Efficient Representation of Quantum Many-body States with Deep Neural Networks

Xun Gao 郜勋

with Prof. Lu-ming Duan

Outline

Neural Network Representation for Many_body State

Deep vs. Shallow Architecture

Limitation of Restricted Bolzmann Machine

Power of Deep Bolzmann Machine

How to train Deep Boltzmann Machine

1

Difficulty: Curse of Dimensionality

Many-Body Problem

1 Particle: $c_0|0\rangle + c_1|1\rangle \in \mathcal{H}$ Dimension is 2

2 Particles: $c_{00}|00\rangle + c_{01}|01\rangle + c_{10}|10\rangle + c_{11}|11\rangle$

$\in \mathcal{H} \otimes \mathcal{H}$ Dimension is 4

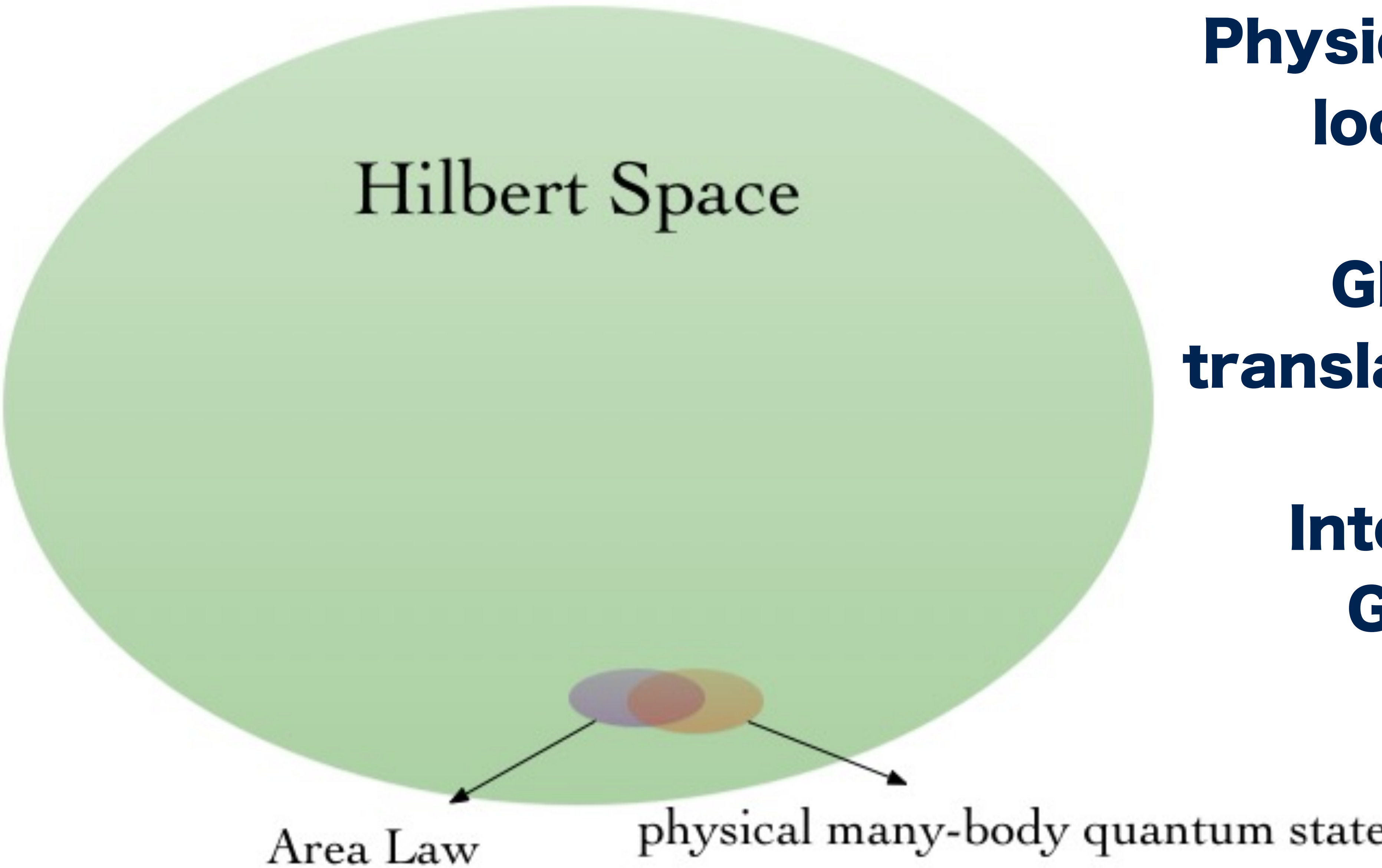
$|00\rangle + |11\rangle$ **Entanglement**

n Particles: $c_{0\ldots0}|0\cdots0\rangle + \cdots + c_{1\ldots1}|1\cdots1\rangle$

$\in \mathcal{H}^{\otimes n}$ Dimension is 2^n

Hilbert space is too large !! How to describe it??

The Physical Relevant Corner



Physical Constraints:
locality (area law)

Global Symmetry:
translational-invariant

Internal Symmetry:
Gauge Symmetry

low description complexity?

Hamiltonian

$$H = \sum_i^n H_i \qquad H_i = I_1 \otimes \cdots \otimes H_i \otimes \cdots \otimes I_n$$

dimension is bounded by constant

low temperature property — **ground state** $|\psi\rangle$ $\qquad \min_{|\psi\rangle} \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}$

thermal equilibrium property — **thermal state** $\qquad \frac{e^{-\beta H}}{\mathrm{tr} e^{-\beta H}}$

time evolution $\qquad e^{-iHt}|\psi_0\rangle \longleftarrow$ some simple state

succinct description but not quite useful

hard to extract information

Solving Many-body Problem

finding ground state is particularly interesting

$$\min_{|\psi\rangle} \langle\psi|H|\psi\rangle - \lambda\langle\psi|\psi\rangle \text{ with } \langle\psi|\psi\rangle = 1$$

QMA-hard for gapless system

intrinsic difficulty

NP-hard for gapped system

if Quantum PCP conjecture is true

QMA-hard

gap between theory and practice:
heuristic algorithm (intuition
& extra information for special instance)
worst-case vs. typical case

1

Some Previous Approach to tackle many-body problem

Mean-field assumption $\rho_k \sim_{n\to\infty} |\psi_1\rangle\langle\psi_1| \otimes \cdots \otimes |\psi_k\rangle\langle\psi_k|$

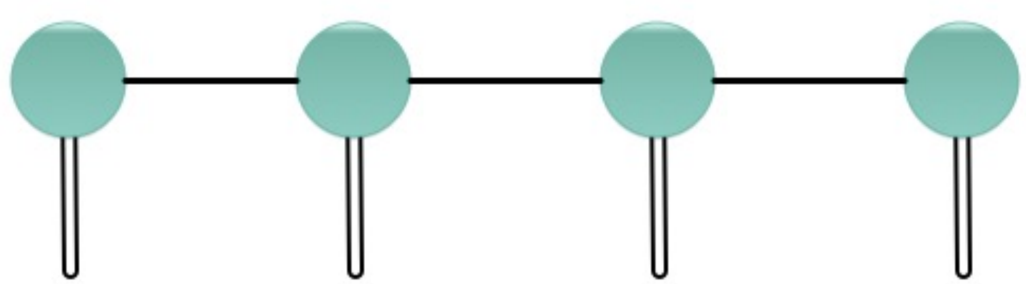
success for bosonic system (Quantum de Finetti)

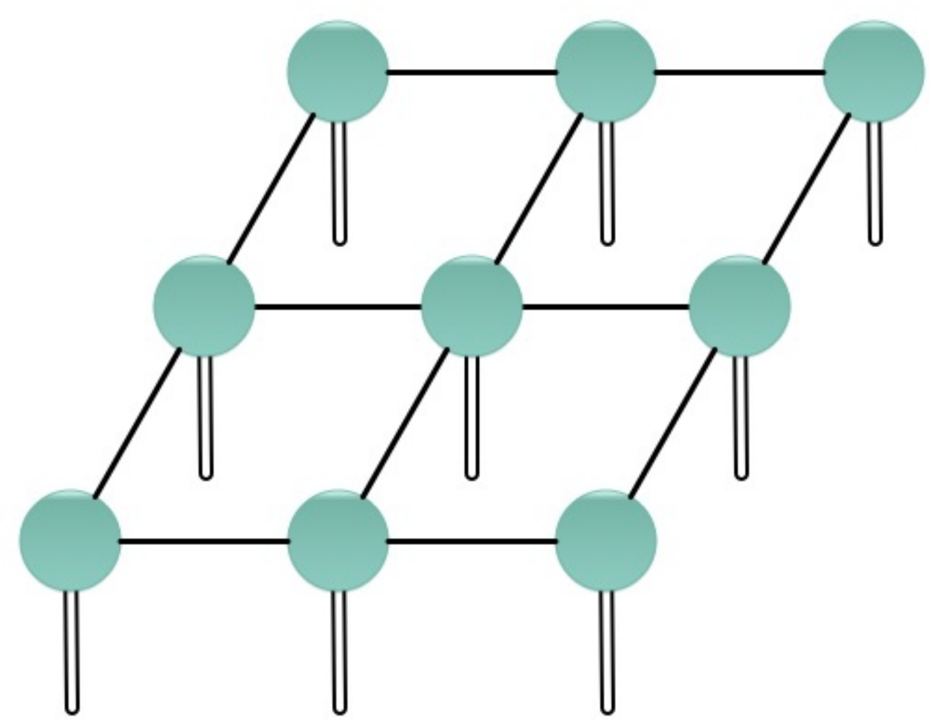
fail for other strong correlated system

Quantum Monte Carlo

$$\frac{\langle\psi(\alpha)|H|\psi(\alpha)\rangle}{\langle\psi(\alpha)|\psi(\alpha)\rangle} \sim \sum_c p_c(\alpha) f_c(\alpha)$$ **sign problem**

Tensor Network





2

Some Previous Approach to tackle many-body problem

very successful for 1D system:

MPS faithfully represent at least gapped system

polynomial time to extract information

heuristic algorithm: DMRG, TEBD, etc

polynomial time algorithm to find ground state

Landau, Zeph, Umesh Vazirani, and Thomas Vidick. "A polynomial time algorithm for the ground state of one-dimensional gapped local Hamiltonians." *Nature Physics* 11.7 (2015): 566-569.

Arad, I., Landau, Z., Vazirani, U., & Vidick, T. (2016). Rigorous RG algorithms and area laws for low energy eigenstates in 1D. *arXiv preprint arXiv:1602.08828*.

2

Some Previous Approach to tackle many-body problem

Schuch, N., Wolf, M. M., Verstraete, F., & Cirac, J. I. (2007). Computational complexity of projected entangled pair states. *Physical review letters*, *98*(14), 140506.

Anshu, A., Arad, I., & Jain, A. (2016). How local is the information in tensor networks of matrix product states or projected entangled pairs states. *Physical Review B*, *94*(19), 195143.

Schwarz, M., Buerschaper, O., & Eisert, J. (2016). Approximating local observables on projected entangled pair states. *arXiv preprint arXiv:1606.06301*.

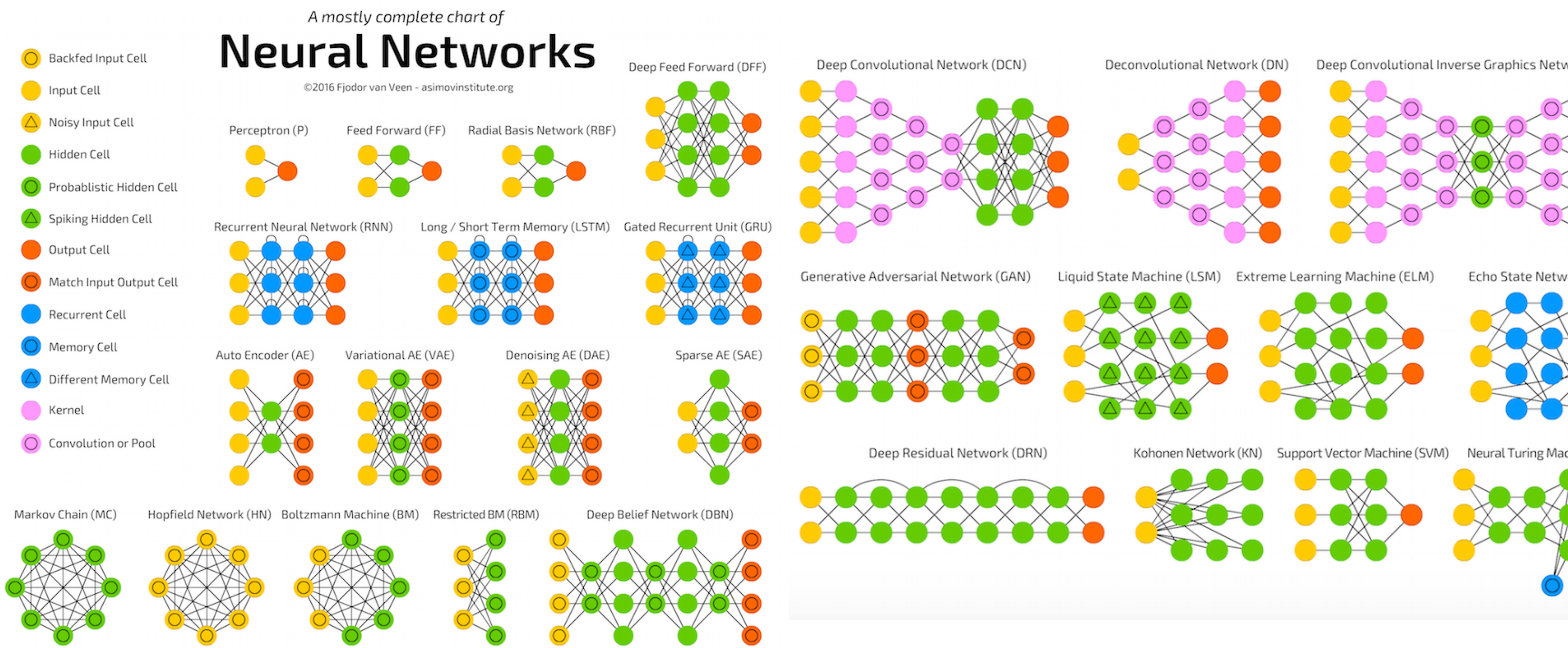
fail for 2D system:

1. unknown whether PEPS is enough
2. extract information is hard
    - i. #P-hard in general case
    - ii. best known approximation algorithm: superpolynomial time under assumptions

1

Neural Network Zoo

Neural Network



http://www.asimovinstitute.org/neural-network-zoo/

natural to use generative model to represent quantum state

1

Restricted Boltzmann Machine (RBM)

Neural Network

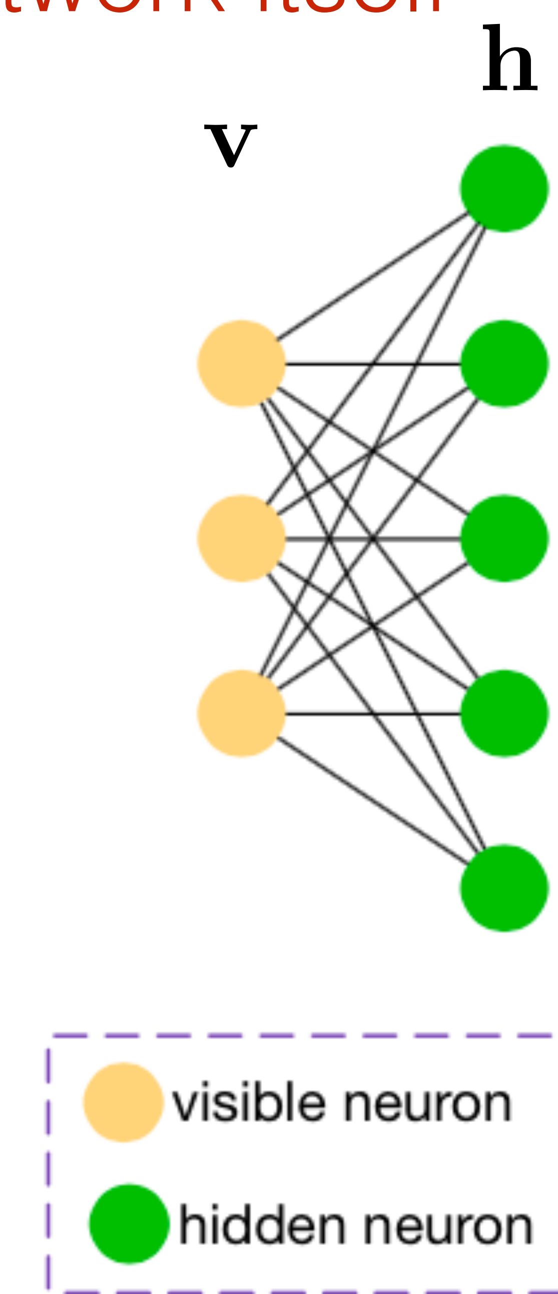
represent quantum state by neural network itself

$$|\psi\rangle = \sum_{\mathbf{v}} \Psi(\mathbf{v})|\mathbf{v}\rangle$$

$$\Psi(\mathbf{v}) = \sum_{\mathbf{h}} e^{\mathcal{W}(\mathbf{v},\mathbf{h})}$$

$$\mathcal{W}(\mathbf{v},\mathbf{h}) = \mathbf{v}^T\mathbf{W}\mathbf{h} + \mathbf{b}^T\mathbf{v} + \mathbf{c}^T\mathbf{h}$$

Weight Function



no intra-layer interactions

2

Restricted Boltzmann Machine (RBM)

Neural Network

Carleo, Giuseppe, and Matthias Troyer. "Solving the quantum many-body problem with artificial neural networks." *Science 2017.*

numerical methods (combined with Monte Carlo)

transverse field Ising: $-h\sum_i \sigma_i^x - \sum_{\langle i,j\rangle} \sigma_i^z \sigma_j^z$

anti-ferromagnetic Heisenberg: $\sum_{\langle i,j\rangle} \sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + \sigma_i^z \sigma_j^z$

better than MPS

Deng, Dong-Ling, Xiaopeng Li, and S. Das Sarma. "Exact machine learning topological states." *arXiv preprint arXiv:1609.09060* (2016).

exact representation for topological states

SPT: 1D cluster state

topological order: toric code

2

Restricted Boltzmann Machine (RBM)

Chen, J., Cheng, S., Xie, H., Wang, L., & Xiang, T. (2017). On the Equivalence of Restricted Boltzmann Machines and Tensor Network States. *arXiv preprint arXiv:1701.04831*.

restricted Boltzmann Machine —> Tensor Network

Tensor Network —> RBM with given architecture

Deng, Dong-Ling, Xiaopeng Li, and S. Das Sarma. "Quantum Entanglement in Neural Network States." *PhysRevX.7.021021*.

area law (local connection)

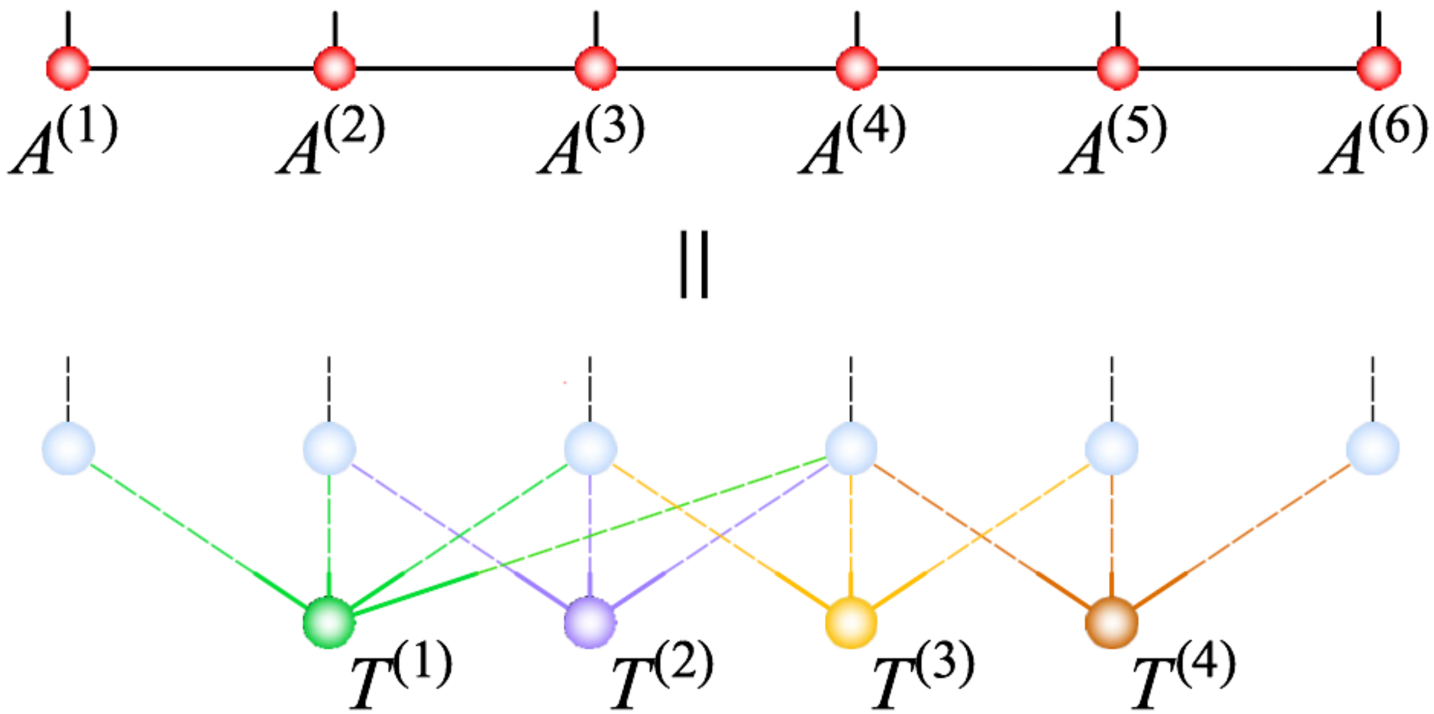
random RBM: entanglement spectrum, not thematize

Huang, Yichen, and Joel E. Moore. "Neural network representation of tensor network and chiral states." *arXiv preprint arXiv:1701.06246* (2017).

using Grassmann number

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

Limitation of RBM



limitation to represent Tensor Network (given architecture)

Copy from

Chen, J., Cheng, S., Xie, H., Wang, L., & Xiang, T. (2017). On the Equivalence of Restricted Boltzmann Machines and Tensor Network States. *arXiv preprint arXiv:1701.04831*.

Limitation of RBM

Universal Approximation Theorem

In order to approximate probability distribution with **k** boolean variables, the number of hidden neurons is at most **2^k** *without constraints on architecture* (analog to bond dimension in MPS)

Le Roux, Nicolas, and Yoshua Bengio. "Representational power of restricted Boltzmann machines and deep belief networks." *Neural computation* 20.6 (2008): 1631-1649.

reasonable to consider:

Efficient Representation:

poly(n) parameters

without constraints on architecture

3

Limitation of RBM

under reasonable complexity assumptions
RBM can not represent the following state efficiently:

state generated by efficient quantum computer
PEPS and other tensor network state
ground state of local gapped Hamiltonian

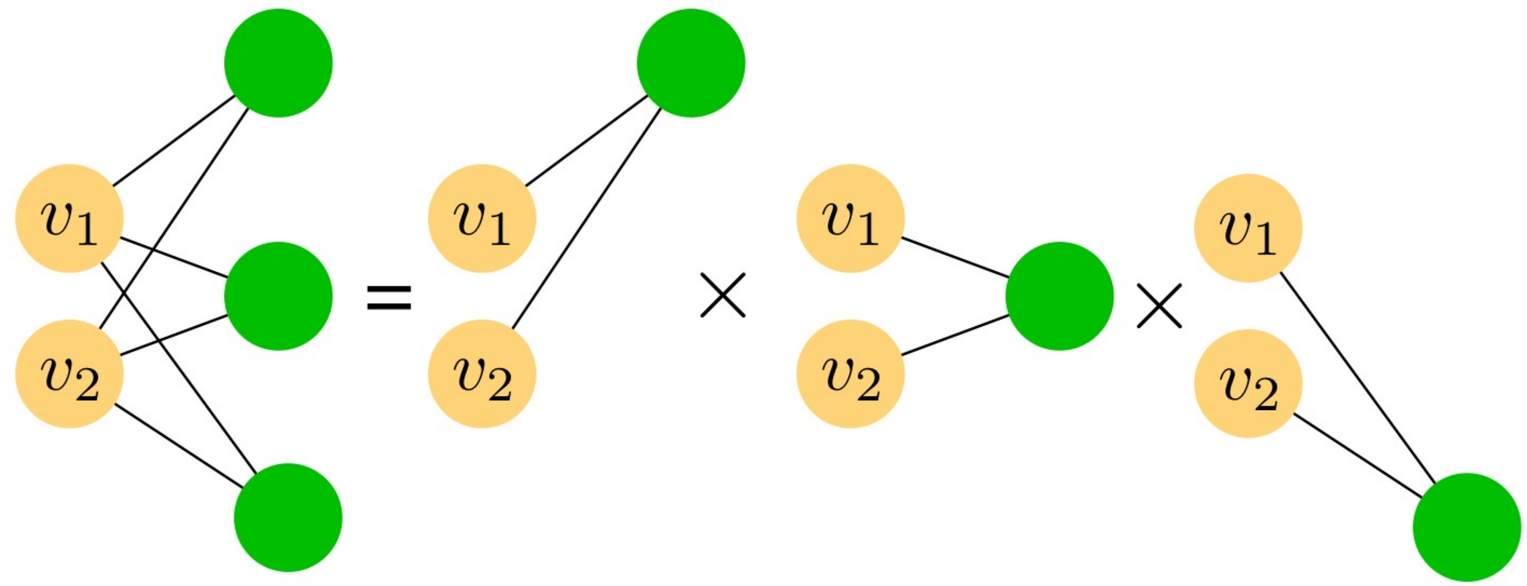
not closed in a quantum phase and dynamics !

otherwise polynomial hierarchy collapse
(a generalization of *P=NP*)

Gao, Xun, and Lu-Ming Duan. "Efficient Representation of Quantum Many-body States with Deep Neural Networks." *arXiv preprint arXiv:1701.05039* (2017).

Proof of Limitation for RBM

$\Psi(\mathbf{v})$ for RBM
can be factorized



RBM:
hidden neurons are conditionally independent

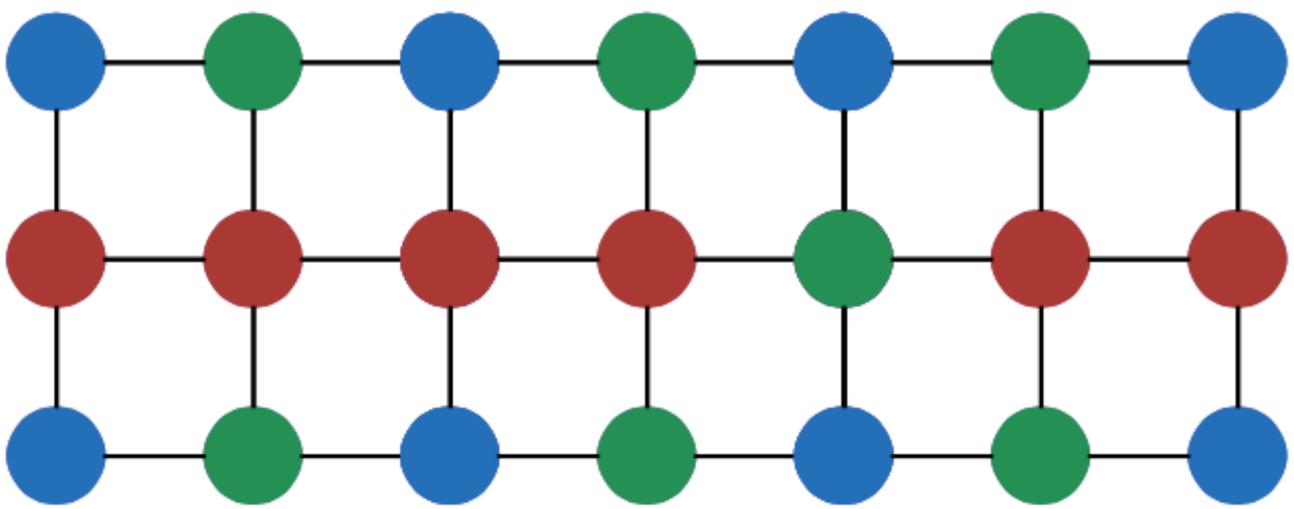
P/poly: can be solved in polynomial size circuit
the circuit not easy to construct

P/poly: $\Psi(\mathbf{v})$
polynomial-size boolean circuit $\Psi(\cdot)$
input of function $\mathbf{v}$

no need to know how to construct RBM (circuit) !

Proof of Limitation for RBM

cluster state: ignore the colors



ground state of

$$-\sum_i \sigma_i^x \prod_{j\in\text{neighbors of } i} \sigma_j^z$$

time evolution on $|+\rangle^{\otimes N}$

$$\mathcal{H} = -\sum_{\langle i,j\rangle} J Z_i Z_j + \sum_i B_i Z_i \qquad J = \frac{\pi}{4}$$

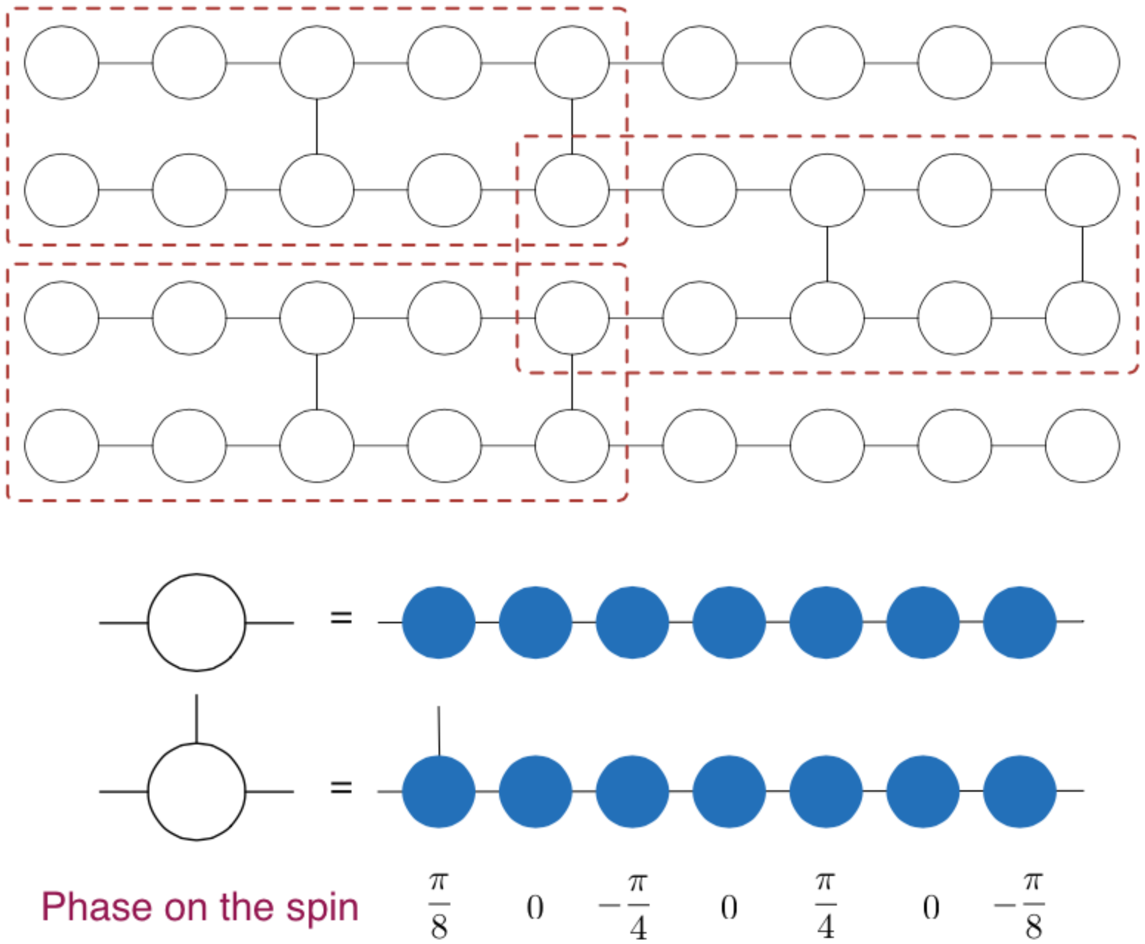
i. state generated from simple dynamics on simple initial state

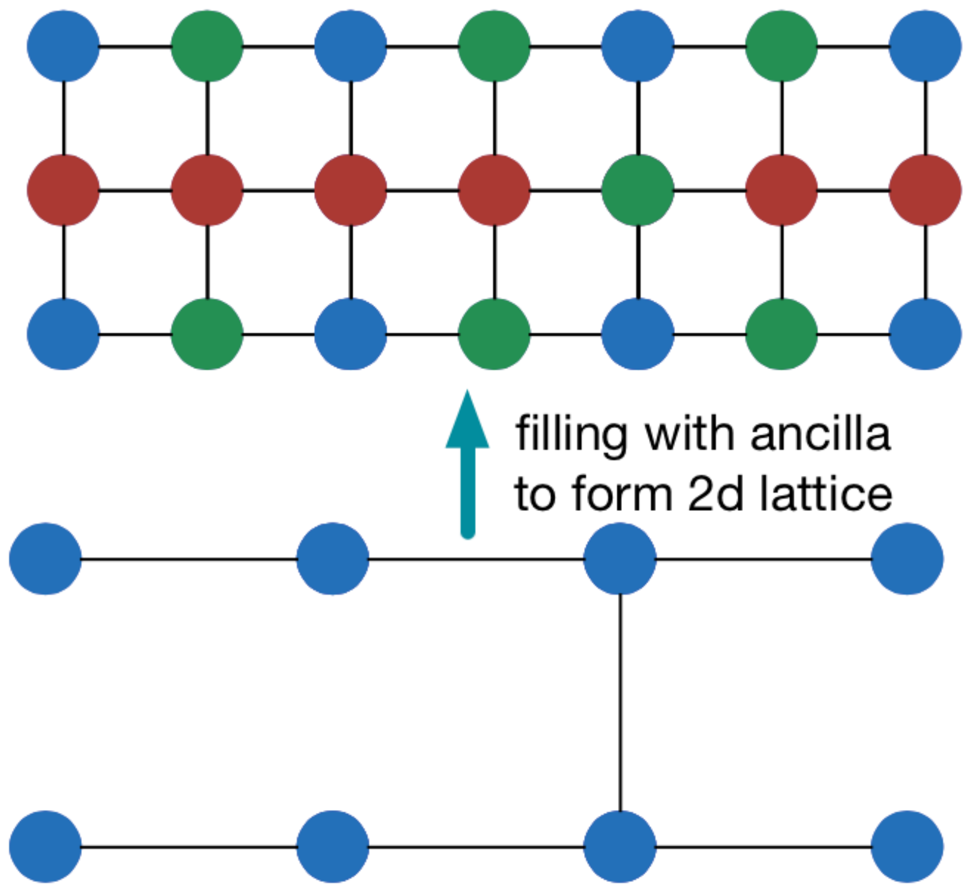
ii. ground state of gapped, local, commuting Hamiltonian

iii. PEPS

Proof of Limitation for RBM

GWD state:





cluster state with single qubit unitary

blue: z-axis with angles, then Hadamard

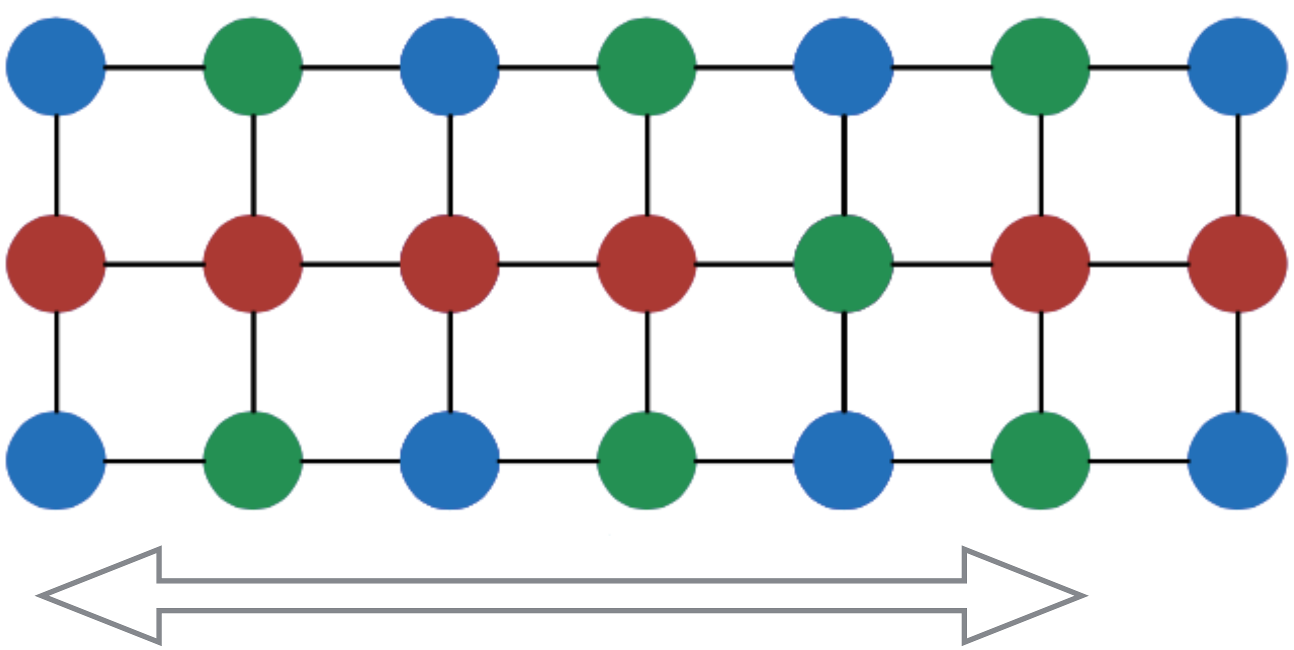
green: Hadamard **red**

Gao, Xun, Sheng-Tao Wang, and Lu-Ming Duan. "Quantum Supremacy for Simulating A Translation-Invariant Ising Spi Model." *Phys.Rev.Lett*. 2017

Proof of Limitation for RBM

translational invariant ! only depends on size

cluster state is resource state in measurement-based quantum computing

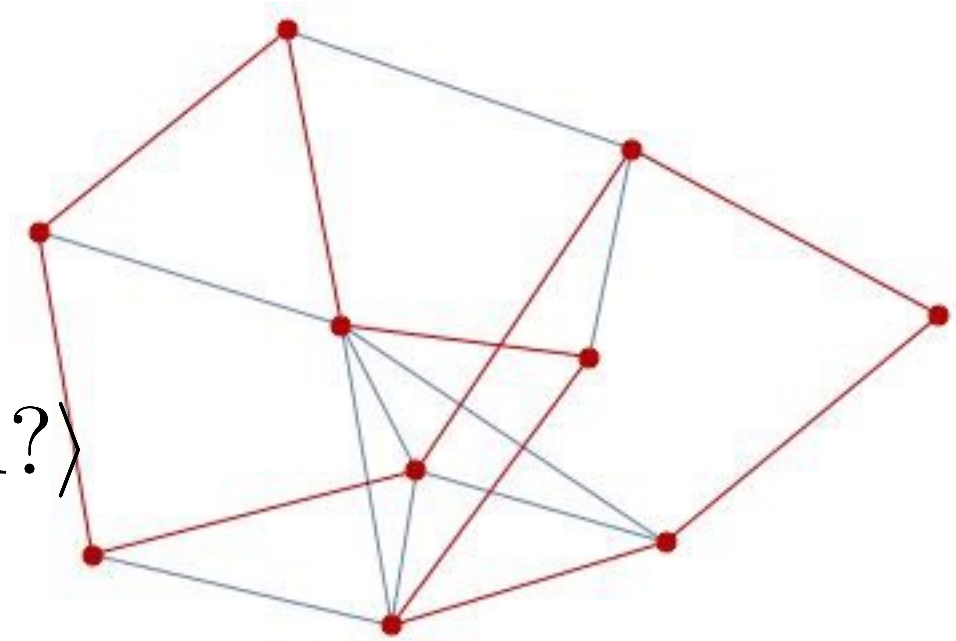


universal quantum computing

#P-hard (much harder than NP): $\Psi(v)$ of GWD state

$$\sum |\text{edge configuration}\rangle|\text{Hamiltonian Path?}\rangle$$

counting the number of True



5

Proof of Limitation for RBM

Neural Network

translational invariant ! only depends on size

if $\Psi(v)$ of $\mathrm{GWD\ state}$ can be represented by RBM efficiently

$\mathsf{P}^{\#P} \subseteq \mathsf{P}^{\mathsf{P/poly}}$

widely believed not true:

polynomial hierarchy collapse !

generalization P=NP, very unlikely

extend to approximate case in terms of trace distance

with another reasonable complexity assumptions

related to quantum supremacy

for sampling random quantum circuit

supported by quantum chaos theory

Google's quantum supremacy plan

Boixo, Sergio, et al. "Characterizing quantum supremacy in near-term devices." *arXiv preprint arXiv: 1608.00263* (2016).

5

Proof of Limitation for RBM

under reasonable complexity assumptions
RBM can not represent the following state efficiently:

state generated by efficient quantum computer
PEPS and other tensor network state
ground state of local gapped Hamiltonian

not closed in a quantum phase and dynamics !

otherwise polynomial hierarchy collapse
(a generalization of $P=NP$)

Characterize RBM state

Although limitation of RBM still useful in practice

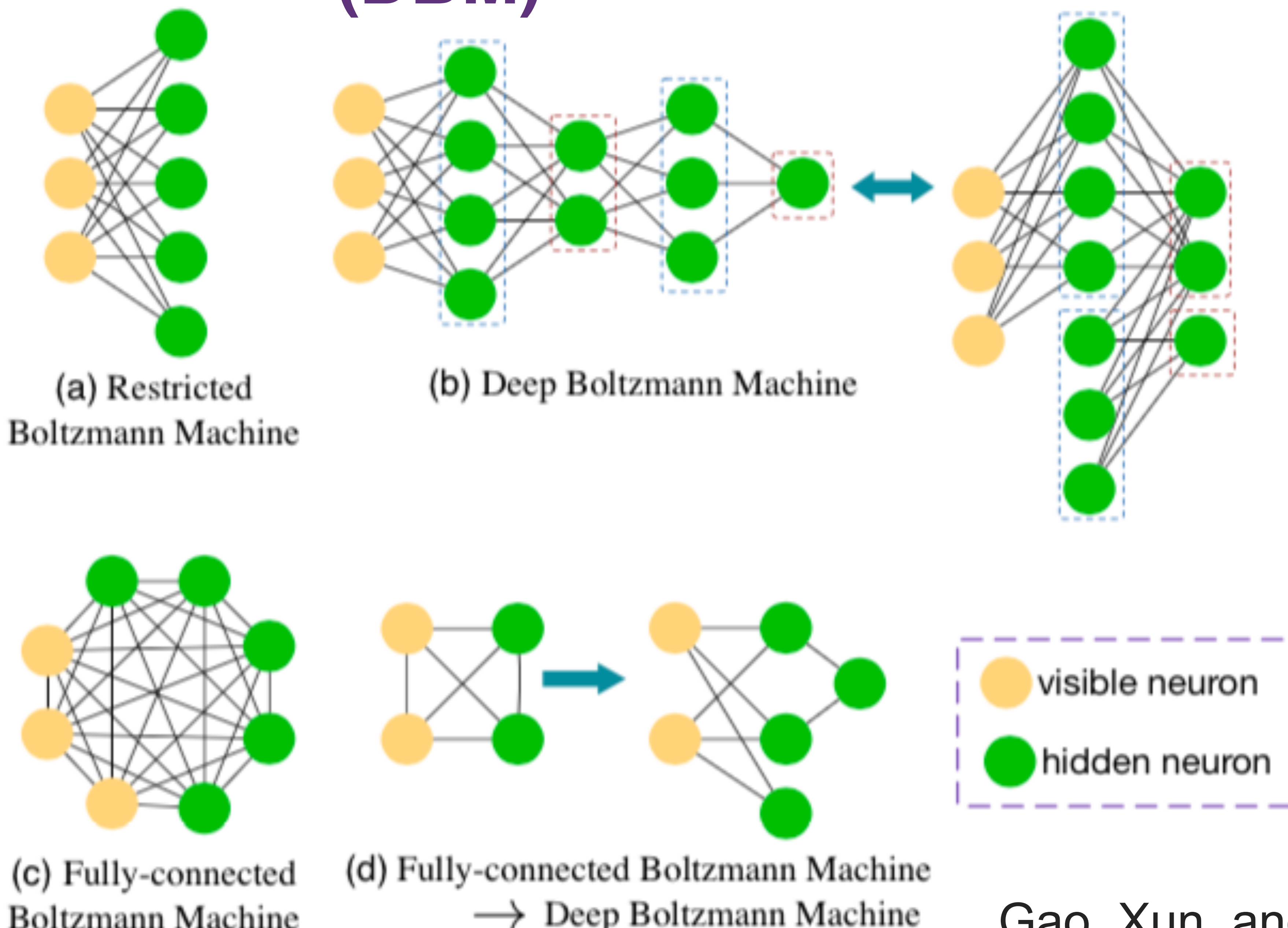
how to characterize the state represented by RBM?

string-bond state?

1

Deep Boltzmann Machine (DBM)

Deep Neural Networks



DBM with 2 hidden layers captures fully-connected Boltzmann Machine and deeper DBM

Gao, Xun, and Lu-Ming Duan. "Efficient Representation of Quantum Many-body States with Deep Neural Networks." *arXiv preprint arXiv: 1701.05039* (2017).

1

Deep Boltzmann Machine (DBM)

Deep Neural Networks

Deep vs. Shallow

DBM can represent the following state efficiently:

quantum computing or dynamics: $O(nT)$ — circuit depth or evolution time

ground state: $O\left(\frac{1}{\Delta}\left(n+\log\frac{1}{\epsilon}\right)m^2\right)$

energy gap

number of local terms in Hamiltonian

gapless, even non-local

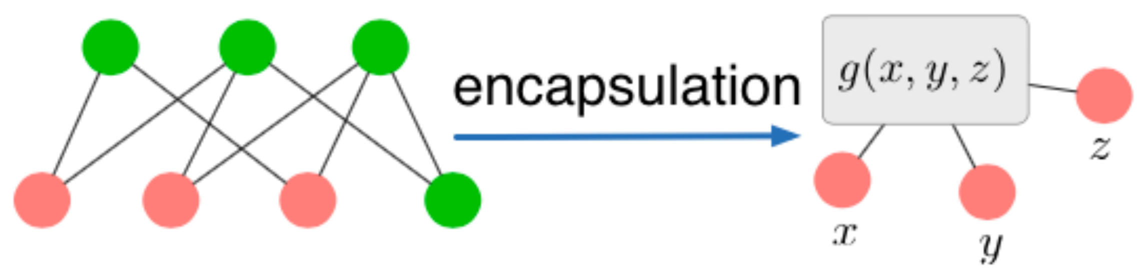
tensor network state: $O(D^{2d}n)$ D: bond dimension

d: coordination number

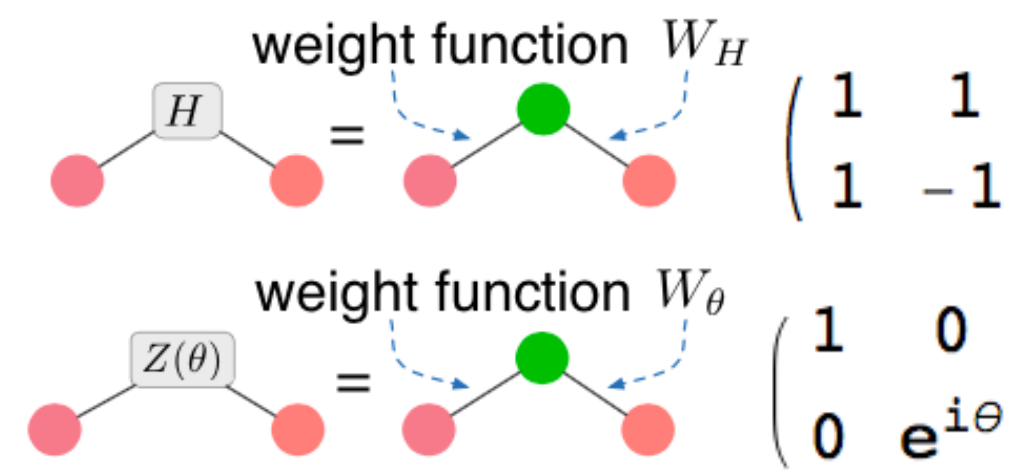
closed in a quantum phase and dynamics

2

Proof of Power of Deep Boltzmann Machine



(a) Gadget

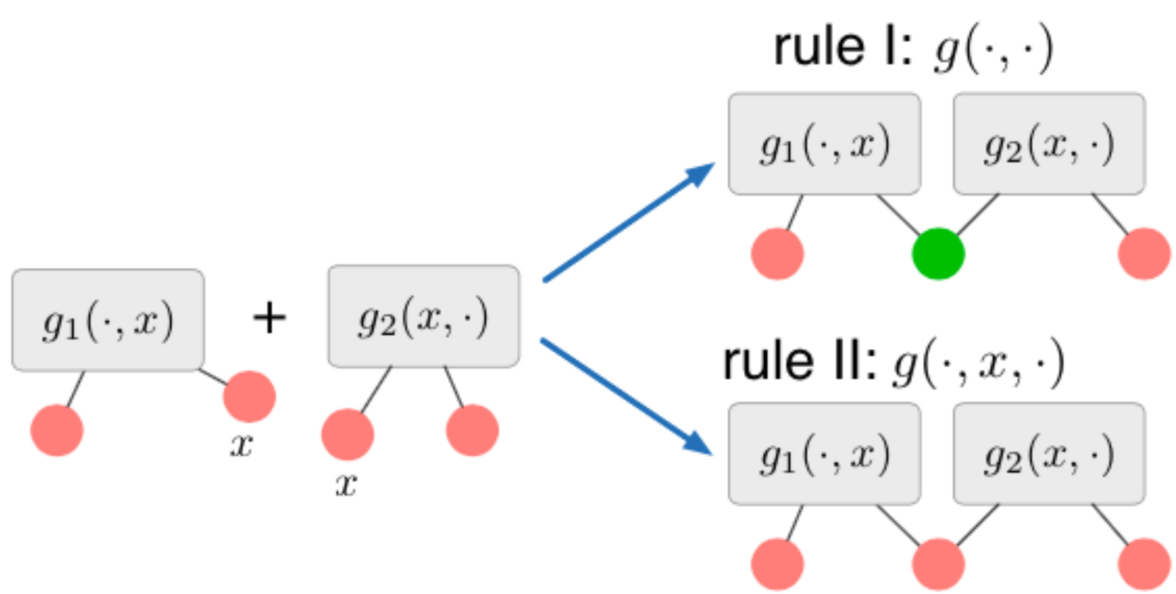


(b) Hadamard Gadget & Phase Gad

gadget: DBM-like function

$$W_\theta(x, h) = \frac{\theta}{2} ix + i\pi xh.$$

$$W_H(x, h) = \frac{\pi}{8} i - \frac{\pi}{2} ix - \frac{\pi}{4} ih + i\pi$$

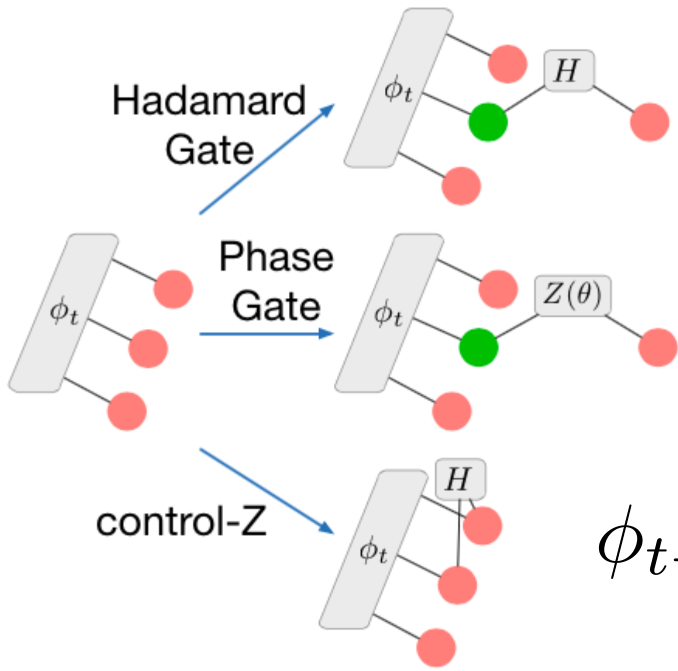


$$\text{rule I: } g(\cdot, \cdot) = \sum_x g_1(\cdot, x) g_2(x, \cdot),$$

$$\text{rule II: } g(\cdot, x, \cdot) = g_1(\cdot, x) g_2(x, \cdot),$$

Proof of Power of Deep Boltzmann Machine

Quantum Computing or Dynamics



(d) Construction of ϕ_{t+1}

rule I simulates matrix multiplication

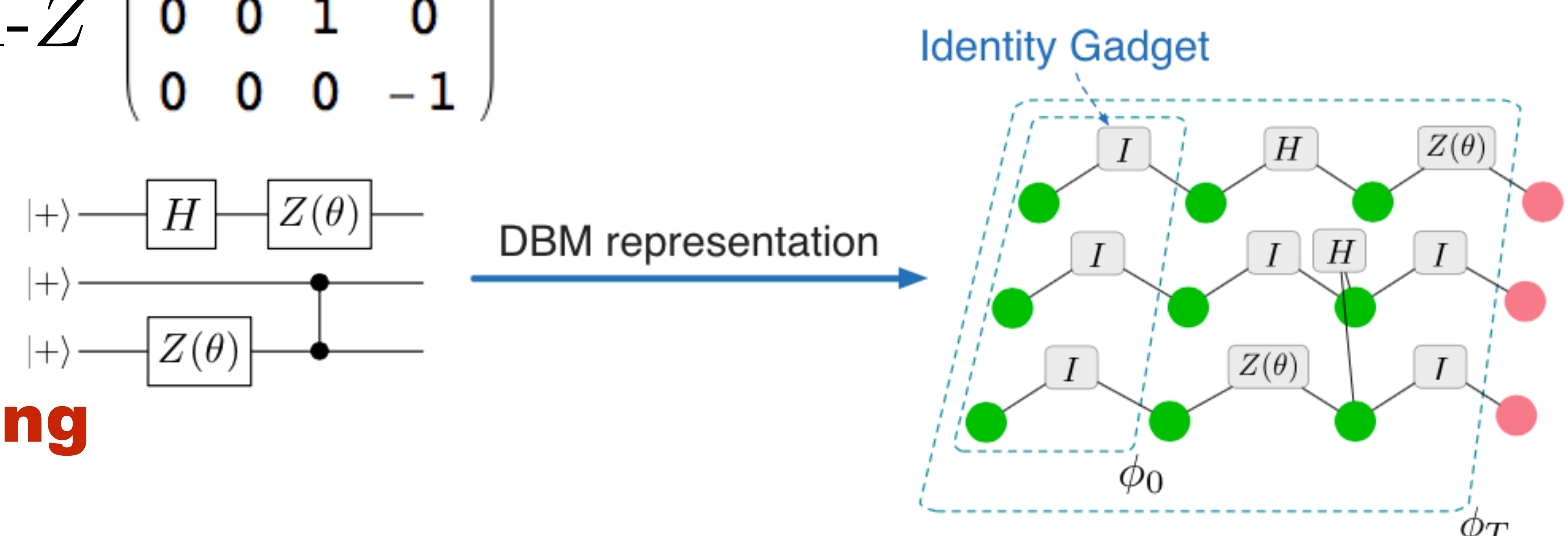
phase gadget is not necessary
absorbed into bias term

$$\phi_{t+1}(\cdots x_i, x_{i+1}\cdots) = (-1)^{x_i x_{i+1}} \phi_t(\cdots x_i, x_{i+1}, \cdots)$$

rule II simply multiplication

$$H, Z(\theta), \text{control-}Z \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

universal set for quantum computing



2

Proof of Power of Deep Boltzmann Machine

Quantum Computing or Dynamics

dynamics: quantum simulation

Lloyd, S. Universal quantum simulators. *Science* **273**, 1073 (1996).
Poulin, D., Qarry, A., Somma, R. & Verstraete, F. Quantum simulation of time-dependent hamiltonians and the convenient illusion of hilbert space. *Physical review letters* **106**, 170501 (2011).
Berry, D. W., Childs, A. M., Cleve, R., Kothari, R. & Somma, R. D. Simulating hamiltonian dynamics with a truncated taylor series. *Physical review letters* **114**, 090502 (2015).

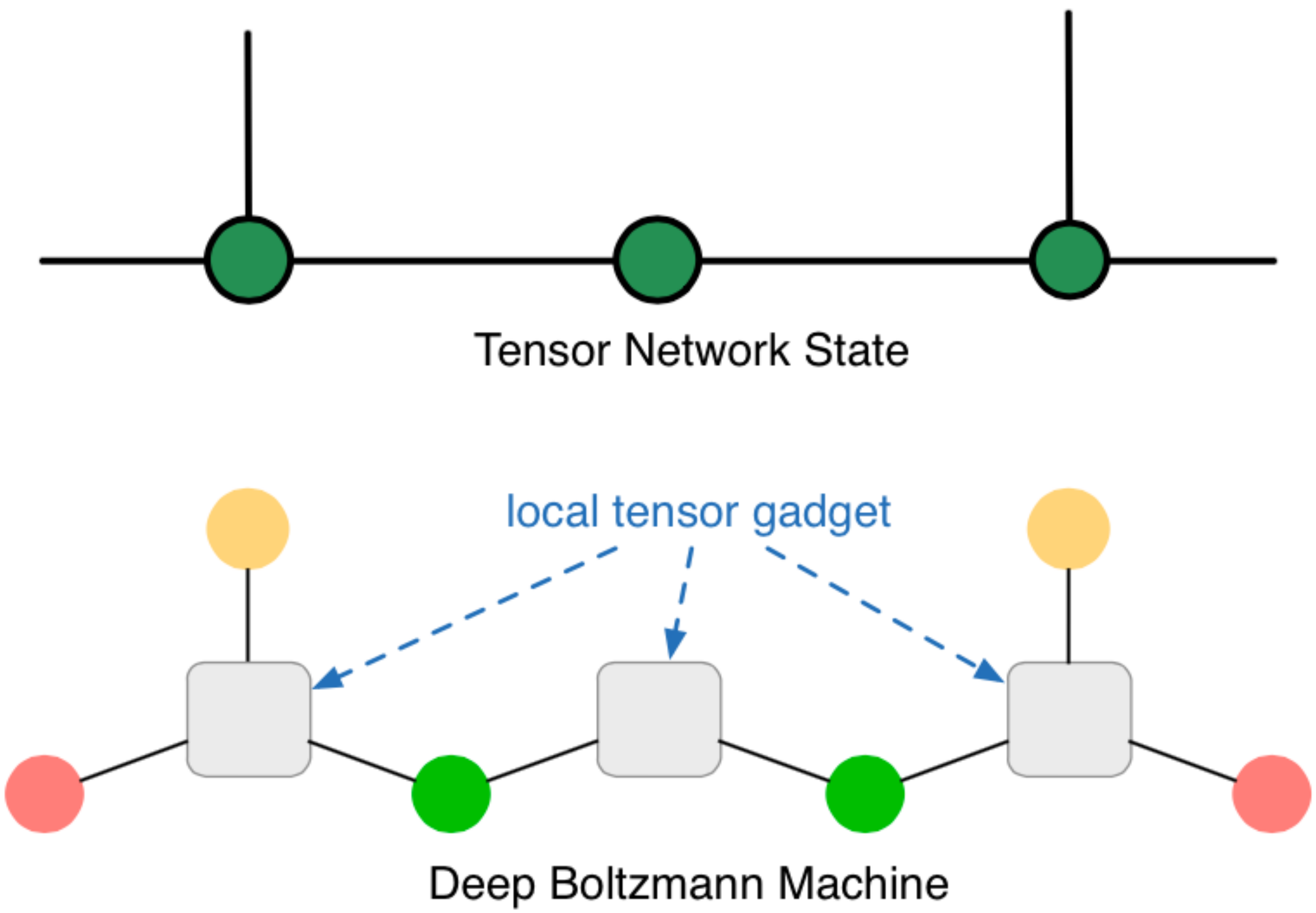
closed under a quantum phase: simulating adiabatic evolution

2

Proof of Power of Deep Boltzmann Machine

Deep Neural Networks

Tensor Network



$A_{x_1 \cdots x_c}$ local tensor

using quantum computing gadget:

$$|A\rangle = \sum_{x_1, \cdots, x_c} A_{x_1 \cdots x_c} |x_1, \cdots, x_c\rangle$$

at most square of Hilbert space dimension

rule I simulates contraction of bond index

2

Proof of Power of Deep Boltzmann Machine

Tensor Network

see also

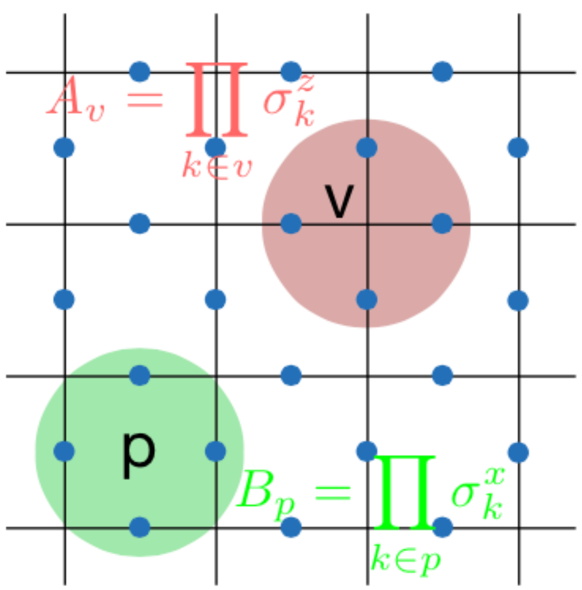
Chen, J., Cheng, S., Xie, H., Wang, L., & Xiang, T. (2017). On the Equivalence of Restricted Boltzmann Machines and Tensor Network States. *arXiv preprint arXiv:1701.04831*.

Huang, Yichen, and Joel E. Moore. "Neural network representation of tensor network and chiral states." *arXiv preprint arXiv:1701.06246* (2017).

why not use universal approximate theorem representing $A_{x_1 \cdots x_c}$ directly ?

physical property is very sensitive to local tensor !

toric code



$$T_{ijkl} = 1 \rightarrow T_{ijkl} = 1, \text{ if } i+j+k+l = 0 \bmod 2$$
$$T_{ijkl} = 0 \rightarrow T_{ijkl} = \epsilon, \text{ if } i+j+k+l = 1 \bmod 2$$

so long as $\epsilon \neq 0$ topological entanglement entropy disappears !

Chen, X., Zeng, B., Gu, Z. C., Chuang, I. L., & Wen, X. G. (2010). Tensor product representation of a topological ordered phase: Necessary symmetry conditions. *Physical Review B*, *82*(16), 165119.

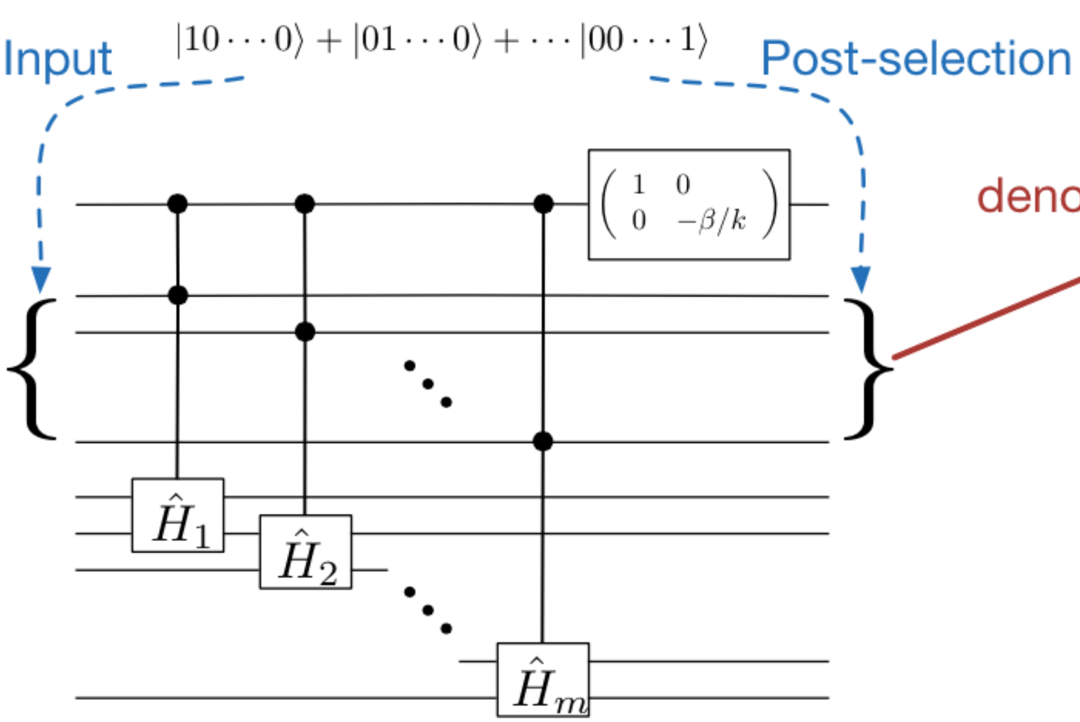
Proof of Power of Deep Boltzmann Machine

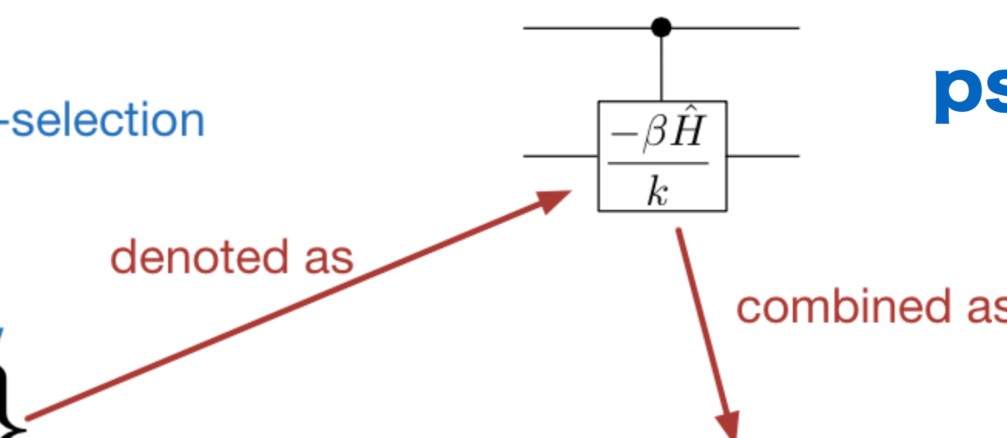
Ground State

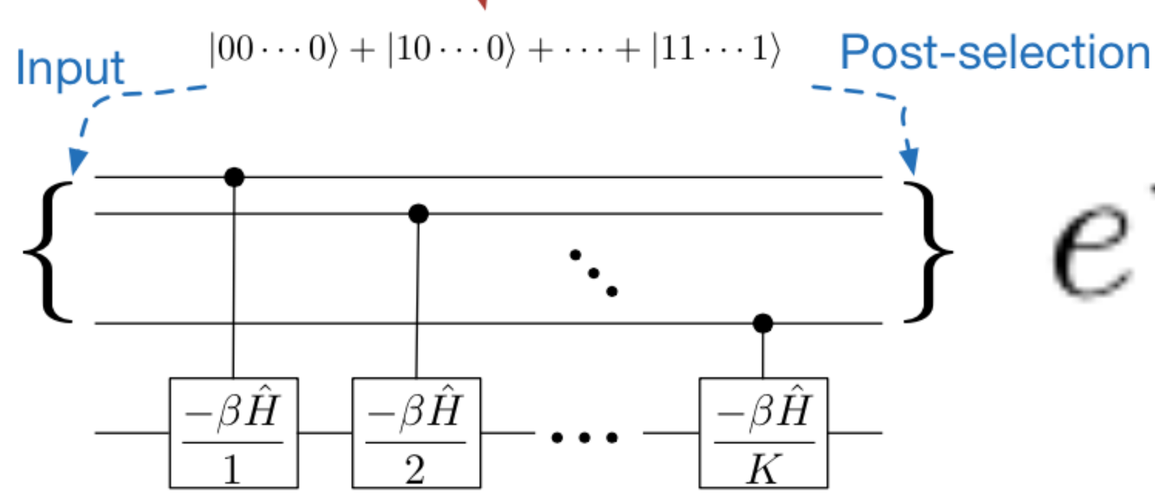
inspired by

Berry, D. W., Childs, A. M., Cleve, R., Kothari, R., & Somma, R. D. (2015). Simulating Hamiltonian dynamics with a truncated Taylor series. *Physical review letters*, *114*(9), 090502.

using Taylor series instead of Trotter decomposition: exponential improvement on precision







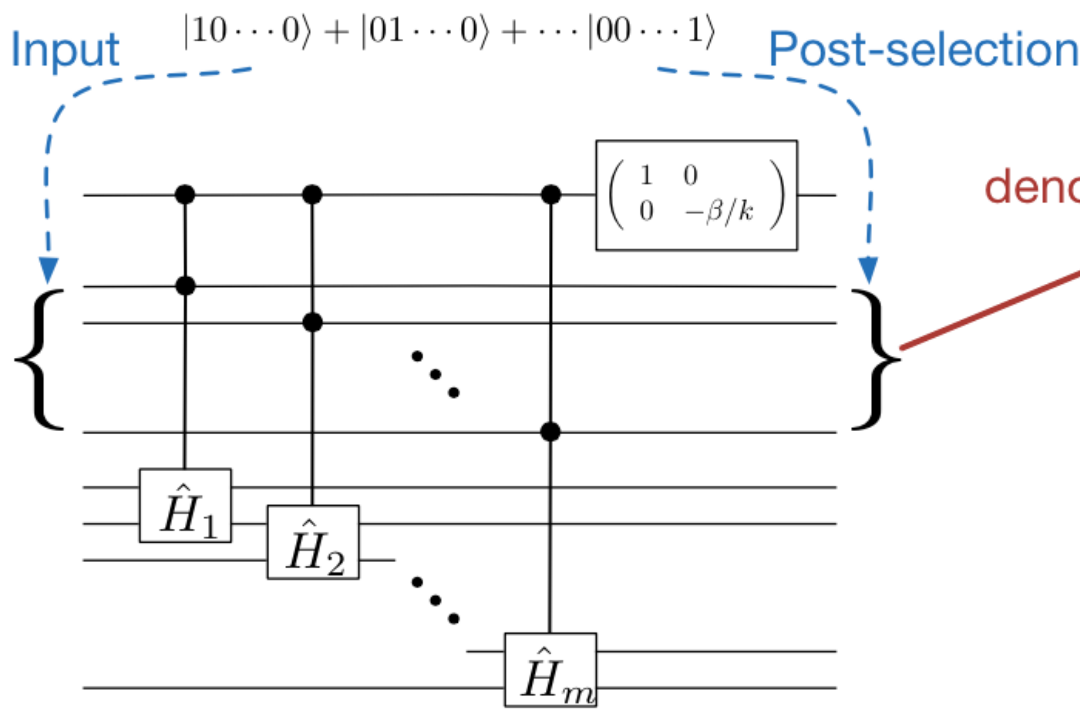
pseudo quantum circuit
non-unitary
also tensor network

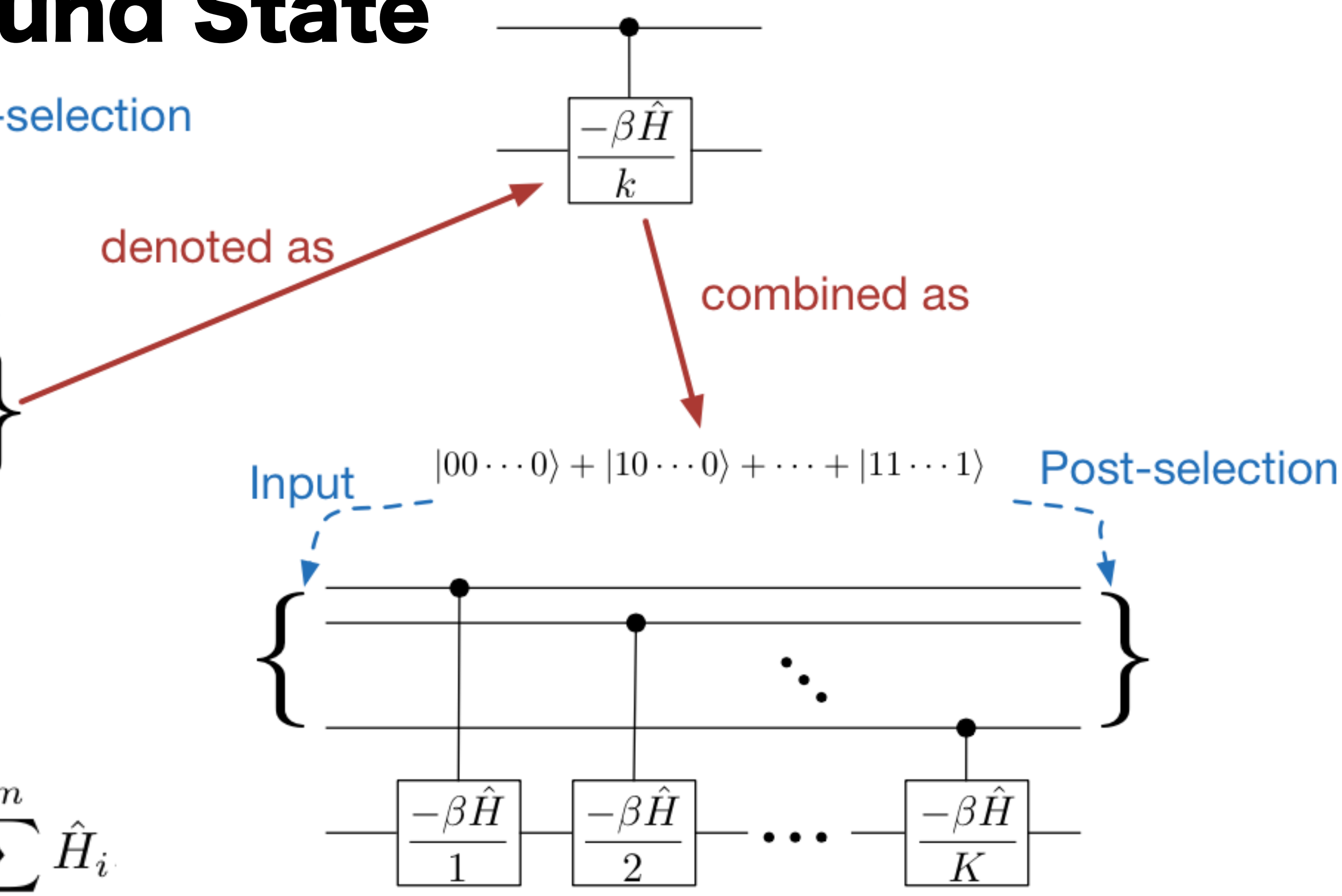
$$e^{-\beta\hat{H}}$$

2

Proof of Power of Deep Boltzmann Machine

Ground State





$$\sum_{i=1}^{m}|i\rangle \xrightarrow{\text{controlled-}\hat H_i} \sum_{i=1}^{m} |i\rangle \hat H_i \xrightarrow{\text{post-selection}} \sum_{i=1}^{m} \hat H_i$$

$$\sum_{k=0}^{K} |k\rangle \xrightarrow{\text{controlled-}(-\beta\hat H/k)} \sum_{k=0}^{K} |k\rangle \frac{(-\beta\hat H)^k}{k!} \xrightarrow{\text{post-selection}} \sum_{k=0}^{K} \frac{(-\beta\hat H)^k}{k!}$$

$$\sum_{i=0}^{2^n-1} |i\rangle|i\rangle = \sum_{i=0}^{2^n-1} |\psi_i\rangle|\psi_i^*\rangle \longrightarrow |\psi_0\rangle\langle\psi_0| + O\left(2^n e^{-\beta\Delta} + \frac{2^n(\beta\|\hat H\|)^K}{K!}\right)$$

3

Potential of Deep Boltzmann Machine

non-local even non-sparse property (compared to PEPS):

proved result for representing ground state beyond 1D

appropriate for highly entanglement state like time-evolution

expected to have less parameters

deeper compared to RBM

harder to extract information or training ?

both #P-complete (computing local observable)

Training DBM

Extract Information

prototype Monte-Carlo algorithm

$$\langle\psi|O|\psi\rangle = \frac{\sum_h p_h f_h}{\sum_h p_h g_h}$$

$\frac{p_{h'}}{p_h}$ easy to compute

Metroplis algorithm $Pr(h \to h') = \min\left(1, \frac{p_{h'}}{p_h}\right)$

f and g are easy to compute

fluctuation is too large?

local minimum?

the same problems also occur in RBM ?

Training DBM

seems hard to train (sign problem)

inevitable! (intrinsic) trade off between
representational power & computational difficulty

even though
our work shows we can benefit a lot from depth
like in deep learning

quasi DBM/RBM
fewer neurons in the second hidden layer

or other deep architecture

Summary

Restricted Boltzmann Machine:

- limitation for dynamics (quantum phase), PEPS, GS
- tool of proof: complexity theory
- conjecture, string-bond state

Deep Boltzmann Machine:

- most of physical states (physical relevant corner?)
- dynamics, tensor network
- ground state (even gapless & non-local)

Training Deep Boltzmann Machine:

- Prototype Monte Carlo
- connection with other fields