

Optimization of the CASSCF wavefunction

Feizhi Ding

April 4, 2015

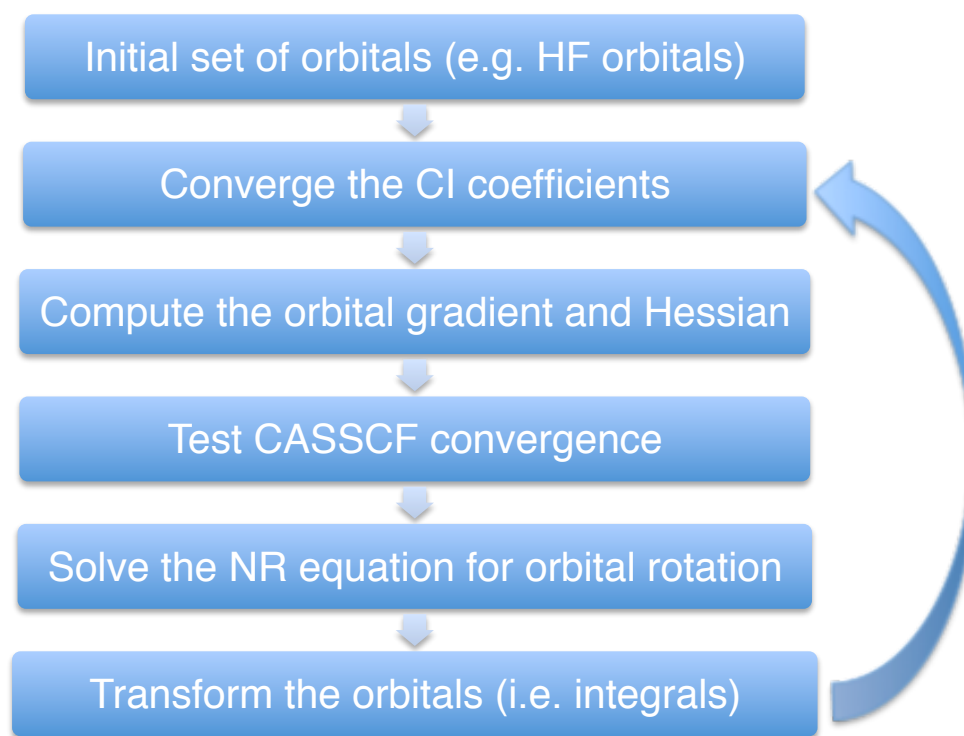


Figure 1: Flow chart for the two-step CASSCF wavefunction optimization.

Organization of the program

The program is organized according to the flow chart shown in Figure 1.

CASSCF is the main program, which calls the following important subroutines:

- Prepare the excitation lists: **StrLst** generates the alpha/beta strings and the excitation lists, which will be used repeatedly for solving the CI eigenvalue problem and the construction of the density matrices. (Section 3.3.1)
- Prepare the initial MO integrals: **LoAOIn** read the AO one- and two-electron integrals from the disk. **InCTrn** transforms AO integrals into MO integrals. Alternatively, the initial MO integrals can be read directly from the disk by calling subroutine **GetMOX**.
- **RHFCor** computes the inactive orbital HF energy ε^c , the inactive orbital Fock matrix \mathbf{h}^c , and the contracted inactive orbital Fock matrix \mathbf{h}^c , in order to solve the CAS-CI eigenvalue problem. (Section 3.2)
- Optimization of the CI coefficients: There are three diagonalization methods implemented: (1) subroutine **Nesbet** uses the Cooper-Nesbet iterative method to find the lowest eigenvalue (i.e. ground state only); (2) subroutine **DavIni** calls the Davidson diagonalization routine **DavDag**; (3) subroutine **DDagCI** directly diagonalizes the CI matrix (expensive, and only works when the CI expansion is smaller than a few thousands).
- Subroutine **ASDMat** forms the one- and two-particle density matrices, and subroutine **GenFck** forms the generalized Fock matrices. The density matrices and generalized Fock matrices will be used for constructing the orbital gradient and Hessian.
- Orbital rotation. There are 4 approaches: (1) do a quasi-Newton step using an approximate, diagonal orbital Hessian, where **OGradS** computes the orbital gradient in square matrix form (sparse!) and **OHessD** computes the diagonal elements of the orbital Hessian; (2) do a Newton-Raphson (NR) step using the full orbital Hessian, where **OGradL** computes the orbital gradient in the compact linearized form and

OHessF computes the full orbital Hessian (indexing is consistent with **OGradL**); (3) Same as (2), except that the state-orbital coupling contributions are added to the orbital Hessian; (4) do a DIIS extrapolation of the orbital rotation matrix from previous steps.

- Subroutine **MatExp** forms the unitary matrix for orbital transformation.
- Transform the one-electron and two-electron integrals. This is done in subroutine **InCTrn**

Notation:

i, j, k, l	inactive orbitals
t, u, v, w	active orbitals
a, b, c, d	virtual orbitals
p, q, r, s, τ	arbitrary orbitals

1 Exponential parameterization of the CASSCF wavefunction

In the CASSCF optimization, the initial wavefunction $|0\rangle$ is described as a linear combination of some configuration state functions (CSFs) or Slater determinants $|K\rangle$, generated by excitations within the active space,

$$|0\rangle = \sum_{K=1}^M c_K |K\rangle \quad (1)$$

where M is the dimension of the expansion.

Optimization of the CASSCF wavefunction consists of two parts: (1) variation of the CI expansion coefficients, and (2) variation of the molecular orbitals.

1.1 Variation of the CI expansion coefficients

There are several ways to parameterize the CI problem. The simplest way is to use the linear expansion coefficients $\{c_K\}$ as the variational parameters. However, this parameterization contains a redundancy because of the normalization constraint

$$\sum_{K=1}^M c_K = 1 \quad (2)$$

Another way is to use an exponential unitary parameterization of the CI space, where the variation of the CI wavefunction $\{c_K\}$ is achieved by

$$|\tilde{0}\rangle = \exp(\hat{S})|0\rangle$$

where the operator \hat{S} is defined as

$$\hat{S} = \sum_{n \neq 0}^M s_n |n\rangle \langle 0| - |0\rangle \langle n| \quad (3)$$

and $\{|n\rangle\}$ are orthogonal complement to the reference state $|0\rangle$.

The advantage of using this exponential parameterization is that it includes the normalization constraint through the exponential unitary operator and the optimization can be easily controlled.

1.2 Variation of the orbitals

The variation of the orbitals $\{\phi_p\}$ can be achieved by orbital rotation through the real unitary matrix \mathbf{U} :

$$\tilde{\phi}_p = \sum_q U_{pq} \phi_q \quad (4)$$

where the \mathbf{U} matrix may be written in terms of an anti-symmetrix matrix \mathbf{x} :

$$\mathbf{U} = \exp(\mathbf{x}) \quad (5)$$

It can be shown that the unitary transformation on the orbitals is equivalent to the following transformation on the CASSCF wavefunction:¹

$$|\tilde{0}\rangle = \exp(\hat{T})|0\rangle$$

where the operator \hat{T} is defined in the formulism of second quantization:

$$\hat{T} = \sum_{pq} x_{pq} \hat{E}_{pq} \quad (6)$$

where \hat{E}_{pq} is the single-excitation operator.

1.3 The CASSCF energy expression

With the above mentioned exponential parameterization for the CI coefficients and orbital rotations, the CASSCF wavefunction can be written as

$$|\tilde{0}\rangle = \exp(\hat{T}) \exp(\hat{S})|0\rangle \quad (7)$$

The total CASSCF energy is then

$$\begin{aligned} E &= \langle \tilde{0} | \hat{H} | \tilde{0} \rangle \\ &= \langle 0 | \exp(-\hat{S}) \exp(-\hat{T}) \hat{H} \exp(\hat{T}) \exp(\hat{S}) | 0 \rangle \end{aligned} \quad (8)$$

Expanding the above expression using the Baker-Campbell-Hausdorff formula, we have

$$\begin{aligned} E &= \langle \tilde{0} | \exp(-\hat{S})(\hat{H} + [\hat{H}, \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \dots) \exp(\hat{S}) | 0 \rangle \\ &= \langle 0 | \hat{H} + [\hat{H}, \hat{S}] + \frac{1}{2}[[\hat{H}, \hat{S}], \hat{S}] + \dots + [\hat{H}, \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \dots \\ &\quad + [[\hat{H}, \hat{T}], \hat{S}] + \dots | 0 \rangle \\ &= \langle 0 | \hat{H} | 0 \rangle + \langle 0 | [\hat{H}, \hat{S}] | 0 \rangle + \langle 0 | [\hat{H}, \hat{T}] | 0 \rangle + \frac{1}{2} \langle 0 | [[\hat{H}, \hat{S}], \hat{S}] | 0 \rangle \\ &\quad + \frac{1}{2} \langle 0 | [[\hat{H}, \hat{T}], \hat{T}] | 0 \rangle + \langle 0 | [[\hat{H}, \hat{T}], \hat{S}] | 0 \rangle + \dots \\ &= E^{(0)} + \mathbf{s}^T \mathbf{g}^c + \mathbf{x}^T \mathbf{g}^o + \frac{1}{2} \mathbf{s}^T \mathbf{H}^{cc} \mathbf{s} + \frac{1}{2} \mathbf{x}^T \mathbf{H}^{oo} \mathbf{x} + \mathbf{x}^T \mathbf{H}^{oc} \mathbf{s} + \dots \end{aligned} \quad (9)$$

where the individule terms are expressed as

$$\langle 0|\hat{H}|0\rangle = E^{(0)} \quad (10)$$

$$\langle 0|[\hat{H}, \hat{S}]|0\rangle = \sum_n s_n g_n^c = \mathbf{s}^T \mathbf{g}^c \quad (11)$$

$$\langle 0|[\hat{H}, \hat{T}]|0\rangle = \sum_{pq} x_{pq} g_{pq}^o = \mathbf{x}^T \mathbf{g}^o \quad (12)$$

$$\langle 0|[[\hat{H}, \hat{S}], \hat{S}]|0\rangle = \sum_n \sum_{n'} s_n s_{n'} H_{nn'}^{cc} = \mathbf{s}^T \mathbf{H}^{cc} \mathbf{s} \quad (13)$$

$$\langle 0|[[\hat{H}, \hat{T}], \hat{T}]|0\rangle = \sum_{pq} \sum_{rs} x_{pq} x_{rs} H_{pq,rs}^{oo} = \mathbf{x}^T \mathbf{H}^{oo} \mathbf{x} \quad (14)$$

$$\langle 0|[[\hat{H}, \hat{T}], \hat{S}]|0\rangle = \sum_{pq} \sum_n x_{pq} H_{pq,n}^{oc} s_n = \mathbf{x}^T \mathbf{H}^{oc} \mathbf{s} \quad (15)$$

where \mathbf{g}^c is called the state gradient, \mathbf{g}^o the orbital gradient, \mathbf{H}^{cc} the state Hessian, \mathbf{H}^{oo} the orbital Hessian, and \mathbf{H}^{oc} the state-orbital coupling Hessian.

Optimization of the CASSCF wavefunction is equivalent to finding the stationary solution of the energy with respect to the CI variation \mathbf{s} and the orbital rotation \mathbf{x} :

$$\frac{\partial E}{\partial \mathbf{s}} = \mathbf{g}^c = 0 \quad (16)$$

$$\frac{\partial E}{\partial \mathbf{x}} = \mathbf{g}^o = 0 \quad (17)$$

If we truncate terms up to second order in \mathbf{s} and \mathbf{x} , we obtain the second-order energy expression:

$$\begin{aligned} E^{(2)} &\approx E^{(0)} + \mathbf{s}^T \mathbf{g}^c + \mathbf{x}^T \mathbf{g}^o + \frac{1}{2} \mathbf{s}^T \mathbf{H}^{cc} \mathbf{s} + \frac{1}{2} \mathbf{x}^T \mathbf{H}^{oo} \mathbf{x} + \mathbf{x}^T \mathbf{H}^{oc} \mathbf{s} \\ &= E^{(0)} + \begin{pmatrix} \mathbf{s}^T & \mathbf{x}^T \end{pmatrix} \begin{pmatrix} \mathbf{g}^c \\ \mathbf{g}^o \end{pmatrix} + \frac{1}{2} \begin{pmatrix} \mathbf{s}^T & \mathbf{x}^T \end{pmatrix} \begin{pmatrix} \mathbf{H}^{cc} & \mathbf{H}^{co} \\ \mathbf{H}^{oc} & \mathbf{H}^{oo} \end{pmatrix} \begin{pmatrix} \mathbf{s} \\ \mathbf{x} \end{pmatrix} \end{aligned} \quad (18)$$

Optimization of this truncated energy expression leads to the Newton-Raphson equation:

$$\begin{pmatrix} \mathbf{H}^{cc} & \mathbf{H}^{co} \\ \mathbf{H}^{oc} & \mathbf{H}^{oo} \end{pmatrix} \begin{pmatrix} \mathbf{s} \\ \mathbf{x} \end{pmatrix} = - \begin{pmatrix} \mathbf{g}^c \\ \mathbf{g}^o \end{pmatrix} \quad (19)$$

2 The two-step optimization approach

Eq. (19) can be either solved simultaneously for both CI variation \mathbf{s} and orbital rotation \mathbf{x} , or solved in a two-step manner where the CI coefficients are optimized for the current set of orbitals and the Newton-Raphson step is taken only for the orbital rotations. Both simultaneous method and two-step approach possess the same local convergence properties, however, the two-step procedure is expected to be more stable since the CI coefficients are optimized more fully at each iteration.

In the two-step procedure, the CI coefficients are optimized first. This amounts to solving $\mathbf{g}^c = 0$, and the Newton-Raphson equation, Eq. (19) becomes

$$\mathbf{H}^{cc}\mathbf{s} + \mathbf{H}^{co}\mathbf{x} = 0 \quad (20)$$

$$\mathbf{H}^{oc}\mathbf{s} + \mathbf{H}^{oo}\mathbf{x} = -\mathbf{g}^o \quad (21)$$

which can be written as a partitioned Newton-Raphson equation for orbital rotations only:

$$(\mathbf{H}^{oo} - \mathbf{H}^{oc}\mathbf{H}^{cc-1}\mathbf{H}^{co})\mathbf{x} = -\mathbf{g}^o \quad (22)$$

The flow chart for the two-step CASSCF wavefunction optimization is given in Figure 1.

The optimization of the CI coefficients will be discussed in Section 3.

Expressions of those matrices used for orbital rotation will be given in Section 5.

3 The CI coefficient optimization

3.1 The projected basis and the CI eigenvalue problem

As discussed above, the first step in the two-step CASSCF optimization is to solve for the stationary point of the energy with respect to the CI variation, i.e.

$$g_n^c = 2\langle 0|\hat{H}|n\rangle = 0 \quad (23)$$

where $|n\rangle$ is an orthogonal complement to $|0\rangle$. Note that the dimension of the state gradient \mathbf{g}^c is $M - 1$.

The choice of the orthogonal complement basis $\{|n\rangle\}$ is arbitrary. For computational convenience, two bases have been used for constructing the orthogonal complement basis: the CI eigenvector basis and the projected basis. The CI eigenvector basis is linearly independent and orthonormal. However, it is not a practical choice when the CI expansion becomes large where full diagonalization of the CI Hamiltonian is impossible.

The projected basis is defined as a projection on to the configurational space:

$$|n\rangle = (1 - |0\rangle\langle 0|)|\tilde{n}\rangle \quad (24)$$

where $|\tilde{n}\rangle$ are the configurational state functions (CSFs) or Slater determinants, and they are orthonormal:

$$\langle \tilde{m} | \tilde{n} \rangle = \delta_{mn} \quad (25)$$

Note that the dimension of the projected basis is M . Therefore it is overcomplete.

The advantage of using the projected basis is that it can be readily constructed and requires no additional computational cost.

In the projected basis, the state gradient becomes

$$\begin{aligned} g_n^c &= 2\langle 0 | \hat{H} (1 - |0\rangle\langle 0|) | \tilde{n} \rangle \\ &= 2\langle 0 | \hat{H} - E_0 | \tilde{n} \rangle \\ &= 2 \sum_m \langle \tilde{m} | \hat{H} - E_0 | \tilde{n} \rangle c_m \end{aligned} \quad (26)$$

and in matrix notation,

$$\mathbf{g}^c = (\mathbf{H} - E_0)\mathbf{c} \quad (27)$$

Therefore, the optimization of the CI coefficients in the project basis is equivalent to

solving the eigenvalue problem:

$$\mathbf{g}^c = 0 \quad \Leftrightarrow \quad \mathbf{H}\mathbf{c} = E_0\mathbf{c} \quad (28)$$

3.2 The CAS-CI eigenvalue problem

For a CASSCF wavefunction, the CI expansion is limited to excitations within the active space. In this case, the Hamiltonian operator \hat{H} can be written as

$$\hat{H} = \varepsilon^c + \hat{H}^{\text{CAS}} \quad (29)$$

where ε^c is the Hartree-Fock energy of the closed shell inactive orbitals and is defined as

$$\varepsilon^c = 2 \sum_i h_{ii} + \sum_{ij} [2 (ii|jj) - (ij|ji)] \quad (30)$$

and \hat{H}^{CAS} is the active space Hamiltonian and is defined as

$$\begin{aligned} \hat{H}^{\text{CAS}} &= \sum_{tu} h_{tu}^c \hat{E}_{tu} + \frac{1}{2} \sum_{tuvw} (tu|vw) (\hat{E}_{tu} \hat{E}_{vw} - \delta_{uv} \hat{E}_{tw}) \\ &= \sum_{tu} h_{tu}^{c'} \hat{E}_{tu} + \frac{1}{2} \sum_{tuvw} (tu|vw) \hat{E}_{tu} \hat{E}_{vw} \end{aligned} \quad (31)$$

where \mathbf{h}^c is the closed shell inactive orbital Fock matrix, and is given by

$$h_{tu}^c = h_{tu} + \sum_i [2 (tu|ii) - (ti|ui)] \quad (32)$$

and $\mathbf{h}^{c'}$ is

$$h_{tu}^{c'} = h_{tu}^c - \frac{1}{2} \sum_v (tv|uv) \quad (33)$$

Therefore, the CAS-CI eigenvalue problem can be written as

$$\mathbf{H}^{\text{CAS}} \mathbf{c} = E^{\text{CAS}} \mathbf{c} \quad (34)$$

where E^{CAS} is defined as

$$E^{\text{CAS}} = E_0 - \varepsilon^c \quad (35)$$

Since the CI expansion is complete within the active space, by writing the CAS-CI eigenvalue problem in the form given in Eq. (34), we can directly apply efficient diagonalization algorithms developed for the full-CI problem.

Subroutine **RHFCor** computes ε^c , \mathbf{h}^c , and $\mathbf{h}^{c'}$.

3.3 The determinant CI approach

3.3.1 Generating strings and excitation lists

For large CAS-CI expansions, direct diagonalization of the Hamiltonian matrix \mathbf{H}^{CAS} becomes impossible and iterative methods such as the Cooper-Nesbet method,² Lanczos method,³ and Davidson algorithm⁴ have to be used. In these iterative procedures, the most time-consuming step is the calculation of the matrix-vector product:

$$\sigma_I = \sum_J^M \langle I | \hat{H}^{\text{CAS}} | J \rangle c_J \quad (36)$$

where I, J are the CI expansion basis functions.

One of the major steps toward the efficient CI is Handy's realization of the computational advances of splitting Slater determinants into alpha and beta strings:

$$|I\rangle = |I_\alpha I_\beta\rangle \quad (37)$$

The construction of the alpha and beta strings can be achieved using either the graphical approach or the method of Knowles and Handy.

The address of any alpha string is computed by the following formula:

$$\text{Address}(I_\alpha) = 1 + \sum_{i=1}^{N_\alpha} Z(i, L_i) \quad (38)$$

where L_i is the occupation of the i -th arc (0 or 1).

Z is the address array and is given by

$$Z(k, l) = \sum_{m=N_{\text{orb}}-l+1}^{N_{\text{orb}}-k} \left[\binom{m}{N_{\alpha}-k} - \binom{m-1}{N_{\alpha}-k-1} \right],$$

$$(N_{\text{orb}} - N_{\alpha} + k \geq l \geq k; \quad k < N_{\alpha}) \quad (39)$$

$$Z(N_{\alpha}, l) = l - N_{\alpha}$$

where k is the electron index and l the orbital index. N_{orb} is the number of active orbitals and N_{α} is the number of active alpha electrons.

The computation of the σ vector involves the calculation of the matrix elements of the single-excitation operators $\langle J_{\alpha} | \hat{E}_{ij}^{\alpha} | I_{\alpha} \rangle$ and $\langle J_{\beta} | \hat{E}_{ij}^{\beta} | I_{\beta} \rangle$. This is done by generating the nonzero excitation lists $\mathcal{L} \{i, j, J_{\alpha}, \text{sgn}(ij), I_{\alpha}\}$ and $\mathcal{L} \{i, j, J_{\beta}, \text{sgn}(ij), I_{\beta}\}$.

For each alpha string I_{α} , the excitation list stores the following elements:

- (1) i, j : the nonzero excitation index pairs;
- (2) J_{α} : the address of the corresponding excited string;
- (3) $\text{sgn}(ij)$: the sign of the matrix element $\langle J_{\alpha} | \hat{E}_{ij}^{\alpha} | I_{\alpha} \rangle$.

The beta excitation list can be calculated in the same fashion (for closed-shell calculations, the same excitation list is used for alpha and beta strings). The list is pre-calculated and stored in main memory, and will be used repeatedly for later calculations.

Subroutine **StrLst** generates the alpha/beta strings and the excitation lists.

3.3.2 Calculating the matrix-vector product

Given the addresses of the alpha string I_{α} and beta string I_{β} , the associated Slater determinant I can be addressed simply as the element of a rectangular array (I_{α}, I_{β}) . In this way, operations on the alpha string alone can be performed for all beta strings in a vector loop, and vice versa.

Separating the alpha and beta strings, the CI matrix-vector product σ can be split into three parts:

$$\sigma_I = \sigma(I_\alpha, I_\beta) = \sigma^\alpha(I_\alpha, I_\beta) + \sigma^\beta(I_\alpha, I_\beta) + \sigma^{\alpha\beta}(I_\alpha, I_\beta) \quad (40)$$

where the first term σ^α involves only the α -type excitations:

$$\sigma^\alpha(I_\alpha, I_\beta) = \sum_{J_\alpha} \sum_{kl} \left[h'_{kl} \langle J_\alpha | \hat{E}_{kl}^\alpha | I_\alpha \rangle + \frac{1}{2} \sum_{K_\alpha} \sum_{ij} (ij|kl) \langle J_\alpha | \hat{E}_{ij}^\alpha | K_\alpha \rangle \langle K_\alpha | \hat{E}_{kl}^\alpha | I_\alpha \rangle \right] c(J_\alpha, I_\beta) \quad (41)$$

Similarly, σ^β only involves the β -type excitations:

$$\sigma^\beta(I_\alpha, I_\beta) = \sum_{J_\beta} \sum_{kl} \left[h'_{kl} \langle J_\beta | \hat{E}_{kl}^\beta | I_\beta \rangle + \frac{1}{2} \sum_{K_\beta} \sum_{ij} (ij|kl) \langle J_\beta | \hat{E}_{ij}^\beta | K_\beta \rangle \langle K_\beta | \hat{E}_{kl}^\beta | I_\beta \rangle \right] c(I_\alpha, J_\beta) \quad (42)$$

For closed-shell calculations, σ^β can be expressed in terms of σ^α through the following relation:

$$\sigma^\beta(I_\alpha, I_\beta) = \sigma^\alpha(I_\beta, I_\alpha) \quad (43)$$

$\sigma^{\alpha\beta}$ contains both α - and β -type single excitations:

$$\sigma^{\alpha\beta}(I_\alpha, I_\beta) = \sum_{J_\alpha, J_\beta} \sum_{ijkl} (ij|kl) \langle J_\alpha | \hat{E}_{ij}^\alpha | I_\alpha \rangle \langle J_\beta | \hat{E}_{kl}^\beta | I_\beta \rangle c(J_\alpha, J_\beta) \quad (44)$$

where the second equality was obtained by permuting i, j with k, l in the summation.

With the pre-calculated excitation lists, $\mathcal{L}\{i, j, J_\alpha, \text{sgn}(ij), I_\alpha\}$ and $\mathcal{L}\{i, j, J_\beta, \text{sgn}(ij), I_\beta\}$, the computation of σ^α , σ^β , and $\sigma^{\alpha\beta}$ can be done using efficient vectorizable loop algorithms.

The algorithm for calculating σ^α is :

Loop over I_α

Initialize an array $F(J_\alpha)$ to zero

Loop over the excitation list $\mathcal{L}\{k, l, K_\alpha, \text{sgn}(kl), I_\alpha\}$

$$F(K_\alpha) = F(K_\alpha) + \text{sgn}(kl)h'_{kl}$$

Loop over the excitation list $\mathcal{L}\{i, j, J_\alpha, \text{sgn}(ij), K_\alpha\}$

$$F(J_\alpha) = F(J_\alpha) + \frac{1}{2}\text{sgn}(ij)\text{sgn}(kl) (ij|kl)$$

end of loop

end of loop

Loop over I_β

$$\sigma^\alpha(I_\alpha, I_\beta) = \sum_{J_\alpha} F(J_\alpha) c(J_\alpha, I_\beta)$$

end of loop

(45)

The computation of $\sigma^\beta(I_\alpha, I_\beta)$ can be done in a similar fashion.

The pseudo-code for $\sigma^{\alpha\beta}(I_\alpha, I_\beta)$ is given as:

For a given I_α, I_β

Loop over the excitation list $\mathcal{L}\{kl, J_\alpha, \text{sgn}(kl), I_\alpha\}$

Loop over the excitation list $\mathcal{L}\{ij, J_\beta, \text{sgn}(ij), I_\beta\}$

$$\sigma^{\alpha\beta}(I_\alpha, I_\beta) = \sigma^{\alpha\beta}(I_\alpha, I_\beta) + \text{sgn}(ij)\text{sgn}(kl) (ij|kl) c(J_\alpha, J_\beta)$$

end of loop

end of loop

(46)

Subroutine **HCMult** computes the the CI matrix-vector product σ using the above algorithm.

3.3.3 Forming the diagonal elements of the CI Hamiltonian

In iterative diagonalization methods such as the Cooper-Nesbet method² and the Davidson algorithm,⁴ one needs to construct the diagonal elements of the CI Hamiltonian.

Within the string representation of the Slater determinants, the diagonal elements of the CI Hamiltonian are expressed as:

$$\begin{aligned} H_{II} = D(I_\alpha, I_\beta) &= \langle I_\alpha I_\beta | \hat{H} | I_\alpha I_\beta \rangle \\ &= D^\alpha(I_\alpha, I_\beta) + D^\beta(I_\alpha, I_\beta) + D^{\alpha\beta}(I_\alpha, I_\beta) \end{aligned} \quad (47)$$

where $D^\alpha(I_\alpha, I_\beta)$, $D^\beta(I_\alpha, I_\beta)$, and $D^{\alpha\beta}(I_\alpha, I_\beta)$ are given as

$$D^\alpha(I_\alpha, I_\beta) = \sum_{kl} \left[h'_{kl} \langle I_\alpha | \hat{E}_{kl}^\alpha | I_\alpha \rangle + \frac{1}{2} \sum_{K_\alpha} \sum_{ij} (ij|kl) \langle I_\alpha | \hat{E}_{ij}^\alpha | K_\alpha \rangle \langle K_\alpha | \hat{E}_{kl}^\alpha | I_\alpha \rangle \right] \quad (48)$$

$$D^\beta(I_\alpha, I_\beta) = \sum_{kl} \left[h'_{kl} \langle I_\beta | \hat{E}_{kl}^\beta | I_\beta \rangle + \frac{1}{2} \sum_{K_\beta} \sum_{ij} (ij|kl) \langle I_\beta | \hat{E}_{ij}^\beta | K_\beta \rangle \langle K_\beta | \hat{E}_{kl}^\beta | I_\beta \rangle \right] \quad (49)$$

$$D^{\alpha\beta}(I_\alpha, I_\beta) = \sum_{ijkl} (ij|kl) \langle I_\alpha | \hat{E}_{ij}^\alpha | I_\alpha \rangle \langle I_\beta | \hat{E}_{kl}^\beta | I_\beta \rangle \quad (50)$$

Based on the above expressions, we develop an algorithm for computing the diagonal CI Hamiltonian, which is a slight modification of the algorithm for calculating the σ vector. This is done in subroutine **FCDiag**.

3.3.4 The Cooper-Nesbet iterative method for obtaining the lowest eigenvalue

The Cooper-Nesbet method is an iterative diagonalization algorithm that can find the lowest eigenvalue for a real symmetric matrix \mathbf{H} :

$$\mathbf{H}\mathbf{c} = E\mathbf{c} \quad (51)$$

In the Cooper-Nesbet method, one eigenvector component (e.g. c_1) is fixed at unity, while each of the other components c_I is adjusted in turn according to:

$$q_I = \sum_I H_{IJ}c_J - Ec_I = \sigma_I - Ec_I \quad (52)$$

$$\Delta c_I = \frac{q_I}{E - H_{II}} \quad (53)$$

$$\Delta E = \frac{q_I \Delta c}{\|\mathbf{c} + \Delta \mathbf{c}\|} \quad (54)$$

with c_I and E being corrected by Δc_I and ΔE , respectively.

We define a macro-iteration as the completion of the sequential update of c_I and E for all components (except for the one that is fixed). Once a macro-iteration is finished, we return to the beginning component and perform another macro-iteration until certain criterion is met. The convergence criterion is chosen such that the maximum change in the CI coefficients in a macro-iteration is smaller than some threshold (e.g. 10^{-6}).

This method is implemented in subroutine **Nesbet**.

4 The active space one- and two-particle density matrices

Once the converged CI coefficients have been obtained, i.e. once the CAS-CI eigenvalue problem Eq. (34) has been solved, the active space one-particle density matrix (1PDM) γ and two-particle density matrix (2PDM) Γ should be constructed. The 1PDM and 2PDM are needed for the calculation of the orbital gradient \mathbf{g}^o and the orbital Hessian \mathbf{H}^{oo} . They are defined as

$$\text{1PDM: } \gamma_{tu} = \langle 0 | \hat{E}_{tu} | 0 \rangle = \sum_{KL} c_K c_L \langle K | \hat{E}_{tu} | L \rangle \quad (55)$$

$$\text{2PDM: } \Gamma_{tuvw} = \langle 0 | \hat{e}_{tuvw} | 0 \rangle = \sum_{KL} c_K c_L \langle K | \hat{E}_{tu} \hat{E}_{vw} - \delta_{uv} \hat{E}_{tw} | L \rangle \quad (56)$$

Subroutine **ASDMat** computes the 1PDM and 2PDM for a specified, converged CI vector. The algorithm is also based on the vectorizable loop structure for computing the CI matrix-vector product σ .

5 Matrix elements for \mathbf{g}^o , \mathbf{H}^{oo} , \mathbf{H}^{oc} , and \mathbf{H}^{cc}

5.1 Orbital gradient \mathbf{g}^o

The orbital gradient \mathbf{g}^o is given by

$$\begin{aligned}
g_{pq}^o &= -(1 - \hat{P}_{pq}) \langle 0 | [\hat{T}_{pq}, \hat{H}] | 0 \rangle \\
&= -2(1 - \hat{P}_{pq}) \left[\sum_r \langle 0 | \hat{E}_{pr} | 0 \rangle h_{qr} + \sum_{rst} \langle 0 | \hat{e}_{prst} | 0 \rangle (qr|st) \right] \\
&= -2(1 - \hat{P}_{pq}) \left[\sum_r \gamma_{pr} h_{qr} + \sum_{rst} \Gamma_{prst} (qr|st) \right] \\
&= -2(F_{pq} - F_{qp})
\end{aligned} \tag{57}$$

where \hat{P}_{pq} is the permutation operator that switches the indices p and q , and \mathbf{F} is the generalized Fock matrix and its elements are given by

$$F_{pq} = \sum_r \gamma_{pr} h_{qr} + \sum_{rst} \Gamma_{prst} (qr|st) \tag{58}$$

It is found that the generalized Fock matrix can be evaluated by blocks, due to the properties of the 1PDM and 2PDM:

$$F_{iq} = F_{iq}^1 \tag{59}$$

$$F_{tq} = F_{tq}^2 \tag{60}$$

$$F_{aq} = 0 \tag{61}$$

where \mathbf{F}^1 and \mathbf{F}^2 are defined as

$$F_{pq}^1 = 2 \left(h_{pq}^c + \sum_{tu} \gamma_{tu} [(pq|tu) - \frac{1}{2} (pt|qu)] \right) \tag{62}$$

$$F_{tq}^2 = \sum_u \gamma_{tu} h_{qu}^c + \sum_{uvw} \Gamma_{tuvw} (qu|vw) \tag{63}$$

Since orbital rotations within each of the inactive, active and virtual spaces are zero, the

nonzero blocks of the orbital gradient are: the inactive-active, the inactive-virtual, and the active-virtual. The upper triangular part of the orbital gradient matrix is:

$$g_{it}^o = -2(F_{it}^1 - F_{ti}^2) \quad (64)$$

$$g_{ia}^o = -2F_{ia}^1 \quad (65)$$

$$g_{ta}^o = -2F_{ta}^2 \quad (66)$$

Since the orbital gradient matrix is anti-symmetric, the lower triangular part can be readily obtained.

Subroutine **GenFck** computes the generalized Fock matrices \mathbf{F}^1 and \mathbf{F}^2 .

Subroutine **OGradS** computes the orbital gradient \mathbf{g}^o in a square matrix form, while subroutine **OGradL** returns the orbital gradient \mathbf{g}^o in a linearized, compact form that only contains the upper triangular part of the orbital gradient matrix.

5.2 Orbital Hessian \mathbf{H}^{oo}

Elements of the orbital Hessian \mathbf{H}^{oo} are given by:

$$\begin{aligned} \frac{1}{2}H_{it,ju}^{oo} = & \gamma_{tu}h_{ij}^c + \delta_{ij}(F_{tu}^1 - F_{tu}^2) - \delta_{tu}F_{ij}^1 + 2[4(ti|uj) - (tj|ui) - (tu|ij)] \\ & - \sum_v \gamma_{tv}[4(vi|uj) - (vj|ui) - (vu|ij)] - \sum_v \gamma_{uv}[4(ti|vj) - (tj|vi) - (tv|ij)] \\ & + \sum_{vw} [(ij|vw) \Gamma_{tuvw} + (iv|jw) (\Gamma_{tvuw} + \Gamma_{tvwu})] \end{aligned} \quad (67)$$

$$\frac{1}{2}H_{ia,jb}^{oo} = 2[4(ai|bj) - (ab|ij) - (aj|bi)] + \delta_{ij}F_{ab}^1 - \delta_{ab}F_{ij}^1 \quad (68)$$

$$\frac{1}{2}H_{ta,ub}^{oo} = \gamma_{tu}h_{ab}^c - \delta_{ab}F_{tu}^2 + \sum_{vw} [(ab|vw) \Gamma_{tuvw} + (av|bw) (\Gamma_{tvuw} + \Gamma_{tvwu})] \quad (69)$$

$$\begin{aligned} \frac{1}{2}H_{it,ja}^{oo} = & \delta_{ij}(F_{at}^1 - \frac{1}{2}F_{ta}^2) + 2[4(ti|aj) - (tj|ai) - (ta|ij)] \\ & - \sum_v \gamma_{tv}[4(vi|aj) - (vj|ai) - (va|ij)] \end{aligned} \quad (70)$$

$$\begin{aligned} \frac{1}{2}H_{it,ua}^{oo} = & -\gamma_{tu}h_{ai}^c + \frac{1}{2}F_{ai}^1 + \sum_v \gamma_{uv}[4(ti|av) - (tv|ai) - (ta|vi)] \\ & - \sum_{vw} [(ia|vw) \Gamma_{tuvw} + (iv|aw) (\Gamma_{tvuw} + \Gamma_{tvwu})] \end{aligned} \quad (71)$$

$$\frac{1}{2}H_{ia,tb}^{oo} = -\frac{1}{2}\delta_{ab}(F_{ti}^1 + F_{ti}^2) + \sum_v \gamma_{tv}[4(ai|bv) - (ab|vi) - (av|bi)] \quad (72)$$

Subroutine **OHessF** calculates the full orbital Hessian, and subroutine **OHessD** only gives the diagonal elements of the orbital Hessian.

5.3 State-orbital coupling Hessian \mathbf{H}^{oc}

The state-orbital coupling Hessian \mathbf{H}^{oc} is given by

$$\begin{aligned} H_{pq,n}^{oc} = & 2(1 - \hat{P}_{pq}) \left[\sum_r h_{pr} \langle 0 | \hat{E}_{rq} + \hat{E}_{qr} | n \rangle + \sum_{rst} (pr|st) \langle 0 | \hat{e}_{qrst} + \hat{e}_{rstq} | n \rangle \right] \\ = & 2(1 - \hat{P}_{pq}) \left[\sum_r h_{pr} \langle 0 | \hat{E}_{qr} | n \rangle + \langle n | \hat{E}_{qr} | 0 \rangle + \sum_{rst} (pr|st) \langle 0 | \hat{e}_{qrst} | n \rangle + \langle n | \hat{e}_{qrst} | 0 \rangle \right] \\ = & 4(1 - \hat{P}_{pq}) \left[\sum_r h_{pr} \gamma_{qr}^{(n)} + \sum_{rst} (pr|st) \Gamma_{qrst}^{(n)} \right] \end{aligned} \quad (73)$$

where \hat{P}_{pq} is the permutation operator that switches the indices p and q , and we have defined the transition 1PDMs and 2PDMs:

$$\gamma_{qr}^{(n)} = \frac{1}{2} \left[\langle 0 | \hat{E}_{qr} | n \rangle + \langle n | \hat{E}_{qr} | 0 \rangle \right] \quad (74)$$

$$\Gamma_{qrst}^{(n)} = \frac{1}{2} \left[\langle 0 | \hat{e}_{qrst} | n \rangle + \langle n | \hat{e}_{qrst} | 0 \rangle \right] \quad (75)$$

Define the transition Fock matrix $\mathbf{F}^{(n)}$:

$$F_{pq}^{(n)} = \sum_r \gamma_{pr}^{(n)} h_{qr} + \sum_{rst} (pr | st) \Gamma_{qrst}^{(n)} \quad (76)$$

Now the state-orbital coupling Hessian can be written as

$$H_{pq,n}^{oc} = -4(F_{pq}^{(n)} - F_{qp}^{(n)}) \quad (77)$$

Similar to the generalized Fock matrix in Eq. (58), the transition Fock matrix can also be split into three blocks:

$$F_{iq}^{(n)} = F_{iq}^{(n)1} = 2 \sum_{tu} \gamma_{tu}^{(n)} [(iq | tu) - \frac{1}{2} (it | qu)] \quad (78)$$

$$F_{tq}^{(n)} = F_{tq}^{(n)2} = \sum_u \gamma_{tu}^{(n)} h_{qu}^c + \sum_{uvw} \Gamma_{tuvw}^{(n)} (qu | vw) \quad (79)$$

$$F_{aq}^{(n)} = 0 \quad (80)$$

As can be seen, $\mathbf{F}^{(n)1}$ and $\mathbf{F}^{(n)2}$ have the same formulas as the matrices \mathbf{F}^1 and \mathbf{F}^2 , except that the transition density matrices are used and $\mathbf{F}^{(n)1}$ does not contain the inactive Hamiltonian \mathbf{h}^c .

According to Eq. (77), the state-orbital coupling Hessian can be constructed in the same way as the orbital gradient Eq. (57):

$$H_{pq,n}^{oc} = 2G_{pq}^{(n)} \quad (81)$$

The only difference is that in Eq. (81) the transition Fock matrices $\mathbf{F}^{(n)1}$ and $\mathbf{F}^{(n)2}$ are used.

Subroutine **TrnDen** calculates the transition density matrices $\boldsymbol{\gamma}^{(n)}$ and $\boldsymbol{\Gamma}^{(n)}$.

Subroutine **TrnFck** computes the transition Fock matrix $\mathbf{F}^{(n)}$.

5.4 The state Hessian \mathbf{H}^{cc}

Elements of the state Hessian \mathbf{H}^{cc} are given by

$$H_{nn'}^{cc} = 2\langle n|\hat{H} - E_0|n'\rangle \quad (82)$$

where n, n' are the orthogonal complement to $|0\rangle$.

If the CI diagonalization is performed before the orbital rotation, that is, $|0\rangle$ and $|n\rangle$ represent the eigenvectors of the CI Hamiltonian:

$$\hat{H}|0\rangle = E_0|0\rangle \quad (83)$$

$$\hat{H}|n\rangle = E_n|n\rangle, \quad n \neq 0 \quad (84)$$

then the state Hessian \mathbf{H}^{cc} becomes diagonal:

$$H_{nn'}^{cc} = 2(E_n - E_0)\delta_{nn'} \quad (85)$$

5.5 The state-orbital coupling contribution to the orbital Hessian

We now evaluate the state-orbital coupling contribution to the orbital Hessian, i.e. the second term in the partitioned orbital Hessian in Eq. (22):

$$\Delta\mathbf{H}^{oo} = -\mathbf{H}^{oc}\mathbf{H}^{cc-1}\mathbf{H}^{co} \quad (86)$$

We assume that the CI diagonalization is performed before the orbital rotation and thus $|0\rangle$ and $|n\rangle$ represent the eigenvectors of the CI Hamiltonian. The coupling correction to

the orbital Hessian can be expressed as

$$\begin{aligned}
\Delta H_{pq,rs}^{oo} &= - \left(\mathbf{H}^{oc} \mathbf{H}^{cc-1} \mathbf{H}^{co} \right)_{pq,rs} \\
&= - \sum_{n \neq 0} \frac{H_{pq,n}^{oc} H_{n,rs}^{co}}{2(E_n - E_0)} \\
&= - \sum_{n \neq 0} \frac{H_{pq,n}^{oc} H_{sr,n}^{oc}}{2(E_n - E_0)} \\
&= - \sum_{n \neq 0} \frac{2G_{pq}^{(n)} \cdot 2G_{sr}^{(n)}}{2(E_n - E_0)} \\
&= 2 \sum_{n \neq 0} \frac{G_{pq}^{(n)} G_{rs}^{(n)}}{E_n - E_0}
\end{aligned} \tag{87}$$

Subroutine **OHessP** calculates the state-orbital coupling contribution $\Delta \mathbf{H}^{oo}$ to the orbital Hessian.

References

- [1] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, Chichester, 2000.
- [2] R. K. Nesbet, J. Chem. Phys. **43**, 311 (1965).
- [3] C. Lanczos, J. Res. Natl. Bur. Stand. **45**, 255 (1950).
- [4] E. R. Davidson, J. Comput. Phys. **17**, 87 (1975).