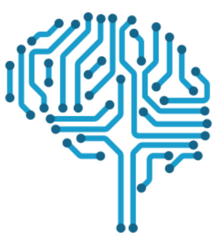
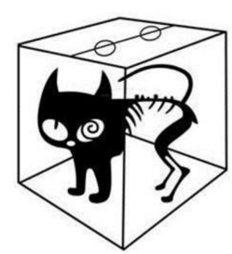
On the Equivalence of Restricted Boltzmann Machines and Tensor Network States

Jing Chen （陈靖）

IOP, CAS

chenjing@iphy.ac.cn





arXiv:1701.04831

Machine learning

AlphaGo

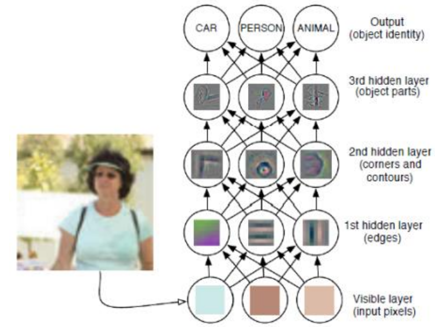


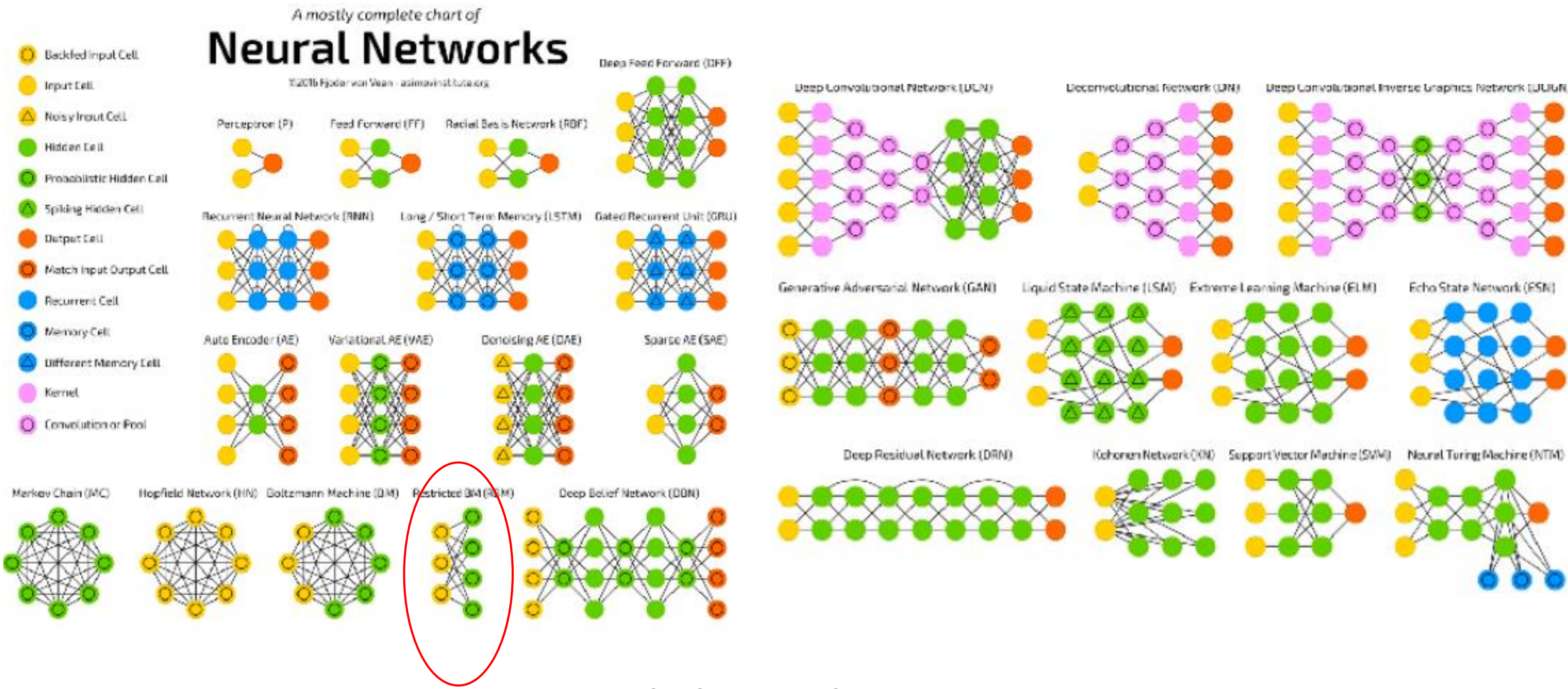
Image Recognition

Amazon recommender

Driverless Car

Finance

Zoo of Neural Network

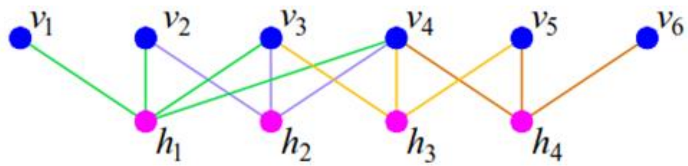


Philosophy:

Connectionism → Intelligence

Restricted Boltzmann Machine (RBM)

Learn



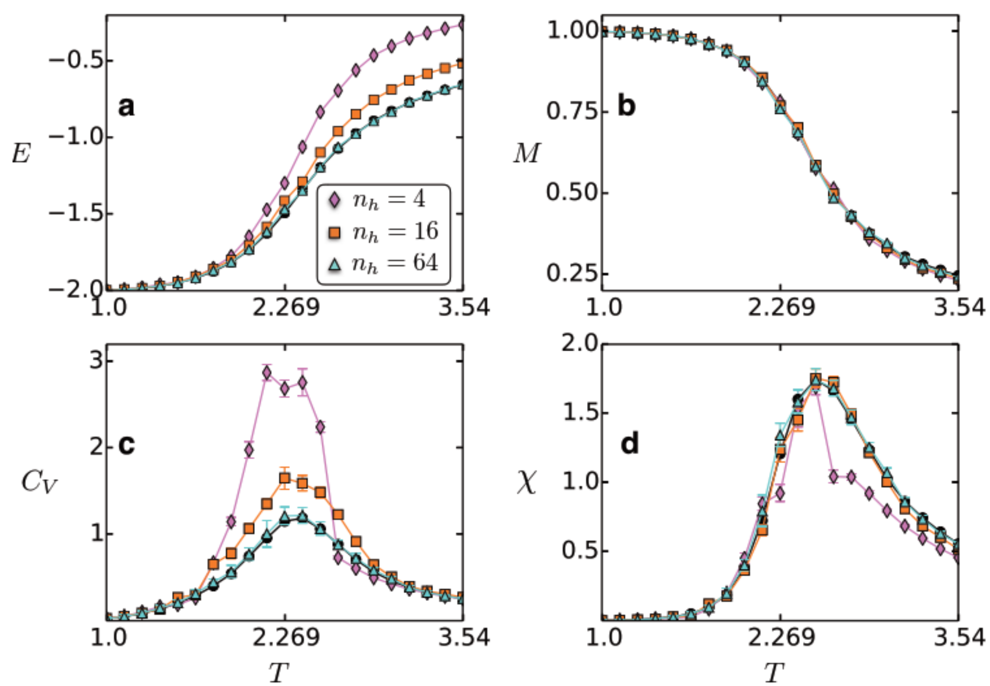
Generate

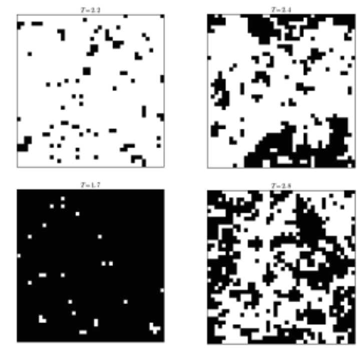
$$(v,h) = -\sum_i a_i v_i - \sum b_j h_j - \sum_{i,j} v_i W_{ij} h_j$$

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

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

Learning Classical Statistic Distribution by RBM



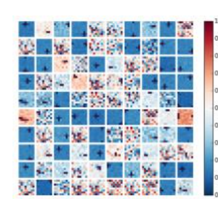


The result is not very good at Tc

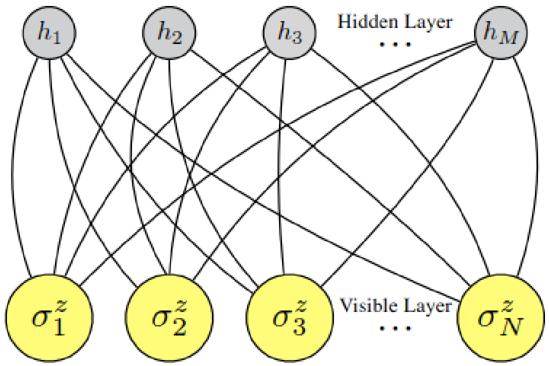
Can RBM represent the distribution well at criticality?

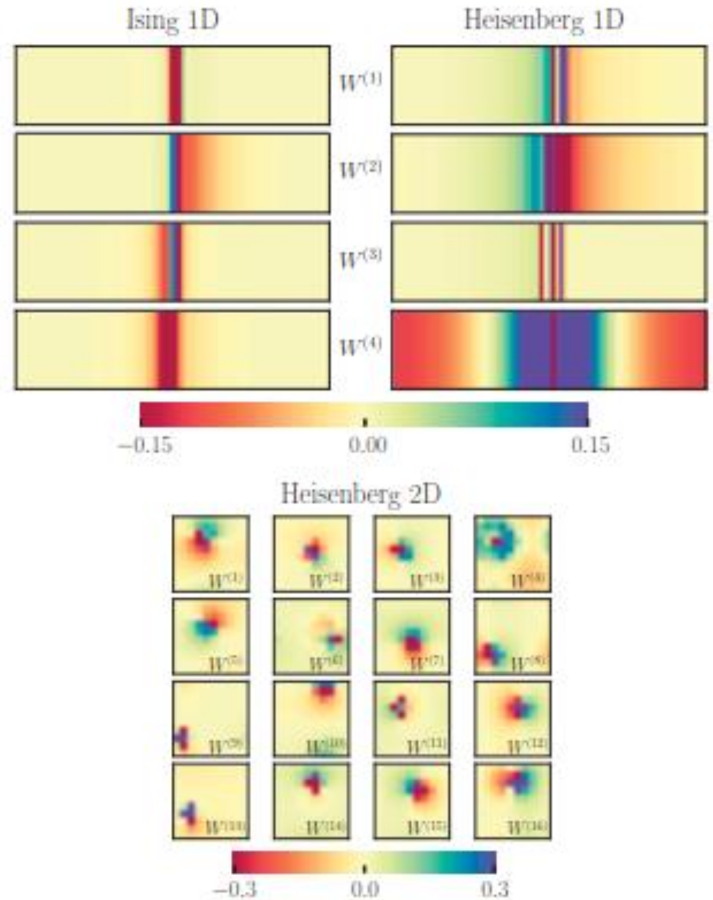
"Learning Thermodynamics with Boltzmann Machines"
G. Torlai, R. G. Melko Phys. Rev. B 94, 165134 (2016)

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



Quantum: RBM as wave function ansatz



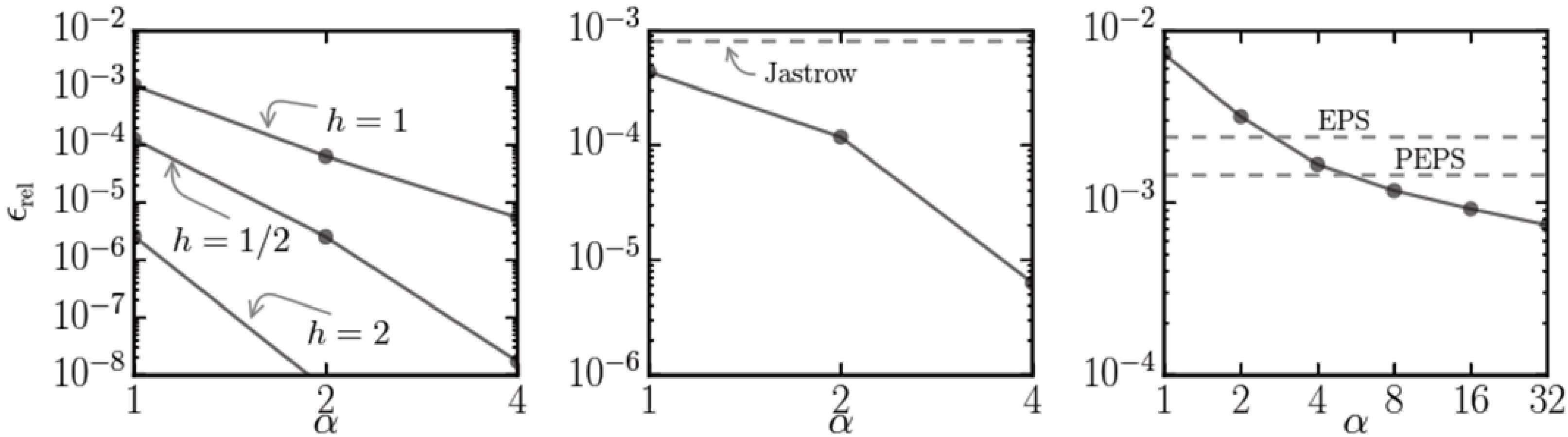


$$\Psi_M(\mathcal{S};\mathcal{W}) = \sum_{\{h_i\}} e^{\sum_j a_j \sigma_j^z + \sum_i b_i h_i + \sum_{ij} W_{ij} h_i \sigma_j^z}$$

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



1d TFIM
L=80, PBC

1d Heisenberg
L=80, PBC

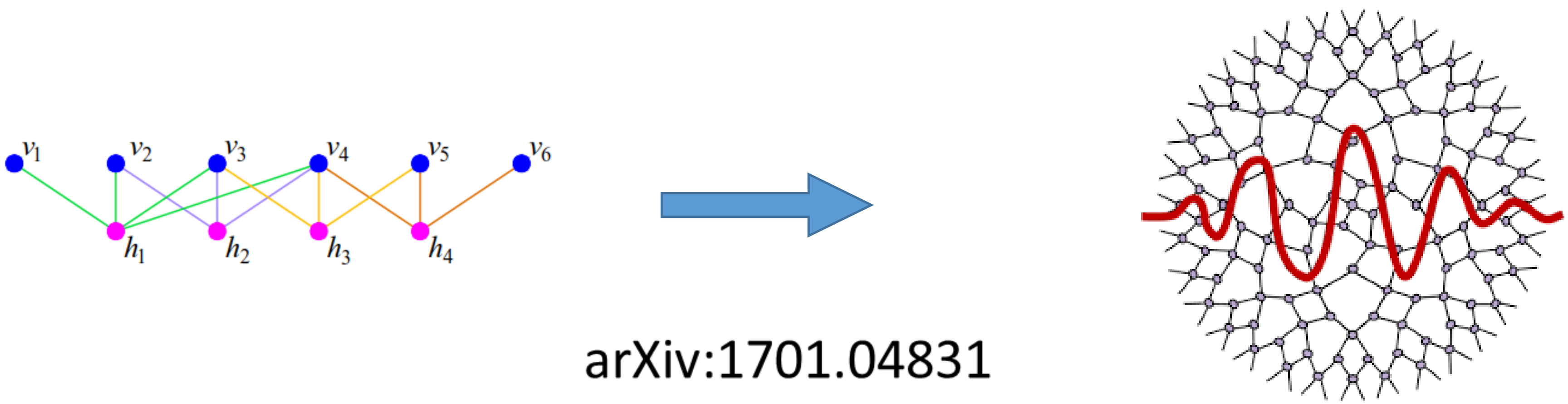
2d Heisenberg
L=10, PBC

“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

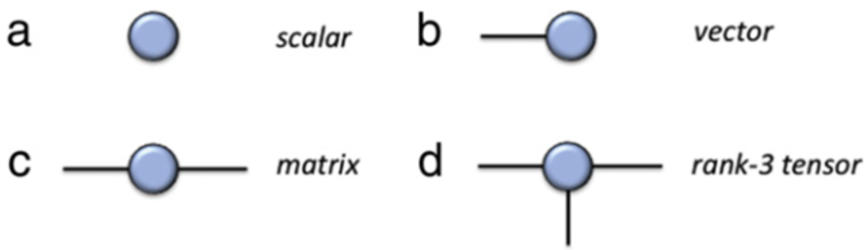
Many-body quantum state tomography with neural networks, Giacomo Torlai, Guglielmo Mazzola, Juan Carrasquilla, Matthias Troyer, Roger Melko, Giuseppe Carleo, arxiv:1703.05334

- 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

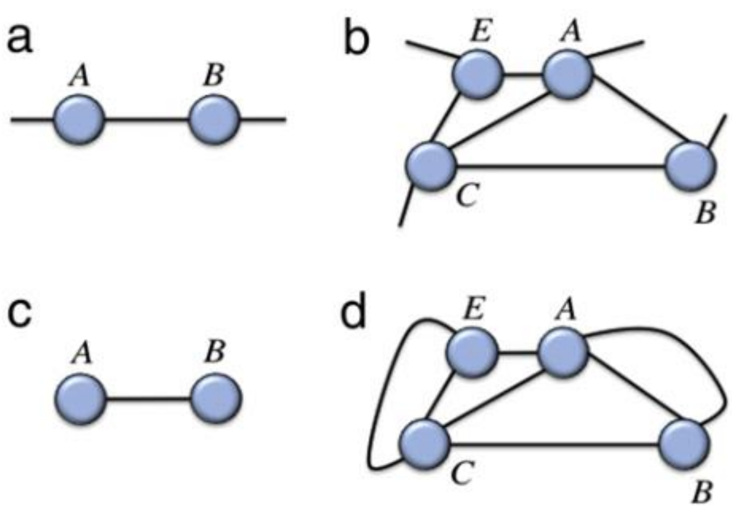
Tensor Network

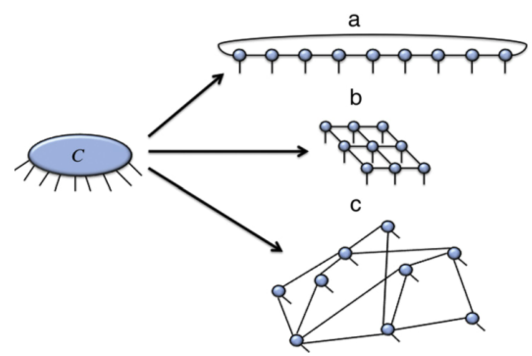


matrix
product

Extend

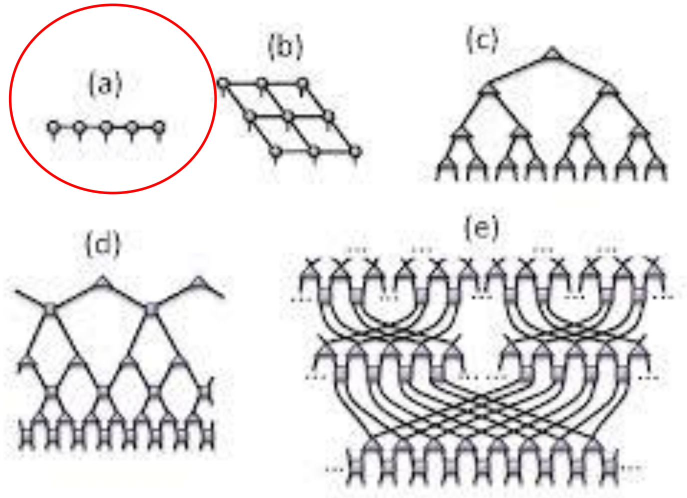
tensor
contraction

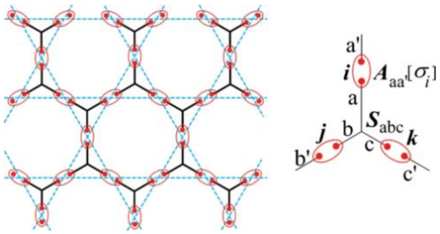


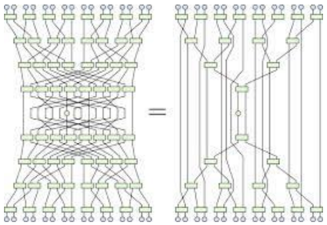


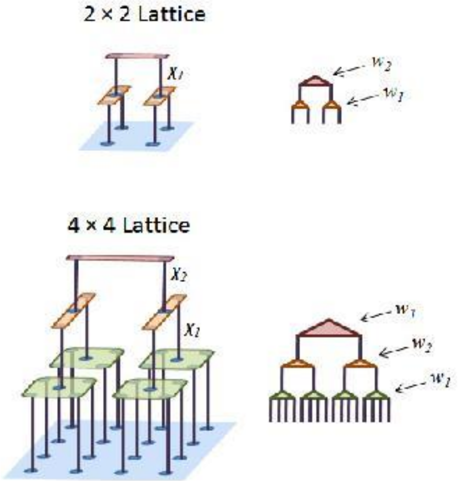
Represent wave functions

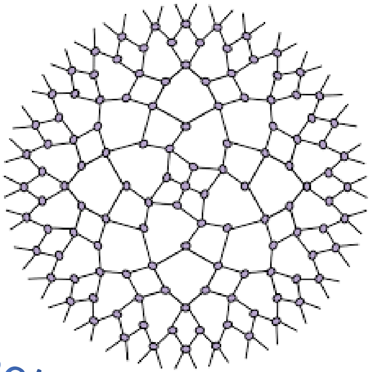
Zoo of Tensor Network State



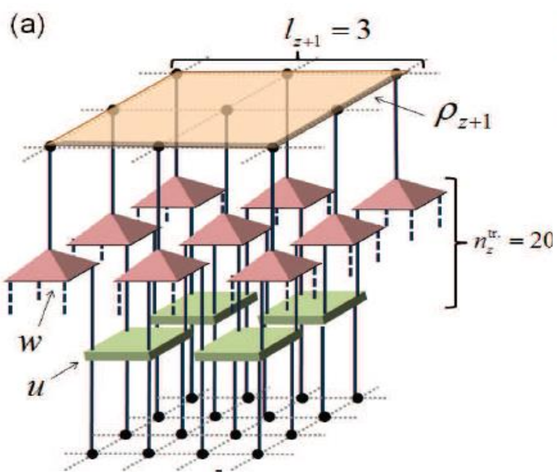


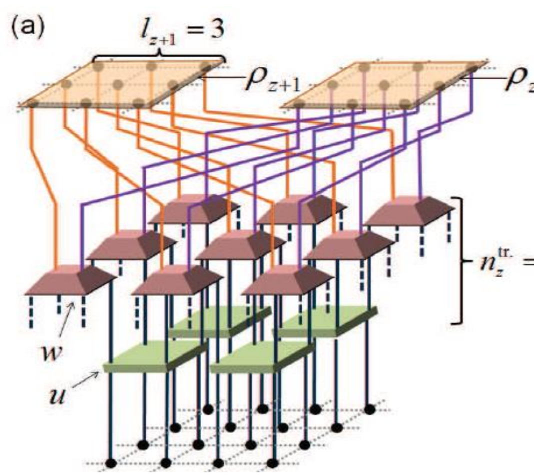






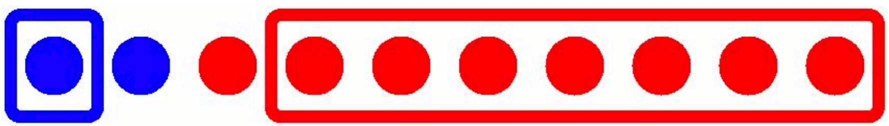
Tensor network family

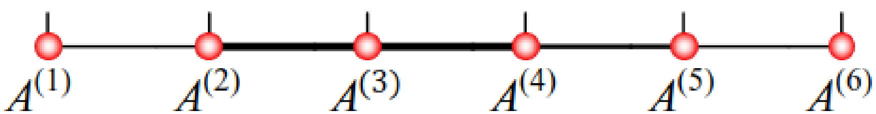




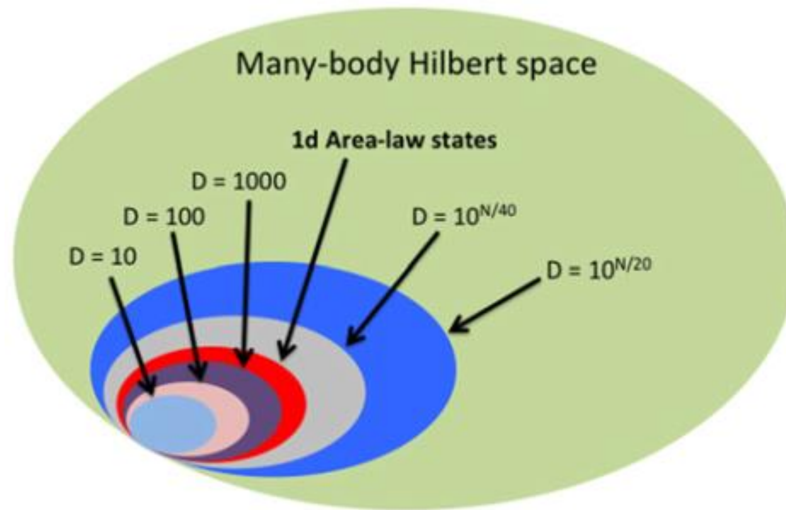
Matrix Product State (MPS)

DMRG wave function ansatz





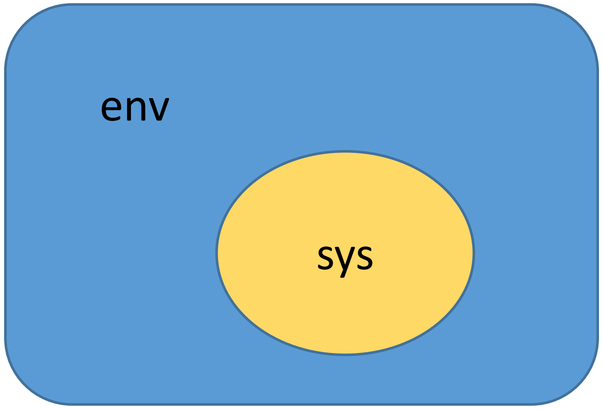
$$\Psi_{\mathrm{MPS}}(v) = \mathrm{Tr} \prod_i A^{(i)}[v_i],$$

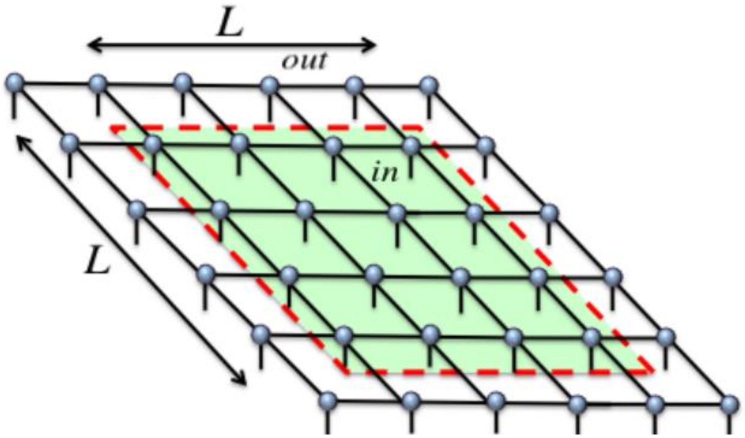


low rank
approximation

Very successful in 1D

Entanglement (Area Law)





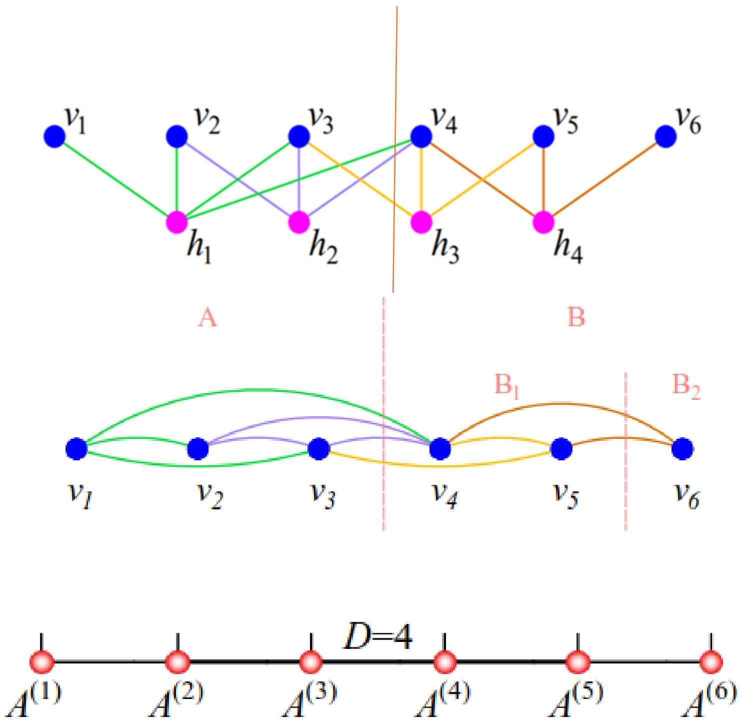
- gapped systems ground state

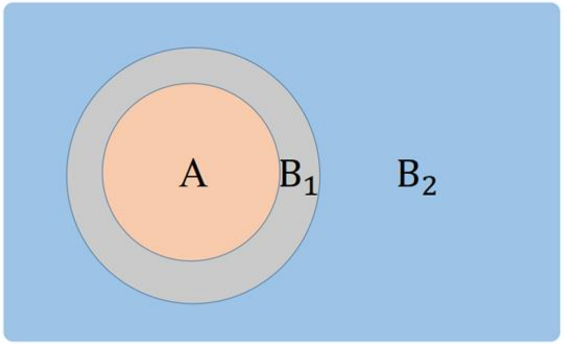
$$S = -Tr_e\left(\rho_{es} \log \rho_{es}\right)$$

$$S \leq n\, ln\, D$$

J Eisert, M Cramer, MB Plenio Reviews of Modern Physics, 2010

The entanglement entropy of RBM





Local Connected

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

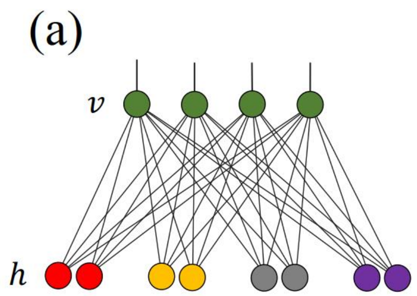
The entanglement depends on the size of B_1

Area Law

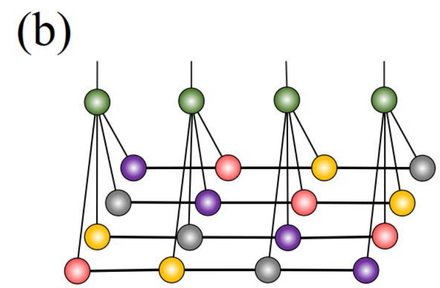
Code: https://github.com/yzcj105/rbm2mps

Good news: Much fewer Variables

Entanglement of shift-invariant RBM



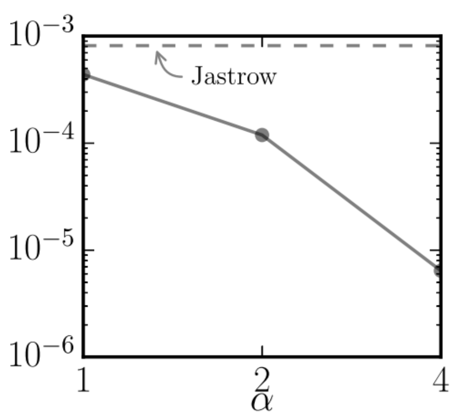
Shifted invariant RBM



MPS

Volume law

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



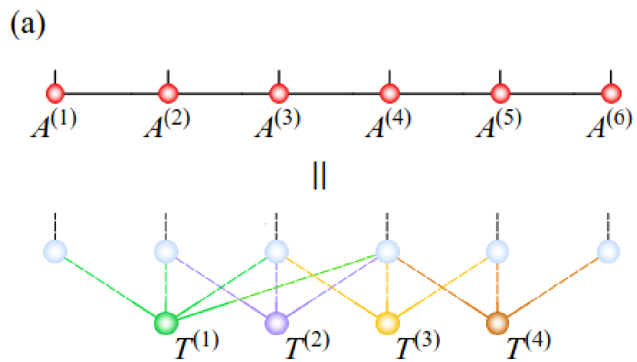
1D Heisenberg L=80 PBC

D=16 MPS

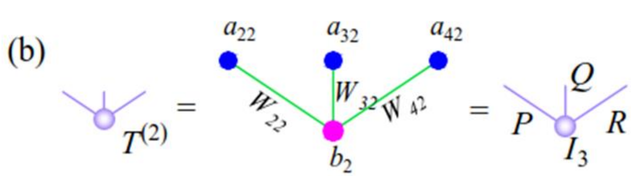
The shift-invariant RBM structure is crucial to the success

3 orders of magnitude fewer variational variables than DMRG

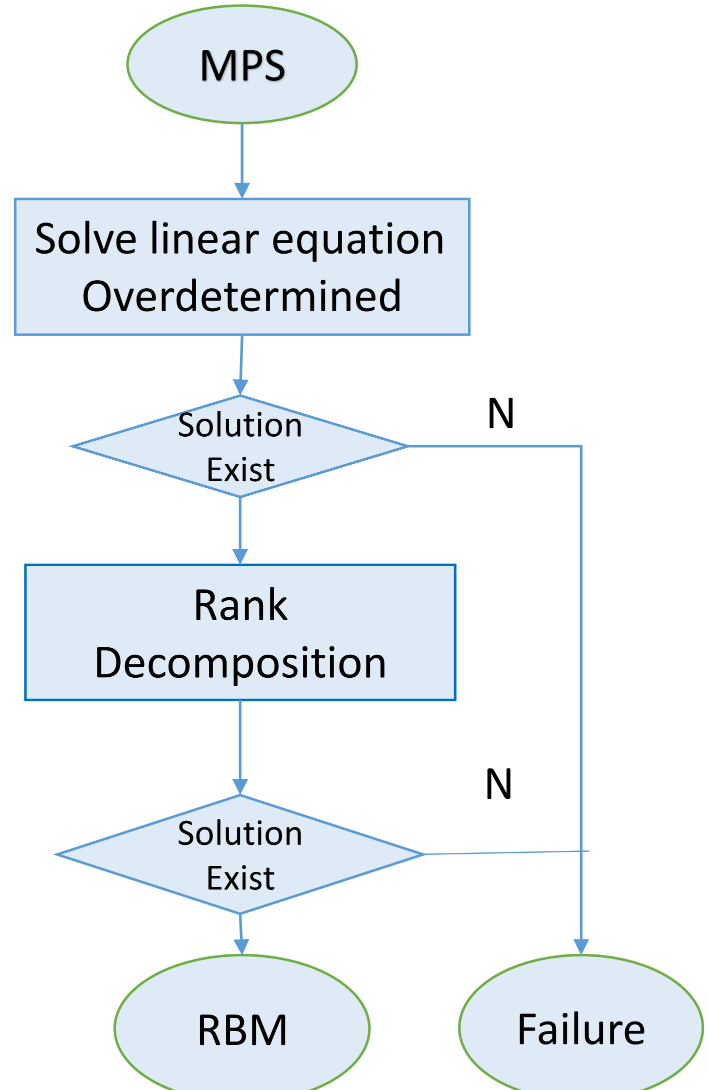
RBM representation of a MPS



$$\mathrm{Tr}\prod_i A^{(i)}[v_i] = T^{(1)}_{v_1v_2v_3v_4}T^{(2)}_{v_2v_3v_4}T^{(3)}_{v_3v_4v_5}T^{(4)}_{v_4v_5v_6}.$$



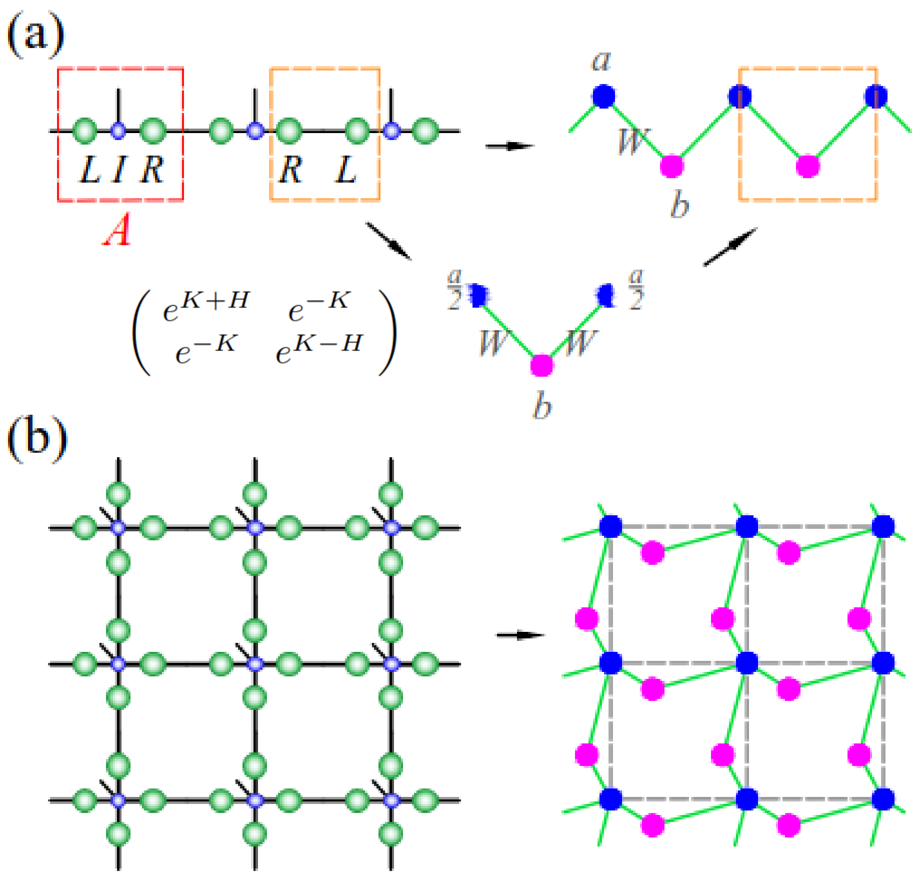
$$T^{(2)}_{v_2v_3v_4} = \sum_{h_2\in\{0,1\}} e^{h_2b_2+\sum_{i\in\{2,3,4\}} v_i(W_{i2}h_2+a_{i2})},$$



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

Local RBM is a subset of TN theoretically but different practically

Explicit RBM of Ising Model



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

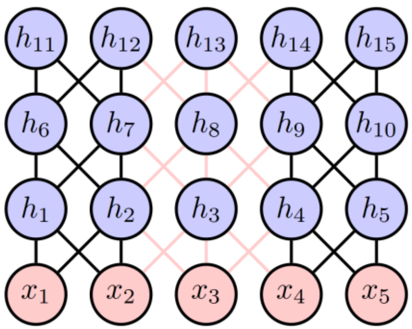
$$W=\ln(4e^{4K}-2)$$

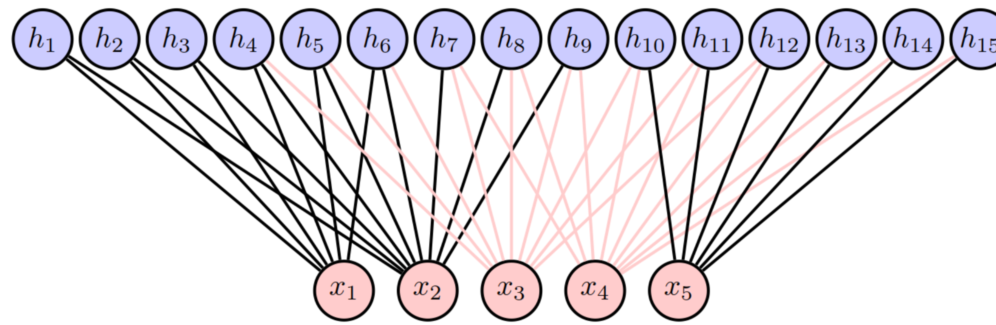
$$a=-8K-2H-4\ln 2$$

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

The RBM can represent Ising model at criticaliy!

Deep or shallow, is a question.





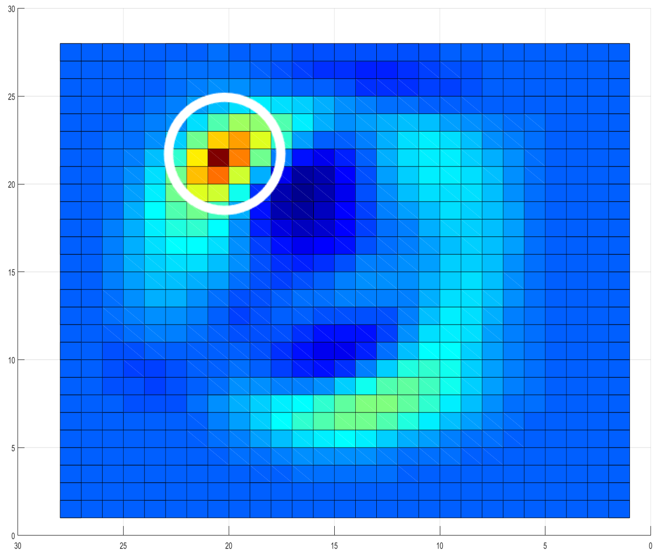
D=16

D=4

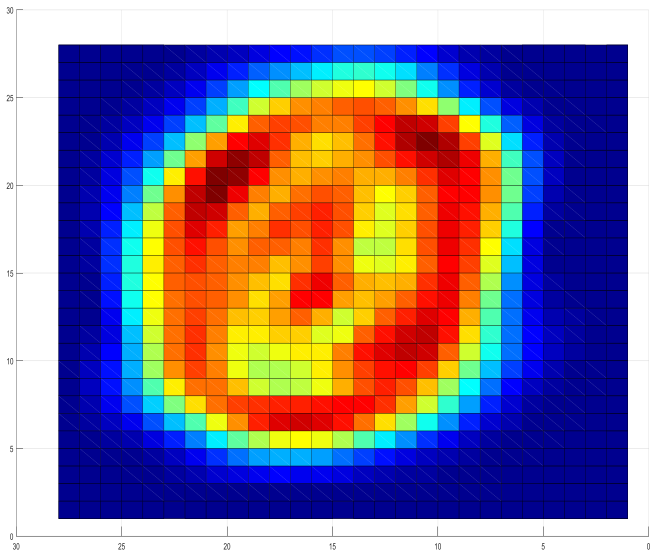
Same number of units and connections.

Deep BM allows more entanglement.

Entanglement and correlation of MNIST datasets



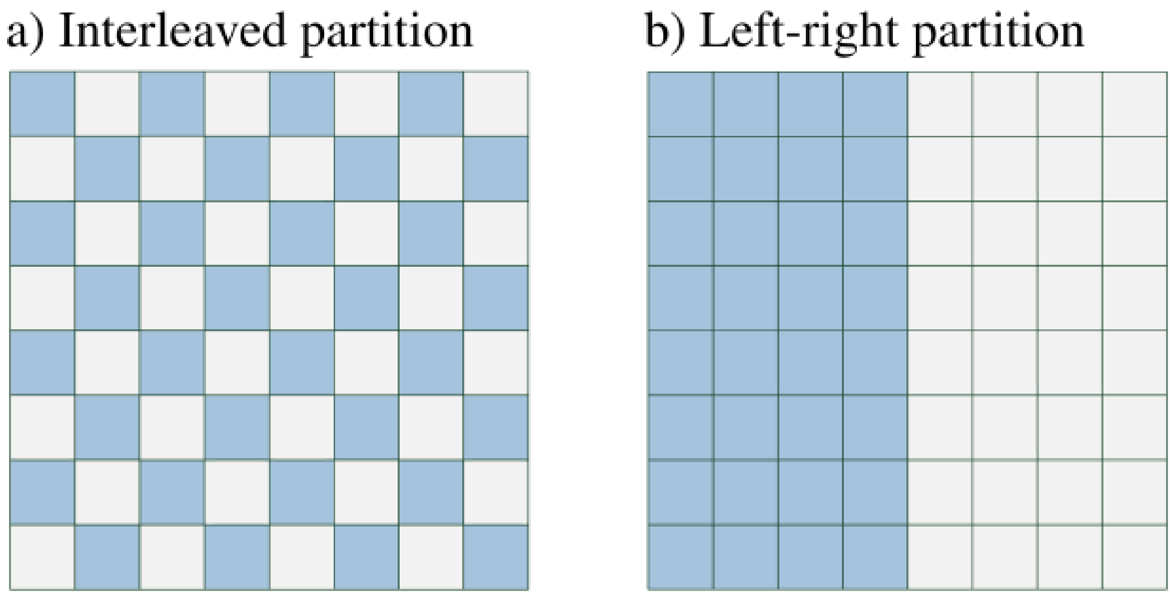
$$\langle S_0 S_{\vec{r}} \rangle$$



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

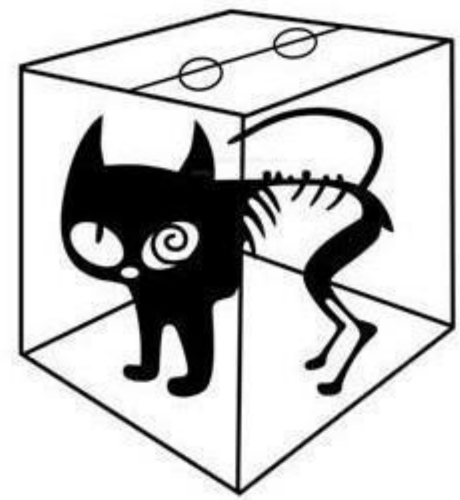
For images, the correlation is local and anisotropic

Introduced to computer science



Deep Learning and Quantum Entanglement: Fundamental Connections with Implications to Network Design
by Y. Levine, D. Yakira, etc. arxiv:1704.01552

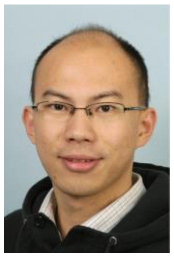
Summary



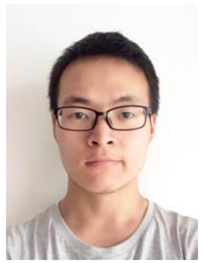
Machine Learning

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, arXiv:1701.05039
- Y. Huang and J. E. Moore, arXiv:1701.06246
- G. Carleo and M. Troyer, Science 355, 602 (2017).