

テンソルネットワークによる情報圧縮と 物性物理への応用

Information compression by tensor networks
and its application to condensed matter physics

東京大学大学院理学系研究科 物理学専攻 大久保 肇
Dept. of Phys., The University of Tokyo, **Tsuyoshi Okubo**

e-mail: t-okubo@phys.s.u-tokyo.ac.jp



東京大学
THE UNIVERSITY OF TOKYO

cdmsi



Computational
Science
Alliance
The University of Tokyo

Self introductions

- 1999 -2008
 - PhD, Kyushu University (Supervisor: Prof. T. Odagaki)
 - Random packing of discs, socio physics, transport problem
- 2008 - 2012
 - Postdoc, Osaka University (Supervisor: Prof. H. Kawamura)
 - **Frustrated magnetism (mainly, classical spin systems)**
- 2012 - 2017
 - Postdoc, ISSP, the University of Tokyo (Supervisor: Prof. N. Kawashima)
 - Quantum spin systems, Tensor network
- 2017 -
 - Project lecturer, The University of Tokyo

Research interest:

Many body physics, phase transition, novel phases of matters, non-equilibrium dynamics, ...

Methods:

Numerical simulations

Contents

- Background
 - Huge data in physics
 - Outline of tensor network states
- Entanglement
 - Entanglement entropy and its area law
- Matrix product state
 - Matrix product state (MPS)
- Generalization to tensor network states
 - Breakdown of MPS representation in higher dimensions
 - Tensor Product States (TPS)
- Application of TPS for frustrated spin systems
- Concluding remarks

Background

Huge data in physics

Many-body problems in physics

- Celestial movement (天体運動)
- Gases, Liquids
- Molecules, Polymers (eg. Proteins), ...
- Electrons in molecules and solids
- Elemental particles (Quantum Chromo Dynamics)
(量子色力学)

In these problems, "systems" contain huge degrees of freedoms:

6 N -dimensional phase space for classical mechanics

$O(e^N)$ -dimensional Hilbert space for quantum system

Why we need information compression?

We can not treat entire data in the present computers.

Available memories in the present computers

Double precision real number
= 8 Byte
 $\sim 10^9$

Personal computers: ~10 GB

Super computers: ~100 GB / node $\sim 10^{10}$

K@RIKEN,
Oakforest-PACS@UTokyo
and Tsukuba Univ,
Sekirei@ISSP, UTokyo

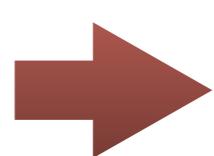
$\sim 1 \text{ PB}$ $\sim 10^{14}$
(whole system)

...

Notice: In quantum system, the size of Hilbert space is $O(e^N)$

Why we need information compression?

We can not treat entire data in the present computers



Try to reduce the "effective" dimension of
(Hilbert) space

By taking proper basis set,
we can represent a quantum state efficiently.

- Krylov subspace
- Matrix product state
- Tensor network states
- ...

Classification of Information Compression by Memory Costs

Linear algebra for huge data: $\vec{v} \in \mathbb{C}^M$

(1) A matrix can be stored

Required memory $\sim O(M^2)$

(2) Although a matrix cannot be stored, vectors can be stored

Required memory $\sim O(M)$

(3) A vector cannot be stored

Required memory $\ll O(M)$

We try to **approximate** a vector in a compact form.

$$M \sim a^N \rightarrow \text{Memory} \sim O(N^x)$$

Exponential **Polynomial**

N : problem size (e.g. system size)

When we efficiently compress a vector?

$$\vec{v} = \sum_{i=1}^M C_i \vec{e}_i \quad \vec{v} \in \mathbb{C}^M$$

If we can find a basis where the coefficients have a structure (correlation).

(1) Almost all C_i are zero (or very small).



We store only a few finite elements $\{(i, C_i)\}$

E.g.

Fourier transformation $\vec{v} = \sum_{k=1}^M D_k \vec{f}_k$

If we can neglect larger wave numbers, we can efficiently approximate the vector with smaller number of coefficients.

Classical state $|\Psi\rangle = |01011\dots00\rangle$

In this case, we know that only a specific C_i is non-zero.

We need only an integer corresponding to the non-zero element.

When we efficiently compress a vector?

$$\vec{v} = \sum_{i=1}^M C_i \vec{e}_i \quad \vec{v} \in \mathbb{C}^M$$

(2) All of C_i are not necessarily independent.

→ We store "**structure**" and "**independent elements**".

$$\{(i, C_i)\}$$

E.g. Product state ("generalized" classical state)

A vector is decomposed into **product of small vectors**.

$$|\Psi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \otimes \cdots$$

e.g.

$$|\phi_1\rangle = \alpha|0\rangle + \beta|1\rangle$$

$$|\phi_1\rangle = |01\rangle - |10\rangle$$

structure: "**product state**"

independent elements: **small vectors**

Tensor network decomposition of a vector

Target:

Exponentially large Hilbert space

$$\vec{v} \in \mathbb{C}^M \text{ with } M \sim a^N$$

+

Total Hilbert space is decomposed as
a product of "local" Hilbert space.

$$\mathbb{C}^M = \mathbb{C}^a \otimes \mathbb{C}^a \otimes \cdots \otimes \mathbb{C}^a$$

eg. quantum many body systems

→ **Tensor network decomposition**

$$v_i = v_{i_1, i_2, \dots, i_N} = \sum_{\{x\}} T^{(1)}[i_1]_{x_1, x_2, \dots} T^{(2)}[i_2]_{x_1, x_3, \dots} \cdots T^{(N)}[i_N]_{x_3, x_{100}, \dots}$$

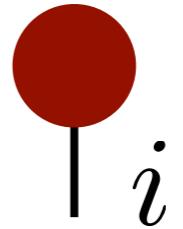
$i_n = 0, 1, \dots, a - 1$: index of local Hilbert space

$T[i]_{x_1, x_2, \dots}$: local tensor for "state" i

Graphical representations for tensor network

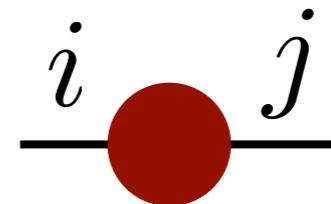
- Vector

$$\vec{v} : v_i$$



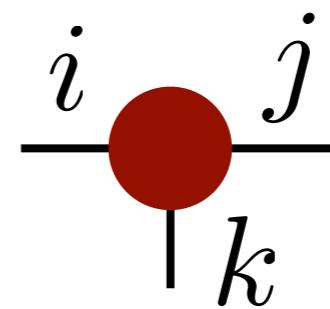
- Matrix

$$M : M_{i,j}$$



- Tensor

$$T : T_{i,j,k}$$



* **n-rank tensor = n-leg object**

When indices are not presented in a graph, it represent a tensor itself.

$$\vec{v} = \text{---} \bullet \text{---}$$

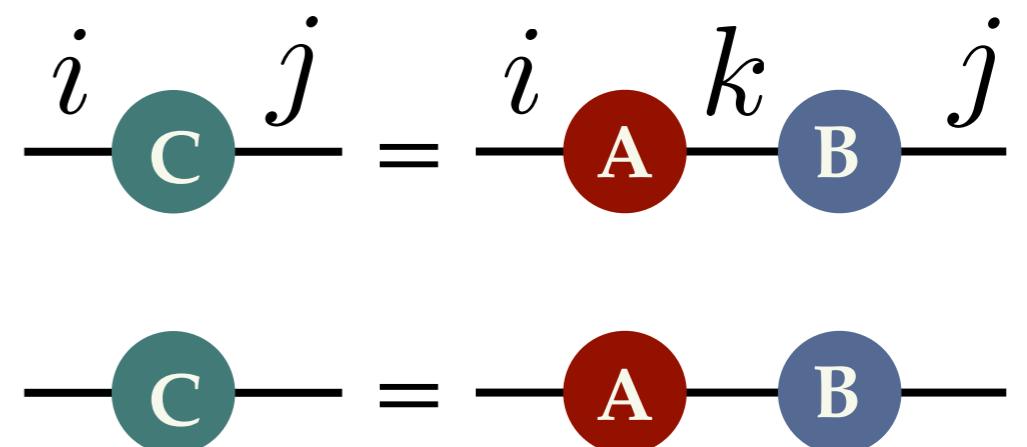
$$T = \text{---} \bullet \text{---}$$

Graphical representations for tensor network

Matrix product

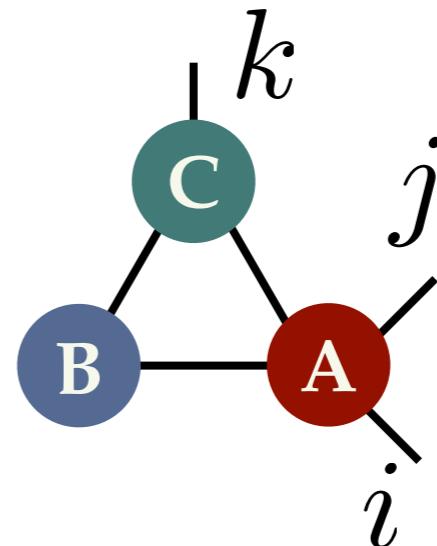
$$C_{i,j} = (AB)_{i,j} = \sum_k A_{i,k} B_{k,j}$$

$$C = AB$$



Generalization to tensors

$$\sum_{\alpha, \beta, \gamma} A_{i,j,\alpha,\beta} B_{\beta,\gamma} C_{\gamma,k,\alpha}$$

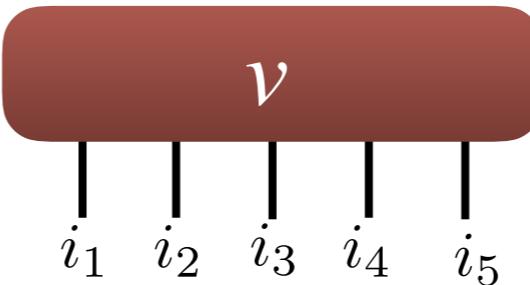


Contraction of a network = Calculation of a lot of multiplications
(縮約)

Graph for a tensor network decomposition

- Vector

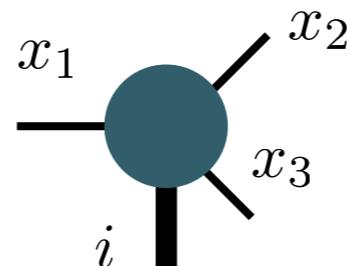
$$v_{i_1, i_2, i_3, i_4, i_5}$$



*Vector looks like a tensor

- Tensor

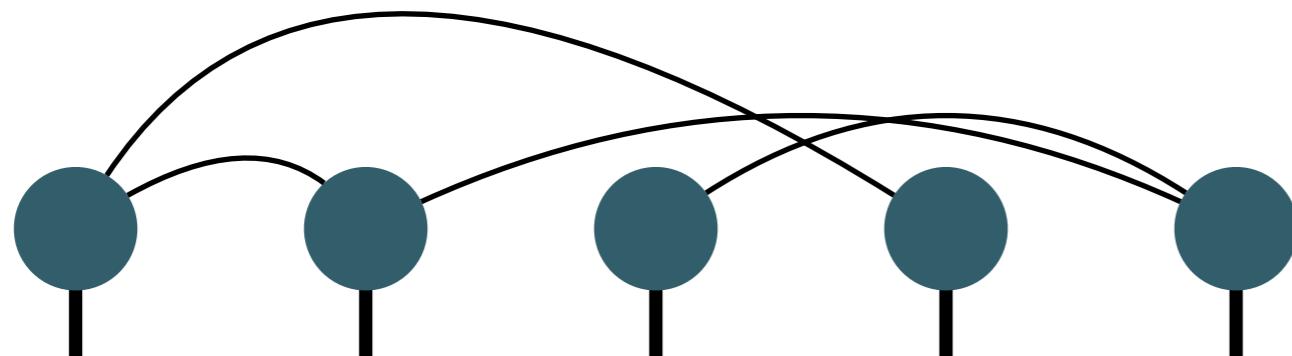
$$T[i]_{x_1, x_2, x_3}$$



*We treat i as an index
of the tensor.

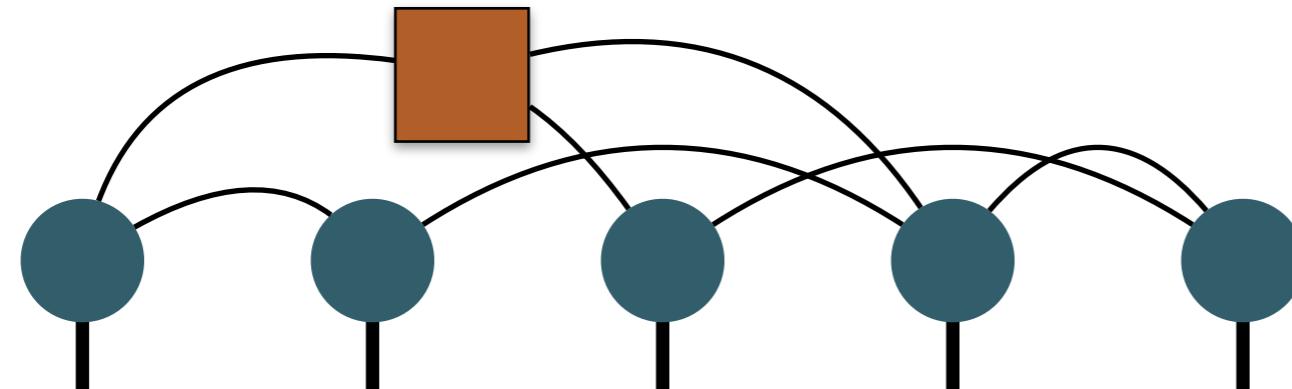
Tensor network decomposition

$$\vec{v} =$$



*We can consider tensors
independent of i .

$$\vec{w} =$$



Background:Singular value decomposition
(Will be skipped)

Singular value decomposition (SVD)

Singular value decomposition (特異値分解)

$$A : M \times N$$

$$A_{ij} \in \mathbb{C}$$

$$A = U \Sigma V^\dagger$$

Unitary **Unitary**

$$U : M \times M \qquad \qquad V : N \times N$$

$$\Sigma = \begin{pmatrix} \Sigma_{r \times r} & 0_{r \times (N-r)} \\ 0_{(M-r) \times r} & 0_{(M-r) \times N-r} \end{pmatrix}$$

$$\Sigma_{r \times r} = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \end{pmatrix}$$

Diagonal matrix with
non-negative real elements

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

Singular values

Properties of SVD

1. Any matrices can be decomposed as SVD: $A = U\Sigma V^\dagger$
 - Singular values are related to the eigenvalues of $A^\dagger A$ and AA^\dagger as
$$\sigma_i = \sqrt{\lambda_i}$$
 - V and U are eigenvectors of $A^\dagger A$ and AA^\dagger , respectively.

2. # of positive singular values **is identical with the rank.**

$$A : M \times N \rightarrow A : \mathbb{C}^N \rightarrow \mathbb{C}^M$$

$$\text{rank}(A) \equiv \dim(\text{img}(A))$$

$$\rightarrow \text{img}(A) = \text{Span}\{A\vec{v}_1, A\vec{v}_2, \dots, A\vec{v}_r\}$$

$$\rightarrow \dim(\text{img}(A)) = r = \# \text{ of positive singular values}$$

Low rank approximation

Low rank approximation (低ランク近似)

Find an approximate matrix

$$A \simeq \tilde{A}$$

with lower rank:

$$\text{rank}(A) > \text{rank}(\tilde{A})$$

→ Through the low rank approximation,
we can reduce amount of the data.

An example of information compressions.

Notice! In order to quantify accuracy of the approximation,
we need a measure of distance between matrices.

Low rank approximation by SVD

Consider a matrix obtained by **neglecting smaller singular values**

$$A = \bar{U} \Sigma_{r \times r} \bar{V}^\dagger \quad \rightarrow \quad \tilde{A} = \tilde{U} \Sigma_{k \times k} \tilde{V}^\dagger \quad (k < r)$$

$$\begin{aligned}\Sigma_{r \times r} &= \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r) \\ \bar{U} &= (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_r) \\ \bar{V} &= (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r)\end{aligned}$$

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$$

$$\text{rank}(A) = r$$

$$\begin{aligned}\Sigma_{k \times k} &= \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_k) \\ \tilde{U} &= (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_k) \\ \tilde{V} &= (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k)\end{aligned}$$

Keep **the largest k singular values**
(and corresponding singular vectors).

$$\text{rank}(\tilde{A}) = k < r$$

This approximation is one of the low rank approximation.

Norm of matrices $\|A\|$

There are two popular norms:

- (1) **Frobenius norm** (フロベニウス ノルム)

$$\|A\|_F = \sqrt{\sum_{i,j} |A_{ij}|^2} = \sqrt{\text{Tr}(A^\dagger A)}$$

*Trace (対角和)

$$\text{Tr}(X) = \sum_i X_{ii}$$

- (2) **Operator norm** (作用素ノルム)

$$\begin{aligned}\|A\|_O &= \inf\{c \geq 0; \|A\vec{x}\| \leq c\|\vec{x}\|\} \\ &= \sigma_1(A)\end{aligned}$$

*inf =infimum (下限)

*We define the norm for a vector as

$$\|\vec{x}\| = \sqrt{\sum_i |x_i|^2}$$

By using these norms, we define the distance between matrices:

$$\|A - \tilde{A}\|$$

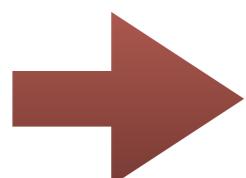
Accuracy of low rank approximation by SVD

Theorem

For $A : M \times N$

$$\min\{\|A - B\|_F : \text{rank}(B) = k\} = \sqrt{\sum_{i=k+1}^{\min(N,M)} \sigma_i^2}$$

$$\min\{\|A - B\|_O : \text{rank}(B) = k\} = \sigma_{k+1}$$



Because the k-rank approximation by SVD gives

$$\|A - \tilde{A}\|_F = \sqrt{\sum_{i=k+1}^{\min(N,M)} \sigma_i^2}, \quad \|A - \tilde{A}\|_O = \sigma_{k+1}$$

it is an "optimal" approximation with rank k .

Entanglement (エンタングルメント)

S=1/2 quantum spin system

Example: Wave function of N spin system



● takes two states $|0\rangle, |1\rangle$
 $(|\uparrow\rangle, |\downarrow\rangle)$

$$\begin{aligned} |\Psi\rangle &= \sum_{\{i_1, i_2, \dots, i_N\}} \Psi_{i_1 i_2 \dots i_N} |i_1\rangle \otimes |i_2\rangle \otimes \dots \otimes |i_N\rangle \\ &= \sum_{\{i_1, i_2, \dots, i_N\}} \Psi_{i_1 i_2 \dots i_N} |i_1 i_2 \dots i_N\rangle \end{aligned}$$

Coefficients = vector: $\vec{\Psi} \in \mathbb{C}^{2^N}$

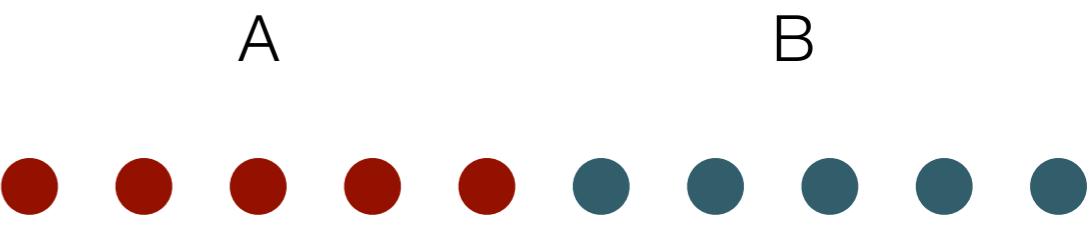
* Inner product: $\langle \Phi | \Psi \rangle = \vec{\Phi}^* \cdot \vec{\Psi}$

Schmidt decomposition for wave function

Wave function: $|\Psi\rangle = \sum_{\{i_1, i_2, \dots, i_N\}} \Psi_{i_1 i_2 \dots i_N} |i_1 i_2 \dots i_N\rangle$

Schmidt decomposition

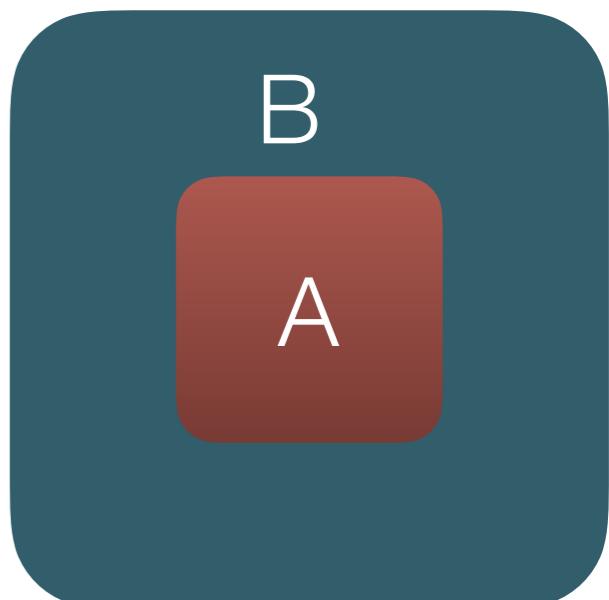
Divide system into two parts, A and B:



→ General wave function can be represented by a superposition of orthonormal basis set.

$$|\Psi\rangle = \sum_{i,j} M_{i,j} |A_i\rangle \otimes |B_j\rangle = \sum_i \lambda_i |\alpha_i\rangle \otimes |\beta_i\rangle$$

$$M_{i,j} \equiv \underbrace{\Psi_{(i_1, \dots), (\dots, i_N)}}_{\text{A} \quad \text{B}}$$



Orthonormal basis: $\langle \alpha_i | \alpha_j \rangle = \langle \beta_i | \beta_j \rangle = \delta_{i,j}$

Schmidt coefficient: $\lambda_i \geq 0$

Relation between SVD and Schmidt decomposition

Singular value decomposition (SVD):

For a $K \times L$ matrix M ,

$$M_{i,j} = \sum_m U_{i,m} \lambda_m V_{m,j}^\dagger$$

Singular values: $\lambda_m \geq 0$

$$\sum_i U_{i,m} U_{m,j}^\dagger = \delta_{i,j}$$

Singular vectors:

$$\sum_i V_{i,m} V_{m,j}^\dagger = \delta_{i,j}$$

Relation to the Schmidt decomposition:

$$|\Psi\rangle = \sum_{i,j} M_{i,j} |A_i\rangle \otimes |B_j\rangle = \sum_m \lambda_m |\alpha_m\rangle \otimes |\beta_m\rangle$$

$$|\alpha_m\rangle = \sum_i U_{i,m} |A_i\rangle$$

$$|\beta_m\rangle = \sum_j V_{m,j}^\dagger |B_j\rangle$$



$$\langle \alpha_i | \alpha_j \rangle = \langle \beta_i | \beta_j \rangle = \delta_{i,j}$$

By using SVD, we can perform Schmidt decomposition

Partial trace and reduced density matrix

For $\vec{x} \in \mathbf{V}_1 \otimes \mathbf{V}_2$ $\dim \mathbf{V}_1 = n_1, \dim \mathbf{V}_2 = n_2$ $|\vec{x}| = 1$

Density matrix: $\rho \equiv \vec{x}\vec{x}^\dagger$ ($\rho = |\vec{x}\rangle\langle\vec{x}|$)

(密度行列)

Note: rank $\rho = 1$ ($\rho_{ij} = x_i x_j^$)

Orthonormal basis: $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_{n_1}\} \in \mathbf{V}_1$ $\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_{n_2}\} \in \mathbf{V}_2$

→ Basis for \vec{x} : $\vec{g}_{i_1, i_2} = \vec{e}_{i_1} \otimes \vec{f}_{i_2}$

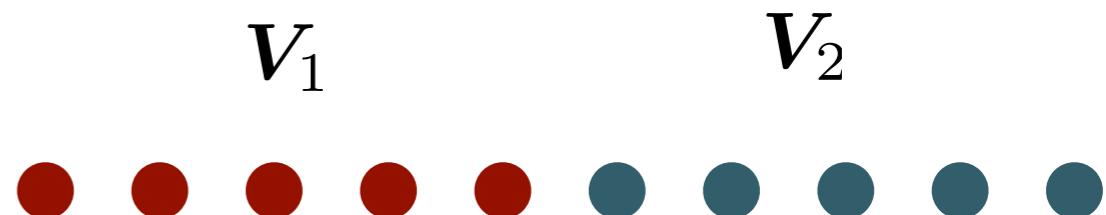
Index: $i = (i_1, i_2)$

Reduced Density matrix:

(縮約密度行列)

$\rho_{\mathbf{V}_1} \equiv \text{Tr}_{\mathbf{V}_2} \rho$: a positive-semidefinite square matrix in \mathbf{V}_1

$$(\rho_{\mathbf{V}_1})_{i_1, j_1} = \sum_{i_2} \rho_{(i_1, \underline{i_2}), (j_1, \underline{i_2})}$$

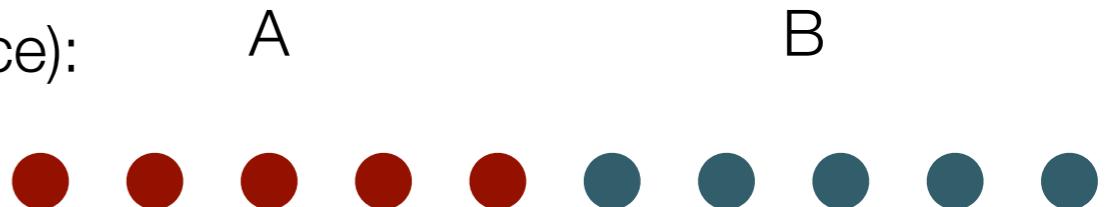


Entanglement entropy

Entanglement entropy:

Reduced density matrix of a sub system (sub space):

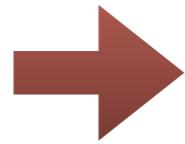
$$\rho_A = \text{Tr}_B |\Psi\rangle\langle\Psi|$$



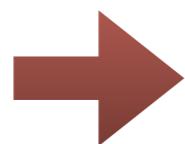
Entanglement entropy = von Neumann entropy of ρ_A

$$S = -\text{Tr}(\rho_A \log \rho_A)$$

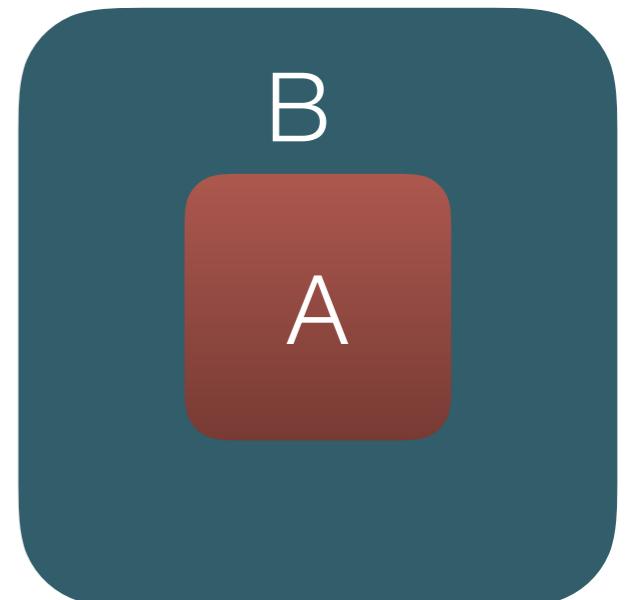
Schmidt decomposition $|\Psi\rangle = \sum_i \lambda_i |\alpha_i\rangle \otimes |\beta_i\rangle$



$$\rho_A = \sum_i \lambda_i^2 |\alpha_i\rangle\langle\alpha_i|$$



$$S = -\sum_i \lambda_i^2 \log \lambda_i^2$$

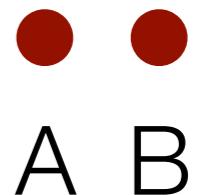


Entanglement entropy is calculated through
the spectrum of Schmidt coefficients

Intuition for EE: two S=1/2 spins

1. $|\Psi\rangle = |\uparrow\rangle \otimes |\downarrow\rangle$

A product state $\rightarrow \lambda = 1, S = 0$



2. $|\Psi\rangle = \frac{1}{2} (|\uparrow\rangle - |\downarrow\rangle) \otimes (|\uparrow\rangle - |\downarrow\rangle)$

Product state : S=0

Another product state $\rightarrow \lambda = 1, S = 0$

3. $|\Psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle)$

Spin singlet

$$\rightarrow \lambda_1 = \lambda_2 = \frac{1}{\sqrt{2}}, S = \log 2$$

Maximally entangled State

4. $|\Psi\rangle = (x|\uparrow\rangle \otimes |\downarrow\rangle + \sqrt{1-x^2}|\downarrow\rangle \otimes |\uparrow\rangle)$

Complicated state

$$\rightarrow \lambda_1 = |x|, \lambda_2 = \sqrt{1-x^2}$$
$$S = x^2 \log x^2 + \sqrt{1-x^2} \log(1-x^2)$$

Large entanglement entropy ~ Large correlation between two parts

Area law of the entanglement entropy **in physics**

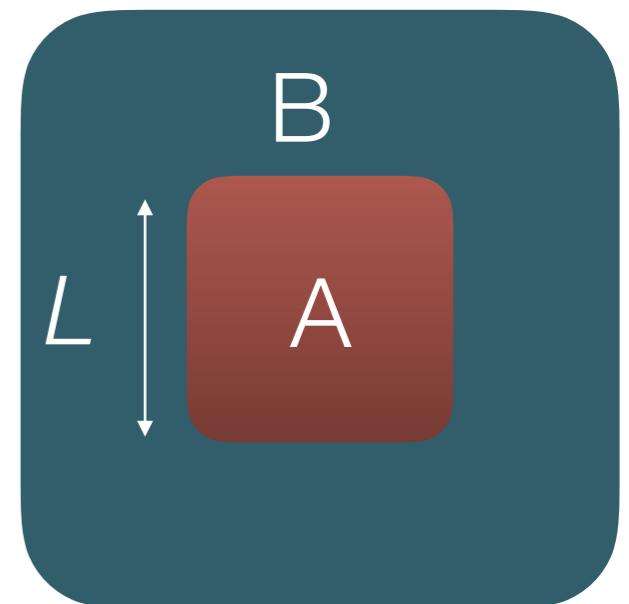
General wave functions:

EE is proportional to its **volume (# of spins)**.

$$S = -\text{Tr}(\rho_A \log \rho_A) \propto L^d \quad (\text{c.f. random vector})$$

Ground state wave functions:

For a lot of ground states, EE is proportional to its area.



J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys, 277, **82** (2010)

$$S = -\text{Tr}(\rho_A \log \rho_A) \propto L^{d-1}$$

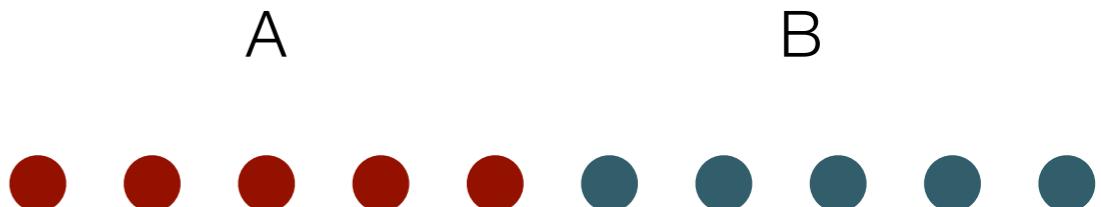
In the case of **one-dimensional system**:

Gapped ground state for **local Hamiltonian**

M.B. Hastings, J. Stat. Mech.: Theory Exp. P08024 (2007)

$$S = O(1)$$

Ground state are in a small part
of the huge Hilbert space



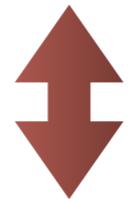
Matrix product state (行列積狀態)

Data compression of wave functions (vectors)

General wave function:

$$|\Psi\rangle = \sum_{\{i_1, i_2, \dots, i_N\}} \Psi_{i_1 i_2 \dots i_N} |i_1 i_2 \dots i_N\rangle$$

Coefficient vector can represent any points in the Hilbert space.

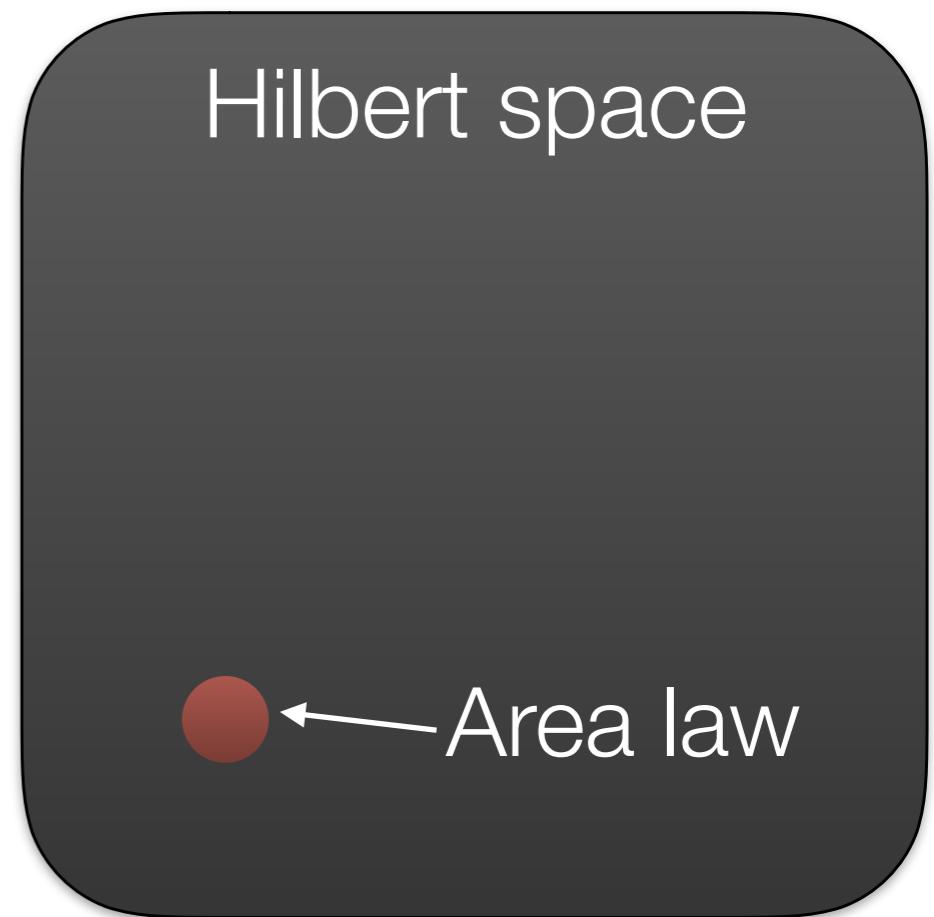


Ground states satisfy the area law.

→ In order to represent the ground state,
we do not need all of a^N elements.

→ Data compression by tensor decomposition:

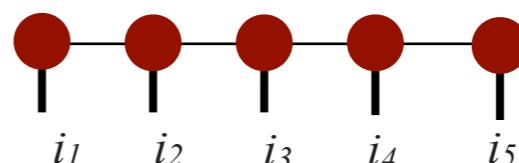
Tensor network states



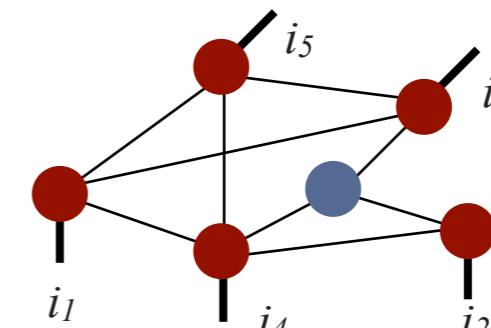
Tensor network state

G.S. wave function: $|\Psi\rangle = \sum_{\{i_1, i_2, \dots, i_N\}} \Psi_{i_1 i_2 \dots i_N} |i_1 i_2 \dots i_N\rangle$

Vector (or N-rank tensor): $\Psi_{i_1 i_2 \dots i_N}$ =  # of Elements = a^N
 "Tensor network" decomposition

- * Matrix Product State (MPS) $A_1[i_1] A_2[i_2] \cdots A_N[i_N] =$ 

- * General network $\text{Tr } X_1[i_1] X_2[i_2] X_3[i_3] X_4[i_4] X_5[i_5] Y$
 X, Y : Tensors
 Tr : Tensor network contraction



By choosing a "good" network, we can express G.S. wave function efficiently.

ex. MPS: # of elements = $2ND^2$

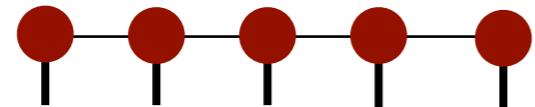
D: dimension of the matrix A

Exponential \rightarrow Linear

*If D does not depend on N...

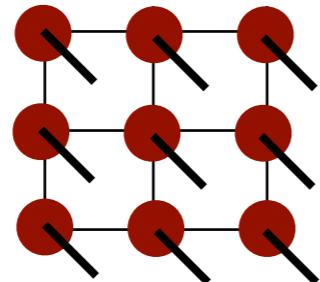
Examples of TNS

MPS:



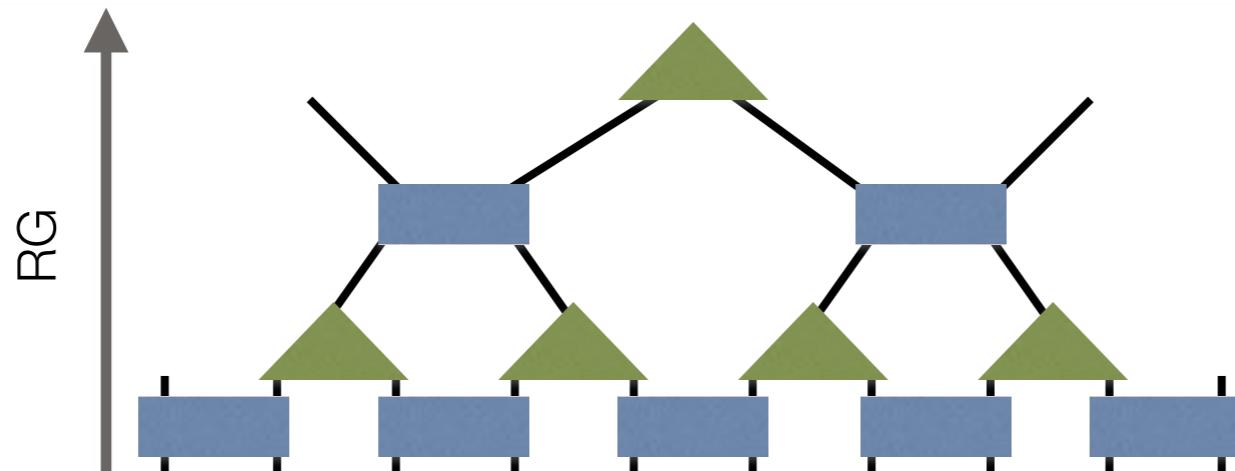
Good for 1-d gapped systems

PEPS, TPS:



For higher dimensional systems
Extension of MPS

MERA:



Scale invariant systems

Good reviews:

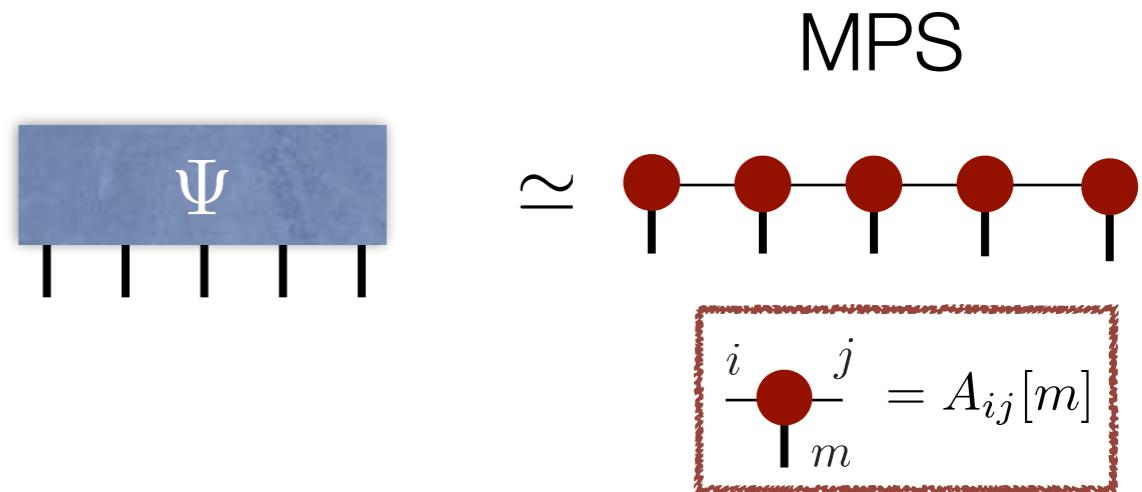
Matrix product state (MPS)

(U. Schollwöck, Annals. of Physics **326**, 96 (2011))
(R. Orús, Annals. of Physics **349**, 117 (2014))

$$|\Psi\rangle = \sum_{\{i_1, i_2, \dots, i_N\}} \Psi_{i_1 i_2 \dots i_N} |i_1 i_2 \dots i_N\rangle$$

$$\Psi_{i_1 i_2 \dots i_N} \simeq A_1[i_1] A_2[i_2] \cdots A_N[i_N]$$

$A[i]$: Matrix for state i



Note:

- MPS is called as "tensor train decomposition" in applied mathematics
(I. V. Oseledets, SIAM J. Sci. Comput. **33**, 2295 (2011))
- A product state is represented by MPS with 1×1 "Matrix" (scalar)

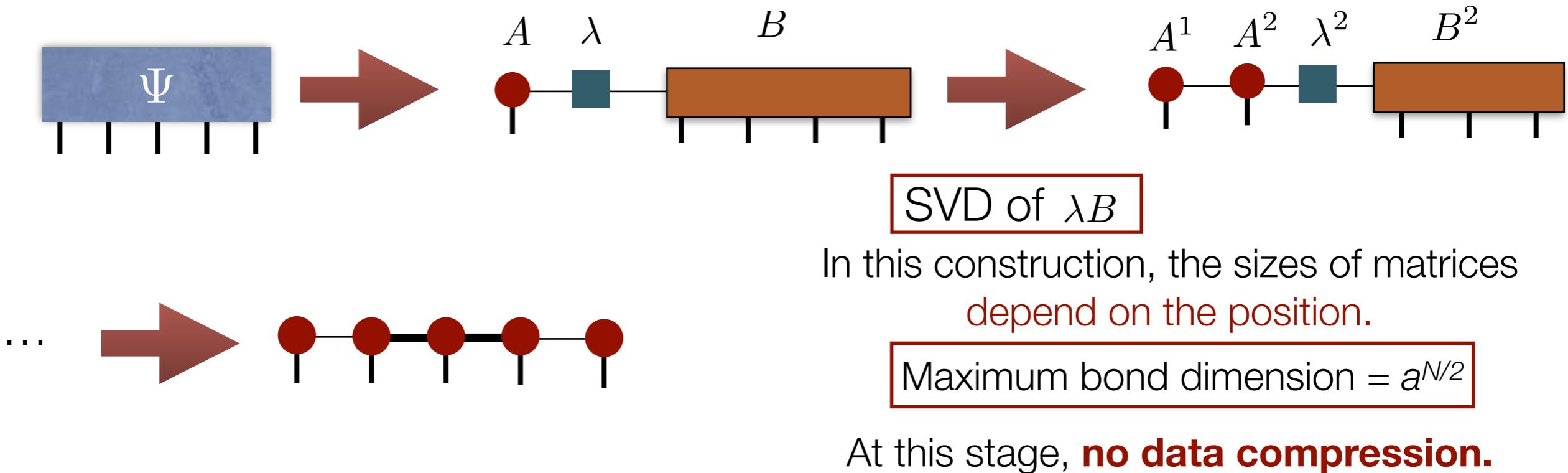
$$|\Psi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \otimes \cdots$$

$$\Psi_{i_1 i_2 \dots i_N} = \phi_1[i_1] \phi_2[i_2] \cdots \phi_N[i_N]$$

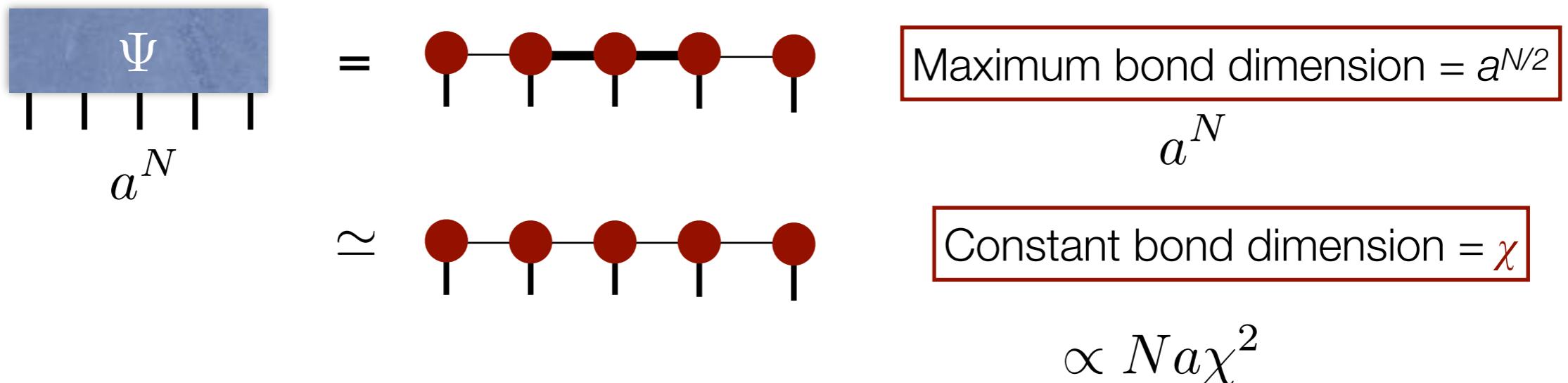
$$\phi_n[i] \equiv \langle i | \phi_i \rangle$$

Matrix product state without approximation

General wave function (or vector) can be represented by MPS **exactly** through successive Schmidt decompositions



Matrix product state: Low rank approximation

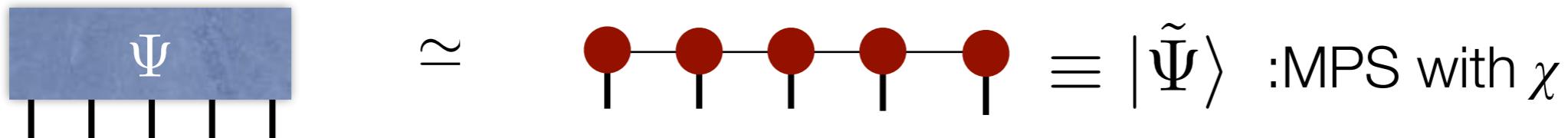


If the entanglement entropy of the system is **O(1)** (independent of N), matrix size " χ " can be small for accurate approximation.

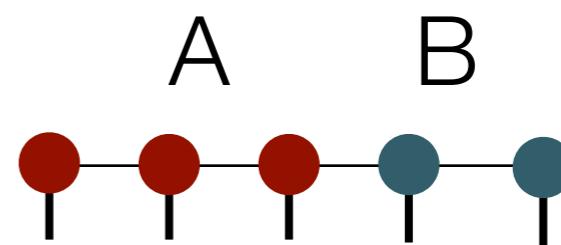
→ MPS is good for gapped 1d systems.

On the other hand, if the EE increases as increase N , " χ " must be increased to keep the same accuracy.

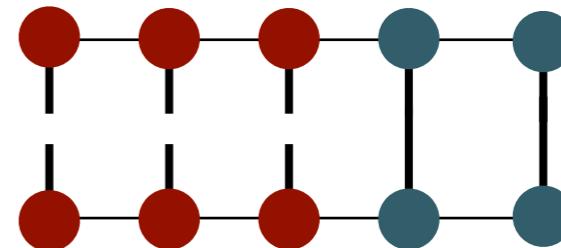
Upper bound of Entanglement entropy



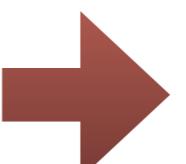
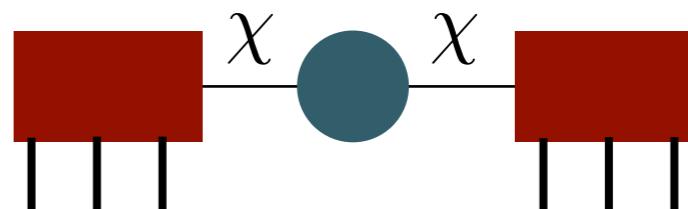
Reduced density matrix of region A:



$$\rho_A = \text{Tr}_B |\tilde{\Psi}\rangle\langle\tilde{\Psi}| =$$



★ Structure of ρ_A :

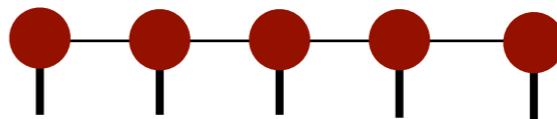


$$\text{rank } \rho_A \leq \chi$$

$$S_A = -\text{Tr } \rho_A \log \rho_A \leq \log \chi$$

Required bond dimension in MPS representation

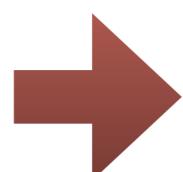
$$S_A = -\text{Tr } \rho_A \log \rho_A \leq \log \chi$$



The upper bound is independent of the "length".

length of MPS \Leftrightarrow size of the problem

$$N \quad a^N$$



| EE of the original vector | Required bond dimension in MPS representation |
|------------------------------|--|
| $S_A = O(1)$ | $\chi = O(1)$ |
| $S_A = O(\log N)$ | $\chi = O(N^\alpha)$ |
| $S_A = O(N^\alpha)$ | $\chi = O(c^{N^\alpha})$ |

Generalization to tensor network states

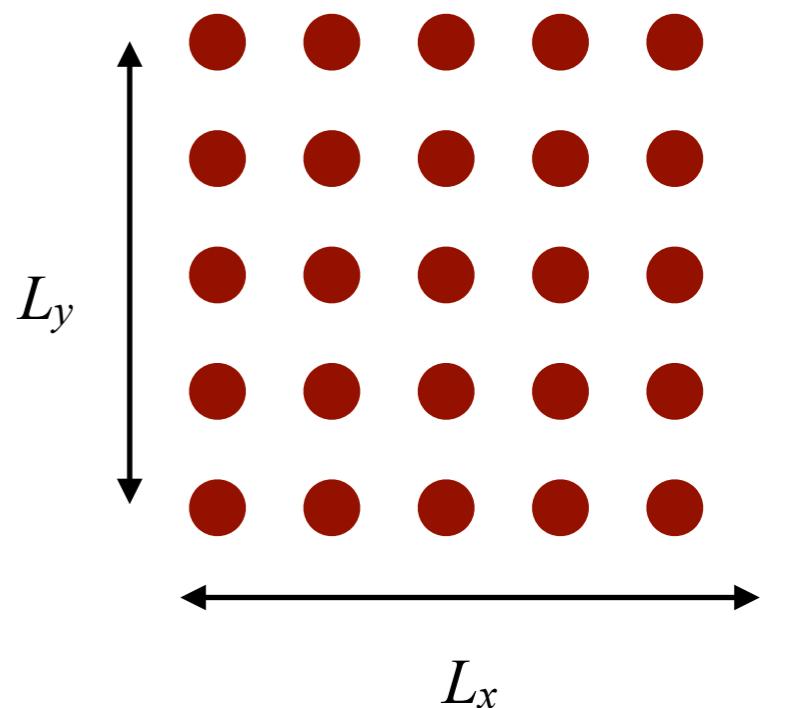
Higher dimensional system

Transverse field Ising model on **square lattice**:

$$\mathcal{H} = - \sum_{\langle i,j \rangle} S_i^z S_j^z - h \sum_{i=1}^N S_i^x$$

$\sum_{\langle i,j \rangle}$: Summation over the nearest neighbor pair

Two-dimensional array



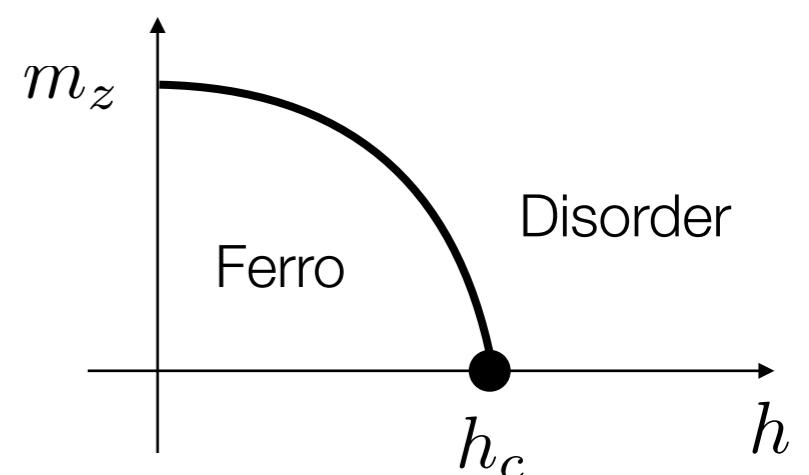
Area law

Even in ferro and disordered phases,
the entanglement entropy depends on size N .

$$S_A \sim \sqrt{N} = L$$

$$N = L_x \times L_y$$

Phase diagram

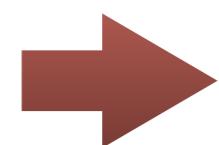


MPS for two-dimensional system

When we apply MPS representation for a square lattice system:

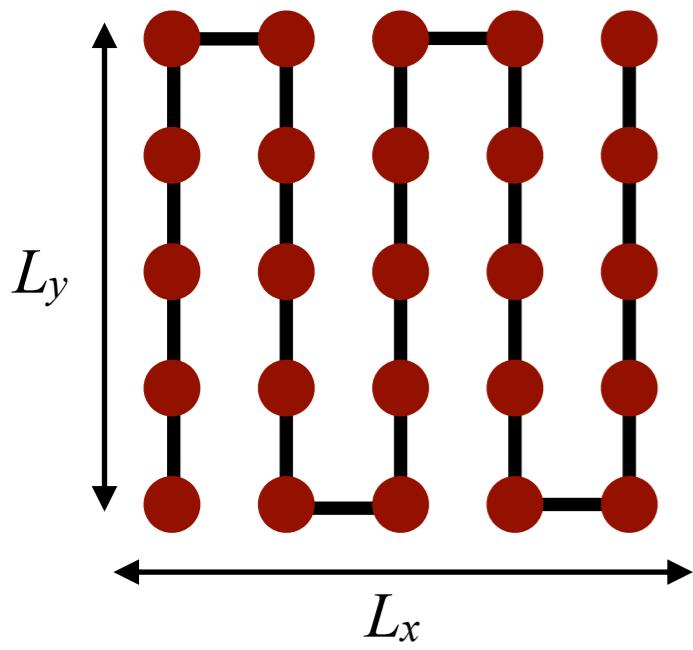
Setting **(1)** $S_A \leq L_x \log \chi$:Satisfying area law?

Setting **(2)** $S_{A'} \leq \log \chi$:Break down of the area law!



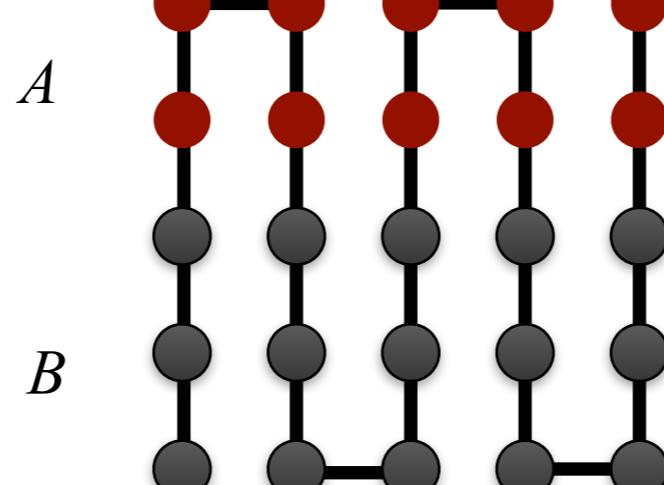
MPS cannot cover the area law of the entanglement entropy in higher ($d = 2, 3, \dots$) dimensions.

Possible MPS
(Snake form)

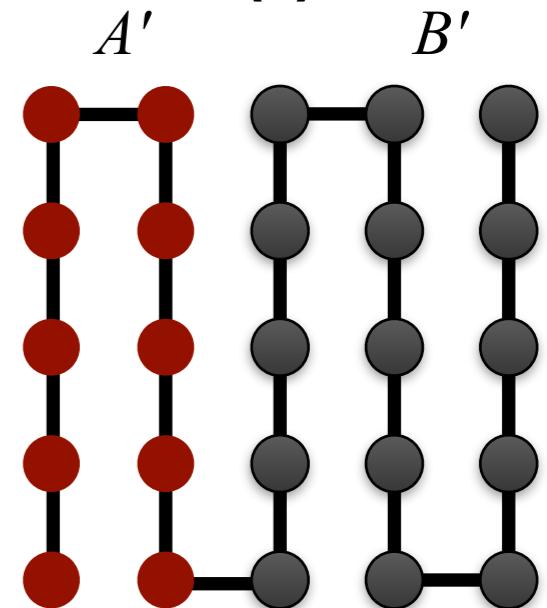


Two settings of **system** and **environment**

(1)



(2)



Entanglement entropy in higher dimensions

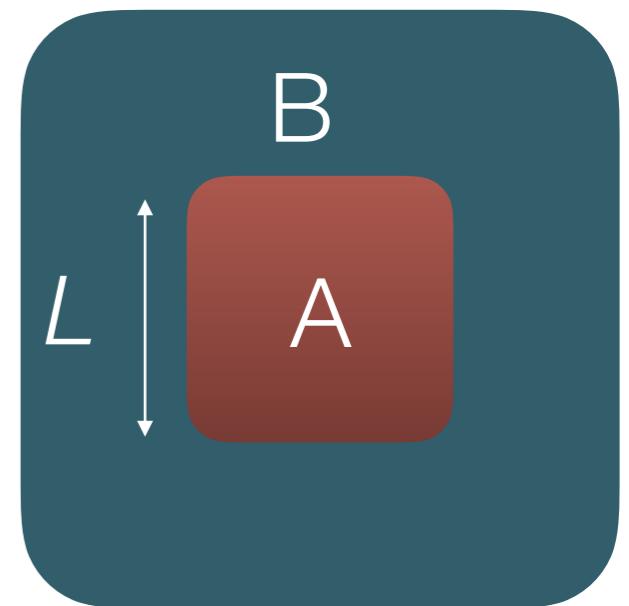
Ground state wave functions:

For a lot of ground states, EE is proportional to its area.

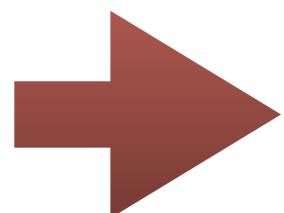
J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys, 277, **82** (2010)

Area law:

$$S = -\text{Tr}(\rho_A \log \rho_A) \propto L^{d-1}$$



In d=1, MPS satisfies the area law.



Q. What is a simple generalization of MPS to $d > 1$?

A. It is Tensor Product State (TPS)!

Tensor Product State (TPS)

TPS (Tensor Product State) (AKLT, T. Nishino, K. Okunishi, ...)

PEPS (Projected Entangled-Pair State)

(F. Verstraete and J. Cirac, arXiv:cond-mat/0407066)

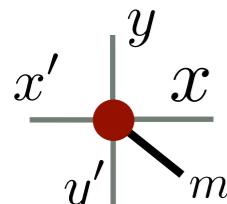
d-dimensional tensor network representation
for the wave function of a d-dimensional quantum system

$$|\Psi\rangle = \sum_{\{m_i=1,2,\dots,m\}} \text{Tr} A_1[m_1] A_2[m_2] \cdots A_N[m_N] |m_1 m_2 \cdots m_N\rangle$$



Tr: tensor network “contraction”

$A_{x_i x'_i y_i y'_i}[m_i]$: Rank 4+1 tensor



$x, y, x', y' = 1, 2, \dots, D$

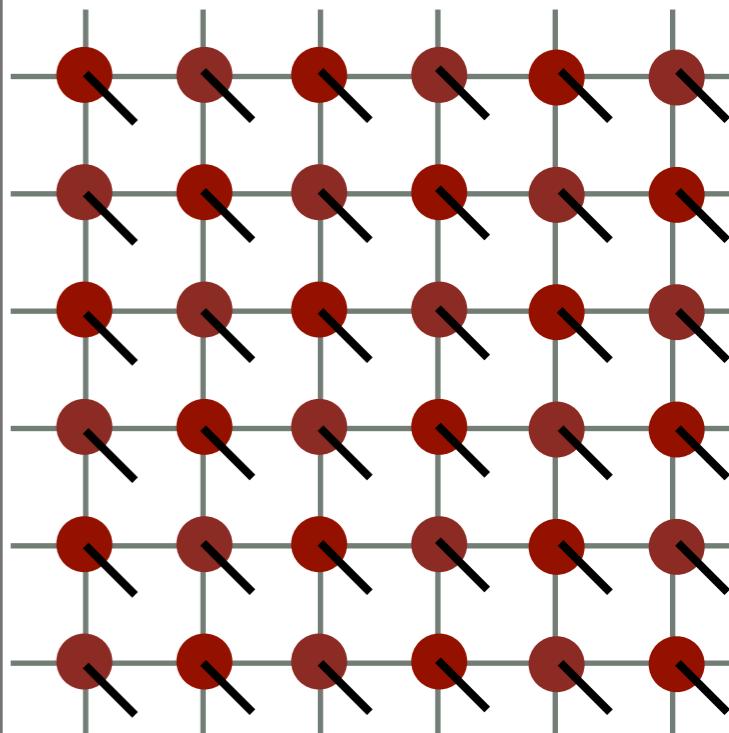
$m_i = 1, 2, \dots, m$

D = “bond dimension”

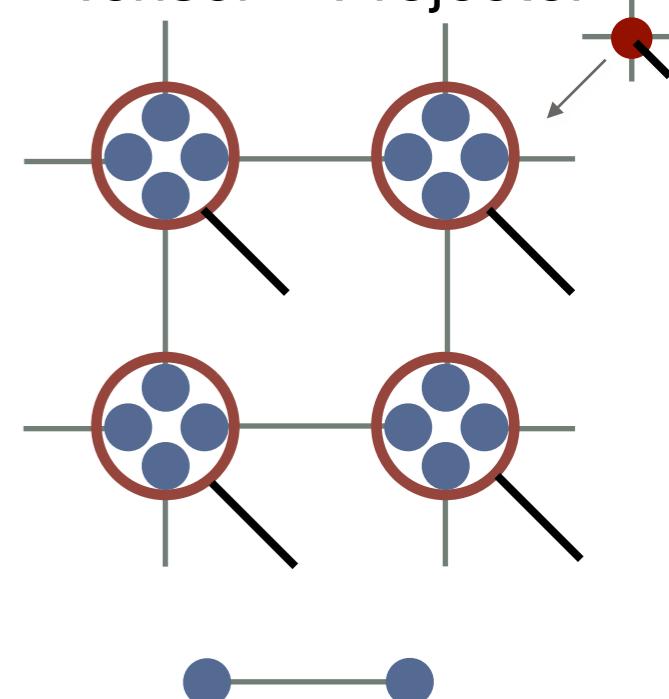
m = dimension of the local Hilbert space

*D can be larger than m. “Virtual state”

TPS on square lattice

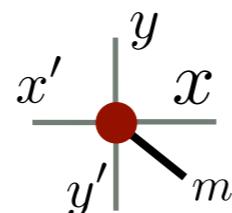
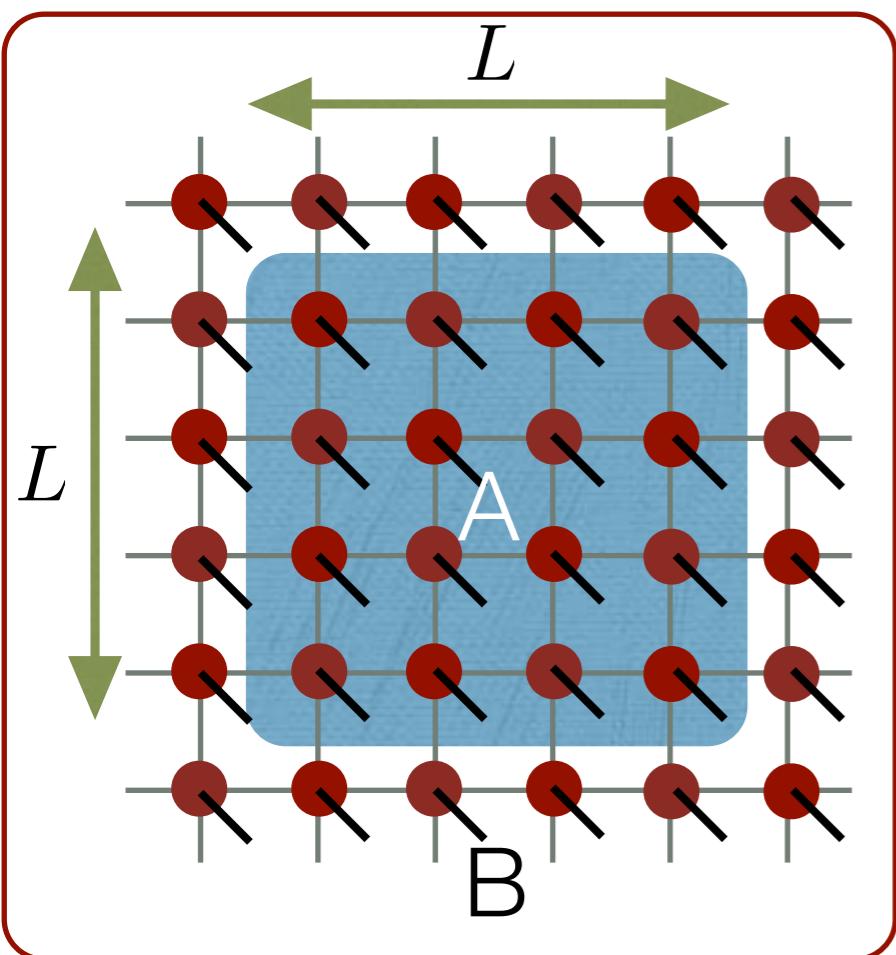


Tensor = Projector



Maximally entangled state
between D-state spins

Entanglement entropy of TPS (PEPS)



Bond dimension = D

of bonds connecting regions A and B

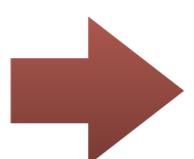
$$N_c(L) = 4L \quad (\text{square lattice})$$

$$N_c(L) = 2dL^{d-1} \quad (\text{d-dimensional hyper cubic lattice})$$

$$\text{rank } \rho_A \leq D^{N_c(L)} \sim D^{2dL^{d-1}}$$

$$S_A = -\text{Tr } \rho_A \log \rho_A \leq 2dL^{d-1} \log D$$

TPS can satisfy the area law even for $d > 1$.



We can efficiently approximate vectors in higher dimensional space by TPS.

* Similar to the MPS in 1d, TPS can approximate infinite system!

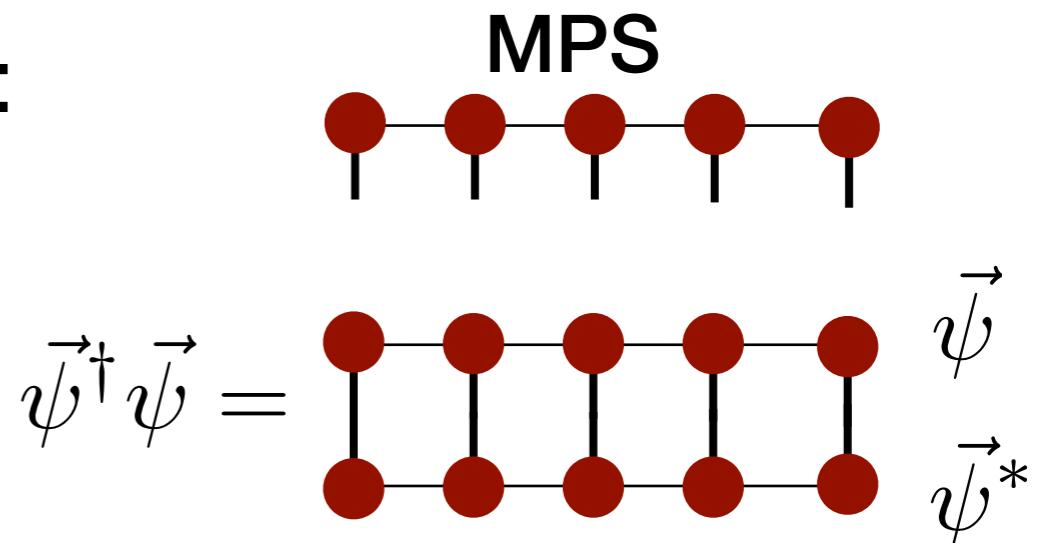
Difference between MPS and TPS

Cost of tensor network contraction:

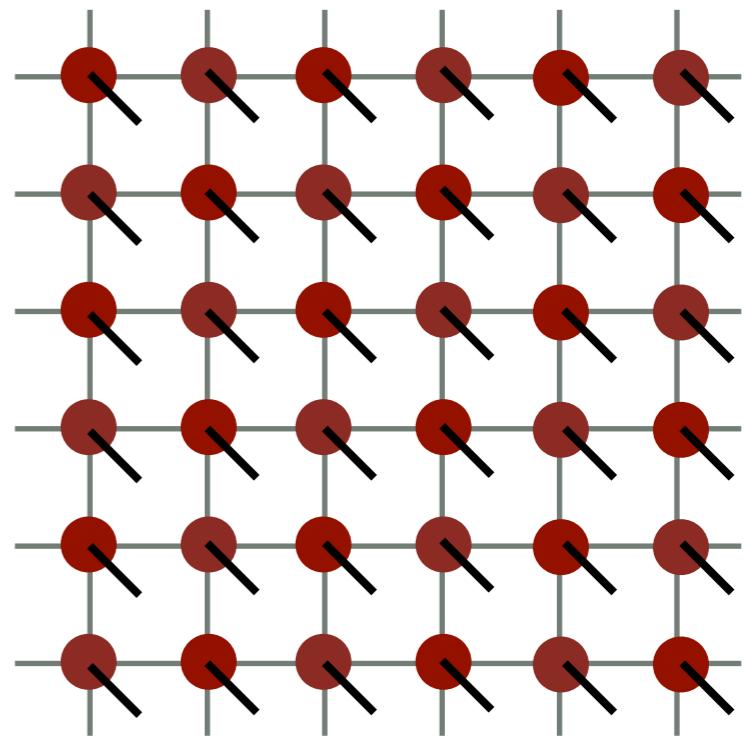
d-dimensional cubic lattice $N = L^d$

MPS: $O(N)$

TPS: $O(e^{L^{d-1}})$



TPS
(PEPS)



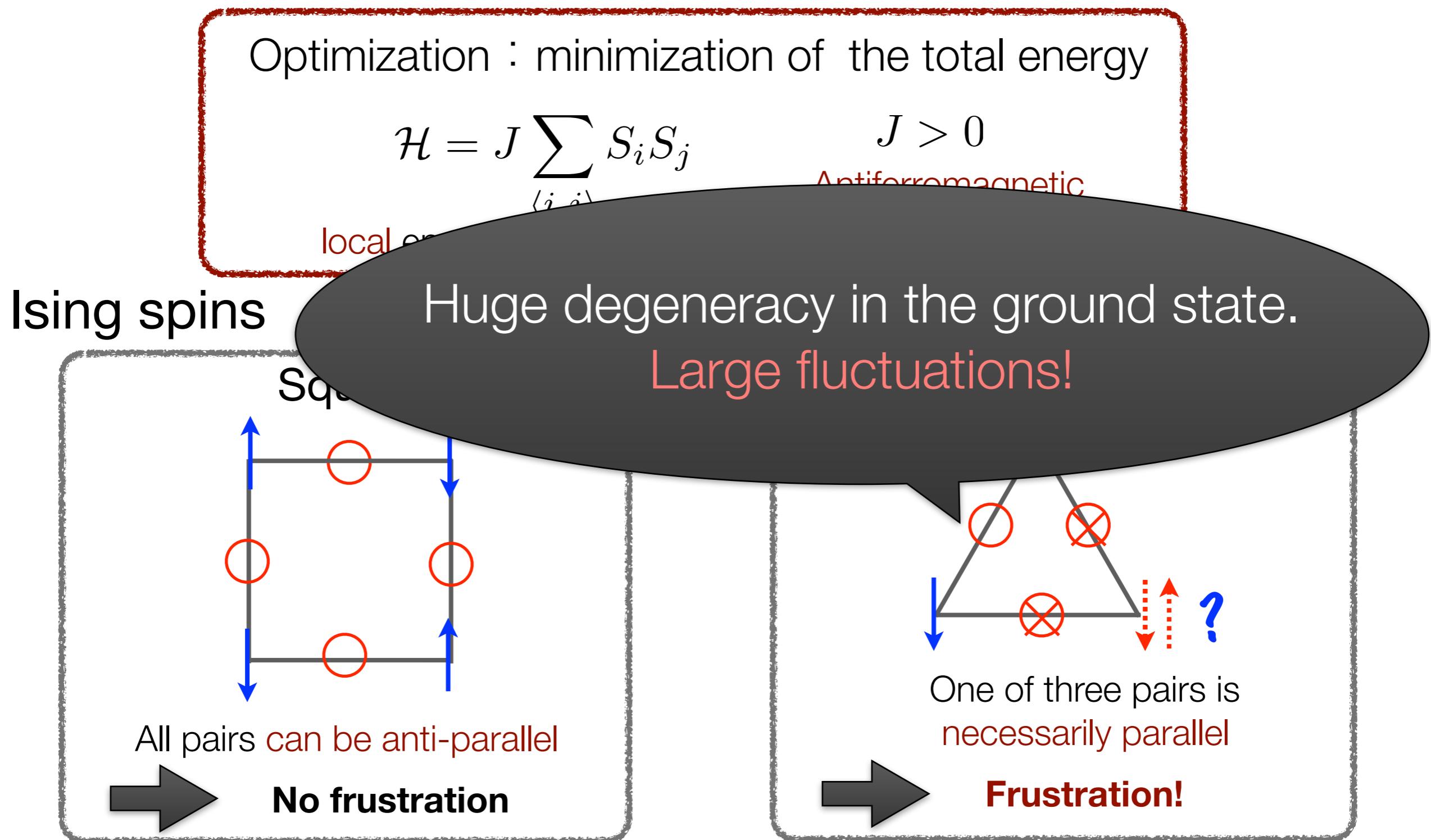
It is **impossible** to perform exact contraction even if we know local tensors in the case of TPS.

In the case of TPS,
usually we **approximately**
calculate the contraction.

Application of TPS to frustrated spin systems

Frustration in spin systems

Frustration : Competition among several optimization conditions

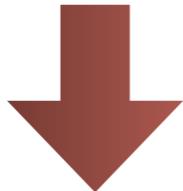


Numerical calculations for frustrated quantum spin systems

★ Problems in numerical calculation for frustrated spin systems

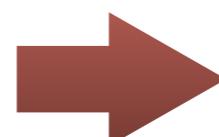
Very few method to approach the thermodynamic limit.

- Quantum Monte Carlo: Poor accuracy due to **sign problem**.
- Exact diagonalization: Limitation to **cluster size**.
- Variational Monte Carlo: **Biased** due to variational wavefunction.



Tensor network method

Use tensor network states as **variational wavefunctions**.



- Applicable even if for **frustrated spin system**.
- It can treat **infinite system** directly.

Variational calculation

Find the smallest eigenvalue and its eigenvector

$$\mathcal{H}\vec{v}_0 = E_0\vec{v}_0$$

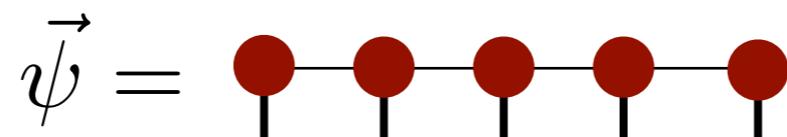


$$\min_{\vec{\psi} \in \mathbb{C}^M} \frac{\vec{\psi}^\dagger (\mathcal{H} \vec{\psi})}{\vec{\psi}^\dagger \vec{\psi}} \left(= \min_{|\psi\rangle} \frac{\langle \psi | \mathcal{H} | \psi \rangle}{\langle \psi | \psi \rangle} \right)$$

Variational calculation using TNS:

Cost function: $F = \frac{\vec{\psi}^\dagger (\mathcal{H} \vec{\psi})}{\vec{\psi}^\dagger \vec{\psi}}$

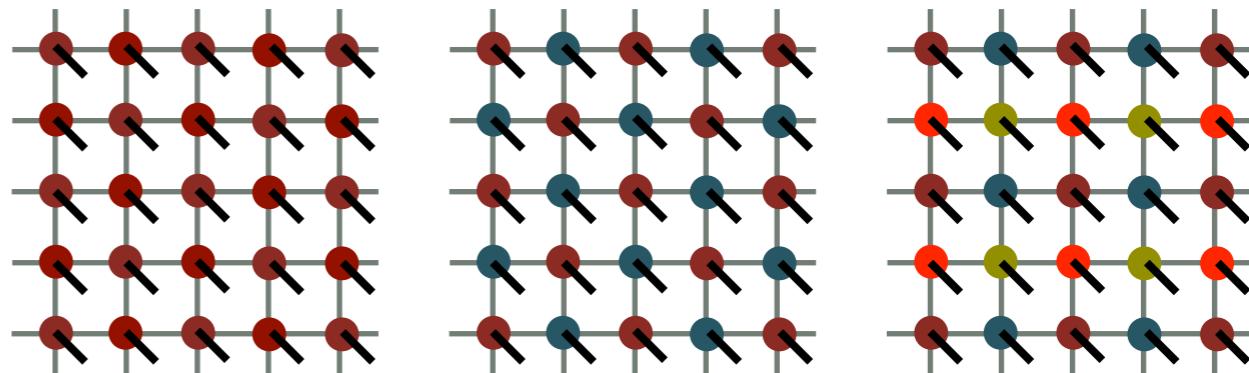
Find the TNS which minimizes F
by optimizing tensors in TNS.



TPS as variational wave functions

- Because TPS satisfies the **area law**, it can represent **a lot of ground states efficiently** when we take enough large bond dimension D.
 - If we assume translational invariance, we can calculate infinite system by finite bond dimension

Examples of translationally invariant systems



- Recently, optimization and contraction methods have been developed rapidly.
 - It can be applicable to novel states, such as spin liquid

Kitaev spin liquid

J. O. Iregui, P. Corboz, and M. Troyer, et al., PRB 90, 195102 (2014).

Typical calculation method

Necessary steps

1. Optimization of the wave functions (tensors)

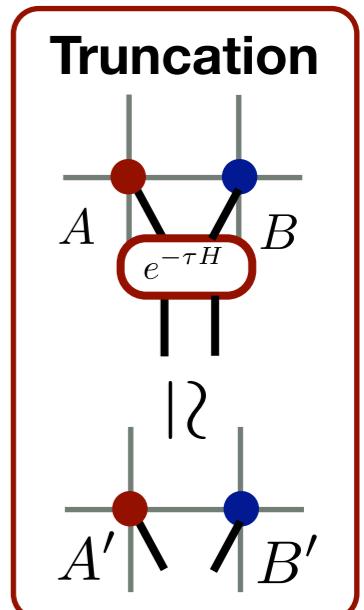
Imaginary time evolution

$$\lim_{M \rightarrow \infty} (e^{-\tau H})^M |\psi\rangle = \text{ground state}$$

Suzuki-Trotter decomposition

$$e^{-\tau H} \simeq e^{-\tau H_x} e^{-\tau H_y} e^{-\tau H'_x} e^{-\tau H'_y} + O(\tau^2)$$

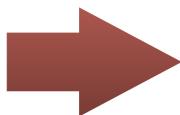
cf. variational optimization : P. Corboz, PRB **94**, 035133 (2016)



2. Calculation of physical quantities

We need the contraction of tensor network.

- It needs **exponentially large computational costs**
- Exact contraction is impossible



Approximate contraction methods

- Transfer matrix, **corner transfer matrix**
real space renormalization, ...

They reduce the cost to polynomials.

Still, we need $O(D^{10}) \sim O(D^{12})$ calculations.

 The contraction of TNS is the bottle neck.

Expectation value

$$\langle \hat{O} \rangle = \frac{\langle \psi | \hat{O} | \psi \rangle}{\langle \psi | \psi \rangle}$$

double layered tensor

Contraction of TNS

$$\langle \Psi | \Psi \rangle = \text{Tr } T_1 T_2 \cdots T_N$$

Main calculations in TPS method

- Tensors are transformed into matrices.
(including transpose operations.)
- Matrix-matrix multiplications: $C = AB \rightarrow \text{BLAS}$
 - The matrices are generally rectangular.
 - $A: D^4 \times D^2, B: D^2 \times D^4 \rightarrow C: D^4 \times D^4$
 - Typically, $D < 20$.
- Low rank approximations by SVD: $M = U \Lambda V^\dagger \rightarrow \text{LAPACK}$.
 - M is a $D^4 \times D^4$ matrix, and we pick up leading D^2 singular values.
 - In practice, we use partial SVD instead of full SVD.

Application1: Honeycomb lattice Kitaev Model

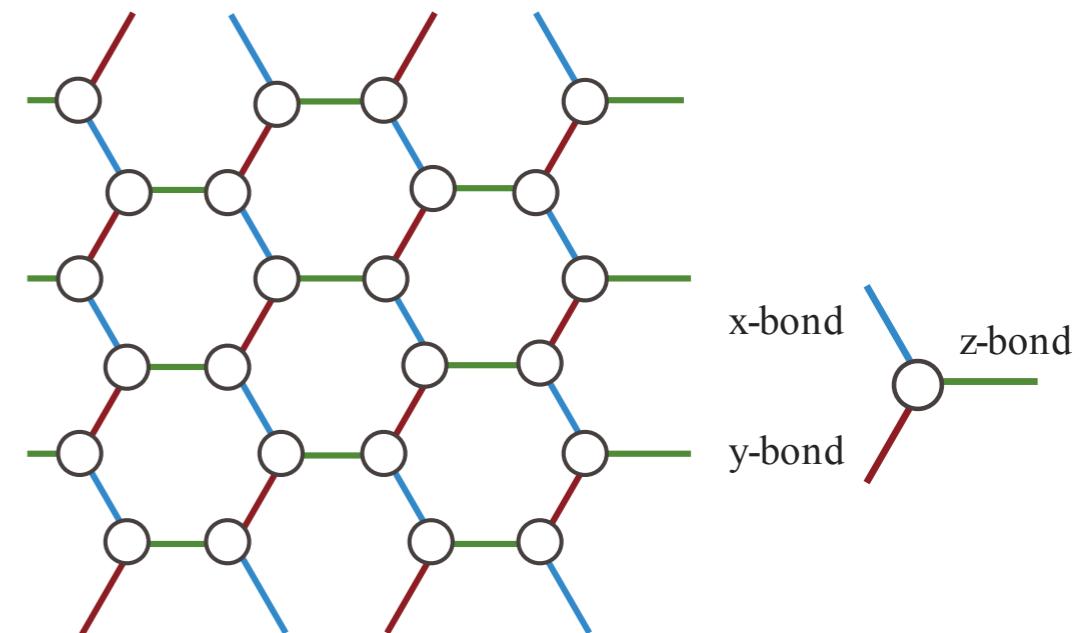
A. Kitaev, Annals of Physics 321, 2 (2006)

Kitaev model

$$\mathcal{H} = - \sum_{\gamma, \langle i,j \rangle_\gamma} J_\gamma S_i^\gamma S_j^\gamma$$

γ :bond direction

Depending on the bond direction, only specific spin components interact.



Exactly solvable by introducing Majorana fermion

Isotropic region (B) : gapless spin liquid

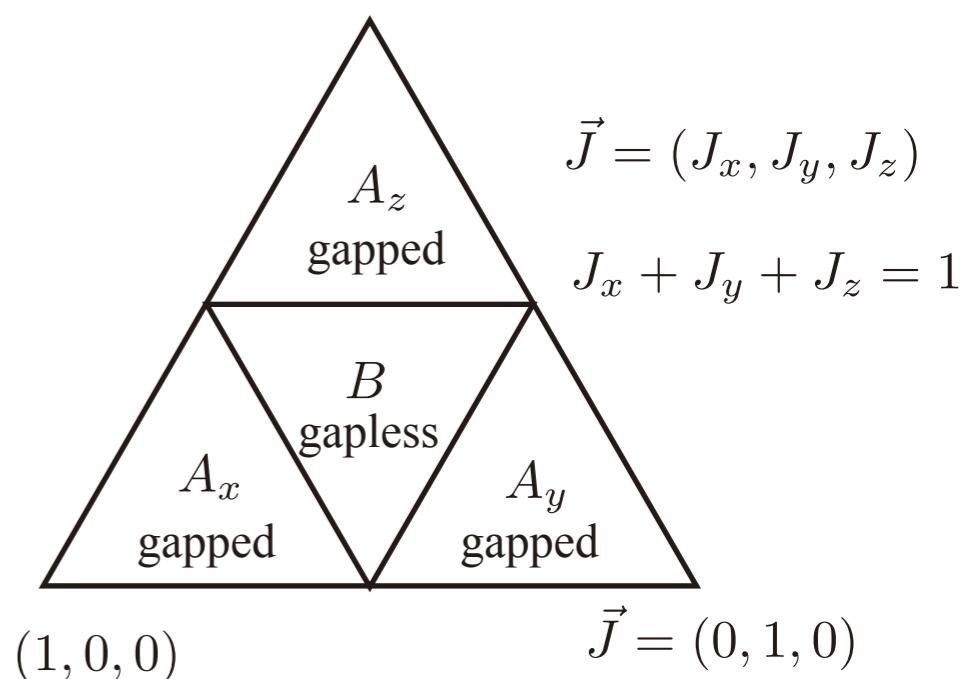
Anisotropic region (A) : gapped spin liquid

Cf. The anisotropic limit corresponds to the Toric code.

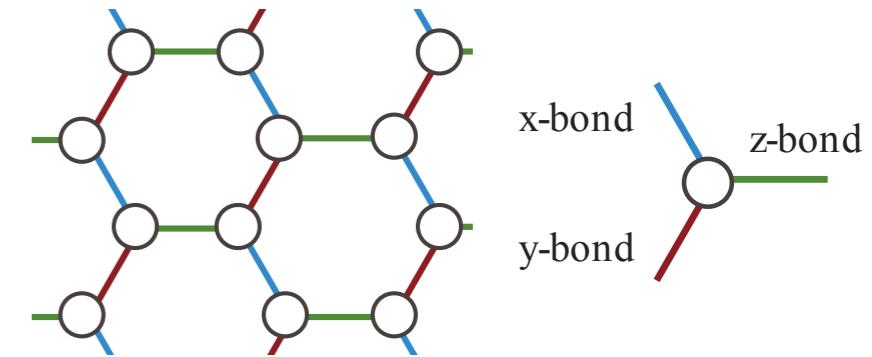
*Recently, researchers have realized that this type of models might appear in real materials.
Hot topic!

Phase diagram

$$\vec{J} = (0, 0, 1)$$



Application1 : Kitaev spin liquid



Honeycomb lattice Kitaev model

At $J_x = J_y = J_z$, the ground state is
a gapless spin liquid.

In the present (super)computers,
we can access around $D=10$ (maybe 16)
by using massively parallel code.

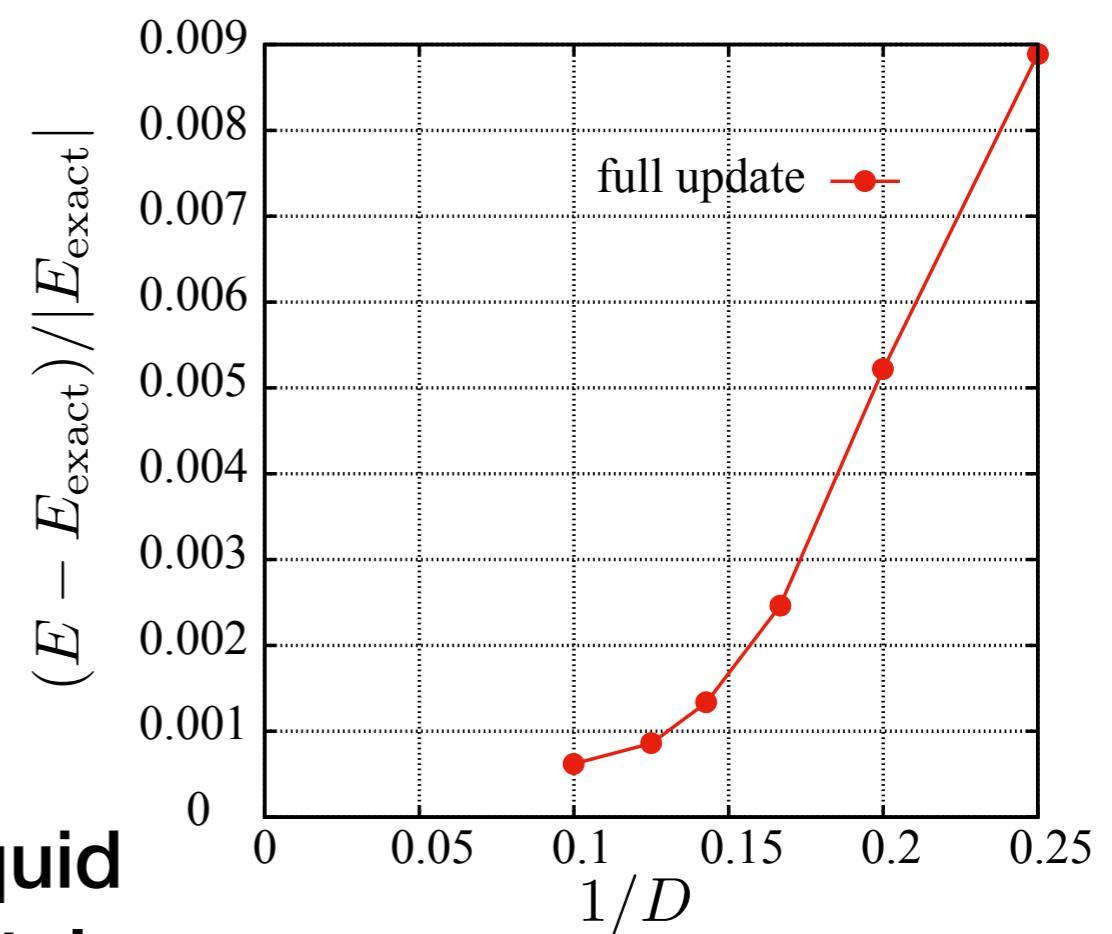
The error of the ground state
energy is **less than 10^{-3}**
for infinite system!

$$\mathcal{H} = - \sum_{\gamma, \langle i,j \rangle_\gamma} J_\gamma S_i^\gamma S_j^\gamma$$

$(\gamma = x, y, z)$

Energy error obtained by iTPS

(T. okubo et al, unpublished)

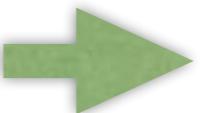


iTPS can represent Kitaev spin liquid
in the thermodynamic limit accurately.

A_2IrO_3

Iridium Oxides

Strong spin-orbit coupling



Effective "spin" moment:

$$J_{\text{eff}} = \frac{1}{2}$$

Na_2IrO_3, Li_2IrO_3

G.Jackeli, et al., PRL 102, 017205 (2009)
J. Chaloupka, et al., PRL 105, 027204 (2010)

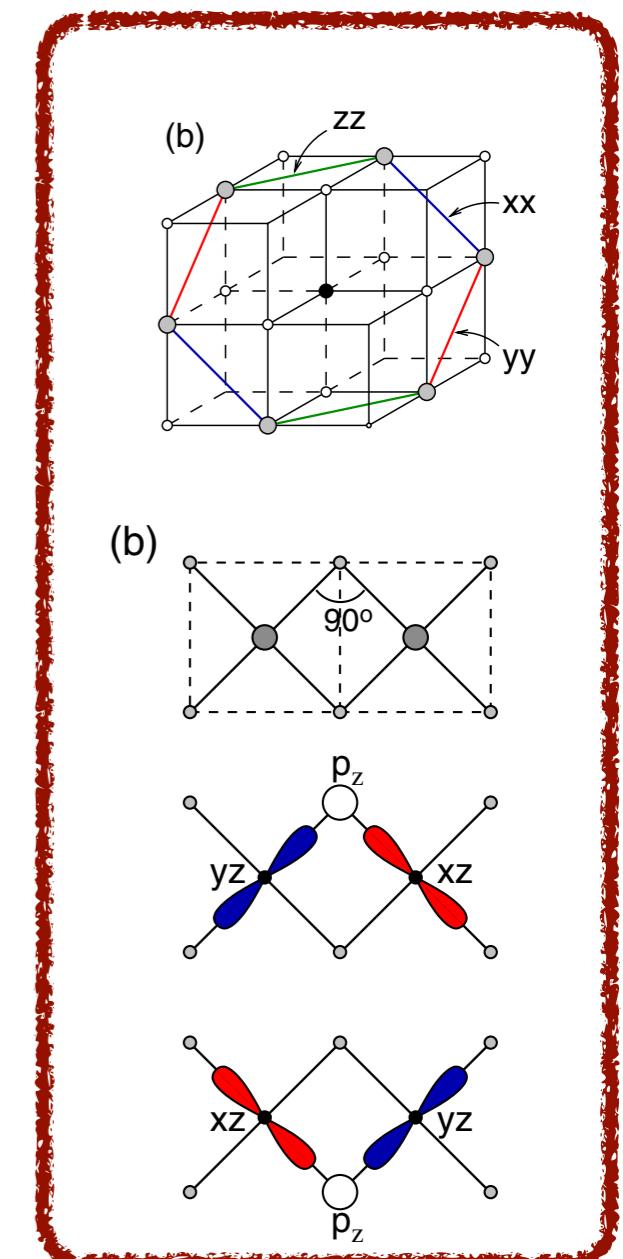
Ir ions form an honeycomb lattice.

Ir - Ir direct exchange: Heisenberg interaction

Ir - O - Ir exchange: Anisotropic Kitaev interaction

Depending on the bond direction, only specific spin component interact.

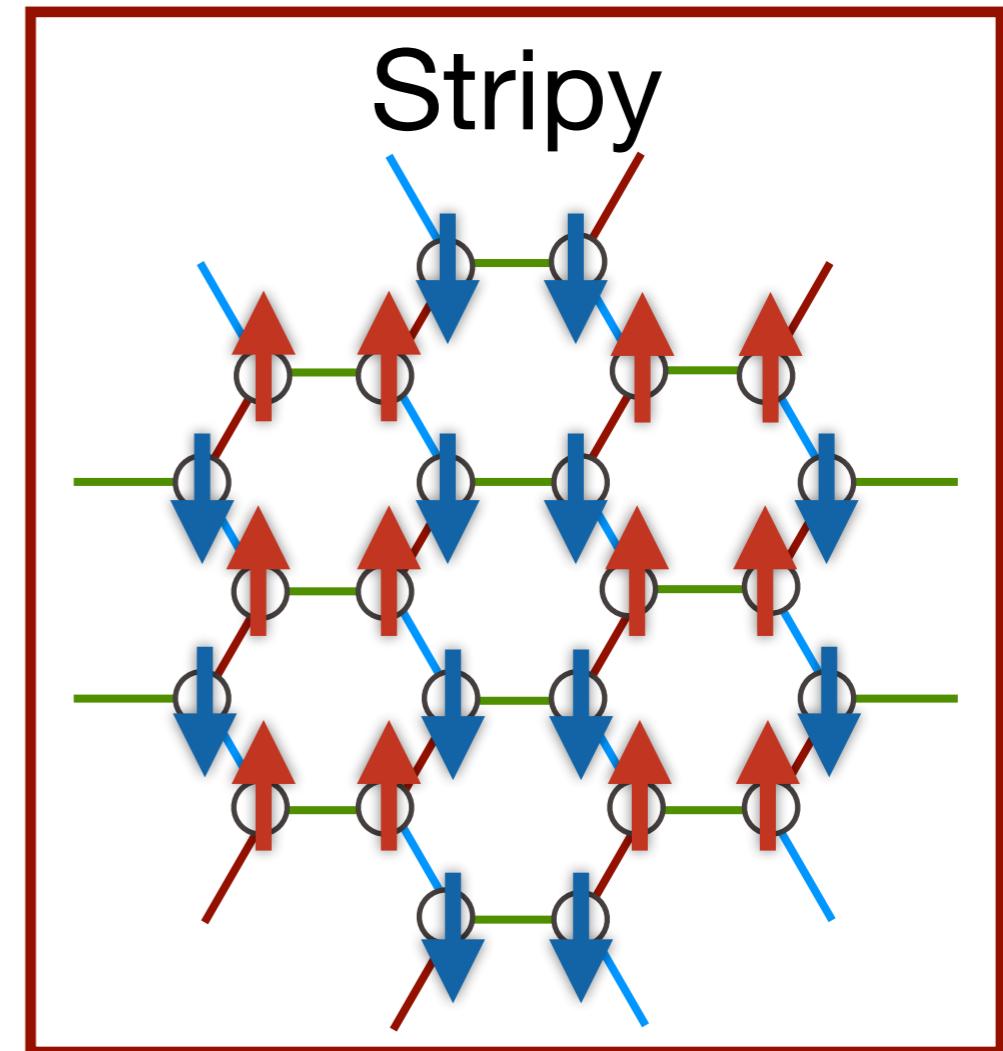
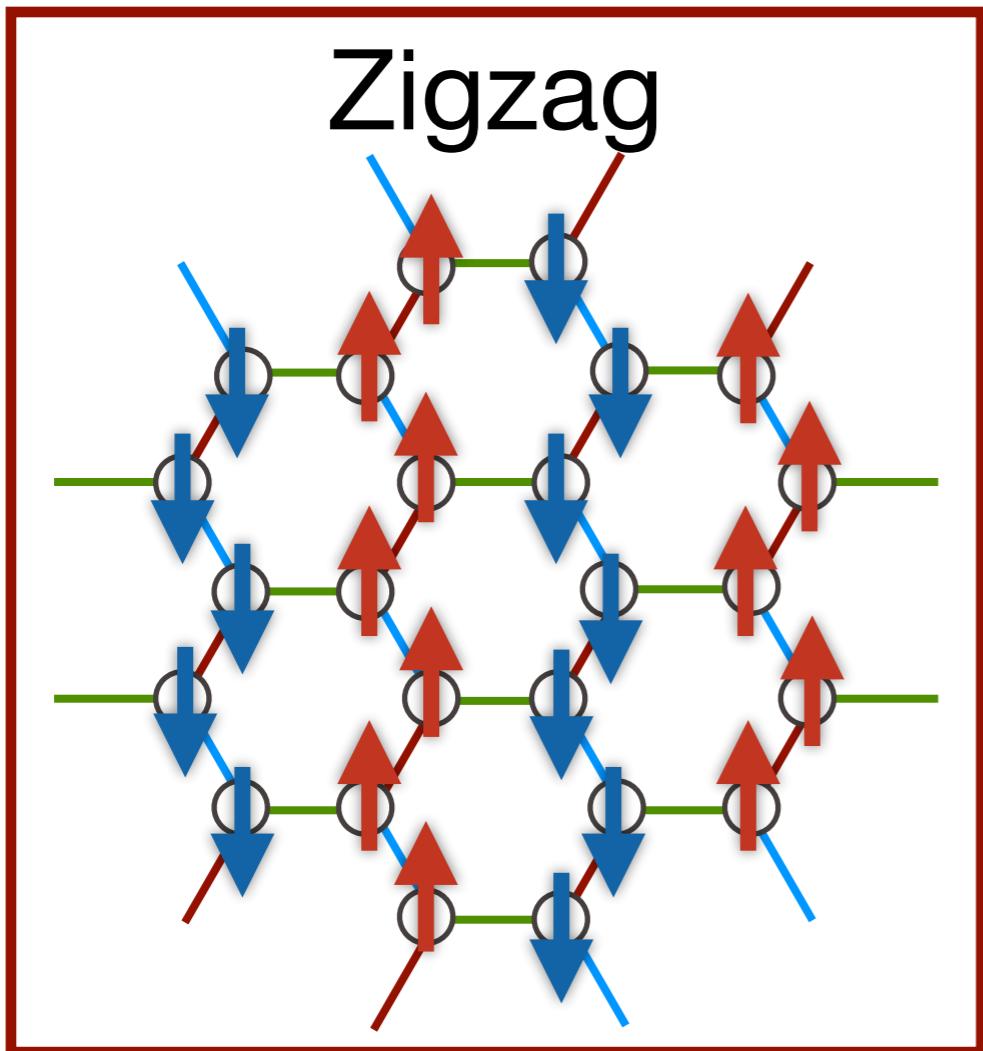
$$H_K^{(\gamma)} = -JS_i^{\gamma}S_j^{\gamma}$$



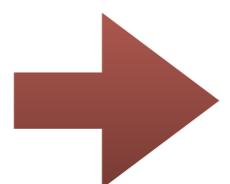
Magnetic order of Na_2IrO_3

In experiments, “Zigzag” ordered has been observed.

(F. Ye et al., PRB 85, 180403(R) (2012))



Combination of Kitaev and Heisenberg interactions
gives stripy order.



In Na_2IrO_3 , other interactions must exist.

ab initio Hamiltonian of Na₂IrO₃

(Y. Yamaji et al. Phys. Rev. Lett. **113**, 107201(2014))

ab initio Hamiltonian

$$\hat{H} = \sum_{\Gamma=X,Y,Z} \sum_{\langle \ell, m \rangle \in \Gamma} \vec{\hat{S}}_\ell^T \mathcal{J}_\Gamma \vec{\hat{S}}_m,$$

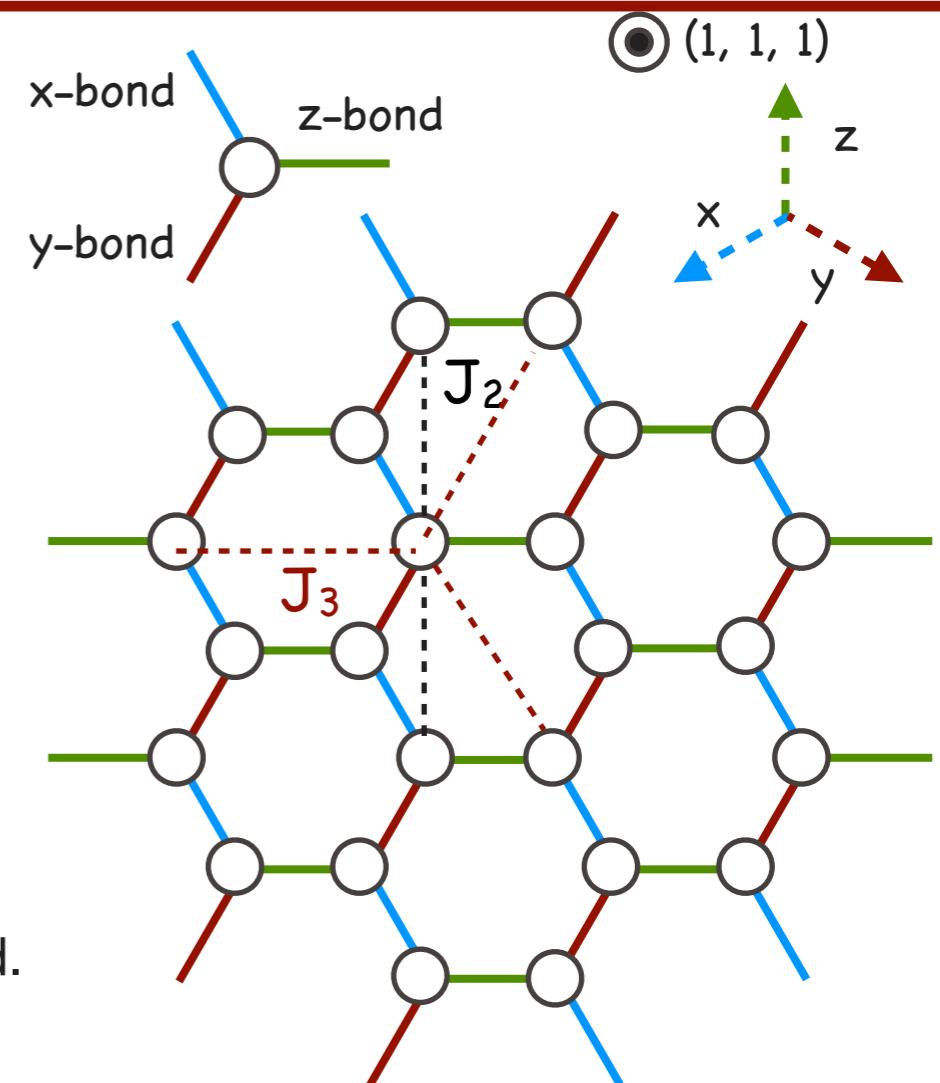
$$\mathcal{J}_Z = \begin{bmatrix} J & I_1 & I_2 \\ I_1 & J & I_2 \\ I_2 & I_2 & K \end{bmatrix}, \mathcal{J}_X = \begin{bmatrix} K' & I_2'' & I_2' \\ I_2'' & J'' & I_1' \\ I_2' & I_1' & J' \end{bmatrix}, \mathcal{J}_Y = \begin{bmatrix} J'' & I_2'' & I_1' \\ I_2'' & K' & I_2' \\ I_1' & I_2' & J' \end{bmatrix},$$

Kitaev coupling **K** and Heisenberg like coupling **J**
+
Off-diagonal couplings **I₁** and **I₂**

It also contains **J₂** and **J₃** interaction term.

J₂ : only “z-bond” which is perpendicular to NN z-bond.

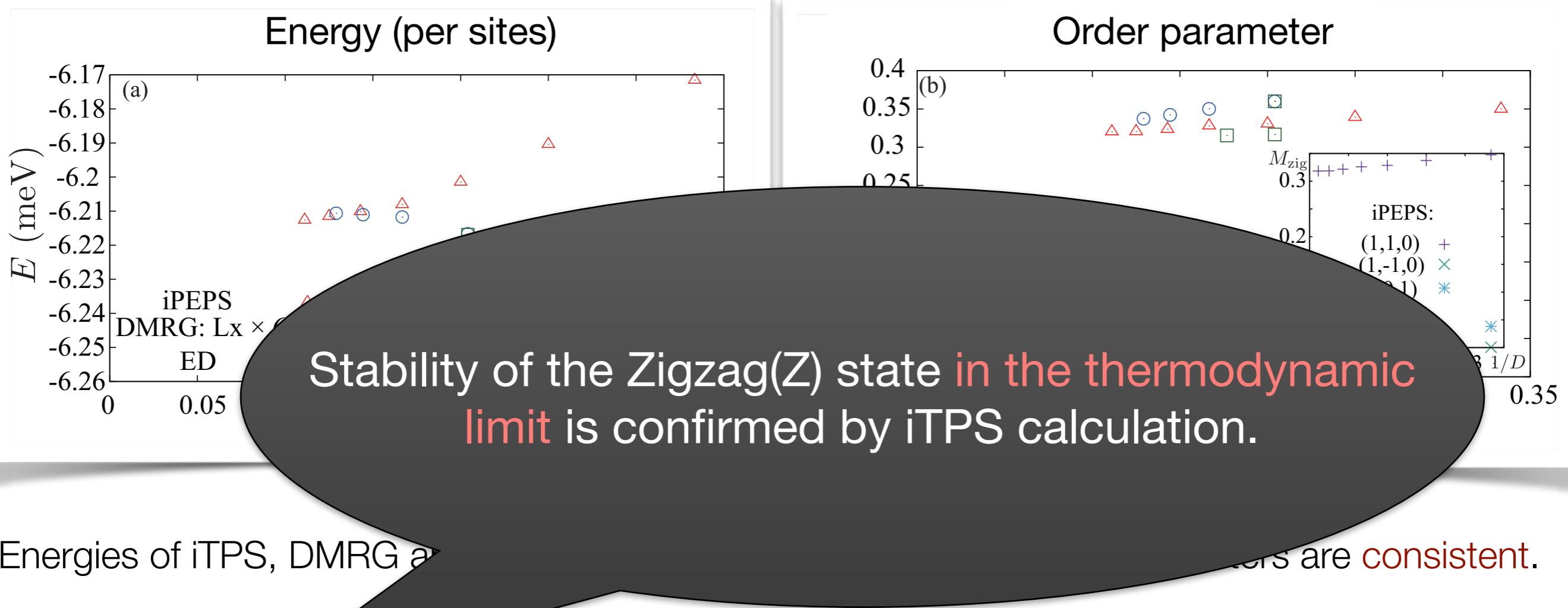
J₃ : all of the three third neighbors



Due to the trigonal distortion, the *ab initio* Hamiltonian contains strong off-diagonal couplings, together with **J₂** and **J₃** interaction

Results: comparison with other methods

T. Okubo *et al*, PRB **96**, 054434 (2017).



- For 4×6 lattice, DMRG and ED give almost same energy.
- Finite D of iTPS and finite L_x of DMRG are overlapped.

- Extrapolations of them are $\langle M \rangle \sim 0.3$
- Spins are almost along $(1,1,0)$ direction, which is consistent with the experimental observations.

Summary

- By choosing proper **tensor network structure**, ground state wave functions can be approximated accurately.
 - To search good tensor networks, **the area law of the entanglement entropy** is important.
 - For one dimensional quantum system, matrix product states (**MPS**) works **very well**.
 - For two or higher dimensional systems, MPS breaks down. In these case, instead, **tensor product states** are good tensor networks.
- Owing to **developments of algorithms and computers**, tensor network methods become powerful method to investigate **frustrated spin system** in two dimensions.
 - Infinite TPS can reproduce the Kitaev spin liquid accurately.
 - It also applicable to realistic Hamiltonian for Na_2IrO_3 .
 - In the case of Shastry-Sutherland lattice model, it successfully reproduce **several magnetization plateaus** consistent with experiments.