Convergent and orthogonality preserving schemes for approximating the Kohn-Sham orbitals

Xiaoying Dai

LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing
Academy of Mathematics and Systems Science
Chinese Academy of Sciences

May 5, 2022

A joint work with Qiao Wang, Liwei Zhang and Aihui Zhou

Outline

1. Motivation
2. An Extended gradient flow model
3. Some orthogonality preserving iteration schemes
4. Numerical experiments
5. Concluding remarks

Outline

Outline

Outline

Outline

1. Motivation
2. An Extended gradient flow model
3. Some orthogonality preserving iteration schemes
4. Numerical experiments
5. Concluding remarks

Outline

1. Motivation
2. An Extended gradient flow model
3. Some orthogonality preserving iteration schemes
4. Numerical experiments
5. Concluding remarks

The total energy for the full interacting many-body system can be expressed as

$$E(\{u_i\}) = \frac{1}{2}\sum_{i=1}^{N}\int_{\mathbb{R}^3}|\nabla u_i(x)|^2 dx + \int_{\mathbb{R}^3} V_{ne}(x)\rho(x)dx + \frac{1}{2}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\frac{\rho(x)\rho(y)}{|x-y|}dxdy + E_{xc}(\rho(x)).$$

- orbitals $\{u_i\}_{i=1}^N$ satisfying

$$\int_{\mathbb{R}^3} u_i u_j = \delta_{ij}, \quad 1 \le i,j \le N$$

- $\rho = \sum_{i=1}^N |u_i|^2$
- $V_{ne}(r) = -\sum_{q=1}^M \frac{Z_q}{|r-R_q|}$
- $E_{xc}(\rho)$ is the exchange-correlation energy

M is the number of atoms, N is the number of electrons, Z_q is the atomic number of the q-th atom, R_q is the position of the q-th atom.

Kohn-Sham DFT Models

- Constrained optimization problem:

$$\min_{U=(u_1,\cdots,u_N)\in(\mathrm{H}^1(\mathbb{R})^3)^N,\int_{\mathbb{R}^3} u_iu_j=\delta_{ij}} E(U). \tag{1}$$

- Nonlinear eigenvalue problem:

$$\begin{cases} \left(-\frac{1}{2}\Delta + V_{eff}(\rho)\right) u_i = \lambda_i u_i \quad \text{in } \mathbb{R}^3, \\ \int_{\mathbb{R}^3} u_iu_j = \delta_{ij}, \quad i,j = 1,2,\cdots,N, \end{cases} \tag{2}$$

where

$$V_{eff}(\rho) = V_{ne}(\rho) + V_H(\rho) + V_{xc}(\rho), \;\; V_{xc}(\rho) = \frac{\delta E_{xc}}{\delta \rho}$$

Kohn-Sham DFT Model

$$\left\{ \begin{array}{rcl} \left(-\frac{1}{2}\Delta + V_{eff}(\rho)\right) u_i & = & \lambda_i u_i \quad \text{in } \mathbb{R}^3, \\ \int_\Omega u_i u_j & = & \delta_{ij}, \quad i,j = 1,2,\cdots,N. \end{array} \right. \tag{3}$$

1. Give initial input charge density ρ_{in}.
2. Compute the effective potential $V_{\text{eff}}(\rho_{\text{in}})$.
3. Find $(\lambda_i, u_i) \in \mathbb{R} \times \mathrm{H}_0^1(\mathbb{R}^3)$ satisfying

$$\left\{ \begin{array}{rcll} \left(-\frac{1}{2}\Delta + V_{eff}(\rho_{\text{in}})\right) u_i & = & \lambda_i u_i & \text{in } \mathbb{R}^3, \\ \int_\Omega u_i u_j & = & \delta_{ij}, & i,j = 1, 2, \cdots, N. \end{array} \right. \tag{4}$$

4. Compute the new output charge density ρ_{out}.
5. Convergence check: if not converged, use some density mixing method to get the new input charge density ρ_{in}, goto step 2; else, stop.

By the SCF iteration, the central computation in solving such nonlinear eigenvalue problems is the repeated solution of algebraic eigenvalue problem

$$Au = \lambda Bu.$$

- Convergence of the SCF iteration
- Solution of the large scale algebraic eigenvalue problem

- Convergence of the SCF iteration
  - Z. Bai, R.-C. Li, and D. Lu (2020)
  - Y. Cai, L.-H. Zhang, Z. Bai, and R.-C. Li (2018)
  - E. Cances (2000, 2001)
  - E. Cances and K. Pernal (2008)
  - L. Lin and C. Yang (2013)
  - X. Liu, Z. Wen, X. Wang, M. Ulbrich, and Y. Yuan (2015)
  - T. Rohwedder and R. Schneider (2011)
  - C. Yang, W. Gao, and J. Meza (2009)
  - . . .

- Solution of large scale eigenvalue problems

- Convergence of SCF iteration
  - Z. Bai, R.-C. Li, and D. Lu (2020)
  - Y. Cai, L.-H. Zhang, Z. Bai, and R.-C. Li (2018)
  - E. Cances (2000, 2001)
  - E. Cances and K. Pernal (2008)
  - L. Lin and C. Yang (2013)
  - X. Liu, Z. Wen, X. Wang, M. Ulbrich, and Y. Yuan (2015)
  - T. Rohwedder and R. Schneider (2011)
  - C. Yang, W. Gao, and J. Meza (2009)
  - ...

  The performance of SCF iteration is unpredictable

  The theoretical analysis is very challenging
- Solution of large scale eigenvalue problems

  Need orthogonalization

- X. Dai, Z. Liu, X. Zhang and A. Zhou (2021)
- X. Dai, L. Zhang and A. Zhou (2019)
- X. Dai, Z. Liu, L. Zhang and A. Zhou (2017)
- B. Gao, X. Liu, X. Chen and Y. Yuan (2018)
- R. Schneider, T. Rohwedder, A. Neelov and J. Blauert (2009)
- M. Ulbrich, Z. Wen, C. Yang, D. Klöckner and Z. Lu (2015)
- C. Yang, J.C. Meza and L. Wang (2007)
- X. Zhang, J. Zhu, Z. Wen and A. Zhou (2014)
- ...

Need orthogonalization

- X. Dai, Z. Liu, X. Zhang and A. Zhou (2021)
- X. Dai, L. Zhang and A. Zhou (2019)
- X. Dai, Z. Liu, L. Zhang and A. Zhou (2017)
- B. Gao, X. Liu, X. Chen and Y. Yuan (2018)
- R. Schneider, T. Rohwedder, A. Neelov and J. Blauert (2009)
- M. Ulbrich, Z. Wen, C. Yang, D. Klöckner and Z. Lu (2015)
- C. Yang, J.C. Meza and L. Wang (2007)
- X. Zhang, J. Zhu, Z. Wen and A. Zhou (2014)
- $\cdots$

Need orthogonalization

- We propose
  - An extended gradient flow model

    Time-dependent problem, preserve orthogonality
  - A class of orthogonality preserving iteration schemes

    Reliability: approximations converge to some ground states
- References:

  - X. Dai, Q. Wang, and A. Zhou, Multiscale Model. Simul., 18(2020), 1621-1663.
  - X. Dai, L. Zhang, and A. Zhou, arXiv:2111.02779, 2021.

- We propose
  - An extended gradient flow model

    Time-dependent problem, preserve orthogonality
  - A class of orthogonality preserving iteration schemes

    Reliability: approximations converge to some ground states
- References:

  - X. Dai, Q. Wang, and A. Zhou, Multiscale Model. Simul., 18(2020), 1621-1663.
  - X. Dai, L. Zhang, and A. Zhou, arXiv:2111.02779, 2021.

- We propose
  - An extended gradient flow model

    Time-dependent problem, preserve orthogonality
  - A class of orthogonality preserving iteration schemes

    Reliability: approximations converge to some ground states
- References:
  - X. Dai, Q. Wang, and A. Zhou, Multiscale Model. Simul., 18(2020), 1621-1663.
  - X. Dai, L. Zhang, and A. Zhou, arXiv:2111.02779, 2021.

Outline

- Inner product matrix

$$\langle U^\top V\rangle = \left((u_i, v_j)_{\mathrm{L}^2(\mathbb{R}^3)}\right)_{i,j=1}^N \in \mathbb{R}^{N\times N}$$

$$\langle \mathcal{F}, U\rangle = \left(\langle \mathcal{F}_i, u_j\rangle\right)_{i,j=1}^N \in \mathbb{R}^{N\times N}$$

where

$$U = (u_1, u_2, \dots, u_N) \in \left(\mathrm{H}^1(\mathbb{R}^3)\right)^N$$

$$V = (v_1, v_2, \dots, v_N) \in \left(\mathrm{H}^1(\mathbb{R}^3)\right)^N$$

$$\mathcal{F} = (\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_N) \in \left((\mathrm{H}^1(\mathbb{R}^3))^N\right)' = \left(\mathrm{H}^{-1}(\mathbb{R}^3)\right)^N$$

- Norm for $U \in (\mathrm{H}^1(\mathbb{R}^3))^N$

$$|||U||| = \mathrm{tr}(\langle U^T U\rangle)$$

- Stiefel manifold

$$\mathcal{M}^N = \left\{U \in \left(\mathrm{H}^1(\mathbb{R}^3)\right)^N \;:\; \langle U^\top U\rangle = I_N\right\}$$

- Distance: for $U, \hat{U} \in \left(\mathrm{H}^1(\mathbb{R}^3)\right)^N$,

$$\mathrm{dist}(U, \hat{U}) = \sum_{j=1}^{N} ||u_j - \hat{u}_j||_{\mathrm{H}^1(\mathbb{R}^3)},$$

- Closed δ-neighborhood of $U \in (\mathrm{H}^1(\mathbb{R}^3))^N$

$$B(U, \delta) = \{\hat{U} \in \left(\mathrm{H}^1(\mathbb{R}^3)\right)^N : \mathrm{dist}(U, \hat{U}) \leq \delta\}$$

$$E(U) = E(UP), \ \forall \ P \in \mathcal{O}^{N\times N}$$

- Grassmann manifold: the quotient of the Stiefel manifold

$$\mathcal{G}^N = \mathcal{M}^N / \sim .$$

$U \sim W \iff \exists \ P \in \mathcal{O}^{N\times N} \ s.t. \ U = WP$

- Equivalence class $[U] = \{UP : P \in \mathcal{O}^{N\times N}\}$
- Tangent space on the Grassmann manifold

$$\mathcal{T}_{[U]}\mathcal{G}^N = \{W \in V^N | W^T U = 0 \in \mathbb{R}^{N\times N}\}$$

- Distance on the Grassmann manifold $\mathcal{G}^N$

$$\text{dist}([U], [\hat{U}]) = \min_{P\in\mathcal{O}^{N\times N}} \text{dist}(U, \hat{U}P),$$

- Closed δ-neighborhood of $[U]$ on $\mathcal{G}^N$

$$B([U], \delta) = \{[\hat{U}] \in \mathcal{G}^N : \hat{U} \in \left(\mathrm{H}^1(\mathbb{R}^3)\right)^N \bigcap \mathcal{M}^N, \text{dist}([U], [\hat{U}]) \le \delta\}$$

- Hamilton operator $\mathcal{H}(\rho) = -\frac{1}{2}\Delta + V_{eff}(\rho)$
- Gradient on $(\mathrm{H}^1(\mathbb{R}^3))^N$

$$\nabla E(U) = \mathcal{H}(\rho)U, \quad U \in (\mathrm{H}^1(\mathbb{R}^3))^N$$

- Gradient on $\mathcal{M}^N$

$$\nabla_G E(U) = \nabla E(U) - U\langle U^T \nabla E(U)\rangle, \ U \in \mathcal{M}^N$$

- Extended gradient on $(\mathrm{H}^1(\mathbb{R}^3))^N$

$$\nabla_G E(U) = \nabla E(U)\langle U^T U\rangle - U\langle U^T \nabla E(U)\rangle, \ U \in (H^1(\mathbb{R}^3))^N$$

- Extended gradient operator on $(\mathrm{H}^1(\mathbb{R}^3))^N$

$$\mathcal{A}_U U = \nabla_G E(U), \ U \in (H^1(\mathbb{R}^3))^N$$

- Hessian on $\mathcal{G}^N$

$$\begin{aligned}\mathrm{Hess}_G E(U)[V, W] &= \mathrm{tr}\big(\langle V^\top \nabla^2 E(U) W\rangle\big) \\ &- \mathrm{tr}\big(\langle V^\top W\rangle\langle U^\top \nabla E(U)\rangle\big), \forall V, W \in \mathcal{T}_{[U]}\mathcal{G}^N\end{aligned}$$

- Extended gradient flow model: Find $U(t) \in (\mathrm{H}^1(\mathbb{R}^3))^N$, such that

$$\begin{cases} \dfrac{\mathrm{d}U(t)}{\mathrm{d}t} = -\nabla_G E(U(t)), \\ U(0) = U_0 \in (\mathrm{H}^1(\mathbb{R}^3))^N, \end{cases} \tag{5}$$

where

$$\nabla_G E(U) = \nabla E(U)\langle U^T U\rangle - U\langle U^T \nabla E(U)\rangle, \ \ U \in (\mathrm{H}^1(\mathbb{R}^3))^N$$

Theorem(Dai, Wang, and Zhou (Multiscale Model. Simul., 2020))

If $U_0 \in \mathcal{M}^N$, that is, $\langle U_0^T U_0 \rangle = I_N$, then the solution of (3)

$$U(t) \in \mathcal{M}^N,$$

$$\frac{\mathrm{d}E\big(U(t)\big)}{\mathrm{d}t} = -\left|\left|\left|\nabla_G E\big(U(t)\big)\right|\right|\right|^2 \leqslant 0, \quad 0 < t < \infty$$

Theorem(Dai, Wang, and Zhou (Multiscale Model. Simul., 2020))

If $U_0 \in \mathcal{M}^N$, that is, $\langle U_0^T U_0 \rangle = I_N$, then

$$\liminf_{t\to\infty} |||\nabla_G E(U(t))||| = 0$$

Theorem(Dai, Wang, and Zhou (Multiscale Model. Simul., 2020))

Suppose the local minimizer $[U^*]$ is the unique critical point of $E(U)$ in $B([U^*], \delta_1)$. If $\nabla E(U)$ is continuous in a neighborhood of the local minimum $U^* \in \mathcal{M}^N$ and the initial value satisfies $E(U_0) \leq (E_0 + E(U^*))/2 \equiv E_1$, where

$$E_0 = \min\{E([\tilde{U}]) \mid [\tilde{U}] \in \overline{B([U^*], \delta_1) \backslash B([U^*], \delta_2)}\}.$$

with δ_2 being a fixed constant satisfying $\delta_2 \in (0, \delta_1]$, then

$$\lim_{t\to\infty} |||\nabla_G E(U(t))||| = 0$$

$$\lim_{t\to\infty} E(U(t)) = E(U^*)$$

$$\lim_{t\to\infty} \mathrm{dist}([U(t)], [U^*]) = 0$$

Theorem(Dai, Wang, and Zhou (Multiscale Model. Simul., 2020))

Suppose the local minimizer $[U^*]$ is the unique critical point of $E(U)$ in $B([U^*], \delta_1)$. If $\nabla E(U)$ is continuous in a neighborhood of the local minimum $[U^*]$, $E(U_0) \leq E_1$, and

$$\text{Hess}_G E(U)[D, D] \geqslant \sigma ||D||^2 \quad \forall [U] \in B([U^*], \delta_3), \ \forall D \in \mathcal{T}_{[U]}\mathcal{G}^N \bigcap (V_{N_g})^N$$

for $\delta_3 \in (0, \delta_1]$ and $\sigma > 0$, then there exists $\widehat{T} > 0$ such that

$$\left|\left|\left|\nabla_G E\big(U(t)\big)\right|\right|\right| \leqslant e^{-\sigma(t-\widehat{T})}$$

$$E\big(U(t)\big) - E(U^*) \leqslant \frac{1}{2\sigma} e^{-2\sigma(t-\widehat{T})}$$

hold for any $t \geqslant \widehat{T}$

Outline

- Finite dimensional space

$$V_{N_g} \subset \mathrm{H}^1(\mathbb{R}^3)$$

- $\mathcal{M}^N_{N_g} = \{U \in (V_{N_g})^N : U^T U = I_N\}$, $\mathcal{G}^N_{N_g} = \mathcal{M}^N_{N_g} / \sim$
- Extended gradient flow model: Find $U(t) \in (V_{N_g})^N$, such that

$$\begin{cases} \dfrac{\mathrm{d}U(t)}{\mathrm{d}t} = -\nabla_G E(U(t)), \\ U(0) = U_0 \in (V_{N_g})^N, \end{cases} \tag{6}$$

where

$$\nabla_G E(U) = \nabla E(U)\langle U^T U\rangle - U\langle U^T \nabla E(U)\rangle,\ U \in (\mathrm{H}^1(\mathbb{R}^3))^N$$

Choose a partition of the interval $[0, +\infty)$

$$0 = t_0 < t_1 < t_2 < \cdots < t_n < t_{n+1} < \cdots$$

Consider the midpoint scheme: find $\{U_n\}_n \subset (V_{N_g})^N$ such that

$$\begin{cases} \dfrac{U_{n+1} - U_n}{\Delta t_n} = -\nabla_G E(U_{n+1/2}), \ \ n = 0, 1, 2, \cdots \\ \langle U_0^T U_0 \rangle = I_N \end{cases} \tag{7}$$

where $U_0 \in (V_{N_g})^N, \Delta t_n = t_{n+1} - t_n$ and

$$U_{n+1/2} = \frac{U_{n+1} + U_n}{2}$$

Midpoint scheme

Algorithm 1: A midpoint scheme

Given $\epsilon > 0$, initial orbitals $U_0 \in \mathcal{M}^{N}_{N_g}$, calculate the gradient $\nabla_G E(U_0)$ and let $n = 0$;

while $|||\nabla_G E(U_n)||| > \epsilon$ **do**

  Choose a suitable $\Delta t_n > 0$ and let $t_{n+1} = t_n + \Delta t_n$;

  Solve

$$\frac{U_{n+1} - U_n}{\Delta t_n} = -\nabla_G E\left(\frac{U_{n+1} + U_n}{2}\right)$$

  to get U_{n+1};

  Let $n = n + 1$, calculate the gradient $\nabla_G E(U_n)$;

Theorem(Dai, Wang, and Zhou (Multiscale Model. Simul.))

If $\{U_n\}$ is obtained by Algorithm 1 and $U_0 \in \mathcal{M}^N_{N_g}$, then

$$U_n \in \mathcal{M}^N_{N_g}$$

Theorem(Dai, Wang, and Zhou (Multiscale Model. Simul.))

Suppose local minimizer $[U^*] \in \mathcal{G}_{N_g}^N$ is the unique critical point in $B\big([U^*], \delta_c\big)$ for some $\delta_c > 0$ and ∇E is Lipschitz continuous in $B\big([U^*], \delta_c\big)$. For sequence $\{U_n\}$ obtained by Algorithm 1, if $[U_0] \in B\big([U^*], \delta_c\big)$ and $\sup\{\Delta t_n : n \in \mathbb{N}\} \leqslant \delta_T$ for δ_T, then

$$\lim_{n\to\infty} |||\nabla_G E(U_n)||| = 0$$

$$\lim_{n\to\infty} E(U_n) = E(U^*)$$

$$\lim_{n\to\infty} \mathrm{dist}([U_n], [U^*]) = 0$$

Theorem: convergence rate

Suppose $[U^*] \in \mathcal{G}^N_{N_g}$ is a local minimizer, ∇E is Lipschitz continuous in $B\big([U^*], \delta_c\big)$ for some $\delta_c > 0$, and

$$\mathrm{Hess}_G E(U)[D, D] \geqslant \sigma |||D|||^2 \quad \forall [U] \in B([U^*], \delta_c),\ \forall D \in \mathcal{T}_{[U]}\mathcal{G}^N_{N_g}$$

for some $\sigma > 0$. If $[U_0]$ is in $B\big([U^*], \delta_c\big)$ and $\{U_n\}$ is obtained by Algorithm 1, then, there exist $N_0 > 1$ and $\delta_T > 0$, s.t.

$$|||\nabla_G E(U_n)||| \leqslant \Big(1 + \frac{L_1\tau}{2}\Big)\Big(\frac{4 + \tau^2 L_1^2 - 2\sigma\tau}{4 + \tau^2 L_1^2 + 2\sigma\tau}\Big)^{(n-N_0+1)/2} \cdot |||\nabla_G E(U_{N_0-1/2})|||,$$

$$E(U_n) - E(U^*) \leqslant \frac{(L+3)(4 + \tau^2 L_1^2 + 2\sigma\tau)}{8\sigma}\Big(\frac{4 + \tau^2 L_1^2 - 2\sigma\tau}{4 + \tau^2 L_1^2 + 2\sigma\tau}\Big)^{n-N_0+1} \cdot |||\nabla_G E(U_{N_0-1/2})|||^2$$

hold for $n \geq N_0$ and $\Delta t_n = \tau \leq \delta_T$

Choose a partition of the interval $[0, +\infty)$

$$0 = t_0 < t_1 < t_2 < \cdots < t_n < t_{n+1} < \cdots$$

Consider the Interpolation based scheme: find $\{U_n\}_{n\in\mathbb{N}} \in (V_{N_g})^N$ such that

$$\begin{cases} \tilde{U}(t) - U_n = -(t - t_n)\mathcal{A}_{U^{\mathrm{Aux}}(t)} \dfrac{U_n + \tilde{U}(t)}{2}, \ t \in [t_n, t_{n+1}), \\ \qquad U_{n+1} = \tilde{U}(t_{n+1}^-). \end{cases} \tag{8}$$

Here $U^{\mathrm{Aux}} : \mathbb{R} \to (V_{N_g})^N$ is a piecewise smooth auxiliary mapping which satisfies $U^{\mathrm{Aux}}(t_n) = U_n$ for all n.

Intepolation based scheme

Algorithm 2: A framework for interpolation based scheme

Given $\epsilon > 0$, initial orbitals $U_0 \in \mathcal{M}^N_{N_g}$, calculate the gradient $\nabla_G E(U_0)$ and let $n = 0$;

while $|||\nabla_G E(U_n)||| > \epsilon$ **do**

  Choose a suitable $\Delta t_n > 0$ and let $t_{n+1} = t_n + \Delta t_n$;

  Define $U^{\mathrm{Aux}}(t), t \in [t_n, t_{n+1})$ such that $U^{\mathrm{Aux}}(t_n) = U_n$;

  Update $U_{n+1} = \lim_{t \to t^-_{n+1}} \tilde{U}(t)$ with $\tilde{U}(t)$ satisfying

$$\tilde{U}(t) - U_n = -(t - t_n)\mathcal{A}_{U^{\mathrm{Aux}}(t)} \frac{U_n + \tilde{U}(t)}{2}, \quad t \in [t_n, t_{n+1});$$

  Let $n = n + 1$, calculate the gradient $\nabla_G E(U_n)$;

Theorem(Dai, Zhang, and Zhou (arXiv:2111.02779, 2021))

If $\{U_n\}_{n\in\mathbb{N}_0}$ is obtained by Algorithm 2 and $U_0 \in \mathcal{M}^N_{N_g}$, then

$$\{U_n\}_{n\in\mathbb{N}_0} \subset \mathcal{M}^N_{N_g}.$$

Theorem(Dai, Zhang, and Zhou (arXiv:2111.02779,2021))

If the sequence $\{U_n\}_{n\in\mathbb{N}}$ produced by Algorithm 2 satisfies

$$E(U_{n+1}) - E(U_n) \leq -\eta \Delta t_n |||\nabla_G E(U_n)|||^2, n \in \mathbb{N}_0 \tag{9}$$

with $\eta > 0$ being a given parameter. Then, there holds

$$\liminf_{n\to\infty} |||\nabla_G E(U_n)||| = 0.$$

In further, assume that the local minimizer $[U^*] \in \mathcal{G}^N_{N_g}$ is the unique critical point in $B\big([U^*], \delta_c\big)$ for some $\delta_c > 0$ and $[U_0] \in B\big([U^*], \delta_c\big)$, then

$$\lim_{n\to\infty} E(U_n) = E(U^*),$$

$$\lim_{n\to\infty} \mathrm{dist}([U_n], [U^*]) = 0.$$

Theorem(Dai, Zhang, and Zhou (arXiv:2111.02779, 2021))

Suppose $[U^*] \in \mathcal{G}^N_{N_g}$ is a local minimizer, ∇E is Lipschitz continuous in $B\big([U^*], \delta_1\big)$ for some $\delta_1 > 0$, and for all $[U] \in B([U^*], \delta_1)$, there hold

$$\mathrm{Hess}_G E(U)[D, D] \geq \sigma |||D|||^2, \forall\ D \in \mathcal{T}_{[U]}\mathcal{G}^N_{N_g},$$

for some $\sigma > 0$. If the sequence $\{U_n\}_{n\in\mathbb{N}_0}$ produced by Algorithm 2 with initial guess $[U_0] \in B([U^*], \delta_1) \subset \mathcal{G}^N_{N_g}$ satisfies

$$E(U_{n+1}) - E(U_n) \leq -\eta \Delta t_n |||\nabla_G E(U_n)|||^2, n \in \mathbb{N}_0 \tag{10}$$

with $\eta > 0$ being a given parameter, and there exists a $\tau > 0$ such that $\Delta t_n > \tau, \forall n \in \mathbb{N}_0$. Then there exists constants $\nu \in (0, 1)$, $C_1, C_2 > 0$, such that

$$\begin{aligned} E(U_n) - E(U^*) &\leq C_1 \nu^n \mathrm{dist}([U_0], [U^*])^2, \\ \mathrm{dist}([U_n], [U^*]) &\leq C_2(\sqrt{\nu})^n \mathrm{dist}([U_0], [U^*]). \end{aligned}$$

- Auxiliary mapping U^{Aux}
- Time steps

- Auxiliary mapping U^{Aux}
- Time steps

- Choice 1: $U^{\text{Aux}}(t) = (1-\alpha_n)U_n + \alpha_n \tilde{U}(t)$, $\alpha_n \in [0,1]$, $t \in [t_n, t_{n+1})$.
  - special case 1: $\alpha_n = 0$, $U^{\text{Aux}}(t) = U_n$, we have

$$U_{n+1} = \left(I_N + \frac{\Delta t}{2}\mathcal{A}_{U_n}\right)^{-1}\left(I_N - \frac{\Delta t}{2}\mathcal{A}_{U_n}\right)U_n, \tag{11}$$

    which is an approximation of Crank-Nicolson Scheme

$$U_{n+1} = \left(I_N + \frac{\Delta t}{2}\mathcal{A}_{U_{n+1}}\right)^{-1}\left(I_N - \frac{\Delta t}{2}\mathcal{A}_{U_n}\right)U_n. \tag{12}$$

  - special case 2: $\alpha_n = \frac{1}{2}$, $U^{\text{Aux}}(t) = \left(\tilde{U}(t) + U_n\right)/2$, we have

$$U_{n+1} = U_n + \Delta t \mathcal{A}_{U^{\text{Aux}}(t)} \frac{U_n + U_{n+1}}{2} \tag{13}$$

    which is exactly the midpoint scheme.

- Choice 2: $U^{\text{Aux}}(t) = U^m_{n+1/2}(t),\ \ t \in [t_n, t_{n+1}), \forall m \in \mathbb{N}_0$, where

$$U^m_{n+1/2}(t) = \left(I + \frac{t - t_n}{2}\mathcal{A}_{U^{m-1}_{n+1/2}(\Delta t_n)}\right)^{-1} U_n,\ \ m = 1, 2, \dots,$$

  and $U^0_{n+1/2} = U_n$.
- Choice 3: Let

$$U^{\text{Aux}}(t) = U_n - m_n(t - t_n)\nabla_G E(U_n),\ \ t \in [t_n, t_{n+1}),$$

  or

$$U^{\text{Aux}}(t) = 2(I + m_n(t - t_n)\mathcal{A}_{U_n})^{-1} U_n - U_n,\ \ t \in [t_n, t_{n+1})$$

  where m_n can be arbitrary real number.

- Auxiliary mapping U^{Aux}
- Time steps

Theorem(Dai, Wang, and Zhou (Multiscale Model. Simul., 2020))

For the midpoint scheme, if the initial guess $[U_0] \in B([U^*], \delta) \subset \mathcal{G}_{N_g}^N$, then there exists an upper bound δT such that for $\Delta t_n \in [0, \delta T]$, there holds

$$E(U_{n+1}) - E(U_n) \leq -\frac{1}{4N} \Delta t_n |||\nabla_G E(U_n)|||^2, n \in \mathbb{N}_0. \tag{14}$$

Remark

This theorem tells us that there indeed exists some time step which satisfy our assumption.

Outline

- Software platform

  PHG, RealSPACES

- Hardware platform

  LSSC-IV

- Kohn-Sham DFT

  $E_{xc} = E_{xc}^{LDA81} \oplus$ full potential

- Finite element discretization
- Midpoint scheme

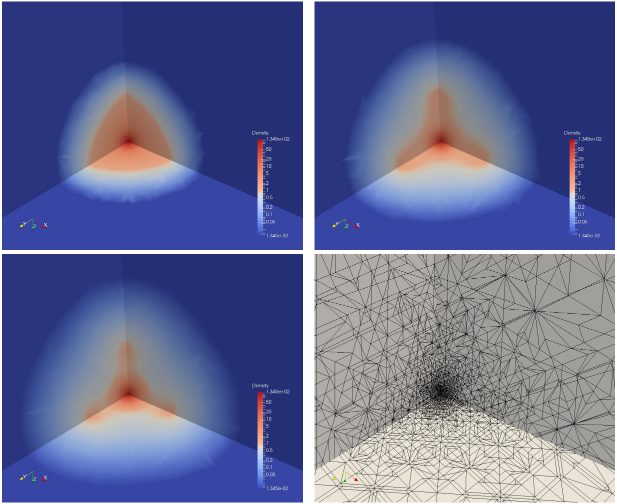
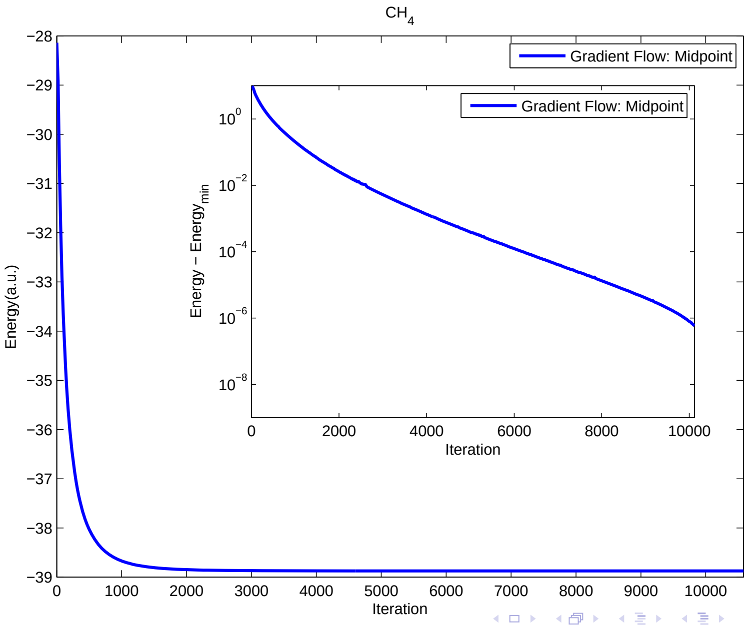
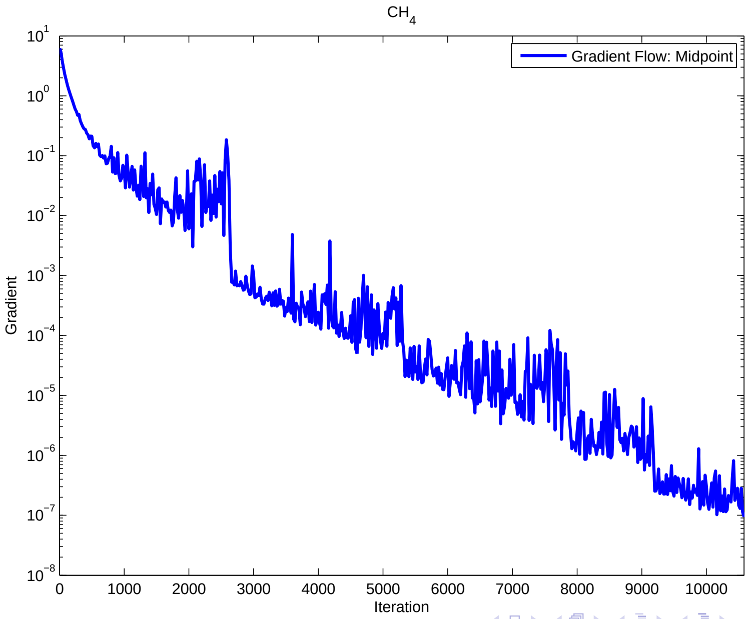


Figure 1: CH_4: Output density(Row 1 Column 1: 0th iteration; Row 1 Column 2: 640th iteration; Row 2 Column 1: 10580th iteration;) and input grid(Row 2 Column 2)

Methane(CH_4): energy



Methane(CH_4):gradient



Ethyne(C_2H_2)

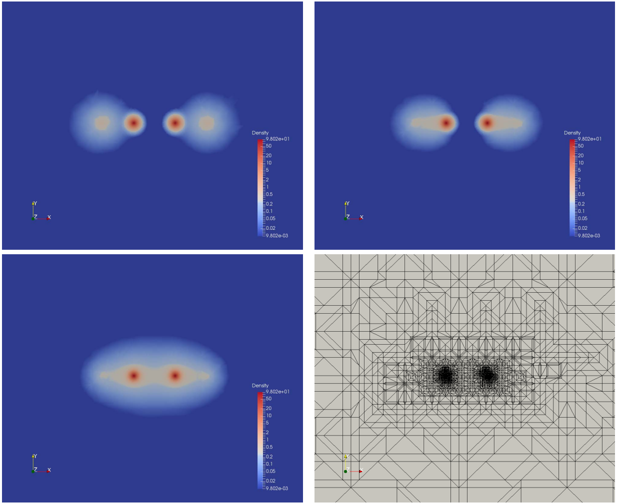


Figure 4: C_2H_2: Output density(Row 1 Column 1: 0th iteration; Row 1 Column 2: 520th iteration; Row 2 Column 1: 15160th iteration;) and input grid(Row 2 Column 2)

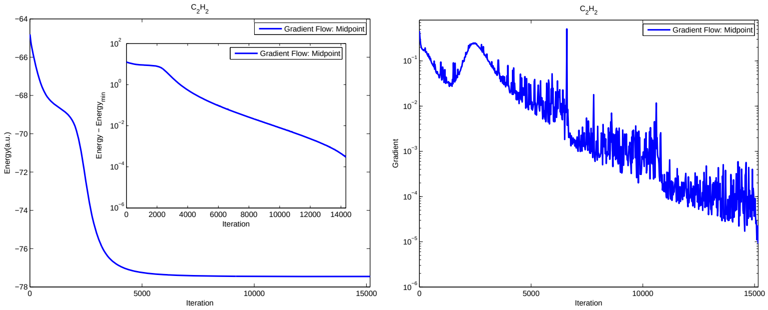


Figure 5: Convergence curves for energy(left) and gradient(right) for C_2H_2

Benzene(C_6H_6)

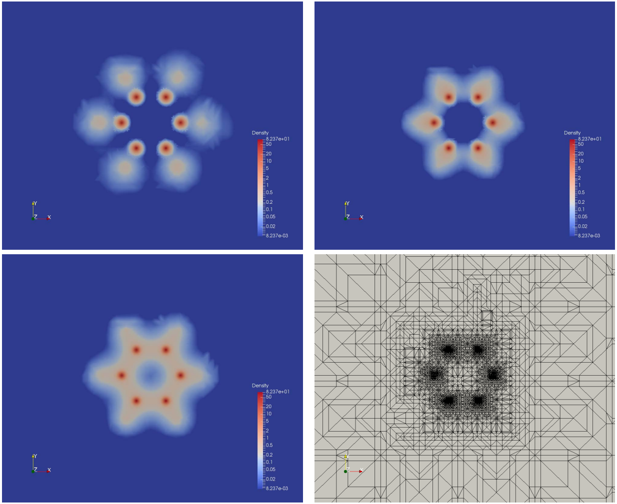


Figure 6: C_6H_6: Output density(Row 1 Column 1: 0th iteration; Row 1 Column 2: 300th iteration; Row 2 Column 1: 7460th iteration;) and input grid(Row 2 Column 2)

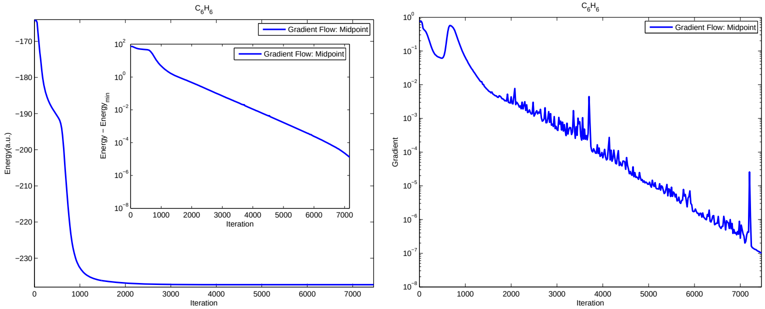


Figure 7: Convergence curves for energy(left) and gradient(right) for C_6H_6

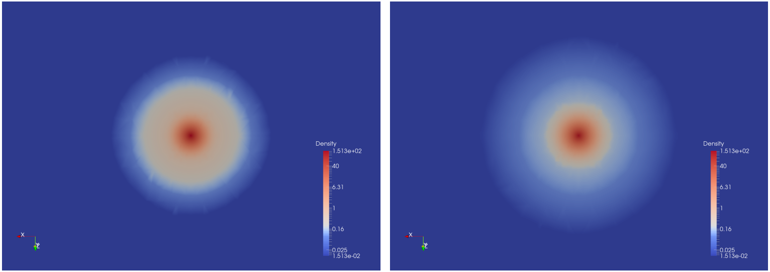


Figure 8: Carbon atom: density obtained at the 0th iteration; density obtained at the 9960th iteration

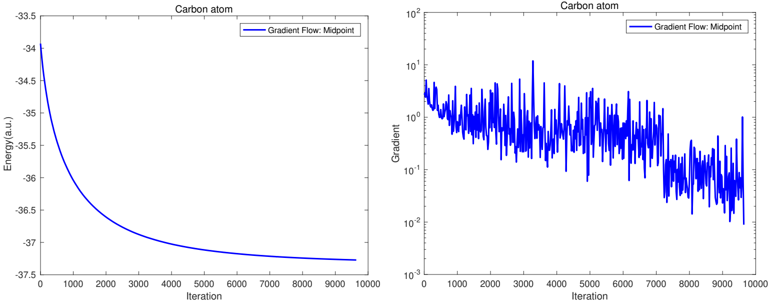


Figure 9: Convergence curves for energy (left) and gradient (right) for carbon atom

Outline

- Propose an extended gradient flow model
- Prove that the new model preserves orthogonality and the flow evolves to the ground state
- Propose and analyze a class of orthogonality preserving schemes for approximating the Kohn-Sham orbitals
- Prove the convergence and derive the local convergence rate of the framework under some mild and reasonable assumptions

References:

- X. Dai, Q. Wang, and A. Zhou, Multiscale Model. Simul., 18(2020), 1621-1663.
- X. Dai, L. Zhang, and A. Zhou, arXiv:2111.02779, 2021.

- Propose an extended gradient flow model
- Prove that the new model preserves orthogonality and the flow evolves to the ground state
- Propose and analyze a class of orthogonality preserving schemes for approximating the Kohn-Sham orbitals
- Prove the convergence and derive the local convergence rate of the framework under some mild and reasonable assumptions

References:

- X. Dai, Q. Wang, and A. Zhou, Multiscale Model. Simul., 18(2020), 1621-1663.
- X. Dai, L. Zhang, and A. Zhou, arXiv:2111.02779, 2021.

Ongoing and future work

Focusing on the extended gradient flow based model, we will study

- Some other orthogolality preserving schemes
- Some more time step choice and analysis
- More typical applications

Thanks !