



Pivot-invariant Knowles Partitioning at the Multireference Level

Student Research (TDK) Thesis

Author: András Gombás¹

¹second year Chemistry MSc student

Advisors: Dr. Ágnes Szabados¹, Dr. Péter R. Surján²

¹ full professor, Department of Inorganic Chemistry

² full professor, Department of Physical Chemistry

Finished: 02/01/2025

Summary

In 2022, Knowles introduced a new approach in single reference many-body perturbation theory [1]. The essence of the idea is to derive the zero-order Hamiltonian of a quantum many-body system from a model by requiring that the zero-order is as close to the exact Hamiltonian as possible. Knowles gave explicit formulae of the procedure that was termed *perturbation-adapted* perturbation theory (PAPT). Since then, adaptation of the theory to the laboratory’s program system to obtain results for a variety of quantum chemical test systems followed by further analysis of the applicability of the method [2] constituted my bachelor’s thesis work [3]. One of the main conclusions of that research was that the superiority of single reference PAAPT over e.g. the Møller-Plesset partitioning for weakly correlated systems encourages towards the extension of the theory to the multireference approach in order to improve the description of strongly correlated systems, too. In 2023, I took part in developing the first multireference adaptation in the general framework of projected multi-configurational perturbation theory (pMCPT) [4] as it required the least alteration of the original approach. Tests showed the enlargement of the inherent pivot-dependence of the pMCPT framework causing severe losses in the parallelity of results to the Full-CI benchmark [5]. At this point the pivot-invariant version of MCPT proposed in 2022 under the name Frame MCPT (fMCPT) [6] was found useful. This allows for removing the artefactual effects associated with pivot change. My work in the project entailed derivation, implementation and debugging of the pMCPT and fMCPT based multireference extended PAAPT methods presented and compared to other similar methods in this thesis. The fMCPT-based multireference adaptation of the Knowles Partitioning is observed to be able to account for a significantly greater part of the correlation than that of the Møller-Plesset partitioning providing a favourably balanced description of weakly and strongly correlated cases. During the development, a straightforward generalized derivation for the equations that define the Knowles Partitioning regardless of the chosen single- or multireference perturbative framework was also formulated. A follow-up report to our previous publication on pMCPT extension [5] is currently in preparation, focusing on the generalized version of the partitioning.

Acknowledgements

I am deeply grateful to my advisors, Ágnes Szabados and Péter Surján, for their unwavering support and guidance throughout my TDK work. Their expertise and patience have been invaluable to me and have played a crucial role in the success of my partaking in the research of the Laboratory of Theoretical Chemistry.

Contents

1	Introduction	6
1.1	Approximations in Quantum Chemistry	6
1.1.1	The non-relativistic Hamiltonian	7
1.1.2	The Born-Oppenheimer Hamiltonian	8
1.1.3	The independent electron approximation	9
1.1.4	The finite basis approximation	9
1.2	Hartree-Fock Theory	10
1.3	Post-Hartree-Fock methods	12
2	Single Reference Approach	14
2.1	Rayleigh-Schrödinger Perturbation Theory	14
2.2	Møller-Plesset Perturbation Theory	15
2.3	Perturbation-adapted Perturbation Theory	17
3	Multireference Approach	21
3.1	Static correlation	21
3.1.1	Geminal Theory	21
3.1.2	Complete Active Space	23
3.2	Dynamic correlation	24
3.2.1	CAS-based Perturbation Theory	25
3.2.1.1	CASPT2	25
3.2.1.2	NEVPT2	26
3.2.2	Multiconfigurational Perturbation Theory	27
3.2.2.1	Projected Multiconfigurational Perturbation Theory	29
3.2.2.2	Frame-based Multiconfigurational Perturbation Theory	33

4	Development of MR-PAPT	36
4.1	MR-PAPT in pMCPT	36
4.1.1	pMCPT-PAPT equations analogous to SR-PAPT	37
4.1.2	Additional pMCPT-PAPT equations	38
4.1.3	On the vanishing third-order energy correction	40
4.1.4	On the redundancy of projection functions	41
4.1.5	Workflow	44
4.2	MR-PAPT in fMCPT	45
4.2.1	Projection functions	47
4.2.2	PAPT Equations	47
4.2.2.1	Efficient coefficient matrix calculation	48
5	Results and Discussion	50
5.1	Potential energy curve of lithium-hydride	50
5.2	Geometry points of berillium-hydride	53
5.3	Size-consistency	56
6	Conclusion and Future Directions	59
	Appendix	65
A	The Davidson technique for repartitioning	65
B	An alternative derivation of the PAPT equations	66
C	Intermediars in the SR-PAPT working code	68
D	Reciprocal vectors in the pMCPT formulae	73
E	Parallelization of the pMCPT-PAPT working code	74

1. Introduction

This Chapter gives a brief introduction to the field of electronic structure theory in quantum chemistry from a bird’s eye view, believed to be practical for an unfamiliar reader. The aim is to set the foundation of some traditional approximations in quantum chemistry. Many of the limitations of simple methods presented in this Chapter are to be overcome by other, more complex methods that still exhibit feasible computational cost if one wishes to achieve quantitative description of a challenging many-body system to a certain precision. Also, the precise meaning behind many of the frequently used simplified quantum chemical terms and phrases is to be clarified.

A reader who is well-acquainted with these basic principles is encouraged to turn to Chapter 2 that focuses on the single reference perturbative framework and the methods of interest to this thesis to begin with the discussion of the Knowles Partitioning itself in Section 2.3.

1.1 Approximations in Quantum Chemistry

Quantum chemistry is a scientific field that applies quantum mechanics to describe chemical systems in order to theoretically account for and predict physical and chemical properties, processes and other phenomena of interest to physical chemistry. A foremost task in quantum chemical modelling is the description of the electronic structure, that is achieved by a systematic application of approximations to make the solution of the time-independent Schrödinger equation

$$\hat{H} |\Psi\rangle = E |\Psi\rangle \tag{1.1}$$

feasible. Thus providing the Ψ wavefunctions as the eigenfunctions and E energy levels as the eigenvalues of the $\hat{H}$ Hamiltonian of a quantum many-body system in a stationary state from which the properties in question can be calculated e.g. as expectation values with the corresponding quantum mechanical operator of the physical quantity. The solution of Eq. (1.1) first requires an explicit formula of the $\hat{H}$ Hamiltonian considered *exact*

within the given framework, allowing for neglects. However, even then with the most reasonable approximations in that regard, Eq. (1.1) can only be solved analytically for simple, special cases e.g. two-body systems. Therefore, in quantum chemistry one has no choice but to resort to some or all of the theoretical approximations and numerical methods discussed in this section.

Ab initio calculation methods for many-electron systems [7] aim to achieve the description to the desired accuracy without introducing empirical parameters, only using fundamental physical principles.

Besides *variational methods* that estimate the energy as an expectation value providing upper bound to the exact one, *many-body perturbation theory* is another ubiquitous tool to calculate approximate solutions to Eq. (1.1) [8].

1.1.1 The non-relativistic Hamiltonian

Let us consider the simplest, hydrogen-atom-like example of quantum physics, a stationary proton and an electron “orbiting” around at r distance. One would write the expression for the $\hat{H}$ Hamiltonian of that system accounting for the $\hat{T}_e$ kinetic and $\hat{V}_{p-e}$ potential energy to calculate wavefunctions and energy levels in SI units as

$$\hat{H} = \hat{T}_e + \hat{V}_{p-e} = -\frac{\hbar^2}{2m_e}\Delta + \frac{e^2}{4\pi\epsilon_0 r} \quad (1.2)$$

with $\hbar$ being the reduced Planck-constant, m_e the mass of an electron, e the elementary charge and ϵ_0 the permittivity of vacuum. Transcribing this formula to atomic units, usual in *non-relativistic* quantum chemistry as

$$\hat{H} = \hat{T}_e + \hat{V}_{p-e} = -\frac{1}{2}\Delta + \frac{1}{r} \quad (1.3)$$

though hides some of the physical concepts behind, equivalently describes the system. Note that this formulation neglects the consequences of Einstein’s theory of special relativity thus introducing the *non-relativistic approximation* generally maintained in this field. Relativistic effects are negligible to the chemical systems and precision of interest characteristic of this thesis.

1.1.2 The Born-Oppenheimer Hamiltonian

The general molecular many-electron $\hat{H}$ Hamiltonian

$$\hat{H}(\mathbf{r}, \mathbf{R}) = \underbrace{-\sum_i^n \frac{1}{2} \Delta_i}_{\hat{T}_e(\mathbf{r})} - \underbrace{\sum_I^N \frac{1}{2m_I} \Delta_I}_{\hat{T}_n(\mathbf{R})} + \underbrace{\sum_{i<j}^n \frac{1}{|r_i - r_j|}}_{\hat{V}_{e-e}(\mathbf{r})} + \underbrace{\sum_{I<J}^N \frac{Z_I Z_J}{|R_I - R_J|}}_{\hat{V}_{n-n}(\mathbf{R})} - \underbrace{\sum_i^n \sum_J^N \frac{Z_J}{|r_i - R_J|}}_{\hat{V}_{e-n}(\mathbf{r}, \mathbf{R})}, \quad (1.4)$$

where Z_I is the charge of nucleus I , depends inseparably on the $\mathbf{r}$ coordinates of the n electrons and $\mathbf{R}$, that of the N nuclei, because of the last $\hat{V}_{e-n}(\mathbf{r}, \mathbf{R})$ term accounting for the potential energy induced by the attraction between electrons and the nuclei.

Similarly, the many-electron wavefunction of the molecule emerging as the solution of Eq. (1.1) written with this Hamiltonian also depends inseparably on these coordinates. However, corresponding to chemical intuition, the *Born-Oppenheimer approximation* separates the wavefunction of electrons and nuclei in a product form [9]

$$\Psi(\mathbf{r}, \mathbf{R}) = \Phi(\mathbf{r}; \mathbf{R}) \cdot \Xi(\mathbf{R}). \quad (1.5)$$

This *adiabatic approximation* assumes that electronic motion adjusts instantaneously to changes in nuclear positions. The reasoning behind this technique is justified by the difference of mass being of four orders of magnitude between the electrons and the nuclei. This allows for setting up a separate equation for $\Phi(\mathbf{r}; \mathbf{R})$ and $\Xi(\mathbf{R})$.

The $\Phi(\mathbf{r}; \mathbf{R})$ electronic wavefunction depends only parametrically on the position of the nuclei collected in $\mathbf{R}$. The $\hat{H}_e$ electronic Hamiltonian in the electronic Schrödinger equation

$$\hat{H}_e(\mathbf{r}; \mathbf{R}) |\Phi(\mathbf{r}; \mathbf{R})\rangle = E(\mathbf{R}) |\Phi(\mathbf{r}; \mathbf{R})\rangle, \quad (1.6)$$

is built of the terms indicated in Eq. (1.4), except for $\hat{T}_n$. Potential energy curves and surfaces can be constructed only in this approximation, by graphing the resulting solutions of Eq. (1.6), i.e. the $E(\mathbf{R})$ electronic energy at certain $\mathbf{R}$ parameters.

On a side note, the well-known limitation of the Born-Oppenheimer approximation [9] is the description of molecular systems exhibiting *conical intersections* [10] or *avoided crossings* [11] in their potential energy surfaces.

Nevertheless, for the rest of this thesis, in which this approximation will be continuously upheld, let us, using quantum chemical jargon, refer to $\Phi(\mathbf{r}; \mathbf{R})$ simply by default as the Ψ *exact many-electron wavefunction*, to $\hat{H}_e$ as the $\hat{H}$ *exact molecular Hamiltonian* and to its eigenvalue problem of Eq. (1.6) as the *Schrödinger equation* with same symbols as in Eq. (1.1), but with different, more restricted quantum chemical meaning.

1.1.3 The independent electron approximation

Wishful chemical thinking when describing the behavior of a single electron in a many-electron system would consider it independently and assume that it is only affected by the average field exerted by the nuclei and other electrons in the system. This could lead to a product of ϕ_i one-electron wavefunctions as an Ansatz for the Ψ many-electron molecular wavefunction (Hartree theory) [12]. However, the *Pauli principle* states that the multi-fermionic wavefunction should be antisymmetrical, i.e. change sign upon exchanging two of its fermions. Consequently, inherently antisymmetrical

$$\Psi(1, 2, \dots, n) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(1) & \phi_2(1) & \phi_3(1) & \cdots & \phi_n(1) \\ \phi_1(2) & \phi_2(2) & \phi_3(2) & \cdots & \phi_n(2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_1(n) & \phi_2(n) & \phi_3(n) & \cdots & \phi_n(n) \end{vmatrix} \quad (1.7)$$

Slater determinants [13] are used instead of the simple product in the *independent electron approximation*. These ϕ_i one-electron wavefunctions are called molecular *spin orbitals* as they depend not only on the spatial coordinates of that one electron but also on its spin coordinate. A many-electron wavefunction of an n -electron molecule written this way is merely an antisymmetrized product form of its n occupied spin orbitals.

Again, for the rest of this thesis, let us refer to such a wavefunction simply as a *determinant*.

1.1.4 The finite basis approximation

The ϕ_i molecular orbitals are constructed as

$$\phi_i = \sum_{\mu=1}^K c_{i\mu} \chi_{\mu} \quad (1.8)$$

linear combinations of a finite K number of χ_{μ} basis functions to make calculations feasible. This inevitably introduces the *finite basis approximation*. However, by systematically expanding the basis set and extrapolating for the so-called complete basis set (CBS) limit, this restriction can be relieved.

It is customary in quantum chemistry to use atomic orbitals centered at each nucleus within the molecule as basis functions. From a theoretical point of view, Slater-type orbitals (STOs) would be advantageous [14] as they both exhibit exponential decay at large distances from the nucleus and fulfill Kato's cusp conditions near the centre [15]. However, integral calculations with STOs are computationally not cost effective. Consequently, upon Boys' suggestion [16], most often Gaussian expansions of Slater-type atomic

orbitals are used. The reason for that is that Pople [17] gave closed form for the integrals by expressing products of Gaussian-type orbitals (GTOs) as linear combination of GTOs thereby significantly reducing computational costs.

Minimal basis sets [7, 18] consider only one basis function for fully filled shells and the entire valence shell. That involves 1 function for s , 3 functions for p shells etc. Each atomic orbital in itself is constructed as a linear combination of primitive Gaussians in most cases. *Split valence basis sets* add flexibility that enables the orbitals to adjust to the molecular environment by representing each valence shell by more than one basis functions. Depending on the number of basis functions per shell, one might use so-called double-, triple-, quadruple-zeta etc. basis sets. That means 2 functions for s , 6 functions for p shells, etc. for double-zeta basis sets. The naming originates from the exponent of STOs, that are usually denoted with ζ . *Polarization functions* alter the spatial distribution of orbitals allowing for the description of chemical bonds that are often indeed polarized. This is achieved by adding a set of l -functions, where l is the smallest azimuthal quantum number not occurring in the valence shell of the atom. Gaussian basis functions with small exponents used as *diffuse functions* to represent the *tail* part of the orbitals can be important to describe anions or dipole moments.

1.2 Hartree-Fock Theory

The simplest, most elemental method that follows from the approximations introduced so far, the *Hartree-Fock Theory* [19, 20] forms the basis of the description of molecular electronic structure theory. Here we consider the *restricted* Hartree-Fock (RHF) method that is formulated for closed shell systems. This restriction allows for the introduction of φ_i molecular *spatial orbitals*, $n/2$ in number, that can accomodate two electrons with opposite spin, so the molecular many-electron wavefunction can be simplified to

$$\Psi(1, 2, \dots, n) = \hat{\mathcal{A}} \left[\underbrace{\varphi_1(1)\alpha(1)}_{\phi_1(1)} \underbrace{\varphi_1(2)\beta(2)}_{\phi_2(2)} \dots \underbrace{\varphi_{n/2}(n-1)\alpha(n-1)}_{\phi_{n-1}(n-1)} \underbrace{\varphi_{n/2}(n)\beta(n)}_{\phi_n(n)} \right], \quad (1.9)$$

where $\hat{\mathcal{A}}$ is the operator ensuring the antisymmetry of the wavefunction (cf. Eq. (1.7)). From this formulation it follows that the RHF determinant always corresponds to a singlet state.

Hartree-Fock theory also decouples the Schrödinger equation of an n -electron system (Eq. (1.6)) to a

$$\hat{F}(i) |\varphi_i(i)\rangle = \varepsilon_i |\varphi_i(i)\rangle \quad \text{for } i \in 1, 2, \dots, n/2 \quad (1.10)$$

system of $n/2$ equations in the restricted, i.e. closed shell case. In Eq. (1.10), the $\hat{F}(i)$ Fockian is an effective one-electron Hamiltonian operator that only contains the average effective field that the electrons are acted upon by, i.e.

$$\hat{F}(i) = \hat{h}(i) + \sum_{j=1}^{n/2} (2\hat{J}_j(i) - \hat{K}_j(i)) \quad (1.11)$$

where $\hat{h}$ contains the one-electron part of the total Hamiltonian, i.e. the kinetic energy of electron i (cf. Eq. (1.4)), $\hat{J}_j(i)$ is the one-electron Coulomb operator accounting for the repulsion of electron i and the charge density corresponding to an electron on orbital j and $\hat{K}$ is the so-called exchange operator stemming from the antisymmetric nature of the determinant. Matrix elements of the Fock operator on the orbitals used as basis, denoted with Greek letters, in accordance with Eq. (1.11) are

$$F_{\mu\nu} = h_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma} \left([\mu\lambda|\nu\sigma] - \frac{1}{2}[\mu\lambda|\sigma\nu] \right), \quad (1.12)$$

where $P_{\lambda\sigma}$ is the

$$P_{\lambda\sigma} = 2 \sum_i^{\text{occ}} c_{i\lambda} c_{i\sigma}^* \quad (1.13)$$

density matrix (cf. Eq. (1.8)) and occ refers to the $n/2$ occupied molecular orbitals. In this formulation, one-electron integrals are represented as $h_{\mu\nu}$, whereas for the two-electron integrals, the [12|12] Dirac convention is used.

The Hartree-Fock equations arise by applying the variational principle for the expectation value of $\hat{H}$ taken with the determinant of Eq. (1.9). The variations of the wavefunction are obtained by varying expansion parameters $c_{i\mu}$ corresponding to the μ th atomic orbital and the i th spatial molecular orbital. In the restricted Hartree-Fock framework, the number of variational parameters is half of that introduced in Eq. (1.8) that was given for spin orbitals.

The solution of Eq. (1.10) can be achieved iteratively. First, a trial wavefunction is constructed by creating a set of φ_i molecular orbitals with some initial coefficients applied in Eq. (1.8). These coefficients can be used to construct the density matrix of Eq. (1.13). Then the matrix representation of the Fockian can be constructed according to Eq. (1.12) using the one- and two-electron integrals. Consequently, the solution of the eigenvalueproblem of the Fockian yields a new set of coefficients to construct the next trial wavefunction and the procedure is repeated until self-consistency is reached.

Orbitals φ_i that diagonalize the Fockian are called canonical molecular orbitals. Orbital energies ε_i appear e.g. in the Koopmans theorem as approximate ionization potentials and electron affinities, and also in the denominators of many-body perturbative

formulae. The total electronic energy is not the sum of these orbital energies but as mentioned can be calculated as an expectation value of the $\hat{H}$ Hamiltonian with the Ψ_{HF} one-determinant molecular wavefunction instead

$$E_{\text{HF}} = \langle \Psi_{\text{HF}} | \hat{H} | \Psi_{\text{HF}} \rangle = \sum_i^{\text{occ}} (h_{ii} + F_{ii}). \quad (1.14)$$

given in spatial orbitals, too.

Brillouin's theorem [21] points out an important feature utilized later as it states that matrix elements are

$$\langle \Psi_0^{\text{HF}} | \hat{H} | K_a^p \rangle = 0, \quad (1.15)$$

where Ψ_0^{HF} is the ground state self-consistent Hartree-Fock wavefunction and K_a^p is any single excitation from it. Simple algebraic proof for that using the Slater-Condon rules yields that such matrix element corresponds to an offdiagonal matrix element of the Fockian that was already diagonalized when the wavefunctions were obtained.

1.3 Post-Hartree-Fock methods

The Hartree-Fock mean-field approach has a foundational role in electronic structure theory, neglecting the correlated behaviour of electrons. *Post-Hartree-Fock* methods, on the other hand, aim to account for the so-called *correlation energy*, the difference of the exact and the Hartree-Fock energy.

The Ψ_{HF} Hartree-Fock determinant classifies the molecular spin orbitals [7] as either occupied or unoccupied, i.e. virtual. Excited determinants, Ψ_K can thusly be formed by placing one or more electrons from occupied orbitals to virtual ones. The exact molecular wavefunction can be expanded on the basis of these n_{CI} determinants

$$|\Phi_{\text{CI}}\rangle = c_{\text{HF}}|\Psi_{\text{HF}}\rangle + \sum_{K=1}^{n_{\text{CI}}} c_K |\Psi_K\rangle, \quad (1.16)$$

the configurational interactions' (CI) space. Note that the variational obtaining of the c coefficient for all possible determinants of the CI-space, the so-called *Full-CI* (FCI) solution indeed yields the exact expansion of the wavefunction in the given basis. Therefore, from a practical point of view, correlation energy can be considered the difference of E_{HF} (see Eq. (1.14)) and E_{FCI} , i.e.

$$E_{\text{FCI}} = \langle \Phi_{\text{FCI}} | \hat{H} | \Phi_{\text{FCI}} \rangle. \quad (1.17)$$

We speak of *weakly correlated* cases where in this expansion, the c_K weight of non-Hartree-Fock determinants is significantly smaller than c_{HF} , that of the Hartree-Fock determinant. In such cases the so-called *dynamic correlation* can be calculated e.g. by single reference perturbative methods (see Ch. 2) using the Hartree-Fock determinant as reference.

However, when there are more than one dominant determinants in the expansion of Eq. (1.16), then merely to achieve qualitative description, the so-called *static correlation* should also be accounted for. This necessitates the inclusion of all determinants of large and comparable weight in the reference function and leads to the multireference approach (see Ch. 3).

It is noted here that this should only be considered a *medium level* of correlation as Scuseria emphasizes [22] that *strong correlation*, a term frequently used in quantum chemical jargon, should refer to those special cases where most or all determinants in the expansion have comparable weight (based on the Kapuy lecture held by Gustavo Scuseria in 2023 at Eötvös Loránd University, Budapest). In this thesis, the nomenclature of Scuseria is not completely adapted. For simplicity, the scenario of multiple dominant determinants in the CI expansion is referred to as strong correlation, i.e. distinguishing cases that call for the consideration of static correlation via application of multireference methods for sufficient accuracy when describing electronic structure from cases where single reference methods are adequate.

Suitable reference functions in the presence of static correlation may be unrestricted Hartree-Fock, as well as multireference, geminal-based and complete active space functions. Systems for which the single-determinantal restricted Hartree-Fock solution is insufficient encompass ionized, excited, near-degenerate and far from equilibrium geometry cases. Correspondingly, to describe many of the systems and processes interesting to chemistry e.g. via potential energy surfaces and hypersurfaces, the method used is required to treat both dynamic and static correlation in balance.

2. Single Reference Approach

As mentioned in Section 1.3 of the previous introductory chapter, for weakly correlated systems, the post-Hartree-Fock consideration of dynamic correlation can be sufficient to achieve the desired quantitative precision. In this Chapter, therefore, the general concepts of single reference perturbative methods are discussed as the original formulation of the Knowles Partitioning refers to this framework.

Here again, the reader familiar with the basics of Rayleigh-Schrödinger Perturbation Theory and the Møller-Plesset partitioning is advised to turn to Section 2.3 to obtain a general understanding of Perturbation-adapted Perturbation Theory (PAPT), aka. the Knowles Partitioning.

Additionally to the detailed presentation, those features of PAPT are also stressed in this Chapter that will play an important role when constructing the working formulae of the multireference extensions (see Chapter 4) and the generalized derivation applicable in many multireference perturbative framework is also pointed out (see Appendix B).

2.1 Rayleigh-Schrödinger Perturbation Theory

Perturbation theory is a tool for generating approximate solution to differential equations [23] by expanding the solution in terms of power series. In this section, non-degenerate time-independent *Rayleigh-Schrödinger perturbation theory* is discussed. The efficiency of these methods arise from a well-chosen partitioning of the total Hamiltonian $\hat{H}$ of the system by identifying a zero-order Hamiltonian of most frequently a model problem for which the eigenvalueproblem is feasible and thus, its complete spectrum can be calculated. All other effects are obtained by introducing a scaling parameter λ in front of the remainder, called perturbation $\hat{W}$, i.e.

$$\hat{H} = \hat{H}^{(0)} + \lambda \hat{W}. \quad (2.1)$$

The emerging Schrödinger equation

$$[\hat{H}^{(0)} + \lambda \hat{W}]|\Psi\rangle = E |\Psi\rangle, \quad (2.2)$$

then can be solved by assuming a series expansion for E and Ψ in increasing powers of λ . The exact k th wavefunction is written as

$$|\Psi_k\rangle = |\Psi_k^{(0)}\rangle + \sum_{i=1}^{\infty} \lambda^i |\Psi_k^{(i)}\rangle. \quad (2.3)$$

The i th order correction, $\Psi_k^{(i)}$ is conveniently expanded in terms of the zero-order solution $\Psi_j^{(0)}$, leading to

$$|\Psi_k\rangle = |\Psi_k^{(0)}\rangle + \sum_{i=1}^{\infty} \sum_{j \neq k} \lambda^i c_{kj}^{(i)} |\Psi_j^{(0)}\rangle. \quad (2.4)$$

The order by order expansion of Eq. (2.1) yields recursive formulae for the $c_{kj}^{(i)}$ coefficient accounting for the contribution of the j th zero-order wavefunction to the k th i -order wavefunction. The exact energy of such a Ψ_k wavefunction can be evaluated as

$$E_k = E_k^{(0)} + \sum_{i=1}^{\infty} \lambda^i E_k^{(i)} = E_k^{(0)} + \sum_{i=1}^{\infty} \lambda^i \langle \Psi_k^{(0)} | \hat{W} | \Psi_k^{(i-1)} \rangle. \quad (2.5)$$

Let us now comment on calculating Ψ_k and E_k according to Eqs. (2.4) and (2.5). In practice, the approximation is constituted by stopping at a finite order. When the expressions are generated to all orders, it is not necessary, that the series indeed provides the exact quantity. A convergence study is necessary to decide on the latter question.

Overall, such perturbative methods are useful when sufficiently accurate results are obtained by low-order truncation of the perturbation series.

2.2 Møller-Plesset Perturbation Theory

When applying Rayleigh-Schrödinger Perturbation Theory for the many-electron problem, the *Møller-Plesset partitioning* [24] is used most frequently. It entails choosing the Fockian as zero-order

$$\hat{H}^{(0)} = \hat{F}, \quad (2.6)$$

and treating the rest, i.e. electron correlation as perturbation.

The sum of occupied orbital energies constitutes the zero-order ground state energy value

$$E_0^{(0)} = \sum_i^{\text{occ}} \varepsilon_i \quad (2.7)$$

associated with the $\Psi_0^{(0)}$ eigenfunction, which is the single Hartree-Fock determinant. The Hartree-Fock energy is, therefore, obtained as the total energy calculated to the first-order, i.e.

$$\begin{aligned} E_0^{(0)} + E_0^{(1)} &= \langle \Psi_0^{(0)} | \hat{H}^{(0)} | \Psi_0^{(0)} \rangle + \langle \Psi_0^{(0)} | \hat{W} | \Psi_0^{(0)} \rangle \\ &= \langle \Psi_0^{(0)} | \hat{H} | \Psi_0^{(0)} \rangle = \langle \hat{H} \rangle = E_{\text{HF}}. \end{aligned} \quad (2.8)$$

Assuming canonical molecular orbitals that correspond to a diagonal Fockian, the formula of the ground state first-order correction wavefunction can be expressed with all possible double excitations from the Hartree-Fock determinant

$$\begin{aligned} |\Psi_0^{(1)}\rangle &= -\frac{1}{2} \sum_{a,b}^{\text{occ}} \sum_{p,q}^{\text{virt}} \sum_{\sigma,\sigma'} \frac{[ab|pq]}{\varepsilon_p + \varepsilon_q - \varepsilon_a - \varepsilon_b} p_{\sigma}^{+} q_{\sigma'}^{+} b_{\sigma'}^{-} a_{\sigma}^{-} |\Psi_0^{(0)}\rangle \\ &= \sum_{a,b}^{\text{occ}} \sum_{p,q}^{\text{virt}} \sum_{\sigma,\sigma'} c_{pq}^{ab} p_{\sigma}^{+} q_{\sigma'}^{+} b_{\sigma'}^{-} a_{\sigma}^{-} |\Psi_0^{(0)}\rangle, \end{aligned} \quad (2.9)$$

where the c_{pq}^{ab} coefficient is introduced as

$$c_{pq}^{ab} = -\frac{1}{2} \frac{[ab|pq]}{\varepsilon_p + \varepsilon_q - \varepsilon_a - \varepsilon_b}, \quad (2.10)$$

and the ε_i canonical orbital energies appear in the denominator. The equations are written using second quantized formalism, with latin letters denoting the spatial molecular orbitals. The index convention

$$a, b, c, \dots, h \in \{\text{occupied}\}$$

$$i, j, k, \dots, o \in \{\text{generic}\}$$

$$p, q, r, \dots, z \in \{\text{virtual}\}$$

is also introduced for the current discussion, later extended when necessary.

The second-order energy correction in this partitioning also resorts to double excitations reading

$$E_0^{(2)} = - \sum_{a,b}^{\text{occ}} \sum_{p,q}^{\text{virt}} \frac{[ab|pq] \left(2[ab|pq] - [ab|qp] \right)}{\varepsilon_p + \varepsilon_q - \varepsilon_a - \varepsilon_b}. \quad (2.11)$$

This partitioning, although being robust, underestimates the correlation energy, it is prone to slow convergence or even divergence [25] and fails with excitations involving near-degenerate (eg. HOMO-LUMO) orbitals (cf. orbital energies in the denominator of Eqs. (2.9) and (2.11)), that are well-known phenomena.

2.3 Perturbation-adapted Perturbation Theory

One might argue that in many-body perturbation theory, the model problem is not necessarily introduced to aid physical interpretation. Following this reasoning, it can be argued that the only purpose of the model used is to have a one-body zero-order operator with the ground state zero-order wavefunction being a determinant as close to the target problem as possible and the excited eigenstates being simple determinants, too.

Introducing more parameters at zero-order by using a $\hat{\bar{F}}$ one-electron operator, whose representation is not diagonal but block-diagonal on the canonical basis could possibly drive the zero-order problem closer to the total problem in question. What is of greater difficulty, is to give a general, straightforwardly applicable formula to obtain these parameters.

Facing this challenge Knowles has suggested in 2022 [1] to consider projection functions as

$$|\Theta_{ab}\rangle = \sum_c^{\text{occ}} \sum_{p,q}^{\text{virt}} c_{ac}^{pq} p^+ q^+ c^- b^- |0\rangle \quad (2.12)$$

$$|\Theta_{pq}\rangle = \sum_{a,b}^{\text{occ}} \sum_r^{\text{virt}} c_{ab}^{pr} q^+ r^+ b^- a^- |0\rangle \quad (2.13)$$

where $|0\rangle$ denotes the Hartree-Fock determinant and c the first-order correction coefficient of e.g. the Møller-Plesset partitioning (see Eq. (2.10) written with spatial orbitals). In the above and for the rest of the thesis, orbital indices refer to spin orbitals when not declared otherwise.

Functions Θ_{aa} and Θ_{pp} emerge as partial sums of the terms of first-order correction wavefunction (cf. Eq. (2.9)). The number of Θ_{ab} and Θ_{pq} functions altogether is the same as the number of parameters Λ_{ab} and Λ_{pq} in

$$\hat{\Lambda} = \sum_{a,b}^{\text{occ}} \Lambda_{ab} b^+ a^- + \sum_{p,q}^{\text{virt}} \Lambda_{pq} q^+ p^-. \quad (2.14)$$

Knowles introduced operator $\hat{\Lambda}$ in order to serve as zero-order. Note that the matrix representation of $\hat{\Lambda}$ is blockdiagonal in the occ-occ and virt-virt sense on the basis of canonical molecular orbitals. Therefore, parameters Λ_{ab} and Λ_{pq} correspond to level shifts and orbital rotations, i.e. they can be interpreted as the matrix elements of the $\hat{\bar{F}}$ transformed Fockian using the Davidson technique [26] (see Appendix A).

Using the projection functions of Eqs. (2.12) and (2.13), the

$$\hat{\Lambda} |\Psi^{(1)}\rangle \approx \hat{H} |\Psi^{(1)}\rangle, \quad (2.15)$$

operation equation, where $|\Psi^{(1)}\rangle$ is e.g. the first-order Møller-Plesset correction wavefunction, can be projected to the first-order interacting space of double excitations to obtain the equation

$$\langle \Theta_{kl}^+ | \hat{\Lambda} - \hat{H} | \Psi^{(1)} \rangle = 0, \quad (2.16)$$

that defines $\hat{\Lambda}$ so that it can be used as a zero-order Hamiltonian in Rayleigh-Schrödinger Perturbation Theory. To preserve hermiticity of the resulting operator, the

$$|\Theta_{kl}^+\rangle = |\Theta_{kl}\rangle + |\Theta_{lk}\rangle, \quad \text{for } k \geq l \quad (2.17)$$

symmetrization is introduced.

Knowles named this method *perturbation-adapted* perturbation theory, because in a sense the zero-order Hamiltonian is not directly taken from a model but constructed with the help of the perturbed function to be as close to the total Hamiltonian as possible (see Eq. (2.16)) within the first-order interaction space.

It is important to note here that since partial sums of the first-order correction wavefunction can be found among the projection functions, they are linearly dependent. This results in the $\mathbf{T}^\Lambda$ coefficient matrix of the

$$\sum_{i \geq j} \underbrace{\langle \Theta_{kl}^+ | [j^+ i^- + i^+ j^-] | \Psi^{(1)} \rangle}_{\mathbf{T}_{kl,ij}^\Lambda} \Lambda_{ij} = \underbrace{\langle \Theta_{kl}^+ | \hat{H} | \Psi^{(1)} \rangle}_{\mathbf{Y}_{kl}^\Lambda} \quad (2.18)$$

PAPT equations being singular.

It was shown [3] that using normal ordered operators, thus constructing

$$\hat{\Lambda}_N = \hat{\Lambda} - \langle 0 | \hat{\Lambda} | 0 \rangle \quad (2.19)$$

with coefficient matrix

$$\mathbf{T}_{kl,ij}^{\Lambda_N} = \langle \Theta_{kl}^+ | \left[[j^+ i^- + i^+ j^-] - \langle 0 | [j^+ i^- + i^+ j^-] | 0 \rangle \right] | \Psi^{(1)} \rangle, \quad (2.20)$$

and inhomogeneous vector

$$\mathbf{Y}_{kl}^{\Lambda_N} = \langle \Theta_{kl}^+ | [\hat{H} - \langle 0 | \hat{H} | 0 \rangle] | \Psi^{(1)} \rangle, \quad (2.21)$$

there is only one zero singular value of $\mathbf{T}^{\Lambda_N}$ which always corresponds to that linear combination of the Λ_{ij} parameters being undefined that is anyway immaterial from the point of view of perturbation, i.e. a constant shift.

It can be argued that the formulae of the Knowles Partitioning seem heuristic. In contrast to that, an alternative, functional minimalization based derivation of the PAPT

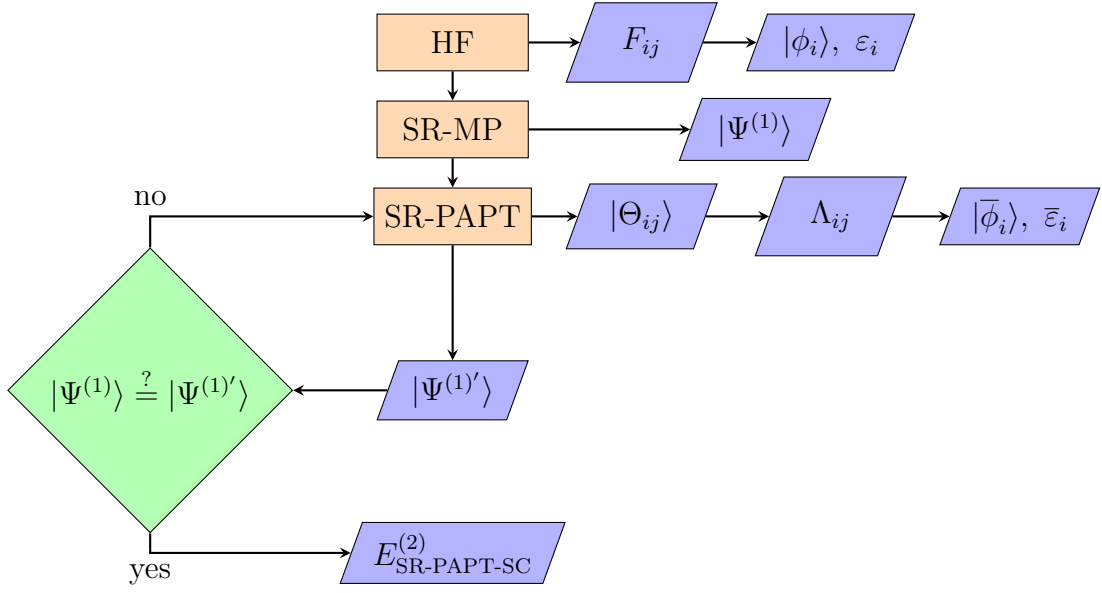


Figure 2.1: The workflow of the SR-PAPT procedure

equations of Eq. (2.18) is a novel, yet unpublished result of us. Its significance is that it makes no recourse to the notion of occupied and virtual orbitals. See Appendix B for more details.

To aid the understanding of the nature of a PAPT calculation in single reference approach, Fig. 2.1 shows an overview of the workflow of the PAPT procedure that has been implemented to the program-system of the laboratory. It starts with solving the Hartree-Fock equations, that way one obtains the Fockian and the ϕ_i canonical orbitals with ε_i canonical orbital energies. From the Fockian the first order Møller-Plesset wavefunction and from that the first set of $|\Theta_{ij}\rangle$ projection functions can be constructed so that the PAPT equations can be solved yielding the Λ_{ij} parameters. The new set of the $\bar{\phi}_i$ proper canonical orbitals (see Appendix A) then emerge as its eigenbasis with the level-shifted $\bar{\varepsilon}_i$ orbital energies as the corresponding eigenvalues. Using $\hat{\Lambda}$, a new first order correction wavefunction can be constructed in the usual manner and the regular many-body formulae can be used to obtain the second and third order energies etc. Our code was written in a way that also allows one to separate the two components of the repartitioning, the level-shifts and the orbital rotations and it was concluded that their effect is synergic [3].

Knowles also suggested that the PAPT procedure might be used iteratively. As Fig. 2.1 indicates, not only the original first order Møller-Plesset wavefunction, but also the newly obtained PAPT correction wavefunction can be used to construct new parameters

and thus a new zero-order until self-consistency is achieved.

From a theoretical point of view, the power of PAPT lies behind the fact that upon solving the PAPT equations written with normal ordered operators, as pointed out by us [3], the third order energy correction vanishes and the first non-zero correction to the second order energy appears at the fourth order. This feature has a resemblance to Feenberg scaling [27], with the difference that self-consistent PAPT in the single reference approach is size-consistent whereas the former is not.

Assessing the performance of SR-PAPT, we found [3] that the method is able to provide results for total energies at third order that are an order of magnitude closer to the Full-CI benchmark than traditional Møller-Plesset results. In fact, PAPT2 second-order results of PAPT are comparable to CCSD in most cases.

With appropriate computer implementation e.g. via the recognition of suitable intermediates (see Appendix C), computational cost of the Knowles partitioning in single reference framework, SR-PAPT scales with the sixth power of the number of basis functions used. The cost and accuracy enables PAPT to stand out from the family of optimized partitionings that have previously been investigated [26].

3. Multireference Approach

Contrary to weakly correlated cases, for systems entailing medium-level or strong correlation, single reference methods, such as presented in the previous Chapter, are insufficient as they cannot account for the static correlation as discussed in Sec. 1.3. In this Chapter, some examples of variational as well as perturbative multireference methods are discussed to take the static and dynamic correlation into consideration. This cannot usually be achieved in one step, therefore the application of a method describing static correlation (see Section 3.1) providing a suitable reference function can be followed by a perturbative method describing dynamic correlation (see Section 3.2) based on that.

The newly developed multireference adaptations (see Chapter 4) of the Knowles Partitioning constitute the latter case. From the family of many multireference perturbative frameworks, Multiconfigurational Perturbation Theory (MCPT, see Section 3.2.2) is the most relevant to this thesis as it provides the general background for the formulation of MR-PAPT in the current work. The other methods like CASPT and NEVPT are presented merely to illustrate the context of other well-known methods of the same league in the literature without any claim to completeness.

3.1 Static correlation

Multiconfigurational Perturbation Theory offers a framework to calculate perturbative corrections to an arbitrary reference function with known expansion on the basis of determinants. Geminal as well as CAS-based methods of calculating static correlation to provide a reference function for such further perturbative calculations are detailed in the following as they were used in the test calculations presented in Chapter 5.

3.1.1 Geminal Theory

Two-electron functions, i.e. so-called *geminals* have been introduced in quantum chemistry [28, 29] to tackle intrapair electron correlation. Starting from the determinant

of Eq. (1.9), it is possible to deduce $n/2$ Φ_i two-electron functions that are antisymmetric. They can be considered as *separated pairs* [30] constituting the n -electron system function [31, 32]

$$\Psi(1, 2, \dots, n) = \hat{\mathcal{A}}[\Phi_1(1, 2) \Phi_2(3, 4) \dots \Phi_{n/2}(n-1, n)]. \quad (3.1)$$

The expansion of geminals in a one-electron basis can be constructed as

$$|\Phi_i(1, 2)\rangle = \sum_{m,n} c_{mn}^i \hat{\mathcal{A}}[\chi_m(1)\chi_n(2)]. \quad (3.2)$$

In the restricted Hartree-Fock case, the extension of Eq. (3.2) extends over $\varphi_{i\alpha}$ and $\varphi_{i\beta}$. Intrapair correlation can be described by including more orbitals in the expansion. The geminal of Eq. (3.2) can be represented in second-quantized formalism in a concise manner as

$$|\Phi_i\rangle = \sum_{m,n} c_{mn}^i m^+ n^+ |\text{vac}\rangle = \Phi_i^+ |\text{vac}\rangle. \quad (3.3)$$

It is important to point out here that this implies that the coefficient c_{mn}^i describing the contribution of the m th and n th basis functions to the i th geminal is given in a matrix form.

The geminal-based many-electron wavefunction with the above introduced Φ_i^+ composite creation operators, can be simply given as

$$|\Psi\rangle = \Phi_{n/2}^+ \dots \Phi_2^+ \Phi_1^+ |\text{vac}\rangle \quad (3.4)$$

due to the inherent antisymmetry of the second quantized formalism.

Optimization of the Φ_i two-electron functions with respect to the energy, i.e. obtaining the c_{mn}^i coefficients variationally is a much more elaborate task than optimizing one-electron orbitals in the Hartree-Fock method. At the same time, one has the benefit that geminal-based calculations are size-consistent, provide an upper bound to the energy of the many-electron system and even provide the exact description of two-electron systems.

Orthogonality of geminals reads

$$\langle \Phi_i | \Phi_j \rangle = \int dx_1 \int dx_2 \Phi_i(x_1, x_2) \Phi_j(x_1, x_2) = 0. \quad (3.5)$$

Geminals can be orthogonal not only if their spatial parts do not interfere, but also when they have non-zero contributions at the same region of the space, their overlap integral can still be zero due to the nodal structure.

The so-called *strong orthogonality* occurs when

$$\int dx_1 \Phi_i(x_1, x_2) \Phi_j(x_1, x_2) = 0 \quad (3.6)$$

is fulfilled.

Arai showed [33] that the set of geminals that obey strong orthogonality can be expanded in disjunct subspaces of one-electron functions, modifying the expression in Eq. (3.3) to

$$\Phi_i^+ = \sum_{m,n \in s_i} c_{mn}^i m^+ n^+ \quad (3.7)$$

where s_i is the subspace of the one-electron basis functions assigned to the i th geminal.

To prove Arai's theorem, one needs to show implications in two directions. One is rather involved but the other is relatively easy. When geminal coefficient matrices correspond to disjunct orbital sets, as in Eq. (3.7), the relation for $i \neq j$

$$\sum_n c_{mn}^i c_{ln}^j = 0 \quad (3.8)$$

holds for all m th and l th basis functions. This ensures Eq. (3.6).

Using antisymmetrized product of strongly orthogonal geminals (APSG) in Eq. (3.4) simplifies calculations as matrix elements between APSG-type wavefunctions can be evaluated in a similar manner to those between single determinants.

The APSG calculation can be initiated after choosing the necessary s_i subspaces of one-electron orbitals. This step is usually performed relying on localization procedures and analyzing the spatial distribution of localized orbitals. The optimal orbitals and c_{mn}^i coefficients can be computed e.g. by solving local two-electron effective Schrödinger equations. The APSG scheme where each valence geminal has two one-electron orbitals is referred to as *generalized valence bond* scheme [34].

All in all, APSG calculations can account for static correlation while several important correlation-based phenomena cannot be described by this theory as it lacks intergeminal dynamic correlation. Yet, computational economy of the APSG wavefunction makes it a viable candidate to serve as the underlying reference function in a multiconfigurational perturbational treatment.

3.1.2 Complete Active Space

To construct a multiconfigurational reference wavefunction, it is customary [7, 35] to separate the space of molecular spatial orbitals into 3 subsets in the following manner:

(i) inactive (occupied) orbitals that are occupied by 2 electrons of opposite spin in all configurations, (ii) virtual orbitals that are unoccupied in all configurations, and (iii) active (variably occupied) orbitals that have different occupancy in different configurations. The orbital energies associated with these orbitals follow the order

$$\varepsilon_i^{\text{inactive}} < \varepsilon_j^{\text{active}} < \varepsilon_k^{\text{virtual}},$$

respectively.

This separation is in accordance with chemical intuition suggesting that occupied Hartree-Fock orbitals that are lower in energy remain occupied during chemical processes, while those closer to the Fermi-level may vary in occupancy as reactions take place such as breaking or forming of chemical bonds etc.

The CI expansion

$$|\Phi_{\text{CAS}}\rangle = \sum_K c_K |K\rangle \quad \forall K \in \{\text{CAS-CI}\} \quad (3.9)$$

of the reference function considers all possible K determinants with n_e^{active} active electrons on the n_o^{active} active orbitals, preserving the total spin quantum number of the system distributed. These determinants span the CAS-CI space. The name Complete Active Space (CAS) refers to including all determinants of the CAS-CI space in Eq. (3.9). Number of active electrons and orbitals are options of the method usually given in a short-hand form as $\text{CAS}(n_e^{\text{active}}, n_o^{\text{active}})$. Correspondingly, $\text{CAS}(2,2)$ has two active electrons on two active orbitals, which makes it identical to the previously discussed APSG method for the case of a single correlated geminal.

Since the number of the determinants in CAS-CI increases exponentially with n_e^{active} and n_o^{active} , computation time limits allow for calculations with a couple of tens of active orbitals and electrons. The necessity of CAS calculations with relatively large active subsets can be justified by the fact that for many strongly correlated systems, the importance of each determinant can significantly vary across the potential surface, while at the same time predicting which determinants will be of greater weight at a given point in the expansion can be difficult.

3.2 Dynamic correlation

Augmenting the CAS-CI space to include all determinants – which may become necessary to describe a chemical process – is not practical due to the prohibitive computational cost. As an alternative, it has been suggested [36] to resort to a rather minimal

CAS expansion and apply a more cost-economical scheme to include the individually small effect of the many determinants left out from CAS-CI.

Dynamic correlation techniques developed for the Hartree-Fock reference have been generalized [37] with this purpose for the multireference situation. Among these, let us focus on a few perturbative methods which can be later compared to the newly developed MR-PAPT methods as well as setting the foundation for the derivation of the multireference adapted working equations of MR-PAPT in the MCPT framework.

3.2.1 CAS-based Perturbation Theory

3.2.1.1 CASPT2

Based on the philosophy mentioned above, Roos et al. devised [36] the Complete Active Space Second Order Perturbation Theory (CASPT2) method in 1992. It became widely used and got implemented into the MOLCAS and MOLPRO quantum chemical program systems. The CASPT2 approach assumes a CAS wavefunction as zero-order function $\Psi^{(0)}$.

The first-order correction wavefunction spans the first-order interacting space of double excitations similarly to Eq. (2.9) as

$$|\Psi^{(1)}\rangle = \sum'_{a,b} \sum'_{p,q} \sum_{\sigma,\sigma'} c_{pq}^{ab} p_{\sigma}^{+} q_{\sigma'}^{+} b_{\sigma'}^{-} a_{\sigma}^{-} |\Psi^{(0)}\rangle, \quad (3.10)$$

but the restrictions of the summations (denoted by ') are different in this case, i.e.

$$a, b \in \{\text{inactive, active}\}$$

$$p, q \in \{\text{active, virtual}\}.$$

The case where all orbitals would be active is excluded from the expansion of Eq. (3.10). The coefficients in Eq. (3.10) can be determined by solving the

$$[\hat{H}^{(0)} - E^{(0)}]|\Psi^{(1)}\rangle = [\hat{H}^{(0)} - \hat{H}]|\Psi^{(0)}\rangle \quad (3.11)$$

system of linear equations e.g. iteratively. The energy at second-order can be obtained by the standard RSPT formula

$$E^{(2)} = \langle \Psi^{(0)} | \hat{H} | \Psi^{(1)} \rangle, \quad (3.12)$$

in accordance with Eq. (2.5).

The partitioning underlying CASPT2 is formulated with the Fockian as

$$\hat{H}^{(0)} = \hat{O}\hat{F}\hat{O} + \hat{P}\hat{F}\hat{P}. \quad (3.13)$$

Writing two separate terms in Eq. (3.13) is necessary since the reference function is usually not an eigenfunction of the Fockian. In the above $\hat{O}$ is the projector to the model space formed by the CAS reference function, and $\hat{P}$ projects to the complementary space, orthogonal to the model space. For the second-order energy, the inclusion of the single and double excitations of $\Psi^{(0)}$ are sufficient. Note, that excited functions obtained by replacement operators acting on $\Psi^{(0)}$ are not orthogonal. This must be taken care of when constructing a matrix representation of Eq. (3.13). The Fockian, constructed from the spin-averaged first-order density matrix and usually modified with small occupation-dependent level-shifts [38], has non-diagonal elements that were not included in $\hat{H}^{(0)}$ in the first version of CASPT2 [39]. It was advantageous regarding computational costs, but caused the energies to change discontinuously for rearranging orbital subspaces even if the occupancies remained unchanged. Later implementation of the CASPT2 method [36] that became e.g. part of the MOLPRO program system, therefore included off-diagonals of the Fockian in the zero-order as well.

In 2011, Pulay concluded [40] that the greatest drawback of CASPT2 is its performance being highly dependent on the declaration of orbital subsets. A minimal choice can be easily identified but unfortunately, it also implies insufficient description of the dynamic correlation. The optimal choice for the system in question, on the other hand, requires great insight to the method itself, i.e. overall, CASPT2 is not fail safe to use as a black-box.

3.2.1.2 NEVPT2

In 2001, Malrieu et al. [41] introduced the n-electron valence state perturbation theory (NEVPT2) that starts from the CAS wavefunction written as the antisymmetrized product form Ansatz

$$\Phi_{\text{CAS}} = \hat{\mathcal{A}}[\Phi_{\text{core}}^{(c)} \Phi_{\text{valence}}^{(v)}] \quad (3.14)$$

indicating separately the c core (inactive) and v valence (active) electrons. The CAS function is obtained by

$$\hat{P}_{\text{CAS}} \hat{H} \hat{P}_{\text{CAS}} |\Phi_{\text{CAS}}\rangle = E^{(0)} |\Phi_{\text{CAS}}\rangle, \quad (3.15)$$

i.e. diagonalizing the Hamiltonian in the CAS-CI space.

Theoretical development proceeds by formulating subspaces of determinants in the CI space. In particular, determinants

$$\hat{\mathcal{A}}[\Phi_{\text{core \& inactive}, l}^{-k} \Phi_{\text{valence}}^{+k}] \quad (3.16)$$

constitute subspace s_{kl} . Determinants in Eq. (3.16) feature a common occupancy of $c - k$ electrons distributed over the core and inactive orbitals and vary in the distribution of $v + k$ valence electrons over the active orbitals.

In the so-called totally uncontracted version of the method, the zero-order Hamiltonian is composed by considering the full Hamiltonian in subspaces s_{kl} and neglecting any interaction element between functions of different subspaces. This corresponds to

$$\hat{H}^{(0)} = \sum_{k,l} \hat{P}_{s_{kl}} \hat{H} \hat{P}_{s_{kl}} \quad (3.17)$$

where $\hat{P}_{s_{kl}}$ projects onto subspace s_{kl} .

Due to the possibly large number of the valence part of the determinants, the dimension of subsets s_{kl} can be computationally prohibitive in spite of the fact that $-2 \leq k \leq 2$ is sufficient e.g. for the second-order energy. This initiated the development of various contracted versions of the theory, that are not to be detailed here further.

The NEVPT2 method is available in the ORCA [42] and MOLPRO [43] program systems. Its advantageous features include size-consistency, invariance to rotations among active space orbitals, spin-conservation, and insensitivity to intruder states.

3.2.2 Multiconfigurational Perturbation Theory

The multireference perturbative frameworks discussed above assume a CAS reference. This Section presents two methods that are applicable to a reference function of arbitrary structure. The approach behind both methods termed *multiconfigurational perturbation theory* (MCPT) together with the methods themselves was devised in the Laboratory of Theoretical Chemistry as a black box tool to obtain perturbative corrections to any reference function that is provided as an expansion over determinants of the CI space. Here, the general working formulae of projected and frame-based MCPT (pMCPT and fMCPT, respectively) are discussed as well as the Møller-Plesset partitioning within this framework. The MCPT framework introduced in this Section will be extensively used later in Chapter 4 to formulate two multireference versions of PAPT.

From a theoretical point of view MCPT's main advantage is in its simple formulation and computational implementation while the size-extensivity violation of pMCPT is an unquestionable drawback [44]. Test calculations on zero-order functions with incomplete model space indicated that pMCPT and fMCPT yield similar results [6]. The former was observed to be superior in absolute energies and the latter in energy differences. Comparison of MCPT-MP to APSG-PT of Surján and Rosta [45], and MRMP of Hirao [46] also indicated similar results. [47] A general agreement was also observed among the similar partitionings in different MR-PT frameworks [44]. In other words, since MCPT is

a general framework rather than a single method, performance of MCPT was observed to strongly depend on the partitioning technique used [48].

A common feature of MCPT and the previously discussed methods is their *single-but-multi* nature. This means that a one-dimensional reference space is associated with the projector

$$\hat{O} = |\Phi\rangle\langle\Phi|, \quad (3.18)$$

where the single Φ reference function is taken as not just one but a linear combination of M model determinants

$$|\Phi\rangle = \sum_K^M c_K |K\rangle = c_0 |0\rangle + \sum_{K \neq 0}^M c_K |K\rangle. \quad (3.19)$$

The separation of one $|0\rangle$ model determinant from the rest has importance later in the presentation of the projected version of MCPT. The reference function is normalized as

$$\langle\Phi|\Phi\rangle = \sum_{K=0}^M c_K^2 = 1. \quad (3.20)$$

Firstly, it is supposed that the

$$\langle\Psi|\Phi\rangle \neq 0 \quad (3.21)$$

overlap of the Φ reference function and the Ψ exact wavefunction is non-zero, i.e. Φ is not a singularly bad approximation to Ψ . Secondly, it is also supposed that the expansion of the Φ reference function upon the K determinants is known, i.e. the values of c_K are obtained beforehand e.g. from generalized valence bond, geminal-based or complete active space calculation.

To fulfill the

$$\hat{H}^{(0)}|\Phi\rangle = E^{(0)}|\Phi\rangle, \quad (3.22)$$

zero-order equation. The zero-order energy associated with Φ is fixed to be the eigenvalue of the total Hamiltonian taken with the reference function

$$E^{(0)} = \langle\Phi|\hat{H}|\Phi\rangle, \quad (3.23)$$

so one vector of the n -electron Hilbert space to construct the spectrum of $\hat{H}^{(0)}$ must be Φ . Adding the M model determinants and the remaining determinants of the CI expansion renders the Hilbert space redundant and non-orthogonal since for any K model determinant the overlap with the reference function reads

$$\langle K|\Phi\rangle = \sum_L^M c_L \langle K|L\rangle = c_K. \quad (3.24)$$

Both methods of our current interest tackle the problem of the overlapping vectors with the reference function by performing Schmidt orthogonalization to obtain *projected determinants*

$$|K'\rangle = \hat{P}|K\rangle = |K\rangle - |\Phi\rangle \underbrace{\langle\Phi|K\rangle}_{c_K} = |K\rangle - c_K|\Phi\rangle, \quad (3.25)$$

where in accordance with Eq. (3.18) that defines the projector of the model space, the projector of the *complementary space* was introduced as

$$\hat{P} = 1 - \hat{O} = 1 - |\Phi\rangle\langle\Phi|. \quad (3.26)$$

Note that the projection in Eq. (3.25) leaves all non-model determinants unchanged.

The methods differ, however, in their approach towards the treatment of the redundancy and the overlap among the model determinants.

3.2.2.1 Projected Multiconfigurational Perturbation Theory

Projected Multiconfigurational Perturbation Theory (pMCPT) overcomes the redundancy by spanning the Hilbert space with vectors of the set

$$\{|\Phi\rangle\} \cup \{|K\rangle \mid K \neq 0\}.$$

Note that the so-called *pivotal determinant* introduced in Eq. (3.19) is excluded above. Usually but not necessarily the determinant of largest weight (i.e. c_K^2) in Φ is chosen as the pivotal determinant. With this separation made, the remaining $K \neq 0$ terms can be considered *excited determinants* with respect to the pivot and the molecular orbitals can be straightforwardly classified based on the pivot to the usual occupied and virtual categories.

The $|K'\rangle$ excited model determinants still overlap with each other. This is treated by introducing $\langle\widetilde{K}'|$ reciprocal vectors of the *biorthogonal formalism* fulfilling

$$\langle\widetilde{K}'|L'\rangle = \delta_{KL}. \quad (3.27)$$

It can be verified by substitution in Eq. (3.27) and using Eq. (3.25) that the two-determinantal reciprocal vectors read

$$\langle\widetilde{K}'| = \langle K| - \frac{c_K}{c_0}\langle 0|. \quad (3.28)$$

(See Appendix D for confirmation.) Note that in this case too, the reciprocal set construction leaves all non-model excited determinants unchanged.

The projector of the complementary space (cf. Eq. (3.26)), therefore, in pMCPT reads

$$\hat{P} = \sum_{K \neq 0} |K'\rangle \langle \widetilde{K'}|. \quad (3.29)$$

The spectral resolution of the zero-order Hamiltonian in the biorthogonal formalism can be expressed as

$$\hat{H}^{(0)} = E_0^{(0)} |\Phi\rangle \langle \Phi| + \sum_{K \neq 0} E_K^{(0)} |K'\rangle \langle \widetilde{K'}|, \quad (3.30)$$

where the partitioning is fixed by choosing a set of values for $E_K^{(0)}$.

As for the ground state energy, up to the first-order it is fixed as

$$E_0^{(0)} + E_0^{(1)} = \langle \Phi | \hat{H} | \Phi \rangle, \quad (3.31)$$

and $E_0^{(0)}$ in accordance with Eq. (3.23) is usually fixed as the reference energy

$$E_{\text{ref}} = E_0^{(0)} = \langle \Phi | \hat{H} | \Phi \rangle \quad (3.32)$$

leaving

$$E_0^{(1)} = 0. \quad (3.33)$$

The $E_K^{(0)}$ energies associated with excited determinants may be obtained e.g. by generalizing the well-known Møller-Plesset or Epstein-Nesbet partitionings of the single reference framework.

For example, it may appear plausible to generalize the Møller-Plesset partitioning as

$$E_K^{(0)} = \langle K | \hat{F} | K \rangle = \sum_i^{\text{occ in } K} F_{ii}, \quad (3.34)$$

where $\hat{F}$ may be a generalized Fockian, reflecting the multi-reference nature of Φ [37].

While the expression in Eq. (3.34) falls back to the sum of orbital energies in the single reference case, in multireference scenario, $\hat{F}$ usually exhibits off-diagonal elements, too.

For this reason, the partitioning based on Eq. (3.34) is named after Davidson and Kapuy, who advocated the use of the diagonal elements of F at zero-order [49].

In the next section, a generalization of the Møller-Plesset partitioning in the pMCPT framework is considered, that uses the full matrix representation of $\hat{F}$, in line with CASPT2.

Once the partitioning is fixed, perturbative corrections are generated by the case of the following biorthogonal perturbative formulae.

The first-order wavefunction, orthogonal to the reference function reads

$$|\Psi^{(1)}\rangle = \sum_{K \neq 0} c_K^{(1)} |K'\rangle, \quad (3.35)$$

where the $c_K^{(1)}$ coefficients are expressed as

$$c_K^{(1)} = -\frac{\langle \widetilde{K}' | \hat{H} | \Phi \rangle}{E_K^{(0)} - E_{\text{ref}}} \quad (3.36)$$

in case of the diagonal zero-order Hamiltonian of Eq. (3.30).

The second-order energy formula is obtained as

$$E_{\text{pMCPPT}}^{(2)} = -\langle \Phi | \hat{H} \hat{Q} \hat{H} | \Phi \rangle, \quad (3.37)$$

where the $\hat{Q}$ reduced resolvent was introduced as

$$\hat{Q} = \sum_{K \neq 0} \frac{|K'\rangle \langle \widetilde{K}'|}{E_K^{(0)} - E_{\text{ref}}}. \quad (3.38)$$

In analogy with Eq. (3.13) the zero-order Hamiltonian in the Møller-Plesset partitioning [50] is

$$\hat{H}^{(0)} = E^{(0)} \hat{O} + \hat{P} \hat{F} \hat{P}, \quad (3.39)$$

which upon the substitution of Eqs. (3.18) and (3.26) yields

$$\hat{H}^{(0)} = E_0^{(0)} |\Phi\rangle \langle \Phi| + \sum_{K, L \neq 0} |K'\rangle \underbrace{\langle \widetilde{K}' | \hat{F} | L' \rangle}_{F_{KL}} \langle \widetilde{L}'|. \quad (3.40)$$

Similarly to Eq. (3.30), projection to the complementary space was necessary to ensure that the zero-order equation is fulfilled, cf. Eq. (3.22).

The *generalized Fockian* appearing in Eq. (3.39) reads

$$F_{pq} = h_{pq} + \sum_{rs} P_{sr} [pr || qs], \quad (3.41)$$

where

$$P_{sr} = \langle 0 | r^+ s^- | 0 \rangle \quad (3.42)$$

is the *first-order density matrix* associated with the pivotal determinant.

Again in line with CASPT2, the first-order wavefunction can be obtained in analogy to Eq. (3.11) as

$$[\hat{H}^{(0)} - E^{(0)}] |\Psi^{(1)}\rangle = [E^{(1)} - \hat{W}] |\Phi\rangle. \quad (3.43)$$

In contrast to Eq. (3.36), the $c_K^{(1)}$ coefficients cannot, in this case, simply be calculated by the formula in Eq. (3.36) as the Fockian is not diagonal in the pMCPT basis. One therefore has to project Eq. (3.43) with reciprocal vectors $\langle \widetilde{L}' |$ and solve the resulting

$$\sum_{K \neq 0} \langle \widetilde{L}' | \hat{F} - E^{(0)} | K' \rangle c_K^{(1)} = -\langle \widetilde{L}' | \hat{H} | \Phi \rangle \quad (3.44)$$

system of linear equations.

Let us introduce in the above elements of the $\mathbf{T}^{\Psi^{(1)}}$ coefficient matrix as

$$\begin{aligned} T_{LK}^{\Psi^{(1)}} &= \langle \widetilde{L}' | \hat{F} - E^{(0)} | K' \rangle \\ &= \langle L | \hat{F} - E^{(0)} | K \rangle - \frac{c_L}{c_0} \langle 0 | \hat{F} - E^{(0)} | K \rangle \\ &\quad - c_K \langle L | \hat{F} - E^{(0)} | \Phi \rangle + \frac{c_K c_L}{c_0} \langle 0 | \hat{F} - E^{(0)} | \Phi \rangle \end{aligned} \quad (3.45)$$

and components of the $\mathbf{Y}^{\Psi^{(1)}}$ inhomogenous vector as

$$\begin{aligned} Y_L^{\Psi^{(1)}} &= -\langle \widetilde{L}' | \hat{H} | \Phi \rangle \\ &= -\langle L | \hat{H} | \Phi \rangle + \frac{c_L}{c_0} \langle 0 | \hat{H} | \Phi \rangle. \end{aligned} \quad (3.46)$$

Using the above, Eq. (3.44) can be concisely written as the matrix equation

$$\sum_{K \neq 0} T_{LK}^{\Psi^{(1)}} c_K^{(1)} = Y_L^{\Psi^{(1)}}, \quad (3.47)$$

from which the $c_K^{(1)}$ coefficients can be obtained by inverting $\mathbf{T}^{\Psi^{(1)}}$.

In the study where the Møller-Plesset partitioning was developed in pMCPT [37], excited determinants contributing to $\Psi^{(1)}$ were restricted for computational economy as doubly excited with respect to the pivotal determinant. In this work, double and single excitations are also allowed with respect to the pivot, leading to

$$|\Psi^{(1)}\rangle = \sum_K^{2x \text{ exc}} c_K^{(1)} |K'\rangle = \frac{1}{4} \sum_{ab}^{\text{occ}} \sum_{pq}^{\text{virt}} c_{pq}^{ab(1)} |K_{ab}^{pq'}\rangle + \sum_a^{\text{occ}} \sum_p^{\text{virt}} c_p^{a(1)} |K_a^{p'}\rangle. \quad (3.48)$$

The notation introduced in the above is used extensively in this thesis, where K_a^p stands for the determinant in which the electron from the a occupied spin orbital of the pivotal determinant is excited to the p virtual one. In an analogous manner, K_{ab}^{pq} is the determinant in which, in addition to the aforementioned a to p replacement, the electron from the b occupied orbital was additionally moved to the q virtual one. The excitations obtained this way preserve the net spin z -component of the pivotal determinant, when the spin indices of the spin orbitals a and b agree with those of p and q .

The inclusion of single excitations with respect to the pivotal determinant can be justified by the fact that as one moves away from the single reference self-consistent Hartree-Fock solution, *Brillouin's theorem* no longer holds.

3.2.2.2 Frame-based Multiconfigurational Perturbation Theory

The pivot-dependence of pMCPT can be eliminated by

- (i). averaging over all the solutions where each of the determinants were selected once as pivotal determinant or
- (ii). using an overcomplete set of projected determinants in the complementary space.

Frame theory can be utilized to perform the latter which led to the formulation of the pivot-independent frame-based Multiconfigurational Perturbation Theory (fMCPT). [6]

The overlapping projected determinants of Eq. (3.25) form a so-called *frame*, similar to the concept of a basis but allowing linear dependence on the constituting vectors. The introduction of frame theory to MCPT, therefore, allows to keep all vectors of the linearly dependent set, thus alleviating the need to distinguish one pivotal determinant among the model determinants to be excluded from the complementary space. Instead the construction of the Moore–Penrose inverse of the singular overlap matrix treats the overlap in a closed form.

First, let us divide the complementary space of Eq. (3.26) to the *model and perpendicular spaces* with the projectors

$$\hat{P}_M = \sum_{K=1}^M |K\rangle\langle K| - \hat{O} \quad (3.49)$$

and

$$\hat{P}_\perp = \sum_{K=M+1}^N |K\rangle\langle K|, \quad (3.50)$$

correspondingly, where M denotes the number of determinants in the model and N that of the Full-CI expansion. Lack of the prime notation on the K determinants of Eq. (3.50) is due to the fact that perpendicular space determinants are orthogonal as they are to each other and to the reference function.

Consequently, the zero-order Hamiltonian reads

$$\hat{H}^{(0)} = E_0|\Phi\rangle\langle\Phi| + \hat{P}_M\hat{X}\hat{P}_M + \hat{P}_\perp\hat{Z}\hat{P}_\perp \quad (3.51)$$

with $\hat{X}$ and $\hat{Z}$ specified later.

Analogously to Eq. (3.25), projected determinants in fMCPT read

$$|K'\rangle = \hat{P}_M|K\rangle = |K\rangle - c_K|\Phi\rangle. \quad (3.52)$$

However, reciprocal vectors are replaced (cf. Eq. (3.28)) in this case by *canonical dual frames*

$$\langle\widetilde{K}'| = \sum_{L=1}^M R_{KL}\langle L'| = \sum_{L=1}^M S_{KL}\langle L'| = \langle K'|, \quad (3.53)$$

where $\mathbf{S}$ is the M dimensional overlap matrix of projected model determinants. The rank of $\mathbf{S}$ is $(M - 1)$ since it is the matrix representation of

$$\hat{P}_M = \sum_{K=1}^M |K'\rangle\langle K'| \quad (3.54)$$

(cf. Eq. (3.49)) reading

$$S_{KL} = \langle K'|L'\rangle = \langle K|\hat{P}_M|L\rangle = \delta_{KL} - c_K c_L. \quad (3.55)$$

In Eq. (3.53), $\mathbf{R} = \mathbf{S}$ is the Moore-Penrose inverse of $\mathbf{S}$.

Consequently, Eq. (3.51) can be expressed as

$$\hat{H}^{(0)} = E_0|\Phi\rangle\langle\Phi| + \sum_{K,L=1}^M \mathbf{X}_{KL}|K'\rangle\langle L'| + \sum_{K,L=M+1}^N \mathbf{Z}_{KL}|K\rangle\langle L|. \quad (3.56)$$

The projected determinants do not form an orthogonal basis hence the form of the model space part of Eq. (3.56), where $\mathbf{X}$ is also of dimension M and only of $(M - 1)$ in rank.

In accordance with Eq. (3.51), the $\hat{X}$ model space part of the zero-order Hamiltonian can either be constructed from an arbitrary operator such as the multireference generalized Fockian (see Eq. (3.41))

$$\hat{X} = \hat{F} - \langle\Phi|\hat{F}|\Phi\rangle \quad (3.57)$$

or the total Hamiltonian

$$\hat{X} = \hat{H} - \langle\Phi|\hat{H}|\Phi\rangle. \quad (3.58)$$

In case of a Møller-Plesset type partitioning, the $\hat{Z}$ perpendicular space part of the zero-order Hamiltonian is given by Eq. (3.57).

In line with Eq. (3.47), the linear matrix equation that defines the $c^{(1)}$ coefficient of the $\Psi^{(1)}$ first-order correction wavefunction can be expressed as

$$\begin{bmatrix} {}^M\mathbf{T}^{\Psi^{(1)}} & 0 \\ 0 & {}^T\mathbf{T}^{\Psi^{(1)}} \end{bmatrix} \begin{bmatrix} {}^M\mathbf{c}^{(1)} \\ {}^T\mathbf{c}^{(1)} \end{bmatrix} = \begin{bmatrix} {}^M\mathbf{Y}^{\Psi^{(1)}} \\ {}^T\mathbf{Y}^{\Psi^{(1)}} \end{bmatrix} \quad (3.59)$$

containing a single zero singular value in the ${}^M\mathbf{T}^{\Psi^{(1)}}$ coefficient matrix of the model space due to the redundancy in the frame. Therefore, the first-order correction coefficient can be obtained from

$$c_L^{(1)} = \sum_K G_{LK} Y_K, \quad (3.60)$$

where $\mathbf{G}^{\Psi(1)}$ is a blockdiagonal matrix that consists of the ${}^M\mathbf{G}^{\Psi(1)}$ block, the Moore-Penrose inverse of

$${}^M\mathbf{T}_{LK}^{\Psi(1)} = \langle L'|\hat{X}|K'\rangle \quad (3.61)$$

for model space determinants L, K and the ${}^T\mathbf{G}^{\Psi(1)}$ block, the regular inverse of

$${}^T\mathbf{T}_{LK}^{\Psi(1)} = \langle L|\hat{Y}|K\rangle \quad (3.62)$$

for perpendicular space determinants L, K .

Correspondingly, the first-order wavefunction correction reads

$$|\Psi^{(1)}\rangle = \sum_L^M {}^M c_L^{(1)} |L'\rangle + \sum_{L=M+1}^N {}^T c_L^{(1)} |L\rangle \quad (3.63)$$

$$= \sum_L^M {}^M c_L^{(1)} |L\rangle - \underbrace{\sum_K^M {}^M c_K^{(1)} c_K}_{\kappa} |\Phi\rangle + \sum_{L=M+1}^N {}^T c_L^{(1)} |L\rangle, \quad (3.64)$$

and can be expressed on the basis of unprojected determinants in the latter expression.

If the zero-order Hamiltonians used in the model and perpendicular spaces $\hat{X}$ and $\hat{Y}$ match, coupling of the model and perpendicular spaces can be considered by constructing the off-diagonal blocks of the $\mathbf{T}^{\Psi(1)}$ coefficient matrix (cf. Eq. (3.59)) to give:

$$\begin{bmatrix} {}^M\mathbf{T}^{\Psi(1)} & {}^{MT}\mathbf{T}^{\Psi(1)} \\ \left({}^{MT}\mathbf{T}^{\Psi(1)}\right)^\top & {}^T\mathbf{T}^{\Psi(1)} \end{bmatrix} \begin{bmatrix} {}^M\mathbf{c}^{(1)} \\ {}^T\mathbf{c}^{(1)} \end{bmatrix} = \begin{bmatrix} {}^M\mathbf{Y}^{\Psi(1)} \\ {}^T\mathbf{Y}^{\Psi(1)} \end{bmatrix}, \quad (3.65)$$

where the offdiagonal blocks are constructed as

$${}^{MT}\mathbf{T}_{LK}^{\Psi(1)} = \langle L'|\hat{X}|K\rangle \quad (3.66)$$

for all model space determinant L and perpendicular space determinant K . In this case $\mathbf{G}^{\Psi(1)}$ is simply the Moore-Penrose inverse of the total $\mathbf{T}^{\Psi(1)}$ coefficient matrix containing the zero singular value due to the redundancy.

4. Development of MR-PAPT

An important feature of the Knowles Partitioning in the single-reference approach (see Sec. 2.3) is that the construction of the zero-order Hamiltonian is based upon the projection of an operator equation to the first-order interacting space of double excitations with respect to the Hartree-Fock determinant. When extending this philosophy to the multireference case, the projected method of multiconfigurational perturbation theory (see Sec. 3.2.2.1) is a suitable choice as it implies the concept of a pivotal determinant offering a plausible way to construct singly, doubly etc. excited determinants with respect to.

On the other hand, to alleviate the pivot-dependence of pMCPT hindering the parallelity of results with the Full-CI benchmark and to treat all determinants of the reference function on an equal footing is found to call for more elaborate considerations. This is addressed in the context of the more recently developed frame-based MCPT method.

As mentioned before, the MCPT methods assume no special structure of the Φ multideterminantal reference wavefunction (see Eq. (3.19)) and this feature is conserved for the MR-PAPT methods based upon them.

4.1 MR-PAPT in pMCPT

The zero-order Hamiltonian is formulated in accordance with the pMCPT version of the Møller-Plesset partitioning (see Eq. (3.39)). In the spirit of PAPT, matrix elements of the one-body operator are considered parameters to be optimized. For this reason the $\hat{F}$ generalized Fockian is replaced with the parameterized, perturbation-adapted $\hat{\Lambda}$ operator in Eq. (3.39) to obtain

$$\hat{H}^{(0)} = E_0^{(0)} \hat{O} + \hat{P} \hat{\Lambda} \hat{P}, \quad (4.1)$$

where the projection operators of the model and complementary space are given in Eqs. (3.18) and (3.29), correspondingly. In terms of projected excited determinants and their

reciprocals, the zero-order Hamiltonian reads

$$\hat{H}^{(0)} = E_0^{(0)} |\Phi\rangle\langle\Phi| + \sum_{\substack{K,L \\ \neq 0}} \sum_{i,j} \Lambda_{ij} |K'\rangle\langle\widetilde{K}'| j^+ i^- |L'\rangle\langle\widetilde{L}'|. \quad (4.2)$$

The degree of freedom when it comes to the partitioning is associated with Λ_{ij} in Eq. (4.2) whereas in Eq. (3.30) these are the $E_K^{(0)}$ values that are arbitrary. The role of PAPT in pMCPT is to fix the parameters of $\hat{\Lambda}$ to values that possibly differ from the elements of the $\hat{F}$ generalized Fockian utilized in the Møller-Plesset partitioning of pMCPT (see Eq. (3.39)).

Though a generalized Brillouin theorem can be formulated for the multireference case [37], it does not set the occupied-virtual block of the generalized Fockian zero. It therefore generates no ground to confine the matrix representation of $\hat{\Lambda}$ blockdiagonal in the occupied and virtual blocks. Therefore, the second quantized expression of $\hat{\Lambda}$ is given simply as

$$\hat{\Lambda} = \sum_{i,j} \Lambda_{ij} i^+ j^-. \quad (4.3)$$

Formulation of the multireference extended version of the PAPT equations (see Eq. (2.18) for the single-reference version) that eventually define the set of these $\{\Lambda_{ij}\}$ parameters in the pMCPT framework is detailed in the following sections.

The formulation of the equations should preserve the most prominent features of SR-PAPT. Therefore it is required that

1. the third-order energy correction vanishes upon self-consistency,
2. parameters of the partitioning are unique apart from a constant shift,
3. in the limit where the reference function is single-determinantal the parameters coincide with the parameters of the single reference version and no occupied-virtual matrix elements of $\hat{\Lambda}$ are generated.

4.1.1 pMCPT-PAPT equations analogous to SR-PAPT

The multireference projection functions with two occupied or two virtual outer indices, using the classification of molecular orbitals reflected by the pivotal determinant, are formulated analogously to the single reference version with the inclusion of single excitations as

$$\langle\Theta_{ab}| = \frac{1}{2} \sum_c^{\text{occ}} \sum_{pq}^{\text{virt}} \widetilde{c}_{ac}^{pq(1)} \langle\widetilde{K}_{bc}^{pq'}| + \sum_p^{\text{virt}} \widetilde{c}_a^{p(1)} \langle\widetilde{K}_b^p| \quad (4.4)$$

and

$$\langle \Theta_{pq} | = \frac{1}{2} \sum_{ab}^{\text{occ}} \sum_r^{\text{virt}} \tilde{c}_{ab}^{pr(1)} \langle \widetilde{K}_{ab}^{qr'} | + \sum_a^{\text{occ}} \tilde{c}_a^{p(1)} \langle \widetilde{K}_a^q |, \quad (4.5)$$

respectively. Since the original SR-PAPT equations (see Eq. (2.18)) contain the projection functions in bra vectors, the $\langle \widetilde{K}' |$ reciprocal vectors, introduced in Eq. (3.28) figure in the projection functions. Correspondingly, the $\tilde{c}_K^{(1)}$ reciprocal coefficients of the biorthogonal formalism in the above are introduced as

$$\tilde{c}_K^{(1)} = - \sum_{L \neq 0} \langle \Phi | \hat{H} | L' \rangle G_{LK}, \quad (4.6)$$

where

$$\mathbf{G} = \left(\mathbf{T}^{\Psi(1)} \right)^{-1} \quad (4.7)$$

is the reduced resolvent and the elements of the $\mathbf{T}^{\Psi(1)}$ coefficient matrix are

$$\mathbf{T}_{KL}^{\Psi(1)} = \langle \widetilde{K}' | \hat{H}^{(0)} - E^{(0)} | L' \rangle, \quad (4.8)$$

in accordance with Eq. (3.45).

To preserve hermiticity, the introduction of

$$\langle \Theta_{ij}^+ | = \langle \Theta_{ij} | + \langle \Theta_{ji} | \quad (4.9)$$

for $i \geq j$ is still necessary.

These projection functions alone would, however, not result in a sufficient number of SR-PAPT analogous pMCPT-PAPT equations

$$\sum_{k \geq l} \underbrace{\langle \Theta_{ij}^+ | [l^+ k^- + k^+ l^-] | \Psi^{(1)} \rangle}_{\mathbf{T}_{ij,kl}^\Lambda} \Lambda_{kl} = \underbrace{\langle \Theta_{ij}^+ | \hat{H} | \Psi^{(1)} \rangle}_{\mathbf{Y}_{ij}^\Lambda}, \quad (4.10)$$

where the indices i, j, k, l constituting the hyperindices ij and kl are of the same occupied or virtual category, respectively. The equations are in the same number as the parameters of the diagonal block, but in this case $\{\Lambda_{kl}\}$ parameters need to be defined in the offdiagonal occupied-virtual blocks, too. This makes the construction of additional equations to define the partitioning necessary.

4.1.2 Additional pMCPT-PAPT equations

To fulfill the third requirement via the construction of the additional pMCPT-PAPT equations that are in the same number as the single excitations and the number of the

parameters of the occupied-virtual blocks, let us partition the equation for the first-order wavefunction for the space of single and double excitations as

$$\begin{pmatrix} \mathbf{T}_{SS}^\Lambda & \mathbf{T}_{SD}^\Lambda \\ \mathbf{T}_{DS}^\Lambda & \mathbf{T}_{DD}^\Lambda \end{pmatrix} \begin{pmatrix} \Psi_S^{(1)} \\ \Psi_D^{(1)} \end{pmatrix} = \begin{pmatrix} (\hat{H}\Phi)_S \\ (\hat{H}\Phi)_D \end{pmatrix}. \quad (4.11)$$

Due to Brillouin's theorem and the lack of $\mathbf{T}_{SD}^\Lambda$ -type coefficient matrix elements, the singles' block of the first-order wavefunction is zero in the single reference Knwoles partitioning. However, this is not the case with the multireference adaptation, since Brillouin's theorem no longer holds and there are $\mathbf{T}_{SD}^\Lambda$ -type coefficient matrix elements. Nevertheless, the equations can be designed to yield Λ_{ap} parameters such that $\Psi_S^{(1)}$ still becomes zero by definition.

Inserting $\Psi_S^{(1)} = 0$ into Eq.(4.11) results in

$$\mathbf{T}_{SD}^\Lambda \Psi_D^{(1)} = (\hat{H}\Phi)_S \quad (4.12)$$

$$\mathbf{T}_{DD}^\Lambda \Psi_D^{(1)} = (\hat{H}\Phi)_D \quad (4.13)$$

which can be combined into

$$\mathbf{T}_{SD}^\Lambda (\mathbf{T}_{DD}^\Lambda)^{-1} (\hat{H}\Phi)_D = (\hat{H}\Phi)_S, \quad (4.14)$$

which is essentially a first-order Brueckner condition [51].

In practice, $(\mathbf{T}_{DD}^\Lambda)^{-1}$ is approximated by its diagonal form leading to the

$$\sum_{bq} \mathbf{T}_{ap,bq}^\Lambda \Lambda_{bq} = \langle \widetilde{K_a^{p'}} | \hat{H}_N | \Phi \rangle \quad \forall a \in \{\text{occ}\} \text{ and } \forall p \in \{\text{virt}\} \quad (4.15)$$

concise form of the additional equations defining the partitioning with

$$\mathbf{T}_{ap,bq}^\Lambda = \sum_{cd}^{\text{occ}} \sum_{rs}^{\text{virt}} \langle \widetilde{K_a^{p'}} | b^+ q^- + q^- b^+ | L_{cd}^{rs'} \rangle \frac{\langle \widetilde{L_{cd}^{rs'}} | \hat{H}_N | \Phi \rangle}{\langle \widetilde{L_{cd}^{rs'}} | \hat{\Lambda}_N | L_{cd}^{rs'} \rangle}. \quad (4.16)$$

Since these newly introduced equations are nonlinear due to Λ_{aa} and Λ_{pp} parameters entering the coefficient matrix of Eq.(4.16) via the denominator, the linearized form of Eq.(4.15) is solved in practice together with Eq. (4.10). First, elements $\mathbf{T}_{ap,bq}^\Lambda$ are constructed with F_{aa} and F_{pp} correspondingly and then the diagonal parameters are updated in course of a self-consistent PAPT iteration. Hence Λ_{aa} and Λ_{pp} become consistent in Eq. (4.10) and (4.15) only upon convergence of a PAPT iteration. The eventual contribution of single excitations to the self-consistent $\Psi^{(1)}$ of pMCPT-PAPT is found to be small, but nonzero.

4.1.3 On the vanishing third-order energy correction

Let us inspect the third-order energy correction since the guideline of extending the PAPT procedure to the multireference case is to set it zero upon self-consistency analogously to SR-PAPT. Starting with the traditional Rayleigh-Schrödinger perturbative formula

$$E^{(3)} = \langle \Psi^{(0)} | \hat{H} | \Psi^{(2)} \rangle, \quad (4.17)$$

the reference function Φ serves as the zero-order wavefunction and the second-order correction vector, orthogonal to Φ , expanded on the basis of K' projected excited determinants reads

$$|\Psi^{(2)}\rangle = \sum_{K \neq 0} c_K^{(2)} |K'\rangle. \quad (4.18)$$

The $c_K^{(2)}$ second-order coefficients are expressed as

$$c_K^{(2)} = - \sum_{L \neq 0} G_{KL} \langle \tilde{L}' | \hat{W} - E^{(1)} | \Psi^{(1)} \rangle, \quad (4.19)$$

with matrix $\mathbf{G}$ introduced in Eq. (4.7). Substituting Eqs. (4.18) and (4.19) into Eq. (4.17) yields

$$E^{(3)} = \sum_{L \neq 0} \left[- \underbrace{\sum_{K \neq 0} \langle \Phi | \hat{H} | K' \rangle G_{KL}}_{\tilde{c}_L^{(1)}} \right] \langle \tilde{L}' | \hat{W} - E^{(1)} | \Psi^{(1)} \rangle, \quad (4.20)$$

where the $\tilde{c}_L^{(1)}$ coefficient introduced in Eq. (4.6) is recognized. Consequently, Eq. (4.21) simply reduces to

$$E^{(3)} = \underbrace{\sum_{L \neq 0} \tilde{c}_L^{(1)} \langle \tilde{L}' | \hat{W} - E^{(1)} | \Psi^{(1)} \rangle}_{\langle \tilde{\Psi}^{(1)} |}, \quad (4.21)$$

where the bra counterpart of $\Psi^{(1)}$ is introduced as

$$\langle \tilde{\Psi}^{(1)} | = \sum_{L \neq 0} \tilde{c}_L^{(1)} \langle \tilde{L}' |. \quad (4.22)$$

This allows to simplify Eq. (4.21) to

$$E^{(3)} = \langle \tilde{\Psi}^{(1)} | \hat{W} - E^{(1)} | \Psi^{(1)} \rangle. \quad (4.23)$$

Since in Eq. (4.23), $\hat{W}$ can be replaced with $\hat{H} - \hat{H}^{(0)}$, the third-order energy correction vanishes if

$$\langle \tilde{\Psi}^{(1)} | \hat{H}^{(0)} | \Psi^{(1)} \rangle = \langle \tilde{\Psi}^{(1)} | \hat{H} - E^{(1)} | \Psi^{(1)} \rangle \quad (4.24)$$

holds. However, such a zero-order Hamiltonian cannot be given in a straightforward manner as there is only one equation, i.e. the problem is underdefined for the number of n_{basis}^2 parameters in the zero-order. One could separate each term of the summation for the singly and doubly excited determinants and set up an equation for each determinant, but then the number of equations would not match the number of parameters in general. In typical applications, the former would be larger than the latter, i.e. the problem would be overdefined.

This logical barrier is elevated by following the philosophy introduced by Knowles in the single reference version. In the same manner, projection functions

$$\langle \Theta_K | = \sum_{L \neq 0} \tilde{c}_L^{(1)} \langle \tilde{L}' | \quad (4.25)$$

with restricted summations are introduced. These correspond to $\langle \Theta_{aa} |$ and $\langle \Theta_{pp} |$ in Eqs. (4.4) and (4.5), respectively.

Once projection functions are expressed with molecular orbital indices for the $a = b$ and $p = q$ case, it is straightforward to obtain the expression for $a \neq b$ and $p \neq q$. In analogy with SR-PAPT, the outer molecular orbital indices associated with $\tilde{c}_L^{(1)}$ and $\langle \tilde{L}' |$ are allowed to differ. Note that projection functions with $a \neq b$ and $p \neq q$ have no contribution to $E^{(3)}$. Replacing $\langle \tilde{\Psi}^{(1)} |$ in Eq. (4.24) with the $\langle \Theta_K |$ projection functions of Eqs. (4.4) and (4.5) yields the multireference version of the single reference analogous PAPT equations of Eq. (4.10).

4.1.4 On the redundancy of projection functions

Similarly to the single-reference case, the set of projection functions is affected by redundancy since

$$\sum_a^{\text{occ}} \langle \Theta_{aa}^+ | = \sum_p^{\text{virt}} \langle \Theta_{pp}^+ |. \quad (4.26)$$

This sets the coefficient matrix of Eq. (4.10) singular. In particular, one singular value of matrix $\mathbf{T}^\Lambda$ is zero as a consequence.

This is concluded to have no effect on the perturbative corrections obtained with the Λ_{ij} parameters set by PAPT [3], if the complementary space contribution of the zero-order Hamiltonian is shifted by its expectation value taken with the pivotal determinant.

Accordingly, Eq. (4.1) is replaced by a shifted zero-order Hamiltonian reading as

$$\hat{H}_s^{(0)} = E_{0,s}^{(0)} \hat{O} - \hat{P} \left[\hat{\bar{F}} - \langle 0 | \hat{\bar{F}} | 0 \rangle \right] \hat{P}. \quad (4.27)$$

In the above, $\hat{\bar{F}}$ represents both the $\hat{F}$ generalized Fockian applied in MR-MP and the $\hat{\Lambda}$ operator appearing in the pMCPT based MR-PAPT working formulae.

Since the complementary space component of $\hat{H}_s^{(0)}$ involves an excitation energy at the one-electron level, it is plausible to set

$$E_{0,s}^{(0)} = 0. \quad (4.28)$$

This ensures that

$$\hat{\bar{F}} - \langle 0 | \hat{\bar{F}} | 0 \rangle$$

enters without modification in the zero-order reduced resolvent. A non-zero $E_0^{(0)}$ would introduce a further shift, c.f. the LHS of Eq. (3.43).

It is noted here, that one may envisage subtracting a quantity different from $\langle 0 | \hat{\bar{F}} | 0 \rangle$ from $\hat{H}^{(0)}$. The version according to Eq. (4.27) was implemented in order to develop the working formulae of the pMCPT based version of MR-PAPT with the least modification of the single reference variant that is constructed with normal-ordered operators. In the single reference case, as presented in Eqs. (2.20) and (2.21), the expectation value that was subtracted from the operators was taken with the single Hartree-Fock determinant. Eq. (4.27) is analogously constructed with the subtraction of the expectation value taken with the single pivotal determinant of the pMCPT framework.

Up to this point the freedom in specifying $\hat{H}^{(0)}$ was considered. Another scalar of energy dimension is up to choice, that is the zero-point of the energy scale. This quantity is set to ensure

$$E^{(1)} = 0, \quad (4.29)$$

in parallel with the MCPT formulae presented in Sec. 3.2.2.

A constant shift in the total Hamiltonian is accordingly introduced as

$$\hat{H}_s = \hat{H} - \langle \Phi | \hat{H} | \Phi \rangle. \quad (4.30)$$

Measuring the energy from $\langle \Phi | \hat{H} | \Phi \rangle$, the shifted zero-order ground state energy brings no contribution in agreement with Eq. (4.27).

Obviously,

$$E^{(1)} = \langle \Phi | \hat{H}_s - \hat{H}_s^{(0)} | \Phi \rangle = 0 \quad (4.31)$$

since both operators have vanishing expectation value with Φ .

For the shifted version of Eqs. (3.44) and (4.38), shifted Fockian and $\hat{\Lambda}$ operators need to be expressed as

$$\begin{aligned}\hat{\bar{F}}_s &= \hat{\bar{F}} - \langle 0 | \hat{\bar{F}} | 0 \rangle \\ &= \sum_{i,j} [i^+ j^- - \langle 0 | i^+ j^- | 0 \rangle] \bar{F}_{ij}.\end{aligned}\quad (4.32)$$

The shifted $c_{K,s}^{(1)}$ coefficients of the first-order wavefunction correction of MR-MP and MR-PAPT in pMCPT arise as

$$\sum_{K \neq 0} \underbrace{\langle \widetilde{L}' | \hat{\bar{F}}_s | K' \rangle}_{\mathbf{T}_{KL}^{\Psi^{(1),s}}} c_{K,s}^{(1)} = - \underbrace{\langle \widetilde{L}' | \hat{H} | \Phi \rangle}_{\mathbf{Y}_L^{\Psi^{(1),s}}}, \quad (4.33)$$

where the shifted $\mathbf{T}^{\Psi^{(1),s}}$ coefficient matrix and $\mathbf{Y}^{\Psi^{(1),s}}$ inhomogeneous vector can be introduced. Note that s is omitted from the subscript of $\hat{H}$ on the RHS of Eq. (4.33) on purpose, since a constant has no effect in between orthogonal bra and ket vectors.

As for the equations that define the PAPT parameters, the shifted version of Eq. (4.10) is introduced as

$$\sum_{i \geq j} \mathbf{T}_{ij,kl}^{\Lambda_s} \Lambda_{kl} = \mathbf{Y}_{kl}^{\Lambda_s}. \quad (4.34)$$

The shifted coefficient matrix is formulated as

$$\mathbf{T}_{ij,kl}^{\Lambda_s} = \sum_{i \geq j} \langle \Theta_{kl}^+ | [j^+ i^- - \langle 0 | j^+ i^- | 0 \rangle + i^+ j^- - \langle 0 | i^+ j^- | 0 \rangle] | \Psi^{(1)} \rangle. \quad (4.35)$$

At difference with the $\mathbf{Y}^{\Psi^{(1)}}$ inhomogeneous vector of the equations that define $\Psi^{(1)}$, the shift of $\hat{H}$ does appear in the inhomogeneous vector of the PAPT-equations reading as

$$\mathbf{Y}_{kl}^{\Lambda_s} = \langle \Theta_{kl}^+ | \hat{H}_s | \Psi^{(1)} \rangle. \quad (4.36)$$

Note however, that $E^{(1)}$ can be omitted in Eq. (4.36), due to Eq. (4.31). This leads to

$$\mathbf{Y}_{kl}^{\Lambda_s} = \langle \Theta_{kl}^+ | [\hat{H} - \langle \Phi | \hat{H} | \Phi \rangle] | \Psi^{(1)} \rangle. \quad (4.37)$$

The constant shifts discussed in this Section are intentionally omitted from the formulae introduced at the beginning of Sec. 4.1.1 in order to keep the presentation transparent.

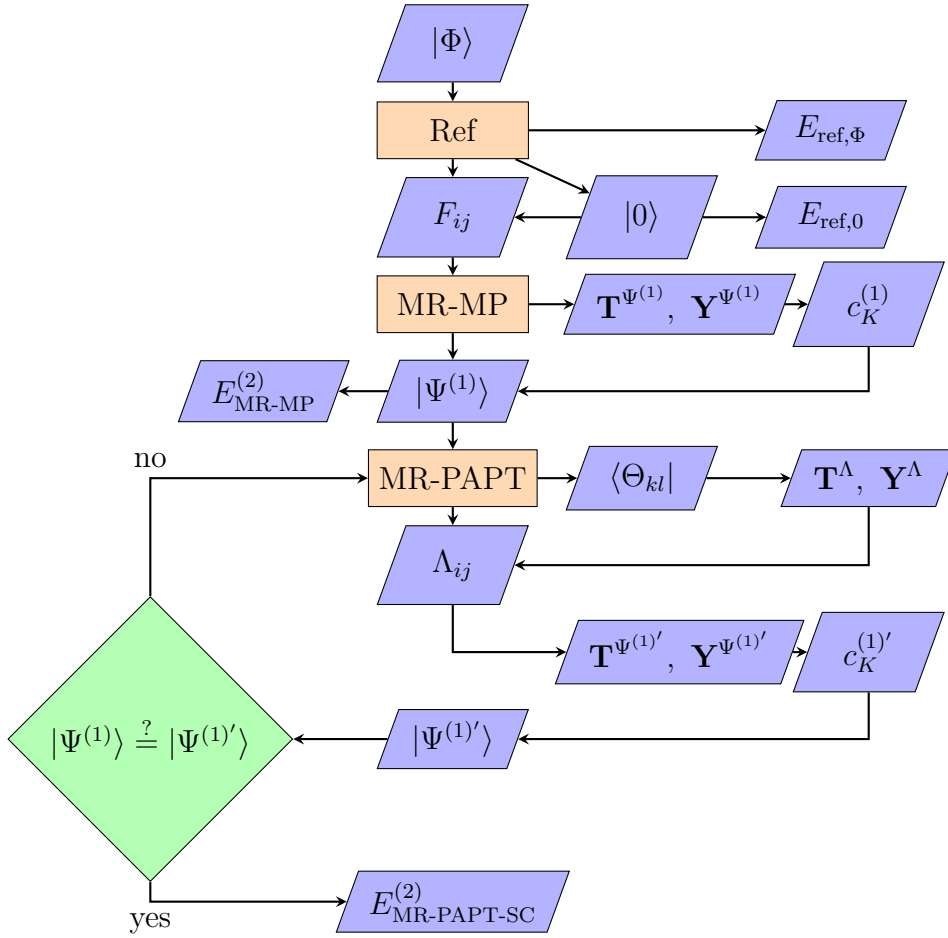


Figure 4.1: The workflow of the MR-PAPT working code

4.1.5 Workflow

A working code was constructed in FORTRAN90 language, the workflow of which is presented on Fig. 4.1. The input of the code is the c_K coefficients and the specification of the associated K determinants as well as the one- and two-electron integrals on the molecular orbital basis. The code applies the pMCPT multireference adaptation of PAPT and solves the emerging equations iteratively until self-consistency is reached. The ground state energy is calculated in each iteration step up to second-order.

As it is illustrated by Fig. 4.1, upon the selection of the pivotal determinant within the Φ reference function, reference energies are calculated and the Fockian is constructed using Eqs. (3.41) and (3.42). Then the MR-MP procedure is initialized using the Fockian, which is constructed using the pMCPT method described in Sec. 3.2.2.1. The MR-MP

procedure is used to create the $\Psi^{(1)}$ first-order correction wavefunction and obtain the underlying $c_K^{(1)}$ coefficients by solving the matrix equation involving $\mathbf{T}^{\Psi^{(1)}}$ and $\mathbf{Y}^{\Psi^{(1)}}$ as shown in Eq. (3.47). Then $\Psi^{(1)}$ is used to initiate the pMCPT based MR-PAPT procedure. It entails construction of the $\langle \Theta_{kl} |$ projection functions followed by the solution of the matrix equation involving $\mathbf{T}^\Lambda$ and $\mathbf{Y}^\Lambda$ – e.g. via the singular-value decomposition of the former – to obtain the set of Λ_{ij} parameters. This defines the PAPT partitioning and allows for the calculation of the $c_K^{(1)'}$ coefficients yielding the $\Psi^{(1)'}$ first-order PAPT correction wavefunction, again through the solution of the corresponding matrix equation

$$\sum_{K \neq 0} \underbrace{\langle \tilde{L} | \hat{\Lambda} - E^{(0)} | K' \rangle}_{\mathbf{T}_{KL}^{\Psi^{(1)'}}} c_K^{(1)'} = - \underbrace{\langle \tilde{L} | \hat{H} | \Phi \rangle}_{\mathbf{Y}_L^{\Psi^{(1)'}}} \quad (4.38)$$

involving $\mathbf{T}^{\Psi^{(1)'}}$ and $\mathbf{Y}^{\Psi^{(1)'}}$. The solution of the PAPT equations is set to be self-consistent iteratively by constructing the projection functions from $\Psi^{(1)'}$ in a consecutive PAPT procedure until the first-order correction wavefunction and correspondingly, the second-order energy converges.

The MR-PAPT working code was tested against the SR-PAPT working code that was in return tested against the implementation of P. Knowles. The bug-free character of the multireference extension was further checked by inspecting and cross-checking the intermediate quantities generated during the MR-PAPT procedure for simple, 2- and 4-electron systems in minimal basis set. Since the routines underlying the MR-PAPT implementations evaluate matrix elements taken with determinants of any number of electrons, it can be positively expected that correcting bugs for few-electron systems ensures bug-free nature for systems with more electrons and/or orbitals, too.

The relatively most time consuming part of the working code was parallelized in MPICH to allow for the calculation of chemical systems in relatively large basis sets. For the details of the parallelization, see Appendix E.

4.2 MR-PAPT in fMCPT

Having assessed the preliminary test results of pMCPT based MR-PAPT (see Sec. 5.1 of the next Chapter) it was concluded that a pivot-independent variant is called for, suitably developed in frame-based MCPT as the general framework. Analogously to the pMCPT formulation, the fMCPT formulation of PAPT derives from the zero-order Hamiltonian of Eq. (3.56) reading as

$$\hat{H}^{(0)} = E_0 |\Phi\rangle \langle \Phi| + \sum_{K,L=1}^M \mathbf{X}_{KL} |K'\rangle \langle L'| + \sum_{K,L=M+1}^N \mathbf{Z}_{KL} |K\rangle \langle L|, \quad (4.39)$$

where the first term is zero due to the energy shift in the Hamiltonian

$$\langle \Phi | \hat{H}^{(0)} | \Phi \rangle = E_0 = 0 \quad (4.40)$$

and matrix elements of the $\hat{\Lambda}$ effective one-body zero-order Hamiltonian are obtained simply as

$$\Lambda_{KL} = \langle K | \hat{\Lambda} | L \rangle, \quad (4.41)$$

holding

$$\langle \Phi | \Lambda | \Phi \rangle = 0. \quad (4.42)$$

In Eq. (4.39), the $\hat{X}$ operator in the model space can either be the total Hamiltonian or $\hat{\Lambda}$ in accordance with Eqs. (3.57) and (3.58), respectively.

With the aim of treating open shell cases, the second quantized form of the generalized, pivot-independent $\hat{\Lambda}$ is (cf. Eq. (4.3)) constructed as

$$\begin{aligned} \hat{\Lambda} = \frac{1}{2} \sum'_{IJ} \left[\Lambda_{IJ}^{\alpha} \left(I_{\alpha}^{+} J_{\alpha}^{-} + J_{\alpha}^{+} I_{\alpha}^{-} - 2P_{IJ}^{\alpha} \right) \right. \\ \left. + \Lambda_{IJ}^{\beta} \left(I_{\beta}^{+} J_{\beta}^{-} + J_{\beta}^{+} I_{\beta}^{-} - 2P_{IJ}^{\beta} \right) \right] \end{aligned} \quad (4.43)$$

for generic spatial molecular orbitals I, J , where density matrices P^{α} and P^{β} read

$$P_{IJ}^{\alpha} = \langle \Phi | I_{\alpha}^{+} J_{\alpha}^{-} | \Phi \rangle \quad (4.44)$$

$$P_{IJ}^{\beta} = \langle \Phi | I_{\beta}^{+} J_{\beta}^{-} | \Phi \rangle. \quad (4.45)$$

The restriction on the summation indicates that only the diagonal block of the Λ parameters is filled according to

$$\Lambda^{\alpha} = \begin{bmatrix} \Lambda_{AB}^{\alpha} & 0 & 0 \\ 0 & \Lambda_{XY}^{\alpha} & 0 \\ 0 & 0 & \Lambda_{PQ}^{\alpha} \end{bmatrix} \quad \text{and} \quad \Lambda^{\beta} = \begin{bmatrix} \Lambda_{AB}^{\beta} & 0 & 0 \\ 0 & \Lambda_{XY}^{\beta} & 0 \\ 0 & 0 & \Lambda_{PQ}^{\beta} \end{bmatrix} \quad (4.46)$$

with a new, augmented index convention for core i.e. doubly occupied in Φ molecular orbitals A, B , active i.e. partially occupied in Φ molecular orbitals X, Y and inactive i.e. unoccupied in Φ molecular orbitals P, Q .

4.2.1 Projection functions

In line with Eq. (4.46), only the corresponding projection functions with core-core, active-active and inactive-inactive outer molecular orbital indices are constructed. The form of the projection functions needs to be fundamentally different from Eqs. (4.4) and (4.5) as determinants of the first-order interacting space cannot in this case be classified simply as singly or doubly excited as there is no underlying pivotal determinant with respect to calculate the excitation levels.

The expressions, derived based on the generalized derivation of the Knowles Partitioning as presented in Appendix B, read

$$\langle \Theta_{IJ}^\alpha | = \sum_{K,L=1}^N \frac{\langle \Phi | \hat{H} \hat{Q} | L' \rangle}{\left(\mathbf{G}^{\Psi(1)} \mathbf{Y}^{\Psi(1)} \right)_L} \langle L' | I_\alpha^+ J_\alpha^- - \underbrace{\langle \Phi | I_\alpha^+ J_\alpha^- | \Phi \rangle}_{P_{IJ}^\alpha} | K' \rangle \langle K' | \quad (4.47)$$

$$\langle \Theta_{IJ}^\beta | = \sum_{K,L=1}^N \frac{\langle \Phi | \hat{H} \hat{Q} | L' \rangle}{\left(\mathbf{G}^{\Psi(1)} \mathbf{Y}^{\Psi(1)} \right)_L} \langle L' | I_\beta^+ J_\beta^- - \underbrace{\langle \Phi | I_\beta^+ J_\beta^- | \Phi \rangle}_{P_{IJ}^\beta} | K' \rangle \langle K' |, \quad (4.48)$$

where I, J are spatial molecular orbital indices and K, L refer to determinants of the first-order interacting space (FOCI). To preserve hermiticity, equations

$$\langle \Theta_{IJ}^\alpha |^+ = \langle \Theta_{IJ}^\alpha | + \langle \Theta_{JI}^\alpha | \quad (4.49)$$

and

$$\langle \Theta_{IJ}^\beta |^+ = \langle \Theta_{IJ}^\beta | + \langle \Theta_{JI}^\beta | \quad (4.50)$$

hold in this case, too.

To expand the projection functions on the basis of unprojected determinants, analogous transformation can be performed to subtracting the κ term in Eq. (3.64).

4.2.2 PAPT Equations

Since the blockdiagonal structure of $\hat{\Lambda}$ is maintained in this formulation, it is sufficient to construct SR-PAPT analogous equations with the fMCPT version of projection

functions as

$$\begin{aligned}
 & \sum_{K,L=1}^N \frac{1}{2} \underbrace{\langle \Theta_{IJ}^{\alpha+} | K_{\alpha}^+ L_{\alpha}^- + L_{\alpha}^+ K_{\alpha}^- - 2P_{KL}^{\alpha} | \Psi^{(1)} \rangle}_{\mathbf{T}_{IJ,KL}^{\Lambda\alpha\alpha}} \Lambda_{KL}^{\alpha} \\
 & + \sum_{K,L} \frac{1}{2} \underbrace{\langle \Theta_{IJ}^{\alpha+} | K_{\beta}^+ L_{\beta}^- + L_{\beta}^+ K_{\beta}^- - 2P_{KL}^{\beta} | \Psi^{(1)} \rangle}_{\mathbf{T}_{IJ,KL}^{\Lambda\alpha\beta}} \Lambda_{KL}^{\beta} = \underbrace{\langle \Theta_{IJ}^{\alpha+} | \hat{H} | \Psi^{(1)} \rangle}_{\mathbf{Y}_{IJ}^{\Lambda\alpha}}
 \end{aligned} \tag{4.51}$$

and

$$\begin{aligned}
 & \sum_{K,L=1}^N \frac{1}{2} \underbrace{\langle \Theta_{IJ}^{\beta+} | K_{\alpha}^+ L_{\alpha}^- + L_{\alpha}^+ K_{\alpha}^- - 2P_{KL}^{\alpha} | \Psi^{(1)} \rangle}_{\mathbf{T}_{IJ,KL}^{\Lambda\beta\alpha}} \Lambda_{KL}^{\alpha} \\
 & + \sum_{K,L=1}^N \frac{1}{2} \underbrace{\langle \Theta_{IJ}^{\beta+} | K_{\beta}^+ L_{\beta}^- + L_{\beta}^+ K_{\beta}^- - 2P_{KL}^{\beta} | \Psi^{(1)} \rangle}_{\mathbf{T}_{IJ,KL}^{\Lambda\beta\beta}} \Lambda_{KL}^{\beta} = \underbrace{\langle \Theta_{IJ}^{\beta+} | \hat{H} | \Psi^{(1)} \rangle}_{\mathbf{Y}_{IJ}^{\Lambda\beta}},
 \end{aligned} \tag{4.52}$$

where I, J, K, L are restricted as follows: I and J are both core, active or inactive indices of spatial molecular orbitals, and independently, K and L are both core, active or inactive indices of spatioal molecular orbitals.

In line with Eq. (4.46), the number of equations defining the PAPT partitioning matches the number of Λ_{IJ} parameters and can be written as a linear matrix equation as

$$\begin{bmatrix} \mathbf{T}^{\Lambda\alpha\alpha} & \mathbf{T}^{\Lambda\beta\alpha} \\ \mathbf{T}^{\Lambda\alpha\beta} & \mathbf{T}^{\Lambda\beta\beta} \end{bmatrix} \begin{bmatrix} \Lambda^{\alpha} \\ \Lambda^{\beta} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}^{\Lambda\alpha} \\ \mathbf{Y}^{\Lambda\beta} \end{bmatrix}. \tag{4.53}$$

4.2.2.1 Efficient coefficient matrix calculation

Upon expansion of the first-order correction wavefunction as

$$|\Psi^{(1)}\rangle = \sum_I \left(\mathbf{G}^{\Psi^{(1)}} \mathbf{Y}^{\Psi^{(1)}} \right)_I |I'\rangle \tag{4.54}$$

and the substitution of the definition of the projection functions, e.g. the LHS of Eq. (4.51) with inactive molecular orbital indices yields

$$\begin{aligned}
 T_{PQ,RS}^{\Lambda\alpha\alpha} &= \frac{1}{2} \langle \Theta_{PQ}^{\alpha+} | \sum_I \left(\mathbf{G}^{\Psi^{(1)}} \mathbf{Y}^{\Psi^{(1)}} \right)_I \cdot \left(R_{\alpha}^+ S_{\alpha}^- + S_{\alpha}^+ R_{\alpha}^- - 2P_{RS}^{\alpha} \right) | I' \rangle \\
 &= \frac{1}{2} \sum_L \left(\mathbf{G}^{\Psi^{(1)}} \mathbf{Y}^{\Psi^{(1)}} \right)_L \langle L' | P_{\alpha}^+ Q_{\alpha}^- + Q_{\alpha}^+ P_{\alpha}^- - 2P_{PQ}^{\alpha} \underbrace{\sum_{K=1}^N |K'\rangle \langle K'|}_{\hat{P}} \cdot \\
 &\quad \cdot \sum_I \left(\mathbf{G}^{\Psi^{(1)}} \mathbf{Y}^{\Psi^{(1)}} \right)_I \cdot \left(R_{\alpha}^+ S_{\alpha}^- + S_{\alpha}^+ R_{\alpha}^- - 2P_{RS}^{\alpha} \right) | I' \rangle
 \end{aligned} \tag{4.55}$$

with $R \neq S$, into which due to its idempotency, another $\hat{P}$ projector of the complementary space can be inserted to give

$$\begin{aligned}
 T_{PQ,RS}^{\Lambda\alpha\alpha} &= \frac{1}{2} \sum_L \left(\mathbf{G}^{\Psi^{(1)}} \mathbf{Y}^{\Psi^{(1)}} \right)_L \langle L' | P_{\alpha}^+ Q_{\alpha}^- + Q_{\alpha}^+ P_{\alpha}^- - 2P_{PQ}^{\alpha} \underbrace{\sum_{K=1}^N |K'\rangle \langle K'|}_{\hat{P}} \cdot \\
 &\quad \cdot \sum_I \left(\mathbf{G}^{\Psi^{(1)}} \mathbf{Y}^{\Psi^{(1)}} \right)_I \underbrace{\sum_{J=1}^N |J'\rangle \langle J'|}_{\hat{P}} \left(R_{\alpha}^+ S_{\alpha}^- + S_{\alpha}^+ R_{\alpha}^- - 2P_{RS}^{\alpha} \right) | I' \rangle,
 \end{aligned} \tag{4.56}$$

simplifying to

$$T_{PQ,RS}^{\Lambda\alpha\alpha} = \frac{1}{2} \langle \Theta_{PQ}^{\alpha+} | \Theta_{RS}^{\alpha+} \rangle. \tag{4.57}$$

In other words, the coefficient matrix of the fMCPT-PAPT equations is simply constructed of the overlap of the projection functions. Analogous derivations can be performed for the other blocks of the $\mathbf{T}^{\Lambda}$ coefficient matrix allowing for efficient calculations without the need to perform parallelization of this section of the working code in contrast to the corresponding part of the pMCPT working code (as presented in Appendix E).

5. Results and Discussion

In this Chapter, three different kinds of preliminary tests of the multireference adaptations of PAPT are discussed entailing bond dissociation potential curves, size-consistency as well as transition from a three-atom molecule configuration to a two-atom molecule and a distant single-atom configuration.

On the example of the bond dissociation curve of the lithium-hydride (see Sec. 5.1), the performance of the firstly developed projected MCPT adaptation of the Knowles Partitioning (pMCPT-PAPT) is benchmarked with respect to the exact Full-CI solution and the Møller-Plesset partitioning in that framework (pMCPT-MP). The findings of this comparison indicating severe losses in parallelity to the Full-CI solution, as mentioned in Sec. 4.2 of the previous Chapter, gave the motivation to develop the pivot-invariant formulation of the Knowles Partitioning.

In Sec. 5.2, the most recently developed, first pivot-independent frame-based MCPT adaptation of the Knowles Partitioning (fMCPT-PAPT) is investigated in comparison to the Full-CI, the Møller-Plesset partitioning in the same framework and also to the projected MCPT adaptation.

The series of these short preliminary tests is concluded with a size-consistency test in Sec. 5.3.

5.1 Potential energy curve of lithium-hydride

First, let us consider the potential curve for the dissociation of the heteronuclear diatomic lithium-hydride molecule. In the simple linear combination of atomic orbitals to molecular orbitals picture, the 4-electron system in equilibrium geometry can be described by a doubly occupied core molecular orbital and a doubly occupied bonding molecular orbital that is polarized towards the lithium atom. The molecule, being stable but highly reactive, can be associated with a bond order of 1 and a bonding energy of 243 kJ/mol, correspondingly.

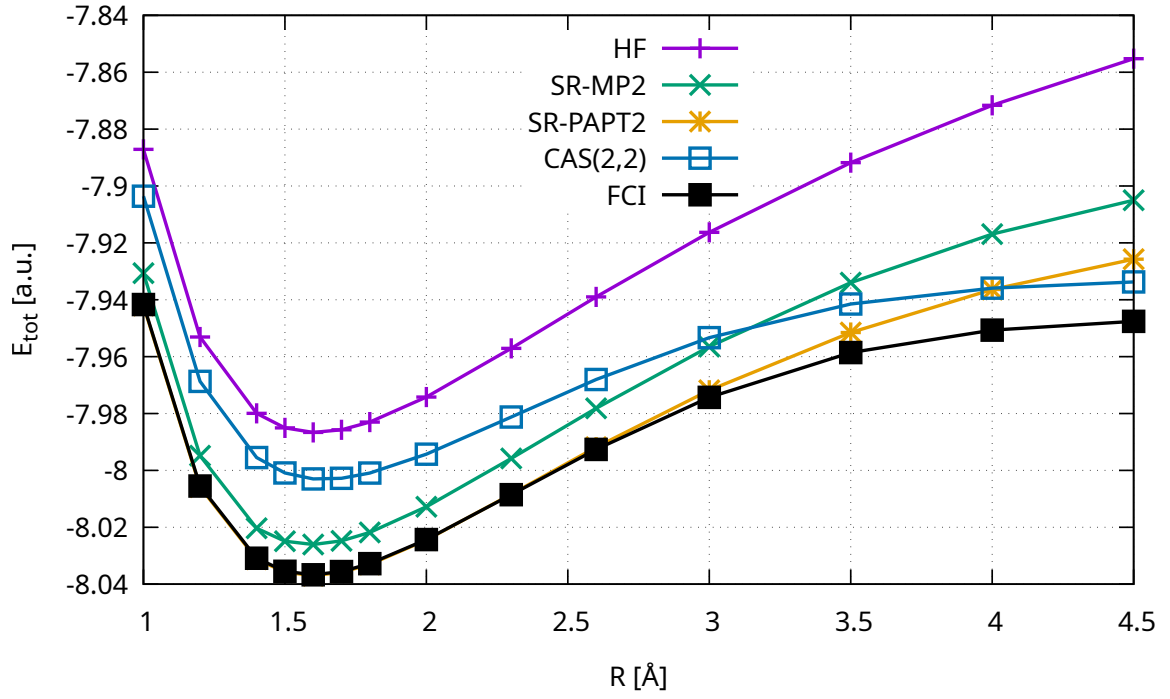


Figure 5.1: Total energies of LiH in cc-pVTZ basis.

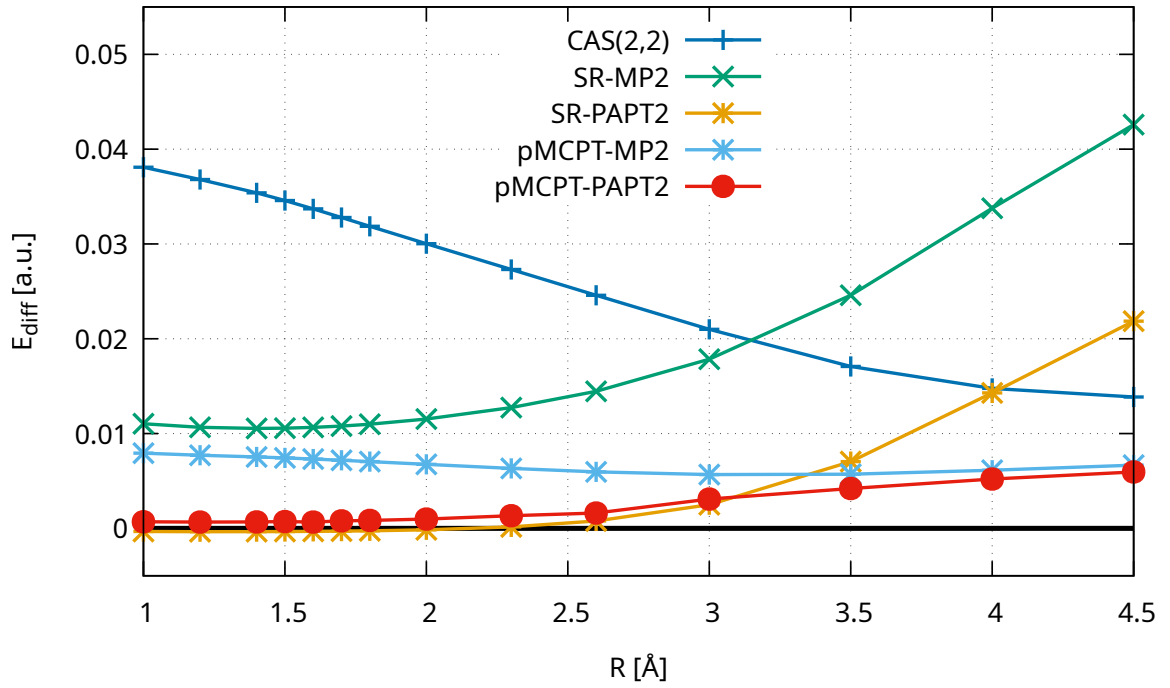


Figure 5.2: Energies differences from Full-CI of LiH in cc-pVTZ basis.

The system is weakly correlated at interatomic distances close to the equilibrium, where the Hartree-Fock determinant is the only dominant term of the CI expansion. At larger distances, the system starts to exhibit some strongly correlated characteristics as the weight of the first (lowest in energy) doubly excited determinant becomes comparable to the weight of the Hartree-Fock determinant.

Fig. 5.1 shows potential energy curves of LiH calculated in correlation-consistent triple-zeta (cc-pVTZ) basis set. It can be observed that the equilibrium bond length is around 1.6 Å. Medium level of correlation is exhibited above cca. 2.5 Å, indicated by the increasing distance between the Hartree-Fock and the Full-CI curves. Single reference second-order perturbative calculations (SR-MP2 and SR-PAPT2) can account for dynamic correlation but run off roughly in parallel with the Hartree-Fock curve in the strongly correlated regime. In Fig. 5.1 the efficiency of SR-PAPT2 is well-manifested, it is able to increase the accuracy of SR-MP2 and provide results that are an order of magnitude closer to Full-CI unless the system necessitates the consideration of static correlation (above cca. 2.5 Å).

As for the multireference methods, a two-electron two-orbital CAS calculation as detailed in Sec. 3.1.2, composed of two determinants can account for the static correlation of the LiH system. This is manifested in Fig. 5.1 in the CAS(2,2) curve staying closer to Full-CI in the dissociated regime than Hartree-Fock. Overall, the CAS(2,2) curve is observed to be more parallel to the Full-CI curve than the single reference curves. At the same time, CAS(2,2) does not account for sufficient dynamic correlation, as reflected by the significant non-parallelity of the difference curve in Fig. 5.2. Still, it can serve as the reference function for the calculation of perturbative corrections (MR-MP2 and MR-PAPT2) using the pMCPT and fMCPT based methods as detailed in Sec. 3.2.2.

This test on the LiH molecule that was conducted, as mentioned in the introduction of this Chapter, before the development of fMCPT-PAPT serves as a good example to stress the features of pMCPT-PAPT with respect to pMCPT-MP and single reference methods that called for the refinement of the multireference generalization recently.

To illustrate the performance of pMCPT-PAPT2 with respect to pMCPT-MP2 and the single reference perturbative methods, Fig. 5.2 shows the energy differences from the Full-CI benchmark of the same dataset as Fig. 5.1. As mentioned before in Sec. 1.3, in order to obtain chemically correct potential curves, a post-Hartree-Fock multireference method is required to treat static and dynamic correlation in balance. In other words, besides running close to the zero-point in Fig. 5.2, the more a method's curve resembles to a straight horizontal line, the better.

Comparing the SR-MP2 curve to the pMCPT-MP2 curves in Fig. 5.2 reveals that

the contribution of static correlation allows for a balanced description of the LiH system. Projected MCPT based MP2 significantly outperforms SR-MP2 throughout the potential curve with ground state energy values of MR-MP2 being consistently only 5 - 10 millihartrees above the Full-CI results.

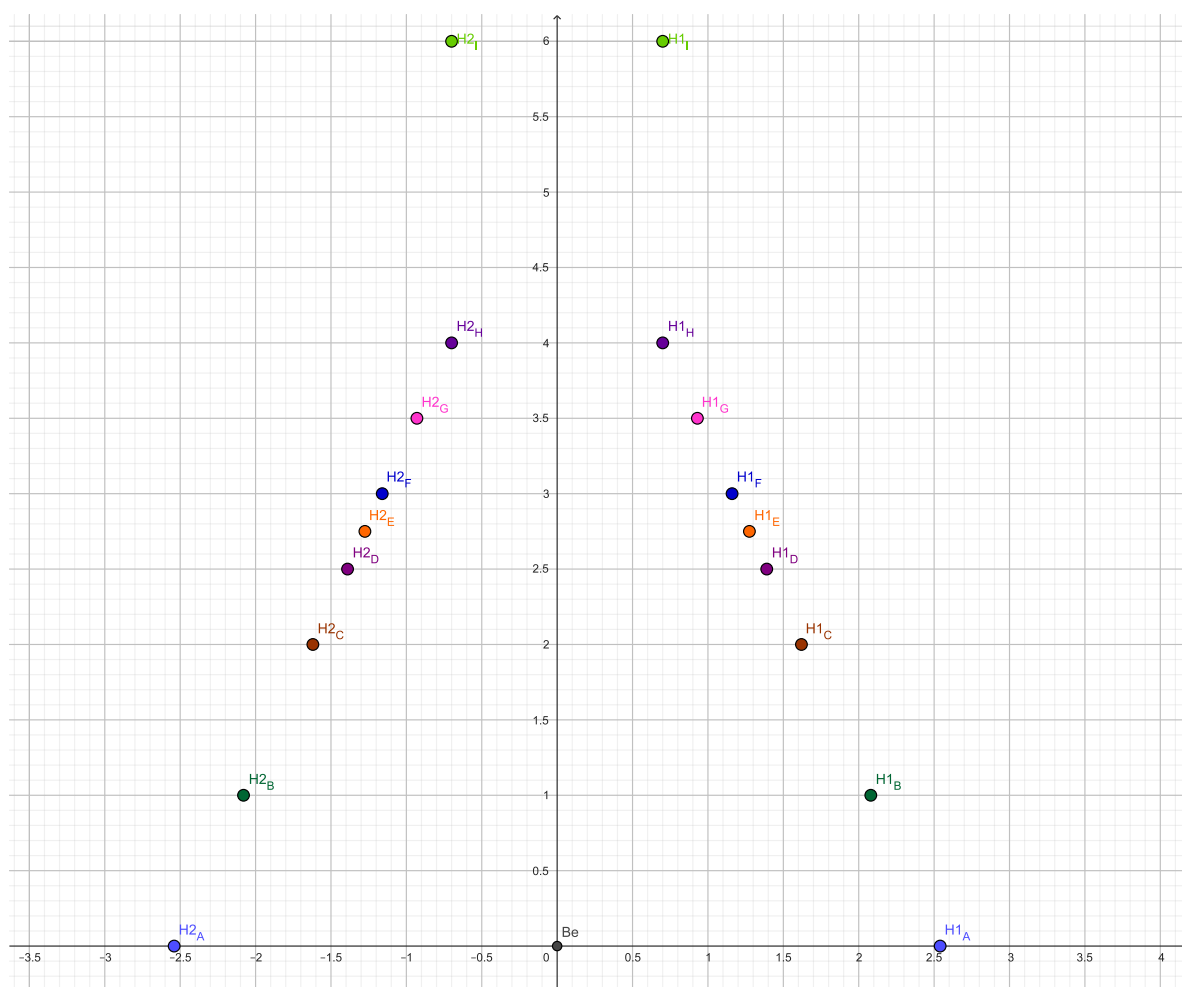
As for pMCPT-PAPT2, in the weakly correlated regime, the little room left for enhancement with respect to SR-PAPT2 is not exploited well. The pMCPT-PAPT2 curve is slightly but consistently worse than SR-PAPT2 at around equilibrium geometry. Comparing with pMCPT-MP2, pMCPT-PAPT2 is observed to bring significant improvement with deviations from Full-CI under the 5 millihartree lower limit of pMCPT-MP2 in all points, and even under 1 millihartree in some points. However, the deviation of pMCPT-PAPT2 from Full-CI is not balanced, as there is an increasing tendency of the error as the system exhibits more strongly correlated characteristics.

This feature of the pMCPT based adaptation of PAPT is most likely to be caused by the enhanced pivot-dependence. The pMCPT method itself is not invariant to the choice of the pivot. In this case, the PAPT equations are formulated utilizing the concept of the pivotal determinant when constructing the parameters that define the zero-order problem. In particular, the restricted summations of the projection functions in Eqs. (4.4) and (4.5) classify the molecular orbitals to be either occupied or virtual with respect to the pivotal determinant.

5.2 Geometry points of berillium-hydride

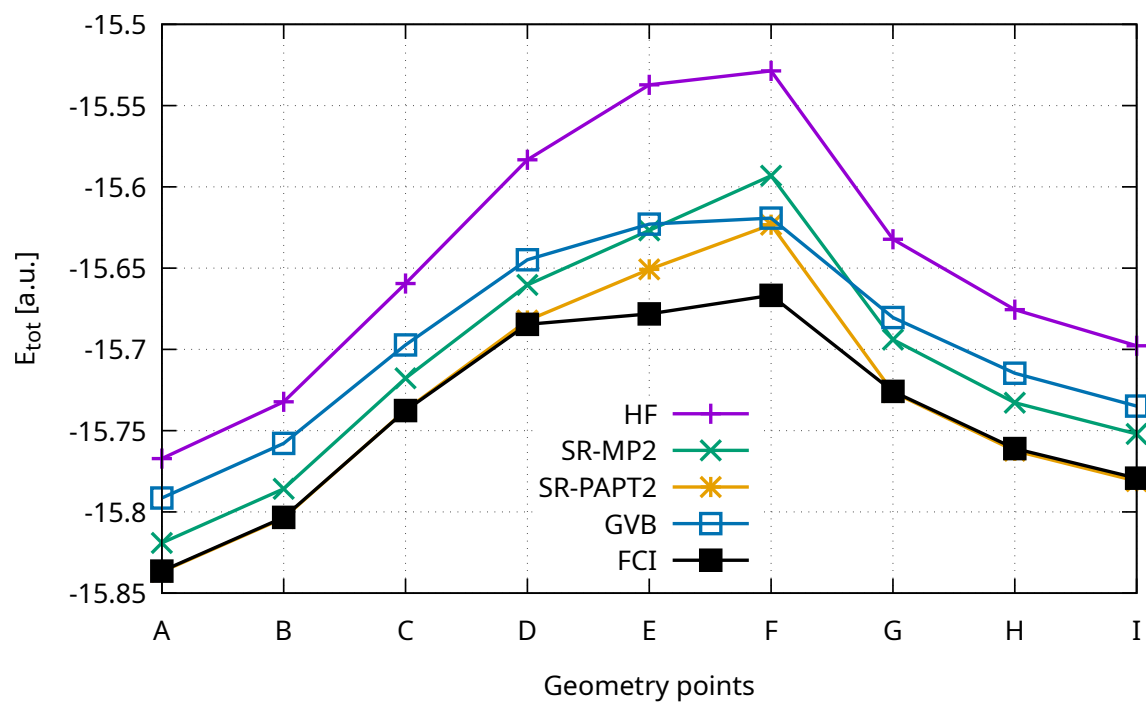
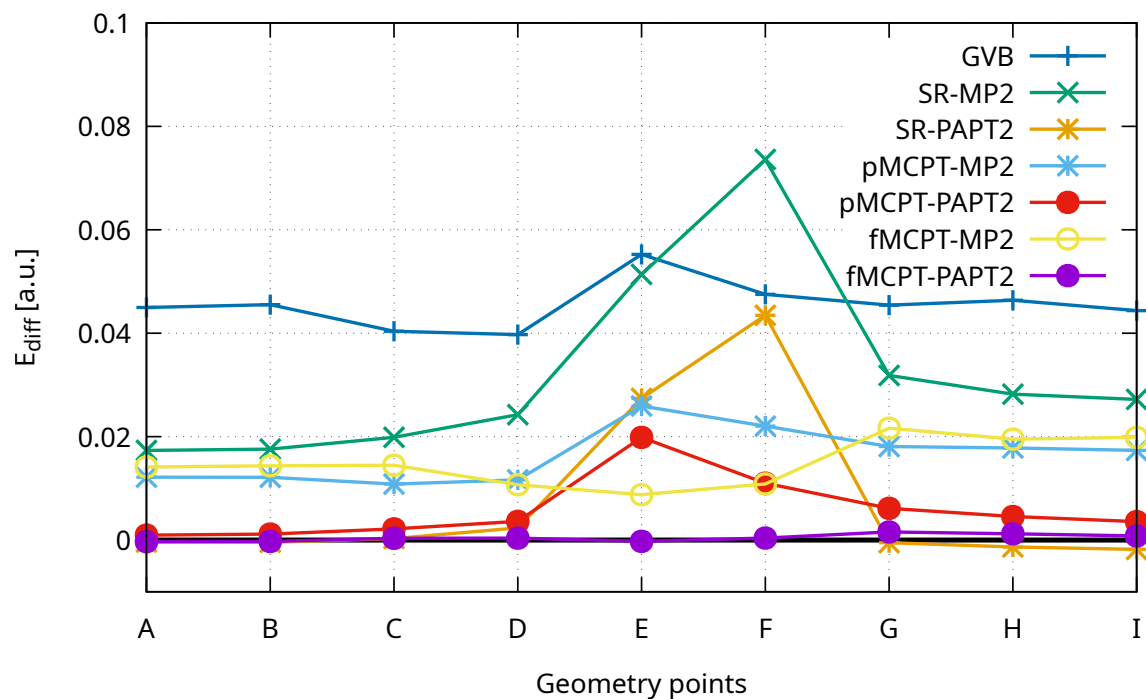
As a modest size test case of covalent bond cleavage and formation, nine standardized geometry points (A-I) were devised by Purvis and Bartlett [52]. The points describe different configurations along the reaction path where the linear BeH₂ molecule is distorted into a H₂ molecule and distant Be atom. This is considered to be a challenging inherent multireference problem of medium-level correlation. In this test, the Be atom lays in the origin and the two H atoms are placed symmetrically to the z-axis with coordinates in atomic unit $(0, \pm 2.54, 0)$, $(0, \pm 2.08, 1.0)$, $(0, \pm 1.62, 2.0)$, $(0, \pm 1.39, 2.5)$, $(0, \pm 1.275, 2.75)$, $(0, \pm 1.16, 3.0)$, $(0, \pm 0.93, 3.5)$, $(0, \pm 0.70, 4.0)$, and $(0, \pm 0.70, 6.0)$, respectively (see Fig. 5.3). The process can also be considered as the perpendicular insertion of a Be atom into a H₂ molecule starting from point I, and proceeding towards point A.

Assessment of the performance of the pMCPT and fMCPT versions of MR-PAPT with respect to other methods, on the geometry points A-I of the BeH₂ molecule in cc-pVDZ basis set is presented in Figs. 5.4 and 5.5. Total energies are displayed in Fig. 5.4 while Fig. 5.5 depicts difference from Full-CI.

Figure 5.3: Geometry points of BeH_2 .

As can be observed in Fig. 5.4, the single reference approach gradually breaks down as the distortion takes place. The reason behind this phenomenon is that two determinants become comparable in weight in the Full-CI expansion. From this perspective, the middle points E and F are the most challenging, exhibiting strongly correlated characteristics illustrated by the increasing distance between the benchmark Full-CI curve and the single reference curves (HF, SR-MP2, SR-PAPT2). A two-geminal based GVB calculation can correct the shape of the single reference curves with respect to Full-CI thanks to the description of static correlation. The fact that GVB is off from Full-CI by cca. 100 millihartree reflects the lack of dynamical correlation.

Evaluation of the MR-PAPT methods is presented in Fig. 5.5 showing that for the weakly correlated points (A-D and G-I) the performance of SR-PAPT2 is superior to the pMCPT version of MR-PAPT2 in running close to Full-CI, but the fMCPT version of MR-PAPT2 runs closest. With errors under 5 millihartree, the fMCPT-PAPT2

Figure 5.4: Total energies of BeH_2 in cc-pVDZ basis.Figure 5.5: Energies differences from Full-CI of BeH_2 in cc-pVDZ basis.

method exhibits balanced description, i.e. being more horizontal than fMCPT-MP2. In the strongly correlated points (E and F), pMCPT-PAPT2 outperforms pMCPT-MP2 but its overall balance is significantly worse which can be attributed to the enhanced pivot-dependence. The pivot-independent fMCPT-PAPT2, on the other hand, shows no deterioration even in these strongly correlated cases allowing for the most accurate and most balanced description in this test.

5.3 Size-consistency

An important feature, not necessarily reflected by quantum chemical methods of electronic structure is that extensive quantities should scale properly when extending the system. Terms size-consistency and size-extensivity [53], although being two distinct concepts, are often used interchangeably to describe that characteristic. Let us refer to the strict additive separability of the energy of an $A + B$ supersystem containing two non-interacting systems A and B , i.e.

$$E(A, B) = E(A) + E(B) \tag{5.1}$$

as size-consistency [54].

Theoretically, perturbative methods can provide size-consistent results with proper choice of zero-order and careful computational implementation [55]. Accordingly, in the single reference approach, both MP and PAPT are considered to be size-consistent. However, the MCPT method, that is used to formulate the multireference extension of MP and PAPT in this thesis, was shown to be slightly size-inconsistent [56].

To assess whether MR-PAPT formulated in the MCPT framework improves or deteriorates on inconsistency, tests on two distant, non-interacting LiH molecules was conducted. Using a correlation-consistent basis set, computational limitations allowed for frozen core calculations for the dimer and the monomer in double-zeta (cc-pVDZ) basis as presented in Table 5.1. Calculations without freezing the core molecular orbitals were performed for the same system in Pople’s 6-31G double-zeta basis set, presented in Table 5.2, but in this latter case again, time did not yet allow for the assessment of fMCPT results.

Inconsistency values for the CAS reference in Tables 5.1 and 5.2 reflect the numerical precision of the optimization procedure. This method is size-consistent with proper choice of active electrons and orbitals. In this case a CAS(2,2) model was adopted for the monomer and a CAS(4,4) was accordingly specified for the non-interacting dimer.

It is found that compared to pMCPT-MP2, the pMCPT-PAPT2 method roughly

Method	Monomer	Dimer	Error
	E_{mon}	E_{dim}	$ E_{\text{dim}} - 2E_{\text{mon}} $
	[E _h]	[E _h]	[mE _h]
CAS	-7.99221261	-15.98442524	$1.9 \cdot 10^{-5}$
pMCPT-MP2-SD	-8.00063141	-16.00072131	0.54
pMCPT-PAPT2-D	-8.00336470	-16.00565683	1.07
pMCPT-PAPT2-SD	-8.00336210	-16.00565163	1.07
fMCPT-MP2	-8.00087976	-16.00191073	0.15
fMCPT-PAPT2	-8.00338645	-16.00685815	0.09

Table 5.1: Size-consistency error test results on LiH monomer and non-interacting dimer in cc-pVDZ basis with frozen core calculations. Bond length of the monomer units is 2.00 Å.

Method	Monomer	Dimer	Error
	E_{mon}	E_{dim}	$ E_{\text{dim}} - 2E_{\text{mon}} $
	[E _h]	[E _h]	[mE _h]
CAS	-7.9892728	-15.9785459	$3.1 \cdot 10^{-4}$
pMCPT-MP2-SD	-7.9910346	-15.9819499	0.12
pMCPT-PAPT2-D	-7.9914200	-15.9826178	0.22
pMCPT-PAPT2-SD	-7.9914537	-15.9825896	0.32

Table 5.2: Size-consistency error test results on LiH monomer and non-interacting dimer in 6-31G basis with frozen core calculations. Bond length of the monomer units is 2.00 Å.

doubles the inconsistency if single excitations are excluded, i.e. no occupied-virtual matrix elements are allowed in $\hat{\Lambda}$. In case of the correlation consistent basis, where the core orbitals are frozen during the test, when including single excitations, inconsistency was not observed to grow. However, in the case of the 6-31G basis set, relatively large core-virtual matrix elements of $\hat{\Lambda}$ roughly triple the inconsistency.

Results of both size-consistency tests are in agreement with the conclusion that the inconsistency of the MCPT framework is an order of magnitude smaller than the second-order correction in the pMCPT and two orders of magnitude smaller in the fMCPT framework. This error is not observed to be significantly enlarged by the multireference adaptations of PAPT. In fact, in Table 5.1 size inconsistency of the Møller-Plesset partitioning is reduced by PAPT in the fMCPT framework.

6. Conclusion and Future Directions

As it was previously observed [3], the Knowles Partitioning (PAPT) is a viable candidate for multireference extension. This statement is further supported by the results presented in Ch. 5 that illustrate the promising properties of SR-PAPT. In particular, SR-PAPT accounts for a significantly greater part of the dynamic correlation than the Møller-Plesset partitioning deteriorating only in strongly correlated scenarios, a direct consequence of the limitation of the single reference approach itself. To develop a multireference version of PAPT, the philosophy of SR-PAPT is found to be straightforwardly extendable in the projected Multiconfigurational PT (pMCPT) framework for an arbitrary multiconfigurational reference function. The working formulae of pMCPT-PAPT presented in Ch. 4 are devised in a way that allows to preserve the most prominent features of SR-PAPT in the multireference approach, e.g. the general superiority to the Møller-Plesset partitioning and the vanishing of the third-order energy ensured by self-consistency. However, the artefactual pivot-dependence introduced by the pMCPT framework causing severe losses in parallelity with Full-CI called for the development of a pivot-independent formulation. This was performed in the frame-based MCPT (fMCPT) framework taking one step away from the original structure of the working formulae with the help of an equivalent derivation of the partitioning technique. The computational cost when optimally coded, is expected to scale in accordance with the single reference version, with the sixth power of the number of basis functions also introducing a linear increase with the number of determinants in the CI expansion of the reference function in the fMCPT adaptation. Preliminary tests performed with a parallelized pMCPT and an fMCPT working code indicate the mitigation of the ill-effect of SR-PAPT in strongly correlated scenarios, capitalizing on the enhanced description of static correlation by the reference function while maintaining the advance from MR-MP in accounting for the dynamic correlation. Additionally, the inherent slight size-inconsistency is not observed to be enlarged. In conclusion, the newly developed fMCPT based multireference adaptation of PAPT is observed to satisfy the need to give a balanced description in the weakly and strongly correlated scenarios while maintaining high level of accuracy. Publication of the findings of the research is among future plans after comparison to other well-known MR-PT methods.

Bibliography

- [1] Peter J. Knowles. Perturbation-adapted perturbation theory. *The Journal of Chemical Physics*, 156:011101, 1 2022.
- [2] András Gombás, Péter R Surján, and Ágnes Szabados. Analysis and Assessment of Knowles’ Partitioning in Many-Body Perturbation Theory. *Journal of Chemical Theory and Computation*, 2024.
- [3] András Gombás. *Analysis of the Knowles-partitioning in Many-Body Perturbation Theory*. Bachelor’s thesis, Eötvös Loránd University, 2023. (Available at: http://coulson.chem.elte.hu/gombas/PAPT_Thesis.pdf).
- [4] András Gombás. *Development of Multiconfigurational Perturbation-adapted Perturbation Theory*. TDK Thesis, 2023. (Available at: http://coulson.chem.elte.hu/gombas/MRPAPT_TDK_Thesis.pdf).
- [5] Ágnes Szabados, András Gombás, and Péter R Surján. Knowles Partitioning at the Multireference Level. *The Journal of Physical Chemistry A*, 2024.
- [6] Á Szabados, Á Margócsy, and PR Surján. Pivot invariance of multiconfiguration perturbation theory via frame vectors. *The Journal of Chemical Physics*, 157(17), 2022.
- [7] T. Helgaker, P. Jørgensen, and J. Olsen. *Molecular Electronic-Structure Theory*. John Wiley & Sons Ltd, England, 2000.
- [8] I. Shavitt and R. J. Bartlett. *Many-Body Methods in Chemistry and Physics*. Cambridge University Press, Cambridge, 2009.
- [9] I. Mayer. *Simple Theorems, Proofs, and Derivations in Quantum Chemistry*. Kluwer, New York, 2003.
- [10] Michael Baer. *Beyond Born-Oppenheimer: electronic nonadiabatic coupling terms and conical intersections*. John Wiley & Sons, 2006.

- [11] Lionel Salem, Claude Leforestier, Gerald Segal, and Ross Wetmore. Avoided surface crossings. *Journal of the American Chemical Society*, 97(3):479–487, Feb 1975.
- [12] D. R. Hartree. *Proc. Camb. Phil. Soc.*, 24:89, 1928.
- [13] J. C. Slater. The Theory of Complex Spectra. *Physical Review*, 34:1293–1322, 11 1929.
- [14] Alexandre L. Magalhães. Gaussian-Type Orbitals versus Slater-Type Orbitals: A Comparison. *Journal of Chemical Education*, 91(12):2124–2127, Dec 2014.
- [15] Tosio Kato. On the eigenfunctions of many-particle systems in quantum mechanics. *Communications on Pure and Applied Mathematics*, 10(2):151–177, 1957.
- [16] Samuel Francis Boys and Maurice Vincent Wilkes. The integral formulae for the variational solution of the molecular many-electron wave equation in terms of Gaussian functions with direct electronic correlation. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 258(1294):402–411, 1960.
- [17] Peter M. W. Gill, Benny G. Johnson, and John A. Pople. Two-electron repulsion integrals over Gaussian s functions. *International Journal of Quantum Chemistry*, 40(6):745–752, 1991.
- [18] R. A. Poirier, R. Kari, and I. G. Csizmadia. *Handbook of gaussian basis sets*. Elsevier, Amsterdam, 1985.
- [19] V. Fock. *Z. Physik*, 61:126, 1930.
- [20] J. C. Slater. 35:210, 1930.
- [21] L. Brillouin. *Actualités Sci. et ind.*, No. 71, No. 159, 1933.
- [22] Thomas M. Henderson, Guo P. Chen, and Gustavo E. Scuseria. Strong–weak duality via Jordan–Wigner transformation: Using fermionic methods for strongly correlated su(2) spin systems. *The Journal of Chemical Physics*, 157(19):194114, 11 2022.
- [23] T. Kato. *Perturbation Theory for Linear Operators*. Springer, Berlin, 1966.
- [24] Chr. Møller and M. S. Plesset. Note on an Approximation Treatment for Many-Electron Systems. *Physical Review*, 46:618–622, 10 1934.
- [25] Matthew L. Leininger, Wesley D. Allen, Henry F. Schaefer, and C. David Sherrill. Is Møller-Plesset perturbation theory a convergent ab initio method? *The Journal of Chemical Physics*, 112(21):9213–9222, June 2000.

- [26] Péter R. Surján, Dóra Köhalmi, and Ágnes Szabados. A note on perturbation-adapted perturbation theory. *The Journal of Chemical Physics*, 156:116102, 3 2022.
- [27] A. T. Amos. Feenberg–Goldhammer Procedure and Geometric Approximation in Hartree–Fock Perturbation Theory. *The Journal of Chemical Physics*, 52(2):603–605, 09 2003.
- [28] A. C. Hurley, J. Lennard-Jones, and J. A. Pople. A220:446, 1953.
- [29] A. C. Hurley. *Electron correlation in small molecules*. Academic, New York, 1976.
- [30] Joe M. Parks and Robert G. Parr. Theory of Separated Electron Pairs. *The Journal of Chemical Physics*, 28(2):335–345, 08 2004.
- [31] P. R. Surján. *Topics in current chemistry*, 203:63–88, 1999.
- [32] P.R. Surján. Geminal approach. *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering, Theoretical and Computational Chemistry*, 137, 2016.
- [33] Tadashi Arai. Theorem on Separability of Electron Pairs. *The Journal of Chemical Physics*, 33(1):95–98, 08 2004.
- [34] Frank W Bobrowicz and William A Goddard III. The self-consistent field equations for generalized valence bond and open-shell hartree—fock wave functions. In *Methods of electronic structure theory*, pages 79–127. Springer, 1977.
- [35] Björn O. Roos, Peter R. Taylor, and Per E.M. Sigbahn. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chemical Physics*, 48(2):157–173, 1980.
- [36] Kerstin Andersson, Per-Åke Malmqvist, and Björn O. Roos. Second-order perturbation theory with a complete active space self-consistent field reference function. *The Journal of Chemical Physics*, 96(2):1218–1226, 01 1992.
- [37] R. McWeeny. *Methods of Molecular Quantum Mechanics*. Academic, London, 1989.
- [38] Giovanni Ghigo, Björn O. Roos, and Per Åke Malmqvist. A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chemical Physics Letters*, 396(1):142–149, 2004.
- [39] Kerstin. Andersson, Per Aake. Malmqvist, Bjoern O. Roos, Andrzej J. Sadlej, and Krzysztof. Wolinski. Second-order perturbation theory with a CASSCF reference function. *The Journal of Physical Chemistry*, 94(14):5483–5488, Jul 1990.

- [40] Peter Pulay. A perspective on the CASPT2 method. *International Journal of Quantum Chemistry*, 111(13):3273–3279, 2011.
- [41] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu. Introduction of n-electron valence states for multireference perturbation theory. *The Journal of Chemical Physics*, 114(23):10252–10264, 06 2001.
- [42] Frank Neese. The ORCA program system. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(1):73–78, 2012.
- [43] Hans-Joachim Werner, Peter J Knowles, Frederick R Manby, Joshua A Black, Klaus Doll, Andreas Heßelmann, Daniel Kats, Andreas Köhn, Tatiana Korona, David A Kreplin, et al. The Molpro quantum chemistry package. *The Journal of chemical physics*, 152(14), 2020.
- [44] Mark R Hoffmann, Dipayan Datta, Sanghamitra Das, Debashis Mukherjee, Agnes Szabados, Zoltán Rolik, and Péter R Surján. Comparative study of multireference perturbative theories for ground and excited states. *The Journal of chemical physics*, 131(20), 2009.
- [45] Edina Rosta and Péter R Surján. Two-body zeroth order Hamiltonians in multireference perturbation theory: The APSG reference state. *The Journal of chemical physics*, 116(3):878–890, 2002.
- [46] K Hirao. State-specific multireference Møller—Plesset perturbation treatment for singlet and triplet excited states, ionized states and electron attached states of H₂O. *Chemical physics letters*, 201(1-4):59–66, 1993.
- [47] Ágnes Szabados. *Perturbációszámítás alapú módszerek molekulák elektronszerkezetének leírására*. PhD thesis, ELTE Eötvös Loránd Tudományegyetem, 2020.
- [48] Rajat K Chaudhuri, Karl F Freed, Gabriel Hose, Piotr Piecuch, Karol Kowalski, Marta Włoch, Sudip Chattopadhyay, Debashis Mukherjee, Zoltán Rolik, Ágnes Szabados, et al. Comparison of low-order multireference many-body perturbation theories. *The Journal of chemical physics*, 122(13), 2005.
- [49] E. Kapuy. On the possibility of deriving equivalent localized basis sets from symmetry adapted ones. *Acta Physica Academiae Scientiarum Hungaricae*, 51:21–24, 5 1981.
- [50] M. Kobayashi and Á. Szabados and H. Nakai and P. R. Surján. *J. Chem. Theory. Comput.*, 6:2024, 2010.
- [51] RK Nesbet. Brueckner’s theory and the method of superposition of configurations. *Physical Review*, 109(5):1632, 1958.

- [52] George D. Purvis, Ron Shepard, Franklin B. Brown, and Rodney J. Bartlett. C2V Insertion pathway for BeH₂: A test problem for the coupled-cluster single and double excitation model. *International Journal of Quantum Chemistry*, 23:835–845, 3 1983.
- [53] Rodney J Bartlett. Many-body perturbation theory and coupled cluster theory for electron correlation in molecules. *Annual review of physical chemistry*, 32(1):359–401, 1981.
- [54] Peter R Taylor. Coupled-cluster methods in quantum chemistry. In *Lecture Notes in Quantum Chemistry II: European Summer School in Quantum Chemistry*, pages 125–202. Springer, 1994.
- [55] Poul Jørgensen. *Second quantization-based methods in quantum chemistry*. Elsevier, 2012.
- [56] Agnes Szabados, Zoltán Rolik, Gábor Tóth, and Péter R Surján. Multiconfiguration perturbation theory: Size consistency at second order. *The Journal of chemical physics*, 122(11), 2005.
- [57] Ernest R. Davidson. Selection of the Proper Canonical Roothaan-Hartree-Fock Orbitals for Particular Applications. I. Theory. *The Journal of Chemical Physics*, 57:1999–2005, 9 1972.
- [58] William Gropp and Ewing Lusk. User’s Guide for mpich, a Portable Implementation of MPI, 1996.

Appendix

A The Davidson technique for repartitioning

To enhance the efficiency of perturbative procedures, one can reconsider the partitioning. Let us first investigate this question out of the scope of perturbation theory, and consider simply the partitioning of the electronic Hamiltonian to one-electron part and a two-electron part as

$$\hat{H} = \sum_i^n \hat{h}(i) + \sum_{i<j}^n \hat{V}(i, j). \quad (1)$$

In second-quantized formalism, Eq. (1) corresponds to

$$\hat{H} = \sum_{i,j} \sum_{\sigma} h_{ij} i_{\sigma}^{+} j_{\sigma}^{-} + \frac{1}{2} \sum_{i,j,k,l} \sum_{\sigma,\sigma'} [ik|jl] i_{\sigma}^{+} k_{\sigma'}^{+} l_{\sigma'}^{-} j_{\sigma}^{-} \quad (2)$$

with one- and two-electron integrals depending on the basis.

Davidson pointed out [57] that the canonical molecular orbitals, which diagonalize the matrix representation of the Fockian (see. Eq. (1.12)), “are still arbitrary”. Since to define a transformed Fockian, one may change the partitioning of Eq. (1) by adding an arbitrary $\hat{\alpha}$ one-electron, preferably hermitian operator to $\hat{h}$ and correct for this in the original two-electron term according to

$$\hat{H} = \sum_i^n \underbrace{[\hat{h}(i) + (n-1) \cdot \hat{\alpha}(i)]}_{\hat{\bar{h}}(i)} + \sum_{i<j}^n \underbrace{[\hat{V}(i, j) - \hat{\alpha}(i) - \hat{\alpha}(j)]}_{\hat{\bar{V}}(i,j)}. \quad (3)$$

The transformed Fockian stemming from Eq. (3) is constructed with dressed integrals

$$\bar{F}_{ij} = \underbrace{h_{ij} + A_{ij}}_{\bar{h}_{ij}} + \sum_{k,l} P_{kl} \left[\underbrace{\left([ik|jl] - \frac{A_{ij}\delta_{kl} + A_{kl}\delta_{ij}}{n-1} \right)}_{\overline{[ik|jl]}} - \frac{1}{2} \underbrace{\left([ij|lk] - \frac{A_{il}\delta_{jk} + A_{jk}\delta_{il}}{n-1} \right)}_{\overline{[ij|lk]}} \right], \quad (4)$$

where A_{ij} is the matrix representation of

$$\hat{A} = (n - 1) \cdot \hat{\alpha}. \quad (5)$$

This transformation leaves the density matrix, the Hartree-Fock energy and wavefunction unchanged as it generates non-zero elements in the occ-occ and virt-virt block of $\bar{F}$. Consequently, $\hat{A}$ defines a new set of orbitals that can be constructed by rotating the original canonical orbitals to obtain $\bar{\varphi}_i$ *proper canonicals*, on the basis of which $\bar{F}$ becomes diagonal.

When applying only the diagonal elements of A , the change affects only the ε_i orbital energies associated with F . Diagonals of A generate *level shifts* according to

$$\bar{\varepsilon}_a = \varepsilon_a + \frac{A_{aa} - 2 \cdot \text{Tr}_{\text{occ}}(A)}{n - 1} \quad (6)$$

for occupied and

$$\bar{\varepsilon}_p = \varepsilon_p - \frac{A_{pp} - 2 \cdot \text{Tr}_{\text{occ}}(A)}{n - 1} \quad (7)$$

for virtual levels.

The choice of $\hat{\alpha}$ and the underlying orbital rotations and level shifts may be guided by a specific interest. One may use it to get localized orbitals, or orbitals optimized for calculating ionization potentials or excitation energies. It is also possible to harness the flexibility provided by $\hat{\alpha}$ to enhance efficiency of many-body perturbation theory as it is shown in the upcoming Section.

B An alternative derivation of the PAPT equations

The interpretation of the physical meaning of the Knowles Partitioning might be guided by the following alternative derivation of PAPT equations based simply on the minimalization of a norm yet to be published.

Let us take functional J as

$$J = \sum_K^{2 \times \text{exc}} \left| \langle K | \hat{H}_N - \hat{\Lambda}_N | \Psi^{(1)} \rangle \right|^2, \quad (8)$$

where determinants K are doubly excited determinants with respect to the Hartree-Fock determinant

$$|K\rangle = |K_{ij}^{ab}\rangle = \{a^+ b^+ j^- i^-\} |0\rangle, \quad (9)$$

the first-order correction wavefunction $\Psi^{(1)}$ is also constructed of doubly excited determinants

$$|\Psi^{(1)}\rangle = \sum_{mn}^{\text{occ}} \sum_{cd}^{\text{virt}} c_{mn}^{cd(1)} |K_{mn}^{cd}\rangle \quad (10)$$

and the operators are

$$\hat{H}_N = \sum_{pq} f_{pq} \{p^+ q^-\} + \frac{1}{4} \sum_{pqrs} [pq||rs] \{p^+ q^+ s^- r^-\}, \quad (11)$$

$$\hat{\Lambda}_N = \frac{1}{2} \sum_{kl}^{\text{occ}} \Lambda_{kl} \left(\{k^+ l^-\} + \{l^+ k^-\} \right) + \frac{1}{2} \sum_{ef}^{\text{virt}} \Lambda_{ef} \left(\{e^+ f^-\} + \{f^+ e^-\} \right). \quad (12)$$

To minimize functional J , the following criteria should be satisfied:

$$\frac{\partial J}{\partial \Lambda_{kl}} = 0 \quad \forall k, l \in \{\text{occ}\} \quad (13)$$

and

$$\frac{\partial J}{\partial \Lambda_{ef}} = 0 \quad \forall e, f \in \{\text{virt}\}. \quad (14)$$

Starting with Eq. (13), the evaluation of the differentiation using the chain rule yields

$$\frac{\partial J}{\partial \Lambda_{kl}} = -\frac{1}{4} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \langle K_{ij}^{ab} | \hat{H}_N - \hat{\Lambda}_N | \Psi^{(1)} \rangle \langle K_{ij}^{ab} | \{k^+ l^-\} + \{l^+ k^-\} | \Psi^{(1)} \rangle. \quad (15)$$

To derive the factor of the middle expression already present in the definition of J in Eq. (8), let us investigate the functional A reading

$$A = -\frac{1}{4} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \langle K_{ij}^{ab} | \{k^+ l^-\} | \Psi^{(1)} \rangle \quad (16)$$

$$= -\frac{1}{4} \sum_{ijmn}^{\text{occ}} \sum_{abcd}^{\text{virt}} c_{mn}^{cd(1)} \langle 0 | \{i^+ j^+ b^- a^-\} \{k^+ l^-\} \{c^+ d^+ n^- m^-\} | 0 \rangle, \quad (17)$$

upon which applying the Generalized Wick's theorem for the expectation value of the product of normal-ordered operators, it simply yields

$$A = c_{ki}^{ab(1)} \delta_{jl}. \quad (18)$$

Following an analogous derivation for the $\{l^+ k^-\}$ -containing term, the final form of Eq. (15) reads

$$\begin{aligned} \frac{\partial J}{\partial \Lambda_{kl}} &= \sum_i^{\text{occ}} \sum_{ab}^{\text{virt}} c_{ki}^{ab(1)} \langle K_{il}^{ab} | \hat{H}_N - \hat{\Lambda}_N | \Psi^{(1)} \rangle \\ &\quad + \sum_i^{\text{occ}} \sum_{ab}^{\text{virt}} c_{li}^{ab(1)} \langle K_{ik}^{ab} | \hat{H}_N - \hat{\Lambda}_N | \Psi^{(1)} \rangle \end{aligned} \quad (19)$$

corresponding to the LHS of PAPT Equations

$$\langle \Theta_{kl}^+ | \hat{H}_N - \hat{\Lambda}_N | \Psi^{(1)} \rangle = 0 \quad \forall k, l \in \{\text{occ}\}. \quad (20)$$

The same procedure can be applied to Eq. (14) with an analogous conclusion corresponding to the LHS of PAPT Equations

$$\langle \Theta_{ef}^+ | \hat{H}_N - \hat{\Lambda}_N | \Psi^{(1)} \rangle = 0 \quad \forall e, f \in \{\text{virt}\}. \quad (21)$$

C Intermediars in the SR-PAPT working code

Computation time is reduced, if in the cycles constructing the normal-ordered version of the

$$Y_{kl}^\Lambda = \langle \tilde{\Theta}_{kl} | \hat{H} | \Psi^{(1)} \rangle \quad (1)$$

$\mathbf{Y}$ inhomogenous vector (considering kl as a hyperindex) to

$$Y_{kl}^{\Lambda_N} = \langle \tilde{\Theta}_{kl} | [\hat{H} - \langle 0 | \hat{H} | 0 \rangle] | \Psi^{(1)} \rangle \quad (2)$$

of the single-reference PAPT equations, suitable intermediars are found, that reduce the overall $(n_{\text{basis}})^7$ cost to $(n_{\text{basis}})^6$.

As for the inhomogenous part with occupied indices

$$\begin{aligned} Y_{ab}^{\Lambda_N} = & -8 \sum_{c,d,p,q}^{\prime} F_{dc} c_{pq}^{ac} \bar{c}_{pq}^{bd} - 4 \sum_{c,d,p,q}^{\prime} F_{db} c_{pq}^{ac} \bar{c}_{pq}^{dc} - 4 \sum_{c,d,p,q}^{\prime} F_{da} c_{pq}^{bc} \bar{c}_{pq}^{dc} \\ & + 4 \sum_{c,p,q,r}^{\prime} F_{pr} \left(c_{rq}^{cb} \bar{c}_{qp}^{ac} + c_{qr}^{cb} \bar{c}_{pq}^{ac} + c_{rq}^{ca} \bar{c}_{qp}^{bc} + c_{qr}^{ca} \bar{c}_{pq}^{bc} \right) \\ & + 4 \sum_{c,d,p,q,r}^{\prime} [pd|cr] \left(\bar{c}_{qp}^{ac} \bar{c}_{qr}^{bd} + \bar{c}_{qp}^{bc} \bar{c}_{qr}^{ad} \right) \\ & - 4 \sum_{c,d,p,q,r}^{\prime} [pd|rc] \left(c_{qp}^{ac} \bar{c}_{qr}^{bd} + c_{pq}^{ac} \bar{c}_{rq}^{bd} + c_{qp}^{bc} \bar{c}_{qr}^{ad} + c_{pq}^{bc} \bar{c}_{rq}^{ad} \right) \\ & + 4 \sum_{c,d,p,q,r}^{\prime} [pd|br] \bar{c}_{pq}^{ac} \bar{c}_{qr}^{cd} - 4 \sum_{c,d,p,q,r}^{\prime} [pd|rb] \left(c_{pq}^{ac} \bar{c}_{qr}^{cd} + c_{qp}^{ac} \bar{c}_{rq}^{cd} \right) \\ & + 4 \sum_{c,d,p,q,r}^{\prime} [pd|ar] \bar{c}_{pq}^{bc} \bar{c}_{qr}^{cd} - 4 \sum_{c,d,p,q,r}^{\prime} [pd|ra] \left(c_{pq}^{bc} \bar{c}_{qr}^{cd} + c_{qp}^{bc} \bar{c}_{rq}^{cd} \right) \\ & + 4 \sum_{d,e,c,p,q}^{\prime} [de||bc] c_{pq}^{de} c_{pq}^{ac} + 4 \sum_{d,e,c,p,q}^{\prime} [de||ac] c_{pq}^{de} c_{pq}^{bc} + 8 \sum_{c,p,q,r,s}^{\prime} [pq||rs] c_{rs}^{bc} c_{pq}^{ac}, \quad (3) \end{aligned}$$

the following form can be introduced:

$$\begin{aligned}
 Y_{ab}^{\Lambda_N} = \sum_{c,p,q} \bigg[& -\text{IM1}(b, c, p, q) \, c_{pq}^{ac} \\
 & - \left[\text{IM2}(a, c, p, q) \, c_{pq}^{bc} + \text{IM2}(b, c, p, q) \, c_{pq}^{ac} \right] \\
 & + \left[\text{IM3}(a, c, p, q) \, \bar{c}_{pq}^{bc} + \text{IM3}(b, c, p, q) \, \bar{c}_{pq}^{ac} \right] \\
 & + \left[\text{IM4}(a, c, p, q) \, \bar{c}_{qp}^{bc} + \text{IM4}(b, c, p, q) \, \bar{c}_{qp}^{ac} \right] \\
 & + \left[\text{IM5}(a, c, p, q) \, \bar{c}_{qp}^{bc} + \text{IM5}(b, c, p, q) \, \bar{c}_{qp}^{ac} \right] \\
 & - \left[\text{IM6}(a, c, p, q) \, c_{qp}^{bc} + \text{IM6}(b, c, p, q) \, c_{qp}^{ac} \right] \\
 & - \left[\text{IM7}(a, c, p, q) \, c_{pq}^{bc} + \text{IM7}(b, c, p, q) \, c_{pq}^{ac} \right] \\
 & + \left[\text{IM8}(a, c, p, q) \, \bar{c}_{pq}^{bc} + \text{IM8}(b, c, p, q) \, \bar{c}_{pq}^{ac} \right] \\
 & - \left[\text{IM9}(a, c, p, q) \, c_{pq}^{bc} + \text{IM9}(b, c, p, q) \, c_{pq}^{ac} \right] \\
 & - \left[\text{IM10}(a, c, p, q) \, c_{qp}^{bc} + \text{IM10}(b, c, p, q) \, c_{qp}^{ac} \right] \\
 & + \left[\text{IM11}(a, c, p, q) \, c_{pq}^{bc} + \text{IM11}(b, c, p, q) \, c_{pq}^{ac} \right] \\
 & + \text{IM12}(b, c, p, q) \, c_{pq}^{ac} \bigg], \tag{4}
 \end{aligned}$$

where the intermediers are

$$\begin{aligned}
 \text{IM1}(b, c, p, q) &= 8 \sum_d' F_{dc} \, \bar{c}_{pq}^{bd} \\
 \text{IM2}(a, c, p, q) &= 4 \sum_d' F_{da} \, \bar{c}_{pq}^{dc} \\
 \text{IM3}(a, c, p, q) &= 4 \sum_r' F_{pr} \, c_{rq}^{ac} \\
 \text{IM4}(a, c, p, q) &= 4 \sum_r' F_{pr} \, c_{qr}^{ac} \\
 \text{IM5}(a, c, p, q) &= 4 \sum_{d,r}' [pd|cr] \, \bar{c}_{qr}^{ad} \\
 \text{IM6}(a, c, p, q) &= 4 \sum_{d,r}' [pd|rc] \, \bar{c}_{qr}^{ad}
 \end{aligned}$$

$$\begin{aligned}
\text{IM7}(a, c, p, q) &= 4 \sum_{d,r}^{\prime} [pd|rc] \bar{c}_{rq}^{ad} \\
\text{IM8}(a, c, p, q) &= 4 \sum_{d,r}^{\prime} [pd|ar] \bar{c}_{qr}^{cd} \\
\text{IM9}(a, c, p, q) &= 4 \sum_{d,r}^{\prime} [pd|ra] \bar{c}_{qr}^{cd} \\
\text{IM10}(a, c, p, q) &= 4 \sum_{d,r}^{\prime} [pd|ra] \bar{c}_{rq}^{cd} \\
\text{IM11}(a, c, p, q) &= 4 \sum_{d,e}^{\prime} [de||ac] c_{pq}^{de} \\
\text{IM12}(a, c, p, q) &= 8 \sum_{r,s}^{\prime} [pq||rs] c_{rs}^{ac}, \tag{5}
\end{aligned}$$

respectively.

The same can be applied to the

$$\begin{aligned}
Y_{pq}^{\Lambda_N} &= +8 \sum_{a,b,r,s}^{\prime} F_{rs} c_{pr}^{ab} \bar{c}_{qs}^{ab} + 4 \sum_{a,b,r,s}^{\prime} F_{qs} c_{pr}^{ab} \bar{c}_{sr}^{ab} + 4 \sum_{a,b,r,s}^{\prime} F_{ps} c_{qr}^{ab} \bar{c}_{sr}^{ab} \\
&\quad - 4 \sum_{a,b,c,r}^{\prime} F_{ca} \left(c_{pr}^{ba} \bar{c}_{rq}^{cb} + c_{pr}^{ab} \bar{c}_{rq}^{bc} + c_{qr}^{ba} \bar{c}_{rp}^{cb} + c_{qr}^{ab} \bar{c}_{rp}^{bc} \right) \\
&\quad + 4 \sum_{a,b,c,r,s}^{\prime} [rc|as] \left(\bar{c}_{qs}^{bc} \bar{c}_{pr}^{ba} + \bar{c}_{ps}^{bc} \bar{c}_{qr}^{ba} \right) \\
&\quad - 4 \sum_{a,b,c,r,s}^{\prime} [rc|sa] \left(c_{pr}^{ba} \bar{c}_{qs}^{bc} + c_{rp}^{ba} \bar{c}_{sq}^{bc} + c_{qr}^{ba} \bar{c}_{ps}^{bc} + c_{rq}^{ba} \bar{c}_{sp}^{bc} \right) \\
&\quad + 4 \sum_{a,b,c,r,s}^{\prime} [qc|bs] \bar{c}_{rp}^{ab} \bar{c}_{rs}^{ac} - 4 \sum_{a,b,c,r,s}^{\prime} [qc|sb] \left(c_{pr}^{ab} \bar{c}_{sr}^{ac} + c_{rp}^{ab} \bar{c}_{rs}^{ac} \right) \\
&\quad + 4 \sum_{a,b,c,r,ss}^{\prime} [pc|bs] \bar{c}_{rq}^{ab} \bar{c}_{rs}^{ac} - 4 \sum_{a,b,c,r,s}^{\prime} [pc|sb] \left(c_{qr}^{ab} \bar{c}_{sr}^{ac} + c_{rq}^{ab} \bar{c}_{rs}^{ac} \right) \\
&\quad + 4 \sum_{a,b,r,s,t}^{\prime} [qr||st] c_{st}^{ab} c_{pr}^{ab} + 4 \sum_{a,b,r,s,t}^{\prime} [pr||st] c_{st}^{ab} c_{qr}^{ab} + 8 \sum_{a,b,c,d,r}^{\prime} [cd||ba] c_{qr}^{cd} c_{pr}^{ba} \tag{6}
\end{aligned}$$

part with virtual indices in a similar manner to obtain:

$$\begin{aligned}
 Y_{pq}^{\Lambda_N} = \sum_{a,b,r} \Bigg[& + \text{IM1}(a, b, q, r) \, c_{pr}^{ab} \\
 & + \left[\text{IM2}(a, b, p, r) \, c_{qr}^{ab} + \text{IM2}(a, b, q, r) \, c_{pr}^{ab} \right] \\
 & - \left[\text{IM3}(a, b, p, r) \, c_{rq}^{ab} + \text{IM3}(a, b, q, r) \, c_{rp}^{ab} \right] \\
 & - \left[\text{IM4}(a, b, p, r) \, c_{qr}^{ab} + \text{IM4}(a, b, q, r) \, c_{pr}^{ab} \right] \\
 & + \left[\text{IM5}(a, b, p, r) \, \bar{c}_{qr}^{ba} + \text{IM5}(a, b, q, r) \, \bar{c}_{pr}^{ba} \right] \\
 & - \left[\text{IM6}(a, b, p, r) \, c_{qr}^{ba} + \text{IM6}(a, b, q, r) \, c_{pr}^{ba} \right] \\
 & - \left[\text{IM7}(a, b, p, r) \, c_{rq}^{ba} + \text{IM7}(a, b, q, r) \, c_{rp}^{ba} \right] \\
 & + \left[\text{IM8}(a, b, p, r) \, \bar{c}_{rq}^{ab} + \text{IM8}(a, b, q, r) \, \bar{c}_{rp}^{ab} \right] \\
 & - \left[\text{IM9}(a, b, p, r) \, c_{rq}^{ab} + \text{IM9}(a, b, q, r) \, c_{rp}^{ab} \right] \\
 & - \left[\text{IM10}(a, b, p, r) \, c_{qr}^{ab} + \text{IM10}(a, b, q, r) \, c_{pr}^{ab} \right] \\
 & + \left[\text{IM11}(a, b, p, r) \, c_{qr}^{ab} + \text{IM10}(a, b, q, r) \, c_{pr}^{ab} \right] \\
 & + \text{IM12}(a, b, q, r) \, c_{pr}^{ba} \Bigg], \tag{7}
 \end{aligned}$$

where the intermediers are

$$\begin{aligned}
 \text{IM1}(a, b, q, r) &= 8 \sum_{s, \substack{\prime \\ s}} F_{rs} \, \bar{c}_{qs}^{ab} \\
 \text{IM2}(a, b, p, r) &= 4 \sum_{s, \substack{\prime \\ s}} F_{ps} \, \bar{c}_{sr}^{ab} \\
 \text{IM3}(a, b, p, r) &= 4 \sum_{c, \substack{\prime \\ c}} F_{ca} \, \bar{c}_{pr}^{bc} \\
 \text{IM4}(a, b, p, r) &= 4 \sum_{c, \substack{\prime \\ c}} F_{ca} \, \bar{c}_{rp}^{bc} \\
 \text{IM5}(a, b, p, r) &= 4 \sum_{c, s, \substack{\prime \\ c, s}} [rc|as] \, \bar{c}_{ps}^{bc} \\
 \text{IM6}(a, b, p, r) &= 4 \sum_{c, s, \substack{\prime \\ c, s}} [rc|sa] \, \bar{c}_{ps}^{bc}
 \end{aligned}$$

$$\begin{aligned}\text{IM7}(a, b, p, r) &= 4 \sum_{c,s}^I [rc|sa] \bar{c}_{sp}^{bc} \\ \text{IM8}(a, b, p, r) &= 4 \sum_{c,s}^I [pc|bs] \bar{c}_{rs}^{ac} \\ \text{IM9}(a, b, p, r) &= 4 \sum_{c,s}^I [pc|sb] \bar{c}_{rs}^{ac} \\ \text{IM10}(a, b, p, r) &= 4 \sum_{c,s}^I [pc|sb] \bar{c}_{sr}^{ac} \\ \text{IM11}(a, b, p, r) &= 4 \sum_{s,t}^I [pr||st] c_{st}^{ab} \\ \text{IM12}(a, b, q, r) &= 8 \sum_{c,d}^I [cd||ba] c_{qr}^{cd},\end{aligned}\tag{8}$$

correspondingly.

This optimization enabled us to run tests in larger basis sets for publication. For comparison, presented are here two such examples.

```
-----  
BeH2 Geometry Point D in cc-pVTZ basis  
-----  
Y without intermediers (~Nbasis^7)  
real    3m4.131s  
user    3m2.308s  
sys     0m1.185s  
Y with intermediers (~Nbasis^6)  
real    0m53.093s  
user    0m51.697s  
sys     0m1.110s
```

Figure 1: Computational times of single-reference BeH₂ molecule in cc-pVTZ basis.

CH4 Tetrahedral in cc-pVTZ basis	

Y without intermediers (~Nbasis^7)	
real	59m19.372s
user	59m12.227s
sys	0m1.082s
Y with intermediers (~Nbasis^6)	
real	6m53.949s
user	6m52.134s
sys	0m1.086s

Figure 2: Computational times of single-reference CH₄ molecule in cc-pVTZ basis.

D Reciprocal vectors in the pMCPT formulae

The overlap of the projected determinants in pMCPT method yields

$$S_{0K} = \langle 0|K' \rangle = \underbrace{\langle 0|K \rangle}_0 - c_K \langle 0|\Phi \rangle = - \sum_L c_K c_L \langle 0|L \rangle = 0 \quad \forall K \neq 0, \quad (1)$$

but

$$S_{KL} = \langle K'|L' \rangle = \underbrace{\langle K|L \rangle}_{\delta_{KL}} - c_L \underbrace{\langle K|\Phi \rangle}_{c_K} - c_K \underbrace{\langle \Phi|L \rangle}_{c_L} + c_K c_L \underbrace{\langle \Phi|\Phi \rangle}_1 = \delta_{KL} - c_K c_L \quad \forall K, L \neq 0. \quad (2)$$

The choice of

$$\langle \widetilde{K}' | = \langle K | - \frac{c_K}{c_0} \langle 0 |, \quad (3)$$

that corresponds to

$$\begin{aligned} \langle \widetilde{K}' | &= \sum_{J \neq 0} (S^{-1})_{JK} \langle J' | \\ &= \langle K' | - \frac{c_K}{c_0^2} \sum_{J \neq 0} c_J \langle J' | \end{aligned} \quad (4)$$

being the reciprocal vectors used in the working formulae of the pMCPT method can, therefore, be justified by

$$\begin{aligned}
 \langle \widetilde{K}' | L' \rangle &= \left(\langle K | - \frac{c_K}{c_0} \langle 0 | \right) \left(| L \rangle - c_L | \Phi \rangle \right) \\
 &= \underbrace{\langle K | L \rangle}_{\delta_{KL}} - \frac{c_K}{c_0} \underbrace{\langle 0 | L \rangle}_0 - \underbrace{\langle K | \Phi \rangle}_{c_K} c_L + \frac{c_K c_L}{c_0} \underbrace{\langle 0 | \Phi \rangle}_{c_0} \\
 &= \delta_{KL} - c_K c_L + c_K c_L = \delta_{KL}
 \end{aligned} \tag{5}$$

the fact that it ensures orthonormality.

E Parallelization of the pMCPT-PAPT working code

The working code that was written in FORTRAN90 language is a modular program system that calculates matrix elements of operators with determinants on the fly. On the one hand, this allows for the straightforward transcription of the developed working formulae into a demo code. This is advantageous as one does not have to deal with rather tedious derivations of matrix elements based on Wick's theorem. On the other hand, such a programming technique results in high computational cost, since instead of direct formulae expressed with quantities indexed by molecular orbitals, one has expressions in terms of determinants of the Hilbert space. The latter usually involve extra do loop when compared with the former. The ideal scaling of computation time with system size is achievable only with the former approach.

A standard FORTRAN90 code compiled with e.g. GNUFORTRAN can be considered serial as it runs on only one thread of the CPU and executes commands one by one, consecutively. A simple time-saving approach e.g. when one conducts a numerical test involving a few sterical configurations of a certain many-body test system, is to launch multiple serial calculations at once to utilize not just one but more threads of the CPU.

Another approach to achieve reduced computational time by parallelization is to use message-passing interface (MPI). This is helpful when a section of the serial code can be identified that does not require consecutive serial evaluation but instead can be separated into parts allowing evaluation on multiple threads simultaneously. The message-passing necessary in this case entails communication between the multiple threads of the CPU. Following this latter approach, the pMCPT based MR-PAPT working code was parallelized using the MPICH [58] implementation of MPI that is to be detailed in this Section.

After conducting a few preliminary tests of the serial working code, it was concluded that the construction of the $\mathbf{T}^A$ coefficient matrix for the MR-PAPT equations (see Eq.

(4.10)) is the slowest. It is one of the arrays used in the MR-PAPT procedure that would exhibit computational time scaling with the sixth power of the number of basis functions in the ideal case. Therefore, in order to be able to calculate results for simple test systems with chemically reasonable basis sets in feasible time, the parallelization of the routine that constructs $\mathbf{T}^\Lambda$ was accomplished by MPICH. Each row of the $\mathbf{T}^\Lambda$ coefficient matrix corresponds to the LHS of an equation with one particular $\langle \Theta_{ij}^+ |$ projection function, so distributing the calculation by rows is an adequate way of parallelization.

Fig. 3 illustrates the parallelized part of the MR-PAPT working code. One among the threads is usually distinguished, having the role of the "master". The master thread coordinates the others, classified as "slaves". Serial parts of the code preceding and following the parallelized part presented on Fig. 3, are evaluated only by the master thread. Slave processes are sent to the beginning of the MR-PAPT routine at the start of the program and are commanded to wait until the master, having calculated all the preceding parts, reaches the parallelized part as well.

Construction of the projection functions and the inhomogeneous vector of the PAPT-equations is also done only by the master thread. Following this, a communication between the threads, coordinated by the master thread takes place, with the aim of passing the necessary variables to the slave threads.

Since slave threads calculate all elements of the $\mathbf{T}^\Lambda$ coefficient matrix in a row, all the determinants within the first-order correction vector need to be available for all slave threads as well as the master thread. Therefore, at the beginning of the parallel execution part, the first-order correction wavefunction that serves as an input to the $\mathbf{T}^\Lambda$ construction routine has to be *broadcasted*, i.e. all slave threads need to make a copy of the arrays and variables obtained from the master thread to their own memory area for their own use.

As for the set of projection functions, each thread requires only that subset of the functions that correspond to the set of rows the given thread is commanded to calculate. Knowing the number of available threads and the number of rows in the coefficient matrix, the master thread *scatters* the total set of projection functions into subsets accordingly. Consequently, scatter-type inter-thread communication entails the distribution of the arrays and variables constituting the subsets of projection functions among all available threads.

When all threads have finished calculation of their batch of the rows of the coefficient matrix, the concluding message-passing task is to *gather* all rows to the memory area of the master thread. Computation then proceeds by calling the DGELSD LAPACK FORTRAN routine to solve the singular matrix equation and obtain the parameters of MR-PAPT.

To illustrate computational time savings accomplished by the parallelization of one

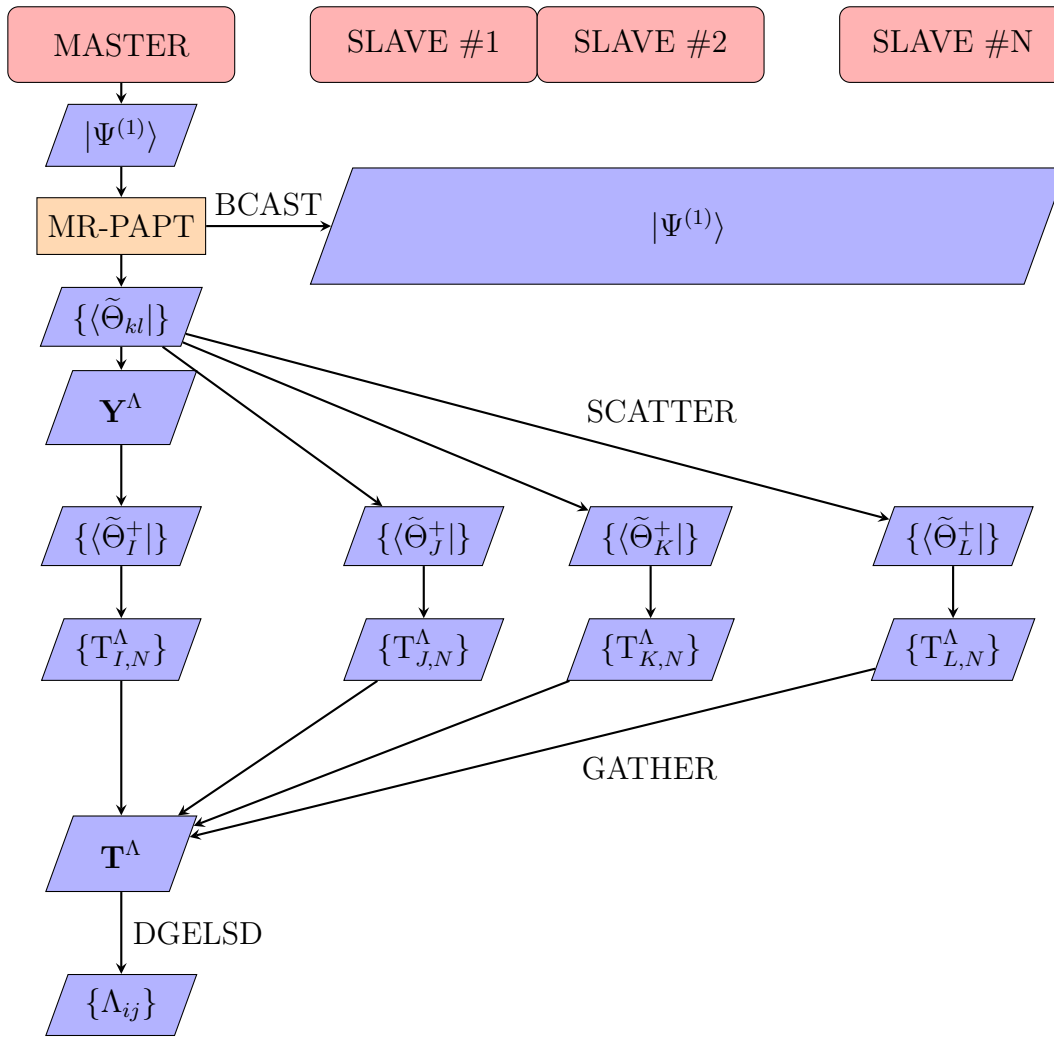


Figure 3: The workflow of the parallelized part of the MR-PAPT working code

particular section of the working code, Fig. 4 presents the decrease of computational time on a logarithmic scale by increasing the number of threads used for the calculations. The test presented was conducted on the ATLAS high-performance computer of ELTE that allows for a maximum number of 128 threads for calculating rows of the $\mathbf{T}^A$ matrix simultaneously.

In an ideal case, where the time spent on message-passing between the threads is negligible with respect to the total computational time, one expects a reciprocal dependence of the t computational time on the n_t number of threads as

$$t(n_t) = \frac{t(1)}{n_t}. \quad (6)$$

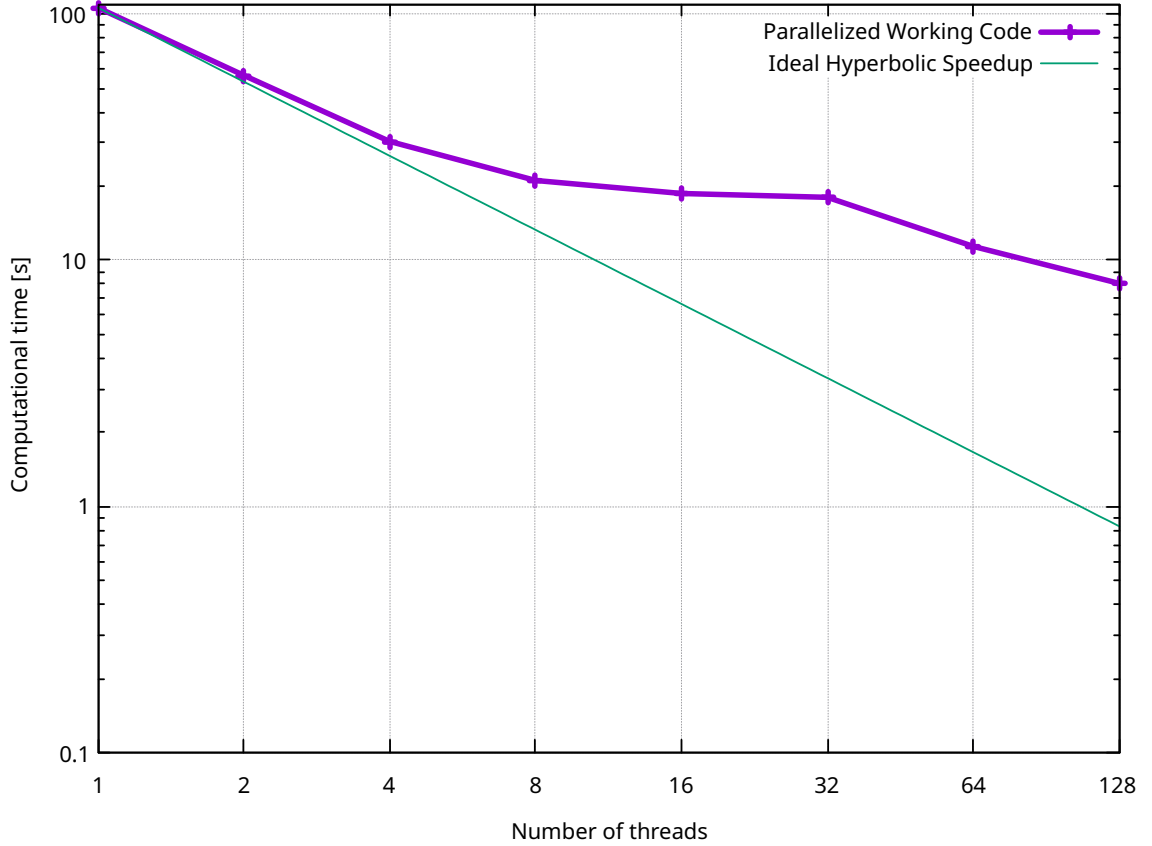


Figure 4: Comparison of ideal and actual computational times of the parallelized working code for a BeH_2 calculation in 6-311G basis using various number of threads.

It is correspondingly observed on Fig. 4 that computational time reduces quite significantly and in an ideal-like manner by increasing the number of threads from 1 to cca. 8 reassuring that parallelization is executed at an adequate section of the working code.