All-mode Renormalization for Tensor Network with Stochastic Noise

PRD107,114515(2023)

Shinji Takeda

(Kanazawa U.)

in collaboration with E. Arai, H. Ohki and M. Tomii

Field Theory Research Team Seminar @R-CCS in Kobe

2023.09.01

Contents

- Introduction of tensor network
  - Why/What's tensor network
  - Tensor renormalization group (TRG)

- All-mode renormalization
  - How to use stochastic noise in tensor decomposition
  - Key idea comes from lattice QCD
  - Numerical results

Introduction of tensor network

Why tensor networks?

- Good
  - Applicable to any models even for complex action

    ≈ no sign problem
  - Extremely large volume (thermodynamic limit)
  - High-precision is attainable in 2D system with simple internal d.o.f.
- Bad
  - Expensive for higher dimensional system

Why tensor networks?

- Good
  - Applicable to any models even for complex action

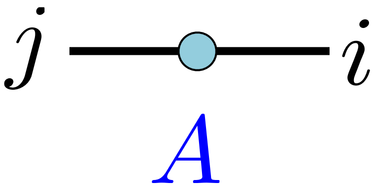
    ≈ no sign problem

Challenge to

- QCD + μ
- θ-term
- Lattice SUSY
- Real-time dynamics
- Chiral gauge theory

Notational rules

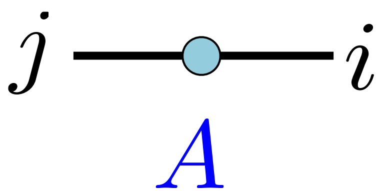
Rank 2 tensor (matrix)



$$= A_{ij}$$

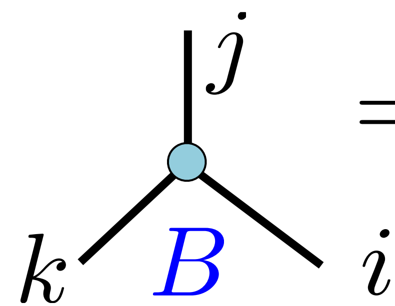
Notational rules

Rank 2 tensor (matrix)



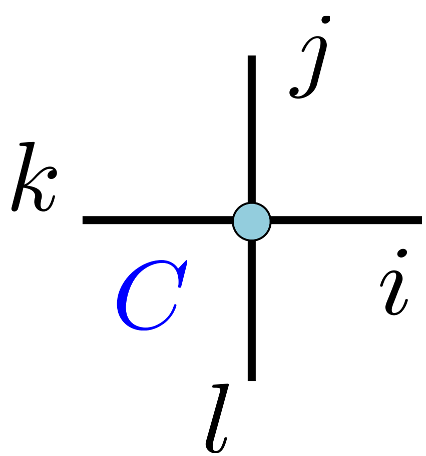
$$= A_{ij}$$

Rank 3 tensor



$$= B_{ijk}$$

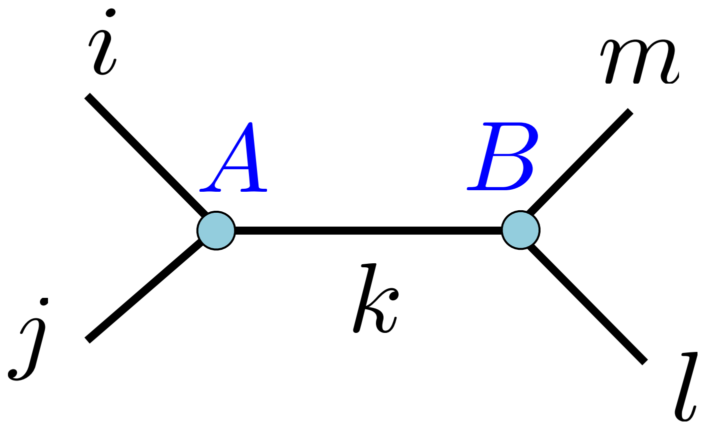
Rank 4 tensor



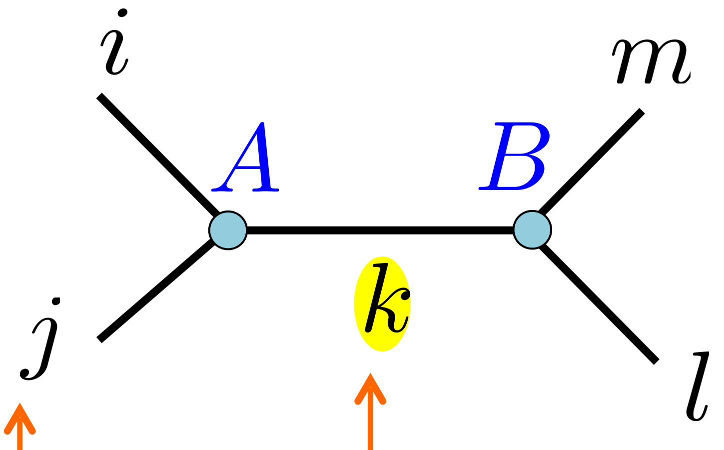
$$= C_{ijkl}$$

Tensor : vertex
index : link

Contraction (summation) rule



Contraction (summation) rule



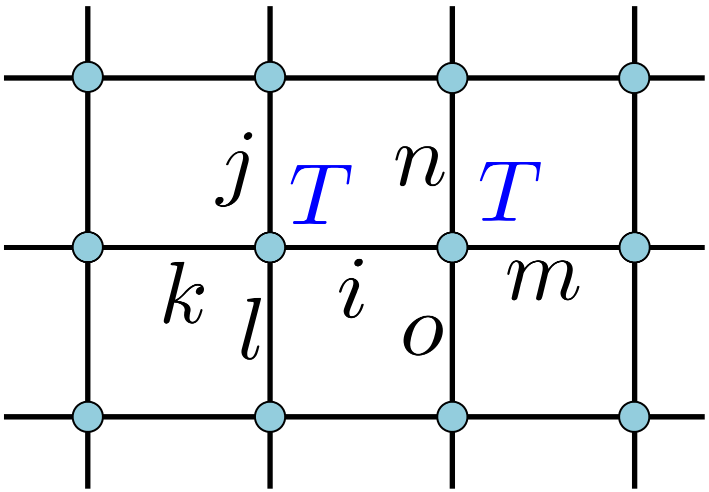
$$= \sum_{k} A_{ijk} B_{klm}$$

index shared by two tensors : summation

index connecting with a single tensor : no summation

What's tensor network?

Example: TN for square lattice

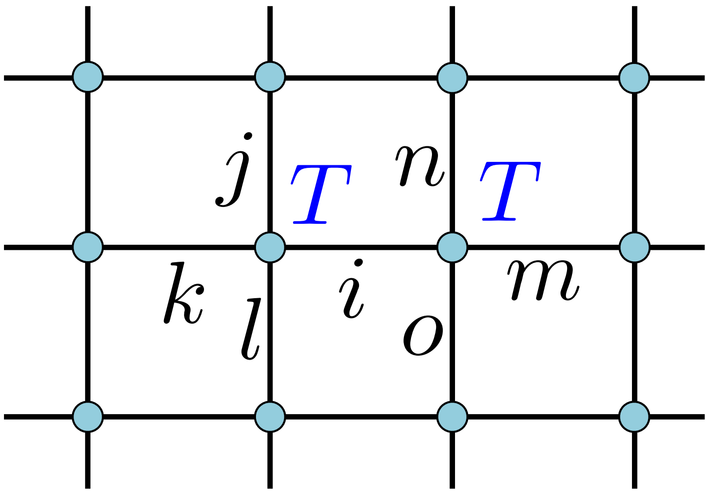


$$= \sum_{\cdots,i,j,k\cdots} \cdots T_{ijkl} T_{mnio} \cdots$$

$\Rightarrow$ Tensor network

What’s tensor network?

Example: TN for square lattice



$$= \sum_{\cdots,i,j,k\cdots} \cdots T_{ijkl} T_{mnio} \cdots$$

$\Rightarrow$ Tensor network

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

Two approaches in TN

	Hamiltonian approach	**Lagrangian approach**
target system	quantum many-body system	Classical statistical system, Path-integral rep. of quantum field theory
TN is used to express	wave function	partition function, path integral
combining with	variational method	coarse-graining (real-space renormalization group)

This talk

Working flow of Lagrangian approach

Target model in continuum space-time

↓ Lattice regularization

Lattice model → MC if no sign problem

↓

TN rep. of Z

$$Z \equiv \int [d\phi] e^{-S[\phi]} \stackrel{?}{=} \sum_{..,i,j,k,l,..} \cdots T_{ijkl} T_{mnio} \cdots$$

Working flow of Lagrangian approach

Target model in continuum space-time

↓ Lattice regularization

Lattice model → MC if no sign problem

↓

TN rep. of *Z*

- Non-trivial step, but
- OK for scalar, gauge, and fermion fields as long as the interaction is local

$$Z \equiv \int [d\phi] e^{-S[\phi]} \stackrel{?}{=} \sum_{..,i,j,k,l,..} \cdots T_{ijkl} T_{mnio} \cdots$$

Working flow of Lagrangian approach

Target model in continuum space-time

Lattice regularization

Lattice model

MC if no sign problem

TN rep. of Z

For each model, tensor is prepared

Bond dimension

cost $\propto \chi^{2V}$ for $1 \leq i, j, \ldots \leq \chi$

$$Z \equiv \int [d\phi] e^{-S[\phi]} = \sum_{\ldots, i, j, k, l, \ldots} \cdots T_{ijkl} T_{mnio} \cdots$$

Working flow of Lagrangian approach

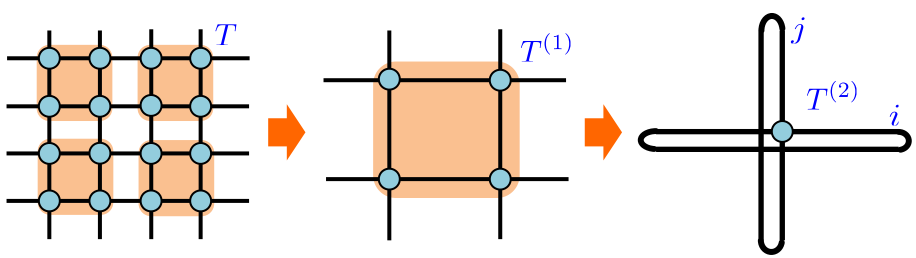
Target model in continuum space-time

Lattice regularization

Lattice model → MC if no sign problem

TN rep. of Z — For each model, tensor is prepared

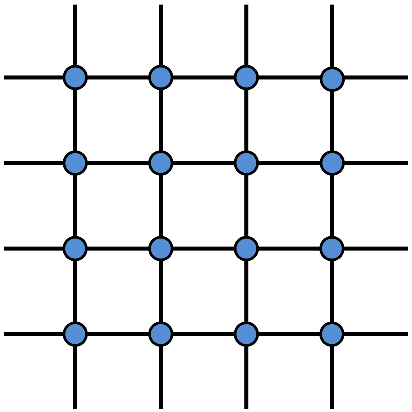
Coarse-graining — like spin-blocking/real space renormalization group



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

Coarse-graining

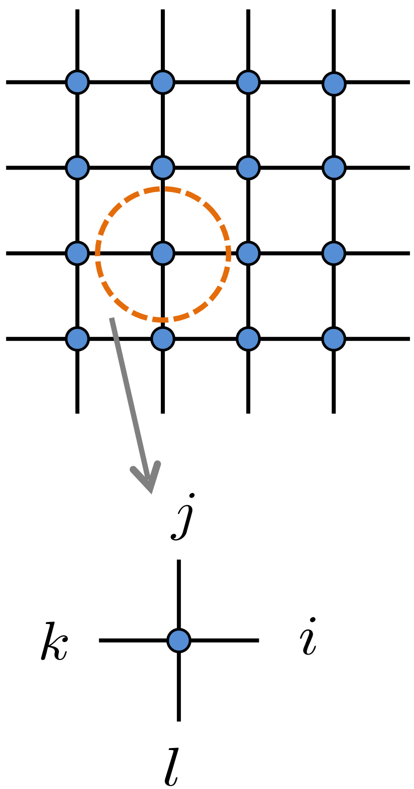
Tensor renormalization group (TRG)
PRL99,120601(2007)



Coarse-graining

Tensor renormalization group (TRG)

PRL99,120601(2007)



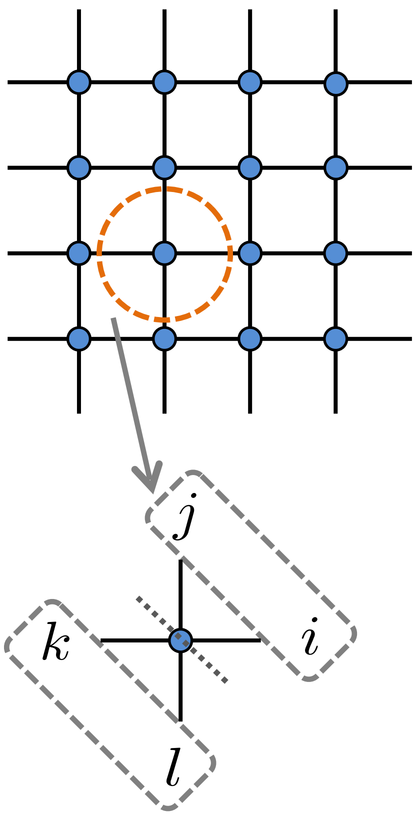
Bond dimension

$$1 \leq i, j, ... \leq \chi$$

$$\Leftrightarrow \quad T_{ijkl}$$

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)



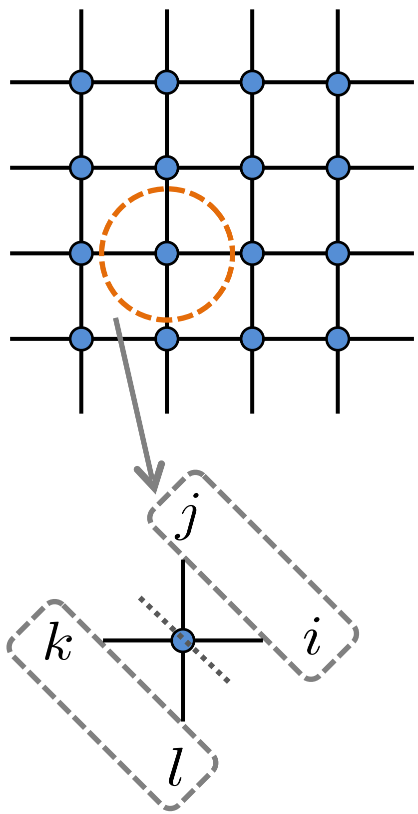
$$1 \leq i, j, ... \leq \chi$$

$\chi^2 \times \chi^2$ matrix

$$\Leftrightarrow \quad T_{ijkl} = M_{(ij)(kl)}$$

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)



$$M \in \mathbb{C}^{\chi^2 \times \chi^2}$$

Singular Value Decomposition(SVD)

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

unitary matrix

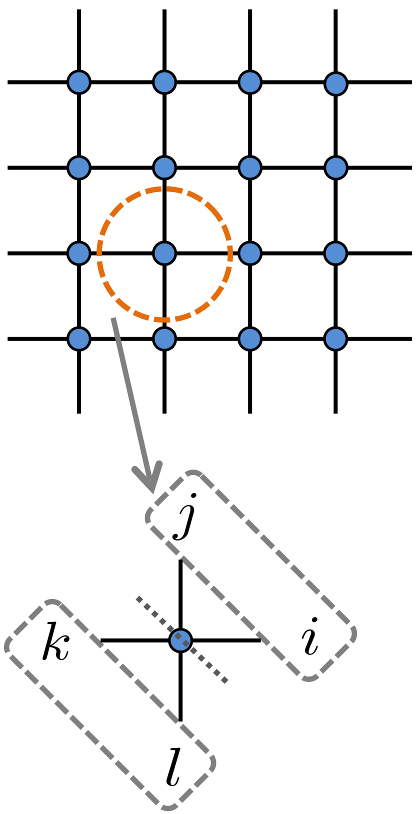
$$\Lambda_1 \geq \Lambda_2 \geq \ldots \geq 0$$

: singular value （non-negative）

$$\Leftrightarrow \quad T_{ijkl} = M_{(ij)(kl)}$$

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)



$$M \in \mathbb{C}^{\chi^2 \times \chi^2} \quad \Rightarrow \text{TN is sign-problem-free}$$

Singular Value Decomposition(SVD)

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

unitary matrix

$$\Lambda_1 \geq \Lambda_2 \geq \ldots \geq 0$$

: singular value （non-negative）

$$\Leftrightarrow \quad T_{ijkl} = M_{(ij)(kl)}$$

SVD

$$= \sum_{m=1}^{\chi^2} u_{(ij)m} \Lambda_m v^\dagger_{m(kl)}$$

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)

$M \in \mathbb{C}^{\chi^2 \times \chi^2}$ ⇒ TN is sign-problem-free

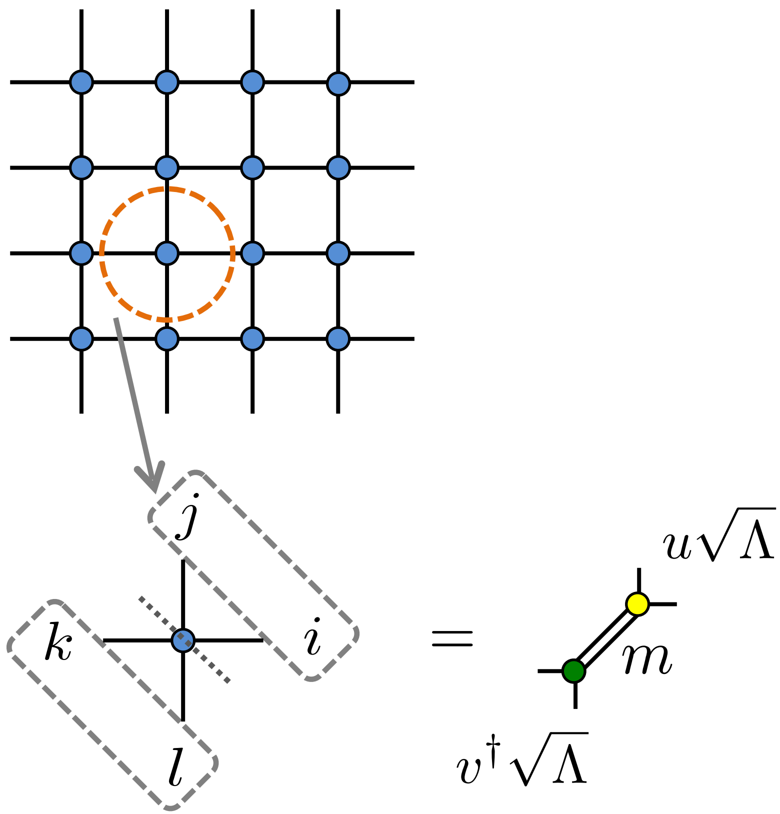
Singular Value Decomposition(SVD)

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

unitary matrix

$$\Lambda_1 \geq \Lambda_2 \geq \ldots \geq 0$$

: singular value （non-negative）



$$\Leftrightarrow \quad T_{ijkl} = M_{(ij)(kl)}$$

SVD

$$= \sum_{m=1}^{\chi^2} u_{(ij)m} \Lambda_m v^\dagger_{m(kl)}$$

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)

$M \in \mathbb{C}^{\chi^2 \times \chi^2}$ $\Rightarrow$ TN is sign-problem-free

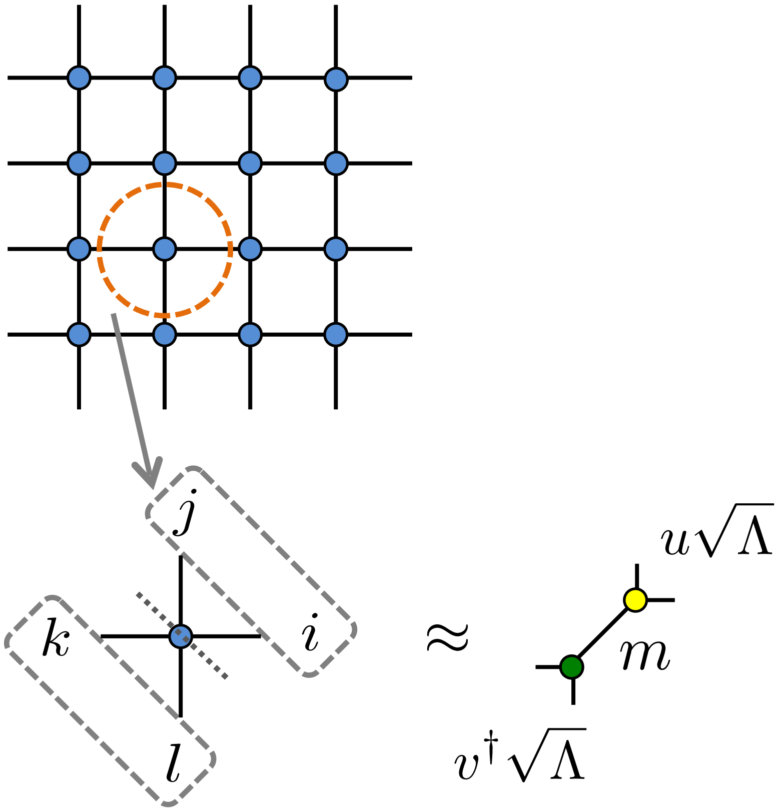
Singular Value Decomposition(SVD)

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

unitary matrix

$$\Lambda_1 \geq \Lambda_2 \geq \ldots \geq 0$$

: singular value （non-negative）



$\Leftrightarrow$

$$T_{ijkl} = M_{(ij)(kl)}$$

truncation

$$\approx \sum_{m=1}^{D_{\mathrm{cut}}} u_{(ij)m} \Lambda_m v^\dagger_{m(kl)}$$

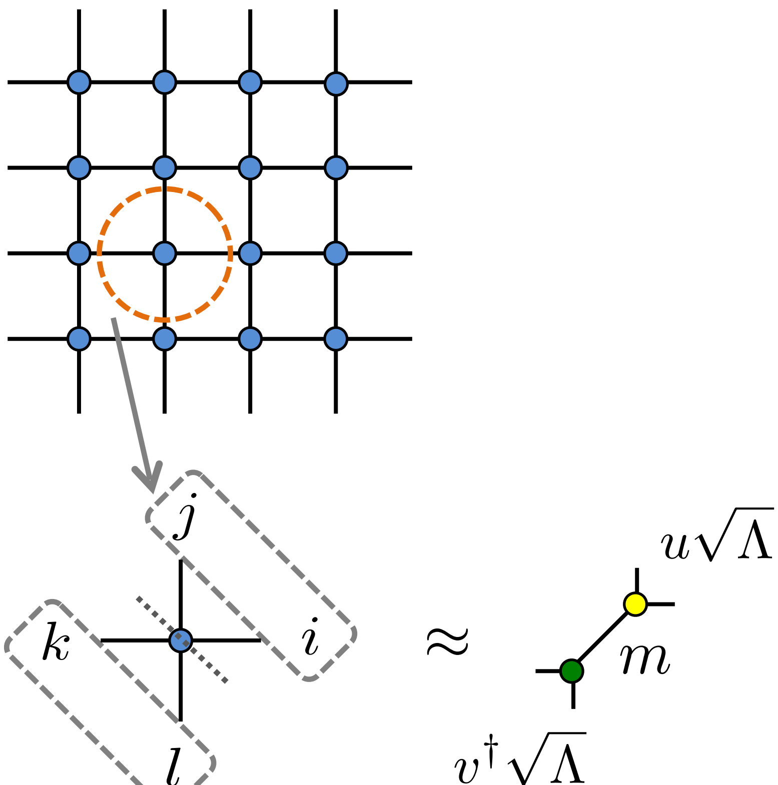
$D_{\mathrm{cut}} < \chi^2$

$D_{\mathrm{cut}} = \chi$: standard choice

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)

$M \in \mathbb{C}^{\chi^2 \times \chi^2}$ $\Rightarrow$ TN is sign-problem-free



Singular Value Decomposition(SVD)

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

unitary matrix

$$\Lambda_1 \geq \Lambda_2 \geq \ldots \geq 0$$

: singular value (non-negative)

$$\Leftrightarrow \quad T_{ijkl} = M_{(ij)(kl)}$$

truncation

$$\approx \sum_{m=1}^{D_{\mathrm{cut}}} u_{(ij)m} \Lambda_m v^\dagger_{m(kl)}$$

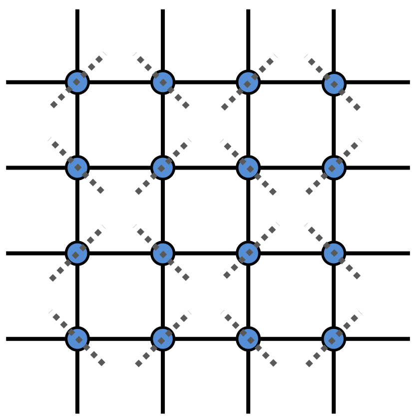
truncation of SVD = information compression

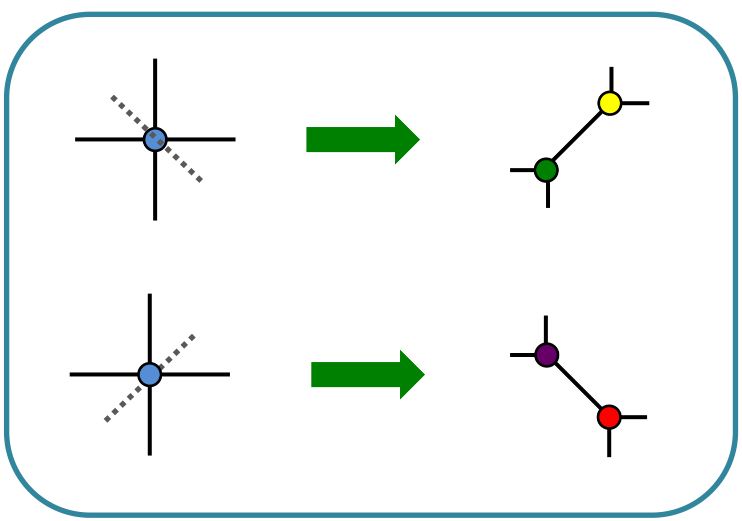
$D_{\mathrm{cut}} < \chi^2$

$D_{\mathrm{cut}} = \chi$: standard choice

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)

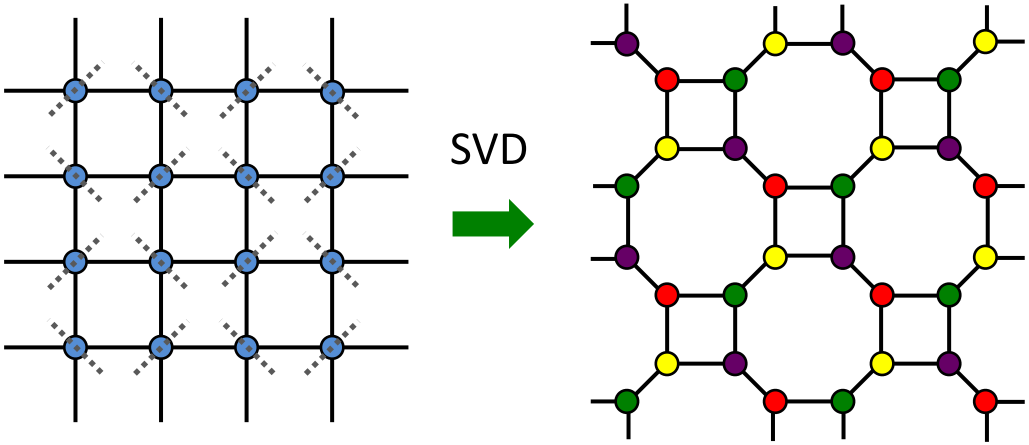


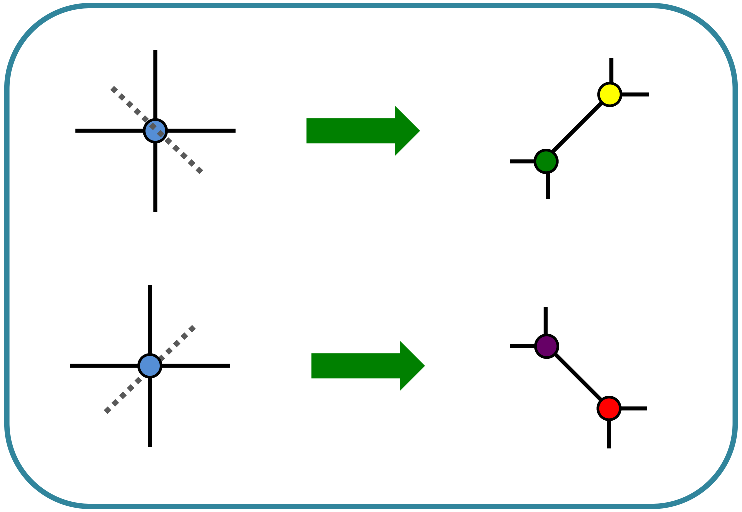


truncated SVD

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)

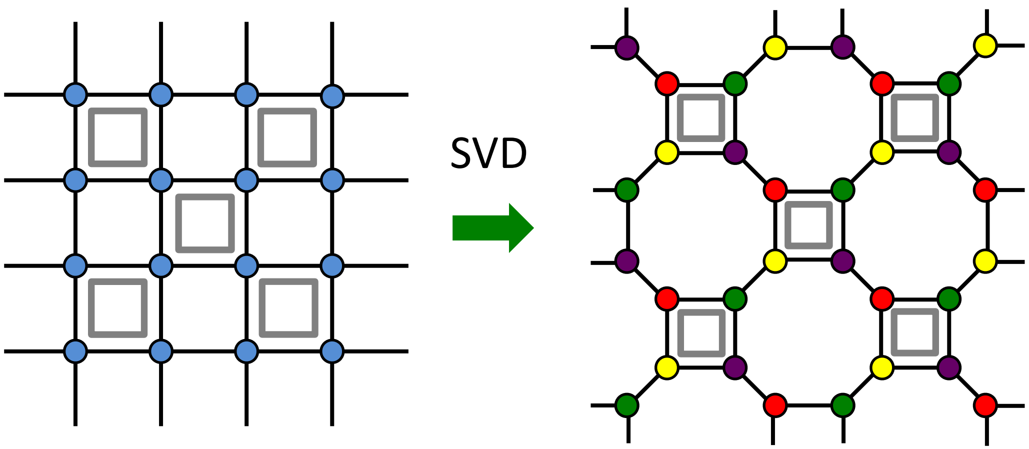


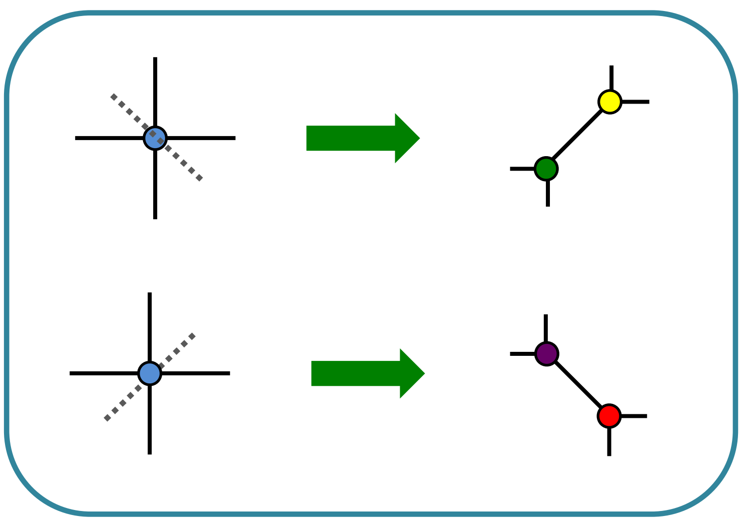


truncated SVD

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)

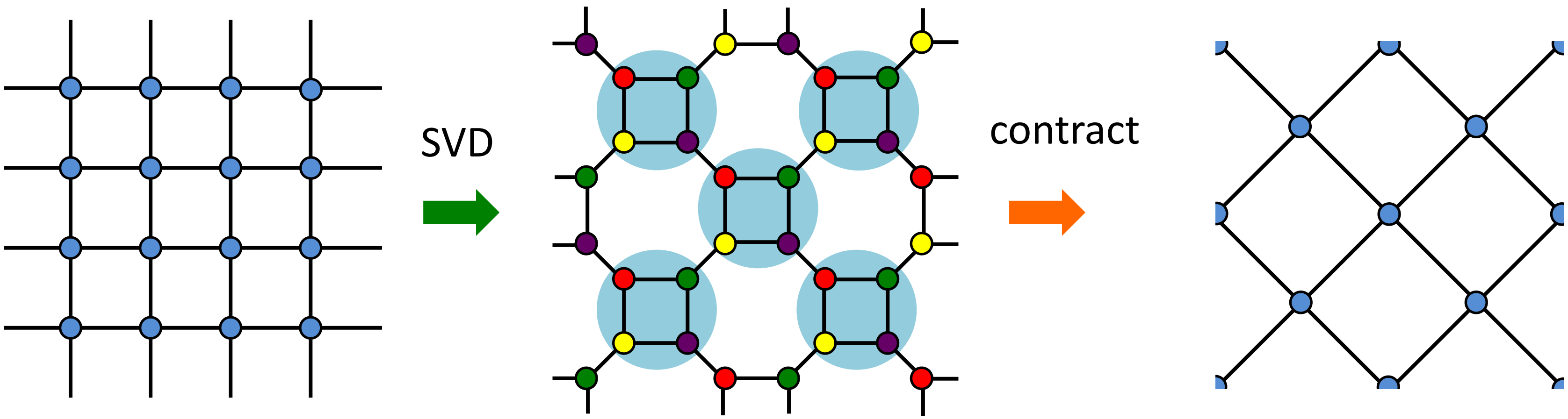




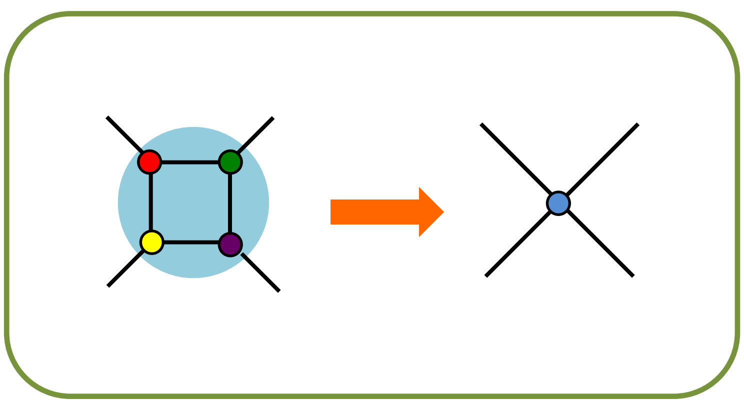
truncated SVD

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)



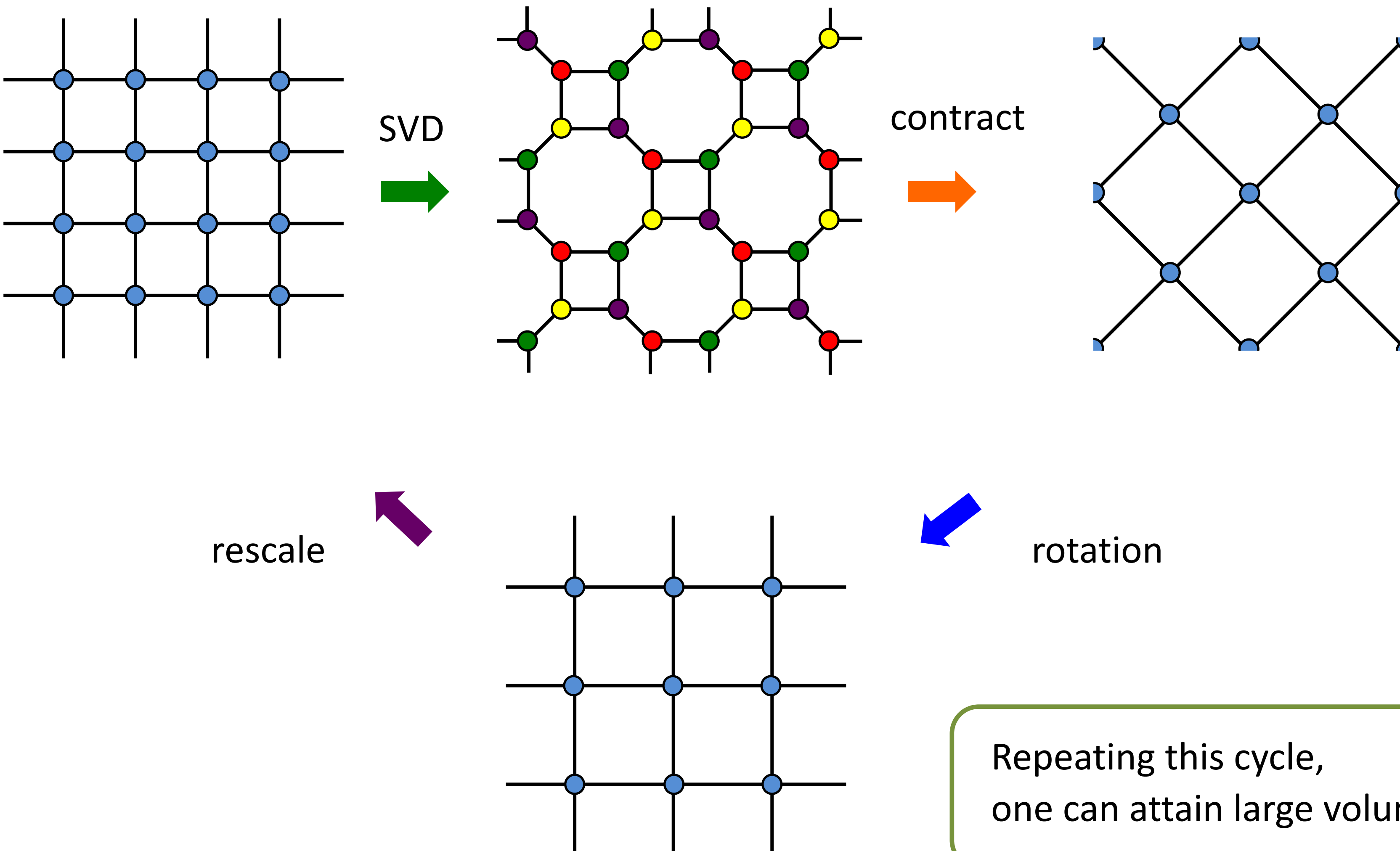
contraction



renormalization!

Coarse-graining

Tensor renormalization group (TRG)
PRL99,120601(2007)



Repeating this cycle,
one can attain large volume

Summary (so far)

- Tensor network is free of sign problem
- Key point of coarse-graining scheme is information compression based on singular value decomposition
- For Langrangian approach, improvement of coarse-graining algorithm is essential
- 4D system with simple d.o.f. is now feasible

All-mode renormalization

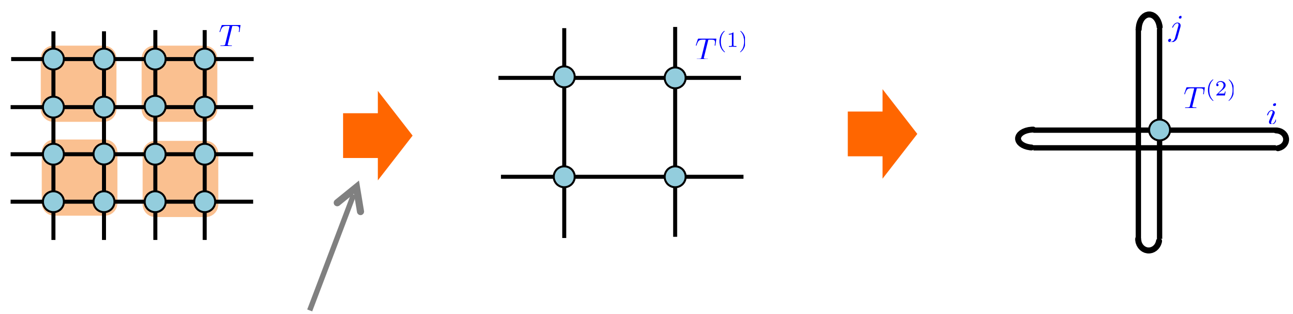
PRD107,114515(2023)

Lagragian approach

$$Z \equiv \int [d\phi] e^{-S[\phi]} = \sum_{..,i,j,k,l,..} \cdots T_{ijkl} T_{mnio} \cdots$$

Periodic BC

=



Coarse-graining by TRG

Information compression based on
truncated singular value decomposition

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

Levin+Nave 2007

Systematic error

- Coarse-graining algorithms for tensor network use truncated singular value decomposition to reduce computational cost
- However, the truncation introduces systematic error Furthermore, iterations accumulate systematic errors
- Can we reduce the systematic errors?
- Stochastic method can remove the systematic errors

Ferris 2015

Truncated SVD

$$M_{ab} = \sum_{s=1}^{R} u_{as} \Lambda_s v_{sb} \qquad (\Lambda_1 \geq \Lambda_2 \geq \cdots \geq \Lambda_R > 0)$$

low mode

$$= \sum_{s=1}^{D_{\mathrm{svd}}} u_{as} \Lambda_s v_{sb} + \cancel{\sum_{s=D_{\mathrm{svd}}+1}^{R} \cdots} \qquad (D_{\mathrm{svd}} < R)$$

$$\Searrow \; O(\Lambda_{D_{\mathrm{svd}}+1})$$

$$\approx \sum_{s=1}^{D_{\mathrm{svd}}} u_{as} \Lambda_s v_{sb}$$

Using stochastic noise

$$M_{ab} = \sum_{s=1}^{R} u_{as}\Lambda_s v_{sb} \qquad (\Lambda_1 \geq \Lambda_2 \geq \cdots \geq \Lambda_R > 0)$$

low mode

$$= \sum_{s=1}^{D_{\rm svd}} u_{as}\Lambda_s v_{sb} + \boxed{\sum_{s=D_{\rm svd}+1}^{R} u_{as}\Lambda_s v_{sb}} \qquad (D_{\rm svd} < R)$$

$$\sum_{s,t}^{R} u_{as}\Lambda_s \underline{\delta_{st}} v_{tb}$$

$$\delta_{st} = \frac{1}{N_{\rm r}} \sum_{r=1}^{N_{\rm r}} \eta_{sr}\eta^*_{tr} + O\left(\frac{1}{\sqrt{N_{\rm r}}}\right)$$

$e^{i2\pi m/N} \quad (m = 1, 2, \cdots, N)$

$N_{\rm r}$: # of noise

(dimension of noise vector)

$\eta_{sr} \in Z_N$: stochastic noise

$\langle \eta_s \eta_t \rangle = \delta_{st} \qquad \langle \eta_s \rangle = 0$

Using stochastic noise

$$M_{ab} = \sum_{s=1}^{R} u_{as}\Lambda_s v_{sb} \qquad (\Lambda_1 \ge \Lambda_2 \ge \cdots \ge \Lambda_R > 0)$$

low mode

$$= \sum_{s=1}^{D_{\rm svd}} u_{as}\Lambda_s v_{sb} + \boxed{\sum_{s=D_{\rm svd}+1}^{R} u_{as}\Lambda_s v_{sb}} \qquad (D_{\rm svd} < R)$$

$$\sum_{s,t}^{R} u_{as}\Lambda_s \delta_{st} v_{tb} \approx \sum_{r=1}^{N_{\rm r}} \underbrace{\left(\sum_{s}^{R} \sqrt{\frac{\Lambda_s}{N_{\rm r}}}\, u_{as}\eta_{sr}\right)}_{S_{3ar}} \underbrace{\left(\sum_{t}^{R} \sqrt{\frac{\Lambda_t}{N_{\rm r}}}\, v_{tb}\eta^*_{tr}\right)}_{S_{1br}}$$

all lower modes are included

$$\delta_{st} = \frac{1}{N_{\rm r}} \sum_{r=1}^{N_{\rm r}} \eta_{sr}\eta^*_{tr} + O\left(\frac{1}{\sqrt{N_{\rm r}}}\right)$$

$N_{\rm r}$: # of noise
(dimension of noise vector)

$\eta_{sr} \in Z_N$: stochastic noise

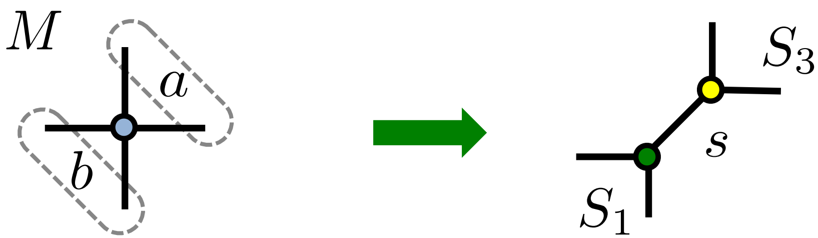
$\langle \eta_s \eta_t \rangle = \delta_{st} \qquad \langle \eta_s \rangle = 0$

SVD + stochastic noise

$$M_{ab} = \sum_{s=1}^{D_{\rm svd}+N_{\rm r}} S_{3as}(\eta) S_{1bs}(\eta) + O\left(\frac{\Lambda_{D_{\rm svd}+1}}{\sqrt{N_{\rm r}}}\right)$$

$D_{\rm cut} = D_{\rm svd} + N_{\rm r}$

$N_{\rm r}$: noise dim.



$$S_{3as}(\eta) = \begin{cases} \sqrt{\Lambda_s} u_{as} & (1 \le s \le D_{\rm svd}) \\ \sum_{i=D_{\rm svd}+1}^{R} \sqrt{\frac{\Lambda_i}{N_r}} u_{ai} \eta_{i-D_{\rm svd}, s-D_{\rm svd}} & (D_{\rm svd}+1 \le s \le D_{\rm svd}+N_r) \end{cases}$$

$$S_{1bs}(\eta) = \begin{cases} \cdots \\ \cdots \end{cases}$$

Larger modes are exactly included

All lower modes are included in combined with stochastic noise

all modes are included

Foley+ 2005

To reduce the error $O\left(\frac{\Lambda_{D_{\rm svd}+1}}{\sqrt{N_{\rm r}}}\right)$ we need $N_{\rm r} \to \infty \iff D_{\rm cut} \to \infty$

This is NOT practically useful

Ensemble of noises

Ensemble of noises $\eta^{[\ell]}$ $(\ell = 1, 2, \cdots, N)$

$$D_{\rm cut} = D_{\rm svd} + N_{\rm r}$$

$$M_{ab} = \underline{\frac{1}{N}\sum_{\ell=1}^{N}} \sum_{s=1}^{D_{\rm svd}+N_{\rm r}} S_{3as}(\underline{\eta^{[\ell]}}) S_{1bs}(\underline{\eta^{[\ell]}}) + O\left(\frac{\Lambda_{D_{\rm svd}+1}}{\sqrt{N_{\rm r}\underline{N}}}\right)$$

$$Z_V = \lim_{N\to\infty} \frac{1}{N}\sum_{\ell=1}^{N} Z(T^{(n)[\ell]})$$ with $$Z(T^{(n)[\ell]}) = \sum_{i,j=1}^{D_{\rm cut}} T^{(n)[\ell]}_{ijij}$$

$$V = 2^n$$

On the error

$$O\left(\frac{\Lambda_{D_{\rm svd}+1}}{\sqrt{N_{\rm r}N}}\right) \to 0$$ for $N \to \infty$ with fixed $D_{\rm cut} = D_{\rm svd} + N_{\rm r}$

$\Rightarrow$ With finite $D_{\rm cut}$, the decomposition becomes exact for $N \to \infty$

moreover, $\boxed{\text{No autocorrelation, Parallelizable}}$

Ferris arXiv:1507.00767

Application to TRG

TRG

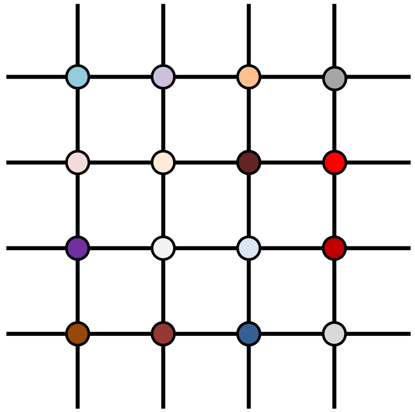
truncated SVD $\Rightarrow$ SVD + stochastic noise

replace

Noise ensemble method

- Position-dependent noise

Ferris arXiv:1507.00767 different probability distribution



noises are generated independently for each site

= position-dependent noise

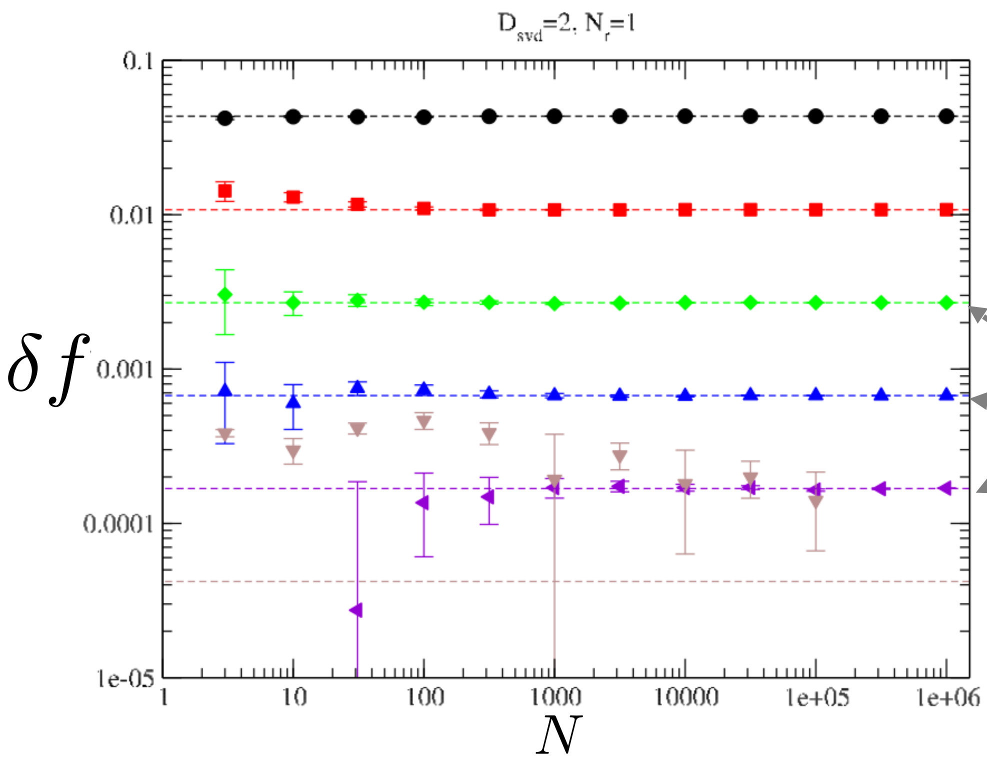
TRG + Position-dependent noise

$$\delta f = \left| \frac{f - f_{\mathrm{exact}}}{f_{\mathrm{exact}}} \right|$$

f : free energy

Onsager 1944

- Ising on square lattice
- $T = T_c$
- $V = 2^n$ (n : # of RG steps)
- $D_{\mathrm{cut}} = D_{\mathrm{svd}} + N_r = 2 + 1$
- N : # of samples



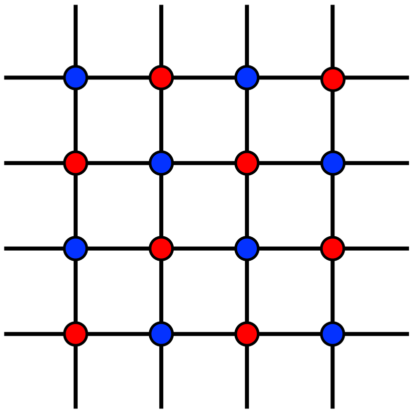
exact value at finite volume

Kaufman 1949

Application to TRG

- Position-dependent noise   Ferris arXiv:1507.00767 different probability distribution
  - No systematic error. Only statistical error
  - Cost : $O\left(D_{\rm cut}^6\ N\ V\right)$ N: # of samples, V: volume   → high cost!
- Common noise
  - Introduce even-odd correlated noise
  - Cost : $O(D_{\rm cut}^6\ N \log V)$
  - New systematic error ~ $1/N_{\rm r}$

($N_{\rm r}$: noise dim.)

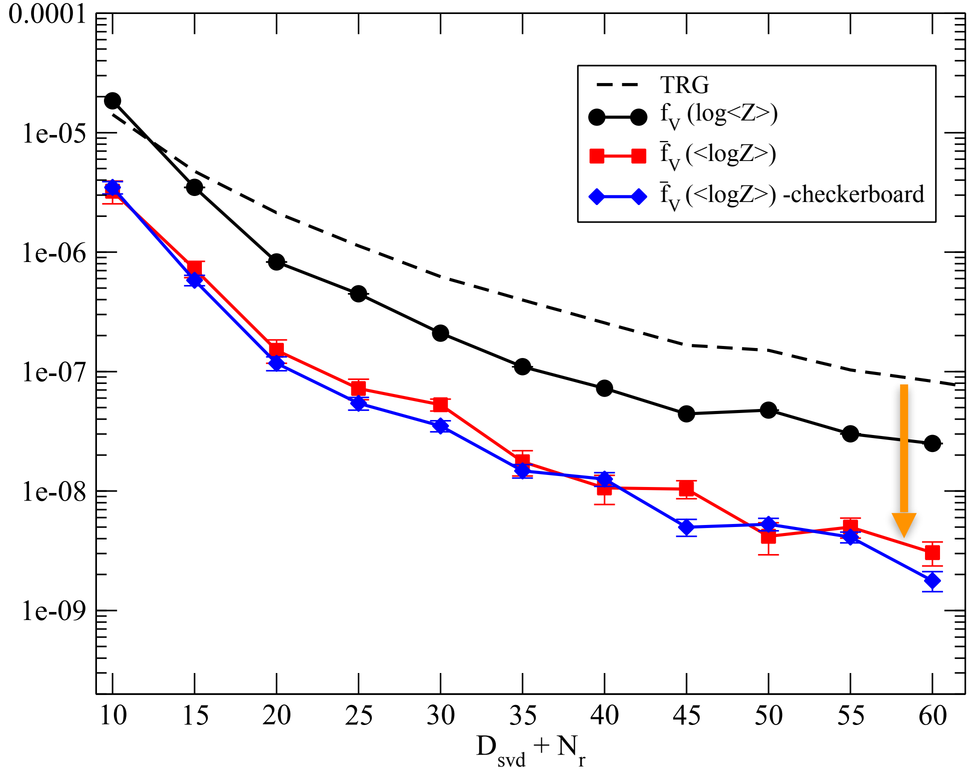


TRG + Common noise

$$\delta f = \left| \frac{f - f_{\mathrm{exact}}}{f_{\mathrm{exact}}} \right|$$

f : free energy

- Ising on square lattice
- $T = T_c$
- $V = 2^{50}$
- $D_{\mathrm{cut}} = D_{\mathrm{svd}} + 4 \quad (N_{\mathrm{r}} = 4)$
- $N = 100$ samples

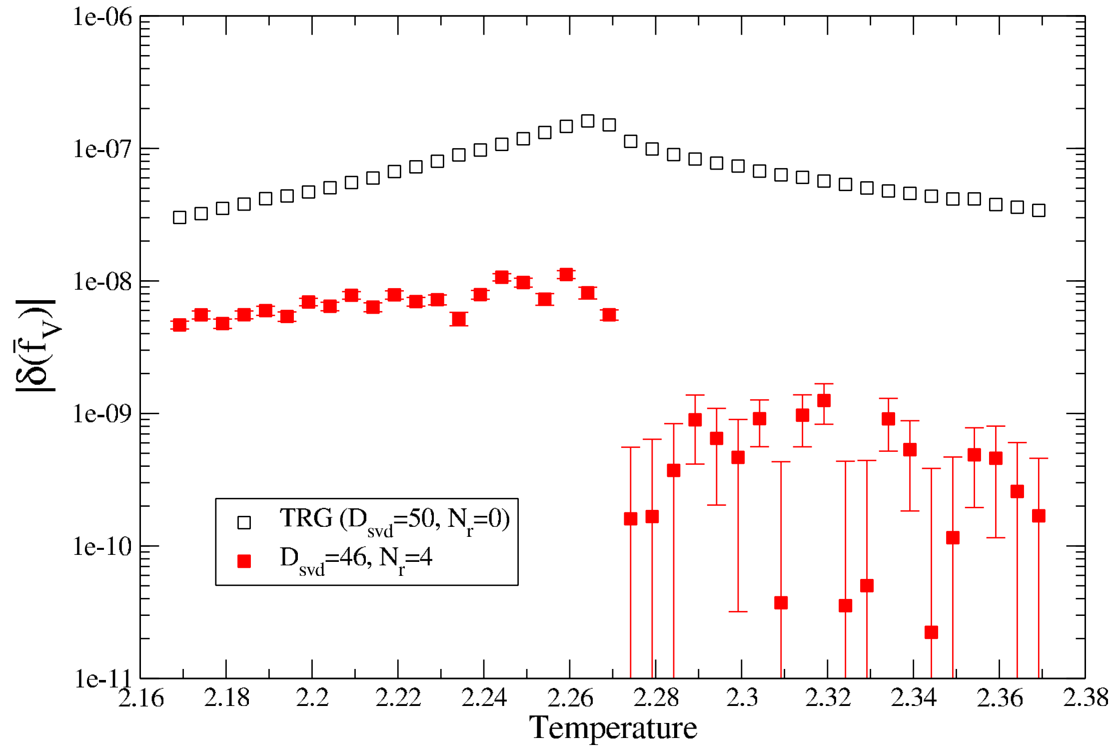


δf

$$D_{\mathrm{cut}} = D_{\mathrm{svd}} + N_r$$

TRG $\Leftrightarrow N_r = 0$

δf vs. Temperature



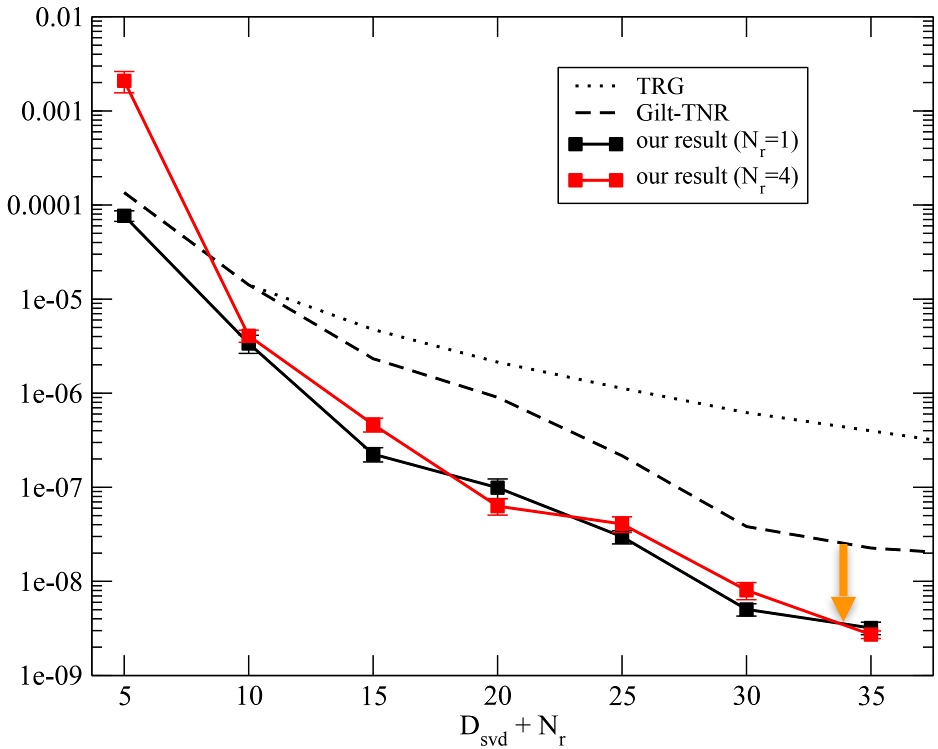
- Ising on square lattice
- TRG + Common noise
- $D_{\rm cut} = 50$
- $V = 2^{30}$
- $N = 100$ samples

Gilt-TNR + common noise

Gilt-TNR : sophisticated coarse-graining algorithm

Hauru+Delcamp+Mizera 2018

δf



- Ising on square lattice
- $T = T_c$
- $V = 2^{51}$
- $D_{\mathrm{cut}} = D_{\mathrm{svd}} + N_{\mathrm{r}}$
- $N = 100$ samples
- $\varepsilon = 8 \times 10^{-8}$: threshold parameter for Gilt

$$D_{\mathrm{cut}} = D_{\mathrm{svd}} + N_r$$

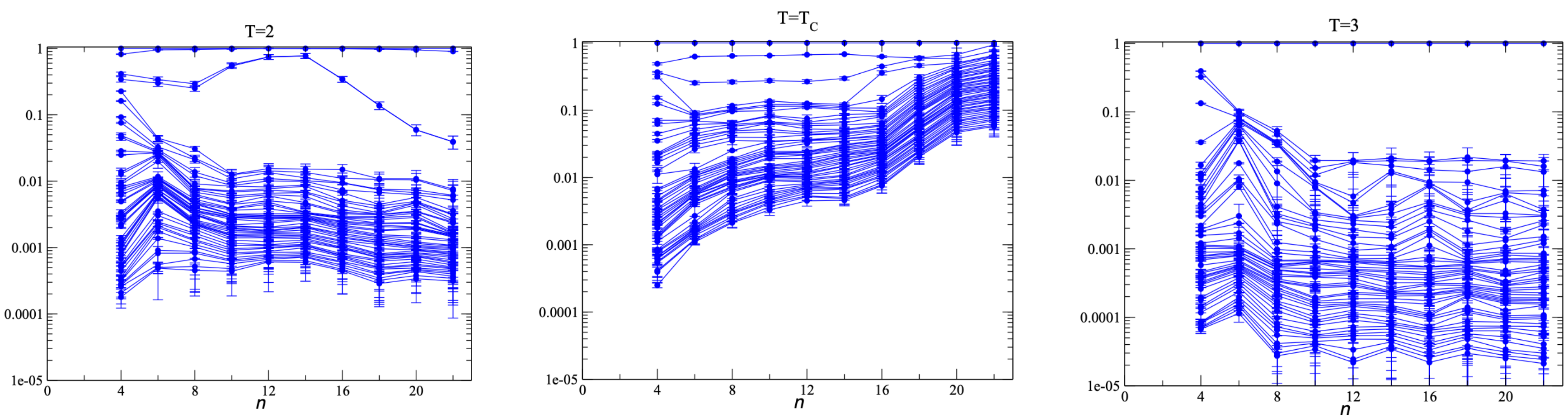
Summary

- Stochastic noise removes systematic error in tensor (matrix) decomposition at the expense of statistical errors
- Position-dependent noise is free of systematic error but expensive
- Common noise is relatively cheap and shows better accuracy and simple error scaling
- Our stochastic method is easily applicable to all truncated SVD steps in coarse-graining algorithms.

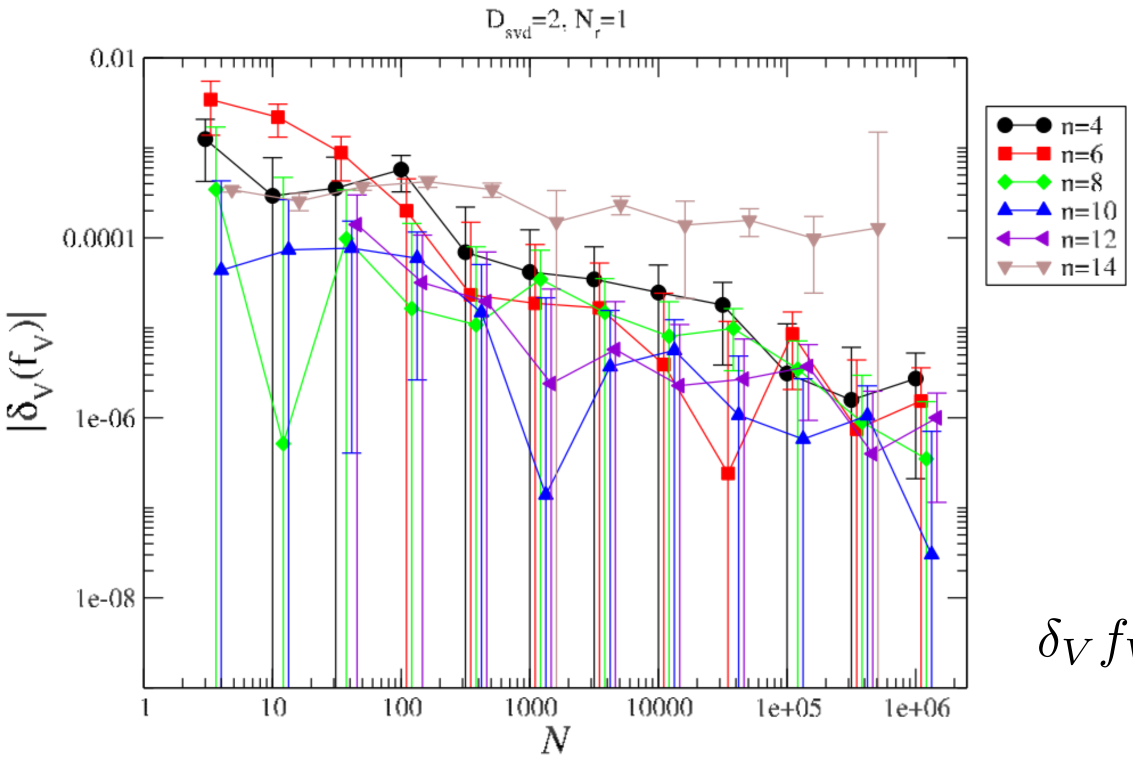
Backup slides

Flow of singular values

- Ising on square lattice
- TRG + position-dep. noise
- n= # of RG step
- $D_{\rm cut} = 7 + 1$
- 500 statistics



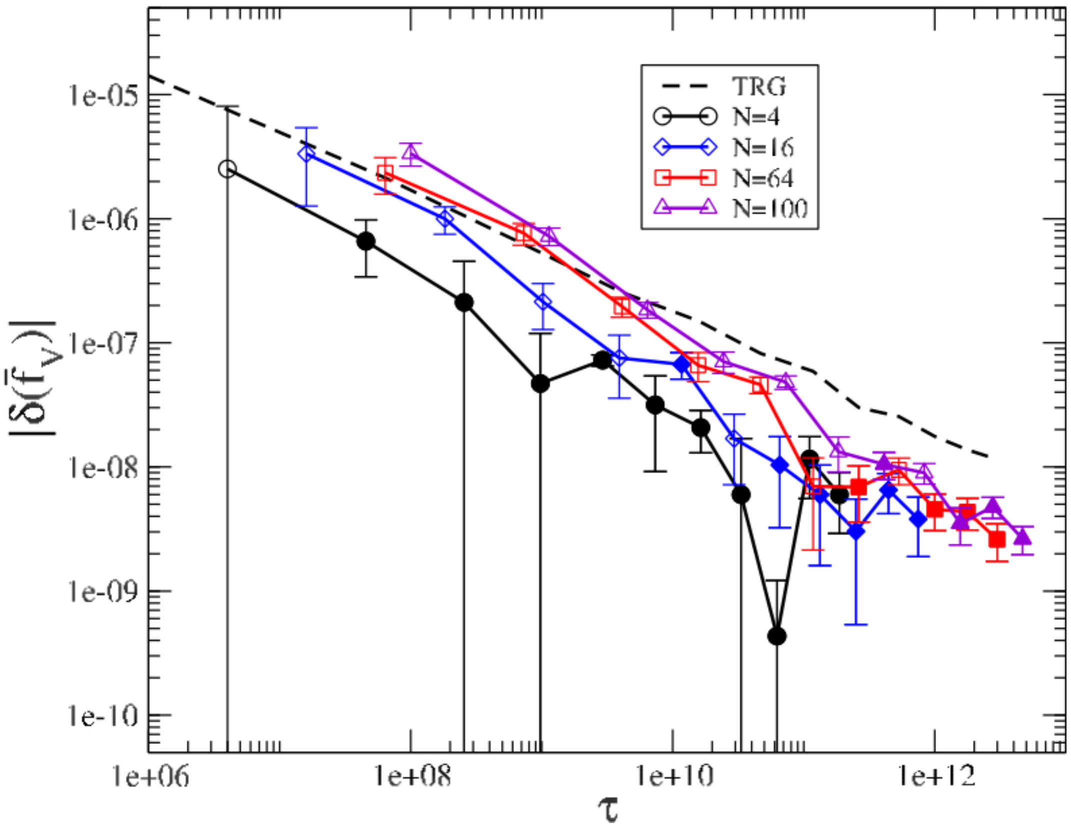
Position-dependent noise



- Ising on square lattice
- $T = T_c$
- $V = 2^{30}$
- $D_{\mathrm{cut}} = 2 + 1$
- N= # of samples

$$\delta_V f_V = \left| \frac{f_V - f_{V,exact}}{f_{V,exact}} \right| \propto \frac{1}{\sqrt{N}}$$

Performance of common noise

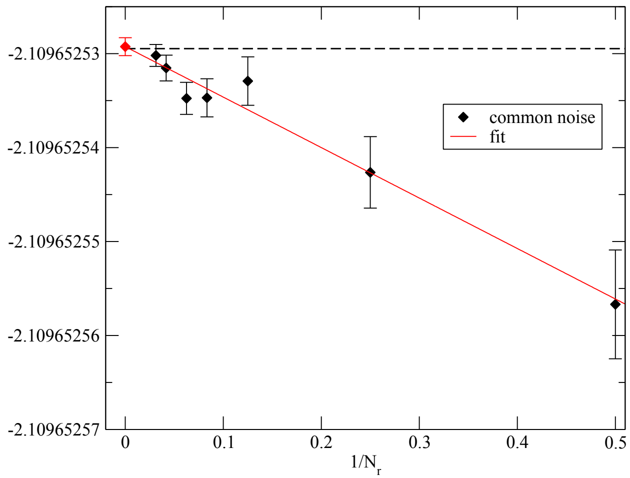


- Ising on square lattice
- $T = T_c$
- $V = 2^{30}$
- $D_{\mathrm{cut}} = D_{\mathrm{svd}} + 4$
- N= # of samples

$$\tau = \begin{cases} D_{\mathrm{cut}}^6 & \text{for TRG} \\ N D_{\mathrm{cut}}^6 & \text{for common noise method.} \end{cases}$$

Systematic error for common noise

f

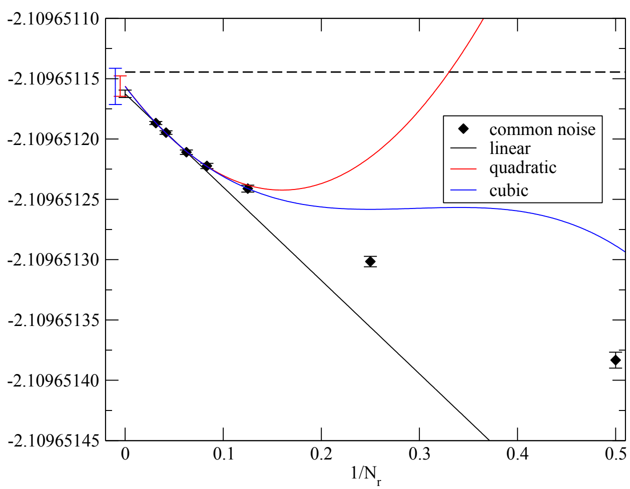


$1/N_r$ scaling is observed

$V = 2^{20}$

- Ising on square lattice
- TRG + Common noise
- $D_{\mathrm{cut}} = 20 + N_{\mathrm{r}}$
- N= 5000 statistics

f



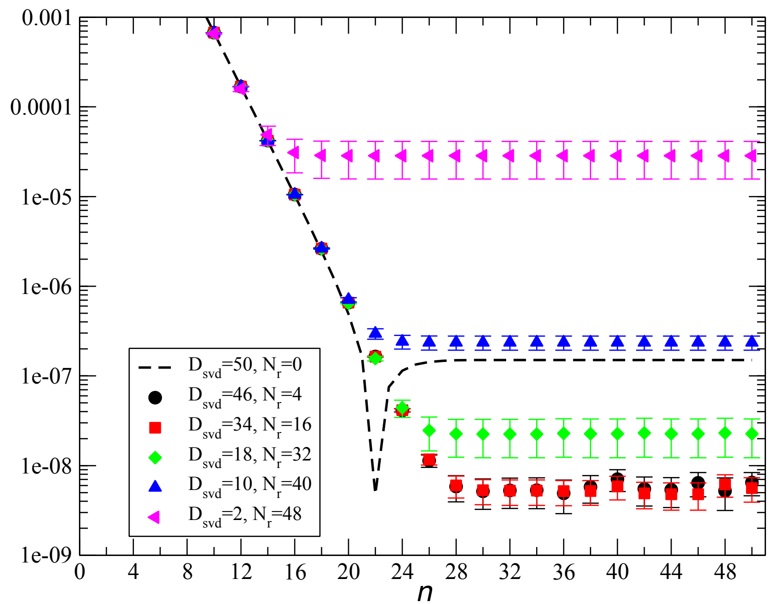
$V = 2^{40}$

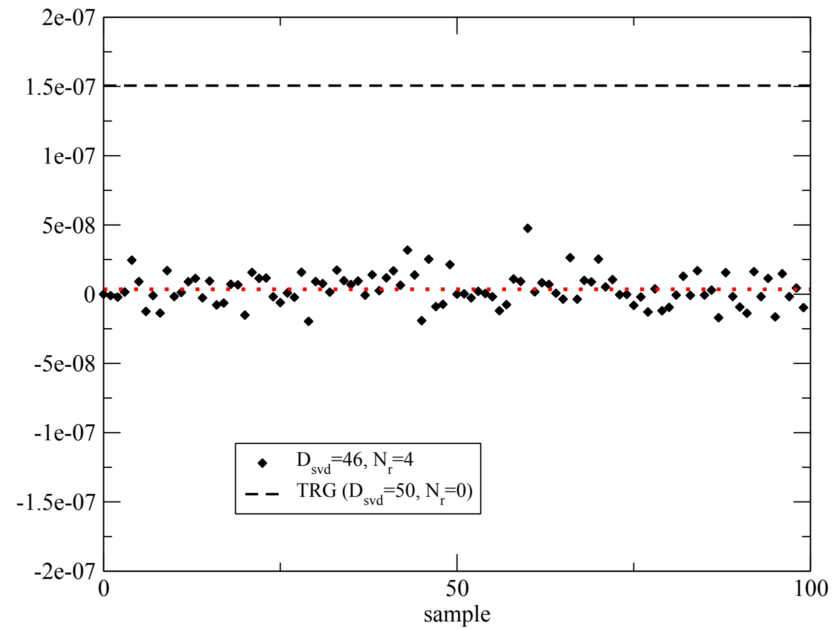
Common noise results

- Ising on square lattice
- TRG + Common noise
- $T = T_c$
- $D_{\mathrm{cut}} = D_{\mathrm{svd}} + N_r$

$N = 50$

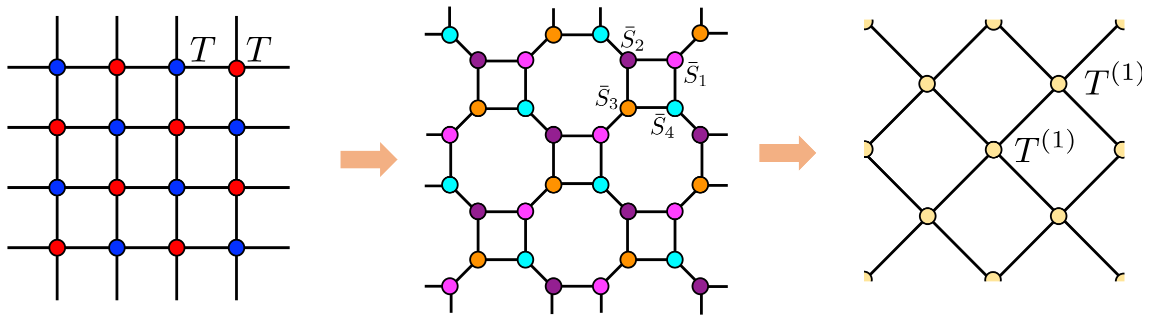
δf





$V = 2^{50}$

Noise correlation for common noise



$$
\begin{aligned}
\frac{1}{N_r^2}\left(\boldsymbol{\eta}\cdot\boldsymbol{\eta}^\dagger\right)_{ab}\otimes\left(\boldsymbol{\eta}\cdot\boldsymbol{\eta}^\dagger\right)_{cd} =& \frac{1}{N_r^2}\sum_{r_1=1}^{N_r}\eta_{ar_1}\eta^*_{br_1}\sum_{r_2=1}^{N_r}\eta_{cr_2}\eta^*_{dr_2} \\
=& \frac{N_r-1}{N_r^2}\sum_{r_1=1}^{N_r}\eta_{ar_1}\eta^*_{br_1}\left(\delta_{cd}+\mathcal{O}(1/\sqrt{N_r})\right) \\
&+\frac{1}{N_r^2}\sum_{r_1=1}^{N_r}\eta_{ar_1}\eta^*_{br_1}\eta_{cr_1}\eta^*_{dr_1}+\mathcal{O}(1/\sqrt{N_r}) \\
=& \delta_{ab}\delta_{cd}+\frac{1}{N_r}\delta_{ad}\delta_{cb}(1-\delta_{ab}\delta_{cd})+\mathcal{O}(1/\sqrt{N_r}),
\end{aligned}
$$

Tensor network rep. of Z

depends on property of field and interaction

- Scalar field (non-compact)
  - Orthonormal basis expansion

    Shimizu mod.phys.lett. A27,1250035(2012), Lay & Rundnick PRL88,057203(2002)
  - Gauss Hermite quadrature Sakai et al., JHEP03(2018)141

    We will see later!
- Gauge field (compact : SU(N), CP(N) etc.)
  - Character expansion : maintain symmetry, better convergence

    Meurice et al., PRD88,056005(2013)
- Fermion field (Dirac/Majorana)

  Shimizu & Kuramashi PRD90,014508(2014), ST & Yoshimura PTEP(2015)043B01
  - Grassmann number $\theta^2=0$ -> finite sum

$$e^{\phi\theta} = 1 + \phi\theta = \sum_{n=0}^{1} (\phi\theta)^n$$

In principle, we can treat any fields