Component Analysis on the Latent Space

The principal components analysis (PCA) is applied on the latent means and standard deviations of the Ising spin configurations obtained after fitting the β-TCVAE, independently to produce linear transformations of the Gaussian parameters that discriminate between the samples using the scikit-learn package [16, 56]. The PCA performs an orthogonal transformation into a new basis set of linearly uncorrelated features, principal components. Each principal component encompasses the largest possible variance across the sample space under an orthogonality constraint [16].

The latent representation characterizes the structure of the Ising configurations, but the principal components of the latent representation shows greater discrimination between the different structural characteristics of the configurations compared to the raw variable representations [50]. The rationale of using the PCA is to provide a more compact representation which characterizes the different phases of the Ising model. As we will show in the results section, the first and the second dominated components already show good signals to distinguish different phases of the anisotropic Ising model.

IV. RESULTS

As explained in the previous section, for the VAE models, the samples were drawn from a multi-dimensional Gaussian distribution parametrized by vectors of means and standard deviations denoted as μ_i and σ_i respectively, where i is the index of the dimension of the distribution. All of the plots in this paper were generated with the MatPlotLib package using a perceptually uniform colormap [57]. In each plot, the coloration of each square pixel on the diagram represents the ensemble average value of the principal components of the latent space mean or variance of the multi-dimensional Gaussian function. We study two different system sizes $N = 64$, and 128. We will focus on the first two principal components of the latent variance and the second principal component of the latent mean. We denote them as τ_0, τ_1, and ν_1 respectively. The first principal component of the latent mean, ν_0 does not capture a clear distinction between the ferromagnetic and paramagnetic phases of the system.

ν_1 are plotted in fig. 2. It is apparent that ν_1 is a resemblance of the magnetization, m, of snapshots of the Ising spin configuration. We note that it is not to be expected that any of the latent variables to have the same value of any physical quantities, such as magnetization. Nonetheless, from the plot of ν_1, it clearly discriminates the ferromagnetic from paramagnetic phases, separating them by the phase transition line. The phase transition line in white corresponds to the analytical solution, eq. A14 in the appendix calculated in the thermodynamic limit. Since the magnetization can be seen as the order parameter for the 2-dimensional Ising model, a reasonable representation of the order parameter is seen to be possibly extracted from the VAE. We note that the simulations and the VAE are done on finite size systems, a true broken symmetry does not occur, that is the reason of the seemingly random values of ν_1 in the ferromagnetic phase. Usually, the amplitude of the magnetization is considered for finite size simulation, as we will show other latent variables from the VAE has similar structure as the amplitude of the magnetization. As the magnetization is a linear feature of the Ising spin configuration, a simpler linear model would be sufficient for extracting the magnetization.

The first principal component of the latent variance τ_0 are plotted in fig 3. The white line is the analytical solution for the phase transition points. One can find that the value of τ_0 remains very small in the upper right region of the figures. This can be considered as the reflection of the small changes of amplitude of the energy or amplitude of the magnetization in the paramagnetic phase. Once the systems approach the critical line from the upper right region of the figure, the value of τ_0 increase sharply. This behavior is again consistent with the amplitude of energy or magnetization.

In particular, we consider the case for the isotropic limit ($K_x = K_y$), that is the classical ($\Gamma = 0$) limit. The critical point is given as $K_c(= K_x = K_y) = \frac{2}{\ln[1+\sqrt{2}]} \sim 2.2721$ [58]. From the fig. 3, we see a sharp change in the τ_0 around the value of K_c. This result is consistent with the accomplishments of prior published works for the isotropic Ising model on a square lattice [23–25, 50].

Given that the paramagnetic samples are essentially noise due to entropic contributions. Therefore, these happens to be easy to discriminate from the rest of the samples using a β-TCAVE model. In view of the fact, the samples with ν_1 values corresponding to nearly zero magnetization and rather high values for τ_0 resemble Gaussian noise with no notable order preference. An interesting question is to look for quantity which may resemble to response function, such as susceptibility and heat capacity.

We plot the second principal component of the latent variance τ_1 in fig. 4. The phase transition line is again included to show the transition in the thermodynamic
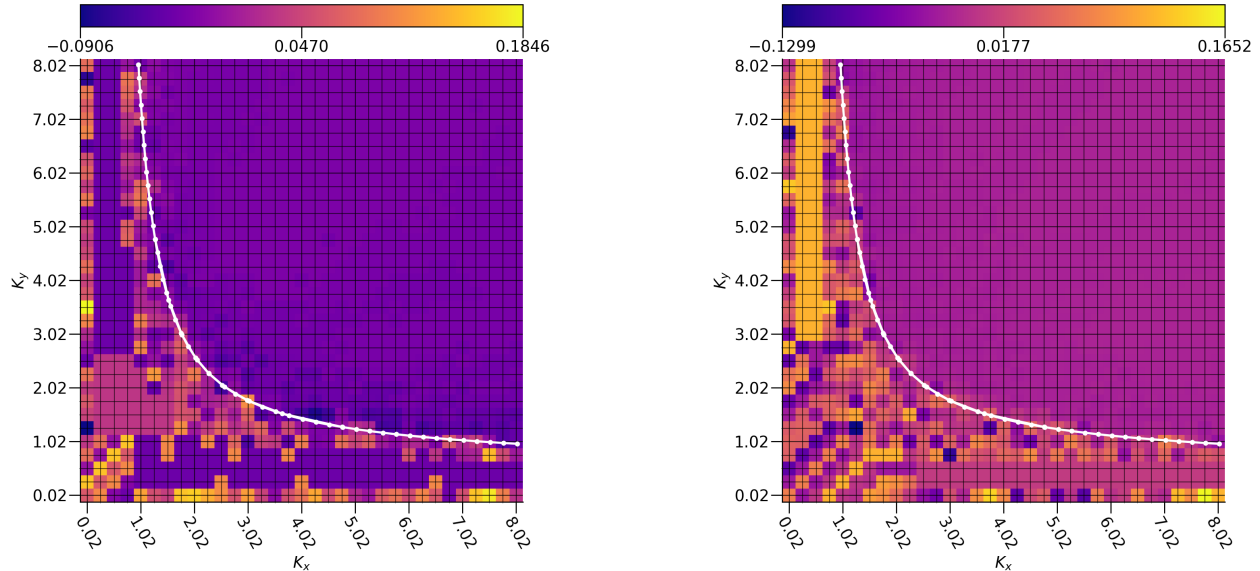
FIG. 2. The second principal component of the latent mean, ν_1, with respect to the different K_x and K_y $(1/(\beta J_x)$ and $1/(\beta J_y))$ for the $2D$ square lattice. The left panel and right panel are for the system sizes $N = M = 64$ and $N = M = 128$, respectively.
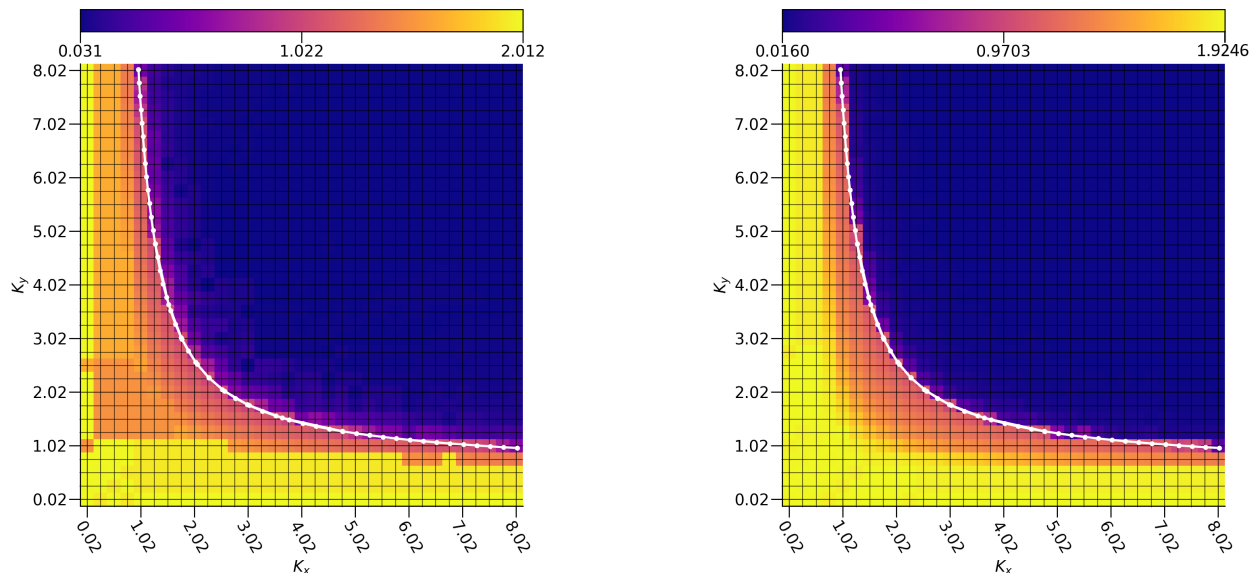

FIG. 3. The first principal component of the latent variance, τ_0, with respect to the different K_x and K_y $(1/(\beta J_x)$ and $1/(\beta J_y))$ for the $2D$ square lattice. The left panel and right panel are for the system sizes $N = M = 64$ and $N = M = 128$, respectively.

limit for comparison. The values of τ_1 are much larger along the analytical critical line than the other combinations of K_x and K_y. Thus τ_1 bears a strong resemblance to the magnetic susceptibility.

From the plots of τ_0, τ_1, and ν_1, we find that the VAE is capable to generate quantities which resembles to the magnetization, amplitude of magnetization/energy, and

magnetic susceptibility/heat capacity, respectively. From these quantities we can draw the phase transition line between the paramagnetic phase and the ferromagnetic phase.

The quality of fitting the Ising spin configurations to the present VAE model can be quantified by the values of the loss functions. The three losses represented
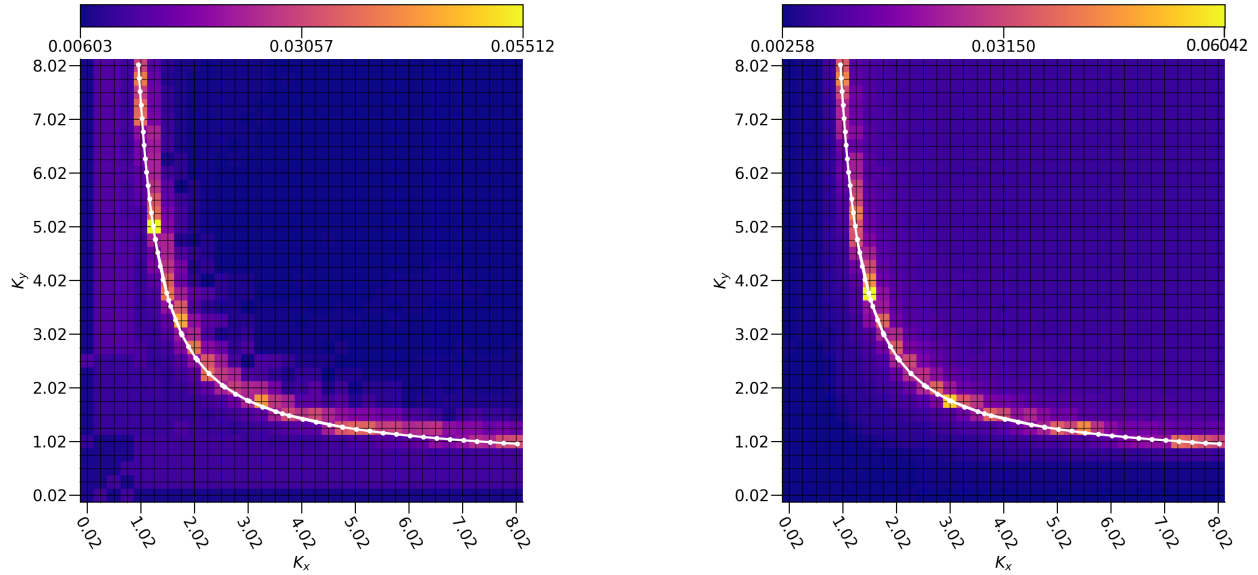
FIG. 4. The second principal component of the latent variance, τ_1, with respect to the different K_x and K_y ($1/(\beta J_x)$ and $1/(\beta J_y)$) for the $2D$ square lattice. The left panel and right panel are for the system sizes $N = M = 64$ and $N = M = 128$, respectively.
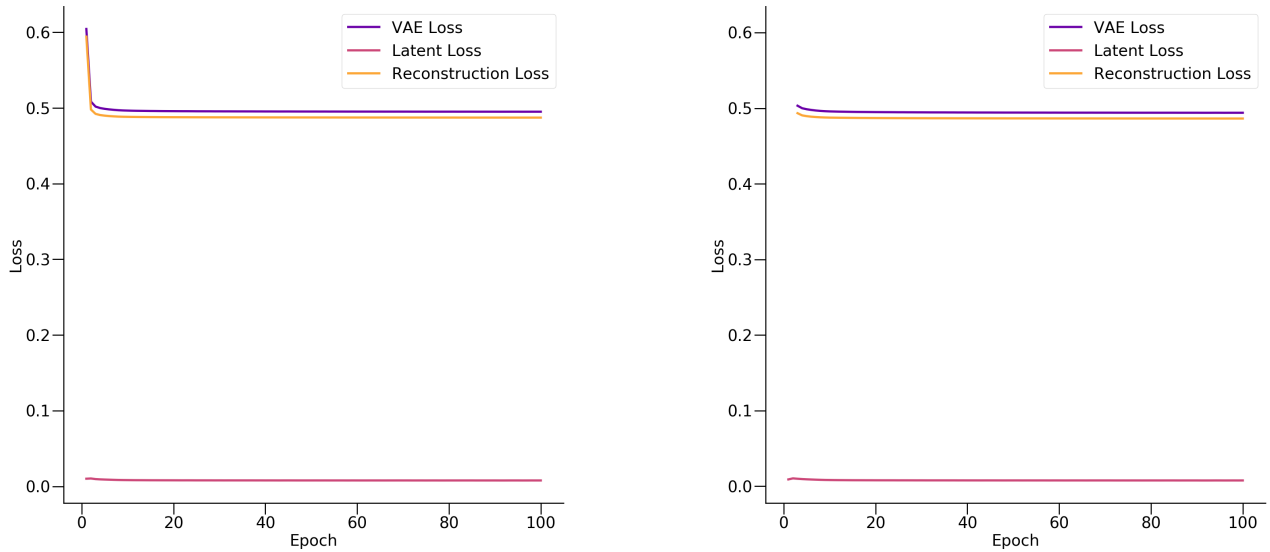

FIG. 5. β-TCVAE loss and the reconstruction loss settle to about 0.5 and the latent loss settles close to 0 for 100 epoch cycles. The left panel and right panel are for the square lattice of size $N = M = 64$ and $N = M = 128$, respectively.

in fig 5 are the VAE loss, reconstruction loss, and the latent loss. The reconstruction loss converges to about 0.5. The latent loss is obtained from the β-TCKLD term. This loss converges quickly to a value near to 0. The total β-TCVAE loss for the 2-dimensional anisotropic Ising model for both lattice size $N = 64$ and $N = 128$ settles quickly around 0.5.

V. DISCUSSION AND CONCLUSIONS

We use the β-TCVAE to extract structural information from raw Ising spin configurations. It exposes interesting derived descriptors of the configurations used to identify the second order phase transition line among other regions. The analysis here shows the interpretation of the extracted sample space as represented by the latent variables. This is done by studying the behavior of the latent variable mappings of the Ising spin configurations with respect to the anisotropic coupling (J_x, J_y) associated with the temperatures.

We find that ν_1, the second principal component of the latent mean, is considered to reflect the magnetization of the 2-dimensional anisotropic Ising model, hence it is interpreted as an indicator for the magnetization exhibited by the configurations. In contrast, τ_0 and τ_1, which are the first two principal components of the latent variance, can be seen as an indicator of the amplitude of magnetization or energy and the magnetic susceptibility or heat capacity. Thus both τ_0 and τ_1 can provide a reasonable estimate of the second order phase transition line.

Since the $d + 1$-dimensional anisotropic Ising model is equivalent to the d-dimensional quantum spin system through the Suzuki-Trotter transformation. This method can be trivially extended to other 1D quantum systems [59].

Moreover, methods for strongly correlated systems, such as the dynamical mean field theory (DMFT) and their cluster generalizations–the dynamical cluster approximation (DCA) and the cellular dynamical mean field theory (CDMFT)–have very similar data structure in their Hirsch-Fye impurity solver [60–62]. It is notoriously difficult to obtain the putative quantum critical point from the paramagnetic solution of the Hubbard model as there is no simple quantity to track the transition [63, 64]. Therefore, the method developed here would be an important tool for analyzing the data from the DMFT, DCA, and CDMFT.

There are many opportunities for further developing this method, not only investigating more complex systems, but also by introducing improvements which are beyond the scope of this work. Finite size scaling turns to be an important approach to address the limitations in solving finite-sized systems for investigating regions near critical phenomena [65, 66]. A correspondence established between the VAE encodings of different system sizes is a challenging argument, as different VAE structures need specific training for each system size, which in turn demands for different hyperparameters and training iterations, counts to obtain equivalent analysis and hence similar answers [65]. Numerical difficulties tend to arise while performing finite-size scaling analysis, because the variation of predicted properties with respect to the system size is hard to isolate from the systemic variation due to different neural networks trained with different hyperparameters is being used to extract the specific macroscopic properties. However, this would play a significant role towards improving VAE characterization of critical phenomena.

Another interesting direction is to use the generative adversarial network (GAN) [67] instead of VAE as the generative model, promising results have been obtained for the isotropic 2D Ising model [68].

VI. ACKNOWLEDGMENT

Appendix A: Self-Duality of the two-dimensional Ising Model

In this appendix, we summarize the derivation of the critical points for the anisotropic 2-dimensiaonl Ising model via the self-duality property of the model [29, 69, 70]. We closely follow the lecture notes by Muramatsu in the following [70].

Considering the partition function of the Ising model (where $K = \frac{1}{\beta J}$)

$$Z = \sum_{\{S_j\}} e^{K \sum_{<j,l>} S_j S_l} = \sum_{\{S_j\}} \prod_{<j,l>} e^{K S_j S_l}$$

$$= \sum_{\{S_j\}} \prod_{<j,l>} \sum_{r=0}^{1} C_r(K)(S_j S_l)^r \quad \text{(A1)}$$

where $C_0(K) = \cosh K$ and $C_1(K) = \sinh K$. Applying the simple transformation, for each bond $< j, l >$, a new Z_2 variable, r is introduced. We label the new variable as r_μ with $\mu \equiv (i, < i, j >)$, labelling it with the site i from which the bond $< i, j >$ emanates. The partition function thus follows

$$Z = \sum_{\{S_j\}} \sum_{\{r_\mu\}} \prod_{<j,l>} C_{r_\mu}(K) \prod_i S_i^{\sum_{<i,j>} r_\mu} \quad \text{(A2)}$$

grouping all products of spins on site i together, $\sum_{<i,j>} r_\mu$, contains all four contributions, resulted due to the bonds connected to site i. Further, we perform explicitly the sum over all spin configurations

$$Z = \sum_{\{r_\mu\}} \prod_{<j,l>} C_{r_\mu}(K) \prod_i \sum_{S_i = \pm 1} S_i^{\sum_{<i,j>} r_\mu}$$

$$= \sum_{\{r_\mu\}} \prod_{<j,l>} C_{r_\mu}(K) \prod_i 2\delta[\mod_2(\sum_{<i,j>} r_\mu)] \quad \text{(A3)}$$

We define a dual lattice, where the vertices of the dual lattice are set in the center of the plaquettes defined by the original lattice. We have vanishing contributions due to the presence of the Kronecker delta in many configurations. Defining the new Z_2 variables $\sigma_i = \pm 1$ on the sites of the dual lattice. We can associate for each link of the original lattice, there is a pair of σ_i's (e.g. on the sites i and j on the dual lattice). Therefore, the variable r_μ is expressed as:

$$r_\mu = \frac{1}{2}(1 - \sigma_i \sigma_j) \tag{A4}$$

where sites i and j on the dual lattice are those where the link crosses r_μ. The sum of r_μ is over the four nearest neighbors of a site i, we have

$$\sum_{<i,j>} r_\mu = 2 - \frac{1}{2}(\sigma_1 \sigma_2 + \sigma_2 \sigma_3 + \sigma_3 \sigma_4 + \sigma_4 \sigma_1) \tag{A5}$$

There are 2^4 possible configurations for the four variables $\sigma_1, ..., \sigma_4$. They are grouped in four cases and all the cases lead to an even number for the sum of r_μ over the nearest neighbor bonds. The choice of variables needs to satisfy the $\delta-$function. The partition function becomes

$$Z = \frac{1}{2} 2^N \sum_{\{\sigma_i\}} \prod_{<j,l>} C_{[(1-\sigma_i\sigma_j)/2]}(K), \tag{A6}$$

where N is the number of sites on the lattice and the product is now over the bonds associated in the dual lattice. The expression of the partition function shows that the weight for each configuration of the σ's is given by the coefficients $C(K)$. Hence, introducing $C(K)$ to a form that resembles a Boltzmann weight.

$$\begin{aligned} C_r(K) &= \cosh(K)[1 + r(\tanh(K) - 1)] \tag{A7} \\ &= \cosh(K) \exp(\ln[1 + r(\tanh(K) - 1)]) \\ &= \cosh(K) \exp(r \ln \tanh(K)) \\ &= \cosh(K) \exp[\frac{1}{2}(1 - \sigma_i \sigma_j) \ln \tanh(K)] \\ &= [\cosh(K) \sinh(K)]^{1/2} \exp(-\frac{1}{2} \ln \tanh K \sigma_i \sigma_j) \end{aligned}$$

The partition function becomes

$$Z = 1/2(\sinh 2\tilde{K})^{-N} \sum_{\{\sigma_i\}} exp(\tilde{K} \sum_{<j,l>} \sigma_j \sigma_l). \tag{A8}$$

There are $2N$ bonds and defining the new coupling constant as

$$\tilde{K} \equiv -\ln \tanh(K/2) \tag{A9}$$

(where $K = \frac{1}{\beta J}$) coupling for the Ising model on the dual lattice. The Ising model is self-dual since the duality

transformations brings it into itself. We consider the free energy per site

$$f(Z) = -\frac{1}{N} \ln Z. \tag{A10}$$

According to the relation between the partition functions of the original and the dual model, we can write

$$f(K) = \ln \sinh(2\tilde{K}) + f(\tilde{K}). \tag{A11}$$

This is a strong constraint on the free energy. Since, $\sinh(2\tilde{K})$ is an analytic function, the equation has a singularity in $f(K)$, which corresponds to a singularity in $f(\tilde{K})$. $\tilde{K}(K)$ is a monotonous function of K, hence it holds $\tilde{K}_c = K_c$. We have

$$K_c = \frac{1}{2} \ln(1 + \sqrt{2}). \tag{A12}$$

The self-duality has allowed us to calculate the exact value of the critical temperature in the 2-dimensional isotropic Ising model($K_x = K_y$), where $K_x = \frac{1}{\beta J_x}$ and $K_y = \frac{1}{\beta J_y}$. Generalizing the results obtained in the isotropic to the anisotropic one, i.e. when couplings $K_x \neq K_y$ in the respective directions. The anisotropic case is as follows

$$\tilde{K}_y \equiv -\frac{1}{2} \ln \tanh(K_x), \tilde{K}_x \equiv -\frac{1}{2} \ln \tanh(K_y). \tag{A13}$$

Given K_x and K_y there is only one critical point, summarizing both the equations above with the following condition for a critical line separating the ordered from the disordered phase in the anisotropic Ising model

$$\sinh(2K_{xc}) \sinh(2K_{yc}) = 1. \tag{A14}$$

Appendix B: β-TCVAE Loss

The expression of the total loss for the VAE is given by

$$L(\phi, \theta; x, z) = L_{RC} + L_{KLD}, \tag{B1}$$

where the reconstruction error (RC) and the Kullback-Leibler divergence (KLD) are defined as

$$L_{RC} = -E_{z \sim q_\phi(z|x)} [\log p_\theta(x|z)] \tag{B2}$$

and

$$L_{KLD} = D_{KL} [q_\phi(z|x)||p(z)] \tag{B3}$$

respectively. If the prior distribution of the latent representation is Gaussian, the VAE provides disentangled factors in the latent representation, which means the significant dimension of the latent space is largely independent of each other. In β-total correlation VAE (β-TCVAE), we try to improve the disentanglement of factors in the representation by decomposing the KLD

term and apply tuning parameters to them independently [49, 50]. Each training sample is identified with a unique integer index $n \in 1, 2, .., N$ and assigned a uniform random variable in this decomposition. The aggregated posterior, $q_\phi(z) = \sum_n q_\phi(z|n)p(n)$ captures the aggregate structure of the latent variables under the distribution of the input, where $q_\phi(z|n) = q_\phi(z|x_n)$ and $q_\phi(z,n) = q_\phi(z|n)p(n) = \frac{1}{N}q_\phi(z|n)$. The decomposition is given as

$$I(z;x) + D_{KL}[q_\phi(z)||\prod_j q_\phi(z_j)]$$
$$+ \sum_j D_{KL}[q_\phi(z_j)||p(z_j)] \quad (B4)$$

The first term is the index-code mutual information (MI), $I(z;x) = D_{KL}[q_\phi(z,n)||q_\phi(z)p(n)]$, between the input and the latent variable, which is based on the empirical input distribution $q_\phi(z,n)$. The second term is a mea-

sure of the dependence between the latent variables and is called the total correlation(TC). It is important to produce representations which penalize the total correlation and hence force the model towards discovering statistically disentangled factors in the input distribution. The third term prevents the individual latent variables in the representation from deviating far from their priors. It is called the dimension-wise KLD [50]. After adding the tuning parameters to the decomposition, the β-TC modified KLD term becomes :-

$$L_{\beta-TC} = \alpha I(z;x) + \beta D_{KL}[q_\phi(z)||\prod_j q_\phi(z_j)]$$
$$+ \gamma \sum_j D_{KL}[q_\phi(z_j)||p(z_j)] \quad (B5)$$

Modulating only the β parameter shows the greatest effect on disentanglement in the latent representation given by empirical evidence [49].

[1] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, B. C. V. Esesn, A. A. S. Awwal, and V. K. Asari, arXiv:1803.01164 (2018).

[2] J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431–434 (2017).

[3] G. Pilania, J. E. Gubernatis, and T. Lookman, Phys. Rev. B **91**, 214302 (2015).

[4] N. Walker, K.-M. Tam, B. Novak, and M. Jarrell, Phys. Rev. E **98**, 053305 (2018).

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

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

[7] S. J. Wetzel and M. Scherzer, Physical Review B **96**, 184410 (2017).

[8] L. Huang and L. Wang, Phys. Rev. B **95**, 035105 (2017).

[9] G. Torlai and R. G. Melko, Phys. Rev. B **94**, 165134 (2016).

[10] P. Mehta and D. J. Schwab, arXiv:1410.3831 (2014).

[11] L. D. Landau, Phys. Z. Sowjet. **11**, 26 (1937).

[12] M. E. Fisher, Rev. Mod. Phys. **46**, 597 (1974).

[13] L. Savary and L. Balents, Rep. Prog. Phys. **80**, 016502 (2016).

[14] J. Schmidhuber, Neural Networks **61**, 85 (2015).

[15] T. O. Ayodele, New advances in machine learning **3**, 19 (2010).

[16] K. Pearson, London Edinburgh Philos. Mag. J. Sci. **2**, 559 (1901).

[17] Y. Wang, H. Yao, and S. Zhao, Neurocomputing **184**, 232 (2016).

[18] J. Wang, H. He, and D. V. Prokhorov, Procedia Comput. Sci. **13**, 120 (2012).

[19] J. Almotiri, K. Elleithy, and A. Elleithy, in *2017 IEEE Long Island Systems, Applications and Technology Conference (LISAT)* (IEEE, 2017) pp. 1–5.

[20] E. Plaut, arXiv:1804.10253 (2018).

[21] D. P. Kingma and M. Welling, arXiv:1312.6114 (2014).

[22] D. P. Kingma and M. Welling, arXiv:1906.02691 (2019).

[23] S. J. Wetzel, Phys. Rev. E **96**, 022140 (2017).

[24] C. Alexandrou, A. Athenodorou, C. Chrysostomou, and S. Paul, The European Physical Journal B **93**, 10.1140/epjb/e2020-100506-5 (2020).

[25] N. Walker, K.-M. Tam, and M. Jarrell, Sci. Rep. **10** (2020).

[26] D. Kim and D.-H. Kim, Phys. Rev. E **98**, 022138 (2018).

[27] A. Morningstar and R. G. Melko, J. Mach. Learn. Res. **18**, 17 (2018).

[28] K. Shiina, H. Mori, Y. Okabe, and H. K. Lee, Sci. Rep. **10** (2020).

[29] H. A. Kramers and G. H. Wannier, Phys. Rev. **60**, 263 (1941).

[30] G. H. Wannier, Rev. Mod. Phys. **17**, 50 (1945).

[31] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Elsevier, 2016).

[32] P. Pfeuty, Ann. Phys. (N. Y.) **57**, 79 (1970).

[33] R. B. Stinchcombe, J. Phys. C: Solid State Phys. **6**, 2459 (1973).

[34] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, 1995).

[35] M. Suzuki, *Quantum Monte Carlo Methods in Equilibrium and Nonequilibrium Systems*, Vol. 74 (Springer Science & Business Media, 2012).

[36] M. Suzuki, Prog. Theor. Phys. **46**, 1337 (1971).

[37] B. K. Chakrabarti and A. Das, in *Quantum Annealing and Other Optimization Methods* (Springer, 2005) pp. 1–36.

[38] G. Van Rossum and F. L. Drake Jr, *Python reference manual* (Centrum voor Wiskunde en Informatica Amsterdam, 1995).

[39] S. van der Walt, S. C. Colbert, and G. Varoquaux, Comput. Sci. Eng. **13**, 22 (2011).

[40] D. D. Team, Dask: Library for dynamic task scheduling (2016).

[41] S. K. Lam, A. Pitrou, and S. Seibert, in *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, LLVM '15 (ACM, New York, NY, USA, 2015)

pp. 7:1–7:6.

[42] K. A. Nicoli, S. Nakajima, N. Strodthoff, W. Samek, K.-R. Müller, and P. Kessel, Phys. Rev. E **101** (2020).

[43] J. Vielhaben and N. Strodthoff, Physical Review E **103** (2021).

[44] C. Doersch, arXiv:1606.05908 (2016).

[45] A. B. L. Larsen, S. K. Sønderby, H. Larochelle, and O. Winther, in *International conference on machine learning* (PMLR, 2016) pp. 1558–1566.

[46] S. Kullback and R. A. Leibler, Ann. Math. Stat. **22**, 79 (1951).

[47] L. Weng, lilianweng.github.io/lil-log (2018).

[48] C. P. Burgess, I. Higgins, A. Pal, L. Matthey, N. Watters, G. Desjardins, and A. Lerchner, arXiv preprint arXiv:1804.03599 (2018).

[49] R. T. Chen, X. Li, R. Grosse, and D. Duvenaud, arXiv preprint arXiv:1802.04942 (2018).

[50] N. Walker, Identifying structure transitions using machine learning methods, louisiana state university thesis (2020).

[51] W. Zhang, K. Itoh, J. Tanida, and Y. Ichioka, Appl. Opt. **29**, 4790 (1990).

[52] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, arXiv:1706.02515 (2017).

[53] S. Ruder, arXiv preprint arXiv:1609.04747 (2016).

[54] L. Banarescu, C. Bonial, S. Cai, M. Georgescu, and K. Griffitt, Networks **25**, 211 (2016).

[55] F. Chollet *et al.*, Keras (2015).

[56] L. van der Maaten and G. Hinton, J. Mach. Learn. Res. **9**, 2579 (2008).

[57] J. D. Hunter, Comput. Sci. Eng. **9**, 90 (2007).

[58] L. Onsager, Phys. Rev. **65**, 117 (1944).

[59] H. Betsuyaku, Prog. Theor. Phys. **73**, 319 (1985).

[60] T. Maier, M. Jarrell, T. Pruschke, and M. H. Hettler, Rev. Mod. Phys. **77**, 1027 (2005).

[61] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

[62] J. E. Hirsch and R. M. Fye, Phys. Rev. Lett. **56**, 2521 (1986).

[63] N. Vidhyadhiraja, A. Macridin, C. Şen, M. Jarrell, and M. Ma, Phys. Rev. Lett. **102**, 206407 (2009).

[64] S. Kellar and K.-M. Tam, arXiv (2020), arXiv:2008.12324 [cond-mat.str-el].

[65] J. Cardy, *Finite-size Scaling*, Current physics (North-Holland, 1988).

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

[67] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial networks (2014), arXiv:1406.2661 [stat.ML].

[68] N. Walker and K.-M. Tam, Mach. learn.: sci. technol. **2**, 025001 (2021).

[69] J. B. Kogut, Rev. Mod. Phys. **51**, 659 (1979).

[70] A. Muramatsu, The Ising Model, Duality, and Transfer Matrix - Lattice Gauge Theory (2009).