Neural Network Representations of Quantum Many-Body States

Brandon Euverman Supervisor: Jean-Sébastien Bernier

University of Northern British Columbia

NSERC CRSNG

UNBC UNIVERSITY OF NORTHERN BRITISH COLUMBIA

Motivation

Quantum many-body systems are challenging to solve due to the exponential growth of the Hilbert space with the number of degrees of freedom. Here, we investigate a method for approximating these systems using neural-network quantum states, as introduced by Carleo and Troyer [1].

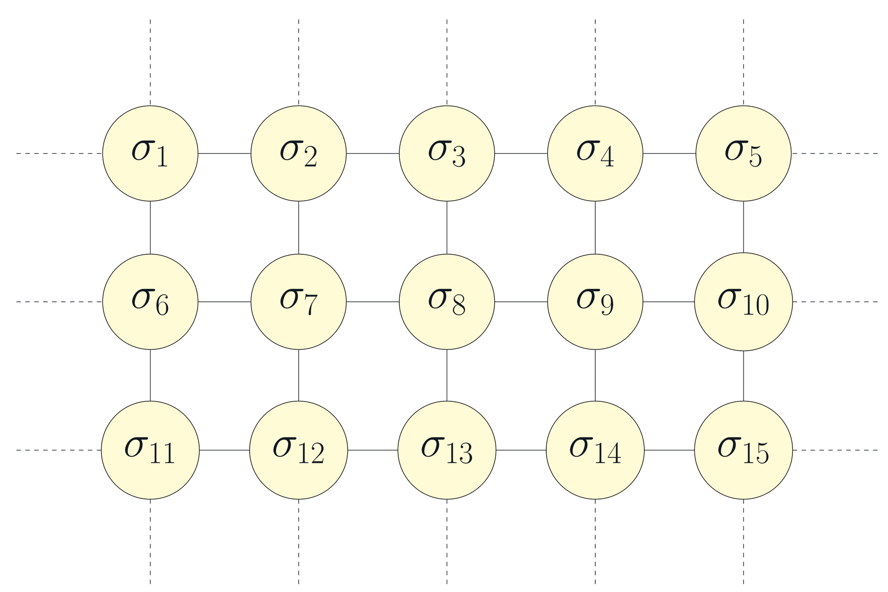

Fig. 1: A sample grid spin-lattice, dashed lines representing additional potential sites.

Model

Consider a system with N spin sites $S = \{\sigma_1, \sigma_2, \ldots, \sigma_N\}$, which take on values $\{-1, 1\}$. Our basis elements are then from $\{-1, 1\}^{\times N}$. We can write our state in this basis as

$$|\Psi(\mathcal{W})\rangle = \sum_{i=1}^{2^N} \langle S_i|\Psi(\mathcal{W})\rangle \, |S_i\rangle = \sum_{i=1}^{2^N} \Psi(S_i; \mathcal{W}) \, |S_i\rangle \, . \tag{1}$$

We represent a wavefunction in this system by way of a Restricted Boltzmann Machine (RBM), which consists of a visible layer of N nodes and a single hidden layer of M nodes. This corresponds to the following variational expression for the state:

$$\Psi(S; \mathcal{W}) = \frac{1}{2^M} \sum_{h_i \in \{-1,1\}} \exp\left(\sum_{j=1}^{N} a_j \sigma_j^z + \sum_{i=1}^{M} b_i h_i + \sum_{i=1}^{M} \sum_{j=1}^{N} W_{ij} h_i \sigma_j^z \right), \tag{2}$$

where the weights $\mathcal{W} = \{a_j, b_i, W_{ij}\}$ fully specify the response of the network to a given input S. In particular, a dictates the visible layer biases, b the hidden layer biases, and W the inter-layer connections. Since this architecture does not include any intra-layer interactions, the hidden layer can be explicitly traced out to read

$$\Psi(S; \mathcal{W}) = \exp\left(\sum_{j=1}^{N} a_j \sigma_j^z \right) \times \prod_{i=1}^{M} \cosh[\theta_i(S)], \tag{3}$$

where the *effective angles* θ_i are given by $\theta_i(S) = b_i + \sum_j W_{ij} \sigma_j^z$. It is convenient for computation to also consider the log of the wave function:

$$\ln[\Psi(S; \mathcal{W})] = \sum_{j=1}^{N} a_j \sigma_j^z + \sum_{i=1}^{M} \ln(\cosh[\theta_i(S)]). \tag{4}$$

For a given Hamiltonian $\mathcal{H}$, we then want to adapt the weights to provide the best possible representation of its ground state. We do this by reinforcement learning through minimizing the expectation value of the energy $E(\mathcal{W}) = \langle \Psi|\mathcal{H}|\Psi\rangle \, / \, \langle \Psi|\Psi\rangle$, with respect to the weights $\mathcal{W}$.

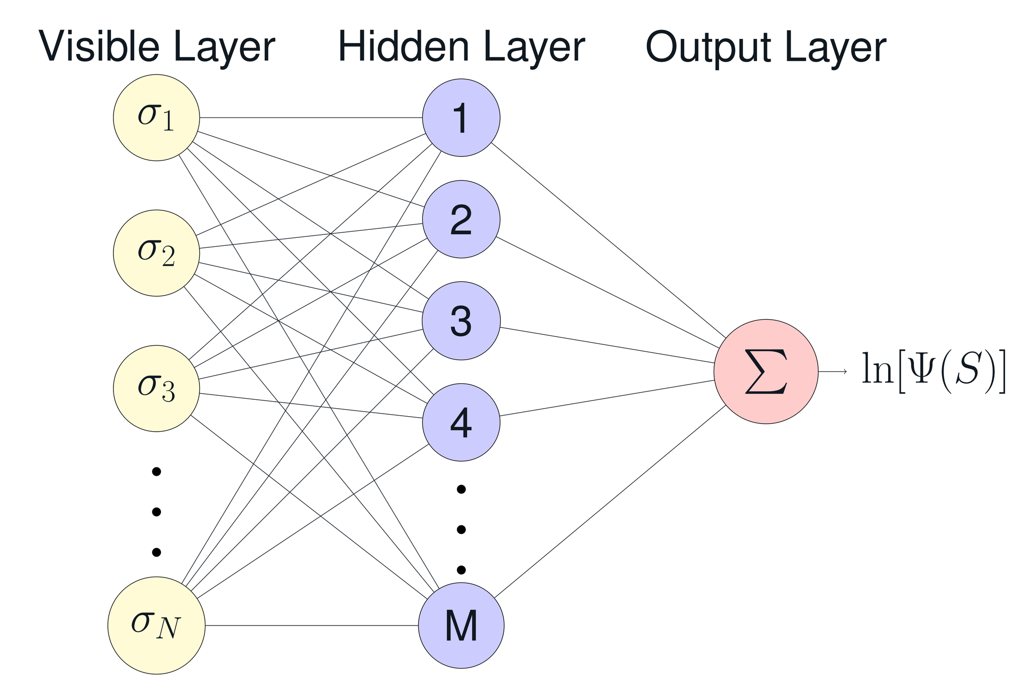

Fig. 2: Illustration of the neural network described by an Restricted Boltzmann Machine (RBM).

At each iteration k, a Monte Carlo sampling of $|\Psi(S; \mathcal{W}_k)|^2$ is obtained for the current set of parameters $\mathcal{W}_k$. Stochastic estimates of the energy gradient are then computed, with which a new set of weights $\mathcal{W}_{k+1}$ are proposed according to a gradient-descent optimization known as Stochastic Reconfiguration [4].

Results

To validate our model we apply it to finding the ground state of the transverse-field Ising model, given by

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z - h_x \sum_i \sigma_i^x \tag{5}$$

where $\langle i, j\rangle$ refers to a sum only over neighbouring sites, and σ^x and σ^z are Pauli matrices. This model is of interest due to it being one of the simplest models that undergoes a phase transition, from ferromagnetic to polarized at $J = h_x$, $J > 0$. This so-called critical point is also the most difficult place at which to analyze the system.

As an example, here we trained on a 25-site chain with periodic boundary conditions, using $J = h_x = 1$ and a hidden layer of size $M = 90$.

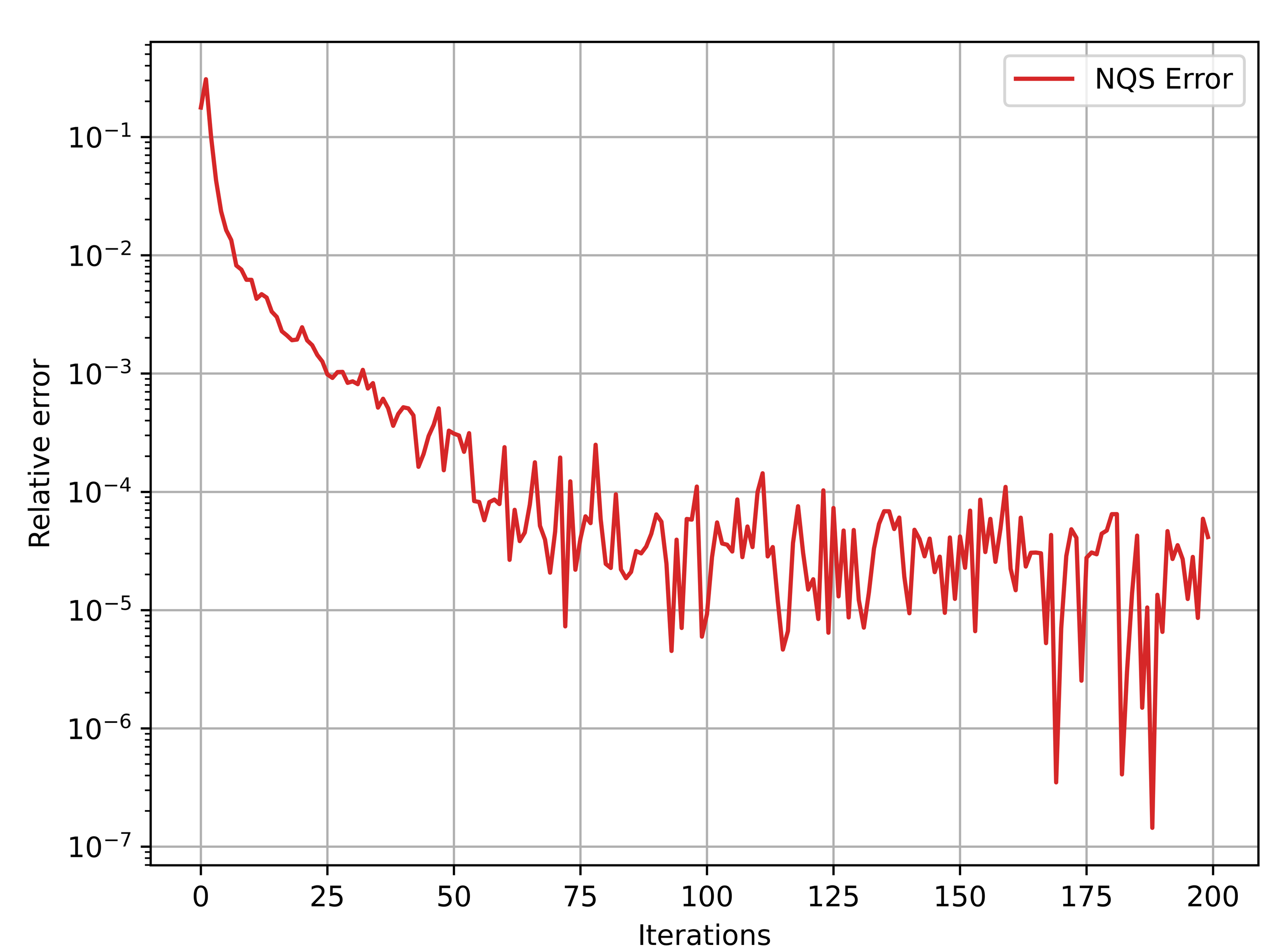

Fig. 3: Approximate energy of the NQS with relative error computed against the exact ground state energy as found by MPS.

To verify the translational-symmetry of our resulting state, we can look at the spin-spin correlations $\langle \sigma_i^\alpha \sigma_j^\alpha \rangle$ in some axis α. Consider a chain lattice of odd length, and fix i to be the middle site. Then for any eigenstate of $\mathcal{H}$ we expect

$$\langle \sigma_{mid}^\alpha \sigma_{mid+j}^\alpha \rangle = \langle \sigma_{mid}^\alpha \sigma_{mid-j}^\alpha \rangle \tag{6}$$

We show here the result of computing these observables on the same system described above:

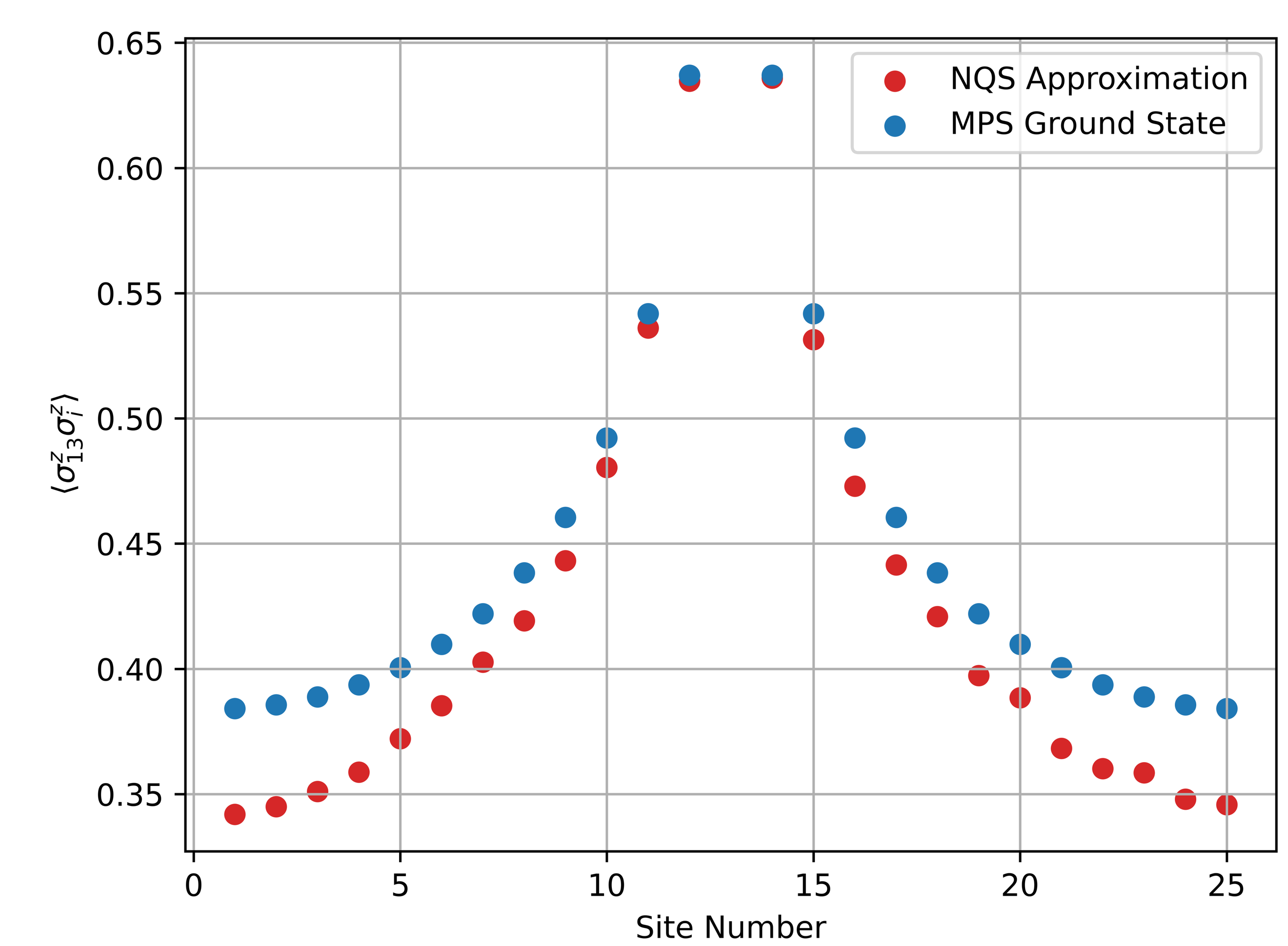

Fig. 4: Spin-spin correlations in the z-axis. We expect such a graph to be symmetric, which is demonstrated by the true ground state (blue). The result of the NQS is shown in red.

Remarks

Neural-network quantum states are a relatively recent advancement that have been shown to be very promising and lacking many of the downsides that come with other contemporary variational techniques. Possible architectures have expanded to include (among others) Recursive Neural Networks [2] and Convolutional Neural Networks [3], and have been applied to excited states and finite temperature states, to name a few [3].

References

[1] Giuseppe Carleo and Matthias Troyer. "Solving the quantum many-body problem with artificial neural networks". In: *Science* 355.6325 (Feb. 2017), pp. 602–606. ISSN: 1095-9203. DOI: 10.1126/science.aag2302. URL: http://dx.doi.org/10.1126/science.aag2302.

[2] Mohamed Hibat-Allah et al. "Recurrent neural network wave functions". In: *Phys. Rev. Res.* 2 (2 June 2020), p. 023358. DOI: 10.1103/PhysRevResearch.2.023358. URL: https://link.aps.org/doi/10.1103/PhysRevResearch.2.023358.

[3] Hannah Lange et al. *From Architectures to Applications: A Review of Neural Quantum States.* 2024. arXiv: 2402.09402 [cond-mat.dis-nn]. URL: https://arxiv.org/abs/2402.09402.

[4] 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 (July 2007). ISSN: 1089-7690. DOI: 10.1063/1.2746035. URL: http://dx.doi.org/10.1063/1.2746035.