

6. Configuration Interaction

I. Configuration Interaction

We *define* the error in the Hartree-Fock model as being due to electron correlation, and we define the correlation energy as

$$E_{corr} = E_{exact} - E_{HF} \quad (6.1)$$

Because the Hartree-Fock energy is always above the exact energy, the correlation energy is always negative,

$$E_{corr} < 0. \quad (6.2)$$

Why do we call the error in the Hartree-Fock model the correlation energy? Let us define the *Hartree-Fock Hamiltonian*,

$$\hat{F} \equiv \sum_{i=1}^N \hat{f}(\mathbf{r}_i) \quad (6.3)$$

where the Fock operator, $\hat{f}(\mathbf{r}_i)$, is defined as before. Recall, however, that when a Hamiltonian takes the form of a sum of one-electron Hamiltonians that the energy eigenvalue is just the sum of the one-electron Hamiltonians' energies and that the wave function is just a product of the one-electron wave functions, so that

$$\hat{F}\psi_{i_1}(\mathbf{z}_1)\psi_{i_2}(\mathbf{z}_2)\dots\psi_{i_N}(\mathbf{z}_N) = \sum_{k=1}^N \varepsilon_{i_k} \psi_{i_1}(\mathbf{z}_1)\psi_{i_2}(\mathbf{z}_2)\dots\psi_{i_N}(\mathbf{z}_N) \quad (6.4)$$

where $\{\psi_{i_k}\}$ and $\{\varepsilon_{i_k}\}$ are any of the Hartree-Fock orbitals. Of course, the Hartree-product wave function in Eq. (6.4) neither respects the indistinguishability of electrons nor the antisymmetry principle, but a linear combination of these wave functions—the Slater determinant—does:

$$\hat{F}\left|\psi_{i_1}\psi_{i_2}\dots\psi_{i_N}\right| = \left(\sum_{k=1}^N \varepsilon_{i_k}\right)\left|\psi_{i_1}\psi_{i_2}\dots\psi_{i_N}\right|. \quad (6.5)$$

The eigenfunctions of \hat{F} that are appropriate for indistinguishable Fermions—like electrons—are the Slater determinants, and because \hat{F} is a Hamiltonian, which is a Hermitian operator, these eigenfunctions form a complete set. We conclude that we can, by considering all the possible ways of forming a Slater determinant of Hartree-Fock orbitals, obtain a complete set of functions in terms of which the exact wave function can be expanded.

We also note that there are no $\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$ terms in the Hartree-Fock Hamiltonian—there are no terms explicitly linking the distance between the electrons. In the Hartree-Fock Hamiltonian: electrons only interact with the “average position” of the other electrons, and there is no “explicit” and “direct” Coulomb repulsion between electrons. Similarly, if we ignore the requirement of antisymmetry, the Hartree product wave function suggests that electrons move entirely independent—the wave function of the total system is merely the product of the wave functions for each electron, and so the probability of observing electron 1 at \mathbf{r}_1 , electron 2 at \mathbf{r}_2 , etc. is just the product of the probability that electron 1 is at \mathbf{r}_1 , electron 2 is at \mathbf{r}_2 , etc., with all electrons moving entirely independently:

$$|\psi_1(\mathbf{z}_1)\psi_2(\mathbf{z}_2)\dots\psi_N(\mathbf{z}_N)|^2 = |\psi_1(\mathbf{z}_1)|^2 |\psi_2(\mathbf{z}_2)|^2 \dots |\psi_N(\mathbf{z}_N)|^2 \quad (6.6)$$

The probability electron 1 is at \mathbf{r}_1 is unaffected by whether or not electron 2 is at \mathbf{r}_2 . When we move from the Hartree-product description to the Slater determinant description, matters are complicated, but similar. In particular, the probability that there is an electron with spin α at \mathbf{x} and an electron with spin β at \mathbf{x}' can be computed as

$$\rho_2(\mathbf{x}, \alpha; \mathbf{x}', \beta) = \left\langle \psi_1 \dots \psi_N \left| \sum_{i=1}^{N-1} \sum_{j=i+1}^N \delta_{\sigma_i \alpha} \delta(\mathbf{r}_i - \mathbf{x}) \delta_{\sigma_j \beta} \delta(\mathbf{r}_j - \mathbf{x}') \right| \psi_1 \dots \psi_N \right\rangle \quad (6.7)$$

which is similar in form to the 2-electron electron-electron repulsion integral, and can be evaluated using the same methods to give

$$\begin{aligned} & \left\langle \psi_1 \psi_2 \dots \psi_N \left| \sum_{i=1}^N \sum_{j=1}^N \delta_{\sigma_i \alpha} \delta(\mathbf{r}_i - \mathbf{x}) \delta_{\sigma_j \beta} \delta(\mathbf{r}_j - \mathbf{x}') \right| \psi_1 \psi_2 \dots \psi_N \right\rangle \\ &= \sum_{i=1}^{N-1} \sum_{j=1}^N \left[\left\langle \psi_i(\mathbf{z}_1) \psi_j(\mathbf{z}_2) \left| \delta_{\sigma_i \alpha} \delta(\mathbf{r}_1 - \mathbf{x}) \delta_{\sigma_j \beta} \delta(\mathbf{r}_2 - \mathbf{x}') \right| \psi_i(\mathbf{z}_1) \psi_j(\mathbf{z}_2) \right\rangle \right. \\ & \quad \left. - \left\langle \psi_i(\mathbf{z}_1) \psi_j(\mathbf{z}_2) \left| \delta_{\sigma_i \alpha} \delta(\mathbf{r}_1 - \mathbf{x}) \delta_{\sigma_j \beta} \delta(\mathbf{r}_2 - \mathbf{x}') \right| \psi_j(\mathbf{z}_1) \psi_i(\mathbf{z}_2) \right\rangle \right] \\ &= \sum_{i=1}^N \sum_{j=1}^N |\psi_i(\mathbf{x})|^2 |\psi_j(\mathbf{x}')|^2 \delta_{\sigma_i \alpha} \delta_{\sigma_j \beta} \\ &= \left(\sum_{i=1}^N \delta_{\sigma_i \alpha} |\psi_i(\mathbf{x})|^2 \right) \left(\sum_{i=1}^N \delta_{\sigma_i \beta} |\psi_i(\mathbf{x}')|^2 \right) \end{aligned} \quad (6.8)$$

which is just the product of the probability of observing an α -spin electron at \mathbf{x} and that of observing a β -spin electron at \mathbf{x}' . When the electrons have the same spin things are slightly more complicated (the second term in the second line in Eq. (6.8) need not be zero), and we have, for example,

$$\begin{aligned}
& \left\langle \left| \psi_1 \psi_2 \dots \psi_N \right| \left| \sum_{i=1}^N \sum_{j=1}^N \delta_{\sigma_i \alpha} \delta(\mathbf{r}_i - \mathbf{x}) \delta_{\sigma_j \alpha} \delta(\mathbf{r}_j - \mathbf{x}') \right| \left| \psi_1 \psi_2 \dots \psi_N \right| \right\rangle \\
&= \sum_{i=1}^{N-1} \sum_{j=1}^N \left[\left\langle \left| \psi_i(\mathbf{z}_1) \psi_j(\mathbf{z}_2) \right| \delta_{\sigma_i \alpha} \delta(\mathbf{r}_1 - \mathbf{x}) \delta_{\sigma_j \alpha} \delta(\mathbf{r}_2 - \mathbf{x}') \left| \psi_i(\mathbf{z}_1) \psi_j(\mathbf{z}_2) \right\rangle \right. \\
&\quad \left. - \left\langle \left| \psi_i(\mathbf{z}_1) \psi_j(\mathbf{z}_2) \right| \delta_{\sigma_i \alpha} \delta(\mathbf{r}_1 - \mathbf{x}) \delta_{\sigma_j \alpha} \delta(\mathbf{r}_2 - \mathbf{x}') \left| \psi_j(\mathbf{z}_1) \psi_i(\mathbf{z}_2) \right\rangle \right] \\
&= \sum_{i=1}^N \sum_{j=1}^N \left| \psi_i(\mathbf{x}) \right|^2 \left| \psi_j(\mathbf{x}') \right|^2 \delta_{\sigma_i \alpha} \delta_{\sigma_j \alpha} - \sum_{i=1}^N \sum_{j=1}^N \psi_i^*(\mathbf{x}) \psi_i(\mathbf{x}') \psi_j(\mathbf{x}) \psi_j^*(\mathbf{x}') \delta_{\sigma_i \alpha} \delta_{\sigma_j \alpha} \\
&= \left(\sum_{i=1}^N \delta_{\sigma_i \alpha} \left| \psi_i(\mathbf{x}) \right|^2 \right) \left(\sum_{i=1}^N \delta_{\sigma_i \alpha} \left| \psi_i(\mathbf{x}') \right|^2 \right) - \sum_{i=1}^N \sum_{j=1}^N \psi_i^*(\mathbf{x}) \psi_i(\mathbf{x}') \psi_j(\mathbf{x}) \psi_j^*(\mathbf{x}') \delta_{\sigma_i \alpha} \delta_{\sigma_j \alpha}.
\end{aligned} \tag{6.9}$$

A small amount of correlation between same spin electrons, then is introduced by requiring that no two electrons with the same spin can be at the same place (Pauli exclusion principle), but we usually call this very special sort of “correlation” between the electrons “exchange”, and reserve the term “correlation” to mean the more sophisticated collective electron motions that represent corrections to both Eqs. (6.8) and (6.9).

How can we include correlation effects? Recall the variational principle:

$$E_0 \equiv \min_{\Psi} \langle \Psi | \hat{H} | \Psi \rangle \tag{6.10}$$

(we assume the wave function is normalized). *We can get the exact ground state energy and exact ground-state wave function, including electron correlation effects, by minimizing the energy with respect to all normalized, antisymmetric, wave functions.* Unfortunately, test all the normalized wave functions more than we can manage. A simple choice was to take a Slater determinant of orbitals, and this gave us the Hartree-Fock equations. However, there are certainly cases (e.g., the ozone molecule or a diyl intermediate) where, because several Slater determinants are very close together in energy, we expect that describing our system with a single Slater determinant will be a poor approximation. In this case, we might take a cue from the configuration interaction method we used to improve the molecular-orbital approximation to the hydrogen molecule and helium atom—write the wave function as a sum of several Slater determinants,

$$\Psi = \sum_k c_k \Phi_k \tag{6.11}$$

Recall that given a Hermitian operator, the set of all eigenfunctions of that operator forms a complete, orthonormal, set. That is, if we consider a complete set of Slater determinants, then the expansion in Eq. (6.11) will be *exact*. (We used the same trick in perturbation theory, for example, when we wrote the wave function for the “perturbed” system as a linear combination of the wave functions of the “unperturbed” system.) In the

configuration interaction method, we choose, along the lines of the previous discussion, all the possible Slater determinants of Hartree-Fock orbitals,

$$\Psi_{exact} \equiv \sum_{\substack{\text{all Slater} \\ \text{dets. of HF} \\ \text{orbitals}}} c_k \Phi_k \quad (6.12)$$

We expect that the “largest component” will be the ground state Hartree-Fock wave function. Often the other coefficients in Eq. (6.12) are very small.

In order to keep track of all the possible Hartree-Fock orbitals, we often write the ground-state Hartree-Fock wave function as Φ_0 , the Slater determinant with an electron “excited” from the i^{th} occupied orbital to the a^{th} unoccupied orbital as Φ_i^a , the “doubly-excited” Slater determinants as Φ_{ij}^{ab} , etc.. With this notation, we can write the exact wave function as

$$\begin{aligned} \Psi = & c_0 \Phi_0 + \underbrace{\sum_{i=1}^N \sum_{a=N+1}^K c_i^a \Phi_i^a}_{\text{single excitations (S)}} + \underbrace{\sum_{i>j=1}^N \sum_{a>b=N+1}^K c_{ij}^{ab} \Phi_{ij}^{ab}}_{\text{double excitations (D)}} \\ & + \underbrace{\sum_{i>j>k=1}^N \sum_{a>b>c=N+1}^K c_{ijk}^{abc} \Phi_{ijk}^{abc}}_{\text{triple excitations (T)}} + \underbrace{\sum_{i>j>k>l=1}^N \sum_{a>b>c>d=N+1}^K c_{ijkl}^{abcd} \Phi_{ijkl}^{abcd}}_{\text{quadruple excitations (Q)}} + \dots \end{aligned} \quad (6.13)$$

The inequality signs reflect the fact that all that matters (to within a negative sign) is the orbitals that are occupied. For example, all of the following excitations describe the same doubly-excited state: $\Phi_{ij}^{ab} = -\Phi_{ji}^{ab} = -\Phi_{ij}^{ba} = \Phi_{ji}^{ba}$.

Because, in practice, one can never compute ALL the unoccupied Hartree-Fock orbitals, the upper limit in the sum is designed to indicate that there are only K occupied orbitals. It should be clear that the number of terms in this expansion grows *very* rapidly as we consider higher-and-higher levels of excitation.

Sometimes it is helpful to abbreviate the indices on the Slater determinants by introducing vectors, \mathbf{i} and \mathbf{a} , whose components are the orbitals from which an electron is removed (i_1, i_2, \dots) and the orbitals to which it is excited (a_1, a_2, \dots), respectively. If $i_k = a_k$, then there is no excitation at all from orbital $\psi_{i_k}(\mathbf{r})$. By convention, we group the electrons that are excited at the beginning of the vector, so that

$$\begin{aligned}
\Psi = & c_{[1,2,3,\dots,N]}^{[1,2,3,\dots,N]} \Phi_{[1,2,3,\dots,N]}^{[1,2,3,\dots,N]} + \underbrace{\sum_{i_1=1}^N \sum_{a=N+1}^K c_{[i_1,i_2,\dots,i_N]}^{[a_1,i_2,\dots,i_N]} \Phi_{[i_1,i_2,\dots,i_N]}^{[a_1,i_2,\dots,i_N]}}_{\text{single excitations (S)}} + \underbrace{\sum_{i_1>i_2=1}^N \sum_{a_1>a_2=N+1}^K c_{[i_1,i_2,i_3,\dots,i_N]}^{[a_1,a_2,i_3,\dots,i_N]} \Phi_{[i_1,i_2,i_3,\dots,i_N]}^{[a_1,a_2,i_3,\dots,i_N]}}_{\text{double excitations (D)}} \\
& + \underbrace{\sum_{i_1>i_2>i_3=1}^N \sum_{a_1>a_2>a_3=N+1}^K c_{[i_1,i_2,i_3,i_4,\dots,i_N]}^{[a_1,a_2,a_3,i_4,\dots,i_N]} \Phi_{[i_1,i_2,i_3,i_4,\dots,i_N]}^{[a_1,a_2,a_3,i_4,\dots,i_N]}}_{\text{triple excitations (T)}} \\
& + \underbrace{\sum_{i_1>i_2>i_3>i_4=1}^N \sum_{a_1>a_2>a_3>a_4=N+1}^K c_{[i_1,i_2,i_3,i_4,i_5,\dots,i_N]}^{[a_1,a_2,a_3,a_4,i_5,\dots,i_N]} \Phi_{[i_1,i_2,i_3,i_4,i_5,\dots,i_N]}^{[a_1,a_2,a_3,a_4,i_5,\dots,i_N]}}_{\text{quadruple excitations (Q)}} + \dots \\
= & \sum_{i,a} c_i^a \Phi_i^a
\end{aligned} \tag{6.14}$$

To compute the configuration interaction wave function, we start with the Schrödinger equation,

$$H\Psi = E\Psi. \tag{6.15}$$

Next, we multiply both sides by a Slater determinant, Φ_i^a :

$$\Phi_i^a \hat{H} \Psi = \Phi_i^a E \Psi. \tag{6.16}$$

Next, we substitute in the form we are using for the wave function, Eq. (6.14), so that

$$\Phi_i^a \hat{H} \sum_{j,b} c_j^b \Phi_j^b = E \Phi_i^a \sum_{j,b} c_j^b \Phi_j^b \tag{6.17}$$

and integrate, obtaining

$$\begin{aligned}
\sum_{j,b} \langle \Phi_i^a | \hat{H} | \Phi_j^b \rangle c_j^b &= E \sum_{j,b} \langle \Phi_i^a | \Phi_j^b \rangle c_j^b \\
&= E \sum_{j,b} \delta_{ij} \delta_{ab} c_j^b \\
&= E c_i^a
\end{aligned} \tag{6.18}$$

Recalling that the “component form” for the matrix eigenvalue problem, $\mathbf{A}\mathbf{c} = \alpha\mathbf{c}$ is just

$$\sum_{j=1}^n A_{ij} c_j = \alpha c_i, \tag{6.19}$$

we recognize that Eq. (6.18) is just a matrix eigenvalue problem, with the Hamiltonian matrix being

$$H_{i,j}^{a,b} \equiv \langle \Phi_i^a | \hat{H} | \Phi_j^b \rangle. \tag{6.20}$$

Finding the lowest-energy eigenvalue from

$$\sum_{i,a} H_{i,j}^{a,b} c_j^b = E c_i^a \tag{6.21}$$

is entirely equivalent to minimizing the energy with respect to all wave functions with the form of Eq. (6.14).

Of course, we will need to truncate the expansion in Eq. (6.14), and not consider any excitations above some order. When we truncate at zeroth-order, we have the Hartree-Fock method. At first order, we have “Configuration-Interaction with Single excitations” (CIS), at second order we have “Configuration Interaction with Single and Double excitations” (CISD), and so on: CISDT (third order), CISDTQ (fourth order), When we do not truncate the expansion (so that we include N -electron excitations) we say we are doing a Full-Configuration-Interaction calculations, which is called Full-CI (FCI). Full configuration interaction calculations are very, very costly, and their cost is related to the binomial coefficient¹

$$\binom{K}{N} \equiv \frac{(K)!}{(K-N)!N!} \quad (6.22)$$

where K is the total number of Hartree-Fock orbitals that were solved for and N is the number of electrons. For sufficiently large K , Full-CI calculations are essentially exact. The success of all other forms of computation tends to be measured based on “how close to Full-CI” they are.

We now consider the possible ways to truncate the expansion.

II. Single Excitations Only (CIS)

When this series is truncated at “zeroth” order, we have the Hartree-Fock method. One might expect that the “next best thing” would be to truncate the expansion at first-order, including only the single excitations. This method is called “configuration interaction singles” (CIS) and is occasionally used for excited states. However, there is a result called Brillouin’s theorem:

¹ The cost is actually a good bit less than this, owing to the fact that configurations with different spins and symmetries do not affect one another. However, the factorial nature of the cost is inherent in full configuration interaction.

Brillouin's Theorem:²

$$\langle \Phi_i^a | \hat{H} | \Phi_0 \rangle = 0 \quad (6.23)$$

which implies that including the single excitations does not help us describe the ground state. (Brillouin's theorem implies that the Hamiltonian cannot "induce single-excitations" from the Hartree-Fock determinant, and so there is no correction from the single-excitations alone.)

To see this in more detail, we consider the Hartree-Fock wave function, Φ_1^1 and just two single excited states, Φ_1^2 and Φ_1^3 . Then the CIS wave function is

$$\Psi_{CIS} \equiv c_1^1 \Phi_1^1 + c_1^2 \Phi_1^2 + c_1^3 \Phi_1^3. \quad (6.24)$$

Examining Eq. (6.21), it is clear that we need to solve the eigenvalue problem

$$\begin{bmatrix} \langle \Phi_1^1 | \hat{H} | \Phi_1^1 \rangle & \langle \Phi_1^1 | \hat{H} | \Phi_1^2 \rangle & \langle \Phi_1^1 | \hat{H} | \Phi_1^3 \rangle \\ \langle \Phi_1^2 | \hat{H} | \Phi_1^1 \rangle & \langle \Phi_1^2 | \hat{H} | \Phi_1^2 \rangle & \langle \Phi_1^2 | \hat{H} | \Phi_1^3 \rangle \\ \langle \Phi_1^3 | \hat{H} | \Phi_1^1 \rangle & \langle \Phi_1^3 | \hat{H} | \Phi_1^2 \rangle & \langle \Phi_1^3 | \hat{H} | \Phi_1^3 \rangle \end{bmatrix} \begin{bmatrix} c_1^1 \\ c_1^2 \\ c_1^3 \end{bmatrix} = E \begin{bmatrix} c_1^1 \\ c_1^2 \\ c_1^3 \end{bmatrix} \quad (6.25)$$

and, from Brillouin's theorem, we have that all the integrals between the Hartree-Fock wave function and the singly excited states are zero, so

$$\begin{bmatrix} H_{1,1}^{1,1} & 0 & 0 \\ 0 & H_{1,1}^{2,2} & H_{1,1}^{2,3} \\ 0 & (H_{1,1}^{2,3})^* & H_{1,1}^{3,3} \end{bmatrix} \begin{bmatrix} c_1^1 \\ c_1^2 \\ c_1^3 \end{bmatrix} = E \begin{bmatrix} c_1^1 \\ c_1^2 \\ c_1^3 \end{bmatrix} \quad (6.26)$$

which has the eigenvalues

$$\begin{aligned} E_0 &= H_{1,1}^{1,1} \equiv E_{HF} \\ E_{1,2} &= \frac{H_{1,1}^{2,2} + H_{1,1}^{3,3}}{2} \pm \frac{1}{2} \sqrt{(H_{1,1}^{2,2} - H_{1,1}^{3,3})^2 + 4 |H_{1,1}^{2,3}|^2} \end{aligned} \quad (6.27)$$

That is, while there is no correction to the ground state energies, there is a correction to the Hartree-Fock excitation energies. Because CIS calculations are not that difficult, they are commonly used for computing excitation energies. This is sometimes called the Tamm-Dancoff approximation by people who like to sound obscure.

A CIS calculation, then, can be denoted "symbolically" as:

² You should be able to derive Brillouin's theorem as the derivation requires nothing more than routine manipulations of integrals of Slater determinants.

$$\begin{bmatrix} H_{i,i}^{i,i} & 0 & 0 \\ 0 & H_{i,i}^{a,a} & H_{i,j}^{a,b} \\ 0 & (H_{i,j}^{a,b})^* & H_{j,j}^{b,b} \end{bmatrix} \begin{bmatrix} c_i^i \\ c_i^a \\ c_i^b \end{bmatrix} = E \begin{bmatrix} c_i^1 \\ c_i^a \\ c_i^b \end{bmatrix} \quad (6.28)$$

where i, j are confined to the occupied orbitals, while a, b are restricted to the unoccupied orbitals.

III. Single + Double Excitations, CISD

The only type of excited state Slater determinant that interacts with the Hartree-Fock wave function *directly* is the double excitations, where

$$\langle \Phi_{HF} | \hat{H} | \Phi_{i,j}^{a,b} \rangle \equiv \langle \psi_i \psi_j | \frac{1}{|r_i - r_j|} | \psi_a \psi_b \rangle - \langle \psi_i \psi_j | \frac{1}{|r_i - r_j|} | \psi_b \psi_a \rangle. \quad (6.29)$$

We can consider the easy case where we consider one single and one double excitation, so that

$$\Psi \approx c_{i_1 i_2}^{i_1 i_2} \Phi_{i_1 i_2}^{i_1 i_2} + c_{i_1 i_2}^{a_1 i_2} \Phi_{i_1 i_2}^{a_1 i_2} + c_{i_1 i_2}^{a_1 a_2} \Phi_{i_1 i_2}^{a_1 a_2} \quad (6.30)$$

and we find the coefficients by solving the eigenvalue problem

$$\begin{bmatrix} H_{i_1 i_2; i_1 i_2}^{i_1 i_2; i_1 i_2} & 0 & H_{i_1 i_2; a_1 a_2}^{i_1 i_2; i_1 i_2} \\ 0 & H_{i_1 i_2; a_1 i_2}^{a_1 i_2; a_1 i_2} & H_{i_1 i_2; a_1 a_2}^{a_1 i_2; a_1 i_2} \\ (H_{i_1 i_2; a_1 a_2}^{i_1 i_2; a_1 a_2})^* & (H_{i_1 i_2; a_1 i_2}^{a_1 i_2; a_1 i_2})^* & H_{i_1 i_2; i_1 i_2}^{a_1 a_2; a_1 a_2} \end{bmatrix} \begin{bmatrix} c_{i_1 i_2}^{i_1 i_2} \\ c_{i_1 i_2}^{a_1 i_2} \\ c_{i_1 i_2}^{a_1 a_2} \end{bmatrix} = E \begin{bmatrix} c_{i_1 i_2}^{i_1 i_2} \\ c_{i_1 i_2}^{a_1 i_2} \\ c_{i_1 i_2}^{a_1 a_2} \end{bmatrix}. \quad (6.31)$$

or, in simpler notation,

$$\begin{bmatrix} E_{HF} & 0 & H_{02} \\ 0 & H_{11} & H_{12} \\ (H_{02})^* & H_{12}^* & H_{22} \end{bmatrix} \begin{bmatrix} c_{i_1 i_2}^{i_1 i_2} \\ c_{i_1 i_2}^{a_1 i_2} \\ c_{i_1 i_2}^{a_1 a_2} \end{bmatrix} = E \begin{bmatrix} c_{i_1 i_2}^{i_1 i_2} \\ c_{i_1 i_2}^{a_1 i_2} \\ c_{i_1 i_2}^{a_1 a_2} \end{bmatrix} \quad (6.32)$$

We could solve the eigenproblem directly in this case (we can write analytic solutions for anything less than a 5x5 problem), but the result is too complicated to give us much insight. To gain insight, we can approximate the solution by guessing the lowest root of the secular equation using Newton's method. This gives

$$E_0 \approx E_{HF} + \frac{(H_{11} - E_{HF}) |H_{02}|^2}{|H_{12}|^2 + |H_{02}|^2 - (H_{11} - E_{HF})(H_{22} - E_{HF})} \quad (6.33)$$

Note that the single-excited state energy is not necessarily inconsequential: the single-excited state interacts with the doubly-excited state, which in turn interacts with the

ground state through the matrix element H_{02} . However, H_{12} is usually rather small, and so the doubly excited states are more important than the single excited states. In the limit where $H_{12} = 0$, then

$$E_0 = \frac{E_{HF} + H_{22}}{2} - \frac{1}{2} \sqrt{(E_{HF} - H_{22})^2 + 4|H_{02}|^2} \quad (6.34)$$

It is clear that the general CISD eigenproblem has the same form, Eq. (6.31), and will obey similar principles of solution. Unlike Hartree-Fock (cost $\sim K^3$) and CIS (cost $\sim K^4$), the cost of CISD increases as K^6 ,³ roughly proportional to the size of the system raised to the sixth power.)

IV. Multireference Singles and Doubles CI (MRSDCI or MRCISD)

As is apparent from the examples we presented, there are times where several different Slater determinants will be required to describe an ground state. For example, for the ozone molecule both the Slater determinant where the $1a_2$ orbital is doubly occupied, $\Phi_{1a_2} \equiv |\psi_1 \dots \psi_{1a_2} \alpha \psi_{1a_2} \beta|$, and the Slater determinant where the $2b_1$ orbital is doubly occupied, $\Phi_{2b_1} \equiv |\psi_1 \dots \psi_{2b_1} \alpha \psi_{2b_1} \beta| = \Phi_{1a_2\alpha, 1a_2\beta}^{2b_1\alpha, 2b_1\beta}$ are very important. In fact, the second configuration is so important that the Hartree-Fock approximation is totally unacceptable: even a very crude model for ozone requires treating both states. Consequently, we might expect that an appropriate model for the wave function could be obtained by considering all the single and double excitations from the Φ_{1a_2} and all the single and double excitations from Φ_{2b_1} :

$$\begin{aligned} \Psi_{MRSDCI} \approx & c_{1a_2} \Phi_{1a_2} + \underbrace{\sum_{i=1}^N \sum_{a=N+1}^K c_i^a \Phi_i^a}_{\text{single exc. from } \Phi_{1a_2}} + \underbrace{\sum_{i>j=1}^N \sum_{a>b=N+1}^K c_{ij}^{ab} \Phi_{ij}^{ab}}_{\text{double exc. from } \Phi_{1a_2} \text{ excitations (D)}} \\ & + \underbrace{\sum_{i<1a_2}^N \sum_{a=N+1}^K c_i^a \Phi_{i, 1a_2\alpha, 1a_2\beta}^{a, 2b_1\alpha, 2b_1\beta}}_{\text{single exc. from } \Phi_{2b_1}} + \underbrace{\sum_{i_1<j<1a_2}^N \sum_{a>b=N+1}^K \Phi_{i,j, 1a_2\alpha, 1a_2\beta}^{a,b, 2b_1\alpha, 2b_1\beta}}_{\text{double exc. from } \Phi_{2b_1}} \end{aligned} \quad (6.35)$$

Notice that because Φ_{2b_1} is a double excitation of Φ_{1a_2} , Equation (6.35) actually includes triple and quadruple excitations terms. However, because the triples and quadruples being considered are very special (only the types listed on the second line in Eq.

³ More accurate, the cost of CISD is proportional to $N^2(K - N)^4$.

(6.35)), the calculations is relatively easy—very easy compared to considering all the triple and quadruple excitations.

When you know what configurations to employ, MRSDCI is among the most accurate and efficient methods. However, it is difficult to learn which configurations should be used (one esteemed computational chemist calls it a “black art”), and there are very few people who can do so reliably. The problem is that picking out the important configurations must be done before the calculation is done, requiring a deep intuition about the system. (Many people would argue that, for instance, that occupation of the higher-energy $7a_1$ and $5b_2$ orbitals should also be considered, as should be the possibility that the electrons are taken from the $6a_1$ orbital.) There are some people with such deep intuition, and there are “pretenders” (like your prof.).

The role of the different configurations in Eq. (6.35) serves to demonstrate the importance of what is often called “static” versus “dynamic” electron correlation. Static, or non-dynamic, electron correlation is due to the non-crossing rule—the “repelling” of nearly degenerate molecular states. Computing the energy using the wave function

$$\Psi \approx c_{1a_2} \Phi_{1a_2} + c_{2b_1} \Phi_{2b_1} \quad (6.36)$$

would approximate the static electron correlation energy in ozone. Similarly, one could approximate the static electron correlation in Beryllium with $\Psi \approx c_{1s^2 2s^2} \Phi_{1s^2 2s^2} + c_{1s^2 2p^2} \Phi_{1s^2 2p^2}$. Because the coefficient of each Slater determinant is rather large, static electron configuration affects the way electrons are distributed in important ways. For example, in the ozone atom, mixing the $2b_1$ configuration plays a key role in reducing the unfavorable positive charge at the central oxygen atom, which has a formal charge of +1 in the Φ_{1a_2} configuration. There is a sense, then, in which static electron correlation affects the way electrons are arranged with respect to the molecular nuclei.

“Dynamic” correlation has no clear provenance in the non-crossing rule, representing the “repulsion” due to states that are energetically far apart. There is a sense in which dynamic correlation is primarily concerned not with how electrons are arranged about the nuclei, but in how electrons “move” with respect to each other, and incorporates the detailed dependence on the interelectronic distance, $|\mathbf{r}_i - \mathbf{r}_j|$, that the electron-electron repulsion term in the exact Hamiltonian requires.

Based on this classification, in cases where there are “near degeneracies”, there are low-lying doubly excited states, which affect the energy a large amount.⁴ This means that there are relatively important quadruple excited states, and we can include these states by moving beyond the Hartree-Fock reference system. The best, and most computationally affordable, methods are multi-reference singles and doubles CI methods, but they require choosing the appropriate state. The general rule is that:

- When static correlation is important, the Hartree-Fock description of a molecule is poor. Static correlation is *always* addressed by including a *few*, nearly degenerate, Slater determinants.
- Dynamic correlation is always important, and is more “universal”—not depending on the degeneracy or near degeneracy of the ground state. It usually takes *many* Slater determinants to address dynamic correlation—single and double excitations from the important “static-correlation” wave function, at the very least.

What happens if you, naively, perform a CISD calculation for a system like Ozone? Well, you are including both of the very important Slater determinants, Φ_{1a_2} and Φ_{2b_1} , and so static correlation is being accounted for. However, you are only considering single and double excitations relative to the Φ_{1a_2} determinant, and so only this state is being “dynamically correlated.” Equivalently, because the single and double excitations from Φ_{2b_1} were neglected, the effects of dynamic correlation on this state were neglected. We conclude, then, that a CISD calculation will obtain the static correlation and a large amount of the dynamic correlation associated with the Φ_{1a_2} determinant, while neglecting the dynamic correlation of the Φ_{2b_1} determinant. We may reasonably suppose, then, that the MRSDCI calculation will obtain significantly more dynamic correlation than the single-reference CISD calculation, and so when static correlation is important, the CISD wave function tends to be dramatically undercorrelated.

⁴ Sometimes the low-lying states do not have suitable symmetry, and do not interact with the ground state for that reason. (Recall that the non-crossing rule only applies to states with the same symmetry.) In most larger molecules, there is no symmetry, however, and low-lying orbitals always imply important static correlation effects.

V. CISD with Triple and Quadruple Excitations, CISDTQ

Observing Eq. (6.35), we recognize that what occurs in MRSDCI is that we include a few of the “important” triple and quadruple excitations. For the non-savant, it is often difficult to determine what these “important” excitations are, and so we might choose to include *all* the triple and quadruple excitations,

$$\begin{aligned} \Psi_{\text{CISDTQ}} = & c_0 \Phi_0 + \underbrace{\sum_{i=1}^N \sum_{a=N+1}^K c_i^a \Phi_i^a}_{\text{single excitations (S)}} + \underbrace{\sum_{i>j=1}^N \sum_{a>b=N+1}^K c_{ij}^{ab} \Phi_{ij}^{ab}}_{\text{double excitations (D)}} \\ & + \underbrace{\sum_{i>j>k=1}^N \sum_{a>b>c=N+1}^K c_{ijk}^{abc} \Phi_{ijk}^{abc}}_{\text{triple excitations (T)}} + \underbrace{\sum_{i>j>k>l=1}^N \sum_{a>b>c>d=N+1}^K c_{ijkl}^{abcd} \Phi_{ijkl}^{abcd}}_{\text{quadruple excitations (Q)}} \end{aligned} \quad (6.37)$$

This will include static correlations due to single and doubly excited Slater determinants and, less importantly, “higher order” dynamical correlation effects.

What are “higher-order” dynamical correlation effects? In a CISD calculation, the dynamics correlation of doubly-excited states has not been considered. That is, we have not treated the “single and double excitations” of the doubly-excited states. Thus, when we incorporate the triple and quadruple excitations, we are both “correlating” any important “static-correlation” Slater determinants for the system and providing correlation corrections to the “dynamic correlation” Slater determinants. This sort of “second order” correlation is important, but it is a lot less important than the “first-order” dynamic correlation due to single and double excitations. In practice, the order of the importance of the various correlation terms is double excit. \gg quadruple excit. $>$ single excit. \approx triple excit.. The next most important terms are the hextuple-excitations, which serve to correlate the quadruple excited states, which help correlate the double excited states, which correlate the ground state directly.

The above “word picture” is most clear when we consider the various sorts of matrix elements. We know that the single excitations do not affect the ground state energy directly,

$$\langle \Phi_i^a | \hat{H} | \Phi_{\text{HF}} \rangle = 0 \quad (6.38)$$

but the double excitations do.

$$\langle \Phi_{\text{HF}} | \hat{H} | \Phi_{ij}^{ab} \rangle \equiv \langle \psi_i \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle - \langle \psi_i \psi_j | \hat{v}_{ee} | \psi_b \psi_a \rangle \quad (6.39)$$

Each term in Eq. (6.39) has the form of two distributions with zero charge interacting, for instance,

$$\langle \psi_i \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle \equiv \delta_{\sigma_i \sigma_a} \delta_{\sigma_j \sigma_b} \iint \frac{(\psi_i^*(\mathbf{r}') \psi_a(\mathbf{r}')) (\psi_j^*(\mathbf{r}') \psi_b(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \quad (6.40)$$

We expect such terms to be smaller than the interaction of a distribution with itself, and can show that:⁵

$$\begin{aligned} \left| \iint \frac{(\psi_i^*(\mathbf{r}) \psi_a(\mathbf{r})) (\psi_j^*(\mathbf{r}') \psi_b(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \right| &\leq \frac{1}{2} \left(\iint \frac{(\psi_i^*(\mathbf{r}) \psi_a(\mathbf{r})) (\psi_i^*(\mathbf{r}') \psi_a(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \iint \frac{(\psi_j^*(\mathbf{r}) \psi_b(\mathbf{r})) (\psi_j^*(\mathbf{r}') \psi_b(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \right) \\ &\leq \left(\frac{K_{ia} + K_{jb}}{2} \right) \end{aligned} \quad (6.41)$$

The “off diagonal” terms in the Hamiltonian matrix then, being a difference between two small terms, is also small.

Because the one- and two-electron integrals of Hartree products differing by more than two orbitals are all zero, triple, quadruple, and other higher-order excitations do not affect the ground state energy directly either,

$$0 = \langle \Phi_{HF} | \hat{H} | \Phi_i^a \rangle = \langle \Phi_{HF} | \hat{H} | \Phi_{ijk}^{abc} \rangle = \langle \Phi_{HF} | \hat{H} | \Phi_{ijkl}^{abcd} \rangle = \dots \quad (6.42)$$

The single excitations affect the double excitations. Using our rules for Hartree products and Slater determinates, we obtain⁶

⁵ This is a sort of Cauchy inequality also. The key is that because $\frac{1}{|\mathbf{r} - \mathbf{r}'|}$ is positive, we can regard

$\iint f(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} f(\mathbf{r}') d\mathbf{r} d\mathbf{r}'$ as a norm, which measures the magnitude of a function. Then the Cauchy inequality holds in the form $2\|xy\| \leq \|x\|^2 + \|y\|^2$, which gives Eq. (6.48).

⁶ The derivation of Brillouin’s theorem is similar to this analysis.

$$\begin{aligned}
\langle \Phi_i^a | \hat{H} | \Phi_{ij}^{ab} \rangle &= \left\langle \left(\dots \psi_{i-1} \psi_a \psi_{i+1} \dots \left| \sum_{p=1}^N \hat{h}(\mathbf{r}_i) + \sum_{q=p+1}^N \hat{v}_{ee}(\mathbf{r}_p, \mathbf{r}_q) \right| \dots \psi_{i-1} \psi_a \psi_{i+1} \dots \psi_{j-1} \psi_b \psi_{j+1} \dots \right) \right\rangle \\
&= \left\langle \left(\dots \psi_{i-1} \psi_a \psi_{i+1} \dots \left| \sum_{i=1}^N \hat{h}(\mathbf{r}_i) \right| \dots \psi_{i-1} \psi_a \psi_{i+1} \dots \psi_{j-1} \psi_b \psi_{j+1} \dots \right) \right\rangle \\
&\quad \sqrt{N!} \left\langle \left(\dots \psi_{i-1} \psi_a \psi_{i+1} \dots \left| \sum_{p=1}^N \sum_{q=p+1}^N \hat{v}_{ee}(\mathbf{r}_p, \mathbf{r}_q) \right| \dots \psi_{i-1} \psi_a \psi_{i+1} \dots \psi_{j-1} \psi_b \psi_{j+1} \dots \right) \right\rangle \\
&= \langle \psi_j | \hat{h} | \psi_b \rangle + \sum_{\substack{k=1 \\ k \neq j \\ k \neq i}}^N \left(\langle \psi_k \psi_j | \hat{v}_{ee} | \psi_k \psi_b \rangle - \langle \psi_k \psi_j | \hat{v}_{ee} | \psi_b \psi_k \rangle \right) \\
&\quad + \left(\langle \psi_a \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle - \langle \psi_a \psi_j | \hat{v}_{ee} | \psi_b \psi_a \rangle \right) \\
&= \langle \psi_j | \hat{h} | \psi_b \rangle + \left[\sum_{k=1}^N \left(\langle \psi_k \psi_j | \hat{v}_{ee} | \psi_k \psi_b \rangle - \langle \psi_k \psi_j | \hat{v}_{ee} | \psi_b \psi_k \rangle \right) \right. \\
&\quad \left. - \left(\langle \psi_j \psi_j | \hat{v}_{ee} | \psi_j \psi_b \rangle - \langle \psi_j \psi_j | \hat{v}_{ee} | \psi_b \psi_j \rangle \right) \right. \\
&\quad \left. - \left(\langle \psi_i \psi_j | \hat{v}_{ee} | \psi_i \psi_b \rangle - \langle \psi_i \psi_j | \hat{v}_{ee} | \psi_b \psi_i \rangle \right) \right] \\
&\quad + \left(\langle \psi_a \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle - \langle \psi_a \psi_j | \hat{v}_{ee} | \psi_b \psi_a \rangle \right) \\
&= \left\langle \psi_j \left| \hat{h} \psi_b(\mathbf{r}) + \sum_{k=1}^N \int \frac{|\psi_k(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \psi_b(\mathbf{r}) - \int \frac{\psi_k^*(\mathbf{r}') \psi_b(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \psi_k(\mathbf{r}) \right. \right\rangle - 0 \\
&\quad + \left(\langle \psi_a \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle - \langle \psi_a \psi_j | \hat{v}_{ee} | \psi_b \psi_a \rangle \right) - \left(\langle \psi_i \psi_j | \hat{v}_{ee} | \psi_i \psi_b \rangle - \langle \psi_i \psi_j | \hat{v}_{ee} | \psi_b \psi_i \rangle \right) \\
&= \langle \psi_j | \hat{f} | \psi_b \rangle + \left(\langle \psi_a \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle - \langle \psi_a \psi_j | \hat{v}_{ee} | \psi_b \psi_a \rangle \right) \\
&\quad - \left(\langle \psi_i \psi_j | \hat{v}_{ee} | \psi_i \psi_b \rangle - \langle \psi_i \psi_j | \hat{v}_{ee} | \psi_b \psi_i \rangle \right) \tag{6.43} \\
&= \langle \psi_j | \varepsilon_b \psi_b \rangle + \left(\langle \psi_a \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle - \langle \psi_a \psi_j | \hat{v}_{ee} | \psi_b \psi_a \rangle \right) \\
&\quad - \left(\langle \psi_i \psi_j | \hat{v}_{ee} | \psi_i \psi_b \rangle - \langle \psi_i \psi_j | \hat{v}_{ee} | \psi_b \psi_i \rangle \right) \\
&= \varepsilon_b \cdot 0 + \left(\langle \psi_a \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle - \langle \psi_a \psi_j | \hat{v}_{ee} | \psi_b \psi_a \rangle \right) \\
&\quad - \left(\langle \psi_i \psi_j | \hat{v}_{ee} | \psi_i \psi_b \rangle - \langle \psi_i \psi_j | \hat{v}_{ee} | \psi_b \psi_i \rangle \right) \\
&= \left(\langle \psi_a \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle - \langle \psi_a \psi_j | \hat{v}_{ee} | \psi_b \psi_a \rangle \right) - \left(\langle \psi_i \psi_j | \hat{v}_{ee} | \psi_i \psi_b \rangle - \langle \psi_i \psi_j | \hat{v}_{ee} | \psi_b \psi_i \rangle \right)
\end{aligned}$$

The first and third terms can be written as an interaction energy between a distribution with charge one and a distribution with charge zero, for instance,

$$\langle \psi_a \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle = \delta_{\sigma_j \sigma_b} \iint \frac{|\psi_a(\mathbf{r})|^2 (\psi_j^*(\mathbf{r}') \psi_b(\mathbf{r}'))}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{6.44}$$

and, using the Cauchy inequality

$$\left| \psi_j^*(\mathbf{r}') \psi_b(\mathbf{r}') \right| \leq \frac{|\psi_j^*(\mathbf{r}')|^2 + |\psi_b(\mathbf{r}')|^2}{2} \tag{6.45}$$

we have that

$$\begin{aligned}
\langle \psi_a \psi_j | \hat{v}_{ee} | \psi_a \psi_b \rangle &= \delta_{\sigma_j \sigma_b} \iint \frac{|\psi_a(\mathbf{r})|^2 (\psi_j^*(\mathbf{r}') \psi_b(\mathbf{r}'))}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\
&< \delta_{\sigma_j \sigma_b} \iint \frac{|\psi_a(\mathbf{r})|^2 |\psi_j^*(\mathbf{r}') \psi_b(\mathbf{r}')|}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\
&< \frac{\delta_{\sigma_j \sigma_b}}{2} \left(\iint \frac{|\psi_a(\mathbf{r})|^2 |\psi_j^*(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \iint \frac{|\psi_a(\mathbf{r})|^2 |\psi_b^*(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \right) \\
&< \delta_{\sigma_j \sigma_b} \left(\frac{J_{aj} + J_{ab}}{2} \right)
\end{aligned} \tag{6.46}$$

Because integrals like Eq. (6.44) represent an interaction energy between a distribution with unit charge and a distribution with zero charge, it should be unsurprising that it is always less than comparable Coulomb integrals. We also note that the first and third integrals have in Eq. (6.43) have the same form but opposite signs; some cancellation between these terms can occur.

The second and fourth terms are interaction energies between two different distributions, both with zero charge. For example,

$$\langle \psi_a \psi_j | \hat{v}_{ee} | \psi_b \psi_a \rangle \equiv \delta_{\sigma_a \sigma_b} \delta_{\sigma_j \sigma_a} \iint \frac{(\psi_a^*(\mathbf{r}) \psi_b(\mathbf{r})) (\psi_j^*(\mathbf{r}') \psi_a(\mathbf{r}'))}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{6.47}$$

As per our discussion of the double excitations, the interaction of a charge distribution with itself is larger than the interaction of the charge density with another distribution, so

$$\begin{aligned}
\left| \iint \frac{(\psi_a^*(\mathbf{r}) \psi_b(\mathbf{r})) (\psi_j^*(\mathbf{r}') \psi_a(\mathbf{r}'))}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \right| &\leq \frac{1}{2} \left(\iint \frac{(\psi_a^*(\mathbf{r}) \psi_b(\mathbf{r})) (\psi_a^*(\mathbf{r}') \psi_b(\mathbf{r}'))}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + \iint \frac{(\psi_j^*(\mathbf{r}) \psi_a(\mathbf{r})) (\psi_j^*(\mathbf{r}') \psi_a(\mathbf{r}'))}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \right) \\
&\leq \left(\frac{K_{ab} + K_{aj}}{2} \right)
\end{aligned} \tag{6.48}$$

In any event, we are led to conclude that the coupling between singly and doubly excited determinants is weak.

The matrix element between single and triple excitations is similar to the matrix element between the Hartree-Fock wave function and the doubles, as is the matrix element between double and quadruple excitations:

$$\begin{aligned}
\langle \Phi_i^a | \hat{H} | \Phi_{ijk}^{abc} \rangle &= \langle \psi_j \psi_k | \hat{v}_{ee} | \psi_b \psi_c \rangle - \langle \psi_j \psi_k | \hat{v}_{ee} | \psi_c \psi_b \rangle \\
\langle \Phi_{ij}^{ab} | \hat{H} | \Phi_{ijkl}^{abcd} \rangle &= \langle \psi_k \psi_l | \hat{v}_{ee} | \psi_c \psi_d \rangle - \langle \psi_k \psi_l | \hat{v}_{ee} | \psi_d \psi_c \rangle.
\end{aligned} \tag{6.49}$$

Similarly, the size of the matrix elements coupling double and triple excitations and triple and quadruple excitations will resemble the interaction between the single and double excitations (but the form is bit more complicated). In the end, the Hamiltonian matrix for which the eigenvalue problem needs to be solved can be written

$$\begin{bmatrix}
E_{HF} & 0 & \langle \Phi_0 | \hat{H} | \Phi_{ij}^{ab} \rangle & 0 & 0 & 0 & \dots \\
0 & \langle \Phi_0 | \hat{H} | \Phi_i^a \rangle & \langle \Phi_i^a | \hat{H} | \Phi_{ij}^{ab} \rangle & \langle \Phi_i^a | \hat{H} | \Phi_{ijk}^{abc} \rangle & 0 & 0 & \dots \\
\langle \Phi_0 | \hat{H} | \Phi_{ij}^{ab} \rangle^* & \langle \Phi_i^a | \hat{H} | \Phi_{ij}^{ab} \rangle^* & \langle \Phi_{ij}^{ab} | \hat{H} | \Phi_{ij}^{ab} \rangle & \langle \Phi_{ij}^{ab} | \hat{H} | \Phi_{ijk}^{abc} \rangle & \langle \Phi_{ij}^{ab} | \hat{H} | \Phi_{ijkl}^{abcd} \rangle & 0 & \dots \\
0 & \langle \Phi_i^a | \hat{H} | \Phi_{ijk}^{abc} \rangle^* & \langle \Phi_{ij}^{ab} | \hat{H} | \Phi_{ijk}^{abc} \rangle^* & \langle \Phi_{ijk}^{abc} | \hat{H} | \Phi_{ijk}^{abc} \rangle & \langle \Phi_{ijk}^{abc} | \hat{H} | \Phi_{ijkl}^{abcd} \rangle & \langle \Phi_{ijk}^{abc} | \hat{H} | \Phi_{ijklm}^{abcde} \rangle & \\
0 & 0 & \langle \Phi_{ij}^{ab} | \hat{H} | \Phi_{ijkl}^{abcd} \rangle^* & \langle \Phi_{ijk}^{abc} | \hat{H} | \Phi_{ijkl}^{abcd} \rangle^* & \langle \Phi_{ijkl}^{abcd} | \hat{H} | \Phi_{ijkl}^{abcd} \rangle & \langle \Phi_{ijkl}^{abcd} | \hat{H} | \Phi_{ijklm}^{abcde} \rangle & \\
0 & 0 & 0 & \langle \Phi_{ijk}^{abc} | \hat{H} | \Phi_{ijklm}^{abcde} \rangle^* & \langle \Phi_{ijkl}^{abcd} | \hat{H} | \Phi_{ijklm}^{abcde} \rangle^* & \langle \Phi_{ijklm}^{abcde} | \hat{H} | \Phi_{ijklm}^{abcde} \rangle & \\
\vdots & \vdots & \vdots & & & & \ddots
\end{bmatrix} \quad (6.50)$$

Equation (6.50) is the foundation for our preceding discussion: quadruple excitations are important because they make the same sorts of corrections to doubly-excited determinants that the doubly-excited determinants make to the Hartree-Fock wave function. Singly and triply excited determinants are also important (but marginally less so). Higher-order excitations are less important: for example, sixth-order corrections correct the quadruply excited determinants, which corrects the doubly excited determinants, which corrects the ground state wave function: this gives a correction to a correction to a correction. For most systems with rather few electrons, these higher excitation levels are not that important. (The corrections are more important for systems with more electrons but, for systems with many electrons, it becomes impossible to perform such accurate calculations, so this is a moot point.)

How expensive is a CISDTQ calculation? A CISDTQ calculations cost is proportional to K^{10} .⁷ This is practically impossible for anything except a small molecule, which is why there is so much interest in multi-reference methods.

VI. Describing “chemistry” with Configuration Interaction

Configuration Interaction calculations can be particularly problematic when we try to consider chemical processes. The first problem is again the problem of static correlation: even if there is only one important configuration near the equilibrium geometry, for different molecular geometries there will often be additional important states. A simple and important case occurs near the transition state in many chemical reactions, wherein the gap between the ground and first excited state is often small. Because static correlation will be

⁷ More precisely and generally, the cost of a configuration interaction calculation including all excited state determinants with 1,2,3,... k excitations is $\text{cost} \sim (N)^k (K - N)^{k+2}$.

important, some quadruple excitations are needed to describe the “reaction path.” For example, for Cope rearrangements that occur via a diyl intermediate, there are two important configurations at for this intermediate, but only one important configuration when the molecule resembles either reactants or products. Describing such reactions, then, requires (at minimum) either CISDTQ or MRSDCI. You can imagine how frustrating it is to try to find the reference determinants for describing chemical processes: one must identify not only the important determinants in the reactant and product state, but also include such determinants as are necessary to describe the static correlation of intermediates and transition states. Moreover, one must identify these states at the beginning of one’s theoretical treatment. In particular, when the purpose of a calculation is to fill gaps in experimental knowledge, it seems unlikely that you will be able to, with any certainty, select the correct reference configurations. You will need to be lucky or good, or otherwise use a treatment with quadruple excitations.

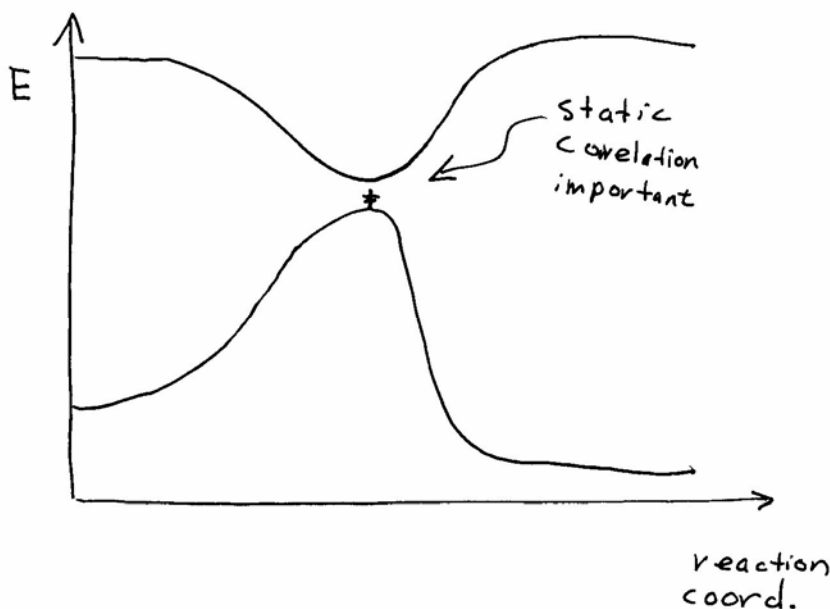


Figure: Static Correlation can be important near transitions states, which are often associated with “avoided crossings” of potential energy curves.

Molecular dissociation processes are particular problematic for two reasons. First of all, static correlation effects can be important as bonds are stretched, since bonding and anti-bonding orbitals tend to approach each others energy as the bond length increases.

Second, there is the problem of *size consistency*. A computational method is said to be size consistent if it predicts that the energy of two atoms or molecules, A and B , that are infinitely far apart is the same as the sum of the energies of the isolated systems,

$$\lim_{R \rightarrow \infty} E \left(\underset{\text{R}}{\text{A-B}} \right) = E(A) + E(B) \quad (6.51)$$

Configuration interaction calculations (except for Full-CI) are not size consistent. For example, consider a CISD calculation on the Helium dimer ($A = B = \text{He}$ in Eq. (6.51)). Since Helium has only two electrons, the energy of the Helium atom is computed exactly by an CISD calculation, which amounts to Full-CI for this system. However, for two Helium atoms very far apart, Full-CI requires a CISDTQ calculation, and the CISD calculation does not allow configurations where both Helium atoms are simultaneously in an excited state. More explicitly, let 1_l denote the α -spin 1s-orbital on the “left” Helium atom and $\bar{1}_l$ denote the β -spin 1s-orbital for the left Helium atom, with similar notation for the “right” atom and also the 2s orbital. We ignore higher excitations. The wave function for the “left hand” Helium atom has four configurations:

$$\Psi_{\text{He}}(z_1, z_2) = c_{1_l \bar{1}_l} \Phi_{1_l \bar{1}_l}^{1_l \bar{1}_l} + c_{1_l \bar{1}_l}^{2_l} \Phi_{1_l \bar{1}_l}^{2_l \bar{1}_l} + c_{1_l \bar{1}_l}^{1_l \bar{2}_l} \Phi_{1_l \bar{1}_l}^{1_l \bar{2}_l} + c_{1_l \bar{1}_l}^{2_l \bar{2}_l} \Phi_{1_l \bar{1}_l}^{2_l \bar{2}_l} \quad (6.52)$$

and the wave function for the Helium dimer has

$$\begin{aligned} \Psi_{\text{He-He}}(z_1, z_2, z_3, z_4) = & c_{1_l \bar{1}_l 1_r \bar{1}_r} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{1}_l 1_r \bar{1}_r} + c_{1_l \bar{1}_l 1_r \bar{1}_r}^{2_l} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{2_l \bar{1}_l 1_r \bar{1}_r} + c_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{2}_l} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{2}_l 1_r \bar{1}_r} + c_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{1}_l 2_r} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{1}_l 2_r \bar{1}_r} \\ & + c_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{1}_l 2_r} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{1}_l 1_r \bar{2}_r} + c_{1_l \bar{1}_l 1_r \bar{1}_r}^{2_l \bar{2}_l 1_r \bar{1}_r} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{2_l \bar{2}_l 1_r \bar{1}_r} + c_{1_l \bar{1}_l 1_r \bar{1}_r}^{2_l \bar{1}_l 2_r \bar{1}_r} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{2_l \bar{1}_l 2_r \bar{1}_r} + c_{1_l \bar{1}_l 1_r \bar{1}_r}^{2_l \bar{1}_l 1_r \bar{2}_r} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{2_l \bar{1}_l 1_r \bar{2}_r} \\ & + c_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{2}_l 2_r \bar{1}_r} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{2}_l 2_r \bar{1}_r} + c_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{2}_l 1_r \bar{2}_r} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{2}_l 1_r \bar{2}_r} + c_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{1}_l 2_r \bar{2}_r} \Phi_{1_l \bar{1}_l 1_r \bar{1}_r}^{1_l \bar{1}_l 2_r \bar{2}_r} \end{aligned} \quad (6.53)$$

However, the exact wave function for two separated Helium atoms (so that the correlation between the electrons in the right-side and left-side atoms is small) is

$$\Psi_{\text{He}_l + \text{He}_r}(z_1, z_2, z_3, z_4) = \mathcal{A} \left[\begin{aligned} & \left(c_{1_l \bar{1}_l} \Phi_{1_l \bar{1}_l}^{1_l \bar{1}_l} + c_{1_l \bar{1}_l}^{2_l} \Phi_{1_l \bar{1}_l}^{2_l \bar{1}_l} + c_{1_l \bar{1}_l}^{1_l \bar{2}_l} \Phi_{1_l \bar{1}_l}^{1_l \bar{2}_l} + c_{1_l \bar{1}_l}^{2_l \bar{2}_l} \Phi_{1_l \bar{1}_l}^{2_l \bar{2}_l} \right) \\ & \times \left(c_{1_r \bar{1}_r} \Phi_{1_r \bar{1}_r}^{1_r \bar{1}_r} + c_{1_r \bar{1}_r}^{2_r} \Phi_{1_r \bar{1}_r}^{2_r \bar{1}_r} + c_{1_r \bar{1}_r}^{1_r \bar{2}_r} \Phi_{1_r \bar{1}_r}^{1_r \bar{2}_r} + c_{1_r \bar{1}_r}^{2_r \bar{2}_r} \Phi_{1_r \bar{1}_r}^{2_r \bar{2}_r} \right) \end{aligned} \right] \quad (6.54)$$

where $\mathcal{A}\Phi_l\Phi_r$ denotes the antisymmetric product of the wave functions for the left- and right-side atoms. However, Eq. (6.54) will include terms contain four excitations, which are not in Eq. (6.53), and so the energy of the two separated Helium atoms (from the wave function in Eq. (6.53)) is greater than the sum of the energies of the two Helium atoms (from the wave function in Eq. (6.54)).

The same argument can clearly be made for any dissociating system: the “dissociated” molecule is described less accurately than its fragments alone, so that, for anything less than Full-CI,

$$\lim_{R \rightarrow \infty} E \left(\underbrace{A-B}_{\text{R}} \right) > E(A) + E(B). \quad (6.55)$$

For small molecules, the size consistency error is often approximately removed using the “Davidson correction” formula,

$$E = E_{\text{CISD}} + (1 - c_0^2)(E_{\text{CISD}} - E_{\text{HF}}), \quad (6.56)$$

where E_{CISD} is the energy from the CISD calculation, E_{HF} is the energy from the Hartree-Fock calculation, and c_0 is the coefficient of the Hartree-Fock Slater determinant in the CISD wave function, cf. Eq. (6.13). For MRSDCI and CISDTQ calculations, the size-consistency error

$$\varepsilon_{\text{size consistency}} \equiv \left(\lim_{R \rightarrow \infty} E \left(\underbrace{A-B}_{\text{R}} \right) \right) - E(A) + E(B) \quad (6.57)$$

is often smaller than the error in the correlation energy, and so the error in the description of molecular dissociation is primarily due to the imperfect treatment of dynamical correlation, not the problem of size consistency.

As a general rule, I prefer not to use configuration interaction calculations when describing molecular dissociation and association reactions.

VII. Interpretation of Configuration-Interaction Calculations

People often ask what a wave function with the form of Eq. (6.13),

$$\begin{aligned} \Psi = c_0 \Phi_0 &+ \underbrace{\sum_{i=1}^N \sum_{a=N+1}^K c_i^a \Phi_i^a}_{\text{single excitations (S)}} + \underbrace{\sum_{i>j=1}^N \sum_{a>b=N+1}^K c_{ij}^{ab} \Phi_{ij}^{ab}}_{\text{double excitations (D)}} \\ &+ \underbrace{\sum_{i>j>k=1}^N \sum_{a>b>c=N+1}^K c_{ijk}^{abc} \Phi_{ijk}^{abc}}_{\text{triple excitations (T)}} + \underbrace{\sum_{i>j>k>l=1}^N \sum_{a>b>c>d=N+1}^K c_{ijkl}^{abcd} \Phi_{ijkl}^{abcd}}_{\text{quadruple excitations (Q)}} + \dots \end{aligned} \quad (6.58)$$

has to do with electron correlation. After all, the problem with the Hartree-Fock theory that the intricate and detailed collective motions of the electrons induced by the electron-electron repulsion term were “averaged over”, so that the electrons move essentially independently. Why does adding excited-state determinants help? Some insight can be gleaned from analysis like that in problem VI of homework set 3, but a hand-waving argument may be instructive.

- The Single excitation, Φ_i^a represents “moving one electron from orbital $\psi_i(\mathbf{z})$ to orbital $\psi_j(\mathbf{z})$ ”, while all the other electrons are “left alone.” So we are

moving one electron, at a time, independently and, if you like to think in “real space”, you can think of this as describing how the wave function changes as we move an electron from a place, \mathbf{r} , where $\psi_i(\mathbf{r})$ is large to a place, \mathbf{r}' , where $\psi_a(\mathbf{r}')$ is large. However, we are only moving one electron at a time—we are moving electrons “independently.” But this is the essence of the Hartree-Fock approximation: “move electrons independently.” We expect, then, that the single excitations do not correct the Hartree-Fock energy. This is Brillouin’s theorem.

- The double excitation, Φ_{ij}^{ab} , describes two electrons moving together. One electron, with spin and position \mathbf{z}_α , moves from the state described by the orbital $\psi_i(\mathbf{z}_\alpha)$ to the state described by the orbital $\psi_a(\mathbf{z}_\alpha)$, at the same time that another electron, with spin and position \mathbf{z}_β , moves from $\psi_j(\mathbf{z}_\beta)$ to $\psi_b(\mathbf{z}_\beta)$. Double excitations describe the way two electrons move together or, if you like, the “correlation” between two different electrons’ motions.
- The triple excitation, Φ_{ijk}^{abc} , describes three electrons, with coordinates $\mathbf{z}_\alpha, \mathbf{z}_\beta, \mathbf{z}_\gamma$, moving together. One electron moves from $\psi_i(\mathbf{z}_\alpha)$ to $\psi_a(\mathbf{z}_\alpha)$, another electron moves from $\psi_j(\mathbf{z}_\beta)$ to $\psi_b(\mathbf{z}_\beta)$, and a third electron moves from $\psi_k(\mathbf{z}_\gamma)$ to $\psi_c(\mathbf{z}_\gamma)$. So triple excitations describe “three body correlations.”
- The quadruple excitation, Φ_{ijkl}^{abcd} , describes four electrons, with coordinates $\mathbf{z}_\alpha, \mathbf{z}_\beta, \mathbf{z}_\gamma, \mathbf{z}_\delta$, moving together. One electron moves from $\psi_i(\mathbf{z}_\alpha)$ to $\psi_a(\mathbf{z}_\alpha)$, another electron moves from $\psi_j(\mathbf{z}_\beta)$ to $\psi_b(\mathbf{z}_\beta)$, a third electron moves from $\psi_k(\mathbf{z}_\gamma)$ to $\psi_c(\mathbf{z}_\gamma)$, and the fourth electron moves from $\psi_l(\mathbf{z}_\delta)$ to $\psi_d(\mathbf{z}_\delta)$. So quadruple excitations describe “four-electron correlations”.