

CHAPTER 10

Multireference wave function methods

In this chapter we will discuss methods can be used to treat bond break process and electronically excited states.

10.1 Multiconfigurational self-consistent-field theory

As we have seen previously, in many cases the Hartree–Fock method is usually captures a large fraction of the total electronic energy. This is usually true of closed-shell molecules and high-spin open-shells. However, we have also encountered cases, like for example, the H_2 molecule the dissociation limit, where a single determinant wave function fails to capture even the correct qualitative behavior of the electrons.

In this section we will discuss multiconfigurational self-consistent-field (MCSCF) methods, which may be viewed as generalizations of Hartree–Fock theory. In MCSCF methods, an electronic state is approximated with a small configuration interaction wave function

$$|\Psi_{\text{MCSCF}}\rangle = \sum_I C_I |\Phi_I\rangle, \quad (10.1)$$

in which both the orbitals (ϕ_i) from which the determinants (Φ_I) and the determinant coefficients (C_I) are variationally optimized. Note, that in general, that they determinants employed to expand the MCSCF wave function (understood as occupation patterns of electrons) are not changed in the optimization process.

The MCSCF energy is obtained by performing a simultaneous optimization of the orbitals and the determinant coefficients

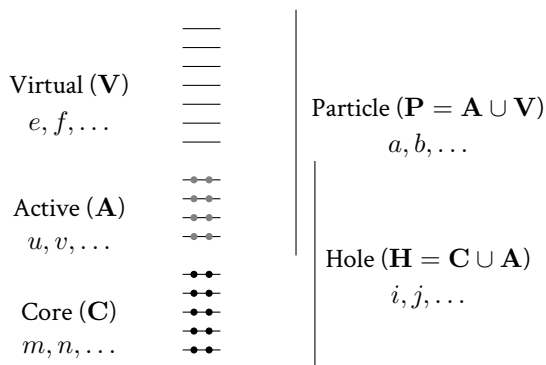
$$E_{\text{MCSCF}} = \min_{C_I, \{\psi_i\}} \langle \Psi_{\text{MCSCF}} | \hat{H} | \Psi_{\text{MCSCF}} \rangle, \quad (10.2)$$

imposing that the determinant vector is normalized ($\sum_I |C_I|^2 = 1$) and the orbitals are orthonormal. It is always implicitly assumed that the MCSCF wave function is constructed using restricted spin orbitals so that the alpha and beta spatial parts of a given orbital are the same.

Different MCSCF methods are defined by the way the determinant are selected. The most widely used approach is the complete-active-space self-consistent-field (CASSCF) method, which basically forms all the MCSCF determinant by performing a FCI in a limited number of orbitals. In CASSCF one starts by partitioning the orbitals into three spaces:

- **Core (or doubly occupied) orbitals (C).** These orbitals are generally the lowest one in energy and are kept doubly occupied in all the determinants in the CASSCF wave function.

- **Active orbitals (A).** These orbitals have variable occupation in the CASSCF wave function. In most cases these are chosen to be valence orbital associated with the particular aspect we want to model. For example, for to describe the wave function for the single bond-breaking reaction $A-B \longrightarrow A+B$ it is sufficient to include a single pair of $A-B$ bonding/antibonding orbitals.
- **Virtual orbitals (V).** These orbitals are not occupied in the CASSCF wave function. These are usually high-energy orbitals that are typically irrelevant to description of a chemical reaction or an excited state.



In CASSCF one generates determinants with the core orbitals occupied and n_A active electrons distributed in the active orbitals. Therefore, a generic CASSCF determinant has the form

$$|\Phi_I\rangle = \underbrace{\hat{a}_u^\dagger \hat{a}_v^\dagger \cdots \hat{a}_z^\dagger}_{\text{active}} \prod_m^C \hat{a}_m^\dagger | \rangle. \quad (10.3)$$

In the literature, and active space in which a given number of electrons (n_{el}) is distributed in a given number of active orbitals (n_A) is abbreviated with $\text{CAS}(n_{\text{el}}, n_A)$. For example, an active space with two electrons in two orbitals is denoted $\text{CAS}(2e, 2o)$.

In CASSCF, the space of determinants is **complete**, in the sense that all possible determinants are included. One consequence of this choice is that the wave function is invariant with respect to unitary transformations of the active orbitals. This means that if we transform the active orbitals via a unitary transformation \mathbf{U}

$$\psi_u \rightarrow \psi'_u = \sum_v \psi_v U_{vu}, \quad (10.4)$$

and consequently all the determinant change according to

$$|\Phi_I\rangle = |\psi_{i_1} \psi_{i_2} \cdots \psi_{i_N}\rangle \rightarrow |\Phi'_I\rangle = |\psi'_{i_1} \psi'_{i_2} \cdots \psi'_{i_N}\rangle, \quad (10.5)$$

there is a corresponding set of CI coefficients C'_I such that the transformed CASSCF wave function is identical to the original one

$$\sum_I C_I |\Phi_I\rangle = \sum_I C'_I |\Phi'_I\rangle. \quad (10.6)$$

This property implies that we are free to work with any unitarily transformed version of the CASSCF active orbitals (e.g., delocalized, localized).

Note that since the CAS space includes all possible determinants, its size grows very quickly with the number of active orbitals and electrons. If we ignore point group symmetry, we can compute the number of determinants with a given number of alpha (n_α) and beta (n_β) electrons in the active orbitals just by counting how many arrangements of electrons are possible for the two spin cases. For each spin case, the number of ways we can distribute n electrons is

given by the number of ways we can choose n spin orbitals to occupy and taking into account that permutations are equivalent. This quantity is given by the binomial coefficient

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}, \quad (10.7)$$

where $n! = 1 \times 2 \times \dots \times n$, and $0! = 1$ by definition. The total number of determinants (d) is given by the product of the number of arrangements of alpha and beta electrons

$$d = \binom{n_{\mathbf{A}}}{n_{\alpha}} \binom{n_{\mathbf{A}}}{n_{\beta}}. \quad (10.8)$$

As shown in Tab. 10.1, this quantity grows very quickly, and therefore, practical computations are typically limited to active spaces of size CAS(16e,16o).

Size of the active space		
Active electrons	Active orbitals	CAS determinants
2	2	4
4	4	36
6	6	400
8	8	4,900
10	10	63,504
12	12	853,776
14	14	11,778,624
16	16	165,636,900
18	18	2,363,904,400
20	20	34,134,779,536

Table 10.1: Number of determinants in a CAS space at half filling. These numbers ignore the reduction in size due to point group symmetry.

Note that one can also define a corresponding CASCI approach in which one generated a CAS space and performs the variational optimization of the CI coefficients **without** optimizing the orbitals.

10.2 Multireference configuration interaction

The CASSCF method can only account for static correlation effects, which in practice is not enough to accurately describe multideterminantal electronic states. The simplest way to include dynamical correlation effects on top of a CASSCF reference is via a configuration interaction wave function. In the multireference configuration interaction (MRCI) approach, one starts with a CASSCF reference and adds excited determinants that account for dynamical correlation effects. The simplest approach is **uncontracted** MRCI with singles and doubles (u-MRCISD), in which one simply augments the CASSCF determinant space with all singly and doubly excited determinants of each reference determinant. The u-MRCISD wave function is given by

$$|\Psi_{\text{u-MRCI}}\rangle = \sum_I^{\text{CAS}} C_I |\Phi_I\rangle + \sum_A^{\text{EXT}} C_A |\Phi_A\rangle, \quad (10.9)$$

where the space of external determinants is generated by applying the operators $\hat{a}_a^\dagger \hat{a}_i$ and $\hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i$ to all the CAS determinants

$$\text{EXT} = \{ \Phi_A | \Phi_A = \hat{a}_a^\dagger \hat{a}_i |\Phi_I\rangle, \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i |\Phi_I\rangle \forall \Phi_I \in \text{CAS} \}. \quad (10.10)$$

In defining the EXT space, the orbital indices i, j run over the core and active orbitals, while a, b run over the active and virtuals. The u-MRCISD method provides very accurate results, but it has a computational scaling that becomes quickly unmanageable when the number of CAS determinants is large. The approximate number of u-MRCISD determinants depends on the number of hole (N_H) and particle (N_P) orbitals and is proportional to

$$N_{\text{det}}^{\text{u-MRCISD}} = d N_H^2 N_P^2, \quad (10.11)$$

which is limited both by the size of the active space (d) and the number of electrons and basis functions used.

A more economical alternative to the u-MRCISD method is the **internally-contracted** MRCISD (ic-MRCISD) method. In ic-MRCISD, one still generates all possible singles and doubles out of a CAS; however, an excitation operator is applied to the entire CAS reference all at once. For example, internally-contracted singly excited configurations are defined as

$$|\Psi_i^a\rangle = \hat{a}_a^\dagger \hat{a}_i |\Psi_{\text{CAS}}\rangle = \sum_I C_I \hat{a}_a^\dagger \hat{a}_i |\Phi_I\rangle. \quad (10.12)$$

The right hand side of this equation shows that $|\Psi_i^a\rangle$ is a linear combination of excited determinants ($\hat{a}_a^\dagger \hat{a}_i |\Phi_I\rangle$) with coefficients fixed to the value of the CAS wave function (C_I).

The ic-MRCISD wave function is defined by the following wave function guess

$$\begin{aligned} |\Psi_{\text{ic-MRCISD}}\rangle &= (1 + \hat{C}) |\Psi_{\text{CAS}}\rangle \\ &= \sum_I C_I |\Phi_I\rangle + \sum_i \sum_a c_i^a \hat{a}_a^\dagger \hat{a}_i |\Psi_{\text{CAS}}\rangle + \frac{1}{4} \sum_{ij} \sum_{ab} c_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i |\Psi_{\text{CAS}}\rangle \\ &= \sum_I C_I |\Phi_I\rangle + \sum_i \sum_a c_i^a |\Psi_i^a\rangle + \frac{1}{4} \sum_{ij} \sum_{ab} c_{ij}^{ab} |\Psi_{ij}^{ab}\rangle, \end{aligned} \quad (10.13)$$

where the coefficients c_i^a and c_{ij}^{ab} are obtained by minimizing the energy of the ic-MRCISD wave function. Therefore, the number of variational parameters in the ic-MRCISD wave function is of the order of $N_H^2 N_P^2$, which is significantly less than the corresponding for u-MRCISD. The lower scaling of ic-MRCISD makes it the methods of choice among all variants of MRCI. Note, however, that there are several very important variants of MRCI that only partially contract the wave function.

10.3 Multireference perturbation theory and coupled cluster theory

It is also possible to generalize perturbation theory to the case of a multideterminantal reference. As for standard Møller–Plesset perturbation theory, one has to select a partitioning of the Hamiltonian into zeroth- plus first-order terms. In the case of multideterminantal references this choice is not straightforward and several choices have been considered. The simplest strategy is to employ a **diagonal** one-body zeroth-order operator, like in Møller–Plesset perturbation theory, for example

$$\hat{H}_0 = \sum_p \epsilon_p \hat{a}_p^\dagger \hat{a}_p, \quad (10.14)$$

however, a more common choice employs a general one-body operator

$$\hat{F} = \sum_{pq} f_{pq} \hat{a}_p^\dagger \hat{a}_q, \quad (10.15)$$

where f_{pq} are matrix elements of the average Fock matrix and are computed as

$$f_{pq} = \langle \Psi^{(0)} | [\hat{a}_p, [\hat{H}, \hat{a}_q^\dagger]]_+ | \Psi^{(0)} \rangle. \quad (10.16)$$

Then the corresponding zeroth-order Hamiltonian is expressed as

$$\hat{H}_0 = \hat{P}_{\text{CAS}} \hat{H} \hat{P}_{\text{CAS}} + \hat{P}_{\text{SD}} \hat{F} \hat{P}_{\text{SD}}, \quad (10.17)$$

where \hat{P}_{CAS} and \hat{P}_{SD} are projection operators defined in terms of the CAS determinants and all singly and doubly excited determinants generated from the CAS (SD). For example, \hat{P}_{CAS} is defined as

$$\hat{P}_{\text{CAS}} = \sum_I^{\text{CAS}} |\Phi_I\rangle \langle \Phi_I|. \quad (10.18)$$

Another way to define \hat{F} is using the Dyal partitioning, which includes two-body terms in the active space. Note that it was also found beneficial to apply a shift to the active orbital energies (IPEA shift) to

Like for MRCI, one can develop various contracted variants of MRPT: uncontracted, partially-contracted, internally-contracted, and strongly contracted. All of these differ in the number of variational degrees of freedom. The most common formulation of MRPT are the second-order CAS perturbation theory (CASPT2) and second-order n -electron valence perturbation theory (NEVPT2). Describing these approaches in detail is beyond the scope of these notes and the interested reader is urged to consult the primary literature.

Many variants of multireference coupled cluster (MRCC) methods have also been proposed. These can be divided in four broad categories. The first is Fock-space multireference coupled cluster methods, which aim to describe states with different number of electrons. The second one is Jeziorksi–Monkhorst methods, which employ an uncontracted ansatz of the form

$$|\Psi_{\text{JM-MRCC}}\rangle = \sum_I^{\text{CAS}} C_I e^{\hat{T}_I} |\Phi_I\rangle. \quad (10.19)$$

The third type is internally-contracted MRCC theories, which generalize the ic-MRCI method and are based on the following ansatz

$$|\Psi_{\text{ic-MRCC}}\rangle = e^{\hat{T}} |\Psi_{\text{CAS}}\rangle = \sum_I^{\text{CAS}} C_I e^{\hat{T}} |\Phi_I\rangle. \quad (10.20)$$

There are also a large number of methods that use the single-reference CC ansatz but modify it so that it can describe some type of wave functions with mild multideterminantal character. Some example include methods-of-moment CC, equation-of-motion spin-flip CC, and the CCSDt/CCSDtq methods.