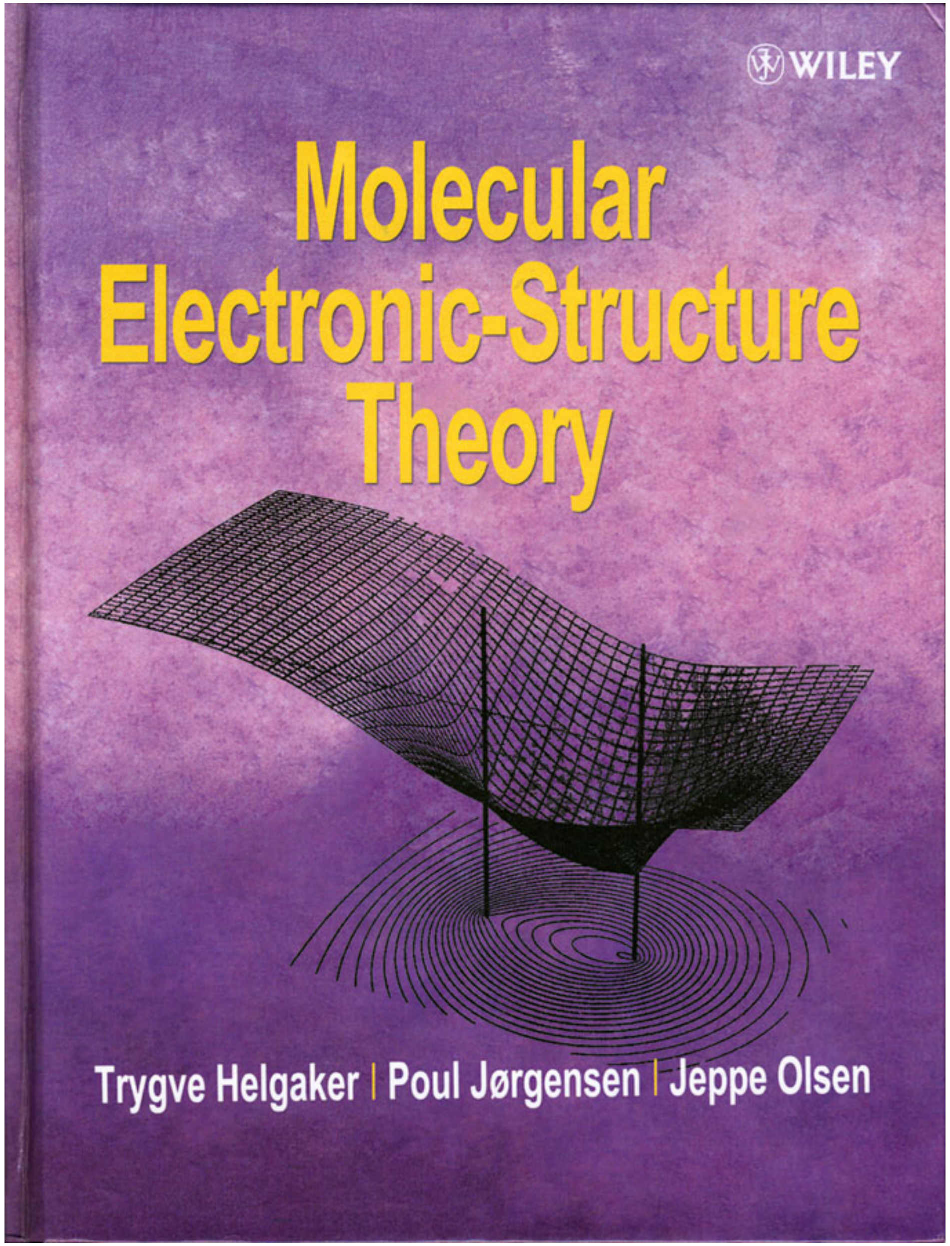
Hartree–Fock and post-Hartree–Fock methods

Emmanuel Fromager

Institut de Chimie de Strasbourg - Laboratoire de Chimie Quantique -
Université de Strasbourg /CNRS

M2 lecture, Strasbourg, France.



WILEY

Molecular
Electronic-Structure
Theory

Trygve Helgaker | Poul Jørgensen | Jeppe Olsen

Notations

- Molecular orbitals: $\phi_p(\mathbf{r}) = \sum_\mu C_{\mu p}\, \chi_\mu(\mathbf{r})$ $\quad\langle\phi_p|\phi_q\rangle = \delta_{pq}$
- *Non-orthogonal* set of atomic orbitals (Gaussian functions): $\langle\chi_\mu|\chi_\nu\rangle = S_{\mu\nu} \neq \delta_{\mu\nu}$
- Hamiltonian in second quantization:

$$\hat{H} = \sum_{p,q} h_{pq}\hat{E}_{pq} + \frac{1}{2}\sum_{p,q,r,s}\langle pr|qs\rangle\Big(\hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps}\Big)$$

where $\quad h_{pq} = \int \mathrm{d}\mathbf{r}\ \phi_p(\mathbf{r}) \times \Big[-\frac{1}{2}\nabla^2_{\mathbf{r}} \ +\ v_{\mathrm{ne}}(\mathbf{r})\Big]\phi_q(\mathbf{r})$

and $\quad \langle pr|qs\rangle \equiv (pq|rs) = \int\int \mathrm{d}\mathbf{r_1}\mathrm{d}\mathbf{r_2}\ \phi_p(\mathbf{r_1})\phi_r(\mathbf{r_2})\,\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\,\phi_q(\mathbf{r_1})\phi_s(\mathbf{r_2})$

Variational and non-variational approximations

- The exact electronic ground state Ψ_0 and its energy E_0 can be obtained two ways:

$$E_0 = \min_{\Psi} \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\langle \Psi_0 | \hat{H} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \qquad \textit{or} \qquad \hat{H} | \Psi_0 \rangle = E_0 | \Psi_0 \rangle$$

- Approximate parametrized ground-state wave function: $\Psi(\boldsymbol{\lambda}_0)$

where $\boldsymbol{\lambda}_0$ denotes the complete set of optimized parameters.

Variational calculation

$$\frac{\partial}{\partial \boldsymbol{\lambda}} \frac{\langle \Psi(\boldsymbol{\lambda}) | \hat{H} | \Psi(\boldsymbol{\lambda}) \rangle}{\langle \Psi(\boldsymbol{\lambda}) | \Psi(\boldsymbol{\lambda}) \rangle} \Bigg|_{\boldsymbol{\lambda} = \boldsymbol{\lambda}_0} = 0$$

$$\downarrow$$

Hartree-Fock (HF)
Configuration Interaction (CI)
Multi-Configurational Self-Consistent Field (MCSCF)

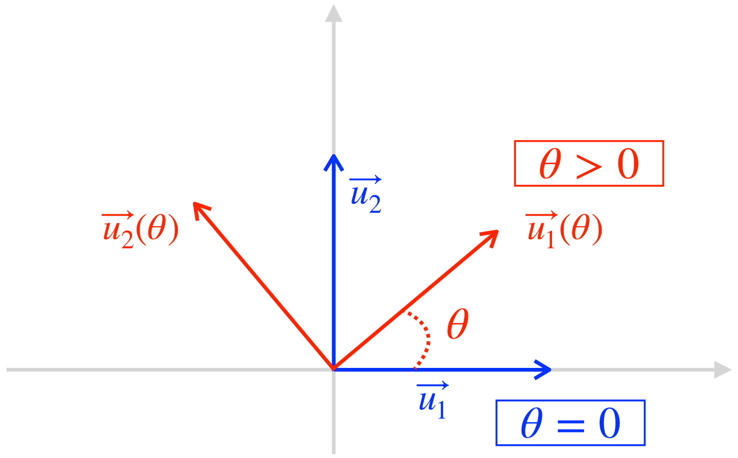
Non-variational calculation

$$\hat{H} | \Psi(\boldsymbol{\lambda}) \rangle - E(\boldsymbol{\lambda}) | \Psi(\boldsymbol{\lambda}) \rangle = 0 \qquad \text{for } \boldsymbol{\lambda} = \boldsymbol{\lambda}_0$$

$$\downarrow$$

Many-Body Perturbation Theory (MBPT)
Coupled Cluster (CC)

Rotation in real space



$$\overrightarrow{u_1}(\theta) = \cos\theta\,\overrightarrow{u_1} + \sin\theta\,\overrightarrow{u_2}$$

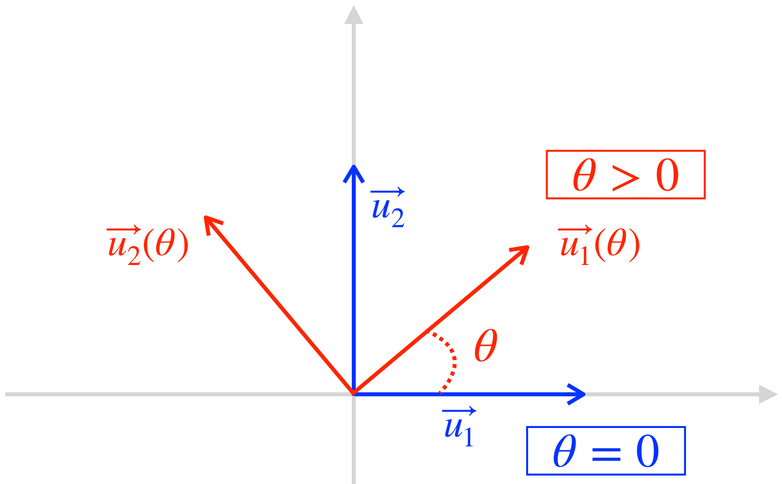
$$\overrightarrow{u_2}(\theta) = -\sin\theta\,\overrightarrow{u_1} + \cos\theta\,\overrightarrow{u_2}$$

Rotation in spin-orbital space

$$|\varphi_1(\theta)\rangle = \cos\theta\,|\varphi_1\rangle + \sin\theta\,|\varphi_2\rangle \equiv \hat{a}^\dagger_{1(\theta)}|\mathrm{vac}\rangle \longleftarrow \hat{a}^\dagger_{1(\theta)} = \cos\theta\,\hat{a}^\dagger_1 + \sin\theta\,\hat{a}^\dagger_2$$

$$|\varphi_2(\theta)\rangle = -\sin\theta\,|\varphi_1\rangle + \cos\theta\,|\varphi_2\rangle \equiv \hat{a}^\dagger_{2(\theta)}|\mathrm{vac}\rangle \longleftarrow \hat{a}^\dagger_{2(\theta)} = -\sin\theta\,\hat{a}^\dagger_1 + \cos\theta\,\hat{a}^\dagger_2$$

Matrix representation of a rotation



$$\vec{u}_1(\theta) = \cos\theta\,\vec{u}_1 + \sin\theta\,\vec{u}_2$$

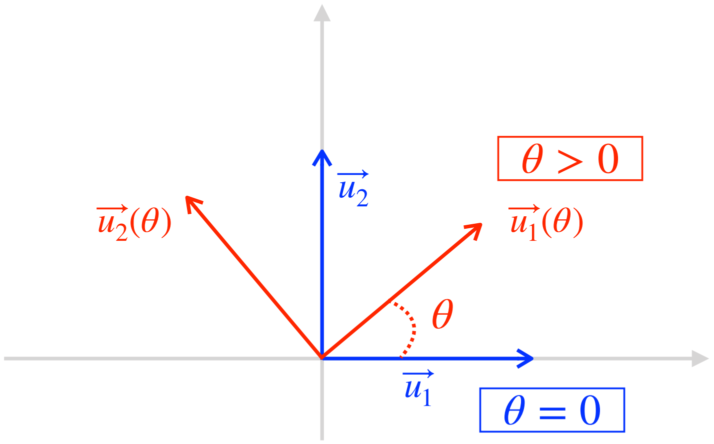
$$\vec{u}_2(\theta) = -\sin\theta\,\vec{u}_1 + \cos\theta\,\vec{u}_2$$

$$\mathbf{U}(\theta) = \begin{array}{cc} \vec{u}_1(\theta) & \vec{u}_2(\theta) \\ \downarrow & \downarrow \\ \left[\begin{array}{c} \cos\theta \\ \sin\theta \end{array}\right. & \left.\begin{array}{c} -\sin\theta \\ \cos\theta \end{array}\right] \end{array} = e^{-\mathbf{K}(\theta)}$$

where

$$\mathbf{K}(\theta) = \begin{bmatrix} 0 & \theta \\ -\theta & 0 \end{bmatrix} = \overset{\textit{anti-hermitian}}{-}\left[\mathbf{K}(\theta)\right]^T = -\mathbf{K}(-\theta)$$

Matrix representation of a rotation



$$\vec{u}_1(\theta) = \cos\theta\,\vec{u}_1 + \sin\theta\,\vec{u}_2$$

$$\vec{u}_2(\theta) = -\sin\theta\,\vec{u}_1 + \cos\theta\,\vec{u}_2$$

$$\mathbf{U}(\theta) = \begin{array}{c} \begin{array}{cc} \vec{u}_1(\theta) & \vec{u}_2(\theta) \\ \downarrow & \downarrow \end{array} \\ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \end{array} = e^{-\mathbf{K}(\theta)}$$

$$\mathbf{K}(\theta) = \begin{bmatrix} 0 & \theta \\ -\theta & 0 \end{bmatrix} = -\left[\mathbf{K}(\theta)\right]^T = -\mathbf{K}(-\theta)$$

$$\boxed{\left[\mathbf{U}(\theta)\right]^{-1}} = e^{+\mathbf{K}(\theta)} = e^{-\mathbf{K}(-\theta)} = \boxed{\mathbf{U}(-\theta)}$$

$$\left[\mathbf{U}(\theta)\right]^{-1} = e^{-\left[\mathbf{K}(\theta)\right]^T} = \left[e^{-\mathbf{K}(\theta)}\right]^T = \boxed{\left[\mathbf{U}(\theta)\right]^T}$$

Spin-orbital rotation

- Let $\{\varphi_P\}_{P=1,2,\ldots}$ denote an orthonormal basis of spin-orbitals and $\{\tilde{\varphi}_P\}_{P=1,2,\ldots}$ another *orthonormal* basis obtained from $\{\varphi_P\}_{P=1,2,\ldots}$ through a unitary transformation:

$$|\varphi_P\rangle \longrightarrow |\tilde{\varphi}_P\rangle = \sum_{Q=1,2,\ldots} U_{QP}|\varphi_Q\rangle$$

- "unitary" means that $\boldsymbol{U}^\dagger \equiv \left[\boldsymbol{U}^T\right]^* = \boldsymbol{U}^{-1}$ where $\boldsymbol{U} \equiv \{U_{PQ}\}_{P,Q=1,2,\ldots}$.

- $\boldsymbol{U}$ can be written as $\boxed{\boldsymbol{U} = e^{-\boldsymbol{\kappa}} \text{ where } \boldsymbol{\kappa}^\dagger = -\boldsymbol{\kappa}}$ $\leftarrow$ $\boldsymbol{U}^\dagger = \left(e^{-\boldsymbol{\kappa}}\right)^\dagger = e^{-\boldsymbol{\kappa}^\dagger} = e^{\boldsymbol{\kappa}} = \boldsymbol{U}^{-1}$

- It can be shown that, in *second quantization*, the unitary transformation of a basis reads as

$$\hat{a}^\dagger_{\tilde{P}} \equiv \sum_Q \left(e^{-\boldsymbol{\kappa}}\right)_{QP} \hat{a}^\dagger_Q = \boxed{e^{-\hat{\kappa}}\, \hat{a}^\dagger_P\, e^{\hat{\kappa}} = \hat{a}^\dagger_{\tilde{P}}} \qquad \text{where} \qquad \hat{\kappa} = \sum_{PQ} \kappa_{PQ}\, \hat{a}^\dagger_P \hat{a}_Q$$

Spin-orbital rotation

- Note that the (so-called) rotation operator $\hat{\kappa}$ is *anti-hermitian*:

$$\hat{\kappa}^{\dagger}=\sum_{PQ}\kappa^{*}_{PQ}\,\hat{a}^{\dagger}_{Q}\hat{a}_{P}=\sum_{PQ}\left[\kappa^{\dagger}\right]_{QP}\hat{a}^{\dagger}_{Q}\hat{a}_{P}=-\sum_{PQ}\kappa_{QP}\,\hat{a}^{\dagger}_{Q}\hat{a}_{P}=-\hat{\kappa}$$

- Unitary transformation of an N-electron *Slater determinant*:

$$|\tilde{P}_1\tilde{P}_2\ldots\tilde{P}_N\rangle=\hat{a}^{\dagger}_{\tilde{P}_1}\hat{a}^{\dagger}_{\tilde{P}_2}\ldots\hat{a}^{\dagger}_{\tilde{P}_N}|\text{vac}\rangle=e^{-\hat{\kappa}}\hat{a}^{\dagger}_{P_1}\,e^{\hat{\kappa}}e^{-\hat{\kappa}}\hat{a}^{\dagger}_{P_2}\,e^{\hat{\kappa}}\ldots e^{-\hat{\kappa}}\hat{a}^{\dagger}_{P_N}\,e^{\hat{\kappa}}|\text{vac}\rangle$$

$$=e^{-\hat{\kappa}}\hat{a}^{\dagger}_{P_1}\hat{a}^{\dagger}_{P_2}\ldots\hat{a}^{\dagger}_{P_N}\,\underbrace{e^{\hat{\kappa}}|\text{vac}\rangle}_{|\text{vac}\rangle}$$

$$\boxed{|\tilde{P}_1\tilde{P}_2\ldots\tilde{P}_N\rangle=e^{-\hat{\kappa}}\,|P_1P_2\ldots P_N\rangle}$$

Spin-restricted orbital rotation

- In a *restricted formalism* the same set of orbitals is used for α and β spins:

$$\hat{\kappa} = \sum_{PQ} \kappa_{PQ}\, \hat{a}^{\dagger}_{P}\hat{a}_{Q} = \sum_{pq}\sum_{\sigma\sigma'} \underbrace{\kappa_{p,\sigma q,\sigma'}}_{\kappa_{pq}\delta_{\sigma\sigma'}}\, \hat{a}^{\dagger}_{p,\sigma}\hat{a}_{q,\sigma'} = \sum_{pq} \kappa_{pq}\hat{E}_{pq}$$

- Since $\kappa_{pq} = -\kappa_{qp}$ (real algebra)

$$\hat{\kappa} = \sum_{p>q} \kappa_{pq}\hat{E}_{pq} - \sum_{p<q} \kappa_{qp}\hat{E}_{pq}$$

$$\boxed{\hat{\kappa} = \sum_{p>q} \kappa_{pq}\left(\hat{E}_{pq} - \hat{E}_{qp}\right)}$$

Hartree-Fock approximation

- For simplicity we consider here the particular case of a non-degenerate singlet closed-shell ground state.

- The HF method consists then in approximating the exact wave function Ψ_0 by a single Slater determinant Φ_0. The orbital space is thus *divided in two*:

doubly occupied molecular orbitals $\phi_i, \phi_j, \ldots$ **unoccupied molecular orbitals** $\phi_a, \phi_b, \ldots$

$$|\Phi_0\rangle = \prod_i^{\text{occ.}} \prod_{\sigma=\alpha,\beta} \hat{a}^{\dagger}_{i,\sigma} |\text{vac}\rangle$$

- The initial set of molecular orbitals is usually not optimized → the optimized HF molecular orbitals will be obtained by means of unitary transformations (orbital rotations).

Hartree-Fock approximation

- Exponential parametrization: $|\Phi(\boldsymbol{\kappa})\rangle = e^{-\hat{\kappa}}\,|\Phi_0\rangle$ with $\hat{\kappa} = \sum_{p>q} \kappa_{pq}\left(\hat{E}_{pq} - \hat{E}_{qp}\right)$

$\boldsymbol{\kappa} = \begin{bmatrix} \vdots \\ \kappa_{pq} \\ \vdots \end{bmatrix}_{p>q}$ denotes the column vector containing all the *to-be-optimized* parameters.

- *occupied-occupied* and *unoccupied-unoccupied* rotations can be *disregarded* (through first order):

$$\hat{\kappa} = \underbrace{\sum_{i>j} \kappa_{ij}\left(\hat{E}_{ij} - \hat{E}_{ji}\right)}_{\hat{\kappa}^{\text{occ.}}} + \sum_{i,a} \kappa_{ai}\left(\hat{E}_{ai} - \hat{E}_{ia}\right) + \underbrace{\sum_{a>b} \kappa_{ab}\left(\hat{E}_{ab} - \hat{E}_{ba}\right)}_{\hat{\kappa}^{\text{unocc.}}}$$

$\hat{\kappa}^{\text{occ.}}|\Phi_0\rangle = \hat{\kappa}^{\text{unocc.}}|\Phi_0\rangle = 0 \rightarrow$ only occupied-unoccupied rotations will be optimized $\rightarrow \boldsymbol{\kappa} = \begin{bmatrix} \vdots \\ \kappa_{ai} \\ \vdots \end{bmatrix}$.

Hartree-Fock approximation

- *Orbital-dependent* Hartree-Fock energy expression:

$$E(\boldsymbol{\kappa}) = \frac{\langle \Phi(\boldsymbol{\kappa})|\hat{H}|\Phi(\boldsymbol{\kappa})\rangle}{\langle \Phi(\boldsymbol{\kappa})|\Phi(\boldsymbol{\kappa})\rangle} = \frac{\langle \Phi_0|e^{-\hat{\kappa}^\dagger}\ \hat{H}\ e^{-\hat{\kappa}}|\Phi_0\rangle}{\langle \Phi_0|e^{-\hat{\kappa}^\dagger}\ e^{-\hat{\kappa}}|\Phi_0\rangle} = \boxed{\langle \Phi_0|e^{\hat{\kappa}}\hat{H}e^{-\hat{\kappa}}|\Phi_0\rangle = E(\boldsymbol{\kappa})}$$

- Variational optimization of $\boldsymbol{\kappa}$: $E^{[1]}_{\boldsymbol{\kappa}_+} = \left.\frac{\partial E(\boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}}\right|_{\boldsymbol{\kappa}_+} = 0$

- Iterative procedure (Newton method):

$$E(\boldsymbol{\kappa}) \approx E(0) + \boldsymbol{\kappa}^T E^{[1]}_0 + \frac{1}{2}\boldsymbol{\kappa}^T E^{[2]}_0 \boldsymbol{\kappa} \quad \rightarrow \quad E^{[1]}_{\boldsymbol{\kappa}_+} \approx E^{[1]}_0 + E^{[2]}_0 \boldsymbol{\kappa}_+ = 0 \quad \rightarrow \quad E^{[2]}_0 \underbrace{\boldsymbol{\kappa}_+}_{} = -E^{[1]}_0$$

Newton step

- Update the HF determinant: $\Phi_0 \leftarrow \Phi(\boldsymbol{\kappa}_+)$

- The HF calculation has converged when the *energy gradient equals zero*: $\boxed{E^{[1]}_0 = 0}$

Fock matrix and canonical HF orbitals

It can be shown *[see the complements]* that $\boxed{E^{[1]}_{0,ai} = -2\langle\Phi_0|\hat{H}\hat{E}_{ai}|\Phi_0\rangle = 0 = -4f_{ia}}$

Brillouin theorem

where the Fock matrix elements are defined as $f_{pq} = h_{pq} + \sum_{rs}\Big(\langle pr|qs\rangle - \frac{1}{2}\langle pr|sq\rangle\Big)D_{rs}$,

$D_{rs} = \langle\Phi_0|\hat{E}_{rs}|\Phi_0\rangle$ ← one-electron density matrix

- Generating the canonical HF orbitals:

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}^{\text{occ.}} & 0 \\ 0 & 0 \end{bmatrix} \quad \mathbf{f} = \begin{bmatrix} \mathbf{f}^{\text{occ.}} & 0 \\ 0 & \mathbf{f}^{\text{unocc.}} \end{bmatrix} \quad \longrightarrow \quad \mathbf{D}' = \mathbf{D}, \quad \mathbf{f}' = \begin{bmatrix} \mathbf{f}'^{\text{occ.}} & 0 \\ 0 & \mathbf{f}'^{\text{unocc.}} \end{bmatrix}$$

$$\mathbf{D}^{\text{occ.}}_{ij} = 2\delta_{ij} \qquad \mathbf{f}^{\text{occ.}}_{ij} = f_{ij}, \quad \mathbf{f}^{\text{unocc.}}_{ab} = f_{ab} \qquad \mathbf{f}'^{\text{occ.}}_{ij} = \delta_{ij}\,\varepsilon_i, \quad \mathbf{f}'^{\text{unocc.}}_{ab} = \delta_{ab}\,\varepsilon_a$$

- Fock operator in second-quantized form using *canonical* orbitals:

$$\hat{F} = \sum_{p,q} f_{pq}\hat{E}_{pq} = \sum_{i} \varepsilon_i\, \hat{E}_{ii} + \sum_{a} \varepsilon_a\, \hat{E}_{aa}$$

- Møller-Plesset partitioning of the Hamiltonian: $\hat{H} = \underbrace{\hat{F}}_{\hat{H}_0} + \underbrace{(\hat{H} - \hat{F})}_{\hat{V}}$ with $\hat{F}|\Phi_0\rangle = \underbrace{2\sum_{i} \varepsilon_i}_{E^{(0)}}|\Phi_0\rangle$

- Using perturbation theory, the HF energy is recovered through *first order*:

$$E^{(0)} + E^{(1)} = E^{(0)} + \langle\Phi_0|\hat{V}|\Phi_0\rangle = \langle\Phi_0|\hat{H}|\Phi_0\rangle = E_{\mathrm{HF}}$$

- The correlation energy is defined as the difference between the exact and HF energies. It is a *second- and higher-order* energy correction:

$$E_c = E_0 - E_{\mathrm{HF}} = E^{(2)} + E^{(3)} + \ldots$$

- Indeed, the HF determinant is an approximation to the exact ground-state wave function:

$$\hat{H}|\Phi_0\rangle = E^{(0)}|\Phi_0\rangle + \hat{V}|\Phi_0\rangle \neq E_0|\Phi_0\rangle \qquad \Longrightarrow \qquad \langle\Phi_I|\hat{H}|\Phi_0\rangle \overset{I>0}{=} \langle\Phi_I|\hat{V}|\Phi_0\rangle \overset{\exists I>0}{\neq} 0$$

Møller-Plesset perturbation theory

$\langle\Phi_i^a|\hat{H}|\Phi_0\rangle=0$ → no projection onto the singly-excited configurations $|\Phi_i^a\rangle = \frac{1}{\sqrt{2}}\hat{E}_{ai}|\Phi_0\rangle$ ...

... but doubly-excited determinants appear when applying $\hat{V}$ to $|\Phi_0\rangle$:

$$\sum_D |D\rangle\langle D|\hat{V}|\Phi_0\rangle = \frac{1}{2}\sum_{a,b,i,j}\langle ab|ij\rangle\hat{E}_{ai}\hat{E}_{bj}|\Phi_0\rangle =$$

$$\sum_{b>a}\sum_{j>i}\sum_{\sigma}\Bigg(\Big(\langle ab|ij\rangle-\langle ab|ji\rangle\Big)\hat{a}^\dagger_{a,\sigma}\hat{a}_{i,\sigma}\hat{a}^\dagger_{b,\sigma}\hat{a}_{j,\sigma}|\Phi_0\rangle + \langle ab|ij\rangle\hat{a}^\dagger_{a,\sigma}\hat{a}_{i,\sigma}\hat{a}^\dagger_{b,-\sigma}\hat{a}_{j,-\sigma}|\Phi_0\rangle$$

$$-\langle ab|ji\rangle\hat{a}^\dagger_{a,\sigma}\hat{a}_{i,-\sigma}\hat{a}^\dagger_{b,-\sigma}\hat{a}_{j,\sigma}|\Phi_0\rangle\Bigg) + \sum_{a}\sum_{j>i}\sum_{\sigma}\langle aa|ij\rangle\hat{a}^\dagger_{a,\sigma}\hat{a}_{i,\sigma}\hat{a}^\dagger_{a,-\sigma}\hat{a}_{j,-\sigma}|\Phi_0\rangle$$

$$+\sum_{i}\sum_{b>a}\sum_{\sigma}\langle ab|ii\rangle\hat{a}^\dagger_{a,\sigma}\hat{a}_{i,\sigma}\hat{a}^\dagger_{b,-\sigma}\hat{a}_{i,-\sigma}|\Phi_0\rangle + \sum_{a,i}\langle aa|ii\rangle\hat{a}^\dagger_{a,\alpha}\hat{a}_{i,\alpha}\hat{a}^\dagger_{a,\beta}\hat{a}_{i,\beta}|\Phi_0\rangle$$

$$\boxed{\langle\hat{a}^\dagger_{a,\sigma}\hat{a}_{i,\sigma}\hat{a}^\dagger_{b,\sigma}\hat{a}_{j,\sigma}|\Phi_0|\hat{V}|\Phi_0\rangle = \langle ab|ij\rangle - \langle ab|ji\rangle \neq 0}$$

- Wave function through *first order*: $|\Psi_0\rangle = |\Phi_0\rangle + \sum_D \frac{|D\rangle\langle D|\hat{V}|\Phi_0\rangle}{E^{(0)} - E_D^{(0)}} + \ldots$

- Energy through *second order* (MP2): $E_0 = E_{\text{HF}} + \sum_D \frac{\langle D|\hat{V}|\Phi_0\rangle^2}{E^{(0)} - E_D^{(0)}} + \ldots$

$$E_0 = E_{\text{HF}} + 4\sum_{b>a}\sum_{j>i} \frac{\langle ab|ij\rangle^2 + \langle ab|ji\rangle^2 - \langle ab|ij\rangle\langle ab|ji\rangle}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} + 2\sum_{a}\sum_{j>i} \frac{\langle aa|ij\rangle^2}{\varepsilon_i + \varepsilon_j - 2\varepsilon_a}$$
$$+2\sum_{b>a}\sum_{i} \frac{\langle ab|ii\rangle^2}{2\varepsilon_i - \varepsilon_a - \varepsilon_b} + \sum_{a,i} \frac{\langle aa|ii\rangle^2}{2\varepsilon_i - 2\varepsilon_a} + \ldots$$

$$\boxed{E_0 = E_{\text{HF}} + \sum_{a,b,i,j} \frac{\langle ab|ij\rangle\Big(2\langle ab|ij\rangle - \langle ab|ji\rangle\Big)}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b} + \ldots}$$

- Note that a correlated wave function *cannot be packed* into a *single Slater determinant*. This is due to the double excitations.

- When single excitations contribute at first order to the wave function, they are not associated to correlation but to orbital relaxation instead: $|\Phi(\boldsymbol{\kappa})\rangle = e^{-\hat{\kappa}}\,|\Phi_0\rangle = |\Phi_0\rangle - \sum_{a,i} \kappa_{ai}\,\hat{E}_{ai}|\Phi_0\rangle + \ldots$

- Note also that, for open-shell atoms, the HF calculation is based on the atomic terms (^{2S+1}L) which consist in linear combinations of Slater determinants which all correspond to the *same configuration* (for example $1s^2 2s^2 2p^3$ for nitrogen). The HF wave function is then multideterminantal but describes one single configuration (no correlation effects).

Configuration Interaction method (CI)

- The idea is to *expand the wave function* in the basis of the *determinants* that are constructed from the canonical (occupied and virtual) *HF orbitals*.
- These determinants are generated by applying single, double, triple, quadruple, ... excitations to the (single reference) HF determinant Φ_0:

$$|\Psi(\mathbf{C})\rangle = C_0|\Phi_0\rangle + \sum_S C_S|S\rangle + \sum_D C_D|D\rangle + \sum_T C_T|T\rangle + \sum_Q C_Q|Q\rangle + \ldots \equiv \sum_\xi C_\xi|\mathrm{det}_\xi\rangle$$

- If no truncation is made in the CI expansion (all excitations included) ⟶ Full CI (FCI) ⟵ *exact for a given one-electron (atomic orbital) basis.*
- Truncated CI models: CIS, CISD, CISDT, CISDTQ, ...
- The coefficients $\mathbf{C} \equiv \left\{C_\xi\right\}_\xi$ are used as *variational* parameters. The energy to be *minimized* is

$$E(\mathbf{C}) = \frac{\langle\Psi(\mathbf{C})|\hat{H}|\Psi(\mathbf{C})\rangle}{\langle\Psi(\mathbf{C})|\Psi(\mathbf{C})\rangle}.$$

- Stationarity condition:

$$\frac{\partial E(\mathbf{C})}{\partial \mathbf{C}} = 0 \;\Leftrightarrow\; \mathbf{HC} = E(\mathbf{C})\mathbf{C} \quad \text{where} \quad \mathbf{H} \equiv \left\{H_{\xi\xi'} = \langle \mathrm{det}_\xi|\hat{H}|\mathrm{det}_{\xi'}\rangle\right\}.$$

H_2 in a minimal basis

EXERCISE:

Show that the Hamiltonian matrix for H_2 can be written in the basis of the two single-determinant states $|1\sigma_g^\alpha 1\sigma_g^\beta\rangle$ and $|1\sigma_u^\alpha 1\sigma_u^\beta\rangle$ as follows,

$$[\hat{H}] = \begin{bmatrix} E_g & K \\ K & E_u \end{bmatrix}, \qquad \text{where}$$

for $\quad i = g, u, \quad E_i = 2h_{ii} + \langle 1\sigma_i 1\sigma_i | 1\sigma_i 1\sigma_i \rangle, \quad h_{ii} = \langle 1\sigma_i | \hat{h} | 1\sigma_i \rangle, \quad K = \langle 1\sigma_u 1\sigma_u | 1\sigma_g 1\sigma_g \rangle.$

Size-consistency problem in truncated CI calculations

- Size-consistency property of a method: multiplicatively separable wave function and additively separable energy, *i.e.*, $\boxed{E(1+2) = E(1) + E(2)}$

  where 1 and 2 denote two non-interacting monomers ($\hat{H} = \hat{H}_1 + \hat{H}_2$).
- Example: H_2 dimer in a minimal basis set
- For the monomer I ($I = 1, 2$), the $1\sigma_{gI}$ and $1\sigma_{uI}$ orbitals only are considered.
- Ground-state HF determinant for the monomer: $|\Phi_0(I)\rangle = \hat{a}^\dagger_{1\sigma_{gI},\alpha}\hat{a}^\dagger_{1\sigma_{gI},\beta}|\text{vac}\rangle$
- Ground-state HF determinant for the dimer: $|\Phi_0(1+2)\rangle = \left(\prod_{I=1}^{2} \hat{a}^\dagger_{1\sigma_{gI},\alpha}\hat{a}^\dagger_{1\sigma_{gI},\beta}\right)|\text{vac}\rangle$

$$E_{\text{HF}}(1+2) = \langle\Phi_0(1+2)|\hat{H}|\Phi_0(1+2)\rangle = \langle\Phi_0(1)|\hat{H}_1|\Phi_0(1)\rangle + \langle\Phi_0(2)|\hat{H}_2|\Phi_0(2)\rangle$$

$$= E_{\text{HF}}(1) + E_{\text{HF}}(2) \qquad \longleftarrow \qquad \text{size-consistent!}$$

Size-consistency problem in truncated CI calculations

- CID corresponds to FCI for the monomer:

$$|\Psi^{\mathrm{CID}}(I)\rangle = \Big(1 + c\hat{D}_I\Big)|\Phi_0(I)\rangle$$

where $\hat{D}_I = \hat{a}^\dagger_{1\sigma_{uI},\alpha}\hat{a}^\dagger_{1\sigma_{uI},\beta}\hat{a}_{1\sigma_{gI},\beta}\hat{a}_{1\sigma_{gI},\alpha}$ $\longleftarrow$ double excitation on monomer I

$$\mathbf{H}^{\mathrm{CID}}(I) - E_{\mathrm{HF}}(I) = \begin{bmatrix} 0 & K \\ K & 2\Delta \end{bmatrix} \quad \longrightarrow \quad \boxed{E^{\mathrm{CID}}(I) = E_{\mathrm{HF}}(I) + \Delta - \sqrt{\Delta^2 + K^2}}$$

- CID is *not* FCI for the dimer: $|\Psi^{\mathrm{CID}}(1+2)\rangle = \Big(1 + c\hat{D}_1 + c\hat{D}_2\Big)|\Phi_0(1+2)\rangle$

$$|\Psi^{\mathrm{FCI}}(1+2)\rangle = \Big(1 + c\hat{D}_1 + c\hat{D}_2 + c_{12}\hat{D}_1\hat{D}_2\Big)|\Phi_0(1+2)\rangle$$

Size-consistency problem in truncated CI calculations

EX5: Show that, for the dimer,

(i) the CID Hamiltonian matrix equals

$$\mathbf{H}^{\mathrm{CID}}(1+2) - E_{\mathrm{HF}}(1+2) = \begin{bmatrix} 0 & K & K \\ K & 2\Delta & 0 \\ K & 0 & 2\Delta \end{bmatrix}$$

(ii) the FCI Hamiltonian matrix equals

$$\mathbf{H}^{\mathrm{FCI}}(1+2) - E_{\mathrm{HF}}(1+2) = \begin{bmatrix} 0 & K & K & 0 \\ K & 2\Delta & 0 & K \\ K & 0 & 2\Delta & K \\ 0 & K & K & 4\Delta \end{bmatrix}$$

(iii) CID is not size-consistent since

$$E^{\mathrm{CID}}(1+2) = E_{\mathrm{HF}}(1+2) + \Delta - \sqrt{\Delta^2 + 2K^2} \neq E^{\mathrm{CID}}(1) + E^{\mathrm{CID}}(2)$$

(iv) FCI is size-consistent and $c_{12} = c^2$

Size-consistency problem in truncated CI calculations

- $E^{\mathrm{FCI}}(1+2) - E^{\mathrm{CID}}(1+2) = \Delta\left(1+\sqrt{1+2\left(\frac{K}{\Delta}\right)^2} - 2\sqrt{1+\left(\frac{K}{\Delta}\right)^2}\right)$

$$= \Delta\left(-\frac{1}{4}\left(\frac{K}{\Delta}\right)^4 + \ldots\right)$$

- FCI wave function written as a Coupled-Cluster wave function (exponential ansatz):

$$|\Psi^{\mathrm{FCI}}(1+2)\rangle = \left(1 + c\hat{D}_1 + c\hat{D}_2 + c^2\hat{D}_1\hat{D}_2\right)|\Phi_0(1+2)\rangle$$

$$= \left(1 + c\hat{D}_1\right)\left(1 + c\hat{D}_2\right)|\Phi_0(1+2)\rangle = e^{c\hat{D}_1}e^{c\hat{D}_2}|\Phi_0(1+2)\rangle$$

$$\boxed{|\Psi^{\mathrm{FCI}}(1+2)\rangle = e^{c\hat{D}_1 + c\hat{D}_2}|\Phi_0(1+2)\rangle}$$

CCD generates quadruple excitations, by means of the exponential, as products of double excitations and thus ensures size-consistency!

Coupled-Cluster model (CC)

- Note that the exponential used in CC enables to describe not only orbital rotations (single excitations) but also *electron correlation*.
- Exponential ansatz in the general case: $\boxed{|\Psi(\mathbf{t})\rangle = e^{\hat{\mathcal{T}}}|\Phi_0\rangle}$ where

$$\hat{\mathcal{T}} = \sum_S t_S \hat{S} + \sum_D t_D \hat{D} + \sum_T t_T \hat{T} + \sum_Q t_Q \hat{Q} + \ldots = \sum_\mu t_\mu \hat{\tau}_\mu$$

$\hat{\tau}_\mu|\Phi_0\rangle = |\mu\rangle \quad \longleftarrow$ excited determinant

$\mathbf{t} = \begin{bmatrix} \vdots \\ t_\mu \\ \vdots \end{bmatrix} \quad \longleftarrow$ CC amplitudes to be optimized

- Truncated and approximate CC models: CCSD, CCSDT, CCSDTQ, CCSD(T), CC2, . . .

Coupled-Cluster model (CC)

- Variational optimization of the CC amplitudes *not convenient*:

$$\frac{\langle\Psi(\mathbf{t})|\hat{H}|\Psi(\mathbf{t})\rangle}{\langle\Psi(\mathbf{t})|\Psi(\mathbf{t})\rangle} = \frac{\langle\Phi_0|e^{\hat{\mathcal{T}}^\dagger}\hat{H}e^{\hat{\mathcal{T}}}|\Phi_0\rangle}{\langle\Phi_0|e^{\hat{\mathcal{T}}^\dagger}e^{\hat{\mathcal{T}}}|\Phi_0\rangle} \longleftarrow \text{not easy to expand } (\hat{\mathcal{T}}^\dagger \neq -\hat{\mathcal{T}})$$

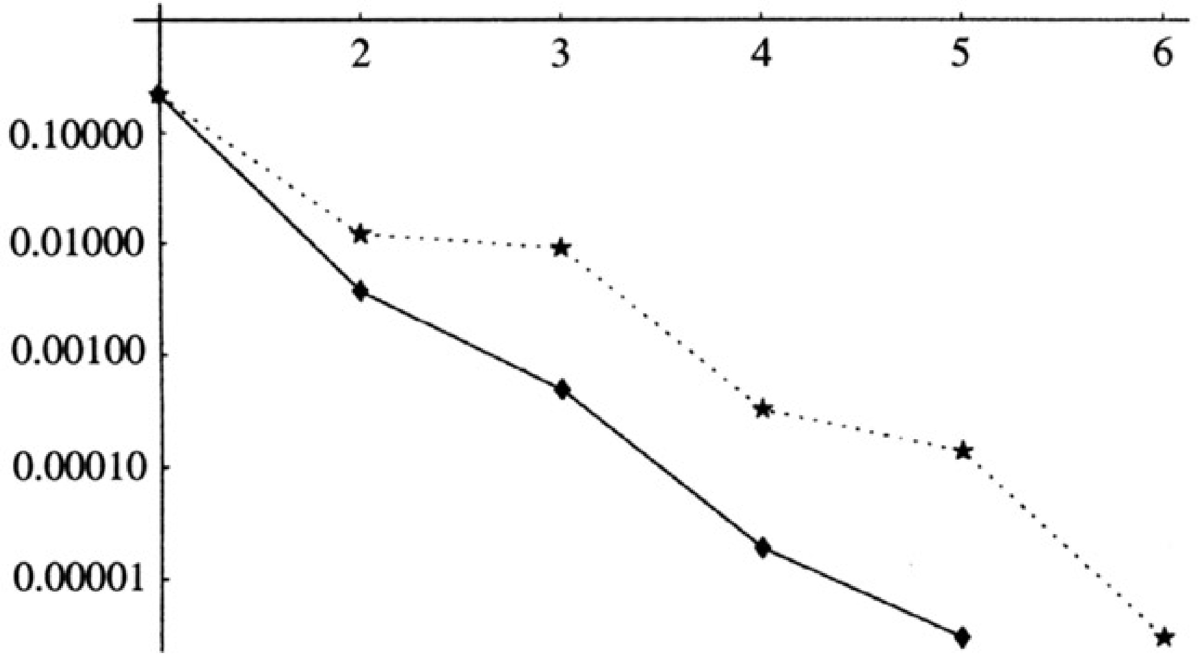
- Non-variational optimization: $\hat{H}|\Psi(\mathbf{t})\rangle = E(\mathbf{t})|\Psi(\mathbf{t})\rangle \longrightarrow \hat{H}e^{\hat{\mathcal{T}}}|\Phi_0\rangle = E(\mathbf{t})e^{\hat{\mathcal{T}}}|\Phi_0\rangle$

- "Linked" formulation: $e^{-\hat{\mathcal{T}}}\hat{H}e^{\hat{\mathcal{T}}}|\Phi_0\rangle = E(\mathbf{t})|\Phi_0\rangle$

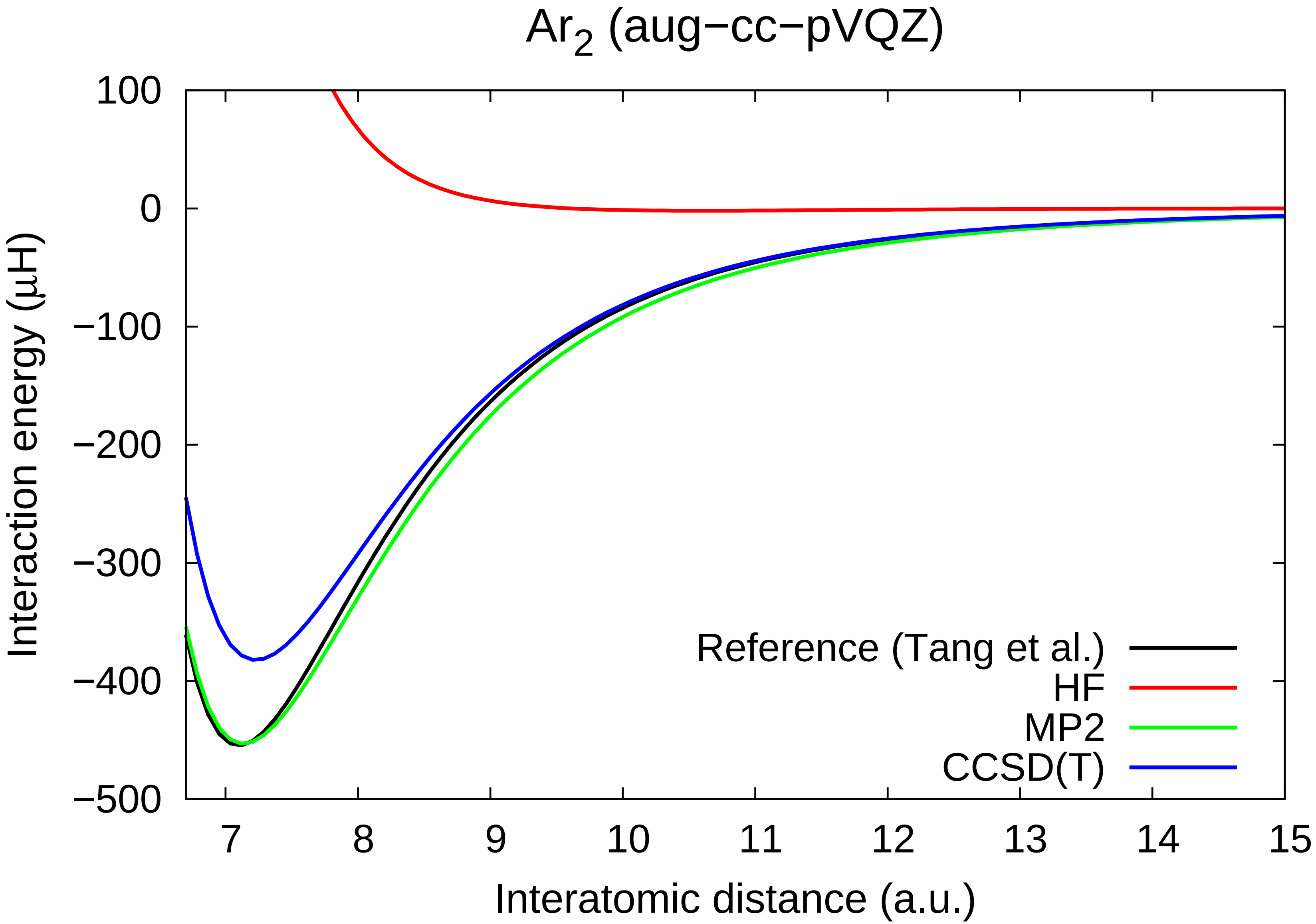
- CC energy: $E(\mathbf{t}) = \langle\Phi_0|e^{-\hat{\mathcal{T}}}\hat{H}e^{\hat{\mathcal{T}}}|\Phi_0\rangle = \langle\Phi_0|\hat{H}e^{\hat{\mathcal{T}}}|\Phi_0\rangle$

$$\boxed{E(\mathbf{t}) = \langle\Phi_0|\hat{H}\left(1 + \sum_D t_D\hat{D} + \frac{1}{2}\Big(\sum_S t_S\hat{S}\Big)^2\right)|\Phi_0\rangle}$$

- CC amplitudes: $\langle\mu|e^{-\hat{\mathcal{T}}}\hat{H}e^{\hat{\mathcal{T}}}|\Phi_0\rangle = 0$ ⟵ *easier to expand* (no terms beyond fourth order!)



The error (with respect to FCI) in the total energy (E_h) of coupled-cluster wave functions (full line) and CI wave functions (dotted line) at different excitation levels for the water molecule at the equilibrium geometry in the cc-pVDZ basis.


Ar$_2$ (aug−cc−pVQZ)
Interaction energy (μH)
100
0
−100
−200
−300
−400
−500
7
8
9
10
11
12
13
14
15
Interatomic distance (a.u.)
Reference (Tang et al.)
HF
MP2
CCSD(T)

Complements

Hartree-Fock approximation

EXERCISE: Using the Taylor expansion of $\hat{f}(x) = e^{-x\hat{A}}\,\hat{B}\,e^{x\hat{A}}$ about $x = 0$, prove the *Baker-Campbell-Hausdorff* (BCH) expansion:

$$e^{-\hat{A}}\,\hat{B}\,e^{\hat{A}} = \hat{B} + \sum_{n=1}^{+\infty} \frac{1}{n!}\,[\![\hat{B},\hat{A}]\!]_n \qquad [\![\hat{B},\hat{A}]\!]_{n+1} = [[\![\hat{B},\hat{A}]\!]_n,\hat{A}], \qquad [\![\hat{B},\hat{A}]\!]_1 = [\hat{B},\hat{A}]$$

$$= \hat{B} + [\hat{B},\hat{A}] + \frac{1}{2}[[\hat{B},\hat{A}],\hat{A}] + \ldots$$

- Analytical formulas for the gradient and the hessian:

$$E(\boldsymbol{\kappa}) = E(0) + \underbrace{\langle\Phi_0|[\hat{\kappa},\hat{H}]|\Phi_0\rangle} + \frac{1}{2}\langle\Phi_0|[\hat{\kappa},[\hat{\kappa},\hat{H}]]|\Phi_0\rangle + \ldots$$

$$\sum_{ai}\kappa_{ai}\langle\Phi_0|[\hat{E}_{ai}-\hat{E}_{ia},\hat{H}]|\Phi_0\rangle \quad\rightarrow\quad E^{[1]}_{0,ai} = \langle\Phi_0|[\hat{E}_{ai}-\hat{E}_{ia},\hat{H}]|\Phi_0\rangle$$

$$= -2\langle\Phi_0|\hat{H}\hat{E}_{ai}|\Phi_0\rangle = 0 \;\textbf{(Brillouin theorem)}$$

Configuration Interaction method (CI)

- Iterative optimization of the CI coefficients C_i:

$$|\Psi(\mathbf{C}^{(0)})\rangle = |\Psi^{(0)}\rangle = \sum_i C_i^{(0)}|i\rangle \quad \longleftarrow$$

normalized starting CI wave function

$$|\Psi(\boldsymbol{\delta})\rangle = \frac{|\Psi^{(0)}\rangle + \hat{Q}|\boldsymbol{\delta}\rangle}{\sqrt{1+\langle\boldsymbol{\delta}|\hat{Q}|\boldsymbol{\delta}\rangle}} \quad \longleftarrow$$

convenient parametrization $\boldsymbol{\delta} = \begin{bmatrix} \vdots \\ \delta_i \\ \vdots \end{bmatrix}$

$$\hat{Q} = 1 - |\Psi^{(0)}\rangle\langle\Psi^{(0)}|, \qquad |\boldsymbol{\delta}\rangle = \sum_i \delta_i|i\rangle, \qquad \langle\Psi^{(0)}|\hat{Q}|\boldsymbol{\delta}\rangle = 0, \qquad \langle\Psi(\boldsymbol{\delta})|\Psi(\boldsymbol{\delta})\rangle = 1$$

- CI energy expression: $E(\boldsymbol{\delta}) = \langle\Psi(\boldsymbol{\delta})|\hat{H}|\Psi(\boldsymbol{\delta})\rangle$

$$= \frac{E(0) + 2\boldsymbol{\delta}^{\mathrm{T}}\mathbf{QHC}^{(0)} + \boldsymbol{\delta}^{\mathrm{T}}\mathbf{QHQ}\boldsymbol{\delta}}{1+\boldsymbol{\delta}^{\mathrm{T}}\mathbf{Q}\boldsymbol{\delta}}$$

where $\mathbf{H}_{ij} = \langle i|\hat{H}|j\rangle$ and $\mathbf{Q} = 1 - \mathbf{C}^{(0)}\mathbf{C}^{(0)\mathrm{T}}$

Configuration Interaction method (CI)

- Variational condition: $E^{[1]}_{\boldsymbol{\delta}_+} = \left.\frac{\partial E(\boldsymbol{\delta})}{\partial \boldsymbol{\delta}}\right|_{\boldsymbol{\delta}_+} = 0$
- Newton method:

$$E(\boldsymbol{\delta}) \approx E(0) + \boldsymbol{\delta}^T E_0^{[1]} + \frac{1}{2}\boldsymbol{\delta}^T E_0^{[2]}\boldsymbol{\delta} \quad \rightarrow \quad E^{[1]}_{\boldsymbol{\delta}_+} \approx E_0^{[1]} + E_0^{[2]}\boldsymbol{\delta}_+ = 0 \quad \rightarrow \quad E_0^{[2]} \underbrace{\boldsymbol{\delta}_+}_{} = -E_0^{[1]}$$

Newton step

EX: Show that the CI gradient and hessian can be expressed as

$$\boxed{E_0^{[1]} = 2\Big(\mathbf{H} - E(0)\Big)\mathbf{C}^{(0)} \quad \text{and} \quad E_0^{[2]} = 2\mathbf{Q}\Big(\mathbf{H} - E(0)\Big)\mathbf{Q}}$$

- Note that $E_0^{[2]}$ cannot be inverted since $E_0^{[2]}\mathbf{C}^{(0)} = 0$
- We can choose $\boldsymbol{\delta}_+$ such that $\mathbf{C}^{(0)\mathrm{T}}\boldsymbol{\delta}_+ = 0 \longrightarrow \underbrace{\left(E_0^{[2]} + 2\alpha\mathbf{C}^{(0)}\mathbf{C}^{(0)\mathrm{T}}\right)}_{G_0^{[2]}}\boldsymbol{\delta}_+ = -E_0^{[1]}$

where $\alpha \neq 0$

Configuration Interaction method (CI)

After some algebra* ... $\boldsymbol{\delta}_{+} = -\left(G_0^{[2]}\right)^{-1} E_0^{[1]} = -\mathbf{C}^{(0)} + \dfrac{\left(\mathbf{H} - E(0)\right)^{-1}\mathbf{C}^{(0)}}{\mathbf{C}^{(0)\mathrm{T}}\left(\mathbf{H} - E(0)\right)^{-1}\mathbf{C}^{(0)}}$

which does not depend on α since $\mathbf{C}^{(0)\mathrm{T}} E_0^{[1]} = 0$

- Update of the CI vector:

$$\mathbf{C}^{(0)} \longrightarrow \mathbf{C}^{(0)} + \mathbf{Q}\boldsymbol{\delta}_{+} = \boxed{\mathbf{C}^{(0)} + \boldsymbol{\delta}_{+} = \frac{\left(\mathbf{H} - E(0)\right)^{-1}\mathbf{C}^{(0)}}{\mathbf{C}^{(0)\mathrm{T}}\left(\mathbf{H} - E(0)\right)^{-1}\mathbf{C}^{(0)}}}$$

  and then normalize.
- The CI calculation has converged when $E_0^{[1]} = 0 \longrightarrow \mathbf{H}\mathbf{C}^{(0)} = E(0)\mathbf{C}^{(0)}$
- This procedure is also known as Rayleigh method.

*T. Helgaker, P. Jørgensen, and J. Olsen, Molecular Electronic- Structure Theory (Wiley, Chichester, 2004), pp. 544-545.