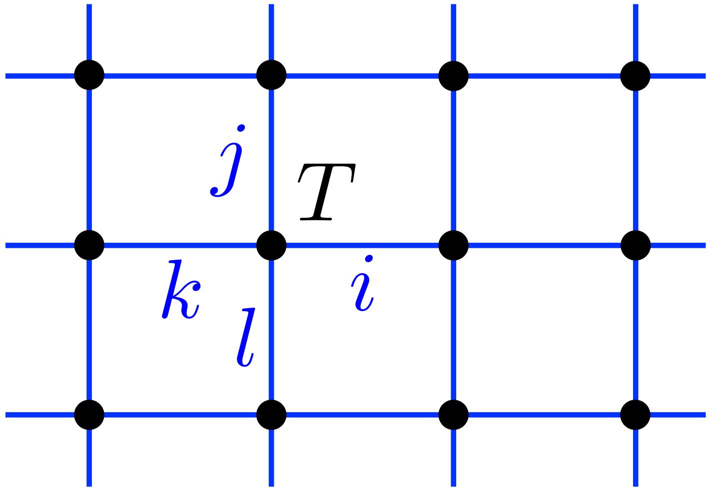
Tensor Networks

Shinji Takeda

LATTICE 2016

July 26, 2016, University of Southampton, UK

What's tensor network?

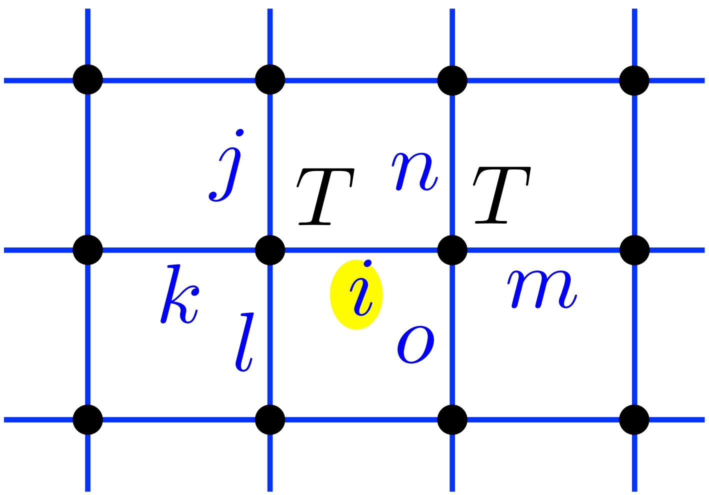


$$T_{ijkl}$$

tensor : lattice point

indices : link

What's tensor network?



$$\Leftrightarrow \sum_{\cdots,i,\cdots} \cdots T_{ijkl} T_{mnio} \cdots$$

contraction

A target quantity (wave function/partition function) is represented by tensor network

Why tensor networks?

Free from the sign problem

$\because$ Probability does not enter

Tensor network approaches

Hamiltonian/Hilbert space	**Lagrangian/Path integral**
Quantum many-body system	Classical many-body system/path integral rep. of quantum system
Wave function of ground state/excited states	Partition function
Variational method	Approximation, Coarse graining
Real time, Out-of-equilibrium, Quantum simulation	Useful in equilibrium system suffering from the sign problem in MC($\mu \neq 0$, $\theta \neq 0$, etc.)
DMRG, MPS, PEPS, MERA, …	TRG, SRG, HOTRG, TNR, Loop-TNR, …

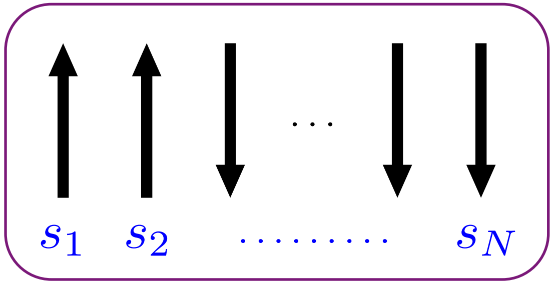
Hamiltonian approach

Matrix Product State (MPS) originated from DMRG White 1992, Schollwoeck 2004

- Ansatz of Wave function (in Tensor network representation)
- Variational method to obtain ground state in 1D quantum sys. (gapped)
- Quantum entanglement is taken into account
- By using information compression (SVD), one can drastically reduce the # of parameters $O(2^N) \rightarrow O(N)$ while keeping accuracy reasonably

2^N elements $\quad s_i = \pm$

$$|\psi\rangle = \sum_{s_1, s_2, \ldots, s_N} \underline{\psi_{s_1, s_2, \ldots, s_N}} |s_1\rangle \otimes |s_2\rangle \otimes \cdots \otimes |s_N\rangle$$



$$\hat{s}|\pm\rangle = \pm|\pm\rangle$$

Hamiltonian approach

Matrix Product State (MPS) originated from DMRG White 1992, Schollwoeck 2004

- Ansatz of Wave function (in Tensor network representation)
- Variational method to obtain ground state in 1D quantum sys. (gapped)
- Quantum entanglement is taken into account
- By using information compression (SVD), one can drastically reduce the # of parameters O(2^N)→O(N) while keeping accuracy reasonably

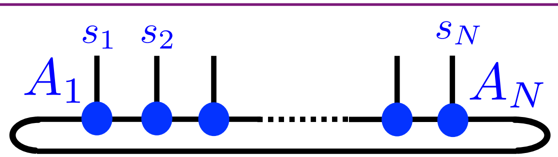
2^N elements $\quad s_i = \pm$

$$|\psi\rangle = \sum_{s_1, s_2, \dots, s_N} \underline{\psi_{s_1, s_2, \dots, s_N}} |s_1\rangle \otimes |s_2\rangle \otimes \cdots \otimes |s_N\rangle$$

$\downarrow$ information compression

$$\underline{\mathrm{tr}\left[A_1^{s_1} A_2^{s_2} \cdots A_N^{s_N}\right]}$$: Matrix product

$2d^2N$ elements $\quad A^s : d \times d$ matrix



Tensor network rep. of wave function

Hamiltonian approach

related with high energy physics

- Schwinger model, periodic BC, spectrum, DMRG
  Byrnes, Sriganesh, Bursill & Hamer, PRD66(2002)013002
- Schwinger model, open BC, spectrum, MPS
  Banuls, Cichy, Cirac, Jansen, JHEP11(2013)158
- Schwinger model, T≠0, MPS+MPO
  Banuls, Cichy, Cirac, Jansen & Saito, PRD92(2015)034519, PRD93(2016)094512
- 2D SU(2), real time dynamics of string breaking, MPS
  Kuehn, Zohar, Cirac & Banuls, JHEP07(2015)130
- Schwinger model with $N_f \neq 1$ and $\mu \neq 0$
  Kuehn, 7/25(Mon.)17:25 TD
- Fermions with long range interaction using MPS
  Szyniszewski, today's Poster

A Hamiltonian approach: TNS as ansatz for states

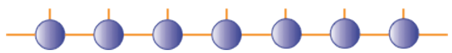
Test bench: Schwinger model (QED2)

from K. Cichy

$$H = -\frac{i}{2a}\sum_n \left(\phi_n^\dagger e^{i\theta_n}\phi_{n+1} - \text{h.c.}\right) + m\sum_n (-1)^n \phi_n^\dagger \phi_n + \frac{ag^2}{2}\sum_n L_n^2$$

At T=0, ground state and low lying excitations can be represented as MPS

Banuls, Cichy, Cirac, Jansen, Saito, JHEP11(2013)158; PoS LAT13, 332

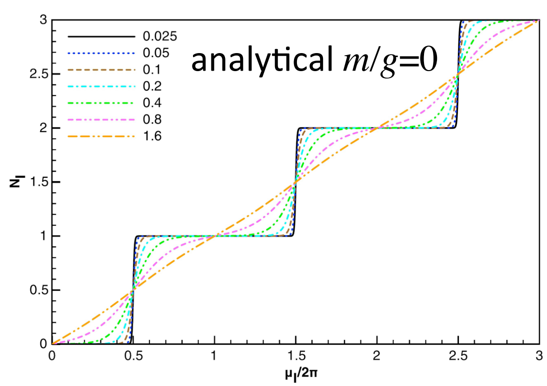


- variational search
- reliable control of systematic errors performed
- attain reliable continuum limit

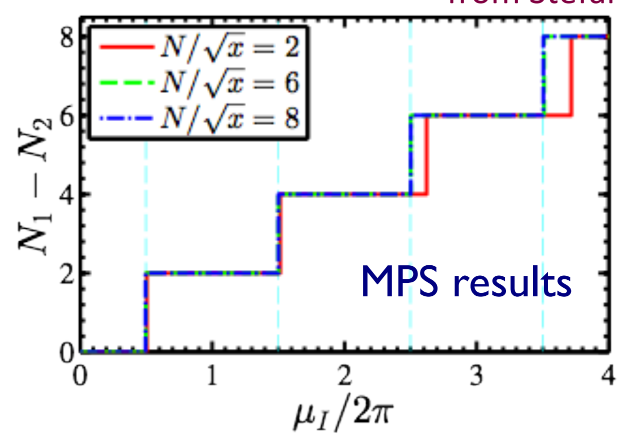
	Vector mass gap		Scalar mass gap	
m/g	MPS with OBC	DMRG	MPS with OBC	SCE
0	**0.56421(9)**	0.56419(4)	**1.1279(12)**	1.11(3)
0.125	**0.53953(5)**	0.53950(7)	**1.2155(28)**	1.22(2)
0.25	**0.51922(5)**	0.51918(5)	**1.2239(22)**	1.24(3)
0.5	**0.48749(3)**	0.48747(2)	**1.1998(17)**	1.20(3)

Sign problem is overcome: multiflavor case with chemical potential

from Stefan Kühn's talk



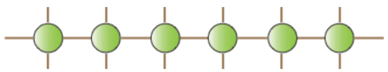
Lohmayer, Narayanan PRD88 (2013) 105030



A Hamiltonian approach: TNS as ansatz for states

from K. Cichy

At finite T, thermal equilibrium states described by MPO



$$\rho_{th}(\beta) \propto e^{-\frac{\beta}{2}H}\mathbf{1}e^{-\frac{\beta}{2}H}$$ density operator

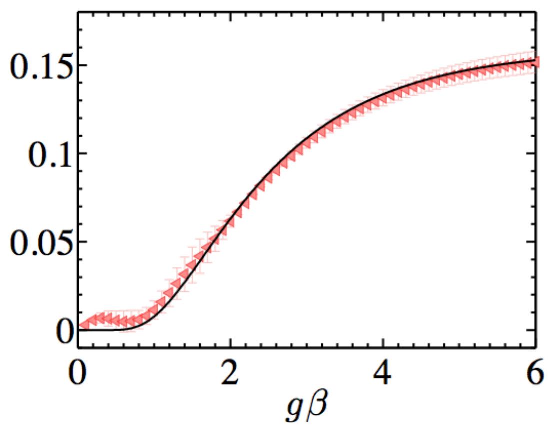
imaginary time (thermal) evolution can be simulated with MPO-MPS techniques

all error sources have been controlled

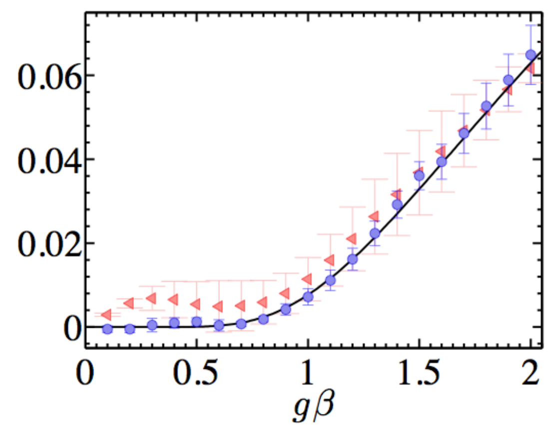
observables

$$\langle O\rangle_{th} \propto \mathrm{tr}(Oe^{-\frac{\beta}{2}H}\mathbf{1}e^{-\frac{\beta}{2}H})$$

e.g. thermal evolution of chiral condensate for $m/g=0$



full physical space:
no truncation of electric flux



truncating the electric flux per link
up to (converged) L_{cut}

Banuls, Cichy, Cirac, Jansen, Saito,
PoS LAT14, 302; Phys. Rev. D92, 034519;
PoS LAT15, 283; Phys. Rev. D93, 094512

Lagrangian approach

$\Downarrow$

Tensor Renormalization Group (TRG)

Levin & Nave PRL99,120601(2007)

Procedures of TRG

① Rewrite Z in tensor network representation

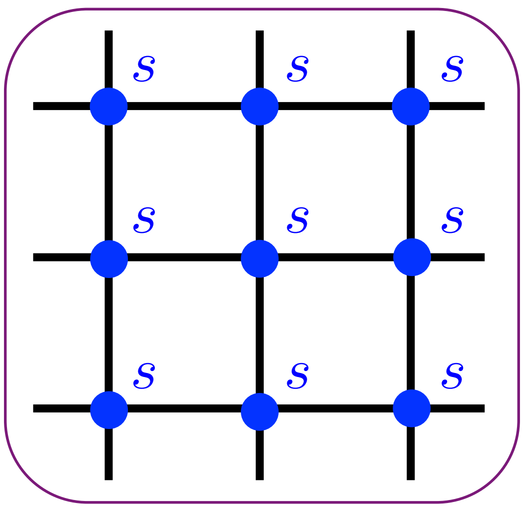
$$Z \equiv \sum_{\{s\}} e^{-\beta H[s]} = \sum_{i,j,k,l,\ldots} \ldots T_{ijkl} T_{mnio} \ldots$$

- Analytic
- exact

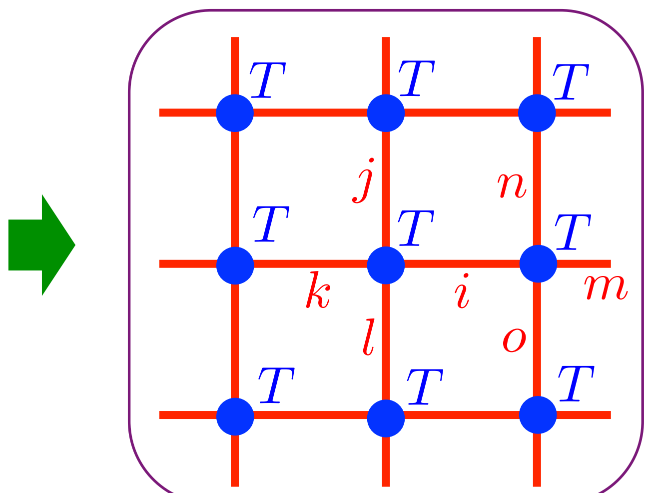
Spins

will be explained soon

New degrees of freedom



e.g. 2D Ising model

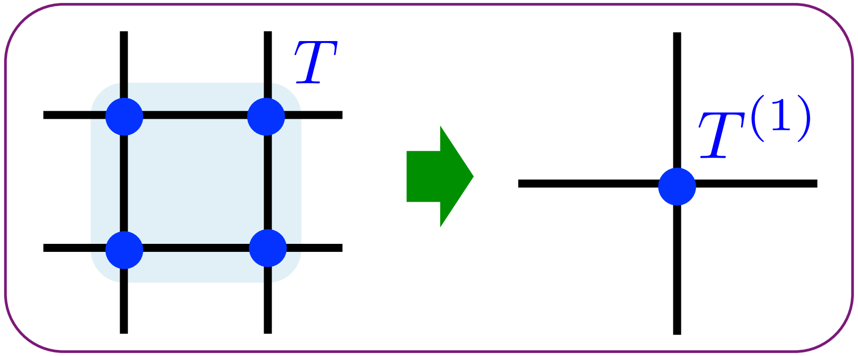


Procedures of TRG

① Rewrite Z in tensor network representation

② Coarse graining Tensor

Blocking of Tensor (like spin-blocking)



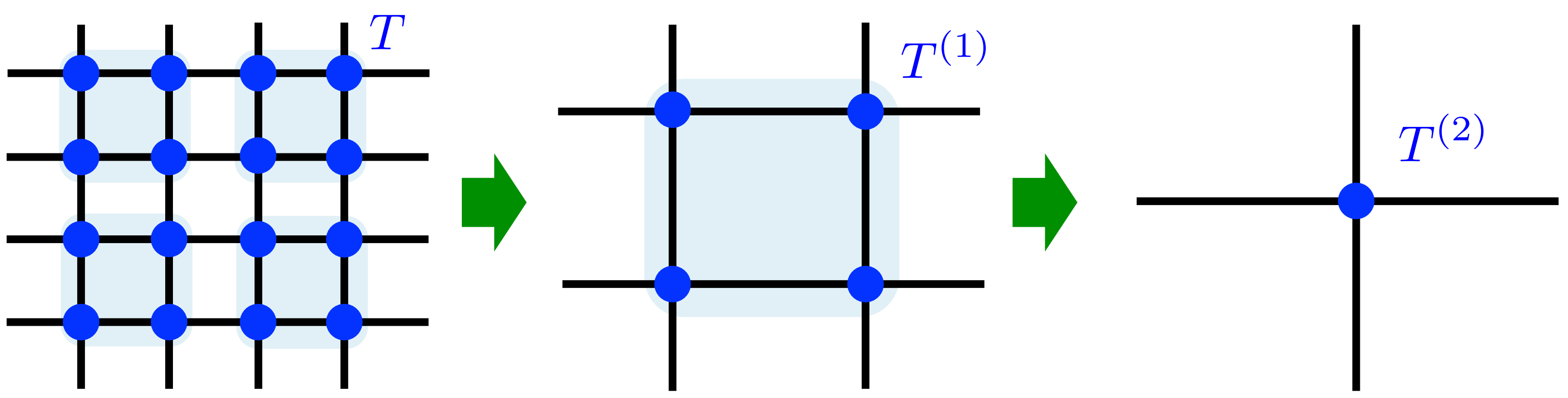
- extracting important information numerically
- selection of information introduces approximation

Procedures of TRG

① Rewrite Z in tensor network representation

② Coarse graining Tensor

③ Repeat the coarse graining and then reduce the number of tensors, finally compute Z by contraction



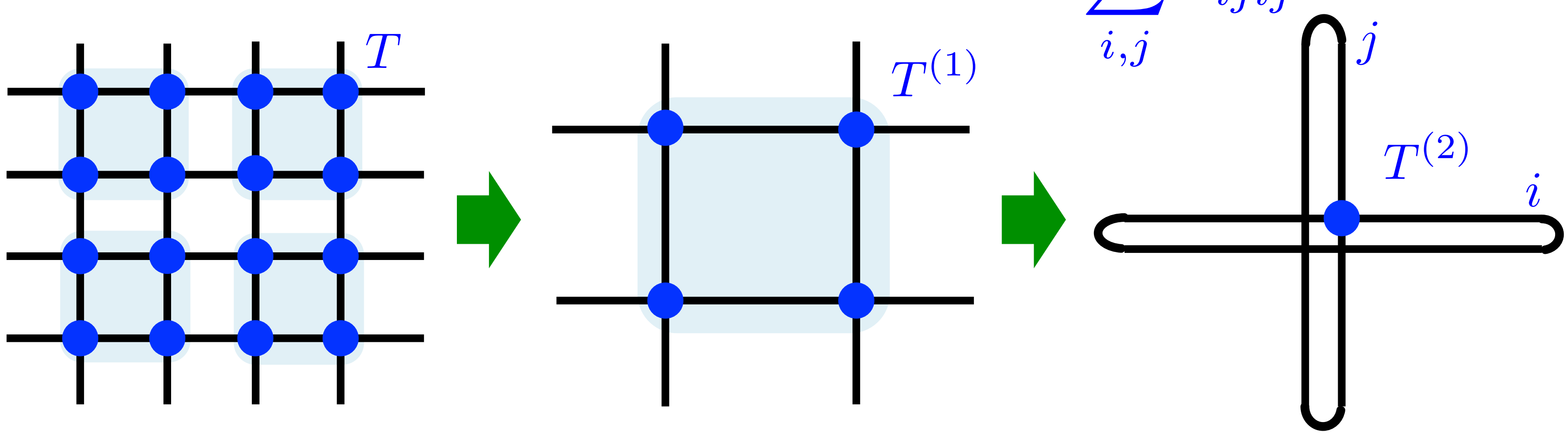
Procedures of TRG

① Rewrite Z in tensor network representation

② Coarse graining Tensor

③ Repeat the coarse graining and then reduce the number of tensors, finally compute Z by contraction

$$Z \approx \sum_{i,j} T^{(n)}_{ijij}$$ for Periodic BC



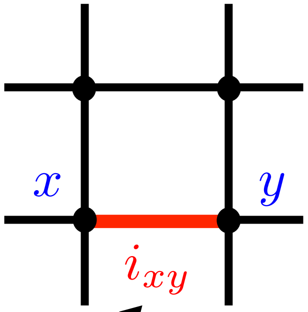
① Tensor network rep.

$$Z \equiv \sum_{\{s\}} e^{-\beta H[s]} = \sum_{i,j,k,l,\dots} \dots T_{ijkl} T_{mnio} \dots$$

1) Expand Boltzmann weight as in High-T expansion
2) Identify integer, which appears in the expansion, as new d.o.f. → index of tensor

$$\begin{aligned} e^{\beta s_x s_y} &= \cosh(\beta s_x s_y) + \sinh(\beta s_x s_y) \\ &= \cosh\beta + s_x s_y \sinh\beta \\ &= \cosh\beta(1 + s_x s_y \tanh\beta) \\ &= \cosh\beta \sum_{i_{xy}=0}^{1} (s_x s_y \tanh\beta)^{i_{xy}} \end{aligned}$$

$s_x = \pm 1$



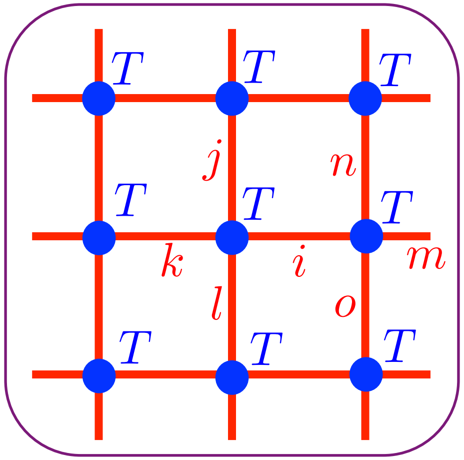
new d.o.f.

e.g. 2D Ising model

① Tensor network rep.

$$Z \equiv \sum_{\{s\}} e^{-\beta H[s]} = \sum_{i,j,k,l,\dots} \dots T_{ijkl} T_{mnio} \dots$$

1) Expand Boltzmann weight as in High-T expansion
2) Identify integer, which appears in the expansion, as new d.o.f. → index of tensor
3) Integrate out spin variable (old d.o.f.)
4) Get tensor network rep. !



$$\begin{bmatrix} T_{0000} & T_{0001} & T_{0010} & T_{0011} \\ T_{0100} & T_{0101} & T_{0110} & T_{0111} \\ T_{1000} & T_{1001} & T_{1010} & T_{1011} \\ T_{1100} & T_{1101} & T_{1110} & T_{1111} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \tanh\beta \\ 0 & \tanh\beta & \tanh\beta & 0 \\ 0 & \tanh\beta & \tanh\beta & 0 \\ \tanh\beta & 0 & 0 & (\tanh\beta)^2 \end{bmatrix}$$

$$\times 2(\cosh\beta)^2$$

e.g. 2D Ising model

① Tensor network rep.

$$Z \equiv \sum_{\{s\}} e^{-\beta H[s]} = \sum_{i,j,k,l,\ldots} \ldots T_{ijkl} T_{mnio} \ldots$$

1) Expand Boltzmann weight as in High-T expansion
2) Identify integer, which appears in the expansion, as new d.o.f. $\rightarrow$ index of tensor
3) Integrate out spin variable (old d.o.f.)
4) Get tensor network rep. !

Basic procedure is common to fermion/other boson system

Boson field in compact group ➞ character expansion ➞ character : new d.o.f.

Meurice et al., PRD88,056005(2013)

① Tensor network rep.

$$\mathcal{Z} = \sum_{..,i,j,k,l,m,n,o,..} \cdots T_{ijkl} T_{mnio} \cdots$$

- So far, we have just rewritten Z
- Next step is to carry out the summation
- But, naïve approach costs ∝ 2^{2V}
- Introduce approximation and reduce the cost while keeping an efficiency by summing important part in Z

① Tensor network rep.

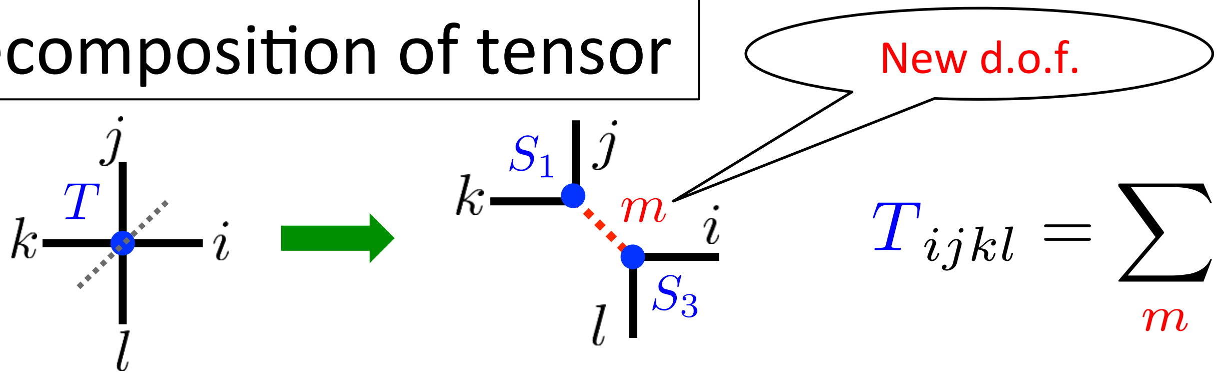
$$\mathcal{Z} = \sum_{..,i,j,k,l,m,n,o,..} \cdots T_{ijkl} T_{mnio} \cdots$$

- So far, we have just rewritten Z
- Next step is to carry out the summation
- But, naïve approach costs ∝ 2^{2V}
- Introduce approximation and reduce the cost while keeping an efficiency by summing important part in Z

Coarse graining (renormalization, blocking)

②Coarse graining

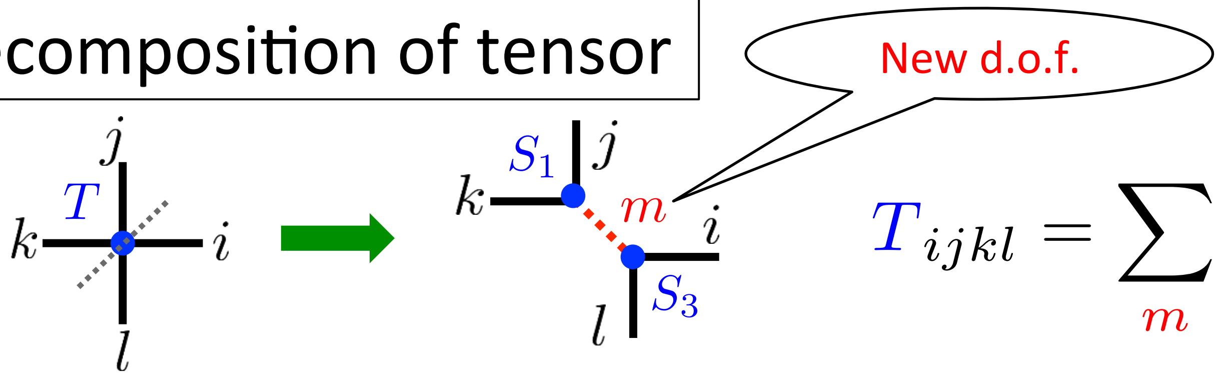
Decomposition of tensor



$$T_{ijkl} = \sum_m (S_1)_{jkm}(S_3)_{lim}$$

② Coarse graining

Decomposition of tensor



$$T_{ijkl} = \sum_m (S_1)_{jkm} (S_3)_{lim}$$

Singular value decomposition (SVD)

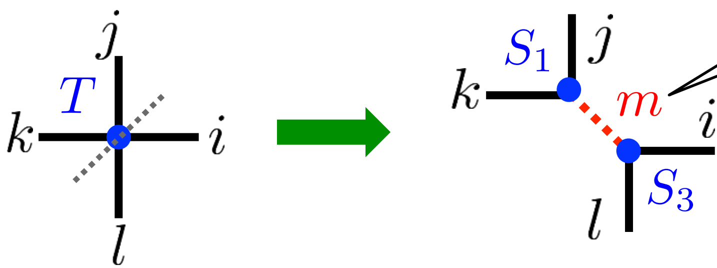
u, v : unitary matrix

$$M_{ab} = \sum_m u_{am} \sigma_m (v^\dagger)_{mb}$$

$\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$: singular values

②Coarse graining

Decomposition of tensor



$$T_{ijkl} = \sum_m (S_1)_{jkm} (S_3)_{lim}$$

Singular value decomposition (SVD)

u, v : unitary matrix

$$M_{ab} = \sum_m u_{am} \sigma_m (v^\dagger)_{mb}$$

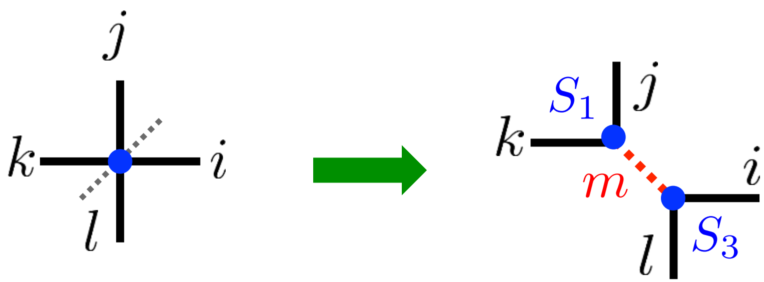
$\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$: singular values

approx.

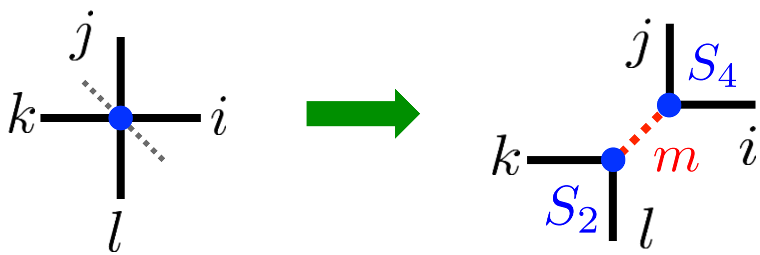
$$T_{ijkl} = M_{(kj),(il)} = \sum_m^{\text{all}} \underline{u_{(kj),m}\sqrt{\sigma_m}} \cdot \underline{\sqrt{\sigma_m} v^\dagger_{m,(il)}} \approx \sum_{m=1}^{D_{\text{cut}}} \underline{(S_1)_{jkm}} \underline{(S_3)_{lim}}$$

Tensor (matrix) is approximated by low-rank tensor = information compression

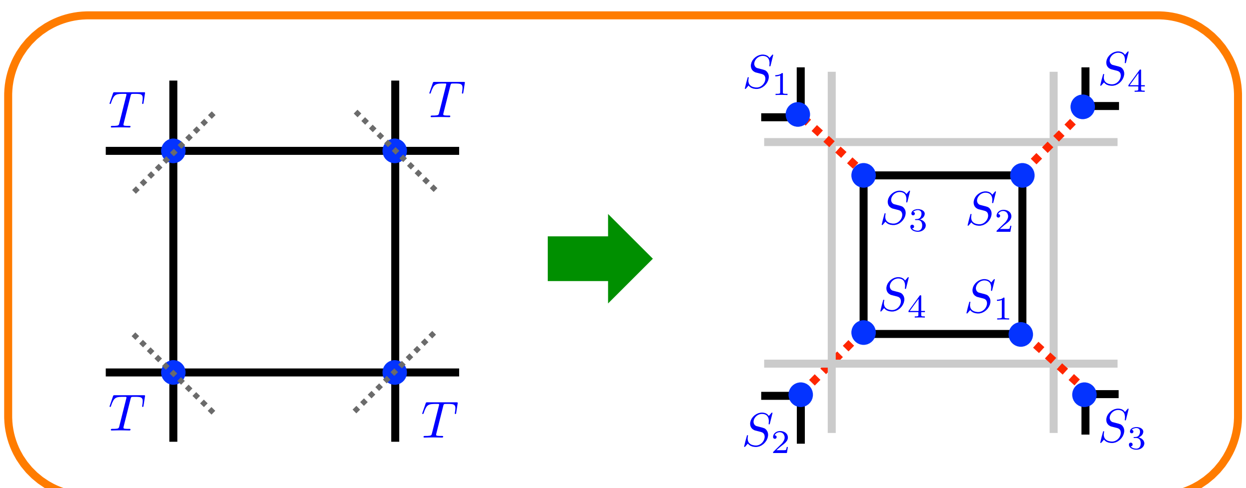
② Coarse graining



$$T_{ijkl} \approx \sum_{m=1}^{D_{\mathrm{cut}}} (S_1)_{jkm} (S_3)_{lim}$$



$$T_{ijkl} \approx \sum_{m=1}^{D_{\mathrm{cut}}} (S_2)_{klm} (S_4)_{ijm}$$



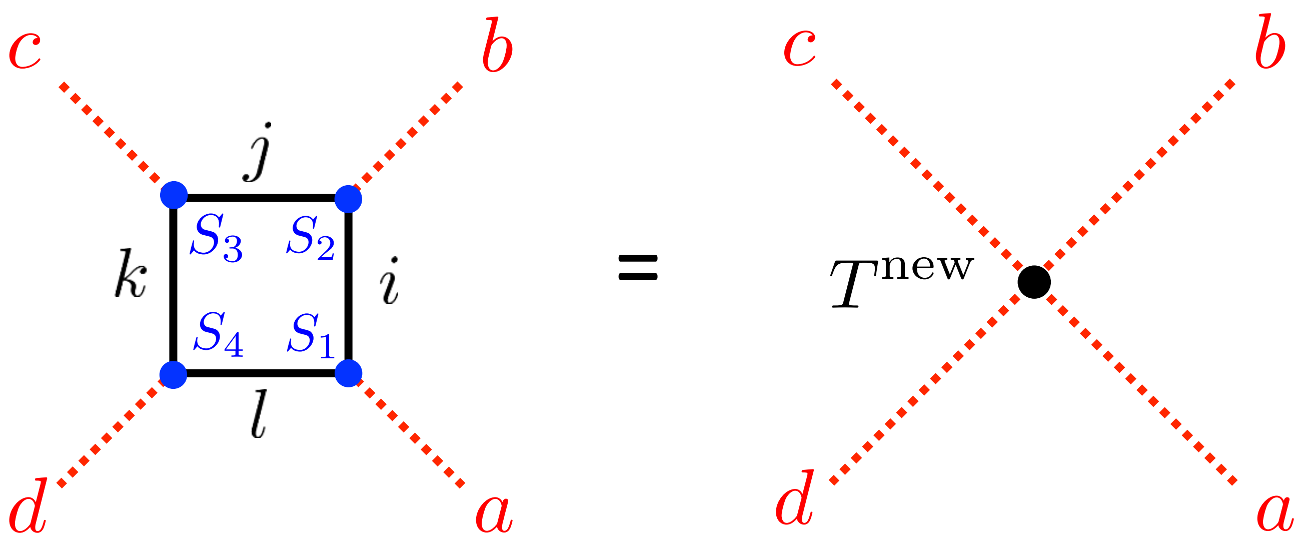
② Coarse graining

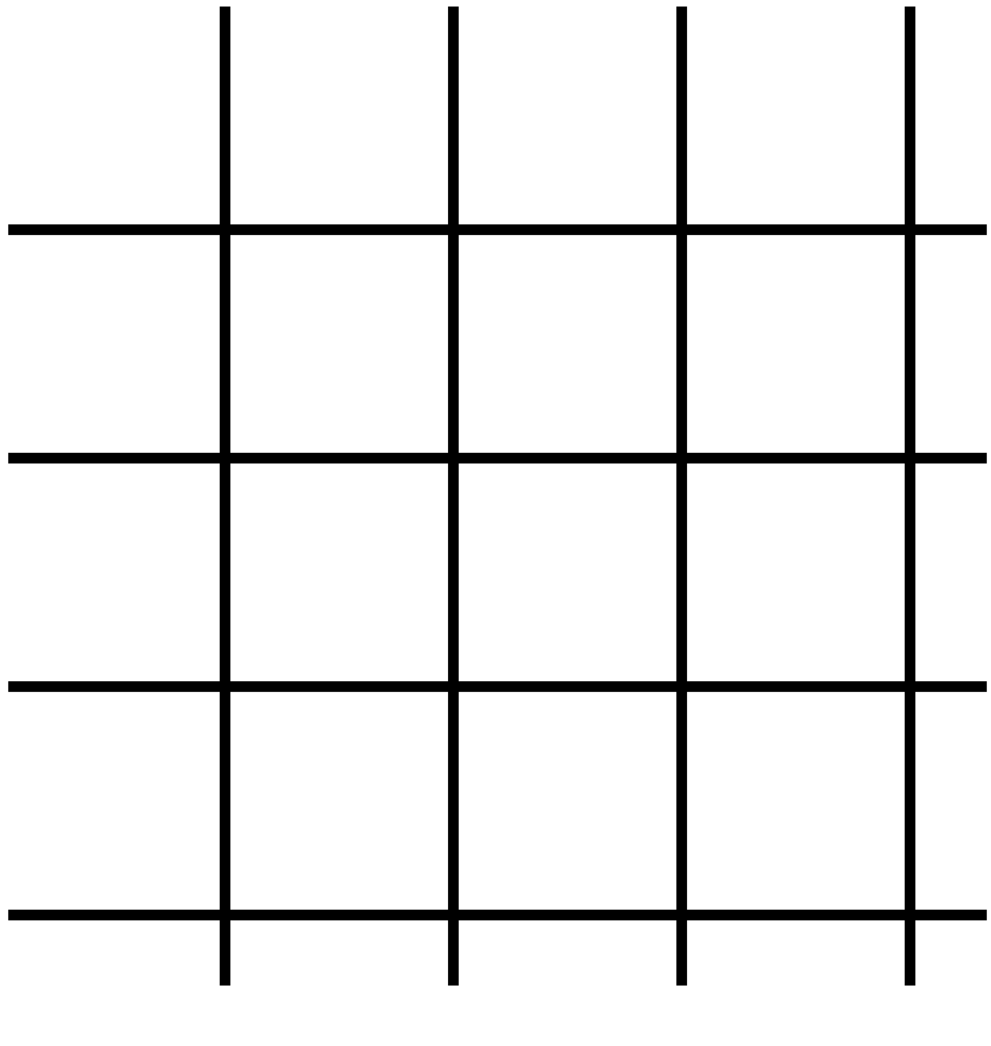
Making new tensor by contraction

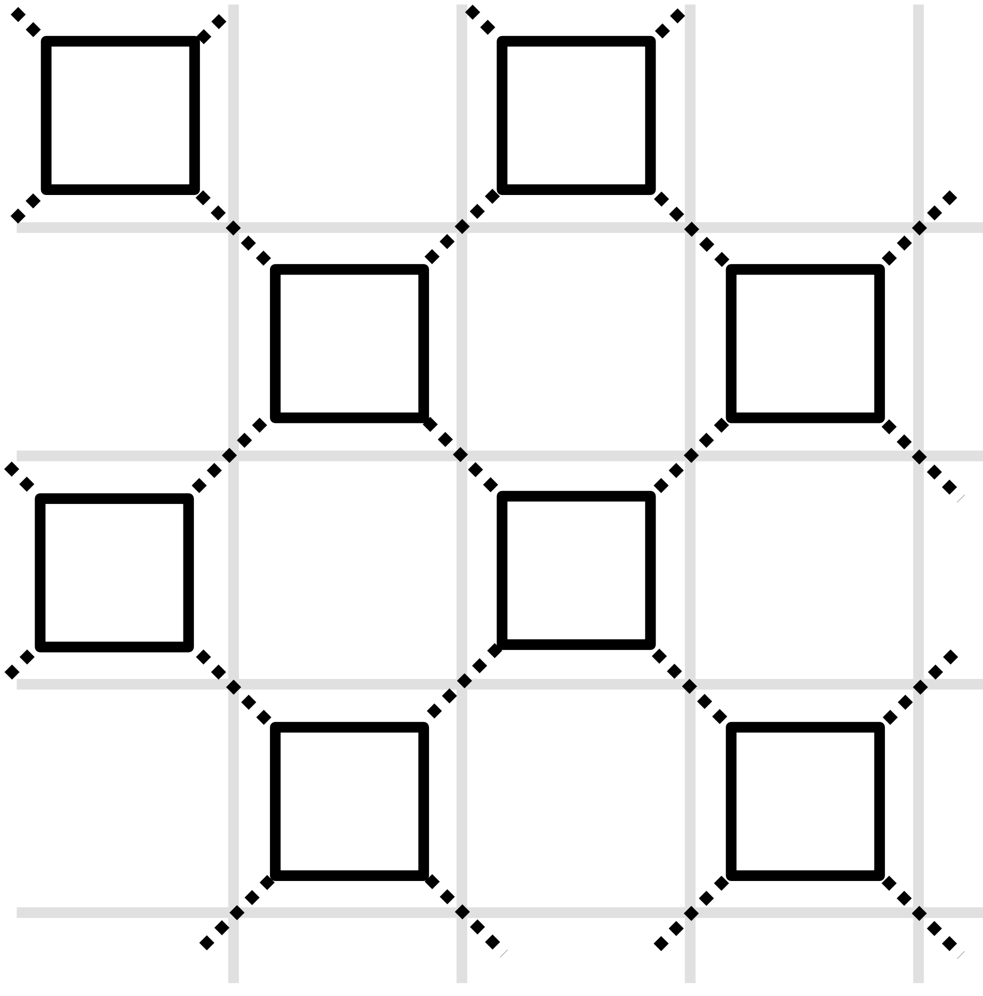
$$\sum_{i,j,k,l}^{\text{all}} (S_1)_{ila}(S_2)_{jib}(S_3)_{kjc}(S_4)_{lkd} = T^{\text{new}}_{abcd}$$

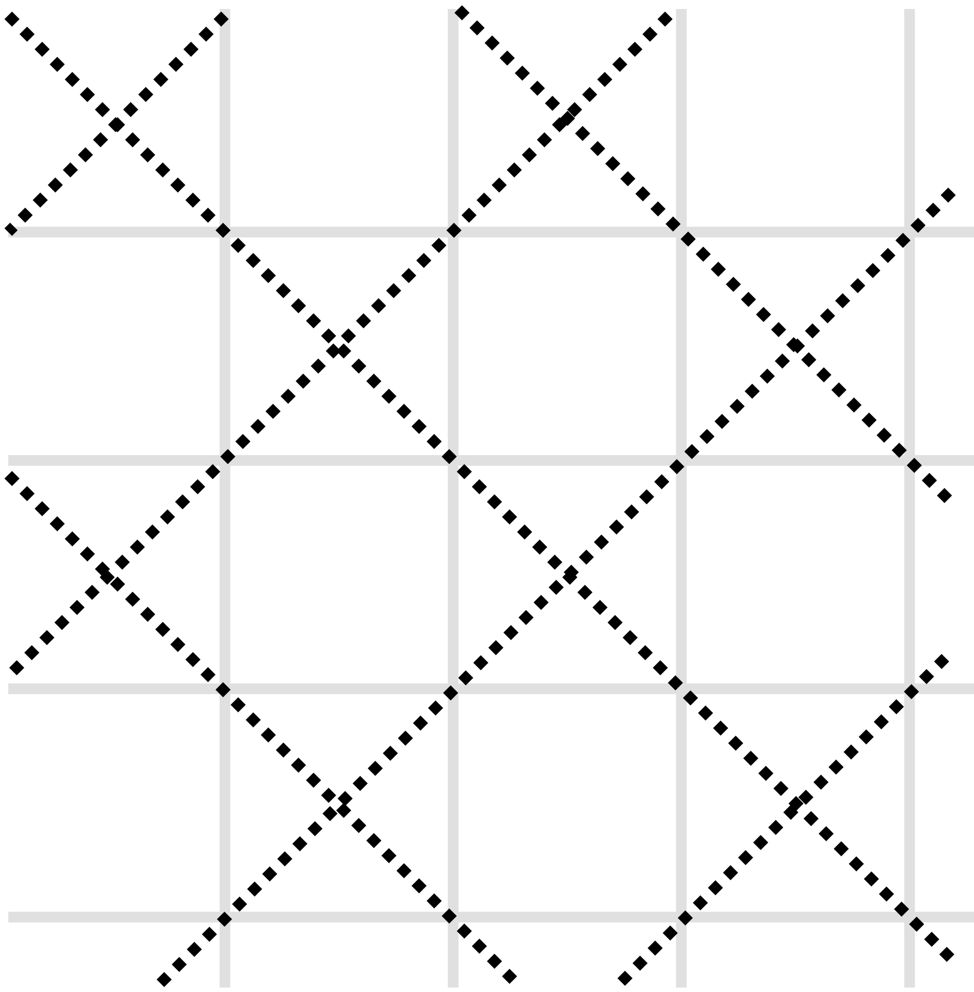
integrate out old d.o.f.

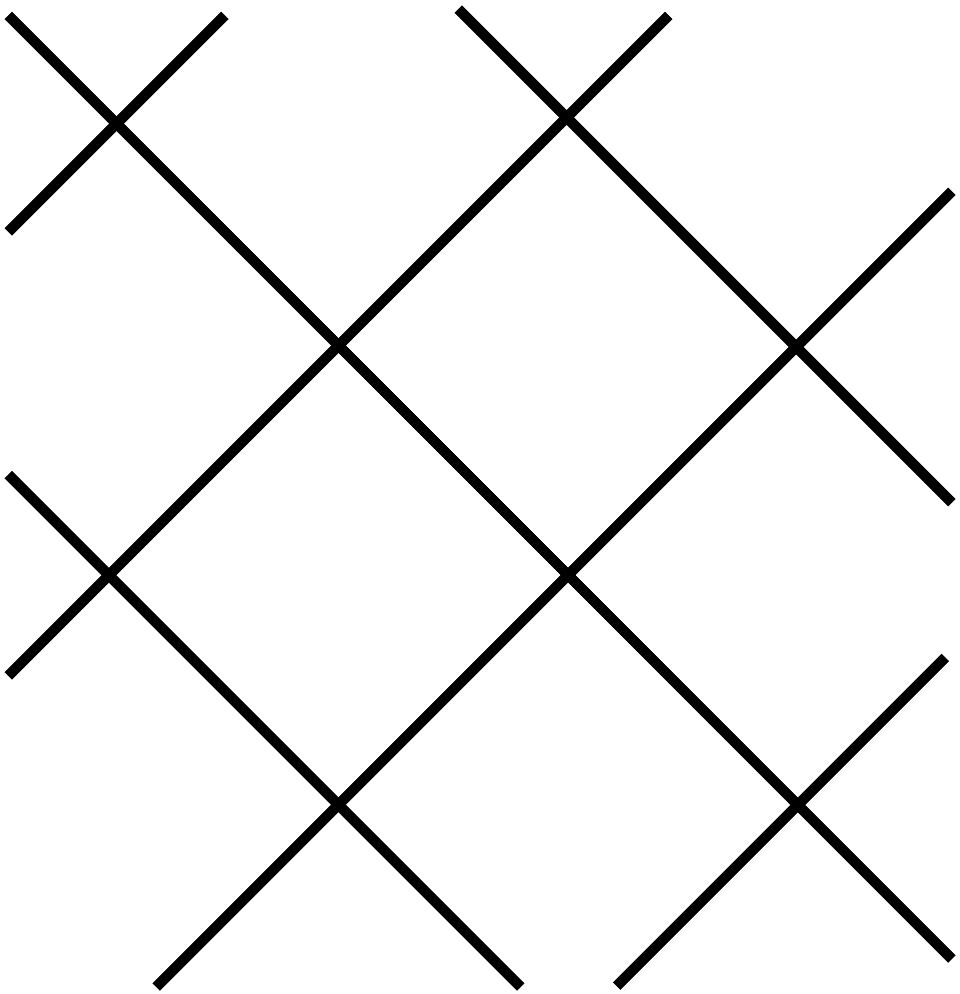
Renormalization-like!

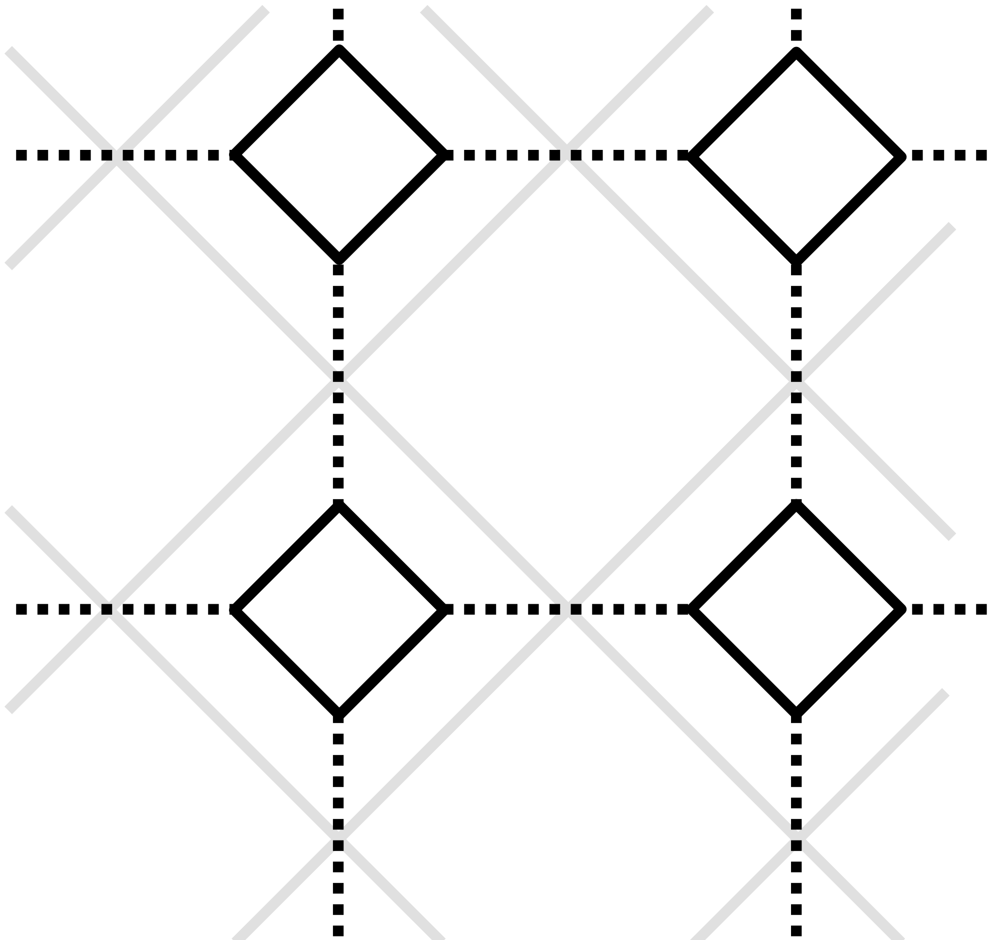


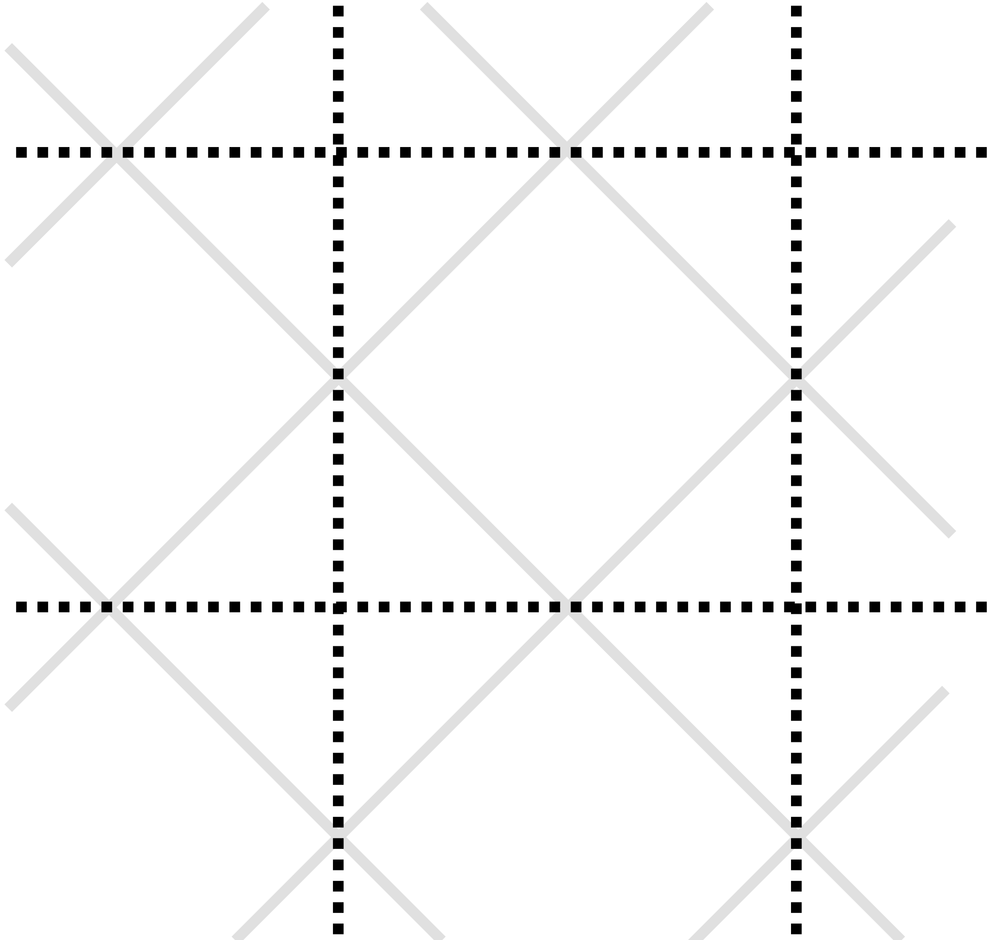


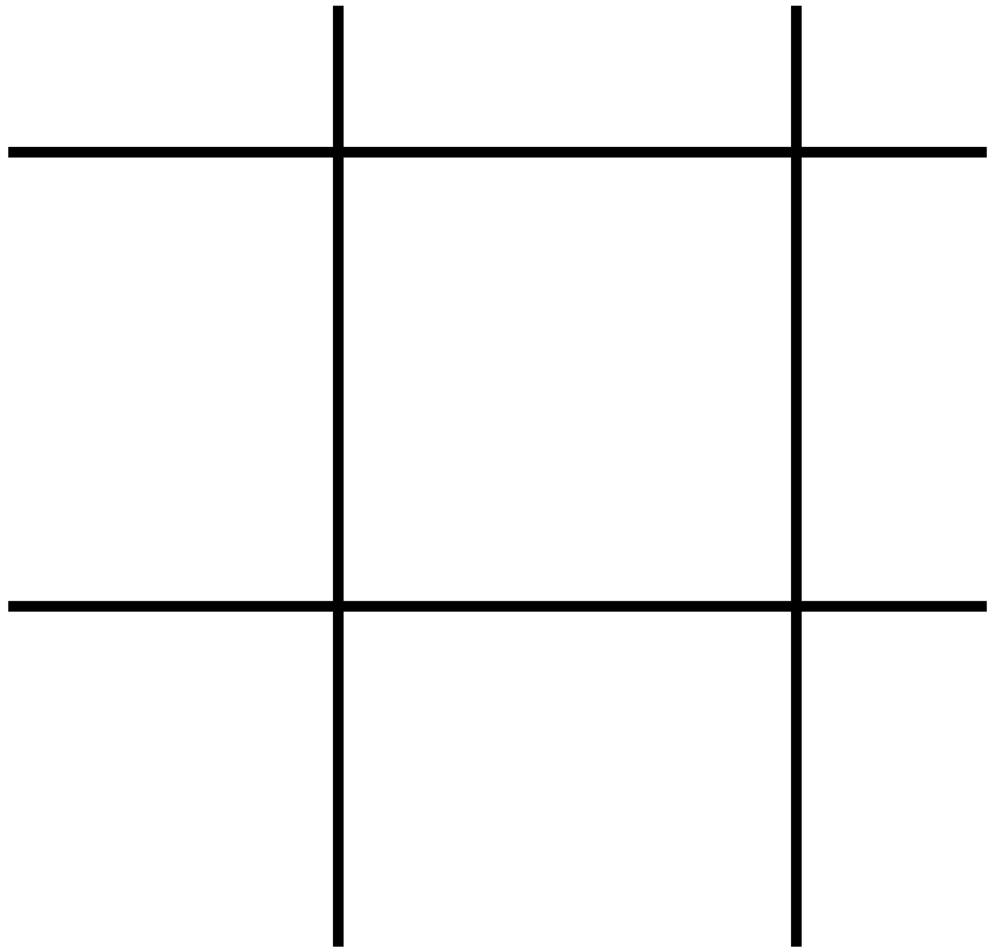


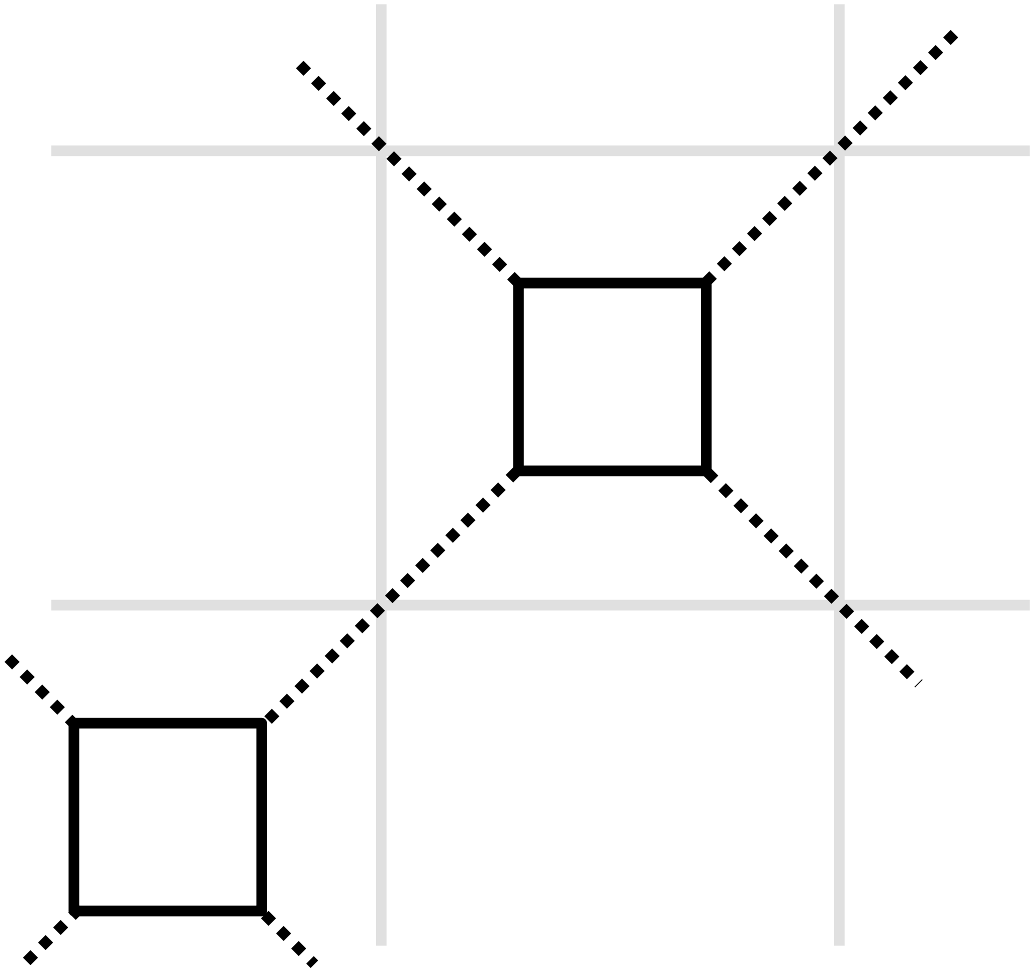


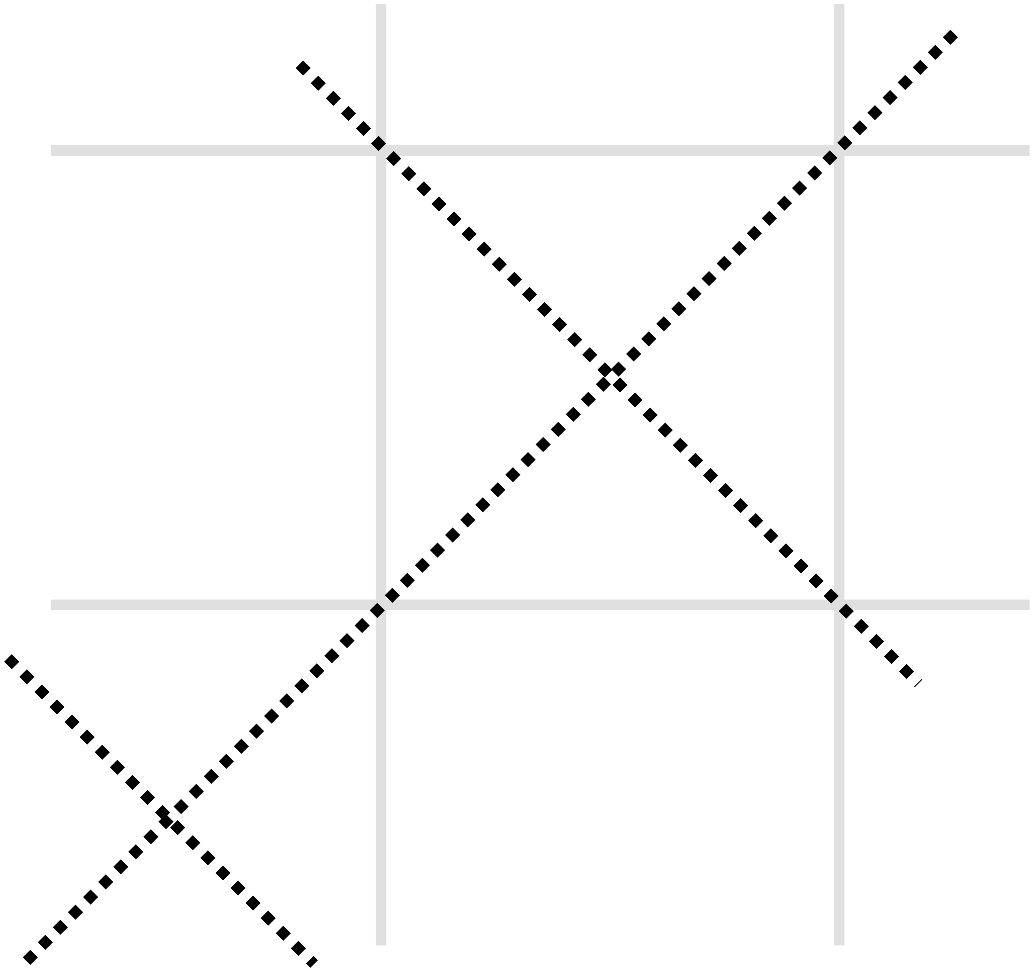


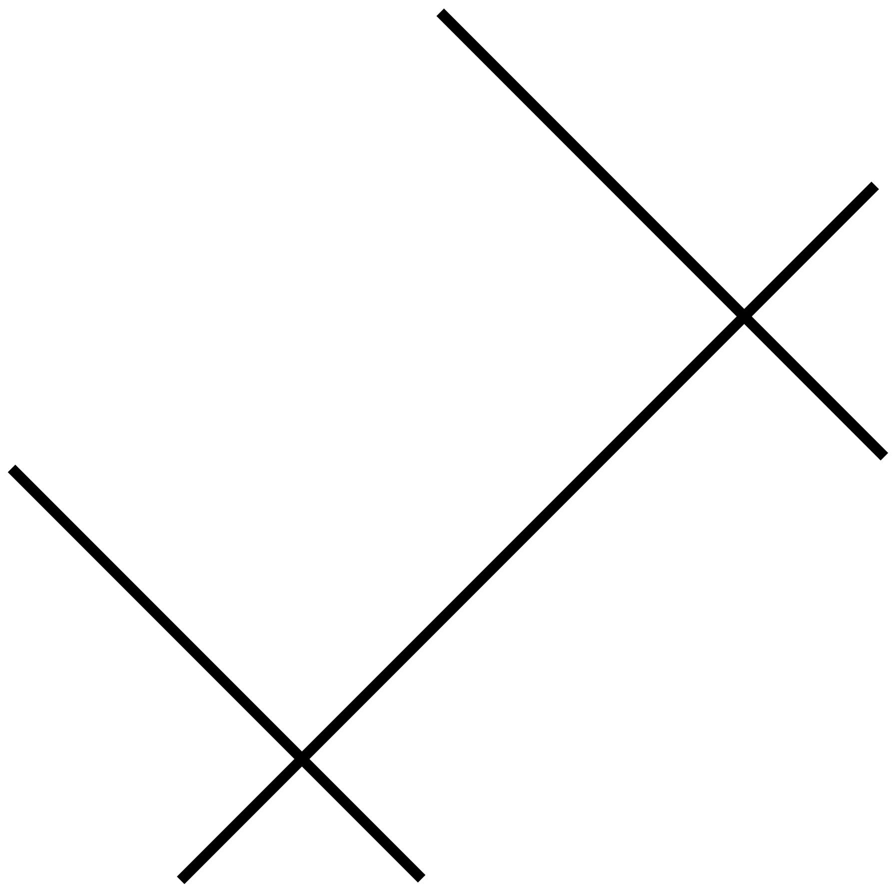


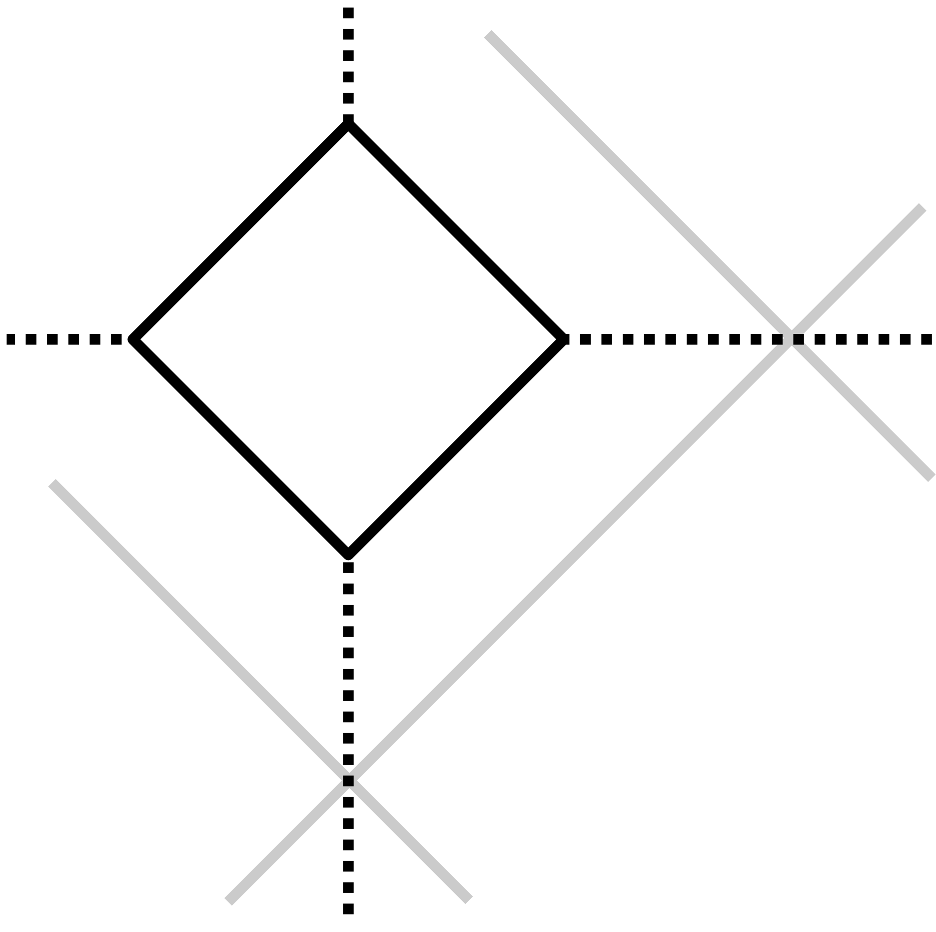


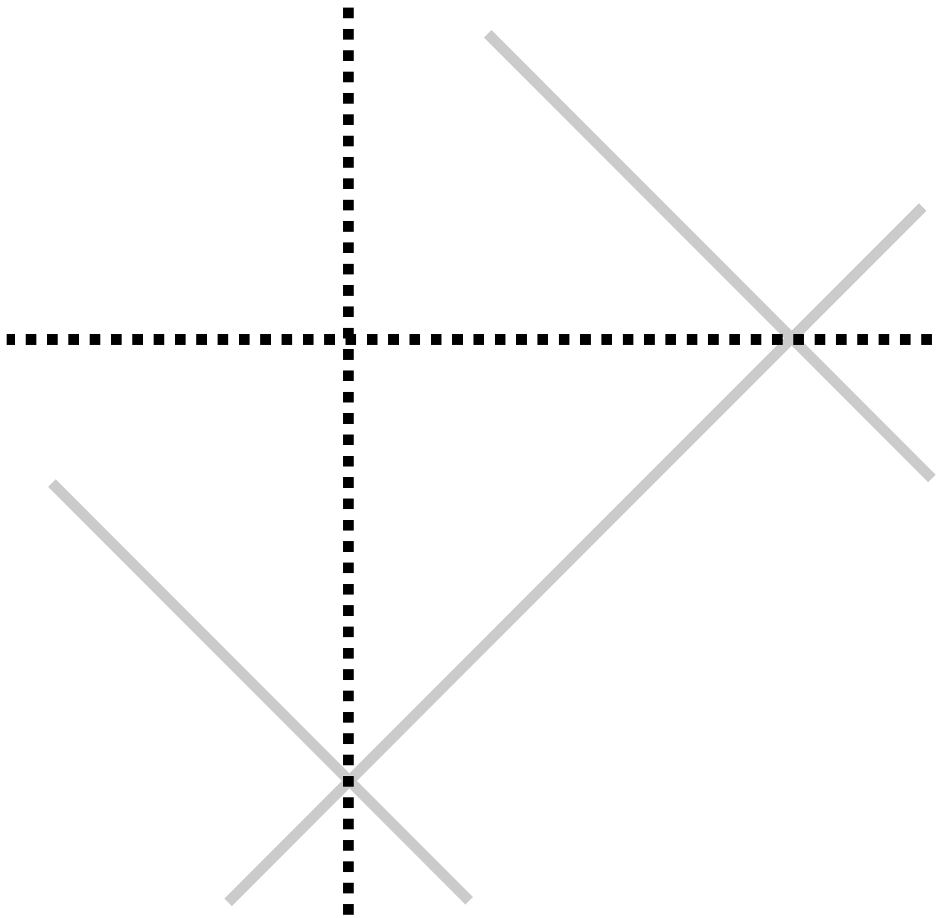


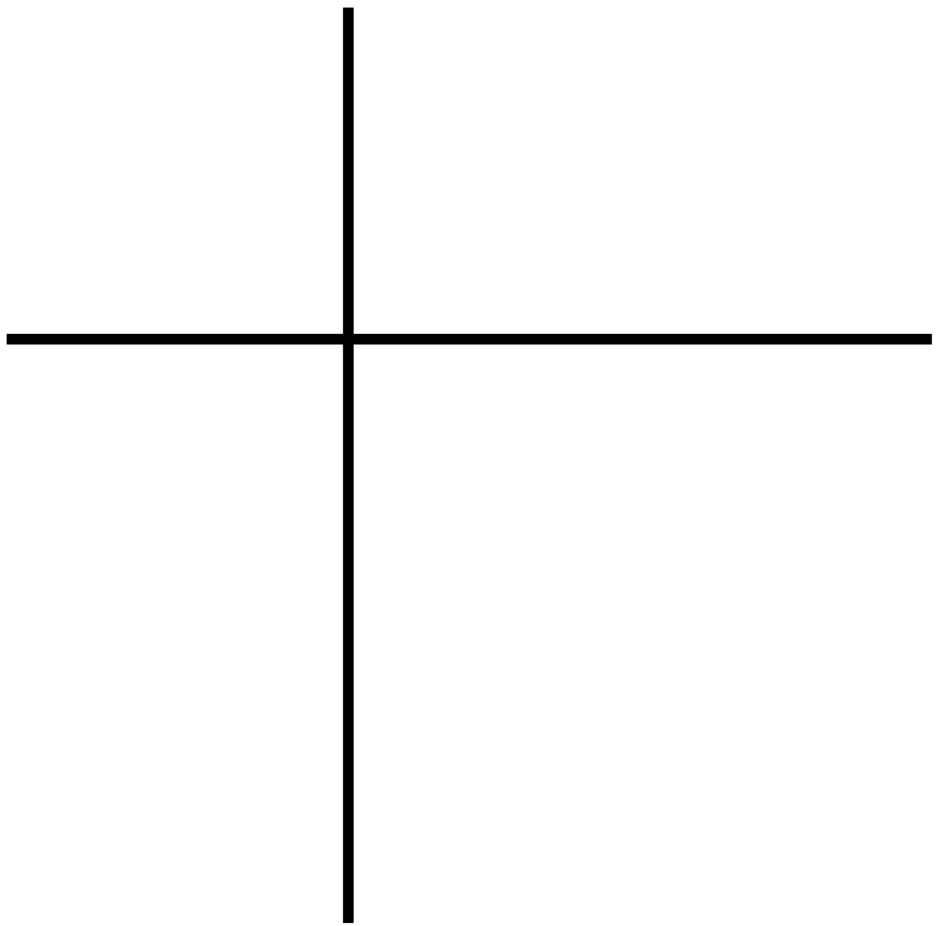


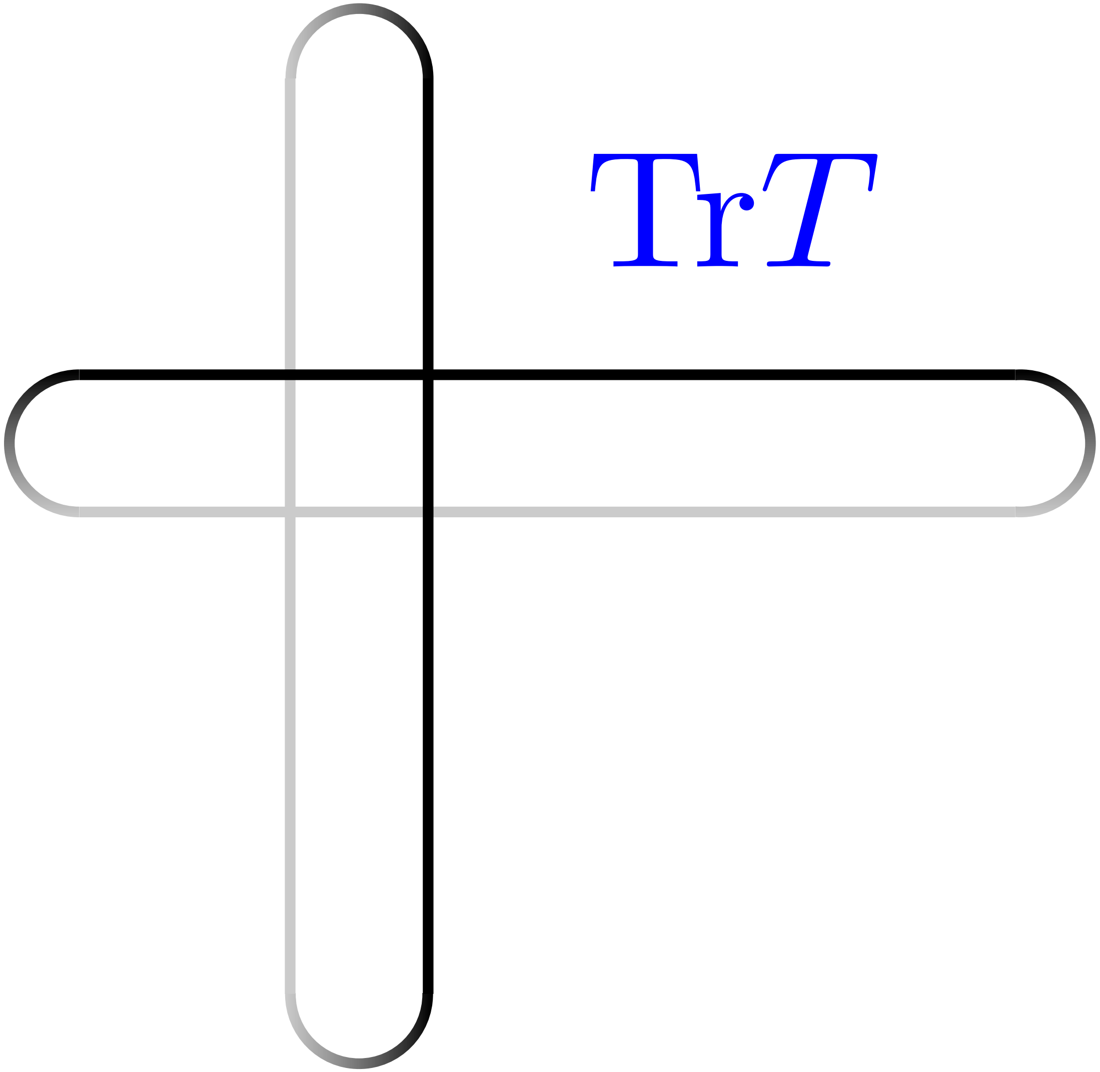




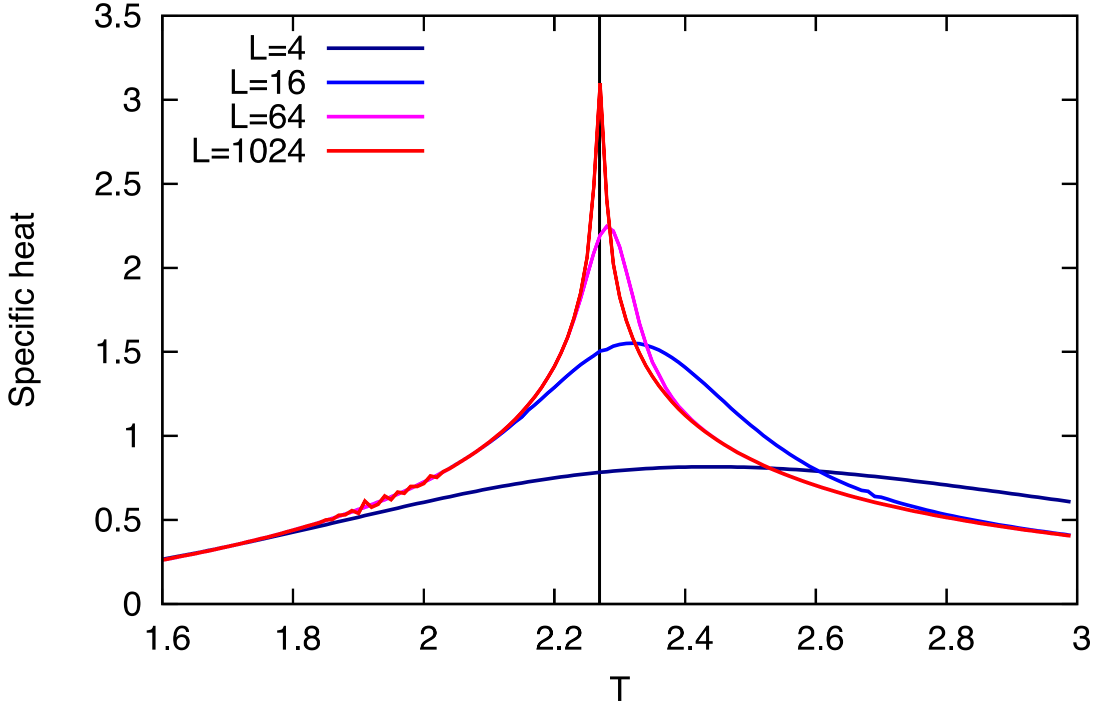







TrT

2D Ising model on square lattice



$$C = \frac{1}{L^2}\frac{\partial}{\partial T}\left(T^2 \frac{\partial \ln \mathcal{Z}}{\partial T}\right)$$

numerical derivative

D_{cut}=32

$$\begin{aligned} T_c &= 2/[\ln(1+\sqrt{2})] \\ &= 2.269... \end{aligned}$$

only one day use of this MBA

Cost ∝ log(Lattice volume) × $(D_{cut})^6$ × [# temperature mesh]

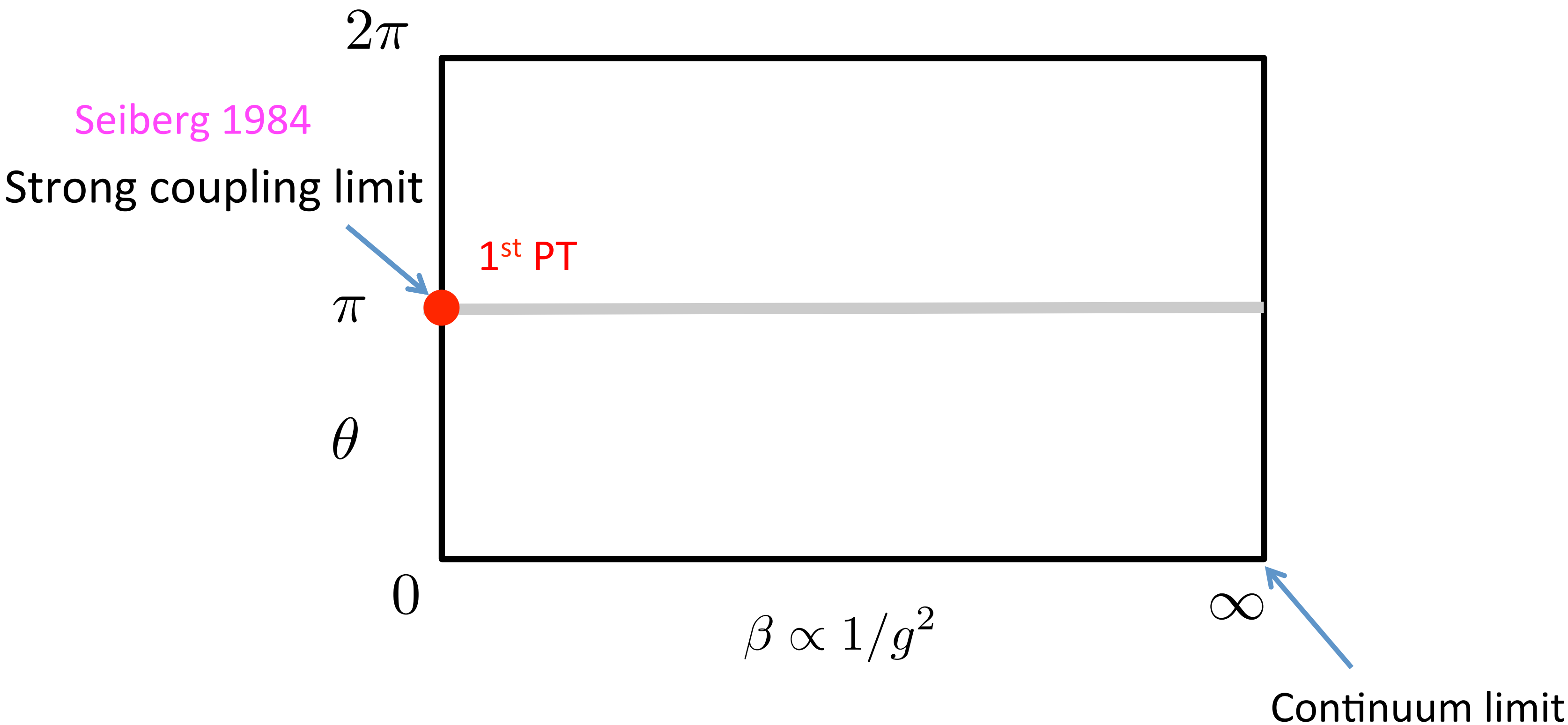
Status of Lagrangian approach

- 2D system
  - Spin model : Ising model Levin & Nave 2007, X-Y model Meurice et al. PRE89,013308(2014), X-Y model with Fisher zero Meurice et al. PRD89,016008(2014), O(3) model Unmuth-Yockey et al. LATTICE2014, X-Y model + μ Meurice et al. PRE93,012138(2016)
  - Abelian-Higgs Bazavov et al. LATTICE2015
  - ϕ^4 theory Shimizu Mod.Phys.Lett.A27,1250035(2012)
  - QED_2, $QED_2 + \theta$ Shimizu & Kuramashi PRD90,014508(2014) & PRD90,074503(2014)
  - Gross-Neveu model + μ ST & Yoshimura PTEP2015,043B01
  - CP(N-1) + θ Kawauchi & ST PRD93,114503(2016), Kawauchi 7/25(Mon)17:05 TD
  - Towards Quantum simulation of O(2) model Zou et al, PRA90,063603
  - TRG and quantum simulation Meurice 7/26(Tue.)15:40 TD
- 3D system, Higher order TRG(HOTRG) : a coarse graining method applicable for any dimensional system Xie et al. PRB86,045139(2012)
  - 3D Ising, Potts model Wan et al. CPL31,070503(2014)
  - 3D Fermion system Sakai 7/26(Tue.)18:10 TD
- Decorated tensor network renormalization Wittrich et al. New J.Phys. 18,053009(2016)

Application to CP(N-1) + θ

Toy model of QCD

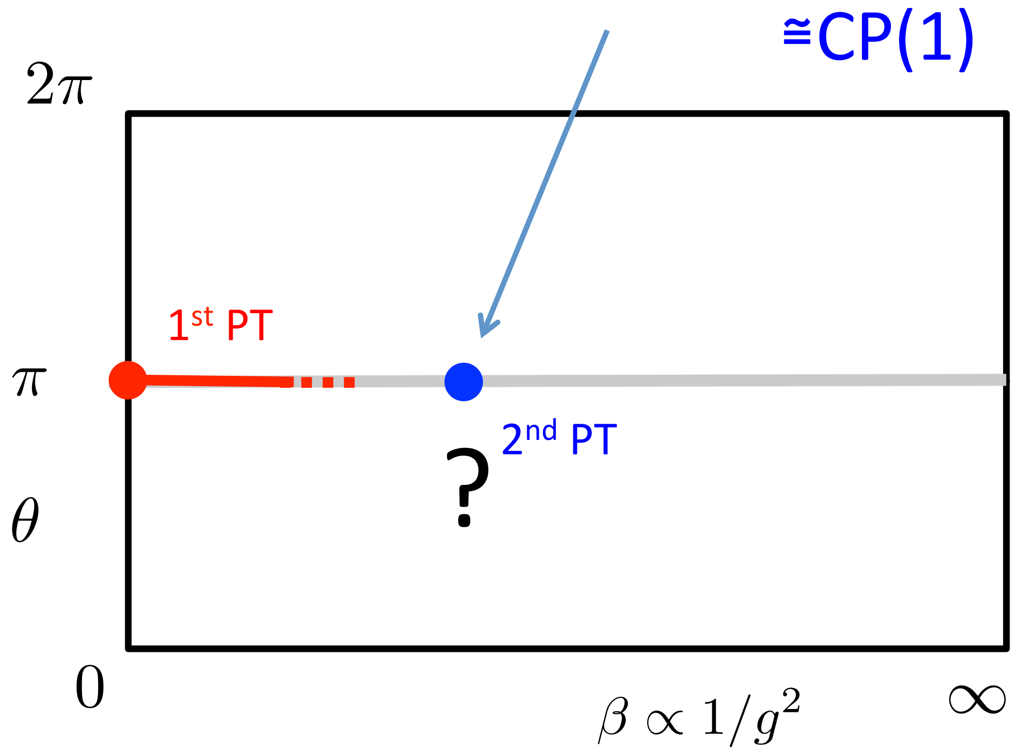
Strong CP problem Schierholtz 1994, Plefka et al. 1997, Imachi et al. 1999, Plefka et al. 1999



Application to CP(1) + θ

Haldane's conjecture : mass gap of O(3) vanishes at θ=π

Haldane 1983



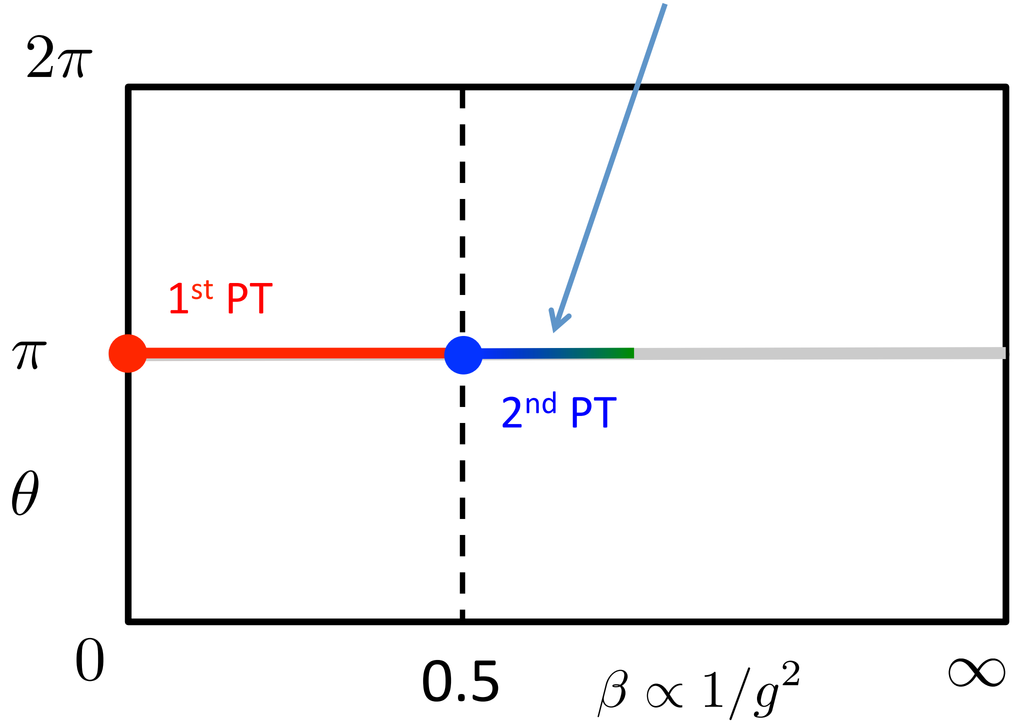
O(3) + θ was intensively studied by MC and Haldane's conjecture is confirmed

Universality class is also consistent with each other

Bietenholtz et al. 1995, Wiese et al. 2012, de Forcrand et al. 2012, Azcoiti et al. 2012, Alles et al. 2014

Application to CP(1) + θ

Azcoiti et al, PRL98,257203(2007)   CP(1) MC, imaginary θ



However, the universality class is not fixed to the expected one (k=1 WZNW model) and critical exponent is continuously changing

Wess & Zumino 1971, Novikov 1981, Witten 1984

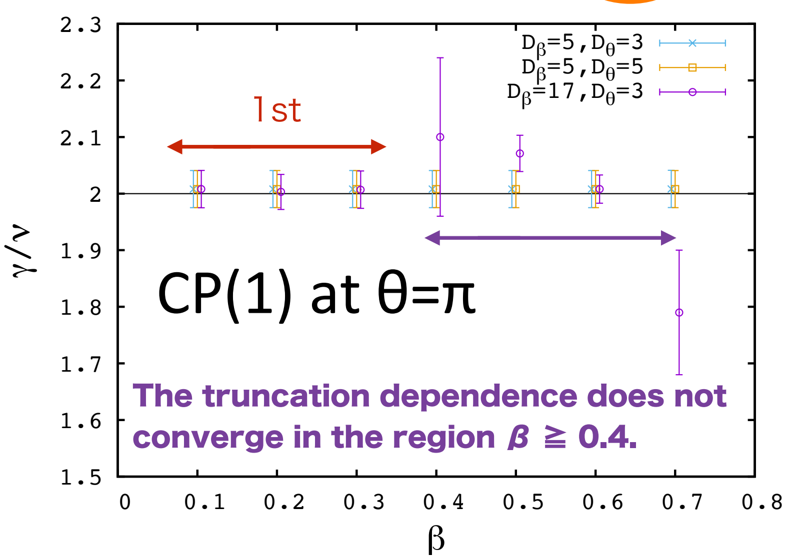
Application to CP(1) + θ

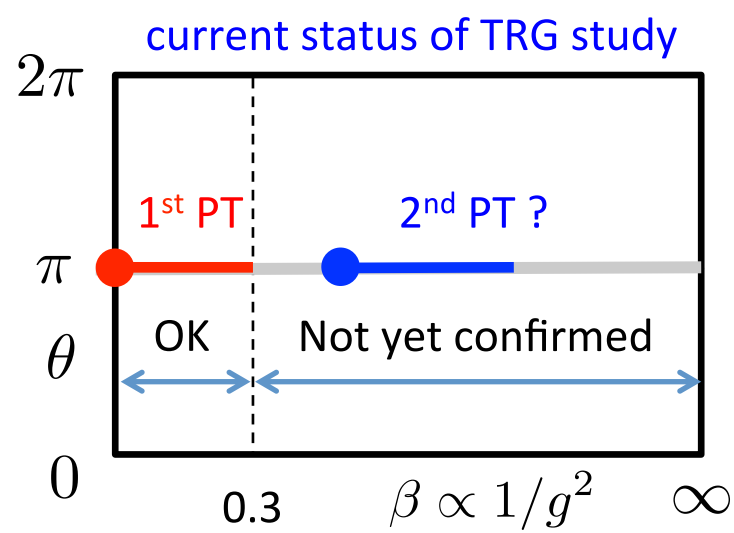
Kawauchi 7/25(Mon)17:05 TD

Tensor network rep. is obtained thanks to character-like expansion
Plefka et al. 1997
TRG is used as coarse graining

Truncation dependence of $\chi_{\max} \propto L^{\frac{\gamma}{\nu}}$

topological susceptibility





TRG does not work well near critical region
⇒ Tensor Network Renormalization should be used in future

Evenbly & Vidal 2014, Gu et al., 2015

MC	TRG
Boltzmann weight is interpreted as probability	Tensor network rep. of partition function (no probability interpretation)
Importance sampling	Compression of tensor by SVD, Variational approach
Statistical errors	Systematic errors
Sign problem may appear	No sign problem ∵ no probability
Critical slowing down	Efficiency of compression gets worse around criticality

can be improved by TNR, Loop-TNR in 2D system

Evenbly & Vidal 2014, Gu et al., 2015

Summary

- Tensor network has No sign problem
- Key of point: information compression based on SVD
- Various 1D/2D systems are under investigation in Hamiltonian/Lagrangian approach
- Higher dimensional system is still hard…

Future prospects

- Long way to go 4D QCD + μ & θ
  - Cost : $O(D_{cut}^{15})$, Memory : $O(D_{cut}^{8})$ for 4D system (HOTRG)
  - Non-Abelian gauge theory

    Character expansion $\Rightarrow$ Tensor network rep. is OK but internal d.o.f. is huge!!!
  - Better coarse graining method in 4D? (TNR, Loop-TNR)
  - Efficient parallelization?
- Low dim. system suffering from the sign problem ?
  - Lattice SUSY, Lattice chiral gauge theory,….
- Combining with stochastic method to reduce cost

Good and Bad points

Good

- Free from the sign problem
- Large volume is easy to do

Bad

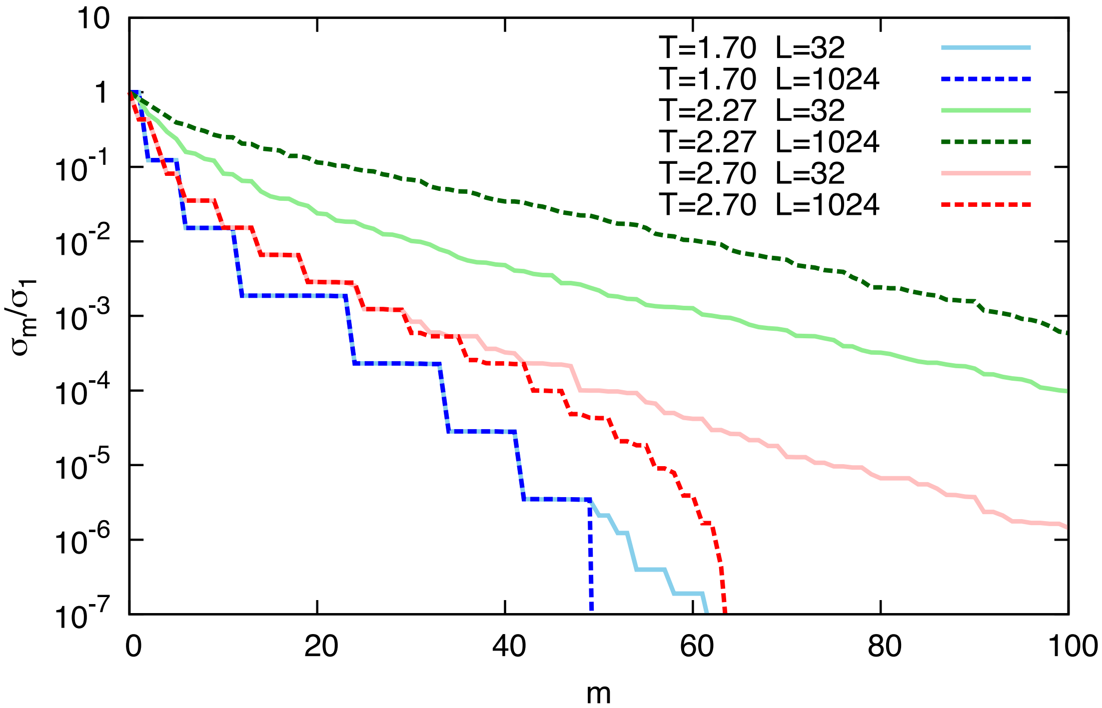
- Higher dimensionality is hard (2, 3 is OK but 4 is hard)
- Larger # of internal degree of freedom is also hard

$$\text{Cost} \propto \log(\text{Volume}) \times (D_{\text{cut}})^{\text{const.} \times \text{Dim.} \times \#\text{DOF}}$$

“Good” algorithm can reduce it

Hierarchy of singular value

2D Ising model



D_{cut}=32

$$T_c = 2/[\ln(1+\sqrt{2})] = 2.269\ldots$$

Entanglement entropy

$$S = -\sum_i \tilde{\sigma}_i \ln \tilde{\sigma}_i$$

normalized singular value

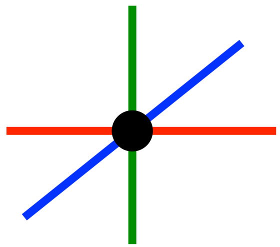
- Off criticality: good hierarchy (small S)
- Near criticality: hierarchy gets worse (large S)

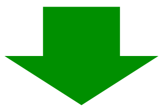
like critical slowing down in MC

Tensor network renormalization (TNR) Evenbly&Vidal 2014 can cure the situation

Numerical aspect of TRG and Task

- Main computation（For HOTRG, n-dim system）
  - Decomposition ⇒ SVD(EVD) : $O(D_{cut}^{6})$
  - Contraction ⇒ matrix-matrix product : $O(D_{cut}^{4n-1})$ Hot spot
- Memory
  - # elements of tensor : $O(D_{cut}^{2n})$
  - internal d.o.f. ⇒ more memory



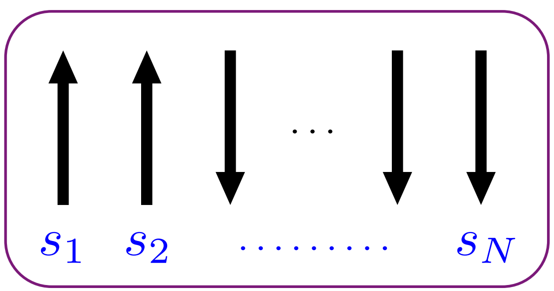


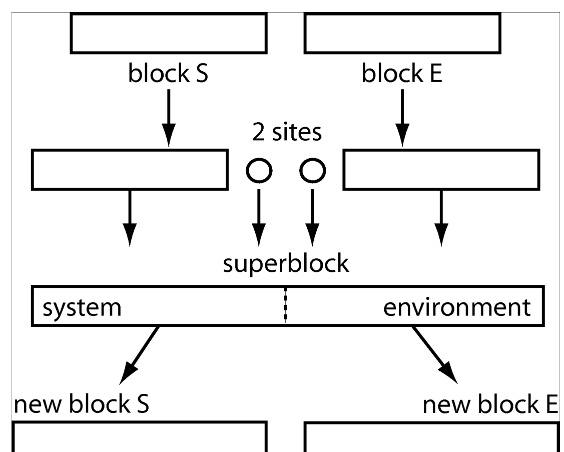
- matrix-matrix product : Level 3 BLAS ≫ SVD
- Better coarse graining with small D_{cut} (highly compression)?

Hamiltonian approach

Density matrix renormalization group (DMRG) White 1992 Schollwoeck 2004

- 1D quantum gapped system (e.g. N sites system)
- Target: Wave function
- Efficient way of choosing GOOD basis by using Schmit decomposition (like SVD)
- small # of basis = information compression
- Before this appears, limited to N=30. But DMRG enables N=100





Improvement of Coarse graining algorithm

- Tensor Entanglement Filtering Renormalization Gu et al. 2008
  - Removing short range correlation (partially)
  - works in off-critical point but not near criticality
- Second TRG Xie et al., 2009
  - Optimization including environment (TRG: locally optimal)
  - works in off-critical point but not near criticality
- Tensor Network Renormalization (TNR) Evenbly & Vidal 2014
  - Firstly remove short correlation (entanglement) by using disentangler, and then coarse graining is performed
  - Even around criticality, sustainable coarse graining is realized
- Loop Tensor Network renormalization (Loop TNR) Gu et al., 2015
  - Robustness around the criticality is comparable with TNR
  - Less cost $O(D_{cut}^6)$ while TNR costs $O(D_{cut}^7)$

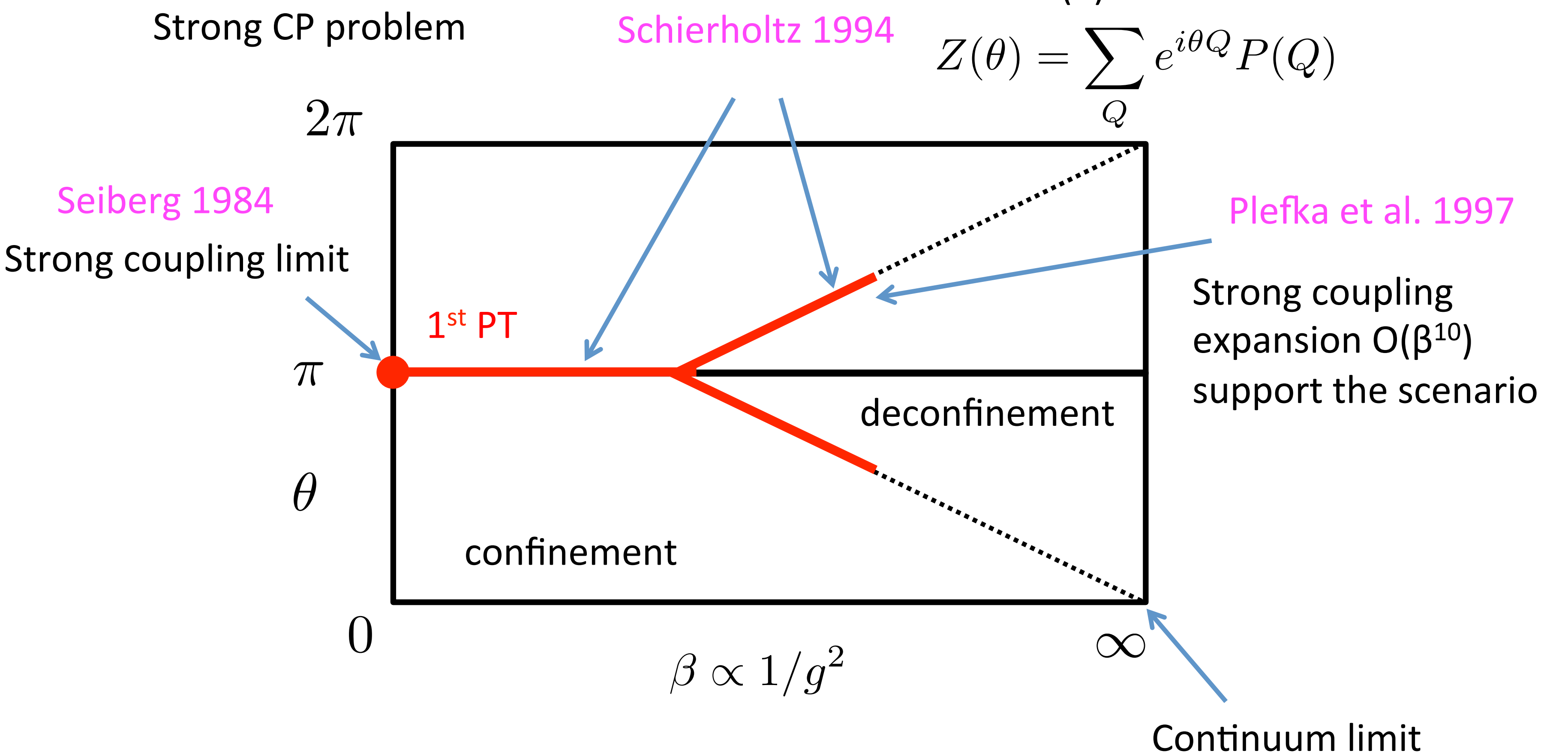
Application to CP(N-1) + θ

toy model of QCD

Strong CP problem

MC CP(3)

$$Z(\theta) = \sum_Q e^{i\theta Q} P(Q)$$



Application to CP(N-1) + θ

toy model of QCD

Strong CP problem

$$Z(\theta) = \sum_Q e^{i\theta Q} P(Q)$$