MACHINE LEARNING
Science and Technology

PAPER

Restricted Boltzmann machine representation for the groundstate and excited states of Kitaev Honeycomb model

Mohammadreza Noormandipour[1] ![ORCID], Sun Youran[2] and Babak Haghighat[2,*]

[1] TCM Group, Cavendish Laboratory, University of Cambridge, J.J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom
[2] Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, People's Republic of China
* Author to whom any correspondence should be addressed.

E-mail: babakhaghighat@tsinghua.edu.cn

Abstract

In this work, the capability of restricted Boltzmann machines (RBMs) to find solutions for the Kitaev honeycomb model with periodic boundary conditions is investigated. The measured groundstate energy of the system is compared and, for small lattice sizes (e.g. 3×3 with 18 spinors), shown to agree with the analytically derived value of the energy up to a deviation of 0.09%. Moreover, the wave-functions we find have 99.89% overlap with the exact ground state wave-functions. Furthermore, the possibility of realizing anyons in the RBM is discussed and an algorithm is given to build these anyonic excitations and braid them for possible future applications in quantum computation. Using the correspondence between topological field theories in $(2+1)$d and 2d conformal field theories, we propose an identification between our RBM states with the Moore-Read state and conformal blocks of the 2d Ising model.

1. Introduction

The Honeycomb model we consider in this work is a two-dimensional lattice spin system (see figure 1) which was studied in detail by Kitaev in 2006 [1] and is famous because of the topological quantum order due to a degenerate gapped groundstate which is persistent to local and finite-sized perturbations. The system also supports both Abelian and non-Abelian topological phases as demonstrated in the original proposal [1]. There is a wide range of applications for this model, from fault-tolerant quantum computation [2] to analytical study of strongly correlated systems [3] and quantum spin liquids [4].

The honeycomb lattice in figure 3, is not a Bravias lattice in its original structure, but it can be considered as a triangular Bravias lattice with a two-spin basis. The direct Bravias lattice and the primitive cells are illustrated in figure 1(a) with the dashed lines. Each primitive cell has a pair of odd and even indexed spins. Just for the sake of simplicity, we will denote each cell with the even indexed (empty circles) spins of the system. The primitive vectors of the lattice, $\mathbf{a}_1$ and $\mathbf{a}_2$ are also shown in figure 1(a) (see also equation (1)). The entire lattice can be tiled and covered with primitive cells using translations composed of different linear combinations of primitive vectors
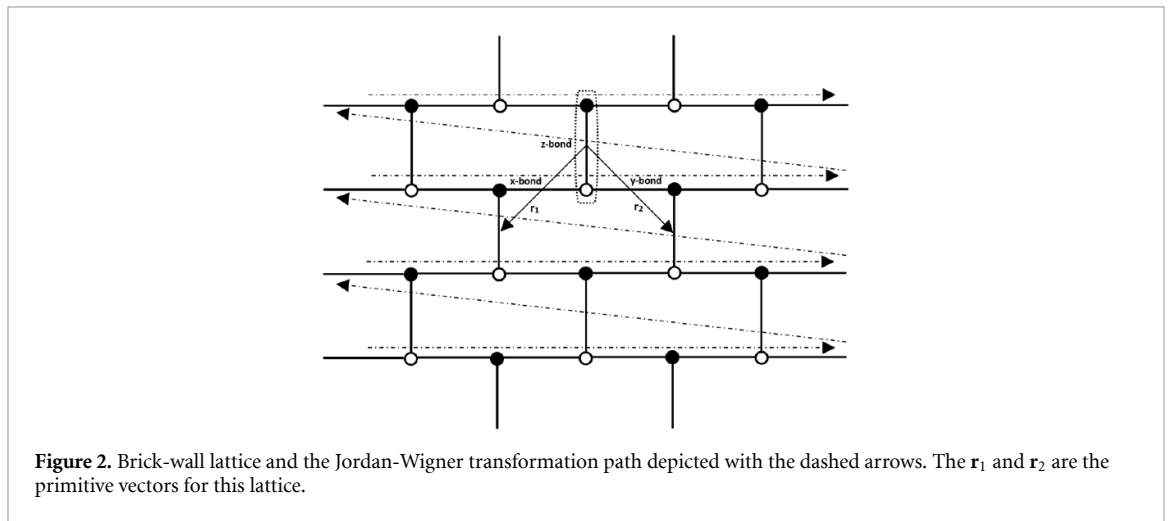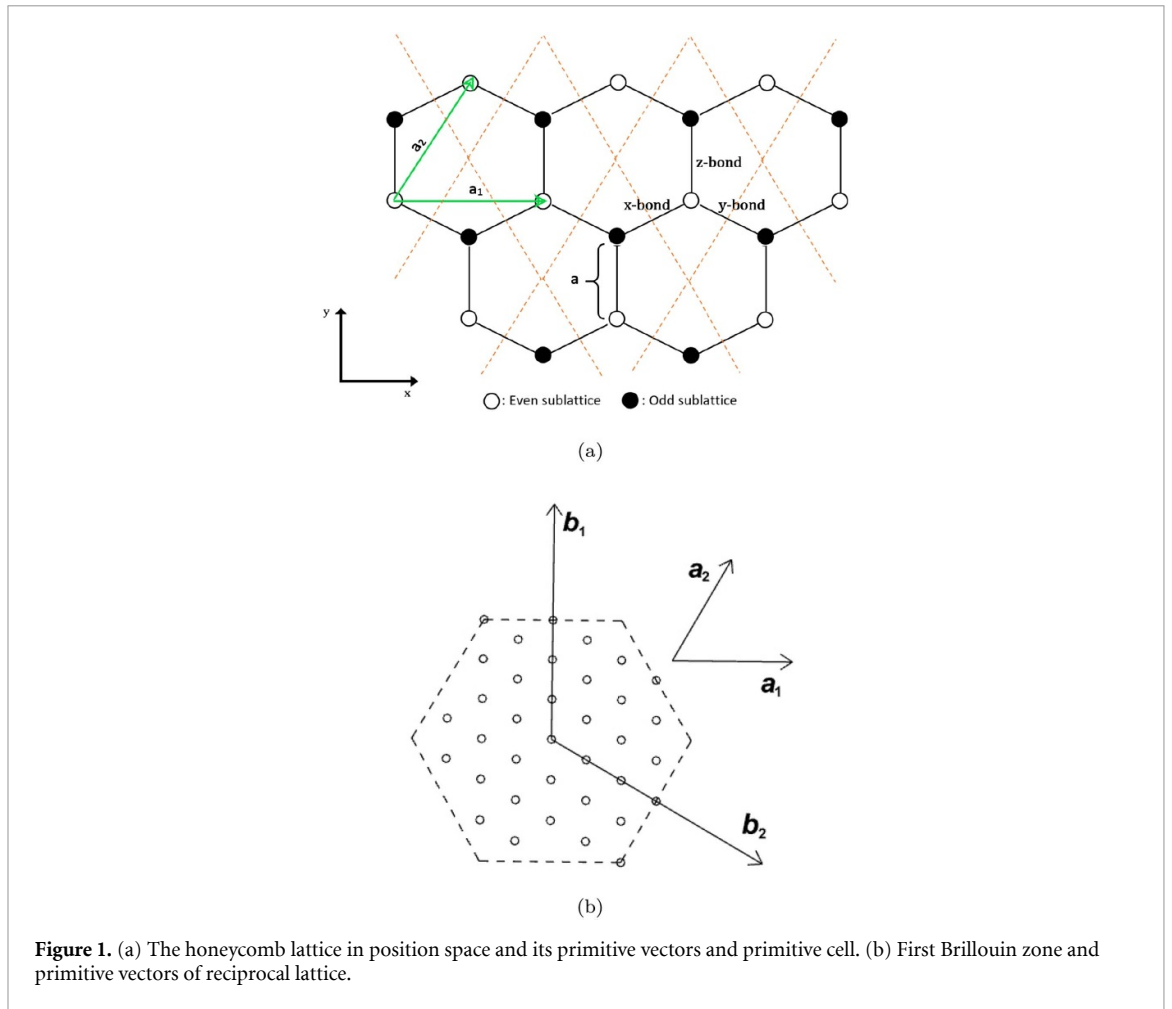
$$\mathbf{a}_1 = \sqrt{3}a\mathbf{e}_x \quad \& \quad \mathbf{a}_2 = \frac{\sqrt{3}}{2}a(\mathbf{e}_x, \sqrt{3}\mathbf{e}_y), \tag{1}$$

where a is the lattice constant. Basis vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ for the reciprocal lattice can be obtained by solving equation (2)

$$\mathbf{a}_i.\mathbf{b}_j = 2\pi\delta_{ij}. \tag{2}$$

The solution is as below

$$\mathbf{b}_1 = \frac{2\pi}{\sqrt{3}a}\left(\mathbf{e}_x - \frac{1}{\sqrt{3}}\mathbf{e}_y\right) \quad \& \quad \mathbf{b}_2 = \frac{4\pi}{3a}\mathbf{e}_y. \tag{3}$$

Figure 1. (a) The honeycomb lattice in position space and its primitive vectors and primitive cell. (b) First Brillouin zone and primitive vectors of reciprocal lattice.

Figure 2. Brick-wall lattice and the Jordan-Wigner transformation path depicted with the dashed arrows. The $\mathbf{r}_1$ and $\mathbf{r}_2$ are the primitive vectors for this lattice.

The reciprocal lattice is depicted in figure 1(b).

The Hamiltonian of the Kitaev model is a nearest-neighbor interaction of Pauli matrices on a honeycomb lattice as written in equation (4), where $r \& r'$ are indices for the nearest neighbor spins. The physics of the system is symmetric under permutation of coupling strengths J_α with $\alpha = x, y, z$ and due to an-isotropic interaction, the model is a frustrated spin system, because a spin cannot satisfy conflicting demands of orientation from its three neighboring sites [1]

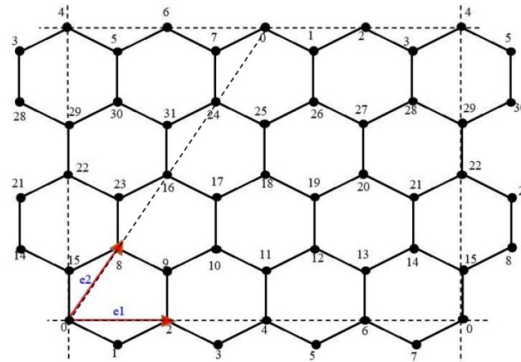$$H = -\sum_\alpha J_\alpha \sum_{\alpha-bonds} \sigma_r^\alpha \sigma_{r'}^\alpha. \tag{4}$$

Figure 3. Kitaev Honeycomb lattice model on a torus: the generators of the lattice in position space are named $\mathbf{e}_1$ and $\mathbf{e}_2$, as shown in the image above. Hence, the lattice points are $\mathbf{L} = n_1\mathbf{e}_1 + n_2\mathbf{e}_2$ with $n_1, n_2 \in \mathbb{Z}$ and the above lattice is drawn for $N_1 = N_2 = 4$ with toric boundary conditions.

Of central interest in this work is the property of the Kitaev model to support non-Abelian chiral spin liquid (CSL) states. A very symmetric way to realize this phase of the model is by setting $J_x = J_y = J_z = 1$ and turning on a small magnetic field which gives rise to a further interaction term in the Hamiltonian (see equation (47)). In the present work, we will restrict ourselves to this non-Abelian phase but set—with the exception of section 5.3—the magnetic field to zero. When CSL systems are placed on topologically non-trivial Riemann surfaces with non-zero genus, one obtains multiple degenerate groundstates that are separated from the rest of the spectrum by an energy gap. This allows to describe the low-energy physics of such models using topological quantum field theories (TQFTs) (see for example [5] for a review). The non-Abelian phase of the Kitaev model is in the universality class of the Ising TQFT. When placed on the torus, which corresponds to periodic boundary conditions of the lattice and is the choice employed in this work, this results in a triply-degenerate groundstate. These three groundstates correspond to the three different ways to obtain a torus by gluing the two ends of a cylinder in the Ising TQFT, namely by inserting the operators $\mathbb{1}$, σ and ψ at the two ends.

For the topological phase discussed above, the Kitaev model has recently been numerically solved in the spin basis to a high degree of approximation using tensor networks [6, 7]. In this paper we employ a different method based on the approach developed in [8] using restricted Boltzmann machines (RBMs) to find groundstates and excited states of the Honeycomb model. For a different application of RBMs to the Kitaev model, see [9]. The RBM as a variational ansatz to find groundstates of spin lattice models was first proposed in [10], where the formalism is applied to the 1d transverse-field Ising model and the 2d anti-ferromagnetic Heisenberg model. The authors of [11] use RBMs and feed-forward neural networks to study excited states of the 1d Heisenberg model and the 1d Bose–Hubbard model. Further applications of RBMs in the literature can be found in [12, 13, 13–16]. An application of RBMs to systems with non-Abelian symmetries was presented in [17], where the authors sample in the basis of irreducible representations instead of individual spins.

In this work, the central focus is on the RBM as a variational ansatz to find accurate groundstate wave-functions for the 3×3 Kitaev Honeycomb lattice as this is the smallest lattice size where the model admits topological order resulting in a triply-degenerate groundstate. Along the way, we also use the techniques developed to solve accurately for groundstates of smaller lattice sizes and explore the accuracy for larger lattices as well as the possibility to realize excited states. It is, however, not the goal of this paper to develop a machine learning architecture which solves efficiently and accurately for groundstates of large lattice systems. As explained in the main body of the paper, such an endeavor would need the simultaneous deployment of several distinct techniques employed and is left for future work.

Apart from a variational ansatz to find groundstates, another deep insight into chiral topological order is provided by the work of Moore and Read [18], where the authors find that many wave-functions can be expressed as conformal blocks. Conformal blocks are chiral correlation functions in certain conformal field theories (CFTs) which are also known to arise as states in the Hilbert space of the $2 + 1$-dimensional topological Chern–Simons theory [19]. This connection has been generalized to describe chiral topological states in lattice systems [20]. In our work, we explore the possibility to realize the three groundstates of the Kitaev model on the torus in terms of conformal blocks of the $SU(2)_2$ WZW CFT and show that our ansatz respects the symmetries of the Kitaev Hamiltonian (4). The consequences of the identification of the RBM states with the conformal black ansatz is then explored in the conclusions.

The outline of this paper is as follows. In section 2, we briefly review the analytical solution of the model. In section 3 the mapping of the spin model to an RBM architecture is explained. In particular, we demonstrate how to calculate the groundstate and excited states of the Honeycomb model using machine learning techniques. To this end we employ both the `NetKet` environment [21] as well our own self-built package using PyTorch to train a RBM in order to find a groundstate via gradient descent. In section 4 realization of vortices is discussed from both analytical and RBM points of view. In brief, a vortex pair can be created by adjusting the parameters of the RBM and subsequent braiding can be achieved by further adjustments of parameters. Section 5 is dedicated to results and relevant discussions and the reconstruction of a given quantum state from simple measurements known as quantum state tomography [22]. In section 6 we propose an identification of our RBM states with the Moore-Read states on the torus which can be computed using conformal blocks of the 2d Ising model. Finally we conclude in section 7.

2. Solution of the model

In this section an exact solution of the system is provided and an explicit form of the groundstate energy is derived where we will be following [1, 23, 24]. First of all, we fermionize the Hamiltonian by performing a one-dimensional Jordan-Wigner transformation [25] defined in equation (5). A deformation of the hexagonal lattice to a brick-wall lattice (see figure 2) clarifies the mechanism of the Jordan-Wigner transformation and why it is one-dimensional [26, 27]. In the brick-wall lattice, each lattice site r is denoted by the coordinates (i, j).

$$\sigma_{i,j}^{+} = 2[\prod_{j'<j}\prod_{i'}\sigma_{i',j'}^{z}][\prod_{i'<i}\sigma_{i',j}^{z}]a_{i,j}^{\dagger}$$

$$\sigma_{i,j}^{-} = 2[\prod_{j'<j}\prod_{i'}\sigma_{i',j'}^{z}][\prod_{i'<i}\sigma_{i',j}^{z}]a_{i,j}$$

$$\sigma_{i,j}^{z} = 2a_{i,j}^{\dagger}a_{i,j} - 1. \tag{5}$$

This transformation maps the Hilbert space of spins to the Hilbert space of spinless complex fermions. Using the fact that $\sigma^{\pm} = \sigma^x \pm i\sigma^y$ one can expand the Hamiltonian in equation (4) into three terms and rewrite them as below [24]

$$\begin{aligned}
\sigma_{i,j}^{x}\sigma_{i+1,j}^{x} &= \prod_{i'<i}\sigma_{i',j}^{z}(a_{i,j}^{\dagger}+a_{i,j})\prod_{i'<i+1}\sigma_{i',j}^{z}(a_{i+1,j}^{\dagger}+a_{i+1,j}) \\
&= (a_{i,j}^{\dagger}+a_{i,j})\sigma_{i,j}^{z}(a_{i+1,j}^{\dagger}+a_{i+1,j}) \\
&= -(a_{i,j}^{\dagger}-a_{i,j})(a_{i+1,j}^{\dagger}+a_{i+1,j}) \\
\sigma_{i,j}^{y}\sigma_{i+1,j}^{y} &= -\prod_{i'<i-1}\sigma_{i',j}^{z}(a_{i-1,j}^{\dagger}-a_{i-1,j})\prod_{i'<i}\sigma_{i',j}^{z}(a_{i,j}^{\dagger}-a_{i,j}) \\
&= -(a_{i-1,j}^{\dagger}-a_{i-1,j})\sigma_{i-1,j}^{z}(a_{i,j}^{\dagger}-a_{i,j}) \\
&= (a_{i-1,j}^{\dagger}+a_{i-1,j})(a_{i,j}^{\dagger}-a_{i,j}) \\
\sigma_{i,j}^{z}\sigma_{i,j+1}^{z} &= (2a_{i,j}^{\dagger}a_{i,j}-1)(2a_{i,j+1}^{\dagger}a_{i,j+1}-1). \tag{6}
\end{aligned}$$

Therefore, the Hamiltonian transforms to

$$\begin{aligned}
H = &+ J_x \sum_{x-links} (a_{i,j}^{\dagger}-a_{i,j})(a_{i+1,j}^{\dagger}+a_{i+1,j}) \\
&- J_y \sum_{y-links} (a_{i-1,j}^{\dagger}+a_{i-1,j})(a_{i,j}^{\dagger}-a_{i,j}) \\
&- J_z \sum_{z-links} (2a_{i,j}^{\dagger}a_{i,j}-1)(2a_{i,j+1}^{\dagger}a_{i,j+1}-1). \tag{7}
\end{aligned}$$

The J_x and J_y terms are quadratic interactions in spinless fermions and are easy to solve, however, the J_z term is a product of number density operators and can be further simplified by introducing the Majorana operators in equation (8) [24]. As we will see, this simplification can be done due to the presence of the conserved quantity called plaquette operator B_p [24].

$$c_{i,j} = i(a_{i,j}^{\dagger}-a_{i,j}), \quad d_{i,j} = a_{i,j}^{\dagger}+a_{i,j} \text{ for } i+j = \text{even} \equiv \circ$$

$$c_{i,j} = a_{i,j}^\dagger + a_{i,j}, \; d_{i,j} = i(a_{i,j}^\dagger - a_{i,j}) \text{ for } i+j = \text{odd} \equiv \bullet. \tag{8}$$

These operators have the following commutation relations

$$
\begin{aligned}
c_{i,j}^2 &= d_{i,j}^2 = 1 \\
\{c_{i,j}, c_{i',j'}\} &= \{d_{i,j}, d_{i',j'}\} = 2\delta_{ii'}\delta_{jj'} \\
\{c_{i,j}, d_{i',j'}\} &= 0.
\end{aligned}
\tag{9}
$$

Then, the J_z term can be rewritten using the Majorana operators

$$
\begin{aligned}
\sigma_{i,j}^z \sigma_{i,j+1}^z &= (2a_{i,j}^\dagger a_{i,j} - 1)(2a_{i,j+1}^\dagger a_{i,j+1} - 1) \\
&= i(id_{i,j+1}d_{i,j})c_{i,j+1}c_{i,j}.
\end{aligned}
\tag{10}
$$

Finally, using the circle indices for the odd and even lattice sites as defined in equation (8), the Hamiltonian transforms to the expression in equation (11).

$$
\begin{aligned}
H = &- iJ_x \sum_{x-links} c_\circ c_\bullet \\
&+ iJ_y \sum_{y-links} c_\bullet c_\circ \\
&- iJ_z \sum_{z-links} (id_\bullet d_\circ) c_\bullet c_\circ.
\end{aligned}
\tag{11}
$$

If we write the Hamiltonian as a sum over unit cells we have

$$
\begin{aligned}
H = i\sum_{\mathbf{r}} \Big[&J_x c_{\bullet,\mathbf{r}} c_{\circ,\mathbf{r}+\mathbf{r}_1} + J_y c_{\bullet,\mathbf{r}} c_{\circ,\mathbf{r}+\mathbf{r}_2} \\
&J_z (id_{\bullet,\mathbf{r}} d_{\circ,\mathbf{r}}) c_{\bullet,\mathbf{r}} c_{\circ,\mathbf{r}} \Big],
\end{aligned}
\tag{12}
$$

where $\mathbf{r}$ is the position vector of z-bonds or unit cells (see figure 2). It makes no difference in the physics if we set $\mathbf{r}$ to be any point along the z-bond, whether we choose an even site or an odd site or some point in the middle, they are all related through translation. What will be important in what follows, is that now since we have grouped a pair of even and odd sites as an unit cell, to evaluate the summation over the unit cells, it is sufficient to just run over the even or odd sites.

Furthermore, the $\alpha_{\mathbf{r}} = (id_{\bullet,\mathbf{r}} d_{\circ,\mathbf{r}})$ operators are defined on each z-bond of the lattice (labeled by $\mathbf{r}$) and they commute with the Hamiltonian and are good quantum numbers. Moreover, it can easily be shown that each plaquette operator can be written as

$$B_p = \sigma_1^y \sigma_2^z \sigma_3^x \sigma_4^y \sigma_5^z \sigma_6^x = \alpha_{61}\alpha_{43}. \tag{13}$$

On the other hand, from Lieb's theorem [28] we know that the groundstate manifold is obtained by setting $B_p = 1, \forall p$. Thus the uniform choice of $\alpha_{\mathbf{r}} = 1, \forall \mathbf{r}$ corresponds to a vortex-free sector, nevertheless all configurations leading to the same sector are equivalent.

We also introduce a Dirac fermion on each z-link using the Majorana operators

$$
\begin{aligned}
d_{\mathbf{r}} &= \frac{1}{2}(c_{\bullet,\mathbf{r}} - ic_{\circ,\mathbf{r}}) \\
d_{\mathbf{r}}^\dagger &= \frac{1}{2}(c_{\bullet,\mathbf{r}} + ic_{\circ,\mathbf{r}}).
\end{aligned}
\tag{14}
$$

Using the inverse transformation we can rewrite the Hamiltonian as

$$
\begin{aligned}
H = \sum_{\mathbf{r}} \Big[&J_x(d_{\mathbf{r}}^\dagger + d_{\mathbf{r}})(d_{\mathbf{r}+\mathbf{r}_1}^\dagger - d_{\mathbf{r}+\mathbf{r}_1}) \\
&+ J_y(d_{\mathbf{r}}^\dagger + d_{\mathbf{r}})(d_{\mathbf{r}+\mathbf{r}_2}^\dagger - d_{\mathbf{r}+\mathbf{r}_2}) \\
&+ J_z \alpha_{\mathbf{r}}(2d_{\mathbf{r}}^\dagger d_{\mathbf{r}} - 1) \Big].
\end{aligned}
\tag{15}
$$

The Hamiltonian is translation invariant and can be transformed to momentum space in order to be diagonalized. We define the Fourier transform for the Dirac fermion through

$$d_{\mathbf{r}} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{+i\mathbf{k}.\mathbf{r}} d_{\mathbf{k}}$$

$$d_{\mathbf{r}}^{\dagger} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}.\mathbf{r}} d_{\mathbf{k}}^{\dagger}. \tag{16}$$

Setting $\mathbf{k} \to \text{-}\mathbf{k}$ in $d_{\mathbf{r}}^{\dagger}$ simplifies the calculations. We then have

$$\mathbf{X} : \frac{J_x}{N} \sum_{\mathbf{r}} \sum_{\mathbf{k},\mathbf{k}'} e^{+i\mathbf{k}.\mathbf{r}} e^{+i\mathbf{k}'.(\mathbf{r}+\mathbf{r}_1)} (d_{-\mathbf{k}}^{\dagger} + d_{\mathbf{k}})(d_{-\mathbf{k}'}^{\dagger} - d_{\mathbf{k}'})$$

$$= J_x \sum_{\mathbf{k}} [-2\cos(k_1) d_{\mathbf{k}}^{\dagger} d_{\mathbf{k}} + i\sin(k_1)(d_{\mathbf{k}}^{\dagger} d_{-\mathbf{k}}^{\dagger} - h.c.)]$$

$$\mathbf{Y} : \frac{J_y}{N} \sum_{\mathbf{r}} \sum_{\mathbf{k},\mathbf{k}'} e^{+i\mathbf{k}.\mathbf{r}} e^{+i\mathbf{k}'.(\mathbf{r}+\mathbf{r}_2)} (d_{-\mathbf{k}}^{\dagger} + d_{\mathbf{k}})(d_{-\mathbf{k}'}^{\dagger} - d_{\mathbf{k}'})$$

$$= J_y \sum_{\mathbf{k}} [-2\cos(k_2) d_{\mathbf{k}}^{\dagger} d_{\mathbf{k}} + i\sin(k_2)(d_{\mathbf{k}}^{\dagger} d_{-\mathbf{k}}^{\dagger} - h.c.)]$$

$$\mathbf{Z} : \frac{J_z}{N} \sum_{\mathbf{r}} \sum_{\mathbf{k},\mathbf{k}'} e^{+i\mathbf{k}.\mathbf{r}} e^{+i\mathbf{k}'.\mathbf{r}} (2 d_{-\mathbf{k}}^{\dagger} d_{\mathbf{k}'}^{\dagger} - 1)$$

$$= J_z \sum_{\mathbf{k}} 2 d_{\mathbf{k}}^{\dagger} d_{\mathbf{k}} - J_z N. \tag{17}$$

For the above equation we used the orthogonality relation in Fourier transformation

$$\sum_{\mathbf{r}} e^{i(\mathbf{k}+\mathbf{k}').\mathbf{r}} = N\delta(\mathbf{k}+\mathbf{k}'). \tag{18}$$

Using all previous transformations we obtain a Hamiltonian which is quadratic in the Dirac fermions living on each of the unit cells along the z-links

$$H = \sum_{\mathbf{k}} [\epsilon_{\mathbf{k}} d_{\mathbf{k}}^{\dagger} d_{\mathbf{k}} + \frac{1}{2}(i\Delta_{\mathbf{k}} d_{\mathbf{k}}^{\dagger} d_{-\mathbf{k}}^{\dagger} - i\Delta_{\mathbf{k}} d_{-\mathbf{k}} d_{\mathbf{k}})] - J_z N$$

$$\epsilon_{\mathbf{k}} = 2[J_z - J_x\cos(k_1) - J_y\cos(k_2)]$$

$$\Delta_{\mathbf{k}} = 2[J_x\sin(k_1) + J_y\sin(k_2)]. \tag{19}$$

By applying a unitary Bogoliubov transformation, we diagonalize the Hamiltonian which can be written as

$$H = \sum_{\mathbf{k}} \frac{1}{2} \begin{pmatrix} \gamma_{\mathbf{k}}^{\dagger} & \gamma_{\text{-}\mathbf{k}} \end{pmatrix} \begin{pmatrix} E_{\mathbf{k}} & 0 \\ 0 & -E_{\mathbf{k}} \end{pmatrix} \begin{pmatrix} \gamma_{\mathbf{k}} \\ \gamma_{\text{-}\mathbf{k}}^{\dagger} \end{pmatrix}$$

$$= \sum_{\mathbf{k}} E_{\mathbf{k}}(\gamma_{\mathbf{k}}^{\dagger} \gamma_{\mathbf{k}} - \frac{1}{2}), \quad E_{\mathbf{k}} = \sqrt{\epsilon_{\mathbf{k}}^2 + \Delta_{\mathbf{k}}^2}, \tag{20}$$

where $\gamma_{\mathbf{k}}^{\dagger}$ and $\gamma_{\mathbf{k}}$ are quasiparticle creation and annihilation operators. The groundstate energy can easily be read off from the above Hamiltonian and is given by $E_{GS} = -\sum_{\mathbf{k}} \frac{E_{\mathbf{k}}}{2}$.

3. RBM representations

3.1. RBM as a neural network quantum state

With the ever growing applications of neural networks in sciences and the emergent new technologies to deploy and build physical neural networks, this is the right time to investigate potential applications of them in condensed matter systems. Recently, a new approach has been proposed for simulating topological quantum states using neural networks [8]. In this section we use the same approach and use a RBM as a variational ansatz for the Honeycomb Kitaev model. There are many reasons why an RBM is chosen. This particular architecture has proved to be effective in many tasks such as dimensional reduction, classification, regression, collaborative filtering, feature learning and topic modeling and moreover a general theorem shows that RBMs are universal approximators of discrete distributions [29].

Having a set of spins on a lattice, $\Xi = (\sigma_1, \sigma_2, \cdots \sigma_N)$ (in our case the Jordan-Wigner chain of spins), our goal is to use an RBM to reduce the dimensionality of the Hilbert space and estimate the energy of the system
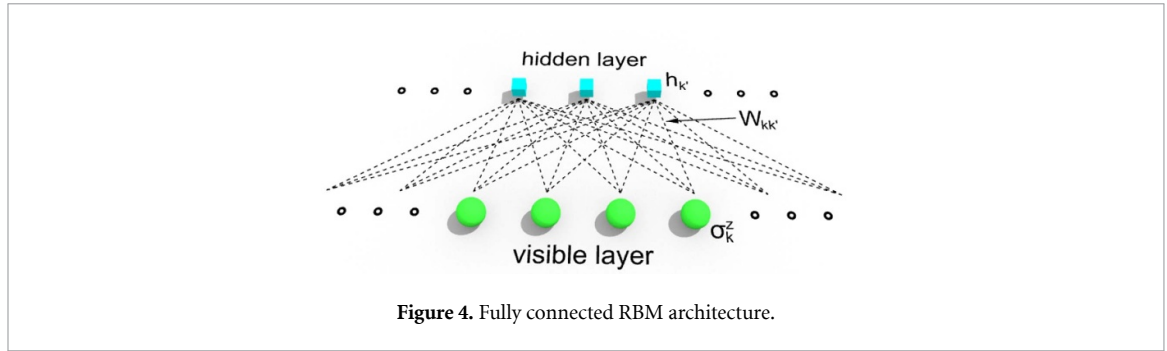
Figure 4. Fully connected RBM architecture.

in both groundstate and excited states to be able to classify different topological phases of the model. An RBM consists of two layers. The first layer (visible layer) has N nodes which are representing the physical spins in the Hamiltonian and the second layer (hidden layer) has M binary valued nodes (architecture of the network is shown in figure 4).

The quantum state of the Honeycomb model, up to an irrelevant normalization factor, can be written as

$$|\Phi\rangle = \sum_{\Xi} \Phi_M(\Xi; \Omega)|\Xi\rangle \tag{21}$$

where

$$\Phi_M(\Xi; \Omega) = \sum_{\{h_k\}} e^{\sum_k a_k \sigma_k^z + \sum_{k'} b_{k'} h_{k'} + \sum_{kk'} W_{kk'} h_k \sigma_{k'}^z}$$

$$= e^{\sum_k a_k \sigma_k^z} \times \prod_{k'} \cosh\left(\sum_k W_{kk'} \sigma_k^z + b_{k'}\right). \tag{22}$$

To obtain the second equality we used the values for $\{h_k\} = \{-1, 1\}^M$ which is the set of possible configurations of hidden layer nodes and $\Omega = (a_k, b_{k'}, W_{kk'})$ is the set of weights and biases of the RBM which should be trained in such a way that the final RBM state represents the desired quantum state of the model (i.e. groundstate or the excited states), where we refer to [21], sections 2.2.3 and 2.2.4, for further details of the variational principle. The combined number of weight and bias parameters and nodes is polynomial in system size and computationally feasible.

The variational calculations for the model with RBM as ansatz is done twofold. One approach we follow is to apply the machinery already developed in the NetKet software package [21], while the second approach we employ is to develop our own package using PyTorch. In the case of NetKet, in order to train the network, the parameter space of the network is sampled using the Metropolis algorithm and the optimization iterations are based on stochastic gradient descent. It turns out that although the sampling algorithm is effective, the ground state energy cannot be reached to a satisfactory precision. The reasons will be explained as we proceed. To circumvent this problem, we develop our own software package using PyTorch which is designed tightly to suit the purposes of the Honeycomb model and avoid possible pitfalls.

3.2. Quantum state tomography

Reconstruction of complex synthetic quantum states from experimental measurement data is computationally exponentially expensive. Therefore, the capability of neural network quantum states such as RBM to accurately and efficiently represent high-dimensional quantum states is beneficial to quantum state tomography. Apart from the numerous applications of QST, in our study case, the reconstructed states of the honeycomb model can later on be used for validation of particular quantum processes, e.g. braiding of vortices in the system, phase transitions, etc at the experimental level and after implementation of the framework developed in section 4. For instance, there are promising proposals for implementation of topological systems in cold atoms. A review has recently been published here [30].

In section 5.4 we present QST results for the groundstate of a small system size of the honeycomb model, as proof of concept. In a normal QST task, the data sets come from experimental measurements, but since we do not have experimental data, we need to first find the exact wave-function for the system through exact diagonalization methods and then sample that wave-function by performing single shot measurements in different bases to build the data set.
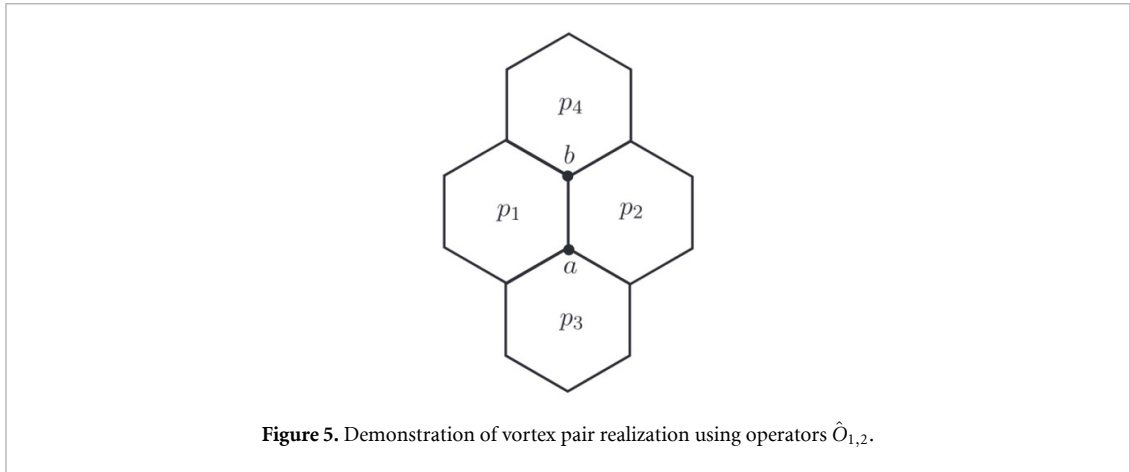
Figure 5. Demonstration of vortex pair realization using operators $\hat{O}_{1,2}$.

4. Braiding

In this section we want to explore the possibility to create anyonic excitations in the RBM representation and describe anyon braiding. Such a representation is important for various reasons, in particular for performing quantum computation [5].

First of all, let us describe the creation of these quasi-particles by starting with a groundstate sector and then subsequently changing the eigenvalue of the B_p operators from $+1$ to -1 locally and in the region where we want to realize the vortices in the system. In an arbitrary configuration of spins (not necessarily in the groundstate sector) it is easy to prove that $\prod B_p = +1, \forall p$. Hence, we can just build the vortices in pairs. For example, if you apply the operator $\hat{O}_1$ to the groundstate, it produces two vortices in plaquettes 1 and 2 (see figure 5) [31].

$$\hat{O}_1 = \exp\left(-i\frac{\pi}{2}\hat{\sigma}_a^z\right). \tag{23}$$

Another possibility is to apply the following operator

$$\hat{O}_2 = \exp\left(-i\frac{\pi}{2}\hat{\sigma}_a^x\right)\exp\left(-i\frac{\pi}{2}\hat{\sigma}_b^y\right), \tag{24}$$

which produces two vortices along the $\hat{z}$ direction in plaquettes 3 and 4.

Let us explain how these operators create vortices from the viewpoint of Kitaev's original Majorana fermion representation [1]. We assume that there are two fermionic modes living on each lattice site corresponding to the four creation and annihilation operators $a_{m,i}^\dagger$ and $a_{m,i}$, where $m \in \{1,2\}$ is the index for the modes and i is the index for the lattice sites. Here unlike before, we use one index for lattice site to avoid unnecessary complications. Then we will separate the imaginary and real parts of these operators, to define the Majorana fermions as below

$$\begin{aligned}
c_i &= a_{1,i} + a_{1,i}^\dagger \\
b_i^x &= i(a_{1,i}^\dagger - a_{1,i}) \\
b_i^y &= a_{2,i} + a_{2,i}^\dagger \\
b_i^z &= i(a_{2,i}^\dagger - a_{2,i}).
\end{aligned} \tag{25}$$

Notice that the spins have a two dimensional space (being up or down) and now that we are representing them with four Majorana modes (two complex fermionic modes), we need to project out the unphysical states. The Fock space of the complex fermionic modes can be represented as $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. We make the following correspondence between the states of spins and fermions

$$\begin{aligned}
|\uparrow\rangle &= |00\rangle \\
|\downarrow\rangle &= |11\rangle.
\end{aligned} \tag{26}$$

One can define the projector P_i on site i [32] to do this job for us

$$P_i = \frac{1 + D_i}{2}$$

$$D_i = (1 - 2a_{1,i}^\dagger a_{1,i})(1 - 2a_{2,i}^\dagger a_{2,i})$$
$$= b_i^x b_i^y b_i^z c_i. \tag{27}$$

One can easily show that the relation between Majorana operators and the original Pauli operators is as follows

$$\sigma_i^\alpha = i b_i^\alpha c_i \quad \text{for} \quad \alpha \in \{x, y, z\}. \tag{28}$$

In fact, by the above projection, we satisfied the extra condition coming from the algebra of Pauli matrices, $-i\sigma_i^x \sigma_i^y \sigma_i^z = b_i^x b_i^y b_i^z c_i = \mathbb{1}$. Therefore, the eigenvalue of the projector P_i for physical states is 1 and for unphysical states is 0.

Using equation (28) the Hamiltonian can be written in terms of Majorana operators as

$$H = \frac{i}{2} \sum_{i,j} A_{ij} c_i c_j$$
$$A_{ij} = J_{ij} u_{ij} \,, \; u_{ij} = i b_i^\alpha b_j^\alpha \text{ with } \alpha \in \{x, y, z\}. \tag{29}$$

The u_{ij} are antisymmetric Hermitian link operators with eigenvalues ± 1

$$u_{ij} = -u_{ji}, \; u_{ij}^2 = 1, \; , u_{ij}^\dagger = u_{ij}. \tag{30}$$

The link operators commute with the Hamiltonian, $[H, u_{ij}] = 0$, so they are local symmetries. In this form, the Hamiltonian is representing a tight binding model of free Majorana fermions hopping along the lattice, with tunneling couplings that depend on the eigenvalues of link operators which can be thought of as a classical $\mathbb{Z}_2$ gauge field. One can assign a configuration $\{u\}$ to all link operators and diagonalize the resulting quadratic Hamiltonian in Majorana operators directly and obtain the spectrum and groundstate energy for $H\{u\}$. But which configuration $\{u\}$ corresponds to the global minimum of the energy? This question has been answered by Lieb [28]. The groundstate lies in the sector with $B_p = 1 \; \forall \, p$ and since the B_p operators are the only gauge invariant objects, every gauge fixing choice $\{u\}$ which leads to this particular configuration for B_p is equivalent. Therefore, our goal is to find a configuration $\{u\}$ which leads to $B_p = 1 \; \forall \, p$.

Next, we define the vortex operator $\hat{W}_p$ as

$$\hat{W}_p = \prod_{(i,j) \in \partial P} u_{ij} \text{ where } \begin{cases} i \in \text{ even sublattice} \\ j \in \text{ odd sublattice} \end{cases}. \tag{31}$$
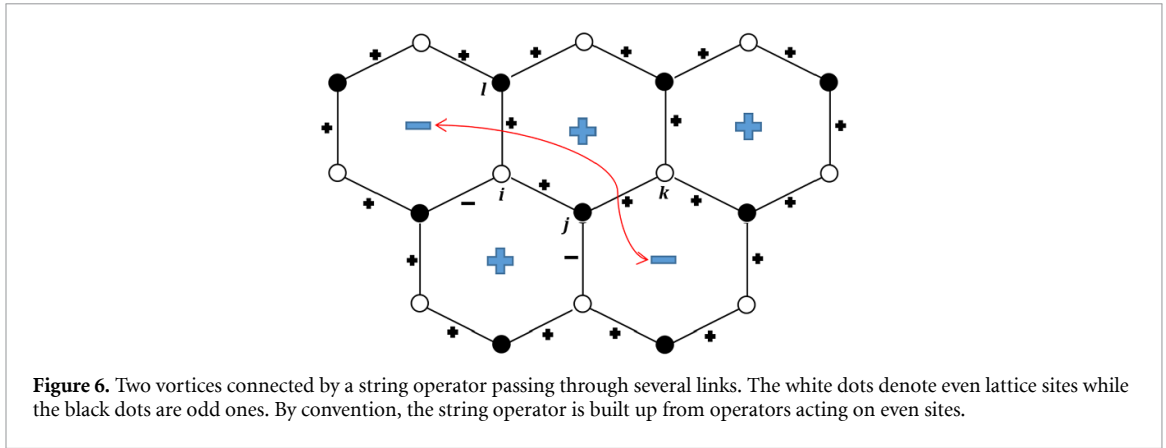
The vortex operator can be simplified as below

$$\hat{W}_p = u_{12}^x u_{32}^y u_{34}^z u_{54}^x u_{56}^y u_{16}^z$$
$$= -(i)^6 c_1^x c_2^x c_2^y c_3^y c_3^z c_4^z c_4^x c_5^x c_5^y c_6^y c_6^z c_1^z$$
$$= \sigma_1^y D_1 \; \sigma_2^z D_2 \; \sigma_3^x D_3 \; \sigma_4^y D_4 \; \sigma_5^z D_5 \; \sigma_6^x D_6, \tag{32}$$

where we used $ic_j^x c_j^y = -\sigma_j^z D_j$ and cyclic permutations. Considering the fact that $[\hat{W}_p, H] = [\hat{W}_p, D_i] = 0$, in the physical subspace where $D_j = 1$ we recover $\hat{W}_p = B_p$. Therefore, the uniform choice of $u_{ij} = 1$ for every link will lead us to the groundstate in the extended (not yet projected) space with $B_p = 1$. In order to find the physical groundstate, we apply the global projector **D** to project out the unphysical states

$$|\psi_w\rangle_{physical} = \mathbf{D} |\psi_u\rangle_{extended}, \quad \text{where} \quad \mathbf{D} = \prod_{i=1}^{N} P_i. \tag{33}$$

Now we are at a point to explain in detail why the operators $\hat{O}_1$ and $\hat{O}_2$ can create vortices in the system. As mentioned earlier, if we flip the sign of the eigenvalue of the vortex operator at p, a vortex will be created in the system. Looking back at equation (31), we see that the vortex configuration $\{\hat{W}_p\}$ is created by fixing the gauge at every link $\{u_{ij}\}$. Therefore, by changing the link operators, we can move between vortex sectors. Starting from a vortex-free sector with all link operators being $+1$, applying $\hat{O}_1$ on site a will leave the σ_a^z untouched while flipping the σ_a^x and σ_a^y. Consequently, this will change the sign of the eigenvalue of the u_{ij} for x-link and y-link attached to site a, leaving the system with $B_{p_1} = -1$ and $B_{p_2} = -1$. We can also move the vortices around in the system, by consecutively applying these operators on the particular sites in a path for

Figure 6. Two vortices connected by a string operator passing through several links. The white dots denote even lattice sites while the black dots are odd ones. By convention, the string operator is built up from operators acting on even sites.

which we want to move the vortices along. As an example, the red path shown in the figure 6 corresponds to the operator

$$\hat{O} = \exp\left(-i\frac{\pi}{2}\sigma_i^z\right)\exp\left(-i\frac{\pi}{2}\sigma_j^x\right). \tag{34}$$

In general, we can define string operators to move the vortices anywhere in the system along the string, similar to the example above. To create such a string operator is straightforward now. If the string passes through a link, we need to apply an operator on one of the sites connected to that $\alpha - link$ which carries σ^α in the exponent. Before continuing, we need to make a convention. A sharp reader might point out that applying the operator $\hat{O}' = \exp\left(-i\frac{\pi}{2}\sigma_l^z\right)\exp\left(-i\frac{\pi}{2}\sigma_k^x\right)$ instead of $\hat{O}$ creates the same pair of vortices in the system. For consistency reasons, we therefore use the convention that if a string is passing through a link, the corresponding string operator is built up by picking the even site on the link. Having said all of this, a generic string operator is defined as

$$\hat{S} = \prod_{\alpha-links} \exp\left(-i\frac{\pi}{2}\hat{\sigma}_\circ^\alpha\right). \tag{35}$$

This string operator will create two vortices at the ending points of the string which passes through the $\alpha - links$. The manipulation of link operators using these string operators is practically equivalent to changing the sign of the couplings J_{ij} corresponding to the link u_{ij} in the definition of the A_{ij} (see equation (29)). Therefore, in the RBM language one can simply redefine the Hamiltonian with the desired signs of J_{ij} for a particular vortex sector and then try to find the RBM representation for this new Hamiltonian through the approach explained in section 3. The only practical disadvantage of this vortex realization approach is that it is computationally expensive as the RBM should be trained again from scratch to find the representation for the excited state. An ideal approach is one where, given the representation for the groundstate, we would be able to directly and without further optimization steps find the representation for the excited state.

In principle, it is possible to directly manipulate the weight and bias parameters of the RBM, $\Omega = (a_k, b_{k'}, W_{kk'})$ to realize the vortices and find the excited state representation. Looking at equation (22), we notice that there is a symmetry between the eigenvalues of physical spins and the value of parameters of the RBM. In other words, flipping a particular spin in site k is equivalent to flipping the sign of the RBM parameter sitting next to the σ_k^α in equation (22) and/or adding a phase to it. In general, for a string operator $\hat{S}_l^\alpha = \exp\left(-i\frac{\pi}{2}\hat{\sigma}_l^\alpha\right)$ we have

$$\hat{S}|\Phi\rangle = (-i\hat{\sigma}_l^\alpha)\sum_\Xi \Phi_M(\Xi;\Omega)|\Xi\rangle$$

$$= \sum_\Xi e^{[(-1)^{\delta(\alpha-z)-1}a_l]\sigma_l^z + \sum_{k\neq l} a_k\sigma_k^z} \times \prod_{k'} \cosh\left((-1)^{\delta(\alpha-z)-1}W_{lk'}\sigma_l^z + \sum_{k\neq l} W_{kk'}\sigma_k^z + b_{k'}\right)(-i\hat{\sigma}_l^\alpha)|\Xi\rangle. \tag{36}$$

Next, one needs to evaluate how $(-i\hat{\sigma}_a^\alpha)$ acts on $|\Xi\rangle$. Here we subdivide the three cases $\alpha = z$, x and y. If $\alpha = z$, spin up states will get a phase $-i$ and spin down states a phase i. The overall effect of this can be absorbed in a shift of a_l by $-i\frac{\pi}{2}$. If $\alpha = x$, then spins up and down are just exchanged and the $-i$ is an overall

phase which can be ignored. If $\alpha = \gamma$, we get a combination of the previous two cases. In equation form, the above explanations take the following form

$$
\begin{aligned}
\hat{S}_l^z &: (W_{lk'}, a_l) \mapsto \left(W_{lk'}, a_l - i\frac{\pi}{2} \right) \\
\hat{S}_l^x &: (W_{lk'}, a_l) \mapsto (-W_{lk'}, -a_l) \\
\hat{S}_l^y &: (W_{lk'}, a_l) \mapsto \left(-W_{lk'}, -a_l - i\frac{\pi}{2} \right).
\end{aligned}
\tag{37}
$$

5. Results and discussion

5.1. Groundstate
In this section, we present the results for the groundstate energy estimate of different system sizes for the proposed mapping discussed in section 3. These results are also compared to the exact analytical expression given in equation (20). The absolute value of the energies is plotted in figure 11 in appendix C. All results for this and the following sections are for the phase $J_x = J_y = J_z = 1$.

We used both `NetKet` and PyTorch to build our RBM. While `NetKet` provides build-in functions to evaluate physical observables, we build our own version of these functions in the PyTorch environment. When training with `NetKet`, the configuration space of the network is sampled for hundreds of times, depending on the lattice size, using the Metropolis algorithm. While when training with PyTorch, the configuration space of the network is sampled from all of the Hilbert space. This is easier to implement and enables us to focus on the explicit RBM representation. The loss function of the network has been minimized using Stochastic Gradient Descent and the learning rates are set to decay using the adaptive moment estimation algorithm. The groundstate energies are obtained by fine-tuning the optimization algorithm parameters.

The numerical results obtained are shown in figure 11 in appendix C. The different methods for computing the groundstate energy listed in the table 1 are as follows:

(a) Analytic formula equation (20) using
   (i) anti-periodic boundary conditions
   (ii) periodic boundary conditions.
(b) The method used by Kitaev and Pachos [2, 23].
(c) Direct diagonalization for small system sizes, using `NetKet`.
(d) RBM results for an architecture with $\alpha = \frac{\text{num}_h}{\text{num}_v} = 2$ and 1k samples and 10k iteration steps, using `NetKet`.
(e) RBM results built using PyTorch with sampling all of the Hilbert space and using projection onto translation invariant states as described below.

Based on the results obtained for the energy of the groundstate of the system, training of the network is tested and optimized by changing the learning rate for different training phases and iteration numbers. The most effective initial learning rate is found to be $\text{lr}_{init} \approx 0.01$. This value is small enough to avoid over-fitting and large enough to train the network effectively. This learning rate is rather large compared to standard machine learning implementations, because the number of parameters in our RBM is relatively small compared to hundreds of millions of parameters used in typical deep learning applications.

A central ingredient of our approach in PyTorch is to use *projection* onto translation invariant states in order to restrict RBM parameters to the physical parameter space. This dramatically improves the results we can achieve. The theoretical reason for this projection is that ground states of our Hamiltonian reside in such translation invariant subspaces of the Hilbert space as explained in section 6. Here we provide the technical details of the implementation. Suppose $|\Phi\rangle$ is the RBM state mentioned in equation (21). Then the state used in calculating the energy is as follows

$$
|\Phi'\rangle = \sum_{m,n} \hat{T}_{m\mathbf{a}_1 + n\mathbf{a}_2} |\Phi\rangle,
\tag{38}
$$

where $\hat{T}_{m\mathbf{a}_1 + n\mathbf{a}_2}$ are operators translating m steps along the $\mathbf{a}_1$ direction and n steps along the $\mathbf{a}_2$ direction, with $\mathbf{a}_1, \mathbf{a}_2$ defined in equation (1). $\hat{T}_{m\mathbf{a}_1 + n\mathbf{a}_2}$ can be realized either as a permutation matrix acting on quantum states or a permutation of parameters in RBM.

The hyperbolic trigonometric functions appearing in the weights of the RBM are numerically unstable. They sometimes lead to divergence and we have to use an extremely small learning rate to avoid the

Table 1. Groundstate energy for different system sizes obtained via different methods explained in the text.

Lattice size	2 × 2	2 × 3	2 × 4	3 × 3	3 × 4	4 × 4	5 × 5	6 × 6	7 × 7	3 × 3 with $K = 0.2$
a.i	−6.4721[1]	−9.8003	−12.5851[1]	−14.2915	−18.9869	−25.1282	−39.3892	−56.7529	−77.1249	
a.ii	−6	−9.2915	−12.4721	−14.2915	−19.0918	−25.4164	−39.3892	−56.2668	−77.1249	
b	−6	−9.2915	−12.4721	−14.2915	−19.0918	−25.4164	−39.3892	−56.2668	−77.1249	
c	−6.9282	−9.8003	−12.9443	−14.2915	−19.0918	—	—	—	—	−17.5260
d	−6.9282	−9.8003	−12.573	−13.7374	−18.0466	−24.0783	−37.305	−52.920	−71.793	
e	−6.9282	−9.8003	−12.9293	−14.2787						−17.393

[1] Notice that the energies obtained via the analytical formula do not match the exact diagonalization results for 2 × 2 and 2 × 4 lattices. This is not a problem because the energy obtained via the analytical formula (method a.i) is still in the spectrum, which means this is a case where Lieb's theorem fails—the vortex free sector no longer contains the ground state. This is an edge case and will not appear when $m, n > 2$.

divergence. To cure this instability, we use the following *pre-train* algorithm. We first perform the substitution

$$\cosh x \mapsto 1 + \frac{x^2}{2}, \ \sinh x \mapsto x, \tag{39}$$

as the polynomials x and x^2 are more stable under numerical changes than the exponential e^x which appears in the hyperbolic trigonometric functions. Then this modified model is trained for several thousand epochs. After we get close enough to the minimum, we change back to the standard RBM representation using cosh and sinh to refine our result. We find that this two-step training shortens the time to reach a given accuracy.

Remember that RBM parameters are complex numbers. However, the meaning of the real part and the imaginary part of parameters are different. The real part contributes by cosh / sinh while imaginary parts show up as arguments of cos / sin, i.e. they are phases. So they should be initialized with different standard deviations in the sampling distribution and trained with different learning rates. In other words they are 'renormalized' differently along the flow.

On the other hand, when working with NetKet, we find that the ground state energy cannot be reached beyond a certain accuracy. Thus the results obtained with our PyTorch algorithm described above are always far more accurate than the best results NetKet can achieve. We believe this is due to the following identity

$$\Pi_i x_i = e^{\sum_i \log x_i}, \tag{40}$$

used by NetKet, in order to accelerate sequential multiplication. This is theoretically correct but numerically unstable as in an RBM representation there are always extremely large or small values of cosh / sinh (this is the means by which an RBM is capable of approximating the correct physical state!) leading to truncation errors when using the functions log and exp. So it is not surprising that NetKet cannot reach a high enough accuracy and gets stuck at a too high energy. We find that when using the identity (40) in our PyTorch implementation, our ground state energy becomes similarly inaccurate as the one obtained by NetKet.

In order to check the effect of the relative number of hidden and visible nodes on the outcome of the network, different values for α were tested, keeping other parameters fixed. The α parameter was changed just by changing the number of hidden nodes for a fixed system size (and hence fixed visible layer size). The results (for a 5×5 lattice) are summarized in the table 4 in appendix C.

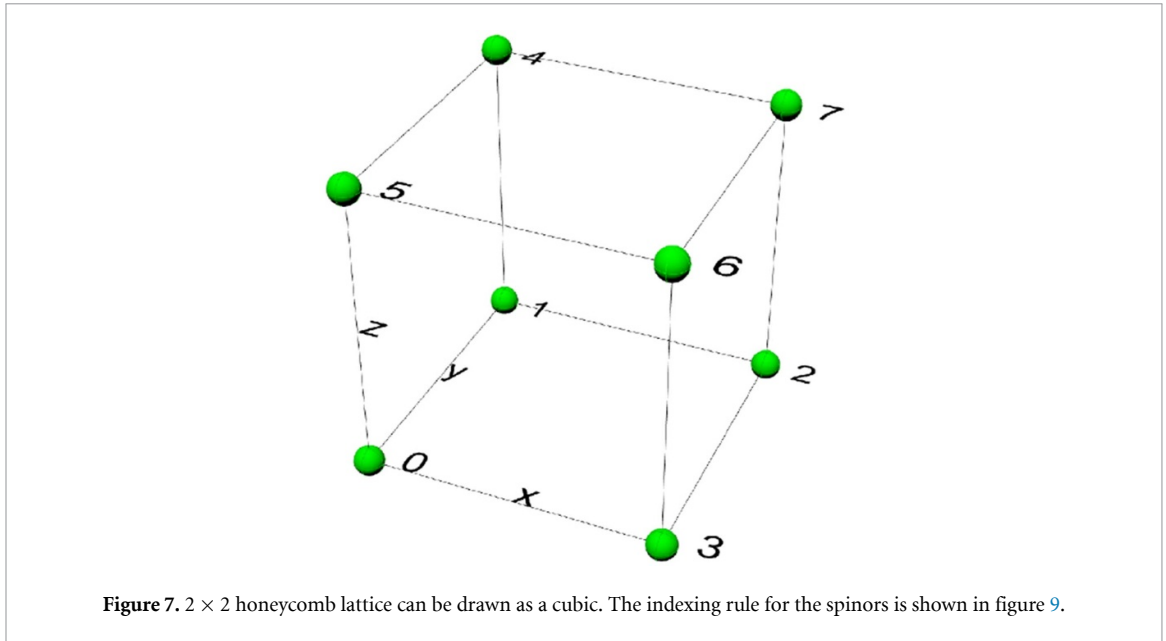### 5.1.1. - The performance of FFNN vs. no-visible-bias RBM
In table 5 in appendix C, we give a brief summary of the comparison results between an FFNN with one layer and a no-visible-bias RBM (referred to as novb-RBM in the table). The two architectures are similar and for the same computational resources (such as CPU and memory) and system size these results have been obtained. The conclusions we reach are:

(a) The FFNN trains faster than the RBM with relatively fewer free parameters and performs better than the RBM with no bias.
(b) But the FFNN's result deviates by a greater margin from the analytic result.

5.1.2. RBM's parameters: an example
To get a better feeling for how the RBM approximates quantum states so well, let us have a closer look at an RBM for a small system size, as shown in equation (41). These are the parameters for an RBM with 5 hidden nodes for a 2×2 system which has 8 spinors. So the parameter matrix $W_{kk'}$ is 5×8 in size. The indexing rule for the spinors is shown in figure 9. The biases a_k and $b_{k'}$ in RBM are set to 0 manually. For illustration purposes, this RBM is trained without the *projection method*. After training, the energy of the resulting parameter set is -6.928201 while the true ground state energy is -6.92820323. Its wave-function has 99.999982% overlap with the groundstate.

$$\text{Re}(W) = \begin{bmatrix} 0.3399 & 0.3517 & 0.3000 & -0.2862 & 0.2862 & 0.3000 & -0.3516 & 0.3398 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ -0.3687 & 0.3906 & -0.6170 & -0.5241 & 0.5241 & -0.6170 & -0.3906 & -0.3687 \\ 0.3236 & -0.3456 & 0.5754 & 0.4805 & 0.4805 & -0.5754 & -0.3456 & -0.3236 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}$$

$$\text{Im}(W) = \begin{bmatrix} 0.7658 & 0.7894 & -0.7439 & -0.7592 & -0.0264 & 0.0412 & -0.0037 & -0.0199 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & -0.7854 & -0.7854 & 0.7854 & 0.7854 \\ 0.2733 & -0.4073 & 0.4872 & 0.1436 & 0.6418 & -0.2982 & -0.3781 & -0.5121 \\ 0.2612 & 1.0566 & -0.7134 & -0.0218 & -0.8072 & -0.0720 & 0.2712 & 0.5242 \\ 0.7854 & 0.7854 & -0.7854 & -0.7854 & -0.7854 & -0.7854 & 0.7854 & 0.7854 \end{bmatrix}. \tag{41}$$

Figure 7. 2×2 honeycomb lattice can be drawn as a cubic. The indexing rule for the spinors is shown in figure 9.

To begin with, notice that, for example, for a state with all spins up, the RBM amplitude is given by

$$\Phi_M(|\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow\rangle) = \prod_{k'=1}^{5} \cosh\left(\sum_{k=1}^{8} W_{kk'}\right), \tag{42}$$

where we are using the convention that spin up corresponds to the value $\sigma^z = 1$ and spin down to $\sigma^z = -1$. If we focus on the factor corresponding to $k' = 5$ and use the identification $\pi/4 \approx 0.7854$, we see from equation (41) that its amplitude is given by

$$\cos\left(\frac{\pi}{4} + \frac{\pi}{4} - \frac{\pi}{4} - \frac{\pi}{4} - \frac{\pi}{4} - \frac{\pi}{4} + \frac{\pi}{4} + \frac{\pi}{4}\right) = \cos(0) = 1, \tag{43}$$

where we have used that

$$\cosh(x + iy) = \cosh(x)\cos(y) + i\sinh(x)\sin(y). \tag{44}$$

On the other hand, if we flip one of the spins, the corresponding amplitude becomes

$$\begin{aligned}
&\Phi_M(|\uparrow\uparrow\uparrow\uparrow\downarrow\uparrow\uparrow\uparrow\rangle) \\
=\ &\cos\left(\frac{\pi}{4} + \frac{\pi}{4} - \frac{\pi}{4} - \frac{\pi}{4} + \frac{\pi}{4} - \frac{\pi}{4} + \frac{\pi}{4} + \frac{\pi}{4}\right) \\
=\ &\cos\left(\frac{\pi}{2}\right) = 0.
\end{aligned} \tag{45}$$

That means the second and fifth line of $\Im(W)$ impose strong restrictions on spinor configurations and in this particular case imply that up and down spins are paired in even numbers. This restriction also happens in the first, third and fourth line of $\Im(W)$, although this is not apparent from first sight. This can be seen from observing that

$$\begin{aligned}
\Im(W)[0,0] - \Im(W)[0,7] &= 0.7857 \approx \pi/4\,, \\
\Im(W)[0,3] - \Im(W)[0,4] &= -0.7856 \approx -\pi/4.
\end{aligned} \tag{46}$$

Furthermore, there are many numerical coincidences for parameters between spinor pairs $(0,7)$ $(1,6)$ $(2,5)$ and $(3,4)$. This reflects the 'parity' symmetry of the 2×2 lattice. More specifically, the 2×2 lattice can be drawn as a cubic as shown in figure 7. Then it is apparent that there are strong correlations between the pairs $(0,7)$ $(1,6)$ $(2,5)$ and $(3,4)$.

This small example shows that an RBM is able to find restrictions and symmetries of our quantum system efficiently. We can imagine that this will also happen for larger systems, although it is hard to observe such symmetries explicitly since the number of parameters becomes too large.

Table 2. Energy of excited states with a pair of vortices in a 3×3 lattice system, measured through different methods as explained in the text.

Spins flipped	Energy (Method A)	Energy (Method B)	Energy (Method C)
σ_8^y	$-11.7335(5)$	$-12.1070(51)$	-13.9146
σ_9^y	$-11.7474(5)$	$-11.8873(57)$	-13.9146
$\sigma_8^x \sigma_8^z$	Unrealizable	$-12.7447(37)$	-13.9146

5.2. Creating Vortex pairs and measuring their energy

As described in section 4, there are three theoretically equivalent methods to create vortices:

(a) modifying parameters of RBM;
(b) creating an auxiliary Hamiltonian and training it again;
(c) transforming into auxiliary fermion representation and fixing u_{ij}s [2, 23].

We applied all three methods to a 3×3 lattice and computed the corresponding energy levels for excited states. The results we get are shown in table 2.

We find that method (a) is more stable than method (b), which might be due to the random initialization of RBM parameters in method (b) for the training process. Both methods, (a) and (b), deviate by a margin from the analytic values, namely method (c). This can be explained by the fact that the state generated by the spin flip operation is not the groundstate of the vortex sector but rather an excited state carrying quasi-particles. We will comment on this further in the next sub-sections about the expectation values of plaquette operators.

5.3. Perturbation with non-trivial magnetic field

As originally discussed by Kitaev [1], one can add the following perturbation to the Hamiltonian (4)

$$\Delta H = -K \sum_{(i,j,k)} \sigma_i^x \sigma_j^y \sigma_k^z, \tag{47}$$

which corresponds to an effective magnetic field with coupling K, and the sum is over all vertices with links (i,j,k). This term leads to an energy gap in the spectrum for asymptotically large lattices and renders the vortex sectors of the model topological. In this phase, the vortices of the honeycomb lattice bind localized Majorana modes and behave as Ising anyons. For the 3×3 lattice, we find that the gap in the spectrum separating the groundstates from excited states becomes larger for non-zero K. For $K = 0.2$, exact diagonalization gives a GS energy of -17.5260. We find that using our RBM approach, with the same number of parameters and training procedures discussed above, we get an energy value of -17.393 (0.75% deviation) which is quite close to the analytical value, as shown in table 1. This shows that the RBM can be an efficient and accurate representation for the non-trivial topological phases of the Kitaev honeycomb model. As in this work we are mainly interested in the groundstate, which already displays topological order for the 3×3 lattice, we leave further work on the phase with non-trivial K for the future.

5.4. Quantum state tomography for groundstate

In this section we apply the technique discussed in section 3.2 to reconstruct the wave function from measurements. In order to generate measurement data set, we sample from an exact groundstate by random measurements in different basis. To this end, we focus on the 2×2 lattice in order to deal with the computational resource requests and also being able to find the exact groundstate wave function through exact diagonalization. The last step is to train the RBM using the data set obtained. The result is shown in figure 8.

There are many intricate parameters which one can adjust during training, such as learning rate, batch size, number of Metropolis samples, and number of single measurements. Only with a suitable choice of parameters and large enough training data set, can one achieve a satisfactory result. For the example depicted in figure 8, in order to build the training data set, the exact wave function was sampled 2000 times for 256 randomly chosen basis x, y, z and the hidden node density α was set equal to 2. With this combinations of parameters, QST leads to an overlap of 97% between the exact wave function and the learned one.

5.4.1. Plaquette Operator expectation values and spin flipping
Starting from the state obtained from tomography as discussed above, one can then realize vortices in the system by modifying parameters of the RBM as described in section 4. In order to verify that modification of
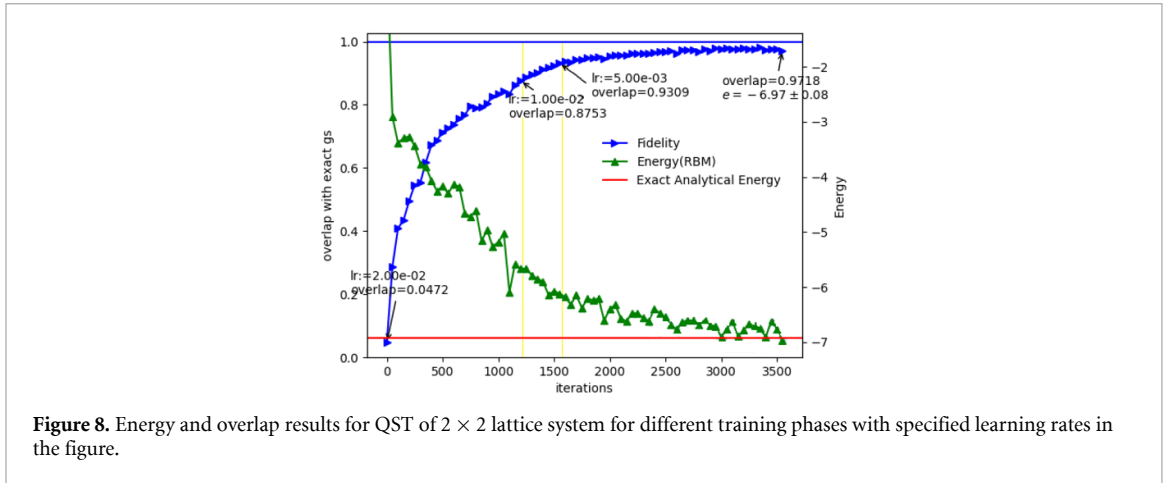
Figure 8. Energy and overlap results for QST of 2×2 lattice system for different training phases with specified learning rates in the figure.
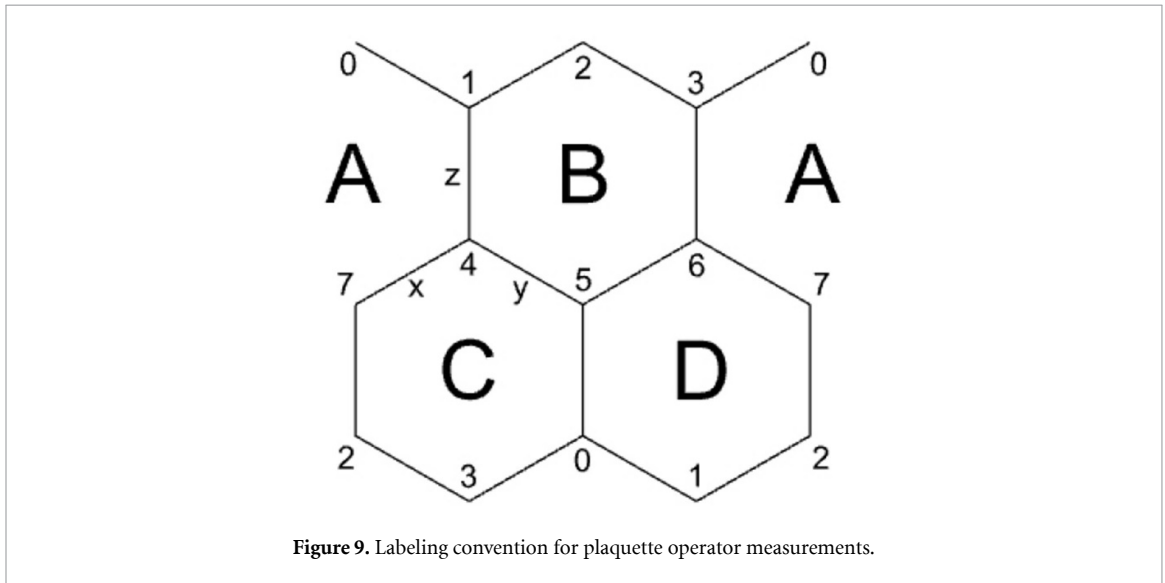

Figure 9. Labeling convention for plaquette operator measurements.

RBM parameters takes us to the right excited state, we measure the relevant plaquette operators in the system. The convention for labeling plaquette operators and sites is as shown in figure 9.

For the trained RBM state after performing QST for groundstate of 2×2 lattice, the expectation values of its four plaquette operators as well as the measured energies are shown in table 3. Moreover, flipping the y- and z-components of the spin at lattice site number 5 (or equivalently, applying $\hat{S}_5^x$) will produce a pair of vortices at B and D plaquettes. Therefore, the expectation value of the $\hat{W}_p$ operator in these plaquettes also flips sign. Similarly, applying $\hat{S}_7^y$, will kill the vortex in plaquette D and produce one in plaquette A. The changes in the sign of expectation values for these plaquettes is shown in the table 3. Finally, by applying $\hat{S}_3^z$ we return back to the no-vortex sector. These results, prove the effectiveness of the framework developed in section 4 for realization of vortices in RBM states, in both QST and usual learning approaches.

Note that the energy after each flip changes by an amount $\Delta E \sim 2$. This can be explained by the fact that each vortex sector has its own groundstate but also excited states with higher energies created by fermionic quasi-particle excitations. For example, the no-vortex sector has a spectrum given by the Hamiltonian in equation (20) and while the groundstate is the lowest lying state in this sector, there will be excited quasi-particle states carrying momentum $\mathbf{k}$ which are created by the $\gamma_{\mathbf{k}}^{\dagger}$ operators. Similarly, vortex sectors with a non-trivial vortex number can be in an excited state by applying quasi-particle creation operators. A rough estimate shows that the difference $\Delta E \sim 2$ is of the order of several such a quasi-particle excitations and thus the vortices we are creating carry such excitations. This also explains why when returning to the no-vortex sector we do not arrive at the original energy.

6. Relation to conformal blocks and Moore-Read state

As originally described in [1], the non-Abelian sector of the Kitaev Honeycomb model which we are addressing in the current paper is in the same universality class as the Ising CFT. That is the topological

Table 3. Expectation values of plaquette operators after spin flips.

Spin flip operator (acting successively)	W_A	W_B	W_C	W_D	Energy
Groundstate	0.9489	0.9517	0.9525	0.9553	−6.078
$\hat{S}_5^x$	0.9489	−0.9517	0.9525	−0.9553	−4.011
$\hat{S}_7^y$	−0.9489	−0.9517	0.9525	0.9553	−1.804
$\hat{S}_3^z$	0.9489	0.9517	0.9525	0.9553	0.061

excitations of the model correspond to the primary operators of the chiral sector of the Ising CFT, namely the vacuum corresponds to the identity operator $\mathbb{1}$, the vortex corresponds to the σ-field and the quasi-particle corresponds to the ψ-field. This relation is part of a broader framework connecting topological phases of $2+1$-dimensional quantum field theories to chiral sectors of WZW CFTs that can be traced back to the work of Witten on Chern–Simons theory and WZW models [19, 33], see [5] for a review. Within this correspondence wave-functions of the $2+1$-dimensional topological theory are mapped to conformal blocks of the WZW model and vice versa. One very well-known application of this paradigm is the description of fractional Quantum Hall states with filling fraction

$$\nu = \frac{p}{k+pn}, \tag{48}$$

using conformal blocks in an $SU(p)_k$ WZW model [34], where p and k are positive integers, and n is an arbitrary positive integer. For $p=1$, these are simply the Laughlin states. The case of Ising anyons corresponds to the choice $p=k=2$ and leads to the famous Moore-Read state [18].

Let us now describe this map in some degree of detail which is suitable for our purposes. The ground state wave-function of the topological theory on the 2-sphere is described by the identification [35]

$$\Psi_0(z_1,\ldots,z_N) = \prod_{i<j}^{N}(z_i-z_j)^{n+k/p}\langle \mathcal{O}_R(z_1)\ldots\mathcal{O}_R(z_N)\rangle e^{-\sum_i |z_i|^2/4l_B^2}, \tag{49}$$

where in order to obtain a unique ground state, the conformal block $\langle \mathcal{O}_R(z_1)\ldots\mathcal{O}_R(z_N)\rangle$ is computed in the representation $R = \mathrm{Sym}_k$, namely the kth symmetric representation of $SU(p)$, which has the property

$$\underbrace{\mathrm{Sym}_k \times \mathrm{Sym}_k \times \ldots \times \mathrm{Sym}_k}_{p} = \mathbb{1}. \tag{50}$$

In the case of quantum hall states, $l_B = \sqrt{\hbar/eB}$ and the z_i are positions of 'electrons'. Specifying to the case $p=k=2$, we obtain the *Pfaffian* state [35]

$$\Psi_0^{\mathrm{Pf}}(\{z_i\}) = \mathrm{Pf}\left(\frac{|ij\rangle_1}{z_i-z_j}\right)\prod_{i<j}(z_i-z_j)^{n+1}e^{-\frac{1}{4l_B^2}\sum_j |z_j|^2}, \tag{51}$$

where we have taken $\mathcal{O}_R$ to be the spin 1 field ψ of the Ising model and the Pfaffian of a $2\,m$-by-$2\,m$ antisymmetric matrix is defined as

$$\mathrm{Pf}(A) = \frac{1}{2^m m!}\sum_{P\in S_{2m}}\mathrm{sgn}(P)\prod_{i=1}^{m}A_{P(2i-1),P(2i)}. \tag{52}$$

The state $|ij\rangle_1$ appearing in the Pfaffian is a spin singlet formed from two spin 1 states

$$|ij\rangle_1 = |1_i\rangle|-1_j\rangle + |-1_i\rangle|1_j\rangle - 2|0_i\rangle|0_j\rangle. \tag{53}$$

We will see the significance of this choice shortly. The non-Abelian sector of the Kitaev model corresponds to the value $n=0$ as quasi-particles carry the same electric charge as an electron

$$e^* = \nu e = e \quad \Rightarrow \quad \nu = 1 \quad \Rightarrow \quad n = 0. \tag{54}$$

With these choices equation (51) is a spin singlet version of the Moore Read state [18]. Next, we would like to interpret equation (51) as a ground state wave-function of the Kitaev model. That such a correspondence is possible for spin lattice models has been previously shown in the case of Laughlin spin-liquid states on

lattices [20, 36]. Passing over to the torus with periodic boundary conditions, which is the situation examined in the present paper, the Pfaffian state lifts by the replacements

$$\frac{1}{z_i - z_j} \mapsto \frac{\vartheta[\alpha](z_i - z_j)}{\vartheta_1(z_1 - z_j)}, \quad z_i = (x_i + iy_i)/L_x, \tag{55}$$

to the following expression involving theta functions [37],

$$\Psi_{0,\alpha}^{\mathrm{Pf}}(z_1, \ldots, z_N) = e^{-\sum_i |z_i|^2/4l_B^2} F_{\mathrm{cm}}^{\alpha}(\sum_i z_i) \mathrm{Pf}\left(\frac{\vartheta[\alpha](z_i - z_j)}{\vartheta_1(z_i - z_j)} |ij\rangle_1\right) \prod_{i<j} \vartheta_1(z_i - z_j), \tag{56}$$

where the modular parameter of the theta-functions is chosen to be the one of the lattice[4], namely in our case $\tau = e^{\pi i/3}$. The center of mass contribution has been modified in order to accommodate for the fact that $\nu = 1$ and, using $N_s = L_x L_y/2\pi l_B^2$, takes the form

$$F_{\mathrm{cm}}^{\alpha}(z) = \vartheta\begin{bmatrix} (N_s - 2)/2 + (1 - 2\epsilon)/2 \\ -(N_s - 2)/2 - (1 - 2\epsilon')/2 \end{bmatrix}(z, \tau), \tag{57}$$

and we refer the reader to appendix B for conventions and definitions of the theta functions. Note that on the torus, the ground state is triply degenerate with the different states corresponding to the values $\alpha = (1/2, 0)$, $(0, 0)$, and $(0, 1/2)$. The transition between these ground states can be understood as creation of vortex/anti-vortex pairs which are then successively transported along the cycles of the torus and then annihilated, as beautifully described in [39]. Moreover, for $B = 0$ we have $l_B \to \infty$ and $N_s = 0$, resulting in periodic boundary conditions for the wave-function.

Now let us make the connection between what we have said above about the Moore-Read state and the RBM ground state wave-functions constructed in this paper. First of all, note that the RBM wave-function is in the spin basis of the electrons at each lattice site. Recall that, as discussed in section 4, the Hilbert space of our spin $\frac{1}{2}$ electrons can be described by two spinless complex fermions a_1 and a_2. In particular, spin up will correspond to both modes unoccupied, while spin down arises from both modes occupied:

$$|\uparrow\rangle = |00\rangle, \quad |\downarrow\rangle = |11\rangle, \tag{58}$$

with $a_1|00\rangle = a_2|00\rangle = 0$ and $|11\rangle = a_1^\dagger a_2^\dagger|00\rangle$. However, we remark here that $|\uparrow\rangle$ and $|\downarrow\rangle$ are not the standard spin up and down basis vectors with respect to σ^z, but rather rotated basis vectors which respect the symmetries of the Honeycomb Hamiltonian which for $J_x = J_y = J_z = 1$ is invariant under permutation of x, y and z. In particular, we make the choice of basis satisfying

$$\langle\uparrow|\sigma^\alpha|\uparrow\rangle = \frac{1}{\sqrt{3}}, \quad \langle\downarrow|\sigma^\alpha|\downarrow\rangle = -\frac{1}{\sqrt{3}}, \tag{59}$$

where $\alpha = z, x, y$. A concrete solution in the σ^z-basis satisfying the above constrains is given by

$$|\uparrow\rangle = \begin{pmatrix} \sqrt{\frac{1}{2} + \frac{1}{2\sqrt{3}}} \\ \sqrt{\frac{1}{2} - \frac{1}{2\sqrt{3}}} e^{i\pi/4} \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} \sqrt{\frac{1}{2} - \frac{1}{2\sqrt{3}}} \\ \sqrt{\frac{1}{2} + \frac{1}{2\sqrt{3}}} e^{i5\pi/4} \end{pmatrix}. \tag{60}$$

The representation (58) is faithful if we restrict to the subspace of fermionic states $|\Psi\rangle$ that satisfy

$$D_i|\Psi\rangle = |\Psi\rangle, \quad D_i = (1 - 2a_{1,i}^\dagger a_{1,i})(1 - 2a_{2,i}^\dagger a_{2,i}). \tag{61}$$

This projects out the unphysical states $|10\rangle$ and $|01\rangle$. As pointed out in [35], the unprojected system can also describe the degrees of freedom of a spin 1 particle

$$|00\rangle = |1\rangle, \quad |11\rangle = |-1\rangle, \quad |10\rangle = |01\rangle = |0\rangle. \tag{62}$$

Thinking reversely, we propose to identify our RBM state $|\Phi\rangle$ with the Pfaffian state by projecting onto physical states [5]

$$|\Phi\rangle \sim \widehat{P}_{\mathrm{physical}} \Psi_{0,\alpha}^{\mathrm{Pf}}, \tag{63}$$

---

[4] After publishing this work as preprint, we were notified by the authors of [38] who had obtained a similar ansatz for rectangular lattices with periodic boundary conditions.

[5] In general, the RBM state obtained after training will be a linear combination of the three groundstates.

where the positions z_i of spin 1 operators ψ are identified with complex lattice point coordinates of lattice site i on the torus. Put another way, an electron at site i of the Honeycomb lattice is represented in the conformal block by $\psi(z_i)$. Note that the number of lattice sites is twice the number of plaquettes and therefore always even such that the Pfaffian is well-defined. Moreover, note that the projection operator $\widehat{P}_{\text{physical}}$ projects our spin 1 singlet onto the following spin $\frac{1}{2}$ tensor product

$$\bigotimes_k \left(\frac{1 + D_k}{2} \right) |ij\rangle_1 = |\uparrow_i\rangle|\downarrow_j\rangle + |\downarrow_i\rangle|\uparrow_j\rangle. \tag{64}$$

The relation (63) should not be thought of as an exact equality but rather as an approximation. The reason is that due to the complexity of the Pfaffian state it is not possible for us, yet, to show that the projected state respects all the symmetries of the Honeycomb Hamiltonian. In fact, our Ansatz (56) is highly symmetric and is invariant under the 120 degree rotational symmetry of the Honeycomb lattice as well as the symmetry operator which permutes the Pauli matrices, as described in appendix A. Together, these symmetries leave the Hamiltonian invariant and should thus also leave the groundstate sector invariant, which in appendix A is used to conjecture a more precise version of the correspondence (63). Moreover, we note that our Ansatz satisfies three key criteria which the ground state wave-function has to respect. Namely, it is invariant under translations by torus lattice generators, it is triply degenerate, and it is lying in a 2^m-dimensional sub-space of the 2^{2m}-dimensional Hilbert space which also happens to be the dimension of the vortex-free sector.

7. Conclusion

In this paper, we have constructed RBM realizations of several sectors of the Kitaev Honeycomb model. In particular, we have focused on the gap-less non-Abelian phase corresponding to the parameter choice $J_x = J_y = J_z$ and used machine learning algorithms to train RBMs to find groundstates and excited states of such systems. We used two different approaches to find these states, namely `NetKet` and PyTorch. We find that using PyTorch and our own software package, we can achieve a very accurate result for the groundstate energy for small lattice sizes. On the other, the RBM trained with `NetKet` does not give an accurate enough result for the energy but the sampling algorithm employed is very efficient and allows `NetKet` to also proper larger lattices. An ideal approach would be to combine `NetKet`'s sampling algorithm with our own groundstate search algorithm. We leave this task for the future. Since our RBM wave-functions have 99.8% overlap with the exact groundstate (this result is for the 3×3 lattice), we find that we can transition to vortex sectors by manipulating RBM parameters as described in section 4. This transitioning into vortex sectors has been also demonstrated starting from a wave-function obtained from quantum state tomography. Here we are able to reconstruct an exact wave function from single-shot measurements and use this method to 'detect' the vortex state of the system. This is useful for applications in topological quantum computation where one would need to 'decode' input data, then perform a quantum gate operation, and finally create an output.

Of course, in the case of the Honeycomb model, we have an exact solution of the spectrum using Kitaev's Majorana fermion representation. However, such a representation, though giving exact values for the energies of given states in the Hilbert space, is at the same time losing information. For example, it is difficult and in some cases unknown how to compute expectation values of string operators composed of Pauli matrices. Moreover, one may lose the Majorana representation in cases where a complicated perturbation is added to the Hamiltonian which obscures such a simple representation. For all these reasons it is desirable to obtain realizations of the quantum state in terms of spin eigenstates and the RBM gave in equation (21) is such a representation.

One important application of our results which we present in section 6, is to relate our RBM groundstate wave-function to the Moore-Read Pfaffian state of the Ising CFT. This relation is part of the broader framework of correspondence between $2 + 1$ dimensional topological field theories and 2d CFTs which goes back to the work of Witten on WZW models and Chern–Simons theory [19, 33]. Our relation is a proposal which passes first non-trivial tests but should be examined in more detail to find further evidence and/or possible proof for it. We leave this for future work. Here, we note that such a correspondence is of great significance as it opens a door for neural networks to represent quantum field theory correlation functions of a high degree of complexity. On the one hand, the parameters of the RBM wave-function grow only polynomially with the spin number, while the CFT correlation function to which it is identified is of exponential complexity in the spin number. We hope that in the future one can generalize this framework/recipe to other models. It would be also very interesting to see whether there is a connection between what we propose, namely using an RBM to compute CFT conformal blocks, and the proposal made in [40] regarding a novel relation between neural networks and quantum field theory.

Data availability statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

Appendix A. Symmetries of the Pfaffian ansatz

One remarkable property of the Ansatz (56) is that the wave-function is invariant under the symmetries of the Honeycomb lattice as well as under the operator which permutes the Pauli matrices. The latter is rather easy to see. Following [6], one can define the operator $\hat{U}_{C_6}$ acting on all lattice sites simultaneously

$$\hat{U}_{C_6} = \bigotimes_i \frac{1}{2}(\sigma_i^0 + i\sigma_i^x + i\sigma_i^y + i\sigma_i^z), \tag{A1}$$

and permutes the Pauli matrices as follows

$$\hat{U}_{C_6}\sigma^x\hat{U}_{C_6}^\dagger = \sigma^z, \quad \hat{U}_{C_6}\sigma^y\hat{U}_{C_6}^\dagger = \sigma^x, \quad \hat{U}_{C_6}\sigma^z\hat{U}_{C_6}^\dagger = \sigma^y. \tag{A2}$$

Now it is easy to see that $\hat{U}_{C_6}$ leaves our spin-pair states invariant, i.e.

$$\hat{U}_{C_6}\left(|\uparrow_i\rangle|\downarrow_j\rangle + |\downarrow_i\rangle|\uparrow_j\rangle\right) = |\uparrow_i\rangle|\downarrow_j\rangle + |\downarrow_i\rangle|\uparrow_j\rangle. \tag{A3}$$

Next, let us show invariance under the $\hat{C}_6$ operator by which we denote the operator generating the 120 degree rotational lattice symmetry. In practice, this operator leads to a re-indexing of sites as depicted in figure 10.

Let us now see how our Ansatz is invariant under the action of this operator. To this end, note that $\hat{C}_6$ acts on differences of site coordinates isotropically, namely by a 120 degree rotation. This can be summarized by the following formula

$$\hat{C}_6(z_i - z_j) = e^{2\pi i/3}(z_i - z_j). \tag{A4}$$

Now note, since $\tau = e^{i\pi/3}$, that this can be rewritten in terms of the complex structure of the torus as follows

$$\hat{C}_6(z_i - z_j) = (\tau - 1)(z_i - z_j). \tag{A5}$$

We can then use the transformation properties of theta functions as reviewed in appendix B to derive the following identity

$$
\begin{aligned}
\vartheta_1(\tau, (\tau-1)(z_i - z_j)) &= \vartheta_1\begin{bmatrix}1/2\\1/2\end{bmatrix}(\tau, (\tau-1)(z_i - z_j)) \\
&= \kappa\left(\begin{bmatrix}1/2\\1/2\end{bmatrix}, \gamma\right)^{-1}(\tau-1)^{-\frac{1}{2}}e^{-i\pi\frac{(z_i-z_j)^2}{\tau-1}}\vartheta_1\left(-\frac{1}{\tau-1}, z_i - z_j\right) \\
&= \kappa\left(\begin{bmatrix}1/2\\1/2\end{bmatrix}, \gamma\right)^{-1}(\tau-1)^{-\frac{1}{2}}e^{-i\pi\frac{(z_i-z_j)^2}{\tau-1}}\vartheta_1(\tau, z_i - z_j),
\end{aligned}
\tag{A6}
$$

where in the first line we have applied the $SL(2,\mathbb{Z})$ transformation

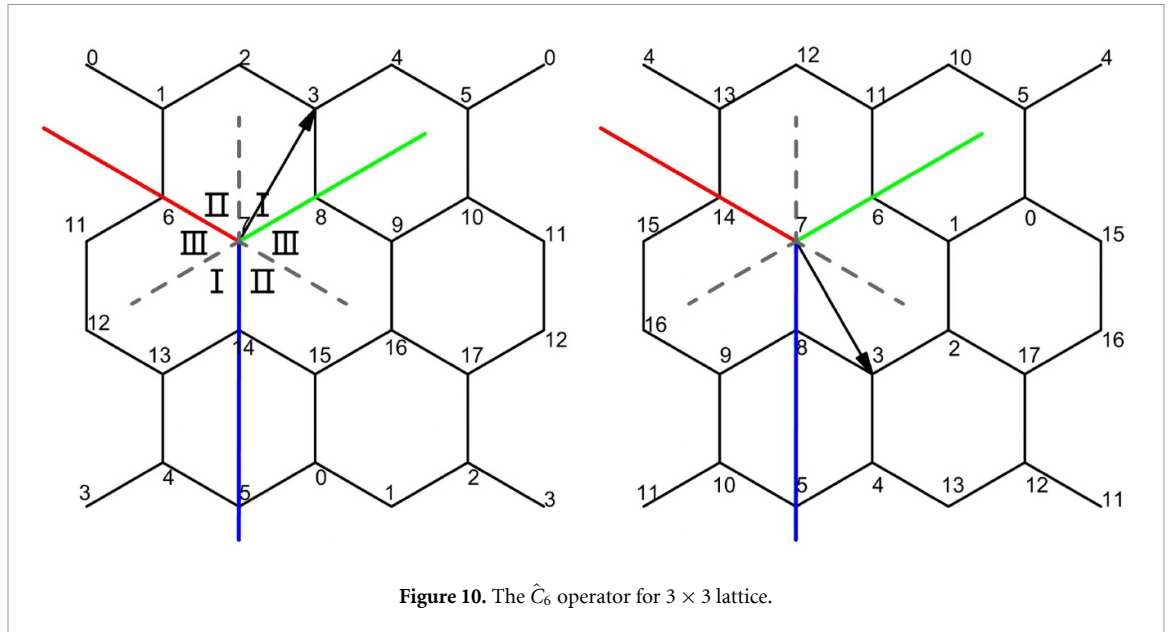$$\gamma = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \tag{A7}$$

Figure 10. The $\hat{C}_6$ operator for 3×3 lattice.

and in the last line we have made use of the identity $\frac{-1}{\tau - 1} = \tau$ for our particular choice of τ corresponding to the honeycomb lattice. Similarly, one can derive the identity

$$
\begin{aligned}
\vartheta[\alpha](\tau, (\tau - 1)(z_i - z_j)) &= \vartheta \begin{bmatrix} \epsilon \\ \epsilon' \end{bmatrix} (\tau, (\tau - 1)(z_i - z_j)) \\
&= \kappa([\alpha], \gamma)^{-1} (\tau - 1)^{-\frac{1}{2}} e^{-i\pi \frac{(z_i - z_j)^2}{\tau - 1}} \vartheta \begin{bmatrix} -\epsilon - \epsilon' - 1/2 \\ \epsilon \end{bmatrix} \left(-\frac{1}{\tau - 1}, z_i - z_j \right) \\
&= \kappa([\alpha], \gamma)^{-1} (\tau - 1)^{-\frac{1}{2}} e^{-i\pi \frac{(z_i - z_j)^2}{\tau - 1}} \vartheta \begin{bmatrix} -\epsilon - \epsilon' - 1/2 \\ \epsilon \end{bmatrix} (\tau, z_i - z_j).
\end{aligned} \tag{A8}
$$

Combining equations (A6) and (A8), we then see that the ratio of the two transforms as follows under the action of the $\hat{C}_6$ operator

$$
\hat{C}_6 \frac{\vartheta[\alpha](\tau, z_i - z_j)}{\vartheta_1(\tau, z_i - z_j)} = \frac{\vartheta[\alpha](\tau, (\tau - 1)(z_i - z_j))}{\vartheta_1(\tau, (\tau - 1)(z_i - z_j))} = \kappa([\alpha], \gamma)^{-1} \kappa \left(\begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}, \gamma \right) \frac{\vartheta[\beta](\tau, z_i - z_j)}{\vartheta_1(\tau, z_i - z_j)}, \tag{A9}
$$

where the resulting spin structure β of the theta function after the transformation is given as follows

$$
\alpha = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \longrightarrow \beta = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix} \longrightarrow \beta = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix} \longrightarrow \beta = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{A10}
$$

Thus we see, that as the two factors of κ just contribute an overall phase which is the same for all summands of the Pfaffian in expression (56), the three groundstate wave functions just get permuted under the action of $\hat{C}_6$.

   This is, however, a too large symmetry group as the groundstate space should be only invariant under the *combined* action of $\hat{C}_6$ and $\hat{U}_{C_6}$. To understand what has been happening and how to modify the Ansatz such that it is only invariant under the combined symmetry, it is useful to have another look at the projection (64). When performing such projections, there is a freedom for the choice of phase factors and it is a-priory not clear which ones to choose. In (64) we have made the simplest choice for such factors. Here, we want to modify our choice and present a more precise conjecture for the relation between the Pfaffian state and the groundstate wave-functions. To this end, note that the Honeycomb lattice can be divided into three conical regions as shown in figure 10. These regions are connected by 120 degree rotations and invariant under inversions. When applying $\hat{P}_{\text{physical}}$, we introduce phases depending on in which region the complex value $x_i - z_j$ lies:

$$
\widehat{P}_{\text{physical}} |ij\rangle_1 = |\uparrow_i\rangle |\downarrow_j\rangle + |\downarrow_i\rangle |\uparrow_j\rangle \qquad \text{if } z_i - z_j \in \text{region I}, \tag{A11}
$$

$$
\widehat{P}_{\text{physical}} |ij\rangle_1 = |\uparrow_i\rangle |\uparrow_j\rangle \qquad \text{if } z_i - z_j \in \text{region II}, \tag{A12}
$$

$$\widehat{P}_{\text{physical}}|ij\rangle_1 = e^{2\pi i/3}|\downarrow_i\rangle|\downarrow_j\rangle \qquad\qquad \text{if } z_i - z_j \in \text{region III}. \tag{A13}$$

Next, note the above three spin states are each eigenvectors of $\hat{U}_{C_6}$ with eigenvalues given as follows

$$\hat{U}_{C_6}(|\uparrow_i\rangle|\downarrow_j\rangle + |\downarrow_i\rangle|\uparrow_j\rangle) = |\uparrow_i\rangle|\downarrow_j\rangle + |\downarrow_i\rangle|\uparrow_j\rangle, \tag{A14}$$

$$\hat{U}_{C_6}|\uparrow_i\rangle|\uparrow_j\rangle = e^{2\pi i/3}|\uparrow_i\rangle|\uparrow_j\rangle, \tag{A15}$$

$$\hat{U}_{C_6}|\downarrow_i\rangle|\downarrow_j\rangle = e^{-2\pi i/3}|\downarrow_i\rangle|\downarrow_j\rangle. \tag{A16}$$

Now it is easy to see that a permutation which moves $\widehat{P}_{\text{physical}}|ij\rangle_1$ from region I to region II is combined with an action of $\hat{U}_{C_6}$ with eigenvalue 1, a permutation which moves the state from region II to region III will be combined with an eigenvalue $e^{2\pi i/3}$ from the action of $\hat{U}_{C_6}$ resulting in a state in region III, and a permutation moving the state from region III to I will lead to a cancellation of phases upon the action of $\hat{U}_{C_6}$ resulting in a target state with the correct phase. Thus it is apparent that the Pfaffian state (56) is invariant (up to a permutation of theta function spin structures as described above) under the combined action of $\hat{C}_6$ and $\hat{U}_{C_6}$ while it is not invariant under individual actions of these operators. This mirrors the symmetry properties of the RBM state under the action of these operators exactly. We therefore conjecture the Pfaffian state to be an exact groundstate under the refined action of $\widehat{P}_{\text{physical}}$ as described above.

Appendix B. Theta functions

We define theta functions with *characteristic* $\left[\begin{smallmatrix}\epsilon\\\epsilon'\end{smallmatrix}\right] \in \mathbb{R}^2$ as follows:

$$\vartheta\begin{bmatrix}\epsilon\\\epsilon'\end{bmatrix}(\tau,z) = \sum_{n\in\mathbb{Z}} \exp 2\pi i\left\{\frac{1}{2}(n+\epsilon)^2 + (n+\epsilon)(z+\epsilon')\right\}. \tag{B1}$$

The classical Jacobi theta functions are related to these by the relations

$$\vartheta_1(\tau,z) = \vartheta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\end{bmatrix}(\tau,z), \tag{B2}$$

$$\vartheta_2(\tau,z) = \vartheta\begin{bmatrix}\frac{1}{2}\\0\end{bmatrix}(\tau,z), \tag{B3}$$

$$\vartheta_3(\tau,z) = \vartheta\begin{bmatrix}0\\0\end{bmatrix}(\tau,z), \tag{B4}$$

$$\vartheta_4(\tau,z) = \vartheta\begin{bmatrix}0\\\frac{1}{2}\end{bmatrix}(\tau,z). \tag{B5}$$

The theta functions are periodic, up to a multiplicative factor, under torus lattice translations. Namely, for all integers m and n we have

$$\vartheta\begin{bmatrix}\epsilon\\\epsilon'\end{bmatrix}(\tau,z+n+m\tau) = \exp 2\pi i\left\{n\epsilon - m\epsilon' - mz - \frac{m^2}{2}\tau\right\}\vartheta\begin{bmatrix}\epsilon\\\epsilon'\end{bmatrix}(\tau,z). \tag{B6}$$

Under integral shifts of the characteristics, the theta functions pick up a phase according to

$$\vartheta\begin{bmatrix}\epsilon+m\\\epsilon'+n\end{bmatrix}(\tau,z) = \exp(2\pi i\epsilon n)\vartheta\begin{bmatrix}\epsilon\\\epsilon'\end{bmatrix}(\tau,z), \tag{B7}$$

and under reflections they satisfy

$$\vartheta\begin{bmatrix}-\epsilon\\-\epsilon'\end{bmatrix}(\tau,z) = \vartheta\begin{bmatrix}\epsilon\\\epsilon'\end{bmatrix}(\tau,-z). \tag{B8}$$

A property of theta functions deeper than the previous identities is the *modular transformation* property. For any characteristic $\chi \in \left[\begin{smallmatrix} \epsilon \\ \epsilon' \end{smallmatrix}\right] \in \mathbb{R}^2$, and any element $\gamma = \begin{pmatrix} a & b \\ c & c \end{pmatrix}$ of $SL(2, \mathbb{Z})$, we have

$$\exp \pi i \left(\frac{-cz^2}{c\tau + d} \right) \vartheta \begin{bmatrix} \epsilon \\ \epsilon' \end{bmatrix} \left(\frac{a\tau + b}{c\tau + d}, \frac{z}{c\tau + d} \right) = \kappa \left(\begin{bmatrix} \epsilon \\ \epsilon' \end{bmatrix}, \gamma \right) (c\tau + d)^{\frac{1}{2}} \vartheta \begin{bmatrix} a\epsilon + c\epsilon' - \frac{ac}{2} \\ b\epsilon + d\epsilon' + \frac{bd}{2} \end{bmatrix} (\tau, z), \tag{B9}$$

for $z \in \mathbb{C}$, $\tau \in \mathbb{H}^2$, where $\kappa \left(\left[\begin{smallmatrix} \epsilon \\ \epsilon' \end{smallmatrix}\right], \gamma \right)$ satisfies the identity

$$\kappa \left(\begin{bmatrix} \epsilon \\ \epsilon' \end{bmatrix}, \gamma \right) = \exp 2\pi i \left\{ -\frac{1}{2} (a\epsilon + c\epsilon') bd - \frac{1}{2} (ab\epsilon^2 + cd\epsilon'^2 + 2bc\epsilon\epsilon') \right\} \kappa \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \gamma \right), \tag{B10}$$

with $\kappa \left(\left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right], \gamma \right)$ an eighth root of unity.

Appendix C. Extended data

Table 4. Numerical results for the density of hidden nodes, $\alpha = \frac{num_h}{num_v}$, and its influence on groundstate energy calculation. The results summarized in this table are for a 5×5 plaquette lattice.

α	Number of hidden nodes	Number of parameters	Energy
0.1	5	305	$-27.7808(59)$
0.3	15	815	$-33.4312(37)$
0.5	25	1325	$-36.6024(22)$
1.0	50	2600	$-37.2464(13)$
1.5	75	3875	$-37.1235(15)$
2.0	100	5150	$-37.3054(14)$
3.0	150	7700	$-37.3037(14)$
4.0	200	10 250	$-37.1007(16)$
6.0	300	15 350	$-37.2989(14)$
8.0	400	20 450	$-37.1710(15)$
10.0	500	25 550	$-37.1946(15)$

Table 5. Comparison between a single-layer FFNN and a novb-RBM performance, in terms of training time and accuracy of energy for a 5×5 plaquette lattice.

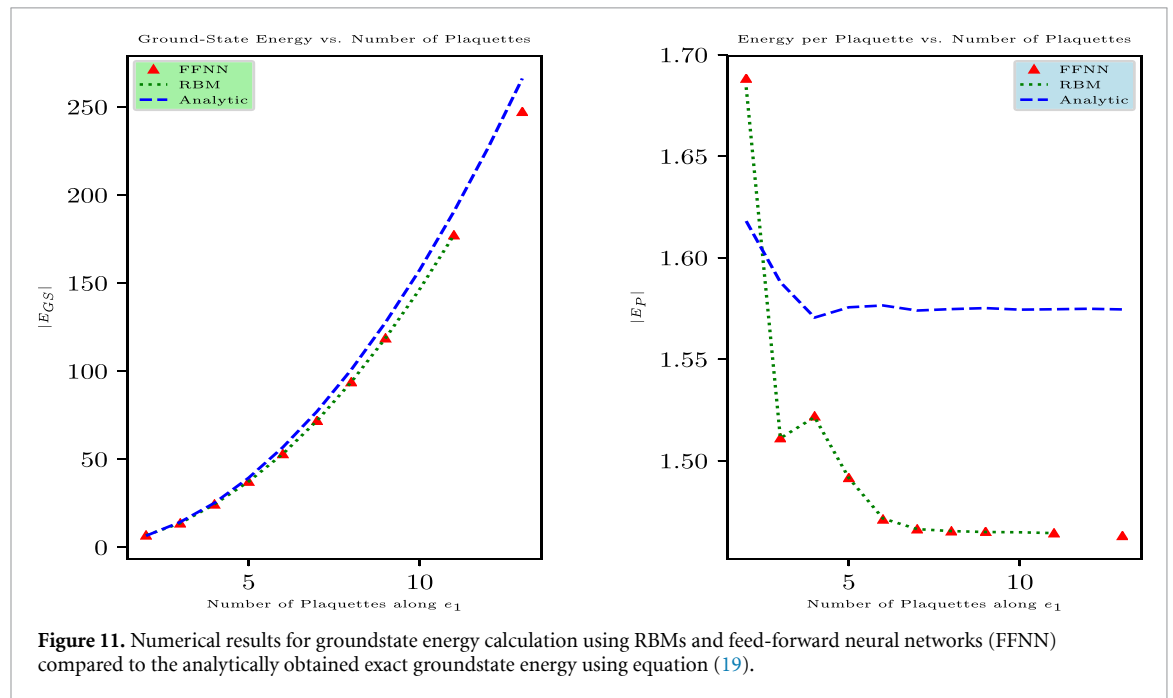| α	# Hidden Nodes	# Parameters	Energy(FFNN)	Energy (novb-RBM)	Time(FFNN)	Time (novb-RBM)
1.0	50	2550	$-37.2401(15)$	$-37.1532(15)$	2783 s	2754 s
2.0	100	5100	$-37.2958(14)$	$-36.8954(18)$	4591 s	5476 s
3.0	150	7650	$-37.3007(14)$	$-37.2191(15)$	6889 s	8331 s

Figure 11. Numerical results for groundstate energy calculation using RBMs and feed-forward neural networks (FFNN) compared to the analytically obtained exact groundstate energy using equation (19).

ORCID iD

Mohammadreza Noormandipour ● https://orcid.org/0000-0001-9294-2035

References

[1] Kitaev A 2006 Anyons in an exactly solved model and beyond *Ann. Phys. NY* **321** 2–111

[2] Kitaev A 2003 Fault-tolerant quantum computation by anyons *Ann. Phys. NY* **303** 2–30

[3] Jackeli G and Khaliullin G 2009 Mott insulators in the strong spin-orbit coupling limit: from Heisenberg to a quantum compass and Kitaev models *Phys. Rev. Lett.* **102** 017205

[4] Tikhonov K S, Feigel'man M V and Kitaev A Y 2011 Power-law spin correlations in a perturbed spin model on a honeycomb lattice *Phys. Rev. Lett.* **106** 067203

[5] Nayak C, Simon S H, Stern A, Freedman M and Das Sarma S 2008 Non-Abelian anyons and topological quantum computation *Rev. Mod. Phys.* **80** 1083

[6] Lee H-Y, Kaneko R, Okubo T and Kawashima N 2019 Gapless Kitaev spin liquid to classical string gas through tensor networks *Phys. Rev. Lett.* **123** 087203

[7] Lee H-Y, Kaneko R, Okubo T and Kawashima N 2020 Abelian and non-Abelian chiral spin liquids in a compact tensor network representation *Phys. Rev. B* **101** 035140

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

[9] Alcalde Puente D and Eremin I M 2020 Convolutional restricted Boltzmann machine aided Monte Carlo: an application to Ising and Kitaev models *Phys. Rev. B* **102** 195148

[10] Carleo G and Troyer M 2017 Solving the quantum many-body problem with artificial neural networks *Science* **355** 602–6

[11] Choo K, Carleo G, Regnault N and Neupert T 2018 Symmetries and many-body excitations with neural-network quantum states *Phys. Rev. Lett.* **121** 167204

[12] Nagy A and Savona V 2019 Variational quantum Monte Carlo method with a neural-network ansatz for open quantum systems *Phys. Rev. Lett.* **122** 250501

[13] Vicentini F, Biella A, Regnault N and Ciuti C 2019 and Variational neural-network ansatz for steady states in open quantum systems *Phys. Rev. Lett.* **122** 250503

[14] Yoshioka N and Hamazaki R 2019 Constructing neural stationary states for open quantum many-body systems *Phys. Rev. B* **99** 214306

[15] Jia Z-A, Zhang Y-H, Wu Y-C, Kong L, Guo G-C and Guo G-P 2019 Efficient machine-learning representations of a surface code with boundaries, defects, domain walls and twists *Phys. Rev. A* **99** 012307

[16] Lu S, Gao X and Duan L-M 2019 Efficient representation of topologically ordered states with restricted Boltzmann machines *Phys. Rev. B* **99** 155136

[17] Vieijra T, Casert C, Nys J, De Neve W, Haegeman J, Ryckebusch J, and Verstraete F 2020 Restricted Boltzmann machines for quantum states with non-Abelian or anyonic symmetries *Phys. Rev. Lett.* **124** 097201

[18] Moore G W and Read N 1991 Nonabelions in the fractional quantum Hall effect *Nucl. Phys. B* **360** 362

[19] Witten E and Field Q 1989 Theory and the jones polynomial *Commun. Math. Phys.* **121** 351

[20] Nielsen A E B, Cirac J I and Sierra G 2012 Laughlin spin-liquid states on lattices obtained from conformal field theory *Phys. Rev. Lett.* **108** 257206

[21] Carleo G *et al* 2019 NetKet: a machine learning toolkit for many-body quantum systems *Software* X **10** 100311

[22] Torlai G, Mazzola G, Carrasquilla J, Troyer M, Melko R and Carleo G 2018 Neural-network quantum state tomography *Nat. Phys.* **14** 447–50

[23] Pachos J K 2012 *Introduction to Topological Quantum Computation* (Cambridge: Cambridge University Press)

[24] Schmoll P and Orús R 2017 Kitaev honeycomb tensor networks: exact unitary circuits and applications *Phys. Rev. B* **95** 045112

[25] Jordan P and Wigner E P 1928 About the Pauli exclusion principle *Z. Phys.* **47** 631–51

[26] Feng X-Y, Zhang G-M and Xiang T 2007 Topological characterization of quantum phase transitions in a spin-1/2 model *Phys. Rev. Lett.* **98** 087204

[27] Chen H-D and Nussinov Z 2008 Exact results of the Kitaev model on a hexagonal lattice: spin states, string and brane correlators and anyonic excitations *J. Phys. A: Math. Theor.* **41** 075001

[28] Lieb E H 1994 Flux phase of the half-filled band *Phys. Rev. Lett.* **73** 2158

[29] Le Roux N and Bengio Y 2008 Representational power of restricted Boltzmann machines and deep belief networks *Neural Comput.* **20** 1631–49

[30] Zhang D W, Zhu Y Q, Zhao Y X, Yan H and Zhu S L 2018 Topological quantum matter with cold atoms *Adv. Phys.* **67** 253–402

[31] Leggett A J 2013 Lecture 26, the kitaev models, **PHYS598PTD** University of Illinois, Department of Physics

[32] Pedrocchi F L, Chesi S and Loss D 2011 Physical solutions of the Kitaev honeycomb model *Phys. Rev. B* **84** 165414

[33] Witten E 1992 On Holomorphic factorization of WZW and coset models *Commun. Math. Phys.* **144** 189

[34] Blok B and Wen X G 1992 Many body systems with nonAbelian statistics *Nucl. Phys. B* **374** 615

[35] Dorey N, Tong D and Turner C 2016 and Matrix model for non-Abelian quantum Hall states *Phys. Rev. B* **94** 085114

[36] Glasser I, Pancotti N, August M, Rodriguez I D and Cirac J I 2018 Neural-network quantum states, string-bond states and chiral topological states *Phys. Rev. X* **8** 011006

[37] Chung S B and Stone M 2007 Explicit monodromy of Moore–Read wavefunctions on a torus *J. Phys. A: Math. Theor.* **40** 4923–47

[38] Zhang H-C, Wu Y-H, Tu H-H and Xiang T 2021 Resonating valence bond realization of spin-1 non-Abelian chiral spin liquid on the torus *Phys. Rev. B* **103** 075130

[39] Oshikawa M, Kim Y B, Shtengel K, Nayak C and Tewari S 2007 Topological degeneracy of non-Abelian states for dummies *Ann. Phys. NY* **322** 1477–98

[40] Halverson J, Maiti A and Stoner K 2020 Neural networks and quantum field theory (arXiv:2008.08601 [cs.LG])