Massive plane-wave calculations in massive simulation cells

Chris-Kriton Skylaris

with Peter D. Haynes, Arash A. Mostofi and Mike C. Payne

Theory of Condensed Matter group
Cavendish Laboratory
Madingley Road
Cambridge CB3 0HE
ONETEP: Density-matrix linear-scaling DFT

Molecular orbitals (MOs)  \[ \rho(r, r') = \sum_n f_n \psi_n(r) \psi_n^*(r') \]

Non-orthogonal Generalised Wannier Functions (NGWFs)  \[ = \sum_{\alpha\beta} \phi_{\alpha}(r) K^{\alpha\beta} \phi_{\beta}^*(r') \]

- Optimise non-orthogonal localised functions \( \{ \phi_{\alpha}(r) \} \) instead of orthogonal extended wavefunctions \( \{ \psi_n(r) \} \)
- Aim: to achieve the same accuracy as traditional plane-wave methods
Our basis set consists of plane-waves combined into spike-like “psinc” functions

\[
D(\mathbf{r}) = \frac{1}{N} \sum_{\mathbf{G}}^{G_{\text{max}}} e^{i\mathbf{G} \cdot \mathbf{r}}
\]

\[
\phi_\alpha(\mathbf{r}) = \sum_{\text{sim. cell}} D_m(\mathbf{r}) C_{m\alpha}
\]

\[
= \sum_{\text{sim. cell}} D(\mathbf{r} - \mathbf{r}_m) C_{m\alpha}
\]

\[
C_{m\alpha} = 0 \text{ if } m \notin \text{ sphere of } \alpha
\]
Linear-scaling formulation

\[ \rho(\mathbf{r}, \mathbf{r}') = \sum_{\alpha\beta} \phi_{\alpha}(\mathbf{r}) K^{\alpha\beta} \phi^*_\beta(\mathbf{r}') \]

Density \[ n(\mathbf{r}) = 2\rho(\mathbf{r}, \mathbf{r}) \]

Energy \[ E[\eta] = E[\{K^{\alpha\beta}\}, \{\phi_{\alpha}(\mathbf{r})\}] = E[\{K^{\alpha\beta}\}, \{C_{m\alpha}\}] \]

Short-ranged: \[ \rho(\mathbf{r}, \mathbf{r}') \to 0 \quad \text{as} \quad |\mathbf{r} - \mathbf{r}'| \to \infty \]

\[ \Rightarrow \text{impose cutoffs:} \quad \phi_{\alpha}(\mathbf{r}) = 0 \quad \text{when} \quad |\mathbf{r} - \mathbf{R}_\alpha| > R_{\text{reg}} \]

\[ K^{\alpha\beta} = 0 \quad \text{when} \quad |\mathbf{R}_\alpha - \mathbf{R}_\beta| > R_{\text{cut}} \]

Idempotent:

\[ \rho^2(\mathbf{r}, \mathbf{r}') = \int \rho(\mathbf{r}, \mathbf{r}'') \rho(\mathbf{r}'', \mathbf{r}') d^3r'' = \rho(\mathbf{r}, \mathbf{r}') \]
Parallelisation Strategy

Demonstration with four processors (0 to 3)

Serial calculation

Parallel calculation

Phase 1
Distribution of atomic data (each processor holds only a subset of the \{\varphi_\alpha\})
⇒ Large number of atoms

Phase 2 (New work - completed in July 2004)
Distribution of simulation cell data (each processor holds only a slice of the charge density and local potentials)
⇒ Large simulation cells
Phase 2: Most challenging aspect - the FFTbox technique

• Fourier transforms in **small regions** of the simulation cell, independent of system size

• Each processor has its own FFTbox

• Construction of the Hamiltonian matrix requires “filling” the FFTbox with “slices” of local potential from different processors

• Construction of the charge density needs the opposite: “Deposit” the contents of each FFTbox to “slices” on different processors

• A complex yet efficient **algorithm for communication** between processors has been developed

Chris-Kriton Skylaris, Electronic Structure Discussion Group, 25/8/2004
Example: Plane-wave calculations in massive simulation cells!

Test - supercell approximation for a 500-atom capped nanotube. KE Cutoff 574eV. $R_{reg} = 8.0 \, a_0$. $R_{cut} = \infty$. 96 processors (Franklin).

Simulation cell volume: $50\text{Å} \times 50\text{Å} \times 50\text{Å}$ $100\text{Å} \times 100\text{Å} \times 100\text{Å}$ $160\text{Å} \times 160\text{Å} \times 160\text{Å}$

Maximum possible simulation cell size before Phase 2 parallelisation

Total energy (Hartree): $-2526.007305$ $-2526.006203$ $-2526.010921$
Single-point total energy calculations on DNA

64 processors, simulation cell of 30Åx30Åx220Å. Wannier function RMS gradient convergence threshold 1.0E-6 Eh*a0^(3/2)

<table>
<thead>
<tr>
<th>N base pairs</th>
<th>N atoms</th>
<th>N iterations</th>
<th>Total time (h)</th>
<th>DeltaE/atom (Eh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>117</td>
<td>26</td>
<td>2.0</td>
<td>1.3E-8</td>
</tr>
<tr>
<td>4</td>
<td>248</td>
<td>29</td>
<td>4.2</td>
<td>9.6E-9</td>
</tr>
<tr>
<td>6</td>
<td>379</td>
<td>29</td>
<td>6.4</td>
<td>6.6E-9</td>
</tr>
<tr>
<td>8</td>
<td>510</td>
<td>29</td>
<td>8.7</td>
<td>6.5E-9</td>
</tr>
<tr>
<td>10</td>
<td>641</td>
<td>29</td>
<td>11.2</td>
<td>7.6E-9</td>
</tr>
<tr>
<td>20</td>
<td>1296</td>
<td>30</td>
<td>26.9</td>
<td>1.1E-8</td>
</tr>
<tr>
<td>30</td>
<td>1951</td>
<td>28</td>
<td>44.6</td>
<td>1.1E-8</td>
</tr>
</tbody>
</table>

Chris-Kriton Skylaris, Electronic Structure Discussion Group, 25/8/2004
Conclusions

ONETEP is a pseudopotential DFT method with Plane-wave accuracy

• **Linear-scaling** w.r.t. to the number of atoms

• Now also able to treat **extremely large simulation cells**

• These simulation cell volumes can fit **many thousands of atoms**, so calculations of this size should now be possible

• Speedups should result from future **improvements** in the new “Phase 2 parallelisation” code

Acknowledgments

Cambridge-Cranfield High Performance Computing Facility (CCHPCF) for access to “Franklin”, their SUN-Fire 900-processor parallel supercomputer