Computational Material Science
Part I

Ito Chao (趙奕сте)
Institute of Chemistry
Academia Sinica
The self-consistent field method in Hartree-Fock theory moves an electron in an average potential of the other electrons. ∴ The instantaneous position of an electron is not influenced by a nearby electron. (Not correct!)

In fact, electrons avoid each other more than Hartree-Fock theory would suggest, giving rise to a lower energy. \[ E_{\text{correlation}} = E_{\text{exact}} - E_{\text{HF}} \]

**How to handle electron correlation**

- Use better wave functions?
  CI, MCSCF, GVB, CC
- Play tricks in the Hamiltonian?
  MP
- Empirical functional for \( E^{xc} \) (exchange and correlation energy)?
  DFT
Cannot do better than the HF wave function with a single determinant…

– Improve wave function by a linear combination of determinants

\[ \Psi = c_0 \Psi_{HF} + c_1 \Psi_1 + c_2 \Psi_2 + \cdots \]

Based on a single reference determinant

c reflect the weight of each determinant and ensure normalization

=> handle dynamical correlation
– Configuration Interaction (CI)

- Full configuration interaction (full CI): overall wavefunction is a linear combination of the ground and excited-state wavefunctions; consider all electrons including all orbitals... Full CI with an infinite basis set is an “exact” solution of the (non-relativistic, Born-Oppenheimer, time-independent) Schrödinger equation
  
  No reoptimization of HF orbitals is required as the CSF set is complete

\[ \Psi = a_0 \Psi_{\text{HF}} + \sum_{i} \sum_{r} a_i^r \Psi_i^r + \sum_{i<j} \sum_{r<s} a_{ij}^{rs} \Psi_{ij}^{rs} + \ldots \]

Determinant with single excitation

Determinant with double excitation

CSF: Configuration State Function

Full CI includes both dynamic and non-dynamic correlation.

Dynamic: individual value of \( a \) may be small, but many \( a \) contribute to dynamic correlation

Non-dynamic: limited number of \( a \) contribute, but the individual value is comparatively large

Problem: Full CI for methanol (CH\(_3\)OH) with 6-31G(d)
14 electrons in 38 orbitals (14, 38) => 2.4 x 10\(^{13}\) coefficients!

Computationally demanding…
• **Truncated configuration interaction (CI):** allow only a limited number of excitations…

Single-determinant reference (assume no non-dynamical correlation and no need to reoptimize the MOs):

\[
\Psi = a_0 \Psi_{HF} + \sum_i \sum_r a_i^r \Psi_i^r + \sum_i \sum_{i<j} \sum_{r<s} a_{ij}^{rs} \Psi_{ij}^{rs} + \cdots
\]

\text{determinant with single excitation} \quad \text{determinant with double excitation}

\( E \) (energy) of \( N \) different CI wave functions (corresponding to different variationally determined sets of coefficients) can be determined from The \( N \) roots of the CI secular equation:

\[
\begin{vmatrix}
H_{11} - E & H_{12} & \cdots & H_{1N} \\
H_{21} & H_{22} - E & \cdots & H_{2N} \\
\vdots & \vdots & \ddots & \vdots \\
H_{N1} & N_{N2} & \cdots & H_{NN} - E
\end{vmatrix} = 0
\]

(7.11)

\[
H_{mn} = \langle \Psi_m | H | \Psi_n \rangle
\]
Not all excitations contribute to energy lowering…

Ex: Brillouin’s theorem -- single excitations do not mix with the ground state.

*Determinants of single excitation state and ground state differ only by one column*

**Ground state of H₂: 1σ₁^2**  
Single excited: 1σ₁^11σ₂^1

\[
\begin{vmatrix}
1σ_g(1)α(1) & 1σ_g(1)β(1) \\
1σ_g(2)α(2) & 1σ_g(2)β(2)
\end{vmatrix} \quad \begin{vmatrix}
1σ_g(1)α(1) & 1σ_u(1)α(1) \\
1σ_g(2)α(2) & 1σ_u(2)α(2)
\end{vmatrix}
\]

Characteristics of determinants:

\[
\begin{vmatrix}
a & c \\
b & d
\end{vmatrix} + \begin{vmatrix}
a & e \\
b & f
\end{vmatrix} = \begin{vmatrix}
a & c + e \\
b & d + f
\end{vmatrix}
\]

Therefore, although \(Ψ = a_0 \text{Determinant}_{\text{ground}} + a_i \text{Determinant}_{\text{single excitation}}\), \(Ψ\) is still effectively a single determinant wavefunction.
Figure 7.4  Structure of the CI matrix as blocked by classes of determinants. The HF block is the (1,1) position, the matrix elements between the HF and singly excited determinants are zero by Brillouin’s theorem, and between the HF and triply excited determinants are zero by the Condon–Slater rules. In a system of reasonable size, remaining regions of the matrix become increasingly sparse, but the number of determinants in each block grows to be extremely large. Thus, the (1,1) eigenvalue is most affected by the doubles, then by the singles, then by the triples, etc

<table>
<thead>
<tr>
<th>$\psi_{HF}$</th>
<th>$\psi^a$</th>
<th>$\psi^{ab}$</th>
<th>$\psi^{abc}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_{HF}$</td>
<td>0</td>
<td>dense</td>
<td>0</td>
</tr>
<tr>
<td>$\psi^a$</td>
<td>0</td>
<td>dense</td>
<td>sparse</td>
</tr>
<tr>
<td>$\psi^{ab}$</td>
<td>sparse</td>
<td>sparse</td>
<td>extremely sparse</td>
</tr>
<tr>
<td>$\psi^{abc}$</td>
<td>0</td>
<td>very sparse</td>
<td>extremely sparse</td>
</tr>
</tbody>
</table>

Not all matrix elements $H_{mn}$ has value!

$\Rightarrow$ Some excitations are more important

$\Rightarrow$ **CID** (only include double excitations)

$\Rightarrow$ **CISD** (single excitations interact with double excitations, but not ground state)
CID

For $H_2$, $\sigma^2$ to $\sigma^*^2$ (determinants based on ground state and doubly excited state are used)

$$\begin{vmatrix}
H_{11} - E & H_{12} \\
H_{21} & H_{22} - E
\end{vmatrix} = 0$$

$H_{12}$: Basically electron-repulsion

$E = \frac{1}{2} \left[ H_{11} + H_{22} \pm \sqrt{(H_{22} - H_{11})^2 + 4H_{12}} \right]$  

\[\text{Doubly excited state energy} \rightarrow \text{positive}\]

\[\text{ground state energy}\]

$\Rightarrow E^-$ lower than the HF energy, the difference is correlation energy

(STO-3G; bond distance of 1.4 a.u. for $H_2$;

$E_{\text{correlation}} = -0.02056$ a.u. = 13 kcal/mol)
The non-size-consistent problem of truncated CI

Ex: CID

For two isolated $\text{H}_2$

$2\text{e}^-$

$2 \times \text{energy of } \text{H}_2 \text{ will have the contribution from quadruple excitation!}$

For $\text{H}_2$ dimer at long separation

$2\text{e}^-$

$2 \times E_A \not< E_{A2}$

The energies from the above two systems are not comparable!
However, for molecules such as trimethylenemethane (TMM)

The singlet TMM can be

$$\Psi_{\text{RHF}} = | \cdots \pi_1^2 \pi_2^2 \pi_3^0 \rangle$$

$$\Psi'_{\text{RHF}} = | \cdots \pi_1^2 \pi_0^0 \pi_2^2 \rangle$$

Different reference determinant needed!

$$\Psi_{\text{MCSCF}} = a_1 | \cdots \pi_1^2 \pi_2^2 \rangle + a_2 | \cdots \pi_1^2 \pi_3^2 \rangle$$

$$\Rightarrow$$ handle non-dynamical correlation

**MCSCF**: Multiconfiguration Self-Consistent Field Theory

*Basis function coefficients and determinant coefficients both optimized!*

Occupation number of orbitals are described as:

$$\text{(occ. no.)}_{i,\text{MCSCF}} = \sum_n (\text{occ. no.})_{i,\pi} \alpha_n^2$$
For a more complete consideration, the singlet TMM can be

\[ \Psi_{MCSCF} = a_1|\cdots\pi_1^0\pi_2^0\pi_3^0\pi_4^0\rangle + a_2|\cdots\pi_1^2\pi_2^0\pi_3^0\pi_4^0\rangle + a_3|\cdots\pi_1^0\pi_2^0\pi_3^2\pi_4^0\rangle 
+ a_4|\cdots\pi_1^0\pi_2^2\pi_3^0\pi_4^0\rangle + a_5 (|\cdots\pi_1^0\pi_2^1\pi_3^1\pi_4^0\rangle + |\cdots\pi_1^2\pi_2^1\pi_3^1\pi_4^0\rangle) \]

The number of singlet configuration state function (CSF) can be formed from the distribution of \( m \) electrons in \( n \) orbitals is:

\[ N = \frac{n!(n+1)!}{\binom{m}{2}!(\frac{m}{2}+1)!\left(n - \frac{m}{2}\right)!\left(n - \frac{m}{2} + 1\right)!} \]

(14,12) for 14 electrons in 12 orbitals: \( N = 169884 \)!
Figure 7.3 Possible assignment of different orbitals in a completely general MCSCF formalism. Frozen orbitals are not permitted to relax from their HF shapes, in addition to having their occupation numbers of zero (virtual) or two (occupied) enforced.
• **Schemes to reduce the number of CSFs**

  • Choose according to symmetry

  • **GVB** (general valence bond): localized orbital; electrons only excites from a bonding orbital to its antibonding orbital

  • **RAS** (restricted active space): allow a limited number of excitations from/to orbitals outside of the CAS space

  • Freeze the shapes of the core orbitals
• **Multireference configuration interaction (MRCI):** MCSCF wave function is used instead of the HF wave function

- Enormous number of matrix elements, only suitable for small systems

- Both dynamical and non-dynamical correlation energy considered

- Good for study a large section of a PES, where significant changes in bonding (and thus correlation energy) are taking place

- MRCISD with large basis sets can be better than full CI with small basis set, illustrating most of the correlation energy can be captured by including limited excitations
– Coupled-cluster theory (CC)

Improve wavefunction in the following fashion

\[ \Psi = e^T \psi_{HF} \]

\[ T = T_1 + T_2 + T_3 + \cdots + T_n \]

\( T_n \) operators generate all possible determinants having \( n \) excitations from the reference. For example,

\[ T_2 \psi_{HF} = \sum_{i<j} \sum_{a<b} t_{ij}^{ab} \psi_{ij}^{ab} \]

• **Truncated CC**: if only double excitation operator considered (CCD)

\[ \psi_{CCD} = e^T \psi_{HF} \]

\[ = \left( 1 + T_2 + \frac{T_2^2}{2!} + \frac{T_2^3}{3!} + \cdots \right) \psi_{HF} \]

Contrary to CID, CCD does not suffer from non-size-consistent!
• Often-used truncated CC: **CCSD, CCSD(T)** [(T) means singles/triples coupling term considered]

• Coupled-cluster theory is not variational

• Single-determinant based; when the **$T_1$ diagnostic of Lee and Taylor** is larger than 0.02, avoid using CCSD and CCSD(T)

\[
T_1 = \sqrt{\frac{\sum_{i} \sum_{a} \rho_{i}^{a}}{n}}
\]

A measure of multireference character
– Rayleigh-Schrödinger Perturbation Theory

Rewrite operator $A$ as

$$A = A^{(0)} + \lambda V$$

*perturbing operator*

*dimensionless parameter (0 to 1)*

*an operator whose eigenfunction can be found*

Expand ground-state eigenfunctions and eigenvalues as Taylor series in $\lambda$

$$\psi_0 = \psi_0^{(0)} + \lambda \frac{\partial \psi_0^{(0)}}{\partial \lambda} \bigg|_{\lambda=0} + \frac{1}{2!} \lambda^2 \frac{\partial^2 \psi_0^{(0)}}{\partial \lambda^2} \bigg|_{\lambda=0} + \frac{1}{3!} \lambda^3 \frac{\partial^3 \psi_0^{(0)}}{\partial \lambda^3} \bigg|_{\lambda=0} + \cdots$$

$$a_0 = a_0^{(0)} + \lambda \frac{\partial a_0^{(0)}}{\partial \lambda} \bigg|_{\lambda=0} + \frac{1}{2!} \lambda^2 \frac{\partial^2 a_0^{(0)}}{\partial \lambda^2} \bigg|_{\lambda=0} + \frac{1}{3!} \lambda^3 \frac{\partial^3 a_0^{(0)}}{\partial \lambda^3} \bigg|_{\lambda=0} + \cdots$$

$\psi_0^{(0)}$: normalized eigenfunction for $A^{(0)}$

$a_0^{(0)}$: eigenvalue for $\psi_0^{(0)}$
For ease of notation, rewrite the taylor series as

\[ \Psi_0 = \Psi_0^{(0)} + \lambda \Psi_0^{(1)} + \lambda^2 \Psi_0^{(2)} + \lambda^3 \Psi_0^{(3)} + \cdots \]

\[ a_0 = a_0^{(0)} + \lambda a_0^{(1)} + \lambda^2 a_0^{(2)} + \lambda^3 a_0^{(3)} + \cdots \]

Superscripts \( n \) are referred to as ‘nth-order corrections’ to the zeroth order term

Subscripts specify which eigenfunction is being considered

We may write

\[(A^{(0)} + \lambda V) |\Psi_0\rangle = a_0 |\Psi_0\rangle \]

as

\[ (A^{(0)} + \lambda V) |\Psi_0^{(0)}\rangle + \lambda |\Psi_0^{(1)}\rangle + \lambda^2 |\Psi_0^{(2)}\rangle + \lambda^3 |\Psi_0^{(3)}\rangle + \cdots = \]

\[ (a_0^{(0)} + \lambda a_0^{(1)} + \lambda^2 a_0^{(2)} + \lambda^3 a_0^{(3)} + \cdots) |\Psi_0^{(0)}\rangle + \lambda |\Psi_0^{(1)}\rangle + \lambda^2 |\Psi_0^{(2)}\rangle + \lambda^3 |\Psi_0^{(3)}\rangle + \cdots \]

For powers 0 ~ 3 of \( \lambda \)

\[ A^{(0)} |\Psi_0^{(0)}\rangle = a_0^{(0)} |\Psi_0^{(0)}\rangle \]

\[ A^{(0)} |\Psi_0^{(1)}\rangle + V |\Psi_0^{(0)}\rangle = a_0^{(0)} |\Psi_0^{(1)}\rangle + a_0^{(1)} |\Psi_0^{(0)}\rangle \]

\[ A^{(0)} |\Psi_0^{(2)}\rangle + V |\Psi_0^{(1)}\rangle = a_0^{(0)} |\Psi_0^{(2)}\rangle + a_0^{(1)} |\Psi_0^{(1)}\rangle + a_0^{(2)} |\Psi_0^{(0)}\rangle \]

\[ A^{(0)} |\Psi_0^{(3)}\rangle + V |\Psi_0^{(2)}\rangle = a_0^{(0)} |\Psi_0^{(3)}\rangle + a_0^{(1)} |\Psi_0^{(2)}\rangle + a_0^{(2)} |\Psi_0^{(1)}\rangle + a_0^{(3)} |\Psi_0^{(0)}\rangle \]
From above equations and some basic rules, one can obtain \( n \)-th order corrections of eigenvalues (Chap. 7.4 of Essentials of Computational Chemistry)

\[
a^{(1)}_0 = \langle \Psi_0^{(0)} | V | \Psi_0^{(0)} \rangle
\]

Higher order wavefunction \( \Psi_0^{(1)} \) can be expressed as a linear combination of the complete set of eigenfunctions of the zeroth-order Hamiltonian \( A^{(0)} \)

\[
\Psi_0^{(1)} = \sum_{i > 0} c_i \Psi_i^{(0)} \quad \text{ where } c_j = \frac{\langle \Psi_j^{(0)} | V | \Psi_0^{(0)} \rangle}{a^{(0)}_j - a^{(0)}_i}
\]

Similarly

\[
a^{(2)}_0 = \sum_{j > 0} \frac{|\langle \Psi_j^{(0)} | V | \Psi_0^{(0)} \rangle|^2}{a^{(0)}_j - a^{(0)}_i}
\]

\[
a^{(3)}_0 = \sum_{j,k > 0} \frac{\langle \Psi_0^{(0)} | V | \Psi_j^{(0)} \rangle \langle \Psi_j^{(0)} | V | \Psi_k^{(0)} \rangle - \delta_{jk} \langle \Psi_0^{(0)} | V | \Psi_0^{(0)} \rangle \langle \Psi_k^{(0)} | V | \Psi_0^{(0)} \rangle}{(a^{(0)}_j - a^{(0)}_i)(a^{(0)}_j - a^{(0)}_k)}
\]

Our focus is to get correlation energy, notice that it can be obtained if we have \( V \), eigenvalues and eigenfunctions of \( A^{(0)} \)

Note: Not variational!
• Single-reference perturbation theory by Møller and Plesset (MPn):
  
  \( n \) is the order at which the perturbation theory is truncated
  (acronym: MBPTn)

Take \( A^{(0)} (=H^{(0)}) \) as sum of the one-electron Fock operator

\[
H^{(0)} = \sum_{i=1}^{n} f_i \\
H^{(0)}\Psi^{(0)} = \sum_{i}^{\text{occ.}} \varepsilon_i \Psi^{(0)}
\]

Error in the above equation: each orbital energy includes the repulsion of the occupying electron(s) with all other electrons --- double counting of e-repulsion. Use correction term \( V \) to correct

\[
V = \sum_{i}^{\text{occ.}} \sum_{j>i}^{\text{occ.}} \frac{1}{r_{ij}} - \sum_{i}^{\text{occ.}} \sum_{j}^{\text{occ.}} \left( J_{ij} - \frac{1}{2} K_{ij} \right)
\]

Here means Coulomb and exchange operators, not Coulomb and exchange integrals.
As \( a_0^{(1)} = \langle \Psi_0^{(0)} | V | \Psi_0^{(0)} \rangle \)

\[
a_0^{(0)} + a_0^{(1)} = \langle \Psi^{(0)} | H^{(0)} | \Psi^{(0)} \rangle + \langle \Psi^{(0)} | V | \Psi^{(0)} \rangle
\]

\[
 = \langle \Psi^{(0)} | H^{(0)} + V | \Psi^{(0)} \rangle
\]

\[
 = \langle \Psi^{(0)} | H | \Psi^{(0)} \rangle
\]

\[
 = E_{\text{HF}} \quad \Rightarrow \text{MP1 (correction to the first order) gives HF energy}
\]

\[
a_0^{(2)} = \sum_i \sum_{j>i} \sum_a \sum_{b>a} \left[ (ia|jb) - (ib|ja) \right]^2 \frac{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}{\varepsilon_i - \varepsilon_a - \varepsilon_b}
\]

\[
a_0^{(0)} + a_0^{(1)} + a_0^{(2)} = E_{\text{MP2}}
\]

\text{MP3} \quad \text{empirical evidence shows rather little improvement over MP2…}

\text{MP4} \quad \text{MP4SDQ (triply excited states ignored; good for species with a large gap in frontier orbitals). With a good basis set, 95% of correlation energy recovered.}
• **Multireference perturbation**

- Using MCSCF wave function instead of a single-determinant RHF or UHF (CASPT2 by Roos.)

- Geometry optimization less straightforward, as analytic gradients not available

- Address dynamical correlation after a separate treatment of non-dynamical correlation
- Performance summary

\[
\begin{align*}
\text{HF} & < \text{MP2} \sim \text{MP3} \sim \text{CCD} < \text{CISD} \\
& < \text{MP4SDQ} \sim \text{QCISD} \sim \text{CCSD} < \text{MP4} \\
& < \text{QCISD(T)} \sim \text{CCSD(T)} \sim \text{BD(T)}
\end{align*}
\]

Table 7.4 Average errors in correlation energies (kcal mol\(^{-1}\)) compared to full CI for various methods applied to HB, H\(_2\)O, and HF at both equilibrium and bond-stretched geometries

<table>
<thead>
<tr>
<th>Level of theory</th>
<th>Equilibrium geometry</th>
<th>Equilibrium and stretched geometries</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP2</td>
<td>10.4</td>
<td>17.4</td>
</tr>
<tr>
<td>MP3</td>
<td>5.0</td>
<td>14.4</td>
</tr>
<tr>
<td>CISD</td>
<td>5.8</td>
<td>13.8</td>
</tr>
<tr>
<td>CCD</td>
<td>2.4</td>
<td>8.0</td>
</tr>
<tr>
<td>MP4SDQ</td>
<td>2.7</td>
<td>7.1</td>
</tr>
<tr>
<td>CCSD</td>
<td>1.9</td>
<td>4.5</td>
</tr>
<tr>
<td>QCISD</td>
<td>1.7</td>
<td>4.0</td>
</tr>
<tr>
<td>MP4</td>
<td>1.3</td>
<td>3.7</td>
</tr>
<tr>
<td>MP5</td>
<td>0.8</td>
<td>3.2</td>
</tr>
<tr>
<td>MP6</td>
<td>0.3</td>
<td>0.9</td>
</tr>
<tr>
<td>CCSD(T)</td>
<td>0.3</td>
<td>0.6</td>
</tr>
<tr>
<td>QCISD(T)</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>CCSD(T)</td>
<td>0.2</td>
<td>0.5</td>
</tr>
<tr>
<td>CCSDTQ</td>
<td>0.01</td>
<td>0.02</td>
</tr>
</tbody>
</table>

QCISD: an variant of CISD to correct the non-size-consistent problem
BD: an variant of CCSD to reduce the single reference problem

4/19/2007 CMS I
There are methods to reduce the scaling...

Table 7.5 Formal scaling behavior, as a function of basis functions $N$, of various electronic structure methods

<table>
<thead>
<tr>
<th>Scaling behavior</th>
<th>Method(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N^4$</td>
<td>HF</td>
</tr>
<tr>
<td>$N^5$</td>
<td>MP2</td>
</tr>
<tr>
<td>$N^6$</td>
<td>MP3, CISD, MP4SDQ, CCSD, QCISD</td>
</tr>
<tr>
<td>$N^7$</td>
<td>MP4, CCSD(T), QCISD(T)</td>
</tr>
<tr>
<td>$N^8$</td>
<td>MP5, CISDT, CCSDT</td>
</tr>
<tr>
<td>$N^9$</td>
<td>MP6</td>
</tr>
<tr>
<td>$N^{10}$</td>
<td>MP7, CISDTQ, CCSDTQ</td>
</tr>
</tbody>
</table>

Non-local DFT
– Overall impression

Dynamic correlation: single reference approach

- truncated CI
- QCI
- MBPT, MPn
- CC

Nondynamic correlation (static correlation): multireference

- MCSCF
- CASSCF
- GVB

When you see MR-CISD, GVB-CISD, CASPT2, MR-MBPT, MR-CC....
=> Trying to consider both nondynamic and dynamic correlations
– Parameterized methods for correlation energy
  • Spin-component scaled MP2 (SCS-MP2 by Grimme)

Different scaling of the opposite-spin ($E_{OS}$) and the same-spin ($E_{SS}$) electron pair contributions to the correlation energy

$$E_c = p_{OS}E_{OS}^{MP2} + p_{SS}E_{SS}^{MP2}$$

$p_{OS}$ and $p_{SS}$ are scaling factors of 6/5 and 1/3

...In Hartree-Fock, the same-spin electron pairs are correlated (Fermi holes), while the opposite-spin pairs are uncorrelated. Therefore, $p_{OS} > p_{SS}$

**TABLE 2: Deviations of Calculated Heats of Formation $\Delta H_f^0$ (298 K) (in kcal/mol) from Experiment$^a$ for the G2/97' Neutral Test Set (160 compounds)**

<table>
<thead>
<tr>
<th></th>
<th>SCS-MP2</th>
<th>SOS-MP2</th>
<th>MP2</th>
<th>B3LYP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean deviation</td>
<td>-0.14</td>
<td>0.07</td>
<td>-0.38</td>
<td>0.33</td>
</tr>
<tr>
<td>Mean absolute deviation</td>
<td>1.18</td>
<td>1.36</td>
<td>1.74</td>
<td>2.12</td>
</tr>
<tr>
<td>Maximum deviation</td>
<td>4.9</td>
<td>7.4</td>
<td>7.3</td>
<td>14.5</td>
</tr>
<tr>
<td>&lt;2 kcal/mol$^d$</td>
<td>76%</td>
<td>76%</td>
<td>66%</td>
<td>58%</td>
</tr>
<tr>
<td>&lt;3 kcal/mol$^d$</td>
<td>90%</td>
<td>87%</td>
<td>80%</td>
<td>72%</td>
</tr>
</tbody>
</table>

Deviation = experiment-theory.

SOS-MP2: only consider opposite-spin; $p_{OS} = 1.3$

• SCS-MP2 reaction energies can be as good as QCISD(T)

• Good for compounds with main group elements

• Does not work for spin-contaminated case


**Multilevel Methods (G2 and G3 by Pople)**

Table 7.6  Steps in G2 and G3 theory for molecules$^{a,b}$

<table>
<thead>
<tr>
<th>Step</th>
<th>G2</th>
<th>G3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>HF/6-31G(d) geometry optimization</td>
<td>HF/6-31G(d) geometry optimization</td>
</tr>
<tr>
<td>(2)</td>
<td>ZPVE from HF/6-31G(d) frequencies</td>
<td>ZPVE from HF/6-31G(d) frequencies</td>
</tr>
<tr>
<td>(3)</td>
<td>MP2(full)/6-31G(d) geometry optimization (all subsequent calculations use this geometry)</td>
<td>MP2(full)/6-31G(d) geometry optimization (all subsequent calculations use this geometry)</td>
</tr>
<tr>
<td>(4)</td>
<td>$E[\text{MP4/6-311+G(d,p)}] - E[\text{MP4/6-311G(d,p)}]$</td>
<td>$E[\text{MP4/6-311G(2df,p)}] - E[\text{MP4/6-311G(2df,p)}]$</td>
</tr>
<tr>
<td></td>
<td>$E[\text{MP4/6-31G(2df,p)}]$ - $E[\text{MP4/6-31G(d,p)}]$</td>
<td>(1) + (2) for ZPVE</td>
</tr>
<tr>
<td>(5)</td>
<td>$E[\text{QCISD(T)/6-311G(d)}]$ - $E[\text{MP4/6-311G(d)}]$</td>
<td>(3) for getting good geometry</td>
</tr>
<tr>
<td>(6)</td>
<td>$E[\text{MP2/6-311+G(3df,2p)}] - E[\text{MP2/6-311G(2df,p)}]$</td>
<td>(4) For estimating the effect of +</td>
</tr>
<tr>
<td></td>
<td>$-E[\text{MP2/6-311G(2df,p)}] - E[\text{MP2/6-311G(d,p)}] + E[\text{MP2/6-311G(d,p)}]$</td>
<td>(5) For the effect of d, f, p</td>
</tr>
<tr>
<td>(7)</td>
<td>$-0.00481 \times \text{(number of valence electron pairs)} - 0.00019 \times \text{(number of unpaired valence electrons)}$</td>
<td>(6) For the effect of better level</td>
</tr>
<tr>
<td></td>
<td>$-0.006386 \times \text{(number of valence electron pairs)} - 0.002977 \times \text{(number of unpaired valence electrons)}$</td>
<td>(7) For the effect of larger basis set</td>
</tr>
<tr>
<td></td>
<td>$E_0 = 0.8929 \times (2) + E[\text{MP4/6-311G(d,p)}] + (4) + (5) + (6) + (7) + (8)$</td>
<td>(8) For the effect of number of electrons</td>
</tr>
</tbody>
</table>

For atoms, G3 energies are defined to include a spin-orbit correction taken either from experiment or other high-level calculations. In addition, different coefficients are used in step (8).

In the G2 method, the 6-311G basis set and its derivatives are not defined for second-row atoms; instead, a basis set optimized by McLean and Chandler (1980) is used.

Available at [http://chemistry.annl.gov/compmet/g3theory.htm](http://chemistry.annl.gov/compmet/g3theory.htm). Defined to use canonical 5 d and 7 f functions.
• For a test set of 148 heats of formation,
  
  average error of G2 theory is 1.6 kcal mol\(^{-1}\)
  
  G3 theory is 0.9 kcal mol\(^{-1}\)

• G3 twice as fast as G2

• Many variants exist, but share the same spirit (using lower level calculations to approximate higher level results):
  
  G3(MP2), G3B3, CBS-4, CBS-q, CBS-Q, CBS-APNO, W1, W2, G3(MCG3), BAC-MP4…
How about DFT?

In general better than HF, often similar to MP2 and higher levels (sometimes better, sometimes worse...)

Table 8.3 Mean and maximum absolute errors (kcal mol\(^{-1}\)) in enthalpies of activation and forward reaction for different methods

<table>
<thead>
<tr>
<th>Level of theory</th>
<th>Activation</th>
<th>Reaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Maximum</td>
</tr>
<tr>
<td><strong>Reaction set 3(^{b,c})</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HF</td>
<td>18.7</td>
<td>26.7</td>
</tr>
<tr>
<td>CASSCF</td>
<td>16.0</td>
<td>34.6</td>
</tr>
<tr>
<td>MP2</td>
<td>4.6</td>
<td>7.6</td>
</tr>
<tr>
<td>CASPT2//CASSCF</td>
<td>2.4</td>
<td>5.7</td>
</tr>
<tr>
<td>CBS-QB3</td>
<td>1.9</td>
<td>4.3</td>
</tr>
<tr>
<td>BPW91</td>
<td>3.7</td>
<td>6.9</td>
</tr>
<tr>
<td>KMLYP</td>
<td>3.2</td>
<td>10.3</td>
</tr>
<tr>
<td>OLYP</td>
<td>3.4</td>
<td>9.0</td>
</tr>
<tr>
<td>OLYP/6-311+G(2d,p)</td>
<td>4.4</td>
<td>13.0</td>
</tr>
<tr>
<td>MPW1K/6-31+G(d,p)</td>
<td>2.2</td>
<td>6.9</td>
</tr>
<tr>
<td>B3LYP</td>
<td>1.7</td>
<td>6.0</td>
</tr>
<tr>
<td>B3LYP/6-31+G(d,p)</td>
<td>2.4</td>
<td>8.1</td>
</tr>
<tr>
<td>B3LYP/6-311+G(2d,p)</td>
<td>2.9</td>
<td>10.1</td>
</tr>
<tr>
<td>O3LYP/OLYP</td>
<td>3.0</td>
<td>9.0</td>
</tr>
</tbody>
</table>

\(^{a}\)See Baker, Muir, and Andzelm (1995).

\(^{b}\)6-31G(d) basis set unless otherwise indicated.

\(^{c}\)Using five spherical d functions instead of the usual six Cartesian functions implied by this basis set name.

\(^{d}\)See Lynch and Truhlar (2003a) and Zhao et al. (2004); 6-31+G(d,p) basis set; the Reaction column refers to the atomization enthalpies for six molecules chosen to be representative of a larger set in a fashion analogous to the H-atom transfer reactions, namely, SiO, S, silane, propyne, glyoxal, and cyclobutane.

\(^{e}\)See Guner et al. (2003, 2004).

For full table see Essentials of Computational Chemistry
Hybrid Methods

High level MO methods afford accurate geometries, energies, vibrational frequencies…. How to extend this accuracy to larger systems?

-- Treat important part at a high level of theory and less important part at a lower level?

Molecular mechanics (MM)
popular tool for treating large systems
most valuable when steric or electrostatic interactions are dominant
no good for electronic properties, bond-breaking/forming

=> many QM/MM hybrid approaches in the literature
Effective Hamiltonian built and additional empirical parameters developed to allow the combined energy to reproduce experiments

\[ H_{\text{eff}} = H_{\text{QM}} + H_{\text{MM}} + H_{\text{QM-MM}} \]
Oniom (Our own \(N\)-layered integrated molecular orbital and molecular mechanics method) by Morokuma

- Conceptually different from common QM/MM methods; similar to G2, G3 methods in spirit
- Also known as IMOMM (integrated MO-MM method), IMOMO (integrated MO-MO method), ONIOM\(n\), \(n\) specify number of layers.
  
  \(\Rightarrow\) IMOMM equals to ONIOM2(MO:MM)
  
  \(\Rightarrow\) IMOMO equals to ONIOM2(MO:MO)

Ideally, one should calculate the real system with high theory level. However, one may face the situation that the high theory level may only be able to handle the model system, while only the low theory level can handle the real system....

**Approximate A, B, C, D with ~A, ~B, ~C, ~D**
5 (= 2n – 1) calculations required when n = 3

Ex: ONIOM3(CCSD(T):MP2:HF)
ONIOM3(HF:PM3:MM3)
In a two-layer calculation:

- Cut bond are replaced with hydrogen atom (Set 2)
- Geometry optimization based on E(ONIOM2)
- If no covalent bond is cut, no linked atom needed e.g., (H₂O)₂

Note:
Geometry is from the real system
Ex: $H_2$ addition to $Pt(P(t-Bu)_3)_2$

Use B3LYP:B3LYP:B3LYP as benchmark
- When B3LYP was used for small model, large improvement on $E_a$.
- When B3LYP used for small and intermediate model, further improvement
- Note the time difference
- When MM3 used for intermediate model, no electronic effect of methyl group considered, so larger error
In an ensemble of $N$ molecules, the instantaneous value of a property, $A$ (e.g. pressure or heat capacity), depends on the positions ($r$) and momenta ($p$) of molecules.

\[ A(p_{1x}, p_{1y}, p_{1z}, p_{2x}, p_{2y}, p_{2z}, \ldots, x_1, y_1, z_1, \ldots, t) = A(p^N(t), r^N(t)) \]

Each combination of $3N$ positions and $3N$ momenta defines a point in the $6N$-dimensional phase space.
Property $A$ fluctuates with time as the molecules move. The measured value of $A$ is the average of the instantaneous values over the time of the measurement. If the time over which the measurement is made increases to infinity, the following integral approaches the ‘true’ average value of the property:

$$A_{\text{average}} = \lim_{T \to \infty} \frac{1}{T} \int_{t=0}^{T} A(\vec{p}^N(t), \vec{r}^N(t))dt$$

Boltzmann and Gibbs replaced the time average of $A$ by an ensemble average: simultaneously consider lots of replications of the system

$$\langle A \rangle = \int \cdots \int A(\vec{p}^N, \vec{r}^N)\rho(\vec{p}^N, \vec{r}^N)d\vec{p}^N d\vec{r}^N$$

probability density of an ensemble

The so-called *ergodic hypothesis*: ensemble average is equal to time average
• Probability density of an ensemble: the probability of finding a configuration with momenta $p^N$ and position $r^N$

Example:
Under the condition of NVT (constant number of particles, volume, temperature), probability is determined by potential energy of the ensemble

$$\rho(\vec{r}^N) = \frac{1}{Z} \exp \left[ -\frac{V(\vec{r}^N)}{k_B T} \right] = \underbrace{\text{Boltzmann factor}}_{\text{vs sum of Boltzmann factors of all ensembles}}$$

$$Z = \int \exp \left( -\frac{V(\vec{r}^N)}{k_B T} \right) d\vec{r}^N$$

Configurational integral $Z$: integration of boltzmann factor for all ensembles
Monte Carlo Simulation (MC)

Use random movements of the particles to generate ensembles and rely on algorithms to sample ensembles according to the Boltzmann distribution (**importance sampling** replace the probability density)

\[
<A> \approx \frac{1}{M} \sum_{i=1}^{M} A(r^N)
\]

– Metropolis Monte Carlo calculations

1. Generate a low energy initial configuration of the ensemble
2. Generate a random move of a particle (or a molecule) and calculate the energy
3. If \( V_{\text{new}} < V_{\text{old}} \) => accept the new configuration, return to step 2
4. If \( V_{\text{new}} > V_{\text{old}} \) => calculate the Boltzmann factor \( \exp\left[-(V_{\text{new}} - V_{\text{old}})/kT\right] \), accept the configuration only if Boltzmann factor > \( \text{rand}(0,1) \), return to step 2
5. After a given number of trials has been performed, calculate the ensemble average based on the above equation

(Note: information on p.244 of the hand-out is incorrect)
• If energy goes up significantly in the new configuration, the value of \( \exp[-(V_{\text{new}}-V_{\text{old}})/kT] \) will be close to 0 and is highly probable to be smaller than the random number between 0,1. Therefore, this new configuration is likely to be rejected.

• Normally the rejection rate is adjusted to 60 to 50% (if nothing is rejected, there is no importance sampling).

• Particles are move by a random amount limited to a maximum value. Adjust the maximum value will have an effect on the rejection rate.

• Basically, millions of configurations will be generated in the equilibration phase and millions will be generated in the averaging phase.
Molecular Dynamics Simulation (MD)

Configurations generated by moving particles according to Newton’s second law.

\[ \vec{f}_i = m_i \vec{a}_i \]

\[ \vec{a}_i = \frac{\vec{f}_i}{m_i} = \frac{d\vec{v}_i}{dt} = \frac{d^2\vec{r}_i}{dt^2} \]

With initial positions assigned, one can calculate potential energy, force, and acceleration. With initial velocities assigned according to Maxwell-Boltzmann distribution for a given temperature T, one can predict new positions and new velocities.

Example: velocity Verlet algorithm

\[ \vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}(t)\Delta t + \frac{\vec{a}(t)}{2}(\Delta t)^2 \]

\[ \vec{v}(t + \Delta t) = \vec{v}(t) + \frac{1}{2}[\vec{a}(t) + \vec{a}(t + \Delta t)]\Delta t \]
• Time step: on the order of femto second (fs; $10^{-15}$ s)
  Force is assumed to be a constant within $\Delta t$
  $\Delta t$ too small: trajectory covers only limited proportion of the phase space
  too large: may cause instability

• Typical simulation time: on the order of pico and nano seconds

• Thermodynamic averages from molecular dynamics:
  \[
  \langle A \rangle \approx \frac{1}{M} \sum_{i=1}^{M} A(p^N, r^N)
  \]

• MD Trajectory: record time evolution of $r$, $v$, $E$ (position, velocity, energy)

• Deterministic: the current status is determined by the previous status
Differences between MD and MC Simulations

- **MD**
  - Deterministic
    - may analyze time-dependent quantities (e.g. diffusion constant)
  - Total energy contains kinetic energy and potential energy
    \[
    KE = \frac{1}{2} \sum mv^2
    \]
  
  Like a movie

- **MC**
  - Non-deterministic
    - (no time dependence)
    - difficult to find the correlation between different configurations
  - Potential energy only

Random snap shots