

(continued from previous page)

```

end
%epnrmr
  GIAO_2el GIAO_2el_RIJCOSX      # Also use RIJCOSX for GIAO integrals
end                               # (This is the default for !RIJCOSX)
*xyzfile 0 1 geometry.xyz

```

Numerical DLPNO-MP2 derivatives

The various truncations in local correlation methods introduce small discontinuities in the potential energy surface. For example, a small displacement may change the sizes of correlation domains, leading to a slightly larger or smaller error from the domain approximation. The default DLPNO-MP2 truncation thresholds are conservative enough, so that these discontinuities should not cause problems in geometry optimizations using analytic gradients.^[352] However, if one wishes to calculate (semi-)numerical derivatives of the DLPNO-MP2 energy, gradient, or properties using finite differences, large errors can occur. Therefore, in these cases it is advisable to keep the pair lists and correlation domains fixed upon displacement. Currently, this can be achieved using the following procedure: first, the calculation at the reference geometry is carried out with the additional setting `StoreDLPNOData=true`:

```

! DLPNO-MP2 def2-SVP def2-SVP/C VeryTightSCF
%base "calc0"
%mp2
  StoreDLPNOData true
end
*xyzfile 0 1 geom0.xyz

```

This will produce the additional files *calc0.MapDLPNO00.tmp*, *calc0.MapDLPNOPre0.tmp*, *calc0.IJLIST.0.tmp*, *calc0.IJLISTSCR.0.tmp*, and *calc0.IJNPNO.0.tmp*, which are needed in the working directory for the next step together with the localized orbitals in *calc0.loc*. The calculation at the displaced geometry is then requested as:

```

! DLPNO-MP2 def2-SVP def2-SVP/C VeryTightSCF
%mp2
  RefBaseName "calc0"
end
*xyzfile 0 1 geom1.xyz

```

The program will use the orbitals from *calc0.loc* as a starting guess for the localization and map the reference orbitals to the new ones based on maximal overlap. The lists of correlated and screened out pairs are read from the files *calc0.IJLIST.0.tmp* and *calc0.IJLISTSCR.0.tmp*, while the domain information (MO-PAO, MO-Aux, etc.) is read from *calc0.MapDLPNO00.tmp* and *calc0.MapDLPNOPre0.tmp*. The number of PNOs for a each pair (stored in *calc0.IJNPNO.0.tmp*) is also kept consistent with the reference calculation: the ones with the highest occupation numbers are kept, disregarding T_{CutPNO} .

This procedure should improve the accuracy and numerical stability for manually calculated geometric derivatives of various DLPNO-MP2 properties (including those that require analytic first or second derivatives at the displaced geometries). For semi-numerical Hessian calculations (NumFreq), it is sufficient to add `StoreDLPNOData=true` as shown below and ORCA will handle the rest. For the sake of numerical stability, it is also recommended to increase `PAOOverlapThresh` and add a PNO level shift for the reasons described in section [Local MP2 Second Derivatives and Response Properties](#).

```

! DLPNO-MP2 def2-SVP def2-SVP/C VeryTightSCF NumFreq
%mp2
  StoreDLPNOData true
  PAOOverlapThresh 1e-5
  TScalePNO_LShift 0.1
end
# geometry definition

```

Note that in case the orbital localization Hessian is (near-)singular, the mapping of orbitals from reference to displaced geometries will likely fail. No solution is presently implemented for this problem.

Multi-Level DLPNO-MP2 calculations

With the DLPNO-MP2 method it is possible to treat the interactions among different fragments of a system with varying accuracy, or exclude some interactions from the electron correlation treatment entirely. A more detailed discussion in the DLPNO-CCSD(T) context is given in section [Multi-Level Calculations](#) and in ref. [355]. Here we just present the technical capabilities of the MP2 module and the required input. Currently, multilayer calculations are only available for closed-shell DLPNO-MP2. Multilayer gradients and response properties are also possible. Fragments must be defined – see [Fragment Specification](#).

```
! DLPNO-MP2 NoFrozenCore TightPNO
%mp2
  LoosePNOFragInter {1 *} # * can be used as a wildcard for either or both indices
  NormalPNOFragInter {1 1} {1 2} # multiple fragment pairs can be listed like this
  TightPNOFragInter {2 3}
  HFFragInter {3 1} {4 2}
  CustomFragInter
    FragPairs {4 4} {3 4} # pair definition is required
    HFOnly false # flag to skip MP2 for these pairs - same as HFFragInter
    FrozenCore false # flag to skip core correlation - requires !NoFrozenCore
    TCutPNO 1e-8 # custom value for these pairs
    TCutDO 1e-2 # custom value for these pairs
    TCutDOij 1e-5 # custom value for these pairs
    TCutPre 1e-6 # custom value for these pairs
  end
end
# geometry and fragment definition
```

Note that a given pair or fragments can only belong to a single layer and definitions later in the input overwrite previous ones. This means that if the above input is used in a 4-fragment calculation, the 1-4 interfragment interactions will be treated with LoosePNO thresholds, the interactions within fragment 1 and with fragment 2 – with NormalPNO thresholds, 2-3 pairs – with TightPNO, 1-3 and 2-4 pairs will be left at the HF level, 3-4 and 4-4 pairs will be treated with $T_{\text{CutPNO}} = 10^{-8}$ and $T_{\text{CutDO}} = 10^{-2}$ (i.e. the NormalPNO defaults), and 2-2 and 3-3 pairs will be left at the global (TightPNO) settings.

3.9.12 Explicitly correlated MP2 calculations

ORCA features an efficient explicit correlation module that is available for MP2 and coupled-cluster calculations (section [Explicitly Correlated Methods: F12-MP2 and F12-CCSD\(T\)](#)). It is described below in the context of coupled-cluster calculations.

3.9.13 MP2 in “Double-Hybrid” Density Functional Theory

A slightly more general form is met in the double-hybrid DFT gradient. The theory is briefly described below.

The energy expression for perturbatively and gradient corrected hybrid functionals as proposed by Grimme is:

$$E = V_{NN} + \langle \mathbf{Ph}^+ \rangle + \frac{1}{2} \int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 - \frac{1}{2} a_x \sum_{\mu\nu\kappa\tau\sigma} P_{\mu\kappa}^\sigma P_{\nu\tau}^\sigma (\mu\nu|\kappa\tau) + c_{\text{DF}} E_{\text{XC}}[\rho_\alpha, \rho_\beta] + c_{\text{PT}} E_{\text{PT}}$$

$$E = E_{\text{SCF}} + c_{\text{PT}} E_{\text{PT}} \quad (3.31)$$

Here V_{NN} is the nuclear repulsion energy and $h_{\mu\nu}$ is a matrix element of the usual one-electron operator which contains the kinetic energy and electron-nuclear attraction terms ($\langle \mathbf{ab} \rangle$ denotes the trace of the matrix product \mathbf{ab}). As usual, the molecular spin-orbitals are expanded in atom centered basis functions ($\sigma = \alpha, \beta$):

$$\psi_p^\sigma(\mathbf{r}) = \sum_{\mu} c_{\mu p}^\sigma \varphi_{\mu}(\mathbf{r}) \quad (3.32)$$

with MO coefficients $c_{\mu p}^\sigma$. The total density is given by (real orbitals are assumed throughout):

$$\rho(\mathbf{r}) = \sum_{i\sigma} |\psi_i^\sigma(\mathbf{r})|^2 = \sum_{\mu\nu\sigma} P_{\mu\nu}^\sigma \varphi_\mu(\mathbf{r}) \varphi_\nu(\mathbf{r}) = \rho^\alpha(\mathbf{r}) + \rho^\beta(\mathbf{r}) \quad (3.33)$$

Where $\mathbf{P} = \mathbf{P}^\alpha + \mathbf{P}^\beta$ and $P_{\mu\nu}^\sigma = \sum_{i\sigma} c_{\mu i}^\sigma c_{\nu i}^\sigma$.

The second term of (3.31) represents the Coulombic self-repulsion. The third term represents the contribution of the Hartree-Fock exchange with the two-electron integrals being defined as:

$$(\mu\nu|\kappa\tau) = \int \int \phi_\mu(\mathbf{r}_1) \phi_\nu(\mathbf{r}_1) r_{12}^{-1} \phi_\kappa(\mathbf{r}_2) \phi_\tau(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (3.34)$$

The mixing parameter a_x controls the fraction of Hartree-Fock exchange and is of a semi-empirical nature. The exchange correlation contribution may be written as:

$$E_{XC}[\rho_\alpha, \rho_\beta] = (1 - a_x) E_X^{\text{GGA}}[\rho_\alpha, \rho_\beta] + b E_C^{\text{GGA}}[\rho_\alpha, \rho_\beta] \quad (3.35)$$

Here, $E_X^{\text{GGA}}[\rho_\alpha, \rho_\beta]$ is the exchange part of the XC- functional in question and $E_C^{\text{GGA}}[\rho_\alpha, \rho_\beta]$ is the correlation part. The parameter b controls the mixing of DFT correlation into the total energy and the parameter c_{DF} is a global scaling factor that allows one to proceed from Hartree-Fock theory ($a_X = 1$, $c_{\text{DF}} = 0$, $c_{\text{PT}} = 0$) to MP2 theory ($a_X = 1$, $c_{\text{DF}} = 0$, $c_{\text{PT}} = 1$) to pure DFT ($a_X = 1$, $c_{\text{DF}} = 0$, $c_{\text{PT}} = 1$) to hybrid DFT ($0 < a_X < 1$, $c_{\text{DF}} = 1$, $c_{\text{PT}} = 0$) and finally to the general perturbatively corrected methods discussed in this work ($0 < a_X < 1$, $c_{\text{DF}} = 1$, $0 < c_{\text{PT}} < 1$). As discussed in detail by Grimme, the B2- PLYP functional uses the Lee-Yang-Parr (LYP) functional as correlation part, the Becke 1988 (B88) functional as GGA exchange part and the optimum choice of the semi-empirical parameters was determined to be $a_X = 0.53$, $c_{\text{PT}} = 0.27$, $c_{\text{DF}} = 1$, $b = 1 - c_{\text{PT}}$. For convenience, we will suppress the explicit reference to the parameters a_X and b in the XC part and rewrite the gradient corrected XC energy as:

$$E_{XC}[\rho^\alpha, \rho^\beta] = \int f(\rho^\alpha, \rho^\beta, \gamma^{\alpha\alpha}, \gamma^{\beta\beta}, \gamma^{\alpha\beta}) d\mathbf{r} \quad (3.36)$$

with the gradient invariants $\gamma^{\sigma\sigma'} = \vec{\nabla}\rho^\sigma \vec{\nabla}\rho^{\sigma'}$. The final term in eq (48) represents the scaled second order perturbation energy:

$$E^{\text{PT2}} = \frac{1}{2} \sum_{i_\alpha < j_\alpha} \langle \mathbf{t}^{i_\alpha j_\alpha} \mathbf{R}^{i_\alpha j_\alpha +} \rangle + \frac{1}{2} \sum_{i_\beta < j_\beta} \langle \mathbf{t}^{i_\beta j_\beta} \mathbf{R}^{i_\beta j_\beta +} \rangle + \sum_{i_\alpha, j_\beta} \langle \mathbf{t}^{i_\alpha j_\beta} \mathbf{R}^{i_\alpha j_\beta +} \rangle \quad (3.37)$$

The PT2 amplitudes have been collected in matrices $\mathbf{t}^{i_\sigma j_{\sigma'}}$ with elements:

$$t_{a_\sigma b_{\sigma'}}^{i_\sigma j_{\sigma'}} = \bar{K}_{a_\sigma b_{\sigma'}}^{i_\sigma j_{\sigma'}} (\varepsilon_i^\sigma + \varepsilon_j^{\sigma'} - \varepsilon_a^\sigma - \varepsilon_b^{\sigma'})^{-1} \quad (3.38)$$

Where the orbitals were assumed to be canonical with orbital energies ε_p^σ . The exchange operator matrices are $\bar{K}_{a_\sigma b_{\sigma'}}^{i_\sigma j_{\sigma'}} = (i_\sigma a_\sigma | j_{\sigma'} b_{\sigma'})$ and the anti-symmetrized exchange integrals are defined as $\bar{K}_{a_\sigma b_{\sigma'}}^{i_\sigma j_{\sigma'}} = (i_\sigma a_\sigma | j_{\sigma'} b_{\sigma'}) - \delta_{\sigma\sigma'} (i_\sigma b_\sigma | j_{\sigma'} a_\sigma)$.

The orbitals satisfy the SCF equations with the matrix element of the SCF operator given by:

$$F_{\mu\nu}^\sigma = h_{\mu\nu} + \sum_{\kappa\tau} P_{\kappa\tau} (\mu\nu|\kappa\tau) - a_X P_{\kappa\tau}^\sigma (\mu\kappa|\nu\tau) + c_{\text{DF}} (\mu|V_{\text{XC}}^\sigma|\nu) \quad (3.39)$$

The matrix elements of the XC-potential for a gradient corrected functional are: [180]

$$(\mu|V_{\text{XC}}^\sigma|\nu) = \int \left\{ \frac{\delta f}{\delta \rho_\alpha} (\varphi_\mu \varphi_\nu) + 2 \frac{\delta f}{\delta \gamma_{\alpha\alpha}} \vec{\nabla} \rho_\alpha \vec{\nabla} (\varphi_\mu \varphi_\nu) + \frac{\delta f}{\delta \gamma_{\alpha\beta}} \vec{\nabla} \rho_\beta \vec{\nabla} (\varphi_\mu \varphi_\nu) \right\} d\mathbf{r} \quad (3.40)$$

The energy in equation (3.31) depends on the MO-coefficients, the PT2-amplitudes and through V_{NN} , V_{eN} (in \hbar) and the basis functions also explicitly on the molecular geometry. Unfortunately, the energy is only stationary with respect to the PT2 amplitudes since they can be considered as having been optimized through the minimization of the Hylleraas functional:

$$E_{\text{PT2}} = \min_{\mathbf{t}} \left\{ \frac{1}{2} \sum_{i_\alpha < j_\alpha} \langle \mathbf{t}^{i_\alpha j_\alpha} \mathbf{R}^{i_\alpha j_\alpha +} \rangle + \frac{1}{2} \sum_{i_\beta < j_\beta} \langle \mathbf{t}^{i_\beta j_\beta} \mathbf{R}^{i_\beta j_\beta +} \rangle + \sum_{i_\alpha, j_\beta} \langle \mathbf{t}^{i_\alpha j_\beta} \mathbf{R}^{i_\alpha j_\beta +} \rangle + \langle \mathbf{D}'^\alpha \mathbf{F}^{\alpha+} \rangle + \langle \mathbf{D}'^\beta \mathbf{F}^{\beta+} \rangle \right\} \quad (3.41)$$

The unrelaxed PT2 difference density is defined as:

$$D'_{ij}{}^\alpha = -\frac{1}{2} \sum_{k_\alpha} \langle \mathbf{t}^{i_\alpha k_\alpha} \mathbf{t}^{k_\alpha j_\alpha} \rangle - \sum_{k_\beta} \langle \mathbf{t}^{i_\alpha k_\beta} \mathbf{t}^{k_\beta j_\alpha} \rangle \quad (3.42)$$

$$D'_{ab}{}^\alpha = \sum_{i_\alpha < j_\alpha} \mathbf{t}^{i_\alpha j_\alpha} \mathbf{t}^{i_\alpha j_\alpha} + \sum_{i_\beta j_\alpha} \mathbf{t}^{i_\beta j_\alpha} \mathbf{t}^{i_\beta j_\alpha} \quad (3.43)$$

With analogous expressions for the spin-down unrelaxed difference densities. Minimization of this functional with respect to the amplitudes yields the second order perturbation energy. The derivative of the SCF part of equation (3.31) with respect to a parameter “ x ” is straightforward and well known. It yields:

$$E_{\text{SCF}}^x = V_{NN}^x + \langle \mathbf{P}^x \rangle + \langle \mathbf{W}^{\text{SCF}} \mathbf{S}^{(x)} \rangle + \sum_{\mu\nu\kappa\tau} \Gamma_{\mu\nu\kappa\tau} (\mu\nu|\kappa\tau)^{(x)} + \sum_{\substack{\sigma \\ (\sigma' \neq \sigma)}} \int \left\{ \frac{\delta f}{\delta \rho_\sigma(\mathbf{r})} \rho_\sigma^{(x)} + 2 \frac{\delta f}{\delta \gamma_{\sigma\sigma'}} \vec{\nabla} \rho_\sigma \vec{\nabla} \rho_{\sigma'} + \frac{\delta f}{\delta \gamma_{\sigma\sigma'}} \vec{\nabla} \rho_{\sigma'} \vec{\nabla} \rho_\sigma^{(x)} \right\} d\mathbf{r} \quad (3.44)$$

Superscript “ x ” refers to the derivative with respect to some perturbation “ x ” while a superscript in parentheses indicates that only the derivative of the basis functions with respect to “ x ” is to be taken. For example:

$$\rho_\sigma^{(x)} = \sum_{\mu\nu} P_{\mu\nu}^\sigma \left\{ \frac{\partial \varphi_\mu}{\partial x} \varphi_\nu + \varphi_\mu \frac{\partial \varphi_\nu}{\partial x} \right\} \\ h_{\mu\nu}^x = \left(\frac{\partial \varphi_\mu}{\partial x} | \hat{h} | \varphi_\nu \right) + \left(\varphi_\mu | \hat{h} | \frac{\partial \varphi_\nu}{\partial x} \right) + \left(\varphi_\mu | \frac{\partial \hat{h}}{\partial x} | \varphi_\nu \right) \quad (3.45)$$

In equation (3.44), \mathbf{S} is the overlap matrix and \mathbf{W}^{SCF} the energy weighted density:

$$W_{\mu\nu}^{\text{SCF}} = W_{\mu\nu}^{\alpha;\text{SCF}} + W_{\mu\nu}^{\beta;\text{SCF}} = - \sum_{i\sigma} c_{\mu i}^\sigma c_{\nu i}^\sigma \varepsilon_i^\sigma \quad (3.46)$$

At this point, the effective two-particle density matrix is fully separable and reads:

$$\Gamma_{\mu\nu\kappa\tau} = \frac{1}{2} P_{\mu\nu} P_{\kappa\tau} - \frac{1}{2} a_x P_{\mu\kappa}^\alpha P_{\nu\tau}^\alpha - \frac{1}{2} a_x P_{\mu\kappa}^\beta P_{\nu\tau}^\beta \quad (3.47)$$

The derivative of the PT2 part is considerably more complex, since E_{PT2} is not stationary with respect to changes in the molecular orbitals. This necessitates the solution of the coupled-perturbed SCF (CP-SCF) equations. We follow the standard practice and expand the perturbed orbitals in terms of the unperturbed ones as:

$$\psi_p^{\sigma;x}(\mathbf{r}) = \sum_q U_{qp}^{\sigma;x} \psi_q^\sigma(\mathbf{r}) \quad (3.48)$$

The occupied-occupied and virtual-virtual blocks of \mathbf{U} are fixed, as usual, through the derivative of the orthonormality constraints:

$$U_{ij}^{\sigma;x} = -\frac{1}{2} S_{ij}^{\sigma(x)} \quad (3.49)$$

$$U_{ab}^{\sigma;x} = -\frac{1}{2} S_{ab}^{\sigma(x)} \quad (3.50)$$

$$U_{ia}^{\sigma;x} = -S_{ia}^{\sigma(x)} - U_{ai}^{\sigma;x} \quad (3.51)$$

Where $S_{pq}^{\sigma(x)} = \sum_{\mu\nu} c_{\mu p}^\sigma c_{\nu q}^\sigma S_{\mu\nu}^{(x)}$. The remaining virtual-occupied block of \mathbf{U}^x must be determined through the solution of the CP-SCF equations. However, as shown by Handy and Schaefer, this step is unnecessary and only a single set of CP-SCF equations (Z-vector equations) needs to be solved. To this end, one defines the Lagrangian:

$$L_{ai}^\alpha = R^\alpha(\mathbf{D}')_{ai} + 2 \sum_{j_\alpha b_\alpha c_\alpha} (a_\alpha c_\alpha | j_\alpha b_\alpha) t_{c_\alpha b_\alpha}^{i_\alpha j_\alpha} - 2 \sum_{j_\alpha k_\alpha b_\alpha} (k_\alpha i_\alpha | j_\alpha b_\alpha) t_{a_\alpha b_\alpha}^{k_\alpha j_\alpha} \\ + 2 \sum_{j_\beta b_\beta c_\alpha} (a_\alpha c_\alpha | j_\beta b_\beta) t_{b_\beta c_\alpha}^{j_\beta i_\alpha} - 2 \sum_{j_\beta k_\alpha b_\beta} (k_\alpha i_\alpha | j_\beta b_\beta) t_{b_\beta a_\alpha}^{j_\beta k_\alpha} \quad (3.52)$$

An analogous equation holds for L_{ai}^β . The matrix elements of the response operator $R^\alpha(\mathbf{D}')$ are best evaluated in the AO basis and then transformed into the MO basis. The AO basis matrix elements are given by:

$$R^\alpha(\mathbf{D}')_{\mu\nu} = \sum_{\kappa\tau} 2D'_{\kappa\tau} (\mu\nu|\kappa\tau) - D'_{\kappa\tau}{}^\alpha [(\mu\kappa|\nu\tau) + (\nu\kappa|\mu\tau)] \\ + \sum_{\zeta} \int \left[\frac{\delta^2 f}{\delta \rho_\alpha \delta \zeta} \zeta(\mathbf{D}') (\phi_\mu \phi_\nu) + \left(2 \frac{\delta^2 f}{\delta \gamma_{\alpha\alpha} \delta \zeta} \vec{\nabla} \rho_{\mathbf{P}}^\alpha + \frac{\delta^2 f}{\delta \gamma_{\alpha\beta} \delta \zeta} \vec{\nabla} \rho_{\mathbf{P}}^\beta \right) \zeta(\mathbf{D}') \vec{\nabla} (\phi_\mu \phi_\nu) \right. \\ \left. + \left(2 \frac{\delta f}{\delta \gamma_{\alpha\alpha}} \vec{\nabla} \rho_{\mathbf{D}'}^\alpha + \frac{\delta f}{\delta \gamma_{\alpha\beta}} \vec{\nabla} \rho_{\mathbf{D}'}^\beta \right) \vec{\nabla} (\phi_\mu \phi_\nu) \right] d\mathbf{r} \quad (3.53)$$

where

$$\zeta(\mathbf{D}') = \rho_{\mathbf{D}'}^\alpha, \rho_{\mathbf{D}'}^\beta, \gamma_{\alpha\alpha}(\mathbf{D}'), \gamma_{\beta\beta}(\mathbf{D}'), \gamma_{\alpha\beta}(\mathbf{D}') \quad (3.54)$$

The ζ -gradient-parameters are evaluated as a mixture of PT2 difference densities and SCF densities. For example:

$$\gamma_{\alpha\alpha}(\mathbf{D}') = 2\vec{\nabla}\rho_{\mathbf{D}'}^\alpha \cdot \vec{\nabla}\rho_{\mathbf{P}'}^\alpha \quad (3.55)$$

With

$$\rho_{\mathbf{D}'}^\alpha(\mathbf{r}) = \sum_{\mu\nu} D'_{\mu\nu}{}^\alpha \phi_\mu(\mathbf{r}) \phi_\nu(\mathbf{r}) \quad (3.56)$$

$$\rho_{\mathbf{P}'}^\alpha(\mathbf{r}) = \sum_{\mu\nu} P_{\mu\nu}^\alpha \phi_\mu(\mathbf{r}) \phi_\nu(\mathbf{r}) \quad (3.57)$$

Having defined the Lagrangian, the following CP-SCF equations need to be solved for the elements of the “Z- vector”:

$$(\varepsilon_a^\sigma - \varepsilon_i^\sigma) Z_{ai}^\sigma + R^\sigma(\mathbf{Z})_{ai} = -L_{ai}^\sigma \quad (3.58)$$

The solution defines the occupied-virtual block of the relaxed difference density, which is given by:

$$\mathbf{D}^\sigma = \mathbf{D}'^\sigma + \mathbf{Z}^\sigma \quad (3.59)$$

For convenience, \mathbf{D}^σ is symmetrized since it will only be contracted with symmetric matrices afterwards. After having solved the Z-vector equations, all parts of the energy weighted difference density matrix can be readily calculated:

$$W_{ij}^{\alpha;\text{PT2}} = -\frac{1}{2}D_{ij}^\alpha (\varepsilon_i^\alpha + \varepsilon_j^\alpha) - \frac{1}{2}R(\mathbf{D})_{ij} - \sum_{k_\alpha a_\alpha b_\alpha} (i_\alpha a_\alpha | k_\alpha b_\alpha) t_{a_\alpha b_\alpha}^{j_\alpha k_\alpha} - \sum_{k_\beta a_\alpha b_\beta} (i_\alpha a_\alpha | k_\beta b_\beta) t_{b_\beta a_\alpha}^{k_\beta j_\alpha} \quad (3.60)$$

$$W_{ab}^{\alpha;\text{PT2}} = -\frac{1}{2}D_{ab}^\alpha (\varepsilon_a^\alpha + \varepsilon_b^\alpha) - \sum_{i_\alpha j_\alpha c_\alpha} (i_\alpha a_\alpha | j_\alpha c_\alpha) t_{b_\alpha c_\alpha}^{i_\alpha j_\alpha} - \sum_{i_\alpha j_\beta c_\beta} (i_\alpha a_\alpha | j_\beta c_\beta) t_{c_\beta b_\alpha}^{j_\beta i_\alpha} \quad (3.61)$$

$$W_{ai}^{\alpha;\text{PT2}} = -2 \sum_{j_\alpha k_\alpha b_\alpha} (k_\alpha i_\alpha | j_\alpha b_\alpha) t_{a_\alpha b_\alpha}^{k_\alpha j_\alpha} - 2 \sum_{j_\beta k_\alpha b_\beta} (k_\alpha i_\alpha | j_\beta b_\beta) t_{b_\beta a_\alpha}^{j_\beta k_\alpha} \quad (3.62)$$

$$W_{ia}^{\alpha;\text{PT2}} = -\varepsilon_i^\alpha Z_{ai}^\alpha \quad (3.63)$$

Once more, analogous equations hold for the spin-down case. With the relaxed difference density and energy weighted density matrices in hand, one can finally proceed to evaluate the gradient of the PT2 part as ($\mathbf{W}^{\text{PT2}} = \mathbf{W}^{\alpha;\text{PT2}} + \mathbf{W}^{\beta;\text{PT2}}$):

$$E_{\text{PT2}}^x = \langle \mathbf{D} \mathbf{h}^x \rangle + \langle \mathbf{W}^{\text{PT2}} \mathbf{S}^{(x)} \rangle + \sum_{\mu\nu\kappa\tau} \Gamma_{\mu\nu\kappa\tau}^{\text{PT2}} (\mu\nu | \kappa\tau)^{(x)} + \sum_{\sigma} \int \left\{ \frac{\delta f}{\delta \rho_\sigma(\mathbf{r})} \rho_\sigma^{(x)} + 2 \frac{\delta f}{\delta \gamma_{\sigma\sigma}} \vec{\nabla} \rho_\sigma^r \cdot \vec{\nabla} \rho_\sigma^{(x)} + \frac{\delta f}{\delta \gamma_{\sigma\sigma}} \vec{\nabla} \rho_{\sigma'}^r \cdot \vec{\nabla} \rho_\sigma^{(x)} \right\} d\mathbf{r} \quad (3.64)$$

The final derivative of eq. (3.31) is of course the sum $E_{\text{SCF}}^x + c_{PT} E_{\text{PT2}}^x$. Both derivatives should be evaluated simultaneously in the interest of computational efficiency.

Note that the exchange-correlation contributions to the gradient take a somewhat more involved form than might have been anticipated. In fact, from looking at the SCF XC-gradient (eq. (3.44)) it could have been speculated that the PT2 part of the gradient is of the same form but with $\rho_{\mathbf{P}}^{\sigma(x)}$ being replaced by \hat{H} , the relaxed PT2 difference density. This is, however, not the case. The underlying reason for the added complexity apparent in equation (3.64) is that the XC contributions to the PT2 gradient arise from the contraction of the relaxed PT2 difference density with the derivative of the SCF operator. Since the SCF operator already contains the first derivative of the XC potential and the PT2 energy is not stationary with respect to changes in the SCF density, a response type term arises which requires the evaluation of the second functional derivative of the XC-functional. Finally, as is well known from MP2 gradient theory, the effective two-particle density matrix contains a separable and a non-separable part:

$$\Gamma_{\mu\nu\kappa\tau}^{\text{PT2}} = D_{\mu\nu} P_{\kappa\tau} - D_{\mu\kappa}^\alpha P_{\nu\tau}^\alpha - D_{\mu\kappa}^\beta P_{\nu\tau}^\beta + \Gamma_{\mu\nu\kappa\tau}^{\text{NS}} \quad (3.65)$$

$$\Gamma_{\mu\nu\kappa\tau}^{\text{NS}} = \sum_{i_\alpha j_\alpha a_\alpha b_\alpha} c_{\mu i}^\alpha c_{\nu a}^\alpha c_{\kappa j}^\alpha c_{\tau b}^\alpha t_{a_\alpha b_\alpha}^{i_\alpha j_\alpha} + \sum_{i_\beta j_\beta a_\beta b_\beta} c_{\mu i}^\beta c_{\nu a}^\beta c_{\kappa j}^\beta c_{\tau b}^\beta t_{a_\beta b_\beta}^{i_\beta j_\beta} + 2 \sum_{i_\alpha j_\beta a_\alpha b_\beta} c_{\mu i}^\alpha c_{\nu a}^\alpha c_{\kappa j}^\beta c_{\tau b}^\beta t_{a_\alpha b_\beta}^{i_\alpha j_\beta} \quad (3.66)$$

Thus, the non-separable part is merely the back-transformation of the amplitudes from the MO to the AO basis. It is, however, important to symmetrize the two-particle density matrix in order to be able to exploit the full permutational symmetry of the AO derivative integrals.

3.9.14 Orbital Optimized MP2 Methods

OO-MP2 Theory

The MP2 energy can be regarded as being stationary with respect to the MP2 amplitudes, since they can be considered as having been optimized through the minimization of the Hylleraas functional:

$$E_{\text{MP2}} = \min_{\mathbf{t}} \left\{ 2 \langle \Psi_1 | \hat{H} | \Psi_0 \rangle + \langle \Psi_1 | \hat{H}_0 - E_0 | \Psi_1 \rangle \right\} \quad (3.67)$$

\hat{H} is the 0th order Hamiltonian as proposed by Møller and Plesset, Ψ_0 is the reference determinant, Ψ_1 is the first-order wave function and $E_0 = E_{\text{HF}} = \langle \Psi_{\text{HF}} | \hat{H} | \Psi_{\text{HF}} \rangle$ is the reference energy. The quantities \mathbf{t} collectively denote the MP2 amplitudes.

The fundamental idea of the OO-MP2 method is to not only minimize the MP2 energy with respect to the MP2 amplitudes, but to minimize the total energy additionally with respect to changes in the orbitals. Since the MP2 energy is not variational with respect to the MO coefficients, no orbital relaxation due to the correlation field is taken into account. If the reference determinant is poor, the low-order perturbative correction then becomes unreliable. This may be alleviated to a large extent by choosing better orbitals in the reference determinant. Numerical evidence for the correctness of this assumption will be presented below.

In order to allow for orbital relaxation, the Hylleraas functional can be regarded as a functional of the wavefunction amplitudes \mathbf{t} and the orbital rotation parameters \mathbf{R} that will be defined below. Through a suitable parameterization it becomes unnecessary to ensure orbital orthonormality through Lagrange multipliers. The functional that we minimize reads:

$$L\{\mathbf{t}, \mathbf{R}\} = E_0[\mathbf{R}] + 2 \langle \Psi_1 | \hat{H} | \Psi_0 \rangle + \langle \Psi_1 | \hat{H}_0 - E_0 | \Psi_1 \rangle \quad (3.68)$$

Ψ_0 is the reference determinant. However, it does no longer correspond to the Hartree-Fock (HF) determinant. Hence, the reference energy $E_0[\mathbf{R}] = \langle \Psi_0[\mathbf{R}] | \hat{H} | \Psi_0[\mathbf{R}] \rangle$ also changes during the variational process and is no longer stationary with respect to the HF MO coefficients. Obviously, $E_0[\mathbf{R}] \geq E_{\text{HF}}$ since the HF determinant is, by construction, the single determinant with the lowest expectation value of the full Hamiltonian.

The reference energy is given as:

$$E_0[\mathbf{R}] = \sum_i \langle i | h | i \rangle + \frac{1}{2} \sum_{ij} \langle ij || ij \rangle \quad (3.69)$$

The first-order wave function excluding single excitations is:

$$|\Psi_1\rangle = \frac{1}{4} \sum_{ijab} t_{ab}^{ij} |\Psi_{ij}^{ab}\rangle \quad (3.70)$$

A conceptually important point is that Brillouin's theorem [356] is no longer obeyed since the Fock matrix will contain off-diagonal blocks. Under these circumstances the first-order wavefunction would contain contributions from single excitations. Since the orbital optimization brings in all important effects of the singles we prefer to leave them out of the treatment. Any attempt to the contrary will destroy the convergence properties. We have nevertheless contemplated to include the single excitations perturbatively:

$$E_{\text{Singles}}^{(2)} = - \sum_{ia} \frac{|F_{ia}|^2}{\varepsilon_a - \varepsilon_i} \quad (3.71)$$

The perturbative nature of this correction would destroy the stationary nature of the total energy and is hence not desirable. Furthermore, results with inclusion of single excitation contributions represent no improvement to the results reported below. They will therefore not be documented below and henceforth be omitted from the OO-MP2 method by default.

The explicit form of the orbital-optimized MP2 Hylleraas functional employing the RI approximation (OO-RI-MP2) becomes:

$$L_{\infty}[\mathbf{t}, \mathbf{R}] = \sum_i \langle i | \hat{h} | i \rangle + \frac{1}{2} \sum_{ij} \langle ij || ij \rangle + \sum_{iaP} (ia|P) \Gamma_{ia}^{'P} + \sum_{ij} D_{ij} F_{ij} + \sum_{ab} D_{ab} F_{ab} \quad (3.72)$$

with:

$$\Gamma'_{ia} = \sum_Q V_{PQ}^{-1} \sum_{jb} (Q|jb) t_{ab}^{ij} \quad (3.73)$$

$$(ia|P) = \int \int \psi_i(\mathbf{r}_1) \psi_a(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \eta_P(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (3.74)$$

$$(P|Q) = \int \int \eta_P(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \eta_Q(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (3.75)$$

Here, $\{\psi\}$ is the set of orthonormal molecular orbitals and $\{\eta\}$ denotes the auxiliary basis set. F_{pq} denotes a Fock matrix element:

$$F_{pq} = \langle p | \hat{h} | q \rangle + \sum_k \langle pk || qk \rangle \quad (3.76)$$

and it is insisted that the orbitals diagonalize the occupied and virtual subspaces, respectively:

$$\begin{aligned} F_{ij} &= \delta_{ij} F_{ii} = \delta_{ij} \varepsilon_i \\ F_{ab} &= \delta_{ab} F_{aa} = \delta_{ab} \varepsilon_a \end{aligned} \quad (3.77)$$

The MP2 like density blocks are,

$$\begin{aligned} D_{ij} &= -\frac{1}{2} \sum_{kab} t_{ab}^{ik} t_{ab}^{jk} \\ D_{ab} &= \frac{1}{2} \sum_{ijc} t_{ac}^{ij} t_{bc}^{ij} \end{aligned} \quad (3.78)$$

where the MP2 amplitudes in the case of a block diagonal Fock matrix are obtained through the condition $\frac{\partial L_\infty}{\partial t_{ab}^{ij}} = 0$:

$$t_{ab}^{ij} = -\frac{\langle ij || ab \rangle}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \quad (3.79)$$

The orbital changes are parameterized by an anti-Hermitian matrix \mathbf{R} and an exponential Ansatz,

$$\begin{aligned} \mathbf{c}^{\text{new}} &= \mathbf{c}^{\text{old}} \exp(\mathbf{R}) \\ \mathbf{R} &= \begin{pmatrix} 0 & \mathbf{R}_{ia} \\ -\mathbf{R}_{ia} & 0 \end{pmatrix} \end{aligned} \quad (3.80)$$

The orbitals changes to second order are,

$$\begin{aligned} \exp(\mathbf{R}) |i\rangle &= |i\rangle + \sum_a \mathbf{R}_{ai} |a\rangle - \frac{1}{2} \sum_{jb} \mathbf{R}_{bi} \mathbf{R}_{bj} |j\rangle + \dots \\ \exp(\mathbf{R}) |a\rangle &= |a\rangle - \sum_i \mathbf{R}_{ai} |i\rangle - \frac{1}{2} \sum_{jb} \mathbf{R}_{aj} \mathbf{R}_{bj} |b\rangle + \dots \end{aligned} \quad (3.81)$$

Through this Ansatz it is ensured that the orbitals remain orthonormal and no Lagrangian multipliers need to be introduced. The first-order expansion of the Fock operator due to the orbital rotations are:

$$F_{pq}[R] = F_{pq}[0] + R_{pq}^{(1)} + \sum_r R_{rp} F_{rq}[0] + R_{rq} F_{pr}[0] \quad (3.82)$$

$$R_{pq}^{(1)} = \sum_{kc} R_{ck} \{ \langle pc || qk \rangle + \langle pk || qc \rangle \} \quad (3.83)$$

The first-order energy change becomes $(h_{pq} \equiv \langle p | \hat{h} | q \rangle, g_{pqrs} \equiv \langle pq || rs \rangle)$:

$$\begin{aligned} L_\infty[\mathbf{t}, \mathbf{R}] &= \sum_{ic} R_{ci} (h_{ci} + h_{ic}) + \frac{1}{2} \sum_{ijc} R_{ci} (g_{ciji} + g_{ijcj}) + R_{cj} (g_{icij} + g_{ijic}) \\ &\quad + 2 \sum_{iacP} R_{ci} (ac|P) \Gamma'_{ia}{}^P - 2 \sum_{ikaP} R_{ak} (ik|P) \Gamma'_{ia}{}^P \\ &\quad - \sum_{ij} D_{ij} \left(R_{ij}^{(1)} + \sum_c (R_{ci} F_{cj} + R_{cj} F_{ic}) \right) \\ &\quad + \sum_{ab} D_{ab} \left(R_{ab}^{(1)} - \sum_k (R_{ak} F_{kb} + R_{bk} F_{ak}) \right) \end{aligned} \quad (3.84)$$

The condition for the energy functional to be stationary with respect to the orbital rotations $\left(\frac{\partial L_\infty[\mathbf{t}, \mathbf{R}]}{\partial R_{ai}} = 0 \right)$, yields the expression for the orbital gradient and hence the expression for the OO-RI-MP2 Lagrangian.

$$\frac{\partial L_\infty[\mathbf{t}, \mathbf{R}]}{\partial R_{ai}} \equiv g_{ai} = 2F_{ai} + 2 \sum_j D_{ij} F_{aj} - 2 \sum_b D_{ab} F_{ib} + R^{(1)}(\mathbf{D})_{ai} \quad (3.85)$$

$$+2 \sum_{cP} (ac|P) \Gamma'_{ia} - 2 \sum_{kP} (ik|P) \Gamma'_{ia}$$

The goal of the orbital optimization process is to bring this gradient to zero. There are obviously many ways to achieve this. In our experience, the following simple procedure is essentially satisfactory. We first build a matrix **B** in the current MO basis with the following structure:

$$\begin{aligned} \mathbf{B}_{ij} &= \delta_{ij} \mathbf{F}_{ii} \\ \mathbf{B}_{ab} &= \delta_{ab} (\mathbf{F}_{aa} + \Delta) \\ \mathbf{B}_{ai} &= \mathbf{B}_{ia} = \mathbf{g}_{ai} \end{aligned} \quad (3.86)$$

where Δ is a level shift parameter. The occupied/occupied and virtual/virtual blocks of this matrix are arbitrary but their definition has a bearing on the convergence properties of the method. The orbital energies of the block diagonalized Fock matrix appear to be a logical choice. If the gradient is zero, the **B**-matrix is diagonal. Hence one obtains an improved set of orbitals by diagonalizing **B**.

In order to accelerate convergence a standard DIIS scheme is used. [333, 357] However, in order to carry out the DIIS extrapolation of the **B**-matrix it is essential that a common basis is used that does not change from iteration to iteration. Since the **B**-matrix itself is defined in the molecular orbitals of the current iteration we choose as a common set of orthonormal orbitals the MOs of the HF calculation. The extrapolation is carried out in this basis and the extrapolated **B**-matrix is transformed back to the current set of MOs prior to diagonalization. Obviously, the same strategy can be used for orbital optimization in any method for which an orbital gradient is available.

For well behaved cases this simple scheme converges in 5-10 iterations. Transition metals and more complicated molecules may require up to 20 iterations and level shifting in order to achieve convergence.

Upon convergence the sum of the matrix **D** and the density of the reference determinant $P_{\mu\nu} = \sum_i c_{\mu i} c_{\nu i}$ form the true one-particle density matrix of the OO-MP2 approach that can be used for property or gradient calculations.

OO-MP2 - Tips, Usage and Examples

By making the Hylleraas functional stationary with respect to the orbital rotations one obtains the orbital-optimized MP2 method that is implemented in ORCA in combination with the RI approximation (OO-RI-MP2). One obtains from these calculations orbitals that are adjusted to the dynamic correlation field at the level of second order many-body perturbation theory. Also, the total energy of the OO-RI-MP2 method is lower than that of the RI-MP2 method itself. One might think of this method as a special form of multiconfigurational SCF theory except for the fact that the Hamiltonian is divided into a 0th order term and a perturbation.

The main benefit of the OO-RI-MP2 method is that it “repairs” the poor Hartree–Fock orbitals to some extent which should be particularly beneficial for systems which suffer from the imbalance in the Hartree–Fock treatment of the Coulomb and the Exchange hole. Based on the experience gained so far, the OO-RI-MP2 method is no better than RI-MP2 itself for the thermochemistry of organic molecules. However, for reactions barriers and radicals the benefits of OO-MP2 over MP2 are substantial. This is particularly true with respect to the spin-component scaled variant of OO-RI-MP2 that is OO-RI-SCS-MP2. Furthermore, the OO-RI-MP2 method substantially reduces the spin contamination in UHF calculations on radicals.

Since every iteration of the OO-MP2 method is as expensive as a RI-MP2 relaxed density calculation, the computational cost is much higher than for RI-MP2 itself. One should estimate about a factor of 10 increase in computational time with respect to the RI-MP2 time of a normal calculation. This may still be feasible for calculations in the range of 1000–2000 basis functions (the upper limit, however, implies very significant computational costs). A full assessment of the orbital optimized MP2 method has been published.[358]

OO-RI-MP2 is triggered either with ! OO-RI-MP2 or ! OO-RI-SCS-MP2 (with spin component scaling) in the simple input line or by `OrbOpt true` in the `%mp2` block. The method comes with the following new variables:

```
%mp2 OrbOpt true    # turns on the orbital optimization
      CalcS2 false  # calculate the S**2 expectation value
                        # in spin-unrestricted calculations
      MaxOrbIter 64  # Max. number of iterations
      MP2Shift 0.1   # Level shift for the procedure
      end
```


The solver is a simple DIIS type scheme with additional level shifting. We have found that it is not really beneficial to first converge the Hartree-Fock equations. Thus it is sensible to additionally use the keyword `! noiter` in order to turn off the standard Hartree-Fock SCF process before entering the orbital optimizations.

The OO-RI-MP2 method is implemented for RHF and UHF reference wavefunctions. Analytic gradients are available.

The density does not need to be requested separately in OO-RI-MP2 calculations because it is automatically calculated. Also, there is no distinction between relaxed and unrelaxed densities because the OO-RI-MP2 energy is fully stationary with respect to all wavefunction parameters and hence the unrelaxed and relaxed densities coincide.

3.9.15 Regularized MP2 and RI-MP2

Regularized MP2 is a variant of second-order Moller-Plesset theory (MP2) introduced by J. Shee, M. Loipersberger, A. Rettig, J. Lee, and M. Head-Gordon [359] that aims to improve its accuracy for systems with π -driven dispersion interactions and dative bonds in transition metal complexes. The approach achieves this by introducing a single-parameter, energy-gap dependent regularization that dampens overestimated pairwise additive contributions, thus renormalizing first-order amplitudes to empirically mimic higher-order correlations.

For this, the standard MP2 energy and thus the standard algorithms are modified. For the σ ($p = 1$) and σ^2 ($p = 2$) regularization, the energy is modified according to

$$E_{\sigma^p\text{-MP2}} = -\frac{1}{4} \sum_{ijab} \frac{|\langle ij||ab \rangle|^2}{\Delta_{ij}^{ab}} (1 - e^{-\sigma(\Delta_{ij}^{ab})^p})$$

which corresponds to regularizing the first-order amplitudes. For the κ regularization, the MP2 energy is modified according to

$$E_{\kappa\text{-MP2}} = -\frac{1}{4} \sum_{ijab} \frac{|\langle ij||ab \rangle|^2}{\Delta_{ij}^{ab}} \left(1 - e^{-\kappa(\Delta_{ij}^{ab})}\right)^2$$

which corresponds to regularizing the first-order amplitudes and the exchange integrals.

Regularized MP2 is available for standard MP2 in “memory mode” (Q1Opt>0) and RI-MP2 (RIJDX, RIJCOSX, RIJK).

The usage of regularized MP2 is controlled by the `DoRegMP2` keyword, the type of regularization can be specified by setting the `RegMP2Type` parameter to 0 for κ , 1 for σ , or 2 for σ^2 . The value of the regularizers can be specified by `RegMP2Kappa` and `RegMP2Sigma` respectively.

```
%mp2
  DoRegMP2      true      # required
  RegMP2Type    0         # kappa regularizer
                  1         # sigma regularizer
                  2         # sigma-squared regularizer
  RegMP2Kappa   1.1       # kappa value
  RegMP2Sigma   0.5       # sigma value
end
```

It is important to note that only single point energies are available and tested for regularized MP2. Density, Gradient, and Hessian calculations are not yet supported.

3.9.16 Keywords

Table 3.34: Simple input keywords selecting MP2 variants.

Keyword	Description
MP2	Canonical MP2
RI-MP2	RI-MP2 (Aliases: RIMP2, MP2RI)
DLPNO-MP2	DLPNO-MP2
SCS-MP2	Spin-component-scaled canonical MP2
SCS-MP2	Spin-opposite-scaled canonical MP2
RI-SCS-MP2	Spin-component-scaled RI-MP2 (Aliases: SCS-RI-MP2, SCS-RIMP2)
RI-SOS-MP2	Spin-opposite-scaled RI-MP2 (Aliases: SOS-RI-MP2, SOS-RIMP2)
DLPNO-SCS-MP2	Spin-component-scaled DLPNO-MP2 (Alias: SCS-DLPNO-MP2)
DLPNO-SOS-MP2	Spin-opposite-scaled DLPNO-MP2 (Alias: SOS-DLPNO-MP2)
RI-SCS-MP2	Spin-opposite-scaled RI-MP2 (Aliases: SOS-RI-MP2, SOS-RIMP2)
OO-RI-MP2	Orbital-optimized RI-MP2 (Aliases: OO-RIMP2, OO-MP2RI)
OO-RI-SCS-MP2	Orbital-optimized spin-component-scaled RI-MP2 (Aliases: OO-SCS-RI-MP2, OO-SCS-RIMP2)
OO-RI-SOS-MP2	Orbital-optimized spin-opposite-scaled RI-MP2 (Aliases: OO-SOS-RI-MP2, OO-SOS-RIMP2)
F12-MP2	MP2 with F12 correction (Alias: MP2-F12)
F12-RI-MP2	RI-MP2 with F12 correction (Aliases: MP2-F12-RI, RI-MP2-F12)
F12/D-RI-MP2	RI-MP2 with F12 correction employing the lower-cost D approximation (Aliases: MP2-F12/D-RI, RI-MP2-F12/D, F12D-RI-MP2, MP2-F12D-RI, RI-MP2-F12D)
F12-DLPNO-MP2	DLPNO-MP2 with F12 correction (Alias: DLPNO-MP2-F12)
F12/D-DLPNO-MP2	DLPNO-MP2 with F12 correction employing the lower-cost D approximation (Aliases: DLPNO-MP2-F12/D, F12D-DLPNO-MP2, DLPNO-MP2-F12D)
LoosePNO	Selects loose thresholds for DLPNO-MP2
NormalPNO	(Default) Selects default thresholds for DLPNO-MP2
TightPNO	Selects tight thresholds for DLPNO-MP2

3.10 Coupled Cluster and CI Theories (MDCI)

Aside from the *second-order Møller-Plesset theory (MP2)*, ORCA features a variety of single-reference correlation methods for single point energies (restricted to a RHF or RKS determinant in the closed-shell case and a UHF or UKS determinant in the open-shell case; quasi-restricted orbitals (QROs)[360] are also supported in the open-shell case). These methods are all fairly expensive but maybe be used in order to obtain accurate results in the case that the reference determinant is a good starting point for the expansion of the many-body wavefunction. The methods are implemented in the `orca_mdci` module, which is the main subject of this section. ‘MDCI’ is an abbreviations for “matrix driven configuration interaction”. The term is rather technical to emphasize that if one wants to implement these methods (CCSD, QCISD etc.) efficiently, one needs to write them in terms of matrix operations, which pretty much every computer can drive at peak performance.

It should be noted that in recent years, a number of highly correlated single reference methods and their properties were added to ORCA using *automatic code generation and reside in the `orca_autoci` module*, which follows similar design principles.[361, 362] Let us first briefly describe the theoretical background of the methods, that we have implemented in ORCA.

3.10.1 Theory

We start from the full CI hierarchy in which the wavefunction is expanded as:

$$|\Psi\rangle = |0\rangle + |S\rangle + |D\rangle + |T\rangle + |Q\rangle + \dots \quad (3.87)$$

where $|0\rangle$ is a single-determinant reference and S, D, T, Q, ... denote the single, double, triple quadruple and higher excitations relative to this determinant at the spin-orbital level. As usual, labels i, j, k, l refer to occupied orbitals in $|0\rangle$, a, b, c, d to unoccupied MOs and p, q, r, s to general MOs. The action of the second quantized excitation operators $a_i^a = a_a^\dagger a_i$ on $|0\rangle$ lead to excited determinants $|\Phi_i^a\rangle$ that enter $|\Psi\rangle$ with coefficients C_a^i . The variational equations are:

$$\langle \Phi_i^a | H - E_0 | 0 + S + D \rangle = E_C C_a^i - \langle \Phi_i^a | H - E_0 | T \rangle \quad (3.88)$$

$$\langle \Phi_{ij}^{ab} | H - E_0 | 0 + S + D \rangle = E_C C_{ab}^{ij} - \langle \Phi_{ij}^{ab} | H - E_0 | T + Q \rangle \quad (3.89)$$

Further equations coupling triples with singles through pentuples etc.

The total energy is the sum of the reference energy $E_0 = \langle 0 | H | 0 \rangle$ and the correlation energy

$$E_C = \langle 0 | H | S + D \rangle \quad (3.90)$$

which requires the exact singles- and doubles amplitudes to be known. In order to truncate the series to singles- and doubles one may either neglect the terms containing the higher excitations on the right hand side (leading to CISD) or approximate their effect thereby losing the variational character of the CI method (CCSD, QCISD and CEPA methods). Defining the one- and two-body excitation operators as $\hat{C}_1 = \sum_{ia} C_a^i a_i^a$, $\hat{C}_2 = \frac{1}{4} \sum_{ijab} C_{ab}^{ij} a_{ab}^{ij}$ one can proceed to approximate the triples and quadruples by the disconnected terms:

$$|T\rangle = \hat{C}_1 \hat{C}_2 |0\rangle \quad (3.91)$$

$$|Q\rangle = \frac{1}{2} \hat{C}_2^2 |0\rangle \quad (3.92)$$

As is well known, the CCSD equations contain many more disconnected contributions arising from the various powers of the \hat{C}_1 operator (if one would stick to CC logics one would usually label the cluster amplitudes with $t_a^i, t_{ab}^{ij}, \dots$ and the n -body cluster operators with \hat{T}_n ; we take a CI point of view here). In order to obtain the CEPA type equations from ((3.88)-(3.92)), it is most transparent to relabel the singles and doubles excitations with a compound label P for the internal indices (i) or (ij) and x for (a) or (ab). Then, the approximations are as follows:

$$\frac{1}{2} \langle \Phi_P^x | (H - E_0) \hat{C}_2^2 | 0 \rangle = \frac{1}{2} \sum_{Q R y z} C_y^Q C_z^R \langle \Phi_P^x | H - E_0 | \Phi_{QR}^{yz} \rangle \quad (3.93)$$

$$\approx C_x^P \sum_{Qy} C_y^Q \langle \Phi_P^x | H | \Phi_{PQ}^{xy} \rangle \quad (3.94)$$

$$= C_x^P \sum_{Qy} C_y^Q \langle 0 | H | \Phi_Q^y \rangle - C_x^P \sum_{Qy \cup Px} C_y^Q \langle 0 | H | \Phi_Q^y \rangle \quad (3.95)$$

$$\approx C_x^P \left(E_C - \sum_{Q \cup P} \varepsilon_Q \right) \quad (3.96)$$

Here the second line contains the approximation that only the terms in which either Qy or Rz are equal to Px are kept (this destroys the unitary invariance) and the fourth line contains the approximation that only “exclusion principle violating” (EPV) terms of internal labels are considered. The notation $Qy \cup Px$ means “ Qy joint with Px ” (containing common orbital indices) and ε_Q is the pair correlation energy. The EPV terms must be subtracted from the correlation energy since they arise from double excitations that are impossible due to the fact that an excitation out of an occupied or into an empty orbital of the reference determinant has already been performed. Inserting eq. (3.96) into eq. (3.89) $C_x^P E_C$ cancels and effectively is replaced by the “partial correlation energy” $\sum_{Q \cup P} \varepsilon_Q$.

The resulting equations thus have the appearance of a diagonally shifted (“dressed”) CISD equation

$\langle \Phi_P^x | H - E_0 + \Delta | 0 + S + D \rangle = 0$. If the second approximation mentioned above is avoided Malrieu’s (SC)²-CISD arises. [307, 308, 309, 310, 363] Otherwise, one obtains CEPA/3 with the shift:

$$-\Delta_{ab}^{ij} = \sum_k (\varepsilon_{ik} + \varepsilon_{jk}) - \varepsilon_{ij} \quad (3.97)$$

CEPA/2 is obtained by $-\Delta_{ab}^{ij} = \varepsilon_{ij}$ and CEPA/1 is the average of the CEPA/2 and CEPA/3. As mentioned by Ahlrichs, [364] no consensus appears to exist in the literature for the appropriate shift on the single excitations. If one proceeds straightforwardly in the same way as above, one obtains:

$$\langle \Phi_i^a | (H - E_0) \hat{C}_1 \hat{C}_2 | 0 \rangle \approx C_a^i \left(E_C - 2 \sum_k \varepsilon_{ik} \right) \quad (3.98)$$

as the appropriate effect of the disconnected triples on the singles. It has been assumed here that only the singles $|\Phi_i^a\rangle$ in \hat{C}_1 contribute to the shift. If $|0\rangle$ is a HF determinant, the effect of the disconnected triples in the doubles projection vanishes under the same CEPA approximations owing to Brillouin’s theorem. Averaged CEPA models are

derived by assuming that all pair correlation energies are equal (except $\varepsilon_{ii} = 0$). As previously discussed by Gdanitz [365], the averaging of CEPA/1 yields $\frac{2}{n}E_C$ and CEPA/3 $E_C \frac{4n-6}{n(n-1)}$ where n is the number of correlated electrons. These happen to be the shifts used for the averaged coupled-pair functional (ACPF [366]) and averaged quadratic coupled-cluster (AQCC [367]) methods respectively. However, averaging the singles shift of eq. (3.98) gives $\frac{4}{n}E_C$. The latter is also the leading term in the expansion of the AQCC shift for large n . In view of the instability of ACPF in certain situations, Gdanitz has proposed to use the AQCC shift for the singles and the original ACPF shift for the doubles and called his new method ACPF/2 [366]. Based on what has been argued above, we feel that it would be most consistent with the ACPF approach to simply use $\frac{4}{n}E_C$ as the appropriate singles shift. We refer to this as NACPF.

It is readily demonstrated that the averaged models may be obtained by a variation of the modified correlation energy functional:

$$E_C = \frac{\langle 0 + S + D | H - E_0 | 0 + S + D \rangle}{1 + g_S \langle S | S \rangle + g_D \langle D | D \rangle} \quad (3.99)$$

with g_S and g_D being the statistical factors $\frac{4}{n}, \frac{2}{n}, \frac{4n-6}{n(n-1)}$, as appropriate for the given method. Thus, unlike the CEPA models, the averaged models fulfill a stationarity principle and are unitarily invariant. However, if one thinks about localized internal MOs, it appears evident that the approximation of equal pair energies must be one of rather limited validity and that a more detailed treatment of the electron pairs is warranted. Maintaining a stationarity principle while providing a treatment of the pairs that closely resembles that of the CEPA methods was achieved by Ahlrichs and co-workers in an ingenious way with the development of the CPF method [368]. In this method, the correlation energy functional is written as:

$$E_C = 2 \sum_{Px} \frac{\langle \Phi_p^x | H | 0 \rangle}{N_P} + \sum_{PQxy} \frac{\langle \Phi_p^x | H - E_0 | \Phi_Q^y \rangle}{\sqrt{N_P N_Q}} \quad (3.100)$$

with

$$N_P = 1 + \sum_Q T_{PQ} \sum_y (C_y^Q)^2 \quad (3.101)$$

The topological matrix for pairs $P = (ij)$ and $Q = (kl)$ is chosen as: [115]

$$T_{PQ} = \frac{\delta_{ik} + \delta_{il}}{2n_i} + \frac{\delta_{jk} + \delta_{jl}}{2n_j} \quad (3.102)$$

with n_i being the number of electrons in orbital i in the reference determinant. The singles out of orbital i are formally equated with $P = (ii)$. At the spin-orbital level, $n_i = 1$, for closed shells $n_i = 2$. Using the same topological matrix in $\Delta_P = \sum_Q T_{PQ} \varepsilon_Q$ one recovers the CEPA/1 shifts for the doubles in eq. (3.98). It is straightforward to obtain the CPF equivalents of the other CEPA models by adjusting the T_{PQ} matrix appropriately. In our program, we have done so and we refer below to these methods as CPF/1, CPF/2 and CPF/3 in analogy to the CEPA models (CPF/1 \equiv CPF). In fact, as discussed by Ahlrichs and co-workers, variation of the CPF-functional leads to equations that very closely resemble the CEPA equation and can be readily implemented along the same lines as a simple modification of a CISD program. Ahlrichs *et al.* argued that the energies of CEPA/1 and CPF/1 should be very close. We have independently confirmed that in the majority of cases, the total energies predicted by the two methods differ by less than 0.1 mEh.

An alternative to the CPF approach which is also based on variational optimization of an energy functional is the VCEPA method [369]. The equations resulting from application of the variational principle to the VCEPA functional are even closer to the CEPA equations than for CPF so that the resulting energies are practically indistinguishable from the corresponding CEPA values. The VCEPA variants are referred to as VCEPA/1, VCEPA/2, and VCEPA/3 in analogy to CEPA and CPF. A strictly size extensive energy functional (SEOI) which is invariant with respect to unitary transformations within the occupied and virtual orbital subspaces is also available [370] (an open-shell version is not implemented yet).

Again, a somewhat critical point concerns the single excitations. They do not account for a large fraction of the correlation energy. However, large coefficients of the single excitations lead to instability and deterioration of the results. Secondly, linear response properties are highly dependent on the effective energies of the singles and their balanced treatment is therefore important. Since the CEPA and CPF methods amount to shifting down the diagonal energies of the singles and doubles, instabilities are expected if the effective energy of an excitation approaches the

reference energy of even falls below it. In the CPF method this would show up as denominators N_P that are too small. The argument that the CPF denominators are too small has led Chong and Langhoff to the proposal of the MCPF method which uses a slightly more elaborate averaging than $(N_P N_Q)^{1/2}$ [371].¹ However, their modification was solely based on numerical arguments rather than physical or mathematical reasoning. In the light of Eq. (3.98) and the performance of the NACPF, it appears to us that for the singles one should use twice the T_{PQ} proposed by Ahlrichs and co-workers. The topological matrix T_{PQ} is modified in the following way for the (very slightly) modified method to which we refer to as NCPF/1:

$$T_{ij,kl} = \frac{\delta_{ik} + \delta_{il}}{2n_i} + \frac{\delta_{jk} + \delta_{jl}}{2n_j} \quad (3.103)$$

$$T_{ij,k} = 0 \quad (3.104)$$

$$T_{i,kl} = 2 \frac{\delta_{ik} + \delta_{il}}{n_i} \quad (3.105)$$

$$T_{i,k} = 0 \quad (3.106)$$

(note that $T_{PQ} \neq T_{QP}$ for this choice). Thus, the effect of the singles on the doubles is set to zero based on the analysis of the CEPA approximations and the effect of the singles on the singles is also set to zero. This is a sensible choice since the product of two single excitations is a double excitation which is already included in the SD space and thus none of them can belong to the outer space. It is straightforward to adapt this reasoning about the single excitations to the CEPA versions as well as to NCPF/2 and NCPF/3.

The aforementioned ambiguities arising from the use of single excitations in coupled-pair methods can be avoided by using correlation-adapted orbitals instead of Hartree-Fock orbitals thus eliminating the single excitations. There are two alternatives: (a) Brueckner orbitals and (b) optimized orbitals obtained from the variational optimization of the electronic energy with respect to the orbitals. Both approaches have already been used for the coupled-cluster doubles (CCD) method [372, 373] and later been extended to coupled-pair methods [374]. In the case of CCD, orbital optimization requires the solution of so-called Λ (or Z vector) equations [375]. There is, however, a cheaper alternative approximating the Z vector by a simple analytical formula [376].

Furthermore, the parametrized coupled-cluster (pCCSD) method of Huntington and Nooijen [377], which combines the accuracy of coupled-pair type methods for (usually superior to CCSD, at least for energies and energy differences) with the higher stability of the coupled-cluster methods, is an attractive alternative. Comprehensive numerical tests [378] indicate that particularly pCCSD(-1,1,1) (or pCCSD/1a) and pCCSD (-1.5,1,1) (or pCCSD/2a) have great potential for accurate computational thermochemistry. These methods can be employed by adding the “simple” keywords pCCSD/1a or pCCSD/2a to the first line of input. As mentioned in section *Local correlation (DLPNO)*, the DLPNO variants of the pCCSD/1a method is also available for RHF and UHF references via the simple keyword DLPNO-pCCSD/1a.

Closed-Shell Equations

Proceeding from spin-orbitals to the spatial orbitals of a closed-shell determinant leads to the actual working equations of this work. Saebo, Meyer and Pulay have exploited the generator state formalism to arrive at a set of highly efficient equations for the CISD problem [379]. A similar set of matrix formulated equations for the CCSD and QCISD cases has been discussed by Werner and co-workers [380] and the MOLPRO implementation is widely recognized to be particularly efficient. Equivalent explicit equations for the CISD and CCSD methods were published by Scuseria *et al.* [381]² The doubles equations for the residual “vector” σ are ($i \leq j$, all a, b):

$$\begin{aligned} \sigma_{ab}^{ij} = & K_{ab}^{ij} + K(\mathbf{C}^{ij})_{ab} + \{\mathbf{F}^V \mathbf{C}^{ij} + \mathbf{C}^{ij} \mathbf{F}^V\}_{ab} - \sum_k \{F_{jk} C_{ab}^{ik} + F_{ik} C_{ab}^{kj}\} + \sum_{kl} K_{kl}^{ij} C_{ab}^{kl} \\ & + \sum_k \{ (2\mathbf{C}^{ik} - \mathbf{C}^{ik+}) (\mathbf{K}^{kj} - \frac{1}{2} \mathbf{J}^{kj}) + (\mathbf{K}^{ik} - \frac{1}{2} \mathbf{J}^{ik}) (2\mathbf{C}^{kj} - \mathbf{C}^{kj+}) \}_{ab} \\ & - \sum_k \{ \frac{1}{2} \mathbf{C}^{ik+} \mathbf{J}^{jk+} + \frac{1}{2} \mathbf{J}^{ik} \mathbf{C}^{kj+} + \mathbf{J}^{jk} \mathbf{C}^{ik} + \mathbf{C}^{kj} \mathbf{J}^{ik+} \}_{ab} \\ & + C_a^i F_b^j + C_b^j F_a^i - \sum_k \{ K_{ka}^{ji} C_b^k + K_{kb}^{ij} C_a^k \} + \{ \mathbf{K}^{ia} \mathbf{C}^j + \mathbf{K}^{ja} \mathbf{C}^i \}_b \\ & - \Delta^{ij} C_{ab}^{ij} \end{aligned} \quad (3.107)$$

¹ This method – although it has been rather extensively used in the past – is not implemented in ORCA. We recommend to use our NCPF/1 instead.

² Our coupled-cluster implementation is largely based on this nice paper. The equations there have been extensively verified to be correct.

The singles equations are:

$$\begin{aligned} \sigma_a^i = & F_a^i + \{\mathbf{F}^V \mathbf{C}^i\}_a - \sum_j F_{ij} C_a^j - \sum_{jkb} (2K_{jb}^{ik} - J_{jb}^{ik}) C_{ba}^{kj} \\ & + \sum_j \{ (2\mathbf{K}^{ij} - \mathbf{J}^{ij}) \mathbf{C}^j + \mathbf{F}^j (2\mathbf{C}^{ij+} - \mathbf{C}^{ij}) + \langle (2\mathbf{K}^{ia} - \mathbf{K}^{ia+}) \mathbf{C}^{ij+} \rangle_a \\ & - \Delta^i C_a^i \end{aligned} \quad (3.108)$$

The following definitions apply:

$$K(\mathbf{C}^{ij})_{ab} = \sum_{cd} (ac|bd) C_{cd}^{ij} \quad (3.109)$$

$$K_{rs}^{pq} = (pr|qs) \quad (3.110)$$

$$J_{rs}^{pq} = (pq|rs) \quad (3.111)$$

$$\langle \mathbf{AB} \rangle = \sum_{pq} A_{pq} B_{qp} \quad (3.112)$$

The two-electron integrals are written in (11|12) notation and \mathbf{F} is the closed-shell Fock operator with \mathbf{F}^V being its virtual sub-block. We do not assume the validity of Brillouin's theorem. The amplitudes C_a^i , C_{ab}^{ij} have been collected in vectors \mathbf{C}^i and matrices \mathbf{C}^{ij} wherever appropriate. The shifts Δ^i and Δ^{ij} are dependent on the method used and are defined in Table 3.35 for each method implemented in ORCA.

Table 3.35: Summary of the diagonal shifts used in various singles- and doubles methods discussed in this chapter. The quantities ε_i and ε_{ij} are the correlation energy increments brought about by the single- and the double excitations respectively. The partial denominators for the CPF type methods N_i and N_{ij} are specified in eq. (3.101).

Method	Doubles Shift	Singles Shift
CISD	E_C	E_C
CEPA/0	0	0
CEPA/1	$\frac{1}{2}(\varepsilon_i + \varepsilon_j) + \frac{1}{4} \sum_k (\varepsilon_{ik} + \varepsilon_{jk})$	$\frac{1}{2}\varepsilon_{ii} + \frac{1}{2} \sum_k \varepsilon_{ik}$
CEPA/2	$\delta_{ij}\varepsilon_i + \varepsilon_{ij}$	$\varepsilon_i + \varepsilon_{ii}$
CEPA/3	$(\varepsilon_i + \varepsilon_j) - \delta_{ij}\varepsilon_i - \varepsilon_{ij} + \frac{1}{2} \sum_k (\varepsilon_{ik} + \varepsilon_{jk})$	$\varepsilon_i + \sum_k \varepsilon_{ik}$
NCEPA/1	$\frac{1}{4} \sum_k (\varepsilon_{ik} + \varepsilon_{jk})$	$\varepsilon_{ii} + \sum_k \varepsilon_{ik}$
NCEPA/2	ε_{ij}	$2\varepsilon_{ii}$
NCEPA/3	$-\varepsilon_{ij} + \frac{1}{2} \sum_k (\varepsilon_{ik} + \varepsilon_{jk})$	$2 \sum_k \varepsilon_{ik}$
CPF/1	$N_{ij} \left\{ \frac{1}{2} \left(\frac{\varepsilon_i}{N_i} + \frac{\varepsilon_j}{N_j} \right) + \frac{1}{4} \sum_k \left(\frac{\varepsilon_{ik}}{N_{ik}} + \frac{\varepsilon_{jk}}{N_{jk}} \right) \right\}$	$N_i \left\{ \frac{1}{2} \frac{\varepsilon_{ii}}{N_{ii}} + \frac{1}{2} \sum_k \frac{\varepsilon_{ik}}{N_{ik}} \right\}$
CPF/2	$N_{ij} \left\{ \delta_{ij} \frac{\varepsilon_i}{N_i} + \frac{\varepsilon_{ij}}{N_{ij}} \right\}$	$N_i \left\{ \frac{\varepsilon_i}{N_i} + \frac{\varepsilon_{ii}}{N_{ii}} \right\}$
CPF/3	$N_{ij} \left\{ \frac{\varepsilon_i}{N_i} (1 - \delta_{ij}) + \frac{\varepsilon_j}{N_j} - \frac{\varepsilon_{ij}}{N_{ij}} + \frac{1}{2} \sum_k \left(\frac{\varepsilon_{ik}}{N_{ik}} + \frac{\varepsilon_{jk}}{N_{jk}} \right) \right\}$	$N_i \left\{ \frac{\varepsilon_i}{N_i} + \sum_k \frac{\varepsilon_{ik}}{N_{ik}} \right\}$
NCPF/1	$\frac{1}{4} N_{ij} \sum_k \left(\frac{\varepsilon_{ik}}{N_{ik}} + \frac{\varepsilon_{jk}}{N_{jk}} \right)$	$N_i \left\{ \frac{\varepsilon_{ii}}{N_{ii}} + \sum_k \frac{\varepsilon_{ik}}{N_{ik}} \right\}$
NCPF/2	$N_{ij} \frac{\varepsilon_{ij}}{N_{ij}}$	$2N_i \frac{\varepsilon_{ii}}{N_{ii}}$
NCPF/3	$N_{ij} \left\{ -\frac{\varepsilon_{ij}}{N_{ij}} + \frac{1}{2} \sum_k \left(\frac{\varepsilon_{ik}}{N_{ik}} + \frac{\varepsilon_{jk}}{N_{jk}} \right) \right\}$	$2N_i \sum_k \frac{\varepsilon_{ik}}{N_{ik}}$
ACPF	$\frac{2}{n} E_C$	$\frac{2}{p} E_C$
ACPF/2	$\frac{2}{n} E_C$	$\left[1 - \frac{(n-3)(n-2)}{n(n-1)} \right] E_C$
NACPF	$\frac{2}{p} E_C$	$\frac{4}{p} E_C$
AQCC	$\left[1 - \frac{(n-3)(n-2)}{n(n-1)} \right] E_C$	$\left[1 - \frac{(n-3)(n-2)}{n(n-1)} \right] E_C$

The QCISD method requires some slight modifications. We found it most convenient to think about the effect of the nonlinear terms as a “dressing” of the integrals occurring in equations (3.107) and (3.108). This attitude is close

to the recent arguments of Heully and Malrieu and may even open interesting new routes towards the calculation of excited states and the incorporation of connected triple excitations.[382] The dressed integrals are given by:

$$\bar{F}_{ik} = F_{ik} + \sum_l \langle \mathbf{C}^{il} (2\mathbf{K}^{kl} - \mathbf{K}^{kl+}) \rangle \quad (3.113)$$

$$\bar{F}_{ab} = F_{ab} - \sum_{kl} \{ \mathbf{C}^{kl} (2\mathbf{K}^{kl} - \mathbf{K}^{kl+}) \}_{ab} \quad (3.114)$$

$$\bar{F}_{kc} = F_{kc} + \sum_l (2\mathbf{K}^{kl} - \mathbf{K}^{kl+}) \mathbf{C}^l \quad (3.115)$$

$$\bar{K}_{kl}^{ij} = K_{kl}^{ij} + \langle \mathbf{K}^{kl} \mathbf{T}^{kl+} \rangle \quad (3.116)$$

$$\bar{K}_{ab}^{ij} = K_{ab}^{ij} + \sum_k \left\{ \mathbf{C}^{ik} \left(\mathbf{K}^{kj} - \frac{1}{2} \mathbf{K}^{jk} \right) + \mathbf{C}^{ki} \mathbf{K}^{kj} \right\}_{ab} \quad (3.117)$$

$$\bar{J}_{ab}^{ij} = J_{ab}^{ij} + \sum_k \{ \mathbf{C}^{ki} \mathbf{K}^{jk} \}_{ab} \quad (3.118)$$

The CCSD method can be written in a similar way but requires 15 additional terms that we do not document here. They may be taken conveniently from our paper about the LPNO-CCSD method [383].

A somewhat subtle point concerns the definition of the shifts in making the transition from spin-orbitals to spatial orbitals. For example, the CEPA/2 shift becomes in the generator state formalism:

$$-\langle \tilde{\Phi}_{ij}^{ab} | \Delta^{ij} | \Psi \rangle = C_{ab}^{ij} \left(\frac{1}{3} \varepsilon_{ij}^{\alpha\alpha} + \frac{2}{3} \varepsilon_{ij}^{\alpha\beta} \right) + C_{ba}^{ij} \left(-\frac{1}{3} \varepsilon_{ij}^{\alpha\alpha} + \frac{1}{3} \varepsilon_{ij}^{\alpha\beta} \right) \quad (3.119)$$

($\tilde{\Phi}_{ij}^{ab}$ is a contravariant configuration state function, see Pulay *et al.* [375]. The parallel and antiparallel spin pair energies are given by:

$$\varepsilon_{ij}^{\alpha\alpha} = \frac{1}{2} \sum_{ab} [K_{ab}^{ij} - K_{ba}^{ij}] (C_{ab}^{ij} - C_{ba}^{ij}) \quad (3.120)$$

$$\varepsilon_{ij}^{\alpha\beta} = \frac{1}{2} \sum_{ab} K_{ab}^{ij} C_{ab}^{ij} \quad (3.121)$$

This formulation would maintain the exact equivalence of an orbital and a spin-orbital based code. Only in the (unrealistic) case that the parallel and antiparallel pair correlation energies are equal the CEPA/2 shift of Table 3.35 arise. However, we have not found it possible to maintain the same equivalence for the CPF method since the electron pairs defined by the generator state formalism are a combination of parallel and antiparallel spin pairs. In order to maintain the maximum degree of internal consistency we have therefore decided to follow the proposal of Ahlrichs and co-workers and use the topological matrix T_{PQ} in equation (3.102) and the equivalents thereof in the CEPA and CPF methods that we have programmed.

Open-Shell Equations

We have used a non-redundant set of three spin cases ($\alpha\alpha$, $\beta\beta$, $\alpha\beta$) for which the doubles amplitudes are optimized separately. The equations in the spin-unrestricted formalism are straightforwardly obtained from the corresponding spin orbital equations by integrating out the spin. For implementing the unrestricted QCISD and CCSD method, we applied the same strategy (dressed integrals) as in the spin-restricted case. The resulting equations are quite cumbersome and will not be shown here explicitly [384].

Note that the definitions of the spin-unrestricted CEPA shifts differ from those of the spin-restricted formalism described above (see Kollmar *et al.* [384]). Therefore, except for CEPA/1 and VCEPA/1 (and of course CEPA/0), for which the spin-adaptation of the shift can be done in a consistent way, CEPA calculations of closed-shell molecules yield slightly different energies for the spin-restricted and spin-unrestricted versions. Since variant 1 is also the most accurate among the various CEPA variants [385], we recommend to use variant 1 for coupled-pair type calculations. For the variants 2 and 3, reaction energies of reactions involving closed-shell and open-shell molecules simultaneously should be calculated using the spin-unrestricted versions only.

A subtle point for open-shell correlation methods is the choice of the reference determinant [386]. Single reference correlation methods only yield reliable results if the reference determinant already provides a good description of the systems electronic structure. However, an UHF reference wavefunction suffers from spin-contamination which can spoil the results and lead to convergence problems. This can be avoided if quasi-restricted orbitals (QROs) are used [360, 382] since the corresponding zeroth-order wavefunction is an eigenfunction of the \hat{S}^2 operator and thus, no severe spin-contamination will appear. The coupled-pair and coupled-cluster equations will be still solved in a spin-unrestricted formalism but the energy will be slightly higher compared to the results obtained with a spin-polarized UHF reference determinant. Furthermore, especially for more difficult systems like e.g. transition metal complexes, it is often advantageous to use Kohn-Sham (KS) orbitals instead of HF orbitals.

3.10.2 Basic Usage

The coupled-cluster and CI methods in `orca_mdci` are available for RHF and UHF references. The implementation is fairly efficient and suitable for large-scale calculations. The most elementary use of this module is fairly simple.

```
! METHOD
# where METHOD is:
#   CCSD CCSD(T) QCISD QCISD(T) CPF/n NCPF/n CEPA/n NCEPA/n
#   (n=1,2,3 for all variants) ACPF NACPF AQCC CISD

! AOX-METHOD
# computes contributions from integrals with 3- and 4-external
# labels directly from AO integrals that are pre-stored in a
# packed format suitable for efficient processing

! AO-METHOD
# computes contributions from integrals with 3- and 4-external
# labels directly from AO integrals. Can be done for integral
# direct and conventional runs. In particular, the conventional
# calculations can be very efficient

! MO-METHOD (this is the default)
# performs a full four index integral transformation. This is
# also often a good choice

! RI-METHOD
# selects the RI approximation for all integrals. Rarely advisable

! RI34-METHOD
# selects the RI approximation for the integrals with 3- and 4-
# external labels
#
# The module has many additional options that are documented
# later in the manual.

! RCSinglesFock
! RIJKSinglesFock
! NoRCSinglesFock
! NoRIJKSinglesFock
# Keywords to select the way the so-called singles Fock calculation
# is evaluated. The first two keywords turn on, the second two turn off
# RIJCOSX or RIJK, respectively.
```

Note

- The same FrozenCore options as for MP2 are applied in the MDCI module.
- Since ORCA 4.2, an additional term, called “4th-order doubles-triples correction” is considered in open-shell CCSD(T). To reproduce previous results, one should use a keyword,

```
%mdci
  Include_4thOrder_DT_in_Triples  false
end
```

- The MDCI module cannot deal with systems without alpha virtual orbitals, even if there are beta virtual orbitals. Normally this only happens when the user uses minimal basis sets.

The computational effort for these methods is high — $O(N^6)$ for all methods and $O(N^7)$ if the triples correction is to be computed (calculations based on an unrestricted determinant are roughly 3 times more expensive than closed-shell calculations and approximately six times more expensive if triple excitations are to be calculated). This restricts the calculations somewhat: on presently available PCs 300–400 basis functions are feasible and if you are patient and stretch it to the limit it may be possible to go up to 500–600; if not too many electrons are correlated maybe even up to 800–900 basis functions (when using AO-direct methods).

Tip

- For calculations on small molecules and large basis sets the MO-METHOD option is usually the most efficient; say perhaps up to about 300 basis functions. For integral conventional runs, the AO-METHOD may even more efficient.
- For large calculations (>300 basis functions) the AO-METHOD option is a good choice. If, however, you use very deeply contracted basis sets such as ANOs these calculations should be run in the integral conventional mode.
- AOX-METHOD is usually slightly less efficient than MO-METHOD or AO-METHOD.
- RI-METHOD is seldom the most efficient choice. If the integral transformation time is an issue than you can select %mdci trafotype trafo_ri or choose RI-METHOD and then %mdci kcopt kc_ao.
- Regarding the singles Fock keywords (RCSinglesFock, etc.), the program usually decides which method to use to evaluate the singles Fock term. For more details on the nature of this term, and options related to its evaluation, see *The singles Fock term*.

To put this into perspective, consider a calculation on serine with the cc-pVDZ basis set — a basis on the lower end of what is suitable for a highly correlated calculation. The time required to solve the equations is listed in [Table 3.36](#). We can draw the following conclusions:

- As long as one can store the integrals and the I/O system of the computer is not the bottleneck, the most efficient way to do coupled-cluster type calculations is usually to go via the full transformation (it scales as $O(N^5)$ whereas the later steps scale as $O(N^6)$ and $O(N^7)$ respectively).
- AO-based coupled-cluster calculations are not much inferior. For larger basis sets (i.e. when the ratio of virtual to occupied orbitals is larger), the computation times will be even more favorable for the AO based implementation. The AO direct method uses much less disk space. However, when you use a very expensive basis set the overhead will be larger than what is observed in this example. Hence, conventionally stored integrals — if affordable — are a good choice.
- AOX-based calculations run at essentially the same speed as AO-based calculations. Since AOX-based calculations take four times as much disk space, they are pretty much outdated and the AOX implementation is only kept for historical reasons.
- RI-based coupled-cluster methods are significantly slower. There are some disk space savings, but the computationally dominant steps are executed less efficiently.
- CCSD is at most 10% more expensive than QCISD. With the latest AO implementation the awkward coupled-cluster terms are handled efficiently.
- CEPA is not much more than 20% faster than CCSD. In many cases CEPA results will be better than CCSD and then it is a real saving compared to CCSD(T), which is the most rigorous.

- If triples are included practically the same comments apply for MO versus AO based implementations as in the case of CCSD.

ORCA is quite efficient in this type of calculation, but it is also clear that the range of application of these rigorous methods is limited as long as one uses canonical MOs. ORCA implements novel variants of the so-called local coupled-cluster method which can calculate large, real-life molecules in a linear scaling time. This will be addressed in Sec. [Local correlation \(DLPNO\)](#).

Table 3.36: Computer times (minutes) for solving the coupled-cluster/coupled-pair equations for Serine (cc-pVDZ basis set)

Method	SCFMode	Time (min)
MO-CCSD	Conv	38.2
AO-CCSD	Conv	47.5
AO-CCSD	Direct	50.8
AOX-CCSD	Conv	48.7
RI-CCSD	Conv	64.3
AO-QCISD	Conv	44.8
AO-CEPA/1	Conv	40.5
MO-CCSD(T)	Conv	147.0
AO-CCSD(T)	Conv	156.7

All of these methods are designed to cover dynamic correlation in systems where the Hartree-Fock determinant dominates the wavefunctions. The least attractive of these methods is CISD which is not size-consistent and therefore practically useless. The most rigorous are CCSD(T) and QCISD(T). The former is perhaps to be preferred, since it is more stable in difficult situations.³ One can get highly accurate results from such calculations. However, one only gets this accuracy in conjunction with large basis sets. It is perhaps not very meaningful to perform a CCSD(T) calculation with a double-zeta basis set (see [Table 3.37](#)). The very least basis set quality required for meaningful results would perhaps be something like def2-TZVP(-f) or preferably def2-TZVPP (cc-pVTZ, ano-pVTZ). For accurate results quadruple-zeta and even larger basis sets are required and at this stage the method is restricted to rather small systems.

Let us look at the case of the potential energy surface of the N₂ molecule. We study it with three different basis sets: TZVP, TZVPP and QZVP. The input is the following:

```
! TZVPP CCSD(T)
%paras R= 1.05,1.13,8
end
* xyz 0 1
N 0 0 0
N 0 0 {R}
*
```

For even higher accuracy we would need to introduce relativistic effects and - in particular - turn the core correlation on.⁴

Table 3.37: Computed spectroscopic constants of N₂ with coupled-cluster methods.

³ The exponential of the T1 operator serves to essentially fully relax the orbitals of the reference wavefunction. This is not included in the QCISD model that only features at most a linear T1T2 term in the singles residuum. Hence, if the Hartree-Fock wavefunction is a poor starting point but static correlation is not the main problem, CCSD is much preferred over QCISD. This is not uncommon in transition metal complexes.

⁴ Note that core correlation is not simply introduced by including the core orbitals in the correlation problem. In addition, special correlation core-polarization functions are needed. They have been standardized for a few elements in the cc-pCVxZ (X=D,T,Q,5,6) basis sets.

Method	Basis set	R_e (pm)	ω_e (cm ⁻¹)	$\omega_e x_e$ (cm ⁻¹)
CCSD(T)	SVP	111.2	2397	14.4
	TZVP	110.5	2354	14.9
	TZVPP	110.2	2349	14.1
	QZVP	110.0	2357	14.3
	ano-pVDZ	111.3	2320	14.9
	ano-pVTZ	110.5	2337	14.4
	ano-pVQZ	110.1	2351	14.5
CCSD	QZVP	109.3	2437	13.5
Exp		109.7	2358.57	14.32

One can see from Table 3.37 that for high accuracy - in particular for the vibrational frequency - one needs both - the connected triple-excitations and large basis sets (the TZVP result is fortuitously good). While this is an isolated example, the conclusion holds more generally. If one pushes it, CCSD(T) has an accuracy (for reasonably well-behaved systems) of approximately 0.2 pm in distances, <10 cm⁻¹ for harmonic frequencies and a few kcal/mol for atomization energies.⁵ It is also astonishing how well the Ahlrichs basis sets do in these calculations — even slightly better than the much more elaborate ANO bases.

Note

The quality of a given calculation is not always high because it carries the label “coupled-cluster”. Accurate results are only obtained in conjunction with large basis sets and for systems where the HF approximation is a good 0th order starting point.

Frozen Core Options

In coupled-cluster calculations the Frozen Core (FC) approximation is applied by default. This implies that the core electrons are not included in the correlation treatment, since the inclusion of dynamic correlation in the core electrons usually affects relative energies insignificantly.

The frozen core option can be switched on or off with ! FrozenCore or ! NoFrozenCore in the simple input. More information and further options are given in section *Frozen Core Options* and in section *Including (semi)core orbitals in the correlation treatment*.

The singles Fock term

In most MDCI calculations, there is an intermediate, which resembles closely to the SCF Fock matrix, and similar methods are available to efficiently calculate it. In the followings, a short discussion will be given of the so-called singles Fock term, which in the closed shell case has the form

$$G(\mathbf{t}_1)_{pq} = \sum_{jb} t_b^j (2(pq|jb) - (pj|qb)) = \sum_{\mu\nu} c_\mu^p c_\nu^q G(\mathbf{t}_1)_{\mu\nu},$$

The singles Fock matrix can be obtained via transformation from its counterpart ($G(\mathbf{t}_1)_{\mu\nu}$) in the atomic orbital (AO) basis

$$G(\mathbf{t}_1)_{\mu\nu} = \sum_{jb} t_b^j (2(\mu\nu|jb) - (\mu j|\nu b)) = \sum_{\kappa\tau} P(\mathbf{t}_1)_{\kappa\tau} (2(\mu\nu|\kappa\tau) - (\mu\kappa|\nu\tau)), \quad (3.122)$$

where

$$P(\mathbf{t}_1)_{\kappa\tau} = \sum_{jb} t_b^j c_\kappa^j c_\tau^b$$

⁵ However, in recent years it became more evident that even CCSD(T) achieves its high apparent accuracy through error cancellations. The full CCSDT method (triples fully included) usually performs worse than CCSD(T). The reason is that the (T) correction undershoots the effects of the triples to some extent and thereby compensates for the neglect of connected quadruple excitations. For very high accuracy quantum chemistry, even these must be considered. The prospects for treating chemically more relevant molecules with such methods is not particularly bright for the foreseeable future...

is the analogue of the SCF density matrix for the singles Fock case. For the singles Coulomb ($J(\mathbf{t}_1)_{\mu\nu}$) case, the density may be symmetrized ($\tilde{P}(\mathbf{t}_1)_{\kappa\tau} = P(\mathbf{t}_1)_{\kappa\tau} + P(\mathbf{t}_1)_{\tau\kappa}$), and one may use the resolution of identity approximation

$$J(\mathbf{t}_1)_{\mu\nu} = \sum_{\kappa\tau} \tilde{P}(\mathbf{t}_1)_{\kappa\tau} (\mu\nu|\kappa\tau) \approx \sum_{AB} \sum_{\kappa\tau} \tilde{P}(\mathbf{t}_1)_{\kappa\tau} (\mu\nu|r_{12}^{-1}|A) V_{AB}^{-1} (B|r_{12}^{-1}|\kappa\tau),$$

where A, B are elements of the RI/DF auxiliary fitting basis. Note that the factor of 2 in ((3.122)) is taken care of by symmetrization. Since we are using a symmetric density, we may use the same efficient routine to evaluate the singles Coulomb term as in the SCF case, see [RI-J](#) and [Split-RI-J](#).

For the exchange case ($K(\mathbf{t}_1)_{\mu\nu}$), one possibility is to use the COSX approximation (see [RIJCOSX](#))

$$K(\mathbf{t}_1)_{\mu\nu} = \sum_{\kappa\tau} P(\mathbf{t}_1)_{\kappa\tau} (\mu\kappa|\nu\tau) \approx \sum_g Q_{\mu g} \sum_{\tau} A_{\nu\tau}(\mathbf{r}_g) \sum_{\kappa} P(\mathbf{t}_1)_{\kappa\tau} X_{\kappa g},$$

The COSX routine is able to deal with asymmetric densities as well, and thus, it can be used here similar to the SCF case.

The other possibility is to use RI for exchange (RIK),

$$K(\mathbf{t}_1)_{\mu\nu} = \sum_{j\kappa\tau} c_{\kappa}^j C(\mathbf{t}_1)_{\tau}^j (\mu\kappa|\nu\tau) \approx \sum_{jAB} (\mu j|r_{12}^{-1}|A) V_{AB}^{-1} (B|r_{12}^{-1}|\nu\tilde{j}),$$

where

$$C(\mathbf{t}_1)_{\tau}^j = \sum_b t_b^j c_{\tau}^b,$$

and the \tilde{j} is an “orbital” transformed using $C(\mathbf{t}_1)$.

Using these approximations, there are two approximations for the total singles Fock term, RIJCOSX called by the simple keyword `RCSinglesFock` and RIJK called by `RIJKSinglesFock`, see [Basic Usage](#). For canonical and LPNO methods, by default the program chooses the same approximation used in the SCF calculation. DLPNO2013 uses RIJCOSX by default, while in DLPNO, the singles Fock term is also evaluated using PNOs via `SinglesFockUsePNOs`, see [Local correlation \(DLPNO\)](#). This behavior can also be changed by keywords in the method block.

```
%method RIJKSinglesFock 1 # 0 false, 1 true
      RCSinglesFock 0 # 0 false, 1 true
end
```

3.10.3 Coupled-Cluster Densities

If one is mainly accustomed to Hartree-Fock or DFT calculations, the calculation of the density matrix is more or less a triviality and is automatically done together with the solution of the self-consistent field equations. Unfortunately, this is not the case in coupled-cluster theory (and also not in MP2 theory). The underlying reason is that in coupled-cluster theory, the expansion of the exponential $e^{\hat{T}}$ in the expectation value

$$D_{pq} = \frac{\langle \Psi | E_p^q | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\langle e^{\hat{T}} \Psi_0 | E_p^q | e^{\hat{T}} \Psi_0 \rangle}{\langle e^{\hat{T}} \Psi_0 | e^{\hat{T}} \Psi_0 \rangle}$$

only terminates if all possible excitation levels are exhausted, i.e., if all electrons in the reference determinant Ψ_0 (typically the HF determinant) are excited from the space of occupied to the space of virtual orbitals (here D_{pq} denotes the first order density matrix, E_p^q are the spin traced second quantized orbital replacement operators, and \hat{T} is the cluster operator). Hence, the straightforward application of these equations is far too expensive. It is, however, possible to expand the exponentials and only keep the linear term. This then defines a linearized density which coincides with the density that one would calculate from linearized coupled-cluster theory (CEPA/0). The difference to the CEPA/0 density is that converged coupled-cluster amplitudes are used for its evaluation. This density is straightforward to compute and the computational effort for the evaluation is very low. Hence, this is a density that can be easily produced in a coupled-cluster run. It is not, however, what coupled-cluster aficionados would accept as a density.