Multireference Configuration Interaction: Methodological Aspects and Applications

Hans Lischka
University of Vienna
The concept of a reference wave function

A simple example: $\text{H}_2 (^1\Sigma_g^+)$

In a minimal basis we have two orbitals - $\sigma_g$ and $\sigma_u$

Ground state MO scheme

At $R \approx R_{\text{min}}$ configuration I will be a qualitatively good representation of the exact wave function. However, for $R \rightarrow \infty$ the second orbital, $\sigma_u$, will become degenerate or quasi-degenerate and configuration I will not describe the system well. Within the given symmetry a second configuration $\sigma_u^2$ has to be taken into account.
Generalization to more electrons

Single reference (SR) case (closed shell)

Reference wave function $\Psi_0$:

$$\Psi_{HF} = |\Phi_1 \Phi_1 \cdots \Phi_n \Phi_n|$$

\[\begin{array}{c}
\vdots \\
\_ \quad \text{virtual} \\
\_ \quad \text{orbitals} \\
\_ \\
\_ \\
\_ \\
\_ \\
\_ \\
\_ \quad M \text{ doubly occupied} \\
\_ \quad \text{Orbitals} \\
\_ \\
\_ \quad M = n
\end{array}\]
Construction of the SR wave function

excitation (substitution) of occupied orbitals by virtual ones

\[
|\Phi_1 \Phi_1 \ldots \Phi_i \Phi_i \ldots \Phi_n \Phi_n| \quad \downarrow \quad \Phi_i \rightarrow \Phi_a
\]

\[
|\Phi_1 \Phi_1 \ldots \Phi_a \Phi_i \ldots \Phi_n \Phi_n|
\]

Single-, double-, triple- … m-tuple excitations

\[
\Psi_i^a, \Psi_{ij}^{ab}, \Psi_{ijk}^{abc} \ldots
\]

Method of configuration interaction (CI):

\[
\Psi_{SR-CI} = c_0 \Psi_0 + \sum_{i,a} c_i^a \Psi_i^a + \sum_{i,j,a,b} c_{ij}^{ab} \Psi_{ij}^{ab} + \ldots
\]

Variation principle (Ritz)

Goal: **Full CI** – inclusion of all possible configurations
Multireference (MR) case

\[ \Psi_{\text{MR-CI}} = \sum c_i \Psi_i \]

Reference wave function \( \Psi_0: \Psi_1 \ldots \Psi_{N_{\text{ref}}} \) m-tuple excitations from \( \Psi_i, i = 1 \ldots N_{\text{ref}} \) into the virtual orbitals creates a set of configurations \( \{\Psi_i\} \)

Select reference wave functions in order to include all quasi-degeneracies
MR-CI approach – general considerations

Stationary, nonrelativistic, clamped-nuclei, electronic Schrödinger equation

\[ H \Psi = E \Psi \]

\[ H = \sum_i h_i + \sum_{i<j} g_{ij} + V_{nn} \]

\[ h_i = -\frac{1}{2} \nabla_i^2 - \sum_A Z_A / R_{Ai} \]

\[ g_{ij} = \frac{1}{r_{ij}} \]

\[ V_{nn} = \sum_{A<B} Z_A Z_B / R_{AB} \]

Variational principle

\[ E[\Psi] = \langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle \]

The linear ansatz for the trial function \( \Psi \)
(Ritz variation principle)

\[ \Psi = \sum_{s=1}^n c_s \Phi_s \]
leads to the following matrix eigenvalue problem

\[ \mathbf{H} \mathbf{c}_i = E_i \mathbf{c}_i, \quad i = 1 \ldots n \]

with

\[ H_{st} = \langle \Phi_s | H | \Phi_t \rangle \]

and assuming orthonormality of the \( \Phi \)'s, i.e.

\[ \langle \Phi_s | \Phi_t \rangle = \delta_{st} \]

Energies are ordered as

\[ E_1 \leq E_2 \leq \ldots \leq E_n \]

Each eigenvalue \( E_p \) is an upper bound to the corresponding exact eigenvalue. As additional terms are added to the expansion each eigenvalue \( E_p^{n+1} \) of the \((n+1)\)-term expansion satisfies the inequalities

\[ E_p^n \leq E_p^{n+1} \leq E_p^n \]

General references:
For practical calculations we have to make several decisions concerning
• the representation of the many-electron basis functions $\Phi_s$
  • Slater determinants
  • Configuration state functions (CSFs) – eigenfunctions of $S^2$
• How to calculate the Hamiltonian matrix $H$ (calculate it at all?). In most applications the dimension of $H$ will be too large for explicit storage.

**Davidson subspace method for diagonalization**

Davisdon 1975, 1993
Algorithm:

A. If the $k$th eigenvalue is wanted, select a zeroth-order orthonormal subspace $v_1, v_2, \ldots v_l$ ($l \geq k$) spanning the dominant components of the first $k$ eigenvectors. Form $Hv_1 \ldots Hv_l$ and $(v_i, Hv_j) = \tilde{H}_{ij}, 1 \leq i \leq j \leq l$. Diagonalize $\tilde{H}$ using a standard method, select the $k$th eigenvalue $\lambda_k^{(l)}$ and the eigenvector $\alpha_k^{(l)}$.

B. Form the residual vector

$$q_M = \sum_{i=1}^{M} \alpha_{i,k}^{(M)} (Hv_i) - \sum_{i=1}^{M} \alpha_{i,k}^{(M)} \lambda_k^{(M)} v_i$$

$M$ is the dimension of $\tilde{H}$.

C. Form $\|q_M\|$ and check for convergence.

D. Form $d_{I,(M+1)} = \left(\lambda_k^{(M)} - H_{II}\right)^{-1} q_{I,M}, I = 1 \ldots N$

(and orthogonalize to the $v_i, i = 1 \ldots M$).

E. $v_{(M+1)} = d_{(M+1)} / \|d_{(M+1)}\|$.

F. Form $Hb_{(M+1)}$

G. Form $\tilde{H}_{i,M+1} = (v_i, Hv_{(M+1)}), i = 1 \ldots M + 1$.

H. Diagonalize $\tilde{H}$ and return to step B with $\alpha_k^{(M+1)}$ and $\lambda_k^{(M+1)}$. 
The main computational work is the matrix-vector multiplication $\mathbf{w} = \mathbf{Hv}$.

Options:

- Compute $\mathbf{H}$, store it on disk, and compute $\mathbf{Hv}_i$ afterwards – conventional CI
- We do not need to construct $\mathbf{H}$, we only need $\mathbf{w}_i = \mathbf{Hv}_i$! This is the basis of the “direct CI” (Roos 1972, Roos 1977)

Matrix elements

$$
\langle \Phi_s | H | \Phi_i \rangle = \sum_{ij} a_{ij}^{st} h_{ij} + \sum_{ijkl} b_{ijkl}^{st} g_{ijkl}
$$

Orbital integrals $h_{ij}$ and $g_{ijkl}$ are defined by

$$
h_{ij} = \langle \varphi_i | h_1 | \varphi_j \rangle \quad \text{and} \quad g_{ijkl} = \langle \varphi_i \varphi_j | g_{12} | \varphi_k \varphi_l \rangle
$$

$$
\mathbf{w}_s = \sum_t \sum_{ij} a_{ij}^{st} h_{ij} \mathbf{v}_t + \sum_t \sum_{ijkl} b_{ijkl}^{st} g_{ijkl} \mathbf{v}_t
$$
The coupling elements are simple (-1 or +1) in case of single determinants (Slater-Condon rules). For CSFs several possibilities exist (Shavitt 1977):

- Spin projection
- Symmetric group
- Unitary group

**Graphical) Unitary Group Approach (G)UGA**

We are using Gelfand states

Hamiltonian: 

\[ H = \sum_{ij} h_{ij} E_{ij} + \sum_{ijkl} (ij|kl) e_{ij,kl} \]

where \( X^+_{i\sigma} \) and \( X_{j\sigma} \) are spinorbital creation and annihilation operators, \( \sigma = \alpha, \beta \), and the sum \( E_{ij} = \sum_{\sigma} X^+_{i\sigma} X_{j\sigma} \) is over the spin.

\[ e_{ijkl} = E_{ij} E_{kl} - \delta_{kj} E_{il} \]
A matrix element of $H$ between two Gelfand states $\langle m' |$ and $| m \rangle$ is given as

$$H_{m,m'} = \langle m' | H | m \rangle = \sum_{ij} h_{ij} \langle m' | E_{ij} | m \rangle + \sum_{ijkl} (ij|kl) \langle m' | e_{ij,kl} | m \rangle$$

Thus, the matrix elements of the one-body and two-body unitary group operators are the coupling coefficients of the corresponding one- and two-electron integrals, respectively.
GUGA

• This presentation describes the **distinct row table (DRT)** and its graphical representation in the graphical unitary group approach (GUGA). Further details can be found, e.g., in:


  • and references therein.

• GUGA slides taken from the contribution of I. Shavitt presented at the COLUMBUS Programmer's Workshop (http://www.univie.ac.at/columbus/workshops/argonne2005/), Argonne National Laboratory, Chicago, USA, August 15th-19th 2005,
Graphical representation of a DRT

6 orbitals
5 electrons
$S = 1/2$

The circles represent the nodes of the graph

The directed walks through the DRT represent all the configuration state functions
What possibilities do we have to connect one node with another one above or below?
There are only 4 possibilities!

Different step numbers $d_i$ are represented by arcs of different slope:
Characterization of a configuration by a step vector: $d_1d_2d_3d_4d_5d_6$
e.g.: $1 \ 0 \ 3 \ 1 \ 0 \ 2$
The step vector characterizes the orbital occupation
**Step Vector**

<table>
<thead>
<tr>
<th>Orbital</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>d_i</td>
<td>1 0</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>so</td>
<td>uo</td>
<td>so</td>
<td>uo</td>
<td>so</td>
<td></td>
</tr>
</tbody>
</table>

- \( d_i = 0 \Rightarrow \text{orbital } i \text{ is unoccupied (uo)} \)
- \( d_i = 1,2 \Rightarrow \text{orbital } i \text{ is singly occupied (so)} \)
- \( d_i = 3 \Rightarrow \text{orbital } i \text{ is doubly occupied (do)} \)

Example:

- Orbital: 1 2 3 4 5 6
- Step Vector: 1 0 3 1 0 2
- Interpretation: so uo do so uo so
Multireference CI singles and doubles

Reference configurations in red
Inactive (reference doubly occupied)
Active (variable occupation in the reference)
External (virtual) orbitals: only singles and doubles
• The internal (active + closed shell) part of the graph is complicated, but relatively small in comparison to the virtual (external) space.
• The graph for the external is space is very simple due to the fact that we allow only double excitations. Moreover, its structure is independent of the internal part. Respective loops (coupling elements) can be computed once and for all.
• The interface between internal and external space is given by the vertices Z (0-excitations), Y (single-excitations), X (double excitations, triplet coupling) and W (double excitations, singlet coupling)
• The loops are split into an internal part and into an external one. The internal part is computed explicitly and either stored on a file (formula file) or recomputed every Davidson iteration. The external part is added “on-the-fly” when the total coupling elements are computed.

• The two-electron integrals are sorted according to the number of internal indices: all (four)-internal, three-internal, two-internal, one internal and all-external.
Algorithm for $w=Hv$

$$w_s = \sum_t \sum_{ij} a_{ij}^{st} h_{ij} v_t + \sum_t \sum_{ijkl} b_{ijkl}^{st} g_{ijkl} v_t$$

In this approach (COLUMBUS program, Lischka 1981, Shepard 1988, Lischka 2001, 2006) the computational scheme is driven by the integral indices

Loop over types of two-electron integrals
   For each set
      Loop over the quadruple of indices
         Get coupling elements, integrals and $v$
         Compute the respective contribution to $w$
         Update $w$
      End loop indices
   End loop integral types

The remaining steps of the Davidson iteration are simple.
The update of $\mathbf{w}$ can be performed via matrix and vector operations if the whole virtual space if the same virtual orbital space is used for excitations from all reference configurations. In this case selection schemes can only be applied at the level of the reference configurations.

Individual selection schemes specific for each reference configurations (e.g. multireference double excitation CI (MRD-CI) Buenker 1968, 1974, 1975, Hanrath 1997) do not have this restriction.
Overall CI scheme

- AO integrals
- Determination of MOs by means of a SCF or MCSCF calculation
- Transformation of the two-electron integrals into the MO basis
- Sorting of the integrals
- **Davidson diagonalization**
- Properties
- Energy gradient and nonadiabatic couplings
- Geometry optimization routines
Overview of different CI approaches

• Selection at the reference configuration level
  – **Uncontracted CI**: single- and double excitations are constructed from each reference CSF and the unique set of configurations ($\Psi_I, I = 1 \ldots N_{CI}$) is taken. The CI coefficient of each CSF is included in the variational procedure.
  – **Contracted MRCI (IC-MRCI)** (Meyer 1977, Werner 1988):

\[
\Psi_{ijp}^{ab} = \frac{1}{2} \left( E_{ai,bj} + pE_{bi,aj} \right) \Psi_0
\]

($p = \pm 1$ for external singlet and triplet pairs)

\[
\Psi_0 = \sum_{R\mu} a_{R\mu} \Psi_{R\mu}
\]

The configurations $\Psi_{ijp}^{ab}$ span exactly the first-order interacting space
## Uncontracted vs. contracted CI

### Uncontracted CI

**Advantages:**
- Larger flexibility of the wave function, full variation, excited states!
- Computation of analytic gradient relatively easy (available in COLUMBUS)

**Disadvantage:**
- Much larger computational effort

### Contracted CI

**Advantages:**
- High computational efficiency

**Disadvantages:**
- Relaxation effects for $\Psi_0$
- Potential energy surfaces for excited states, avoided crossings,…
- Analytic gradient difficult (not available)
Overview of MRD-CI

Original work: Buenker 1968, 1974, 1975

- Selection scheme based on perturbation theory – energy threshold $T$
- Extrapolation scheme $T \to 0$

Conventional CI program

- Hanrath 1997:
  DIESEL-MR-CI (direct internal external separated individually selecting MR-CI): direct, division of internal/external space

Advantage of MRD-CI: great flexibility, relatively large molecules

Disadvantage: Extrapolation, gradient for extrapolated energy
When do we use which method?

• MR-CI, MR-CI+Q, MR-AQCC (Szalay 1993, 1995) in multireference cases, otherwise use single reference method

• When geometry optimization at the CI level is not required or when pointwise optimization is possible – IC-MRCI or MRD-CI

• Difficult cases: robust method is required for description of large portions of the energy surfaces – uncontracted MR-CI or MR-AQCC

Special feature: analytic energy gradients and nonadiabatic couplings (conical intersections)
Characterization of Internal Orbital spaces

Active orbitals: usually as CAS
Auxiliary orbitals: for the description of Rydberg states (single excitations or individual configurations)
Selection of CSF space

Three options:

1. a) The reference configurations are chosen to have the same symmetry as the state symmetry
   b) Only those singles and doubles are constructed, which have a nonzero matrix element with one of the references (interacting space restriction (Bunge 1970))

2. Restriction 1b is lifted

3. All possible reference configurations are constructed within the specified orbital-occupation restrictions. Singles and doubles from reference with "wrong" symmetry can have the correct symmetry.

Condition 3 allows consistent calculations using different subgroups of the actual molecular symmetry.
COLUMBUS project

- Set of programs for high-level ab initio calculations
- Methods: MCSCF, MR-CISD, MR-ACPF/AQCC, Spin-orbit CI
- Focus: multireference calculations on ground and excited states
- Recent achievements: analytic MR-CI gradient, nonadiabatic couplings, parallel CI
- COLUMBUS, an ab initio electronic structure program, release 5.9.1 (2006).
- Web page: http://www.univie.ac.at/columbus/
COLUMBUS program package


- **User-friendly interface**
- Public domain – free of charge
- Distribution of source code
References