Massive plane-wave calculations in massive simulation cells

Chris-Kriton Skylaris

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

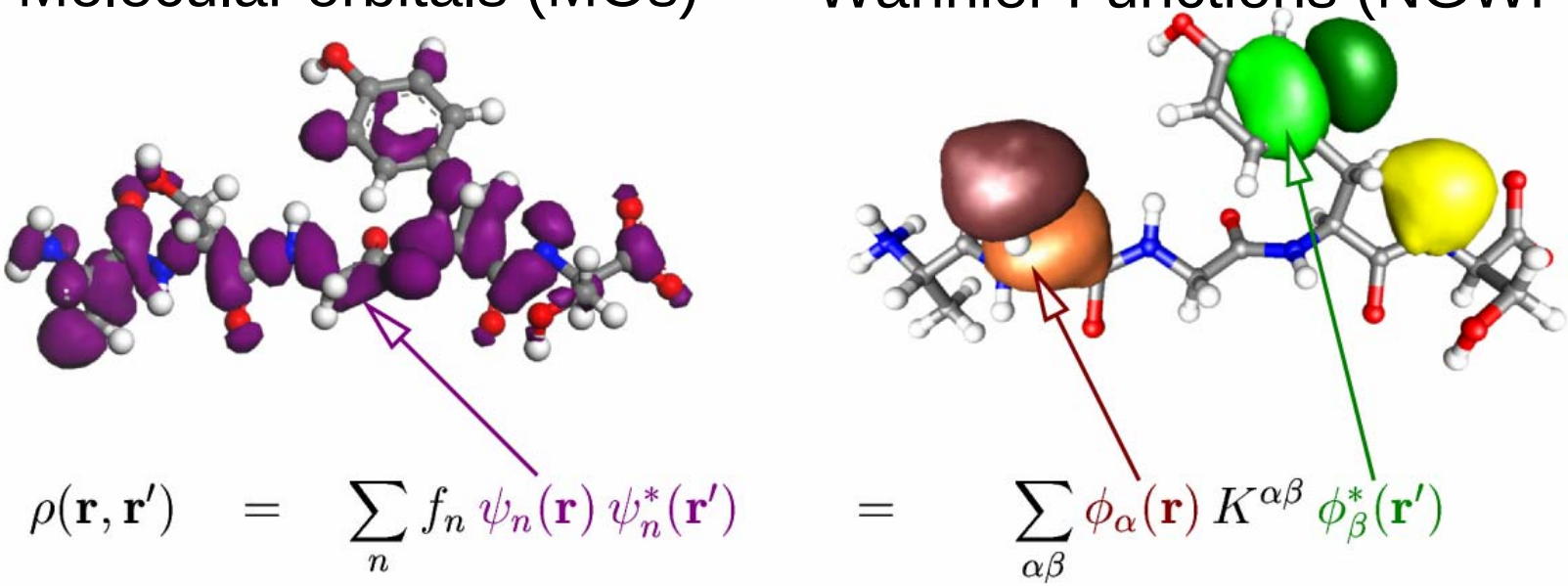
Theory of Condensed Matter group
Cavendish Laboratory
Madingley Road
Cambridge CB3 0HE

UNIVERSITY OF CAMBRIDGE
CAVENDISH LABORATORY

ONETEP: Density-matrix linear-scaling DFT

Molecular orbitals (MOs)

Non-orthogonal Generalised Wannier Functions (NGWFs)

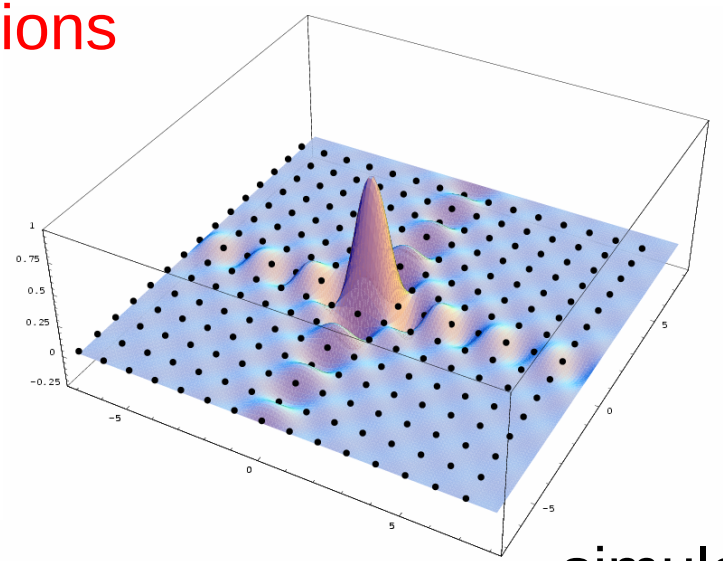


$$\rho(\mathbf{r},\mathbf{r}') = \sum_n f_n\,\psi_n(\mathbf{r})\,\psi_n^*(\mathbf{r}') = \sum_{\alpha\beta}\phi_\alpha(\mathbf{r})\,K^{\alpha\beta}\,\phi_\beta^*(\mathbf{r}')$$

- Optimise non–orthogonal localised functions $\{\phi_\alpha(\mathbf{r})\}$ instead of orthogonal extended wavefunctions $\{\psi_n(\mathbf{r})\}$ } linear scaling
- Aim: to achieve the same accuracy as traditional plane–wave methods

Our basis set consists of plane-waves combined into spike-like “psinc” functions

$$D(\mathbf{r}) = \frac{1}{N} \sum_{\mathbf{G}}^{\mathbf{G}_{max}} e^{i\mathbf{G}\cdot\mathbf{r}}$$

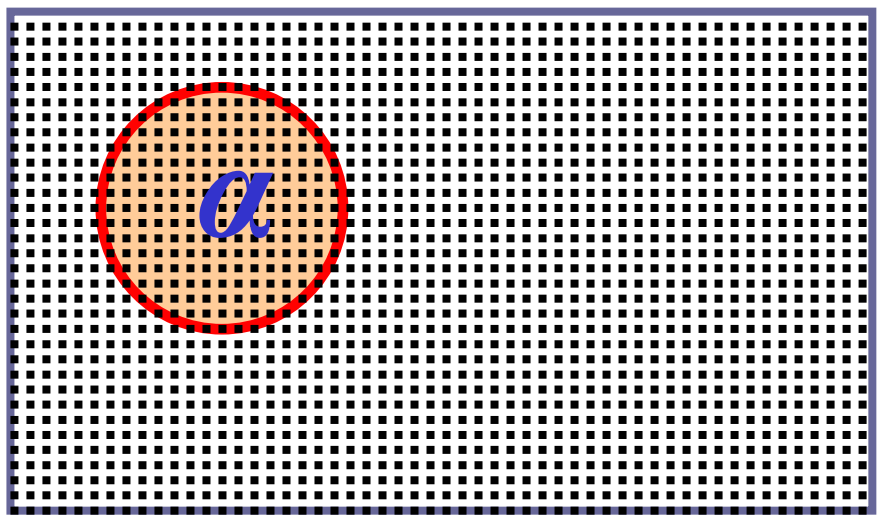


$$\phi_\alpha(\mathbf{r}) = \sum_{m}^{\text{sim. cell}} D_m(\mathbf{r}) C_{m\alpha}$$

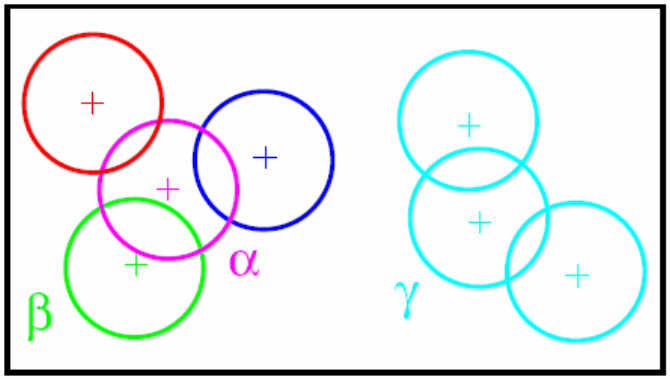
$$= \sum_{m}^{\text{sim. cell}} D(\mathbf{r} - \mathbf{r}_m) C_{m\alpha}$$

$$C_{m\alpha} = 0 \text{ if } m \notin \text{sphere of } \alpha$$

simulation cell



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

Short−ranged: $\rho(\mathbf{r},\mathbf{r}') \longrightarrow 0$ as $|\mathbf{r}-\mathbf{r}'| \longrightarrow \infty$

$\Rightarrow$ impose cutoffs: $\phi_\alpha(\mathbf{r}) = 0$ when $|\mathbf{r}-\mathbf{R}_\alpha| > R_{\rm reg}$

$K^{\alpha\beta} = 0$ when $|\mathbf{R}_\alpha - \mathbf{R}_\beta| > R_{\rm cut}$

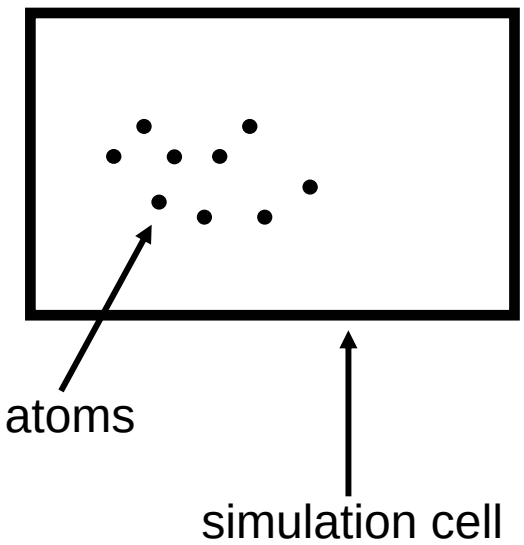
Idempotent:

$$\rho^2(\mathbf{r},\mathbf{r}') = \int \rho(\mathbf{r},\mathbf{r}'')\,\rho(\mathbf{r}'',\mathbf{r}')\,d^3r'' = \rho(\mathbf{r},\mathbf{r}')$$

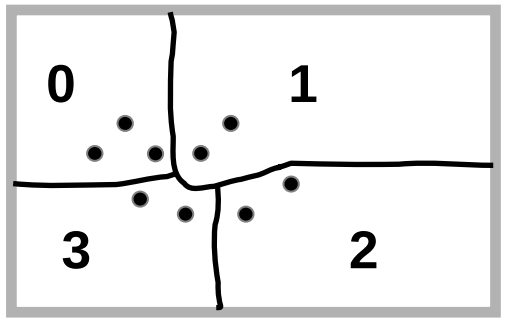
Parallelisation Strategy

Demonstration with four processors (0 to 3)

Serial calculation



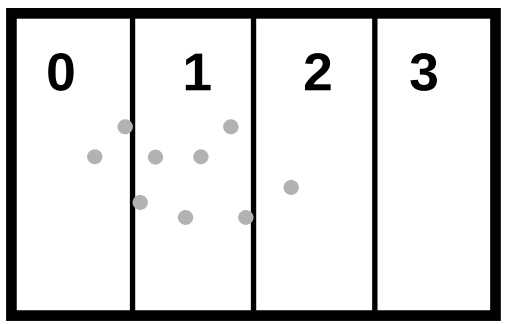
Parallel calculation



Phase 1

Distribution of atomic data (each processor holds **only a subset** of the $\{\varphi\alpha\}$)

⇒ **Large number of atoms**

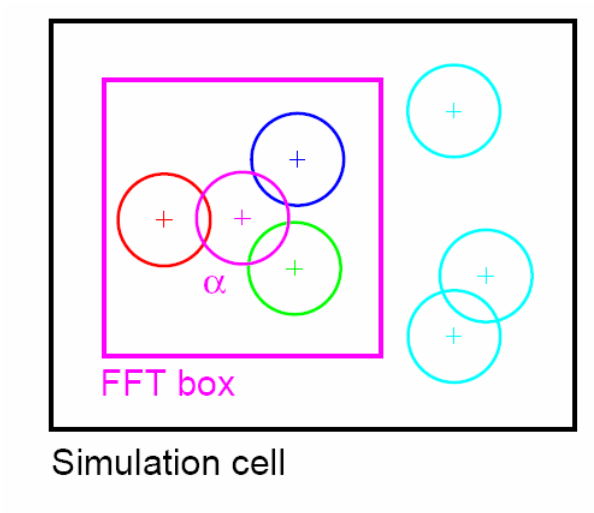


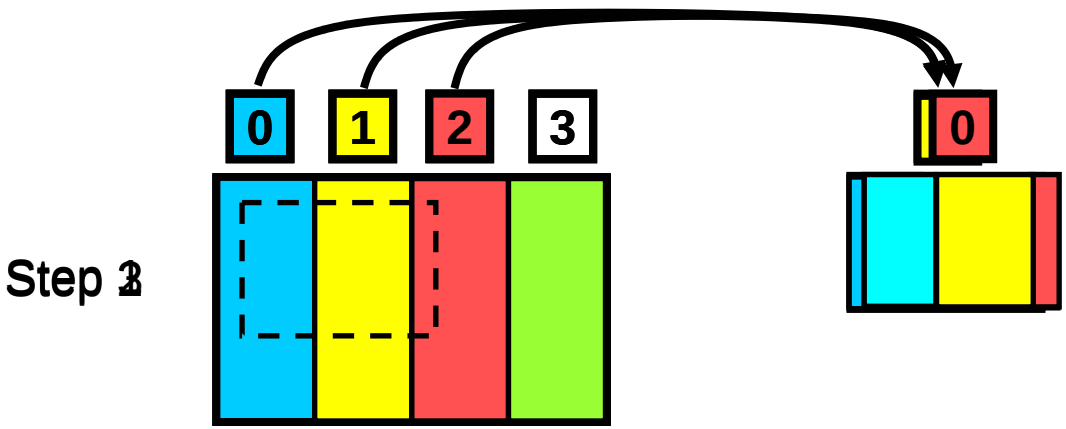
Phase 2 (**New work - completed in July 2004**)

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

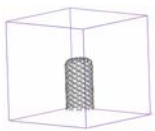
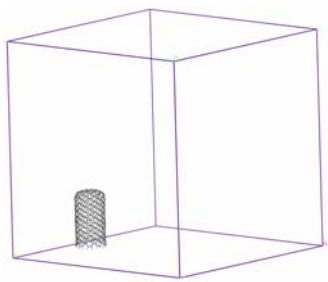
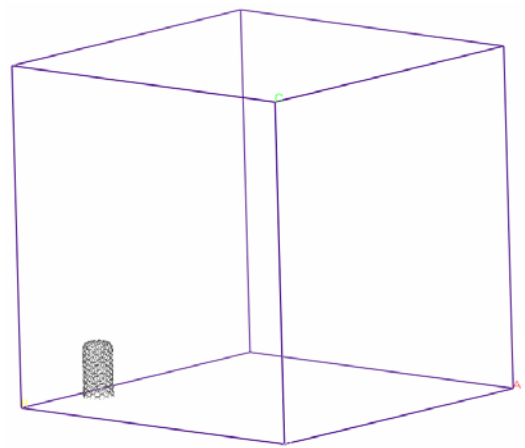




- Fourier transforms in **small regions** of the simulation cell, **independent of system size**
- Each processor has **its own** FFTbox
- Construction of the Hamiltonian matrix requires **"filling" the FFTbox with "slices" of local potential** from different processors
- Construction of the **charge 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

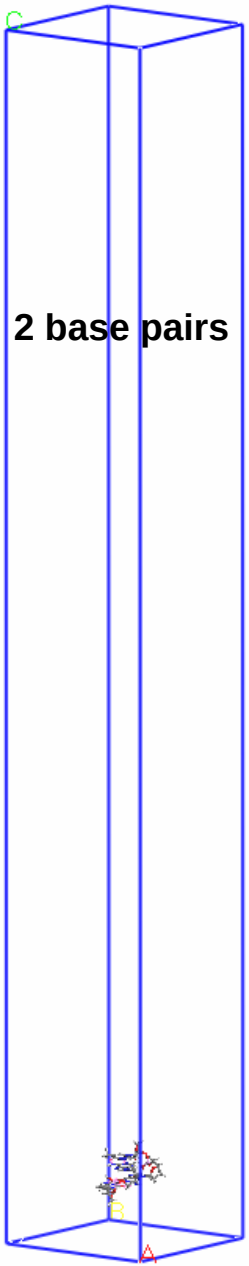
Example: Plane-wave calculations in massive simulation cells!

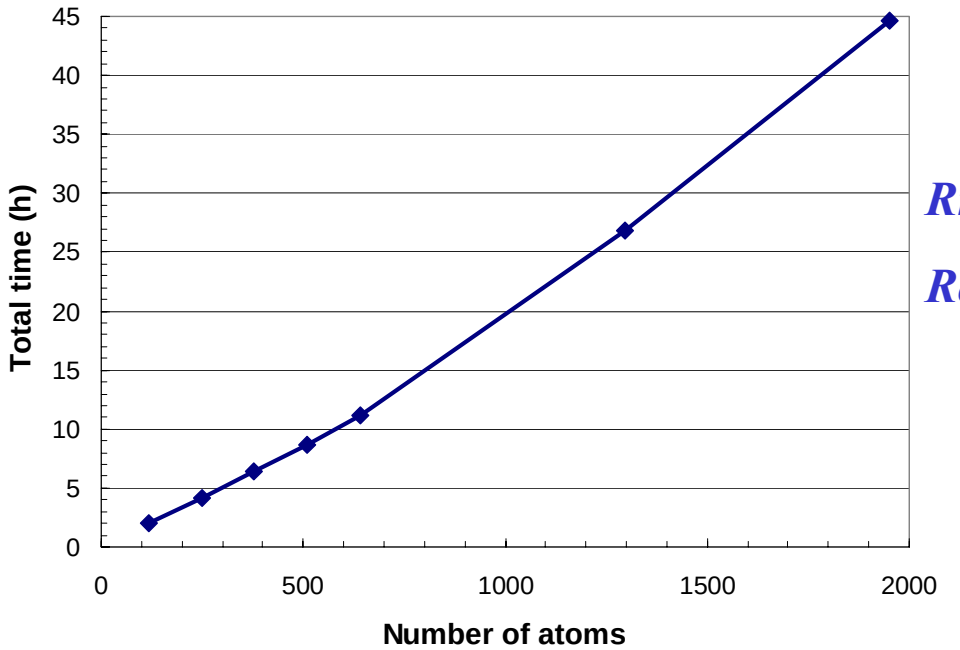
Test - supercell approximation for a 500-atom capped nanotube.
KE Cutoff 574eV. $Rreg$ =8.0 a0. $Rcut$=∞. 96 processors (Franklin).

Simulation cell volume:	**50Å x 50Å x 50Å**	**100Å x 100Å x 100Å**	**160Å x 160Å x 160Å**
			
	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

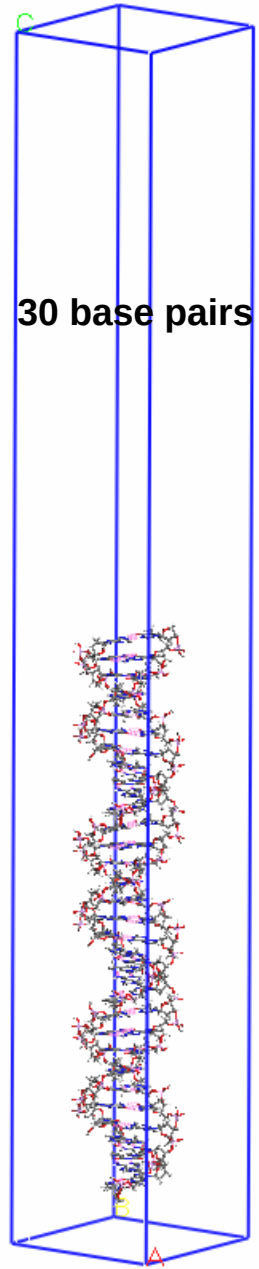
64 processors, simulation cell of 30Åx30Åx220Å. Wannier function RMS gradient convergence threshold 1.0E-6 Eh*a0^(3/2)





Rreg=6.3a0

Rcut=25.0a0

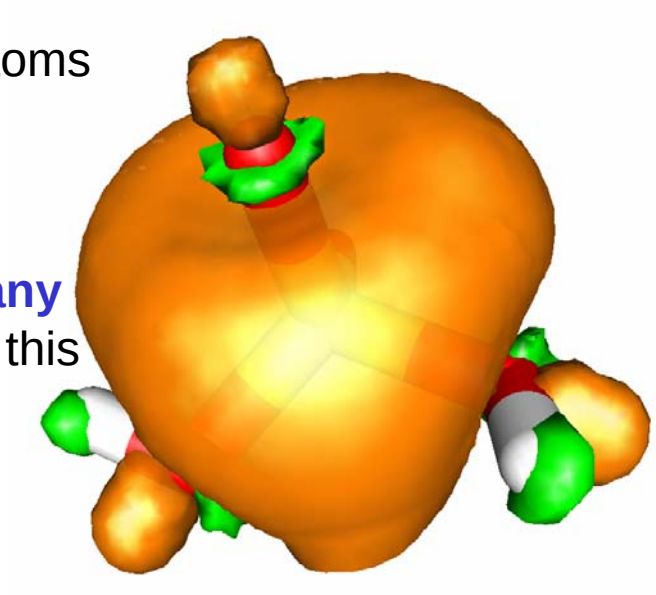


N base pairs	N atoms	N iterations	Total time (h)	DeltaE/atom (Eh)
2	117	26	2.0	1.3E-8
4	248	29	4.2	9.6E-9
6	379	29	6.4	6.6E-9
8	510	29	8.7	6.5E-9
10	641	29	11.2	7.6E-9
20	1296	30	26.9	1.1E-8
30	1951	28	44.6	1.1E-8

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