UNIVERSITY OF
Southampton
School of Chemistry

Introduction to ONETEP

Chris-Kriton Skylaris

Outline

- Density matrix reformulation of DFT
- Localised function representation of density matrix
- Linear-scaling with localised functions
- Linear-scaling with large basis set accuracy
  - NGWFs, density kernel
  - Plane waves and psinc basis set
  - FFT box
  - Parallel scaling
- Compilation and hardware requirements
- Running a simple calculation
- Functionality available
- Examples of ONETEP applications

ONETEP Developers' group (ODG)) of core developers

Jacek Dziedzic,[1,2] Peter D. Haynes,[3] Nicholas D. M. Hine,[4] Arash A. Mostofi,[3]
Mike C. Payne[5] and Chris-Kriton Skylaris[1]

__2__ Gdansk University of Technology

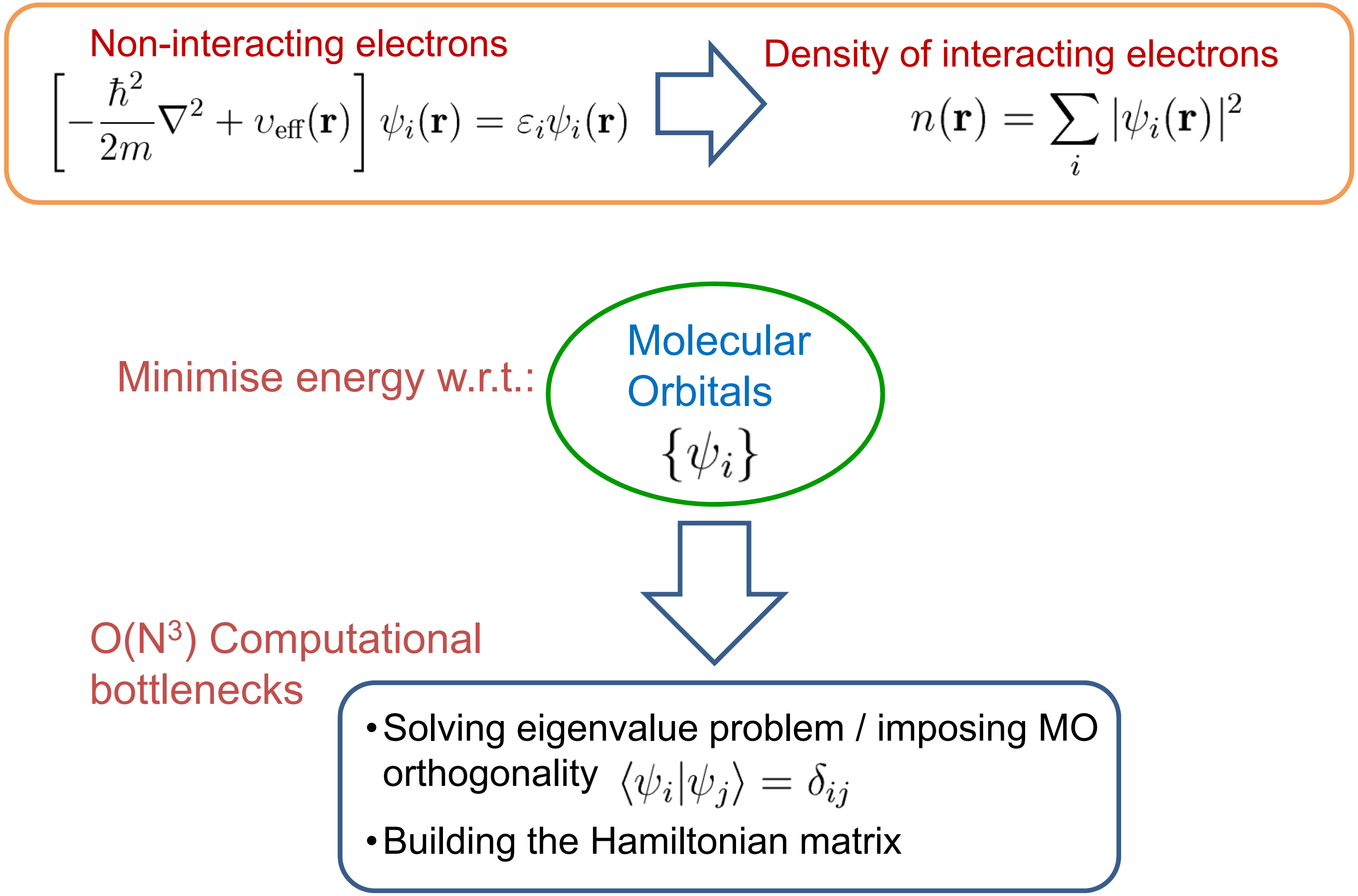
__4__ University of Warwick

__5__ University of Cambridge

__3__ Imperial College London

__1__ University of Southampton

Computational bottlenecks in DFT



Non-interacting electrons

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + v_{\text{eff}}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r})$$

Density of interacting electrons

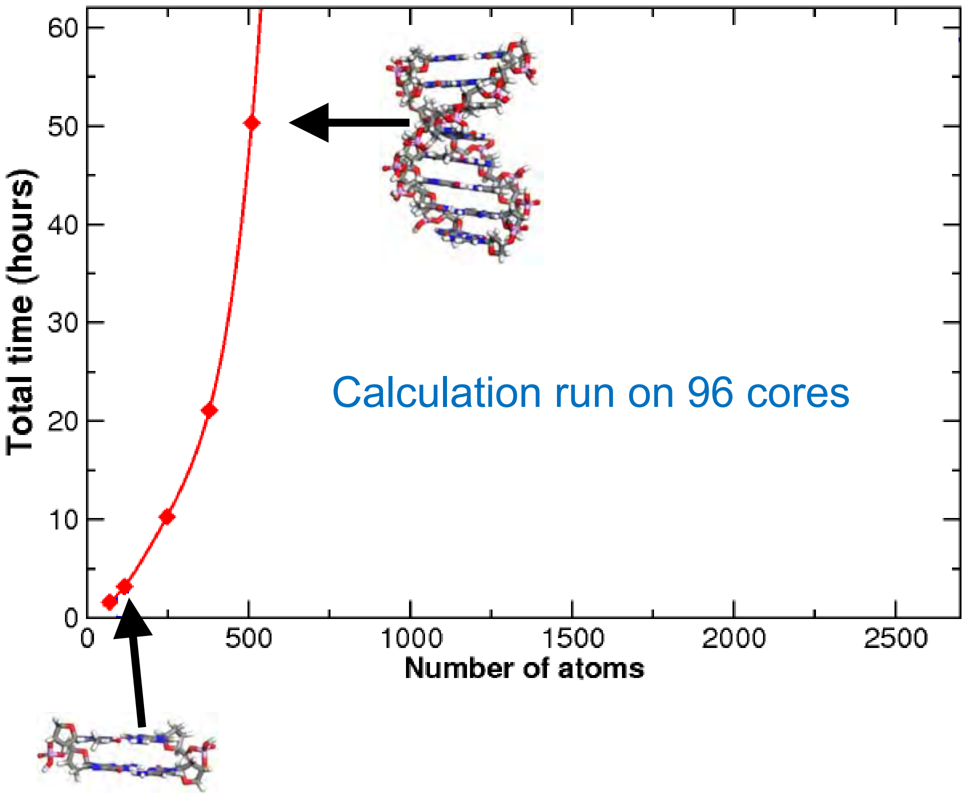
$$n(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2$$

Minimise energy w.r.t.: Molecular Orbitals $\{\psi_i\}$

O(N^3) Computational bottlenecks

- Solving eigenvalue problem / imposing MO orthogonality $\langle\psi_i|\psi_j\rangle = \delta_{ij}$
- Building the Hamiltonian matrix

Computational cost of DFT: cubic-scaling



- Not suitable for biomolecules/nanostructures with **thousands of atoms**
- A **linear-scaling reformulation of DFT** is needed

Linear-scaling DFT

- Physical principle

Nearsightedness of electronic matter
W. Kohn, *Phys. Rev. Lett*. **76**, 3168 (1996)

In molecules with non-zero band gap, the density matrix decays exponentially

$$\rho(\mathbf{r},\mathbf{r}') \sim e^{-\gamma|\mathbf{r}-\mathbf{r}'|} \to 0 \quad \text{as} \quad |\mathbf{r}-\mathbf{r}'| \to \infty$$

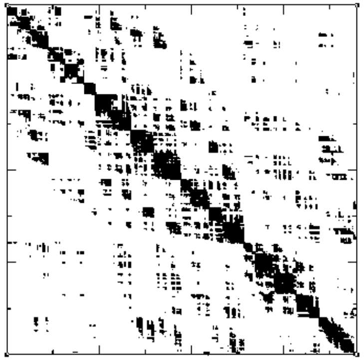
- Linear-scaling approaches

Truncate exponential "tail"

$$\rho(\mathbf{r},\mathbf{r}') = 0 \quad \text{when} \quad |\mathbf{r}-\mathbf{r}'| > r_{\text{cut}}$$

- Practical implementation

- Localised orbitals
- No diagonalisation
- Energy minimisation
- Sparse matrices and algorithms – O(N) memory and CPU cost



Density matrix DFT: energy expression

DFT energy with molecular orbitals

$$E_{\mathrm{DFT}} = \sum_i f_i \int \psi_i^*(\mathbf{r}) \left(-\frac{1}{2}\nabla^2\right) \psi_i(\mathbf{r}) d\mathbf{r} + \int V_{ext}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{xc}[n]$$

Density matrix

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

Molecular orbitals

Localised orbitals

Density

$$n(\mathbf{r}) = \rho(\mathbf{r}, \mathbf{r})$$

DFT energy with density matrix

$$E_{\mathrm{DFT}} = -\sum_{\alpha,\beta} K^{\alpha\beta} \int \phi_\beta^*(\mathbf{r}) \frac{1}{2} \nabla^2 \phi_\alpha(\mathbf{r}) d\mathbf{r} + \int V_{\mathrm{ext}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \int \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{xc}[n]$$

Idempotency condition on the density matrix

Density matrix

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

Molecular orbitals

Localised orbitals

MO orthonormality

$$\langle \psi_n | \psi_m \rangle = \delta_{nm}$$

Orbital occupancies

$$f_n = 1 \text{ or } 0$$

Idempotency condition

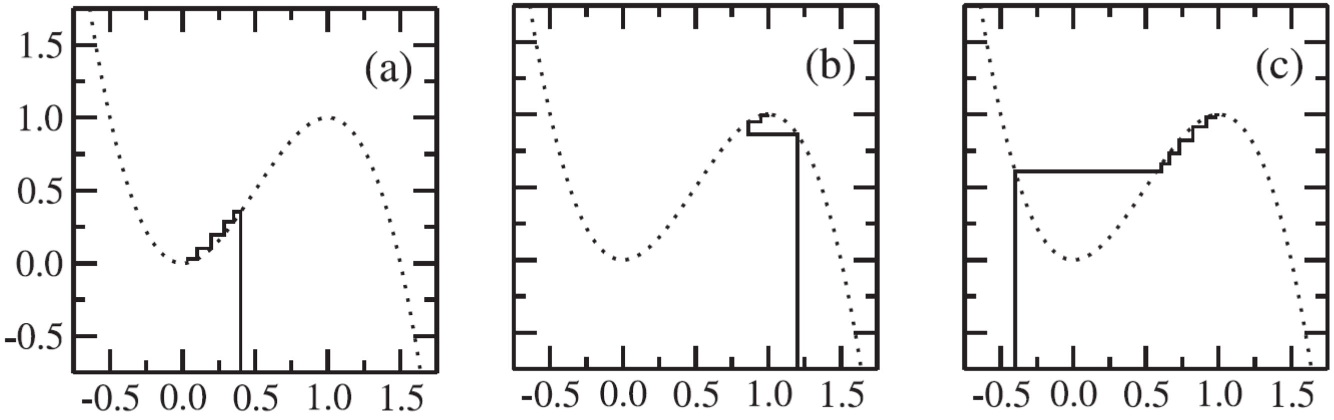
$$\mathbf{KSK} = \mathbf{K}$$

$$S_{\alpha\beta} = \langle \phi_\alpha | \phi_\beta \rangle$$

Maintaining idempotency: Purification transformation

$$f_{new} = 3f_{old}^2 - 2f_{old}^3$$

R. McWeeny, *Rev. Mod. Phys.* **32**, 335 (1960)



- Improves the idempotency of a nearly idempotent density matrix
- Forces occupancies in interval [0,1] ("weak" idempotency)
- Quadratic convergence. Example:
  - $f^{(1)} = 1.1$
  - $f^{(2)} = 3\times1.1^3 - 2\times1.1^2 = 0.968$
  - $f^{(3)} = 3\times0.968^3 - 2\times0.968^2 = 0.997$
  - Etc..
- But, diverges if the initial occupancies are not within certain bounds

UNIVERSITY OF Southampton
School of Chemistry

Energy expression containing a purification transformation

- If we minimise $E_{DFT}(\mathbf{K})$ w.r.t. **K** without any constraint we will obtain minus infinity as there is nothing to stop it converging to a **K** that has infinite occupancies
- We need to minimise subject to the impotency constraint for **K**
- This constraint is difficult to impose explicitly, but we can impose it implicitly by writing **K** as a purification transformation of an **auxiliary density matrix L**

$$\mathbf{K} = 3\mathbf{LSL} - 2\mathbf{LSLSL}$$

- Then, $E_{DFT}(\mathbf{K}(\mathbf{L}))$ is a function of **L** and EDFT(**L**) and is minimised (unconstrained) w.r.t. **L**. Of course, at convergence **K=L**
- This is the basis of the Li-Nunes-Vanderbilt (LNV) method
- It is not robust: the minimisation diverges often as at some point **L** will fall outside the convergence interval of the purification transformation. Also charge conservation is not guaranteed

Our robust LNV approach

P. D. Haynes, C.-K. Skylaris, A. A. Mostofi, and M. C. Payne, *J. Phys.: Condens. Matter* **20**, 294207 (2008)

- We minimise the following expression

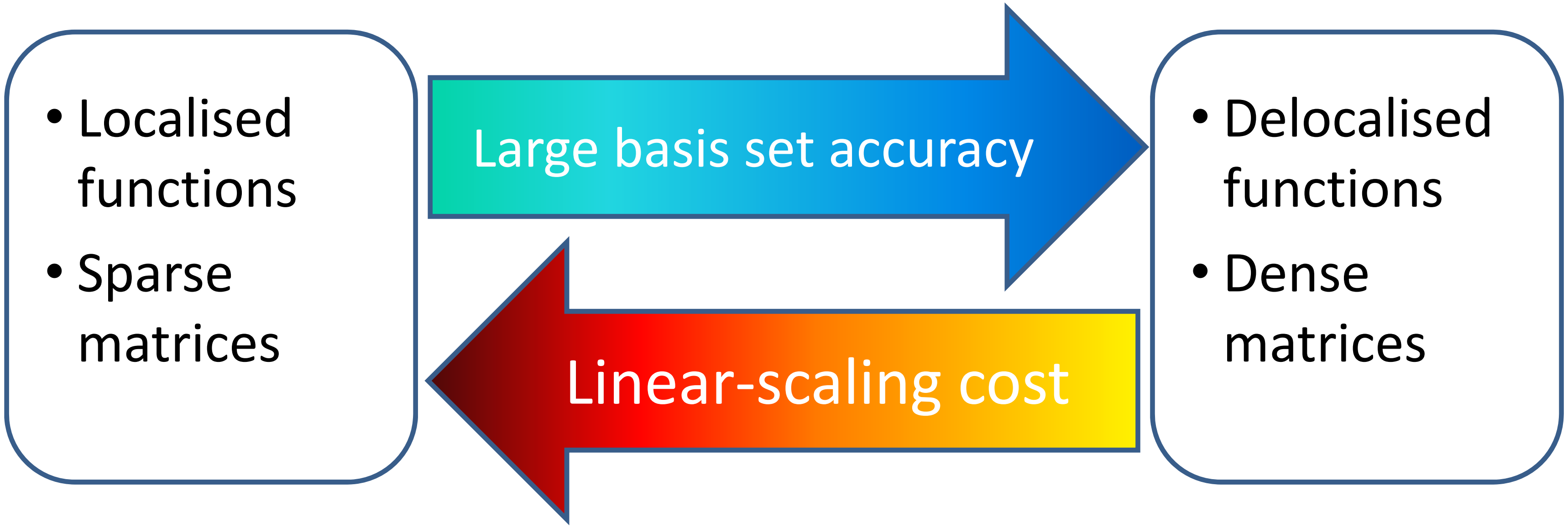
$$E_{\mathrm{DFT}}\left(\frac{N_e}{\mathrm{tr}(\mathbf{KS})}\mathbf{K}\right) = E_{\mathrm{DFT}}(\mathbf{L}) \qquad \mathbf{K} = 3\mathbf{LSL} - 2\mathbf{LSLSL}$$

which preserves the number of electrons by scaling the purified density matrix.

- The purification transformation is never allowed to break down:
  - Iteratively monitor the lowest and highest occupancy values
  - If occupancies are detected to be outside of the convergence interval of the purification we switch to an globally convergent penalty functional expression until the occupancies are back within the purification convergence range

Linear-scaling DFT in practice

Conflicting requirements



- Localised functions
- Sparse matrices

Large basis set accuracy

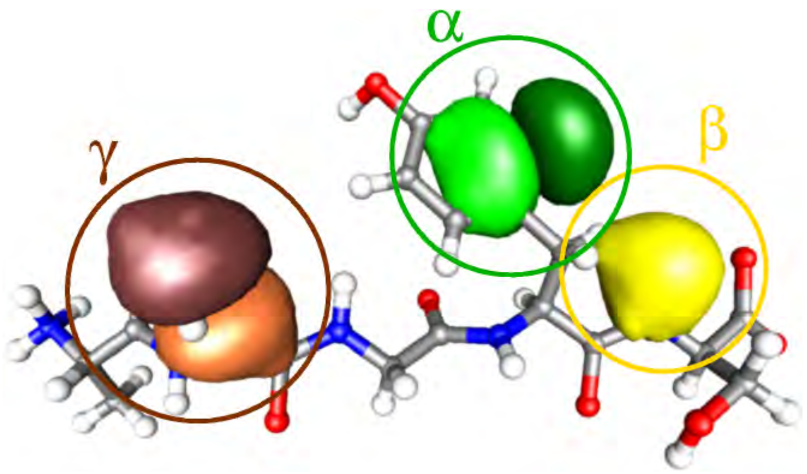
Linear-scaling cost

- Delocalised functions
- Dense matrices

How can we have linear-scaling cost with “cubic-scaling accuracy” ?

The ONETEP linear-scaling approach

- Use a minimal number of $\{\phi_\alpha\}$
- Optimise both **K** and $\{\phi_\alpha\}$
- Aim is to have linear-scaling DFT with large basis set accuracy



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

Non-orthogonal Generalised Wannier Functions (**NGWFs**)

Density kernel

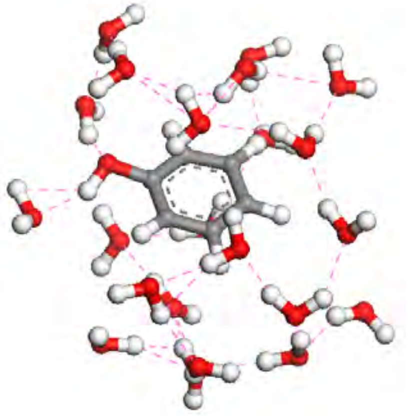
- NGWFs confined to spherical regions
- Density kernel **K** sparse by truncation

C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, O. Dieguez, M. C. Payne, *Phys. Rev. B* **66**, 035119 (2002).

C.-K. Skylaris, P. D. Haynes, A. A. Mostofi and M. C. Payne, *J. Chem. Phys.* **122**, 084119 (2005).

C.-K. Skylaris, P. D. Haynes, A. A. Mostofi and M. C. Payne, *Phys. Stat. Sol. (b)* **243**, 973 (2006).

Linear-scaling with near-complete basis set accuracy



Example: binding energy calculation

ONETEP linear-scaling DFT

NGWF radii (Å)	# NGWFs	BE (kcal/mol)
2.9	166	**-11.93**
3.2	166	**-12.86**
3.7	166	**-8.25**
4.2	166	**-7.06**
4.8	166	**-7.04**

Cubic-scaling DFT

Basis set	# AOs	BE + BSSE (kcal/mol)
3-21G	361	**-12.55**
6-31G*	535	**-8.95**
6-311++G**	1017	**-7.39**
cc-pVTZ	1765	**-7.04**

C.-K. Skylaris, O. Dieguez, P. D. Haynes and M. C. Payne, *Phys. Rev. B* **66**, 073103 (2002).

P. D. Haynes, C.-K. Skylaris, A. A. Mostofi and M. C. Payne, *Chem. Phys. Lett.* **422** 345 (2006).

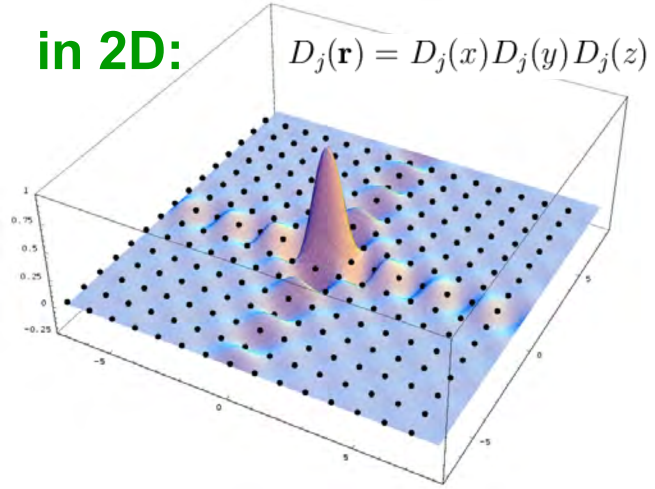
Psinc basis set

- Real linear combinations of plane waves
- Highly localised
- Orthogonal

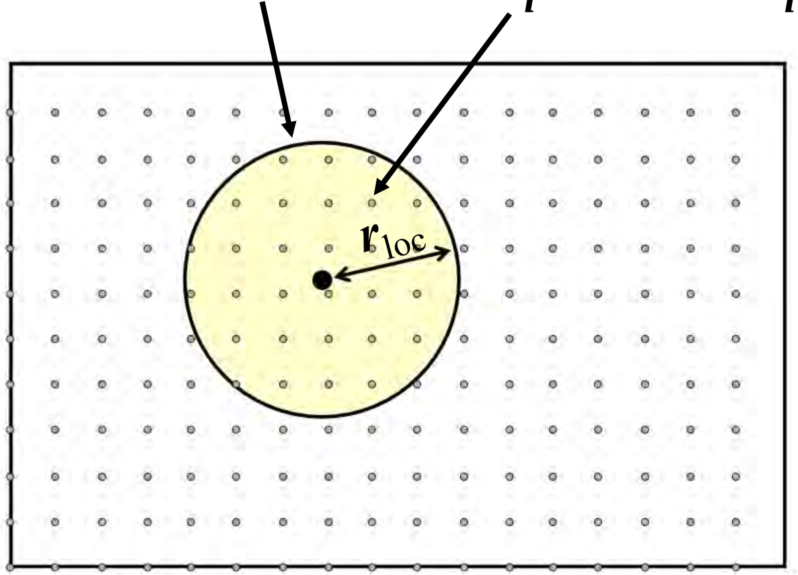
$$D_j(\mathbf{r}) = \frac{1}{N} \sum_{\mathbf{G}}^{\mathbf{G}_{\max}} e^{i\mathbf{G}\cdot(\mathbf{r}-\mathbf{r}_j)}$$

$$\phi_\alpha(\mathbf{r}) = \sum_{i \in \mathrm{LR}_\alpha} D_i(\mathbf{r})\, C_{i,\alpha}$$

in 2D:



$$D_j(\mathbf{r}) = D_j(x)D_j(y)D_j(z)$$

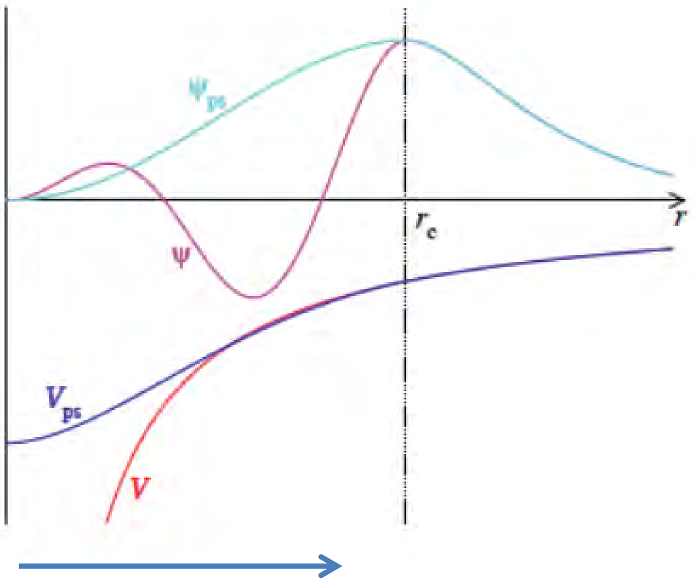


NGWF localisation sphere

$\mathbf{r}_i$ centre of $D_i(\mathbf{r})$

- A. A. Mostofi, P. D. Haynes, C.-K. Skylaris and M. C. Payne, *J. Chem. Phys.* **119**, 8842 (2003)
- D. Baye and P. H. Heenen, *J. Phys. A: Math. Gen.* **19**, 2041 (1986)

Pseudopotentials

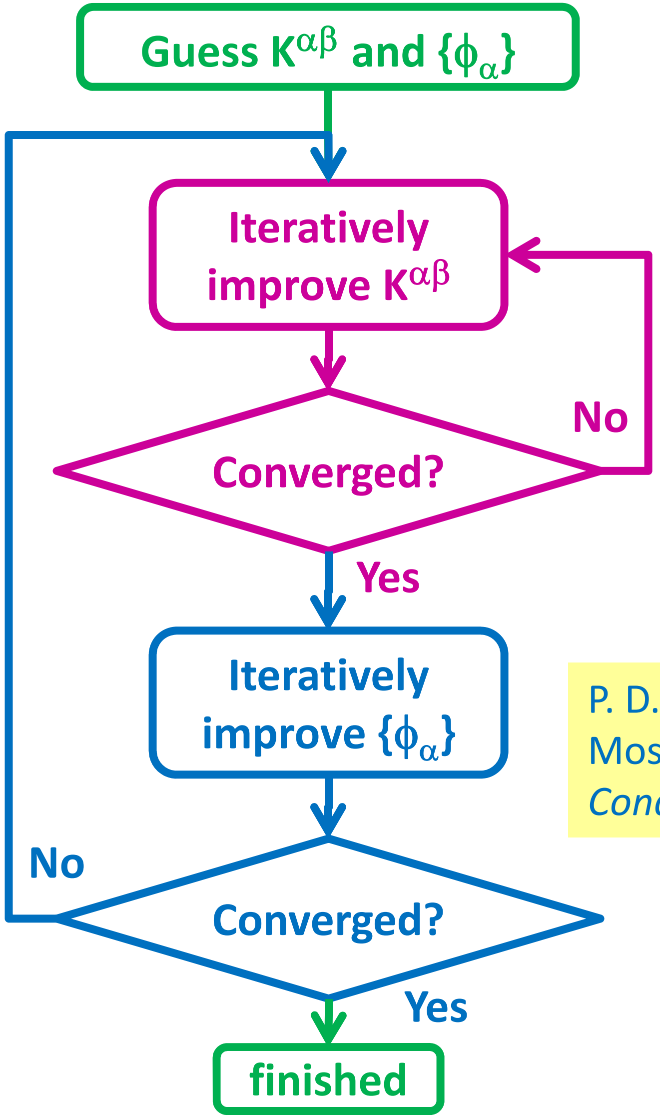


Distance from nucleus

- Replace the strong Coulomb nuclear potential V by a much weaker *pseudopotential* V_{ps} which mimics the effect of the core electrons
- Only keep the valence electrons, in smooth valence wave functions Ψ_{ps} without oscillations in the core region
- Eliminating the 1/r singularity at the nucleus and the rapid oscillations of the valence wavefunctions near it make it possible to use a plane wave basis set and converge with a reasonable kinetic energy cut-off
- Even hydrogen atoms (obviously, no core electrons) have pseudopotentials in plane wave calculations!

- Beyond the core radius r_c, the V_{ps} and Ψ_{ps} match the all-electron calculation
- Determining r_c, as well as which electrons to remove as “core” and which to keep as “valence” is an art and requires experience
- Small r_c leads to “hard” pseudopotentials which are more transferable but require more plane waves (higher E_{max}) and therefore more computational effort
- There are several recipes for generating pseudopotentials and libraries of these are publicly available, see for example http://opium.sourceforge.net/ and http://www.sas.upenn.edu/rappegroup/research/pseudo-potential-gga.html
- Before selecting a pseudopotential it is crucial to test it thoroughly, ideally against all-electron calculations

Energy optimisation in ONETEP



Guess $K^{\alpha\beta}$ and $\{\phi_\alpha\}$

Iteratively improve $K^{\alpha\beta}$

Converged?

No

Yes

Iteratively improve $\{\phi_\alpha\}$

Converged?

No

Yes

finished

$$E = E[\mathbf{K}, \{\phi_\alpha\}]$$

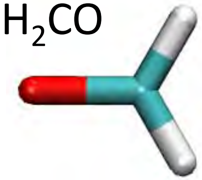
- **Outer loop:** Optimise total (interacting) energy *E* w.r.t. to **K and $\{\phi_\alpha\}$**

- **Inner loop**: Optimise total (interacting) energy *E* w.r.t **K** for fixed **$\{\phi_\alpha\}$** while imposing **idempotency** and **normalisation**

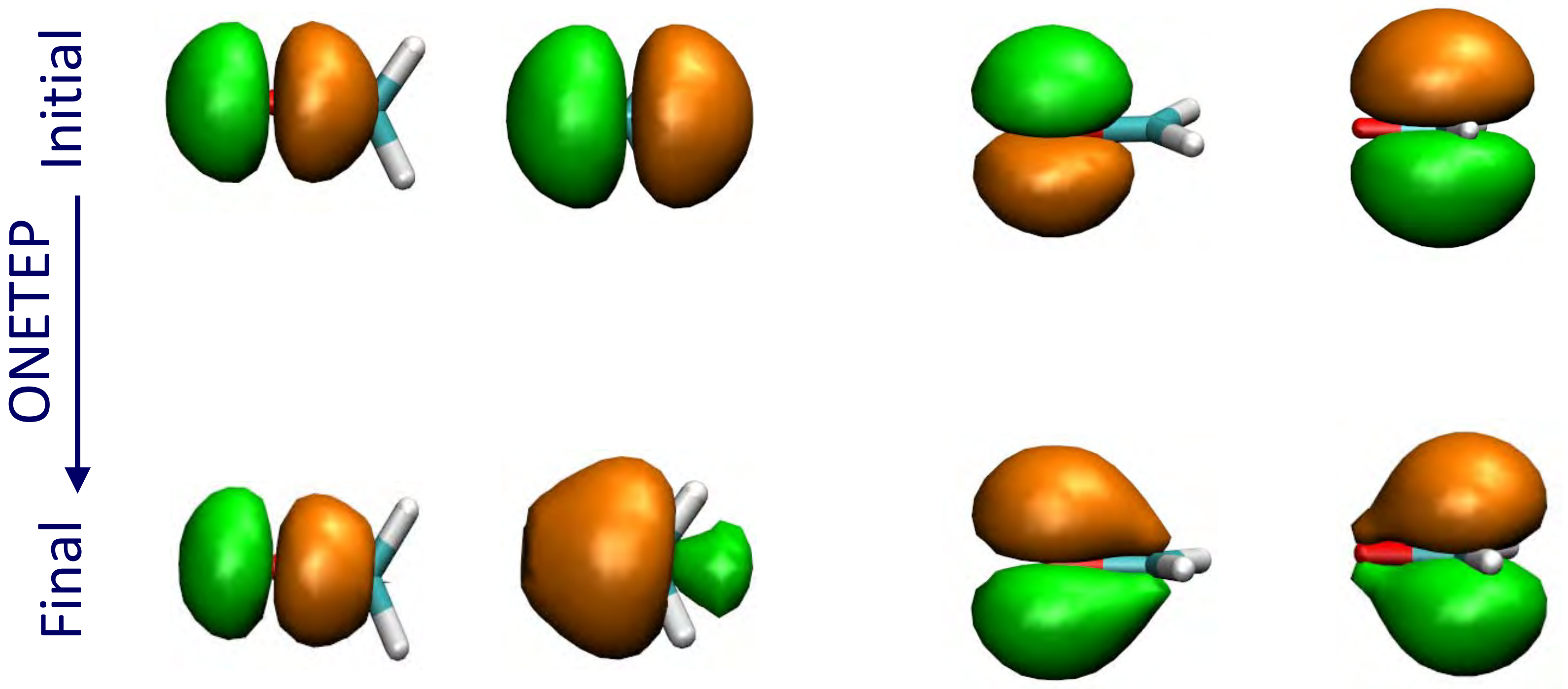
P. D. Haynes, C.-K. Skylaris, A. A. Mostofi and M. C. Payne, *J. Phys. Condens*. Matter **20**, 294207 (2008)

NGWF optimisation

formaldehyde, H_2CO



O *p* C p O *p* C p

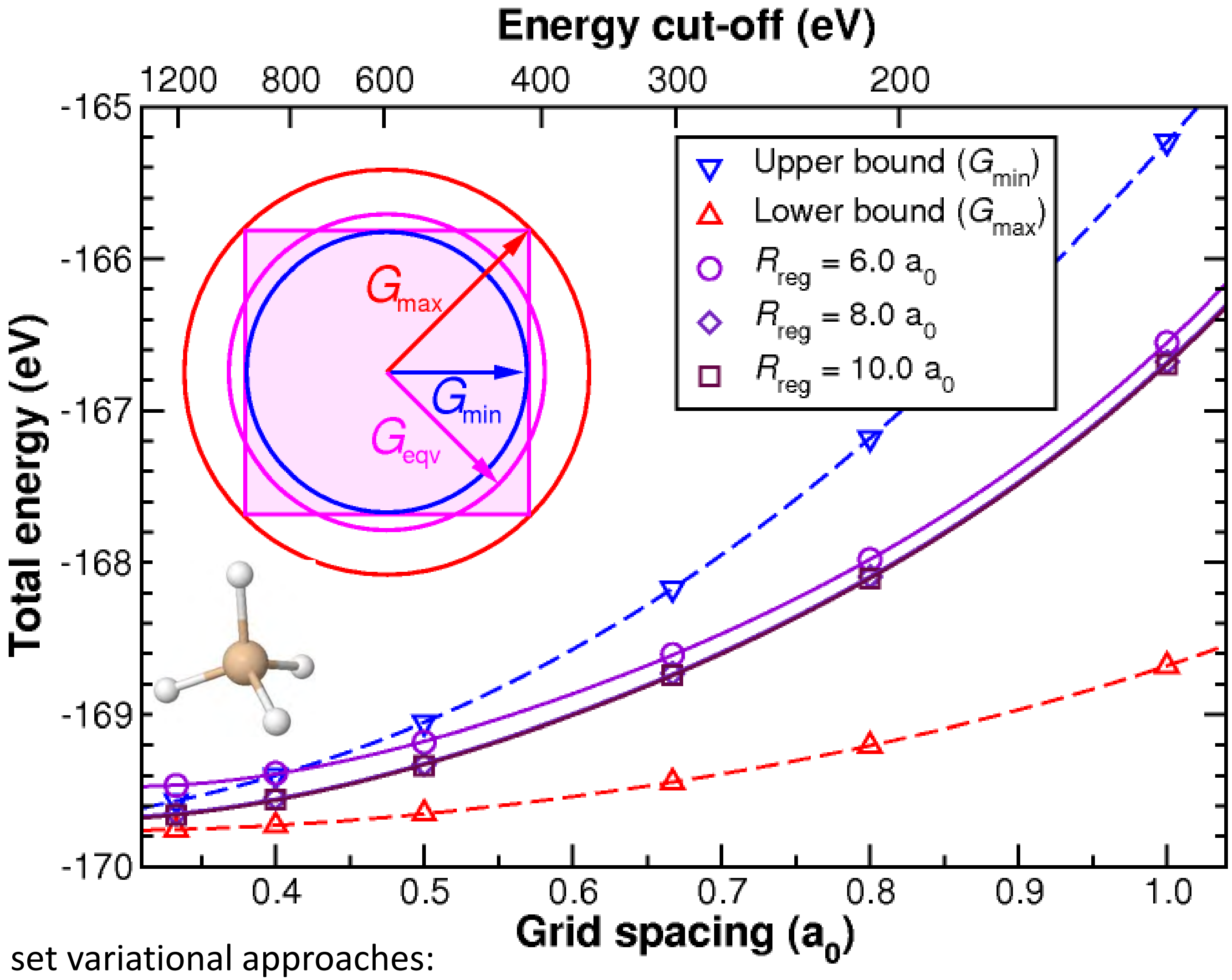


Initial

ONETEP

Final

Psinc basis energy cut-off

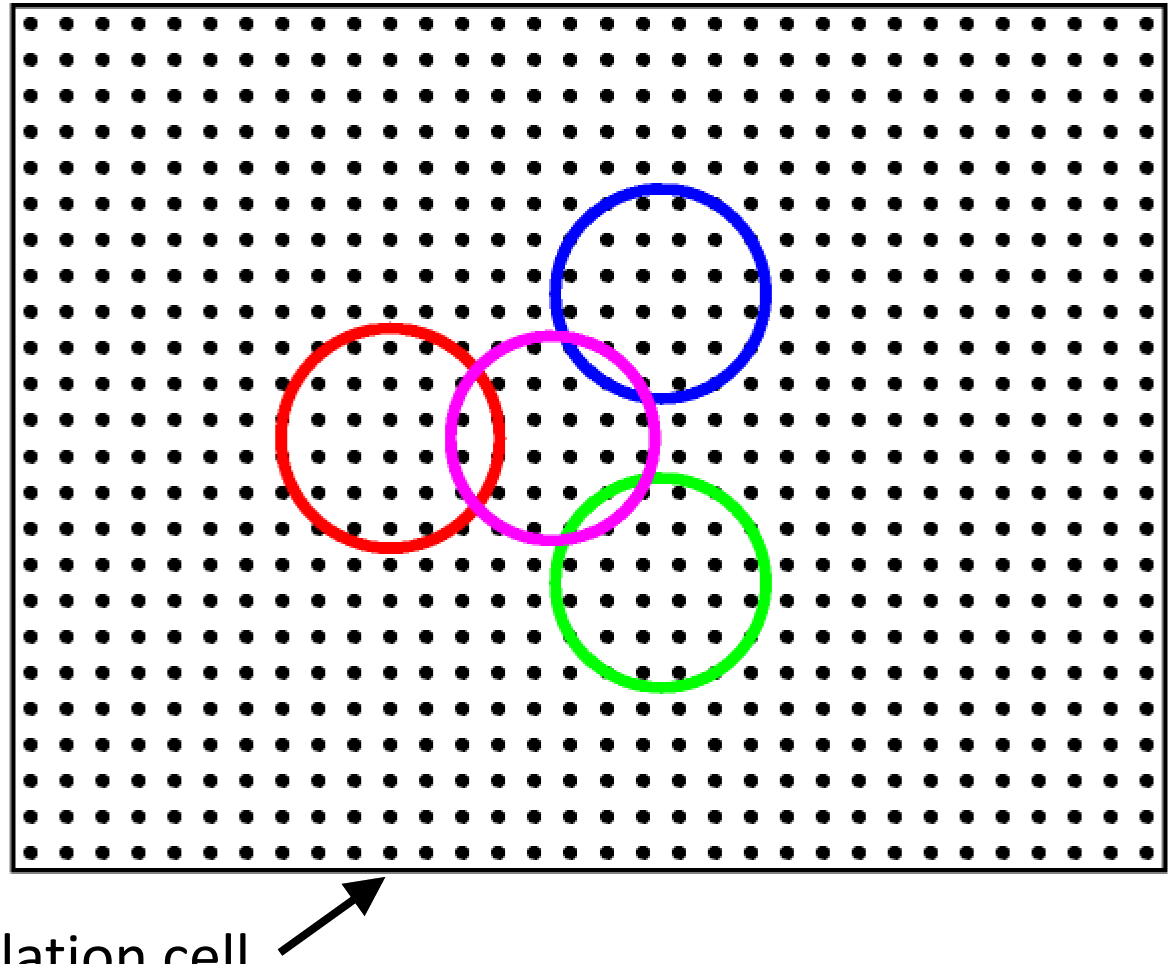


Basis set variational approaches:

C.-K. Skylaris, O. Dieguez, P. Haynes and M. C. Payne, *Phys. Rev. B* **66**, 073103 (2002).

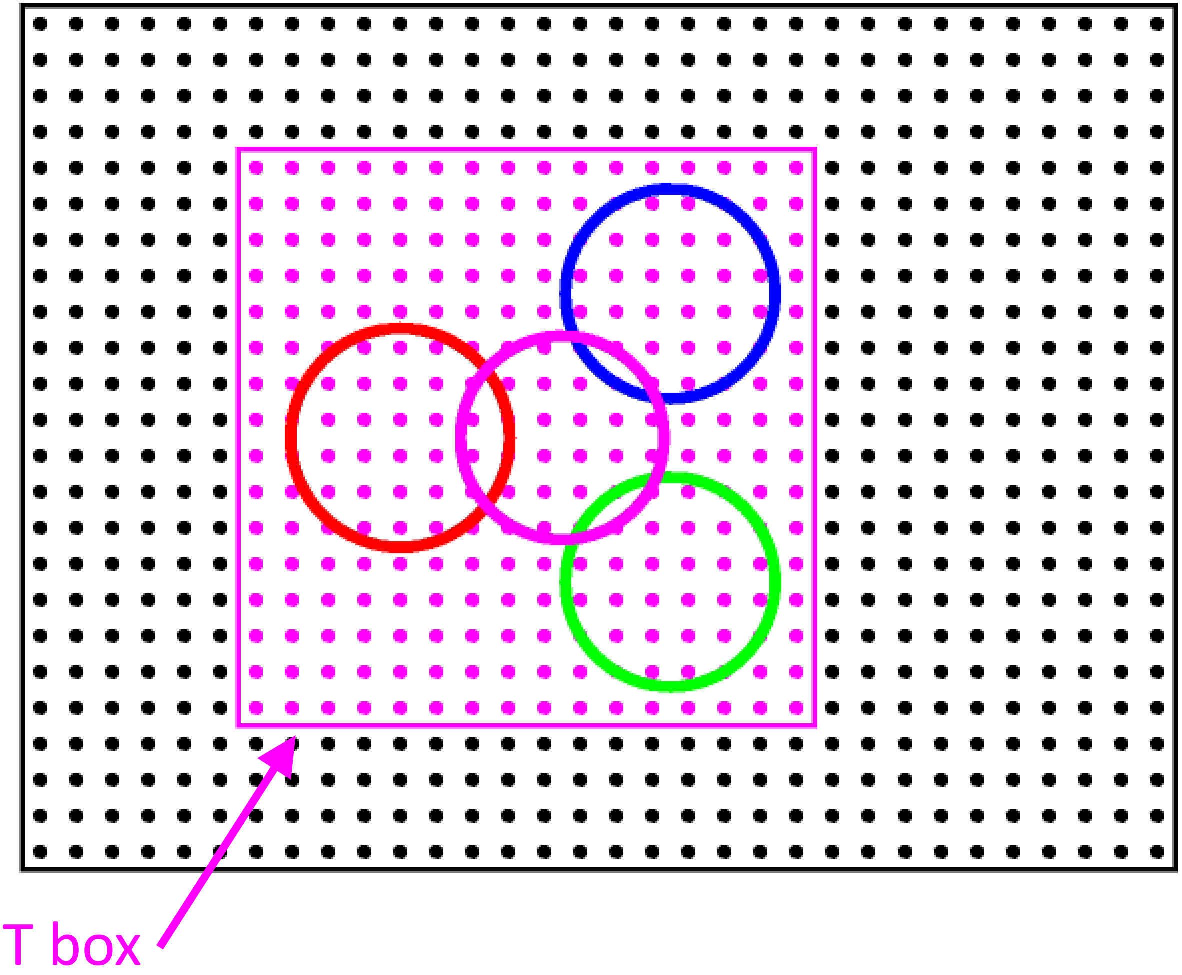
UNIVERSITY OF Southampton School of Chemistry

FFT box technique



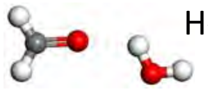
simulation cell

FFT box technique

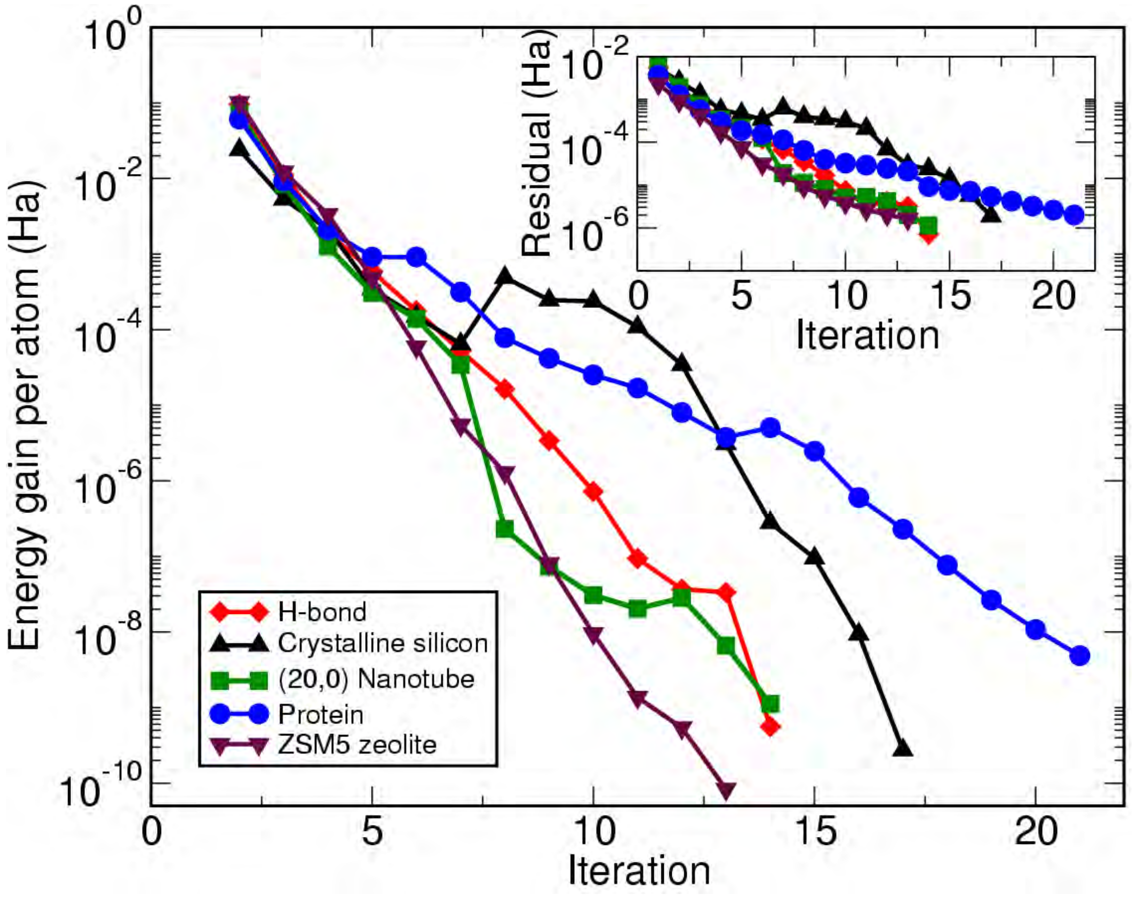


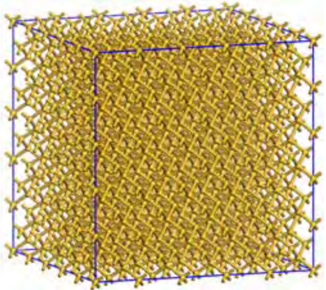
C.-K. Skylaris, A. A. Mostofi, P. D. Haynes, C. J. Pickard & M. C. Payne, *Comp. Phys. Comm.* **140**, 315 (2001)
A. A. Mostofi, C.-K. Skylaris, P. D. Haynes & M. C. Payne, *Comp. Phys. Comm.* **147**, 788 (2002)

True linear scaling: kinetic energy preconditioning

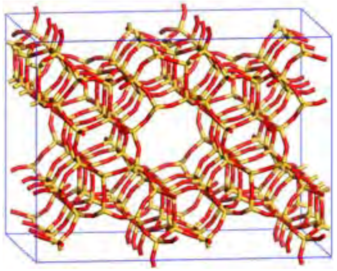


H-bond (7 atoms)

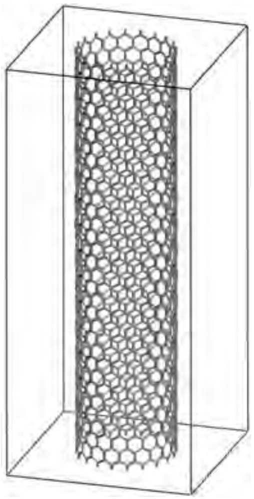




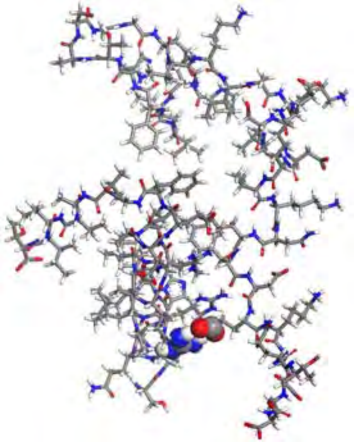
Crystalline silicon
(1000 atoms)



ZSM5 zeolite
(576 atoms)



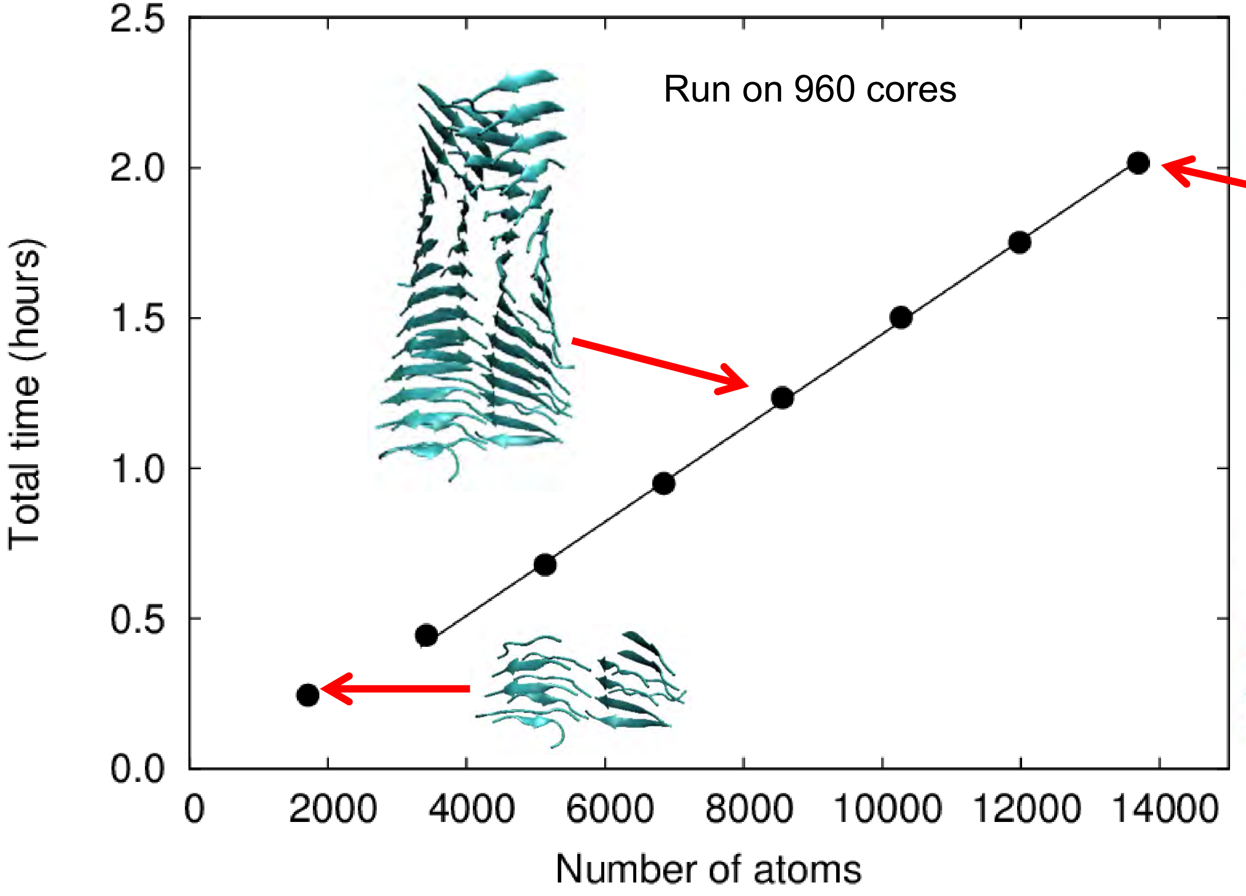
(20, 0) Nanotube
(1280 atoms)

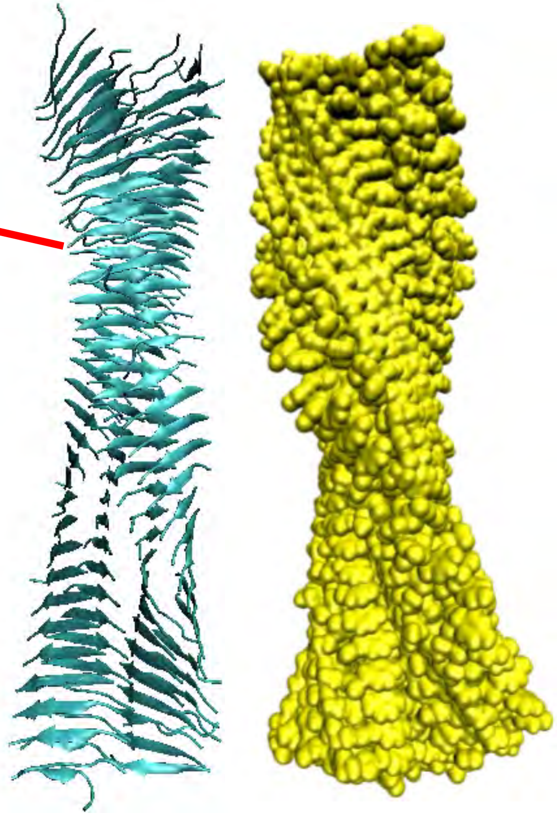


Protein (988 atoms)

Skylaris, Haynes, Mostofi & Payne, *J. Phys.: Condens. Matter* **17**, 5757 (2005)

Linear-scaling with the number of atoms
Example: ONETEP calculations on amyloid fibril proteins





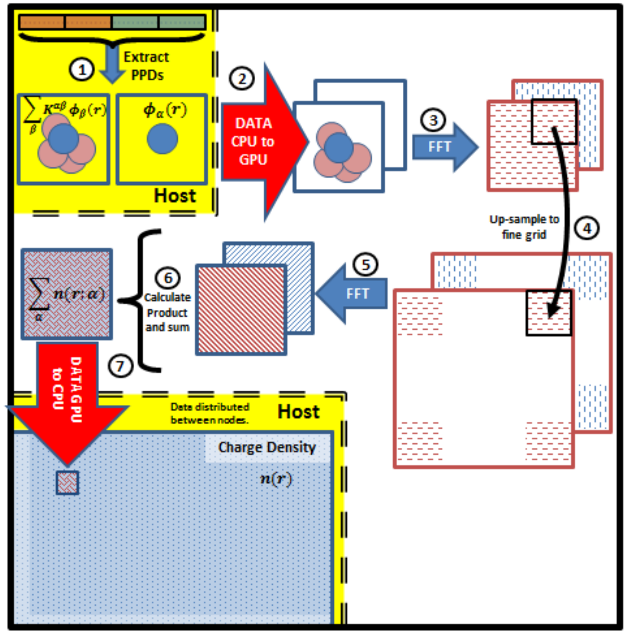
Electronic density from ONETEP calculation

Structures of the amyloid fibril kindly provided by the authors of
*J. T. Berryman, S. E. Radford and S. A. Harris, Biophysical Journal, **97** 1 (2009)*

Parallelisation

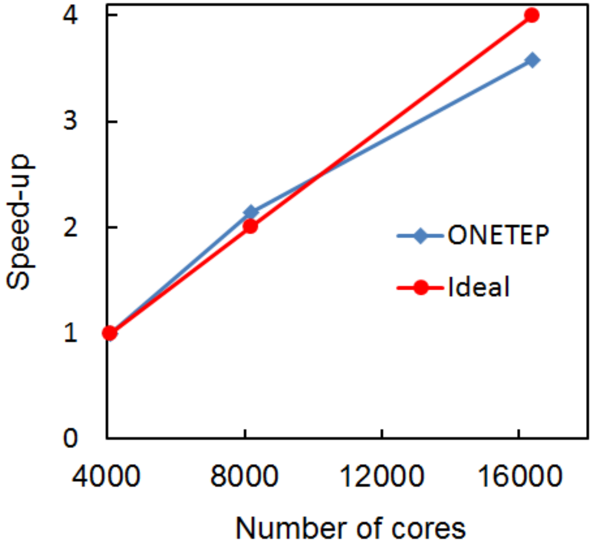
Graphical Processor Units (GPUs)

"Porting ONETEP to Graphical Processing Unit based Coprocessors. 1. FFT box Operations"
K. Wilkinson and C.-K. Skylaris, *J. Comp. Chem.* **34** (2013) 2446.



Combining MPI and OpenMP

"Hybrid MPI-OpenMP Parallelism in the ONETEP Linear-Scaling Electronic Structure Code: Application to the Delamination of Cellulose Nanofibrils"
K. Wilkinson, N. D. M Hine and C.-K. Skylaris, *J. Chem. Theor. Comput.* **10** (2014) 4782.



- Tests on"Blue Joule", 1.2 PetaFLOPS, IBM Blue Gene/Q
- 114,688 compute cores

UNIVERSITY OF Southampton
School of Chemistry

Compiling ONETEP

Simple multi-platform build system, needs:

- Fortran 2003/2008 compiler
- BLAS and LAPACK (or SCALAPACK) numerical libraries
- FFT library: vendor-supplied or FFTw
  - www.fftw.org
- MPI library for parallel version
- OpenMP

Running ONETEP

- Parallel computer
  - Minimum 1 GB per processor (core)
  - Typically distribute 10-100 atoms per processor
  - Cross-over >100 atoms

- Prepare input file: free format
  - Documentation at www.onetep.org

- Supply pseudopotential files (`.recpot` format)

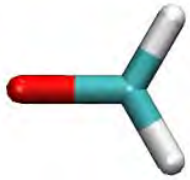
- We suggest that instead of running ONETEP directly, you run the **onetep_launcher** script which sets up correctly the runtime environment and does important checks

Input file

- Keywords of different types:
  - Integer
  - Boolean
  - String
  - Real
  - Physical (real + unit)
  - Block data e.g. atomic positions, delimited by **`%block`** and **`%endblock`**
- Atomic units by default (hartree and bohr)

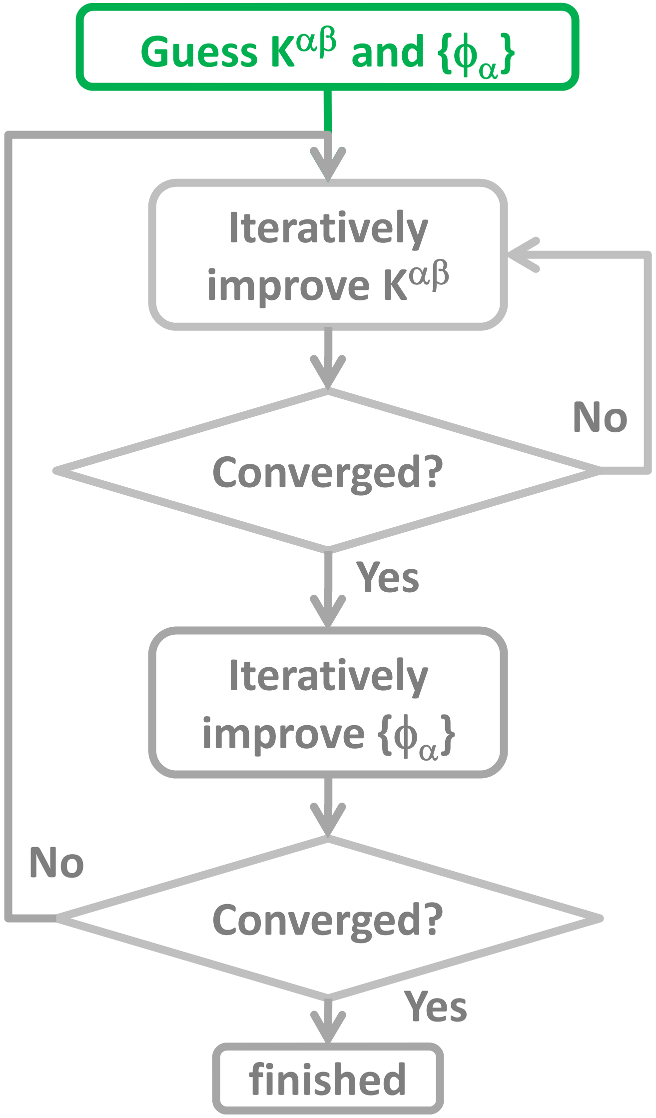
Example input file: formaldehyde

```
! Example input file for the ONETEP program
! Formaldehyde molecule

cutoff_energy 600 eV

%block   lattice_cart
  48.00      0.00      0.00
   0.00     48.00      0.00
   0.00      0.00     48.00
%endblock  lattice_cart

%block  positions_abs
O      24.887507   23.896975   22.647313
C      27.731659   23.667449   22.643306
H      28.655157   21.721170   22.637547
H      28.955467   25.440371   22.646039
%endblock  positions_abs

%block species
O    O    8    4    6.5
C    C    6    4    6.5
H    H    1    1    6.5
%endblock species

%block species_pot
O    oxygen.recpot
C    carbon.recpot
H    hydrogen.recpot
%endblock species_pot
```



ONETEP calculation outline

- Initialisation phase:
  - Construct initial NGWFs (STOs or PAOs)
  - Construct initial charge density (atomic superposition) and effective potential
  - Construct initial Hamiltonian
  - Obtain initial density kernel using iterative approach



Guess $K^{\alpha\beta}$ and $\{\phi_\alpha\}$

Iteratively improve $K^{\alpha\beta}$

Converged?

No

Yes

Iteratively improve $\{\phi_\alpha\}$

Converged?

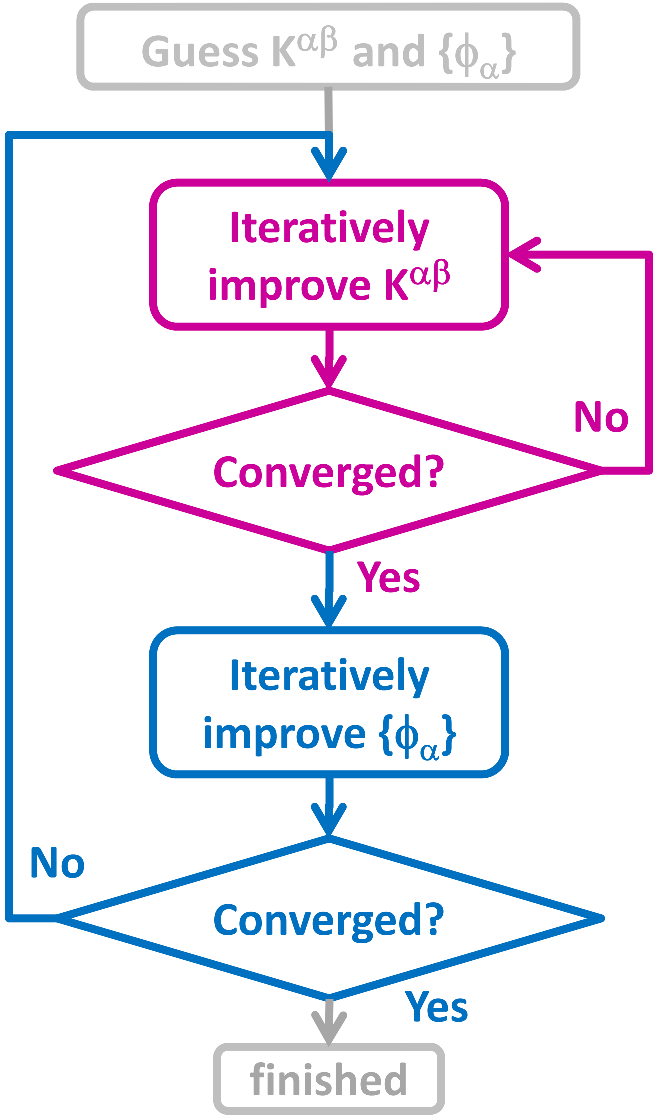
No

Yes

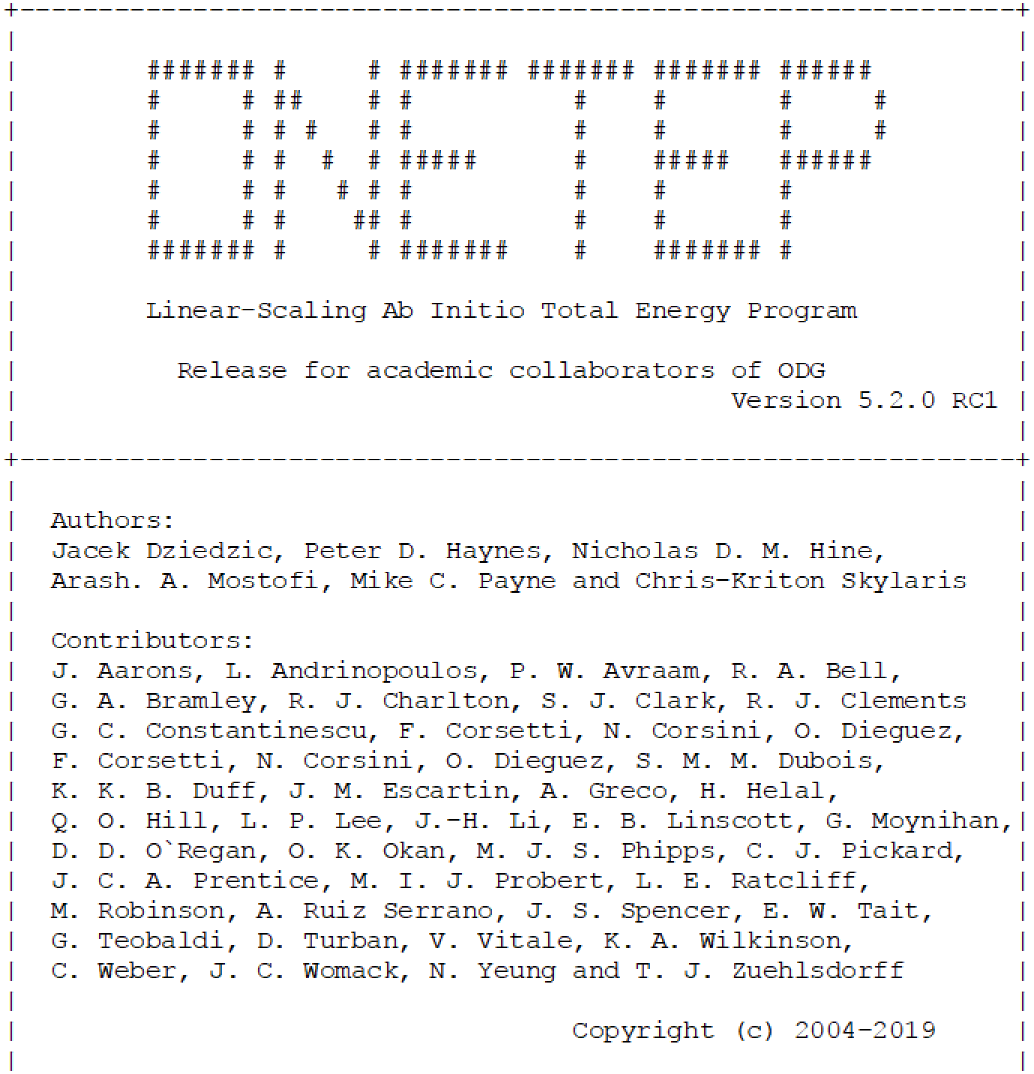
finished

ONETEP calculation outline continued

- Main optimisation phase:
  - Combination of nested self-consistent loops
  - Outer loop optimises the NGWFs (density kernel fixed)
  - Inner loop optimises the density kernel (NGWFs fixed) using Density Matrix Minimisation approaches

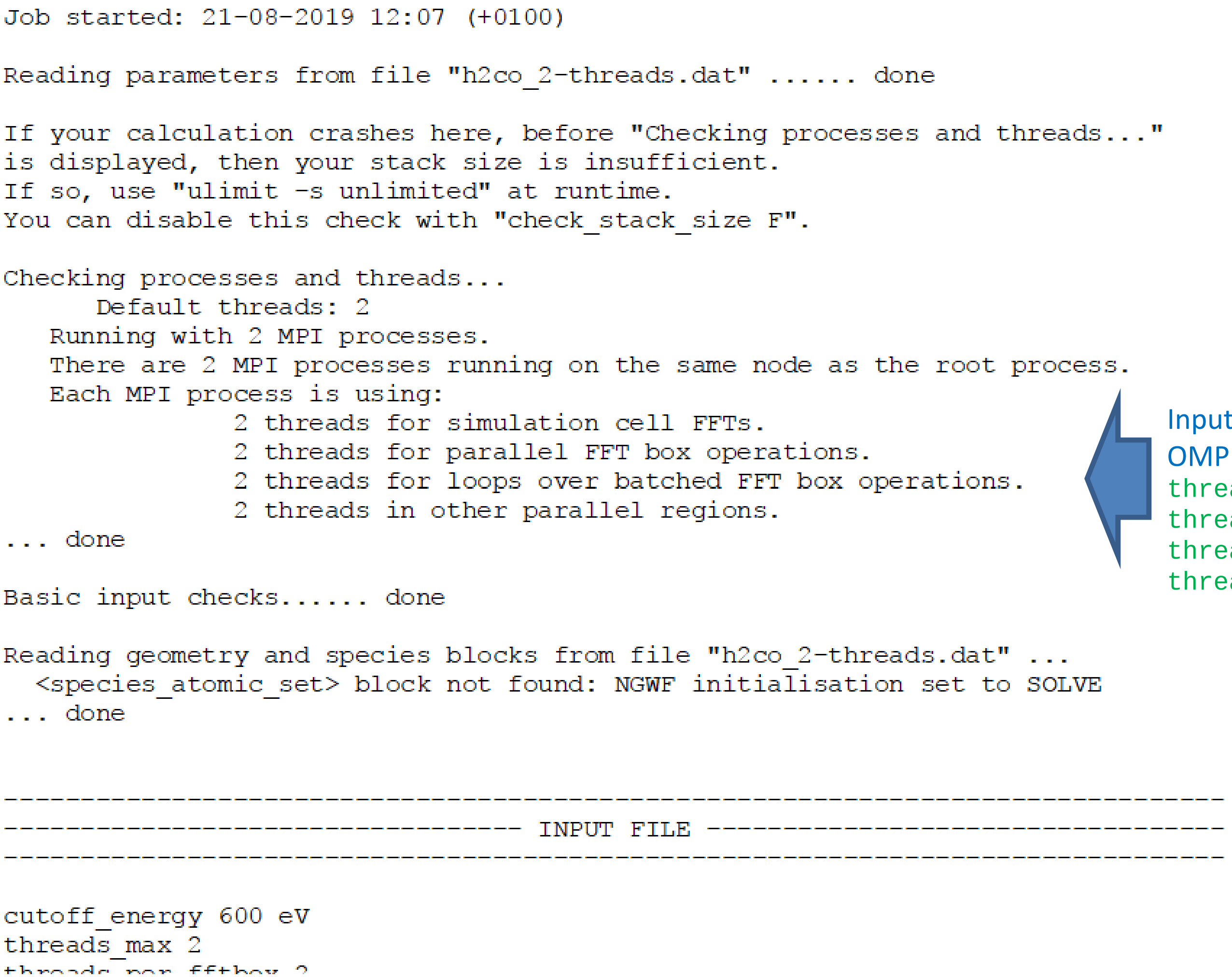


Example output file: formaldehyde



```
+---------------------------------------------------------------+
|                                                               |
|        ####### #     # ####### ####### ####### ######         |
|        #     # ##    # #          #    #       #     #        |
|        #     # # #   # #          #    #       #     #        |
|        #     # #  #  # #####      #    #####   ######         |
|        #     # #   # # #          #    #       #              |
|        #     # #    ## #          #    #       #              |
|        ####### #     # #######    #    ####### #              |
|                                                               |
|        Linear-Scaling Ab Initio Total Energy Program          |
|                                                               |
|          Release for academic collaborators of ODG            |
|                                             Version 5.2.0 RC1 |
|                                                               |
+---------------------------------------------------------------+
|                                                               |
|  Authors:                                                     |
|  Jacek Dziedzic, Peter D. Haynes, Nicholas D. M. Hine,        |
|  Arash. A. Mostofi, Mike C. Payne and Chris-Kriton Skylaris   |
|                                                               |
|  Contributors:                                                |
|  J. Aarons, L. Andrinopoulos, P. W. Avraam, R. A. Bell,       |
|  G. A. Bramley, R. J. Charlton, S. J. Clark, R. J. Clements   |
|  G. C. Constantinescu, F. Corsetti, N. Corsini, O. Dieguez,   |
|  F. Corsetti, N. Corsini, O. Dieguez, S. M. M. Dubois,        |
|  K. K. B. Duff, J. M. Escartin, A. Greco, H. Helal,           |
|  Q. O. Hill, L. P. Lee, J.-H. Li, E. B. Linscott, G. Moynihan,|
|  D. D. O`Regan, O. K. Okan, M. J. S. Phipps, C. J. Pickard,   |
|  J. C. A. Prentice, M. I. J. Probert, L. E. Ratcliff,         |
|  M. Robinson, A. Ruiz Serrano, J. S. Spencer, E. W. Tait,     |
|  G. Teobaldi, D. Turban, V. Vitale, K. A. Wilkinson,          |
|  C. Weber, J. C. Womack, N. Yeung and T. J. Zuehlsdorff       |
|                                                               |
|                                   Copyright (c) 2004-2019     |
|                                                               |
```

- Parallel strategy, calculation parameters

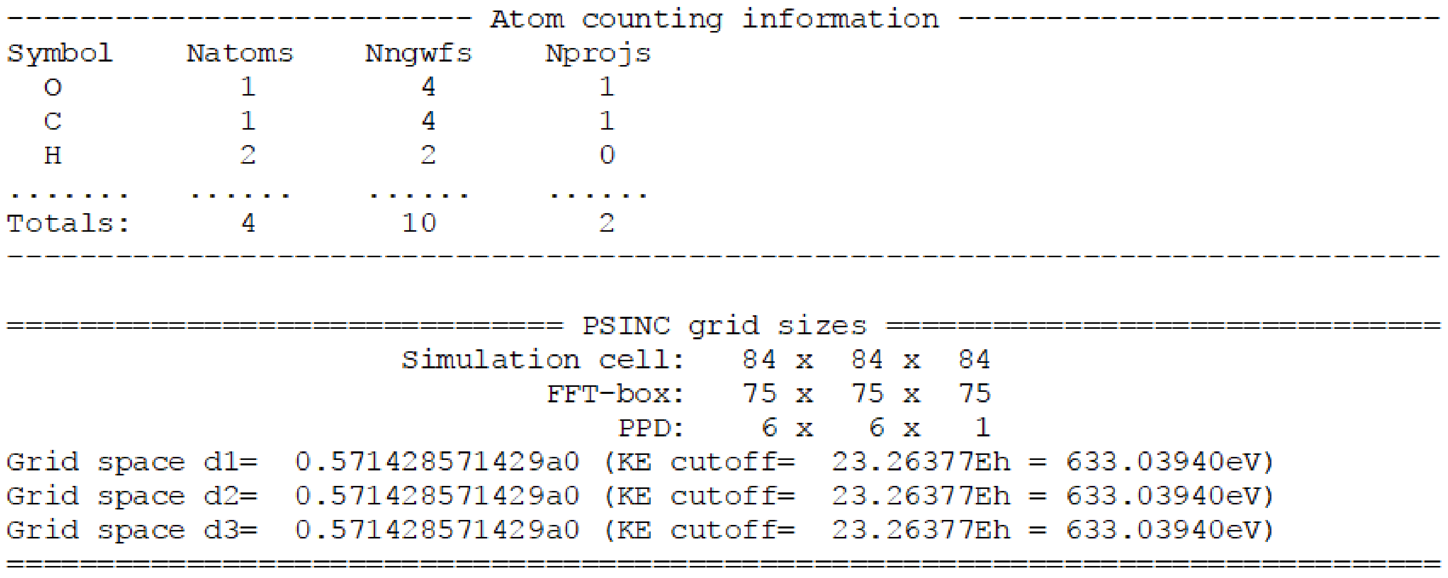


```
Job started: 21-08-2019 12:07 (+0100)

Reading parameters from file "h2co_2-threads.dat" ...... done

If your calculation crashes here, before "Checking processes and threads..."
is displayed, then your stack size is insufficient.
If so, use "ulimit -s unlimited" at runtime.
You can disable this check with "check_stack_size F".

Checking processes and threads...
      Default threads: 2
   Running with 2 MPI processes.
   There are 2 MPI processes running on the same node as the root process.
   Each MPI process is using:
               2 threads for simulation cell FFTs.
               2 threads for parallel FFT box operations.
               2 threads for loops over batched FFT box operations.
               2 threads in other parallel regions.
... done

Basic input checks...... done

Reading geometry and species blocks from file "h2co_2-threads.dat" ...
  <species_atomic_set> block not found: NGWF initialisation set to SOLVE
... done


--------------------------------------------------------------------------------
---------------------------------- INPUT FILE ----------------------------------
--------------------------------------------------------------------------------

cutoff_energy 600 eV
threads_max 2
threads_per_fftbox 2
```

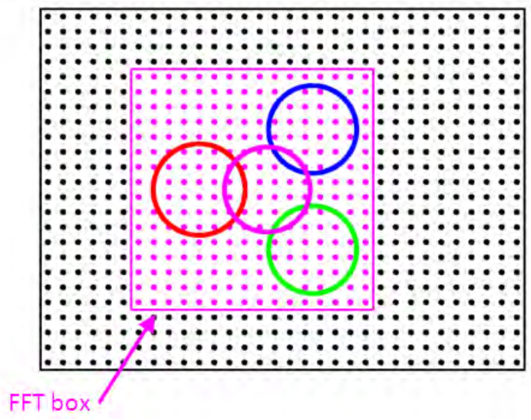
Input file parameters for OMP threads

```
threads_max 2
threads_per_fftbox 2
threads_num_fftboxes 2
threads_per_cellfft 2
```

Example output file: formaldehyde



```
-------------------------- Atom counting information --------------------------
Symbol     Natoms     Nngwfs     Nprojs
  O           1          4          1
  C           1          4          1
  H           2          2          0
.......    ......     ......     ......
Totals:       4         10          2
-------------------------------------------------------------------------------

=============================== PSINC grid sizes ===============================
                      Simulation cell:   84 x  84 x  84
                              FFT-box:   75 x  75 x  75
                                  PPD:    6 x   6 x   1
Grid space d1=  0.571428571429a0 (KE cutoff=  23.26377Eh = 633.03940eV)
Grid space d2=  0.571428571429a0 (KE cutoff=  23.26377Eh = 633.03940eV)
Grid space d3=  0.571428571429a0 (KE cutoff=  23.26377Eh = 633.03940eV)
================================================================================
```



- Simulation cell, FFT box, grid spacing

Example output file: formaldehyde

- NGWF initialisation to pseudo-atomic orbitals (PAOs)

```
Atom SCF Calculation for O  : Z (AE atom) =   8 : Z (PS atom) =   6
Config String:
Orbitals (num,occ):  2     2.00 4.00
Orbitals   (num,l):  2        0    1
Atom SCF converged after  33 iterations to a total energy of    -15.71589674

Atom SCF Calculation for C  : Z (AE atom) =   6 : Z (PS atom) =   4
Config String:
Orbitals (num,occ):  2     2.00 2.00
Orbitals   (num,l):  2        0    1
Atom SCF converged after  33 iterations to a total energy of     -5.33577996

Atom SCF Calculation for H  : Z (AE atom) =   1 : Z (PS atom) =   1
Config String:
Orbitals (num,occ):  1     1.00
Orbitals   (num,l):  1        0
Atom SCF converged after  30 iterations to a total energy of     -0.43623914
NGWF initialisation ...... done
```

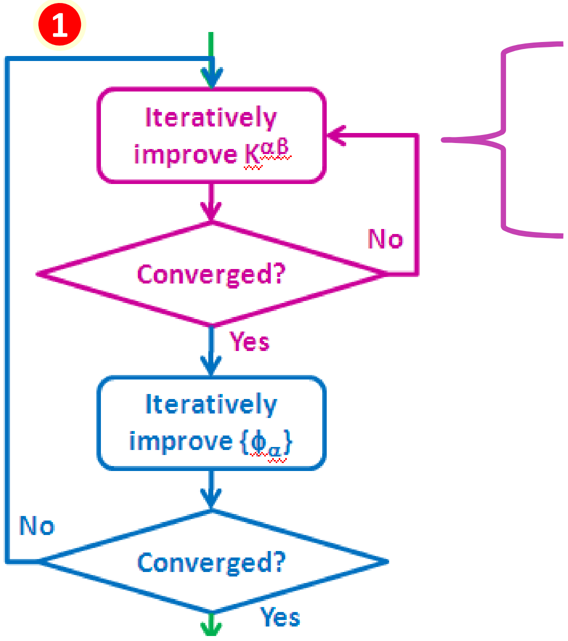
Example output file: formaldehyde

- Initialise and refine density kernel

```
Up spin density kernel initialisation ...... done

Writing density kernel to file "h2co.dkn" ... done

>>> Optimising kernel for current NGWFs:
   iter  |      energy (Eh)       | rms gradient |  commutator   |   dE (Eh)
    1        -2.230601586111E+01     5.3808E-02     4.4840E-02
    2        -2.233073648394E+01     3.8281E-02     3.1901E-02    -2.4721E-02
    3        -2.235530086009E+01     1.1889E-02     9.9074E-03    -2.4564E-02
    4        -2.235864046835E+01     6.8691E-03     5.7242E-03    -3.3396E-03
    5        -2.235945547230E+01     3.9727E-03     3.3106E-03    -8.1500E-04
Finished density kernel iterations (   5)

Writing density kernel to file "h2co.dkn" ... done


>>> Density kernel optimised for the current NGWF basis:
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     Total energy                = -2.23594554723032E+01  Eh
     Estimated bandgap           =   5.4645E-02  Eh
     RMS occupancy error         =   5.0205E-13
     [H,K] commutator            =   3.3106E-03
     Occupancy bounds            = [  0.000:  1.000]
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

Example output file: formaldehyde

- SCF iterations



①

```
################################################################################
########################### NGWF CG iteration  003 #############################
################################################################################

>>> Checking for convergence of NGWFs:   NOT CONVERGED
  ========================================================================
         NGWF RMS gradient =  5.2368E-04 >  2.0000E-06  | above tolerance
  ========================================================================

>>> Improving NGWFs using line search:
  ========================================================================
   Predicted total energy    =      -2.26058828338753E+01 Eh
   Predicted gain in energy  =      -6.90722410908862E-03 Eh
  ========================================================================

Writing NGWFs to file "h2co.tightbox_ngwfs"....  done

>>> Optimising kernel for current NGWFs:
   iter  |      energy (Eh)       | rms gradient |  commutator   |   dE (Eh)
    1        -2.260585039399E+01     4.3060E-03     3.5884E-03
    2        -2.260614202529E+01     8.8690E-04     7.3908E-04    -2.9163E-04
    3        -2.260616377801E+01     2.2375E-04     1.8646E-04    -2.1753E-05
    4        -2.260616700389E+01     5.2027E-05     4.3355E-05    -3.2259E-06
    5        -2.260616709879E+01     1.7239E-05     1.4365E-05    -9.4900E-08
Finished density kernel iterations (   5)

Writing density kernel to file "h2co.dkn" ... done


>>> Density kernel optimised for the current NGWF basis:
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
     Total energy                =  -2.26061670987944E+01   Eh
     Estimated bandgap           =   7.1917E-02  Eh
     RMS occupancy error         =   8.2047E-14
     [H,K] commutator            =   1.4365E-05
     Occupancy bounds            = [  0.000:  1.000]
     ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~



################################################################################
########################### NGWF CG iteration  004 #############################
################################################################################
```

Example output file: formaldehyde

- End of SCF calculation

```
>>> Checking for convergence of NGWFs:
           .........................................................
           |         *** NGWF optimisation converged ***           |
           | RMS NGWF gradient =    0.00000126658643               |
           | Criteria satisfied:                                   |
           | -> RMS NGWF gradient lower than set threshold.        |
           =========================================================

================================================================================
           ---------------- ENERGY COMPONENTS (Eh) ----------------
           | Kinetic                    :       14.91662629423981 |
           | Pseudopotential (local)    :      -75.55456193656947 |
           | Pseudo (non-coul chg cor)  :        0.00000000000000
           | Pseudopotential (non-local):        3.08867264674192 |
           | Hartree                    :       29.55304098295274 |
           | Exchange-correlation       :       -5.51703677107650 |
           | Ewald                      :       10.90369328705708 |
           | Total                      :      -22.60956549665441 |
           ---------------------------------------------------------
           Integrated density           :       11.99999999999941
================================================================================


                        <<<<< CALCULATION SUMMARY >>>>>
|ITER|    RMS GRADIENT    |     TOTAL ENERGY     |  step    |    Epredicted
    0     0.00339284991963    -22.35945547230321   0.578496    -22.56473206954799
    1     0.00109819484603    -22.56842090396507   0.774264    -22.59768874673372
    2     0.00052368029265    -22.59897560976616   0.779737    -22.60588283387525
    3     0.00030645887473    -22.60616709879438   0.778478    -22.60852120360797
    4     0.00017075786499    -22.60861681158951   0.618320    -22.60917256889675
    5     0.00010955112072    -22.60917385344275   0.827718    -22.60948077698849
    6     0.00006504197690    -22.60948592159049   0.577777    -22.60956236994081
    7     0.00000603288289    -22.60956529599148  -0.003358    -22.60956529222391
    8     0.00000662328177    -22.60956547322916   0.099059    -22.60956549599337
    9     0.00000126658643    -22.60956549665442  <-- CG
```

Summary of Functionality

Total energies

- Various exchange-correlation functionals:
    - LDA (Ceperley-Alder-Perdew-Zunger, Vosko-Wilk-Nusair, PW92)
    - GGA (PW91, PBE, revPBE, RPBE, BLYP, XLYP, WC)
    - Non-local exchange-correlation functionals for dispersion (e.g. Langreth and Lundqvist, VV10)
    - Meta-GGAs, e.g. PKZB, B97m-v
    - Hartree-Fock exchange and hybrid functionals
- Spin polarisation
- DFT+D (empirical dispersion)
- DFT+U
- Charge-constrained DFT
- Finite temperature DFT for metallic systems (Ensemble DFT, AQUA-FOE)
- Embedding one level of theory inside another (e.g. mGGA in LDA), using EMFT

Summary of Functionality

Excited states

- Conduction NGWFs
- LR-TDDFT
- LR-TDDFT with the Tamm-Dancoff approximation (TDA)

Boundary conditions

- Periodic boundary conditions
- Open boundary conditions (Cut-off Coulomb, Martyna-Tuckerman or real-space open boundaries)
- Implicit solvent model
- Electrostatic embedding

Core electrons

- Norm conserving pseudopotentials
- Projector Augmented wave (PAW) approach (all electron)

Summary of Functionality

Atomic forces

- Geometry optimisation
- Transition state search
- *Ab initio* molecular dynamics

Visualisation

- NGWFs
- Molecular Orbitals
- Density and potentials
- Electron Localisation Function (ELF)

Atomic orbital basis set option

- Instead of NGWFs construct and use SZ, SZP, DZ, DZP, etc atomic orbital basis sets

Summary of Functionality

Electronic properties

- Density of states (DOS), local DOS, projected DOS
- Atomic charges (Mulliken, IH, Density derived electrostatic and chemical (DDEC))
- Dipoles and higher moments
- Optimisation of separate NGWF set for accurate conduction bands and optical absorption spectra
- Natural Bond Orbital (NBO) analysis (Natural Atomic Orbitals in ONETEP and interface to NBO5.9 program)
- Energy Decomposition Analysis (EDA)
- Electron transport
- Distributed multipole analysis

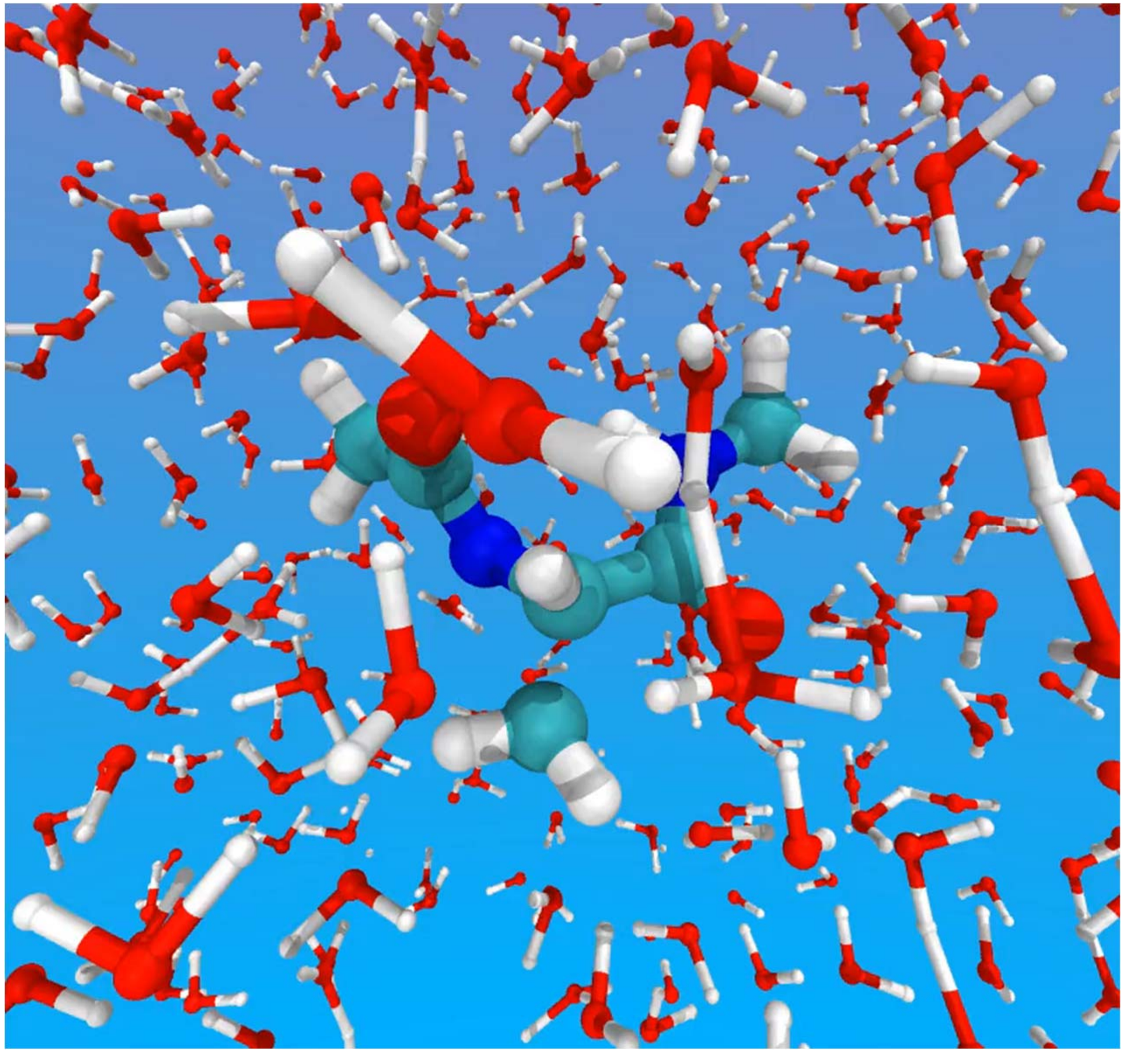
Ab initio Molecular Dynamics with ONETEP

- Peptide in water
- 1522 atoms
- 2256 cores
- Time step of 0.5 fs

Fully anharmonic IR spectra of molecules in solution

V. Vitale, J. Dziedzic, S. M.-M. Dubois, H. Fangohr, and C.-K. Skylaris, *J. Chem. Theor. Comput.* **11** 2231 (2015).

V. Vitale, J. Dziedzic, A. Albaugh, A. N. M. Niklasson, T. Head-Gordon, and C.-K. Skylaris. *J. Chem. Phys.* **146** 124115 (2017)



UNIVERSITY OF Southampton
School of Chemistry

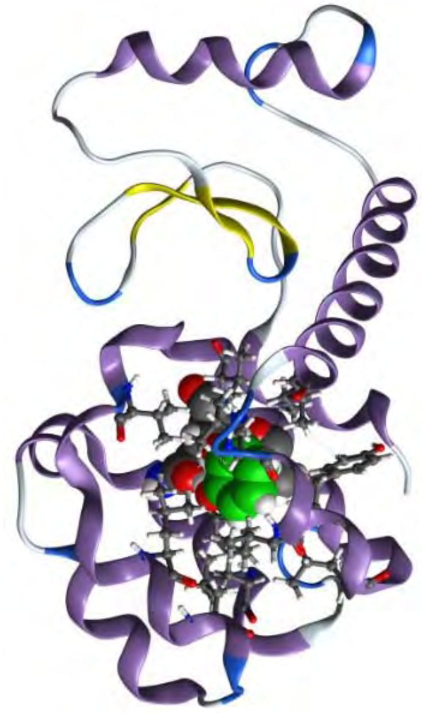
Applications to drug design: T4 Lysozyme L99A/M102Q protein

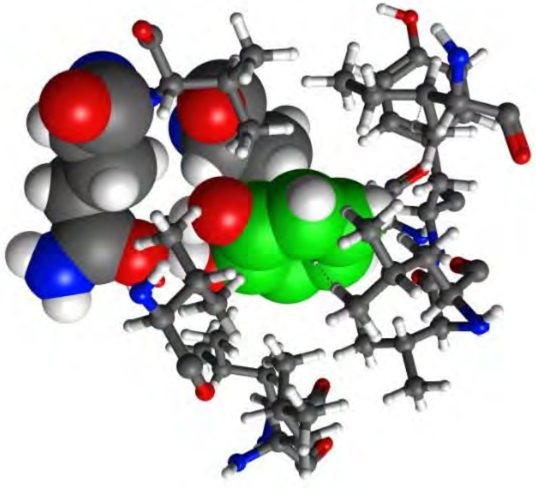
- 2616 atoms
- Polar binding site

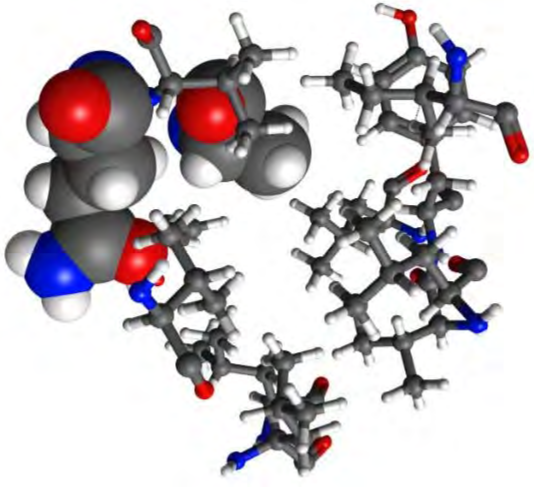
S. J. Fox, J. Dziedzic, T. Fox, C. S. Tautermann, and C.-K. Skylaris,
Proteins **82** (2014) 3335-3346

C. Cave-Ayland, C.-K. Skylaris, and J. W. Essex,
J. Phys. Chem. B **119** (2015) 1017

S. Mansour, A. S. Tocheva, C. Cave-Ayland, M. M. Machelett, B. Sander, N. M. Lissin, P. E. Molloy, M. S. Baird, G. Stübs, N. W. J. Schröder, R. R. Schumann, J. Rademann, A. D. Postle, B. K. Jakobsen, B. G. Marshall, R. Gosain, P. T. Elkington, T. Elliott, C.-K. Skylaris, J. W. Essex, I. Tews, and S. D. Gadola.
PNAS. **113** (2016) E1266

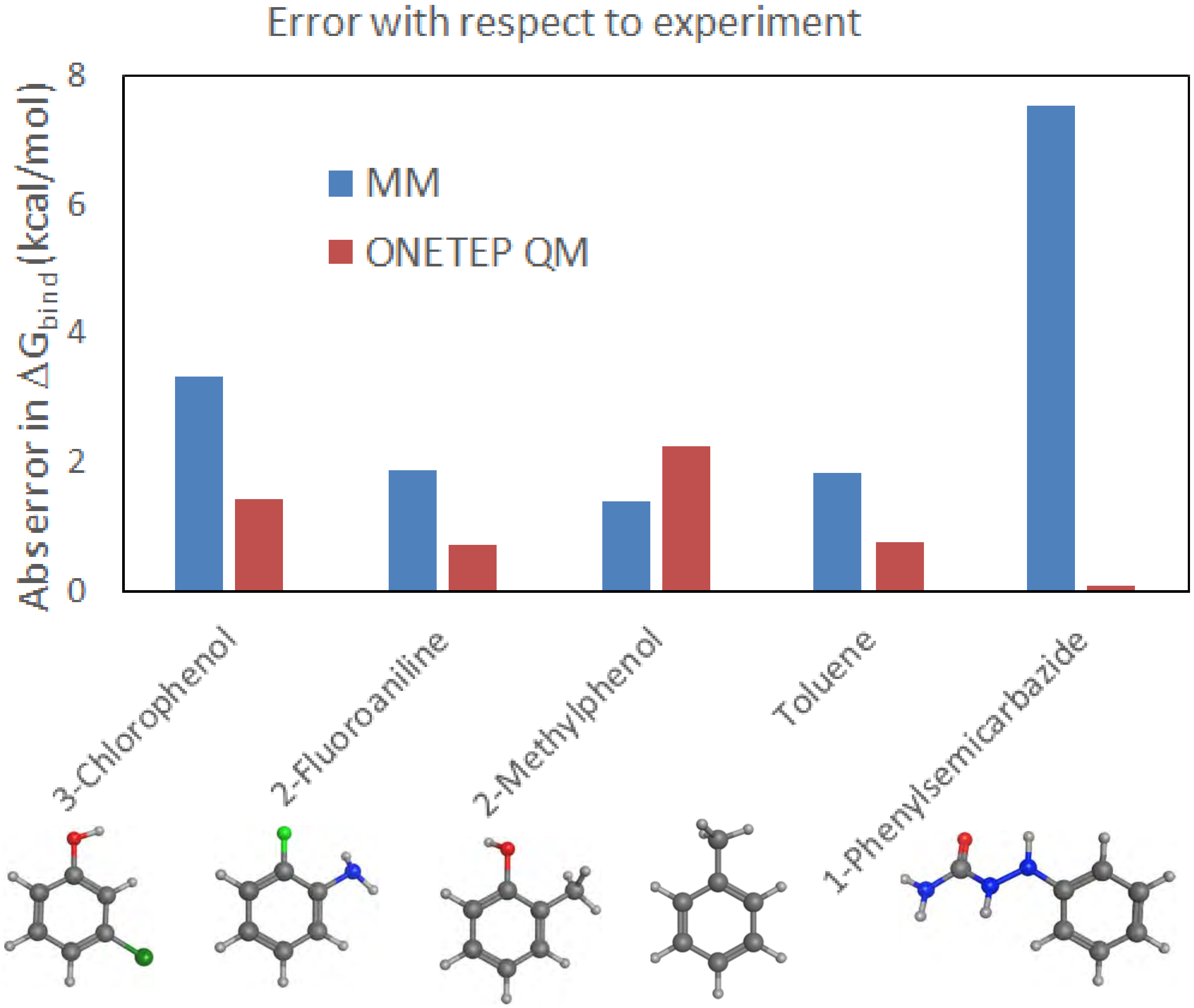






Complex of T4 Lysozyme L99A/M102Q and catechol

Free energies of binding: T4 Lysozyme L99A/M102Q



UNIVERSITY OF Southampton
School of Chemistry

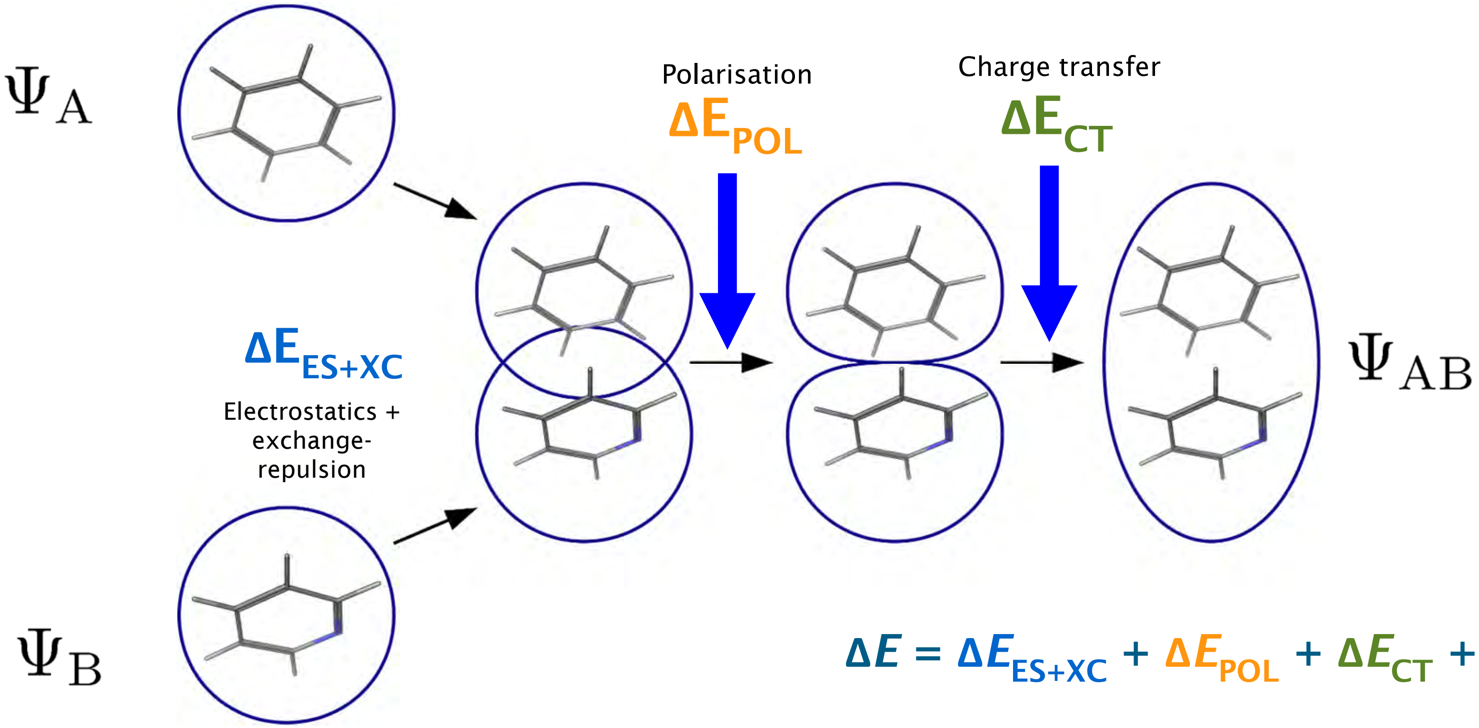
Energy decomposition analysis (EDA) in ONETEP

M. J. S. Phipps, T. Fox, C. S. Tautermann and C.-K. Skylaris, Chem. Soc. Rev. **44**(2015) 3177-3211.
M. J. S. Phipps, T. Fox, C. S. Tautermann, and C.-K. Skylaris, *J. Chem. Theory Comput.*, 2016 **DOI:** 10.1021/acs.jctc.6b00272

Decompose any host-guest binding energy to chemically relevant components

Can be used to

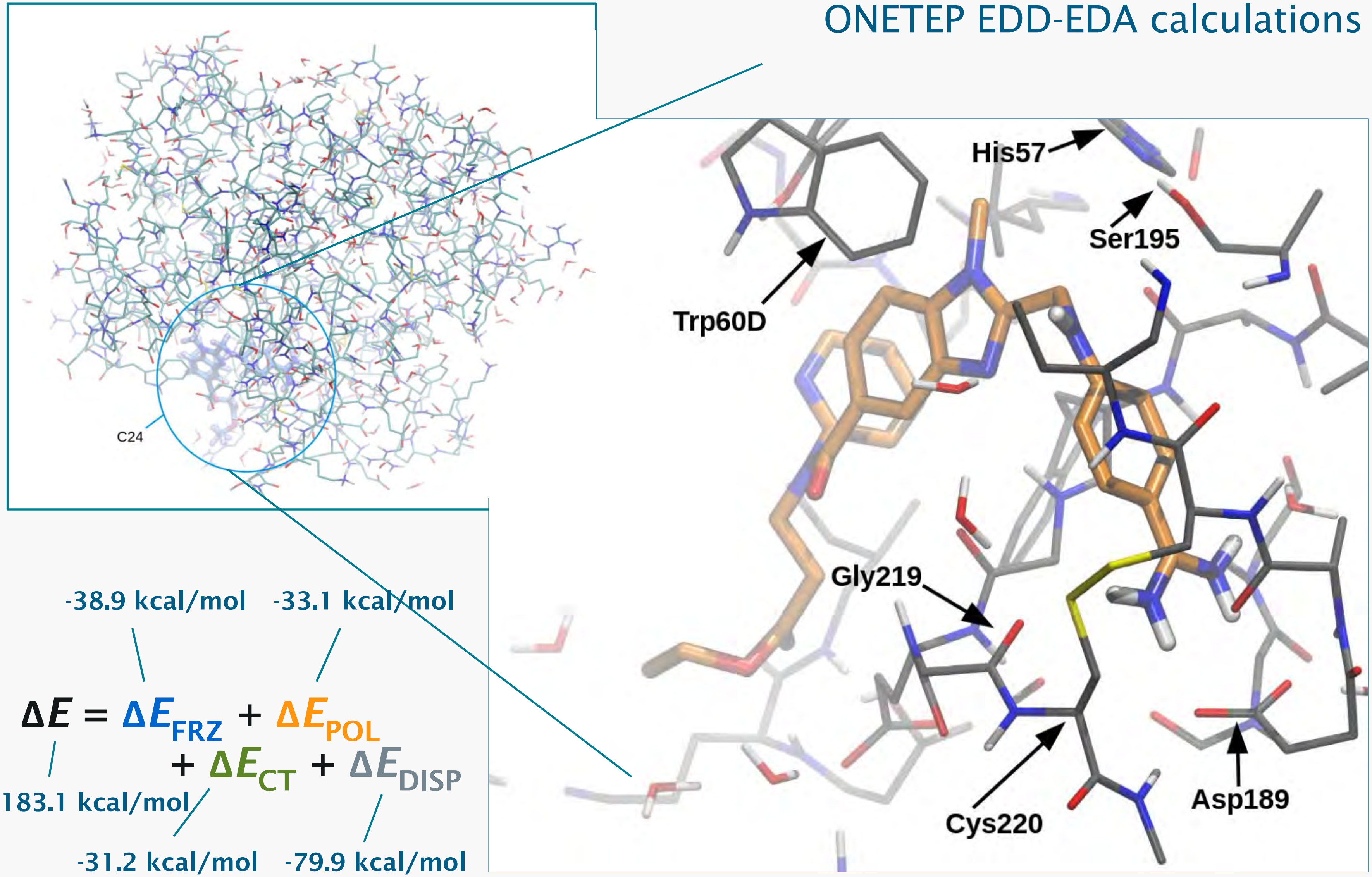
- Understand the nature of a particular interaction
- Inform the design of new ligands
- Optimise more approximate simulation methods (e.g. new MM approaches)



$$\Delta E = \Delta E_{ES+XC} + \Delta E_{POL} + \Delta E_{CT} + \ldots$$

Thrombin-inhibitor complex (4975 atoms)

ONETEP EDD-EDA calculations



-38.9 kcal/mol -33.1 kcal/mol

$$\Delta E = \Delta E_{FRZ} + \Delta E_{POL} + \Delta E_{CT} + \Delta E_{DISP}$$

-183.1 kcal/mol

-31.2 kcal/mol -79.9 kcal/mol

NANO LETTERS

Nano Lett., **2017**, *17* (7), pp 4003

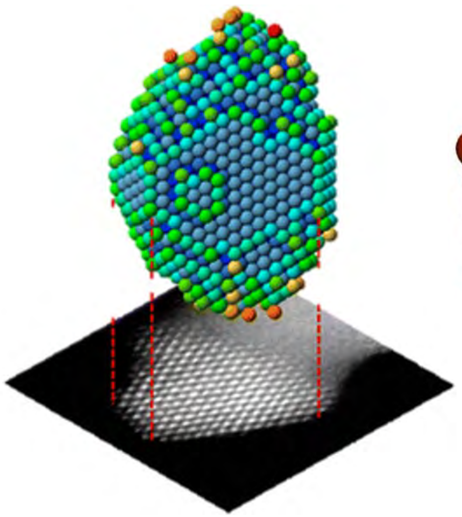
Letter

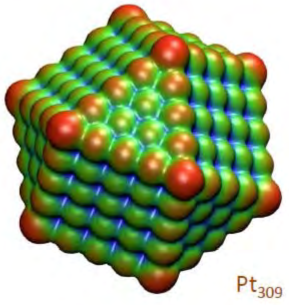
pubs.acs.org/NanoLett

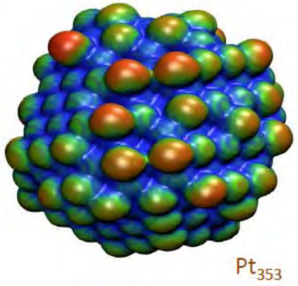
Predicting the Oxygen-Binding Properties of Platinum Nanoparticle Ensembles by Combining High-Precision Electron Microscopy and Density Functional Theory

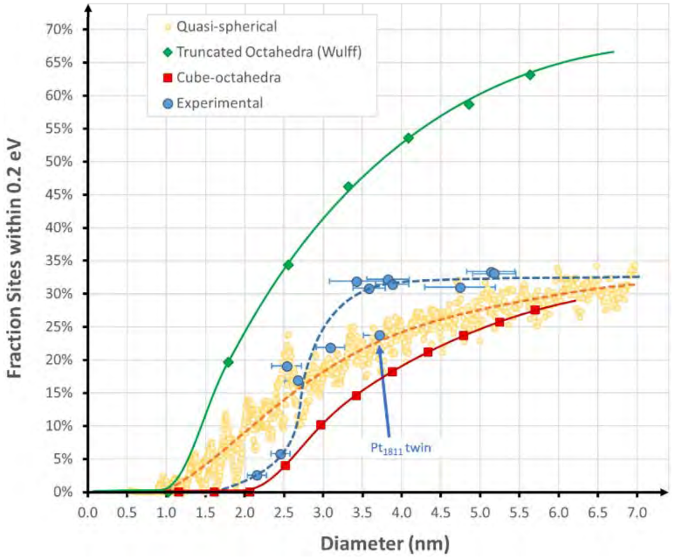
Jolyon Aarons,[†] Lewys Jones,[*,‡] Aakash Varambhia,[‡] Katherine E. MacArthur,[§] Dogan Ozkaya,[‖] Misbah Sarwar,[‖] Chris-Kriton Skylaris,[†] and Peter D. Nellist[‡]

[†] Department of Chemistry, University of Southampton, Highfield SO17 1BJ, United Kingdom
[‡] Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, United Kingdom
[§] Ernst Ruska Center for Microscopy and Spectroscopy with Electrons, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
[‖] Johnson Matthey Technology Centre, Sonning Common, Reading, RG4 9NH, United Kingdom




Pt$_{309}$


Pt$_{353}$

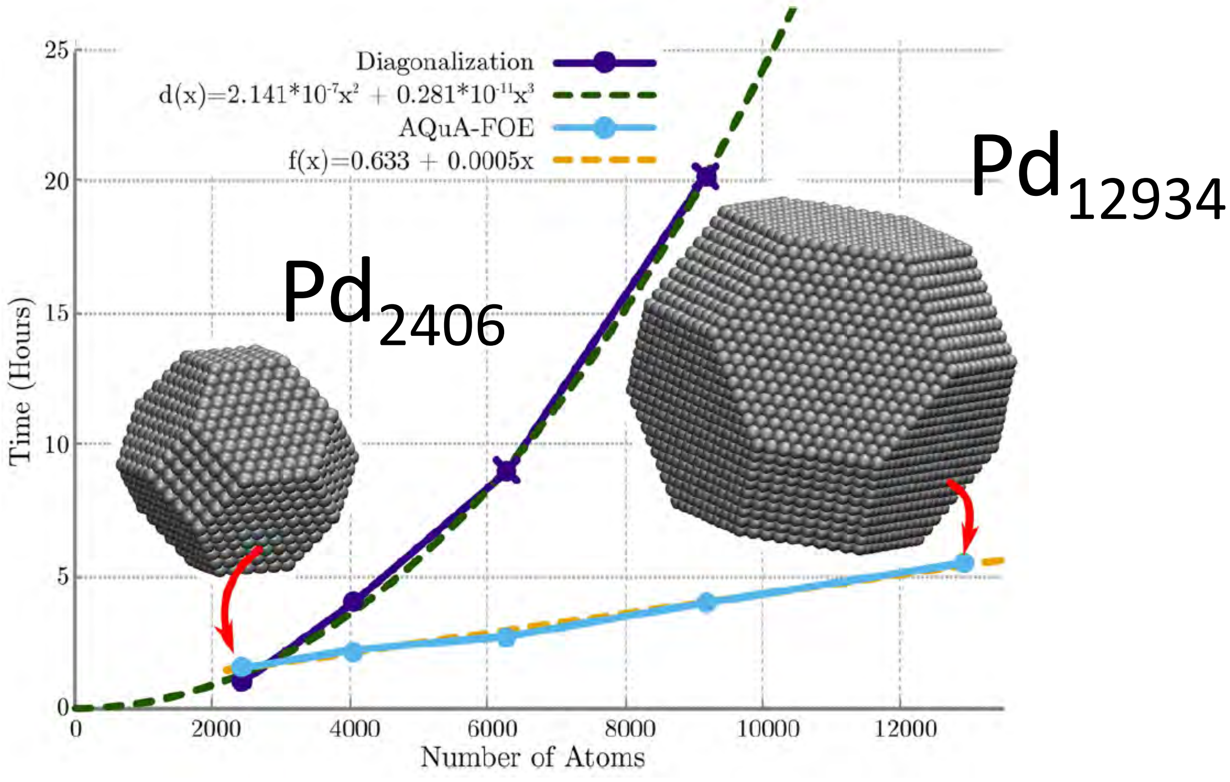


New linear-scaling method for metallic systems in ONETEP

Annealing and QUenching Algorithm FOE (AQuA-FOE)

J. Aarons and C.-K. Skylaris, *J. Chem. Phys.* **148**, 074107 (2018)

- Construct the density matrix from a polynomial expansion of the Hamiltonian
- Number of terms kept constant by using high electronic temperature and then annealing to the target (low) temperature
- Comparison of AQUA-FOE with EDFT (cubic –scaling diagonalization based calculation)



Pd_{2406}

Pd_{12934}

More information

User manuals, keywords, tutorials, publications:
www.onetep.org

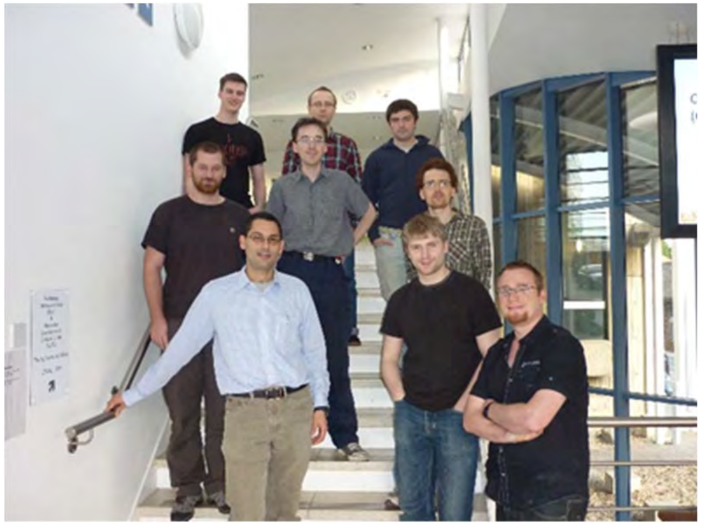
See also *J. Chem. Phys.* **122**, 084119 (2005)

Karl Wilkinson prize 2019

The 2019 Karl Wilkinson prize winner is Subhayan Roychoudhury for the paper:

Wannier-function-based constrained DFT with nonorthogonality-correcting Pulay forces in application to the reorganization effects in graphene-adsorbed pentacene
Subhayan Roychoudhury, David D. O'Regan, and Stefano Sanvito.
Physical Review B **97**, 205120 (2018)

Karl Wilkinson (19/2/1981-25/8/2017) was a bright scientist, a friend, and a continuing collaborator in the ONETEP community. This prize has been established to celebrate his significant contribution to ONETEP and recognise the contributions of others in the ONETEP community



UNIVERSITY OF Southampton
School of Chemistry