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

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
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
Outline

1 Motivation

2 An Extended gradient flow model

3 Some orthogonality preserving iteration schemes

4 Numerical experiments

5 Concluding remarks
Kohn-Sham density functional theory (DFT)

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

\[
E(\{u_i\}) = \frac{1}{2} \sum_{i=1}^{N} \int_{\mathbb{R}^3} |\nabla u_i(x)|^2 \, dx + \int_{\mathbb{R}^3} V_{ne}(x) \rho(x) \, dx
\]

\[
+ \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x)\rho(y)}{|x-y|} \, dx \, dy + E_{xc}(\rho(x)).
\]

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

\[
\int_{\mathbb{R}^3} u_i u_j = \delta_{ij}, \quad 1 \leq i, j \leq N
\]

- \( \rho = \sum_{i=1}^{N} |u_i|^2 \)

- \( V_{ne}(r) = - \sum_{q=1}^{M} \frac{Z_q}{|r-R_q|} \)

- \( E_{xc}(\rho) \) 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:

\[ \min \limits_{U=(u_1,\ldots,u_N)\in(H^1(\mathbb{R}^3))^N,\int_{\mathbb{R}^3} u_i u_j=\delta_{ij}} E(U). \]  

- Nonlinear eigenvalue problem:

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

where

\[ V_{\text{eff}}(\rho) = V_{ne}(\rho) + V_H(\rho) + V_{xc}(\rho), \quad V_{xc}(\rho) = \frac{\delta E_{xc}}{\delta \rho}. \]
Nonlinear eigenvalue problem

Kohn-Sham DFT Model

\[
\begin{align*}
(\frac{-1}{2} \Delta + V_{\text{eff}}(\rho)) u_i &= \lambda_i u_i \quad \text{in } \mathbb{R}^3, \\
\int_{\Omega} u_i u_j &= \delta_{ij}, \quad i, j = 1, 2, \ldots, N.
\end{align*}
\]
Self-consistent field (SCF) iterations

1. Give initial input charge density $\rho_{in}$.
2. Compute the effective potential $V_{eff}(\rho_{in})$.
3. Find $(\lambda_i, u_i) \in \mathbb{R} \times H_0^1(\mathbb{R}^3)$ satisfying
   \begin{align*}
   \left\{ \begin{array}{l}
   \left( -\frac{1}{2} \Delta + V_{eff}(\rho_{in}) \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, \ldots, N.
   \end{array} \right.
   \end{align*}

4. Compute the new output charge density $\rho_{out}$.
5. Convergence check: if not converged, use some density mixing method to get the new input charge density $\rho_{in}$, 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
Nonlinear eigenvalue problem

- **Convergence of the SCF iteration**
  - Z. Bai, R.-C. Li, and D. Lu (2020)
  - E. Cances (2000, 2001)
  - 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)
- E. Cances (2000, 2001)
- 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
Constrained optimization problem

- ...
Constrained optimization problem

- ... 

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:

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:

A new model

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:

Outline

1. Motivation

2. An Extended gradient flow model

3. Some orthogonality preserving iteration schemes

4. Numerical experiments

5. Concluding remarks
Basic notation

- **Inner product matrix**

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

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

where

\[
U = (u_1, u_2, \ldots, u_N) \in \left( H^1(\mathbb{R}^3) \right)^N
\]

\[
V = (v_1, v_2, \ldots, v_N) \in \left( H^1(\mathbb{R}^3) \right)^N
\]

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

- **Norm** for \( U \in \left( H^1(\mathbb{R}^3) \right)^N \)

\[
|||U||| = \text{tr}(\langle U^\top U \rangle)
\]

- **Stiefel manifold**

\[
\mathcal{M}^N = \{ U \in \left( H^1(\mathbb{R}^3) \right)^N : \langle U^\top U \rangle = I_N \}\]
Basic notation

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

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

- **Closed \( \delta \)-neighborhood** of \( U \in (H^1(\mathbb{R}^3))^N \)

\[
B(U, \delta) = \{ \hat{U} \in (H^1(\mathbb{R}^3))^N : \text{dist}(U, \hat{U}) \leq \delta \}.
\]
\[ E(U) = E(UP), \ \forall \ P \in O^{N \times N} \]

- **Grassmann manifold**: the quotient of the Stiefel manifold
  \[ \mathcal{G}^N = \mathcal{M}^N / \sim. \]

- \( U \sim W \iff \exists \ P \in O^{N \times N} \text{ s.t. } U = WP \)

- **Equivalence class** \([U] = \{ UP : P \in O^{N \times N} \}\)

- **Tangent space** on the Grassmann manifold
  \[ T_{[U]} \mathcal{G}^N = \{ W \in V^N | W^T U = 0 \in \mathbb{R}^{N \times N} \}\]

- **Distance on the Grassmann manifold** \( \mathcal{G}^N \)
  \[ \text{dist}([U], [\hat{U}]) = \min_{P \in O^{N \times N}} \text{dist}(U, \hat{UP}), \]

- **Closed \( \delta \)-neighborhood of \([U]\) on \( \mathcal{G}^N \)
  \[ B([U], \delta) = \{ [\hat{U}] \in \mathcal{G}^N : \hat{U} \in (H^1(\mathbb{R}^3))^N \cap \mathcal{M}^N, \text{dist}([U], [\hat{U}]) \leq \delta \} \]
Gradient and extend gradient

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

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

- **Gradient on** \(\mathcal{M}^N\)

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

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

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

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

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

- **Hessian on** \(\mathcal{G}^N\)

\[
\text{Hess}_G E(U)[V, W] = \text{tr}(\langle V^T \nabla^2 E(U) W \rangle) - \text{tr}(\langle V^T W \rangle \langle U^T \nabla E(U) \rangle), \quad \forall V, W \in T_U \mathcal{G}^N
\]
Extended gradient flow model

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

$$\begin{cases}
\frac{dU(t)}{dt} = -\nabla_{gE}(U(t)), \\
U(0) = U_0 \in (H^1(\mathbb{R}^3))^N,
\end{cases}$$

(5)

where

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

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

$$\frac{dE(U(t))}{dt} = -\left\| \nabla_G E(U(t)) \right\|^2 \leq 0, \quad 0 < t < \infty$$
Asymptotical behavior

**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

$$\liminf_{t \to \infty} ||\nabla G E(U(t))|| = 0$$
Asymptotical behavior

**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

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

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

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

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

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

\[
\text{Hess}_G E(U)[D, D] \geq \sigma \| D \|^2 \quad \forall [U] \in B([U^*], \delta_3), \quad \forall D \in \mathcal{T}_U \mathcal{G}^N \bigcap (\mathcal{V}_N^g)^N
\]

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

\[
\left\| \nabla_G E(U(t)) \right\| \leq e^{-\sigma(t-\hat{T})}
\]

\[
E(U(t)) - E(U^*) \leq \frac{1}{2\sigma} e^{-2\sigma(t-\hat{T})}
\]

hold for any \( t \geq \hat{T} \).
Outline

1 Motivation

2 An Extended gradient flow model

3 Some orthogonality preserving iteration schemes

4 Numerical experiments

5 Concluding remarks
Extended gradient flow model in finite dimensional space

- **Finite dimensional space**

  \[ V_{N_g} \subset H^1(\mathbb{R}^3) \]

- \( \mathcal{M}_N^{N_g} = \{ U \in (V_{N_g})^N : U^T U = I_N \} \), \( G_{N_g}^{N} = \mathcal{M}_N^{N_g} / \sim \)

- Extended gradient flow model: Find \( U(t) \in (V_{N_g})^N \), such that

\[
\begin{cases}
\frac{dU(t)}{dt} = -\nabla G E(U(t)), \\
U(0) = U_0 \in (V_{N_g})^N,
\end{cases}
\]

where

\[
\nabla G E(U) = \nabla E(U) \langle U^T U \rangle - U \langle U^T \nabla E(U) \rangle, \quad U \in (H^1(\mathbb{R}^3))^N
\]
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 \(\{U_n\}_n \subset (V_{Ng})^N\) such that

\[
\begin{cases}
\frac{U_{n+1} - U_n}{\Delta t_n} = -\nabla G E(U_{n+1/2}), & n = 0, 1, 2, \cdots \\
\langle U_0^T U_0 \rangle = I_N
\end{cases}
\]

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

\[
U_{n+1/2} = \frac{U_{n+1} + U_n}{2}
\]
Algorithm 1: A midpoint scheme

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

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

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

4. Solve

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

   to get $U_{n+1}$;

5. 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 \( U_0 \in \mathcal{M}_{N_g}^N \), then

\[ U_n \in \mathcal{M}_{N_g}^N \]
Midpoint scheme: convergence

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

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

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

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

$$
\lim_{n \to \infty} \text{dist}([U_n], [U^*]) = 0
$$
Midpoint scheme: convergence rate

**Theorem: convergence rate**

Suppose $[U^*] \in G_{Ng}^N$ is a local minimizer, $\nabla E$ is Lipschitz continuous in $B([U^*], \delta_c)$ for some $\delta_c > 0$, and

$$\text{Hess}_G E(U)[D, D] \geq \sigma \|D\|^2 \quad \forall [U] \in B([U^*], \delta_c), \ \forall D \in T[U] G_{Ng}^N$$

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

$$\|\|\nabla G E(U_n)\|\| \leq \left(1 + \frac{L_1 \tau}{2}\right) \left(\frac{4 + \tau^2 L_1^2 - 2\sigma \tau}{4 + \tau^2 L_1^2 + 2\sigma \tau}\right)^{(n-N_0+1)/2} \cdot \|\|\nabla G E(U_{N_0-1/2})\|\|,$$

$$E(U_n) - E(U^*) \leq \frac{(L + 3)(4 + \tau^2 L_1^2 + 2\sigma \tau)}{8\sigma} \left(\frac{4 + \tau^2 L_1^2 - 2\sigma \tau}{4 + \tau^2 L_1^2 + 2\sigma \tau}\right)^{n-N_0+1} \cdot \|\|\nabla G E(U_{N_0-1/2})\|\|^2$$

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

$$\begin{cases} 
\tilde{U}(t) - U_n = -(t - t_n)A_{U_n} \frac{U_n + \tilde{U}(t)}{2}, \quad t \in [t_n, t_{n+1}), \\
U_{n+1} = \tilde{U}(t_{n+1}). 
\end{cases} \tag{8}$$

Here $U^{Aux} : \mathbb{R} \rightarrow (V_{N_g})^N$ is a piecewise smooth auxiliary mapping which satisfies $U^{Aux}(t_n) = U_n$ for all $n$. 
Algorithm 2: A framework for interpolation based scheme

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

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

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

4. Define $U_{\text{Aux}}(t), t \in [t_n, t_{n+1})$ such that $U_{\text{Aux}}(t_n) = U_n$;

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

\[
\tilde{U}(t) - U_n = -(t - t_n) A_{U_{\text{Aux}}(t)} \frac{U_n + \tilde{U}(t)}{2}, \quad t \in [t_n, t_{n+1});
\]

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

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

(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 $[U^*] \in \mathcal{G}_{\mathcal{N}_g}^N$ is the unique critical point in $B([U^*], \delta_c)$ for some $\delta_c > 0$ and $[U_0] \in B([U^*], \delta_c)$, then

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

$$\lim_{n \to \infty} \text{dist}([U_n], [U^*]) = 0.$$
Intepolation based scheme: convergence rate

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

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

\[
\text{Hess}_G E(U)[D, D] \geq \sigma |||D|||^2, \quad \forall \ D \in \mathcal{T}_U \mathcal{G}^N_{Ng},
\]

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

\[
E(U_{n+1}) - E(U_n) \leq -\eta \Delta t_n |||\nabla G E(U_n)|||^2, \quad n \in \mathbb{N}_0 \tag{10}
\]

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

\[
E(U_n) - E(U^*) \leq C_1 \nu^n \text{dist}(U_0, U^*)^2, \\
\text{dist}(U_n, U^*) \leq C_2 (\sqrt{\nu})^n \text{dist}(U_0, U^*).
\]
Specific schemes

- Auxiliary mapping $U^{\text{Aux}}$
- Time steps
Specific schemes

- 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

  \[
  U_{n+1} = (I_N + \frac{\Delta t}{2} A_{U_n})^{-1} (I_N - \frac{\Delta t}{2} A_{U_n}) U_n,
  \]

  which is an approximation of Crank-Nicolson Scheme

  \[
  U_{n+1} = (I_N + \frac{\Delta t}{2} A_{U_{n+1}})^{-1} (I_N - \frac{\Delta t}{2} A_{U_n}) U_n.
  \]  

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

  \[
  U_{n+1} = U_n + \Delta t A_{U^{\text{Aux}}(t)} \frac{U_n + U_{n+1}}{2}
  \]

  which is exactly the midpoint scheme.
Auxiliary mapping $U^{\text{Aux}}$

- **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

  $$U^m_{n+1/2}(t) = \left( I + \frac{t - t_n}{2} A_{U^m_{n+1/2}(\Delta t_n)} \right)^{-1} U_n, \; m = 1, 2, \ldots,$$

  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

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

  where $m_n$ can be arbitrary real number.
Specific schemes

- Auxiliary mapping $U^{\text{Aux}}$
- Time steps
A possible time step

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

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

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

**Remark**

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

1 Motivation

2 An Extended gradient flow model

3 Some orthogonality preserving iteration schemes

4 Numerical experiments

5 Concluding remarks
Calculation of ground state energy

- **Software platform**
  PHG, RealSPACES

- **Hardware platform**
  LSSC-IV

- **Kohn-Sham DFT**
  \[ E_{xc} = E_{xc}^{LDA81} \oplus \text{full potential} \]

- **Finite element discretization**

- **Midpoint scheme**
Methane ($\text{CH}_4$)

Figure 1: $\text{CH}_4$: 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

Figure 2: Convergence curves for energy of CH$_4$. The graph shows the energy (in a.u.) as a function of iteration, with a logarithmic scale for the energy axis. The energy decreases significantly with increasing iteration, indicating a rapid convergence to a minimum energy state.
Methane (CH$_4$): gradient
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\textsubscript{2}H\textsubscript{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$_6$H$_6$
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

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

- 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:

Conclusions

- 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:

Ongoing and future work

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

- Some other orthogonality preserving schemes
- Some more time step choice and analysis
- More typical applications
Thanks!