Derivation of the Configuration Interaction Singles (CIS) Method for Various Single Determinant References and Extensions to Include Selected Double Substitutions (XCIS)

C. David Sherrill
School of Chemistry and Biochemistry
Georgia Institute of Technology

September 1996

1 Introduction

The configuration interaction expansion of an approximate solution to the electronic Schrödinger equation is typically written

$$|\Psi\rangle = c_0|\Phi_0\rangle + \sum_{ia} c_i^a|\Phi_i^a\rangle + \sum_{i<j,a<b} c_{ij}^{ab}|\Phi_{ij}^{ab}\rangle + \sum_{i<j<k,a<b<c} c_{ijk}^{abc}|\Phi_{ijk}^{abc}\rangle + \ldots \quad (1)$$

$|\Phi_0\rangle$ is the so-called "reference," typically obtained from a Hartree-Fock self-consistent-field (SCF) procedure as the best single Slater determinant (or configuration state function, CSF) which describes the electronic state of interest. $|\Phi_i^a\rangle$ is the determinant formed by replacing spin-orbital i in $|\Phi_0\rangle$ with spin orbital a, etc. These notes follow the convention that i, j, k denote orbitals occupied in the reference, a, b, c denote orbitals unoccupied in the reference, and p, q, r are general indices. The widely-employed CI singles and doubles (CISD) wavefunction includes only those N-electron basis functions which represent single or double substitutions relative to the reference state and typically accounts for about 95% of the correlation energy for small molecules near their equilibrium geometries.

Head-Gordon, Pople, and others have advocated the use of configuration interaction with only single substitutions (CIS) as the starting point for investigations of excited electronic states. In their 1992 paper, Foresman, Head-Gordon, Pople, and Frisch [1] list the following desirable properties of CIS: well defined (and differentiable), applicable to large systems, size-consistent, variational, and providing directly comparable (i.e. orthogonal) electronic state solutions. They go on to present equations for the CIS energy and gradient when the reference is a single determinant obtained from an SCF procedure. These notes present a derivation of the CIS energy for general and several specific types of single-determinant references. Later, we discuss some extensions which are necessary for reliable treatments of open-shell systems.

2 CIS Energy Equations

Since there are only two types of determinants according to excitation level (i.e., the reference and single excitations), there are only two relevant types of matrix elements, $\langle\Phi_0|\hat{H}|\Phi_i^a\rangle$ and $\langle\Phi_i^a|\hat{H}|\Phi_j^b\rangle$. Assuming that the determinants are made up of a common set of orthonormal spin orbitals, these matrix elements may be evaluated using Slater's rules. The first is given by

$$\langle\Phi_0|\hat{H}|\Phi_i^a\rangle = h_{ia} + \sum_{k\in\Phi_0}\langle ik||ak\rangle = F_{ia}, \tag{2}$$

where the Fock matrix element F_{pq} is defined as

$$F_{pq} = h_{pq} + \sum_{k\in\Phi_0}\langle pk||qk\rangle. \tag{3}$$

The other relevant matrix elements are of the form $\langle\Phi_i^a|\hat{H}|\Phi_j^b\rangle$. The singly excited determinants may differ from each other by two spin orbitals if $i \neq j$ and $a \neq b$. If so, the determinants are already in maximum coincidence and the matrix element is of the form

$$\langle\cdots a\cdots j\cdots|\hat{H}|\cdots i\cdots b\cdots\rangle \tag{4}$$

or

$$\langle\cdots j\cdots a\cdots|\hat{H}|\cdots b\cdots i\cdots\rangle \tag{5}$$

The matrix elements are $\langle aj||ib\rangle$ and $\langle ja||bi\rangle$, respectively, and these antisymmetrized integrals are equal to each other (and also to $-\langle ja||ib\rangle$ and $-\langle aj||bi\rangle$).

For the case $i = j$, $a \neq b$, the matrix elements are

$$\langle\Phi_i^a|\hat{H}|\Phi_i^b\rangle = \langle\cdots a\cdots|\hat{H}|\cdots b\cdots\rangle = h_{ab} + \sum_{k\in\Phi_0, k\neq i}\langle ak||bk\rangle = F_{ab} - \langle ai||bi\rangle. \tag{6}$$

Likewise, for the case $i \neq j$, $a = b$, the matrix elements are

$$\langle\Phi_i^a|\hat{H}|\Phi_j^a\rangle = \langle\cdots i\cdots a\cdots|\hat{H}|\cdots a\cdots j\cdots\rangle = -h_{ij} - \sum_{k\in\{\Phi_0+a\}}\langle ik||jk\rangle = -F_{ij} - \langle ia||ja\rangle. \tag{7}$$

Finally, when $i = j$ and $a = b$,

$$\begin{aligned}\langle\Phi_i^a|\hat{H}|\Phi_i^a\rangle &= \sum_{k\in\Phi_0} h_{kk} + \frac{1}{2}\sum_{k,l\in\Phi_0}\langle kl||kl\rangle - h_{ii} + h_{aa} - \sum_{k\in\Phi_0}\langle ki||ki\rangle + \sum_{k\in\Phi_0}\langle ka||ka\rangle - \langle ia||ia\rangle \quad (8)\\ &= E_0 - F_{ii} + F_{aa} - \langle ia||ia\rangle,\end{aligned}$$

where $E_0 = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle$; if $|\Phi_0\rangle$ is obtained by an SCF procedure, then this is the SCF energy.

Using the permutational symmetries of the antisymmetrized two-electron integrals, the two-electron terms for the preceding four cases can be combined (rearranging the integral $\langle ia||ja\rangle$ requires the assumption that the orbitals are real). This yields the final, compact result

$$\langle \Phi_i^a | \hat{H} | \Phi_j^b \rangle = E_0 \delta_{ij} \delta_{ab} + F_{ab} \delta_{ij} - F_{ij} \delta_{ab} + \langle aj || ib \rangle. \tag{9}$$

This is equation (11) of Maurice and Head-Gordon [2], who extended the CIS method to the case of restricted open-shell (ROHF) and unrestricted (UHF) reference determinants. Note that E_0 occurs along the diagonal of the entire matrix $\mathbf{H}$; this means that we can subtract E_0 before diagonalizing and add it back later to each of the eigenvalues. If all matrix elements $F_{ia} = 0$, as they often are, then the reference determinant does not mix with any of the excited determinants and $|\Phi_0\rangle$ is already an eigenfunction of the CIS Hamiltonian with eigenvalue E_0; furthermore, the eigenvalues of the CIS Hamiltonian less the E_0 diagonal terms represent excitation energies. From this point onward, E_0 will be subtracted from the Hamiltonian.

Given the above matrix elements, it remains to write down the CIS energy expression. Recall that the CIS wavefunction is expanded as

$$|\Psi\rangle = c_0 |\Phi_0\rangle + \sum_{ia} c_i^a |\Phi_i^a\rangle. \tag{10}$$

Assuming real CI coefficients, the energy is given by

$$E_{\text{CIS}} = E_0 + 2 \sum_{ia} c_0 c_i^a F_{ia} + \sum_{iab} c_i^a c_i^b F_{ab} - \sum_{ija} c_i^a c_j^a F_{ij} + \sum_{ijab} c_i^a c_j^b \langle aj || ib \rangle. \tag{11}$$

For a closed-shell SCF reference $|\Phi_0\rangle$, off-diagonal terms of the Fock matrix vanish, and the expression becomes

$$E_{\text{CIS}} = E_{\text{SCF}} + \sum_{ia} (c_i^a)^2 (\epsilon_i - \epsilon_j) + \sum_{ijab} c_i^a c_j^b \langle aj || ib \rangle, \tag{12}$$

which matches equation (2.15) of Foresman *et al.* [1] once the two-electron integral is rearranged. Of course, this equation is only useful once the CI coefficients are known. In general, the lowest several eigenvectors are of interest in a CIS study; these can be obtained by iteratively diagonalizing the CIS Hamiltonian using Davidson's method [3] or the Davidson-Liu Simultaneous Expansion Method [4]. Iterative solution calls for diagonalization of the Hamiltonian in a small subspace of trial vectors, with the set of vectors being expanded every iteration until convergence. This requires calculation of the following quantities, usually called the σ vectors:

$$\sigma_I = \sum_J H_{IJ} c_J. \tag{13}$$

One σ vector must be computed for each $\mathbf{c}$ in the set of trial vectors. For CIS, σ can be written

$$\begin{aligned}\sigma_0 &= \sum_{jb} \langle \Phi_0 | \overline{H} | \Phi_j^b \rangle c_j^b && (14)\\ \sigma_i^a &= \langle \Phi_{ia} | \overline{H} | \Phi_0 \rangle c_0 + \sum_{jb} \langle \Phi_i^a | \overline{H} | \Phi_j^b \rangle c_j^b, && (15)\end{aligned}$$

where the bar over H is a reminder that E_0 has been subtracted from the Hamiltonian. These expressions can be expanded to

$$\begin{aligned}\sigma_0 &= \sum_{jb} c_j^b F_{jb} && (16)\\ \sigma_{ia} &= c_0 F_{ia} + \sum_{jb} c_j^b \left[F_{ab}\delta_{ij} - F_{ij}\delta_{ab} + \langle aj || ib \rangle \right]. && (17)\end{aligned}$$

The above equations make it clear that σ can be computed directly from the one- and two-electron integrals without the need to explicitly compute or store the one- and two-electron coupling coefficients γ_{pq}^{IJ} and Γ_{pqrs}^{IJ} as separate quantities; this makes the CIS method a *direct CI* procedure. As noted by Foresman *et al.* [1], the CIS iterations can actually be performed in a "double-direct" fashion; i.e., the integrals can also be computed on-the-fly as needed. As shown by Maurice and Head-Gordon [2], the contribution of the two-electron integrals to σ_i^a can be written as a Fock-like matrix,

$$\begin{aligned}\tilde{F}_{ia} &= \sum_{jb} c_j^b \langle aj || ib \rangle && (18)\\ &= \sum_{\mu\nu} C_{\mu a}^* C_{\nu i} \sum_{\lambda\sigma} \langle \mu\lambda || \nu\sigma \rangle \sum_{jb} C_{\lambda j}^* c_j^b C_{\sigma b}, && (19)\end{aligned}$$

where $C_{\mu i}$ are the coefficients defining the transformation from atomic orbitals (AOs) to molecular orbitals (MOs). Evaluation of $\tilde{F}_{ia}$ can be carried out as a series of matrix multiplies. The pseudodensity matrix,

$$\tilde{P}_{\lambda\sigma} = \sum_{jb} C_{\lambda j}^* c_j^b C_{\sigma b}, \qquad (20)$$

can be multiplied by the two-electron integrals as they are formed in the atomic orbital basis to yield the AO Fock-like matrix,

$$\tilde{F}_{\mu\nu} = \sum_{\lambda\sigma} \langle \mu\lambda || \nu\sigma \rangle \tilde{P}_{\lambda\sigma}, \qquad (21)$$

which is transformed back into the MO basis by

$$\tilde{F}_{ia} = \sum_{\mu\nu} C_{\mu a}^* \tilde{F}_{\mu\nu} C_{\nu i}. \qquad (22)$$

2.1 Restricted Hartree-Fock References

Now consider the case of a closed-shell RHF reference determinant. In this case, $F_{pq} = \delta_{pq}\epsilon_p$, so that σ_0 becomes zero. σ_{ia} simplifies to

$$\sigma_{ia} = c_i^a(\epsilon_a - \epsilon_i) + \sum_{jb} c_j^b \langle aj||ib\rangle. \tag{23}$$

Conserving M_s requires that the spins of j and b are equal. Therefore,

$$\sigma_{ia} = c_i^a(\epsilon_a - \epsilon_i) + \sum_{jb} c_j^b \langle aj||ib\rangle + \sum_{\bar{j}\bar{b}} c_{\bar{j}}^{\bar{b}} \langle a\bar{j}||i\bar{b}\rangle. \tag{24}$$

After integrating over spin, this becomes in chemists' notation

$$\sigma_{ia} = c_i^a(\epsilon_a - \epsilon_i) + \sum_{jb} c_j^b[(ai|jb) - (ab|ji)] + \sum_{\bar{j}\bar{b}} c_{\bar{j}}^{\bar{b}}(ai|jb). \tag{25}$$

Time-reversal symmetry imposes certain conditions on the CI coefficients. In alpha and beta string notation [5], for $M_s = 0$ cases,

$$c(I_\alpha, I_\beta) = (-1)^S c(I_\beta, I_\alpha). \tag{26}$$

An analogous equation also holds for σ. This means that for a closed shell RHF reference, $c_i^a = c_{\bar{i}}^{\bar{a}}$ ($\sigma_i^a = \sigma_{\bar{i}}^{\bar{a}}$) for singlets, and $c_i^a = -c_{\bar{i}}^{\bar{a}}$ ($\sigma_i^a = -\sigma_{\bar{i}}^{\bar{a}}$) for triplets; thus only half of the CI coefficients must be computed explicitly. These sign rules are also evident from the observation that $|\Phi_i^a\rangle$ and $|\Phi_{\bar{i}}^{\bar{a}}\rangle$ are not spin eigenfunctions, but that the total CI wavefunction will be a spin eigenfunction (if the required determinants are present in the CI space). Using the determinant sign convention of Szabo and Ostlund [6], spin eigenfunctions (or CSFs) associated with the above determinants are

$$\begin{aligned} |{}^1\Phi_i^a\rangle &= \frac{1}{\sqrt{2}}\left(|\Phi_i^a\rangle + |\Phi_{\bar{i}}^{\bar{a}}\rangle\right) && (27) \\ |{}^3\Phi_i^a\rangle &= \frac{1}{\sqrt{2}}\left(|\Phi_i^a\rangle - |\Phi_{\bar{i}}^{\bar{a}}\rangle\right). && (28) \end{aligned}$$

Using these relationships between CI coefficients, we obtain

$$\begin{aligned} {}^1\sigma_i^a(\text{RCIS}) &= c_i^a(\epsilon_a - \epsilon_i) + \sum_{jb} c_j^b\,[2(ai|jb) - (ab|ji)] && (29) \\ {}^3\sigma_i^a(\text{RCIS}) &= c_i^a(\epsilon_a - \epsilon_i) - \sum_{jb} c_j^b(ab|ji). && (30) \end{aligned}$$

Furthermore, ${}^1\sigma_{\bar{i}}^{\bar{a}}(\text{RCIS}) ={}^1 \sigma_i^a(\text{RCIS})$ and ${}^3\sigma_{\bar{i}}^{\bar{a}}(\text{RCIS}) = -{}^3\sigma_i^a(\text{RCIS})$.

Consider how equations (20)-(22) change when spin is explicitly accounted for. There will be two pseudodensity matrices,

$$\tilde{P}^{\alpha}_{\lambda\sigma} = \sum_{jb} C^{*}_{\lambda j} c^{b}_{j} C_{\sigma b} \tag{31}$$

$$\tilde{P}^{\beta}_{\lambda\sigma} = \sum_{\overline{j}\overline{b}} C^{*}_{\lambda j} c^{\overline{b}}_{\overline{j}} C_{\sigma b}. \tag{32}$$

Due to equation (26), $\tilde{P}^{\alpha}_{\lambda\sigma} = \tilde{P}^{\beta}_{\lambda\sigma}$ for singlets, and $\tilde{P}^{\alpha}_{\lambda\sigma} = -\tilde{P}^{\beta}_{\lambda\sigma}$ for triplets. Thus it is necessary to form only one of the Fock-like matrices, $\tilde{F}^{\alpha}_{\mu\nu}$ or $\tilde{F}^{\beta}_{\mu\nu}$. The former is constructed as

$${}^{1}\tilde{F}^{\alpha}_{\mu\nu} = \sum_{\lambda\sigma} \left[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)\right] \tilde{P}^{\alpha}_{\lambda\sigma} \tag{33}$$

$${}^{3}\tilde{F}^{\alpha}_{\mu\nu} = -\sum_{\lambda\sigma} (\mu\sigma|\lambda\nu) \tilde{P}^{\alpha}_{\lambda\sigma} \tag{34}$$

Finally $\tilde{F}^{\alpha}_{ia}$ is constructed according to eq (22). In terms of these quantities, the σ vector can be written

$${}^{1}\sigma^{a}_{i}(\mathrm{RCIS}) = {}^{1}\sigma^{\overline{a}}_{\overline{i}}(\mathrm{RCIS}) = c^{a}_{i}(\epsilon_a - \epsilon_i) + {}^{1}\tilde{F}^{\alpha}_{ia} \tag{35}$$

$${}^{3}\sigma^{a}_{i}(\mathrm{RCIS}) = -{}^{3}\sigma^{\overline{a}}_{\overline{i}}(\mathrm{RCIS}) = c^{a}_{i}(\epsilon_a - \epsilon_i) + {}^{3}\tilde{F}^{\alpha}_{ia}, \tag{36}$$

2.2 Unrestricted Hartree-Fock References

Now consider the case when $|\Phi_0\rangle$ is obtained by the UHF procedure. Once again, the Fock matrix is diagonal: $F_{pq} = \delta_{pq}\epsilon_p$. However, it is customary to split the Fock matrix into two matrices, one for α and one for β spin orbitals (mixed terms such as $F_{\overline{p}q}$ and $F_{p\overline{q}}$ are zero). As for the RHF case, $\sigma_0 = 0$ and σ^{a}_{i} can be written

$$\sigma^{a}_{i} = c^{a}_{i}(\epsilon_a - \epsilon_i) + \sum_{jb} c^{b}_{j} \langle aj||ib\rangle + \sum_{\overline{j}\overline{b}} c^{\overline{b}}_{\overline{j}} \langle a\overline{j}||i\overline{b}\rangle. \tag{37}$$

Unfortunately, for a UHF reference equations (26)-(28) no longer hold, so that the above equation can be simplified only to

$$\sigma^{a}_{i} = c^{a}_{i}(\epsilon_a - \epsilon_i) + \sum_{jb} c^{b}_{j} \left[(ai|jb) - (ab|ji)\right] + \sum_{\overline{j}\overline{b}} c^{\overline{b}}_{\overline{j}} (ai|\overline{j}\overline{b}). \tag{38}$$

Even though spin has been integrated out, the overbars must be kept in the above equation because the spatial parts of spin orbitals i and $\overline{i}$ are not necessarily equal. An analogous equation holds for $\sigma^{\overline{a}}_{\overline{i}}$,

$$\sigma^{\overline{a}}_{\overline{i}} = c^{\overline{a}}_{\overline{i}}(\epsilon_{\overline{a}} - \epsilon_{\overline{i}}) + \sum_{\overline{j}\overline{b}} c^{\overline{b}}_{\overline{j}} [(\overline{a}\overline{i}|\overline{j}\overline{b}) - (\overline{a}\overline{b}|\overline{j}\overline{i})] + \sum_{jb} c^{b}_{j} (\overline{a}\overline{i}|jb), \tag{39}$$

and there is no general relationship between σ_i^a and $\sigma_{\bar{i}}^{\bar{a}}$. The pseudodensity matrices $\tilde{P}^{\alpha}_{\lambda\sigma}$ and $\tilde{P}^{\beta}_{\lambda\sigma}$ are defined as in eq. (31) and (32), but they are no longer simply related. Thus it is necessary to compute two Fock-like matrices, according to

$$\begin{aligned} \tilde{F}^{\alpha}_{\mu\nu} &= \sum_{\lambda\sigma} \left[(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)\right] \tilde{P}^{\alpha}_{\lambda\sigma} + (\mu\nu|\lambda\sigma)\tilde{P}^{\beta}_{\lambda\sigma} & (40) \\ \tilde{F}^{\beta}_{\mu\nu} &= \sum_{\lambda\sigma} \left[(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)\right] \tilde{P}^{\beta}_{\lambda\sigma} + (\mu\nu|\lambda\sigma)\tilde{P}^{\alpha}_{\lambda\sigma}. & (41) \end{aligned}$$

In contrast to RCIS, there are no longer separate singlet and triplet cases, since the UCIS eigenfunctions are not CSFs. After transforming to the MO basis by

$$\begin{aligned} \tilde{F}^{\alpha}_{ia} &= \sum_{\mu\nu} C^{*}_{\mu a}\tilde{F}^{\alpha}_{\mu\nu}C_{\nu i} & (42) \\ \tilde{F}^{\beta}_{\bar{i}\bar{a}} &= \sum_{\mu\nu} C^{*}_{\mu\bar{a}}\tilde{F}^{\beta}_{\mu\nu}C_{\nu\bar{i}}, & (43) \end{aligned}$$

the expressions for σ become

$$\begin{aligned} \sigma_i^a(\mathrm{UCIS}) &= c_i^a(\epsilon_a - \epsilon_i) + \tilde{F}^{\alpha}_{ia} & (44) \\ \sigma_{\bar{i}}^{\bar{a}}(\mathrm{UCIS}) &= c_{\bar{i}}^{\bar{a}}(\epsilon_{\bar{a}} - \epsilon_{\bar{i}}) + \tilde{F}^{\beta}_{\bar{i}\bar{a}}. & (45) \end{aligned}$$

2.3 Restricted Open-Shell Hartree-Fock References

A single-determinant restricted open-shell Hartree-Fock (ROHF) wavefunction describing a high-spin open-shell system will be an eigenfunction of $\hat{S}^2$ (i.e., a CSF). This is easy to verify by direct application of the $\hat{S}^2$ operator, which is

$$\begin{aligned} \hat{S}^2 &= \hat{S}\cdot\hat{S} = \sum_i^N \sum_j^N \hat{s}(i)\cdot\hat{s}(j) \\ &= \hat{S}_-\hat{S}_+ + \hat{S}_z + \hat{S}_z^2, & (46) \end{aligned}$$

where

$$\hat{S}_z = \sum_i^N \hat{s}_z(i) \qquad (47)$$

$$\hat{S}_z^2 = \sum_i^N \hat{s}_z^2(i) \qquad (48)$$

$$\hat{S}_\pm = \sum_i^N \hat{s}_\pm(i) \qquad (49)$$

and

$$\begin{array}{lll} \hat{s}_z(i)a_i^\dagger|\rangle = \frac{1}{2}a_i^\dagger|\rangle & \hat{s}_+(i)a_i^\dagger|\rangle = 0 & \hat{s}_-(i)a_i^\dagger|\rangle = a_{\bar{i}}^\dagger|\rangle \\ \hat{s}_z(i)a_{\bar{i}}^\dagger|\rangle = -\frac{1}{2}a_{\bar{i}}^\dagger|\rangle & \hat{s}_-(i)a_{\bar{i}}^\dagger|\rangle = 0 & \hat{s}_+(i)a_{\bar{i}}^\dagger|\rangle = a_i^\dagger|\rangle. \end{array} \tag{50}$$

The high-spin ROHF wavefunction can be written as

$$|\Phi_{ROHF}\rangle = \prod_i^{\text{docc}} a_i^\dagger a_{\bar{i}}^\dagger \prod_t^{\text{socc}} a_t^\dagger|\rangle, \tag{51}$$

where t and u will denote open-shells. Applying $\hat{S}^2$, this becomes

$$\hat{S}^2|\Phi_{ROHF}\rangle = \left[\hat{S}_-\hat{S}_+ + \hat{S}_z + \hat{S}_z^2\right] \prod_i^{\text{docc}} a_i^\dagger a_{\bar{i}}^\dagger \prod_t^{\text{socc}} a_t^\dagger|\rangle. \tag{52}$$

This is easy to evaluate:

$$\hat{S}_z \prod_i^{\text{docc}} a_i^\dagger a_{\bar{i}}^\dagger \prod_t^{\text{socc}} a_t^\dagger|\rangle = \left[\frac{1}{2}(N_\alpha - N_\beta)\right] \prod_i^{\text{docc}} a_i^\dagger a_{\bar{i}}^\dagger \prod_t^{\text{socc}} a_t^\dagger|\rangle \tag{53}$$

$$\hat{S}_z^2 \prod_i^{\text{docc}} a_i^\dagger a_{\bar{i}}^\dagger \prod_t^{\text{socc}} a_t^\dagger|\rangle = \left[\frac{1}{2}(N_\alpha - N_\beta)\right]^2 \prod_i^{\text{docc}} a_i^\dagger a_{\bar{i}}^\dagger \prod_t^{\text{socc}} a_t^\dagger|\rangle. \tag{54}$$

The raising operator $\hat{S}_+$ yields zero when acting on $|\Phi_{ROHF}\rangle$: raising operators applied to α electrons always yield zero, and raising operators applied to the β electrons yield α spin orbitals which are already occupied (so the determinant vanishes by the Pauli principle). Hence the final result is

$$\begin{aligned} \hat{S}^2|\Phi_{ROHF}\rangle &= \left[\frac{1}{4}(N_\alpha - N_\beta)^2 + \frac{1}{2}(N_\alpha - N_\beta)\right]|\Phi_{ROHF}\rangle \\ &= \left[\frac{1}{2}N_s\left(\frac{1}{2}N_s + 1\right)\right]|\Phi_{ROHF}\rangle, \end{aligned} \tag{55}$$

where $N_s = N_\alpha - N_\beta$.

Now it is worthwhile to consider how to form CSFs for the single excitations. Determinants which promote electrons from the singly-occupied space to the virtual space, as well as determinants which promote β electrons in the doubly-occupied orbitals to the singly-occupied orbitals, are already spin-adapted (the proof is completely analogous to that above for the ROHF reference determinant). The only other relevant single excitations are those which move an electron from a doubly-occupied orbital to a virtual orbital. These determinants are *not* spin-adapted, as we will proceed to demonstrate. Consider the action of $\hat{S}^2$ on the determinant $|\Phi_i^a\rangle$:

$$\left[\hat{S}_-\hat{S}_+ + \hat{S}_z + \hat{S}_z^2\right] a_a^\dagger a_i \prod_j^{\text{docc}} a_j^\dagger a_{\bar{j}}^\dagger \prod_t^{\text{socc}} a_t^\dagger|\rangle. \tag{56}$$

The result of $\hat{S}_z + \hat{S}_z^2$ is easily determined to be

$$\left[\hat{S}_z + \hat{S}_z^2\right] |\Phi_i^a\rangle = \left[\frac{N_s}{2}\left(\frac{N_s}{2} + 1\right)\right] |\Phi_i^a\rangle. \tag{57}$$

Now all that remains is the raising and lowering operators. These are somewhat more involved and require that attention be paid to the sign.

$$\hat{S}_+ |\Phi_i^a\rangle = \hat{S}_+ a_a^\dagger a_i \prod_j^{\text{docc}} a_j^\dagger a_{\bar{j}}^\dagger \prod_t^{\text{socc}} a_t^\dagger |\rangle. \tag{58}$$

By arguments similar to those presented above, all raising operators yield zero except for $\hat{s}_+(\bar{i})$. Using the anticommutation relations for creation and annihilation operators,

$$\begin{aligned} \hat{s}_+(\bar{i}) a_a^\dagger a_i |\Phi_0\rangle &= \hat{s}_+(\bar{i}) \left(a_{\bar{i}}^\dagger a_{\bar{i}} + a_{\bar{i}} a_{\bar{i}}^\dagger\right) a_a^\dagger a_i |\Phi_0\rangle \\ &= \hat{s}_+(\bar{i}) a_{\bar{i}}^\dagger a_{\bar{i}} a_a^\dagger a_i |\Phi_0\rangle \\ &= a_i^\dagger a_{\bar{i}} a_a^\dagger a_i |\Phi_0\rangle \\ &= -a_a^\dagger a_{\bar{i}} |\Phi_0\rangle. \end{aligned} \tag{59}$$

The $\hat{S}_-$ operator can now affect electrons in any of the following orbitals: a, i, and any of the open-shell orbitals.

$$\begin{aligned} \hat{S}_- \left[-a_a^\dagger a_{\bar{i}} |\Phi_0\rangle\right] &= -a_{\bar{a}}^\dagger a_{\bar{i}} |\Phi_0\rangle + a_a^\dagger a_i |\Phi_0\rangle + \sum_t^{\text{socc}} a_a^\dagger a_t a_{\bar{t}}^\dagger a_{\bar{i}} |\Phi_0\rangle \\ &= |\Phi_i^a\rangle - |\Phi_{\bar{i}}^{\bar{a}}\rangle + \sum_t^{\text{socc}} |\Phi_{t\bar{i}}^{a\bar{t}}\rangle. \end{aligned} \tag{60}$$

Thus overall,

$$\hat{S}^2 |\Phi_i^a\rangle = \left[\frac{N_s}{2}\left(\frac{N_s}{2} + 1\right) + 1\right] |\Phi_i^a\rangle - |\Phi_{\bar{i}}^{\bar{a}}\rangle + \sum_t^{\text{socc}} |\Phi_{t\bar{i}}^{a\bar{t}}\rangle. \tag{61}$$

The analogous equation for $|\Phi_{\bar{i}}^{\bar{a}}\rangle$ is

$$\hat{S}^2 |\Phi_{\bar{i}}^{\bar{a}}\rangle = \left[\frac{N_s}{2}\left(\frac{N_s}{2} + 1\right) + 1\right] |\Phi_{\bar{i}}^{\bar{a}}\rangle - |\Phi_i^a\rangle - \sum_t^{\text{socc}} |\Phi_{t\bar{i}}^{a\bar{t}}\rangle. \tag{62}$$

Clearly a spin eigenfunction can be constructed as

$$|{}^{(N_s+1)}\Phi_i^a\rangle = \frac{1}{\sqrt{2}}\left(|\Phi_i^a\rangle + |\Phi_{\bar{i}}^{\bar{a}}\rangle\right). \tag{63}$$

Then the operation of $\hat{S}^2$ is

$$\hat{S}^2 |{}^{(N_s+1)}\Phi_i^a\rangle = \left[\frac{N_s}{2}\left(\frac{N_s}{2} + 1\right)\right] |{}^{(N_s+1)}\Phi_i^a\rangle. \tag{64}$$

Now that the relevant CSFs have been obtained, they can be used to define the ROHF convergence criteria: the final ROHF wavefunction will not mix with any of the singly substituted CSFs. Thus

$$\begin{array}{rl} \langle \Phi_{ROHF}|\hat{H}|\Phi_t^a\rangle = 0 & \text{socc} \rightarrow \text{virt} \\ \langle \Phi_{ROHF}|\hat{H}|\Phi_{\bar{i}}^{\bar{t}}\rangle = 0 & \text{docc} \rightarrow \text{socc} \\ \langle \Phi_{ROHF}|\hat{H}|^{(N_s+1)}\Phi_i^a\rangle = 0 & \text{docc} \rightarrow \text{virt}, \end{array} \tag{65}$$

which implies the following conditions on the Fock matrix elements

$$F_{ta} = 0 \tag{66}$$

$$F_{\bar{i}\bar{t}} = 0 \tag{67}$$

$$F_{ia} = -F_{\bar{i}\bar{a}}. \tag{68}$$

Using these results, we can write down expressions for the σ vectors. Since determinants $|\Phi_i^a\rangle$ must enter with the same coefficients as $|\Phi_{\bar{i}}^{\bar{a}}\rangle$, $\sigma_0 = 0$ once again. Furthermore, since the ROHF reference cannot mix with any other singly excited configurations, the c_0 contribution to σ_{ia} and $\sigma_{\bar{i}\bar{a}}$ may be safely ignored. We will therefore consider four cases: σ_i^a, $\sigma_{\bar{i}}^{\bar{a}}$, σ_t^a, and $\sigma_{\bar{i}}^{\bar{t}}$, where once again t, u represent singly occupied orbitals.

The equation for σ_i^a is readily seen to be

$$\begin{aligned} \sigma_i^a &= \sum_{jb} c_j^b \left[F_{ab}\delta_{ij} - F_{ij}\delta_{ab} + \langle aj||ib\rangle\right] + \sum_{\bar{j}\bar{b}} c_{\bar{j}}^{\bar{b}} \langle a\bar{j}||i\bar{b}\rangle \\ &+ \sum_{tb} c_t^b \left[-F_{it}\delta_{ab} + \langle at||ib\rangle\right] + \sum_{\bar{j}\bar{t}} c_{\bar{j}}^{\bar{t}} \langle a\bar{j}||i\bar{t}\rangle. \end{aligned} \tag{69}$$

Separating the Fock operator terms from the two-electron integrals, and integrating out spin, this yields

$$\begin{aligned} \sigma_i^a &= \sum_{jb} c_j^b \left[F_{ab}\delta_{ij} - F_{ij}\delta_{ab}\right] - \sum_{tb} c_t^b F_{it}\delta_{ab} + \sum_{jb} c_j^b \left[2(ai|jb) - (ab|ji)\right] \\ &+ \sum_{tb} c_t^b \left[2(ai|tb) - (ab|ti)\right] + \sum_{\bar{j}\bar{t}} c_{\bar{j}}^{\bar{t}} (ai|jt), \end{aligned} \tag{70}$$

where we have used the relation $c_j^b = c_{\bar{j}}^{\bar{b}}$. The analogous equation for β spins is

$$\begin{aligned} \sigma_{\bar{i}}^{\bar{a}} &= \sum_{\bar{j}\bar{b}} c_{\bar{j}}^{\bar{b}} \left[F_{\bar{a}\bar{b}}\delta_{\bar{i}\bar{j}} - F_{\bar{i}\bar{j}}\delta_{\bar{a}\bar{b}}\right] + \sum_{\bar{j}\bar{t}} c_{\bar{j}}^{\bar{t}} F_{\bar{a}\bar{t}}\delta_{\bar{i}\bar{j}} + \sum_{\bar{j}\bar{b}} c_{\bar{j}}^{\bar{b}} \left[2(ai|jb) - (ab|ji)\right] \\ &+ \sum_{tb} c_t^b (ai|tb) + \sum_{\bar{j}\bar{t}} c_{\bar{j}}^{\bar{t}} \left[2(ai|jt) - (at|ji)\right], \end{aligned} \tag{71}$$

If the equality $c_j^b = c_{\bar{j}}^{\bar{b}}$ is to be maintained, we must have $\sigma_i^a = \sigma_{\bar{i}}^{\bar{a}}$. It is then computationally convenient to form these quantities as

$$\sigma_i^a = \sigma_{\bar{i}}^{\bar{a}} = \frac{1}{2}(\sigma_i^a + \sigma_{\bar{i}}^{\bar{a}}). \tag{72}$$

Thus

$$\begin{aligned}
2\sigma_i^a = 2\sigma_{\bar{i}}^{\bar{a}} &= \sum_b c_i^b \left[F_{ab} + F_{\bar{a}\bar{b}}\right] - \sum_j c_j^a \left[F_{ij} + F_{\bar{i}\bar{j}}\right] - \sum_t c_t^a F_{it} + \sum_{\bar{t}} c_{\bar{i}}^{\bar{t}} F_{\bar{a}\bar{t}} \\
&+ \sum_{jb} c_j^b \left[2(ai|jb) - (ab|ji)\right] + \sum_{\bar{j}\bar{b}} c_{\bar{j}}^{\bar{b}} \left[2(ai|jb) - (ab|ji)\right] \\
&+ \sum_{tb} c_t^b \left[2(ai|tb) - (ab|ti)\right] + \sum_{\bar{j}\bar{t}} c_{\bar{j}}^{\bar{t}} \left[2(ai|jt) - (at|ji)\right]. \qquad (73)
\end{aligned}$$

Now it is clear that the two-electron integrals can be treated all together. To begin condensing the notation once again, let us define the following quantities:

$$\begin{aligned}
\tilde{P}_{\lambda\sigma}^{DV} &= \sum_{jb} C_{\lambda j}^* c_j^b C_{\sigma b} && (74) \\
\tilde{P}_{\lambda\sigma}^{\overline{DV}} &= \sum_{\bar{j}\bar{b}} C_{\lambda\bar{j}}^* c_{\bar{j}}^{\bar{b}} C_{\sigma\bar{b}} && (75) \\
\tilde{P}_{\lambda\sigma}^{SV} &= \sum_{tb} C_{\lambda t}^* c_t^b C_{\sigma b} && (76) \\
\tilde{P}_{\lambda\sigma}^{\overline{DS}} &= \sum_{\bar{j}\bar{t}} C_{\lambda\bar{j}}^* c_{\bar{j}}^{\bar{t}} C_{\sigma\bar{t}} && (77) \\
\tilde{P}_{\lambda\sigma}^{+} &= \frac{1}{2}\left[\tilde{P}_{\lambda\sigma}^{DV} + \tilde{P}_{\lambda\sigma}^{\overline{DV}} + \tilde{P}_{\lambda\sigma}^{SV} + \tilde{P}_{\lambda\sigma}^{\overline{DS}}\right] && (78)
\end{aligned}$$

where $\tilde{P}_{\lambda\sigma}^{+}$ the same as that defined in eq. (17) of Maurice and Head-Gordon [2]. Then σ_i^a can be evaluated as

$$\sigma_i^a = \sigma_{\bar{i}}^{\bar{a}} = \frac{1}{2}\left\{\sum_b c_i^b \left[F_{ab} + F_{\bar{a}\bar{b}}\right] - \sum_j c_j^a \left[F_{ij} + F_{\bar{i}\bar{j}}\right] - \sum_t c_t^a F_{it} + \sum_{\bar{t}} c_{\bar{i}}^{\bar{t}} F_{\bar{a}\bar{t}}\right\} + \tilde{F}_{ia}^{+}, \tag{79}$$

where

$$\begin{aligned}
\tilde{F}_{ia}^{+} &= \sum_{\mu\nu} C_{\mu i}^* \tilde{F}_{\mu\nu} C_{\nu a} && (80) \\
\tilde{F}_{\mu\nu}^{+} &= \sum_{\lambda\sigma} \left[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)\right] \tilde{P}_{\lambda\sigma}^{+} && (81)
\end{aligned}$$

Next consider the term σ_t^a:

$$\begin{aligned}
\sigma_t^a &= \sum_{jb} \left[\langle aj||tb\rangle - F_{tj}\delta_{ab}\right] c_j^b + \sum_{\bar{j}\bar{b}} \langle a\bar{j}||t\bar{b}\rangle c_{\bar{j}}^{\bar{b}} + \sum_{ub} \left[F_{ab}\delta_{tu} - F_{tu}\delta_{ab} + \langle a\bar{j}||t\bar{u}\rangle c_u^b\right] + \sum_{\bar{j}\bar{u}} \langle a\bar{j}||t\bar{u}\rangle c_{\bar{j}}^{\bar{u}} \\
&= \sum_{jb} c_j^b \left[-F_{tj}\delta_{ab} + (at|jb) - (ab|jt)\right] + \sum_{\bar{j}\bar{b}} c_{\bar{j}}^{\bar{b}} (at|jb) \\
&+ \sum_{ub} c_u^b \left[(at|ub) - (ab|ut) + F_{ab}\delta_{tu} - F_{tu}\delta_{ab}\right] + \sum_{\bar{j}\bar{u}} c_{\bar{j}}^{\bar{u}} (at|ju) \\
&= \tilde{F}_{ta} - \sum_j c_j^a F_{tj} + \sum_b c_t^b F_{ab} - \sum_u c_u^a F_{tu} \qquad (82)
\end{aligned}$$

It is computationally more efficient to evaluate σ_t^a at the same time as σ_i^a. The value of σ_t^a computed in this manner must of course be corrected for the difference in the formulas for σ_t^a and σ_i^a, but this correction scales as only $\mathcal{O}(N^2)$ (see Maurice and Head-Gordon [2]). The two-electron part is thus computed by

$$\tilde{F}_{ta} = \tilde{F}_{ta}^{+} - \frac{1}{2}\sum_{ub} c_u^b (ab|ut) + \frac{1}{2}\sum_{\bar{j}\bar{u}} c_{\bar{j}}^{\bar{u}} (at|ju) \qquad (83)$$

Similar considerations apply to $\sigma_{\bar{i}}^{\bar{t}}$, which is

$$\begin{aligned}
\sigma_{\bar{i}}^{\bar{t}} &= \sum_{jb} c_j^b \langle \bar{t}j||\bar{i}b\rangle + \sum_{\bar{j}\bar{b}} c_{\bar{j}}^{\bar{b}} \left[F_{\bar{t}\bar{b}}\delta_{\bar{i}\bar{j}} + \langle \bar{t}\bar{j}||\bar{i}\bar{b}\rangle\right] + \sum_{ub} c_u^b \langle \bar{t}u||\bar{i}b\rangle + \sum_{\bar{j}\bar{u}} c_{\bar{j}}^{\bar{u}} \left[F_{\bar{t}\bar{u}}\delta_{\bar{i}\bar{j}} - F_{\bar{i}\bar{j}}\delta_{\bar{t}\bar{u}} + \langle \bar{t}\bar{j}||\bar{i}\bar{u}\rangle\right] \\
&= \sum_{\bar{b}} c_{\bar{i}}^{\bar{b}} F_{\bar{t}\bar{b}} + \sum_{\bar{u}} c_{\bar{i}}^{\bar{u}} F_{\bar{t}\bar{u}} - \sum_{\bar{j}} c_{\bar{j}}^{\bar{t}} F_{\bar{i}\bar{j}} + \sum_{jb} c_j^b (ti|jb) + \sum_{\bar{j}\bar{b}} c_{\bar{j}}^{\bar{b}} \left[(ti|jb) - (tb|ji)\right] \\
&+ \sum_{ub} c_u^b (ti|ub) + \sum_{\bar{j}\bar{u}} c_{\bar{j}}^{\bar{u}} \left[(ti|ju) - (tu|ji)\right] \\
&= \sum_{\bar{b}} c_{\bar{i}}^{\bar{b}} F_{\bar{t}\bar{b}} + \sum_{\bar{u}} c_{\bar{i}}^{\bar{u}} F_{\bar{t}\bar{u}} - \sum_{\bar{j}} c_{\bar{j}}^{\bar{t}} F_{\bar{i}\bar{j}} + \tilde{F}_{\bar{i}\bar{t}} \qquad (84)
\end{aligned}$$

where $\tilde{F}_{\bar{i}\bar{t}}$ is actually evaluated according to

$$\tilde{F}_{\bar{i}\bar{t}} = \tilde{F}_{\bar{i}\bar{t}}^{+} + \frac{1}{2}\sum_{ub} c_u^b (ti|ub) - \frac{1}{2}\sum_{\bar{j}\bar{u}} c_{\bar{j}}^{\bar{u}} (tu|ji) \qquad (85)$$

3 Extensions of CIS to Include Certain Double Substitutions (XCIS)

The ROCIS method appears to be superior to UCIS for open-shell molecules [2]. Nevertheless, ROCIS is not as reliable for open-shell cases as RCIS is for closed-shell cases. As explained by

Maurice and Head-Gordon [7], a careful analysis of the failures of ROCIS indicated that certain double substitutions which are neglected in UCIS and ROCIS can of crucial importance in open-shell systems. The spin-adapted configurations are of the form:

$$|\tilde{\Phi}_i^a(1)\rangle = \frac{1}{\sqrt{6}}\left(|\Phi_{\bar{i}}^{\bar{a}}\rangle - |\Phi_i^a\rangle\right) + \frac{2}{\sqrt{6}}|\Phi_{t\bar{i}}^{a\bar{t}}\rangle. \tag{86}$$

Although the third determinant is a double substitution as far as spin-orbitals are concerned, it is only a single substitution when spatial orbital occupations are considered. Hence, it is very reasonable to assume that this CSF may be of comparable importance to the singles included in ROCIS. The extended CIS method (XCIS) is a spin-adapted CI method including these "extended single" substitutions.

It is helpful to first verify that eq. (86) is indeed an eigenfunction of $\hat{S}^2$. Using previous results for ROCIS, it is trivial to see that

$$\hat{S}^2\left(|\Phi_{\bar{i}}^{\bar{a}}\rangle - |\Phi_i^a\rangle\right) = \left[\frac{N_s}{2}\left(\frac{N_s}{2}+1\right)+2\right]\left(|\Phi_{\bar{i}}^{\bar{a}}\rangle - |\Phi_i^a\rangle\right) - 2\sum_t^{socc}|\Phi_{t\bar{i}}^{a\bar{t}}\rangle. \tag{87}$$

It remains to be seen what is the effect of $\hat{S}^2$ acting on $|\Phi_{t\bar{i}}^{a\bar{t}}\rangle$.

$$\begin{aligned}
\hat{S}^2|\Phi_{t\bar{i}}^{a\bar{t}}\rangle &= \hat{S}^2 a_a^\dagger a_{\bar{t}}^\dagger a_{\bar{i}} a_t \prod_i^{docc} a_i^\dagger a_{\bar{i}}^\dagger \prod_u^{socc} a_u^\dagger|\rangle \\
&= \hat{S}_-\hat{S}_+|\Phi_{t\bar{i}}^{a\bar{t}}\rangle + \left[\frac{N_s}{2}\left(\frac{N_s}{2}+1\right)\right]|\Phi_{t\bar{i}}^{a\bar{t}}\rangle.
\end{aligned} \tag{88}$$

The factor $\hat{S}_+$ acts on unpaired β spins; thus,

$$\begin{aligned}
\hat{S}_+ a_a^\dagger a_{\bar{t}}^\dagger a_{\bar{i}} a_t|\Phi_0\rangle &= \hat{s}_+(\bar{t}) a_a^\dagger a_{\bar{t}}^\dagger a_{\bar{i}} a_t|\Phi_0\rangle \\
&= -a_t^\dagger a_a^\dagger a_{\bar{i}} a_t|\Phi_0\rangle \\
&= -a_a^\dagger a_{\bar{i}}|\Phi_0\rangle.
\end{aligned} \tag{89}$$

The result of $-\hat{S}_- a_a^\dagger a_{\bar{i}}|\Phi_0\rangle$ has already been worked out in eq. (60). Thus overall,

$$\hat{S}^2|\Phi_{t\bar{i}}^{a\bar{t}}\rangle = \left[\frac{N_s}{2}\left(\frac{N_s}{2}+1\right)\right]|\Phi_{t\bar{i}}^{a\bar{t}}\rangle + |\Phi_i^a\rangle - |\Phi_{\bar{i}}^{\bar{a}}\rangle + \sum_u |\Phi_{u\bar{i}}^{a\bar{u}}\rangle, \tag{90}$$

and it is easy to see that

$$\hat{S}^2\left(|\Phi_{\bar{i}}^{\bar{a}}\rangle - |\Phi_i^a\rangle + 2|\Phi_{t\bar{i}}^{a\bar{t}}\rangle\right) = \left[\frac{N_s}{2}\left(\frac{N_s}{2}+1\right)\right]\left(|\Phi_{\bar{i}}^{\bar{a}}\rangle - |\Phi_i^a\rangle + 2|\Phi_{t\bar{i}}^{a\bar{t}}\rangle\right). \tag{91}$$

Hence $|\tilde{\Phi}_i^a(1)\rangle$ (eq. 86) is a CSF.

References

[1] J. B. Foresman, M. Head-Gordon, J. A. Pople, and M. J. Frisch, J. Phys. Chem. **96**, 135 (1992).

[2] D. Maurice and M. Head-Gordon, Int. J. Quantum Chem. Symp. **29**, 361 (1995).

[3] E. R. Davidson, J. Comp. Phys. **17**, 87 (1975).

[4] B. Liu, The simultaneous expansion method for the iterative solution of several of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices, Technical Report LBL-8158, Lawrence Berkeley Laboratory, University of California, Berkeley, 1978.

[5] N. C. Handy, Chem. Phys. Lett. **74**, 280 (1980).

[6] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. McGraw-Hill, New York, 1989.

[7] D. Maurice and M. Head-Gordon, J. Phys. Chem. **100**, 6131 (1996).