

## Including Electron Correlation in Molecular Orbital Theory

### Dynamical vs. Non-dynamical Electron Correlation

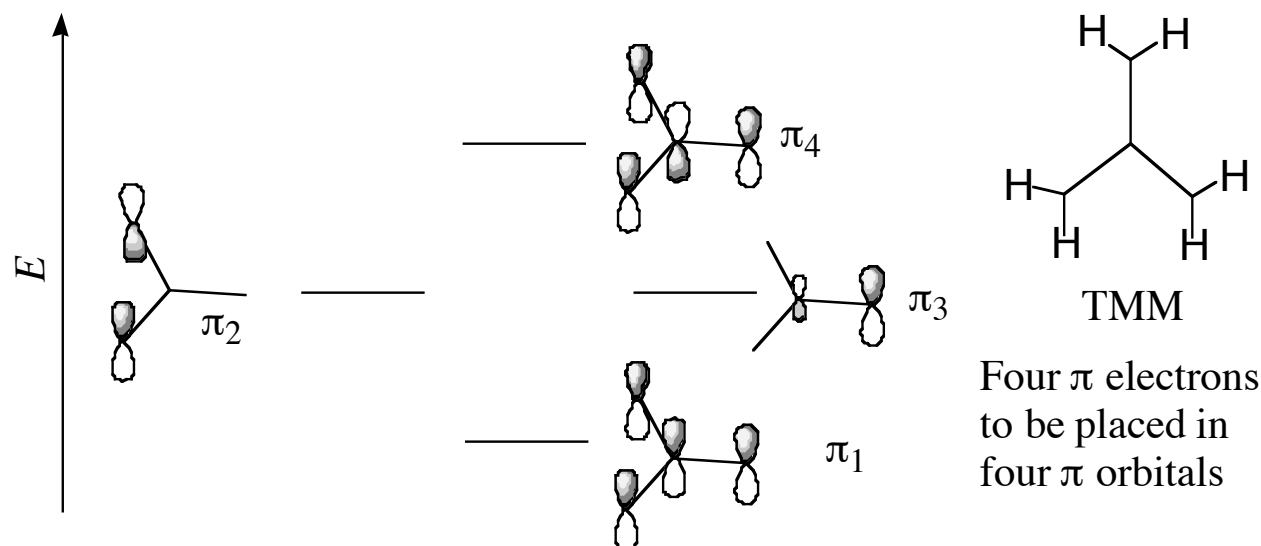
How we might modify the HF wave function to obtain a lower electronic energy? By the variational principle, such a construction would be a more accurate wave function. We cannot do better than the HF wave function with a single determinant – one obvious choice is to construct a wave function as linear combination of multiple determinants:

$$\Psi = c_0 \Psi_{HF} + c_1 \Psi_1 + c_2 \Psi_2 + \dots$$

where the coefficients  $c$  reflect the weight of each determinant in the expansion and also ensure normalization.

For the majority of chemical species, the main error in the HF approximation derives from ignoring the correlated motion of each electron with every other – ‘dynamical correlation’ – refers to the dynamical character of the electron-electron interaction. For most systems the HF wave function dominates in the linear combination; even though the correlation energy may be large, it tends to be made up from a sum of individually small contributions from other determinants.

However, in some cases one of or more of the other determinants may have coefficients of similar magnitude to that for the HF wave function.



Consider an example: trimethylenemethane. In  $D_{3h}$  symmetry, TMM has two degenerate frontier orbitals for which only two electrons are available. Following a molecular analog of Hund's rule, the molecule has a triplet ground state, but here we are concerned with the best closed-shell singlet. If we carry out a closed-shell HF calculation, one or other of the degenerate frontier pairs will be chosen to be occupied, the calculation will optimize the shapes of all of the occupied orbitals, and we will find a best possible single-Slater-determinantal wave function formed from these MOs. But it is obvious that an equally good wave function might have been formed if the original guess had chosen to populate the other of the two degenerate frontier orbitals – we might expect each of these two different HF determinants to contribute with equal weights to the overall wave function expansion. This kind of electron correlation, where different determinants have similar weights because of near (or exact) degeneracy in frontier orbitals, is called 'non-dynamical correlation'. This emphasizes that the error here is not so much that the HF approximation ignores the correlation motion of the electrons, but rather that the HF process is constructed in an intrinsically single-determinantal fashion.

## Multi-configuration self-consistent field theory

Let's say that we have carried out a HF calculation where the frontier orbital chosen to

be occupied was  $\pi_2$ :

$$\Psi_{HF} = |\cdots \pi_1^2 \pi_2^2 \pi_3^0 \rangle$$

and orbital  $\pi_3$  will be empty (a virtual orbital). We might generate the alternative determinant by keeping the same MOs but simply switching the occupation numbers:

$$\Psi_{\pi_2 \rightarrow \pi_3} = |\cdots \pi_1^2 \pi_2^0 \pi_3^2 \rangle$$

Or, we can require the HF calculation to populate  $\pi_3$  initially.

Then

$$\Psi'_{HF} = |\cdots \pi_1'^2 \pi_2'^0 \pi_3'^2 \rangle$$

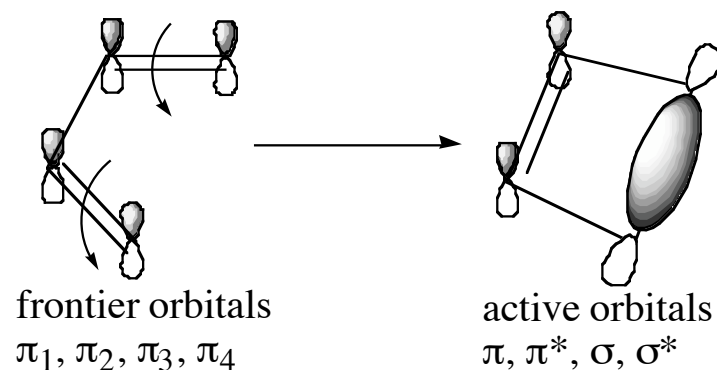
Since different orbitals were occupied during the SCF process, the shapes of all orbitals will be different comparing one HF function to the other. From the nature of the system, we would really like  $\pi_2$  and  $\pi_3$  to be treated equivalently during the orbital optimization process. We would like to find the best orbital shapes for these MOs so as to minimize the energy of the two-configuration wave function:

$$\Psi_{MCSCF} = a_1 |\cdots \pi_1^2 \pi_2^2 \rangle + a_2 |\cdots \pi_1^2 \pi_3^2 \rangle$$

where  $a_1$  and  $a_2$  account for normalization and relative weighting. Such a wave function – ‘multiconfiguration self-consistent-field’ (MCSCF) wave function.

Selection of orbitals to include in an MCSCF requires a consideration of chemistry. In the TMM example, a two-configuration wave function is probably not a very good choice. When we consider the orbitals, which belong to a  $\pi$  system, it is a good idea to include all of them, because usually they are all fairly close to one another in energy – a more complete active space for TMM would include all four  $\pi$  orbitals and the possible ways to distribute four  $\pi$  electrons within them. MCSCF active space choices are abbreviated as ‘( $m,n$ )’ where  $m$  is the number of electrons and  $n$  is a number of orbitals – (4,4) for TMM.

Sometimes, we study reactions that involve substantial change in bonding – it is critical to make a consistent choice of active orbitals. Example: cyclization of 1,3-butadiene to cyclobutene:



The frontier orbitals of butadiene are associated with the  $\pi$  system, however, the cyclization reaction transforms two  $\pi$  bonds into one different  $\pi$  bond and one new  $\sigma$  bond – a consistent (4,4) choice in cyclobutene would involve the  $\pi$  and  $\pi^*$  orbitals and the  $\sigma$  and  $\sigma^*$  orbitals of the new single bond.

How to generally allow the distribution of the electrons in the active space? It is clear that we want to include all possible electron distributions (or ‘configuration state functions’, CSF, which refer to the molecular spin state and the occupation numbers of the orbitals). If we again consider TMM, we can include the following CSFs:

$$\Psi_{MCSCF} = a_1 |\cdots \pi_1^2 \pi_2^2 \pi_3^0 \pi_4^0 \rangle + a_2 |\cdots \pi_1^2 \pi_2^0 \pi_3^2 \pi_4^0 \rangle + a_3 |\cdots \pi_1^2 \pi_2^0 \pi_3^0 \pi_4^2 \rangle \\ + a_4 |\cdots \pi_1^0 \pi_2^2 \pi_3^2 \pi_4^0 \rangle + a_5 \left( |\cdots \pi_1^2 \pi_2^1 \bar{\pi}_3^1 \pi_4^0 \rangle + |\cdots \pi_1^2 \bar{\pi}_2^1 \pi_3^1 \pi_4^0 \rangle \right)$$

where the electron in a singly occupied orbital has  $\alpha$  spin unless the orbital has a bar over it, in which case it has  $\beta$  spin.

We may try to decide arbitrarily (or based on some chemical considerations), which CSFs to include in some particular expansion for MCSCF calculations. An alternative approach – to include all possible configurations in the expansion.

In general, the number  $N$  of singlet CSFs that can be formed from the distribution of  $m$  electrons in  $n$  orbitals is

$$N = \frac{n!(n+1)!}{\left(\frac{m}{2}\right)!\left(\frac{m}{2}+1\right)!\left(n-\frac{m}{2}\right)!\left(n-\frac{m}{2}+1\right)!}$$

(4,4) –  $N = 20$ . If we include all possible arrangements of electrons in MCSCF, the method is called ‘complete active space’ SCF – CASSCF or CAS.  $N$  increases very quickly with the active space size. CASSCF calculations of methanol ( $\text{CH}_3\text{OH}$ ) including all valence electrons: (14,12) –  $N = 169884$  – very expensive task.

### **Full Configuration Interaction**

What if we carry out a CASSCF calculation for all electrons including all orbitals in the complete active space – ‘full configuration interaction’ or ‘full CI’. Within the choice of basis set, this is the best possible calculation that can be done, because it considers the contribution of every possible CSF. Thus, a full CI with an infinite basis set is an ‘exact’ solution of the Schrödinger equation (non-relativistic, Born-Oppenheimer, time-independent).

The number of CSFs in a full CI can be staggeringly large. The trouble is not the number of electrons, which is a constant, but the number of basis functions. Methanol with the 6-31G(d) basis set – the total number of basis functions is 38 – (14,38) active space – we must optimize  $2.4 \times 10^{13}$  expansion coefficients, and this is really a rather small basis set for chemical purposes.

Full CI calculations with large basis sets are usually carried out for only the smallest of molecules. In larger systems, the practical restriction to smaller basis sets makes full CI calculations less chemically interesting, but, as an optimal limit, they permit evaluation of the quality of other methods for including electron correlation.

## Configuration Interaction

### Single-determinant reference

If we consider all possible excited configurations that can be generated from the HF determinant, we have a full CI, but such a calculation is too demanding to accomplish. What if we reduce the CI problem by allowing only a limited number of excitations? How many we should include? Let's rewrite our wave function:

$$\Psi = a_0 \Psi_{HF} + \sum_i \sum_r^{occ. \ vir.} a_i^r \Psi_i^r + \sum_{i < j} \sum_{r < s}^{occ. \ vir.} a_{ij}^{rs} \Psi_{ij}^{rs} + \dots$$

$i$  and  $j$  are occupied MOs in the HF 'reference' wave function,  $r$  and  $s$  are virtual MOs in  $\Psi_{HF}$  and the additional CSFs are generated by exciting an electron from the occupied orbital(s) indicated by subscripts into the virtual orbitals indicated by superscripts. Thus, the second term (after  $\Psi_{HF}$ ) in the summation includes all possible single electronic excitations, and the third includes all possible double excitations, etc.

If we do not have any problem with non-dynamical correlation, we do not need to reoptimize MOs in each determinant. Then, the problem is reduced to determining the expansion coefficients for each excited CSF that is included.

The energies  $E$  of  $N$  different CI wave functions (corresponding to different variationally determined sets of coefficients) can be determined from the  $N$  roots of the CI secular equation:

$$\begin{vmatrix} H_{11} - E & H_{12} & \cdots & H_{1N} \\ H_{21} & H_{22} - E & \cdots & H_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1} & H_{N2} & \cdots & H_{NN} - E \end{vmatrix} = 0$$

$$H_{mn} = \langle \Psi_m | H | \Psi_n \rangle \quad H - \text{the Hamiltonian operator}$$

A special case – the matrix element between the HF determinant and a singly excited CSF:

$$H_{1n} = \langle \Psi_{HF} | H | \Psi_i^r \rangle = \langle \phi_r | F | \phi_i \rangle$$

$F$  is the Fock operator and  $i$  and  $r$  are the occupied and virtual HF orbitals in the single excitation. Since these orbitals are eigenfunctions of the Fock operator, we have

$$\langle \phi_r | F | \phi_i \rangle = \varepsilon_i \langle \phi_r | \phi_i \rangle = \varepsilon_i \delta_{ir}$$

All matrix elements between the HF determinants and singly excited determinants are zero – Brillouin theorem (1934).

How we truncate the CI expansion?

- 1) CIS – only single excitations are included – no use for ground states but the simplest method (corresponds to HF) for excited states.
- 2) CID – only double excitations are included.
- 3) CISD – single and double excitations are included – the most popular single-reference CI method.

The scaling for CISD –  $N^6$  in the large basis limit. Some orbitals can be frozen in the generation of excited states – a popular choice is to keep the core orbitals frozen. One of the most appealing features of CISD – it is variational – the CISD energy represents an upper bound on the exact energy. The main unattractive feature – it is not ‘size consistent’. This means that if you calculate the CISD energy of two separated non-interacting fragments together, it will not be exactly equal to the sum of the CISD energies calculated separately for the two fragments. This creates problems and inaccuracies in calculations of, for example, bond rupture energies.

## **Multireference**

The formalism for multireference configuration interaction (MRCI) is similar to that for single-reference CI – instead of the HF wave function serving as a reference, an MCSCF wave function is used. The significant improvement of virtual orbitals usually makes CI more rapidly convergent. Nevertheless, the number of matrix elements needed in MRCI calculations is enormous – they are usually undertaken only for small systems. Typically, MRCI is useful to study a large section of PES, where significant changes in bonding occur and a sophisticated method is needed to accurately predict dynamical and non-dynamical correlation energy. Usually is truncated to singles and doubles – MRCISD.



## Perturbation Theory

When we try to solve an eigenvalue equation (like the Schrödinger equation), it is sometimes worthwhile to create a simplified (more tractable) operator by removing some particularly unpleasant portion of the original operator. Using exact eigenfunctions and eigenvalues of the simplified operator, it is possible to estimate the eigenfunctions and eigenvalues of the more complete operator – perturbation theory is based on this idea. In general case, we have some operator  $\mathbf{A}$  that we can write as

$$\mathbf{A} = \mathbf{A}^{(0)} + \lambda \mathbf{V}$$

$\mathbf{A}^{(0)}$  – an operator for which we can find eigenfunctions

$\mathbf{V}$  – a perturbing operator

$\lambda$  - a dimensionless parameter (between 0 and 1), maps  $\mathbf{A}^{(0)}$  to  $\mathbf{A}$

If we expand our ground-state eigenfunctions and eigenvalues as Taylor series in  $\lambda$ , we have

$$\Psi_0 = \Psi_0^{(0)} + \lambda \left. \frac{\partial \Psi_0^{(0)}}{\partial \lambda} \right|_{\lambda=0} + \frac{1}{2!} \lambda^2 \left. \frac{\partial^2 \Psi_0^{(0)}}{\partial \lambda^2} \right|_{\lambda=0} + \frac{1}{3!} \lambda^3 \left. \frac{\partial^3 \Psi_0^{(0)}}{\partial \lambda^3} \right|_{\lambda=0} + \dots$$
$$a_0 = a_0^{(0)} + \lambda \left. \frac{\partial a_0^{(0)}}{\partial \lambda} \right|_{\lambda=0} + \frac{1}{2!} \lambda^2 \left. \frac{\partial^2 a_0^{(0)}}{\partial \lambda^2} \right|_{\lambda=0} + \frac{1}{3!} \lambda^3 \left. \frac{\partial^3 a_0^{(0)}}{\partial \lambda^3} \right|_{\lambda=0} + \dots$$

$a_0^{(0)}$  is the eigenvalue for  $\Psi_0^{(0)}$  - the appropriate normalized ground-state eigenfunction for  $\mathbf{A}^{(0)}$ .

We can write the equations simpler:

$$\Psi_0 = \Psi_0^{(0)} + \lambda \Psi_0^{(1)} + \lambda^2 \Psi_0^{(2)} + \lambda^3 \Psi_0^{(3)} + \dots \quad a_0 = a_0^{(0)} + \lambda a_0^{(1)} + \lambda^2 a_0^{(2)} + \lambda^3 a_0^{(3)} + \dots$$

The terms with superscripts ( $n$ ) are called  $n$ th-order corrections.

$$\begin{aligned} (A^{(0)} + \lambda V) |\Psi_0\rangle &= a |\Psi_0\rangle \\ (A^{(0)} + \lambda V) |\Psi_0^{(0)} + \lambda \Psi_0^{(1)} + \lambda^2 \Psi_0^{(2)} + \lambda^3 \Psi_0^{(3)} + \dots\rangle &= \\ (a_0^{(0)} + \lambda a_0^{(1)} + \lambda^2 a_0^{(2)} + \lambda^3 a_0^{(3)} + \dots) |\Psi_0^{(0)} + \lambda \Psi_0^{(1)} + \lambda^2 \Psi_0^{(2)} + \lambda^3 \Psi_0^{(3)} + \dots\rangle \end{aligned}$$

This equation is valid for any choice of  $\lambda$  between 0 and 1.

Therefore, it can be expanded as

$$\begin{aligned} A^{(0)} |\Psi_0^{(0)}\rangle &= a_0^{(0)} |\Psi_0^{(0)}\rangle \\ A^{(0)} |\Psi_0^{(1)}\rangle + V |\Psi_0^{(0)}\rangle &= a_0^{(0)} |\Psi_0^{(1)}\rangle + a_0^{(1)} |\Psi_0^{(0)}\rangle \\ A^{(0)} |\Psi_0^{(2)}\rangle + V |\Psi_0^{(1)}\rangle &= a_0^{(0)} |\Psi_0^{(2)}\rangle + a_0^{(1)} |\Psi_0^{(1)}\rangle + a_0^{(2)} |\Psi_0^{(0)}\rangle \\ A^{(0)} |\Psi_0^{(3)}\rangle + V |\Psi_0^{(2)}\rangle &= a_0^{(0)} |\Psi_0^{(3)}\rangle + a_0^{(1)} |\Psi_0^{(2)}\rangle + a_0^{(2)} |\Psi_0^{(1)}\rangle + a_0^{(3)} |\Psi_0^{(0)}\rangle, \quad \text{etc.} \end{aligned}$$

Our goal – to determine the various  $n$ th order corrections. We can eventually do that expression the eigenfunction corrections as a linear combination of the complete set of eigenfunctions of the initial simplified operator  $\mathbf{A}^{(0)}$ :

$$\Psi_0^{(1)} = \sum_{i>0} c_i \Psi_i^{(0)}$$

Then, we can find

$$a_0^{(1)} = \langle \Psi_0^{(0)} | V | \Psi_0^{(0)} \rangle$$

$$c_j = \frac{\langle \Psi_j^{(0)} | V | \Psi_0^{(0)} \rangle}{a_0^{(0)} - a_j^{(0)}}$$

When the first-order eigenvalue and wave function corrections are in hand, we can carry out similar operations to determine the second-order corrections, then the third-order, etc.:

$$a_0^{(2)} = \sum_{j>0} \frac{\left| \langle \Psi_j^{(0)} | V | \Psi_j^{(0)} \rangle \right|^2}{a_0^{(0)} - a_j^{(0)}}$$

### Single-reference (Møller-Plesset) perturbation theory

We consider the use of perturbation theory where the complete operator  $\mathbf{A}$  is the Hamiltonian  $\mathbf{H}$ . Møller and Plesset proposed choices for  $\mathbf{A}^{(0)}$  and  $\mathbf{V}$  – their theory is referred to as MP $n$ , where  $n$  is the order at which the theory is truncated: MP2, ..

In the MP approach,  $\mathbf{H}^{(0)}$  is taken as the sum of the one-electron Fock operators:

$$H^{(0)} = \sum_{i=1}^n f_i$$

$\Psi^{(0)}$  – the HF wave function – a Slater determinant formed from the occupied orbitals.

One can show that the eigenvalue of  $\mathbf{H}^{(0)}$  is the sum of the occupied orbital energies:

$$H^{(0)} \Psi^{(0)} = \sum_i^{occ.} \varepsilon_i \Psi^{(0)}$$

The correction term  $\mathbf{V}$  that will return us to the correct Hamiltonian and allow us to use perturbation theory to improve the HF wave function and eigenvalues is then

$$V = \sum_i^{occ.} \sum_{j>i}^{occ.} \frac{1}{r_{ij}} - \sum_i^{occ.} \sum_{j>i}^{occ.} \left( J_{ij} - \frac{1}{2} K_{ij} \right)$$

The first term is the proper way to compute electron repulsion and the second term is how it is computed from summing over the Fock operators for the occupied orbitals where  $J$  and  $K$  are the Coulomb and exchange operators.

Now we calculate the first-order correction  $a^{(1)}$ :

$$a^{(0)} + a^{(1)} = \langle \Psi^{(0)} | H^{(0)} | \Psi^{(0)} \rangle + \langle \Psi^{(0)} | V | \Psi^{(0)} \rangle = \langle \Psi^{(0)} | H^{(0)} + V | \Psi^{(0)} \rangle = \langle \Psi^{(0)} | H | \Psi^{(0)} \rangle = E_{HF}$$

This is because Hartree-Fock energy is actually calculated not as an eigenvalue of the HF Hamiltonian but by applying the correct Hamiltonian (the expectation value) to the HF wave function. Thus, the Hartree-Fock energy is the energy correct through first-order in Møller-Plesset perturbation theory – MP1 does not advance us beyond the HF level in determining the energy. The next step is the second-order correction.

The full expression for the second-order energy correction is

$$a^{(2)} = \sum_i^{occ.} \sum_{j>i}^{occ.} \sum_a^{vir.} \sum_{b>a}^{vir.} \frac{[(ij | ab) - (ia | jb)]^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}$$

The sum of  $a^{(0)}$ ,  $a^{(1)}$ , and  $a^{(2)}$  defines the MP2 energy.

MP2 calculations can be done reasonably fast because the above equation can be efficiently evaluated. The scaling behavior of the MP2 method is  $\sim N^5$ , where  $N$  is the number of basis functions. Analytic gradients and second derivatives are available for MP2 – it can conveniently be used to explore PESs. MP2 and all orders of MP $n$  are size-consistent – a desirable feature.

However, the Møller-Plesset formalism is potentially dangerous in design. Perturbation theory works best when the perturbation is small because the Taylor expansions are then expected to be quickly convergent. But, in the case of MP theory, the perturbation is the full electron-electron repulsion energy, which is a rather large contributor to the total energy – there is no reason to expect that an MP2 calculation will give a good value for the correlation energy. Also, the MP $n$  methodology is not variational – it is possible that the MP2 estimate for the correlation energy will be too large instead of too small.

If one wants to improve convergence, one can proceed to higher orders in perturbation theory. At third order, it is still true that only matrix elements involving doubly excited determinants need to be evaluated – MP3 is not too much more expensive than MP2. However, MP3 calculations tend to offer rather little improvement over MP2. Analytic gradients are not available for third and higher orders of perturbation theory. At the MP4 level, integrals involving triply quadruply excited determinants appear. The evaluation of the terms involving triples is the most costly – scales as  $N^7$ .

If one simply chooses to ignore the triples, the method scales more favorably – MP4SDQ. Usually, the cost of accounting for the triples is roughly equal to that for the rest of the calculation – triples double the time. In closed-shell systems with large frontier orbital separations, the contributions from the triples tend to be rather small – ignoring them may be worthwhile in terms of efficiency. However, when the frontier orbital separation drops, the contribution of the triples can become very large, and major errors in interpretation can arise from ignoring their effect.

Empirically, MP4 calculations can be quite good – typically account for more than 95% of the correlation energy with a good basis set. However, a typical behavior of  $MP_n$  results is oscillatory and it is difficult to extrapolate a convergent result from  $MP_1 = HF$ ,  $MP_2$ ,  $MP_3$ , and  $MP_4$  values.

## Multireference perturbation theory

The generalization of  $MP_n$  theory to the multireference case involves the obvious choice of using an MCSCF wave function for  $\Psi^{(0)}$  instead of single-determinant HF wave function – the CASPT2 method. Most multireference methods described to date have been limited to second order in perturbation theory. Analytic gradients are available for CASPT2. While some third order results have begun to appear, similar to the single-reference case, they do not offer much improvement over the second order.

A good feature of multireference perturbation theory – it can correct for the deficiencies due to the use of an incomplete active space. For example, the relative energies for various electronic states of TMM were found to vary widely depending on whether a (2,2), (4,4), or (10,10) active space was used, however, relative energies from corresponding CASPT2 calculations agreed well with one another.

## Coupled-cluster theory

One of the more mathematically elegant techniques for estimating the electron correlation – coupled-cluster (CC) theory (Cizek 1966). The central idea of CC theory – the full-CI wave function (the ‘exact’ one within the basis set approximation) can be described as

$$\Psi = e^{\mathbf{T}} \Psi_{HF}$$

The cluster operator  $\mathbf{T}$  is defined as

$$\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3 + \dots + \mathbf{T}_n$$

where  $n$  is the total number of electrons and the various  $\mathbf{T}_i$  operators generate all possible determinants having  $i$  excitations from the reference.

For example,

$$T_2 = \sum_{i < j} \sum_{a < b}^{occ. \text{ vir.}} t_{ij}^{ab} \Psi_{ij}^{ab}$$

where the amplitudes  $t$  are determined by the constraint that the  $\Psi = e^T \Psi_{HF}$  equation be satisfied. The expansion of  $\mathbf{T}$  ends at  $n$  because no more than  $n$  excitations are possible.

When the expansions are not truncated, both CC and CI approaches end up in full CI. What is then the advantage of CC? The answer lies in the consequences associated with truncation of  $\mathbf{T}$ . Let's make the approximation  $\mathbf{T} = \mathbf{T}_2$ . In that case, Taylor expansion of the exponential function gives

$$\Psi_{CCD} = e^T \Psi_{HF} = (1 + T_2 + \frac{T_2^2}{2!} + \frac{T_2^3}{3!} + \dots) \Psi_{HF}$$

CCD implies coupled cluster with only the double-excitation operator. The first two terms in parentheses,  $1 + \mathbf{T}_2$ , define the CID method. The remaining terms involve products of excitation operators. Each generation of  $\mathbf{T}_2$  generates double excitations, so the product of two applications (the square of  $\mathbf{T}_2$ ) generates quadruple excitations.

Similarly, the cube of  $\mathbf{T}_2$  generates hextuple substitutions, etc. The failure to include these excitations makes CI non-size-consistent. Using the exponential of  $\mathbf{T}$  ensures size consistency.

The computational problem is determination of cluster amplitudes  $t$  for all of the operators included in the particular approximation. When the amplitudes are obtained, the coupled-cluster energy is computed as

$$\langle \Psi_{HF} | H | e^T \Psi_{HF} \rangle = E_{CC}$$

In practice, the cost of including single excitation ( $\mathbf{T}_1$ ) in addition to doubles is worth the increase in accuracy – the CCSD model. The scaling behavior of CCSD –  $N^6$ . Inclusion of  $\mathbf{T}_3$  – CCSDT, but this is very computationally costly ( $N^8$ ). One can instead use perturbation theory to evaluate the contribution of triples – this method is called CCSD(T). The **CCSD(T)** level has become the **gold-standard** level of theory for single-reference calculations. (Note that coupled cluster theory is not, however, variational.)

The related method – quadratic configuration interaction including singles and doubles (QCISD). It was originally developed by Pople and co-workers (1987) to correct for size-consistency errors in CISD, but was later shown to be almost equivalent to CCSD. The QCISD(T) method includes the same perturbative correction for triples as that used in CCSD(T). Typically, CCSD and QCISD give close results and the same is true for CCSD(T) and QCISD(T). But, in certain challenging systems the more complete coupled-cluster methods have been found to be more robust. Unless, a wave function exhibits a strong multireference character, CCSD(T) – the best choice. A measure of the multireference

character  $T_1$  diagnostic should be smaller than 0.02

$$T_1 = \sqrt{\frac{\sum_i \sum_a^{occ. vir.} (t_i^a)^2}{n}}$$



## Practical Issues in Application

### Basis set convergence

Basis set flexibility is crucial to accurately describing the molecular wave function. This becomes more true for methods including electron correlation. The correlated wave function is a linear combination of determinants – excited determinants include occupation of orbitals that are virtual in the HF determinant – the excited states are more dependent on basis-set completeness. This prompted the developers of basis sets to optimize basis sets at the correlated levels – cc-pVnZ basis sets.

### Sensitivity to reference wave function

When a non-trivial degree of multireference character exists, perturbation theory is particularly sensitive to this feature and can give untrustworthy results. TMM example: we take the single-configuration wave function and consider the MP2 energy from the double excitation taking both electrons from occupied orbital  $\pi_2$  to virtual orbital  $\pi_3$ . The MP2

energy correction is

$$a^{(2)} = \sum_i^{occ.} \sum_{j>i}^{occ.} \sum_a^{vir.} \sum_{b>a}^{vir.} \frac{[(ij|ab) - (ia|jb)]^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}$$

Since the  $\pi_2$  and  $\pi_3$  orbitals are formally degenerate, the denominator is zero – the perturbation energy expression for the energy associated with this term is infinite!

As a general rule, whenever the frontier orbital separation becomes small, the magnitude of the MP $n$  energy terms will become large because of their inverse dependence on orbital energy separation – perturbation theory will be very slowly convergent. CCSD is similarly sensitive to multireference character. Inclusion of triples in the CCSD wave function is usually very effective in correcting for a single-reference treatment of a weakly to moderately multireference problem.

The most common way to include the triples is by perturbation theory, CCSD(T), and this level too can be unstable if singles amplitudes are large. In such a case, so-called Brueckner doubles BD(T) calculations, which eliminate the singles amplitudes by orbital localization, can perform better.

### **Price/Performance Summary**

For a typical equilibrium structure, the HF level of theory predicts bond lengths that are usually a little too short. As a rule, the MP2 level is an excellent choice for geometry optimizations of minima that include correlation energy – significant improvements can be obtained at fairly reasonable cost. Examination of bond lengths in 108 molecules containing 2 to 8 atoms: at the MP2/6-31G(d,p) level the average error was 0.015 Å – improvement as compared to the HF level (0.021 Å). If one observes a large change in some geometric property on going from the HF to the MP2 level, it is worthwhile to investigate the predictions from still higher levels of theory – clearly the perturbations are large and it is quite possible that MP2 does not provide convergence in this particular property.

Energetics: MP2 is considered as a very efficient level of theory for energy differences between minima. Often, the error in such differences is reduced by 25-50% from HF to MP2. Example – 45 isomerization energies – on going from HF/6-31G(d) to MP2/6-31G(d) the average error was reduced from 2.9 to 1.9 kcal/mol. For 11 glucose conformers, the average error in conformational energy is reduced from 0.6 to 0.4 kcal/mol. However, the structures considered in these examples are already fairly well described by HF wave functions. A rough quality ordering of methods: **HF < MP2 ~ MP3 ~ CCD < CISD < MP4SDQ ~ QCISD ~ CCSD < MP4 < QCISD(T) ~ CCSD(T) ~ BD(T)**

Average errors in correlation energies (kcal/mol) compared to full CI for various methods applied to HB, H<sub>2</sub>O, and HF at equilibrium and bond-stretched geometries

Level of theory	Equilibrium	Equilibrium & stretched
MP2	10.4	17.4
MP3	5.0	14.4
CISD	5.8	13.8
CCD	2.4	8.0
MP4SDQ	2.7	7.1
CCSD	1.9	4.5
QCISD	1.7	4.0
MP4	1.3	3.7
MP5	0.8	3.2
MP6	0.3	0.9
CCSD(T)	0.3	0.6
QCISD(T)	0.3	0.5
CCSDT	0.2	0.5
CCSDTQ	0.01	0.02

## Formal scaling behavior, as a function of basis functions $N$ , of various electronic structure methods

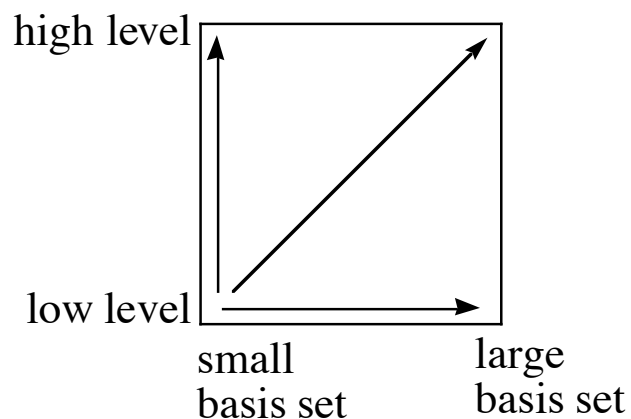
Scaling behavior	Method(s)
$N^4$	HF
$N^5$	MP2
$N^6$	MP3, CISD, MP4SDQ, CCSD, QCISD
$N^7$	MP4, CCSD(T), QCISD(T)
$N^8$	MP5, CISDT, CCSDT
$N^9$	MP6
$N^{10}$	MP7, CISDTQ, CCSDTQ

### **Parametrized methods**

Wouldn't be nice to get the right answer for any problem in general? Although methods like full CI and CCSDTQ, when used in conjunction with large and flexible basis sets, are very accurate as solutions of the Schrödinger equation, they simply cannot be applied to more than the smallest fraction of chemically interesting systems because of their computational expense. And, with scaling behavior on the order of  $N^{10}$ , this situation is unlikely to change anytime soon – practitioners of ab initio theory have returned to the idea of introducing parameters to improve predictive accuracy.

## Multilevel methods

Pople's 'model chemistries: Gaussian 2 (G2) and Gaussian 3 (G3) theories. The main idea – the additivity of basis set corrections at different levels of theory.



Carry out high level calculations (QCISD(T)) with a moderate basis set, then add corrections for the basis set expansions at low levels, MP2 and MP4. Add an empirical correction, which depends on the number of broken electron pairs in the reaction. Empirical correction has only one parameter chosen to reproduce experimental heats of formation for a test set of molecules.

## Steps in G2 and G3 theories for molecules

G2	G3
(1) HF/6-31G(d) geometry optimization	HF/6-31G(d) geometry optimization
(2) ZPE from HF/6-31G(d) frequencies	ZPE from HF/6-31G(d) frequencies
(3) MP2(full)/6-31G(d) geometry optimization	MP2(full)/6-31G(d) geometry optimization
All subsequent calculations use this geometry	
(4) $E[\text{MP4}/6\text{-}311\text{+G(d,p)}] - E[\text{MP4}/6\text{-}311\text{G(d,p)}]$	$E[\text{MP4}/6\text{-}31\text{+G(d,p)}] - E[\text{MP4}/6\text{-}31\text{G(d)}]$
(5) $E[\text{MP4}/6\text{-}311\text{G(2df,p)}] - E[\text{MP4}/6\text{-}311\text{G(d,p)}]$	$E[\text{MP4}/6\text{-}31\text{G(2df,p)}] - E[\text{MP4}/6\text{-}31\text{G(d)}]$
(6) $E[\text{QCISD(T)}/6\text{-}311\text{G(d)}] - E[\text{MP4}/6\text{-}311\text{G(d)}]$	$E[\text{QCISD(T)}/6\text{-}31\text{G(d)}] - E[\text{MP4}/6\text{-}31\text{G(d)}]$
(7) $E[\text{MP2}/6\text{-}311\text{+G(3df,2p)}] - E[\text{MP2}/6\text{-}311\text{G(2df,p)}] - E[\text{MP2}/6\text{-}311\text{+G(d,p)}] + E[\text{MP2}/6\text{-}311\text{G(d,p)}]$	$E[\text{MP2(full)}/\text{G3large}] - E[\text{MP2}/6\text{-}31\text{G(2df,p)}] - E[\text{MP2}/6\text{-}31\text{+G(d,p)}] + E[\text{MP2}/6\text{-}31\text{G(d,p)}]$
(8) $-0.00481n_\beta - 0.00019n_\alpha$	$-0.006386n_\beta - 0.002977n_\alpha$
‘Higher-level correction’ (HLC) – depends on the number of valence electron pairs ( $n_\beta$ ) and the number of unpaired valence electrons ( $n_\alpha$ )	
$E_0 = E[\text{MP4}/6\text{-}311\text{G(d,p)}] + (4)$	$E[\text{MP4}/6\text{-}31\text{G(d)}] + (4)$
(5) + (6) + (7) + (8) + 0.8929×ZPE	(5) + (6) + (7) + (8) + 0.8929×ZPE

The goal of most multilevel models – absolute energy prediction. Single-level models can still be useful both qualitatively and quantitatively for relative energy differences because of cancellation of errors.

Modified G2 theory (**G2M**; Mebel, Morokuma, Lin – 1995):

- 1) Geometry optimization and frequency calculations – density functional B3LYP/6-311G\*\*
- 2) QCISD(T) has been replaced with generally more robust CCSD(T)

Similar modifications of G3 are also available – G3(CC)//B3LYP

The latest ‘model chemistry’ – **G4** (2007)

- 1) Geometry optimization and vibrational frequencies – B3LYP/6-31G(2df,p)
- 2) HF limit calculations:  $E_{\text{HF/aug-cc-pVnZ}} = E_{\text{HF/limit}} + B \exp(-\alpha n)$
- 3)  $E(\text{G4}) = E(\text{MP4/6-31G(d)}) + \Delta E(+)$  +  $\Delta E(2\text{df,p})$  +  $\Delta E(\text{CC})$  +  $\Delta E(\text{G3LargeXP})$  +  $\Delta E(\text{HF})$  +  $\Delta E(\text{SO})$  +  $\Delta E(\text{HLC})$

$$\Delta E(\text{HF}) = E(\text{HF/limit}) - E(\text{HF/G3LargeXP})$$

$\Delta E(\text{SO})$  – spin-orbit correction; taken into account only for atoms based on experiment.

CCSD(T) extrapolation to the complete basis set limit – **CCSD(T)/CBS**

$$E_{\text{CCSD(T)/cc-pVnZ}} = E_{\text{CCSD(T)/CBS}} + B \exp(-\alpha n)$$

Cheaper extrapolation schemes based only on smaller basis set calculations:

$$E(\text{CBS}) = (E(\text{VTZ}) - E(\text{VDZ}) * 2.5^3 / 3.5^3) / (1 - 2.5^3 / 3.5^3) \quad \text{D-T extrapolation}$$

$$E(\text{CBS}) = (E(\text{VQZ}) - E(\text{VTZ}) * 2.75^3 / 3.75^3) / (1 - 2.75^3 / 3.75^3) \quad \text{T-Q extrapolation}$$