Multiconfigurational and multireference methods

Peter R. Taylor and Jeppe Olsen

ESQC-19, Torre Normanna, September 2019

September 11-13, 2019
Contents of the lectures

Wednesday, Sept. 11
- Molecular orbitals for H\(_2\) using RHF and UHF
- Introduction to the MCSCF method
- The CASSCF method
- Examples of some simple CASSCF calculations
- Excited states

Thursday, Sept. 12
- RASSCF, optimization, CASSI/RASSI

Friday, Sept. 13
- Adding dynamic correlation

P.R. Taylor and J. Olsen, Multiconfigurational Methods
A reminder - HF and CI

Hartree-Fock

- Minimize the energy of one Slater determinant or a symmetry/spin-defined combination of Slater determinants
- The individual unoccupied orbitals have a limited meaning

CI: Configuration Interaction

- Expand wave function as a linear combination of Slater determinants $|0\rangle = \sum I C_I |SD_I\rangle$
- The orbitals are fixed
- The parameters $C_I$ are chosen to minimize the energy
- Leads to eigenvalue problem $HC = EC, H_{IJ} = \langle SD_I | \hat{H} | SD_J \rangle$
- If all determinants of a given basis is included, the Schrödinger equation projected on the basis is solved - full CI (FCI)
Molecular Orbitals for $\text{H}_2$

**RHF**

The molecule

$$H_A \quad R \quad H_B$$

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

- MOs defined by symmetry:
  - $1\sigma_g = N_g(1s_A + 1s_B)$
  - $1\sigma_u = N_u(1s_A - 1s_B)$.
  - $N_g$ and $N_u$ are normalization constants.

- Closed-shell RHF configuration $1\sigma_g^2$,
  - $\Phi_{\text{RHF}} = |1\sigma_g \alpha 1\sigma_g \beta| (a_{1\sigma_g \alpha}^+ a_{1\sigma_g \beta}^+ |\text{vac}\rangle)$.
  - “restricted HF” (RHF) wave function (proper eigenfunctions of spin and spatial operator).
Molecular Orbitals for $\text{H}_2$

$\text{RHF}, \Psi = \Phi_{\text{RHF}} = |1\sigma_g \alpha 1\sigma_g \beta| \ (a_{1\sigma_g}^+ \alpha a_{1\sigma_g}^+ \beta |\text{vac}\rangle)$

At equilibrium

$\Phi_{\text{RHF}}$ is a good approximation to the exact wave function.

$R \rightarrow \infty$

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

- $1\sigma_g \rightarrow \frac{1}{\sqrt{2}} (1s_A + 1s_B)$

- $\Phi_{\text{RHF}} \rightarrow$
  
  $\frac{1}{2} (|1s_A \alpha 1s_B \beta| + |1s_B \alpha 1s_A \beta| + |1s_A \alpha 1s_A \beta| + |1s_B \alpha 1s_B \beta|)$,

- Unphysical 50:50 mixture of “covalent” and “ionic” terms.

- Consequence of the electrons in HF do not correlate the motion
Molecular Orbitals for $\text{H}_2$

RHF

Potential curves using RHF and FCI for $\text{H}_2$, cc-pVQZ basis

P.R. Taylor and J. Olsen, Multiconfigurational Methods

ESQC-19
Molecular Orbitals for $\text{H}_2$

**UHF**

**Orbitals and wave function (Coulson and Fischer-Hjalmars)**

- MOs are not symmetry adapted.
- $\psi_1 = N(1s_A + \lambda 1s_B)$, $\psi_2 = N(\lambda 1s_A + 1s_B)$.
- $\Phi_{\text{UHF}} = |\psi_1 \psi_2|$.
- Not an eigenfunction of $\hat{S}^2$.

**$\lambda = ?$**

- Optimize energy with respect to $\lambda$ at each value of $R$.
- Near $R_e \lambda = 1$, whereas as $R \to \infty$, $\lambda \to 0$.
- $\Phi_\lambda$ for optimum $\lambda$ is an example of an **unrestricted** Hartree-Fock (UHF) wave function.
Molecular Orbitals for $\text{H}_2$

UHF

Potential curves using UHF, RHF and FCI for $\text{H}_2$, cc-pVQZ basis
Molecular Orbitals for $H_2$

**UHF**

**Advantages**

- Wave function and energy at dissociation: dissociates to two H atoms.
- Wave function becomes RHF wave function around equilibrium $\rightarrow$ good approximation here.

**Disadvantages**

- Does not display inversion symmetry $\rightarrow$ symmetry-broken approximation.
- $\Phi_{\text{UHF}}$ is an eigenfunction of $\hat{S}_z$ but not of $\hat{S}^2$ $\rightarrow$ spin-broken approximation.
- Large errors in region where bonds are broken.
The symmetry problem: to break or not to break

Symmetry and HF

Do worry-restrict

- Argument: symmetry-properties are important, so keep them.
- Once a symmetry property is lost, it may be hard to regain it.

Do not worry-unrestricted

- We are making so many approximations, so why worry?
- (Life is too short for symmetry adaptation.)
- As correlation is added, symmetry will be recovered.
- Gives a very simple wave function - a single determinant.

Projected and extended Hartree-Fock

- Project on correct symmetry (Projected HF) and perhaps reoptimize (Extended HF). Recent work by G. Scuseria.
The symmetry problem: to break or not to break

UHF in practice

- UHF is cheap and straightforward. Deficiencies are easy to monitor (look at $\langle S^2 \rangle$).
- Doubles the number of parameters compared to RHF. Not a big problem in UHF, but increases the work 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. 11

- Molecular orbitals for $H_2$ using RHF and UHF
- Introduction to the MCSCF method
- The CASSCF method
- Examples of some simple CASSCF calculations
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: \( |\tilde{0}\rangle = \sum C_I |\tilde{I}\rangle \) where \( |I\rangle \) are Slater determinants (ONV’s) or configuration state functions.
- Optimize the orbitals and the CI coefficients \( C_I \).
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'.
MCSCF
Introduction

What can be expected from MCSCF? (correlation)

• MCSCF does not describe the short-range terms that arise as $r_{12} \to 0$: *dynamical correlation*.

• (Convergence of dynamical correlation in an orbital-based expansion requires huge orbital-expansions.)

• MCSCF includes the *nondynamical* (or “static”) correlation that comes from configurational near-degeneracies or from gross deficiencies in the RHF wave function.

• (Static correlation is converged if the near-degenerate orbitals are included)

• Aims at describing part of correlation energy that changes, not the complete correlation energy
Wednesday, Sept. 11

- Molecular orbitals for \( \text{H}_2 \) using RHF and UHF
- The symmetry problem
- Introduction to the MCSCF method
- The CASSCF method
- Examples of some simple CASSCF calculations
- Excited states
The CASSCF method

Simple to define

- Number of orbitals (per symmetry) and number of electrons in the active space.
- There is usually also 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

Our \( \text{H}_2 \) cas

2 electrons in 2 active orbitals

\[ \Phi_0 = c_g |1\sigma_g \overline{1\sigma_g}| + c_u |1\sigma_u \overline{1\sigma_u}| \]

- Both CI coefficients and orbitals are optimized.
- 2 in 2 CAS.
- Describes wave function well both at equilibrium and at dissociation.
- Well defined spatial and spin symmetry.
- More complicated to optimize than UHF.
The CASSCF method

Potential curves using UHF, RHF, CAS 2in2 and FCI for $H_2$, cc-pVQZ basis

P.R. Taylor and J. Olsen, Multiconfigurational Methods
# The CASSCF method

The size of the configuration expansion

<table>
<thead>
<tr>
<th>k</th>
<th>Number of Slater determinants (aka ONV’s) for 2k electrons in 2k orbitals</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>36</td>
</tr>
<tr>
<td>6</td>
<td>400</td>
</tr>
<tr>
<td>8</td>
<td>4.900</td>
</tr>
<tr>
<td>10</td>
<td>63.504</td>
</tr>
<tr>
<td>12</td>
<td>853.776</td>
</tr>
<tr>
<td>14</td>
<td>11.778.896</td>
</tr>
<tr>
<td>16</td>
<td>165.636.896</td>
</tr>
<tr>
<td>18</td>
<td>2.363.904.260</td>
</tr>
<tr>
<td>20</td>
<td>34.134.777.856</td>
</tr>
<tr>
<td>24</td>
<td>7.312.459.672.336</td>
</tr>
</tbody>
</table>

- # SD’s = \( \binom{2k}{k}^2 \)
- For large \( k \): # SD’s = \( \frac{16^k}{k\pi} \)
- Adding one Cr atom (2k=6) \( \sim \) 18 years of computer development
- Standard CI: largest calc: 24 in 24, practical: 20 in 20
- DMRG and Monte Carlo CI allow larger spaces.
The CASSCF method

Choosing active spaces

Simple diatomics

- Sometimes the problem seems and is simple: H\textsubscript{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\textsubscript{2} ground state will need \((\sigma_g, \pi_u, \pi_g, \sigma_u)^6\) or the 2p orbitals.

- C\textsubscript{2} will need the 2s orbitals as well, because of the 2s/2p near-degeneracy in C atom.

- F\textsubscript{2}: \((\sigma_g, \sigma_u)^2\) (the bonding/antibonding pair) and \((\sigma_g, \pi_u, \pi_g, \sigma_u)^{10}\) gives only a qualitative result (bonding…). The 3p orbitals are needed!

- Even diatomics are not straightforward.
The CASSCF method
Choosing active spaces

A bit on polyatomic molecules
- In most polyatomics, the full valence space is too large.
- Identify the orbitals involved in the process.
- Spectroscopy of C\textsubscript{6}H\textsubscript{6}: use the six $\pi$ MOs.
- Breaking a CH or CC bond in a hydrocarbon: use $\sigma, \sigma^*$
- May have to refine this choice (we will see how). Clearly very far from a black-box approach!

Even larger active spaces
- More than about 20 in 20 is very demanding.
- Consider alternative expansions, RAS, GAS.
- Or alternative ways for FCI: DMRG, Monte Carlo CI.
Wednesday, Sept. 11

- Molecular orbitals for $\text{H}_2$ using RHF and UHF
- The symmetry problem
- Introduction to the MCSCF method
- The CASSCF method
- Examples of some simple CASSCF calculations
- Excited states
Examples of CASSCF calculations

The $\text{N}_2$ ground state

- Curves have been shifted to be identical at infinite separation.
- Qualitative, but not quantitative correct.
Examples of CASSCF calculations
The Cr$_2$ ground state

MCSCF vs accurate potential curve

- Wave function is actually rather accurate..

P.R. Taylor and J. Olsen, Multiconfigurational Methods
Examples of CASSCF calculations

$N_2$ excitation energies (eV)

<table>
<thead>
<tr>
<th>Excitation</th>
<th>MCSCF</th>
<th>MCSCF (av)</th>
<th>Accurate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^3\Sigma_u^+ \leftarrow ^1\Sigma_g^+$</td>
<td>7.91</td>
<td>7.76</td>
<td>7.57</td>
</tr>
<tr>
<td>$^1\Sigma_u^+ \leftarrow ^1\Sigma_g^+$</td>
<td>11.36</td>
<td>11.15</td>
<td>10.32</td>
</tr>
</tbody>
</table>

$^3\Sigma_u^+$ and $^1\Sigma_g^+$ dissociate to $^4S + ^4S$; $^1\Sigma_u^+$ dissociates to $^2D + ^2D$.

Differential dynamical correlation error.

P.R. Taylor and J. Olsen, Multiconfigurational Methods ESQC-19
Examples of CASSCF calculations

M(CO)$_n$ binding energies

- Dissociation

  \[ \text{M(CO)}_n \rightarrow \text{M} + n\text{CO}. \]

- 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!
Examples of CASSCF calculations
The Mn$_4$O$_5$Ca complex using DMRG

- CAS with 44 electrons in 35 orbitals
- The first detailed investigation into the bonding and oxidation states
Contents

Wednesday, Sept. 11

- Molecular orbitals for H\textsubscript{2} using RHF and UHF
- The symmetry problem
- Introduction to the MCSCF method
- The CASSCF method
- Examples of some simple CASSCF calculations
- Excited states

P.R. Taylor and J. Olsen, Multiconfigurational Methods
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, example: LiH

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

- Orbitals for the excited state differs very much from the 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.

P.R. Taylor and J. Olsen, Multiconfigurational Methods
# 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:
  \[
  E_{\text{aver}} = \sum_{I=1}^{M} \omega_I E_I
  \]
- States will have identical orbitals— but different CI coefs.
- A standard single state program is easily modified to perform state-averaged MCSCF calculations

### Problem

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

P.R. Taylor and J. Olsen, Multiconfigurational Methods

ESQC-19
MCSCF Calculations on Excited States

State-averaged MCSCF

An Example: N,V, states of C₂H₄

- Simplest choice $\pi_u, \pi_g$ 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: $\langle \pi_g | z^2 | \pi_g \rangle = 9.13$
- A single $\pi_g$ MO cannot do both jobs! At least two $\pi_g$-orbitals are needed (and the V state is actually more difficult than this).

P.R. Taylor and J. Olsen, Multiconfigurational Methods
Contents

Thursday, Sept. 12

- Extensions of the CASSCF method
- Optimization of CASSCF wave functions
- The CASSI/RASSI method
Extensions of the CASSCF method

The RAS construction: Three active orbital spaces

**RAS1 orbitals:** Max number of holes

**RAS2 orbitals:** $n_{occ}$ varies

**RAS3 orbitals:** Max number of electrons

Many type of CI expansions within a common framework.
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 ORMAS (occupation restricted multiple active spaces)

It is all right now, in fact it is a GAS
Jagger/Richards
Common aspects of the RASSCF and GASSCF methods

**Advantages**
- Much larger number of active orbitals may now be employed
- Expansions that are tailored a given problem may be devised
- Part of the dynamic correlation may be calculated

**Disadvantages**
- Dynamic correlation is not obtained with high accuracy - an active space of valence orbitals $+ 1s1p1d$ per (light) atom recovers about 70% of correlation energy
- Convergence of the optimization procedure is typically much slower due to strong coupling between CI and orbital changes

P.R. Taylor and J. Olsen, Multiconfigurational Methods
Contents

Thursday, Sept. 12

- Extensions of the CASSCF method
- **Optimization of CASSCF wave functions**
- The CASSI/RASSI method
Optimization of MCSCF Wave Functions

The wave function to be optimized

\[ |\tilde{0}\rangle = \sum_m |\tilde{m}\rangle c_m \]

The optimization problem

- Determine the MOs and the MC coefficients using the variational principle
- \[ \delta E = \delta \left( \frac{\langle 0 | \hat{H} | 0 \rangle}{\langle 0 | 0 \rangle} \right) = 0 \]
Optimization of MCSCF Wave Functions

The energy

Non-Relativistic Hamiltonian (Second Quantization)

\[ \hat{H} = \hat{h} + \hat{g} + h_{\text{nuc}} = \sum_{pq} h_{pq} E_{pq} + \frac{1}{2} \sum_{pqrs} g_{pqrs} (E_{pq} E_{rs} - \delta_{qr} E_{ps}) + h_{\text{nuc}} \]

Contribution from one-electron operator

\[ \langle 0 | \hat{h} | 0 \rangle = \sum_{pq} h_{pq} D_{pq} \]

\[ D_{pq} = \langle 0 | E_{pq} | 0 \rangle : \text{The first order reduced density matrix.} \]

Contribution from the two-electron operator

\[ \langle 0 | \hat{g} | 0 \rangle = \sum_{pqrs} g_{pqrs} P_{pqrs} \]

\[ P_{pqrs} = \frac{1}{2} \langle 0 | E_{pq} E_{rs} - \delta_{qr} E_{ps} | 0 \rangle : \text{second order density matrix} \]
Optimization of MCSCF Wave Functions

The energy

Summary

- Total MCSCF energy from sum over occupied orbitals

\[
E = \langle \tilde{0} | \hat{H} | \tilde{0} \rangle = \sum_{pq} h_{pq} D_{pq} + \sum_{pqrs} g_{pqrs} P_{pqrs} + h_{nuc}.
\]

- Can be written as summations over active orbitals

\[
E = \langle \tilde{0} | \hat{H} | \tilde{0} \rangle = \sum_{vw} F_{vw}^{I} D_{vw} + \sum_{vwxy} g_{vwxy} P_{vwxy} + E_{core}.
\]

- The molecular orbital coefficients appear in the \( h_{pq} \) and \( g_{pqrs} \).
- The CI coefficients appear in \( D \) and \( P \).
- Energy depends on MOs \( \tilde{\phi} \) and CI coefficients \( c \).
Optimization of MCSCF Wave Functions

Exponential parameterization

- MOs are orthonormal and we wish to preserve this.
- A unitary transformation of the orbitals \( \tilde{\varphi} = \varphi U \), \( U^\dagger U = 1 \), ensures orthonormality.
- Any unitary matrix can be written in the form \( U = \exp T \) with \( T^\dagger = -T \), (\( T \) is anti-hermitian).
- Real MOs: orthogonal transformation and antisymmetric matrix \( T^T = -T \).
- Creation operators are transformed as \( \hat{a}_p^\dagger = \exp(\hat{T})a_p^+ \exp(-\hat{T}) \), \( \hat{T} = \sum_{pq} T_{pq} E_{pq} = \sum_{p>q} T_{pq}(E_{pq} - E_{qp}) \).
- For CAS: only inactive \( \rightarrow \) active and secondary, active \( \rightarrow \) secondary excitations are included (nonredundant excitations).
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 $|0\rangle$: $|\tilde{0}\rangle = e^{\hat{S}} |0\rangle$.
- $|\tilde{0}\rangle$ remains normalized.
Optimization of MCSCF Wave Functions
The wave function ansatz and energy function

\[ |\tilde{0}\rangle = e^{\hat{T}} e^{\hat{S}} |0\rangle \]

Comments
- Transformation of both orbital and configuration space.
- The operators \( \hat{T}, \hat{S} \) are antihermitian.
- Wave function is normalized.
- The parameters \( T_{pq}, p > q \) and \( S_{K0} \) can vary independently.
- Other forms are normally used.

The energy function

\[ E(T, S) = \langle \tilde{0} | \hat{H} | \tilde{0} \rangle = \langle 0 | e^{-\hat{S}} e^{-\hat{T}} \hat{H} e^{\hat{T}} e^{\hat{S}} |0\rangle. \]
## Optimization of MCSCF Wave Functions

### The wave function ansatz and energy function

#### The energy function (from last slide)

\[
E(T, S) = \langle \tilde{0} | \hat{H} | \tilde{0} \rangle = \langle 0 | e^{-\hat{S}} e^{-\hat{T}} \hat{H} e^{\hat{T}} e^{\hat{S}} | 0 \rangle
\]

#### The optimal energy

- Vary \( T \) and \( S \) such that the energy becomes stationary.

\[
\frac{\partial E}{\partial T_{pq}} = 0 \quad \frac{\partial E}{\partial S_{K0}} = 0
\]

#### Nonlinear set of equations

- Must be solved iteratively.
- Large arsenal of methods from numerical analysis, including
  1. The Newton-Raphson method.
The Newton-Raphson method in general

- We wish to find a stationary point of a function $E(p)$, where $p$ is a set of parameters that can be freely varied.
- Start with a guess, which for simplicity here we set to zero $p_0 = 0$.
- Expand $E$ through second order around this point

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

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

- $g$ 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 $E^{(2)}$.
- $\frac{\partial E^{(2)}}{\partial p_i} = 0 \rightarrow g + Hp = 0$ ($p = -H^{-1}g$)
- For this $p$, the new $g$ and $H$ are constructed,..
- Continue until convergence: $|g| \approx 0$.

Comments

- Approximating $E$ with $E^{(2)}$ → only valid for small $p$ → problems far from convergence.
- Converges quadratically when $p$ 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

\[
|\tilde{0}\rangle = e^{\hat{T}}e^{\hat{S}}|0\rangle
\]

\[
E = \langle\tilde{0}|\hat{H}|\tilde{0}\rangle = \langle0|e^{-\hat{S}}e^{-\hat{T}}\hat{H}e^{\hat{T}}e^{\hat{S}}|0\rangle
\]

The BCH expansion for operators

\[
e^{-\hat{A}}\hat{B}e^{\hat{A}} = B + [\hat{B},\hat{A}] + \frac{1}{2}[\hat{B},[\hat{A},\hat{A}]] + \cdots.
\]

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

\[
E^{(2)}(\hat{T}, \hat{S}) = \langle0|\hat{H}|0\rangle + \langle0|[\hat{H}, \hat{T}] + [\hat{H}, \hat{S}]|0\rangle
\]

\[
+ \langle0|\frac{1}{2}[[\hat{H}, \hat{T}], \hat{T}] + \frac{1}{2}[[\hat{H}, \hat{S}], \hat{S}] + [[\hat{H}, \hat{T}], \hat{S}]|0\rangle
\]
Optimization of MCSCF Wave Functions
Taylor Expansion of the MCSCF Energy

From last slide

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

The orbital gradient

\[ \hat{T} = \sum_{p>q} T_{pq} (E_{pq} - E_{qp}) = \sum_{p>q} T_{pq} E_{pq}^- \]

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

\[ g_{pq}^0 \text{ vanishes trivially for e.g inactive-inactive excitations.} \]

The Extended Brillouin Theorem for CASSCF wave functions

\[ \text{For the adjoint of a non-redundant rotations } E_{qp} | 0 \rangle = 0, \text{ so} \]

\[ g_{pq}^o = \langle 0 | [\hat{H}, (E_{pq} - E_{qp})] | 0 \rangle = 2 \langle 0 | \hat{H} E_{pq} | 0 \rangle. \]

\[ \text{At convergence the matrix elements between the wave-function and single excited states vanish.} \]
Optimization of MCSCF Wave Functions
The orbital gradient

Calculated using (yet) a Fock-matrix

- $g_{pq}^o = 2(F_{pq} - F_{qp})$.
- Most demanding part: active(v) $\rightarrow$ secondary(a).
- $F_{va} = \sum_w F_{wa}^l D_{vw} + \sum_{wxy} d_{vwxy} g_{awxy}$.
- Sum is over active orbitals.
- Requires integrals with 3 active indices.
- If integrals and density-matrices are known: scales as $N_{act}^4 N_{sec}$.
- No problem in constructing and storing.
Optimization of MCSCF Wave Functions
Taylor Expansion of the MCSCF Energy

From last slide

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

The CI gradient

- \[ \hat{S} = \sum_{K \neq 0} S_{K0} (|K\rangle \langle 0| - |0\rangle \langle K|), \langle K| 0 \rangle = 0. \]
- \[ \frac{\partial E}{\partial S_{K0}} = g^c_K = \langle 0 | ( [\hat{H}, |K\rangle \langle 0| - |0\rangle \langle K|] ) | 0 \rangle = 2 \langle K| \hat{H} | 0 \rangle \]
- \[ |K\rangle \text{ is orthogonal complement to } |0\rangle \text{ so: } g^c = 0 \rightarrow (1 - |0\rangle \langle 0|) \hat{H} | 0 \rangle = 0. \]
- \[ HC = EC. \]
- The optimal CI-coefficients constitute a solution to the standard CI eigenvalue problem.
Optimization of MCSCF Wave Functions
Taylor Expansion of the MCSCF Energy

From last slide

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

The Hessian: Partition into orbital (o) and CI (c) parts

\[ H = \begin{pmatrix} H^{cc} & H^{co} \\ H^{oc} & H^{oo} \end{pmatrix} \]
### Optimization of MCSCF Wave Functions

The Newton-Raphson Equations for MCSCF

#### Form

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

#### Number of elements in $H$ for medium scale calculation

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

\[
\begin{align*}
H^{oo} & : 4000 \times 4000 = 16 \times 10^6 \text{ - not problematic.} \\
H^{oc} & : 4000 \times 10^6 = 4 \times 10^9 \text{ - may be stored.} \\
H^{cc} & : 10^6 \times 10^6 = \times 10^{12} \text{ - difficult to store, (but is sparse).}
\end{align*}
\]

**Cannot be solved using standard methods like LU decomposition.**
Optimization of MCSCF Wave Functions

Solving the Newton-Raphson Equations for MCSCF

Form

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

To large to solve using standard decomposition methods.

Make approximations and/or use iterative methods

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

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

Further approximations Approximate \(H^{oo}\); done in the super-CI approach.
Optimization of MCSCF Wave Functions
Solving the Newton-Raphson Equations for MCSCF

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.
Optimization of MCSCF Wave Functions

Excited states

- **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).

P.R. Taylor and J. Olsen, Multiconfigurational Methods

ESQC-19
Thursday, Sept. 12

- MCSCF calculations on excited states
- Extensions of the CASSCF method
- Optimization of CASSCF wave functions
- The CASSI/RASSI method
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 determine transitions between the states, we must calculate the the *transition moment*: \( \langle X|\hat{\mu}|Y\rangle \)
- \( \hat{\mu} \) is for example the dipole operator: \( \hat{\mu} = \sum_{pq} \vec{\mu}_{pq} E_{pq} \).
- Transition matrix element is \( \langle X|\hat{\mu}|Y\rangle = \sum_{pq} D_{pq}^{XY} \vec{\mu}_{pq} \).
- \( D_{pq}^{XY} = \langle X|E_{pq}|Y\rangle \) is a transition density matrix.

P.R. Taylor and J. Olsen, Multiconfigurational Methods
RASSI/CASSI: The RASSCF (CASSCF) State Interaction Method

The Nonorthogonality Problem

- The calculation of the transition density matrix is easy when the two states are described in the same MO basis.
- The same holds when the MOs are not the same but they are biorthonormal: $\langle p^X | q^Y \rangle = \delta_{pq}$
- For CASSCF or RASSCF wave functions, the CI-coefficients may be counter-rotated, so a wave functions may be re-expressed in the biorthonormal basis.
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 $D_{XY}$ and $P_{XY}$.
- 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

\[
\begin{pmatrix}
H_{XX} - E1 & H_{XY} - ES_{XY} \\
H_{YX} - ES_{XY} & H_{YY} - E1
\end{pmatrix}
\begin{pmatrix}
c_X \\
c_Y
\end{pmatrix} = 0
\]

- The resulting states are orthogonal, and non-interacting through the Hamiltonian.
- Hundreds of XASSCF states can be handled.

P.R. Taylor and J. Olsen, Multiconfigurational Methods

ESQC-19
Contents

Friday, Sept. 13

- Introduction to dynamic correlation
- The multi-reference CI method
- Nearly size-extensive multi-reference methods
- Perturbation theory, general and multireference methods
- CASPT2
- NEVPT2
Introduction to Dynamical correlation

And it is

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

Characteristica

- Wave function is non-differentiable when the coordinates of two electrons are identical
- At these points, the convergence of the wave function in an orbital-based expansion is very slow.
- Using optimized orbitals does not help for the last %
Introduction to Dynamical correlation

The Coulomb cusp for He

- Left: One electron is fixed. The other electron moves in two-dim-plane.
- Right: One electron is fixed. The other electron moves in circle with the same radius as fixed electron.
Introduction to Dynamical correlation

**Methods to determine dynamic correlation**

**Single-reference (SR)** Based on excitations out from a reference state containing a single determinant/CSF

**Multi-reference (MR)** Based on excitations out from a reference state containing a several determinant/CSF

- **Variational** SRCl 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

The challenge have been to develop methods that accurate include both static and dynamic correlation

P.R. Taylor and J. Olsen, Multiconfigurational Methods
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, is not size-extensive
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. 13

- Introduction to dynamic correlation
- **The multi-reference CI method**
- Nearly size-extensive multi-reference methods
- Perturbation theory, general and multireference methods
- CASPT2
- NEVPT2

P.R. Taylor and J. Olsen, Multiconfigurational Methods
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 configuration state functions (CSFs). $i,j$ are occupied orbitals and $x,y$ occupied or external orbitals.
- Obtained from operators $\hat{E}_{xi}$, $\hat{E}_{xi}\hat{E}_{yj}$ operating on each $\Phi(I)$.
- The above set of CSFs is redundant - same CSFs obtained several times.
- E.g., $\text{H}_2$ 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

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

- Includes both dynamic and static correlation energy
- Linear expansion \( \rightarrow \) problems with size-extensivity

The parameters \( C \)

- Are determined using the variational principle
- Leads to the eigenvalue problem \( (H - E1)C = 0 \)
The Multireference CI Method

The Direct CI Method

**Idea**
- It is not possible to store the Hamiltonian matrix $H$
- Obtain instead selected roots using iterative methods

**Efficient direct CI methods require**
- Efficient numerical algorithms to obtain the lowest eigenvalues
- Efficient routines for the calculation of $HC$ from integrals

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

P.R. Taylor and J. Olsen, Multiconfigurational Methods
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 $\Psi_0$, instead of to the configurations in the reference space.
- The 'configurations' obtained in this way are much more complicated objects.
- There are much fewer coefficients to optimize: $C_{ij}^{xy}(I) \rightarrow C_{ij}^{xy}$, more-or-less independent of the number of CSFs in $\Psi_0$.
- 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 $\Psi_0$ during the iterations.
- Occasional problems, particularly with singles coefficients and these are sometimes then uncontracted.
- MOLPRO: internally contracted MRCI
- MOLCAS: no contraction in MRCI
- A number of new programs coming (BAGEL)

P.R. Taylor and J. Olsen, Multiconfigurational Methods
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. 15

- Introduction to dynamic correlation
- The multi-reference CI method
- Nearly size-extensive multi-reference methods
- Perturbation theory, general, MP2 and multireference methods
- CASPT2
- NEVPT2
Nearly size-extensivene multireference methods
Multireference ACPF and relatives

Idea
- Use MRCI form of wave function: \( \Psi = \Psi_0 + \Psi_a + \Psi_e \)
- \( \Psi_a \) is in reference space, \( \Psi_e \) contain the external excitations
- Modify energy-expression

The energy expression
\[
E = \frac{\langle \Psi_0 + \Psi_a + \Psi_e | H - E_0 | \Psi_0 + \Psi_a + \Psi_e \rangle}{1 + g_a \langle \Psi_a | \Psi_a \rangle + g_e \langle \Psi_e | \Psi_e \rangle}
\]
(Numerator is size-extensive)

\( g \rightarrow \) methods (\( N \) is number of electrons)
- **MRCI**: \( g_a = 1, \ g_e = 1 \).
- **MRACPF**: \( g_a = 1, \ g_e = 2/N \).
- **MRCEPA(0)**: \( g_a = 0, \ g_e = 0 \).
- **MRLCCM**: no \( \Psi_a, \ g_e = 0 \).
- **QDVPT**: \( g_a = 1, \ g_e = 0 \).

P.R. Taylor and J. Olsen, Multiconfigurational Methods
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 forms of corrections:
  \[ \Delta E_{\text{corr}} = (E_{\text{MRCI}} - E_{\text{REF}}) \left\{ 1 - \sum_R (c_{R}^{\text{MRCI}})^2 \right\} \]
  \[ \Delta E_{\text{corr}} = (E_{\text{MRCI}} - E_{\text{REF}}) \left\{ 1 - \sum_R (c_{R}^{\text{MRCI}} c_{R}^{\text{REF}}) \right\} \]
- Both trivial to compute.

**Problems**
- Weak theoretical foundation- but works often anyhow!!
Contents

Friday, Sept. 13

- Introduction to dynamic correlation
- The mult-reference CI method
- Nearly size-extensive multi-reference methods
- Perturbation theory, general and multireference methods
- CASPT2
- NEVPT2
Multi-reference perturbation theory
Rayleigh-Schrödinger Perturbation Theory

### Split Hamiltonian

\[ \hat{H} = \hat{H}_0 + \lambda \hat{V}. \]
\[ \hat{H}_0 |0_0\rangle = E_0 |0_0\rangle \]

### First-order correction to wave function and second-order correction to energy

\[ |0_1\rangle = \sum_{\mu} C_{\mu}^1 |\mu\rangle, \langle \mu |0_0\rangle = 0 \]
\[ C_{\mu}^1 = -\sum_{\nu} (\hat{H}_0 - E_0 S)^{-1}_{\mu \nu} \langle \nu | \hat{V} |0_0\rangle \]
\[ (\hat{H}_0)_{\mu \nu} = \langle \mu | \hat{H}_0 |\nu\rangle, S_{\mu \nu} = \langle \mu |\nu\rangle \]
\[ E_2 = \langle 0_1 | \hat{V} |0_0\rangle \]
Multi-reference perturbation theory

Various forms

Many forms, differs by choices of

1. Reference state
2. Form of $\hat{H}_0$
3. Form of wave function corrections
4. (Use of intermediate Hamiltonian, buffer states ...)

Reference state

- CASSCF/RASSCF/GASSCF...
- Incomplete spaces

Is usually not convergent, but low-order corrections (2+3) are useful

P.R. Taylor and J. Olsen, Multiconfigurational Methods
Multi-reference perturbation theory

Various forms

Form of $\hat{H}_0$

- Fock-type one-electron operator in all orbital spaces (CASPT)
- Fock-type one-electron operator in inactive and secondary spaces, full two-electron Hamilton operator in active space (NEVPT)

Form of wave function corrections

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

Friday, Sept. 13

- Introduction to dynamic correlation
- The multi-reference CI method
- Nearly size-extensive multi-reference methods
- Perturbation theory, general and multireference methods
- CASPT2
- NEVPT2
### Reference state, $|0_0\rangle$

- CAS state - with or without optimized orbitals

### Form of first-order correction

- $|0_1\rangle = \sum_{\mu} C^1_{\mu} O_{\mu} |0_0\rangle$
- $O_{\mu}$ contains the single and double excitations from $|0_0\rangle$
- Only SD space in first-order correction: interacts with $|0_0\rangle$

### 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
- Projections are needed for a zero-order eigenvalue equation
- $\hat{P}_0$ projects on $|0_0\rangle$, $\hat{P}_K$ projects on other states in the CASpace, $\hat{P}_{SD}$ projects on single and doubles, $\hat{P}_X$ on the rest.
CASPT2
The Fock Operator, simplest choice

Defined as

\[ \hat{f} = \sum_{pq} f_{pq} E_{pq} \]

\[ f_{pq} = \sum_{\sigma} \langle 0_0 | [a_p \sigma, [H, a_q^{\dagger} \sigma]_+] | 0_0 \rangle = h_{pq} + \sum_{rs} D_{rs} [(pq|rs) - \frac{1}{2}(pr|qs)]. \]

Becomes standard MP \( \hat{H}_0 \) if |0\rangle is HF wave function

Properties of \( f \)

<table>
<thead>
<tr>
<th>Inactive orbitals</th>
<th>( f_{pp} = -IP_p )</th>
</tr>
</thead>
<tbody>
<tr>
<td>External orbitals</td>
<td>( f_{pp} = EA_p )</td>
</tr>
<tr>
<td>( n_p^{occ} = 1 ):</td>
<td>( f_{pp} = \frac{1}{2}(-IP_p + EA_p) )</td>
</tr>
</tbody>
</table>

Unpaired electrons are favoured, so binding energies are too low.
**CASPT2**

**Fock matrix Elements and the First-Order Equation**

**The first-order equation**

- \((H_0 - E_0 S)C^1 = -V\)

- Eight types of double excitations

- Equations may be split into 8 blocks, one for each of 8 excitation types

- The coupling of the blocks is introduced in a second iterative step.

**Fock matrix Elements, \(H_0\)**

- Contain up to fourth-order density matrices:
  \[
  \langle pqrs | \hat{F} | p'q'r's' \rangle = \sum mn f_{mn} \langle 00 | \hat{E}_{sr} \hat{E}_{qp} \hat{E}_{mn} \hat{E}_{p'q'} \hat{E}_{r's'} | 00 \rangle
  \]

- May be rearranged to require three-body density

---

P.R. Taylor and J. Olsen, Multiconfigurational Methods

ESQC-19
CASPT2
Computational steps in a CASPT2 calculation

1. Perform a CASSCF calculation, single-state or state-average.
2. Determine orbitals that block-diagonalizes the Fock matrix.
3. Transform integrals to obtain 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 $H_0$.
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.
- A large coefficient may arise from artifacts of $\hat{H}_0$
- This is the *intruder state* problem in CASPT2.
- May be removed using level-shifting and/or other choices of $\hat{F}$
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 Hamiltonians: $(H_{eff})_{ij} = \delta_{ij}E_i + \langle \Psi_i | \hat{H} | \chi_j \rangle$
- Recommended for treating several states of one symmetry.
Advantages of CASPT2

- A CAS wave function is the reference so static correlation may be included.
- 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.
- May be extended to RAS and GAS wave functions.
- Requires often tuning of the level-shift parameter.
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).
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...?)
### N₂ ground-state spectroscopic constants

<table>
<thead>
<tr>
<th>Method</th>
<th>$r_e$ (Å)</th>
<th>$\omega_e$ (cm⁻¹)</th>
<th>$D_e$ (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASSCF</td>
<td>1.102</td>
<td>2329</td>
<td>8.91</td>
</tr>
<tr>
<td>CASPT2</td>
<td>1.100</td>
<td>2312</td>
<td>9.43</td>
</tr>
<tr>
<td>ACPF</td>
<td>1.099</td>
<td>2330</td>
<td>9.79</td>
</tr>
<tr>
<td>CCSD(T)</td>
<td>1.099</td>
<td>2332</td>
<td>9.82</td>
</tr>
<tr>
<td>Expt</td>
<td>1.098</td>
<td>2359</td>
<td>9.90</td>
</tr>
</tbody>
</table>
Contents

Friday, Sept. 13

- Introduction to dynamic correlation
- The multi-reference CI method
- Nearly size-extensive multi-reference methods
- Perturbation theory, general and multireference methods
- CASPT2
- NEVPT2
NEVPT2

Use Dyall’s zero-order operator

\[ \hat{H}_0 = \sum_{ij} f_{ij} E_{ij} + \sum_{ab} f_{ab} E_{ab} + \sum_{vw} h_{vw} E_{vw} + \]
\[ \frac{1}{2} \sum_{vwxy} g_{vwxy} (E_{vw} E_{xy} - \delta_{wx} E_{vy}) \]

- Fock-operator in inactive (indices ij) and secondary orbitals (indices ab)
- Full Hamiltonian in active orbitals (indices vwxy)
- Requires five-body terms over active orbitals.

Expansion of first-order wave-function

- Advantageous to use internal contraction
- Due to the accurate treatment of interactions between active orbitals, the CAS CI coefficients may be re-optimized
- Overall formalism very similar to the CASPT formalism.
NEVPT2

Pro and cons compared to CASPT-theoretical aspects

- More complicated and involved computations due to the 2-electron term over active orbitals - Not decisive these days.
- Treats the orbitals given rise to static correlation accurately.
- But treats the different types of orbitals differently.
- Allows relaxation of the CAS-coefficients
- Less developed codes (no multistate version?, No RAS/GAS extensions).
- zero-order eigenvalue equation automatically fulfilled.
- Rigourously size-extensive.
- No (very few) intruder states.

P.R. Taylor and J. Olsen, Multiconfigurational Methods
Comparing RASPT2(12,22) to DMRG-CASPT2(12,28), the difference can be situated in the extra orbitals included in the active space or the restrictions to the excitation level that are applied in the RAS reference compared to the DMRG reference wave function. Unfortunately, there is no DMRG-CASPT2-(12,22) data available, which could give more insight into the origin of the discrepancy and rule out an effect of the extra 6 active orbitals. Although DMRG-NEVPT2(12,22) uses the same active space as RASPT2(12,22), the difference is again not only limited to the RAS restrictions but also the type of zeroth-order Hamiltonian. This means that with the current data, we cannot pinpoint the origin of the much larger dissociation energy obtained with RASPT2 compared to either DMRG-CASPT2 or DMRG-NEVPT2.

Compared to experiment, both DMRG-based methods give the best agreement, within 0.13 eV of the experimental value from Simard et al. RASPT2(12,22) on the other hand overestimates the stability of the chromium−chromium bond compared to experiment by 0.33 eV. We should note at this point that the experimental value of 1.56 eV we chose for comparison is the most recently obtained, and other values for the dissociation energy are available as well. For example, when taking 1.47 eV instead as the experimental reference, it is DMRG-NEVPT2(12,22) and MR-AQCC that have the best agreement, both within 0.12 eV from the experimental $D_e$. Guo et al. also investigated the effect of including semicore correlation already in the reference wave function by performing a set of DMRG-NEVPT2(28,20) calculations by adding to the original (12,12) active space the 3s and 3p orbitals of each chromium atom. Curiously, even though a large upward shift of the PES by almost 0.6 eV compared to DMRG-NEVPT2(12,12) is observed, the authors conclude that the effect is not as important as the double shell effect. We tried to reproduce this by performing a RASPT2(28,20) where the 3s and 3p orbitals were added to RAS1, allowing up to double excitations into the (12,12) RAS2 space. However, we instead found that the effect was negligible. We should note at this point that the SC-NEVPT2 method, used for the DMRG-NEVPT2 calculations, is the so-called "strongly contracted" variant, where the first-order wave function is expanded in a smaller space.

### Table 2. Equilibrium Distance ($r_e$), Dissociation Energy ($D_e$), and Harmonic Vibration Wavenumber ($\omega_e$) for the Ground State of $\text{Cr}_2$

<table>
<thead>
<tr>
<th>Method</th>
<th>$r_e$ (Å)</th>
<th>$D_e$ (eV)</th>
<th>$\omega_e$ (cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DMRG-CASPT2(12,28), CBS(QZ/5Z)</td>
<td>1.658</td>
<td>1.69</td>
<td>565</td>
</tr>
<tr>
<td>DMRG-NEVPT2(12,22), CBS(QZ/5Z)</td>
<td>1.651</td>
<td>1.45</td>
<td>423</td>
</tr>
<tr>
<td>RASPT2(12,22), CBS(QZ/5Z)</td>
<td>1.666</td>
<td>1.89</td>
<td>489</td>
</tr>
<tr>
<td>CASPT2(12,12), CBS(QZ/5Z)</td>
<td>1.681</td>
<td>1.85</td>
<td>770</td>
</tr>
<tr>
<td>MR-AQCC CBS(TZ/QZ)</td>
<td></td>
<td>1.35</td>
<td>453</td>
</tr>
<tr>
<td>Experiment</td>
<td></td>
<td>1.56$^\pm$0.06</td>
<td>480.6</td>
</tr>
</tbody>
</table>

$a$ $r_e$ and $D_e$ were obtained from fitting second-order interpolating splines through the single-energy data points. $\omega_e$ was obtained using the VIBROT program, using 10 vibrational levels, providing the spline functions as PES. Details of the analysis can be found in the Supporting Information.
The theoretical advances of the NEVPT formalism do not in general lead to improved accuracy. CASPT has a tendency to overshoot correlation effects, which often leads to very good results in smaller basis sets. For complicated transition metal systems, the current advice is to stick to the CASPT method. I am not sure that the final verdict on the relative merits has been made.