Neural Network Representation of Tensor Network and Chiral States

Yichen Huang (黃溢辰)[1] and Joel E. Moore[2]

[1]Institute for Quantum Information and Matter
California Institute of Technology

[2]Department of Physics
University of California, Berkeley

KITS Workshop on Machine Learning and Many-Body Physics
July 6, 2017

Background

In machine learning, neural network is the most widely used method. It is a variational approach in that the learning process is to perform optimization over the parameters in the neural network.

Types of neural network: feedforward neural network, convolutional neural network, (restricted) Boltzmann machine, etc.

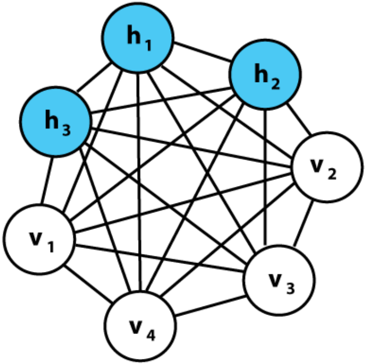
Recently, Carleo and Troyer simulated quantum many-body systems with a variational approach based on restricted Boltzmann machines.

The most popular variational approach in the study of many-body physics is based on tensor network states.

It is important to compare and understand the connection between the representational power of neural versus tensor network states.

Boltzmann machine (BM)

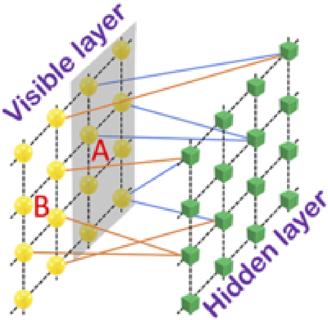
$$\sum_{\substack{\{s_1,s_2,\ldots,s_{|V|+|H|}\}\\ \in\{\pm1\}^{\otimes(|V|+|H|)}}} e^{-\sum_{j=1}^{|V|+|H|} h_j s_j - \sum_{j,k=1}^{|V|+|H|} w_{jk} s_j s_k} |\{s_1, s_2, \ldots, s_{|V|}\}\rangle$$



$V = \{v_1, v_2, \ldots v_{|V|}\}$ is the set of visible units carrying classical Ising variables $s_1, s_2, \ldots s_{|V|}$. $H = \{h_1, h_2, \ldots h_{|H|}\}$ is the set of hidden units carrying $s_{|V|+1}, s_{|V|+2}, \ldots, s_{|V|+|H|}$. Figure taken from `https://en.wikipedia.org/wiki/Boltzmann_machine`

Locality

$$\sum_{\substack{\{s_1,s_2,\ldots,s_{|V|+|H|}\}\\ \in\{\pm 1\}^{\otimes(|V|+|H|)}}} e^{-\sum_{j=1}^{|V|+|H|} h_j s_j - \sum_{j,k=1}^{|V|+|H|} w_{jk} s_j s_k} |\{s_1, s_2, \ldots, s_{|V|}\}\rangle$$



A restricted Boltzmann machine (RBM) is a BM such that $w_{jk} = 0$ for any edge that connects a visible unit $j \leq |V|$ with a hidden unit $k \geq |V| + 1$.
A (restricted) Boltzmann machine is local if there are only short-range connections.
Figure taken from Deng et al. PRX 7, 021021, 2017.

References and related work

G. Carleo and M. Troyer. Solving the quantum many-body problem with artificial neural networks. Science 355, 602, 2017.

References and related work

G. Carleo and M. Troyer. Solving the quantum many-body problem with artificial neural networks. Science 355, 602, 2017.

D.-L. Deng, X. Li, and S. Das Sarma. Exact machine learning topological states. arXiv:1609.09060; Quantum entanglement in neural network states. Phys. Rev. X 7, 021021, 2017.

J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang. On the equivalence of restricted Boltzmann machines and tensor network states. arXiv:1701.04831.

X. Gao and L.-M. Duan. Efficient representation of quantum many-body states with deep neural networks. arXiv:1701.05039.

Y. Huang and J. E. Moore. Neural network representation of tensor network and chiral states. arXiv:1701.06246.

Motivations

We study the representational power of Boltzmann machines.

In terms of representational power, how are neural network states compared with tensor network states? Do all tensor network states have neural network representations?

How much computational time does it take to convert a tensor network state into a neural network state? How much space does such a neural network representation occupy?

Can we find a physically interesting class of states, which have neural network representations, but are hard (or even impossible) to describe using tensor networks?

Neural and tensor network representations

We provide a polynomial-time algorithm that constructs a (local) neural network representation for any (local) tensor network state.[1]

The construction is almost optimal: the number of parameters in the neural network representation is almost linear in the number of nonzero parameters in the tensor network representation.

---

[1] See X. Gao and L.-M. Duan, arXiv:1701.05039 for a similar result.

[2] N. Le Roux and Y. Bengio. Neural Comput. 20, 1631, 2008.

Neural and tensor network representations

We provide a polynomial-time algorithm that constructs a (local) neural network representation for any (local) tensor network state.[1]

The construction is almost optimal: the number of parameters in the neural network representation is almost linear in the number of nonzero parameters in the tensor network representation.

Example: Any n-qubit matrix product state of bond dimension D has a local neural network representation with $2nD^2$ hidden units and $4nD^2 \log_2 D$ parameters.

[1] See X. Gao and L.-M. Duan, arXiv:1701.05039 for a similar result.

[2] N. Le Roux and Y. Bengio. Neural Comput. 20, 1631, 2008.

Neural and tensor network representations

We provide a polynomial-time algorithm that constructs a (local) neural network representation for any (local) tensor network state.[1]

The construction is almost optimal: the number of parameters in the neural network representation is almost linear in the number of nonzero parameters in the tensor network representation.

Example: Any n-qubit matrix product state of bond dimension D has a local neural network representation with $2nD^2$ hidden units and $4nD^2 \log_2 D$ parameters.

Using the universal approximation theorem, we convert every tensor in the network to a RBM. Then, we contract these RBMs in the same way as the tensors are contracted.

Universal approximation theorem[2]: Any tensor M can be arbitrarily well approximated by a RBM, provided that the number of hidden units is the number of nonzero elements in M.

[1]See X. Gao and L.-M. Duan, arXiv:1701.05039 for a similar result.

[2]N. Le Roux and Y. Bengio. Neural Comput. 20, 1631, 2008.

Tensor network representation of chiral states

Early attempts did not impose locality[3] or just target at expectation values of local observables rather than the wave function.[4]

Recent progress[5] shows that local tensor networks can describe chiral topological states, but the examples there are all gapless.

Dubail and Read's no-go theorem: For any chiral local free-fermion tensor network state, any local parent Hamiltonian is gapless.

[3] Z.-C. Gu, F. Verstraete, and X.-G. Wen. arXiv:1004.2563
[4] B. Beri and N. R. Cooper. Phys. Rev. Lett. 106, 156401, 2011.
[5] T. B. Wahl, H.-H. Tu, N. Schuch, and J. I. Cirac. Phys. Rev. Lett. 111, 236805, 2013; J. Dubail and N. Read. Phys. Rev. B 92, 205307, 2015.

Tensor network representation of chiral states

Early attempts did not impose locality[3] or just target at expectation values of local observables rather than the wave function.[4]

Recent progress[5] shows that local tensor networks can describe chiral topological states, but the examples there are all gapless.

Dubail and Read's no-go theorem: For any chiral local free-fermion tensor network state, any local parent Hamiltonian is gapless.

It seems difficult to obtain a local tensor network representation for gapped chiral topological states. This is an open problem in the community.

---

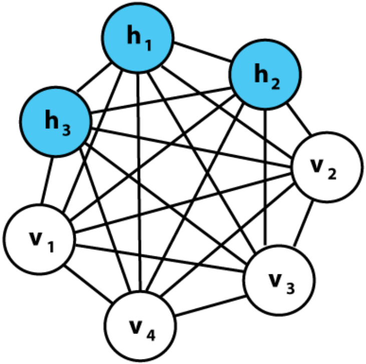
[3] Z.-C. Gu, F. Verstraete, and X.-G. Wen. arXiv:1004.2563

[4] B. Beri and N. R. Cooper. Phys. Rev. Lett. 106, 156401, 2011.

[5] T. B. Wahl, H.-H. Tu, N. Schuch, and J. I. Cirac. Phys. Rev. Lett. 111, 236805, 2013; J. Dubail and N. Read. Phys. Rev. B 92, 205307, 2015.

Fermionic neural network states

$$\left(\int \mathrm{d}\xi_{|V|+1} \mathrm{d}\xi_{|V|+2} \cdots \mathrm{d}\xi_{|V|+|H|} e^{-\sum_{j,k=1}^{|V|+|H|} w_{jk} \xi_j \xi_k} \right) |0\rangle$$



$V = \{v_1, v_2, \ldots v_{|V|}\}$ is the set of visible units carrying Grassmann numbers $\xi_1, \xi_2, \ldots \xi_{|V|}$.
$H = \{h_1, h_2, \ldots h_{|H|}\}$ is the set of hidden units carrying $\xi_{|V|+1}, \xi_{|V|+2}, \ldots, \xi_{|V|+|H|}$.
We identify $\xi_1, \xi_2, \ldots \xi_{|V|}$ with $c_1^\dagger, c_2^\dagger, \ldots, c_{|V|}^\dagger$.
Locality in fermionic BMs can be defined in the same way as in spin systems.

$p + ip$ superconductor

$$H = \sum_{\vec{x}\in\mathbf{Z}^2} (c^\dagger_{\vec{x}+\vec{i}} c_{\vec{x}} + c^\dagger_{\vec{x}+\vec{j}} c_{\vec{x}} + c^\dagger_{\vec{x}+\vec{i}} c^\dagger_{\vec{x}} + i c^\dagger_{\vec{x}+\vec{j}} c^\dagger_{\vec{x}} + \mathrm{h.c.}) - 2\mu \sum_{\vec{x}\in\mathbf{Z}^2} c^\dagger_{\vec{x}} c_{\vec{x}}$$

$$= \int_{\mathrm{BZ}} \mathrm{d}\vec{k} H_{\vec{k}}, \quad H_{\vec{k}} = \Delta_{\vec{k}} c^\dagger_{\vec{k}} c^\dagger_{-\vec{k}} + \Delta^*_{\vec{k}} c_{-\vec{k}} c_{\vec{k}} + 2M_{\vec{k}} c^\dagger_{\vec{k}} c_{\vec{k}}$$

$$\Delta_{\vec{k}} = \sin k_x + i \sin k_y, \quad M_{\vec{k}} = \cos k_x + \cos k_y - \mu$$

$$E_{\vec{k}} = 2\sqrt{|\Delta_{\vec{k}}|^2 + M^2_{\vec{k}}}, \quad |g.s.\rangle = e^{\frac{1}{2}\int_{\mathrm{BZ}} \mathrm{d}\vec{k} \frac{2\Delta_{\vec{k}}}{E_{\vec{k}} - 2M_{\vec{k}}} c^\dagger_{\vec{k}} c^\dagger_{-\vec{k}}} |0\rangle$$

This model represents topological superconductors with opposite chirality for $-2 < \mu < 0$ and $0 < \mu < 2$, respectively.

Neural network representation of chiral states

Approximate representations are acceptable, and we have to work with a finite system size in order to rigorously quantify the error of the approximation and the succinctness of the representation.

Example: A gapped ground state in a chain of n spins can be approximated (in the sense of 99% fidelity) by a matrix product state with bond dimension $2^{\tilde{O}(\log^{3/4} n)}$.[6]

[6] I. Arad, A. Kitaev, Z. Landau, and U. Vazirani. arXiv:1301.1162.

Neural network representation of chiral states

Approximate representations are acceptable, and we have to work with a finite system size in order to rigorously quantify the error of the approximation and the succinctness of the representation.

Example: A gapped ground state in a chain of n spins can be approximated (in the sense of 99% fidelity) by a matrix product state with bond dimension $2^{\tilde{O}(\log^{3/4} n)}$.[6]

On a square lattice of size $n \times n$, we construct a $O(\log n)$-local neural network state $|n.n.s.\rangle$ with one visible and one hidden unit per site such that

$$|\langle n.n.s|g.s.\rangle| > 1 - 1/\operatorname{poly} n.$$

The same result can be established for the ground state of any translationally invariant gapped free-fermion systems.

[6] I. Arad, A. Kitaev, Z. Landau, and U. Vazirani. arXiv:1301.1162.

Summary

We provide a polynomial-time algorithm that constructs a (local) neural network representation for any (local) tensor network state.

The construction is almost optimal in the sense that the number of parameters in the neural network representation is almost linear in the number of nonzero parameters in the tensor network representation.

Summary

We provide a polynomial-time algorithm that constructs a (local) neural network representation for any (local) tensor network state.

The construction is almost optimal in the sense that the number of parameters in the neural network representation is almost linear in the number of nonzero parameters in the tensor network representation.

Despite the difficulty of representing (gapped) chiral topological states with local tensor networks, we construct a quasi-local neural network representation for a chiral p-wave superconductor.

This is an explicit and physically interesting example that neural network states may "go beyond" tensor network states.

Outlook

Numerical study of chiral topological order with neural network states.[7]

The parton approach may allow us to construct quasi-local neural network representation for strongly interacting chiral topological states (e.g., the fractional quantum Hall state).

What is the representational power of neural network states based on RBM?

[7] I. Glasser et al. The geometry of neural network states, string-bond states and chiral topological order. preceding talk in this workshop.