計算科学における情報圧縮

Information Compression in Computational Science **2017.12.27**

#12:テンソルネットワーク表現への発展

Tensor network representation

理学系研究科 物理学専攻 大久保 毅 Department of Physics, **Tsuyoshi Okubo**

Outline

- Breakdown of MPS representation
 - Critical system
 - Higher dimensional system
- Tensor Network for critical systems
 - Multi-scale Entanglement Renormalization Ansatz (MERA)
- Tensor Network for higher dimensions
 - Tensor Product State (TPS)
- Report problems

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



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

Phase transition

Transverse field Ising chain:

$$\mathcal{H} = -\sum_{i=1}^{N-1} S_i^z S_{i+1}^z - h \sum_{i=1}^N S_i^x$$

Ground state |1

$$h = 0$$
 : Ferromagnetic state

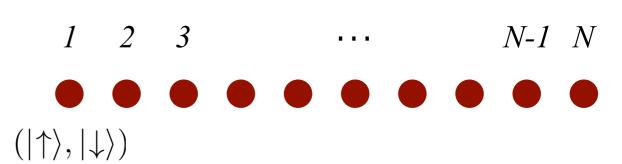
$$h \to \infty$$
: Disordered state (Field induced ferro)

In between these two limits, there is a phase transition.

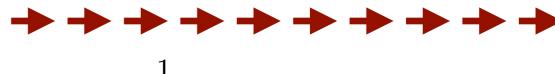
At the phase transition, order parameter becomes zero.

(秩序変数)

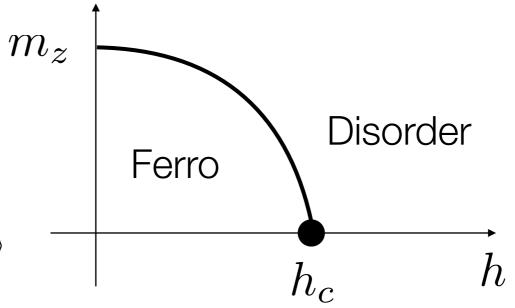
(Spontaneous)
$$m_z = \frac{1}{N} \sum_i \langle \Psi | S_i^z | \Psi \rangle$$
 (自発磁化)







$$= \frac{1}{\sqrt{2}} \left(|\uparrow\rangle + |\downarrow\rangle \right)$$



Critical point and correlation length

$$h = h_c$$
: Critical point (臨界点)

Behavior of a correlation function:

$$0 \le h < h_c$$
 :Ferromagnetic state

$$\langle \Psi | S_i^z S_{i+r}^z | \Psi \rangle \sim C e^{-\frac{r}{\xi}} + m_z^2$$

 $h_c < h$:Disordered state

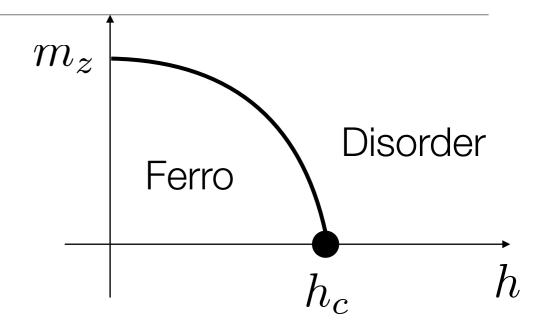
$$\langle \Psi | S_i^z S_{i+r}^z | \Psi \rangle \sim e^{-\frac{r}{\xi}}$$

$$h = h_c$$
 :Critical point

$$\langle \Psi | S_i^z S_{i+r}^z | \Psi \rangle \sim r^{-2p}$$

Ferro

Correlation length diverges at critical point!



$$\xi = \xi(h)$$
: Correlation length (相関長)

Scale invariance at the critical point

 $h = h_c$: Critical point (臨界点)

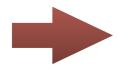
$$C(r) \equiv \langle \Psi | S_i^z S_{i+r}^z | \Psi \rangle \sim r^{-2p}$$

Power low decay!

After a scale transformation r' = br

$$C(r') = C(br) = b^{-2p}C(r)$$

Change in the correlation function is only a constant factor.



If we scale spins as $\tilde{S}_i^z = b^p S_i^z$

the correlation function becomes

$$\tilde{C}(r') \equiv \langle \Psi | \tilde{S}_i^z \tilde{S}_{i+r'}^z | \Psi \rangle = C(r)$$

This property is called as "scale invariance". (スケール不変性)

Physics (properties) in different scale is essentially same.

DMRG (variational MPS) calculation of TFI model

Ö. Legeze, and G. Fáth, Phys. Rev. B 53, 14349 (1996)

Errors of the ground and the 1st excited states energies varying system size *N*.

For a fixed dimension m,

300 10

N

40 100

N

300

Ferro and disordered states:

The errors are almost independent of N.

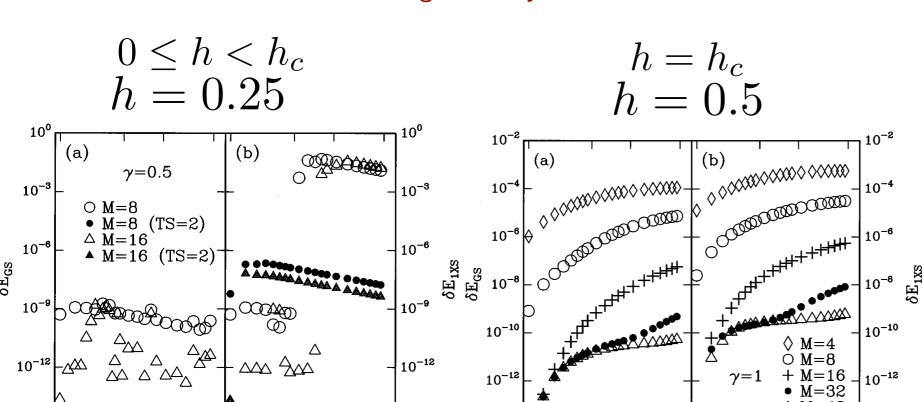
Critical point:

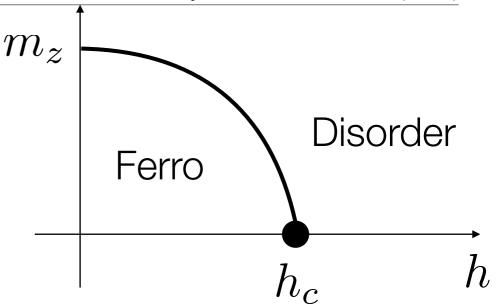
The errors gradually increases as increase N.

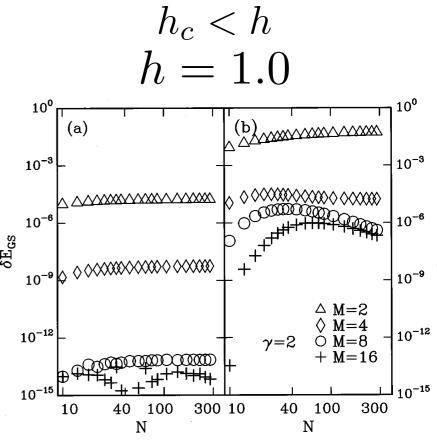
300 10

N

Ν



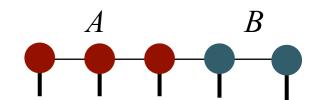




Entanglement entropy of TFI model

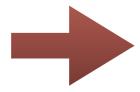
Entanglement entropy:

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



State	EE of the original vector	Required bond dimension
Ferro or Disordered	$S_A = O(1)$	$\chi = O(1)$
Critical	$S_A = O(\log N)$	$\chi = O(N^{\alpha})$

We need polynomially large bond dimension for critical system!



More efficient tensor network for critical systems?

Key point: Scale invariance of the system

Higer dimensional system

Transverse field Ising model on square lattice:

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

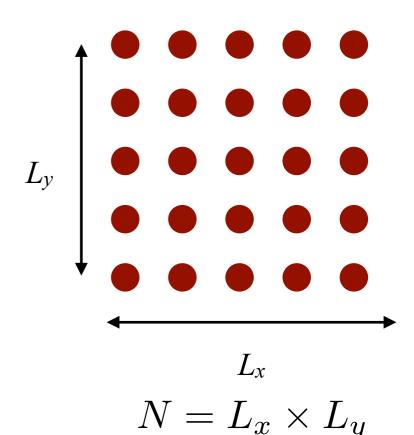
Summation over the $\langle i,j \rangle$ nearest neighbor pair

Area law

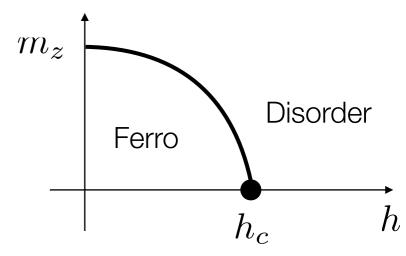
Even in ferro and disordered phases, the entanglement entropy depends on size N.

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

Two-dimensional array



Phase diagram



MPS for two-dimensional system

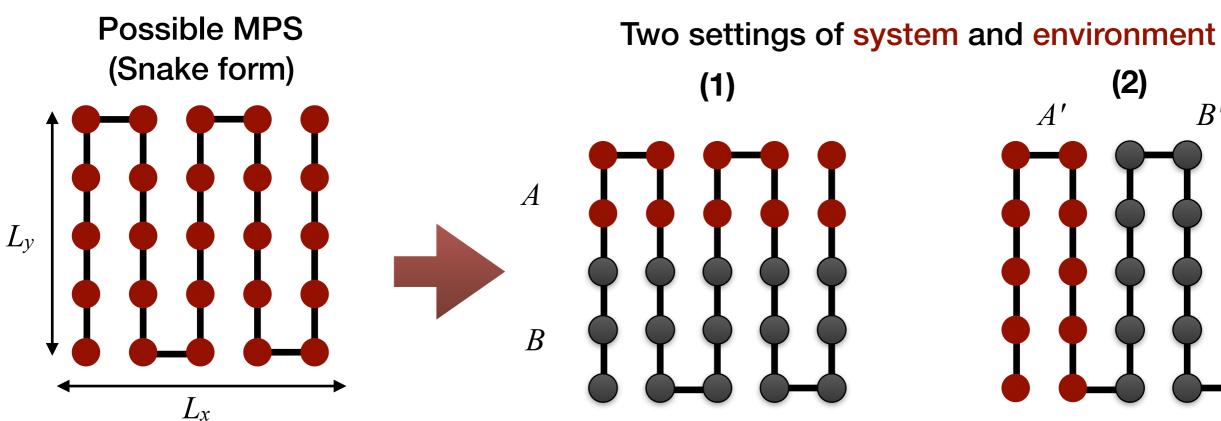
When we apply MPS representation for a square lattice system:

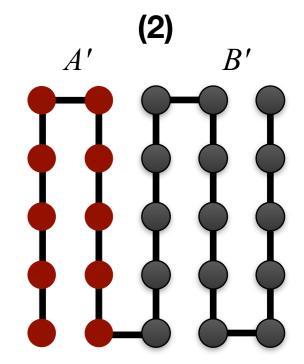
:Satisfying area law? Setting (1) $S_A \leq L_x \log \chi$

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, ...) dimensions.





MPS for two-dimensional system: comment

MPS can treat "rectangular" or "quasi one dimensional" lattice.

In setting (1), MPS can satisfy the area low partially.

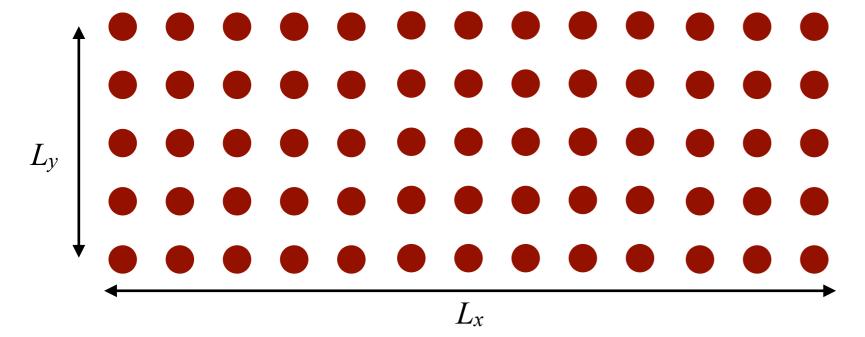


We can increase L_x easily with keeping L_y constant.

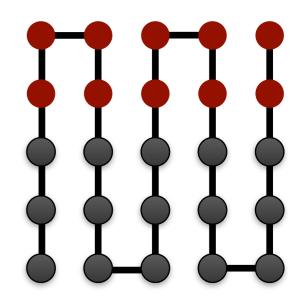
$$\chi = O(e^{L_y})$$

$$L_y \lesssim 10, L_x \gg L_y$$

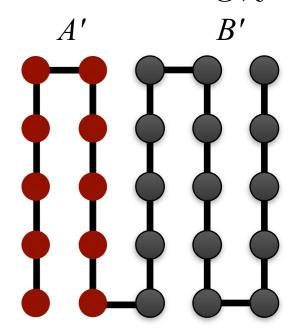
Quasi one dimensional system ("strip" or "cylinder")



(1) $S_A \leq L_x \log \chi$



(2) $S_{A'} \leq \log \chi$



Tensor network for critical systems: Multi-scale Entanglement Renormalization Ansatz

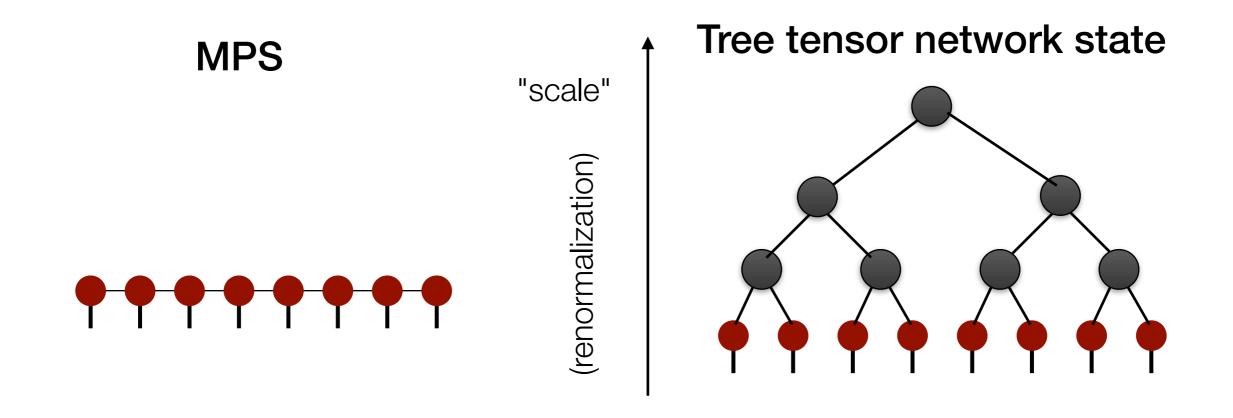
Hierarchical structure: tree tensor network

Critical system



Scale invariance

A simple scale invariant tensor network: tree tensor network



Notice:

Unitary tensors

Unitary tensor

$$U_{ij}^{kl} = \bigcup_{i=1}^{k} \bigcup_{j}^{l} \sum_{i,j} U_{ij}^{kl} (U^{\dagger})_{k'l'}^{ij} = \delta_{kk'} \delta_{ll'}$$

$$(U^{\dagger})_{kl}^{ij} = (U_{ij}^{kl})^*$$

$$\sum_{i,j} U_{ij}^{kl} (U^{\dagger})_{k'l'}^{ij} = \delta_{kk'} \delta_{ll'}$$

$$\sum_{k,l} (U^{\dagger})^{ij}_{kl} U^{kl}_{i'j'} = \delta_{ii'} \delta_{jj}$$

$$\sum_{k,l} (U^{\dagger})^{ij}_{kl} U^{kl}_{i'j'} = \delta_{ii'} \delta_{jj'}$$

$$U$$

$$i'$$

$$U$$

$$i'$$

$$i'$$

$$i'$$

Isometric tensors

Isometric tensor (half unitary tensor) = Isometry

$$W_{ij}^k = \sum_{i,j} W_{ij}^k (W^\dagger)_{k'}^{ij} = \delta_{kk'}$$

Unitarity condition only for "bottom" legs.

Isometry works as a "projector" from the bottom space to the top space.

$$\dim(\text{bottom}) \ge \dim(\text{top})$$

It is also related to the "renormalization" of degree of freedoms.

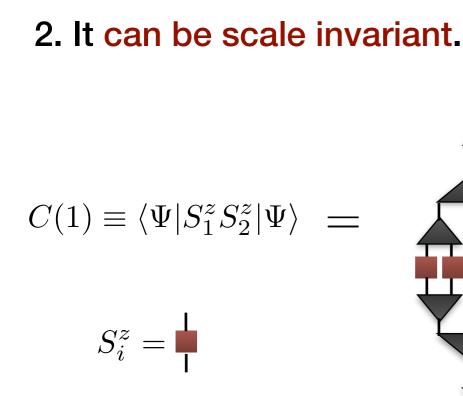
We pick up "important" degree of freedoms by isometries.

Isometric tree tensor network and its scale invariance

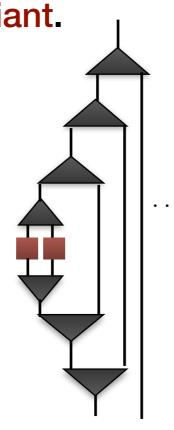
Consider an (infinite) tree tensor network consists of identical isometries as a wave function.

Properties:

1. It is normalized as $\langle \Psi | \Psi \rangle = 1$ (Trivial from the definition of the isometry)



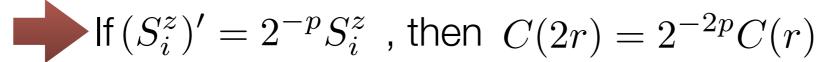
spin



$$C(2) \equiv \langle \Psi | S_1^z S_3^z | \Psi \rangle =$$

$$(S_i^z)' =$$

renormalized" spin



Scale invariant!

Entanglement entropy of TTN

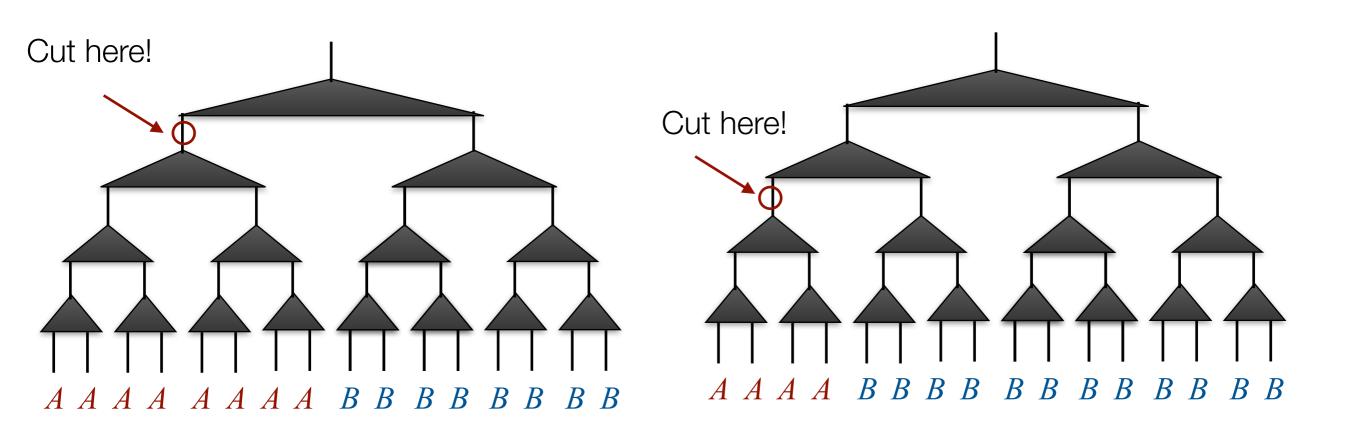
Entanglement entropy of tree tensor networks (TTN):

Due to the tree structure, two regions are connected by only "one bond".

(or a few)



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

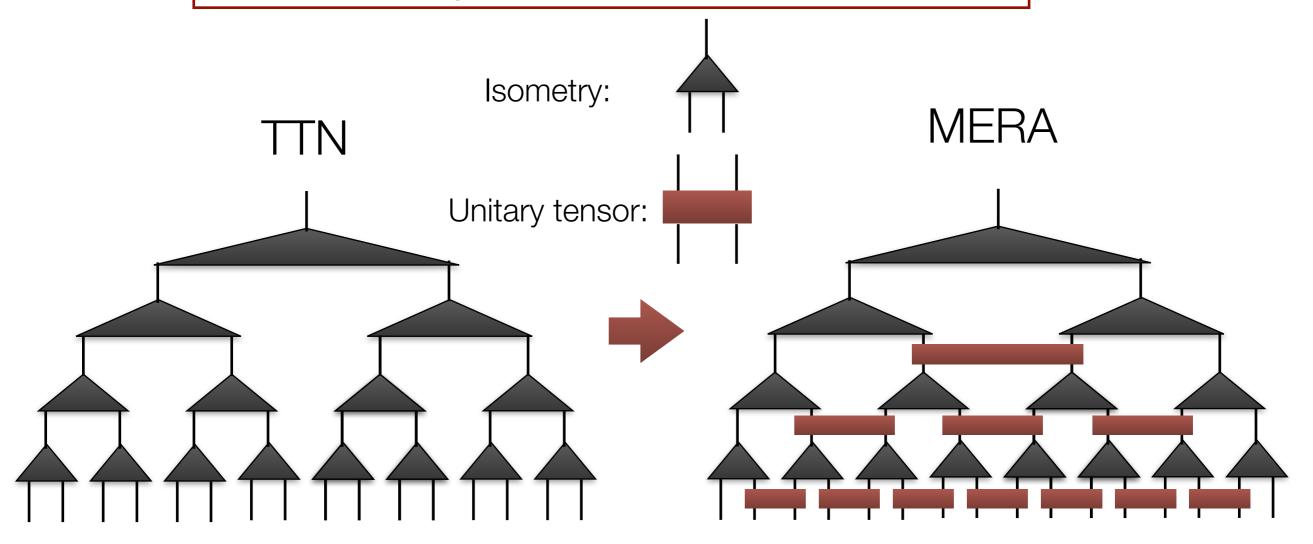


MERA

- (G. Vidal, Phys. Rev. Lett. 99, 220405 (2007))
- (G. Vidal, Phys. Rev. Lett. 101, 110501 (2008))

Multi-scale Entanglement Renormalization Ansatz (MERA)

Before applying isometry, insert a unitary tensor.

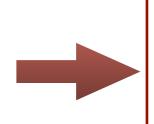


Mormalization

Scale invariance (if we set the identical tensors)

Entanglement entropy of MERA

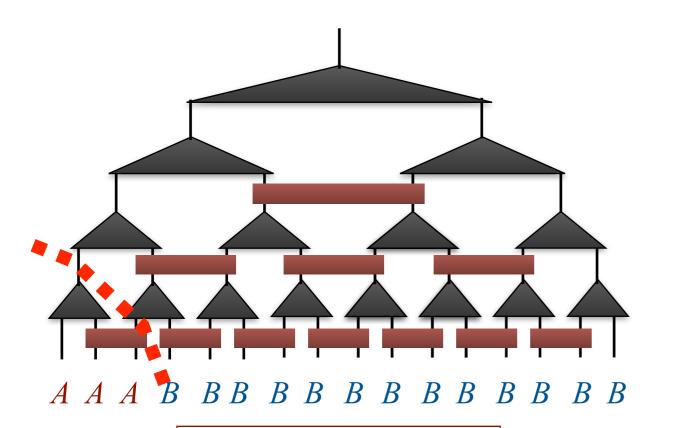
Due to the unitary matrices, # of bonds connecting two regions logarithmically increase.

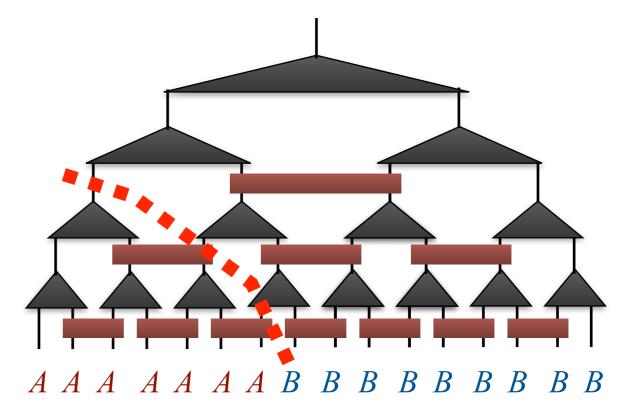


rank
$$\rho_A \le \chi^{N_c(N)} \sim \chi^{\log N}$$

 $S_A = -\text{Tr } \rho_A \log \rho_A \le (\log \chi) \log N$

 $N_c(N)$ # of minimum cut for a N-site region





Minimum # of cuts = 2

Minimum # of cuts = 3

Application of MERA

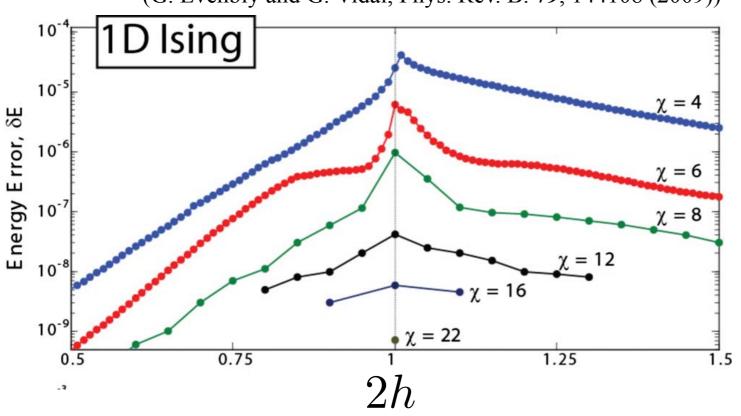
Transverse field Ising chain:

$$\mathcal{H} = -\sum_{i=1}^{N-1} S_i^z S_{i+1}^z - h \sum_{i=1}^N S_i^x$$

Energy errors:

MERA (Infinite chain)

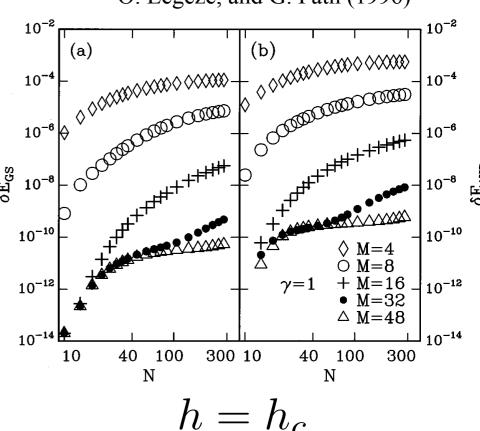
(G. Evenbly and G. Vidal, Phys. Rev. B. 79, 144108 (2009))



MERA can represent very large (Infinite) critical system!

DMRG (finite chain)

Ö. Legeze, and G. Fáth (1996)



Interesting topics related to MERA

- By using scale invariance of MERA, we can calculate properties of critical system accurately.
 (G. Evenbly and G. Vidal, Phys. Rev. B. 79, 144108 (2009))
 (R.N.C. Pfeifer, (G. Evenbly and G. Vidal, Phys. Rev. A. 79, 040301(R) (2009))
 - Critical exponents and Operator product expansion coefficients in the Conformal Filed Theory (CFT)
- We can consider MERA in higher dimensions
 - It is scale invariant but satisfies the area low

(G. Evenbly and G. Vidal, Phys. Rev. Lett. 102, 180406 (2009))

 For the system with logarithmic correction in the EE, such as metal, "branching MERA" has been proposed.

(G. Evenbly and G. Vidal, Phys. Rev. Lett. 112, 220502 (2014))

(G. Evenbly and G. Vidal, Phys. Rev. B. **89**, 235113 (2014))

- Relation between MERA and other fields
 - Wavelet transform (G. Evenbly and S. R. White, Phys. Rev. Lett. 112, 140403 (2016))
 - AdS/CFT (quantum gravity, black hole)

(M. Nozaki, S. Ryu, and T. Takayanagi, J. High Energy Phys. 10, 193 (2012))

Tensor network for higher dimensional systems: Tensor Product State (Projected Entangled Pair State)

Entanglement entropy in higher dimensions

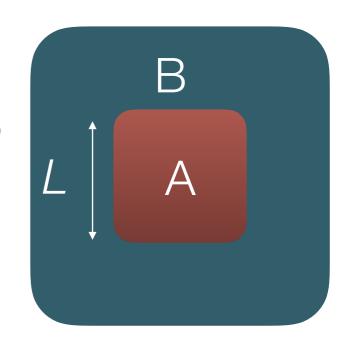
Ground state wave functions:

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 low:

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



In d=1, MPS satisfies the area low.



- 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

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

Tr: tensor network "contraction"

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

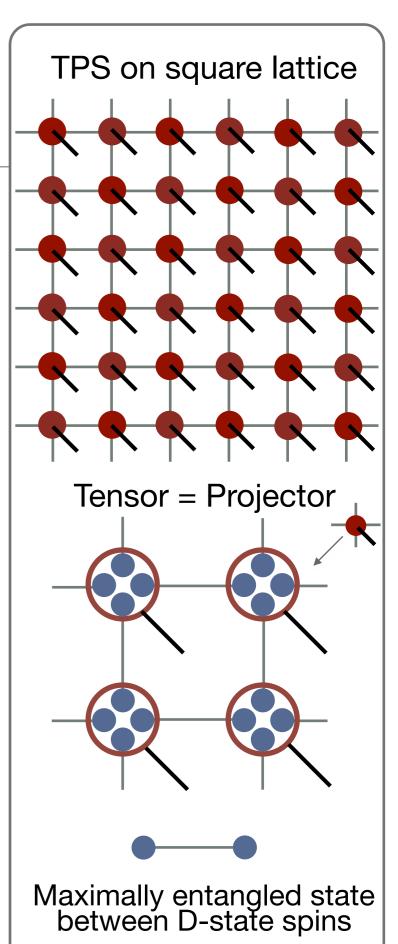
$$x'$$
 x x x

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

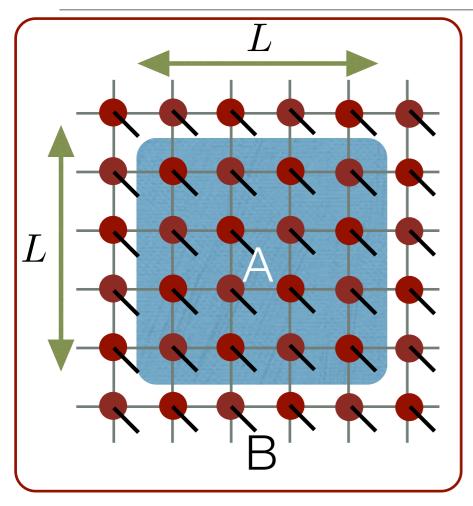
$$m_i = 1, 2, \dots m_i$$

x,y,x',y'=1,2,...D D= "bond dimension" m=1,2,...m m= dimension of the local Hilbert space

*D can be larger than m. "Virtual state "



Entanglement entropy of TPS (PEPS)



$$x'$$
 y
 m

Bond dimension = D

of bonds connecting regions A and B

$$N_c(L) = 4L$$
 (square lattice)



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

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 satisfies 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!

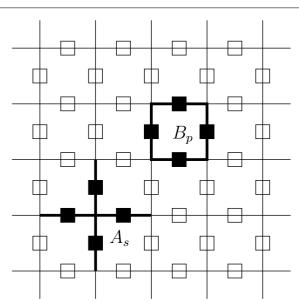
Example: Ground state represented by TPS

Toric code model

(A. Kitaev, Ann. Phys. **303**, 2 (2003).

$$\mathcal{H} = -\sum_{s} A_{s} - \sum_{p} B_{p}$$

$$A_s = \prod_{j \in \text{star}(s)} \sigma_j^x \qquad B_p = \prod_{j \in \partial p} \sigma_j^z.$$





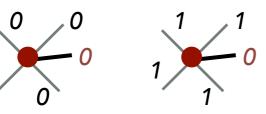
Its ground state is so called Z_2 spin liquid state.

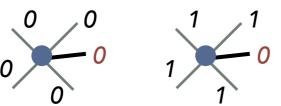
"Spin liquid" is a novel phase different from conventional magnetic orders.

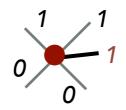
It can be represented by D=2 TPS.

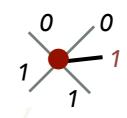
(F. Verstraete, et al, Phys. Rev. Lett. 96, 220601 (2006).

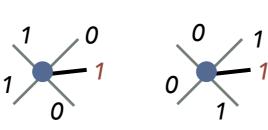
0,1: eigen state of σ_x

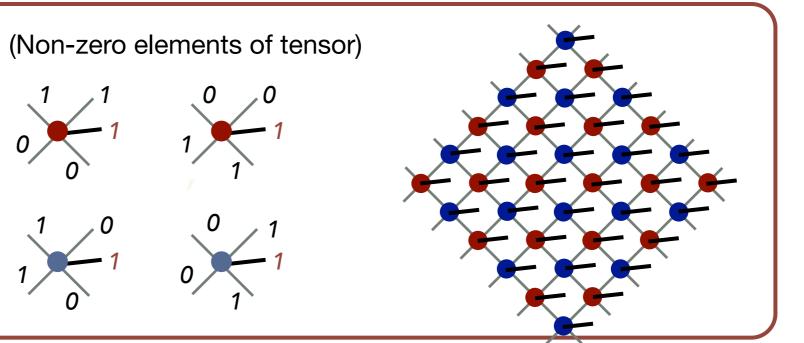












Difference between MPS and TPS

Cost of tensor network contraction:

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

MPS: O(N)

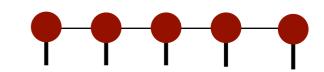
TPS: $O(e^{L^{d-1}})$



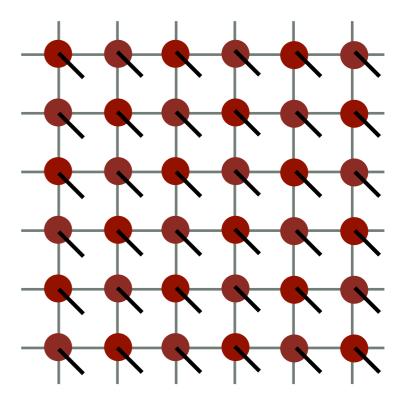
It is impossible to perform exact contraction even if we know local tensors in the case of TPS.

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

MPS



TPS (PEPS)



Example of approximate contraction: CTM method

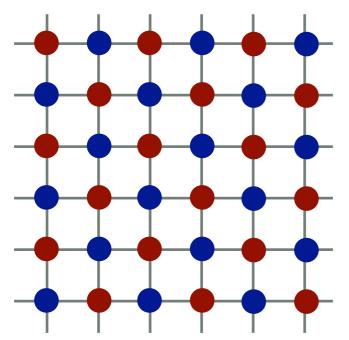
For (infinite) open boundary system

(T. Nishino and K. Okunishi, JPSJ 65, 891 (1996))

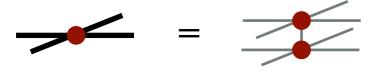
(R. Orus et al, Phys. Rev. B 80, 094403 (2009))

Infinite PEPS

(with a translational invariance)



Double tensor





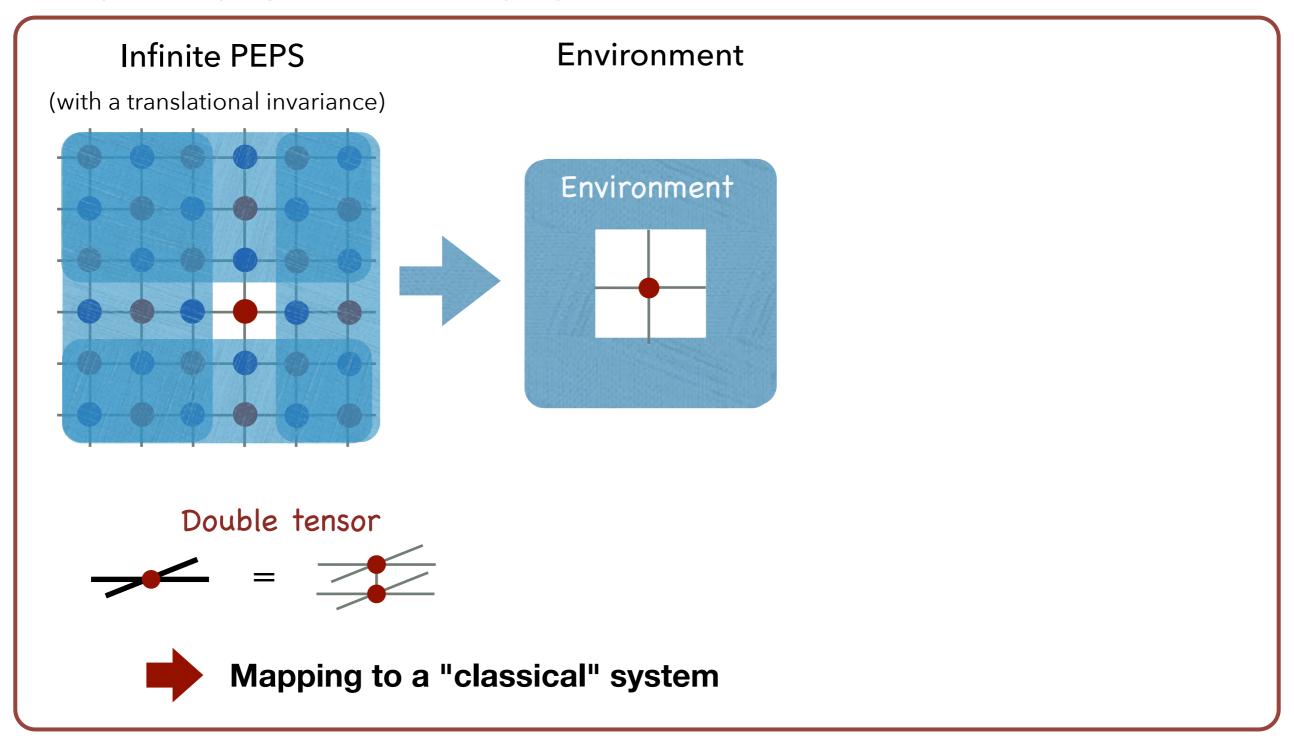
Mapping to a "classical" system

Example of approximate contraction: CTM method

For (infinite) open boundary system

(T. Nishino and K. Okunishi, JPSJ 65, 891 (1996))

(R. Orus et al, Phys. Rev. B 80, 094403 (2009))

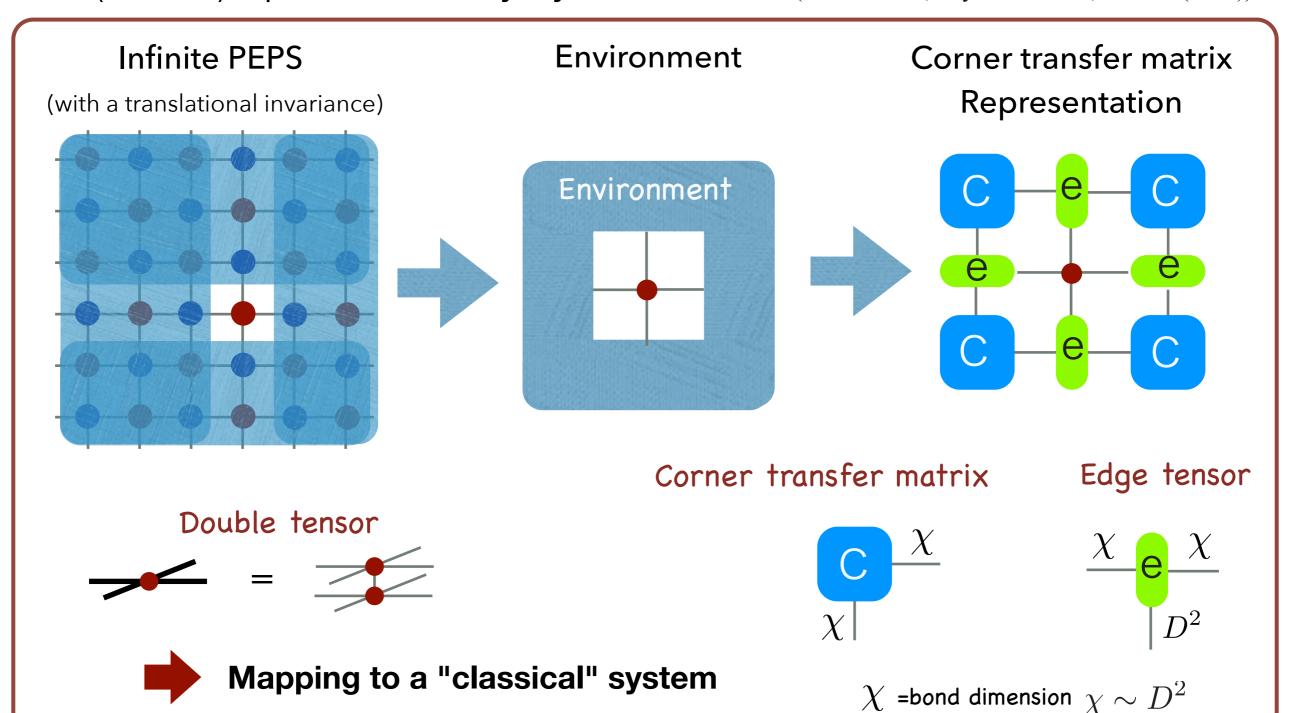


Example of approximate contraction: CTM method

For (infinite) open boundary system

(T. Nishino and K. Okunishi, JPSJ 65, 891 (1996))

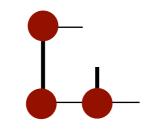
(R. Orus et al, Phys. Rev. B 80, 094403 (2009))



Cost of (approximate) contraction

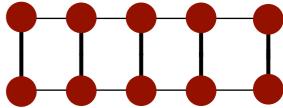
MPS:

$$\chi - \frac{m}{\chi} \quad (m \ll \chi)$$



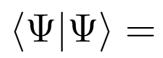
$$O(\chi^3)$$





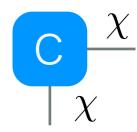
TPS:

$$D \mid_{D}$$





When we use CTM environment in 2D,



$$\chi$$
 e χ



$$D^2 D^2$$

$$O(\chi^2 D^6), O(\chi^3 D^4) \sim O(D^{10}) \quad (\chi \sim D^2)$$

We can treat very small bond dimensions in TPS!

Application of TPS to eigenvalue problem

For calculation of minimum eigenvalues and its eigenvector, we can use similar techniques to those in MPS

Variational method:

(P. Corboz, Phys. Rev. B **94**, 035133 (2016))

(L. Vanderstraeten, Phys. Rev. B 94, 155123 (2016))

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

Imaginary time evolution:

$$|\Psi_{\rm GS}\rangle \propto \lim_{\beta \to \infty} e^{-\beta \mathcal{H}} |\Psi_0\rangle$$

For a initial state $\langle \Psi_{\rm GS} | \Psi_0 \rangle \neq 0$

(H. G. Jiang *et al*, Phys. Rev. Lett. **101**, 090603 (2008)) (J. Jordan *et al*, Phys. Rev. Lett. **101**, 250602 (2008))

Example of application: Honeycomb lattice Kitaev Model

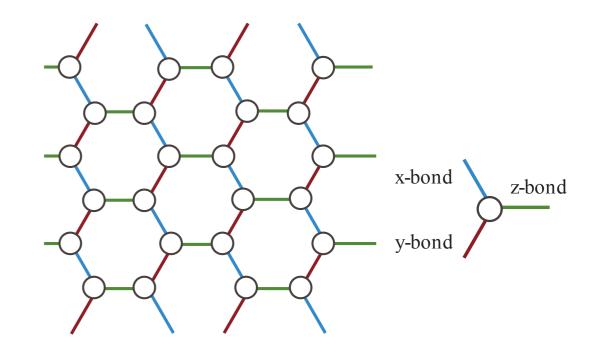
A. Kitaev, Annals of Physics 321, 2 (2006)

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.



Phase diagram

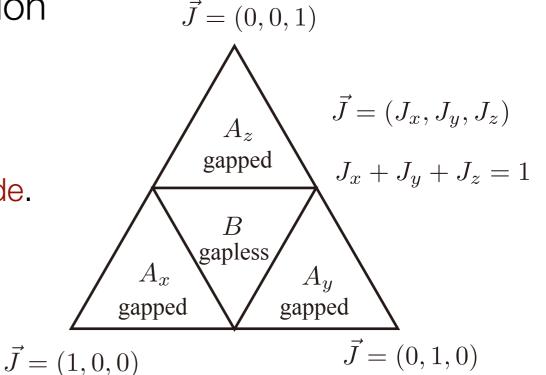
Exactly solvable by introducing Majorana fermion

Isotropic region (B): gapless spin liquid

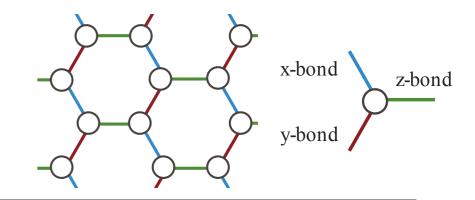
Anisotropic region (A): gapped spin liquid

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!







Honeycomb lattice Kitaev model

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

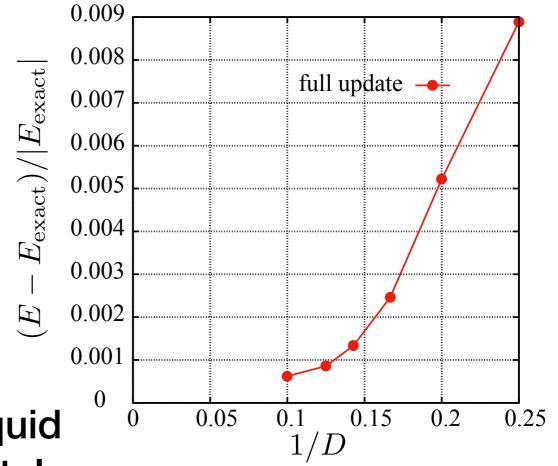
In the present (super)computers, we can access around D=10 (maybe 16) by using massively parallel code.

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

$$\mathcal{H} = -\sum_{\substack{\gamma, \langle i,j \rangle_{\gamma} \\ (\gamma = x, y, z)}} J_{\gamma} S_{i}^{\gamma} S_{j}^{\gamma}$$

Energy error obtained by iTPS

(T. okubo et al, unpublished)





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

Interesting topics related to TPS

Application to itinerant electron system, which may break the area law

```
(P. Corboz et al, Phys. Rev. B. 81, 165104 (2010))
(P. Corboz, Phys. Rev. B. 93, 045116 (2016))
```

- Characterization of topologies in wave function
 - Symmetric tensor network and modular matrix (J.-W. Mei et al, Phys. Rev. B. 95, 235107 (2017))
- Application to three dimensions
 - So far, there is no practical calculations for non-trivial models.
 - Mainly, due to the scaling: $O(D^{18})$?

Report problems

- Pleas solve both of the problems:
 - 1. Report problem (Partial SVD)
 - 2. Report problem (MPS)
- Please include your name and student id in your report.
- Please subit through ITC LMS
 (If you have any troubles, please send us email: t-okubo@phys.s.u-tokyo.ac.jp
- Deadline is 2018/1/10 (23:30)

Report Problem: Partial SVD 1

SVD-1. (compulsory) From given sets of partial rating matrices, infer the complete rating matrices by using PSVD.py. At least, inset01.in and inset11.in should be used. In your report, k_u and k_o you used and the inferred matrices should be included.

-We prepared 2 complete rating matrices: A random matrix and a correlated matrix. From these matrices, we randomly sample and make two sets of rating matrices.

inset01.in, inset02.in, inset03.in, inset04.in are sampled from a matrix, and inset11.in, inset12.in, inset13.in, inset14.in from the other matrix.

-PSVD.py receives an input file, a rank for partial SVD *f*, and maximum iteration *imax* for the steepest gradient.

python ./PSVD.py inputfile f imax

Report Problem: Partial SVD 2

SVD-2. (optional) Infer which set of rating matrix is sampled from a random rating matrix. Show the basis (根拠) for your inference (推定).

- -For example, you may compare the inferred rating matrices from (inset01.in, inset02.in, inset03.in, inset04.in) or (inset11.in, inset12.in, inset13.in, inset14.in) and measure variance of the inferred matrix elements.
- -Variance may depends on the hyperparameters k_u and k_o

Revisit: Singular Value Decomposition of Partially Unknown Matrix

As a review,

Chih-Chao Ma, A Guide to Singular Value Decomposition for Collaborative Filtering.

Mathematical formulation

Minimize the cost function

$$E = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} I_{ij} [V_{ij} - p(U_i, O_j)]^2$$

$$+\frac{k_u}{2}\sum_{i=1}^n\|U_i\|_2^2+\frac{k_o}{2}\sum_{j=1}^m\|O_j\|_2^2$$
 -L₂ regularization

Feature vectors

$$U_i = (U_{1i}, U_{2i}, \dots, U_{fi})^T$$
$$O_j = (O_{1j}, O_{2j}, \dots, O_{fj})^T$$

A prediction function for the rating (You need to choose)

$$p(U_i, O_j) = \begin{cases} 1 & \text{if } U_i^T O_j < 0 \\ 1 + U_i^T O_j & \text{if } 0 < U_i^T O_j < 4 \\ 5 & \text{if } 4 < U_i^T O_j \end{cases}$$

Revisit: Singular Value Decomposition of Partially Unknown Matrix

A predicted rating matrix

$$p(U_i, O_j) = \begin{bmatrix} 4.92 & 4.39 & 2.29 & 2.96 & 4.93 & 3.23 \\ 2.57 & 2.02 & 1.07 & 1.78 & 2.94 & 1.60 \\ 3.01 & 3.82 & 3.92 & 2.05 & 2.55 & 3.94 \\ 4.98 & 4.91 & 3.11 & 3.00 & 4.64 & 3.85 \\ 4.55 & 3.85 & 2.01 & 2.78 & 4.92 & 2.98 \end{bmatrix}$$

Report problem (MPS)

- 1.Try MPS approximation of vectors (compulsory)
 - 1. Prepare random vector m^N
 - 2. Make exact MPS from it.
 - 3. By using low rank approximation based on SVD, make approximate MPS.
 - 4. Calculate norm between exact and approximate MPSs Norm = $||\vec{v}_{ex} \cdot \vec{v}_{ap}^*||$
 - 5. Vary N and m and investigate (and discuss) behavior of norm This can be done by sample python code Report_Random.py Usage:

python Report_Random.py -N N -m m -chi chi_max

```
Koto:Report_MPS okubo$ python Report_Random.py -N 10 -m 2 -chi 10
Parameters:N,m,chi_max= 10 2 10
Truncation: chi_max = 10
norm between exact and truncated MPS [ 0.73742445]
```

Report problem (MPS)

- 1. Try MPS approximation of spin model (optional)
 - 1. Prepare GS of a N site spin model
 - 2. Make exact MPS from it.
 - 3. By using low rank approximation based on SVD, make approximate MPS.
 - 4. Calculate energies and norm between exact and approximate MPSs
 - 5. Vary N, m and model parameters, and investigate (and discuss) behavior of norm and energy difference This can be done by sample python code Report_Model.py

python Report_Model.py -N N -m m -chi $chi \; max$ -Jz Jz -Jx Jx -hx hx

```
okubo$ python2.7 Report_Model.py -N 10 -m 2 -chi 10 -Jz -1.0 -Jxy 0.0 -hx 0.5 Model parameters:Jz,Jxy,hx= -1.0 0.0 0.5 Parameters:N,m,chi_max= 10 2 10 Ground state energy per bond= -0.343930277768 Energy of Exact MPS -0.343930277768 Truncation: chi_max = 10 Energy of MPS with truncation -0.343930277768 norm between exact and truncated MPS [ 1.]
```

Next week

第1回: 現代物理学における巨大なデータ

第2回: 情報圧縮と繰り込み

第3回: 情報圧縮の数理1 (線形代数の復習)

第4回: 情報圧縮の数理2 (特異値分解と低ランク近似)

第5回: 情報圧縮の数理3 (スパース・モデリングの基礎)

第6回: 情報圧縮の数理4 (クリロフ部分空間法の基礎)

第7回: 物質科学における情報圧縮

第8回: スパース・モデリングの物質科学への応用

第9回: クリロフ部分空間法の物質科学への応用

第10回: 行列積表現の基礎

第11回: 行列積表現の応用

第12回: テンソルネットワーク表現への発展

第13回: テンソルネットワーク繰り込みと低ランク近似の応用

"Tensor network renormalization " (We will assign report problem on 1/11.)