Adaptive Local Basis Set for Kohn-Sham Density Functional Theory

Lin Lin

Joint work with Weinan E, Jianfeng Lu and Lexing Ying

Computational Research Division, Lawrence Berkeley National Lab
Kohn-Sham density functional theory

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

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

- **Efficient:** Single particle theory
- **Accurate:** Exact ground state energy for exact \( V_{xc}[\rho] \), [Hohenberg-Kohn, 1964], [Kohn-Sham, 1965]
## Dilemma: Choice of basis (discretization)

<table>
<thead>
<tr>
<th>Basis</th>
<th>Example</th>
<th>DOF / atom</th>
<th>Convergence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform basis</td>
<td>Planewave</td>
<td>500~10000 (pseudopotential, even larger for full-potential)</td>
<td>Simple</td>
</tr>
<tr>
<td></td>
<td>Finite difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Finite element</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contracted basis</td>
<td>Gaussian orbitals</td>
<td>4~100</td>
<td>Complicated</td>
</tr>
<tr>
<td></td>
<td>Atomic orbitals</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Q: Combine the advantage of both?
Why so many basis functions are needed

- Large $|\nabla \rho|$ near the position of the nuclei

Large number of planewaves to resolve the small wavelength.

Electron density $\rho$
Along one [001] slice of a Si unit cell (using HGH pseudopotential)

Gradient $|\nabla \rho|$
Atomic orbitals

• **Observation:** the fast oscillation near the nuclei is relatively inert to the change of the environment: atomic orbitals.

• Atomic orbitals as basis functions:
  • Linear combination of atomic orbitals (LCAO)
  • Variational optimized atomic orbitals (SIESTA, AIMS etc.)
  • Gaussian type orbitals (QCHEM, CP2K etc.)

• Mixed (enriched) basis functions:
  • Augmented plane-wave (APW), Linear APW (LAPW), Full potential LAPW (FLAPW)
  • Enriched finite element [Sukumar and Pask, 2009]
Problems of the atomic orbitals

• Atomic orbitals do not directly reflect the environmental effect

• Atomic orbitals
  • Carefully chosen library of the basis functions.
  • Fine tuning of the parameters.
  • Overcompleteness and incompleteness, ill conditioned matrices.

• Mixed basis functions
  • Fine tuning
  • Relatively large number of basis functions per atom.
Local eigenfunctions as basis functions

\[ \rho_{in} \Rightarrow H_{eff}[\rho_{in}] \]

\[ H_{eff,Q_1}\tilde{\phi}_{1,j} = \varepsilon_{1,j}\tilde{\phi}_{1,j} \]

\[ \ldots \]

\[ H_{eff,Q_M}\tilde{\phi}_{M,j} = \varepsilon_{M,j}\tilde{\phi}_{M,j} \]

Construct a Hamiltonian

Small Eigenvalue problem

\[ \rho_{out} \]

SCF
Automatic dimension reduction

- Construct the local basis functions adaptively by solving a small part of the system.

- Local solve to obtain discontinuous basis function.
- Discontinuous Galerkin (DG) framework to patch basis functions together.
Adaptive local basis functions

- Element: $E_k$, Buffer: $Q_k$.
- Solving the linear eigenvalue problem (NO pseudo-hydrogen saturation etc.) inside the buffer region with periodic boundary conditions.

$$H_{\text{eff}, Q_k} \tilde{\varphi}_{k,j} = \varepsilon_{k,j} \tilde{\varphi}_{k,j}.$$
Adaptive local basis functions

- Restrict $\tilde{\varphi}_{k,j}$ from buffer to element $\varphi_{k,j}$
- Orthogonalize $\varphi_{k,j}$: adaptive local basis functions.

\[ H_{\text{eff},Q_k} \tilde{\varphi}_{k,j} = \varepsilon_{k,j} \tilde{\varphi}_{k,j}. \]

Red: Element $E_k$; Red+Blue: Buffer $Q_k$. 
Discontinuous Galerkin method

- Finite element method with **discontinuous** basis functions.
- **Inter-element continuity** (consistency): e.g. Penalty on the inter-element jump [Arnold, 1982]
Kohn-Sham energy functional

Fix the electron density $\rho$, and consider the linear problem (i.e. one step SCF iteration)

$$H_{\text{eff}} = -\frac{1}{2} \Delta + V_{\text{eff}} + \sum_{\ell} \gamma_{\ell} |b_{\ell}\rangle \langle b_{\ell}|$$

with orthonormality constraints $\langle \psi_i, \psi_j \rangle = \delta_{ij}$.

The effective one-body potential

$$V_{\text{eff}}[\rho](x) = V_{\text{ext}}(x) + \int \frac{\rho(y)}{|x - y|} \, dy + \epsilon_{xc}'[\rho(x)].$$
Kohn-Sham energy functional

The linear problem minimizes the following quadratic functional

\[ E_{\text{eff}}(\{\psi_i\}) = \frac{1}{2} \sum_{i=1}^{N} \int |\nabla \psi_i(x)|^2 \, dx + \int V_{\text{eff}}(x) \rho(x) \, dx \]

\[ + \sum_{\ell} \gamma_{\ell} \sum_{i=1}^{N} |\langle b_{\ell}, \psi_i \rangle|^2 , \]
Discontinuous Galerkin method

\[ E_{DG}(\{\psi_i\}) = \frac{1}{2} \sum_{i=1}^{N} \langle \nabla \psi_i, \nabla \psi_i \rangle_T + \langle V_{\text{eff}}, \rho \rangle_T + \sum_{\ell} \gamma_{\ell} \sum_{i=1}^{N} |\langle b_{\ell}, \psi_i \rangle_T|^2 \]

\[ - \sum_{i=1}^{N} \langle \{\nabla \psi_i\}, [\psi_i] \rangle_S + \frac{\alpha}{\hbar} \sum_{i=1}^{N} \langle [\psi_i]_S, [\psi_i]_S \rangle_S. \]

average and jump operators across surfaces \( S \):

\[ \{\{ q \} \} = \frac{1}{2}(q_1 + q_2) \quad \text{on } S. \]

\[ [u] = u_1 n_1 + u_2 n_2 \quad \text{on } S. \]

- Interior penalty method [Arnold, 1982]
Euler-Lagrange equation

Linear eigenvalue problem to obtain the coefficients of the adaptive local basis functions

\[
\sum_{k,j} \left( \frac{1}{2} \left\langle \nabla \varphi_{k', j'}, \nabla \varphi_{k, j} \right\rangle_T - \frac{1}{2} \left\langle \left\{ \varphi_{k', j'} \right\}, \left\{ \nabla \varphi_{k, j} \right\} \right\rangle_S \right. \\
- \frac{1}{2} \left\langle \left\{ \nabla \varphi_{k', j'} \right\}, \left\{ \varphi_{k, j} \right\} \right\rangle_S + \frac{\alpha}{\hbar} \left\langle \left\{ \varphi_{k', j'} \right\}, \left\{ \varphi_{k, j} \right\} \right\rangle_S + \left\langle \varphi_{k', j'}, V_{\text{eff}} \varphi_{k, j} \right\rangle_T \\
+ \sum_{\ell} \gamma_\ell \left\langle \varphi_{k', j'}, b_\ell \right\rangle_T \left\langle b_\ell, \varphi_{k, j} \right\rangle_T \right) c_{i; k, j} = \varepsilon_i \sum_{k, j} \left\langle \varphi_{k', j'}, \varphi_{k, j} \right\rangle c_{i; k, j}.
\]

- Kohn-Sham energies: $\varepsilon_i$.
- Kohn-Sham orbitals: $\psi_i(r) = \sum_{k \in T} \sum_{j} c_{i; k, j} \varphi_{k, j}(r)$
- Electron density: $\rho(r) = 2 \sum_{i=1}^{N/2} |\psi_i(r)|^2$. 
Flowchart of adaptive local basis functions

\[ \rho_{in} \Rightarrow H_{eff}[\rho_{in}] \]

\[ H_{eff,Q_1} \Phi_{1,j} = \varepsilon_{1,j} \Phi_{1,j} \]

\[ \ldots \]

\[ H_{eff,Q_M} \Phi_{M,j} = \varepsilon_{M,j} \Phi_{M,j} \]

Construct DG Hamiltonian

DG Eigenvalue problem

\[ \rho_{out} \]

SCF
Numerical examples

• Quantification of the error

$$\Delta E = \frac{|E_{GLB} - E_{DG}|}{N_{Atom}}$$

• LDA, Hartwigen-Goedecker-Hutter (HGH) pseudopotential.
• Legendre-Gauss-Lobatto (LGL) grid inside each element for numerical integration.
• Disordered system
Illustration of basis

• Disordered Na in 3D
Disordered Quasi-1D system

8 Na atoms

32 Si atoms
Disordered Quasi-1D system

![Graph 1](image1.jpg)

![Graph 2](image2.jpg)
Disordered Quasi-2D and 3D bulk system

Quasi 2D: 32 Na atoms

3D Bulk: 128 Na atoms
Parallel scalability

3D densely packed Na
16 basis per atom

<table>
<thead>
<tr>
<th>Atom#</th>
<th>Proc#</th>
<th>Global (s)</th>
<th>DG (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>64</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>432</td>
<td>216</td>
<td>88</td>
<td>14</td>
</tr>
</tbody>
</table>

- Dense Eigensolver
- Generate basis
- DG Overhead
- $\propto (\text{Atom#})^{1.64}$
Effect of the penalty parameter $\alpha$

**Graphs:**

- **Graph 1:**
  - Title: Error/Atom (au)
  - X-axis: $\alpha$
  - Y-axis: Error/Atom (au)
  - Data points for Na: $\alpha \approx 0.66$
  - Data points for Si: $\alpha \approx 0.58$

- **Graph 2:**
  - Title: Error/Atom (au)
  - X-axis: $\alpha$
  - Y-axis: Error/Atom (au)
  - Data points for Na: $\alpha \approx 0.66$
  - Data points for Si: $\alpha \approx 0.58$
Initial results on force

Disordered Si:
15 basis / atom
3D: Still not satisfactory

- Deterioration of the efficiency of ALB w.r.t. the dimensionality
**Element orbitals (EO)**

- The ALBs are already good basis functions (small error of the total energy per atom).
- Recover the continuous orbitals from discontinuous basis functions to reduce the DOF.
- Include basis functions from neighboring elements by linear combinations of ALBs.
- Each basis function is not strictly localized in the element, but centered around the element

[LL-Ying, PRB, 85, 235144 (2012)]
Requirement of EO

- ALBs are always **discontinuous** across the element boundaries, while qualitatively the EOs should be a **continuous** function since the atomic orbitals are continuous

- EOs should have **low energy**

- EOs should be **localized around its associated element** to avoid potential linear dependence
Step 1: Select candidate functions

- $I_k$ the index set of ALBs in $Q_k$
- $H_k = H(I_k, I_k)$
- Solve eigenvalue problems to get \textbf{candidate functions} for element orbitals
  $$H_k M_k = M_k \Delta_k$$

- In practice, the energy ($\Delta_k$) of the candidate functions should not exceed the largest energy of the ALB in $E_k$.

- Element: $E_k$
- Extended element: $Q_k$
Step 2: Localization for element orbitals

- Weight function $w_k(x)$

- Weight matrix $W_k(k', j'; k'', j'') = \langle u_{k', j'}, w_k u_{k'', j''} \rangle_T$

- Solve another eigenvalue problems to perform localization
  $$(M_k^t W_k M_k) L_k = L_k \Gamma_k$$

- Localization technique and weighting technique
  [Garcia-Cervera-Lu-Xuan-E, PRB (2009); Junquera-Paz-Sanchez-Portal-Artacho, PRB (2011)]

- Element: $E_k$

- Extended element: $Q_k$
Step 3: Generalized eigenvalue problem

- Coefficient matrix of element orbitals in terms of the adaptive local basis functions ($S_k$: selection matrix, $S_k(I_k, I_k) = I$ and zero otherwise)
  \[ C_k = S_k M_k L_k \]

- Global coefficient matrix
  \[ C = (C_1, \ldots, C_M) \]

- Generalized eigenvalue problem
  \[ (C^T H C) V = (C^T C) V \Lambda \]
Flowchart of element orbitals

\[ \rho_{\text{in}} \Rightarrow H_{\text{eff}} [\rho_{\text{in}}] \]

\[ H_{\text{eff},Q_1} \tilde{\phi}_{1,j} = \varepsilon_{1,j} \tilde{\phi}_{1,j} \]

\[ \ldots \]

\[ H_{\text{eff},Q_M} \tilde{\phi}_{M,j} = \varepsilon_{M,j} \tilde{\phi}_{M,j} \]

Construct a DG Hamiltonian H

Construct Element orbital and \( C^T HC \) and \( C^T C \)

Generalized eigenvalue problem

\[ \rho_{\text{out}} \]
3D Disordered Na system

- 432 Na atoms ($6 \times 6 \times 6$ b.c.c. cells + disorderedness in $\Gamma$ point only)
- Error is quantified by direct comparison of the total energy with ABINIT, which is a standard electronic structure software.
Shape of the element orbitals

- Fixed 27 Na atoms (yellow balls)

- First 9 orbitals (red: + part; blue: - part)

- s, p, d type orbitals but centered around elements but with the right multiplicity (1,3,5)

- "Holes" around atoms even for s-orbitals ⇒ Environmental effect
Lattice constant calculation

- Na: 3D 432 atoms, 4 EOs / atom
- Si: 3D 512 atoms, 10 EOs / atom
Graphene with a Si defect

- Cyan balls: C atoms; Gold balls: Si atoms
- 12 EOs / atom, Error of the total energy: 0.38 meV / atom
Graphene with many Si defects

Atomic configuration
Cyan ball: C; Gold ball: Si

Electron density

- 512 atoms. Dimension: 10.000 × 64.432 × 74.400 a.u.
- Total energy error 4.79 meV / atom
Computational cost of EO

- Total time per SCF
- LOBPCG for ALB
- Construct element orbitals
- Dense DG solve
- DG overhead
- Other
Conclusion

- Adaptive local basis functions (ALB); Element orbitals (EO)

- Automatic dimension reduction

- Atomic and environmental effect

- Accurate and small number of basis per atom.

- Local Hamiltonian matrix. Parallel computing.
Future work

• Change the dense matrix linear algebra to sparse matrix linear algebra to save memory and computational cost.

• Combined with the pole expansion and selected inversion algorithm (PEXSI) for low order scaling computation (at most $O(N^2)$ scaling) of KSDFT.

• Numerical integration is still expensive, which calls for more effective pseudopotential (USPP or PAW type) and intra-element parallelization for hard atoms

• Discontinuous or continuous basis functions?

• More accurate force: optimization?

Acknowledgment

Luis Alvarez fellowship supported by LBNL and DOE. SciDAC.

Thank you for your attention!