NICHOLAS COPERNICUS UNIVERSITY
IN TORUŃ
Faculty of Physics and Astronomy
Department of Computer Methods

Jarosław Meller

New computational algorithms based on the Configuration Interaction method

Doctoral Thesis written under guidance of prof. dr hab. W. Duch
at the Department of Computer Methods.

Toruń, February 1996

Acknowledgments

When I started this work, encouraged by prof. Włodzisław Duch, I was not aware of what the most important thing in the graduate period is. After four years I may state, without any doubts, it is a chance to crystallize one's own personality and view of the human beings efforts to understand the nature and themselves.

Such a unique chance comes from the special character of the scientific community a young graduate student enters. I had also a chance to take benefit of this special situation meeting on my way many wonderful people, excellent teachers and scientists.

During that time my debts have especially accumulated to the supervisor of this thesis, prof. Włodzisław Duch. I would like to thank him for inspiration, his invaluable help and patience for hundreds of my silly questions. He has always had time for me, despite his extremely rich activities, introducing me step by step to the world of Quantum Chemistry and to the world of Science. I would also like to stress that he not only directed my work but in fact we have done the work on Sup-CI method together.

The above remark applies as well to the second part of my thesis, devoted to a dressed CI formulation of various CC type approximations. This work has been done together with prof. Jean-Paul Malrieu during my three visits to Toulouse and further cooperation. The thanks that I offer him here is an adequate measure of what his filling me with eagerness and constant care of my progress have meant to the final result of my imperfect efforts. My stays were financed by the Tempus project, french government and Polish Committee for Science (KBN), which are gratefully acknowledged.

I would be remiss not to thank professors K. Jankowski, J. Karwowski and J. Wasilewski. Attending their seminars and asking frequently, I have gained much of my knowledge of Quantum Chemistry. All my friends and colleagues at the Department of Computer Methods and Institute of Physics of NCU in Toruń and at the Laboratoire de Physique Quantique of UPS in Toulouse, have also my gratitude for all discussions and their support.

Contents

Part I

Preliminaries

Preface

"Concepts, like men, are gregarious"
Ernest Gellner, "Plough, Sword and Book. The Structure of Human History."

Although, E. Gellner refers to a culture, which he defines as a system of notions, interrelated and interdependent in various complex ways, I find his metaphor matching very well the content and background of this work. It is devoted to the development of new methods of the Quantum Chemistry or Theoretical Atomic and Molecular Physics, if one prefers. Not depending on the actual (and historically determined) nomenclature, it means theoretical methods of research on the electronic structure of the matter.

After 70 years (starting from the very beginning of the Quantum Mechanics) of the development of this discipline and 50 years since the invention of the computer - its nowadays inevitable companion - it has reached high degree of sophistication and impressive predictive capabilities. Due to 'black box' quantum chemistry packages, one can routinely describe the structure of chemical compounds, reaction paths and spectra, and many other interesting properties and characteristics [1] – [3].

Nevertheless, despite the parallel progress in computer technology, further formal development of methods and increasing of effectiveness of the computer algorithms of Quantum Chemistry is highly desired. We shall always be pushed to attack larger and larger systems with growing accuracy.

It seems that the project of linear Quantum Chemistry (i.e. using algorithms for which the computational complexity scales linearly with the number of electrons) remains still quite exotic. The problem of bad scaling has been however constantly inducing efforts resulting in various combinations in the spirit of the so-called direct algorithms and other numerically efficient schemes.

Another area of vital importance for contemporary Quantum Chemistry is further development of existing and well established methods for special and difficult cases as for example open shell systems and excited states.

This work has been inspired by both challenges and tries to offer some new solutions to these problems. They are formulated on the grounds of well known and standard Configuration Interaction method, that had been devised and employed to account for the electron correlation i.e. to go beyond the one particle approximation. On the other hand the methods presented in this thesis are far from the methodological purity and are based on effective combination of concepts, coming from different approaches to the electron correlation problem. Certainly all these approaches are interrelated and this fact has been used many times to introduce new schemes and algorithms.

The Superdirect Configuration Interaction (Sup-CI) method, proposed independently by Bendazzolli [42] and in a more general form by W. Duch [39] - the supervisor of this thesis - uses Perturbation Theory to built effective short linear expansion of the electronic wave function. This idea has been previously used with great success in various diagonalization algorithms for instance. The Sup-CI method in third order is a matter of the first part of this work. It is tested on a series of benchmark problems and its efficient implementation is discussed. The formulas for matrix elements occurring in this approximation are derived.

The 'dressing' of the set of linear equations of the Configuration Interaction (CI) method, which allows to incorporate non-linear terms of the Coupled Cluster (CC) method, has been used in the first implementations of the Coupled Electron Pair Approximation method. Recently, it has been rediscovered and generalized by J.P. Malrieu and his collaborators [86], giving theoretical and practical bridge between CI and CC methods. It appeared to be a generator of a series of efficient algorithms of the Coupled Cluster type, formulated in terms of Configuration Interaction method. Some of them are presented in the second part of the thesis, with the obvious accent on those, which the author of this work has contributed to. Because one is facing a collection of (gregarious) ideas in this case, I shall leave further comments to the subsequent parts of the work.

Most of the presented schemes have been formulated in terms of multiconfigurational expansions and are directed to the description of systems having different electronic structures at different conformations as for instance dissociating molecules, of the open shell systems and excited states. Multiconfigurational procedures are certainly more complex than their single reference counterparts, but real chemical systems and processes inevitably require some measurements of a non-dynamical correlation. All those general issues will be discussed in a more detailed manner in the first chapter.

Coming back to the Gellner's metaphor, I would like to point out that it may be interpreted in two - contradictory to some extent - ways. One may regard that it expresses a natural and fruitful tendency in any systematic search for the logically consistent knowledge, especially in science. This tendency manifests itself by two rules of scientific investigation. First - do not try to go too far as long as it is possible. Second - combining known things you may always find something new. Its deeper meaning suggests however that only straying one may really make a progress.

This is consistent with the character of a scientific revolution or a change of paradigm, as described by Kuhn. Men are conformistic. They usually choose what has already been chosen. They tend to exclude all different individuals from their communities. It may however happen that a deep crisis may force them to accept previously rejected options or to invent new solutions. The same concepts do.

Looking at this work from a wider perspective of the theoretical physics, one has to admit that it has nothing to do with any fundamental question that might change our view of Quantum Mechanics. The aim of this thesis is to widen the range of standard computational options enabling chemists to predict behavior of real chemical systems.

The Quantum Paradigm still seems not to be seriously threatened and we seem to remain in the relatively calm, intermediate period. I am (gregarious) man and I may only hope that my modest contribution to the development of Quantum Chemistry will not cause a butterfly effect.

Seriously speaking, my hope is that at least some of the algorithms I have contributed to, will find their way to the world of real applications in chemistry and molecular physics or will inspire some further formal developments resulting in more mature methods. It seems to me that combining traditional approaches may not only serve in increasing of the efficiency of calculations, but it may also bring a new light on the mutual relationships among those deeply interrelated approaches.

Chapter 1

Introduction

Some basic aspects of the description of electronic states of chemical molecules and methodological background of the developments presented here shall be considered in the following chapter.

1.1 Roots

In Quantum Mechanics any microscopical system may be characterized by the mathematical quantity called a wave function (usually denoted as $\Psi(\mathbf{r}, t)$). The wave function has probabilistic interpretation: $\Psi^*(\mathbf{r}, t)\Psi(\mathbf{r}, t)$ means the probability density of finding objects, a given system consists of, at points given by $\mathbf{r}$ and at a moment t. It fulfills time dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t} = \hat{H}\Psi(\mathbf{r},t) = \left[\hat{T} + V(\mathbf{r},t)\right]\Psi(\mathbf{r},t) \tag{1.1}$$

A differential operator $\hat{H}$ is called Hamiltonian and its particular form is depended on the system considered. We shall consider only non-relativistic Hamiltonians.

When the potential V is not depended on time, one can separate variables, obtaining time independent Schrödinger equation

$$\hat{H}\Psi(\mathbf{r}) = \left[\hat{T} + V(\mathbf{r})\right]\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{1.2}$$

Square integrible (i.e. localized in space) solutions of this equation define a Hilbert space $\mathcal{H}$. The bounded states of a system are represented by functions belonging to $\mathcal{H}$ and thus may be expanded in terms of eigenfunctions of $\hat{H}$, that form a basis of the space $\mathcal{H}$. A bounded state, which is a superposition of eigenfunctions having the same eigenvalue E, is called a *stationary state*. A set of eigenvalues E_n associated with all stationary states is in turn called an *energetic spectrum* of a system [4].

The goal of Quantum Chemistry is to solve the Schrödinger equation (1.2) for a given chemical system. More precisely, one of the main objectives of Quantum Chemistry is to find (not necessarily all) stationary states of electron cloud in the potential associated with the interactions between nuclei and electrons, a given molecule consists of. Speaking in general terms, the methods presented in this thesis are aiming at that goal as well.

The study of time dependent phenomena (e.g. arising in the interaction between time-dependent external field and the electronic system) in general require solving the time dependent Schrödinger equation, but this may often be done in the basis of time independent eigenstates of eq. (1.2) [5].

There are only a few known analytic solutions of the Schrödinger equation for such simple systems as hydrogen atom. Therefore various approximations are inevitable. They not only allow us to solve the Schrödinger equation but also play a fundamental role in understanding (modeling) of molecular structure.

Since a large difference in mass between electrons and nuclei it is well grounded to postulate the separation of nuclei and electrons motions [5]. This assumption is known as Born-Oppenheimer approximation (or adiabatic approximation if one includes some additional terms that do not couple nuclei and electronic states [5]) and leads to great practical and conceptual simplifications. It allows to introduce the electronic Hamiltonian $\hat{H}_e$

$$\hat{H}_e(\mathbf{r}_e, \mathbf{R}) = \hat{T}(\mathbf{r}_e) + V(\mathbf{r}_e, \mathbf{R}) \tag{1.3}$$

and electronic wave functions that depend on the (fixed at a given conformation) positions of nuclei $\mathbf{R}$ as parameters only. In the above equation $\mathbf{r}_e$ denotes a vector of electronic coordinates. Since the potential $V(\mathbf{r}_e, \mathbf{R})$ consists of one-electron, separable terms and two-electron non-separable terms

$$V(\mathbf{r}_e, \mathbf{R}) = \sum_k (\hat{t}(\mathbf{r}_k) + \sum_A \frac{Z_A}{r_{kA}(\mathbf{R})}) + \sum_{k<l} \frac{1}{r_{kl}} = \hat{h}_1 + \hat{h}_2 \tag{1.4}$$

one may define the one- and two-particle parts (denoted as $\hat{h}_1$ and $\hat{h}_2$ respectively) in $\hat{H}_e$ [6]. The summation over A refers to the atomic centers and atomic units are used in the eq. (1.4).

As a consequence of the separation of the electronic and nuclei states the electronic energy (eigenvalue associated with electronic wave function) plays the role of a potential energy in the motion of nuclei. This in turn allows to introduce the concept of the Potential Energy Surface (PES) [5]. Therefore the electronic structure is not only important for the description of the electronic phenomena (as for instance excitation energies or transition probabilities) but it serves also as a starting point for further investigation of vibrational and rotational phenomena.

Another model central to contemporary Quantum Chemistry is the one-particle approximation. Within this model every electron is associated with a one-particle function, called an *electron spin orbital* ψ_i. One may look for one-particle functions that define the energetically lowest many-particle wave function, which is defined as antisymmetrized product (determinant) of one-particle functions. It brings substantial reduction of the complexity of the problem and leads to a set of well known integro-differential Hartree-Fock equations for one-electron problems (a single electron interacting with an averaged field of all electrons)

$$\hat{f}\psi_i = e_i\psi_i; \qquad \hat{f} = \hat{h} + \sum_j (\hat{J}_j - \hat{K}_j) \tag{1.5}$$

where e_i is called one-particle energy and the Fock operator $\hat{f}$ is a sum of one-particle Hamiltonian and Coulomb and exchange operators respectively. For N electrons, from a set of spin orbitals with N lowest one-electron energies one can built the best single determinant approximation to the ground state of the N-electron system described by an electronic Hamiltonian $\hat{H}_e$ [6].

In case of atoms and linear molecules the above equations may be directly solved by numerical methods [12]. Because the operator $\hat{f}$ is defined by spin orbitals ψ_i, one has to solve eqs. (1.5) iteratively, in a self-consistent way. Unfortunately, this procedure is not as yet applicable to molecules consisting of a larger number of atoms [12] and further simplifications are required.

1.2 Molecular orbital approximation

The N-particle Hilbert space $\mathcal{H}^N$ has very nice and simple structure in the one-particle approximation. It is built of the one-particle function products i.e. it has the form of a tensor product

$$\mathcal{H}^N = \otimes^N (\mathcal{H}^1) \tag{1.6}$$

where $\mathcal{H}^1$ denotes the one-particle Hilbert space of all localized one-particle functions ($\dim(\mathcal{H}^1) = \infty$). When dealing with fermions, only the antisymmetric subspace ${}^A\mathcal{H}^N$ has to be taken into account (we shall skip further the A index). Solving the Hartree-Fock equations in this space one obtains the so-called *Hartree-Fock limit* for the energy of N-particle system [6].

As was pointed out in the previous section, except for the simplest linear molecules, algebraization of the problem is necessary to obtain solutions of the Hartree-Fock equations in case of chemical systems. It was Roothan, who introduced the expansion of molecular orbitals in the finite basis of atomic orbitals [10]. Therefore we say about Hartree-Fock-Roothan equations or LCAO MO (Linear Combination of Atomic Orbitals, Molecular Orbitals) method.

From the formal point of view it means that one-particle Hilbert space $\mathcal{H}^1$ has finite dimension n, where n is the dimension of the basis set used and certainly the N-particle Hilbert space is also of finite dimension. As the one-particle basis approaches completeness ($n \to \infty$) the method approaches Hartree-Fock limit. The acronym SCF (Self-Consistent Field) method is often used to distinguish the Hartree-Fock method in the finite basis set.

It is common to introduce a set of $k = n/2$ spatial basis functions, *electron orbitals*. The spatial parts of the spin orbitals with the α (and β respectively) spin function are expanded in terms of these atomic orbitals, giving finally n molecular spin orbitals. Assuming the same spatial parts for α and β spins, we get the *restricted* HF (RHF) method. Relaxing this constraint one obtains *unrestricted* HF (UHF) method [6].

In practice, one has to restrict the dimension of the basis set severely (to about 100 in routine calculations). Quality of the basis is of great importance. The art of constructing of such basis for molecular *ab initio* calculations has been developing since the early days of Quantum Chemistry. The most common choice is to expand molecular orbitals as linear combinations of atomic gaussian functions (i.e. with exponents and coefficients of primitive functions optimized in atomic calculations), leading to simplifications in evaluation of multicentered integrals [6].

In contradistinction to *ab initio* methods the so-called *semi-empirical* methods mean procedures that use some parameters fitted to known empirical data and not only basic physical constants [14]. The use of atomic functions with optimized parameters for the construction of one-particle basis functions makes the distinction between *ab initio* and semi-empirical methods actually only traditional. Nevertheless, the semi-empirical methods usually assume a simplified form of a Hamiltonian, as the Hückel one, and are in this sense less refined [7].

The Hartree-Fock approximation is important not only for its own sake but as a starting point to the so-called post Hartree-Fock methods that try to incorporate the instantaneous interactions among the electrons - to include the electron correlation, as quantum chemists say.

1.3 Electron correlation problem

Although the Hartree-Fock (or molecular orbital) approximation has proved its high predictive capabilities, it is very often inadequate for description of chemical systems and processes because they depend on energy differences (which are relatively small) rather than on energies themselves.

In extreme cases like that of the F_2 molecule, the HF method is unable to describe the bonding [8]. The problem cannot be removed even by using the numerical solutions of the HF equations (HF limit). Single determinantal description, even in the infinite Hilbert space, is not adequate and one has to go beyond the one-particle approximation.

When each electron is assumed to move in the field of nuclei and the average field of all other electrons, one should expect an error coming from the discarding of the instantaneous effects in the Coulomb interactions, arising from the interelectronic potential $e_1 e_2 / r_{12}$.

The electronic *correlation energy* is essentially a measure of the error of the HF method. It is defined by the difference

$$E_{corr} = E - E_{HF} \tag{1.7}$$

where E is the exact eigenvalue of the Hamiltonian $\hat{H}$ under consideration [9]. Because the Hamiltonian employed normally in Quantum Chemistry does not contain the relativistic effects, which are always present in the experimental energy, the correlation energy has no physical meaning. It only indicates how good (or bad) the HF approximation is.

It is worth to note that working in a given basis set ($\mathcal{H}^1$) of finite dimension we may only talk about HF energy E_{SCF} and the 'exact' energy E in the resulting N-particle space $\mathcal{H}^N$. Despite that, the correlation energy (in a given basis) remains very useful for the evaluation of the quality of the post Hartree-Fock methods.

When dealing with finite dimensional basis sets another problem arises, connected to the Coulomb correlation and leading to a slow convergence of the post HF methods. The restricted flexibility of such an approach causes errors even when going beyond the one-particle approximation. Taking into account all determinants in N-particle space $\mathcal{H}^N$ one is not able to describe properly such special effects as the Coulomb hole, arising for $r_{12} \rightarrow 0$. The solution for this problem requires methods that explicitly take into account the interelectron distances r_{12} and face serious computational difficulties [11].

One should not forget that, in practice, the correlation energy is a formal parameter rather than a measure of physical correlation in the system. The latter one is partially included in the HF method and partially is not included even in E of eq. (1.7), when employing finite dimensional approximations.

In the next section we shall briefly consider different post HF approaches. An important remark should be made before we go further. All the considerations presented here refer to *molecular ab initio* methods and not to *density* methods. Various methods oriented at electron density instead of the wave functions undertake recently an intensive development. They are commonly referred to as the Density Functional Theory (DFT) and offer an interesting alternative to the traditional molecular approach (at least for the lowest states in a given symmetry) [13].

1.4 Going beyond the Hartree-Fock method

A given choice of n atomic basis functions for an N-electron system implies that one obtains in the SCF procedure n (usually orthonormal for the sake of computational efficiency) molecular spin orbitals. They form a basis of the finite dimensional space $\mathcal{H}^1$. All their antisymmetrized products are eigenfunctions of N-particle Fock operator $\hat{F} = \sum_i^N \hat{f}_i$ and form a basis of N-particle Hilbert space $\mathcal{H}^N$. The number of all such products is given by $n!/N!(n-N)!$ and may be very large even for the systems of a moderate size. For 28 basis functions and 10 electrons (of the water molecule for instance) it reaches 10^7. Although this number is significantly decreased (by about one order of magnitude) when excluding determinants characterized by unwanted total spin projection and further reduction is possible by the full spin and space symmetry adaptation, it shows the potential computational difficulties.

Those spin orbitals that occur in the energetically lowest, so-called Hartree-Fock determinant Ψ_{HF}, are defined as *occupied* spin orbitals. The rest is called *virtual* spin orbitals. Since one can describe each determinant with respect to the HF one (by the number of occupied spin orbitals replaced by virtual spin orbitals), we may define single (S), double (D), triple (T), quadruple (Q) ... N-tuple substitutions or excitations

$$\mathcal{H}^N = \{|\Psi_{HF}\rangle\} \oplus \mathcal{H}_S^N \oplus \mathcal{H}_D^N \oplus \cdots \oplus \mathcal{H}_N^N \tag{1.8}$$

From now we shall use more convenient language of the *state vectors* $|\Psi_{HF}\rangle$ in the abstract Hilbert space rather than the explicit coordinate representation $\Psi_{HF}(\mathbf{r}) = \langle \mathbf{r}|\Psi_{HF}\rangle$. The partitioning of (1.8) is a useful tool to avoid problems associated with huge dimensions of $\mathcal{H}^N$ by defining some truncated schemes. When employing orbitals rather than spin orbitals and spin adapted basis rather than determinantal one, it is necessary to use other partitioning, based for instance on the concept of the interaction space, since the excitation level may not be uniquely defined [16].

Going beyond the Hartree-Fock method means that we take into account not only the ground HF state. A given state is expressed as a linear combination of a number of N-electron functions. Among those state functions (being in case of symmetry and spin adaptation combinations of determinants) one can distinguish those that have significantly large coefficients (comparable to the coefficient of the ground state determinant) in the expansion of a given state in terms of N-particle basis and those that have relatively small coefficients or weights in the total wave function. Such a distinction is never sharp nor unique. Nevertheless, some intuitive heuristics usually allow us to justify a given choice.

In general *non-dynamical* or structural correlation effects and *dynamical* correlation effects are distinguished [16], manifested by relatively large or small coefficients respectively. A more strict distinction follows the definition of Sinanoglu [15] which says that the term non-dynamical correlation refers to a minimal qualitatively correct description of the separation of a molecule into fragments. Thus, the non-dynamical correlation arises in a situation, when some configurations are of different importance at various molecular conformations and all of them have to be included in the zeroth order treatment to obtain correct description of the PES. It also may happen that even at equilibrium the single determinant treatment is entirely not adequate (as in the F_2 case mentioned before) and a multireference treatment, accounting for the non-dynamical correlation, is necessary. Some excited states are also inherently of the multireference character.

The above situation implies that all methods used in Quantum Chemistry, primarily devised for single reference zeroth order description, has now its multireference counterpart. It concerns also the SCF method itself, which has its generalized version: Multiconfigurational SCF (MCSCF) method. In the MCSCF scheme the orbitals are variationally optimized not for single determinant, but for linear combination of determinants. The choice of the reference space is often based on the concept of the Complete Active Space (CAS) i.e. a space composed of all possible excitations within a given set of (chemically active) orbitals. It is equivalent to the FCI treatment (see subsection 1.4.2) in this small subspace of $\mathcal{H}^N$.

One should recall another and related distinction, important from the conceptual and practical point of view. A *closed shell* state of N-electron system is a state described by a single electronic configuration, consisting of completely occupied shells. Any other state is called an open shell state. This includes cases with a single modeling configuration but involving partially occupied shells and states that cannot be described in a qualitatively correct way using single reference approach [16].

Once the non-dynamical part of the correlation energy is taken care of by small MCSCF or CASSCF calculations the post Hartree-Fock methods account for the dynamical part of it. This is computationally more demanding. Moreover, as mentioned before, the problem with quantum chemical methods is that there is no clear distinction between the dynamical and the non-dynamical parts of the correlation energy in different molecular conformations. In the next three subsections we shall restrict to the single reference approaches unless otherwise stated.

1.4.1 Perturbation Theory

Perturbation Theory prevails in many branches of physics. It has been also used as one of the basic tools of the molecular quantum mechanics.

On the grounds of molecular orbital approximation, a special role is played by the Rayleigh-Schrödinger expansion with the Fock operator as a zeroth order approximation to the exact Hamiltonian [17]

$$\hat{H} = \hat{H}_0 + V \ ; \qquad \hat{H}_0 = \hat{F}; \qquad V = \sum_{i<j} r_{ij}^{-1} - \sum_i u_i \tag{1.9}$$

where u_i is a potential of the motion of ith electron, obtained in the one-particle approximation. Then, the expansion of the wave function for a state $|\Psi\rangle$ with a dominating contribution from $|\Psi_0\rangle$ and the corresponding expression for the energy are given order by order

$$|\Psi\rangle = \hat{\Omega}_{PT}|\Psi_0\rangle = (1 + \hat{\Omega}^{(1)} + \hat{\Omega}^{(2)} + \cdots)|\Psi_0\rangle \tag{1.10}$$

$$E = E^{(0)} + E^{(1)} + E^{(2)} + \cdots \tag{1.11}$$

where the indices in the parenthesis (i) refer to orders in the pertubative expansions and $|\Psi^{(i)}\rangle = \hat{\Omega}^{(i)}|\Psi_0\rangle$; $|\Psi^{(0)}\rangle = |\Psi_0\rangle$. The zeroth order description is obtained using the one-particle approximation $\hat{F}|\Psi_0\rangle = E_0|\Psi_0\rangle$; $E^{(0)} = E_0$, as indicated by the partitioning of the total Hamiltonian. The operator $\hat{\Omega}_{PT}$ is called a *wave operator* and occurs in the theory of effective Hamiltonians (see Sec. 4.1). Acting on the model functions it produces the exact functions. In this and the two next subsections it will simply denote a generator of the particular (here perturbative) expansion characteristic for a given post Hartree-Fock method.

One reason for the success of the Rayleigh-Schrödinger approach is that the exact energy is not explicitly present in this formalism and therefore it can be applied to a group of states simultaneously and leads to the energy-independent effective operators [17]. Second reason is probably more important. Using the so-called diagrammatic technique (graphical representation similar to that introduced by Feynman in field theory) one can conveniently derive formulas for energies and properties at subsequent orders in terms of sums of products of molecular integrals [17]. Using this technique, it was also formally shown by Goldstone [18] that the so-called *unlinked* terms, that have non-physical i.e. non-linear dependence on the number of electrons in the system disappear from the expansion and they do not need to be considered at subsequent orders.

When a particular method scales properly (i.e. linearly) with the size in a homogeneous system we say that it is *size-extensive* [19] [21]. In case of the methods formulated in the second quantization language size-extensivity is insured by the requirement of linked expression for the energy and is very often simply identified with the lack of unlinked terms [20]. The Møller-Plesset Many Body Perturbation Theory (MBPT) i.e. the RSPT with the partitioning of the Hamiltonian of the form (1.9), first proposed by Møller and Plesset, is size-extensive, even when truncated at finite order, taking benefit of the linked diagram theorem [18].

Related, although different notion which deals with the problem of correct scaling with size is *size-consistency.* As defined by Pople *et. al.* [22] a method is considered size-consistent if the energy of a (super)system $A \cdots B$ composed of two (sub)systems A and B far apart is equal to the sum of subsystem energies computed separately by the same method. Hence we require that for any molecule AB the following separability condition is satisfied

$$E(AB) \overset{r_{AB}\to\infty}{\longrightarrow} E(A\cdots B) = E(A) + E(B) \tag{1.12}$$

where A and B may be open or closed shell fragments.

The above condition for a supersystem $A \cdots B$ may be satisfied by a given method only when the separable (i.e. correctly dissociating) reference function is used. Thus size-consistency imposes additional requirement on the zeroth order description and refers to a desired model of the dissociation processes. While using separable zeroth order description, size-extensivity ensures size-consistency. The reverse is however not true. In the interaction regions size-extensivity is still well defined as lack of unlinked terms, whereas size-consistency i.e. additivity of a supersystem energy is not (the same concerns atoms).

There is a lot of confusion in the literature, regarding the terms size-extensivity, size-consistency and separability. There is no canonical (commonly accepted) distinction among these terms. Therefore it is necessary to evoke explictly the definitions exploited here (for further discussion see Sec. 4.1).

Perturbation methods that have been very successful and are routinely used are second, third and fourth order Møller-Plesset MBPT. This is a well defined theory that can be unambiguously applied in completely automated way [1]. The success of the MBPT method depends on the proper zeroth-order description of the system. Perturbation theory can be applied safely for the ground states of closed-shell molecules around equilibrium geometries.

Unfortunately, the generalization of PT to multireference situations (referred to as multireference PT – MRPT) is not straightforward. There is no well defined zeroth order Hamiltonian in this case. Some MRPT schemes have been formulated, but they rely on a rather arbitrary choice of $\hat{H}_0$ or are iterative. All of them become quite expensive already at the third order [80].

1.4.2 Configuration Interaction method

Expressing the wave function as a linear combination of (all or a part of) determinants belonging to $\mathcal{H}^N$

$$|\Psi\rangle = \sum_{L=0}^{d} c_L |\Psi_L\rangle; \qquad d \leq \dim(\mathcal{H}^N) \tag{1.13}$$

and applying the Ritz variational principle, one obtains conceptually the simplest approach to electron correlation problem, known as the method of *configuartion interaction* (CI) or *superposition of configurations* [31]. It leads to well known eigenvalue problem

$$\mathbf{Hc} = E\mathbf{Sc} \tag{1.14}$$

where usually the overlap matrix $\mathbf{S} = \mathbf{I}$. When all determinants are employed we talk about *full configuration interaction* (FCI), otherwise about limited CI.

The method is variational i.e. E is an upper bound to the exact energy (of the considered Hamiltonian) [31]. As the one-particle basis approaches completness the FCI energy approaches the exact energy. Since the FCI result is the best one can get in a given basis set it is often used as a benchmark result for other calculations.

Variational methods are much more robust although computationally more demanding, since they are iterative and their complexity is at least of the third order (n^6). They are also not so easy to use as MBPT methods: except for the most commonly used configuration interaction method with singly and doubly excited configurations (CISD) out of a single reference, roughly equivalent in accuracy (at least when the size-inextensivity errors are small) as well as computational complexity to the MBPT(3), they require experience in the selection of configuration space and interpretation of their results.

While MBPT methods may give us energies and properties directly, variational methods almost always compute wave functions. For large basis sets or highly excited configurations vectors of wave function coefficients become prohibitively long, making the computer memory, rather than the time of computations, the main barrier of further progress. The full CI (FCI) techniques used in recent years produce vectors of dimension up to $10^6 - 10^9$ [29] and an order of magnitude increase is in sight [30].

Nevertheless, only limited CI schemes (CISD, CISDTQ) remain practical, despite tremendous progress in both computer and CI technology. Thus, in the expansion of the CI wave function, which may be rewritten in the intermediate normalization as

$$|\Psi\rangle = \hat{\Omega}_{CI}|\Psi_0\rangle = (1 + \hat{C}_1 + \hat{C}_2 + \cdots + \hat{C}_N)|\Psi_0\rangle \tag{1.15}$$

one is usually tempted to truncate it at a given level of excitations. The k-body excitation operators $\hat{C}_k$ generate k-tuply excited functions (with respect to $|\Psi_0\rangle$) with the proper coefficients i.e.

$$\hat{C}_k|\Psi_0\rangle = \sum_{|\Psi_L\rangle \in \mathcal{H}_K^N} c_L|\Psi_L\rangle \tag{1.16}$$

where the summation over L is restricted to a proper subspace $\mathcal{H}_K^N$ consisting of k-tuply excited functions only. In the second quantized form they become

$$\hat{C}_k = \sum_{\substack{a_1 < \cdots < a_k \\ m_1 < \cdots < m_k}} c_{m_1 \cdots m_k}^{a_1 \cdots a_k} \hat{e}_{m_1 \cdots m_k}^{a_1 \cdots a_k} \tag{1.17}$$

where the indices m_i refer to occupied spin orbitals, a_i to the virtual ones and

$$\hat{e}_{m_1 \cdots m_k}^{a_1 \cdots a_k} = a_{a_1}^+ \cdots a_{a_k}^+ a_{m_k} \cdots a_{m_1} \tag{1.18}$$

Any truncated CI suffers from very serious formal drawback. It is neither size-extensive nor size-consistent [23] [20]. It means that one cannot properly describe using approximate CI schemes, processes like dissociation or extended systems like crystals. Because cancellation of unlinked effects that are present in truncated CI schemes is only possible by mixing different categories of excitations, a remedy for this problem may be achieved by an exponential factorization of the wave operator [20]. Nevertheless, one should note that the Multireference CI (MRCI) method, which is direct generalization of the single reference counterpart, allows in principle to remove substantial part of the size-consistency error.

1.4.3 Coupled Cluster method

The Coupled Cluster (CC) method, originally formulated in the framework of nuclear physics by Coester and Kümmel, was later introduced to molecular electronic theory by Čižek [64]. The CC method is non-variational but it is size-extensive. The CC wave function has an exponential structure

$$|\Psi\rangle = \hat{\Omega}_{CC}|\Psi_0\rangle = e^{\hat{T}}|\Psi_0\rangle \tag{1.19}$$

where the cluster operator $\hat{T}$ is given as

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_N \tag{1.20}$$

$$\hat{T}_k|\Psi_0\rangle = \sum_{|\Psi_L\rangle \in \mathcal{H}_K^N} t_L^k|\Psi_L\rangle \tag{1.21}$$

with the coefficients t_L^k called cluster amplitudes, which are unknown quantities to be determined. The index k in the amplitude t_L^k says that this amplitude is associated with the promotion of k electrons from k occupied spin orbitals to k virtual ones and not with a product of two promotions defining finally the same function $|\Psi_L\rangle$. Such products occur due to the exponential structure of the CC expansion and as a result higher excited functions appear also with the coefficients being products of lower rank amplitudes. For example the total coefficient in the CC expansion of a quadruply excited function $|\Psi_{L_0}\rangle$ will involve $t^4_{L_0}$ and products of the type $t^2_{L_1}t^2_{L_2}$ as well, where L_1 and L_2 denote some doubly excited functions defining in terms of excitations decomposition of the function L_0. The additional index k becomes obsolate when using the second quantized definition of the cluster operator

$$\hat{T}_k = \sum_{\substack{a_1 < \cdots < a_k \\ m_1 < \cdots < m_k}} t^{a_1 \cdots a_k}_{m_1 \cdots m_k} \hat{e}^{a_1 \cdots a_k}_{m_1 \cdots m_k} \tag{1.22}$$

One way to look at the CC expansion is based on the comparison with the MBPT method. When rearranging all terms not order by order but according to the rank of k-body effects included one finds that it may be realized by means of the exponential expression (1.19) for the wave function. Moreover, as shown by Hubbard [27], there is a CC counterpart of the linked energy diagram theorem in MBPT stating that only the *connected* diagrams has to be retained in the CC expansion [20] [28].

The working equations of the CC method are usually obtained by projections against functions defining the corresponding CI space [28]. Premultiplying the Schrödinger equation $\hat{H}|\Psi\rangle = E|\Psi\rangle$ by e^{-T} one obtains

$$e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle = E|\Psi_0\rangle \tag{1.23}$$

or equivalently, using the identity [28]

$$\hat{H}|e^{\hat{T}}\Psi_0\rangle = (\hat{H}e^{\hat{T}})_c|e^{\hat{T}}\Psi_0\rangle = e^{\hat{T}}(\hat{H}_N e^{\hat{T}})_c|\Psi_0\rangle \tag{1.24}$$

and subtracting $\langle\Psi_0|\hat{H}|\Psi_0\rangle$

$$(\hat{H}_N e^{\hat{T}})_c|\Psi_0\rangle = \Delta E|\Psi_0\rangle \tag{1.25}$$

where $\hat{H}_N$ indicates the normal product form of the Hamiltonian and $(\ )_c$ means that only connected terms are included [20]. Projecting eq. (1.25) onto $|\Psi_0\rangle$ and the set of the excited state functions, one obtains the expression for the correlation energy $\Delta E = E - \langle\Psi_0|\hat{H}|\Psi_0\rangle$ and the amplitudes of subsequent excitation, respectively. Certainly one might also project eq. (1.23) and then using the Hausdorff-Cambell formula one obtains an equivalent set of equations for cluster amplitudes [5].

The simplest CC type approximation is the CCD method. Within this scheme one postulates that $\hat{T} = \hat{T}_2$. The corresponding equations for the energy and amplitudes of doubly excited determinants (defined by a set of the one-particle indices $|^{ab}_{ij}\rangle$ where i, j are occupied indices which are replaced by the virtual indices a, b) have the form

$$\Delta E = \langle\Psi_0|\hat{H}_N(1 + \hat{T}_2)|\Psi_0\rangle_c \tag{1.26}$$

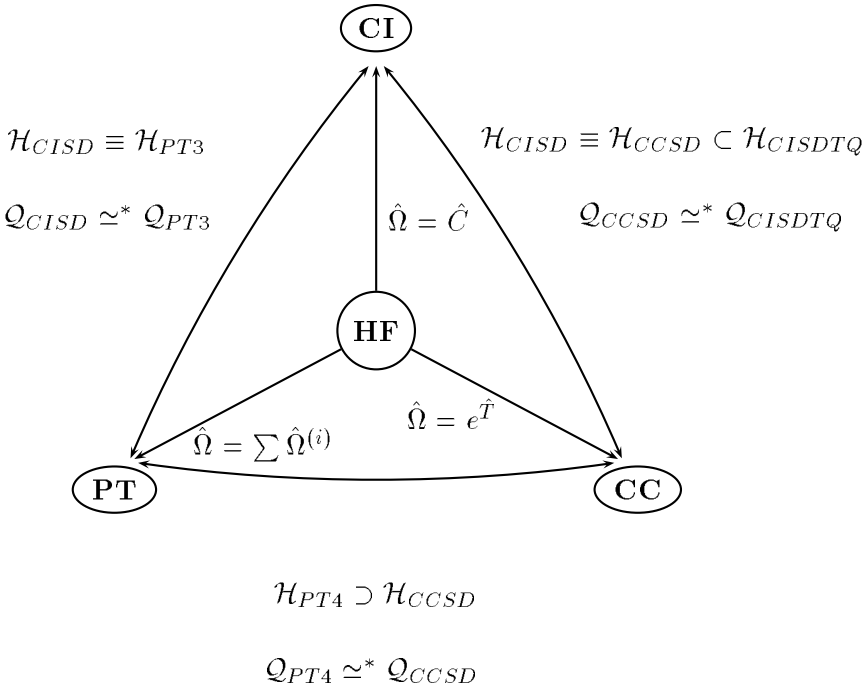
$$0 = \langle^{ab}_{ij}|\hat{H}_N(1 + \hat{T}_2 + \frac{1}{2}\hat{T}_2^2)|\Psi_0\rangle_c \tag{1.27}$$

As may be seen from the above formulas, all the linked quadruple excitation energy diagrams that arise from the disconnected part of the wave operator $\hat{T}_2^2$ are included in CCD. The resulting energy is accurate up to the fourth order of MP(DQ) i.e. Møller-Plesset PT in a space of all double and quadruple excitations from the closed shell ground state (for open shells that may require multireference treatment the comparison is not so obvious) and moreover includes infinite summation of all pair effects [24]. The computational complexity of the CC(S)D procedure is of the order n^6 [25]. Neglecting most of the non-linear terms while retaining those which are necessary to restore size-consistency one obtains various approximations of the CEPA (Coupled Electron Pairs Approximation) type [71].

From the formal point of view the CC method is certainly superior to the CI method. It is size-extensive, has nice invariance properties and is highly accurate even when truncated at low rank operator level. On the other hand the coupled cluster approach is more difficult to generalize to the multireference situations than the CI approach. The impressive development of Multireference Coupled Cluster (MRCC) methods has led to mature formalisms and tractable approximations [65]. An increasing number of applications to atomic [77] and molecular [77] [84] [97] systems is encountered in the recent literature. However the computational complexity of these methods is very high and there are still many specific problems that have to be overcome [78] in order to reach the 'black box' level.

1.4.4 Wind rose

Figure 1.1: Standard molecular *ab initio* methods



Some of the previously presented observations about the basic approaches to the electron correlation and their most typical approximations are put together in the schematic diagram 1.1. $\mathcal{H}_X$ indicates here the subspace of $\mathcal{H}^N$, the problem is projected onto (within a given approximation). $\mathcal{Q}_{\mathcal{X}}$ indicates the range of interactions (mixing of different $\mathcal{H}_{\mathcal{S}}$, $\mathcal{H}_{\mathcal{D}}$, ... subspaces) that is actually taken into account and is related to the quality of a given approximation.

The relation $\simeq$ may be read as 'approximations of comparable quality' i.e. including the same range of interactions (although not necessarily at the same level of accuracy). The star $*$ denotes 'practically comparable approximations' in the sense of comparable quality of results, provided that each method is used in the range of its applicability (e.g. when perturbation series is quickly convergent, size-consistency error is small or the cluster assumption is reasonable for PT, CI and CC respectively). PTN means Nth order of the MBPT method (for the energy). Parametrization of the wave operator (simply as a generator of the expansion, not necessarily implying the use of the theory of effective Hamiltonians) is indicated by arrows labels.

Perhaps the relationship between PT4 and CCSD requires more detailed comment. In fact the restriction of $\hat{T}$ to $\hat{T}_1$ and $\hat{T}_2$ is an approximation which omits some important 4^{th} order corrections to the energy (i.e. the linked contributions of the triples) but includes some higher order effects: 5^{th} (resp. 6^{th}) order corrections coming from $\hat{T}_1^2\hat{T}_2$ and $\hat{T}_1^3$ (resp. $\hat{T}_1^4$). Thus, one can expect similar results in PT4 and CCSD, except when some $\hat{T}_3$ terms have very large amplitudes [24].

Finally, one should remember that the above schematical presentation refers to the single reference methods and is much more approximate in case of multireference schemes (in this case only Hilbert space CC approch may be compared directly to MRCI for instance - see Chap. 5).

1.5 Introduction *par excellence*

After brief presentation of the methodological background of this thesis, it is time to introduce more precisely the methods, it is conserned with. Figure 1.2 shows their relative positions with respect to the three main (molecular) approaches to correlation energy i.e. CI, CC and PT with arrows indicating the 'migration' of ideas.

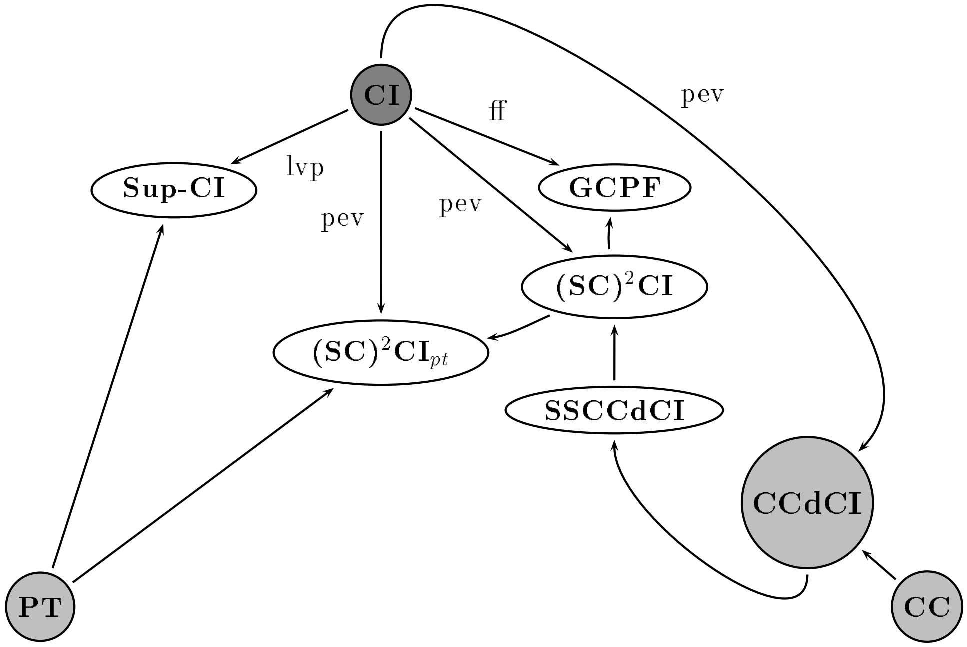
The CI node is here treated in a special way with arrows going out described by the following labels: *lvp* - linear variational principle, *ff* - functional form of the method, *pev* - pseudoeigenvalue problem. All these terms and relationships between the new algorithms and standard approaches will be described in details in the subsequent parts of the thesis.

For the sake of simplicity, the MR acronyms, indicating multireference procedures are omitted. It has also deeper motivation. All the methods, even though they may be formally single reference (i.e. there exists a state function which plays distinguished role) work in multireference spaces i.e. generated as single and double substitutions from a set of reference state functions.

There are many procedures (e.g. SSCC method proposed by Adamowicz [92], some versions of MRPT [44]) that similarily work in multireference spaces, although they employ single reference formalism. To avoid mistakes one should call them single reference methods in multireference spaces (SR(MR)), as proposed by Duch [44]. All the methods presented below are at least of the SR(MR) type.

- The Superdirect Configuration Interaction or **Sup-CI** method. It has been proposed in-

Figure 1.2: New algorithms: methodological perspective



dependently by Duch [39] and by Bendazzolli *et. al.* [42]. It has the usual versatility and stability of the CI methods with computational efficiency typical to that of the many body methods, such as the many-body perturbation theory (MBPT). Hamilton operator is projected into a space of a few trial vectors, such as Krylov, Nesbet or Møller-Plesset perturbation correction vectors. In this space Hamiltonian matrix elements may be directly computed in the many-body fashion, as weighted sums of integral products over orbital indices and such formulas up to the third order are derived in the second chapter. Variation-perturbation method based on the first order wave function is equivalent to the Sup-CI method with a single correction vector of the Møller-Plesset type. The Sup-CI method in third order is consisdered in this work.

- The Coupled Cluster method through the pseudoeigenvalue problem or **CCdCI** ('-d-' standing for 'dressed'). It was demonstrated by Malrieu *et. al.* [104] that a set of non-linear equations for every approximate CC method may be translated into a set of 'dressed' linear equations of the corresponding CI approximation. In other words the CC methods may be formulated in terms of the pseudoeigenvalue problem. Then, using very stable and efficient Davidson-like diagonalization procedures one may obtain its solution in an iterative, self-consistent manner or directly, using perturbative arguments for the non-linear, coefficient-dependent, dressing terms. Another, very important advantage of such reformulation is great flexibility, comparable to that of pure CI. Selected schemes that employ only a part of a given class of excitations (instead of all of them taken into account in CC once a given $\hat{T}_i$ is included) and treat the remaining small contributions by lower order methods are easily obtainable.

One considers in this work various dressings of the MRCI matrix by non-linear terms appearing in CCSD or CEPA methods, in a way insuring size-consistency of such defined procedures. The state-specific size-consistent multireference procedure of the CC type (referred to as **SS-CCdCI**) as well as its generalization to state-universal or multistate procedure (**MS-CCdCI**) are presented in the chapters 5 and 6.

Approximate schemes of the CEPA or MRCEPA type: Size-consistent Self-consistent Configuration Interaction or **(SC)^{2}CI** method and Generalized Coupled Pair Functional or **GCPF** method, and Multireference or **MR(SC)^{2}CI** method, respectively, are also introduced (the (SC)^{2}CI method for pedagogical reasons - the author has not contributed to its invention but has clarified the relationship between this scheme and conventional CEPA method). The character of approximations introduced to CCdCI to obtain such CEPA type procedures will be discussed in the Chapter 4.

- The acronym **(SC)^{2}CI(pt)** denotes Size-consistent Self-consistent combination of selected CI and perturbation theory, which is actually a selected CEPA type procedure with an account of a pertubative type for the influence of remaining small (i.e. not selected) doubles.

None of the algorithms presented here matches all basic characteristics of the pure CI method. Nevertheless, there is a common denominator occurring in all of them: at least a part of CI machinery used to maximize the efficiency of the method.

From the formal point of view the pseudoeigenvalue problem cannot be regarded as arising from a linear combination of a set of state functions. However, the procedure for solving this problem is the same as for the original set of equations of the CI method. The Sup-CI method in turn, although it employs compact linear expansion of the wave function and variational principle, similarily to MBPT avoids the explicit construction of the long vector of the wave function coefficients and uses matrix elements expressed in terms of weighted sums of integral products over orbital indices. In light of the above, the title of this thesis reflects my subjective point of view on the matter (coming from a tradition I have been growing up). I hope hewever that it expresses the right perspective in this methodologically confusing situation.

Theoretical methods of quantum chemistry have reached a high degree of sophistication in recent years. A good deal of this sophistication has been passed on to research chemists. Thanks to such "black box" packages as GAUSSIAN [1], HONDO [2] or GAMESS [3] *ab initio* systems of programs, computational chemistry is finding its way to the real world applications.

The simplest molecular *ab initio* methods, such as the Hartree-Fock and the second order perturbation theory, are the most frequently used. These methods are not only the least expensive but also the easiest to use. Application of more sophisticated methods requires much more understanding of the theoretical techniques and more computer resources. However, real chemical processes, such as multiple bond breaking, chemical reactions or quasidegenerate situations, are not easy to describe adequately using simple methods. There is a great need for reliable and computationally inexpensive methods that could treat the more complicated chemical processes in qualitatively and quantitatively right way. My hope is that at least some of the algorithms I present in the subsequent parts of this work will find their way to the world of real applications.

Part II

Superdirect Configuration Interaction method

Chapter 2

Superdirect Configuration Interaction method

The Superdirect Configuration Interaction (Sup-CI) method [39] is presented in this chapter. A version in which third order contributions are computed [57] for a relatively small (10–100) space of reference and correction vectors has been implemented using traditional approach to computing of matrix elements and several numerical tests on benchmark problems are included. Different points of view on the superdirect CI method as well as selection of the best "effective first order spaces" and size-extensivity corrections in Sup-CI are discussed.

The MBPT like formulas for the matrix elements have been also derived, using the Symmetric Group Approach (SGA) to CI and an algebra of circular operators, invented for this purpose by Duch [56]. They are presented in the next chapter.

2.1 Introduction

Correlation energy, i.e. the difference between the exact non-relativistic energy and the Hartree-Fock results, is hard to calculate if we want the dynamical part of it (the non-dynamical part may be taken care of by small MCSCF or CASSCF calculations). From the practical point of view we are usually not so much interested in obtaining the exact results as in qualitatively correct description. What we really want to see in the case of stretched chemical bond is the potential curve which is parallel to that experimentally obtained or at least to the FCI potential curve. Such description should be possible with a small number of reference configurations to take care of the non-dynamical part of correlation and third-order perturbative corrections to include a substantial part of dynamical correlation.

In the *configuration interaction* method (CI) solution of the Schrödinger equation is obtained by expressing the wavefunction in the form of a linear combination of N–electron functions, so–called *Configuration State Functions* (CSFs). From a formal point of view this means, that the Hamilton operator is projected into a finite subspace of the Hilbert space, spanned by the CSFs. In this space one can find the elements of the matrix representation of the Hamiltonian, and then solve the eigenvalue problem [31] (see also Sec. 1.4).

This conceptually simple method of describing the electron correlation in practice has very serious limitations, connected with the slow convergence and thus the length of the CI expansion. Many schemes have been devised and employed in order to overcome these limitations. One may divide them in 2 classes: improving the CI techniques to treat very long expansions, or selecting CSFs to reduce the length of the expansion.

Since the Hamilton operator contains one and two-body interactions only it should be possible to reduce the N-particle problem to the 2-particle equations. The *direct* CI method proposed by Roos [32] was the first step in this direction, getting rid of the N-particle matrix elements but leaving the coefficients of the N-particle CSFs.

Direct CI has dominated development of CI algorithms enabling very long expansions in terms of CSFs. It became a standard method since over two decades, with various implementations based on graphical unitary and symmetric group approach [33, 34]. In this scheme solving the matrix eigenvalue problem is coupled with simultaneous matrix elements evaluation, without explicit construction of the matrix $\mathbf{H}$. The next step in this spirit is to avoid not only construction of the matrix, but also of the long eigenvectors.

Several CI methods belonging to the second class, aiming at reduction of the length of CI expansion, were formulated. In the simplest case they reduce the number of CSFs by performing numerical selection to find those CSFs that contribute the most to the wavefunction sought. In more complicated cases CSFs are not constructed as products of spinorbitals but are made from more complex (two or more-particle) functions such as geminals or explicitly correlated functions. In CI method Hamiltonian is always projected into the space of selected CSFs.

Projection into the space of CSFs build from explicitly correlated functions gives more accurate results than projection into the space (of similar dimension) of CSFs build from one-particle functions. However, computational complexity due to the complicated form of matrix elements is increased when explicitly correlated functions are used [11].

Another possibility exists: projection into relatively small number of N-particle functions constructed as linear combination of CSFs. This approach leads to various forms of contracted CI [38] and to the Sup-CI method. It has been recently formulated by Duch [39] as an improvement over the Connected Moment Expansion (CMX) [40, 41], and subsequently traced back to an earlier (expressed in a different language, starting from quite different point of view) work of Bendazzoli *et.al*, known as the FAST CI [42].

In the Sup-CI approach one avoids construction of both the Hamiltonian matrix and the eigenvector, obtaining the energies and other properties from equations formulated directly at the 2-particle level. The Hamilton operator is projected into a space of a few trial vectors, such as Krylov or Nesbet correction vectors used in diagonalization methods. In this small space (dramatically smaller in comparison to the space of CSFs in classical approach) matrix elements are computed directly as sums of integral products over orbital indices (using MBPT techniques), without explicitly generating an eigenvector.

The aim of the Sup-CI method is to combine the efficiency and simplicity of the many-body perturbation theory with the robustness of variational methods, optimizing the tradeoff between the desired simplicity of the method and its potential for the best description of chemical reactions involving breaking of molecular bonds. Since most chemical processes require multireference treatment the superdirect CI approach (at the lowest level) within a multireference scheme is used. In essence one obtains variational results with perturbationally selected and in some cases optimized zeroth-order multireference states, at the cost comparable to that of the third-order perturbation theory.

In the next section the essence of the superdirect approach will be summarized and the theoretical details described. Some aspects of the implementation of the method are discussed in the second section. In the third section some illustrative calculations on water, acetylene and NH_2 molecule bond breaking are reported.

2.2 Theory

2.2.1 Sup-CI ansatz

Starting from some zeroth-order solution $|0\rangle$ to the Schrödinger equation we may create a number of correction vectors improving the quality of this vector

$$|V(0)\rangle = R_V|0\rangle \tag{2.1}$$

where the operator R_V is an arbitrary function of the Hamiltonian or of some partition of the Hamiltonian. In the simplest case we define Krylov, Nesbet and Møller-Plesset first-order correction vectors [31] as:

$$\begin{aligned} |K(0)\rangle &= R_K|0\rangle = \hat{Q}_0\hat{H}|0\rangle && (2.2)\\ |N(0)\rangle &= R_N|0\rangle = (\hat{H}_D - E_{SCF})^{-1}\hat{Q}_0\hat{H}|0\rangle && (2.3)\\ |M(0)\rangle &= R_M|0\rangle = (\hat{H}_0 - E_0)^{-1}\hat{Q}_0\hat{H}|0\rangle && (2.4) \end{aligned}$$

The Epstein-Nesbet and the Møller-Plesset vectors are made from the Krylov vector dividing its elements by different denominators - in the first case diagonal elements $\hat{H}_D$ of the Hamiltonian, in the second case elements of the Fock operator $\hat{H}_0$. We could introduce more vectors of this type, for example by adding some scaling parameter that will define intermediate cases between Nesbet-Epstein and Møller-Plesset partitioning, or by taking higher powers in the denominator [42]

$$|M^k(0)\rangle = (\hat{H}_0 - E_0)^{-k}\hat{Q}_0\hat{H}|0\rangle \tag{2.5}$$

According to Olsen [29] zeroth-order vectors of the following form should be useful:

$$\begin{aligned} |O_N(0)\rangle &= (\hat{H}_D - E_{SCF})^{-1}|0\rangle && (2.6)\\ |O_M(0)\rangle &= (\hat{H}_0 - E_0)^{-1}|0\rangle && (2.7) \end{aligned}$$

If the Hamiltonian is expressed via the unitary group generators [33, 34]

$$\hat{H} = \sum_{ij}(i|j)\hat{E}_{ij} + \frac{1}{2}\sum_{ijkl}(ij|kl)(\hat{E}_{ij}\hat{E}_{kl} - \delta_{jk}\hat{E}_{il}) \tag{2.8}$$

one can easily partition it into physically meaningful terms, defining for example core-valence $\hat{H}(c-v)$, valence-valence $\hat{H}(v-v)$, core-external $\hat{H}(c-e)$ and valence-external $\hat{H}(v-e)$ operators:

$$\hat{H} = \hat{H}(c-v) + \hat{H}(v-v) + \hat{H}(c-e) + \hat{H}(v-c) \tag{2.9}$$

Correction vectors may be accordingly created for each of these operators (although in this case selection of correction vectors for the best convergence of the final wavefunction remains an open question) with the goal of improving and stabilizing the solution of the Schrödinger equation projected into the space of these vectors:

$$\mathbf{H}|\Psi\rangle = E\mathbf{S}|\Psi\rangle \tag{2.10}$$

where matrices **H** and **S** are defined in the space of the trial vectors:

$$\{\ |0\rangle = R_0|0\rangle,\ |K(0)\rangle,\ |N(0)\rangle,\ |M(0)\rangle\ ...\} \tag{2.11}$$

2.2.2 Sup-CI from different perspectives

There are several ways of looking at the Sup-CI.

1. The original point of view [39] was to look at the Sup-CI as the next logical step from classical CI through direct CI to complete reduction of equations to the single-particle level, leading to non-iterative approximation to the MR-CISD. It was motivated by the use of the expectation values of the Hamiltonian powers in the CMX expansion [40] on the one hand, and by the desire to avoid very long CI expansions, leading to logistic problems with storage of the bilion-term eigenvector [29].

2. Sup-CI is equivalent to the first few iterations in a Davidson-like diagonalization method in which more than one type of correction vectors may be used. In particular the use of the Krylov vectors leads to the convergence characteristic of the Lanczos method and the use of the Epstein-Nesbet vectors to the Davidson method [43]. The second iteration in Davidson diagonalization gives the lowest-order Sup-CI results; the complexity of such calculation is of the third-order or n^6. Multireference Sup-CI corresponds, however, to a new diagonalization method in which small matrix is created not only from the single new correction vector added to k-th iteration correction vector – in case of full configuration interaction requiring n^{4k} operations to compute elements of the new iteration vector and n^{4k+2} to compute the energy – but also correction vectors constructed from $H|I\rangle$ for the most important $|I\rangle$ components. These new vectors are constructed at the cost of n^4 operations only but computation of the elements of the small matrix involving these correction vectors require in general n^6 arithmetic operations.

3. Sup-CI is also a generalization of variation-perturbation theory. In the variation-perturbation theory variational results are obtained in the space of correction vectors obtained by perturbation theory. Instead of using higher-order correction vectors more correction vectors of the same order but of different type are added to the variational space, selectively exploring the space of highly excited CSFs. Second-order perturbation theory wave function is defined in the space of quadruply exicted configurations relatively to the zeroth-order reference function; in Sup-CI correction vectors are singly and doubly excited relatively to a number of CSF. The space of CSFs explored in Sup-CI is similar to that of the PTMR method [44] but the results are variational.

4. Perhaps the most fruitful point of view on the Sup-CI is to note that it belongs to the contracted CI type of methods. The Hamiltonian operator is projected into a space of contracted CSFs (CCSFs) to reduce the number of variational parameters: in the externally contracted CI [38] all CSFs sharing the same internal orbitals are contracted into one contracted configuration

state function (CCSF):

$$\begin{aligned}
|\Psi_{ECCI}\rangle &= \sum_{I}^{(N)} C_I|\Phi_I\rangle + \sum_{I}^{(N-1)} C_I \sum_{a} \tilde{C}_I^a|\Phi_I^a\rangle + \sum_{I}^{(N-2)} C_I \sum_{ab} \tilde{C}_I^{ab}|\Phi_I^{ab}\rangle \\
\tilde{C}_I^{ab} &= \frac{\langle 0|\hat{H}|\Phi_I^{ab}\rangle}{E_0 - \langle\Phi_I^{ab}|\hat{H}|\Phi_I^{ab}\rangle} \qquad (2.12)
\end{aligned}$$

The number of variational parameters left is equal to the number of internal paths for N, $N-1$ and $N-2$ electrons in the corresponding graph [33],[34].

In the internally contracted case all CSFs that do not contain any external orbitals are combined into one reference function from which excitations are formed. The general form of the internally contracted wavefunction is:

$$\begin{aligned}
|\Psi_{ICCI}\rangle &= C_0|\Psi_0\rangle + \sum_{ia} C_i^a|\Psi_i^a\rangle + \sum_{ijab} C_{ij}^{ab}|\Psi_{ij}^{ab}\rangle \\
\Psi_{ij}^{ab} &= (E_{ai}E_{bj} \pm E_{aj}E_{bi})|\mathrm{MCSCF}\rangle \qquad (2.13)
\end{aligned}$$

with the sign ± depending on singlet or triplet coupling of orbitals a, b. The number of variational parameters is almost independent of the number of the CSFs in the MCSCF state. Unfortunately full implementation of this idea requires computation of the 5-th order density matrices [37].

In the Sup-CI contractions are formed in a way similar to the externally contracted case, but disregarding internal-external division of orbitals:

$$\begin{aligned}
|\Psi_{SCCI}\rangle &= \sum_{I,k} C_{I,k}\hat{R}_k|\Phi_I\rangle = \sum_{I,k} C_{I,k} \sum_{ijab} C_{ij;k}^{ab}|\Phi_{I,ij}^{ab}\rangle \\
C_{ij;k}^{ab} &= \langle\Phi_{I,ij}^{ab}|\hat{R}_k|\Phi_I\rangle \qquad (2.14)
\end{aligned}$$

where k numbers different types of $\hat{R}$ operators.

5. We project the Hamiltonian into the small space of first-order correction vectors effective in different iterative diagonalization methods. Convergence of the Sup-CI is therefore strongly correlated with the convergence of the diagonalization methods. Effective Hamiltonians [45] aim at projection of exact wavefunction onto a finite model space. Since ”exact” refers to ”exact in a given one particle-basis”, i.e. to the full CI wavefunction, Sup-CI Hamiltonians may also be regarded as effective Hamiltonians. The Sup-CI method is a particular method of finding the effective Hamiltonian not by Bloch or des Cloizeaux formulation (in practice always solved by perturbation theory) but by searching first for the effective spaces, i.e. best combinations of the CSFs, and than solving variational problem.

6. From computational point of view the formulas in Sup-CI are of the 3-rd order and should be similar to the 3-rd order CIPSI procedure [53]. Unfortunately such a procedure has not yet been developed.

2.2.3 Convergence of the method.

Adding more correction vectors guarantees not only, by the variational principle, that the energy will decrease, but also serves to stabilize the method when some of the vectors do not give a proper correction to the zeroth-order state, for instance when low-order perturbation theory breaks down. The Krylov vector plays such stabilizing role since it has no "dangerous" denominators, although its influence on energy is not significant. In practice if the space of reference configurations describes the wave function in qualitatively correct way one type of correction vectors is enough to achieve high accuracy and stable behaviour of the method at different molecular conformations [57]. A single second-order correction vector may be more effective in decreasing the energy than a set of first-order vectors but it is also much more expensive to compute.

Convergence of variational series with the order of correction vectors may be quite rapid: in [39] it is shown that taking sixth order Krylov vectors gives in case of calculations for water around equilibrium geometry essentially exact results. For many molecules results of variation-perturbation calculations are better than the results of perturbation theory in the corresponding order. Using first-order Møller-Plesset function the best linear combination with the reference state (i.e. Sup-CI in the space $\{|0\rangle,\ |M(0)\rangle\}$) leads to the following formula:

$$E_{s1} = \frac{E_3 - E_2 - \sqrt{(E_3 - E_2)^2 + 4\langle\psi_1|\psi_1\rangle E_2^2}}{2\langle\psi_1|\psi_1\rangle} \tag{2.15}$$

Variational correlation energy is expressed here in terms of the second and third order perturbation energies and the norm of the first-order wavefunction. This is the simplest version of the superdirect CI with a single correction vector obtained from perturbation theory.

Sup-CI energies are always better than the variation-perturbation results and may be computed starting from the restricted as well as the unrestricted Hartree-Fock formalism. Variational energies obtained from equation (2.15) are usually very close to, and sometimes even slightly better than, the third order Møller-Plesset energies (examples are given below). Although it has been known for a long time now that variation-perturbation method is capable of high accuracy the method has been rarely used, parhaps because it has been labeled as "not-extensive". However, it is quite obvious that variational energies lower than the non-variational ones are more accurate and are "better" in the sense of being more uniform at different molecular conformations.

For some molecules single reference variation-perturbation energies E_{sk} are already lower than the perturbation theory at the corresponding $2k+1$ order. For example, unrestricted MP3 correlation energy calculations [46] for lithium gave -32.21 mH and the variation-perturbation result is E_{s1}=-32.68 mH, calculation on carbon $C(^5S)$ gave -64.10 mH and E_{s1}=-64.43 mH, carbon $C(^3P)$ gave -109.25 mH compared to E_{s1}=-109.34 mH, for BeH UMP3 correlation energy is -70.10 mH and E_{s1}=-70.65 mH. The calculations for BH molecule around equilibrium geometry in the DZP basis give E_{s1} lower on about 2 mH than E_3 and E_{s2} lower on 0.8 mH than E_5. Similar differences have been noted for calculations on the Be_2 molecule in 7s3p1d basis [57].

To test higher order convergence Sup-CI calculations for water in equilibrium geometry and with the bonds stretch symmetrically to twice their equilibrium values have been performed, with one type of correction vectors selected for each series of calculations [57]. For comparison with previous calculations [41, 44, 47] 6-21G basis set was used. In Tables 2.1 and 2.2 effectiveness of different correction vectors in variation-perturbation method is compared.

Table 2.1: Convergence of variation-perturbation method, H_2O.

Order	MP	V-P Krylov	V-P MP	V-P EN
2	120.865			
3	124.167	110.016	120.482	119.868
4	129.017			
5	129.505	123.078	129.252	129.205
6	129.940			
7	130.016	127.180	130.008	130.009
8	130.063			
9	130.076	128.713	130.077	130.078
10	130.082			
11	130.084	129.456	130.084	130.084
12	130.085			
13	130.085	129.775	130.085	130.085
14	130.085			
15	130.085	129.922	130.085	130.085
Exact	130.085	130.085	130.085	130.085

Calculation on water in equilibrium distance (from [39]), 6-21G basis set [47]. Results for different correction vectors. Minus correlation energy given in millihartrees.

In this case convergence with Krylov vectors is rather slow (although much faster than convergence of the CMX expansion [40, 41] using the same expectation values). The differences between Møller-Plesset perturbation theory and variational results computed with Epstein-Nesbet correction vectors are negligible, with Sup-CI based on Møller-Plesset correction vectors converging slightly slower. For stretched geometry convergence of variational series is much slower (see Table 2.2), with Krylov vectors it is very slow and convergence of perturbative series is quite erratic. As we shall see it is much better to use first-order correction vectors for a few reference configurations instead of higher-order correction vectors.

It should be noted that higher order perturbation theory functions lead to slightly different variational energies than the correction vectors obtained in iterative methods. For example, in Table 2.2 the Møller-Plesset (Epstein-Nesbet) variational correlation energy obtained from the 3-rd iteration of Davidson method is -266.912 (-277.382 mH), and from the 4-th iteration -290.635 mH (-292.932 mH), while the best energy obtained from the first and second-order perturbation wavefunction is -269.849 mH (-274.979) and with the third-order wavefunctions -288.870 mH (-292.003). This suggest that a diagonalization method based on Nesbet first-order correction vectors used iteratively, as Davidson method does [43], converges as rapidly as the second-order method based on the vectors obtained from perturbation theory.

Recently a number of papers devoted to the restricted Møller-Plesset theory for open-shell molecules and MCSCF reference functions appeared [48]. While multireference Møller-Plesset theory is formulated via iterative approach and cannot easily be reduced to the MBPT form

Table 2.2: Same as Table 2.1 but for bonds stretched to $2R_e$.

Order	MP	V-P Krylov	V-P MP	V-P EN
2	241.643			
3	235.520	135.715	204.372	211.023
4	281.985			
5	286.204	179.735	266.912	277.382
6	298.573			
7	299.605	217.156	290.635	292.932
8	303.671			
9	301.351	242.831	295.950	297.252
10	302.768			
11	300.577	262.244	298.362	298.658
12	300.854			
13	299.898	274.407	299.224	299.406
14	299.875			
15	299.729	283.161	299.610	299.762
Exact	299.864	299.864	299.864	299.864

the open-shell methods are in this respect analogous to the single-reference case [80]. Since convergence properties of these new methods is similar results obtained by Knowles *et.al* should be typical [48]. One may expect that using semi-canonical orbitals of Knowles *et.al* [48] will significantly improve convergence of perturbational as well as variational series. However, one can notice that even if perturbation series is divergent variational series may converge quite well.

For the ${}^2B_1$ state of NH_2 molecule at 2 R_e geometry starting from usual ROHF Hamiltonian, as implemented in the GAMESS system [3], leads to the divergent perturbation series (see Table 2.5), for example E_8 in Møller-Plesset series is -61625 Hartree and in Epstein-Nesbet series is +0.055 Hartree). Despite this divergence of non-variational energies and similar divergence of expectation values computed with the perturbative wavefunction (third order values are +2.566 in the Møller-Plesset and +0.12849 in Epstein-Nesbet case) as well as the diverging norms of correction vectors using perturbative wavefunctions for variational expansion leads to satisfactory convergence of correlation energy: with Møller-Plesset vectors the first 3 energies are: $E_s = -0.1656,\ -0.1864,\ -0.1962$ and with the Epstein-Nesbet vectors much better convergence is obtained: $E_s = -0.1765,\ -0.2337,\ -0.2512$ (exact result is -0.2643) [57]. It is clear that when several types of correction vectors are included in the superdirect CI calculations convergence should be reliable.

How can one improve convergence of the Sup-CI, especially for the open shell cases and excited states? It is not difficult to include simplified higher-order corrections using diagrammatic techniques for summation of certain terms (like ladder diagrams) in computation of $\langle 0|R_i^k H R_j^l|0\rangle$ matrix elements. Although such summations have not been included in the calculations presented here, it has been verified that adding second-order correction vectors with pair excitations only may significantly reduce the remaining error and such terms should be included in future versions of the method.

2.2.4 Size-extensivity corrections

Corrections for size extensivity [49, 50], although not an ideal solution, remove part of the error due to neglecting higher order excitations in singly and doubly excited CI. It is worthwhile to note that these corrections may be easily applied in the Sup-CI method. The variational energy E_{s1}, Eq. (2.15) may be corrected for size-extensivity errors using the correction formulas given in [49, 50]. First, C_0 coefficient should be computed:

$$C_0 = 1\Big/\sqrt{1 + \left(\frac{E_{s1}}{E_2}\right)^2 \langle\psi_1|\psi_1\rangle} \tag{2.16}$$

The formula for Davidson correction is:

$$\Delta E_{DC} = (1 - C_0^2)E_{s1} = \frac{E_{s1}^3\langle\psi_1|\psi_1\rangle}{E_2^2 + E_{s1}^2\langle\psi_1|\psi_1\rangle} \tag{2.17}$$

Renormalized Davidson correction formula is:

$$\Delta E_{RD} = \frac{1 - C_0^2}{C_0^2}E_{s1} = \frac{E_{s1}^3}{E_2^2}\langle\psi_1|\psi_1\rangle \tag{2.18}$$

The formula for the Davidson-Silver correction is identical with the first term of the CMX expansion [40, 41] and with the $E[2/1]$ Padé approximant [51] for E_2 and E_3 energies:

$$\begin{aligned} E_{s1} + \Delta E_{DS} &= E_{s1} + \frac{1 - C_0^2}{2C_0^2 - 1}E_{s1} \\ &= E_2^2/(E_2 - E_3) = E[2/1] \end{aligned} \tag{2.19}$$

The same formula is obtained as the linear approximation to the quadratic equation for the variation-perturbation energy from which E_{s1} in Eq. (2.15) was obtained.

Multireference versions of these formulas may also be applied to the Sup-CI results, including the qubic correction [52]. Since in practice all these corrections gave very similar results in test calculations the original Davidson correction, eq. (2.17) are always given [57]. However, it is clear that corrections in Sup-CI will be significantly smaller than in MR-CI for the same reference space. The reference energy is the same for both methods but MR-CI total correlation energy is larger and the norm of the reference configurations in the final function smaller.

2.2.5 Selection of reference space

Selection of configurations for the zeroth-order state is of great importance. This can be done by estimating energy lowering due to configurations at a given order of perturbation theory or their coefficients in the corresponding wavefunction. The simplest approach, scaling like n^4 (where n is the number of orbitals), estimates the importance of configurations by second-order perturbation theory relatively to the Hartree-Fock state, and is capable of selecting only those configurations that are no more than doubly excited relatively to the Hartree-Fock state. However, for some molecules this is not sufficient even around their equilibrium geometry, and in any case exploring potential surfaces one frequently finds highly excited configurations dominating at stretched geometries. Therefore a second-order method was used: small CIPSI calculations to construct reference space iteratively are performed before Sup-CI.

Starting from some arbitrary zeroth-order state and performing second order perturbation calculations one obtains (taking configurations with coefficients larger than a given threshold) new reference space. Diagonalizing Hamiltonian matrix in the reference space the new multiconfigurational zeroth-order state is obtained and the whole procedure is repeated until all important configurations are included. The final perturbation energy is an approximation to the full CI limit. The CIPSI method has proved its usefulness for the study of medium-size molecular systems [53]. It is closely related to various selected CI methods (cf. [54]). The inexpensive CIPSI scheme (especially if it is done via intermediate projection algorithm [55]) for constructing the reference space seems to be a reasonable choice for low order methods.

Second selection method was employed to test the effectiveness of the iterative CIPSI scheme as a generator of reference CSFs. Reference configurations are selected as those corresponding to the CSF coefficients in the renormalized first-order wavefunction larger than a given threshold. It allows to select all important doubly excited configurations only. This method is not really suitable for generation of reference configurations because it does not select higher-order configurations, but it is the simplest and the least expensive therefore comparison with CIPSI selections scheme was made in calculations for water.

2.3 Implementation of Sup-CI

One can develop the superdirect approach in several directions. In [39] estimation of higher order Hamiltonian expectation matrix elements was mentioned. Another choice is to test an inexpensive third-order method based on a relatively small zeroth-order space of CSFs. Diagramatic version of such method, restricted to the case of several singlet reference configurations has been developed by Bendazzoli *et.al* [42]. The Sup-CI may use larger reference spaces and different correction vectors generated from arbitrary open-shell configurations selected using second-order perturbation theory.

Matrix elements of H in the basis of correction vectors may be written down in the form similar to the expressions of the many-body perturbation theory. General open-shell second order formula was derived [56] as an example of application of the algebra of unitary group generators and other elements appearing in the superdirect method are derived here (see Chap. 3) along the same lines. Taking only the first-order correction vectors the complexity of the energy expressions is of the third order. For second order correction vectors elements of the fifths order would appear – for general open-shell CSFs they would be very difficult to derive and program. Since the final matrix is small it is easy to take care of the near-linear dependencies among the trial vectors. In this way we obtain variational results without the need to store the list of variational wavefunction coefficients, hence the name "superdirect".

The multireference superdirect method may be implemented in several ways. After the selection procedure the zeroth-order vector $|0\rangle$ is a linear combination of N_{ref} reference CSFs

$$|0(C^{(0)})\rangle = \sum_{i=1}^{N_{ref}} C_i^{(0)}|w_i^{(0)}\rangle \tag{2.20}$$

From this reference state linearly independent correction vectors $\hat{R}_V|0\rangle$ are generated, one for each operator $\hat{R}_V$. Assuming that $\hat{R}_0 = I$, taking MCSCF as the zeroth-order function and

$$\begin{aligned} |\Psi\rangle &= \sum_{i=0}^{L} C_i|\Phi_i\rangle = \sum_{i=0}^{L} C_i\hat{R}_i|0(C^{(0)})\rangle \\ &= C_0|0(C^{(0)})\rangle + C_1\hat{R}_1|0(C^{(0)})\rangle + \ldots C_L\hat{R}_L|0(C^{(0)})\rangle \end{aligned} \tag{2.21}$$

as the trial function for variational procedure Sup-CI method that is similar to the internally contracted CI [38] is obtained. To increase the number of variational parameters L one can use different types of excitations (core-valence, valence-valence, core-virtual, valence-virtual, one and two-electron) forming the $\hat{R}_V$ operators.

Since the energy depends on the $C^{(0)}$ coefficients a better approximation is obtained by optimization of all these coefficients to minimize the energy obtained from the $(L+1)$-dimensional generalized eigenvalue problem:

$$\tilde{\mathbf{H}}\mathbf{C} = E\tilde{\mathbf{S}}\mathbf{C} \tag{2.22}$$

where $\tilde{\mathbf{H}} = \tilde{\mathbf{H}}(C^{(0)})$ is the projection of the Hamiltonian into the space of the first order wavefunctions and $\tilde{\mathbf{H}} = \tilde{\mathbf{H}}(C^{(0)})$ of the overlap matrix. Although in this formulation the dimension of the eigenvalue problem is very small we have to solve the nonlinear optimization problem for $E(C^{(0)})$ and in practice there is no way to avoid computation of all matrix elements of the type:

$$\tilde{H}_{ik,lj} = \langle w_i|\hat{R}_k H \hat{R}_l|w_j\rangle \tag{2.23}$$

where $\tilde{H}_{i0,0j}$ are the usual CI matrix elements.

The "direct optimization" in which all LN_{ref} coefficients $C_k^{(0)}C_j$ are taken as independent, leading to larger $\tilde{\mathbf{H}}$ matrices requires smaller computational effort as the "indirect optimization" described above.

In calculations reported in the Sec. 2.4 three types of correction vectors are used: Krylov vectors (suggested by the gradient optimization of energy functional) and first-order perturbation theory corrections for both Møller-Plesset and Epstein-Nesbet partitionings. Each reference CSF is therefore associated with the following set of vectors:

$$\{\ |w_j\rangle = R_0|w_j\rangle,\ |K(w_j)\rangle,\ |N(w_j)\rangle,\ |M(w_j)\rangle\} \tag{2.24}$$

where j extends over all reference CSFs. From the computational point of view formulas for $\langle w_i|\hat{R}_k H \hat{R}_l|w_j\rangle$ matrix elements for all first order correction vectors are closely related, for fixed i, j differing only in the denominators. Most of these matrix elements involve summation of two indices only (if w_i is quadruply excited relatively to w_j), some four indices and only diagonal $\langle w_i R_k|H|R_l w_i\rangle$ elements are of the third order (i.e. involve six-fold summations). Therefore one may expect approximately linear scaling of the timing of the Sup-CI program with the number of reference configurations. The overall complexity of the method should be $N_{ref}n^6$.

The timing should be faster than a single iteration of the MR-CI procedure when computation of all matrix elements is programmed in the superdirect way (i.e. involving only summation over orbital indices, without reference to the MR-CI wavefunction). It is easy to generate more variational parameters increasing accuracy of the method without significantly increasing computational complexity: separating Hamiltonian operator according to Eq. (2.9) we get more correction vectors but no new matrix elements, increasing only the size of the small matrix in final diagonalization.

So far we have discussed calculation of energies only. In principle calculation of properties via Sup-CI method is equally simple: since we do have a compact wavefunction it is enough to derive appropriate matrix elements. Since the wavefunction in Eq. (2.22) is expressed as a combination of non-orthogonal functions contribution of a given function to the norm of the wavefunction is calculated as:

$$\begin{aligned} ||\Phi_i||^2 &= C_i\langle\Phi|\Psi\rangle = \sum_j C_i S_{ij} C_j \\ ||\Psi||^2 &= \sum_i ||\Phi_i||^2 = \mathbf{C^T}\tilde{\mathbf{S}}\mathbf{C} = 1 \end{aligned} \tag{2.25}$$

Finally let us stress that the Sup-CI method has to converge in a monotonic way to the exact energy when the number of variational parameters (reference configurations) is increased. The real convergence of the method is tested in a series of calculations described in the next section.

2.4 Test calculations

I would like to stress that all calculations reported here have not been yet done using the efficient formulas for matrix elements, presented in the next chapter. Their efficient implementation requires more time than the test calculations and is a matter of a future work. Instead the modified SGGA-CI program [34] has been used, extended to give perturbation theory and variation-perturbation results [44] from 4 iterations of the CI procedure. Energies are computed up to 8-th order and wavefunctions are computed up to 4-th order. The program is also capable of computing CIPSI and Sup-CI results, calculating matrix elements $\langle w_i|\hat{R}_k H \hat{R}_l|w_j\rangle$ of the Hamiltonian, although not so efficiently as the many-body version should compute them. All reported results are obtained with the properly spin-adapted CSFs. GAMESS system of programs [3] was used for generation of orbitals. All calculations were run on a personal computer PC-486 and some on a notebook computer PC-486 and are described in a more detailed manner in the ref. [57].

In some calculations even higher order perturbation or variation-perturbation calculations fail if the single reference start is used. For example Sup-CI at 7-th order (with third order correction vectors) gives for calculations on water (reported in details below), with the bonds stretched to twice their equilibrium distance, an error of the order of 10 mH with respect to the full CI limit, whereas at equilibrium geometry the error is 1/10 mH [57]. In such cases perturbation theory converges very slowly and the number of iterations in CI or CC procedures is quite large. It shows the importance of well defined zeroth-order state.

In order to check how the present inexpensive approach is capable of describing bond stretching processes, the multireference superdirect method on rather difficult cases, stretching two bonds in water and and an open-shell NH_2 molecule and the triple carbon–carbon bond of acetylene, is tested. These are rather challenging computations for the low-order methods.

2.4.1 Water

Some Sup-CI results have been already presented in Tables 2.1 and 2.2, in which high-order convergence rates for Møller-Plesset perturbation theory and variation-perturbation results are compared in calculations on water in 6-21G basis set. Multireference Sup-CI calculations on water were also done using the same double zeta basis set as in the benchmark calculations of Saxe *et.al* [58], for 3 geometries, with the bond length R_e, 1.5 R_e, 2 R_e [57]. The geometry of Bauschlicher and Taylor [58] is employed and, for comparison with calculations of Duch and Diercksen [44], the lowest orbital is not correlated.

In the first set of calculations configurations with coefficients larger than a given threshold in the first order Epstein-Nesbet function were included in the reference space (this partitioning gave always larger reference spaces than Møller-Plesset). This choice allows us to compare CIPSI selection scheme with the simplest second-order selection scheme. Two values of selection threshold are used, typical for CIPSI and MRD-CI calculations: 0.05 and 0.03. Depending on the respective geometry our reference space is composed of 10, 18 and 15 reference functions for the smaller threshold and 26, 34 and 26 reference CSFs for the larger one. Since SGGA program treats simultaneously all CSFs differing only by spin couplings and sharing the same orbital configuration some open-shell reference configurations are associated with more than one CSF (only one of them has to give large contribution). After some experimentation, the following correction vectors are finally included for each reference configuration $|I\rangle$: Epstein-Nesbet $|N(I)\rangle$, Møller-Plesset $|M(I)\rangle$ and a vector

$$|D(I)\rangle = (H_D - E_I)^{-1} H_D |N(I)\rangle \tag{2.26}$$

where H_D is the diagonal part of the Hamiltonian. Thus the size of the small matrix that is diagonalized is equal to 4 times the number of reference functions. In practice Epstein-Nesbet correction vectors recovered most of the correlation energy, with other correction vectors contributing no more than 1 mH in all cases, with their total contribution to the wavefunction norm being not larger than 0.001.

In this small space, using only first order correction vectors, very encouraging results were obtained – for every geometry more than 96.5% with the larger threshold and more than 95.5% of the full CI correlation energy with the smaller threshold was recovered. One cannot expect that the dynamical correlation energy will be well reproduced for this case but as one can see from Table 2.3 and Figure 2.1, especially with the larger number of references (smaller threshold) the results are comparable to the 5-th order MBPT results, with the error of 2.6, 4.5 and 11.2 millihartree at R_e, 1.5 R_e, and 2 R_e geometries respectively. Multireference Davidson corrections were also computed, giving in this case significant reduction of the error to 1.5, 3.1 and 5.7 millihartree, or more than 98% of correlation energy at all points.

One should stress that the first set of calculations was done only for comparison as the iterative CIPSI scheme of reference space selection is much better. It allows to take into account

Table 2.3: Results for water in DZ basis set

Method	R_e	$1.5R_e$	$2R_e$
FCI	-76.143399	-75.992471	-75.886986
Escf	137.7	204.4	307.6
E2	8.2	20.2	50.5
E3	7.6	27.2	74.3
E$_{s1}$	12.8	40.4	103.6
CI(SD)	7.7	22.6	60.6
E4	1.1	6.4	15.2
E5	0.9	5.1	15.9
E$_{s2}$	1.3	8.3	35.0
E6	0.1	2.0	3.7
CI(SDTQ)	0.2	1.1	4.2
E7	0.1	1.0	-0.9
E$_{s3}$	0.1	1.7	12.0
E8	0.0	0.5	-0.7
Superdirect and CIPSI results			
$S_{PT2}^{0.05}$	5.9	7.0	14.1
+ Davidson corr.	3.4	4.5	7.0
$S_{PT2}^{0.03}$	2.6	4.5	11.2
+ Davidson corr.	1.5	3.1	5.7
CIPSI 0.05	-7.2	-2.9	-2.7
$S_{CIPSI}^{0.05}$	5.9	5.8	9.3
+ Davidson corr.	3.4	3.8	6.4
CIPSI 0.03	1.5	-1.0	2.3
$S_{CIPSI}^{0.03}$	2.6	3.6	3.2
+ Davidson corr.	1.5	2.7	2.1

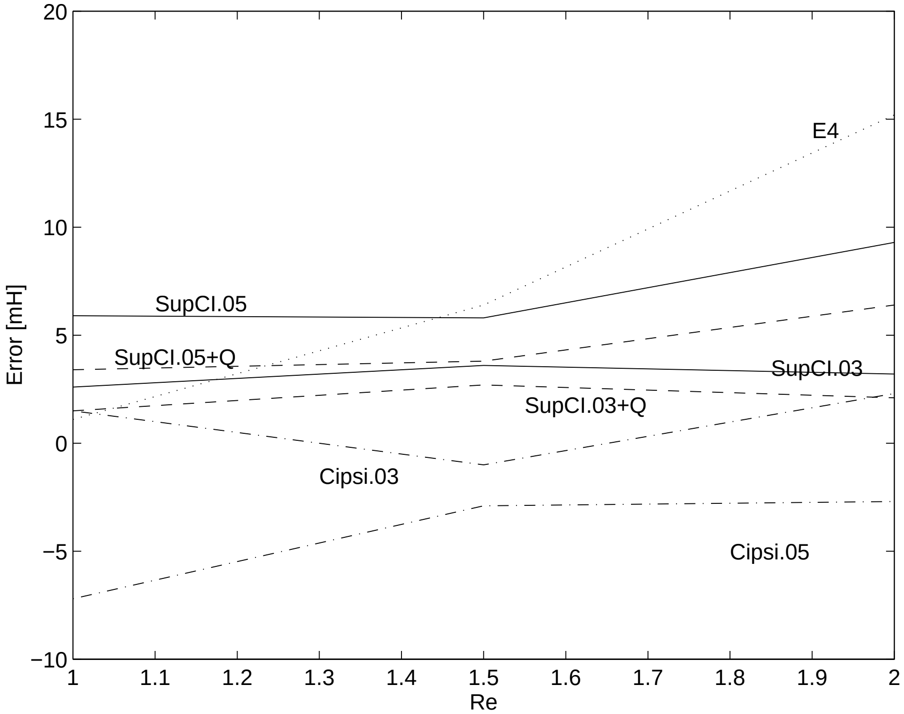
Geometry from Bauschlicher and Taylor [58], $1s$ orbital frozen. Energy differences in millihartrees, with respect to the full CI energy (except first row). Perturbation theory results and variation-perturbation results E_{sk} are obtained with the Møller-Plesset partitioning. $S_{CIPSI}^{0.03}$ means Sup-CI with the threshold 0.03 and modified CIPSI reference selection scheme, $S_{PT2}^{0.05}$ denotes Sup-CI with the threshold 0.05 and selection using first order wavefunction [57].

the most important configurations at the lowest possible level – using second order perturbation theory. However, we cannot expect that Sup-CI will remove the inadequacy of the reference space completly. The second set of calculations was done using modified CIPSI algorithm to construct the reference space. Epstein-Nesbet or Møller-Plesset first order correction vectors are built from the zeroth-order vector, obtained from the previous iteration by diagonalization of the Hamiltonian in the former reference space. In case of Epstein-Nesbet partitioning many configurations relatively strongly interact with the Hartree-Fock function (in comparison with the Møller-Plesset case). That implies selection of larger number of configurations already in the first iteration. Since Epstein-Nesbet and Møller-Plesset correction vectors are subject to intermediate normalization we compute their norm, storing at the same time information about all configurations with coefficients larger than the threshold, and renormalize coefficients of these selected configurations at the end of CIPSI calculations. Since renormalization reduces their magnitude some of the selected configurations are dropped. The full CIPSI vector is never stored, only the largest components.

Moreover, dropping from the reference set configurations already selected is not allowed if their coefficient became smaller than the threshold in the next iteration. Tests performed with such an option showed that cyclic solutions were possible. For example, 2 CSFs selected in the first iteration gave 8 CSFs in the third iteration, but this time renormalization will bring the number of CSFs with coefficient above the threshold back to the same 2 configurations: large number of reference CSFs leads to a large number of singly and doubly excited CSFs in the CIPSI wavefunction, and their combined weight reduces the norm of the reference configurations in the renormalized wavefunction. Increasing the number of CSFs until no new important configurations are found escapes from this cyclic atractor and leads to very rapid convergence, with 2-3 CIPSI iterations being sufficient for selection.

As one can see from Table 2.3 and Figure 2.1, the errors at different geometries are quite uniform with this selection scheme. In the tables E_n is the Møller-Plesset MBPT(n) energy; Epstein-Nesbet energies are in general worse (for $2R_e$ divergent). With both thresholds, T=0.05 and T=0.03, energies are remarkably parallel to the full CI results. For T=0.05 the errors are 5.9, 5.8 and 9.3 mH and for T=0.03 they are reduced to 2.6, 3.6 and 3.2 millihartree at R_e, $1.5R_e$ and $2R_e$ geometries respectively. With the Davidson corrections we have in the later case 1.5, 2.7 and 2.1 mH. The large improvement of results (comparing to the previous selection scheme) for $2R_e$ was obtained because new quadruply excited configurations appeared in our reference set. Even for a larger threshold 0.05 the superdirect results are quite good and for stretched bonds better than MBPT(5) and similar to CI-SDTQ.

Figure 2.1: Results for water in DZ basis set, three geometries



Selection of references by CIPSI. Energy relative to FCI. E_4 is Møller-Plesset energy.

With the threshold 0.03 the accuracy of the description of dissociation is better than CI-SDTQ and much better than the MBPT(5). This time the maximum dimension of the Hamiltonian matrix is equal to 160, but almost as good results are obtained from 80-dimensional matrix with Epstein-Nesbet correction vectors only. Norms of different types of correction vectors in the final wavefunction give an idea about the relative importance of these corrections. For R_e with 0.03 threshold for selection of reference configurations by CIPSI 26 configurations are selected, giving the following contributions to the norm of the wavefunction:

$$|\Phi_{HF}|^2 = 0.9517;\ |\Phi_{Ref}|^2 = 0.0287;$$
$$|\Phi_{EN}|^2 = 0.0194;\ |\Phi_{MP}|^2 = 0.0002;\ |\Phi_D|^2 = 0.0001$$

and the contributions to correlation energy: -78.485 mH from the reference space, -55.851 mH from Epstein-Nesbet correction vectors, -0.475 mH from Møller-Plesset vectors and -0.265 mH from Φ_D vectors. For $2R_e$ stretched bonds 39 CSFs are selected with 0.03 threshold, giving the norms:

$$|\Phi_{HF}|^2 = 0.5754;\ |\Phi_{Ref}|^2 = 0.4043;$$
$$|\Phi_{EN}|^2 = 0.0199;\ |\Phi_{MP}|^2 = 0.0002;\ |\Phi_D|^2 = 0.0002$$

and the contributions to correlation energy: -253.457 mH from the reference space, -50.160 mH from Epstein-Nesbet correction vectors, -0.453 mH from Møller-Plesset vectors and -0.409 mH from Φ_D vectors. In this case the weight of the Hartree-Fock solution in the final norm is very low, making serious problems to all single-reference methods. For the $2R_e$ case the selection procedure brings some quadruply excited configurations into the reference set from the CIPSI wavefunction. These configurations are of great importance for the performance of the method.

In Table 2.4 one may see the comparison between CIPSI and Sup-CI results for water in the same DZ basis set, but with all electrons correlated, in the original geometry of Saxe *et.al* [58], used also by Cimiraglia and Persico and by Evangelisti, Daudey and Malrieu [53] in their studies on the convergence of CIPSI algorithms. CIPSI results of this work are given in Table 2.4 and Fig. 2.2 and are similar to the results of [53]. Smaller CIPSI errors in the Epstein-Nesbet partitioning are due to the use of spin eigenfunctions by the SGGA program instead of determinants used by the original CIPSI program. The given here CIPSI energies are obtained with the Møller-Plesset and Epstein-Nesbet partitioning techniques, while Sup-CI energies use both type of correction vectors. For Møller-Plesset partitioning the E_0 single-reference energy has been corrected by the correlation energy obtained by diagonalization of the Hamiltonian in the reference space instead of the baricentric energy [53].

Results presented in Table 2.4 compare the rate of convergence of the perturbative CIPSI method and variational Sup-CI method for growing number of the reference configurations. Good agreement of CIPSI for 0.03 threshold is rather fortuitious: we have verified that even for 0.001 threshold (66 reference functions) the error of the CIPSI method is still around +0.7 mH, with 0.0001 threshold (196 reference functions) it is -0.02 mH but decreasing the threshold to 0.0000125 (751 reference functions) still leads to an error of +0.6 mH. Similar behaviour is observed at other geometries.

Thus behaviour of CIPSI even at the very low threshold levels is not predictable. In contrast to this erratic behaviour much better convergence is obtained with variational Sup-CI procedure

Table 2.4: CIPSI, CASPT2 and Sup-CI results for water at equilibrium

Threshold	0.5	0.05	0.04	0.03	0.02
No. of refs. CIPSI-MP	1	4	7	15	37
Δ E(CIPSI^{MP})	8.551	14.967	15.581	14.385	9.572
Refs. CIPSI-EN & Sup-CI	1	9	13	26	50
CI in the ref. space	148.028	107.282	94.692	71.793	47.695
Δ E(CIPSI^{EN})	-30.650	-6.404	-3.239	1.400	0.355
Δ E(Sup-CI)	12.370	6.377	5.152	3.071	1.463
Δ E(Sup-CI)+Dav. corr.	7.379	3.478	2.940	1.796	0.984
CASPT2 el./orb./refs	4/4/12	6/6/142	8/8/492		
	8.9	9.1	3.1		

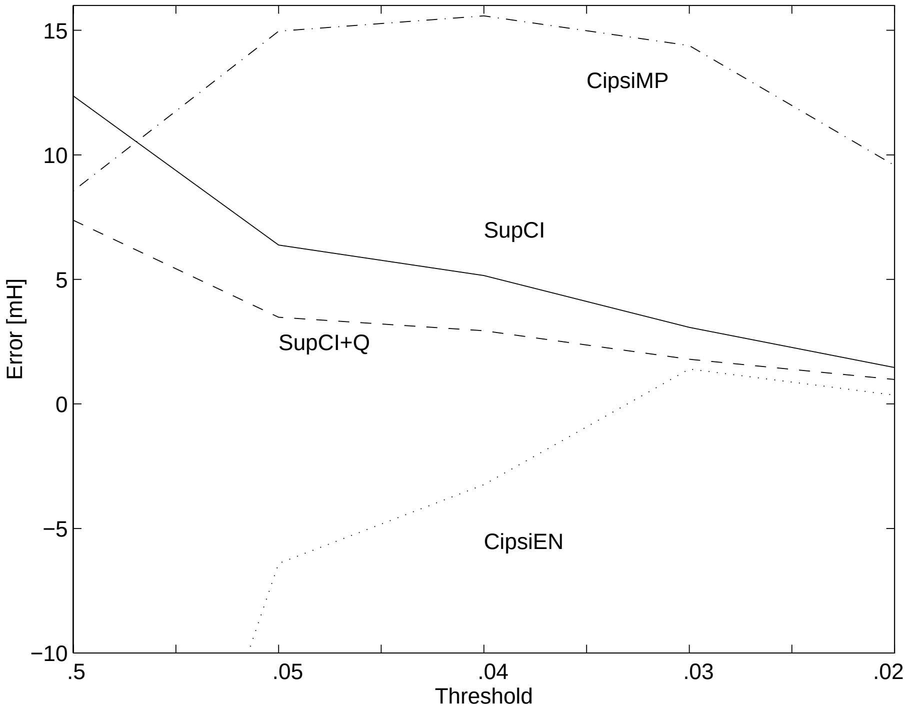
C: CIPSI; S: Superdirect CI
Results from [57]. Geometry from Saxe *et.al* [58]. Energy differences in millihartrees, with respect to the full CI limit of -76.157866. CASPT2 results from [59], number of electrons in active orbitals and number of reference CAS configurations is given.

– it is not *a priori* obvious, because Sup-CI is variational but third-order while CIPSI is non-variational but of the second order. There is no problem in Sup-CI with overshooting the full CI limit (as is the case with CIPSI for Epstein-Nesbet partitioning, in general more realiable of the two partitionings used), because the method is variational.

One could also compare these results with those of the CAS-PT2 method developed recently [59] and similar second-order calculations of McDouall *et.al.* [60]. Three different CAS spaces were selected, corresponding to 12, 142 and 492 reference CSFs. The errors at equilibrium for the first two CAS spaces are around 9 mH and in the largest case, for CAS with 8 electrons in 8 orbitals, are reduced to about 3 mH.

As one can see in Table 2.4 and Figure 2.2 Sup-CI errors are reduced to such level with much smaller number of reference configurations. Numerical selection of reference spaces is much more effective than taking all CAS configurations. Already with 25 numerically selected reference CSFs the same accuracy is achieved using Sup-CI method as with the 492 reference CSFs in the CAS-PT2 method. The variation for different points along the potential surface is also smaller, although RHF orbitals are much worse than the corresponding CAS orbitals. The accuracy of Sup-CI with the CAS orbitals is much better, as we shall also see in case of the NH_2 calculations. In fact there is no reason why CAS-PT2 should not behave in the same way as CIPSI with CAS space as reference.

Figure 2.2: Comparison of convergence of Sup-CI and CIPSI methods



Errors of Sup-CI and CIPSI relative to FCI correlation energy for different thresholds. Results from calculation on water in the DZ basis set and the equilibrium geometry.

2.4.2 Acetylene

Calculation on acetylene was done in STO-3G basis, with carbon $1s$ orbitals frozen. Acetylene is much more difficult case because the presence of the triple bond complicates description already at the equilibrium geometry.

Increasing the equilibrium bond length by half already decreases the coefficient of the SCF function in the normalized first order wave function to 0.9 and to 0.78 in the final CIPSI normalized first order function. The Davidson diagonalization procedure in MR-CI converges very slowly in this case, therefore the superdirect results will also be slowly convergent and rather far from full CI. Using the second-order wavefunction in the superdirect procedure with a single reference (at a cost comparable to the full 5-th order of of perturbation theory) still gives an error of 45 mH and going to third-order function still misses almost 18 mH.

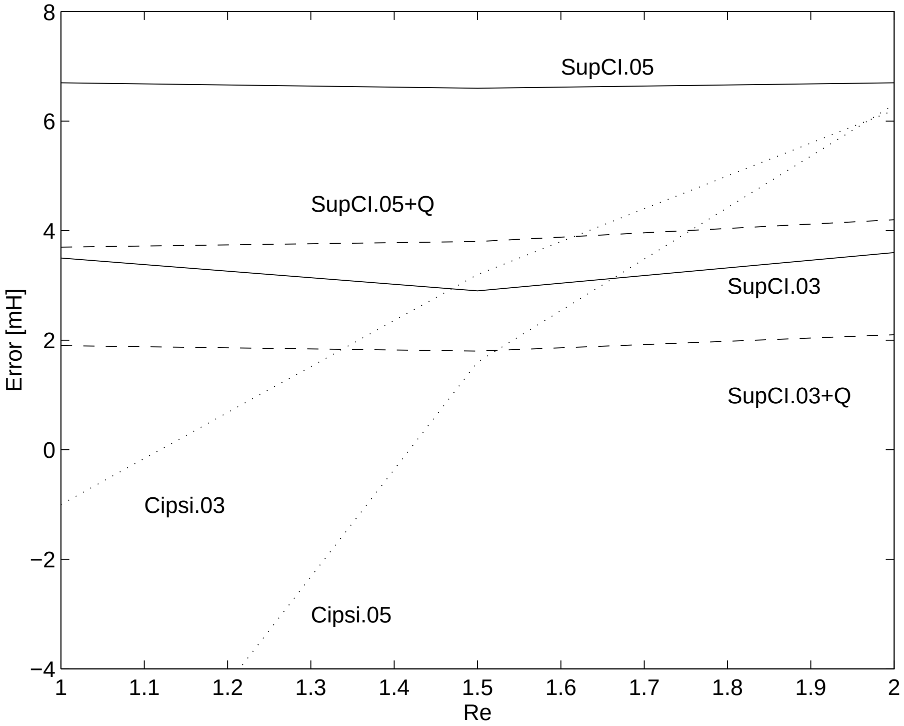
However, in spite of this poor reproduction of the dynamical part of correlation energy by the single reference methods and the failure of CIPSI at thresholds 0.05 to 0.03 to obtain accurate correlation energy Sup-CI results are quite good indeed. The results are closer to the full CI only in the case of CIPSI selection and they remain nearly constant for both geometries, giving an error of 3.9 and 5.8 mH at R_e and $1.5R_e$ geometries respectively (or 1.5 and 2.3 mH with the Davidson correction).

One can find more details on these calculations in ref. [57]. We shall look in the next paragraph at the superdirect CI results for open-shell system.

2.4.3 NH$_2$

Calculations for $^2B_1$ and $^2A_1$ open-shell states of NH_2 will be briefly reported in this subsection. This time the reference benchmark were calculations done by Bauschlicher *et.al* [61]. The same double zeta basis set and geometries have been used (with ROHF energies reproduced within a few μH). For $^2B_1$ state the SCF reference configuration is $1a_1^2 2a_1^2 3a_1^2 1b_1^1 1b_2^2$ at all geometries. For $^2A_1$ state all geometries correspond to the $3a_1 \rightarrow 1b_1$ excitation relative to the $^2B_1$ configuration. The lowest orbital was frozen in these calculations.

Figure 2.3: Results for NH_2 in the DZ basis set, $^2B_1$ state, three geometries



Selection of references by CIPSI. Energy relative to FCI.

All results are summarized in Table 2.5 ($^2B_1$ state) and Table 2.6 ($^2A_1$ state). The superdirect results converge quite well and are remarkably stable, giving at all geometries almost the same errors (see Fig.2.3 and Table 2.5). Such stable behaviour of the Sup-CI results for stretched geometries is rather surprising, because convergence of the Davidson procedure in full CI calculations is very slow, giving for $2R_e$ after 10 iterations an error of 0.27 mH. Results at the 0.03 threshold level are already quite close to CI-SDTQ, and the error after adding Davidson correction is even lower than CI-SDTQ. The coefficient of SCF function in the CI-SDTQ function is in this case 0.71 – it means that the ROHF description is really poor.

Perturbation theory with the standard partitionings (Møller-Plesset and Epstein-Nesbet)

Table 2.5: Results for NH_2, $^2B_1$ state, DZ basis set.

Method	R_e	$1.5R_e$	$2R_e$
FCI	-55.646028	-55.534809	-55.449427
MBPT results			
E2	-12.1	-31.5	-99.4
E3	7.7	32.2	121.5
E4	-2.6	-20.6	-166.0
E5	1.4	14.4	165.2
Variational perturbation results			
E_{s1} MP	19.5	38.2	98.3
E_{s1} EN	8.6	29.9	87.8
E_{s2} MP	19.1	36.4	77.4
E_{s2} EN	0.8	5.4	30.6
E_{s3} MP	17.6	35.5	68.1
E_{s3} EN	0.1	0.9	13.1
Variational results			
Escf	102.2	161.0	264.3
CI(SD)	4.6	16.4	55.1
CI(SDTQ)	0.1	0.5	2.6
Superdirect and CIPSI results			
$S_{0.05}$	6.7	6.6	6.7
+ Davidson corr.	3.7	3.8	4.2
CIPSI 0.05	-8.2	1.6	6.3
$S_{0.03}$	3.5	2.9	3.6
+ Davidson corr.	1.9	1.8	2.1
CIPSI 0.03	-1.0	3.2	6.2

Geometry from Bauschlicher *et.al* [61], $1s$ orbital frozen. Results from [57]. Energy differences in millihartrees, with respect to the FCI energy (except first row). Perturbation theory results in Møller-Plesset partitioning are divergent for all 3 points, and in Epstein-Nesbet partitioning, as given below, diverge only for $2R_e$ case. Modified CIPSI scheme for selection of the reference configurations used.

diverges for open-shell systems, although low-order results have deceptively reasonable values. In a very similar calculation for NH_2, $^2B_1$ state in STO-3G basis set Nobes *et.al* [62] showed that for $2R_e$ the contribution of 25-th order is still larger than 0.1 mH for every tested method i.e. different versions of RHF based MP theory and UHF based MP theory. Results in Table 2.5 show also how dangerous it is to believe in the results of the low-order perturbation theory.

As one can see from Table 2.6 the errors are smaller for $^2A_1$ state, except for $2R_e$ geometry where ROHF becames quite inadequate as a zeroth-order description. The Davidson procedure converges even slower than in the previous case: after 10 iterations the error is still 2 mH. CIPSI results with thresholds 0.05 and 0.03 are also quite poor, indicating that the reference space is still too small. In spite of such slow convergence Sup-CI with the threshold 0.03 gives an error of 4.5 mH (50 references), a significant reduction from 14.2 mH error at 0.05 threshold (16 references).

This example shows that large reference space may compensate even serious inadequacies of the single-particle description. Nevertheless, one should try to improve the method to include more correlation in smaller reference space. An obvious way in this direction is by better selection of the molecular orbital space: using the simplest valence CAS-SCF orbitals one can reduce the error for $2R_e$ geometry, $^2B_1$ state in calculations with the threshold 0.05 from 6.7 mH (4.2 mH

Table 2.6: Results for NH_2, $^2A_1$ state, DZ basis set.

Method	R_e	$1.5R_e$	$2R_e$
FCI	-55.603404	-55.449846	-55.355766
Escf	98.0	138.3	200.7
MBPT results			
E2	-13.6	-26.9	-20.7
E3	8.1	23.6	14.1
E4	-3.1	-14.0	-20.3
E5	1.8	10.1	45.5
Variational perturbation results			
E_{s1} MP	46.9	60.2	132.4
E_{s1} EN	8.1	21.0	49.8
E_{s2} MP	23.3	50.3	108.8
E_{s2} EN	0.8	3.4	19.6
E_{s3} MP	22.4	50.0	105.8
E_{s3} EN	0.1	0.6	8.9
Variational results			
CI(SD)	4.3	12.0	32.6
CI(SDTQ)	0.1	0.4	2.1
Superdirect and CIPSI results			
$S_{0.05}$	4.7	5.8	14.2
+ Davidson corr.	2.3	3.4	11.4
CIPSI 0.05	-3.9	-1.1	14.9
$S_{0.03}$	2.7	2.3	4.5
+ Davidson corr.	1.5	1.5	3.3
CIPSI 0.03	0.9	1.6	8.0

Geometry from Bauschlicher *et.al* [61], 1*s* orbital frozen. Results from [57]. Energy differences in millihartrees, with respect to the FCI energy given in the first row. Modified CIPSI selection scheme is used. Perturbation theory results are given in Epstein-Nesbet partitioning, Møller-Plesset energies are divergent at all geometries.

with correction) to 1.6 mH (0.9 mH), and for the 0.03 threshold the error drops from 3.6 mH (2.1 corrected) to only 1.3 mH (0.7 mH). Indeed, all results presented in this paper are much more accurate if CAS-SCF or simple MC-SCF orbitals are used, but the goal here was to test the Sup-CI method performing the simplest (and still the most common) calculations rather than to present the best results.

2.5 Discussion

The results reported in the previous section are very encouraging. Comparing these results to those of FAST CI approach of Bendazzoli *et.al* [42] we note that adding new types of correction vectors and relaxing restrictions on the types of reference CSFs gives results that are superior to all single reference methods, including CI and many-body perturbation theory. On the other hand the method properly programmed should be an order of magnitude more efficient than multireference CI giving results of similar quality. The work is in progress now on a black-box Sup-CI program that should be as easy to use as perturbation theory. Explicit formulas for matrix elements are rather lengthy and they are given in a compact notation in the appendix. The most efficient approach to their implementation is a matter of investigation.

CIPSI procedure for selection of the reference space applied at the preliminary stage of potential energy calculations will identify all relevant configurations and allow to fix one reference space for all geometry points used, increasing reliability of calculations. Such a combined CIPSI-Sup-CI procedure may be very useful for computations of molecular properties and potential energy surfaces. The calculations on water and NH_2 convince that convergence of the variation-perturbation method in higher orders is superior to the convergence of the perturbation series.

Since the method is variational one may increase the number of first-order correction vectors obtained from different reference CSFs or introduce higher order corrections to improve the results. The latter one seems to be however difficult and one may think of using approximate second-order correction vectors, for example by allowing only pair–pair interactions in the Hamiltonian used to create second-order correction vectors.

Selection of the one-particle basis set in the variation-perturbation method has been discussed by Bendazolli *et.al* [42]. Their conclusion was that MCSCF orbitals are the proper choice. Parametrization of the Hartree-Fock operator shows that the convergence rate of low-order CI methods may be substantially increased. Indeed, preliminary studies of HF method with scaling parameters for Coulomb and exchange integrals (g-Hartree-Fock method and $\lambda\bar{\lambda}$ method [63] suggest, that the canonical HF function may be the worst choice in all cases except for single reference CISD. Although introduction of λ parameters to the HF equations is trivial it is not clear how to predict *a priori* their best value. Orbitals obtained in this way may also lead to erratic convergence of perturbative series, but that should not influence variational Sup-CI results.

The biggest problem that remains in Sup-CI as well as MR-CI is the size-extensivity error due to neglecting of the higher-order excitations. Although *a posteriori* corrections applied to Sup-CI energies remove a part of this error it is desirable to find more accurate corrections for this method.

Chapter 3

Matrix elements in the Sup-CI method

The outline of the SGA approach and an algebra for the Sup-CI matrix elements are presented in the first two sections of this chapter. The algorithm based on these techniques is described in the fourth section and applied then to the derivation of the MBPT-like expressions for the matrix elements occuring in third order Sup-CI. Some of the formulas are explicitly given.

3.1 Symmetric group approach to CI methods

In the direct CI (DCI) methods one avoids storing the CI matrix. The product of this matrix and the vector of CI coefficients, which is formed in any diagonalization procedure directed at diagonalizations of large CI matrices [43] is constructed directly from the list of two electron integrals [32]. It requires however sophisticated mathematical tools to implement such an algorithm in an efficient way. The formal problems appearing in DCI schemes have been solved by Paldus and Shavitt using the Unitary Group Approach (UGA) [33] and by Duch and Karwowski within the Symmetric Group Approach (SGA) [34]. These formal developments may also provide a tool for deriving the MBPT-like formulas i.e. expressions in terms of sums of products of two electron integrals for the matrix elements between contracted functions, appearing in the superdirect CI mode [56].

In the following we shall briefly recall the basic principles of the SGA, which is then used to evaluate matrix elements occurring in the third-order Sup-CI method.

3.1.1 Hamiltonian and its symmetry

The non-relativistic spin-free electronic Hamiltonian (1.3) may be represented in the finite Hilbert space $\mathcal{H}^N$ in the following form [34]

$$\hat{H} = \sum_{ij}^{k}(i|j)\hat{E}_{ij} + \frac{1}{2}\sum_{ijkl}^{k}(ij|kl)(\hat{E}_{ij}\hat{E}_{kl} - \delta_{jk}\hat{E}_{il}) \tag{3.1}$$

The summation indices $i, j, \ldots$ refer to orthonormal orbitals $\{|\phi_i\rangle\}_{i=1}^{k}$, which span the one-particle orbital space $\mathcal{H}_o^1$. The one- and two-electron integrals are defined as follows

$$(i|j) = \langle\phi_i(1)|\hat{h}_1(1)|\phi_j(1)\rangle; \qquad (ij|kl) = \langle\phi_i(1)\langle\phi_k(2)|\hat{h}_2(1,2)|\phi_l(2)\rangle|\phi_j(1)\rangle \tag{3.2}$$

The operators $\hat{E}_{kl}$ (for simplicity of notation we shall often simply write E_{kl}) are called the *replacement operators* since they replace $|\phi_l\rangle$ by $|\phi_k\rangle$ when acting on a product of orbitals

$$E_{kl} = \sum_{i}^{N} |\phi_k(i)\rangle\langle\phi_l(i)| \tag{3.3}$$

They are generators of the unitary group $U(k)$ which is an invariance group of $\mathcal{H}_o^1$. This fact forms a basis for the UGA approach to the CI methods [33]. The replacement operators satisfy the following commutation rule

$$[E_{ij}, E_{kl}] = \delta_{kj}E_{il} - \delta_{il}E_{kj} \tag{3.4}$$

and may be expressed in the second quantization language in terms of anihilation and creation operators $E_{ij} = \sum_\sigma a_{i\sigma}^{+} a_{j\sigma}$.

Since all kinds of spin operators commute with $\hat{H}$ the total spin S and its projection M are good quantum numbers and the eigenfunctions of $\hat{H}$ should be chosen also as eigenfunctions of $\hat{S}^2$ and $\hat{S}_z$. It imposes additional restrictions on the variational coefficients of the CI expansion in terms of determinants i.e. determinantal basis of $\mathcal{H}^N$. They may be taken into account by an explicit expansion in terms of conveniently chosen spin adapted basis. Such a basis consists in general of the linear combinations of a number of Slater determinants corresponding to a given set of orbitals and differing only in their spin parts, referred to as the *configuration state functions* (CSFs) [31]. The CSFs will be denoted $|\lambda; SM, l\rangle$. They fulfill equations

$$P|\lambda; SM, l\rangle = \sigma(P)|\lambda; SM, l\rangle \tag{3.5}$$

$$\hat{S}^2|\lambda; SM, l\rangle = S(S+1)|\lambda; SM, l\rangle \tag{3.6}$$

$$\hat{S}_z|\lambda; SM, l\rangle = M|\lambda; SM, l\rangle \tag{3.7}$$

where P is a permutation operator of electron coordinates and $\sigma(P)$ is its parity. λ stands for a set of the orbital indices used to construct a given CSF or in other words for an orbital configuration. The index $l = 1, \ldots f(S,s)$ distinguishes independent eigenfunctions of $\hat{S}^2$ and $\hat{S}_z$ belonging to the same values of S and M. Their number depends on the number of singly occupied orbitals s in λ and is given by $f(S,s) = (s!(2S+1))/(a!(1-a)!)$; $a = s/2 - S$ [34]. The CI expansion for the kth state characterized by given S and M numbers is then written

$$|\Psi; SM, k\rangle = \sum_{\lambda} \sum_{l=1}^{f(S,s)} C_{\lambda l}^{SMk} |\lambda; SM, l\rangle \tag{3.8}$$

Equivalently we may say that the Schrödinger equation is projected onto a proper spin adapted (S, M)-subspace of the Hilbert space $\mathcal{H}^N$

$$\mathcal{H}^N = \bigoplus_{S=[0,1/2]}^{N/2} \bigoplus_{M=-S}^{S} \mathcal{H}^N(S, M) \tag{3.9}$$

Another basic invariance property of the Hamilton operator considered here is associated with the indistinguishability of electrons. It implies that the Hamiltonian (1.3) is symmetric in electron coordinates and as a result it is invariant with respect to the symmetric group S_N. The antisymmetry of the N-electron fermion functions (1.2) is equivalent to the assumption that they should transform according to the one-dimensional, so-called sign representation of S_N. This is already taken into account in eq. (3.5).

The molecular electronic Hamiltonian (1.3) may be also invariant with respect to a certain point group G if we fix the distribution of the nuclei. Then a symmetry adapted CSF (SACSF) has to transform according to an irreducible representation Γ of the group G. In the following we shall assume that the considered point group is abelian i.e. all irreducible representations are one-dimensional and each CSF corresponding to a given orbital configuration of a proper symmetry spans such a representation. Thus for a given Γ, the set of SACSFs $|\lambda^\Gamma; SM, l\rangle$ is a subset of CSFs $|\lambda; SM, l\rangle$. In general case each SACSF would be rather a linear combination of CSFs [31]. For simplicity we shall further drop the point group symmetry label.

3.1.2 Separation of the spin part

The SGA specific step consists in construction of separate N-particle orbital and spin spaces (before the antisymmetrization is performed) [34]. The total one-particle Hilbert space $\mathcal{H}^1$ is a product of the k-dimensional orbital space and two-dimensional spin space $\mathcal{H}^1_s = \{|\theta_i\rangle\}_{i=1,2}$

$$\mathcal{H}^1 = \mathcal{H}^1_o \otimes \mathcal{H}^1_s \tag{3.10}$$

Then one may construct the N-particle space as

$$\mathcal{H}^N = \hat{\mathcal{A}}[\mathcal{H}^N_o \otimes \mathcal{H}^N_s] = \hat{\mathcal{A}}[(\mathcal{H}^1_o \otimes^N) \otimes (\mathcal{H}^1_s \otimes^N)] \tag{3.11}$$

Equivalently, we shall assume that the N-electron CSFs are of the form

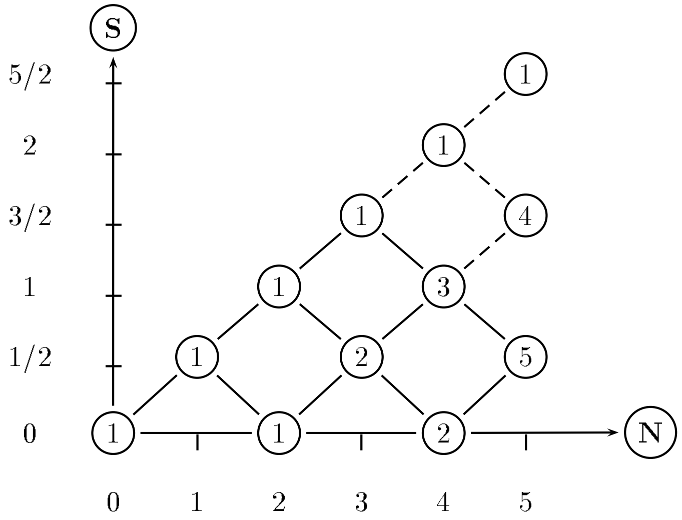
$$|\lambda; SM, l\rangle = \xi_\lambda \hat{\mathcal{A}}[|\lambda\rangle |SM, l\rangle] \tag{3.12}$$

where $|SM, l\rangle \in \mathcal{H}^N_s$ is the pure spin function being an eigenfunction of $\hat{S}^2$ and $\hat{S}_z$ and $|\lambda\rangle$ is a spin independent orbital function defined as a product of N (orthonormalized) orbitals corresponding to the configuration λ, s_λ of them being singly and d_λ doubly occupied. The antisymmetrizer operator and the normalization constant are respectively given by

$$\hat{\mathcal{A}} = \frac{1}{N!} \sum_{P \in S_N} \sigma(P) P; \qquad \xi_\lambda = (N!/2^{d_\lambda})^{1/2} \tag{3.13}$$

In the following we shall assume that the spin functions are geminally antisymmetric spin functions i.e. they are antisymmetric with respect to the transpositions within doubles. In other words the spins of those electron pairs that correspond to doubles are coupled in the N-electron spin functions to two-electron singlets, insuring the antisymmetry of the resulting CSFs [34]. For the remaining s_λ singly occupied orbitals one has $f(S, s)$ different coupling schemes i.e. with each orbital configuration λ there is associated a vector of spin functions $|\mathbf{SM}\rangle$ with the components $|SM, l\rangle$, $l = 1, \dots f(S, s)$.

Figure 3.1: Branching diagram for N=5, solid lines: S=1/2



All the spin functions that span the N-particle spin space or the different coupling schemes within a given set of spins may be characterized using the so-called branching diagram [34]. Each path on this diagram, passing from the $(0,0)$ node to a given (N,S) node uniquely represents certain spin function $|SM,l\rangle \in \mathcal{H}_s^N$, $l = 1,\dots f(S,N)$. The projection of the spin M does not influence its shape. The numbers in the nodes indicate a number of possible paths (spin functions) from the origin $(0,0)$ to a given node. An example of the branching diagram for $N = 5$ electrons is displayed in the figure 3.1.

It should be noted that the spin functions coupled according to the branching diagram are not the geminally antisymmetric functions. To get the latter ones one may use the so-called reversed branching diagram, introduced by Duch [34].

Let us consider now the permutational symmetry of the pure spin functions. For each $P \in S_N$ the N-electron spin space $\mathcal{H}_s^N$ is closed under its action. More specifically, since the $\hat{S}^2$ and $\hat{S}_z$ operators are symmetric in the coordinates of the electrons they commute with an arbitrary permutation. Thus for each $P \in S_N$ the new function $P|SM,l\rangle$ is an eigenfunction of $\hat{S}^2$ and $\hat{S}_z$ with unchanged eigenvalues $S(S+1)$ and M. It is in general a linear combination of the spin functions.

Thus we may use the permutational symmetry and irreducible representations of S_N to classify also the spin functions and for further factorization of matrix representations of the spin-free Hamiltonian. In fact such a factorization, as indicated in eq. (3.9) follows from the corresponding decomposition of the spin space.

A linear space that carries a (irreducible) representation of a group G is called a (irreducible) G-module. The $\mathcal{H}_s^N$ space is a reducible S_N-module. It may be decomposed into a direct sum of irreducible S_N-modules $\mathcal{H}_s^N(S,M)$ characterized by the total spin and the spin projection quantum numbers

$$\mathcal{H}_s^N = \bigoplus_{S=[0,1/2]}^{N/2} \bigoplus_{M=-S}^{S} \mathcal{H}_s^N(S,M) \tag{3.14}$$

Alternatively, spin functions belonging to a pair of S, M $(S \geq |M|)$ quantum numbers form a basis for an irreducible representation of the symmetric group S_N or more precisely of an irreducible S_N-module $\mathcal{H}_s^N(S,M)$. The irreducible representation that is carried by this module is associated with the matrix representation i.e. the set of matrices $\{\mathbf{U}_S^N(P);\ P \in S_N\}$ where

$$U_S^N(P)_{kl} = \sigma(P)\langle SM,k|P|SM,l\rangle \qquad k,l = 1,\ldots f(S,N) \tag{3.15}$$

It may be easily proved [36] that these representation matrices are M-independent. For orthogonal spin functions these matrices are also unitary.

3.1.3 Hamiltonian matrix elements

Due to the separation of the spin and orbital parts in the CSFs one may perform separate integration over spin and orbital variables in the resulting Hamiltonian matrix elements. For the future use it is convenient to consider matrix elements of the powers of the Hamiltonian $\hat{H}^p$

$$H^p_{k(\lambda)l(\mu)} = \langle\lambda; SM,k|\hat{H}^p|\mu; SM,l\rangle = \xi_\lambda\xi_\mu\langle SM,k|\langle\lambda|\hat{\mathcal{A}}^\dagger\hat{H}^p\hat{\mathcal{A}}|\mu\rangle|SM,l\rangle =$$

$$\xi_\lambda\xi_\mu/(N!)\sum_{P\in S_N}\sigma(P)\langle SM,k|P|SM,l\rangle\langle\lambda|\hat{H}^pP|\mu\rangle \tag{3.16}$$

where we have used the fact that $\hat{\mathcal{A}}^\dagger = \hat{\mathcal{A}}$ and $\hat{\mathcal{A}}^2 = 1/(N!)\hat{\mathcal{A}}$.

The last formula may be expressed in a somewhat different manner, following the proposal of Karwowski [35]. Let Π_λ be an invariance group of $|\lambda\rangle$ generated by the transpositions within doubles in λ and let D_λ stands for its dimension. Then for two orbital configurations λ and μ and a given permutation P one may define a subgroup of S_N called the double coset $\Pi_\lambda P\Pi_\mu$. Every group may be decomposed into a number distinct double cosets having no elements in common.

The antisymmetrizer may be also decomposed into the double cosets contributions [34]

$$\hat{\mathcal{A}} = \frac{1}{N!}\sum_q \sigma(P_q)\frac{D^q_{\lambda\mu}}{D_\lambda D_\mu}\sum_{\nu\in\Pi_\lambda;\,\tau\in\Pi_\mu}\nu P_q\tau \tag{3.17}$$

where the summation runs over distinct double cosets $\Pi_\lambda P_q\Pi_\mu$ of the dimensions $D^q_{\lambda\mu}$, generated by P_q . Moreover one may replace the permutations acting on the electron coordinates in the orbital integrals by their hermitian conjugate i.e. the same permutations but acting on the orbital indices. Assuming that spins associated with the doubly occupied orbitals are coupled to singlets (Singlet Coupled Pairs - SCP) and the spin functions are generated according to the reversed branching diagram, the full representation matrices may be replaced by the small rectangular blocks $[\mathbf{U}_S^N(P_q)]^{fg}$ defined by the sets of singly occupied orbitals in $|\lambda\rangle$ and $|\mu\rangle$ respectively.

Finally one gets in the matrix block notation for $k = 1, \dots f = f(S, s_\lambda)$ and $l = 1, \dots g = f(S, s_\mu)$

$$\mathbf{H}^p_{(\lambda)(\mu)} = 2^{-(d_\lambda + d_\mu)/2} \sum_q D^q_{\lambda\mu} [\mathbf{U}^N_S(P_q)]^{fg} \langle P_q\lambda | \hat{H}^p | \mu \rangle \tag{3.18}$$

As have been proved by Kotani *et. al.* [36] further (and essential) simplification is possible since all $[\mathbf{U}^N_S(P_q)]^{ff}$ blocks may be reduced to $\mathbf{U}^s_S$ matrices, corresponding to appropriate permutations of singles only. The final numerical factor appearing in eq. (3.18) takes simple form and may be tabulated for several different cases and in practice absorbed in the definition of representation matrices [34].

From the practical point of view it is crucial that only at most one of the permutations P_q in eq. (3.18), called line-up permutation leads to complete coincidence between $|P_q\lambda\rangle$ and each chain (product) of replacement operators $\Upsilon|\mu\rangle$ occurring in $\hat{H}^p$, giving non-zero contribution in terms of products of electron integrals. For such a permutation one has to find then the proper spin integral i.e. relevant matrix $[\mathbf{U}^{s_{max}}_S(P_q)]^{fg}$, $s_{max} = max(s_\lambda, s_\mu)$. Several very efficient and suitable procedures for evaluating the relevant matrices of the representations of the symmetric group have been proposed by Duch and Karwowski [34].

3.2 Circular operators

Let us introduce for further considerations the following convention distinguishing the occupation of the one-particle functions (orbitals) in the N-particle functions by the respective indices:

arbitrary $n_{index} = 0, 1, 2$	unoccupied $n_{index} = 0$	singly occupied $n_{index} = 1$	doubly occupied $n_{index} = 2$
e, f, g, h	a, b, c, d	s, t, u, v	i, j, k, l
m, n, p, q		w, x, y, z	

Let $\Upsilon_{efgh\cdots pq}$ denote the product (chain) of replacement operators

$$\Upsilon_{efgh\cdots mnpq} = \hat{E}_{ef}\hat{E}_{gh} \cdots \hat{E}_{mn}\hat{E}_{pq} \tag{3.19}$$

and let Υ denote a chain with non-specified indices. If the subsequent orbitals we excite from are the same as those we excite to in the nearest right hand side neighbor, regarding additionally the first and the last index as neighbors, we shall call such a chain of replacement operators a *circular* operator and denote $\mathcal{E}_{efh\cdots mnq}$

$$\mathcal{E}_{efh\cdots mne} = \Upsilon_{effh\cdots mnne} = \hat{E}_{ef}\hat{E}_{fh} \cdots \hat{E}_{mn}\hat{E}_{ne} \tag{3.20}$$

Notice that each index must appear even number of times in Υ if it is to be a circular operator. The sums over all indices of circular operators are known as the Casimir invariants of the $U(k)$ group. The first Casimir invariant $\sum_e \hat{E}_{ee}$ gives N acting on any N-electron orbital function.

Consider now matrix elements of the type

$$\Upsilon^{(\lambda)(\mu)} = \gamma [\mathbf{U}^N_S(P_q)]^{fg} \langle P_q\lambda | \Upsilon | \mu \rangle \tag{3.21}$$

where γ is an appropriate product of one- or two-electron integrals associated with a chain Υ occuring in the $\mathbf{H}^p_{(\lambda)(\mu)}$. The matrix elements of the powers of the Hamiltonian (3.18) are in general sums of a number of such terms. The necessary condition for non-vanishing of the $\Upsilon^{(\lambda)(\mu)}$ is the same occupation scheme in $|\lambda\rangle$ and $\Upsilon|\mu\rangle$ i.e. any of the (singly or doubly occupied) orbitals appearing in $|P_q\lambda\rangle$ must also appear (as singly or doubly occupied respectively) in $\Upsilon|\mu\rangle$. If the occupation schemes are the same one may choose a permutation P_q such that it leads to complete coincidence between $|P_q\lambda\rangle$ and $\Upsilon|\mu\rangle$.

Let us restrict now to the diagonal matrix elements. If $|\lambda\rangle = |\mu\rangle$ we shall use $|0\rangle$ to denote the total, orbital or spin N-electron function (or their set associated with given S, M numbers) depending on the context. Thus we shall write

$$\langle 0|\hat{H}^p|0\rangle = \sum_{\Upsilon} \gamma(\Upsilon)\langle 0|P(\Upsilon)|0\rangle\langle P(\Upsilon)0|\Upsilon 0\rangle = \sum_{\Upsilon'} \gamma(\Upsilon)\langle 0|P(\Upsilon)|0\rangle \tag{3.22}$$

where all the numerical factors as well as the sign factors are absorbed either in γ or in the spin integrals and $\sum_{\Upsilon'}$ denotes summation with Υ giving vanishing contributions $\langle P(\Upsilon)0|\Upsilon 0\rangle = 0$ excluded. According to our previous considerations $\langle 0|P(\Upsilon)|0\rangle$ means square matrix block. The above equation is a good starting point for the derivation of the matrix elements occurring in the Sup-CI method.

Let us notice that the summation over distinct chains of replacement operators giving non-vanishing contributions to the matrix element in eq. (3.22) may be expressed as a sum over circular operators only. Indeed, if Υ is to give a non-zero contribution it cannot change the occupation of any orbital in $|0\rangle$ i.e. each index m must appear $2i$ times in Υ, i times as left index E_{mp} and i times as right one E_{qm}. It means that commuting the replacement operators one may end up with the circular chain. The additional terms that may appear by applying the commutation rule (3.4) are of two types

$$\cdots \hat{E}_{ef}\hat{E}_{he}\cdots \longrightarrow -\cdots \hat{E}_{hf}\cdots; \qquad \cdots \hat{E}_{ef}\hat{E}_{fh}\cdots \longrightarrow \cdots \hat{E}_{eh}\cdots \tag{3.23}$$

and may be then also permuted to the circular form. Thus we may write

$$\langle 0|\hat{H}^p|0\rangle = \sum_{\Upsilon'} \gamma(\Upsilon) \sum_{\mathcal{E}(\Upsilon)} \langle 0|P(\mathcal{E}(\Upsilon))|0\rangle \tag{3.24}$$

where the second sum runs over all circular operators arising from a given chain Υ. This will be our general strategy: to specify all non-vanishing chains of operators in the orbital integral and then transform them to sums of circular chains. Now, we must learn what permutation is generated by a given circular chain. Then in the final step a proper spin integral may be evaluated.

Some general properties of the circular operators may be directly derived from the commutation relations (3.4). As noticed by Duch [56] they provide a tool for an alternative and convenient reduction of the diagonal spin integrals to the occupation numbers or to the permutations associated with a chain of operators involving singly occupied orbitals only.

Let us first distinguish *closed* circular operators involving any orbital index twice only i.e. through a single pair: excited from, excited to (when expressing the circular operator explicitly in terms of replacement operators). An arbitrary circular operator may be expressed in terms of closed circular operators only - as a product of closed circular operators or simply product of closed chains or a sum of a number of such products:

$$\mathcal{E}_{efhe\cdots mne} = \mathcal{E}_{efhe}\mathcal{E}_{e\cdots mne}; \qquad \mathcal{E}_{efh\cdots fg\cdots mne} = \mathcal{E}_{fhe\cdots f}\mathcal{E}_{efg\cdots mne} + \mathcal{E}_{eh\cdots f\cdots mne} \tag{3.25}$$

For the closed chains the following general rule is valid

$$\mathcal{E}_{ef\cdots m\cdots ne} = \mathcal{E}_{m\cdots nef\cdots m} + \mathcal{E}_{ef\cdots ne} - \mathcal{E}_{m\cdots ne\cdots m} \tag{3.26}$$

In the simplest case one gets for instance

$$\mathcal{E}_{efe} = \mathcal{E}_{fef} + \mathcal{E}_{ee} - \mathcal{E}_{ff} = \mathcal{E}_{fef} + \hat{n}_e - \hat{n}_f \tag{3.27}$$

where $\hat{n}_e = 0, 1, 2$ denotes the occupation number operator for the orbital $|\psi_e\rangle$ with values $n_e = 0, 1, 2$. Since excitation from unoccupied or to doubly occupied orbitals are forbidden we have also in general

$$\mathcal{E}_{a\cdots a}|0\rangle = \mathcal{E}_{\cdots i\cdots}|0\rangle = 0 \tag{3.28}$$

where indices a and i refer to unoccupied and doubly occupied orbitals in $|0\rangle$, respectively. Notice that $|0\rangle$ may be here either pure orbital function or the total function.

Using eqs. (3.26) and (3.28) one then easily finds for the matrix elements that the unoccupied orbital indices may always be simply removed

$$\langle 0|\mathcal{E}_{e\cdots abc\cdots e}|0\rangle = \langle 0|\mathcal{E}_{e\cdots e}|0\rangle \tag{3.29}$$

whereas the doubly occupied indices always contribute through the terms of the type $\langle 0|\mathcal{E}_{ie\cdots fi}|0\rangle$ only and may also be eliminated

$$\langle 0|\mathcal{E}_{iai}|0\rangle = n_i = 2; \qquad \langle 0|\mathcal{E}_{isi}|0\rangle = n_i - n_s = 1 \tag{3.30}$$

$$\langle 0|\mathcal{E}_{isti}|0\rangle = \langle 0|\mathcal{E}_{stis} + \mathcal{E}_{iti} - \mathcal{E}_{sts}|0\rangle = 1 - \langle 0|\mathcal{E}_{sts}|0\rangle \tag{3.31}$$

$$\cdots$$

In this way all expressions for the diagonal matrix elements of the closed chains may be reduced to the matrix elements of the chains involving singly occupied indices only.

When eliminating the indices of the doubly occupied orbitals we had to evaluate the spin integrals appearing in $\langle 0|\mathcal{E}_{ii}|0\rangle$ and $\langle 0|\mathcal{E}_{ss}|0\rangle$. Since such chains introduce identity permutation only they always give spin integral equal to one (as many times as a given orbital occurs in $|0\rangle$). The non-trivial spin integrals will appear when considering the chains $\langle 0|\mathcal{E}_{st\cdots s}|0\rangle$. We have to describe the permutations associated with such circular operators.

Acting on the orbital product $\cdots|\psi_s\rangle\cdots|\psi_t\rangle\cdots$ with $\mathcal{E}_{sts}$ one may verify that [56] this operator is equivalent to:

$$\mathcal{E}_{sts} = 1 + (s, t) \tag{3.32}$$

where (s, t) denotes the transposition of sth and tth orbitals. For longer chains:

$$\mathcal{E}_{stus} = (1 + (s, t))(1 + (s, u)) \tag{3.33}$$

$$\mathcal{E}_{stuws} = (1 + (s,t))(1 + (s,u))(1 + (s,w)) \tag{3.34}$$

Let us then denote by $\langle s,t\rangle$, $\langle s,t,u\rangle$ etc. the appropriate spin integrals

$$\langle s,t\rangle = \langle 0|(s,t)|0\rangle \tag{3.35}$$

$$\langle s,t,u\rangle = \langle 0|(s,t)(s,u)|0\rangle = \langle 0|(s,t,u)|0\rangle = \langle s,t\rangle\langle s,u\rangle \tag{3.36}$$

where now $|0\rangle$ stands for the set of spin functions and $\langle s,t\rangle\langle s,u\rangle$ is a product of two (representation) matrices. Thus finally we obtain

$$\langle 0|\mathcal{E}_{sts}|0\rangle = 1 + \langle s,t\rangle; \qquad \langle 0|\mathcal{E}_{stus}|0\rangle = (1 + \langle s,t\rangle)(1 + \langle s,u\rangle) = 1 + \langle s,t\rangle + \langle s,u\rangle + \langle s,t,u\rangle \tag{3.37}$$

and in general

$$\langle 0|\mathcal{E}_{stu\cdots ws}|0\rangle = (1 + \langle s,t\rangle)(1 + \langle s,u\rangle)\cdots(1 + \langle s,w\rangle) \tag{3.38}$$

i.e. the diagonal elements of the circular operators with more than two different indices may always be expressed through the elements of operators with lower number of indices.

The basic integrals $\langle s,t\rangle$ corresponding to the transpositions of spins numbers s and t are equal to ±1 in the determinantal basis. In case of spin adapted basis they are appropriate representation matrices of the transpositions (s,t) and are easily calculated from the branching diagram [34]. Now, we are ready to apply the above derived machinery to the calculation of the matrix elements occurring in the Sup-CI method with the arbitrarily complicated reference functions in the spin adapted basis.

3.3 Matrix elements in the third order Sup-CI

In the following all types of the relevant matrix elements will be specified. In the multireference third order Sup-CI method we need formulas up to the third order i.e. matrix elements of the type:

$$H_{ik,lj} = \langle w_i|\hat{R}_k\hat{H}\hat{R}_l|w_j\rangle; \qquad S_{ik,lj} = \langle w_i|\hat{R}_k\hat{R}_l|w_j\rangle \tag{3.39}$$

where indices i,j mean different reference (open shell) functions and k,l mean different 'resolvent' operators, which may take one of the following forms

$$\hat{R}_I = \hat{\mathbf{1}}; \qquad \hat{R}_K = \hat{Q}_0\hat{H}; \qquad \hat{R}_X = \hat{D}_X\hat{Q}_0\hat{H} \tag{3.40}$$

where in the latter one ($X = N$ or $X = M$ where N refers to the Epstein–Nesbet scheme and M refers to the Møller–Plesset scheme) the respective perturbative denominators appear

$$\hat{D}_N = (\hat{H}_d - E_{SCF})^{-1}; \qquad \hat{D}_M = (\hat{H}_0 - E_0)^{-1} \tag{3.41}$$

From now we shall use $|0\rangle$ and $|0'\rangle$ to distinguish different reference functions and R_1, R_2 for different resolvent operators (except the identity operator $\hat{R}_I$).

Those elements that contain Hamiltonian in the first power are the usual CI matrix elements derived in SGA formalism by Duch and Karwowski [34]. We shall distinguish 'diagonal' type of the matrix elements considered, for which $|0\rangle = |0'\rangle$ but the resolvent operators (i.e. denominators) may be different. The non-diagonal type is defined by $|0\rangle \neq |0'\rangle$. In the diagonal case

H_{0R_1,R_20} with the second power of $\hat{H}$ (i.e. when $R_1 = I$, $R_2 \neq I$) one gets PT2-like or k^4-like (i.e. involving fourth-fold summations) formula and with the third power ($R_1 \neq I$, $R_2 \neq I$)the third order or k^6-like formula. In the non-diagonal case the complexity of the evaluation of those matrix elements is certainly lower since more indices in the summation are fixed.

Table 3.1: Matrix elements types in the third order Sup-CI

compl.	ordinary CI type	non-diagonal	diagonal
$< n^4$	$H_{0I,I0}$; $H_{0I,I0'}$ S_{0I,R_10}; $S_{0I,R_10'}$	$H_{0I,R_10'}$; $S_{0R_1,R_20'}$ $H_{0R_1,R_20'(3,4,5,6)}$	
n^4		$H_{0R_1,R_20'(2)}$	H_{0I,R_10}; S_{0R_1,R_20}
n^6			H_{0R_1,R_20}

For the future use it is useful to distinguish more precisely different categories of the matrix elements of the general form (3.39). They are gathered in Table 3.1. The non-diagonal type matrix elements are specified (if useful) by the relative excitations of the ket function with respect to the bra function e.g. $0'(5,6)$ means all reference functions $|0'\rangle$ which differ by 5 or 6 orbitals with respect to $|0\rangle$ (quintuply or sextuply excited configurations).

3.4 Algorithm and formulae

The formula for the diagonal matrix elements of the second order type i.e. H_{0I,R_10} (thus S_{0R_1,R_20} as well) has been derived by Duch [56], using techniques presented in the previous subsections. This formula does not include some simple one-particle terms, but they may be easily added. We shall derive now the third order diagonal formula $\langle 0|\hat{R}_1\hat{H}\hat{R}_2|0\rangle$. Notice that if $\hat{R}_1 = \hat{R}_2 = \hat{R}_M$ we get the core of E_3 MBPT formula. Then, the non-diagonal matrix elements will be considered.

3.4.1 Contractions

In the following we shall divide the set of arbitrarily occupied indices into two subsets - of those that we may excite from ($n_{index} = 2, 1$): e, f, m, n; and of those we may excite to ($n_{index} = 1, 0$): g, h, p, q. Let r stands for an arbitrary orbital index. Furthermore, since we shall explicitly differentiate the possible relative occupation schemes in all intermediate projections it should be recalled [34] that one may extract from the definition of the Hamiltonian those terms which do not vanish between CSFs differing by a certain number of orbital indices. Denoting by $\hat{B}_0, \hat{B}_1^{eg}, \hat{B}_2^{efgh}$ the parts of $\hat{H}$ (including the one-particle terms) connecting states differing on 0,1,2 orbitals ($\langle 0|\cdots|0\rangle$; $|^g_e\rangle$; $|^{gh}_{ef}\rangle$) respectively, one gets [34]

$$\hat{B}_0 = \frac{1}{2}\sum_e (ee|ee)\hat{n}_e(\hat{n}_e - 1) + \sum_{e<f}[(ee|ff)\hat{n}_e\hat{n}_f + (ef|ef)(\hat{E}_{ef}\hat{E}_{fe} - \hat{n}_e)] + \sum_e (e|e)\hat{n}_e \quad (3.42)$$

$$\hat{B}_1^{eg} = \sum_r (eg|rr)(\hat{n}_r - \delta_{gr})\hat{E}_{eg} + \sum_{r\neq e,g} (er|gr)\hat{E}_{rg}\hat{E}_{er} + (e|g)\hat{E}_{eg} \tag{3.43}$$

$$\hat{B}_2^{efgh} = 2^{-\delta_{ef}\delta_{gh}}(eg|fh)\hat{E}_{eg}\hat{E}_{fh} + (1-\delta_{ef})(1-\delta_{gh})(eh|fg)E_{eh}E_{fg} \tag{3.44}$$

Let us now consider the most complicated matrix element of the diagonal, n^6 type

$$\langle 0|\hat{R}_1\hat{H}\hat{R}_2|0\rangle = \sum_{L\neq 0}\sum_{K\neq 0} D_1(L)D_2(K)\langle 0|\hat{H}|L\rangle\langle L|\hat{H}|K\rangle\langle K|\hat{H}|0\rangle \tag{3.45}$$

where L, K denote N-electron basis function (CSFs) and are subject to the conditions $L \in SD(0)$; $K \in SD(0) \cap SD(L)$ i.e. are at most doubly excited with respect to $|0\rangle$ and $|L\rangle$ respectively. $D_i(L)$ are appropriate denominators.

Dividing the intermediate sums into contributions coming from the singly and doubly excited configurations one may rewrite the r.h.s. of the eq. (3.45) as

$$\sum_{L\in D(0)} D_1(L)\langle 0|\hat{B}_2|L\rangle \; \{ \sum_{K\in D(0)\cap D(L)} D_2(K)\langle L|\hat{B}_2|K\rangle\langle K|\hat{B}_2|0\rangle +$$

$$\sum_{K\in D(0)\cap S(L)} D_2(K)\langle L|\hat{B}_1|K\rangle\langle K|\hat{B}_2|0\rangle + D_2(L)\langle L|\hat{B}_0|L\rangle\langle L|\hat{B}_2|0\rangle \; \}$$

For simplicity contributions from the singles with respect to $|0\rangle$ are for a moment omitted. The most complicated term in the above expression is

$$S1 = \sum_{L\in D(0)}\sum_{K\in D(0)\cap D(L)} D_2(K)D_1(L)\langle 0|\hat{B}_2|L\rangle\langle L|\hat{B}_2|K\rangle\langle K|\hat{B}_2|0\rangle \tag{3.46}$$

or in the more explicit notation

$$S1 = \sum_{m\leq n}\sum_{p\leq q}\sum_{e\leq f}\sum_{g\leq h} D_1^{mnpq}D_2^{efgh}\langle 0|\hat{B}_2^{mnpq}|{}^{pq}_{mn}\rangle\langle{}^{pq}_{mn}|\hat{B}_2^{uvwz}|{}^{gh}_{ef}\rangle\langle{}^{gh}_{ef}|\hat{B}_2^{ghef}|0\rangle \tag{3.47}$$

Because the non-vanishing contributions come from the double excitations only $|{}^{gh}_{ef}\rangle \in D(L)$, $L = |{}^{pq}_{mn}\rangle$, among indices e, f, g, h one pair only may differ from the indices m, n, p, q. This gives six possible contractions of eight orbital indices, reducing the sums to 6-fold only:

$$\begin{array}{lcl}
e = m, \;\; f = n & & p, q, g, h \\
g = p, \;\; h = q & & e, f, m, n \\
e = m, \;\; g = p & \text{for which } u, v, w, z \text{ are respectively} & f, q, n, h \\
e = m, \;\; h = q & & f, p, n, g \\
f = n, \;\; g = p & & e, q, m, h \\
f = n, \;\; h = q & & e, p, m, g
\end{array}$$

Finally, $S1$ may be rewritten as

$$S1 = \sum_{m\leq n}\sum_{p\leq q} D_1^{mnpq}\langle 0|\hat{B}_2^{mnpq}|{}^{pq}_{mn}\rangle\langle{}^{pq}_{mn}|$$

$$
\begin{aligned}
\{ \ & \sum_{g\le h} D_2^{mngh} |\hat{B}_2^{pqgh}|{}^{gh}_{mn}\rangle\langle{}^{gh}_{mn}|\hat{B}_2^{ghmn}|0\rangle + \sum_{e\le f} D_2^{efmn}|\hat{B}_2^{efmn}|{}^{pq}_{ef}\rangle\langle{}^{pq}_{ef}|\hat{B}_2^{pqef}|0\rangle + \\
& + \sum_e \sum_g [\ D_2^{mepg}|\hat{B}_2^{eqng}|{}^{pg}_{me}\rangle\langle{}^{pg}_{me}|\hat{B}_2^{pgme}|0\rangle + D_2^{enpg}|\hat{B}_2^{eqmg}|{}^{pg}_{en}\rangle\langle{}^{pg}_{en}|\hat{B}_2^{pgen}|0\rangle + \\
& \quad + D_2^{megq}|\hat{B}_2^{epng}|{}^{gq}_{me}\rangle\langle{}^{gq}_{me}|\hat{B}_2^{gqme}|0\rangle + D_2^{engq}|\hat{B}_2^{epmg}|{}^{gq}_{en}\rangle\langle{}^{gq}_{en}|\hat{B}_2^{gqen}|0\rangle\ ]\ \}
\end{aligned}
\tag{3.48}
$$

The third part of the above expression has been obtained due to the exchange of the summation indices. Notice that the summations over e and g are not completely free – they cannot be the same as their counterparts in the first intermediate configuration $|{}^{pq}_{mn}\rangle$ e.g. D_2^{mepg} implies that $e \neq n$, $g \neq q$ whereas D_2^{engq} implies that $e \neq m$, $g \neq p$. On the other hand e may be equal to n ($n \neq m$) and/or g may be equal to q ($q \neq p$) in the latter case, implying that the overall complexity of such terms is lower than n^6 (although they still involve two doubly excited configurations being also doubly excited with respect to another). In the next section we shall extract all the terms with some of the summation indices occuring many repetetively. Recall also that any orbital configuration $|{}^{gq}_{en}\rangle$ represents a number of CSFs with different spin couplings.

Table 3.2: H^3 diagonal matrix element – CSC

For explanation of symbols see text below.

$< 0|R_1 H R_2|0 > =$

d1(m,p) d2(m,p) $<$ | b1(m,p) b0(m,p) b1(p,m) | $>$ +

d1(m,p) [d2(e,p) $<$ | b1(m,p) b1(e,m) b1(p,e) | $>$

+ d2(m,g) $<$ | b1(m,p) b1(p,g) b1(g,m) | $>$] +

d1(m,p) d2(m,f,p,h) $<$ | b1(m,p) b1(f,h) b2(p,h,m,f) | $>$

+ d1(m,p) d2(e,g) $<$ | b1(m,p) b2(p,e,m,g) b1(g,e) | $>$ +

d1(m,p) [d2(e,f,p,h) $<$ | b1(m,p) b2(e,f,m,h) b2(p,h,e,f) | $>$

+ d2(m,f,g,h) $<$ | b1(m,p) b2(p,f,g,h) b2(g,h,m,f) | $>$] +

d1(m,n,p,q) d2(m,n,p,q) $<$ | b2(m,n,p,q) b0(m,n,p,q)b2(p,q,m,n) | $>$

+ d1(m,n,p,q) d2(n,q) $<$ | b2(m,n,p,q) b1(p,m) b1(q,n) | $>$

+ d1(m,n,p,q) [d2(e,n,p,q) $<$ | b2(m,n,p,q) b1(e,m) b2(p,q,e,n) | $>$

+ d2(m,n,g,q) $<$ | b2(m,n,p,q) b1(p,g) b2(g,q,m,n) | $>$] +

d1(m,n,p,q) [d2(m,g) $<$ | b2(m,n,p,q) b2(p,q,n,g) b1(g,m) | $>$

+ d2(e,p) $<$ | b2(m,n,p,q) b2(q,e,m,n) b1(p,e) | $>$] +

d1(m,n,p,q) [d2(e,f,p,q) $<$ | b2(m,n,p,q) b2(e,f,m,n) b2(p,q,e,f) | $>$

+ d2(m,n,g,h) $<$ | b2(m,n,p,q) b2(p,q,g,h) b2(g,h,m,n) | $>$

+ d2(e,n,g,q) $<$ | b2(m,n,p,q) b2(p,e,m,g) b2(g,q,e,n) | $>$]

The complete expression for $< 0|R_1 H R_2|0 >$ with contributions of the singles and all possible contraction schemes is given in Table 3.2. A compact summation convention (CSC) is used in this table and in all automatically generated expressions i.e. generated by means of the symbolic algebra program written for this purpose in the Maple language. According to this convention all sums are associated with the ordered orbital indices occuring as arguments of the denominators functions. We shall use a simplified notation: $D_i^{mp} \equiv di(m,p)$; $\hat{B}_1^{mp} \equiv b1(m,p)$; $D_i^{mnpq} \equiv di(m,n,p,q)$; $\hat{B}_2^{mnpq} \equiv b2(m,n,p,q)$. Thus for example

$$d1(m,p)d2(m,p)\cdots \longrightarrow \sum_m \sum_p d1(m,p)d2(m,p)\cdots \tag{3.49}$$

$$d1(m,p)d2(e,f,p,h)\cdots \longrightarrow \sum_m \sum_p \sum_{e\leq f} \sum_h d1(m,p)d2(e,f,p,h)\cdots \tag{3.50}$$

$$d1(m,n,p,q)d2(e,f,p,q)\cdots \longrightarrow \sum_{m\leq n} \sum_{p\leq q} \sum_{e\leq f} d1(m,n,p,q)d2(e,f,p,q)\cdots \tag{3.51}$$

$$d1(m,n,p,q)d2(e,n,g,q)\cdots \longrightarrow \sum_{m\leq n} \sum_{p\leq q} \sum_e \sum_g d1(m,n,p,q)d2(e,n,g,q)\cdots \tag{3.52}$$

The general rule is clearly visible from the above examples. Certainly, orbital indices we excite from and we excite to have to be different - otherwise contributions are already included in the formulas involving lower rank $\hat{B}_i$ operators. Furthermore, some terms in CSC represent actually a number of terms according to the **following rule**: if there is one contracted index among the indices we excite from or we excite to in the second denominator (i.e. an index equal to any index occuring in the corresponding part of the first denominator) it generates two terms with two different contractions when the first denominator contains two indices in the corresponding part. For example

$$\begin{aligned}
&\mathrm{d1}(\,m,n,p,q\,)\mathrm{d2}(\,e,n,g,q\,) < |\,\mathrm{b2}(\,m,n,p,q\,)\,\mathrm{b2}(\,p,e,m,g\,)\,\mathrm{b2}(\,g,q,e,n\,)\,| > = \\
&\mathrm{d1}(\,m,n,p,q\,)[\;\mathrm{d2}(\,e,n,g,q\,) < |\,\mathrm{b2}(\,m,n,p,q\,)\,\mathrm{b2}(\,p,e,m,g\,)\,\mathrm{b2}(\,g,q,e,n\,)\,| > \\
&+\mathrm{d2}(\,e,n,p,g\,) < |\,\mathrm{b2}(\,m,n,p,q\,)\,\mathrm{b2}(\,q,e,m,g\,)\,\mathrm{b2}(\,p,g,e,n\,)\,| > \\
&+\mathrm{d2}(\,m,e,g,q\,) < |\,\mathrm{b2}(\,m,n,p,q\,)\,\mathrm{b2}(\,p,e,n,g\,)\,\mathrm{b2}(\,g,q,m,e\,)\,| > \\
&+\mathrm{d2}(\,m,e,p,g\,) < |\,\mathrm{b2}(\,m,n,p,q\,)\,\mathrm{b2}(\,q,e,n,g\,)\,\mathrm{b2}(\,p,g,m,e\,)\,| >]
\end{aligned}$$

Finally the expression for $S1$, eq. (refs1contr), in the CSC notation is

$$\begin{aligned}
S1 \;=\; &\mathrm{d1}(\,m,n,p,q\,)[\;\mathrm{d2}(\,e,f,p,q\,) < |\,\mathrm{b2}(\,m,n,p,q\,)\,\mathrm{b2}(\,e,f,m,n\,)\,\mathrm{b2}(\,p,q,e,f\,)\,| > \\
&+\mathrm{d2}(\,m,n,g,h\,) < |\,\mathrm{b2}(\,m,n,p,q\,)\,\mathrm{b2}(\,p,q,g,h\,)\,\mathrm{b2}(\,g,h,m,n\,)\,| > \\
&+\mathrm{d2}(\,e,n,g,q\,) < |\,\mathrm{b2}(\,m,n,p,q\,)\,\mathrm{b2}(\,p,e,m,g\,)\,\mathrm{b2}(\,g,q,e,n\,)\,| >]
\end{aligned}$$

The formula for $\langle 0|R_1 H R_2|0\rangle$, with all the contractions and the terms represented in the CSC implicitly expanded, is given in the Appendix (in Table 8.1).

Notice that expressions in Table 3.2 do not include the intermediate projections since for every chain of operators with a certain contraction scheme one may replace the product of integrals

$$\langle 0|\hat{B}_2^{mnpq}|^{pq}_{mn}\rangle\langle^{pq}_{mn}|\hat{B}_2^{pqgh}|^{gh}_{mn}\rangle\langle^{gh}_{mn}|\hat{B}_2^{ghmn}|0\rangle \tag{3.53}$$

by a single integral

$$\langle 0|\hat{B}_2^{mnpq}\hat{B}_2^{pqgh}\hat{B}_2^{ghmn}|0\rangle = \langle 0|\hat{B}_2^{mnpq}\hat{\mathbf{1}}\hat{B}_2^{pqgh}\hat{\mathbf{1}}\hat{B}_2^{ghmn}|0\rangle =$$

$$= \langle 0|\hat{B}_2^{mnpq}|{}^{pq}_{mn}\rangle\langle{}^{pq}_{mn}|\hat{B}_2^{pqgh}|{}^{gh}_{mn}\rangle\langle{}^{gh}_{mn}|\hat{B}_2^{ghmn}|0\rangle \tag{3.54}$$

The last step follows from the fact that only CSFs associated with the orbital configurations given by the respective $\hat{B}_2$ operators survive in the identity operators resolved in the basis of all CSFs. As a result all the intermediate states may be simply removed in all contracted chains of operators, but diagonal elements of higher order operators (involving up to 12 replacement operators) have to be computed.

3.4.2 Chains

According to our general strategy, as specified by the equation (3.24) we shall now expand explicitly all the chains of operators to transform them further to the circular chains. In order to apply the commutation rule (3.4) one has to first extract all repetitions of indices. They occur due to summation over pairs e.g. $\sum_{p\leq q}$ and due to the presence of the singly occupied indices one excites from and excite to as well. The latter ones will be called: from/to open shell contributions (FTOS).

Consider first summations over pairs of indices. Any such summation generates certainly two sums with the different chains of orbitals involving $p < q$ and $p = q$ respectively. Thus for example $\sum_{m\leq n}\sum_{p\leq q}\sum_{e\leq f}$ generates eight (2^3) sums with the distinct chains of operators. The corresponding pairs of indices must be different by the definition of the doubly excited configurations $|{}^{pq}_{mn}\rangle$ and $|{}^{pq}_{ef}\rangle$. Furthermore for the doubly excited configurations (relatively to each other) $e \neq m, n$ and $f \neq m, n$ and there are no other repetitions. The same concerns $\sum_{m\leq n}\sum_{p\leq q}\sum_{g\leq h}$.

Table 3.3: S1 expanded with respect to the repetitions of indices – FTOS not included

$$\begin{aligned}
S1 = {}& \\
& \mathrm{d1}(m,n,p,q)\,[\mathrm{d2}(e,f,p,q) + \mathrm{d2}(e,e,p,q) \\
& + \mathrm{d2}(m,n,g,h) + \mathrm{d2}(m,n,g,g) + \mathrm{d2}(e,n,g,q) \\
& + \mathrm{d2}(n,n,g,q) + \mathrm{d2}(e,n,q,q) + \mathrm{d2}(n,n,q,q)]\ + \\
& \mathrm{d1}(m,m,p,q)\,[\mathrm{d2}(e,f,p,q) + \mathrm{d2}(e,e,p,q) \\
& + \mathrm{d2}(m,m,g,h) + \mathrm{d2}(m,m,g,g) + \mathrm{d2}(e,m,g,q) + \mathrm{d2}(e,m,q,q)]\ + \\
& \mathrm{d1}(m,n,p,p)\,[\mathrm{d2}(e,f,p,p) + \mathrm{d2}(e,e,p,p) \\
& + \mathrm{d2}(m,n,g,h) + \mathrm{d2}(m,n,g,g) + \mathrm{d2}(e,n,g,p) + \mathrm{d2}(n,n,g,p)]\ + \\
& \mathrm{d1}(m,m,p,p)\,[\mathrm{d2}(e,f,p,p) + \mathrm{d2}(e,e,p,p) \\
& + \mathrm{d2}(m,m,g,h) + \mathrm{d2}(m,m,g,g) + \mathrm{d2}(e,m,g,p)]
\end{aligned}$$

The third 6-fold summation, namely $\sum_{m\leq n}\sum_{p\leq q}\sum_{e}\sum_{g}$ is more complicated. As specified in (3.48) it is associated with four contraction schemes of the type

$$\langle 0|\hat{B}_2^{mnpq}|{}^{pq}_{mn}\rangle\langle{}^{pq}_{mn}|\hat{B}_2^{pemg}|{}^{gq}_{en}\rangle\langle{}^{gq}_{en}|\hat{B}_2^{gqen}|0\rangle \tag{3.55}$$

In the above case one gets for the e, g indices the following constraints only: $g \neq e$; $e \neq m$; $g \neq p$. Thus one has to take into account additional possibilities:

$$e = n \text{ if } n \neq m; \quad g = q \text{ if } q \neq p; \quad (e = n \text{ and/or } g = q) \text{ if } (n \neq m \text{ and } q \neq p)$$

Finally it gives nine different sums (4+2+2+1 instead of four sums only) as long as we omit the FTOS type of repetitions. Similarily, one can expand the three other chains associated with the considered summation. At that level $S1$ has the expansion as given in Table 3.3. All integrals are dropped for simplicity.

Recall that according to the CSC convention the $\mathrm{d1}(m,n,p,q)\,\mathrm{d2}(e,n,g,q)$ term represents four distinct chains of the replacement operators with the same summations indices whereas $\mathrm{d1}(m,m,p,q)\,\mathrm{d2}(e,m,g,q)$ as well as $\mathrm{d1}(m,n,p,p)\,\mathrm{d2}(e,n,g,p)$ represent two associated chains.

As we have seen before each of the four chains of operators represented in CSC by the single expression $\mathrm{d1}(m,n,p,q)\,\mathrm{d2}(e,n,g,q)$ generates additional chains with e and g indices equal to some (depending on the case) of the m, n and p, q indices respectively. Therefore, each of the five additional terms in Table 3.3 represents a number of terms that may come from the four distinct chains and the CSC must be extended to take that fact into account.

The new terms arise according to the **following rule**: if there is a pair of contracted indices (we excite from or we excite to) in the second denominator and both indices in this pair are the same it generates two terms when the corresponding pair in the first denominator consists of the different indices – in the second term the different index from the corresponding pair in the first denominator replaces the original one. For example

$$\mathrm{d1}(m,n,p,q)\,\mathrm{d2}(n,n,g,q) = \mathrm{d1}(m,n,p,q)\,[\,\mathrm{d2}(n,n,g,q) + \mathrm{d2}(m,m,g,q)\,] \tag{3.56}$$

where any of the two terms on the right hand side still represents two distinct chains according to the previously defined rules of the CSC. Thus, the explicit expressions are

$$\mathrm{d1}(m,m,p,q)\,\mathrm{d2}(e,m,q,q) = \mathrm{d1}(m,m,p,q)\,[\,\mathrm{d2}(e,m,q,q) + \mathrm{d2}(e,m,p,p)\,] \tag{3.57}$$

$$\begin{gathered}\mathrm{d1}(m,n,p,q)\,\mathrm{d2}(n,n,g,q) = \\ \mathrm{d1}(m,n,p,q)\,[\,\mathrm{d2}(n,n,g,q) + \mathrm{d2}(m,m,g,q) + \mathrm{d2}(n,n,p,g) + \mathrm{d2}(m,m,p,g)\,]\end{gathered} \tag{3.58}$$

$$\begin{gathered}\mathrm{d1}(m,n,p,q)\,\mathrm{d2}(n,n,q,q) = \\ \mathrm{d1}(m,n,p,q)\,[\,\mathrm{d2}(n,n,q,q) + \mathrm{d2}(n,n,p,p) + \mathrm{d2}(m,m,q,q) + \mathrm{d2}(m,m,p,p)\,]\end{gathered} \tag{3.59}$$

As mentioned before there are some other possible repetitions due to the presence of the singly occupied indices i.e. the FTOS cases. Consider first the following term

$$\mathrm{d1}(m,p)\,\mathrm{d2}(e,g) < |\,\mathrm{b1}(m,p)\,\mathrm{b2}(p,e,m,g)\,\mathrm{b1}(g,e)\,| > \tag{3.60}$$

appearing in the expression for the $\langle 0|R_1 H R_2|0\rangle$ (see Table 3.2). It corresponds to two singly excited configurations $|{}^{p}_{m}\rangle$ and $|{}^{g}_{e}\rangle$ being doubly excited with respect to each other. Since the only restrictions that exclude double occurring of the indices are: $p \neq m$, $g \neq e$, $e \neq m$ and $g \neq p$ in the open shell case we have to take into account the possible repetitions $e = p = s$ and/or $g = m = t$. Thus when distinguishing all the distinct sets of the orbital indices involved we get except the term (3.60) itself (with all indices different) three other terms

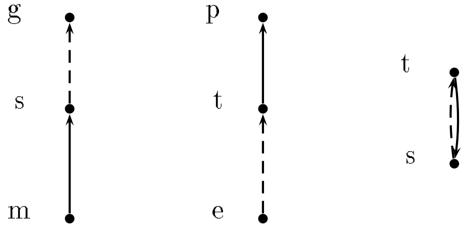
$$\mathrm{d1}(\,m,s\,)\,\mathrm{d2}(\,s,g\,) < |\,\mathrm{b1}(\,m,s\,)\,\mathrm{b2}(\,s,s,m,g\,)\,\mathrm{b1}(\,g,s\,)\,| >$$

$$\mathrm{d1}(\,t,p\,)\,\mathrm{d2}(\,e,t\,) < |\,\mathrm{b1}(\,t,p\,)\,\mathrm{b2}(\,p,e,t,t\,)\,\mathrm{b1}(\,t,e\,)\,| >$$

$$\mathrm{d1}(\,t,s\,)\,\mathrm{d2}(\,s,t\,) < |\,\mathrm{b1}(\,t,s\,)\,\mathrm{b2}(\,s,s,t,t\,)\,\mathrm{b1}(\,t,s\,)\,| >$$

They are graphically represented by the respective diagrams on the Fig. 3.2. Solid and dashed lines are used to distinguish excitations defining the two different configurations. The arrows points the nodes (indices) we excite to.

Figure 3.2: FTOS for single excitations



For the doubly excited configurations, being also doubly excited with respect to each other i.e. for the terms occurring in the expression for $S1$ one gets unfortunately as many as 20 additional terms which are gathered in Table 3.4.

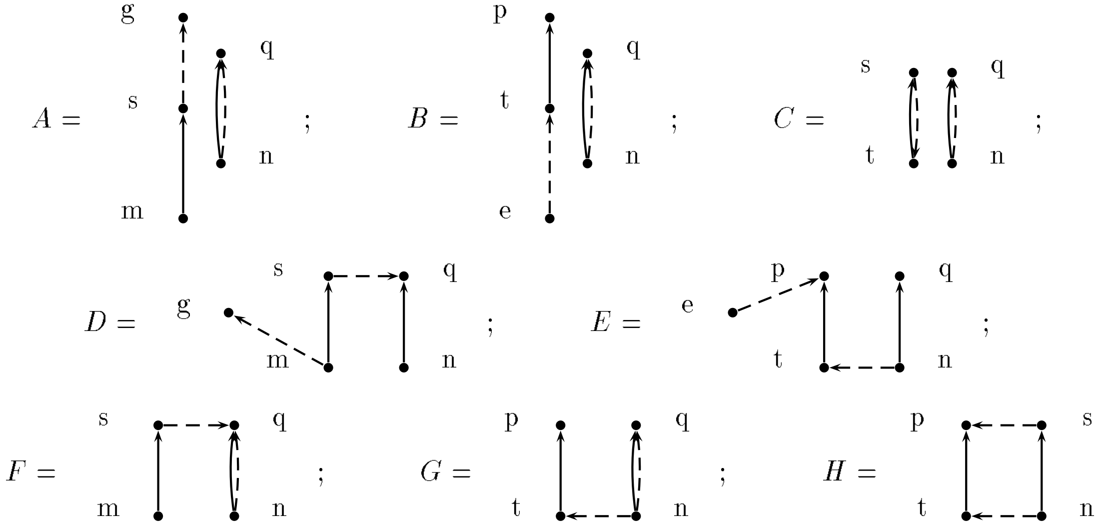
A large number of such terms follows from the fact that they arise independently from the different chains of indices in Table 3.3. For each distinct chain and associated summation scheme one has to consider all possible appearances of the indices that involve excitation from and excitation to. In order to simplify this procedure one may again introduce a graphical representation of the different FTOS terms as given in Figure 3.3. Actually the number of terms that arise from the graphs in this figure is equal to 36 but some of them are equivalent i.e. they involve an equivalent chain of indices for the same summation scheme. Let us learn how to generate different FTOS contributions from those graphs.

As an example the term $\mathrm{d1}(\,m,n,p,q\,)\,\mathrm{d2}(\,e,n,g,q\,)$, $m \neq n$ and $p \neq q$ is taken. As we know it corresponds to four chains. For each chain we get three different FTOS terms since for two free indices there are three possible situations: one or another index or both of them being singly occupied. Thus we get

$$\mathrm{d1}(\,m,n,s,q\,)\,\mathrm{d2}(\,s,n,g,q\,) \longrightarrow \sum_{m<n}\sum_{s}\sum_{q}\sum_{g} \longrightarrow A$$

$$\mathrm{d1}(\,t,n,p,q\,)\,\mathrm{d2}(\,e,n,t,q\,) \longrightarrow \sum_{t}\sum_{n}\sum_{p<q}\sum_{e} \longrightarrow B$$

Figure 3.3: FTOS for double excitations



$$\mathrm{d1}(\,t, n, s, q\,)\,\mathrm{d2}(\,s, n, t, q\,) \quad \longrightarrow \quad \sum_t \sum_n \sum_s \sum_q \quad \longrightarrow \quad C$$

for the term $\mathrm{d1}(\,m, n, p, q\,)\,\mathrm{d2}(\,e, n, g, q\,)$ itself. Notice that in the first of the above terms the summation indices are subject to the following restrictions: $q \neq s$ and $g \neq q$ - otherwise the relative excitation would be different from two and $g \neq s$ by the definition of the doubly excited configuration $|{}^{gq}_{sn}\rangle$. Thus there are no more chains represented by this term according to the CSC (one free index g, $g \neq s$). It concerns the two remaining terms as well.

For $\mathrm{d1}(\,m, n, p, q\,)\,\mathrm{d2}(\,m, e, g, q\,)$ we get in turn

$$\begin{aligned}
\mathrm{d1}(\,m, n, s, q\,)\,\mathrm{d2}(\,m, s, g, q\,) \quad &\longrightarrow \quad \sum_{m<n} \sum_s \sum_q \sum_g \quad \longrightarrow \quad D \\
\mathrm{d1}(\,m, t, p, q\,)\,\mathrm{d2}(\,m, e, t, q\,) \quad &\longrightarrow \quad \sum_m \sum_t \sum_{p<q} \sum_e \quad \longrightarrow \quad E\binom{p\leftrightarrow q}{m\leftrightarrow n} \\
\mathrm{d1}(\,m, t, s, q\,)\,\mathrm{d2}(\,m, s, t, q\,) \quad &\longrightarrow \quad \sum_m \sum_t \sum_s \sum_q \quad \longrightarrow \quad H\binom{p\leftrightarrow q}{m\leftrightarrow n}
\end{aligned}$$

$X\binom{p\leftrightarrow q}{m\leftrightarrow n}$ denotes a diagram X with the simultaneous exchanges of the indices $m \leftrightarrow n$ and $p \leftrightarrow q$. One may easily verify that this operation generates non-equivalent diagrams only if it involves terms with the summation over pairs. Thus, for example $H\binom{p\leftrightarrow q}{m\leftrightarrow n} \equiv H$. On the other hand since the second term involves summation over pairs $E\binom{p\leftrightarrow q}{m\leftrightarrow n} \not\equiv E$. Indeed, in $E \equiv \mathrm{d1}(\,t, n, p, q\,)\,\mathrm{d2}(\,e, n, t, p\,)$ the index p (and not q) is distinguished with respect to the sum $\sum_{p<q}$ in the second denominator. Analogously, we get for $\mathrm{d1}(\,m, n, p, q\,)\,\mathrm{d2}(\,m, e, p, g\,)$

$$\begin{aligned}
\mathrm{d1}(\,m, n, p, s\,)\,\mathrm{d2}(\,m, s, p, g\,) \quad &\longrightarrow \quad \sum_{m<n} \sum_p \sum_s \sum_g \quad \longrightarrow \quad A\binom{p\leftrightarrow q}{m\leftrightarrow n} \\
\mathrm{d1}(\,m, t, p, q\,)\,\mathrm{d2}(\,m, e, p, t\,) \quad &\longrightarrow \quad \sum_m \sum_t \sum_{p<q} \sum_e \quad \longrightarrow \quad B\binom{p\leftrightarrow q}{m\leftrightarrow n}
\end{aligned}$$

$$\mathrm{d1}(\,m,t,p,s\,)\,\mathrm{d2}(\,m,s,p,t\,) \;\longrightarrow\; \sum_{m}\sum_{t}\sum_{p}\sum_{s} \;\longrightarrow\; C\binom{p\leftrightarrow q}{m\leftrightarrow n} \equiv C$$

and for $\mathrm{d1}(\,m,n,p,q\,)\,\mathrm{d2}(\,e,n,p,g\,)$

$$\begin{aligned}
\mathrm{d1}(\,m,n,p,s\,)\,\mathrm{d2}(\,s,n,p,g\,) &\;\longrightarrow\; \sum_{m<n}\sum_{p}\sum_{s}\sum_{g} \;\longrightarrow\; D\binom{p\leftrightarrow q}{m\leftrightarrow n}\\
\mathrm{d1}(\,t,n,p,q\,)\,\mathrm{d2}(\,e,n,p,t\,) &\;\longrightarrow\; \sum_{t}\sum_{n}\sum_{p<q}\sum_{e} \;\longrightarrow\; E\\
\mathrm{d1}(\,t,n,p,s\,)\,\mathrm{d2}(\,s,n,p,t\,) &\;\longrightarrow\; \sum_{t}\sum_{n}\sum_{p}\sum_{s} \;\longrightarrow\; H
\end{aligned}$$

All the remaining terms may be obtained by joining in all possible manners the (not singly occupied) vertices and applying the exchange of indices $m \leftrightarrow n$ and $p \leftrightarrow q$. Only the terms non-equivalent to any other previously obtained have to be included. In the alphabetical order we get first 6 terms from A and from B as well. Notice that the diagrams F and G arise from A and B respectively absorbing two additional terms in both cases - they are distinguished in order to illustrate how the new terms may arise by joining the respective vertices. Then we get 2 terms (1 independent) from C, 6 (2 independent) from D and E as well, 4 (2 independent) from F and G and 2 (1 independent) from H.

Some examples of equivalent terms are:

$$A(m=n) \equiv D(m=n); \qquad A(g=q) \equiv F$$

$$A(m=n,g=q) \equiv A(m=n,g=q)\binom{p\leftrightarrow q}{m\leftrightarrow n} \equiv F(m=n)\binom{p\leftrightarrow q}{m\leftrightarrow n} \equiv F(m=n) \equiv D(m=n,g=q)$$

$$F\binom{p\leftrightarrow q}{m\leftrightarrow n} \equiv F; \qquad G\binom{p\leftrightarrow q}{m\leftrightarrow n} \equiv G$$

All the distinct FTOS terms in the $S1$ expression are explicitly given in Table 3.4.

Table 3.4: FTOS terms in S1

$$\begin{aligned}
&S1(FTOS) =\\
&\quad \mathrm{d1}(\,m,n,s,q\,)\,\mathrm{d2}(\,s,n,g,q\,) + \mathrm{d1}(\,m,m,s,q\,)\,\mathrm{d2}(\,s,m,g,q\,) + \mathrm{d1}(\,m,n,s,q\,)\,\mathrm{d2}(\,m,s,g,q\,) \;+\\
&\quad \mathrm{d1}(\,t,n,p,q\,)\,\mathrm{d2}(\,e,n,t,q\,) + \mathrm{d1}(\,t,n,p,p\,)\,\mathrm{d2}(\,e,n,t,p\,) + \mathrm{d1}(\,t,n,p,q\,)\,\mathrm{d2}(\,e,n,p,t\,) \;+\\
&\quad \mathrm{d1}(\,t,n,s,q\,)\,\mathrm{d2}(\,s,n,t,q\,) + \mathrm{d1}(\,t,n,p,s\,)\,\mathrm{d2}(\,s,n,p,t\,) \;+\\
&\quad \mathrm{d1}(\,m,n,p,s\,)\,\mathrm{d2}(\,s,n,p,g\,) + \mathrm{d1}(\,m,m,p,s\,)\,\mathrm{d2}(\,s,m,p,g\,) + \mathrm{d1}(\,m,n,p,s\,)\,\mathrm{d2}(\,m,s,p,g\,) \;+\\
&\quad \mathrm{d1}(\,m,n,s,q\,)\,\mathrm{d2}(\,s,n,q,q\,) + \mathrm{d1}(\,m,m,s,q\,)\,\mathrm{d2}(\,s,m,q,q\,) + \mathrm{d1}(\,m,n,s,q\,)\,\mathrm{d2}(\,m,s,q,q\,) \;+\\
&\quad \mathrm{d1}(\,m,t,p,q\,)\,\mathrm{d2}(\,m,e,t,q\,) + \mathrm{d1}(\,m,t,p,p\,)\,\mathrm{d2}(\,m,e,t,p\,) + \mathrm{d1}(\,m,t,p,q\,)\,\mathrm{d2}(\,m,e,p,t\,) \;+\\
&\quad \mathrm{d1}(\,t,n,p,q\,)\,\mathrm{d2}(\,n,n,t,q\,) + \mathrm{d1}(\,t,n,p,p\,)\,\mathrm{d2}(\,n,n,t,p\,) + \mathrm{d1}(\,t,n,p,q\,)\,\mathrm{d2}(\,n,n,p,t\,)
\end{aligned}$$

In contradistinction to the case of doubly excited configurations one gets only a few terms in case of one singly excited and another doubly excited configurations. As a matter of fact the

FTOS terms arise only if those configurations are doubly excited relatively to each other. They may be obtained from the D and E diagrams while removing the g and e vertices respectively. Thus we have only four such terms.

All the distinct chains of indices associated with appropriate summation schemes are gathered in the Appendix - see Table 8.2. When all chains of the different indices (and thus of the replacement operators) are uniquely defined one may use the commutations rules (3.4) to make those chains circular. Substituting the definitions of all quantities involved one gets an expression in terms of sums over orbital indices with some chains of the replacement operators.

3.4.3 Factorizations

The most straightforward strategy would be now to extract all possible circular chains of the replacement operators and to evaluate for all the different occupations schemes the corresponding spin integrals. This simple approach leads however to a large number of final spin integrals that have to be considered. Moreover these integrals would involve long cycles and require significant effort for the computation of the representation matrices. In the following we shall demonstrate how to reduce complexity of the above direct approach by means of various conveniently chosen intermediate states and resulting factorizations of the spin integrals.

Let us consider the first term in the expression (3.48) for $S1$ which we shall denote $S1\Sigma_1$, expanding explicitly the second and the third $\hat{B}_2$ operators

$$S1\Sigma_1 = \sum_{m\leq n}\sum_{p\leq q} D_1^{mnpq}\langle 0|\hat{B}_2^{mnpq}|{}^{pq}_{mn}\rangle\langle{}^{pq}_{mn}|A|0\rangle \tag{3.61}$$

where using $\delta'_{xy} = 1 - \delta_{xy}$ and $\delta''_{xywz} = -(\delta_{xy} + \delta_{wz})$

$$\begin{aligned} A &= 2^{\delta''_{pqmn}}\sum_{g} D_2^{mngg}\ (pg|qg)(gm|gn)\hat{E}_{pg}\hat{E}_{qg}\hat{E}_{gm}\hat{E}_{gn} + \sum_{g<h} D_2^{mngh}B \\ B &= (pg|qh)(gm|hn)\hat{E}_{pg}\hat{E}_{qh}\hat{E}_{gm}\hat{E}_{hn} + \delta'_{mn}\ (pg|qh)(gn|hm)\hat{E}_{pg}\hat{E}_{qh}\hat{E}_{gn}\hat{E}_{hm} + \\ &+ \delta'_{pq}(ph|qg)(gm|hn)\hat{E}_{ph}\hat{E}_{qg}\hat{E}_{gm}\hat{E}_{hn} + \delta'_{mn}\delta'_{pq}(ph|qg)(gn|hm)\hat{E}_{ph}\hat{E}_{qg}\hat{E}_{gn}\hat{E}_{hm} \end{aligned} \tag{3.62}$$

One has to distinguish in the above expression different repetitions of indices, as specified in Table 3.3. As an example we may first expand explicitly the term involving the second chain of A for $n \neq m$ and $q \neq p$ which is

$$\sum_{m<n}\sum_{p<q}\sum_{g<h} D_1^{mnpq}D_2^{mngh}\ (pg|qh)(gm|hn)\langle 0|\hat{B}_2^{mnpq}\hat{E}_{pg}\hat{E}_{qh}\hat{E}_{gm}\hat{E}_{hn}|0\rangle \tag{3.63}$$

where according to the definition of $\hat{B}_2^{mnpq}$ the integral splits in two products of the two-electron orbital integrals and spin integrals

$$\langle 0|\hat{B}_2^{mnpq}\hat{E}_{pg}\hat{E}_{qh}\hat{E}_{gm}\hat{E}_{hn}|0\rangle =$$

$$(mp|nq)\langle 0|\hat{E}_{mp}\hat{E}_{nq}\hat{E}_{pg}\hat{E}_{qh}\hat{E}_{gm}\hat{E}_{hn}|0\rangle +\ (mq|np)\langle 0|\hat{E}_{mq}\hat{E}_{np}\hat{E}_{pg}\hat{E}_{qh}\hat{E}_{gm}\hat{E}_{hn}|0\rangle \tag{3.64}$$

All the indices involved in the definition of the above spin integrals are different and we may transform them into integrals with circular chains. The results are

$$\langle 0|\hat{E}_{mp}\hat{E}_{nq}\hat{E}_{pg}\hat{E}_{qh}\hat{E}_{gm}\hat{E}_{hn}|0\rangle = \langle 0|\mathcal{E}_{mpgm}\mathcal{E}_{nqhn}|0\rangle \tag{3.65}$$

and respectively

$$\langle 0|\hat{E}_{mq}\hat{E}_{np}\hat{E}_{pg}\hat{E}_{qh}\hat{E}_{gm}\hat{E}_{hn}|0\rangle = \langle 0|\mathcal{E}_{mqhnpgm} - \mathcal{E}_{mqhpgm}|0\rangle \tag{3.66}$$

The above expressions, which are not so simple in general case, may be generated automatically using special function written for this purpose in the Maple language. It is a part of the general program for evaluation of the matrix elements appearing in the Sup-CI method. The Maple code of this function is given in the Appendix since it has potentially wider applications.

The integrals in eqs (3.65), (3.66) are quite complicated. Nevertheless the main problem is associated with the large number of possible occupation schemes that have to be considered. For each distinct summation scheme defined by a certain contraction and certain repetitive occurring of indices we have to take into account all partial sums involving different occupation schemes separately. Otherwise the spin integrals could not be evaluated using the technique described previously.

We have three pairs of the summation indices in the case considered above, for which the possible occupations are:

$$\begin{array}{rcl} m,\ n & \rightarrow & (2,2);\ (2,1);\ (1,1) \\ p,\ q;\ \ g,\ h & \rightarrow & (1,1);\ (1,0);\ (0,0) \end{array} \tag{3.67}$$

The assumed order of orbital indices is the following: first the doubly occupied indices, then the singly occupied and the unoccupied ones, which are later than the singly occupied. Thus, from the expression (3.63) we get 27 partial sums that have to be separately treated since they involve integrals with different indices being doubly or singly occupied or unoccupied. These sums run certainly over the respective subsets of all indices e.g. the occupations $(2,2)(0,0)(0,0)$ are taken into account through the partial sum

$$\sum_{i<j}\sum_{a<b}\sum_{c<d}\cdots \tag{3.68}$$

where according to our convention i, j designate doubly occupied indices and a, b, c, d unoccupied indices respectively.

Consider now another extreme case when $m = n$, $p = q$, $g = h$ and the associated chain of operators is:

$$\frac{1}{2}\langle 0|\hat{E}_{mp}\hat{E}_{mp}\hat{E}_{pg}\hat{E}_{pg}\hat{E}_{gm}\hat{E}_{gm}|0\rangle \tag{3.69}$$

Making it circular we get

$$\frac{1}{2}\langle 0|\mathcal{E}_{mpgm}\mathcal{E}_{mpgm} - \mathcal{E}_{mpgm}\mathcal{E}_{gpg} + \mathcal{E}_{mpgm}\mathcal{E}_{mgm} + \mathcal{E}_{mpgm}\mathcal{E}_{mpm} - 2\mathcal{E}_{mpgm}\mathcal{E}_{pgp} - 2\mathcal{E}_{mpgm}|0\rangle \tag{3.70}$$

The number of possible occupation schemes is however equal to one in this case.

The total number of partial summations that have to be extracted is equal to $27+27+9+1 = 64$. They arise as follows: 27 terms for all indices different, 9 terms for each set of six indices with one pair of equal indices and all the remaining being different e.g. $m = m$, $p < q$, $g < h$ $(3*9 = 27)$, 3 terms for each set with two pairs of equal indices $(3*3 = 9)$ and 1 term for three pairs of equal indices. Similarly one gets 64 partial summations that has to be extracted from the second general contraction scheme in the expression (3.48) for $S1$

$$\sum_{m\leq n}\sum_{p\leq q}\sum_{e\leq f}\cdots \tag{3.71}$$

For the third contraction scheme

$$\sum_{m\leq n}\sum_{p\leq q}\sum_{e}\sum_{g}\cdots \tag{3.72}$$

situation is more complicated since as we have seen some additional repetitions appear (e.g. of the FTOS type). Despite the fact that some repetitions occur and the number of possible occupation schemes is limited the total number of the resulting partial summation that have to be separately considered is unfortunately large because of large number of such additional terms. While excluding these additional repetitions (i.e. assuming that $e \neq g$ and both e and g are not equal to any other summation index) one gets again 64 partial summations. When $m < n$, $p < q$ the possible occupations are

$$\begin{aligned} m,\ n &\rightarrow (2,2);\ (2,1);\ (1,1) \\ p,\ q &\rightarrow (1,1);\ (1,0);\ (0,0) \\ e,\ g &\rightarrow (2,0);\ (2,1);\ (1,0);\ (1,1) \end{aligned} \tag{3.73}$$

giving 36 partial summations. For $m = n$, $p < q$ and $m < n$, $p = q$ one gets 12 summations in each case and for $m = n$, $p = q$ there are 4 summations.

There are five (one with all indices fixed) additional summation schemes with associated chains of operators when relaxing constraints for the summation indices e and g - see Table 3.3. They are the following

$$\begin{aligned} &\mathrm{d1}(\,m,n,p,q\,)\,[\mathrm{d2}(\,n,n,g,q\,) + \mathrm{d2}(\,e,n,q,q\,) + \mathrm{d2}(\,n,n,q,q\,)] \\ &\mathrm{d1}(\,m,m,p,q\,)\ \mathrm{d2}(\,e,m,q,q\,) \\ &\mathrm{d1}(\,m,n,p,p\,)\ \mathrm{d2}(\,n,n,g,p\,) \end{aligned} \tag{3.74}$$

One may easily check that they generate respectively, 18, 18, 0, 12 and 12 partial summations. One may also verify that 20 additional contractions with repetitive occurring of singly occupied indices of the FTOS type generate 80 partial summations. Thus finally $S1$ generates as many as $192 + 60 + 80 = 332$ partial summations in the direct approach described above. Fortunately, as we have seen the number of FTOS and other additional terms is large in the case of $S1$ only, and all other terms give rise to a small number of partial summations.

There is however an alternative solution to the 'direct approach' described above, which allows for significant reduction of the number of different partial summations that have to be separately considered. Introducing two ghost orbitals a, b ($n_a = n_b = 0$ for every function corresponding to $|0\rangle$) one may write down the identity:

$$\hat{E}_{ij} = \hat{E}_{ia}\hat{E}_{aj} \tag{3.75}$$

Let us come back to the expression for $S1\Sigma_1$. Using eq. (3.75) one may shift g, h indices to gather them in every chain of the type $\hat{E}_{pg}\hat{E}_{qh}\hat{E}_{gn}\hat{E}_{hm}$. For example

$$\hat{E}_{pg}\hat{E}_{qg}\hat{E}_{gm}\hat{E}_{gn} = \hat{E}_{pa}\hat{E}_{ag}\hat{E}_{qb}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{am}\hat{E}_{gb}\hat{E}_{bn} = \hat{E}_{pa}\hat{E}_{qb}\hat{E}_{ag}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{gb}\hat{E}_{am}\hat{E}_{bn} \tag{3.76}$$

Inserting then identity operators as specified below one gets

$$\begin{aligned}
\langle{}^{pq}_{mn}|\hat{E}_{pg}\hat{E}_{qg}\hat{E}_{gm}\hat{E}_{gn}|0\rangle &= \langle{}^{pq}_{mn}|\hat{E}_{pa}\hat{E}_{qb}\hat{\mathbf{1}}\hat{E}_{ag}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{gb}\hat{\mathbf{1}}\hat{E}_{am}\hat{E}_{bn}|0\rangle = \\
&= \langle{}^{pq}_{mn}|\hat{E}_{pa}\hat{E}_{qb}|{}^{ab}_{mn}\rangle\langle{}^{ab}_{mn}|\hat{E}_{ag}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{gb}|{}^{ab}_{mn}\rangle\langle{}^{ab}_{mn}|\hat{E}_{am}\hat{E}_{bn}|0\rangle = \\
&= \langle{}^{pq}_{mn}|\hat{E}_{pa}\hat{E}_{qb}\hat{E}_{am}\hat{E}_{bn}|0\rangle\langle{}^{ab}_{mn}|\hat{E}_{ag}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{gb}|{}^{ab}_{mn}\rangle
\end{aligned} \tag{3.77}$$

provided that all functions associated with a given configuration are taken into account.

From the above follows that

$$S1\Sigma_1 = \sum_{m\le n}\sum_{p\le q} D_1^{mnpq}\ \{$$

$$\langle 0|\hat{B}_2^{mnpq}\hat{E}_{pa}\hat{E}_{qb}\hat{E}_{am}\hat{E}_{bn}|0\rangle \times \{\ 2^{\delta''_{pqmn}}\sum_{g} D_2^{mngg}(pg|qg)(gm|gn)\langle 0'|\hat{E}_{ag}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{gb}|0'\rangle +$$

$$+\sum_{g<h} D_2^{mngh}[(pg|qh)(gm|hn)\langle 0'|\hat{E}_{ag}\hat{E}_{bh}\hat{E}_{ga}\hat{E}_{hb}|0'\rangle + \delta'_{pq}(ph|qg)(gm|hn)\langle 0'|\hat{E}_{ah}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{hb}|0'\rangle\ ]\ \}+$$

$$+\langle 0|\hat{B}_2^{mnpq}\hat{E}_{pa}\hat{E}_{qb}\hat{E}_{an}\hat{E}_{bm}|0\rangle\delta'_{mn} \times \sum_{g<h} D_2^{mngh}[(pg|qh)(gn|hm)\langle 0'|\hat{E}_{ag}\hat{E}_{bh}\hat{E}_{ga}\hat{E}_{hb}|0'\rangle +$$

$$+\delta'_{pq}(ph|qg)(gn|hm)\langle 0'|\hat{E}_{ah}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{hb}|0'\rangle\ ]\ \} \tag{3.78}$$

where $|0'\rangle = |{}^{ab}_{mn}\rangle$.

Since the integrals with m, n, p, q and g, h indices involved are separated one may significantly simplify evaluation of the $S1\Sigma_1$ term. Notice that in all different products of orbital and spin integrals the terms resulting from integrals of the type $\langle 0|\hat{B}_2^{mnpq}\hat{E}_{pa}\hat{E}_{qb}\hat{E}_{an}\hat{E}_{bm}|0\rangle$, defined by the occupation of m, n, p, q indices only, are multiplied by a factor which involves summation over $g \le h$ and one of the following spin integrals

$$\langle 0'|\hat{E}_{ah}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{hb}|0'\rangle; \qquad \langle 0'|\hat{E}_{ag}\hat{E}_{bh}\hat{E}_{ga}\hat{E}_{hb}|0'\rangle; \qquad \langle 0'|\hat{E}_{ag}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{gb}|0'\rangle \tag{3.79}$$

which are equivalent to

$$\langle 0'|\mathcal{E}_{ahbga} - \mathcal{E}_{ahga}|0'\rangle = A; \qquad \langle 0'|\mathcal{E}_{aga}\mathcal{E}_{bhb}|0'\rangle = B; \qquad \langle 0'|\mathcal{E}_{aga}\mathcal{E}_{bgb} + \mathcal{E}_{agba} - \mathcal{E}_{aga}|0'\rangle = C \tag{3.80}$$

respectively. Because $n_a = n_b = 1$ in $|0'\rangle$ we have 3 possible occupation schemes in case of A and B with the corresponding values of the integrals.

n_g, n_h	A	B
$0, 0$	$\langle a, b\rangle$	1
$1, 0$	$\langle a, b\rangle + \langle a, b, g\rangle$	$1 + \langle a, g\rangle$
$1, 1$	$(1 + \langle a, h\rangle)(1 + \langle a, g\rangle)\langle a, b\rangle$	$(1 + \langle a, g\rangle)(1 + \langle b, h\rangle)$

Let A_i, B_i denote the value from the ith row of this table. Notice that for instance A_3 may be rewritten as $(1 + \langle a, t\rangle)(1 + \langle a, s\rangle)\langle a, b\rangle$ since in this case both g and h are singly occupied indices ($g \equiv s$, $h \equiv t$). There is only one possible occupation scheme in case of the third integral C, namely $n_g = 0$ and therefore $g \equiv c$ where c denotes unoccupied index. Thus

$$C = \langle 0'|\mathcal{E}_{aa}\mathcal{E}_{bb} + \mathcal{E}_{aba} - \mathcal{E}_{aa}|0'\rangle = 1 + \langle a, b\rangle \tag{3.81}$$

In light of the above it is clear that we may separately extract four partial summations from the sum over pair $g \leq h$ and evaluate easily the spin integrals A, B or C playing then a role of a numerical factor in each partial sum. Thus, the expression for $S1\Sigma_1$ (3.78) reads

$$\begin{aligned} S1\Sigma_1 \quad = \quad & \sum_{m\leq n}\sum_{p\leq q} D_1^{mnpq}\langle 0|\hat{B}_2^{mnpq}\hat{E}_{pa}\hat{E}_{qb}\hat{E}_{am}\hat{E}_{bn}|0\rangle \times \\ \times \quad & \{\ 2^{\delta''_{pqmn}}\sum_{c} D_2^{mncc}(pc|qc)(cm|cn)C + \\ + \quad & \sum_{c<d} D_2^{mncd}[(pc|qd)(cm|dn)B_1 + \delta'_{pq}(pd|qc)(cm|dn)A_1] + \\ + \quad & \sum_{s}\sum_{d} D_2^{mnsh}[(ps|qd)(sm|dn)B_2 + \delta'_{pq}(pd|qs)(sm|dn)A_2] + \\ + \quad & \sum_{s<t} D_2^{mnst}[(ps|qt)(sm|tn)B_3 + \delta'_{pq}(pt|qs)(sm|tn)A_3]\} + R \end{aligned} \tag{3.82}$$

where R designates the remaining terms which may be reduced in the same way.

Now, we may separately consider different occupations of the indices m, n, p, q. We shall illustrate how to extract the subsequent partial summations using as an example the first of two different integrals of the type $\langle 0|\hat{B}_2^{mnpq}\hat{E}_{pa}\hat{E}_{qb}\hat{E}_{an}\hat{E}_{bm}|0\rangle$ appearing in eq. (3.78). Let I^{mnpq} denote $I^{mnpq} = \sum_{m\leq n}\sum_{p\leq q}\langle 0|\hat{B}_2^{mnpq}\hat{E}_{pa}\hat{E}_{qb}\hat{E}_{am}\hat{E}_{bn}|0\rangle$. Four cases have to be distinguished generating 16=1+3+$\bar{3}$+9 separate summations:

1. $m = n = i$; $p = q = c$

This certainly induces only one partial summation

$$I^{iicc} = \sum_{i}\sum_{c}\frac{1}{2}(ic|ic)\langle 0|\hat{E}_{ic}\hat{E}_{ic}\hat{E}_{ca}\hat{E}_{cb}\hat{E}_{ai}\hat{E}_{bi}|0\rangle = \sum_{i}\sum_{c}\frac{1}{2}(ic|ic) \times 4$$

2. (3.) $m = n = i$; $p < q$ ($m < n$; $p = q = c$ respectively)

Here occupation of one pair of indices is specified and three different partial summations (in both cases) arise: $\sum_i\sum_{c<d}$; $\sum_i\sum_s\sum_d$ and $\sum_i\sum_{s<t}$ ($\sum_{i<j}\sum_c$; $\sum_i\sum_s\sum_c$ and $\sum_{s<t}\sum_c$ respectively). Notice that once the distinct chains with $m = n$ or $p = q$ are specified we may gather terms

corresponding to partial summations $\sum_i \sum_s \sum_d$ and $\sum_i \sum_s \sum_c$. We skip the explicit forms of the integrals involved.

4. $m < n$; $p < q$

Nine partial summations arise with integrals that may be easily derived from the general formula

$$\begin{aligned} I^{mnpq} &= (mp|nq)\langle 0|\hat{E}_{mp}\hat{E}_{nq}\hat{E}_{pa}\hat{E}_{qb}\hat{E}_{am}\hat{E}_{bn}|0\rangle + (mq|np)\langle 0|\hat{E}_{mq}\hat{E}_{np}\hat{E}_{pa}\hat{E}_{qb}\hat{E}_{am}\hat{E}_{bn}|0\rangle = \\ &= (mp|nq) \times \langle \mathcal{E}_{mpam}\mathcal{E}_{nqbn}\rangle + (mq|np) \times \langle \mathcal{E}_{mqbnpam} - \mathcal{E}_{mqbpam}\rangle = \\ &= (mp|nq) \times \langle \mathcal{E}_{mpm}\mathcal{E}_{nqn}\rangle + (mq|np) \times \langle \mathcal{E}_{mqnpm} - \mathcal{E}_{mqpm}\rangle \end{aligned} \tag{3.83}$$

These partial summations are:

$$\begin{array}{ccccc} \sum\limits_{i<j}\sum\limits_{s<t} & ; & \sum\limits_{i<j}\sum\limits_{s}\sum\limits_{d}\,; & & \sum\limits_{i<j}\sum\limits_{c<d} \\ \sum\limits_{i}\sum\limits_{s}\sum\limits_{t<u} & ; & \sum\limits_{i}\sum\limits_{s}\sum\limits_{t}\sum\limits_{d}\,; & & \sum\limits_{i}\sum\limits_{s}\sum\limits_{c<d} \\ \sum\limits_{s<t}\sum\limits_{u<w} & ; & \sum\limits_{s<t}\sum\limits_{u}\sum\limits_{d}\,; & & \sum\limits_{s<t}\sum\limits_{c<d} \end{array} \tag{3.84}$$

where the summations over subsequent singly occupied indices are subject to the following restrictions $w \neq s, t$; $u \neq s, t$; $t \neq s$.

The complexity of the problem decreases when using the factorizations of the type of eq. (3.77). Instead of 64 partial summations, implied by the different combinations of the occupation numbers for the orbitals m, n, p, q, g, h, with different spin integrals to evaluate one can consider only 16 partial summations for the indices m, n, p, q and then 4 summation for the pair g, h. Moreover, thanks to factorization the longest chain has only 5 indices $\mathcal{E}_{stuws}$ instead of 7. Similar reduction of complexity may be achieved by various factorizations for all summations and spin integrals considered here.

Recall that the second general contraction scheme in $S1$ (see (3.48)) which shall be denoted $S1\Sigma_2$ has the following form

$$S1\Sigma_2 = \sum_{m \leq n}\sum_{p \leq q} D_1^{mnpq}\langle 0|\hat{B}_2^{mnpq}|{}^{pq}_{mn}\rangle\langle{}^{pq}_{mn}|A|0\rangle \tag{3.85}$$

where the quantity A is defined as

$$\begin{aligned} A &= 2^{\delta''_{pqmn}}\sum_{e} D_2^{eepq}\ (em|en)(pe|qe)\hat{E}_{em}\hat{E}_{en}\hat{E}_{pe}\hat{E}_{qe} + \sum_{e<f} D_2^{efpq} B \qquad (3.86) \\ B &= (em|fn)(pe|qf)\hat{E}_{em}\hat{E}_{fn}\hat{E}_{pe}\hat{E}_{qf} + \delta'_{mn}\ (en|fm)(pe|qf)\hat{E}_{en}\hat{E}_{fm}\hat{E}_{pe}\hat{E}_{qf} + \\ &+ \delta'_{pq}\ (em|fn)(pf|qe)\hat{E}_{em}\hat{E}_{fn}\hat{E}_{pf}\hat{E}_{qe} + \delta'_{mn}\delta'_{pq}\ (en|fm)(pf|qe)\hat{E}_{en}\hat{E}_{fm}\hat{E}_{pf}\hat{E}_{qe} \end{aligned}$$

Consider now as an example the following term

$$\sum_{m \leq n}\sum_{p \leq q} D_2^{mnpq}\langle 0|\hat{B}_2^{mnpq}|{}^{pq}_{mn}\rangle\langle{}^{pq}_{mn}|\ \sum_{e<f} D_2^{efpq}\ (em|fn)(pe|qf)\hat{E}_{em}\hat{E}_{fn}\hat{E}_{pe}\hat{E}_{qf}|0\rangle \tag{3.87}$$

Inserting again two ghost orbitals a and b and applying the commutation rules (3.4) to gather indices m, n, p, q and e, f we get

$$\hat{E}_{em}\hat{E}_{fn}\hat{E}_{pe}\hat{E}_{qf} = \hat{E}_{ea}\hat{E}_{am}\hat{E}_{fb}\hat{E}_{bn}\hat{E}_{pa}\hat{E}_{ae}\hat{E}_{qb}\hat{E}_{bf} =$$

$$= \hat{E}_{pa}\hat{E}_{qb}\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}\hat{E}_{bn}\hat{E}_{am} - A - B + C \tag{3.88}$$

where

$$\begin{aligned} A &= \hat{E}_{pm}\hat{E}_{qb}\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}\hat{E}_{bn} && (3.89)\\ B &= \hat{E}_{pa}\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}\hat{E}_{qn}\hat{E}_{am} && (3.90)\\ C &= \hat{E}_{pm}\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}\hat{E}_{qn} && (3.91) \end{aligned}$$

Let L designate the longest chain in eq. (3.88) i.e. $L = \hat{E}_{pa}\hat{E}_{qb}\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}\hat{E}_{bn}\hat{E}_{am}$. Inserting then identity operators and retaining the non-vanishing projections only one gets

$$\begin{aligned} \langle {}^{pq}_{mn}|L|0\rangle &= \langle {}^{pq}_{mn}|\hat{E}_{pa}\hat{E}_{qb}|{}^{ab}_{mn}\rangle\langle {}^{ab}_{mn}|\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}|{}^{ab}_{mn}\rangle\langle {}^{ab}_{mn}|\hat{E}_{am}\hat{E}_{bn}|0\rangle = \\ &= \langle {}^{pq}_{mn}|\hat{E}_{pa}\hat{E}_{qb}\hat{E}_{am}\hat{E}_{bn}|0\rangle\langle {}^{ab}_{mn}|\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}|{}^{ab}_{mn}\rangle && (3.92)\\ \langle {}^{pq}_{mn}|A|0\rangle &= \langle {}^{pq}_{mn}|\hat{E}_{pm}\hat{E}_{qb}|{}^{b}_{n}\rangle\langle {}^{b}_{n}|\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}|{}^{b}_{n}\rangle\langle {}^{b}_{n}|\hat{E}_{bn}|0\rangle = \\ &= \langle {}^{pq}_{mn}|\hat{E}_{pm}\hat{E}_{qb}\hat{E}_{bn}|0\rangle\langle {}^{b}_{n}|\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}|{}^{b}_{n}\rangle && (3.93)\\ \langle {}^{pq}_{mn}|B|0\rangle &= \langle {}^{pq}_{mn}|\hat{E}_{pa}|{}^{aq}_{mn}\rangle\langle {}^{aq}_{mn}|\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}|{}^{aq}_{mn}\rangle\langle {}^{aq}_{mn}|\hat{E}_{qn}\hat{E}_{am}|0\rangle = \\ &= \langle {}^{pq}_{mn}|\hat{E}_{pa}\hat{E}_{qn}\hat{E}_{am}|0\rangle\langle {}^{aq}_{mn}|\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}|{}^{aq}_{mn}\rangle && (3.94)\\ \langle {}^{pq}_{mn}|C|0\rangle &= \langle {}^{pq}_{mn}|\hat{E}_{pm}|{}^{q}_{n}\rangle\langle {}^{q}_{n}|\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}|{}^{q}_{n}\rangle\langle {}^{q}_{n}|\hat{E}_{qn}|0\rangle = \\ &= \langle {}^{pq}_{mn}|\hat{E}_{pm}\hat{E}_{qn}|0\rangle\langle {}^{q}_{n}|\hat{E}_{ea}\hat{E}_{fb}\hat{E}_{ae}\hat{E}_{bf}|{}^{q}_{n}\rangle && (3.95) \end{aligned}$$

Although the number of the resulting spin integrals involving the separate pair of indices e, f is larger than for $S1\Sigma_1$ their structure is very simple for $n_a = 0$ or/and $n_b = 0$ in a given $|0'\rangle$. Finally, again only 16 partial summations appear involving separately treated 4 partial summations for pair e, f.

Let us consider now as the last example factorization of the third term in expression for $S1$ (3.48), namely the term associated with the general contraction

$$\sum_{m\leq n}\sum_{p\leq q}\sum_{e}\sum_{g}\cdots \tag{3.96}$$

As previously we may demonstrate that the pair e, g may be separately treated introducing proper intermediate projections. The chains that have to be considered are now of the type

$$\langle 0|\hat{B}_2^{mnpq}|{}^{pq}_{mn}\rangle\langle {}^{pq}_{mn}|\hat{E}_{em}\hat{E}_{pg}\hat{E}_{ge}\hat{E}_{qn}|0\rangle \tag{3.97}$$

This time the ghost orbitals are inserted as follows

$$\hat{E}_{em}\hat{E}_{pg}\hat{E}_{ge}\hat{E}_{qn} = \hat{E}_{ea}\hat{E}_{am}\hat{E}_{pa}\hat{E}_{ag}\hat{E}_{gb}\hat{E}_{be}\hat{E}_{qn} =$$

$$= \hat{E}_{pa}\hat{E}_{ea}\hat{E}_{ag}\hat{E}_{gb}\hat{E}_{be}\hat{E}_{am}\hat{E}_{qn} - A \tag{3.98}$$

where

$$A = \hat{E}_{ea}\hat{E}_{ag}\hat{E}_{gb}\hat{E}_{be}\hat{E}_{pm}\hat{E}_{qn} \tag{3.99}$$

Thus, inserting the intermediate projections one gets

$$\begin{aligned}\langle {}^{pq}_{mn}|\hat{E}_{pa}\hat{E}_{ea}\hat{E}_{ag}\hat{E}_{gb}\hat{E}_{be}\hat{E}_{am}\hat{E}_{qn}|0\rangle &= \langle {}^{pq}_{mn}|\hat{E}_{pa}|{}^{aq}_{mn}\rangle\langle {}^{aq}_{mn}|\hat{E}_{ea}\hat{E}_{ag}\hat{E}_{gb}\hat{E}_{be}|{}^{aq}_{mn}\rangle\langle {}^{aq}_{mn}|\hat{E}_{am}\hat{E}_{qn}|0\rangle = \\ &= \langle {}^{pq}_{mn}|\hat{E}_{pa}\hat{E}_{am}\hat{E}_{qn}|0\rangle\langle {}^{aq}_{mn}|\hat{E}_{ea}\hat{E}_{ag}\hat{E}_{gb}\hat{E}_{be}|{}^{aq}_{mn}\rangle \end{aligned} \tag{3.100}$$

$$\begin{aligned}\langle {}^{pq}_{mn}|A|0\rangle &= \langle {}^{pq}_{mn}|\hat{E}_{ea}\hat{E}_{ag}\hat{E}_{gb}\hat{E}_{be}|{}^{pq}_{mn}\rangle\langle {}^{pq}_{mn}|\hat{E}_{pm}\hat{E}_{qn}|0\rangle = \\ &= \langle {}^{pq}_{mn}|\hat{E}_{pm}\hat{E}_{qn}|0\rangle\langle {}^{pq}_{mn}|\hat{E}_{ea}\hat{E}_{ag}\hat{E}_{gb}\hat{E}_{be}|{}^{pq}_{mn}\rangle \end{aligned} \tag{3.101}$$

Let us finally remark that the above technique may be mechanized. One can easily specify which indices should be gathered. Then, using the commutations rules one gets the desired chains with chosen indices separated. A few rules, except for presented above, for inserting the ghost orbitals depending on the case, are sufficient. All other terms (i.e. not only those included in $S1$) may also be factorized in a similar way. Nevertheless, from the computational point of view, one may gain little when factorizing terms of complexity lower than n^5.

3.4.4 Non-diagonal elements

In the following subsection we shall briefly discuss the evaluation of expressions for the matrix elements in the non-diagonal (in the sense defined in Sec. 3.3) cases. As displayed in Table 3.1 the different types of the non-diagonal matrix elements in the third order Sup-CI are

compl.	type
$< n^4$	$H_{0I,R_10'}$; $S_{0R_1,R_20'}$; $H_{0R_1,R_20'(3,4,5,6)}$
n^4	$H_{0R_1,R_20'(2)}$
n^5	$H_{0R_1,R_20'(1)}$

The most complicated type i.e. $H_{0R_1,R_20'(1)}$ involves 5-fold summation. All these matrix elements may be derived using the same machinery as in the diagonal cases. Only minor revisions are necessary provided that one is able to transform evaluation of the general non-diagonal spin integrals into evaluation of diagonal integrals considered previously.

The reduction of the non-diagonal expressions to the diagonal ones is based on the following observation. Let C denote a certain chain of replacement operators. Then we may consider evaluation of the integral

$$\langle 0|C|{}^{p}_{m}\rangle \tag{3.102}$$

There is a number of spin functions associated with the configuration $|{}^{p}_{m}\rangle$ (defined with respect to certain configuration $|0\rangle$). Since $\hat{E}_{pm}$ generates in general different set of functions when acting on $|0\rangle$ [56] we cannot simply replace $\langle 0|C|{}^{p}_{m}\rangle$ by $\langle 0|C\hat{E}_{pm}|0\rangle$. Relatively simple solution to this problem may however be found. Acting with $\hat{E}_{pm}\hat{E}_{mp}$ on $|{}^{p}_{m}\rangle$ and inserting identity one gets

$$\hat{E}_{pm}\hat{E}_{mp}|{}^{p}_{m}\rangle = \hat{E}_{pm}\mathbf{1}\hat{E}_{mp}|{}^{p}_{m}\rangle = \langle 0|\hat{E}_{mp}|{}^{p}_{m}\rangle\hat{E}_{pm}|0\rangle \tag{3.103}$$

Inserting another identity

$$\hat{\mathbf{1}}\hat{E}_{pm}\hat{E}_{mp}|{}^{p}_{m}\rangle = \langle{}^{p}_{m}|\hat{E}_{pm}\hat{E}_{mp}|{}^{p}_{m}\rangle|{}^{p}_{m}\rangle = \langle 0|\hat{E}_{mp}|{}^{p}_{m}\rangle\hat{E}_{pm}|0\rangle \tag{3.104}$$

Let us denote now the relevant spin integrals as $A(1) = \langle{}^{p}_{m}|\hat{E}_{pm}\hat{E}_{mp}|{}^{p}_{m}\rangle$ and $B(1) = \langle 0|\hat{E}_{mp}|{}^{p}_{m}\rangle$. Recall that these integrals are in fact (representation) matrices of the dimensions determined by the number of spin functions associated with a given orbital configuration. Since $A(1)$ is a square matrix it can be reverted. Thus, we finally get that

$$|{}^{p}_{m}\rangle = A(1)^{-1}B(1)\hat{E}_{pm}|0\rangle \tag{3.105}$$

where the right hand side is certainly a vector of the same dimension as the left hand side vector. The generalization of the above trick to doubly excited functions $|{}^{pq}_{mn}\rangle$ and higher excited configurations is straightforward. At most sextuply excited functions (relatively to each other) may appear in the third order method. Thus, the longest non-diagonal integral involve six replacement operators

$$B(6) = \langle 0|\hat{E}_{mp}\cdots\hat{E}_{nq}|{}^{p\cdots q}_{m\cdots n}\rangle \tag{3.106}$$

The non-diagonal elements $H_{0R_1,R_2 0'(i)}$ correspond to a pair of reference functions $|0\rangle$ and $|0'(i)\rangle$ which are i-fold excited with respect to each other. The number of reference functions is limited and their relative structure may be used to evaluate all integrals of the type $B(i)$ and tabulate them in the simple loop over pairs of references in the preliminary step of the procedure. The integrals $B(i)$ are simple (all occupations are fixed) and may be evaluated directly from the definition or in another convenient way. Thus, the problem of evaluation of general non-diagonal integrals reduces by means of (3.105) to evaluation of a limited number of simple integrals of the type (3.106).

We may specify now how to modify the expressions for diagonal type of matrix elements considered here in order to extract the formulas in non-diagonal cases. Let us turn to matrix elements of the type $ND = H_{0I,R_1 0'(i)}$. In this case $i = 1, 2, 3, 4$. Dropping for simplicity the denominator associated with R_1 we may write more explicitly

$$ND = \sum_{K\in SD(0);\ K\in SD(0')} \langle 0|\hat{H}|K\rangle\langle K|\hat{H}|0'\rangle \tag{3.107}$$

where $|0'\rangle$ is at most quadruply excited with respect to $|0\rangle$. Let the chain of replacement operators corresponding to the configuration $|0'\rangle$ defined with respect to $|0\rangle$ be denoted by $C(0')$. For example $C(|{}^{p}_{m}\rangle) = \hat{E}_{pm}$. The relevant matrices shall be denoted by $B(0')$ and $A(0')$. In the above notation

$$ND = \sum_{K\in SD(0);\ K\in SD(0')} A(0')^{-1}B(0')\langle 0|\hat{H}|K\rangle\langle K|\hat{H}C(0')|0\rangle \tag{3.108}$$

Thus, for a rectangular problem we may first find the square diagonal representation matrix (or integral as it has been called throughout above considerations) D and then transform it into a rectangular one by the matrix multiplication $A^{-1}BD$. This means that it is sufficient to consider diagonal integrals D.

In case of matrix elements of ND type we have specifically four cases that have to be taken into account. When $|0'\rangle$ is quadruply excited with respect to $|0\rangle$ summation over K reduces to 36 doubly excited configurations which involve two pairs of indices defining $|0'\rangle$ with respect to $|0\rangle$. If $|0'\rangle$ is triply excited with respect to $|0\rangle$ only one pair and one index in the relative definition of $|0'\rangle$ are fixed. Therefore, except for 9 singly excited configurations and 9 doubly excited with all indices involved in the relative definition of $|0'\rangle$ gathered in two short summations, two other summations over one free index (belonging to the set of occupied or unoccupied indices respectively) and the other three indices involved in the relative definition of $|0'\rangle$ will appear for the doubly excited configurations.

When $|0'\rangle$ is doubly excited with respect to $|0\rangle$ we may extract the different summations schemes and the corresponding chains of operators from the formula for the diagonal matrix element derived previously

$$D = \sum_{L \in SD(0)} \sum_{K \in SD(0);\ K \in SD(L)} \langle 0|\hat{H}|L\rangle\langle L|\hat{H}|K\rangle\langle K|\hat{H}|0\rangle \tag{3.109}$$

Notice that removing the first summation i.e. the first denominator in the CSC convention employed in the previous section and the first integral $\langle 0|\hat{H}|L\rangle$ we get

$$\sum_{K \in SD(0);\ K \in SD(L)} \langle L|\hat{H}|K\rangle\langle K|\hat{H}|0\rangle \tag{3.110}$$

which is equivalent to ND if we fix the vacuum level at $|0'\rangle$ and if we insert $C(0')$ and multiply by $A(0')^{-1}B(0')$. This means that simply removing the first denominator and the first integral i.e. involving the first B_2 operator in the diagonal type of formulas one gets the formula for ND.

Consider now another type of non-diagonal elements, which is of complexity n^5 if $0'$ is singly excited with respect to $|0\rangle$

$$\sum_{L \in SD(0)} \sum_{K \in SD(0');\ K \in SD(L)} \langle 0|\hat{H}|L\rangle\langle L|\hat{H}|K\rangle\langle K|\hat{H}|0'\rangle \tag{3.111}$$

Formulas for such matrix elements may be also derived by simple modifications of the diagonal formulas. As previously, we have to insert the corresponding chain $C(0')$ and multiply by $A(0')^{-1}B(0')$. However, instead of removing some parts of the diagonal formulas we shall only restrict summations occurring in the diagonal case this time. Let for example $|0'\rangle$ be a doubly excited with respect to $|0\rangle$ configuration. Then, the summation over configurations K subject in general to conditions $K \in SD(0')$; $K \in SD(L)$ may be separated into the following cases:

1. $K \in SD(0)$ - this part may be extracted from the diagonal formula by restricting the summation over K, defined in CSC by second denominator, to configurations involving at least two indices of $|0'\rangle$ when $K \in D(0)$ and at least one index of $|0'\rangle$ when $K \in S(0)$. Notice that since we had only two free summation indices in $K \in SD(0)$; $K \in L(0)$ in the diagonal formula there are no longer free indices and 6-fold summation reduces to 4-fold summation with different chains of operator depending on the actual contraction of indices between $K \in D(0)$ and $|0'\rangle \in D(0)$. In case of singles there was one free index in the diagonal case which is now absorbed as well i.e. it becomes equal to one of indices defining $|0'\rangle$.

2. $K \in TQ(0)$ - since $K \in SD(0')$ and in turn $|0'\rangle \in D(0)$ the relative definition of K with respect to $|0\rangle$ must involve two pairs of indices defining $|0'\rangle$ with respect to $|0\rangle$. Thus we have only one pair (for $K \in T(0)$) or two pairs (for $K \in Q(0)$) of not fixed indices with the contractions resulting from the restriction $K \in SD(L)$. Considering doubles only i.e. when $K \in Q(0)$; $K \in D(0')$ and $L \in D(0)$ one has: if L and $|0'\rangle$ have no indices in common L can contribute to (3.111) only through K being composition of $|0'\rangle$ and L, if L and $|0'\rangle$ have some indices in common L contributes to (3.111) through configurations K involving the indices of L that are not in common. In consequence all the summations are at most 4-fold.

The above restrictions may be easily taken into account in the procedure for generating all different contractions in the diagonal case. All non-diagonal matrix elements may be treated according to the above presented general prescription. One should remark that although the resulting chains of replacement operator are in general longer, the subsets of not fixed indices are smaller and in fact these integrals are much simpler than the integrals considered previously, corresponding to the diagonal type of matrix elements.

3.5 Discussion

As we have seen the evaluation of general open shell third order Sup-CI formulas is quite complicated. The number of terms that arise is large and their structure may be quite complex. Developing the Sup-CI (Fast CI in the original work) method for a limited case of several singlet open shell functions Bendazzoli *et. al.* [42] had to consider 63 MBPT diagrams. For an arbitrary reference functions the problem is certainly more complex. Nevertheless, one can manage this problem using symbolic algebra program based on the algorithm described in this chapter. Various intermediate projections may decrease the number of different partial summation that have to be considered. Graphical techniques may simplify evaluation of some terms.

Additional effort is however required to develop further the Maple code for generating Sup-CI matrix elements. First of all one should look for a possible occurring of equivalent summation schemes to gather more terms together and reduce the number of separate terms. This is crucial for an efficient implementation of these formulas.

It seems that the algebra of generators [56], described in the first part of the chapter and then used to evaluate the relevant matrix elements, would not be useful in orders higher than third. This is because this technique requires explicit extraction of different occupation schemes to find the spin integrals. The number of possible occupation schemes may be very large in higher orders in general open shell case. One could certainly restrict oneself and consider only reference functions of certain simplified structure. The other possible direction of the future development is to include only some higher order effects, which would not require significant effort. Such possibility has been already suggested in Chap. 2.

Part III

Size-consistent modifications of the CI method

Chapter 4

Dressing of CI matrices as a bridge between CI and CC methods

4.1 Introductory notes

The Coupled Cluster approach to the correlation problem [64] offers a rigorous and elegant solution. Assuming the exponential parametrization of the wave operator

$$\hat{\Omega} = e^{\hat{T}} \tag{4.1}$$

one obtains immediately for a system composed of two subsystems A and B a multiplicatively separable wave function provided that the cluster operator is additively separable

$$\hat{T}_{A\cdots B} = \hat{T}_A + \hat{T}_B \tag{4.2}$$

and a separable zeroth order description is used

$$|0_{A\cdots B}\rangle = |0_A\rangle|0_B\rangle \tag{4.3}$$

Assuming also that the molecular orbitals are localized either on A or on B one avoids problems with antisymmetry. As a result of (4.2) and (4.3) one obtains additively separable energy and the CC method is size-consistent:

$$\begin{aligned} E_{A\cdots B} &= \langle 0_{A\cdots B}|(\hat{H}_A + \hat{H}_B)e^{T_{A\cdots B}}|0_{A\cdots B}\rangle = \\ &= \langle 0_A|\hat{H}_A e^{T_A}|0_A\rangle + \langle 0_B|\hat{H}_B e^{T_B}|0_B\rangle = E_A + E_B \end{aligned} \tag{4.4}$$

The additivity of cluster operators (4.2), even truncated at arbitrary level of excitations, is insured by the connectedness of the $\hat{T}$ operator, the condition stating that in the definition of $\hat{T}$

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_N \tag{4.5}$$

there are no disconnected terms i.e. parts which could be resolved into products of two or more lower $\hat{T}_k$ operators [19]. At separation it insures that the cross terms, involving excitations on both systems, will vanish and the eq. (4.2) holds.

This special feature of the CC method is probably even more transparent when associated with the MBPT linked diagram theorem [18]. Instead of eq. (1.25) one may directly project the Schödinger equation

$$(\hat{H} - \langle\Psi_0|\hat{H}|\Psi_0\rangle)|\Psi\rangle = \Delta E|\Psi\rangle \tag{4.6}$$

against $|\Psi_0\rangle$ and excited determinants, obtaining an alternative set of equations for the energy and amplitudes. While restricted to $\hat{T}_1$ and $\hat{T}_2$ they have the following form

$$\Delta E = \langle\Psi_0|\hat{H}_N(1 + \hat{T}_1 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_2)|\Psi_0\rangle \tag{4.7}$$

$$t^a_m\Delta E = \langle{}^a_m|\hat{H}_N(1 + \hat{T}_1 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_2 + \hat{T}_1\hat{T}_2 + \frac{1}{3!}\hat{T}_1^3)|\Psi_0\rangle \tag{4.8}$$

$$(t^{ab}_{mn} + t^a_m t^b_n - t^b_m t^a_n)\Delta E = \langle{}^{ab}_{mn}|\hat{H}_N(1 + \hat{T}_1 + \frac{1}{2}\hat{T}_1^2 + \hat{T}_2 + \hat{T}_1\hat{T}_2 + \frac{1}{3!}\hat{T}_1^3 + \frac{1}{2}\hat{T}_1^2\hat{T}_2 + \frac{1}{2}\hat{T}_2^2 + \frac{1}{4!}\hat{T}_1^4)|\Psi_0\rangle \tag{4.9}$$

which is equivalent to that of the explicitly connected equations because all disconnected terms (i.e. having a disconnected diagramatic form) on the right hand sides of eqs. (4.8) and (4.9), occuring due to quadratic terms of the type $\hat{T}_2^2$, will exactly cancell the energy dependent left hand sides [25].

For any arbitrarily chosen level of truncation of the cluster operator $\hat{T}$ the strict cancellation of the terms defined above may no longer take place, but the connected cluster theorem says that only connected terms (i.e. having a connected diagrammatic representation) need to be considered, implying that only linked terms will appear in the expression for the energy and thus finally insuring the size-extensivity [26].

The relevant terminology might be confusing. Here the convention used by Lindgren [17] is employed. The terms linked and unlinked refer to closed energy diagrams whereas terms connected and disconnected refer to diagrammatic representation of operators (e.g. of the wave operator) in the second quantization language. Thus, by an unlinked diagram one means a diagram which does contain a closed disconnected part, and by a disconnected one a diagram which has open disconnected part.

In contrast the CI linear expansion does not insure the size-extensivity, when truncated at some level. For a system of N non-interacting H_2 molecules one obtains non-linear dependence of the energy with respect to N (in the limit $N \rightarrow \infty$ the energy is proportional to $\sqrt{N}$) [23]. By projections of eq. (1.15) one is led to the following set of equations for the energy and CI coefficients (when limiting the expansion to the singly and doubly excitated determinants $\hat{\Omega}_{CI} = 1 + \hat{C}_1 + \hat{C}_2$)

$$\Delta E = \langle\Psi_0|\hat{H}_N(1 + \hat{C}_1 + \hat{C}_2)|\Psi_0\rangle \tag{4.10}$$

$$c^a_m\Delta E = \langle{}^a_i|\hat{H}_N(1 + \hat{C}_1 + \hat{C}_2)|\Psi_0\rangle \tag{4.11}$$

$$c^{ab}_{mn}\Delta E = \langle{}^{ab}_{ij}|\hat{H}_N(1 + \hat{C}_1 + \hat{C}_2)|\Psi_0\rangle \tag{4.12}$$

Because, as was pointed out before, the left hand side terms have wrong dependence on the number of electrons it is crucial to cancel them at least approximately to correct the improper behavior of the CI method. Such a cancellation is only possible by implicit mixing of different categories of excitations e.g. for CID it is necessary to account for some effects of quadruples to restore correct scaling. In fact this is what all methods that have been devised to modify the CI method for size-extensivity do, starting from the simple a posteriori Davidson-like corrections [49] to CC type procedures such as the Quadratic Configuration Interaction (QCI) method [68] [70] or differnet CEPA type modifications [71] that incorporate to some extent the CC terms of the type $\hat{T}_2^2$ and may restore strict size-extensivity.

The methods presented in this chapter certainly follow the same direction. They generally employ the Coupled Cluster assumption saying that one can reasonably approximate the higher excitations effects in terms of products of lower rank excitations and they use the fact that the cluster assumption coupled with the exponential form of the wave function leads to a cancellation of unlinked effects.

Since the CC approach has natural links with MBPT [19], the second quantization language is the most natural for CC methods. However, since algorithms directed at CI method are considered in this thesis and moreover, in most cases multireference schemes, the choice to use consequently the second quantization formalism seems not to be the most fruitful. Adding some CC derived terms to the corresponding CI matrix (which is actually the definition of a *dressing* of CI matrices) may confuse readers used to the MBPT language but hopefully this shall not become a serious obstacle in understanding this part of the thesis.

Another general remark should be also made before we go further. Recall that for size-consistency we have to insure first of all the correct separation of a reference $|0_{AB}\rangle$

$$|0_{AB}\rangle \overset{r_{AB}\to\infty}{\longrightarrow} |0_{A\cdots B}\rangle = |0_A\rangle|0_B\rangle \tag{4.13}$$

which is the case of a single RHF reference determinant when the molecule dissociates into closed shell fragments or UHF reference determinant otherwise (at least in some cases). The UHF introduces a spin contamination and therefore in general is not a proper choice [21] [16].

The separability of electronic states is not a trivial question, except the simplest (and rather rare) case when a ground, closed shell state dissociates into two ground, closed shell states of the fragments [16]. The situation gets complicated when multireference description is inevitable. If at least one of the dissociation products requires multireference description the model space for AB system (S_{AB}) no longer consists of a single reference function as well, and is usually built as a proper spin and symmetry subspace of the Complete Active Space (CAS) i.e. a set of all determinants obtainable by excitations of valence electrons within valence orbitals. The active (valence) one-particle space is chosen to provide qualitatively correct description of the separation into fragments [16]. The last requirement may imply very large model spaces in a sense of the model spaces occuring in the theory of effective Hamiltonians. Intermediate Hamiltonians used later here impose the corresponding CI spaces as model spaces. But even keeping relatively small complete model spaces one may face the so-called intruder states problem due to the lack of a good energy separation between states derived from the model space states [45] [87] (by switching on the dynamical correlation) and other states.

The CAS zeroth order description is separable [16] i.e. for each reference space state $|m_{AB}^{S_{AB}}\rangle$ there exist subsystem states $|p_A^{S_A}\rangle$ and $|q_B^{S_B}\rangle$ such that

$$|m_{A\cdots B}^{S_{A\cdots B}}\rangle = |p_A^{S_A}\rangle|q_B^{S_B}\rangle \tag{4.14}$$

where $|m_{A\cdots B}^{S_{A\cdots B}}\rangle$ is defined by the separation process

$$|m_{AB}^{S_{AB}}\rangle \overset{r_{AB}\to\infty}{\longrightarrow} |m_{A\cdots B}^{S_{A\cdots B}}\rangle \tag{4.15}$$

Introducing the dynamical correlation one may however face convergence problems due to the possible occurance of avoided crossings at some conformations. Thus, when we consider a given state $|m_{AB}\rangle$ of the AB system, derived from its reference space counterpart $|m_{AB}^{S_{AB}}\rangle$ our method may not converge at all or may 'jump' to a potential curve of another (nearly crossing) state, excluding anyway the correct dissociation.

In the subsequent considerations a number of various (size-consistent) modifications of CI method shall be presented. In all these cases the size-consistency i.e. additivity of the energies at large separation is formally verified, at least for some special structure of the model spaces (in particular separable reference spaces are required - see next section). I would like to stress that this does not mean that the methods are also size-extensive. By the physically oriented definition of size-consistency [22] one may escpape from the inherently MBPT linked diagram theorem and difficulties that appear when using the MBPT language because of the self-consistent character of the new proposals. It means however that size-extensivity (defined as a lack of unlinked terms) is not verified.

Showing that a (projected onto a supersystem model space) product of subsystem solutions for a pair of subsystem states p and q is also a solution of the equations for the supersystem with an additive eigenvalues we formally demonstrate the size-consistency. This concerns however only the states derived from the reference space and not some arbitrary states. When tracing a given root number k for AB problem one should remember that it may be an intruder state, for which the demonstration is not valid and one may obtain a non-additive result for it. Another difficulty arises in connection with the above mentioned convergence problems and possible switching between two solutions. One should not forget about these restrictions when considering the separability condition as implying smoth dissociation of a set of states of system AB.

We shall use the following notation in the current and three next chapters. The capital letters $I, J, \ldots$ will designate N-electron states playing the role of reference functions. The indices i, j, k, l shall designate many-electron states if used to specify state functions (usually some excitations with respect to the reference functions). The indices m, n, p, q and a, b, c, d will refer to occupied and virtual (one-electron) spin orbitals respectively. Thus consequently the excitation operators $\hat{e}_{mn}^{ab}$ and $\hat{e}_i$ are defined as follows

$$\hat{e}_{mn}^{ab}|0\rangle = |{}_{mn}^{ab}\rangle; \qquad \hat{e}_i|0\rangle = |i\rangle \tag{4.16}$$

When considering multireference situations in the spirit of the Hilbert space approach (see Chap. 5) the $\hat{e}_i$ excitation operators will be by default defined with respect to the vacuum given by the reference function $|I\rangle$ they act on $\hat{e}_i|I\rangle$. Otherwise, the vacuum level will be explicitly defined by second subscript e.g. $\hat{e}_{iI}$. In both cases such excitation operator designates the relative excitation, which generates $|i\rangle$ when acting on $|I\rangle$.

4.2 Self-consistent state-specific intermediate Hamiltonians

In the theory of effective Hamiltonians one defines a small subspace of the total Hilbert space $M \subset \mathcal{H}^N$ of projector $\hat{P}_M$, which is called a *model space*

$$\mathcal{H}^N = M \oplus M_Q \tag{4.17}$$

where M_Q is a complementary space and in practice $k = \dim(M) << \dim(\mathcal{H}^N)$. The choice of a model space corresponds to the choice of k lowest roots of the exact Schrödinger equation

$$\hat{H}|\Psi_i\rangle = E_i|\Psi_i\rangle \qquad i = 1, \ldots, k \tag{4.18}$$

for which we want to reproduce exact eigenvalues E_i in the model space i.e. to built an *effective* Hamiltonian $\hat{H}_{\text{eff}}$ having the same eigenvalues when acting on projection of the exact states onto the model space

$$\hat{H}_{\text{eff}}\hat{P}_M|\Psi_i\rangle = E_i\hat{P}_M|\Psi_i\rangle \qquad i = 1, \ldots, k \tag{4.19}$$

Introducing the so-called *wave operator* $\hat{\Omega}$

$$|\Psi_i\rangle = \hat{\Omega}\hat{P}_M|\Psi_i\rangle \qquad i = 1, \ldots, k \tag{4.20}$$

one gets for the effective Hamiltonian $\hat{H}_{\text{eff}} = \hat{P}_M\hat{H}\hat{\Omega}$. Using then Bloch equation (see table below the Fig. 4.1) one may find the wave operator, hence the effective Hamiltonian and diagonalize it in the model space to find desired eigenvalues [74] [45].

In practice the above procedure is only applicable when we have reasonable starting approximation to exact eigenstates. Thus we define a set of k functions $|\Phi_i\rangle$, which span the model space M of projector $\hat{P}_M = \sum_{i=1,\ldots k}|\Phi_i\rangle\langle\Phi_i|$ and from the formal point of view M must be chosen in such a way to insure that the projections of the exact eigenstates will be non-vanishing and lineary indepenedent [45]. To obtain reasonable results those projections should be also reasonable zeroth order approximation to the exact states.

The effective Hamiltonians may be safely applied only when the model and outer space are well separated energetically [87]. Otherwise they suffer from the intruder state problem. To avoid this problem Malrieu *et. al.* [87] have proposed *intermediate* Hamiltonians, which are effective Hamiltonians that are supposed to reproduce only a part of the exact eigenvalues E_i when diagonalizing the intermediate effective Hamiltonian $\hat{H}_{int}$ in the model space M. Tracing only well separated roots should in principle avoid convergence problems connected with the presence of intruders.

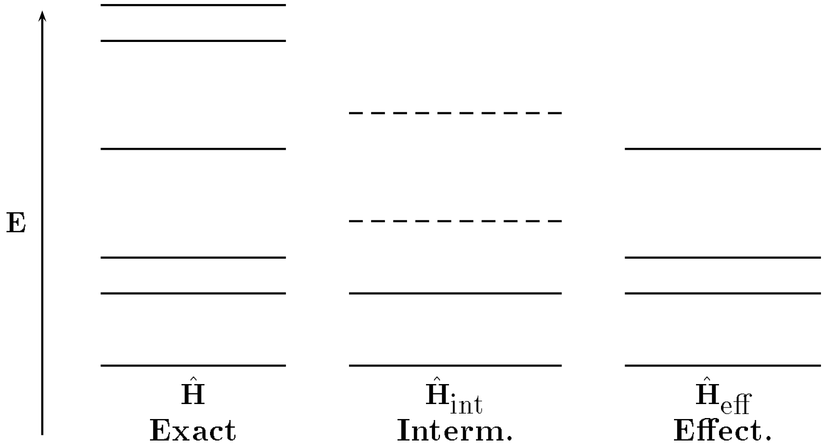
The outline of the intermediate Hamiltonian theory may be summarized as follows. The model space M is divided into the so-called *main* model space M_m (of the dimension $l < k$) and its orthogonal complement M_{int} called an *intermediate* model space

$$M = M_m \oplus M_{int} \tag{4.21}$$

and then one tries to built an intermediate Hamiltonian $\hat{H}_{int}$ satisfying

$$\hat{H}_{int}\hat{P}_M|\Psi_i\rangle = E_i\hat{P}_M|\Psi_i\rangle \qquad i = 1, \ldots, l < k \tag{4.22}$$

Figure 4.1: Effective Hamiltonians vs Intermediate Hamiltonians



Hamiltonians:	Effective	Intermediate
Spaces	$\mathcal{H}^N = M \oplus M_Q$	$M = M_m \oplus M_{int}$
Operators	$\hat{H}_{\mathrm{eff}} = \hat{P}_M \hat{H} \hat{\Omega}$	$\hat{H}_{int} = \hat{P}_M \hat{H} \hat{R}$
Equations	$\hat{H}\hat{\Omega} = \hat{\Omega}\hat{H}_{\mathrm{eff}}$	$\hat{R}\hat{\Omega} = \hat{\Omega}$

One may introduce a counterpart of the wave operator, denoted in the original work by $\hat{R}$ [87] and find some necessary conditions that have to be satisfied by $\hat{R}$. However, these cinditions are not sufficient for a unique definition of $\hat{R}$ [87]. This formal drawback paradoxically may open a way for variuos interesting proposals that would use the theory of intermediate Hamiltonians as a general frame, say general heuristic scheme.

Various dressing procedures aiming generally at making differently truncated CI schemes size-consistent have been recently proposed by the group of Toulouse [115]. Some of them as well as several new ones are presented here. They take benefit of some degree of freedom which is present in the theory of intermediate Hamiltonians and fix their form by introducing a dressing operator which is supposed to account for the outer space determinants effects. This in turn opens a way to very flexible procedures taking into acount CC type non-linear terms.

Modifying CI equations in the dressing procedures we shall define a model space as the corresponding CI space and then distinguishing one or a few lowest roots we shall define a proper main model space. An intermediate Hamiltonian $\hat{H}_{int}$ is postulated to take the form

$$\hat{H}_{int} = \hat{P}_M(\hat{H} + \hat{\Delta})\hat{P}_M \tag{4.23}$$

where the dressing operator $\hat{\Delta}$ is defined by the equation (4.22) [114]. Now we may adjust the general definitions to particular situations considered in this work.

In the following we shall consider a multireference space S spanned by the reference determinants I. We shall also consider the determinants i which do not belong to S, obtained by the single and double substitutions on the references. These determinants span a space s. The corresponding projectors are

$$\hat{P}_S = \sum_{I \in S} |I\rangle\langle I| \tag{4.24}$$

$$\hat{P}_s = \sum_{i \in s} |i\rangle\langle i| \tag{4.25}$$

The determinants i may interact with one or several reference determinants ($\langle i|\hat{H}|I\rangle \neq 0$), depending on their nature and the structure of the reference space. Hereafter the model space will be chosen as the union $M = S \oplus s$ and the associated projector will be labeled $\hat{P}_{S\oplus s}$.

For the state $|\Psi_m\rangle$ (ground or excited state) one can write its expansion in terms of the N-electron basis $\{|I\rangle,\ |i\rangle,\ |\alpha\rangle\}_{I\in S,\, i\in s,\, \alpha\notin S\oplus s}$

$$|\Psi_m\rangle = \sum_{I\in S} C_I^m |I\rangle + \sum_{i\in s} c_i^m |i\rangle + \sum_{\alpha\notin S\oplus s} c_\alpha^m |\alpha\rangle \tag{4.26}$$

Now, we would like to define an intermediate effective Hamiltonian $\hat{P}_{S\oplus s}(\hat{H} + \hat{\Delta}^m)\hat{P}_{S\oplus s}$, where $\hat{\Delta}^m$ is a dressing operator, such that

$$\hat{P}_{S\oplus s}(\hat{H} + \hat{\Delta}^m)\hat{P}_{S\oplus s}|\Psi_m\rangle = E_m \hat{P}_{S\oplus s}|\Psi_m\rangle \tag{4.27}$$

$|\Psi_m\rangle$ and E_m being the exact eigenstate and eigenvalue

$$\hat{H}|\Psi_m\rangle = E_m|\Psi_m\rangle \tag{4.28}$$

As usually for the intermediate effective Hamiltonians [87], we demand the exact energy when acting on the projection of the exact eigenvector onto the model space. The state-specific intermediate Hamiltonians are only required to give *one* exact root (one-dimensional main model space). It should be stressed that contrary to the usual effective Hamiltonians the model and the reference space are different.

Inserting (4.26) into (4.28) and multiplying by $\langle I|$ on the left hand side we obtain

$$\sum_{\substack{J\in S\\ J\neq I}} H_{IJ} C_J^m + (H_{II} - E_m) C_I^m + \sum_{i\in s} H_{Ii} c_i^m = 0 \tag{4.29}$$

which is the Ith row of the matrix representation of the eigenvalue problem (4.28). Since the elements $\langle I|\hat{H}|\alpha\rangle = 0$ this equation is the same for $\hat{H}$ and for $\hat{P}_{S\oplus s}\hat{H}\hat{P}_{S\oplus s}$, so that there is no dressing to introduce in the $P_S H P_{S\oplus s}$ block of the matrix

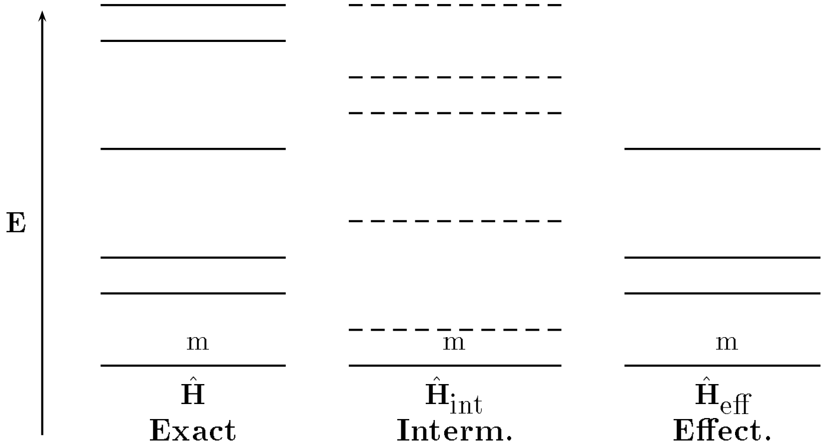
$$\hat{P}_S \hat{\Delta}^m \hat{P}_{S\oplus s} = 0 \tag{4.30}$$

Multiplying (4.28) by $\langle i|$ belonging to the space of the singles and doubles, the eigenequation becomes

$$\sum_{I\in S} H_{iI} C_I^m + \sum_{\substack{j\in s\\ j\neq i}} H_{ij} c_j^m + (H_{ii} - E_m) c_i^m + \sum_{\alpha\notin S\oplus s} H_{i\alpha} c_\alpha^m = 0 \tag{4.31}$$

To obtain a correct dressing one has simply to transform the last summation of the above equation into a proper matrix element of a dressing operator and include it in one of the first

Figure 4.2: Self-consistent Intermediate Hamiltonians



$$\boxed{\hat{H}_{int} = \hat{P}_{S\oplus s}(\hat{H} + \hat{\Delta}(E_m; \mathbf{C}_m))\hat{P}_{S\oplus s}}$$

three terms (i.e. effectively shift the last summation to the model space). One may for instance define a diagonal dressing [89]

$$\Delta^m_{ii} = (\sum_{\alpha \notin S\oplus s} H_{i\alpha} c^m_\alpha)(c^m_i)^{-1} \tag{4.32}$$

$$\Delta^m_{ii} = 0 \quad \text{if } i \neq j \tag{4.33}$$

such that $\hat{P}_{S\oplus s}(\hat{H} + \hat{\Delta}^m)\hat{P}_{S\oplus s}$ has $\hat{P}_{S\oplus s}|\Psi_m\rangle$ as the eigenvector for the energy E_m.

Another dressing operator is obtained assuming that one can write the coefficients of the outer space determinants as a sum over references

$$c^m_\alpha = \sum_{J \in S} c^m_{\alpha J} \tag{4.34}$$

Then one may define a dressing of the first columns of the matrix [105] (more precisely of the $P_{S\oplus s} H P_S$ block)

$$\Delta'^m_{iJ} = (\sum_{\alpha \notin S\oplus s} H_{i\alpha} c^m_{\alpha J})(C^m_J)^{-1} \tag{4.35}$$

$$\Delta'^m_{ij} = 0 \tag{4.36}$$

The two dressings lead to a common eigenvector of $\hat{P}_{S\oplus s}(\hat{H} + \hat{\Delta}^m)\hat{P}_{S\oplus s}$ and $\hat{P}_{S\oplus s}(\hat{H} + \hat{\Delta}'^m)\hat{P}_{S\oplus s}$, namely $|\tilde{\Psi}_m\rangle = \hat{P}_{S\oplus s}|\Psi_m\rangle$ with the eigenvalue E_m, but all other eigenstates are different. It should be emphasized that postulating different forms of dressing we actually impose additional conditions that allow to define an intermediate effective Hamiltonian uniquely [89].

The above proposal is purely academic if one does not have a reasonable evaluation of the coefficients c^m_α of the outer space determinants. Remember that these determinants are here triples and quadruples with respect to at least one reference $I \in S$ and singles or doubles with respect to some of the singles and doubles $i \in s$.

The strategy to approximate the outer space coefficients will be grounded on the CC assumption saying that higher-body effects may be reasonably described in terms of two-body (and one-body) cluster amplitudes.

Before we go into details, let us only notice that the above defined dressing operators, given in the matrix representations by eqs. (4.32) and (4.33) or by eqs. (4.35) and (4.36) may be postulated to take an operator form

$$\hat{\Delta}^m = \hat{P}_s \hat{H} \hat{\omega}^m \hat{P}_S \tag{4.37}$$

where the operator $\hat{\omega}$ generates the outer space determinants. Then one may assume that all α's will be generated with respect to a selected reference $|0\rangle$ only or that they are generated succesively with respect to all references. Certainly in both cases we may go into limit of one-dimensional reference space when these two choices become equivalent. In the first case one may thus write

$$\hat{\omega}^m = \sum_{\alpha \notin S \oplus s} \omega^m_\alpha |\alpha\rangle\langle 0| \tag{4.38}$$

whereas in the second (assuming (4.34))

$$\hat{\omega}^m = \sum_{\alpha \notin S \oplus s} \sum_{I \in S} \omega^m_{\alpha I} |\alpha\rangle\langle I| \tag{4.39}$$

In consistency with the matrix representations the 'amplitudes' are given by

$$\omega^m_\alpha = \frac{c^m_\alpha}{C^m_0} \quad \text{and} \quad \omega^m_{\alpha I} = \frac{c^m_{\alpha I}}{C^m_I} \tag{4.40}$$

and may be exponentially factorized as in the Coupled Cluster theory.

4.3 CEPA method as a dressing of CI matrix

We shall leave now the deductive perspective for a while and recall the known fact that Correlated Electron Pair Approximation (CEPA) type methods [71] [73] may be formulated in a self-consistent manner in terms of the pseudoeigenvalue problem, to use the usual CI machinery for practical implementations [73] [72]. In the following we shall restrict the discussion to the CEPA(2) method [73].

The simplest reasonable structure of the cluster operator leads to CCD approximation, first introduced by Čižek under the name Correlated Pair Many Electron Theory (CPMET) [64]. In the traditional formulation of CCD/CPMET method the equations for the amplitudes of doubles are usually written in the following form [6]

$$\langle {}^{ab}_{mn}|\hat{H}|0\rangle + \sum_{p<q} \sum_{c<d} \langle {}^{ab}_{mn}|\hat{H}|{}^{cd}_{pq}\rangle c^{cd}_{pq} + \sum_{\substack{p<q \\ pq \neq mn}} \sum_{\substack{c<d \\ cd \neq ab}} \langle {}^{ab}_{mn}|\hat{H}|{}^{abcd}_{mnpq}\rangle \; c^{ab}_{mn} \star c^{cd}_{pq} =$$

$$= E_0 c_{mn}^{ab} + \sum_{p<q}\sum_{c<d}\langle 0|\hat{H}|_{pq}^{cd}\rangle c_{pq}^{cd} c_{mn}^{ab} \tag{4.41}$$

where $c_{mn}^{ab} \star c_{pq}^{cd}$ denotes a sum of 18 different products of the coefficients of doubles into wich a given quadruple may be factorized with proper signs defined by the antisymmetry requirement [6]. The above equation is certainly equivalent to eq. (4.9) if we put there $\hat{T}_1 = 0$.

Different CEPA methods corresponds to different approximations of the left hand side non-linear term $LN = \sum_{\substack{p<q\\ pq\neq mn}} \sum_{\substack{c<d\\ cd\neq ab}} \langle {}^{ab}_{mn}|\hat{H}|^{abcd}_{mnpq}\rangle \; c_{mn}^{ab} \star c_{pq}^{cd}$. The CEPA(2) method for instance is defined by [73]

$$LN^{CEPA(2)} = c_{mn}^{ab} \sum_{\substack{p<q\\ pq\neq mn}} \sum_{c<d}\langle 0|\hat{H}|_{pq}^{cd}\rangle c_{pq}^{cd} \tag{4.42}$$

Notice, that the restriction on the summation over virtual spin orbitals (particles) is released. This is due to the definition of the *pair energy* ϵ_{ij}

$$\epsilon_{mn} = \sum_{c<d}\langle 0|\hat{H}|_{mn}^{cd}\rangle c_{mn}^{cd} \tag{4.43}$$

which is defined as a sum over *all* excitation from a given pair m, n of occupied indices. Using eq. (4.43) one may finally write the equation for CEPA(2) amplitude

$$\langle {}^{ab}_{mn}|\hat{H}|0\rangle + \sum_{p<q}\sum_{c<d}\langle {}^{ab}_{mn}|\hat{H}|_{pq}^{cd}\rangle c_{pq}^{cd} = (E_0 + \epsilon_{mn}) c_{mn}^{ab} \tag{4.44}$$

It is clear already here that the term $LN^{CEPA(2)}$ (4.42) is to be added to the ith line ($|i\rangle \equiv |^{ab}_{mn}\rangle$) of CID matrix to obtain an equivalent set of equations. Certainly when we start from the pure CI coefficients they do not satisfy such dressed equations. However the pure CI coefficients define first approximation to the CC amplitudes and non-linear term $LN^{CEPA(2)}$. At convergence of this self-consistent process one obtains the CEPA amplitudes that satisfy eqs. (4.44). Thus the CEPA(2) dressing of the CID matrix may be viewed simply as a way of solving (by a pseudoeigenvalue formulation) CEPA(2) equations.

In fact the pseudoeigenvalue formulation was used for solving CEPA equations in a self-consistent manner already in the seventies [73]. Let us illustrate it using as an example work of Alrichs [72]. Following ref. [72] we may introduce now a matrix block notation, which may be used to study the relationship between CEPA and CID. Let the index μ denotes a pair of occupied indices $\mu \equiv mn$. Then by $H_{0\mu}$ we shall denote a block vector $H_{0\mu} \equiv (H_{0\mu_1}, \ldots, H_{0\mu_m})$ where m is equal to the number of excitations from pair μ and by $H_{\mu\mu}$ a block matrix

$$H_{\mu\mu} \equiv \begin{pmatrix} H_{\mu_1\mu_1} & \cdots & H_{\mu_1\mu_m} \\ & & \vdots \\ & & H_{\mu_m\mu_m} \end{pmatrix}$$

Using this notation we can rewrite the CID and CEPA(2) equations. Let the total number of pairs will be equal to n. The CID eigenvalue problem $\mathbf{Hc} = E_{ci}\mathbf{c}$ is equivalent to a set of coupled matrix equations [72]

$$\begin{pmatrix} E_0 & H_{0\mu} \\ H_{\mu 0} & H_{\mu\mu} \end{pmatrix} \begin{pmatrix} 1 \\ c_\mu^{ci} \end{pmatrix} - (E_0 + \epsilon_\mu^{ci}) \begin{pmatrix} 1 \\ c_\mu^{ci} \end{pmatrix} = - \sum_{\substack{\nu=1,n\\ \nu\neq\mu}} \begin{pmatrix} 0 \\ H_{\mu\nu} c_\nu^{ci} - \epsilon_\nu^{ci} c_\mu^{ci} \end{pmatrix} \tag{4.45}$$

where $E_{ci} = E_0 + \sum_\mu \epsilon_\mu^{ci}$; $\epsilon_\mu^{ci} = H_{0\mu} c_\mu^{ci}$. The corresponding CEPA equation is different by the absence of the 'wrong' term $\sum_{\nu \neq \mu} \epsilon_\nu c_\mu$ and reads

$$\begin{pmatrix} E_0 & H_{0\mu} \\ H_{\mu 0} & H_{\mu\mu} \end{pmatrix} \begin{pmatrix} 1 \\ c_\mu^{cepa} \end{pmatrix} - (E_0 + \epsilon_\mu^{cepa}) \begin{pmatrix} 1 \\ c_\mu^{cepa} \end{pmatrix} = - \sum_{\substack{\nu=1,n \\ \nu \neq \mu}} \begin{pmatrix} 0 \\ H_{\mu\nu} c_\nu^{cepa} \end{pmatrix} \tag{4.46}$$

Comparing eq. (4.45) and eq. (4.46) one immediately notices that the latter one is a result of a diagonal dressing of CID matrix by the quantities $\Delta_{\mu\mu} = \sum_{\nu \neq \mu} \epsilon_\nu$ More precisely all diagonal elements in every block $H_{\mu\mu}$ obtain one common shift $\Delta_{\mu\mu}$.

Thus the CEPA(2) equations in the form of (4.46) form a basis for an iterative proces: starting from the CID equations one may evaluate the CI approximations to diagonal shifts $\Delta_{\mu\mu}^{ci} = \sum_{\nu \neq \mu} \epsilon_\nu^{ci}$, solving then such dressed equations one obtains next approximation etc. At convergence one obtains CEPA solution.

In the Alhrichs' original work the iterative process goes through a ssries of small diagonalizations for pair problems. One starts from IEPA approximation (for which the left hand side of eq. (4.46) is simply equal to zero) and then adds the terms that couple different pairs i.e. the set of CEPA equations may be solved by a self-consistent dressing of IEPA small (pairs) CI problems. For this special case such approach might be more convinient, but going beyond the pair approximation one may no longer use it. Dressing of a total CI matrix is more universal (although for very large CI problems one may be tempted to use again small diagonalizations dressed in a more refined manner - see next section)

4.4 Size-consistent self-consistent CI or exact CEPA method

In light of the above considerations dressing appears as a particular method of solving CEPA equations. This is true and is interesting by itself: machinery developed for solving the CI eigenproblem [43] is very stable even in badly degenerated situations and one may take benefit of that when facing convergence problems with the standard methods of solving various non-linear CC equations.

However, reversing this logic one gains a lot of freedom. Starting from a given CI approximation one may look for different kinds of dressing that may restore size-extensivity or an approximate size-extensivity. This may lead to new formal developments which are somewhere between traditional CC and CI approximations similar (or even superior) in the quality but less complex than traditional CC methods [115].

Certainly the problem of adding some pertubative or non-linear CC type terms to CI equations to make such modified CI method size-extensive has been considered many times for the last two decades. Of course, doing that one looses the property of variational upper bound of the resulting energy [20]. Satisfying one formal requirement another one is spoiled and it seems that there is no escape from this contradiction.

The first proposals to make CI method approximately size-extensive were guided by a pertubative analysis and in the case of CISD approximately accounted for the so-called renormalization term in the fourth order of MBPT which gives rise to diagrams that should be excluded in the linked diagram expresion for the fourth order MBPT energy. This leads to various forms of a posteriori (Davidson-type) corrections [49].

Including some quadratic terms arising from $\hat{T}_2^2$ one obtains various CEPA methods [71]. Adding singles and higher excitations one may consider at each level all terms that are necessary to restore size-extensivity of the corresponding CI. The Quadratic Configuration Interaction (QCI) method proposed by Pople *et al* [68] belongs to this family of methods. In the ref. [70] one may find the review of single reference 'dressings' of this type. Nevertheless all these computational schemes use traditional CC machinery i.e. a set of non-linear equation (comparing to the traditional CC approximation some terms are omitted) is solved using Reduced Linear Equation (RLE) or other methods [69]. It seems that except the recent efforts of Malrieu *et. al.* [86] [115] [85] [105] such trials have never been systematically studied in terms of the pseudoeigevalue problem with the use of CI machinery and additional freedom in choosing terms for a dressing.

In the family of dressings proposed by the Toulouse group the first was Size-consistent Self-consistent Configuration Interaction ((SC)^{2}CI) method [86]. It will be briefly recalled in the remaining part of this section.

Let us first consider purely single reference case i.e. let the reference space S consists of one closed shell determinant $|0\rangle$ only. The s space is assumed to contain all single and double excitations from $|0\rangle$ and the relevant outer space determinants $|\alpha\rangle$ are triples and quadruples. Thus we shall dress single reference CISD matrix. The dressing concerns the ground state only (it is state-specific dressing), so that we shall also assume that $|0\rangle$ is a reasonable approximation to the ground state. In the following the state index m will be omitted.

The diagonal dressing of eqs. (4.32) and (4.33) is employed in the ref. [86]. Further, it is assumed that the outer space determinants coefficients may be approximated by the products of the coefficients of determinants belonging to s. Thus we write

$$c_\alpha = c_i c_j \tag{4.47}$$

for $|\alpha\rangle = \hat{e}_j \hat{e}_i |0\rangle$ where $\hat{e}_j$ and $\hat{e}_i$ are defined by $|j\rangle = \hat{e}_j|0\rangle$ and $|i\rangle = \hat{e}_i|0\rangle$ respectively for some determinants $i, j \in s$. Assumption (4.47) leads to the following form of the diagonal dressing (4.32)

$$\Delta_{ii} = (\sum_{\alpha \notin S \oplus s} H_{i\alpha} c_\alpha)(c_i)^{-1} = \sum_{\substack{j \\ \hat{e}_j \hat{e}_i \neq 0}} H_{0j} c_j \tag{4.48}$$

since $H_{i\alpha} = \langle 0\ {}^{\dagger}\hat{e}_i | \hat{H} | \hat{e}_j \hat{e}_i 0 \rangle = H_{0j}$. When the intermediate normalization is assumed

$$E_{corr} = \sum_j H_{0j} c_j \tag{4.49}$$

and the dressing of eq. (4.48) reads $E_{corr} + EPV_i$ where EPV_i denotes a sum over contributions from the conjoint or Exclusion Principle Violating (EPV) terms

$$EPV_i = - \sum_{\substack{j \\ \hat{e}_j \hat{e}_i = 0}} H_{0j} c_j \tag{4.50}$$

From the practical point of view it is sufficient to calculate efficiently the EPV contributions.

Restricting ourselves to double excitations in s space only we may compare the (SC)^{2}CID dressing directly with that of CEPA(2) presented in the previous section. The latter one for the ith row, $|i\rangle \equiv |^{ab}_{mn}\rangle$ reads (compare eq. (4.42)

$$\Delta^{cepa}_{ii} = \sum_{\substack{p<q \\ pq\neq mn}} \sum_{c<d} \langle 0|\hat{H}|^{cd}_{pq}\rangle c^{cd}_{pq} \tag{4.51}$$

The new dressing takes into account *all* EPV effects

$$\Delta^{(sc)^2cid}_{ii} - \Delta^{cepa}_{ii} = - \sum_{\substack{p<q \\ pq\neq mn}} \langle 0|\hat{H}|^{ab}_{pq}\rangle c^{ab}_{pq} \tag{4.52}$$

and therefore the (SC)^{2}CID method may be regarded as an exact CEPA method. As the two dressings differ only by EPV terms all the unlinked effects are removed in both cases and the (SC)^{2}CID method is strictly size-extensive [86]. Notice also that for a two electron system $\Delta^{(sc)^2cid}_{ii}$ vanishes. The CID is equivalent to FCI in this case and should not be dressed. Thus the (SC)^{2}CID method has correct behavior in this limit similarly as the CEPA(2) method - the difference given by eq. (4.52) is equal to zero in such a case.

The form of the (SC)^{2}CI dressing (eq. (4.48) remains unchanged when also single excitations are included as well as some selected higher excitations appear [86]. The (SC)^{2}CISD method is a particular case of MR(SC)^{2}CI procedure considered further. Therefore the proof of size-consistency of the MR(SC)^{2}CI method, given in Sec. 5.2, is applicable to (SC)^{2}CISD method.

In a general case with some selected singles, doubles and higher excitations with respect to $|0\rangle$ generated as singles and doubles with respect to a set of references $|I\rangle \in S$ ($|0\rangle \in S$) one still may treat all those excitations as derived from the ground state determinant only. It corresponds to the choice (4.38) of the parametrization of the dressing operator. The dressing has now the form

$$\Delta_{ii} = \Big(\sum_{\alpha\notin S\oplus s} H_{i\alpha}c_\alpha\Big)(c_i)^{-1} = \sum_{\substack{j \\ \hat{e}_j\hat{e}_i\neq 0;\ \hat{e}_j\hat{e}_i|0\rangle\notin S\oplus s}} H_{0j}c_j \tag{4.53}$$

or equivalently

$$\Delta_{ii} = E_{corr} + EPV_i + R_i \tag{4.54}$$

where the term R_i takes care of possible redundancy effects

$$R_i = - \sum_{\substack{j \\ \hat{e}_j\hat{e}_i|0\rangle\in S\oplus s}} H_{0j}c_j \tag{4.55}$$

The outer space sextuple excitation for instance will appear in the dressing of model space quadruples with the coefficients being products of quadruples and proper doubles (i.e. defined by the decompositon of our sextuple including a given quadruple) coeffients. For selected CI procedures the dressing is selected as well i.e. we may only dress by this part of interacting outer space determinants which are obtainable as products of model space excitations. Because some higher excitations are present in the model space one has to take care of possible redundancies R_i: e.g. product of two doubles may also belong to the model space.

In general case the method may no longer be strictly size-extensive. For such Selected (SC)^{2}CI procedure one can however show (see ref. [86] and further *erratum* in ref. [85]) size-consistency as long as the selection procedure introduces only localized excitations consistently for the super- and sub-systems i.e. when a given $0_A j_B$ appears in the supersystem expansion j_B is also selected in the separate treatment of B. Let us finally remark that in such generalized (SC)^{2}CI method one may in fact work in a multireference space with a distinguished role of one of references (single reference dressing of MRCI matrix). Thus, according to the typology presented in the first chapter the (SC)^{2}CI method belongs to SR(MR) family.

The results of applications reported so far (mostly using the multireference i.e. SR(MR) variants) show that the (SC)^{2}CI method offers an interesting alternative to CCSD method for describing the PES of closed shell systems. It gives results of comparable quality and remains stable even for extremely stretched bonds where CCSD usually fails [86] [75]. The Selected (SC)^{2}CI method takes into account the most important linked contributions arising from $\hat{T}_2^2$ and other products of $\hat{T}_1$ and $\hat{T}_2$ which are included in the CCSD scheme, provided that the selection procedure introduces the most important triple and quadruple excitations (in this way certainly also $\hat{T}_3$, $\hat{T}_4$ contributions from those selected triples and quadruples are included). Actually all those linked contributions may also be included in the full CC type dressing as we discuss in Sec. 4.6.

Regarding the practical implementation, one does not perform explicit summation over all EPV contributions (the number of which is approximately proportional to n^3, n being the size of the basis set). One takes benefit of a trick proposed in the previous work for the infinite summation of EPV diagrams [113]. One-, two- and three-indices arrays store the contributions to the correlation energy of each spin orbital r

$$e_1(r) = \sum_{i(r)} c_i \langle 0|H|i\rangle \tag{4.56}$$

of each pair of spin orbitals r, s

$$e_2(r,s) = \sum_{i(r,s)} c_i \langle 0|H|i\rangle \tag{4.57}$$

and of each triplet of spin orbitals r, s, p

$$e_3(r,s,p) = \sum_{i(r,s,p)} c_i \langle 0|H|i\rangle \tag{4.58}$$

where summations over all determinants i are restricted as follows: $i(r)$ indicates that only determinants i defined by $\hat{e}_i$ involving r are taken, $i(r,s)$ is resricted to $\hat{e}_i$ involving r and s and respectively $i(r,s,p)$ is resricted to $\hat{e}_i$ involving r, s and p. Using these quantities, the calculation of the EPV$_i$ term becomes straigthforward since it only requires summations over the holes and particles of $|i\rangle \equiv |^{ab}_{mn}\rangle$

$$EPV_i = -e_1(m) - e_1(n) - e_1(a) - e_1(b) + e_2(m,n) + e_2(a,b) + e_2(m,a) + e_2(m,b) +$$

$$e_2(n,a)+e_2(n,b)-e_3(m,n,a)-e_3(m,n,b)-e_3(m,a,b)-e_3(n,a,b)+H_{0i}c_i \tag{4.59}$$

The above trick is used in all methods introduced in the subsequent chapters whenever the EPV contributions have to be evaluated.

It is worth noting finally that IEPA small CI problems may also be dressed in the same manner. For a given pair μ the model space is then defined by all determinants a pair μ consists of. All these determinants obtain then exactly the same dressing as in (SC)^{2}CI method including all EPV contributions. The only difference concerns the ground state determinant which is dressed in the diagonalization for the μth pair by an effect of outer space determinants i.e. by the quantity $\Delta_{00}^{\mu} = E_{corr} - \epsilon_{\mu}$. The mutual dressing of small CI problems provides a full set of the coefficients of the doubles.

4.5 Functional form of a dressing or generalized CPF method

We have been assuming so far that the dressing concerns the Hamiltonian matrix $\mathbf{H}$ in the eigenvalue problem $\mathbf{Hc} = E\mathbf{c}$. Introducing a generalized eigenvalue problem

$$\mathbf{H}\mathbf{c}^m = E_m\mathbf{S}\mathbf{c}^m \tag{4.60}$$

one may however consider a dressing of the overlap matrix. Let us recall that the term we want to effectively take into account by the dressing is $\sum_{\alpha\notin S\oplus s} H_{i\alpha}c_{\alpha}^m$. Defining an effective state-specific change (dressing) of the diagonal elements of the overlap matrix

$$^{\mathrm{eff}}S_{ii}^m = 1 - \Big(\sum_{\alpha\notin S\oplus s} H_{i\alpha}c_{\alpha}^m\Big)/(E_m c_i^m) \tag{4.61}$$

one reproduces the exact ith row of the matrix representation of the eigenproblem (4.28) in the normalized basis as ith row of (4.60)

$$\sum_{I\in S} H_{iI}C_I^m + \sum_{\substack{j\in s\\ j\neq i}} H_{ij}c_j^m + (H_{ii} - E_m\,{}^{\mathrm{eff}}S_{ii}^m)c_i^m = 0 \tag{4.62}$$

Notice that using the definition of the diagonal dressing Δ_{ii}^m one is led to

$$^{\mathrm{eff}}S_{ii}^m = 1 - \Delta_{ii}^m/E_m \tag{4.63}$$

So far the dressing Δ_{ii}^m remains unspecified and may take different forms i.e. single or multireference, CEPA or total CC type, as discussed in the next section.

Let us first consider the single reference situations i.e. we shall in general assume that the outer space determinants coefficients are defined with respect to a certain determinant $|0\rangle$ as specified by eq. (4.38). This includes all versions of the (SC)^{2}CI dressing for SRCI as well as for MRCI matrices when the redundancy contributions have to be substracted

$$\Delta_{ii} = E_{corr} + EPV_i + R_i \tag{4.64}$$

For the single reference ground state dressing of the (SC)^{2}CISD method we simply had $\Delta_{ii} = E_{corr} + EPV_i$. Assuming that the zero of the energy is equal to $E_0 = \langle 0|H|0\rangle$ one may write

$$^{\mathrm{eff}}S_{ii} = -(EPV_i + R_i)/E_{corr} \tag{4.65}$$

Let us recall that the expectation value of the energy $E_m = \langle \Psi_m | \hat{H} | \Psi_m \rangle / \langle \Psi_m | \Psi_m \rangle$ for the single reference ground state in a CI model space $S \oplus s$ may be rewritten in the intermediate normalization (with respect to E_0) as

$$E_{corr} = \frac{\langle \Phi_0 | \hat{H} - E_0 | \Phi_0 \rangle}{\langle \Phi_0 | \Phi_0 \rangle} = \frac{\langle 0 + \Psi_c | \hat{H} - E_0 | 0 + \Psi_c \rangle}{1 + \langle \Psi_c | \Psi_c \rangle} \tag{4.66}$$

where $|\Phi_0\rangle = \hat{P}_{S \oplus s} |\Psi_0\rangle = |0\rangle + |\Psi_c\rangle$.

The Coupled Pair Functional (CPF) method [76] and its further generalizations to multireference situations the Averaged Coupled Pair Functional (ACPF) [99] modify the norm $\langle \Psi_c | \Psi_c \rangle$ to achieve approximate size-extensivity of CISD and MRCISD method. This is based on the observation that the effect of higher excitations (necessary to restore size-extensivity) manifests in a partial cancellation of the norm denominator in (4.66). In other words the outer space effect may be effectively taken into account by a proper change of the normalization [76].

In the CPF method one defines pair functions (in the usual CEPA sense) and then norm of each pair function is multiplied by a factor g_μ chosen to satisfy certain limit conditions. In the original work [76] the numerator of eq. (4.66) is divided into parts coming from different pairs and then denominator is actually modified by $g_{\mu\nu}$ factors depending on the relation between pairs μ and ν. For the non-interacting pairs the $\mu\nu$ $(\mu \neq \nu)$ cross terms must vanish and by the requirement of a correct description of two electron (pair) systems and supersystems of identical non-interacting pair systems one is led to the choice

$$g_\mu = g = \frac{1}{n_p} = \frac{2}{N} \tag{4.67}$$

where n_p denotes number of pairs $2n_p = N$ and N is the number of electrons. The averaged common factor $g = 2/N$ was used in the ACPF methods where the notion of the electron pair is no longer employed and the method may be defined for MRCISD functional with respect to a multireference zeroth order energy [99]. In the single reference case the difference between CPF and ACPF comes simply from the less refined statistical limit employed by the latter method. The single reference ACPF energy functional reads

$$F_c^{\text{acpf}}[\Psi_c] = \frac{\langle 0 + \Psi_c | \hat{H} - E_0 | 0 + \Psi_c \rangle}{1 + g \langle \Psi_c | \Psi_c \rangle} \tag{4.68}$$

Both CPF and ACPF functionals are not bounded from below by the lowest eigenvalue of $\hat{H}$ but they are bounded by some finite real numbers [99].

Notice that in analogy to CPF or ACPF functionals one obtains a pseudofunctional form (since g_i are depenedent on c_i) of the (SC)^{2}CI or other single reference dressing while taking into account the effective change of the overlap matrix

$$F_{c_i}[E_{corr}] = \frac{\langle 0 + \sum_i c_i i | \hat{H} - E_0 | 0 + \sum_j c_j j \rangle}{1 + \sum_i g_i c_i^2} \tag{4.69}$$

Summations over i and j indices run in generally over determinants belonging to $S \oplus s$ space except that the ground state determinant and the factors g_i are given by

$$g_i = {}^{\text{eff}}S_{ii} = -(EPV_i + R_i)/E_{corr} \tag{4.70}$$

with R_i and E_i^{lkd} terms equal to zero in case of (SC)^{2}CISD dressing since $|\Psi_c\rangle$ contains only singles and doubles with respect to $|0\rangle$.

Let us verify the limit behavior of g_i factors defined in this way. For this purpose we shall test the (SC)^{2}CID dressing for a system of n_p identical non-interacting closed shell electron pairs (e.g. separated H_2 molecules). Thus, let the determinant $|i\rangle \equiv |^{ab}_{mn}\rangle$ belong to a pair $\mu \equiv mn$. Then certainly all excitations in μ are not possible on i whereas charge transfer excitations involving indices of different pairs have vanishing matrix elements and we simply have $EPV_i = \epsilon_\mu$. The supersystem correlation energy is equal to the sum of localized pair energies and finally

$$g_i = \frac{EPV_i}{E_{corr}} = \frac{\epsilon_\mu}{\sum_{\nu=1}^{n_p} \epsilon_\mu} = \frac{2}{N} \tag{4.71}$$

one gets the ACPF factor as expected. Let us notice that CEPA(2) dressing would lead to the same limit. It shows that limit requirements used in ACPF are very weak. On the other hand puting all EPV's as zero (CEPA(0) approximation) one gets the functional form of CEPA(0) found already by Čižek [76] [64]

$$F_c^{\text{cepa0}}[\Psi_c] = \langle 0 + \Psi_c|\hat{H} - E_0|0 + \Psi_c\rangle \tag{4.72}$$

Although the CPF renormalization is more refined than the ACPF one it also averages the exact determinant specific contributions $g_i = EPV_i/E_{corr}$. For heterogeneous electronic systems the g_i factors may vary within a given pair since they may involve excitations on orbitals of very different importance. This suggests the following generalization of the CPF method (which goes in the opposite direction comparing to the MRAQCC method [106]) referred to as Generalized Coupled Pair Functional (GCPF)

$$F_c^{\text{gcpf}}[\Psi_c] = \frac{\langle 0 + \Psi_c|\hat{H} - E_0|0 + \Psi_c\rangle}{1 + \sum_i g_i^{(0)} c_i^2} \tag{4.73}$$

where the determinant dependent $g_i^{(0)}$ factors are fixed as numbers before the dressing procedure (i.e. evaluated from the CI coefficients) or after the first or subsequent iteration of the dressing. It means that one has to diagonalize the CI matrix first (dressed or not) and then use such evaluated $g_i^{(0)}$ factors in the GCPF step. In this way one obtains a functional (and not only pseudofunctional) form of dressing differing from the CPF one by the fact that all diagonal elements of overlap matrix are different. It means that the density matrix and gradients may be evaluated analogously, as described in ref. [76]. Let us notice that g_i quantities should not strongly depend on the dressing steps since they are relative quantities. Fixing them from the pure CI coefficients (by simple evaluation of EPV contributions) should not bring large error. The merits and limitations of such a scheme have to be however confirmed by numerical applications.

The generalized CPF method is different from the original CPF by determinant-specific change of the norm. Simultaneously one avoids restrictions to pair and single reference closed shell theories as in ACPF. By considering the definition of eq. (4.70) it is clear that we may add for instance linked CC type corrections as well. The possibility of generalization to multireference situation becomes more transparent when one comes back to the general notation and writes

$$g_i^m = {}^{\text{eff}}S_{ii}^m \tag{4.74}$$

where ${}^{\mathrm{eff}}S_{ii}^{m} = 1 - \Delta_{ii}^{m}/E_m$. Then assuming that the outer space coefficients may be factorized into parts coming from different references - in consistency with eq. (4.39) one may write $\Delta_{ii}^{m} = \sum_{I \in S} \Delta_{ii,I}^{m}$. The diagonal element of the effective overlap matrix may be written as

$$ {}^{\mathrm{eff}}S_{ii}^{m} = (E_{corr}^{m} - \sum_{I \in S} \Delta_{ii,I}^{m})/E_{corr}^{m} \tag{4.75}$$

provided that the zero of energy is taken as E_m^0, obtained by the diagonalization in the reference space S. This means that we introduce a new basis in the reference space $|I_m^{'}\rangle = \sum_{I \in S} \tilde{C}_I^m |I\rangle$ where the $\tilde{C}_I^m$ coefficients come from the diagonalization. Then we shall employ (as in multireference version of ACPF) the MRCI expectation value for the energy (with respect to E_m^0) [99]

$$\begin{aligned} E_{corr}^{m} &= \frac{\langle \Phi_m | \hat{H} - E_m^0 | \Phi_m \rangle}{\langle \Phi_m | \Phi_m \rangle} = \\ &= \frac{\langle I_m^{'} + Q_S(I_m^{'}) + \Psi_c^m | \hat{H} - E_m^0 | I_m^{'} + Q_S(I_m^{'}) + \Psi_c^m \rangle}{1 + \langle Q_S(I_m^{'}) | Q_S(I_m^{'}) \rangle + \langle \Psi_c^m | \Psi_c^m \rangle} \end{aligned} \tag{4.76}$$

where this time $|\Phi_m\rangle = \hat{P}_{S \oplus s} |\Psi_m\rangle = |I_m^{'}\rangle + |Q_S(I_m^{'})\rangle + |\Psi_c\rangle$ and $Q_S(I_m^{'})$ denotes an orthogonal complement of $|I_m^{'}\rangle$ in the S space. Introducing the reference dependent and state dependent dressing of eq. (4.75) one may write in general

$$F_{c,m}^{\mathrm{gcpf}}[Q_S(I_m^{'}), \Psi_c] = \frac{\langle I_m^{'} + Q_S(I_m^{'}) + \Psi_c | \hat{H} - E_m^0 | I_m^{'} + Q_S(I_m^{'}) + \Psi_c \rangle}{1 + \langle Q_S(I_m^{'}) | Q_S(I_m^{'}) \rangle + \sum_{i \in s} g_i^{(0),m} c_i^2} \tag{4.77}$$

where the functional $F_{c,m}^{\mathrm{gcpf}}$ is state dependent since the renormalization factors g_i^m are state-specific. If all the coefficients g_i^m are assumed to be equal $2/N$ we again get the ACPF functional in its multireference version.

One should mention that it has been shown [67] [66] that even keeping the coefficient dependent factors g_i, which is equivalent to some CEPA or CC approximations, one may also define a functional which furnishes the starting equations (of the CEPA or CC type) when making it stationary with respect to variations of the coefficients. Nevertheless, it can be only done by the introduction of new parameters (Lagrangian multipliers) [67], which make the calculations of derivatives very expensive [66].

4.6 Coupled Cluster type of dressing

So far we have considered different approximations that may be located somewhere in between CI and CC approaches. It has been however suggested [104] that also the CC method itself (truncated at some level) may be transferred into a dressed CI problem. Speaking more precisely the CC non-linear terms may be all (and not only partially as in CEPA type dressing) added to the corresponding CI equations. In principle CC method equations may be solved by a self-consistent dressing of the corresponding CI matrix or, in other words, by a pseudoeigenvalue formulation.

Let us recall that in the CEPA method (dressing) actually only single product $c_i c_j$ ($|i\rangle \equiv |{}^{ab}_{mn}\rangle$, $|j\rangle \equiv |{}^{cd}_{pq}\rangle$) of the coefficients is taken into account. Including the whole non-linear term on the left hand side of eq. (4.41) one obtains the CCD equation through a pseudoeigenvalue problem with the following dressing

$$\hat{\Delta}_{ii} = (c^{ab}_{mn})^{-1} \sum_{\substack{p<q \\ pq \neq mn}} \sum_{\substack{c<d \\ cd \neq ab}} \langle {}^{ab}_{mn}|\hat{H}|{}^{abcd}_{mnpq}\rangle \, c^{ab}_{mn} \star c^{cd}_{pq} \tag{4.78}$$

or in equivalent notation

$$\hat{\Delta}_{ii} = (c_i)^{-1} \sum_{\substack{j \\ \hat{e}_j \hat{e}_i |0\rangle = |\alpha\rangle}} \langle 0|\hat{H}|j\rangle \sum_{\substack{(k,l) \\ P^{\alpha}_{kl} \hat{e}_k \hat{e}_l |0\rangle = |\alpha\rangle}} (-1)^{\sigma(P^{\alpha}_{kl})} c_k c_l \tag{4.79}$$

where P^{α}_{kl} denotes the permutation which leads to coincidence between $\hat{e}_k \hat{e}_l |0\rangle$ and $|\alpha\rangle$.

In the ref. [104] a column dressing rather than the diagonal one is employed to avoid division by c_i . The results obtained in such way are equivalent to the traditional CCD with the accuracy of 1 μH at all conformations (for which the traditional CCD method converges) of T-shaped Li_4 cluster used to test the numerical equivalence of those two methods of solving the CCD equations. Moreover through an eigenvalue problem formulation it is possible to get results when near degenaracies appear and the usuall methods for solving non-linear equations diverge [104].

In order to generalize the above procedure to CCSD one has to clearly distinguish the CI and CC amplitudes: c_i and t_i respectively. For the exact state $|\Psi_0\rangle$ and for the exact CI coefficients and CC amplitudes one gets a well-known cascade of equations defining amplitudes of the subsequent k-body operators $\hat{t}_i$ from the coefficients c_i and the amplitudes of l-body operators ($l < k$).

$$c^{a}_{m} = t^{a}_{m} \tag{4.80}$$

$$c^{ab}_{mn} = t^{ab}_{mn} + t^{a}_{m} t^{b}_{n} - t^{b}_{m} t^{a}_{n} \tag{4.81}$$

$$\ldots$$

As long as the s space consists of doubly excited determinants only the CID coefficients are good approximations to CCD amplitudes $t^{ab\,(0)}_{mn} = c^{ab\,(cid)}_{mn}$ and we actually do not need introduce different notation - the working equations define the amplitudes we deal with. In case of CCSD approximation we write in turn

$$t^{a\,(0)}_{m} = c^{a\,(ci)}_{m} \tag{4.82}$$

$$t^{ab\,(0)}_{mn} = c^{ab\,(ci)}_{mn} - t^{a\,(0)}_{m} t^{b\,(0)}_{n} + t^{b\,(0)}_{m} t^{a\,(0)}_{n} \tag{4.83}$$

to reformulate CCSD equations into the dressed eigenvalue problem. The set of eqs. (4.83) has to be solved after each iteration of the dressing procedure. Then the new dressing by the outer space triples and quadruples may be evaluated. Since the singles and triples are now present the dressing may be written in general as

$$\hat{\Delta}_{ii} = (c_i)^{-1} \langle i|\hat{H}(\hat{T}_1\hat{T}_2 + \frac{1}{3!}\hat{T}_1^3 + \frac{1}{2}\hat{T}_1^2\hat{T}_2 + \frac{1}{2}\hat{T}_2^2 + \frac{1}{4!}\hat{T}_1^4)|0\rangle \tag{4.84}$$

instead of (4.79). When the self-consistency is achieved we end up with the CCSD amplitudes.

As pointed out in ref. [104] the above procedure may be generalized to CCSDT and higher levels. Thus any single reference CC approximation may be transformed by the dressing of the corresponding CI matrix into an intermediate effective Hamiltonian pseudoeigenvalue problem to solve its working equations. Notice also that adding the linked terms arising from the quadratic terms of eq. (4.84) to the (SC)^{2}CI dressing considered in the previous section in the context of GCPF methods one obtains a pseudofunctional form of the CC dressing. Fixing ratios

$$g_i = (E_{corr} - \Delta_{ii})/E_{corr} \tag{4.85}$$

at some iteration one gets an approximate functional form of the CC method through an effective change of the norm of the CI energy functional.

Let us now consider the reverse situation. Starting from an arbitrary CI space $S \oplus s$ we define it as a model space of certain intermediate effective Hamiltonian which is supposed to reproduce the exact ground state energy. In order to built this effective operator we shall exponentialize the model space expansion $\hat{P}_{S\oplus s}|\Psi_0\rangle$ to take into account the effect of outer space determinants also exponentially factorized. In other words we shall define CC approximation adjusted to a given CI problem. Thus, we shall assume the following

$$|\Psi_0\rangle = \hat{\Omega}|0\rangle = e^{\hat{T}}|0\rangle \tag{4.86}$$

where the cluster operator $\hat{T}$ defined as

$$\hat{T} = (\sum_{I\in S,\, I\neq 0} t_I \hat{e}_I + \sum_{i\in s} t_i \hat{e}_i) + \sum_{\alpha\notin S\oplus s} t_\alpha \hat{e}_\alpha = \hat{T}_{S\oplus s} + \hat{T}_{out} \tag{4.87}$$

will be truncated in the following way

$$\hat{T}_{out} = 0; \qquad \hat{T} = \hat{T}_{S\oplus s} \tag{4.88}$$

instead of the traditional truncation according to the excitation level.

The coefficients of the model space determinants in the expansion of $|\Psi_0\rangle$ are defined by

$$c_i = \langle i|e^{\hat{T}}|0\rangle = \langle i|e^{\hat{T}_{S\oplus s}}|0\rangle \tag{4.89}$$

Eqs. (4.89) are certainly equivalent to the hierarchy of eqs. (4.80), (4.81) restricted to $S \oplus s$. Hence all the model space cluster amplitudes $\{t_I\}_{I\in S,\, I\neq 0}$; $\{t_i\}_{i\in s}$ are uniquely defined by the CI coefficients of $\hat{P}_{S\oplus s}|\Psi_0\rangle$. The outer space coefficients $c_\alpha = \langle\alpha|e^{\hat{T}}|0\rangle$ are approximated in terms of model space decompositions since we truncate the cluster operator to the model spaces $\hat{T} = \hat{T}_{S\oplus s}$

$$c_\alpha = \langle\alpha|e^{\hat{T}_{S\oplus s}}|0\rangle \tag{4.90}$$

Multiplying eq. (4.86) by $\hat{\mathbf{1}}$ we get

$$|\Psi_0\rangle = \hat{P}_{S\oplus s}\hat{\Omega}|0\rangle + \hat{P}_{out}\hat{\Omega}|0\rangle \tag{4.91}$$

Notice that the exponentialization of the second term of the right hand side of eq. (4.91) means that we epxonentially factorize our dressing operator (4.38) since the term $\hat{P}_{out}\hat{\Omega}$ corresponds to

$$\hat{\omega}^m = \sum_{\alpha \notin S \oplus s} \omega^m_\alpha |\alpha\rangle\langle 0| \tag{4.92}$$

Since we know the approximate form of the outer space coefficients the dressing term which should be added to the diagonal of $(S \oplus s)$CI matrix is evident. The ith row is for instance dressed according to eq. (4.32) for instance. All outer space determinants α that interact with i and do not belong to $S \oplus s$ contribute to the dressing with coefficients given by (4.90).

Certainly, in practice we start again from the pure $(S \oplus s)$CI coefficients as zeroth order approximation to the corresponding amplitudes (through the cascade of eqs. (4.89)) to evaluate the first dressing correction etc. until self-consistency is achieved. The size-consistency of such formulated procedure will be discussed in Part C of this chapter. This solution has been recently proposed by Adamowicz and Malrieu [111].

The single reference CC methods such as CCSD, CCSD(T) (CCSD with perturbative inclusion of Triples) are very efficient and routinely used in atomic and molecular calculations. The CCSD(T) method works well not only when dealing with well defined single reference case but often is able to reproduce reasonable PES [25]. Nevertheless when some double excitations become important (have large amplitudes) during changes of conformation it may diverge. In such situations, using the corresponding CCdCI procedure one may take benefit of the Davidson-like diagonalization procedures to obtain solutions in extreme regimes where the traditional ways of solving the CC equations fail. Moreover, various single reference dressing schemes appear to be very flexible and allow introducing lower approximations combined with CI spaces generated with respect to a set of references to account for the non-dynamical correlation.

However, in degenerate situations there is no escape from the inherently multireference description. The MRCC methods still face problems and are expensive. Therefore different multireference dressing schemes are of special interest.

Chapter 5

State-specific size-consistent multireference procedures

Using the theory of state-specific self-consistent intermediate Hamiltonians several dressings of a Multireference (MR) Singles and Doubles Configuration Interaction Hamiltonian matrix, which insure size-consistency (in the sense defined in Part A of this chapter) are proposed here. These methods are based on a Coupled Cluster (CC) type of factorization of coefficients of triples and quadruples. The most refined approach leads to a dressed CI formulation of a state-specific MRCC method [105]. Simpler dressings lead to revised formulations of the recently proposed MR Self-consistent Size-Consistent CI algorithm [85] and its lower approximation of the MRCEPA(0) type [91].

5.1 Introduction

The research on the multireference post Hartree–Fock methods is probably the most active in the contemporary Quantum Chemistry. The study of bond breaking and thus chemical reactivity must necessarily use such approaches. The well established perturbative or non-perturbative single reference methods may be safely applied to the ground state near equilibrium. Most of the excited states also demand a multiconfigurational zeroth order description and the study of excited potential energy surfaces requires very flexible approaches.

Generalization of the CI method for multireference cases is straightforward and the multireference configuration interaction (MRCI) method has been the most popular tool in molecular spectroscopy for a long time. In the MRCI method one may in principle remove an essential part of the extensivity error (due to the inclusion of higher order excitations), but it still suffers from the lack of rigorous extensivity, which is crucial when we deal with extended systems. The usual aposteriori Davidson-type corrections that take into account to some extent the unlinked contributions which should be removed [49] can only partially resolve the problem and many attempts have been made to modify the method to obtain nearly extensive results (for recent review see ref. [66]).

On the other hand the coupled cluster approach is more difficult to generalize to the multireference situations than the CI method. Several Multireference Coupled Cluster (MRCC) methods have been developed [65] and increasing number of applications to atomic [77] and molecular [77] [84] [97] systems is encountered in the recent literature. However, the computational complexity of these methods is very high and they are still facing many specific problems, as for instance the generalization of the Fock space methods to an arbitrary open shell reference state [78].

One should also mention another very active and rapidly growing field, namely the research on multireference perturbation theory (MRPT) [80]. Unfortunately MRPT methods become expensive already at third order, although second order results on huge CAS reference spaces are routinely available [81] and offer an interesting alternative to those of MRCC and MRCI. Nevertheless these methods are irrelevant in our present considerations.

Additional complexity that inevitably occurs in MRCC theories follows from the lack of the unambiguously specified separation of hole and particle states. One can in general introduce a fixed vacuum state (e.g. one of the references) for all reference functions or a reference-dependent vacuum with independent (although coupled) expansions with respect to each reference. As a result the genuine MRCC theories may be classified according to the above distinction, imposing different definitions of the wave operator $\hat{\Omega}$ [65]. Moreover both schemes are inherently of the multistate character. In the so-called Fock space approach (referred to also as the valence universal CC approach - VUCC) one has to consider a manifold of states with different number of valence electrons [79], while in the Hilbert space approach (state universal CC approach - SUCC) all states (with a constant number of valence electrons) corresponding to the different linear combinations of reference functions must be simultaneously handled [83].

In light of the above remarks the research on the single state or state-specific approximations is of a great practical meaning. Several such MRCC schemes have been originally proposed (see ref. [92]). They employ in general the state-specific wave operator acting on a single reference state but of the multideterminantal form.

Aforementioned procedures that rely on the underlying configuration space and try to modify MRCI method in order to obtain extensive results form another group of methods, which may be viewed as a group of approximate MRCC methods of the MRCEPA type [82] [66]. All those schemes are grounded in the cluster assumption and inherently use the CC factorization of the coefficients of triples and quadruples, reproducing to some extent the hierarchy of single reference CEPA methods. They are state specific similarily as the CI method.

The computational schemes presented in the subsequent parts of this chapter are aiming at obtaining a general state-specific MRCC theory starting from the recently proposed MR 'dressed' CI method (MR$(SC)^2$CI) [85], which may be considered as an exact MRCEPA procedure. The previously formulated single reference counterpart referred to as the size-consistent self-consistent CI method ($(SC)^2$CI) [86] insures the extensivity by a proper cancellation of unlinked effects. It was shown [90] that adding the linked effects (arising from the triples and quadruples as in CCSD method) one may eventually end up with a 'dressed' CI formulation of the corresponding SRCCSD approximation. As in the single reference case it is possible to include the linked contributions of triples and quadruples as well, obtaining 'dressed' CI formulation of a state-specific MRCCSD type approximation.

5.2 Multireference Self-consistent Size-consistent Configuration Interaction method

The recently proposed state specific self-consistent dressing of a multireference CI matrix and a simplified scheme of MRCEPA(0) type have been implemented by the author of this thesis and tested on a series of benchmark problems (H_4, H_2O, CH_2^+) [91]. The proof is given that this method is size-consistent provided that the reference space is separable (i.e. it contains all products of localized reference states). The structure of the method and its results are compared to those of other size-consistent multireference schemes. Its formal derivation in terms of CC expansion shall be presented in Sec. 5.3.

5.2.1 Theory

For a long time the multireference configuration interaction (MRCI) methods and especially the MRCISD, limited to the single and double excitations from the references, have been the most popular tool in molecular spectroscopy. The method is flexible and robust but it is not size-extensive nor separable. It would be worth to find a strategy insuring the size-extensivity and the separability of a modified MRCI algorithm.

The recently proposed self-consistent dressing of the CI matrix which makes the ground state description size-consistent [86], has been recalled in Chap. 4. However, the dressing was a single reference dressing with the ground state single determinant $|0\rangle$ playing special role and a generalization was needed. Its principle has been given in a recent paper [85] and its content is briefly described below.

A) The method

The references are labeled I,J, ... and define a reference space S of projector $\hat{P}_S$. The singles and doubles with respect to references are labeled i,j,k, ... and define the space s of projector $\hat{P}_s$. The method uses the concept of intermediate state-specific Hamiltonian [87] according to the general prescription given in Chap. 4. Here the MRCISD matrix is dressed to account for some non-linear terms. The diagonal dressing operator $\hat{\Delta}$ ($\Delta_{ij} = 0,\ i \neq j$) is employed, taking in the MR(SC)^{2}CI method the following form (for the desired state $|\Psi_m\rangle$)

$$\Delta_{ii}^m = \sum_I (h_I^m + EPV(i,I)^m + R(i,I)^m)\rho_{iI}^m \tag{5.1}$$

where h_I^m designate the effective energy shifts of the references

$$h_I^m = (\sum_{i \in s} c_i^m H_{Ii})C_{Im}^{-1} \tag{5.2}$$

whereas the terms

$$EPV(i,I)^m = -(\sum_{\substack{\hat{e}_k \\ \hat{e}_k|i\rangle=0}} c_k^m H_{I,\hat{e}_k I})C_{Im}^{-1} \tag{5.3}$$

$$R(i,I)^m = -(\sum_{\substack{\hat{e}_k \\ \hat{e}_k|i\rangle \in S\oplus s}} c_k^m H_{I,\hat{e}_k I}) C_{Im}^{-1} \tag{5.4}$$

take care of all exclusion effects (EPV) and redundancy effects (R) respectively [85]. Each determinant i receives a parentage ratio ρ_{iI}^m from all references $|I\rangle$ according to the following definition

$$\rho_{iI}^m = \frac{H_{iI}C_{Im}}{\sum_J H_{iJ}C_{Jm}} \tag{5.5}$$

This parentage relies on the first order perturbation arguments with the function $\hat{P}_S|\Psi_m\rangle$ taken as the zeroth order function [85]. The coefficients C_{Im} and c_i^m are here those of the solution of the dressed MRCISD eigenequation

$$|\Psi_m\rangle = \sum_{I\in S} C_{Im}|I\rangle + \sum_{i\in s} c_i^m|i\rangle \tag{5.6}$$

Further, the state index m will be conveniently put as superscript or subscript, depending on the number of other super- and subscripts.

The effective Hamiltonian matrix to be diagonalized is the dressed matrix

$$\hat{P}_{S\oplus s}(\hat{H} + \hat{\Delta}^m)\hat{P}_{S\oplus s} \tag{5.7}$$

Of course $\hat{\Delta}^m$ is dependent on the coefficients and the method is iterative. Notice that the method is uncontracted since it does not fix the components of the wavefunction in the reference space.

B) MRCEPA(0) type approximation

The above procedure may be viewed as a generalized multireference CEPA method, which takes into account all EPV effects. Actually the correct treatment of EPV and redundancy effects is either time or memory consuming and usually these terms are approximated to some extent [98] [107] [106]. Indeed, it sometimes happens that neglect of some of these effects may lead to accurate results due to a compensation of errors (namely with the neglect of higher order effects such as the linked effects of the triples). If we forget the EPV terms in eq. (5.1) we obtain a method that may be considered as an MR-CEPA(0)

$$\Delta_{ii}^{\prime m} = \sum_I (h_I^m + R(i,I)^m)\rho_{iI}^m \tag{5.8}$$

For practical efficiency one may be tempted to neglect also the redundancies and use

$$\Delta_{ii}^{\prime\prime m} = \sum_I h_I^m \, \rho_{iI}^m \tag{5.9}$$

The neglect of the redundancies may introduce undesirable but rare unlinked effects.

Notice that when the effect of the dressing is calculated using the eigenvector $|\Psi_m^{ci}\rangle$ of the undressed MRCISD matrix (at first iteration of the dressing procedure), we obtain a generalized Davidson-like correction

$$\Delta E^m = \langle\Psi_m^{ci}|\Delta^m|\Psi_m^{ci}\rangle = \sum_i \Delta_{ii}^{\prime\prime m} c_{im}^{(ci)\,2} \tag{5.10}$$

C) Generalization of the proof of size-consistency

Ref. [85] has only given a proof of size-consistency (called in ref. [85] separability) of MR(SC)^{2}CISD method when localized MO's are used and when the reference space involves only excitations of one subsystem. Actually it is possible to prove a more general theorem, stating that the method is separable (i.e. size-consistent according to the definition given in Chap. 1) if and only if the reference space is 'separable' i.e. it contains all products of subsystem references, insuring that any function belonging to S space is multiplicatively separable.

The MO's are supposed to be localized on non interacting subsystems A and B. Let us call I_A and J_B the references for independent description of A and B, S_A and S_B the corresponding reference spaces. A separable model space for the supersystem is a space built of *all* determinants $I_A J_B$

$$S_{AB} = S_A \otimes S_B \tag{5.11}$$

The SD space s_{AB} may be then written as a direct sum

$$s_{AB} = (s_A \otimes S_B) \oplus (S_A \otimes s_B) \tag{5.12}$$

i.e. it is composed of determinants of the forms $i_A J_B$ and $I_A j_B$, where $i_A \in s_A$, $j_B \in s_B$. Notice that the determinants $i_A j_B$ do not appear in the SD space for the supersystem. Although $i_A j_B$ determinants, where i_A and j_B are single substitutions in subsystems, are formally doubles, but they do not interact with references $I_A J_B$ of the supersystem $A \cdots B$ i.e. $\langle i_A | \langle j_B | \hat{H}_A + \hat{H}_B | I_A \rangle | J_B \rangle = 0$.

Consider now two eigenstates $|\Psi_p^A\rangle$ and $|\Psi_q^B\rangle$ of the dressed subsystems corresponding to the pth and qth roots, with E_p^A and E_q^B respectively, and let us form their product $|\Psi_m^{AB}\rangle$ (projected onto the supersystem model space $M_{AB} = S_{AB} \oplus s_{AB}$)

$$|\Psi_m^{AB}\rangle = \hat{P}_{M_{AB}} |\Psi_p^A\rangle |\Psi_q^B\rangle \tag{5.13}$$

We shall prove that this function, with coefficients satisfying (further we shall omit the upper indices for subsystems)

$$C_{I_A J_B, m}^{AB} = C_{I_A, p}^{A} C_{J_B, q}^{B}; \qquad c_{i_A J_B, m}^{AB} = c_{i_A, p}^{A} C_{J_B, q}^{B}; \qquad c_{I_A j_B, m}^{AB} = C_{I_A, p}^{A} c_{j_B, q}^{B} \tag{5.14}$$

is an eigenfunction of the dressed supersystem MRCISD matrix with an additively separable energy

$$E_m^{AB} = E_p^A + E_q^B \tag{5.15}$$

Hence, we shall prove that $|\Psi_m^{AB}\rangle$ corresponds to a certain (size-consistent) root m (with E_m^{AB}) of the supersystem eigenvalue problem.

Acting on such a trial function the dressing operator (5.1) may be expressed in a special way. For simplicity of notation we shall write the diagonal matrix elements with a unique index between square brackets (e.g. $\Delta_{[i]} \equiv \Delta_{ii}$). From eqs. (5.2) and (5.14) it follows that

$$h_{I_A J_B}^m = h_{I_A}^p + h_{J_B}^q \tag{5.16}$$

Using in turn eq. (5.5) it is easy to show that

$$\rho_{(i_A J_B)(I_A J_B)}^m = \rho_{i_A I_A}^p \tag{5.17}$$

Then using eqs. (5.16) and (5.17) one may demonstrate that

$$\Delta^{m,AB}_{[i_A J_B]} = \Delta^{p}_{[i_A]} + h^{q}_{J_B} \tag{5.18}$$

To prove it notice that determinant $i_A J_B$ has only parents of the type $I_A J_B$ and due to (5.17)

$$\Delta^{m,AB}_{[i_A J_B]} = \sum_{I_A} (h^{m}_{I_A J_B} + EPV(i_A J_B, I_A J_B)^m + R(i_A J_B, I_A J_B)^m)\rho^{p}_{i_A I_A} \tag{5.19}$$

Moreover

$$EPV(i_A J_B, I_A J_B)^m = EPV(i_A, I_A)^p \tag{5.20}$$

$$R(i_A J_B, I_A J_B)^m = R(i_A, I_A)^p \tag{5.21}$$

Therefore

$$\Delta^{m,AB}_{[i_A J_B]} = \sum_{I_A} (h^{p}_{I_A} + h^{q}_{J_B} + EPV(i_A, I_A)^p + R(i_A, I_A)^p)\rho^{p}_{i_A I_A}$$

$$= \Delta^{p}_{[i_A]} + h^{q}_{J_B} \tag{5.22}$$

since $\sum_{I_A} \rho^{p}_{i_A I_A} = 1$.

The subsystems eigenequation for i_A and J_B are respectively

$$(H_{[i_A]} + \Delta^{p}_{[i_A]} - E^{A}_{p})c^{p}_{i_A} + \sum_{I_A} H_{i_A I_A} C^{p}_{I_A} + \sum_{j_A \neq i_A} H_{i_A j_A} c^{p}_{j_A} = 0 \tag{5.23}$$

$$(H_{[J_B]} - E^{B}_{q})C^{q}_{J_B} + \sum_{K_B \neq J_B} H_{J_B K_B} C^{q}_{K_B} + \sum_{k_B} H_{J_B k_B} c^{q}_{k_B} = 0 \tag{5.24}$$

the last term in eq. (5.24) being $h^{q}_{J_B}$. Multiplying eq. (5.23) by $C^{q}_{J_B}$ and eq. (5.24) by $c^{p}_{i_A}$ and adding them one obtains the eigenequation for the determinant $i_A J_B$ and the eigenstate m in the dressed supersystem matrix

$$(H_{[i_A J_B]} + \Delta^{m,AB}_{[i_A J_B]} - E^{AB}_{m})c^{AB}_{i_A J_B, m} + \sum_{I_A} H_{i_A I_A} C^{AB}_{I_A J_B, m} +$$

$$+ \sum_{K_B \neq J_B} H_{J_B K_B} C^{AB}_{i_A K_B, m} + \sum_{j_A \neq i_A} H_{i_A j_A} c^{AB}_{j_A J_B, m} = 0 \tag{5.25}$$

provided that relations (5.14) - (5.15) hold. Hence our trial function is an eigenfunction of the dressed supersystem hamiltonian with the eigenvalue $E^{AB}_{m} = E^{A}_{p} + E^{B}_{q}$. □

Notice that if some of the determinants $I_A J_B$ are not included in the reference space (non-separable model space) determinants $i_A J_B$ are dressed in a way that breaks the separability because of inconsistent parentage in the super- and subsystems.

5.2.2 Computational aspects

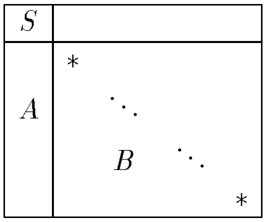
A) Practical implementation

Some aspects of the practical implementation of the multireference dressed CI (MR(SC)^{2}CI) procedure were already discussed in ref. [85]. They concern especially the CAS reference space when relevant simplifications are possible. Nevertheless, the general version of the MR(SC)^{2}CI algorithm has been implemented, without any restrictions for the reference space. The possibility to start for instance from the numerically selected references is a great advantage comparing to other multireference size-consistent schemes. It demands however careful construction of the program to avoid too high time and memory requirements (mainly due to complexity of redundancy effects evaluation).

Basically the MR(SC)^{2}CI procedure, as implemented by the author of this thesis [91], may be illustrated as follows:

1. Diagonalize the usual MRCISD matrix and for the desired eigenstate m use the coefficients $\{C_I^m, c_i^m\}$ to modify the matrix in the dressing step.

2. In the loop over block A (see figure below) find column contributions $h_I^m C_I^m$ and the parentage denominators $\rho_{iI}^m H_{iI}^{-1} = 1/\sum_J C_J^m H_{iJ}$. Build a vector containing non-zero H_{iI} elements and auxilliary addressing vectors to locate elements H_{iI}.

3. For a given determinant i find $\Delta_{ii}^{''m} = \sum_I h_I^m \; \rho_{iI}^m$ (loop over parents of i). In case of MRCEPA(0) (further labeled MR(SC)^{2}CI$_0$) go to 6.

4. In the loop over block B (which defines the overall complexity of the program) the redundancies and partially EPV's are subtracted by the analysis of interactions: a given H_{ij} and the relative operation $\hat{e}_{rel}$ defines (if possible) $k = \hat{e}_{rel} I$ determinant, giving an EPV contribution on common parents of i and j and R contribution on the remaining parents of i.
   It is crucial to find k (locate the corresponding matrix element H_{kI}, to undress by its effect) without next loop over determinants. In the special addressing vector one can keep the number of k (row of the matrix) in the cell given by the position of $\hat{e}_{rel}$ in the imaginary loop that generates all singles and doubles from I. Then having address of a given $\hat{e}_{rel}$ with respect to I one gets H_{kI}.

5. Subtract the rest of EPV's using one index e-arrays (actually two indices arrays $e(x, I)$, x being a hole or particle and I being a reference) - analogously to ref. [86]

6. Add such evaluated shifts Δ_{ii}^m to the diagonal (in case of direct CI algorithm first column would be more convenient [89]) of the CI matrix and diagonalize it. Repeat 2,..,6 with the new coefficients.

As one can see the memory requirements are specified by the dimension of the vector containing H_{iI}'s, whereas the complexity is roughly proportional to $N_{ref}n^4 \times \bar{N}_{inter} \times \bar{N}_p^2$ (loop in 4. to find redundancy contributions) where $\bar{N}_p$ means the averaged number of parents ($\bar{N}_p < N_{ref}$), $\bar{N}_{inter}$ the averaged number of interactions (non zero elements in rows of P_sHP_s; $\bar{N}_{inter} << N_{ref}n^4$) and n is the dimension of the basis set.

B) Convergence problems

In the formalism two possibly dangerous denominators appear. The first one is the C_{Im}^{-1} factor in eq. (5.2). As mentioned before this factor disappears in the final dressing. The second denominator appears in the definition of ρ_{iIm} (eq. (5.5)) and there is a danger of numerical instability or divergence if the quantity

$$A = \sum_J H_{iJ} C_J^m \tag{5.26}$$

is small (and may eventually become zero). Looking at the eigenequation for c_i:

$$\sum_{J \in S} H_{iJ} C_J^m + \sum_{\substack{k \neq i \\ k \in s}} H_{ik} c_k^m = (E_m - H_{ii}) c_i^m \tag{5.27}$$

we see that the definition of the parentage coefficient ρ_{iI}^m assumed that the coefficient c_i^m is essentially determined by the first summation

$$c_i^m = \frac{\sum_{J \in S} H_{iJ} C_J^m}{(E_m - H_{ii})} = \frac{\langle i | \hat{H} | \hat{P}_S \Psi_m \rangle}{(E_m - H_{ii})} \tag{5.28}$$

i.e. by the first order of perturbation theory from $\hat{P}_S|\Psi_m\rangle$ (otherwise also the second summation would appear in the definition of ρ_{iI}^m - see ref. [105]). If this assumption is valid, then the quantity A is small only when c_i^m is small too, in which case the dressing Δ_{ii}^m will have a negligable influence and may be forgotten.

A more troublesome situation happens when

$$|A| < |\sum_k H_{ik} c_k^m| \tag{5.29}$$

i.e. when the coefficient c_i^m is not determined by the first order perturbation or in other words by the interaction with the reference determinants, but by the second order (i.e. the interactions occurring in P_sHP_s). In this case the parentage becomes irrelevant and actually it is better to remove the determinant i from the first order interaction space. This leads to a selection procedure: if there exists a reference I (for a given determinant i) such that

$$|\frac{c_i^m H_{iI}}{\sum_{J \in S} H_{iJ} c_J^m}| > \eta \tag{5.30}$$

where η is a threshold equal to about one, then the determinant i is excluded from the s space. The above inequality comes from the analysis of the coefficients and parentage in terms of coupled cluster expansion, and it simply expresses the fact that all amplitudes should be smaller than one (see Sec. 5.3).

In practice in all numerical tests one observes that:

i) the number of determinants to be excluded remains small

ii) they do not contribute much to the energy

iii) the results are weakly dependent (to about 0.1 mH) on the precise value of the threshold η

For example, the MR(SC)^{2}CI energies for water molecule at stretched geometry (2R$_e$ - see Table 5.1) are 308.57, 308.59 and 308.61 mH with thresholds set to 0.5, 1.0 and 2.0 respectively. The number of removed determinants decreases from 24 with threshold 0.5 to 21 with the biggest one (the total number of MRCI determinants is 8616 in this case). The number of iterations of dressing procedure increases in turn from 3 to 8, showing that convergence is influenced by large amplitudes. On the other hand the speed of convergence is not dependent on the size of the problem. In fact the threshold was not always necessary (as for water at equilibrium geometry) but for the sake of consistency it was kept in all calculations.

In Sec. 5.3 we shall discuss a revised version of the parentage which starts directly from the psedoamplitudes and avoids the dangerous denominators in the present form.

5.2.3 Results

$(H_2)_2$, $(H_4)_2$ and $(H_2O)_2$ dimers have been used to verify the separability of the method in tests of the program. In all cases the localized MO's (subsystems separated at very large distance) and reference spaces of the CAS type located either on one (i.e. on the second system only the single reference determinant was involved) or on both subsystems have been used. In all those tests the additivity of the MR(SC)^{2}CI energy has been confirmed within tens of microhartrees.

H$_4$ The H_4 model has been extensively used in studies on single and multireference coupled cluster methods [95]. Already in the minimal basis set it represents an example of a non-trivial system owing to near degeneracies occurring in various geometrical arrangements. Here we use the rectangular arrangement called in ref. [95] P4, consisting of two parallel H_2 molecules with a fixed (stretched to 2 a.u.) bond length. The distance R between the two molecules is varied in this model, changing the degree of quasidegeneracy [95].

Using the minimal basis set and the *HOMO, LUMO* orbitals with two electrons as CAS reference space makes the dressing irrelevant. There is nothing left out of CASCISD space. Therefore the model has been modified slightly [91], by adding four additional $1s$ atomic orbitals: two bond centered and two located between H_2 molecules, forming a rectangular shape with $1s$ orbitals in the middle of every side. Such a system has the same properties as the original one - at the square geometry exact degeneracy occurs. However the precise values of energies are certainly slightly different (correlation energies are different on about 3.5 mH at some geometries). It makes the direct comparison impossible but in this case we are rather interested in the qualitative observations such as the percentage of the correlation energy reproduced and the general behavior (stability) of the method near degeneracy and in other regions.

Only distances R larger than 2 are considered here. The results are given together with the results of SSCCdCI method in Table 5.3 (see Sec. 5.3). The MRCI calculations with determinants $(\psi_1)^2(\psi_2)^2$ and $(\psi_1)^2(\psi_3)^2$ (ψ_i, $i = 1, \ldots, 8$ corresponds to the increasing orbital energy) as references are done, starting from the HF MO's for the triplet state. As one can see MR(SC)^{2}CI method adds about 0.1 mH to the corresponding MRCI energy and finally gives almost constant (99.7-99.9 %) part of the correlation energy independently on the degree of quasidegeneracy [91]. Two or three dressing iterations were sufficient to achieve the convergence in all cases. The results will be compared to those of different MRCC schemes in the next section.

The approximate MR(SC)^{2}CI$_0$ scheme gives quite significantly shifted results at all listed geometries, lower on about 5 mH, e.g. at $R = 2.001$ the MR(SC)^{2}CI$_0$ energy is equal to 119.118 mH with the first iteration (i.e. the new Davidson-like correction referred to as NQ) giving 118.940. This reflects the relative importance of EPV contributions for that small system.

H_2O Symmetrical bonds stretching in water molecule is very well known benchmark [108] that has been used to test and calibrate many methods. In the Table 5.1 the results of CASSCF, MRCI (MRCI + Davidson correction denoted by MRCI$_Q$) and MRLCC of Laidig and Bartlett [106], MRCPA(0) and MRCPA(2) of Tanaka *et al* [98], MRACPF of Gdanitz and Ahlrichs (reported in ref. [66]) and MRAQCC of Szalay and Bartlett [66] are collected (for explanation of acronyms see Discussion 5.2.4) together with the results for the new MR(SC)^{2}CI and MR(SC)^{2}CI$_0$ (as well as NQ) algorithms [91].

Comparing different forms of MRCEPA(0) scheme one notices that MR(SC)^{2}CI$_0$ is in the best agreement with FCI results (the error of the MRLCC method is smaller at equilibrium geometry but the potential curve deviates more from FCI and the errors for stretched bonds are larger). The behavior of MRCPA(0) is somewhat intriguing in the light of other results of MRCEPA(0) type approximation (see also similar calculations in DZP basis set by Ruttink *et al* [107]), underestimating the FCI results more than MRCPA(2). This seems to be in opposition to what we know about the cancellation of two errors, namely neglect of higher orders linked contributions of triples and quadruples and neglect of EPV terms, which usually manifests in more accurate results of this kind of approximations.

Regarding the 'exact CEPA' results of MR(SC)^{2}CI method one notices that it removes about 20 % of the MRCI errors in this case and it almost precisely reproduces the shape of the MRCI potential curve. Comparing to ACPF method and its more rigorous modification MRAQCC, one can estimate the influence of approximated treatment of EPV terms, which furnishes however better results due to the cancellation of errors mentioned above. The new extensivity corrections NQ are a little bit worse than the usual Davidson corrections but they are remarkably parallel to the FCI curve.

As was mentioned before the reference space does not have to be of the CAS type. To illustrate this another set of calculations for water molecule in the same basis set and the same geometries but with numerically selected reference determinants has been done. CIPSI procedure

Table 5.1: Results for water in DZ basis set.

Method	R_e	$1.5R_e$	$2R_e$
FCI	-76.15787	-76.01452	-75.90525
SCF	148.0	211.0	310.1
$CASSCF$	95.0	90.2	78.0
$MRCI$	2.1	2.3	2.0
$MRCI_Q$	-0.5	-0.3	0.1
$MRCPA(0)$	0.7	1.0	1.1
$MRCPA(2)$	0.1	0.5	0.8
$MRLCC$	-0.3	0.8	1.1
$MRACPF$	0.02	0.2	0.5
$MRAQCC$	0.6	0.8	0.9
This work			
NQ	-0.5	-0.4	-0.4
$MR(SC)^2CI_0$	-0.6	-0.3	-0.2
$MR(SC)^2CI$	1.7	1.9	1.5

Geometry from Saxe *et al* [108]. Energy errors in millihartrees, with respect to the full CI energy given in the first row. CAS reference space as specified in ref. [106].
MRCPA(0), MRCPA(2) - ref. [98]; *MRLCC* -ref. [106]; *CASSCF, MRCI,* $MRCI_Q$ - ref. [106] and this work; *MRACPF, MRAQCC* - ref. [66]; *SCF, FCI* - ref. [108].

[53] was used to generate the MRCI spaces: in the first step threshold η is set to choose references and in the second step all singles and doubles are produced with zero threshold. Two different selection thresholds $\eta = 0.05$ and $\eta = 0.03$ give MRCI results (with respect to FCI) 4.3 mH (5 ref.), 3.0 mH (24 ref.) and 5.9 mH (28 ref.) respectively in the first case and 1.5 mH (21 ref.), 1.9 mH (40 ref.) and 5.3 mH (49 ref.) with the second threshold. Numerical selection with usual HF orbitals provides here less balanced description of the potential surface, which is then reflected in the MR(SC)^{2}CI results being respectively: 2.4, 1.4 and 4.0 with the larger threshold (MR(SC)^{2}CI$_0$ approximation behaves surprisingly well in this case giving -0.8, -1.1 and -0.8 mH) and 0.5, 0.7 and 2.7 mH with the smaller one.

CH$_2$ The five lowest states (1-5) $^2A_1$ of CH_2, for which the FCI results are known [100], were used as a test of ACPF method by Gdanitz and Ahlrichs [99]. As in the ref. [100] one starts from the CASSCF orbitals optimized for the average of the five states and then MRCISD calculations are performed with all (76) determinants in the CAS space as references. The results are displayed in Table 5.2.

The MR(SC)^{2}CI method gives better results (except for the first root) than those of ACPF method. Even for the fifth root the error is only 1.3 mH (for this root ACPF fails because of much lower contribution of the reference determinants to the MRCI wavefunction - the same

Table 5.2: Results for CH_2 (1-5 $^2A_1$ states).

$Root$	1	2	3	4	5
FCI	−38.65031	−38.21354	−38.17007	−38.14301	−38.04032
$CASSCF$	68.5	80.0	77.8	82.0	101.2
$MRCI$	0.7	1.2	1.1	1.2	2.6
$MRCI_Q$	−1.1	−2.6	−2.0	−2.8	−8.8
$MRACPF$	−0.2	−0.7	−1.1	−0.8	−10.5
			$This\ work$		
NQ	−1.3	−3.9	−2.7	−8.1	1.7
$MR(SC)^2CI_0$	−1.4	−4.5	−2.6	−2.6	0.02
$MR(SC)^2CI$	0.4	0.6	0.5	0.2	1.3

Energy errors in millihartrees, with respect to the full CI energy given in the first row. Basis set, geometry and reference space from Bauschlicher *et al* [100].
MRACPF - ref. [99]; *CASSCF, MRCI,* $MRCI_Q$ - ref. [100] and this work; *FCI* - ref. [100].

concerns the generalized Davidson correction). The MR(SC)^{2}CI$_0$ approximation overestimates the FCI energy (except the last root) and is in worse agreement with it than the MRCI energies. As one see from the table the new form of extensivity correction (NQ) shows quite erratic behavior: for the third and fourth root the convergence of MR(SC)^{2}CI$_0$ is not monotonous and the first iteration gives results that are lower than those at convergence. For the fourth root NQ overestimates the FCI result by about 8 mH.

Summarizing briefly the numerical results one may state that MR(SC)^{2}CI method gives typically about 1/3 of the MRCI error and never exceeds the FCI energy. The MR(SC)^{2}CI$_0$ energies are also in quite good agreement with the exact ones, exceeding them by not more than 4.5 mH in all tested cases. The newly proposed extensivity correction may overestimate the FCI result quite severely as observed for CH_2. It suggests that one could rather use for such corrections estimates in terms of ratios of GCPF method, which are expected not to be much sensitive to the iteration process.

5.2.4 Discussion

The MR(SC)^{2}CI method should be compared with other size-consistent multireference algorithms (for recent and comprehensive review of the MRCEPA type of procedures see ref. [66]). First of all one should stress the fact that it cannot compete in principle with the MRCC methods since it only dresses the SD diagonal energies H_{ii} to eliminate the unlinked effects of the triples and quadruples, taking correctly into account all EPV effects, while CC methods also incorporate their linked contributions. For the single reference case these linked effects have also been treated through an approximate 'total dressing' [90]. A dressed-CI formulation of new state-specific MRCC algorithm is introduced in the next section.

The method most related to MR(SC)^{2}CI is probably that proposed by Ruttink *et. al.* [107]. Starting from a multiconfigurational function of the CAS type, they apply excitation operators and introduce diagonal energy shifts to functions obtained in this way. The energy shifts are the same for a given number of inactive holes and inactive particles in the excitation in a manner which is somewhat similar to that proposed in ref. [85] for the special case of CAS type reference function. This takes properly into account the possible redundancy effects. Some EPV terms are also treated correctly as a by-product of this procedure, but for the single reference case, when there are no redundancies it is equivalent to CEPA(0). In light of the above the MR(SC)^{2}CI method appears to be different in

i) its uncontracted character, revising the content of the reference space (which moreover does not have to be of the CAS type, although in such a case it may be not separable)

ii) accurate treatment of exclusion effects

Another quite similar proposal is due to Tanaka *et. al.* [98]. They develop a contracted scheme as well and the perturbative evaluation of the amplitudes. Finally the EPV's are not all included, as in the CEPA(2) method. The numerical tests of Tanaka and Ruttink's schemes are very encouraging.

Other MRCEPA schemes have been proposed recently by Fulde and Stoll [103] through a cummulant expansion and by Fink and Staemmler [101]. Both methods are contracted and make simplifications in the treatment of EPV's. The version proposed by Fink and Staemmler introduces a specific shift of each excited configuration in an averaged way.

Somewhat further from the present approach one might mention the QDVPT (quasi degenerate variation perturbation theory) of Cave and Davidson [102] which introduces a unique energy shift of the diagonal energies of the s space determinants, the ACPF method of Gdanitz and Ahlrichs [99] and MRAQCC method by Szalay and Bartlett [106] where the size-consistency of the CI is restored by a proper change of the normalization, as discussed previously. The last method improves the multireference linearized coupled cluster scheme (MRLCC) by inclusion of the EPV quadratic terms that violate the Pauli principle through the occupied MO's only. The ACPF method is not contracted (see Part A of this chapter).

The MR(SC)^{2}CI method is a unique proposal which conserves the symmetry of exclusion effects with respect to occupied and virtual MO's. One of its main advantages is its uncontracted character. Problems arise where the interaction between determinants in S and s spaces changes completely the content of the wavefunction in the reference space. A famous example is the LiF problem where the valence CASSCF level predicts a curve crossing at 8 bohr, which is pushed by the dynamical correlation to 12 bohr, changing completely the valence content of the wavefunction in a large domain of distances. Any contracted scheme will lead to dramatic artifacts in such problems [94]. Second very important feature of MR(SC)^{2}CI method is its flexibility: any (e.g. numerically selected) reference space is allowed. Another advantage is that one can use the usual CI machinery.

5.3 State-specific dressing of CC type of MRCISD matrix

5.3.1 Theory

In the following we shall use the same notation as previously and consider a multireference space S spanned by reference determinants I and a space s consisting of the determinants i obtained by the single and double substitutions $\hat{e}_k$ on the references that do not belong to S. The corresponding projectors are $\hat{P}_S$ and $\hat{P}_s$. Hereafter the model space will be chosen as the union $S \oplus s$ and the associated projector will be labeled $\hat{P}_{S\oplus s}$.

Let us also recall that for a state $|\Psi_m\rangle$ (ground or excited state)

$$|\Psi_m\rangle = \sum_{I\in S} C_I^m |I\rangle + \sum_{i\in s} c_i^m |i\rangle + \sum_{\alpha\notin S\oplus s} c_\alpha^m |\alpha\rangle \tag{5.31}$$

one can define an intermediate effective Hamiltonian $\hat{H}_{int} = \hat{P}_{S\oplus s}(\hat{H} + \hat{\Delta}^m)\hat{P}_{S\oplus s}$, where $\hat{\Delta}^m$ is a dressing operator, such that $\hat{P}_{S\oplus s}|\Psi_m\rangle$ is an eigenstate of $\hat{H}_{int}$ with the eigenvalue E_m, where $|\Psi_m\rangle$ and E_m are the exact eigenstate and eigenvalue $\hat{H}|\Psi_m\rangle = E_m|\Psi_m\rangle$. If one can write the coefficients of the outer space determinants as a sum over references

$$c_\alpha^m = \sum_{J\in S} c_{\alpha J}^m \tag{5.32}$$

then one may introduce a dressing of the first columns of the matrix (of the $P_{S\oplus s} H P_S$ block)

$$\Delta_{iJ}^{\prime m} = (\sum_{\alpha\notin S\oplus s} H_{i\alpha} c_{\alpha J}^m)(C_J^m)^{-1} \tag{5.33}$$

A) Coupled Cluster type of dressing

The above proposal is purely academic if one does not have a reasonable evaluation of the coefficients c_α^m of the outer space determinants. Remember that these determinants are here triples and quadruples with respect to at least one reference $I \in S$ and singles or doubles with respect to some of the singles and doubles $i \in s$.

The strategy to approximate the outer space coefficients is grounded in the CC assumption that the higher-body effects may be reasonably described in terms of two-body (and one-body) cluster amplitudes. In the following we shall assume a separate cluster expansion for every reference determinant $I \in S$, implying for the wave operator

$$\hat{\Omega}^m = \sum_I \hat{\Omega}_I^m \tag{5.34}$$

where $\hat{\Omega}_I^m$ has the form

$$\hat{\Omega}_I^m = e^{\hat{T}_I^m} \hat{P}_I \tag{5.35}$$

The index m means that we shall try to define a state-specific ansatz.

Since the MRCISD matrix is dressed the cluster operators are truncated and involve single and double excitations only

$$\hat{T}_I^m = \hat{T}_{1,I}^m + \hat{T}_{2,I}^m \tag{5.36}$$

In the second quantized form

$$\hat{T}^m_{1,I} = \sum_{\substack{a \\ p}} d^a_p(I)^m \hat{e}^a_p(I); \qquad \hat{T}^m_{2,I} = \sum_{\substack{a < b \\ p < q}} d^{ab}_{pq}(I)^m \hat{e}^{ab}_{pq}(I) \tag{5.37}$$

where the vacuum is given by $|I\rangle$ and defines the occupied and unoccupied indices. The symbols d^a_p, d^{ab}_{pq} are used instead of t^a_p, t^{ab}_{pq} since the coefficients d^a_p are not the independent variables and can be only interpreted as pseudoamplitudes. In the following we shall not use the second quantized notation but the notation introduced already in Chap. 4. If all singly and doubly excited determinants from the references are included in the model space we may equivalently define the cluster operators as follows

$$\hat{T}^m_{1,I} = \sum_{k \in S(I)} d^m_{kI} \hat{e}_k; \qquad \hat{T}^m_{2,I} = \sum_{k \in D(I)} d^m_{kI} \hat{e}_k \tag{5.38}$$

where $S(I)$ $(D(I))$ designates the set of all singly (doubly) excited with respect to $|I\rangle$ determinants and the summmations run over all determinants k belonging to $S(I)$ or $D(I)$ respectively. According to the convention defined in Chap. 4 the excitation operator $\hat{e}_k$ is by default defined with respect to the vacuum given by the reference function $\hat{e}_k$ acts on

$$\hat{e}_k |I\rangle = |k\rangle \tag{5.39}$$

and is determined as the relative excitation generating determinant $|k\rangle$ when acting on $|I\rangle$.

In the following we shall also assume that the disconnected part coming from $(\hat{T}^m_{1,I})^2$ for doubles may be neglected. This means that we have for the coefficients of determinants belonging to the s space

$$c^m_i = \sum_{\substack{I \\ |i\rangle = \hat{e}_i |I\rangle}} d^m_{iI} C^m_I \tag{5.40}$$

For the outer space determinants we shall have decompositions into products of two model space pseudoamplitudes d^m_{kI} only. Thus, the coefficients of the outer space determinants are approximated as

$$c^m_\alpha = \sum_{I \in S} 1/2 \sum_{\substack{(k,l) \\ |\alpha\rangle = \hat{e}_l \hat{e}_k |I\rangle}} d^m_{lI} d^m_{kI} C^m_I \tag{5.41}$$

where the second sum runs over all possible decompositions of $|\alpha\rangle$ with respect to a given reference determinant. More precisely, the summation over pairs (k, l) is restricted to such determinants k, l that define excitations (with respect to I) $\hat{e}_k$ and $\hat{e}_l$ such that

$$\hat{e}_\alpha |I\rangle = \hat{e}_l \hat{e}_k |I\rangle \tag{5.42}$$

Since the summation indices are not ordered here the factor 1/2 must appear.

Using eq. (5.41) one is able to introduce a column dressing according to eq. (4.35). The partitioning of the outer space coefficients with respect to references is clearly given by

$$c^m_{\alpha J} = 1/2 \sum_{\substack{(k,l) \\ |\alpha\rangle = \hat{e}_l \hat{e}_k |J\rangle}} d^m_{kJ} d^m_{lJ} C^m_J \tag{5.43}$$

leading to the column dressing of the form

$$\Delta'^{m}_{iJ} = 1/2 \sum_{\substack{(k,l) \\ |\alpha\rangle = \hat{e}_l \hat{e}_k |J\rangle}} H_{Jj}\, d^m_{kJ} d^m_{lJ} \tag{5.44}$$

where $|j\rangle$ and the matrix element H_{Jj} are given by the particular decomposition of $|\alpha\rangle$ involving $|i\rangle$

$$|\alpha\rangle = \hat{e}_j \hat{e}_i |J\rangle = \hat{e}_j |i\rangle \tag{5.45}$$

The crucial thing here is the definition of the state-specific cluster amplitudes d^m_{kI}. In the next section we shall partition the dressed MRCI coefficients c_i with respect to references to obtain pseudoamplitudes d^m_{iI}.

The procedure will be iterative: starting from the undressed CI coefficients (weighted to obtain initial pseudoamplitudes) we shall dress the MRCI matrix and diagonalize it. Using the new pseudoamplitudes (defined by the new coefficients) the new dressing may be evaluated. The process is repeated until self-consistency is achieved. In fact this new procedure, which will be referred to as State Specific Coupled Cluster type of dressing of the (MR)CI(SD) matrix (SS-CCdCI), is a generalization of the MR(SC)^{2}CI scheme [85] to a 'total' dressing that includes also linked effects of triples and quadruples (which are present in CCSD method) with respect to the reference determinants.

B) Reference-specific state-specific amplitudes

As one see from the previous subsection, we actually have in common with the state universal (or Hilbert space) approach, that we employ separated cluster ansätze for references [83] (for more detailed comparison see Sec. 5.3.1). In this approach a given determinant $i \in s$ is associated with a set of excitations with respect to its parents, by which it is obtainable, and the corresponding set of amplitudes coming from expansions for the parent references. All these amplitudes 'enter' the corresponding CI coefficient (as in eq. (5.40)).

Consider now the reference space of the CAS type. In such a case the most numerous inactive double excitations generate determinants having one parent only. The corresponding reference-specific amplitudes are thus uniquely defined by the proper coefficients. The semi-active excitations introduce determinants which might have many parents. One may hope that for such determinants an appropiate scheme of partitioning of their (dressed) MRCI coefficients will provide a reasonable definition of the reference-dependent amplitudes.

Let us try now to define state-specific and reference-specific psedoamplitudes for all substitutions leading from S to s space. Let us remark that the coefficients of the eigenvector $|\tilde{\Psi}_m\rangle$ of the dressed matrix $\tilde{H} = H + \Delta'^{m}$ satisfy the eigenequation (we shall omit the tilde accents for the actual coefficients)

$$\sum_{I \in S} \tilde{H}_{iI} C^m_I + \sum_{\substack{j \in s \\ j \neq i}} H_{ij} c^m_j + (H_{ii} - E_m) c^m_i = 0 \tag{5.46}$$

In the following we shall only dress the H_{iI} columns ($\tilde{H}_{iI} \neq H_{iI}$). This suggests two possibilities. One may either

i) consider the second sum of eq. (5.46) as negligible comparing to the first one (since the C_I's are supposed to be larger than c_i's, being respectively of order zero and order one in terms of a pertubative expansion from $|\tilde{\Psi}_m^0\rangle = \sum_{I\in S} C_I^m |I\rangle$); thus

$$c_i^m \simeq \frac{\sum_{I\in S} \tilde{H}_{iI} C_I^m}{E_m - H_{ii}} \tag{5.47}$$

This suggests to postulate $c_i^m \simeq \sum_{I\in S} \tilde{H}_{iI} C_I^m / \Delta\varepsilon_i^m$ or directly for the amplitudes

$$d_{iI}^m = \frac{\tilde{H}_{iI}}{\Delta\varepsilon_i^m} \tag{5.48}$$

where $\Delta\varepsilon_i^m$ is an effective energy difference. In order to obey strictly the basic eq. (5.46) $\Delta\varepsilon_i^m$ should be defined as

$$\Delta\varepsilon_i^m = (E_m - H_{ii}) - (\sum_{\substack{j\in s \\ j\neq i}} H_{ij} c_j^m)(c_i^m)^{-1} = \sum_{J\in S} \tilde{H}_{iJ} C_J^m (c_i^m)^{-1} \tag{5.49}$$

ii) or consider the full eq. (5.46) and write

$$c_i^m = \frac{\sum_{I\in S} \tilde{H}_{iI} C_I^m + \sum_{\substack{j\in s \\ j\neq i}} H_{ij} c_j^m}{(E_m - H_{ii})} = \frac{\sum_{I\in S} (\tilde{H}_{iI} + \sum_{\substack{j\in s \\ j\neq i}} H_{ij} d_{jI}^m) C_I^m}{(E_m - H_{ii})} \tag{5.50}$$

Hence the set of linear (assuming that one may replace $\tilde{H}_{iJ}$ by H_{iJ}) equations

$$d_{iI}^m = \frac{H_{iI} + \sum_{\substack{j\in s \\ j\neq i}} H_{ij} d_{jI}^m}{E_m - H_{ii}} \tag{5.51}$$

Although the second definition avoids the denominators of eq. (5.48), which may be problematic when some ratios c_j^m / c_i^m (see eq. (5.49)) become large, it is rather impractical, since it introduces M-body dressing operator (d_{iI} will not disappear even if $H_{iI} = 0$). Regarding the first possibility, for practical proposals one may be tempted to replace $\tilde{H}_{iJ}$ in the definition of eqs. (5.48) and (5.49) by H_{iJ}. This gives

$$d_{iI}^m = \frac{H_{iI} c_i^m}{\sum_{J\in S} H_{iJ} C_J^m} \tag{5.52}$$

and further we shall assume that the reference-specific pseudoamplitudes d_{iI}^m are given by eq. (5.52).

The above definition corresponds to the weighting of MRCI coefficients proposed for MR-CEPA type dressing (see ref. [85]). If one postulates that pseudoamplitudes d_{iI}'s come from a normalized, reference dependent partitioning of c_i coefficient

$$c_i^m = \sum_I \varrho_{iIm} c_i^m; \qquad \sum_I \varrho_{iIm} = 1 \tag{5.53}$$

and on the other hand $c_i^m = \sum_I d_{iI}^m C_I^m$ and the definition of eq. (5.52) is employed, one is led to

$$\varrho_{iIm} = \frac{H_{iI} C_I^m}{\sum_{J \in S} H_{iJ} C_J^m} \tag{5.54}$$

which is partitioning suggested by eq. (5.46), proposed already in ref. [85]. However, as mentioned before, the resulting d_{iI}'s may become undesirably large (or even infinite) when the quantity $\sum_{J \in S} H_{iJ} C_J^m$ becomes small (or even zero) due to cancelation of contributions with different signs.

In MR(SC)^{2}CI (see ref. [91]) procedure the amplitudes are chacked and when some of them become large, the underlying determinants are removed from the MRCI space. As the numerical tests show, such a selection concerns only very small part (usually much less than 1%) of all determinants and, insuring nice convergence, does not influence practically the final results, which are very encouraging [91]. Therefore this definition is kept as a basic one, although several other definitions may also be proposed. One might use the equal weight approximation for instance ($d_{iI}^m = c_i^m / (C_I^m N_p)$ where N_p denotes the number of parents of i), for the rare determinants governed by the second sum of eq. (5.46).

Another possibility, which scales the quantities d_{iI} not only on the strength of the interaction H_{iI} but also on the energy gap between H_{ii} and H_{II}, is to start from amplitudes optimized variationally in 2 dimensional CI space spanned by $|I\rangle$ and $|i\rangle$

$$\begin{array}{cc} & \begin{array}{cc} |I\rangle & |i\rangle \end{array} \\ \begin{array}{c} |I\rangle \\ |i\rangle \end{array} & \begin{pmatrix} H_{II} & H_{Ii} \\ H_{iI} & H_{ii} \end{pmatrix} \end{array}$$

and then normalize them, in order to satisfy eq. (5.40)

$$d_{iI}^m = d_{iI}^{1,m} c_i^m / C_I^m; \qquad d_{iI}^{1,m} = \frac{|d_{iI}^{0,m}|}{\sum_J |d_{iJ}^{0,m}|}; \qquad d_{iI}^{0,m} = \frac{c_i^0}{C_I^0} \tag{5.55}$$

where C_I^0, c_i^0 are variationally optimized coefficients. The primitive amplitudes $d_{iI}^{0,m}$

$$d_{iI}^{0,m} = \frac{\Delta - \sqrt{\Delta^2 + 4H_{iI}^2}}{2H_{iI}} \quad \text{when } H_{iI} \neq 0; \quad 0 \ \text{ otherwise} \tag{5.56}$$

where $\Delta = H_{ii} - H_{II}$, depend on both H_{iI} and $H_{ii} - H_{II}$ (with the latter one in the numerator). These amplitudes remain (in absolute values) smaller than one, provided that we always take the root with C_I^0 larger than c_i^0. They never diverge – if one refers to Quasi Degenarate Perturbation Theory [87] the wave operator $\Omega_I^{(1)}$ at first order defines $d_{iI} = H_{iI}/(H_{II} - H_{ii})$, which diverges when $H_{II} = H_{ii}$.

Anyway, when some single or double substitutions have much larger coefficients than those of reference determinants $c_i >> C_I$ (i.e. we actually face the intruder states problem) some amplitudes might again become very large and destroy the convergence. The remedy for that is the carefull and reasonable choice of the reference space. Let us notice however that when $c_i \simeq C_I$ i.e. when c_i is of the same same order of magnitude as C_I^0 nothing wrong should happen.

It is not always possible to choose the reference space in such a way to avoid the intruders. Therefore some dumping of large amplitudes would be desirable. There are three situations, which should be distinguished and may be treated in a different manner:

i) $c_i >> C_I$ and I is the only parent of i. Then there is no risk of small denominator and the definiton (5.52) may be applied.

ii) consider the following model situation: let determinant i has two parents 1 and 2, $c_i >> C_1$ and $c_i \simeq C_2$. Then we may shift the large amplitude d_{i1} on ref. 2 as follows

$$c_i = d_{i1}C_1 + d_{i2}C_2; \qquad d_{i1} = 1 + \Delta_{i1}$$

$$d'_{i1} = 1; \qquad d'_{i2} = d_{i2} + \Delta_{i1}\frac{C_1}{C_2}$$

where $\Delta_{i1}(C_1/C_2)$ remains small since $C_2 >> C_1$.

iii) i has many parents, but for all of them the relation $c_i >> C_I$ hold. In such a (rare) situation i might be treated as having no parents. With this special status i would be dressed according to (SC)^{2}CI scheme, with respect to a reference determinant K chosen such that $C_K > c_i$ and i being triple or quadruple from K.

To choose the best definition of pseudoamplitudes further numerical tests of the method are needed.

C) Comparison with MRCC formalisms

The MRCISD matrix dressed in this way may be considered as providing an eigenvalue (pseudoeigenvalue, due to the dependence on the coefficients) formulation of a state-specific Multireference Coupled Cluster Singles and Doubles (MRCCSD) method. To remain closer to the rigorous CC approaches one may distinguish the amplitudes of the single substitutions

$$a^m_{kI} = d^m_{kI} \tag{5.57}$$

if $\hat{e}_k$ is a single substitution with respect to I ($\hat{T}^m_{1,I} = \sum_{k\in S(I)} a^m_{kI}\hat{e}_k$) and introduce the proper amplitudes of the double substitutions

$$b^m_{lI} = d^m_{lI} \; - 1/2 \sum_{\substack{(i,j) \\ \hat{e}_l = \hat{e}_i\hat{e}_j}} a^m_{iI}a^m_{jI} \tag{5.58}$$

where $i, j \in S(I)$. Eq. (5.44) takes only into account $\hat{T}_1\hat{T}_2$ and $\hat{T}_2^2$. In order to strictly obey the assumed exponential form of $\hat{\Omega}$ one should also introduce the amplitudes of the triples coming from the third power of single substitutions ($\hat{T}_1^3$) and the amplitudes of quadruples coming from $\hat{T}_1^2\hat{T}_2$ and $\hat{T}_1^4$. However, one may notice that this rigorous formulation differs from the simplest one only by effects of orders higher than two in the wave function and higher than four in the energy. In fact the restriction of $\hat{T}$ to $\hat{T}_1$ and $\hat{T}_2$ is an approximation which omits some important 4^{th} order corrections for the energy i.e. the linked contributions of the triples. The exact treatment of 5th (resp. 6th) order corrections coming from $\hat{T}_1^2\hat{T}_2$ and $\hat{T}_1^3$ (resp. $\hat{T}_1^4$) is not worthwhile in view of this error, except if some $\hat{T}_1$ terms have very large amplitudes.

The present state-specific formulation differs deeply from the dominant trends in the MRCC research. Most of them (as mentioned in the introduction) belong to two families which assume a very universal structure of the wave operator. The first one (VUCC or Fock space CC) exploits the Valence Universal wave operator [88] and is less important in our considerations. The second family (SUCC or Hilbert space CC) is based on the reference dependent choice of the vacuum. The wave operator is expressed as a sum of independent wave operators defined with respect to reference determinants

$$\hat{\Omega} = \sum_{K}^{d} e^{\hat{T}_K} \hat{P}_K \tag{5.59}$$

Such an ansatz was originally proposed by Jeziorski and Monkhorst [83]. Its basic feature is that we have as many independent amplitudes as the number of references times the number of amplitudes in the single reference case. As a result we have to consider the manifold of all d states simultaneously to build enough equations. The so-obtained amplitudes are reference-dependent but they are state universal i.e. they are solutions of the whole system of equations (for all roots in the model space).

The philosophy, which is behind the SS-CCdCI procedure, is a very modest one, since one defines a state-specific expansion for the wave operator $\hat{\Omega}$. The ansatz we exploit here has a form similar to that of SUCC approach

$$\hat{\Omega}^m = \sum_{I \in S} e^{\hat{T}_I^m} \hat{P}_I \tag{5.60}$$

where the summation over I is restricted to S space only (and does not concern the whole model space $S \oplus s$) and $\hat{T}_I^m$ denote state-specific operators with reference-dependent pseudoamplitudes obtained by a proper partitioning (different for different states) of dressed CI coefficients with respect to reference determinants.

In the theory of traditional effective Hamiltonians the effect of outer space determinants is first included in the effective Hamiltonian (via the wave operator), which is then diagonalized in the model space to provide d exact roots. Here we not only restrict (as usually for the intermediate effective Hamiltonians) the manifold of states to be reproduced to the main model space. Another basic difference follows from the fact that the wave operator of eq. (5.60) is not built explicitly. However the self-consistent dressing of MRCISD matrix, as defined in Sec. 2B, is equivalent to the traditional scheme

$$\hat{H}_{\text{eff}} \hat{P}_S |\Psi_m\rangle = \hat{P}_S \hat{H} \hat{\Omega}^m \hat{P}_S |\Psi_m\rangle = E_m \hat{P}_S |\Psi_m\rangle \tag{5.61}$$

$$\hat{P}_{S\oplus s}(\hat{H} + \hat{\Delta}^m)\hat{P}_{S\oplus s}|\Psi_m\rangle = E_m \hat{P}_{S\oplus s}|\Psi_m\rangle \tag{5.62}$$

From the practical point of view it means that instead of a set of non-linear equations we may use the CI machinery for the pseudoeigenvalue problem, which is hoped to be much less troublesome. On the other hand, as the model space is now formally much larger, one has to take care of redundancy effects coming from $\hat{T}_1\hat{T}_2$, $\hat{T}_2^2$ ..., since for instance some quadruples with respect to one of the references may still belong to the s space, being doubles from other references. This may complicate the algorithm, especially for incomplete reference spaces.

Several other state-specific MRCC schemes have been originally proposed (see ref. [92]). In the very beginning of the MRCC development Sinanoglu and Silverstone [93] proposed a state-specific expansion, but it was rapidly recognized that their formulation introduced an exceedingly large number of amplitudes which could not be defined. Linearized version of MRCC Hilbert space method (LMRCC) with prediagonalization, proposed by Laidig and Bartlett in the early 80's [106], in which the manifold of states is decoupled, also belongs to this category. Recently Piecuch and Adamowicz [92] have formulated another practical proposal referred to as State Selective (SS) CC method, which employs the single reference formalism with a state-specific selection of higher rank excitations introduced by model space determinants. Li and Paldus [92] have proposed another, spin-adapted SSCC method using unitary group approach. The SS-CCdCI scheme is truly multireference i.e. the amplitudes are reference dependent, there is no determinant playing distinguished role and it does not need prediagonalization.

D) Size-consistency

Starting from the the MRCISD coefficients, which are not size-extensive, one repeats the dressing procedure until self-consistency is achieved. One may demonstrate that the method is size-consistent i.e. at convergence the desired eigenstate of the dressed MRCISD matrix provides additively separable energy for a supersystem $A \cdots B$ composed of two non-interacting systems A and B.

We shall employ the same assumptions and notation as for the generalization of the proof of separability of MR(SC)^{2}CI method. Thus, the MO's are supposed to be localized on subsystems A and B and the reference space for the supersystem is supposed to be separable

$$S_{AB} = S_A \otimes S_B \tag{5.63}$$

The Singles and Doubles from the references form space s_{AB} and the supersystem model space is $M_{AB} = S_{AB} \oplus s_{AB}$. Again $i_A j_B$ determinants, where i_A and j_B are single substitutions in subsystems, which are formally doubles, but they do not interact with references $I_A J_B$ shall not be included in the model space. In order to restore size-consistency in case of these special doubles included one would have to take into account the disconnected part of the coefficients of doubles in eq. (5.40).

Consider now, as previously, two eigenstates $|\Psi_p^A\rangle$ and $|\Psi_q^B\rangle$ of the dressed subsystems corresponding to the pth and qth roots, with E_p^A and E_q^B respectively, and let us form their product $|\Psi_m^{AB}\rangle$ (projected onto the supersystem model space M_{AB})

$$|\Psi_m^{AB}\rangle = \hat{P}_{M_{AB}} |\Psi_p^A\rangle |\Psi_q^B\rangle \tag{5.64}$$

We shall prove that this function, with coefficients satisfying eqs. (5.14) (thus being products of the corresponding subsystem determinants coefficients) is an eigenfunction of the dressed supersystem MRCISD matrix with an additively separable energy

$$E_m^{AB} = E_p^A + E_q^B \tag{5.65}$$

The demonstration might follow the proof of separability of MR(SC)^{2}CI method, in which the dressing operator is factorized into non-interacting subsystems contributions. It is possible however to prove a more general statement which may be applied to both methods (and other similar procedures) in order to show their size-consistency.

Definition 5.3.1 *We shall call an intermediate, effective Hamiltonian* $\hat{H} + \hat{\Delta}^m$ **separable**, *if it accepts function* Ψ_m^{AB} *as an eigenfunction with an additive eigenvalue* $E_m^{AB} = E_p^A + E_q^B$.

Let now the indices ρ, τ refer to localized determinants belonging to S, s or outer spaces.

Definition 5.3.2 *The outer space coefficients are* **factorizable** *if for each supersystem determinant* $\alpha_{AB} = \rho_A\tau_B$ *its coefficient is the product of coefficients of the corresponding subsystem determinants:* $c^m_{\alpha_{AB}} = c^p_{\rho_A} c^q_{\tau_B}$.

Statement 5.3.1 *If the outer space coefficients are factorizable the* $\hat{H} + \hat{\Delta}^m$ *is separable.*

Proof Let, according to our assumptions, the coefficients of determinants belonging to the model space M_{AB} be factorized as specified in eq. (5.14) and those of the outer space determinants be factorizable according to definition 5.3.2. $\hat{H} + \hat{\Delta}^m$ is diagonalized in the model space, so that we have to consider equations for two types of determinants, namely I_AJ_B and i_AJ_B. For the latter one we have (for simplicity of notation we shall write the diagonal matrix elements with a unique index between square brackets - e.g. $H_{[I]} \equiv H_{II}$)

$$(H_{[i_AJ_B]} - E_m^{AB})c^{AB}_{i_AJ_B,m} + \sum_{I_AK_B} H_{i_AJ_B,I_AK_B} C^{AB}_{I_AK_B,m} + \sum_{\alpha} H_{i_AJ_B,\alpha} c^{AB}_{\alpha,m} +$$

$$+ \sum_{K_B\neq J_B} H_{J_BK_B} C^{AB}_{i_AK_B,m} + \sum_{j_A\neq i_A} H_{i_Aj_A} c^{AB}_{j_AJ_B,m} = 0 \tag{5.66}$$

where α may represent two types of outer space determinants, namely those which are triples or quadruples on A (resp. B) i.e. of the type α_AJ_B (resp. $I_A\alpha_B$) and those which are singles or doubles on both A and B, i.e. of the type i_Aj_B. Eq. (5.66) is satisfied for $E_m^{AB} = E_p^A + E_q^B$ when the assumed factorization of the inner and outer space coefficients holds

$$[(H_{[i_A]} - E_p^A)c^p_{i_A} + \sum_{I_A} H_{i_AI_A} C^p_{I_A} + \sum_{\alpha_A} H_{i_A\alpha_A} c^p_{\alpha_A} + \sum_{j_A\neq i_A} H_{i_Aj_A} c^p_{j_A}]C^q_{J_B} +$$

$$[(H_{[J_B]} - E_q^B)C^q_{J_B} + \sum_{K_B\neq J_B} H_{J_BK_B} C^q_{K_B} + \sum_{k_B} H_{J_Bk_B} c^q_{k_B}]c^p_{i_A} = 0 \tag{5.67}$$

since the quantities in the square brackets [ ] are zero (the are the eigenequations for i_A and J_B respectively, in the dressed separated subsystems). For the lines corresponding to the reference functions I_AJ_B analogous derivation is straightforward. □

Thus, as long as we approximate the c_α's in a separable manner, the method is size-consistent and statement 5.3.1 provides a universal tool, which enables to check size-consistency of dressing procedures. We shall prove now that our definition of outer space determinants coefficients used for CC type dressing has the desirable property. The proof goes through the following steps

i) notice that

$$d^m_{i_AJ_B,I_AJ_B} = d^p_{i_AI_A} \tag{5.68}$$

This is a direct consequence of the definition of the amplitudes (5.52) and of the fact that i_AJ_B interacts only with references of the type K_AJ_B

$$d^m_{i_AJ_B,I_AJ_B} = \frac{H_{i_AI_A} \cdot c^p_{i_A} C^q_{J_B}}{\sum_{K_A} H_{i_AK_A} \cdot C^p_{K_A} C^q_{J_B}} \tag{5.69}$$

It is easy to show that amplitudes resulting from variationally optimized primitives (see eq. (5.55)) satisfy eq. (5.68) as well. Since

$$\Delta = H_{[i_AJ_B]} - H_{[I_AJ_B]} = H_{[i_A]} - H_{[I_A]} \tag{5.70}$$

we have

$$d^{0,m}_{i_AJ_B,I_AJ_B} = d^{0,m}_{i_A,I_A} \tag{5.71}$$

and then

$$d^{1,m}_{i_AJ_B,I_AJ_B} = \frac{|d^{0,m}_{i_AJ_B,I_AJ_B}|}{\sum_{K_A} |d^{0,m}_{i_AJ_B,K_AJ_B}|} = \frac{|d^{0,m}_{i_A,I_A}|}{\sum_{K_A} |d^{0,m}_{i_A,K_A}|} \tag{5.72}$$

Therefore eq. (5.68) holds again.

ii) as pointed out previously there are two types of outer space determinants, namely α_AJ_B (resp. $I_A\alpha_B$) and i_Aj_B. Using the definition of the c_α's (eq. (5.41)) and eqs (5.68), (5.40) one may show for both types the desired factorization

$$c^m_\alpha = \sum_{I_AK_B} \sum_{\substack{(k,l) \\ |\alpha\rangle = \hat{e}_k\hat{e}_l|I_AK_B\rangle}} d^m_{(\hat{e}_kI_AJ_B)I_AK_B} d^m_{(\hat{e}_lI_AJ_B)I_AK_B} C^m_{I_AK_B} =$$

for α_AJ_B

$$\begin{aligned} &= \sum_{I_A} \sum_{\substack{(k,l) \\ |\alpha_AJ_B\rangle = \hat{e}_{k_A}\hat{e}_{l_A}|I_AJ_B\rangle}} d^m_{(\hat{e}_{k_A}I_AJ_B)I_AJ_B} d^m_{(\hat{e}_{l_A}I_AJ_B)I_AJ_B} C^m_{I_AJ_B} = \\ &\sum_{I_A} \sum_{\substack{(k,l) \\ |\alpha_A\rangle = \hat{e}_{k_A}\hat{e}_{l_A}|I_A\rangle}} d^p_{(\hat{e}_{k_A}I_A)I_A} d^p_{(\hat{e}_{l_A}I_A)I_A} C^p_{I_A} C^q_{J_B} = c^p_{\alpha_A} C^q_{J_B} \end{aligned} \tag{5.73}$$

and for i_Aj_B

$$\begin{aligned} &= \sum_{I_AK_B} \sum_{\substack{(k,l) \\ |i_Aj_B\rangle = \hat{e}_{k_A}\hat{e}_{l_B}|I_AK_B\rangle}} d^m_{(\hat{e}_{k_A}I_AK_B)I_AK_B} d^m_{(\hat{e}_{l_B}I_AK_B)I_AK_B} C^m_{I_AK_B} = \\ &\sum_{I_A} d^p_{i_AI_A} C^p_{I_A} \sum_{K_B} d^q_{j_BK_B} C^q_{K_B} = c^p_{i_A} c^q_{j_B} \end{aligned} \tag{5.74}$$

Thus the dressing is separable and our trial function (eq. (5.64) is an eigenfunction of the dressed supersystem hamiltonian with the eigenvalue $E^{AB}_m = E^A_p + E^B_q$.

5.3.2 Numerical example

In the following the same modifications of H_4 model in the rectangular arrangement called in ref.[95] P4 as in the Sec. 5.2.3 are used to illustrate the efficiency of the present proposal [105].

The results are given in Table 5.3. The MRCI results are already quite good in this case and the dressings do not contribute too much. As one can see the SS-CCdCI method adds from 0.07 mH (at 2.001 and 10.0 au) to 0.17 mH (at R =5.0 au) to the corresponding MRCI energy and finally gives almost constant (99.64-99.98 %) part of the correlation energy independently on the degree of quasidegeneracy. This is an important message. The MRCC results remain between MRCI and FCI ones in the whole region: for strong and weak quasidegeneracy as well. Near square geometry the MRCC correction is about 1/5 of the difference between MRCI and FCI, whereas for large geometries it riches 4/5. However, the difference between MRCC and MR(SC)^{2}CI is biggest in the intermediate region (it is equal to 0.96 mH at R =5.0) where the total contribution of MRCC method (with respect to MRCI) is biggest as well.

Table 5.3: Results for H_4.

$R[au]$	$SCF[H]$	$MRCI$	$MR(SC)^2CI$	$SSCCdCI$ $(\%E_{corr})$	FCI
2.001	−1.947095	115.754	115.806	115.826 (99.70)	116.178
2.01	−1.947978	113.230	113.282	113.302 (99.70)	113.655
2.1	−1.955194	95.153	95.208	95.233 (99.64)	95.574
2.5	−1.961392	73.125	73.184	73.233 (99.68)	73.467
3.0	−1.937631	66.150	66.215	66.292 (99.78)	66.438
4.0	−1.866259	60.346	60.418	60.510 (99.87)	60.586
5.0	−1.806237	57.896	57.973	58.073 (99.94)	58.107
6.0	−1.765281	56.706	56.778	56.866 (99.97)	56.885
10.0	−1.708607	55.254	55.286	55.327 (99.98)	55.339

Modified P4 model (minimal basis set with additional $1s$ atomic orbitals located between H atoms) with two determinantal reference space. Energy differences in millihartrees, with respect to the SCF energy given in the first column - all signs reversed (except SCF energies). At R = 2.001 the coefficients of references are respectively $C_0 = 0.692$ and $C_{22}^{33} = -0.688$ [105].

At infinite separation (with both references localized on one H_2) both dressings give just the sum of single H_2 FCI energies (including the presence of bond centered orbitals), confirming separability (notice that CISD is equivalent to FCI for this 2 electron system). Two or three dressing iterations were sufficient to insure the convergence in all cases without any screening of large amplitudes (the definition of eq. (5.52) for d_{iI}'s were employed).

The VUCC (valence universal coupled cluster) or SUCC (state universal coupled cluster) results for the original P4 model have been recently reported [96] [97]. The exact values of correlation energies for the original P4 model are very similar to those presented here (the difference is not larger than 3.5 mH at all geometries). However, there is a qualitative difference in behavior of VUCC and SUCC results compared to SS-CCdCI procedure. The valence universal method gives 106.9, 107.9, 97.4, 64.8 and 50.3 (VU-CCSD/A of ref. [96]) percent of the correlation energy respectively at $R = 2.002$, 2.1, 3.0, 5.0, 10.0. In the case of SUCC (MRCCSD-3 of ref. [97]) one obtains 100.5, 100.0, 100.9, 107.0 and 109.9 percent of the correlation energy at the same geometries. As discussed in ref. [97] the linear MRCC theory (L-MRCCSD) has singularities at two geometries.

As one can see, in case VUCC and SUCC method one is not able to describe properly various degrees of quasidegeneracy (it concerns especially large geometries where the second and third determinants become relatively close) using reference space with two determinants. The SS-CCdCI approximated scheme does not face similar difficulties. As one can see from Table 5.3 there is no problem with stability of our multireference method in the whole region: for strong and weak quasidegeneracy as well. The error with respect to FCI results never exceeds 0.4 mH.

5.3.3 MRCEPA type of dressing revisited

The formalism introduced in this section may also be applied to MRCEPA type of dressing of the MRCISD matrix i.e. to the MR(SC)^{2}CI method. Using psedoamplitudes d^m_{kI} one may reformulate the MR(SC)^{2}CI procedure in the new language.

As we already know in the MR(SC)^{2}CI scheme the outer space contributions are effectively shifted to the model space $S \oplus s$ through a diagonal multireference dressing. Its form is fixed by assumption that one can weight MRCI coefficients with respect to reference determinants

$$\rho^m_{iI} = \frac{H_{iI} C^m_I}{\sum_J H_{iJ} C^m_J} \tag{5.75}$$

Then the energy shifts of the references are transfered to the determinants $i \in s$ according to the parentage ratios.

Now we can introduce the amplitudes d^m_{kI}, recognizing them easily in the definition of parentage ratios (5.75). It enables us to reformulate the method in this new language. In the following we shall neglect for simplicity the EPV and R terms.

As a direct consequence of definitions (5.52) and (5.75) we have

$$\rho^m_{iI} = \frac{H_{iI} C^m_I c^m_i}{\sum_J H_{iJ} C^m_J} (c^m_i)^{-1} = d^m_{iI} (c^m_i)^{-1} C^m_I \tag{5.76}$$

It implies that

$$\begin{aligned} \Delta^m_{ii} &= \sum_I h^m_{II} \rho_{iIm} = (c^m_i)^{-1} \sum_{I \in P(I)} d^m_{iI} \Big(\sum_{k \in s} c^m_k H_{Ik}\Big) = \\ &= (c^m_i)^{-1} \sum_{I \in P(i)} \sum_{k \in s} \sum_{J \in P(k)} d^m_{iI} d^m_{kJ} C^m_J H_{Ik} \end{aligned} \tag{5.77}$$

where by $P(i)$ we denote a set of parents of i and we used relation (5.44) $c^m_k = \sum_{J \in P(k)} d^m_{kJ} C_J$. Comparing this with eqs. (4.32) and (5.33) one sees that we postulate for the coefficient of the outer space determinant $|\alpha\rangle = \hat{e}_k \hat{e}_i |I\rangle$, obtained from $|i\rangle = \hat{e}_i |I\rangle$ by substitution $\hat{e}_k$; $|k\rangle = \hat{e}_k |I\rangle$, the following form of its partitioning

$$c^m_{\alpha I} = \sum_J d^m_{iI} d^m_{kJ} C^m_J \tag{5.78}$$

$$c_{\alpha}^{m} = \sum_{I} c_{\alpha I}^{m} \tag{5.79}$$

Chapter 6

Multistate dressing of CC type of the MRCI matrix

6.1 Introduction

We have been considering state-specific dressing procedures so far. The state-specific coupled cluster type dressing presented in the Sec. 5.3 is based on the partition of the state-specific CI coefficients, which are weighted to obtain state-specific reference-dependent pseudoamplitudes d_{iI}^{m}. In this way one decouples the Jeziorski–Monkhorst multistate ansatz into a state-specific decontracted expansion.

Some arguments rationalizing possible definitions of the partition of the CI coefficients giving reasonable approximation to the real CC amplitudes have been presented in the previous part of this chapter. It is a matter of numerical tests to choose the optimal weighting. However all the definitions of such obtained pseudoamplitudes must contain some arbitrariness.

To avoid this drawback one may propose another solution which is a generalization of the SS-CCdCI procedure to the multistate or state-universal dressing of MRCI matrix. The Multistate Coupled Cluster type dressing of Configuration Interaction matrix or MS-CCdCI procedure will start from a set of CI eigenstates to define a state independent dressing in terms of the state-universal CC amplitudes, coming from an adjusted CC approximation with the cluster operator truncated to excitations remaining in the CI space.

The new multistate CC type dressing is a generalization of the exponentialization of an arbitrarily truncated CI expansion and of the resulting CC type of dressing of the corresponding CI matrix with respect to the ground state introduced in Sec. 4.6. It was suggested and formally developed by J.P. Malrieu and the author of this thesis in collaboration with L. Adamowicz and R. Caballol [111].

6.2 The method

Let us consider a set $\mathcal{M}$ of M exact states $\mathcal{M} = \{|\Psi_m\rangle; m = 1, \ldots M\}$ having the expansions

$$|\Psi_m\rangle = \sum_{I \in S} C_I^m |I\rangle + \sum_{i \in s} c_i^m |i\rangle + \sum_{\alpha \notin S \oplus s} c_\alpha^m |\alpha\rangle \qquad m = 1, \ldots M \tag{6.1}$$

where S denotes a reference space and the s space may contain single, double and some higher excitations from the references.

Let us then assume that the reference space S was chosen in such a way that it consists of M determinants and diagonalization of the Hamiltonian in S space provides reasonable approximations to the exact states.

Now, we would like to define a multistate intermediate effective Hamiltonian $\hat{P}_{S\oplus s}(\hat{H}+\hat{\Delta})\hat{P}_{S\oplus s}$, where $\hat{\Delta}$ is a multistate dressing operator, such that

$$\hat{P}_{S\oplus s}(\hat{H}+\hat{\Delta})\hat{P}_{S\oplus s}|\Psi_m\rangle = E_m\hat{P}_{S\oplus s}|\Psi_m\rangle \qquad m=1,\dots M \tag{6.2}$$

$|\Psi_m\rangle$ and E_m being the exact eigenstates and eigenvalues

$$\hat{H}|\Psi_m\rangle = E_m|\Psi_m\rangle \qquad m=1,\dots M \tag{6.3}$$

We demand M exact energies (M-dimensional main model space) when acting on the M projections of the exact eigenvectors onto the model space.

The ith rows in the matrix representation of the eigenequations (6.3) are

$$\sum_{I\in S} H_{iI}C_I^m + \sum_{\substack{j\in s\\ j\neq i}} H_{ij}c_j^m + (H_{ii}-E_m)c_i^m + \sum_{\alpha\notin S\oplus s} H_{i\alpha}c_\alpha^m = 0 \qquad m=1,\dots M \tag{6.4}$$

To define a desired dressing procedure i.e. to reproduce these equations as dressed CI equations one has to transform the last summation of the above equations into proper matrix elements of a dressing operator, included in one of the first three terms. In other words one has to effectively shift the last summation into the model space.

In the following we shall postulate a dressing operator limited to the M columns involving the reference determinants

$$\Delta_{iI} \neq 0 \qquad I\in S \tag{6.5}$$

$$\Delta_{ik}^m = 0 \qquad k\in s \tag{6.6}$$

such that the eqs. (6.2) hold i.e. $\hat{P}_{S\oplus s}(\hat{H}+\hat{\Delta})\hat{P}_{S\oplus s}$ accepts $\hat{P}_{S\oplus s}|\Psi_m\rangle$ as eigenvectors with the energies E_m. The multistate dressing operator $\hat{\Delta}$ should carry an additional index $\hat{\Delta}^{\mathcal{M}}$ but for simplicity of notation we shall drop the $\mathcal{M}$ index in most cases.

Let us denote the terms to be included in the dressing by v_i^m

$$v_i^m = \sum_{\alpha\notin S\oplus s} H_{i\alpha}c_\alpha^m \tag{6.7}$$

Taking into account eqs. (6.2) and (6.5) - (6.6) one obtains

$$v_i^m = \sum_{I\in S} \Delta_{iI}C_I^m \qquad m=1,\dots M \tag{6.8}$$

Notice that if the s space involves all Singles and Doubles with respect to the reference determinants the quantities v_I^m and Δ_{IJ} are zero and the dressing operator only concerns $(N-M)M$ matrix elements between the S and s spaces, $N=\dim(s)$. In a general case, for example for numerical selection of the s space, the $\hat{P}_S\hat{H}\hat{P}_S$ block of the CI matrix would also remain undressed by the effect of the not selected Singles and Doubles because their approximated coefficients (in terms of model space ones) are not available.

Since we want to calculate v_i^m quantities we must estimate the coefficients c_α^m of the outer space determinants, coupled with the model space. As suggested in the introduction it will be based again on the CC arguments. More precisely we shall employ separate CC ansätze from the M reference determinants in the spirit of the state-universal MRCC scheme of Jeziorski–Monkhorst [83].

As pointed out in the part devoted to the state-specific solution, the Jeziorski–Monkhorst MRCC expansions for the M states read (FCC indicates not truncated expansion)

$$|\Psi_{FCC}^m\rangle = \sum_{I\in S} C_I^m e^{\hat{T}_I}|I\rangle \qquad m = 1,\dots M \tag{6.9}$$

in consistency with the postulated form of the wave operator

$$\hat{\Omega} = \sum_I \hat{\Omega}_I \tag{6.10}$$

where $\hat{\Omega}_I$ is

$$\hat{\Omega}_I = e^{\hat{T}_I}\hat{P}_I; \qquad \hat{T}_I = \sum_i t_{iI}\hat{e}_{iI} \tag{6.11}$$

Therefore the state-specific coefficients are products of state-specific coefficients C_I^m of the reference determinants and state independent amplitudes t_{iI}.

Here we shall assume the same form of the expansion of the wave function. There are however two basic differences in the present proposal. Now, in eqs. (6.9) the C_I^m coefficients will come from the diagonalization of the intermediate effective Hamiltonian in the model space $S\oplus s$ and not in the reference space S, which plays a role of a model space in the traditional approach. The truncation scheme of the cluster operators $\hat{T}_I$ is also different. As for the full exponentialization of an arbitrarily truncated CI expansion with respect to the ground state introduced in Sec. 4.6 we shall write

$$\hat{T}_I = \sum_{i\in s} t_{iI}\hat{e}_{iI} + \sum_{\alpha\notin S\oplus s} t_{\alpha I}\hat{e}_{\alpha I} = \hat{T}_{I,s} + \hat{T}_{I,out} \tag{6.12}$$

Then the cluster operators $\hat{T}_I$ are truncated by putting

$$\hat{T}_{I,out} = 0 \tag{6.13}$$

rather than according to the traditional excitation level scheme, giving

$$|\Psi_{CC}^m\rangle = \sum_{I\in S} C_I^m e^{\hat{T}_{I,s}}|I\rangle \qquad m = 1,\dots M \tag{6.14}$$

Our next step will be to evaluate CC amplitudes t_{iI} from the CI coefficients. For the exact expansions i.e. exact coefficients and amplitudes we have for the model space determinants $c_i^m = \langle i|\Psi_{FCC}^m\rangle$ or more explicitly

$$c_i^m = \sum_{I\in S} C_I^m \langle i|e^{\hat{T}_{I,s}}|I\rangle = \sum_{I\in S} C_I^m d_{iI} \tag{6.15}$$

where $d_{iI} = \langle i|e^{\hat{T}_{I,s}}|I\rangle$. The state independent pseudoamplitudes d_{iI} satisfy the known cascade of equations defining amplitudes of the subsequent k-body operators $\hat{e}_{iI}$ from the coefficients $d_{iI}^{(k)}$ and the amplitudes of l-body operators ($l < k$).

$$d_{iI}^{(1)} = t_{iI}^{(1)} \tag{6.16}$$

$$d_{iI}^{(2)} = t_{iI}^{(2)} + \frac{1}{2} \sum_{\substack{(j,k) \\ \hat{e}_{jI}\hat{e}_{kI}=\hat{e}_{iI}}} t_{kI}^{(1)} t_{jI}^{(1)} \tag{6.17}$$

$$\dots$$

Once we know the pseudoamplitudes d_{iI} we may easily find the true CC amplitudes t_{iI}. In order to determine d_{iI} coefficients one has to solve the set of linear equations (6.15).

For the outer space determinants in turn one has $c_\alpha^m = \sum_{I\in S} C_I^m \langle\alpha|e^{\hat{T}_I}|I\rangle$. This may be approximated in terms of the model space decompositions according to our truncation scheme $\hat{T}_{I,out} = 0$ by

$$c_\alpha^m = \sum_{I\in S} C_I^m \langle\alpha|e^{\hat{T}_{I,s}}|I\rangle = \sum_{I\in S} C_I^m d_{\alpha I} \tag{6.18}$$

where $d_{\alpha I} = \langle\alpha|e^{\hat{T}_{I,s}}|I\rangle$. The equations for the approximate CC amplitudes are defined by substituting c_i^m of eq. (6.15) and c_α^m of eq. (6.18) into eq. (6.4). On the other hand, taking into account eqs. (6.18) and (6.7) one may write in general

$$v_i^m = \sum_{I\in S} (\sum_{\alpha\notin S\oplus s} H_{i\alpha} d_{\alpha I}) C_I^m \tag{6.19}$$

and the matrix elements of the dressing operator are given by

$$\Delta_{iI} = \sum_{\alpha\notin S\oplus s} H_{i\alpha} d_{\alpha I} \tag{6.20}$$

Let us now assume that we start from the truncated $(S \oplus s)$CI problem having only approximate pure CI coefficients. They will be used as a starting point of the dressing procedure i.e. we shall assume that the CI expansions are first approximations to the CC ones i.e.

$$\hat{P}_{S\oplus s} |\Psi_{CC}^{m[0]}\rangle = |\Psi_{CI}^m\rangle \qquad m = 1, \dots M \tag{6.21}$$

Comparing the explicit definitions of these two expansions one gets

$$\hat{P}_{S\oplus s} \sum_{I\in S} C_I^{m[0]} e^{\hat{T}_{I,s}^{[0]}} |I\rangle = \sum_{I\in S} C_I^{m[ci]} |I\rangle + \sum_{i\in s} c_i^{m[ci]} |i\rangle \qquad m = 1, \dots M \tag{6.22}$$

and then by projecting on the left by $\langle I|$ and $\langle i|$ respectively, one obtains $C_I^{m[0]} = C_I^{m[ci]}$ and

$$c_i^{m[ci]} = \sum_{I\in S} C_I^{m[ci]} d_{iI}^{[0]} \qquad m = 1, \dots M \tag{6.23}$$

Then through the previously defined cascade of equations one may determine amplitudes $t_{iI}^{[0]}$, outer space coefficients $c_\alpha^{m[0]}$ and the dressing terms $v_i^{m[0]}$. The latter ones are to be added to the CI matrix through the dressing operator as specified in eq. (6.20). The so-dressed CI matrix is diagonalized furnishing the new set of dressed eigenvectors.

The procedure defined above is iterated until self-consistency is achieved. At convergence of the dressing procedure one obtains the desired amplitudes of the $|\Psi_{CC}^m\rangle$ CC expansions from the dressed CI coefficients resulting from the diagonalization of the dressed CI matrix. Certainly one gets simultaneously all M energies.

6.3 Discussion

The actual $S \oplus s$ space may be generated with respect to a larger reference space S', $S \subset S'$. Thus in general we shall consider two different partitions of the $S \oplus s$ subspace of the total Hilbert space

$$S \oplus s \equiv S' \oplus s' \tag{6.24}$$

$$S \subset S'; \qquad s' \subset s \tag{6.25}$$

where S' is a CI reference space and s' is a space of excited with respect to reference determinants defining the starting MRCI spaces. When interested in M roots only, $M < \dim(S')$, we shall distinguish a M-dimensional subspace $S \subset S'$ which could be called a generator space since the multireference expansions will be defined with respect to it. Nevertheless it will be called simply a reference space. In order to avoid confusion we shall denote the reference space of the MRCI problem as S' if it is different from S. The s space contains an orthogonal complement of S (in S') and the s' space.

As indicated above one may for instance consider a CASCISD matrix and a reference generators subspace $S \subset S' \equiv$ CAS with the corresponding manifold of states included in the full manifold derived from the CAS space. Nevertheless for logical and physical consistency it may be desirable to assume that if $S' \oplus s'$ is a MRCI space then all the reference determinant belong to S i.e. $S \equiv S'$ and $s \equiv s'$, in consistency with the state-universal scheme of Jeziorski–Monkhorst. Certainly, instead of enlarging the set of desired roots to the full CAS manifold one may rather restrict the MRCI space to the minimal $S \oplus s$, sufficient for desired accuracy of description of a given chemical system.

Regarding the practical implementation of the MSCCdCI procedure let us first remark that the use of the block Davidson-like diagonalization procedure [43] for simultaneous evaluation of M lowest eigenvectors of the (dressed or not dressed) CI matrix is desired for practical efficiency and is a natural choice. It is impossible to predict *a priori* the behavior of the block Davidson-like diagonalization method while adding the new multistate dressing terms to CI matrix. This may be only verified by numerical tests. Nevertheless, on the grounds of the experience accumulated so far concerning various dressing schemes one may expect good convergence as long as the dressing introduces reasonable corrections (i.e. relatively small with respect to the dominating matrix elements in the undressed eigenproblem).

One of the basic steps of the new procedure is determining pseudoamplitudes d_{iI} through the solution of the set of linear equation

$$\tilde{\mathbf{C}}\mathbf{d}^i = \mathbf{c}^i \tag{6.26}$$

where $(\mathbf{c}^i)^T = [c_i^1, c_i^2, \ldots, c_i^M]$ and $(\mathbf{d}^i)^T = [d_{i1}, d_{i2}, \ldots, d_{iM}]$ are vectors of the state-specific CI coefficients c_i^m and state independent pseudoamplitudes d_{iI} respectively. The $\tilde{\mathbf{C}}$ matrix contains the coefficients of the reference determinants C_I^m in expansions for states $m = 1, \ldots M$ and reads

$$\tilde{\mathbf{C}} = \begin{bmatrix} C_1^1 & C_2^1 & \cdots & C_M^1 \\ C_1^2 & C_2^2 & \cdots & C_M^2 \\ \vdots & & \ddots & \vdots \\ C_1^M & C_2^M & \cdots & C_M^M \end{bmatrix} \tag{6.27}$$

The small set of linear equation (6.26) has to be solved for every model space determinant i. It is an ideal task for parallel computers to perform the N matrix multiplications

$$\mathbf{d}^i = (\tilde{\mathbf{C}})^{(-1)}\mathbf{c}^i \tag{6.28}$$

N being the number of the model space functions. On the other hand these matrix manipulations do not significantly increase the computational cost as long as M remains small, which is true in most cases of potential applications. One may also remark that the $\tilde{\mathbf{C}}$ matrix may be easily reversed as long as the vectors $\mathbf{C}^m = [C_1^m, C_2^m, \ldots, C_M^m]$, $m = 1, \ldots M$ (rows of $\tilde{\mathbf{C}}$) are not linearly dependent.

Another important aspect of determining pseudoamplitudes d_{iI} via the set of equations (6.28) should be pointed out. The quantity d_{iI} is defined even if $H_{iI} = 0$ i.e. when determinant i is more than doubly excited with respect to a given reference determinant I. It means that many-body operators will appear in the wave operators $\hat{\Omega}_I$ increasing the complexity of the cascade of equations (6.17) for determining the real CC amplitudes t_{iI} from d_{iI}. If $|i\rangle$ is for instance a quadruple excitation from $|I\rangle$ the pseudoamplitude d_{iI} contains the amplitude t_{iI} traditionally included in CCSDTQ (by $\hat{T}_4$) and products of amplitudes associated with all decompositions of $|i\rangle$ in terms of other model space excitations defined with respect to $|I\rangle$. This is the price to pay for the adjusting of the CC expansion to the CI model space.

In light of the above it is clear that the method is practical only if the reference space S (the number of roots) and the CI space are sufficiently reduced. To avoid this bottleneck one could consider an approximate scheme of the MRCEPA type that would account only for the disjoint decompositions of the outer space excitations. It would then consist in writing

$$d_{\alpha I} = d_{iI}t_{jI}; \qquad |\alpha\rangle = \hat{e}_{jI}\hat{e}_{iI}|I\rangle \tag{6.29}$$

instead of $d_{\alpha I} = \langle\alpha|e^{\hat{T}_{I,s}}|I\rangle$ and would require only partial decomposition (up to the amplitudes of Doubles) of quantities d_{kI}.

It is worth to compare the new proposal to the previously defined computational schemes of the state-specific character. The new multistate procedure may be reduced to the single state or state-specific while restricting the number of desired roots to one ($M = 1$), provided that there is a dominating determinant in the expansion for this root.

Consider for instance the MRCI space $S' \oplus s'$ assuming that one of the determinants belonging to the CI reference space S' is a reasonable approximation to a given mth exact state and will form our one-dimensional reference generator subspace S. All other model space determinants will receive single pseudoamplitudes d_{iK} with respect to the reference determinant $|K\rangle \in S \equiv \{|K\rangle\}$ from which one can get real amplitudes of the single reference (more precisely SR(MR)) CC expansion for the mth state.

Thus the new MS-CCdCI algorithm in the limit of one state is different from the state-specific SS-CCdCI procedure for the same state, in which one introduces state-specific weighting of CI coefficients with respect to all CI references in S'. If the chosen state m is the ground state one obtains the ground state CC type dressing of an arbitrarily truncated CI matrix introduced in Sec. 4.6.

The size-consistency of the MSCCdCI procedure will be discussed in the next section.

6.4 Size-consistency

One may demonstrate that the method is size-consistent i.e. at convergence the desired M eigenstates of the dressed MRCI matrix furnish M additively separable energies for a supersystem $A \cdots B$ composed of two non-interacting systems A and B, under the following constraints:

- The MO's are supposed to be localized on subsystems A and B
- The reference space for the supersystem is supposed to be separable i.e. it is the tensorial product of subsytem reference spaces

$$S_{AB} = S_A \otimes S_B \tag{6.30}$$

- The respective subsystem spaces s_A and s_B contain all the Singles and Doubles from the references in S_A and S_B and may also contain some determinants, which are more than Doubles with respect to all references (selected for instance as Single or Double with respect to larger CI reference spaces S'_A or S'_B)
- The supersystem model space is $M_{AB} = S_{AB} \oplus s_{AB}$ where the space s_{AB} has the following structure

$$s_{AB} = (s_A \otimes S_B) \oplus (S_A \otimes s_B) \oplus P(s_A \otimes s_B) \tag{6.31}$$

  where $P(s_A \otimes s_B)$ denotes a subset of $s_A \otimes s_B$

$$\emptyset \subseteq P(s_A \otimes s_B) \subseteq s_A \otimes s_B \tag{6.32}$$

  Thus s_{AB} contains all the determinants of the form $i_A J_B$ and $I_A j_B$ and a part of composite excitations $i_A j_B \in s_A \otimes s_B$ introduced by the actual truncation scheme e.g. doubles composed of localized singles.

Consider now $M = P \cdot Q$ pairs of eigenstates $|\Psi_p^A\rangle$ and $|\Psi_q^B\rangle$ of the dressed subsystems corresponding to the pth, $p = 1, \ldots P$ and qth, $q = 1, \ldots Q$ roots, with E_p^A and E_q^B energies respectively. Let us form their products $|\Psi_m^{AB}\rangle$ (projected onto the supersystem model space M_{AB})

$$|\Psi_m^{AB}\rangle = \hat{P}_{M_{AB}}|\Psi_p^A\rangle|\Psi_q^B\rangle \qquad m = 1,\ldots M \tag{6.33}$$

where $p = 1,\ldots P;\quad q = 1,\ldots Q$. In the expansions of trial functions of the supersystem all the coefficients are products of the corresponding subsystem coefficients

$$C_{I_AJ_B,m}^{AB} = C_{I_A,p}^{A}C_{J_B,q}^{B}; \qquad c_{i_AJ_B,m}^{AB} = c_{i_A,p}^{A}C_{J_B,q}^{B}; \qquad c_{I_Aj_B,m}^{AB} = C_{I_A,p}^{A}c_{j_B,q}^{B} \tag{6.34}$$

and for $i_Aj_B \in M_{AB}$

$$c_{i_Aj_B,m}^{AB} = c_{i_A,p}^{A}c_{j_B,q}^{B} \tag{6.35}$$

We shall prove that the functions $|\Psi_m^{AB}\rangle$, $m = 1,\ldots M$ are eigenfunctions of the dressed supersystem MRCI matrix with additively separable energies

$$E_m^{AB} = E_p^A + E_q^B \qquad m = 1,\ldots M \tag{6.36}$$

where again $p = 1,\ldots P;\quad q = 1,\ldots Q$.

The demonstration will follow the proof of separability of SS-CCdCI method. We shall first generalize to multistate situations the previously introduced statement specifying sufficient condition for the size-consistency of the dressing procedures as factorizability of the coefficients of the outer space determinants. Then the desired property of the approximation to the coefficients of the outer space determinants will be verified.

The definitions of the separable dressing and factorizable coefficients may be generalized for the search of M states as follows

Definition 6.4.1 *We shall call a multistate intermediate, effective Hamiltonian* $\hat{H} + \hat{\Delta}^{\mathcal{M}}$ $\mathcal{M}$-**separable**, *if it accepts functions* Ψ_m^{AB} *as the eigenfunctions with the additive eigenvalues* $E_m^{AB} = E_p^A + E_q^B$ *for each* $m = 1,\ldots M$.

Definition 6.4.2 *The outer space coefficients are* $\mathcal{M}$-**factorizable** *if for each supersystem determinant* $\alpha_{AB} = \rho_A\tau_B$

$$c_{\alpha_{AB}}^{m} = c_{\rho_A}^{p}c_{\tau_B}^{q} \qquad m = 1,\ldots M$$

The indices ρ, τ refer to localized determinants belonging to S, s or outer spaces, as previously.

Statement 6.4.1 *If the outer space coefficients are* $\mathcal{M}$*-factorizable the* $\hat{H}+\hat{\Delta}^{\mathcal{M}}$ *is* $\mathcal{M}$*-separable.*

Proof We shall follow the logic of the proof of the statement 5.3.1. Let the coefficients of the determinants belonging to the model space M_{AB} be factorized as specified in eqs. (6.34) – (6.35) and those of the outer space determinants be $\mathcal{M}$-factorizable according to def. (6.4).

Now, three types of determinants belong to the model space, namely I_AJ_B, i_AJ_B and some i_Aj_B. The presence of the latter ones in the model spaces means that we have to consider another set of rows of the matrix representation of the pseudoeigenproblems (6.2)

$$(H_{[i_Aj_B]} - E_m^{AB})c_{i_Aj_B}^{m,AB} + \sum_{I_A} H_{i_AI_A}c_{I_Aj_B}^{m,AB} + \sum_{K_B} H_{j_BK_B}c_{i_AK_B}^{m,AB} +$$

$$+\sum_{k_A\neq i_A} H_{i_Ak_A}c^{m,AB}_{k_Aj_B} + \sum_{l_B\neq j_B} H_{j_Bl_B}c^{m,AB}_{i_Al_B} + \sum_{\alpha_A} H_{i_A\alpha_A}c^{m,AB}_{\alpha_Aj_B} + \sum_{\alpha_B} H_{j_B\alpha_B}c^{m,AB}_{i_A\alpha_B} = 0 \tag{6.37}$$

for each $m = 1,\dots M$. The summations over determinants k_Aj_B and i_Al_B respectively include both the model and outer space determinants of this form. Recall also that the diagonal matrix elements are denoted by a unique index between square brackets - e.g. $h_{[I]} \equiv h_{II}$.

Eq. (6.37) is satisfied for $E^{AB}_m = E^A_p + E^B_q$ when the assumed factorization of the model and outer space coefficients holds

$$[(H_{[i_A]} - E^A_p)c^p_{i_A} + \sum_{I_A} H_{i_AI_A}C^p_{I_A} + \sum_{\alpha_A} H_{i_A\alpha_A}c^p_{\alpha_A} + \sum_{k_A\neq i_A} H_{i_Ak_A}c^p_{k_A}]c^q_{j_B} +$$

$$[(H_{[j_B]} - E^B_q)c^q_{j_B} + \sum_{K_B} H_{j_BK_B}C^q_{K_B} + \sum_{\alpha_B} H_{j_B\alpha_B}c^q_{\alpha_B} + \sum_{l_B\neq j_B} H_{j_Bl_B}c^q_{l_B}]c^p_{i_A} = 0 \tag{6.38}$$

since the quantities in the square brackets [ ] are zero for each $p = 1,\dots P$, $q = 1,\dots Q$. They are simply the rows of the eigeneproblem for i_A and j_B respectively, in the dressed separated subsystems eigenequations.

For the rows corresponding to the reference functions I_AJ_B and i_AJ_B determinants the derivation is analogous to the previously considered in the context of the SSCCdCI procedure. The only difference is that in the i_AJ_Bth rows the model space determinants i_Aj_B resulting from the composite excitations will certainly not appear in the sum over outer spaces determinants, but in the separate term without any influence on the final conclusion since it may be factorized into subsystem contributions as the other terms. □

We shall prove now that our definition of the coefficients of the outer space determinants used for the multistate CC type of dressing has the desirable property. The proof goes through the following steps

i) Recall that the coefficients of the outer space determinants are given by eq. (6.18), which for the supersystem problem reads

$$c^m_{\alpha_{AB}} = \sum_{I_AK_B\in S_{AB}} C^{m,AB}_{I_AK_B} d_{\alpha_{AB},I_AK_B} \tag{6.39}$$

Since we have assumed that the coefficients of the reference determinants are $\mathcal{M}$-factorizable it is sufficient to prove that the supersystem pseudoamplitudes $d_{\alpha_{AB}I_AK_B}$ are $\mathcal{M}$-factorizable.

ii) Notice that there are in general six types of the supersystem outer spaces determinants: α_AK_B, $I_A\alpha_B$, α_Aj_B, $i_A\alpha_B$, $\alpha_A\alpha_B$ and $i_Aj_B \notin M_{AB}$. Thus it is sufficient to show that

$a)\quad d_{\alpha_AK_B,I_AK_B} = d_{\alpha_AI_A}$; $\qquad b)\quad d_{\alpha_Aj_B,I_AK_B} = d_{\alpha_AI_A}d_{j_BK_B}$;

$c)\quad d_{\alpha_A\alpha_B,I_AK_B} = d_{\alpha_AI_A}d_{\alpha_BK_B}$; $\qquad d)\quad d_{i_Aj_B,I_AK_B} = d_{i_AI_A}d_{j_BK_B}$;

iii) Consider now the supersystem model space pseudoamplitudes. For $i_A K_B$ we have

$$c^m_{i_A K_B} = \sum_{I_A J_B \in S_{AB}} C^m_{I_A J_B} d_{i_A J_B, I_A J_B} \qquad m = 1, \dots M \tag{6.40}$$

where on the right hand side the terms with $J_B = K_B$ may only survive, whereas for i_A in the subsystem problem

$$c^p_{i_A} = \sum_{I_A \in S_A} C^p_{I_A} d_{i_A I_A} \qquad p = 1, \dots P \tag{6.41}$$

Subsequently multiplying all equations of the above set of linear equations by $C^q_{K_B}$, $q = 1, \dots Q$ we see that $d_{i_A K_B, I_A K_B}$ and $d_{i_A I_A}$ satisfy the same set of linear equations. Thus, since the model space coefficients $c^m_{i_A K_B}$ and $C^m_{I_A K_B}$ are assummed to be $\mathcal{M}$-factorizable

$$d_{i_A K_B, I_A K_B} = d_{i_A I_A} \tag{6.42}$$

Similarly one can check that for $k_A l_B \in M_{AB}$ one has

$$d_{k_A l_B, I_A K_B} = d_{k_A I_A} d_{l_B K_B} \tag{6.43}$$

iv) From the model space pseuamplitudes we may then get the real cluster amplitudes. Through the cascade of equations (6.16), (6.17), ... we get for $i_A K_B$

$$t^{(1)}_{i_A K_B, I_A K_B} = d^{(1)}_{i_A K_B, I_A K_B} = t^{(1)}_{i_A I_A} \tag{6.44}$$

$$t^{(2)}_{i_A K_B, I_A K_B} = d^{(2)}_{i_A K_B, I_A K_B} - \frac{1}{2} \sum_{\substack{(k,l) \\ \hat{e}_{k_A I_A} \hat{e}_{l_A I_A} = \hat{e}_{i_A I_A}}} t^{(1)}_{k_A I_A} t^{(1)}_{l_A I_A} = t^{(2)}_{i_A I_A} \tag{6.45}$$

$$\dots$$

For the second type of model space determinants, namely $k_A l_B$ one gets that all amplitudes $t^{(k)}_{k_A l_B, I_A K_B}$ are zero. It is clear that for the amplitudes of singles $t^{(1)}_{k_A l_B, I_A K_B} = 0$ because such composite excitations are at least doubles. Then, because of eqs. (6.43) and (6.16), assuming additionally that the charge transfer excitations that would introduce another decompositions have zero amplitudes, we get

$$t^{(2)}_{k_A l_B, I_A K_B} = d^{(2)}_{k_A l_B, I_A K_B} - t^{(1)}_{k_A I_A} t^{(1)}_{l_B I_B} = 0 \tag{6.46}$$

It may be easily generalized for the amplitudes of an arbitrary k-body operator. This fact is crucial for separability.

v) Let us introduce now the partition of the cluster operators $\hat{T}_{I_A K_B, s_{AB}}$ into parts with localized subsystem excitations and composite excitations

$$\hat{T}_{I_A K_B, s_{AB}} = \hat{T}_{I_A K_B, s_{AB}}(A) + \hat{T}_{I_A K_B, s_{AB}}(B) + \hat{T}_{I_A K_B, s_{AB}}(AB) \tag{6.47}$$

where for instance the part involving only excitations localized on subsystem A is

$$\hat{T}_{I_A K_B, s_{AB}}(A) = \sum_{k_A} t_{k_A K_B, I_A K_B} \hat{e}_{k_A I_A} \tag{6.48}$$

Since, according to iv) $t_{k_A l_B, I_A K_B} = 0$ we have

$$\hat{T}_{I_A K_B, s_{AB}}(AB) = \sum_{k_A l_B} t_{k_A l_B, I_A K_B} \hat{e}_{k_A l_B, I_A K_B} = 0 \tag{6.49}$$

Moreover

$$\hat{T}_{I_A K_B, s_{AB}}(A) = \hat{T}_{I_A, s_A}; \qquad \hat{T}_{I_A K_B, s_{AB}}(B) = \hat{T}_{K_B, s_B} \tag{6.50}$$

vi) From the definition of the outer space pseudoamplitudes

$$d_{\alpha_{AB} I_A K_B} = \langle \alpha_{AB} | e^{\hat{T}_{I_A K_B, s_{AB}}} | I_A K_B \rangle$$

follows that they are expressed in terms of sums of products of CC supersystem model space amplitudes.

Let us first explicitly consider the first case, namely a) of ii)

$$\begin{aligned} d_{\alpha_A K_B, I_A K_B} = \langle \alpha_A K_B | e^{\hat{T}_{I_A K_B, s_{AB}}} | I_A K_B \rangle = \langle \alpha_A | e^{\hat{T}_{I_A K_B, s_{AB}}(A)} | I_A \rangle = \\ \sum_{\substack{(k_A, l_A, \dots) \\ |\alpha_A\rangle = \hat{e}_{k_A I_A} \hat{e}_{l_A I_A} |I_A\rangle}} t_{k_A K_B, I_A K_B} t_{l_A K_B, I_A K_B} \cdot \ \cdots \end{aligned} \tag{6.51}$$

In this case certainly only excitations localized on subsystem A may survive by projecting on the left. Finally taking into account eqs. (6.44), (6.45), ... we get

$$d_{\alpha_A K_B, I_A K_B} = d_{\alpha_A, I_A} \tag{6.52}$$

For the other types of supersystem determinants i.e. $\alpha_A j_B$, $\alpha_A \alpha_B$ and $i_A j_B \notin M_{AB}$ one needs the eq. (6.49) of v), which states that all composite excitations have vanishing cluster amplitudes. We shall also use eq. (6.50). Let us consider for instance the determinants $i_A j_B \notin M_{AB}$ i.e. the case d) of ii)

$$\begin{aligned} d_{i_A j_B, I_A K_B} &= \langle i_A j_B | e^{\hat{T}_{I_A K_B, s_{AB}}(A) + \hat{T}_{I_A K_B, s_{AB}}(B)} | I_A K_B \rangle = \\ &= \langle i_A | \langle j_B | e^{\hat{T}_{I_A K_B, s_{AB}}(A)} e^{\hat{T}_{I_A K_B, s_{AB}}(B)} | I_A \rangle | K_B \rangle = d_{i_A I_A} d_{j_B K_B} \end{aligned} \tag{6.53}$$

In the same way one may demonstrate that the sufficient conditions as specified in ii) are satisfied for the other types of outer space determinants.

Thus the MS-CCdCI dressing is separable and our trial functions (eq. (6.33) are eigenfunctions of the dressed supersystem hamiltonian with the eigenvalues $E_m^{AB} = E_p^A + E_q^B$.

Chapter 7

Size-consistent self-consistent combination of selected CI and PT

The computational cost of the dressing procedures presented in the previous chapters may increase too rapidly with the size of the considered systems to apply them to large molecules, even in case of the relatively simple methods of the type of MRCEPA. Further approximations may be still necessary. The flexibility of variuos dressing procedures opens a way to such approximated schemes.

In principle one may dress any kind of CI matrices e.g. even two by two matrix and then join it with a dressing of other CI problems. An example of such combined scheme, suggested by Malrieu and developed in collaboration with Heully and the author of this thesis [116], is used to illustrate the potential merits of such lower order methods.

The present chapter contains a brief presentation of a consistent combination of two computational schemes, namely a Selected (SC)^{2}CI method and a non divergent 2nd order perturbation evaluation (see below). The method treats Large Doubles (and any set of higher excited determinants) in a variational manner and the Small Doubles in a perturbative mode, with reciprocal dressings of the CEPA type and exact treating of the EPV terms. Strict separability into closed shell subsystems is insured if the localized MOs are used and the selection procedure introduces only localized excitations. The results and short discussion of some illustrative calculations (NH_3 in DZP basis set and H_2O in DZ basis set) are also presented.

7.1 The method

In the following we shall restrict ourselves again to the closed shell ground state problem. Let us consider a two dimensional CI problem spanned by the ground state determinant and a doubly excited determinant with respect to the ground state. This defines a CI model space $\{|0\rangle, |i\rangle\}$ and the outer space containing the rest of excited determinants. When we diagonalize two by two problems it provides a non-perturbative evaluation of the coefficients of doubly excited determinants called in ref. [115] an Independent Excitation Approximation (IEA). Then we may dress the small CI matrices

$$\begin{array}{c} |0\rangle \\ |i\rangle \end{array} \quad \begin{pmatrix} H_{00}+\delta_{00} & H_{0i} \\ H_{i0} & H_{ii}+\delta_{ii} \end{pmatrix}$$

by the effect of outer space detrminants using diagonal energy shifts

$$\delta_{00} = \sum_{j \neq i} c_j \langle 0|H|j\rangle \tag{7.1}$$

and

$$\delta_{ii} = \sum_{\substack{j \\ \hat{e}_j|i\rangle \neq 0}} c_j \langle 0|H|j\rangle \tag{7.2}$$

where $|i\rangle = \hat{e}_i|0\rangle$ certainly represents a redundancy effect. These dressings incorporate the effects of the doubles $|j\rangle = \hat{e}_j|0\rangle$ for $|0\rangle$ (except i) and of quadruples $\hat{e}_j|i\rangle$ for $|i\rangle$. When all coefficients of the doubles are evaluated from such dressed two by two problems one is led to a self-consistent and non-divergent evaluation of the 2nd order perturbation type correlation energies [112]. The cost of this self-consistent 2nd order (dressed IEA) procedure is a few times that of a classical MP2 calculation and the so far reported [112] behavior of the potential energy for single bond breaking is very encouraging. However, this simple method could not be applied to the breaking of multiple bonds.

As expressed in Chap. 4 the (SC)^{2}CI method may be applied to any selected CI including arbitrary fractions of the various classes of excitations with respect to the ground state determinant. It means that one can apply this procedure to improve results of a numerically selected CI method. A large fraction of doubles may be eliminated from the CI expansion because of small coefficients. Then one might think of obtaining c_i by a perturbative techique if $|c_i|$ is small. The traditional PT may however diverge and one may use the dressed IEA amplitudes instead, joining the two self-consistent size-consistent processes.

The algorithm proposed here is only relevant when selected doubles are included in the variational CI step, otherwise it reduces to the (SC)^{2}CI algorithm. On the other side it reduces to the dressed IEA method when the number of selected deteminants falls down to 2. The method goes continously from dressed IEA to (SC)^{2}CI by increasing the number of the doubles included in the variational CI. It is always possible to add the important triples and quadruples to this selected space. The method will be referred to as (SC)^2(CI+PT2) [116].

From the practical point of view for the determination of c_i it is sufficient to consider the shifted matrix

$$\begin{pmatrix} 0 & H_{0i} \\ H_{i0} & H_{ii}-H_{00}+\delta_{ii}-\delta_{00} \end{pmatrix}$$

where

$$\delta_{ii}-\delta_{00} = -\sum_{\substack{j \\ \hat{e}_j|i\rangle = 0}} c_j \langle 0|H|j\rangle + c_i \langle 0|H|i\rangle$$

and then the dressing reduces to the summation over EPV terms. Regarding the CI matrix we do the same as in the (SC)^{2}CI procedure, so that the dressing has the form

$$\delta_{ii} = \sum_{\substack{j \\ \hat{e}_j|i\rangle \neq 0;\ \hat{e}_j|i\rangle \notin S}} c_j \langle 0|H|j\rangle$$

with the only difference that now summation over doubles j runs over all of them independently on the origin of the c_j coefficient.

Thus we have a unique and complete set of coefficients of all Doubles, coming from either small 2×2 or large CI diagonalizations that can be used in the mutual dressing of all matrices (small or large ones).

The above procedure is size-consistent as long as the selection for the supersystem $A \cdots B$ does not introduce composite excitations resulting from the simultaneous excitations on A and B. The demonstration follows the same logic as that given in ref. [86] (with the later *erratum* in the ref. [85]) except for the fact that the small coefficients result from 2×2 diagonalizations and is subject to the same general assumptions that the MOs are localized on either A or B and the intersystem double excitations have zero amplitudes.

Table 7.1: Results for ammonia in DZP basis set.

Var. space			*Small Doubles*			*Energies*		
dim	n^{LD}	E^{var}	n^{SD}	ϵ^{mp2}	ϵ^{scpt}	$\tilde{E}$	$\tilde{E} + \epsilon^{mp2}$	$\widehat{E}$
397	392	−135.6	5426	−62.0	−73.5	−138.8	−200.8	−210.9
1440	1414	−180.0	4401	−17.3	−20.3	−185.0	−202.3	−204.8
2807	2633	−194.4	3182	−4.4	−5.1	−199.7	−204, 1	−204.7
3381	3061	−196.9	2765	−2.3	−2.7	−202.2	−204.5	−204.8
4957	3947	−200.1	1861	−0.4	−0.5	−205.1	−205.5	−205.6
13588	4892	−203.2	918	0.0	0.0	−207.3	−207.3	−207.3

Geometry and basis set are from Knowles and Handy [117], 1s orbital frozen. Energy differencies in milihartrees, with respect to the SCF energy. The full CI energy is -209.9 mH [117]. n^{LD} means the number of large doubles and n^{SD} the number of small ones. E^{var} refers to undressed CI, $\tilde{E}$ to (SC)^{2}CI energy (dressed only by large doubles) and $\widehat{E}$ refers to (SC)^2(CI+PT) energy. ϵ^{mp2} and ϵ^{scpt} denote the usual MP2 and SCPT contributions of small doubles, respectively [116].

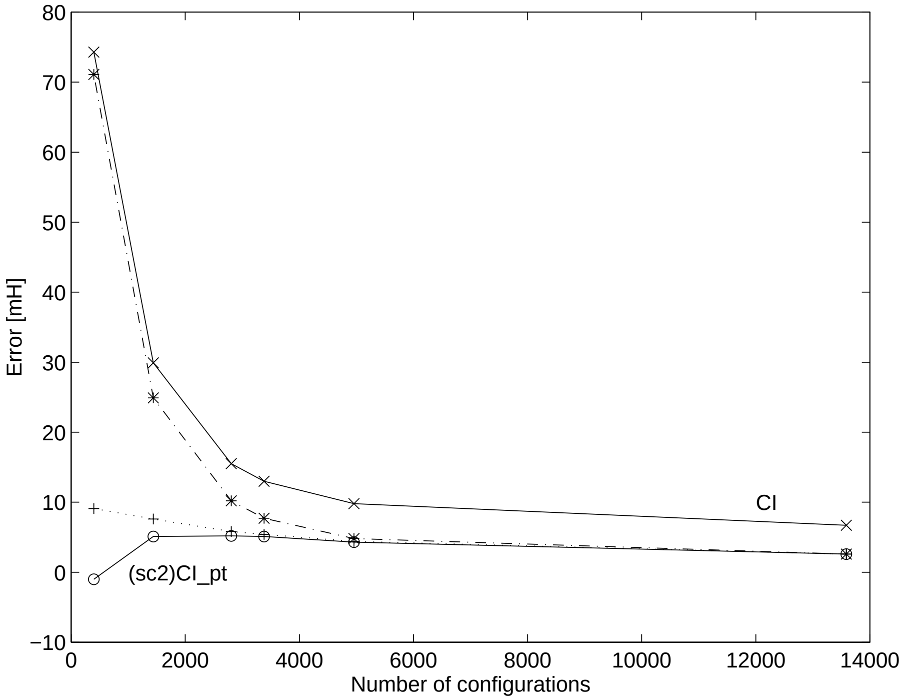
7.2 Test calculations

Summation over EPV terms uses $e(a, \ldots)$ quantities (4.59). The calculation of the diagonal dressing δ_{ii} becomes straigthforward then since it only requires summations over the holes and particles of i. In practice all determinants are dressed by all possible double excitations. The scheme of the algorithm is given in Appendix 7.3.

The efficiency of the above proposal has been tested through three sets of computations [116]. The first one concerns the NH_3 molecule in DZP basis set for which an estimate of the full CI energy has been proposed a few years ago [117]. The results appear in Table 7.1 and Fig. 7.1 for different variational spaces selected according to the CIPSI scheme. These spaces contain from 5% to 85% of all doubles, plus the most important triples and quadruples. The error never exceeds 5 mH.

Comparing the results from ref. [86] and from Table 7.1 one can notice that including 2633 doubles and 174 most important triples and quadruples in the S space and treating the other 3182 doubles in a perturbative mode one gets a better energy (-0.205 a.u.) than including all doubles in the SDCI (-0.197 a.u.) or in the $(SC)^2$SDCI (-0.203 a.u.) method. This example proves the importance of including the most significant higher excitations in the variational process while treating small doubles at a low level. As one can see from Table 1 small doubles still give 5.1 mH. In cases with much larger number of double excitations, where only a fraction of them could be treated variationally and where any MRMP2 calculations would be very difficult (if possible), the advantage of the combination of $(SC)^2$CI and $(SC)^2$PT2 may be even more pronounced.

Figure 7.1: NH_3 correlation energies (relative to FCI) for different dimensions of the variational space



'x' – undressed CI; '*' – $(SC)^2$CI; '+' – $(SC)^2$CI $+\epsilon^{MP2}$; 'o' – $(SC)^2$(CI+PT)

In order to test the behavior of the (SC)^{2}CI+PT method for bond breaking process, the H_2O molecule in the DZ basis set has been examinated. The results are given in Table 7.2.

The CIPSI MRMP2 and MREN2 methods give similar or slightly smaller errors (in the absolute numbers) comparing to the present results. But the perturbative CIPSI step is much more expensive and the present proposal is free from the uncertainty coming from the choice of $\hat{H}_0$.

Table 7.2: Results for water in DZ basis set.

	Var. space			*Small Doubles*		*Energies*		
	dim	n^{LD}	E^{var}	n^{SD}	ϵ^{scpt}	C^{MP}	C^{EN}	$\widehat{E}$
$1r_e$	1	0	0	841	−177.0	−130.5	−183.3	−177.0
	8	7	−32.9	834	−139.7	−134.5	−161.5	−172.0
	32	31	−69.0	810	−90.9	−136.1	−150.9	−159.9
	82	81	−97.1	760	−50.5	−140.2	−147.3	−148.6
	767	364	−142.5	477	−1.6	−147.3	−148.0	−146.2
	6478	709	−147.4	132	−0.01	− − −	− − −	−147.9
	CISD	=	−140.2			*CISDTQ*	=	−147.8
$1.5r_e$	1	0	0	841	−256.2	−191.1	−288.8	−256.2
	33	30	−112.5	811	−115.9	−199.5	−218.0	−226.5
	54	50	−136.6	791	−81.0	−201.5	−213.3	−218.0
	117	107	−160.8	734	−37.4	−204.1	−211.2	−202.6
	1329	366	−189.1	613	−6.6	−209.9	−210.6	−209.0
	6965	683	−209.2	158	−0.02	− − −	− − −	−210.1
	CISD	=	−188.6			*CISDTQ*	=	−209.9
$2r_e$	1	0	0	841	−348.1	−257.3	−519.8	−348.1
	38	32	−212.1	809	−102.9	−303.3	−314.4	−307.8
	73	61	−234.0	780	−52.9	−302.3	−312.1	−294.4
	123	89	−253.0	752	−34.8	−303.3	−309.4	−297.5
	1483	360	−298.7	481	−1.2	−308.4	−309.4	−305.3
	6551	649	−304.6	192	−0.02	− − −	− − −	−307.2
	CISD	=	−249.6			*CISDTQ*	=	−305.7

Geometry and basis set from Saxe *et. al.* [108]. Energy differencies in milihartrees, with respect to the SCF energy. The full CI energies are -148.0, -211.0, -310.1 mH for $1r_e$, 1.5_e and $2r_e$ respectively [108]. n^{LD} means the number of large doubles, whereas n^{SD} the number of small ones. E^{var} refers to undressed CI, $\widehat{E}$ to (SC)^2(CI+PT) energy. ϵ^{scpt} denotes SCPT contributions of small doubles, whereas C^{MP} and C^{EN} denote Møller–Plesset and Epstein–Nesbet CIPSI energies. Single reference CISD and CISDTQ energies are also given for comparison.

The separability property has been verified as a practical test for the program by calculating two water molecules at very large distances, using localized MOs. When the selection introduces only doubles in the variational space, for the supersystem AB as well as for A and B, the variational space is the simple union of those coresponding to the subsystems and the additivity follows from the proof given in ref. [86] and [112]. In the case where using the same thresholds for selections the variational space contains some intermolecular quadruple products of intramolecular doubles the results are no longer separable (contrary to what is stated in ref. [116]).

7.3 Discussion

It is worth comparing the present procedure to some popular multireference MP2 methods, although the perturbative part is based on a single reference. Standard MRMP2 methods proceed first to the diagonalization of the $P_S H P_S$ matrix and perturb later the multiconfigurational vector, while here we take into account an outer space effect before diagonalization. This point is crucial to insure size-extensivity [86].

Obviously some effects treated in MRMP2 methods are neglected in the present scheme, mainly the MP3 contribution incorporating the interactions between small and large doubles. The MRMP2 methods calculate all matrix elements between the model space determinants i and the outer space determinants α: $\langle i|H|\alpha\rangle$ while here one uses explicitely the matrix elements of the first row $\langle 0|H|\alpha\rangle$. Therefore the computational cost of the new procedure is simply that of the diagonalization plus that of an MP2, while in MRMP2 techniques the bottleneck is the perturbative step with the cost that increases rapidly with the number of reference determinants.

Let us remark finally that although the test calculations have not been performed on real scale problems since they concerned medium basis sets and small numbers of electrons, they illustrate the efficiency and possibilities of the method. The reciprocal dressing of the variational and perturbative type treatments improves significantly the results when the contribution of the small doubles remains important. This should be the case for the relevant domains of application of the method.

The above procedure may be generalized to open shell and excited states multireference situations. As proposed by Malrieu *et. al.* [115] one may generalize Independent Excitation Approximation (IEA) to MRIEA method. This generalization concerns CAS reference spaces and defines relatively small variational problems including the reference functions and excitations involving a given fixed set of inactive orbitals. Then all those excitations obtain a unique diagonal shift taking into account the outer space effects. The procedure goes through a set of such reciprocally dressed diagonalizations. Nevertheless the MRIEA method has not been implemented as yet and its usefulness must be evaluated.

Perspectives and conclusions

Remark: towards large scale problems

The tremendous development of computer technology and parallel progress in methodology of Quantum Chemistry have made possible accurate investigations of small to medium size molecular structures using *ab initio* methods. Unfortunately the computational costs of such methods increase with the size of the molecular system much faster than linearly - in case of methods accounting for the correlation energy usually faster than n^4, n being the number of electrons in the system. Simple and efficient procedures that enable to predict with reasonable accuracy various electronic phenomena are highly desirable. They should allow to attack larger systems using large basis sets. The efforts reported in the previous chapters were motivated by such reasoning. The new algorithms presented here result from looking for a compromise between contradictory requirements i.e. desired precision and computational cost.

Development of general formalisms of the MRCC type is a point of reference for further development of hopefully more practical proposals of the MRCEPA type. The same concerns the Sup-CI method. Although in principle one could include higher order effects, the multireference third order version implemented here seems to be an optimal choice. Nevertheless, as the computational cost increases so rapidly with the size of the considered systems, even such relatively simple methods can be only applied to moderate systems and further approximations may be necessary.

Regarding the Superdirect Configuration Interaction Method some simplified correction vectors could be used to built the effective short CI expansion as suggested in Chap. 2. The flexibility of various dressing procedures opens even greater possibilities for some approximated schemes due to the fact that one may dress any kind of CI matrices and combine it consistently with a dressing of other CI problems.

Summary

In this work several computational schemes for the electron correlation problem were formally derived and implemented. The following paragraphs contain summary of the thesis and some final remarks concerning the new methods of the Configuration Interaction type, which have been developed by the author of this thesis in collaboration with prof. W. Duch and prof. J.P. Malrieu.

The Superdirect Configuration Interaction method

The Superdirect Configuration Interaction (Sup-CI) method proposed by Duch [39] was presented. The method uses compact CI expansion in terms of contracted functions of the perturbative type. A version in which third order contributions are computed for a relatively small (10–100) space of reference and first order correction vectors has been implemented using modification of SGGA-CI program for matrix elements computations [57]. Several numerical tests on benchmark problems are included in Sec. 2.4. The MBPT like formulas for the matrix elements have been also derived, using the Symmetric Group Approach (SGA) to CI and an algebra of circular operators, invented for this purpose by Duch [56]. Various formal and graphical developments have been introduced by the author of this thesis in order to simplify the evaluation of these matrix elements according to general algorithm proposed by Duch. The derivation of general open shell third order Sup-CI matrix elements is the subject of Chap. 3.

As long as the first order correction vectors are used the Sup-CI method is an approximation of the MRCISD method. The Sup-CI method may be useful for computations of molecular properties and potential energy surfaces, since the relative errors for different geometries are much smaller than the absolute errors. CIPSI procedure for selection of the reference space applied at the preliminary stage of potential energy calculations will identify all relevant configurations and allow to fix one reference space for all geometry points used, increasing reliability of calculations. The results reported in Chap. 2 are superior to all single reference methods, including CI and many-body perturbation theory. On the other hand the method properly programmed should be an order of magnitude more efficient than multireference CI giving results of similar quality.

One should stress that from a formal point of view increasing the number of first-order correction vectors obtained from different reference CSFs leads to convergence to the exact solution of Schrödinger equation. Another, faster but computationally more demanding, way of improving the results would involve going beyond the first order corrections by computing approximations to the higher-order Hamiltonian expectation values, either by dividing integrals into different classes or by statistical sampling techniques. The most direct approach is to use approximate second-order correction vectors, for example by allowing only pair–pair interactions in the Hamiltonian used to create second-order correction vectors. Since the method is variational even approximate correction vectors must improve the energy. However, in calculation for water it has been verified that about one half of the improvement for second-order correction vectors comes from interactions involving integrals with all four indices different, which are not so easy to include.

The biggest problem that remains in Sup-CI as well as MR-CI is the size-extensivity error due to neglecting of the higher-order excitations. Although *a posteriori* corrections applied to Sup-CI energies remove a part of this error it is desirable to find more accurate corrections for this method. Another problem is the complexity of the resulting formulas presented in Chap. 3 and in the Appendix. Efficient programming of such formulas is a challenge comparable to programming of MBPT(5) method or MRPT methods.

Dressing of CI matrices as a bridge between CI and CC methods

Chapter 4 plays basically an auxiliary role. It introduces the concept of dressing of CI matrices based on the theory of the intermediate Hamiltonians [87] [86]. Roughly speaking dressing means adding non-linear terms, such as the non-linear terms appearing in CC methods, to the corresponding CI matrix. This may be viewed as size-consistent modification of the CI method in case of simplified dressings of the type of CEPA method or as particular method of solving the non-linear equations of the CC type through a dressed, pseudolinear set of CI equations.

Methodological background of this new approach was presented. Some previous applications are traced and relationships between different methods that may be regarded as particular forms of a dressing are explained by the author of this thesis using the existing literature. Moreover a new proposal, namely Generalized Coupled Pair Functional (GCPF) was proposed. The GCPF method is a special form of dressing of the CEPA type of the overlap matrix. It has been suggested and developed by Malrieu and the author of this thesis.

Mutireference state-specific size-consistent procedures

A partition of the coefficients of the Singles and Doubles from a multireference space, as obtained from a MRCISD wave function has been proposed. This partition makes possible the definition of state-specific reference-dependent amplitudes for the single and double excitations, compatible with a partition of the state-specific operator $\hat{\Omega}^m$ into a sum of reference-specific state-specific operators

$$\hat{\Omega}^m = \sum_{I \in S} \hat{\Omega}^m_I \tag{7.3}$$

Then it is possible to assume an exponential form for each of the $\hat{\Omega}^m_I$'s

$$\hat{\Omega}^m_I = e^{\hat{T}^m_I} \hat{P}_I \tag{7.4}$$

and one obtains directly a coupled cluster expansion of the multireference wave function. The so-obtained leading contributions to the coefficients of the Triples and Quadruples (i.e. of $\hat{T}^2$ type) are used to dress the CISD matrix according to the state-specific intermediate Hamiltonians formalism, in the iterative, self-consistent manner. This proposal may be viewed as a dressed CISD formulation of the CCSD method for the multireference case, as previously proposed for the single reference [90]. One should also note that this procedure decouples the Jeziorski-Monkhorst multistate ansatz [83] into a state-specific decontracted expansion. The new method is referred to as State Specific Coupled Cluster type of dressing of the MRCI matrix (SS-CCdCI) [105]. The results of a preliminary application of this method to the H_4 problem, which happens to be difficult for VU and SU MRCC approaches, are very encouraging. They are presented in Chap. 5. Nevertheless, it is certainly desirable to refine the definition of the reference dependent state-specific amplitudes within the general framework proposed here.

The state-specific CC type of dressing was inspired by the previous formulation of MR(SC)^{2}CI method, proposed by Malrieu *et. al.* [85]. The "weighted genealogy" of this method is logically connected to the partition of the coefficients of single and double excitations with respect to their parent references exploited here. Convergence problems that arise when using such a weighting may be overcome. The numerical results of the MR(SC)^{2}CI method, implemented by the author of this thesis, are very promising. They are presented in Chap. 5.

Mutireference state-universal size-consistent procedures

The state-specific coupled cluster type of dressing is based on the partition of the state-specific CI coefficients. All definitions of such a procedure must however contain some arbitrariness. To avoid this drawback a generalization of the SS-CCdCI procedure to the Multistate or State-Universal Coupled Cluster type of dressing of the MRCI matrix has been proposed [111]. The MS-CCdCI procedure starts from a set of CI eigenstates to define a state independent dressing in terms of the state-universal CC amplitudes. More precisely, one assumes again the separate cluster ansätze in the spirit of the Jeziorski–Monkhorst Hilbert space MRCC method but with the truly multistate amplitudes and truncation of the cluster operators to the corresponding CI model space.

In order to define the CC amplitudes one solves a set of linear equations obtained by comparison of the linear CI expansions for a set of the exact states and the corresponding CC expansions in terms of products of the state-universal amplitudes and state-specific reference coefficients. One may then consider for any truncated CI expansion (actually for a set of CI expansions for a set of states) its exponentialization i.e. an adjusted CC approximation with the cluster operators truncated to excitations remaining in the CI space. The products of such excitations define through the CC expansion the approximate amplitudes of the outer space determinants (i.e. those that not belong to the CI space) amplitudes and the dressing terms that has to be added to the CI equations in order to obtain the CC equations. At convergence one obtains amplitudes of CC approximation adjusted to the CI space from the dressed CI coefficients.

In the MS-CCdCI one has to take into account all the possible decompositions of *all* model space determinants (not only singly and doubly excited ones) when considering expansions with respect to subsequent references in the Jeziorski–Monkhorst ansatz spirit. Thus, the amplitudes of triples or quadruples with respect to a given reference may appear in the dressing of the MR-CISD matrix. This certainly increases the complexity of the method. One may however hope that the approximate scheme of the MRCEPA type, which requires decompositions involving doubly excited functions only and is therefore relatively cheap, will be an interesting alternative to the existing multireference size-consistent procedures. The MS-CCdCI method and its approximated MS-CEPA version have not been implemented as yet. Both approximations have been formally developed by Malrieu and the author of this thesis in collaboration with Adamowicz and Caballol [111]. The proof of size-consistency of MS-CCdCI procedure and other proofs of size-consistency included in this thesis come also from the author of this thesis.

The $(SC)^2$CI+PT method

Size-consistent self-consistent combination of selected CI and PT [116] is an effective combination of the dressing of CEPA type of selected CI matrix and a number of two by two dressed diagonalizations providing coefficients of the not selected doubly excited functions. This procedure is more flexible and more powerful than an approximate SDCI going through a partition into large doubles and small doubles. The advantage come from the size-consistency, but even more important is the incorporation of higher order effects, including the most important triples and quadruples. This should be especially important for many problems in quantum chemistry where a subset of highly correlated electrons appears, while the electronic correlation of the remaining electrons may be treated at a lower level of accuracy. One may think of chemical reactions where only a few bonds are strongly affected, or of the π electrons "above" the σ electrons in conjugated molecules.

Another field of application where size-extensivity is strongly desired is the interaction between molecules. At present the methods used for computation of molecular interactions are either the single reference MBPTn expansions (in practice MP2 in most cases) or single reference CC methods. The former ones are of poor reliability (especially if some internal doubles have large coefficients) whereas the second ones are very expensive, except for CCSD which may not be sufficient.

In comparison with MRPT2 methods, the present scheme may be seen as more approximate since it neglects the interaction between small and large doubles, but it is formally more rigorous and certainly much cheaper for a given size of the variational space since we do not generate all the determinants interacting with the selected configurations. The bottleneck is no longer the perturbative step, allowing for larger dimensions of the CI space (especially if one uses Direct Selected algorithms). The $(SC)^2CI+PT$ method has been implemented by the author of this thesis. Several numerical tests on benchmark problems are presented in Chap. 7.

Appendix

A: List of papers the Ph.D. thesis is based on

- W. Duch and J. Meller; *On multireference superdirect configuration interaction in third order*, International Journal of Quantum Chemistry, **50**, 243 (1994)
- J. Meller, J. L. Heully and J. P. Malrieu; *Size-consistent self consistent combination of selected CI and perturbation theory*, Chemical Physics Letters, **218**, 276 (1994)
- J. Meller, J.P. Malrieu and J.L. Heully; *Multireference size-consistent self consistent CI method: a few applications to ground and excited states*, Chemical Physics Letters, **244**, 440 (1995)
- J. Meller, J.P. Malrieu and R. Caballol; *State-specific Coupled Cluster type dressing of Multireference Singles and Doubles Configuration Interaction matrix*, Journal of Chemical Physics, in press (1996)
- L. Adamowicz, R. Caballol, J.P. Malrieu and J. Meller; *A general bridge between CI and CC methods: a multistate solution*, Chemical Physics Letters, submitted for publication

In preparation:

- J. Meller and J.P. Malrieu, *A Generalized CPF method through a modified dressing of the CI matrix*
- J. Meller and W. Duch, *General open shell third order MBPT formulae*

B: Formulae for the Sup-CI matrix elements

Table 8.1: H^3 type diagonal matrix element: contractions
For explanations see Chap. 3

$$\begin{aligned}
< 0|R_1 H R_2|0 > := {} & \mathrm{d1}(m,p)\,\mathrm{d2}(m,p) < |\,\mathrm{b1}(m,p)\,\mathrm{b0}(m,p)\,\mathrm{b1}(p,m)\,| > \\
& + \mathrm{d1}(m,p)(\mathrm{d2}(e,p) < |\,\mathrm{b1}(m,p)\,\mathrm{b1}(e,m)\,\mathrm{b1}(p,e)\,| > \\
& + \mathrm{d2}(m,g) < |\,\mathrm{b1}(m,p)\,\mathrm{b1}(p,g)\,\mathrm{b1}(g,m)\,| >) \\
& + \mathrm{d1}(m,p)\,\mathrm{d2}(m,f,p,h) < |\,\mathrm{b1}(m,p)\,\mathrm{b1}(f,h)\,\mathrm{b2}(p,h,m,f)\,| > \\
& + \mathrm{d1}(m,p)\,\mathrm{d2}(e,g) < |\,\mathrm{b1}(m,p)\,\mathrm{b2}(p,e,m,g)\,\mathrm{b1}(g,e)\,| > + \\
& \mathrm{d1}(m,p)(\mathrm{d2}(e,f,p,h) < |\,\mathrm{b1}(m,p)\,\mathrm{b2}(e,f,m,h)\,\mathrm{b2}(p,h,e,f)\,| > \\
& + \mathrm{d2}(m,f,g,h) < |\,\mathrm{b1}(m,p)\,\mathrm{b2}(p,f,g,h)\,\mathrm{b2}(g,h,m,f)\,| >) + \\
& \mathrm{d1}(m,n,p,q)\,\mathrm{d2}(m,n,p,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b0}(m,n,p,q) \\
& \mathrm{b2}(p,q,m,n)\,| > + \mathrm{d1}(m,n,p,q)(\\
& \mathrm{d2}(n,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(p,m)\,\mathrm{b1}(q,n)\,| > \\
& + \mathrm{d2}(n,p) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(q,m)\,\mathrm{b1}(p,n)\,| > \\
& + \mathrm{d2}(m,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(q,m)\,\mathrm{b1}(p,n)\,| > \\
& + \mathrm{d2}(m,p) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(p,m)\,\mathrm{b1}(q,n)\,| >) + \\
& \mathrm{d1}(m,n,p,q)(\\
& \mathrm{d2}(e,n,p,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(e,m)\,\mathrm{b2}(p,q,e,n)\,| > \\
& + \mathrm{d2}(m,e,p,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(e,n)\,\mathrm{b2}(p,q,m,e)\,| > \\
& + \mathrm{d2}(m,n,g,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(p,g)\,\mathrm{b2}(g,q,m,n)\,| > \\
& + \mathrm{d2}(m,n,p,g) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(q,g)\,\mathrm{b2}(p,g,m,n)\,| >) + \\
& \mathrm{d1}(m,n,p,q)(\\
& \mathrm{d2}(m,g) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(p,q,n,g)\,\mathrm{b1}(g,m)\,| > \\
& + \mathrm{d2}(n,g) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(p,q,m,g)\,\mathrm{b1}(g,n)\,| > \\
& + \mathrm{d2}(e,p) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(q,e,m,n)\,\mathrm{b1}(p,e)\,| > \\
& + \mathrm{d2}(e,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(p,e,m,n)\,\mathrm{b1}(q,e)\,| >) + \\
& \mathrm{d1}(m,n,p,q)(\\
& \mathrm{d2}(e,f,p,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(e,f,m,n)\,\mathrm{b2}(p,q,e,f)\,| > \\
& + \mathrm{d2}(m,n,g,h) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(p,q,g,h)\,\mathrm{b2}(g,h,m,n)\,| > \\
& + \mathrm{d2}(e,n,g,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(p,e,m,g)\,\mathrm{b2}(g,q,e,n)\,| > \\
& + \mathrm{d2}(e,n,p,g) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(q,e,m,g)\,\mathrm{b2}(p,g,e,n)\,| > \\
& + \mathrm{d2}(m,e,g,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(p,e,n,g)\,\mathrm{b2}(g,q,m,e)\,| > \\
& + \mathrm{d2}(m,e,p,g) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(q,e,n,g)\,\mathrm{b2}(p,g,m,e)\,| > \\
&)
\end{aligned}$$

Table 8.2: H^3 type diagonal mat. el. - distinct chains

CSC expanded with respect to indices repetitions. FTOS contributions included, except for the S1 FTOS contributions given in Table 3.4 and S1 terms from (3.74). For explanations see Chap. 3.

$$
\begin{aligned}
&\mathrm{d1}(m,n,s,q)\,\mathrm{d2}(s,n,g,q) < |\,\mathrm{b2}(m,n,s,q)\,\mathrm{b2}(s,s,m,g)\\
&\quad \mathrm{b2}(g,q,s,n)\,| > + \mathrm{d1}(m,m,s,q)\,\mathrm{d2}(s,m,g,q) < |\,\mathrm{b2}(m,m,s,q)\\
&\quad \mathrm{b2}(s,s,m,g)\,\mathrm{b2}(g,q,s,m)\,| > + \mathrm{d1}(m,m,p,q)(\\
&\quad \mathrm{d2}(e,f,p,q) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{b2}(e,f,m,m)\,\mathrm{b2}(p,q,e,f)\,| >\\
&\quad + \mathrm{d2}(e,e,p,q) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{b2}(e,e,m,m)\,\mathrm{b2}(p,q,e,e)\,| >\\
&\quad + \mathrm{d2}(m,m,g,h) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{b2}(p,q,g,h)\,\mathrm{b2}(g,h,m,m)\,| >\\
&\quad + \mathrm{d2}(m,m,g,g) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{b2}(p,q,g,g)\,\mathrm{b2}(g,g,m,m)\,| >\\
&\quad + \mathrm{d2}(e,m,g,q) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{b2}(p,e,m,g)\,\mathrm{b2}(g,q,e,m)\,| >)\\
&\quad + \mathrm{d1}(m,m,p,p)(\\
&\quad \mathrm{d2}(e,m,p,p) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{b1}(e,m)\,\mathrm{b2}(p,p,e,m)\,| >\\
&\quad + \mathrm{d2}(m,m,g,p) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{b1}(p,g)\,\mathrm{b2}(g,p,m,m)\,| >)\\
&\quad + \mathrm{d1}(m,n,p,q)(\\
&\quad \mathrm{d2}(m,g) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(p,q,n,g)\,\mathrm{b1}(g,m)\,| >\\
&\quad + \mathrm{d2}(e,p) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(q,e,m,n)\,\mathrm{b1}(p,e)\,| >) +\\
&\quad \mathrm{d1}(m,m,p,q)(\\
&\quad \mathrm{d2}(m,g) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{b2}(p,q,m,g)\,\mathrm{b1}(g,m)\,| >\\
&\quad + \mathrm{d2}(e,p) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{b2}(q,e,m,m)\,\mathrm{b1}(p,e)\,| >) +\\
&\quad \mathrm{d1}(m,n,p,p)(\\
&\quad \mathrm{d2}(m,g) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{b2}(p,p,n,g)\,\mathrm{b1}(g,m)\,| >\\
&\quad + \mathrm{d2}(e,p) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{b2}(p,e,m,n)\,\mathrm{b1}(p,e)\,| >) +\\
&\quad \mathrm{d1}(m,m,p,p)(\\
&\quad \mathrm{d2}(m,g) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{b2}(p,p,m,g)\,\mathrm{b1}(g,m)\,| >\\
&\quad + \mathrm{d2}(e,p) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{b2}(p,e,m,m)\,\mathrm{b1}(p,e)\,| >) +\\
&\quad \mathrm{d1}(m,n,p,q)(\\
&\quad \mathrm{d2}(e,f,p,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(e,f,m,n)\,\mathrm{b2}(p,q,e,f)\,| >\\
&\quad + \mathrm{d2}(e,e,p,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(e,e,m,n)\,\mathrm{b2}(p,q,e,e)\,| >\\
&\quad + \mathrm{d2}(m,n,g,h) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(p,q,g,h)\,\mathrm{b2}(g,h,m,n)\,| >\\
&\quad + \mathrm{d2}(m,n,g,g) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(p,q,g,g)\,\mathrm{b2}(g,g,m,n)\,| >\\
&\quad + \mathrm{d2}(e,n,g,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b2}(p,e,m,g)\,\mathrm{b2}(g,q,e,n)\,| >)\\
&\quad + \mathrm{d1}(m,t,p,q)\,\mathrm{d2}(m,t) < |\,\mathrm{b2}(m,t,p,q)\,\mathrm{b2}(p,q,t,t)\,\mathrm{b1}(t,m)\,| >\\
&\quad + \mathrm{d1}(m,t,p,p)\,\mathrm{d2}(m,t) < |\,\mathrm{b2}(m,t,p,p)\,\mathrm{b2}(p,p,t,t)\,\mathrm{b1}(t,m)\,| >
\end{aligned}
$$

$+\,\mathrm{d1}(m,n,p,s)\,\mathrm{d2}(s,p) < |\,\mathrm{b2}(m,n,p,s)\,\mathrm{b2}(s,s,m,n)\,\mathrm{b1}(p,s)\,| >$
$+\,\mathrm{d1}(m,m,p,s)\,\mathrm{d2}(s,p) < |\,\mathrm{b2}(m,m,p,s)\,\mathrm{b2}(s,s,m,m)$
$\mathrm{b1}(p,s)\,| > +\,\mathrm{d1}(t,n,p,q)\,\mathrm{d2}(e,n,t,q) < |\,\mathrm{b2}(t,n,p,q)$
$\mathrm{b2}(p,e,t,t)\,\mathrm{b2}(t,q,e,n)\,| > +\,\mathrm{d1}(t,n,p,p)\,\mathrm{d2}(e,n,t,p) < |$
$\mathrm{b2}(t,n,p,p)\,\mathrm{b2}(p,e,t,t)\,\mathrm{b2}(t,p,e,n)\,| > +\,\mathrm{d1}(t,n,s,q)$
$\mathrm{d2}(s,n,t,q) < |\,\mathrm{b2}(t,n,s,q)\,\mathrm{b2}(s,s,t,t)\,\mathrm{b2}(t,q,s,n)\,| > +$
$\mathrm{d1}(m,n,p,p)($
$\mathrm{d2}(e,f,p,p) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{b2}(e,f,m,n)\,\mathrm{b2}(p,p,e,f)\,| >$
$+\,\mathrm{d2}(e,e,p,p) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{b2}(e,e,m,n)\,\mathrm{b2}(p,p,e,e)\,| >$
$+\,\mathrm{d2}(m,n,g,h) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{b2}(p,p,g,h)\,\mathrm{b2}(g,h,m,n)\,| >$
$+\,\mathrm{d2}(m,n,g,g) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{b2}(p,p,g,g)\,\mathrm{b2}(g,g,m,n)\,| >$
$+\,\mathrm{d2}(e,n,g,p) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{b2}(p,e,m,g)\,\mathrm{b2}(g,p,e,n)\,| >)$
$+\,\mathrm{d1}(m,m,p,p)($
$\mathrm{d2}(e,f,p,p) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{b2}(e,f,m,m)\,\mathrm{b2}(p,p,e,f)\,| >$
$+\,\mathrm{d2}(e,e,p,p) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{b2}(e,e,m,m)\,\mathrm{b2}(p,p,e,e)\,| >$
$+\,\mathrm{d2}(m,m,g,h) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{b2}(p,p,g,h)\,\mathrm{b2}(g,h,m,m)\,| >$
$+\,\mathrm{d2}(m,m,g,g) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{b2}(p,p,g,g)\,\mathrm{b2}(g,g,m,m)\,| >$
$+\,\mathrm{d2}(e,m,g,p) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{b2}(p,e,m,g)\,\mathrm{b2}(g,p,e,m)\,| >)$
$+\,\mathrm{d1}(m,n,p,q)\,\mathrm{d2}(n,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(p,m)\,\mathrm{b1}(q,n)\,| >$
$+\,\mathrm{d1}(m,s)\,\mathrm{d2}(m,s,g,g) < |\,\mathrm{b1}(m,s)\,\mathrm{b2}(s,s,g,g)\,\mathrm{b2}(g,g,m,s)$
$| > +\,\mathrm{d1}(m,p)(\mathrm{d2}(e,p) < |\,\mathrm{b1}(m,p)\,\mathrm{b1}(e,m)\,\mathrm{b1}(p,e)\,| >$
$+\,\mathrm{d2}(m,g) < |\,\mathrm{b1}(m,p)\,\mathrm{b1}(p,g)\,\mathrm{b1}(g,m)\,| >) +\,\mathrm{d1}(m,s)$
$\mathrm{d2}(m,s,g,h) < |\,\mathrm{b1}(m,s)\,\mathrm{b2}(s,s,g,h)\,\mathrm{b2}(g,h,m,s)\,| > +$
$\mathrm{d1}(m,n,p,p)($
$\mathrm{d2}(e,n,p,p) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{b1}(e,m)\,\mathrm{b2}(p,p,e,n)\,| >$
$+\,\mathrm{d2}(m,n,g,p) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{b1}(p,g)\,\mathrm{b2}(g,p,m,n)\,| >)$
$+\,\mathrm{d1}(t,p)\,\mathrm{d2}(e,e,p,t) < |\,\mathrm{b1}(t,p)\,\mathrm{b2}(e,e,t,t)\,\mathrm{b2}(p,t,e,e)\,| >$
$+\,\mathrm{d1}(m,m,p,p)\,\mathrm{d2}(m,p) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{b1}(p,m)^2\,| >$
$+\,\mathrm{d1}(m,n,p,p)\,\mathrm{d2}(n,p) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{b1}(p,m)\,\mathrm{b1}(p,n)\,| >$
$+\,\mathrm{d1}(m,m,p,q)\,\mathrm{d2}(m,q) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{b1}(p,m)\,\mathrm{b1}(q,m)\,| >$
$+\,\mathrm{d1}(m,m,p,q)($
$\mathrm{d2}(e,m,p,q) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{b1}(e,m)\,\mathrm{b2}(p,q,e,m)\,| >$
$+\,\mathrm{d2}(m,m,g,q) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{b1}(p,g)\,\mathrm{b2}(g,q,m,m)\,| >)$
$+\,\mathrm{d1}(m,p)\,\mathrm{d2}(m,p) < |\,\mathrm{b1}(m,p)\,\mathrm{b0}(m,p)\,\mathrm{b1}(p,m)\,| > +$
$\mathrm{d1}(m,n,p,q)($
$\mathrm{d2}(e,n,p,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(e,m)\,\mathrm{b2}(p,q,e,n)\,| >$
$+\,\mathrm{d2}(m,n,g,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b1}(p,g)\,\mathrm{b2}(g,q,m,n)\,| >) +$
$\mathrm{d1}(m,n,p,q)\,\mathrm{d2}(m,n,p,q) < |\,\mathrm{b2}(m,n,p,q)\,\mathrm{b0}(m,n,p,q)$
$\mathrm{b2}(p,q,m,n)\,| >$

$$
\begin{aligned}
&+\mathrm{d1}(m,p)\,\mathrm{d2}(m,f,p,h) < |\,\mathrm{b1}(m,p)\,\mathrm{b1}(f,h)\,\mathrm{b2}(p,h,m,f)\,| > \\
&+\mathrm{d1}(m,p)\,\mathrm{d2}(e,g) < |\,\mathrm{b1}(m,p)\,\mathrm{b2}(p,e,m,g)\,\mathrm{b1}(g,e)\,| > + \\
&\mathrm{d1}(m,p)(\mathrm{d2}(e,f,p,h) < |\,\mathrm{b1}(m,p)\,\mathrm{b2}(e,f,m,h)\,\mathrm{b2}(p,h,e,f)\,| > \\
&+\mathrm{d2}(e,e,p,h) < |\,\mathrm{b1}(m,p)\,\mathrm{b2}(e,e,m,h)\,\mathrm{b2}(p,h,e,e)\,| > \\
&+\mathrm{d2}(m,f,g,h) < |\,\mathrm{b1}(m,p)\,\mathrm{b2}(p,f,g,h)\,\mathrm{b2}(g,h,m,f)\,| > \\
&+\mathrm{d2}(m,f,g,g) < |\,\mathrm{b1}(m,p)\,\mathrm{b2}(p,f,g,g)\,\mathrm{b2}(g,g,m,f)\,| >) \\
&+\mathrm{d1}(m,s)\,\mathrm{d2}(s,g) < |\,\mathrm{b1}(m,s)\,\mathrm{b2}(s,s,m,g)\,\mathrm{b1}(g,s)\,| > \\
&+\mathrm{d1}(t,p)\,\mathrm{d2}(e,t) < |\,\mathrm{b1}(t,p)\,\mathrm{b2}(p,e,t,t)\,\mathrm{b1}(t,e)\,| > \\
&+\mathrm{d1}(t,s)\,\mathrm{d2}(s,t) < |\,\mathrm{b1}(t,s)^2\,\mathrm{b2}(s,s,t,t)\,| > \\
&+\mathrm{d1}(t,p)\,\mathrm{d2}(e,f,p,t) < |\,\mathrm{b1}(t,p)\,\mathrm{b2}(e,f,t,t)\,\mathrm{b2}(p,t,e,f)\,| >
\end{aligned}
$$

Table 8.3: H^3 type diag. mat. el. - 2nd and 3d H expanded

CSC expanded with respect to indices repetitions. FTOS contributions included, except for the S1 FTOS contributions given in Table 3.4 and S1 terms from (3.74). $c1(m,p) = \sum_k (mp|kk)(\hat{n}_k - \delta_{mk}) + (m|p)$ and $c2(m,p,k) = \sum_{k \neq m,p} (mk|pk)$ are factors arising from $\hat{B}_1$. For explanations see Chap. 3.

$$\begin{aligned}
&\mathrm{d1}(m,m,p,q)\Big(\mathrm{d2}(e,f,p,q) < |\,\mathrm{b2}(m,m,p,q)(\\
&\quad \mathrm{i1}(e,m,f,m)\,\mathrm{i1}(p,e,q,f)\,[e,m,f,m,p,e,q,f] \\
&\quad +\mathrm{i1}(e,m,f,m)\,\mathrm{i1}(p,f,q,e)\,[e,m,f,m,p,f,q,e])| > + \frac{1}{2} \\
&\quad \mathrm{d2}(e,e,p,q) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{i1}(e,m,e,m)\,\mathrm{i1}(p,e,q,e) \\
&\quad [e,m,e,m,p,e,q,e]\,| > + \mathrm{d2}(m,m,g,h) < |\,\mathrm{b2}(m,m,p,q)(\\
&\quad \mathrm{i1}(p,g,q,h)\,\mathrm{i1}(g,m,h,m)\,[p,g,q,h,g,m,h,m] \\
&\quad +\mathrm{i1}(p,h,q,g)\,\mathrm{i1}(g,m,h,m)\,[p,h,q,g,g,m,h,m])| > + \frac{1}{2} \\
&\quad \mathrm{d2}(m,m,g,g) < |\,\mathrm{b2}(m,m,p,q)\,\mathrm{i1}(p,g,q,g)\,\mathrm{i1}(g,m,g,m) \\
&\quad [p,g,q,g,g,m,g,m]\,| > + \mathrm{d2}(e,m,g,q) < |\,\mathrm{b2}(m,m,p,q)(\\
&\quad \mathrm{i1}(p,m,e,g)\,\mathrm{i1}(g,e,q,m)\,[p,m,e,g,g,e,q,m] \\
&\quad +\mathrm{i1}(p,m,e,g)\,\mathrm{i1}(g,m,q,e)\,[p,m,e,g,g,m,q,e] \\
&\quad +\mathrm{i1}(p,g,e,m)\,\mathrm{i1}(g,e,q,m)\,[p,g,e,m,g,e,q,m] \\
&\quad +\mathrm{i1}(p,g,e,m)\,\mathrm{i1}(g,m,q,e)\,[p,g,e,m,g,m,q,e])| >\Big) + \\
&\quad \mathrm{d1}(m,n,p,p)\Big(\mathrm{d2}(e,f,p,p) < |\,\mathrm{b2}(m,n,p,p)(\\
&\quad \mathrm{i1}(e,m,f,n)\,\mathrm{i1}(p,e,p,f)\,[e,m,f,n,p,e,p,f] \\
&\quad +\mathrm{i1}(e,n,f,m)\,\mathrm{i1}(p,e,p,f)\,[e,n,f,m,p,e,p,f])| > + \frac{1}{2} \\
&\quad \mathrm{d2}(e,e,p,p) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{i1}(e,m,e,n)\,\mathrm{i1}(p,e,p,e) \\
&\quad [e,m,e,n,p,e,p,e]\,| > + \mathrm{d2}(m,n,g,h) < |\,\mathrm{b2}(m,n,p,p)(\\
&\quad \mathrm{i1}(p,g,p,h)\,\mathrm{i1}(g,m,h,n)\,[p,g,p,h,g,m,h,n] \\
&\quad +\mathrm{i1}(p,g,p,h)\,\mathrm{i1}(g,n,h,m)\,[p,g,p,h,g,n,h,m])| > + \frac{1}{2} \\
&\quad \mathrm{d2}(m,n,g,g) < |\,\mathrm{b2}(m,n,p,p)\,\mathrm{i1}(p,g,p,g)\,\mathrm{i1}(g,m,g,n) \\
&\quad [p,g,p,g,g,m,g,n]\,| > + \mathrm{d2}(e,n,g,p) < |\,\mathrm{b2}(m,n,p,p)(\\
&\quad \mathrm{i1}(p,m,e,g)\,\mathrm{i1}(g,e,p,n)\,[p,m,e,g,g,e,p,n] \\
&\quad +\mathrm{i1}(p,m,e,g)\,\mathrm{i1}(g,n,p,e)\,[p,m,e,g,g,n,p,e] \\
&\quad +\mathrm{i1}(p,g,e,m)\,\mathrm{i1}(g,e,p,n)\,[p,g,e,m,g,e,p,n] \\
&\quad +\mathrm{i1}(p,g,e,m)\,\mathrm{i1}(g,n,p,e)\,[p,g,e,m,g,n,p,e])| >\Big) + \\
&\quad \mathrm{d1}(m,m,p,p)\Big(\mathrm{d2}(e,f,p,p) < |\,\mathrm{b2}(m,m,p,p)\,\mathrm{i1}(e,m,f,m) \\
&\quad \mathrm{i1}(p,e,p,f)\,[e,m,f,m,p,e,p,f]\,| > + \frac{1}{4}\mathrm{d2}(e,e,p,p) < | \\
&\quad \mathrm{b2}(m,m,p,p)\,\mathrm{i1}(e,m,e,m)\,\mathrm{i1}(p,e,p,e)\,[e,m,e,m,p,e,p,e]
\end{aligned}$$

$$
\begin{array}{l}
| > + \mathrm{d2}(m,m,g,h) < | \mathrm{b2}(m,m,p,p)\, \mathrm{i1}(p,g,p,h)\, \mathrm{i1}(g,m,h,m) \\
{[p,g,p,h,g,m,h,m]} \,| > + \frac{1}{4}\mathrm{d2}(m,m,g,g) < | \mathrm{b2}(m,m,p,p) \\
\mathrm{i1}(p,g,p,g)\, \mathrm{i1}(g,m,g,m)\, [p,g,p,g,g,m,g,m] \,| > + \\
\mathrm{d2}(e,m,g,p) < | \mathrm{b2}(m,m,p,p)(\\
\mathrm{i1}(p,m,e,g)\, \mathrm{i1}(g,e,p,m)\, [p,m,e,g,g,e,p,m] \\
+ \mathrm{i1}(p,m,e,g)\, \mathrm{i1}(g,m,p,e)\, [p,m,e,g,g,m,p,e] \\
+ \mathrm{i1}(p,g,e,m)\, \mathrm{i1}(g,e,p,m)\, [p,g,e,m,g,e,p,m] \\
+ \mathrm{i1}(p,g,e,m)\, \mathrm{i1}(g,m,p,e)\, [p,g,e,m,g,m,p,e])| >\Big) + \\
\mathrm{d1}(m,p)\, \mathrm{d2}(m,f,p,h) < | \mathrm{b1}(m,p)(\\
\mathrm{c1}(f,h)\, \mathrm{i1}(p,m,h,f)\, [f,h,p,m,h,f] \\
+ \mathrm{c1}(f,h)\, \mathrm{i1}(p,f,h,m)\, [f,h,p,f,h,m] \\
+ \mathrm{c2}(f,h,r2)\, \mathrm{i1}(p,m,h,f)\, [r2,h,f,r2,p,m,h,f] \\
+ \mathrm{c2}(f,h,r2)\, \mathrm{i1}(p,f,h,m)\, [r2,h,f,r2,p,f,h,m])| > \\
+ \mathrm{d1}(m,p)\, \mathrm{d2}(m,p) < | \mathrm{b1}(m,p)\, \mathrm{b0}(m,p)\, \mathrm{b1}(p,m) \,| > + \\
\mathrm{d1}(m,p)(\mathrm{d2}(e,p) < | \mathrm{b1}(m,p)(\mathrm{c1}(e,m)\, \mathrm{c1}(p,e)\, [e,m,p,e] \\
+ \mathrm{c1}(e,m)\, \mathrm{c2}(p,e,r3)\, [e,m,r3,e,p,r3] \\
+ \mathrm{c2}(e,m,r2)\, \mathrm{c1}(p,e)\, [r2,m,e,r2,p,e] \\
+ \mathrm{c2}(e,m,r2)\, \mathrm{c2}(p,e,r3)\, [r2,m,e,r2,r3,e,p,r3])| > + \\
\mathrm{d2}(m,g) < | \mathrm{b1}(m,p)(\mathrm{c1}(p,g)\, \mathrm{c1}(g,m)\, [p,g,g,m] \\
+ \mathrm{c1}(p,g)\, \mathrm{c2}(g,m,r3)\, [p,g,r3,m,g,r3] \\
+ \mathrm{c2}(p,g,r2)\, \mathrm{c1}(g,m)\, [r2,g,p,r2,g,m] \\
+ \mathrm{c2}(p,g,r2)\, \mathrm{c2}(g,m,r3)\, [r2,g,p,r2,r3,m,g,r3])| >) + \\
\mathrm{d1}(t,p)\, \mathrm{d2}(e,f,p,t) < | \mathrm{b1}(t,p)(\\
\mathrm{i1}(e,t,f,t)\, \mathrm{i1}(p,e,t,f)\, [e,t,f,t,p,e,t,f] \\
+ \mathrm{i1}(e,t,f,t)\, \mathrm{i1}(p,f,t,e)\, [e,t,f,t,p,f,t,e])| > + \frac{1}{2}\mathrm{d1}(t,p) \\
\mathrm{d2}(e,e,p,t) < | \mathrm{b1}(t,p)\, \mathrm{i1}(e,t,e,t)\, \mathrm{i1}(p,e,t,e) \\
{[e,t,e,t,p,e,t,e]} \,| > + \mathrm{d1}(m,s)\, \mathrm{d2}(m,s,g,h) < | \mathrm{b1}(m,s)(\\
\mathrm{i1}(s,g,s,h)\, \mathrm{i1}(g,m,h,s)\, [s,g,s,h,g,m,h,s] \\
+ \mathrm{i1}(s,g,s,h)\, \mathrm{i1}(g,s,h,m)\, [s,g,s,h,g,s,h,m])| > + \frac{1}{2} \\
\mathrm{d1}(m,s)\, \mathrm{d2}(m,s,g,g) < | \mathrm{b1}(m,s)\, \mathrm{i1}(s,g,s,g)\, \mathrm{i1}(g,m,g,s) \\
{[s,g,s,g,g,m,g,s]} \,| > + \mathrm{d1}(m,n,p,q)\, \mathrm{d2}(m,n,p,q) < | \\
\mathrm{b2}(m,n,p,q)\, \mathrm{b0}(m,n,p,q)\, \mathrm{b2}(p,q,m,n) \,| > + \mathrm{d1}(m,p)(\\
\mathrm{d2}(e,f,p,h) < | \mathrm{b1}(m,p)(\\
\mathrm{i1}(e,m,f,h)\, \mathrm{i1}(p,e,h,f)\, [e,m,f,h,p,e,h,f] \\
+ \mathrm{i1}(e,m,f,h)\, \mathrm{i1}(p,f,h,e)\, [e,m,f,h,p,f,h,e] \\
+ \mathrm{i1}(e,h,f,m)\, \mathrm{i1}(p,e,h,f)\, [e,h,f,m,p,e,h,f] \\
+ \mathrm{i1}(e,h,f,m)\, \mathrm{i1}(p,f,h,e)\, [e,h,f,m,p,f,h,e])| > + \\
\mathrm{d2}(e,e,p,h) < | \mathrm{b1}(m,p)\, \mathrm{i1}(e,m,e,h)\, \mathrm{i1}(p,e,h,e) \\
{[e,m,e,h,p,e,h,e]} \,| > + \mathrm{d2}(m,f,g,h) < | \mathrm{b1}(m,p)(\\
\mathrm{i1}(p,g,f,h)\, \mathrm{i1}(g,m,h,f)\, [p,g,f,h,g,m,h,f] \\
+ \mathrm{i1}(p,g,f,h)\, \mathrm{i1}(g,f,h,m)\, [p,g,f,h,g,f,h,m] \\
+ \mathrm{i1}(p,h,f,g)\, \mathrm{i1}(g,m,h,f)\, [p,h,f,g,g,m,h,f] \\
+ \mathrm{i1}(p,h,f,g)\, \mathrm{i1}(g,f,h,m)\, [p,h,f,g,g,f,h,m])| > + \\
\mathrm{d2}(m,f,g,g) < | \mathrm{b1}(m,p)\, \mathrm{i1}(p,g,f,g)\, \mathrm{i1}(g,m,g,f) \\
{[p,g,f,g,g,m,g,f]} \,| >) + \mathrm{d1}(m,p)\, \mathrm{d2}(e,g) < | \mathrm{b1}(m,p)(
\end{array}
$$

$$
\begin{array}{l}
\mathrm{i1}(p,m,e,g)\,\mathrm{c1}(g,e)\,[p,m,e,g,g,e] \\
+\,\mathrm{i1}(p,m,e,g)\,\mathrm{c2}(g,e,r3)\,[p,m,e,g,r3,e,g,r3] \\
+\,\mathrm{i1}(p,g,e,m)\,\mathrm{c1}(g,e)\,[p,g,e,m,g,e] \\
+\,\mathrm{i1}(p,g,e,m)\,\mathrm{c2}(g,e,r3)\,[p,g,e,m,r3,e,g,r3])| > + \\
\mathrm{d1}(m,s)\,\mathrm{d2}(s,g) < |\,\mathrm{b1}(m,s)(\\
\mathrm{i1}(s,m,s,g)\,\mathrm{c1}(g,s)\,[s,m,s,g,g,s] \\
+\,\mathrm{i1}(s,m,s,g)\,\mathrm{c2}(g,s,r3)\,[s,m,s,g,r3,s,g,r3])| > +\,\mathrm{d1}(t,p) \\
\mathrm{d2}(e,t) < |\,\mathrm{b1}(t,p)(\mathrm{i1}(p,t,e,t)\,\mathrm{c1}(t,e)\,[p,t,e,t,t,e] \\
+\,\mathrm{i1}(p,t,e,t)\,\mathrm{c2}(t,e,r3)\,[p,t,e,t,r3,e,t,r3])| > +\,\mathrm{d1}(t,s) \\
\mathrm{d2}(s,t) < |\,\mathrm{b1}(t,s)\,\Big(\frac{1}{2}\,\mathrm{i1}(s,t,s,t)\,\mathrm{c1}(t,s)\,[s,t,s,t,t,s] \\
+\,\frac{1}{2}\,\mathrm{i1}(s,t,s,t)\,\mathrm{c2}(t,s,r3)\,[s,t,s,t,r3,s,t,r3]\Big)\,| > + \\
\mathrm{d1}(m,n,p,q)\,\mathrm{d2}(n,q) < |\,\mathrm{b2}(m,n,p,q)(\\
\mathrm{c1}(p,m)\,\mathrm{c1}(q,n)\,[p,m,q,n] \\
+\,\mathrm{c1}(p,m)\,\mathrm{c2}(q,n,r3)\,[p,m,r3,n,q,r3] \\
+\,\mathrm{c2}(p,m,r2)\,\mathrm{c1}(q,n)\,[r2,m,p,r2,q,n] \\
+\,\mathrm{c2}(p,m,r2)\,\mathrm{c2}(q,n,r3)\,[r2,m,p,r2,r3,n,q,r3])| > + \\
\mathrm{d1}(m,m,p,q)\,\mathrm{d2}(m,q) < |\,\mathrm{b2}(m,m,p,q)(\\
\mathrm{c1}(p,m)\,\mathrm{c1}(q,m)\,[p,m,q,m] \\
+\,\mathrm{c1}(p,m)\,\mathrm{c2}(q,m,r3)\,[p,m,r3,m,q,r3] \\
+\,\mathrm{c2}(p,m,r2)\,\mathrm{c1}(q,m)\,[r2,m,p,r2,q,m] \\
+\,\mathrm{c2}(p,m,r2)\,\mathrm{c2}(q,m,r3)\,[r2,m,p,r2,r3,m,q,r3])| > + \\
\mathrm{d1}(m,n,p,p)\,\mathrm{d2}(n,p) < |\,\mathrm{b2}(m,n,p,p)(\\
\mathrm{c1}(p,m)\,\mathrm{c1}(p,n)\,[p,m,p,n] \\
+\,\mathrm{c1}(p,m)\,\mathrm{c2}(p,n,r3)\,[p,m,r3,n,p,r3] \\
+\,\mathrm{c2}(p,m,r2)\,\mathrm{c1}(p,n)\,[r2,m,p,r2,p,n] \\
+\,\mathrm{c2}(p,m,r2)\,\mathrm{c2}(p,n,r3)\,[r2,m,p,r2,r3,n,p,r3])| > + \\
\mathrm{d1}(m,m,p,p)\,\mathrm{d2}(m,p) < |\,\mathrm{b2}(m,m,p,p)\,\big(\mathrm{c1}(p,m)^{2}\,[p,m,p,m] \\
+\,\mathrm{c1}(p,m)\,\mathrm{c2}(p,m,r3)\,[p,m,r3,m,p,r3] \\
+\,\mathrm{c2}(p,m,r2)\,\mathrm{c1}(p,m)\,[r2,m,p,r2,p,m] \\
+\,\mathrm{c2}(p,m,r2)\,\mathrm{c2}(p,m,r3)\,[r2,m,p,r2,r3,m,p,r3]\big)\,| > + \\
\mathrm{d1}(m,m,p,p)(\mathrm{d2}(e,m,p,p) < |\,\mathrm{b2}(m,m,p,p)(\\
\mathrm{c1}(e,m)\,\mathrm{i1}(p,e,p,m)\,[e,m,p,e,p,m] \\
+\,\mathrm{c2}(e,m,r2)\,\mathrm{i1}(p,e,p,m)\,[r2,m,e,r2,p,e,p,m])| > + \\
\mathrm{d2}(m,m,g,p) < |\,\mathrm{b2}(m,m,p,p)(\\
\mathrm{c1}(p,g)\,\mathrm{i1}(g,m,p,m)\,[p,g,g,m,p,m] \\
+\,\mathrm{c2}(p,g,r2)\,\mathrm{i1}(g,m,p,m)\,[r2,g,p,r2,g,m,p,m])| >) + \\
\mathrm{d1}(m,n,p,q)(\mathrm{d2}(e,n,p,q) < |\,\mathrm{b2}(m,n,p,q)(\\
\mathrm{c1}(e,m)\,\mathrm{i1}(p,e,q,n)\,[e,m,p,e,q,n] \\
+\,\mathrm{c1}(e,m)\,\mathrm{i1}(p,n,q,e)\,[e,m,p,n,q,e] \\
+\,\mathrm{c2}(e,m,r2)\,\mathrm{i1}(p,e,q,n)\,[r2,m,e,r2,p,e,q,n] \\
+\,\mathrm{c2}(e,m,r2)\,\mathrm{i1}(p,n,q,e)\,[r2,m,e,r2,p,n,q,e])| > + \\
\mathrm{d2}(m,n,g,q) < |\,\mathrm{b2}(m,n,p,q)(\\
\mathrm{c1}(p,g)\,\mathrm{i1}(g,m,q,n)\,[p,g,g,m,q,n] \\
+\,\mathrm{c1}(p,g)\,\mathrm{i1}(g,n,q,m)\,[p,g,g,n,q,m] \\
+\,\mathrm{c2}(p,g,r2)\,\mathrm{i1}(g,m,q,n)\,[r2,g,p,r2,g,m,q,n]
\end{array}
$$

$$\begin{array}{l}
+\,\mathrm{c2}(p,g,\mathit{r2})\,\mathrm{i1}(g,n,q,m)\,[\mathit{r2},g,p,\mathit{r2},g,n,q,m])|>)\,+\\
\mathrm{d1}(m,m,p,q)(\mathrm{d2}(e,m,p,q)<|\,\mathrm{b2}(m,m,p,q)(\\
\mathrm{c1}(e,m)\,\mathrm{i1}(p,e,q,m)\,[e,m,p,e,q,m]\\
+\,\mathrm{c1}(e,m)\,\mathrm{i1}(p,m,q,e)\,[e,m,p,m,q,e]\\
+\,\mathrm{c2}(e,m,\mathit{r2})\,\mathrm{i1}(p,e,q,m)\,[\mathit{r2},m,e,\mathit{r2},p,e,q,m]\\
+\,\mathrm{c2}(e,m,\mathit{r2})\,\mathrm{i1}(p,m,q,e)\,[\mathit{r2},m,e,\mathit{r2},p,m,q,e])|>\,+\\
\mathrm{d2}(m,m,g,q)<|\,\mathrm{b2}(m,m,p,q)(\\
\mathrm{c1}(p,g)\,\mathrm{i1}(g,m,q,m)\,[p,g,g,m,q,m]\\
+\,\mathrm{c2}(p,g,\mathit{r2})\,\mathrm{i1}(g,m,q,m)\,[\mathit{r2},g,p,\mathit{r2},g,m,q,m])|>)\,+\\
\mathrm{d1}(m,n,p,p)(\mathrm{d2}(e,n,p,p)<|\,\mathrm{b2}(m,n,p,p)(\\
\mathrm{c1}(e,m)\,\mathrm{i1}(p,e,p,n)\,[e,m,p,e,p,n]\\
+\,\mathrm{c2}(e,m,\mathit{r2})\,\mathrm{i1}(p,e,p,n)\,[\mathit{r2},m,e,\mathit{r2},p,e,p,n])|>\,+\\
\mathrm{d2}(m,n,g,p)<|\,\mathrm{b2}(m,n,p,p)(\\
\mathrm{c1}(p,g)\,\mathrm{i1}(g,m,p,n)\,[p,g,g,m,p,n]\\
+\,\mathrm{c1}(p,g)\,\mathrm{i1}(g,n,p,m)\,[p,g,g,n,p,m]\\
+\,\mathrm{c2}(p,g,\mathit{r2})\,\mathrm{i1}(g,m,p,n)\,[\mathit{r2},g,p,\mathit{r2},g,m,p,n]\\
+\,\mathrm{c2}(p,g,\mathit{r2})\,\mathrm{i1}(g,n,p,m)\,[\mathit{r2},g,p,\mathit{r2},g,n,p,m])|>)\,+\\
\mathrm{d1}(m,t,p,q)\,\mathrm{d2}(m,t)<|\,\mathrm{b2}(m,t,p,q)(\\
\mathrm{i1}(p,t,q,t)\,\mathrm{c1}(t,m)\,[p,t,q,t,t,m]\\
+\,\mathrm{i1}(p,t,q,t)\,\mathrm{c2}(t,m,\mathit{r3})\,[p,t,q,t,\mathit{r3},m,t,\mathit{r3}])|>\,+\\
\mathrm{d1}(m,t,p,p)\,\mathrm{d2}(m,t)<|\,\mathrm{b2}(m,t,p,p)\,\Bigg(\\
\frac{1}{2}\,\mathrm{i1}(p,t,p,t)\,\mathrm{c1}(t,m)\,[p,t,p,t,t,m]\\
+\,\frac{1}{2}\,\mathrm{i1}(p,t,p,t)\,\mathrm{c2}(t,m,\mathit{r3})\,[p,t,p,t,\mathit{r3},m,t,\mathit{r3}]\Bigg)\,|>\,+\\
\mathrm{d1}(m,n,p,s)\,\mathrm{d2}(s,p)<|\,\mathrm{b2}(m,n,p,s)(\\
\mathrm{i1}(s,m,s,n)\,\mathrm{c1}(p,s)\,[s,m,s,n,p,s]\\
+\,\mathrm{i1}(s,m,s,n)\,\mathrm{c2}(p,s,\mathit{r3})\,[s,m,s,n,\mathit{r3},s,p,\mathit{r3}])|>\,+\\
\mathrm{d1}(m,m,p,s)\,\mathrm{d2}(s,p)<|\,\mathrm{b2}(m,m,p,s)\,\Bigg(\\
\frac{1}{2}\,\mathrm{i1}(s,m,s,m)\,\mathrm{c1}(p,s)\,[s,m,s,m,p,s]\\
+\,\frac{1}{2}\,\mathrm{i1}(s,m,s,m)\,\mathrm{c2}(p,s,\mathit{r3})\,[s,m,s,m,\mathit{r3},s,p,\mathit{r3}]\Bigg)\,|>\,+\\
\mathrm{d1}(m,m,p,p)(\mathrm{d2}(m,g)<|\,\mathrm{b2}(m,m,p,p)(\\
\mathrm{i1}(p,m,p,g)\,\mathrm{c1}(g,m)\,[p,m,p,g,g,m]\\
+\,\mathrm{i1}(p,m,p,g)\,\mathrm{c2}(g,m,\mathit{r3})\,[p,m,p,g,\mathit{r3},m,g,\mathit{r3}])|>\,+\\
\mathrm{d2}(e,p)<|\,\mathrm{b2}(m,m,p,p)(\\
\mathrm{i1}(p,m,e,m)\,\mathrm{c1}(p,e)\,[p,m,e,m,p,e]\\
+\,\mathrm{i1}(p,m,e,m)\,\mathrm{c2}(p,e,\mathit{r3})\,[p,m,e,m,\mathit{r3},e,p,\mathit{r3}])|>)\,+\\
\mathrm{d1}(m,n,p,q)(\mathrm{d2}(m,g)<|\,\mathrm{b2}(m,n,p,q)(\\
\mathrm{i1}(p,n,q,g)\,\mathrm{c1}(g,m)\,[p,n,q,g,g,m]\\
+\,\mathrm{i1}(p,n,q,g)\,\mathrm{c2}(g,m,\mathit{r3})\,[p,n,q,g,\mathit{r3},m,g,\mathit{r3}]\\
+\,\mathrm{i1}(p,g,q,n)\,\mathrm{c1}(g,m)\,[p,g,q,n,g,m]\\
+\,\mathrm{i1}(p,g,q,n)\,\mathrm{c2}(g,m,\mathit{r3})\,[p,g,q,n,\mathit{r3},m,g,\mathit{r3}])|>\,+\\
\mathrm{d2}(e,p)<|\,\mathrm{b2}(m,n,p,q)(\\
\mathrm{i1}(q,m,e,n)\,\mathrm{c1}(p,e)\,[q,m,e,n,p,e]\\
+\,\mathrm{i1}(q,m,e,n)\,\mathrm{c2}(p,e,\mathit{r3})\,[q,m,e,n,\mathit{r3},e,p,\mathit{r3}]
\end{array}$$

$+\,\mathrm{i1}(q,n,e,m)\,\mathrm{c1}(p,e)\,[q,n,e,m,p,e]$
$+\,\mathrm{i1}(q,n,e,m)\,\mathrm{c2}(p,e,r3)\,[q,n,e,m,r3,e,p,r3])|>)\,+$
$\mathrm{d1}(m,m,p,q)(\mathrm{d2}(m,g)<|\,\mathrm{b2}(m,m,p,q)($
$\mathrm{i1}(p,m,q,g)\,\mathrm{c1}(g,m)\,[p,m,q,g,g,m]$
$+\,\mathrm{i1}(p,m,q,g)\,\mathrm{c2}(g,m,r3)\,[p,m,q,g,r3,m,g,r3]$
$+\,\mathrm{i1}(p,g,q,m)\,\mathrm{c1}(g,m)\,[p,g,q,m,g,m]$
$+\,\mathrm{i1}(p,g,q,m)\,\mathrm{c2}(g,m,r3)\,[p,g,q,m,r3,m,g,r3])|>\,+$
$\mathrm{d2}(e,p)<|\,\mathrm{b2}(m,m,p,q)($
$\mathrm{i1}(q,m,e,m)\,\mathrm{c1}(p,e)\,[q,m,e,m,p,e]$
$+\,\mathrm{i1}(q,m,e,m)\,\mathrm{c2}(p,e,r3)\,[q,m,e,m,r3,e,p,r3])|>)\,+$
$\mathrm{d1}(m,n,p,p)(\mathrm{d2}(m,g)<|\,\mathrm{b2}(m,n,p,p)($
$\mathrm{i1}(p,n,p,g)\,\mathrm{c1}(g,m)\,[p,n,p,g,g,m]$
$+\,\mathrm{i1}(p,n,p,g)\,\mathrm{c2}(g,m,r3)\,[p,n,p,g,r3,m,g,r3])|>\,+$
$\mathrm{d2}(e,p)<|\,\mathrm{b2}(m,n,p,p)($
$\mathrm{i1}(p,m,e,n)\,\mathrm{c1}(p,e)\,[p,m,e,n,p,e]$
$+\,\mathrm{i1}(p,m,e,n)\,\mathrm{c2}(p,e,r3)\,[p,m,e,n,r3,e,p,r3]$
$+\,\mathrm{i1}(p,n,e,m)\,\mathrm{c1}(p,e)\,[p,n,e,m,p,e]$
$+\,\mathrm{i1}(p,n,e,m)\,\mathrm{c2}(p,e,r3)\,[p,n,e,m,r3,e,p,r3])|>)\,+$
$\mathrm{d1}(m,n,p,q)(\mathrm{d2}(e,f,p,q)<|\,\mathrm{b2}(m,n,p,q)($
$\mathrm{i1}(e,m,f,n)\,\mathrm{i1}(p,e,q,f)\,[e,m,f,n,p,e,q,f]$
$+\,\mathrm{i1}(e,m,f,n)\,\mathrm{i1}(p,f,q,e)\,[e,m,f,n,p,f,q,e]$
$+\,\mathrm{i1}(e,n,f,m)\,\mathrm{i1}(p,e,q,f)\,[e,n,f,m,p,e,q,f]$
$+\,\mathrm{i1}(e,n,f,m)\,\mathrm{i1}(p,f,q,e)\,[e,n,f,m,p,f,q,e])|>\,+$
$\mathrm{d2}(e,e,p,q)<|\,\mathrm{b2}(m,n,p,q)\,\mathrm{i1}(e,m,e,n)\,\mathrm{i1}(p,e,q,e)$
$[e,m,e,n,p,e,q,e]\,|>+\,\mathrm{d2}(m,n,g,h)<|\,\mathrm{b2}(m,n,p,q)($
$\mathrm{i1}(p,g,q,h)\,\mathrm{i1}(g,m,h,n)\,[p,g,q,h,g,m,h,n]$
$+\,\mathrm{i1}(p,g,q,h)\,\mathrm{i1}(g,n,h,m)\,[p,g,q,h,g,n,h,m]$
$+\,\mathrm{i1}(p,h,q,g)\,\mathrm{i1}(g,m,h,n)\,[p,h,q,g,g,m,h,n]$
$+\,\mathrm{i1}(p,h,q,g)\,\mathrm{i1}(g,n,h,m)\,[p,h,q,g,g,n,h,m])|>\,+$
$\mathrm{d2}(m,n,g,g)<|\,\mathrm{b2}(m,n,p,q)\,\mathrm{i1}(p,g,q,g)\,\mathrm{i1}(g,m,g,n)$
$[p,g,q,g,g,m,g,n]\,|>+\,\mathrm{d2}(e,n,g,q)<|\,\mathrm{b2}(m,n,p,q)($
$\mathrm{i1}(p,m,e,g)\,\mathrm{i1}(g,e,q,n)\,[p,m,e,g,g,e,q,n]$
$+\,\mathrm{i1}(p,m,e,g)\,\mathrm{i1}(g,n,q,e)\,[p,m,e,g,g,n,q,e]$
$+\,\mathrm{i1}(p,g,e,m)\,\mathrm{i1}(g,e,q,n)\,[p,g,e,m,g,e,q,n]$
$+\,\mathrm{i1}(p,g,e,m)\,\mathrm{i1}(g,n,q,e)\,[p,g,e,m,g,n,q,e])|>)$

C: Example of Maple code for algebra of generators

Function *gcir* transforms an arbitrary chain of replacement operators into a linear combination of circular chains (sse Chap. 3). In order to get the following chain $\hat{E}_{ag}\hat{E}_{bg}\hat{E}_{ga}\hat{E}_{gb}$ as a linear combination of circular chains one has to put:

$$gcir([a,g,b,g,g,a,g,b])$$

The output has the form

$$\& + ([a,g,a,b,g,b],[a,g,b,a],(\& - [a,g,a]])$$

which means $\mathcal{E}_{aga}\mathcal{E}_{bgb} + \mathcal{E}_{agba} - \mathcal{E}_{aga}$. *aGroup* is extracted from the Maple library and slightly modified. It defines an abelian group with multiplication designated by $\&+$ and inversion denoted by $\&-$.

```
`define/aGroup`:=
proc(OpName,Identity,Inverse)
local x;
options `Copyright 1993 by Waterloo Maple Software`;
    if
    nargs <> 3 or not type(OpName,string) or not type(Inverse,{name,procedure})
     then
        ERROR(`invalid arguments`)
    fi;
    proc()
    local i,t,x;
    options remember;
        t := [args];
        t := map(
            proc(y)
                if type(y,function) and op(0,y) = DOpName then op(y) else y fi
            end                                                               ,
            t);
        t := sort(t);
        t := map(proc(y) if y <> DIdentity then y fi end,t);
        for i while i < nops(t) do
            if t[i] = DInverse(t[i+1]) then
                t := subsop(i = NULL,i+1 = NULL,t); i := i-1
            fi
        od;
        if nops(t) = 0 then RETURN(DIdentity) fi;
        if nops(t) = 1 then RETURN(t[1]) fi;
        RETURN(subs(x = op(t),'DOpName(x)'))
    end;
    if type(OpName,protected) then
        unprotect('OpName'); lprint(`Warning: new definition for`,OpName)
    fi;
    OpName :=
        subs('DOpName' = OpName,'DIdentity' = Identity,'DInverse' = Inverse,");
    if type(op(Inverse),name) then
```

```
        proc(a)
        options remember;
            if nargs <> 1 then ERROR(`invalid arguments`) fi;
            if type(a,function) and op(0,a) = DInverse then RETURN(op(a)) fi;
            if a = DIdentity then RETURN(DIdentity) fi;
            'DInverse(args)'
        end;
        Inverse := subs('DInverse' = Inverse,'DIdentity' = Identity,");
        proc(a)
            if nargs = 1 and type(a,function) and op(0,a) = DOpName then
                if 1 < nops(a) then
                    DOpName(expand(DInverse(DOpName(op(2 .. nops(a),a)))),
                        expand(DInverse(op(1,a))))
                else DInverse(args)
                fi
            fi
        end;
        `expand/`.Inverse := subs('DOpName' = OpName,'DInverse' = Inverse,")
    else
        if Inverse(Identity) <> Identity or Inverse(Inverse(x)) <> x then
            ERROR(`Inverse operator defined incorrectly`)
        fi
    fi;
    NULL
end;

define(aGroup(`&+`, [], `&-`));

delta1:=proc(x) local i,j,k,l,nl,nl1,nl2;
  if type(x,list) and nops(x)=4 then
     i:=x[1]; j:=x[2]; k:=x[3]; l:=x[4];
     if j=k then nl1:=[i,l] else nl1:=[ ] fi;
     if i=l then nl2:= &- [k,j] else nl2:=[ ] fi;
     nl1 &+ nl2
  else
     ERROR(`Wrong number of type of arguments`)
  fi
end;

 find12:=proc(x,i,j) local n,k;
       n:=nops(x);
           for k to n/2 do
                 if x[2*k-1]=j and x[2*k]=i then  RETURN(k) fi
           od;
           for k to n/2 do
                 if x[2*k-1]=j then  RETURN(k) fi
           od;
       RETURN(0)
 end;

shiftqtop:=proc(xlist,p,q) local i,j,k,n,bp,aq,bpq,xx,xxn,cl;
    if not type(xlist,list) then ERROR(`Invalid argument`) fi;
    xx:=xlist;
    n:=nops(xlist);
    if p<1 or 2*q>n or p>q then ERROR(`Invalid argument`) fi;
    bp:=[op(1 .. 2*p, xx)];
    if 2*q=n then aq:= [] else aq:=[op(2*q+1 .. n, xx)] fi;
    bpq:=[op(2*p+1 .. 2*q-2, xx)];
    k:=nops(bpq)/2;
    xxn:= [];
    for i to k do
     cl:=delta1([bpq[2*k-2*i+1],bpq[2*k-2*i+2],xx[2*q-1],xx[2*q]]);
```

```
     if not cl=[] then
       if op(0,cl)=`&+` then cl:=[op(cl)] else cl:=[cl] fi;
       if 2*k-2*i=0 then cbpqb:= [ ]
             else cbpqb:=[op(1 .. 2*k-2*i, bpq)] fi;
       if 2*k-2*i+2=nops(bpq) then cbpqe:= [ ]
             else cbpqe:=[op(2*k-2*i+3 .. nops(bpq), bpq)] fi;
       xxn:=[op(xxn), op(map( proc(y,bp,aq,cbpqb,cbpqe)
          if op(0,y)=`&-` then &-([op(bp),op(cbpqb),op(op(y)),op(cbpqe),op(aq)])
          else [op(bp),op(cbpqb),op(y),op(cbpqe),op(aq)] fi end, cl,
                                                          bp,aq,cbpqb,cbpqe))]
     fi
    od;
    cl:= [op(bp),op(2*q-1 .. 2*q,xx),op(bpq),op(aq)];
    &+ ( cl, op(xxn) );
end;


 findxx:=proc(x) local n,k;
       n:=nops(x);
           for k to n/2 do
                 if x[2*k-1]=x[2*k] then  RETURN(k) fi
           od;
       RETURN(0)
 end;

shiftxx:=proc(xlist,p) local i,j,k,n,bp,aq,xx,xxn,cl;
    if not type(xlist,list) then ERROR(`Invalid argument`) fi;
    xx:=xlist;
    n:=nops(xlist);
    if p<1 or p>n then ERROR(`Invalid argument`) fi;
    if p=n then RETURN(xx) fi;
    if p=1 then bp:= [] else bp:=[op(1 .. 2*p-2, xx)] fi;
    aq:=[op(2*p+1 .. n, xx)];
    k:=nops(aq)/2;
    xxn:= [];
    for i to k do
     cl:=delta1([xx[2*p-1],xx[2*p],aq[2*i-1],aq[2*i]]);
     if not cl=[] then
       if op(0,cl)=`&+` then cl:=[op(cl)] else cl:=[cl] fi;
       if i=k then cbpqe:= [ ]
             else cbpqe:=[op(2*i+1 .. 2*k, aq)] fi;
       if i=1 then cbpqb:= [ ]
             else cbpqb:=[op(1 .. 2*i-2, aq)] fi;
       xxn:=[op(xxn), op(map( proc(y,bp,cbpqb,cbpqe)
          if op(0,y)=`&-` then &-([op(bp),op(cbpqb),op(op(y)),op(cbpqe)])
          else [op(bp),op(cbpqb),op(y),op(cbpqe)] fi end, cl, bp,cbpqb,cbpqe))]
     fi
    od;
    cl:= [op(bp),op(aq),op(2*p-1 .. 2*p,xx)];
    &+ ( cl, op(xxn) );
end;

 findxx2:=proc(x) local n,k,i,xx;
       if op(0,x)=`&-` then xx:=op(x) else xx:=x fi;
       n:=nops(xx);
       k:=findxx(xx);
       if k=0 or k=n/2 then RETURN(0) fi;
       for i to n/2-k do
         if not xx[2*k+2*i-1]=xx[2*k+2*i] then  RETURN(k) fi
       od;
       RETURN(0)
 end;
```

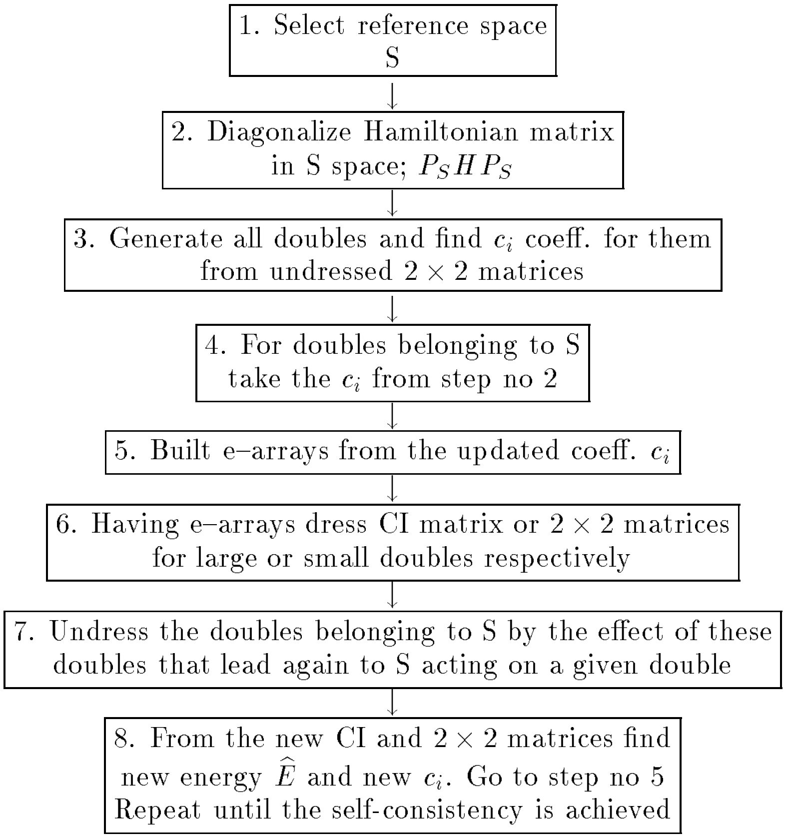
```
shiftxx2:=proc(x,p) local xx;
    if p=0 then RETURN(x) fi;
    if op(0,x)=`&-` and type(op(x),list) then
       xx:=shiftxx(op(x),p);
       if op(0,xx)=`&+` then xx:=[op(xx)]
                        else xx:=[xx] fi;
       xx:=map(proc(y) &- (y) end, xx)
    else
       if type(x,list) then
            xx:=shiftxx(x,p);
            if op(0,xx)=`&+` then xx:=[op(xx)]
                             else xx:=[xx] fi
       else ERROR(`Invalid argument`) fi
    fi;
    &+ (op(xx));
end;

extractxx:=proc(xlist) local i,j,k,n,xx,it;
    if not type(xlist,list) then ERROR(`Invalid argument`) fi;
    if findxx(xlist)=0 then RETURN(xlist) fi;
    xx:=[xlist]; i:=1;
    while i>0 do
      ii:=[seq(j, j=1 .. nops(it))];
      xx:=map(proc(y) local zz;
                  zz:=shiftxx2(y,findxx2(y));
                  if op(0,zz)=`&+` then [op(zz)]
                                   else [zz] fi
              end, xx);
      n:=nops(xx);
      xx:=[seq(op(op(j,xx)), j=1 .. n)];
      it:=map(findxx2,xx);
      i:=max(op(it));
    od;
    &+ (op(xx));
end;


getinter:=proc(x) local n,k,i,c,ll;
   n:=nops(x);
   ll:=[0,0];
   for k to n/2-1 do
     c:=x[2*k];
     for i from k+2 to n/2-1 do
       if x[2*i]=c and x[2*i+1]=c then
         ll:=subsop(1=k,2=i,ll);
       fi
     od;
   od;
   ll;
end;

getinter1:=proc(x) local k,i,j,ll,c,l,yes1;
   ll:=getinter(x);
   k:=ll[1]; i:=ll[2];
   if k=0 then RETURN(0) fi;
   yes1:=true;
   for j to k do
     c:=x[2*j-1];
     if j>1 then
       yes1:=true;
       if x[2*j-2]=c then yes1:=false fi
```

```
     fi;
     for l from k+1 to i-1 do
        if x[2*l]=c and yes1 then RETURN(0) fi
     od;
   od;
   RETURN(k)
end;

findcirc:=proc(xlist) local n,k,i,yes1,xx;
   if op(0,xlist)=`&-` then xx:=op(xlist) else xx:=xlist fi;
   if not type(xx,list) then ERROR(`Invalid argument`) fi;
   n:=nops(xx);
   for k to n/2-1 do
     if not xx[2*k]=xx[2*k+1] then
        yes1:=false;
        for i to k do
           if xx[2*k]=xx[2*k-2*i+1] then
              if i=k then yes1:=true
              else
                   if not xx[2*k-2*i+1]=xx[2*k-2*i] then
                                            yes1:=true fi
              fi
           fi
        od;
        if not yes1 then RETURN(k) fi
     fi;
   od;
   getinter1(xx);
end;

 find12bis:=proc(x,p) local xx,i,j,k,l,n;
       if p=0 then RETURN(0) fi;
       if op(0,x)=`&-` then xx:=op(x) else xx:=x fi;
       n:=nops(xx);
       i:=xx[2*p-1]; j:=xx[2*p];
       xx:=[seq(op(2*p+2+k,xx), k=1 .. n-2*p-2)];
       l:=find12(xx,i,j);
       if l>0 then RETURN(l+1+p) fi;
       RETURN(0)
 end;

shiftqtop2:=proc(x,p,q) local xx;
    if p=0 or q=0 then RETURN(x) fi;
    if op(0,x)=`&-` and type(op(x),list) then
       xx:=shiftqtop(op(x),p,q);
       if op(0,xx)=`&+` then xx:=[op(xx)]
                        else xx:=[xx] fi;
       xx:=map(proc(y) &- (y) end, xx)
    else
       if type(x,list) then
            xx:=shiftqtop(x,p,q);
            if op(0,xx)=`&+` then xx:=[op(xx)]
                                 else xx:=[xx] fi
       else ERROR(`Invalid argument`) fi
    fi;
    &+ (op(xx));
end;

getcirc:=proc(xlist) local i,j,k,n,xx,it;
    if not type(xlist,list) then ERROR(`Invalid argument`) fi;
    if findcirc(xlist)=0 then RETURN(xlist) fi;
    xx:=extractxx(xlist);
```

```
    if op(0,xx)=`&+` then xx:=[op(xx)] else xx:=[xx] fi; i:=1;
    while i>0 do
      xx:=map(proc(y) local zz,ll;
                  ll:=findcirc(y); lprint(ll);
                  zz:=shiftqtop2(y,ll,find12bis(y,ll));
                  if op(0,zz)=`&+` then [op(zz)]
                                   else [zz] fi
              end, xx);
      n:=nops(xx); lprint(xx);
      xx:=[seq(op(op(j,xx)), j=1 .. n)]; lprint(xx);
      xx:=map(proc(y) local zz;
                if op(0,y)=`&-` then
                  zz:=extractxx(op(y));
                  if op(0,zz)=`&+` then zz:=[op(zz)]
                                   else zz:=[zz] fi;
                  map(proc(y) &- (y) end,zz)
                else
                  zz:=extractxx(y);
                  if op(0,zz)=`&+` then [op(zz)]
                                   else [zz] fi
                fi
              end,xx);
      n:=nops(xx); lprint(xx);
      xx:=[seq(op(op(j,xx)), j=1 .. n)]; lprint(xx);
      it:=map(findcirc,xx);
      i:=max(op(it));
    od;
    &+ (op(xx));
end;
shortcirc:=proc(x) local i,j,l,n,xx,xxl,yes1;
   yes1:=false;
   if op(0,x)=`&-` then xx:=op(x); yes1:=true else xx:=x fi;
   l:=xx[1]; xxl:=[l];
   n:=nops(xx)/2;
   for i to n-1 do
      if xx[2*i]=l then
       xxl:=[op(xxl),xx[2*i]];
       l:=xx[2*i+1]; xxl:=[op(xxl),l];
      else
        if xx[2*i]=xx[2*i+1] then xxl:=[op(xxl),xx[2*i]]
        else ERROR(`NON-CIRCULAR chain`) fi
      fi
   od;
   xxl:=[op(xxl),xx[2*n]];
   if yes1 then &- (xxl) else xxl fi
end;
conj:=proc(xlist) local i,n,xx,xxl,yes1;
  if op(0,xlist)=`&-` then xx:=op(xlist); yes1:=true
  else xx:=xlist; yes1:=false fi;
  n:=nops(xx); xxl:=[];
  for i to n do
    xxl:=[op(xxl),xx[n+1-i]];
  od;
  if yes1 then &- (xxl) else xxl fi
end;
multpr:=proc(xlist) local i,j,k,n,l,xx,xxk,xxl;
  n:=nops(xlist);
  xxl:=[]; xx:=xlist;
  while n>0 do
    l:=1; xxk:=[];
    for i to n-1 do
      if xx[1]=xx[i+1] or xx[1]=conj(xx[i+1]) then l:=l+1
```

```
      elif xx[1]=&-(xx[i+1]) or xx[1]=&-(conj(xx[i+1])) then
           l:=l-1
      else xxk:=[op(xxk),xx[i+1]]  fi;
    od;
    if l=1 then xxl:=[op(xxl),xx[1]]
    else if not l=0 then xxl:=[op(xxl),l*xx[1]] fi fi;
    xx:=xxk; n:=nops(xx);
  od;
  &+ (op(xxl));
end;
gcir:=proc(xlist) local i,j,k,n,xx,it;
   if not type(xlist,list) then ERROR(`Invalid argument`) fi;
   n:=nops(xlist); n:=n mod 2;
   if not n=0 then ERROR(`Improper length of the chain`) fi;
   xx:=getcirc(xlist);
   n:=nops(xx);
   if op(0,xx)=`&+` then xx:=[op(xx)] else xx:=[xx] fi;
   xx:= map(proc(y) shortcirc(y) end,xx);
   multpr(xx)
end;
```

D: The scheme of the $(SC)^2(CI+PT)$ program



1. In practice two iterations of CIPSI were done, so that some doubles and most important triples and quadruples were included in S.

6. For a given doubly excited determinant $|i\rangle = D_i^+|0\rangle$ one uses the e–arrays as described in Chap. 7 to evaluate δ_{ii}, instead of explicit summation over EPV (with respect to i) contributions. Then the corresponding diagonal element in 2×2 or CI matrix is modified.

7. The undressing is done by the explicit decomposition of all selected triples and quadruples, as described in the ref. [86].

8. The contribution of small doubles is added to $\langle 0|H|0\rangle$ element of the dressed CI matrix, which is then diagonalized to obtain next value of $\widehat{E}$ and the coefficients c_i of large doubles.

List of Tables

List of Figures

Bibliography

[1] J.A. Pople, M. Head-Gordon, D.J. Fox, K. Raghavachari and L.A. Curtiss, J.Chem.Phys. **90**, 5622 (1989)

[2] M. Dupuis, J.D. Watts, H.O. Villar and G.J.B. Hurst, Comp.Phys.Comm. **52**, 415 (1989)

[3] M.W. Schmidt, J.A. Boatz, K.K. Baldrige, S. Koseki, M.S. Gordon, S.T. Elbert and B. Lam, QCPE Bulletin Vol.**7**, 115 (1987)

[4] I. Białynicki–Birula, M. Cieplak and J. Kamiński, *Teoria kwantów: mechanika falowa* (in polish), PWN, Warszawa 1991

[5] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular electronic structure theory*

[6] A. Szabo and N. Ostlund, *Modern Quantum Chemistry*, McGraw–Hill, New York 1989

[7] A. Gołębiowski, *Elementy mechaniki i chemii kwantowej* (in polish), PWN, Warszawa 1982

[8] W. Kołos, *Chemia Kwantowa* (in polish), PWN, Warszawa 1986, p. 145

[9] P.O. Löwdin, Int.J.Quantum Chem. **55**, 77 (1995)

[10] C.C.J. Roothan, Rev. Mod. Phys. **23**, 69 (1951)

[11] E.A. Hylleraas, Z.Physik **54**, 347 (1929); W. Kutzelnigg, Theor.Chim.Acta **68**, 445 (1985); W. Woźnicki in it Theory of Electronic Shell in Atoms and Molecules, ed. A. Jucys, Vilnius 1971, p. 103; G. Pestka and W. Woźnicki, Chem.Phys.Lett. submitted for publication

[12] L. Laaksonen, P. Pyykkö and D. Sundholm, Comp.Phys.Rep. **4**, 313 (1986)

[13] W. Kohn and L.J. Sham, Phys. Rev. A **140**, 1133 (1965); R.G. Parr and W. Yang, *Density Functional Theory for Atoms and Molecules*, Oxford 1989; N.C. Handy, to be published

[14] J. Sadlej, *Obliczeniowe metody chemii kwantowej* (in polish), PWN, Warszawa 1988

[15] A. Hurley, it Electron Correlation in Small Molecules, Academic Press, London 1976

[16] J. Wasilewski, Dissertation, Nicholas Copernicus University, Toruń 1990

[17] I. Lindgren and J. Morisson, *Atomic many-body theory*, Springer, Berlin 1986

[18] K.A. Brueckner, Phys. Rev. **97**, 1353 (1953); **100**, 1353 (1955); J. Goldstone, Proc.R.Soc. London Ser. A, 239 (1957)

[19] R.J. Bartlett, Ann.Rev.Phys.Chem. **32**, 359 (1981)

[20] R.J. Bartlett, J.Phys.Chem. **93**, 1697 (1989)

[21] R.J. Bartlett and G.D. Purvis III, Phys. Scripta **21**, 255 (1980)

[22] J.A. Pople, J.S. Binkley and R. Seeger, Int.J.Quantum Chem.Symp. **10**, 1 (1976); J.A. Pople, R. Seeger and R. Krishnan, Int.J.Quantum Chem.Symp. **11**, 149 (1977); J.A. Pople, R. Krishnan, H.B. Schlegel and J.S. Binkley, Int.J.Quantum Chem. **14**, 545 (1978); J.A. Pople, in: Energy, structure and reactivity. Eds. D.W. Smith, W.B. McRae (Wiley, N.Y. 1973), pp. 51–61;

[23] E. Davidson in *Realtivistic and Electron Correlation Effects in Molecules and Solids*, Ed. G.L. Malli, Plenum Press 1994, p. 105

[24] R.J. Bartlett and G.D. Purvis III, Int.J.Quantum Chem. **14**, 561 (1978)

[25] G.D. Purvis III and R.J. Bartlett, J.Chem.Phys. **76**, 1910 (1982)

[26] J.Paldus, J.Chem.Phys. **67**, 303 (1977); I. Lindgren, J.Phys. B **7**, 2441 (1974)

[27] J. Hubbard, Proc.R.Soc. London Ser. A, 240 (1957), ibid 243, 244 (1958)

[28] K. Jankowski, J. Paldus and P. Piecuch, Theor.Chim.Acta(Berl.) **80**, 223 (1991)

[29] J. Olsen, P. Jørgensen and J. Simons, Chem.Phys.Lett. **169**, 463 (1990)

[30] R. Harrison and E. Stahlberg, Parallel Computing (1994)

[31] I. Shavitt, in *Methods of Electronic Structure Theory, Vol 3,* H.F.Schaefer III, Ed, Plenum Press, New York 1977, pp 189-275.

[32] B.O. Roos, Chem.Phys.Lett. **15**, 153 (1972); B.O. Roos and P. E. M. Siegbahn, in: *Methods of Electronic Structure Theory*, ed. H. F. Schaefer III (Plenum, New York, 1977), p. 277

[33] J. Paldus, J.Chem.Phys. **61**, 5321 (1974); F.A. Matsen and R. Pauncz, The Unitary Group in Quantum Chemistry. (Elsevier, Amsterdam, 1986); P.E.M. Siegbahn, J.Chem.Phys. **72**, 1647 (1980); R.D. Kent and M. Schlesinger, Int.J.Quantum Chem. **30**, 737 (1986); H. Lischka, R. Shepard, F.B. Brown and I. Shavitt, Int.J.Quantum Chem. **S15**, 91 (1981); P. Saxe, D.J. Fox, H.F. Schaefer and N.C. Handy, J.Chem.Phys. **77**, 5584 (1982)

[34] W. Duch and J. Karwowski, Int.J.Quantum Chem. **22**, 783 (1982); Comp.Phys.Rep. **2**, 92 (1985); Theor.Chim.Acta(Berl.) **71**, 187 (1987); W. Duch, Graphical Representation of Model Spaces, *Lecture Notes in Chemistry*, Vol. **42** (Springer, 1986); A. Guldberg, S. Rettrup, G.L. Bendazzoli and P. Palmieri, Int.J.Quantum Chem. **S21**, 513 (1987)

[35] J. Karwowski, Theor.Chim.Acta(Berl.) **29**, 151 (1973)

[36] M. Kotani, A. Amemiya, E. Ishiguro and T. Kimura, it Tables of Molecular Integrals, Maruzen, Tokyo 1955

[37] K.J. Werner and P. Knowles, J.Chem.Phys. **89**, 5803 (1988)

[38] P.E.M. Siegbahn, Int.J.Quantum Chem. **18**, 1229 (1980); **23**, 1869 (1983)

[39] W. Duch, Chem.Phys.Lett. **162**, 56 (1989)

[40] J. Cioslowski, Phys.Rev A **36**, 374 (1987); Phys.Rev.Lett. **58**, 83 (1987); T. Yoshida and K. Iguchi, J.Chem.Phys. **91**, 4249 (1989)

[41] P.J. Knowles, Chem.Phys.Lett. **134**, 512 (1987)

[42] G.L. Bendazzoli, G. Fano, G. Ortolani and P. Palmieri, J.Chem.Phys. **76**, 2498 (1982); G.L. Bendazzoli, P. Palmieri, G. Fano and G. Ortolani, J.Chem.Phys. **83**, 674 (1985)

[43] E.R. Davidson, J.Comp.Phys. **17**: 87 (1975)

[44] W. Duch and G.H.F. Diercksen, Phys.Rev A **46**, 95 (1992)

[45] P. Durand and J-P Malrieu, *Effective Hamiltonians and Pseudo-Operators as Tools for Rigorous Modelling*, in: *Ab initio* methods in Quantum Chemistry I, K.P. Lawley, ed (John Wiley & Sons, 1987)

[46] J.A. Pople, J.S. Binkley and R. Seeger, Int.J.Quantum Chem. **S11**, 149 (1977)

[47] N.C. Handy, P.J. Knowles and K. Somasundram, Theor.Chim.Acta(Berl.) **68**,87 (1985); P.J.Knowles, K. Somasundram, N.C. Handy and K. Hirao, Chem.Phys.Lett. **113**,8 (1985)

[48] P.J. Knowles, J.S. Andrews, R.D. Amos, N.C. Handy and J.A. Pople, Chem.Phys.Lett. **186**, 130 (1991); W.J. Lauderdale, J.F. Stanton, J. Gauss, J.D. Watts and R.B. Bartlett, Chem.Phys.Lett. **187**, 21 (1991); R.B. Murphy and R.P. Messmer, Chem.Phys.Lett. **183**, 443 (1991)

[49] E.R. Davidson, in *The world of quantum chemistry*, ed. R. Daudel and B. Pullman (Reidel, Dordrecht 1974); S.R. Langhoff and E.R. Davidson, Int.J.Quantum Chem. **8**, 61 (1974)

[50] J.A. Pople, in: Energy, structure and reactivity. Eds. D.W. Smith, W.B. McRae (Wiley, N.Y. 1973), pp. 51–61

[51] S. Wilson, Correlation Energy in Molecules (Clarendon Press, Oxford 1982)

[52] W. Duch and G.H.F. Diercksen, J.Chem.Phys. **101**, 3018 (1994)

[53] S. Evangelisti, J-P. Daudey and J-P. Malrieu Chem.Phys. **75**, 91 (1983); F. Illas, J. Rubio and J.M. Ricart, J.Chem.Phys. **89**, 6376 (1988) R. Cimiraglia and M. Perscico, J.Comp.Chem. **8**, 39 (1987)

[54] R.J. Buenker, S.D. Peyerimhoff and W. Butscher, Molecular Physics **35**, 771 (1978); R.J. Harrison, J.Chem.Phys. **94**, 5021 (1991)

[55] W. Duch, Int.J.Quantum Chem. **S24**: 683 (1990)

[56] W. Duch, J.Chem.Phys. **91**, 2552 (1989)

[57] W. Duch and J. Meller, Int.J.Quantum Chem. **50**, 243 (1994)

[58] P. Saxe, H.F. Schaefer III and N.C. Handy, Chem.Phys.Lett. **79**, 202 (1981); C.W. Bauschlicher, Jr and P.R. Taylor, J.Chem.Phys. **85**, 2779 (1986)

[59] K. Andersson, P-Å. Malmqvist, B.O. Roos, A.J. Sadlej and K. Wolinski, J.Phys.Chem. **94**, 5483 (1990)

[60] J.J.W. McDouall, K. Peasly and M.A. Robb, Chem.Phys.Lett. **148**, 183 (1988)

[61] C. Bauschlicher, S.R. Langhoff, P. Taylor, N. Handy and P.J. Knowles, J.Chem.Phys. **85**, 1469 (1986)

[62] R.H. Nobes, J.A. Pople, L. Radom, N.C. Handy and P.J. Knowles, Chem.Phys.Lett. **138**, 481 (1987)

[63] J.P. Connerade and K. Dietz, Comm.at Mol.Phys. **19**, 283 (1987); F. Tarantelli and L.S. Cederbaum, J.Chem.Phys. **89**, 4170 (1988)

[64] F. Coester, Nucl.Phys. **1**, 421 (1958); F. Coester and H. Kümmel, Nucl.Phys. **17**, 477 (1960); J. Čižek, J.Chem.Phys. **45**, 4256 (1966); Adv.Chem.Phys. **14**, 35 (1969); R.J. Bartlett, J.Chem.Phys. **93**, 1697 (1989)

[65] J. Paldus, in *Realtivistic and Electron Correlation Effects in Molecules and Solids*, Ed. G.L. Malli, Plenum Press 1994, p. 234; I. Lindgren and D. Mukherjee, Phys.Rep. **151**, 93 (1987); D. Mukhopadhyay Jr. and D. Mukherjee, Chem.Phys.Lett. **163**, 171 (1989); L. Meissner and R.J. Bartlett, J.Chem.Phys. **92**, 561 (1990)

[66] P.G. Szalay and R.J. Bartlett, J.Chem.Phys. **103**, 3600 (1995); for particular contributions see [82]

[67] P. Jørgensen and T. Helgaker, J.Chem.Phys. **89**, 1560 (1988); T. Helgaker and P. Jørgensen, Theor.Chim.Acta **75**, 111 (1989)

[68] J.A. Pople, M. Head-Gordon and K. Raghavachari, J.Chem.Phys. **87**, 5968 (1987)

[69] J. Paldus J. Čižek and B. Jeziorski, J.Chem.Phys. **90**, 4356 (1989)

[70] D. Cremer and Z. He, Theor.Chim.Acta **88**, 47 (1994)

[71] S. Koch and W. Kutlenigg, Theor.Chim.Acta **59**, 387 (1981)

[72] R. Ahlrichs, H. Lischka, V. Staemmler and W. Kutzelnigg, J.Chem.Phys. **62**, 1235 (1975)

[73] P. Čársky and M. Urban, it Ab Inition Calculations, Springer–Verlag, Berlin 1980

[74] C. Bloch, Nucl.Phys. **6**, 329 (1958)

[75] R. Caballol and J.P. Malrieu, Chem.Phys.Lett. **188** 126 (1992)

[76] R. Alhrichs, P. Scharf and C. Eberhardt, J.Chem.Phys. **82**, 890 (1985)

[77] U. Kaldor, Theor.Chim.Acta **80**, 427 (1991); S.R. Hughes and U. Kaldor, Physical Review A **47**, 4705 (1993); see also [96]

[78] R.J. Bartlett, Theor.Chim.Acta **80**, 71 (1991)

[79] B. Jeziorski and J. Paldus, J.Chem.Phys. **90**, 2714 (1989); D. Mukherjee and S. Pal, Adv. Qauntum Chem. **20**, 291 (1990); U. Kaldor, Theor.Chim.Acta **80**, 427 (1991); S.R. Hughes and U. Kaldor, Physical Review A **47**, 4705 (1993)

[80] J.P. Malrieu, J.L. Heully and A. Zaitsevski, Theor.Chim.Acta **90**, 167 (1995); K. Wolinski and P. Pulay, J.Chem.Phys. **90**, 3647 (1989)

[81] K. Andersson, P.Å. Malmqvist and B.O. Roos, J.Chem.Phys. **96**, 1218 (1992)

[82] R.J. Gdanitz and R. Ahlrichs, Chem.Phys.Lett. **143**, 413 (1988); T. Sakai and K. Tanaka, Theor.Chim.Acta **85**, 451 (1993); R. Fink and V. Staemmler, Theor.Chim.Acta **87**, 129 (1993); R.J. Cave and E.R. Davidson, J.Chem.Phys. **89**, 6978 (1988); W.D. Laidig and R.J. Bartlett, Chem.Phys.Lett. **104**, 424 (1984); P.G. Szabo and R.J. Bartlett, Chem.Phys.Lett. **214**, 481 (1993); P.J.A. Ruttink, J.H. van Lenthe, R. Zwaans and G.C. Groenenboom, J.Chem.Phys. **94**, 7212 (1991)

[83] B. Jeziorski and H.J. Monkhorst, Physical Review A **24**, 1668 (1981); B. Jeziorski and J. Paldus, J.Chem.Phys. **88**, 5673 (1988); L. Meissner, S.A. Kucharski and R.J. Bartlett, J.Chem.Phys. **91**, 6187 (1989)

[84] A. Balkova, S.A. Kucharski, L. Meissner and R.J. Bartlett, Theor.Chim.Acta **80**, 335 (1991)

[85] J.P. Malrieu, J.P. Daudey and R. Caballol, J.Chem.Phys. **101**, 8908 (1994)

[86] J.P. Daudey, J.L. Heully and J.P. Malrieu, J.Chem.Phys. **99**, 1240 (1993)

[87] J.P. Malrieu, Ph. Durand and J.P. Daudey, J.Phys.A: Math.Gen. **18**, 809 (1985)

[88] I. Lindgren, Inter.J.Quantum Chem. S **12**, 33 (1978)

[89] J.L. Heully and J.P. Malrieu, Chem.Phys.Lett. **199**, 545 (1992)

[90] J. Sanchez-Marin, D. Maynau and J.P. Malrieu, Theor.Chim.Acta **87**, 107 (1993); I. Nebot-Gil, J. Sanchez-Marin, J.L. Heully, J.P. Malrieu and D. Maynau, Chem.Phys.Lett. **234**, 45 (1995), J.Chem.Phys. **103**, 2576 (1995)

[91] J. Meller, J.P. Malrieu and J.L. Heully, Chem.Phys.Lett. **244**, 440 (1995)

[92] N. Oliphant and L. Adamowicz, J.Chem.Phys. **94**, 1229 (1991); P. Piecuch, N. Oliphant and L. Adamowicz, J.Chem.Phys. **99**, 1875 (1993); P. Piecuch and L. Adamowicz, J.Chem.Phys. **100**, 5792 (1994); A. Balkova and R.J. Bartlett, Chem.Phys.Lett. **193**, 364 (1992); X. Li and J. Paldus, Inter.J.Quantum Chem. S **27**, 269 (1993)

[93] H.J. Silverstone and O. Sinanoglu, J.Chem.Phys. **44**, 1989, 3608 (1966)

[94] C.W. Bauschlicher and S.R. Langhoff, J.Chem.Phys. **89**, 4246 (1988); I. Garcia-Cuesta, L. Serrano-Andres, A.S. de Meras and I. Nebot-Gil, Chem.Phys.Lett. **199**, 535 (1992)

[95] K. Jankowski and J. Paldus, Inter.J.Quantum Chem. **18**, 1243 (1980); K. Jankowski, J. Paldus, I. Grabowski and K. Kowalski, J.Chem.Phys. **97**, 7600 (1992); K. Jankowski and K. Kowalski, Inter.J.Quantum Chem. **50**, 353 (1994)

[96] P. Sterna, Master Thesis, Nicholas Copernicus University, Torun 1994; K. Jankowski and P. Sterna, to be published

[97] J. Paldus, P. Piecuch, L. Pylypow and B. Jeziorski, Physical Review A **47**, 2739 (1993)

[98] T. Sakai and K. Tanaka, Theor.Chim.Acta **85**, 451 (1993); K. Tanaka, T. Sakai and H. Terashima, *ibid* **76**, 213 (1989)

[99] R.J. Gdanitz and R. Alhrichs, Chem.Phys.Lett. **143**, 413 (1988)

[100] C.W. Bauschlicher and P.R. Taylor, J.Chem.Phys. **86**, 2844 (1987)

[101] R. Fink and V. Staemmler, Theor.Chim.Acta **87**, 129 (1993)

[102] R.J. Cave and E.R. Davidson, J.Chem.Phys. **89**, 6978 (1988)

[103] Fulde and Stoll, J.Chem.Phys. **97**, 4248 (1992)

[104] I. Nebot-Gil, J. Sanchez-Marin, J.L. Heully, J.P. Malrieu and D. Maynau, Chem.Phys.Lett. **234**, 45 (1995)

[105] J. Meller, J.P. Malrieu and R. Caballol, J.Chem.Phys. in press (1996)

[106] W.D. Laidig and R.J. Barttlet, Chem.Phys.Lett. **104**, 424 (1984); P.G. Szabo and R.J. Barttlet, Chem.Phys.Lett. **214**, 481 (1993)

[107] P.J.A. Ruttink, J.H. van Lenthe, R. Zwaans and G.C. Groenenboom, J.Chem.Phys. **94**, 7212 (1991)

[108] P. Saxe, H.F. Schaefer III and N.C. Handy, Chem.Phys.Lett. **79**, 202 (1981); R.J. Harrison and N.C. Handy, Chem.Phys.Lett. **95**, 386 (1983)

[109] B. Jeziorski and H.J. Monkhorst, Physical Review A **24**, 1668 (1981)

[110] R.H. Nobes, J.A. Pople, L. Radom, N.C. Handy and P.J. Knowles, J.Chem.Phys. **85**, 1469 (1986)

[111] L. Adamowicz, R. Caballol, J.P. Malrieu and J. Meller, Chem.Phys.Lett., submitted for publication

[112] M.B. Lepetit and J.P. Malrieu, Chem.Phys.Lett. **208**, 503 (1993)

[113] M.B. Lepetit and J.P. Malrieu, J.Chem.Phys. **87**, 5937 (1987)

[114] J.P. Malrieu, I. Nebot-Gil, J. Sanchez-Marin, and D. Maynau, J.Chem.Phys. **100**, 1441 (1994)

[115] J.P. Malrieu, I. Nebot-Gil, J. Sanchez-Marin, and D. Maynau, J.Chem.Phys. submitted for publication

[116] J. Meller, J.L. Heully and J.P. Malrieu, Chem.Phys.Lett. **218**, 276 (1994)

[117] P.J. Knowles and N.C. Handy, J.Chem.Phys. **91**, 2396 (1989)

[118] S. Evangelisti, J.P. Daudey and J.P. Malrieu, Chem.Phys **75**, 91 (1983); R.J. Harrison, J.Chem.Phys. **94**, 5021 (1991)