Elliptic Preconditioner for Accelerating the Self Consistent Field Iteration of Kohn-Sham Density Functional Theory

Lin Lin

Joint work with Chao Yang

Computational Research Division, Lawrence Berkeley National Lab
Electronic structure theory

Main goal: Given atomic positions $\{R_\alpha\}_{\alpha=1}^{M}$, Compute the ground state electron energy $E_e(\{R_\alpha\})$

Ground state electron wavefunction $\Psi_e(r_1, \ldots, r_N; \{R_\alpha\})$

$$\Psi_e = \frac{1}{2} \sum_{i=1}^{N} \Delta_i - \sum_{\alpha=1}^{M} \sum_{j=1}^{N} \frac{Z_\alpha}{|r_i - R_\alpha|} + \frac{1}{2} \sum_{i,j=1}^{N} \frac{1}{|r_i - r_j|} \Psi_e = E_e(\{R_\alpha\})\Psi_e$$

Curse of dimensionality

The fundamental laws necessary to the mathematical treatment of large parts of physics and the whole of chemistry are thus fully known, and the difficult lies only in the fact that application of these laws leads to equations that are too complex to be solved.

—P. Dirac, 1929
Pople diagram

Density functional theory

Table 1. Physical Review Articles with more than 1000 Citations Through June 2003

<table>
<thead>
<tr>
<th>Publication</th>
<th># cites</th>
<th>Av. age</th>
<th>Title</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR 140, A1133 (1965)</td>
<td>3227</td>
<td>26.7</td>
<td>Self-Consistent Equations Including Exchange and Correlation Effects</td>
<td>W. Kohn, L. J. Sham</td>
</tr>
<tr>
<td>PR 136, B864 (1964)</td>
<td>2460</td>
<td>28.7</td>
<td>Inhomogeneous Electron Gas</td>
<td>P. Hohenberg, W. Kohn</td>
</tr>
<tr>
<td>PR 108, 1175 (1957)</td>
<td>1364</td>
<td>20.2</td>
<td>Theory of Superconductivity</td>
<td>J. Bardeen, L. N. Cooper, J. R. Schrieffer</td>
</tr>
<tr>
<td>PRL 19, 1264 (1967)</td>
<td>1306</td>
<td>15.5</td>
<td>A Model of Leptons</td>
<td>S. Weinberg</td>
</tr>
<tr>
<td>PRB 12, 3060 (1975)</td>
<td>1259</td>
<td>18.4</td>
<td>Linear Methods in Band Theory</td>
<td>O. K. Anderson</td>
</tr>
<tr>
<td>PR 124, 1866 (1961)</td>
<td>1178</td>
<td>28.0</td>
<td>Effects of Configuration Interaction of Intensities and Phase Shifts</td>
<td>U. Fano</td>
</tr>
<tr>
<td>PRB 13, 5188 (1976)</td>
<td>1023</td>
<td>20.8</td>
<td>Special Points for Brillouin-Zone Integrations</td>
<td>H. J. Monkhorst, J. D. Pack</td>
</tr>
</tbody>
</table>


[S. Redner, Citation Statistics from 110 Years of Physical Review]
Density functional theory

<table>
<thead>
<tr>
<th>Publication</th>
<th># cites</th>
<th>Av. age</th>
<th>Title</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PR 140, A1133 (1965)</td>
<td>3227</td>
<td>26.7</td>
<td>Self-Consistent Equations Including Exchange and Correlation Effects</td>
<td>W. Kohn, L. J. Sham</td>
</tr>
<tr>
<td>PR 136, B864 (1964)</td>
<td>2460</td>
<td>28.7</td>
<td>Inhomogeneous Electron Gas</td>
<td>P. Hohenberg, W. Kohn</td>
</tr>
<tr>
<td>PR 108, 1175 (1957)</td>
<td>1364</td>
<td>20.2</td>
<td>Theory of Superconductivity</td>
<td>J. Bardeen, L. N. Cooper, J. R. Schrieffer</td>
</tr>
<tr>
<td>PRL 19, 1264 (1967)</td>
<td>1306</td>
<td>15.5</td>
<td>A Model of Leptons</td>
<td>S. Weinberg</td>
</tr>
<tr>
<td>PRB 12, 3060 (1975)</td>
<td>1259</td>
<td>18.4</td>
<td>Linear Methods in Band Theory</td>
<td>O. K. Anderson</td>
</tr>
<tr>
<td>PR 124, 1866 (1961)</td>
<td>1178</td>
<td>28.0</td>
<td>Effects of Configuration Interaction of Intensities and Phase Shifts</td>
<td>U. Fano</td>
</tr>
<tr>
<td>PRB 13, 5188 (1976)</td>
<td>1023</td>
<td>20.8</td>
<td>Special Points for Brillouin-Zone Integrations</td>
<td>H. J. Monkhorst, J. D. Pack</td>
</tr>
</tbody>
</table>

[Source: S. Redner, Citation Statistics from 110 Years of Physical Review]
Kohn-Sham density functional theory

- Electronic structure theory: Ground state energy $E_{tot}$.

\[
\min E_{tot}\left[\{\psi_i\}_{i=1}^N\right] = \frac{1}{2} \sum_{i=1}^{N} \int dx |\nabla \psi_i(x)|^2 + \int dx V_{ion}(x)\rho(x) + \frac{1}{2} \int dx \int dx' \frac{\rho(x)\rho(x')}{|x - x'|} + E_{xc}[\rho]
\]

\[
\rho(x) = \sum_{i=1}^{N} |\psi_i(x)|^2, \quad \int dx \psi_i^*(x)\psi_j(x) = \delta_{ij}, \quad x \in \mathbb{R}^3
\]

- In principle exact many body ground state energy, [Hohenberg-Kohn,1964], [Kohn-Sham, 1965]

- Best compromise between efficiency and accuracy. Most widely used electronic structure theory for condensed matter systems.
Strategies for solving KSDFT

• Nonlinear minimization problem
  • [Car-Parrinello, 1985]
  • [Teter-Payne-Allan, 1989]
  • [Payne-Teter-Allen-Arias-Joannopoulos, 1992]
  • [Yang-Meza-Wang, 2007]

• Euler-Lagrange equation
  • [Anderson, 1965]
  • [Pulay, 1980]
  • [Kerker, 1981]
  • [Kresse-Furthmüller, 1996]
  • [Kudin-Scuseria-Cances, 2002]
  • [Anglade-Gonze, 2008]
Kohn-Sham equation

\[ H[\rho] \psi_i(x) = \left( -\frac{1}{2} \Delta + V_{\text{ion}} + \int dx' \frac{\rho(x')}{|x - x'|} + V_{xc}[\rho] \right) \psi_i(x) = \varepsilon_i \psi_i(x) \]

\[ \rho(x) = \sum_{i=1}^{N} |\psi_i(x)|^2, \quad \int dx \, \psi_i^*(x) \psi_j(x) = \delta_{ij} \]

- \( \{\varepsilon_i\}_{i=1}^{N} \): Kohn-Sham energies
- \( \{\psi_i\}_{i=1}^{N} \): Kohn-Sham eigenfunctions (orbitals)
- Nonlinear eigenvalue problem
Self consistent field iteration

\[ H[\rho_{in}] \]

\[ \rho_{in} \quad \text{Update} \quad \rho_{out} \]

Charge mixing  Preconditioning
costly: \( O(N^3) \) scaling.
Example: Metal

Anderson: Generic mixing method

Kerker: Preconditioner designed for metals

This work: the same as Kerker.
Example: Insulator

**Anderson**: Generic mixing method

**Kerker**: Preconditioner designed for metals

**This work**: the same as Anderson.
Example: Metal + Insulator

Anderson: Generic mixing method

Kerker: Preconditioner designed for metals

This work: Best performance
Outline

- Fixed point iteration and its convergence
- Advanced mixing method and preconditioner
- New method: elliptic preconditioner
- Numerical results
Fixed point iteration for density or potential

\[ \mathcal{V}[^\rho] = V_{ion} + \int dx' \frac{\rho(x')}{|x - x'|} + V_{xc}[\rho] \]

Density fixed point

\[ \cdots \Rightarrow \rho_k \Rightarrow V_k \equiv \mathcal{V}[\rho_k] \Rightarrow \rho_{k+1} \Rightarrow V_{k+1} \Rightarrow \cdots \]

Potential fixed point

Potential fixed point: do not have the positivity constraint that \( \rho \geq 0 \).
Kohn-Sham map

\[ V_{out} = \mathcal{V}[F[V_{in}]] \]

- Fixed point iteration problem

\[ V^* = \mathcal{V}[F[V^*]] \]

- Kohn-Sham map: Fermi operator

\[ \rho(x) = F[V](x) = \frac{2}{1+e^{\beta(H[V]-\mu)}}(x, x) \]

- \( \beta = 1/k_B T \): inverse temperature
- \( \mu \): Chemical potential
Linearization

\[ V_{k+1} = \mathcal{V}[F[V_k]] \]

Linearization around \( V^* \). Error \( \delta V_k = V_k - V^* \)

\[ \delta V_{k+1} = V_{k+1} - V^* = \mathcal{V}[F[V_k]] - \mathcal{V}[F[V^*]] = \frac{\delta \mathcal{V}}{\delta \rho} \cdot \frac{\delta F}{\delta \mathcal{V}} \cdot \delta V_k \]

- \( \frac{\delta \mathcal{V}}{\delta \rho} (x, y) = \frac{1}{|x-y|} + \frac{\delta V_{xc}}{\delta \rho} (x, y) \equiv v_c(x, y) + K_{xc}(x, y) \)
- \( \frac{\delta F}{\delta \mathcal{V}} (x, y) \equiv \chi(x, y) \): independent particle polarizability matrix
Convergence of the linearized fixed point iteration

Linear algebra notation: treat $V$ as vectors and $\frac{\delta V}{\delta \rho}$ and $\chi$ as matrices

$$
\delta V_{k+1} = \frac{\delta V}{\delta \rho} \chi \delta V_k = \left( \frac{\delta V}{\delta \rho} \chi \right)^k \delta V_1
$$

Condition for contraction map

$$
\sigma \left( \frac{\delta V}{\delta \rho} \chi \right) < 1
$$

Not always satisfied.
Simple mixing

\[ V_{k+1} = V_k - \alpha (V_k - \mathcal{V}[F[V_k]]) \equiv V_k - \alpha r_k \]

\[ \delta V_{k+1} = \delta V_k - \alpha \left( I - \frac{\delta \mathcal{V}}{\delta \rho} \chi \right) \delta V_k \equiv (I - \alpha \varepsilon) \delta V_k \]

\[ = (I - \alpha \varepsilon)^k \delta V_1 \]

- \( \varepsilon = I - \frac{\delta \mathcal{V}}{\delta \rho} \chi \): Dielectric operator.

- Contraction map

\[ \sigma (I - \alpha \varepsilon) < 1 \]
Convergence of simple mixing

\[ |1 - \alpha \lambda(\varepsilon)| < 1 \Rightarrow 0 < \alpha < \frac{2}{\lambda(\varepsilon)} \]

- Simple mixing converge requires \( \lambda(\varepsilon) > 0 \)

- Guarantee convergence with a proper choice of \( \alpha \). Most stable among all mixing methods.

- Convergence can be very slow, and we can do better.

- The scaling of \( \lambda(\varepsilon) \) with respect to the system size.
Estimation of $\lambda(\varepsilon)$

- Random phase approximation (test-charge) for

$$\frac{\delta V}{\delta \rho}(x, y) = v_c(x, y) + K_{xc}(x, y) \approx v_c(x, y)$$

- Test function (long-wave mode, $\Omega$: Volume)

$$\phi_q(x) = \frac{1}{\sqrt{\Omega}} e^{iqx}, \quad \int dy \, v_c(x, y) \phi_q(y) = \frac{4\pi}{q^2} \phi_q(x)$$

- Estimate the bound of $\lambda$ by Ritz value

$$\int dx \, dy \, \phi_q^*(x) \varepsilon(x, y) \phi_q(y) \approx 1 - \frac{4\pi}{q^2} \int dx \, dy \, \phi_q^*(x) \chi(x, y) \phi_q(y)$$
Estimation of $\lambda(\varepsilon)$

- Adler-Wiser formula [Alder, 1962; Wiser, 1963]

$$\chi(x, y) = 2 \sum_{n=1}^{N} \sum_{m=N+1}^{\infty} \frac{\psi_n(x)\psi^*_m(x)\psi^*_n(y)\psi_m(y)}{\varepsilon_n - \varepsilon_m}$$

$$\varepsilon_n - \varepsilon_m \leq 0 \Rightarrow \chi \leq 0 \Rightarrow \lambda(\varepsilon) \geq 0$$

- $E_g = \varepsilon_{N+1} - \varepsilon_N$

$$\lim_{N \to \infty} E_g > 0: \text{ Insulator}$$

$$\lim_{N \to \infty} E_g = 0: \text{ Metal}$$
Estimation of $\lambda(\varepsilon)$

\[
\hat{\chi}(q, q) = \int dx \int dy \phi_q^*(x) \chi(x, y) \phi_q(y)
\]
\[
= 2 \sum_{n} \sum_{m} \int dx \phi_q^*(x) \psi_n(x) \psi_m^*(x) \int dy \phi_q(y) \psi_n^*(y) \psi_m(y) \frac{\varepsilon_n - \varepsilon_m}{\varepsilon_n - \varepsilon_m}
\]

- **Metal:** $\lim_{q \to 0} \hat{\chi}(q, q) = -\gamma < 0$

- **Insulator:** $\lim_{q \to 0} \hat{\chi}(q, q) \sim -\beta q^2 \to 0$ (isotropic material)

[Pick-Cohen-Martin, 1970; Goedecker 1999]
Convergence of simple mixing

- $\Omega = [0, L]^d$, the longest wave mode $q = \frac{2\pi}{L}$

- Metal

$$\lambda(\varepsilon) \sim 1 + \frac{4\pi \gamma}{q^2} = 1 + \frac{\gamma L^2}{\pi} \Rightarrow \text{“Charge sloshing”}$$

Simple mixing converge $\Rightarrow \alpha < \frac{2}{\lambda(\varepsilon)} \sim \frac{2\pi}{\gamma L^2} \to 0$

- Insulator

$$\lambda(\varepsilon) \sim 1 + 4\pi \beta$$

Simple mixing converge $\Rightarrow \alpha < \frac{2}{\lambda(\varepsilon)} \sim \frac{2}{1 + 4\pi \beta}$

Independent of system size
Example of charge sloshing

Calculated using KSSOLV [Yang-Meza-Lee-Wang, 2008]
Example of charge sloshing
Outline

• Fixed point iteration and its convergence

• Advanced mixing method and preconditioner

• New method: elliptic preconditioner

• Numerical results
Beyond simple mixing

Approximate $\varepsilon$

Advanced mixing
- Anderson mixing
- Pulay mixing (DIIS)
- Broyden mixing
- etc

Preconditioning
- HIJ
- Kerker
- Extrapolar
- TFvW
- etc
Advanced mixing methods

- Generic for solving $x = f(x)$
- Broyden mixing [Johnson, 1988]

- $r_k = V_k - V[F[V_k]]$
- $s_k = V_k - V_{k-1}, \ y_k = r_k - r_{k-1}$, find $C$ such that
  \[
  \min_C \frac{1}{2} \| C - C_{k-1} \|_F^2
  \]
  s.t. $S_k = CY_k$

  \[
  S_k = (s_k \ s_{k-1} \cdots s_{k-l}), \quad Y_k = (y_k \ y_{k-1} \cdots y_{k-l})
  \]
Advanced mixing methods

Solution $C = C_{k-1} + (S_k - C_{k-1}Y_k)Y_k^+, \ Y_k^+ = (Y_k^T Y_k)^{-1}Y_k^T$

Special case: $C_{k-1} = C_0$
In particular, choose $C_0 = \alpha I$, Anderson’s method [Anderson 1965]

$$V_{k+1} = V_k - C_k r_k$$

$$= V_k - S_k Y_k^+ r_k + \alpha(I - Y_k Y_k^+) r_k$$
Example of advanced mixing method
Preconditioning

\[ \delta V_{k+1} = (I - \alpha \varepsilon) \delta V_k \]

- Choose \( \alpha \) as a **preconditioner matrix**.
- Ideal preconditioner \( \alpha = \varepsilon^{-1} \).

\[ \tilde{r}_k = \alpha r_k \]

\[ V_{k+1} = V_k - \tilde{r}_k \]
More general understanding of preconditioner

- Preconditioner $\Leftrightarrow$ Choice of $C_0$ in Broyden’s method

$$V_{k+1} = V_k - C_k r_k$$

$$= V_k - S_k Y_k^+ r_k + C_0 (I - Y_k Y_k^+) r_k$$
Flowchart of preconditioner + Anderson mixing

Start from $V_k$

1. Compute the residue $r_k = V_k - \mathcal{V}[F[V_k]]$. If $e_k = \frac{\|r_k\|_2}{\|V_k\|_2} < \epsilon$, stop.

2. Evaluate the preconditioning direction
   \[ \tilde{r}_k = C_0 (I - Y_k Y_k^+) r_k \]

3. Update $V_{k+1} = V_k - S_k Y_k^+ r_k + \tilde{r}_k$. Go to step 1.
HIJ Approach

• Compute $\varepsilon$ by brute force [Ho-Ihm-Joannopoulos, 1982]

$$\chi(x, y) = 2 \sum_{n=1}^{N} \sum_{m=N+1}^{\infty} \frac{\psi_n(x)\psi_m^*(x)\psi_n^*(y)\psi_m(y)}{\varepsilon_n - \varepsilon_m}$$

• $O(N^4)$ scaling for constructing the preconditioner, compared with $O(N^3)$ scaling for solving each Kohn-Sham step.
Approximating $\varepsilon$ for metallic system

- Kerker preconditioner [Kerker, 1981], most successful and widely used for simple metal

- Motivation $\hat{\varepsilon}(q, q) \sim \frac{1}{q^2}$ as $q \to 0$, and should be damped.

- $\mathcal{F}(\tilde{r}_k)(q) = \frac{q^2}{q^2 + 4\pi \gamma} \mathcal{F}(r_k)(q)$, $\mathcal{F}$: Fourier transform
  $\tilde{r}_k = \mathcal{F}^{-1}(\mathcal{F}(\tilde{r}_k))$

- Effective for simple metals
- Fail for insulator, metal + insulator
Kerker mixing for insulator is not good
Outline

• Fixed point iteration and its convergence

• Advanced mixing method and preconditioner

• New method: elliptic preconditioner

• Numerical results
Another view of preconditioner

\[ \tilde{r}_k = \varepsilon^{-1} r_k = (1 - v_c \chi)^{-1} r_k = (v_c^{-1} - \chi)^{-1} v_c^{-1} r_k \]

Here \( v_c(x, y) = \frac{1}{|x-y|} \), and \( v_c^{-1} = -\frac{1}{4\pi} \Delta \)

Solve the equation

\[ (-\Delta - 4\pi \chi)\tilde{r}_k = -\Delta r_k \]

\( \chi \) is semi negative definite \( \Rightarrow -\Delta - 4\pi \chi \) is a semi positive definite operator
Kerker mixing revisit

- \( \mathcal{F}(\tilde{r}_k)(q) = \frac{q^2}{q^2 + 4\pi \gamma} \mathcal{F}(r_k)(q) \), \( \mathcal{F} \): Fourier transform
  \[ \tilde{r}_k = \mathcal{F}^{-1}(\mathcal{F}(\tilde{r}_k)) \]

- This is equivalent to solving
  \[ (-\Delta - 4\pi \chi)\tilde{r}_k = -\Delta r_k \]

by approximating \( \chi \) using a constant \(-\gamma\) (Thomas-Fermi screening length)
Advanced mixing only

- For insulator $\hat{\chi}(q, q) \sim -\beta q^2$. $\chi \sim \beta \Delta$

\[-(1 + 4\pi\beta)\Delta \tilde{r}_k = -\Delta r_k \Rightarrow \tilde{r}_k = \frac{1}{1 + 4\pi\beta} r_k\]

- Rescale by a constant factor $\frac{1}{1+4\pi\beta}$ related to the optimal simple mixing coefficient.
Unified method: elliptic preconditioner

[LL-Yang, arXiv:1206.2225]

\[ A\tilde{r}_k = (-\nabla \cdot (a(x)\nabla) + 4\pi b(x))\tilde{r}_k = -\Delta r_k \]

- \( a(x) = 1 + 4\pi\beta, \ b(x) = 0 \), simple insulator with optimal mixing coefficient
- \( a(x) = 1, \ b(x) = \gamma \), simple metal
- \( a(x), b(x) \) can be spatially dependent for metal + insulator.
- \( a(x) \geq 1, b(x) \geq 0 \ \Rightarrow \ A \) is an elliptic operator and semi-positive definite.

- Standard method to achieve \( O(N) \) scaling for elliptic preconditioner: Multigrid, FMM, H-matrix, HSS etc.
Flowchart of elliptic preconditioner + Anderson mixing

Start from $V_k$

1. Compute the residue $r_k = V_k - \mathcal{V}[F[V_k]]$. If $e_k = \frac{\|r_k\|_2}{\|V_k\|_2} < \epsilon$, stop.

2. Evaluate the preconditioning direction

\[
(-\nabla \cdot (a(x)\nabla) + 4\pi b(x))\tilde{r}_k = -\Delta(I - Y_k Y_k^+)r_k
\]

3. Update $V_{k+1} = V_k - S_k Y_k^+ r_k + \tilde{r}_k$. Go to step 1.
Outline

• Fixed point iteration and its convergence
• Advanced mixing method and preconditioner
• New method: elliptic preconditioner
• Numerical results
Reduced Hartree-Fock 1D model

- Reduced Hartree-Fock model [Solovej, 1991, Cances-Deleurence-Lewin, 2008]
  \[ H = -\frac{1}{2} \Delta + \int dy \ K(x, y) (\rho(y) + m(y)) \]

- Yukawa type kernel \( K(x, y) = \frac{2\pi}{\kappa \varepsilon_0} e^{-\kappa |x-y|} \),
  \[ -\Delta V + \kappa^2 V = \frac{4\pi}{\varepsilon_0} \delta(x) \]

- Pseudo-charge for \( V_{ion} \): \( m(x) = \sum_i m_i (x - R_i) \)
  \[ m_i (x - R_i) = -\frac{Z_i}{\sqrt{2\pi \sigma_i^2}} e^{-\frac{(x-R_i)^2}{2\sigma_i^2}} \]
Insulator

\[ \xi \]

\[ \rho(x) \]

- Occupied
- Unoccupied
Insulator

Charge mixing (Anderson)

Preconditioning (Kerker)

This work
Metal
Metal

Charge mixing (Anderson)

Preconditioning (Kerker)

This work
Metal + Insulator
Metal + Insulator: choice of $a(x)$ and $b(x)$
Metal+Insulator

Charge mixing (Anderson)

Preconditioning (Kerker)

This work

\[ ||r_k|| / ||V_k|| \]

iteration

Anderson

Kerker+Anderson

Elliptic+Anderson

\[ ||r_k|| / ||V_k|| \]

iteration

\[ ||r_k|| / ||V_k|| \]

iteration
3D Example

- Sodium (Na) bar with a large amount of vacuum (Metal+Vacuum)
- Nanowire, Nanotube, Surface system
- Solved by KSSOLV package [Yang-Meza-Lee-Wang, 2009]
Convergence

(a) Na 32 atoms

(b) Na 64 atoms

\[ \frac{\|r_k\|}{\|V_k\|} \]

iteration

\[ \frac{\|r_k\|}{\|V_k\|} \]

iteration
Conclusion

- Self consistent iteration: charge mixing and preconditioning
- Different behavior of metal and insulator. $\lambda(\varepsilon)$ is the key.

- Elliptic preconditioner:
  Unified treatment of metal and insulator.  
  Unified treatment of charge mixing, Kerker preconditioner.  
  Effective for metal, insulator, and large inhomogeneous metallic system with metal+insulator. 
  Simple to implement.
Acknowledgment

• Luis Alvarez fellowship supported by LBNL and DOE.

• Discussion with Roberto Car, Eric Cances, Weinan E, Weiguo Gao, Jianfeng Lu, Lin-Wang Wang and Lexing Ying