

DFTB+ Goes Open Source

Ben Hourahine¹, Bálint Aradi²,
Thomas Frauenheim²



2. University of Bremen, Germany

Motivation and computational cost

Method	Eform. (eV)	Time
Orthogonal TB	3.2	1
Self-Consistent Orthogonal TB	3.2	13
Non-orthogonal TB	3.1	2 (~DFTB)
SC non-orthogonal TB	3.4	19 (~SCC-DFTB)
3 centre non-orthogonal TB	4.3	72
3 centre SC non-orthogonal TB	4.4	87
LCAO LDA DFT	3.8	44
SC LCAO LDA DFT	4.1	110

Si vacancy formation 64 atoms periodic Γ -point minimal basis

A. P. Horsfield and A. M. Bratkovsk, J. Phys. Cond. Mat. 12 R1–R24 (2000).



DFTB

Semi-empirical DFTB modelling (DFT-*lite* descended from the Harris functional)

- **Non-orthogonal tight-binding** minimal sp (d,f,...) basis
- Second order charge and spin contributions – behaves like (S)GGA/LDA
- Parameterised integrals/repulsives but no integration after that (Open parameters : <http://www.dftb.org>)
- **Standard DFT-like properties** including thermodynamic levels (Janak's theorem works), vibrational modes/intensities/Raman x-sections, ...
- Time dependent and transport extensions (following DFT)
- **Simple picture of what's happening** – Mulliken populations actually mean something

J. Phys. Chem. A 111, issue 26 (2007)
pssb 249 issue 2 (2012)



DFTB+ 17.1



- DFTB1, DFTB2 and DFTB3 for clusters and periodic systems
- Spin polarized calculations with collinear (and non-collinear) spin
- Geometry optimization with constraints (Cartesian)
- Molecular dynamics (NVE, NPH, NVT and NPT ensembles)
- Dispersion correction (van der Waals interaction)
- Plotting of charge distribution, molecular orbitals, etc.
- OpenMP parallelisation

B. Aradi, B. Hourahine, and Th. Frauenheim. DFTB+, a sparse matrix-based implementation of the DFTB method, J. Phys. Chem. A, 111 5678 (2007)

<https://github.com/dftbplus>

<http://www.dftbplus.org/>





- Treatment of s , p , d and f -electrons (including LDA+U, spin orbit coupling)
- External electrical fields, QM/MM coupling via fields
- Time dependent DFTB in the Casida-formulation
- Automatic code validation (autotest system)
- Documentation – user guide, tutorials, developer guide
- LGPL license release (16/06/2017)

B. Aradi, B. Hourahine, and Th. Frauenheim. DFTB+, a sparse matrix-based implementation of the DFTB method, J. Phys. Chem. A, 111 5678 (2007)

<https://github.com/dftbplus>

<http://www.dftbplus.org/>



DFTB – underlying idea

Sum of neutral **confined** atomic densities (Confinement dictated by physical environment)

$$n_0(r) = \sum_{\alpha} n_0^{\alpha} |r - R_{\alpha}|$$

$$E_{\text{tot}} = E_{\text{bs}}[n_0] + E_{\text{rep}}[n_0] + E_{\text{2nd}}[n_0, \delta n^2]$$

$E_{\text{bs}}[n_0]$ and $E_{\text{2nd}}[n_0, \delta n^2]$ **calculated explicitly** (approximations: yes, adjustable parameters: no)

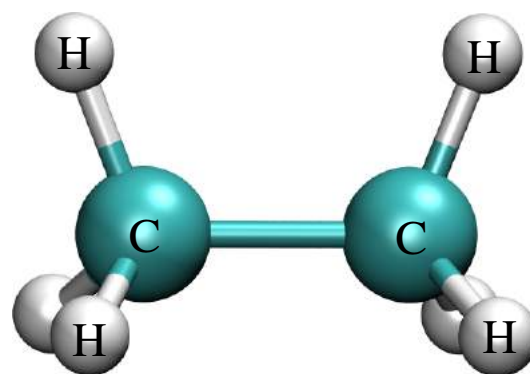
Superposition of pairwise interactions

$$E_{\text{rep}} = \frac{1}{2} \sum_{\alpha} \sum_{\beta \neq \alpha} E_{\text{rep}}^{\alpha\beta} (|R_{\alpha} - R_{\beta}|)$$

Deviation from *ab initio* calculation tabulated as function of distance **in advance**

$$E_{\text{rep}}^{\alpha\beta}(R_{\alpha\beta}) = E_{\text{ab initio}}(R_{\alpha\beta}) - [E_{\text{bs}} + E_2](R_{\alpha\beta})$$

$E_{\text{rep}}[n_0]$ **fitted to *ab initio* calculations**
(corrects errors due to approximations)



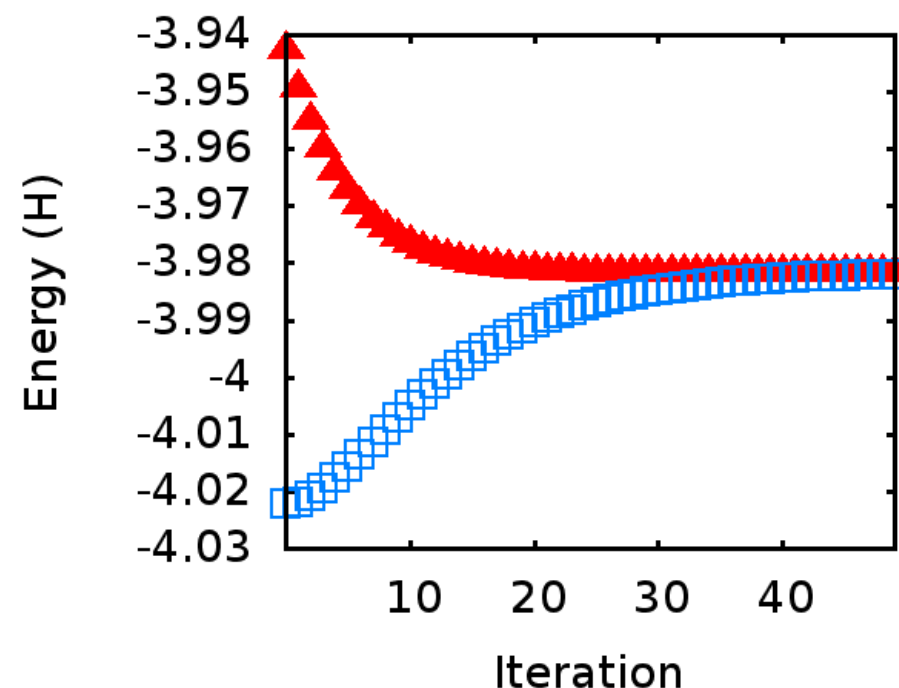
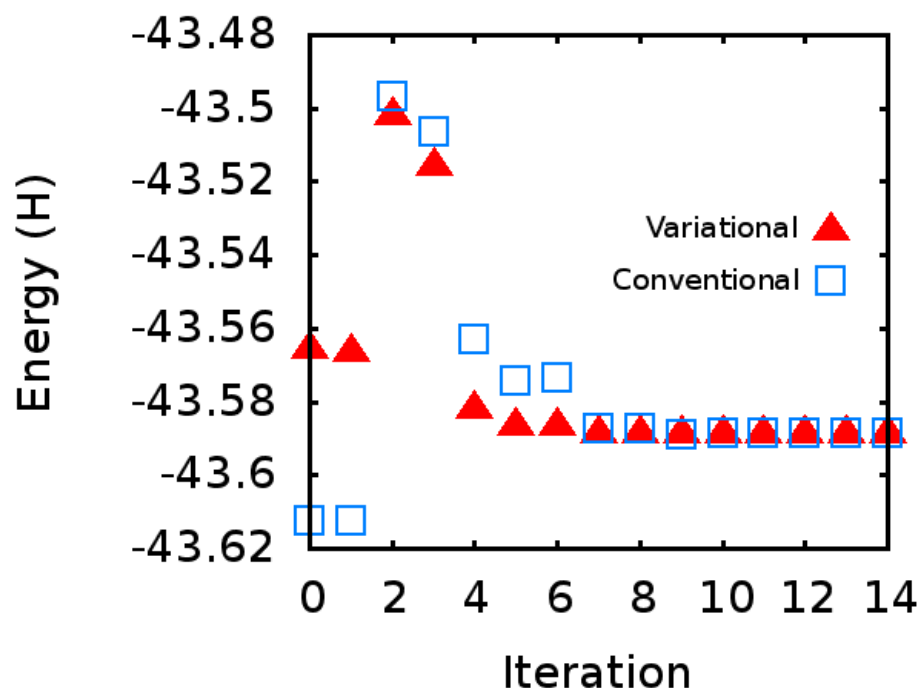
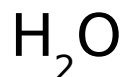
Repulsives needed: $\left\{ \begin{array}{l} \text{C} - \text{C} \\ \text{C} - \text{H} \\ \text{H} - \text{H} \end{array} \right.$

Charge self consistency (DFTB2)

$$E_{tot} = \sum_i f_i \sum_{\mu\nu} c_{i\mu} c_{i\nu} H_{\mu\nu}^{bs} + \frac{1}{2} \sum_{\alpha\beta} \gamma_{\alpha\beta} \Delta q_{\alpha} \Delta q_{\beta} + E_{rep}$$

$$E_{tot} = \sum_i f_i \varepsilon_i - \frac{1}{2} \sum_{\alpha\beta} \gamma_{\alpha\beta} \Delta q_{\alpha}^{in} \Delta q_{\beta}^{out} + E_{rep} \quad - \text{old SCC-DFTB}$$

At self-consistency, $\Delta n_{in} = \Delta n_{out}$, so the two forms agree, but before this :



Hourahine et al. *J. Phys. Chem A* 111, 5671 (2007)

<http://www.dftbplus.org/>

ScotCHEM17



SCC-DFTB (DFTB2) – accuracy

	DFTB	cc-pVDZ	cc-pVTZ	G2
$\text{H}_2 + \text{CH} \equiv \text{CH} \rightarrow \text{CH}_2 = \text{CH}_2$	-38.8	-41.0	-37.9	-40.1
$\text{H}_2 + \text{CH}_2 = \text{CH}_2 \rightarrow \text{CH}_3 - \text{CH}_3$	-37.0	-28.5	-26.5	-30.5
$3\text{H}_2 + \text{HCN} \rightarrow \text{NH}_3 + \text{CH}_4$	-47.9	-47.1	-48.5	-53.7
$\text{H}_2 + \text{CO} \rightarrow \text{H}_2\text{CO}$	+10.9	+0.1	+0.7	+3.1
$2\text{H}_2 + \text{CO} \rightarrow \text{CH}_3\text{OH}$	-7.8	-12.6	-13.1	-15.5
$\text{H}_2 + \text{CH}_3\text{OH} \rightarrow \text{CH}_4 + \text{H}_2\text{O}$	-25.7	-20.6	-24.6	-26.2
$2\text{H}_2 + \text{N}_2 \rightarrow \text{NH}_2 - \text{NH}_2$	+32.7	+31.6	+31.9	+30.7
$\text{H}_2 + \text{NH}_2 - \text{NH}_2 \rightarrow 2\text{NH}_3$	-47.4	-38.3	-43.1	-46.7
$\text{H}_2 + \text{H}_2\text{O}_2 \rightarrow 2\text{H}_2\text{O}$	-81.7	-63.3	-71.1	-82.8
$2\text{H}_2 + \text{CO}_2 \rightarrow \text{H}_2\text{O} + \text{H}_2\text{CO}$	+15.6	+26.6	+21.9	+14.5
$\text{CH}_4 + \text{CO} \rightarrow \text{CH} \equiv \text{CH} + \text{H}_2\text{O}$	+55.9	+51.5	+43.7	+44.3
$\text{CH}_4 + \text{H}_2\text{CO} \rightarrow \text{CH}_2 = \text{CH}_2 + \text{H}_2\text{O}$	+6.2	+10.4	+5.2	+1.1
$\text{CH}_4 + \text{CH}_3\text{OH} \rightarrow \text{CH}_3 - \text{CH}_3 + \text{H}_2\text{O}$	-12.1	-5.4	-7.6	-10.8
$2\text{CH}_4 + \text{N}_2 \rightarrow \text{NH}_2 - \text{NH}_2 + \text{CH}_2 = \text{CH}_2$	+83.3	+75.4	+75.4	+76.6
$\text{CH}_4 + \text{H}_2\text{O}_2 \rightarrow \text{CH}_3\text{OH} + \text{H}_2\text{O}$	-56.0	-42.6	-46.4	-56.6
$2\text{NH}_3 + \text{CH} \equiv \text{CH} \rightarrow \text{NH}_2 - \text{NH}_2 + \text{CH}_2 = \text{CH}_2$	+8.7	-2.8	+5.1	+6.7
$2\text{NH}_3 + \text{CH}_2 = \text{CH}_2 \rightarrow \text{NH}_2 - \text{NH}_2 + \text{CH}_3 - \text{CH}_3$	+10.4	+9.7	+16.6	+16.2
$\text{NH}_3 + \text{HCN} \rightarrow \text{CH}_4 + \text{N}_2$	-33.1	-40.5	-37.4	-37.6
$\text{NH}_3 + \text{CO} \rightarrow \text{HCN} + \text{H}_2\text{O}$	+14.3	+13.8	+10.8	+12.0
$2\text{NH}_3 + \text{H}_2\text{CO} \rightarrow \text{NH}_2 - \text{NH}_2 + \text{CH}_3\text{OH}$	+28.8	+25.5	+29.3	+28.2
$\text{H}_2 + \text{H}_2\text{CO} \rightarrow \text{CH}_3\text{OH}$	-18.7	-12.8	-13.8	-18.6
$\text{Oxirane} + \text{H}_2\text{O} \rightarrow \text{OH} - \text{CH}_2 - \text{CH}_2 - \text{OH}$	-30.9	-20.5	-16.7	-19.6
$\text{Oxirane} + \text{NH}_3 \rightarrow \text{NH}_2 - \text{CH}_2 - \text{CH}_2 - \text{OH}$	-33.1	-20.4	-17.5	-22.6
$\text{HNCO} + \text{H}_2\text{O} \rightarrow \text{NH}_2 - \text{COOH}$	-4.3	-17.3	-11.1	-16.1
$\text{CH}_2 = \text{NH} + \text{CH}_4 + \text{NH}_3 \rightarrow 2\text{CH}_3\text{NH}_2$	-2.7	+1.1	+4.3	-0.6
$\text{H}_2\text{CO} + \text{CH}_4 + \text{H}_2\text{O} \rightarrow 2\text{CH}_3\text{OH}$	+7.1	+7.9	+10.9	+7.6
$\text{HCN} + 2\text{CH}_4 + 2\text{NH}_3 \rightarrow 3\text{CH}_3\text{NH}_2$	+17.4	+14.1	+21.4	+14.6
$\text{CO} + 2\text{CH}_4 + 2\text{H}_2\text{O} \rightarrow 3\text{CH}_3\text{OH}$	+43.6	+28.6	+36.2	+36.8

With respect to G2 set:

mean ave. dev.: 4.3 kcal/mole

Krueger, *et al.*, J. Chem. Phys. 122 (2005) 114110.

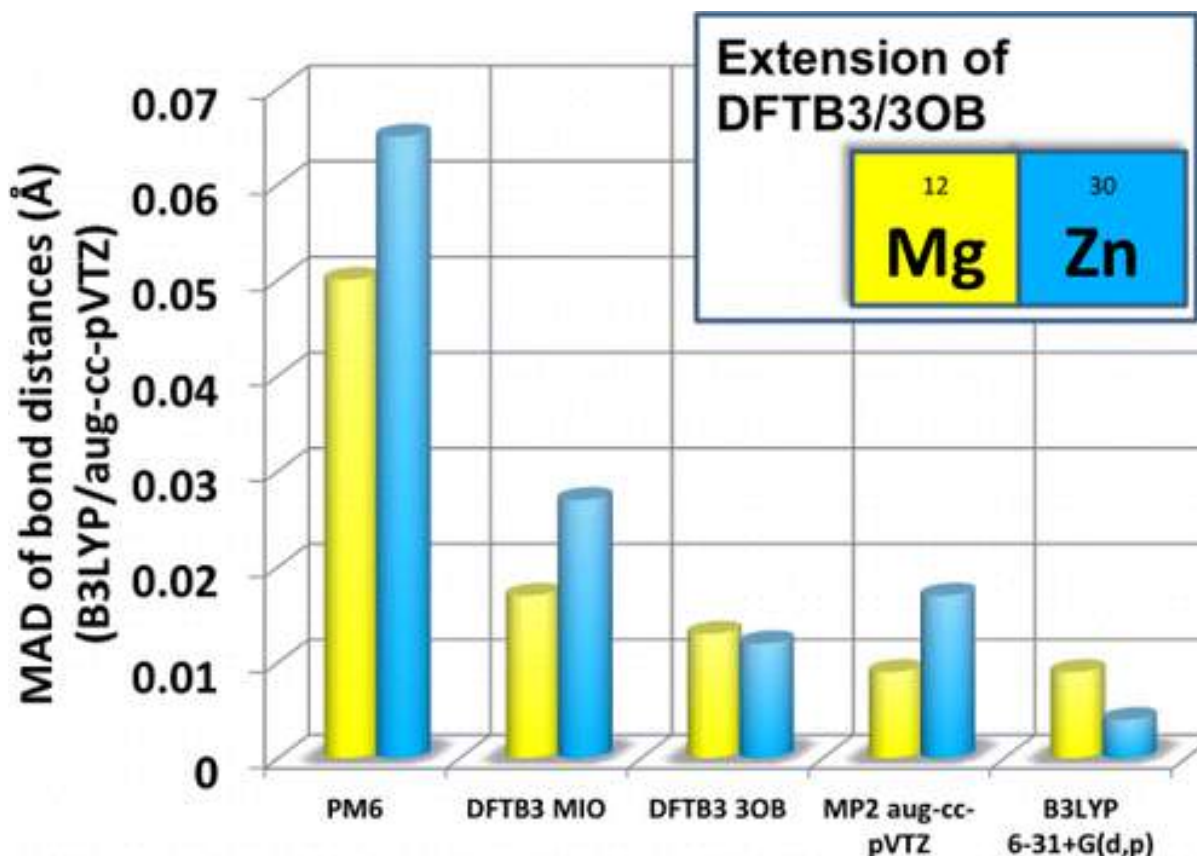
<http://www.dftbplus.org/>

ScotCHEM17



DFTB3 – accuracy

Expand the DFT energy to 3rd order in charge fluctuations



Small molecule
test set

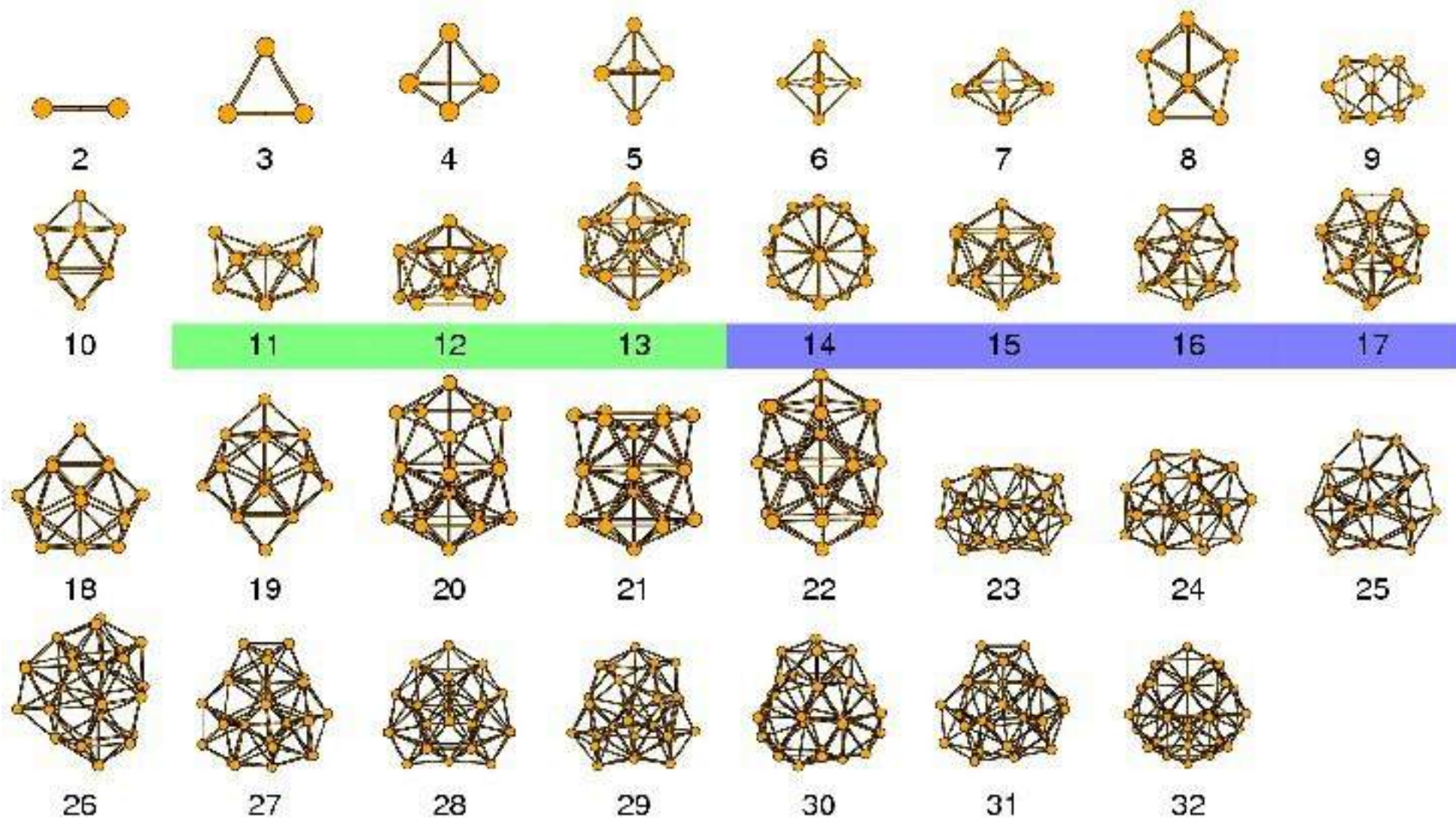
Paper includes
peptides and
other larger
systems at
DFTB3 level

Typical mean absolute deviation (MAD) of ~3–5 kcal/mol

X. Lu, M. Gaus, M. Elstner, Q. Cui, *J. Phys. Chem B* **2015**, 119, 1062–1082.



Spin

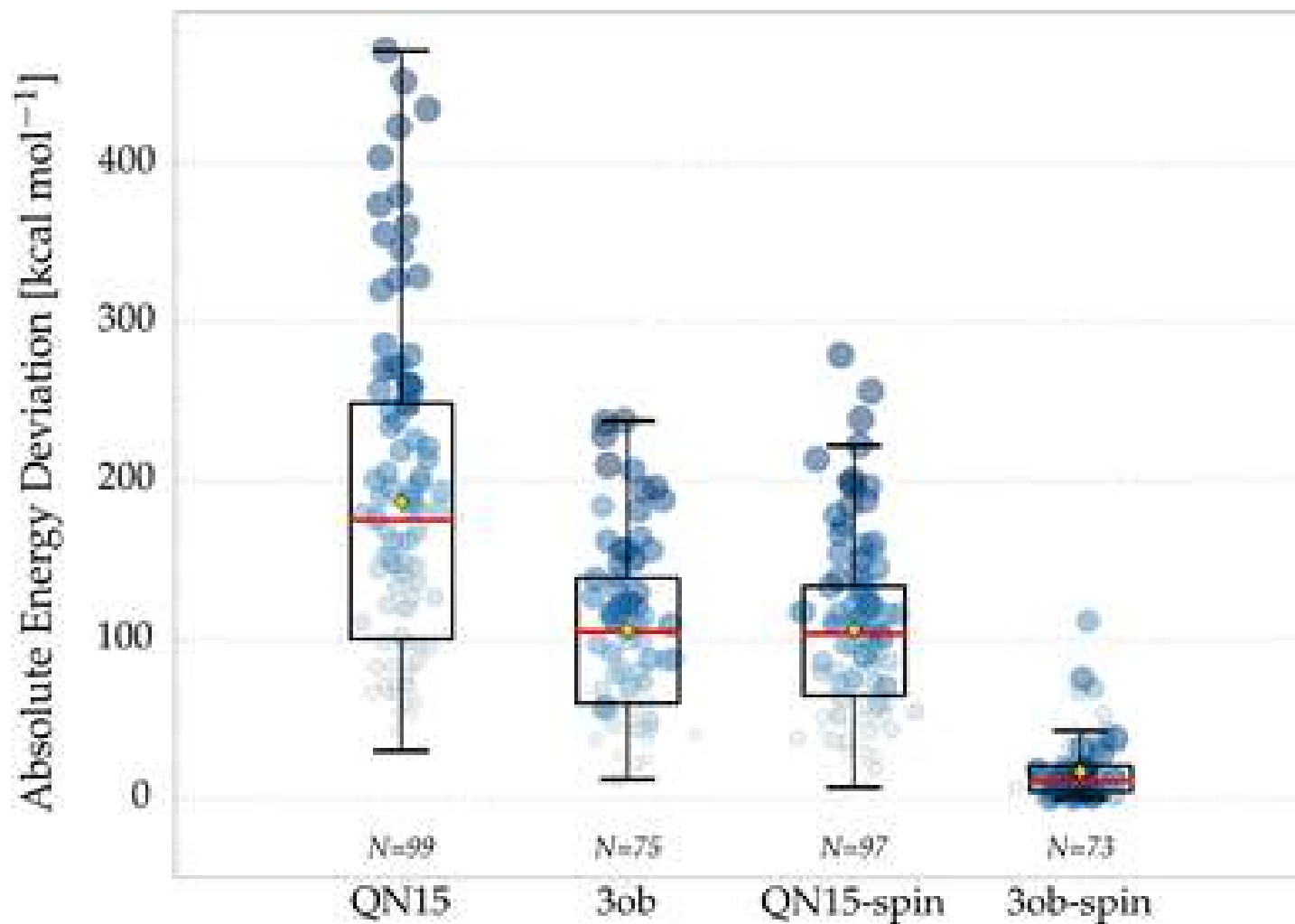


Köhler et al. Chem. Phys 309 23 (2005)

<http://www.dftbplus.org/>

ScotCHEM17

SDFTB2 - accuracy



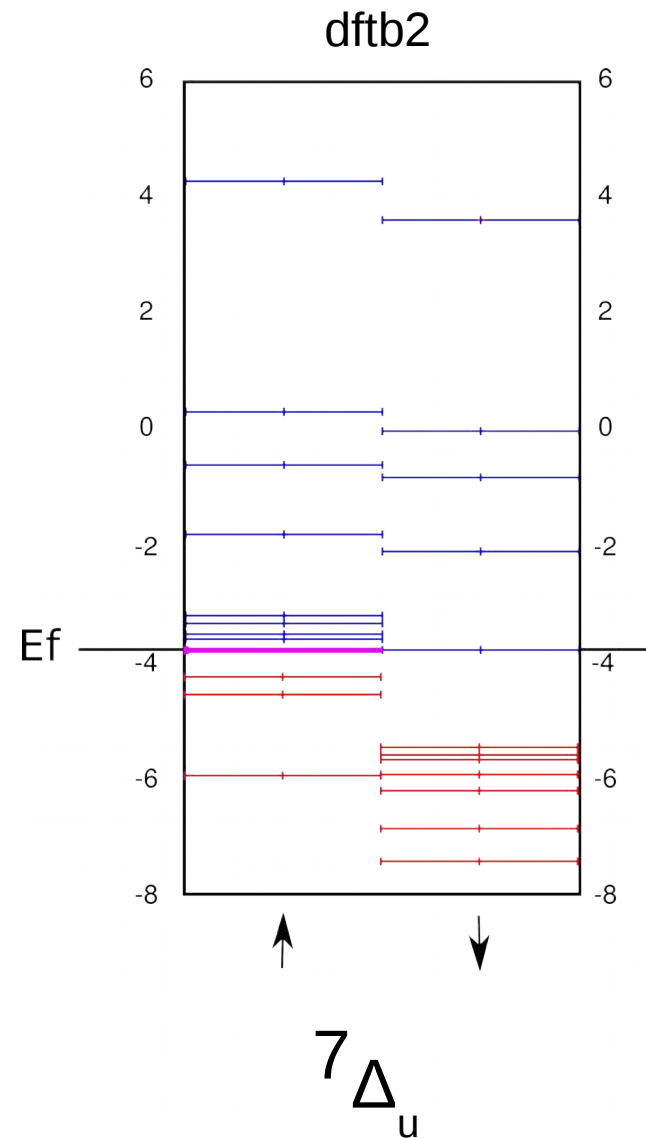
GMTKN30 test set

3ob set technically
intended for
DFTB3 level model

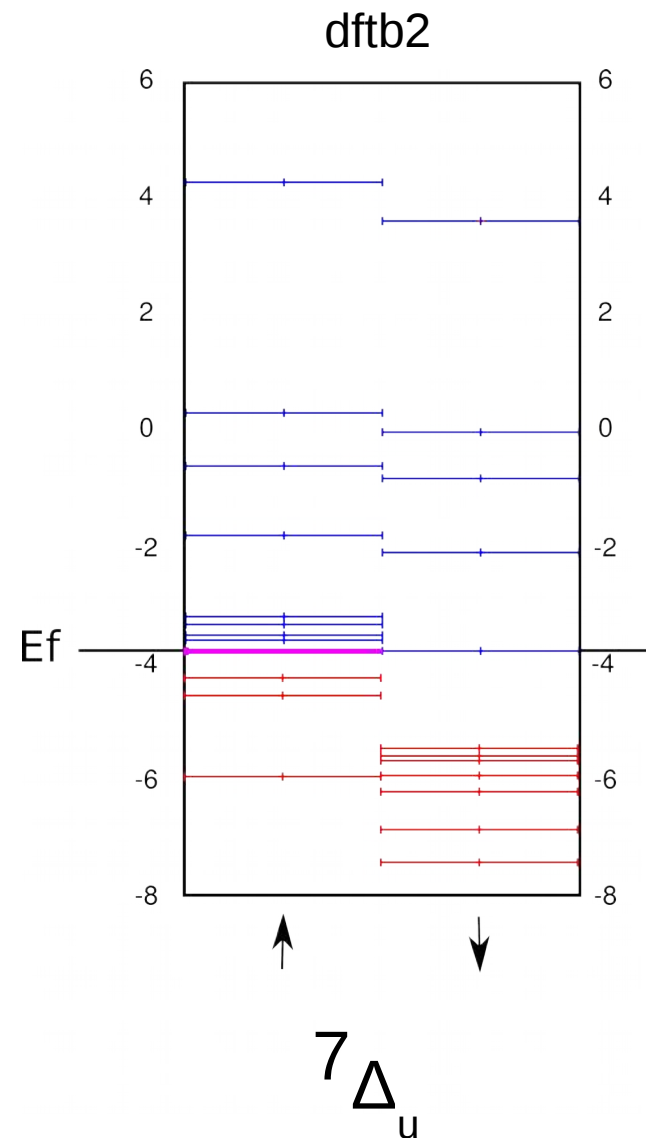
Melix, P., Oliveira, A.F., Rüger, R. et al. Theor Chem Acc (2016) 135: 232.



Fe₂ molecule ground state



Fe₂ molecule ground state



Wrong symmetry

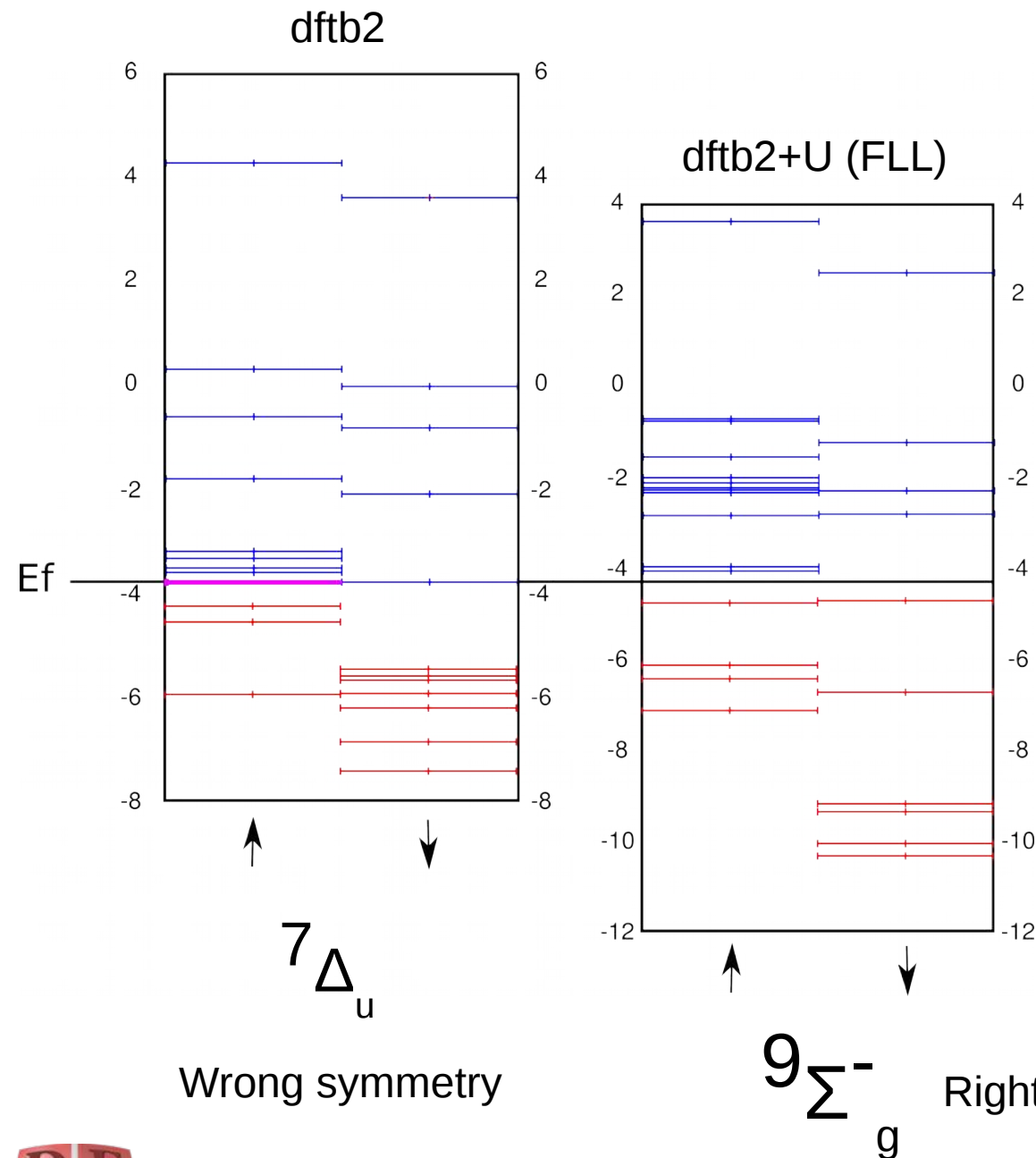
Same behaviour as

PBE see:

“Density Functional Theory in Transition-Metal Chemistry: Relative Energies of Low-Lying States of Iron Compounds and the Effect of Spatial Symmetry Breaking”

Sorkin, Iron, Truhlar, J. Chem. Theory Comput., **4** 307 (2008).

Fe₂ molecule ground state



Same behaviour as
PBE / PBE+U, see:

“Density Functional Theory in
Transition-Metal Chemistry:
Relative Energies of Low-Lying
States of Iron Compounds and
the Effect of Spatial Symmetry
Breaking”

Sorkin, Iron, Truhlar, J. Chem.
Theory Comput., **4** 307 (2008).



NiO

Nickel oxide is the 'famous' example for the failure of mean field functionals in DFT. It's a rock salt wide gap insulator with anti-ferromagnetic ordering of Ni spins (μ).

	μ	E_g	$E_g + \Delta$
SCC-DFTB	0.0	0.0	
spin-SCC-DFTB	0.31	0.0	
DFTB+U	1.42	2.03	4.94
DFTB+pSIC	1.22	1.28	3.81
LSDA-SIC	1.16	~2.86	
LDA+U	1.71	2.73	
B3LYP	1.67		4.1
Expt	1.64-1.90	4.0-4.3	

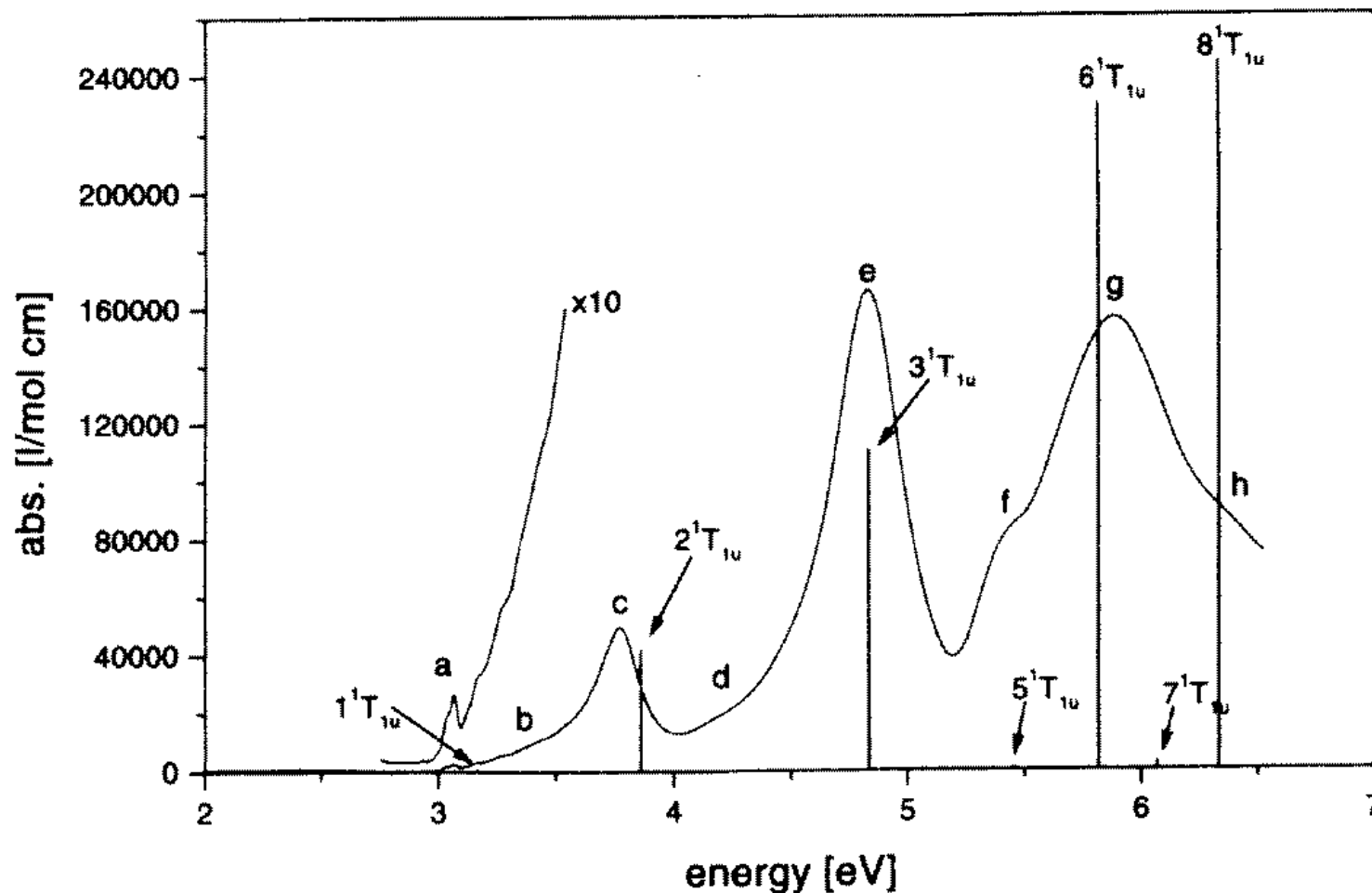


Hourahine et al. J. Phys. Chem A 111, 5671 (2007)

<http://www.dftbplus.org/>

Casida excited states in DFTB

TD-DFTB Spectra of C_{60}



DFTB has been used for Casida-type excitations for ~ 15 years.



Niehaus et al. Phys. Rev. B 63, 085108 (2001)

<http://www.dftbplus.org/>

ScotCHEM17

Spin-orbit – topological insulators

Dirac cones like graphene, but on opposite sides of the whole solid – linked electron momenta and spin for states – only possible to change electron direction for surface states by tunnelling through whole solid.

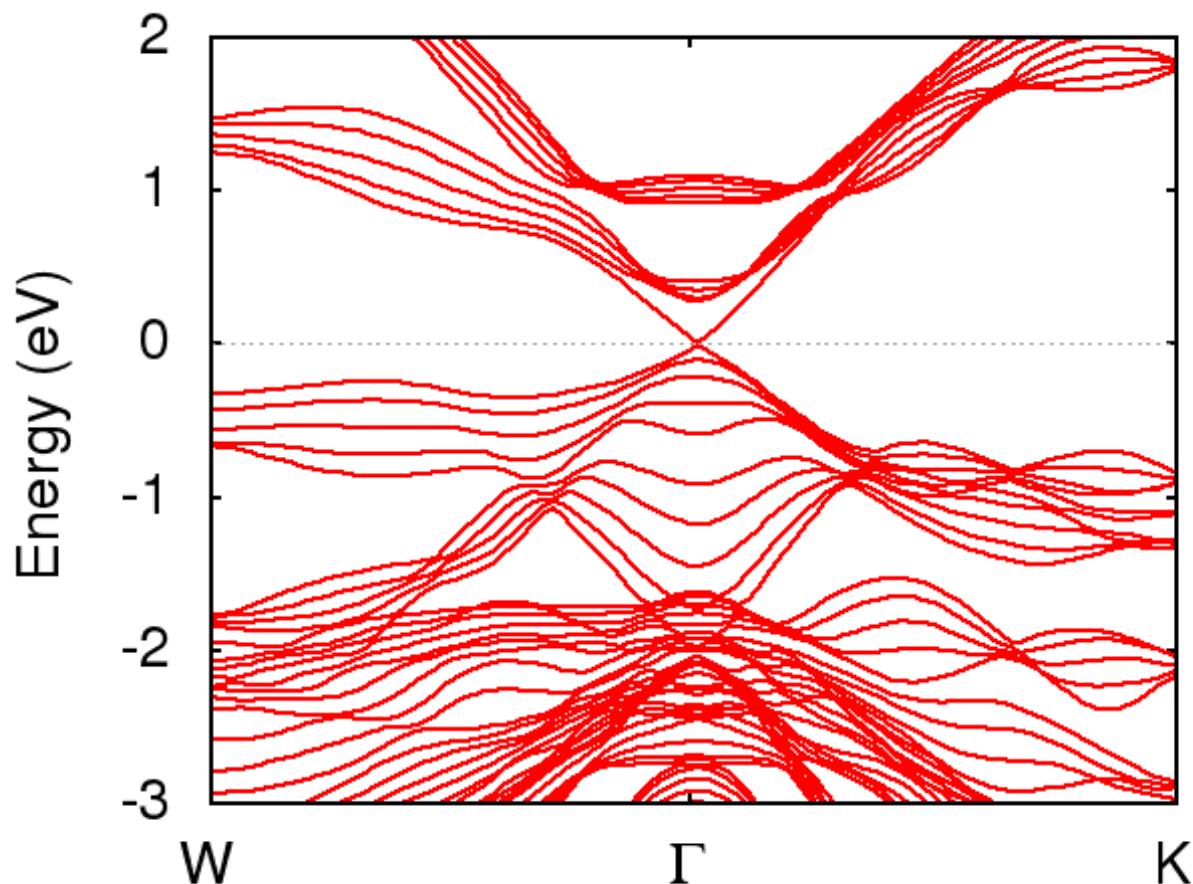


Insulating in the bulk, but with protected metallic surface states.
Consequence of non-trivial topology of wavefunction
→ odd number of band inversions in the bulk electronic states

Kane and Mele, Phys. Rev. Lett. 95, 226801 (2005)

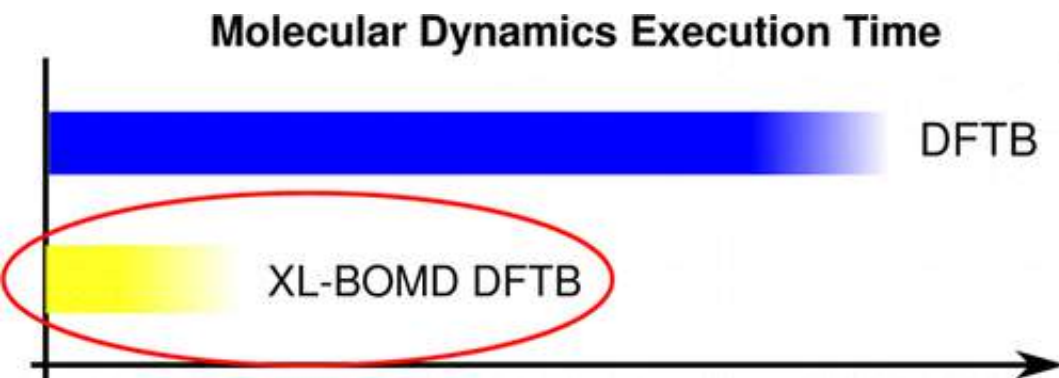
Topological insulator Bi_2Se_2

DFTB⁺
surface
state
calculation
For 3
formula
unit thick
slab - i.e.
primitive
cell +
vacuum



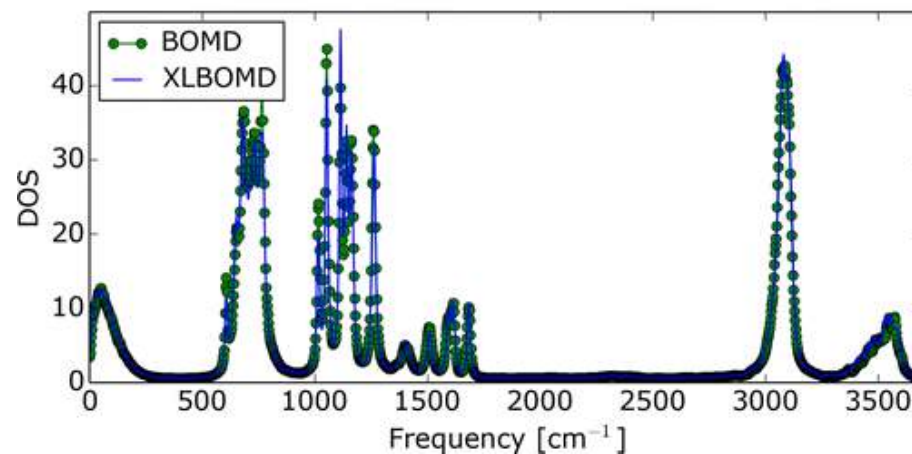
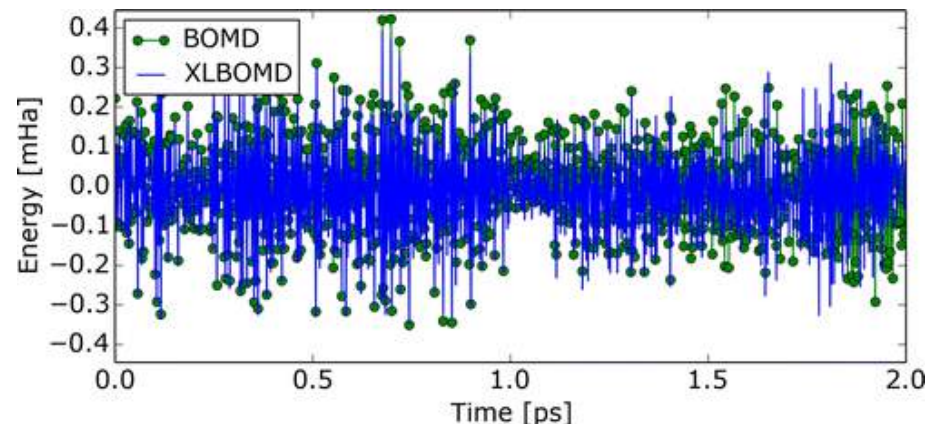
Protected surfaces states as expected...

Fast DFTB2 NVE dynamics



1 SCC step per dynamics step

~10x speed up for DFTB2 MD at the same accuracy



imidazol crystal with 16 imidazol
molecules / supercell

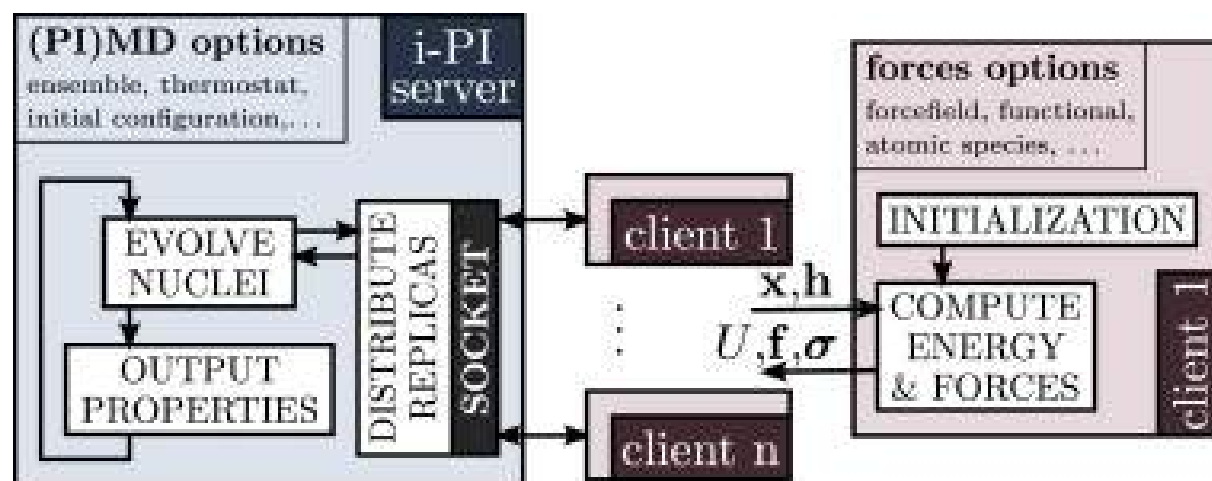
Bálint Aradi, Anders M. N. Niklasson, and Thomas Frauenheim
J. Chem. Theory Comput., 2015, 11 (7), pp 3357–3363

<http://www.dftbplus.org/>

ScotCHEM17



Path integral and REMD dynamics



DFTB+ has an interface to the i-PI open source code

– They support several sampling modes for molecular dynamics:

- Free energy landscapes
- Quantum nuclear dynamics

<https://github.com/i-pi>

Computer Physics Communications

Volume 185, Issue 3, March 2014, Pages 1019-1026

i-PI: A Python interface for ab initio path integral molecular dynamics simulations

Michele Ceriotti, Joshua More and David E. Manolopoulos

Interface to ASE code as well

<https://wiki.fysik.dtu.dk/ase/ase/calculators/dftb.html>



<http://www.dftbplus.org/>

ScotCHEM17

What's next?

DFTB+ 17.1 is the current release (as of today).

17.2

- MPI parallel version re-integrated to main code
- 1D, 2D and helical boundary conditions released
- libDFTB (CHARMM and GROMACS integration)

17.3

- Non-Equilibrium Green's function transport re-integrated to main code
- Coupled perturbed and Sternheimer response modules released

18.1

- Hybrid functionals released
- linear scaling and $O(N^2)$ solvers released
- Δ SCF excitations

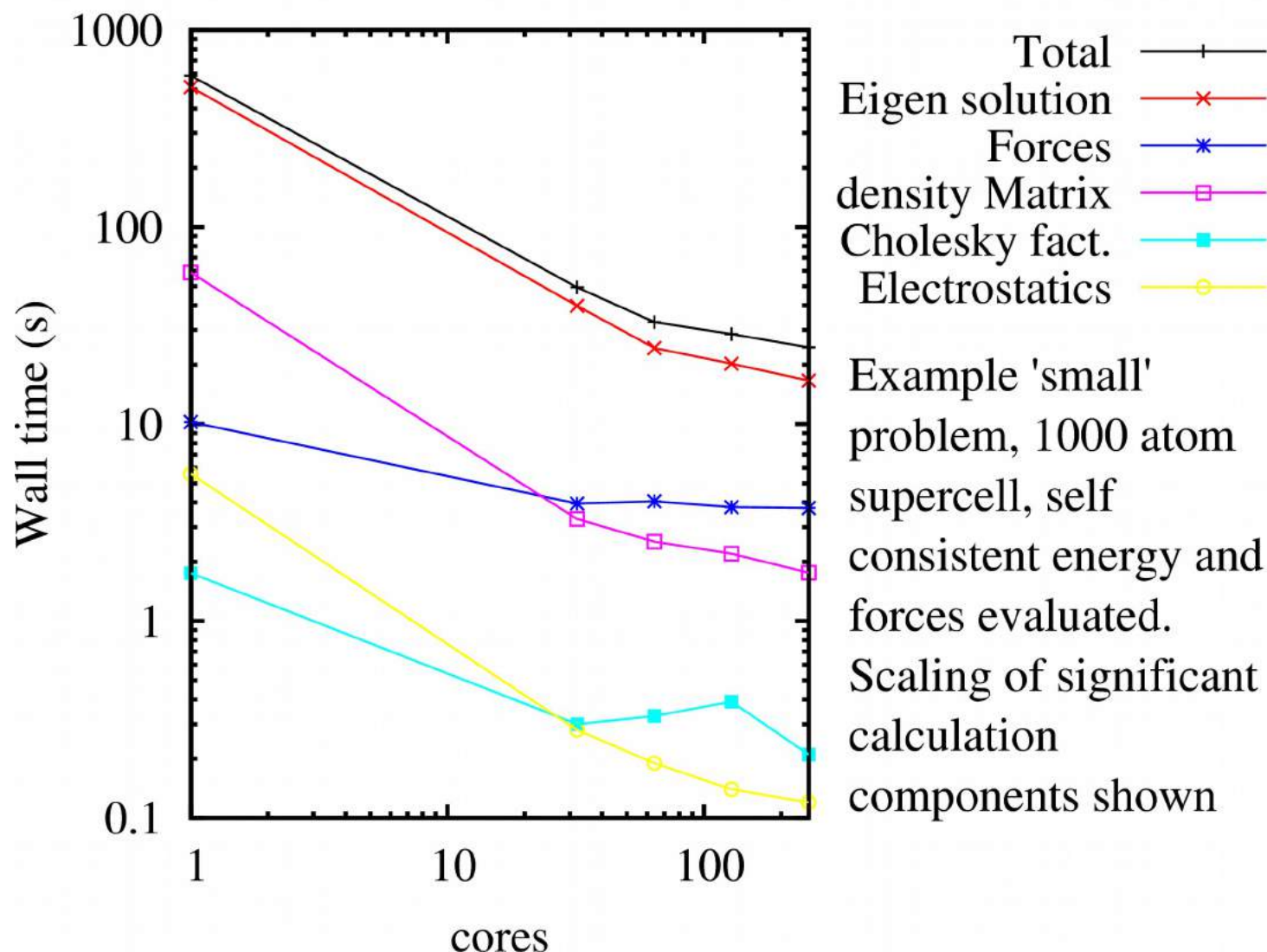
18.2

- Parameterization tools packaged with DFTB+



MPI Parallel scaling

1000 atom SiC supercell (11 SCC cycles + forces)

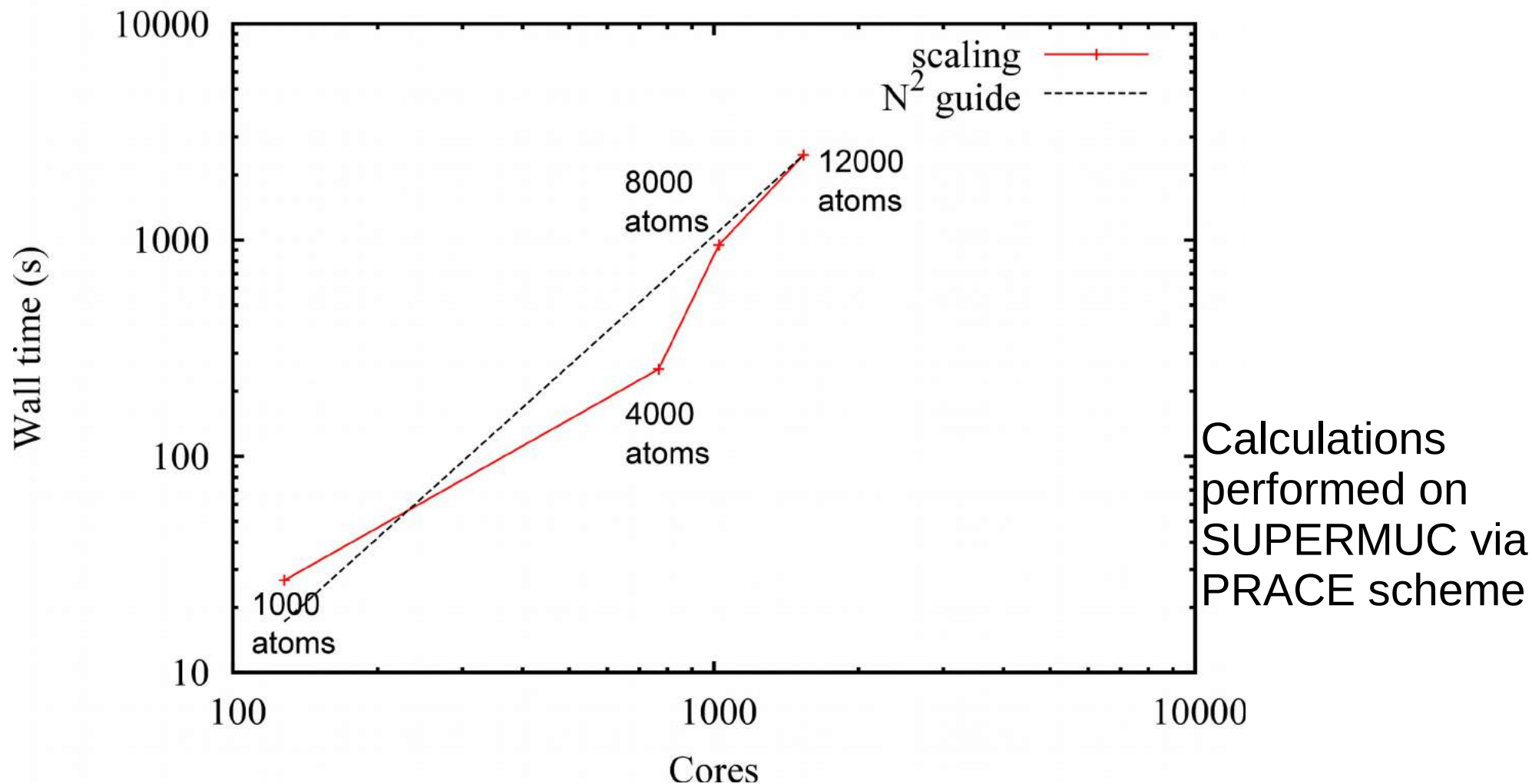


11 SCC cycles,
gamma point
sampling, increase
problem size with
core count

Example 'small'
problem, 1000 atom
supercell, self
consistent energy and
forces evaluated.
Scaling of significant
calculation
components shown

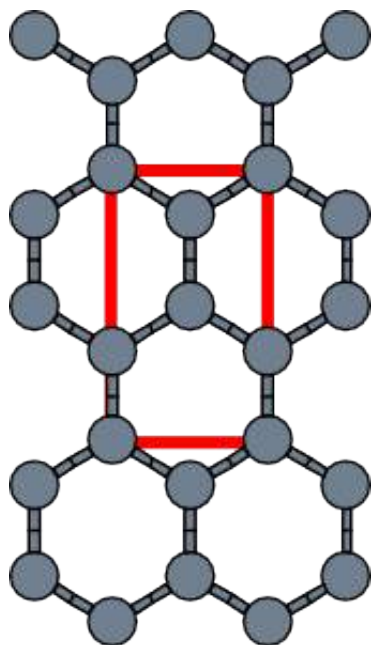
Calculations
performed on
SUPERMUC via
PRACE scheme

MPI Parallel scaling

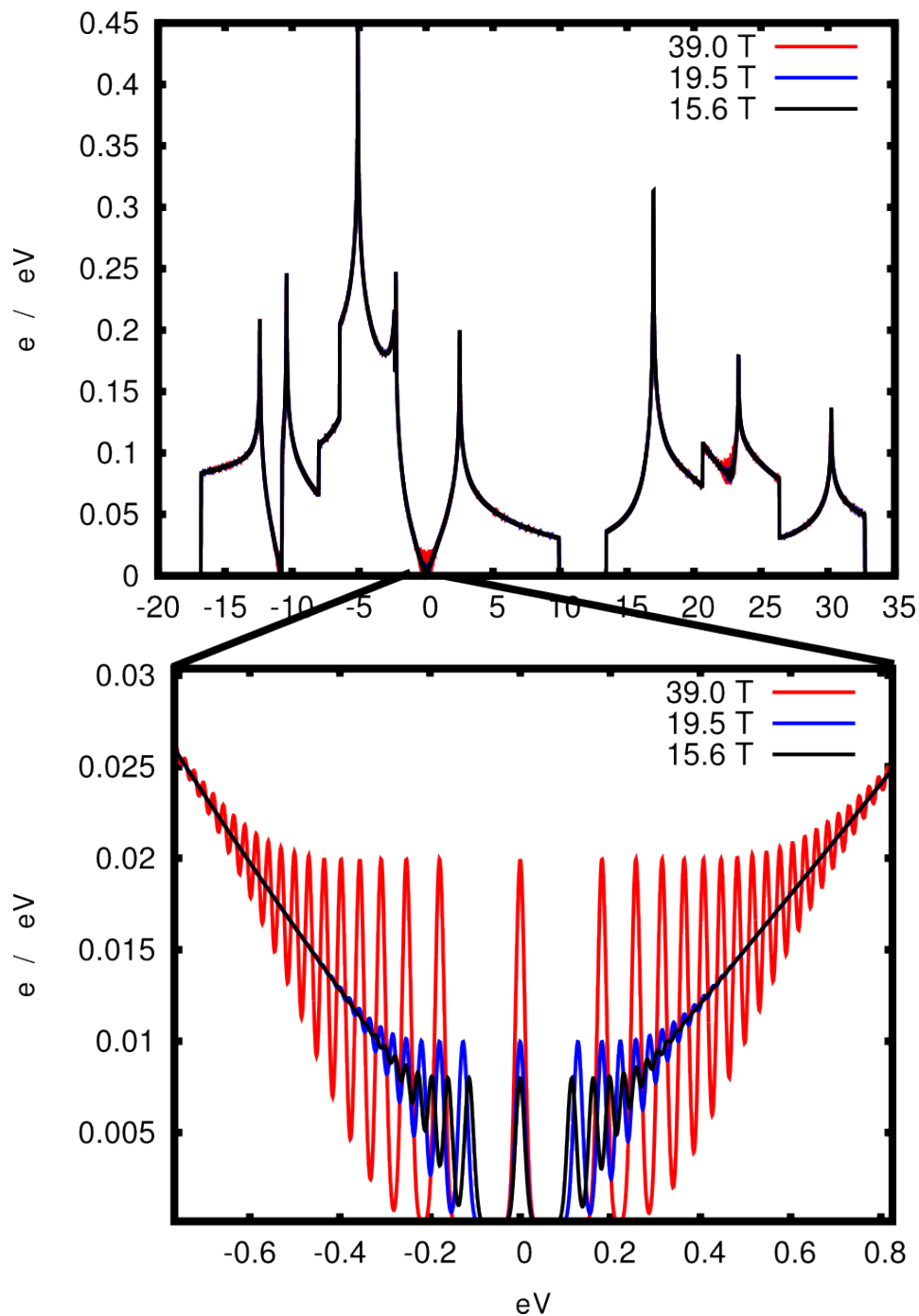


11 SCC cycles, gamma point sampling, increase problem size with core count (called “weak scaling”)

Graphene



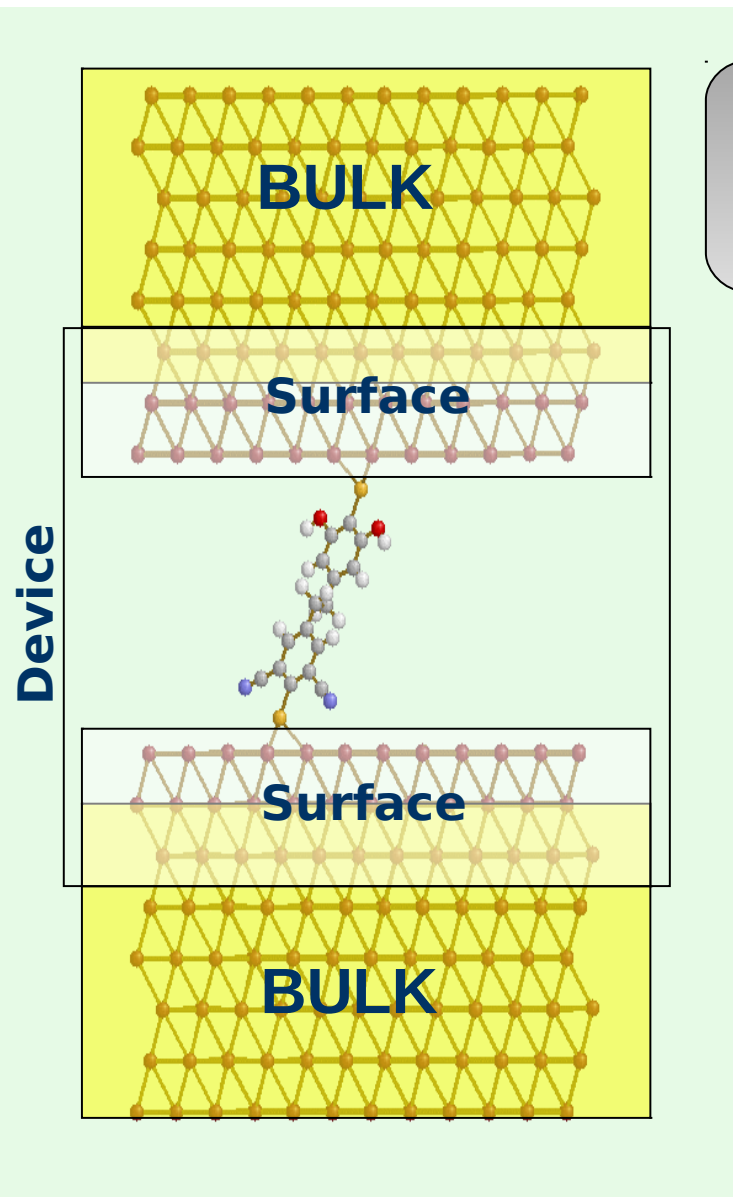
Landau levels in
magnetic field



Field	length	Atoms
39 T	0.8 μm	8000
19.5 T	1.7 μm	16000
15.6 T	2.1 μm	20000

Calculations
performed on
SUPERMUC via
PRACE scheme

Self-consistent quantum transport



$$\rho_{\mu\nu} = \frac{1}{2\pi i} \int dE G_{\mu\nu}^<$$

Density Matrix

$$\Delta q_{\mu} = q_{\mu}^0 - \sum_{\nu} \rho_{\mu\nu} S_{\mu\nu}$$

Mulliken charge Correction

$$\delta n \rightarrow \nabla^2 \delta V_H = -4\pi \delta n$$

$$\rightarrow \delta H \rightarrow \delta G^< \rightarrow \delta n'$$

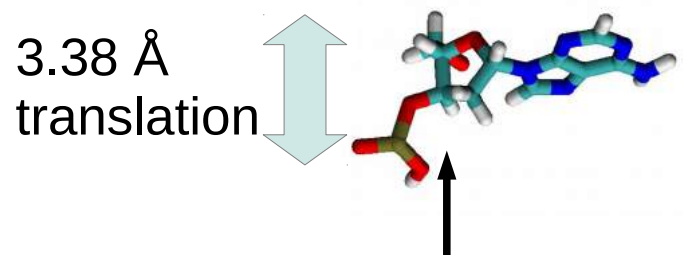
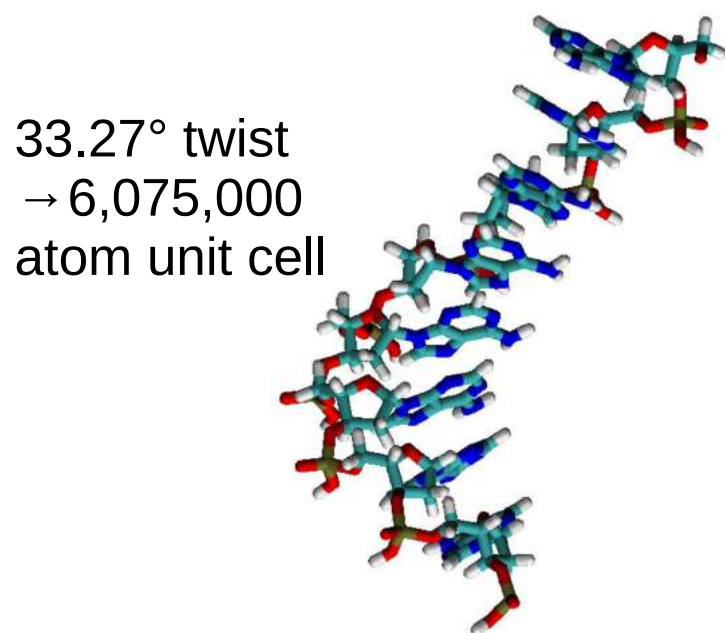
SCC-loop

$$I = \frac{2e}{h} \int Tr[\Sigma^<(E)G^>(E) - \Sigma^>(E)G^<(E)]dE$$

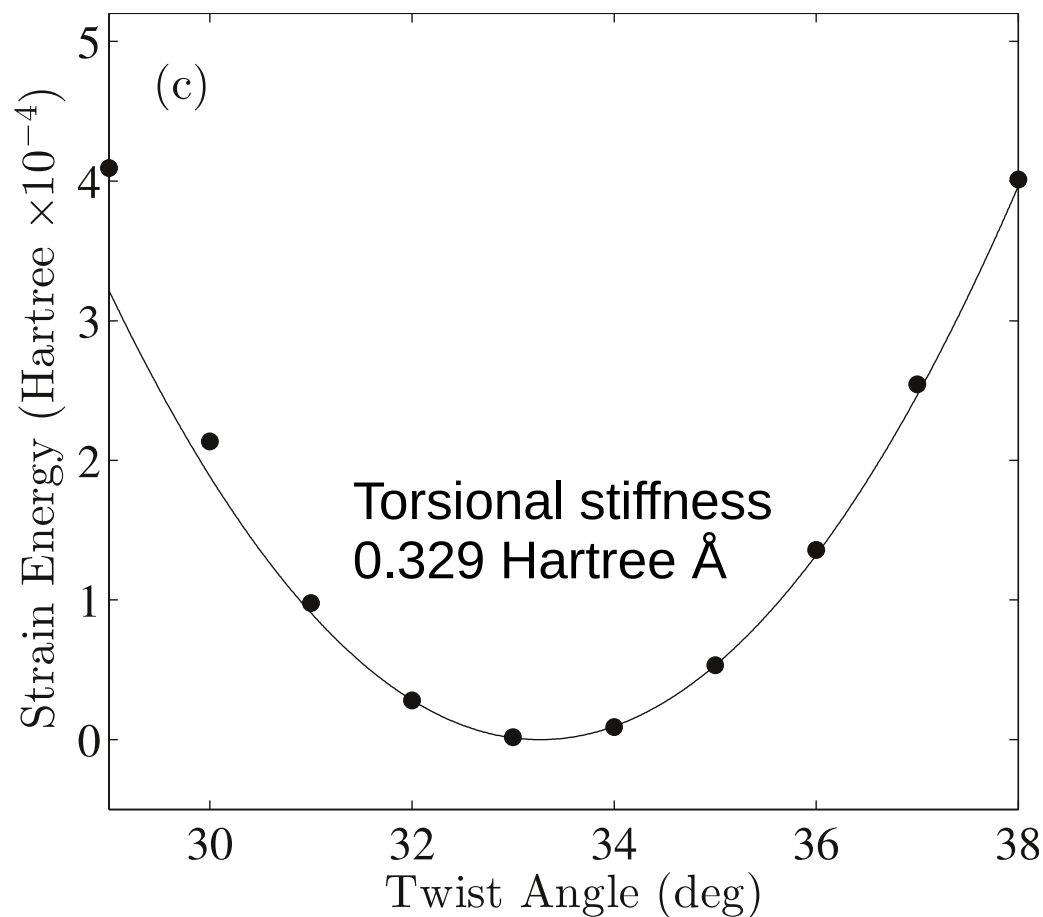
Di Carlo et. al. Physica B, 314, 86 (2002)

Single helix DNA – pure Adenosine in vacuum

Ewald sum of r^6 dispersion interactions included



Minimal (objective) unit
cell – 108 atoms



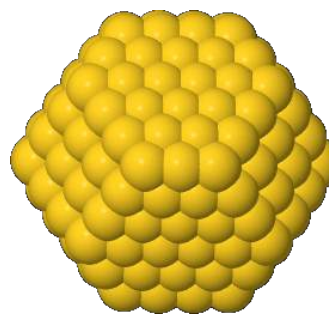
Nikiforov, Hourahine, Aradi, Frauenheim and
Dumitrică, J. Chem. Phys. 139 094110 (2013)

<http://www.dftbplus.org/>

ScotCHEM17

Coupled perturbed TD-response

$$|\dot{\psi}_i^{\pm}\rangle = - \sum_j^{empty} \frac{\langle \psi_j | \partial H / \partial E_k | \psi_i \rangle}{\varepsilon_i - \varepsilon_j + \eta \pm \omega} |\psi_j\rangle$$

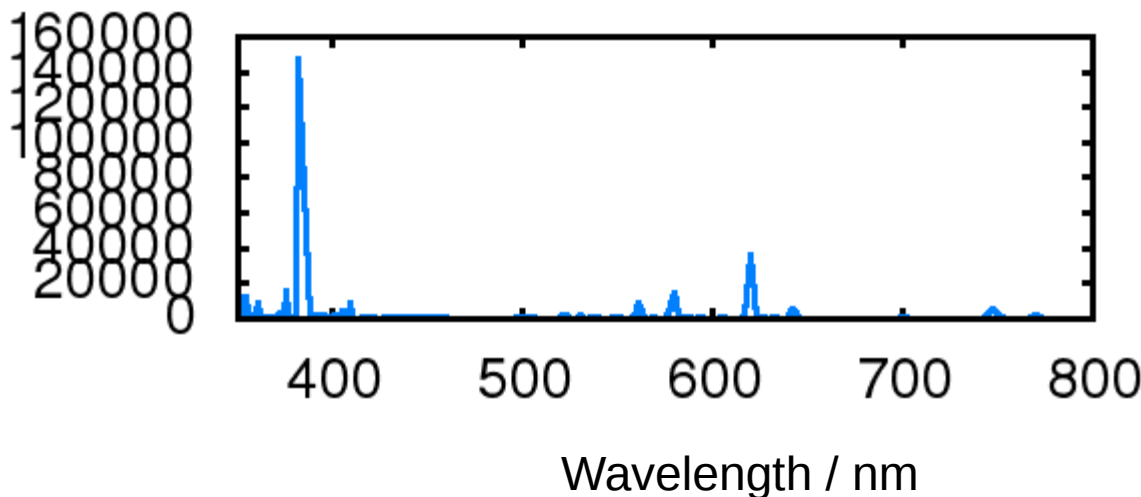


Au_{405}

Wulff
construction
and relax

Spectra in ~2 days
(48 quite old cores)
static structure, but
with L·S spin orbit
coupling

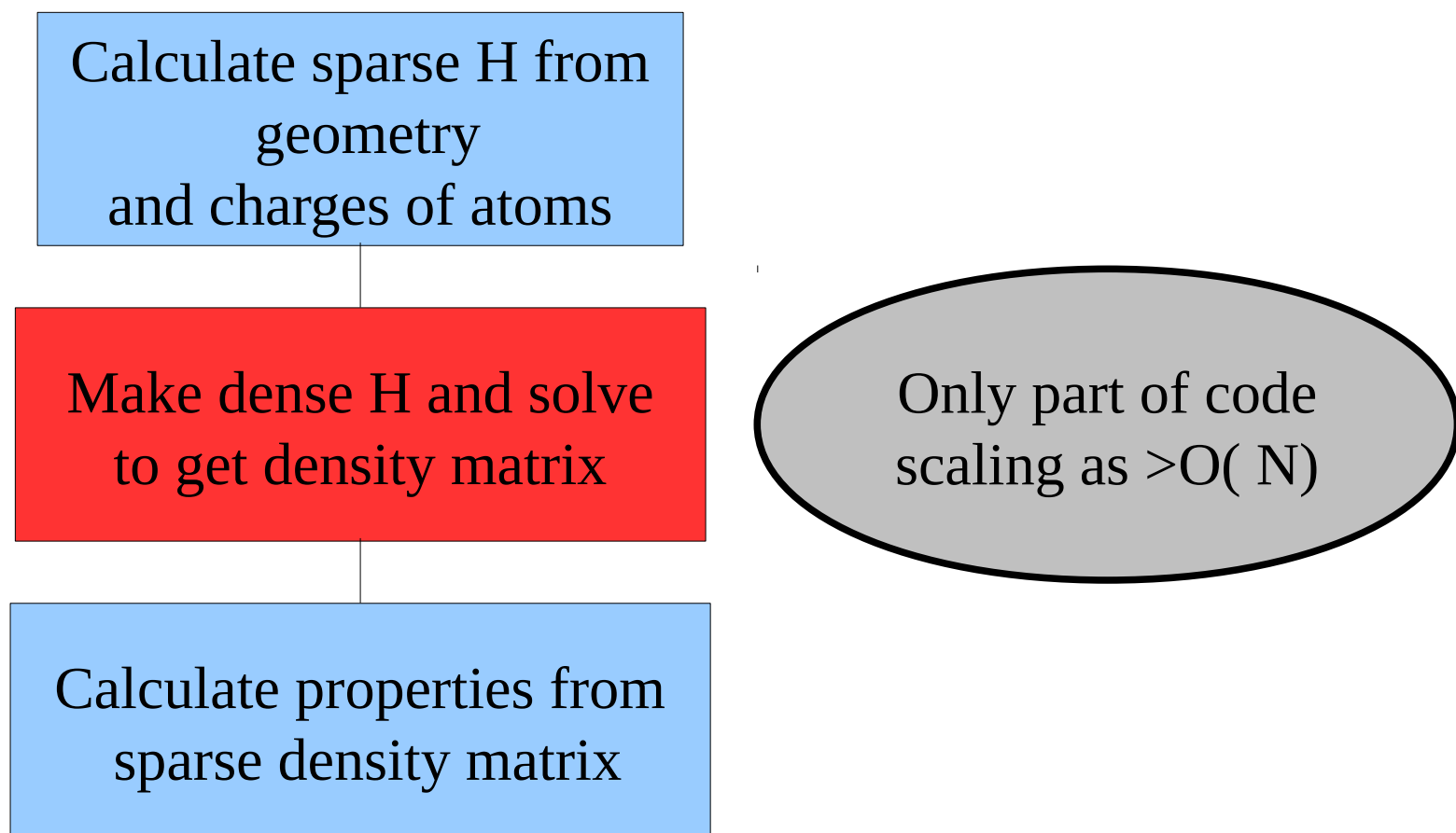
Charge
response



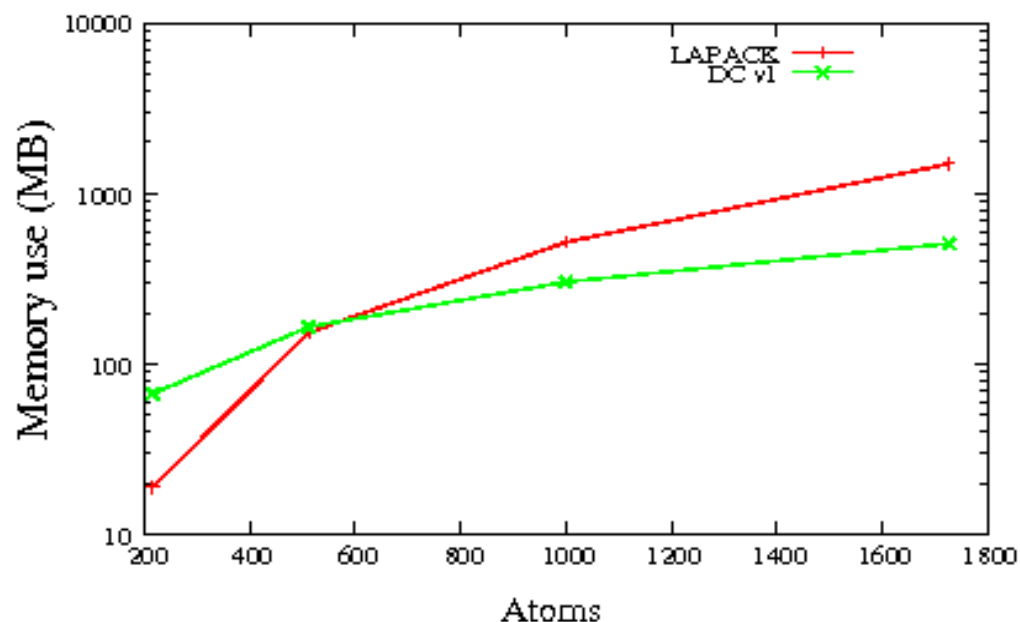
$O(n^3)$ scaling for sum over states, but can be broken into a few GEMM operations for calculation over all $\{i\}$ states simultaneously.

Linear scaling

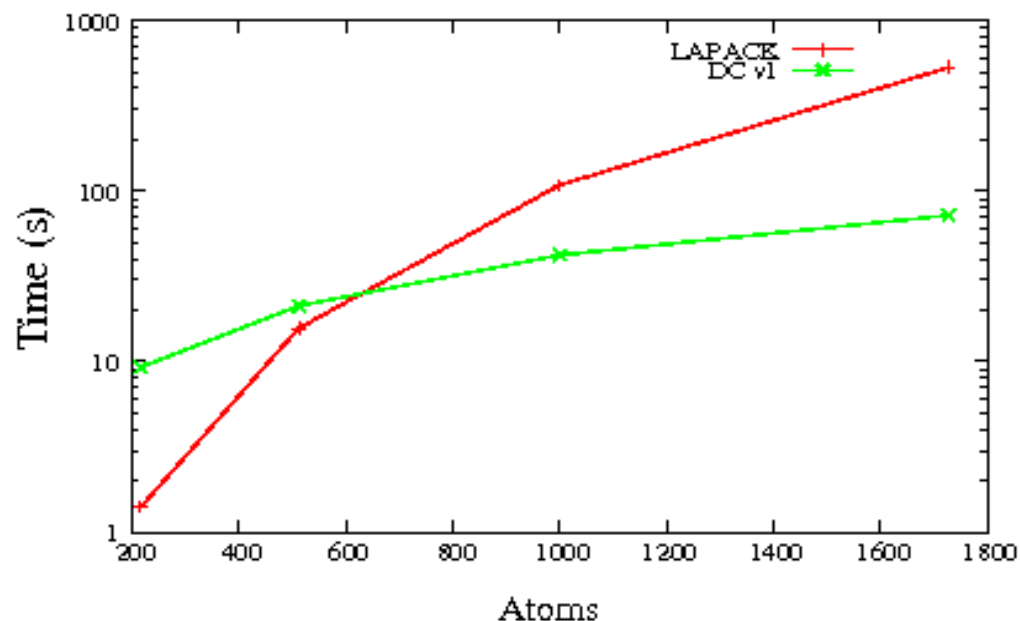
Internally DFTB+ is designed to use a flexible sparse representation to solve the quantum mechanics of the system. Only at one point in its calculation is there a dense (hence expensive) calculation to solve the electronic Hamiltonian.



Simple divide and conquer linear scaling



First results from (yet another) linear scaling DFTB implementation – Yang's divide and conquer
Energies and forces for Si supercells at Γ point.



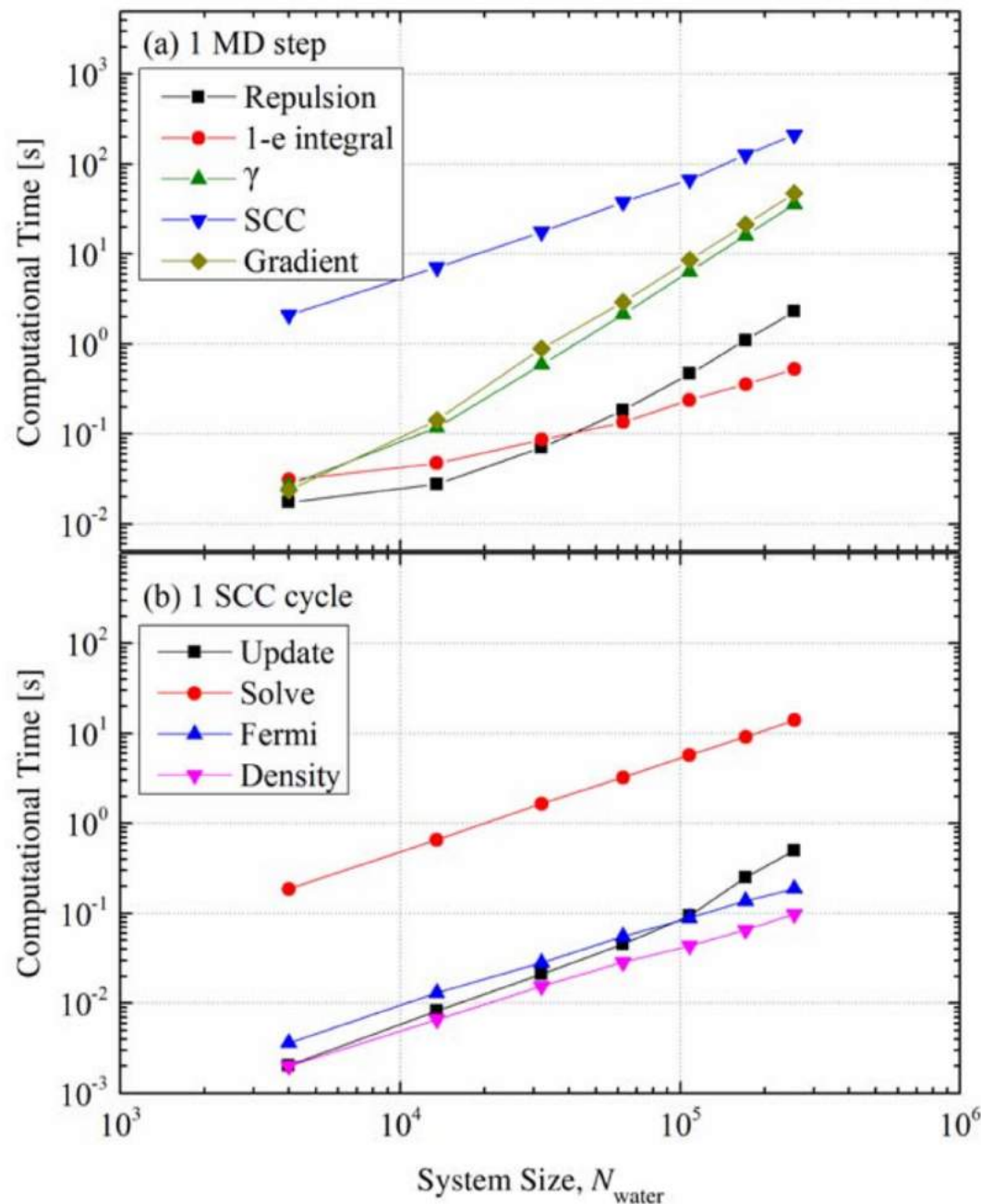
Alternatives also under development

- use of PEXI library
- Galli functional

Current size limits: DC-DFTB-K code

Liquid water simulation
using
Divide and conquer to
build the density matrix,
and a (very) large
supercomputer

Nishizawa et al. Journal of
Computational Chemistry
2016, 37, 1983–1992



And in conclusion



<http://www.dftbplus.org>

Looking forward to supporting more users and new developers

DFT connection

Density Functional Theory, Kohn-Sham theory:

$$E = E[n(\mathbf{r})] \quad n(\mathbf{r}) = \sum_i f_i |\psi_i(\mathbf{r})|^2$$

Foulkes-Haydock: $n(\mathbf{r}) = n_0(\mathbf{r}) + \delta n(\mathbf{r})$

Total energy up to second order in the charge fluctuation:

$$E_{\text{tot}} = E_{\text{bs}}[n_0] + E_{\text{rep}}[n_0] + E_{\text{2nd}}[n_0, \delta n^2]$$

$$\left. \begin{aligned} E_{\text{bs}}[n_0] &= \sum_i f_i \left\langle \psi_i \left| \frac{-\nabla^2}{2} + v_{\text{eff}}[n_0] \right| \psi_i \right\rangle; v_{\text{eff}}[n_0] = v_{\text{ext}} + \frac{\int n_0(r')}{|r-r'|} dr' + v_{\text{xc}}[n_0] \\ E_{\text{rep}}[n_0] &= \frac{-1}{2} \int \frac{n_0(r)n_0(r')}{|r-r'|} dr' dr + E_{\text{xc}}[n_0] - \int v_{\text{xc}}[n_0] n_0 dr + \frac{\sum Z_\alpha Z_\beta}{|R_\alpha - R_\beta|} \\ E_{\text{2nd}}[n_0, \delta n] &= \frac{1}{2} \int \frac{1}{|r-r'|} + \left[\frac{\delta^2 E_{\text{xc}}}{\delta n \delta n} \right]_{n_0} \delta n \delta n' \end{aligned} \right\} \begin{array}{l} \text{Depends} \\ \text{only on} \\ n_0(\mathbf{r}) \end{array}$$

Depends on charge fluctuation
mixes both n^{in} and n^{out}
- stationary energy expression



SCC-DFTB

Charge fluctuations in monopole approximation

(as Mulliken charges):

Second order term:

$$E_{2nd} = \frac{1}{2} \sum_{\alpha} \sum_{\beta \neq \alpha} \gamma_{\alpha\beta} \Delta q_{\alpha} \Delta q_{\beta} \quad \Delta q_{\alpha} = q_{\alpha} - q_{\alpha}^0 \quad q_{\alpha} = \sum_i f_i \sum_{\mu \in \alpha} \sum_{\nu} c_{i\mu} c_{i\nu} S_{\mu\nu}$$

XC contribution described with
chemical hardness of the atoms

$$\gamma_{\alpha\beta} = \frac{1}{R_{\alpha\beta}} - s(R_{\alpha\beta}, U_{\alpha}, U_{\beta}) \text{ From } ab \text{ initio atomic calc.}$$

Mulliken charges depend
on the one-electron
wavefunctions

$$H_{\mu\nu}^{2nd} = \frac{1}{2} S_{\mu\nu} \sum_{\xi} (\gamma_{\alpha\xi} + \gamma_{\beta\xi}) \Delta q_{\xi}$$

$$c_{i\nu} (H_{\mu\nu} - \epsilon_i S_{\mu\nu}) = 0$$

must be solved **self consistently**

$$\rightarrow E_{tot} = \sum_i f_i \sum_{\mu\nu} c_{i\mu} c_{i\nu} H_{\mu\nu}^{bs} + \frac{1}{2} \sum_{\alpha\beta} \gamma_{\alpha\beta} \Delta q_{\alpha} \Delta q_{\beta} + E_{rep} \quad \left. \vphantom{\sum_i} \right\} \text{ Depends only on } n^{\text{out}}$$



Spin

Spin polarization is an additional term to the SCC-DFTB energy expression

- SCC contribution to the total energy is resolved by angular atomic momentum
- Mulliken spin populations used – shell resolved magnetisation
- Atomic spin constants $W_{ll'}$ similar to the atomic Hubbard U_l (Stoner-I like)

$$E_{tot} = \sum_i f_i \sum_{\mu\nu} c_{i\mu} c_{i\nu} H_{\mu\nu}^0 + \frac{1}{2} \sum_{ab} \gamma_{kk'} \Delta q_k \Delta q_{k'} + \frac{1}{2} \sum_a W_{ll'} \Delta p_l \Delta p_{l'} + E_{rep}$$

$k, k' \in (a, b); l, l' \in a$

$$H_{\mu\nu}^{\uparrow} = \frac{1}{2} S_{\mu\nu} \sum_{\xi} (\gamma_{\alpha\xi} + \gamma_{\beta\xi}) \Delta q_{\xi} \quad H_{\mu\nu}^{\downarrow} = -H_{\mu\nu}^{\uparrow}$$

(generalising to non-collinear spin – swap spinors for wavefunctions and dot product for multiplication)

Köhler et al., Chem. Phys., vol. 309, pp. 23 (2005)

Köhler et al. J. Phys. Chem. A 111 5622 (2007)

