Topological Phases and the Negative Sign Problem in Neural-Network Quantum State Simulation

Benjamin Safvati
University of California, Berkeley
Physics Honors Thesis

May 17, 2019

Approved: _____
Advisor: Professor of Physics Joel Moore, Ph.D.

Contents

1. Introduction 1

2. The Haldane Phase and Hidden $\mathbb{Z}_2 \times \mathbb{Z}_2$ Symmetry 3

2.1. The Kennedy-Tasaki Transformation 5

2.2. Spin-1 Bilinear-Biquadratic Hamiltonian 7

3. Computational Methods 9

3.1. Variational Ground State Optimization 9

3.2. Stochastic Reconfiguration Gradient Descent 10

3.3. Markov Chain Monte Carlo Sampling 11

3.4. Restricted Boltzmann Machines 12

4. Simulation Results 14

4.1. AKLT Model 14

4.2. Spin-1 Heisenberg Antiferromagnet 19

5. Conclusions 21

6. References 22

1 Introduction: Finding Efficient Descriptions of Many-Body Physics

The emergent properties of interacting quantum particles are of great interest in condensed matter physics for the engineering of exotic quantum materials [18] and understanding of phenomena unexplained by the symmetry breaking paradigm in Ginzburg-Landau theory. Novel phases of matter, like those seen in the fractional quantum hall effect [20] and high-temperature superconductors [9], are indiscernible from the physics of individual particles, motivating new approaches to accurately simulate the quantum behavior of bulk ensembles of particles with strong interactions. The accessible states of many-body quantum systems notoriously live in a Hilbert space exponentially large with respect to system size, and so accessing a complete description of the many-body wavefunction is found to be unattainable for classical simulation techniques.

Modern approaches to the many-body problem have found success by employing efficient variational models that represent the wave function with polynomial scaling in system size. Tensor Network States, which represent the wavefunction by the product of variational tensor elements, excel at representing the ground states of realistic Hamiltonians in nature, where interactions between particles are local and the entanglement entropy of a sub-region scales with the sub-system boundary, rather than the volume of the region [13]. This constraint exponentially reduces the space of relevant states and motivates the design of Tensor Network States with varied entanglement structures for different systems, including Matrix Product States (MPS) for local lattice Hamiltonians in 1D with a non-zero energy gap, Projected Entangled Pair States (PEPS) for higher dimensional systems, and the Multiscale Entanglement Renormalization Ansatz (MERA) for systems near criticality [29]. More recently, Tensor Network States have been adopted outside of condensed matter physics for their elucidative geometric interpretation of entanglement structure in the ADS/CFT correspondence [32] and as general tools for describing complex correlations [12].

A more recent technique developed by Carleo and Troyer [4] uses a Neural Network representation of the wavefunction called a Neural Quantum State (NQS), where the complex amplitudes in a configuration basis are found as a function of the configuration and a set of hidden "neurons" that connect between sites and learn the relevant particle correlations. This machine learning-inspired ansatz was then used for a feedback-based stochastic optimization algorithm that accurately found the ground state of various local spin-1/2 Hamiltonians. Whereas the entanglement scaling law in a system acts as a barrier for efficient Tensor Network simulation, Neural Network quantum states have been shown to exhibit maximal volume-law entanglement scaling with the number of hidden units growing only linearly with system size [10]. Other work has made explicit transformations between the two variational models [6, 15], showing how a NQS with only local connections displays area-law entanglement and can be mapped to a MPS, whereas a Neural Network with dense connections has no efficient Tensor Network structure. [10]

Neural Networks are promising as generative models for quantum many-body states, but it is not clear from a computational perspective what kinds of systems can and can't be accurately described by a NQS with a reasonable number of variational parameters. The Affleck-Kennedy-Lieb-Tasaki (AKLT) model, a spin-1 Hamiltonian with a well-defined ground state, has a simple Tensor Network description, yet cannot be efficiently represented by a NQS [11], showing that the choice of variational model is highly problem-dependent and can be motivated by the physics of the system. In the AKLT model, a non-local "hidden" ordering of the system's ground state prevents efficient ground state simulation with a NQS, whereas the size of the exact MPS representation scales linearly with system size [6]. Interestingly, this model also suffers from the negative sign problem in Quantum Monte Carlo (QMC) simulation methods [28], where negative configuration probabilities lead to inaccurate sampling of the phase space. Because this problem is NP-hard, general solutions are unavailable and system-dependent. Yet for a wide class of spin-1 local Hamiltonians including the AKLT model it has been shown [28] that a non-local unitary transformation, first described by Kennedy and Tasaki [19], combined with a local basis change removes the negative sign problem and allows for accurate simulation of the transformed model. We shall later describe this transformation as removing the topological entanglement structure of the ground state, with confirmation from studies of entanglement spectra. [27]

Motivated by an interest in the generalizability of Neural Networks for the classical simulation of quantum spin systems with non-local orders and topological entanglement, this thesis implements the NQS extension to higher spins and studies numerical performance of Neural Quantum States in simulating a class of spin-1 Hamiltonians characterized by bilinear and biquadratic interactions. This includes the Heisenberg antiferromagnetic chain and the AKLT model, whose ground state is a typical example of symmetry-protected topological order. We will determine the effect of the negative sign problem on neural network simulation, and show how a disentangling unitary transformation [27] that removes the sign problem maps the non-local order in the original Hamiltonian into a more tractable ferromagnetic order. This transformation makes the Hamiltonian suitable for neural network simulation and leads to a highly accurate estimation of the ground state energy. This work demonstrates an obstacle to neural network many-body simulation and provides a better understanding of the relationship between ground state entanglement, non-local hidden order, and the capabilities of the NQS wavefunction representation.

2 The Haldane Phase and Hidden $\mathbb{Z}_2 \times \mathbb{Z}_2$ Symmetry

In order to accurately assess NQS simulation performance for systems with symmetry-protected topological phases and non-local order parameters, we focus our attention on 1D chains of spin-1 particles. A fascinating conjecture by Haldane demonstrated that identical Hamiltonians for particles of different spin can have drastically different system properties. Specifically he claimed that for integer spin, the antiferromagnetic Heisenberg model

$$H_{AFH} = \sum_i \vec{S_i} \cdot \vec{S_{i+1}}$$

has a unique ground state, nonzero energy gap above the ground state energy, and exponentially decaying spin correlations, whereas for half-integer spin this same model has ground state degeneracy, no energy gap, and power-law decay of correlations [22]. These properties for integer spin systems are now considered representative of a *Haldane* phase of quantum spin chains. Aside from its massive interest in the condensed matter community, this new quantum phase of matter is also of great interest for computational physicists. Because exponential decay of correlations implies an entanglement entropy area law, identification of this phase would provide a useful constraint on the quantum state's complexity, particularly for Tensor Network models.

The first integer spin chain proven to have a non-zero energy gap with a well-defined ground state is the spin-1 AKLT model [1],

$$H_{AKLT} = \sum_i \vec{S_i} \cdot \vec{S_{i+1}} + \frac{1}{3}(\vec{S_i} \cdot \vec{S_{i+1}})^2$$

similar to the AFH spin chain but with an added biquadratic interaction. This Hamiltonian was first imagined by mapping spin-1 particles to the symmetrized subspace of two spin-1/2 particles. Specifically, if we denote the spin-1 single particle basis by $\{|+\rangle, |0\rangle, |-\rangle\}$ and the spin-1/2 two-particle basis by $\{|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle\}$, then the transformation we consider is

$$|+\rangle \rightarrow |\uparrow\uparrow\rangle$$

$$|0\rangle \rightarrow \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$$

$$|-\rangle \rightarrow |\downarrow\downarrow\rangle$$

The creators of this model were motivated to identify a spin-1 Hamiltonian for which the ground state could be expressed as the zero eigenvector of the projection operator onto the total spin-2 subspace of neighboring pairs of spins. In this case, the ground state condition is that for any spin-1/2 particles that are neighbors and part of separate $s = 1$ symmetrized subspaces, the spins must exist in the singlet state

$$\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle).$$

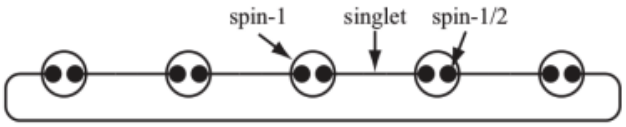

Figure 1: The Valence Bond Solid ground state of the spin-1 AKLT model. This state is defined by mapping each particle to the symmetrized subspace of two spin-1/2 particles, with neighboring spin-1/2 particles forming singlet states between each spin triplet pair. The case above represents periodic boundary conditions where the ground state is unique and translationally invariant, but for open boundary conditions the spin-1/2 edge spins on either side lead to a fourfold degeneracy in the ground state space that decays exponentially with the length of the chain.

This procedure derives the valence bond solid structure of the ground state of the AKLT Hamiltonian (visualized in figure 1). The rigorously defined model is the first demonstration of a Haldane phase in an integer Heisenberg spin chain, with a unique ground state (for periodic conditions), a non-zero energy gap, and exponential decay of correlations. Because these properties then lead to an entanglement area law for the AKLT model, this motivates the existence of an efficient Matrix Product State description [13, 16]. However as shown in the results section, these states cannot be efficiently described by a Neural Quantum State, even with this entanglement scaling constraint, implying that the limiting factors of NQS simulation are unclear.

Another characteristic of spin-1 systems in the Haldane phase is the existence of a nonzero non-local order parameter [26] over arbitrary distances, even in the presence of exponentially decaying spin correlations. The classical determination of phase transitions was understood under the lens of breaking some local symmetry, such as the loss of translational and rotational symmetry when water freezes into ice. In this framework the goal of condensed matter physicists is to identify the relevant local order parameter for a given symmetry so that one can determine the presence of different symmetries in a system's varied phases. Although Landau's paradigm lost its universality with the discovery of topological orders that persist at zero temperature, the discovery of a new kind of hidden order in these spin systems motivates a view of the Haldane Phase in terms of the breaking of a hidden symmetry.

For the class of antiferromagnetic spin-1 Heisenberg Hamiltonians including the AKLT model, there is a non-vanishing hidden string order O^{α}_{string} for a given ground state of a Hamiltonian H

$$O^{\alpha}_{string}(H) = \lim_{|i-j|\to\infty} \langle -S^{\alpha}_i e^{(i\pi \sum_{k=i+1}^{j-1} S^{\alpha}_k)} S^{\alpha}_j \rangle_H, \quad \alpha = (x, y, z)$$

where $\langle \dots \rangle_H$ is the expectation of an observable when measured from the ground state space of H. Intuitively, this parameter determines the presence of long-range antiferromagnetic order with $S^z = 0$ particles in between alternating $+$ and $-$ sites. The exponential term in the string order expression takes the sum of all non-zero spin values between the sites of interest and multiplies a factor of -1 if the sum total is an odd integer, thus evaluating to $+1$ for perfect variable-length antiferromagnetic order.

By mapping a NQS to Matrix Product States of optimal bond dimension, Chen et al. [6] illustrated how the presence of such a non-local order would require an inefficient amount of variational parameters for a NQS to keep track of the long-range correlations with an arbitrary set of zero spins between them. For an ensemble in the Haldane phase, configurations with hidden antiferromagnetic order like $\{+00...00-\}$ and $\{-00...00+\}$ become equally probable after fixing any set of zero spins and so trying to separate the wavefunction amplitudes with a local basis transformation cannot alleviate the simulation problem. Thus while Neural Network models can effectively capture volume law entanglement scaling, nontrivial *topological* entanglement structure in the ground state can affect Neural Network efficiency even when efficient Tensor Network representations exist. We will realize a NQS simulation of spin-1 chains in the Haldane Phase in the results section, where the convergence of the ground state energy is inaccurate whenever we also estimate a non-zero string order parameter, verifying the trouble caused by this hidden quantum phase.

2.1 The Kennedy-Tasaki Transformation

In the following section we describe a non-local unitary transformation inspired by the hidden string order parameter that explicitly maps the non-local order to a localized ferromagnetic order while preserving the energy levels of the original model. It is clear that the NQS wavefunction can efficiently capture the local correlations of the ferromagnetic phase, so clever application of this transformation would allow one to approximate the ground state energy and infer the phase structure of models that would otherwise be intractable for a NQS. Before realizing an explicit operator expression for the Kennedy-Tasaki (KT) transformation, we describe the action of this transformation on the eigenstates of the spin-1 configuration basis. Just as the string factor counts the nonzero spin values between sites, this transformation travels across the spin chain (from left to right for the case of open boundary conditions) and flips a non-zero spin if an odd number of non-zero spins were already traversed. A few examples below demonstrate how this transformation turns long range antiferromagnetic order into a ferromagnetic chain:

$$|0+0--0+-+0-\rangle \rightarrow |0+0+-0---0+\rangle$$

$$|+00-+0-000+\rangle \rightarrow |+00++0+000+\rangle$$

$$|-+0-00+0-0+\rangle \rightarrow |--0-00-0-0-\rangle$$

Notice that the first state does not transform into any meaningful phase, whereas the latter two states both exhibit long range antiferromagnetic correlations and so the resulting states are ferromagnetic.

This transformation is particularly useful for characterization of symmetries in the AKLT model, where the Valence Bond Solid state is defined as the equal superposition of all states displaying long range hidden order [19]. The KT transformation completely removes the hidden order that characterizes the Haldane phase, so the non-local action of this operator acts to change the state's phase structure. Thus we expect that this transformation also affects the symmetries of the Hamiltonian, since transitions between different phases are often accompanied by the breaking of

some symmetry. We will see a direct connection between ground state degeneracy and the hidden symmetry breaking by the action of this operator, further suggesting how these operations on the Hamiltonian can account for structure in the ground state that isn't easily captured by the simulation.

Note that although a non-zero string order parameter is proven to exist in the ground state of the spin-1 antiferromagnetic Heisenberg Chain, it is not the most general indicator of the varied symmetry-protected topological phases of 1D gapped systems, and a more generalized treatment of the phases has been clarified by the study of symmetries in varied spin representations and group cohomology theory [7].

Because several Heisenberg antiferromagnetic chains are in the Haldane phase, we choose the simple Hamiltonian H_{AFH} to understand the hidden symmetry-breaking picture of this non-local transformation. The KT transformation U has been written explicitly as the product of "pair disentanglers" [30] $U_{i,j}$ on every pair in the spin chain so that

$$U = \prod_{i<j} U_{i,j} = \prod_{i<j} e^{i\pi S_i^z S_j^x}.$$

This form of the unitary operator allows us to decompose the transformation to the pairwise local level, meaning the transformed Hamiltonian will also allow a 2-local description. When applied to the Heisenberg antiferromagnet, the new Hamiltonian $\tilde{H}_{AFH}$ becomes

$$\tilde{H}_{AFH} = UHU^{-1} = U(\sum_i \vec{S_i} \cdot \vec{S_{i+1}})U^{-1}$$

$$= \sum_j S_j^x (e^{i\pi S_{j+1}^x}) S_{j+1}^x + S_j^y (e^{i\pi (S_j^z + S_{j+1}^x)}) S_{j+1}^y + S_j^z (e^{i\pi S_j^z}) S_{j+1}^z.$$

Now for the spin-1 pauli matrices S^z and S^x we have

$$e^{i\pi S^x} S^x = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix} = -S^x$$

and similarly $e^{i\pi S^z} S^z = -S^z$. Thus the transformed Hamiltonian now has the form

$$\tilde{H}_{AFH} = \sum_j -S_j^x S_{j+1}^x + S_j^y (e^{i\pi (S_j^z + S_{j+1}^x)}) S_{j+1}^y - S_j^z S_{j+1}^z.$$

At first sight this new Hamiltonian appears far more complicated than the original, and in fact while the original Hamiltonian has global rotational spin symmetry, this new Hamiltonian is only globally symmetric for discrete π rotations along the x, y, and z axes. The three rotations then generate a group isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$, and so the previously hidden symmetry breaking in the original model now becomes manifest with the ferromagnetic order in the new system. This condition also relates very closely to the fourfold near-degeneracy of the ground states of these models, where spin-1/2 edge spin degrees of freedom at each end of the chain (for open boundary conditions) lead to a small energy splitting in the ground

state space that falls exponentially with chain length. In the original Hamiltonian this symmetry acts to protect the hidden order and thus originates the meaning of a symmetry-protected topological phase, of which there are numerous classified for various symmetries in gapped quantum systems [31].

The KT transformation has more recently been shown to act as the "perfect disentangler" of the Valence Bond Solid chain [27] by study of its effect on the change in the ground state's entanglement spectrum. Single particles in the VBS chain are maximally entangled with entanglement entropy $S = \ln 3$, whereas for the transformed Hamiltonian the entropy goes to zero and the state can be written in a single-particle basis. The topological aspects of the ground state can thus be filtered out by an appropriate non-local unitary operation, and for more general symmetry-protected topological phases the proper transformation can be extrapolated from the type of hidden symmetry and its corresponding string order parameter.

2.2 Spin-1 Bilinear-Biquadratic Hamiltonian

The KT transformation and similar disentangling operations on a many-body Hamiltonian have notably been used in Quantum Monte Carlo to remove the negative sign problem [28, 25] and allow for accurate estimation of ground state thermodynamic properties. Clearly there is a relationship between topological entanglement in the ground state and the negative sign problem, as both are present in the Haldane phase and both can be removed for a large class of spin-1 Hamiltonians by the transformation described above. We introduce a class of spin-1 Hamiltonians called the bilinear-biquadratic (BLBQ) chain, parameterized by α as done in [28].

$$H = \sum_i h_{i,i+1}$$

$$h_{i,i+1} = \vec{S_i} \cdot \vec{S_{i+1}} + \alpha[(\vec{S_i} \cdot \vec{S_{i+1}})^2 - 1]$$

where $\vec{S_i} = (S_i^x, S_i^y, S_i^z)$ is the vector of spin-1 pauli matrices on the ith lattice site and the parameter α determines the exact system under study. Notice that $\alpha = 0$ reduces to the Heisenberg antiferromagnet, whereas $\alpha = 1/3$ is equivalent to the AKLT model. Both of these spin chains lie within the Haldane phase of this parent Hamiltonian, and also contain positive off-diagonal elements in the standard $\{|+\rangle, |0\rangle, |-\rangle\}^{\otimes 2}$ basis.

$$h_{i,i+1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha-1 & 0 & 1-\alpha & 0 & \alpha & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-\alpha & 0 & \alpha & 0 & 1-\alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \alpha & 0 & 1-\alpha & 0 & \alpha-1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Because the KT transformation can be formulated in terms of pair operations, the transformed Hamiltonian can be written in terms of a transformed local interaction term as well. In addition we incorporate a local basis change V found in [27] such that the new Hamiltonian is sign problem-free in the new basis that we label $\{|1\rangle, |2\rangle, |3\rangle\}$. The single particle basis transformation on site i is defined as V_i such that

$$\begin{bmatrix} |1\rangle \\ |2\rangle \\ |3\rangle \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} |+\rangle \\ |0\rangle \\ |-\rangle \end{bmatrix}.$$

Then the complete local transformation on the Hamiltonian is simply the tensor product

$$V = \prod_i V_i$$

The resulting Hamiltonian is negative-sign free for $\alpha < -1$ and has the same energy eigenvalues because both transformations are unitary.

$$\tilde{H} = VUHU^\dagger V^\dagger = \sum_i \tilde{h}_{i,i+1}$$

$$\tilde{h}_{i,i+1} = \begin{bmatrix} \alpha & 0 & 0 & 0 & \alpha-1 & 0 & 0 & 0 & \alpha-1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha-1 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & \alpha-1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ \alpha-1 & 0 & 0 & 0 & \alpha-1 & 0 & 0 & 0 & \alpha \end{bmatrix}$$

This Hamiltonian will be the focus of our simulation efforts, and we will see how neural quantum states are unable to find the ground state energy of systems with the negative sign problem due to the nonlinearity of the parameter space. On the transformed Hamiltonian, however, these networks perform exceedingly well, even for sparsely connected models.

3 Computational Methods

In this section we go over the optimization procedure and model architecture for simulating the spin-1 BLBQ Hamiltonian on 1D chains with neural networks. The goal is to develop an understanding of the effect of topological entanglement on the NQS simulation of the ground state, and how the negative sign problem is realized in this variational model.

3.1 Variational Ground State Optimization

The starting point of any ground state numerical optimization algorithm is the variational principle in Quantum Mechanics. This method utilizes the quadratic structure of the energy expectation to show that this expression is bounded below by the lowest energy eigenvalue of the Hamiltonian, which is the ground state energy. This allows us to formulate the energy expectation as our objective function for finding the ground state of the Hamiltonian with the minimization problem

$$|\Psi_0\rangle = \min_{|\Psi\rangle} \langle\Psi| H |\Psi\rangle .$$

Although this quadratic form still suffers from exponentially large dimensions relative to system size, a landmark result by McMillan [24] shows how this expectation can be calculated numerically by averaging over the probability distribution of the target quantum state. We show this by algebraic manipulation of the objective function, noting that we can rewrite the quadratic form as a sum over pairs of configurations (S, S') where the element $H_{S,S'} = \langle S| H |S'\rangle$ is explicitly operated on in the calculation below

$$\frac{\langle\Psi| H |\Psi\rangle}{\langle\Psi| |\Psi\rangle} = \frac{\sum_{S,S'} \psi(S)^* \langle S| H |S'\rangle \psi(S')}{\sum_S |\psi(S)|^2}$$

$$\frac{\langle\Psi| H |\Psi\rangle}{\langle\Psi| |\Psi\rangle} = \frac{\sum_S |\psi(S)|^2 (\sum_{S'} H_{S,S'} \frac{\psi(S')}{\psi(S)})}{\sum_S |\psi(S)|^2}.$$

In this form the expectation value of the energy can be interpreted as a statistical average over the probability distribution determined by the quantum state. The quantity $E_{loc}(S) = \sum_{S'} H_{S,S'} \frac{\psi(S')}{\psi(S)}$, known as the local energy of configuration S, is efficiently computed for local Hamiltonians where the sparsity of non-zero elements is guaranteed.

Because neural networks often settle at local minima near the global minimum in the parameter space, a useful test to determine whether the ground state has been found is to estimate the energy variance by generating samples from the trained network. Because the ground state is an eigenstate of the Hamiltonian, the energy variance should be exactly zero:

$$\sigma_H^2 = \langle H^2\rangle - \langle H\rangle^2 = \langle\psi_0| H^2 |\psi_0\rangle - (\langle\psi_0| H |\psi_0\rangle)^2$$

$$\sigma_H^2 = E_0^2 \langle\psi_0| |\psi_0\rangle - E_0^2 = 0.$$

This property of the ground state also suggests that as the model learns the ground state, the variance will decrease, subsequently improving the next round of training.

3.2 Stochastic Reconfiguration Gradient Descent

From this energy objective function the direction of steepest descent towards the ground state can be found by taking the gradient with respect to each variational parameter in the quantum state model, allowing us to update the parameters iteratively. Shaping the model parameters into a vector $\theta \in \mathbb{C}^{N_{var}}$ where N_{var} is the number of model parameters, the model parameters are updated according to the following rule

$$\theta^{t+1} = \theta^t - \gamma F^t$$

where F is the gradient of the energy function, referred to as the generalized forces of the system in analogy with Lagrangian Mechanics. The learning rate γ scales the update, and by slowly decreasing this value the temperature of the system can be gradually brought down and the system can settle to the global minimum. Define the variational derivative of the ith parameter as

$$O_i(S) = \frac{1}{\psi(S)} \frac{\partial}{\partial \theta_i} \psi(S)$$

(this is the energy gradient for a model resembling the Boltzmann distribution, which will be true for our model). Then the generalized force for the ith parameter can be framed as an expectation over the quantum state probability distribution (see [4] for a derivation), where

$$F_i = \langle E_{loc} O_i^* \rangle - \langle O_i^* \rangle \langle E_{loc} \rangle.$$

This gradient expression iteratively updates the parameters by finding an approximation of the energy gradient based on a random sampling of the full data set. This process is referred to as stochastic gradient descent because the full gradient is never explicitly solved, and the updates calculated have some variance due to the finite sampling. The variance falls as the square root of the number of samples, so that larger sample sizes in general lead to more accurate gradient updates. This formulation, however, fails to capture the notion of distance within the Hilbert Space that the quantum state lives in. In particular, this descent method is implicitly defined on a Euclidean metric for the parameter space [5], whereas in describing the quantum state the update should be measured on the Fubini-Study metric [4] of the Hilbert Space. This transformation of the gradient update accounts for the quantum geometry of the model, where small steps in the parameter space can correspond to large updates to the wavefunction amplitudes that the model computes as a function of those parameters. This transformation is achieved by the Stochastic Reconfiguration algorithm as formulated by Sorella et al. [34] with the new parameter update

$$\theta^{t+1} = \theta^t - \gamma S^{-1} F^t$$

with the matrix S defined as

$$S_{i,j} = \langle O_i^* O_j \rangle - \langle O_i^* \rangle \langle O_j \rangle.$$

This matrix performs the transformation of the updates to the appropriate lengths induced by the Hilbert Space that quantum states occupy. This data augmentation is similar to other work in finding the "natural gradient" [2] for a set of parameters

based on the geometry of the data space. This correction to the parameter updates leads to more consistent convergence of the model to the correct ground state energies for all models learned.

In practice the inversion of the reconfiguration matrix S is prone to instability in the presence of stochastic noise, and so a regularization matrix λI is added to S that leads to more accurate parameter updates. Adding a multiple of identity to the matrix decreases the condition number of the matrix, a measure of instability when inverting a matrix that is found as the ratio of the largest and smallest eigenvalues. The regularization parameter λ should be gradually lessened as the model approaches the ground state to converge toward the true model without regularization.

3.3 Markov Chain Monte Carlo Sampling

The accuracy of the gradient descent method depends on the set of sample configurations $\{S_i\}$ used to compute the expectations involved in the stochastic parameter updates. Because the model approximates the ground state $|\Psi_0\rangle$, we want the expectation to be over the complex probability distribution determined by the amplitudes of $|\Psi_0\rangle$. Starting with randomly assigned parameters, this distribution can be approached asymptotically with the Metropolis-Hastings algorithm. To create samples, random walks in the space of spin configurations are taken and the new configurations are accepted with probability determined by the ratio of quantum state amplitudes. This process has a Markov chain interpretation [4] where a new spin configuration is added to the chain with probability

$$P(S^{k+1}) = \min\left(1, \left|\frac{\psi(S^{k+1})}{\psi(S^k)}\right|^2\right).$$

After several thousand iterations we have a large enough training set to compute parameter updates using the Stochastic Reconfiguration Algorithm. Samples should be acquired after every some amount of Metropolis steps, allowing proper time for the Markov Chain to mix so that the samples are uncorrelated. The theory requires that the samples should be accepted from the chain each time the autocorrelation time has passed, but in practice taking a number of steps equal to the number of variational parameters will suffice for proper sample mixing.

An improvement to this sampling method called parallel tempering [8] can improve mixing times and force the model to traverse a wider sampling space at some greater computational cost. To accomplish this, a number of replicas of the same network are created at different temperatures such that $T = 1$ is the lowest temperature and corresponds to the true network. Higher temperature models are more willing to accept configurations that are less probable according to the current parameters, so this helps the model to escape local minima. At each iteration a Metropolis-Hastings step is taken for each of the N_r replicates generating a list of configurations where the kth sample comes from the network at temperature

$T(k) = \frac{N_r}{N_r - k + 1}$. Then by iterating over the even indices i, samples x_i and x_{i+1} are swapped with probability

$$P_{swap} = e^{(\beta_{i+1} - \beta_i)[2\Re(\ln(\frac{\psi(x_i)}{\psi(x_{i+1})})]}$$

where β is the inverse temperature. The same swapping procedure is performed over all odd indices, then the current sample for the system at $T = 1$ is accepted for that step of the sampling. Because swaps to lower temperatures are more likely, the swapping action mixes in samples outside the nearby parameter space and allows for improved convergence to the global minimum. In practice this method didn't lead to any significant improvements in network convergence, but incorporating new techniques [17] that optimize the choice of temperatures and number of replicas could lead to some improvements for nonlinear optimization landscapes. This technique can also be made more efficient by parallelization of the sampling for each replica, which was not done here.

As the model parameters are gradually updated in the direction of lowest energy, the estimates for the state amplitudes will also improve. Then the next Markov Chain of spin configurations generated by the network will more closely resemble the sampling distribution of the ground state, leading to improved approximation of the gradient updates. This feedback process is readily implemented with the Restricted Boltzmann Machine (RBM) neural network described below, so that we may iteratively approach the exact ground state of various phases of the spin-1 BLBQ. Importantly, this simulation is performed in times polynomial with respect to the number of variational parameters, and so we are able to avoid the exploding dimensionality of the system as the number of particles increases.

3.4 Restricted Boltzmann Machines

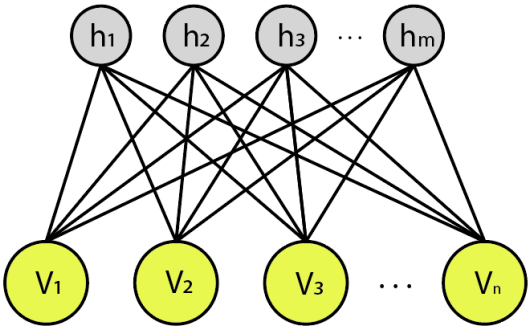

Figure 2: A Restricted Boltzmann Machine (RBM) with a visible layer representing the spin degrees of freedom at each site on the spin chain, as well as a hidden layer of neurons that keep track of correlations between spins. As a generative model, the RBM models a joint distribution over all of the network units, and the wavefunction amplitudes can then be found by tracing out the hidden units.

The choice of neural network architecture is highly non-generic and different models have been shown to faithfully represent the many-body wavefunction for var-

ious kinds of systems [4, 33]. For our purposes a practical choice is the Restricted Boltzmann Machine (RBM) [14], an energy-based model that resembles the canonical ensemble in statistical mechanics. This network consists of a layer of n visible neurons $\sigma_1, ..., \sigma_n$ that correspond to the spin degrees of freedom attached to a layer of m hidden units $h_1, ..., h_m$ that model the relevant correlations in the spin chain. Then the energy of a given configuration is found as

$$E(\sigma, h; \theta) = -\sum_{j=1}^{n} a_j \sigma_j - \sum_{i=1}^{m} b_i h_i - \sum_{i=1,\, j=1}^{m,\, n} W_{ij} h_i \sigma_j$$

where $\theta = (a_j, b_i, W_{ij})$ are the parameters of the neural network. A quantum state can then be expressed as the Boltzmann weight for the energy of that configuration,

$$\psi(\sigma) = \sum_{\{h_i\}} e^{-E(\sigma, h; \theta)} = \sum_{\{h_i\}} \exp\left(\sum_{j=1}^{n} a_j \sigma_j + \sum_{i=1}^{m} b_i h_i + \sum_{i=1,\, j=1}^{m,\, n} W_{ij} h_i \sigma_j \right).$$

The parameters are chosen to be complex, allowing us to compute both a magnitude and phase when modelling the ground state. The RBM is of special interest because there are no interactions between neurons in the same layer, which means the hidden units h_i can be explicitly traced out to identify the probability of the various spin configurations. The RBM thus models the wavefunction by finding a joint complex probability distribution over the visible and hidden layers, with the complexity of the model determined by the number of hidden units. After summing over the hidden units, an explicit representation of the quantum state amplitudes $\psi(\sigma)$ in terms of just the network parameters and visible spins is given by:

$$\psi(\sigma) = \exp\left(\sum_{i=1}^{n} a_i \sigma_i \right) \prod_{i=1}^{m} \left(2 \cosh\left[b_i + \sum_{j=1}^{n} W_{ij} \sigma_j \right] \right).$$

An important value is $\alpha \equiv m/n$, the ratio of hidden units to visible units, which characterizes the density of connections between the visible and hidden layer. In principle, by increasing α the representative ability of the RBM can be systematically improved, but the added complexity may also lead to overfitting of the model to the current training set. This ruins the generalizability of the network and leads to inaccurate estimates of observables [14].

For most machine learning algorithms, the success of the model depends on the careful tuning of hyperparameters, parameters of the model that are chosen rather than learned. This includes the number of samples generated in each training step, the regularization parameter, the learning rate, and the number of hidden units. In our simulations, the best results have been found with an initial rate of $\gamma(0) = .8$ and decay $\gamma(k) = \frac{\gamma(0)}{\sqrt{k}}$ where k is the iteration number. This gradually decreasing learning rate allows for larger parameter updates initially when the system is far from the global minimum, and smaller updates later to improve the accuracy of the model output when it is approaching the solution. Similarly, the regularization parameter λ decreases geometrically with iteration number so that $\lambda(k) = 100(.9)^k$. The number of hidden units and training samples necessary to accurately learn a given Hamiltonian's ground state has no clear answer; the simulation results presented below aim to understand what aspects of a Hamiltonian entail greater model complexity or a larger training set for accurate ground state representation.

4 Simulation Results

We will begin by attempting to naively approach the various spin-1 models in their Haldane phase with the RBM architecture and ternary visible units valued by the S^z eigenstates. A chain of N spins has 3^N possible complex amplitudes for the network to learn, and in the presence of hidden string order it is believed that the model will face difficulties learning the correct ground state [6]. These simulations are all done locally on my Macbook Pro 2015, and each run would finish in less than 2 hours. In all cases below the spin chains have periodic boundary conditions.

4.1 AKLT Model

The Hamiltonian used for simulation of the AKLT model is

$$H_{AKLT} = \sum_i \vec{S}_i \cdot \vec{S_{i+1}} + \frac{1}{3}(\vec{S}_i \cdot \vec{S_{i+1}})^2 - \frac{1}{3}.$$

The true ground state energy per spin in this form is $E_0 = -1$, but the NQS algorithm fails to reach this value for any simulation test performed for this paper.

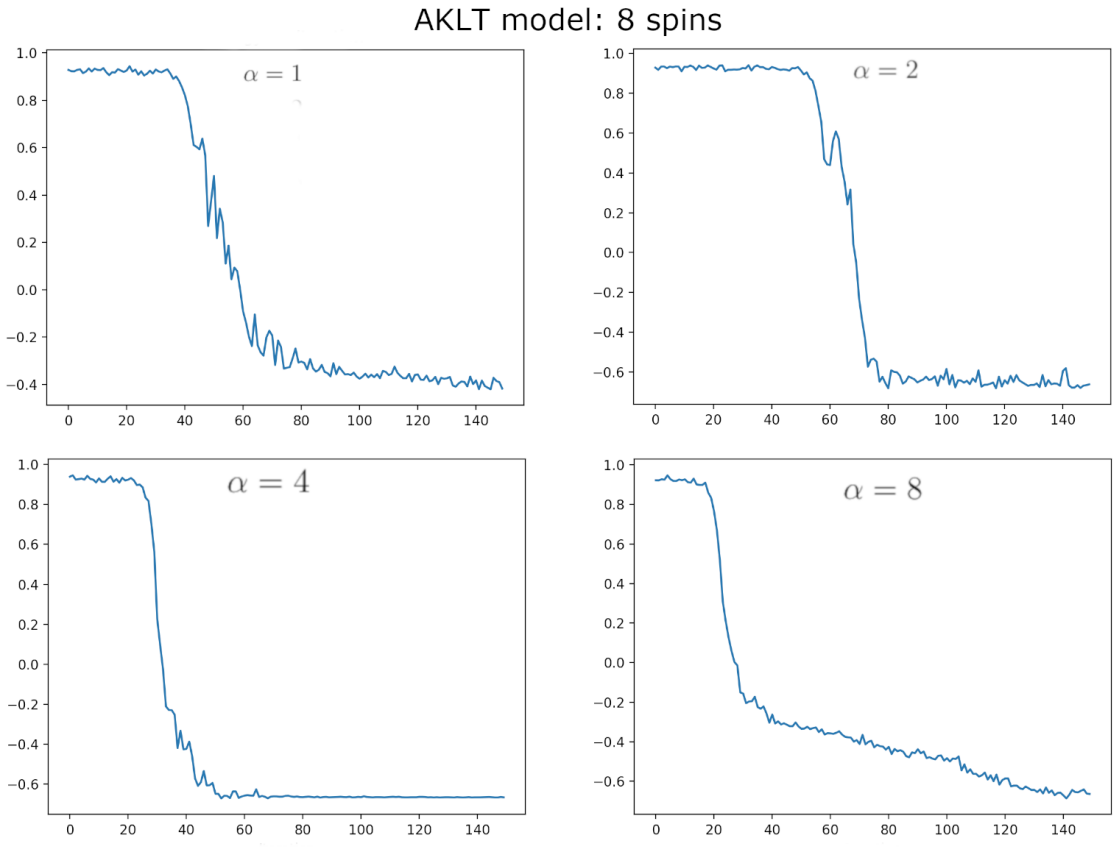

Figure 3: The plots above show convergence of the ground state energy after training of the NQS, demonstrating that increased connectivity in the RBM does little to improve the simulation accuracy on the original AKLT Hamiltonian. For a spin chain with only 8 spins, the model is unable to overcome the irregularity of the optimization landscape regardless of how many hidden units are used to represent the state.

By systematically increasing the ratio of hidden units to visible units, labelled by α, it became clear that Neural Networks are not well suited to this problem. For a system of 8 spins, we take 150 optimization steps, generating 300 Markov Chain Monte Carlo samples for a training set at each step for varying number of hidden units, and find that increasing the model complexity does little to improve the estimate of the ground state energy. The plots in figure 3 report the estimated energy per particle of the ground state of the Hamiltonian as a function of the iteration number. Although the networks converge in all cases, they appear to be stuck in a suboptimal local minimum, which is confirmed by the non-zero energy variance measured in all cases. Looking at the spin correlations generated by the trained network, there is an antiferromagnetic Neel order that shouldn't be present in the AKLT model, further proof that the model is getting stuck somewhere in the parameter space. The negative sign problem evidently corresponds to a great degree of nonlinearity that traps the optimization algorithm from approaching the true VBS state.

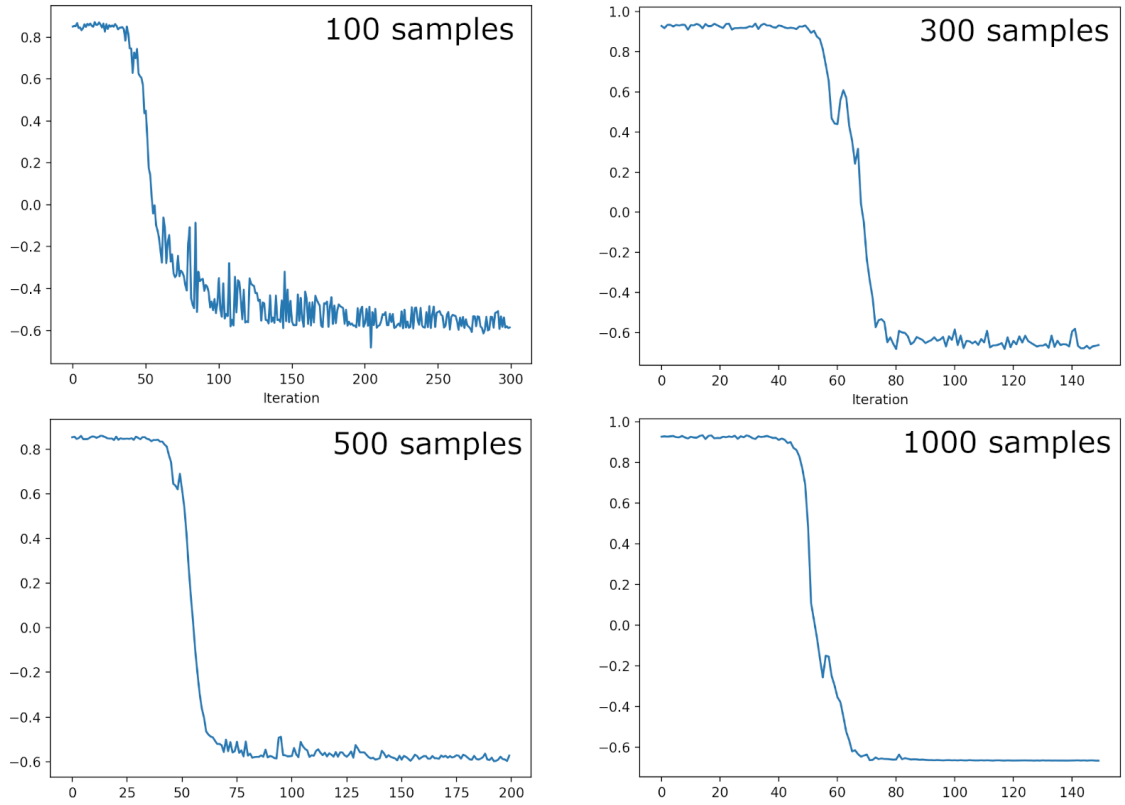

Figure 4: The plots above show how the size of the training set generated at each optimization step affect the convergence of the ground state energy. While higher sample sizes do reduce the statistical fluctuations in the predicted energies, each model approaches nearly the same energy in each case.

It is possible that the training sample set is producing gradients that are too noisy, and so we also vary the size of the training set produced at each optimization step to see if less noisy gradients can navigate the parameter space more precisely. Unfortunately, the plots in figure 4 indicate that there is still no improvement to the energy estimation with larger sample sizes, and the inaccurate antiferromagnetic

order is still present in the estimated ground state.

After applying the KT transformation to the AKLT model, the ground state energy should be identical, yet now the entanglement structure has been removed and the hidden order has been revealed as a local ferromagnetic order. There is still a negative sign problem without the addition of a local basis change, so by studying the performance of the NQS on the Hamiltonian $UH_{AKLT}U^\dagger$ it can be distinguished whether the hidden string order is responsible for the model's failure or if the negative sign problem in Quantum Monte Carlo is manifesting itself here more directly.

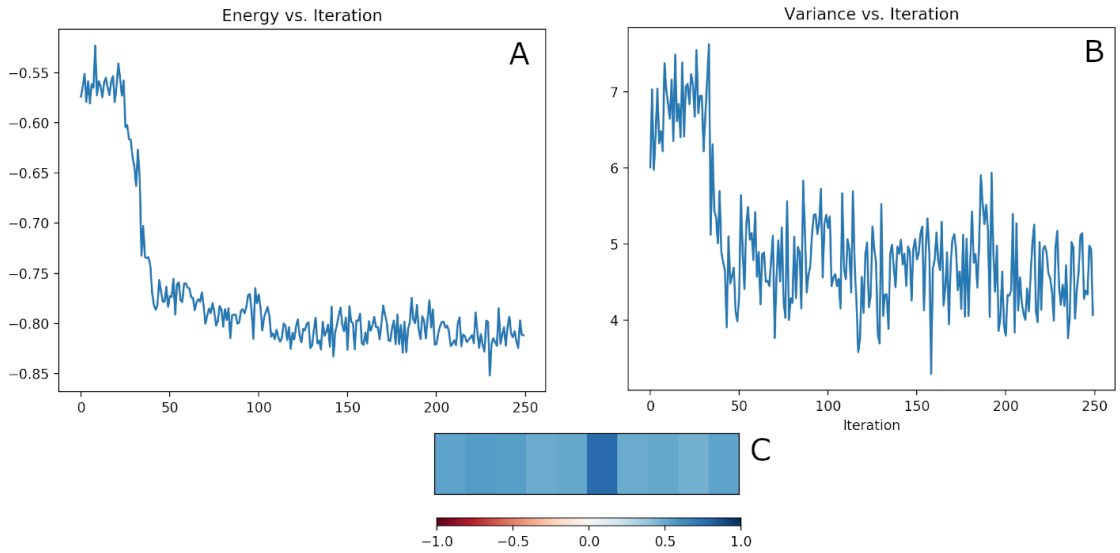

Figure 5: Simulation of the transformed Hamiltonian $UH_{AKLT}U^\dagger$ on 8 spins with hidden unit density $\alpha = 2$ and 500 samples per optimization step. (A) The energy fails to converge to the analytically calculated value $E_0 = -1$, but converges to an estimate that is noticeably improved from the original Hamiltonian. (B) The variance in the energy as the network learns drops then converges to some finite non-zero value. This is a reasonable indication that the network hasn't reached the global minimum, because the estimated state cannot be a true eigenstate of the many-body Hamiltonian. (C) A map of the spin correlations between spin 0 (framed in the middle due to periodic boundary conditions, dark blue site) and neighboring sites. Notice the ferromagnetic ordering is correctly predicted by the network even though the predicted ground state energy is incorrect.

After this transformation, the neural network still fails to converge to the true ground state energy, although the value it does converge to improves upon the estimate of the original Hamiltonian. The variance measurement provides a useful look into the predicted state's overlap with the true ground state, and as can be seen above the non-zero variance in the final estimate proves the model is settling on some local minimum. Because the negative sign problem was alleviated only partially whereas the hidden string order was in theory removed completely by the KT transformation, this lends evidence to the idea that the negative sign problem is relevant for Neural Quantum States.

To fully remove the positive off-diagonal elements from the AKLT model, a local basis change is applied in addition to the nonlocal KT transformation. Simulation of $\tilde{H}_{AKLT} = VUH_{AKLT}U^{\dagger}V^{\dagger}$ should in theory be highly efficient because the ground state of the transformed model in the new basis is composed of disentangled single spin states. This hypothesis proved true, and in the simulations below it is clear that neural networks can very easily converge to the true ground state energy even with a fractional hidden unit density on spin chains as large as length 80, shown below. For all cases studied the NQS ansatz is able to consistently converge to the global minimum with errors of order 10^{-4}. With increasing of the hidden unit density this error can be further decreased.

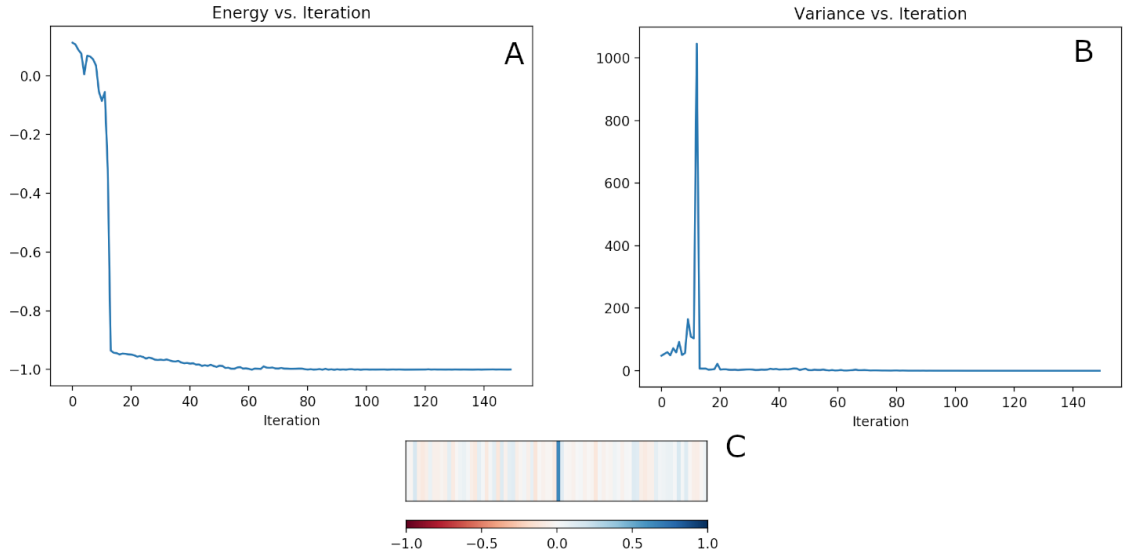

Figure 6: Simulation of the transformed AKLT Hamiltonian $\tilde{H}_{AKLT} = VUH_{AKLT}U^{\dagger}V^{\dagger}$ on a spin chain of length 80, with hidden unit density $1/4$ and 100 samples per optimization step. (A) The energy quickly approaches the true ground state energy after only about 20 parameter updates, suggesting this new ground state has a very efficient NQS description. This is especially surprising given only 20 hidden units are used to model a many-body state characterized by 3^{80} complex amplitudes. (B) The variance in energy confirms that the wavefunction estimated by this NQS is an eigenstate of the Hamiltonian. The sharp peak at around the 15th iteration corresponds to the point at which the model converges to the global minimum, and mirrors the fact that there exist increasing fluctuations near a critical point. (C) The spin correlations with respect to the center spin indicate that all of the spins are uncorrelated with each other, which agrees with the theory that all of the spins should be perfectly disentangled by the non-local transformation.

The NQS efficiently reaches the ground state energy even when the number of variational parameters is limited and the system size is made exceedingly large relative to what is tractable with exact diagonalization. Importantly, without the introduced transformations the AKLT model would be impossible to probe using QMC or neural network methods for any size spin chain. The removal of the negative sign problem was the key to achieving accurate simulation results, meaning the local basis change V needed to be included with the KT transformation for the algorithm to work. This local unitary operation has no obvious benefit to the

NQS simulation like in the case of the KT transformation other than removing the off-diagonal positive matrix elements, further hinting the connection between the negative sign problem and the capabilities of this simulation technique.

As a final test that the NQS ansatz is accurately modelling the AKLT model and its subsequent transformations, we numerically approximate the string order in each case and show how the KT transformation causes this hidden order to vanish. Because the hidden string order parameter is expressed in the limit of infinite distance between spins, the closest numerical approximation on a periodic spin chain of N spins is found by measuring the hidden string order in the S^z basis between sites 0 and $N/2$.

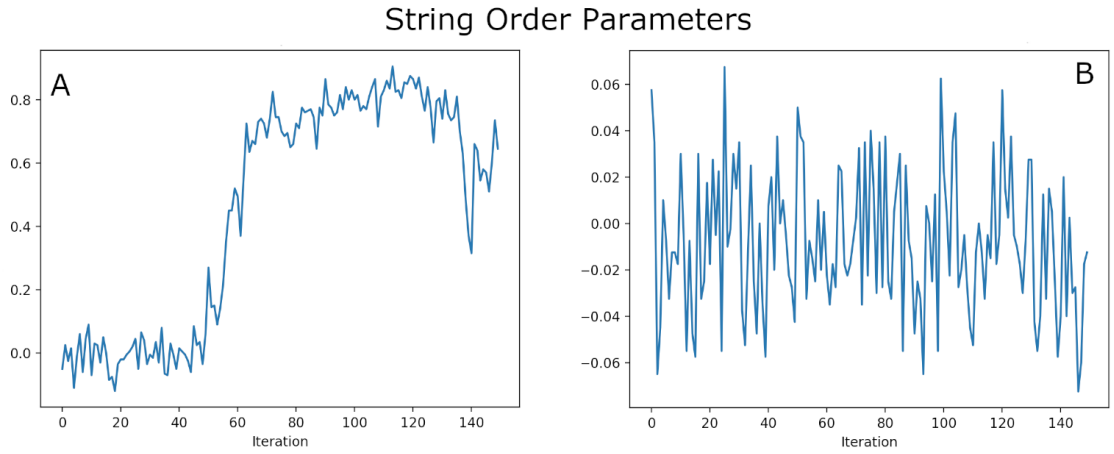

Figure 7: Estimation of the hidden string order parameter present in the Haldane phase for (A) the AKLT Hamiltonian and (B) the same Hamiltonian after applying the disentangling Kennedy-Tasaki transformation and a local basis change. The disappearance of the hidden order in (B) demonstrates that the learned model agrees with theory and that the removal of the Haldane phase has been accurately accounted for. Thus this transformation allows one to study the statistical properties of the AKLT model after dealing with the effects of the symmetry-protected topological entanglement with the non-local transformation.

Although the NQS trained on the original Hamiltonian fails to capture the true ground state energy, it does identify a non-zero hidden order parameter in every simulation we have conducted. Because the model does not reach the true ground state, however, the value of the estimated hidden order parameter is not entirely reliable, and the determination of the phase structure of this model becomes at best qualitative. An indication that simulation of the transformed AKLT Hamiltonian was successful comes from the vanishing of the string order parameter in the transformed Hamiltonian that is present in the original model.

4.2 Spin-1 Heisenberg Antiferromagnet

In a similar fashion to above we extend our study of NQS performance in the Haldane phase by simulating the spin-1 Heisenberg antiferromagnet. As predicted, simulation of this Hamiltonian suffers immensely in the presence of hidden order and the negative sign problem, meaning the NQS only approaches the correct ground state energy after both of these properties have been removed. The ground state energy for this Hamiltonian has been approximated to be $E_0 = -1.388$, yet the NQS converges to values closer to -1.0 when simulating the untransformed Hamiltonian, regardless of the hidden unit density or number of samples generated. The data below shows again how the algorithm has become trapped in a local minimum, but after the KT transformation has been applied the optimization landscape is sufficiently smoothed out to allow for efficient simulation.

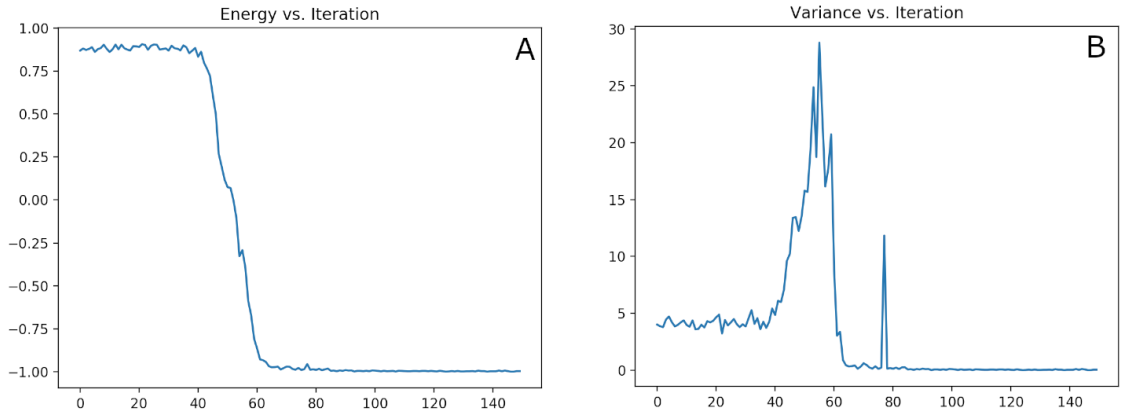

Figure 8: NQS simulation of the energy (A) and its variance (B) for 8 spins under the antiferromagnetic Heisenberg Hamiltonian with hidden unit density 2 and 300 samples gathered per optimization step. The nearly-vanishing energy variance indicates that the wavefunction has high overlap with an energy eigenstate. But because the converging energy value is different from the approximated ground state energy $E_0 = -1.388$ found in literature, this tell us that the model is being confused from the true ground state by the presence of higher energy eigenvalues occupying nearby local minima.

Even without the KT transformation, the neural network converges to an approximate eigenstate of the Hamiltonian as can be seen by the near-zero variance in the energy in the figure above. Yet in this case the calculated ground state energy doesn't match existing numerical approximations. Additionally, the samples generated from the estimated wavefunction have a local antiferromagnetic order, identical to the results of simulations of H_{AKLT} where we claimed the model was stuck in a local minimum. This result suggests that higher energy states of the Hamiltonian are affecting the gradient updates by leading the algorithm toward a local minimum that may not be in the direction of the desired global minimum eigenvalue.

Unlike the AKLT model, the Heisenberg antiferromagnet becomes negative sign problem-free without the local basis change, and in our test runs there is no noticeable difference in NQS performance if we include the basis change V or not for this

system. Now studying the Hamiltonian after the KT transformation, we see that the trained NQS exhibits the characteristic ferromagnetic order associated with the hidden symmetry breaking, just as in the case of the transformed AKLT model. We also see that convergence properties of this model are not as useful as in the transformed AKLT model: whereas the RBM efficiently approached the energy of the transformed AKLT ground state after less than 100 steps and with limited hidden units, the transformed antiferromagnetic spin chain approaches the true ground state energy only when the hidden density is relatively large and the number of optimization steps is extended to 500 and possibly greater for more precise results.

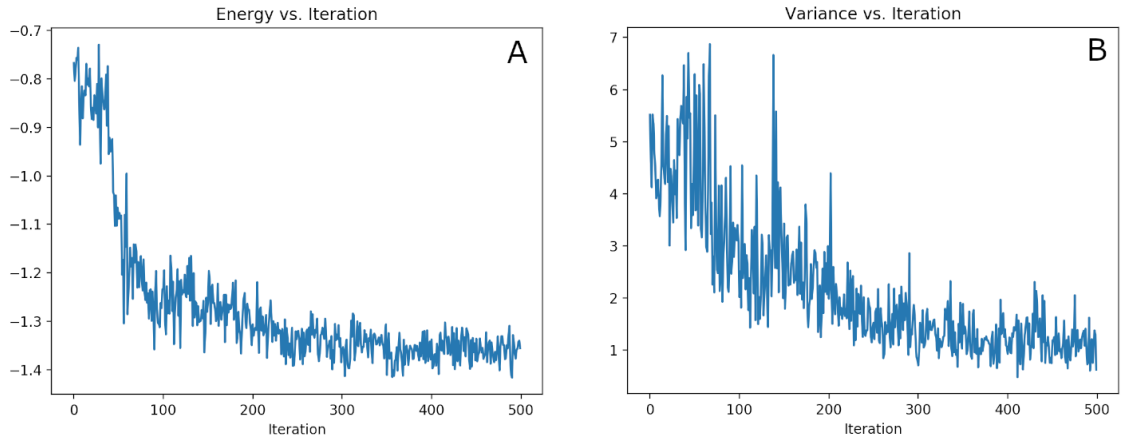

Figure 9: NQS simulation of the energy (A) and its variance (B) for 10 spins under the antiferromagnetic Heisenberg Hamiltonian after application of the KT transformation. The network has hidden unit density 4 and gathers 100 samples per optimization step. The noise in the plots arises from the small size of the training sets and can be systematically reduced by taking more samples. With even more hidden units or a longer optimization time more precise estimates of the ground state energy can be attained.

The KT transformation is clearly necessary for the accurate simulation of spin-1 chains in their Haldane phase, yet the presence of any negative sign problem in NQS simulation cannot be ignored; in every case of neural network simulation of a Hamiltonian with the negative sign problem, the estimated ground state energies were inaccurate until unitary transformations were applied to remove the sign problem. Subsequent simulations with the transformed models were much more forgiving, particularly for the AKLT state where the final form of the Hamiltonian had a simple product state description. While the KT transformation had a well-defined action on the topological entanglement structure of the original Hamiltonians, the role of the unitary operator V is much less clear, but can perhaps be motivated by study of the AKLT ground state and determination of a basis for which each particle is in an equal superposition of all of the new basis states. For other Hamiltonians with a topological phase protected by hidden $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry, the necessary operation on the Hamiltonian to remove the non-local order could be realized by the KT transformation and similar local basis changes.

5 Conclusions

It comes as no surprise that a classical machine would have problems simulating strongly correlated quantum systems. Neural networks, though possessing surprising generalization abilities and intrinsically non-local connections, fail at accurately representing certain spin-1 models with drastically more efficient tensor network descriptions. While these networks have demonstrated highly efficient means of representing highly entangled quantum states, the topological entanglement and hidden order in the AKLT state have been shown in this work to significantly hinder the accuracy of NQS simulation. The relationship between topological phases in the ground state and the negative sign problem is implicated by Kennedy and Tasaki's non-local unitary transformation, and further research is necessary to determine how the negative sign problem appears in other systems with symmetry-protected topological phases.

This investigation of Neural Network performance for topologically ordered systems is by no means comprehensive and has been approached by others. Lu, Gao, and Duan [23] constructed exact representations of several topologically ordered phases, including a unary representation of the spin-1 particles that can more efficiently decode the appropriate basis states. The problem of finding appropriate basis states may have limited the efficiency of simulating the antiferromagnetic Heisenberg chain after the KT transformation, as the necessary basis for the ground state to be written as a product state was unknown. For the AKLT model this unitary basis change was explicitly given, and the resulting simulation was extremely efficient. The ability to learn a relevant basis for a problem is appealing for extending this disentangling methodology to other forms of symmetry-protected topological phases, where knowledge of a hidden order parameter can motivate the design of nonlocal unitary operators that separate the topological entanglement structure from the Hamiltonian.

There are also open questions on the choice of neural network architecture and the number of hidden layers necessary for sufficient model complexity. Representability theorems for Deep Boltzmann Machines [3] (DBMs), which are RBMs with extra hidden layers, have guaranteed that wide classes of lattice Hamiltonians can be simulated efficiently with complexity scaling polynomially with system size. Surprisingly, these DBM constructions do not rely on any stochastic optimization, but rather are designed deterministically with a constructive algorithm that builds the necessary network bonds. This may lend itself to an explicit construction of the string bonds that can't be efficiently identified with RBMs. Another landmark result came with the use of convolutional neural networks for solving the spin-1/2 $J_1 - J_2$ antiferromagnetic Heisenberg model on square lattices [21]. This kind of neural network relies on symmetries of the Hamiltonian to reduce the number of variational parameters [4], and was able to approximate ground state energies for the Hamiltonian similar to the state of the art in Tensor Network simulation, even in places with significant frustration and the negative sign problem. Incorporating convolutional structures for the simulation of symmetry-protected topological phases would be an interesting prospect for future work in this area.

References

[1] Ian Affleck et al. "Valence Bond Ground States in Isotropic Quantum Antiferromagnets". In: *Condensed Matter Physics and Exactly Soluble Models* (1988), pp. 253–304. DOI: 10.1007/978-3-662-06390-3_19.

[2] Shun-Ichi Amari. "Natural Gradient Works Efficiently in Learning". In: *Neural Computation* 10.2 (1998), pp. 251–276. DOI: 10.1162/089976698300017746.

[3] Giuseppe Carleo, Yusuke Nomura, and Masatoshi Imada. "Constructing exact representations of quantum many-body systems with deep neural networks". In: *Nature Communications* 9.1 (2018). DOI: 10.1038/s41467-018-07520-3.

[4] Carleo et al. *Solving the Quantum Many-Body Problem with Artificial Neural Networks*. June 2016. URL: https://arxiv.org/abs/1606.02318.

[5] Michele Casula, Claudio Attaccalite, and Sandro Sorella. "Correlated geminal wave function for molecules:An efficient resonating valence bond approach". In: *The Journal of Chemical Physics* 121.15 (2004), pp. 7110–7126. DOI: 10.1063/1.1794632.

[6] Jing Chen et al. "Equivalence of restricted Boltzmann machines and tensor network states". In: *Physical Review B* 97.8 (2018). DOI: 10.1103/physrevb.97.085104.

[7] Xie Chen et al. "Symmetry protected topological orders and the group cohomology of their symmetry group". In: *Physical Review B* 87.15 (2013). DOI: 10.1103/physrevb.87.155114.

[8] Choo et al. *Symmetries and many-body excited states with neural-network quantum states*. July 2018. URL: https://arxiv.org/abs/1807.03325.

[9] Elbio Dagotto. "Correlated electrons in high-temperature superconductors". In: *Reviews of Modern Physics* 66.3 (1994), pp. 763–840. DOI: 10.1103/revmodphys.66.763.

[10] Deng et al. *Quantum Entanglement in Neural Network States*. May 2017. URL: https://arxiv.org/abs/1701.04844.

[11] Ivan Glasser et al. "Neural-Network Quantum States, String-Bond States, and Chiral Topological States". In: *Physical Review X* 8.1 (2018). DOI: 10.1103/physrevx.8.011006.

[12] Glasser et al. *Supervised learning with generalized tensor networks*. June 2018. URL: https://arxiv.org/abs/1806.05964.

[13] M B Hastings. "An area law for one-dimensional quantum systems". In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.08 (2007). DOI: 10.1088/1742-5468/2007/08/p08024.

[14] Geoffrey E. Hinton. "A Practical Guide to Training Restricted Boltzmann Machines". In: *Lecture Notes in Computer Science Neural Networks: Tricks of the Trade* (2012), pp. 599–619. DOI: 10.1007/978-3-642-35289-8_32.

[15] Huang et al. *Neural network representation of tensor network and chiral states*. Jan. 2017. URL: https://arxiv.org/abs/1701.06246.

[16] Vahid Karimipour and Laleh Memarzadeh. "Matrix product representations for all valence bond states". In: *Physical Review B* 77.9 (2008). DOI: 10.1103/physrevb.77.094416.

[17] Helmut G Katzgraber et al. "Feedback-optimized parallel tempering Monte Carlo". In: *Journal of Statistical Mechanics: Theory and Experiment* 2006.03 (2006). DOI: 10.1088/1742-5468/2006/03/p03018.

[18] B. Keimer and J. E. Moore. "The physics of quantum materials". In: *Nature Physics* 13.11 (2017), pp. 1045–1055. DOI: 10.1038/nphys4302.

[19] Tom Kennedy and Hal Tasaki. "HiddenZ2Z2symmetry breaking in Haldane-gap antiferromagnets". In: *Physical Review B* 45.1 (1992), pp. 304–307. DOI: 10.1103/physrevb.45.304.

[20] R. B. Laughlin. "Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations". In: *Physical Review Letters* 50.18 (1983), pp. 1395–1398. DOI: 10.1103/physrevlett.50.1395.

[21] Xiao Liang et al. "Solving frustrated quantum many-particle models with convolutional neural networks". In: *Physical Review B* 98.10 (2018). DOI: 10.1103/physrevb.98.104426.

[22] Elliott Lieb, Theodore Schultz, and Daniel Mattis. "Two Soluble Models of an Antiferromagnetic Chain". In: *Condensed Matter Physics and Exactly Soluble Models* (2004), pp. 543–601. DOI: 10.1007/978-3-662-06390-3_35.

[23] Sirui Lu, Xun Gao, and L.-M. Duan. "Efficient representation of topologically ordered states with restricted Boltzmann machines". In: *Physical Review B* 99.15 (2019). DOI: 10.1103/physrevb.99.155136.

[24] W. L. Mcmillan. "Ground State of LiquidHe4". In: *Physical Review* 138.2A (1965). DOI: 10.1103/physrev.138.a442.

[25] Tota Nakamura. "Vanishing of the negative-sign problem of quantum Monte Carlo simulations in one-dimensional frustrated spin systems". In: *Physical Review B* 57.6 (1998). DOI: 10.1103/physrevb.57.r3197.

[26] Marcel Den Nijs and Koos Rommelse. "Preroughening transitions in crystal surfaces and valence-bond phases in quantum spin chains". In: *Physical Review B* 40.7 (1989), pp. 4709–4734. DOI: 10.1103/physrevb.40.4709.

[27] Kouichi Okunishi. "Topological disentangler for the valence-bond-solid chain". In: *Physical Review B* 83.10 (2011). DOI: 10.1103/physrevb.83.104411.

[28] Kouichi Okunishi and Kenji Harada. "Symmetry-protected topological order and negative-sign problem forSO(N)bilinear-biquadratic chains". In: *Physical Review B* 89.13 (2014). DOI: 10.1103/physrevb.89.134422.

[29] Román Orús. "A practical introduction to tensor networks: Matrix product states and projected entangled pair states". In: *Annals of Physics* 349 (2014), pp. 117–158. DOI: 10.1016/j.aop.2014.06.013.

[30] M Oshikawa. "Hidden Z2*Z2symmetry in quantum spin chains with arbitrary integer spin". In: *Journal of Physics: Condensed Matter* 4.36 (1992), pp. 7469–7488. DOI: 10.1088/0953-8984/4/36/019.

[31] Frank Pollmann et al. “Symmetry protection of topological phases in one-dimensional quantum spin systems”. In: *Physical Review B* 85.7 (2012). DOI: 10.1103/physrevb.85.075125.

[32] Mukund Rangamani and Tadashi Takayanagi. “AdS/CFT and Tensor Networks”. In: *Holographic Entanglement Entropy Lecture Notes in Physics* (2017), pp. 221–234. DOI: 10.1007/978-3-319-52573-0_14.

[33] Hiroki Saito and Masaya Kato. “Machine Learning Technique to Find Quantum Many-Body Ground States of Bosons on a Lattice”. In: *Journal of the Physical Society of Japan* 87.1 (2018), p. 014001. DOI: 10.7566/jpsj.87.014001.

[34] Sandro Sorella, Michele Casula, and Dario Rocca. “Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods”. In: *The Journal of Chemical Physics* 127.1 (2007), p. 014105. DOI: 10.1063/1.2746035.