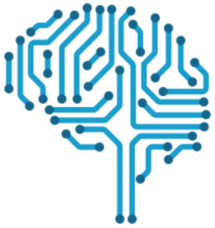
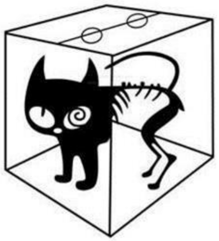
From Quantum Entanglement to Machine Learning

Jing Chen （陈靖）

IOP, CAS

yzcj105@126.com





arXiv:1701.04831

Outline

- Tensor Network
- Machine learning
- Connections

History

Wilson

NRG
1975

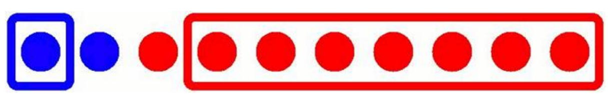
→

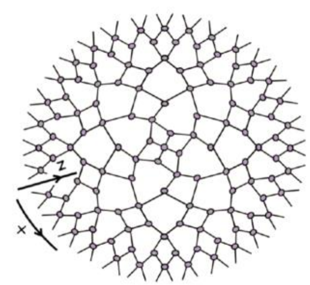
White

DMRG
1992

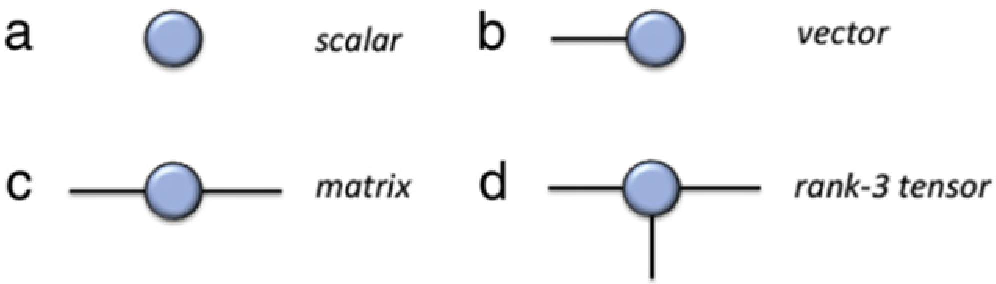
→

Tensor
Network



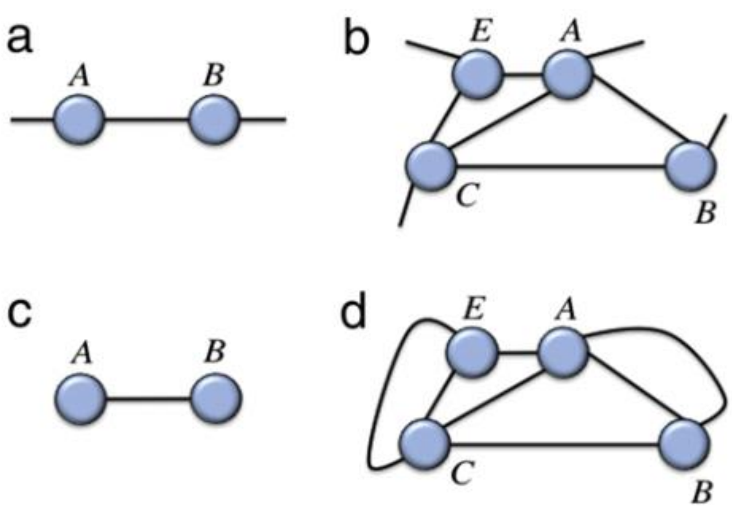


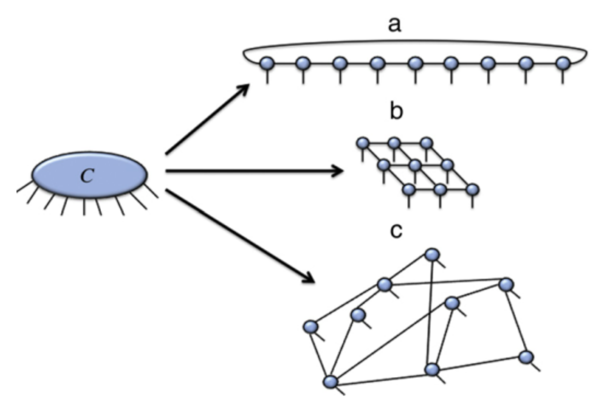
Tensor Network



matrix → Extend → tensor

product → contraction



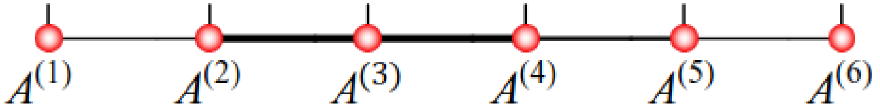


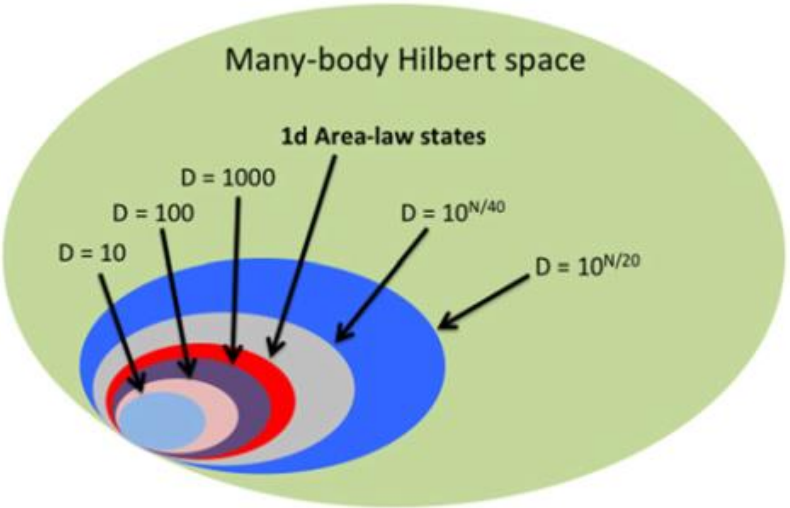
Represent wave functions

R. Orus Annals of Physics 349 (2014) 117-158

Matrix Product State (MPS)

DMRG wave function ansatz



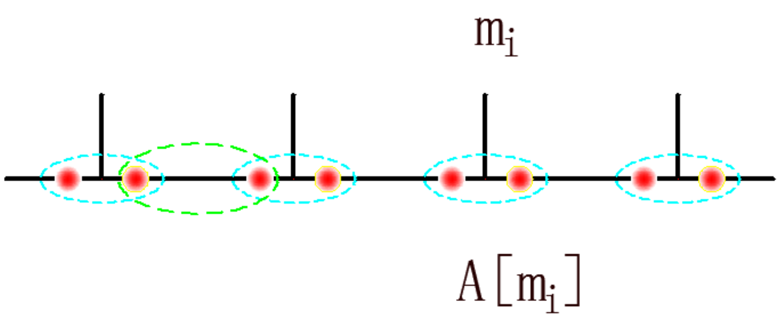


low rank
approximation

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

Very successful in 1D

MPS expression of AKLT state



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

$$A[+] = \sqrt{\frac{2}{3}}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

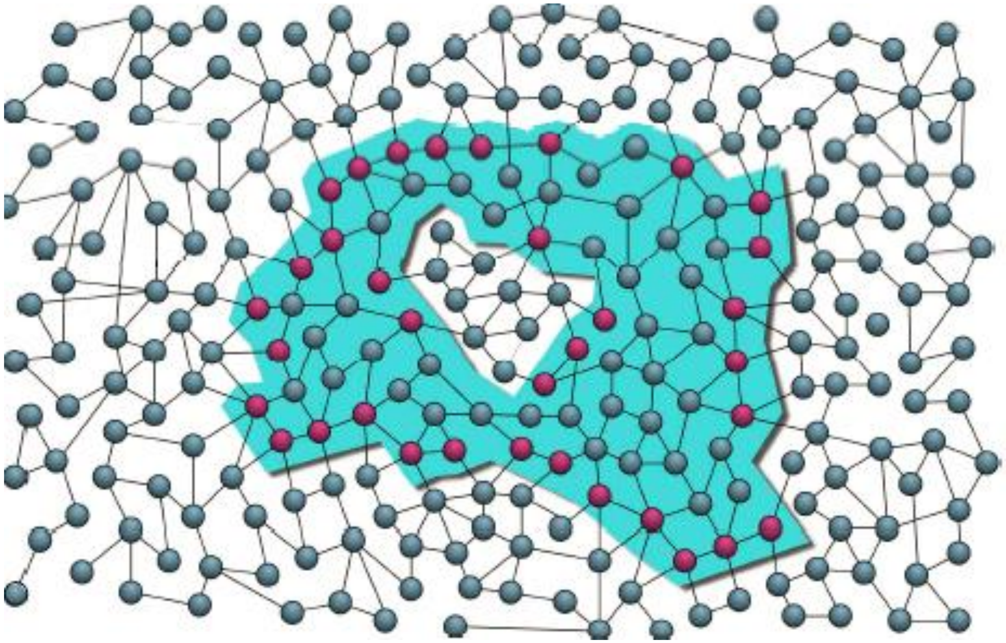
$$A[0] = \sqrt{\frac{1}{3}}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A[-] = \sqrt{\frac{2}{3}}\begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}$$

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

Phys. Rev. Lett. 59, 799

Entanglement (Area Law)

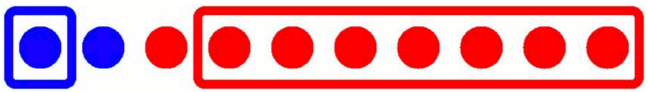


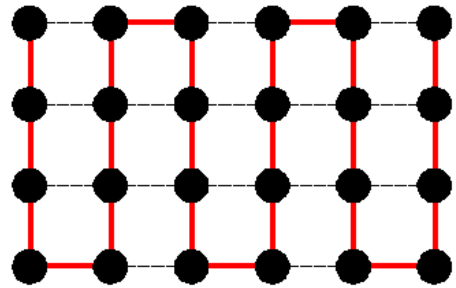
- gapped systems ground state

$$S = -Tr_e \left(\rho_{es} \log \rho_{es} \right)$$

J Eisert, et.al. MB Plenio RMP, 2010

1D -> 2D DMRG





$S \sim \ln D$

$S \sim W \ln D$

$D \sim Const$

$D \sim \exp(W)$

More challenge due to large entanglement

Large size

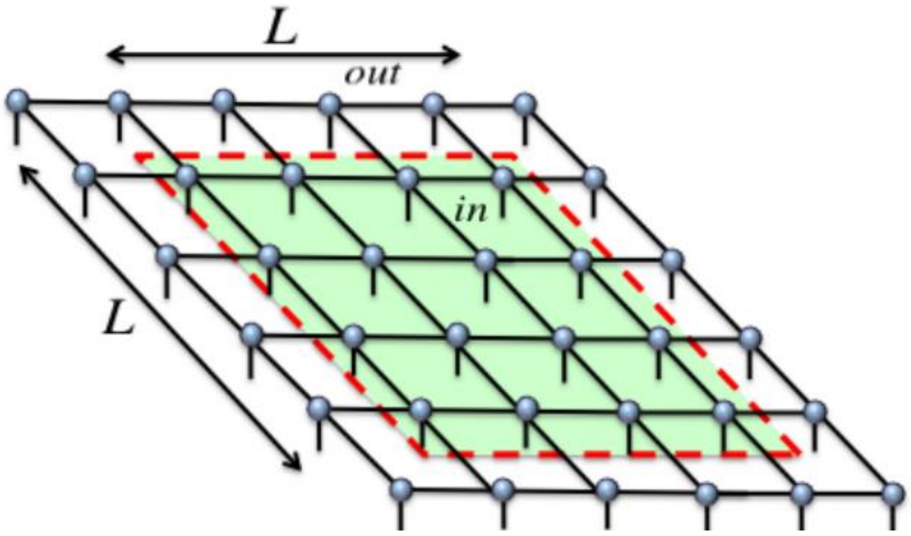
↓

Large D

↓

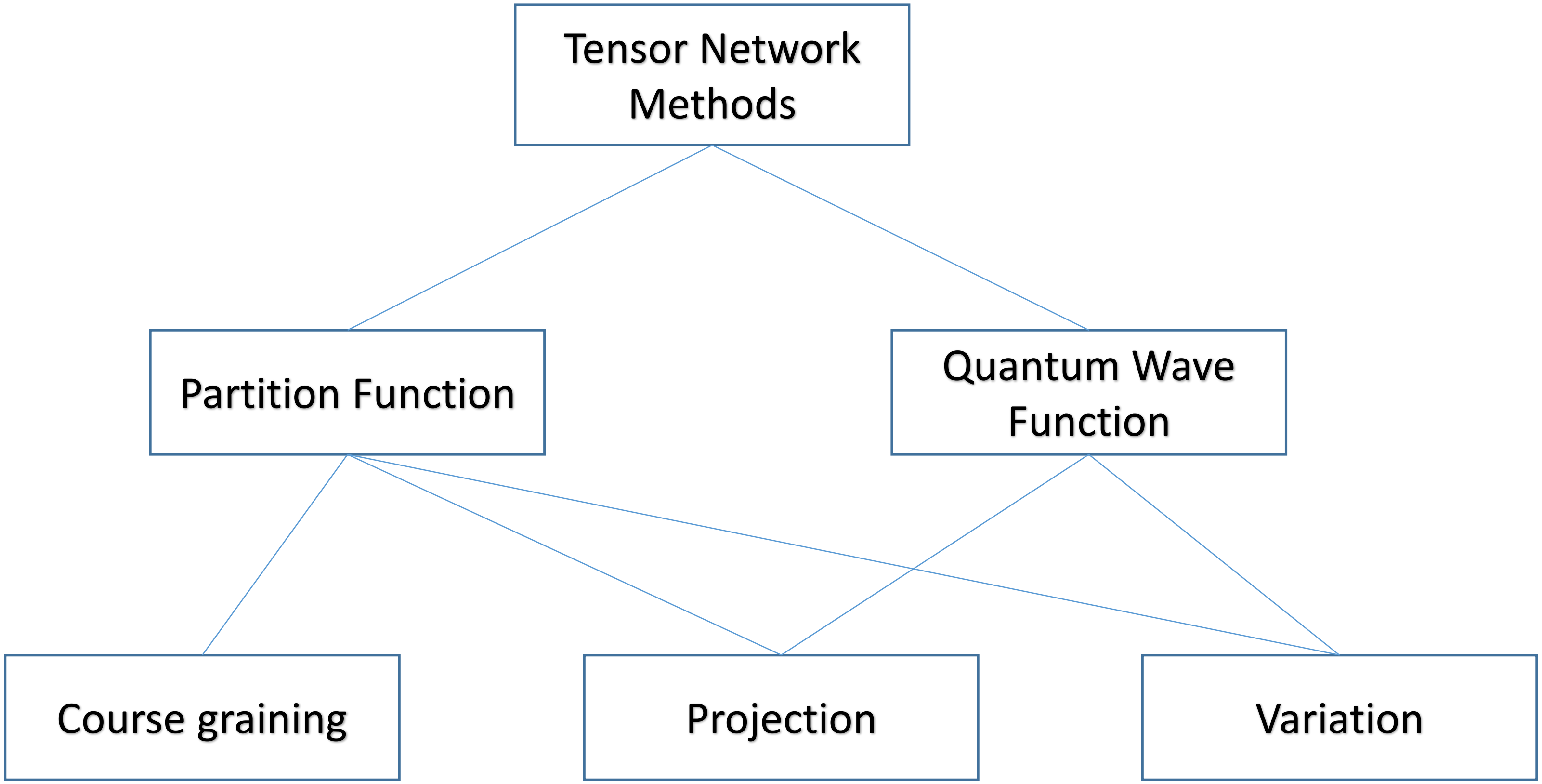
Difficulty in 2D

2D extension: PEPS



$$S \leq n \ln D$$

- Fullfill the Area Law

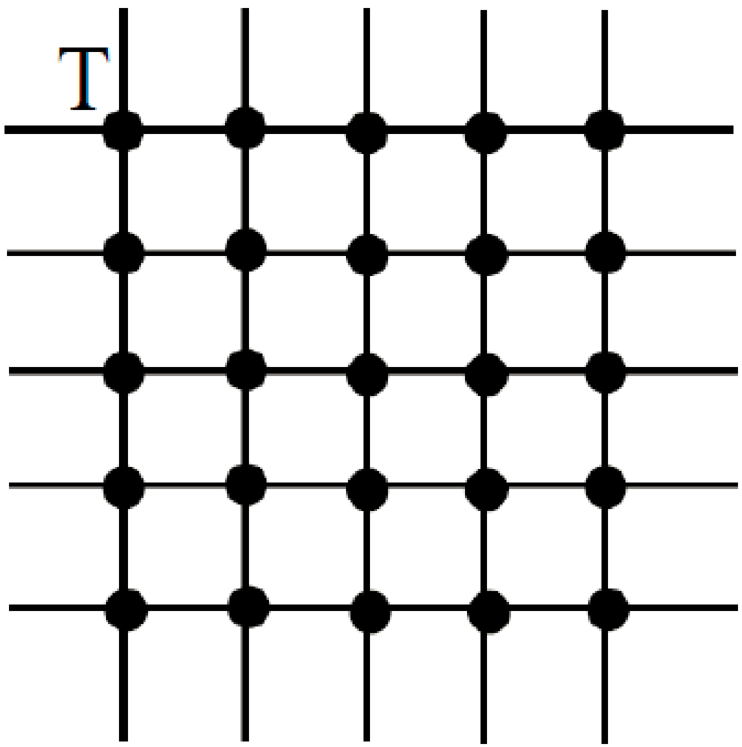


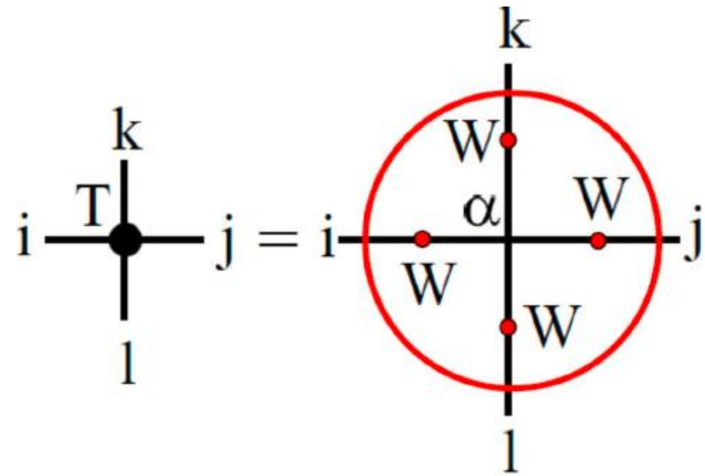
TRG, SRG,
HOTRG,HOSRG
TNR loop-TNR …

TEBD, CTMRG

DMRG, PEPS,

Coarse graining e.g. HOTRG



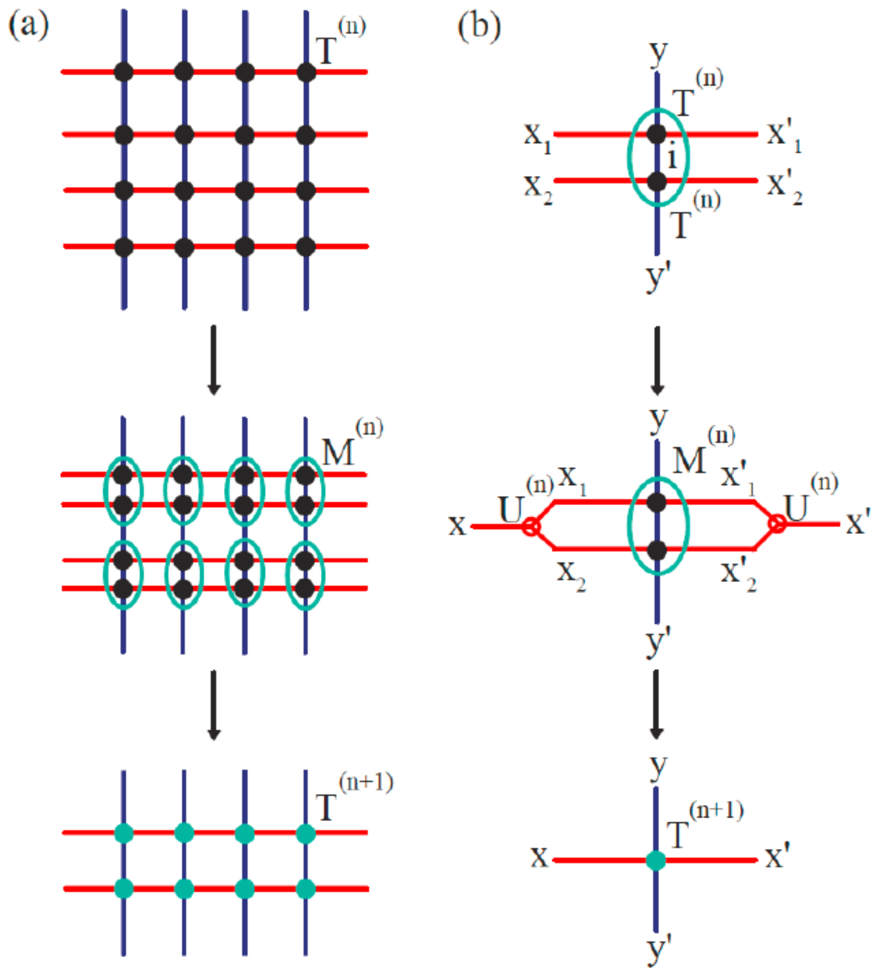


$$A = \begin{bmatrix} e^{\beta} & e^{-\beta} \\ e^{-\beta} & e^{\beta} \end{bmatrix}$$

$$A = WW^{\dagger}$$

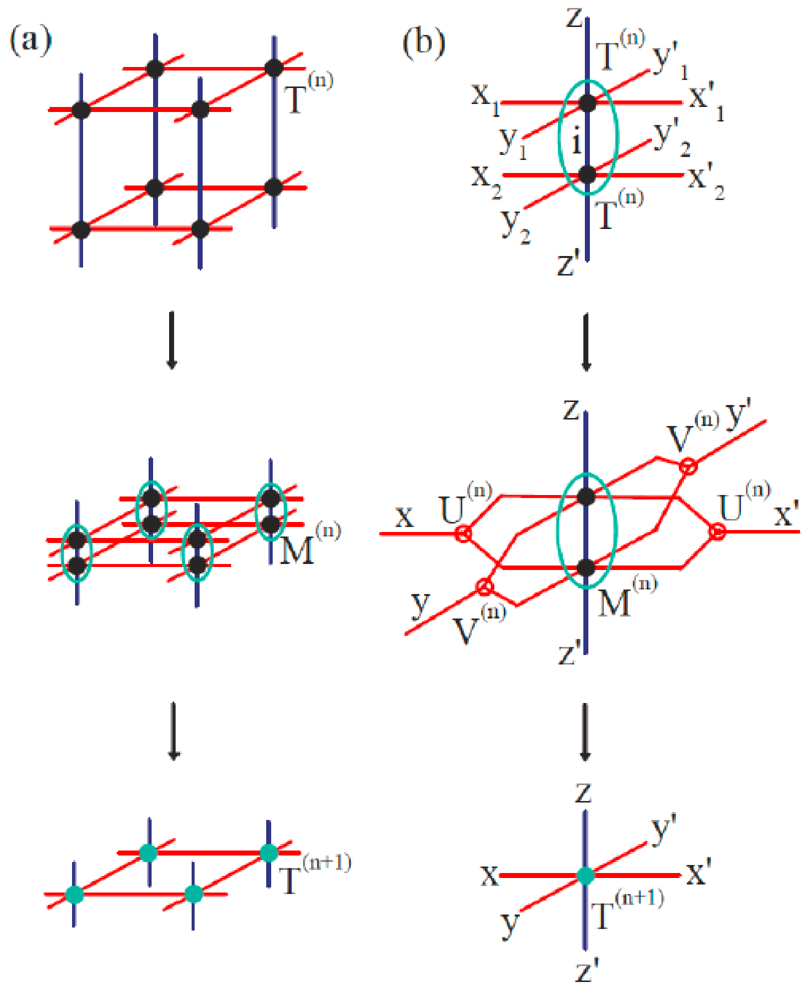
Z. Y. Xie PhD thesis

Coarse graining e.g. HOTRG



Z. Y. Xie et.al. Phys. Rev. B 86, 045139

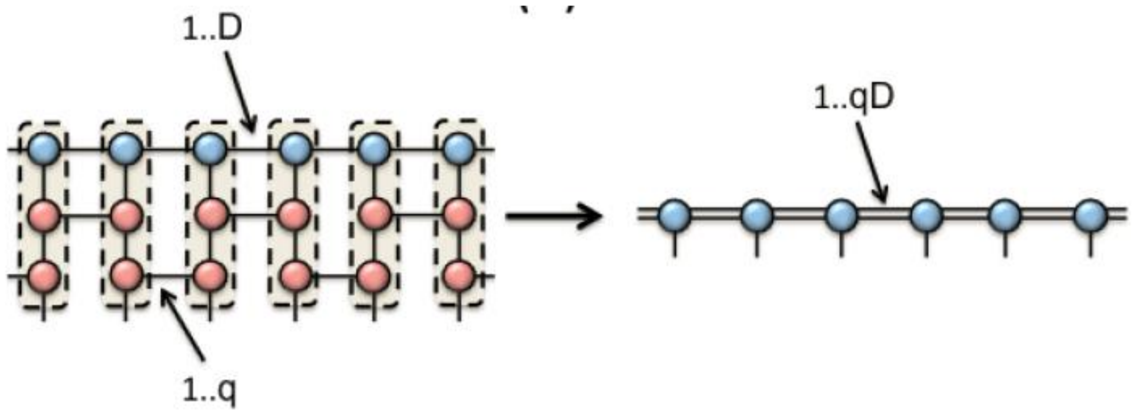
Coarse graining e.g. HOTRG



Z. Y. Xie et.al. Phys. Rev. B 86, 045139

Projection Method: TEBD

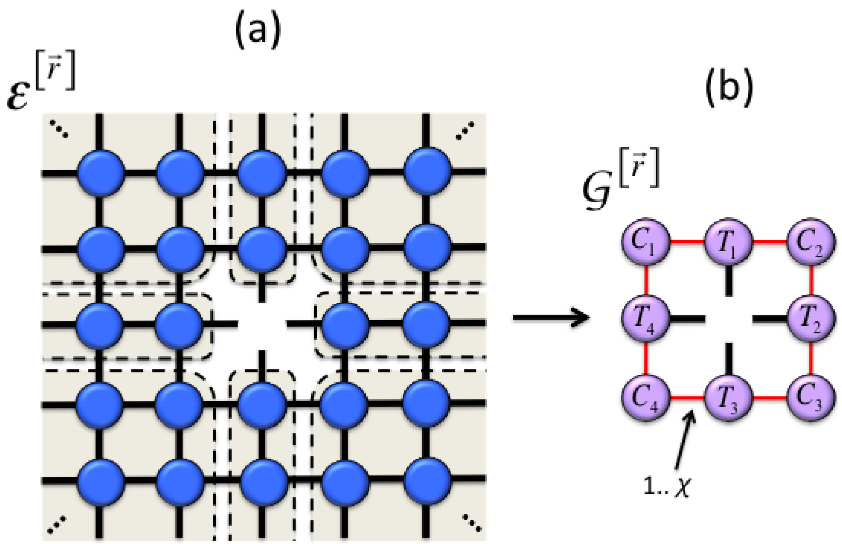
- $|\psi_G\rangle = \lim_{\beta\to\infty} e^{-\beta H}|\psi_G\rangle = \lim_{N\to\infty}(e^{-\tau H})^N|\psi_0\rangle$
- = $\lim_{N\to\infty}(e^{-\tau H_A}e^{-\tau H_B})^N|\psi_0\rangle + o(\tau^2)$



G. Vidal PRL 91, 147902 (2003)

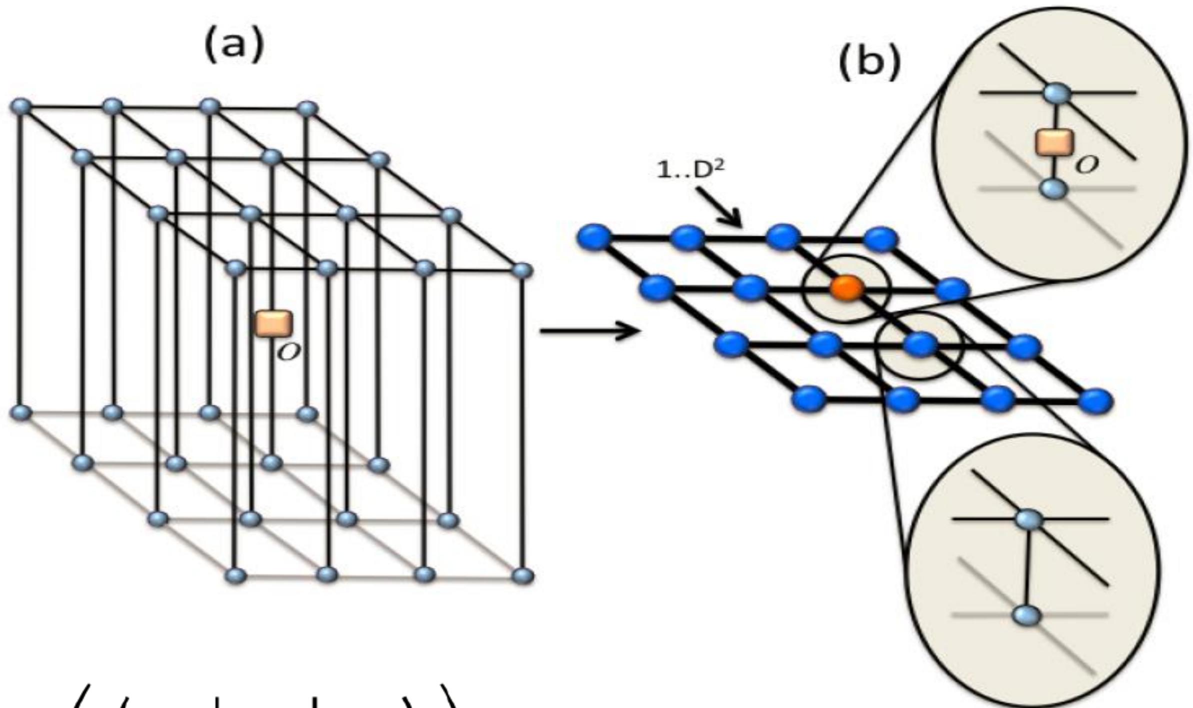
Projection Method: CTMRG

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



T. Nishino J. Phys. Soc. Jpn. 65, pp. 891-894 (1996)

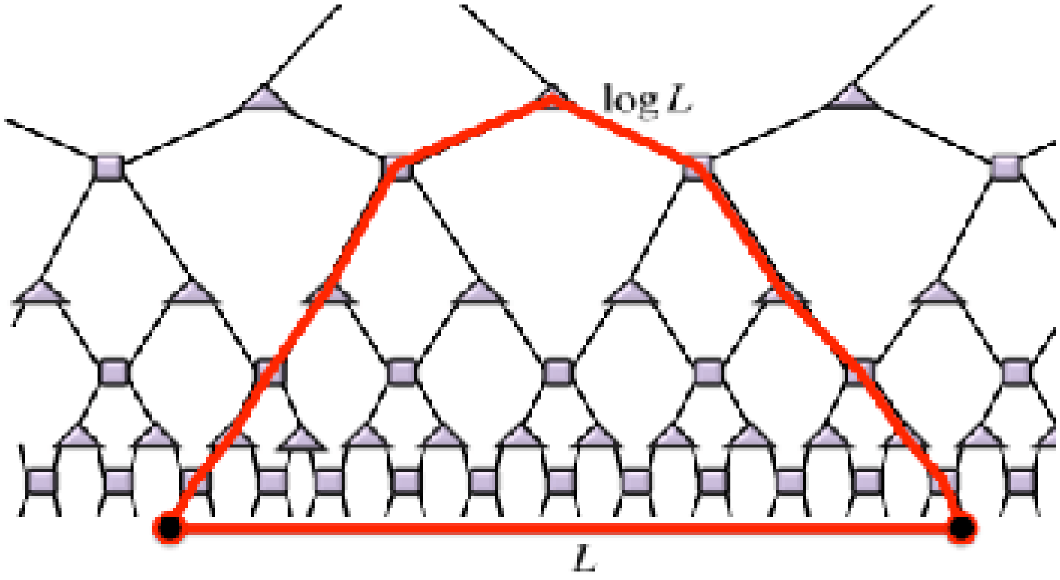
Variational wave function



$$\min\left(\frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle}\right)$$

R. Orus Annals of Physics 349 (2014) 117-158

MERA

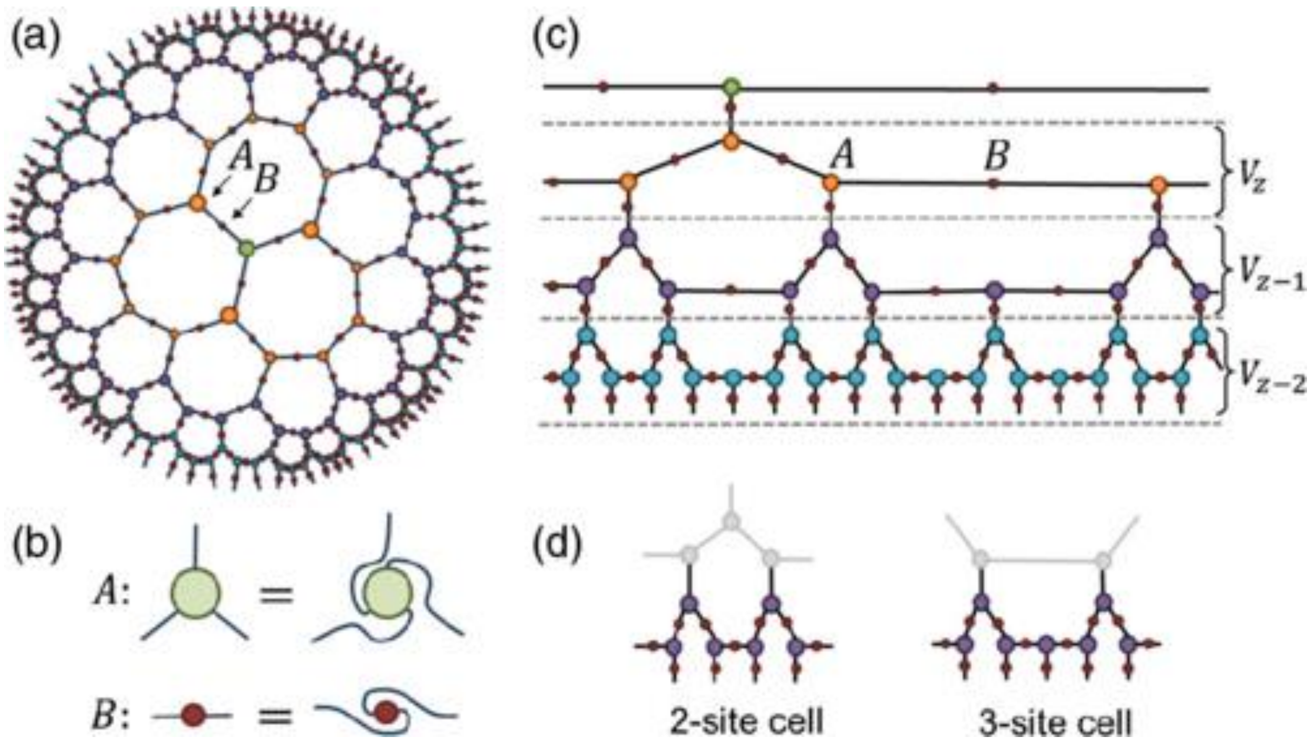


$$S \approx L \ln L$$

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

G. Vidal Phys. Rev. Lett. 99, 220405

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

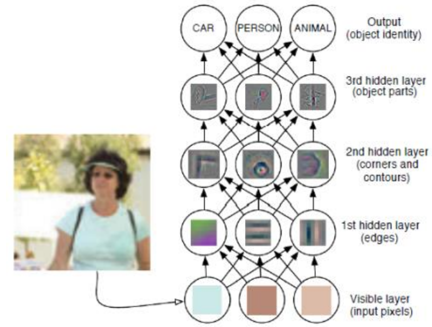


Image Recognition

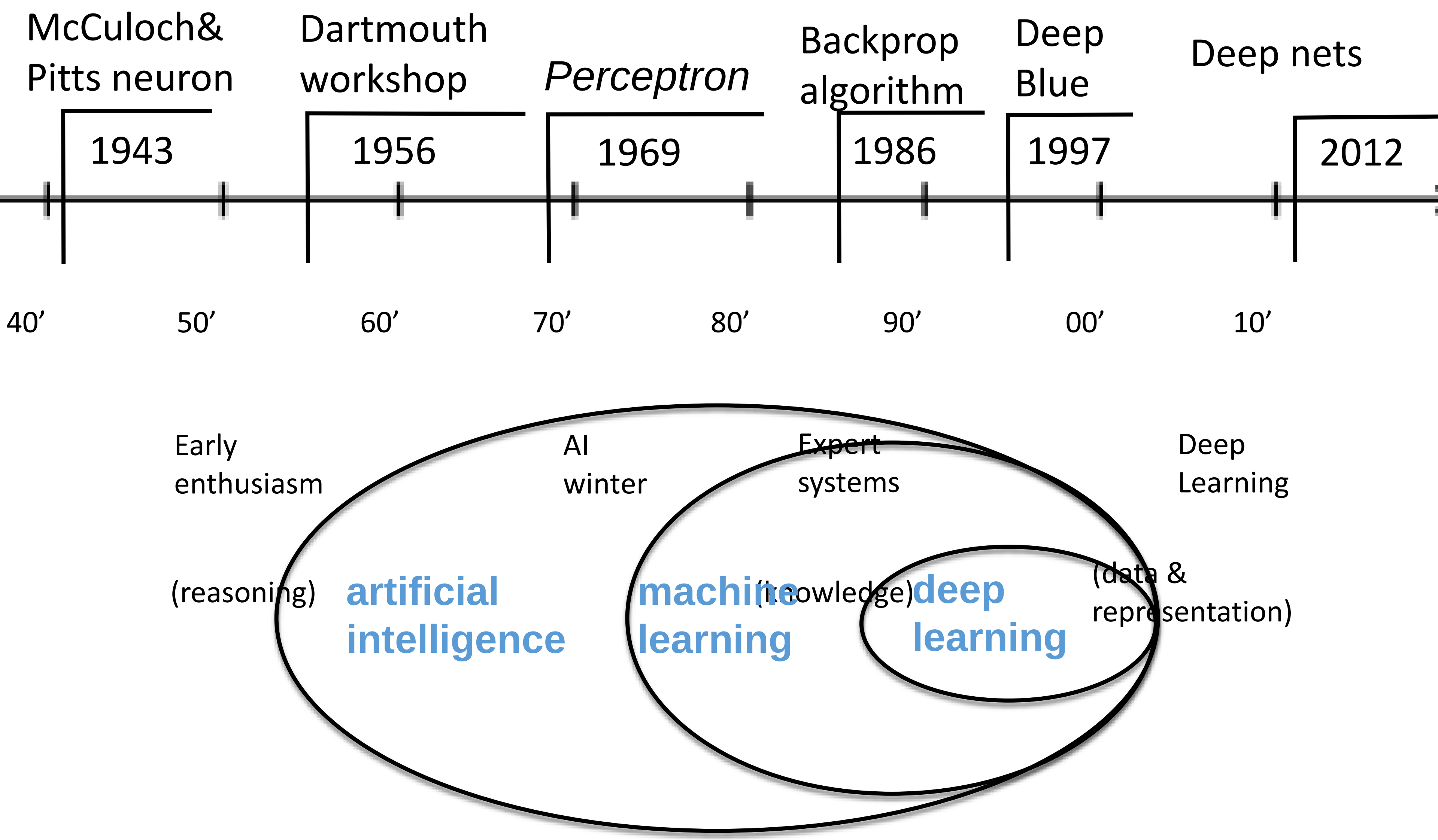
AlphaGo

Amazon recommender

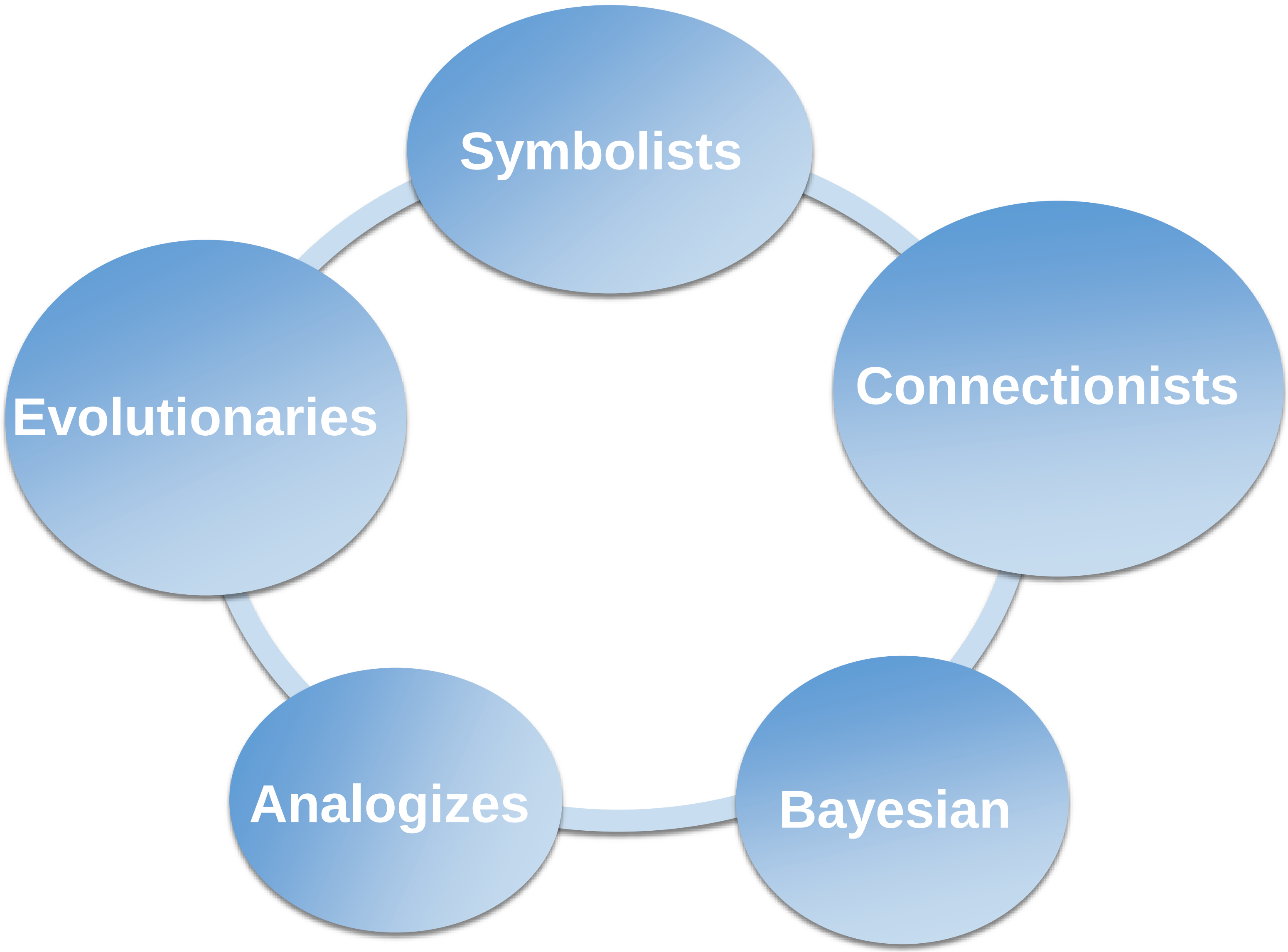
Driverless Car

Finance

Timeline of AI research

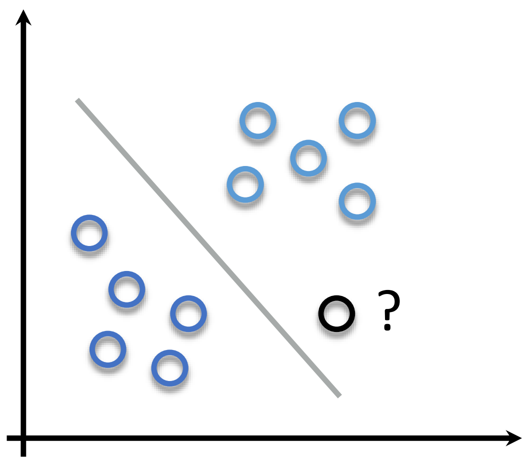


Five schools of ML



Machine Learning 101

Supervised learning

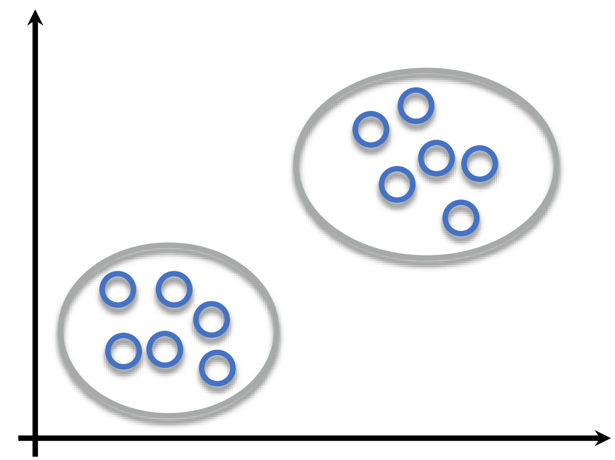


Classification

Spam detection

Image recognition

Unsupervised learning



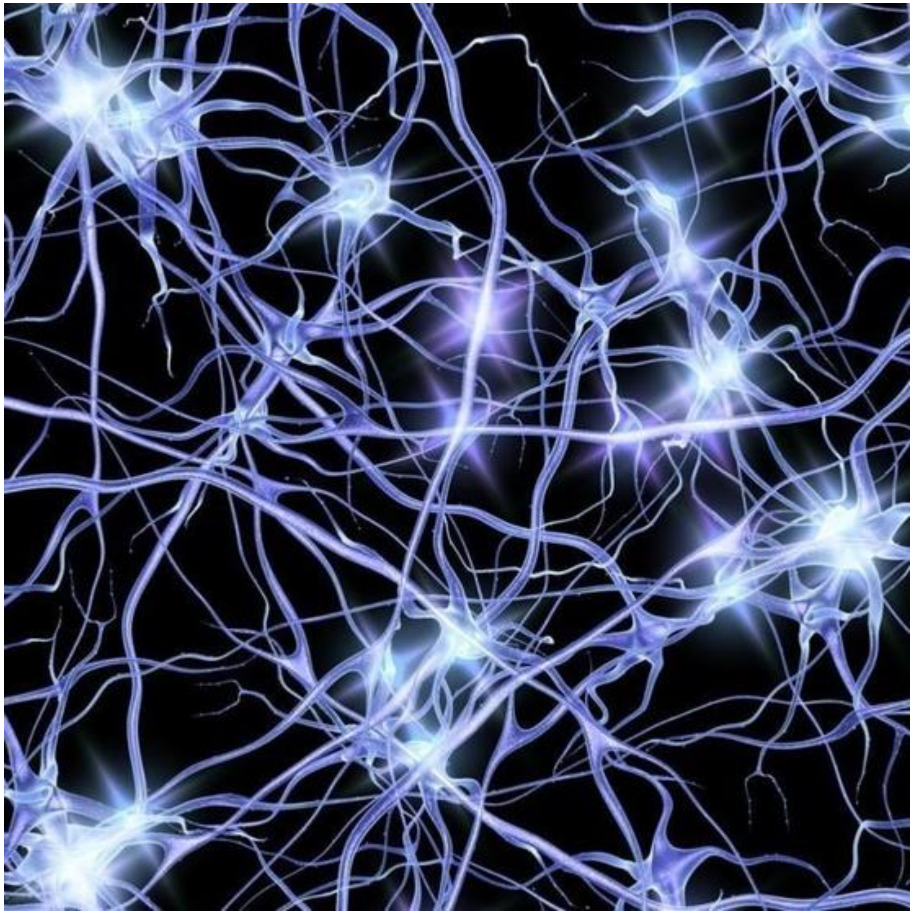
Clustering

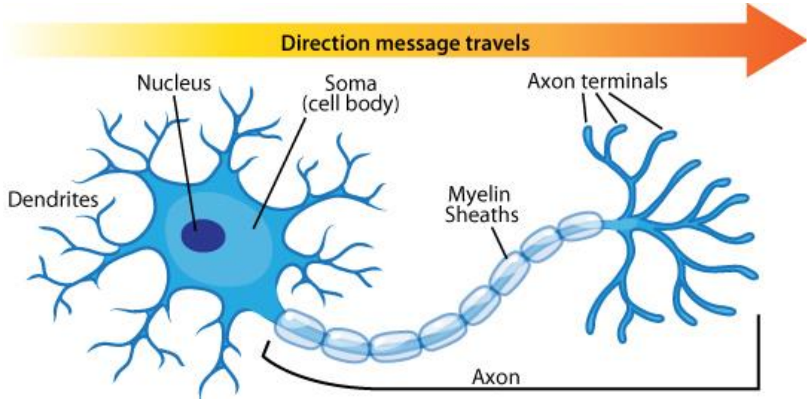
Online advertising

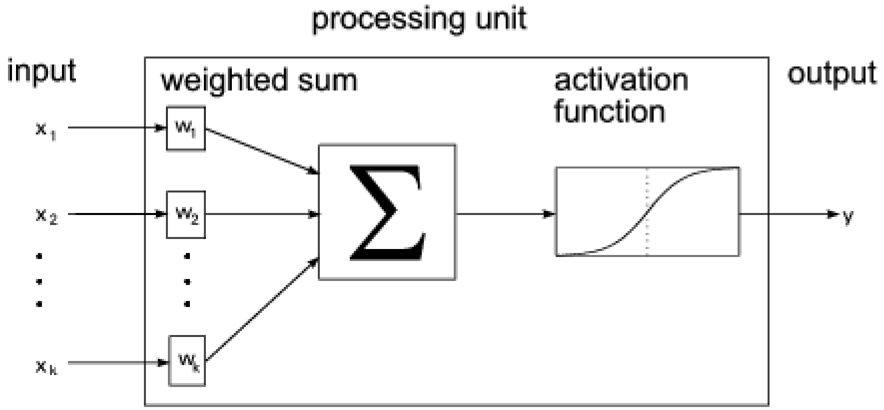
Anomaly detection

Generative learning

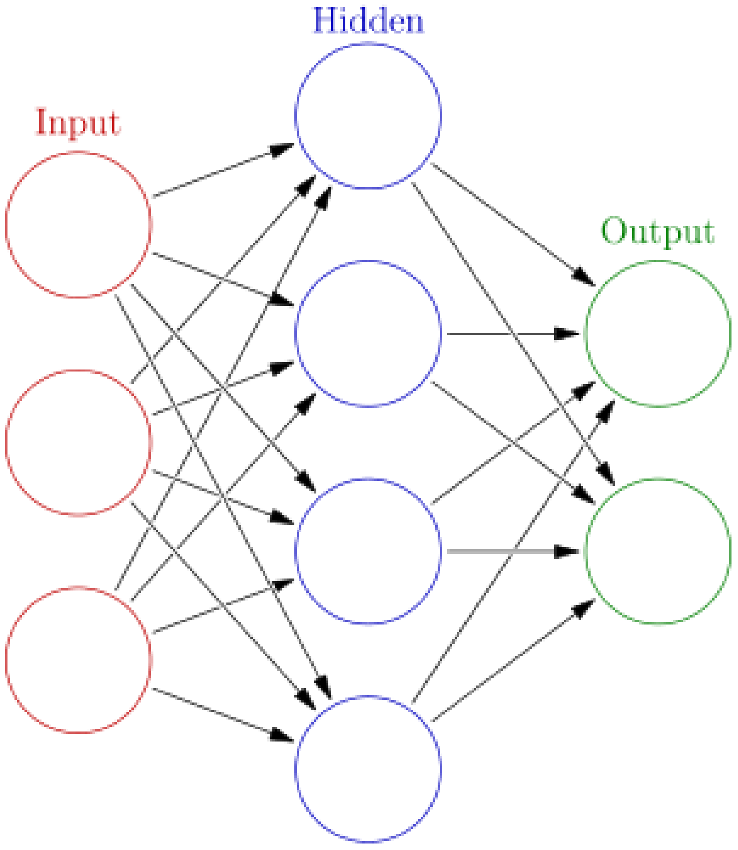
Neural Network







Neural Network

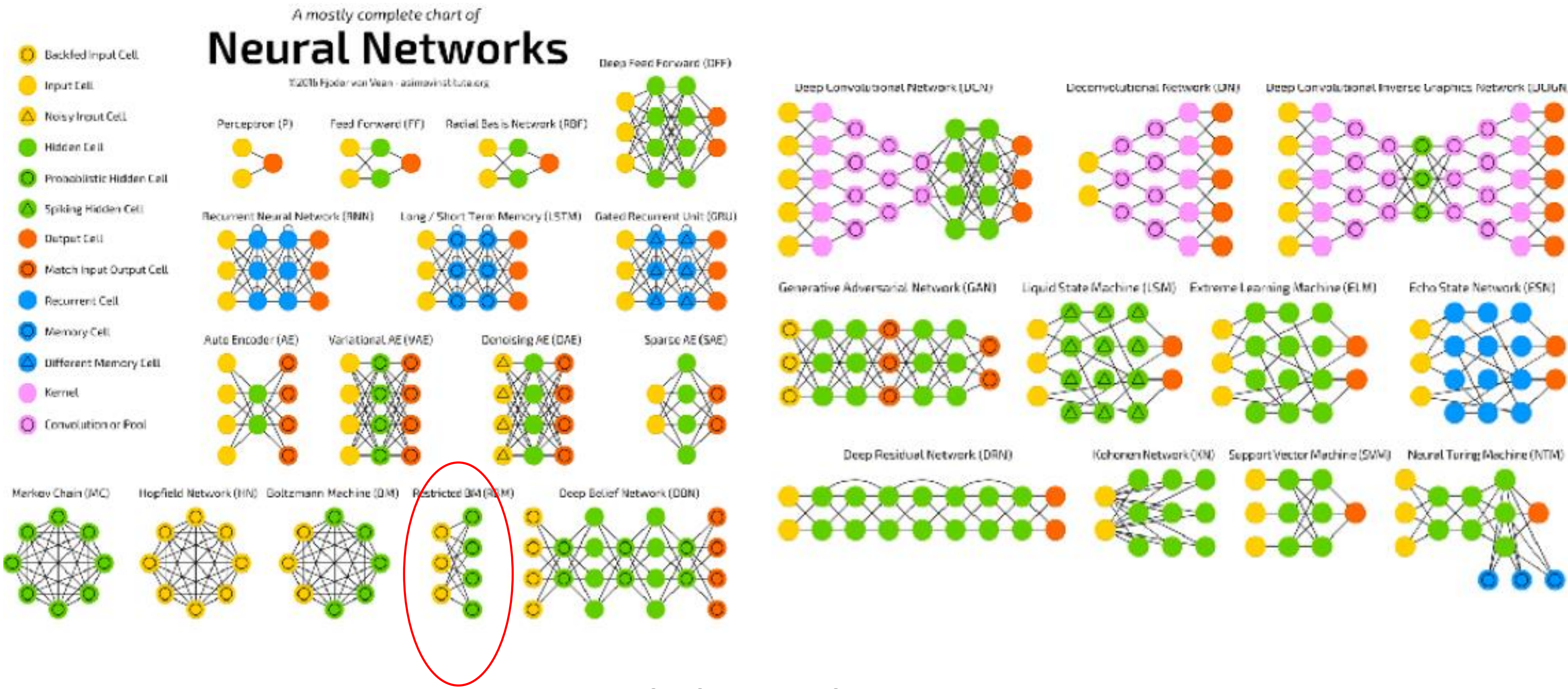


Cat

Dog

[Playground]

Zoo of Neural Network

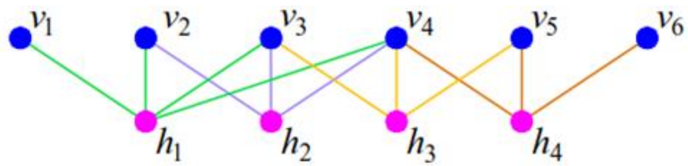


Philosophy:

Connectionism → Intelligence

Restricted Boltzmann Machine (RBM)

Learn



Generate

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

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

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

Universal approximation theorem

Formal statement [edit]

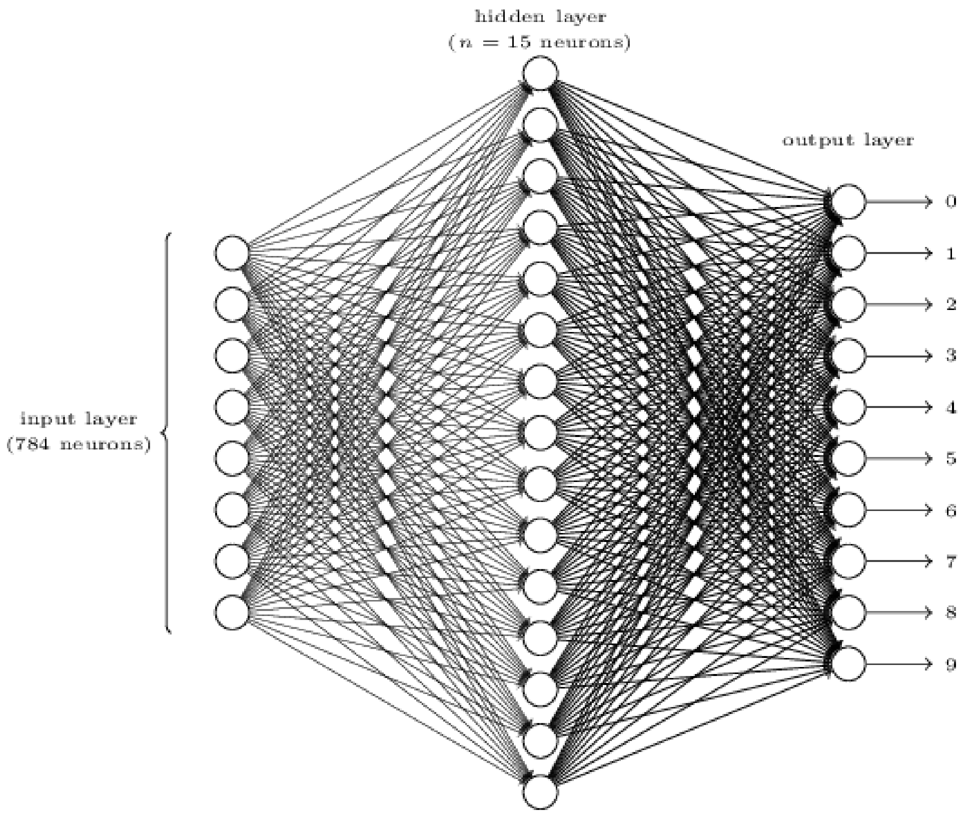
The theorem[2][3][4][5] in mathematical terms:

> 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 $i = 1, \cdots, N$, such that we may define:
>
> $$F(x) = \sum_{i=1}^{N} v_i \varphi \left(w_i^T x + b_i \right)$$
>
> as an approximate realization of the function f where f is independent of φ; 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

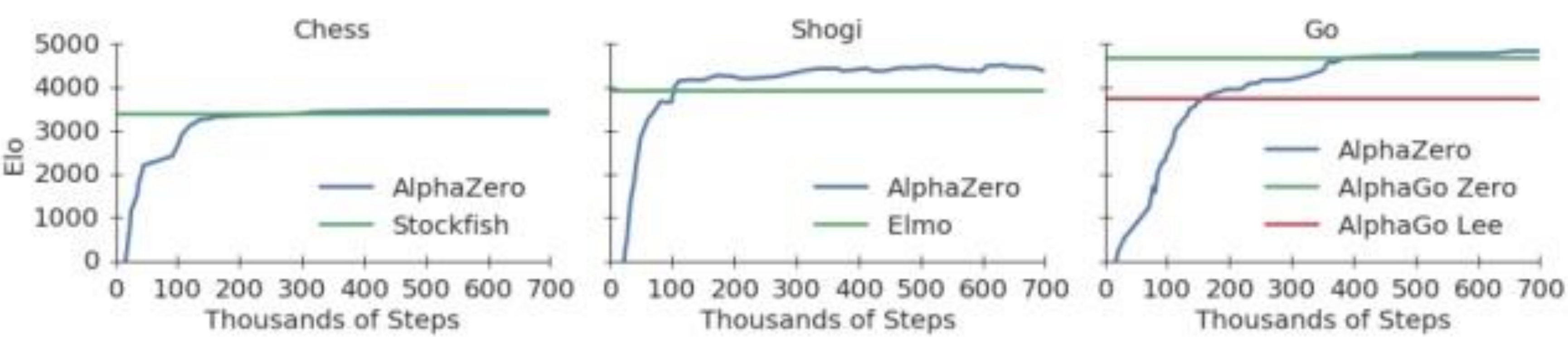
- Simulate any function with enough units



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

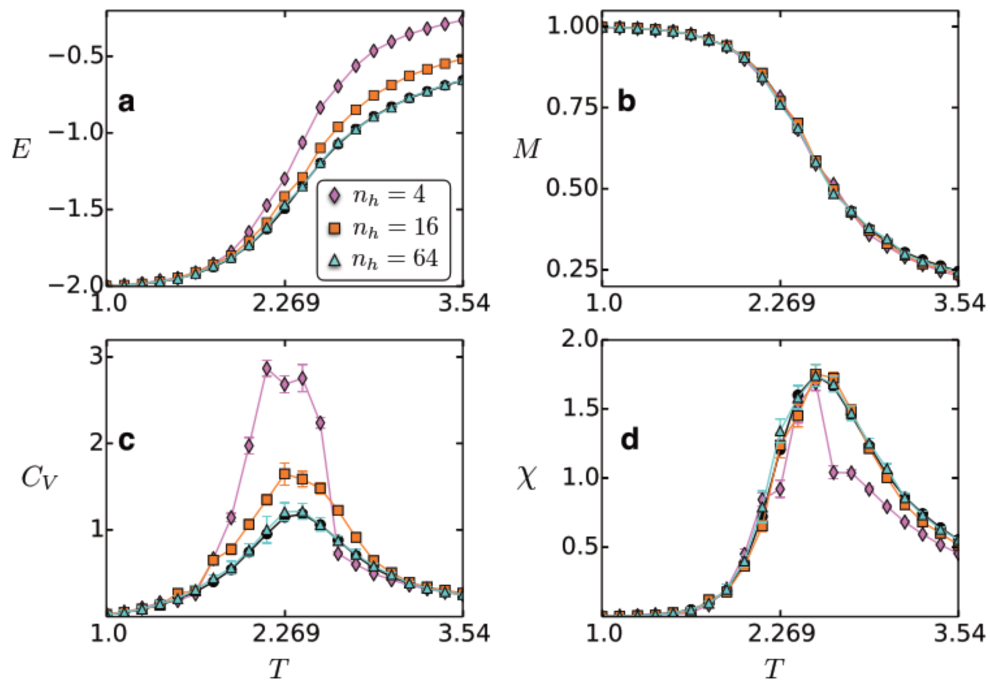
- Tensor Network
- Machine learning
- Connections

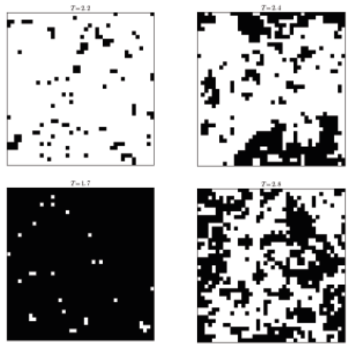
ML applied to physical problems

- 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



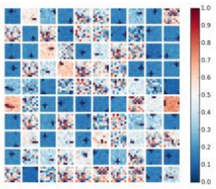


The result is not very good at Tc

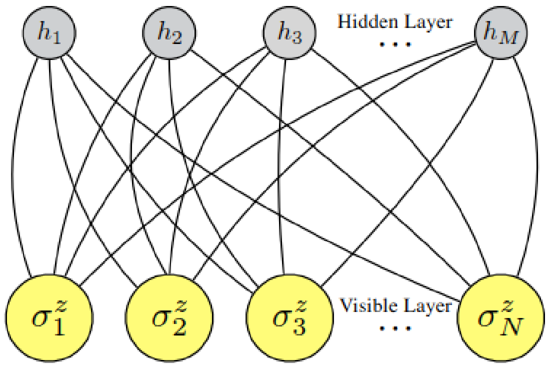
Can RBM represent the distribution well at criticality?

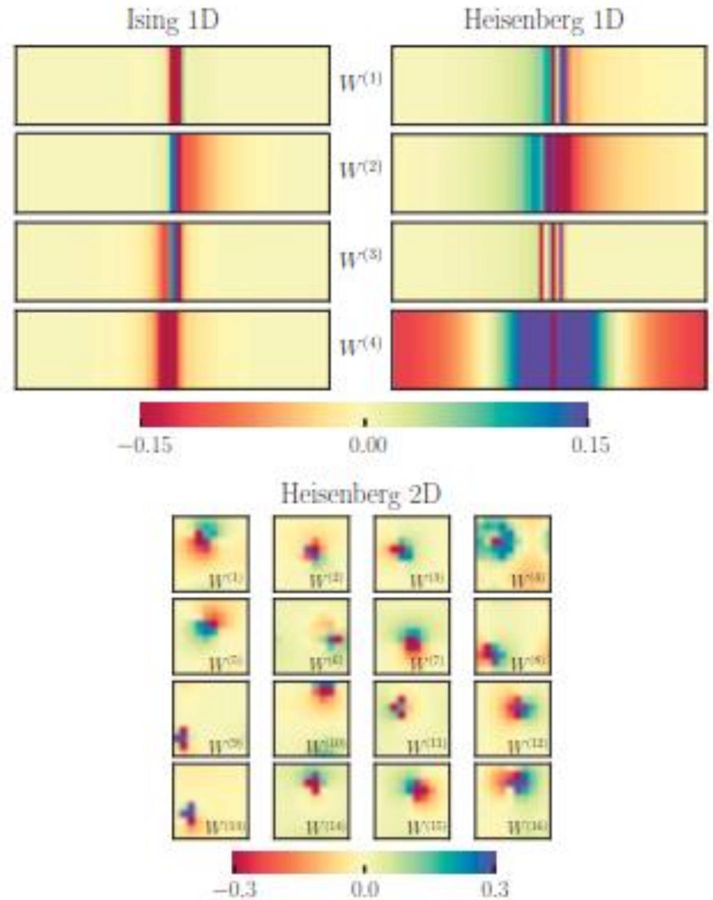
"Learning Thermodynamics with Boltzmann Machines"
G. Torlai, R. G. Melko Phys. Rev. B 94, 165134 (2016)

Accelerated Monte Carlo simulations with restricted Boltzmann machines
L Huang, L Wang Phys. Rev. B 95, 035105



Quantum: RBM as wave function ansatz



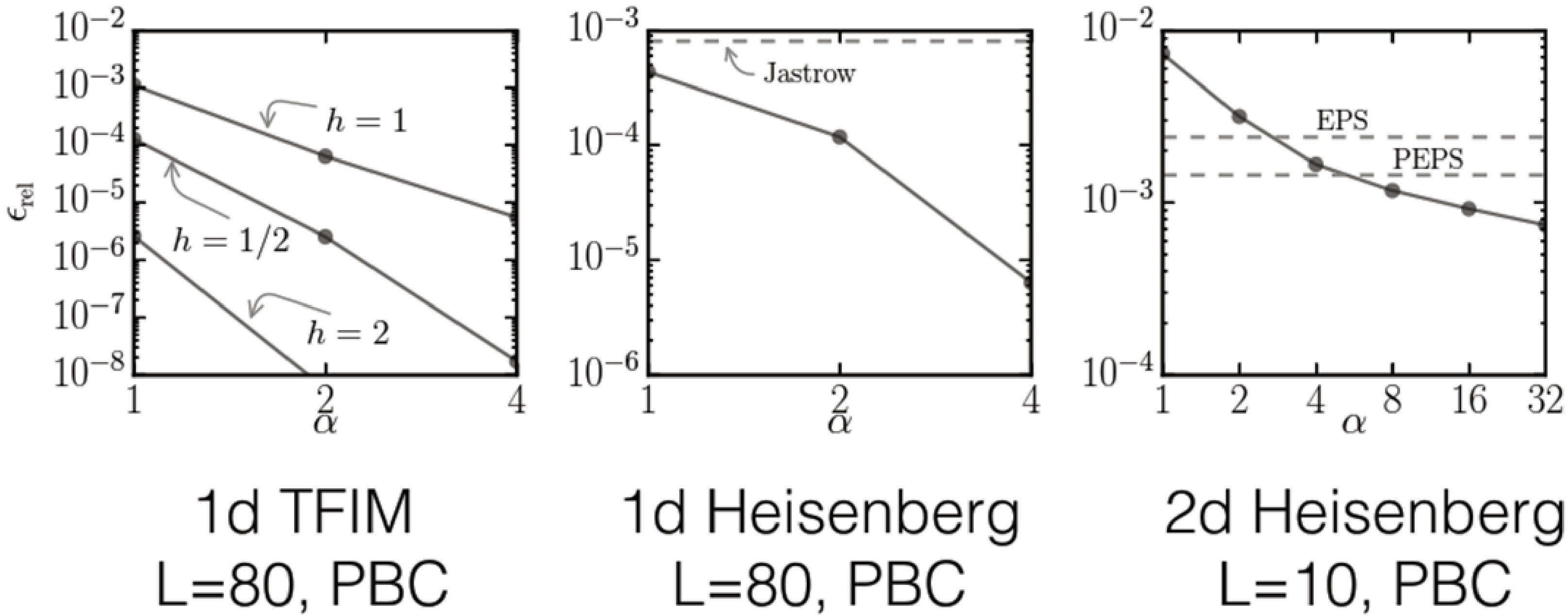


$$\Psi_M(\mathcal{S};\mathcal{W}) = \sum_{\{h_i\}} e^{\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} W_{ij} h_i \sigma_j^z}$$

Complex W,a,b

“Solving the quantum many-body problem with artificial neural networks” by G. Carleo and M. Troyer, Science **355**, 602 (2017).

RBM as wave function ansatz

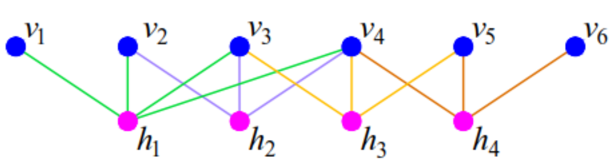


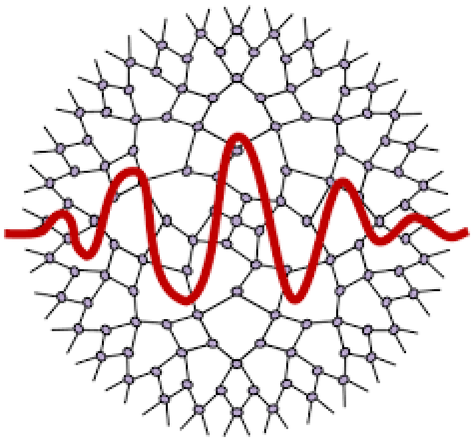
"Solving the quantum many-body problem with artificial neural networks" by G. Carleo and M. Troyer, Science **355**, 602 (2017).

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

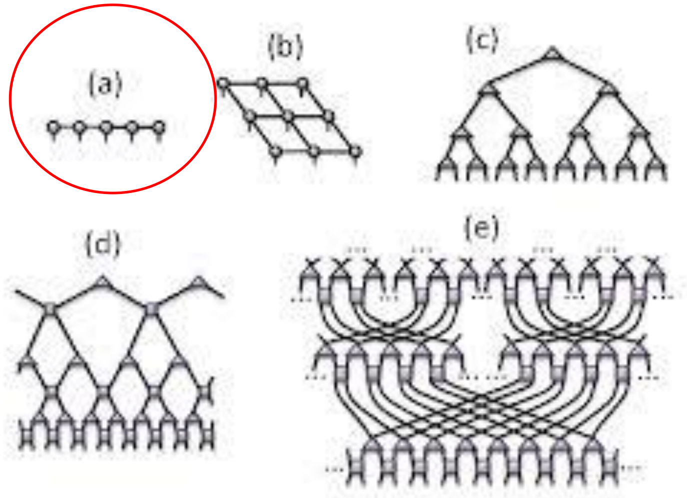
- How is the expressive power of RBM ?
- Does RBM satisfy the area law ?
- Can RBM represent critical possibility distribution ?
- Why is RBM wave function successful ?

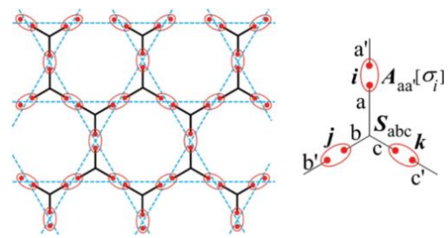


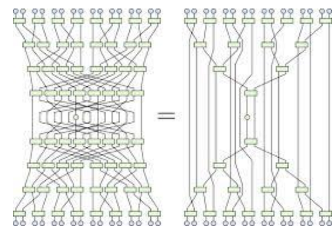


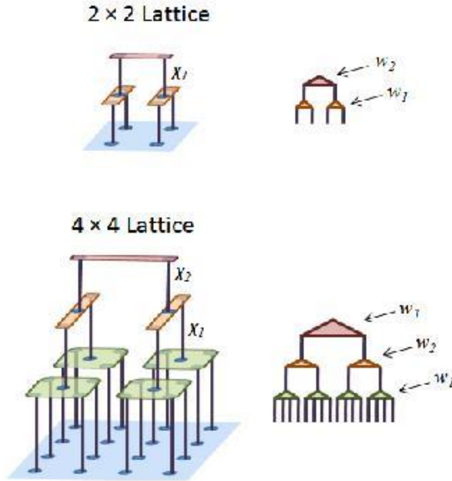
arXiv:1701.04831

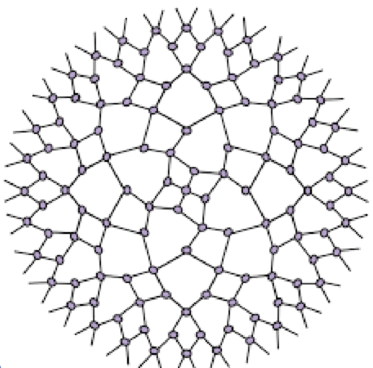
Zoo of Tensor Network State



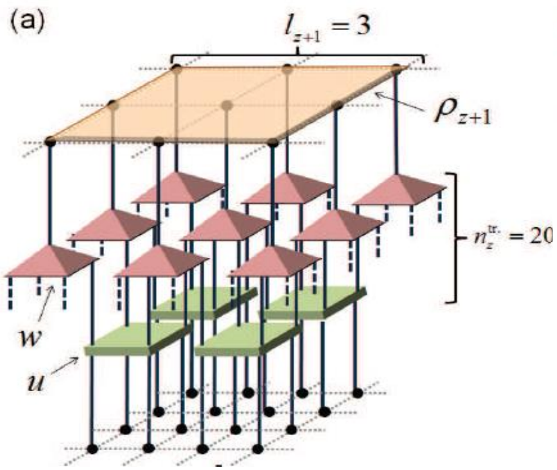


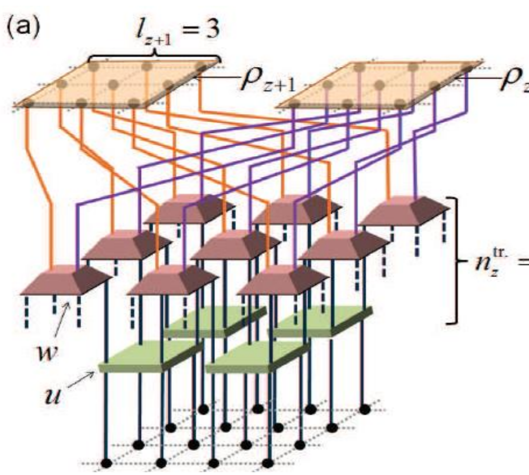




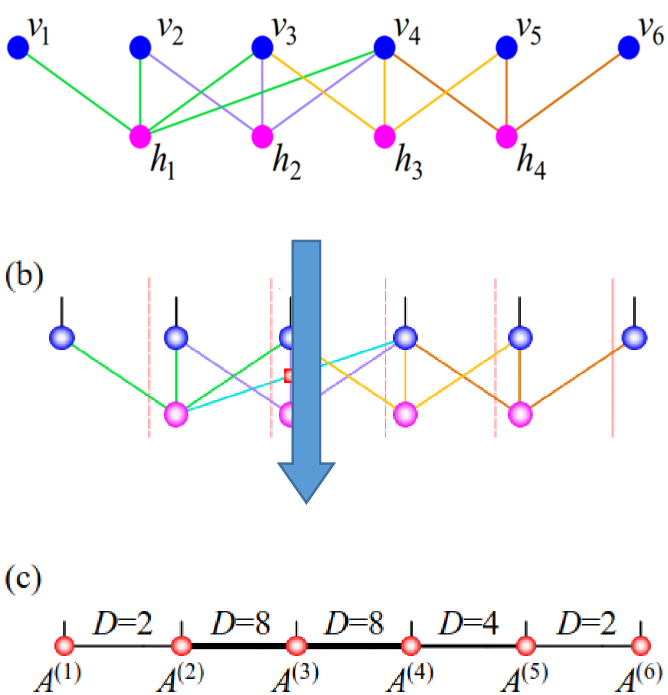


Tensor network family





TNS representation of RBM

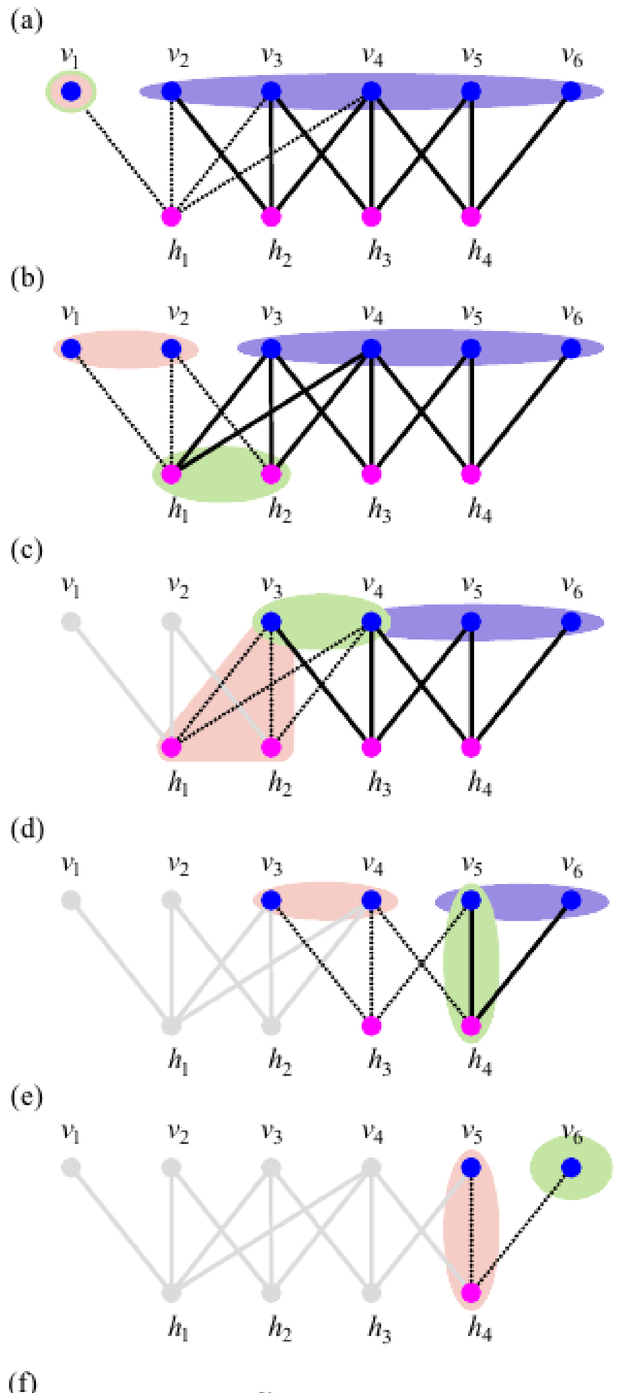


$$M^{(ij)} = \begin{pmatrix} 1 & 1 \\ 1 & e^{W_{ij}} \end{pmatrix},$$
$$\Lambda_v^{(i)} = \text{diag}\left(1, e^{a_i}\right),$$
$$\Lambda_h^{(j)} = \text{diag}\left(1, e^{b_j}\right),$$

$$D = 2^n$$

n is # of cuts

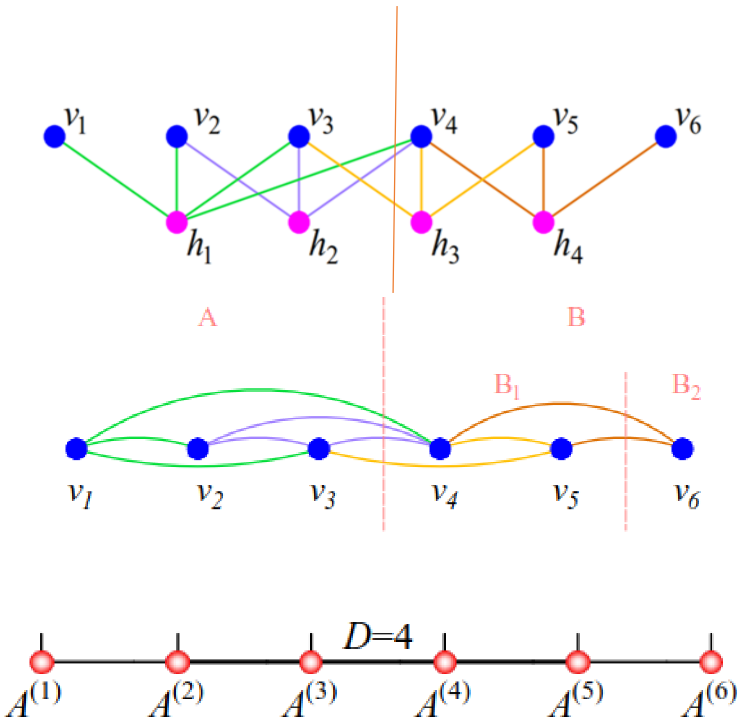
This is the direct way, but not the most efficient way

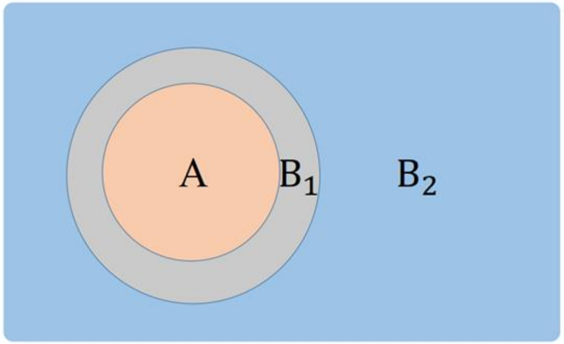


Get the optimal tensor on the fly

Can be generalized to other models.

The entanglement entropy of RBM





Local Connected

$$S_{\max} \sim m^{\frac{1}{d}} L^{d-1},$$

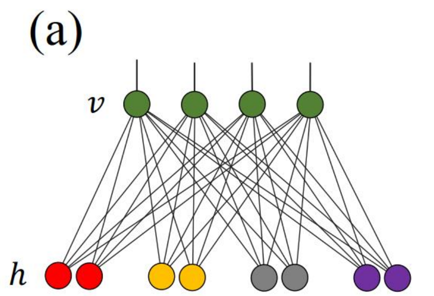
Area Law

The entanglement depends on the size of B_1

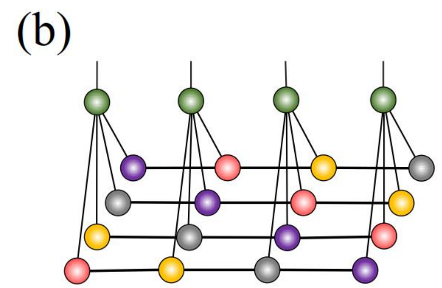
Code: https://github.com/yzcj105/rbm2mps

Good news: Much fewer Variables

Entanglement of shift-invariant RBM

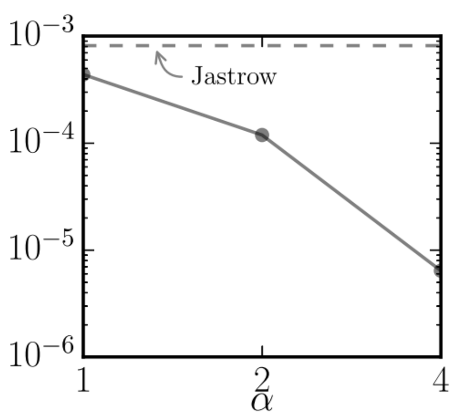


Shifted invariant RBM



MPS

$$\Psi(v) = \prod_{\mathcal{T}} \Psi_{\text{RBM}}(\mathcal{T}\{v_i\}),$$



1D Heisenberg L=80 PBC

Volume law

D=16 MPS

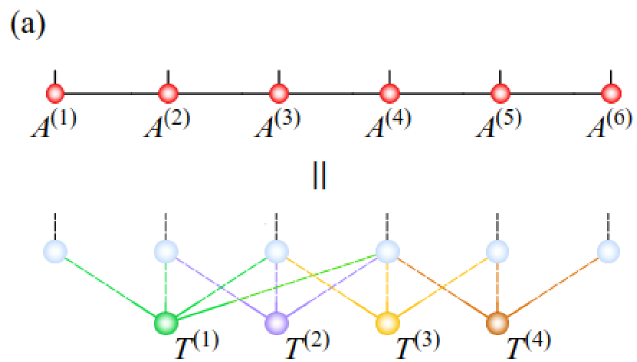
The shift-invariant RBM structure is crucial to the success

3 orders of magnitude fewer variational variables than DMRG

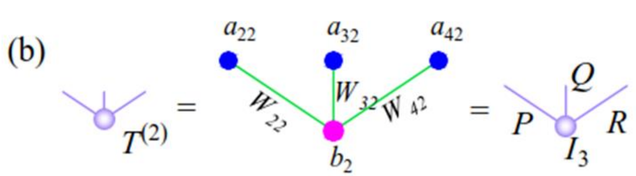
Example:2D System	PEPS	RBM
Long term interactions	Passed by the sites between , increass D	Connected directly
N-body interactions	tensor with D^N elements	N weights
Sampling of the physical freedom	Contraction of a 2D TN	Just a summation in the exponent
Philosophy	Contraction	Product

RBM is a subset of TN theoretically but different practically

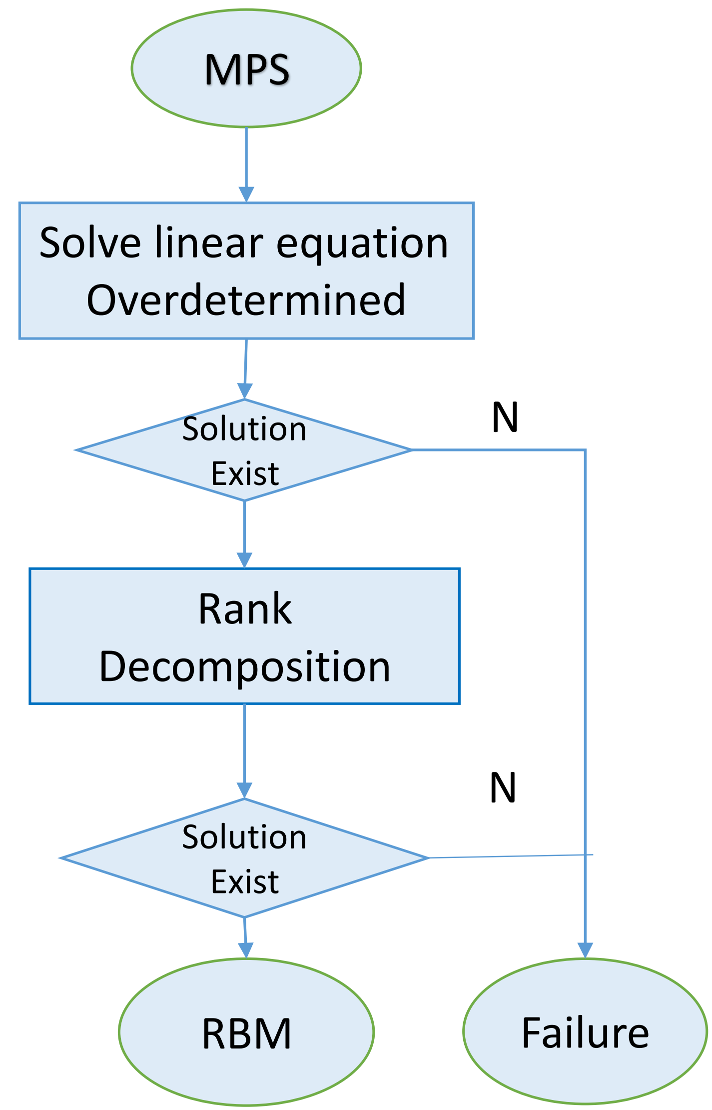
RBM representation of a MPS



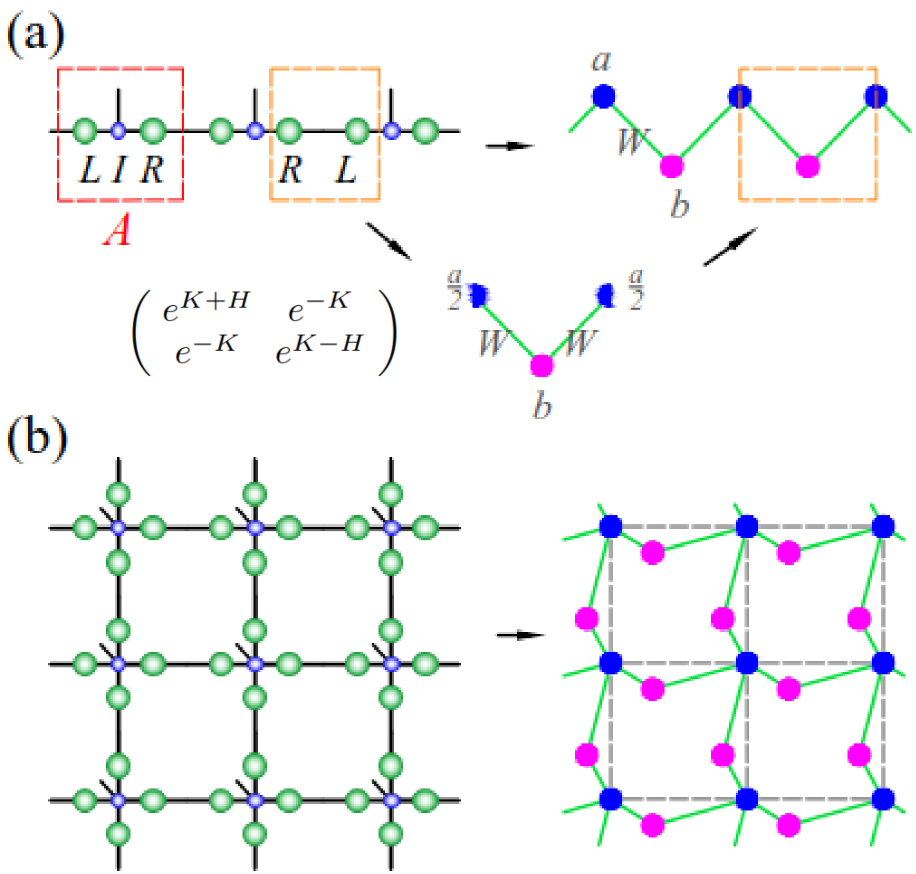
$$\mathrm{Tr}\prod_i A^{(i)}[v_i] = T^{(1)}_{v_1 v_2 v_3 v_4} T^{(2)}_{v_2 v_3 v_4} T^{(3)}_{v_3 v_4 v_5} T^{(4)}_{v_4 v_5 v_6}.$$



$$T^{(2)}_{v_2 v_3 v_4} = \sum_{h_2 \in \{0,1\}} e^{h_2 b_2 + \sum_{i \in \{2,3,4\}} v_i (W_{i2} h_2 + a_{i2})},$$



Explicit RBM of Ising Model



$$Z = \sum_{\{s_i\}} \exp\left(K \sum_{\langle i,j \rangle} s_i s_j + H \sum_i s_i\right)$$

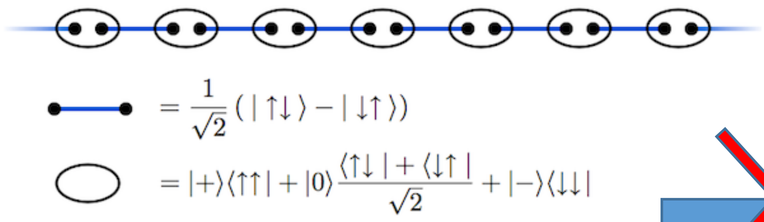
$$W = \ln(4e^{4K} - 2)$$

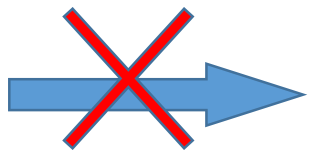
$$a = -8K - 2H - 4\ln 2$$

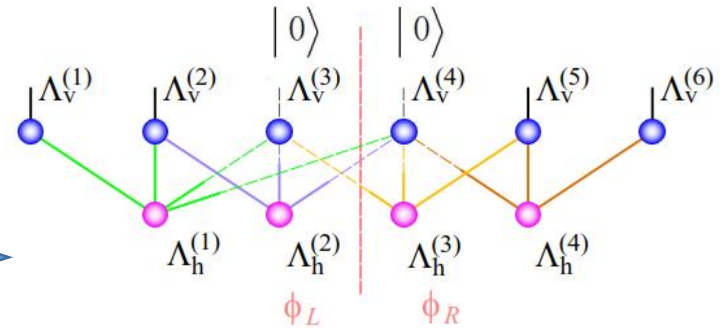
$$b = -\ln(e^{4K} - 1) - 2\ln 2$$

The RBM can represent Ising model at criticaliy!

AKLT







+ 0 -- + 0 – 0 0 + 0 -- 0

Hidden order

+ -- + – + --

$$\Psi(\cdots 00 \cdots) = \phi_L \phi_R$$

… + 0 0 0 0 -- …

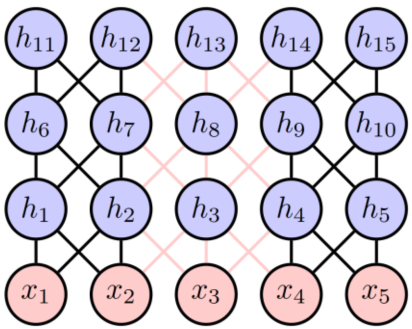
couples

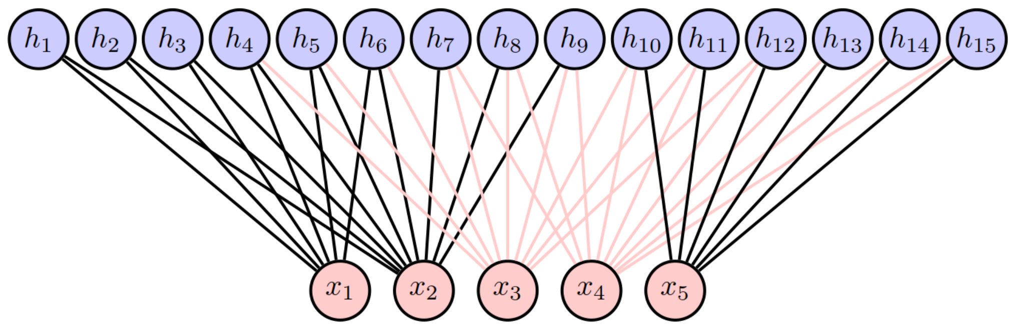
… -- 0 0 0 0 + …

No matter how long the sequences of 0 is, it is always the superposition

Once the units of interface region are fixed, the wave function decouples

Deep or shallow, is a question.





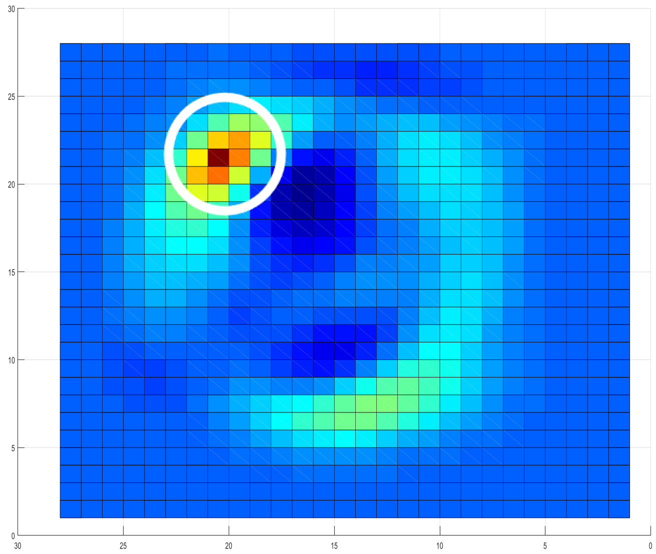
D=16

D=4

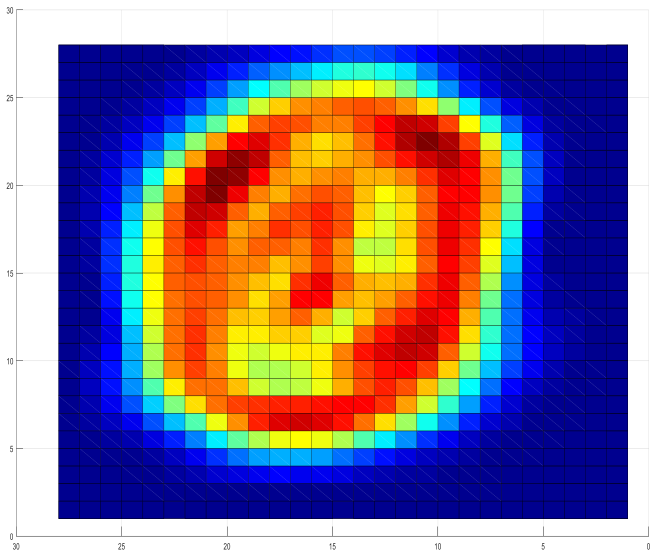
Same number of units and connections.

Deep BM allows more entanglement.

Entanglement and correlation of MNIST datasets



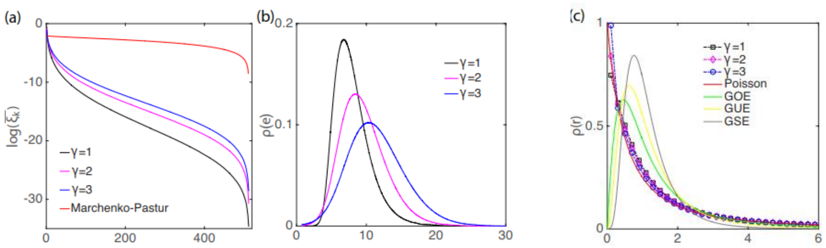
$$\langle S_0 S_{\vec{r}} \rangle$$



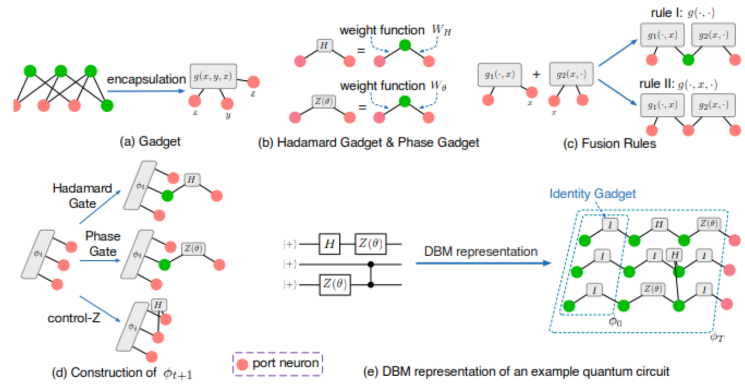
$$\sum_{\vec{p}} \langle S_{\vec{p}} S_{\vec{r}} \rangle$$

For images, the correlation is local and anisotropic

Work related

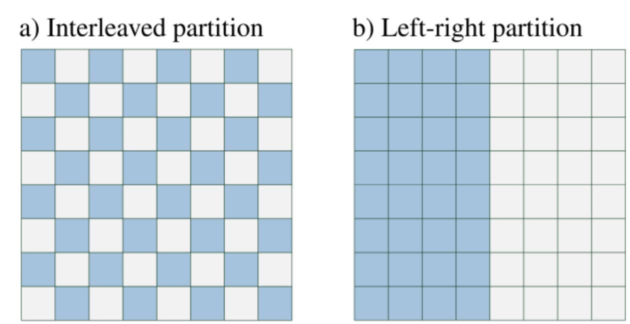


Quantum Entanglement in Neural Network States by D.-L. Deng, X. Li, and S. D. Sarma, arXiv:1701.04844



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

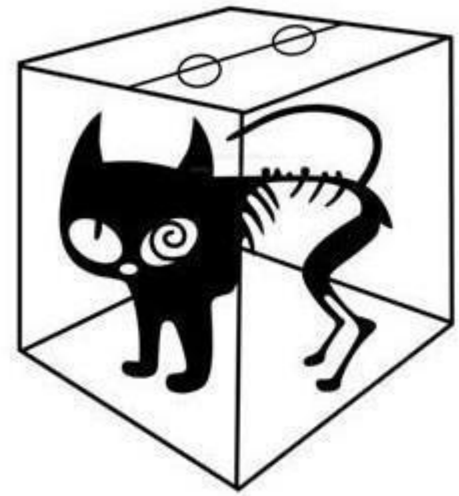


Deep Learning and Quantum Physics : A Fundamental Bridge
by Y. Levine, D. Yakira, etc. arxiv:1704.01552

Following work on RBM wave function

- G. Carleo and M. Troyer, Science 355, 602 (2017).
- R Kaubruegger et.al. arXiv:1710.04713
- Stephen R. Clark arXiv: 1710.03545.
- Ivan Glasser et.al. arXiv: 1710.04045.
- Yusuke Nomura et.al. Phys. Rev. B 96, 205152 (2017).

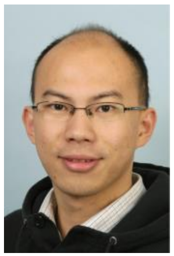
Summary



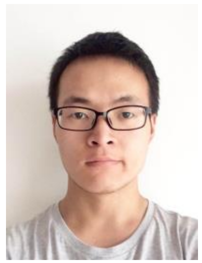
Machine Learning

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

- J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, arXiv:1701.04831
- D.-L. Deng, X. Li, and S. D. Sarma, Phys. Rev. X 7, 021021 (2017)
- X. Gao and L.-M. Duan, Nature Communications 8, 662 (2017).
- Y. Huang and J. E. Moore, arXiv:1701.06246