THE UNIVERSITY of York

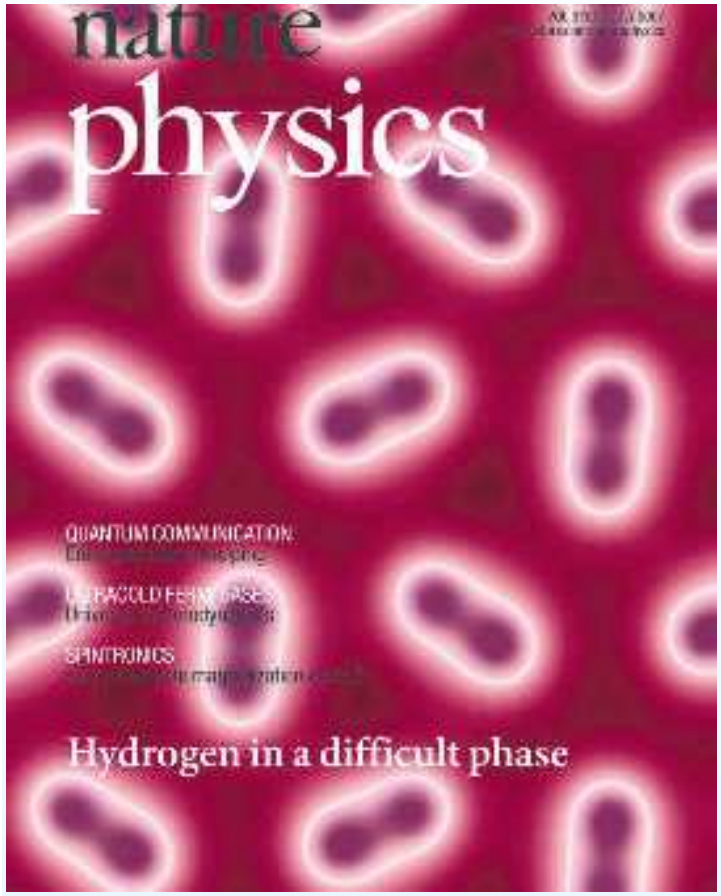
Convergence – how to get a good answer!

Matt Probert
Condensed Matter Dynamics Group
Department of Physics,
University of York, U.K.
http://www-users.york.ac.uk/~mijp1

Research results

- First-principles methods may be used for subtle, elegant and accurate computer experiments and insight into the structure and behaviour of matter.

Structure of phase III of solid hydrogen
Chris J. Pickard and Richard J. Needs
Nature Physics **3**, 473 - 476 (2007)



Research results

- First-principles methods may be used for subtle, elegant and accurate computer experiments and insight into the structure and behaviour of matter.

… or worthless nonsense!

Overview of lecture

- Approximations and convergence
- CASTEP details
- Examples

Approximations and convergence

Approximations

- Distinguish *physical* approximations
  - Born-Oppenheimer Approximation
  - Level of Theory and XC functional
- and convergable, *numerical* approximations
  - basis-set size
  - Integral evaluation cutoffs
  - numerical approximations - FFT grid
  - Iterative schemes: number of iterations and exit criteria (tolerances)
  - system size

Approximations

- Scientific integrity and reproducibility:
  - All methods should agree answer to (for example) “What is lattice constant of silicon in LDA approximation” if sufficiently well-converged.
- No *ab-initio* calculation is ever fully-converged!
  - Need to appreciate how to acceptable level of convergence for given level of theory

Basis set size

- Basis set is fundamental approximation to get correct shape of wavefunction.
- How accurate does it need to be?
  - Variational principle guarantees we get an upper bound on true ground state E_0
  - With energy error $\delta E \sim |\delta\psi|^2$
  - With plane waves can increase size and accuracy by adding waves with higher ***G***
  - Much harder with Gaussian basis – need to add *ad hoc* ‘diffuse’ or ‘polarization’ functions

Properties

- Fortunately well converged *properties* may frequently be computed using an incomplete basis.
- Size of planewave basis set governed by single parameter:
  - **`cut_off_energy`** (in energy units) or **`basis_precision=COARSE/MEDIUM/FINE`** in CASTEP `.param` file
- NB Though E is *monotonic* in E_{cut} it is not necessarily *regular*.

Error cancellation

- Consider energetics of simple chemical reaction: $MgO_{(s)} + H_2O_{(g)} \rightarrow Mg(OH)_{2,(s)}$
- Reaction energy computed as

$$\triangle E = E_{product} - \sum E_{reactants} = E_{Mg(OH)_2} - E\left(E_{MgO} + E_{H_2O}\right)$$

- What if compute at E_{cut}=500 & 4000 eV?
  - Change in energy of each component: ΔE(MgO) = -0.021 eV, ΔE(H_2O) = -0.566 eV, ΔE($Mg(OH)_2$) = -0.562 eV, error in ΔE = -0.030 eV
  - i.e. errors cancel in final result if same E_{cut} etc
  - *Energy differences* converge faster than E

Planewave convergence

- Properties of a plane-wave basis.
  - Cutoff determines highest representable spatial Fourier component of density.
  - Energy *differences* converge faster than total energies. Electron density varies most rapidly near nuclei, where min. effect on bonding
- E_{cut} (and G_{max}) depend on each element
  - Simulation cutoff is max of all elements in calc
- Required cutoff is *system-independent* but not *property-independent*.

Pseudopotentials

- With efficient pseudopotentials *rate* of convergence with cutoff is not always smooth.
  - Can get plateaus, etc
  - Sometimes cause is over-optimization of the pseudopotential at too low a desired cutoff.
  - In which case choose criterion for convergence energy to be plateau values.
- Absolute energy convergence is rarely desirable. Force and stress convergence is much more useful criterion.

FFT grid

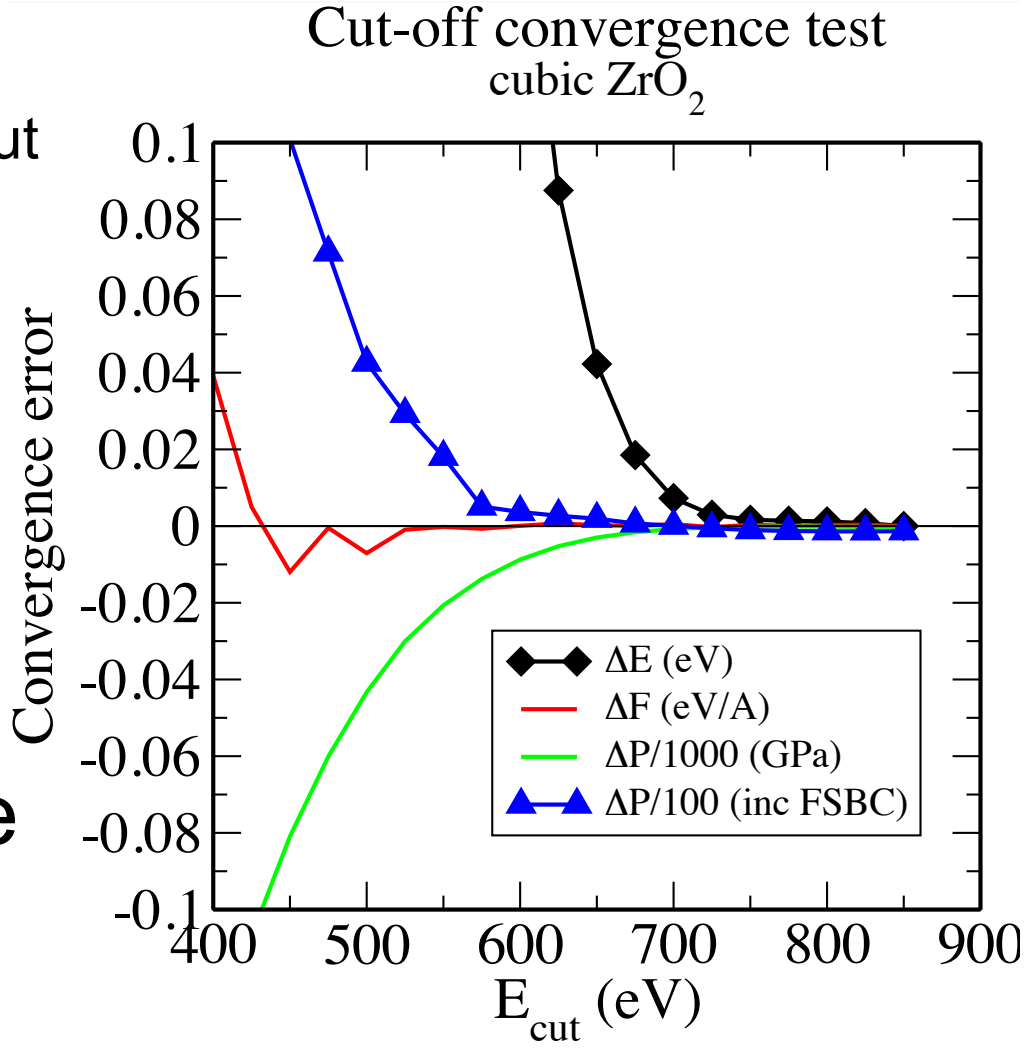
- FFT grid should be large enough to handle all **G**-vectors of density $\rho(\boldsymbol{r})$, within cutoff:
  - $|\mathbf{G}| \leq 2|\mathbf{G}_{max}|$ - the **`grid_scale`**
  - Guaranteed to avoid ”aliasing” errors in case of LDA and pseudopotentials without additional core-charge density.
  - CASTEP v25 has default **`grid_scale=2.0`**
  - Older versions used 1.75 by default as faster and less RAM – OK for LDA
  - GGA and higher may need `grid_scale>2`

FFT grid and ultrasoft

- CASTEP can use ultrasoft pseudopotentials
  - See advanced course for more info on USPs
- Finer grid needed for USP augmentation charges or core-charge densities.
  - CASTEP has a second, finer grid for core/augmentation charges
  - Set by parameter **`fine_grid_scale`** (multiplier of G_{max}) or **`fine_gmax`** (inverse length G_{fine}) – keep G_{max} for orbitals.
- `fine_gmax` is transferable to other systems but `fine_grid_scale` is not.

Force convergence

- CASTEP can do MD or geom opt with forces
- Forces usually converge at lower E_{cut} than total energy as nuclear region unimportant
- Pressure/stress different again
- Test the convergence of each property!!!



Iterative solvers

- Parameter **`elec_energy_tol`** specifies when SCF energy converged.
  - Optimizer also exits if **max_SCF_cycles** reached
  - i.e. *you must always check that it really did find the ground state!*

- How accurate does SCF convergence need to be?

Solver tolerances

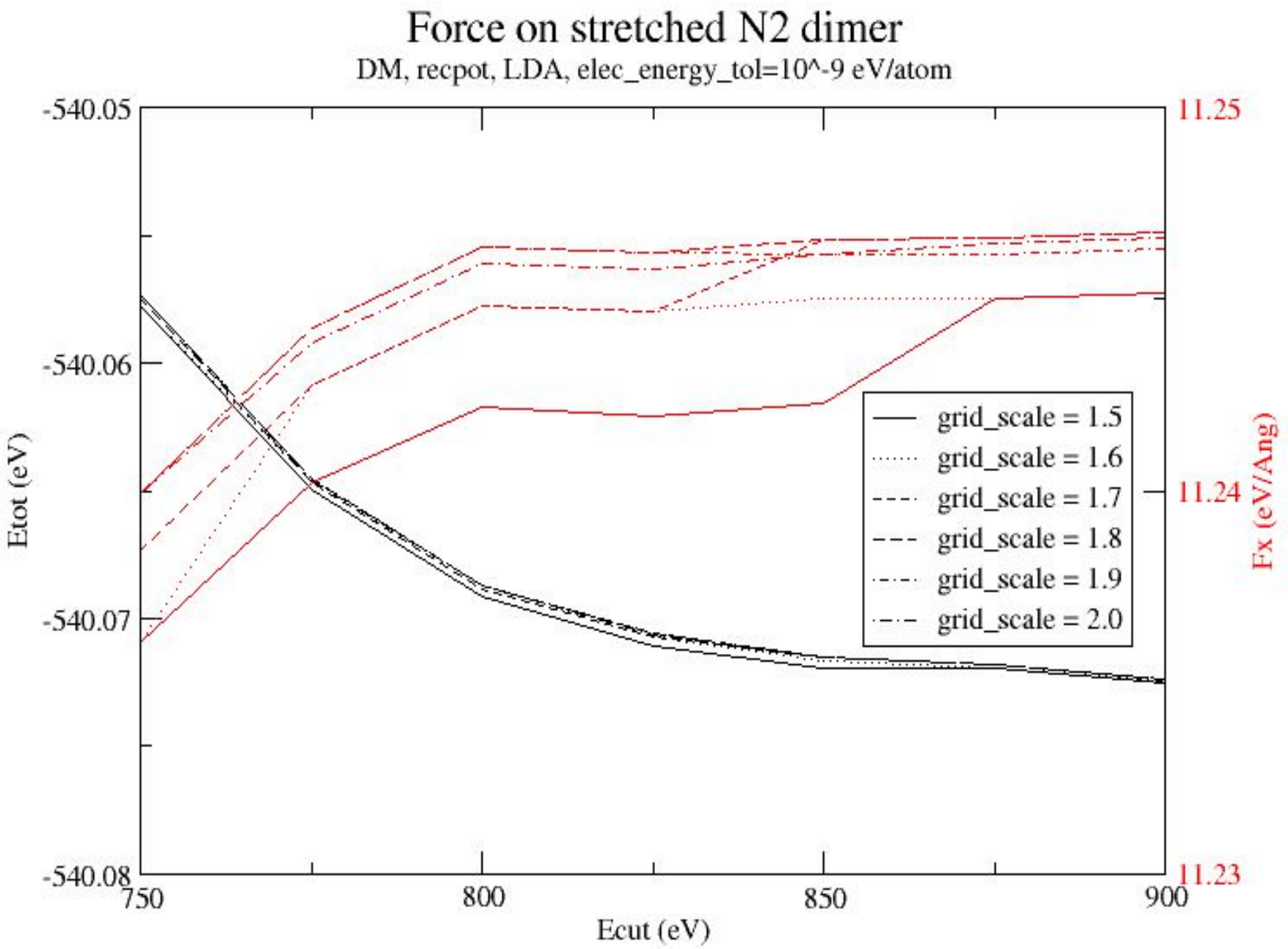
- **Energetics:** same accuracy of result.
- **Geometry/MD:** need much smaller (tighter) `elec_energy_tol` to converge forces.
  - Cost of higher tolerance is only a few additional SCF iterations.
  - Or use `elec_force_tol` to exit SCF only when force convergence criterion satisfied.
- Inaccurate forces are common cause of geometry optimization failure.

Quality of Forces/Stresses

- Can only find a good structure if have reliable forces/stresses
- Stresses converge more slowly than energies as increase number of plane waves
  - CONVERGENCE!
- Should also check degree of SCF convergence

  elec_energy_tol (fine quality ≤ 10^{-6} eV/atom)
  can also set **elec_force_tol** – useful with DM

  Stresses converge slower than forces!

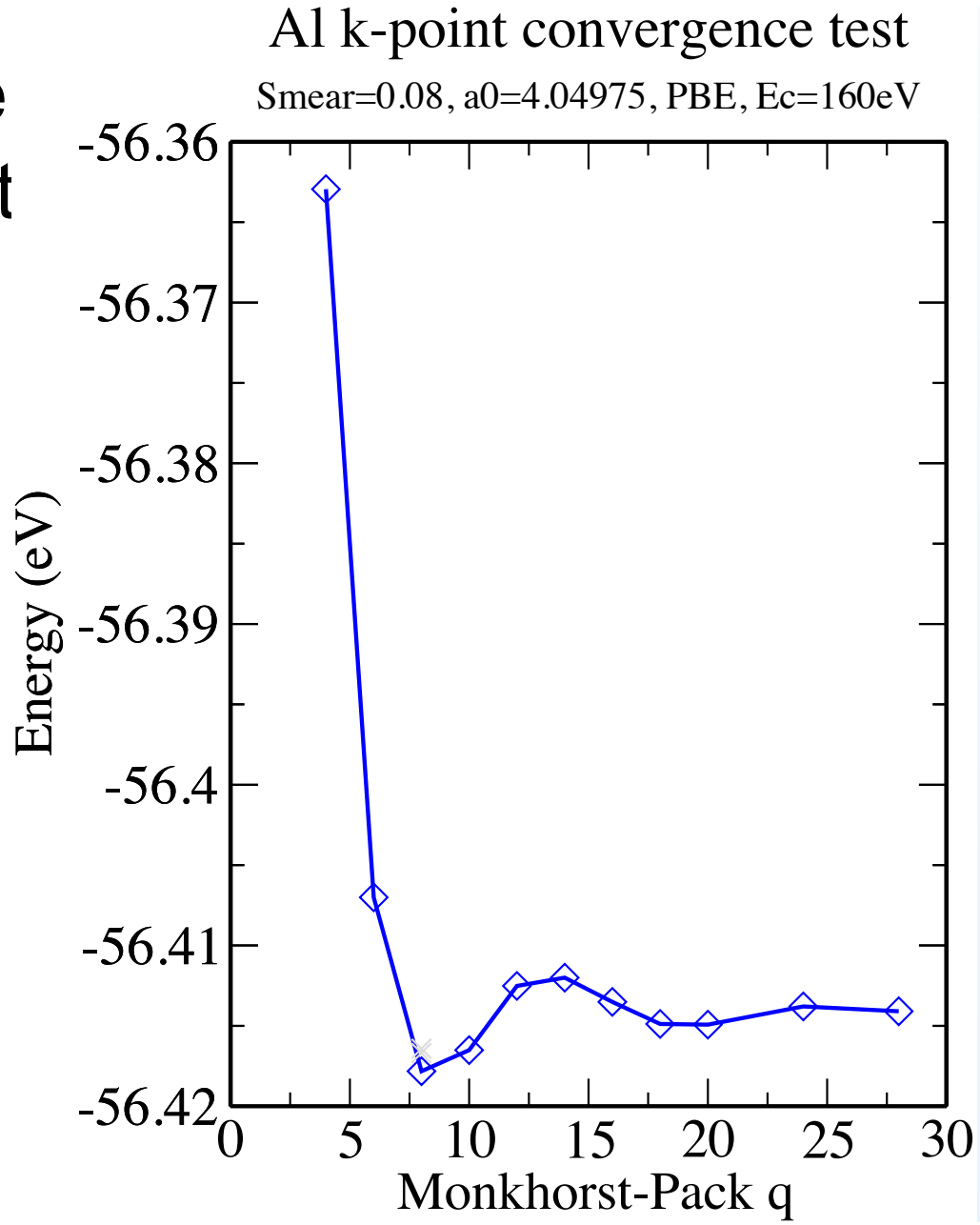

Force on stretched N2 dimer
DM, recpot, LDA, elec_energy_tol=10^-9 eV/atom
-540.05
-540.06
-540.07
-540.08
Etot (eV)
11.25
11.24
11.23
Fx (eV/Ang)
750
800
850
900
Ecut (eV)
grid_scale = 1.5
grid_scale = 1.6
grid_scale = 1.7
grid_scale = 1.8
grid_scale = 1.9
grid_scale = 2.0

K-point convergence

- Brillouin Zone sampling is very important
- Usually set in `<seed>.cell` with keyword **`kpoint_mp_grid p q r`** or **`kpoint_mp_spacing X`** with optional offset **`kpoint_mp_offset`**.
  - Convergence is **not** variational and frequently oscillates.
- Finite-temperature *smearing* can accelerate convergence, but must extrapolate result back to T=0K

K-point convergence of Al

- Even simple metals like Al require dense k-point meshes in primitive cell
- Metals require more k-points than insulators due to occupancy discontinuity at $E=E_F$
- Can get some error cancellation with insulators but only if *identical* cells



Strategies for convergence testing

Strategies for testing convergence

- Tests should be simple and fast
  - Make use of transferability of `cut_off_energy` and `grid_scale` and `kpoint_mp_spacing` or use *length scaling* for k-point grids convergence
  - Test on *representative* small systems
  - Test E_{cut} with coarse k-point grid
  - Test k-point grids with coarse E_{cut}
- Work with fixed geometry as optimization depends on details of convergence
- Accuracy of forces a good proxy for phonon convergence

Automation of convergence tests

- Recently added **`castepconv`** to the standard set of CASTEP tools
  - A suite of Python programs to automate convergence testing
  - Can do cutoff energy, k-point grid and fine_gmax
  - And graph the results

Golden rules of convergence test

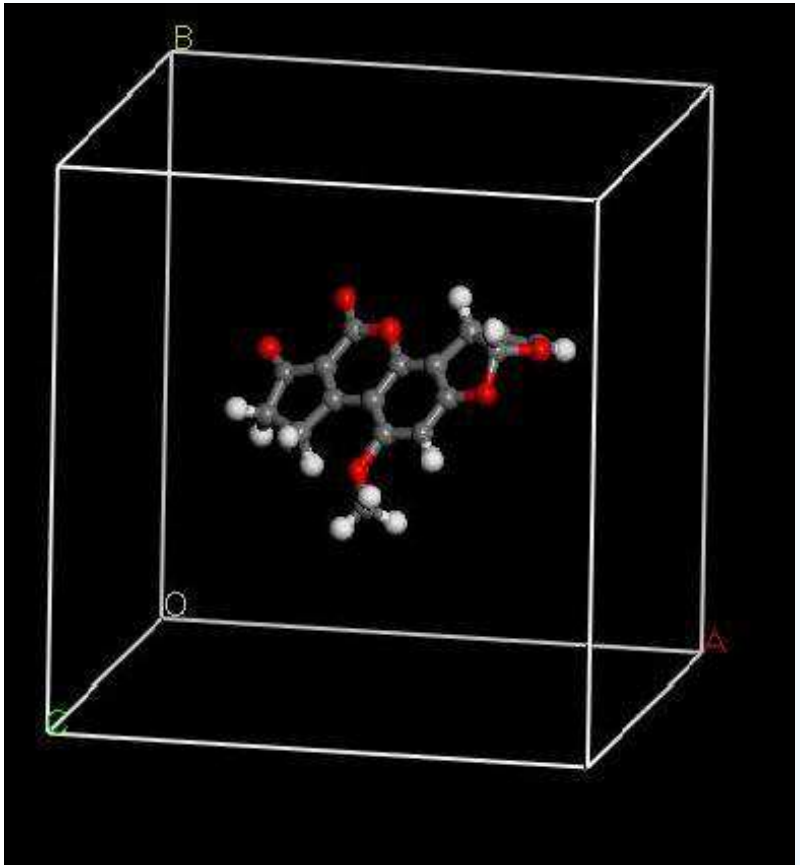
- Change one parameter at a time
- Do not use too high an E_{cut} – are you converging augmentation charges? Check `fine_gmax` instead
- Use forces & stresses as proxies for other more expensive properties (e.g. phonons)
- Write your convergence criteria in paper
- Convergence is when parameter of interest stops changing, NOT when run out of resources or when agree with experiment!

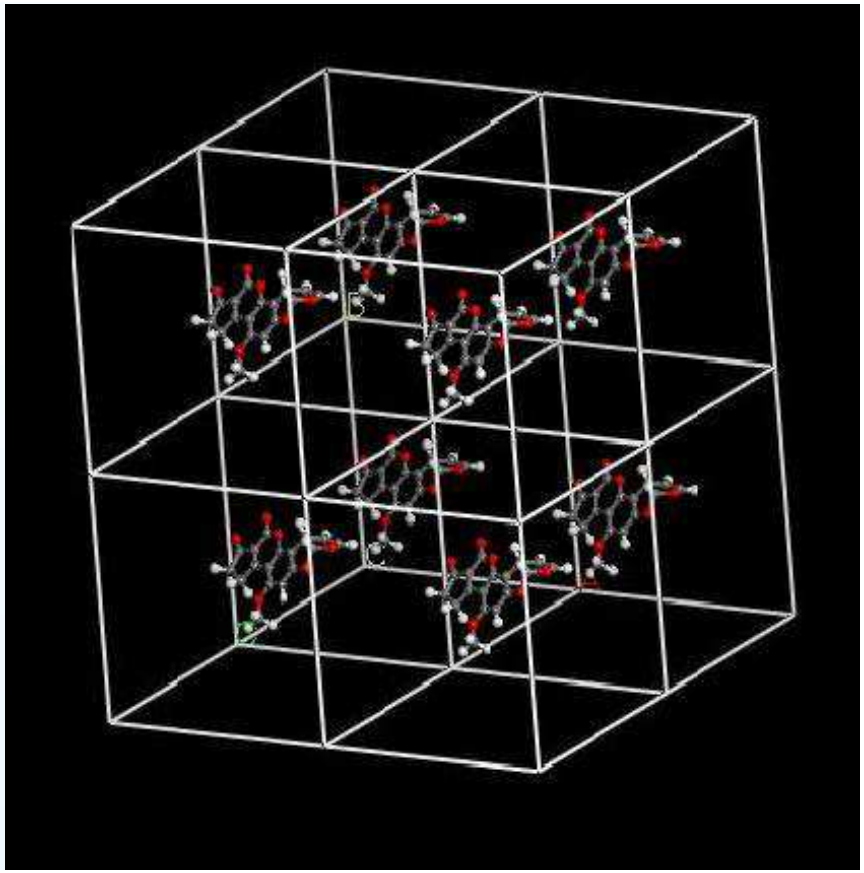
THE UNIVERSITY of York

Structural calculations

Modelling molecules

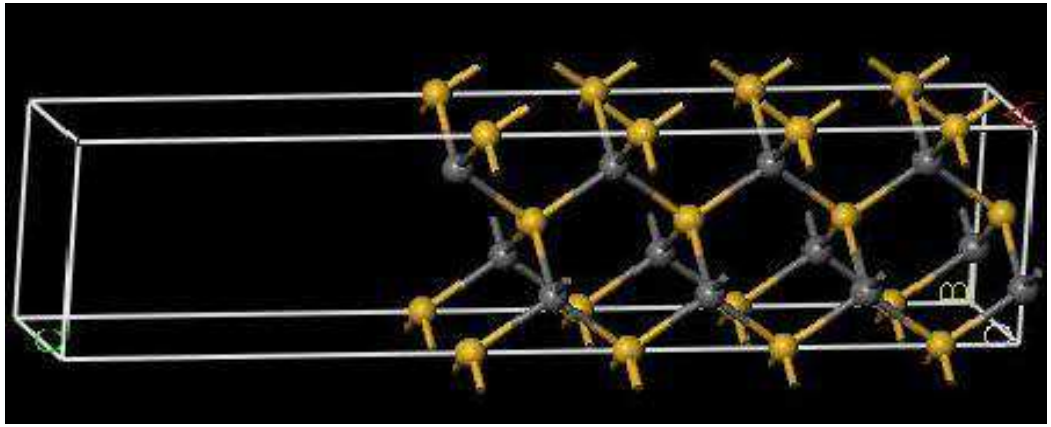
- CASTEP has 3D periodic boundary conditions
  - Surround molecule by vacuum space
  - Periodic array …

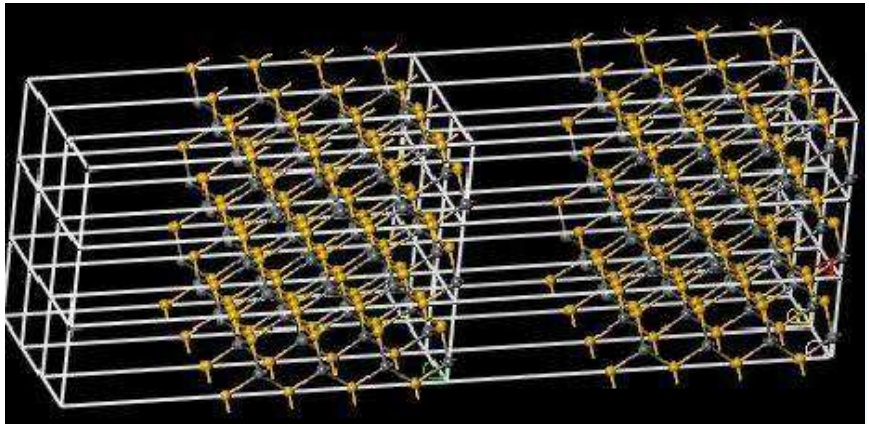




Modelling surfaces

- Ditto – add vacuum to separate surfaces





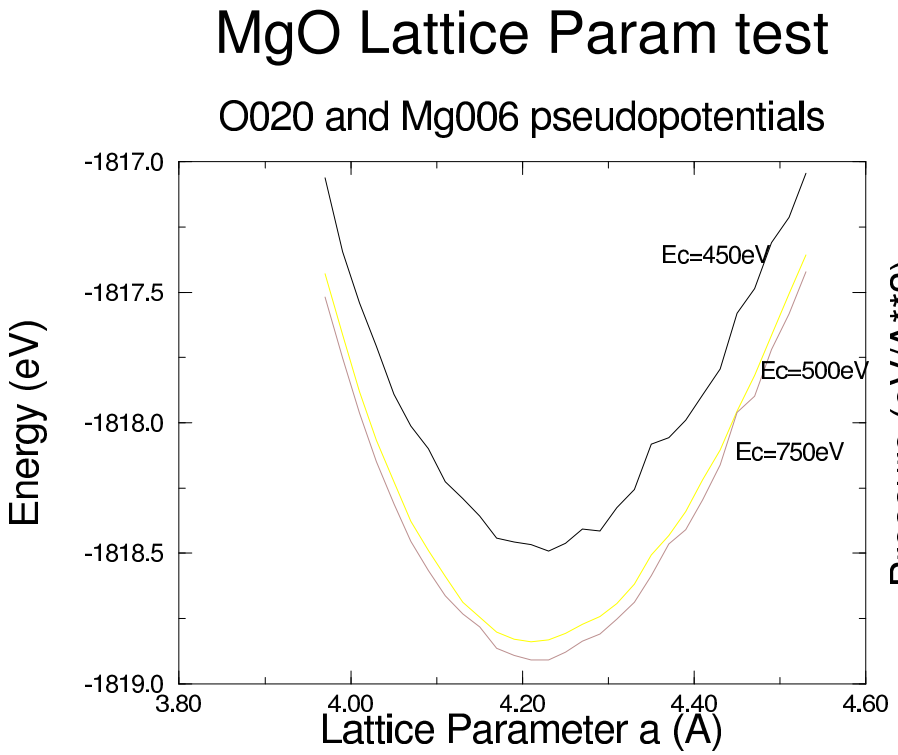
  - But how much vacuum? Must converge it!
- And for slab, need enough layers to get to bulk-like environment
  - Another convergence parameter!

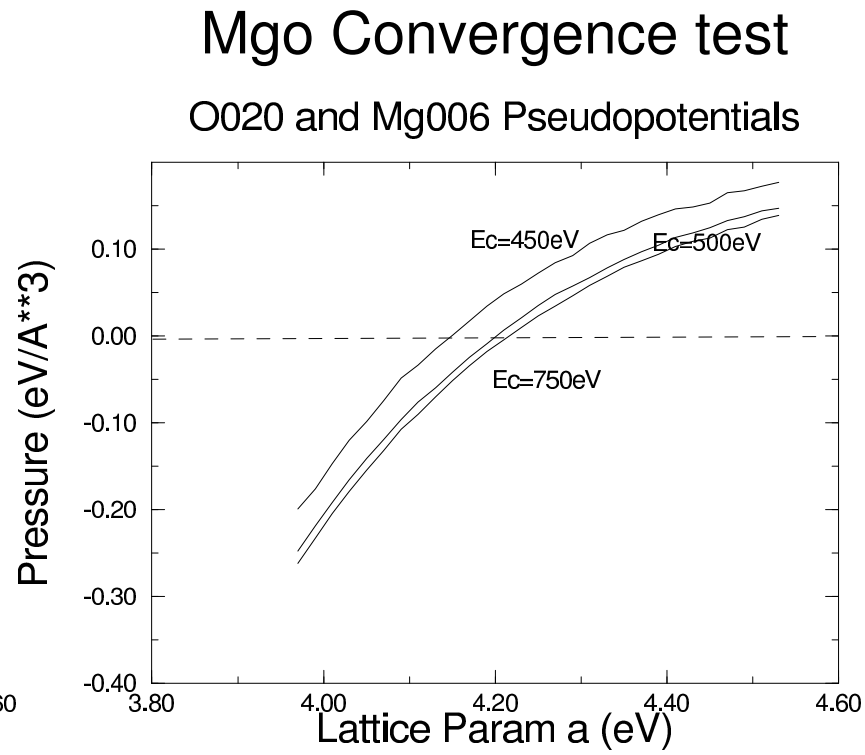
Dipolar surfaces

- Beware polar materials and dipole surfaces
  - Surface energy does not converge with slab thickness!
  - Use same cell for bulk and slab where possible to get k-point error cancellation
  - If not possible then must use absolute converged k-point set as no error cancellation
  - Can use `dipole_correction = static/self-consistent` with `dipole_dir=x/y/z/all` to correct electrostatics

Variable cell calculations

- Can get equilibrium lattice parameter from min E with vol or from where stress=0
  - Stress is automated with geom opt but converges less well than min E with E_{cut}





Finite Basis Correction

- "Jagged" E vs V curve due to change in number of plane waves when change volume
  - G_{min} changes with real-space cell size
  - Correct using Francis-Payne method (J. Phys. Conden. Matt **2**, 4395 (1990))
- Use **`finite_basis_corr=MANUAL/AUTO`** with additional **`finite_basis_npoints`** at **`finite_basis_spacing`** values of Ecut
- Or specify **`basis_dE_dlogE`** instead.

Number of plane waves

- Two possibilities when changing cell size:
  - *fixed basis size* (`fixed_NPW=true`)
    - Number of plane-waves NPW is constant
    - Cell expansion *lowers* G_{max} and K.E. of each plane wave, and therefore lowers effective E_{cut}.
    - Easy to code but easy to get erroneous results.
    - Need very well-converged cutoff for success.
  - *fixed cutoff* => physically correct
    - reset basis every time change cell parameters
    - vary NPW so as to keep G_{max} and E_{cut} fixed

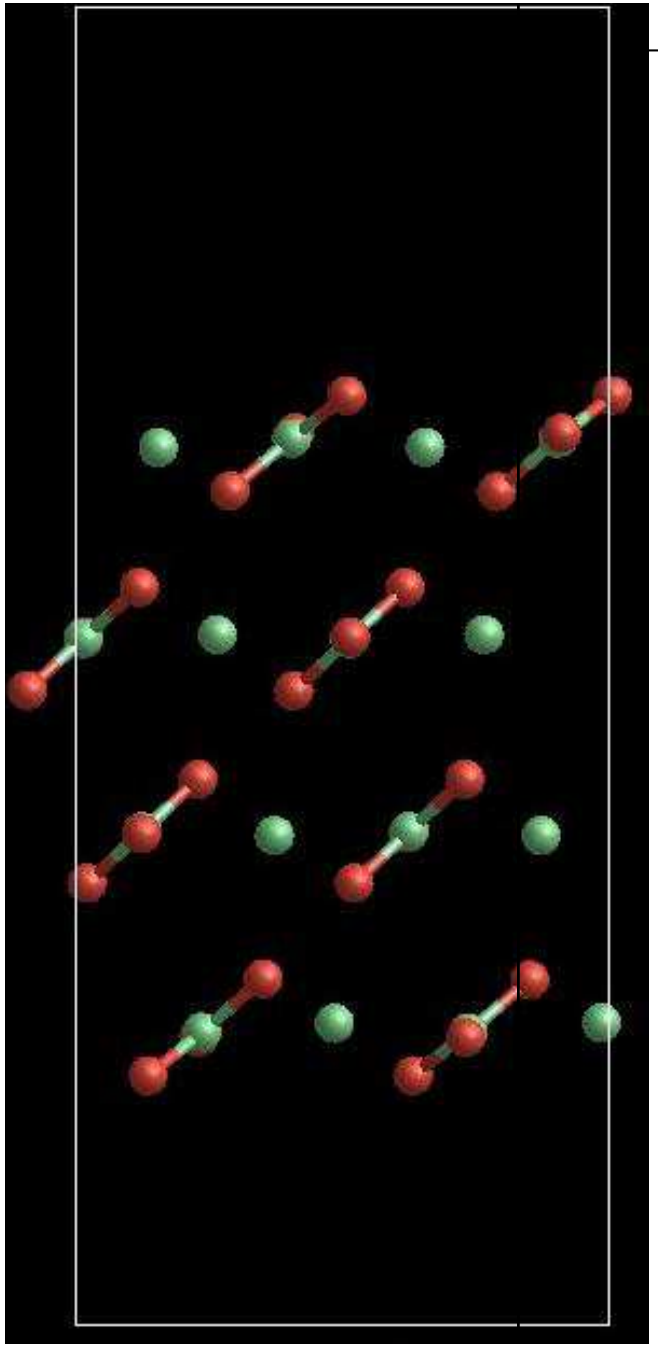
Surface modelling

- Surfaces can be modelled as a *slab* cleaved from bulk crystal. Can calculate
  - Surface *energy* or *free energy*
  - Surface reconstructions
  - Energies of steps
  - Adsorption energies and structures of adsorbates
  - Surface chemical reaction energies.

Choice of surface

- Choice of slab: usually need to make both surfaces identical.
- Surfaces related by symmetry operation are more easily geometry optimised.
- Simulation cell should *not* be optimised. (use **`fix_all_cell=T`** in `.cell` file).
  - In-plane cell parameters usually fixed at values from relaxed bulk crystal.
  - Perpendicular direction has extra vacuum

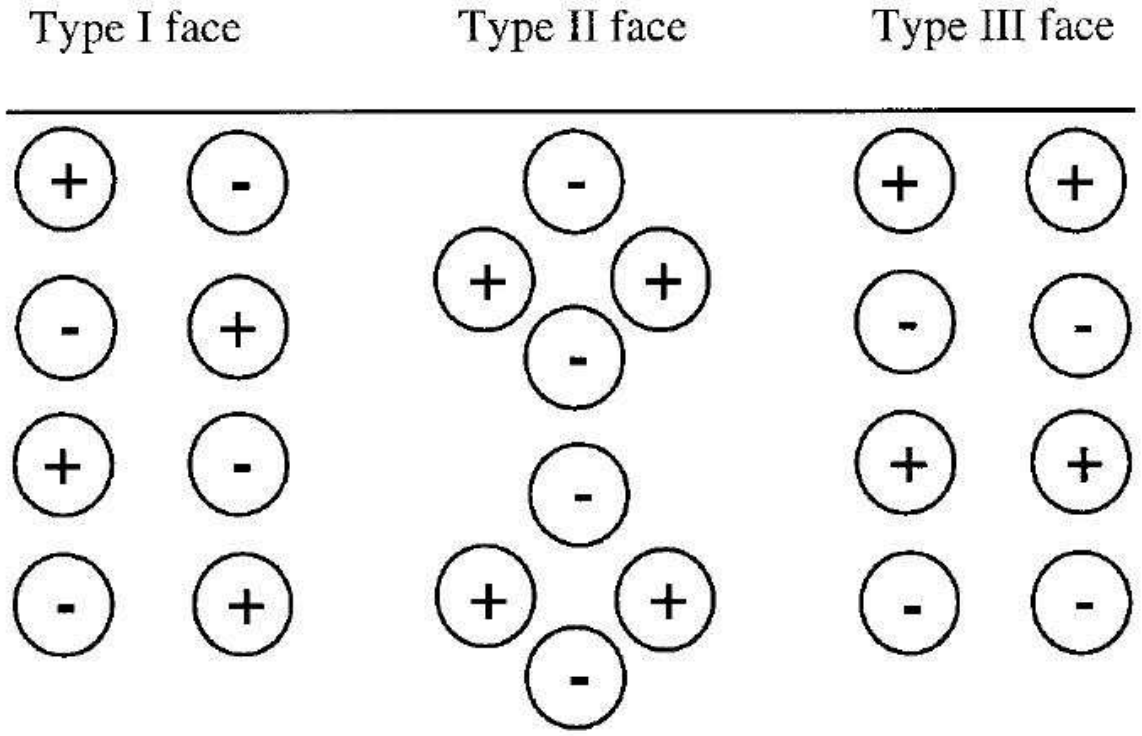
Surface convergence



- E_{surf} is very sensitive to convergence of total energies.
  - Sometime get k-point error cancellation of slab vs bulk if use non-primitive bulk cell with same in-plane vectors as slab. (*not* $CaCO_3 10\bar{1}4$ shown).
  - Only 1 k-point in perp to slab.
- Must converge slab thickness and vacuum gap

Polarization in slabs

- Electric dipoles perpendicular to surface raise theoretical difficulties
  - Energy *does not converge* with slab thickness
  - Three types of polar surface:



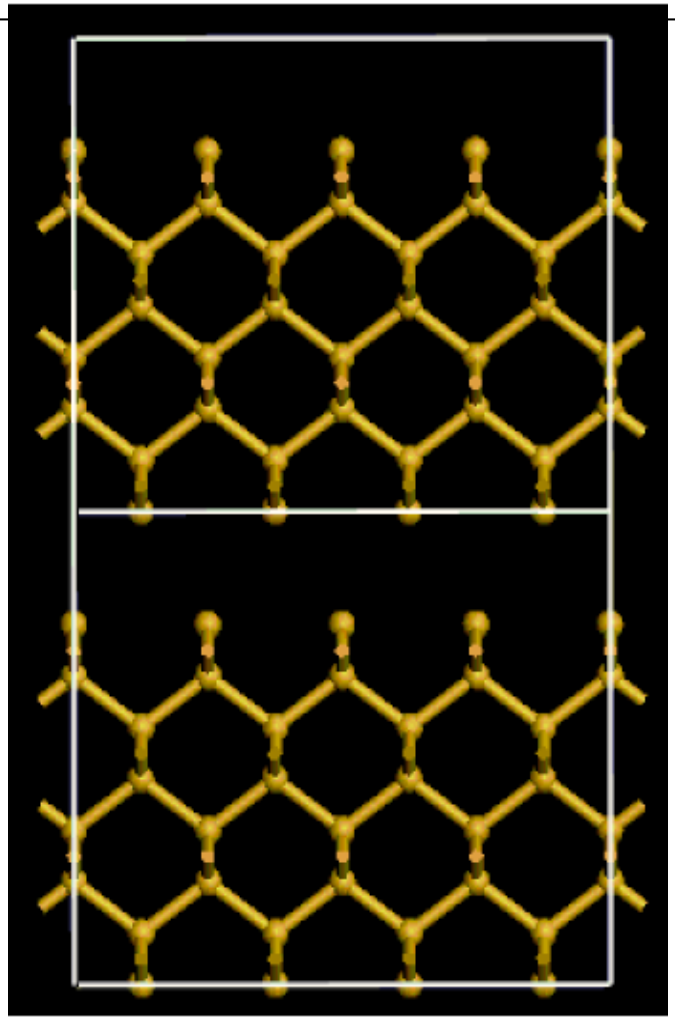
Polarization in slabs

- In classical charge model, Type III unstable and must always reconstruct.
- In *ab initio* calculation, surfaces can instead become metallic.
- Polar molecules on surface also raise dipole-dipole problems.
  - Use CASTEP dipole correction scheme
- Can sometimes use double surface with inversion symmetry
  - Dipoles cancel.

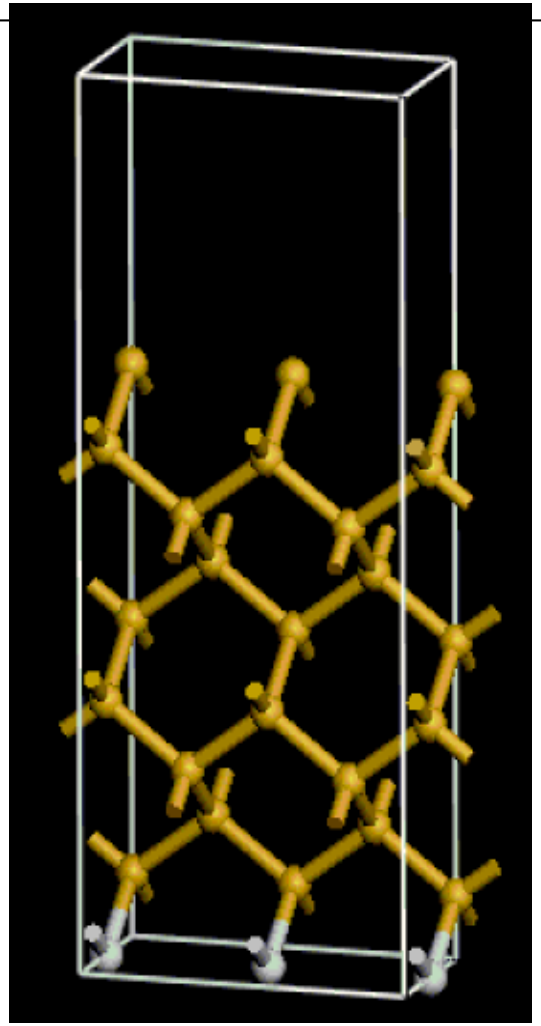
Defect calculations

- Can model defects by *supercell* approach:
  - Some tricky convergence issues to get defect-defect interaction to zero.
  - Only need to converge energy to a few meV, but need accurate *forces* to get relaxation OK.
  - Local strain around defect decays as 1/R. Can use classical models if know suitable pot
- Charged defects can be modelled using periodic interaction correction terms. Also correction terms for higher order multipoles.

Example: Si(100) Surface Reconstruction



Si(100) supercell with vacuum gap



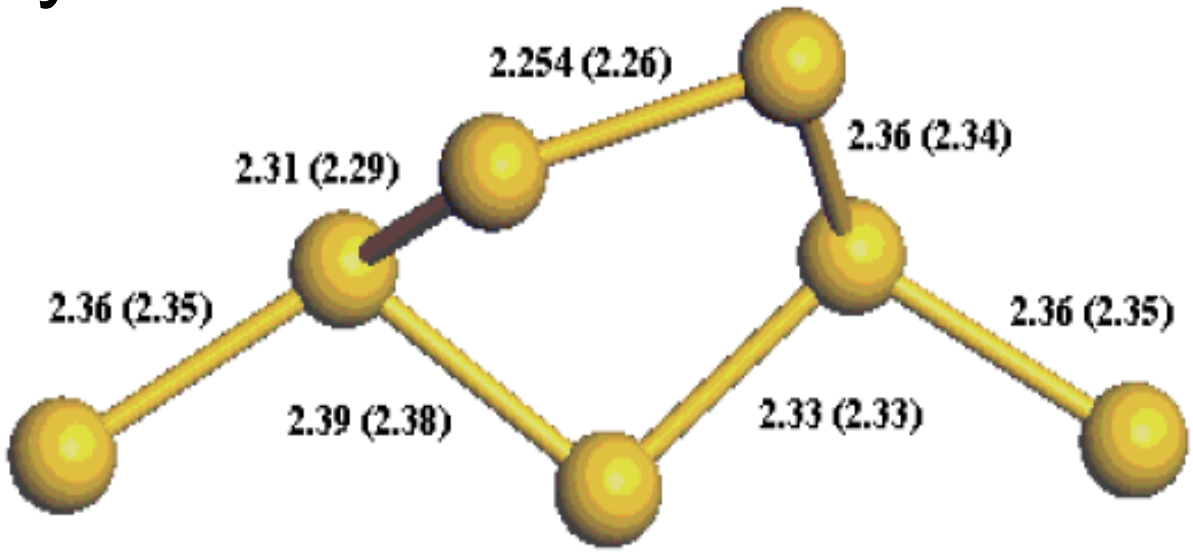
… with added hydrogen passivation

9 layers of Silicon and 7 Å vacuum

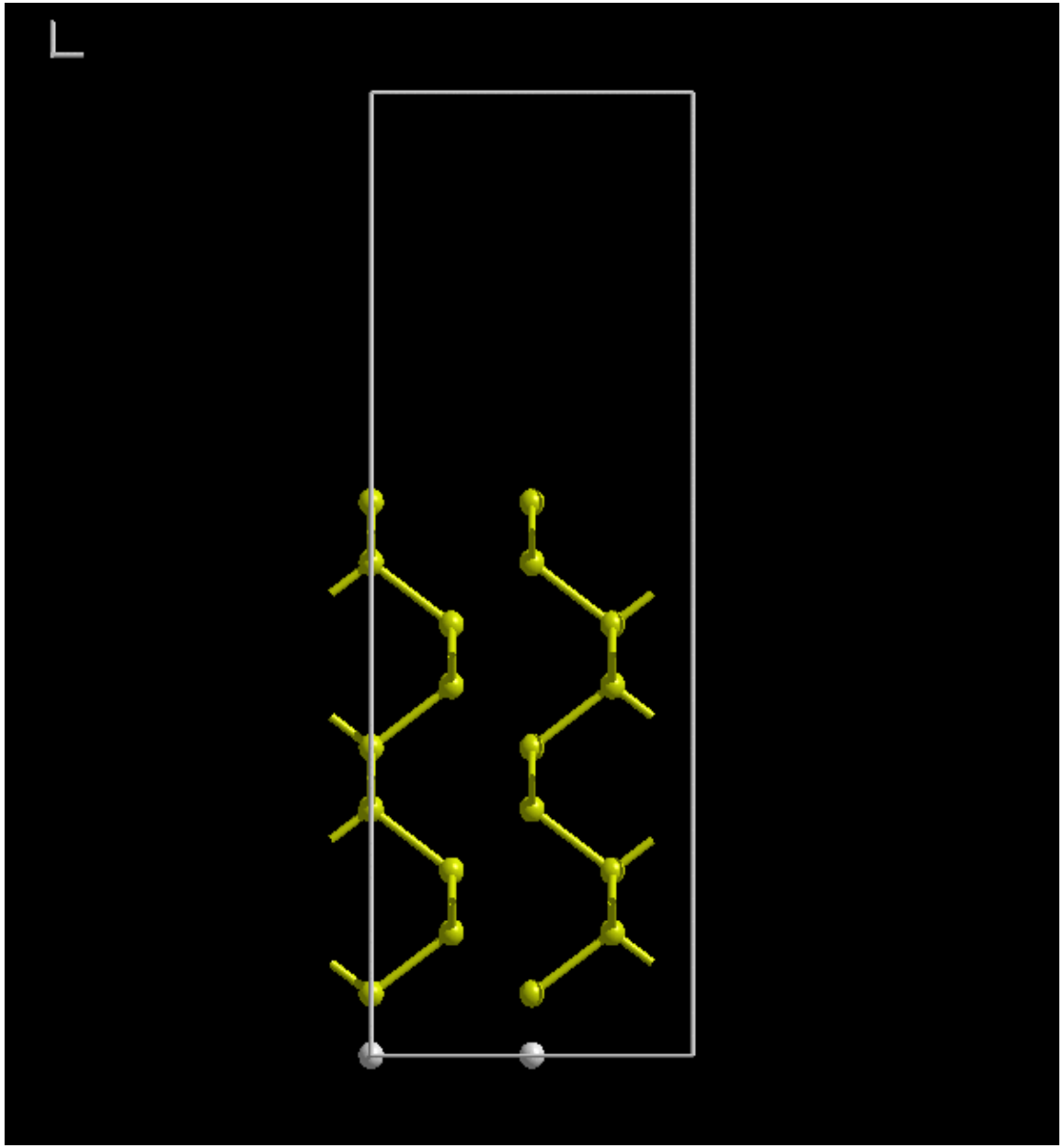
Convergence Tests

- Converge cut-off energy → 370 eV
- Converge k-point sampling → 9 k-points
- Converge number of bulk layers → 9 layers
- Converge vacuum gap → 9 Å
- only then see asymmetric dimerisation:

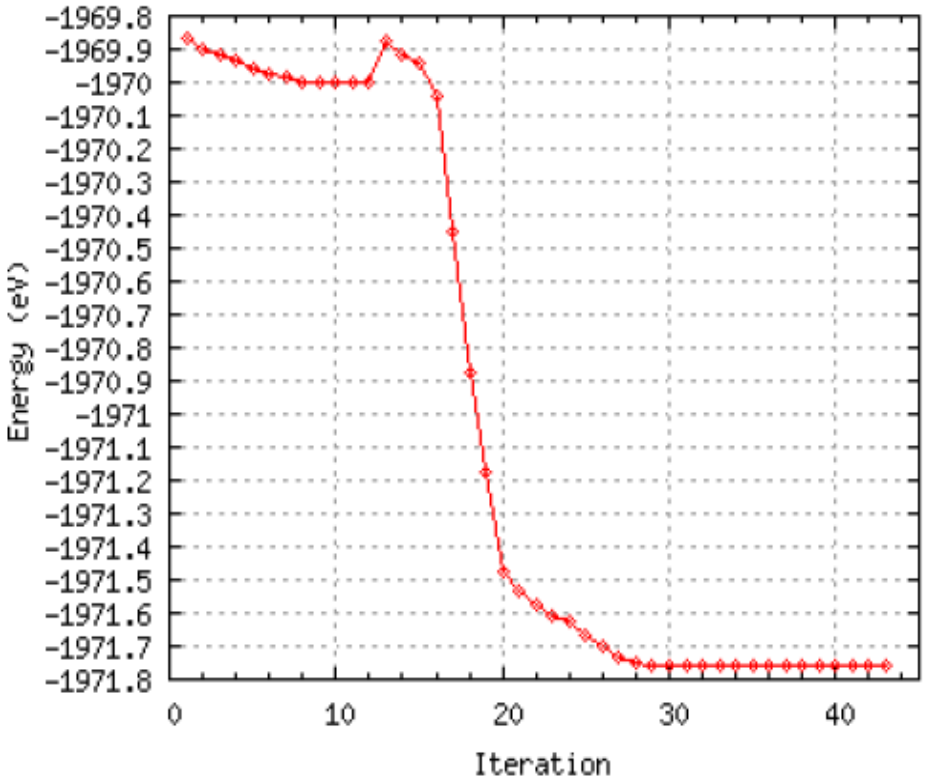
()=best lit. value

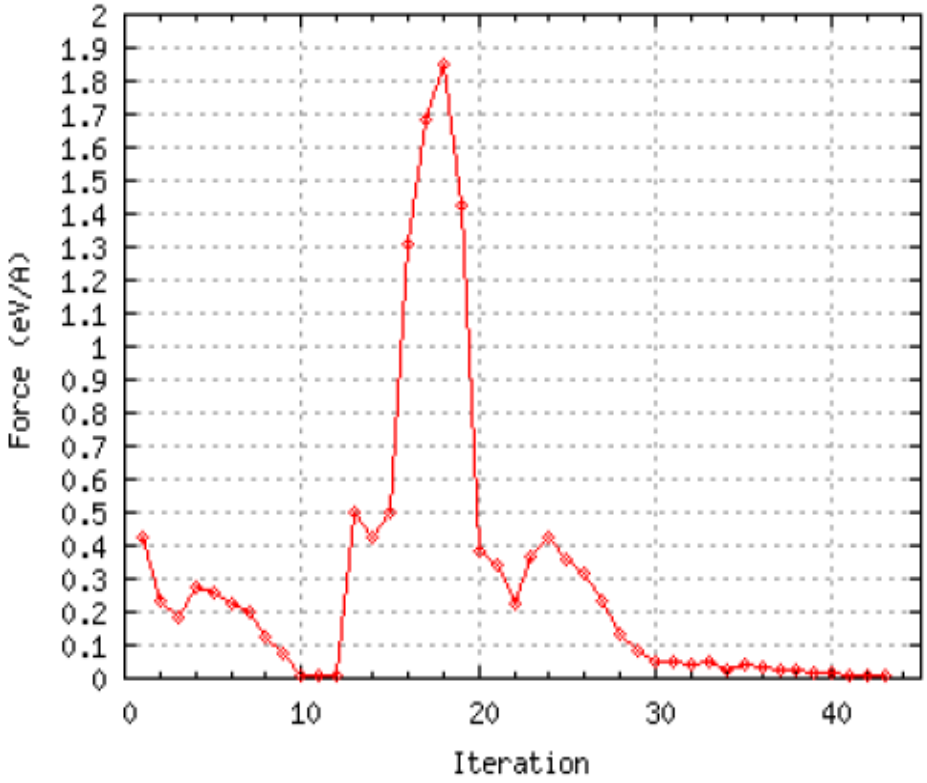


Si(100) – The Movie



Si(100) – BFGS Progress





- Initial slow decrease in energy due to surface layer compression.
- Then small barrier to dimer formation overcome around iteration 14.
- Then rapid energy drop due to dimerisation.
- Final barrier to asymmetric dimerisation overcome around iteration 24.

Summary

Summary

- First principles materials modelling can give highly accurate property values.
  - Need to know how convergence error depends on basis set, kpoints, grid, etc.
  - Degree of convergence required depends on scientific question asked …
  - An over-converged calculation is costly.
  - An under-converged calculation is useless.
  - Be systematic: use automated scripts such as *castepconv*

References

- Anne E Mattson *et al, "Designing meaningful density functional theory calculation in materials science - a primer"* Model. Sim. Mater. Sci Eng. **13** R1-R31 (2005)
- M Leslie and M Gillan, "*The energy and elastic dipole tensor of defects in ionic crystals calculated by the supercell method*", J. Phys. Cond. Mat. **18**, 973 (1985)
- G Makov and MC Payne, "*Periodic boundary conditions in ab initio calculations*", Phys. Rev. B **51**, 4014 (1995)
- MIJ Probert and MC Payne, "*Improving the convergence of defect calculations in supercells: An ab initio study of the neutral silicon vacancy*", Phys. Rev. B **67** 075204 (2003)