Tensor Network Skeletonization

Lexing Ying
Department of Mathematics
Stanford University
Big Data Meets Computation, IPAM, UCLA, 2/1/17

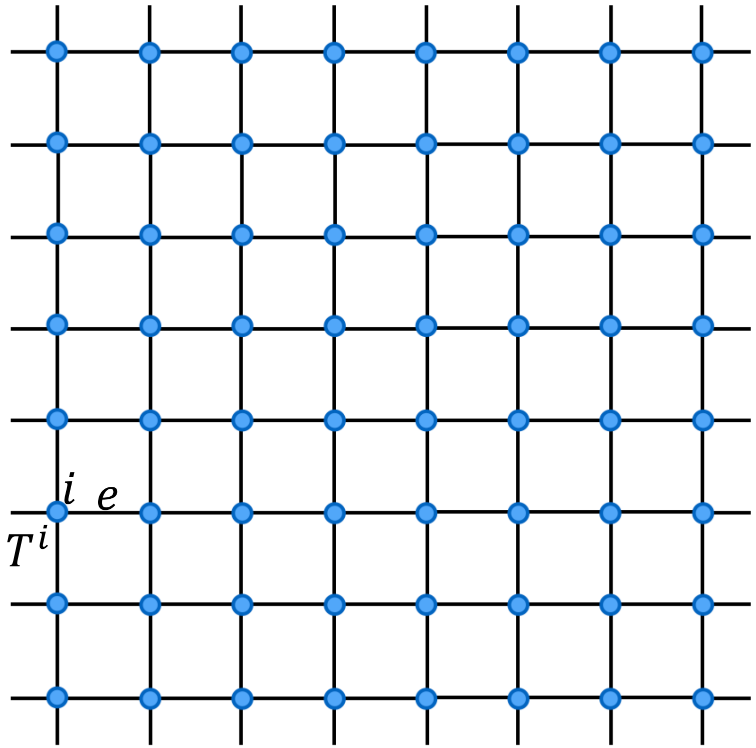
What are tensor networks?

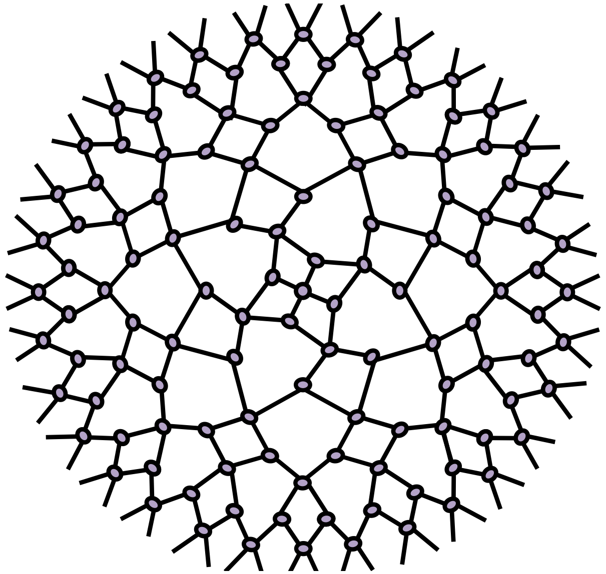
What are tensor networks

- A triple $\left(V, E, \{T^i\}_{i \in V}\right)$
    - $V = \{i\}$: the vertex set
        - d_i: degree
    - $E = \{e\}$: the edge set
        - χ_e: bond dim
        - $E = E_I \cup E_B$ (int. vs. bdry.)
    - T^i is a d_i-tensor at $i \in V$
        - dims $= \{\chi_e\}$ of adj. edges
- Represent a function

$$\mathrm{tr}_{E_I}\left(\bigotimes_{i \in V} T^i\right)$$

    - All edges in E_I are contracted
    - A high-dim function or a number when $E_B = \emptyset$



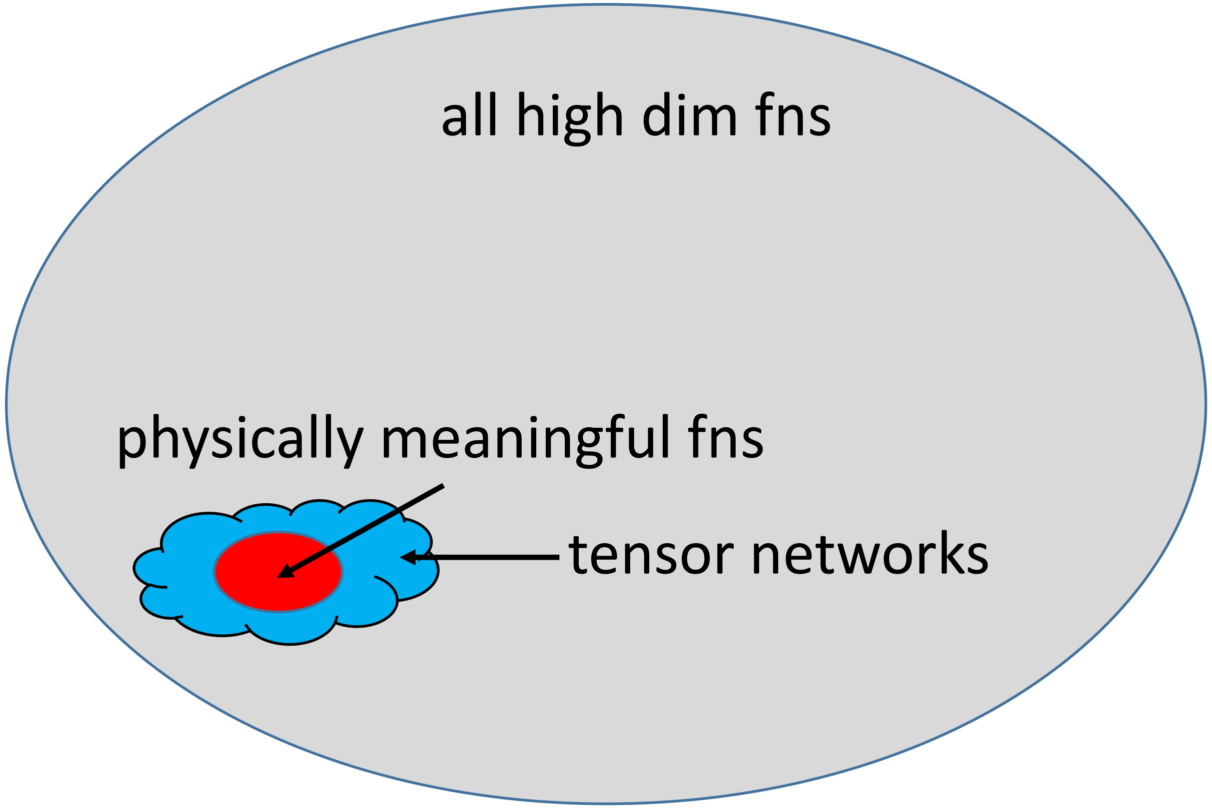


Why tensor networks

- Represent high-dim functions and probability distributions

- Existing methods
  - Independent or weakly-dependent
  - Mean field approach
  - Limited power for “hard or critical” physical systems

- Tensor networks
  - Powerful and flexible
  - Encode geometry



- Examples
  - DMRG/MPS for 1D quantum ground state
  - PEPS for 2D quantum ground state

Example: 2D statistical Ising model

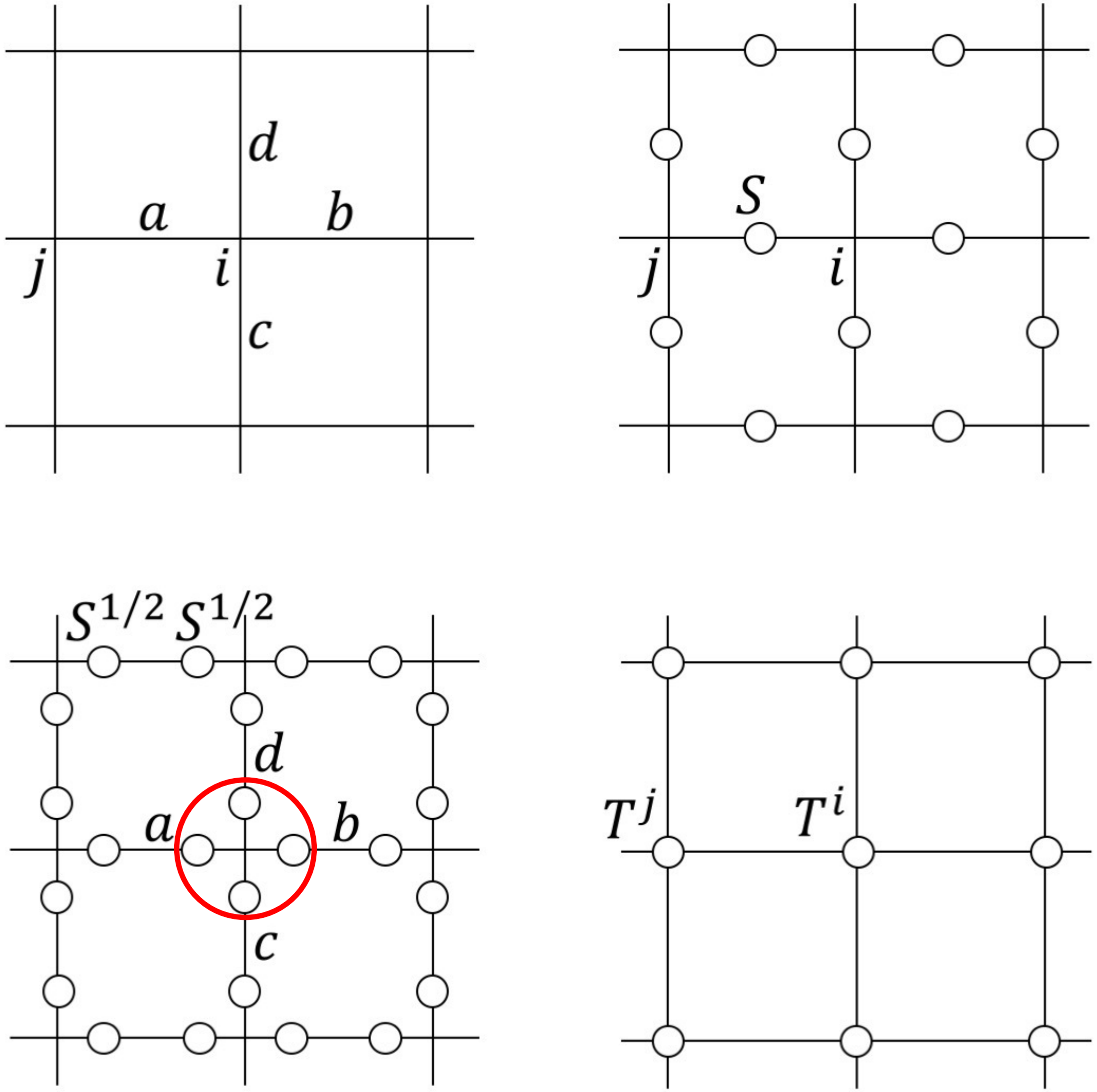
$$S = \begin{pmatrix} e^{\beta} & e^{-\beta} \\ e^{-\beta} & e^{\beta} \end{pmatrix}$$

- Ising model with $N = n \times n$ spins on a Cartesian grid
  - $\sigma = \{\sigma_1, \ldots, \sigma_N\}$: spin configuration ± 1
  - $\beta = 1/T$: inverse temperature

$$Z_N(\beta) = \sum_{\sigma} e^{-\beta H_N(\sigma)}$$

$$H_N(\sigma) = -\sum_{(ij)\in E} \sigma_i \sigma_j$$

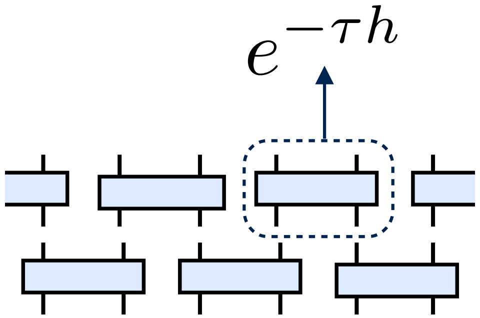
- Tensor network
  - V: sites of Cartesian grid
  - E: edges between neighbors (periodic)

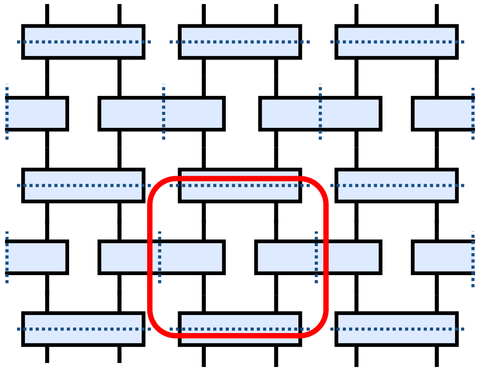


Example: 1D quantum Ising model

- 1D spin chain of length N
- Hamiltonian H is local

$$H = \sum_r h_{r,r+1}.$$





- Ground states

$$|\Phi_0\rangle \sim \lim_{\beta\to\infty} e^{-\beta H}|\Phi_{any}\rangle$$

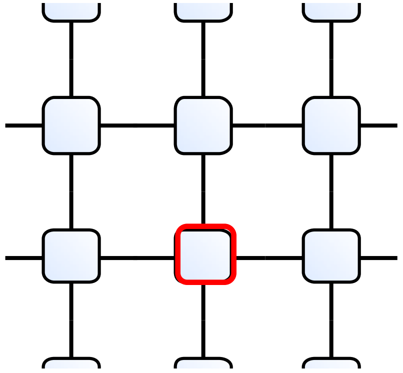
- Suzuki-Trotter decomposition

$$e^{-\beta H} = \left(e^{-\tau H}\right)^{(\beta/\tau)}$$

$$e^{-\tau H} \approx e^{-\tau H_{\mathrm{odd}}} e^{-\tau H_{\mathrm{even}}}$$

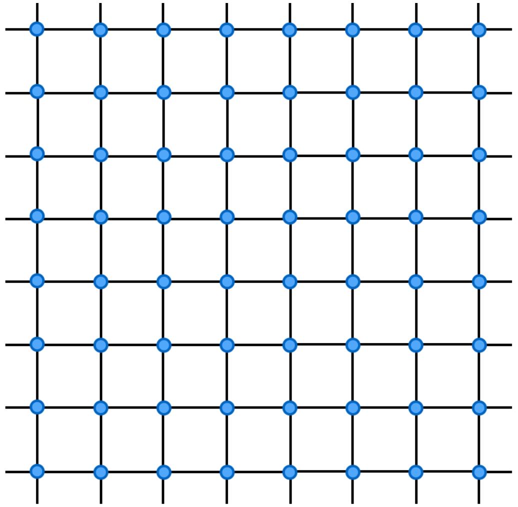
$$e^{-\tau H_{\mathrm{odd}}} = \prod_{\mathrm{odd}\ r} e^{-\tau h_{r,r+1}}$$

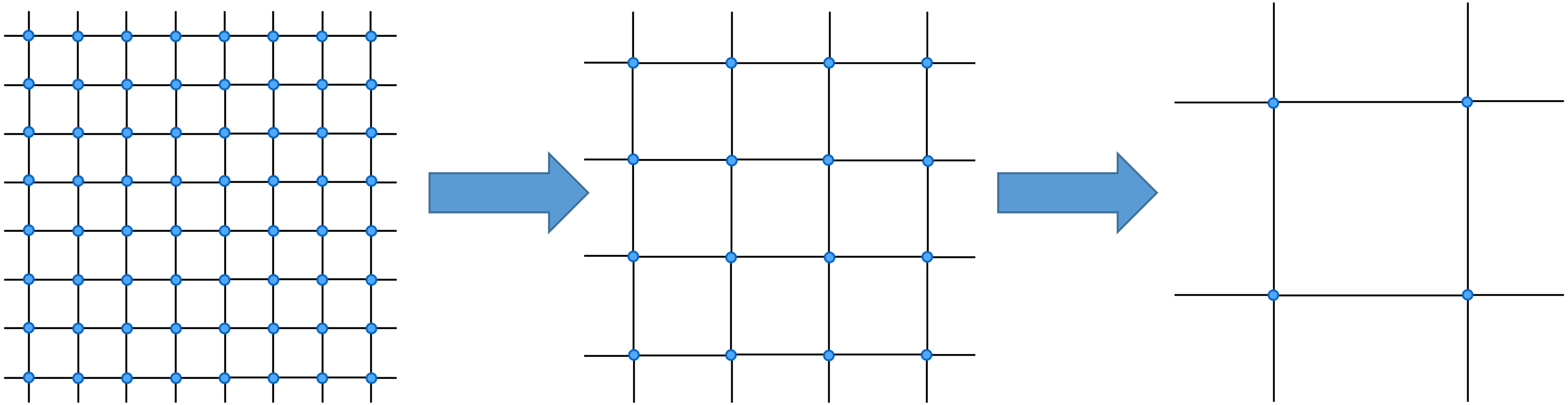
$$e^{-\tau H_{\mathrm{even}}} = \prod_{\mathrm{even}\ r} e^{-\tau h_{r,r+1}}$$



This talk

- Computation of tensor networks
    - New approach: tensor network skeletonization (TNS)
    - An exercise to understand what physicists have done and try to improve on them
- Example: 2D statistical Ising model $Z_N(\beta) = \mathrm{tr}_E \left(\bigotimes_{i \in V} T^i \right)$
    - Critical behavior at $\beta_c = 1/T_c$
- Main theme: renormalization or upscaling

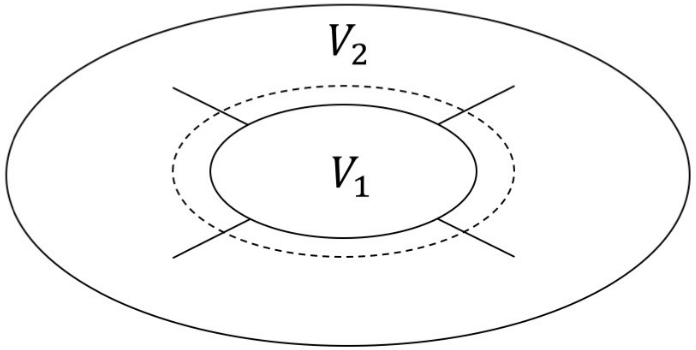




Basic tool

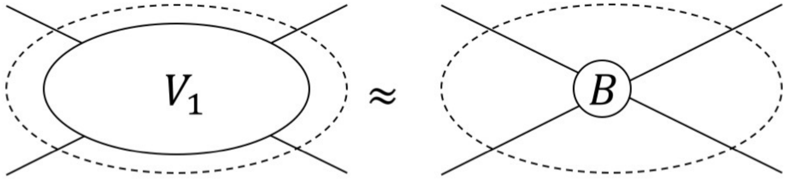
- Local replacements

$$V = V_1 \cup V_2, \quad E = E_1 \cup E_2 \cup E_{12},$$

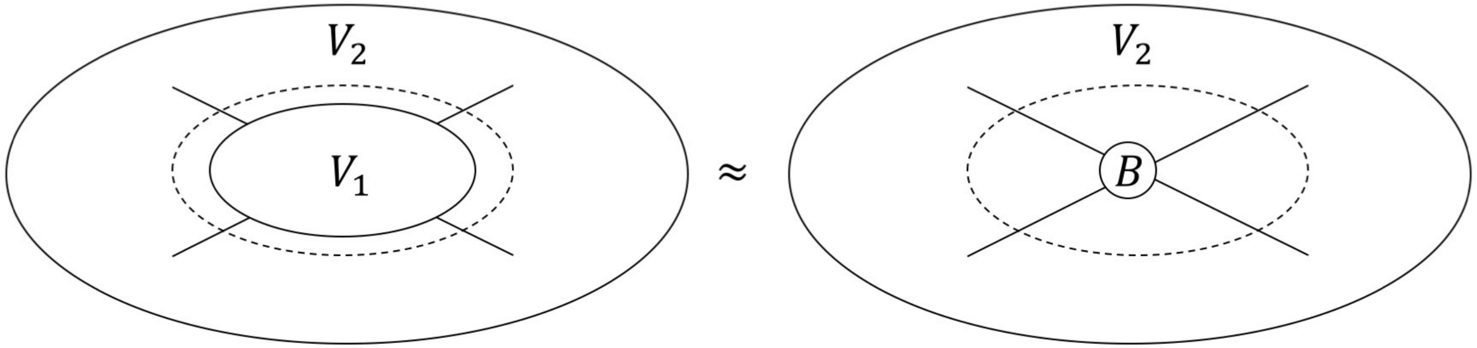


$$\mathrm{tr}_E \left(\bigotimes_{i \in V} T^i \right) = \mathrm{tr}_{E_2 \cup E_{12}} \left(\left(\bigotimes_{i \in V_2} T^i \right) \bigotimes \mathrm{tr}_{E_1} \left(\bigotimes_{i \in V_1} T^i \right) \right)$$

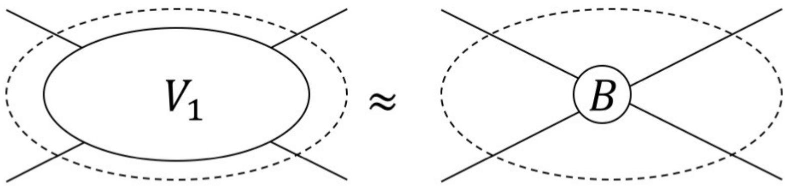
$$B \approx \mathrm{tr}_{E_1} \left(\bigotimes_{i \in V_1} T^i \right)$$



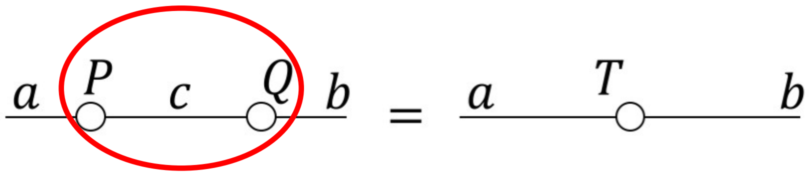
$$\mathrm{tr}_E \left(\bigotimes_{i \in V} T^i \right) \approx \mathrm{tr}_{E_2 \cup E_{12}} \left(\left(\bigotimes_{i \in V_2} T^i \right) \bigotimes B \right)$$



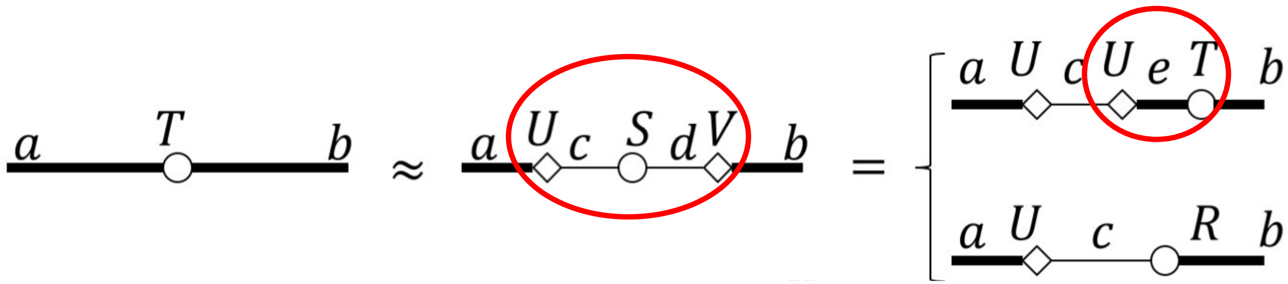
Local replacements



- Tensor contraction

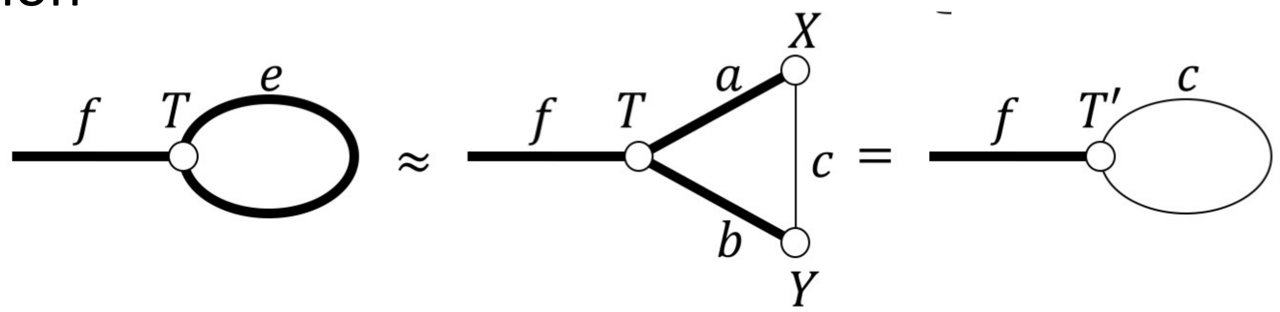


- Projection

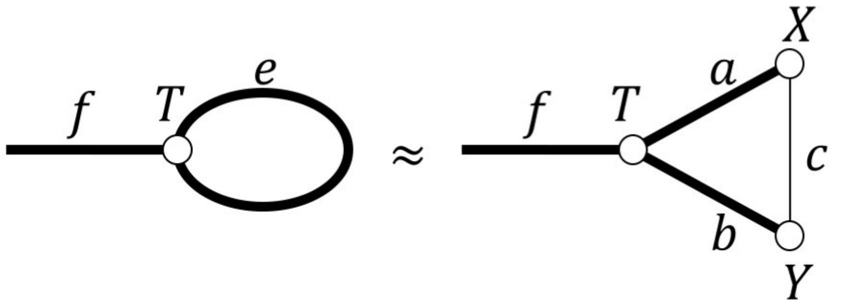


$R = U'T$

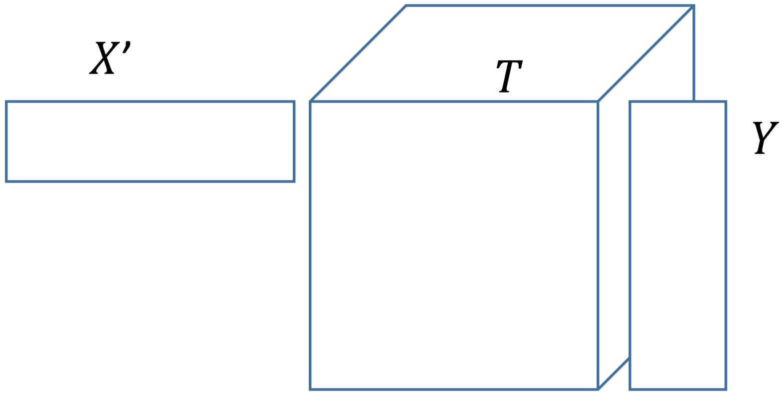
- Structure-preserving skeletonization
  - Loop simplification
  - Key to TNS



Structure-preserving skeletonization (SPS)



$$\operatorname{tr}_e T \approx \operatorname{tr}_{abc}(X \otimes T \otimes Y), \quad \sum_e T_{eef} \approx \sum_{abc} X_{ac} T_{abf} Y_{bc},$$



- Key to removing short-range correlation in a loop
- In matrix notation:

$$\operatorname{tr}_e T \approx \operatorname{tr}_c(X^* T Y)$$

- Solve a regularized problem

$$\min_{X,Y} \| \operatorname{tr}_e T - \operatorname{tr}_c(X^* T Y)\|_2^2 + \alpha(\|X\|_F^2 + \|Y\|_F^2),$$

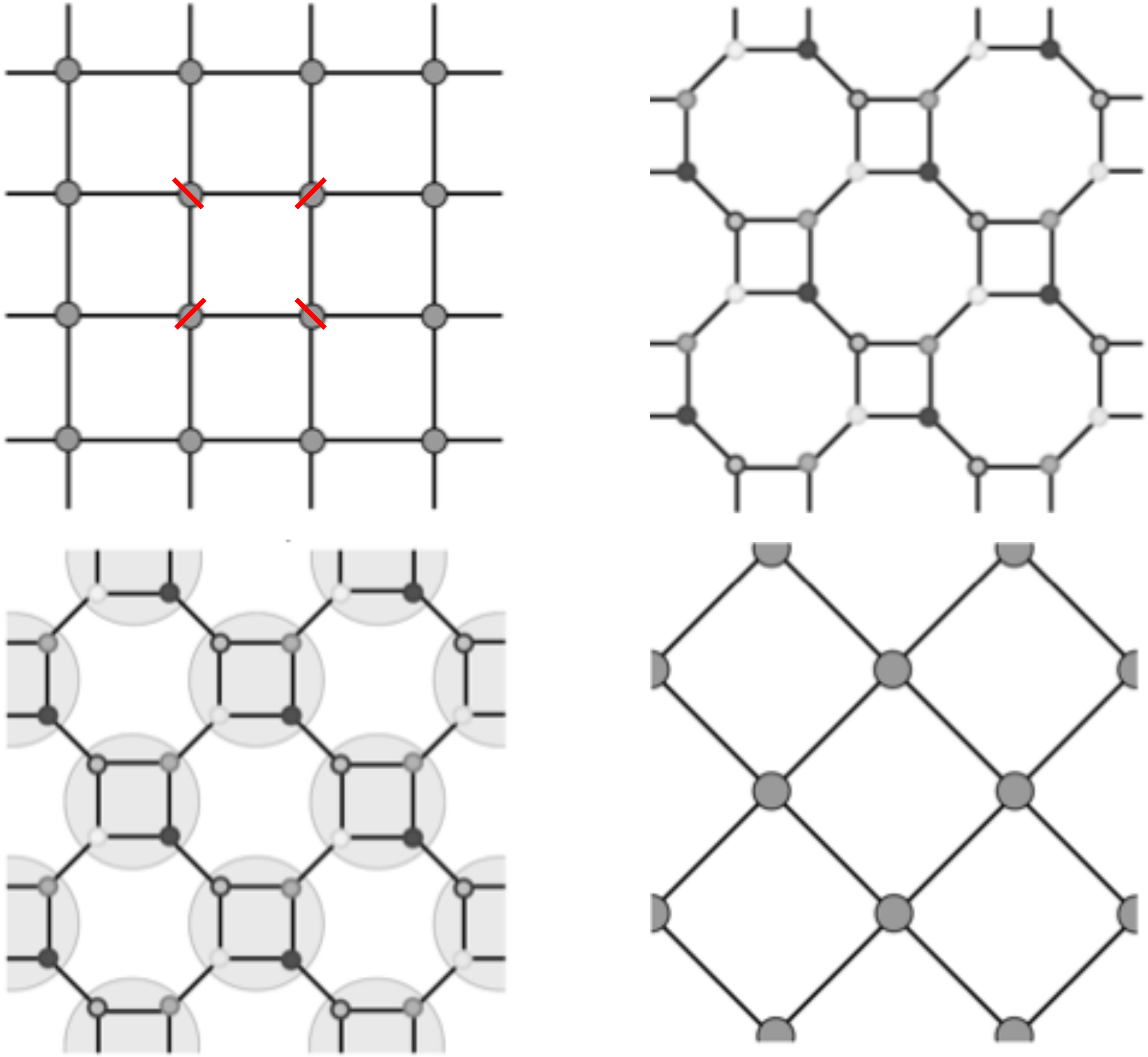
- Use alternating least square algorithm with reasonable initial guess (W. Yin)

$$X^{(n+1)} = \operatorname{argmin}_X \| \operatorname{tr}_e T - \operatorname{tr}_c(X^* T Y^{(n)})\|_2^2 + \alpha \|X\|_F^2$$
$$Y^{(n+1)} = \operatorname{argmin}_Y \| \operatorname{tr}_e T - \operatorname{tr}_c((X^{(n+1)})^* T Y)\|_2^2 + \alpha \|Y\|_F^2.$$

Previous work

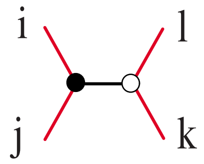
Previous work: Tensor renormalization group (TRG)

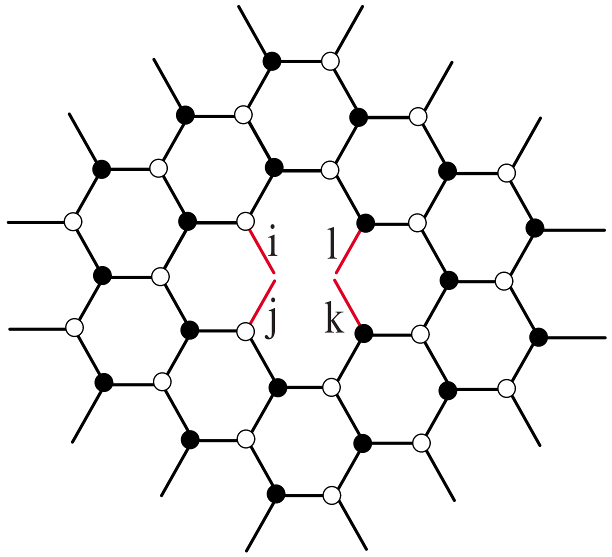
- Levin & Nave (2007)
  - Arguably the first practical algorithm for (>1D) tensor networks
  - Rotate by 45 degrees
  - Cannot remove corner-double-line (CDL) tensors
  - Cannot remove short-range correlation

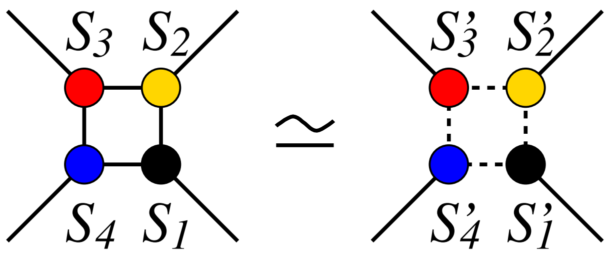


Previous work

- Second renormalization group
  - Xie et al (2009), Xiang's group at CAS
  - Considers environment effect when performing projection
  - Extends to 3D models
- Tensor-entanglement-filtering renormalization (TEFR)
  - Gu & Wen (2009)
  - Points out the importance of removing correlation in loops
  - Simple iterative algorithms
  - 2D systems, rotate 45 degrees

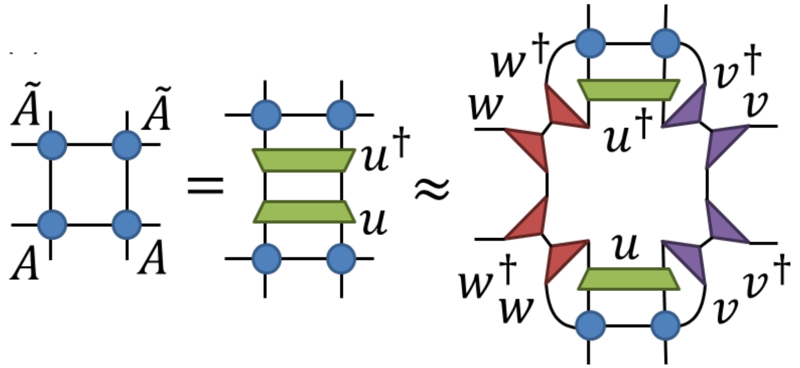


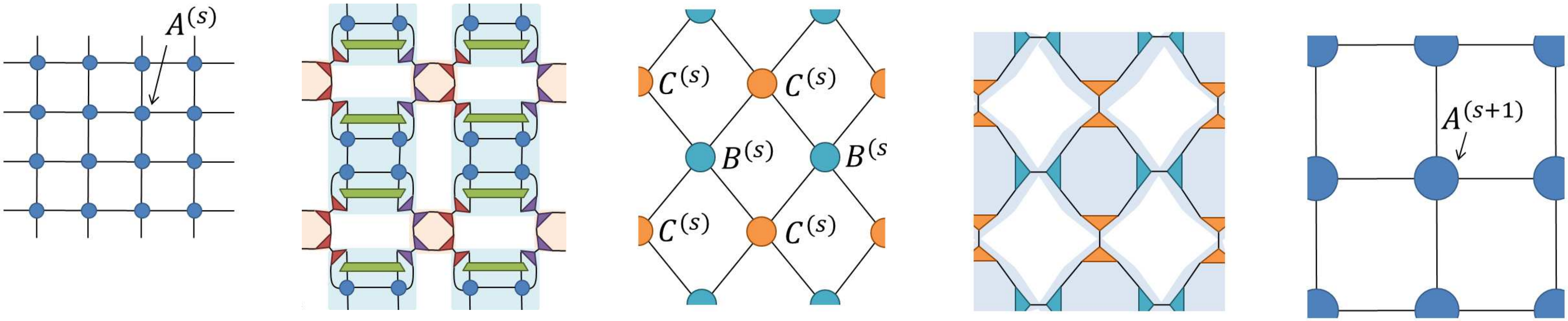




Recent work: Tensor network renormalization (TNR)

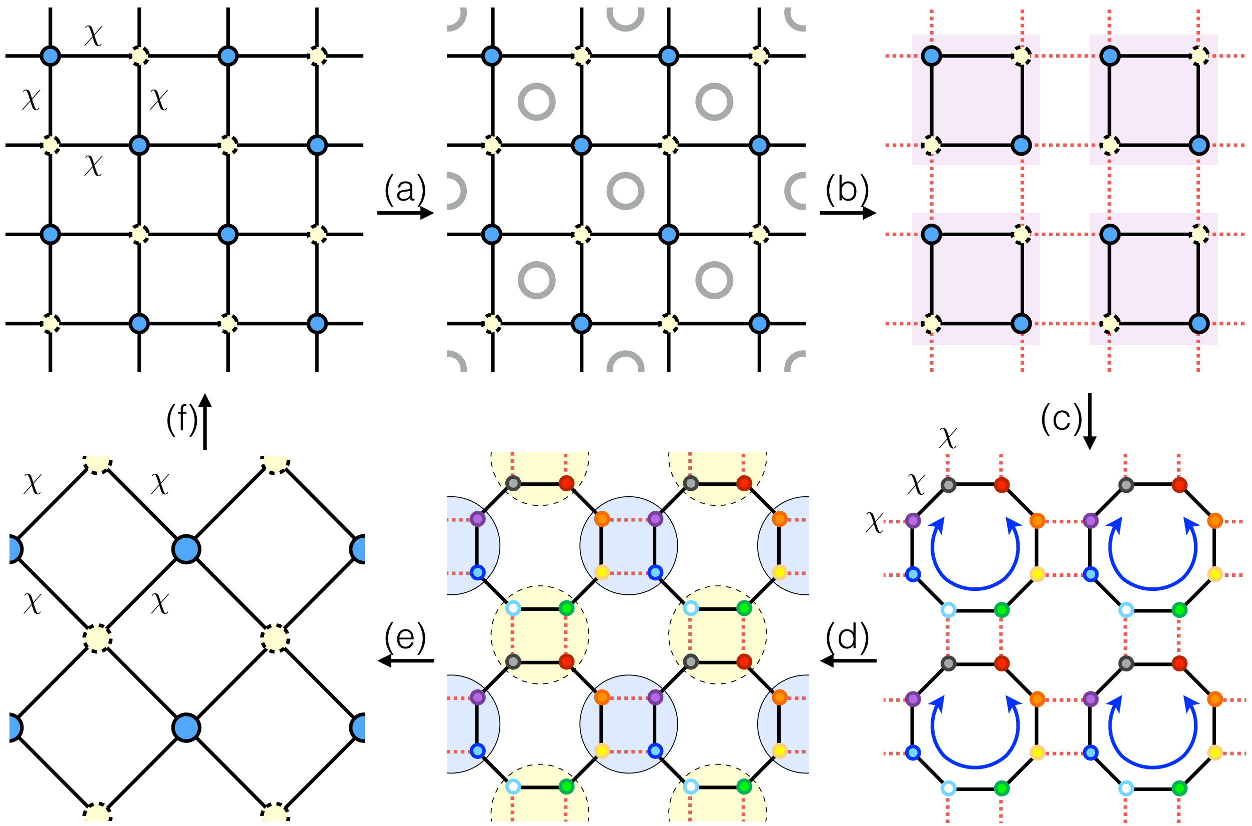
- Evenbly & Vidal (2015)
  - Apply disentangler to remove short-range correlation in loops
  - 2D systems: rotate by 45 degrees
  - Related with MERA, AdS/CFT correspondence, holograph principle





Very recent work: Loop-TNR

- Yang, Gu, & Wen (2016)
- Built on top of TEFR
- Remove short-range correlation in loop more effectively
- Use iterative QR algorithms and MPS algorithms for optimization



Comparison

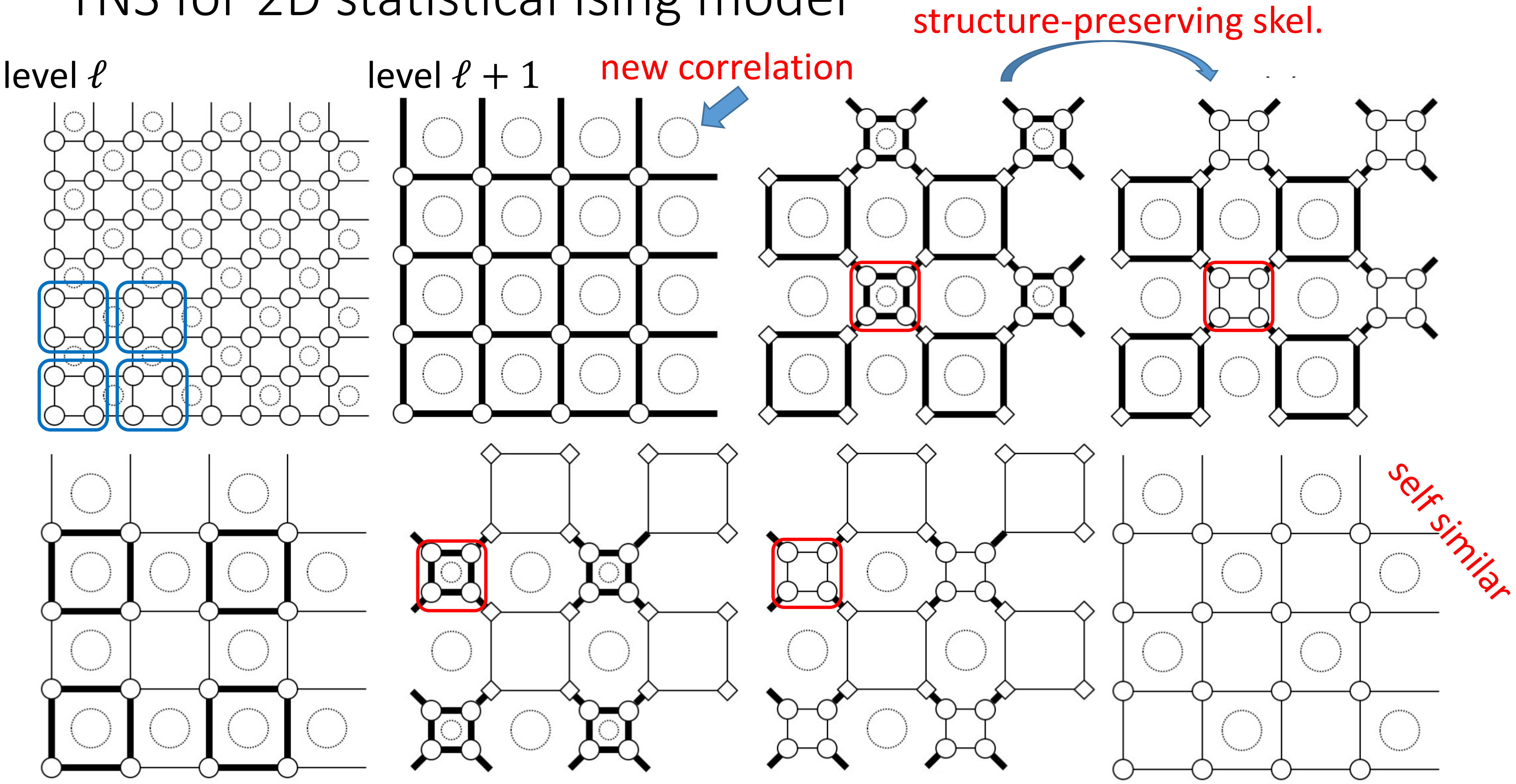
	Off-critical	Critical	Higher dimensions
TRG	✓		
SRG	✓		✓
TEFR	✓		
TNR	✓	✓	
Loop-TNR	✓	✓	?
TNS	✓	✓	✓

Tensor network skeletonization

Tensor network skeletonization (TNS)

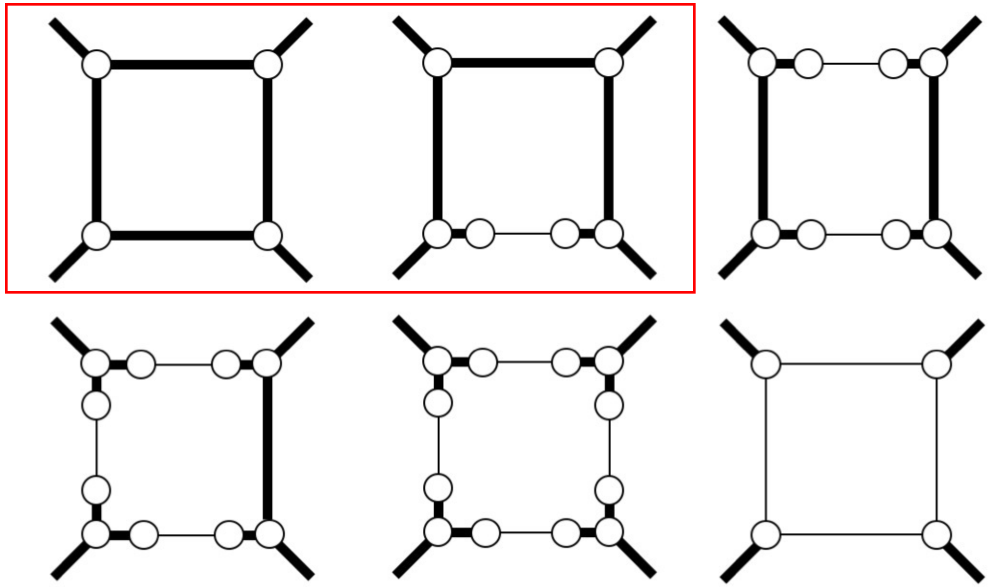
- A renormalization/upscaling approach
- Structure-preserving skeletonization for removing short-range correlation in loops
- Preserve the Cartesian structure (hence the name ”skeletonization”)
- Extends easily to high dimensional models

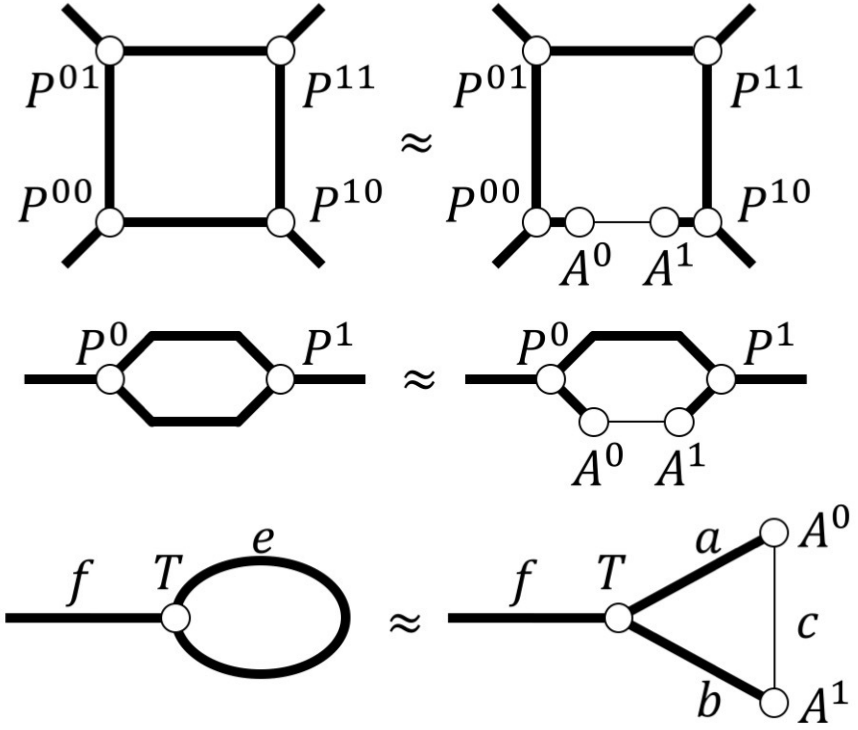
TNS for 2D statistical Ising model



SPS for short-range correlation removal

- For a loop strucutre
  - Consider edges one by one
  - For each edge, insert two tensors on the edge and reduce bond dimension

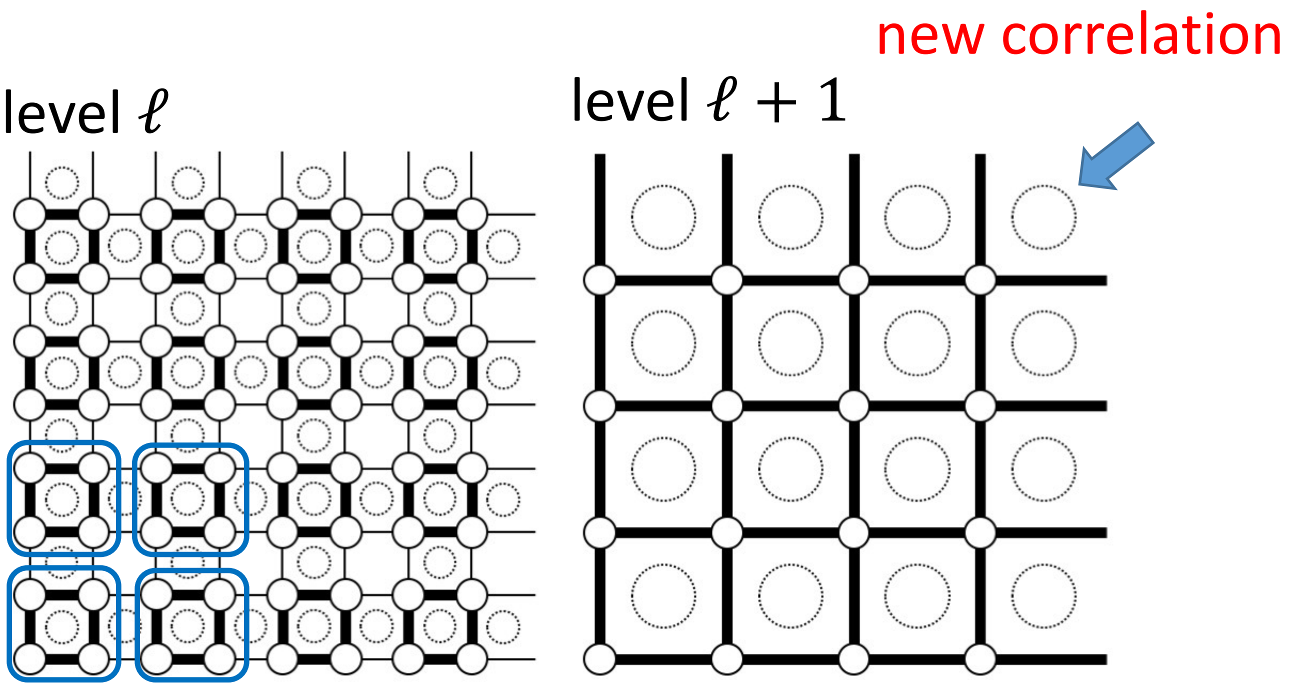


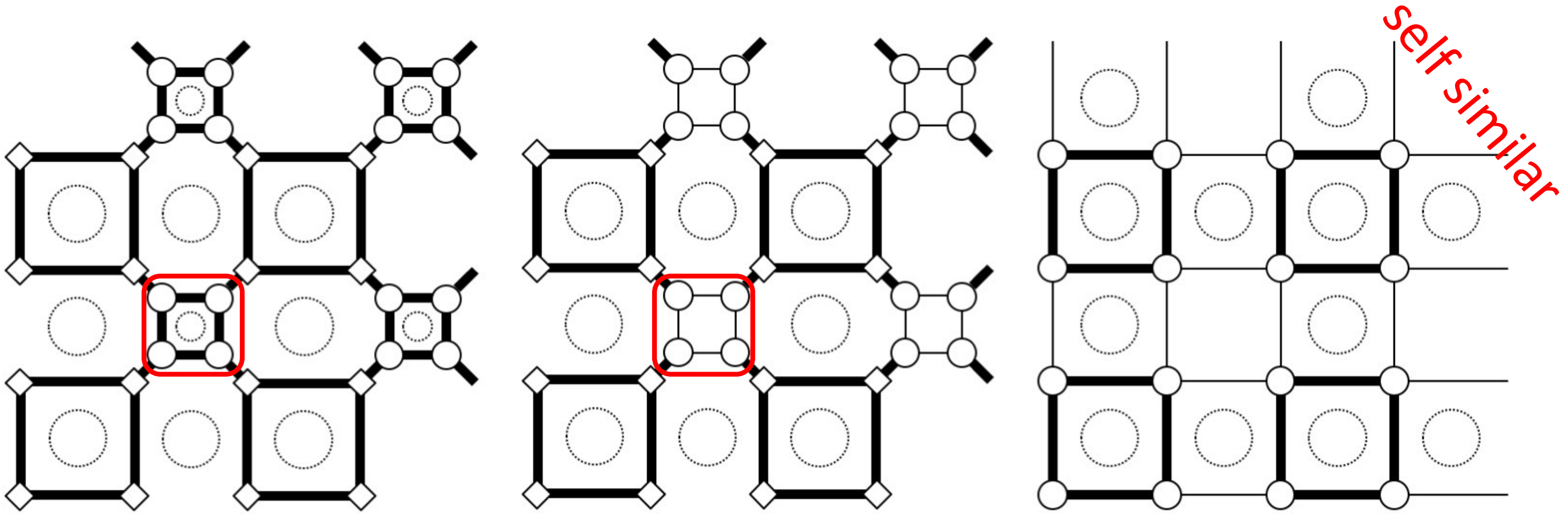


Use structure-preserving skeletonization

A modified version

- When we only care about evaluating the partition function
- Simpler and sometimes faster

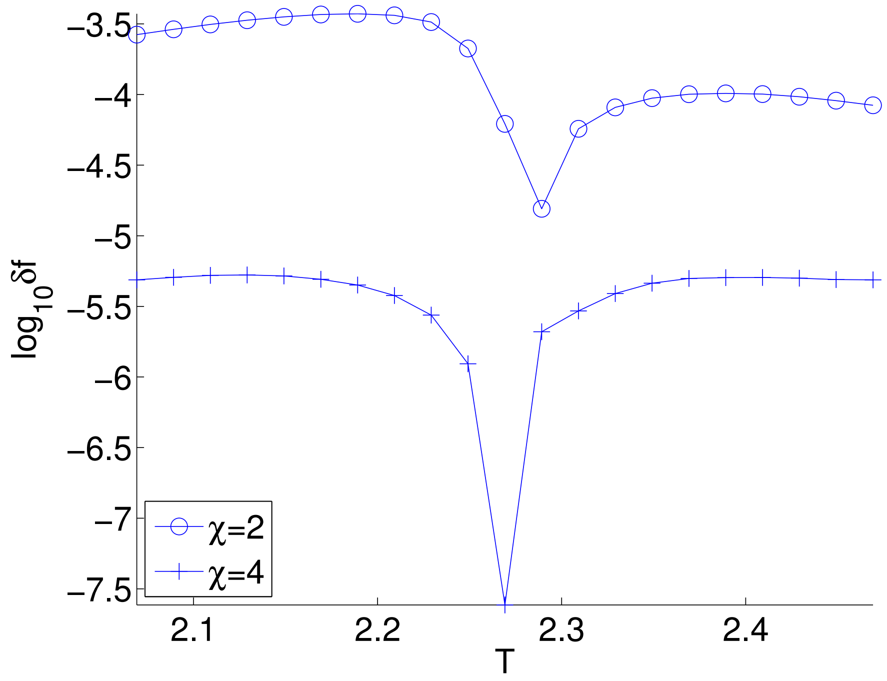




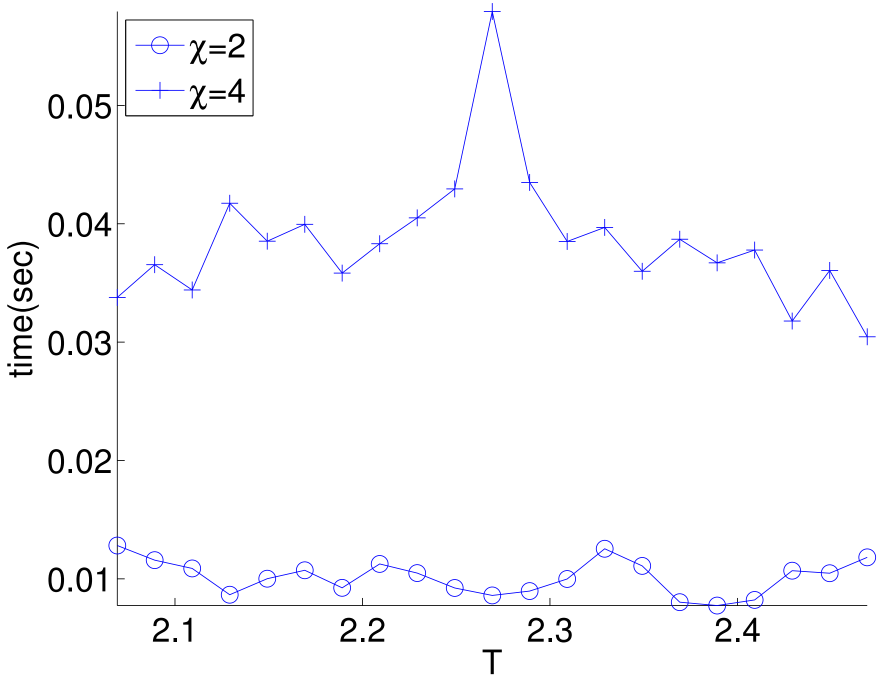
Partition function (2D statistical Ising model)

- $N = n \times n$ periodic Cartesian lattice, up to $n = 2^{15}$
- Free energy per site

$$f_N(\beta) = \left(-\frac{1}{\beta}\log Z_N(\beta)\right)/N \qquad T_c = 2/\ln(1+\sqrt{2})$$



relative error



time per iteration

Observables

- Average magnetization

$$\langle \sigma_i \rangle_{N,B}(\beta) = \frac{\sum_\sigma \sigma_i e^{-\beta H_{N,B}(\sigma)}}{\sum_\sigma e^{-\beta H_{N,B}(\sigma)}}$$

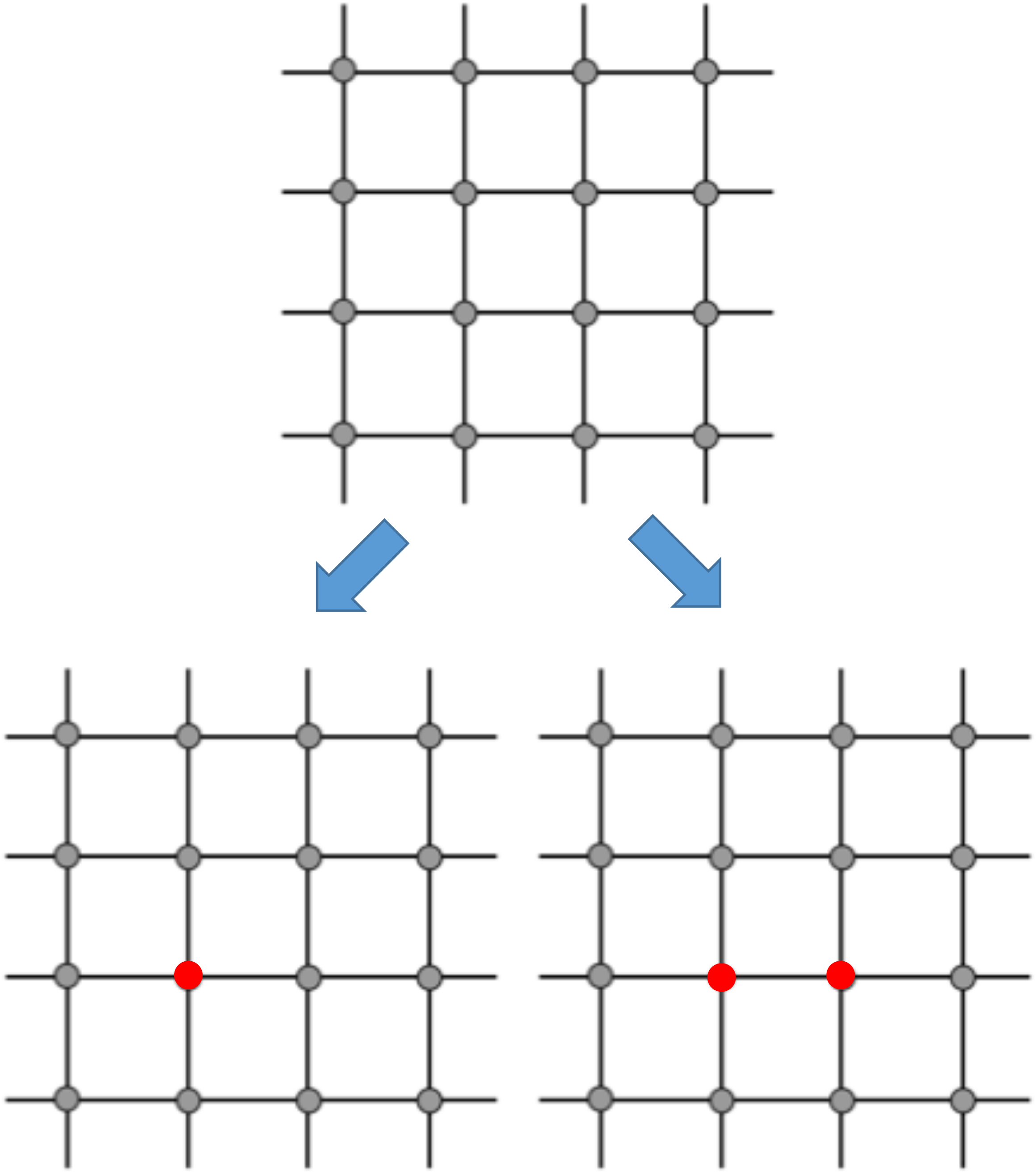
- Internal energy per site

$$u_N(\beta) = 2\frac{\sum_\sigma (\sigma_i \sigma_j) e^{-\beta H_N(\sigma)}}{\sum_\sigma e^{-\beta H_N(\sigma)}}$$

- Requires computation of

$$\sum_\sigma (\sigma_i \sigma_j) e^{-\beta H_N(\sigma)}, \quad \sum_\sigma \sigma_i e^{-\beta H_{N,B}(\sigma)}$$

- Impurity tensor method
    - Modified 1 or 2 tensors in network

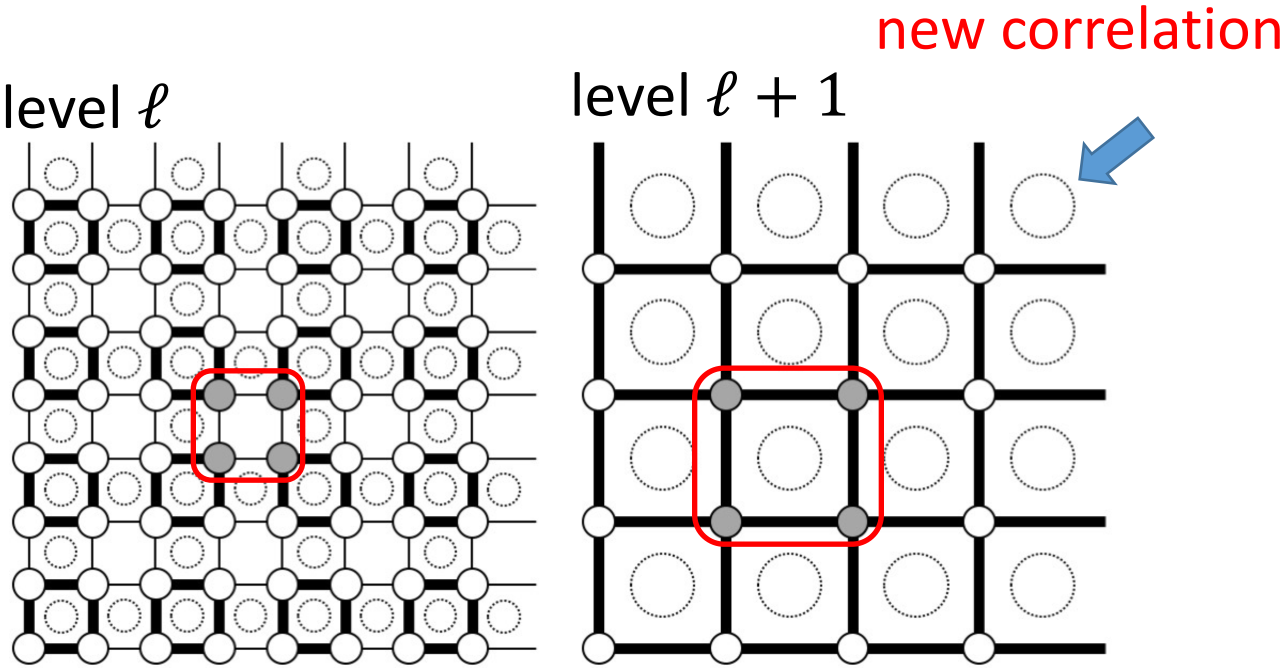


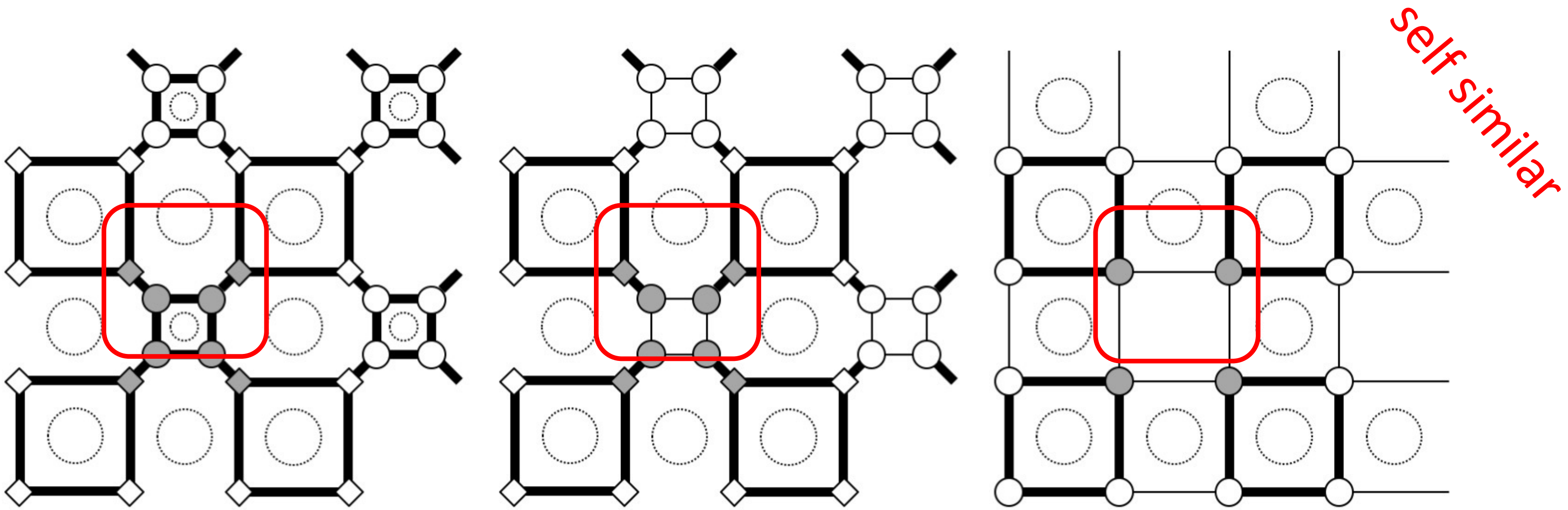
Impurity tensor method

- The modified version works with minimal changes

$$\sum_{\sigma}(\sigma_i\sigma_j)e^{-\beta H_N(\sigma)}, \quad \sum_{\sigma}\sigma_i e^{-\beta H_{N,B}(\sigma)}$$

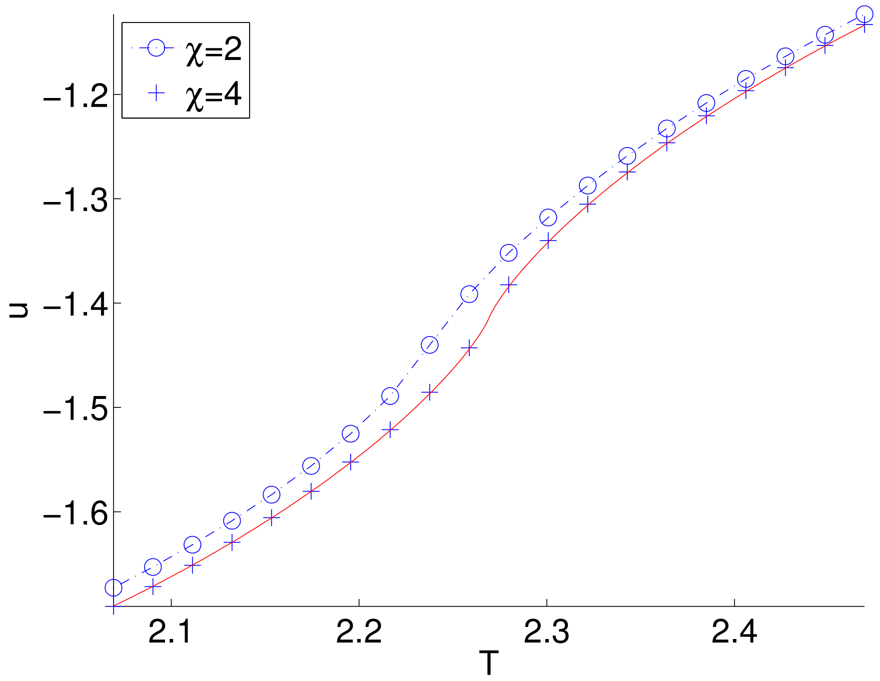
- White tensors exactly the same
- Only gray ones are modified



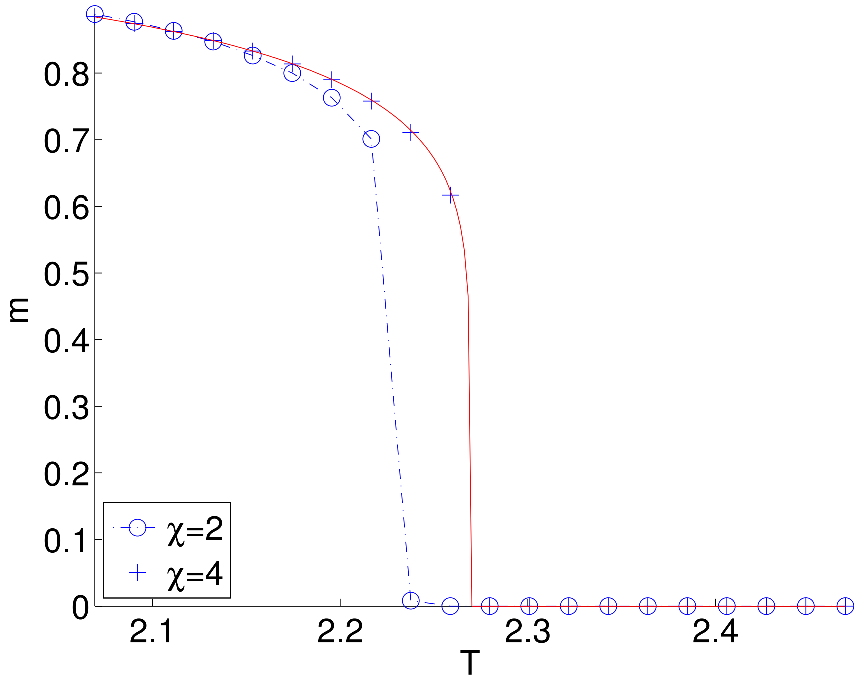


Average magnetization and internal energy

- $N = n \times n$ periodic Cartesian lattice, up to $n = 2^{15}$

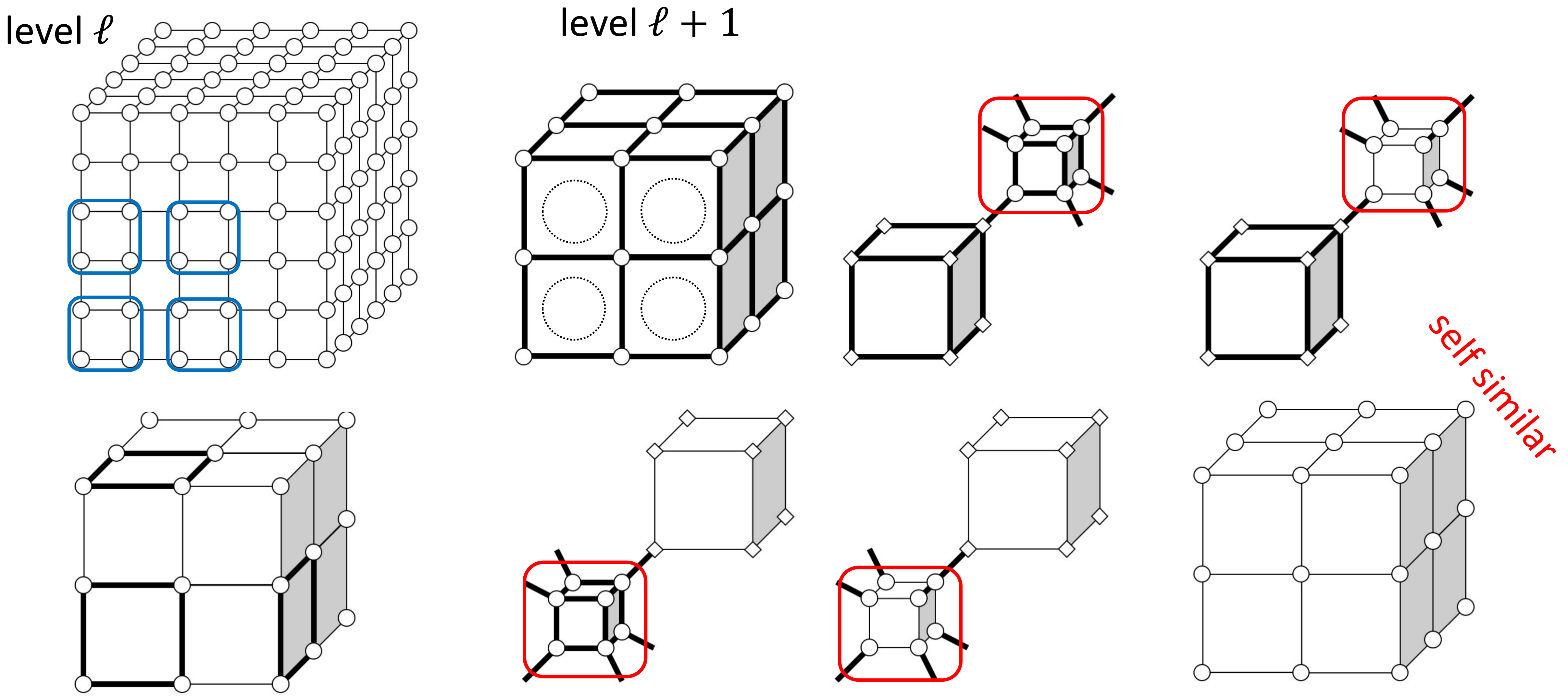


internal energy per site



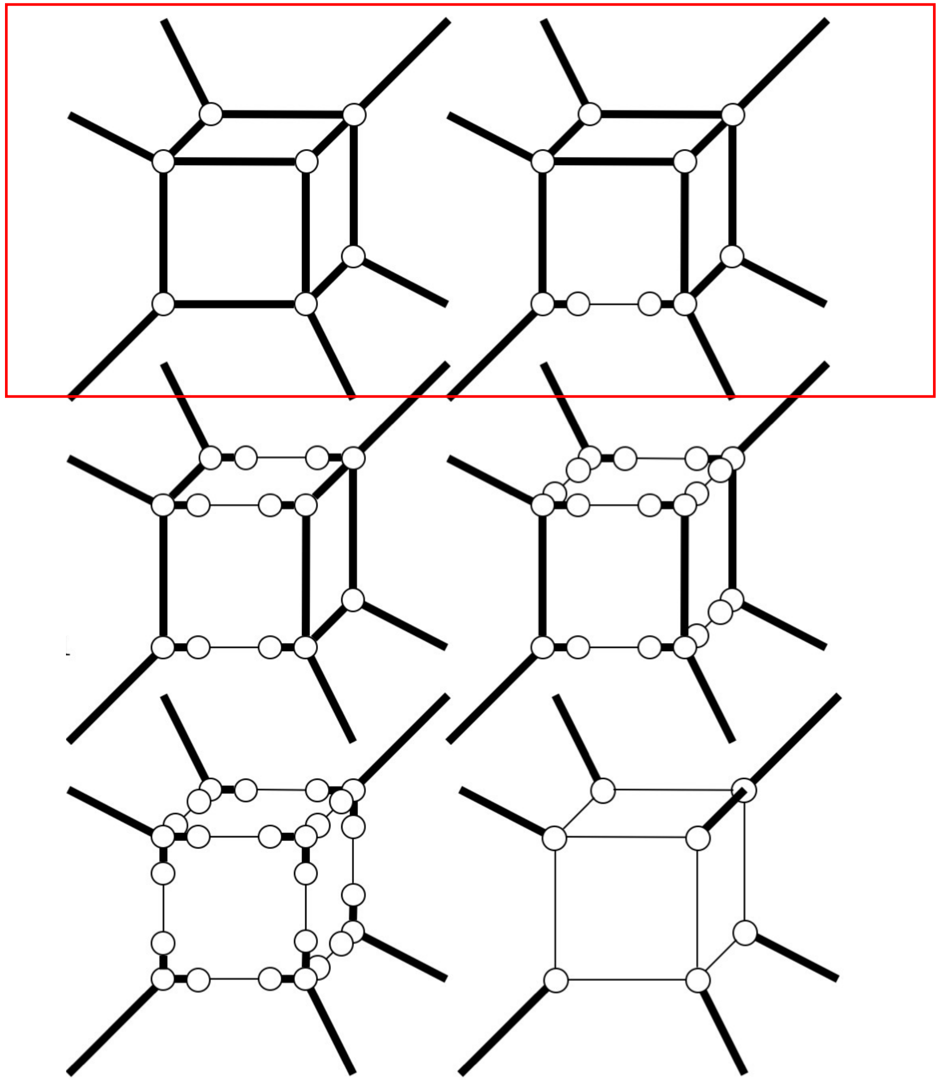
average magnetization

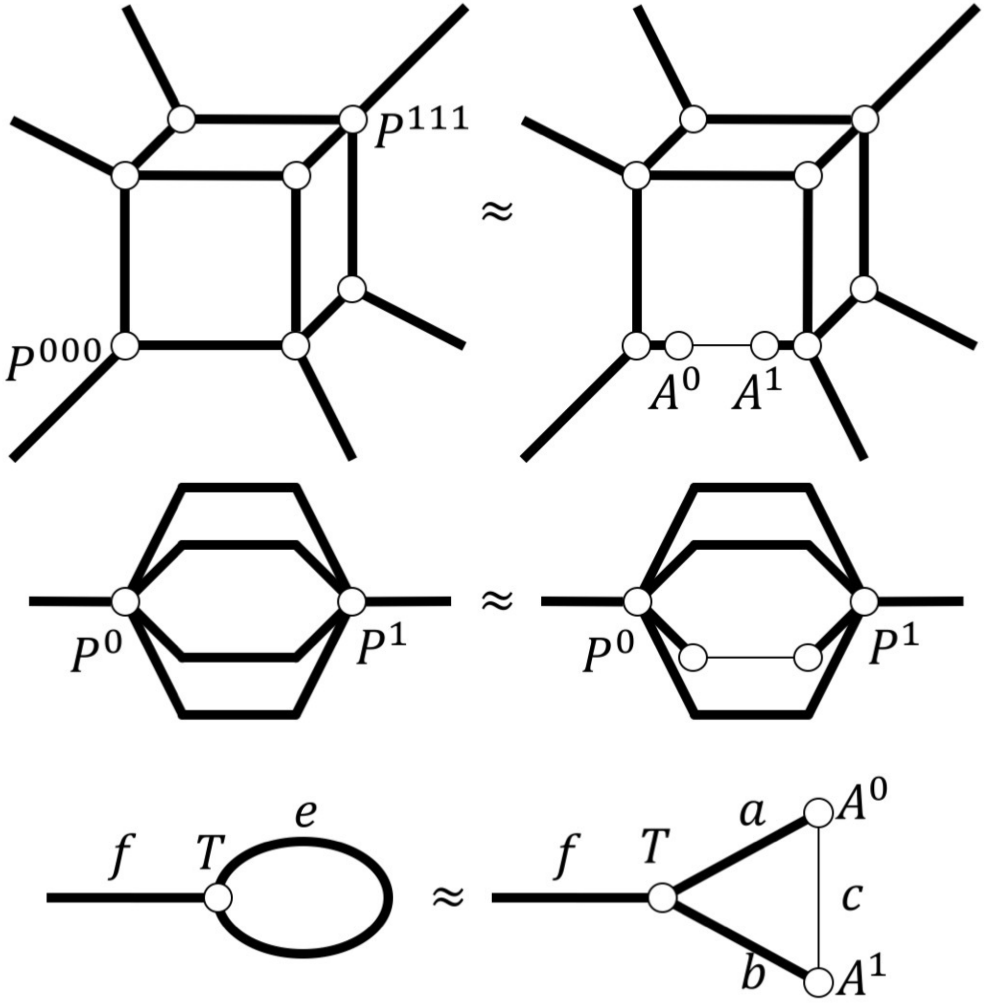
TNS for 3D statistical Ising model



SPS for short-range correlation removal (3D)

- Consider 12 edges one by one





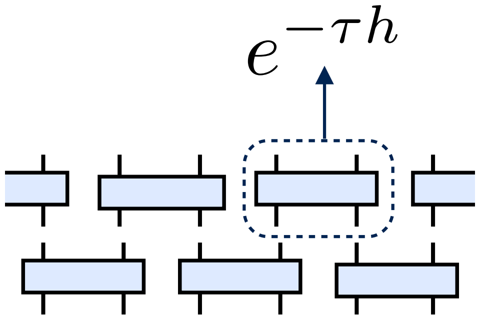
Use structure-preserving skeletonization

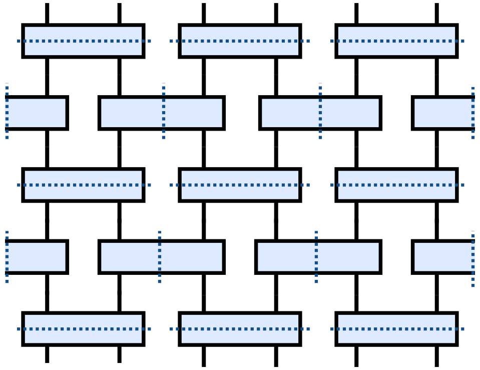
Representing ground states

Ground state of 1D quantum Ising model

- 1D Ising chain
- Hamiltonian H is local

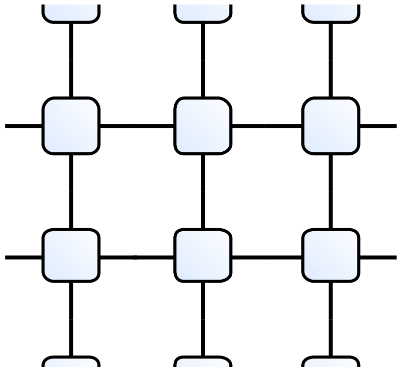
$$H = \sum_r h_{r,r+1}.$$





- Ground states

$$|\Phi_0\rangle \sim \lim_{\beta\to\infty} e^{-\beta H}|\Phi_{any}\rangle$$



- Suzuki-Trotter decomposition

$$e^{-\beta H} = \left(e^{-\tau H}\right)^{(\beta/\tau)}$$

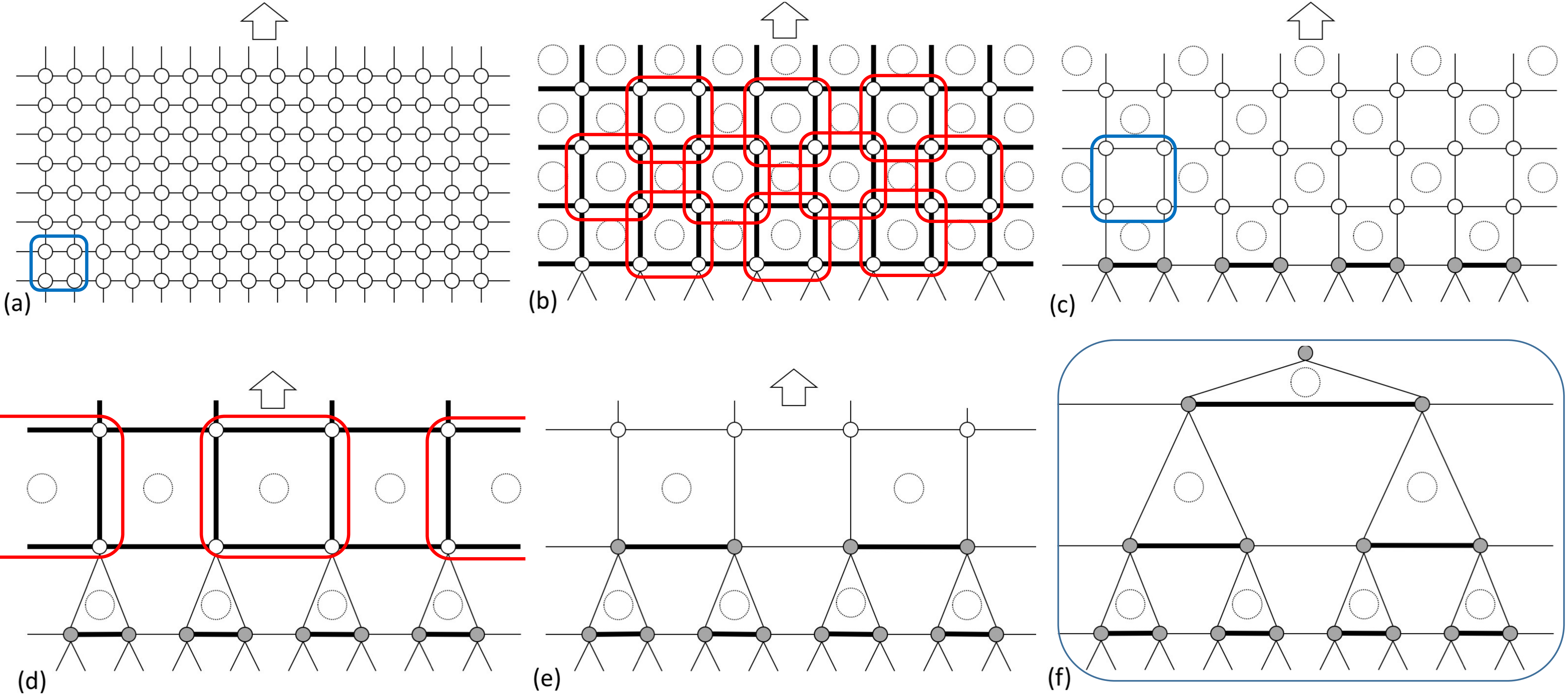
$$e^{-\tau H} \approx e^{-\tau H_{\text{odd}}} e^{-\tau H_{\text{even}}}$$

$$e^{-\tau H_{\text{odd}}} = \prod_{\text{odd } r} e^{-\tau h_{r,r+1}}$$

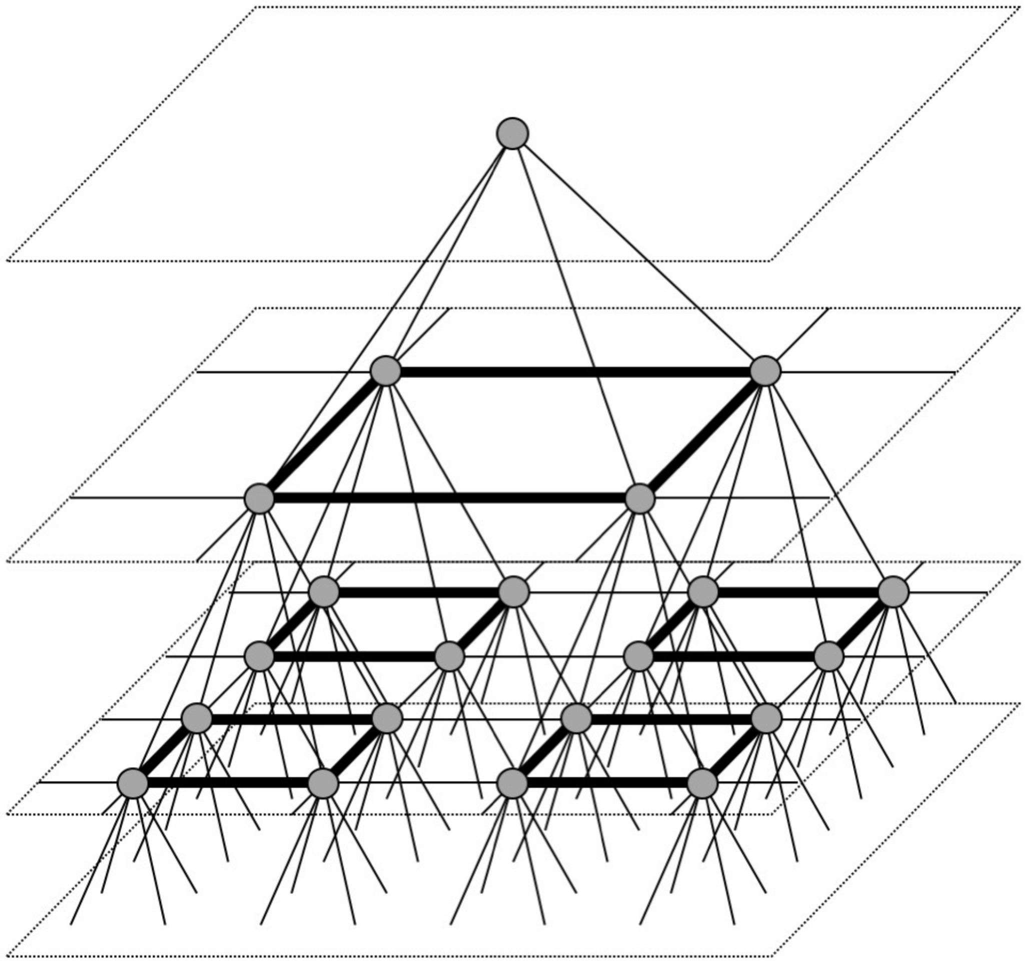
$$e^{-\tau H_{\text{even}}} = \prod_{\text{even } r} e^{-\tau h_{r,r+1}}$$

Ground state through TNS

Similar to how TNR produces MERA



Ground state for 2D quantum Ising model

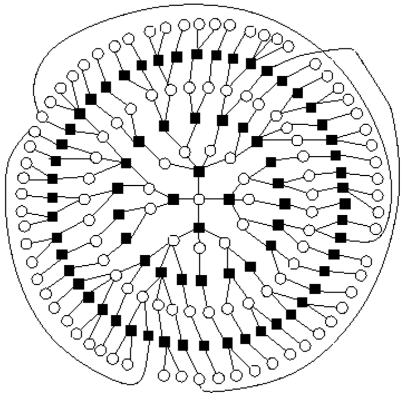


What is next?

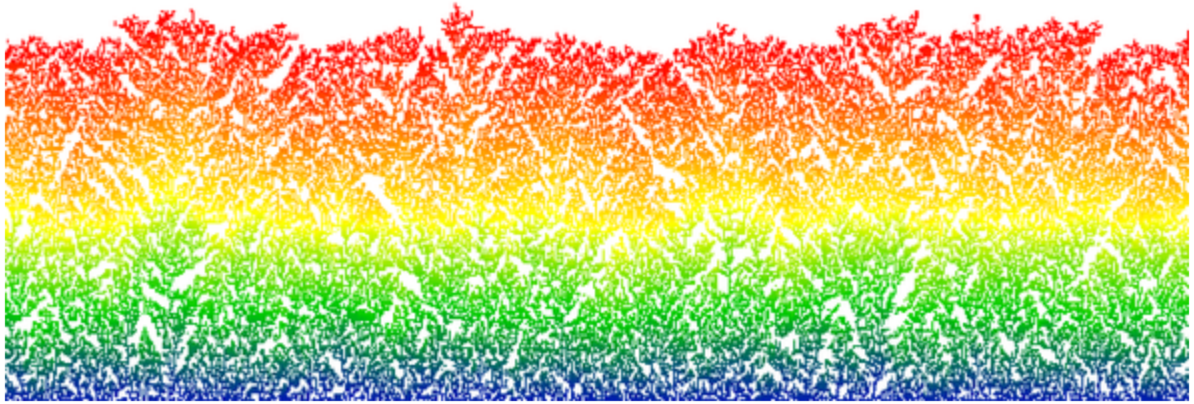
Future work

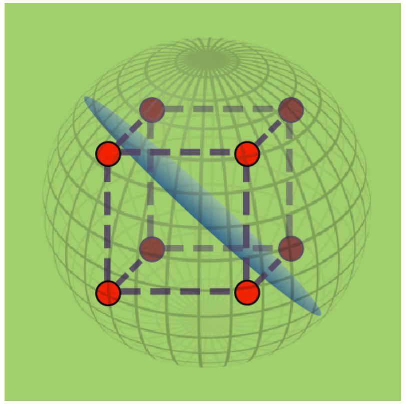
- More efficient implementations for TNS
- Quantuam ground state representations
- Disordered systems
- 4D statistical and 3+1 quantum spin systems

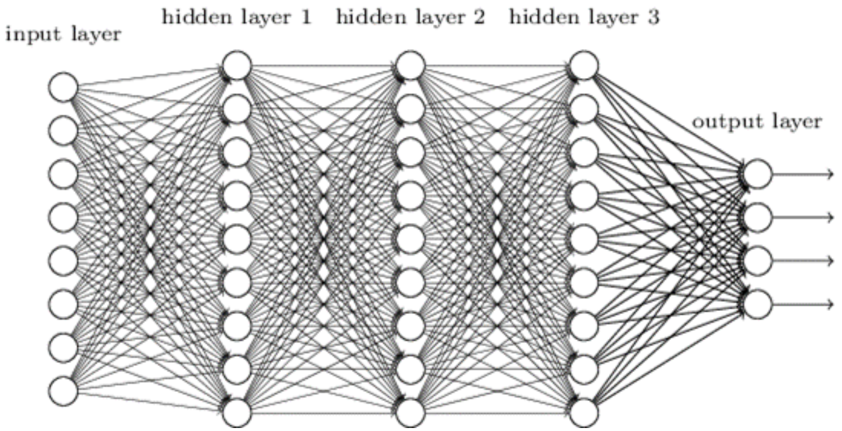
High dimensional functions and probabilities



- Graphical models in machine learning
- Uncertainty quantification
- Connection to deep learning networks
- Discrete/boolean analysis







Summary

- TNS: a new renormalization/upscaling method
- Structure-preserving skeletonization removes short-range correlation
- Extends to 3D, etc.
- Represents ground states effectively

- Reference
  - L. Ying. Tensor network skeletonization, arXiv:1607.00050v1

Thank you