Multiconfigurational and multireference methods

Peter R. Taylor and Jeppe Olsen

September, 2011

Contents of the three lectures

Wednesday, Sept. 21
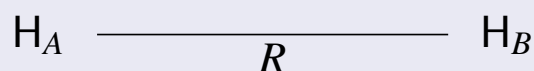
Thursday, Sept. 22

Friday, Sept. 23

- Introduction to dynamic correlation
- The multi reference CI method
- A bit on Multireference CC
- Other nearly degenerate multi reference methods
- Perturbation theory, general, MP2 and multireference methods
- CASPT2

Molecular Orbitals for H$_2$
RHF

The molecule

$$H_A \underline{\hspace{4cm}}_{R} \quad H_B$$

Groundstate in a minimal basis: $(1s_A, 1s_B)$

- MOs given by symmetry:

$$\sigma_g = N_g(1s_A + 1s_B) \qquad \sigma_u = N_u(1s_A - 1s_B)$$

with normalization constants N_g and N_u.

- Closed-shell HF configuration σ_g^2, "restricted HF" (RHF) wave function

$$\Phi_1 = |\sigma_g\alpha\sigma_g\beta| \quad (a_{\sigma_g\alpha}^+ a_{\sigma_g\beta}^+ |vac\rangle)$$

Molecular Orbitals for H$_2$
RHF

At equilibrium

Φ_1 is a good approximation to exact wave function

$R \to \infty$

- Correct/exact wave function:
$\Phi_{cov} = \frac{1}{\sqrt{2}}(|1s_A\alpha 1s_B\beta| + |1s_B\alpha 1s_A\beta|)$

- $\sigma_g \to \frac{1}{\sqrt{2}}(1s_A + 1s_B)$ and thus

$$\Phi_1 \to \frac{1}{2}(|1s_A\alpha 1s_B\beta| + |1s_B\alpha 1s_A\beta| $$
$$+ |1s_A\alpha 1s_A\beta| + |1s_B\alpha 1s_B\beta|),$$

an unphysical 50:50 mixture of "covalent" and "ionic" terms.

Molecular Orbitals for H$_2$

RHF

Why can two separated H-atoms not be described with RHF?

- σ_g has the same weights on both atoms $\to$ 50 % chance that an electron is at a given atom
- Both electrons are in σ_g $\to$ 25 % chance of finding both electrons at a given atom, 50 % chance of fining the electrons at different atoms
- So the problem depends on the Restricted Hartree-Fock wave function *per se*

Improved behavior may be obtained using UHF

Molecular Orbitals for H$_2$

UHF

The (first) wave function of Coulson and Fischer

- Write a new determinant Φ_λ in terms of new spin-orbitals $\psi_1 = N(1s_A + \lambda 1s_B)$ and $\psi_2 = N(\lambda 1s_A + 1s_B)$ as

$$\Phi_\lambda = |\psi_1 \overline{\psi_2}|.$$

- Optimize energy with respect to λ at each value of R.
- Near R_e $\lambda \approx 1$, whereas as $R \to \infty$, $\lambda \to 0$.
- Φ_λ for optimum λ is an example of an unrestricted Hartree-Fock (UHF) wave function.
- The case $\lambda = 1$ gives a stationary energy for any R, but this is not necessarily a minimum.

Molecular Orbitals for H$_2$
UHF

The (first) wave function of Coulson and Fischer, advantages

- Wave function and energy behave as expected at dissociation: dissociates to two H atoms
- Wave function becomes RHF wave function around equilibrium → good approximation here

The (first) wave function of Coulson and Fischer, disadvantages

- Φ_λ does not display inversion symmetry → symmetry-broken approximation
- Φ_λ is an eigenfunction of $\hat{S}_z$ but not of $\hat{S}^2$ → spin-broken approximation
- Gives problems in region where bond is broken

Molecular Orbitals for H$_2$
UHF

Rewrite of $\Phi_{UHF} = \Phi_\lambda$

$$\Phi_{UHF} =$$
$$\underbrace{c_1^2 |\sigma_g \overline{\sigma}_g| - c_2^2 |\sigma_u \overline{\sigma}_u|}_{\text{gerade sym, singlet}} - \underbrace{c_1 c_2 (|\sigma_g \overline{\sigma}_u| - |\sigma_u \overline{\sigma}_g|)}_{\text{ungerade sym, triplet}}$$

- $c_1 = N(1+\lambda)/2N_g$ and $c_2 = N(1-\lambda)/2N_u$
- Unless $\lambda = \pm 1$ there is both spin-and symmetry-breaking
- Is correct at dissociation because all states in Φ_{UHF} are degenerate in this limit.

Molecular Orbitals for H_2

Arguments for a two-configuration wave function

- The exact wave function at $R = \infty$ is

$$\Phi_{diss} = \frac{1}{\sqrt{2}}(|\sigma_g\overline{\sigma}_g| - |\sigma_u\overline{\sigma}_u|).$$

- Why not generalize this for any R to

$$\Phi_0(R) = c_g(R)|\sigma_g\overline{\sigma}_g| + c_u(R)|\sigma_u\overline{\sigma}_u|)$$

and optimize energy with respect to $c_g(R)$ and $c_u(R)$?
- Two-configuration wave function: spin eigenfunction, symmetry-adapted.
- More parameters to optimize than in UHF!

Contents

Wednesday, Sept. 21

- Molecular orbitals for H_2 at equilibrium and dissociation using RHF and UHF
- The symmetry problem
- Introduction to the MCSCF method
- The CASSCF method and its extensions

The symmetry problem: to break or not to break

Symmetry?

- By symmetry we mean any operation that leaves the Hamiltonian invariant, e.g., spin (if $\hat{H}$ is spin-free), spatial symmetry, permutations of identical particles. . .
- Hamiltonian $\hat{H}$ invariant to symmetry-operator $\hat{O}$
  $\rightarrow \hat{H} = \hat{O}\hat{H}\hat{O}^{-1} \rightarrow \hat{O}\hat{H} = \hat{H}\hat{O}$
- A symmetry operator therefore commutes with the Hamiltonian

The symmetry problem: to break or not to break
Exact wave functions

Symmetry for exact wave functions

- Operators that commute have a set of common eigenfunctions
- The solutions to Schrödinger equation may therefore be choosen as eigenfunctions for the symmetry-operators

Degenerate eigenvalues gives problems

- The fact that two operators A, B commute does not ensure that any eigenfunction of A is also a eigenfunction of B
- Example: $|\sigma_g \alpha \sigma_u \beta|$ is an eigenfunction of $\hat{S}_z$ (eigenvalue 0)
- $[\hat{S}_Z, \hat{H}] = 0$, but $|\sigma_g \alpha \sigma_u \beta|$ is not eigenfunction of $\hat{H}$
- Instead: We can make linear combinations of the eigenfunctions of $\hat{S}_Z$ with eigenvalue 0 to obtain eigenfunction of $\hat{H}$

The symmetry problem: to break or not to break
Approximate wave functions

Linear expansions (CI)

- If the space is closed under the action of symmetry-operator, then there is a common set of eigen functions for the symmetry-operator and the projected Hamiltonian
- Problems again with degenerate eigenvalues

Non linear expansions (SCF, MCSCF, CC)

- Only way to ensure correct symmetry is to ensure that only the wave-function only is varied over space of the correct symmetry
- Examples: Ensure orbitals have well defined symmetry, that the total spin is restricted to the requested spin

The symmetry problem: to break or not to break
Symmetry and HF

Do worry-restrict

- It may be argued that the symmetry-properties are important, so keep them
- And once a symmetry property is lost, it may be hard to regain it
- Requires to explicit work with function that are symmetry-adapted- typically requiring several Slater determinants
- Makes thus complications in the form of the wave function

The symmetry problem: to break or not to break
Symmetry and HF

Do not worry- unrestricted

- Argument: We are making so many approximations, so why worry
- (Life is too short for symmetry adaptation)
- As we improve the level of say correlation, we will approach the correct symmetry
- Gives very simple starting point- typical single determinant

The symmetry problem: to break or not to break
Symmetry and HF

Third choice: Do UFH and clean up

- Hence the terms *restricted* and *unrestricted*.
- We could remove the "contaminating" terms in UHF, recovering the symmetries. The contaminants can only raise the energy of the lowest state, so this removal, by projection, say, gives an energy that cannot be worse. *Projected* UHF.
- This PUHF wave function is however no longer optimum: re-optimize and further improve energy. *Extended* HF.

The symmetry problem: to break or not to break
UHF in practice

- A molecular UHF wave function dissociates to atomic UHF wave functions.

- None of these UHF wave functions are in general spin or symmetry eigenfunctions.

- No (single configuration) Hartree-Fock method can be used straightforwardly on other than the lowest state of a given spin or symmetry, so optimizing excited states, like singlet diradicals, is a problem

The symmetry problem: to break or not to break
UHF in practice

- Response methods can sometimes be used but these have their own issues.

- For some systems (N_2 near r_e) the lowest energy solution to the UHF equations is actually the RHF wave function: at some critical distance there is a bifurcation in the potential curve where the broken-symmetry UHF solution falls below the RHF solution!

The symmetry problem: to break or not to break
UHF in practice

- UHF is cheap and straightforward. Deficiencies are easy to monitor (look at $\langle S^2 \rangle$).
- Double the number of parameters of RHF (not a big problem but it greatly increases the work later in correlated calculations built on UHF).
- Loss of spin and spatial symmetries (corrected as correlation is introduced).
- Cannot do low spin biradicals etc
- Cannot do general excited states
- Possible bifurcations and non-analytic behaviour of potential curves and surfaces.

Contents

Wednesday, Sept. 21

- Molecular orbitals for H$_2$ at equilibrium and dissociation using RHF and UHF
- The symmetry problem
- Introduction to the MCSCF method
- The CASSCF method and its extensions

MCSCF
Introduction

The Multi configurational Self Consistent Field Method

- Introduce a small number of orbitals (active orbitals), say 10 -20, where the occupation are allowed to vary.
- The active orbitals are those which (for some geometry) has occupation numbers significantly larger than 0 and smaller than two
- Select the configurations that will be included.
- Form of wave function: $\left|\tilde{0}\right\rangle = \sum_I C_I \left|\tilde{I}\right\rangle$ where $\left|I\right\rangle$ are Slater determinants (ONV's) or configuration state functions
- Optimize the orbitals and the CI coefficients C_I

MCSCF
Introduction

What can we expect from MCSCF?

- Obviously describing *electron correlation* effects.
- But not the short-range problems that arise as $r_{12} \to 0$: the so-called *dynamical correlation*.
- (Because convergence of dynamical correlation in an orbital-based expansion is very slow.)
- Will describe the *nondynamical* (or "static") correlation that comes from configurational near-degeneracies or from deficiencies in the Hartree-Fock orbitals.

MCSCF
Introduction

The configuration spaces

- In the early years of MCSCF (up to about 1980) this was a big problem.

- Ad hoc choices of configurations (very useful when you already know the answer...)

- More systematic approaches, e.g., all configurations required for proper dissociation. Easy for diatomics, but dissociation to what in a polyatomic?

- In spectroscopy, the configuration(s) configurations could include those for a simple orbital model of excitation (say, $n \rightarrow \pi^*$), but can easily miss important configurations this way.

MCSCF
Introduction

A choice of the configuration spaces: CASSCF

- First simplifying assumption: we will choose a subspace of the MOs, the *active space*, from which all configurations will be built.

- Then include *all* configurations generated by allocating electrons to these orbitals: a *full CI* in the active space.

- We still have to choose the active orbitals, but then configuration generation is automatic.

- A nice naive approach would be the valence AOs for all atoms, but this quickly gets very large...

- MCSCF optimization using this sort of *complete active space* (CAS) — a CASSCF calculation — is 'easy'

Contents

Wednesday, Sept. 21

- Molecular orbitals for H$_2$ at equilibrium and dissociation using RHF and UHF
- The symmetry problem
- Introduction to the MCSCF method
- The CASSCF method and its extensions

The CASSCF method

Is simple to define

- Number of orbitals (per symmetry) and number of electrons in the active space
- (There will in general also be a number of double occupied orbitals: inactive orbitals)

The active orbital space

- Should include all orbitals where the occupation number changes significantly during a process (like a reaction, excitation, ionization), or where the occupation number differs significantly from two or zero.

The CASSCF method
The Orbital Spaces for CAS Wave Functions

Inactive, active, and secondary orbitals

Unoccupied Orbitals

Active Orbitals

Inactive Orbitals
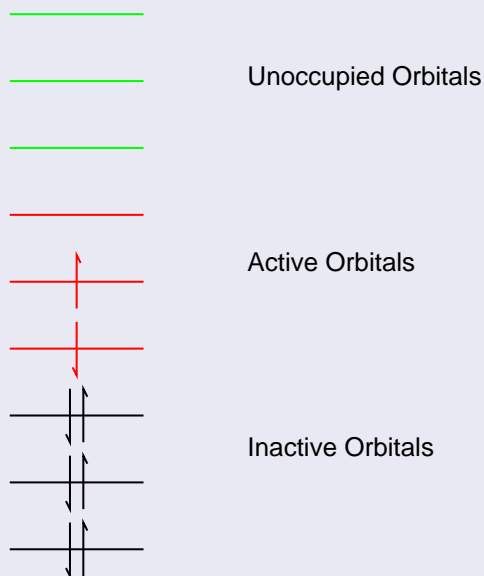
The CASSCF method
The size of the configuration expansion

Number of Slater determinants (aka ONV's) for 2k electrons in 2k orbitals

2k	# SD's $= \left(\begin{array}{c} 2k \\ k \end{array} \right)^2$
2	4
4	36
6	400
8	4.900
10	63.504
12	853.776
14	11.778.896
16	165.636.896
18	2.363.904.260

The CASSCF method
The size of the configuration expansion

Scaling considerations

- \# SD's: $\begin{pmatrix} 2k \\ k \end{pmatrix}^2$

- For large k Stirlings approximation gives: \# SD's $= \frac{16^k}{k\pi}$

- Adding two electrons and orbitals increase \# SD's with a factor of 16

- Corresponds to 6 years of computer developments

- Adding one Cr atom (6 electrons in 6 orbitals to active space) requires thus 18 years of computer development

The conceptual simplicity of the CASSCF method comes at a price

The CASSCF method
Choosing active spaces

Simple diatomics

- Sometimes the problem seems and is simple: H$_2$ ground-state potential curve will need $(\sigma_g, \sigma_u)^2$, or the two $1s$ orbitals, as the active space.

- Bigger diatomics seem similarly easy: N$_2$ ground state will need $(\sigma_g, \pi_u, \pi_g, \sigma_u)^6$ or the $2p$ orbitals.

- C$_2$ will need the $2s$ orbitals as well, because of the $2s/2p$ near-degeneracy in C atom.

- F$_2$: neither $(\sigma_g, \sigma_u)^2$ (the bonding/antibonding pair) nor $(\sigma_g, \pi_u, \pi_g, \sigma_u)^{10}$ gives other than a qualitative result (bonding...). The $3p$ orbitals are needed (because of F$^-$)!

- Even diatomics are not straightforward.

The CASSCF method
Choosing active spaces

A bit on polyatomic molecules

- In most polyatomics the full valence shell that works for e.g., N$_2$ (or C$_2$, depending on the definition of "valence") will be too large.
- Identify the orbitals involved in the process.
- Spectroscopy of C$_6$H$_6$: use the six π MOs.
- Breasking a CH or CC bond in a hydrocarbon: use σ, σ^*
- May have to refine this choice (we will see how). Clearly very far from a black-box approach!

The CASSCF method
Choosing active spaces

Even larger active spaces

- More than about 15–16 active MOs (and 15 or 16 active electrons) are time-consuming

The RAS construction: Three active orbital spaces

RAS1 orbitals: Max number of holes

RAS2 orbitals: n_{occ} varies

RAS3 orbitals: Max number of electrons

Contains many type of CI expansions, see later

The RASSCF method
RAS Orbital Spaces

Unoccupied Orbitals

RAS 3 Orbitals

RAS 2 Orbitals

RAS 1 Orbitals

Inactive Orbitals

The RASSCF method
Some Typical RAS Wave Functions

- Closed Shell SCF (RAS1, RAS2, RAS3 empty).
- SDTQ...CI with a closed shell reference function (RAS2 empty).
- CASSCF (RAS1 and RAS3 empty).
- Multireference SDCI with a CASSCF reference (max two holes in RAS1 and max two electrons in RAS3).
- "Polarization" CI (max one hole in RAS1, one electron in RAS3).

The GASSCF method
Some further developments of active spaces

The Generalized Active Space

- Allows an arbitrary number of active orbital spaces
- Allows arbitrary types of occupation restrictions
- An american cousin is called ORMAS(occupation restricted multiple active spaces)

It is all right now, in fact it is a GAS
Jagger/Richards

Optimization of MCSCF Wave Functions

The wave function

$$\Psi = \sum_m \Phi_m c_m$$

or

$$\left|0'\right\rangle = \sum_m \left|m'\right\rangle c_m$$

The optimization problem

Determine the MOs and the MC coefficients using the variational principle

$$E = \frac{\langle 0'|\hat{H}|0'\rangle}{\langle 0'|0'\rangle}$$

Optimization of MCSCF Wave Functions
The energy

Non-Relativistic Hamiltonian (Second Quantization)

- Hamiltonian

$$\hat{H} = \sum_{ij} h_{ij}\hat{E}_{ij} + \frac{1}{2}\sum_{ijkl} g_{ijkl}(\hat{E}_{ij}\hat{E}_{kl} - \delta_{jk}\hat{E}_{il})$$

$$h_{ij} = \int \phi_i^*(\mathbf{x})\hat{h}(\mathbf{x})\phi_j(\mathbf{x})d\mathbf{x}$$

$$g_{ijkl} = \int \phi_i^*(\mathbf{x}_1)\phi_j(\mathbf{x}_1)G(\mathbf{x}_1,\mathbf{x}_2)\phi_k^*(\mathbf{x}_2)\phi_l(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2$$

are the one- and two-electron integrals.
- "Excitation operator" $\hat{E}_{ij} = \hat{a}_{i\alpha}^\dagger\hat{a}_{j\alpha} + \hat{a}_{i\beta}^\dagger\hat{a}_{i\beta}$

Optimization of MCSCF Wave Functions
The energy

Contribution from one-electron operator

- One-electron operator: $\hat{h} = \sum_{ij} h_{ij}\hat{E}_{ij}$
- Matrix elements: $\langle m|\hat{h}|n\rangle = \sum_{i,j} h_{ij}\langle m|\hat{E}_{ij}|n\rangle = \sum_{i,j} h_{ij}D_{ij}^{mn}$,
- D_{ij}^{mn} are the *one-electron coupling coefficients*.
- The energy contribution is:

$$\langle 0|\hat{h}|0\rangle = \sum_{ij} h_{ij}D_{ij},$$

- $D_{ij} = \langle 0|\hat{E}_{ij}|0\rangle = \sum_{mn} c_m^* c_n D_{ij}^{mn}$ are elements of the *first order reduced density matrix* or 1-matrix.

Optimization of MCSCF Wave Functions
The energy

Contribution from the two-electron operator

- Two-electron operator $\hat{g} = \frac{1}{2}\sum_{ijkl} g_{ijkl}(\hat{E}_{ij}\hat{E}_{kl} - \delta_{jk}\hat{E}_{il})$
- $\langle m|\,\hat{g}\,|n\rangle = \sum_{ijkl} g_{ijkl}P^{mn}_{ijkl}$
- P^{mn}_{ijkl} are the *two-electron coupling coefficients*.
- The energy contribution is: $\langle 0|\,\hat{g}\,|0\rangle = \sum_{ijkl} g_{ijkl}P_{ijkl}$
- $P_{ijkl} = \sum_{m,n} c^*_m c_n P^{mn}_{ijkl}$ are elements of the *second order reduced density matrix* or 2-matrix.

Optimization of MCSCF Wave Functions
The energy

Summary

- Total MCSCF energy

$$E = \langle 0|\,\hat{H}\,|0\rangle = \sum_{ij} h_{ij}D_{ij} + \sum_{ijkl} g_{ijkl}P_{ijkl} + h_{nuc}.$$

- The molecular orbital coefficients appear in the one- and two-electron integrals h_{ij} and g_{ijkl}.
- The CI coefficients appear in $\mathbf{D}$ and $\mathbf{P}$.
- Energy depends on MOs φ and CI coefficients c

$$E = E\{\varphi, \mathbf{c}\}$$

Optimization of MCSCF Wave Functions
Unitary Transformation of MOs

Exponential parameterization

- MOs are orthonormal (no loss of generality) and we wish to preserve this.

- Consider a unitary transformation of the orbitals $\varphi' = \varphi \mathbf{U}$, where $\mathbf{U}^\dagger \mathbf{U} = \mathbf{1}$.

- Any unitary matrix can be written in the form. $\mathbf{U} = \exp \mathbf{T}$ with $\mathbf{T}^\dagger = -\mathbf{T}$, that is, as the exponential of an anti-Hermitian matrix.

- Real MOs, *orthogonal* transformation and *antisymmetric* or *skew-symmetric* matrix $\mathbf{T}^T = -\mathbf{T}$.

- Creation operators are transformed as $a_{\tilde{i}}^+ = \exp(\hat{T}) a_i^+ \exp(-\hat{T})$, $\hat{T} = \sum_{ij} T_{ij} \hat{E}_{ij}$

Optimization of MCSCF Wave Functions
Unitary Transformation of CI vector

An exponential ansatz for normalized expansions

- The wave function is $|0\rangle = \sum_m |m\rangle C_m$ with $\sum_m |C_m|^2 = 1$.

- The *complementary space* $|K\rangle$ is orthogonal to $|0\rangle : \langle 0|k\rangle = 0$

- Define $\hat{S} = \sum_{K \neq 0} S_{K0}(|K\rangle \langle 0| - |0\rangle \langle K|)$

- S_{K0} are variational parameters and $\hat{S}^\dagger = -\hat{S}$.

- unitary transformation of $|0\rangle$: $|0'\rangle = e^{\hat{S}} |0\rangle$

- $|0'\rangle$ remains normalized.

Optimization of MCSCF Wave Functions
The wave function ansatz and energy function

$$|0'\rangle = e^{\hat{T}} e^{\hat{S}} |0\rangle$$

Comments

- Transformation of both orbital and configuration space
- Wave function is normalized
- The parameters **T** and **S** can vary freely.

The energy function

- $E(\mathbf{T},\mathbf{S}) = \langle 0| e^{-\hat{S}} e^{-\hat{T}} \hat{H} e^{\hat{T}} e^{\hat{S}} |0\rangle$

Optimization of MCSCF Wave Functions
The wave function ansatz and energy function

The optimal energy

- Vary **T** and **S** such that the energy becomes stationary

$$\frac{\partial E}{\partial T_{ij}} = 0 \quad \frac{\partial E}{\partial S_{K0}} = 0$$

Nonlinear set of equations

- Must be solved iteratively
- Large arsenal of methods from numerical analysis, including
  1. The Newton Raphson method
  2. Approximate /Quasi Newton methods

Optimization of MCSCF Wave Functions
The Newton-Raphson Optimization Method

The Newton-Raphson method in general

- We wish to find a stationary point a function $E(\mathbf{p})$, where $\mathbf{p}$ is a set of parameters that can be freely varied.
- Start with a guess, which for simplicity here we set to zero $\mathbf{p}_0 = \mathbf{0}$.
- Expand E through second order around this point

$$E(\mathbf{p}) \approx E^{(2)} = E(0) + \sum_i \left(\frac{\partial E}{\partial p_i}\right)_0 p_i + \frac{1}{2}\sum_{ij} p_i \left(\frac{\partial^2 E}{\partial p_i \partial p_j}\right)_0 p_j$$

$$= E(0) + \mathbf{g}^\dagger \mathbf{p} + \frac{1}{2}\mathbf{p}^\dagger \mathbf{H}\mathbf{p}$$

- g is the gradient vector and H is the Hessian matrix

Optimization of MCSCF Wave Functions
The Newton-Raphson Optimization Method

The Newton-Raphson method in general

- An approximation to the stationary point is found by finding the stationary point of $E^{(2)}$.
- $\frac{\partial E^{(2)}}{\partial p_i} = 0 \rightarrow \mathbf{g} + \mathbf{H}\mathbf{p} = \mathbf{0}$ ($\mathbf{p} = -\mathbf{H}^{-1}\mathbf{g}$)
- For this p, a new $\mathbf{g}, \mathbf{H}$ is constructed,..
- Continue untill convergence: $\mathbf{g}_N \approx \mathbf{0}$)

Comments

- Approximated E with $E^{(2)} \rightarrow$ only valid for small $\mathbf{p} \rightarrow$ problems far from convergence
- Converges quadratically when $\mathbf{p}$ is small
- The linear equations must often be solved using iterative methods

Optimization of MCSCF Wave Functions

Taylor Expansion of the MCSCF Energy

The wave function and energy

- $|0'\rangle = e^{\hat{T}} e^{\hat{S}} |0\rangle$

- $E = \langle 0'| \hat{H} |0'\rangle = \langle 0| e^{-\hat{S}} e^{-\hat{T}} \hat{H} e^{\hat{T}} e^{\hat{S}} |0\rangle$

Expand through second order in $\hat{T}, \hat{S}$ using the BCH expansion

$$E^{(2)}(\mathbf{T}, \mathbf{S}) = \langle 0| \hat{H} |0\rangle + \langle 0| [\hat{H}, \hat{T}] + [\hat{H}, \hat{S}] |0\rangle$$
$$+ \langle 0| \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{S}], \hat{S}] + [[\hat{H}, \hat{T}], \hat{S}] |0\rangle$$

Optimization of MCSCF Wave Functions

Taylor Expansion of the MCSCF Energy

From last slide

$E^{(2)}(\mathbf{T}, \mathbf{S}) = \langle 0| \hat{H} |0\rangle + \langle 0| [\hat{H}, \hat{T}] + [\hat{H}, \hat{S}] |0\rangle +$
$\langle 0| \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{S}], \hat{S}] + [[\hat{H}, \hat{T}], \hat{S}] |0\rangle$

The orbital gradient

- $\hat{T} = \sum_{ij} T_{ij}(\hat{E}_{ij} - \hat{E}_{ji}) = \sum_{ij} T_{ij} \hat{E}_{ij}^{-}$

- $\frac{\partial E}{\partial T_{ij}} = g_{ij}^{o} = \langle 0| [\hat{H}, \hat{E}_{ij}^{-}] |0\rangle$

- The stationarity requirement $g_{ij}^{o} = 0$ is the *Extended Brillouin Theorem*.

- g_{ij}^{0} vanishes trivilly for some rotations: for example inactive-inactive. These rotations are redundant and are not included.

Optimization of MCSCF Wave Functions
Taylor Expansion of the MCSCF Energy

From last slide

$E^{(2)}(\mathbf{T}, \mathbf{S}) = \langle 0| \hat{H} |0\rangle + \langle 0| [\hat{H}, \hat{T}] + [\hat{H}, \hat{S}] |0\rangle +$
$\langle 0| \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{S}], \hat{S}] + [[\hat{H}, \hat{T}], \hat{S}] |0\rangle$

The CI gradient

- $\hat{S} = \sum_{K \neq 0} S_{K0}(|K\rangle \langle 0| - |0\rangle \langle K|), \ \langle K|0\rangle = 0$
- $\frac{\partial E}{\partial S_K} = g_K^c = 2\langle K| \hat{H} |0\rangle$
- $|K\rangle$ is orthogonal complement to $|0\rangle$ so:
  $\mathbf{g}^c = 0 \rightarrow (1 - |0\rangle \langle 0|)\hat{H} |0\rangle = 0$
- $\mathbf{HC} = E\mathbf{C}$
- The optimal CI-coefficients constitute an solution to the standard CI eigenvalue problem

Optimization of MCSCF Wave Functions
Taylor Expansion of the MCSCF Energy

From last slide

$E^{(2)}(\mathbf{T}, \mathbf{S}) = \langle 0| \hat{H} |0\rangle + \langle 0| [\hat{H}, \hat{T}] + [\hat{H}, \hat{S}] |0\rangle +$
$\langle 0| \frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{S}], \hat{S}] + [[\hat{H}, \hat{T}], \hat{S}] |0\rangle$

The Hessian

$$\mathbf{H} = \left(\begin{array}{cc} \mathbf{H}^{cc} & \mathbf{H}^{co} \\ \mathbf{H}^{oc} & \mathbf{H}^{oo} \end{array} \right)$$

The cc block

- $H_{KL}^{cc} = 2(\langle K| \hat{H} |L\rangle - \delta_{KL} \langle 0| \hat{H} |0\rangle) = 2(H_{KL} - E_0 \delta_{KL})$
- Is the Hamiltonian matrix (in the $|K\rangle$ basis) shifted with the energy

Optimization of MCSCF Wave Functions
Taylor Expansion of the MCSCF Energy

The Orbital Part of the Hessian Matrix

- From $\frac{1}{2}\langle 0|\,[[\hat{H},\hat{T}],\hat{T}]\,|0\rangle$ we obtain

$$H^{oo}_{ij,kl} = \langle 0|\,\hat{E}^-_{ij}\hat{E}^-_{kl}\hat{H}\,|0\rangle + \langle 0|\,\hat{H}\hat{E}^-_{ij}\hat{E}^-_{kl}\,|0\rangle - 2\,\langle 0|\,\hat{E}^-_{ij}\hat{H}\hat{E}^-_{kl}\,|0\rangle$$

- Is dense (no zero elements)

The Coupling Part of the Hessian Matrix

- From $\frac{1}{2}\langle 0|\,[[\hat{H},\hat{T}],\hat{S}]\,|0\rangle$ we obtain

$$H^{co}_{K,ij} = H^{oc}_{ij,K} = 2\,\langle K|\,[\hat{H},\hat{E}^-_{ij}]\,|0\rangle\,.$$

- Is sparse

Optimization of MCSCF Wave Functions
The Newton-Raphson Equations for MCSCF

Form

$$\begin{pmatrix} \mathbf{H}^{cc} & \mathbf{H}^{co} \\ \mathbf{H}^{oc} & \mathbf{H}^{oo} \end{pmatrix} \begin{pmatrix} \mathbf{S} \\ \mathbf{T} \end{pmatrix} = - \begin{pmatrix} \mathbf{g}^c \\ \mathbf{g}^o \end{pmatrix}$$

Number of elements in **H** for medium scale calculation

- Assume: 200 orbitals, 20 occupied orbitals, 10^6 SD's
- Number of orbital rotations: $20 \times 200 = 4000$

  $\mathbf{H}^{oo}$: $4000 \times 4000 = 16 \times 10^6$ - not problematic

  $\mathbf{H}^{oc}$: $4000 \times 10^6 = 4 \times 10^9$ - difficult to store, (but is sparse)

  $\mathbf{H}^{oc}$: $10^6 \times 10^6 = \times 10^{12}$ - difficult to store, (but is sparse)

Cannot be solved using standard methods like LU decomposition

Optimization of MCSCF Wave Functions

Solving the Newton-Raphson Equations for MCSCF

Form

$$\left(\begin{array}{cc} \mathbf{H}^{cc} & \mathbf{H}^{co} \\ \mathbf{H}^{oc} & \mathbf{H}^{oo} \end{array} \right) \left(\begin{array}{c} \mathbf{S} \\ \mathbf{T} \end{array} \right) = - \left(\begin{array}{c} \mathbf{g}^{c} \\ \mathbf{g}^{o} \end{array} \right)$$

- To large to solve using standard decomposition methods

Make approximations and/or use iterative methods

Iterative full second-order Set op a scheme to calculate Hessian times vector directly.

Decouple CI and Orbital part Neglect orbital-configuration coupling. Construct H^{oo} explicitly and solve configuration part by standard CI

Further approximations Approximate H^{oo},done in the super-CI approach

Various codes

- Virtually no package offers *only* a naive NR: can switch off CI/orbital coupling, use damping techniques, etc. Diagonalization of augmented Hessian (norm-extended optimization).
- *Trust-region* approaches that offer guaranteed convergence:
  - Dalton CASSCF/RASSCF (second-order) NEO/NR, coupling.
  - Molpro general MCSCF (second-order+ in MOs), coupling.
- *Quasi-Newton* approaches:
  - Molcas CASSCF/RASSCF (No coupling, first-order+ in MOs)
- Note: Number of iterations does not indicate complexity

Contents

Thursday, Sept. 22

- Optimization of CASSCF wave functions
- MCSCF calculations on excited states
- The CASSI/RASSI method
- A few examples

MCSCF Calculations on Excited States
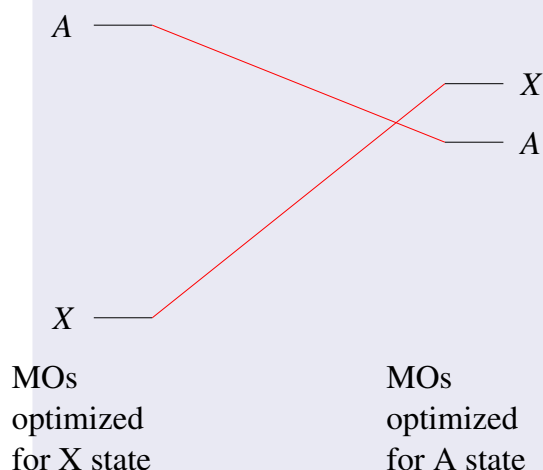Separate calculations on the individual states

Problems

- Calculations on states that are not the lowest of their spin and spatial symmetry are more difficult. . .
  1. Convergence to a saddle point (normally the Hessian has $n-1$ negative eigenvalues in state n).
  2. *Root flipping* may occur- the excited state become the lowest root in the CI
  3. Converged MCSCF wave functions for two roots of the same symmetry are in general not orthogonal.

MCSCF Calculations on Excited States
Separate calculations on the individual states

Root flipping

- LiH: Ground state $c_1 1\sigma^2 2\sigma^2 + c_2 1\sigma^2 3\sigma^2$ ($X^1\Sigma^+$), excited state $1\sigma^2 2\sigma^1 3\sigma^1$ ($A^1\Sigma^+$).

A

X

A

X

MOs optimized for X state

MOs optimized for A state

- Orbitals for excited state differs very much from ground state orbitals $\rightarrow$ excited state becomes lowest state
- Can be handled by methods using the complete Hessian
- However, the upper energy property is lost for the excited state

MCSCF Calculations on Excited States
Separate calculations on the individual states

Issues to be considered

1. Sometimes possible, but not always (root flipping and convergence problems).
2. Hessian eigenvalues should be checked if possible.
3. Local minima for the optimization can arise.
4. States of the same symmetry are generally not orthogonal.
5. How to compute transition properties?

MCSCF Calculations on Excited States
State-averaged MCSCF

Purpose

Obtain orbitals that describe several states

Procedure

- Introduce a (weighted) average of the energies of M states:
  $E_{aver} = \sum_{I=1}^{M} \omega_I E_I$
- all states will have the same orbitals- but different CI coefs.
- The average energy in terms of density matrices:
  $E_{aver} = \sum_{ij} h_{ij} \tilde{D}_{ij} + \sum_{ijkl} g_{ijkl} \tilde{P}_{ijkl}$
- $\tilde{D}_{ij} = \sum_{I=1}^{M} \omega_I D_{ij}^I, \quad \tilde{P}_{ijkl} = \sum_{I=1}^{M} \omega_I P_{ijkl}^I$.
- By calculating more roots in the CI, the same program can be used for average orbital MCSCF calculations

MCSCF Calculations on Excited States
State-averaged MCSCF

Advantages

1. Orthogonality: $\langle I|J \rangle = 0$.
2. Normally much better convergence.
3. Easy to compute transition properties.
4. One calculation for all states.

Disadvantages

1. MOs in different states may be *very* different!
2. May therefore require large active spaces

MCSCF Calculations on Excited States
State-averaged MCSCF

An Example: N,V, states of C_2H_4

- Simplest choice π_u, π_g with two active electrons.
- N state: $C_1(\pi_u)^2 + C_2(\pi_g)^2$
- V state: $(\pi_u \pi_g)_{(S=0)}$
- For the ground (N) state: $\langle \pi_g | z^2 | \pi_g \rangle = 1.69$
- For the excited (V) state: $\langle \pi_g | z^2 | \pi_g \rangle = 9.13$
- The same π_g MO will not do both jobs! At least two are needed (and the V state is actually more difficult than this).

MCSCF Calculations on Excited States
Excited states in practice

- Dalton: only separate calculation on each state. Second-order procedure allows convergence to excited states
- Molcas: separate calculations, or averaging over states of same spin and spatial symmetry.
- Molpro: separate calculations, or averaging over any mixture of states (spin, symmetry and charge can all be different).
- Averaging is often used to ensure nonabelian symmetry- (can be done in a much simpler and more efficient manner)

Contents

Thursday, Sept. 22

- Optimization of CASSCF wave functions
- MCSCF calculations on excited states
- The CASSI/RASSI method
- A few examples

RASSI/CASSI: The RASSCF (CASSCF) State Interaction Method

Matrix elements between different XASSCF(X=C,R) states

- Assume we have obtained XASSCF wave function, $|X\rangle$ and $|Y\rangle$, for two electronic states
- To understand transitions between the states, we must calculate the the *transition moment*: $\langle X| \hat{\mu} |Y\rangle$
- what $\hat{\mu}$ for example is the dipole operator: $\hat{\mu} = \sum_{pq} \vec{\mu}_{pq} \hat{E}_{pq}$.
- Transition matrix element is $\langle X| \hat{\mu} |Y\rangle = \sum_{pq}^{XY} D_{pq}^{XY} \vec{\mu}_{pq}$
- $D_{pq}^{XY} = \langle X| \hat{E}_{pq} |Y\rangle = \sum_{mn} C_m^X C_n^Y D_{pq}^{mn}$ is a transition density matrix.

RASSI/CASSI: The RASSCF (CASSCF) State Interaction Method

The Nonorthogonality Problem

- Calculation of the coupling coefficients D_{pq}^{mn} is easy when the two states are described in the same MO basis.
- In that case D_{pq}^{mn} are the normal one-electron coupling coefficients.
- The same holds when the MO's are not the same but they are *biorthonormal*: $\left\langle p^X \middle| q^Y \right\rangle = \delta_{pq}$
- It is not always possible to transform two MO sets to biorthonormal form without changing the space of the wave function.
- It *is* possible for CASSCF or RASSCF wave functions.

RASSI/CASSI: The RASSCF (CASSCF) State Interaction Method

The CI-transformation technique of Prof. Malmqvist

The idea

- A CI-vector is given for a given set of orbitals
- The orbitals are now changed to a new basis
- Change the CI-coefficients so the state with transformed orbitals is identical to the original state

Complexity of the problem

- For a pair of states, the operation count for the transformation corresponds to CI with a one-electron operator
- That is: Peanuts..

RASSI/CASSI: The RASSCF (CASSCF) State Interaction Method

The algorithm allows

- The computation of transition densities $\mathbf{D}^{XY}$ and $\mathbf{P}^{XY}$.
- Also: overlap integrals: $\langle A|B\rangle$ and Hamiltonian matrix elements $\langle X|\hat{H}|Y\rangle$.
- Allows the solution of the XAS *state interaction* secular problem $\begin{pmatrix} \mathbf{H}_{XX} - E\mathbf{1} & \mathbf{H}_{XY} - E\mathbf{S}_{XY} \\ \mathbf{H}_{YX} - E\mathbf{S}_{YX} & \mathbf{H}_{YY} - E\mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{c}_X \\ \mathbf{c}_Y \end{pmatrix} = 0$
- The resulting states are orthogonal, and non-interacting through the Hamiltonian.
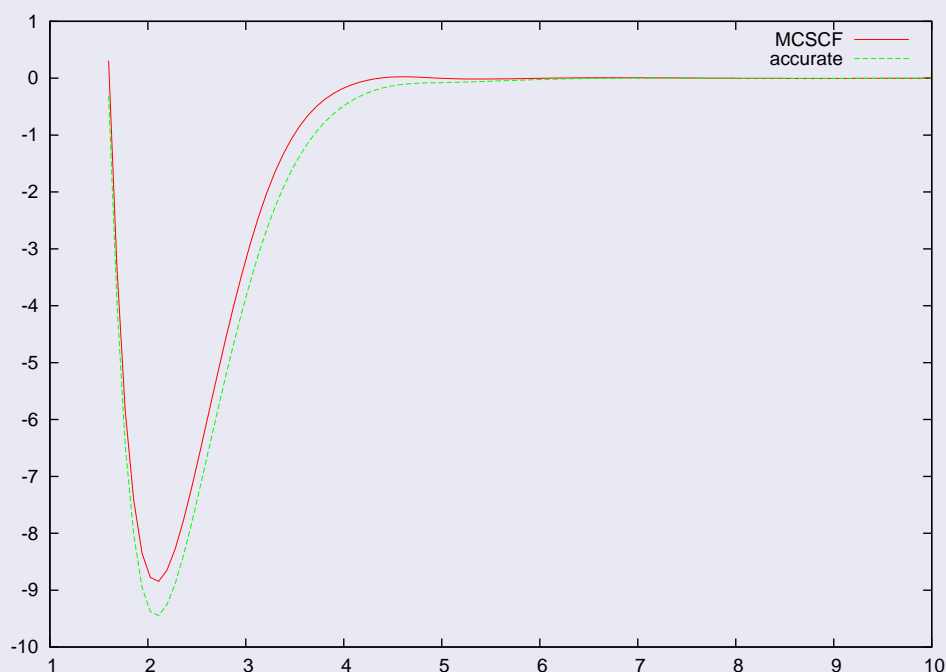- Hundreds of XASSCF states can be handled.

Contents

Thursday, Sept. 22

- Optimization of CASSCF wave functions
- MCSCF calculations on excited states
- The CASSI/RASSI method
- A few examples

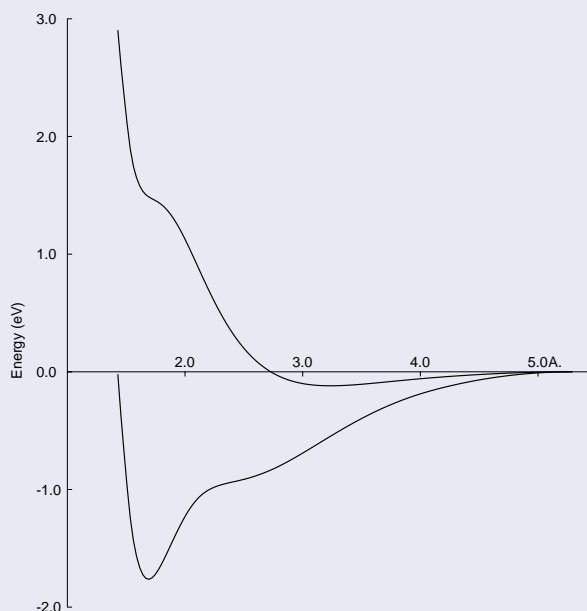# Examples of CASSCF calculations
The N$_2$ ground state

Ground state potential curve

Examples of CASSCF calculations
The Cr$_2$ ground state

MCSCF vs accurate potential curve

Examples of CASSCF calculations
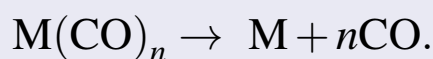
N_2 excitation energies (eV)

	MCSCF	MCSCF(av)	accurate
$^3\Sigma_u^+ \leftarrow {}^1\Sigma_g^+$	7.91	7.76	7.57
$^1\Sigma_u^+ \leftarrow {}^1\Sigma_g^+$	11.36	11.15	10.32

- $^3\Sigma_u^+$ and $^1\Sigma_g^+$ dissociate to $^4S + {}^4S$; $^1\Sigma_u^+$ dissociates to $^2D + {}^2D$.
- Differential dynamical correlation error.

Examples of CASSCF calculations

M(CO)$_n$ binding energies

- Dissociation

$$M(CO)_n \rightarrow M + nCO.$$

- For Ni(CO)$_4$, Fe(CO)$_5$ and Cr(CO)$_6$ CASSCF predicts very little binding (about 100 kJ/mol) compared to accurate results of 550–650.

- Bond lengths too long.

- In fact, much of the CASSCF binding comes from basis set superposition error even in very large basis sets!

Introduction to Dynamical correlation
Wishing list for correlation methods

Size extensivity, definition

- Consider two molecules A, B, infinite apart
- Perform calculation with method X on the two molecules separately $\rightarrow E_A, E_B$
- Perform also calculation with method X on supermolecule containing both A,B $\rightarrow E_{AB}$
- If $E_A + E_B = E_{AB}$ then method X is *size-extensive*

Size-extensive methods

- RHF, UHF, FCI, CC, Perturbation theory (not CASPT2..)
- Active spaces may be choosen, so CAS is size-extensive
- CI, Including MRCI, standard CASPT are not size-extensive

Introduction to Dynamical correlation
Wishing list for correlation methods

Definition of size-consistency

- Consider calculation on a molecule AB with method X
- Increase the distance between A and B, and calculate energy
- Compare the energy and wave function of system A in the limit of infinite distances with that of a calculation on system A by itself.
- If the two energies are identical, method X is size-consistent

Size-consistent methods

- UHF, UHF based correlation methods (-CI)
- Active spaces may be choosen, so CAS is size-consistent
- RHF, RHF based correlation methods are not size-consistent

Introduction to Dynamical correlation

And it is

The short-range correlation that arise as $r_{12} \to 0$.

Characteristica

- Convergence in an orbital-based expansion is very slow.
- Using optimized orbitals does not help for the last %
- Use orbitals from the MCSCF calculation

Methods to determine dynamic correlation

Variational	SRCI and MRCI
Perturbation	Møller-Plesset and various Multireference schemes
Projection	SR or MR Coupled cluster
	Methods containing two-electron functions (F12, geminals, ..) may be combined with the above

Introduction to Dynamical correlation
Combined with static correlation

- We want a method to treat dynamical correlation built on top of MCSCF methods.
- (Or use UHF-based methods and hope...)
- Need "multireference" methods for CC, CI, PT.
- Should preferable be both size-extensive and size-consistent
- Or perhaps some sort of DFT on top of MCSCF (not discussed here).

Contents

Friday, Sept. 23

- Introduction to dynamic correlation
- The multi reference CI method
- A bit on Multireference CC
- Other nearly degenerate multi reference methods
- Perturbation theory, general, MP2 and multireference methods
- CASPT2

The Multireference CI Method

Introduction

The wave function space

- Select a number of reference configurations $\Phi(I)$ based on an MCSCF calculation.
- Generate all singly, $\Phi(I)_i^x$ and doubly, $\Phi(I)_{ij}^{xy}$ excited configurations. i, j are occupied orbitals and x, y occupied or external orbitals.
- Excitation operators $\hat{E}_{xi}$, $\hat{E}_{xi}\hat{E}_{yj}$ operating on each $\Phi(I)$.
- Except in the single-reference case this set of operators may generate redundancies and nonorthogonalities among the excited configurations.
- E.g., H₂ two-reference, $\hat{E}_{x\sigma_g}\hat{E}_{y\sigma_g}$ and $\hat{E}_{x\sigma_u}\hat{E}_{y\sigma_u}$ give the same doubly-excited configuration.

The Multireference CI Method
Introduction

The wave function

- A linear combination of the configurations

$$\Psi_{MRCI} = \sum_I \left[C(I)\Phi(I) + \sum_{ix} C_i^x(I)\Phi(I)_i^x + \sum_{ijxy} C_{ij}^{xy}(I)\Phi(I)_{ij}^{xy} \right]$$

The parameters C

- Are determined using the variational principle
- Leads to the eigen value problem $(\mathbf{H} - E\mathbf{S})\mathbf{C} = 0$

The Multireference CI Method
The Direct CI Method

Idea

- It is not possible to store the Hamiltonian matrix $\mathbf{H}$
- Obtain instead selected roots using iterative methods

Efficient direct CI methods require

- Efficient numerical algorithms
- Efficient routines for the calculation of $\mathbf{HC}$ from integrals

$\sigma = \mathbf{HC}$

- $\hat{H} = \sum_{pq} h_{pq}\hat{E}_{pq} + \frac{1}{2}\sum_{pqrs}(pq|rs)(\hat{E}_{pq}\hat{E}_{rs} - \delta_{qr}\hat{E}_{ps})$
- $\sigma_\mu = \sum_v \left[\sum_{pq} h_{pq}A_{pq}^{\mu v} + \frac{1}{2}\sum_{pqrs}(pq|rs)A_{pqrs}^{\mu v}\right] C_v$
- $A_{pq}^{\mu v}$ and $A_{pqrs}^{\mu v}$ are the *direct CI coupling coefficients*.

The Multireference CI Method
Internal contraction

Problem with MRCI
Number of configurations runs easily into millions and billions

Internal contraction
- Apply excitation operators to the MCSCF wave function Ψ_0, instead of to the configurations in the reference space
- The 'configurations' obtained in this way are much more complicated objects.
- There are far fewer coefficients to optimize: no I dependence $\to c_{ij}^{xy}$, more-or-less independent of the number of CSFs in Ψ_0.
- This approach is termed *internally contracted* MRCI.

The Multireference CI Method
Internal contraction

The pro and cons
- Internal contraction reduces variational freedom and will raise the energy.
- This is rarely an issue, especially in implementations that "relax" the coefficients in Ψ_0 during the iterations.
- Occasional problems with particularly singles coefficients, and these are sometimes then uncontracted.
- Molpro: internally contracted MRCI
  Molcas: no contraction in MRCI.

The Multireference CI method

Advantages

- Probably the most accurate method available for small molecules.
- Balanced calculations for several electronic states.

Disadvantages

- MRCI is not size-extensive.
- The size of the uncontracted CI expansion grows quickly with the number of reference configurations.
- Even with internal contraction large multiconfiguration reference functions become intractable.
- Becomes less and less practical for large molecules.

Contents

Friday, Sept. 23

- Introduction to dynamic correlation
- The multi reference CI method
- A bit on Multireference CC
- Other nearly degenerate multi reference methods
- Perturbation theory, general, MP2 and multireference methods
- CASPT2

Multireference CC?

- This is not easy. The expansion of $\exp(-T)H\exp(T)$ in commutators does not terminate after five terms, as in closed-shell CC (or UHF CC).
- In fact, the termination is messily excitation-level dependent, making the equations highly nonlinear and the work substantial.
- A lot of advances in recent years
- Or with single-reference high-excitation level approaches, eliminating some terms. E.g., to dissociate N$_2$ needs up to six-fold excitation in the valence space — start with CCSDTQ5678 (!) but eliminate terms that are more than (say) doubles with respect to the valence-space configurations.
- Remains a very active, but challenging area.

Contents

Friday, Sept. 23

- Introduction to dynamic correlation
- The multi reference CI method
- A bit on Multireference CC
- Other nearly degenerate multi reference methods
- Perturbation theory, general, MP2 and multireference methods
- CASPT2

Other (nearly) size-extensive multireference methods
Multireference ACPF and relatives

Idea

- Use MRCI form of wave function: $\Psi = \Psi_0 + \Psi_a + \Psi_e$
- Ψ_a is in reference space, Ψ_e contain the external excitations
- Modify Energy-expression

The energy expression

- $E = \frac{\langle \Psi_0 + \Psi_a + \Psi_e | H - E_0 | \Psi_0 + \Psi_a + \Psi_e \rangle}{1 + g_a \langle \Psi_a | \Psi_a \rangle + g_e \langle \Psi_e | \Psi_e \rangle}$

Various choices of $g \rightarrow$ various methods

MRCI: $g_a = 1$, $g_e = 1$.   MRLCCM: no Ψ_a, $g_e = 0$.

MRACPF: $g_a = 1$, $g_e = 2/N$.   QDVPT: $g_a = 1$, $g_e = 0$.

MRCEPA(0): $g_a = 0$, $g_e = 0$.

Other (nearly) size-extensive multireference methods
Multireference Davidson corrections

Idea

- Perform a MRCI calculation as usual
- Add a correction to the obtained correlation energy to obtain a better (lower) energy

Two forms

- Two different corrections:
  $\Delta E_{\text{corr}} = (E_{\text{MRCI}} - E_{\text{REF}})\{1 - \sum_R (c_R^{\text{MRCI}})^2\}$
  $\Delta E_{\text{corr}} = (E_{\text{MRCI}} - E_{\text{REF}})\{1 - \sum_R (c_R^{\text{MRCI}} c_R^{\text{REF}})\}$
- Both trivial to compute.

Problems

- Weak theoretical foundation- but works often anyhow!!

Contents

Friday, Sept. 23

- Introduction to dynamic correlation
- The multi reference CI method
- A bit on Multireference CC
- Other nearly degenerate multi reference methods
- Perturbation theory, general, MP2 and multireference methods
- CASPT2

Multi reference perturbation theory
Rayleigh-Schrödinger Perturbation Theory

Divide Hamiltonian

$\hat{H} = \hat{H}_0 + \lambda \hat{H}_1.$

Expand the wave function and energy in λ

- $\Psi = \Psi_0 + \lambda \Psi_1 + \lambda^2 \Psi_2 + \ldots$
- $E = E_0 + \lambda E_1 + \lambda^2 E_2 + \ldots.$

Insert in Schrödinger equation to obtain

$$(\hat{H}_0 - E_0)|\Psi_0\rangle = 0$$
$$(\hat{H}_0 - E_0)|\Psi_1\rangle = (E_1 - \hat{H}_1)|\Psi_0\rangle$$
$$(\hat{H}_0 - E_0)|\Psi_2\rangle = (E_1 - \hat{H}_1)|\Psi_1\rangle + E_2|\Psi_0\rangle$$

Multi reference perturbation theory
Rayleigh-Schrödinger Perturbation Theory

From last slide
- $(\hat{H}_0 - E_0)|\Psi_1\rangle = (E_1 - \hat{H}_1)|\Psi_0\rangle$
- $(\hat{H}_0 - E_0)|\Psi_2\rangle = (E_1 - \hat{H}_1)|\Psi_1\rangle + E_2|\Psi_0\rangle$

Normalization
- Impose intermediate normalization $\langle\Psi_0|\Psi_n\rangle = 0 (n > 0)$

To obtain the second-order energy
- $E_2 = \langle\Psi_0|\hat{H}_1|\Psi_1\rangle$
- $(\hat{H}_0 - E_0\hat{1})|\Psi_1\rangle = -\hat{H}_1|\Psi_0\rangle$

Multi reference perturbation theory
Rayleigh-Schrödinger Perturbation Theory, second order

From last slide
- $E_2 = \langle\Psi_0|\hat{H}_1|\Psi_1\rangle$
- $(\hat{H}_0 - E_0\hat{1})|\Psi_1\rangle = -\hat{H}_1|\Psi_0\rangle$

First-order correction
- Expand Ψ_1 in a basis: $\Psi_1 = \sum_\mu C_v \Phi_v$
- $\Psi_0 = \Phi_0$, $\langle\Phi_\mu|\Phi_v\rangle = \delta_{\mu v}$.
- Insert in first-order equation, project with Φ_μ

$$\sum_v \left(E_0\delta_{\mu v} - \langle\Phi_\mu|\hat{H}_0|\Phi_v\rangle\right) C_v = \langle\Phi_\mu|\hat{H}_1|\Phi_0\rangle.$$

Multi reference perturbation theory
Rayleigh-Schrödinger Perturbation Theory, second order

Diagonal representation

- If the Φ_μ are eigenfunctions of $\hat{H}_0$ with eigenvalues E_μ we obtain trivially:

$$C_\mu = -\frac{\langle \Phi_\mu | \hat{H}_1 | \Psi_0 \rangle}{E_\mu - E_0}$$

with the second-order energy:

$$E_2 = -\sum_\mu \frac{|\langle \Phi_\mu | \hat{H}_1 | \Psi_0 \rangle|^2}{E_\mu - E_0}.$$

Multi reference perturbation theory
Various forms

Many forms, differs by choices of

1. Reference state
2. Form of H_0
3. Form of wave function corrections
4. (Use of intermediate Hamiltonian, buffer states …)

Reference state

- CASSCF/RASSCF…
- Incomplete spaces

Multi reference perturbation theory
Various forms

Form of $\hat{H}_0$

- Use of Fock-type one-electron operator in general (CASPT)
- Use of Fock-type one-electron operator in inactive and secondary spaces, full two-electron Hamilton operator in actice space (NEVPT)

Form of wave function corrections

- Complete internal contraction - MOLCAS CASPT, NEVPT
- Partial internal contraction - MOLPRO CASPT
- No internal contraction: many approaches

Contents

Friday, Sept. 23

- Introduction to dynamic correlation
- The multi reference CI method
- A bit on Multireference CC
- Other nearly degenerate multi reference methods
- Perturbation theory, general, MP2 and multireference methods
- CASPT2

CASPT2

Form of reference state, correction, and $\hat{H}_0$

Reference state

- CAS state - with or without optimized orbitals

Form of correction

- Divide the N-electron space into

  The reference function: $\qquad \Psi_0 = |CASSCF\rangle = |0\rangle$

  The complementary CAS space: $\qquad |K\rangle$

  SD substitutions from the reference: $\quad |pqrs\rangle = \hat{E}_{pq}\hat{E}_{rs}|\Psi_0\rangle$

  The remaining configuration space: $\quad |X\rangle$.

- Only the SD space in CASPT2: interacts with $|0\rangle$

The zero-order Hamiltonian

- $\hat{H}_0 = \hat{P}_0\hat{F}\hat{P}_0 + \hat{P}_K\hat{F}\hat{P}_K + \hat{P}_{SD}\hat{F}\hat{P}_{SD} + \hat{P}_X\hat{F}\hat{P}_X$

- $\hat{F}$ is a Fock-type one-electron operator, several choices

CASPT2

The Fock Operator

Defined as

- $\hat{f} = \sum_{pq} f_{pq} E_{pq}$

- $f_{pq} = \sum_\sigma \langle\Phi_0|\,[a_{p\sigma},[H,a_{q\sigma}^\dagger]_+]\,|\Phi_0\rangle =$
  $h_{pq} + \sum_{rs} D_{rs}\left[(pq|rs) - \frac{1}{2}(pr|qs)\right]$.

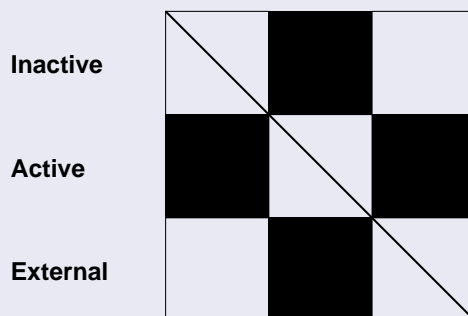- Becomes standard MP $\hat{H}_0$ if reference is a closed shell system

Properties of **f**

Inactive orbitals	$f_{pp} = -IP_p$
External orbitals	$f_{pp} = EA_p$
$n_p^{occ} = 1$:	$f_{pp} = -\frac{1}{2}(IP_p + EA_p)$

- Unpaired electrons are favoured, so binding energies are too

CASPT2

Structure of the Fock matrix

- The Fock matrix can be made diagonal *within* each orbital subspace

Inactive

Active

External

- Elements in nondiagonal block are in general nonvanishing

CASPT2

Fock Matrix Elements and the First-Order Equation

Fock Matrix Elements

- Contain up to fourth-order density matrices:
  $\langle pqrs | \hat{F} | p'q'r's' \rangle = \sum_{mn} f_{mn} \langle \Psi_0 | \hat{E}_{sr} \hat{E}_{qp} \hat{E}_{mn} \hat{E}_{p'q'} \hat{E}_{r's'} | \Psi_0 \rangle$
- May be rearranged to give involve a three-body density

The first-order equation

- $(\mathbf{H}_0 - E_0 \mathbf{S})\mathbf{C} = -\mathbf{V}$
- There is a total of eight different types of double excitations
- Equations may be split into a 8 blocks, one for each of 8 excitation types
- The coupling of the blocks may then be introduced in a second iterative step.

CASPT2
Computational steps in a CASPT2 calculation

1. Perform a CASSCF calculation, single-state or state-average.
2. Transform the Fock matrix to pseudo-diagonal form.
3. Transform two-electron integrals with at least two indices corresponding to occupied orbitals (second order transformation).
4. Compute **S** and some additional matrices and diagonalize them.
5. Compute the second-order energy in the "diagonal" approximation.
6. Solve the large system of linear equations introducing the coupling arising from the non-diagonal blocks of the Fock matrix **F**.

CASPT2

Intruder States in CASPT2

- CASPT2 will only be an adequate method when the perturbation is small.
- All large CI coefficients should thus be included in the CAS space.
- When large coefficients appear in the first-order wave function (weight of the reference function is small), the active space should in general be increased.
- When the interaction of a specific state with the reference function is small, the effect of that state may be removed using a level shift technique.
- This is the *intruder state* problem in CASPT2.

CASPT2
The Reference Weight

- Write the first-order wave function as: $|\Psi'\rangle = |\Psi_0\rangle + |\Psi_1\rangle$
- Define S_1 by $\langle\Psi'|\Psi'\rangle = 1 + S_1$.
- Introduce $\omega = 1/(1+S_1)$
- The normalized function is $|\Psi\rangle = \sqrt{\omega}\,|\Psi_0\rangle + \sqrt{1-\omega}\,|\Psi_1\rangle$
- Since CASPT2 is nearly size-extensive, ω will decrease when the number of electrons increases.

Typical values of ω

- Order of magnitude: $\omega = (1+\alpha)^{-N/2}$, $\alpha \approx 0.015$
- $N = 10 \rightarrow \omega = 0.93$, $N = 100 \rightarrow \omega = 0.48$.
- Values much smaller than these estimates indicate an intruder

CASPT2
Level-shifted CASPT2

- Introduce a level shift ε in the first-order equation
- $(\hat{H}_0 - E_0 + \varepsilon)\tilde{\Psi}_1 = (E_1 - \hat{H}_1)\Psi_0$
- (tilde denoting a level-shifted quantity)
- For a diagonal $\hat{H}_0$: $\tilde{E}_2 = E_2 + \varepsilon\sum_\mu |\tilde{C}_\mu|^2 \left(1 + \frac{\varepsilon}{\varepsilon_\mu - E_0}\right)$.
- Assume that $\varepsilon_\mu - E_0 \gg \varepsilon$ and neglecting quadratic terms: $E_2 \approx \tilde{E}_2 - \varepsilon\left\langle\tilde{\Psi}_1\middle|\tilde{\Psi}_1\right\rangle$.
- This technique removes the effect of the intruder state without large changes od the contributions from the other states.

CASPT2
Multi-State (MS) CASPT2

Idea

- Diagonalize approx. Hamiltonian in space of CASPT2 states
- First perturb, then diagonalize scheme
- Natural extension of XASSI (X=C,R)

The method

- Assume a number of CASSCF states, $E_i, \Psi_i, i = 1, N,$
- The CASPT first-order wave functions are denoted $\chi_i, i = 1, N.$
- Use $\Psi_i + \chi_i$ as basis functions in a approx. var. calc.
- Effective Hamiltonian has elements:
  $(H_{eff})_{ij} = \delta_{ij} E_i + \langle \Psi_i | \hat{H} | \chi_j \rangle$
- Always recommended when several states of the same symmetry are considered.

CASPT2

Advantages of CASPT2

- A CAS wave function is the reference, so very general cases may be treated.
- The formalism is independent of the size of the CAS CI space. Thus large expansions for Ψ_0 may be used.
- The size of the contracted SD space is never large and is independent of the CAS CI space.
- The formalism is nearly size-extensive. Therefore a large number of electrons may be correlated (more than 100 in practical applications).
- The method has the same orbital invariance as the CASSCF method.

CASPT2

Applications of CASPT2

- Energy surfaces for ground and excited states.
- Electronic spectroscopy, including all types of excited states (singly, doubly, etc. excited, valence and Rydberg states, charge transfer, etc.).
- The whole periodic system from H to Pu (scalar relativity in CASSCF, spin-orbit with RASSI).
- Radicals and biradicals, positive and negative ions.
- Large molecules where MRCI is not applicable (calculations on systems with up to 50 atoms have been performed).

CASPT2

Problems with CASPT2

- The active space may become prohibitively large (today's limit 14-16 active orbitals).
- The intruder state problem.
- The zeroth-order Hamiltonian has a (small) systematic error (has recently been fixed with a shifted Hamiltonian, the IPEA shift).
- Transition properties are obtained at the CASSCF level, not CASPT2.
- It is not a "black box" method. (But is this really a problem/disadvantage. . . ?)

CASPT2

Applications

N$_2$ ground-state spectroscopic constants

	r_e (Å)	ω_e (cm^{-1})	D_e (eV)
CASSCF	1.102	2329	8.91
CASPT2	1.100	2312	9.43
ACPF	1.099	2330	9.79
CCSD(T)	1.099	2332	9.82
Expt	1.098	2359	9.90

CASPT2

Applications

Cr$_2$ potential curve