On the efficient implementation of many-body theories

Ph.D. dissertation

Zoltán Rolik

Supervisor: Prof. Péter Surján

Chemistry Doctoral School

Head of the School: Prof. György Inzelt

Theoretical and Physical Chemistry, Structural Chemistry Program

Head of the Program: Prof. Péter Surján

Eötvös Loránd University, Budapest

Institute of Chemistry,

Laboratory of Theoretical Chemistry

2010

Contents

Introduction

This dissertation collects our results in the field of many-body theory that we obtained in the past few years while working at the Laboratory of Theoretical Chemistry, Eötvös University, Budapest.

In the first Chapter we introduce some essential concepts in many-body theory like Bloch-equation and the diagrammatic approach. We discuss a multiconfiguration-function-based perturbation method, called multiconfiguration perturbation theory (MCPT), developed in our laboratory. The presentation is focused on the third-order implementation of the theory using an automatized implementation. Numerical results up to third-order are computed for molecular systems which show strong multi-reference character. Advantages and disadvantages of MCPT are discussed in detail. Results presented in this Chapter where partly published in [1-3].

The second Chapter summarizes our work in the field of Multipartitioning Many-Body Perturbation Theory (MP MBPT). This method is one of the possible generalizations of the single-reference many-body perturbation theory, showing several advantageous properties. We investigate the symmetry behavior of MP MBPT and suggest a way to correct the symmetry breaking present in the original theory. We also study in detail the connectivity and scaling properties of MP MBPT at high orders. The theory is implemented up to fourth-order and tested on various multi-reference systems to demonstrate its convergence and its robustness against intruder states. Our results in this field are published in [4] and a further manuscript is under preparation [5].

Finally, an efficient full configuration interaction (CI) algorithm is presented in the third Chapter. This algorithm makes full use of the sparsity of the CI vector. The motivation of this study is the experience, that both CPU time and storage requirement can be successfully reduced if omitting zero elements (i.e. components below threshold) of

the CI vector. The algorithm explores important subspaces of the full CI space during iteration. Definition of ”importance” of a determinant is based on an approximation of contribution for the energy.

Numerical results show that the computation cost can be significantly reduced accepting errors far below the chemical accuracy. These results are published in [6].

Chapter 1

Multi-reference Perturbation Theories

1.1 Introduction

Single-reference many-body perturbation theory (MBPT) [7] that starts from a single determinant zero order state, has a limited field of application due to the well known problems of the inadequacy of a single determinant in describing several chemically important situations. The success of MBPT applied to closed shell systems around equilibrium geometry motivated an extensive research towards multi-reference perturbation theories (MRPT). The central idea of these theories is to replace the single determinant reference function by a multideterminantal expansion of moderate size. This reference function captures the most essential correlation effect of the system (so-called static correlation) for the qualitatively correct description but lacks the contribution from myriads of excited determinants (so-called dynamic correlation). The large number of various formalism emerged in this field [1,8-21] shows that the generalization of MBPT that would be perfect in every respect, is not a trivial task.

In this Section a brief theoretical overview is given of the main concepts and difficulties of multi-reference perturbation theories. It is not possible here to show all aspects of the different approaches, and only some of them will be discussed in detail. In this section we also introduce the main technical concepts necessary in the field of MBPT (e.g. model space, effective Hamiltonian, extensivity) and fix the notations used throughout the thesis.

In the following subsection the general framework of Bloch equation and the effective Hamiltonian theory will be introduced which gives the basics of many MRPT theories like the multipartitioning Møller-Plesset perturbation theory which will be discussed in Chapter 2.

1.1.1 Remarks on connectivity and scaling properties

Among the several many-body approximation those are preferred which hold as many properties of the exact solution as possible. Approximations with the proper spin and spatial symmetry properties or with orbital rotational invariance are desired. Similarly, an approximation is preferred which shows the same scaling properties with the system size as the exact solution has.

In this section we discuss scaling property connected to non-interacting systems, which is usually called size-consistency and the scaling property connected to large systems, which is often called size-extensivity. Both properties are strongly related to the connectivity of mathematical formulae of the approximations.

An expression constructed from two multi-indexed quantity (e.g. v_{pqrs} and u_{mn}) is connected if there is at least one summation to a common index of them: $\sum_m v_{pqrm}u_{mn}$. An expression composed from more than two terms is connected if there is no sub-product which is not connected to the remaining terms.

The connectedness is an important feature of a many-particle theory since it is strongly related to the scaling properties like size-consistency and size-extensivity [22, 23]. A method is size-consistent if its application to systems contains two non-interacting subspace gives the same energies than the sum of the energies come from the application on the subsystems. Considering a system of two non-interacting parts and applying localized orbitals to describe the subsystems each connected diagram and thus the energy can be written as a sum of the contributions of the two subsystems, since the integrals connecting orbitals from both subsystems are zero. If the method is size-inconsistent, dangerous unphysical effects can enter to the description of large systems. For example in a dilute gas the size-consistency error scales quadratically with the number of particles instead of the expected linear scaling of the energy [24].

The later property, i.e. linear scaling of the energy with the system size is called size-

extensivity. Of course if the system contains interacting subsystems this property can be expected only for the application to sufficiently (infinitely) large systems. According to work of Brueckner[25, 26], Goldstone[27] and Bartlett[23] connected theories are also size-extensive (see also [28]).

1.1.2 Bloch-equation and the effective Hamiltonian

MRPT theories can be particularly important in the cases where the qualitatively correct description of the many-electron system needs linear combination of determinants. Sometimes more than one (target) states are of interest where some of them can have multi-reference character. Let us separate the many electron space into a model space P and its orthogonal complement Q, called outer space or orthogonal space. The model space should contain the most important determinants of each target state. In terms of orthogonal vectors $|J\rangle$ the model space projector is given by

$$\hat{P} = \sum_{J \in P} |J\rangle\langle J| \tag{1.1}$$

and the projector of the orthogonal space is $\hat{Q} = \hat{I} - \hat{P}$. Denoting the exact eigenstates of our interest by Ψ_k and their model space projection by $\hat{P}\Psi_k$ we introduce the wave operator $\hat{\Omega}$ that maps any $\hat{P}\Psi_k$ to Ψ_k:

$$\Psi_k = \hat{\Omega}\hat{P}\Psi_k \ . \tag{1.2}$$

Multiplying Eq.(1.2) by $\hat{P}$, one sees that

$$\hat{P} = \hat{P}\hat{\Omega}\hat{P} \ . \tag{1.3}$$

Substituting form (1.2) of Ψ_k on the left hand side of the Schrödinger-equation $\hat{H}\Psi_k = E_k\Psi_k$ and multiplying by $\hat{P}$ one readily obtains

$$\hat{H}_{\text{eff}}\ \hat{P}\Psi_k = E_k\ \hat{P}\Psi_k \tag{1.4}$$

where

$$\hat{H}_{\text{eff}} = \hat{P}\hat{H}\hat{\Omega}\hat{P} \tag{1.5}$$

is the effective Hamiltonian whose eigenvalues are exact and the eigenvectors are projections of the exact eigenvectors into the model space. According to the generalized

Bloch-equation[29, 30] the effective Hamiltonian satisfies $\hat{H}\hat{\Omega}\hat{P} = \hat{\Omega}\hat{H}_{\text{eff}}\hat{P}$, writing in more detail

$$\hat{H}\hat{\Omega}\hat{P} = \hat{\Omega}\hat{P}\hat{H}\hat{\Omega}\hat{P}. \tag{1.6}$$

The solution of the Bloch-equation can be obtained by using the main concept of the perturbation theory, namely the partitioning of the Hamiltonian for a zero-order and a perturbation-operator :

$$\hat{H} = \hat{H}_0 + \hat{V}. \tag{1.7}$$

Here we suppose that the perturbation operator $\hat{V}$ is "small" compared to the zero-order Hamiltonian. In this case the main contribution to the wave operator arises from $\hat{H}^{(0)}$ and the corrections connecting to the perturbation-operator can be ordered by the order of the perturbation:

$$\hat{\Omega} = \hat{P} + \hat{\Omega}^{(1)} + \hat{\Omega}^{(2)} + \dots \tag{1.8}$$

Upper index of $\hat{\Omega}$ refers to the overall power of $\hat{V}$ in the expression. Applying Eq.(1.7) and Eq.(1.1) in the Bloch-equation (1.6), we get:

$$\left(\hat{H}^{(0)} + \hat{V}\right)\hat{\Omega}\hat{P} = \sum_{J\in P} \hat{\Omega}|J\rangle\langle J|\left(\hat{H}^{(0)} + \hat{V}\right)\hat{\Omega}\hat{P}\,. \tag{1.9}$$

Supposing that $\hat{H}^{(0)}$ is diagonal on the basis of states $|J\rangle$ we can substitute $\hat{H}^{(0)}$ by $E_J^{(0)}$ on the right hand side of the above equation: $|J\rangle$:

$$\left(\hat{H}^{(0)} + \hat{V}\right)\hat{\Omega}\hat{P} = \sum_{J\in P} \hat{\Omega}|J\rangle\langle J|\left(E_J^{(0)} + \hat{V}\right)\hat{\Omega}\hat{P}\,. \tag{1.10}$$

where $E_J^{(0)}$ denotes the eigenvalue of $\hat{H}^{(0)}$ corresponding to $|J\rangle$. Reordering the above formula by moving the terms involving the perturbation-operator to the right we obtain

$$\sum_{J\in P}\left(\hat{H}^{(0)} - E_J^{(0)}\right)\hat{\Omega}|J\rangle\langle J| = \hat{\Omega}\hat{P}\hat{V}\hat{\Omega}\hat{P} - \hat{V}\hat{\Omega}\hat{P}\,. \tag{1.11}$$

Multiplying this formula from the right by projector $|J\rangle\langle J|$ and from the left by the so-called reduced resolvent operator $\left(E_J^{(0)} - \hat{Q}\hat{H}^{(0)}\right)^{-1}\hat{Q}$ we get:

$$\hat{Q}\hat{\Omega}|J\rangle\langle J| = \left(E_J^{(0)} - \hat{H}^{(0)}\hat{Q}\right)^{-1}\hat{Q}\left(\hat{V}\hat{\Omega} - \hat{\Omega}\hat{P}\hat{V}\hat{\Omega}\right)|J\rangle\langle J|\,. \tag{1.12}$$

To obtain the Q-P block of the wave operator the above formula is summed up for all model space states to result

$$\hat{Q}\hat{\Omega}\hat{P} = \sum_{J\in P} \hat{R}_J \left(\hat{V}\hat{\Omega} - \hat{\Omega}\hat{P}\hat{V}\hat{\Omega}\right) |J\rangle\langle J| \,, \tag{1.13}$$

where the shorthand $\hat{R}_J$ is introduced for the reduced resolvent:

$$\hat{R}_J = (E_J^{(0)} - \hat{Q}\hat{H}^{(0)})^{-1}\hat{Q}. \tag{1.14}$$

Eq.(1.13) is a recursive equation for the wave operator. By taking the first approximation $\hat{\Omega} = \hat{P}$ and iterating the expression, one can generate terms of Eq.(1.8) as follows:

$$\begin{aligned}
\hat{\Omega}^{(0)} &= \hat{P} \\
\hat{\Omega}^{(1)} &= \sum_{I\in P} \hat{R}_I \hat{V}|I\rangle\langle I| \\
\hat{\Omega}^{(2)} &= \sum_{I\in P} \hat{R}_I \left(\hat{V}\hat{\Omega}^{(1)} - \hat{\Omega}^{(1)}\hat{P}\hat{V}\hat{P}\right)\hat{P}_I \\
&= \sum_{I\in P} \hat{R}_I \hat{V}\hat{R}_I \hat{V}\hat{P}_I - \sum_{I\in P} \hat{R}_I \sum_{J\in P} \hat{R}_J \hat{V}\hat{P}_J \hat{V}\hat{P}_I \\
\hat{\Omega}^{(3)} &= \sum_{I\in P} \hat{R}_I \left(\hat{V}\hat{\Omega}^{(2)} - \hat{\Omega}^{(1)}\hat{P}\hat{V}\hat{\Omega}^{(1)} - \hat{\Omega}^{(2)}\hat{P}\hat{V}\hat{P}\right)\hat{P}_I \\
&= \sum_{I\in P} \hat{R}_I \hat{V}\hat{R}_I \hat{V}\hat{R}_I \hat{V}\hat{P}_I - \sum_{I\in P} \hat{R}_I \hat{V}\hat{R}_I \sum_{J\in P} \hat{R}_J \hat{V}\hat{P}_J \hat{V}\hat{P}_I \\
&- \sum_{I\in P} \hat{R}_I \sum_{J\in P} \hat{R}_J \hat{V}\hat{P}_J \hat{V}\hat{R}_I \hat{V}\hat{P}_I - \sum_{I\in P} \hat{R}_I \sum_{J\in P} \hat{R}_J \hat{V}\hat{R}_J \hat{V}\hat{P}_J \hat{V}\hat{P}_I \\
&+ \sum_{I\in P} \hat{R}_I \sum_{J\in P} \hat{R}_J \sum_{L\in P} \hat{R}_L \hat{V}\hat{P}_L \hat{V}\hat{P}_J \hat{V}\hat{P}_I
\end{aligned} \tag{1.15}$$

With the use of the wave operator up to a given order n the effective Hamiltonian corrections up to order $(n+1)$ can be constructed by the application of Eq.(1.5) and Eq.(1.7), e.g.:

$$\begin{aligned}
\hat{H}_{\text{eff}}^{(0)} &= \hat{P}\hat{H}^{(0)}\hat{\Omega}^{(0)} = \hat{P}\hat{H}^{(0)}\hat{P} \\
\hat{H}_{\text{eff}}^{(1)} &= \hat{P}\hat{V}\hat{\Omega}^{(0)} = \hat{P}\hat{V}\hat{P} \\
\hat{H}_{\text{eff}}^{(2)} &= \hat{P}\hat{V}\hat{\Omega}^{(1)} = \sum_{I\in P} \hat{P}\hat{V}\hat{R}_I \hat{V}\hat{P}
\end{aligned} \tag{1.16}$$

It is interesting to note that the sum of the zero and first-order effective Hamiltonian is the Hamiltonian projected to the model space: $\hat{H}_{\text{eff}}^{(0)} + \hat{H}_{\text{eff}}^{(1)} = \hat{P}\hat{H}\hat{P}$.

As it is shown above the zero-order approximation of effective Hamiltonian is corrected in a perturbative fashion. To obtain the approximate energies of the the system, the effective Hamiltonian has to be diagonalized. Perturbation theories based on an effective Hamiltonian formalism are often called *perturb then diagonalize* approaches, for this reason.

1.1.3 Quasi-degenerate perturbation theory

The first remarkable MRPT was the quasi-degenerate perturbation theory (QDPT) introduced by Brandow [8] and developed by several others [30, 31, 32]. This theory inspired many others and it has an important connection with the multipartitioning Møller-Plesset PT which is the topic of Chapter 2. The term 'quasi-degenerate' refers to the fact, that quasi-degenerate levels are collected into the model space, and their mutual interaction is accounted for by diagonalization instead of direct perturbation.

Brandow's QDPT is based on the Bloch-equation and seeks a perturbative solution as outlined in the previous subsection. In QDPT formalism the many electron model space is chosen to be a complete active space (CAS) . This means that the molecular orbitals are gathered into three groups called *core*, *active* and *external* orbitals where the core orbitals are doubly occupied while the external orbitals are unoccupied in all determinants of the CAS. The *inactive* denomination refers both to *core* and *external* orbitals. The CAS is defined by the determinants with all possible occupation of active orbitals by active electrons, which do not reside in the core. In QDPT determinants are used to expand the model space. The zero-order Hamiltonian is a diagonal one-particle operator

$$\hat{H}^{(0)} = \sum_p \epsilon_p \hat{p}^+ \hat{p}^- \ , \tag{1.17}$$

similar to that used in single-reference Møller-Plesset theory[33]. The zero-order energy of determinant $|I\rangle$ is given by

$$E_I^{(0)} = \sum_{i \in I} \epsilon_i \ . \tag{1.18}$$

where $i \in I$ denotes that the dummy index is restricted to orbitals occupied in determinant $|I\rangle$. One-particle energies ϵ_i have been defined in many ways. In the simplest case the ϵ's

are chosen to be the diagonal elements of the Fock matrix defined by determinant $|I\rangle$

$$\epsilon_q = h_{qq} + \sum_{i \in I} \langle iq||iq\rangle \tag{1.19}$$

where h_{qq} is a one-electron integral and $\langle iq||iq\rangle = \langle iq|iq\rangle - \langle iq|qi\rangle$ is the antisymmetric two-electron integral in $\langle 12|12\rangle$ convention. Another frequently used method is to take the diagonal elements of a generalized Fock matrix defined by the one-particle density matrix $P_{pq} = \langle \Phi_k | \hat{p}^+ \hat{q}^- | \Phi_k \rangle$ of a given eigenstate Φ_k of the CAS Hamiltonian (i.e. the Hamiltonian projected to the model space):

$$\epsilon_q = h_{qq} + \sum_{rs} P_{rs} \langle rq||sq\rangle \tag{1.20}$$

As inferred from Eq.(1.16), diagonalization of CAS Hamiltonian gives the energies and wave functions in space P up to order 1. The first-order correction for the QDPT wave operator from stems from Eq.(1.15) by applying Eq.(1.17) to get:

$$\hat{\Omega}^{(1)} = \sum_{I \in P} \left\{ \sum_{\substack{ai \\ \hat{a}^+ \hat{i}^- |I\rangle \in Q}} \frac{h_{ai}}{\epsilon_i - \epsilon_a} \hat{a}^+ \hat{i}^- + \frac{1}{2} \sum_{\substack{a<b; i<j \\ \hat{a}^+ \hat{b}^+ \hat{i}^- \hat{j}^- |I\rangle \in Q}} \frac{\langle ab|ji\rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \hat{a}^+ \hat{b}^+ \hat{i}^- \hat{j}^- \right\} |I\rangle\langle I| \tag{1.21}$$

where $|I\rangle$ is a determinant in the CAS, $\hat{a}^+ \hat{i}^- |I\rangle$ and $\hat{a}^+ \hat{b}^+ \hat{i}^- \hat{j}^- |I\rangle$ are determinants of the Q space, and the Hamiltonian operator is written as:

$$\hat{H} = \sum_{pq} h_{pq} \hat{p}^+ \hat{q}^- + \frac{1}{2} \sum_{pqrs} \langle pq|sr\rangle \hat{p}^+ \hat{q}^+ \hat{r}^- \hat{s}^- . \tag{1.22}$$

In the above and further on we use the following nomenclature for one-electron orbitals:

$a, b, \ldots$	virtual
$i, j, \ldots$	occupied
$p, q, \ldots$	general

with respect to the Fermi vacuum $|I\rangle$. It is worth to note here that the perturbational denominators belong to the energy difference of a model space determinant and an outer

space determinant according to Eq.(1.14). As a consequence at least one of the one-particle indices of any perturbation denominator must be inactive.

An attractive feature of QDPT is that it can be expressed in a diagrammatic way, similar to that used in single-reference many-body PT (MBPT)[7]. For this reason the denomination multi-reference many-body PT (MR MBPT) is also frequently applied [32]. Application of diagrams in the many-body framework will be discussed in detail in Sections 1.3.1 and 2.3.2. Diagrammatic formulation of a theory is advantageous on one hand because it facilitates efficient implementation. On the other hand it offers a way to analyze the extensivity property which is closely related to the connectedness of diagrams. Using the diagrammatic approach Brandow [8] (see also Lindgren [30]) showed that if the model space is a CAS, the linked-cluster theorem remains true in QDPT. This means that only connected diagrams yield contribution to the effective Hamiltonian matrix elements (see also Section 2.3.2). As shown by Meissner and Jankowski[34], additive separability of the energy over non-interacting subsystems follows from connectedness, if using localized orbitals. The source of the restriction for localized orbitals is the fact that MR MBPT is not invariant to the rotation of orbitals, since it relies on zero-order Hamiltonian diagonal in the determinantal basis. Rotational invariance holds among degenerate orbitals only.

The main drawback of MR MBPT theory is the intruder state problem, i.e. close to zero denominators which give nonphysically large contribution to the effective Hamiltonian matrix elements and thus hinder the convergence. According to Eq.(1.14), denominators contain zero-order energy differences of determinants from CAS space and determinants from the orthogonal space. In the general case a CAS contains not only the low lying states (or determinants) close in energy to the target state(s) but also states (determinants) with high energy and less physical meaning. These high energy functions usually overlap in energy with the outer space determinants and their difference gives close to zero denominator. The appearance of high energy determinants and thus the intruder state problem is especially probable in the case when a large CAS space is used. In other words, the problem in Møller-Plesset partitioning is connected with the fact, that particle energies of the active orbitals can appear with positive and negative sign in the denominators. This means that there is no well defined gap which would separate the occupied and virtual space like in the single-reference case. In higher-orders the denom-

inators are constructed from larger number of ϵ's which frequently became close to zero. To eliminate this problem the application of incomplete model space was intensively studied in MR MBPT e.g. by Hose and Kaldor [31] and by Meissner and Bartlett[32]. By applying an incomplete model space high energy determinants can be excluded from the model space but special care has to be taken not to loose size-extensivity[34, 35].

A more sophisticated way to eliminate the intruder state problem is the intermediate Hamiltonian theory proposed by Malrieu [9], where the model space is split into a main space and an intermediate space. The effective Hamiltonian is defined over the entire model space, but only the eigenvalues corresponding to the main space are considered reliable. States falling to the intermediate space are not sought after. The intermediate Hamiltonian theory is formulated so that the energy differences of the main space and the orthogonal space appear in the denominators and the energies of the intermediate states can not contribute. If the main space is energetically well separated from the orthogonal space, the theory gives intruder-free solutions.

An alternative workaround for the intruder state problem – also proposed by Malrieu and coworkers – is the application of different partitioning for each column of the effective Hamiltonian[19]. This approach – termed by Multipartitioning Møller-Plesset PT – keeps most of the beneficial properties of MR MBPT. We discuss this theory in detail in Chapter 2.

1.1.4 An alternative QDPT formulation

To suppress the intruder state problem in the approach of Davidson [15] and Nakano [36] the $\Phi_k^{(0)} = \sum_I C_I^k |I\rangle$ eigenvectors of the CAS Hamiltonian are applied. The zero-order Hamiltonian is chosen diagonal on the basis of $\Phi_k^{(0)}$, $k = 1, ..., m$ space in the model states and also diagonal in the orthogonal space, on the basis of determinants $|J\rangle$:

$$\hat{H}_0 = \sum_k E_k^{(0)} |\Phi_k^{(0)}\rangle\langle\Phi_k^{(0)}| + \sum_J E_J^{(0)} |J\rangle\langle J|. \tag{1.23}$$

The $E_k^{(0)}$ energies are composed of one-particle energies as a generalization of the Møller-Plesset partitioning: $E_k^{(0)} = \sum_p \epsilon_p \langle\Phi_k^{(0)}|\hat{p}^+\hat{p}^-|\Phi_k^{(0)}\rangle$ and
$E_J^{(0)} = \sum_p \epsilon_p \langle J|\hat{p}^+\hat{p}^-|J\rangle$, where ϵ_p's are derived from the generalized Fockian (1.20). The first-order wave operator correction follows from Eq.(1.8) and Eq.(1.23) and gets the

form[36]:

$$\hat{\Omega}^{(1)} = \sum_{I\in P} \sum_{\substack{ai \\ \hat{a}^+\hat{i}^-|I\rangle \in Q}} h_{ai} \sum_k C_I^k \frac{\hat{a}^+\hat{i}^-}{\epsilon_i - \epsilon_a + E_k^{(0)} - E_I^{(0)}} |I\rangle\langle\Phi_k^{(0)}| \tag{1.24}$$

$$+\frac{1}{2}\sum_{I\in P} \sum_{\substack{a<b;i<j \\ \hat{a}^+\hat{b}^+\hat{i}^-\hat{j}^-|I\rangle \in Q}} \langle ab|ji\rangle \sum_k C_I^k \frac{\hat{a}^+\hat{b}^+\hat{i}^-\hat{j}^-}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b + E_k^{(0)} - E_I^{(0)}} |I\rangle\langle\Phi_k^{(0)}|$$

Here the summations over one-particle indices are restricted so that determinants $\hat{a}^+\hat{i}^-|I\rangle$ and $\hat{a}^+\hat{b}^+\hat{i}^-\hat{j}^-|I\rangle$ belong to the outer space. The projector on the right hand side of Eq.(1.8) is written in the form $|\Phi_k^{(0)}\rangle\langle\Phi_k^{(0)}| = \sum_I C_I^k |I\rangle\langle\Phi_k^{(0)}|$ to be able to express the energies of the outer space determinants $\hat{a}^+\hat{i}^-|I\rangle$ and $\hat{a}^+\hat{b}^+\hat{i}^-\hat{j}^-|I\rangle$ by $E_I^{(0)} - \epsilon_i + \epsilon_a$ and $E_I^{(0)} - \epsilon_i - \epsilon_j + \epsilon_a + \epsilon_b$, respectively. Comparing the above formula with Eq.(1.21) the most important difference is the $E_k^{(0)} - E_I^{(0)}$ denominator shift which appears in Davidson-Nakano's formalism. This shift has large value if $|I\rangle$ is a model determinant with high energy. If the target states are energetically well separated from the outer space, this shift can ensure non-zero denominators thereby decreasing sensitivity to intruder states as confirmed by experience [36].

Unfortunately this MRPT approach is not size-consistent, as will be shown later. Still, the beneficial effect of the denominator shift can be utilized in a size-consistent way as shown by Finley [37]. His diagrammatic CASPT will be discussed in more detail in Chapter 2.

1.1.5 Multi-reference Møller-Plesset (MRMP) theory

Our consideration so far involved effective Hamiltonians, constructed by a perturbation philosophy. Another main branch of the application of PT for the electron correlation problem is the so-called *diagonalize, then perturb* approach. These multi-reference theories are also based on a CAS subspace of limited size and the CAS eigenvectors are used to form the zero-order states within this subspace. Perturbation theory is used subsequently to improve the CAS states. By this approach the resolvent is kept diagonal and the zero-order energies corresponding to the active states are the CAS eigenvalues.

One of the simplest method of the diagonalize then perturb class is Hirao's MRMP [38]. It can be looked upon as a state-specific version of Nakano's PT with the zero-

order operator defined in Eq.(1.23). As already mentioned, the effective Hamiltonian approach based on Eq.(1.23) is not size-consistent. As a consequence Hirao's MRMP does not fulfill this property either. A similar problem may arise in the Multiconfigurational Perturbation Theory (MCPT) to be discussed in Section 1.2. To see the source of the size-consistency harming behavior in detail, let us look at the second-order energy of Hirao's MRMP:

$$E_k^{(2)} = \sum_{L \in Q} \frac{\langle \Phi_k^{(0)} | \hat{H} | L \rangle \langle L | \hat{H} | \Phi_k^{(0)} \rangle}{E_k^{(0)} - E_L^{(0)}}. \tag{1.25}$$

Supposing that the system contains two non-interacting subsystems denoted by A and B, the Hamiltonian and the energies of the composite system are sums over subsystems: $\hat{H} = \hat{H}_A + \hat{H}_B$, $E_k^{(0)} = E_{k_A}^{(0)} + E_{k_B}^{(0)}$ and if localized orbitals are used the zero-order (CAS) wave function is the product of subsystems functions, $|\Phi_k^{(0)}\rangle = |\Phi_{k_A}^{(0)} \Phi_{k_B}^{(0)}\rangle$. The determinants of the orthogonal space are also of product form: $|L\rangle = |L_A L_B\rangle$. The model space of the composite system $\hat{P}$ can be constructed by the direct product of the model spaces of subsystems $\hat{P}_A$ and $\hat{P}_B$. The complement of the direct product of $\hat{P}_A$ and $\hat{P}_B$ gives projector $\hat{Q}$. The second-order energy correction follows from these considerations and from Eq.(1.25):

$$E_{AB}^{(2)} = \left\{ \sum_{|L_A L_B\rangle \in Q} \frac{\langle \Phi_{k_A} \Phi_{k_B} | \hat{H}_A | L_A L_B \rangle \langle L_A L_B | \hat{H}_A | \Phi_{k_A} \Phi_{k_B} \rangle}{E_{k_A} + E_{k_B} - E_{L_A} - E_{L_B}} + \right.$$
$$\left. \sum_{|L_A L_B\rangle \in Q} \frac{\langle \Phi_{k_A} \Phi_{k_B} | \hat{H}_A | L_A L_B \rangle \langle L_A L_B | \hat{H}_B | \Phi_{k_A} \Phi_{k_B} \rangle}{E_{k_A} + E_{k_B} - E_{L_A} - E_{L_B}} \right\} + \left\{ A \leftrightarrow B \right\}. \tag{1.26}$$

In the first term $\hat{H}_A$ can not act on $|L_B\rangle$ thus the $\langle \Phi_{k_B} | L_B \rangle$ scalar product appears. As a result, the summation over $|L_B\rangle$ can be restricted to P_B and $|L_A\rangle$ must lie in Q_A to ensure that $|L_A L_B\rangle$ is in subspace Q. Based on the previous argument, the second term of Eq.(1.26) is trivially zero. Introducing this into Eq.(1.26) one gets:

$$E_{AB}^{(2)} = \sum_{L_A \in Q_A; L_B \in P_B} \frac{\langle \Phi_{k_A} \Phi_{k_B} | \hat{H}_A | L_A L_B \rangle \langle L_A L_B | \hat{H}_A | \Phi_{k_A} \Phi_{k_B} \rangle}{E_{k_A} + E_{k_B} - E_{L_A} - E_{L_B}} + (A \leftrightarrow B)$$
$$= \sum_{L_A \in Q_A; L_B \in P_B} {C_{L_B}^{k_B}}^2 \frac{\langle \Phi_{k_A} | \hat{H}_A | L_A \rangle \langle L_A | \hat{H}_A | \Phi_{k_A} \rangle}{E_{k_A} + E_{k_B} - E_{L_A} - E_{L_B}} + (A \leftrightarrow B) \tag{1.27}$$

To separate the size-consistency complying and violating terms one can approximate the $(E_{k_A} + E_{k_B} - E_{L_A} - E_{L_B})^{-1}$ expression supposing that $E_{k_B} - E_{L_B}$ (where L_B belongs to the

CAS space) is small compared to $E_{k_A} - E_{L_A}$ (where L_A belongs to the orthogonal space). Up to the first-order one can write

$$\frac{1}{E_{k_A} - E_{L_A}} - \frac{E_{k_B} - E_{L_B}}{(E_{k_A} - E_{L_A})^2} \tag{1.28}$$

thus the lowest order approximation of the size-consistency error is:

$$E_{AB}^{(2)}(\text{size-cons. err.}) = \sum_{L_A \in Q_A; L_B \in P_B} C_{L_B}^{k_B\,2} \quad (E_{k_B} - E_{L_B}) \frac{\langle \Phi_{k_A} | \hat{H}_A | L_A \rangle \langle L_A | \hat{H}_A | \Phi_{k_A} \rangle}{(E_{k_A} - E_{L_A})^2} + (A \leftrightarrow B) \tag{1.29}$$

The above expression can be considered small in the case where the zero-order function is well separated from the orthogonal space in both non-interacting subspaces. This holds irrespective of the possible close-lying levels within the model space. The reason is the fact, that if $|L_B\rangle$ is a level quasi-degenerate with $|k_B\rangle$ then $C_{L_B}^{k_B}$ is large but $E_{k_B} - E_{L_B}$ is small. On the other hand, if levels $|k_B\rangle$ and $|L_B\rangle$ are far in energy, then $E_{k_B} - E_{L_B}$ is large but $C_{L_B}^{k_B}$ is certainly small in this case.

It is also apparent from Eq.(1.29) that in the case where excitations taking from P to Q space are localized on the non-interacting partners and $E_{k_B} - E_{L_B} = 0$ holds, the size-consistency error disappears.

1.1.6 Internally contracted approaches

An important class of MRPT theories uses a non-diagonal zero-order Hamiltonian and thus works with a non-diagonal resolvent operator. These theories rely on an iterative procedure to get the PT corrections, similarly to the approach adapted in localized Møller-Plesset PT [39-42]. Since the generalized Fockian

$$\hat{F} = \sum_{i,j=1}^{N} f_{ij} \hat{i}^{+} \hat{j}^{-}, \tag{1.30}$$

with the matrix elements being

$$f_{kl} = h_{kl} + \sum_{nm} P_{nm} \langle kn || lm \rangle \, . \tag{1.31}$$

is non-diagonal in the MR case, one must work with a non-diagonal resolvent if $\hat{H}^0$ is based on it. In Eq.(1.31) matrix P_{nm} is the density matrix belonging to the reference

Φ_0. In the general case a multiconfigurational CAS state is not an eigenfunction of the generalized Fock operator, thus $\hat{F}$ can not be considered as a zero-order operator in the Rayleigh-Schrödinger framework. One can however easily construct a proper zero-order Hamiltonian by the definition:

$$\hat{H}_0 = E^{(0)}|\Phi_0\rangle\langle\Phi_0| + \hat{Q}\hat{F}\hat{Q}, \tag{1.32}$$

where $\hat{Q}$ is the idempotent operator which projects to the orthogonal space:

$$\hat{Q} = \hat{I} - |\Phi_0\rangle\langle\Phi_0| \, . \tag{1.33}$$

To recover the Møller-Plesset zero-order in the single-reference case, it is practical to choose the zero-order energy $E^{(0)}$ as the expectation value of the generalized Fockian with the zero-order function:

$$E^{(0)} = \langle\Phi_0|\hat{F}|\Phi_0\rangle. \tag{1.34}$$

As a consequence of the appearance of projector $\hat{Q}$ in Eq.(1.32) the zero-order Hamiltonian is generally not diagonal on the determinant basis. The first-order perturbation correction for Φ_0 using zero-order Hamiltonian (1.32) reads

$$|\Psi^{(1)}\rangle = (E^{(0)} - \hat{Q}\hat{H}_0)^{-1}\hat{Q}\hat{V}|\Phi_0\rangle, \tag{1.35}$$

where $\hat{V} = \hat{H} - \hat{H}_0$ is the perturbation operator. Efficiency of the calculation of the resolvent operator $(E^{(0)} - \hat{Q}\hat{H}_0)^{-1}\hat{Q}$ depends on the structure of matrix $\hat{H}_0$.

The strongly block diagonal structure reduces the dimensionality of the diagonalization process. To utilize this, Wolinski and Pulay [11] suggested to use the zero-order Hamiltonian

$$\hat{H}_0 = E^{(0)}|\Phi_0\rangle\langle\Phi_0| + \hat{Q}_S\hat{F}\hat{Q}_S + \hat{Q}_D\hat{F}\hat{Q}_D + \dots \, , \tag{1.36}$$

where $\hat{Q}_S$ and $\hat{Q}_D$ project to the subspace of single and double excitations, respectively. Space $\hat{Q}_S$ is generated by the action of the spin-averaged excitation operator

$$\hat{E}_{pq} = \sum_{\sigma=\alpha,\beta} \hat{p}^+_\sigma \hat{q}_\sigma \tag{1.37}$$

on the reference function and a product of two such on operators are applied for space $\hat{Q}_D$. The functions generated this way are called internally contracted (IC) states. The advantage of IC functions over determinants interacting with Φ_0 is that IC states are much

less in number. The inversion of the resolvent operator can be done in an iterative fashion. Since the internally contracted basis is not orthogonal and generally linearly dependent, canonical orthogonalization[43] is used to determine an orthogonal basis and eliminate the redundancy.

Up to now we discussed multiconfigurational theories based on a CAS space. This property can be serious drawback on its own since the calculation cost of the CAS problem is exponentially growing with the model space dimension. Of the theories presented so far, it is the method of Wolinsky and Pulay which is capable to treat zero-order states of a general form. The non-diagonal $\hat{H}_0$ (1.32) was applied by Murphy and Messmer [44] for the case where the zero-order function Φ_0 is a general valence bond (GVB)[45] wave function. Utilizing the fact that a GVB wave function is usually significantly shorter than a CAS function, they adopted a determinant basis to expand the interacting subspace. This greatly simplifies the implementation.

To cure the difficulties originating in the inversion of the resolvent operator, for the case when the zero-order function is a CAS state, Roos and his coworkers applied an $\hat{H}_0$ which is composed of smaller non-zero blocks[46]. Since the zero-order wave function arises from a CAS calculation, the first-order wave function correction lies entirely in the $\hat{Q}_{SD}$ space, which is orthogonal to the CAS state. To reduce the computation cost of the inversion, Roos *et al* divide the $\hat{Q}_{SD}$ subspace into subspaces according to the different type of internally contracted excitations. It is possible to define an *internal*, a *semi internal* and an *external* subspace where the number of electrons in the external orbitals are zero, one or two, respectively. These three subspaces can be further divided according to the number of core-active excitations (zero, one or two). As a result, altogether eight distinct subspaces can be defined in $\hat{Q}_{SD}$. Choosing $\hat{F}$ in a diagonal form: $\hat{F} = \sum_p \epsilon_p \hat{E}_{pp}$ would be most favorable from the computation point of view, since this choice excludes all interactions between vectors spanning space $\hat{Q}_{SD}$. Unfortunately this zero-order gives unsatisfactory results at second-order thus the application of the non-diagonal

$$\hat{F} = \sum_{pq} f_{pq} \hat{E}_{pq} \tag{1.38}$$

operator was later introduced[47]. Most of the interaction between different subspaces of $\hat{Q}_{SD}$ remain zero and the inversion of the resolvent operator can be performed in an iterative manner. In this procedure treatment of internal and semi internal subspaces needs

the calculation of third and fourth-order density matrices, due to the appearance of the overlap matrix. To avoid the computationally demanding construction of higher-order density matrices Werner used configuration state functions (CSF) as basis in the internal and semi internal subspaces and implemented the CASPT theory up to third-order[16].

The exact wave functions and energies of a many electron system are independent of the molecular orbitals applied. Approximate wave functions – due to the various restrictions on the molecular orbital occupancies – may only partially fulfill or fully ruin this invariance. A CAS function used as the zero-order state in the different perturbation approaches shows invariance to rotation within the core, the active and the external orbitals' subspace. A well behaving MRPT method should keep this property. The MRPT framework by Wolinski and Pulay may or may not keep the orbital invariance shown by the CAS function. Crucial in this point of view is the definition of the Fockian. A diagonal form of $\hat{F}$ destroys orbital invariance, while application of definition (1.38) in formula (1.32) or (1.36) defines an orbital invariant zero-order Hamiltonian. This means that the rotational invariance of the CAS functions is kept at any order.

Size-consistency property of the MRPT formulation by Wolinski-Pulay and by Roos was discussed in detail by [48, 49] . Theoretical and numerical examinations were reported claiming size-consistency of the Wolinski-Pulay formalism. The approach by Roos is known to be size-inconsistent.

1.2 Multiconfigurational perturbation theory (MCPT)

1.2.1 Original formalism of MCPT

To present a new multiconfigurational perturbation formalism[1] we start with a function $|0\rangle$, that can be written as a weighted sum of a principal determinant $|\mathrm{HF}\rangle$ and several other Slater determinants $|K\rangle$:

$$|0\rangle \;=\; d_{\mathrm{HF}}|\mathrm{HF}\rangle + \sum_{K\neq\mathrm{HF}} d_K|K\rangle \,. \tag{1.39}$$

It is practical to choose $|\mathrm{HF}\rangle$ as the largest component in absolute value in $|0\rangle$. This reference function defines the projector onto a one-dimensional reference space

$$\hat{P} \;=\; |0\rangle\langle 0| \tag{1.40}$$

and its orthogonal complement $\hat{Q} = 1 - \hat{P}$.

In the spirit of perturbation theory, we consider $|0\rangle$ as the zero-order ground state function and look for perturbation corrections to it. For this end we define a formal zero-order Hamiltonian

$$\hat{H}^0 \;=\; E_0|0\rangle\langle 0| + \sum_{K\neq \mathrm{HF}} E_K|K'\rangle\langle\widetilde{K'}| \,. \tag{1.41}$$

where $|K'\rangle$, $|K\rangle \neq |\mathrm{HF}\rangle$ represent an overlapping set of projected determinants

$$|K'\rangle \;=\; \hat{Q}|K\rangle \;=\; |K\rangle - d_K|0\rangle \tag{1.42}$$

with the overlap matrix

$$S_{K'L'} \;= \langle K'|L'\rangle = \;\langle K|\left(1-\hat{P}\right)|L\rangle \;=\; \delta_{KL} - d_K d_L \,, \tag{1.43}$$

and $\langle\widetilde{K'}|$-s are biorthogonal to $|K'\rangle$, i.e. $\langle\widetilde{K'}|L'\rangle = \delta_{KL}$:

$$\langle\widetilde{K'}| = \sum_{L\neq \mathrm{HF}} S^{-1}_{K'L'}\langle L'| \tag{1.44}$$

where $S^{-1}_{K'L'}$ is a shorthand for the elements of the inverse of the overlap matrix (1.43). This inverse can be given analytically due to the simple structure of $S_{K'L'}$:

$$S^{-1}_{K'L'} = \;\delta_{KL} + \sum_{L\neq \mathrm{HF}} \frac{d_K d_L}{d^2_{\mathrm{HF}}}, \tag{1.45}$$

and the connection between the direct and biorthogonal elements is:

$$\langle\widetilde{K'}| = \;\langle K'| + \sum_{L\neq \mathrm{HF}} \frac{d_K d_L}{d^2_{\mathrm{HF}}}\langle L'| = \langle K| - \frac{d_K}{d_{\mathrm{HF}}}\langle \mathrm{HF}| \,. \tag{1.46}$$

Here and further on tildes will denote reciprocal (biorthogonal) vectors.

The perturbation operator is defined as

$$\hat{V} = \hat{H} - \hat{H}^0 \tag{1.47}$$

and low-order PT corrections are straightforward to construct based on standard biorthogonal PT:

$$E^{(2)} \;=\; -\sum_{K\neq \mathrm{HF}} \frac{\langle 0|\hat{H}|K'\rangle\langle\widetilde{K'}|\hat{H}|0\rangle}{E_K - E_0} \,, \tag{1.48}$$

$$E^{(3)} \;=\; \sum_{K\neq \mathrm{HF}}\sum_{L\neq \mathrm{HF}} \frac{\langle 0|\hat{H}|K'\rangle\langle\widetilde{K'}|\hat{V}|L'\rangle\langle\widetilde{L'}|\hat{H}|0\rangle}{(E_K - E_0)\,(E_L - E_0)} \,, \tag{1.49}$$

etc.

The zero-order ground state energy is most practically chosen as $E_0 = \langle 0|\hat{H}|0\rangle$, while the zero-order excited energies, E_K-s, are parameters of the theory. By fixing these numbers, one defines the partitioning in the MCPT framework. Several choices for E_K-s have been discussed previously[1, 2, 50].

There are two classical options for the definition of $E'_K s$. The first possibility – following Møller and Plesset – is to choose the zero-order excitation energies as the sum and difference of one-particle energies of the orbitals where electrons are removed/attached when generating determinant $|K\rangle$ starting from $|HF\rangle$:

$$E_K = E_0 + \Delta_{i,a}, \text{ with } |K\rangle = \hat{a}^+\hat{i}^-|\text{HF}\rangle, \Delta_{i,a} = \epsilon_a - \epsilon_i \ , \quad i \in |\text{HF}\rangle, \ a \notin |\text{HF}\rangle. \tag{1.50}$$

Energies of the higher excitations are defined in a similar way. One particle energies ϵ_q are simply taken as the diagonal element of the Fock operator corresponding to the reference determinant |HF⟩, already shown in Eq.(1.19). In multi-reference problems, however, the application of the diagonal elements of the generalized Fock operator Eq.(1.31) seems to be more suitable :

$$\epsilon_i = h_{ii} + \sum_{jk} \langle ij||ik\rangle P_{kj} \ , \tag{1.51}$$

with

$$P_{kj} = \langle 0|a_j^+ a_k|0\rangle \ . \tag{1.52}$$

We will use the Davidson-Kapuy (DK) denomination [41, 42, 51, 52, 53] for the cases where the diagonal part of the ordinary or generalized Fockian is used to construct the zero-order Hamiltonian. DK denomination is applied to distinguish this possibility from Møller-Plesset partitioning. The second simple option for the partitioning is the application of Epstein-Nesbet denominators where the energy belonging to any state is simply the expectation value of the Hamiltonian:

$$E_0 = \langle 0|\hat{H}|0\rangle, \ E_K = \langle K|\hat{H}|K\rangle \ . \tag{1.53}$$

The above definition of $\hat{H}^0$ differs from the zero-order Hamiltonian used by Wolinski and Pulay [11] and Roos [46, 12] (see also Section 1.1.6) in two essential points. First,

it does not apply projectors onto n-fold excited subspaces. Second, its reduced resolvent $\hat{X} = \left(E_0 - \hat{Q}\hat{H}^0\right)^{-1} \hat{Q}$ can be specified explicitly as

$$\hat{X} = \sum_{K=1} \frac{|K'\rangle\langle\tilde{K}'|}{E_0 - E_K} \tag{1.54}$$

That is, $\hat{X}$ is diagonal on the biorthogonal basis, while the inverse of the overlap matrix is known explicitly. Thus, no iterative procedure is required to get the perturbation corrections. This simplicity, on the other hand, is achieved by a non-Hermitian zero-order Hamiltonian and resolvent.

As already mentioned, a similar iteration-free procedure was formerly proposed by Hirao in his multi-reference Møller-Plesset perturbation theory[13, 54]. In Hirao's approach the overlap between the target state and the excited determinants in the CAS space was handled with a numerical diagonalization procedure while the treatment of the orthogonal space is essentially the same. It is important to emphasize here that MCPT does not suppose that the zero-order function comes from a CAS calculation. It can stem from any methods for example from GVB[45, 55] or APSG[56]. It is also to be noted that the MCPT framework as detailed here is inherently size-inconsistent. Size-inconsistency may partly originate from the denominators. This can be eliminated with a proper choice of zero-order energies, as already discussed in Section 1.1.4 (see Eq.(1.29)). A second source of size-inconsistency is the appearance of projector $\hat{P}$ in the zero-order operator (1.41) which can induce a coupling in $\hat{H}^0$ between non-interacting subsystems. This is similar to the problem encountered in the CASPT scheme[48, 49, 57].

1.2.2 Reformulation of MCPT: SC2-MCPT

In order to diminish the consistency violation of the MCPT framework one needs to redefine the zero-order Hamiltonian so that projector $\hat{P}$ is excluded[3]. To reach this goal, let us use unprojected Slater-determinants $|K\rangle$ instead of $|K'\rangle$ in $\hat{H}^0$:

$$\hat{H}^0_{\mathrm{SC2}} \;=\; E_0|0\rangle\langle\tilde{0}| + \sum_{K\neq\mathrm{HF}} E_K|K\rangle\langle\tilde{K}| \;. \tag{1.55}$$

Here (1.55) of the zero-order Hamiltonian $\langle\tilde{0}|$ and $\langle\tilde{K}|$-s stand for the reciprocal (biorthogonal) vectors of the overlapping set $\{|0\rangle\} \cup \{|K\rangle \,|\, K \neq \mathrm{HF}\}$. Perturbation theory defined

by Eq.(1.55) is called SC2-MCPT referring to the size-consistent property of the second-order energy. As it will be shown later, one has to introduce strict restrictions to the form of E_K energies to obtain a size-consistent second-order energy. To construct the tilded vectors, let us build the metric matrix of the overlapping set:

$$S_{KL} \;=\; \delta_{KL} \;+\; d_K\,\delta_{L0}\,(1-\delta_{K0}) \;+\; d_L\,\delta_{K0}\,(1-\delta_{L0}) \;. \tag{1.56}$$

(Case $K = 0$ designates the multiconfigurational reference state $|0\rangle$.) The inverse of overlap matrix (1.56) can be expressed by the closed formula

$$S^{-1}_{KL} = \delta_{KL} - \delta_{K0}\delta_{L0} + e_K e_L \tag{1.57}$$

with $e_0 = d_{\mathrm{HF}}^{-1}$, $e_I = -d_I d_{\mathrm{HF}}^{-1}$ for $I \neq 0$ and S^{-1}_{KL} is a shorthand for the elements of the inverse of the overlap matrix (1.56). Inverse (1.57) results the reciprocal vectors (see Eq.(1.44))

$$\langle\tilde{0}| \;=\; \frac{1}{d_{\mathrm{HF}}}\langle \mathrm{HF}| \tag{1.58}$$

and

$$\langle\tilde{K}| \;=\; \langle K| - \frac{d_K}{d_{\mathrm{HF}}}\langle \mathrm{HF}| \,. \tag{1.59}$$

The zero-order operator (1.55) has the advantage over (1.41) that it lacks projector $\hat{P}$, therefore coupling between two independent subsystems can no longer emerge from it. The zero-order ground state energy in this scheme is most practically taken as

$$E_0 \;=\; \langle\tilde{0}|\hat{H}|0\rangle \,, \tag{1.60}$$

so that $E^{(1)}$ vanishes. The zero-order excited energies E_K-s can be chosen in the spirit discussed in Section 1.2.1.

Similarly to the original formulation the reduced resolvent corresponding to the zero-order Hamiltonian (1.55) is diagonal in the biorthogonal formulation:

$$\hat{X}_{\mathrm{SC2}} \;=\; \sum_{K\neq \mathrm{HF}} \frac{|K\rangle\langle\tilde{K}|}{E_K - E_0} \,. \tag{1.61}$$

Corrections E_0 in the SC2 variant of MCPT take the form:

$$E^{(2)}_{\mathrm{SC2}} \;=\; -\sum_{K\neq \mathrm{HF}} \frac{\langle\tilde{0}|\hat{H}|K\rangle\langle\tilde{K}|\hat{H}|0\rangle}{E_K - E_0} \,, \tag{1.62}$$

$$E^{(3)}_{\mathrm{SC2}} \;=\; \sum_{K\neq \mathrm{HF}} \frac{\langle\tilde{0}|\hat{H}|L\rangle\langle\tilde{L}|\hat{V}|K\rangle\langle\tilde{K}|\hat{H}|0\rangle}{(E_L-E_0)(E_K-E_0)}\;, \tag{1.63}$$

etc.

Comparison of the second-order formulae Eq.(1.48) and Eq.(1.62) reveals that the latter is computationally cheaper than the former, since a sum for excited configurations is present in $\langle 0|$ not like in $\langle\tilde{0}|$. This might also give a warning that formula (1.62) may yield smaller corrections than (1.48) – this however is not found in our numerical tests to be presented in Section 1.4.

At the same time, numerical studies indicate, that dependence of formulae (1.62) and (1.63) on the Fermi-vacuum choice is much expressed than the dependence of (1.48) and (1.49). This unfavorable property can lead to difficulties in the description of potential energy surfaces when the dominant determinant is changing as will be shown in Section 1.4. A possible solution for this problem is discussed by Szabados and Surján [58], who suggested an averaging over the possible Fermi-vacua.

To examine the size-consistent or -inconsistent nature of the SC2 variant of MCPT, let us look first at the zero-order quantities for a joint system AB. We suppose, that the reference function is product separable:

$$|0\rangle \;=\; |0_A 0_B\rangle \tag{1.64}$$

just like its reciprocal vector

$$\langle\tilde{0}| \;=\; \frac{\langle \mathrm{HF}_A \mathrm{HF}_B|}{d_{\mathrm{HF}_A} d_{\mathrm{HF}_B}} \;=\; \langle\widetilde{0_A}\;\widetilde{0_B}| \tag{1.65}$$

giving rise to the additively separable zero-order ground state energy

$$E_{0,AB} \;=\; \langle\widetilde{0_A}\;\widetilde{0_B}|\hat{H}_A+\hat{H}_B|0_A 0_B\rangle \;=\; E_{0,A}+E_{0,B}\;. \tag{1.66}$$

Zero-order excited state *ket* vectors are excited determinants, where the excitation may take place on one system, or the other, or both: $|\mathrm{HF}_A K_B\rangle$, or $|K_A \mathrm{HF}_B\rangle$, or $|K_A L_B\rangle$. Unfortunately neither $|\mathrm{HF}_A K_B\rangle$ nor $|K_A \mathrm{HF}_B\rangle$ is a product of a zero-order vector on system A and another on system B, since the vector $|\mathrm{HF}\rangle$ is not contained in the expansion set [*cf.* Eq.(1.55)]. This has unfavorable consequences on the consistency property of the energy from third-order on.

Looking at zero-order excited *bra* vectors and using Eq.(1.59), one finds:

$$\langle\widetilde{\mathrm{HF}_A K_B}| \;=\; \langle \mathrm{HF}_A K_B| - \frac{d_{K_B}}{d_{\mathrm{HF}_B}}\langle \mathrm{HF}_A \mathrm{HF}_B| \;=\; \langle \mathrm{HF}_A \widetilde{K_B}| \tag{1.67}$$

similarly

$$\langle\widetilde{K_A \mathrm{HF}_B}| \;=\; \langle \widetilde{K_A} \mathrm{HF}_B| \tag{1.68}$$

and

$$\langle\widetilde{K_A L_B}| \;=\; \langle K_A L_B| - \frac{d_{K_A} d_{L_B}}{d_{\mathrm{HF}_A} d_{\mathrm{HF}_B}}\langle \mathrm{HF}_A \mathrm{HF}_B| \;\neq\; \langle \widetilde{K_A}\; \widetilde{L_B}| \,. \tag{1.69}$$

Apart from the constant d_{HF_A}, $\langle\widetilde{\mathrm{HF}_A K_B}|$ is the product of zero-order functions $\langle\widetilde{0_A}|$ and $\langle\widetilde{K_B}|$, which is desirable. However, this is not the case for $\langle\widetilde{K_A L_B}|$.

Zero-order excited state energies are considered in the form

$$E_K \;=\; E_0 + \Delta_K \,, \tag{1.70}$$

with Δ_K being constructed of one-particle energies that characterize the excitation taking from $|\mathrm{HF}\rangle$ to $|K\rangle$. By this Møller-Plesset (MP) type construction one can avoid emergence of a coupling between independent subsystems in the energy denominators, since excited state energies then look:

$$E_{\mathrm{HF}_A K_B} \;=\; E_0 + \Delta_{K_B} \,, \tag{1.71}$$

$$E_{K_A \mathrm{HF}_B} \;=\; E_0 + \Delta_{K_A} \,, \tag{1.72}$$

and

$$E_{K_A L_B} \;=\; E_0 + \Delta_{K_A} + \Delta_{L_B} \,. \tag{1.73}$$

Using the above zero-order functions and energies it is easy to see that the zero-order Hamiltonian (1.55) is not additive over non interacting subsystems A and B. Full size-consistency of the SC2-MCPT scheme therefore can not be expected. Still, we shall show that the first non vanishing corrections behave correctly. Let us start with the second-order energy

$$\begin{aligned} E^{(2)}_{AB} \;=\; &- \sum_{\substack{K_A L_B \\ |\mathrm{HF}_A \mathrm{HF}_B\rangle \neq |K_A K_B\rangle}} \frac{\langle\widetilde{0_A 0_B}|\hat{H}_A|K_A L_B\rangle\langle\widetilde{K_A L_B}|\hat{H}_A|0_A 0_B\rangle}{E_{K_A L_B} - E_0} \\ - &\sum_{\substack{K_A L_B \\ |\mathrm{HF}_A \mathrm{HF}_B\rangle \neq |K_A K_B\rangle}} \frac{\langle\widetilde{0_A 0_B}|\hat{H}_A|K_A L_B\rangle\langle\widetilde{K_A L_B}|\hat{H}_B|0_A 0_B\rangle}{E_{K_A L_B} - E_0} + \{A \leftrightarrow B\} \,. \end{aligned} \tag{1.74}$$

At this point we can substitute the expressions from Eq.(1.58), Eq.(1.65) and Eq.(1.67) then apply Eq.(1.70) and Eq.(1.73.). Utilizing that $\hat{H}_A$ does not act on vectors of system B,

$$E^{(2)}_{AB} = -\sum_{\substack{K_A,L_B \\ |\mathrm{HF}_A\mathrm{HF}_B\rangle\neq|K_AK_B\rangle}} \frac{\langle \mathrm{HF}_A\mathrm{HF}_B|\hat{H}_A|K_A\mathrm{HF}_B\rangle\langle\widetilde{K_A}\mathrm{HF}_B|\hat{H}_A|0_A0_B\rangle}{d_{\mathrm{HF}_A}d_{\mathrm{HF}_B}\Delta_{K_A}}$$
$$-\sum_{\substack{K_A,L_B \\ |HF_AHF_B\rangle\neq|K_AK_B\rangle}} \frac{\langle \mathrm{HF}_A\mathrm{HF}_B|\hat{H}_A|K_A\mathrm{HF}_B\rangle\langle\widetilde{K_A}\mathrm{HF}_B|\hat{H}_B|0_A0_B\rangle}{d_{\mathrm{HF}_A}d_{\mathrm{HF}_B}\Delta_{K_A}} + \{A \leftrightarrow B\}\,. \tag{1.75}$$

Since $\langle \mathrm{HF}_B|0_B\rangle = d_{\mathrm{HF}_B}$ the first term clearly does not depend on index B. The second term is zero due to the $\langle\tilde{K_A}|0_A\rangle = 0$, ($K_A \neq \mathrm{HF}_A$) coming from the biorthogonalality relation. As a result of the above derivation, the reformulation of the MCPT theory leads to size-consistent second-order:

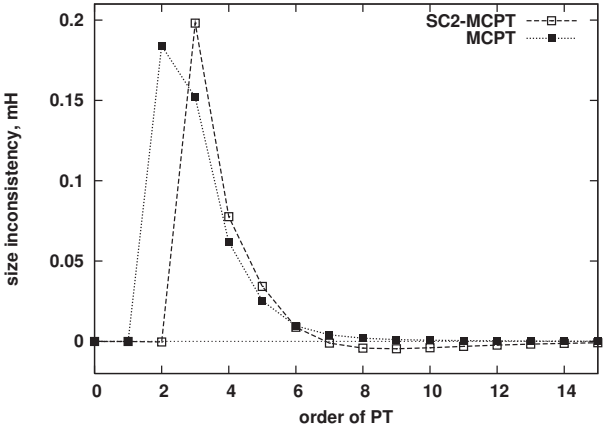
$$E^{(2)}_{AB} = E^{(2)}_A + E^{(2)}_B. \tag{1.76}$$

A similar analysis for the energy of order three or higher reveals the appearance of inter-system terms, leading to size-inconsistency.

As an illustration, two H_2 molecules in 6-311G** basis set in a rectangular arrangement were selected for numerical size-consistency check. Internuclear distance of the individual H_2 molecules was 1.0 Å , the two molecules systems were put 100 Å away from each other. A simple (2,2) CAS was computed as reference for the stretched H_2 molecule, and a (4,4) CAS was prepared for the non-interacting dimer.

Size-consistency violation as a function of the order of PT is plotted in Fig. 1.1 for this system. Here one sees, that consistency violation of the SC2-MCPT scheme remains slightly larger than that of MCPT at every order from third-order on. Size-inconsistency of the MCPT scheme diminishes faster, still the violation of the two schemes fall into the same order of magnitude at the third and higher-orders.

Figure 1.1: Size-inconsistency of MCPT and SC2-MCPT schemes on the example of two stretched H_2 molecules 100 Åfrom each other. For geometry see text. Numbers displayed are in mH, basis set used is 6-311G**. Inconsistency is computed as $E_{H_2...H_2} - 2E_{H_2}$.



1.3 Implementation of third-order MCPT and SC2-MCPT formalism

1.3.1 Wick's theorem and the diagrammatic representation

Theories discussed in the previous sections can be coded in a configuration driven algorithm which means that the *do loops* in the implementation run over the single determinantal states. In this way the second-order correction can be implemented with maximum efficiency by calculating quantity $\hat{H}|K\rangle$ for all the determinants appearing in the reference function $|0\rangle$. The resultant vector contains one- and two-fold excitations of the determinants appearing in the reference function $|0\rangle$. The leading term of the second-order energy correction stemming from Eq.(1.48) can be written as:

$$E^{(2)}_{leading} = \sum_{K \neq HF} \sum_{I,J} c_I c_J \frac{\langle I|\hat{H}|K\rangle\langle K|\hat{H}|J\rangle}{E_0 - E_K} \tag{1.77}$$

which shows that after the division by the denominators scalar products have to be calculated. In this way the calculation cost is proportional to the dimension of the interacting subspace of $|0\rangle$, which is approximately $NN_v^2N_o^2$, where N is the number of determinants in the reference function, N_v is the number of virtuals, N_o is the number of occupied orbitals. Though the configuration driven approach is simple, in higher-order it becomes impractical. To show this let us consider the leading term of formula (1.49):

$$E^{(3)}_{leading} = \sum_{K,L\neq HF} \sum_{I,J} c_I c_J \frac{\langle I|\hat{H}|K\rangle\langle K|\hat{H}|L\rangle\langle L|\hat{H}|J\rangle}{(E_0 - E_K)(E_0 - E_L)} \,. \tag{1.78}$$

If the main part of the first-order wave function correction

$$\Psi^{(1)}_{leading} = -\sum_{J} \sum_{L\neq \mathrm{HF}} \frac{|L\rangle\langle L|}{E_L - E_0} \hat{H}|J\rangle c_J \tag{1.79}$$

is already calculated, the calculation cost of $\hat{H}\Psi^{(1)}_{leading}$ is roughly proportional to $NN_v^4N_o^4$. However, this calculation is done partly in vain since only those elements of $\hat{H}\Psi^{(1)}_{leading}$ are needed for the calculation of Eq.(1.78) which may interact with $|0\rangle$ through $\hat{H}$. To avoid the calculation of unnecessary terms i.e. $\langle K|\hat{H}|\Psi^{(1)}_{leading}\rangle$ matrix elements where $\langle K|\Psi^{(1)}_{leading}\rangle = 0$, one should restrict the second quantized operator indices in operator $\hat{H}$. To achieve this, an orbital driven approach is to be followed.

To derive the formulae necessary for a sum-over-orbitals implementation it is important to clarify the mathematical background of the calculation of matrix elements of second quantized operator products. For simplicity we first consider the case where the expectation value of an operator product is evaluated:

$$\langle I|\hat{p}^+\hat{q}^+\ldots\hat{r}^-\hat{s}^-|I\rangle. \tag{1.80}$$

Operators in (1.80) can be classified as occupied (denoted by $i, j, k, \ldots$) or virtual (denoted by $a, b, c, \ldots$) operator with respect to determinant $|I\rangle$. Using the anti commutation relations of fermions

$$\{\hat{p}^-, \hat{q}^-\} = 0, \;\; \{\hat{p}^+, \hat{q}^+\} = 0, \;\; \{\hat{p}^+, \hat{q}^-\} = \delta_{pq} \tag{1.81}$$

one can move the occupied annihilation and virtual creation operators to the left and occupied creation and virtual annihilation operators to the right. After the rearrangement

the constant term will give the value of the matrix element of (1.80), since all other terms will give zero by definition. As an example let us see an eight membered operator product:

$$\begin{aligned}
&\hat{i}^+\hat{j}^+\hat{a}^+\hat{c}^-\hat{b}^+\hat{d}^-\hat{k}^-\hat{l}^- = -\hat{a}^+\hat{b}^+\hat{k}^-\hat{l}^-\hat{i}^+\hat{j}^+\hat{c}^-\hat{d}^- + \hat{a}^+\hat{k}^-\hat{l}^-\hat{i}^+\hat{j}^+\hat{d}^-\delta_{bc} - \hat{a}^+\hat{b}^+\hat{k}^-\hat{i}^+\hat{c}^-\hat{d}^-\delta_{jl} \\
&-\hat{a}^+\hat{b}^+\hat{l}^-\hat{j}^+\hat{c}^-\hat{d}^-\delta_{ik} + \hat{a}^+\hat{b}^+\hat{l}^-\hat{i}^+\hat{c}^-\hat{d}^-\delta_{jk} + \hat{a}^+\hat{b}^+\hat{k}^-\hat{j}^+\hat{c}^-\hat{d}^-\delta_{il} + \hat{a}^+\hat{l}^-\hat{j}^+\hat{d}^-\delta_{ik}\delta_{bc} \\
&+\hat{a}^+\hat{k}^-\hat{j}^+\hat{d}^-\delta_{il}\delta_{bc} + \hat{a}^+\hat{k}^-\hat{i}^+\hat{d}^-\delta_{bc}\,\delta_{jl} - \hat{a}^+\hat{l}^-\hat{j}^+\hat{d}^-\delta_{bc}\,\delta_{il} - \hat{a}^+\hat{l}^-\hat{i}^+\hat{d}^-\delta_{bc}\,\delta_{jk} \\
&-\hat{a}^+\hat{b}^+\hat{c}^-\hat{d}^-\delta_{jk}\delta_{il} + \hat{a}^+\hat{b}^+\hat{c}^-\hat{d}^-\delta_{jl}\delta_{ik} + \hat{a}^+\hat{d}^-\delta_{bc}\delta_{ik}\delta_{jl} - \hat{a}^+\hat{d}^-\delta_{bc}\delta_{il}\delta_{jk}. \qquad (1.82)
\end{aligned}$$

To simplify the calculation of the above formula and to get an easy-to-memorize calculation method, contraction of operators and normal ordered operator products are introduced [59]. Contractions of creation and annihilation operators are defined by the following rules:

$$\hat{a}^-\hat{b}^+ = \delta_{ab}, \quad \hat{i}^+\hat{j}^- = \delta_{ij}, \quad \hat{b}^+\,\hat{a}^- = 0, \quad \hat{j}^-\,\hat{i}^+ = 0 \qquad (1.83)$$

and all other contractions are zero. An operator product normal ordered with respect to reference determinant $|I\rangle$ is denoted by curly bracket:

$$\{\hat{i}^+\hat{j}^+\hat{a}^+\hat{c}^-\hat{b}^+\hat{d}^-\hat{k}^-\hat{l}^-\}_I = -\hat{a}^+\hat{b}^+\hat{k}^-\hat{l}^-\hat{i}^+\hat{j}^+\hat{c}^-\hat{d}^-. \qquad (1.84)$$

In the normal ordered form of an operator product, by definition, all virtual creation and all occupied annihilation operators are moved to the left of virtual annihilation and occupied creation operators. The string is multiplied by the sign of the permutation necessary to achieve the rearrangement. Note that the even permutation gets positive while the odd permutation receives negative sign. Note also that one can freely permute the operators in curly bracket, only the sign of the permutation has to be taken into account.

When deriving formula (1.82) the anti-commutation rule $\hat{p}^+\hat{q}^- = \delta_{pq} - \hat{q}^-\hat{p}^+$ was extensively used. The idea of contraction of operators is introduced to take into account the Kronecker-delta while the second term with the sign is taken into account by the application of the curly bracket. Rewriting the right hand side of Eq.(1.82) using contractions and normal ordered operators one obtains:

$$\begin{aligned}
\hat{i}^+\hat{j}^+\hat{a}^+\hat{c}^-\hat{b}^+\hat{d}^-\hat{k}^-\hat{l}^- = \{\hat{i}^+\hat{j}^+\hat{a}^+\hat{c}^-\hat{b}^+\hat{d}^-\hat{k}^-\hat{l}^-\}_I + \sum_{\textit{single contr.}} \{\overbrace{\hat{i}^+\hat{j}^+\hat{a}^+\hat{c}^-\hat{b}^+\hat{d}^-\hat{k}^-\hat{l}^-}\}_I \\
+ \sum_{\textit{double contr.}} \{\overbrace{\hat{i}^+\hat{j}^+\hat{a}^+\hat{c}^-\hat{b}^+\hat{d}^-\hat{k}^-\hat{l}^-}\}_I + \sum_{\textit{triple contr.}} \ldots + \sum_{\textit{full contr.}} \ldots \qquad (1.85)
\end{aligned}$$

Note, that free permutation in contracted normal ordered strings affects only the operators not involved in any contraction. It is also important that the application of operator contraction and rearrangement into normal order do not commute. In the above equation first operator contractions have to be calculated then the normal ordering acts of the remaining operator product is achieved. Contraction of any two operators in normal ordered form is zero by definition. As a consequence, any contraction within a normal ordered operator product gives zero. Summations in the above expression run over the possible singly contracted, doubly contracted etc. terms. The result obtained for the above example can be formulated in a general manner as follows. A general product of second quantized operators can be expressed by the summation of all the possible normal ordered quantities constructed by contracting the operators in the product in all possible way. This is the Wick's theorem[59]. Application of Wick's theorem to formula (1.80) yields

$$\langle I|\hat{p}^+\hat{q}^+\ldots\hat{r}^-\hat{s}^-|I\rangle = \sum_{full\ contr} \langle I|\{\overbrace{\hat{p}^+\hat{q}^+\ldots\hat{r}^-\hat{s}^-}\}_I|I\rangle\,, \tag{1.86}$$

where expectation values of normal ordered operator products with the reference determinant are not included since they all give zero. In other words, the expectation value of a second quantized operator product calculated with a single determinant involves only the fully contracted terms. It is easy to understand that the sign of a fully contracted term can be calculated based on the number of crossings of the contraction lines.

In practice we usually meet products of normal ordered operators like

$$\{\hat{p}^+\ldots\hat{r}^-\}_I\{\hat{q}^+\ldots\hat{s}^-\}_I\,. \tag{1.87}$$

This formula as a whole can be rewritten into normal ordered form by using Wick's theorem and noting that any contractions connecting two operators within the same bracket give zero. This leads to the generalized Wick's theorem [59] which can be written for the product of two normal ordered operator as:

$$\begin{aligned}\{\hat{p}^+\ldots\hat{r}^-\}_I\{\hat{q}^+\ldots\hat{s}^-\}_I &= \{\{\hat{p}^+\ldots\hat{r}^-\}_I\{\hat{q}^+\ldots\hat{s}^-\}_I + \\ &+ \sum_{single\ contr.} \{\{\overbrace{\hat{p}^+\ldots\hat{r}^-}\}_I\{\hat{q}^+\ldots\hat{s}^-\}_I + \ldots + \sum_{full\ contr.} \{\overbrace{\hat{p}^+\ldots\hat{r}^-}\}_I\{\overbrace{\hat{q}^+\ldots\hat{s}^-}\}_I\,.\end{aligned} \tag{1.88}$$

A similar rule applies to products of more than two normal ordered terms.

An important application of Wick's theorem is the separation of the electronic Hamiltonian into the fully contracted and the remaining normal ordered part:

$$\hat{H} = \hat{H}_N + \langle I|\hat{H}|I\rangle, \tag{1.89}$$

where the first term is the Hamiltonian normal ordered with respect to determinant $|I\rangle$ and the second is the Hartree–Fock energy belonging to determinant $|I\rangle$. The explicit form of the normal ordered Hamiltonian contains the elements of the Fock matrix defined by the reference determinant $|I\rangle$. Without the long but straightforward derivation we quote the expression of the normal order Hamiltonian:

$$\hat{H}_N = \sum_{pq} f^I_{pq}\{\hat{p}^+\hat{q}^-\}_I + \frac{1}{4}\sum_{pqrs}\langle pq||rs\rangle\{\hat{p}^+\hat{q}^+\hat{s}^-\hat{r}^-\}_I \tag{1.90}$$

Here we introduced the Fock matrix defined by the occupied orbitals in determinant $|I\rangle$

$$f^I_{pq} = h_{pq} + \sum_{i\in I}\langle pi||qi\rangle \tag{1.91}$$

and the antisymmetric two electron integrals $\langle pq||rs\rangle = \langle pq|rs\rangle - \langle pq|sr\rangle$ are written in the $\langle 12|12\rangle$ convention.

We now set out to derive formulae of a sum-over-orbital expression of MCPT. This can be achieved by calculating full contractions of normal ordered operator products stemming from $\hat{H}_N$. As we will see, to derive the working equations of a sum-over-orbitals implementation of a many-body approach, full contractions of normal ordered products can be processed by a computer. To simplify the treatment and reduce the large number of terms, a diagrammatic representation is useful.

To demonstrate the diagrammatic representation, we briefly discuss its application to single-reference Many-Body Perturbation Theory (MBPT)[7]. MBPT can be considered as a special case of the QDPT where the CAS space contains only the Hartree–Fock determinant. In this simplest case all perturbation operators $\hat{V}$ in Eqs.(1.15) and (1.16) can be changed into its normal ordered counterpart $\hat{V}_N$. The normal ordering is now defined with respect to the Hartree–Fock determinant $|I\rangle$. In the single-reference case there is no summation over J and L in Eq.(1.15) and all projectors $\hat{P}$ are indexed by I. Since $\hat{P}_I\hat{V}_N\hat{P}_I$ is zero, at third-order only the first term, at fourth-order only the first and the third terms give contributions. We can finally substitute the form of $\hat{P}_I$ expressed with

$|I\rangle$ the right hand side of Eqs.(1.15) and write $\Psi^{(n)}$ wave function corrections instead of the wave operator components. As a result the different orders of the wave function look:

$$\begin{aligned}
\Psi^{(0)} &= |I\rangle \\
\Psi^{(1)} &= \hat{R}_I \hat{V}_N |I\rangle \\
\Psi^{(2)} &= \hat{R}_I \hat{V}_N \hat{R}_I \hat{V}_N |I\rangle \\
\Psi^{(3)} &= \hat{R}_I \hat{V}_N \hat{R}_I \hat{V}_N \hat{R}_I \hat{V}_N |I\rangle - \hat{R}_I^2 \hat{V}_N \hat{P}_I \hat{V}_N \hat{R}_I \hat{V}_N |I\rangle
\end{aligned} \tag{1.92}$$

Using Hartree–Fock canonical orbitals, the zero-order Hamiltonian is given by Eq.(1.19) where $\epsilon_q = f^I_{qq}$. Since f^I_{pq} is diagonal, the perturbation operator is of the form

$$\hat{V} = \frac{1}{4} \sum_{pqrs} \langle pq||sr\rangle \hat{p}^+ \hat{q}^+ \hat{r} \hat{s}. \tag{1.93}$$

Using relation $E^{(0)} = \langle \Psi^{(0)} | \hat{H}_0 | \Psi^{(0)} \rangle$ and $E^{(n)} = \langle \Psi^{(n-1)} | \hat{V}_N | \Psi^{(0)} \rangle$ one obtains the Hartree–Fock energy as, the zero-order plus the first-order energy:

$$E_I^{(0)} + E_I^{(1)} = \langle I | \hat{H}_0 + \hat{V} | I \rangle \tag{1.94}$$

and the higher-order terms takes the form

$$E^{(2)} = \langle I | \hat{V}_N \hat{R}_I \hat{V}_N | I \rangle, \tag{1.95}$$

$$E^{(3)} = \langle I | \hat{V}_N \hat{R}_I \hat{V}_N \hat{R}_I \hat{V}_N | I \rangle, \tag{1.96}$$

$$E^{(4)} = \langle I | \hat{V}_N \hat{R}_I \hat{V}_N \hat{R}_I \hat{V}_N \hat{R}_I \hat{V}_N | I \rangle - \langle I | \hat{V}_N \hat{R}_I^2 \hat{V}_N | I \rangle \langle I | \hat{V}_N \hat{R}_I \hat{V}_N | I \rangle. \tag{1.97}$$

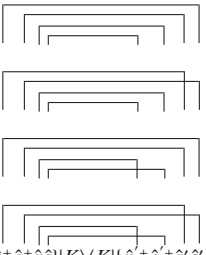
To introduce the diagrammatic technique, let us take a closer look at the second-order expression:

$$\begin{aligned}
E_I^{(2)} &= \sum_{K \neq I} \frac{\langle I | \hat{V}_N | K \rangle \langle K | \hat{V}_N | I \rangle}{E_I^{(0)} - E_K^{(0)}} \\
&= \frac{1}{16} \sum_{K \neq I} \sum_{pqrs} \sum_{p'q'r's'} \langle pq||sr\rangle \langle p'q'||s'r'\rangle \frac{\langle I | \{\hat{p}^+ \hat{q}^+ \hat{r}^- \hat{s}^-\} | K \rangle \langle K | \{\hat{p}'^+ \hat{q}'^+ \hat{r}'^- \hat{s}'^-\} | I \rangle}{E_I^{(0)} - E_K^{(0)}}.
\end{aligned} \tag{1.98}$$

Using the generalized Wick's theorem (1.88) one has to construct all possible full contractions of the creation and annihilation operators. It can be realized in four different

Figure 1.2: Possible contractions in the second-order MBPT energy.

ways, i.e.:



$$\frac{1}{16}\langle pq||sr\rangle\langle p'q'||s'r'\rangle\langle I|\{\hat{p}^+\hat{q}^+\hat{r}\hat{s}\}|K\rangle\langle K|\{\hat{p}'^+\hat{q}'^+\hat{r}'\hat{s}'\}|I\rangle \tag{1.99}$$

where the sums are omitted from Eq.(1.98). The second and third full contractions themselves give a minus sign, but together with the antisymmetric integrals all four expressions yield the same result. The value altogether is

$$E_I^{(2)} = \frac{1}{4}\sum_{abij}\frac{\langle ij||ab\rangle\langle ab||ij\rangle}{\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b}. \tag{1.100}$$

Using a different graphical representation the four different contractions can be symbolized by four graphs, shown in Fig. 1.2. The horizontal lines (vertices) symbolize the antisymmetric integrals in formula (1.99). The upward or downward oriented lines connecting the vertices symbolize the contractions. Upgoing lines labeled by a and b are virtual, downgoing lines denoted by i and j are occupied orbitals. The vertices of two-electron integrals are defined so that there is one incoming and one outgoing contraction line on both sides of the vertex. Outgoing arrows correspond to creation operators, incoming arrows refer to annihilation operators. Some examples of possible one- and two-electron vertices are shown in Fig. 1.3. Due to the antisymmetry of the two electron integrals the creation operators or the annihilation operators can be freely exchanged without sign change. Using this flexibility the incoming and outgoing lines of a vertex can be exchanged among

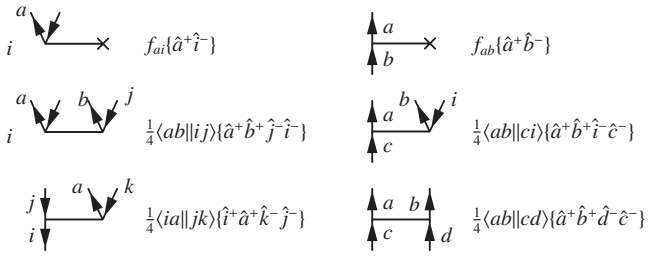


Figure 1.3: Examples for diagrammatic representation of the one- and two-electron operators

themselves. For example the following diagrams

(1.101)

represent the same operator. As a result of the above flexibility the four diagrams in Fig. 1.2 give the same result, thus one can consider them to be equivalent and draw only the first one. The general convention is to use the diagram with the maximum number of loops from the equivalent topologies[60]. The contributions of the equivalent diagrams can be taken into account by using a prefactor determined by the topology of the diagram. There are various different conventions of drawing diagrams, the types presented here are called antisymmetrized Goldstone-diagrams.

To show the determination of the multiplicative factor, consider the symmetries of the two electron part of the Hamiltonian at Eq.(1.90). Using the fermion anti-commutator rule Eq.(1.81) and the properties of the antisymmetrized two-electron integrals it is easy to see that when establishing a contraction like

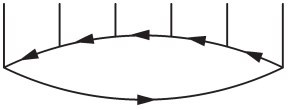
$$\ldots\frac{1}{4}\sum_{pqrs}\langle pq||sr\rangle\ldots\{\hat{p}^{''+}\hat{q}^{''+}\hat{r}^{''-}\hat{s}^{''-}\}_I\{\hat{p}^+\hat{q}^+\hat{r}^-\hat{s}^-\}_I\{\hat{p}^{'+}\hat{q}^{'+}\hat{r}^{'-}\hat{s}^{'-}\}_I\ldots \tag{1.102}$$

one can draw three other contractions yielding the same algebraic expression, similarly to the example shown in Eq.(1.99). In the case of a product containing n pieces of two-electron normal ordered terms of the Hamiltonian, one can generate the 4^n equivalent terms to a possible full contraction by using the above mentioned symmetry. The factor 4^n would exactly cancel the factor $(\frac{1}{4})^n$ coming from Eq.(1.90) if there are no equivalent

contractions. Equivalent contractions i.e. two-particle or hole lines starting from the same vertex and ending in the same vertex alter the picture. If two vertices are connected by two equivalent contractions the number of distinct contractions is only two instead of four and the factor of $\frac{1}{4}$ is not fully cancelled. As a result each equivalent pair of (hole or particle) contraction lines gives a factor of $\frac{1}{2}$.

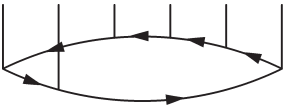
The sign of a contraction can be also read from the diagram as follows. First we suppose that the diagram is in the maximum loops form, which ensures that in each two-particle interaction term two operators (one creation and one annihilation operator) belong to the same loop while the the remaining two operators belong to another loop. One can also suppose for any vertex, that the creation-annihilation pair which belong to the same loop are on the same end, i.e. left or right. This form of a diagram can be achieved by contraction line shifts of the type shown in (1.101), and it involves no sign change. Having a diagram in such a form, it is clear that its sign is a product of loop contributions, since any creation-annihilation pair of a loop is separated by even number of other operators. Using this fact it is enough to determine the sign rule for a single loop constructed by the multiplication of one-particle like terms. The following examples help to understand the general rule:

$$\{\hat{p}_1^+\hat{q}_1^-\}\{\hat{p}_2^+\hat{q}_2^-\}\{\hat{p}_3^+\hat{q}_3^-\}\{\hat{p}_4^+\hat{q}_4^-\}\{\hat{p}_5^+\hat{q}_5^-\}\{\hat{p}_6^+\hat{q}_6^-\}\,. \tag{1.103}$$



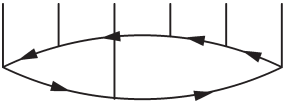
Contractions are also illustrated by the diagrammatic form where now the left going lines are virtual orbitals (particles) and the right going lines are occupied orbitals (holes). In the above example the sign of the loop is positive since there is no crossing of contraction lines while in the second example

$$\{\hat{p}_2^+\hat{q}_2^-\}\{\hat{p}_1^+\hat{q}_1^-\}\{\hat{p}_3^+\hat{q}_3^-\}\{\hat{p}_4^+\hat{q}_4^-\}\{\hat{p}_5^+\hat{q}_5^-\}\{\hat{p}_6^+\hat{q}_6^-\}\,. \tag{1.104}$$

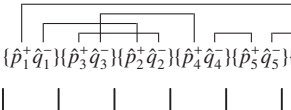


there is one more occupied orbital and the sign becomes negative. One can consider that in the above diagrammatic representation the sign comes from the contraction line crossing the interaction vertex. The following two examples similarly give negative sign:

$$\{\hat{p}_2^+\hat{q}_2^-\}\{\hat{p}_3^+\hat{q}_3^-\}\{\hat{p}_1^+\hat{q}_1^-\}\{\hat{p}_4^+\hat{q}_4^-\}\{\hat{p}_5^+\hat{q}_5^-\}\{\hat{p}_6^+\hat{q}_6^-\}\,. \tag{1.105}$$

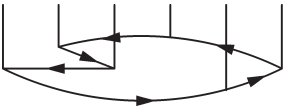


$$\{\hat{p}_1^+\hat{q}_1^-\}\{\hat{p}_3^+\hat{q}_3^-\}\{\hat{p}_2^+\hat{q}_2^-\}\{\hat{p}_4^+\hat{q}_4^-\}\{\hat{p}_5^+\hat{q}_5^-\}\{\hat{p}_6^+\hat{q}_6^-\}\,. \tag{1.106}$$



The last example shows the case of two additional hole lines giving positive sign:

$$\{\hat{p}_2^+\hat{q}_2^-\}\{\hat{p}_4^+\hat{q}_4^-\}\{\hat{p}_3^+\hat{q}_3^-\}\{\hat{p}_5^+\hat{q}_5^-\}\{\hat{p}_1^+\hat{q}_1^-\}\{\hat{p}_6^+\hat{q}_6^-\}\,. \tag{1.107}$$



Drawing further examples one can confer that a loop with odd number of hole lines contributes positive sign while a loop with even hole lines produces negative sign. Each loop

give similar contributions thus the final result is $(-1)^{l+h}$ where the l is the number of loops and h is the number of hole lines.

Up to now we omitted the reduced resolvent operator $\hat{X}$ which separates the two normal ordered operators in Eq.(1.98). This projection like operator can be expressed using the Fermi vacuum in the form:

$$\hat{X} = \sum_{ai} \frac{\hat{a}^+\hat{i}^-|I\rangle\langle I|\hat{i}^+\hat{a}^-}{\epsilon_i - \epsilon_a} + \frac{1}{4}\sum_{abij} \frac{\hat{a}^+\hat{b}^+\hat{i}^-\hat{j}^-|I\rangle\langle I|\hat{i}^+\hat{j}^+\hat{b}^-\hat{a}^-}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} + \dots \tag{1.108}$$

Applying the above formula for the last contraction of Eq.(1.99) and keeping only the relevant terms from Eq.(1.108) we obtain:

$$\frac{1}{16}\frac{\langle pq||sr\rangle\langle p'q'||s'r'\rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}\langle I|\{\hat{p}^+\hat{q}^+\hat{r}^-\hat{s}^-\}\{\hat{a}^+\hat{b}^+\hat{j}^-\hat{i}^-\}|I\rangle\langle I|\{\hat{i}^+\hat{j}^+\hat{b}^-\hat{a}^-\}\{\hat{p}'^+\hat{q}'^+\hat{r}'^-\hat{s}'^-\}|I\rangle \tag{1.109}$$

where summation on all indices is implied. As we can see all indices enter into operator $\hat{X}$ from the right and leave to the left. As a result, a denominator appears in operator $\hat{X}$ with one-particle energies constructed of indices which enter the reduced resolvent. For simplicity, contractions with the reduced resolvent operator are usually not denoted, but are taken into account by dividing by the proper denominator. In the diagrams the denominators are denoted by horizontal lines crossing the contraction lines. According to the above discussion the denominators are given by formula

$$D = \sum \epsilon_{down} - \sum \epsilon_{up}, \tag{1.110}$$

where ϵ_{down} are the one-particle energies of the down going indices crossing the denominator line and ϵ_{up} are the one-particle energies of the up going lines.

Fig. 1.4.a and Fig. 1.4.b show only the topology of the second and third-order diagrams. These type of diagrams, which do not contain arrows, are called skeleton diagrams. At second-order only one diagram can be generated from the skeleton diagram. At third-order three different diagrams can be generated, since there are three different possibilities for the orientation of the loops. The connected fourth-order diagrams are more numerous, three examples are shown in Figs. 1.4.c. In Figs. 1.4.d we show all fourth-order disconnected diagrams. Diagrams belonging to the first term of Eq.(1.97) term contain a reduced resolvent operator at each intervertex level. Since in diagram d_3 there is no excitation with respect to the Fermi-vacuum at the second intervertex level, this

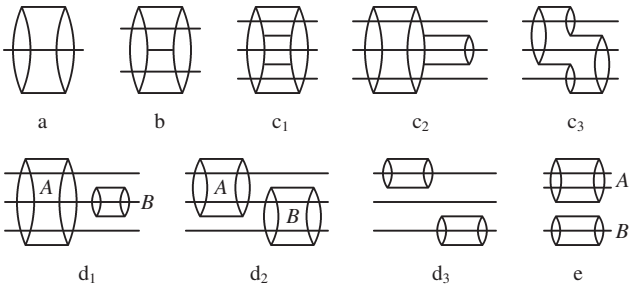


Figure 1.4: Topology of MBPT diagrams. Diagrams a and b show the topology of the second and third-order diagrams, respectively. Diagrams c_1, c_2 and c_3 are examples for the connected fourth-order MBPT diagrams, while diagrams d_1, d_2 and d_3 are the fourth-order disconnected MBPT diagrams. Diagram e is the renormalization diagram (second term in the last row of Eq.(1.92).

diagram has zero value. Apart from the minus sign, Fig. 1.4.e shows the renormalization diagram belonging to the second term of the last equation of (1.92). This renormalization diagram is also disconnected.

According to the linked cluster theorem, disconnected terms in MBPT can not give contribution to the energy[8, 27, 61]. To show the validity of the linked cluster theorem at fourth-order, consider the sum of diagrams d_1 and d_2:

$$\sum_{A,B} \frac{K(A)L(B)}{D_A^2(D_A+D_B)} + \sum_{A,B} \frac{K(A)L(B)}{D_A(D_A+D_B)D_B} \equiv \sum_{A,B} \frac{K(A)L(B)}{D_A^2 D_B}, \tag{1.111}$$

where A and B are hyper indices which denote the four orbital indices appearing at a given intervertex level according to Figs. d_1 and d_2, and K(A) and L(B) are products of two integrals depending either on A or on B. This result equals the negative of the renormalization term thus the cancellation of disconnected terms is complete. Eq.(1.111) is a simple application of the Frantz-Mills factorization theorem[62].

At higher-orders the diagrammatic representation can bring unphysical terms where at a given intervertex level there are two or more particle or hole lines with the same index i.e. two electrons are removed from a given orbital of the Fermi-vacuum or two electrons are taken to a given orbital, respectively. See for example Fig. 1.5.a and Fig. 1.5.c. These unphysical terms are called exclusion principle violating (EPV) terms and it is obvious that they should not give contribution to the energy (or any physical quantity).

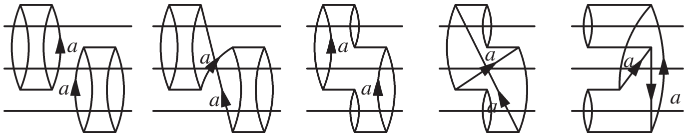


Figure 1.5: Examples for the EPV terms. Diagrams d and e are equal.

A transparent way to get rid of them is to restrict the summations but this would lead to technical difficulties. Fortunately, there is no need for such a restriction since these terms automatically cancel each other out. To show this, let us take diagrams in Figs. 1.5 a and b. The first diagram is clearly a disconnected diagram which gives an EPV term if two-particle indices have the value a, as Figs. 1.5.a shows. Exchanging the ending of these particle lines we get Figs. 1.5.b. Doing so the number of loops is reduced by one, the value of the second diagram therefore is of opposite sign and the sum of these two diagrams gives zero. In Fig. 1.5.c another EPV term is shown where exchanging the ending of lines labeled by a leads to Fig. 1.5.d which can be drawn in a more transparent form (as shown in Fig. 1.5.e.). This transformation gives again a sign change and the consequent cancellation of these terms. The above process ensures the automatic cancellation of EPV terms in every case which allows the automatic summation for each contraction line without any restrictions.

Formulae of multiconfiguration perturbation theory, e.g. (1.78) contains terms where determinant $\langle I|$ at the far left end is different from determinant $|J\rangle$ at the right end. In this case it is usual to consider determinant $\langle I|$ as an excited determinant with respect to Fermi vacuum $\langle J|$ in the form

$$\langle I| = \langle J|\{\hat{i}_1^+\hat{i}_2^+ \dots \hat{i}_n^+\hat{a}_n^- \dots \hat{a}_2^-\hat{a}_1^-\}_I. \tag{1.112}$$

As a result, the left hand side of the working formulae are closed by a normal ordered product like (1.112) instead of an interaction vertex, if J is different from I.

Such expressions can be represented by open diagrams, the open legs being indexed by the orbital indices in formula (1.112). The unclosed loops of an open diagram contribute to the sing rule, i.e. when counting loops, one must consider both closed and unclosed loops. When generating contributions of open diagrams, one must take into ac-

PT order	no open leg	one pair of open legs	two pairs of open legs
2nd	2	8	9
3rd	14	78	135
4th	213	1568	3416

Table 1.1: Number of diagrams at different perturbational orders, including contain both connected and disconnected ones. Diagrams with zero legs are the closed diagrams, diagrams with one (two) pair of open legs describe one (two)-fold excited bra determinant with respect to the ket determinant.

count permutation of open legs, if there are at least four of them. The permutation affects hole indices i_k and particle indices a_k among themselves. The resulting diagrams have to be multiplied by the parity of the permutation operators. An example for this is shown in connection with Eq.(1.118) in the following subsection.

Finally, it is worth to summarize the diagrammatic rules as follows.

- Each vertex represents a one- or two-particle integrals. Oriented lines connecting vertices are the contractions of virtual (upgoing) or occupied (downgoing) lines and there are summations for these indices in the algebraic form.

- Horizontal lines at the intervetex levels give denominators according to Eq.(1.110).

- Each equivalent pair of contraction lines (starting and ending at the same vertex, same orientation) gives a factor $\frac{1}{2}$.

- Sign comes from the number of hole lines (h) and the number of loops (l) according to $(-1)^{h+l}$.

- Off-diagonal matrix elements of the Hamiltonian are represented by open diagrams where the open lines are indexed by spin-orbital labels making the difference between the two determinants at the two ends of the expression. All permutations of open lines have to be considered with their proper sign. Open loops also contribute to the sign.

1.3.2 Automatization of derivations and code generation

In course of the diagrammatic implementation of MCPT and SC2-MCPT we followed the approach first used by Kaldor [31] where the Fermi-vacuum is defined by the rightmost determinant of the perturbation formula and it varies as the rightmost determinant changes. Starting with Eq.(1.78), there are two summations over the determinants in the reference state. Once determinants $|I\rangle$ and $|J\rangle$ are fixed, at order three one has to calculate terms like

$$\sum_{K,L\neq HF} \frac{\langle I|\hat{H}_N^{(J)}|K\rangle\langle K|\hat{H}_N^{(J)}|L\rangle\langle L|\hat{H}_N^{(J)}|J\rangle}{\Delta_{HF}^K \Delta_{HF}^L} \,. \tag{1.113}$$

To demonstrate the calculation process, let us take an example: a simple term from the third-order contribution where the first and the second Hamiltonian contribute a one-electron integral and the third Hamiltonian contributes a two-electron integral

$$T_{IJ}^3(112) = \frac{1}{4} f_{kl}^{(J)} f_{mn}^{(J)} \langle pq||sr\rangle\langle J|\{\hat{a}_2^+\hat{a}_1^+\hat{i}_2\hat{i}_1\}\{\hat{k}^+\hat{l}\}\hat{X}\{\hat{m}^+\hat{n}\}\hat{X}\{\hat{p}^+\hat{q}^+\hat{r}\hat{s}\}|J\rangle, \tag{1.114}$$

where the Einstein convention is used, i.e. there are summations for indices occurring twice. The second reduced resolvent operator $\hat{X}$ can be easily processed since all contractions of the rightmost operators have to be drawn above it:

$$T_{IJ}^3(112) = \frac{1}{4} \frac{f_{kl}^{(J)} f_{mn}^{(J)} \langle pq||sr\rangle}{\Delta_{HF}^J + \Delta_{r,s}^{p,q}} \langle J|\{\hat{a}_2^+\hat{a}_1^+\hat{i}_2^-\hat{i}_1^-\}\{k^+l\}\hat{X}\{\hat{m}^+\hat{n}^-\}\{\hat{p}^+\hat{q}^+\hat{r}^-\hat{s}^-\}|J\rangle \,. \tag{1.115}$$

where for the denominator a shorthand notation is introduced:

$$\Delta_{i_1,\ldots i_n}^{a_1,\ldots a_n} = \epsilon_{i_1} + \ldots + \epsilon_{i_n} - \epsilon_{a_1} - \ldots - \epsilon_{a_n} \tag{1.116}$$

Here it is taken into account that the denominator is defined with respect to the leading determinant of the multiconfiguration reference state while occupied and virtual indices are defined with respect to $|J\rangle$. The zero-order Hamiltonian is defined by Eq.(1.51). Denominators are calculated in two steps. At the first step the denominator of determinant $|J\rangle$ is determined (denoted by Δ_{HF}^J) then in the second step the contribution with respect to determinant $|J\rangle$ is calculated (denoted by $\Delta_{i_1,\ldots i_n}^{a_1,\ldots a_n}$):

$$E_{HF} - E_{\hat{p}^+\hat{q}^+\hat{r}^-\hat{s}^-|J\rangle} = \underbrace{E_J - E_{\hat{p}^+\hat{q}^+\hat{r}^-\hat{s}^-|J\rangle}}_{\Delta_{r,s}^{p,q}} + \underbrace{E_{HF} - E_J}_{\Delta_{HF}^J} \tag{1.117}$$

Taking into consideration in Eq.(1.115) the full contractions the result is the following:

$$T^3_{IJ}(112) = P^{a_1}_{a_2} P^{i_1}_{i_2} (-1)^{(p_a+p_i)} \Bigg\{ \frac{1}{4} \frac{\langle a_1 a_2 || i_2 i_1 \rangle}{\Delta^J_{HF} + \Delta^{a_1,a_2}_{i_1,i_2}} \sum_{ai} \frac{f_{ia} f_{ai}}{\Delta^J_{HF} + \Delta^{a,a_1,a_2}_{i,i_1,i_2}} \qquad (1.118)$$

$$d$$

$$+ \frac{1}{2} \sum_{a} \frac{\langle a_2 a || i_2 i_1 \rangle}{\Delta^J_{HF} + \Delta^{a,a_2}_{i_1,i_2}} \Bigg(\sum_{i} \frac{f_{ia} f_{a_1 i}}{\Delta^J_{HF} + \Delta^{a,a_1,a_2}_{i,i_1,i_2}} - \sum_{b} \frac{f_{a_1 b} f_{ab}}{\Delta^J_{HF} + \Delta^{b,a_2}_{i_1,i_2}} \Bigg)$$

$$a_2 \qquad\qquad a_1$$

$$+ \frac{1}{2} \sum_{i} \frac{\langle a_1 a_2 || i_2 i \rangle}{\Delta^J_{HF} + \Delta^{a_1,a_2}_{i,i_2}} \Bigg(\sum_{j} \frac{f_{ji_1} f_{ij}}{\Delta^J_{HF} + \Delta^{a_1,a_2}_{j,i_2}} - \sum_{a} \frac{f_{ia} f_{ai_1}}{\Delta^J_{HF} + \Delta^{a,a_1,a_2}_{i,i_1,i_2}} \Bigg)$$

$$b_1 \qquad\qquad b_2$$

$$+ \frac{1}{2} \sum_{ij} \frac{\langle a_1 a_2 || i j \rangle}{\Delta^J_{HF} + \Delta^{a_1,a_2}_{i,j}} \Bigg(\frac{f_{ji_1} f_{ii_2}}{\Delta^J_{HF} + \Delta^{a_1,a_2}_{j,i_2}} \Bigg) + \frac{1}{2} \sum_{ab} \frac{\langle ab || i_2 i_1 \rangle}{\Delta^J_{HF} + \Delta^{a,b}_{i_1,i_2}} \Bigg(\frac{f_{a_1 a} f_{a_2 b}}{\Delta^J_{HF} + \Delta^{a,a_2}_{i_1,i_2}} \Bigg)$$

$$e \qquad\qquad f$$

$$+ \sum_{ai} \frac{\langle a a_2 || i_2 i \rangle}{\Delta^J_{HF} + \Delta^{a,a_2}_{i,i_2}} \Bigg(\frac{f_{a_1 i_1} f_{ia}}{\Delta^J_{HF} + \Delta^{a_2}_{i_2}} - \frac{f_{i i_1} f_{a_1 a}}{\Delta^J_{HF} + \Delta^{a_1,a_2}_{i,i_2}} + \frac{f_{ia} f_{a_1 i_1}}{\Delta^J_{HF} + \Delta^{a,a_1,a_2}_{i,i_1,i_2}} - \frac{f_{a_1 a} f_{i i_1}}{\Delta^J_{HF} + \Delta^{a,a_2}_{i_1,i_2}} \Bigg) \Bigg\}$$

$$c_3 \qquad\qquad c_2 \qquad\qquad c_4 \qquad\qquad c_1$$

To implement MCPT and SC2-MCPT theory at third-order the Terms of Eq.(1.118) are also represented in diagrammatic form in Fig. 1.6. Diagrams having common lowest interaction vertex are denoted by the same letter for example a_1 and a_2. There are some disconnected diagrams in Fig. 1.6 for example c_3, c_4 and d.

At third-order both the contraction and the coding of diagrams represents significant difficulty because of the large number of possible diagrams. The number of diagrams at different perturbation orders are shown in Table (1.1). To implement MCPT and SC2-MCPT theory at third-order the derivation and the implementation of the diagrams were automatized. A symbolic algebraic code was developed to construct the possible contractions of normal ordered operator products using Wick's theorem. The input of the code contains the number of normal ordered products, and the number of creation/annihilation operators constituting each product, together with their spin. This input is produced by a script which generates all possible normal ordered products which can appear at a given order of PT. After the determination of fully contracted expressions of a given normal

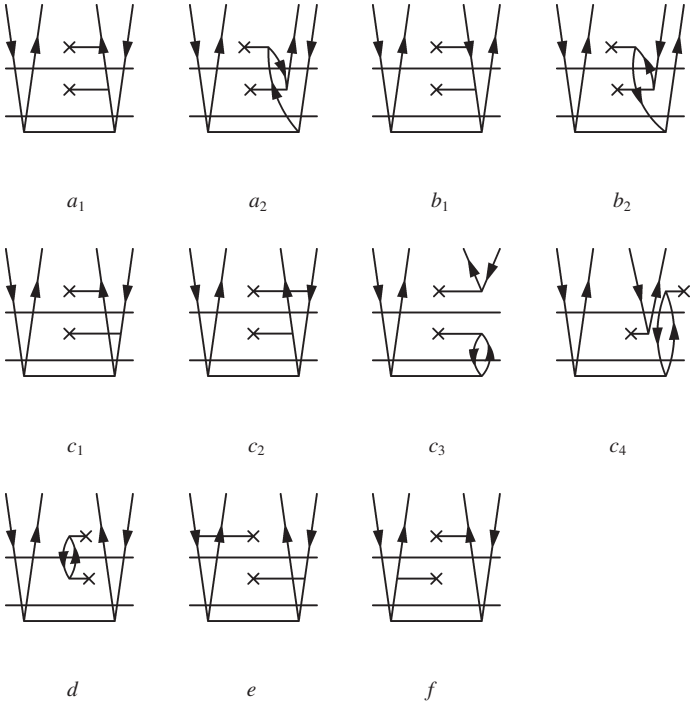


Figure 1.6: Examples for third-order diagrams appearing in multiconfigurational perturbation theories.

ordered product, the code gathers the equivalent terms into diagrams. As a result, the list of spin labeled diagrams with their coefficients is obtained. For example when evaluating term $T^3_{IJ}(112)$, the diagrams coded by the computer look

		1	
Aa(a) Ia(i) |	Aa(a) Ia(i) |	Va(1) Va(2) Oa(2) Oa(1)	0.25
		1	
Aa(a) Ia(i) |	Va(1) Ia(i) |	Va(2) Aa(a) Oa(2) Oa(1)	0.5
		2	
Va(1) Aa(a) |	Aa(a) Aa(b) |	Va(2) Aa(b) Oa(2) Oa(1)	-0.5
		1	
Ia(i) Oa(1) |	Ia(i) Ia(j) |	Va(1) Va(2) Oa(2) Ia(j)	0.5
		2	
Aa(a) Ia(i) |	Aa(a) Oa(1) |	Va(1) Va(2) Oa(2) Ia(j)	-0.5
		1	
Ia(i) Oa(1) |	Ia(j) Oa(2) |	Va(1) Va(2) Ia(j) Ia(i)	0.5
		1	
Va(1) Oa(1) |	Aa(a) Ia(i) |	Aa(a) Va(2) Oa(2) Ia(i)	0.5
		1	
Va(1) Oa(1) |	Aa(a) Ia(i) |	Aa(a) Va(2) Oa(2) Ia(i)	1.0
		2	
Ia(i) Oa(1) |	Va(1) Aa(a) |	Aa(a) Va(2) Oa(2) Ia(i)	-1.0
		3	
Aa(a) Ia(i) |	Va(1) Oa(1) |	Aa(a) Va(2) Oa(2) Ia(i)	1.0
		4	
Va(1) Aa(a) |	Ia(i) Oa(1) |	Aa(a) Va(2) Oa(2) Ia(i)	-1.0

In the above list vertical lines separate the integrals. Letters A and V indicate dummy and fixed virtual orbitals, respectively while I and O indicate dummy and fixed occupied orbitals, respectively. The second letters refer to the spin. In the above list all spins are alpha. Letters or numbers in brackets indicate the numbering or labeling of the given orbital.

For the efficient implementation, factorization of the perturbation terms is essential.

For this reason the symbolic algebraic code reorders the diagrams to be ready for factorized implementation. Such an ordering simplifies the factorization of diagrams in the second step where an automatized generation of the FORTRAN code occurs. For example at third-order those diagrams are collected into one block, which have the same integral at the first vertex (e.g. one-body term with alpha spin or two body term with mixed spin etc.). Numbers separating different lines in the above list are the serial numbers at the given block. These blocks could have sub-blocks containing diagrams which have the same structure up to the second integral from the bottom but this factorization is not used yet.

At the next stage of the automatized implementation a second code was written which uses the diagrams, as input and generates the FORTRAN code itself. Similar solutions were already used in quantum chemistry by Li and Paldus [63] and Hirata [64]. Kállay and Surján [65] used a different approach in the coupled-cluster framework where the diagrams were generated based on topological considerations and the implementation was achieved by a general string based algorithm.

The machinery outlined above was initially developed to implement the MCPT theories. However, it can be easily adapted to any high-order multi-reference many-body PT. For example this machinery was used to implement the multi-reference multipartitioning many-body PT up to fourth-order as will be shown in Chapter 2.

In the above example we considered Davidson-Kapuy denominators. Implementation of Epstein-Nesbet denominators is also possible and has been incorporated in the code. For a given Fermi-vacuum $|J\rangle$ and a given excited determinant $|K\rangle$ the EN denominator looks

$$\Delta^K_{HF}(EN) = E_{HF} - E_J + E_J - E_K = \Delta^J_{HF}(EN) + \langle J|H|J\rangle - \langle K|H|K\rangle \qquad (1.119)$$

where $\langle J|H|J\rangle - \langle K|H|K\rangle$ can be easily constructed once the indices making the difference between $|J\rangle$ and $|K\rangle$ are known, since according to the Slater rules [66] the denominator depends only on these indices.

1.4 Numerical results

Some illustrative applications are reported in this section to show the performance of the MCPT and SC2-MCPT schemes. Our examples contain the dissociation potential curve of diatomic molecules N_2 and F_2, the insertion of a Be atom in between two H atoms to form a BeH_2 molecule and finally the distortion of the C_2H_4 molecule.

1.4.1 N_2 and F_2

Results for the N_2 molecule using 6-311G** basis are presented in Fig. 1.7. Reference functions serving as starting point of the perturbation procedures are APSG functions[45, 56]. Two orbitals were assigned to each non-core geminal, producing GVB-type[45] reference states. In these calculations we applied Møller – Pressed like energy denominators (see Eq.(1.19)) using the density matrix of the leading determinant. Corrections by MCPT and SC2-MCPT are plotted at second at third-order. Apart form the MCPT and SC2-MCPT formulations, the second-order result obtained by the PT scheme of Rosta and Surján[21] is also shown in Fig. 1.7, labeled as APSG-PT2. In this method, following Dyall's idea[17], a two-body zero-order Hamiltonian is applied, whose eigenvectors are the APSG states. A state-selective multiconfigurational coupled-cluster (SS-MRCC)[67] energy was computed at some geometries as benchmark.

The F_2 molecule was also treated in the 6-311G** basis. Here again full CI reference was not affordable. To test the perturbative results, we also computed a multi-reference average quadratic coupled-cluster (MR AQCC) [68]. Curves corresponding to the reference energies MCPT-0 and SC2-MCPT-0 are missing from the plot for purpose: they lie too far from the PT corrected lines to be displayed together in one plot. For the F_2 molecule we used simple (2,2) CAS functions as zero-order ground state reference and Davidson-Kapuy (DK) one-particle energies calculated from the leading determinant.

A zoom into the region at around equilibrium geometry is shown to illustrate the situation more clearly both for N_2 and F_2.

On the example of homonuclear diatomic dissociation curves one can observe that zero-order energies in MCPT and SC2-MCPT formulations do not differ significantly in numerical terms. Second and third-order results in the MCPT formulation lie close to each other. Third-order MCPT slightly worsens the second-order in Fig. 1.7 . On the

other hand, SC2-MCPT second and third-order curves differ significantly. Third-order SC2-MCPT is worse than second-order, showing a bump at around 2 Å. If comparing the second-order of the two MCPT formulations, we see a notable decrease in energy: SC2-MCPT2 represents a significant improvement upon the MCPT2 potential curve. The second-order SC2-MCPT2 lies very close to APSG PT2 in this example at around equilibrium. Unfortunately APSG-PT2 starts to deviate from the good shape at around 2 Å due to the quasi degenerate character of the reference function, that slowly builds up upon dissociation.

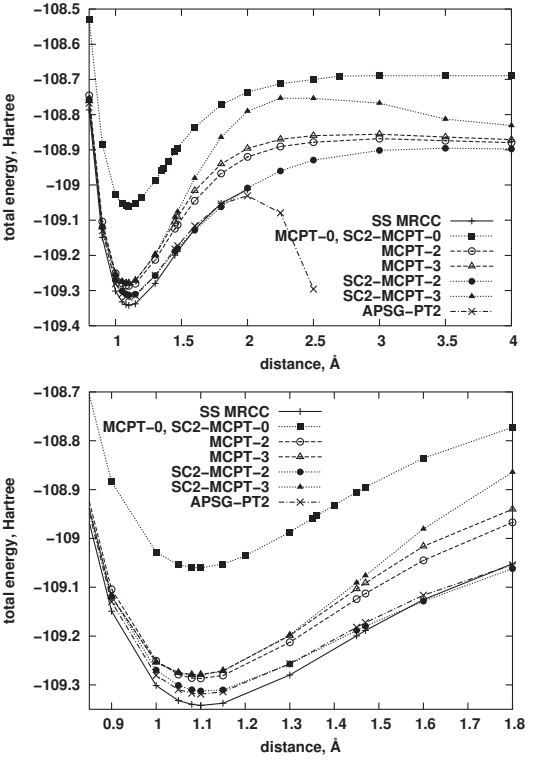
The example of the F_2 molecule shown in Fig. 1.8 is somewhat different from the case of the N_2 molecule. Here we see a rather large deviation of MCPT second and third-order results, third-order improving on both the shape and the minimum value of the second-order curve. In the case of SC2-MCPT, second-order is hard to distinguish from MR AQCC at around equilibrium, but it gets worse as the dissociation takes place. Again at difference with the example of the N_2 molecule, we see a better shaped curve at third-order than at second, though the minimum is far better at second than at third-order.

1.4.2 BeH_2

In the following example (see Fig. 1.9) application of both Davidson-Kapuy and Epstein-Nesbet denominators are tested. DK MCPT and DK SC2-MCPT acronyms are applied for calculations where the zero order energies are calculated from the generalized-Fock matrix Eq.(1.51) while EN MCPT and EN SC2-MCPT abbreviations refer to the application of Epstein-Nesbet denominators from Eq.(1.119).

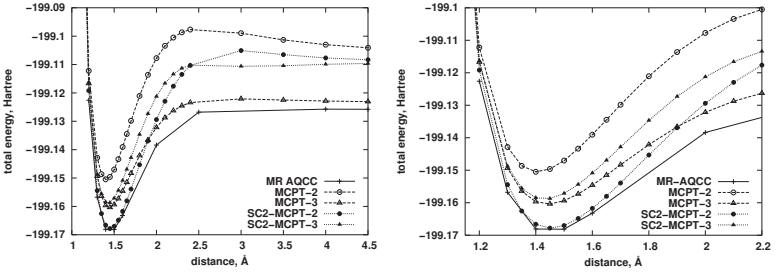
The C_{2v} insertion of a Be atom in between two H atoms with a simultaneous increase of the H-H distance presents various difficulties at different regions along the insertion path and has been a good test case of numerous multi-reference theories[36, 69, 70, 71]. Nuclear arrangements (points A to I) along this path were taken from the work of Purvis and Bartlett[72]. The Be atom is put at the origin (0,0,0), the two H atoms lie symmetric to the z axis, with coordinates in atomic units $(0, \pm 2.54, 0)$, $(0, \pm 2.08, 1.0)$, $(0, \pm 1.62, 2.0)$, $(0, \pm 1.39, 2.5)$, $(0, \pm 1.275, 2.75)$, $(0, \pm 1.16, 3.0)$ $(0, \pm 0.93, 3.5)$, $(0, \pm 0.70, 4.0)$, and finally $(0, \pm 0.70, 6.0)$ at points A, B, C, D, E, F, G, H and finally I.

Figure 1.7: Dissociation potential curve of the N_2 molecule in 6-311G** basis set as obtained by MCPT methods, by a Dyall type PT (APSG-PT [21]) and by SSMRCC. The reference function is APSG. For notations see text. (top) Potential curves displayed in a wide range of diatomic distance. (bottom) Potential curve displayed at around equilibrium geometry.



Dunning's double zeta basis set[73] was applied for the hydrogen atom, and the basis of Purvis et al.[72] was used for the Be atom with the *p* function decontracted, leaving the most compact primitive (exponent 5.693880) alone and contracting the remaining

Figure 1.8: Dissociation potential curve of the F_2 molecule in 6-311G** basis set. (a) Potential curves displayed in a wide range of diatomic distance. (b) Potential curves at around equilibrium geometry.



(a) (b)

two into a second p function (exponents 1.555630, 0.171855 and coefficients 0.144045, 0.949692 respectively). This gives a valence double zeta basis for the system with 13 orbitals altogether.

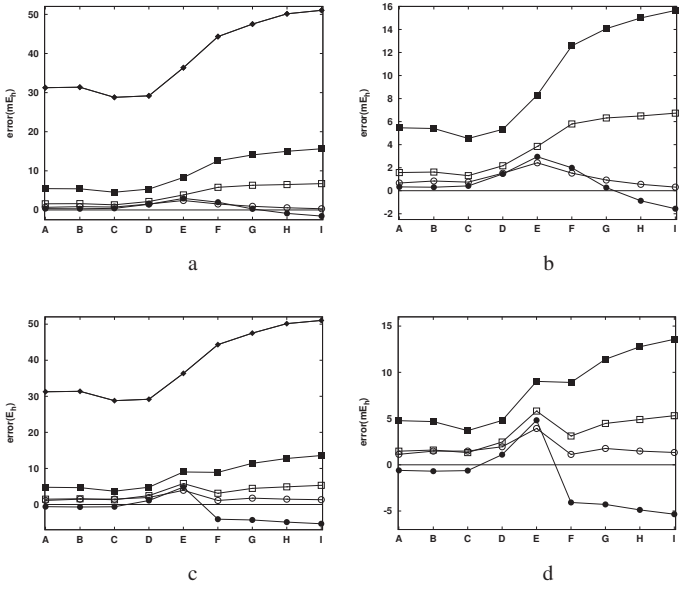
All four valence electrons were taken as active in a CAS(4,4) reference function. Two of the active orbitals (number 2 and 4) are of a_1 symmetry, the other two (number 3 and 5) belong to the b_1 irreducible.

Points D, E and F are particularly interesting from the electron correlation point of view and have been shown to be heavily affected by the intruder state problem in the effective Hamiltonian type MRPT derived from the Bloch-equation Eq.(1.6). At these points the ground state of the system is an open shell triplet, and multi-reference character of the lowest singlet wave function is increased. Major change in the nature of the wave function occurs between points E and F where closed shell determinants $|123123\rangle$ and $|124124\rangle$ change role.

The CAS results at points E,F,G,H and I give less accurate description of the electronic system thus the non-parallelity error of the zero-order results are relatively large (see Figs. 1.9 a and c). Second-order methods reduce significantly both the error of the total energy and also the non-parallelity error, although the improvement of the latter is, as usual, smaller. In the DK MCPT case both the second and the third order curves keep the shape

of the initial CAS results i.e. the error is smaller at first four points and larger at the last four points while the non-parallelity error is reduced (see Fig. 1.9.b). EN MCPT gives a surprisingly good estimation at second-order, although at points D,E and F the error at second-order still larger than at points A,B and C. The third-order gives no or just a small improvement in this case. DK SC2MCPT shows similar trends as DK MCPT with an exception at point E (see Fig. 1.9.d). The second-order EN SC2MCPT is successful at point A,B,C and D but the error is significantly larger at F,G,H and I points. There is a change in Fermi-vacuum between points E and F which explains the 10mH jump in the curve between these two points. For EN SC2MCPT third-order gives a small parallelity error.

Figure 1.9: Errors of total energies of the BeH_2 system obtained (a) MCPT, (b) MCPT around 0 mH, (c) SC2MCPT, (d) SC2-MCPT around 0 mH. Symbols correspond to: CAS (⧫), 2nd order EN (●), 2nd order DK (■), 3rd order EN (○), 3rd order DK (□). Labels A to I refer to the geometry. For coordinates and basis set see text.



1.4.3 C_2H_4

The ethylene molecule is studied in Dunning's double zeta polarized basis set[73] giving 50 basis functions for 16 electrons. We followed the potential upon distorting the dihedral angle (φ), bringing the molecule from D_{2h} symmetry (0^o) to D_{2d} (90^o) and back to D_{2h} (180^o). A relatively small, CAS(2,2) function was applied as zero order function, the two active orbitals being of symmetry b_1 and b_2, becoming degenerate at 90^o, belonging to the e irreducible of D_{2d}. Orbitals b_1 and b_2 change role at $\varphi = 90^o$ in the sense that b_1 figures in the principal determinant for φ between 0^o and 90^o while b_2 appears in the principal determinant for φ between 90^o and 180^o.

Instead of FCI, an Adamowicz type state selective MRCCSDT method[67] was computed, with two-hole, two-particle active indices as the reference. This method, which includes full triple excitations, is much more reliable than second-order PT-s.

Total energies plotted in Fig. 1.10 bracket the MRCCSDT[2+2] curve with DK partitioning results lying above and EN partitioning shooting below. All second-order PT-s bring a correction of around -300 mH to the CAS energy. The shape of PT corrected curves around the top of the barrier, and consequently barrier height varies a lot among second-order theories.

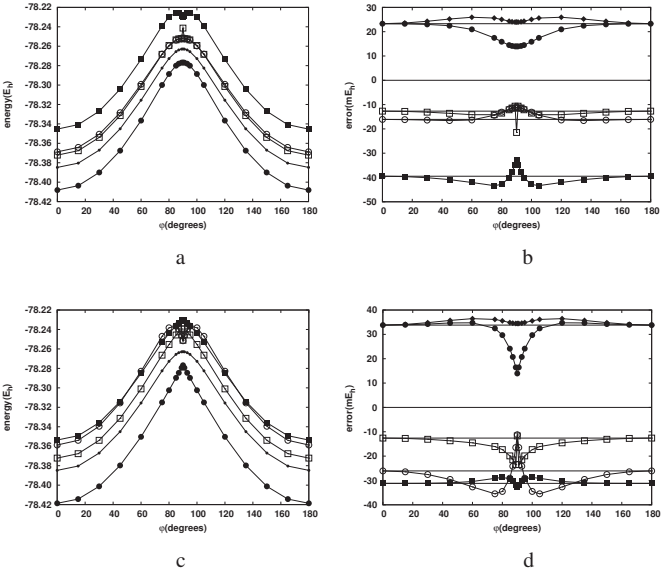
Most notable in Fig. 1.10 is that at second-order most of the MCPT methods give a completely wrong barrier shape with a cusp-like. This applies to SC2-MCPT both in DK and EN partitioning, while MCPT shows this effect in DK partitioning. The problem stems from the choice of Fermi-vacuum, and causes serious trouble, questioning the applicability of these methods for the description of the ethylene torsion process. The strange shape of the barrier top is produced by the intersection of two continuous curves of PT calculations were either the one or the other of the two dominant determinants of the CAS vector were used as the Fermi vacuum. Since neither of these curves have a maximum at D_{2d} geometry, a cusp-like intersection appears at 90^o. Apparently SC2-MCPT is more sensitive to this problem, than MCPT. MCPT in the DK partitioning proves also useless, showing a curve with double maximum. The reason behind this behavior is the orbital degeneracy at 90^o, which can be cured by the application of suitable level shifts.

Far from the 90^o dihedral angle the third-order results give significant improvement, in the DK SC2-MCPT and MCPT cases, while they inherit the singular nature of second-

order DK MCPT and SC2-MCPT at 90^o.

The only exception from the above discussed irregular behavior is the MCPT theory with EN denominators, which gives a smooth curve at $\varphi = 90^o$, similarly to the reference CAS function (see Fig. 1.10.b). Even in this case the non-parallelity error at second-order is still a magnitude larger than the non-parallelity error of the reference state but the third-order correction provides significant improvement. We note at this point that the strong dependence on the Fermi-vacuum can be alleviated by averaging over all choices for a principle dependent, as shown in [58].

Figure 1.10: Total energies of the C_2H_4 system obtained by MCPT (Fig. a) and SC2MCPT (Fig. c) and the non-parallelity error of MCPT (Fig. b) and SC2MCPT (Fig. d). For reference a MRCCSDT calculation was used (see text). Symbols correspond to: CAS (⧫), MRCCSDT (·), 2nd order EN (●), 2nd order DK (■), 3rd order EN (○), 3rd order DK (□). CAS results are shifted by -289 mH and -278 mH in Figs. c and d. Horizontal line in Fig. b and d are drawn to help to assess the non-parallelity of the curves.



1.5 Summary

In this Chapter we briefly introduced the concept of multi-reference perturbation theories and commented on their main difficulties like the intruder state problem and size-extensivity. To avoid the diagonalization of large CAS spaces and to give perturbative corrections for general zero order functions like GVB, APSG, an alternative MR PT approach is introduced. A conspicuous feature of this theory (called MCPT) is that it applies a non-Hermitian zero order Hamiltonian, which can be formulated in a diagonal form using a biorthogonal basis set. This biorthogonal basis can be easily constructed without the application of any numerical procedure, like diagonalization or inversion. Unfortunately MCPT is not size-consistent. To handle this problem we modified the formalism (SC2-MCPT) to keep the size-consistency at least at second-order in energy.

Implementation of MCPT and SC2-MCPT is performed up to the third-order. The main difficulties of an efficient implementation originate from the large number of diagrams appearing at the third-order. These difficulties were avoided by automatizing both the derivation and the coding of the corresponding diagrams. To reduce the calculation cost, factorization of diagrams was also implemented.

Performance of the second and third-order theories can be tuned using different definitions for the one partitioning. Usually both MCPT and SC2-MCPT theories give significant improvement at the second and third-order. In cases where two determinants of the zero order function are degenerate these theories can fail. The reason behind this phenomenon is twofold. On the one hand the definition of the Fermi-vacuum can change along a potential energy surface which can lead to discontinuity, on the other hand, the denominator of the reduced resolvent operator in Eq.(1.54) and Eq.(1.61) may contain zero as energy difference it belonging to a pair of degenerate zero-order determinants.

Both problems are under investigation in the laboratory of Péter Surján presently. The effective Hamiltonian theory presented in Chapter 2 represents an improvement over MCPT in difficult situations due to the absence of the principal determinant and the application of multipartitioning.

Chapter 2

Multipartitioning many-body perturbational theory (MP MBPT)

2.1 Introduction

The original MBPT framework has various appealing features worth to keep in a multi-reference extension. It is size-extensive, size-consistent, conserves the symmetry (spin and spatial) of the zero-order function and can be defined invariant to the rotation of occupied and virtual orbitals among themselves. From this point of view there is a marked difference between MR extensions of MBPT. Thanks to the linked-cluster theorem, Bloch-equation-based approaches (like QDPT) have the potential to be size-extensive. Multi-reference function based 'diagonalize then perturb' theories often violate this requirement as already discussed in Section 1.1. Effective Hamiltonian approaches (see Subsections 1.1.2 and 1.1.3) therefore seem preferable for the reason of extensivity. There is however an opposite indication: sensitivity of the theory to zero-order functions lying close in energy, the so-called intruder effect. Targeting many roots at the same time, effective Hamiltonian theories are more prone to intruders, since quasi-degeneracy of any of the model functions destroys the reliability of the approximation. To reduce the sensitivity to intruder states, attempts were taken to eliminate the high lying CAS states from the model space and apply general model space but this approach can lead to size-extensivity problem [32]. A well-established workaround for this problem is to keep the dimension of the model space, but focus on one single or just a few states of the effective Hamiltonian,

leading to state-specific or intermediate Hamiltonian theories, respectively (see Section 1.1.3).

Close to zero denominators are often avoided both in single-reference and multi-reference cases by denominator shifts [74-76]. Although this approach helps to improve the convergence of the PT series, it may lead to size-consistency problem. In diagrammatic CAS PT(D-CASPT), proposed by Finley[20] and implemented by Finley and Witek [77], orbital dependent level shifts are applied to avoid size-inconsistency.

Finley's modification of QDPT is closely related to the multipartitioning Møller-Plesset PT defined by Malrieu et al. [19, 78, 79]. In this approach the partitioning, i.e. splitting of the Hamiltonian for a zero-order and a perturbation, is changed from column to column when building the effective Hamiltonian matrix. This flexibility allows one to concentrate on just one level when defining a quasi-degeneracy avoiding zero-order and tune the partitioning accordingly. At the same time the theory remains size-consistent.

In this Chapter the multipartitioning Møller-Plesset PT will be discussed in detail. Our contribution to this field involves formal results related to symmetries and scaling properties[4, 5] of higher-order corrections as well as an efficient diagram-based fourth-order implementation that enables to treat systems of moderate size[5].

The structure of this Chapter is as it follows. First we introduce the basics of the theory in Sections 2.2.1 and 2.2.2, and show its relation to MR MBPT. To emphasize the strong connection with MR MBPT, multipartitioning many-body theory (MP MBPT) denomination will be used instead of multipartitioning Møller-Plesset PT.

It will be demonstrated that the lack of rotational invariance within degenerate orbital subspaces leads to symmetry contamination of the original theory. This was observed by Malrieu and co-workers in connection with spin-symmetry and led them to redefine the zero-order Hamiltonian[78]. In Section 2.2.3 it will be argued that symmetry breaking may affect spatial symmetry as well. We suggest a way to remove spatial symmetry contamination[4]. In Section 2.2.4 D-CASPT will be briefly discussed, pointing out its connection to the multipartitioning approach.

Sections 2.3.1 and 2.3.2 are devoted to a detailed discussion on diagrammatics, connectedness and consistency. In the original formulation of MP MBPT the zero-order Hamiltonian is defined by as many set of one-particle energies as the number of determinants in the CAS space. This is the most general formulation and it gives connected

second-order, as already shown by [19]. In Section 2.3.2 it will be shown that in higher-orders connectedness is only ensured if a fixed set of one-particle energies are applied in the construction of the partitioning dependent Hamiltonian. In this way connectedness of the theory can be kept at the third-order both in the *spin-adapted* and the *general spatial symmetry-adapted* case. At fourth-order of the *symmetry-adapted* theory some disconnected diagrams of the effective Hamiltonian are not entirely cancelled. The reason behind this unfavorable property is that both the *spin* and *symmetry-adapted* zero-order Hamiltonian can connect subsystems at infinite distance from each other. In the *spin-adapted* case, size-consistency holds, i.e. a dissociation process can be correctly described using localized orbitals. In the *symmetry adapted* case, size-consistency with localized orbitals can be fulfilled depending on whether or not the infinitely separated subsystems are identical. In Sections 2.3.1 and 2.3.2 the theoretical background of an efficient fourth-order implementation of MP MBPT is described. The implementation will shortly be presented in Section 2.3.4. Finally some promising numerical results will be shown in Section 2.4.

2.2 Hilbert space formulation

2.2.1 Multipartitioning multi-reference MBPT

A perturbational solution of the generalized Bloch-equation

$$\hat{H}\hat{\Omega}\hat{P} = \hat{\Omega}\hat{P}\hat{H}\hat{\Omega}\hat{P} \tag{2.1}$$

(where the notations and the above equation are introduced in Section 1.1.2) applying multiple partitionings was introduced by Zaitevskii and Malrieu[19]. They proposed to apply a partitioning determined by the model function appearing at the right end of the expressions. To see this, let us first consider just term $|I\rangle\langle I|$ of the rightmost projectors in Eq.(2.1)

$$\hat{H}\hat{\Omega}|I\rangle\langle I| = \hat{\Omega}\hat{P}\hat{H}\hat{\Omega}|I\rangle\langle I| \,, \tag{2.2}$$

and let us split the Hamiltonians as $\hat{H} = \hat{H}_I^{(0)} + \hat{V}_I$ to get

$$\left(\hat{H}_I^{(0)} + \hat{V}_I\right)\hat{\Omega}|I\rangle\langle I| = \hat{\Omega}\hat{P}\left(\hat{H}_I^{(0)} + \hat{V}_I\right)\hat{\Omega}|I\rangle\langle I| \,. \tag{2.3}$$

The form of $|I\rangle$ dependent $\hat{H}_I^{(0)}$ will be discussed in Section (2.2.2). Supposing that $\hat{H}_I^{(0)}$ is diagonal on the basis of states $|J\rangle$ and making use of Eq.(1.3) we can substitute $\hat{H}_I^{(0)}$ by $E_I^{(0)}$ on the right hand side of the above equation. Here $E_I^{(0)}$ denotes the eigenvalue of $\hat{H}_I^{(0)}$ corresponding to $|I\rangle$. Equation (2.3) in a rearranged form then looks

$$\left(\hat{V}_I\hat{\Omega} - \hat{\Omega}\hat{P}\hat{V}_I\hat{\Omega}\right)|I\rangle\langle I| = \left(E_I^{(0)} - \hat{H}_I^{(0)}\right)\hat{\Omega}|I\rangle\langle I| \,. \tag{2.4}$$

Multiplying this equation by the resolvent operator in the outer space, i.e. by $Q/(E_I^{(0)} - \hat{H}_I^{(0)})$ and performing the summation over the model states, one gets

$$\hat{Q}\hat{\Omega}\hat{P} = \sum_{I\in P} \frac{\hat{Q}}{E_I^{(0)} - \hat{H}_I^{(0)}} \left(\hat{V}_I\hat{\Omega} - \hat{\Omega}\hat{P}\hat{V}_I\hat{\Omega}\right)|I\rangle\langle I|. \tag{2.5}$$

This is a recursive equation for the wave operator which is solved by introducing an order by order expansion

$$\hat{\Omega}\hat{P} = (\hat{P} + \hat{\Omega}^{(1)} + \hat{\Omega}^{(2)} + \dots)\hat{P} \tag{2.6}$$

where $\hat{\Omega}^{(n)}$ involves n perturbation operators. Using the facts that $\hat{Q}\hat{P} = 0$, $\left[\hat{Q}, \hat{H}_I^{(0)}\right] = 0$ and the shorthand notations $\hat{X}_I = \hat{Q}/(E_I^{(0)} - \hat{H}_I^{(0)})$ and $\hat{P}_I = |I\rangle\langle I|$, we easily obtain from Eq.(2.5) the following low-order expressions of the wave operator

$$\begin{aligned}
\hat{\Omega}^{(1)} &= \sum_{I\in P} \hat{X}_I\hat{V}_I|I\rangle\langle I| \\
\hat{\Omega}^{(2)} &= \sum_{I\in P} \hat{X}_I \left(\hat{V}_I\hat{\Omega}^{(1)} - \hat{\Omega}^{(1)}\hat{P}\hat{V}_I\hat{P}\right)\hat{P}_I \\
&= \sum_{I\in P} \hat{X}_I\hat{V}_I\hat{X}_I\hat{V}_I\hat{P}_I - \sum_{I\in P}\hat{X}_I \sum_{J\in P} \hat{X}_J\hat{V}_J\hat{P}_J\hat{V}_I\hat{P}_I \\
\hat{\Omega}^{(3)} &= \sum_{I\in P} \hat{X}_I \left(\hat{V}_I\hat{\Omega}^{(2)} - \hat{\Omega}^{(1)}\hat{P}\hat{V}_I\hat{\Omega}^{(1)} - \hat{\Omega}^{(2)}\hat{P}\hat{V}_I\hat{P}\right)\hat{P}_I \\
&= \sum_{I\in P} \hat{X}_I\hat{V}_I\hat{X}_I\hat{V}_I\hat{X}_I\hat{V}_I\hat{P}_I - \sum_{I\in P}\hat{X}_I\hat{V}_I\hat{X}_I \sum_{J\in P} \hat{X}_J\hat{V}_J\hat{P}_J\hat{V}_I\hat{P}_I \\
&- \sum_{I\in P}\hat{X}_I \sum_{J\in P} \hat{X}_J\hat{V}_J\hat{P}_J\hat{V}_I\hat{X}_I\hat{V}_I\hat{P}_I - \sum_{I\in P}\hat{X}_I \sum_{J\in P} \hat{X}_J\hat{V}_J\hat{X}_J\hat{V}_J\hat{P}_J\hat{V}_I\hat{P}_I \\
&+ \sum_{I\in P}\hat{X}_I \sum_{J\in P}\hat{X}_J \sum_{L\in P} \hat{X}_L\hat{V}_L\hat{P}_L\hat{V}_J\hat{P}_J\hat{V}_I\hat{P}_I
\end{aligned} \tag{2.7}$$

Expansion (2.6) of the wave operator generates an order by order expansion of the effective Hamiltonian via Eq.(1.5) in the form

$$\hat{H}_{\text{eff}}^{(0)} + \hat{H}_{\text{eff}}^{(1)} + \dots = \sum_{I\in P} \hat{P}\left(\hat{H}_I^{(0)} + \hat{V}_I\right)\left(\hat{P} + \hat{\Omega}^{(1)} + \dots\right)\hat{P}_I \,. \tag{2.8}$$

Equating terms of order zero we get $\hat{H}^{(0)}_{\rm eff} = \sum_I \hat{P}\hat{H}^{(0)}_I\hat{P}_I$. Utilizing $\hat{P}\hat{\Omega}^{(n)} = 0$ for $n \neq 0$ (a consequence of Eq.(1.3) and Eq.(2.6)), at first order we obtain $\hat{H}^{(1)}_{\rm eff} = \sum_I \hat{P}\hat{V}_I\hat{P}_I$. The sum of the two lowest orders gives the usual expression

$$\hat{H}^{(0)}_{\rm eff} + \hat{H}^{(1)}_{\rm eff} = \hat{P}\hat{H}\hat{P} \, . \tag{2.9}$$

For $n \neq 0$ the general form of the effective Hamiltonian corrections are

$$\hat{H}^{(n)}_{\rm eff} = \sum_{I\in P} \hat{P}\hat{V}_I\hat{\Omega}^{(n-1)}\hat{P}_I. \tag{2.10}$$

Explicit form of the second-order correction is

$$\hat{H}^{(2)}_{\rm eff} = \sum_{I\in P} \hat{P}\hat{V}_I\hat{X}_I\hat{V}_I\hat{P}_I \, , \tag{2.11}$$

while the third and fourth-order term is given by

$$\hat{H}^{(3)}_{\rm eff} = \sum_{I\in P} \hat{P}\hat{V}_I\hat{X}_I\hat{V}_I\hat{X}_I\hat{V}_I\hat{P}_I - \sum_{I\in P} \hat{P}\hat{V}_I\hat{X}_I \sum_{J\in P} \hat{X}_J\hat{V}_J\hat{P}_J\hat{V}_I\hat{P}_I \tag{2.12}$$

$$\begin{aligned}\hat{H}^{(4)}_{\rm eff} &= \sum_{I\in P} \hat{P}\hat{V}_I\hat{X}_I\hat{V}_I\hat{X}_I\hat{V}_I\hat{X}_I\hat{V}_I\hat{P}_I - \sum_{I\in P} \hat{P}\hat{V}_I\hat{X}_I\hat{V}_I\hat{X}_I \sum_{J\in P} \hat{X}_J\hat{V}_J\hat{P}_J\hat{V}_I\hat{P}_I \\ &- \sum_{I\in P} \hat{P}\hat{V}_I\hat{X}_I \sum_{J\in P} \hat{X}_J\hat{V}_J\hat{P}_J\hat{V}_I\hat{X}_I\hat{V}_I\hat{P}_I - \sum_{I\in P} \hat{P}\hat{V}_I\hat{X}_I \sum_{J\in P} \hat{X}_J\hat{V}_J\hat{X}_J\hat{V}_I\hat{P}_J\hat{V}_I\hat{P}_I \\ &+ \sum_{I\in P} \hat{P}\hat{V}_I\hat{X}_I \sum_{J\in P} \hat{X}_J \sum_{L\in P} \hat{X}_L\hat{V}_L\hat{P}_L\hat{V}_J\hat{P}_J\hat{V}_I\hat{P}_I\end{aligned} \tag{2.13}$$

There is an intimate relation between this theory and the MR MBPT studied by Brandow, Kaldor, Bartlett and others [8, 31, 32]. In fact, MR MBPT can be viewed as a special case of the MP MBPT, where a uniform partition of the Hamiltonian is applied, independent of the reference state:

$$\hat{H} = \hat{H}^{(0)} + \hat{V}. \tag{2.14}$$

Omitting the I dependence of $\hat{H}^{(0)}$ in operator $\hat{X}_I$ leads to operator $\hat{R}_I = \hat{Q}/(E^{(0)}_I - \hat{H}^{(0)})$, already introduced in Eq.(1.15). Terms of the effective Hamiltonian in MR MBPT(see Eq.(1.15) and Eq.(1.7)) can be obtained by replacing all $\hat{X}_I$ by $\hat{R}_I$ in Eq.(2.11), Eq.(2.12) and Eq.(2.13).

2.2.2 Zero order Hamiltonian

Møller-Plesset partitioning for MR PT is straightforward to apply in a simplest manner, provided that one defines a one-particle energy set. Several experiments on different definitions for orbital energies had the common experience that it is hard to make a suitable choice [19,78-81]. Close to zero energy denominators notoriously appear and cause singularities on potential energy surfaces. Application of level-shifts to avoid this problem is prevalent in MR theories[76] just like in the SR framework[74, 75]. Freed and coworkers suggested a remedy by forcing the degeneracy of valence-energy levels[82]. An alternative solution is to let orbital energies depend on the actual model space determinant under consideration. This determinant the one which appears at the rightmost of the wave operator or the effective Hamiltonian expression. This is the key idea leading to the introduction of multiple partitionings. The corresponding zero-order Hamiltonian can be written in the form[19]

$$\hat{H}_I^{(0)} = \sum_{i\in I} \epsilon_i^{\oplus}(I)\, \hat{\imath}^+\hat{\imath} + \sum_{a\notin I} \epsilon_a^{\ominus}(I)\, \hat{a}^+\hat{a}^- \ , \tag{2.15}$$

where i is occupied and a is not occupied in determinant $|I\rangle$. Notation $\oplus$ and $\ominus$ in superscript is used to stress that physical interpretation of orbital energies as ionization potentials (IP) for the levels occupied in $|I\rangle$ and electron affinities (EA) for the levels unoccupied in $|I\rangle$ are kept from SR theory. The intention applying generalized IPs and EAs as orbital energies is to preserve the large gap between them which is provided in well-behaving SR MBPT cases. It is not necessary but straightforward to define orbital energies as Koopmans-like IPs and EAs in the form[19]

$$\begin{aligned} \epsilon_i^{\oplus}(I) &= \langle I|\hat{H}|I\rangle - \langle I|\hat{\imath}^+\hat{H}\hat{\imath}^-|I\rangle, \\ \epsilon_a^{\ominus}(I) &= \langle I|\hat{a}^-\hat{H}\hat{a}^+|I\rangle - \langle I|\hat{H}|I\rangle \ . \end{aligned} \tag{2.16}$$

Note that $\epsilon_a^{\oplus}(I)$ for a level unoccupied in $|I\rangle$ as well as $\epsilon_i^{\ominus}(I)$ for a level occupied in $|I\rangle$ is zero. With the zero order Hamiltonian of Eq.(2.15) denominators of the PT expressions have the form

$$\langle I|\hat{H}_I^{(0)}|I\rangle - \langle J|\hat{H}_I^{(0)}|J\rangle = \sum_i{}' \epsilon_i^{\oplus}(I) - \sum_i{}' \epsilon_a^{\ominus}(I) \tag{2.17}$$

where the IP of orbital i appears if an electron is removed from it when stepping from determinant $|I\rangle$ to $|J\rangle$ while the EA comes into play for an orbital a where an electron is

put upon exciting $|I\rangle$ to get $|J\rangle$. By definition Eq.(2.16) IPs and EAs are well separated in energy only if the corresponding determinant $|I\rangle$ is important in the low-lying roots of the complete active space (CAS) Hamiltonian. For highly excited determinants of the model space the gap between IPs and EAs may unfortunately disappear. For this reason, definition Eq.(2.16) is unpractical in the general case. The zero-order Hamiltonian of Eq.(2.15) has the further drawback of violating extensivity from the third-order as discussed in Section 2.3.2. We will refer to the MP MBPT formalism using the zero-order Hamiltonian of Eq.(2.15) as *genuine* MP MBPT .

Both of the above problems can be cured by a simplified formulation where only one set of IPs and one set of EAs are defined according to[79]

$$\epsilon_i^{\oplus_\mu} = \frac{\langle\Psi_\mu|\hat{H}|\Psi_\mu\rangle}{\langle\Psi_\mu|\Psi_\mu\rangle} - \frac{\langle\Psi_\mu|\hat{\imath}^+\hat{H}\hat{\imath}^-|\Psi_\mu\rangle}{\langle\Psi_\mu|\hat{\imath}^+\hat{\imath}^-|\Psi_\mu\rangle},$$

$$\epsilon_a^{\ominus_\mu} = \frac{\langle\Psi_\mu|\hat{a}^-\hat{H}\hat{a}^+|\Psi_\mu\rangle}{\langle\Psi_\mu|\hat{a}^-\hat{a}^+|\Psi_\mu\rangle} - \frac{\langle\Psi_\mu|\hat{H}|\Psi_\mu\rangle}{\langle\Psi_\mu|\Psi_\mu\rangle}\ . \tag{2.18}$$

This definition is an MR generalization of IPs and EAs. Expectation values are taken with the multi-reference CAS vector Ψ_μ in which we are primarily interested. Being dependent on index μ, the above expressions introduce a state-specific character in the theory. If more states at a time are targeted, orbital energies can be defined by averaging the above quantities with the formula[79]

$$\epsilon_i^{\oplus} = \frac{\sum\limits_\mu \epsilon_i^{\oplus_\mu}\langle\Psi_\mu|\hat{\imath}^+\hat{\imath}^-|\Psi_\mu\rangle}{\sum\limits_\mu\langle\Psi_\mu|\hat{\imath}^+\hat{\imath}^-|\Psi_\mu\rangle} \tag{2.19}$$

for IPs and analogously for EAs. Either definitions ensure considerable energy gap between IPs and EAs, provided that neither of states Ψ_μ is a high lying root of the CAS problem in energy. By both definitions the corresponding zero-order Hamiltonian is of the form

$$\hat{H}_I^{(0)} = \sum_{i\in I}\epsilon_i^{\oplus}\,\hat{\imath}^+\hat{\imath}^- + \sum_{a\notin I}\epsilon_a^{\ominus}\,\hat{a}^+\hat{a}^-\ . \tag{2.20}$$

We will use the designation *simple* MP MBPT for the method based on Eq.(2.20) as a zero-order. Note that the multipartitioning feature still remains when using Eq.(2.20) since the summations are restricted by determinant $|I\rangle$. We will see in Section 2.3.2 that the theory based on Eq.(2.20) is extensive.

2.2.3 Spin and symmetry considerations

The Hamiltonian of a system shows spin symmetry and may show spatial symmetry. Eigenvalues of the electronic Hamiltonian can be classified according to the irreducible representations of the molecular spin and spatial symmetry group. In particular we are interested in transformations which rotate elements of a multidimensional irreducible (e.g. e, t, etc. representations in spatial symmetry or doublet, triplet etc. representations in spin symmetry) among each other. These transformations leave the exact eigenvalue intact, a consequence of the fact that the Hamiltonian commutes with such an operator. A good approximation method should have the same property. In this section we investigate whether multipartitioning MBPT is well-behaving in this respect.

Let us first consider a general symmetry operator which rotates one-electron orbitals. Supposing that the orbitals are *symmetry-adapted*, this operation only transforms those which are degenerate, i.e. orbitals belonging to one given e.g. e representation or spin-orbitals which have the same spatial part. This operator can be written as

$$\hat{U} = \sum_s \sum_{k,l \in s} U_{kl}\, \hat{k}^+ \hat{l}^- \tag{2.21}$$

where s runs over degenerate orbital subspaces (i.e. $1a_1$ and $2a_1$ are considered separately), matrix U_{kl} is unitary and k, l refer to spin orbitals. Considering the commutation of operator $\hat{U}$ with the zero order Hamiltonian of Eq.(2.20) one obtains by simple second-quantized algebra

$$\begin{aligned} \left[\hat{U}, \hat{H}_I^{(0)}\right] \;=\; & \sum_s \sum_{k,l \in s} U_{kl} \left[N_k^I N_l^I \left(\epsilon_l^{\oplus} - \epsilon_k^{\oplus}\right) \hat{k}^+ \hat{l}^- \;+\; (1 - N_k^I)(1 - N_l^I) \left(\epsilon_l^{\ominus} - \epsilon_k^{\ominus}\right) \hat{k}^+ \hat{l}^- \right. \\ & + \left. N_k^I (1 - N_l^I) \left(\epsilon_l^{\ominus} - \epsilon_k^{\oplus}\right) \hat{k}^+ \hat{l}^- \;+\; (1 - N_k^I) N_l^I \left(\epsilon_l^{\oplus} - \epsilon_k^{\ominus}\right) \hat{k}^+ \hat{l}^- \right] . \end{aligned} \tag{2.22}$$

Notation N_k^I is introduced for occupation number of orbital k in determinant $|I\rangle$. The above result tells, that a good zero-order Hamiltonian should have uniform IPs in degenerate subspaces as well as uniform EAs, moreover IPs and EAs should be equal. In other words, multipartitioning is to be abandoned in degenerate orbital subspaces.

As for spin-symmetry, the above result means that dependence of ϵ's on the spin-index is to be removed in every respect. This can be achieved by formulating the zero-order order Hamiltonian in terms of spin-summed second-quantized $\hat{E}$ operators

$$\hat{E}_i^i \;=\; \hat{i}_\alpha^+ \hat{i}_\alpha^- \;+\; \hat{i}_\beta^+ \hat{i}_\beta^- \tag{2.23}$$

e.g. in the form

$$\hat{H}_J^{(0)} = \sum_{\substack{p \\ \text{2x occ in } |J\rangle}} \epsilon_p^{\oplus} \hat{E}_p^p + \sum_{\substack{q \\ \text{1x occ in } |J\rangle}} \epsilon_q^{\otimes} \hat{E}_q^q + \sum_{\substack{r \\ \text{empty in } |J\rangle}} \epsilon_r^{\ominus} \hat{E}_r^r \,. \tag{2.24}$$

This definition has two shortcomings concerning the half-filled orbitals in $|J\rangle$: (i) it applies orbital energies which depend on determinant $|J\rangle$, whatever definition one takes for $\epsilon_q^{\otimes}$; (ii) it destroys the multipartitioning character. Point (i) is problematic from the point of view of extensivity, as discussed in Section 2.3.2. Point (ii) is a shortcoming because the potential of the theory to be robust against intruders is lost this way.

To conserve the multipartitioning feature and ensure spin-symmetry at the same time, Zaitevskii and Malrieu proposed the following zero-order Hamiltonian[78]:

$$\hat{H}_{\text{sp}}^{(0)}(J) = \sum_{\substack{p \\ \text{2x occ in } |J\rangle}} \epsilon_p^{\oplus} \hat{E}_p^p + \sum_{\substack{q \\ \text{1x occ in } |J\rangle}} \left(\epsilon_q^{\oplus} \hat{E}_q^q + \frac{1}{2} \left(\epsilon_q^{\ominus} - \epsilon_q^{\oplus} \right) \hat{E}_q^q (\hat{E}_q^q - 1) \right) + \sum_{\substack{r \\ \text{empty in } |J\rangle}} \epsilon_r^{\ominus} \hat{E}_r^r \,. \tag{2.25}$$

Although this operator has two electron terms at difference with Eq.(2.20), it conserves the property that pure determinants are its eigenfunctions:

$$\hat{H}_{\text{sp}}^{(0)}(J) = \sum_L \left[\sum_{\substack{p \text{ occ in } |L\rangle \\ N_p^L \leq N_p^J}} N_p^L \epsilon_p^{\oplus} + \sum_{\substack{p \text{ occ in } |L\rangle \\ N_p^L > N_p^J}} \left(N_p^J \epsilon_p^{\oplus} + (N_p^L - N_p^J) \epsilon_p^{\ominus} \right) \right] |L\rangle\langle L| \tag{2.26}$$

with N_p^L denoting the occupancy of spatial orbital p in determinant $|L\rangle$. The reason for the simple structure Eq.(2.26) is the simple form of the two-electron operator in Eq.(2.25), which is in fact just the square of a spin-summed particle-number operator. Expression (2.25) is apparently spin-adapted due to the application of $\hat{E}$ operators.

Using definition (2.25), a spatial orbital q which is half filled in determinant $|J\rangle$ but doubly occupied in determinant $|L\rangle$ contributes an EA, i.e. $-\epsilon_q^{\ominus}$ to the energy denominator $\langle J|\hat{H}_{\text{sp}}^{(0)}(J)|J\rangle - \langle L|\hat{H}_{\text{sp}}^{(0)}(J)|L\rangle$. On the other hand if orbital q is unoccupied in determinant $|L\rangle$, its contribution to the energy denominator is $\epsilon_q^{\oplus}$. The fact that – depending on its role in the excitation – singly occupied orbital q contributes either EA or IP to the denominator is the consequence of multipartitioning. In general the energy denominator when using Eq.(2.25) can be expressed as

$$\langle J|\hat{H}_{\text{sp}}^{(0)}(J)|J\rangle - \langle L|\hat{H}_{\text{sp}}^{(0)}(J)|L\rangle = \sum_{p \text{ if } N_p^L < N_p^J} \left(N_p^J - N_p^L \right) \epsilon_p^{\oplus} - \sum_{r \text{ if } N_r^L > N_r^J} \left(N_r^L - N_r^J \right) \epsilon_r^{\ominus} \,. \tag{2.27}$$

We now turn our attention to the question of general symmetry, i.e. combined spatial and spin-symmetry[4]. Based on the result of Eq.(2.22) neither of the zero-order Hamiltonians discussed so far show this property. We are going to start from Eq.(2.25) and keep the essence of Eq.(2.27). This means that regarding a degenerate subspace say s, the denominator will depend only on the change in occupancy of the given subspace. In particular, subspace s will contribute as many $\epsilon_s^{\oplus}$ as the number of electrons removed from s, and as many $-\epsilon_s^{\ominus}$ as the number of electrons put into the subspace.

The starting point of the following derivation is the observation that the second term in the round brackets in Eq.(2.25) has a role only if orbital s is doubly occupied in $|L\rangle$ (and singly occupied in $|J\rangle$). Let us consider a general M fold degenerate orbital subspace called s. Analogously to Eq.(2.23) let us introduce operator $\hat{E}_s$ in the form

$$\hat{E}_s \;=\; \sum_{k\in s} \hat{k}^{+}\hat{k}^{-} \tag{2.28}$$

since we will be interested only in overall occupancy of subspace s. Index k refers to spin-orbitals. Taking the case where this subspace is occupied with $M-1$ electrons, the corresponding term of the zero-order operator is analogous to the half-filled term of Eq.(2.25)

$$\epsilon_s^{\oplus}\hat{E}_s - (\epsilon_s^{\oplus} - \epsilon_s^{\ominus})\,\frac{\hat{E}_s\left(\hat{E}_s-1\right)\ldots\left(\hat{E}_s-M+1\right)}{M!}\,. \tag{2.29}$$

Considering the action of the above operator on determinant $|L\rangle$, the second term has a role only if $N_s^L = M$ (i.e. $|L\rangle$ contains a subspace s filled by M electrons). In this case the contribution of the whole term to the eigenvalue is $(M-1)\epsilon_s^{\oplus} + \epsilon_s^{\ominus}$. In all other cases the above expression will contribute $N_s^L\epsilon_s^{\oplus}$ to the eigenvalue.

It is apparent that different occupancies of subspace s require similar considerations. Take the case where $N_s^J = M-2$. The corresponding term of the Hamiltonian is

$$\epsilon_s^{\oplus}\hat{E}_s - (\epsilon_s^{\oplus} - \epsilon_s^{\ominus})\,\frac{\hat{E}_s\left(\hat{E}_s-1\right)\ldots\left(\hat{E}_s-M+2\right)}{M!}\Big[2(\hat{E}_s-M+1) \;-\; M(\hat{E}_s-M)\Big]\,. \tag{2.30}$$

The second term now takes care for the cases where the occupancy in $|L\rangle$ exceeds N_s^J by one or by two. If $N_s^L = M$ then we get $(M-2)\epsilon_s^{\oplus} + 2\epsilon_s^{\ominus}$, the factor of two in (2.30) is necessary to have two EAs. If $N_s^L = M-1$ then the contribution to the eigenvalue is $(M-2)\epsilon_s^{\oplus} + \epsilon_s^{\ominus}$. In all other cases only the first term has nonzero effect and produces $N_s^L\epsilon_s^{\oplus}$.

For clarity we take down the $N_s^J = M-3$ case as well, which generates the following term in the zero-order operator:

$$\epsilon_s^{\oplus}\hat{E}_s - (\epsilon_s^{\oplus} - \epsilon_s^{\ominus})\,\frac{\hat{E}_s\left(\hat{E}_s-1\right)\ldots\left(\hat{E}_s-M+3\right)}{M!}\Big[3(\hat{E}_s-M+2)(\hat{E}_s-M+1) \\ -2M(\hat{E}_s-M+2)(\hat{E}_s-M) \;+\; \frac{M(M-1)}{2}(\hat{E}_s-(M-1))(\hat{E}_s-M)\Big]\,. \qquad (2.31)$$

At this point the structure of the terms allows us to deduce the general form of the symmetry-adapted zero-order Hamiltonian:

$$\hat{H}_J^{(0)} = \sum_{M}^{2,4,\ldots}\sum_{F=0}^{M}\sum_{\substack{s\\ N_s^J=M-F}}\left\{\epsilon_s^{\oplus}\hat{E}_s - (\epsilon_s^{\oplus}-\epsilon_s^{\ominus})\,\frac{1}{M!}\sum_{G=0}^{F-1}(F-G)(-1)^G\binom{M}{G}\prod_{\substack{R=0\\ R\neq G}}^{M}\left(\hat{E}_s-M+R\right)\right\} \qquad (2.32)$$

It can be easily verified that this expression reduces to the simpler form Eq.(2.25) if $M=2$. We refer to the method using the zero-order Hamiltonian Eq.(2.32) as *symmetry-adapted* MP MBPT . It is interesting to observe the formal appearance of M-electron operators, since in general M can be larger than two (e.g. if the spatial degeneracy is two-fold then the corresponding dimension is 4, spins included). Zero order Hamiltonian (2.32) is however still simple in the sense that it is diagonal on the basis of determinants:

$$\hat{H}_J^{(0)} \;=\; \sum_s\left(\sum_{\substack{L\\ N_s^L\leq N_s^J}} N_s^L\epsilon_s^{\oplus}|L\rangle\langle L| + \sum_{\substack{L\\ N_s^L>N_s^J}}(N_s^J\epsilon_s^{\oplus} + \left(N_s^L-N_s^J\right)\epsilon_s^{\ominus})|L\rangle\langle L|\right). \qquad (2.33)$$

The many-electron operator feature of Eq.(2.32) therefore is to be markedly distinguished from the partitionings in MR PT, where genuine two-particle interaction is involved at zero-order[17, 21].

As it will be shown in Subsection.2.3.2, using the diagonal form of $\hat{H}_J^{(0)}$ one can modify the perturbation formulae by changing the $\hat{H}_J^{(0)}$ and $\hat{V}_J^N$. As a result of this, terms into $\hat{H}_J$ and only the denominators refer to the actual form of the zero-order Hamiltonian. In the symmetry adapted case these denominators are

$$\langle J|\hat{H}_J^{(0)}|J\rangle - \langle L|\hat{H}_J^{(0)}|L\rangle \;=\; \sum_{s\text{ if }N_s^L<N_s^J}\left(N_s^J-N_s^L\right)\epsilon_s^{\oplus} - \sum_{s\text{ if }N_s^L>N_s^J}\left(N_s^L-N_s^J\right)\epsilon_s^{\ominus}. \qquad (2.34)$$

2.2.4 Diagrammatic CASPT

Finley's D-CASPT [20, 77] gives a solution to the intruder state problem which is similar to Malrieu's multipartitioning approach. In D-CASPT orbital dependent denominator shifts Δ_i^a and Δ_{ij}^{ab} are introduced into the QDPT first order wave operator (1.21) as follows:

$$\hat{\Omega}^{(1)} = \sum_{I\in P} \left\{ \sum_{\substack{ai \\ \hat{a}^+\hat{i}^-|I\rangle \in Q}} \frac{h_{ai}}{\epsilon_i - \epsilon_a + \Delta_i^a} \hat{a}^+\hat{i}^- + \frac{1}{2} \sum_{\substack{a<b;i<j \\ \hat{a}^+\hat{b}^+\hat{i}^-\hat{j}^-|I\rangle \in Q}} \frac{\langle ab|ji\rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b + \Delta_{ij}^{ab}} \hat{a}^+\hat{b}^+\hat{i}^-\hat{j}^- \right\} |I\rangle\langle I| \tag{2.35}$$

To get a reliable definition for the denominator shifts, Finley started from Hirao's state-specific formalism (1.24) where a shift $\Delta_I = E_k^{(0)} - E_I^{(0)}$ is introduced in the denominator:

$$\hat{\Omega}^{(1)} = \sum_{I\in P} \left\{ \sum_{\substack{ai \\ \hat{a}^+\hat{i}^-|I\rangle \in Q}} \sum_k C_I^k \frac{h_{ai}}{\epsilon_i - \epsilon_a + \Delta_I} \hat{a}^+\hat{i}^- + \frac{1}{2} \sum_{\substack{a<b;i<j \\ \hat{a}^+\hat{b}^+\hat{i}^-\hat{j}^-|I\rangle \in Q}} \sum_k C_I^k \frac{\langle ab|ji\rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b + \Delta_I} \hat{a}^+\hat{b}^+\hat{i}^-\hat{j}^- \right\} |I\rangle\langle \Phi_k^{(0)}| \tag{2.36}$$

Appearance of term Δ_I in the above formula is responsible for the intruder free behavior. To benefit from this experience, denominator shifts Δ_i^a and Δ_{ij}^{ab} are chosen similarly to Δ_I by applying a weighted sum, according to:

$$\Delta_i^a = \frac{1}{n_i^a} \sum_I (2 - N_a^I) N_i^I \Delta_I C_I^2 \tag{2.37}$$

$$\Delta_{ij}^{ab} = \frac{1}{n_{ij}^{ab}} \sum_I (2 - N_a^I)(2 - N_b^I) N_i^I N_j^I \Delta_I C_I^2 \quad , \tag{2.38}$$

where N_a^I is the occupation of spatial orbital a in determinant $|I\rangle$ ($N_a^I = 0, 1, 2$), C_I is the coefficient of determinant $|I\rangle$ in the reference state and finally

$$n_i^a = \sum_I (2 - N_a^I) N_i^I C_I^2, \tag{2.39}$$

$$n_{ij}^{ab} = \sum_I (2 - N_a^I)(2 - N_b^I) N_i^I N_j^I C_I^2 \tag{2.40}$$

are normalization factors. For an open shell orbital the spatial index can be both occupied and virtual. To avoid spin contamination, the denominator shift is defined so that the spin flip on an open shell gives no contribution to the denominator, i.e. $\Delta_{im}^{am} = \Delta_{mi}^{am} = \Delta_i^a$. This is similar to *spin adapted* MP MBPT.

Having a look at Eqs.(2.35) and comparing with (2.7) and Eq.(2.17) it is seen that with a proper definition of Δ_i^a and Δ_{ij}^{ab}, second-order D-CASPT and MP MBPT can be related:

$$\epsilon_i^{\oplus} - \epsilon_a^{\ominus} = \epsilon_i - \epsilon_a + \Delta_i^a \tag{2.41}$$

$$\epsilon_i^{\oplus} + \epsilon_j^{\oplus} - \epsilon_a^{\ominus} - \epsilon_b^{\ominus} = \epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b + \Delta_{ij}^{ab}. \tag{2.42}$$

Due to the formal equivalence with MP MBPT second-order MP MBPT [78], second-order D-CASPT is also extensive.

2.3 Formulation in terms of diagrams

2.3.1 Normal ordering of operators

To introduce the diagrammatic representation first we address the question of normal ordering. Following Kaldor[31] the Fermi vacuum is chosen to be the determinant which closes the perturbational formulae from the right.

In MR MBPT[32] one can simply exchange $\hat{V}$ for $\hat{V}_I^N$ in the effective Hamiltonian expressions, with $\hat{V}_I^N = \hat{V} - \langle I|\hat{V}|I\rangle$ denoting the perturbation operator, normal ordered according to determinant $|I\rangle$. Apart from $\hat{V}_I^N$, formulae (2.12) and (2.13) contain operators $\hat{V}_J^N$ and $\hat{V}_L^N$ also. To use the diagrammatic approach, operators in PT formulae should be normal ordered by the same Fermi-vacuum. Hence each appearance of $\hat{V}_J^N$ or $\hat{V}_L^N$ is to be

rewritten for $\hat{V}_I^N$ in Eqs.(2.12) and (2.13). Since both $\hat{H}_J^{(0)}$ and $\hat{H}_I^{(0)}$ are diagonal on the basis of states $|I\rangle$, one can write

$$\hat{X}_J \hat{V}_J^N \hat{P}_J = \hat{X}_J \hat{H}_J^N \hat{P}_J = \hat{X}_J \hat{H} \hat{P}_J = \hat{X}_J \hat{V}_I^N \hat{P}_J \, . \tag{2.43}$$

Simplicity of the above expression is due to the fact that diagonal part of operators $\hat{V}_I^N$ or $\hat{V}_J^N$ do not contribute. In the general case $\hat{V}_L^N$ can be changed into $\hat{V}_I^N$ by the relation $\hat{V}_J^N = \hat{V}_I^N + \hat{H}_I^{(0)N} - \hat{H}_J^{(0)N}$, where $\hat{H}_I^{0N}$ is the zero order Hamiltonian, normal ordered according to determinant $|I\rangle$. Since the form of $\hat{V}_I^N$ itself may be rather complicated we further reorganize the expressions so that the zero-order Hamiltonian does not appear in the numerator of the expressions. This can be easily achieved by applying the formula

$$\hat{V}_I^N = \hat{H}_I^N - \hat{H}_I^{(0)N} \tag{2.44}$$

where $\hat{H}_I^{(0)N}$ is diagonal. The second-order effective Hamiltonian trivially becomes:

$$\hat{H}_{\text{eff}}^{(2)} = \sum_{I \in P} \hat{P} \hat{H}_I^N \hat{X}_I \hat{H}_I^N \hat{P}_I \tag{2.45}$$

In the third-order, by using Eq.(2.44) and Eq.(2.43), expression (2.12) takes the form:

$$\begin{aligned} \hat{H}_{\text{eff}}^{(3)} &= \sum_{I \in P} \hat{P} \hat{H}_I^N \hat{X}_I \hat{H}_I^N \hat{X}_I \hat{H}_I^N \hat{P}_I - \sum_{I \in P} \hat{P} \hat{H}_I^N \hat{X}_I \sum_{J \in P} \hat{X}_J \hat{H}_I^N \hat{P}_J \hat{H}_I^N \hat{P}_I \\ &- \sum_{I \in P} \hat{P} \hat{H}_I^N \hat{X}_I \hat{H}_I^{(0)N} \hat{X}_I \hat{H}_I^N \hat{P}_I + \sum_{I \in P} \hat{P} \hat{H}_I^N \hat{X}_I \sum_{J \in P} \hat{X}_J \hat{H}_I^N \hat{P}_J \hat{H}_I^{(0)N} \hat{P}_I \end{aligned} \tag{2.46}$$

The last term is zero since $\hat{H}_I^{(0)N}$ is normal ordered and diagonal. In the third term relation

$$\hat{H}_I^{(0)N} \hat{X}_I = (\hat{H}_I^{(0)} - \langle I | \hat{H}_I^{(0)} | I \rangle) \hat{X}_I = -\hat{Q} \tag{2.47}$$

can be used to reveal that the third term is simply the second-ordered effective Hamiltonian correction $\hat{H}_{\text{eff}}^{(2)}$. The working equation of the third-order Hamiltonian correction finally becomes

$$\hat{H}_{\text{eff}}^{(3)} = \sum_{I \in P} \hat{P} \hat{H}_I^N \hat{X}_I \hat{H}_I^N \hat{X}_I \hat{H}_I^N \hat{P}_I - \sum_{I \in P} \hat{P} \hat{H}_I^N \hat{X}_I \sum_{J \in P} \hat{X}_J \hat{H}_I^N \hat{P}_J \hat{H}_I^N \hat{P}_I + \hat{H}_{\text{eff}}^{(2)} \, . \tag{2.48}$$

In the above derivation the only information we used about the zero-order Hamiltonian was its diagonal form. The formulae of the effective Hamiltonian are therefore applicable, irrespective of the actual choice for $\hat{H}_I^{(0)}$ (e.g. genuine, simple etc.), as far as it is diagonal

on the basis of determinants. Specification of the diagonal elements of $\hat{H}_I^{(0)}$ affects only the values of denominators. Using the same procedure as explained above, the fourth-order effective Hamiltonian correction is first rewritten by making use of relations $\hat{P}\hat{V}_I^N\hat{X}_I = \hat{P}\hat{H}_I^N\hat{X}_I$ and $\hat{P}_J\hat{V}_I^N\hat{P}_I = \hat{P}_J\hat{H}_I^N\hat{P}_I$ in Eq.(2.13):

$$\begin{aligned}
\hat{H}_{\text{eff}}^{(4)} &= \sum_{I\in P}\hat{P}\hat{H}_I^N\hat{X}_I\hat{V}_I^N\hat{X}_I\hat{V}_I^N\hat{X}_I\hat{H}_I^N\hat{P}_I - \sum_{I\in P}\hat{P}\hat{H}_I^N\hat{X}_I\hat{V}_I^N\hat{X}_I\sum_{J\in P}\hat{X}_J\hat{H}_I^N\hat{P}_J\hat{H}_I^N\hat{P}_I \\
&- \sum_{I\in P}\hat{P}\hat{H}_I^N\hat{X}_I\sum_{J\in P}\hat{X}_J\hat{H}_I^N\hat{P}_J\hat{H}_I^N\hat{X}_I\hat{H}_I^N\hat{P}_I - \sum_{I\in P}\hat{P}\hat{H}_I^N\hat{X}_I\sum_{J\in P}\hat{X}_J\hat{V}_J^N\hat{X}_J\hat{H}_I^N\hat{P}_J\hat{H}_I^N\hat{P}_I \\
&+ \sum_{I\in P}\hat{P}\hat{H}_I^N\hat{X}_I\sum_{J\in P}\hat{X}_J\sum_{L\in P}\hat{X}_L\hat{H}_I^N\hat{P}_L\hat{H}_J^N\hat{P}_J\hat{H}_I^N\hat{P}_I. \qquad (2.49)
\end{aligned}$$

In a second step Eqs.(2.43), (2.47) and

$$\hat{H}_J^N = \hat{H}_I^N + \langle I|\hat{H}|I\rangle - \langle J|\hat{H}|J\rangle \qquad (2.50)$$

are applied, lower-order terms like $\hat{H}_{\text{eff}}^{(2)}$ and $\hat{H}_{\text{eff}}^{(3)}$ are collected and finally the terms containing the scalar $\langle I|\hat{H}|I\rangle - \langle J|\hat{H}|J\rangle$ cancel to get:

$$\begin{aligned}
\hat{H}_{\text{eff}}^{(4)} &= \sum_{I\in P}\hat{P}\hat{H}_I^N\hat{X}_I\hat{H}_I^N\hat{X}_I\hat{H}_I^N\hat{X}_I\hat{H}_I^N\hat{P}_I - \sum_{I\in P}\hat{P}\hat{H}_I^N\hat{X}_I\hat{H}_I^N\hat{X}_I\sum_{J\in P}\hat{X}_J\hat{H}_I^N\hat{P}_J\hat{H}_I^N\hat{P}_I \\
&- \sum_{I\in P}\hat{P}\hat{H}_I^N\hat{X}_I\sum_{J\in P}\hat{X}_J\hat{H}_I^N\hat{P}_J\hat{H}_I^N\hat{X}_I\hat{H}_I^N\hat{P}_I - \sum_{I\in P}\hat{P}\hat{H}_I^N\hat{X}_I\sum_{J\in P}\hat{X}_J\hat{H}_I^N\hat{X}_J\hat{H}_I^N\hat{P}_J\hat{H}_I^N\hat{P}_I \\
&+ \sum_{I\in P}\hat{P}\hat{H}_I^N\hat{X}_I\sum_{J\in P}\hat{X}_J\sum_{L\in P}\hat{X}_L\hat{H}_I^N\hat{P}_L\hat{H}_I^N\hat{P}_J\hat{H}_I^N\hat{P}_I + 2\hat{H}_{\text{eff}}^{(3)} - \hat{H}_{\text{eff}}^{(2)}\,. \qquad (2.51)
\end{aligned}$$

Equations (2.45), (2.48) and (2.51) are the working equations of our implementation. These equations support an efficient implementation, because operators are normal ordered with respect to the same determinant and the zero-order does not figure in the numerator. The latter point is important since this way we do not have to consider contractions of the complicated symmetry-adapted zero-order Hamiltonian of Eq.(2.15). Equations (2.45), (2.48) and (2.51) hold for any zero-order that is diagonal on the basis of determinants, hence also for MR MBPT. Difference between the different methods is hidden in the denominators.

2.3.2 Diagrammatics

Diagrammatic approach of many-body theory for the single reference case has been already discussed in Section 1.3.1. Generalization for the MR case is possible and was

discussed in detail in the context of MR MBPT in, e.g., [83]. As it will be shown, the main differences as compared to the single determinantal case is that the off-diagonal matrix elements of the effective Hamiltonian are represented by open diagrams and the renormalization terms generate new type of diagrams. These results can be also applied in the MP MBPT framework with some modifications in connection with the calculation of the denominator.

A matrix element of the effective Hamiltonian taken with determinants is written e.g. for the second-order as:

$$\langle K|\hat{H}^{(2)}_{\text{eff}}|I\rangle = \langle I|\hat{Z}^{\cdots}_{\cdots}\ \hat{V}^N_I \hat{X}_I \hat{V}^N_I|I\rangle\ . \tag{2.52}$$

In the above operator $\hat{Z}^{\cdots}_{\cdots}$ is a string of second-quantized operators, whose action to the left on $\langle I|$ results determinant $\langle K|$:

$$\langle K| = \langle I|\hat{Z}^{\cdots}_{\cdots}\ . \tag{2.53}$$

Dots in upper (lower) scripts following $\hat{Z}$ are to be replaced by orbital indices from which electrons are removed (where electrons are put) when $\langle I|$ is transformed to $\langle K|$. It is easy to see that $\hat{Z}$ is a normal ordered operator according to determinant $|I\rangle$. When diagonal matrix elements are calculated, operator $\hat{Z}$ is the identity and one can draw closed diagrams. Open diagrams emerge for non-diagonal elements.

Generalized Wick-theorem states that only fully contracted expressions of a normal ordered operator product contribute to the expectation value taken with the Fermi-vacuum [84]. Extensivity of a many-body theory is ensured if there is no disconnected term among various full contractions in the energy expression. In the MR MBPT case the matrix elements of $\hat{H}_{\text{eff}}$ are connected[30]. As shown by Meissner and Jankowski[34], the connectedness of matrix elements leads to a connected energy, when using a complete model space. Without special considerations, incomplete model space leads to extensivity violation.

Starting at second-order, Fig. 2.1.a and b show the closed Goldstone-type skeleton diagrams contributing to the diagonal elements of the effective Hamiltonian. Intervertex straight lines represent operator $\hat{X}_I$ in these diagrams. One must bear in mind that $\hat{X}_I$ operator projects to the Q space, therefore at least one of the lines must correspond to a non-active index, i.e core or inactive virtual. Particle and hole lines crossing the $\hat{X}_I$ line

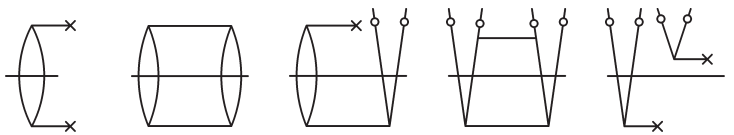


Figure 2.1: Second-order diagrams of the effective Hamiltonian.

give rise the denominator according to the following rule:

$$\Delta = \sum \epsilon^{\oplus}_{up} - \sum \epsilon^{\ominus}_{down}, \tag{2.54}$$

where $\epsilon^{\ominus}_{up}$ refers to the up going particle (virtual) lines and $\epsilon^{\oplus}_{down}$ to the down going hole (occupied) lines, with respect to the given Fermi-vacuum. In the case of symmetry- or spin-adapted formulation a second restriction has to be applied, since denominators depend only on the net occupation number changes according to Eq.(2.34). In this manner a given one-particle subspace s contributes to the denominator in the following way:

$$\Delta_s = (n_s^{up} - n_s^{down})\epsilon^{S}, \text{ where } \quad S = \text{sign}(n_s^{up} - n_s^{down}). \tag{2.55}$$

where n_s^{up} and n_s^{down} denote the number of particle and hole lines which cross the denominator line.

Fig. 2.1.c and d show diagrams of the off-diagonal matrix elements of the effective Hamiltonian. Open lines at the top are necessarily labeled by active indices, since the effective Hamiltonian lies entirely in space P. To emphasize this restriction, an empty circle are drawn to these lines.

As an example, a disconnected second-order diagram is also shown in Fig. 2.1.e (see Section 2.3.3). One readily sees that all diagrams of this short give zero value, because exclusively active lines cross the intervertex level where a Q space projection takes place. Consequently the second-order MP MBPT is connected irrespective of the actual $H_I^{(0)}$ specification (i.e. *genuine*, *single*, *spin-adapted* or *general symmetry-adapted*).

Some examples for diagrams corresponding to the first (regular) and the second (renormalization) term of Eq.(2.48) are shown in
Fig. 2.2.1(a-d) and Fig. 2.2.2(a-d). In the latter diagrams, boxes appear which indicate the special way of determining the denominator: in the *general* or the *simple* formalism

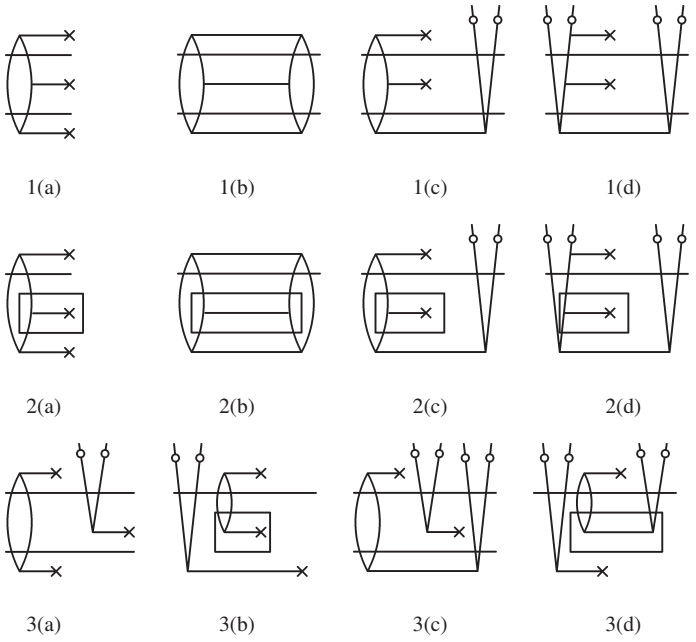


Figure 2.2: Selected examples for nonzero skeleton diagrams of the third-order effective Hamiltonian (2.48).

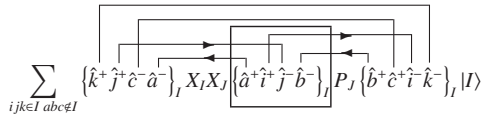


Figure 2.3: Illustration of a renormalization denominator. The perturbation denominator appearing in operator $\hat{X}_J$ contains only those one-particle indices which make the difference between determinant $|J\rangle$ and determinant $\left\{\hat{a}^+\hat{i}^+\hat{j}^-\hat{b}^-\right\}_I|J\rangle$. The up going/down going lines of the corresponding diagram are the lines directed to the left/right. The box over operator $\left\{\hat{a}^+\hat{i}^+\hat{j}^-\hat{b}^-\right\}_I$ has the same role as the renormalization box in the diagram. Creation operators within the box generate outgoing and annihilation operators generate incoming contraction lines. The denominator at $\hat{X}_J$ is $\epsilon_j^{\oplus} + \epsilon_b^{\oplus} - \epsilon_i^{\ominus} - \epsilon_a^{\ominus}$.

lines entering (leaving) the box contribute an orbital energy with positive (negative) sign:

$$\Delta = \sum \epsilon_{in}^{\oplus} - \sum \epsilon_{out}^{\ominus}. \tag{2.56}$$

The source of this diagrammatic rule is the multi-reference nature of the theory. Operator $\hat{X}_J$ in the second term of Eq.(2.48) involves a zero-order Hamiltonian determined with respect to determinant $|J\rangle$ whose nature gets revealed at the first intervertex level, algebraically at the appearance of $\hat{P}_J$ in the expression (see Fig. 2.3).

When drawing renormalization diagrams in the *genuine* or *simple* formulation, one has to take care to include only those contraction lines in a box, which are attached to vertices appearing between the top and bottom line of the box (see Fig. 2.3). It is interesting to note that this question is irrelevant in either the absence of multipartitioning or in *spin* or *general symmetry-adapted* case. In MR MBPT no distinction is made between EA type and IP type orbital energies, therefore the same line entering and leaving the box contributes zero to the denominator. In the *spin* or *general symmetry-adapted* MP MBPT each spin orbital is assigned to a subspace and denominators are constructed based on occupation number changes of the subspaces, Δ_s where:

$$\Delta_s = (n_s^{up} - n_s^{down})\epsilon^{\circledS}, \text{ where } \quad S = \text{sign}(n_s^{up} - n_s^{down}). \tag{2.57}$$

with n_s^{out} and n_s^{in} denoting the number of particle and hole lines which leave and enter the denominator box.

In the MR MBPT formalism each renormalization denominator brings a factor of -1 as a consequence of the second term of Eq.(2.5). This can be easily checked on the example of Eqs.(2.48) and (2.51).

Fig. 2.4 shows prototypes of the fourth-order diagrams where only closed diagrams are printed. A detailed description of high-order MR MBPT diagrams can be found in [32]. Diagrams in Fig. 2.4.b-e illustrate the four different kinds of renormalization terms that appear at order four in accordance with Eq.(2.51).

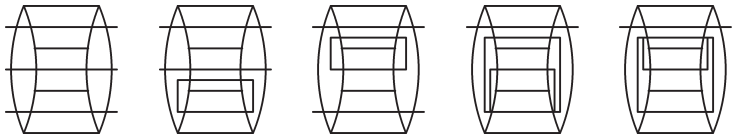


Figure 2.4: Prototypes of fourth-order diagrams.

Finally, it is practical to collect diagram rules which are additional to the diagram rules discussed in Section 1.3.1.

- In the *spin-* and *symmetry-adapted* MP MBPT formalism horizontal denominator lines indicate the application of Eq.(2.55) instead of Eq.(2.54).

- Renormalization diagrams contain one or more denominator boxes which contribute to the denominator according to Eq.(2.56) in the *genuine* or *simple* MP and Eq.(2.57) in the *spin-adapted* or *general symmetry-adapted* cases.

- Each denominator box multiplies the value of the diagram by -1.

- At the horizontal lines of denominators or at the top of denominator boxes at least one of the contraction lines must be inactive.

- At the bottom of denominator boxes all contraction lines have to be active.

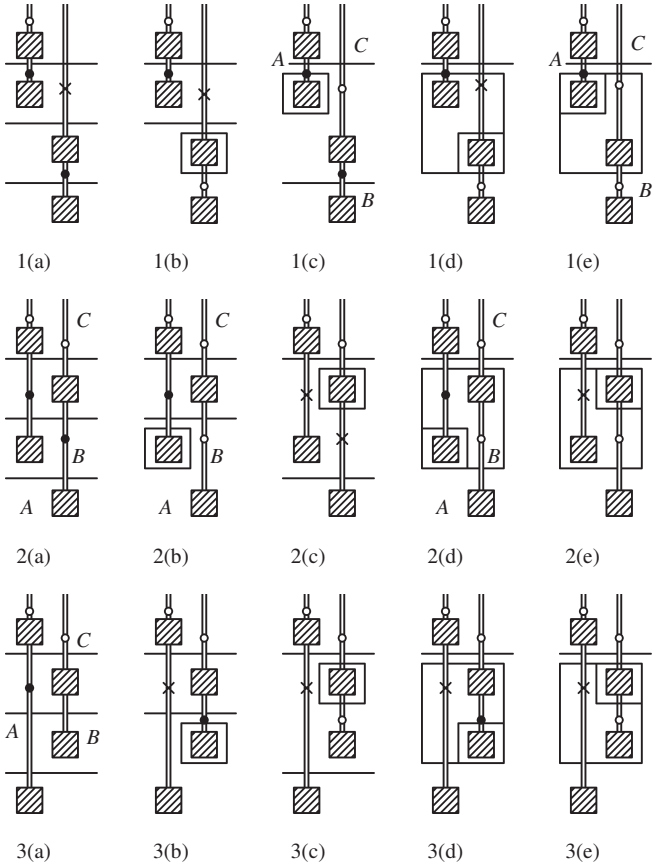


Figure 2.5: Structure of fourth-order diagrams having at least two non-interacting parts where both parts contain two vertices.

2.3.3 Scaling properties

In order to get a size-extensive theory, disconnected diagrams should be zero or should cancel each other. This is true at any order of MR MBPT, as shown by Lindgren[30]. As it is already published by Zaitevskii and Malrieu[78], the *genuine* MP formulation is size-consistent at the second-order. In this section we discuss the extensivity and size-consistency for different formulations of MP MBPT at higher-orders. As we will see, the different MP MBPT formulations show rather different behavior from this point of view.

Fig. 2.1.e shows an example for the second-order disconnected diagrams. One readily sees that all diagrams of this short give zero value, because exclusively active lines cross the intervertex level where a Q space projection takes place. Consequently the second-order MP MBPT is connected irrespective of the actual $H_I^{(0)}$ specification (i.e. *genuine*, *single*, *spin-adapted* or *general symmetry-adapted*).

At second-order the lack of disconnected terms leads to size-consistency of the *genuine*, *simple* and *spin-adapted* formalism if localized orbitals are used. This result was previously obtained in an algebraic way by Zaitevskii and Malrieu[78]. At difference with other partitionings, separability of the second-order of the *general symmetry-adapted* theory may be problematic. Although extensivity of the *symmetry-adapted* formalism is ensured at second-order, it can show separability error if the symmetry of the subsystems is different than the composite system.

Examples for third-order disconnected diagrams are shown in Fig. 2.2.3-4. In order to get a size-extensive theory , disconnected diagrams should be zero or should be cancel each other. In the case of zero-order Hamiltonian of *genuine* MP (2.15) this does not take place. To show this we write the algebraic expression corresponding to diagrams Fig. 2.2.3(a) and Fig. 2.2.3(b). Using notations of Eq.(1.90) we get

$$\sum_{i\in I}\sum_{a\notin I}\frac{f_{ia}f_{\mu\nu}f_{ai}}{\left(\epsilon_i^{\oplus}(I)-\epsilon_a^{\ominus}(I)\right)\left(\epsilon_i^{\oplus}(I)+\epsilon_\mu^{\oplus}(I)-\epsilon_a^{\ominus}(I)-\epsilon_\nu^{\ominus}(I)\right)} \tag{2.58}$$

$$-\sum_{i\in I}\sum_{a\notin I}\frac{f_{\mu\nu}f_{ia}f_{ai}}{\left(\epsilon_i^{\oplus}(I_\mu^\nu)-\epsilon_a^{\ominus}(I_\mu^\nu)\right)\left(\epsilon_i^{\oplus}(I)+\epsilon_\mu^{\oplus}(I)-\epsilon_a^{\ominus}(I)-\epsilon_\nu^{\ominus}(I)\right)} \tag{2.59}$$

where in the denominator of the renormalization term determinant $\nu^+\mu^-|I\rangle$ appears, symbolized by I_μ^ν. If the value of $\epsilon_i^{\oplus}(I)$ depends on determinant $|I\rangle$ then the two terms do not cancel each other.

Simple MP, *spin-* and *general symmetry-adapted* formulations form a different case, since the determinant in round brackets governs only whether the upper index is $\oplus$ or $\ominus$. In fact, reference to I or I^{γ}_{μ} can be omitted from the above formulae and thereby cancellation is ensured. By partitioning (2.20) the final expression corresponding to diagrams 3(a) and 3(b) in Fig. 2.2 takes the form

$$\sum_{i\in I}\sum_{a\notin I}\frac{f_{ia}f_{\mu\nu}f_{ai}}{\left(\epsilon_i^{\oplus}-\epsilon_a^{\ominus}\right)\left(\epsilon_i^{\oplus}+\epsilon_\mu^{\oplus}-\epsilon_a^{\ominus}-\epsilon_\nu^{\ominus}\right)} \tag{2.60}$$

$$-\sum_{i\in I}\sum_{a\notin I}\frac{f_{\mu\nu}f_{ia}f_{ai}}{\left(\epsilon_i^{\oplus}-\epsilon_a^{\ominus}\right)\left(\epsilon_i^{\oplus}+\epsilon_\mu^{\oplus}-\epsilon_a^{\ominus}-\epsilon_\nu^{\ominus}\right)}\,. \tag{2.61}$$

It can be checked analogously, that at third-order all disconnected diagram pairs cancel each other for *simple*, *spin-*, and *general symmetry-adapted* partitioning.

We have already shown that *genuine* MP MBPT is not extensive even at the third-order. Below we shall show that the *simple* and *spin-adapted* MP MBPT formalism keeps the extensivity in higher-orders while the *general symmetry-adapted* formulation does not.

To see this let us consider first the *simple* MP MBPT at order four. Fig. 2.5 shows the fourth-order diagrams separated into at least two (or more) disconnected parts, for the case where the independent parts contain two vertices. Shaded boxes symbolize general interactions, double lines denote contraction lines which start or end in a vertex or simply cross without interaction. To emphasize the restrictions due to the model or orthogonal space projections in Fig. 2.5, full circles are drawn to the interactions with at least one non-active index and empty circles denote the fully active set of indices. There is finally a notation (cross) for the case when no choice can produce a nonzero diagram since both a model and an orthogonal projection act on a given determinant. Studying Fig. 2.5 one can conclude that only six diagrams 1(c), 1(e), 2(a), 2(b), 2(d) and 3(a) may have non-zero value. Let us introduce three hyper-indices A, B and C where for example $A = (i, j, \ldots \in I\ a, b, \ldots \notin I)$ and $\Delta_A = \epsilon_i^{\oplus} + \epsilon_j^{\oplus} \ldots - \epsilon_a^{\ominus} - \epsilon_b^{\ominus} \ldots$. To show the cancellation of diagrams 1(c), 2(a), 2(b) and 3(a), we can utilize the Frantz-Mills factorization theorem

[62]. The sum of the diagrams 2(a), 3(a) and 2(b) can be simplified as:

$$\sum_{A,B\in\bullet\ C\in\circ} \frac{K(A)L(B,C)}{(\Delta_A+\Delta_C)(\Delta_A+\Delta_B)\Delta_B} + \sum_{A\in\bullet\ C\in\circ\ B} \frac{K(A)L(B,C)}{(\Delta_A+\Delta_C)(\Delta_A+\Delta_B)\Delta_A}$$
$$- \sum_{A\in\bullet\ B,C\in\circ} \frac{K(A)L(B,C)}{(\Delta_A+\Delta_C)(\Delta_A+\Delta_B)\Delta_A} = \sum_{A,B\in\bullet\ C\in\circ} \frac{K(A)L(B,C)}{(\Delta_A+\Delta_C)\Delta_A\Delta_B}\,, \tag{2.62}$$

where $K(A)$ is an integral indexed by A and $L(B,C)$ is the contracted product of one- and two-electron integrals. It is important to note that there is no restriction on index B in diagram 3(a) (second term of the above equation). As a result of Eq.(2.62), we get the negative of diagram 1(c), thus these four diagrams cancel each other. The remaining two nonzero diagrams 1(e) and 2(d) also cancel each other. For the case when one nonzero disconnected part contains only one vertex similar considerations can be taken. The corresponding diagrams are given in Fig. 2.6. It is easy to see that all diagrams cancel pairwise: Fig. 2.6.1(a) and Fig. 2.6.2(b), Fig. 2.6.3(a) and Fig. 2.6.2(d), Fig. 2.6.3(b) and Fig. 2.6.2(e) and finally Fig. 2.6.1(e) and Fig. 2.6.3(d) cancel each other. Rules of cancellation of disconnected diagrams having three (or four) non-interacting parts follows trivially from the above statements.

In the above derivation we used the Frantz-Mills factorization theorem where we supposed the additivity of denominators. This property can be formulated by the relation

$$D^{a_1,\ldots,a_m}_{i_1,\ldots,i_m} + D^{a_{m+1},\ldots,i_n}_{i_{m+1},\ldots,i_n} = D^{a_1,a_2,\ldots,a_n}_{i_1,i_2,\ldots,i_n} \tag{2.63}$$

where $D^{a_1,a_2,\ldots,a_n}_{i_1,i_2,\ldots,i_n}$ is a denominator composed of $a_1, a_2, \ldots a_n$ virtual and $i_1, i_2, \ldots i_n$ occupied indices and $1 < m < n$. This is essential for the application of Frantz-Mills factorization theorem and also for the extensivity. In the *simple* MP MBPT formalism Eq.(2.63) valid. Taking into account that MP MBPT and MR MBPT use the same set of diagrams, the proof of the linked cluster theorem published by Lindgren[30] for MR MBPT remains true for the *simple* MP MBPT formalism.

In the *spin-* and *general symmetry-adapted* MP MBPT formulations the cancellation of diagram pairs Fig. 2.5.1(e) and 2.5.2(d), Fig. 2.6.1(a) and Fig. 2.6.2(b), Fig. 2.6.3(a) and Fig. 2.6.2(d), Fig. 2.6.3(b) and Fig. 2.6.2(e), Fig. 2.6.1(e) and Fig. 2.6.3(d) occurs similarly as in the previous case. At the application of the Frantz-Mills factorization theorem, one should be more careful since Eq.(2.63) is not valid by these operators. To show this, suppose that a_1 and i_n active orbitals belong to the same degenerate subspace, s.

Using Eqs.(2.55) and (2.57) the valid form of Eq.(2.63) for the *general symmetry-adapted* formulation is

$$D^{a_1,\ldots,a_m}_{i_1,\ldots,i_m} \; (+) \; D^{a_{m+1},\ldots,i_n}_{i_{m+1},\ldots,i_n} = D^{a_2,\ldots,a_n}_{i_1,i_2,\ldots,i_{n-1}} \tag{2.64}$$

where (+) denotes the special operation which in addition to the summation also contains the elimination of occupied-virtual index pairs belonging to the same degenerate subspace. E.g. if a_1 and i_n active orbitals belong to the same subspace, they do not contribute to $D^{a_1,a_2,\ldots,a_n}_{i_1,i_2,\ldots,i_n}$ but $\epsilon^{\ominus}_{a_1}$ and $\epsilon^{\oplus}_{i_n}$ do appear in $D^{a_1,\ldots,a_m}_{i_1,\ldots,i_m}$ and $D^{a_{m+1},\ldots,i_n}_{i_{m+1},\ldots,i_n}$, respectively. As a result, Eq.(2.62) has to be modified and the sum of diagrams 2(a), 3(a) and 2(b) takes the following form:

$$\begin{aligned} &\sum_{A,B\in\bullet\; C\in\circ} \frac{K(A)L(B,C)}{(\Delta_A(+)\Delta_C)(\Delta_A(+)\Delta_B)\Delta_B} + \sum_{A\in\bullet\; C\in\circ\; B} \frac{K(A)L(B,C)}{(\Delta_A(+)\Delta_C)(\Delta_A(+)\Delta_B)\Delta_A} \\ &- \sum_{A\in\bullet\; B,C\in\circ} \frac{K(A)L(B,C)}{(\Delta_A(+)\Delta_C)(\Delta_A(+)\Delta_B)\Delta_A} = \sum_{A,B\in\bullet\; C\in\circ} \frac{K(A)L(B,C)(\Delta_A+\Delta_B)}{(\Delta_A(+)\Delta_C)(\Delta_A(+)\Delta_B)\Delta_A\Delta_B} \,, \end{aligned} \tag{2.65}$$

while the value of diagram 1(c) is

$$- \sum_{A,B\in\bullet\; C\in\circ} \frac{K(A)L(B,C)}{(\Delta_A(+)\Delta_C)\Delta_A\Delta_B} \,. \tag{2.66}$$

We can see that the cancellation is not complete, since terms on the right hand side of Eq.(2.65) are not equal with expression (2.66). The remaining disconnected term,

$$\sum_{A,B\in\bullet\; C\in\circ} \frac{K(A)L(B,C)}{(\Delta_A(+)\Delta_C)\Delta_A\Delta_B} \frac{\Delta_A+\Delta_B-(\Delta_A(+)\Delta_B)}{\Delta_A(+)\Delta_B} \,, \tag{2.67}$$

give non-zero contribution if one or more degenerate subspaces in the Fermi-vacuum are partly occupied. For example, in the *spin-adapted* case for a closed shell Fermi-vacuum all disconnected terms are cancelled. On the other hand if there are open shell active orbitals in the Fermi-vacuum, called p, with a single electron α and $p_\alpha \in A$ and $p_\beta \in B$, the remaining disconnected term is

$$\sum_{p\; A',B'\in\bullet\; C\in\circ} \frac{K(p_\alpha,A')L(p_\beta,B',C)}{(\Delta_A(+)\Delta_C)\Delta_A\Delta_B} \frac{\epsilon^{\oplus}_p-\epsilon^{\ominus}_p}{\Delta_{A'}+\Delta_{B'}} \,, \tag{2.68}$$

where A' is the hyper index A without p_α and B' is hyper index B without p_β. In general we can say that the two connected units of a disconnected diagram are linked by orbitals belonging to the same partly occupied degenerate subspace.

According to the previous analysis connectivity in the strict sense does not hold at fourth and higher-order if either the *spin-adapted* or the *symmetry adapted* formulation is used. These theories, however, still can be size-extensive or size-consistent if the zero-order Hamiltonian does not connect localized orbitals belonging to different subsystems at large or infinite separation.

Size-consistency of *spin-adapted* formalism

If *spin-adapted* theory is used with localized orbitals and hyper indices A and B belong to different subsystems at infinite separation, relation $\Delta_A + \Delta_B = \Delta_A(+)\Delta_B$ is obviously satisfied. Terms in Eq.(2.65) are then cancelled by the proper terms of Eq.(2.66). Although the disconnected diagrams of type 1(c), 2(a), 2(b) and 3(a) do not cancel entirely each other, products with indices of both subsystems do not give contribution. This means that the *spin-adapted* theory is size-consistent.

Size-consistency of general *symmetry-adapted* formalism

Similarly it can be shown that the *symmetry-adapted* formalism can lead to size-consistent result with localized orbitals if there is no symmetry transformation which would connect partly occupied degenerate active orbitals localized on different systems at infinite separation. Otherwise the *symmetry-adapted* approach is not size-consistent.

Size-extensivity of *spin-adapted formalism*

If localized orbitals are used in the *spin-adapted* theory, extensivity holds due to the fact that the two connected units of a disconnected diagram are always linked by at least an active orbital, singly occupied in the Fermi-vacuum. According to the discussion in Section 1.1.1, this property leads to the conclusion that the *spin-adapted* formalism is size-extensive when using localized orbitals. To investigate the case of delocalized orbitals, let us focus on the two connected units of a disconnected diagram. We know, that if there is a sum over each index of a connected diagram, the result will be proportional the system size. We are going to consider larger systems now, where orbital energies become closely spaced, that summation over orbital indices can be replaced by integration:

$$\sum_q \rightarrow \int \frac{dN}{d\epsilon} d\epsilon = \int \rho d\epsilon. \tag{2.69}$$

The density of one-particle states around ϵ_q is proportional with the size of the system N, $\rho_n(\epsilon) \sim N$. Let us ignore for a moment the sum over orbital q which connects the subdiagrams, suppose that its energy is ϵ_q and denote the contribution of a connected subdiagram from orbitals around ϵ_q by $dD_\alpha(\epsilon)$. This means, that sum over all indices are

evaluated when obtaining $dD_\alpha(\epsilon)$, apart from ϵ_{q_a}. As for q, the sum extends over those indices which lie in $d\epsilon$ interval to ϵ_q. Supposing that the number of orbitals around ϵ_q is proportional with the density of one-particle states we have $dD_\alpha(\epsilon) = R_\alpha(\epsilon)\rho_n(\epsilon)d\epsilon$. Here $R_\alpha(\epsilon)$ is the contribution of the single orbital q to the value of the subdiagram denoted by α. Obviously $dD_\alpha(\epsilon)$ is also proportional with N and $R_\alpha(\epsilon)$ does not scale with the system size. For a given q the contribution of the disconnected diagram for the energy is proportional with $R_\alpha(\epsilon)R_\beta(\epsilon)$, i.e. $dD_{\alpha\beta}(\epsilon) = R_\alpha(\epsilon)R_\beta(\epsilon)\rho_n(\epsilon)d\epsilon$ which is proportional with N. This means, that after the integration for the one-particle energy the whole contribution becomes proportional with the system-size. As a consequence, disconnected diagrams appearing in the *spin-adapted* formalism give size-extensive contribution even if delocalized orbitals are used.

Size-extensivity of general *symmetry-adapted* formalism

In the *symmetry-adapted* case, orbitals partly occupied in a degenerate subspace appear in both connected units of a disconnected diagram. Using an analysis similar to the above discussion, the contribution of a set of degenerate active orbitals $(s_1, \ldots, s_n)$ for a disconnected diagram with two subdiagrams is as follows:

$$dD(\epsilon) = \sum_{i\in I, a\notin I} R_{s_a}(\epsilon)R_{s_i}(\epsilon)\rho(\epsilon)d\epsilon, \tag{2.70}$$

where $R_{s_a}(\epsilon)$ is the contribution of a virtual orbital for one of the subdiagrams and $R_{s_i}(\epsilon)$ is the contribution of an occupied orbital for the other subdiagram. If n does not scale with the system dimension N, then $dD(\epsilon)$ is a linear function of N and size-extensivity is ensured. If the dimension of a degenerate subspace generated by symmetry is proportional with the system size, $dD(\epsilon)$ will then scale quadratically or cubically with system size. Systems with translational symmetry, finite systems with periodic boundary conditions or ring shaped systems are examples for the case where the size-extensivity of the *symmetry-adapted* formalism does not hold.

To summarize our results we can say that the genuine MP MBPT formalism is size-consistent and size-extensive at second-order but it is not true for higher-orders. The *simple* formalism is size-consistent and size-extensive at any order, but harms both the

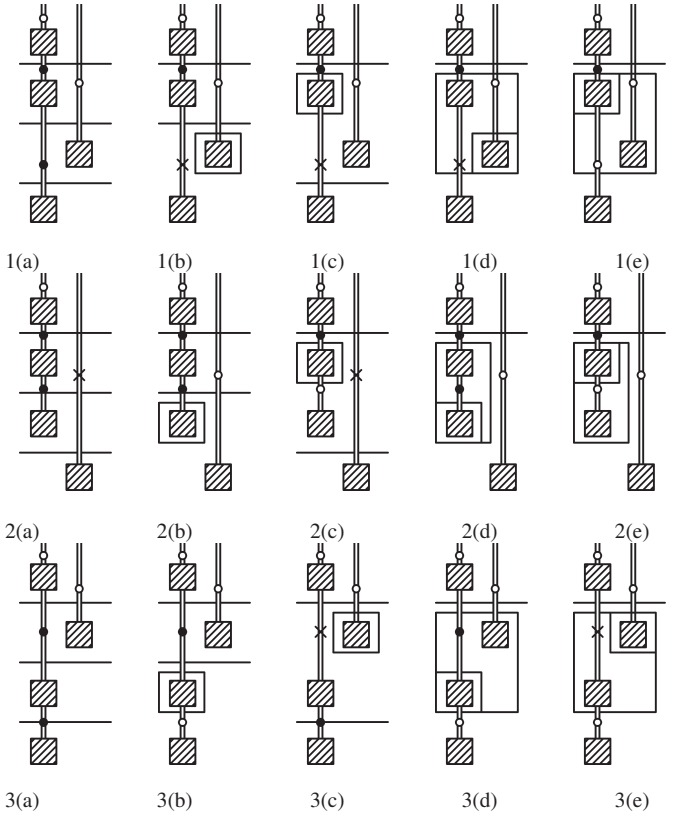


Figure 2.6: Disconnected fourth-order terms where there is one alone standing vertex.

spin and the spatial symmetry. In the *spin-adapted* and *general symmetry-adapted* formulation the definition of the zero-order Hamiltonian is not connected in a strict sense. Definition of connectedness is frequently used as a synonym of extensivity. To keep the well established relation between connectedness and extensivity, we should consider a diagram build of two disjoint units connected, if both subdiagrams contain the same orbital with different spin. Moreover, if we use localized orbitals to describe for example a dissociation process and the two units of a disconnected diagram contain orbitals localized to the same fragment, one can also consider this product connected. In this sense the spin-adapted theory is always connected, size-extensive and size-consistent. This is not true for the *general symmetry-adapted formulation*, since in this case the symmetry can in general connect fragments at infinite distance from each other. The *general symmetry-adapted formulation* can be still size-consistent in special cases, where the symmetry can not link the fragments at infinite separation, e.g. in the description of the dissociation process of a heteronuclear diatomic molecule. One further problem in the *general symmetry-adapted formulation* is that the symmetry can change along a potential energy surface. In order to get a smooth PES, the lowest symmetry has to be used all along the process (see for example Fig. 2.11).

2.3.4 Implementation

Adaptation of the code generating tools of MCPT for the MP MBPT formalism a quick and straightforward implementation of MP MBPT. This was easily achieved by minor modification of the scripts and codes used for implementing MCPT. Although MP MBPT can be calculated for medium size systems by the code produced in this way, its performance is by far not optimal.

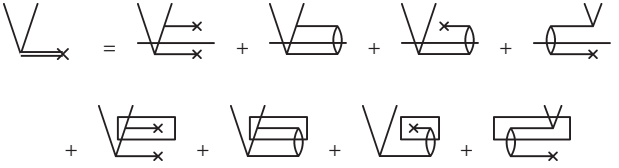
Code developments were motivated by two aspects: calculation time and code length reduction. The initial third-order code scaled as $N_{CAS}N_o^2N_v^4$, while the fourth-order code scaled as $N_{CAS}N_o^2N_v^6$ where N_{CAS}, N_o and N_v are the size of the CAS space, number of occupied orbitals and number of virtuals, respectively. This scaling can be significantly reduced by using properly chosen intermediate quantities. The second aspect is the length of the source code. This was several hundred-thousand lines initially, but it also gets significantly shorter by introducing the intermediers.

At second-order there is no need to apply intermediers. Simple subroutines are called to calculate the diagrams shown for example in Fig. 2.1. The second-order effective Hamiltonian correction, $\hat{H}^{(2)}_{\text{eff}}$ is stored, since it will be needed for higher-order corrections according to Eq.(2.48).

In what follows we focus our attention on the construction of intermediers. This affects only those diagrams, where there is an intervertex level that describes a single or a double excitation but not more. There are numerous diagrams that are not of this sort. These are implemented similar to the procedure discussed in Section 1.3.2. In the discussion below we will elaborate only on those diagrams, where it is possible to introduce intermediers with two or four open lines, called one- or two-particle intermediers, respectively. Three or more particle intermediers are easy to introduce following the lines presented below.

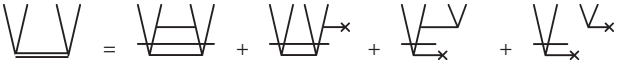
At any order the most expensive diagrams are the closed diagrams, which belong to the diagonal matrix elements well known form single reference theory [60]. To reduce the scaling of the calculation, certain diagrams of the effective Hamiltonian matrix elements are constructed from their wave operator component by closing the wave operator diagram with a Hamiltonian diagram shown in Fig. 1.3. Wave-operator diagrams with no more than two occupied and two virtual indices represent the simplest intermediers. They can be easily stored even for large basis sets since this vector is significantly shorter than the whole integral list.

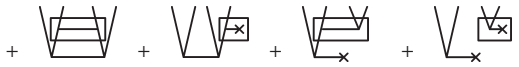
The wave operator diagrams used as intermediers at order three are calculated within the outermost loop which runs over the determinants in the model space. The structure of these second-order one-particle intermediers is as follows:



(2.71)

The form of the second-order two-particle intermediers is the following:



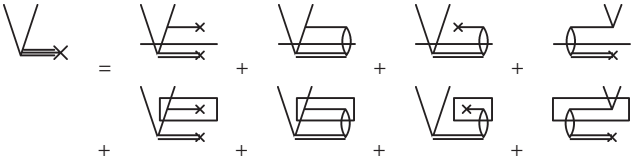


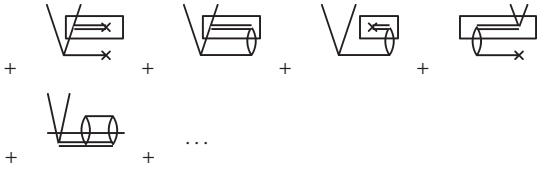
(2.72)

By the use of the above intermediers the third-order contribution to the effective Hamiltonian matrix can be calculated by the routines written to calculate the second order effective Hamiltonian diagrams shown in Fig. 2.1. One has to simply load the above intermediers instead of the integral list appearing at the bottom of diagrams in Fig. 2.1. It is worth to note here that only those terms contribute to the third-order which end in a Q space determinant at the top, in Eqs.(2.71) and (2.72). It is still practical to evaluate all these intermediates without any restrictions, since they will be useful at the fourth-order calculation.

The calculation time of the most expensive intermedier (first term on the right hand side of Eq.(2.72)) scales like $N_o^2 N_v^4$, which would be the scaling without using intermediers. Introduction of intermediers therefore does not give a reduction in the sixth power scaling of the third-order. It is by the factorization, where one gains computation time. Apart from the diagrams discussed so far, there appear at third-order others, which contain more than four contraction lines at the second intervertex level. Some examples for the bottom part of these diagrams are shown in Eqs.(2.75) and (2.76). They contain less contractions but have more open lines. For this reason their scaling is not the most drastic. The worst diagram of this type (first term on the right hand side of Eq.(2.75)) scales like $N_{av} N_{ao} N_o^2 N_v^3$, where N_{av} and N_{ao} is the number of active virtual and active occupied orbitals, respectively. This is lower than $N_o^2 N_v^4$ as far as the number of virtuals is significantly larger than the number of active orbitals. As mentioned before, diagrams with more than four contraction lines at the second intervertex level are computed by the native routines, generated automatically.

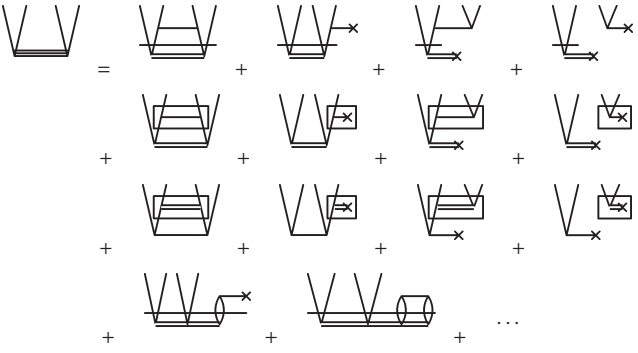
The significance of intermediers at fourth-order is even larger, due to the reduction of the scaling of calculation time. The third-order intermediers used in the fourth-order calculations are the following:





(2.73)

The form of third-order two-particle intermediers is as follows:



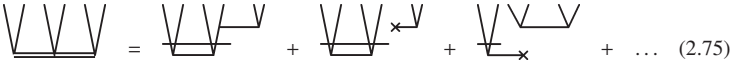
(2.74)

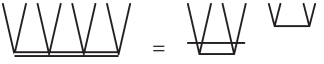
where the renormalization counterparts of the diagrams describing more than two-fold excitation at the second intervertex level are not shown but also included in the sum. Some of the fourth-order diagrams contain one- or two-fold excitation at the second intervertex level but at the third intervertex level a three- or four-fold excitation appears. The contribution of these diagrams are calculated in subroutines used in the third-order. Only the integral list is changed into the intermediers from Eqs.(2.71) and (2.72) at the bottom level of diagrams.

Substituting the first rows of diagrams from Eqs.(2.71) and (2.72) into terms in the first rows of Eqs.(2.73) and (2.74) one obtains diagrams without renormalization denominators like in Fig. 2.4.a. Diagrams with a renormalization denominator at the second intervertex level (e.g. Fig. 2.4.b.) arise from the substitution of diagrams from the second rows of Eqs.(2.71) and (2.72) into the first rows of Eqs.(2.73) and (2.74). Diagrams having a renormalization box around the third interaction vertex like Fig. 2.4.c can be derived from the first row of Eqs.(2.71) and (2.72) plugged into the second row of Eqs.(2.73) and (2.74). Due to the contradictory restrictions at the second intervertex level, substitution

of diagrams from the first row of Eqs.(2.71) or (2.72) into the second rows of Eqs.(2.73) and (2.74) leads to zero. Fourth order diagrams with two renormalization boxes like those in Fig. 2.4.d and Fig. 2.4.e result from the third rows of Eq.(2.71) and (2.72) being substituted into the first and second rows of Eqs.(2.73) and (2.74). Here the Fermi vacuum of second-order intermedier defined by the determinant appears at the first intervertex level.

The last rows of Eqs.(2.73) and (2.74) show terms having three- and four-fold excitation at the second intervertex level. The definitions of these quantities are:


(2.75)


(2.76)

where the renormalization counterparts of the diagrams are not shown but included in the sum. Intermediers defined in (2.76) are explicitly calculated but not stored since their memory requirement rapidly grows with the basis set size. Whenever an intermedier of this sort is calculated, with a given set of uncontracted indices, its contributions to Eqs.(2.73) and (2.74) is right away calculated and the intermedier is dropped afterwards.

Using the intermediers of order three, the effective Hamiltonian matrix elements can be calculated by a final contraction with the Hamiltonian diagrams in Fig. 1.3. This last step scales like $N_o^2 N_v^2$ which is negligible compared to the cost of the construction of certain intermediers. The two most demanding intermediers are the first one in Eq.(2.74) with $N_o^2 N_v^4$ property and the last term in Eq.(2.74) with $N_o^4 N_v^4$ scaling. This latter diagram determines the scaling of the presented implementation which is finally proportional to $N_{CAS} N_o^4 N_v^4$. This scaling can be reduced to $N_{CAS} N_o^4 N_v^3$ by introducing the so-called vertical factorization [60].

2.4 Numerical examples

To demonstrate the applicability of higher-order MP MBPT corrections, calculations for molecules H_2O, N_2, C_2H_4, and BeH_2 are presented.

	CAS[4/4]	2nd order	3rd order	4th order	FCI
$1\ ^1A_1$	0.2137	0.0168	0.0081	0.0012	-76.2464
$1\ ^3A_1$	0.2678	-0.0024	0.0160	0.0002	-75.8920
$2\ ^1A_1$	0.2654	-0.0021	0.0154	0.0002	-75.8657
$5\ ^1A_1$	0.4771	-0.0415	0.0301	0.0069	-75.4608

Table 2.1: Error of state energies of the H_2O molecule at different perturbation orders of MP MBPT and the FCI total energies, in Hartree. *General symmetry-adapted* MP MBPT formulation, with state-averaged one-particle energies. Averaging involves states 1^1A_1 , 1^3A_1 and 2^1A_1. See text for geometry and basis set.

	CAS[4/4]	2nd order	3rd order	4th order
$1\ ^3A_1$	0.0541	-0.0192	0.0079	-0.0010
$2\ ^1A_1$	0.0517	-0.0189	0.0073	-0.0010
$5\ ^1A_1$	0.2634	-0.0583	0.0220	0.0057

Table 2.2: Error of excitation energies of the H_2O molecule in different perturbation orders of MP MBPT, in Hartree. *General symmetry-adapted* MP MBPT, with state-averaged one-particle energies. Averaging involves states 1^1A_1 , 1^3A_1 and 2^1A_1. See text for geometry and basis set.

2.4.1 H_2O

On the example of the H_2O molecule we compare results obtained with different zero-order Hamiltonians. One IP and one EA set is constructed either in a state-selective or in a state-averaged manner by the definition of Eq.(2.18) or Eq.(2.19), respectively. We also compare the *simple* and the *general symmetry-adapted* formulation, defined by Eq.(2.20) and Eq.(2.25), respectively.

Calculation for the H_2O molecule was performed using Pople type 6-311G* basis set, close to equilibrium geometry, with parameters $R = 0.9393$Å and $\Theta = 107.5°$. The core orbital with two electrons was kept frozen. As a zero order approximation a four-electrons-four-orbitals CAS was constructed, using Hartree–Fock canonical orbitals. The active space was defined by orbitals $(3a_1)$, $(1b_2)$, $(4a_1)$ and $(2b_1)$. The calculations being restricted to the totally symmetric states, the model space was of dimension ten. As bench-

	CAS[4/4]	2nd order	3rd order	4th order	FCI
$1\ ^1A_1$	0.2137	0.0168	0.0081	0.0012	-76.2464
$1\ ^3A_1$	0.2678	-0.0008	0.0165	0.0008	-75.8920
$2\ ^1A_1$	0.2654	-0.0018	0.0155	0.0002	-75.8657
$5\ ^1A_1$	0.4771	-0.0415	0.0301	0.0070	-75.4608

Table 2.3: Error of state energies of the H_2O molecule in different perturbation orders of MP MBPT and the FCI total energies, in Hartree. *Simple* MP MBPT formulation was used with state-averaged one-particle energies. Averaging involved states 1^1A_1 , 1^3A_1 and 2^1A_1. See text for geometry and basis set.

	CAS[4/4]	2nd order	3rd order	4th order
$1\ ^3A_1$	0.0541	-0.0176	0.0084	-0.0004
$2\ ^1A_1$	0.0517	-0.0186	0.0074	-0.0010
$5\ ^1A_1$	0.2634	-0.0583	0.0220	0.0058

Table 2.4: Error of excitation energies of the H_2O molecule in different perturbation orders of MP MBPT, in Hartree. *Simple* MP MBPT formulation, with state-averaged one-particle energies. Averaging involves states 1^1A_1 , 1^3A_1 and 2^1A_1. See text for geometry and basis set.

marks, FCI energies were computed for all states. FCI energies were calculated by a code implemented in our laboratory using the algorithm published by Olsen[86]. Assignation of states obtained by MP MBPT is based on a comparison of the leading configuration of FCI vectors and the right hand eigenvector of the approximate effective Hamiltonian. Of the ten eigenvalues of $\hat{H}_{\text{eff}}$, four states proved to be good estimations of one of the FCI states. Since orbitals $(1a_1)$ and $(2a_1)$ are fully occupied in all configurations, we restrict ourselves to indicating the occupation of only six electrons for simplicity. The leading configuration belonging to the $1\ ^1A_1$ state is $(1b_1)^2(3a_1)^2(1b_2)^2$. The first totally symmetric, open shell singlet and triplet states are dominated by the $(1b_1)^2(3a_1)^1(2b_2)^2(4a_1)^1$ configuration and finally the leading term of fifth $5\ ^1A_1$ state is $(1b_1)^2(3a_1)^2(4a_1)^2$.

Table (2.1) shows the energy error of *general symmetry-adapted* state averaged results and the FCI total energies. The averaging involved the three lowest CAS eigenstates. Second-order MP MBPT reduces the error of CAS energies below ten percent as demonstrated by Table (2.1). For the $1\ ^3A_1$ and $2\ ^1A_1$ states the second-order gives a surprisingly good estimation. For these states the third-order is less successful but the fourth-order significantly reduces the error of the second order. Approximations for the $1\ ^1A_1$ and $5\ ^1A_1$ states smoothly converge with increasing order of perturbation. The approximation of the three states, involved in averaging when computing the one-particle energies, is more successful than the approximation of the 5^1A_1 state. Excitation energies obtained at different orders are collected in Table (2.2). In this point of view the third-order results give remarkable improvement compared to the second-order, for each excited states examined. Fourth order brings further notable improvement for each state, similarly to the case of the total energy. Tables (2.3) and (2.4) show the errors of the *simple* formulation. Based on this data one may conclude, that there are no significant differences between the *simple* and *general symmetry-adapted* results.

Significant difference reveals however when examining the singlet-triplet splitting of the states 1^3A_1 and 2^1A_1 as shown in Table (2.5). The *simple* formalism at the second-order reduces the error more than fifty per cent but the third-order gives no further improvement. At the fourth order the remaining error is 6 mE_h. As compared to this, second-order *general symmetry-adapted* estimation of the singlet-triplet splitting is very precise, it has approximately 2 mE_h error. The third-order enlarges this 6 mE_h but the fourth-order gives only 0.7 mE_h error.

	Simple	Spin-adapted
CAS[4/4]	0.023862	0.023862
2nd order	0.025280	0.026528
3rd order	0.025283	0.025729
4th order	0.025682	0.026220
FCI	0.026289	0.026289

Table 2.5: Singlet-triplet separation of open shell states of the H_2O molecule with the main configuration $(1a_1)^2(2a_1)^2(1b_1)^2(3a_1)^1(2b_2)^2(4a_1)^1$, by *simple* and *spin-adapted* MP MBPT calculations. See text for geometry and basis set.

	CAS[4/4]	2nd order	3rd order	4th order	FCI
$1\ ^1A_1$	0.2137	0.0118	0.0075	0.0004	-76.2464
$1\ ^3A_1$	0.2678	0.0038	0.0277	0.0059	-75.8920
$2\ ^1A_1$	0.2654	0.0019	0.0241	0.0030	-75.8657
$5\ ^1A_1$	0.4771	-0.0082	0.0696	0.0301	-75.4608

Table 2.6: Error of state energies for H_2O by MP MBPT and FCI total energies in Hartree. Spin-adapted, state-specific calculations. The target state is the $1\ ^1A_1$ state. See text for geometry and basis set.

To analyze the effect of the different one-particle energies on the results, state-specific calculations were also performed. Tables (2.6) and (2.7) show the results for the case when the ionization potentials and electron affinities were calculated with the CAS state 1^1A_1, using Eq.(2.18). As expected, the state-specific results for the ground state give better estimations than the state-averaged results, at the same time the estimation of total energies of excited states and the excitation energies are worsened. Interestingly, second-order approximations of the excited states remain reasonably good even if one-particle energies are focused on the ground state. Third and fourth-order approximations on the other hand are considerably spoiled.

	CAS[4/4]	2nd order	3rd order	4th order
$1\ ^3A_1$	0.0541	-0.0080	0.0202	0.0055
$2\ ^1A_1$	0.0517	-0.0099	0.0166	0.0026
$5\ ^1A_1$	0.2634	-0.0200	0.0621	0.0297

Table 2.7: Error of excitation energies for H_2O by MP MBPT in Hartree. Spin-adapted, state-specific calculations. The target state is the $1\ ^1A_1$ state. See text for geometry and basis set.

2.4.2 N_2

As a second example a series of calculations for the potential energy surface of the nitrogen molecule is presented. Larsen *et al.* [85] presented FCI results for the ground state and several excited states at different geometries in spherical cc-pVDZ basis set of Dunning *et al.* keeping the 1s core orbital of nitrogen atoms frozen. At internuclear distance larger than 1.8 Å for the $^1\Sigma_g^+$ ground state full-CI energies were calculated by a code implemented in our laboratory using the algorithm published by Olsen[86]. Because of the π_x and π_y orbital degeneracy, this is a perfect test system to compare the *general symmetry-adapted* zero-order Hamiltonian and the pure *spin-adapted* zero-order Hamiltonian.

Fig. 2.7 shows parallelity of the error curves and Fig. 2.9.a shows the total energy along the dissociation curve of the $^1\Sigma_g^+$ ground state of the N_2 molecule. The reference state was provided by a six-electron-six-orbital CAS calculation and the pseudo-canonical orbitals were used in the perturbation calculations. In Fig. 2.7 the zero-order CAS results are shifted by 140mE_h. The perturbative corrections were calculated by the *general symmetry-adapted* MP MBPT formalism using D_{2h} symmetry. The CAS gives a qualitatively correct description of the dissociation process and serves as a good basis for a perturbative approach. The second-order MP MBPT gives a significant improvement to the zero-order results. The third and the fourth-order corrections further reduce the remaining error. At 7Å the FCI and perturbation energies are converged to their dissociation values within five digit. In the right hand side of Fig. 2.7 results from a calculation using localized orbitals are also shown. In this case the degeneracy of p_x, p_y and p_z orbitals, localized to a given nitrogen atom, was utilized in the *general symmetry-adapted* zero-order Hamiltonian. These calculation serves better energies at the infinite separation compar-

ing to the delocalized calculation where only the two dimensional π_x, π_y degeneracy was utilized. If one uses the *general symmetry-adapted* zero-order with delocalized orbitals formulation. Since the *general symmetry-adapted* MP formulation is invariant to the rotation of the degenerate orbitals one can use (see Section 2.3.3), the only reason behind the energy differences is the different symmetry adaptation of the zero-order Hamiltonian.

The *simple* and *spin-adapted* formulations were also tested using the pseudo canonical CAS orbitals. The results are shown in Fig. 2.9.a. At second-order the *spin-adapted* formulation already gives better results than the *simple* formulation while the *general symmetry-adapted* formulation brings further improvement. At third and fourth-order the *simple* and *spin-adapted* MP MBPT results are rather similar to the *general symmetry-adapted* approach. A slight difference appears only for larger internuclear distance.

Since the MP MBPT generally not invariant to the rotation of one-particle basis set, dependence of the theories on one-electron orbitals is an interesting question. To get some numerical experience, we computed the general symmetry-adapted results with HF canonical orbitals. In this case the zero-order approximation is less successful, and as a consequence, perturbative correction are less balanced. As seen in Fig. 2.9, the results are surprisingly good at short internuclear distance, giving better results than the same theory with pseudo-canonical CAS orbitals. Parallelity of the error curve is improved in each order but in general it remains worse than the pseudo-canonical counterpart, and typically less accurate for large internuclear distance.

The lowest ${}^1\Sigma_u^-$, ${}^1\Pi_g$ and ${}^1\Delta_u$ excited states were also computed by the *general symmetry-adapted* formalism with pseudo-canonical CAS orbitals and Hartree–Fock canonical orbitals. The results are shown in Fig. 2.8.b-d and Fig. 2.9.b-d. Pseudo-canonical CAS orbitals were optimized separately for each state. General trends shown by the methods are similar for the excited states as for the ground state. At each point the energy and also the shape of the potential curve is improving for order by order of MP MBPT.

Errors of vertical excitation energies by the *general symmetry-adapted* formalism as a function of the geometry is shown in Fig. 2.10. The convergence of vertical excitation energies is slower than the convergence of individual state energies.

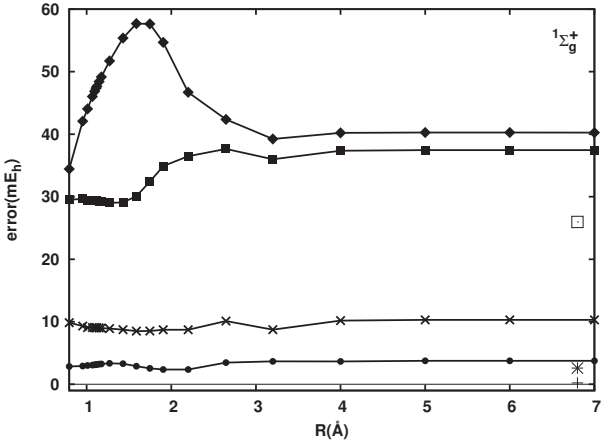


Figure 2.7: Convergence of MP MBPT for $^1\Sigma_g^+$ ground of N_2 using delocalized pseudo-canonical CAS orbitals. The errors ($E_{\text{MP MBPT}} - E_{\text{FCI}}$) of dissociation curves are plotted. CAS energies are shifted by -140mE$_h$. Notations: diamond – CAS energies, square – second-order MP MBPT energies, cross – third-order MP MBPT energies, full circle – fourth-order energies. On the right hand side results for infinitely separated nitrogen atoms are also shown, where localized orbitals were used and the degeneracy of p orbitals are utilized in the zero-order Hamiltonian. Empty square – second-order MP MBPT energy, star – third-order MP MBPT energy, plus sign – fourth-order MP MBPT energy. See text for basis set.

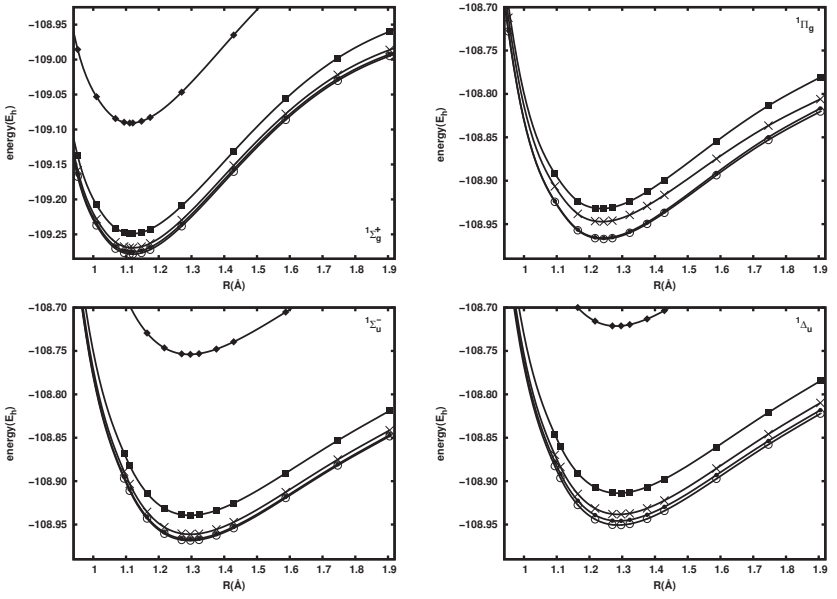


Figure 2.8: Convergence of MP MBPT for various states of N_2 as a function of internuclear distance. Dissociation curves of the N_2 molecule for $^1\Sigma_g^+$ ground and the lowest $^1\Pi_u^-$, $^1\Sigma_u^-$, $^1\Delta_u^-$ excited states. Notations: diamond – CAS energies, square – second-order MP MBPT energies, cross – third-order MP MBPT energies, full circle – fourth-order energies, thin lines – *general symmetry-adapted* formalism with canonical orbitals, thick lines – *general symmetry-adapted* formalism with pseudo-canonical CAS orbitals. See text for basis set.

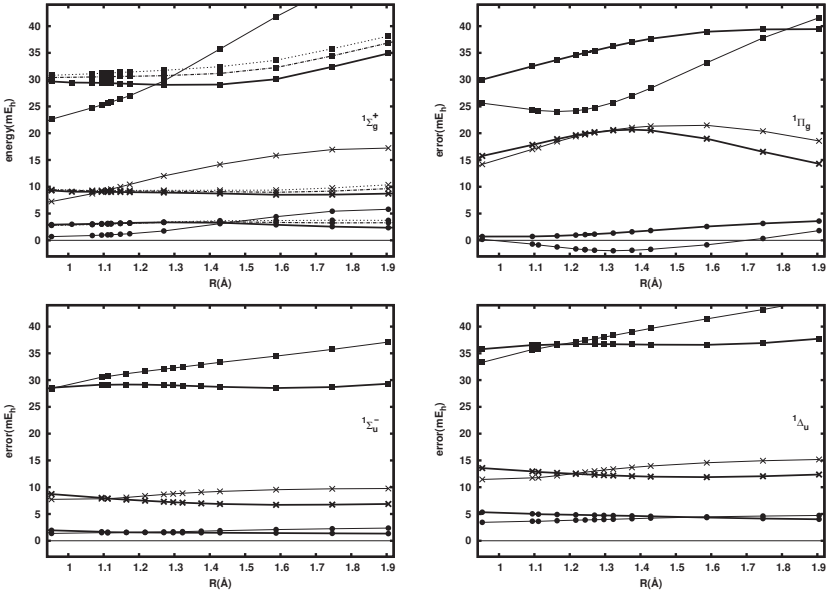


Figure 2.9: Convergence of MP MBPT for various states of N_2 as a function of internuclear distance. Error of different approximations. Notations: diamond – CAS, square – second-order MP MBPT energies, cross – third-order MP MBPT, full circle – fourth-order, empty circle – FCI, thin lines – *general symmetry-adapted* formalism with HF-canonical orbitals, thick lines – *general symmetry-adapted* formalism with pseudo-canonical CAS orbitals, dots – *simple formalism* with pseudo-canonical CAS orbitals, dashed line – *spin-adapted* formalism with pseudo-canonical CAS orbitals. See text for basis set.

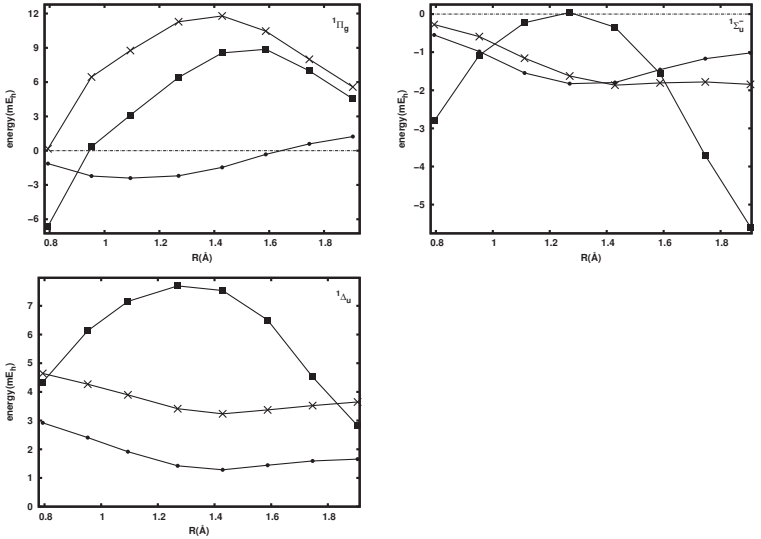


Figure 2.10: Error of excitation energies of different orders of MP MBPT for various states of N_2 as a function of internuclear distance. Notations: diamond – CAS energies, square – second-order MP MBPT energies, cross – third-order MP MBPT energies, full circle – fourth-order energies, thin lines – *symmetry-adapted* formalism with HF canonical orbitals, thick lines – *symmetry-adapted* formalism with pseudo-canonical CAS orbitals. See text for basis set.

2.4.3 C_2H_4

As another example the ethylene torsional potential curve is shown in Fig. 2.11. For the calculations Dunning's double zeta polarized basis set[73] was used. A two by two CAS calculation provided the reference function using pseudo-canonical orbitals. For this system MCPT and SC2-MCPT calculations were already shown in Section (1.4). In the lack of spatial orbital degeneracy *spin-adapted* MP MBPT was applied using D_2 symmetry along the whole potential energy curve. At 90° torsional angle the molecular symmetry point group is the higher D_{2d} and a two dimensional degeneracy of spacial orbitals appears. At this point both D_2 and D_{2d} symmetry were used to construct the zero-order Hamiltonian. In order to have a benchmark to the perturbation calculations, an Adamowicz type MR CCSDT [67] curve was computed. The error curves with respect to the MR CCSDT calculation ($E_{\text{MP MBPT}} - E_{\text{MR CCSDT}}$) are also calculated and the results are shown at Fig. 2.11.

In Fig. 2.11 we see, that the zero-order approximation is qualitatively correct and the perturbation orders gradually improve upon it, showing smooth convergence. If considering the higher symmetry at the apex of the curve, we get a better estimation but a necessarily discontinuous potential curve. The parallelity error of the CAS is small compared to the *spin-adapted* second-order MP MBPT result.

As the examples of Section 1.4 show, the most dangerous point is at the 90° where the orbital degeneracy can lead to intruder state problem. Such an effect does not appear here.

2.4.4 BeH_2

A typical multi-reference problem is the description of BeH_2 molecule with different bond lengths and bond angles. Detailed description of this model can be found in Section 1.4. Here we used the same geometries and basis set for the calculations.

The zero-order description is less accurate for points F, G, H and I, which is responsible for the 20 mE_h parallelity error. This error is reduced by the second, third and fourth-order of *spin-adapted* MP MBPT to 10 mE_h, 5 mE_h and, 3 mE_h, respectively. These results are rather similar to those we obtained in the MP MCPT case.

We have not found any irregular behavior in these calculations either.

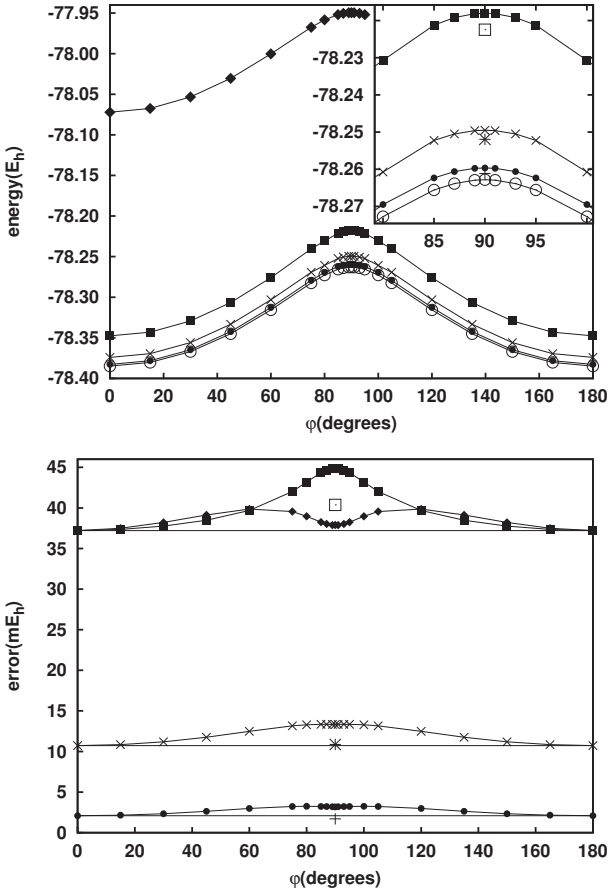


Figure 2.11: Energy of ground state of ethylene molecule for different torsion angles. Different orders of MP MBPT are compared to an Adamowicz type MR CCSDT[67]. Notations: diamond – CAS energies, square – second-order MP MBPT energies, cross – third-order MP MBPT energies, full circle – fourth-order energies, empty circle – MR CCSDT. Enlarged picture at 90° also shows results for D_{2d} symmetry, empty square – second-order MP MBPT energy, star – third order MP MBPT energy, plus sign – fourth-order MP MBPT energy. The second figure shows the parallelity of the error curve. CAS results are shifted by -275.4mE_h. See text for basis set.

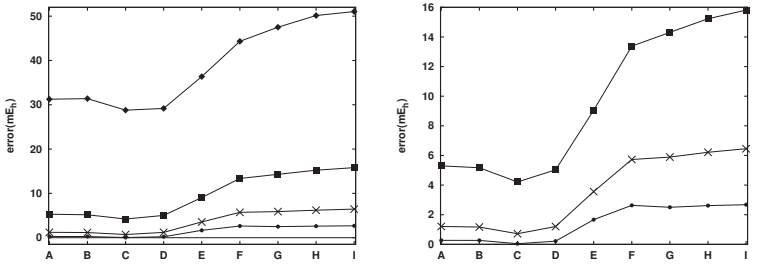


Figure 2.12: Error of different orders of *spin-adapted* MP MBPT calculations for the BeH_2 molecule. Notations: diamond – CAS energies, square – second-order MP MBPT energies, cross – third-order MP MBPT energies, full circle – fourth-order energies. See Section 1.4 for geometries and basis set.

2.5 Summary

Many-body theories using an effective Hamiltonian face numerical instability due to the intruder state problem if the determinants of the model space and determinants in the orthogonal space are not separated energetically. In single-reference MBPT the physical content of one-particle energies – i.e. the energies of the occupied orbitals are ionization potentials while the energies of the virtuals are electron affinities – ensures the energetical separation of the Hartree-Fock determinant and the virtual determinants. In the framework of MP MBPT, proposed by Malrieu and his coworkers[19], the energetical separation of the model space and the orthogonal space is ensured through the energetical separation of occupied and virtual one-particle energies. This is done by applying as many zero-order Hamiltonians as the number of determinants and by defining the one-particle energies as ionization potentials and electron affinities with respect to the given model space determinant.

This approach raises the theoretical question of spin and symmetry contamination. The former problem was solved by Malrieu and coworkers[78], while the latter is discussed by us [4]. We have shown, that the orbital degeneracy can lead to spin or spatial symmetry contamination which can be avoided by using a zero-order Hamiltonian invariant to orbital rotations with in degenerate subspaces.

Scaling properties of energy with system size, like separability and size-extensivity

was studied at second-order by Malrieu and his coworkers[19] and at higher-orders by us[4, 5]. We have pointed out that using different one-particle energy sets for each model space determinant can not be size-extensive from the third-order of theory. We have also shown that certain disconnected diagrams appear in the *spin-adapted* and *general symmetry-adapted* formulations at fourth and higher-orders. In the *spin-adapted* case the connected units of the disconnected diagrams are "linked" by at least one common spatial orbital. This means that at a given intervertex level one of the connected units contains an orbital α and the other contains the same orbital with β spin. In this way the *spin-adapted* MP MBPT is *size-extensive*. Similarly the *general symmetry-adapted* MP MBPT can be also *size-extensive* if the number of active orbitals in any of the degenerate subspaces does not scale with the system size.

Several numerical results have been already published at second-order [79, 19, 78, 87] which prove the intruder free behavior of MP MBPT. We have implemented MP MBPT approaches up to fourth-order using the diagrammatic approach worked out for MR MBPT. The implementation is performed using the automatized code generator tools developed for third-order MCPT. To improve the performance intermedier quantities are introduced. The most recent implementation is still not perfect, it leaves room for further introduction of intermediers. With the present form of the codes it is possible to perform calculations for medium size systems. The numerical results show that third and fourth-orders give significant improvement to the second-order energies, showing an apparently smooth convergence. This is opposite to previous theoretical arguments implying possible divergence of the MP MBPT series[79].

Chapter 3

Sparse Full-CI algorithm

3.1 Introduction

In quantum chemistry different parametrizations are used to describe the wave function. A given parametrization usually prescribes a way to get a hierarchy of possible approximations of the exact wave function. In principle such a hierarchy permits a systematic improvement of the quality of a given description. Well known examples are the CI and the CC hierarchies where any member of the hierarchy is connected to a chosen excitation level in either the wave function or the cluster expansion. In case of CI type functions it is usual to classify determinants according to their excitation level. If one is interested in the eventual importance of a given determinant in the Full-CI (FCI) function, a different classification scheme can be set up that may lead to a new hierarchy of approximations. It is widely known that excitation level and eventual importance in FCI may show just loose connection. For this reason an approach based on the importance of determinants in FCI becomes of value when only a relatively small number of determinants can be treated as compared to the FCI space, still a highly accurate description of the molecular system is needed.

Importance of a given determinant can be defined by, e.g., its magnitude in the FCI function or its contribution to the FCI energy in an *a posteriori* manner. In practice one requires an *a priori* selection rule to pick just the determinants whose importance is greater than given threshold. This Chapter describes an algorithm which finds important determinants in an *a priori* way and makes use of the space spanned by them to give an

approximation to the FCI solution[6]. The algorithm treats all excitation levels on the same footing and selects important determinants by using a threshold for the estimated contribution to the FCI energy in the next iteration step. Apart from small test cases, the algorithm is not capable of handling explicitly all determinants in the FCI space. Right on the contrary, due to limitations in system memory unavoidable even for medium sized systems, the algorithm is intentionally designed for treating just a small fraction of the FCI space and aims to reach the closest possible approximation to the FCI energy within these circumstances. In its purpose, applicability and realization the algorithm shows close relation to the FCI method, hence the Sparse FCI (SFCI) denomination is used.

Although the significance of CI type methods was evident from the beginning of quantum chemistry, practical (large-scale) CI calculation became possible only after Roos and Siegbahn introduced the so-called Direct CI algorithm[88]. In this technique, following a Lanczos-type philosophy, the CI eigenvector is obtained iteratively, without the need of ever loading the entire Hamiltonian matrix into computer memory. In each iteration step the Hamiltonian is let to act on the trial CI vector. The procedure makes use of the second quantized form of the Hamiltonian, and needs the integral list to obtain the resulting vector. When increasing system size, quartic scaling of the integral list appears negligible as compared to the exponential scaling of the FCI vector. Hence, memory requirement of a FCI calculation is determined by the need of storing the FCI vector.

It was realized a long time ago, that the length of the CI vector can be greatly reduced by projecting onto symmetry-adapted subspaces, corresponding, e.g., to spin quantum numbers. Efficient methods like the unitary group approach [89, 90, 91] exist which exploit this possibility, and they are of great help when performing truncated CI. As shown by Handy, pure determinants are, however, preferable to symmetry-adapted configuration state functions (CSF-s)[92] in the FCI case, because coupling coefficients (the weighting factors of Hamiltonian integrals) can only be 0 or $\pm$ 1 if a determinantal expansion of the FCI vector is adopted. For this reason, the present procedure – similarly to most of the current FCI implementations – applies a determinant-based representation.

Recent years in quantum chemistry have seen a steady progress in the size of the largest system ever computed by FCI, which is over 10^{10} determinants presently. Rossi et al.[93] reported a frozen core calculation for the N_2 molecule with 34 basis functions, using almost 10 billion determinants in 1994. A calculation on the CN anion in cc-pVDZ

basis performed by Thøgersen and Olsen[94] in 2004 included about 20 billion determinants. A series of calculations published by Gan et al.[95] in 2006 treated 45 billion determinants in some cases. The largest FCI calculation ever done according to our present knowledge was performed by Gan and Harrison[96] on the C_2 system in 2005, containing 65 billion determinants. This admirable expansion in system size is primarily thanks to the adaptation of the FCI algorithm to parallel computer architectures[97, 98, 99, 100, 101] and extensive use of powerful supercomputers.

In ordinary FCI algorithms like the ones referred to above, the storage of the FCI vector appears a waste of memory regarding the fact that many of its elements are zero (i.e. below a threshold) and can be safely neglected while still preserving the accuracy required. Making use of this fact may open a way to further enlargement of systems tractable by the FCI approach. One obstacle in the way of a sparse FCI code is the need of an effective algorithm that evaluates the action of the Hamiltonian on a sparse trial vector. Efforts in this line were published by Knowles [102, 103] and Mitrushenkov [104, 105]. These algorithms avoid explicit treatment of the unimportant determinants in vector C when performing the linear transformation $\sigma = \hat{H}C$, hence a reduction in computational time as well as smaller disk and memory storage is manifested by them.

In the framework of the FCI procedure Handy[92] introduced breaking down each determinant into an α and β string according to spin. Olsen et al.[86, 106] and Zarrabian et al.[107] further improved the idea by treating the $(\alpha\alpha)$, $(\beta\beta)$ and $(\alpha\beta)$ part of the Hamiltonian independently. Their algorithm scales like $N_{\mathrm{FCI}}\, occ^2\, virt^2$, where occ and $virt$ are the number of occupied and virtual orbitals. A string based direct CI algorithm following the philosophy of Olsen has been shown to be useful also in CI calculations restricted to a selected subspace[108].

In what follows we introduce a FCI algorithm that applies Olsen's idea for the linear transformation $\sigma = \hat{H}C$ and stores both C and σ as sparse vectors. The performance of the linear transformation shows some similarities with Mitrushenkov's dynamic CI algorithm [104], however, in our case the σ vector contains a small portion of the whole FCI space and the factorization of the present algorithm is more efficient. Due to the fact that the result of the action of $\hat{H}$ can be attained only in a subspace of the FCI space, the algorithm provides an *approximation* to the FCI energy. The error committed is connected to how large portion of the FCI space can be taken into account when representing vectors C

and σ.

Organization of Chapter 3 is as follows. Section 3.2 discusses the iteration algorithms previously for obtaining the FCI solution both with and without utilizing on sparsity. The essence of the linear transformation performed with sparse vectors is presented in Section 3.3. Numerical experience and results are collected in Section 3.5. Finally Section 3.6 gives a short conclusion and outlook.

3.2 Iteration scheme

3.2.1 The Davidson algorithm

The sparse algorithm we are going to discuss in Section 3.2.4 is a modification of the Davidson algorithm[109] thus it is practical to briefly summarize the fundamentals of the general theory.

Suppose that we have a trial function written as a linear combination of orthonormal vectors denoted by $|K\rangle$, with combination coefficients c_K. The Rayleigh-quotient gives the energy corresponding to the trial function:

$$E(\underline{c}) = \frac{\sum_{KL} c_K H_{KL} c_L}{\sum_J c_J^2} \,. \tag{3.1}$$

According to the variational principle, minimization of Eq. (3.1) gives the exact wave function and the energy of the ground state if the basis set is complete. If $\underline{c} + \underline{\delta}$ is the exact coefficient vector of the ground state, we have:

$$\left.\frac{dE(\underline{c})}{dc_I}\right|_{c_I+\delta_I} = 0. \tag{3.2}$$

Solving this equation at the first order, one obtains an approximation to δ_I:

$$\delta_I = \frac{\sum_K (H_{IK} - E(\underline{c})\delta_{KI}) c_K}{E(\underline{c}) - H_{II}} \,. \tag{3.3}$$

The above expression has a central role in the Davidson algorithm.

At each iteration step of the Davidson process, a subsequent element of an orthogonal basis set of the many-electron space is determined. In this basis the matrix elements of the Hamiltonian are calculated and it is used to expand the approximate FCI function, $\Psi^{(i)}$.

Let us assume $\Phi^{(i)}$ to be the i^{th} basis vector. To generate the $(i+1)^{\text{th}}$ basis vector, a linear transformation is performed first:

$$\sigma^{(i)} = \hat{H}\Phi^{(i)}. \tag{3.4}$$

Using $\sigma^{(i)}$ one can determine an additional row and column of the Hamiltonian matrix according to $H_{ij} = \langle\sigma^{(i)}|\Phi^{(j)}\rangle$, $j = 1, \ldots, i$. After the diagonalization of the i by i H matrix one can gets a new approximation to the energy denoted by E_i energy and the wave function, written as $\Psi^{(i)} = \sum_{j=1}^{i} c^{(j)}\Phi^{(j)}$. At the next step, a correction to $\Psi^{(i)}$ is calculated using Eq.(3.3):

$$\delta\Psi^{(i)} = \hat{R}(\hat{H} - E_i)\Psi^{(i)}, \tag{3.5}$$

where $\hat{R} = \sum_I \frac{|I\rangle\langle I|}{E_i - H_{II}}$ and the sum runs over the states expand the many-eletron space. Finally $\delta\Psi^{(i)}$ is orthogonalized to $\Phi^{(j)}$, $j = 1, \ldots, i$ and subsequently normalized to give the $(i+1)^{\text{th}}$ basis vector, $\Phi^{(i+1)}$. The iteration process stops when the last basis vector gives a negligible correction to $\Psi^{(i)}$.

As shown by Knowles and Handy [110], in the case where single determinants are used to expand the many-electron space, a modification of the denominator is needed to keep $\delta\Psi^{(i)}$ an eigenfunction of $\hat{S}^2$. Similarly to the case of the spin-adapted MP MBPT (2.2.3), one should avoid to apply different denominators for determinants with the same spatial occupation numbers. Knowles and Handy have solved this problem by using $\tilde{H}_{II}$ in Eq.(3.5) instead of H_{II} where $\tilde{H}_{II}$ is defined to depend only on the spatial occupancy, an averaged exchange energy term as detailed in [110].

3.2.2 Knowles' sparse algorithm

From our present point of view the progress made by Knowles is of special interest [102, 103], which takes into account the sparsity of basis vectors. In this approach important determinants are picked up among those which constitute a given basis vector, and unimportant ones are neglected. This induces a nonzero overlap among basis vectors and necessitates the treatment of an overlapping basis. The procedure of Knowles affects the Davidson algorithm at three main points:

- a generalized eigenvalue problem needs to be solved for the i by i H matrix, and the approximate $\Psi^{(i)}$ is expanded in a non-orthogonal basis

- having calculated $\delta\Psi^{(i)}$ only those determinants I are retained whose estimated contribution to the energy is greater than a given threshold. Such an estimation can be based on the perturbative formula

$$\Delta E_I \approx \frac{\langle I|\hat{H} - E_i|\Psi^{(i)}\rangle^2}{E_i - H_{II}} \tag{3.6}$$

- orthogonalization of $\delta\Psi^{(i)}$ to $\Phi^{(j)}$, $j = 1, \ldots, i$ is skipped in order to avoid the emergence of unimportant determinants. Simply $\delta\Psi^{(i)}$ is normalized to give the $(i+1)^{\text{th}}$ basis vector, $\Phi^{(i+1)}$.

It is important to emphasize here that even if $\Psi^{(i)}$ is sparse, $\sigma^{(i)} = \hat{H}\Psi^{(i)}$ is usually not a sparse vector. All its elements are needed to construct the accurate matrix elements of the i by i H matrix otherwise the variational property of the calculation would be lost. Since only important determinants contribute to $\Phi^{(i+1)}$ the I/O cost is significantly reduced. Although the convergence of this scheme is slower than the ordinary Davidson algorithm, keeping basis vectors short can decrease also the computation cost at the linear transformation step.

3.2.3 Mitrushenkov's sparse algorithm

Basis vectors $\Phi^{(i)}$ have to be stored on disk, and all of them are needed to represent the approximate wave function Ψ at a given iteration step. To avoid the large disk requirement and I/O cost necessary, several modifications of the above algorithm has been introduced [96, 97, 104, 106]. The common philosophy behind these methods is that instead of basis vectors $\Phi^{(k)}$ the latest $\Psi^{(i)}$ function is stored. At each iteration step the wave function correction $\delta\Psi$ is calculated and used to improve $\Psi^{(i)}$, then it is dropped.

As an example we briefly discuss Mitrushenkov's algorithm which is a modified version of the one developed by Olsen *et al* [106].

In this approach a two by two Hamiltonian is contructed at the i^{th} iteration step, on the basis states $\Psi^{(i)}$ and $\Psi^{(i-1)}$. Since these states are usually not orthogonal to each other, a generalized eigenvalue problem has to be solved. As a result, one gets the approximate FCI energy E_i, and one can replace $\Psi^{(i)}$ and $\sigma^{(i)}$ by $\alpha_i\Psi^{(i)} + \alpha_{i-1}\Psi^{(i-1)}$ and $\alpha_i\sigma^{(i)} + \alpha_{i-1}\sigma^{(i-1)}$, where α_i and α_{i-1} comes from the generalized eigenvalue problem. To construct $\Psi^{(i+1)}$,

$$\delta\Psi^{(i)} = \hat{R}(\hat{H} - E_i - \delta E)\Psi^{(i)} \tag{3.7}$$

with $\Psi^{(i+1)}$ is chosen proportional to $\Psi^{(i)} + \delta\Psi^{(i)}$, where

$$\delta E = \frac{\langle \Psi^{(i)} | \hat{R}(\hat{H} - E_i - \delta E) | \Psi^{(i)} \rangle}{\langle \Psi^{(i)} | \hat{R} | \Psi^{(i)} \rangle}. \tag{3.8}$$

gives an estimation of the energy correction introduced by $\Psi^{(i+1)}$. If the convergence is not achieved, the process starts from the beginning. The key point in the derivation of Eqs.(3.7) and (3.8) is to ensure the orthogonality of $\delta\Psi^{(i)}$ to $\Psi^{(i)}$. For details we refer to [104, 106].

In the above algorithm only two long vectors have to be stored, which significantly reduces the disk requirement and the I/O cost compared to the original Davidson algorithm.

3.2.4 A new sparse full-CI algorithm

Calculation time of the linear transformation could be further reduced if the effect of $\hat{H}$ was evaluated only in a certain well chosen subspace of the FCI space:

$$\sigma^{(i)} = \hat{P}_S \hat{H} \Phi^{(i)}. \tag{3.9}$$

Straightforward substitution of the action of $\hat{H}$ by formula (3.9) in the Davidson procedure is however not possible in the FCI framework. Simply neglecting unimportant determinants from $\hat{H}\Phi^{(i)}$ by the application of $\hat{P}_S$ would lead to erroneous Hamiltonian and overlap matrix elements and would cause the violation of the variation principle (unless $\Phi^{(i)}$ belongs to subspace $\hat{P}_S$).

In order to make use of the advantages of (3.9) a new FCI iteration scheme was designed [6]. At step i, we start from a given approximation for the wave function $\Phi^{(i)}$ and a well chosen subspace $S^{(i)}$. Subspace $S^{(i)}$ contains determinants which can hopefully give significant improvement to $\Phi^{(i)}$. For example, at the first step $\Phi^{(1)}$ can be the Hartree–Fock determinant and subspace $S^{(1)}$ is the full set of doubly excited determinants. Choosing subspace $S^{(i+1)}$ for the next iteration step will be discussed at the end of iteration step i. As a start of step i we construct another subspace $S^{(i)'}$ which collects

a) determinants constituting $\Phi^{(i)}$

b) those determinants belonging to $S^{(i)}$ which fulfill the hope of being important.

To perform the selection of item b) the interaction of $\Phi^{(i)}$ is calculated with all elements of $S^{(i)}$ and only those determinants are kept whose energy contribution based on the second order estimation

$$\Delta E_I \approx \frac{\langle I|\hat{H}|\Phi\rangle^2}{E - H_{II}} \tag{3.10}$$

is larger than a given threshold. (For technical reasons $S^{(i)}$ is not orthogonalized to $\Phi^{(i)}$. Determinants that are present both in $\Phi^{(i)}$ and $S^{(i)}$ are among the most important ones according to formula (3.10), but they are anyway selected at point a).) To calculate estimation (3.10) one needs to act with $\hat{H}$ on $\Phi^{(i)}$. This part of the algorithm is called the groping step, referring to the exploration of important regions of subspace $S^{(i)}$.

Iteration step i is continued by picking up the contribution of determinants from $S^{(i)'}$. This part of the algorithm is called the corrective step, indicating the point of improving $\Phi^{(i)}$ by determinants belonging to $S^{(i)'}$. For this end we essentially take one step by the Davidson procedure entering the algorithm *in medias res*. In brief, a $\delta\Phi$, belonging to subspace $S^{(i)'}$ is right away constructed utilizing the outcome of the groping step and the previous iteration step. Following this, $\delta\Phi$ is orthogonalized to $\Phi^{(i)}$ and a 2 by 2 Hamiltonian matrix is built in the space of $\Phi^{(i)}$ and the orthogonalized $\delta\Phi$. This involves calculation of the effect of $\hat{H}$ on $\delta\Phi$. The Hamiltonian matrix is diagonalized and its lower root is subjected to picking just the important components, based on formula (3.10), to obtain the next approximation to the FCI wave function $\Phi^{(i+1)}$. In principle more steps could also be taken by the Davidson algorithm to gain the main contribution of the $S^{(i)'}$ subspace, but a single step was found to be sufficient in practice.

At this point it is apparent that at each iteration step we apply the Davidson procedure in a different subspace $S^{(i)'}$. For this reason iteration steps of the proposed algorithm are independent and energies give an upper bound to the exact value in each iteration step. The algorithm allows a gradual growth of the subspace in which the estimated FCI function is represented. Independence of iteration steps on one hand gives the possibility to use restricted formula (3.9) for the linear transformation. On the other hand slower convergence is expected for the lack of collecting several basis vectors of the Davidson subspace.

Apart from $\Phi^{(i+1)}$, subspace $S^{(i+1)}$ also has to be specified to start iteration step $i + 1$. Determination of subspace $S^{(i+1)}$ is an important part of the algorithm since this decides which determinants have the chance to enter via the $(i+1)^{\text{th}}$ groping step. Because explicit

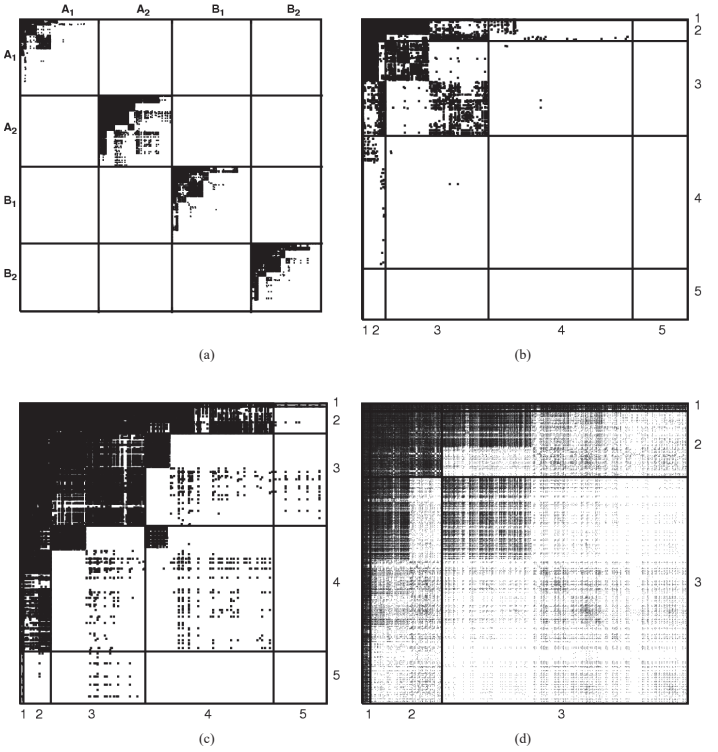


Figure 3.1: Crude presentation of the coefficient matrix of the H_2O molecule, marking nonzero elements by dots. The matrix is the outcome of an SFCI calculation, showing 7 μE_h error as compared to the exact FCI result. Subplots depict (a) full coefficient matrix; (b) A_1,A_1 block; (c) A_2,A_2 block; (d) densest section (top left corner) of the A_2,A_2 block. Numbers indicate excitation levels. See Section 3.5 for more details.

treatment of all determinants of the FCI space is to be avoided, some extra information is needed on how to find regions of the configuration space interesting for us. At this point it is useful to make the observation that picking elements of a FCI coefficient matrix C which give larger energy contribution than a given threshold, as computed by formula (3.10), shows a well defined structure (see Fig. 3.1). One finds that certain strings and string combinations are more important than others. We therefore introduce a selection rule based on strings and include every determinant into $S^{(i+1)}$ for which at least one of the strings is important. The expression we use to define the importance of string I_α,

$$W(I_\alpha) = \sum_{\{I_\beta, (I_\alpha, I_\beta) \in S^{(i)'}\}} \frac{\langle I_\alpha I_\beta | \hat{H} - E | \Phi \rangle^2}{E - E_{I_\alpha I_\beta}} \tag{3.11}$$

is simply the sum of approximated energy contributions of determinants $\langle I| = \langle I_\alpha I_\beta|$, I_β being such that $I_\alpha I_\beta$ is an element of $S^{(i)'}$. Formula (3.11) can be interpreted as a strategy of selecting determinants that resemble presently important ones in the sense that either their α or β part is already "successful".

We conclude this section by a step by step summary of the algorithm as follows:

1. Groping step. Restricted linear transformation $\hat{P}_{S^{(i)}} \hat{H} \Phi^{(i)}$, where the target vector is in subspace $S^{(i)}$, *i.e.* containing at least one important string according to (3.11).

2. Pick up important determinants from $S^{(i)}$ by condition (3.10). Construct subspace $S^{(i)'}$ as the union of important determinants from $S^{(i)}$ and those constituting $\Phi^{(i)}$.

3. Corrective step. One iteration by the Davidson algorithm.

   a) Calculation of $\delta\Phi = P_{S^{(i)'}} \hat{R}(\hat{H} - E)\Phi^{(i)}$.

   b) Orthogonalization of $\delta\Phi$ to $\Phi^{(i)}$:
   $\Phi^{\perp} \sim \delta\Phi - \Phi^{(i)} \langle \Phi^{(i)} | \delta\Phi \rangle$.

   c) Linear transformation in $S^{(i)'}$ space: $\hat{P}_{S^{(i)'}} \hat{H} \Phi^{\perp}$.

   d) Diagonalize 2 by 2 Hamiltonian matrix built in the basis of $\Phi^{(i)}$ and $\Phi^{\perp}$.

   e) Take the eigenvector of the lower root and drop unimportant determinants from it to obtain $\Phi^{(i+1)}$.

4. Obtain $S^{(i+1)}$ by formula (3.11).

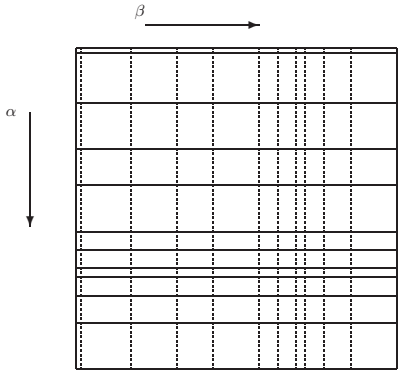


Figure 3.2: Structure of the σ vector at the groping step for the $M_s = 0$ case, ignoring symmetry. Solid lines show the determinants with important α strings selected by formula (3.11). Coefficients belonging to dashed lines are obtained by taking the same element (or its negative) from the corresponding important row.

The size of the subspace used for the FCI wave function expansion is an external input of the algorithm and it is primarily limited by system memory available. The threshold used at step 2 is a number fitted to the maximum size of the subspace and may be dynamically varied during the calculation.

3.3 Computational considerations

The most time consuming and computationally less trivial part of the algorithm is the calculation of linear transformation (3.9). In one iteration step two linear transformations have to be performed, one at the groping and another in the corrective step. Superficially it may seem that a third linear transformation is also present under step 3a). This transformation, however, needs not be explicitly performed as it can be calculated based on the groping step and the outcome of the previous corrective step. Product vector σ shows a different sparse structure at the groping and in the corrective step, since only contiguous rows are filled at the groping step while nonzero elements occupy σ in a totally

non-contiguous manner at the corrective step. (Contiguous columns of σ are filled based on symmetry considerations as a last action of the groping step, as detailed below.) This difference in the structure of σ calls for different strategies when calculating the effect of the Hamiltonian. Differences in the basic loop structures are outlined in Section 3.4, here we make some general remarks of technical relevance.

Considering the α string of a given determinant as row index and β occupation string as column index, it is usual to represent coefficient vectors C and σ as matrices:

$$\Phi \;=\; \sum_{IJ} C_{IJ}|I_\alpha, I_\beta\rangle$$

Following ordinary sparse matrix technology[111], we store only nonzero elements of both matrices by filling them into double precision vectors row by row. A supplementary integer vector of the same length is needed to give column indices, as well as another short integer vector to indicate the borderline of consecutive rows in the double precision and column index vectors. It is the storage of sparse matrices C and σ that gives the dominant memory requirement of the present procedure. There are three long double precision arrays to be stored at a time (C, σ and an auxiliary vector) together with their integer column and row index vectors.

Indexing arrays represent an obvious overhead in sparse matrix technology, considerable benefit is therefore obtained only for truly low populated matrices. A further caveat in connection with sparse matrix technology is the problem of vectorization. Since determinants of a given subspace do not generally correspond to a contiguous section of the coefficient matrix, there is no straightforward way to keep the vectorizable feature of Olsen's algorithm. Fortunately, as we shall demonstrate by sample calculations, in the proposed SFCI algorithm both of the above disadvantages are compensated by the extreme sparsity of arrays σ and C. This permits reduction of not only the CPU time but also the I/O cost of the procedure.

At the groping step, due to the special structure of array σ, it becomes needless to store column indices. This integer array requires half the memory occupied by the double precision array storing the coefficients. Dropping the integer vector frees memory which can be utilized to let array σ accommodate more elements, by about a factor 1.5. Letting $\hat{H}$ to address only appropriately selected rows at the groping step has the downside that spin symmetry of σ must be restored subsequently. For the singlet case this is performed

by transposing important rows to generate elements in important columns as well (see the dashed lines in Fig. 3.2). Prior to transposition, unimportant elements of important rows are dropped, which frees enough memory for double precision elements of the columns and the sparse index vector necessary from this point on for σ. Apart from a singlet wave function, the $M_S = 0$ triplet case is also trivial to handle. For other spin quantum numbers special considerations are needed at the groping step.

Most of the small molecules for which a FCI calculation is manageable also have spatial symmetry around equilibrium geometry, which can be used to reduce computational cost. For this end strings have to be reordered according to symmetry producing a coefficient matrix in block form. It is possible to work with symmetry blocks independently, in the following manner. Sparse matrix C is stored in main memory in its full length, but only one symmetry block (e.g., A_1,A_1 or B_1,B_1 etc. for the total symmetric i.r.) of array σ is dealt with at a time. This way (i) only symmetry allowed elements of σ are computed, reducing the number of multiplications, and (ii) memory is freed due to the reduced length of σ. Memory gain from point (ii) is exploited to enlarge subspace S at the groping step, which has the beneficial effect of accelerating convergence of the algorithm. Memory saving due to symmetry and elimination of the long index vector at the groping step permits to use a relatively long vector σ. For example it contains 12 % of the FCI space while array C is populated about 1 % for the SFCI calculation for molecule N_2, see the calculation involving 10^8 determinants in Table 3.3 in Section 3.5. This has to be compared with the 1 % length of vector σ in the corrective step in the same calculation.

The algorithm of the linear transformation performed between sparse vectors essentially follows the ideas of Olsen[86]. Making use of the sparsity of the initial and product vectors enhances a reduction of the calculation time of the procedure. In principle the number of double precision multiplications necessary in our sparse algorithm is less by a factor of $(N_{\mathrm{SFCI}}/N_{\mathrm{FCI}})^2$ as compared to the original procedure. Loss of computer efficiency on the other hand, as a consequence of the sparse representation of the coefficient matrices (e.g., no vectorization of the innermost loop) has a decelerating effect and admits an eventual $N_{\mathrm{SFCI}}/N_{\mathrm{FCI}}$ factor gain according to experience presented in Section 3.5. The sparse linear transformation is discussed in more detail in the forthcoming Section.

3.4 Linear transformation in the sparse algorithm

In this section we detail the algorithm followed when calculating the linear transformation (action of the Hamiltonian operator) on sparse trial vectors. As our sparse algorithm is a generalization of Olsen's method [106], we shall review the latter pointing out differences with the former. We use asterisks to indicate points where the two implementations differ. At the groping step the calculation of the linear transformation is rather similar to the algorithm used by Mitrushenkov[104, 105]. The key step in Olsen's procedure is the decomposition of the

$$\hat{H} = \sum_{kl} h_{kl}\hat{E}_{kl} + \frac{1}{2}\sum_{ijkl}(ij|kl)\left(\hat{E}_{ij}\hat{E}_{kl} - \delta_{jk}\hat{E}_{il}\right) \tag{3.12}$$

Hamiltonian and matrix σ according to the spin of the excitation operators:

$$\hat{H} = \hat{H}^{\alpha\alpha} + \hat{H}^{\beta\beta} + \hat{H}^{\alpha\beta} \tag{3.13}$$

$$\sigma(I_\alpha, I_\beta) = \sigma^{\alpha\alpha}(I_\alpha, I_\beta) + \sigma^{\beta\beta}(I_\alpha, I_\beta) + \sigma^{\alpha\beta}(I_\alpha, I_\beta). \tag{3.14}$$

where h_{kl} and $(ij|kl)$ denote the one- and two-particle integrals the latter in (11|22) notation and $\hat{E}_{kl}$ is a generator of the linear group which can be decomposed into an alpha and a beta part: $\hat{E}_{kl} = \hat{E}^{\alpha}_{kl} + \hat{E}^{\beta}_{kl}$.

A Same spin excitation

The explicit form of $\sigma^{\alpha\alpha}$ can be given in the following way:

$$\sigma^{\alpha\alpha}(I_\alpha, I_\beta) = \sum_{J_\alpha}\left\{\sum_{kl}\langle J_\alpha|\hat{E}^{\alpha}_{kl}|I_\alpha\rangle\tilde{h}_{kl} + \frac{1}{2}\sum_{ijkl}\langle J_\alpha|\hat{E}^{\alpha}_{ij}\hat{E}^{\alpha}_{kl}|I_\alpha\rangle(ij|kl)\right\}C(J_\alpha, I_\beta) \tag{3.15}$$

where the $\tilde{h}_{kl} = h_{kl} - \frac{1}{2}\sum_j(kj|jl)$ notation has been introduced. It is easy to realize that the expression in curly brackets is independent of I_β. Introducing an intermediate quantity denoted by $F_{I_\alpha}(J_\alpha)$, the factorized expression is written as:

$$\sigma^{\alpha\alpha}(I_\alpha, I_\beta) = \sum_{J_\alpha} F_{I_\alpha}(J_\alpha)C(J_\alpha, I_\beta) \tag{3.16}$$

Performing the multiplications in (3.16) is the most time consuming part of the calculation of $\sigma^{\alpha\alpha}$. For a given I_α there are approximately $\frac{1}{4}\,occ^2\,virt^2$ J_α's for which $F_{I_\alpha}(J_\alpha)$ is

nonzero. The calculation cost is therefore proportional to $\frac{1}{4} N_{\mathrm{FCI}}\, occ^2\, virt^2$, without symmetry considerations. To calculate $\sigma^{\beta\beta}$ one can proceed on the basis of an expression similar to (3.15). However, for singlet/triplet states, in Olsen's algorithm it is enough to calculate only $\sigma^{\alpha\alpha}$ explicitly, since $\sigma^{\beta\beta}$ can be obtained by the relation

$$\sigma^{\beta\beta}(J, I) = \pm\sigma^{\alpha\alpha}(I, J) \tag{3.17}$$

At the groping step of the sparse implementation index I_α is restricted to selected strings, moreover multiplications in (3.16) are performed only for nonzero elements of array C. At the corrective step I_α goes through the whole set of strings but at (3.16) a multiplication is calculated only if both $|I_\alpha, I_\beta\rangle$ and $|J_\alpha, I_\beta\rangle$ are elements of subspace $S^{(i)'}$.

At the groping step implementation of $\sigma^{\alpha\alpha}$ is as follows:

Loop over $I_\alpha \in$ { important strings } (*) $\sim N_{\mathrm{SFCI}}/\sqrt{N_{\mathrm{FCI}}}$
  $\underline{F} = 0, \underline{V} = 0$
  Loop over $(k_\alpha^+ l_\alpha)$
    $K_\alpha = \pm(k_\alpha^+ l_\alpha) I_\alpha$
    $F(K_\alpha) = F(K_\alpha) \pm \tilde{h}_{kl}$
    Loop over $(i_\alpha^+ j_\alpha)$
      $J_\alpha = \pm(i_\alpha^+ j_\alpha) K_\alpha$
      $F(J_\alpha) = F(J_\alpha) \pm (ij|kl)$
    End of loop over $(i_\alpha^+ j_\alpha)$
  End of loop over $(k_\alpha^+ l_\alpha)$
  Loop over J_α, (max. 2-fold ex. wrt I_α) $\sim \frac{1}{4}\, virt^2\, occ^2$
    Loop over $I_\beta \in C_{J_\alpha}$ (**) $\sim N_{\mathrm{SFCI}}/\sqrt{N_{\mathrm{FCI}}}$
      $V(I_\beta) = C(J_\alpha, I_\beta) F(J_\alpha)$
    End of loop over I_β
  End of loop over J_α
  Loop over I_β
    $\sigma^{\alpha\alpha}(I_\alpha, I_\beta) = \sigma^{\alpha\alpha}(I_\alpha, I_\beta) + V(I_\beta)$
  End of loop over I_β
End of loop over I_α

Notation $I_\beta \in C_{J_\alpha}$ indicates that string I_β has to be present in row J_α of matrix C.

At (*) it is the structure of array σ that gives rise to the restriction. Restriction at (**), emerging due to the sparseness of array C, reduces number of multiplications at the innermost loop. At the same time it makes it impossible to vectorize the loop since indices I_β selected by the condition do not fill vector V contiguously.

Focusing on the cost of the innermost loop, we give the time requirement of critical loops in the expressions aligned to the right. Length of the sparse CI vector N_{SFCI} appearing in the formulae is directly determined by the memory passed to the groping step. In the non sparse case the number of strings is given by $\sqrt{N_{\text{FCI}}} = \binom{nb}{occ}$ disregarding symmetry, where *nb* stands for the number of basis functions. Based on this we estimate the number of strings in a given row in the sparse case by simple rational relation as $N_{\text{SFCI}}/\sqrt{N_{\text{FCI}}}$. This gives a rough measure of the number of important strings as well as the average number of beta strings present in row J_α of matrix C, as indicated on the right margin of the innermost loop. Altogether the number of multiplications in the innermost loop is proportional to

$$\frac{1}{4} N_{\text{FCI}} \left(\frac{N_{\text{SFCI}}}{N_{\text{FCI}}} \right)^2 occ^2\, virt^2 \; . \tag{3.18}$$

In our present implementation of $\sigma^{\alpha\alpha}$ efficient use of restriction at (**) was not achieved, overall scaling of the routine shows linear dependence on $\frac{N_{SFCI}}{N_{FCI}}$ instead of the second power indicated in (3.18).

Apart from restricting indices of arrays σ and C in the loops, the present sparse implementation affects the calculation of $\sigma^{\beta\beta}$, too. In contrast to Olsen's algorithm, $\sigma^{\beta\beta}$ has to be explicitly computed at the groping step, since spin symmetry of σ is intentionally destroyed to gain memory in return. Making use of relation (3.17), $\sigma^{\beta\beta}$ is computed as

$$\sigma^{\beta\beta}(I_\alpha, I_\beta) = \sigma^{\alpha\alpha}(I_\beta, I_\alpha) = \sum_{J_\alpha} F_{I_\beta}(J_\alpha) C(J_\alpha, I_\alpha) \; . \tag{3.19}$$

It can be shown that routine $\sigma^{\beta\beta}$ involves the same number of multiplications as $\sigma^{\alpha\alpha}$. Note that $\sigma^{\beta\beta}(I_\alpha, I_\beta)$ is still stored in *row* I_α and *column* I_β of the product matrix (see Fig. 3.2).

At the corrective step implementation of $\sigma^{\alpha\alpha}$ is as follows:

Loop over I_α $\qquad\sim \sqrt{N_{\text{FCI}}}$

  $\underline{F}$=0

  Loop over $(k_\alpha^+ l_\alpha)$

    $K_\alpha = \pm(k_\alpha^+ l_\alpha) I_\alpha$

$F(K_\alpha) = F(K_\alpha) \pm \tilde{h}_{kl}$

Loop over $(i^+_\alpha j_\alpha)$

$J_\alpha = \pm(i^+_\alpha j_\alpha)K_\alpha$

$F(J_\alpha) = F(J_\alpha) \pm (ij|kl)$

End of loop over $(i^+_\alpha j_\alpha)$

End of loop over $(k^+_\alpha l_\alpha)$

Loop over J_α, (max. 2-fold ex. wrt I_α) $\sim \frac{1}{4}\, occ^2\, virt^2$

Loop over I_β, ($I_\beta \in \sigma_{I_\alpha}$ and $I_\beta \in C_{J_\alpha}$) (*) $\sim N^2_{\mathrm{SFCI}}/N_{\mathrm{FCI}}$

$\sigma^{\alpha\alpha}(I_\alpha, I_\beta) = \sigma^{\alpha\alpha}(I_\alpha, I_\beta) + C(J_\alpha, I_\beta)F(J_\alpha)$

End of loop over I_β

End of loop over J_α

End of loop over I_α

In principle the number of multiplications in the innermost loop is proportional to $N^2_{\mathrm{SFCI}}/N_{\mathrm{FCI}}$ (average number of strings in row J_α of C times the average reduction due to imposing I_β to be also an element of row I_α of σ). In practice however comparison of two ordered lists must be performed in the innermost loop, the time requirement being proportional to the length of the lists, i.e. $N_{\mathrm{SFCI}}/\sqrt{N_{\mathrm{FCI}}}$. For this reason overall calculation time of $\sigma^{\alpha\alpha}$ in the corrective step is proportional to $\frac{1}{4}\, N_{\mathrm{SFCI}}\, occ^2\, virt^2$, while the number of multiplications scales as (3.18).

Explicit calculation of the effect of $\hat{H}^{\beta\beta}$ is not needed in the corrective step. Vector $\sigma^{\beta\beta}$ can be obtained directly from relation (3.17).

B Different spin excitation

The formula for different spin excitation

$$\sigma^{\alpha\beta}(I_\alpha, I_\beta) = \sum_{J_\alpha J_\beta} \sum_{ijkl} (ij|kl)\langle J_\alpha|\hat{E}^\alpha_{ij}|I_\alpha\rangle\langle J_\beta|\hat{E}^\beta_{kl}|I_\beta\rangle C(J_\alpha, J_\beta)$$

is factorized in Olsen's algorithm by sorting the terms as

$$\begin{aligned}\sigma^{\alpha\beta}(I_\alpha, I_\beta) &= \sum_{kl}\sum_{J_\alpha}\left\{\sum_{ij}(ij|kl)\langle J_\alpha|\hat{E}^\alpha_{ij}|I_\alpha\rangle\right\} \\ &\times \left\{\sum_{J_\beta}\langle J_\beta|\hat{E}^\beta_{kl}|I_\beta\rangle C(J_\alpha, J_\beta)\right\}.\end{aligned}$$

For a given kl the quantity in the second curly brackets does not depend on I_α, therefore it can be evaluated prior to the calculation of the term in the first curly brackets. Introducing intermediate quantity $\sigma_{kl}^{\alpha\beta}(I_\alpha, I_\beta)$ the expression can be rewritten as

$$\sigma^{\alpha\beta}(I_\alpha, I_\beta) = \sum_{kl} \sigma_{kl}^{\alpha\beta}(I_\alpha, I_\beta)$$

To calculate $\sigma_{kl}^{\alpha\beta}(I_\alpha, I_\beta)$ one first sets up a list of all (I_β, J_β) pairs for which $\langle J_\beta|\hat{E}_{kl}^{\beta}|I_\beta\rangle$ is nonzero. Length of such a list is much shorter than the total number of strings, it is exactly $\binom{nb-1}{occ-1}$ if $k = l$ and $\binom{nb-2}{occ-1}$ if $k \neq l$. String pairs connected by excitation $\hat{E}_{kl}^{\beta}$ are indexed by I, right and left members of the pairs are stored in index vectors $R(I) = I_\beta$, $L(I) = J_\beta$ and the sign of the excitation is put into $sign(I) = \langle J_\beta|\hat{E}_{kl}^{\beta}|I_\beta\rangle$. Noticing that for a given kl and fixed I_β the sum in the second curly brackets can extend at most for a single term, one can write

$$\begin{aligned}\sigma_{kl}^{\alpha\beta}(I_\alpha, R(I)) &= \sum_{J_\alpha} \left\{ \sum_{ij} (ij|kl)\langle J_\alpha|\hat{E}_{ij}^{\alpha}|I_\alpha\rangle \right\} \\ &\quad \times sign(I) C(J_\alpha, L(I))\end{aligned}$$

For a given kl, intermediate quantity $C'(J_\alpha, I) = sign(I)C(J_\alpha, L(I))$ is calculated for all possible I and J_α. After this, intermediate $F_{I_\alpha}(J_\alpha) = \sum_{ij}(ij|kl)\langle J_\alpha|\hat{E}_{ij}^{\alpha}|I_\alpha\rangle$ is computed for all I_α and J_α. Finally the two intermediates are multiplied in the innermost loop according to:

$$\sigma_{kl}^{\alpha\beta}(I_\alpha, R(I)) = \sum_{J_\alpha} F_{I_\alpha}(J_\alpha) C'(J_\alpha, I). \tag{3.20}$$

Number of multiplications in the innermost loop in Olsen's algorithm is approximately $N_{\mathrm{FCI}}\, occ^2\, virt^2$. Contributions to this cost come from loop for kl and I, which together produce $nb\ \binom{nb-1}{occ-1}$ terms for the $k = l$ case plus $nb(nb-1)\ \binom{nb-2}{occ-1}$ terms for the $k \neq l$ case. This is to be multiplied by the number of I_α's which is $\sqrt{N_{\mathrm{FCI}}}$, times the number of possible J_α's, which is given by $occ\, virt$.

In the $M_S = 0$ case the above computational cost can be divided by two using restriction $(ij) \leq (kl)$, which requires just a small modification of the integral list[86]. This reduction can be exploited at the corrective step of the SFCI algorithm, but not at the groping step.

In the sparse implementation restrictions at the groping step affect strings I_α that are considered to be important as well as indices I that point to nonzero element in row J_α of matrix C'. At the corrective step string I_α is unrestricted but I must be such that both $R(I)$ and $L(I)$ point to a nonzero element in the corresponding matrices, i.e. $R(I)$ is a valid element in row I_α of matrix σ and $L(I)$ is present in row J_α of matrix C'.

At the groping step implementation of $\sigma^{\alpha\beta}$ is as follows:

Loop over $(k^+_\beta l_\beta)$ $\sim nb$ if $k = l$, $\sim nb(nb-1)$ otherwise

$L(I) = \pm(k^+_\beta l_\beta)R(I)$

$C'(J_\alpha, I) = \pm C(J_\alpha, L(I))$, $\forall I$ and $\forall J_\alpha \in C_{L(I)}$

Loop over $I_\alpha \in$ { important strings }(*) $\sim N_{\mathrm{SFCI}}/\sqrt{N_{\mathrm{FCI}}}$

Loop over $(i^+_\alpha j_\alpha)$

$J_\alpha = \pm(i^+_\alpha j_\alpha)I_\alpha$

$F(J_\alpha) = F(J_\alpha) \pm (ij|kl)$

End of loop over $(i^+_\alpha j_\alpha)$

Loop over J_α ($F(J_\alpha) \neq 0$) $\sim occ\ virt$

Loop over $I \in C'_{J_\alpha}$ (**) $\sim \frac{N_{\mathrm{SFCI}}}{N_{\mathrm{FCI}}}\binom{nb-1}{occ-1}$ if $k = l$, $\sim \frac{N_{\mathrm{SFCI}}}{N_{\mathrm{FCI}}}\binom{nb-2}{occ-1}$ otherwise

$V(I) = V(I) + F(J_\alpha)C'(J_\alpha, I)$

End of loop over I

End of loop over J_α

$\sigma^{\alpha\beta}(I_\alpha, R(I)) = \sigma^{\alpha\beta}(I_\alpha, R(I)) + V(I), \forall I$

End of loop I_α

End of loop $(k^+_\beta l_\beta)$

As before, notation $I \in C'_{J_\alpha}$ means that I has to be a valid element in row J_α of matrix C'. Note, that vectorization of the innermost loop is again ruined because of the restriction imposed on index I.

Estimated cost of critical loops indicated in right aligned expressions give rise to approximately $N_{\mathrm{FCI}}\left(\frac{N_{\mathrm{SFCI}}}{N_{\mathrm{FCI}}}\right)^2 occ^2\ virt^2$ multiplications in the innermost loop. Time measurements on examples presented in Section 3.5 indicate however, that overall scaling of our implementation is linear in $\frac{N_{SFCI}}{N_{FCI}}$, pointing to a worse than ideal use of sparsity at the innermost loop.

At the corrective step implementation of $\sigma^{\alpha\beta}$ is as follows:

Loop over $(k_\beta^+ l_\beta)$ $\sim nb$ if $k = l$,

$\sim nb(nb-1)$ otherwise

$L(I) = \pm(k_\beta^+ l_\beta)R(I)$

$C'(J_\alpha, I) = \pm C(J_\alpha, L(I))$, $\forall I$ and $\forall J_\alpha \in C_{L(I)}$

Loop over I_α $\sim \sqrt{N_{\rm FCI}}$

Loop over $(i_\alpha^+ j_\alpha)$

$J_\alpha = \pm(i_\alpha^+ j_\alpha)I_\alpha$

$F(J_\alpha) = F(J_\alpha) \pm (ij|kl)$

End of loop over $(i_\alpha^+ j_\alpha)$

If$(R(I) \in \sigma_{I_\alpha})LOG(I) = .TRUE.$, $\forall I$ (*)

Loop over J_α $(F(J_\alpha) \neq 0)$ $\sim occ\ virt$

Loop over $I \in C'_{J_\alpha}$ (**) $\sim \frac{N_{\rm SFCI}}{N_{\rm FCI}} \binom{nb-1}{occ-1}$ if $k = l$,

$\sim \frac{N_{\rm SFCI}}{N_{\rm FCI}} \binom{nb-2}{occ-1}$ otherwise

If $(LOG(I))$ $\sim \frac{N_{\rm SFCI}}{N_{\rm FCI}}$

$V(I) = V(I) + F(J_\alpha)C'(J_\alpha, I)$

End of loop over I

End of loop over J_α

$\sigma^{\alpha\beta}(I_\alpha, R(I)) = \sigma^{\alpha\beta}(I_\alpha, R(I)) + V(I), \forall I$

End of loop I_α

End of loop $(k_\beta^+ l_\beta)$

Accounting for all critical loops, number of multiplications in the innermost loop is given by $N_{\rm FCI}\left(\frac{N_{\rm SFCI}}{N_{\rm FCI}}\right)^2 occ^2\, virt^2$. The rate determining step however is again rather the comparison of two ordered lists at (*) that leads to a final time requirement proportional to $\frac{N_{SFCI}}{N_{FCI}}$.

3.5 Numerical experience and results

To illustrate the efficiency and limitations of the SFCI algorithm pilot numerical calculations are presented on the example of the H_2O and N_2 molecule. All calculations reported were performed on an AMD Opteron(tm) 250 machine. The 8 Gbyte system memory of this computer determines the upper limit to the length of the SFCI vector presently

tractable in our laboratory. Apart from details and outcome of SFCI calculations we provide results and timings obtained by our implementation of Olsen's original FCI algorithm as well.

An example providing insight into the structure of the sparse wave function is given by the calculation performed on the water molecule in 6-311G basis at Hartree–Fock optimized geometry (r = 0.9455 Å, $\theta = 111.881^o$). Full use of C_{2v} symmetry of the molecule was made, along the lines discussed in Section 3.3. There are 19 basis functions for 10 electrons in this system generating almost 34 million determinants belonging to the total symmetric i.r. In this small example a maximum of half million determinants were let to enter array C at any iteration step. Within these circumstances the SFCI algorithm provides −76.174816 a.u. for total energy, which is in 7 μE_h error with the −76.174823 a.u. value of the exact FCI energy. Compared with Olsen's original algorithm, there is one sixth reduction in the average computation time of an iteration step: it takes 263 seconds in Olsen's algorithm and 41 seconds in the SFCI algorithm. Number of iteration steps is another factor to be taken into account. In the original implementation, applying the Davidson algorithm 7 iteration steps are needed while the SFCI algorithm converges only in 11 steps. There still remains a considerable time reduction in the overall time requirement of the SFCI calculation, with an energy error on the order of a few microhartree as a price to pay.

Structure of the wave function, i.e matrix C is shown in Fig. 3.1 by putting a mark in the place of nonzero determinants. (Regarding the accuracy of this calculation and the crudeness of the representation in Fig. 3.1, we can safely state that Fig. 3.1 reflects the structure of the exact FCI wave function truncated at 10^{-11} a.u. energy threshold calculated according to formula (3.10).) First of all, block structure of matrix C generated by reordering strings according to i.r. is apparent in Fig. 3.1a. As the ground state belongs to i.r. A_1, matrix C adopts a block diagonal form. Population of different blocks is not balanced: symmetry block A_2,A_2 containing the Hartree–Fock determinant is the most populated, around 2 percent. The least populated block is A_1,A_1 containing less than 0.6 percent of all possible elements. Within symmetry blocks strings are ordered by excitation level (with respect to the Hartree–Fock determinant), which is marked on the axes in Figs. 3.1b-d. Fig. 3.1b and Fig. 3.1c show the structure of individual blocks A_1,A_1 and A_2,A_2 respectively. Fig. 3.1d gives an insight into the fine structure of the densest portion of the

N_{SFCI}	$N_{\mathrm{SFCI}}/N_{\mathrm{FCI}}$	$\Delta_{SFCI}(mE_h)$	It. steps	$t_{SFCI}(s)$	$\Delta_{\overline{SFCI}}(mE_h)$	$t_{\overline{SFCI}}(s)$
1		233.0091				
$0.5 \cdot 10^6$	0.1 %	0.1556	44	107	0.1320	1100
$1 \cdot 10^6$	0.2 %	0.0600	26	149	0.0553	1200
$2 \cdot 10^6$	0.4 %	0.0207	21	223	0.0189	1300
$5 \cdot 10^6$	1 %	0.0047	15	440	0.0044	1500
$1 \cdot 10^7$	2 %	0.0013	13	840	0.0013	2000
$1.5 \cdot 10^7$	3 %	0.0006	13	1200	0.0006	2200
$2 \cdot 10^7$	4 %	0.0003	12	1700		
$4 \cdot 10^7$	9 %	0.0000	13	3600		
$6 \cdot 10^7$	13 %	0.0000	13	5800		
$4.5 \cdot 10^8$		-76.2654082	9	14000		

Table 3.1: Convergence of SFCI and $\overline{\mathrm{SFCI}}$ algorithms to the exact energy of the H_2O molecule as the sparse vector is allowed to lengthen. Number of determinants in the FCI space and the exact FCI energy is indicated in a separate row at the bottom, together with characteristics of the calculation performed by Olsen's original algorithm. Energy errors by SFCI and $\overline{\mathrm{SFCI}}$ are designated as Δ_{SFCI} and $\Delta_{\overline{SFCI}}$ respectively. Notation t_{SFCI} and $t_{\overline{SFCI}}$ refer to the wall-clock time of an average iteration step. Column 4 collects number of iteration steps till the energy difference between consecutive steps becomes less than 10^{-7} a.u. The $\overline{\mathrm{SFCI}}$ algorithm took 13 iteration steps to converge in all cases tabulated.

A_2,A_2 block. Note the stripy pattern of the figures. This supports the somewhat intuitive selection formula (3.11) of the groping step, suggesting that certain strings can be indeed considered more important than others. An overall tendency of diminishing importance with increasing excitation level is also reflected in Figs. 3.1.

As a consequence of the fact that determinant selection into subspace S is controlled by formula (3.11), it may happen that some determinants never enter, not even the groping step, hence they are completely excluded from the resulting SFCI vector. Examining the above 6-311G water calculation from this point of view one finds that the number of determinants never treated amount to 20 million, 60 % of the FCI space. This observation raises the question whether it would be beneficial to ensure that each determinant has the chance at least once to enter the groping step. In order to answer the question we

N_{SFCI}	0	1	2	3	4	5
$0.5 \cdot 10^6$	1	70	3466	52440	306637	76882
$1 \cdot 10^6$	1	72	3493	68204	503141	266018
$2 \cdot 10^6$	1	72	3503	78640	715310	677886
$5 \cdot 10^6$	1	72	3503	84704	935212	1760678
$10 \cdot 10^6$	1	72	3503	86494	1049201	3280704
$20 \cdot 10^6$	1	72	3503	87230	1130157	5443656
$40\cdot$	1	72	3503	87568	1178226	7635726
$452 \cdot 10^6$	1	72	3503	87720	1199982	9194616

N_{SFCI}	6	7	8	9	10
$0.5 \cdot 10^6$	29352	734	39	0	0
$1 \cdot 10^6$	114658	3468	217	0	0
$2 \cdot 10^6$	460259	14086	1122	0	0
$5 \cdot 10^6$	1971420	110732	12589	8	0
$10 \cdot 10^6$	4841974	422112	63904	408	24
$20 \cdot 10^6$	10570073	1685538	357357	4688	262
$40\cdot$	20838768	7015076	2141973	37950	3947
$452 \cdot 10^6$	40596500	103890720	150252168	112650624	33807600

Table 3.2: Number of determinants in the SFCI wave function of the H_2O molecule, classified according to excitation level with respect to the Hartree–Fock determinant. A separate row at the bottom collects corresponding values of the FCI wave function.

performed test calculations where matrix σ at the groping step included the *whole* FCI space in *all* iterations while matrix σ at the corrective step and matrix C in both steps was kept as short as before. The number of determinants involved this way (denoted $\overline{\text{SFCI}}$) is larger than in SFCI. Since both methods are variational, the energy obtained by $\overline{\text{SFCI}}$ lies lower than SFCI. The energy decrease in triple-zeta water case is $2\,\mu E_h$, i.e. gives -76.174818 a.u. total energy, at the price of an increase of an average iteration step's time to 120 seconds, and no decrease in the number of iteration steps till convergence. Regarding that the calculation took considerably longer while the order of magnitude of the error did not diminish, we consider it unworthy to pursue the $\overline{\text{SFCI}}$ strategy. At the same time, it is reassuring, that the difference between $\overline{\text{SFCI}}$ and SFCI energies, which can be attributed to the error of formula (3.11), is not a predominant part of the total error of the SFCI energy. The error still present in the $\overline{\text{SFCI}}$ energy can be attributed to the truncated nature of the space where the SFCI vector is represented.

To illustrate the efficiency of the SFCI algorithm we carried out a series of calculations still on the example of the water molecule in 6-311G* basis set (24 functions), at geometry $r = 0.9394$ Å, $\theta = 107.5^o$. In this series we gradually increased the allowed length of the final SFCI wave function from calculation to calculation, starting from 0.1 % of the FCI space and finishing at 13 %. Errors of total energies and number of iteration steps for each calculation are collected in Table 3.1. Results of both SFCI and $\overline{\text{SFCI}}$ computations are indicated to assess the error of formula (3.11) on the basis of a broader range of examples. Olsen's original algorithm is still feasible for this small system, therefore all values can be compared with the exact FCI result. In Table 3.1 we see that 0.1 mE_h accuracy is safely obtained by considering only 0.2 % of the FCI space, and microhartree accuracy is achieved at 2 % population. Extending the allowed length of the SFCI vector from 0.1 % to 1 % is accompanied by a spectacular drop in the number of iteration steps, from 44 to 15. From this point on the number of iterations in the SFCI procedure changes only slightly and it stabilizes at a somewhat larger value than the number of iterations needed for Davidson's algorithm. Calculation time of an iteration step on average is roughly one hundredth of the time of a step in Olsen's original algorithm at 0.2 % population and roughly one tenth at 3 % population of the final SFCI vector. These data reflect the linear speedup as compared to the original Olsen algorithm discussed in Section 3.4. Altogether, a linear regression fits well for the computation times collected in Table 3.1

as a function of the population of the SFCI vector. It is interesting to note however, that over 15 % population of matrix C the original, non-sparse algorithm of Olsen takes over the sparse implementation in terms of wall-clock time. In the non-sparse calculation presented in Table 3.1 the linear transformation itself needed 79 % of the full calculation time while in the sparse case 73 % of the full calculation time was spent on the two linear transformations in the smallest SFCI example and 80 % in the largest one. Regarding $\overline{\text{SFCI}}$ results, one sees hardly any improvement in energy over 1 % population. Letting the whole FCI space be tested in the groping step appears important for extremely short SFCI vectors, between 0.1 % and 0.4 % population. However, even in these cases correction by $\overline{\text{SFCI}}$ does not affect the first significant digit of the error of SFCI. Above 3 % population there was no point in computing $\overline{\text{SFCI}}$ numbers.

Table 3.2 gives an insight into the composition of the SFCI wave function for the same series of calculations, by showing the number of nonzero elements (selected based on formula (3.10)) belonging to a given excitation level. Analyzing the entries of the table one observes a gradual shift in the excitation level that accommodates the most elements: at 0.2 % population these are the 4-fold excitations, at 0.4 % about the same number of determinants belong to 4-fold and 5-fold excitations, at 1 % population 5-fold and 6-fold excitations are the most numerous and 6-fold excitations remain to be the greatest in number till 13 % population of the SFCI vector.

As a final example we consider the N_2 molecule near equilibrium geometry, in a relatively large atomic natural orbital (ANO) basis set in [4s3p1d] contraction[112] at geometry $r = 2.1$ Bohr. There are 34 basis functions and 10 valence electrons in this system, generating almost 10^{10} determinants of A_1 symmetry. This is already too large for the original algorithm to be calculated with our facilities. As a reference, we quote the parallel FCI results of Evangelisti et al.[99, 93] on this system. Table 3.3 again presents energy errors with increasing population of the final SFCI vector.

Trends shown by the table are very similar to the previous example: microhartree accuracy is obtained by 2 % population, while for millihartree accuracy already 0.01 % population is sufficient. Linear scaling of the calculation time with the N_{SFCI} population is plotted in Fig. 3.3. An SFCI vector around two hundred million elements is the largest example we present here, where memory consumption of the SFCI calculation was 8 Gbyte and storage of the sparse coefficient matrix needed 2.3 Gbyte memory. Regarding

N_{SFCI}	$N_{\mathrm{SFCI}}/N_{\mathrm{FCI}}$	$\Delta_{SFCI}(mE_h)$	$t_{SFCI}(s)$
1		348.772	
$1 \cdot 10^6$	0.01 %	0.969	1050
$2 \cdot 10^6$	0.02 %	0.477	1250
$5 \cdot 10^6$	0.05 %	0.197	1750
$10 \cdot 10^6$	0.10 %	0.092	2700
$25 \cdot 10^6$	0.26 %	0.032	4800
$50 \cdot 10^6$	0.52 %	0.012	8700
$100 \cdot 10^6$	1.03 %	0.004	17000
$200 \cdot 10^6$	2.07 %	0.001	34000
$9.7 \cdot 10^9$		-109.325905	

Table 3.3: Convergence of the SFCI algorithm towards the exact energy of the N_2 molecule as the sparse vector is allowed to lengthen. Number of determinants in the FCI space and the exact FCI energy computed by Evangelisti et al.[93] is indicated in a separate row at the bottom. Notation t_{SFCI} refers to the wall-clock time of an average iteration step. All calculations shown here converged within 20 iterations.

disk manipulations, there were 12 I/O operations in each iteration step in the calculation reported by Evangelisti[99], and each I/O operation treated double precision quantities amounting to 40 Gbyte. Altogether about 500 Gbyte information was moved between the memory and the hard disk in an iteration. Compared to this, there are a few times more I/O operations in one iteration step in our case, but one I/O treats merely 2.3 Gbyte in the largest case presented.

3.6 Summary

In this Chapter a sparse matrix based CI algorithm was introduced. The speed of a FCI algorithm is determined by the performance of the linear transformation using the Hamiltonian operator. The feasibility of a FCI computation is determined by the length and number of CI vectors to be stored. In the algorithm presented, both the initial and the product vector is kept sparse, when computing the action of Hamiltonian. In other words not only the approximation of the wave function is kept sparse but also the result of the

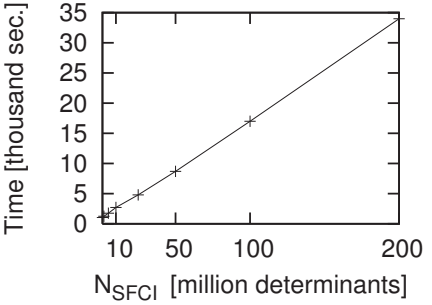


Figure 3.3: Computational time of an average iteration step in the case of N_2 molecule as a function of SFCI population.

linear transformations are calculated to a small subspace of the FCI space. In those cases where the FCI vectors are sparse, these properties considerably reduce the computation time without introducing significant error. To restrict the subspace where the product vector of the linear transformation lies, we have estimated the importance of the occupation strings at each iteration step in a perturbative manner.

The illustrative calculations presented above already reflect that the SFCI algorithm represents a practical tool to obtain approximate correlation energies of molecular systems in their ground state. An advantageous feature of the algorithm is the externally tunable amount of the calculation (i.e. memory requirement). This can be exploited either by adapting the calculation to the accuracy needed in the problem under consideration, or by fitting the calculation to the memory limits determined by computational facilities. In the case where there is a prescribed accuracy for the calculation to reach, it may be of use to stop further refinement of the SFCI space once the energy improvements get below threshold, and quickly finish the calculation by an ordinary Davidson algorithm iterated till convergence (a few steps) in the already fixed SFCI subspace. Parallel implementation of the algorithm, to provide flexible fitting of the calculation to computer clusters as well, is under development.

To assess the error of an SFCI calculation in the lack of an exact FCI result, estimations based on previous experience can be useful. A more rigorous way to attach an approximate error bar is to perform a SFCI calculation with increased population. Alter-

natively a second-order perturbation theory estimation can be computed to obtain an error estimation[105].

Regarding excited states of molecular systems, the SFCI calculation can safely reach the bottom level in each irreducible representation. More experience is needed to investigate whether excited levels of any of the i.r. can be obtained by starting from a well chosen initial guess function. Excited state calculation following the ordinary Davidson strategy is not a viable route for the SFCI procedure, since Davidson steps, which are part of SFCI iteration steps are independent and work with a 2 by 2 matrix at most. Applying orthogonality condition to an already obtained ground state SFCI function in order to get an excited state, involves yet unexplored risks due to the truncated nature of the ground state SFCI vector.

Acknowledgments

I am very grateful to my supervisor, Péter Surján for all his support. I wish to express my special thanks to Ágnes Szabados for the effort she made to help me to complete this work. I would like to thank Debashis Mukherjee for the fruitful discussions. I also want to thank my boss, Mihály Kállay for his help during the last year. I would like to thank all the past and present members of the Péter Surján's Research Family (PSRF) for the friendly atmosphere. I wish to thank Ilona Halupka for her support and understanding during the weekdays. Finally I would like to thank my mother and my brother for all their efforts and the background.

This work was partly supported by the Deák Ferenc Scholarship of the Ministry of Education and Culture(0032/2009).

Bibliography

[1] Z. Rolik, Á. Szabados, and P. R. Surján. *J. Chem. Phys.*, 119:1922, 2003.

[2] P.R. Surján, Z. Rolik, Á. Szabados, and D. Kőhalmi. *Ann. Phys. (Leipzig)*, 13:223–231, 2004.

[3] Á Szabados, G. Tõth Z. Rolik, and P. R. Surján. *J. Chem. Phys.*, 122:114104, 2005.

[4] Z. Rolik and Á. Szabados. *Int. J. Quantum Chem.*, 109:2554, 2009.

[5] Á. Szabados Z. Rolik and P. R. Surján. *In preparation*.

[6] Z. Rolik, Á. Szabados, and P. R. Surján. *J. Chem. Phys.*, 128:144101, 2008.

[7] R. J. Bartlett. *Annu. Rev. Phys. Chem.*, 32:359, 1981.

[8] B. H. Brandow. *Rev.Mod.Phys*, 39:771, 1967.

[9] J.-P. Malrieu, P. Durand, and J. P. Daudey. *J. Phys. A*, 18:809, 1985.

[10] M.G. Sheppard and K. Freed. *J. Chem. Phys.*, 75:4507, 1981.

[11] K. Wolinski and P. Pulay. *J. Chem. Phys.*, 90:3647, 1989.

[12] K. Andersson, P.-Å. Malmqvist, and B. O. Roos. *J. Chem. Phys.*, 96:1218, 1992.

[13] K. Hirao. *Chem. Phys. Letters*, 190:374, 1992.

[14] R. B. Murphy and R. P. Messmer. *J. Chem. Phys.*, 97:4170, 1992.

[15] P. M. Kozlowski and E. R. Davidson. *Chem. Phys. Letters*, 222:615–620, 1994.

[16] H-J. Werner. *Mol. Phys*, 89:645–661, 1996.

[17] K.G. Dyall. *J. Chem. Phys.*, 102:4909, 1995.

[18] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu. *J. Chem. Phys.*, 114:10252–10264, 2001.

[19] A. Zaitevskii and J-P. Malrieu. *Chem. Phys. Letters*, 233:597, 1995.

[20] J. P. Finley. *J. Chem. Phys.*, 108:1081, 1998.

[21] E. Rosta and P.R. Surján. *J. Chem. Phys.*, 116:878–890, 2002.

[22] J.A. Pople, J.S. Binkley, and R. Seeger. *Int. J. Quantum Chem.*, s10:1, 1976.

[23] R. J. Bartlett and G. D. Purvis. *Int. J. Quantum Chem.*, 14:561, 1978.

[24] J. M. Rintelman, I. Adamovic, S. Varganov, and M. S. Gordon. *J. Chem. Phys.*, 122:44105, 2005.

[25] K.A. Brueckner. *Phys. Rev.*, 97:1353, 1955.

[26] K.A. Brueckner. *Phys. Rev.*, 100:36, 1955.

[27] J. Goldstone. *Proc. R. Soc. A*, 239:267, 1957.

[28] N. H. March, W. H. Young, and S. Sampanthar. *The Many-Body Problem in Quantum Mechanics*, page 72. Cambridge University Press, London, 1967.

[29] V. Kvasnička. *Adv. Chem. Phys.*, 36:345, 1977.

[30] I. Lindgren. *J. Phys. B*, 7:2241, 1974.

[31] G. Hose and U. Kaldor. *J. Phys. B*, 12:3827, 1979.

[32] L. Meissner and R. J. Bartlett. *J. Chem. Phys.*, 91:4800, 1989.

[33] C. Møller and M.S. Plesset. *Phys. Rev.*, 46:618, 1934.

[34] L. Meissner and K. Jankowski. *Int. J. Quantum Chem.*, 36:705, 1989.

[35] D. Pahari, S. Chattopadhyay, S. Das, and D. Mukherjee. *Chem. Phys. Letters*, 381:223, 2003.

[36] H. Nakano. *J. Chem. Phys.*, 99:7983, 1993.

[37] J. P. Finley. *J. Chem. Phys.*, 108:1081, 1998.

[38] K. Hirao. *Chem. Phys. Letters*, 201:59, 1993.

[39] P. Pulay and S. Saebø. *Theor. Chim. Acta*, 69:357, 1986.

[40] S. Saebø and P. Pulay. *J. Chem. Phys.*, 86:914, 1987.

[41] E. R. Davidson and C. F. Bender. *J. Chem. Phys.*, 56:4334, 1972.

[42] E. R. Davidson. *J. Chem. Phys.*, 57:1999, 1972.

[43] I. Mayer. *Simple Theorems, Proofs, and Derivations in Quantum Chemistry*, page . Kluwer, New York, 2003.

[44] R. B. Murphy and R. P. Messmer. *Chem. Phys. Letters*, 183:443, 1991.

[45] F. W. Bobrowicz and W. A. Goddard-III. *Methods of Electronic Structure Theory*, page 79. Plenum, New York, 1977.

[46] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski. *J. Phys. Chem.*, 94:5483, 1990.

[47] K. Andersson, P.-Å. Malmqvist, and B.O. Roos. *J. Chem. Phys.*, 96:1218, 1992.

[48] H. J. J. van Dam, J. H. van Lenthe, and P. Pulay. *Mol. Phys.*, 93:431, 1998.

[49] H.J.J. van Dam, J.H. van Lenthe, and P.J.A. Ruttnik. *Int. J. Quantum Chem.*, 72:549–558, 1999.

[50] P.R. Surján and Á. Szabados. in: *Fundamental World of Quantum Chemistry: A Tribute to the Memory of Per-Olov Löwdin*, page . Kluwer, New York, 2005.

[51] E. Kapuy, F. Bartha, F. Bogár, and C. Kozmutza. *Theor. Chim. Acta*, 72:337, 1987.

[52] E. Kapuy, F. Bartha, C. Kozmutza, and F. Bogár. *J. Mol. Struct.* (THEOCHEM), 170:59, 1988.

[53] E. Kapuy, F. Bartha, F. Bogár, Z. Csépes, and C. Kozmutza. *Int. J. Quantum Chem.*, 37:139, 1990.

[54] K. Hirao. *Int. J. Quantum Chem.*, S26:517, 1992.

[55] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry*. McGraw-Hill, New York, 1989.

[56] P. R. Surján. *Topics in Current Chemistry*, 203:63, 1999.

[57] T. Helgaker, P. Jørgensen, and J. Olsen. *Molecular Electronic-Structure Theory*. John Wiley & Sons Ltd, England, 2000.

[58] Á. Szabados and P.R. Surján. *Advances in the Theory of Atomic and Molecular Systems: Conceptual and Computational Advances in Quantum Chemistry, Progress in Theoretical Chemistry and Physics*, 19:257, 2009.

[59] J. Paldus and J. Čížek. *Adv. Quantum Chem.*, 9:105–197, 1975.

[60] S. Kucharski and R. Bartlett. *Advances in Quantum Chemistry*, 18:281, 1986.

[61] P. G. H. Sandars. *Adv. Chem. Phys.*, 14:365, 1969.

[62] L. H. Frantz and R. Mills. *Nucl. Phys.*, 15:16, 1960.

[63] X. Li and J. Paldus. *J. Chem. Phys.*, 101:8812, 1994.

[64] S. Hirata. *J. Phys. Chem.*, 107:9887, 2003.

[65] M. Kállay and P. R. Surján. *J. Chem. Phys.*, 115:2945, 2001.

[66] I. Mayer. *Simple Theorems, Proofs, and Derivations in Quantum Chemistry*. Kluwer, New York, 2003.

[67] M. Kállay, P.G. Szalay, and P.R. Surján. *J. Chem. Phys.*, 117:980–989, 2002.

[68] P. G. Szalay and R. J. Bartlett. *Chem. Phys. Letters*, 214:481, 1993.

[69] L. Füsti-Molnár and P.G. Szalay. *J. Phys. Chem.*, 100:6288–6297, 1996.

[70] N.B. Amor and D. Maynau. *Chem. Phys. Letters*, 286:211–220, 1998.

[71] U.S. Mahapatra, B. Datta, and D. Mukherjee. *J. Chem. Phys.*, 110:6171–6188, 1999.

[72] G.D. Purvis, R. Shepard, F.B. Brown, and R.J. Bartlett. *Int. J. Quantum Chem.*, 23:835–845, 1983.

[73] T. H. Dunning Jr. *J. Chem. Phys.*, 53:2829, 1970.

[74] P. R. Surján and Á. Szabados. *Int. J. Quantum Chem.*, 69:713, 1998.

[75] P. R. Surján, D. Kőhalmi, and Á. Szabados. *Collect. Czech. Chem. Commun.*, 68:331–339, 2003.

[76] N. Forsberg and P.-Å. Malmqvist. *Chem. Phys. Letters*, 274:196, 1997.

[77] J. P. Finley and H. A. Witek. *J. Chem. Phys.*, 112:3958, 2000.

[78] A. Zaitevskii and J-P. Malrieu. *Theor. Chim. Acta*, 96:269–276, 1997.

[79] A. Zaitevskii and J-P. Malrieu. *Chem. Phys. Letters*, 250:366, 1996.

[80] A. Zaitevskii and R. Cimiraglia. *Int. J. Quantum Chem.*, 73:395, 1999.

[81] A. Izmailov and A. Zaitevskii. *Int. J. Quantum Chem.*, 96:202, 2004.

[82] K. F. Freed. *Lecture Notes in Chemistry*, volume 52, page 1. Springer, Berlin, 1989.

[83] S. Kucharski and R. Bartlett. *Int. J. Quantum Chem.*, S 22:383, 1988.

[84] I. Lindgren and J. Morrison. *Atomic Many-Body Theory*. Springer, Berlin, 1986.

[85] H. Larsen, J. Olsen, P. Jørgensen, and O. Christiansen. *J. Chem. Phys.*, 113:6677, 2000.

[86] J. Olsen, B. O. Roos, P. Jørgensen, and H. J. Aa. Jensen. *J. Chem. Phys.*, 89:2185, 1988.

[87] A. Izmailov and A. Zaitsevskii. *Int. J. Quantum Chem.*, 96:202, 2004.

[88] B.O. Roos and P.E.M. Siegbahn. In H.F. Schaefer, editor, in: *Modern Theoretical Chemistry*, volume 3, chapter 7. Plenum, New York, 1977.

[89] M. Moshinsky. *Group Theory and the Many-Body Problem*. Gordon and Breach, New York, 1968.

[90] J. Paldus. In H. Eyring and D.J. Henderson, editors, in: *Theoretical Chemistry: Advances and Perspectives*, volume 2, page 131. Academic, New York, 1976.

[91] I. Shavitt. *Chem. Phys. Letters*, 63:421, 1979.

[92] N. C. Handy. *Chem. Phys. Letters*, 74:280, 1980.

[93] E. Rossi, G.L. Bendazzoli, S. Evangelisti, and D. Maynau. *Chem. Phys. Letters*, 1310:530, 1999.

[94] L. Thøgersen and J. Olsen. *Chem. Phys. Letters*, 393:36, 2004.

[95] Z. Gan, D. J. Grant, R.J. Harrison, and D. A. Dixon. *J. Chem. Phys.*, 125:124311, 2006.

[96] Z. Gan and R. J. Harrison. *Proceedings of the 2005 ACM/IEEE SC—05 Conference (SC'05)*, 2005.

[97] G. L. Bendazzoli and S. Evangelisti. *J. Chem. Phys.*, 98:3141, 1993.

[98] E. Rossi, G. L. Bendazzoli, and S. Evangelisti. *J. Comput. Chem.*, 19:658, 1998.

[99] R. Ansaloni, G. L. Bendazzoli, S. Evangelisti, and E. Rossi. *Computer Physics Communications*, 128:496, 2000.

[100] Z. Gan, Y. Alexeev, M. S. Gordon, and R. A. Kendall. *J. Chem. Phys.*, 119:47, 2003.

[101] M. Klene, M. A. Robb, M.J. Frisch, and P. Celani. *J. Chem. Phys.*, 113:5653, 2000.

[102] P.J. Knowles. *Chem. Phys. Letters*, 155:513, 1989.

[103] P.J. Knowles and N.C. Handy. *J. Chem. Phys.*, 91:2396, 1989.

[104] A.O. Mitrushenkov. *Chem. Phys. Letters*, 217:559, 1994.

[105] Yuri Yu. Dmitriev A.O. Mitrushenkov. *Chem. Phys. Letters*, 235:410, 1995.

[106] J. Olsen, P. Jørgensen, and J. Simons. *Chem. Phys. Letters*, 169:463, 1990.

[107] S. Zarrabian, C.R. Sarma, and J. Paldus. *Chem. Phys. Letters*, 155:183, 1988.

[108] A. Povill and J. Rubio. *Theor. Chim. Acta*, 92:305, 1995.

[109] E. R. Davidson. *J. Comp. Phys.*, 17:87, 1975.

[110] P. J. Knowles and N. C. Handy. *Comp. Phys. Commun.*, 54:75, 1989.

[111] S. Pissanetzky. *Sparse Matrix Technology*. Academic Press, London, 1984.

[112] P.-O. Widmark, P.-Å. Malmqvist, and B.O. Roos. *Theor. Chim. Acta*, 77:291, 1990.