Coupled-Electron Pair Approximations (CEPAs)

The CEPA methods are techniques for computing dynamic correlation energy on top of a Hartree-Fock (HF) reference. For convenience, define a correlation energy E_c and a correlated Hamiltonian $\hat{H}_\mathrm{c}$:

$$E_c = E_\mathrm{e} - E_\mathrm{HF} \qquad\qquad \hat{H}_\mathrm{c} = \hat{H}_\mathrm{e} - E_\mathrm{HF}$$

Obeying the convention that indices $i, j, k...$ correspond to occupied orbitals and $a, b, c, ...$ to virtual orbitals, we also define a doubles excitation operator:

$$\hat{T}_2|\Phi\rangle = \frac{1}{4}\sum_{ijab} t_{ij}^{ab}|\Phi_{ij}^{ab}\rangle.$$

With these definitions and a Hartree-Fock reference $|\Phi\rangle$, the CCD equations are:

$$E_c = \langle\Phi|\hat{H}_c\hat{T}_2|\Phi\rangle \tag{1}$$

$$E_c t_{ij}^{ab} = \langle\Phi_{ij}^{ab}|\hat{H}_c|\Phi\rangle + \langle\Phi_{ij}^{ab}|\hat{H}_c\hat{T}_2|\Phi\rangle + \frac{1}{2}\langle\Phi_{ij}^{ab}|\hat{H}_c\hat{T}_2{}^2|\Phi\rangle \tag{2}$$

where (1) is called the CCD energy expression and (2) is the CCD amplitude equations for the excitation $ij \to ab$. Both expressions can be simplified using Slater rules or Wick's theorem.

The simplified energy expression (1) is:

$$E_c = \cancel{\langle\Phi|\hat{H}_c|\Phi\rangle} + \frac{1}{4}\sum_{klcd} t_{kl}^{cd}\langle\Phi|\hat{H}_c|\Phi_{kl}^{cd}\rangle$$

$$E_c = \frac{1}{4}\sum_{klcd}\langle kl||cd\rangle t_{kl}^{cd}$$

For the amplitude equations, we can split up the $\hat{T}_2^2$ term into linked and unlinked contributions.

$$\begin{aligned}\frac{1}{2}\langle\Phi_{ij}^{ab}|\hat{H}_c\hat{T}_2^2|\Phi\rangle &= \frac{1}{2}\left(\frac{1}{4}\right)^2\sum_{klcd}\sum_{mnef} t_{kl}^{cd}t_{mn}^{ef}\langle\Phi_{ij}^{ab}|\hat{H}_c|\rangle\Phi_{klmn}^{cdef} \\ &= \Delta_U + \Delta_L\end{aligned}$$

where the unlinked (Δ_U) and linked (Δ_L) contributions are defined as follows.

$$\begin{aligned}\Delta_U &= \frac{1}{4}\sum_{klcd}\langle kl||cd\rangle t_{kl}^{cd}t_{ij}^{ab} \\ \Delta_L &= -\frac{1}{2}\hat{P}(ab)\sum_{klcd}\langle kl||cd\rangle t_{ij}^{ac}t_{kl}^{bd} - \frac{1}{2}\sum_{klcd}\hat{P}(ij)\langle kl||cd\rangle t_{ik}^{ab}t_{jl}^{cd} \\ &\quad + \frac{1}{4}\sum_{klcd}\langle kl||cd\rangle t_{ij}^{cd}t_{kl}^{ab} + \hat{P}(ij)\sum_{klcd}\langle kl||cd\rangle t_{ik}^{ac}t_{jl}^{bd}\end{aligned}$$

Note that $\Delta_U = E_c t_{ij}^{ab}$, the left-hand side of equation (2). Thus the amplitude equation (2) simplifies to

$$\cancel{E_c t_{ij}^{ab}} = \langle\Phi_{ij}^{ab}|\hat{H}_c|\Phi\rangle + \langle\Phi_{ij}^{ab}|\hat{H}_c\hat{T}_2|\Phi\rangle + \cancel{\Delta_U} + \Delta_L$$

Applying these simplifications, the energy and amplitude equations can be written as

$$E_c = \frac{1}{4}\sum_{klcd}\langle kl||cd\rangle t_{kl}^{cd} \tag{3}$$

$$0 = \langle\Phi_{ij}^{ab}|\hat{H}_c|\Phi\rangle + \langle\Phi_{ij}^{ab}|\hat{H}_c\hat{T}_2|\Phi\rangle + \Delta_L \tag{4}$$

The CCD expression for Δ_L is nonlinear in the t-amplitudes (they are multiplied by each other). This is what makes the coupled-cluster equations very complex and difficult to parallelize in a production code. Thus, the goal of the CEPA methods is to include a linear approximation for Δ_L.

CEPA(0)

The simplest CEPA technique, known as CEPA(0), simply assumes $\Delta_L = 0$, yielding the equations

$$E_c = \frac{1}{4}\sum_{klcd}\langle kl||cd\rangle t_{kl}^{cd} \tag{5}$$

$$0 = \langle \Phi_{ij}^{ab}|\hat{H}_c|\Phi\rangle + \langle \Phi_{ij}^{ab}|\hat{H}_c\hat{T}_2|\Phi\rangle \tag{6}$$

Note that the energy expression has not changed, but the amplitudes in Equation 6 have been simplified to include only linear terms. The motivation for this simplification is that the linked contributions to the $\hat{T}_2^2$ matrix element mostly cancel each other out. This has proven to be a fair assumption; CEPA_0 often outperforms CCSD in the equilibrium region, albeit due to error cancellation.

CEPA(1)

For the CEPA(1) method, we introduce the concept of pair energies:

$$\epsilon_{ij} = \frac{1}{2}\sum_{ab} c_{ij}^{ab}\langle\Phi|\hat{H}_c|\Phi_{ij}^{ab}\rangle$$

Physically we can think of the pair energy as the energy contributions of those configurations that have electrons excited from the same occupied orbitals, summed over all possible excitations into virtual orbitals. Using this notation, the CEPA(1) technique approximates the unlinked contributions as follows:

$$\Delta_L = c_{ij}^{ab}\left(\epsilon_{ij} + \frac{1}{2}\sum_k (\epsilon_{ik} + \epsilon_{kj})\right)$$

Notably, approximation has the desired property of being exact for separated electron pairs. The resulting amplitude equations are given by

$$c_{ij}^{ab}\left(\epsilon_{ij} + \frac{1}{2}\sum_k (\epsilon_{ik} + \epsilon_{kj})\right) = \langle \Phi_{ij}^{ab}|\hat{H}_c|\Phi\rangle + \langle \Phi_{ij}^{ab}|\hat{H}_c\hat{T}_2|\Phi\rangle. \tag{7}$$