The approximate SCC-DFTB method: performance and challenges

Marcus Elstner
Institute of Physical Chemistry
Karlsruhe Institute of Technology
Approximate density-functional theory: SCC-DFTB
Self consistent-charge density functional tight-binding

O-LCAO; 2-center approximation: approximate DFT
http://theory.chm.tu-dresden.de

efficient parametrization scheme: DFTB
www.bccms.uni-bremen.de

charge self-consistency: SCC-DFTB
www.tu-bs.de/pci

www.dftb.org

approximate DFT
Some relevant publications

- PROTEINS 44 (2001) 484 (O(N))
Extensions

**TD-DFTB, GW**
T. Niehaus
Bremen

**QM/MM ‘multicale’**
Q.Cui,
Madison

**Dispersion Correction**
P. Hobza
Prague

**Molecular Electronics**
A. di Carlo, A. Peccia,
Rome

**SCC-DFTB**

**O(N) DFTB, continuum model**
W. Yang, H. Liu
Duke

**NMR, CPMD**
T. Heine,
Bremen

**LDA+U**
B. Houraine
Glasgow

**transition metals, anal. 2nd deriv**
K. Morokuma, H. Witek
Atlanta

B. Aradi (Bremen)  DFTB+ code =>  www.DFTB.org
Semi-empirical /approximate methods

approximation, neglect and parametrization of interaction integrals from ab-initio and DFT methods

-HF-based:
   CNDO, INDO, MNDO, AM1, PM3, MNDO/d, OM1, OM2

-DFT-based:
   **SCC-DFTB**
   DFT- 3-center- tight binding (Sankey)
   (Fireballs --- > Siesta DFT code)

about 3 orders of magnitude faster than DFT-GGA (medium sized basis set)
DFTB is derived from DFT

inherits the problems of DFT:

• VdW interactions => empirical dispersion
• TD-DFT failures => limited use of TD-DFTB
• overpolarizability
• overbinding
• single reference method
• ...
DFTB is derived from DFT

inherits the problems of DFT:

- VdW interactions => empirical dispersion
- TD-DFT failures => limited use of TD-DFTB
- overpolarizability
- overbinding
- single reference method
- ...

but also the strenghts of DFT

- conceptual simplicity
- good geometries
- reasonable vib. frequencies
- ...

=> MNDO-type and DFTB methods complement each other
Tight Binding (TB) and Density Functional TB

TB methods as approximate DFT

- minimal basis
- approximation of Matrix elements
- DFT double counting: repulsive potential
DFT total energy

Starting point is the DFT total energy \((\int' = \int dr'):\)

\[
E[\rho] = T + \int v_{\text{ext}}(r)\rho + \frac{1}{2} \int \frac{\rho \rho'}{|r - r'|} + E_{\text{xc}}[\rho] + \frac{1}{2} \sum_{\alpha\beta} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}}
\]

\(\rho = \rho(r)\) is the ground state density.

Choose a reference (initial) density \(\rho_0\), which is built from neutral atomic densities:

\[
\rho_0 = \sum_{\alpha} \rho_\alpha
\]
Harris functional \quad (PRB 31, 1770)

\( \rho_1, \rho_2 \): densities of weakly interacting fragments

\[
\rho_0 = \rho_1 + \rho_2 \\
\Delta \rho = \rho - \rho_0
\]

Solve KS-equations NON-selfconsistently using \( \rho_0 \rightarrow \epsilon_i^H \)

Total energy:

\[
E[\rho] = E[\rho_0] + O(\Delta \rho^2)
\]

\[
E[\rho] \approx \sum_{i}^{occ} \epsilon_i^H - \frac{1}{2} \int \frac{\rho_0 \rho'_0}{|r-r'|} + E_{xc}[\rho_0] - \int v_{xc}(r)\rho_0 dr + \frac{1}{2} \sum_{\alpha \beta} \frac{Z_{\alpha} Z_{\beta}}{R_{\alpha \beta}}
\]
Tight binding theory

Minimal LCAO basis: \( \phi_i = \sum_\mu c_{i\mu} \eta_\mu \)

DFT Hamiltonian: \( \hat{H}[\rho_0] = \hat{T} + v_{\text{eff}}[\rho_0] \)

KS-equation:
\[
\sum_\mu c_{i\mu} \hat{H}[\rho_0] |\eta_\mu> = \epsilon_i \sum_\mu c_{i\mu} |\eta_\mu>
\]

Multiplication with \( < \eta_\nu | \)
\[
\sum_\mu c_{i\mu} < \eta_\nu | \hat{H}[\rho_0] |\eta_\mu> = \epsilon_i \sum_\mu c_{i\mu} < \eta_\nu | \eta_\mu>
\]

or in matrix notation

\[
HC = SC\epsilon
\]
Tight binding theory

→ single diagonalization of Hamilton matrix in AO basis $H_{\mu\nu}$ leads to KS eigenvalues $\epsilon_i^H$. Total energy:

$$E[\rho] \approx \sum_{i}^{\text{occ}} \epsilon_i^H - \frac{1}{2} \int \frac{\rho_0 \rho'_0}{|r - r'|} + \frac{1}{2} \sum_{\alpha\beta} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} + E_{xc}[\rho_0] - \int v_{xc}(r) \rho_0 \, dr$$

Tight-Binding (TB) theory:

$$E[\rho] \approx \sum_{i}^{\text{occ}} \epsilon_i^H + E_{\text{rep}}$$

- fit $H_{\mu\nu}$ e.g. in order to reproduce band-structure of solids
- approximate $E_{\text{rep}}$ by exponentials
- correct to 1st order in $\Delta \rho$ (Foulkes & Haydock, PRB 39, 12520)
Repulsive Potential

Expand reference density: \( \rho_0 = \sum_\alpha \rho_\alpha \)

DFT double counting term:

\[
E^{DC}[\rho_0] = -\frac{1}{2} \sum_{\alpha \beta} \int \frac{\rho_\alpha \rho_\beta'}{|r-r'|} + \frac{1}{2} \sum_{\alpha \beta} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} + E_{xc}[\rho_0] - \int v_{xc}(r) \rho_0 dr
\]

- Coulomb terms cancel each other: decay exponentially with \( R_{\alpha\beta} \)
- XC- terms decay exponentially as well in LDA, GGA

\[
\rightarrow E^{DC}[\rho_0] \approx \sum U_{\alpha\beta}(R_{\alpha\beta})
\]

Fit \( E^{DC} \) by fast decaying pair potentials.
Density Functional Tight Binding: DFTB

Compute Hamilton matrix in AO basis $H_{\mu\nu}$ from DFT-GGA (PBE)

Introduce Optimized LCAO basis (Eschrig 1988)

$$
\left[ -\frac{1}{2} \nabla^2 + v_{\text{eff}} [\rho_{\text{atom}}] + \left( \frac{r}{r_0} \right)^2 \right] \eta_\mu = \epsilon_\mu \eta_\mu
$$

- Confined basis functions $\eta_\mu$ (usually minimal basis)
  
  One parameter per atomtype $X : r_0^X \approx 2 r_{\text{cov}}^X$

- 'Atomic' densities $\rho_\alpha = \sum_\mu |\eta_\mu|^2$

  Different confinement for densities: $r_D^X = 7 - 14\text{a.u.}$
DFTB: electronic part

Compute overlap $S_{\mu\nu}$ and Hamilton $H_{\mu\nu}$ matrix

1. $S_{\mu\nu} = \langle \eta_\mu | \eta_\nu \rangle$

2. $H_{\mu\mu} = \epsilon_\mu$ of 'free' (DFT-PBE) atom

3. 2-center approximation for off-diagonal:

$$H_{\mu\nu} = \langle \eta_\nu | \hat{H}[\rho_0] | \eta_\mu \rangle \approx \langle \eta_\nu | \hat{H}[\rho_\alpha + \rho_\beta] | \eta_\mu \rangle$$

$\mu \in \alpha$, $\nu \in \beta$

Alternative: $H_{\mu\nu} \approx \langle \eta_\nu | \hat{H}[\rho_\alpha] + \hat{H}[\rho_\beta] | \eta_\mu \rangle$
Slater-Koster rules

Reduction to primitive form

\[ E_{Sx} = \langle s | H | x \rangle = \cos \theta V_{\text{sp}} - \sin \theta V_{\text{spin}} = 0 \]

Rotate wave function to a more appropriate coordinate system

\[ \psi_x = \cos \theta \psi_{x'} - \sin \theta \psi_{y'} \]

Walter A. Harrison, *Electronic Structure and the Properties of Solids*
DFTB: repulsive part

\[ E[\rho] = \sum_{i}^{\text{occ}} \epsilon_i^H + \sum_{\alpha\beta} U_{\alpha\beta} \rightarrow U_{\alpha\beta} = E^{DFT}[\rho](R_{\alpha\beta}) - \sum_{i}^{\text{occ}} \epsilon_i^H(R_{\alpha\beta}) \]
What are the ‘free’ parameters?

• confinement for wave-function: \( r_0^X \)
• confinement for density: \( r_D^X \)
• repulsive potential \( U_{\alpha\beta} \):
  - 6th order polynomial (spline)
  - 20-40 DFT data points
DFTB: parameters

What are the ‘free’ parameters?

• confinement for wave-function: \( r_0 \)
• confinement for density: \( r_D \)
• repulsive potential \( U_{\alpha\beta} \):
  - 6th order polynomial (spline)
  - 20-40 DFT data points

Foulkes and Haydock (PRB 39, 12520):
choice of initial density
Non-selfconsistent TB

Solve KS equations non-selfconsistently (1 iteration)

\[
\begin{bmatrix}
-\frac{1}{2} \nabla^2 + v_{\text{eff}}[\rho_{\text{in}}]
\end{bmatrix} \psi_i = \epsilon_i \psi_i 
\rightarrow 
\rho_{\text{out}} = \sum_i |\psi_i|^2
\]

Energy is correct to first order in \( \Delta \rho = \rho - \rho_{\text{in}} \)
Non-selfconsistent TB

Solve KS equations non-selfconsistently (1 iteration)

\[
\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}[\rho_{\text{in}}]\right] \Psi_i = \epsilon_i \Psi_i \quad \rightarrow \quad \rho_{\text{out}} = \sum_i |\Psi_i|^2
\]

Energy is correct to first order in \(\Delta \rho = \rho - \rho_{\text{in}}\)

- No charge transfer between atoms \(\Rightarrow\) very good results for homonuclear systems (Si, C), hydrocarbons etc.
- Complete transfer of one charge between atoms \(\Rightarrow\) Also does not fail for ionic systems (e.g. NaCl):
  - Harrison
  - Slater, (Theory of atoms and molecules)
- Problematic case: everything in between
Non-selfconsistent DFTB

• HCOOH: C=O and C-O bond lengths equalized
• H$_2$N-CH=O and peptides: N-C and C=O bond lengths equalized

non-CT systems:
• CO2 vibrational frequencies
• C=C=C=C=C.. chains, dimerization, end effects
Non-selfconsistent DFTB

- HCOOH: C=O and C-O bond lengths equalized
- $\text{H}_2\text{N-CH}=\text{O}$ and peptides: N-C and C=O bond lengths equalized

Non-CT systems:
- CO$_2$ vibrational frequencies
- C=C=C=C=C.. chains, dimerization, end effects

Problem: charge transfer between atoms overestimated due to electronegativity differences between atom

=> need balancing force:
onsite e-e interaction of excess charge is missing!
How to account for CT between atoms?

\[
\left[-\frac{1}{2} \nabla^2 + v_{\text{eff}}[\rho_0]\right] \psi_i = \epsilon_i \psi_i
\]

how to approximate gs charge density?

\[
\left[-\frac{1}{2} \nabla^2 + v_{\text{eff}}[\rho]\right] \psi_i = \epsilon_i \psi_i
\]
How to account for CT between atoms?

\[
\left[ -\frac{1}{2} \nabla^2 + v_{\text{eff}}[\rho_0] \right] \psi_i = \epsilon_i \psi_i
\]

how to approximate gs charge density?

\[
\left[ -\frac{1}{2} \nabla^2 + v_{\text{eff}}[\rho] \right] \psi_i = \epsilon_i \psi_i
\]

Expansion of \( v_{\text{eff}} \) around reference density

\[
v_{\text{eff}}[\rho] = v_{\text{eff}}[\rho_0] + \int \frac{\delta v_{\text{eff}}}{\delta \rho} \delta \rho + \ldots
\]

\[H_{\mu\mu} = H_{\mu\mu}^0 + ??\]
Self Consistent Charge Density Functional Tight Binding: SCC-DFTB

Second order expansion of DFT total energy

- approximation for difference density $\Delta \rho = \rho - \rho_0$
- approximation of DFT kernel

Second order expansion of the DFT total energy functional with respect to the charge density fluctuations $\delta \rho$ around a given reference density $\rho_0$ ($\rho_0' = \rho_0(\vec{r}')$, $\int' = \int d\vec{r}'$):

$$
E = \sum_{i}^{\text{occ}} \langle \psi_i | \hat{H}^0 | \psi_i \rangle + \frac{1}{2} \int \int' \left( \frac{1}{|\vec{r} - \vec{r}'|} + \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \bigg|_{n_0} \right) \Delta \rho \Delta \rho'.
$$

$$
- \frac{1}{2} \int \int' \frac{\rho_0' \rho_0}{|\vec{r} - \vec{r}'|} + E_{xc}[\rho_0] - \int V_{xc}[\rho_0] n_0 + E_{cc}
$$

Reference density: $\rho_0 = \sum \rho_0^\alpha$

Superposition of atomic densities
LCAO basis $\Psi_i = \sum c^i_\mu \eta_\mu$:

$$\langle \Psi_i | \hat{H}^0 | \Psi_i \rangle = \sum c^i_\mu c^i_\nu H^0_{\mu\nu}$$

$$H^0_{\mu\nu} = \begin{cases} 
\epsilon_\mu : \mu = \nu \\
\langle \phi_\mu | H_{KS}[\rho_A + \rho_B] | \phi_\nu \rangle : \mu \in A, \nu \in B \\
0 : \text{otherwise}
\end{cases}$$

- 'Special' minimal basis set $\phi_\mu$ and initial densities from atomic KS eqs.
- $H^0_{\mu\nu}$ and $S_{\mu\nu}$ calculated and stored $\rightarrow$ no integral evaluation during program runtime.
Second order contribution

\[ E^2[\rho, \rho_0] = \frac{1}{2} \iint \left( \frac{1}{|r - r'|} + \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \bigg|_{n_0} \right) \Delta \rho \Delta \rho' \]

Represent by atomic contributions:

\[ \Delta \rho = \sum_{\alpha} \Delta \rho_{\alpha} \]

Monopole approximation

\[ \Delta \rho_{\alpha} \approx \Delta q_{\alpha} F_{00}^\alpha Y_{00} \]

\[ E^2[\rho, \rho_0] \approx \frac{1}{2} \sum_{\alpha \beta} \Delta q_{\alpha} \Delta q_{\beta} \iint \left( \frac{1}{|r - r'|} + \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \bigg|_{n_0} \right) F_{00}^\alpha F_{00}^\beta Y_{00}^2 \]

Basic assumption:

- Only transfer of net charge between atoms.
- Size and shape of atom (in molecule) not changed
Second order contribution

\[ E^2[\rho, \rho_0] \approx \frac{1}{2} \sum_{\alpha \beta} \Delta q_{\alpha} \Delta q_{\beta} \int \int' \left( \frac{1}{|\vec{r} - \vec{r}'|} + \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \bigg|_{n_0} \right) F_{00}^\alpha F_{00}^\beta Y_{00}^2 \]

Consider two limiting cases:

A) Large distance: \(|\vec{r} - \vec{r}'| \approx R_{\alpha\beta} \to \infty\)

\[ E^2[\rho, \rho_0] \to \frac{1}{2} \sum_{\alpha \beta} \frac{\Delta q_{\alpha} \Delta q_{\beta}}{R_{\alpha\beta}} \]

B) 'On-site' e-e repulsion: \(R_{\alpha\beta} \to 0\), i.e. \(\alpha = \beta\)

\[ E^2[\rho, \rho_0] \to \frac{1}{2} \sum_{\alpha} \frac{\partial^2 E_{\alpha}^{at}}{\partial \Delta q_{\alpha}^2} \Delta q_{\alpha}^2 = \frac{1}{2} \sum_{\alpha} U_{\alpha} \Delta q_{\alpha}^2 \]

New (Hubbard) parameter \(U_{\alpha}\): computed from DFT (PBE)
Second order contribution

Derive analytical function which interpolates between these two cases

\[
E^2[\rho, \rho_0] \approx \frac{1}{2} \sum_{\alpha, \beta} \Delta q_\alpha \Delta q_\beta \int \int' \left( \frac{1}{|\vec{r} - \vec{r}'|} + \left. \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \right|_{n_0} \right) F^\alpha_{00} F^\beta_{00} Y^2_{00}
\]

\[
E^2[\rho, \rho_0] \approx \frac{1}{2} \sum_{\alpha, \beta} \Delta q_\alpha \Delta q_\beta \gamma_{\alpha\beta}
\]

\(\gamma_{\alpha\beta}\) also used to approximate TD-DFT kernel

Second order contribution

Consider spherical atomic charge densities

\[ F_{00}^{\alpha}(r) = C \exp(- (r - R_{\alpha}) / \sigma) \quad \sigma \approx R_{\text{cov}} \]

Compute Coulomb integrals: analytic expression for \( \gamma_{\alpha\beta} \)

\[ \int \int' \left( \frac{1}{|\vec{r} - \vec{r}'|} + \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \right) F_{00}^{\alpha} F_{00}^{\beta} \rightarrow \gamma_{\alpha\beta} \]

- deviates from \( 1/r \) for small distances due to overlap: \( \gamma_{\alpha\alpha} = U_{\alpha} \)
- For \( R_{\alpha\beta} = 0: \quad 1/\sigma = 3.2 U_{\alpha} \)

I.e. \( U_{\alpha} \) is a measure for the 'atomic size'

Electron-Electron interaction

- Assumes that atomic size is inverse to chemical hardness ($\eta = 0.5 \, U_\alpha$)
- 'Contains' the size of neutral atoms, modeled as inverse of $U_\alpha$
- Different versions for $\gamma$: Mataga-Nishimoto < Klopman-Ohno < DFTB
- Different needs for different properties? Excited states vs. H-bonding: is there a compatibility problem?

\[ \gamma_{\alpha\beta} = \frac{1}{\sqrt{R^2 + \frac{1}{4} (U^{-1}_\alpha + U^{-1}_\beta)^2}} \]

\[ \text{atom size } \sigma: \quad \frac{1}{\sigma} = 3.2 \, U \]
Repulsive energy terms

\[-\frac{1}{2} \iiint \frac{\rho'_0 \rho_0}{|\vec{r} - \vec{r}'|} + E_{xc}[\rho_0] - \int V_{xc}[\rho_0] n_0 + E_{cc} \rightarrow \sum_{\alpha\beta} U_{\alpha\beta}\]

SCC-DFTB total energy:

\[E = \sum_i \sum_{\mu\nu} c_i^\mu c_i^\nu H_{\mu\nu}^0 + \frac{1}{2} \sum_{\alpha\beta} \gamma_{\alpha\beta} \Delta q_{\alpha} \Delta q_{\beta} + \sum_{\alpha\beta} U_{\alpha\beta}\]
Determination of the repulsive potential

\[ E_{\text{rep}}(R_{\alpha\beta}) = E^{\text{DFT}}(R_{\alpha\beta}) - \left[ \sum \sum c^i_{\mu} c^j_{\nu} H^0_{\mu\nu} + E^2 \right] (R_{\alpha\beta}) \]

strategy:
- calculate \( E_{\text{DFT}} \) and \( E_{\text{el}} \) for various C-C distances
- fit the difference to polynomial/splines
SCC-DFTB total energy

\[ E = \sum_{i}^{\text{occ}} \langle \psi_i | \hat{H}^0 | \psi_i \rangle + \frac{1}{2} \int \int' \left( \frac{1}{|\vec{r} - \vec{r}'|} + \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \right) |_{n_0} \Delta \rho \Delta \rho'. \]

\[ - \frac{1}{2} \int \int' \frac{\rho'_0 \rho_0}{|\vec{r} - \vec{r}'|} + E_{xc}[\rho_0] - \int V_{xc}[\rho_0] n_0 + E_{cc} \]

- minimal basis
- neglect of crystal field and three-center terms
- initial density fixed
- second order expansion
- monopole approximation
- gamma
- two-body approximation
- fit procedure

\[ E = \sum_{i} \sum_{\mu \nu} c^i_{\mu} c^i_{\nu} H^0_{\mu \nu} + \frac{1}{2} \sum_{\alpha \beta} \gamma_{\alpha \beta} \Delta q_{\alpha} \Delta q_{\beta} + \sum_{\alpha \beta} U_{\alpha \beta} \]
Can DFTB be better than DFT?

a) $H_{\mu\nu}$ from PBE => reproduce problems of DFT
   - band gap
   - overpolarizability
   - CT excited states
   - correlation: VdW forces, ...

   Plus additional problems from minimal basis and approximations

=> empirical TB can be better than DFT (e.g. band structure, ...)
Can DFTB be better than DFT?

a) $H_{\mu\nu}$ from PBE $\Rightarrow$ reproduce problems of DFT
   - band gap
   - overpolarizability
   - CT excited states
   - correlation: VdW forces, ...

   Plus additional problems from minimal basis and approximations

$\Rightarrow$ empirical TB can be better than DFT (e.g. band structure, ...)

b) $E = E_{el} + E_{rep}$ $\Rightarrow$ parametrize $E_{rep}$

$\Rightarrow$ properties dependent on $E$ can be better than DFT
   - heats of formation
   - vib. frequencies
   - reaction barriers

However: limited flexibility of DFTB formalism

$\Rightarrow$ conflicts in parameter optimization
Tests for organic and biological molecules

• reaction energies, atomization energies
• geometries
• vib. frequencies

• H-bonds
• VdW

• proton affinities

• peptide structures, DNA bases, sugars ...
Performance for small organic molecules
(mean absolute deviations)

• Reaction energies\(^a\): \(\sim 5\) kcal/mole
• Bond lengths\(^a\): \(\sim 0.014\) Å
• Bond angles\(^b\): \(\sim 2^\circ\)
• Vib. Frequencies\(^c\): \(\sim 6-7\) %


<table>
<thead>
<tr>
<th>Reaction</th>
<th>DFTB</th>
<th>cc-pVDZ</th>
<th>cc-pVTZ</th>
<th>G2</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂ + CH≡CH → CH₂=CH₂</td>
<td>−38.8</td>
<td>−41.0</td>
<td>−37.9</td>
<td>−40.1</td>
</tr>
<tr>
<td>H₂ + CH₂=CH₂ → CH₃–CH₃</td>
<td>−37.0</td>
<td>−28.5</td>
<td>−26.5</td>
<td>−30.5</td>
</tr>
<tr>
<td>3H₂ + HCN → NH₃ + CH₄</td>
<td>−47.9</td>
<td>−47.1</td>
<td>−48.5</td>
<td>−53.7</td>
</tr>
<tr>
<td>H₂ + CO → H₂CO</td>
<td>+10.9</td>
<td>+0.1</td>
<td>+0.7</td>
<td>+3.1</td>
</tr>
<tr>
<td>2H₂ + CO → CH₃OH</td>
<td>−7.8</td>
<td>−12.6</td>
<td>−13.1</td>
<td>−15.5</td>
</tr>
<tr>
<td>H₂ + CH₃OH → CH₄ + H₂O</td>
<td>−25.7</td>
<td>−20.6</td>
<td>−24.6</td>
<td>−26.2</td>
</tr>
<tr>
<td>2H₂ + N₂ → NH₂–NH₂</td>
<td>+32.7</td>
<td>+31.6</td>
<td>+31.9</td>
<td>+30.7</td>
</tr>
<tr>
<td>H₂ + NH₂–NH₂ → 2NH₃</td>
<td>−47.4</td>
<td>−38.3</td>
<td>−43.1</td>
<td>−46.7</td>
</tr>
<tr>
<td>H₂ + H₂O₂ → 2H₂O</td>
<td>−81.7</td>
<td>−63.3</td>
<td>−71.1</td>
<td>−82.8</td>
</tr>
<tr>
<td>2H₂ + CO₂ → H₂O + H₂CO</td>
<td>+15.6</td>
<td>+26.6</td>
<td>+21.9</td>
<td>+14.5</td>
</tr>
<tr>
<td>CH₄ + CO → CH≡CH + H₂O</td>
<td>+55.9</td>
<td>+51.5</td>
<td>+43.7</td>
<td>+44.3</td>
</tr>
<tr>
<td>CH₄ + H₂O → CH₂=CH₂ + H₂O</td>
<td>+6.2</td>
<td>+10.4</td>
<td>+5.2</td>
<td>+1.1</td>
</tr>
<tr>
<td>CH₄ + CH₃OH → CH₃–CH₂ + H₂O</td>
<td>−12.1</td>
<td>−5.4</td>
<td>−7.6</td>
<td>−10.8</td>
</tr>
<tr>
<td>2CH₄ + N₂ → NH₂–NH₂ + CH₂=CH₂</td>
<td>+83.3</td>
<td>+75.4</td>
<td>+75.4</td>
<td>+76.6</td>
</tr>
<tr>
<td>CH₄ + H₂O₂ → CH₃OH + H₂O</td>
<td>−56.0</td>
<td>−42.6</td>
<td>−46.4</td>
<td>−56.6</td>
</tr>
<tr>
<td>2NH₃ + CH≡CH → NH₂–NH₂ + CH₂=CH₂</td>
<td>+8.7</td>
<td>−2.8</td>
<td>+5.1</td>
<td>+6.7</td>
</tr>
<tr>
<td>2NH₃ + CH₂=CH₂ → NH₂–NH₂ + CH₂=CH₂</td>
<td>+10.4</td>
<td>+9.7</td>
<td>+16.6</td>
<td>+16.2</td>
</tr>
<tr>
<td>NH₃ + HCN → CH₄ + N₂</td>
<td>−33.1</td>
<td>−40.5</td>
<td>−37.4</td>
<td>−37.6</td>
</tr>
<tr>
<td>NH₃ + CO → HCN + H₂O</td>
<td>+14.3</td>
<td>+13.8</td>
<td>+10.8</td>
<td>+12.0</td>
</tr>
<tr>
<td>2NH₃ + H₂O → NH₂–NH₂ + CH₃OH</td>
<td>+28.8</td>
<td>+25.5</td>
<td>+29.3</td>
<td>+28.2</td>
</tr>
<tr>
<td>H₂ + H₂O → CH₃OH</td>
<td>−18.7</td>
<td>−12.8</td>
<td>−13.8</td>
<td>−18.6</td>
</tr>
<tr>
<td>Oxirane + H₂O → OH–CH₂–CH₂–OH</td>
<td>−30.9</td>
<td>−20.5</td>
<td>−16.7</td>
<td>−19.6</td>
</tr>
<tr>
<td>Oxirane + NH₃ → NH₂–CH₂–CH₂–OH</td>
<td>−33.1</td>
<td>−20.4</td>
<td>−17.5</td>
<td>−22.6</td>
</tr>
<tr>
<td>HNCO + H₂O → NH₂–COOH</td>
<td>−4.3</td>
<td>−17.3</td>
<td>−11.1</td>
<td>−16.1</td>
</tr>
<tr>
<td>CH₂=NH + CH₄ + NH₃ → 2CH₃NH₂</td>
<td>−2.7</td>
<td>+1.1</td>
<td>+4.3</td>
<td>−0.6</td>
</tr>
<tr>
<td>H₂CO + CH₄ + H₂O → 2CH₃OH</td>
<td>+7.1</td>
<td>+7.9</td>
<td>+10.9</td>
<td>+7.6</td>
</tr>
<tr>
<td>HCN + 2CH₄ + 2NH₃ → 3CH₃NH₂</td>
<td>+17.4</td>
<td>+14.1</td>
<td>+21.4</td>
<td>+14.6</td>
</tr>
<tr>
<td>CO + 2CH₄ + 2H₂O → 3CH₃OH</td>
<td>+43.6</td>
<td>+28.6</td>
<td>+36.2</td>
<td>+36.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reaction</th>
<th>DFTB</th>
<th>cc-pVDZ</th>
<th>cc-pVTZ</th>
<th>G2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{H}_2 + \text{CH} \equiv \text{CH} \rightarrow \text{CH}_2 \equiv \text{CH}_2 )</td>
<td>-38.8</td>
<td>-41.0</td>
<td>-37.9</td>
<td>-40.1</td>
</tr>
<tr>
<td>( \text{H}_2 + \text{CH}_2 \equiv \text{CH} \rightarrow \text{CH}_3 \equiv \text{CH}_3 )</td>
<td>-37.0</td>
<td>-28.5</td>
<td>-26.5</td>
<td>-30.5</td>
</tr>
<tr>
<td>( 3\text{H}_2 + \text{HCN} \rightarrow \text{NH}_3 + \text{CH}_4 )</td>
<td>-47.9</td>
<td>-47.1</td>
<td>-48.5</td>
<td>-53.7</td>
</tr>
<tr>
<td>( \text{H}_2 + \text{CO} \rightarrow \text{H}_2 \text{CO} )</td>
<td>+10.9</td>
<td>+0.1</td>
<td>+0.7</td>
<td>+3.1</td>
</tr>
<tr>
<td>( 2\text{H}_2 + \text{CO} \rightarrow \text{CH}_3 \text{OH} )</td>
<td>-7.8</td>
<td>-12.6</td>
<td>-13.1</td>
<td>-15.5</td>
</tr>
<tr>
<td>( \text{H}_2 + \text{CH}_3 \text{OH} \rightarrow \text{CH}_4 + \text{H}_2 \text{O} )</td>
<td>-25.7</td>
<td>-20.6</td>
<td>-24.6</td>
<td>-26.2</td>
</tr>
<tr>
<td>( 2\text{H}_2 + \text{NH}_3 \rightarrow \text{NH}_2 \text{NH} )</td>
<td>+32.7</td>
<td>+33.6</td>
<td>+34.9</td>
<td>+30.7</td>
</tr>
<tr>
<td>( \text{NH}_3 + \text{CO} \rightarrow \text{HCN} + \text{H}_2 \text{O} )</td>
<td>+14.3</td>
<td>+13.8</td>
<td>+10.8</td>
<td>+12.0</td>
</tr>
<tr>
<td>( 2\text{NH}_3 + \text{H}_2 \text{CO} \rightarrow \text{NH}_2 \equiv \text{NH}_2 + \text{CH}_3 \text{OH} )</td>
<td>+28.8</td>
<td>+25.5</td>
<td>+29.3</td>
<td>+28.2</td>
</tr>
<tr>
<td>( \text{H}_2 + \text{H}_2 \text{CO} \rightarrow \text{CH}_3 \text{OH} )</td>
<td>-18.7</td>
<td>-12.8</td>
<td>-13.8</td>
<td>-18.6</td>
</tr>
<tr>
<td>( \text{Oxirane} + \text{H}_2 \text{O} \rightarrow \text{OH} \equiv \text{CH}_2 \equiv \text{CH}_2 \equiv \text{OH} )</td>
<td>-30.9</td>
<td>-20.5</td>
<td>-16.7</td>
<td>-19.6</td>
</tr>
<tr>
<td>( \text{Oxirane} + \text{NH}_3 \rightarrow \text{NH}_2 \equiv \text{CH}_2 \equiv \text{CH}_2 \equiv \text{OH} )</td>
<td>-33.1</td>
<td>-20.4</td>
<td>-17.5</td>
<td>-22.6</td>
</tr>
<tr>
<td>( \text{HNCO} + \text{H}_2 \text{O} \rightarrow \text{NH}_2 \equiv \text{COOH} )</td>
<td>-4.3</td>
<td>-17.3</td>
<td>-11.1</td>
<td>-16.1</td>
</tr>
<tr>
<td>( \text{CH}_2 \equiv \text{NH} + \text{CH}_4 + \text{NH}_3 \rightarrow 2\text{CH}_3 \text{NH}_2 )</td>
<td>-2.7</td>
<td>+1.1</td>
<td>+4.3</td>
<td>-0.6</td>
</tr>
<tr>
<td>( \text{H}_2 \text{CO} + \text{CH}_4 + \text{H}_2 \text{O} \rightarrow 2\text{CH}_3 \text{OH} )</td>
<td>+7.1</td>
<td>+7.9</td>
<td>+10.9</td>
<td>+7.6</td>
</tr>
<tr>
<td>( \text{HCN} + 2\text{CH}_4 + 2\text{NH}_3 \rightarrow 3\text{CH}_3 \text{NH}_2 )</td>
<td>+17.4</td>
<td>+14.1</td>
<td>+21.4</td>
<td>+14.6</td>
</tr>
<tr>
<td>( \text{CO} + 2\text{CH}_4 + 2\text{H}_2 \text{O} \rightarrow 3\text{CH}_3 \text{OH} )</td>
<td>+43.6</td>
<td>+28.6</td>
<td>+36.2</td>
<td>+36.8</td>
</tr>
</tbody>
</table>

With respect to G2:
- mean ave. dev.: 4.3 kcal/mole
- mean dev.: 1.5 kcal/mole
### Mean Absolute Errors in Calculated Heats of Formation for Neutral Molecules Containing the Elements C, H, N and O (kcal/mol).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>AM1</th>
<th>PM3</th>
<th>PDDG/PM3</th>
<th>SCC-DFTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrocarbons</td>
<td>254</td>
<td>5.6</td>
<td>3.6</td>
<td>2.6</td>
<td>4.8</td>
</tr>
<tr>
<td>All Molecules</td>
<td>622</td>
<td>6.8</td>
<td>4.4</td>
<td>3.2</td>
<td>5.8</td>
</tr>
<tr>
<td>Training Set</td>
<td>134</td>
<td>5.9</td>
<td>4.1</td>
<td>2.6</td>
<td>6.9</td>
</tr>
<tr>
<td>Test Set</td>
<td>488</td>
<td>7.0</td>
<td>4.4</td>
<td>3.4</td>
<td>5.5</td>
</tr>
</tbody>
</table>

SCC-DFTB Tests 3: Sattelmeyer & Jorgensen

*J. Phys. Chem. A 2006, 110, 13551*
**SCC-DFTB Tests 3: Sattelmeyer & Jorgensen**

Mean Absolute Errors in Calculated Heats of Formation for Neutral Molecules Containing the Elements C, H, N and O (kcal/mol).

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>AM1</th>
<th>PM3</th>
<th>PDDG/PM3</th>
<th>SCC-DFTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrocarbons</td>
<td>254</td>
<td>5.6</td>
<td>3.6</td>
<td>2.6</td>
<td>4.8</td>
</tr>
<tr>
<td>All Molecules</td>
<td>622</td>
<td>6.8</td>
<td>4.4</td>
<td>3.2</td>
<td>5.8</td>
</tr>
<tr>
<td>Training Set</td>
<td>134</td>
<td>5.9</td>
<td>4.1</td>
<td>2.6</td>
<td>6.9</td>
</tr>
<tr>
<td>Test Set</td>
<td>488</td>
<td>7.0</td>
<td>4.4</td>
<td>3.4</td>
<td>5.5</td>
</tr>
</tbody>
</table>

$$\Delta E_t = E_{mol}^{tot} - \sum_A E_{el}^A + \sum_A \Delta H_f^A$$

*Note: $E_{el}^A$ fitted!*

**J. Phys. Chem. A 2006, 110, 13551**
### SCC-DFTB Tests 3: Sattelmeyer & Jorgensen

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>AM1</th>
<th>PM3</th>
<th>PDDG/PM3</th>
<th>SCC-DFTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bond lengths (Å)</td>
<td>218</td>
<td>0.017</td>
<td>0.012</td>
<td>0.013</td>
<td>0.012</td>
</tr>
<tr>
<td>Bond angles (deg.)</td>
<td>126</td>
<td>1.5</td>
<td>1.7</td>
<td>1.9</td>
<td>1.0</td>
</tr>
<tr>
<td>Dihedral angles (deg.)</td>
<td>30</td>
<td>2.8</td>
<td>3.2</td>
<td>3.7</td>
<td>2.9</td>
</tr>
<tr>
<td>Dipole moments (D)</td>
<td>47</td>
<td>0.23</td>
<td>0.25</td>
<td>0.23</td>
<td>0.39</td>
</tr>
</tbody>
</table>

- N-O bond parameters need refinement
- Problem with sulfur: parameters need refinement
N-O bond parameters need refinement

DFTB shows overbinding
('build in' due to technical reasons)
nevertheless, reaction energies are quite reasonable i.e. overbinding is systematic
Other tests: Otte, Scholten & Thiel JPCA 111, 5753

<table>
<thead>
<tr>
<th>property(^a)</th>
<th>(N)^{\text{(b)}}</th>
<th>MNDO</th>
<th>AM1</th>
<th>PM3</th>
<th>OM1</th>
<th>OM1</th>
<th>OM2</th>
<th>OM3</th>
<th>DFTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta H_f) (kcal/mol)</td>
<td>140</td>
<td>6.3</td>
<td>5.5</td>
<td>4.2</td>
<td>3.5</td>
<td>3.1</td>
<td>2.9</td>
<td>7.7(^c)</td>
<td></td>
</tr>
<tr>
<td>(R) (pm)</td>
<td>242</td>
<td>1.4</td>
<td>1.7</td>
<td>1.1</td>
<td>1.2</td>
<td>1.6</td>
<td>2.0</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td>(\theta) (deg)</td>
<td>101</td>
<td>2.6</td>
<td>1.9</td>
<td>2.1</td>
<td>1.8</td>
<td>2.2</td>
<td>1.8</td>
<td>1.3</td>
<td></td>
</tr>
<tr>
<td>IP (eV)</td>
<td>52</td>
<td>0.46</td>
<td>0.35</td>
<td>0.42</td>
<td>0.32</td>
<td>0.26</td>
<td>0.45</td>
<td>3.82</td>
<td></td>
</tr>
<tr>
<td>(\mu) (D)</td>
<td>63</td>
<td>0.35</td>
<td>0.26</td>
<td>0.27</td>
<td>0.25</td>
<td>0.28</td>
<td>0.27</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>(\nu) (cm(^{-1}))</td>
<td>112</td>
<td>241</td>
<td>172</td>
<td>151</td>
<td>189</td>
<td>155</td>
<td>120</td>
<td>90</td>
<td></td>
</tr>
</tbody>
</table>

- geometries good
- vibrational frequencies ok
### SCC-DFTB frequencies:

Accuracy for vib. freq., problematic case e.g.:

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Expt.</th>
<th>BLYP/cc-pVTZ</th>
<th>DFTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH≡CH</td>
<td>1974</td>
<td>2009</td>
<td>2115</td>
</tr>
<tr>
<td>CH₂=CH₂</td>
<td>1621</td>
<td>1640</td>
<td>1824</td>
</tr>
<tr>
<td>CH₃–CH₃</td>
<td>995</td>
<td>957</td>
<td>1132</td>
</tr>
</tbody>
</table>

Special fit for vib. Frequencies:

Mean av. Err.: 59 cm⁻¹ → 33 cm⁻¹  for CH

Hydrogen bonding

Hydrogen bonds of 57 complexes

<table>
<thead>
<tr>
<th></th>
<th>AM1</th>
<th>OM2</th>
<th>SCC-DFTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
<td>2.8</td>
<td>1.5</td>
<td>2.7</td>
</tr>
<tr>
<td>R</td>
<td>0.25</td>
<td>0.20</td>
<td>0.08</td>
</tr>
<tr>
<td>$\theta$</td>
<td>33.7°</td>
<td>12.1°</td>
<td>6.2°</td>
</tr>
</tbody>
</table>

Energies: DFTB systematically too low!

Geometries:

- DFTB scatters around B3LYP values
- AM1 0.12 Å too long
- OM2 0.14 Å too short
Secondary-structure elements for Glycine und Alanine-based polypeptides


$N = 1$ (6 stable conformers)

$\beta$-helix stabilization by internal H-bonds between $i$ and $i+3$

$\alpha_R$-helix stabilization by internal H-bonds between $i$ and $i+4$

DFTB very good for:

- relative energies
- geometries
- vib. freq. o.k.!
Secondary-structure elements for Glycine und Alanine-based polypeptides


N = 1 (6 stable conformers)

\[
\begin{align*}
\alpha_R\text{-helix} \\
3_{10} \text{- helix}
\end{align*}
\]

stabilization by internal H-bonds between i and i+3

between i and i+4

DFTB very good for:

- relative energies
- geometries
- vib. freq. o.k.!

\[ \rightarrow \text{main problem for DFT(B): dispersion!} \]

\[ \rightarrow \text{AM1, PM3, MNDO quite bad} \]

\[ \rightarrow \text{OM2 much improved (JCC 22 (2001) 509)} \]
Glycine and Alanine based polypeptides in vacuo

Relative energies, structures and vibrational properties: N=1-8

N = 1  (6 stable conformers)

ΔE  relative energies (kcal/mole)

B3LYP
(6-31G*)
MP2
MP4-BSSE
SCC-DFTB

Ace-Ala-Nme

MP4-BSSE: Beachy et al, BSSE corrected at MP2 level
Secondary-structure elements for Alanine-based polypeptides

Otte, Scholten & Thiel JPCA 111, 5753

TABLE 5: Mean Absolute Deviations for the Peptide Test Set

<table>
<thead>
<tr>
<th></th>
<th>$N$</th>
<th>AM1</th>
<th>OM2</th>
<th>DFTB</th>
</tr>
</thead>
<tbody>
<tr>
<td>relative energies (kcal/mol)</td>
<td>22</td>
<td>2.0</td>
<td>1.7</td>
<td>1.1</td>
</tr>
<tr>
<td>backbone H-bond lengths (Å)</td>
<td>67</td>
<td>0.22</td>
<td>0.34</td>
<td>0.26</td>
</tr>
<tr>
<td>backbone dihedral angles (deg)</td>
<td>190</td>
<td>17.0</td>
<td>12.0</td>
<td>9.0</td>
</tr>
</tbody>
</table>
Weak interactions
Improving SCC-DFTB

1) DFT problems
   - VdW
   - LDA+U
   - GW

2) DFTB approximations
Observation in the 1990’s (Kohn, Pulay, Perez-Jorda, Yang...):

a) DFT-GGA does not cover dispersion interaction
b) not only correlation, but also exchange is problematic, e.g.:
   - B88-X too repulsive
   - PW91-X already (artificially) attractive (LDA anyway)

=> a solution has to worry about correlation as well as exchange
Dispersion: VdW interactions

Observation in the 1990’s (Kohn, Pulay, Perez-Jorda, Yang...):

a) DFT-GGA does not cover dispersion interaction
b) not only correlation, but also exchange is problematic, e.g.:
   - B88-X too repulsive
   - PW91-X already (artificially) attractive (LDA anyway)

=> a solution has to worry about correlation as well as exchange

Wu et al 2001, JCP 115, 8748
Dispersion: VdW interactions

Observation in the 1990’s (Kohn, Pulay, Perez-Jorda, Yang...):

a) DFT-GGA does not cover dispersion interaction
b) not only correlation, but also exchange is problematic, e.g.:

- B88-X too repulsive
- PW91-X already (artificially) attractive (LDA anyway)

=> a solution has to worry about correlation as well as exchange

Pavel Hobzas’ suggestion: put dispersion on top of DFTB like HF+dis
Dispersion: VdW interactions

Observation in the 1990’s (Kohn, Pulay, Perez-Jorda,Yang...):

a) DFT-GGA does not cover dispersion interaction
b) not only correlation, but also exchange is problematic, e.g.:
   - B88-X too repulsive
   - PW91-X already (artificially) attractive (LDA anyway)

=> a solution has to worry about correlation as well as exchange

Pavel Hobzas’ suggestion: put dispersion on top of DFTB like HF+dis

Consider the TB energy:

\[ E[\rho] \approx \sum_{i}^{\text{oocc}} \epsilon_i^H + E_{\text{rep}} \]
Dispersion

\( E_{\text{rep}} \) is short-ranged, cutoff between 2-3 Å !!

a) Therefore, we can add empirical dispersion without expecting double counting to be severe!

b) Confined basis set in DFTB: overlap-region of fragment densities reduced (also Pauli-repulsion underestimated!)
Dispersion forces - Van der Waals interactions
Elstner et al. JCP 114 (2001) 5149

Problem: in DFT (LDA/GGA) dispersion forces not included

\[ E_{\text{tot}} = E_{\text{SCC-DFTB}} - \sum_{\alpha\beta} f(R_{\alpha\beta}) \cdot C_{6\alpha\beta}^6 / R_{\alpha\beta}^6 \]

add an empirical term: \( C_{6\alpha\beta}^6 \) via Slater-Kirkwood combination rules of atomic polarizibilities after Halgreen, JACS 114 (1992) 7827.

damping \( f(R_{\alpha\beta}) = [1-\exp(-3(R_{\alpha\beta}/R_0)^7)]^3 \)
\[ R_0 = 3.8\text{Å (für O, N, C)} \]
DNA-base pair stacking only correct with


stacking energies in MP2/6-31G* (0.25), BSSE-corrected (+ MP2-values)

⇒ Hartree-Fock, very poor description
⇒ AM1, PM3, repulsive interaction (2-10) kcal/mole
⇒ MM-force fields strongly scatter in results

vertical dependence

twist-dependence
Large structures: need dispersion

O(N)-QMMM-MD: with W. Yang

DNA not stable without dispersion

Protein structures!

Instable within DFT-GGA

Liu et al. PROTEINS 44 (2001) 484

==> add dispersion to DFT, otherwise helices unfold
DFT + empirical dispersion: 1st generation

1) Problem of unbalanced Ex:
2) Problem of Ec?

→ Large variation in results when adding dispersion

Wu and Wang JCP 116 (2002) 515
Zimmerli et al. JCP 120 (2004) 2693
DFT + empirical dispersion: 1st generation

1) Problem of unbalanced Ex:
2) Problem of Ec?? Which one to choose?

→ Large variation in results when adding dispersion

Wu and Wang JCP 116 (2002) 515
Zimmerli et al. JCP 120 (2004) 2693

DFT + empirical dispersion: 2nd generation

Grimme JCC 25 (2004) 1463:
scale BLYP + disp with 1.4
scale PW91 + disp with 0.7
**DFT + empirical dispersion: 1st generation**

1) Problem of unbalanced Ex:

2) Problem of Ec?? Which one to choose?

→ Large variation in results when adding dispersion

- Wu and Wang JCP **116** (2002) 515
- Zimmerli et al. JCP **120** (2004) 2693

---

**DFT + empirical dispersion: 2nd generation**

- Grimme JCC **25** (2004) 1463:
  - scale BLYP + disp with 1.4
  - scale PW91 + disp with 0.7

![Dispersion Energy Equation](image)

- \( E_{\text{disp}} = -s_6 \sum_{i=1}^{N_{\text{at}}-1} \sum_{j=i+1}^{N_{\text{at}}} \frac{C_{ij}}{R_{ij}^6} f_{\text{dmp}}(R_{ij}) \)

---

**3rd generation**: revPBE, XLYP and s6=1
Improving SCC-DFTB

\[ E = \sum_{i}^{\text{occ}} \langle \psi_i | \hat{H}^0 | \psi_i \rangle + \frac{1}{2} \iint' \left( \frac{1}{|\mathbf{r} - \mathbf{r}'|} + \left. \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \right|_{n_0} \right) \Delta \rho \Delta \rho' \]

\[ - \frac{1}{2} \iint' \frac{\rho' \rho_0}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[\rho_0] - \int V_{xc}[\rho_0] n_0 + E_{cc} \]

- minimal basis
- neglect of crystal field and three-center terms
- initial density fixed
- second order expansion
- monopole approximation
- gamma
- two-body approximation
- fit procedure
Approximations

1. repulsive energy
   - two-body approximation: seems to work ok, since bond angles quite good
   - short ranged repulsive: leads to overbinding

2. ‘gamma’ approximation
   - assumes hardness-atomic size relation, which is not generally valid

3. second-order approximation
   - localized charges: extension to third order

4. minimal basis
   - problems of N-proton affinities, hybridization?
   - VdW distances too short: better with DZ-basis?

5. neglect of three center/crystal field integrals
   - seems to be compensated by repulsive potential, however
   - transferability between single, double and triple bonds?
   - highly coordinated phases?

6. monopole approximation
   - improved transferability?

7. fixed initial density
   - improved transferability?
Inspection of Gamma function

\[ E^{2nd} = \frac{1}{2} \int \int' \left( \frac{1}{|r - r'|} + \frac{\delta^2 E^{xc}[\rho]}{\delta \rho \delta \rho'} \right) \Delta \rho \Delta \rho' \]

\[ \Rightarrow \quad E^\gamma = \frac{1}{2} \sum_{ab} \Delta q_a \Delta q_b \gamma_{ab} \]
Inspection of Gamma function

\[
E^{2nd} = + \frac{1}{2} \int \int' \left( \frac{1}{|r - r'|} + \frac{\delta^2 E_{\rho \delta \rho}}{\rho \delta \rho} \right) \Delta \rho \Delta \rho'
\]

??

\[
E^\gamma = \frac{1}{2} \sum_{ab} \Delta q_a \Delta q_b \gamma_{ab}
\]
Inspection of Gamma function

\[ E^{2nd} = +\frac{1}{2} \int \int' \left( \frac{1}{|r - r'|} + \frac{\delta^2 E_{\rho}}{\rho \delta \rho} \right) \Delta \rho \Delta \rho' \]

\[ \Delta \rho = \sum_a \Delta \rho_a \quad \Delta \rho_a \cong \Delta q_a F_a^{00} \gamma^{00} \]

\[ F_a^{00} = \frac{\tau_a}{8\pi} \exp(-\tau_a |r - R_a|) \quad \gamma_{ab} = \frac{1}{r_{ab}} - S(\tau_a, \tau_b, r_{ab}) \]
Inspection of Gamma function

\[ E^{2nd} = + \frac{1}{2} \int \int' \left( \frac{1}{|r - r'|} + \frac{\delta^2 E^{xc}[\rho]}{\delta \rho \delta \rho'} \bigg|_{\rho^0} \right) \Delta \rho \Delta \rho' \]

\[ \Delta \rho = \sum_a \Delta \rho_a \]
\[ \Delta \rho_a \cong \Delta q_a F^{00}_a Y^{00} \]

\[ F^{00}_a = \frac{\tau_a}{8\pi} \exp(-\tau_a |r - R_a|) \]

\[ \gamma_{ab} = \frac{1}{r_{ab}} - S(\tau_a, \tau_b, r_{ab}) \]

\[ \gamma_{aa} = \frac{5}{16} \tau_a = U_a \]

Hubbard parameter = twice the chemical hardness

\[ \gamma_{ab} = \frac{1}{r_{ab}} - S(U_a, U_b, r_{ab}) \]
Inspection of Gamma function

\[ E^{2nd} = \frac{1}{2} \int \int' \left( \frac{1}{|r - r'|} + \frac{\partial^2 E_{xc}[\rho]}{\partial \rho \partial \rho'} \bigg|_{\rho^0} \right) \Delta \rho \Delta \rho' \]

\[ \Delta \rho = \sum_a \Delta \rho_a \]

\[ \Delta \rho_a \approx \Delta q_a F_{a}^{00} \Sigma^{00} \]

\[ F_{a}^{00} = \frac{\tau_a}{8\pi} \exp(-\tau_a |r - R_a|) \]

\[ \gamma_{ab} = \frac{1}{r_{ab}} - S(\tau_a, \tau_b, r_{ab}) \]

\[ \gamma_{aa} = \frac{5}{16} \tau_a = U_a \]

Hubbard parameter

= twice the chemical hardness
Inspection of Gamma function

\[ E^{2nd} = \frac{1}{2} \int \int' \left( \frac{1}{|r - r'|} + \frac{\delta^2 E^{xc}[\rho]}{\delta \rho \delta \rho'} \bigg|_{\rho^0} \right) \Delta \rho \Delta \rho' \]

\[ E^\gamma = \frac{1}{2} \sum_{ab} \Delta q_a \Delta q_b \gamma_{ab} \]

\[ \Delta \rho = \sum_a \Delta \rho_a \quad \Delta \rho_a \cong \Delta q_a F_a^{00} \gamma^{00} \]

\[ F_a^{00} = \frac{\tau_a}{8\pi} \exp(-\tau_a |r - R_a|) \]

\[ \gamma_{ab} = \frac{1}{r_{ab}} - S(\tau_a, \tau_b, r_{ab}) \]

\[ r \to 0 \quad \text{taking into account} \quad \frac{\delta^2 E^{xc}[\rho]}{\delta \rho \delta \rho'} \bigg|_{\rho^0} \]

\[ \gamma_{aa} = \frac{5}{16} \tau_a = U_a \]

Hubbard parameter = twice the chemical hardness

Average distance electron-nucleus

\[ R_a^{cov} = \int_0^\infty (r - R_a) F_a^{00} \, d(r - R_a) = \frac{1}{\tau_a} = \frac{5}{16 U_a} \]

\[ 3.2 \times R_a^{cov} = \frac{1}{U_a} \]

\[ U_a = \frac{5}{16} \tau_a \]

\[ \gamma_{ab} = \frac{1}{r_{ab}} - S(U_a, U_b, r_{ab}) \]
deviation from $1/R$ for small $R$

$R=0:\ 1/\sigma = 3.2 \ U_{\text{hubbard}}$

$\rho_\alpha \sim \exp\left(-\frac{r - R_\alpha}{\sigma}\right)$

Is this valid throughout the periodic table?

What is the relation between 'atomic size' and chemical hardness?
**U vs $R_{\text{cov}}$: Hydrogen atom**

U-Hubbard

![Graph showing U vs $R_{\text{cov}}$ for Hydrogen, Oxygen, Nitrogen, Carbon, and Silicon atoms.](image)
\[ \text{Gamma requires: } 3.2 \times R_{\text{cov}} = \frac{1}{U} \]

\[ \text{‘size’ of H overestimated based on hardness value: H has same ‘size’ as N!} \]

In DFTB, H is 0.73 Å instead of 0.33 Å!
On-site interaction and coulomb scaling: H

- $U_H$ for the on-site interaction of H should not be changed!
- However, $U_H$ is a bad measure for the size of H!

→ Leads to too 'large' H-atoms! I.e. coulomb interaction is damped too fast due to 'artificial' overlap effect!
→ modify coulomb-scaling for H!
Modified Gamma for H-bonding
change only X-H interaction!

\[ \gamma_{\alpha\beta} = \frac{1}{R - S} \rightarrow \gamma_{\alpha\beta} = \frac{1}{R - S} \cdot f_{damp} \]

\[ \gamma_{\alpha\beta} = \frac{1}{\sqrt{R^2 + \frac{1}{4} (U^{-1}_\alpha + U^{-1}_\beta)^2 \exp(-\eta R^2)}} \]
Modified Gamma for H-bonding

-Water dimer: 3.3 kcal
  4.6 kcal

standard DFTB: H-bonds ~ 1-2 kcal too low
mod Gamma: ~0.3-0.5 kcal too low


introduce one parameter!
Improving SCC-DFTB

\[ E = \sum_{i}^{\text{occ}} \langle \psi_i | \hat{H}^0 | \psi_i \rangle + \frac{1}{2} \int \int' \left( \frac{1}{|\vec{r} - \vec{r}'|} + \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \bigg|_{n_0} \right) \Delta \rho \Delta \rho' \]

\[ - \frac{1}{2} \int \int' \frac{\rho'_0 \rho_0}{|\vec{r} - \vec{r}'|} + E_{xc}[\rho_0] - \int V_{xc}[\rho_0] n_0 + E_{cc} \]

- minimal basis
- neglect of crystal field and three-center terms
- initial density fixed
- second order expansion
- monopole approximation
- gamma
- two-body approximation
- fit procedure
Third order expansion

\[ \frac{1}{6} \int \int' \int'' \left[ \frac{\delta^3 E[\rho]}{\delta \rho \delta \rho' \delta \rho''} \right]_{\rho_0} \Delta \rho \Delta \rho' \Delta \rho'' \]

monopole expansion
onsite only

\[ E^{3rd} = \frac{1}{6} \sum_{\alpha} \frac{\partial U_{\alpha}}{\partial q_{\alpha}} \Delta q_{\alpha}^3 \]

compute Hubbard derivative \( \partial U/\partial q \) from DFT

2nd+3rd onsite terms: \( U(q) = U(q_0) + 0.5 \frac{\partial U}{\partial q} \Delta q \)

Hubbard becomes charge dependent

=> DFTB atoms can ‘change size’ dependent on charge state
Third order expansion

Example: \( \text{H}_2\text{O} \rightarrow \text{OH}^- + \text{H}^+ \)

Description of \( \text{OH}^- \):
- charge very localized
- ‘size’ of \( \text{O} \) changes significantly

is the approximation of a constant Hubbard value (chemical hardness) appropriate?

Deprotonation energy
- B3LYP/6-311++G(2d2p): 397 kcal/mole
- SCC-DFTB: 424 kcal/mole
Third order DFTB

<table>
<thead>
<tr>
<th></th>
<th>6-31G**</th>
<th>6-31+G**</th>
<th>6-311++g(2d,2p)</th>
<th>SCC-DFTB</th>
<th>Uq</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂O</td>
<td>435.1</td>
<td>395.8</td>
<td><strong>397.2</strong></td>
<td>424.1</td>
<td>392.1</td>
</tr>
<tr>
<td>CH₃OH</td>
<td>406.3</td>
<td>389.4</td>
<td><strong>389.3</strong></td>
<td>397.5</td>
<td>386.4</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>367.8</td>
<td>354.5</td>
<td>355.4</td>
<td>361.6</td>
<td>356.5</td>
</tr>
<tr>
<td>HCOOH</td>
<td>366.1</td>
<td>348.2</td>
<td>349.4</td>
<td>363.7</td>
<td>353.7</td>
</tr>
<tr>
<td>H₃COOH</td>
<td>369.8</td>
<td>352.9</td>
<td>353.7</td>
<td>365.5</td>
<td>355.8</td>
</tr>
</tbody>
</table>

B3LYP vs SCC-DFTB and 3rd order correction Uq:
- basis set dependence
- localized charges on small anions

-U(q): changes “size” of atom: $R_{\text{cov}} \sim 1/U$

3rd order expansion

\[ \frac{1}{6} \int \int \int \left[ \frac{\delta^3 E^{xc}[\rho]}{\delta \rho \delta \rho' \delta \rho''} \right]_{\rho_0} \Delta \rho \Delta \rho' \Delta \rho'' = \frac{1}{6} \int \int \int \left[ \frac{\delta^2 E^{xc}[\rho]}{\delta \rho \delta \rho'} \right]_{\rho_0} \Delta \rho \Delta \rho' \Delta \rho'' \]

\[ E^\Gamma = \frac{1}{6} \sum_{abc} \Delta q_a \Delta q_b \Delta q_c \frac{d\gamma_{ab}}{dq_c} \bigg|_{q_0^c} \]

\[ = \frac{1}{3} \sum_{ab} \Delta q_a^2 \Delta q_b \Gamma_{ab} \]

\[ \Gamma_{ab} = \frac{\partial \gamma_{ab}}{\partial q_a} \bigg|_{q_0^a} = \frac{\partial \gamma_{ab}}{\partial U_a} \frac{\partial U_a}{\partial q_a} \bigg|_{q_0^a} \]

\[ \Gamma_{aa} = \frac{1}{2} \frac{\partial U_a}{\partial q_a} \bigg|_{q_0^a} \]

Charge dependence of U for Carbon atom

U\textsuperscript{d}: parameter calculated from DFT analytically
3rd order expansion

\[
\frac{1}{6} \int \int \int \int \left[ \frac{\delta^3 E^{xc}[\rho]}{\delta \rho \delta \rho' \delta \rho''} \right] \rho_0 \Delta \rho \Delta \rho' \Delta \rho'' = \frac{1}{6} \int \int \int \int
\]

\[
E^\Gamma = \frac{1}{6} \sum_{abc} \Delta q_a \Delta q_b \Delta q_c \frac{d \gamma_{ab}}{dq_c} \bigg|_{q_c^0}
\]

\[
\gamma_{ab} = \begin{cases} 
\frac{1}{r} - S^f & r \neq 0, \alpha \neq \beta \\
\frac{5}{16} \alpha & r = 0
\end{cases}
\]

\[
\alpha = \frac{16}{5} U_a, \beta = \frac{16}{5} U_b, \quad r = |R_b - R_a|
\]

\[
S^f = e^{-\alpha r} f(\alpha, \beta, r) + e^{-\beta r} f(\beta, \alpha, r)
\]

\[
S^g = e^{-\alpha r} g(\alpha, r)
\]

\[
f(\alpha, \beta, r) = \frac{\alpha \beta^4}{2(\alpha^2 - \beta^2)^2} - \frac{\beta^6 - 3\alpha^2 \beta^4}{(\alpha^2 - \beta^2)^3 r}
\]

\[
g(\alpha, r) = \frac{1}{48r} \left( 48 + 33 \alpha r + 9 \alpha^2 r^2 + \alpha^3 r^3 \right)
\]

\[\Gamma_{ab} = \frac{\partial \gamma_{ab}}{\partial q_a} \bigg|_{q_a^0} = \frac{\partial \gamma_{ab}}{\partial U_a} \frac{\partial U_a}{\partial q_a} \bigg|_{q_a^0}
\]

\[\Gamma_{aa} = \frac{1}{2} \frac{\partial U_a}{\partial q_a} \bigg|_{q_a^0}
\]

U_d: parameter calculated from DFT analytically.
3rd order expansion

\[
\frac{1}{6} \int \int' \int'' \left[ \frac{\delta^3 E^{xc}[\rho]}{\delta \rho \delta \rho' \delta \rho''} \right]_{\rho^0} \Delta \rho \Delta \rho' \Delta \rho'' = \frac{1}{6} \int \int' \int'' \frac{\delta}{\delta \rho''} \left[ \frac{\delta^2 E^{xc}[\rho]}{\delta \rho \delta \rho'} \right]_{\rho^0} \Delta \rho \Delta \rho' \Delta \rho''
\]

\[
E^\Gamma = \frac{1}{6} \sum_{abc} \Delta q_a \Delta q_b \Delta q_c \frac{d \gamma_{ab}}{dq_c} \bigg|_{q_c^0}
= \frac{1}{3} \sum_{ab} \Delta q_a^2 \Delta q_b \Gamma_{ab}
\]

\[
\Gamma_{ab} = \left. \frac{\partial \gamma_{ab}}{\partial q_a} \right|_{q_a^0} = \left. \frac{\partial \gamma_{ab}}{\partial q_a} \frac{\partial U_a}{\partial q_a} \right|_{q_a^0}
\]

\[
\Gamma_{aa} = \frac{1}{2} \left. \frac{\partial U_a}{\partial q_a} \right|_{q_a^0}
\]

\[U_d^3: \text{parameter} \quad \text{Calculated from DFT}\]

Charge dependence of $U$ for Carbon atom

![Graph showing the charge dependence of $U$ for Carbon atom.](image)
## Performance 3rd order

Parameter sets for 3rd order including improved gamma

<table>
<thead>
<tr>
<th>set</th>
<th>$U^d_{\text{CHNO}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Calculated</td>
</tr>
<tr>
<td>1</td>
<td>Fitted</td>
</tr>
</tbody>
</table>

Fitted to 22 Binding energies and 23 Proton affinities
Performance 3rd order

Parameter sets for 3rd order including improved gamma

<table>
<thead>
<tr>
<th>set</th>
<th>$U^d_{\text{CHNO}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Calculated</td>
</tr>
<tr>
<td>1</td>
<td>Fitted</td>
</tr>
</tbody>
</table>

Fitted to 22 Binding energies and 23 Proton affinities

Geometries for neutral molecules

No significant change!

MAD for G2 molecule test set (CHNO)
Bond lengths 0.014Å
Angles 0.9°
Performance 3rd order

Parameter sets for 3rd order including improved gamma

<table>
<thead>
<tr>
<th>set</th>
<th>U^d_{CHNO}</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Calculated</td>
</tr>
<tr>
<td>1</td>
<td>Fitted</td>
</tr>
</tbody>
</table>

Fitted to 22 Binding energies and 23 Proton affinities

Geometries for neutral molecules

No significant change!

MAD for G2 molecule test set (CHNO)
Bond lengths 0.014 Å
Angles 0.9°

Effect for charged systems...

B3LYP 1.567 Å 0.971 Å
2nd 1.614 Å 1.004 Å
3rd-0 1.572 Å 0.968 Å

... and hydrogen bonds

PBE 1.917 Å
2nd 1.889 Å
3rd-0 1.829 Å
Performance 3rd order

Parameter sets for 3rd order including improved gamma

<table>
<thead>
<tr>
<th>set</th>
<th>$U_{\text{CHNO}}^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Calculated</td>
</tr>
<tr>
<td>1</td>
<td>Fitted</td>
</tr>
</tbody>
</table>

Fitted to 22 Binding energies and 23 Proton affinities

Effect for charged systems...

B3LYP | 1.567 Å | 0.971 Å |
2nd   | 1.614 Å | 1.004 Å |
3rd-0 | 1.572 Å | 0.968 Å |

PBE   | 1.917 Å |
2nd   | 1.889 Å |
3rd-0 | 1.829 Å |

Proton Transfer Barriers: Deviation to G3B3 (kcal/mol)

<table>
<thead>
<tr>
<th>$r_{\text{OO}}$(Å)</th>
<th>G3B3</th>
<th>2nd</th>
<th>3rd-0</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>0.3</td>
<td>-0.3</td>
<td>-0.3</td>
</tr>
<tr>
<td>2.6</td>
<td>2.2</td>
<td>-1.3</td>
<td>-0.9</td>
</tr>
<tr>
<td>2.7</td>
<td>5.1</td>
<td>-1.1</td>
<td>-0.6</td>
</tr>
<tr>
<td>2.8</td>
<td>8.8</td>
<td>-1.3</td>
<td>-0.9</td>
</tr>
</tbody>
</table>

... and hydrogen bonds

MAD for G2 molecule test set (CHNO)
Bond lengths 0.014 Å
Angles 0.9°

Geometries for neutral molecules

No significant change!

for 2nd and 3rd
### Performance 3rd order

Deviation 22 Binding energies (kcal/mol) for Hydrogen bonds

<table>
<thead>
<tr>
<th>molecule</th>
<th>G3B3</th>
<th>SCC 3rd-diag calc</th>
<th>SCC 3rd-diag fit</th>
<th>3rd calc</th>
<th>3rd fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>2H2O</td>
<td>-4.9</td>
<td>+1.6</td>
<td>-0.0</td>
<td>+0.2</td>
<td>-0.0</td>
</tr>
<tr>
<td>3H2O</td>
<td>-15.1</td>
<td>+5.5</td>
<td>-0.5</td>
<td>+0.2</td>
<td>-0.3</td>
</tr>
<tr>
<td>4H2O</td>
<td>-27.4</td>
<td>+9.7</td>
<td>+0.8</td>
<td>+1.8</td>
<td>+0.8</td>
</tr>
<tr>
<td>5H2O</td>
<td>-36.3</td>
<td>+13.3</td>
<td>+1.8</td>
<td>+3.0</td>
<td>+1.3</td>
</tr>
<tr>
<td>2H2O(H⁺)</td>
<td>-33.9</td>
<td>+4.5</td>
<td>+2.4</td>
<td>+3.4</td>
<td>+0.9</td>
</tr>
<tr>
<td>3H2O(H⁺)</td>
<td>-57.3</td>
<td>+10.4</td>
<td>+5.3</td>
<td>+6.6</td>
<td>+3.7</td>
</tr>
<tr>
<td>4H2O(H⁺)</td>
<td>-77.2</td>
<td>+13.9</td>
<td>+6.4</td>
<td>+7.9</td>
<td>+5.0</td>
</tr>
<tr>
<td>5H2O(H⁺)</td>
<td>-91.9</td>
<td>+18.3</td>
<td>+7.2</td>
<td>+9.1</td>
<td>+6.2</td>
</tr>
<tr>
<td>2H2O(-H⁺)</td>
<td>-27.4</td>
<td>-5.1</td>
<td>-3.4</td>
<td>-1.5</td>
<td>-5.9</td>
</tr>
<tr>
<td>3H2O(-H⁺)</td>
<td>-48.6</td>
<td>-2.6</td>
<td>-6.5</td>
<td>-3.8</td>
<td>-8.4</td>
</tr>
<tr>
<td>4H2O(-H⁺)</td>
<td>-66.7</td>
<td>+0.3</td>
<td>-5.0</td>
<td>-1.8</td>
<td>-7.2</td>
</tr>
<tr>
<td>5H2O(-H⁺)</td>
<td>-86.3</td>
<td>+6.1</td>
<td>-7.7</td>
<td>-4.1</td>
<td>-7.8</td>
</tr>
<tr>
<td>NH₃(H₂O) d</td>
<td>-6.6</td>
<td>+3.2</td>
<td>+2.0</td>
<td>+2.2</td>
<td>+2.1</td>
</tr>
<tr>
<td>NH₄⁺(H₂O) d</td>
<td>-20.4</td>
<td>+0.6</td>
<td>-1.1</td>
<td>-0.7</td>
<td>-1.3</td>
</tr>
<tr>
<td>6H2O_book</td>
<td>-45.8</td>
<td>+16.7</td>
<td>+1.5</td>
<td>+3.2</td>
<td>+1.7</td>
</tr>
<tr>
<td>6H2O_cage</td>
<td>-46.6</td>
<td>+17.2</td>
<td>+0.3</td>
<td>+2.1</td>
<td>+1.5</td>
</tr>
<tr>
<td>6H2O_prism</td>
<td>-47.2</td>
<td>+17.6</td>
<td>+0.1</td>
<td>+2.0</td>
<td>+1.3</td>
</tr>
<tr>
<td>6H2O_ring</td>
<td>-44.7</td>
<td>+16.5</td>
<td>+2.4</td>
<td>+3.9</td>
<td>+1.5</td>
</tr>
<tr>
<td>methylimidazole(-H⁺)(H₂O)</td>
<td>-15.9</td>
<td>+4.1</td>
<td>+1.4</td>
<td>+1.5</td>
<td>+1.2</td>
</tr>
<tr>
<td>methylimidazole(H₂O)_1</td>
<td>-6.2</td>
<td>+2.4</td>
<td>+1.8</td>
<td>+1.9</td>
<td>+1.9</td>
</tr>
<tr>
<td>methylimidazole(H₂O)_2</td>
<td>-8.2</td>
<td>+3.5</td>
<td>+2.0</td>
<td>+2.2</td>
<td>+1.9</td>
</tr>
<tr>
<td>methylimidazoleH⁺(H₂O)</td>
<td>-16.0</td>
<td>+3.3</td>
<td>+2.2</td>
<td>+2.3</td>
<td>+2.5</td>
</tr>
<tr>
<td>MUE</td>
<td>8.0</td>
<td>2.8</td>
<td>3.0</td>
<td>2.9</td>
<td>2.7</td>
</tr>
<tr>
<td>MSE</td>
<td>+7.3</td>
<td>+0.6</td>
<td>+1.9</td>
<td>+0.1</td>
<td>+1.1</td>
</tr>
<tr>
<td>MAX</td>
<td>18.3</td>
<td>7.7</td>
<td>9.1</td>
<td>8.4</td>
<td>8.1</td>
</tr>
</tbody>
</table>
Performance of DFT with small basis set

<table>
<thead>
<tr>
<th>molecule</th>
<th>G3B3</th>
<th>without counterpoise correction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>PBE 6-31+G(d,p)</td>
</tr>
<tr>
<td>2H2O</td>
<td>-4.9</td>
<td>-1.8</td>
</tr>
<tr>
<td>3H2O</td>
<td>-15.1</td>
<td>-5.0</td>
</tr>
<tr>
<td>4H2O</td>
<td>-27.4</td>
<td>-8.7</td>
</tr>
<tr>
<td>5H2O</td>
<td>-36.3</td>
<td>-11.3</td>
</tr>
<tr>
<td>2H2O(H+)</td>
<td>-33.9</td>
<td>-5.4</td>
</tr>
<tr>
<td>3H2O(H+)</td>
<td>-57.3</td>
<td>-7.7</td>
</tr>
<tr>
<td>4H2O(H+)</td>
<td>-77.2</td>
<td>-8.7</td>
</tr>
<tr>
<td>5H2O(H+)</td>
<td>-91.9</td>
<td>-10.6</td>
</tr>
<tr>
<td>2H2O(-H+)</td>
<td>-27.4</td>
<td>-5.3</td>
</tr>
<tr>
<td>3H2O(-H+)</td>
<td>-48.6</td>
<td>-8.6</td>
</tr>
<tr>
<td>4H2O(-H+)</td>
<td>-66.7</td>
<td>-11.5</td>
</tr>
<tr>
<td>5H2O(-H+)</td>
<td>-86.3</td>
<td>-10.4</td>
</tr>
<tr>
<td>NH3(H2O)</td>
<td>-6.6</td>
<td>-2.3</td>
</tr>
<tr>
<td>NH4+(H2O)</td>
<td>-20.4</td>
<td>-2.7</td>
</tr>
<tr>
<td>6H2O_books</td>
<td>-45.8</td>
<td>-13.1</td>
</tr>
<tr>
<td>6H2O_cage</td>
<td>-46.6</td>
<td>-11.5</td>
</tr>
<tr>
<td>6H2O_prism</td>
<td>-47.2</td>
<td>-10.8</td>
</tr>
<tr>
<td>6H2O_ring</td>
<td>-44.7</td>
<td>-14.1</td>
</tr>
<tr>
<td>methylimidazole(-H+)(H2O)</td>
<td>-15.9</td>
<td>-2.6</td>
</tr>
<tr>
<td>methylimidazole(H2O)_1</td>
<td>-6.2</td>
<td>-0.6</td>
</tr>
<tr>
<td>methylimidazole(H2O)_2</td>
<td>-8.2</td>
<td>-1.0</td>
</tr>
<tr>
<td>methylimidazole(H+)(H2O)</td>
<td>-16.0</td>
<td>-1.0</td>
</tr>
<tr>
<td>MUE</td>
<td></td>
<td>7.0</td>
</tr>
<tr>
<td>MSE</td>
<td></td>
<td>-7.0</td>
</tr>
<tr>
<td>MAX</td>
<td></td>
<td>14.1</td>
</tr>
</tbody>
</table>
Performance 3rd order

<table>
<thead>
<tr>
<th>molecule</th>
<th>G3B3</th>
<th>SCC</th>
<th>3rd-diag</th>
<th>3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>calc</td>
<td>fit</td>
</tr>
<tr>
<td>H₂O</td>
<td>398.4</td>
<td>+16.3</td>
<td>+5.8</td>
<td>-1.6</td>
</tr>
<tr>
<td>2H₂O</td>
<td>375.9</td>
<td>+9.6</td>
<td>+2.3</td>
<td>-3.2</td>
</tr>
<tr>
<td>3H₂O</td>
<td>365.0</td>
<td>+8.1</td>
<td>-0.4</td>
<td>-5.6</td>
</tr>
<tr>
<td>4H₂O</td>
<td>359.1</td>
<td>+7.0</td>
<td>-0.1</td>
<td>-5.2</td>
</tr>
<tr>
<td>5H₂O</td>
<td>348.4</td>
<td>+9.2</td>
<td>-3.7</td>
<td>-8.7</td>
</tr>
<tr>
<td>CH₃OH</td>
<td>392.6</td>
<td>-5.7</td>
<td>+5.8</td>
<td>-0.7</td>
</tr>
<tr>
<td>CH₂CH₂OH</td>
<td>388.3</td>
<td>-1.5</td>
<td>+9.3</td>
<td>+2.6</td>
</tr>
<tr>
<td>CH₃CH₂CH₂OH</td>
<td>387.6</td>
<td>-2.2</td>
<td>+8.7</td>
<td>+1.9</td>
</tr>
<tr>
<td>CH₃-CH(OH)-CH₃</td>
<td>385.6</td>
<td>+1.4</td>
<td>+11.7</td>
<td>+4.6</td>
</tr>
<tr>
<td>HCOOH</td>
<td>351.2</td>
<td>+1.7</td>
<td>+14.2</td>
<td>+7.1</td>
</tr>
<tr>
<td>CH₂COOH</td>
<td>355.1</td>
<td>+1.1</td>
<td>+12.7</td>
<td>+5.6</td>
</tr>
<tr>
<td>CH₃CH₂COOH</td>
<td>354.5</td>
<td>+1.0</td>
<td>+13.1</td>
<td>+6.0</td>
</tr>
<tr>
<td>C₆H₅OH</td>
<td>356.7</td>
<td>-4.7</td>
<td>+11.0</td>
<td>+5.2</td>
</tr>
<tr>
<td>p-CH₃-C₆H₄OH</td>
<td>357.9</td>
<td>-5.6</td>
<td>+10.5</td>
<td>+4.5</td>
</tr>
<tr>
<td>p-NO₂-C₆H₄OH</td>
<td>334.6</td>
<td>-9.3</td>
<td>+5.2</td>
<td>-0.7</td>
</tr>
<tr>
<td>H₃O⁺</td>
<td>171.2</td>
<td>-0.4</td>
<td>+9.0</td>
<td>+5.6</td>
</tr>
<tr>
<td>2H₂O(H⁺)</td>
<td>200.2</td>
<td>-3.3</td>
<td>+6.6</td>
<td>+2.3</td>
</tr>
<tr>
<td>3H₂O(H⁺)</td>
<td>213.4</td>
<td>-5.3</td>
<td>+3.2</td>
<td>-0.9</td>
</tr>
<tr>
<td>4H₂O(H⁺)</td>
<td>221.1</td>
<td>-4.7</td>
<td>+3.3</td>
<td>-0.7</td>
</tr>
<tr>
<td>5H₂O(H⁺)</td>
<td>226.7</td>
<td>-5.3</td>
<td>+3.8</td>
<td>-0.4</td>
</tr>
<tr>
<td>CH₃OH⁺</td>
<td>186.8</td>
<td>-8.3</td>
<td>+6.1</td>
<td>+2.0</td>
</tr>
<tr>
<td>H₂COOH⁺</td>
<td>177.1</td>
<td>-11.8</td>
<td>+4.3</td>
<td>+0.5</td>
</tr>
<tr>
<td>CH₃CHOH⁺</td>
<td>190.2</td>
<td>-10.1</td>
<td>+6.5</td>
<td>+2.4</td>
</tr>
<tr>
<td>MUE</td>
<td></td>
<td>5.8</td>
<td>6.8</td>
<td>3.4</td>
</tr>
<tr>
<td>MSE</td>
<td></td>
<td>-1.0</td>
<td>+6.5</td>
<td>+1.0</td>
</tr>
<tr>
<td>MAX</td>
<td></td>
<td>16.3</td>
<td>14.2</td>
<td>8.7</td>
</tr>
</tbody>
</table>

Deviation 23
Proton affinities (kcal/mol) for acidic oxygen

MSE =5.2 kcal: this reflects the ‘overbinding’ in the O-H repulsive potential:

i.e. 3rd order ‘works nearly perfect’ fitting resolves errors from Erepa
## Performance of DFT with small basis set

<table>
<thead>
<tr>
<th>molecule</th>
<th>G3B3</th>
<th>6-31+G(d,p)</th>
<th>6-311(2d,2p)</th>
<th>G3large</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2O</td>
<td>398.4</td>
<td>-4.5</td>
<td>31.4</td>
<td>-3.2</td>
</tr>
<tr>
<td>2H2O</td>
<td>375.9</td>
<td>-8.0</td>
<td>12.6</td>
<td>-6.9</td>
</tr>
<tr>
<td>3H2O</td>
<td>365.0</td>
<td>-8.3</td>
<td>7.4</td>
<td>-7.2</td>
</tr>
<tr>
<td>4H2O</td>
<td>359.1</td>
<td>-7.4</td>
<td>4.7</td>
<td>-6.9</td>
</tr>
<tr>
<td>5H2O</td>
<td>348.4</td>
<td>-3.6</td>
<td>0.1</td>
<td>-3.8</td>
</tr>
<tr>
<td>CH3OH</td>
<td>392.6</td>
<td>-7.6</td>
<td>4.2</td>
<td>-8.3</td>
</tr>
<tr>
<td>CH3CH2OH</td>
<td>388.3</td>
<td>-7.7</td>
<td>3.3</td>
<td>-8.5</td>
</tr>
<tr>
<td>CH3CH2CH2OH</td>
<td>387.6</td>
<td>-7.5</td>
<td>3.0</td>
<td>-7.8</td>
</tr>
<tr>
<td>CH3-CH(OH)-CH3</td>
<td>385.6</td>
<td>-7.3</td>
<td>3.1</td>
<td>-7.7</td>
</tr>
<tr>
<td>HCOOH</td>
<td>351.2</td>
<td>-5.0</td>
<td>9.5</td>
<td>-4.0</td>
</tr>
<tr>
<td>CH3COOH</td>
<td>355.1</td>
<td>-4.3</td>
<td>8.2</td>
<td>-3.5</td>
</tr>
<tr>
<td>CH3CH2COOH</td>
<td>354.5</td>
<td>-4.2</td>
<td>6.9</td>
<td>-3.4</td>
</tr>
<tr>
<td>C6H5OH</td>
<td>356.7</td>
<td>-6.2</td>
<td>2.5</td>
<td>-5.0</td>
</tr>
<tr>
<td>p-CH3-C6H4OH</td>
<td>357.9</td>
<td>-6.2</td>
<td>1.8</td>
<td>-5.1</td>
</tr>
<tr>
<td>p-NO2-C6H4OH</td>
<td>334.6</td>
<td>-9.1</td>
<td>0.5</td>
<td>-7.2</td>
</tr>
<tr>
<td>H3O+</td>
<td>171.2</td>
<td>0.4</td>
<td>5.4</td>
<td>-0.7</td>
</tr>
<tr>
<td>2H2O(H+)</td>
<td>200.2</td>
<td>4.0</td>
<td>9.6</td>
<td>2.1</td>
</tr>
<tr>
<td>3H2O(H+)</td>
<td>213.4</td>
<td>3.0</td>
<td>5.5</td>
<td>1.2</td>
</tr>
<tr>
<td>4H2O(H+)</td>
<td>221.1</td>
<td>0.3</td>
<td>3.7</td>
<td>-0.3</td>
</tr>
<tr>
<td>5H2O(H+)</td>
<td>226.7</td>
<td>-0.2</td>
<td>5.9</td>
<td>-0.8</td>
</tr>
<tr>
<td>CH3OH2+</td>
<td>186.8</td>
<td>-1.0</td>
<td>2.2</td>
<td>-1.8</td>
</tr>
<tr>
<td>H2COOH+</td>
<td>177.1</td>
<td>-2.2</td>
<td>1.0</td>
<td>-1.8</td>
</tr>
<tr>
<td>CH3CHOH+</td>
<td>190.2</td>
<td>0.5</td>
<td>3.7</td>
<td>1.1</td>
</tr>
<tr>
<td>MUE</td>
<td>-4.0</td>
<td>5.9</td>
<td>4.3</td>
<td></td>
</tr>
<tr>
<td>MSE</td>
<td>9.1</td>
<td>31.4</td>
<td>8.5</td>
<td></td>
</tr>
<tr>
<td>MAX</td>
<td>4.7</td>
<td>5.9</td>
<td>4.3</td>
<td></td>
</tr>
</tbody>
</table>
Improving SCC-DFTB

\[ E = \sum_{i}^{\text{occ}} \langle \psi_i | \hat{H}^0 | \psi_i \rangle + \frac{1}{2} \iint' \left( \frac{1}{|\vec{r} - \vec{r}'|} + \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \bigg|_{n_0} \right) \Delta \rho \Delta \rho'. \]

- minimal basis
- neglect of crystal field and three-center terms
- initial density fixed

\[ E = \sum \sum c_i^\mu c_i^\nu \mathbb{H}_{\mu\nu}^0 + \frac{1}{2} \sum_{\alpha\beta} \gamma_{\alpha\beta} \Delta q_\alpha \Delta q_\beta + \sum_{\alpha\beta} U_{\alpha\beta} \]

- second order expansion
- monopole approximation
- gamma

- two-body approximation
- fit procedure
Determination of the repulsive potential

\[ E_{\text{rep}}(R_{\alpha/\beta}) = E^{\text{DFT}}(R_{\alpha/\beta}) - \left[ \sum_i \sum_{\mu} c_i^\mu c_i^\mu H^0_{\mu\mu} + E^2 \right] (R_{\alpha/\beta}) \]

original strategy:
- calculate \( E_{\text{dft}} \) and \( E_{\text{el}} \) for various C-C distances
- fit the difference to polynomial/splines
Determination of the repulsive potential

- Segments of $E_{\text{rep}}$ described by 4th order splines

- No fit to DFT bond-stretching curves

- Set up system of linear equations, which relate parameters to molecular energies and forces

- Input:
  - Atomization energies
  - Equilibrium structures (force = 0)
  - Can be experimental info

- Solve directly for $a_{k,i}$

M. Gaus et al, JCP accepted
Determination of the repulsive potential

- Segments of $E_{\text{rep}}$ described by 4th order splines

- No fit to DFT bond-stretching curves

- Set up system of linear equations, which relate parameters to molecular energies and forces

- Input:
  - Atomization energies
  - Equilibrium structures (force = 0)

- Solve directly for $a_{k,i}$

$\triangle E_t = E_{\text{tot}}^{\text{mol}} - \sum_{A} E_{\text{el}}^A + \sum_{A} \triangle H_f^A$

Note: $E_{\text{el}}^A$ can be fitted!

M. Gaus et al, JCP accepted
Application to hydrocarbons

Training set:
energy+ forces: hydrogen, methand, ethyne, ethene, ethane
only forces: cyclopropen, 2-butyne

‘adjustment set’: 15 molecules in total, energy, geometry, vib. frequency
adjust division points
Application to hydrocarbons: energy and geometry

<table>
<thead>
<tr>
<th>Property $^a$</th>
<th>N $^b$</th>
<th>opt1</th>
<th>opt2</th>
<th>opt3</th>
<th>opt4</th>
<th>mio</th>
<th>PBE$^c$</th>
<th>B3LYP$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta H_f$ (kcal/mol)</td>
<td>39</td>
<td>4.8</td>
<td>5.8</td>
<td>2.2</td>
<td>7.1</td>
<td>55.4$^d$</td>
<td>26.0</td>
<td>7.4</td>
</tr>
<tr>
<td>$\Delta H_f^{\text{max}}$ (kcal/mol)</td>
<td>19.2</td>
<td>29.0</td>
<td>16.3</td>
<td>25.2</td>
<td>114.8$^d$</td>
<td>81.3</td>
<td>17.3</td>
<td></td>
</tr>
<tr>
<td>$\Delta H_f$ (kcal/mol)</td>
<td>39</td>
<td>4.3</td>
<td>4.0</td>
<td>2.1</td>
<td>3.7</td>
<td>4.1</td>
<td>1.8</td>
<td>3.4</td>
</tr>
<tr>
<td>$\Delta H_f^{\text{max}}$ (kcal/mol)</td>
<td>19.7</td>
<td>25.1</td>
<td>16.2</td>
<td>20.5</td>
<td>21.7</td>
<td>7.4</td>
<td>9.6</td>
<td></td>
</tr>
<tr>
<td>$r$ (Å)</td>
<td>196</td>
<td>0.008</td>
<td>0.008</td>
<td>0.008</td>
<td>0.007</td>
<td>0.011</td>
<td>0.008</td>
<td>0.003</td>
</tr>
<tr>
<td>$r_{\text{max}}$ (Å)</td>
<td></td>
<td>0.155</td>
<td>0.267</td>
<td>0.038</td>
<td>0.142</td>
<td>0.035</td>
<td>0.020</td>
<td>0.016</td>
</tr>
<tr>
<td>$a$ (deg)</td>
<td>177</td>
<td>0.6</td>
<td>0.7</td>
<td>0.6</td>
<td>0.6</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>$a_{\text{max}}$ (deg)</td>
<td></td>
<td>14.4</td>
<td>19.9</td>
<td>8.2</td>
<td>13.4</td>
<td>4.8</td>
<td>1.9</td>
<td>1.9</td>
</tr>
<tr>
<td>$d$ (deg)</td>
<td>5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.3</td>
<td>2.1</td>
<td>1.5</td>
<td>0.7</td>
<td>1.0</td>
</tr>
<tr>
<td>$d_{\text{max}}$ (deg)</td>
<td></td>
<td>6.3</td>
<td>6.3</td>
<td>5.5</td>
<td>8.8</td>
<td>6.7</td>
<td>2.0</td>
<td>2.6</td>
</tr>
</tbody>
</table>

- ‘old’ set mio: no ground breaking improvement
- change of confinement radii lead to small variations in performance
  => DFTB quite robust with respect to changes in electronic parameters

G3/99
Application to hydrocarbons: vib. frequency

<table>
<thead>
<tr>
<th>Molecule</th>
<th>BLYP(^a)</th>
<th>opt1</th>
<th>opt2</th>
<th>opt3</th>
<th>opt4</th>
<th>opt5</th>
<th>opt6</th>
<th>mio</th>
<th>M2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methane</td>
<td>23</td>
<td>55</td>
<td>30</td>
<td>47</td>
<td>34</td>
<td>41</td>
<td>30</td>
<td>65</td>
<td>44</td>
</tr>
<tr>
<td>Ethyne</td>
<td>32</td>
<td>30</td>
<td>83</td>
<td>101</td>
<td>57</td>
<td>47</td>
<td>54</td>
<td>70</td>
<td>28</td>
</tr>
<tr>
<td>Ethene</td>
<td>19</td>
<td>40</td>
<td>60</td>
<td>50</td>
<td>39</td>
<td>36</td>
<td>36</td>
<td>49</td>
<td>34</td>
</tr>
<tr>
<td>Ethane</td>
<td>20</td>
<td>40</td>
<td>34</td>
<td>28</td>
<td>29</td>
<td>26</td>
<td>24</td>
<td>52</td>
<td>28</td>
</tr>
<tr>
<td>Allene</td>
<td>13</td>
<td>53</td>
<td>74</td>
<td>73</td>
<td>54</td>
<td>52</td>
<td>45</td>
<td>58</td>
<td>39</td>
</tr>
<tr>
<td>Cyclopropane</td>
<td>17</td>
<td>43</td>
<td>56</td>
<td>52</td>
<td>42</td>
<td>35</td>
<td>37</td>
<td>55</td>
<td>33</td>
</tr>
<tr>
<td>Propene</td>
<td>21</td>
<td>27</td>
<td>32</td>
<td>29</td>
<td>26</td>
<td>21</td>
<td>19</td>
<td>42</td>
<td>20</td>
</tr>
<tr>
<td>1,3-Butadiene</td>
<td>18</td>
<td>35</td>
<td>48</td>
<td>49</td>
<td>33</td>
<td>31</td>
<td>27</td>
<td>46</td>
<td>21</td>
</tr>
<tr>
<td>Bicyclo[2,1,0]pentane</td>
<td>50</td>
<td>40</td>
<td>46</td>
<td>45</td>
<td>40</td>
<td>41</td>
<td>34</td>
<td>52</td>
<td>42</td>
</tr>
<tr>
<td>Spiropentane</td>
<td>32</td>
<td>52</td>
<td>58</td>
<td>59</td>
<td>47</td>
<td>41</td>
<td>40</td>
<td>68</td>
<td>52</td>
</tr>
<tr>
<td>Cyclohexane</td>
<td>27</td>
<td>31</td>
<td>25</td>
<td>23</td>
<td>28</td>
<td>21</td>
<td>24</td>
<td>49</td>
<td>23</td>
</tr>
<tr>
<td>Benzene</td>
<td>17</td>
<td>53</td>
<td>67</td>
<td>70</td>
<td>46</td>
<td>42</td>
<td>34</td>
<td>58</td>
<td>34</td>
</tr>
<tr>
<td>1,3,5-Hexatriene</td>
<td>19</td>
<td>39</td>
<td>49</td>
<td>51</td>
<td>37</td>
<td>36</td>
<td>33</td>
<td>46</td>
<td>24</td>
</tr>
<tr>
<td>Cubane</td>
<td>22</td>
<td>52</td>
<td>47</td>
<td>63</td>
<td>52</td>
<td>64</td>
<td>43</td>
<td>96</td>
<td>35</td>
</tr>
<tr>
<td><strong>MAD (total)(^b)</strong></td>
<td>25</td>
<td>42</td>
<td>48</td>
<td>50</td>
<td>40</td>
<td>38</td>
<td>33</td>
<td>58</td>
<td>32</td>
</tr>
</tbody>
</table>

- ‘old’ set mio: reasonable improvement for vib. freq.
- optimization conflict => limited flexibility of DFTB model
  improvements not via ‘brute force fitting’ but new physics
Improving SCC-DFTB

\[ E = \sum_{i}^{\text{occ}} \langle \psi_i | \hat{H}^0 | \psi_i \rangle + \frac{1}{2} \int \int' \left( \frac{1}{| \vec{r} - \vec{r}' |} + \frac{\delta^2 E_{xc}}{\delta \rho \delta \rho'} \right)_{n_0} \Delta \rho \Delta \rho'. \]

\[ - \frac{1}{2} \int \int' \frac{\rho'_0 \rho_0}{| \vec{r} - \vec{r}' |} + E_{xc}[\rho_0] - \int V_{xc}[\rho_0] n_0 + E_{cc} \]

- minimal basis
- neglect of crystal field and three-center terms
- initial density fixed

\[ E = \sum_{i} \sum_{\mu \nu} c_i^\mu c_i^\nu \hat{H}_{\mu \nu}^0 + \frac{1}{2} \sum_{\alpha \beta} \gamma_{\alpha \beta} \Delta q_{\alpha} \Delta q_{\beta} + \sum_{\alpha \beta} U_{\alpha \beta} \]

- second order expansion
- monopole approximation
- gamma

- two-body approximation
- fit procedure
SCC-DFTB

- accuracy often close to DFT-PBE -- but optimization conflict

- confined minimal AO basis:
  slight underestimation of Pauli-repulsion
  but NO BSSE ==> no problem for large systems

- fixed initial density => partial change in third order

- contains all DFT-GGA failures

• minimal basis
• neglect of crystal field and three-center terms
• initial density fixed

• second order expansion
• monopole approximation
• gamma

• two-body approximation
• fit procedure

\[ E = \sum_i \sum_{\mu\nu} c_i^\mu c_i^\nu H_{\mu\nu}^0 + \frac{1}{2} \sum_{\alpha\beta} \gamma_{\alpha\beta} \Delta q_\alpha \Delta q_\beta + \sum_{\alpha\beta} U_{\alpha\beta} \]