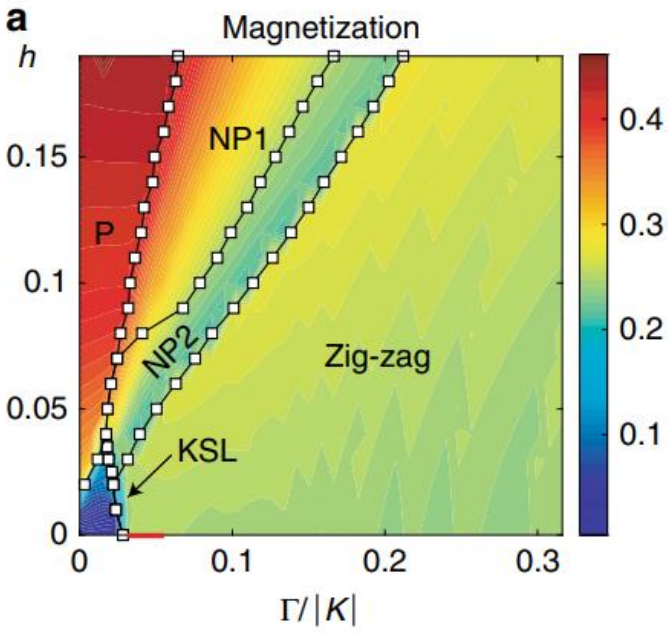
Differentiable programming tensor networks for Kitaev magnets

Xing-Yu Zhang
IOP, CAS

arXiv: 2304.01551
github.com/XingyuZhang2018/TeneT.jl

Motivation

- Kitaev materials
  - α-$RuCl_3$
  - K-Γ and K-J-Γ-Γ'
  - Complex phases
- Computational method
  - QMC sign problem
  - DMRG finite size effect

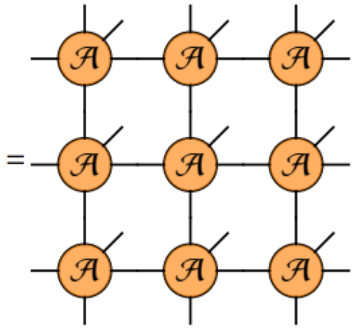


Lee, Hyun-Yong, NC 11.1 (2020): 1639.

Method to 2D quantum system

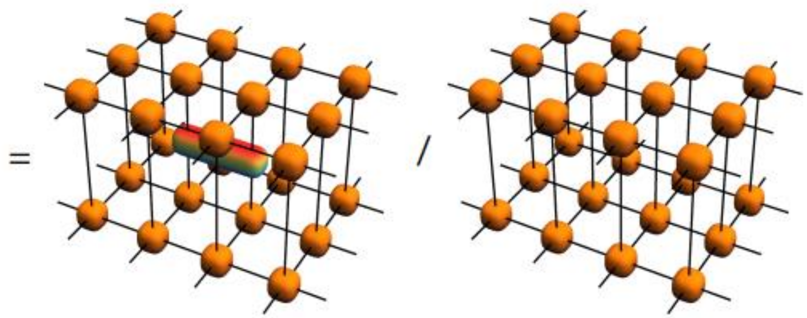
- 2D ansatz

$$|\Psi(\mathcal{A})\rangle = \sum_{\{S_r\}} \mathrm{Tr} \prod_{\mathbf{r}} \mathcal{A}^{S_r}[\mathbf{r}]\, |\{S_r\}\rangle$$



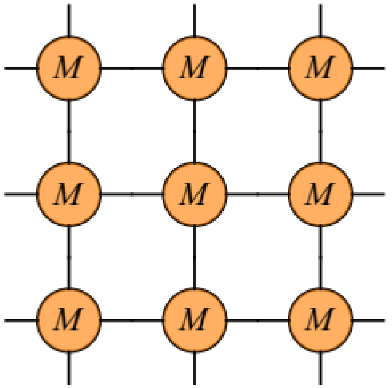
- Energy

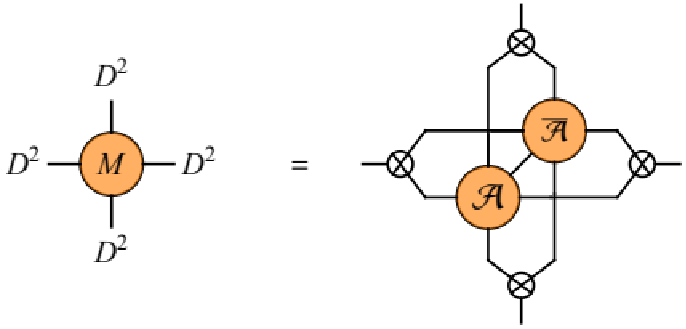
$$E(\mathcal{A}) = \langle \Psi(\bar{\mathcal{A}}) | H | \Psi(\mathcal{A}) \rangle / \langle \Psi(\bar{\mathcal{A}}) \mid \Psi(\mathcal{A}) \rangle$$



- Contraction

$$\langle \Psi(\bar{\mathcal{A}}) \mid \Psi(\mathcal{A}) \rangle =$$



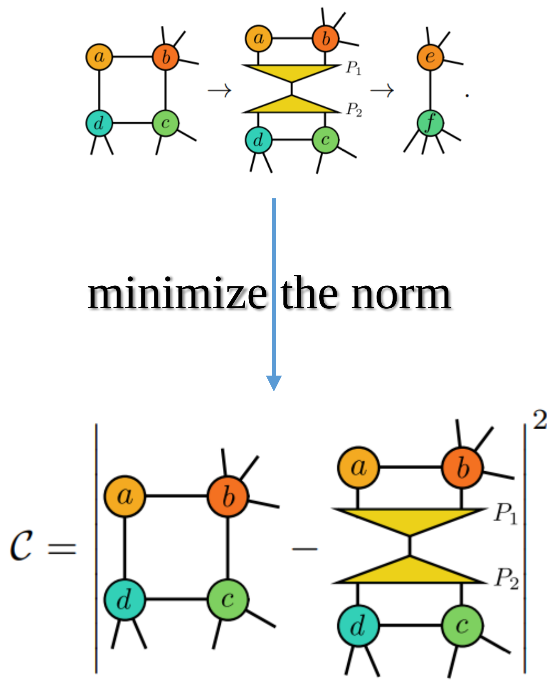


Problems

- Contraction: 2D with large unit cell
  - CTMRG: power method
  - VUMPS: non-Hermitian transfer matrix
- Optimization: iPEPS
  - Imaginary time evolution(ITE): artificial expertise
  - Gradient optimization: graph summations or automatic differentiation (AD) careful design

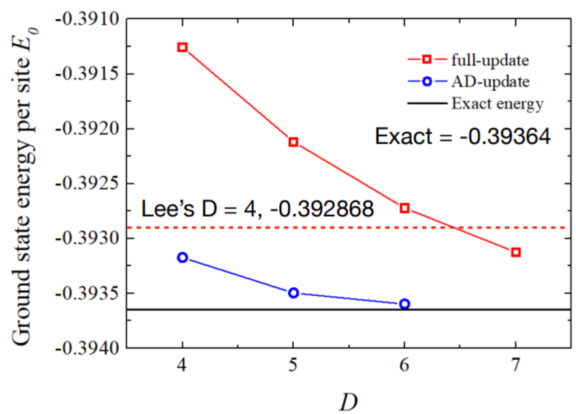
Problem of CTMRG

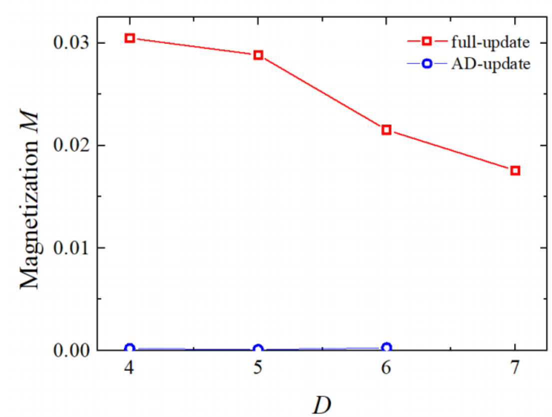
- Power method: slow
- No general truncation principle



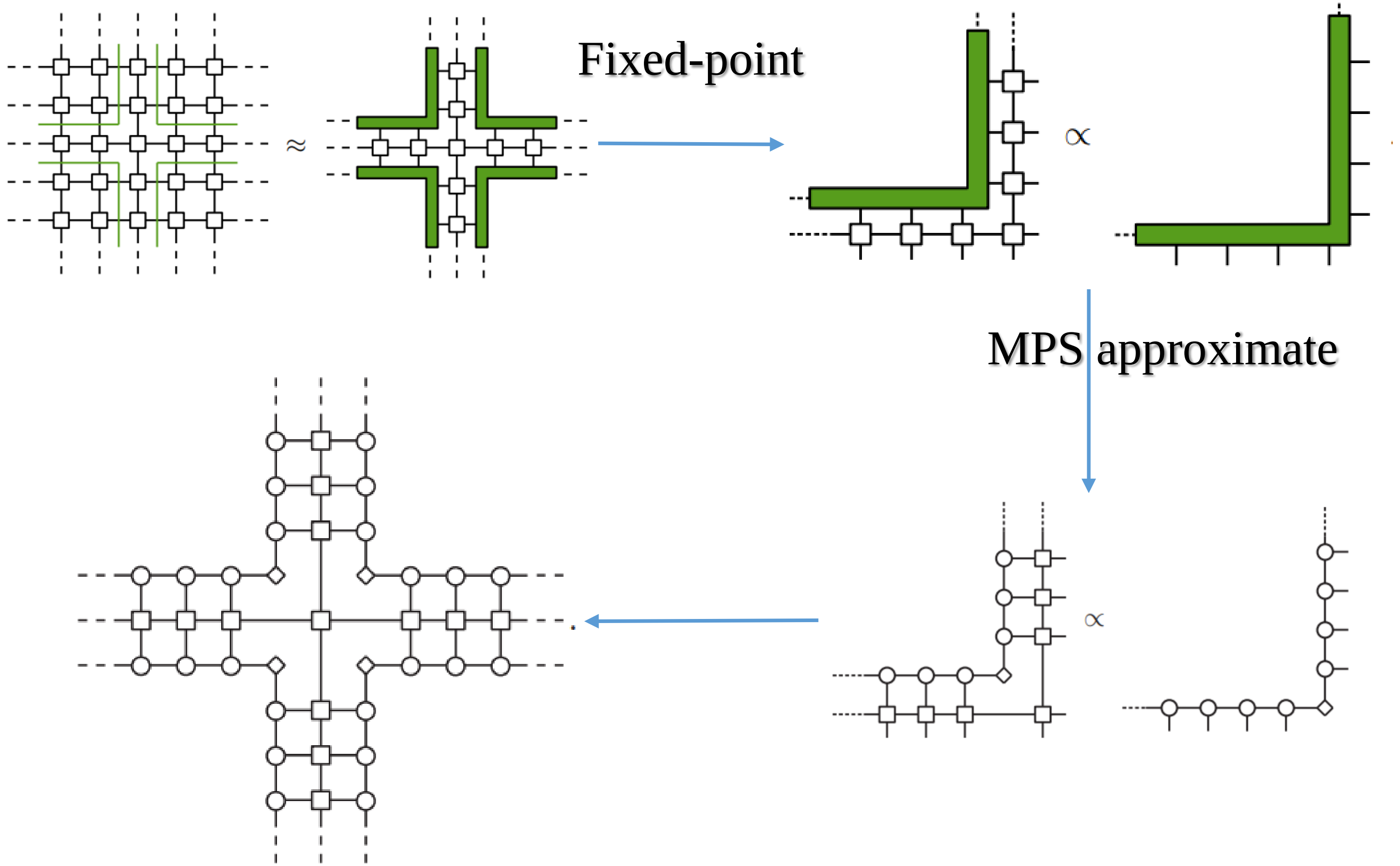
https://journals.aps.org/prl/supplemental/10.1103/PhysRevLett.113.046402/suppl.pdf

Iregui, *et. al.*, PRB 90, 195102 (2014)

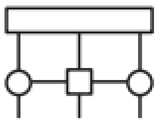
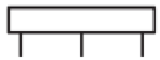




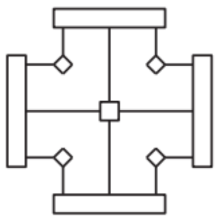
Channel environments



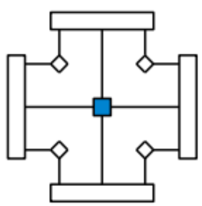
Gradient

 $= \lambda \times$ 

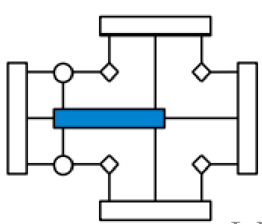
overlap

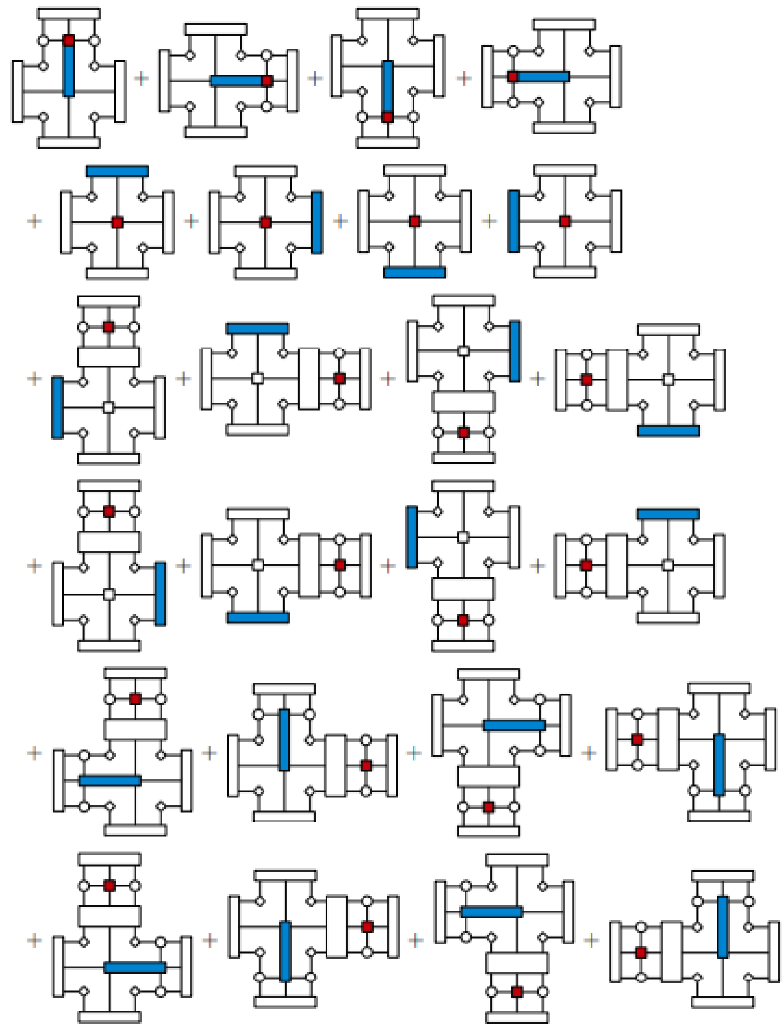
$\langle \Psi(A) | \Psi(A) \rangle =$ 

observable

$\langle \Psi(A) | O | \Psi(A) \rangle =$ 

two-site

$\langle \Psi(A) | O | \Psi(A) \rangle =$ 

grad = 

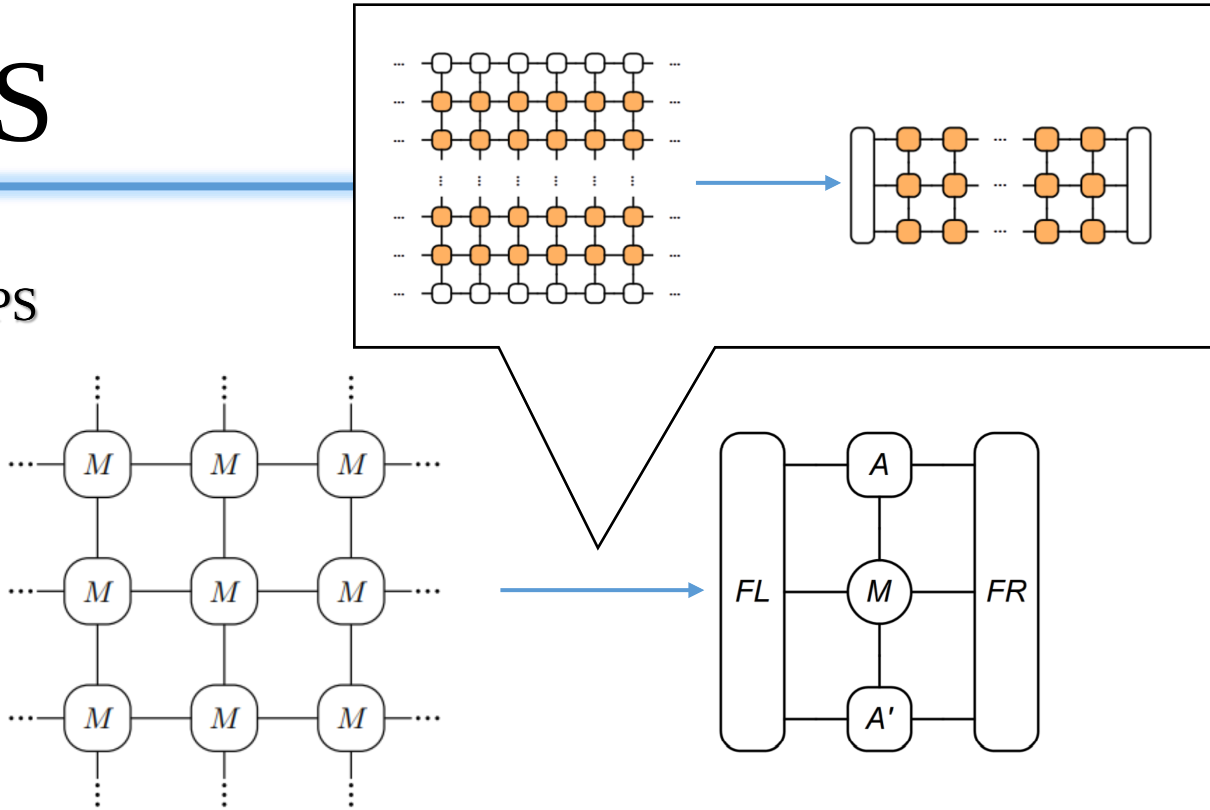
L Vanderstraeten, PRB 94, 155123 (2016)

AD-VUMPS

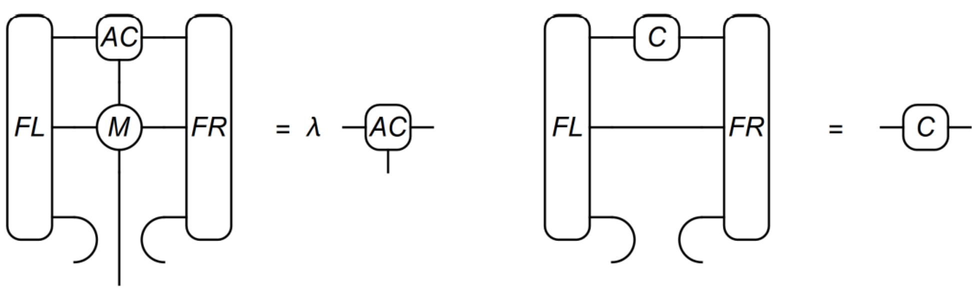
variational uniform matrix product states

VUMPS

Boundary MPS

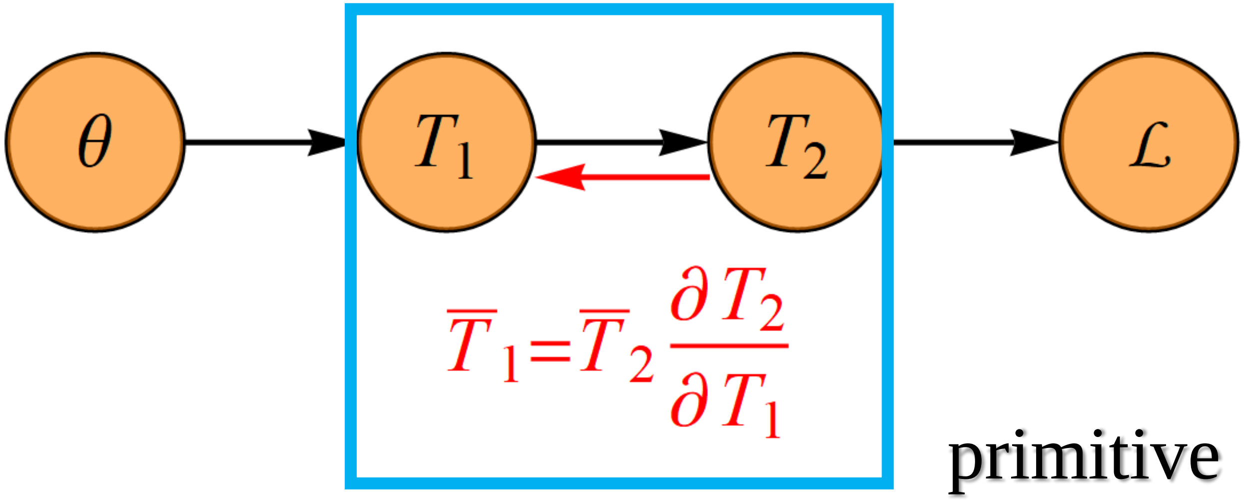


Fixed point



L Vanderstraeten, SciPost Phys. Lect. Notes 7 (2019)

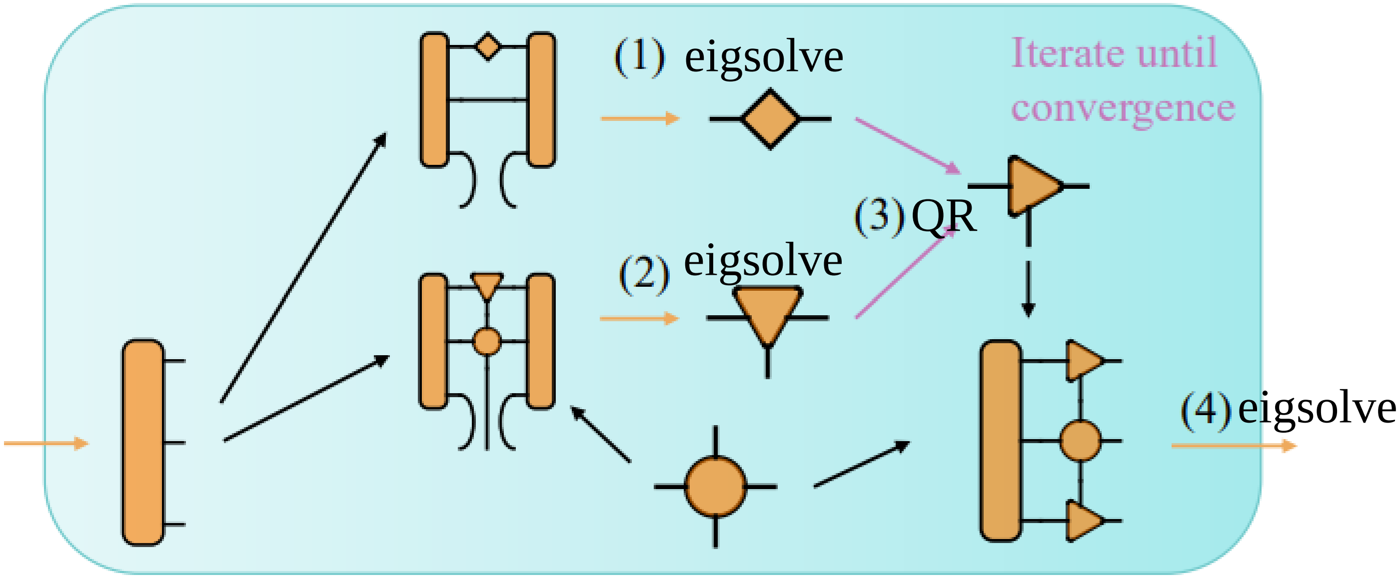
Computation graph and adjoint



$$\frac{\partial \mathcal{L}}{\partial \theta} = \frac{\partial \mathcal{L}}{\partial T_2} \frac{\partial T_2}{\partial T_1} \frac{\partial T_1}{\partial \theta}$$

adjoint $\bar{T} = \partial \mathcal{L} / \partial T$

Computation graph of VUMPS



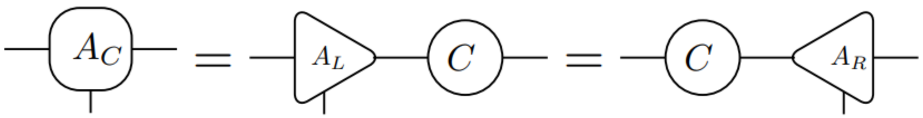
Key factors of AD:

1. Numerical stability
2. Matrix free linear algebra: Krylov, GMRES

Adjoint of QR decomposition

- $\bar{M} = [\bar{Q} + Q\,\mathrm{Hermitian}(R\bar{R}^\dagger - \bar{Q}^\dagger Q)](R^\dagger)^{-1}$

- $C = Q_C \cdot R_C$

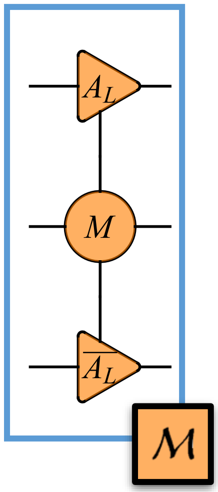


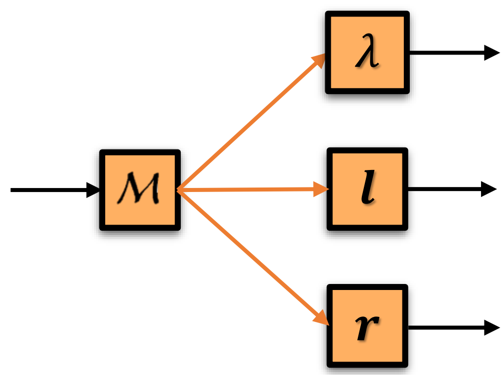
- Trick: add $\delta = 10^{-12}$ to R's diagonal elements for the stability of inverse

Adjoint of dominant eigen solver

- dominant eigen equation

$$\boldsymbol{l}^T \mathcal{M} = \lambda \boldsymbol{l}^T, \quad \mathcal{M}\boldsymbol{r} = \lambda \boldsymbol{r}, \quad \boldsymbol{l}^T \boldsymbol{r} = 1$$





- Adjoint related to linear solver

$$\bar{\mathcal{M}} = \bar{\lambda}\boldsymbol{l}\boldsymbol{r}^T - \boldsymbol{l}\xi_{\boldsymbol{l}}^T - \xi_{\boldsymbol{r}}\boldsymbol{r}^T$$

Xie, PRB 101, 245139 (2020)

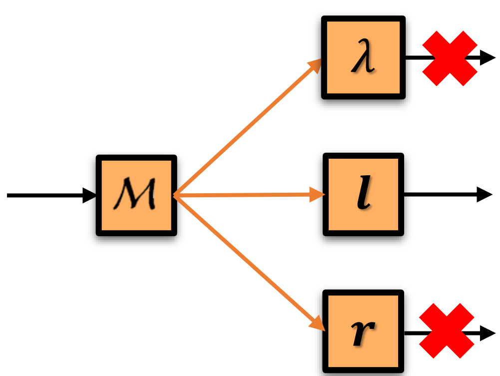
Gauge invariant case

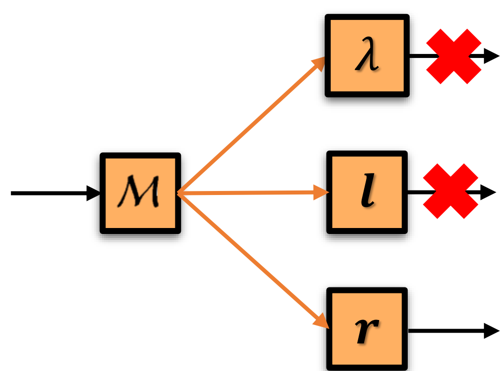
- For physical observable

$$\eta_\lambda = \boldsymbol{l}^T \bar{\boldsymbol{l}} = \boldsymbol{r}^T \bar{\boldsymbol{r}}$$

Only l or r, $\eta_\lambda = 0$

$$\bar{\mathcal{M}} = -\boldsymbol{l}\xi_{\boldsymbol{l}}^T \quad \text{or} \quad \bar{\mathcal{M}} = -\xi_{\boldsymbol{r}}\boldsymbol{r}^T$$





Simplify linear solver

$$(\mathcal{M}-\lambda I)\xi_l = \left(1-\boldsymbol{r}\boldsymbol{l}^T\right)\bar{\boldsymbol{l}}, \quad \boldsymbol{l}^T\boldsymbol{\xi}_l = 0$$

$$\left(\mathcal{M}^T-\lambda I\right)\xi_r = \left(1-\boldsymbol{l}\boldsymbol{r}^T\right)\bar{\boldsymbol{r}}, \quad \boldsymbol{r}^T\xi_r = 0$$

Only $\boldsymbol{l}$ $\quad$ $\boldsymbol{l}^T\bar{\boldsymbol{l}} = \boldsymbol{r}^T\bar{\boldsymbol{r}} = 0$

$$(\mathcal{M}-\lambda I)\boldsymbol{\xi}_l = \bar{l}, \quad \boldsymbol{l}^T\boldsymbol{\xi}_l = 0$$

$$\left(\mathcal{M}^T-\lambda I\right)\boldsymbol{\xi}_r = 0 \quad \boldsymbol{r}^T\boldsymbol{\xi}_r = 0$$

Numerical stability

- Linear solver

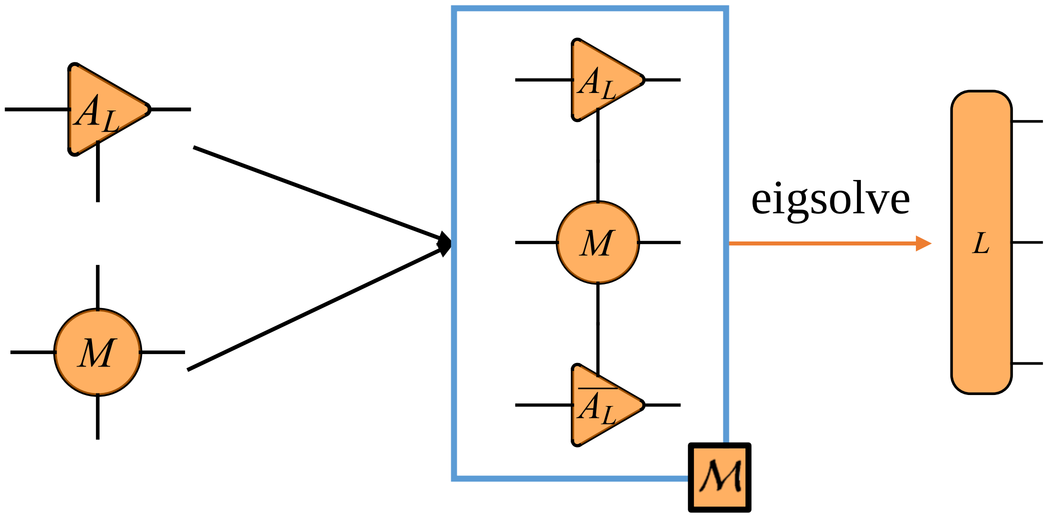
$$(\mathcal{M} - \lambda I)\xi_l = \bar{\boldsymbol{l}}, \quad \boldsymbol{l}^T \xi_l = 0$$

$$\boldsymbol{l}^T \bar{\boldsymbol{l}} = \boldsymbol{r}^T \bar{\boldsymbol{r}} = 0$$

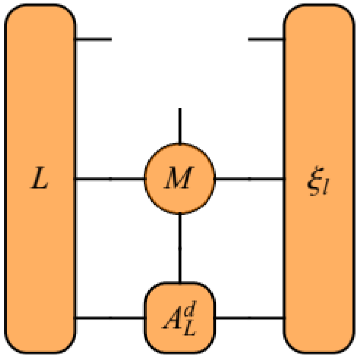
- subtract part of $\bar{\boldsymbol{l}}$ which is not orthogonal to $\boldsymbol{l}$

$$\bar{\boldsymbol{l}} \to \bar{\boldsymbol{l}} - \boldsymbol{l}^T \bar{\boldsymbol{l}} \cdot \boldsymbol{l}$$

Matrix free linear algebra

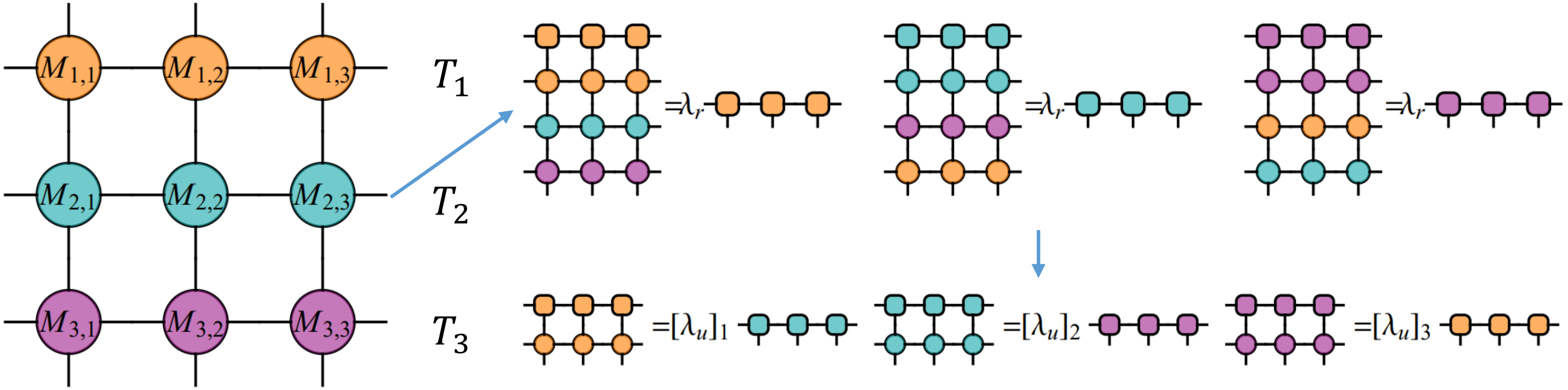


$$\overline{A_L} = \overline{\mathcal{M}} \cdot \frac{\partial \mathcal{M}}{\partial A_L} = -$$

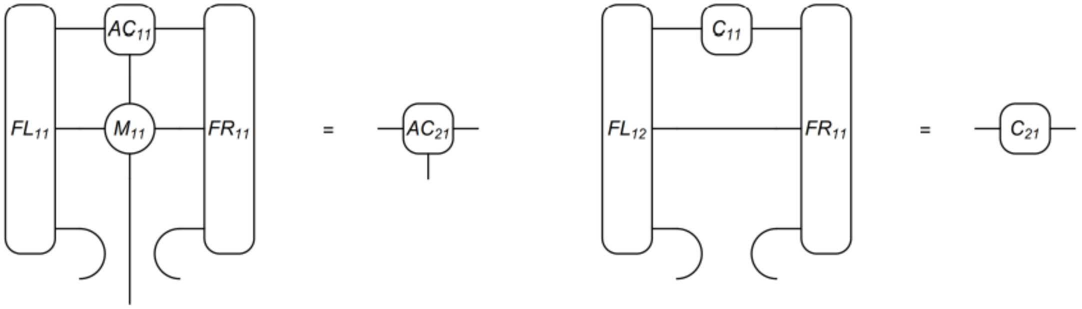


Toward to large unit cell

Large unit cell



Each bulk has its environment

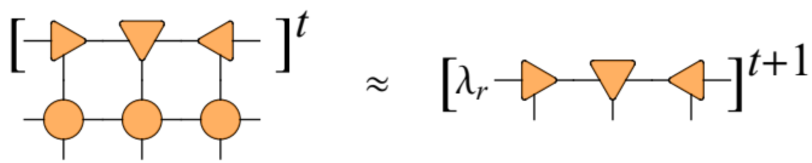


Large cell→Power method for non-hermitian

Nietner, Quantum 4, 328 (2020).

Non-hermitian VUMPS

- One step power

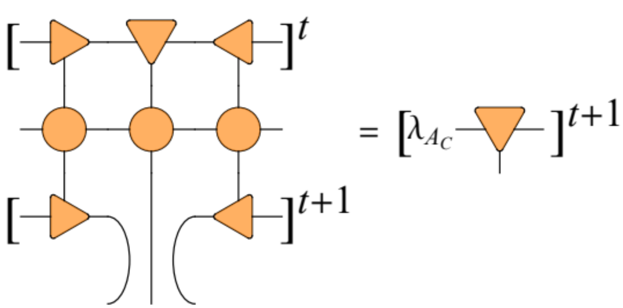


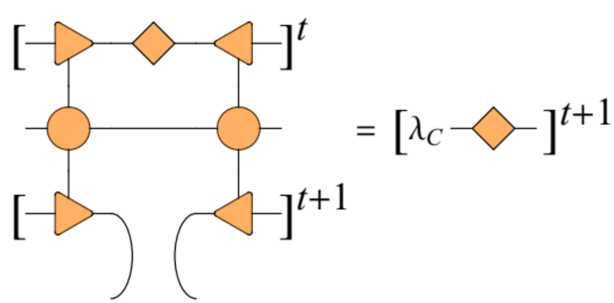
- Maximum fidelity

$$\arg\max_{A^{t+1}} \left(\log \frac{\langle \psi(\bar{A}^{t+1}) | T | \psi(A^{t}) \rangle \langle \psi(\bar{A}^{t}) | T^{\dagger} | \psi(A^{t+1}) \rangle}{\langle \psi(\bar{A}^{t+1}) | \psi(A^{t+1}) \rangle} \right)$$

Bram Vanhecke, arXiv:2001.11882(2021)

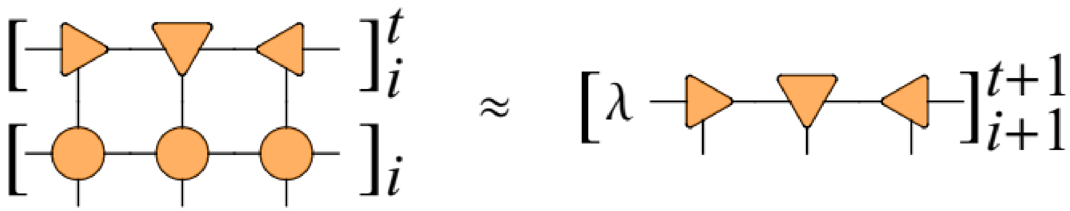
- Fixed-point function



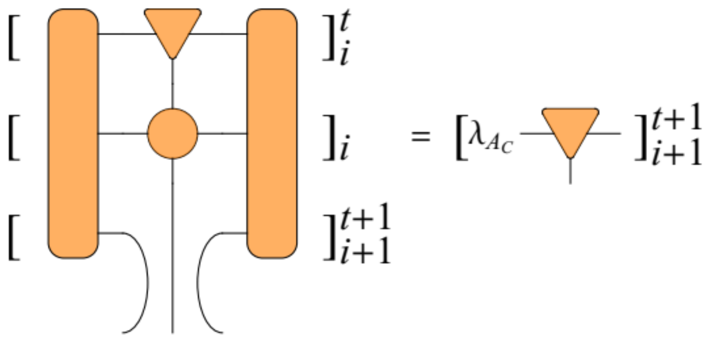


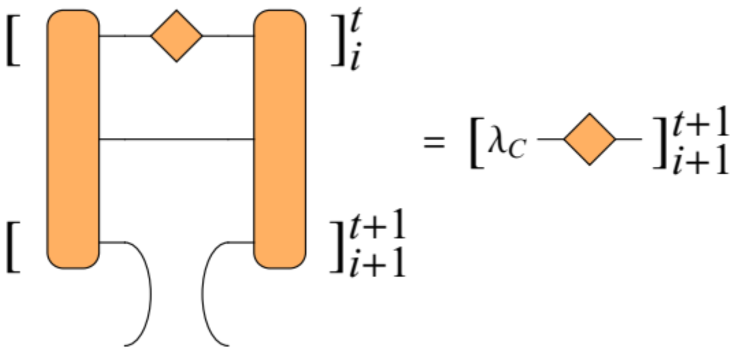
Large unit cell VUMPS

- One step power

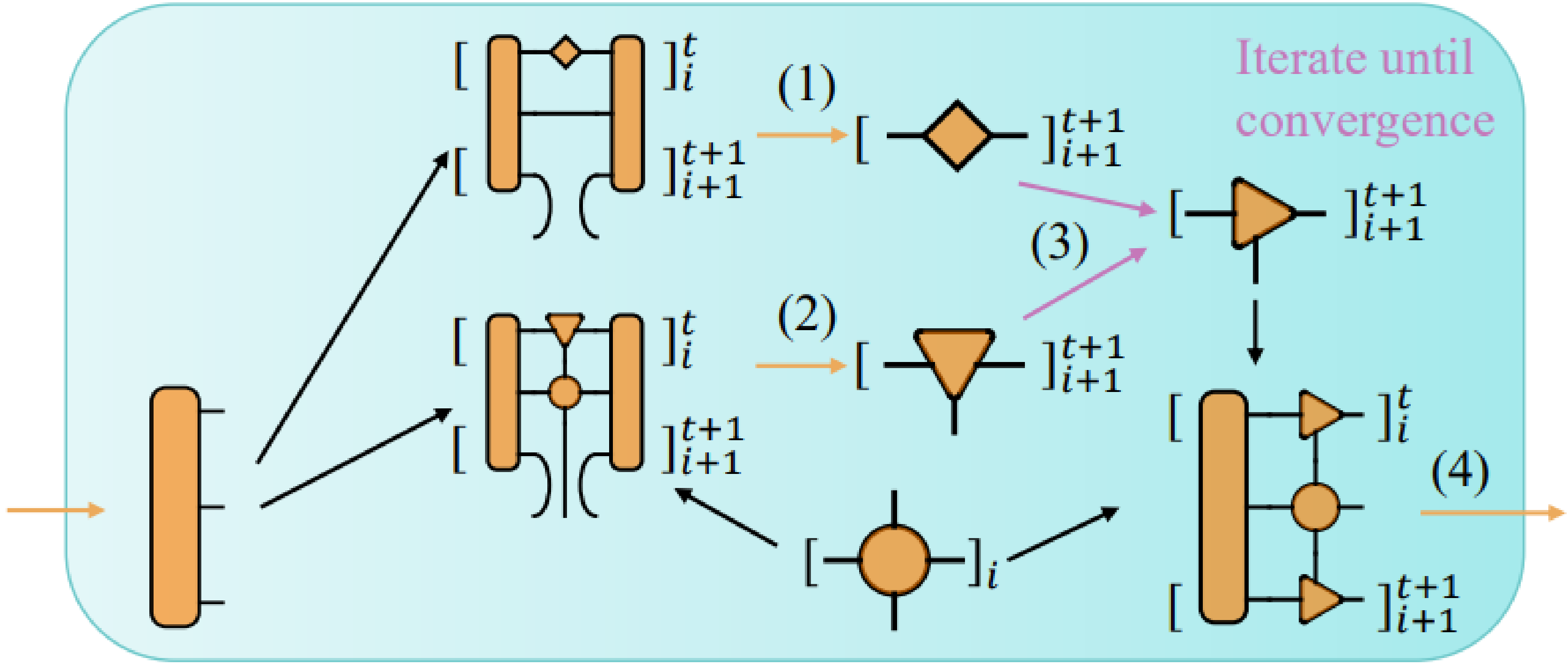


- Fixed-point function





Computation graph



Applications

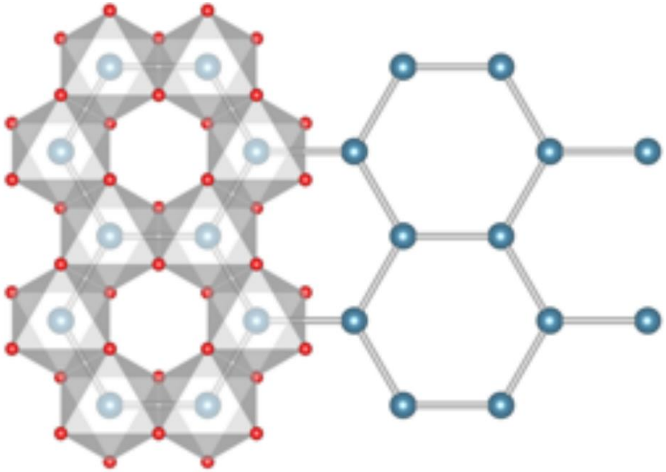
Kitaev models

Model

- Hamiltonian

$$\hat{H} = \sum_{\langle i,j\rangle_\gamma} K S_i^\gamma S_j^\gamma + J \mathbf{S}_i \cdot \mathbf{S}_j + \Gamma \left(S_i^\alpha S_j^\beta + S_i^\beta S_j^\alpha \right)$$
$$+ \Gamma' \left(S_i^\alpha S_j^\gamma + S_i^\gamma S_j^\alpha + S_i^\beta S_j^\gamma + S_i^\gamma S_j^\beta \right)$$

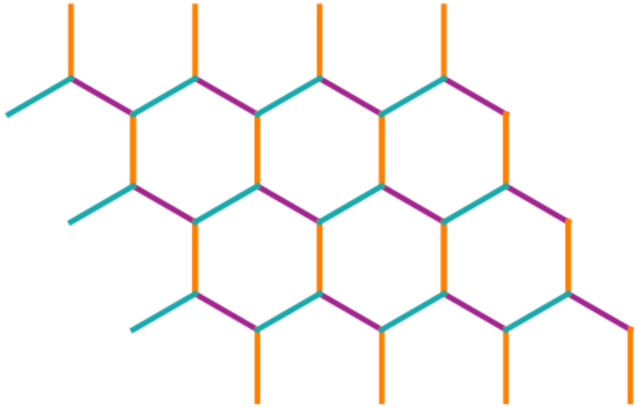
  - J Heisenberg term
  - Γ spin-orbit couplings
  - Γ' lattice distortions

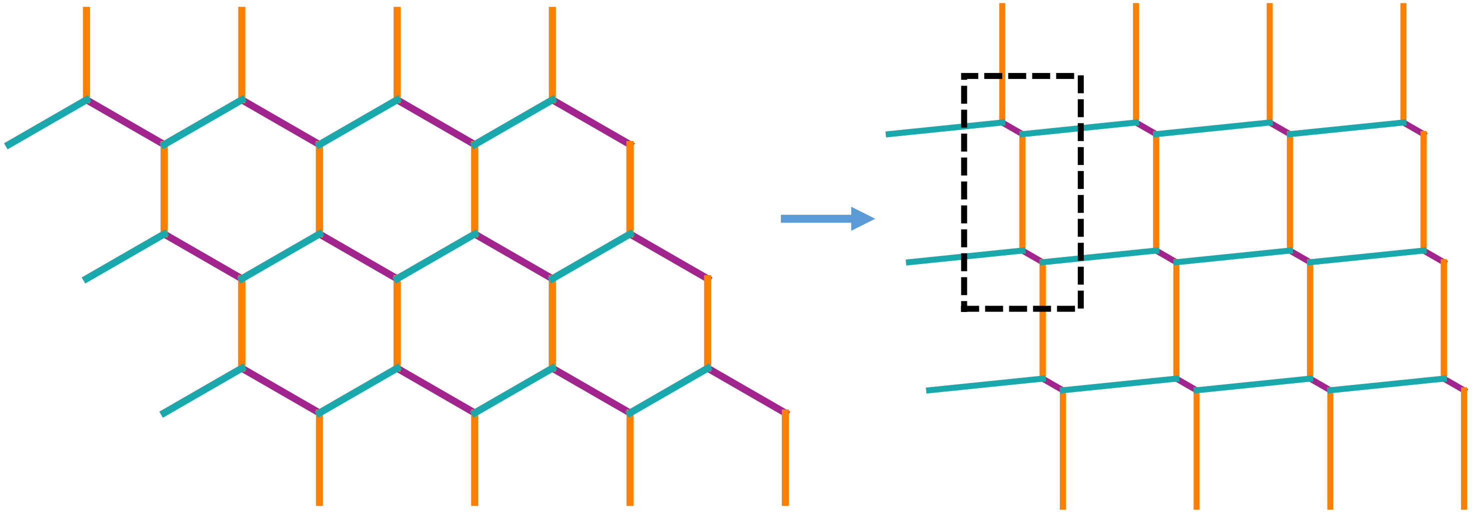


S.Trebst arXiv:1701.07056(2017)

- Rich phase diagram
  - ferromagnetic (FM), antiferromagnetic (Neel), zigzag (ZZ), stripy phases
  - 6-site, 18-site unit

honeycomb lattice to square lattice

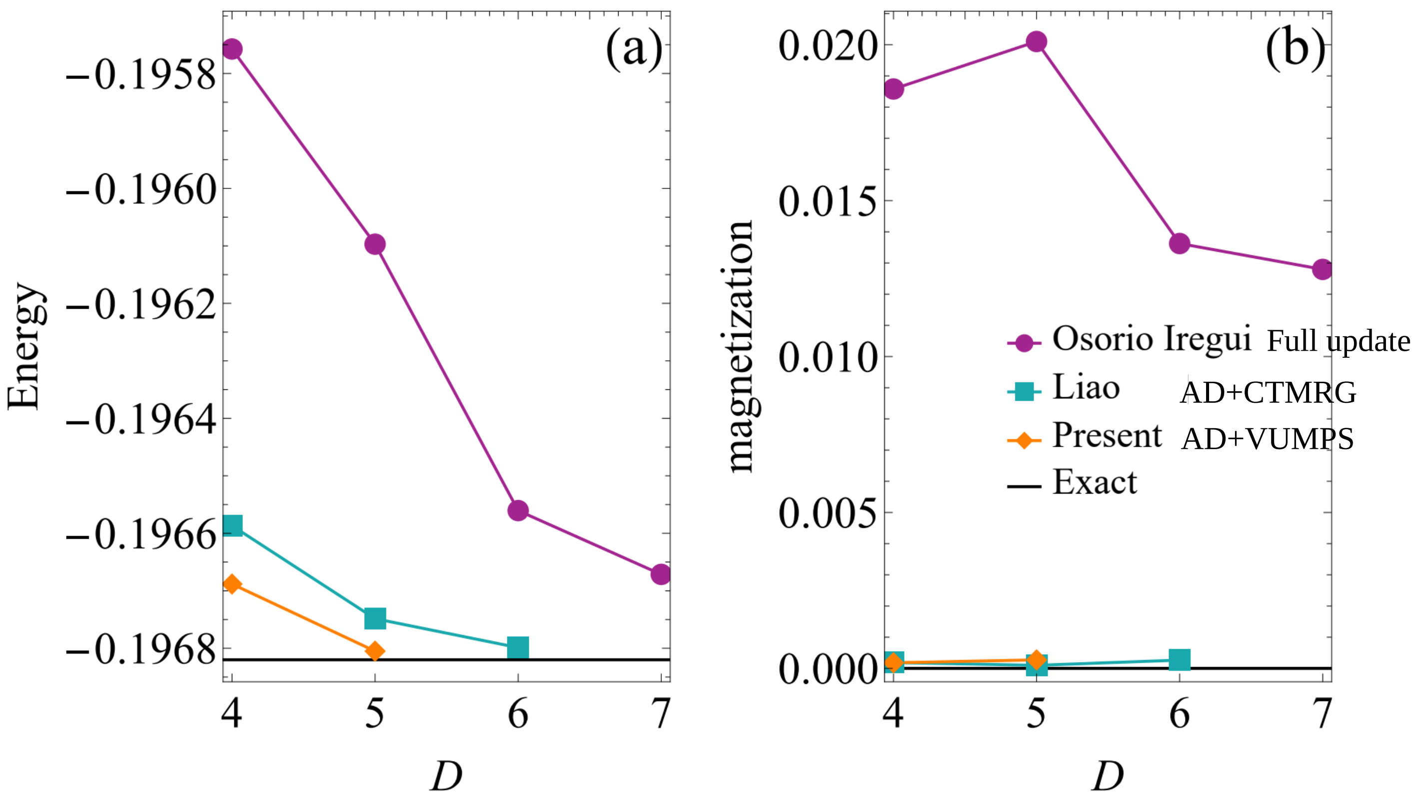




Honeycomb Kitaev

Lower energy

zero mag

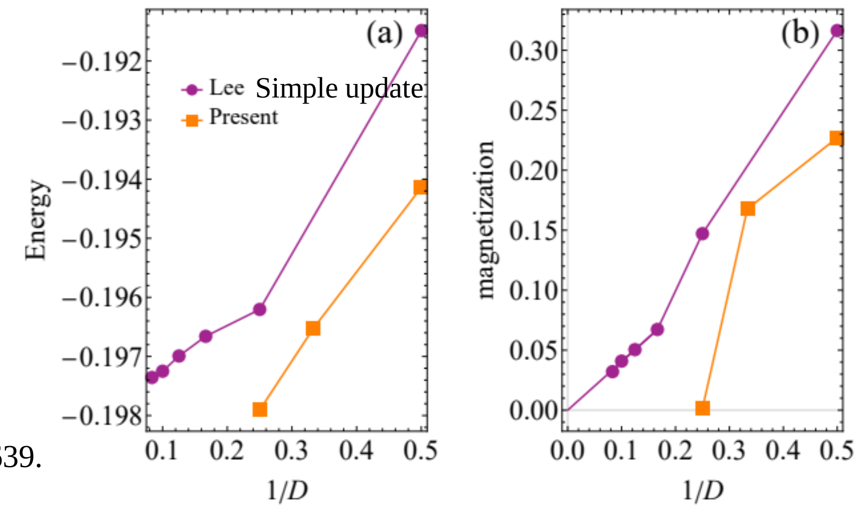


Better optimization rather than larger D!

Liao, Differentiable Programming Tensor Network, talk,2019
Osorio Iregui, Phys. Rev. B 90, 239903 (2014)

K-Γ model

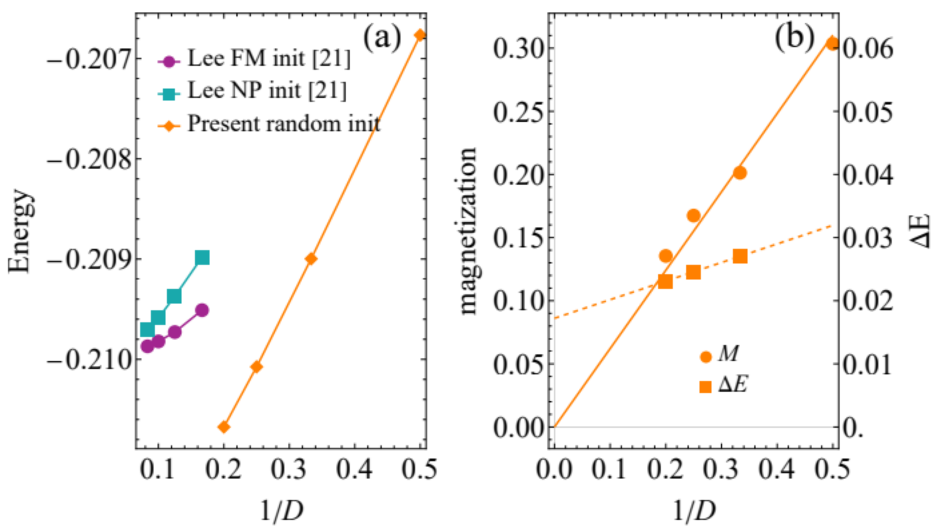
- $\Gamma = 0.03$



Lee, Hyun-Yong, et al. Nature communications 11.1 (2020): 1639.

Kitaev spin liquid

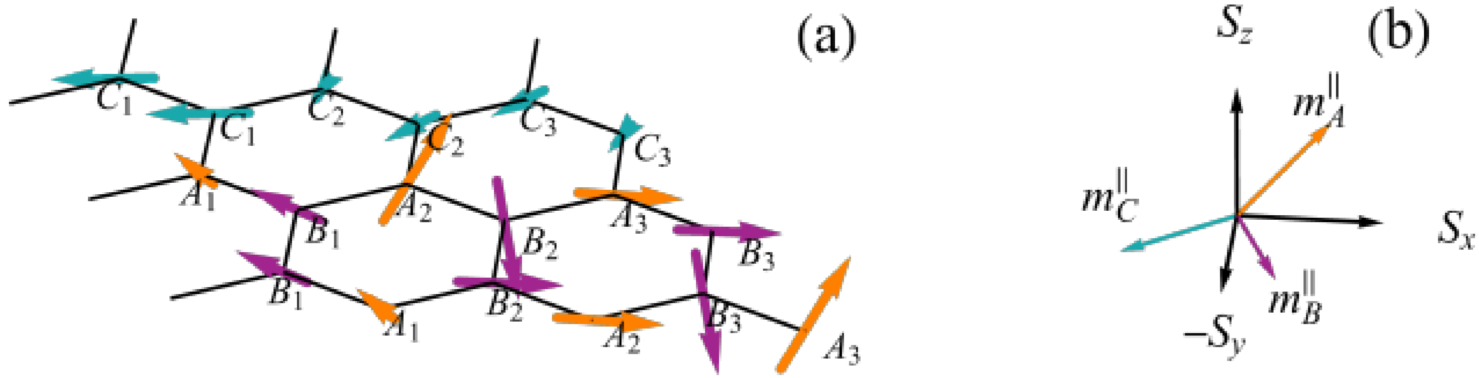
- $\Gamma = 0.095$



Nematic

K-Γ model for $\Gamma = 1$

- 18-site configuration
- $E = -0.3520$ at D = 4 AD-VUMPS, a random initial
- $E = -0.3518$ at D = 8 ITE-CTMRG, hundreds of initials



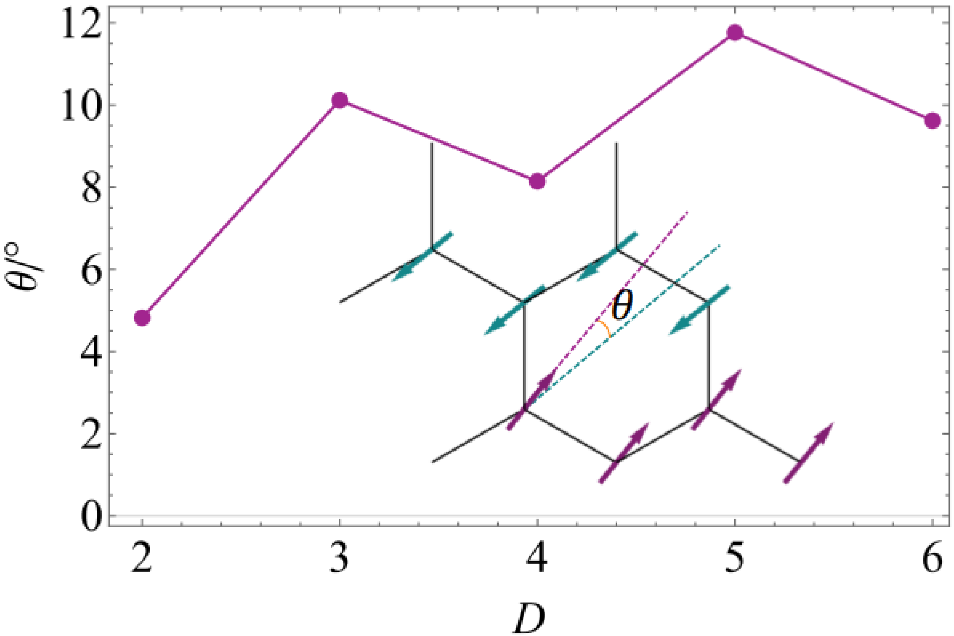
Jheng-Wei Li, arXiv.2206.08946

K-J-Γ-Γ' model

- parameters $K = -25$meV, $\Gamma = 0.3|K|$, $\Gamma' = -0.02|K|$ and $J = -0.1|K|$

H. Li, Nature Communications 12, 1 (2021).

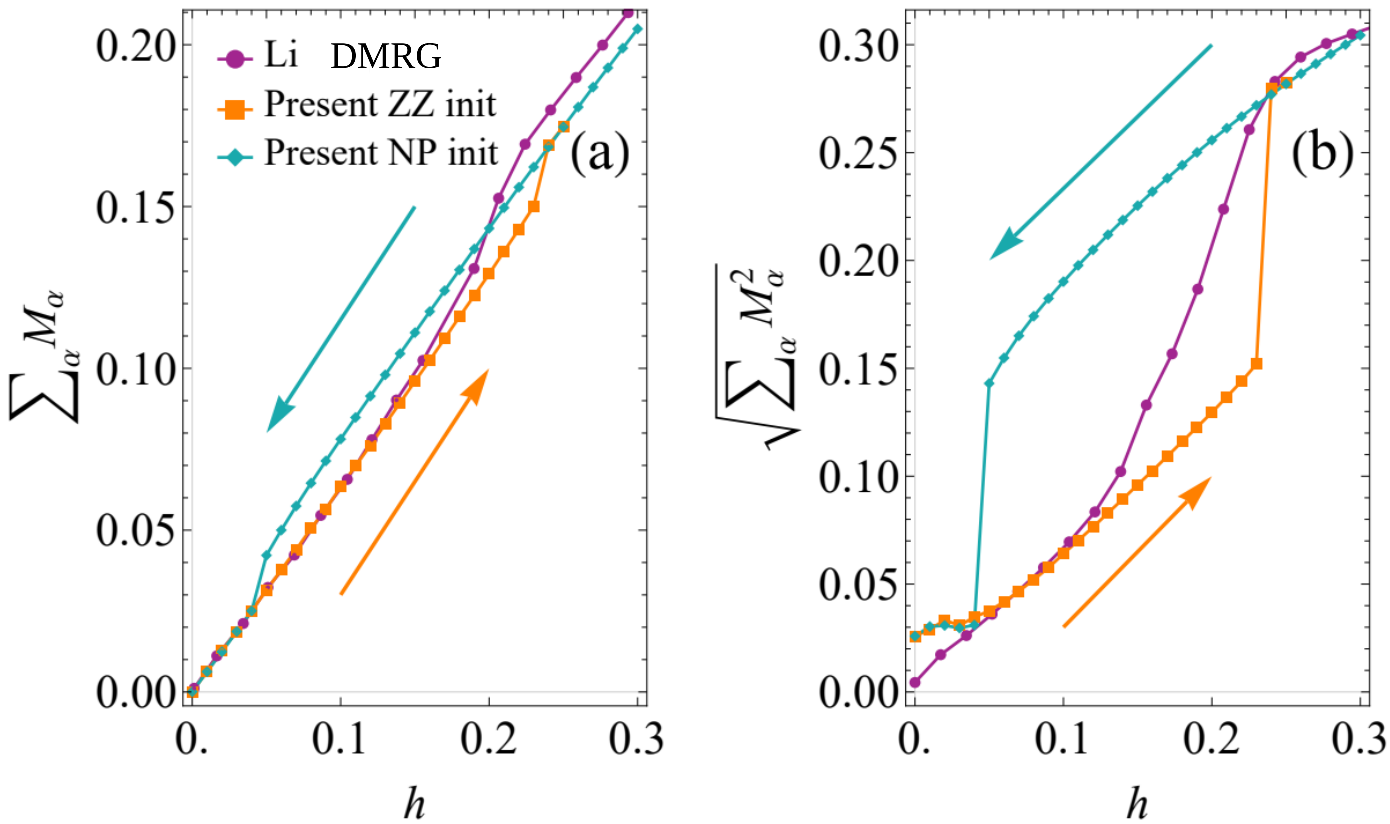
- non-collinear Zigzag



Out of plane critical point

- The obvious jump

H. Li, Nature Communications 12, 1 (2021).



U1-symmetry tensor for fermion

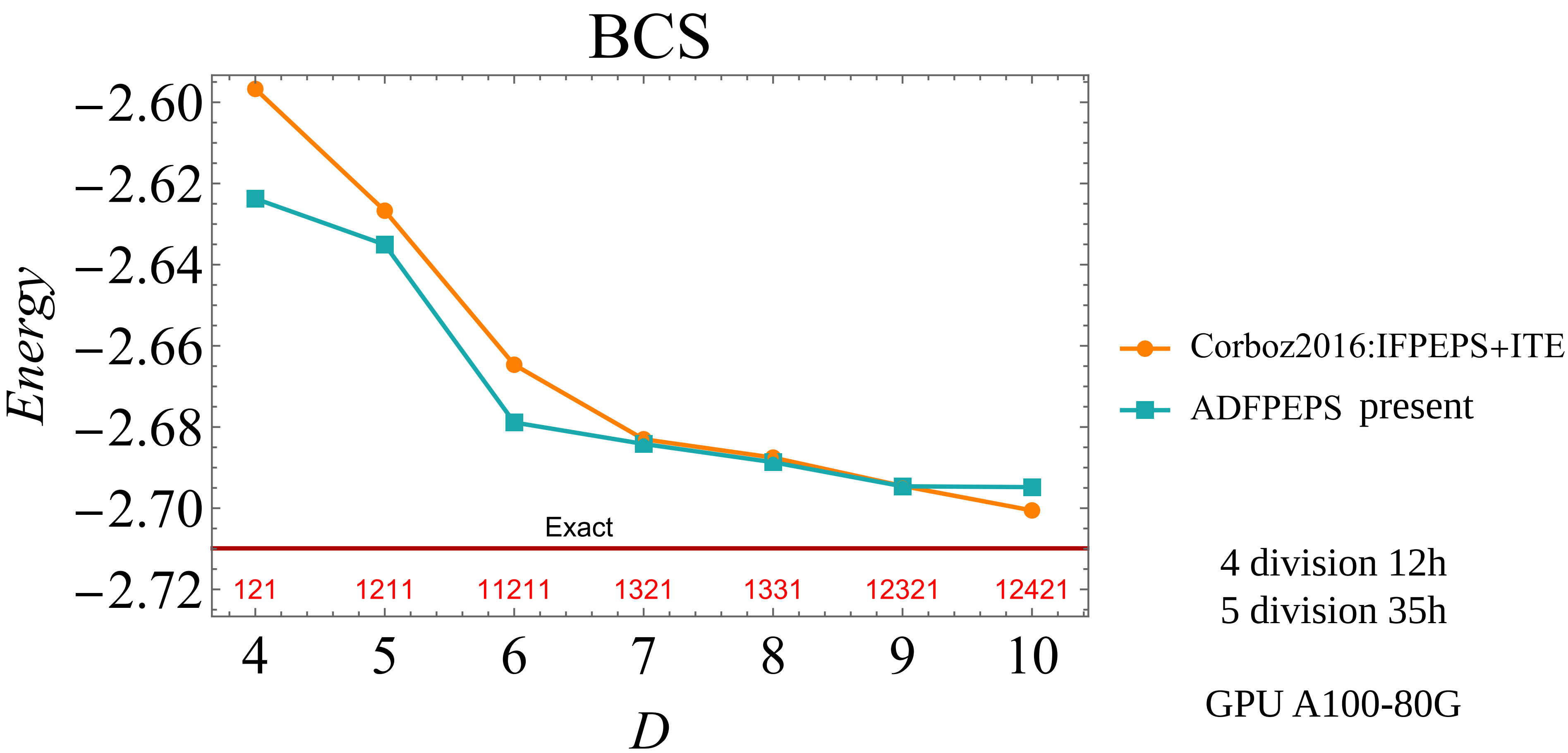
- $Z2 \rightarrow U1$

$$
\begin{array}{cc}
 & \begin{array}{cc} \quad 0 \quad & \quad 1 \quad \end{array} \\
\begin{array}{c} 0 \\ 1 \end{array} &
\begin{pmatrix}
\cdot & \cdot & \cdot & & & \\
\cdot & \cdot & \cdot & & & \\
\cdot & \cdot & \cdot & & & \\
 & & & \cdot & \cdot & \cdot \\
 & & & \cdot & \cdot & \cdot \\
 & & & \cdot & \cdot & \cdot
\end{pmatrix}
\end{array}
\qquad
\begin{array}{cc}
 & \begin{array}{ccc} \quad 0 \quad & \quad 1 \quad & \quad 2 \quad \end{array} \\
\begin{array}{c} 0 \\ 1 \\ 2 \end{array} &
\begin{pmatrix}
\cdot & \cdot & \cdot & & & & & & \\
\cdot & \cdot & \cdot & & & & & & \\
\cdot & \cdot & \cdot & & & & & & \\
 & & & \cdot & \cdot & \cdot & & & \\
 & & & \cdot & \cdot & \cdot & & & \\
 & & & \cdot & \cdot & \cdot & & & \\
 & & & & & & \cdot & \cdot & \cdot \\
 & & & & & & \cdot & \cdot & \cdot \\
 & & & & & & \cdot & \cdot & \cdot
\end{pmatrix}
\end{array}
$$

- Tensor contraction with OMEinsum.jl
  - Rebuilt permutedims and reshape
  - Efficiency: same with C in CPU and 30~40 faster in GPU
- Q: Physical bond division is decided by Hamiltonian, but virtual bond division is arbitrary.

Bogoliubov-de Gennes (BdG)

$$\hat{H} = -t\sum_{\langle i,j,\sigma\rangle}(\hat{c}^{\dagger}_{i\sigma}\hat{c}_{j\sigma} + \mathrm{H.c.}) + \sum_{\langle \mathrm{i,j}\rangle}\gamma_{\mathrm{ij}}(\hat{\mathrm{c}}^{\dagger}_{\mathrm{i}\uparrow}\hat{\mathrm{c}}^{\dagger}_{\mathrm{j}\downarrow} - \hat{\mathrm{c}}^{\dagger}_{\mathrm{i}\downarrow}\hat{\mathrm{c}}^{\dagger}_{\mathrm{j}\uparrow} + \mathrm{H.c.})$$



4 division 12h
5 division 35h

GPU A100-80G

Corboz P. PRB, 2016, 93(4): 045116.

Summary

- feasible, stable and efficient backward of VUMPS
- highly accurate results in Kitaev type frustrated systems
- add U1-symmetry to explore fermion system
    - the U1 block division

Thank you for your attention!