From Quantum Entanglement to Machine Learning

Jing Chen （陈靖）
IOP, CAS
yzcj105@126.com
arXiv:1701.04831
Outline

• Tensor Network

• Machine learning

• Connections
History

Wilson
NRG 1975

White
DMRG 1992

Tensor Network
Tensor Network

Represent wave functions

Matrix Product State (MPS)

DMRG wave function ansatz

\[ \Psi_{\text{MPS}}(v) = \text{Tr} \prod_i A^{(i)}[v_i], \]

low rank approximation

Very successful in 1D
MPS expression of AKLT state

\[ \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \]

\[ |\psi\rangle = Tr \left( \cdots \hat{A}[m_i] \hat{A}[m_{i+1}] \cdots \right) |\cdots m_i m_{i+1} \cdots \rangle \]

\[ A[+] = \sqrt{\frac{2}{3}} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \]
\[ A[0] = \sqrt{\frac{1}{3}} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \]
\[ A[-] = \sqrt{\frac{2}{3}} \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix} \]
Entanglement (Area Law)

- gapped systems ground state

\[ S = -\text{Tr}_e (\rho_{es} \log \rho_{es}) \]

J. Eisert, et.al. MB Plenio RMP, 2010
1D -> 2D DMRG

\[ S \sim \ln D \quad \text{vs} \quad S \sim W \ln D \]

\[ D \sim \text{Const} \quad \text{vs} \quad D \sim \exp(W) \]

More challenge due to large entanglement
2D extension: PEPS

\[ S \leq n \ln D \]

- Fullfill the Area Law
Tensor Network Methods

Partition Function
- Course graining
  - TRG, SRG, HOTRG, HOSRG
  - TNR loop-TNR ...

Quantum Wave Function
- Projection
  - TEBD, CTMRG
- Variation
  - DMRG, PEPS,
Coarse graining e.g. HOTRG

\[
A = \begin{bmatrix}
e^\beta & e^{-\beta} \\
e^{-\beta} & e^\beta
\end{bmatrix}
\]

\[
A = WW^\dagger
\]

Z. Y. Xie PhD thesis
Coarse graining e.g. HOTRG

Coarse graining e.g. HOTRG

Projection Method: TEBD

\[ |\psi_G\rangle = \lim_{\beta \to \infty} e^{-\beta H} |\psi_G\rangle = \lim_{N \to \infty} (e^{-\tau H})^N |\psi_0\rangle \]
\[ = \lim_{N \to \infty} (e^{-\tau H_A} e^{-\tau H_B})^N |\psi_0\rangle + o(\tau^2) \]

Projection Method: CTMRG

• $|\psi_G\rangle = \lim_{\beta \to \infty} e^{-\beta H} |\psi_G\rangle = \lim_{N \to \infty} (e^{-\tau H})^N |\psi_0\rangle$

• $= \lim_{N \to \infty} (e^{-\tau H_A} e^{-\tau H_B})^N |\psi_0\rangle + o(\tau^2)$

Variational wave function

\[ \min \left( \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \right) \]
MERA

$S \approx L \ln L$

- The entanglement grows faster than area law
- Excellent in critical point
- Power law correlation decay

MERA and AdS-CFT: Holography

A tool to simulate the problem in AdS

Glen Evenbly Phys. Rev. Lett. 119, 141602
Outline

• Tensor Network

• Machine learning

• Connections
Machine learning

Image Recognition

AlphaGo

Driverless Car

Finance

Amazon recommender
Five schools of ML

- Symbolists
- Connectionists
- Analogizes
- Bayesian
- Evolutionaries
Machine Learning 101

Supervised learning

Classification
- Spam detection
- Image recognition

Unsupervised learning

Clustering
- Online advertising
- Anomaly detection

Generative learning
Neural Network
Neural Network

Playground
Zoo of Neural Network

Philosophy:

Connectionism → Intelligence
Restricted Boltzmann Machine (RBM)

\[(v, h) = - \sum_i a_i v_i - \sum_j b_j h_j - \sum_{i,j} v_i W_{ij} h_j\]

\[P(v, h) = \frac{1}{\mathcal{Z}} e^{-E(v, h)}\]

Theano deep learning tutorial  
http://www.deeplearning.net/tutorial/rbm.html#rbm
Universal approximation theorem

Formal statement [edit]

The theorem\cite{2} in mathematical terms:

Let $\varphi(\cdot)$ be a nonconstant, bounded, and monotonically-increasing continuous function. Let $I_m$ denote the $m$-dimensional unit hypercube $[0, 1]^m$. The space of continuous functions on $I_m$ is denoted by $C(I_m)$. Then, given any function $f \in C(I_m)$ and $\varepsilon > 0$, there exists an integer $N$, real constants $v_i, b_i \in \mathbb{R}$ and real vectors $w_i \in \mathbb{R}^m$, where $i = 1, \ldots, N$, such that we may define:

$$F(x) = \sum_{i=1}^{N} v_i \varphi (w_i^T x + b_i)$$

as an approximate realization of the function $f$ where $f$ is independent of $\varphi$; that is,

$$|F(x) - f(x)| < \varepsilon$$

for all $x \in I_m$. In other words, functions of the form $F(x)$ are dense in $C(I_m)$. 

Wikipedia
Universal approximation theorem

- Simulate any function with enough units

![Diagram of a neural network with an input layer, a hidden layer with 15 neurons, and an output layer.](image)
Machine Learning

• Can be generalized to other problems. E.g. AlphaGo
• Learn features automatically

• Needs much data.
• Little theoretical analysis
• Requires powerful hardware. GPU and TPU
Outline

• Tensor Network

• Machine learning

• Connections
ML applied to physical problems

• Material and Chemistry Discovery
• Density Functional Theory
• Phase Transitions
• Representing Quantum States
• Quantum Information and Computation
• Algorithmic Innovations

http://wangleiphy.github.io/mlrefs.html
Learning Classical Statistic Distribution by RBM

The result is not very good at $T_c$

Can RBM represent the distribution well at criticality?

“Learning Thermodynamics with Boltzmann Machines”

Accelerated Monte Carlo simulations with restricted Boltzmann machines
L Huang, L Wang Phys. Rev. B 95, 035105
Quantum: RBM as wave function ansatz

\[ \Psi_M(S; W) = \sum_{\{h_i\}} e^{\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} W_{ij} h_i \sigma_j^z} \]

Complex \( W, a, b \)

RBM as wave function ansatz


A Neural Decoder for Topological Codes, Giacomo Torlai, Roger G. Melko, arxiv:1610.04238

Many-body quantum state tomography with neural networks, Giacomo Torlai, Guglielmo Mazzola, Juan Carrasquilla, Matthias Troyer, Roger Melko, Giuseppe Carleo, arxiv:1703.05334
• How is the expressive power of RBM?
• Does RBM satisfy the area law?
• Can RBM represent critical possibility distribution?
• Why is RBM wave function successful?

arXiv:1701.04831
Zoo of Tensor Network State
TNS representation of RBM

This is the direct way, but not the most efficient way

\[ D = 2^n \]

\( n \) is # of cuts
Get the optimal tensor on the fly

Can be generalized to other models.
The entanglement entropy of RBM

The entanglement depends on the size of $B_1$

Code: [https://github.com/yzcj105/rbm2mps](https://github.com/yzcj105/rbm2mps)

Good news: Much fewer Variables
Entanglement of shift-invariant RBM

The shift-invariant RBM structure is crucial to the success.

3 orders of magnitude fewer variational variables than DMRG
<table>
<thead>
<tr>
<th>Example: 2D System</th>
<th>PEPS</th>
<th>RBM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long term interactions</td>
<td>Passed by the sites between, increase $D$</td>
<td>Connected directly</td>
</tr>
<tr>
<td>N-body interactions</td>
<td>Tensor with $D^N$ elements</td>
<td>$N$ weights</td>
</tr>
<tr>
<td>Sampling of the physical freedom</td>
<td>Contraction of a 2D TN</td>
<td>Just a summation in the exponent</td>
</tr>
<tr>
<td>Philosophy</td>
<td>Contraction</td>
<td>Product</td>
</tr>
</tbody>
</table>

RBM is a subset of TN theoretically but different practically.
RBM representation of a MPS

\[
\text{Tr} \prod_i A^{(i)} [v_i] = T_{v_1v_2v_3v_4}^{(1)} T_{v_2v_3v_4}^{(2)} T_{v_3v_4v_5}^{(3)} T_{v_4v_5v_6}^{(4)}.
\]

\[
T_{v_2v_3v_4}^{(2)} = \sum_{h_2 \in \{0,1\}} e^{h_2b_2 + \sum_{i \in \{2,3,4\}} v_i (W_{i2}h_2 + a_{i2})}.
\]
Explicit RBM of Ising Model

\[
W = \ln(4e^{4K} - 2)
\]
\[
a = -8K - 2H - 4\ln 2
\]
\[
b = -\ln(e^{4K} - 1) - 2\ln 2
\]

The RBM can represent Ising model at criticality!
Once the units of interface region are fixed, the wave function decouples.

No matter how long the sequences of 0 is, it is always the superposition.
Deep or shallow, is a question.

Same number of units and connections.

Deep BM allows more entanglement.
Entanglement and correlation of MNIST datasets

\[ \langle S_0 S_{\vec{r}} \rangle \quad \sum_{\vec{p}} \langle S_{\vec{p}} S_{\vec{r}} \rangle \]

For images, the correlation is local and anisotropic
Work related


Efficient Representation of Quantum Many-body States with Deep Neural Networks by X. Gao and L.-M. Duan, arXiv:1701.05039

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

Deep Learning and Quantum Physics: A Fundamental Bridge by Y. Levine, D. Yakira, etc. arxiv:1704.01552
Following work on RBM wave function

• R Kaubruegger et.al. arXiv:1710.04713
Summary

Machine Learning

Quantum Physics
Collaborators

Lei Wang 王磊
Song Cheng 程嵩
Haidong Xie 谢海东
Tao Xiang 向涛

Acknowledgement:

Giuseppe Carleo ETH Zurich
Dongling Deng 邓东灵 University of Maryland
Xiaopeng Li 李晓鹏 Fudan University
Chen Fang 方辰 Institute of Physics
Xun Gao 郜勋 Tsinghua University
E. Miles Stoudenmire UC Irvine
Hong-Hao Tu 涂鸿浩 Ludwig-Maximilians University Munich
Yi-feng Yang 杨义峰 IOP,CAS
Ivan Glasser MPIQO
Reference

• Y. Huang and J. E. Moore, arXiv:1701.06246