Parallel Coordinate Descent Methods for Full Configuration Interaction

Yuejia Zhang   Weiguo Gao   Yingzhou Li

Fudan University

ICIAM, August 2023

The Many-Body Time-Independent Schrödinger Equation

- Ground-state wavefunction $|\boldsymbol{\Phi_0}\rangle$ and energy E_0 of a chemical system given by
$$\hat{H}|\boldsymbol{\Phi_0}\rangle = E_0|\boldsymbol{\Phi_0}\rangle,$$
where $\hat{H} = -\frac{1}{2}\sum_{i=1}^{n_{\mathrm{elec}}} \nabla_i^2 + V(x_1, x_2, \ldots, x_{n_{\mathrm{elec}}})$.
- Application: Molecular property calculations
- Computational Methods:
    - Density Functional Theory (DFT)
    - Many-Body Perturbation Theory (MBPT)
    - Coupled Cluster (CC)
    - **Full Configuration Interaction (FCI)**: for strongly-correlated system

Configuration Interaction Method

- Based on one-electron spin–orbitals $\{\chi_p\}_{p=1}^{n_{\text{orb}}}$ from Hartree–Fock procedure
- Wavefunction approximated as linear combination of anti-symmetrized tensor products (Slater determinants)

$$|\mathbf{\Phi_0}\rangle = \sum_{i=1}^{N_{\text{FCI}}} c_i |D_i\rangle = \sum_{i=1}^{N_{\text{FCI}}} c_i |\chi_{p_i}\chi_{p_j}\cdots\chi_{p_k}\rangle$$

- FCI variational space dimension: $N_{\text{FCI}} = \binom{n_{\text{orb}}}{n_{\text{elec}}}$
- Schrödinger equation transformed to 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 Matrix Properties

Entry: $H_{ij} = \langle D_i | \hat{H} | D_j \rangle$, not guaranteed to be non-negative.

- Symmetric: $H_{ij} = H_{ji}$.
- Sparse: $H_{ij} = 0$ if $|D_i\rangle$ and $|D_j\rangle$ differ by more than two orbitals. H has $O(n_{\text{elec}}^2 n_{\text{orb}}^2)$ entries per row.
- Ground-state eigenvalue $E_0 < 0$, ground-state eigenvector $\mathbf{v_0}$ sparse in the sense of truncation.

Memory usage

Table: Different Molecule Systems and Storage cost

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

Solution: Select a part of configurations!

Related Work and Our Contribution

- Related work:
  - Configuration Interaction by Perturbatively Selecting Iteration (CIPSI)
  - Adaptive Configuration Interaction (ACI)
  - Adaptive Sampling Configuration Interaction (ASCI)
  - Heat-Bath Configuration Interaction (HCI)
  - Semistochastic HCI (SHCI)
- Our contribution:
  - Performs configuration selection using coordinate descent
  - Visits important determinants efficiently
  - Captures the significant part of FCI space for ground state approximation

Coordinate Descent FCI (CDFCI)

FCI eigenvalue problem to unconstrained minimization problem

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

- The only two local minimizers are $\pm\sqrt{E_0}\mathbf{v_0}$.
- Ensures convergence to the ground state wavefunction, given a good starting point (Hartree–Fock ground state).

Coordinate gradient descent method

- Minimizes computational costs by avoiding operations with the entire Hamiltonian matrix.
- In each iteration, only *one coordinate* of the optimizing vector is updated.
- Computation for updating involves only *one column* of the Hamiltonian matrix.

CDFCI Framework

Store $\mathbf{c}$ and $\mathbf{b} = H\mathbf{c}$ in memory.
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 $\mathbf{c}^{(\ell)} = \mathbf{c}^{(\ell-1)} + \alpha^{(\ell)} \mathbf{e}_{i^{(\ell)}}$, $\mathbf{b}^{(\ell)} = \mathbf{b}^{(\ell-1)} + \alpha^{(\ell)} H_{:,i^{(\ell)}}$.

Remark: Gradient
$\nabla f(\mathbf{c}) = 4H\mathbf{c} + 4\mathbf{c}^\top \mathbf{c}\mathbf{c} = 4\mathbf{b} + 4\mathbf{c}^\top \mathbf{c}\mathbf{c}$.

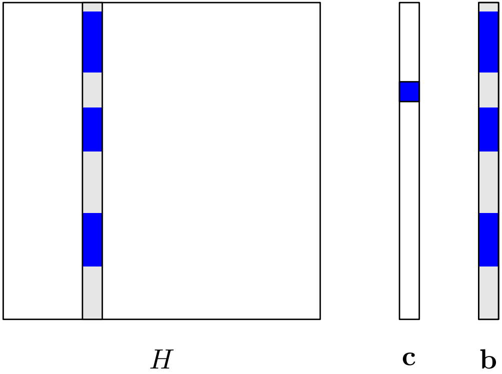


Figure: Update for one coordinate.

CDFCI Framework - for Two Coordinates?

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 stepsize
   $\alpha^{(\ell)} = \arg\min_\alpha f(\mathbf{c}^{(\ell-1)} + \alpha \mathbf{e}_{i^{(\ell)}})$,
   $\beta^{(\ell)} = \arg\min_\beta f(\mathbf{c}^{(\ell-1)} + \beta \mathbf{e}_{j^{(\ell)}})$.
3. Update
   $\mathbf{c}^{(\ell)} = \mathbf{c}^{(\ell-1)} + \alpha^{(\ell)} \mathbf{e}_{i^{(\ell)}} + \beta^{(\ell)} \mathbf{e}_{j^{(\ell)}}$,
   $\mathbf{b}^{(\ell)} = \mathbf{b}^{(\ell-1)} + \alpha^{(\ell)} H_{:,i^{(\ell)}} + \beta^{(\ell)} H_{:,j^{(\ell)}}$.

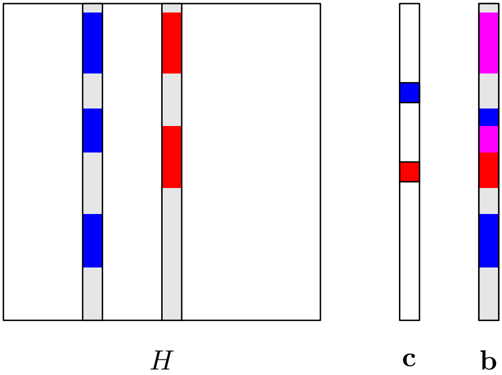


Figure: Update for two coordinates.

CDFCI Framework - Exact Line Search?

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 stepsize $\alpha^{(\ell)}, \beta^{(\ell)} = \arg\min_{\alpha,\beta} f(\mathbf{c}^{(\ell-1)} + \alpha \mathbf{e}_{i^{(\ell)}} + \beta \mathbf{e}_{j^{(\ell)}})$.
3. Update
   $\mathbf{c}^{(\ell)} = \mathbf{c}^{(\ell-1)} + \alpha^{(\ell)} \mathbf{e}_{i^{(\ell)}} + \beta^{(\ell)} \mathbf{e}_{j^{(\ell)}}$,
   $\mathbf{b}^{(\ell)} = \mathbf{b}^{(\ell-1)} + \alpha^{(\ell)} H_{:,i^{(\ell)}} + \beta^{(\ell)} H_{:,j^{(\ell)}}$.

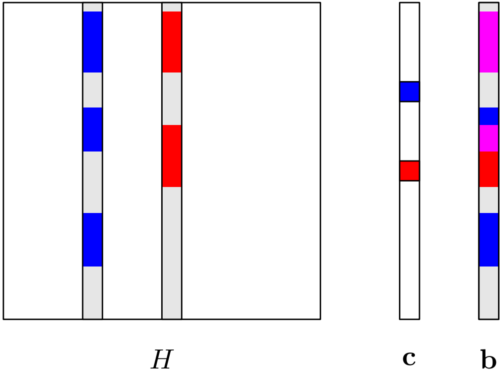


Figure: Update for two coordinates.

Add a Scalar γ for Exact Line Search

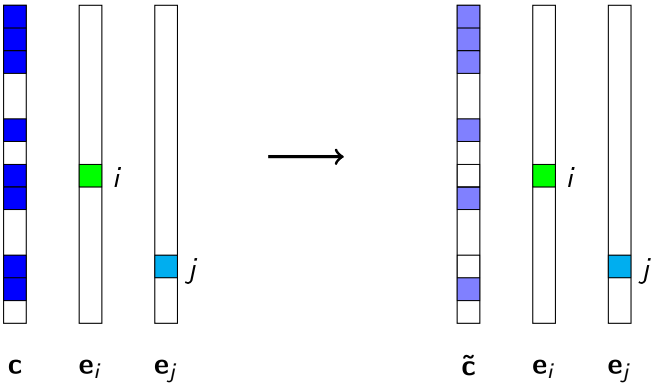
Modify the minimization problem from $\min_{\alpha,\beta} f(\mathbf{c} + \alpha \mathbf{e}_i + \beta \mathbf{e}_j)$ to

$$\min_{\gamma,\alpha,\beta} f(\gamma \mathbf{c} + \alpha \mathbf{e}_i + \beta \mathbf{e}_j) = f\left(\begin{bmatrix} \mathbf{c} & \mathbf{e}_i & \mathbf{e}_j \end{bmatrix} \begin{bmatrix} \gamma \\ \alpha \\ \beta \end{bmatrix}\right)$$

$$= \left\| H + \begin{bmatrix} \mathbf{c} & \mathbf{e}_i & \mathbf{e}_j \end{bmatrix} \begin{bmatrix} \gamma \\ \alpha \\ \beta \end{bmatrix} \begin{bmatrix} \gamma & \alpha & \beta \end{bmatrix} \begin{bmatrix} \mathbf{c}^T \\ \mathbf{e}_i^T \\ \mathbf{e}_j^T \end{bmatrix} \right\|_F^2 .$$

Matrix Orthogonalization

Construct $\begin{bmatrix} \tilde{\mathbf{c}} & \mathbf{e}_i & \mathbf{e}_j \end{bmatrix}$, where $\|\tilde{\mathbf{c}}\|_2 = 1$, $(\tilde{\mathbf{c}}, \mathbf{e}_i) = 0$, $(\tilde{\mathbf{c}}, \mathbf{e}_j) = 0$.



Add γ and $\tilde{\mathbf{c}}$ for Exact Line Search

Modify the minimization problem from $\min_{\alpha,\beta} f(\mathbf{c} + \alpha\mathbf{e}_i + \beta\mathbf{e}_j)$ to

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

Extension to Multi Coordinate Descent FCI

- Select a set of coordinates $I = \{i_1, \ldots, i_k\}, 1 \leq i_j \leq N_{\mathrm{FCI}}$ based on gradient $\nabla f(\mathbf{c}) = 4H\mathbf{c} + 4\mathbf{c}^{\mathrm{T}}\mathbf{c}\mathbf{c}$.
- Denote $\mathcal{E}_I = [e_{i_1}, \ldots, e_{i_k}] \in \mathbb{R}^{N_{\mathrm{FCI}} \times k}$.
- The update is given by

$$\mathbf{c} \leftarrow \gamma\mathbf{c} + \mathcal{E}_I\mathbf{a}.$$

- The values of γ and $\mathbf{a}$ are given by the eigenvector of

$$\begin{bmatrix} \tilde{\mathbf{c}}^{\mathrm{T}} \\ \mathcal{E}_I^{\mathrm{T}} \end{bmatrix} H \begin{bmatrix} \tilde{\mathbf{c}} & \mathcal{E}_I \end{bmatrix} \in \mathbb{R}^{(k+1)\times(k+1)}$$

  corresponding to the minimal eigenvalue $\lambda_{\min}$, which is the current energy estimate.

Implementation Details

- Each thread updates one coordinate: $\mathbf{\Delta b} = H_{:,j} a_j$.
- **c** and **b** are stored in a hash table that allows parallel read-write operations.
- The scaling factor γ is multiplied onto the scale of **c** at each iteration.

Scalability

Table: Speedup of mCDFCI for different number of coordinates (threads) of H_2O in cc-pVDZ basis

		Wall Time (s)	Speedup on k Cores					
Energy	Error	on Single Core	2	4	8	16	32	64
-76.2318601	10^{-2}	9.0	1.9×	3.7×	5.0×	7.0×	10.4×	13.1×
-76.2408601	10^{-3}	292.9	2.0×	3.7×	4.8×	8.2×	14.3×	20.3×
-76.2417601	10^{-4}	1837.6	2.0×	3.5×	6.1×	8.3×	16.2×	22.5×
-76.2418501	10^{-5}	9016.7	2.1×	4.1×	8.0×	12.3×	21.1×	29.0×
-76.2418591	10^{-6}	32931.7	2.1×	4.5×	9.3×	16.2×	29.1×	40.7×

Overall Speedup

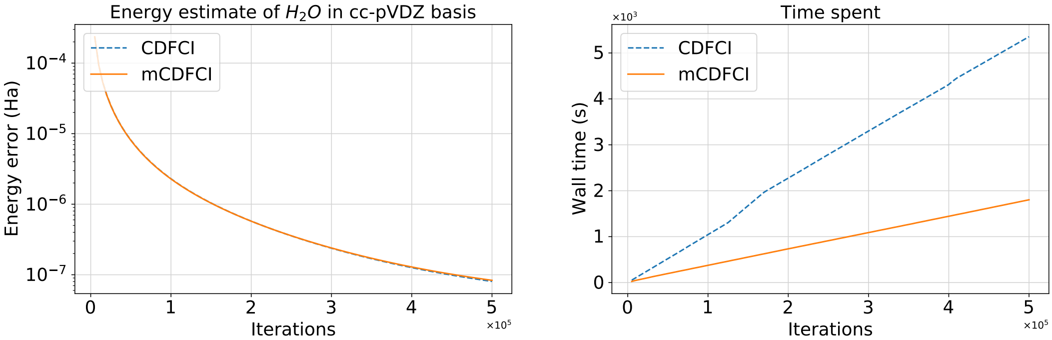


Figure: Speedup of mCDFCI compared with CDFCI(2019), both in 64 threads. We perform k coordinates ($k = 64$) descent per iteration for the original CDFCI.

References I

Adam A Holmes, Norm M Tubman, and CJ Umrigar.
Heat-bath configuration interaction: An efficient selected configuration interaction algorithm inspired by heat-bath sampling.
Journal of chemical theory and computation, 12(8):3674–3680, 2016.

Bernard Huron, JP Malrieu, and P Rancurel.
Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions.
The Journal of Chemical Physics, 58(12):5745–5759, 1973.

Jeffrey B Schriber and Francesco A Evangelista.
Adaptive configuration interaction for computing challenging electronic excited states with tunable accuracy.
Journal of chemical theory and computation, 13(11):5354–5366, 2017.

References II

Sandeep Sharma, Adam A Holmes, Guillaume Jeanmairet, Ali Alavi, and Cyrus J Umrigar.
Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory.
Journal of chemical theory and computation, 13(4):1595–1604, 2017.

Norm M Tubman, Joonho Lee, Tyler Y Takeshita, Martin Head-Gordon, and K Birgitta Whaley.
A deterministic alternative to the full configuration interaction quantum monte carlo method.
The Journal of chemical physics, 145(4), 2016.

Zhe Wang, Yingzhou Li, and Jianfeng Lu.
Coordinate descent full configuration interaction.
Journal of Chemical Theory and Computation, 15(6):3558–3569, 2019.
PMID: 31042383.

Thanks for Your Attention!