Analytic Gradients of Configuration Interaction Energies

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

February 2008 Last Updated: August 2009

1 General Expressions for CI Gradients

A configuration interaction (CI) wavefunction $|\Psi\rangle$ is expressed as a linear combination of N-electron basis functions (typically Slater determinants or spin-adapted configuration state functions):

$$|\Psi\rangle = \sum_I c_I |\Phi_I\rangle. \tag{1}$$

The energy is an upper bound to the true nonrelativistic energy, according to the variational principle, and is given by

$$\begin{aligned} E &= \langle \Psi | \hat{H} | \Psi \rangle \\ &= \sum_{IJ} c_I^* c_J \langle \Phi_I | \hat{H} | \Phi_J \rangle \\ &= \sum_{IJ} c_I^* c_J H_{IJ}. \end{aligned} \tag{2}$$

The configuration interaction method is discussed in great detail in several reviews[1,2,3] and in lecture notes.[4,5] In these notes we will consider analytic gradients of the CI energy.

The first derivative of the energy with respect to some parameter x may be written

$$\frac{\partial E}{\partial x} = \sum_{IJ} \left(\frac{\partial c_I^*}{\partial x} c_J H_{IJ} + c_I^* \frac{\partial c_J}{\partial x} H_{IJ} + c_I^* c_J \frac{\partial H_{IJ}}{\partial x} \right). \tag{3}$$

Using the fact that the converged CI vectors $\mathbf{c}$ satisfy the Schrödinger equation in the selected N-electron space, $\mathbf{Hc} = E\mathbf{c}$, or

$$\sum_J H_{IJ} c_J = E c_I, \tag{4}$$

and using the properties of Hermitian operators, we obtain

$$\frac{\partial E}{\partial x} = E \sum_I \left(\frac{\partial c_I^*}{\partial x} c_I + c_I^* \frac{\partial c_I}{\partial x} \right) + \sum_{IJ} c_I^* c_J \frac{\partial H_{IJ}}{\partial x}. \tag{5}$$

This expression can be simplified further because the CI vector is normalized,

$$\sum_I c_I^* c_I = 1. \tag{6}$$

Differentiating the normalization condition with respect to x yields the useful result

$$\sum_I \left(\frac{\partial c_I^*}{\partial x} c_I + c_I^* \frac{\partial c_I}{\partial x} \right) = 0. \tag{7}$$

Substituting this into the most recent expression for $\partial E/\partial x$ yields simply

$$\frac{\partial E}{\partial x} = \sum_{IJ} c_I^* c_J \frac{\partial H_{IJ}}{\partial x}. \tag{8}$$

Hence, the fact that the CI coefficients c_I are determined variationally means that there is no need to consider their derivatives when evaluating the gradient of the CI energy, and the CI wavefunction obeys the Hellmann-Feynman theorem with respect to the expansion coefficients c_I. Furthermore, these considerations hold regardless of the particular choice of molecular orbitals.

Now it remains to consider the form of the terms $\partial H_{IJ}/\partial x$. Using the one- and two-particle coupling coefficients from configuration interaction theory,

$$\begin{aligned} H_{IJ} &= \langle \Phi_I | \hat{H} | \Phi_J \rangle \\ &= \sum_{pq} \gamma_{pq}^{IJ} h_{pq} + \sum_{pqrs} \Gamma_{pqrs}^{IJ} (pq|rs). \end{aligned} \tag{9}$$

Note: some authors leave a factor of $\frac{1}{2}$ *before the two-particle coupling coefficient. We have followed a majority of publications by absorbing the* $\frac{1}{2}$ *into* Γ*. The* DETCI *program leaves the* $\frac{1}{2}$ *factor pulled out. Thus replace every* Γ *in these notes with* $\frac{1}{2}\Gamma$ *to compare to* DETCI*. Our notation here is consistent with Yamaguchi et al.*[6] Recall that γ_{pq}^{IJ} and Γ_{pqrs}^{IJ} are given by[7,8,9,3]

$$\gamma_{pq}^{IJ} = \langle \Phi_I | \hat{E}_{pq} | \Phi_J \rangle, \tag{10}$$

$$\Gamma_{pqrs}^{IJ} = \frac{1}{2} \langle \Phi_I | \hat{E}_{pq} \hat{E}_{rs} - \delta_{qr} \hat{E}_{ps} | \Phi_J \rangle, \tag{11}$$

and have the properties

$$\gamma^{IJ}_{pq} = (\gamma^{JI}_{qp})^* \tag{12}$$

$$\Gamma^{IJ}_{pqrs} = \Gamma^{IJ}_{rspq} = (\Gamma^{JI}_{srqp})^* = (\Gamma^{JI}_{qpsr})^*. \tag{13}$$

Armed with this more detailed notation, we can proceed with the task of writing down an explicit expression for the CI gradients. Equation (8) may be expanded as

$$\frac{\partial E}{\partial x} \;=\; \sum_{IJ} c_I^* c_J \left(\sum_{pq} \gamma^{IJ}_{pq} \frac{\partial h_{pq}}{\partial x} + \sum_{pqrs} \Gamma^{IJ}_{pqrs} \frac{\partial (pq|rs)}{\partial x} \right). \tag{14}$$

Derivatives of the coupling coefficients are generally zero, so we have omitted such terms. Given this, we can simplify the expression by summing over I and J to yield the standard CI one- and two-particle density matrices,

$$\gamma_{pq} \;=\; \sum_{IJ} c_I^* c_J \gamma^{IJ}_{pq} \tag{15}$$

$$\Gamma_{pqrs} \;=\; \sum_{IJ} c_I^* c_J \Gamma^{IJ}_{pqrs}. \tag{16}$$

Hence the energy gradient can be written more compactly as

$$\frac{\partial E}{\partial x} \;=\; \sum_{pq} \gamma_{pq} \frac{\partial h_{pq}}{\partial x} + \sum_{pqrs} \Gamma_{pqrs} \frac{\partial (pq|rs)}{\partial x}. \tag{17}$$

1.1 Density Contributions from Inactive Core Orbitals

It is possible to reduce the storage requirements of the one- and two-particle density matrices in the case of inactive core or inactive virtual orbitals. By inactive core, we mean that the orbital is required to remain doubly occupied in all N-electron functions which span the CI space. This is called the “frozen core” approximation in the context of CI wavefunctions, but our discussion here will also pertain to multiconfigurational self-consistent-field (MCSCF) wavefunctions, in which the orbitals are constrained to remain doubly occupied but are not “frozen” since they are optimized during the course of the MCSCF procedure. Likewise, inactive virtuals are not allowed to be occupied in the CI wavefunction, but they too are optimized in an MCSCF procedure.

Any element of either density matrix with an index corresponding to an inactive virtual orbital must be zero. Hence, the indices of the density matrices can be restricted to run over only populated (i.e., excluding inactive virtual) orbitals. Furthermore, the density matrices become sparse when at least one index belongs to the inactive core. Let us use $t, u, v, \ldots$ to denote active orbitals (i.e., all orbitals but the inactive core and inactive virtuals), and $i', j', k', \ldots$ to denote inactive core orbitals. Then the one-particle density matrix obeys (assuming real orbitals)

$$\begin{aligned} \gamma_{i'j'} &= 2\delta_{i'j'}, \qquad (18)\\ \gamma_{i't} &= 0, \qquad (19) \end{aligned}$$

and the two-particle density matrix elements are zero if one or more indices belong to the inactive core space, except for the following elements:

$$\begin{aligned} \Gamma_{i'j'k'l'} &= 2\delta_{i'j'}\delta_{k'l'} - \delta_{j'k'}\delta_{i'l'} \qquad (20)\\ \Gamma_{tui'i'} &= \Gamma_{i'i'tu} = \gamma_{tu} \qquad (21)\\ \Gamma_{ti'i'u} &= \Gamma_{i'tui'} = -\frac{1}{2}\gamma_{tu}. \qquad (22) \end{aligned}$$

2 Gradients with respect to Nuclear Coordinates

To continue evaluating this expression, we need to know how the molecular orbitals change for a given perturbation δx of coordinate x. Here we will focus on Cartesian coordinates of the atoms, as required in geometry optimization. We expect a change not only in the coefficients that describe the molecular orbital in terms of atomic orbitals, but also in the underlying atomic orbitals themselves. To first order,[6]

$$\phi_p \rightarrow \phi_p + \delta x \left(\sum_q \tilde{U}^x_{qp} \phi_q + \phi^{\overline{x}}_p \right) + \mathcal{O}(\delta x^2), \qquad (23)$$

where $\tilde{U}^x_{qp}$ are the standard coupled-perturbed coefficients and where $\phi^{\overline{x}}_p$ is the molecular orbital evaluated with differentiated atomic orbitals

$$\phi^{\overline{x}}_p = \sum_\mu C_{\mu p} \left(\frac{\partial \chi_\mu}{\partial x} \right). \qquad (24)$$

Given this, it is straightforward to show that the derivatives of the one- and two-electron integrals are[6]

$$\frac{\partial h_{pq}}{\partial x} = \sum_r h_{rq} \tilde{U}^x_{rp} + \sum_r h_{pr} \tilde{U}^x_{rq} + \sum_{\mu\nu} C_{\mu p} C_{\nu q} \frac{\partial h_{\mu\nu}}{\partial x} \qquad (25)$$

$$\begin{aligned}\frac{\partial(pq|rs)}{\partial x} &= \sum_t \left[\tilde{U}^x_{tp}(tq|rs) + \tilde{U}^x_{tq}(pt|rs) + \tilde{U}^x_{tr}(pq|ts) + \tilde{U}^x_{ts}(pq|rt)\right] \\ &+ \sum_{\mu\nu\rho\sigma} C_{\mu p}C_{\nu q}C_{\rho r}C_{\sigma s}\frac{\partial(\mu\nu|\rho\sigma)}{\partial x}. \qquad (26)\end{aligned}$$

It is convenient to distinguish the parts of these expressions due to mixing of molecular orbitals from those due to the changes in atomic orbitals. Following Kobayashi *et al.*,[10] we will denote the purely AO parts (assuming real orbitals) as

$$h^{\overline{x}}_{pq} = \sum_{\mu\nu} C_{\mu p}C_{\nu q}\frac{\partial h_{\mu\nu}}{\partial x} \qquad (27)$$

$$(pq|rs)^{\overline{x}} = \sum_{\mu\nu\rho\sigma} C_{\mu p}C_{\nu q}C_{\rho r}C_{\sigma s}\frac{\partial(\mu\nu|\rho\sigma)}{\partial x}. \qquad (28)$$

These quantities are called the *skeleton* (or *core*) derivative integrals and denoted h^x_{pq} and $(pq|rs)^x$, respectively, by Yamaguchi *et al.*[6] We will reserve these symbols (without overbars on x) as a compact notation for the full derivatives $\partial h_{pq}/\partial x$ and $\partial(pq|rs)/\partial x$, respectively.

Hence the analytic energy gradient for a CI wavefunction separates into two types of terms. The skeleton terms reflect the changes in the atomic orbitals, and they have the same form as the full energy expression, except that derivative integrals replace the regular integrals. The remaining terms reflect changes in the molecular orbitals and involve the orbital response coefficients $\tilde{U}^x_{pq}$, which could be obtained by the coupled-perturbed Hartree-Fock equations if the molecular orbitals were Hartree-Fock orbitals. We may write the gradient as

$$\begin{aligned}\frac{\partial E}{\partial x} &= \sum_{pq}\gamma_{pq}h^{\overline{x}}_{pq} + \sum_{pqrs}\Gamma_{pqrs}(pq|rs)^{\overline{x}} \qquad (29) \\ &+ \sum_{pq}\gamma_{pq}\sum_r h_{rq}\tilde{U}^x_{rp} + \sum_{pq}\gamma_{pq}\sum_r h_{pr}\tilde{U}^x_{rq} \\ &+ \sum_{pqrs}\Gamma_{pqrs}\sum_t\left[\tilde{U}^x_{tp}(tq|rs) + \tilde{U}^x_{tq}(pt|rs) + \tilde{U}^x_{tr}(pq|ts) + \tilde{U}^x_{ts}(pq|rt)\right].\end{aligned}$$

From (13), it is clear that $\Gamma_{pqrs} = \Gamma_{rspq}$. Furthermore, for real wavefunctions, $\gamma_{pq} = \gamma_{qp}$, and $\Gamma_{pqrs} = \Gamma_{rspq} = \Gamma_{srqp} = \Gamma_{qpsr}$. This allows us to simplify the gradient expression to

$$\begin{aligned}\frac{\partial E}{\partial x} &= \sum_{pq}\gamma_{pq}h^{\overline{x}}_{pq} + \sum_{pqrs}\Gamma_{pqrs}(pq|rs)^{\overline{x}} \qquad (30) \\ &+ 2\sum_{pq}\gamma_{pq}\sum_r h_{rq}\tilde{U}^x_{rp} + 2\sum_{pqrs}\left(\Gamma_{pqrs} + \Gamma_{qprs}\right)\sum_t (tq|rs)\tilde{U}^x_{tp}.\end{aligned}$$

We may write the energy gradient even more compactly as[11]

$$\frac{\partial E}{\partial x} = \sum_{pq} \gamma_{pq} h_{pq}^{\overline{x}} + \sum_{pqrs} \Gamma_{pqrs} (pq|rs)^{\overline{x}} + 2 \sum_{pq} \tilde{U}_{pq}^{x} X_{pq}, \tag{31}$$

by defining the Lagrangian X as

$$X_{pq} = \sum_{r} h_{pr} \gamma_{qr} + \sum_{rst} (pr|st) \left(\Gamma_{qrst} + \Gamma_{rqst} \right). \tag{32}$$

Note that many other authors typically write the Lagrangian as

$$X_{pq} = \sum_{r} h_{pr} \gamma_{qr} + 2 \sum_{rst} (pr|st) \tilde{\Gamma}_{qrst}. \tag{33}$$

This expression is equivalent to the present one (32) if $\tilde{\Gamma}$ is symmetrized to exhibit the full eight-fold permutational symmetry of the integrals (i.e., so that $\tilde{\Gamma}_{qrst} = \tilde{\Gamma}_{rqst}$). This can be accomplished by literally setting $\tilde{\Gamma}_{qrst} = \frac{1}{2} \left(\Gamma_{qrst} + \Gamma_{rqst} \right)$. For the purpose of contracting Γ against the two-electron integrals, no error is introduced by symmetrizing Γ; individual terms will become different, but the overall sum is unchanged because of the symmetry of the two-electron integrals. However, other uses of Γ may not allow this symmetrization.

2.1 Independent Pairs and Gradients of MCSCF Wave Functions

MCSCF wave functions represent a special case for CI gradients, because the orbitals are, by definition, those that minimize the CI energy. Here we assume that the orbitals are fully optimized (i.e., none are fixed in their SCF form — the orbitals may be "restricted" in their occupations, but they are not "frozen" and prevented from being optimized).

If the orbitals are optimized, then, to first order, no rotation between any pair of orbitals will change the CI energy. In a CASSCF or RASSCF, the energy is already invariant to rotations of orbitals within the same RAS/CAS space, and the orbital optimization ensures that the energy is also stationary with respect to rotations between orbitals belonging to different orbital subspaces. If the energy is invariant (to first order) to rotations between orbitals p and q, then the Lagrangian is symmetric with respect to this pair: $X_{pq} = X_{qp}$. (In fact, one can also view the MCSCF procedure as *the process of making the Lagrangian symmetric.*) The symmetric nature of the MCSCF Lagrangian has a very important consequence when coupled with one other fact.

The orthonormality constraint on the molecular orbitals,

$$\langle \phi_p | \phi_q \rangle = \delta_{pq}, \tag{34}$$

leads to the following first-order condition on the coupled-perturbed Hartree-Fock coefficients:

$$\tilde{U}^x_{pq} + \tilde{U}^x_{qp} + S^{\overline{x}}_{pq} = 0. \tag{35}$$

With these two facts ready, we can tremendously simplify the Lagrangian contribution to the MCSCF gradient:

$$\begin{aligned} 2\sum_{pq} X_{pq}\tilde{U}^x_{pq} &= \sum_{pq} X_{pq}\left(\tilde{U}^x_{pq} - \tilde{U}^x_{qp}\right) + \sum_{pq} X_{pq}\left(\tilde{U}^x_{pq} + \tilde{U}^x_{qp}\right) && (36)\\ &= \sum_{pq} \tilde{U}^x_{pq}\left(X_{pq} - X_{qp}\right) - \sum_{pq} X_{pq}S^{\overline{x}}_{pq} && (37)\\ &= -\sum_{pq} X_{pq}S^{\overline{x}}_{pq}. && (38)\end{aligned}$$

This is a remarkable result — there is *no need* to solve the coupled-perturbed Hartree-Fock equations (explicitly or implicitly) to obtain the $\tilde{U}^x_{pq}$ factors, because they disappear! Note once again that this happens because the Lagrangian is symmetric for MCSCF. The Lagrangian will also be symmetric for full CI (because the full CI energy is invariant to any orbital rotations), but only if no orbitals are frozen.

From the above equations, we can see that we need to evaluate the effect of orbital rotation, through the Lagrangian and the coupled-perturbed Hartree-Fock coefficients, only for pairs of orbitals p and q for which $X_{pq} \neq X_{qp}$. Such pairs of orbitals are called the *independent pairs.* In practice, one would use the Z-vector trick to avoid explicit computation of the CPHF coefficients.

2.2 Efficient Computation of Energy Gradients

2.2.1 Backtransformation of the Density Matrices

As any CI program must compute (either explicitly or implicitly) the one- and two-electron coupling coefficients (γ^{IJ}_{pq} and Γ^{IJ}_{pqrs}), it is straightforward to obtain the one- and two-particle density matrices γ_{pq} and Γ_{pqrs} via (15) and (16). However, the derivative integrals $h^{\overline{x}}_{pq}$ and $(pq|rs)^{\overline{x}}$ are originally computed in the atomic orbital (AO) basis. So, one must either transform the derivative integrals into the MO basis, or *backtransform* the density matrices into the AO basis. The latter alternative is much more efficient, since the number of derivative integrals is greater than the number of density matrix elements by a factor of $3N_A$, where N_A is the number of atoms. Note that the backtransformation of the density matrices is not really a true backwards transformation, in the sense that, if applied to the MO integrals, it would not yield the AO integrals back again. It is clear from (31) that one needs to use transposed MO coefficient matrices, and *not* the inverses of these matrices.

2.2.2 Z-Vector Treatment of the Orbital Response Contributions

The term $2\sum_{pq}\tilde{U}^{x}_{pq}X_{pq}$ is most efficiently evaluated using the Z-vector approach of Handy and Schaefer[12] to avoid having to explicitly solve for the coupled-perturbed coefficients $\tilde{U}^{x}_{pq}$ for every perturbation x. Instead, a set of perturbation-independent equations are solved, and then very inexpensive dot products are taken to compute the contributions of the individual perturbations.

First, let us re-write the expression $2\sum_{pq}\tilde{U}^{x}_{pq}X_{pq}$ so that we can work only in terms of the $\tilde{U}^{x}_{pq}$ that are really necessary. We can break down this term as:

$$2\sum_{pq}\tilde{U}^{x}_{pq}X_{pq} = 2\sum_{p>q}\tilde{U}^{x}_{pq}X_{pq} + 2\sum_{p}\tilde{U}^{x}_{pq}X_{pp} + 2\sum_{p<q}\tilde{U}^{x}_{pq}. \tag{39}$$

Using the orthonormality constraint of the orbitals (35), $\tilde{U}^{x}_{pp} = -\frac{1}{2}S^{\overline{x}}_{pp}$, and $\tilde{U}^{x}_{qp} = -\tilde{U}^{x}_{pq} - S^{\overline{x}}_{pq}$. Thus we arrive at

$$2\sum_{pq}\tilde{U}^{x}_{pq}X_{pq} = 2\sum_{p>q}\tilde{U}^{x}_{pq}(X_{pq} - X_{qp}) - \sum_{p}S^{\overline{x}}_{pp}X_{pp} - 2\sum_{p>q}S^{\overline{x}}_{pq}X_{qp}. \tag{40}$$

The last two terms may be handled in various ways, but since we are contracting elements of the Lagrangian with a symmetric matrix ($\mathbf{S}^{\overline{x}}$), it may be advantageous to rearrange these terms so that the contraction involves a symmetric matrix $\tilde{\mathbf{X}}$ obtained from the relevant components of $\mathbf{X}$. In particular, the contribution of the last two terms can be written as

$$-\sum_{pq}S^{\overline{x}}_{pq}\tilde{X}_{pq}, \text{ where } \tilde{X}_{pq} = \begin{cases} X_{pq} & p \leq q \\ X_{qp} & p > q \end{cases} \tag{41}$$

This allows subsequent code to take advantage of the symmetric nature of the derivative overlap matrices and employ only the upper triangle or only the lower triangle, if desired.

Our rearrangement of the Lagrangian contribution may thus be rewritten as

$$2\sum_{pq}\tilde{U}^{x}_{pq}X_{pq} = 2\sum_{p>q}^{CI-IP}\tilde{U}^{x}_{pq}\Delta X_{pq} - \sum_{pq}S^{\overline{x}}_{pq}\tilde{X}_{pq}, \tag{42}$$

where the sum in the first term is only over the CI "independent pairs," i.e., those pairs of orbital rotations for which the CI energy changes. If an orbital rotation between orbitals p and q does not change the CI energy, then the Lagrangian $\mathbf{X}$ will be symmetric with respect to that rotation (i.e., $X_{pq} = X_{qp}$), making $\Delta X_{pq} = 0$, zeroing out that contribution. Such "redundant" pairs are not necessary. Only those pairs of orbitals for which a rotation changes the CI energy are required. To take the simplest example of a CISD out of a closed-shell RHF reference, with no orbitals frozen, the CI independent pairs are only those mixing one occupied and one unoccupied orbital. The CISD energy is invariant to mixings among occupied orbitals or among virtual orbitals, making such rotations redundant. In a RAS CI, only rotations which mix an orbital in one RAS space with an orbital in another RAS space will change the CI energy. For any kind of CI, rotations mixing frozen and unfrozen orbitals will also change the CI energy, and these need to be included in the list of CI independent pairs.

Once the list of independent pairs is created, it remains to evaluate $2\sum_{p>q}^{CI-IP} \tilde{U}^x_{pq} \Delta X_{pq}$. In principle, one could explicitly solve the CPHF equations to obtain $\tilde{U}^x_{pq}$ and then contract against differences in Lagrangian elements. However, it is possible for CI first derivatives to bypass the explicit solution of $\tilde{U}^x_{pq}$. The CPHF equations have the form

$$(\mathbf{\Xi} - \mathbf{A})\tilde{\mathbf{U}}^x = \mathbf{B}^x_0, \tag{43}$$

where $\tilde{\mathbf{U}}^x$ is the set of CPHF coefficients $\tilde{U}^x_{ij}$ for each perturbation x. We will treat $\tilde{\mathbf{U}}^x$ and $\mathbf{B}^x_0$ as vectors (with compound index pq), and matrices $\mathbf{\Xi}$ and $\mathbf{A}$ as a matrix with compound indices pq and rs. $\mathbf{\Xi}$ is a diagonal matrix of SCF orbital eigenvalue differences,

$$\Xi_{pq,rs} = \delta_{pr}\delta_{qs}(\epsilon_q - \epsilon_p). \tag{44}$$

For a closed-shell SCF reference function, the elements of matrix $\mathbf{A}$ are

$$A_{pq,rs} = 4(pq|rs) - (pr|qs) - (ps|qr). \tag{45}$$

The (perturbation-dependent) elements of $\mathbf{B}^x_0$ are

$$B^x_{0,pq} = f^{\overline{x}}_{pq} - S^{\overline{x}}_{pq}\epsilon_q - \sum_{ij}^{d.o.} S^{\overline{x}}_{ij}\left[2(pq|ij) - (pi|qj)\right]. \tag{46}$$

The matrix elements of the skeleton (core) Fock matrix are

$$f^{\overline{x}}_{pq} = h^{\overline{x}}_{pq} + \sum_{k}^{d.o.} \left[2(pq|kk)^{\overline{x}} - (pk|qk)^{\overline{x}}\right]. \tag{47}$$

It will be convenient below to write the CPHF equations even more concisely, as

$$\tilde{\mathbf{A}}\tilde{\mathbf{U}}^x = \mathbf{B}^x_0, \tag{48}$$

where

$$\tilde{\mathbf{A}} = \boldsymbol{\Xi} - \mathbf{A}. \tag{49}$$

The elements of $\mathbf{U}^x$ are not well-defined for rotations which do not change the Hartree-Fock reference energy ("SCF redundant pairs"). Hence, the CPHF equations are solved first for the SCF independent pairs. Let us consider the case of a closed-shell, RHF reference, where the independent pairs will be all pairs (a, i), with a a virtual orbital and i an occupied orbital. We solve for the CPHF coefficients $\mathbf{U}^x_{ai}$ of the SCF independent pairs a, i using the CPHF equation 48, where the relevant parts of $\tilde{\mathbf{A}}$ and $\mathbf{B}^x_0$ are

$$\tilde{A}_{ai,bj} = \delta_{ab}\delta_{ij}(\epsilon_i - \epsilon_a) - \left[4(ai|bj) - (ab|ij) - (aj|ib)\right], \tag{50}$$

$$B^x_{0,ai} = f^{\overline{x}}_{ai} - S^{\overline{x}}_{ai}\epsilon_i - \sum_{jk}^{d.o.} S^{\overline{x}}_{jk}\left[2(ai|jk) - (aj|ik)\right]. \tag{51}$$

Subsequently, the elements for the SCF-redundant pairs may be determined using

$$\tilde{U}^x_{pq} = \frac{1}{(\epsilon_q - \epsilon_p)}\left(\sum_a^{vir}\sum_i^{d.o.} A_{pq,ai}\tilde{U}^x_{ai} + B^x_{0,pq}\right). \tag{52}$$

As pointed out by Yamaguchi *et al.*,[6] the equation above is defined using the diagonal nature of the Fock matrix, making it inapplicable in cases of orbital degeneracies. In practice, we will not need to use this equation except for pairs that are CI independent pairs, and by definition this means that the two orbitals p and q will not be degenerate (degenerate orbitals should always be placed in the same orbital subspace). Now that the relevant CPHF equations have been given, we will proceed to show how to avoid the explicit use of $\tilde{\mathbf{U}}^x$.

The gradient contribution from the non-symmetric part of the Lagrangian may be broken up into SCF-independent and SCF-redundant pairs:

$$\begin{aligned}
2\sum_{p>q}^{CI-IP} \tilde{U}^x_{pq}\Delta X_{pq} &= 2\sum_{p>q}^{\substack{CI-IP \\ SCF-IP}} \tilde{U}^x_{pq}\Delta X_{pq} + 2\sum_{p>q}^{\substack{CI-IP \\ SCF-redun}} \tilde{U}^x_{pq}\Delta X_{pq} && (53)\\
&= 2\sum_a^{vir}\sum_i^{d.o.} \tilde{U}^x_{ai}\Delta X_{ai} + 2\sum_{j>k}^{CI-IP} \tilde{U}^x_{jk}\Delta X_{jk} + 2\sum_{b>c}^{CI-IP} \tilde{U}^x_{bc}\Delta X_{bc}, && (54)
\end{aligned}$$

where we have assumed here that all SCF independent pairs (a, i) are also CI independent pairs. This is not necessarily true (e.g., second-order CI may have a RAS II space mixing occupieds and virtuals), but it is not problematic to include these pairs in the CI independent pair list because the contributions will be zero anyway because ΔX_{ai} will be zero for these terms. However, it is important to include all these pairs because of a combination of terms coming up momentarily; the code which makes a list of CI independent pairs needs to be aware of this.

Using equation (52), this can be expanded as:

$$\begin{aligned} 2\sum_{p>q}^{CI-IP} \tilde{U}^x_{pq}\Delta X_{pq} &= 2\sum_a^{vir}\sum_i^{d.o.} \tilde{U}^x_{ai}\Delta X_{ai} \qquad (55) \\ &+ 2\sum_{j>k}^{CI-IP} \frac{\Delta X_{jk}}{\epsilon_k - \epsilon_j}\left(\sum_a^{vir}\sum_i^{d.o.} A_{jk,ai}\tilde{U}^x_{ai} + B^x_{0,jk}\right) \\ &+ 2\sum_{b>c}^{CI-IP} \frac{\Delta X_{bc}}{\epsilon_c - \epsilon_b}\left(\sum_a^{vir}\sum_i^{d.o.} A_{bc,ai}\tilde{U}^x_{ai} + B^x_{0,bc}\right) \\ &= 2\sum_a^{vir}\sum_i^{d.o.} \tilde{U}^x_{ai}\Delta X'_{ai} \qquad (56) \\ &+ 2\sum_{j>k}^{CI-IP} Z_{jk}B^x_{0,jk} + 2\sum_{b>c}^{CI-IP} Z_{bc}B^x_{0,bc}, \end{aligned}$$

where

$$\begin{aligned} \Delta X'_{ai} &= \Delta X_{ai} + \sum_{j>k}^{CI-IP} \frac{\Delta X_{jk}}{\epsilon_k - \epsilon_j} A_{jk,ai} + \sum_{b>c}^{CI-IP} \frac{\Delta X_{bc}}{\epsilon_c - \epsilon_b} A_{bc,ai}, \qquad (57) \\ Z_{jk} &= \frac{\Delta X_{jk}}{\epsilon_k - \epsilon_j}, \qquad (58) \\ Z_{bc} &= \frac{\Delta X_{bc}}{\epsilon_c - \epsilon_b}. \qquad (59) \end{aligned}$$

We can rewrite the first term in equation (56) as a dot product between two vectors (where the vectors run over all the SCF independent pairs a, i):

$$2\Delta \mathbf{X}'^T \tilde{\mathbf{U}}^x. \qquad (60)$$

Now let us rewrite the CPHF equations as

$$\tilde{\mathbf{U}}^x = \tilde{\mathbf{A}}^{-1}\mathbf{B}^x_0. \qquad (61)$$

With the CPHF coefficients in this form, equation (60) may be expressed as

$$\begin{aligned} 2\Delta\tilde{\mathbf{X}}^T\tilde{\mathbf{U}}^x &= 2\Delta\tilde{\mathbf{X}}^T\tilde{\mathbf{A}}^{-1}\mathbf{B}^x_0 \qquad (62) \\ &= 2\mathbf{Z}^T\mathbf{B}^x_0, \end{aligned}$$

where

$$\mathbf{Z}^T = \Delta\tilde{\mathbf{X}}^T\tilde{\mathbf{A}}^{-1}, \tag{63}$$

or

$$\tilde{\mathbf{A}}^T\mathbf{Z} = \Delta\tilde{\mathbf{X}}. \tag{64}$$

Hence the explicit computation of CPHF coefficients $\tilde{U}^x_{ai}$ can be avoided by computing the perturbation-independent $\mathbf{Z}$ vector instead. In the above equation, $\Delta\tilde{\mathbf{X}}$ and $\mathbf{Z}$ have lengths equal to the number of SCF independent pairs a, i; $\mathbf{A}$ here is a square matrix.

Referring back to equation (56), the second and third terms are also, like our rearranged first term, of the form $\mathbf{Z}^T\mathbf{B}^x_0$. We can thus combine all the terms from equation (56) into the general form

$$2\sum_{p>q}^{CI-IP}\tilde{U}^x_{pq}\Delta X_{pq} = 2\sum_{p>q}^{CI-IP} Z_{pq}B^x_{0,pq}, \tag{65}$$

where

$$\begin{aligned}
\sum_{a}^{vir}\sum_{i}^{d.o.}\tilde{A}^T_{bj,ai}Z_{ai} &= \Delta\tilde{X}_{bj}, && (66)\\
Z_{jk} &= \frac{\Delta X_{jk}}{\epsilon_k-\epsilon_j}, && (67)\\
Z_{bc} &= \frac{\Delta X_{bc}}{\epsilon_c-\epsilon_b}. && (68)
\end{aligned}$$

One way to evaluate this term is to solve for the $\mathbf{Z}$ vector and then contract it with derivative $\mathbf{B}^x_0$ matrices as they are computed. Another approach is to avoid explicit computation of $\mathbf{B}^x_0$, and instead fold these contributions into the one- and two-particle density matrices and the Lagrangian before they are backtransformed for contraction against derivative one and two-electron integrals and overlap integrals. If one pursues the first approach, the relevant equations are already presented above. To pursue the second approach, let us expand equation (65) using the definition (46) for $\mathbf{B}^x_0$:

$$\begin{aligned}
2\sum_{p>q}^{CI-IP} Z_{pq}B^x_{0,pq} &= 2\sum_{p>q}^{CI-IP} Z_{pq}\Bigg(h^{\overline{x}}_{pq} + \sum_{k}^{d.o.}\left[2(pq|kk)^{\overline{x}} - (pk|qk)^{\overline{x}}\right] && (69)\\
&\quad -S^{\overline{x}}_{pq}\epsilon_q - \sum_{ij}^{d.o.} S^{\overline{x}}_{ij}\left[2(pq|ij) - (pi|qj)\right]\Bigg).
\end{aligned}$$

Collecting the terms that multiply $h_{pq}^{\overline{x}}$, $(pq|rs)^{\overline{x}}$, and $S_{pq}^{\overline{x}}$ yields the following modified one-particle density matrix, two-particle density matrix, and Lagrangian:

$$\begin{aligned}
\gamma'_{pq} &= \gamma_{pq} + 2Z_{pq}\delta_{p>q}\delta_{(pq)\in\mathrm{CI-IP}}, &(70)\\
\Gamma'_{pqkk} &= \Gamma_{pqkk} + 4Z_{pq}\delta_{p>q}\delta_{(pq)\in\mathrm{CI-IP}}, &(71)\\
\Gamma'_{pkqk} &= \Gamma_{pkqk} - 2Z_{pq}\delta_{p>q}\delta_{(pq)\in\mathrm{CI-IP}}, &(72)\\
\tilde{X}'_{pq} &= \tilde{X}_{pq} + 2Z_{pq}\epsilon_q\delta_{p>q}\delta_{(pq)\in\mathrm{CI-IP}} + \delta_{p\in d.o.}\delta_{q\in d.o.}\sum_{r>s}^{CI-IP} 2Z_{rs}\left[2(pq|rs) - (pr|qs)\right]. &(73)
\end{aligned}$$

With these modifications, the total energy gradient may be written in a form similar to equation (31) as

$$\frac{\partial E}{\partial x} = \sum_{pq}\gamma'_{pq}h_{pq}^{\overline{x}} + \sum_{pqrs}\Gamma'_{pqrs}(pq|rs)^{\overline{x}} - \sum_{pq}\tilde{X}'_{pq}S_{pq}^{\overline{x}}. \tag{74}$$

References

[1] I. Shavitt, The method of configuration interaction, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer, pages 189–275. Plenum Press, New York, 1977.

[2] W. Duch, J. Mol. Struct. THEOCHEM **80**, 27 (1991).

[3] C. D. Sherrill and H. F. Schaefer, The configuration interaction method: Advances in highly correlated approaches, in *Advances in Quantum Chemistry*, edited by P.-O. Löwdin, volume 34, pages 143–269. Academic Press, New York, 1999.

[4] P. E. M. Siegbahn, The configuration interaction method, in *Lecture Notes in Quantum Chemistry: European Summer School in Quantum Chemistry*, edited by B. O. Roos, volume 58 of *Lecture Notes in Chemistry*, pages 255–293. Springer-Verlag, New York, 1992.

[5] C. D. Sherrill, An introduction to configuration interaction theory, 1995, http://vergil.chemistry.gatech.edu/notes/ci/.

[6] Y. Yamaguchi, Y. Osamura, J. D. Goddard, and H. F. Schaefer, *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab initio Molecular Electronic Structure Theory.* Oxford, New York, 1994.

[7] J. Paldus, J. Chem. Phys. **61**, 5321 (1974).

[8] J. Paldus, Int. J. Quantum Chem. Symp. **9**, 165 (1975).

[9] I. Shavitt, Int. J. Quantum Chem. Symp. **12**, 5 (1978).

[10] R. Kobayashi, N. C. Handy, R. D. Amos, G. W. Trucks, M. J. Frisch, and J. A. Pople, J. Chem. Phys. **95**, 6723 (1991).

[11] J. E. Rice, R. D. Amos, N. C. Handy, T. J. Lee, and H. F. Schaefer, J. Chem. Phys. **85**, 963 (1986).

[12] N. C. Handy and H. F. Schaefer, J. Chem. Phys. **81**, 5031 (1984).