CP2K User Meeting 2016

— HowTo: Filter Matrix Diagonalisation REPEAT method for charge fitting

Lianheng Tong
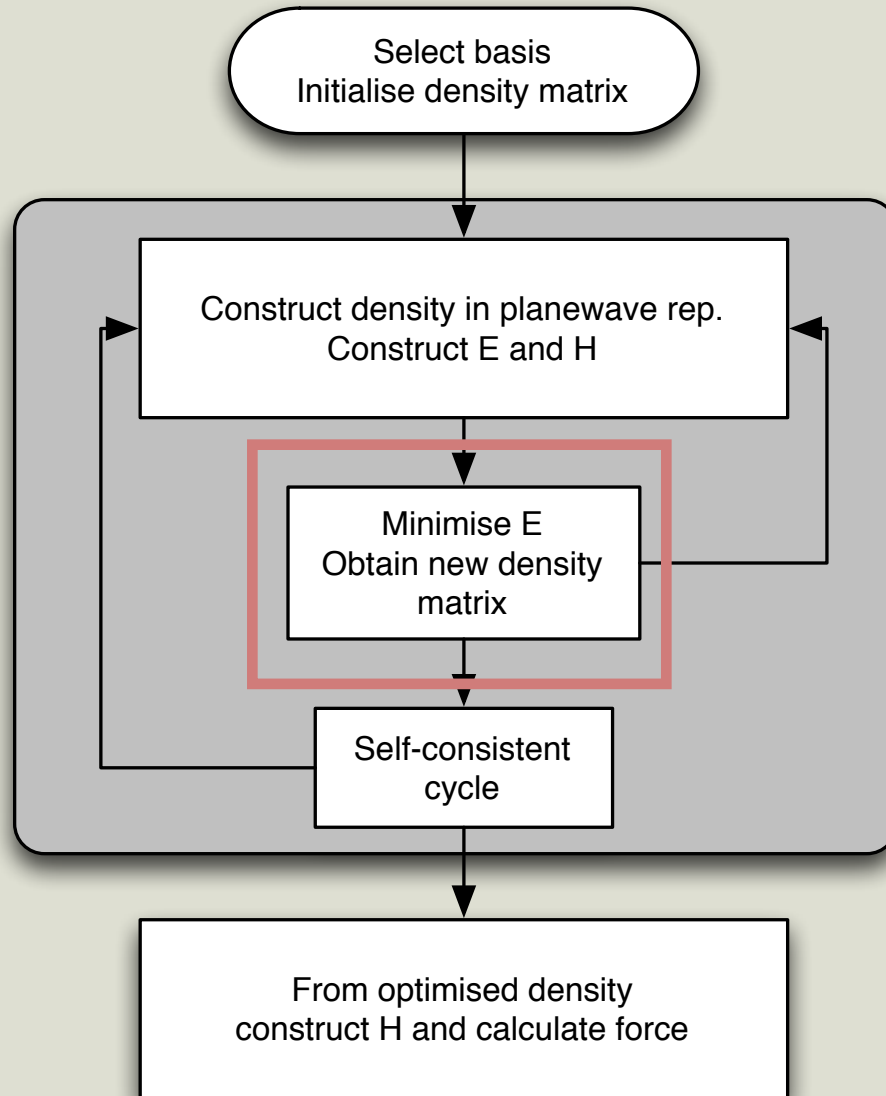
6th Feb. 2015

Faculty of Natural & Mathematical Sciences

Department of Physics

KING'S College LONDON

Self-Consistent Field Calculation

Select basis
Initialise density matrix

Construct density in planewave rep.
Construct E and H

Minimise E
Obtain new density matrix

Self-consistent cycle

From optimised density construct H and calculate force

Problem:

- Metallic systems: band-gap ~ 0.0 eV : conduction band contributing to ground state density

- In a Gamma point problem (CP2K):

$$n(\mathbf{r}) = \sum_n f_n(E_n)\Psi_n(\mathbf{r})\Psi_n^*(\mathbf{r})$$

- Occupancy function is discontinuous (Fermi-Dirac function), a very small change in energy from a SCF step can lead to a jump in the electron charge density, as orbitals from the conduction band at near the Fermi-level gets either included or excluded.

  - Leads to instability in SCF calculations

- A common solution is to introduce smearing, $f_n(E_n)$ becomes continuous.

  - This requires the calculation of energy eigenvalues: not available to methods that directly minimises E with respect to density: (order N, or Orbital Transform methods).

Diagonalisation: Generalised Eigenvalue Problem

$$\hat{H}|\psi_n\rangle = \epsilon_n|\psi_n\rangle$$

$$\mathbb{1} \qquad\qquad\qquad \mathbb{1}$$

$$\sum_{j\beta}\langle\phi_{i\alpha}|\hat{H}|\phi_{j\beta}\rangle\langle\phi_{j\beta}|\psi_n\rangle = \epsilon_n\sum_{j\beta}\langle\phi_{i\alpha}|\phi_{j\beta}\rangle\langle\phi_{j\beta}|\psi_n\rangle$$

atom   shell

$$H_{i\alpha,j\beta} \qquad\qquad\qquad\qquad S_{i\alpha,j\beta} \qquad \psi_{j\beta,n}$$
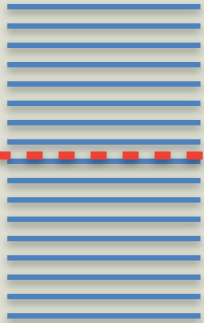
Functional of the density

$$\hat{P} = \sum_n f_n|\psi_n\rangle\langle\psi_n| \qquad\qquad \rho(\mathbf{r}) = \langle\mathbf{r}|\hat{P}|\mathbf{r}\rangle$$
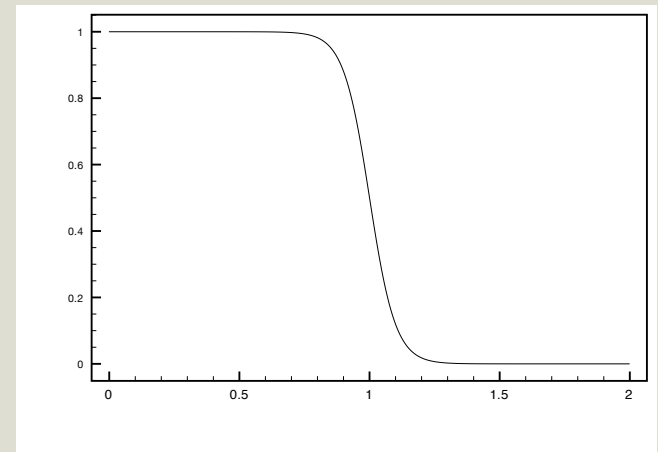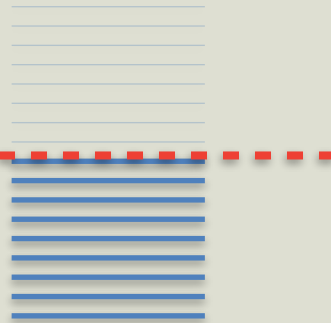
For Ground State we do not need full Hilbert space

Scalar

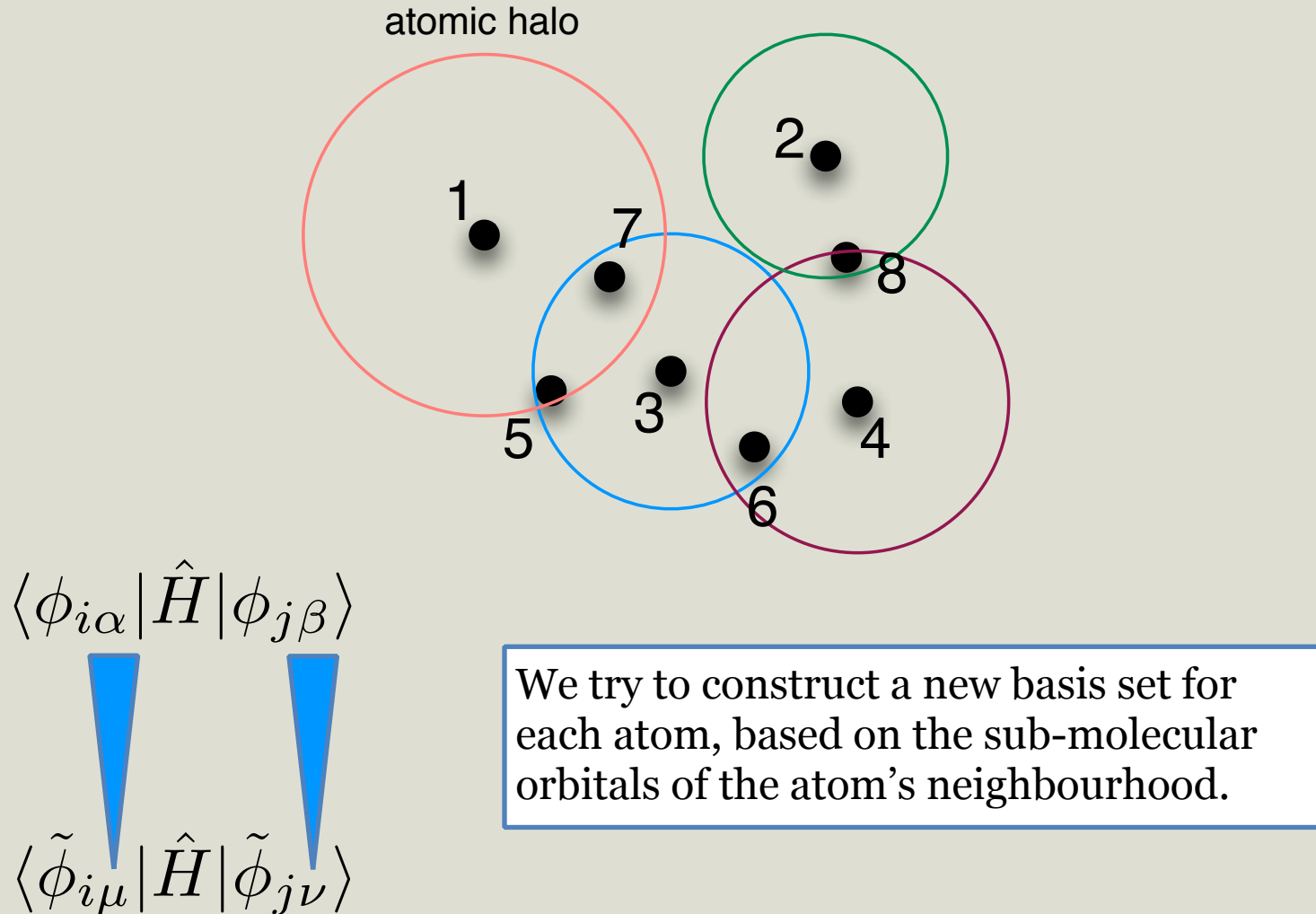$$\sum_n f_n |\psi_n\rangle \langle \psi_n | v\rangle = \sum_{n,i\alpha} f_n \psi_{i\alpha,n}^* v_{i\alpha} |\psi_n\rangle$$

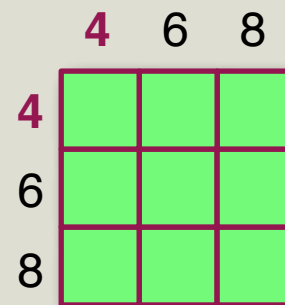$$\rho = \sum_n f_n \Psi_n \Psi_n^*$$

$$E = E[\rho]$$

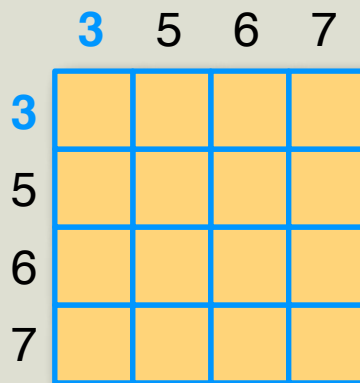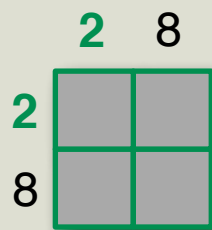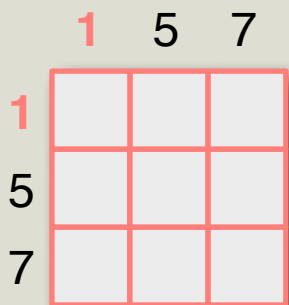- The density matrix projects any state into a space spanned by the lower energy eigenstates

- The smaller set of lower energy eigenstates give exact ground state energy (i.e. a smaller basis set)

- Unfortunately, we do not know the eigenstates a priori

We construct atomic orbitals in situ

atomic halo

$$\langle \phi_{i\alpha} | \hat{H} | \phi_{j\beta} \rangle$$

$$\langle \tilde{\phi}_{i\mu} | \hat{H} | \tilde{\phi}_{j\nu} \rangle$$

We try to construct a new basis set for each atom, based on the sub-molecular orbitals of the atom's neighbourhood.

Rayson and Briddon, Phys. Rev. B 80, 205104 (2009)

Construction of sub-system Hamiltonians
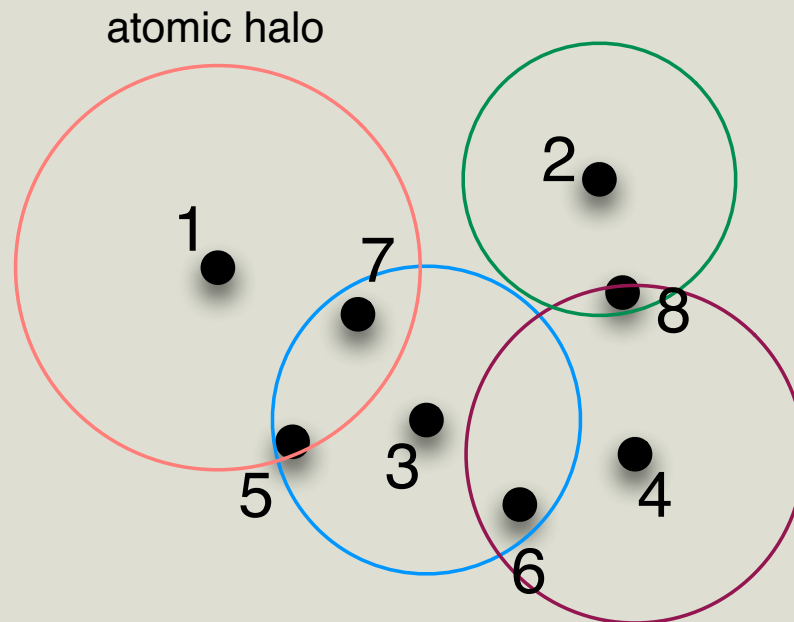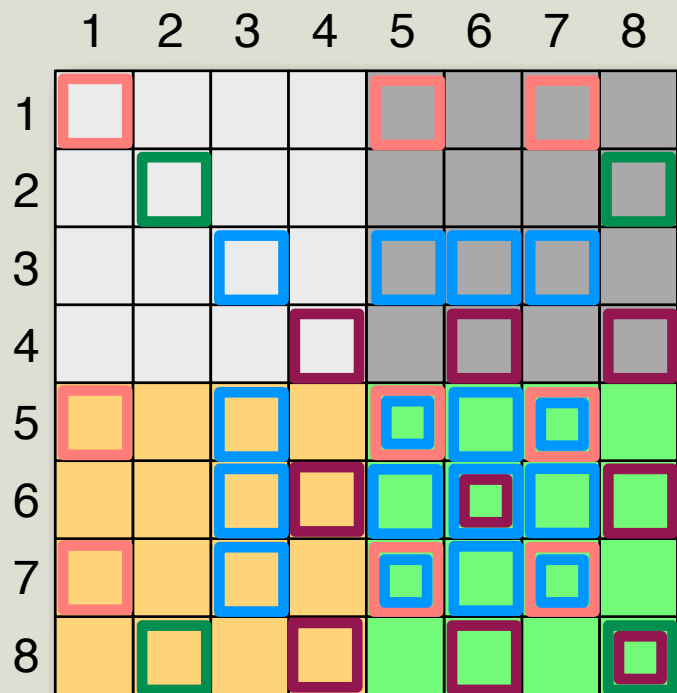
Atomic Density Matrices: Projectors for our new basis

$$\tilde{P} = \sum_n F_n |\tilde{\psi}_n\rangle\langle\tilde{\psi}_n|$$

A high temperature Fermi-Dirac function to include more higher energy orbitals, as we are in an interactive environment
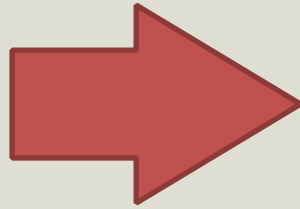
$\tilde{H}_i$

Diagonalise

$$|\tilde{\phi}_{i\mu}\rangle = \tilde{P}|t_{i\mu}\rangle = \sum_n F_n |\tilde{\psi}_n\rangle\langle\tilde{\psi}_n|t_{i\mu}\rangle$$

Optimised for the atom within its environment

$$|\tilde{\psi}_n\rangle = \sum_{j\beta} C_{j\beta,n}|\phi_{j\beta}\rangle, \qquad \|\mathbf{R}_i - \mathbf{R_j}\| < r_{\mathrm{cut}}(i)$$

Still localised around position of atom i

Overall Scheme Of SCF Calculation
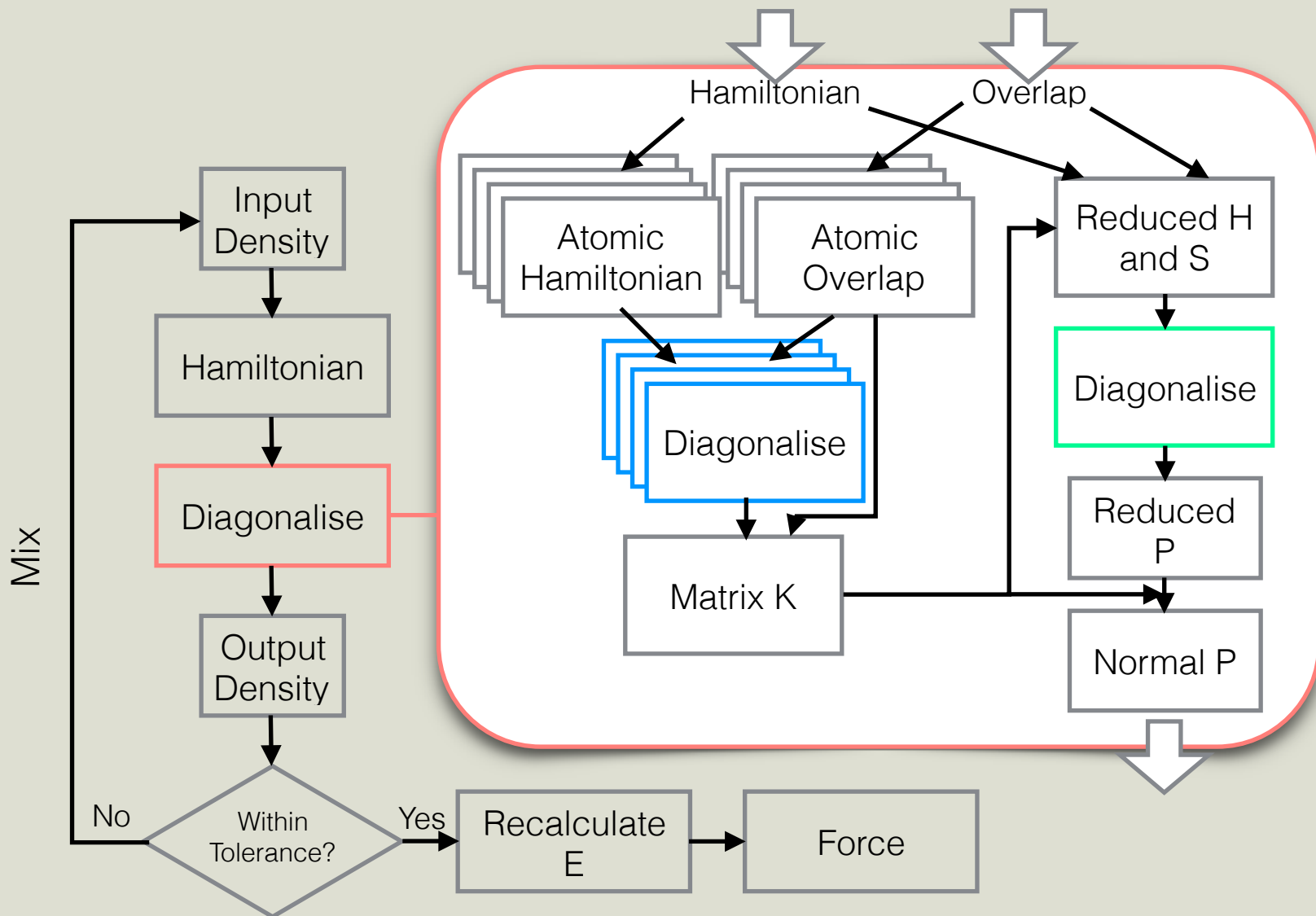
Filter matrix method input

```
&SCF
  &DIAGONALIZATION
     ALGORITHM FILTER_MATRIX
    &FILTER_MATRIX
       FILTER_TEMPERATURE 300000
       AUTO_CUTOFF_SCALE  0.5
    &END FILTER_MATRIX
  &END DIAGONALIZATION
  &PRINT
    &FILTER_MATRIX  ON
    &END FILTER_MATRIX
  &END PRINT
&END SCF
```

Accuracy

Al slab 1372 atoms

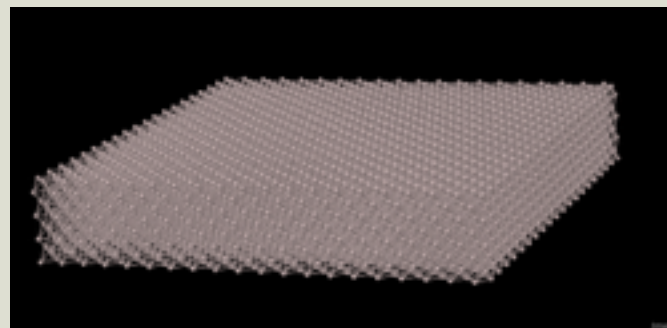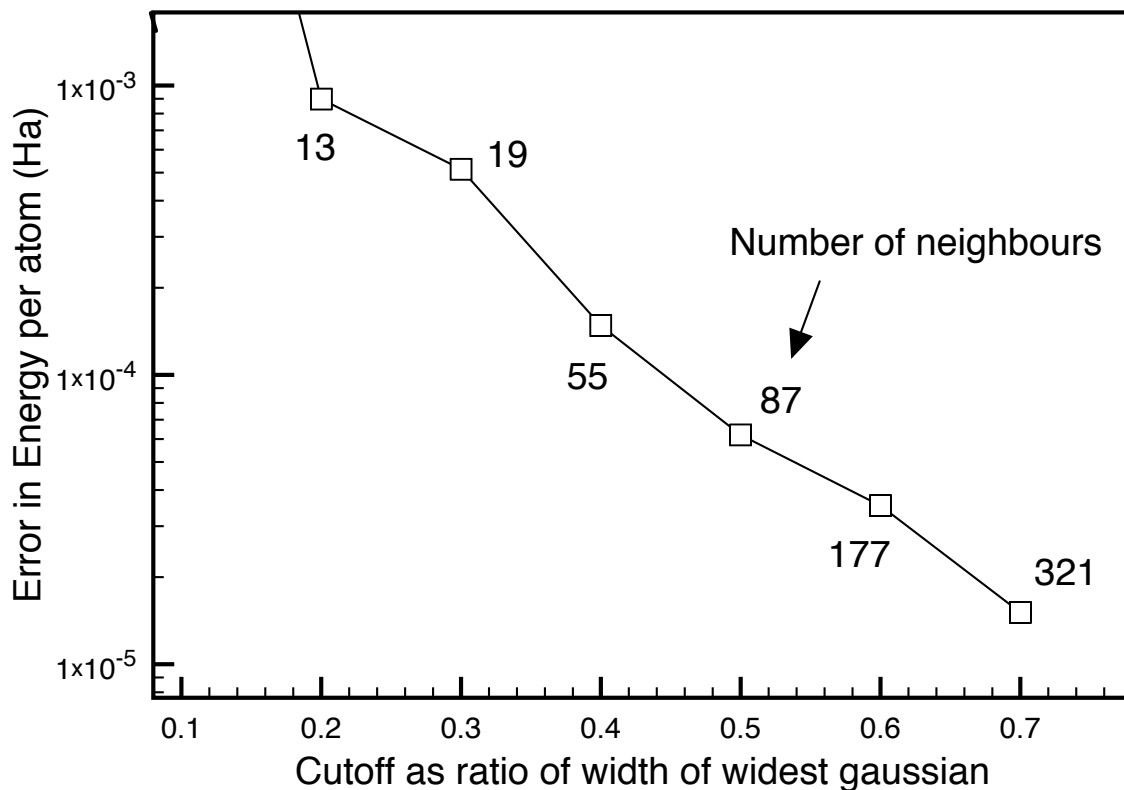Graph: x-axis "Cutoff as ratio of width of widest gaussian" (0.1 to 0.7), y-axis "Error in Energy per atom (Ha)" (1×10^{-5} to 1×10^{-3}). Data points labelled with "Number of neighbours": 13, 19, 55, 87, 177, 321.

Basis set: DZVP-MOLOPT-SR-GTH

Pseudopotential: GTH-PADE-q3

Fermi-Dirac smearing: 300K

Computed on Cray XC30 (ARCHER), gfortran linked with Cray LibSci

Cutoff is defined as the ratio of the extent of the broadest primary basis function. The larger the cutoff the more accurate is the result, but also becomes more expensive quickly

Cost comparison

Single SCF step cost, Al slab
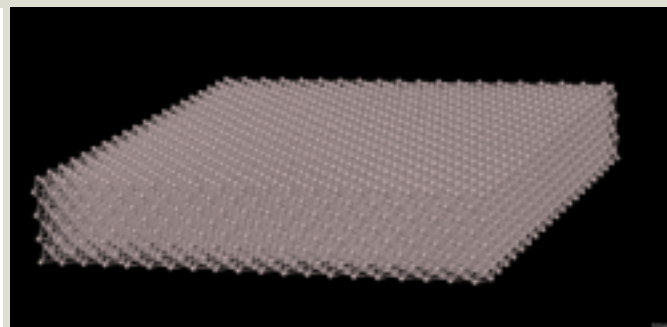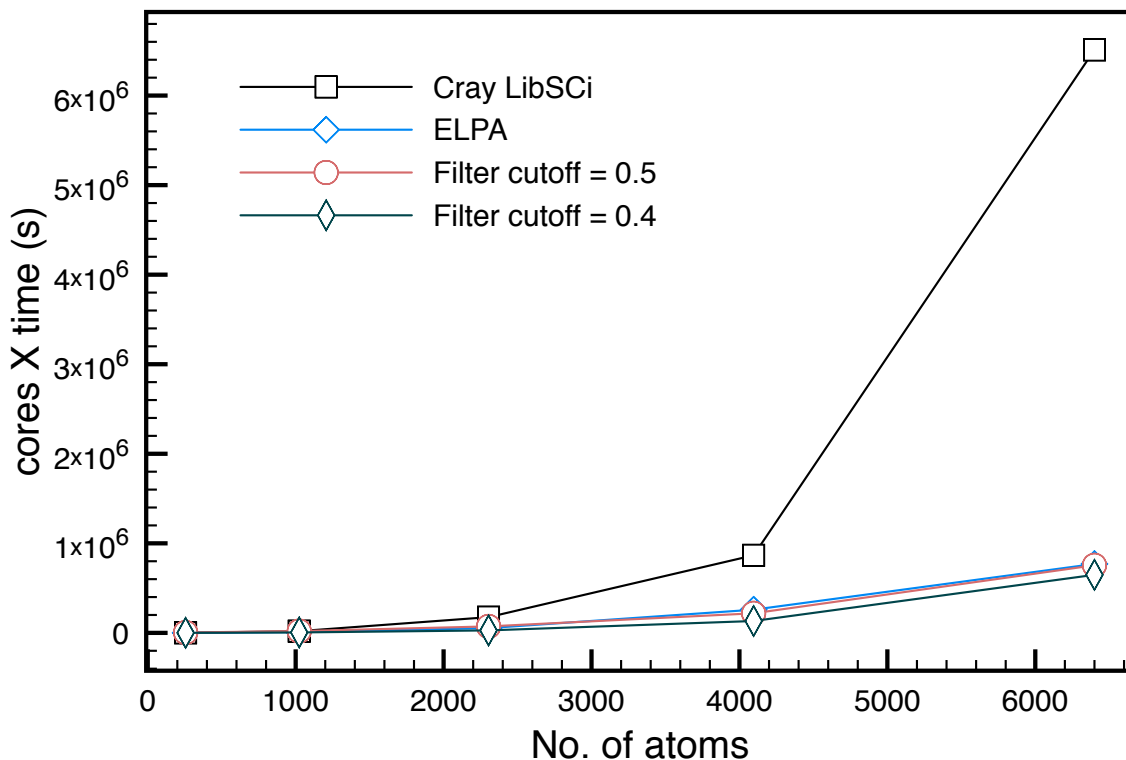
Legend:
- Cray LibSCi (square)
- ELPA (diamond)
- Filter cutoff = 0.5 (circle)
- Filter cutoff = 0.4 (diamond)

Y-axis: cores X time (s), values 0, 1×10^6, 2×10^6, 3×10^6, 4×10^6, 5×10^6, 6×10^6

X-axis: No. of atoms, values 0, 1000, 2000, 3000, 4000, 5000, 6000

Basis set: DZVP-MOLOPT-SR-GTH

Pseudopotential: GTH-PADE-q3

Fermi-Dirac smearing: 300K

Computed on Cray XC30 (ARCHER), gfortran linked with Cray LibSci

Filter basis reduction: DZP to SZ

Cost comparison

Single SCF step cost, random H2O box

Basis set: O and H
TZV2P-GTH

Pseudopotential:
GTH-PADE

Filter basis reduction:
TZV2P to SZ

Computed on Cray XC30
(ARCHER), gfortran linked
with Cray LibSci

MD comparison
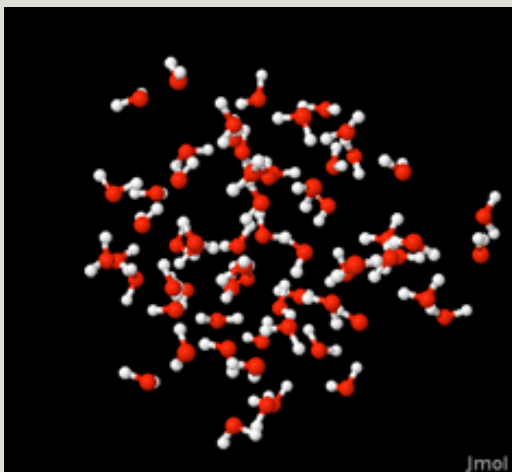
Basis set: O and H
TZV2P-GTH

Pseudopotential:
GTH-PADE

Initial temperature:
300K
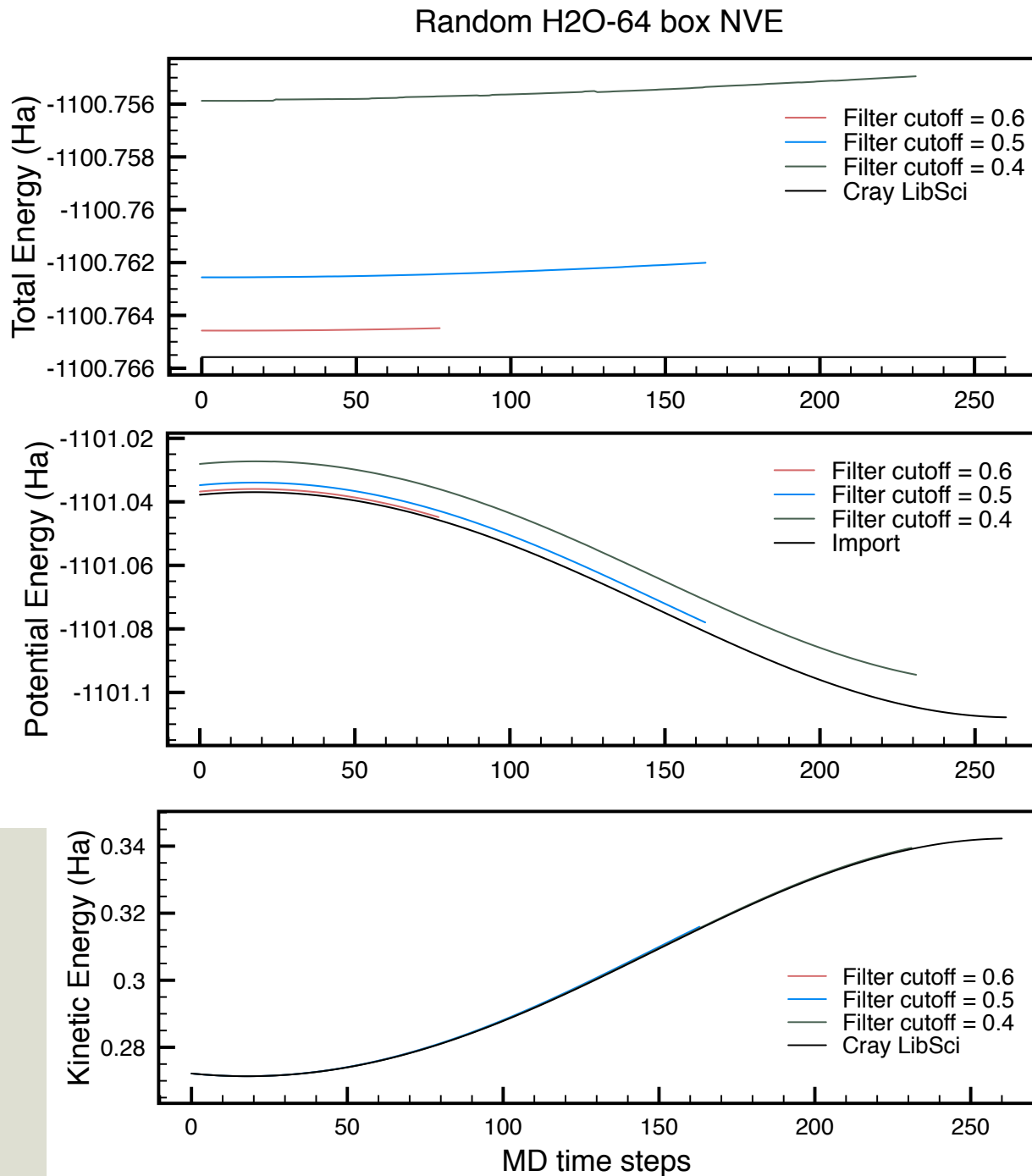
Electrostatic Potential Derived Atomic Charges

- Used for classical forcefield or Monte Carlo simulations for large numbers of atoms
  - Treatment of electrostatic interactions still remains a challenge for transferable forcefields
  - We need an assortment of method to derive atomic charges in an interactive system
    - Electrostatic potential (ESP) derived charges results from calculation from static DFT calculation, by fitting atomic charges to reproduce the electronic static potential
    - In a standard ESP approach restraints and constraints are often required during the fitting process to make the atomic charges to have physically meaningful values. These approaches are referred to as RESP methods.

Electrostatic Potential Derived Atomic Charges

- In CP2K's RESP implementation: electron density of point charges are represented by gaussian functions with fixed width

$$n(\mathbf{r}) = \sum_a q_a \left(\frac{\alpha}{\pi}\right)^{\frac{3}{2}} \exp\left(-\alpha \|\mathbf{r} - \mathbf{R}_a\|^2\right)$$

- The classical static potential due to point charges are then given as the solution of the Poisson's equation:

$$V_{\mathrm{RESP}}(\mathbf{r}) = \sum_a q_a \iiint \mathrm{d}^3\mathbf{r}' \frac{g_a(\mathbf{r}, \mathbf{R}_a)}{\|\mathbf{r} - \mathbf{r}'\|}$$

- We try to fit $V_{\mathrm{RESP}}(\mathbf{r})$ w.r.t. the DFT static potential using the parameters q_a

Electrostatic Potential Derived Atomic Charges

- Fitting process involves the minimisation of the residual function:

Restraints   Constraints

$$R = \frac{1}{N} \sum_{k}^{N} (V_{\text{DFT}}(\mathbf{r}_k) - V_{\text{RESP}}(\mathbf{r}_k))^2 + \beta \sum_{j} (q_j - t_j)^2 + \lambda \sum_{j} q_j$$

- Restraints: q_a should be preferably stay within a stated range, by introducing (quadratic) penalty functions

- Constraints: q_a must obey certain rules, such as the conservation of total charge, by using Lagrange multipliers

REPEAT modification to fitting function

$$R = \frac{1}{N} \sum_{k}^{N} (V_{\text{DFT}}(\mathbf{r}_k) - \boxed{V_{\text{RESP}}(\mathbf{r}_k)})^2 + \beta \sum_{j} (q_j - t_j)^2 + \lambda \sum_{j} q_j$$

$$R = \frac{1}{N} \sum_{k}^{N} (V_{\text{DFT}}(\mathbf{r}_k) - \boxed{(V_{\text{RESP}}(\mathbf{r}_k) + \delta_V)})^2 + \beta \sum_{j} (q_j - t_j)^2 + \lambda \sum_{j} q_j$$

$$\delta_V = \frac{1}{N} \sum_{k}^{N} (V_{\text{DFT}}(\mathbf{r}_k) - V_{\text{RESP}}(\mathbf{r}_k))$$

- For periodic systems, the base level of static potential is ill-defined and arbitrary. It is the shape of the potential surface that matters.

- For REPEAT method, instead of fitting the value of the potential, we will only fit the variance (shape) of the electro-static surface. Coupled with spherical sampling, this leads to a more reasonable set of charges without the need of restraints.

REPEAT Input

```
   &PROPERTIES
     &RESP
      # RESTRAIN_HEAVIES_TO_ZERO F
      USE_REPEAT_METHOD T
       STRIDE 2
       &SPHERE_SAMPLING
          AUTO_RMAX_SCALE 100.0
          AUTO_RMIN_SCALE 1.0
          AUTO_VDW_RADII_TABLE UFF
       &END SPHERE_SAMPLING
       &PRINT
          &PROGRAM_RUN_INFO
          &END PROGRAM_RUN_INFO
       &END PRINT
     &END RESP
&END PROPERTIES
```

```
   &PROPERTIES
     &RESP
      USE_REPEAT_METHOD T
       STRIDE 2
       &SPHERE_SAMPLING
          RMIN_KIND 1.964 Mg
          RMIN_KIND 2.503 C
          RMIN_KIND 1.876 H
          RMIN_KIND 2.275 O
          RMAX 60
       &END SPHERE_SAMPLING
       &PRINT
          &PROGRAM_RUN_INFO
          &END PROGRAM_RUN_INFO
       &END PRINT
     &END RESP
&END PROPERTIES
```

REPEAT Output

```
STARTING RESP FIT

 Calculating charges using REPEAT method.
 Number of explicit constraints:                              0
 Number of explicit restraints:                              0
 Constrain total charge                                      T
 Restrain heavy atoms                                        F
 Stride:                                       2     2     2
 Rmax [angstrom]:                                     60.00000
 Box min [angstrom]:              0.00000   0.00000   0.00000
 Box max [angstrom]:             26.17006  22.66394  20.84412
 Width of Gaussian charge distribution [angstrom^-2]:  11.24900
 Number of potential fitting points found:             1469347
```

```
 RESP charges:
 Type |   Atom   |     Charge

 RESP       1  Mg     1.678470
 RESP       2  O     -0.964519
 RESP       3  C      0.482668
 RESP       4  C     -0.230431
 RESP       5  H      0.185626
 RESP       6  O     -0.990872
 RESP       7  O     -0.783722
 RESP       8  O     -0.980735
 RESP       9  C      1.003359
 RESP      10  C     -0.367970
```