Linearized Coupled Cluster Corrections to Antisymmetrized Product of Strongly Orthogonal Geminals: role of dispersive interactions

Tamás Zoboki, Ágnes Szabados, and Péter R. Surján*

Institute of Chemistry, Laboratory of Theoretical Chemistry, H-1518 Budapest 112, P.O.B. 32

E-mail: surjan@chem.elte.hu

Abstract

A linearized multi-reference Coupled Cluster (MR-LCC) theory is formulated based on the Antisymmetrized Product of Strongly Orthogonal Geminals (APSG) reference state. The role of dispersive interbond interactions is discussed. The presented theory has lead to qualitatively correct potential curves for the case when both OH bonds dissociate in H_2O, a result that cannot be achieved by adding only perturbative corrections to APSG. The potential curve obtained for the He...He problem practically coincides with the full CI (FCI) result, showing the unexpected accuracy of the MR-LCC approach in this case.

1 Introduction

Reference states in quantum chemical calculations are used as zeroth order states obtained at a simpler level to be corrected by perturbative or more sophisticated theories. Coupled Cluster (CC)

---

*To whom correspondence should be addressed theory[1–6] for example offers a powerful tool to describe molecular electronic structure for single reference (SR) problems, i.e. when the electronic wave function can be qualitatively described by a single Slater-determinant. Many chemical processes require the use of two or more determinants even for a qualitative description. Just like Multi-Reference (MR) Perturbation Theory (PT),[7–15] the MR generalization of CC theory is a widely discussed problem, for which several ideas have been proposed, each having their advantages and disadvantages.[16–33]

A central problem in MR-CC theory is the construction of permitted excitations and their amplitudes by ensuring the same number of amplitude equations as the number of unknowns, preserving at the same time the commutation of excitation operators. The latter condition is essential to ensure extensivity manifested in the SR case by the ansatz

$$|\Psi\rangle = e^{\hat{T}}|\mathrm{HF}\rangle,$$

where $|\mathrm{HF}\rangle$ is the SR state (typically the Hartree-Fock determinant) and $\hat{T}$ is the excitation operator

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + ... = \sum_{i}^{occ}\sum_{p}^{virt} t_i^p p^+ i^- + \sum_{i<j}^{occ}\sum_{p<q}^{virt} t_{ij}^{pq} p^+ q^+ j^- i^- + ... \quad , \tag{1}$$

where $i, j, ...$ refer to occupied and $p, q, ...$ refer to virtual indices relative to the Fermi vacuum $|\mathrm{HF}\rangle$ and $\hat{T}_1, \hat{T}_2, ...$ describe single, double, ... excitations, respectively. Creation/annihilation indices categorized into disjoint subsets results in the important fact that terms in $\hat{T}_1, \hat{T}_2, ...$ commute, ensuring extensivity of CC corrections. The simple categorization is clearly impossible if the reference state consists of several determinants. Of possible workarounds we mention the Jeziorski-Monkhorst Ansatz[34]

$$|\Psi\rangle = \sum_{\mu} C_{\mu}\, e^{\hat{T}_{\mu}} |\Phi_{\mu}\rangle,$$

which deals with the MR reference state $|\Phi\rangle = \sum_{\mu} C_{\mu} |\Phi_{\mu}\rangle$, and cluster operators $\hat{T}_{\mu}$ are defined similar to eq.(1), the occupancies referring to that of $|\Phi_{\mu}\rangle$. Special attention has to be paid in these theories to eliminate possible redundancies in excitations and amplitudes.[35] One may obtain a theory free of the redundancy problem by picking up a dominant component $|\Phi_0\rangle$ in $|\Phi\rangle$ and define occupancies with respect to this pivot determinant.[36–38] The coupled-cluster correction developed specifically for the GVB wavefunction by Head-Gordon and co-workers also relies on the concept of the pivot.[39–41] The alternative approach of devising an internally contracted theory has received much interest lately.[42–45]

The MR problem leads to complications also in PT. In SR-PT the formalism remains simple since $|\mathrm{HF}\rangle$ and excited determinants altogether form eigenvectors of a one-body operator, the Fockian ($\hat{F}$). If the reference state $|\Phi\rangle$ is of multi-reference character, no one-body operator emerges as a zero-order Hamiltonian, whose eigenvectors could be easily constructed. To resolve this problem, Dyall[46] was the first to propose the use of two-body zero-order Hamiltonian. In our Laboratory,[47–51] we have developed a MRPT using this philosophy applied to the APSG[52–62] reference state.

A method "between" simple PT and the more sophisticated CC theory is Linearized Coupled Cluster (LCC) formulation.[63–69] It emerges from CC by dropping all nonlinear terms in the amplitude equations. It was shown that LCC equations can also be derived by optimizing the partition via level shifts in a PT framework.[70,71] This latter statement holds both in the SR and MR cases.[70–73]

The present paper outlines a linearized CC (LCC) theory based on the APSG reference state.

The formulation ensures extensivity, provides a uniquely soluble set of amplitude equations and is free from the deficiencies of the PT presented in Ref. 51 concerning description of simultaneous single bond dissociations. Our theory is related to that of Li et al.,[74,75] with notable differences. The present formulation relies strongly on the quasi-particle picture of geminals[47–50,76–85] and strictly maintains the size-consistency of the approach. Another difference is the way geminal subspaces are chosen: we start from localized molecular orbitals and optimize them variationally, while Li et al.[74,75] use the "maximum similarity rule" introduced in Ref. 74.

2 Theoretical Background

2.1 Geminals

A generalization of the one-determinant Hartree-Fock wave function is the Antisymmetrized Product of Strongly Orthogonal Geminals (APSG) wave function composed of two-electron orbitals ψ_i:

$$|\Psi\rangle = \psi_1^+ \psi_2^+ \cdots \psi_N^+ |\mathrm{vac}\rangle, \tag{2}$$

The two-electron operators (ψ_i^+) create *geminals*, and they are expanded in terms of one-electron operators corresponding to spin orbital χ_μ:

$$\psi_i^+ = \sum_{\mu<\nu} C^i_{\mu\nu} \chi_\mu^+ \chi_\nu^+ . \tag{3}$$

In this respect, expression (2) is a multi-determinantal wave function in terms of one-electron orbitals containing some proportion of the correlation effects. Because of the special structure of eqs.(2-3), this is called *intra-geminal correlation*.

The two-electron orbitals $\psi_i(1,2) = \psi_i^+|\mathrm{vac}\rangle$ are constrained to satisfy the strong orthogonality condition

$$\int \psi_i(1,2)\psi_k(1,2)dr_1 = 0 \qquad i \neq k, \tag{4}$$

which, according to Arai,[86] is equivalent to expanding the geminals in mutually exclusive subsets of one-electron orbitals, called Arai subspaces.

In APSG theory, both the geminal coefficients $C^i_{\mu\nu}$ and one-electron basis orbitals χ_μ are optimized variationally, the latter via an MCSCF-type iteration. The geminal coefficients come from solving the set of self-consistent equations arising from the variational principle:

$$\hat{H}_i^{\mathrm{eff}}\psi_i^p = E_i^p\psi_i^p, \tag{5}$$

where i is the geminal index, $p \in \{0,1,...,n_i\}$ the state index and $\hat{H}_i^{\mathrm{eff}}$ is an effective Hamiltonian defined over the spin orbital basis functions assigned to geminal i:

$$\hat{H}_i^{\mathrm{eff}} = \sum_{\mu,\nu\in i} h^{\mathrm{eff}}_{\mu\nu}\chi_\mu^+\chi_\nu^- + \frac{1}{2}\sum_{\mu,\nu,\lambda,\sigma\in i}[\mu\nu|\lambda\sigma]\chi_\mu^+\chi_\nu^+\chi_\sigma^-\chi_\lambda^-. \tag{6}$$

In the above formula $\mathbf{h}^{\mathrm{eff}}$ is the 'effective core' responsible for the inter-geminal electrostatic and exchange interaction:

$$h^{\mathrm{eff}}_{\mu\nu} = h_{\mu\nu} + \sum_{k\neq i}\sum_{\lambda,\sigma} P^{k0}_{\lambda\sigma}[\mu\lambda||\nu\sigma],$$

where $P^{k0}_{\lambda\sigma}$ is the density matrix of the ground-state geminal ψ_k^0:

$$P^{k0}_{\lambda\sigma} = \langle\psi_k^0|\chi_\sigma^+\chi_\lambda^-|\psi_k^0\rangle, \tag{7}$$

and $[\mu\lambda||\nu\sigma]$ is the antisymmetrized two-electron integral written in the $[12||12]$ convention. For more details, we refer to Ref. 61.

2.2 Perturbation theory based on geminals

While APSG theory includes only intra-geminal correlation effects, inter-geminal corrections can be introduced via perturbation theory.[47–51] The basic concepts of this formulation are summarized below. We start from the second quantized Hamiltonian written in terms of the energy-optimized spatial orbitals

$$\hat{H} = \sum_{m,n}\sum_{\sigma} h_{mn} m_{\sigma}^{+} n_{\sigma}^{-} + \frac{1}{2}\sum_{m,n,l,s}\sum_{\sigma,\sigma'} [mn|ls] m_{\sigma}^{+} n_{\sigma'}^{+} s_{\sigma'}^{-} l_{\sigma}^{-}, \tag{8}$$

where m,n,l,s stand for the spatial part and σ,σ' for the spin part of spin orbital χ_{μ}. Because of strong orthogonality, each basis orbital belongs to one of the mutually exclusive orbital subsets. So instead of the sum over basis orbitals, we can introduce first a summation over the subsets (geminals) i, then a summation over the orbitals m belonging to that specific subset:

$$\sum_{m} \equiv \sum_{i}\sum_{m\in i}$$

leading to a natural separation of the full Hamiltonian (8) in geminal indices:[47,48]

$$\hat{H} = \sum_{i}\hat{H}_i + \sum_{i,j}{}' \hat{H}_{ij} + \sum_{i,j,k}{}' \hat{H}_{ijk} + \sum_{i,j,k,l}{}' \hat{H}_{ijkl},$$

where the prime after the sums indicates that the summation indices cannot coincide. The one-geminal Hamiltonian will obviously be

$$\hat{H}_i = \sum_{m,n\in i} \sum_{\sigma} h_{mn} m_\sigma^+ n_\sigma^- + \frac{1}{2} \sum_{m,n,l,s\in i} \sum_{\sigma,\sigma'} [mn|ls] m_\sigma^+ n_{\sigma'}^+ s_{\sigma'}^- l_\sigma^-$$

The two-geminal part becomes:[48]

$$\hat{H}_{ij} = \hat{H}_{\text{1-CT}} + \hat{H}_{\text{disp}} + \hat{H}_{\text{2-CT}},$$

where the terms refer to different 'geminal-geminal interactions'. The first term decreases the electron number by one on one geminal while increasing it by one on the other, this is a 'delocalization interaction' or a 'one electron charge transfer' (1-CT) term:

$$\hat{H}_{\text{1-CT}} = \sum_{m\in i} \sum_{n\in j} \sum_{\sigma} h_{mn} m_\sigma^+ n_\sigma^- + \sum_{m,n,l\in i} \sum_{s\in j} \sum_{\sigma,\sigma'} [mn|ls] m_\sigma^+ n_{\sigma'}^+ s_{\sigma'}^- l_\sigma^- + \sum_{m,l,s\in i} \sum_{n\in j} [mn|ls] m_\sigma^+ n_{\sigma'}^+ s_{\sigma'}^- l_\sigma^-$$

The second term does not change the electron number on the geminals, it describes 'dispersive interaction':

$$\hat{H}_{\text{disp}} = \frac{1}{2} \sum_{m,s\in i} \sum_{n,l\in j} \sum_{\sigma,\sigma'} \Big([mn|ls] m_\sigma^+ n_{\sigma'}^+ s_{\sigma'}^- l_\sigma^- - [mn|sl] m_\sigma^+ n_{\sigma'}^+ s_\sigma^- l_{\sigma'}^- \Big) \tag{9}$$

The last term of the two-geminal Hamiltonian annihilates both electrons on one geminal while putting two extra electrons on the other geminal, it is a 'two electron charge transfer term' (2-CT):

$$\hat{H}_{\text{2-CT}} = \frac{1}{2} \sum_{m,n\in i} \sum_{l,s\in j} \sum_{\sigma,\sigma'} [mn|ls] m_\sigma^+ n_{\sigma'}^+ s_{\sigma'}^- l_\sigma^-$$

Similar formulas can be derived for the three- and four-geminal Hamiltonian as well. With these formulas a standard Rayleigh-Schrödinger PT can be outlined with $\hat{H}_0 = \sum_i \hat{H}_i^{\text{eff}}$ (cf. eq.(6)) being the zero-order Hamiltonian, and $\hat{H}-\hat{H}_0$ the perturbation operator.[48,50] Of various possible corrections, we detail here the second order dispersion correction emerging from the perturbation operator of eq.(9):

$$
\begin{aligned}
E^{(2)}_{disp} = &-\sum_{\substack{i,j\\ i<j}} \sum_{p,q}^{S^{p,q}=0} \frac{\left|\sum\limits_{m,s\in i}\sum\limits_{n,l\in j}\left([mn|sl]-\frac{1}{2}[mn|ls]\right)P^{ip}_{sm}P^{jq}_{ln}\right|^2}{E^q_j+E^p_i-E^0_i-E^0_j} - \\
&-\frac{1}{4}\sum_{\substack{i,j\\ i<j}} \sum_{p,q}^{\substack{S^{p,q}=1;\\ M^{p,q}_S=0}} \frac{\left|\sum\limits_{m,s\in i}\sum\limits_{n,l\in j}[mn|ls]P^{ip}_{sm}P^{jq}_{ln}\right|^2}{E^q_j+E^p_i-E^0_i-E^0_j} - \\
&-\sum_{\substack{i,j\\ i<j}} \sum_{p,q}^{\substack{M^{p,q}_S=\pm 1\\ M^p_S+M^q_S=0}} \frac{\left|\sum\limits_{m,s\in i}\sum\limits_{n,l\in j}[mn|ls]\,{}^{\pm}P^{ip}_{sm}\,{}^{\mp}P^{jq}_{ln}\right|^2}{E^q_j+E^p_i-E^0_i-E^0_j},
\end{aligned}
\tag{10}
$$

where E^p_i are the eigenvalues of the geminal Hamiltonian (5), S^p and M^p_S stand for the quantum number of two- electron $\hat{S}^2$ and $\hat{S}_z$ operators corresponding to the p-th state of the geminal, P^{ip}_{ms} is an element of the transition density matrix defined as

$$
P^{ip}_{sm} = \sum_{\sigma}\langle\psi^0_i|m^+_\sigma s^-_\sigma|\psi^p_i\rangle
$$

and ${}^{\pm}P^{ip}_{sm}$ is the spin-flip transition density matrix:

$$
{}^{\pm}P^{ip}_{sm} = \sum_{\sigma}\langle\psi^0_i|m^+_\sigma s^-_{\bar{\sigma}}|\psi^p_i\rangle
$$

with ψ^p_i is an $M_S\neq 0$ spin-polarized triplet geminal, $+$ referring to $M_S=+1$ and $-$ to $M_S=-1$.

(Note that in previous papers[48–50] the second order correction (10) was displayed erroneously).

It is apparent that this formula has a serious drawback: it exhibits a singularity when the ground state of two geminals, e.g. i and j become degenerate, $E_i^p = E_i^0$ and $E_j^q = E_j^0$. This is perfectly the case when i and j label two dissociating bonds (geminals). As shown in,[51] this feature makes the second order formula inapplicable to describe processes like the symmetric dissociation of H_2O, see also curve 'APSG + PT2' in Fig.2.

In what follows we elaborate a theory free from this deficiency.

3 MR-LCC Theory with the APSG reference state

Here we investigate whether an MR-LCC treatment can cure the problem of dispersion correction for simultaneous dissociation of two single bonds. Previously, an MR-LCC formulation was used to describe 'one-electron delocalization' corrections.[84]

The time-independent Schrödinger equation with the exponential CC ansatz is written with the APSG reference state as:

$$\hat{H}e^{\hat{T}}|\mathrm{APSG}\rangle = Ee^{\hat{T}}|\mathrm{APSG}\rangle.$$

Projecting the Schrödinger equation by $\langle\mathrm{APSG}|e^{-\hat{T}}$ and using the fact that the APSG wave function is normalized, the following energy equation is obtained:

$$\langle\mathrm{APSG}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\mathrm{APSG}\rangle = E \tag{11}$$

Restricting ourselves to Linearized Coupled Cluster (LCC) approximation means a truncation of the Baker-Campbell-Hausdorff expansion after the second term (linear in $\hat{T}$):

$$e^{-\hat{T}}\hat{H}e^{\hat{T}} \approx \hat{H} + [\hat{H},\hat{T}] \tag{12}$$

The cluster operator can be expanded in terms of excitation operators $\hat{X}_k$:

$$\hat{T} = \sum_k t_k \hat{X}_k, \tag{13}$$

where $\hat{X}_k$ will be specified later, and t_k-s are the 'amplitudes', for which the MR-LCC amplitude equations are to be solved. Multiplying the Coupled Cluster Schrödinger equation

$$e^{-\hat{T}}\hat{H}e^{\hat{T}}|\mathrm{APSG}\rangle = E|\mathrm{APSG}\rangle$$

from the left by $\langle\mathrm{APSG}|X_j^\dagger$ and applying the linearization approximation (12) we obtain

$$\langle\mathrm{APSG}|\hat{X}_j^\dagger[\hat{T},\hat{H}]|\mathrm{APSG}\rangle = \langle\mathrm{APSG}|\hat{X}_j^\dagger\hat{H}|\mathrm{APSG}\rangle\ ,$$

assuming $\hat{X}_j^\dagger|\mathrm{APSG}\rangle = 0$. Substituting the expansion of the cluster operator (13), we get a linear equation system for the cluster amplitudes t_i

$$\sum_i \underbrace{\langle\mathrm{APSG}|\hat{X}_j^\dagger(\hat{X}_i\hat{H} - \hat{H}\hat{X}_i)|\mathrm{APSG}\rangle}_{A_{ji}} t_i = \underbrace{\langle\mathrm{APSG}|\hat{X}_j^\dagger\hat{H}|\mathrm{APSG}\rangle}_{b_j} \tag{14}$$

symbolized by

$$\mathbf{A}\ \mathbf{t} = \mathbf{b}\ .$$

These are the MR-LCC amplitude equations.

Before solving eq.(14), a decision is to be made on the type of excitations included in the cluster operator. It is possible e.g. to resort to dispersive amplitudes only. In this case the excitations appearing in eq.(13) are specified as:

$$\hat{T} = \sum_{k}^{disp} t_k \hat{X}_k = \sum_{i<j} \sum_{p}^{n_i} \sum_{q}^{n_j} t_{ij}^{pq} \; \psi_{ip}^{+} \psi_{jq}^{+} \psi_{j0}^{-} \psi_{i0}^{-} . \tag{15}$$

In (15) p and q run over the excited states of geminals i and j respectively (n_i is the number of excited states of geminal i), and ψ_{ip}^{+} and ψ_{jq}^{+} refer to the excited state creation operators of the corresponding geminal while ψ_{i0}^{-} and ψ_{j0}^{-} annihilate ground-state geminals (they are both solutions of eqs.(5)). When working with dispersive amplitudes only, it is the 'dispersive' Hamiltonian which contributes to inhomogeneous vector **b** of the LCC equations. This is a direct consequence of the fact, that $\hat{H}_{\text{disp}}$ conserves the particle number on each of the geminals and so does the cluster operator (15). The dispersive PT correction of eq.(10) is closely related to the dispersive-only MR-LCC theory: it arises by neglecting all off-diagonal terms in matrix **A** of eq.(14).

With the (15) definition of the cluster operator we have $\langle \text{APSG} | \hat{T} = 0$, the energy formula (11) consequently simplifies to the form

$$E = \langle \text{APSG} | \hat{H} + \hat{H}\hat{T} | \text{APSG} \rangle \; .$$

Terms of the BCH-expansion with second or higher power of $\hat{T}$ do not contribute to the energy, since the 'dispersive' Hamiltonian can interact with two excited geminals at most.

Based on analogy with single reference LCC theory, one may include in the cluster operator all types of excitations which generate a function that has nonzero interaction with the APSG wavefunction via the Hamiltonian (the so-called first order interacting subspace). This treatment involves dispersive as well as charge transfer states, see Ref. 51 for a categorization of the various excitations. Parametrization of the cluster operator in terms of geminal creation/annihilation operators can be carried out analogously to the dispersive case. Alternatively, one may formulate the theory using determinantal and/or internally contracted excited states. As far as excited functions span the same space, their actual form is irrelevant, due to the invariance of LCC theory to a unitary transformation.[87]

4 Applications

A numerical study on the performance of MR-LCC theory based on the APSG wavefunction is presented in this section, on the two-geminal example of two He atoms and the water molecule. The latter is treated in the frozen core approximation, hence involves four geminals. Apart from total energies, the second derivative of the potential curves at around equilibrium is also examined.

We study the influence of the dimension of Arai subspaces, the effect of dispersive excitations by themselves and the role of excitations involving spin-polarized triplet geminals. Acronym 'MRLCC|DISP' is used to refer to the dispersive parametrization (15) of the cluster operator. Dispersive excited states may involve spin-polarized triplet geminals (coupled to an overall singlet). When spin-polarized geminal states are omitted, 'noSP' appears in the acronym. Abbreviation 'MRLCC' refers to a calculation where all states constituting the first order interacting subspace (including dispersives) are considered when solving the MR-LCC equations. This approach is termed full MR-LCC at some points. Occasionally, the results are compared with the dispersive second-order PT correction of eq.(10), denoted as 'PT2'. FCI values are used as benchmark, in case of water symmetric dissociation FCI refers to frozen core full CI calculations.

4.1 The He...He interaction

The interaction of two helium atoms is investigated in cc-pVDZ basis set,[88] with all orbitals assigned to either of the geminals. This gives 5 as the dimension of each Arai subspace.

In Fig.1. the APSG energies are presented along with MR-LCC results as compared to the FCI energies. The MR-LCC curve is depicted by restricting the excitations to dispersive ones ("APSG-MRLCC|DISP"), as well as the full MR-LCC curve ("APSG-MRLCC").

Fig.1. reflects the fact that the APSG method describes only intra-geminal correlation, thus it cannot describe inter-geminal dispersion. (The shallow well depth apparent in Fig.1. is a genuine basis set superposition effect). Adding the correction of dispersive amplitudes in an MR-LCC scheme, approximately half of the expected well depth is recapitulated. An interesting feature of this example is that a full MR-LCC calculation provides a potential curve which cannot be distinguished by visual inspection from the FCI result, the exact one in the given basis set. (The difference between these results appears in the ninth digit.) Inter-geminal (in this example also interatomic) correlations tend to zero at infinite distance, the APSG energy therefore matches the FCI result in this limit.

4.2 Symmetric dissociation of water

4.2.1 Role of dispersive amplitudes

As a second example the symmetric dissociation of a water molecule is investigated. As seen in Fig.2. the APSG wave function alone describes this process qualitatively well. Second order dispersive perturbative corrections lead to a divergent energy curve, see 'APSG + PT2' in Fig.2. This can be corrected if accounting for dispersive excitations by the MR-LCC theory: the curve

’APSG-MRLCC|DISP’ shows an acceptable agreement with a slight overestimation of the absolute energy for large OH distances. At around equilibrium, the second order results are closer to the FCI ones than those obtained by MR-LCC, see Fig.3. Having a closer look at the energy values, it is apparent that while CC-correction improves the absolute energies, it does not necessarily improve relative energies.

4.2.2 Role of spin-polarized states

Among dispersive amplitudes, those corresponding to spin-polarized excitations were found to have an especial importance. These are states that contain $\hat{X}_k$ excitation operators in Eq.(15) with $\psi_{ip}^{+} = \sum\limits_{m<n\in i} C_{mn}^{ip} m_{\alpha}^{+} n_{\alpha}^{+}$ being an $M_S = +1$ geminal creation operator and $\psi_{jq}^{+} = \sum\limits_{m<n\in j} C_{mn}^{jq} m_{\beta}^{+} n_{\beta}^{+}$ an $M_S = -1$ one, or vice versa. The two spin-polarized triplets with opposite M_S quantum numbers can be coupled into a singlet four-electron state. Omitting these spin-flip triplets, one obtains the curve denoted by ’APSG-MRLCC|DISP_noSP’ in Fig.2. Apart from badly overestimating the dissociation energy, this curve goes through a singularity at around 3.0 Å, rendering the potential curve meaningless. This is a typical example of an LCC singularity caused by a zero eigenvalue of coefficient matrix **A**.

4.2.3 Role of charge transfer and virtual excitations

As Figs. 2. and 3. reflect, dispersion correction alone is enough to describe water dissociation without a divergent potential energy surface. Moreover, dispersion correction accounts for much part of the correlation effect in the example of Fig.2. This is not a general observation however, the error committed by ’APSG-MRLCC|DISP’ depends largely on the number of orbitals assigned to geminals. To show this, in the next example only two orbitals are assigned to the geminals describing the OH bonds, while lone pairs remain at the Hartree-Fock level, with only one function per geminal. This is called the Generalized Valence Bond (GVB) function[89] (also called as *best orbital description* in Ref. 53). In Fig.4. it is again visible that dispersive excitations described by MR-LCC give a qualitatively correct result. However the total energy shows hardly any improvement at around equilibrium, while a rough 40% of the error of GVB is accounted for in the dissociation limit.

Taking into account all excitations yielding the first order interacting subspace, curve labeled 'GVB-MRLCC' in Fig.4. is obtained. Apart from dispersion, this method takes into account the effect of 'charge transfer' (i.e. annihilating one or two orbitals of a geminal and creating orbitals on other geminals and also excitations to those orbitals that do not belong to the subset of any of the geminals). Inspection of Fig.4. reveals that the total energy is much improved by 'GVB-MRLCC' at around equilibrium. The curve however shows a hump at intermediate bond distance and results in an overestimation of the correlation energy in the dissociation limit. It is interesting to investigate the effect of omitting spin-polarized excitations on the 'GVB-MRLCC' results. The curve produced by this approach is labeled 'GVB-MRLCC_noSP' in Fig.4. Apparently, removal of spin-polarized geminal states cures the shape of the potential energy curve. On the other hand regarding the dissociation energy, neither 'GVB-MRLCC' nor 'GVB-MRLCC_noSP' is accurate. Fig.5. shows that at equilibrium bond distance we are quite close to the FCI results, while at infinite distance both approaches are erroneous by tens of millihartrees.

Staying with a GVB reference function described above, Fig.6. displays the water dissociation profile calculated in 6-31G* basis set. One can observe that 'dispersive states' by themselves (GVB-MRLCC|DISP) produce a curve of qualitatively correct shape, though minor improvement in total energy. When all states of the first-order interacting space are considered, we experience again a hump, but the effect in the polarized basis is considerably greater (GVB-MRLCC). The breakdown of the GVB-MRLCC curve at large bond distance is similar to the behavior observed for the imperfect pairing (IPP) approximation by Head-Gordon and co-workers.[39] The analysis of a two-geminal dissociation case has revealed that the reason behind the ill-behavior of IPP is the presence of a spurious singlet-singlet coupling term in the wavefunction.[40] The reason of the erroneous curve obtained by GVB-MRLCC may also lie with the incorrect spin state of the fragments, as described by APSG in the dissociated limit. Omitting spin-polarized geminals (GVB-MRLCC_noSP) from the full MR-LCC treatment cures the shape of the curve on this example also.

4.2.4 Force constants for water

Table 1. shows force constants of the water molecule, calculated by taking 5 equidistant points along the symmetric stretching, near equilibrium with $2 \cdot 10^{-4}$ Å difference in bond length. Either 'APSG' or 'GVB' is regarded as reference. In the former case each virtual orbital is given to one of the geminals (a full geminal calculation), while in case of 'GVB' only two orbitals are assigned to the bonding geminals. Comparing force constants of the geminal based methods we see that 'APSG' overestimates the force constant, 'GVB' does the opposite. Dispersive excitations described by MR-LCC (APSG-MRLCC|DISP) give a better picture than a PT correction (APSG + PT2). Keeping the 'GVB' function as the reference, a full MR-LCC correction considerably improves the results (GVB-MRLCC).

5 Conclusion

A linearized MR-LCC method based on the APSG reference state was introduced. The theory is size-extensive and no redundancy problems appear, the system of amplitude equations is uniquely soluble. The theory is able to correct the deficiencies of the PT formulation based on the dispersive states, namely the incorrect description of simultaneous single bond dissociations due to zero denominators in the dissociation limit. Addition of single and double excitations of the determinants present in the APSG reference further improves the results. In case of the He...He interaction the MR-LCC results get very close to the FCI calculations. Omission of the spin-polarized states from the dispersive Coupled Cluster method also leads to divergences in the symmetric dissociation of water. When performing full MR-LCC calculations (this case was investigated for the GVB wave function only), we experienced the opposite thing, spin-polarized triplet geminal states were responsible for incorrect dissociation behaviour. Among the methods investigated, this latter one gives the smallest error. It is however necessary to perform further studies on a wider set of examples, before a decision on the preferred correction scheme to APSG could be made.

6 Acknowledgements

The authors would like to thank for the grants TAMOP-4.2.2./B-10/1-2010-0030, OTKA-K81588 and OTKA-K81590 for financial support.

References

(1) Čížek, J. *J. Chem. Phys.* **1966**, *45*, 4256.

(2) Cizek, J.; Paldus, J. *Int. J. Quantum Chem.* **1971**, *5*, 359.

(3) Lee, Y. S.; Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1984**, *81*, 5906.

(4) Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1987**, *86*, 7041.

(5) Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1992**, *97*, 4282.

(6) Shavitt, I.; Bartlett, R. *Many-Body Methods in Chemistry and Physics*; Cambridge University Press: Cambridge, 2009; pp 251–430.

(7) Cave, R. J.; Davidson, E. R. *J. Chem. Phys.* **1993**, *88*, 5770.

(8) Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983.

(9) Hirao, K. *Chem. Phys. Letters* **1993**, *201*, 59.

(10) Kozlowski, P. M.; Davidson, E. R. *Chem. Phys. Letters* **1994**, *222*, 615–620.

(11) Pariser, O.; Ellinger, Y. *Chem. Phys.* **1996**, *205*, 323–349.

(12) Chaudhuri, R. K.; Freed, K. F. *J. Chem. Phys.* **1997**, *107*, 6699.

(13) van Dam, H. J. J.; van Lenthe, J. H.; Pulay, P. *Mol. Phys.* **1998**, *93*, 431.

(14) Chen, F.; Davidson, E.; Iwata, S. *Int. J. Quantum Chem.* **2002**, *86*, 256.

(15) Pahari, D.; Chattopadhyay, S.; Das, S.; Mukherjee, D. *Chem. Phys. Letters* **2003**, *381*, 223.

(16) Balkova, A.; Kucharski, S. A.; Meissner, L.; Bartlett, R. J. *Theor. Chim. Acta* **1991**, *80*, 335.

(17) Piecuch, P.; Paldus, J. *Theor. Chim. Acta* **1992**, *83*, 69.

(18) Paldus, J.; Piecuch, P.; Pylypow, L.; Jeziorski, B. *Phys. Rev. A* **1993**, *47*, 2738.

(19) Berkovic, S.; Kaldor, U. *J. Chem. Phys.* **1993**, *98*, 3090.

(20) Piecuch, P.; Oliphant, N.; Adamowicz, L. *J. Chem. Phys.* **1993**, *99*, 1875.

(21) Piecuch, P.; Paldus, J. *J. Chem. Phys.* **1994**, *101*, 5875.

(22) Balkova, A.; Bartlett, R. J. *J. Chem. Phys.* **1994**, *101*, 8972.

(23) Piecuch, P.; Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1999**, *110*, 6103.

(24) Kowalski, K.; Piecuch, P. *J. Mol. Struct.* (THEOCHEM) **2001**, *547*, 191.

(25) Li, X.; Paldus, J. *J. Chem. Phys.* **2004**, *120*, 5890.

(26) Lyakh, D. I.; Ivanov, V. V.; Adamowicz, L. *J. Chem. Phys.* **2005**, *122*, 024108.

(27) Gour, J. R.; Piecuch, P.; Włoch, M. *J. Chem. Phys.* **2005**, *123*, 134113.

(28) Hanrath, M. *J. Chem. Phys.* **2005**, *123*, 84102.

(29) Hanrath, M. *Chem. Phys. Letters* **2006**, *420*, 426.

(30) Mašik, J.; Hubač, I. *Coll.Czech.Chem.Commun* **1997**, *62*, 829.

(31) Pittner, J.; Nachtigall, P.; Čársky, P.; Mašik, J.; Hubač, I. *J. Chem. Phys.* **1999**, *110*, 10275.

(32) Hubač, I.; Pittner, J.; Čársky, P. *J. Chem. Phys.* **2000**, *112*, 8785.

(33) Pittner, J. *J. Chem. Phys.* **2003**, *118*, 10976.

(34) Jeziorski, B.; Monkhorst, H. J. *Phys. Rev. A* **1981**, *24*, 1668.

(35) Chattopadhyay, S.; Mahapatra, U.; Mukherjee, D. *J. Chem. Phys.* **1999**, *111*, 3820–3831.

(36) Piecuch, P.; Oliphant, N.; Adamowicz, L. *J. Chem. Phys.* **1993**, *99*, 1875.

(37) Piecuch, P.; Adamowicz, L. *J. Chem. Phys.* **1994**, *100*, 5792.

(38) Kállay, M.; Szalay, P.; Surján, P. *J. Chem. Phys.* **2002**, *117*, 980–989.

(39) Voorhis, T. V.; Head-Gordon, M. *Chem. Phys. Letters* **2000**, *317*, 575–580.

(40) Voorhis, T. V.; Head-Gordon, M. *J. Chem. Phys.* **2001**, *115*, 7814–7821.

(41) Voorhis, T. V.; Head-Gordon, M. *J. Chem. Phys.* **2002**, *117*, 9190–9201.

(42) Yanai, T.; Chan, G. K.-L. *J. Chem. Phys.* **2006**, *124*, 194106.

(43) Hanauer, M.; Köhn, A. *J. Chem. Phys.* **2011**, *134*, 204111.

(44) Evangelista, F. A.; Gauss, J. *J. Chem. Phys.* **2011**, *134*, 114102.

(45) Chen, Z.; Hoffmann, M. R. *J. Chem. Phys.* **2012**, *137*, 014108.

(46) Dyall, K. *J. Chem. Phys.* **1995**, *102*, 4909.

(47) Surján, P. R. *Phys. Rev. A* **1984**, *30*, 43–50.

(48) Surján, P. R.; Mayer, I.; Lukovits, I. *Phys. Rev. A* **1985**, *32*, 748.

(49) Surján, P. R. *Int. J. Quantum Chem.* **1994**, *52*, 563–574.

(50) Rosta, E.; Surján, P. R. *Int. J. Quantum Chem.* **2000**, *80*, 96–104.

(51) Rosta, E.; Surján, P. R. *J. Chem. Phys.* **2002**, *116*, 878–890.

(52) Hurley, A. C.; Lennard-Jones, J.; Pople, J. A. *Proc. Roy. Soc. (London)* **1953**, *A220*, 446.

(53) Parks, J. M.; Parr, R. G. *J. Chem. Phys.* **1957**, *28*, 335.

(54) Kapuy, E. *Acta Phys. Hung.* **1958**, *9*, 237.

(55) Kapuy, E. *Acta Phys. Hung.* **1959**, *10*, 125.

(56) Kapuy, E. *Acta Phys. Hung.* **1960**, *12*, 185.

(57) Kutzelnigg, W. *J. Chem. Phys.* **1964**, *40*, 3640.

(58) Kapuy, E. *Chem. Phys. Letters* **1968**, *3*, 43.

(59) Paldus, J. *J. Chem. Phys.* **1972**, *57*, 638.

(60) Paldus, J.; Sengupta, S.; Čížek, J. *J. Chem. Phys.* **1972**, *57*, 652.

(61) Surján, P. R. *Topics in current chemistry* **1999**, *203*, 63–88.

(62) Kutzelnigg, W. *Chem. Phys.* **2012**, *401*, 119.

(63) Čížek, J. *J. Chem. Phys.* **1966**, *45*, 4256.

(64) Čížek, J. *Adv. Chem. Phys.* **1969**, *14*, 35.

(65) Bartlett, R. J.; Shavitt, I. *Chem. Phys. Letters* **1977**, *50*, 190.

(66) Bartlett, R. J.; Shavitt, I. *Chem. Phys. Letters* **1978**, *57*, 157.

(67) Bartlett, R. J.; Purvis, G. D. *Int. J. Quantum Chem.* **1978**, *14*, 561.

(68) Ahlrichs, R.; Scharf, P. *Adv. Chem. Phys.* **1987**, *67*, 501.

(69) Szalay, P. G. *Recent Advances in Computational Chemistry* **1997**, *3*, 81.

(70) Szabados, Á.; Surján, P. R. *Chem. Phys. Letters* **1999**, *308*, 303.

(71) Surján, P. R.; Szabados, A. *J. Chem. Phys.* **2000**, *112*, 4438.

(72) Witek, H. A.; Nakano, H.; Hirao, K. *J. Chem. Phys.* **2003**, *118*, 8197–8206.

(73) Witek, H. A.; Nakano, H.; Hirao, K. *J. Comput. Chem.* **2003**, *24*, 1390–1400.

(74) Li, S.; Ma, J.; Jiang, Y. *Int. J. Quantum Chem.* **2000**, *78*, 153.

(75) Li, S.; Ma, J.; Jiang, Y. *J. Chem. Phys.* **2003**, *118*, 5736.

(76) Girardeau, M. *J. Math. Phys* **1963**, *4*, 1096.

(77) Kvasnićka, V. *Czech. J. Phys.* **1982**, *B32*, 947.

(78) Surján, P. R. *Croatica Chimica Acta* **1984**, *57*, 833–854.

(79) Valdemoro, C. *Phys. Rev. A* **1985**, *31*, 2114.

(80) Valdemoro, C. *Phys. Rev. A* **1985**, *31*, 2123.

(81) Poirier, R. A.; Surján, P. R. *J. Comput. Chem.* **1987**, *8*, 436–441.

(82) Surján, P. R. *Croatica Chemica Acta* **1989**, *62*, 579.

(83) Surján, P. R. In *Theoretical Models of Chemical Bonding, Part 2, The Concept of the Chemical Bond*; Maksic, Z. B., Ed.; Springer: HEIDELBERG, 1989; pp 205–256.

(84) Surján, P. R. *Int. J. Quantum Chem.* **1995**, *55*, 109–116.

(85) Surján, P. R.; Kállay, M.; Szabados, Á. *Int. J. Quantum Chem.* **1998**, *70*, 571.

(86) Arai, T. *J. Chem. Phys.* **1960**, *33*, 95.

(87) Gdanitz, R.; Ahlrichs, R. *Chem. Phys. Letters* **1988**, *143*, 413.

(88) Woon, D. E.; Dunning Jr., T. *J. Chem. Phys.* **1994**, *100*, 2975.

(89) Bobrowicz, F. W.; Goddard-III, W. A. In *Methods of Electronic Structure Theory*; Schaefer-III, H. F., Ed.; Plenum: New York, 1977; p 79.

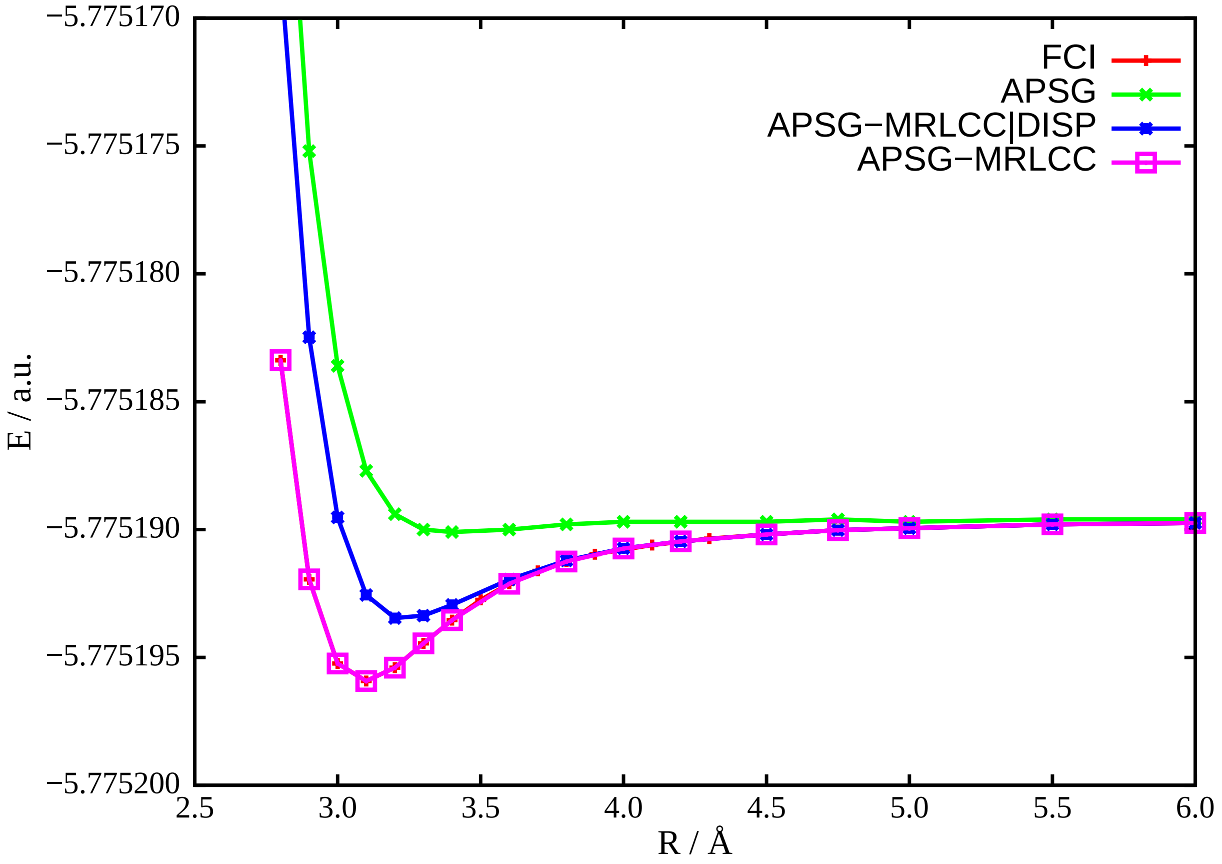

Figure 1: Energy of the helium dimer as a function of interatomic distance in cc-pVDZ basis set. There are 5 functions on each helium in the APSG Wavefunction. See text for acronyms.

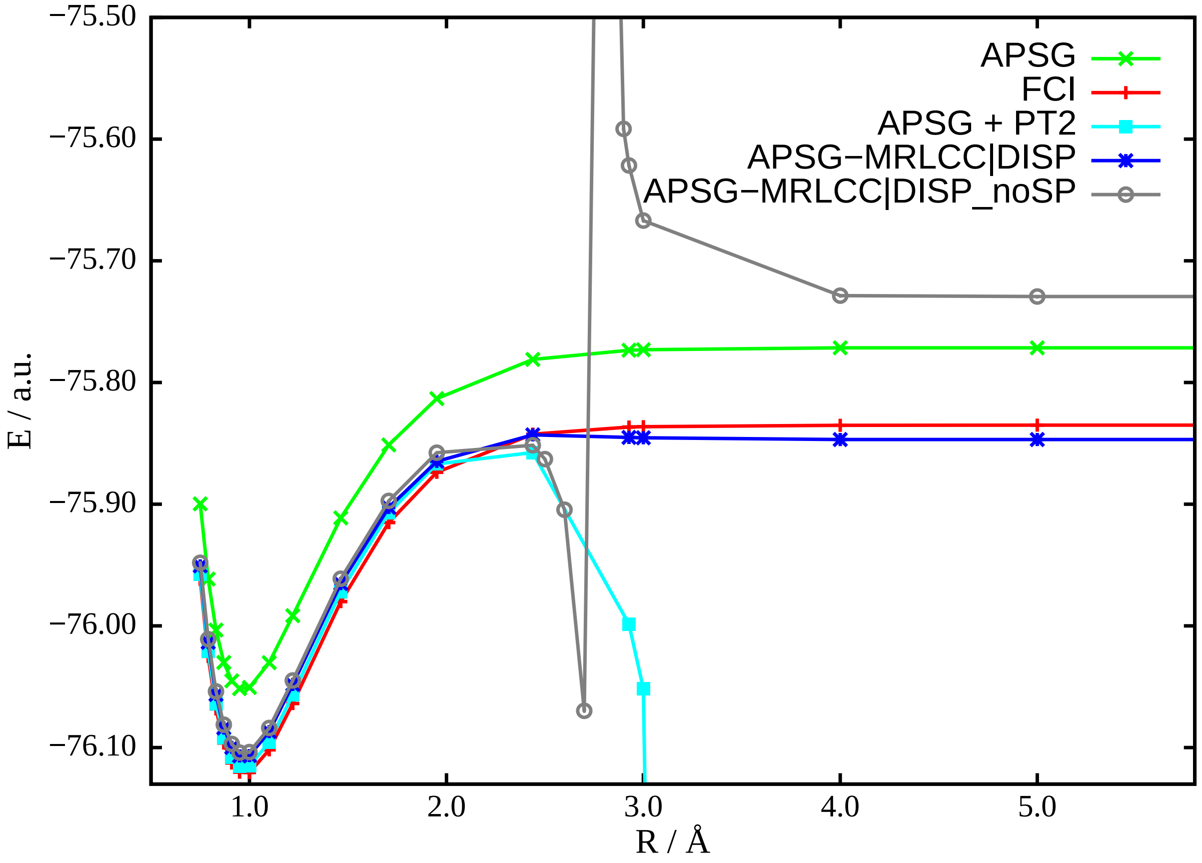

Figure 2: Water symmetric dissociation in 6-31G basis set with bond angle fixed at ∡ (H-O-H) = 110.6°. An APSG wavefunction is assumed as reference. Frozen core approximation is applied. The APSG reference has 4 orbitals on the bonding and 2 on the non-bonding geminals. See text for acronyms.

Table 1: Force constants for water symmetric stretching in 6-31G basis set, ∡ (H-O-H) = 110.6°

. See text for acronyms. GVB reference contains two orbitals on the bonding geminals and one on the non-bondings, APSG reference has 4 orbitals on the bonds and 2 on the non-bonding geminals.

Method	Force constant / $[E_h/a_0^2]$
APSG	1.06
GVB	0.86
APSG + PT2	0.92
APSG-MRLCC|DISP	1.01
GVB-MRLCC	0.96
frozen core FCI	0.98

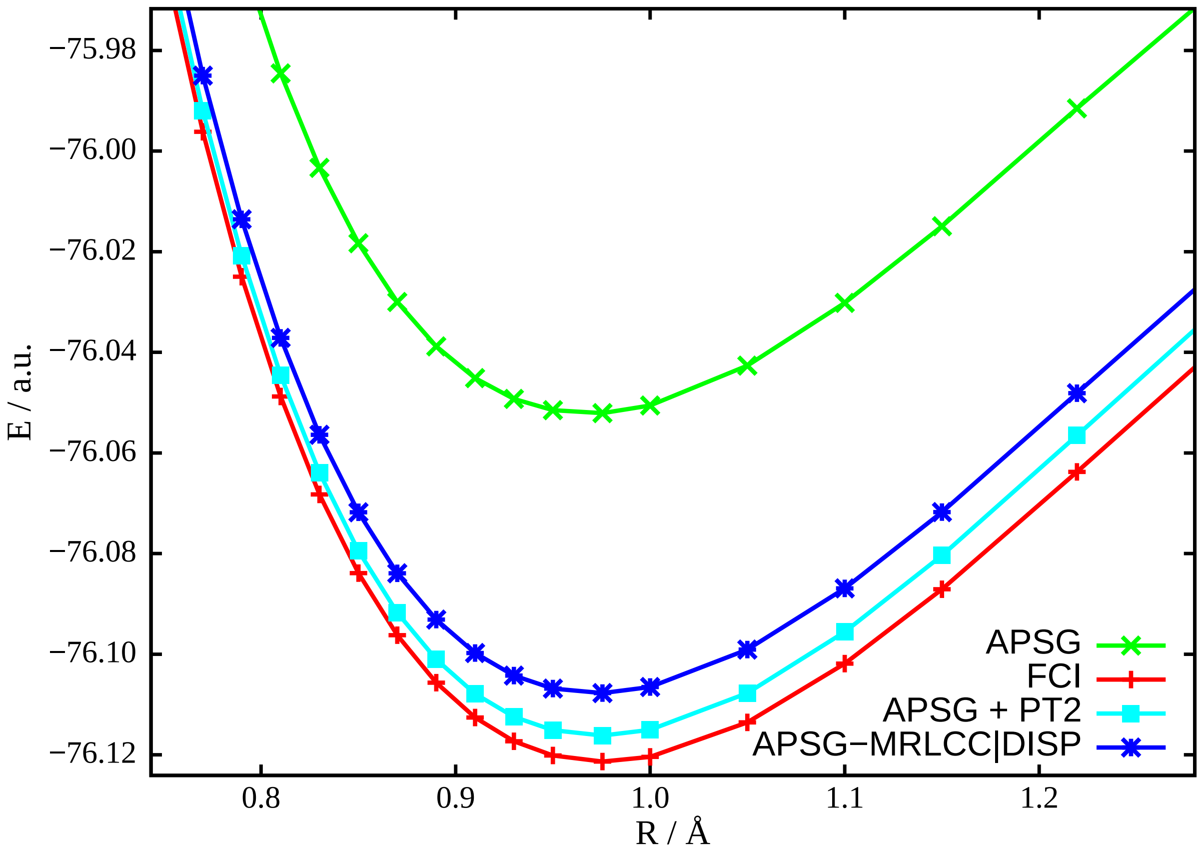

Figure 3: Water symmetric dissociation around equilibrium in 6-31G basis set. See text for acronyms and the legend of Fig.2. for other particulars.

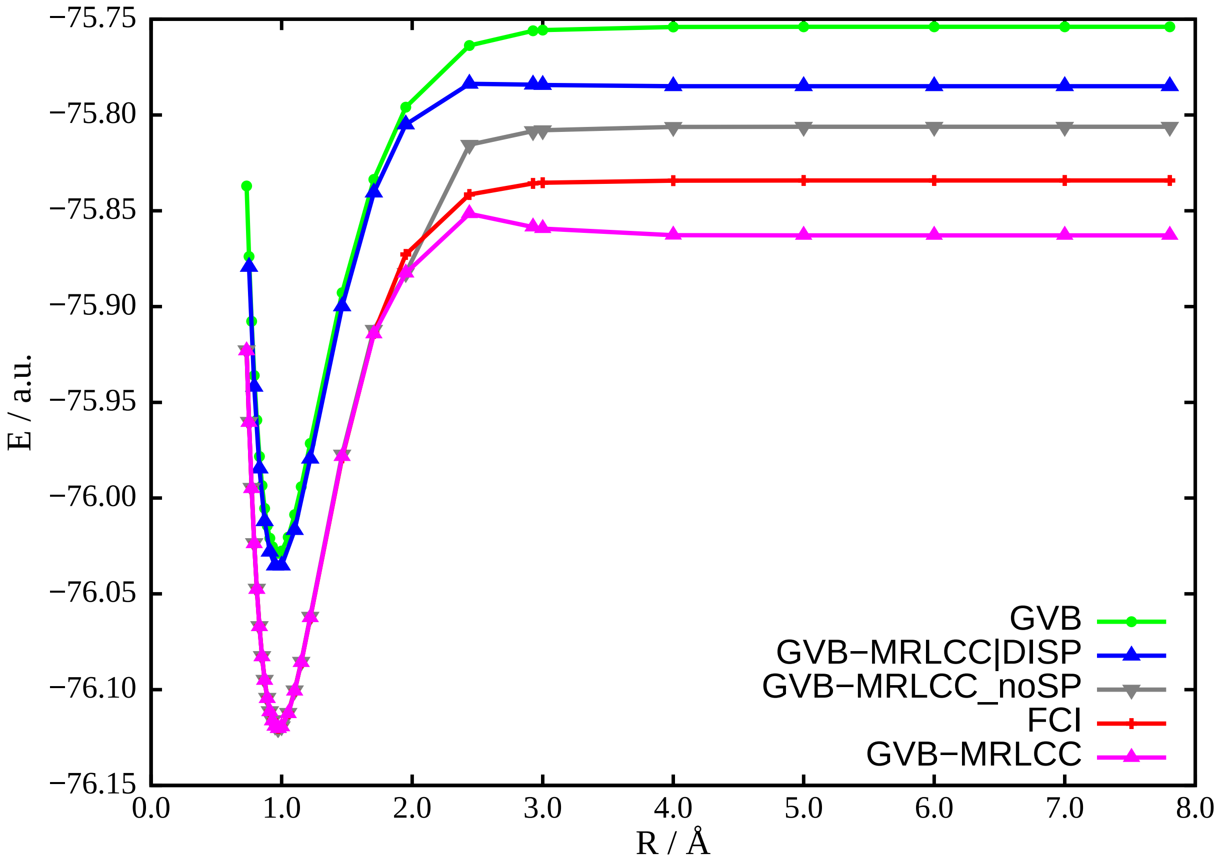

Figure 4: Water symmetric dissociation in 6-31G basis set with bond angle fixed at ∡ (H-O-H) = 110.6°. A GVB wavefunction is assumed as reference and frozen core approximation is applied. The GVB reference contains two orbitals on the bonding and one on the non-bonding geminals. See text for acronyms.

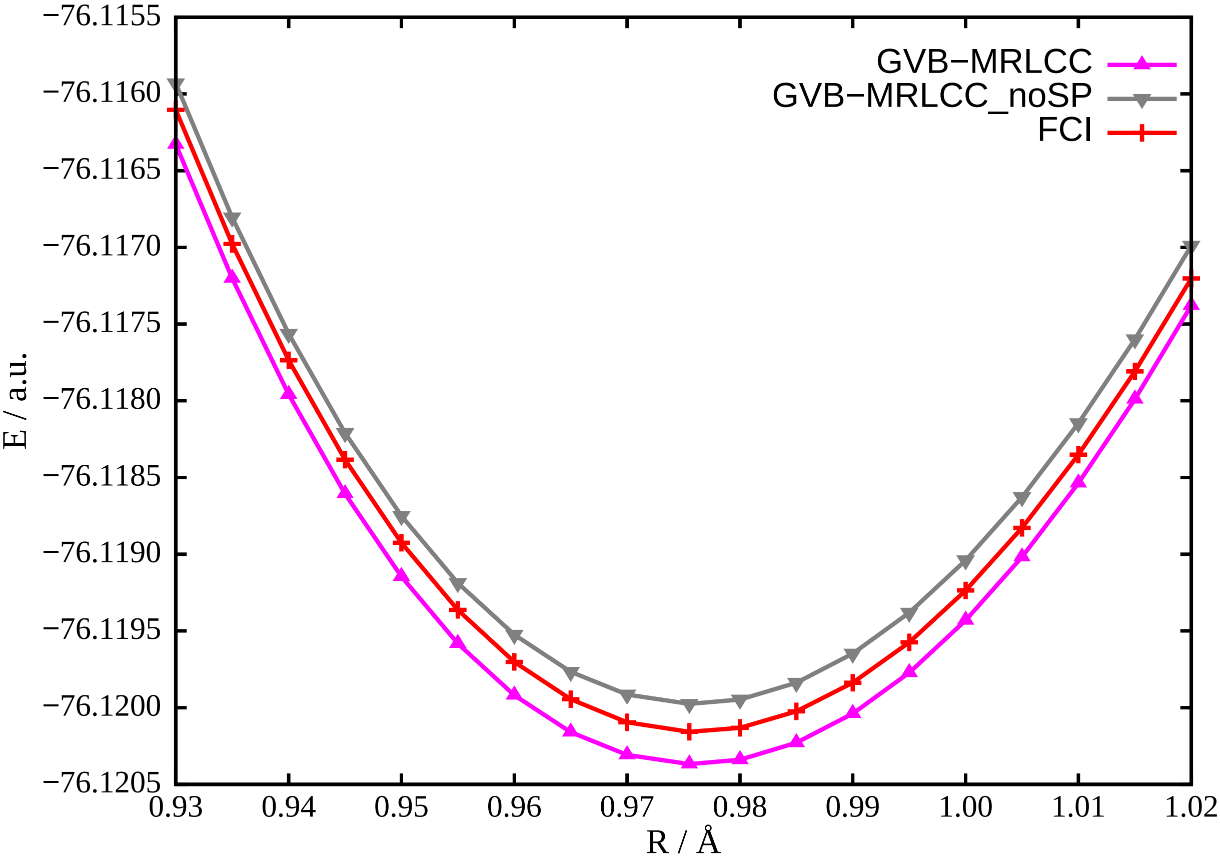

Figure 5: Water symmetric dissociation around equilibrium in 6-31G basis set. See text for acronyms and the legend of Fig.4. for other particulars.

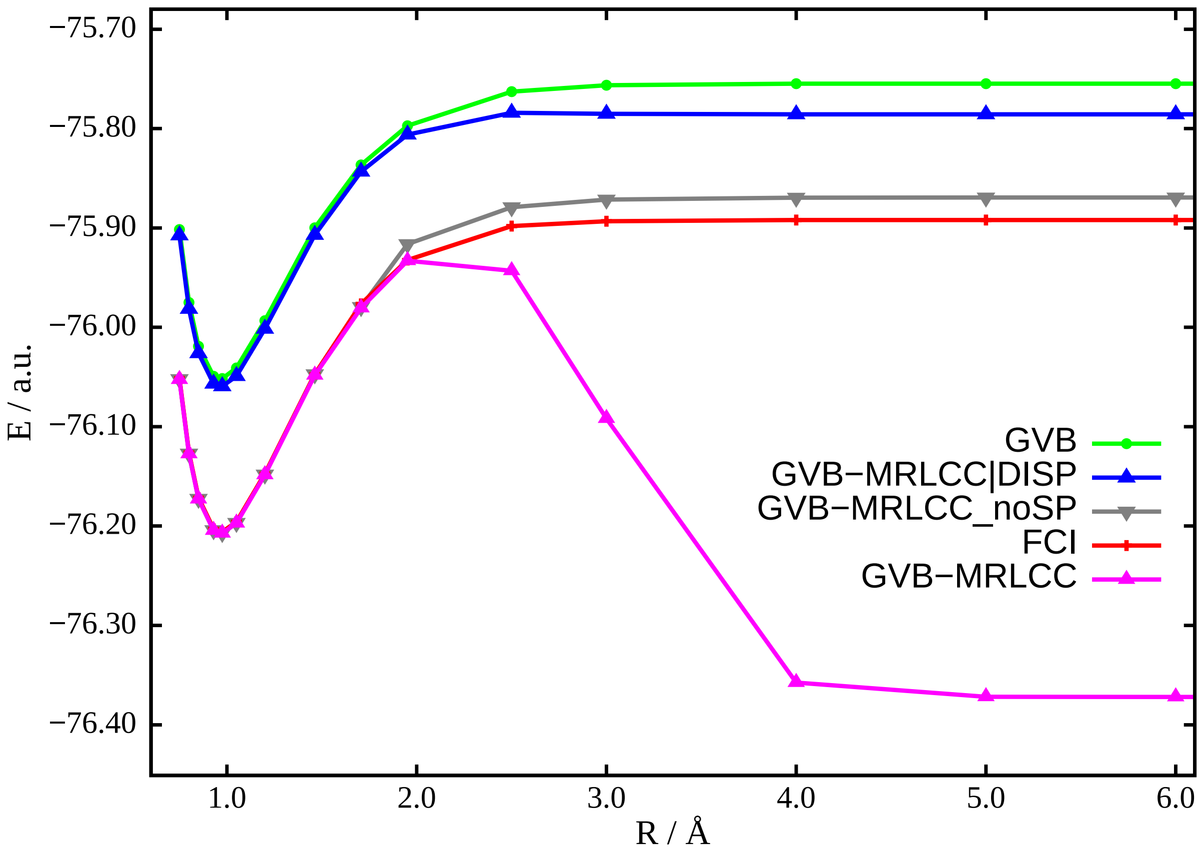

Figure 6: Water symmetric dissociation in 6-31G* basis set with bond angle fixed at ∡ (H-O-H) = 110.6°. A GVB wavefunction is assumed as reference and frozen core approximation is applied. The GVB reference contains two orbitals on the bonding and one on the non-bonding geminals. See text for acronyms.