Lecture 5 – DFT practicalities

Thomas Durrant
(thomas.durrant.14@ucl.ac.uk)

TYC Materials Modeling Course

Overview

1. SCF convergence
2. Geometry optimisation
3. Exchange correlation functionals

SCF convergence

Solving the KS equations

Finding solutions of the Kohn-Sham equations is a chicken and egg problem.

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{\mathrm{ks}}\left(\mathbf{r}\right)\right)\phi_i\left(\mathbf{r}\right) = \varepsilon_i\phi_i\left(\mathbf{r}\right) \tag{1}$$

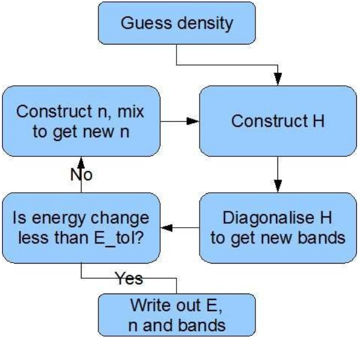
The Kohn-Sham potential v_{ks} depends on the density $n\left(\mathbf{r}\right)$. So if we find the $n\left(\mathbf{r}\right)$ that minimises the energy for a particular $v_{\mathrm{ks}}\left(\mathbf{r}\right)$, $v_{\mathrm{ks}}\left(\mathbf{r}\right)$ will change as a result.

We seek a self consistent value of $n\left(\mathbf{r}\right)$, such that the energy is minimised and $v_{\mathrm{ks}}\left(\mathbf{r}\right)$ remains unchanged.

Self consistent field (SCF)

With the density n we can calculate the Hamiltonian $\hat{H}$, and then solve to get the density n.

Will the input and output densities be the same?



Guess density

Construct n, mix to get new n

Construct H

No

Is energy change less than E_tol?

Diagonalise H to get new bands

Yes

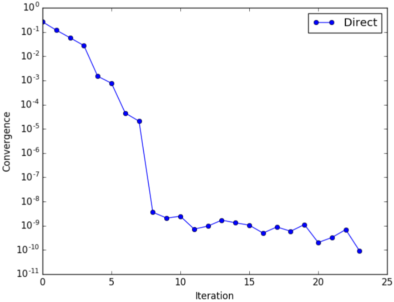
Write out E, n and bands

Hasnip, Phil. ”Bands-parallelism in Castep A dCSE Project.” (2008).

Density mixing

The simplest approach is just to mix the input and output densities

$$n^{\text{new}}\left(\mathbf{r}\right) = \frac{1}{2}\left[n^{\text{input}}\left(\mathbf{r}\right) + n^{\text{output}}\left(\mathbf{r}\right)\right] \tag{2}$$

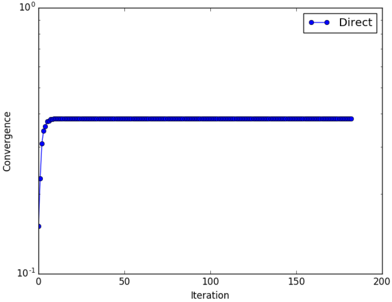


Example of direct mixing for a 2x2x2 supercell of pristine MgO

Density mixing

Not guaranteed to converge and doesn't work for more complex examples due to **charge sloshing instabilities**

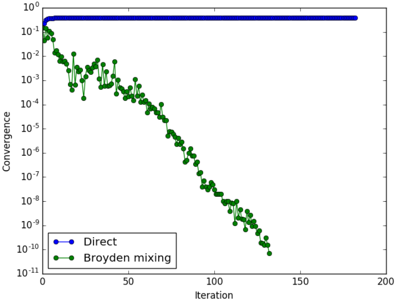
$$n^{\text{new}}(\mathbf{r}) = \frac{1}{2}\left[n^{\text{input}}(\mathbf{r}) + n^{\text{output}}(\mathbf{r})\right] \tag{3}$$



Example of direct mixing for a 2x2x2 MgO containing a V_O^+ defect

Density mixing

More complex mixing algorithms keep a record of all densities calculated and mix many densities together. **Broyden** and **Pulay** mixing are the most commonly used algorithms



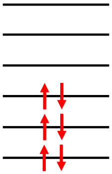
Example of direct mixing for a 2x2x2 MgO containing a V_O^+ defect

Smearing

$$n(\mathbf{r}) = \sum_i O_i(T)\,|\psi_i|^2 \tag{4}$$

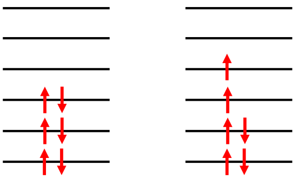
- If a band lies near to the Fermi level it can be fully occupied in one iteration and fully deoccupied in the next, leading to quite unstable behavior. This is always going to happen for metals
- One solution is to allow partial occupancy of bands near the Fermi level, normally using the Fermi-Dirac distribution to fill them
- In this case, a temperature is used to define the distribution. This isn't a physical temperature, but represents a convergence parameter where we are interested in the $T \rightarrow 0$ limit
- This can help with SCF convergence in other contexts e.g. when there are degenerate levels which are not full

Spin state



Singlet
(**S**=0)

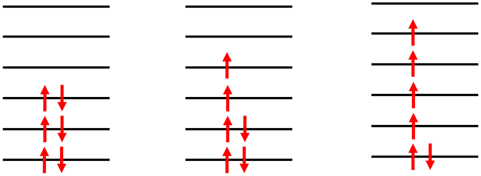
Spin state



Singlet
(**S**=0)

Triplet
(**S**=2)

Spin state



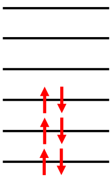
Singlet
(**S**=0)

Triplet
(**S**=2)

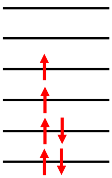
Quintet
(**S**=4)

Spin state

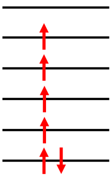
Even electron number



Singlet
(**S**=0)

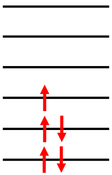


Triplet
(**S**=2)



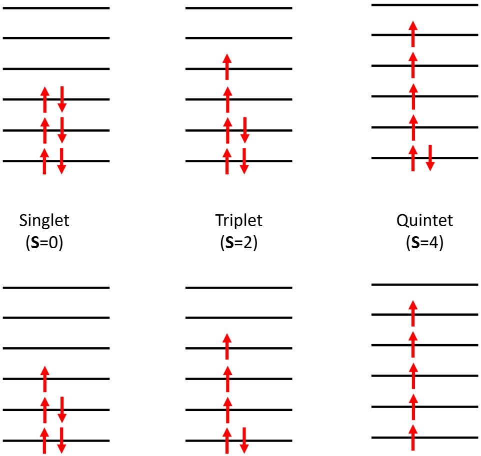
Quintet
(**S**=4)

Odd electron number



Doublet
(**S**=1)

Spin state



Even electron number

Singlet
(**S**=0)

Triplet
(**S**=2)

Quintet
(**S**=4)

Odd electron number

Doublet
(**S**=1)

Quartet
(**S**=3)

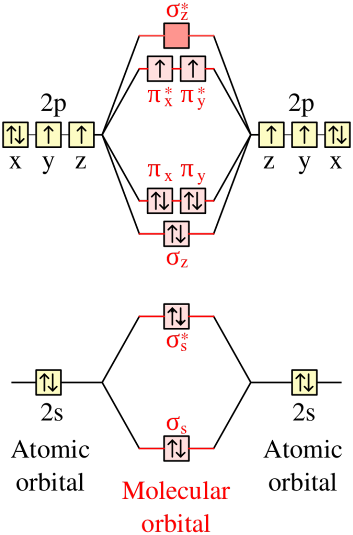
Sextuplet
(**S**=5)

Notes on spin

- You should expect to get integer spin state. Fractional spin states are most commonly seen when something is wrong
- A DFT code will normally guess that systems with an even number of electrons are singlets and that systems with odd numbers of electrons are doublets
- Although there are methods that will optimise the spin state, they are not very good and normally get stuck on the initial guess. You will normally need to try higher spin states by hand if you think they might be lower in energy.

One example

Both atomic and molecular oxygen have a triplet ground state



An example with castep

```
PARAM FILE
task               singlePoint

charge             0.0

spin_polarized     True
spin               0.0

xc_functional      pbe

cut_off_energy     800.0

elec_method        dm
mixing_scheme      Broyden

max_scf_cycles     100

fix_occupancy      False
smearing_scheme    FermiDirac
smearing_width     300.0 K

elec_energy_tol    1e-10

continuation       Default
```

```
CELL FILE
%block lattice_abc
  3.9992 3.9992 3.9992
  90    90    90
%endblock lattice_abc

%block species_pot
    F  NCP
    Li NCP
%endblock species_pot

%block positions_frac
     F 0.5 0.5 0.5
     F 0.5 0.0 0.0
     F 0.0 0.5 0.0
     F 0.0 0.0 0.5
     Li 0.0 0.0 0.0
     Li 0.0 0.5 0.5
     Li 0.5 0.0 0.5
     Li 0.5 0.5 0.0
 %endblock positions_frac

kpoint_mp_grid 10 10 10
symmetry_generate
```

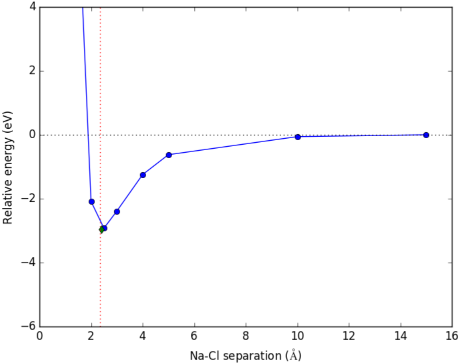
An example with castep

```
OUTPUT
***

Calculating total energy with cut-off of  800.000 eV.
------------------------------------------------------------------------ <-- S
SCF loop      Energy           Fermi           Energy gain       Timer   <-- S
                               energy          per atom          (sec)   <-- S
------------------------------------------------------------------------ <-- S
Initial  -3.42723677E+003  0.00000000E+000                         3.67  <-- S
      1  -3.39578702E+003 -2.92342352E+000  -3.93121806E+000       3.92  <-- S
      2  -3.39756307E+003 -2.90081175E+000   2.22006128E-001       4.13  <-- S
      3  -3.39663167E+003  1.14915823E+000  -1.16425420E-001       4.42  <-- S
      4  -3.39524930E+003 -3.15806975E-001  -1.72795993E-001       4.72  <-- S
      5  -3.39531163E+003 -2.28144440E-001   7.79138978E-003       5.04  <-- S
      6  -3.39531477E+003 -1.82010920E-001   3.91929501E-004       5.34  <-- S
      7  -3.39531479E+003 -1.85547434E-001   2.74457074E-006       5.67  <-- S
      8  -3.39531479E+003 -1.85890679E-001   4.91507229E-007       5.98  <-- S
      9  -3.39531479E+003 -1.85836244E-001   3.19157996E-008       6.30  <-- S
     10  -3.39531479E+003 -1.96826016E-001   1.11551466E-008       6.62  <-- S
     11  -3.39531479E+003 -1.96808343E-001   7.25588684E-010       6.90  <-- S
     12  -3.39531479E+003 -1.96810548E-001   9.64519125E-010       7.15  <-- S
     13  -3.39531479E+003 -1.96812737E-001   5.11890224E-011       7.39  <-- S
     14  -3.39531479E+003 -1.96810682E-001   1.78847387E-012       7.63  <-- S
------------------------------------------------------------------------ <-- S

Integrated Spin Density     =    0.274386E-14 hbar/2
Integrated |Spin Density|   =    0.143708E-05 hbar/2

Final energy, E             =  -3395.314791301     eV

***
```

Geometry optimisation

Geometry Optimisation

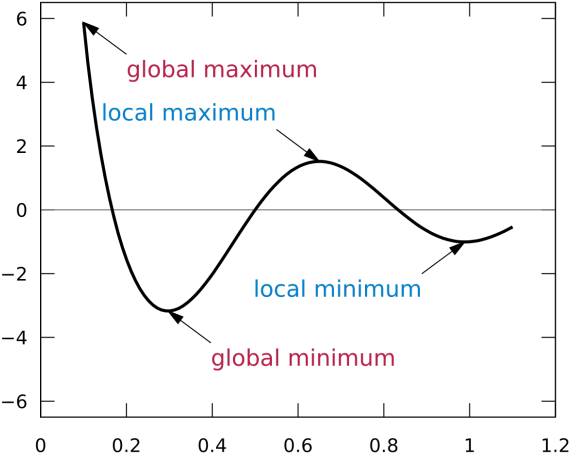
We can either calculate energies given input atomic positions, or we can geometry optimise to find low energy structures



For a NaCl dimer , the only structural parameter is the separation. The blue points are energy calculations for different separations, the green diamond the result of a geometry optimisation.

Geometry Optimisation

Geometry optimisation only finds local minima, not global minima



Geometry Optimisation

These local minima can represent allotropes, which are stable. Carbon serves as an example. A geometry optimisation would return these structures, rather than find the global minimum (diamond)

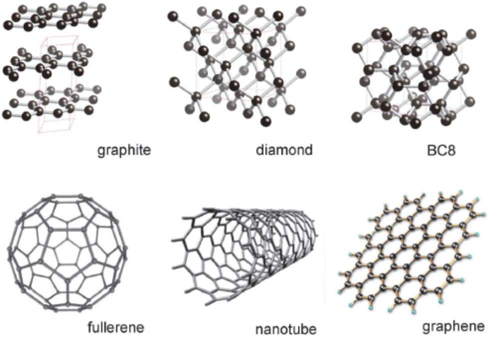


Figure 1. Structures of selected allotropes of carbon.

Oganov, Artem R., et al. ”Structure, bonding, and mineralogy of carbon at extreme conditions.” Reviews in Mineralogy and Geochemistry 75.1 (2013): 47-77.

Geometry optimisation methods

- **Conjugate gradient (CG)** – very robust method that can be quite slow
- **BFGS** – Normally best method to use, quite robust and fast for most problems
- **L-BFGS** – Low memory version of BFGS that can be useful for large systems
- **FIRE** – More modern method, worth trying for difficult cases

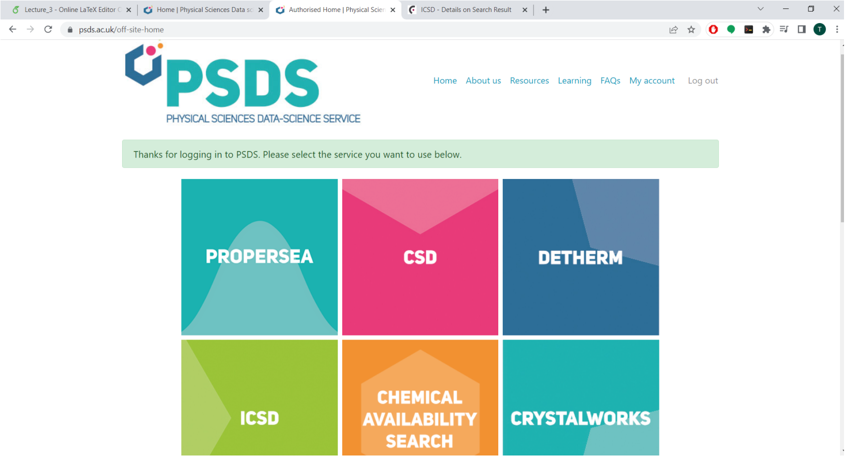
Performance will depend on your application and it is worth testing if you can speed a calculation up by changing algorithm.

Types of geometry optimisation

- **Geometry optimisation** – Minimise the total energy by moving atoms and minimising forces
- **Cell optimisation** – Minimise the stress tensor by modifying the cell vectors (lengths and angles)
- **Both** – Attempt both kinds of optimisation at the same time

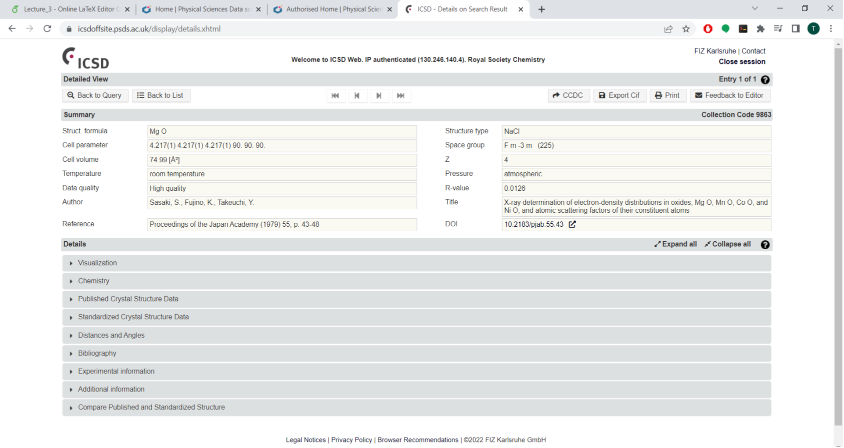
Different DFT codes treat geometry optimisation and cell optimisation differently, some will treat them as different types of calculation (cp2k) and others as types of geometry optimisation (Castep)

Comparison with experiment



Thanks for logging in to PSDS. Please select the service you want to use below.

Comparison with experiment



Welcome to ICSD Web. IP authenticated (130.246.140.4). Royal Society Chemistry

Detailed View Entry 1 of 1

Summary Collection Code 9863

Struct. formula	Mg O	Structure type	NaCl
Cell parameter	4.217(1) 4.217(1) 4.217(1) 90. 90. 90.	Space group	F m -3 m (225)
Cell volume	74.99 [Å³]	Z	4
Temperature	room temperature	Pressure	atmospheric
Data quality	High quality	R-value	0.0126
Author	Sasaki, S.; Fujino, K.; Takeuchi, Y.	Title	X-ray determination of electron-density distributions in oxides, Mg O, Mn O, Co O, and Ni O, and atomic scattering factors of their constituent atoms
Reference	Proceedings of the Japan Academy (1979) 55, p. 43-48	DOI	10.2183/pjab.55.43

Details

- Visualization
- Chemistry
- Published Crystal Structure Data
- Standardized Crystal Structure Data
- Distances and Angles
- Bibliography
- Experimental information
- Additional information
- Compare Published and Standardized Structure

Legal Notices | Privacy Policy | Browser Recommendations | ©2022 FIZ Karlsruhe GmbH

Exchange correlation functionals

Jacob’s ladder of DFT functionals

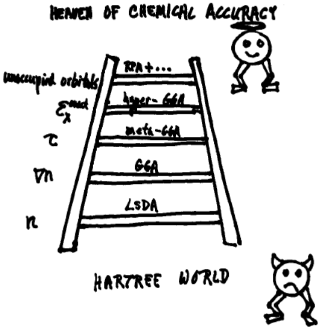


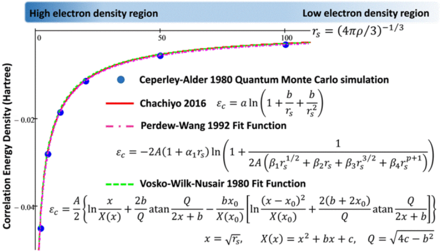
Figure 2. The Jacob’s ladder of density functional approximations to the exchange-correlation energy adds local ingredients successively, leading up in five steps from the Hartree world ($E_{xc} = 0$) of weak or no chemical bonding to the heaven of chemical accuracy (with errors in energy differences of order 1 kcal/mol=0.0434 eV).

Perdew, John P., et al. ”Some fundamental issues in ground-state density functional theory: A guide for the perplexed.” Journal of chemical theory and computation 5.4 (2009): 902-908.

Local density approximation (LDA)

XC energy is locally approximated with the value of a homogeneous electron gas of the same density. The computationally cheapest.

$$E_x^{\mathrm{LDA}}[n] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} \int n(\mathbf{r})^{\frac{4}{3}}\, d\mathbf{r} \qquad (5)$$



Only rung of the ladder that doesn't include any empirical parameters

GGAs

e.g. PBE, PW91, BLYP

$$E_{\mathrm{x}}^{\mathrm{GGA}} = E_{\mathrm{x}}^{\mathrm{GGA}}\left[n\left(\mathbf{r}\right), \nabla n\left(\mathbf{r}\right)\right] \tag{6}$$

- Negligible computational cost increase on LDA, but greater range of available parameters
- The PBE functional is generally considered to be the best general purpose GGA
- Other GGAs might be better for a specific purpose, but tend to be weaker in other areas
- There doesn't appear to be much scope to make better GGAs (PBE is from 1996)

Meta-GGAs

e.g. TPSS, SCAN, M06-L

$$E_{\mathrm{x}}^{\mathrm{mGGA}} = E_{\mathrm{x}}^{\mathrm{mGGA}}\left[n(\mathbf{r}), \nabla n(\mathbf{r}), \nabla^2 n(\mathbf{r}), \tau(\mathbf{r})\right] \tag{7}$$

- Quite small computational cost increase on GGA, but making Meta-GGAs that are better than GGAs is incredibly hard. They are often less transferable
- Earlier attempts like the TPSS functional (2003) have not been used much in practice
- Newer options like M06-L (2006) and SCAN (2015) seem to be more promising

Hybrid functionals

$$E_{\mathrm{x}}^{\mathrm{hybrid}} = (a) E_{\mathrm{x}}^{\mathrm{HF}} + (1 - a) E_{\mathrm{x}}^{\mathrm{GGA}} \left[n(\mathbf{r}), \nabla n(\mathbf{r}) \right] \tag{8}$$

e.g. PBE0, B3LYP, HSE

- Use of non-local exchange term is **extremely computationally expensive**. Essentially guaranteed to be the most expensive part of the resultant DFT calculation
- One important choice is how much non-local exchange to use
- Functionals also differ in how they treat long range exchange
- Started out as an empirical method, but now well justified through the adiabatic connection theorem (e.g. see the hybrid functionals section of this review https://journals.aps.org/rmp/abstract/10.1103/RevModPhys.80.3)

B3LYP

B3LYP is an empirical functional that has been fitted to reproduce experimental atomisation energies, ionisation potentials, proton affinities and atomic energies, for a test set of atoms and simple molecules.

$$
\begin{aligned}
E_{\mathrm{xc}}^{\mathrm{B3LYP}} &= E_{\mathrm{x}}^{\mathrm{LDA}} + a_0 \left(E_{\mathrm{x}}^{\mathrm{HF}} - E_{\mathrm{x}}^{\mathrm{LDA}} \right) + a_x \left(E_{\mathrm{x}}^{\mathrm{GGA}} - E_{\mathrm{x}}^{\mathrm{LDA}} \right) && (9) \\
&+ E_{\mathrm{c}}^{\mathrm{LDA}} + a_c \left(E_{\mathrm{c}}^{\mathrm{GGA}} - E_{\mathrm{c}}^{\mathrm{LDA}} \right) && (10)
\end{aligned}
$$

The fit parameters are $a_0{=}0.20$, $a_x{=}0.72$ and $a_c{=}0.81$.

- B3LYP is more reliable for lighter elements, as this was what it was fit to
- The most popular hybrid functional in Chemistry

PBE0

The PBE0 functional uses $\frac{1}{4}$ exact exchange, which is justified as corresponding to MP4 theory.

$$E_{\mathrm{xc}}^{\mathrm{PBE0}} = \frac{1}{4}E_{\mathrm{x}}^{\mathrm{HF}} + \frac{3}{4}E_{\mathrm{x}}^{\mathrm{PBE}} + E_{\mathrm{c}}^{\mathrm{PBE}} \tag{11}$$

- Significantly more costly to apply to solids than B3LYP is to molecules, as it it quite hard to converge with respect to **k**-points. (B3LYP would have the same issue if it was applied to solids)
- Probably the most popular hybrid functional in Physics

HSE

The HSE functional is designed to be better behaved with respect to **k**-point sampling. Long range exact exchange is turned off using the parameter ω.

$$E_{xc}^{HSE} = \frac{1}{4}E_x^{HF,SR}(\omega) + \frac{3}{4}E_x^{PBE,SR}(\omega) + E_x^{PBE,LR} + E_c^{PBE} \tag{12}$$

- The standard value of ω=0.2 is used. The functional would reduce to PBE0 for ω=0
- Screening exchange in this way reduces the computational cost significantly for solids
- Although motivated by computational arguments, HSE does appear to outperform PBE0 for ‘typical semiconductors’

Performance for solids

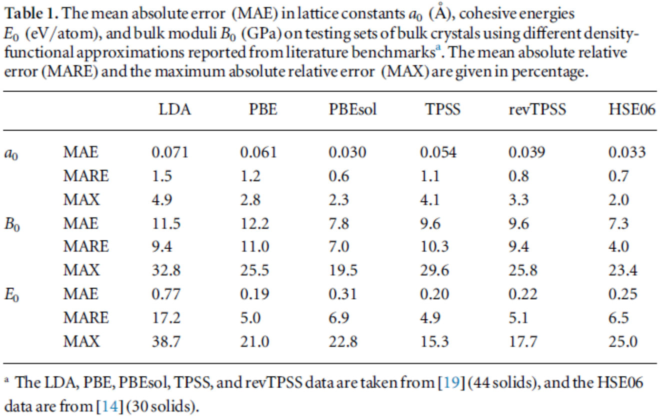


Table 1. The mean absolute error (MAE) in lattice constants a_0 (Å), cohesive energies E_0 (eV/atom), and bulk moduli B_0 (GPa) on testing sets of bulk crystals using different density-functional approximations reported from literature benchmarks[a]. The mean absolute relative error (MARE) and the maximum absolute relative error (MAX) are given in percentage.

		LDA	PBE	PBEsol	TPSS	revTPSS	HSE06
a_0	MAE	0.071	0.061	0.030	0.054	0.039	0.033
	MARE	1.5	1.2	0.6	1.1	0.8	0.7
	MAX	4.9	2.8	2.3	4.1	3.3	2.0
B_0	MAE	11.5	12.2	7.8	9.6	9.6	7.3
	MARE	9.4	11.0	7.0	10.3	9.4	4.0
	MAX	32.8	25.5	19.5	29.6	25.8	23.4
E_0	MAE	0.77	0.19	0.31	0.20	0.22	0.25
	MARE	17.2	5.0	6.9	4.9	5.1	6.5
	MAX	38.7	21.0	22.8	15.3	17.7	25.0

[a] The LDA, PBE, PBEsol, TPSS, and revTPSS data are taken from [19] (44 solids), and the HSE06 data are from [14] (30 solids).

Zhang, Guo-Xu, et al. "Performance of various density-functional approximations for cohesive properties of 64 bulk solids." New Journal of Physics 20.6 (2018): 063020.

Performance for solids

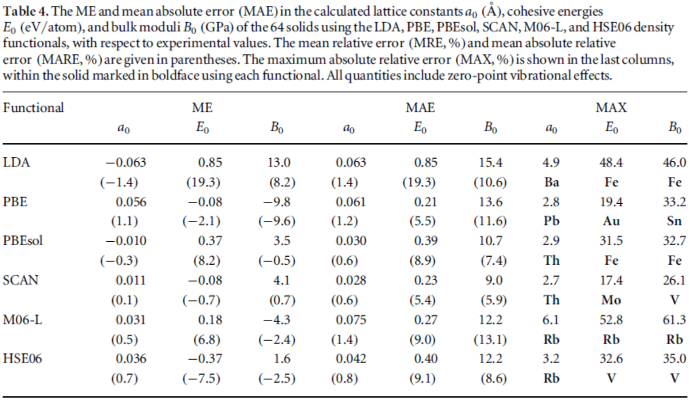
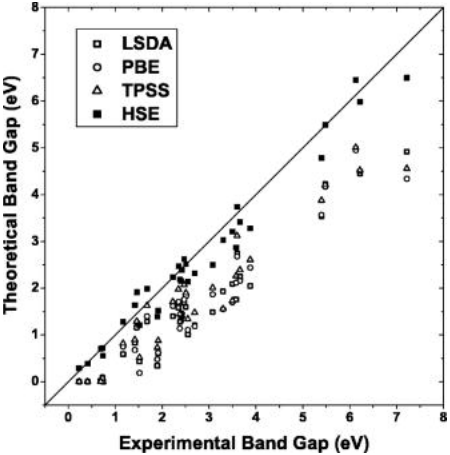


Table 4. The ME and mean absolute error (MAE) in the calculated lattice constants a_0 (Å), cohesive energies E_0 (eV/atom), and bulk moduli B_0 (GPa) of the 64 solids using the LDA, PBE, PBEsol, SCAN, M06-L, and HSE06 density functionals, with respect to experimental values. The mean relative error (MRE, %) and mean absolute relative error (MARE, %) are given in parentheses. The maximum absolute relative error (MAX, %) is shown in the last columns, within the solid marked in boldface using each functional. All quantities include zero-point vibrational effects.

Functional	ME			MAE			MAX		
	a_0	E_0	B_0	a_0	E_0	B_0	a_0	E_0	B_0
LDA	−0.063	0.85	13.0	0.063	0.85	15.4	4.9	48.4	46.0
	(−1.4)	(19.3)	(8.2)	(1.4)	(19.3)	(10.6)	**Ba**	**Fe**	**Fe**
PBE	0.056	−0.08	−9.8	0.061	0.21	13.6	2.8	19.4	33.2
	(1.1)	(−2.1)	(−9.6)	(1.2)	(5.5)	(11.6)	**Pb**	**Au**	**Sn**
PBEsol	−0.010	0.37	3.5	0.030	0.39	10.7	2.9	31.5	32.7
	(−0.3)	(8.2)	(−0.5)	(0.6)	(8.9)	(7.4)	**Th**	**Fe**	**Fe**
SCAN	0.011	−0.08	4.1	0.028	0.23	9.0	2.7	17.4	26.1
	(0.1)	(−0.7)	(0.7)	(0.6)	(5.4)	(5.9)	**Th**	**Mo**	**V**
M06-L	0.031	0.18	−4.3	0.075	0.27	12.2	6.1	52.8	61.3
	(0.5)	(6.8)	(−2.4)	(1.4)	(9.0)	(13.1)	**Rb**	**Rb**	**Rb**
HSE06	0.036	−0.37	1.6	0.042	0.40	12.2	3.2	32.6	35.0
	(0.7)	(−7.5)	(−2.5)	(0.8)	(9.1)	(8.6)	**Rb**	**V**	**V**

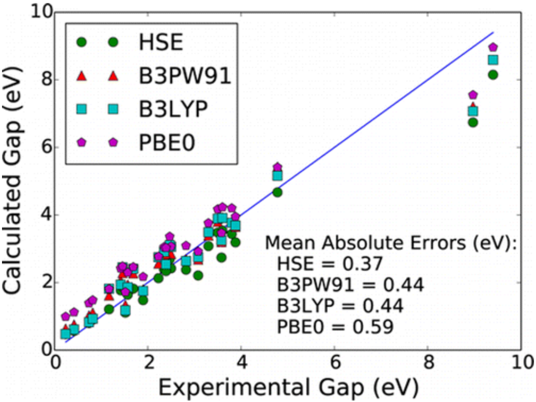
Zhang, Guo-Xu, et al. ”Performance of various density-functional approximations for cohesive properties of 64 bulk solids.” New Journal of Physics 20.6 (2018): 063020.

Performance for solids



Heyd, Jochen, et al. ”Energy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional.” The Journal of chemical physics 123.17 (2005): 174101.

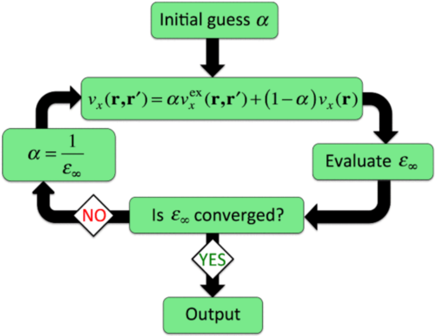
Performance for solids



Garza, Alejandro J., and Gustavo E. Scuseria. "Predicting band gaps with hybrid density functionals." The journal of physical chemistry letters 7.20 (2016): 4165-4170.

Hybrid exchange

The amount of exchange needed to describe a material seems to be related to dielectric screening (itself a material specific property)



Initial guess α

$v_x(\mathbf{r},\mathbf{r}') = \alpha v_x^{\text{ex}}(\mathbf{r},\mathbf{r}') + (1-\alpha) v_x(\mathbf{r})$

Evaluate ε_∞

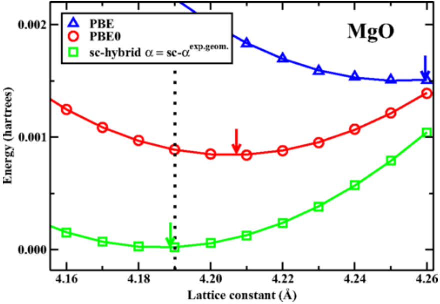
Is ε_∞ converged?

NO

$\alpha = \dfrac{1}{\varepsilon_\infty}$

YES

Output



Skone, Jonathan H., Marco Govoni, and Giulia Galli. ”Self-consistent hybrid functional for condensed systems.” Physical Review B 89.19 (2014): 195112.

Failures of XC functionals

- You can't mix and match XC functionals in a systematic way, so hard to describe systems where different XC functionals are appropriate (e.g. molecule on metal, metal/semiconductor interface)
- Very poor description of long-range interactions, such as van der Waals (empirical corrections e.g. D3)
- Better at molecular binding energies (does chemical reaction happen?) than barriers (how fast does chemical reaction happen?)
- Routes to XC functional improvement seem to require calculations to become really expensive (e.g. double hybrid)

Conclusions

- Density mixing is an important part of practical SCF calculations
- Geometry optimisation can find local minima in the energy landscape, which may or may not be the global minimum
- Local functionals are computationally cheap and good for many properties. Difficult to produce general purpose functionals better than PBE
- Hybrid functionals require significant computational resources, but offer improved band gaps. However, there is a lot of flexibility in how much exchange should be included