多体問題の計算科学

Computational Science for Many-Body Problems

#10 Linear algebra of large and sparse matrices forquantum many-body problems

15:10-16:40 June 22, 2021

Quantum many-body problems and huge sparse matrices

- O. Application of QMC
- 1. Quantum mechanics by linear algebra
- 2. Linear algebra by comupter
- 3. Quantum many-body problems by linear algebra
- 4. Eigenvalue problems of large & sparse matrices

Application of QMC

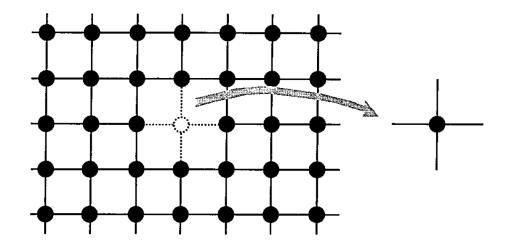
- -Dynamical mean-field theory
- -Quantum chemistry

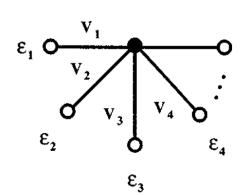
Dynamical Mean Field and Impurity Model

As a review, A. Georges, G. Kotliar, W. Krauth, and Marcelo J. Rozenberg, Rev. Mod. Phys. 68, 13 (1996).

Mapping from lattice model to impurity model

$$\hat{H} = \sum_{\langle i,j \rangle,\sigma} (\hat{c}_{i\sigma}^{\dagger} \hat{c}_{j\sigma} + \hat{c}_{j\sigma}^{\dagger} \hat{c}_{i\sigma}) + U \sum_{i} \hat{c}_{i\uparrow}^{\dagger} \hat{c}_{i\uparrow} \hat{c}_{i\uparrow} \hat{c}_{i\downarrow}$$





$$\hat{H} = \sum_{i=1}^{\infty} \epsilon_j \hat{c}_{j\sigma}^{\dagger} \hat{c}_{j\sigma} + \sum_{i=1}^{\infty} v_j (\hat{c}_{j\sigma}^{\dagger} \hat{d}_{\sigma} + \hat{d}_{\sigma}^{\dagger} \hat{c}_{j\sigma}) + U \hat{d}_{\uparrow}^{\dagger} \hat{d}_{\uparrow} \hat{d}_{\downarrow}^{\dagger} \hat{d}_{\downarrow}$$

T. Ma Rev.

G. Ko Rev.

Dynamical Mean Field and Impurity Model

As a review, A. Georges, G. Kotliar, W. Krauth, and Marcelo J. Rozenberg, Rev. Mod. Phys. 68, 13 (1996).

Self-consistent impurity model

$$\hat{H} = \sum_{j,\sigma} \epsilon_j \hat{c}_{j\sigma}^{\dagger} \hat{c}_{j\sigma} + \sum_{j,\sigma} v_j (\hat{c}_{j\sigma}^{\dagger} \hat{d}_{\sigma} + \hat{d}_{\sigma}^{\dagger} \hat{c}_{j\sigma}) + U \hat{d}_{\uparrow}^{\dagger} \hat{d}_{\uparrow} \hat{d}_{\downarrow}^{\dagger} \hat{d}_{\downarrow}$$

-Green's function of impurity model

$$G^{\text{imp}}(\omega) = \frac{1}{\omega + i\delta + \mu - \Sigma^{\text{imp}}(\omega) - \Delta(\omega)}$$

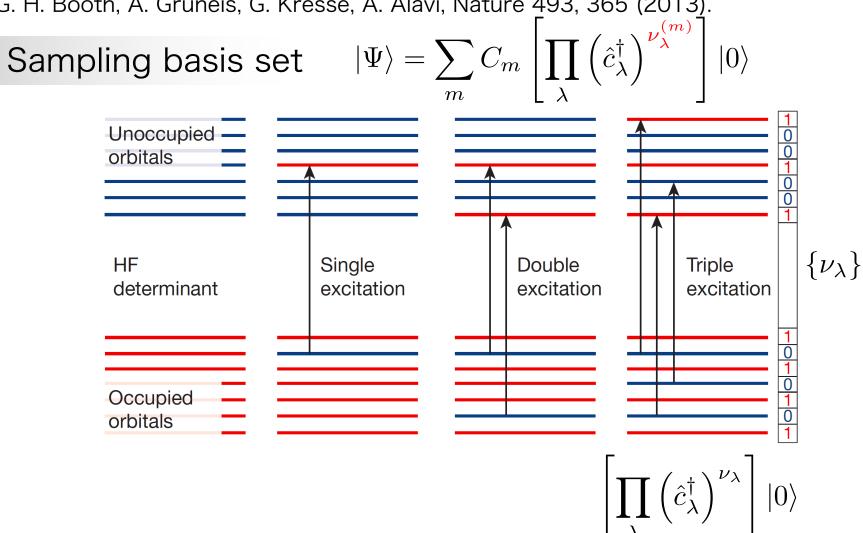
-Dynamical mean-field
$$\Delta(\omega) = \omega + i\delta + \mu - \Sigma^{\mathrm{imp}}(\omega) - \left[\frac{1}{N}\sum_{\vec{k}}\frac{1}{\omega + i\delta + \mu - \epsilon(\vec{k}) - \Sigma^{\mathrm{imp}}(\omega)}\right]^{-1}$$

$$= \sum_{i}\frac{|v_{j}|^{2}}{\omega + i\delta - \epsilon_{j}}$$

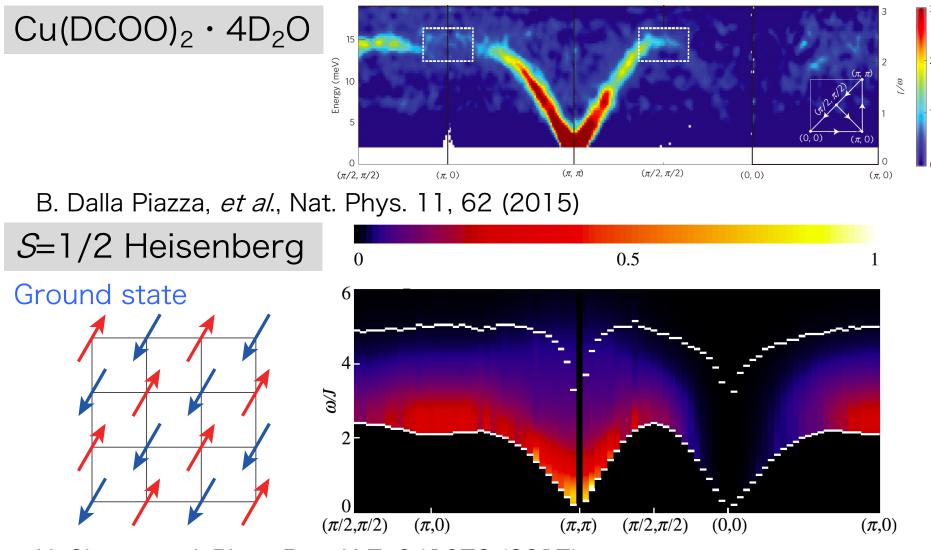
You can find applications of DMFT in G. Kotliar, et al., Rev. Mod. Phys. 78, 865 (2006).

Configuration Interaction Monte Carlo

G. H. Booth, A. Grüneis, G. Kresse, A. Alavi, Nature 493, 365 (2013).



Example of Applications: Spin Excitations in Square Lattice Heisenberg model



H. Shao, et al., Phys, Rev, X 7, 041072 (2017)



Naïvely, linear partial differential equations are rewritten by Linear equations

Schrödinger equation represented by partial diff. eq.

$$i\hbar \frac{d}{dt}\psi(\vec{r},t) = \left[-\frac{\hbar^2}{2m}\vec{\nabla}^2 + V(\vec{r})\right]\psi(\vec{r},t)$$

Stationary solution: $\psi(\vec{r},t) = \phi(\vec{r})e^{-iEt/\hbar}$

$$\left[-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{r}) \right] \phi(\vec{r}) = E\phi(\vec{r})$$

Schrödinger equation represented by linear eqs.

$$\left[-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{r}) \right] \phi(\vec{r}) = E\phi(\vec{r})$$

Expanded by orthonormal basis

$$\phi(\vec{r}) = \sum_{m} c_m u_m(\vec{r})$$

$$\int d^3r \ u_\ell^*(\vec{r})u_m(\vec{r}) = \delta_{\ell,m}$$
$$\int d^3r \phi^*(\vec{r})\phi(\vec{r}) = \sum_m |c_m|^2$$

Matrix representation

$$H_{\ell m} = \int d^3 r u_{\ell}^*(\vec{r}) \left[-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{r}) \right] u_m(\vec{r})$$

$$\sum_{m} H_{\ell m} c_m = E c_{\ell}$$

$$\sum_{m} H_{\ell m} c_m = E c_{\ell}$$

Hermitian matrix $H_{\ell m} = H_{m\ell}^*$

- -Diagonalizable by unitary matrices
- -Real eigenvalues

$$\sum_{m} H_{\ell m} U_{m\alpha} = U_{\ell \alpha} E_{\alpha}$$

$$\sum_{m} (U^{\dagger})_{\beta m} U_{m\alpha} = \sum_{m} (U_{m\beta})^* U_{m\alpha} = \delta_{\beta,\alpha}$$

$$\sum_{m} H_{\ell m} c_m = E c_{\ell}$$

Vector representation of expectation value

$$\frac{\int d^3 r \phi^*(\vec{r}) \hat{O} \phi(\vec{r})}{\int d^3 r \phi^*(\vec{r}) \phi(\vec{r})} = \frac{\sum_{\ell,m} c_{\ell}^* c_m \int d^3 r u_{\ell}^*(\vec{r}) \hat{O} u_m(\vec{r})}{\sum_{l} |c_n|^2} \\
= \frac{\sum_{\ell,m} c_{\ell}^* O_{\ell m} c_m}{\sum_{l} |c_n|^2}$$

Linear Algebra by Computer

Linear Algebra by Computer

$$\sum_{m} H_{\ell m} U_{m\alpha} = U_{\ell \alpha} E_{\alpha}$$

Hermitian matrix $H_{\ell m}=H_{m\ell}^*$

LAPACK (Linear Algebra PACKage)

http://www.netlib.org/lapack/explore-html/index.html

```
subroutine zheev (character
                                                                   JOBZ,
                                                                   UPLO,
                                      character
zheev
                                                                   N.
                                      integer
z: double complex
                                      complex*16, dimension( lda, *)
he: hermitian
                                                                   LDA,
                                      integer
ev: eigenvalue & eigenvector
                                      double precision, dimension(*)
                                                                   WORK,
                                      complex*16, dimension( * )
                                                                   LWORK,
                                      integer
                                                                   RWORK,
                                      double precision, dimension(*)
                                                                   INFO
                                      integer
```

Linear Algebra by Computer

LAPACK (Linear Algebra PACKage)

http://www.netlib.org/lapack/explore-html/index.html

- -Language: Fortran
 C & C++ can call LAPACK
- -License: Modified BSD license
- -Parallelized version: ScaLAPACK
- cf.) intel MKL (commercial library)

- -Transformation
- -Eigenvalue
- -Singular value

Hamiltonian in 2nd quantization form

Many-body electrons confined in one-body potential

(No spin-orbit coupling)

$$\hat{H} = \sum_{\sigma} \int d^3r \hat{\phi}_{\sigma}^{\dagger}(\vec{r}) \left[-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{r}) \right] \hat{\phi}_{\sigma}(\vec{r})$$

$$+ \frac{1}{2} \sum_{\sigma,\sigma'} \int d^3r \int d^3r' \hat{\phi}_{\sigma}^{\dagger}(\vec{r}) \hat{\phi}_{\sigma}(\vec{r}) v(|\vec{r} - \vec{r}'|) \hat{\phi}_{\sigma'}^{\dagger}(\vec{r}') \hat{\phi}_{\sigma'}(\vec{r}')$$

Field operator

$$\hat{\phi}_{\sigma}(\vec{r}) = \sum_{\ell} u_{\ell}(\vec{r}) \hat{a}_{\ell\sigma}$$

$$\int d^3r \ u_\ell^*(\vec{r})u_m(\vec{r}) = \delta_{\ell,m}$$

Fermions

$$\{\hat{a}_{\ell\sigma}, \hat{a}_{m\tau}^{\dagger}\} = \hat{a}_{\ell\sigma}\hat{a}_{m\tau}^{\dagger} + \hat{a}_{m\tau}^{\dagger}\hat{a}_{\ell\sigma} = \delta_{\ell,m}\delta_{\sigma,\tau}$$
$$\{\hat{a}_{\ell\sigma}, \hat{a}_{m\tau}\} = \{\hat{a}_{\ell\sigma}^{\dagger}, \hat{a}_{m\tau}^{\dagger}\} = 0$$

Bosons

$$[\hat{a}_{\ell\sigma}, \hat{a}_{m\tau}^{\dagger}] = \hat{a}_{\ell\sigma} \hat{a}_{m\tau}^{\dagger} - \hat{a}_{m\tau}^{\dagger} \hat{a}_{\ell\sigma} = \delta_{\ell,m} \delta_{\sigma,\tau}$$
$$[\hat{a}_{\ell\sigma}, \hat{a}_{m\tau}] = [\hat{a}_{\ell\sigma}^{\dagger}, \hat{a}_{m\tau}^{\dagger}] = 0$$

$$\hat{H} = \sum_{\sigma} \int d^3r \hat{\phi}_{\sigma}^{\dagger}(\vec{r}) \left[-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{r}) \right] \hat{\phi}_{\sigma}(\vec{r})
+ \frac{1}{2} \sum_{\sigma,\sigma'} \int d^3r \int d^3r' \hat{\phi}_{\sigma}^{\dagger}(\vec{r}) \hat{\phi}_{\sigma}(\vec{r}) v(|\vec{r} - \vec{r}'|) \hat{\phi}_{\sigma'}^{\dagger}(\vec{r}') \hat{\phi}_{\sigma'}(\vec{r}')$$

→ General Hamiltonian with two-body interactions

$$\hat{H} = \sum_{\ell,m,\sigma} K_{\ell m} \hat{a}_{\ell \sigma}^{\dagger} \hat{a}_{m \sigma} + \sum_{\ell_1,\ell_2,m_1,m_2} \sum_{\sigma,\sigma'} I_{\ell_1 \ell_2 m_1 m_2} \hat{a}_{\ell_1 \sigma}^{\dagger} \hat{a}_{\ell_2 \sigma} \hat{a}_{m_1 \sigma'}^{\dagger} \hat{a}_{m_2 \sigma'}$$

$$K_{\ell m} = \int d^3 r u_{\ell}^*(\vec{r}) \left[-\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{r}) \right] u_m(\vec{r})$$

$$I_{\ell_1 \ell_2 m_1 m_2} = \frac{1}{2} \int d^3 r \int d^3 r' u_{\ell_1}^*(\vec{r}) u_{\ell_2}(\vec{r}) v(|\vec{r} - \vec{r}'|) u_{m_1}^*(\vec{r}') u_{m_2}(\vec{r}')$$

Fock space of N-particle fermions expanded by

$$|\Phi\rangle = \sum_{\ell_1, \ell_2, \dots, \ell_N} \sum_{\sigma_1, \sigma_2, \dots, \sigma_N} C_{\ell_1 \ell_2 \dots \ell_N} \hat{a}_{\ell_1 \sigma_1}^{\dagger} \hat{a}_{\ell_2 \sigma_2}^{\dagger} \cdots \hat{a}_{\ell_N \sigma_N}^{\dagger} |\text{vac}\rangle$$

Orthonormalized many-body basis

$$\{\ell_j, \sigma_j\} = \{\ell_1, \sigma_1, \ell_2, \sigma_2, \dots, \ell_N, \sigma_N\}$$

$$|\{\ell_j, \sigma_j\}\rangle = \hat{a}_{\ell_1 \sigma_1}^{\dagger} \hat{a}_{\ell_2 \sigma_2}^{\dagger} \cdots \hat{a}_{\ell_N \sigma_N}^{\dagger} |\text{vac}\rangle$$

$$|\{m_j, \tau_j\}\rangle = \hat{a}_{m_1 \tau_1}^{\dagger} \hat{a}_{m_2 \tau_2}^{\dagger} \cdots \hat{a}_{m_N \tau_N}^{\dagger} |\text{vac}\rangle$$

$$\langle \{m_j, \tau_j\} | \{\ell_j, \sigma_j\} \rangle = \begin{cases} 0 & (\{m_j, \tau_j\} \cup \{\ell_j, \sigma_j\} \neq \{\ell_j, \sigma_j\}) \\ 1 & (\{m_j, \tau_j\} \cup \{\ell_j, \sigma_j\} = \{\ell_j, \sigma_j\}) \end{cases}$$

Common important formula between Hilbert and Fock spaces

Closure by orthonormalized basis

$$1 = \sum_{\mu} |\mu\rangle\langle\mu|$$

$$\langle \mu | \nu \rangle = \delta_{\mu,\nu}$$

$$\left(\sum_{\mu} |\mu\rangle\langle\mu|\right) \times |\Phi\rangle = \left(\sum_{\mu} |\mu\rangle\langle\mu|\right) \times \sum_{\nu} d_{\nu}|\nu\rangle$$

$$= \sum_{\mu} d_{\nu}|\nu\rangle$$

$$= |\Phi\rangle$$

Schrödinger equation $\hat{H}|\Phi\rangle=E|\Phi\rangle$

Hermitian
$$\hat{H}^\dagger = \hat{H}$$
 $H_{\mu\nu} = H_{\nu\mu}^*$

Many-body orthonormalized basis $\langle \mu | \nu \rangle = \delta_{\mu,\nu}$

Closure
$$1 = \sum_{\mu} |\mu\rangle\langle\mu|$$

$$\langle \mu | \times \hat{H} | \Phi \rangle = \langle \mu | \times E | \Phi \rangle$$

$$\Leftrightarrow \sum_{\nu} \langle \mu | \hat{H} | \nu \rangle \langle \nu | \Phi \rangle = E \langle \mu | \Phi \rangle$$

Rewritten Schrödinger equation

$$\sum_{\nu} H_{\mu\nu} d_{\nu} = E d_{\mu}$$

$$H_{\mu\nu} = \langle \mu | \hat{H} | \nu \rangle$$

$$|\Phi\rangle = \sum_{\mu} d_{\mu} |\mu\rangle$$

Eigenvalue Problems of Large and Sparse Matrices

Sparse Matrix

- Particle or orbital number: *N*
- Fock space dimension: exp[Nx const.]
- # of terms in Hamiltonian: Polynomial of N
- → # of matrix elements of Hamiltonian matrix: (Polynomial of M) x exp[N x const.]

For sufficiently large N, (Polynomial of M) x exp[Nx const.] $<< (exp[Nx const.])^2$

Then, the Hamiltonian matrix is sparse

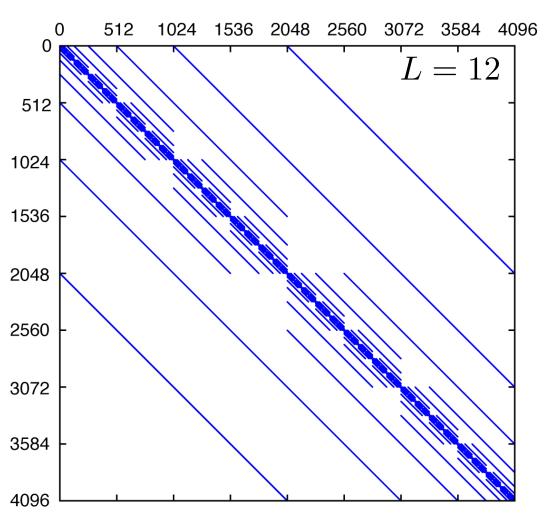
Larger TFIM Revisit

$$\hat{H} = J \sum_{i=0}^{L-1} \hat{S}_i^z \hat{S}_{i+1}^z - \Gamma \sum_{i=0}^{L-1} \hat{S}_i^x$$

-Non-commutative

$$\left[\sum_{i=0}^{L-1} \hat{S}_{i}^{z} \hat{S}_{i+1}^{z}, \sum_{i=0}^{L-1} \hat{S}_{i}^{x}\right] \neq 0$$

- →Quantum fluctuations or Zero point motion
- -Sparse # of elements $\propto O(2^{L})$
- -Solvable
- -Hierarchical matrix?



Why Hamiltonian Matrices are Sparse?

Example: TFIM
$$\hat{H} = J \sum_{i=0}^{\infty} \hat{S}_i^z \hat{S}_{i+1}^z - \Gamma \sum_{i=0}^{\infty} \hat{S}_i^x$$

-Diagonal elements

$$\left[J \sum_{i=0}^{L-1} \hat{S}_{i}^{z} \hat{S}_{\text{mod}(i+1,L)}^{z} \right] |\sigma_{0}\sigma_{1} \cdots \sigma_{L-1}\rangle = \left[\frac{J}{4} \sum_{i=0}^{L-1} (-1)^{\sigma_{i} + \sigma_{\text{mod}(i+1,L)}} \right] \frac{1}{2} |\sigma_{0}\sigma_{1} \cdots \sigma_{L-1}\rangle$$

-Off diagonal elements

$$\hat{S}^x_{\ell} | \sigma_0 \sigma_1 \cdots \sigma_{\ell} \cdots \sigma_{L-1} \rangle = \frac{1}{2} | \sigma_0 \sigma_1 \cdots \overline{\sigma_{\ell}} \cdots \sigma_{L-1} \rangle \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0 \sigma_1 \cdots \sigma_{L-1} \rangle}}} \prod_{\substack{j \leq 1024 \text{ 1536 2048 2560 3072 3584 4096} \\ |\sigma_0$$

Each term in sum of operators maps $|I
angle_{
m d}$ to $|J
angle_{
m d}$

→# of nonzero elements < # of terms in the hamiltonian

Computational and Memory Costs

Matrix-vector product of dense matrix

$$v_i = \sum_{j=0}^{N_{\rm H}-1} A_{ij} u_j$$

Computational: $O((Fock space dimension)^2)$

Memory: $O((Fock space dimension)^2)$

Matrix-vector product of large and sparse matrix

Computational: O(Fock space dimension)

Memory: O(Fock space dimension)

Hamiltonian is not stored in memory

Algorithm for Eigenvalue Problems of Large & Sparse Matrix: Power Method

Min. Eigenvalue of hermitian

Initial vector:
$$|v_1\rangle = \sum_{n=0}^{\infty} c_n |n\rangle$$

Prameter:
$$\max_{n} \{E_n\} \leq \Lambda$$

$$\hat{H}|n\rangle = E_n|n\rangle$$

$$\langle n'|n\rangle = \delta_{n',n}$$

$$E_0 \le E_1 \le \cdots$$

$$\lim_{m \to +\infty} \frac{(\Lambda - \hat{H})^m |v_1\rangle}{\sqrt{\langle v_1 | (\Lambda - \hat{H})^{2m} |v_1\rangle}} = |0\rangle$$

$$(\Lambda - \hat{H})^m |v_1\rangle = \sum_n (\Lambda - E_n)^m c_n |n\rangle$$

$$\sum_{\substack{n > 0 \\ m \to +\infty}} (\Lambda - E_n)^{2m} |c_n|^2$$

$$\lim_{\substack{n > 0 \\ (\Lambda - E_0)^{2m} |c_0|^2}} = 0$$

Advanced Algorithm: Krylov Subspace Method

Krylov subspace

$$\mathcal{K}_m(\hat{H},|v_1\rangle) = \operatorname{span}\{|v_1\rangle, \hat{H}|v_1\rangle, \dots, \hat{H}^{m-1}|v_1\rangle\}$$

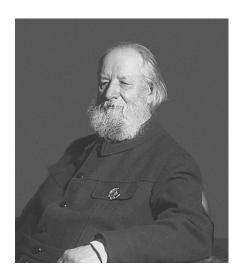
Shift invariance:

$$\mathcal{K}_m(\hat{H},|v_1\rangle) = \mathcal{K}_m(\hat{H}+z\mathbf{1},|v_1\rangle)$$

Krylov subspace method:

- -Lanczos method (symmetric/hermitian), Arnoldi method (general matrix)
- -Conjugate gradient method (CG method) (many variation)

Krylov Subspace Method for Sparse and Huge Matrices



Alexey Krylov
Aleksey Nikolaevich Krylov
1863-1945
Russian naval engineer and applied mathematician

Krylov subspace

$$A \in \mathbb{C}^{L \times L}$$

$$\mathcal{K}_n(A, \vec{b}) = \operatorname{span}\{\vec{b}, A\vec{b}, \dots, A^{n-1}\vec{b}\}$$

Numerical cost to construct K_n : $\mathcal{O}(\text{nnz}(A) \times n)$

Numerical cost to orthogonalize K_n : $\mathcal{O}(L \times n^2)$

Cornelius Lanczos 1950 Walter Edwin Arnoldi 1951 *nnz: Number of non-zero entries/elements

Krylov Subspace Method

from SIAM News, Volume 33, Number 4

The Best of the 20th Century: Editors Name Top 10 Algorithms

By Barry A. Cipra

1950: Magnus Hestenes, Eduard Stiefel, and Cornelius Lanczos, all from the Institute for Numerical Analysis at the National Bureau of Standards, initiate the development of **Krylov subspace iteration methods**.

These algorithms address the seemingly simple task of solving equations of the form Ax = b. The catch, of course, is that A is a huge $n \times n$ matrix, so that the algebraic answer x = b/A is not so easy to compute. (Indeed, matrix "division" is not a particularly useful concept.) Iterative methods—such as solving equations of $a + b - Ax_i$ with a simpler matrix $a + b - Ax_i$ that it is ideally "close" to a - b lead to the study of Krylov subspaces. Named

the form $Kx_{i+1} = Kx_i + b - Ax_i$ with a simpler matrix K that's ideally "close" to A—lead to the study of Krylov subspaces. Named for the Russian mathematician Nikolai Krylov, Krylov subspaces are spanned by powers of a matrix applied to an initial "remainder" vector $r_0 = b - Ax_0$. Lanczos found a nifty way to generate an orthogonal basis for such a subspace when the matrix is symmetric. Hestenes and Stiefel proposed an even niftier method, known as the conjugate gradient method, for systems that are both symmetric and positive definite. Over the last 50 years, numerous researchers have improved and extended these algorithms. The current suite includes techniques for non-symmetric systems, with acronyms like GMRES and Bi-CGSTAB. (GMRES and Bi-CGSTAB premiered in SIAM Journal on Scientific and Statistical Computing, in 1986 and 1992, respectively.)

Initial:
$$\beta_1 = 0$$
, $|v_0\rangle = 0$
for $j = 1, 2, ..., m$ do
 $|w_j\rangle = \hat{H}|v_j\rangle - \beta_j|v_{j-1}\rangle$
 $\alpha_j = \langle w_j|v_j\rangle$
 $|w_j\rangle \leftarrow |w_j\rangle - \alpha_j|v_j\rangle$
 $\beta_{j+1} = \sqrt{\langle w_j|w_j\rangle}$
 $|v_{j+1}\rangle = |w_j\rangle/\beta_{j+1}$

$$\alpha_j = \langle v_j | \hat{H} | v_j \rangle$$

$$\beta_j = \langle v_{j-1} | \hat{H} | v_j \rangle = \langle v_j | \hat{H} | v_{j-1} \rangle \leftarrow \text{Confirm}$$

Orthogonalization

$$|v_{j}\rangle = \frac{\hat{H}|v_{j-1}\rangle - \sum_{\ell=1}^{j-1} |v_{\ell}\rangle\langle v_{\ell}|\hat{H}|v_{j-1}\rangle}{\langle v_{j}|\hat{H}|v_{j-1}\rangle}$$

$$\langle v_{\ell} | \hat{H} | v_{j-1} \rangle = \begin{cases} 0 & (\ell \leq j-3) \\ \beta_{j-1} & (\ell = j-2) \\ \alpha_{j-1} & (\ell = j-1) \end{cases} \leftarrow \text{Confirm}$$

Initial:
$$\beta_1 = 0$$
, $|v_0\rangle = 0$
for $j = 1, 2, ..., m$ do

$$|w_j\rangle = \hat{H}|v_j\rangle - \beta_j|v_{j-1}\rangle$$

$$\alpha_j = \langle w_j|v_j\rangle$$

$$|w_j\rangle \leftarrow |w_j\rangle - \alpha_j|v_j\rangle$$

$$\beta_{j+1} = \sqrt{\langle w_j|w_j\rangle}$$

$$|v_{j+1}\rangle = |w_j\rangle/\beta_{j+1}$$

$$\alpha_{j} = \langle v_{j} | \hat{H} | v_{j} \rangle$$

$$\langle v_{j} | v_{k} \rangle = \delta_{j,k}$$

$$\beta_{j} = \langle v_{j-1} | \hat{H} | v_{j} \rangle = \langle v_{j} | \hat{H} | v_{j-1} \rangle$$

Hamiltonian projected onto m D Krylov subsace

$$H_{m} = \begin{pmatrix} \alpha_{1} & \beta_{2} & & & & & & & & & \\ \beta_{2} & \alpha_{2} & \beta_{3} & & & & & & \\ & \beta_{3} & \alpha_{3} & \ddots & & & & & \\ & & \ddots & \ddots & \beta_{m-1} & & & \\ & & & \beta_{m-1} & \alpha_{m-1} & \beta_{m} & \\ & & & \beta_{m} & \alpha_{m} \end{pmatrix}$$

Eigenvalues of projected Hamiltonian

→ Approximate eigenvalues of original Hamiltonian

Lanczos Method: # of Vectors Required

Initial:
$$\beta_1 = 0$$
, $|v_0\rangle = 0$
for $j = 1, 2, ..., m$ do
$$|w_j\rangle \leftarrow \hat{H}|v_j\rangle - \beta_j|v_{j-1}\rangle \qquad |v_{j-1}\rangle \rightarrow |w_j\rangle, |v_j\rangle$$

$$\alpha_j = \langle w_j|v_j\rangle \qquad |w_j\rangle, |v_j\rangle$$

$$|w_j\rangle \leftarrow |w_j\rangle - \alpha_j|v_j\rangle \qquad |w_j\rangle, |v_j\rangle$$

$$\beta_{j+1} = \sqrt{\langle w_j|w_j\rangle} \qquad |w_j\rangle, |v_j\rangle$$

$$|v_{j+1}\rangle = |w_j\rangle/\beta_{j+1} \qquad |w_j\rangle \rightarrow |v_{j+1}\rangle, |v_j\rangle$$

Convergence of Lanczos Method

Yousef Saad, *Numerical Methods for Large Eigenvalue Problems* (2nd ed) The Society for Industrial and Applied Mathematics 2011

Assumption:
$$\lambda_1 > \lambda_2 > \cdots > \lambda_n$$

Convergence theorem for the largest eigenvalue

$$0 \leq \lambda_1 - \lambda_1^{(m)} \leq (\lambda_1 - \lambda_n) \left[\frac{\tan \theta(|v_1\rangle, |0\rangle)}{C_{m-1}(1+2\gamma_1)} \right]^2$$

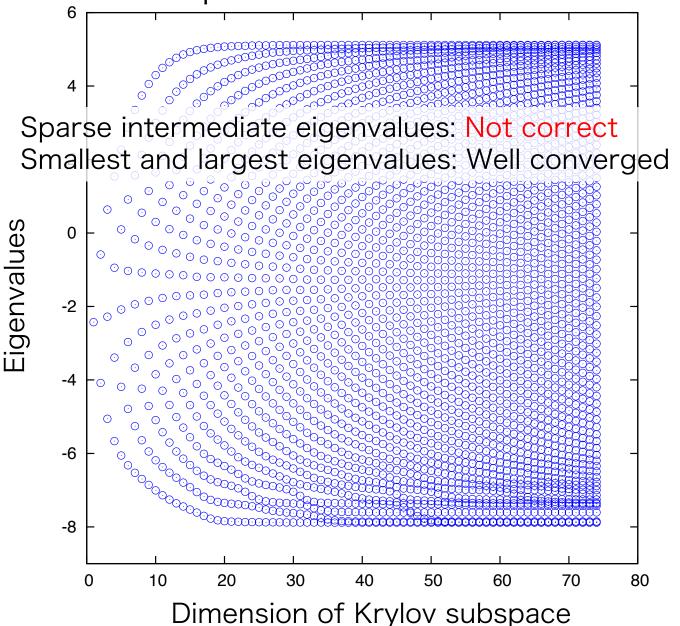
$$\sim 4(\lambda_1 - \lambda_n) \left[\tan \theta(|v_1\rangle, |0\rangle) \right]^2 e^{-4\sqrt{\gamma_1}m}$$

$$\gamma_1 = \frac{\lambda_1 - \lambda_2}{\lambda_2 - \lambda_n}$$

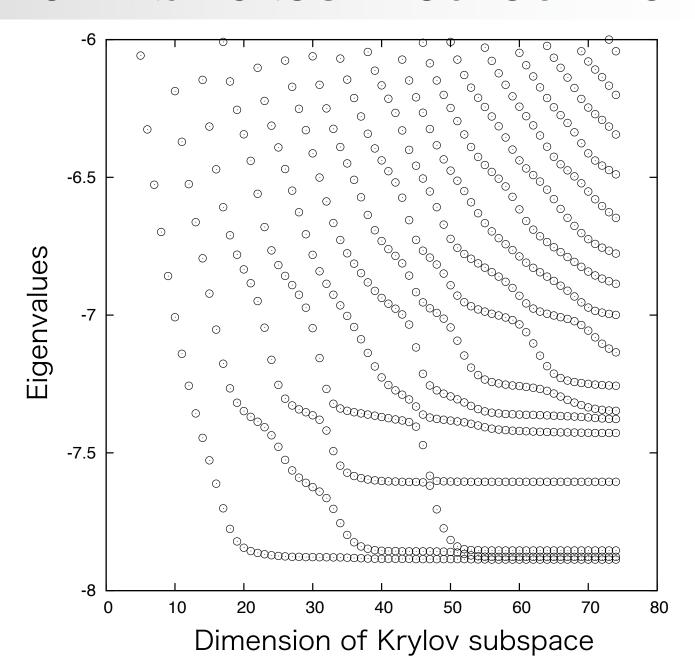
$$C_k(t) = \frac{1}{2} \left[\left(t + \sqrt{t^2 - 1} \right)^k + \left(t + \sqrt{t^2 - 1} \right)^{-k} \right]_{36}$$

How Lanczos Method Works

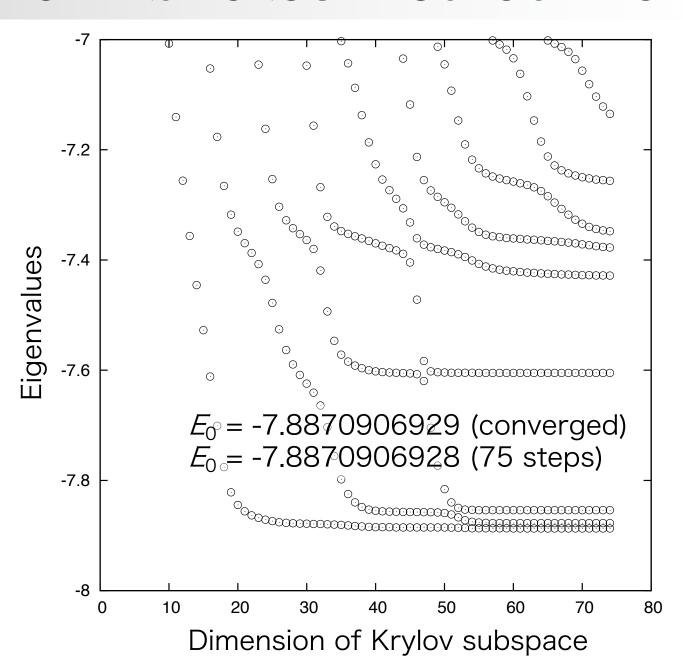
24 site cluster of Kitaev- Γ model (frustrated S=1/2 spins) Dimension of Fock space: $2^{24}=16777216$



How Lanczos Method Works



How Lanczos Method Works



Exercise on Jupyter Notebook

-Lanczos method for spin models implemented in vmc_helium4_2021.ipynb

Lecture Schedule

```
#1 Many-body problems in physics
#2 Why many-body problem is hard to solve
#3 Classical statistical model and numerical simulation
#4 Classical Monte Carlo method and its applications
#5 Molecular dynamics and its application
#6 Extended ensemble method for Monte Carlo methods
#7 Quantum lattice models and numerical approaches
#8 Quantum Monte Carlo methods
#9 Applications of quantum Monte Carlo methods
   -Path integral & applications
#10 Linear algebra of large and sparse matrices for
   quantum many-body problems
#11 Krylov subspace methods and their applications to
   quantum many-body problems
#12 Large sparse matrices and quantum statistical mechanics
```

42

#13 Parallelization for many-body problems