Parallel Multi-Coordinate Descent Methods for FCI

Yuejia Zhang[1] Weiguo Gao[1,2,3,4] Yingzhou Li[1,3,4]

[1]School of Mathematical Sciences, Fudan University
[2]School of Data Science, Fudan University
[3]Shanghai Key Laboratory for Contemporary Applied Mathematics, Shanghai
[4]Key Laboratory of Computational Physical Sciences (MOE), Shanghai

Problem: Searching for Ground-State

The ground-state of a chemical system given by the many-body time-independent Schrödinger Equation

$$\hat{H}|\mathbf{\Phi_0}\rangle = E_0|\mathbf{\Phi_0}\rangle,$$

where $|\mathbf{\Phi_0}\rangle = \mathbf{\Phi_0}(r_1, \ldots, r_{n_{\text{elec}}}), r_i \in \mathbb{R}^3$.
Under Born–Oppenheimer approximation, the Hamiltonian operator with n_{nuc} nuclei and n_{elec} electrons is

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{n_{\text{elec}}} \nabla_i^2 + \sum_{i=1}^{n_{\text{elec}}} V_{\text{ext}}(r_i; \{R_I\}_{I=1}^{n_{\text{nuc}}}) + \sum_{i<j}^{n_{\text{elec}}} \frac{1}{\|r_i - r_j\|}.$$

Full Configuration Interaction (FCI) Method

Start from one-electron spin-orbitals $\{\chi_p\}_{p=1}^{n_{\text{orb}}}$, obtained from Hartree–Fock procedure.
Each configuration is a *Slater determinant* (or, *anti-symmetrized tensor products*):

$$|D_i\rangle = |\chi_{p_1}\chi_{p_2}\cdots\chi_{p_n}\rangle,$$

with $n = n_{\text{elec}}$ and $1 \le p_1 < p_2 < \cdots < p_n \le n_{\text{orb}}$.
Schrödinger equation can be transformed to FCI eigenvalue problem

$$H\mathbf{c} = E_0\mathbf{c}, \quad H \in \mathbb{R}^{N_{\text{FCI}}\times N_{\text{FCI}}}, \quad \mathbf{c} \in \mathbb{R}^{N_{\text{FCI}}}.$$

FCI variational space dimension: $N_{\text{FCI}} = \binom{n_{\text{orb}}}{n_{\text{elec}}}$. Combinatorial explosion!

Molecule	Basis	Electrons	Spin–Orbitals	Dimension	Memory
H_2O	cc-pVDZ	10	48	$\sim 10^8$	~ 1 GB
N_2	cc-pVDZ	14	56	$\sim 10^{11}$	~ 1 TB
N_2	cc-pVTZ	14	120	$\sim 10^{16}$	~ 100 PB
Cr_2	Ahlrichs	48	84	$\sim 10^{22}$	-

Sparsity of Hamiltonian Matrix H

$H_{ij} = \langle D_i|\hat{H}|D_j\rangle$.

- If $|D_i\rangle = a_r^\dagger a_p |D_j\rangle$, $H_{ij} = \langle r|\hat{h}|p\rangle + \sum_k \langle rk||pk\rangle$.
- If $|D_i\rangle = a_r^\dagger a_s^\dagger a_p a_q |D_j\rangle$, $H_{ij} = \langle rs||pq\rangle$.
- Otherwise, $H_{ij} = 0$.

Consequence: H has $O(n_{\text{elec}}^2 n_{\text{orb}}^2)$ entries per row.

Coordinate Descent FCI (CDFCI)[1]

Consider the unconstrained minimization problem

$$\min_{\mathbf{c}\in\mathbb{R}^{N_{\text{FCI}}}} f(\mathbf{c}) = \min_{\mathbf{c}} \|H + \mathbf{c}\mathbf{c}^{\mathrm{T}}\|_F^2,$$

with gradient $\nabla f(c) = 4H\mathbf{c} + 4(\mathbf{c}^{\mathrm{T}}\mathbf{c})\mathbf{c}$, and Hessian $\nabla^2 f(c) = 4H + 8\mathbf{c}\mathbf{c}^{\mathrm{T}} + 4(\mathbf{c}^{\mathrm{T}}\mathbf{c})I$.
Its stationary points are $0, \pm\sqrt{-\lambda_1}\mathbf{v}_1, \ldots, \pm\sqrt{-\lambda_m}\mathbf{v}_m$ $(\cdots < \lambda_m < 0 < \lambda_{m+1} < \cdots)$.
Among them, only two are local minimizers $\pm\sqrt{-\lambda_1}\mathbf{v}_1$ (which are also global minimizers), the others are all strict saddle points.
This ensures convergence to the ground state $\pm\mathbf{c}$, given a good starting point (e.g., Hartree–Fock ground state).

Algorithm:

Initialize $\mathbf{c}^{(0)}$, $\mathbf{b}^{(0)} = H\mathbf{c}^{(0)}$.
For iteration $\ell = 1, 2, \ldots$

1. Select coordinate

$$i^{(\ell)} = \arg\max_i |\nabla_i f(\mathbf{c}^{(\ell-1)})|.$$

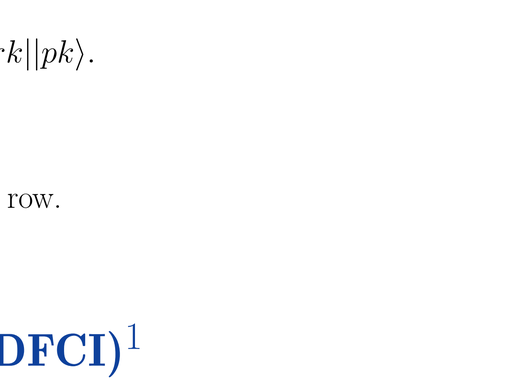
2. Find stepsize by exact line search

$$\alpha^{(\ell)} = \arg\min_{\alpha} f(\mathbf{c}^{(\ell-1)} + \alpha\mathbf{e}_{i^{(\ell)}}).$$

3. Update

$$\mathbf{c}^{(\ell)} = \mathbf{c}^{(\ell-1)} + \alpha^{(\ell)}\mathbf{e}_{i^{(\ell)}},$$
$$\mathbf{b}^{(\ell)} = \mathbf{b}^{(\ell-1)} + \alpha^{(\ell)}H_{:,i^{(\ell)}}$$

.



[1] Z. Wang, Y. Li, J. Lu, J. Chem. Theory Comput., 2019.

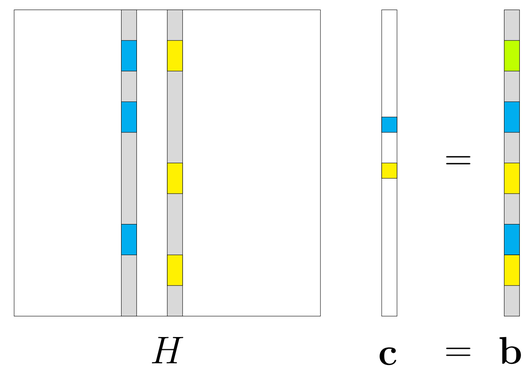
Multi-Coordinate Descent FCI (mCDFCI)[2]

Algorithm:

Initialize $\mathbf{c}^{(0)}$, $\mathbf{b}^{(0)} = H\mathbf{c}^{(0)}$.
For iteration $\ell = 1, 2, \ldots$

1. Select coordinate

$$i^{(\ell)} = \arg\max_i |\nabla_i f(\mathbf{c}^{(\ell-1)})|,$$
$$j^{(\ell)} = \arg\max_{j\neq i^{(\ell)}} |\nabla_j f(\mathbf{c}^{(\ell-1)})|.$$



2. Find scaling factor and stepsizes

$$\gamma^{(\ell)}, \alpha^{(\ell)}, \beta^{(\ell)} = \arg\min_{\gamma,\alpha,\beta} f(\gamma\mathbf{c}^{(\ell-1)} + \alpha\mathbf{e}_{i^{(\ell)}} + \beta\mathbf{e}_{j^{(\ell)}}).$$

3. Update

$$\mathbf{c}^{(\ell)} = \gamma^{(\ell)}\mathbf{c}^{(\ell-1)} + \alpha^{(\ell)}\mathbf{e}_{i^{(\ell)}} + \beta^{(\ell)}\mathbf{e}_{j^{(\ell)}},$$
$$\mathbf{b}^{(\ell)} = \gamma^{(\ell)}\mathbf{b}^{(\ell-1)} + \alpha^{(\ell)}H_{:,i^{(\ell)}} + \beta^{(\ell)}H_{:,j^{(\ell)}}.$$

The update rule is modified to enable exact line search. Scaling factor γ and stepsizes α, β are obtained from

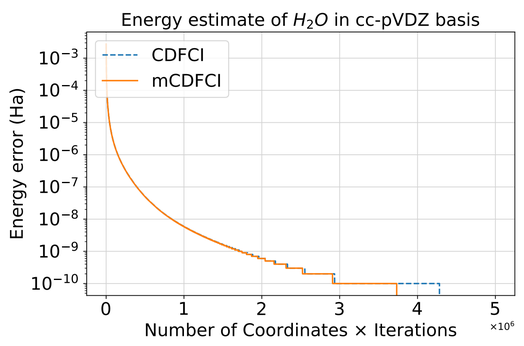
$$\arg\min_{\gamma,\alpha,\beta} f(\gamma\tilde{\mathbf{c}} + \alpha\mathbf{e}_i + \beta\mathbf{e}_j) = f(\begin{bmatrix}\tilde{\mathbf{c}} & \mathbf{e}_i & \mathbf{e}_j\end{bmatrix}\begin{bmatrix}\gamma\\ \alpha\\ \beta\end{bmatrix})$$
$$= \left\| H + \begin{bmatrix}\tilde{\mathbf{c}} & \mathbf{e}_i & \mathbf{e}_j\end{bmatrix}\begin{bmatrix}\gamma\\ \alpha\\ \beta\end{bmatrix}\begin{bmatrix}\gamma & \alpha & \beta\end{bmatrix}\begin{bmatrix}\tilde{\mathbf{c}}^{\mathrm{T}}\\ \mathbf{e}_i^{\mathrm{T}}\\ \mathbf{e}_j^{\mathrm{T}}\end{bmatrix}\right\|_F^2$$
$$= \left\| \underbrace{\begin{bmatrix}\tilde{\mathbf{c}}^{\mathrm{T}}\\ \mathbf{e}_i^{\mathrm{T}}\\ \mathbf{e}_j^{\mathrm{T}}\end{bmatrix} H \begin{bmatrix}\tilde{\mathbf{c}} & \mathbf{e}_i & \mathbf{e}_j\end{bmatrix}}_{\in\mathbb{R}^{3\times 3}} + \begin{bmatrix}\gamma\\ \alpha\\ \beta\end{bmatrix}\begin{bmatrix}\gamma & \alpha & \beta\end{bmatrix}\right\|_F^2 + \text{constant}.$$

This framework can be easily extended to multi-coordinate updates.

[2] Y. Zhang, W. Gao, Y. Li, J. Chem. Theory Comput., 2025.

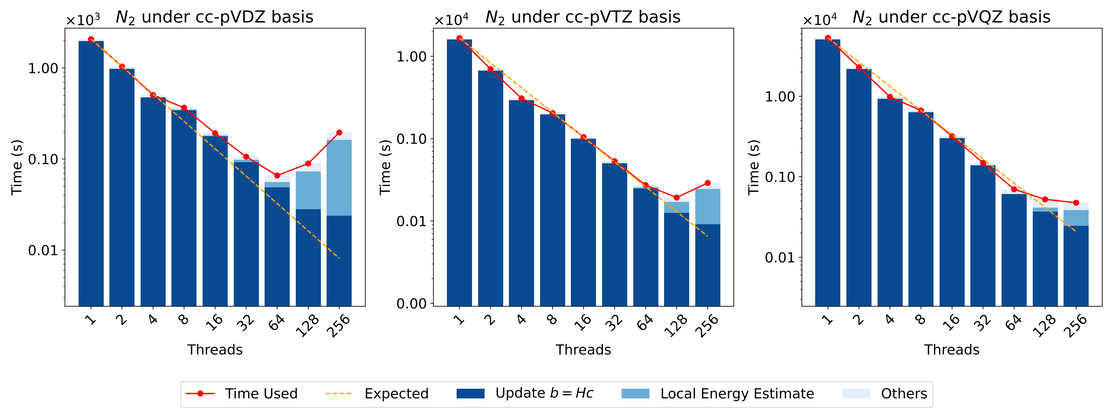
Improvement Over CDFCI

With OpenMP parallelization enabled, mCDFCI achieves a 2 to 3× speedup over CDFCI, thanks to the increased workload per iteration that can be executed concurrently.



Strong Scalability

Our method mCDFCI exhibits strong scalability as the number of threads increases from 1 to 256.



LaTeX TikZposter