Theoretical Chemistry Accounts manuscript No.
(will be inserted by the editor)

Multiple bond breaking with APSG based correlation methods

Comparison of two approaches

Ádám Margócsy · Piotr Kowalski · Katarzyna Pernal · Ágnes Szabados

The work is dedicated to the memory of János Ángyán.

Received: xxx / Accepted: xxx

Abstract Antisymmetrized product of strongly orthogonal geminals (APSG) Ansatz is a computationally economic wavefunction class with favourable for-

Ádám Margócsy

Doctoral School of Chemistry, Faculty of Science, ELTE Eötvös Loránd University, H-1518 Budapest, POB 32, Hungary

E-mail: margocsy@coulson.chem.elte.hu

Piotr Kowalski - Katarzyna Pernal

Institute of Physics, Łodz University of Technology, ul. Wolczanska 219, 90-924, Łodz, Poland

E-mail: pernalk@gmail.com

Ádám Margócsy - Ágnes Szabados

Laboratory of Theoretical Chemistry, Institute of Chemistry, Faculty of Science, ELTE Eötvös Loránd University, H-1518 Budapest, POB 32, Hungary

E-mail: szabados@chem.elte.hu

mal properties. These include extensivity, variational determination of the wavefunction parameters or qualitatively correct description of single bond dissociation. Breaking multiple bonds or non-isolated single bonds results in fragments of incorrect spin state when computed by APSG. This has been identified as a potential problem in APSG based linearized coupled-cluster approach (LCC). An alternative correction scheme based on the extended random phase approximation (ERPA) is investigated from this point of view, in parallel with LCC.

The two methods are compared formally. Potential energy curves and atomic spin by APSG based LCC and ERPA are presented on illustrative examples for multiple bond breaking. Origin of the marked difference between the behaviour of LCC and ERPA is explored.

Keywords antisymmetrized product of strongly orthogonal geminals · beyond perfect pairing · triplet geminals · extended random phase approximation · linearized coupled cluster · local spin

1 Introduction

Practical computational tools for the interplay of strong and weak correlation have been sought after for decades, representing a challenge to date. Wavefunction based electronic structure methods approach these problem in two steps: strong first, weak next. Multireference extension of the self consistent field (SCF) approach is often applied in the first step, considering a full configuration interaction in the space of selected active orbitals (complete active space,

CAS).[1] With the aid of parallel execution, the limit of tractability of CAS has been recently pushed to the order of 20 electrons and 20 active orbitals.[2] Factorial scaling however hinders progress along this line, stimulating many attempts towards reduced cost alternatives to the CAS methodology. Numerical approximation based on e.g. the density matrix renormalization idea[3, 4] or on stochastic algorithms[5] represent a highly successful recent line of research. Efforts have been also invested into designing more economic wavefunction Ansätze, based on occupancy restrictions in subsets of active orbitals, e.g. restricted active space (RAS)[6], generalized active space (GAS)[7] or occupation restricted multiple active space (ORMAS)[8]. Constructing the wavefunction of two-electron building blocks is another useful concept, investigated by many.[9–17]

Among pair function approaches, wavefunctions exploiting the so-called strong orthogonality (SO) criterion represent a special class. Two-electron functions (geminals) $\psi(\mathbf{x}_1, \mathbf{x}_2)$ and $\phi(\mathbf{x}_1, \mathbf{x}_2)$ are termed strongly orthogonal if

$$\int \psi(\mathbf{x}_1, \mathbf{x}_2)\phi(\mathbf{x}_1, \mathbf{x}_2')\mathrm{d}\mathbf{x}_1 \equiv 0 \tag{1}$$

holds for $\psi \neq \phi$. According to Arai[18] SO is equivalent to expanding geminals in mutually orthogonal subspaces of one-electron orbitals. Geminal ψ_γ is consequently expressed, in second-quantized notation as

$$\psi_\gamma^+ = \sum_i^{(\gamma)} C_i \; i_\beta^+ i_\alpha^+ \; , \tag{2}$$

γ referring to the geminal subspace. Form (2) assumes that spatial orbitals $|i\rangle$ are natural orbitals of the many-electron function, constructed as antisymmetrized product of strongly orthogonal geminals (APSG)

$$|\mathrm{APSG}\rangle = \prod_{\gamma} \psi_{\gamma}^{+}|\mathrm{vac}\rangle \ . \tag{3}$$

Note, that APSG includes singlet geminals, i.e. $S = 0, M_S = 0$ for each geminal. When geminal subsets are at most two-dimensional, APSG agrees with the generalized valence bond (GVB) function, assuming perfect pairing (PP).

The APSG wavefunction has many attractive features, e.g. it can be expressed in an exponential form[19] facilitating proper separability as the system breaks up for non-interacting subsystems. Parameters of the theory (configuration interaction (CI) coefficients C_i collected in matrix $\mathbf{C}$ and orbitals $|i\rangle$) can be determined based on the variational principle. Wavefunction optimization with respect to CI coefficients leads to separate eigenvalue problems of the effective geminal Hamiltonians[17]

$$\hat{H}^{(\gamma)} = \sum_{ij}^{(\gamma)} h_{ij}^{\mathrm{eff}} \sum_{\sigma} i_{\sigma}^{+} j_{\sigma}^{-} \ + \ \frac{1}{2} \sum_{ijkl}^{(\gamma)} \langle ij|kl\rangle \sum_{\sigma\sigma'} i_{\sigma}^{+} j_{\sigma'}^{+} l_{\sigma'}^{-} k_{\sigma}^{-} \tag{4}$$

where σ stands for the spin-index, two-electron integrals $\langle ij|kl\rangle$ are written in the 1212 convention and the effective one-electron matrix reads

$$h_{ij}^{\mathrm{eff}} = h_{ij} \ + \sum_{\gamma \neq \gamma_i, \gamma_j} \sum_{kl}^{(\gamma)} P_{lk} \left[2\langle ik|jl\rangle - \langle ik|lj\rangle \right] \tag{5}$$

with h_{ij} being the standard notation for one-electron integrals, and the one-electron reduced density matrix (1-RDM) reading as

$$P_{lk} = \left(\mathbf{C}^{\dagger}\mathbf{C}\right)_{kl} \ = \ \delta_{kl}|C_k|^2 \ . \tag{6}$$

In the above γ_i denotes the geminal subspace accommodating orbital $|i\rangle$. Diagonal nature of the one-particle density is utilized in the second equation of Eq.(6), yielding natural occupation numbers

$$n_k = |C_k|^2 \ . \tag{7}$$

Though k and l refer to spatial orbitals, Eq.(6) gives the spin-dependent one-particle density, spin being suppressed in the notation as P_{lk} is the same for α and β spin. Diagonal element of the spin-summed one-particle density, D_{kk} is accordingly given by $2n_k$. Orbital optimization can be performed based on the orbital gradient derived from the antisymmetric part of the generalized Fockian[17]. Efficient numerical realization of the optimization has been recently explored by Tarumi et al.[20] and by Chatterjee[21].

Being composed of two-electron building blocks, for which the CI problem is solved, the APSG wavefunction incorporates intra-geminal correlation in an exact manner. Processes where a geminal breaks up for two fragments are consequently described properly. Electron interaction between geminals is included in APSG in a mean-field manner, via dependence of the effective geminal Hamiltonian Eq.(4) on the one-particle density of all other geminals.

Recent interest in pair function based methods is partly explained by computational economy. Another source of motivation is provided by the success of pair function methods describing essential (strong) correlation[22–25]. The special structure of APSG has been inspiring various generalizations of single-reference methodologies for capturing dynamical correlation[26–35]. Natural orbital functionals have been also proposed relying on the idea of restricted

occupation in orbital subspaces[36,37], in particular the PNOF5 energy functional has been explicitly shown to match the GVB energy provided that phase factors of the orbitals belonging to the same geminal are opposite[38]. Many of these methods profit from the relatively simple structure of APSG[39] as compared to e.g. CAS.

We here focus on two correlation approaches based on APSG. One is a generalization of the extended random phase approximation (ERPA)[40] and the correlation energy correction derived from it[35,41,42]. The other is a formulation of the linearized coupled-cluster (LCC) method based on the APSG reference function[33]. The investigation focuses on a drawback of the APSG wavefunction, showing up when breaking multiple bonds or single bonds attached to the same atom. Description of such a situation has been known as a weak point of GVB[43], it has in fact been motivating studies for stepping beyond the PP approximation[26,27,44–47]. In spite of the incorrect nature of the wavefunction in multiple bond breaking situations, potential energy curves may be acceptable, as observed for APSG and ERPA-APSG. At difference with this, serious ill-effect has been reported with LCC-APSG and the problem has been tied to the flaw in the reference function[48]. Analysis of the spin structure has been found a valuable source of information in course of this study[49], as well as in formulating constraints for cumulant reconstruction schemes[50].

In what follows, the shortcoming in APSG when breaking multiple bonds is presented. Since atomic spin provides a suitable tool for detecting the problem,

formulae for extracting atomic spin from the molecular wavefunction are briefly summarized. A short account of ERPA-APSG and APSG based LCC theory is followed by an analysis of numerical examples and concluding remarks.

2 Theory

2.1 Multiple bond breaking by APSG

As a case study, let us investigate symmetric dissociation of H_2O, assuming two-dimensional subsets on OH geminals and one-dimensional on all others. The APSG wavefunction accordingly takes the form

$$|\mathrm{APSG}\rangle = \psi^+_{\mathrm{OH2}}\psi^+_{\mathrm{OH1}}|\Xi\rangle \ , \tag{8}$$

Ξ standing for the closed shell part involving the core and lone pairs of atom O. Focusing on the limit where atoms are non interacting with each other, natural orbitals of OH geminals, conforming with Eq.(2) arise as symmetric and antisymmetric combination of atomic orbitals with equal weights. Eigenvalues of the 1-RDM of Eq.(6) being degenerate and equal to 1/2, $\mathbf{P}$ remains unaffected upon expressing ψ^+_{OH} on orbitals localized on atoms (note, that only the 1-RDM is invariant, elements of $\mathbf{C}$ change.). Localized orbitals facilitate making a link with a valence bond (VB) structure reflecting the correct spin coupling in the dissociation limit. This can be written as

$$|T_{\mathrm{O}}T_{\mathrm{HH}}\rangle = \frac{1}{\sqrt{3}}\Big[\binom{3}{+1}\psi^+_{\mathrm{O}}\big)\ \binom{3}{-1}\psi^+_{\mathrm{HH}}\big) \ + \ \binom{3}{-1}\psi^+_{\mathrm{O}}\big)\ \binom{3}{+1}\psi^+_{\mathrm{HH}}\big) \ - \ \binom{3}{0}\psi^+_{\mathrm{O}}\big)\ \binom{3}{0}\psi^+_{\mathrm{HH}}\big)\Big]|\Xi\rangle \tag{9}$$

where ψ_{O} is assigned to two sp hybrid orbitals of atom O, ψ_{HH} is built with each H atom contributing an s orbital. Multiplicity of the geminal function is

indicated in top left index, M_S in bottom left. Making use of the expression of ψ_{OH} with localized orbitals one can relate $|\mathrm{APSG}\rangle$ and the correct VB structure[49], bringing into play another VB structure involving the singlet state of oxygen:

$$|S_{\mathrm{O}}S_{\mathrm{HH}}\rangle = \left({}^1\psi^+_{\mathrm{O}}\right)\,\left({}^1\psi^+_{\mathrm{HH}}\right)|\Xi\rangle\ . \tag{10}$$

In fact, $|\mathrm{APSG}\rangle$ is the combination of the two, with the following coefficients in limiting case:

$$|\mathrm{APSG}\rangle = -\frac{\sqrt{3}}{2}|T_{\mathrm{O}}T_{\mathrm{HH}}\rangle\ -\ \frac{1}{2}|S_{\mathrm{O}}S_{\mathrm{HH}}\rangle\ . \tag{11}$$

Nonzero contribution of the singlet state of atom O in the dissociation limit is an indication of APSG failing to describe the essential change in the spin coupling during the dissociation process. It is however not necessary to reassign geminal subspaces, which would inevitably cause discontinuity on a potential surface. Admitting triplet geminals for dissociating bonds is a cure that allows conservation of geminal subspace assignation[11,44,45,47,49].

2.1.1 Extracting atomic spin from the molecular wavefunction

Spin-squared expectation value of atomic fragments can be straightaway computed in the dissociation limit based on the fragment spin-squared operator defined in the spirit of Clark and Davidson[51], providing

$$\langle T_{\mathrm{HH}}T_{\mathrm{O}}|\hat{S}^2_O|T_{\mathrm{O}}T_{\mathrm{HH}}\rangle = 2\ , \tag{12}$$

$$\langle \mathrm{APSG}|\hat{S}^2_O|\mathrm{APSG}\rangle = \frac{3}{2}\ , \tag{13}$$

a clear numerical indication of the problem. In the following we use atomic spin to measure the qualitative defect of APSG and corrections built upon it. Aiming at an indicator applicable along the entire dissociation process raises the problem of extracting fragment spin for interacting fragments. Atomic spin being non-measurable in a molecule, the quantity may be grabbed in several ways leading to alternative definitions[51–54]. In recent years Mayer and coworkers set four essential requirements for fragment spin and developed methodologies for fulfilling these[55, 56]. Presently we follow the Mulliken-type decomposition of the proposed schemes. Spin density being zero in all models considered in this work, atomic and diatomic components of the spin-squared expectation value are obtained as

$$\langle \hat{S}^2 \rangle_A = \frac{3}{4} \sum_{\mu \in A} \left[2\mathbf{DS} - (\mathbf{DS})^2 \right]_{\mu\mu} + \frac{1}{2} \sum_{\mu,\nu \in A} \sum_{\tau,\rho} \left[\Lambda_{\mu\nu\rho\tau} - \Lambda_{\mu\nu\tau\rho} \right] S_{\rho\mu} S_{\tau\nu} \quad (14)$$

$$\langle \hat{S}^2 \rangle_{AB} = \frac{1}{2} \sum_{\mu \in A} \sum_{\nu \in B} \sum_{\tau,\rho} \left[\Lambda_{\mu\nu\rho\tau} - \Lambda_{\mu\nu\tau\rho} \right] S_{\rho\mu} S_{\tau\nu} \ , \quad (15)$$

where A, B refer to atoms, indices μ, ν, ρ, τ designate atomic orbitals, $\mathbf{S}$ denotes the overlap matrix, $\mathbf{D} = 2\mathbf{P}$ stands for the spin-less 1-RDM and finally the spin-less cumulant is expressed with the 2-RDM as

$$\Lambda_{\mu\nu\rho\tau} = \Gamma_{\mu\nu\rho\tau} - D_{\mu\rho} D_{\nu\tau} + 2 P_{\mu\tau} P_{\nu\rho} \ . \quad (16)$$

See below the expression of the RDM's for the the respective theoretical models.

In the calculations presented, we make explicit use of the requirement of Mayer and coworkers which states that closed shell restricted wave functions should not give rise to local spin. This allows separating the closed shell, doubly

occupied portion of the wavefunction, in the spirit of the frozen core approach and showing that doubly filled orbitals contribute zero to $\langle \hat{S}^2 \rangle_A$.

In addition to fragment spin, a delocalization index was examined in Ref.[50], which should tend to zero betwen two atoms in the limit of completely broken bond. This quantity is well behaving already at the level of APSG or equivalently PNOF5[50] and it correctly tends to zero in the dissociation limit also when obtained by ERPA-APSG or LCC-APSG. For this reason delocalization index is not commented on further in this report.

2.2 ERPA-APSG

One starting point of formulating the ERPA-APSG correction is the following expression of electron-electron interaction

$$E_{ee} = \frac{1}{2} \int \int \frac{\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \, \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2 \tag{17}$$

where $\rho^{(2)}(\mathbf{x}_1, \mathbf{x}_2)$ stands for the pair density. Intrageminal interaction, already accounted for by APSG is to be excluded from Eq.(17). This is achieved by expressing the pair density with one-particle density and its fluctuation, and restricting the density fluctuation contribution to nonequivalent pairs of geminals. By a series of elaborate arguments one is led to the expression[35]

$$E^{\text{ERPA-APSG}} = E^{\text{APSG}} + \sum_{i>j} \sum_{k>l} \left(1 - \delta_{\gamma_i \gamma_j} \delta_{\gamma_k \gamma_l} \delta_{\gamma_i \gamma_k}\right) \langle ik|jl\rangle \tag{18}$$

$$\left[\frac{1}{4} \sum_{\nu} \langle \text{APSG}|E_j^i + E_i^j|\nu\rangle \langle \text{APSG}|E_l^k + E_k^l|\nu\rangle - \frac{1}{2} \left[n_i(1-n_j) + n_j(1-n_i) \right] \delta_{ik} \delta_{jl} \right] .$$

In the above E^{APSG} is the expectation value of the Hamiltonian taken with the wavefunction of Eq.(3). Orbitals are assumed natural and ordered in

decreasing or equal value of natural occupation number, n_i given by Eq.(7). Degeneracy in natural occupancies may lead to $n_i = n_j$ for $i \neq j$, necessitating special consideration, *vide infra*. Index ν refers to excited states and E_i^j stands for the spin-summed excitation operator

$$E_j^i = \sum_\sigma i_\sigma^+ j_\sigma^- \ . \tag{19}$$

Transition density matrices are derived in the spirit of equation of motion theory[57] as

$$\langle \mathrm{APSG}|E_j^i|\nu\rangle = \langle \mathrm{APSG}|[E_j^i, \hat{O}_\nu^\dagger]|\mathrm{APSG}\rangle \tag{20}$$

$\hat{O}_\nu^\dagger$ being the excitation operator parametrized as

$$\hat{O}_\nu^\dagger = \sum_{i>j} (X_\nu)_{ij} E_j^i \ + \ \sum_{i>j} (Y_\nu)_{ij} E_i^j \tag{21}$$

according to an extension of RPA for APSG. Utilizing the expressions given in Ref.[40] one finds for the sum of transition density matrices appearing in $E^{\mathrm{ERPA\text{-}APSG}}$

$$\langle \mathrm{APSG}|E_j^i + E_i^j|\nu\rangle = 2(n_i - n_j)\left((Y_\nu)_{ij} - (X_\nu)_{ij}\right) \ , \ i > j \tag{22}$$

Finally, introducing index pair ij as hyper row index and index ν as column index, the concise equation determining amplitude matrices $\mathbf{X}$ and $\mathbf{Y}$ reads

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B} & \mathbf{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & -\mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \boldsymbol{\omega}^{\mathrm{diag}} \ . \tag{23}$$

In Ref.[40] elements of matrices $\mathbf{A}$ and $\mathbf{B}$ are given explicitly, elements of the overlap matrix are implicit, reading as

$$S_{ij,kl} = -(n_i - n_j)\delta_{ik}\delta_{jl} \ \ , \ i > j \ , k > l \ . \tag{24}$$

Matrix $\boldsymbol{\omega}$ contains positive excitation energies in its diagonal. Amplitude vectors are orthogonal and normalized as

$$\left(\mathbf{X}^{\dagger}\mathbf{S}\mathbf{X}\right)_{\mu\nu} - \left(\mathbf{Y}^{\dagger}\mathbf{S}\mathbf{Y}\right)_{\mu\nu} = \frac{1}{2}\delta_{\mu\nu} \ . \tag{25}$$

Some technical comments on computing ERPA-APSG are due at this point. Being related to the orbital Hessian[58], matrices $\mathbf{A}$ and $\mathbf{B}$ are symmetric provided that orbitals are optimized. We carry out symmetrization before solving Eq.(23) in order to reduce numerical errors stemming from residual orbital gradient. A threshold of 10^{-9} on the antisymmetric part of the generalized Fockian is used to control orbital optimization of APSG, since $\langle \hat{S}^2 \rangle$ was found too sensitive with a 10^{-7} orbital gradient threshold.

Solution of Eq.(23) can be found by transformation to an eigenvalue problems of half size, reading as[35,42]

$$\left[\widetilde{\mathbf{A}}_{+}^{1/2}\widetilde{\mathbf{A}}_{-}\widetilde{\mathbf{A}}_{+}^{1/2}\right]\widetilde{\mathbf{A}}_{+}^{-1/2}\widetilde{\mathbf{Y}}_{-} = \widetilde{\mathbf{A}}_{+}^{-1/2}\widetilde{\mathbf{Y}}_{-}(\boldsymbol{\omega}^{\mathrm{diag}})^2 \tag{26}$$

with

$$\widetilde{\mathbf{Y}}_{-} = \mathbf{L}(\mathbf{Y}-\mathbf{X}) \ , \tag{27}$$

$$\widetilde{\mathbf{A}}_{-} = \mathbf{L}^{-1}(\mathbf{A}-\mathbf{B})\mathbf{L}^{-1} \ , \tag{28}$$

$$\widetilde{\mathbf{A}}_{+} = \mathbf{K}^{-1}(\mathbf{A}+\mathbf{B})\mathbf{K}^{-1} \ , \tag{29}$$

$$\mathbf{K}_{ij,kl} = (C_i + C_j)\delta_{ik}\delta_{jl} \ \ , \ i>j \ , k>l \ , \tag{30}$$

$$\mathbf{L}_{ij,kl} = (C_i - C_j)\delta_{ik}\delta_{jl} \ \ , \ i>j \ , k>l \ . \tag{31}$$

The above equations as well as Eq.(24) reflects that excitation among orbitals of equal natural occupancy $n_i = n_j$, or equivalently $|C_i| = |C_j|$ is to

be avoided. In practice, orbital pairs admitted as $i > j$ resort to virt-act, act-core, virt-core and act-act. A threshold of 0.005 is used to define orbital categories, i.e. $n_i < 0.005$ is considered virtual, $n_i > 0.995$ characterizes the core set. Excitations between active orbitals need further inspection: these are allowed whenever the two orbitals $i > j$ belong to different geminals and $n_i/n_j \leq 0.99$. Altering selection of orbital pairs (elementary excitations) entering Eq.(21) along a potential surface may lead to discontinuities. In order to avoid this, elementary excitations prone to cause discontinuity at some region are preferably discarded all along the potential surface.

Active orbital pairs i and j belonging to the same geminal represent an exceptional case, being admissible regardless of their occupation number. The reason behind is that by proper transformation of Eq.(23) and relying on the stationary conditions of the parameters of APSG $(n_i - n_j)$ can be eliminated from Eq.(23). Whether or not $(n_i - n_j)$ equals zero is consequently immaterial.

The overlap matrix becoming (near) singular typically manifests in the stability conditions being (nearly) violated, but this is not the only possible source of instability. Such situations, characterized by small or negative eigenvalues of $(\mathbf{A} + \mathbf{B})$ and/or $(\mathbf{A} - \mathbf{B})$ [58] are to be avoided. This is achieved by identifying elementary excitations contributing to the problematic root of Eq.(23) and discarding the corresponding orbital pairs from Eq.(21), again along the entire surface.

Wrapping up the description of orbital pair selection it may be worth mentioning that the scheme essentially agrees with the procedure applied in e.g.[41,

42]. Discarding elementary excitations in order to avoid discontinuities is an ingredient not commented on before. This becomes imperative with multiple bond breaking which was not examined in Refs.[41,42]

2.2.1 Density matrices behind ERPA-APSG

One way of interpreting the ERPA-APSG formula is that it stems from a correction beyond mean-field to the intergeminal interaction incorporated in the APSG two-particle density. Identifying the ERPA-APSG two-particle RDM, the energy expression

$$E^{\text{ERPA-APSG}} = \sum_{ij} h_{ij} D_{ji} \; + \; \frac{1}{2} \sum_{ijkl} \langle ik|jl\rangle \Gamma^{\text{ERPA-APSG}}_{jlik} \tag{32}$$

remains valid. The ERPA-APSG 2-RDM reads[42]

$$\begin{aligned} \Gamma^{\text{ERPA-APSG}}_{jlik} = \Gamma^{\text{APSG}}_{jlik} \; + \; & 2(1 - \delta_{\gamma_i\gamma_j}\delta_{\gamma_k\gamma_l}\delta_{\gamma_i\gamma_k}) \\ & \left[\frac{1}{4} \sum_{\nu} \langle \text{APSG}|E^i_j|\nu\rangle\langle \text{APSG}|E^l_k|\nu\rangle - \frac{1}{2} n_i(1-n_k)\delta_{kj}\delta_{li} \right] \end{aligned} \tag{33}$$

with the APSG 2-RDM being

$$\begin{aligned} \Gamma^{\text{APSG}}_{jlik} \; = \; \sum_{\sigma\sigma'} \langle \text{APSG}|i^+_\sigma k^+_{\sigma'} l^-_{\sigma'} j^-_\sigma|\text{APSG}\rangle = 2n_i n_k(1-\delta_{\gamma_i\gamma_k})(2\delta_{ij}\delta_{kl} - \delta_{il}\delta_{kj}) \\ + \; 2C_i C_j \delta_{\gamma_i\gamma_j}\delta_{ik}\delta_{jl} \; . \end{aligned} \tag{34}$$

Interestingly the sum relation between one- and two-particle RDM's

$$\sum_l \Gamma^{\text{ERPA-APSG}}_{jlil} = (N-1) \; D_{ji} \tag{35}$$

also persists, as the correction term of $\boldsymbol{\Gamma}^{\text{ERPA-APSG}}$ beyond APSG has no contribution to the partial trace. The 1-RDM behind ERPA-APSG accordingly

agrees with that of APSG, either based on Eq.(32) or on Eq.(35). In the above N stands for the number of electrons in the system, $\mathbf{D} = 2\mathbf{P}$ denotes the spin-summed 1-RDM and spin summation is implied in $\boldsymbol{\Gamma}^{\text{ERPA-APSG}}$ as well.

2.3 APSG based LCC

An alternative idea of deriving a correction to APSG is based on the multireference extension of LCC[59] or, the simplest variant of the coupled pair approximation (CEPA0)[60] or, perturbation theory (PT) with suitable partitioning[33,61–63]. Based on different grounds, these approaches arrive to the common expression of the corrected wavefunction

$$|\Psi\rangle = |\text{APSG}\rangle \; + \; \sum_{K,L}{}' |\Phi_K\rangle G_{KL} \langle\Phi_L|\hat{H}|\text{APSG}\rangle \; , \tag{36}$$

with $\mathbf{G}$ related to

$$M_{KL} = \langle\Phi_K|E_{APSG} - \hat{H}|\Phi_L\rangle \tag{37}$$

as

$$\mathbf{G} = \mathbf{M}^{-1} \; . \tag{38}$$

Functions Φ_K are assumed orthogonal to $|\text{APSG}\rangle$ (hence the prime on the sum in Eq.(36)) and orthonormal among themselves. The energy correction is expressed by substituting Ψ in the non symmetric formula

$$E_{LCC} = \langle\text{APSG}|\hat{H}|\Psi\rangle \; = \; \sum_{K,L}{}' \langle\text{APSG}|\hat{H}|\Phi_K\rangle G_{KL} \langle\Phi_L|\hat{H}|\text{APSG}\rangle \; . \tag{39}$$

When generating excited functions Φ_K it is sufficient to consider a subspace of the CI space, interacting with APSG via the Hamiltonian. Exploiting the

geminal structure, one can identify excited functions with or without charge transfer among geminal subspaces. In the case where electron transfer occurs between geminals, determinants serve as Φ_K . Excited functions conserving the number of electrons in geminal subspaces involved in APSG are necessarily multideterminantal, to satisfy orthogonality to APSG. These states are generated by Mayer's orthogonalization,[64,65] performed in the subspace of determinants contributing to APSG. The LCC method applied here agrees with previous formulations[33,48]. Representation of the reference differs, as open shell determinants with small but nonzero coefficients were included in Ref.[48], while presently a strictly natural orbital expansion is assumed.

In order to determine atomic spin, density matrices are obtained from the LCC corrected wavefunction as

$$D_{ji}^{\text{LCC-APSG}} = \langle\Psi|E_j^i|\Psi\rangle\langle\Psi|\Psi\rangle^{-1} \tag{40}$$

$$\Gamma_{jlik}^{\text{LCC-APSG}} = \sum_{\sigma\sigma'}\langle\Psi|i_\sigma^+ k_{\sigma'}^+ l_{\sigma'}^- j_\sigma^-|\Psi\rangle\langle\Psi|\Psi\rangle^{-1} \ . \tag{41}$$

2.4 Triplet geminals

The aim of the present study is to compare the effect of triplet geminals in the two correction schemes: ERPA-APSG and LCC-APSG. As shown in Section 2.1 triplet state of geminals involved in APSG are essential to obtain correct fragment spin in a multiple bond breaking procedure. In connection with LCC-APSG it has been found that triplet geminals may play a special role.[48] In certain cases they need to be treated in advance, as part of strong correlation, in order not to spoil the second correction step, incorporating weak correlation.

Here we are mainly concerned with the question whether a similar effect can be identified with ERPA-APSG.

Appearance of the particular triplet geminals in the two correction schemes deserves a brief remark. In case of LCC-APSG they appear among non charge transfer excited states, i.e. among Φ_K's where the number of electrons in geminal subspaces is the same as in APSG. Since these states are generated by Mayer's orthogonalization,[64,65] presence of triplet geminals is not immediately apparent. One can however rely on the manifest unitary invariance of the LCC scheme and assume that the space of non charge transfer excitations, interacting with APSG, is spanned by geminal product functions. These are expressible in the form[1]

$$\psi^+_{\theta^*}\,\psi^+_{\eta^*}\prod_{\gamma\neq\theta,\eta}\psi^+_\gamma|\mathrm{vac}\rangle\ . \tag{42}$$

In the above, θ^* refer to an excited state of geminal θ. Appearance of triplet geminals is obvious in this form, since excited states θ^* and η^* comprise singlet and triplet, $M_S=-1,0,+1$ for the latter.

Presence of triplet geminals behind $E^{\text{ERPA-APSG}}$ follows immediately from $\hat{O}^\dagger_\nu$ conserving overall spin but not the spin of individual geminals. Tracking the contribution of a particular triplet geminal is more subtle. Though solely single excitations are generated by the ERPA excitation operator Eq.(21), a correspondence with the double excitation involving ring coupled cluster (rCCD) formalism indeed exists. Interest in this topic has been reviving re-

[1] Intrageminal excitations within a single geminal are non-interacting with APSG[17] they can therefore be ignored.

cently,[66,67] with valuable contributions made by János Ángyán and coworkers.[68,69] In the single determinantal limiting case, formulated with spin-orbitals RPA and rCCD amplitude equations can be exactly matched, with $\sum_\nu (Y_\nu)_{ia}(X_\nu^{-1})_{jb}$ providing the coupled cluster amplitude of the double excitation $i,j \to a,b$. Extension of the correspondence between RPA and rCCD to the multireference case is less trivial.[32] Moreover, ERPA-APSG operates with an adiabatic connection inspired energy formula[70–72] instead of a coupled cluster type expression. Going into details is out of the scope of this study. We suffice stating that triplet geminals are incorporated in ERPA-APSG via the correction term involving a quadratic expression in $\mathbf{X}$ and $\mathbf{Y}$. In particular, it is an intergeminal excitation amplitude e.g. from geminal θ to η multiplied by an amplitude corresponding to a transition from η to θ that involves the contribution of triplet geminals ψ_{θ^*} and ψ_{η^*}.

3 Illustrative examples

3.1 Symmetric dissociation of H_2O

In order to allow comparison of ERPA and LCC corrected APSG energetic data, a minimalistic GVB reference was constructed, with 2 orbitals assigned to the dissociating OH bonds. This model, labeled (OH corr) in Fig.1 corresponds to the geminal assignation of Section 2.1. Potential energy curves presented in Fig.1 show a considerable improvement over APSG by either ERPA or LCC correction, as compared to Full CI (FCI). Fig.1(b) focuses on the corrected

curves beyond cca. twice the equilibrium geometry. Incorrect shape of LCC-APSG is a manifestation of the breakdown of the PP approximation[48]. [2] Curves of ERPA-APSG are free from any ill effect as apparent in Fig.1. Orbital pairs admitted in Eq.(21) for ERPA-APSG notably do not include intergeminal OH excitations since they violate condition $n_i/n_j \leq 0.99$ beyond cca. 4.5 Å .

In case of (OH corr) further orbital pairs have to be excluded from Eq.(21) due to near instability of the reference. Another interpretation of the problem is that lone pair to bonding geminal excitations are not well balanced, bonding geminals being correlated but lone pairs not. This results in two ERPA roots of exceedingly small excitation energy, with main contributions from elementary excitations between the lone pair of π symmetry and natural orbitals of the OH geminals. There are four such orbital pairs, which are excluded in addition to intergeminal pairs in results labeled "ERPA-APSG(OH corr)". Interestingly, the problematic ERPA roots exhibit a local character beyond cca 1.5 Å when performing the same analysis with Boys localized orbitals on OH geminals (instead of naturals). [3] Dominant contributors in this expansion are elementary excitations between the lone pair of π symmetry and the Boys orbitals localized on atom oxygen. There are two such orbital pairs, excluded in results labeled

[2] The ill-effect is less drastic here than in Ref.[48], where divergence of LCC-APSG was reported. The difference originates in open shell components of the reference being present in Ref.[48], see the comment at the end of Section 2.3.

[3] Note that the ERPA-APSG energy is manifestly invariant to a rotation of orbitals which conserves the orbital pairing admitted in Eq.(21). Transformation of amplitudes $(X_\nu)_{mn}$ and $(Y_\nu)_{mn}$ follows from the transformation of orbitals.

"ERPA-APSG(OH corr), Boys", in addition to intergeminal pairs. Due to strong mixing with other pairs, the "ERPA-APSG(OH corr), Boys" curve can not be continued below cca 1.5 Å .

As Fig.1(b) shows, exclusion based on Boys localized OH orbitals improves upon "ERPA-APSG(OH corr)" on the order of 10 mE_h. Two elementary excitations are responsible for this effect, occurring between the lone pair of π symmetry and the Boys orbitals on atoms hydrogen. These are omitted in "ERPA-APSG(OH corr)" but kept in "ERPA-APSG(OH corr), Boys".

Potential curves obtained with all valence geminals assigned two orbitals are labeled (valence corr) in Fig. 1. This brings an improvement on the order of tens of mE_h at the APSG level, but only a a couple of mE_h at the level of ERPA-APSG. Near instability not affecting the (valence corr) APSG calculation, orbital pair exclusion resorts to intergeminal OH excitations, which is inevitable. Though not warranted based on instability arguments, elimination of four additional orbital pairs is also checked, labeled by "ERPA-APSG(valence corr), Boys". Lone pairs being correlated, the corresponding orbitals are localized rather than symmetric. For this reason both lone pairs are considered when discarding excitations connecting lone pairs and the Boys orbitals of OH geminals localized on atom oxygen. This makes four orbital pairs instead of two, observed with the (OH corr) scheme. Elimination of these four additional orbital pairs results in a considerable energy lowering, getting below FCI by tens of mE_h's as reflected by Fig. 1(b).

Atomic spin, plotted in Fig.2 shows a marked difference between ERPA and LCC corrected APSG. In accordance with Ref.[48], LCC improves only up to a given OH bond distance but fails to set right the qualitative defect of APSG upon further bond elongation. At difference with LCC-APSG, ERPA-APSG does not improve upon the local spin of APSG in any region, it essentially follows the trend of APSG both for the (OH corr) and the (valence corr) scheme, c.f. panel (a) and (b) respectively. Based on Section 2.4, OH intergeminal excitations missing from ERPA-APSG curves depicted in Fig.2 can be expected to have a beneficial effect on local spin. Their correction in fact points in the right direction, if admitted. For example at 3 Å for (valence corr) one gets the value 1.55 for $\langle \hat{S}^2 \rangle_O$, instead of 1.48 plotted in Fig. 2(b). The improvement is modest in view of the FCI result and even this has to be discarded in favour of a continuous potential surface.

There is a remarkable difference in atomic spin values depicted in Fig. 2(a) between ERPA-APSG(OH corr) and "ERPA-APSG(OH corr), Boys", the latter setting local spin essentially right in the dissociation limit. A qualitative picture, rationalizing this effect can be given based on the cartoons in Fig. 4. Taking into account that the square of excitation amplitudes figure in the ERPA-APSG correction, one can vaguely associate product of single excitation amplitudes with doubles' amplitude. For example, if the π lone pair to oxygen sp^2 hybrid excitation is applied twice a Lewis structure depicted in Fig. 4(b) can be generated from the APSG structure of Fig. 4(a). The structure of Fig. 4(b) corresponds to low spin on atom oxygen, the effect of which is missing

both from ERPA-APSG(OH corr) and from "ERPA-APSG(OH corr), Boys". Focusing now on the π lone pair to hydrogen s excitation, its combination with the sp^2 one pair to hydrogen s excitation results in the structure of Fig. 4(c). The latter corresponds to a high spin on atom oxygen, an effect accounted for by "ERPA-APSG(OH corr), Boys" but not by ERPA-APSG(OH corr). Based on this coarse picture, oxygen spin lowering and increasing effects can be understood to compensate in ERPA-APSG(OH corr) while only local spin increasing effect is kept in "ERPA-APSG(OH corr), Boys". Significant overshooting in local spin by "ERPA-APSG(valence corr), Boys" observable in Fig. 2(b) can be rationalized along the same lines. In this case both lone pairs give rise to a structure of the type of Fig. 4(b), implying low spin on atom oxygen and being eliminated from the ERPA-APSG 2-RDM.

Correction in atomic spin by "ERPA-APSG(OH corr), Boys" comes at the price of considerable violation of total spin, as reflected in Fig.3. In the "ERPA-APSG(OH corr), Boys" scheme $\langle \hat{S}^2 \rangle$ rises from the value of $-2.9 \cdot 10^{-3}$ at 1.5 Å to 0.45 at 7.0 Å. The effect is exaggerated in "ERPA-APSG(valence corr), Boys", $\langle \hat{S}^2 \rangle$ starting from the value of 0.66 at 2.0 Å and reaching 0.99 at 7.0 Å. Though not strictly zero, spin contamination of ERPA-APSG remains relatively small, as illustrated by Fig.3. In terms of numbers $\langle \hat{S}^2 \rangle$ by "ERPA-APSG(OH corr)" remains between -0.66 and -0.53 in the distance range shown while the limits for $\langle \hat{S}^2 \rangle$ by "ERPA-APSG(valence corr)" are $-3.8 \cdot 10^{-2}$ and $-8.9 \cdot 10^{-3}$. Curves of APSG, LCC-APSG and FCI are omitted from Fig.3 for clarity, as they all lie on the zero axis.

Spin components by APSG, ERPA-APSG and FCI are collected in Table 3.1 at two points in the stretched regime to complete the picture. From Table 3.1 one can deduce that the dominant effect of ERPA appears in the spin of oxygen at both geometries shown. Interatomic components and $\langle \hat{S}^2 \rangle_H$ are slightly modified by ERPA at 1.5 Å and practically not at 7.0 Å OH distance, both for "(OH corr)" and for "(OH corr), Boys". An overcorrection of $\langle \hat{S}^2 \rangle_O$ by "ERPA-APSG, Boys" is brought about at 1.5 Å while it is essentially set right at 7.0 Å . The "ERPA-APSG, Boys" scheme affecting other spin components at 1.5 Å is a consequence of the non negligible contribution of H atomic orbitals to the Boys orbitals at this distance. While total spin violation is relatively small with ERPA-APSG, there is nothing to compensate the increased oxygen spin in the "ERPA-APSG, Boys" scheme especially at 7.0 Å OH distance. As a consequence total spin of "ERPA-APSG, Boys" gets deteriorated with roughly the same amount by which $\langle \hat{S}^2 \rangle_O$ gets ameliorated. This is apparent from the comparison of Figs.2 and 3 also.

3.2 Double bond dissociation of C_2H_4

Energetic data are again compared on a small basis example, allowing the computation of LCC-APSG and FCI. Only the dissociating CC bonds are correlated at the level of APSG, with two orbitals assigned to each, all other geminal subspaces are kept one dimensional. Total energy profiles, presented in Fig.5(a) show that both ERPA and LCC improve upon APSG considerably. Panel (a) of Fig.5 shows results with symmetry adapted orbitals constituting

	R(OH) = 1.5 Å			R(OH) = 5.5 Å		
	O	H_1	H_2	O	H_1	H_2
APSG				APSG		
O	0.426	-0.213	-0.213	1.500	-0.750	-0.750
H_1		0.213	0.000		0.750	0.000
H_2			0.213			0.750
ERPA-APSG(OH corr)				ERPA-APSG(OH corr)		
O	0.357	-0.208	-0.208	1.451	-0.751	-0.751
H_1		0.214	-0.010		0.750	0.000
H_2			0.214			0.750
ERPA-APSG(OH corr), Boys				ERPA-APSG(OH corr), Boys		
O	0.582	-0.245	-0.245	1.951	-0.751	-0.751
H_1		0.214	-0.017		0.750	0.000
H_2			0.214			0.750
FCI				FCI		
O	0.402	-0.201	-0.201	2.000	-1.000	-1.000
H_1		0.191	0.010		0.750	0.250
H_2			0.191			0.750

Table 1 Atomic $\langle \hat{S}^2 \rangle_A$ and diatomic components $\langle \hat{S}^2 \rangle_{AB}$ for the water molecule in 6-31G* basis at $\angle(HOH) = 109.5^o$. See table heading for R(OH) . Only OH geminals are correlated at the APSG level, with 2 orbitals assigned to each. Benchmark is provided by FCI. See text for elementary excitations participating in Eq.(21) of ERPA-APSG schemes.

the CC geminals. Panel (b) of Fig.5 focuses on the intermediate to long distance range, where an APSG solution with symmetry breaking, banana type orbitals on CC geminals could be found. The latter represents the lowest energy

APSG solution beyond cca. 3 Å, getting about 6 mE_h below the symmetry adapted APSG solution at 6 Å.

In parallel with the H_2O symmetric dissociation example, a non physical hump appears on the LCC-APSG curve at around twice the equilibrium distance both in panel (a) and (b) in Fig. 5. The hump is more expressed with banana type orbitals, amounting to cca. 80 mE_h while it is around 7 mE_h with orbitals belonging to irreducible representations (ir. rep.).

The curve by ERPA-APSG is somewhat more off from FCI in Fig. 5(a) than LCC, it is however free from any qualitative failure. Comparison of panels (a) and (b) shows that the ERPA-APSG energy with banana type orbitals on CC geminals is very close to the result with CC geminals built with ir. rep. orbitals. The ERPA-APSG results depicted in Fig. 5 exclude intergeminal CC excitations from Eq.(21) in the interest of a continuous curve. Instability issues were not detected with this model.

Spin of fragment CH_2 is shown in Fig.6 with a full valence complete active space, CAS(12,12) serving as benchmark in lack of RDMs at the level of FCI. It is notable that fragment spin is slightly negative by CAS(12,12) in the compressed bond range, an effect already observed in the family of unsaturated hydrocarbons[56]. [4]

The curve of APSG tending to an incorrect limit in Fig.6 is analogous to the example of H_2O just like the erratic shape of LCC-APSG. With symmetry

[4] Turning to Löwdin orthogonalized atomic orbitals or a decomposition scheme in 3D space has been found to fix negative atomic spin in Ref.[56]. This investigation is not pursued here our main interest lying with the spin of fragments formed upon dissociation.

adapted orbitals LCC-APSG clearly improves local spin up to cca. 2.8 Å, c.f. Fig.6(a) while it is rather poor with banana type orbitals on CC geminals, c.f. Fig.6(b).

The correction in local spin brought about by ERPA-APSG with symmetry adapted orbitals, Fig.6(a), is nonnegligible around equilibrium, ERPA-APSG being closer to the CAS(12,12) local spin than uncorrected APSG. Unfortunately the correction is marginal in the stretched regime also. Including elementary excitations between CC geminals the picture could be slightly improved. The value of $\langle \hat{S}^2 \rangle_{CH_2}$ reaches 1.71 with ir. rep. orbitals on CC geminals and 1.57 with banana orbitals at around 5 Å in Fig.6 if admitting these orbital pairs in Eq.(21). Since these pairs can not be included without causing a discontinuity beyond 5 Å, their effect is deliberately ignored. Total spin by ERPA-APSG, presented in Fig.7 shows small negative values, similarly to the case of H_2O. Total spin with banana orbitals is slightly better, varying in the range of $-9.94 \cdot 10^{-2}$ (at 2.2 Å) and $-8.37 \cdot 10^{-2}$ (at 6.4 Å). Limits for $\langle \hat{S}^2 \rangle$ with ir. rep. orbitals in the range shown are -0.105 (at 1.0 Å) and $-8.78 \cdot 10^{-2}$ (at 6.4 Å). Closing the subject of C_2H_4 we mention that correlating all valence geminals at the APSG level (with 2 orbitals assigned to them), improves the limiting value of local spin slightly. It reaches the value of 1.50 by ERPA-APSG with intergeminal excitations excluded.

3.3 Rectangle to square transformation of H_4

As a final example the rectangle to square transformation of the H_4 system is presented in Fig.8. At difference with the previous examples the challenge of this system is not the triplet recoupling of two spins residing on the fragment formed upon dissociation. Lewis structures implied in this process both involve singlet coupled pairs. The difficulty here lies with the considerable local spin building up on atom H at square geometry, which is poorly described by APSG. Moreover the APSG energy curve exhibits a characteristic cusp at 90^o, as a consequence of switching between two Lewis structures.

Results shown in Fig.8 indicate that both LCC and ERPA bring a considerable correction in energy. Exclusion of intergeminal elementary excitations is not an issue with ERPA in this case, bonds not getting stretched enough for quasi degeneracy in occupation numbers building up around 1/2. The FCI result is approximated better by LCC than by ERPA, both in absolute value and regarding curve shape, but neither succeed in correcting the cusp at 90^o. As Fig.8(a) shows, ERPA-APSG produces a cusp similar to APSG, while LCC-APSG exhibits a local minimum of cca. 0.8 mE_h, remaining non-differentiable at 90^o. The peculiar shape of LCC-APSG is a consequence of an intruder (i.e. a small eigenvalue of matrix $\mathbf{M}$) appearing close to square geometry eventually causing a divergence if following the higher energy APSG solution beyond 90^o. Inspecting atomic spin curves, plotted in Fig.8(b) one can observe that LCC-APSG gets the order of magnitude of $\langle \hat{S}^2 \rangle_H$ right while ERPA-APSG remains

in the same range as APSG. Similarly to total energy, local spin curves of both LCC and ERPA inherit the cusp at 90^o.

4 Conclusion

Comparison of LCC and ERPA correction to APSG on the example of breaking multiple bonds or single bonds attached to the same atom allows to draw the following conclusions. Incorrect description of spins recoupling to triplet on the dissociation product does not affect potential energy curves by APSG or by ERPA-APSG but causes a defective shape of LCC-APSG curves. Special, separate treatment of triplet geminals has been previously shown to be a remedy in case of LCC.

With ERPA-APSG the way of approach in a multiple bond breaking situation is more simple. One only has to consider that elementary excitations between dissociating geminals may become (close to) zero norm during the process, due to (quasi) degeneracy in orbital occupation number building up around the value of 1/2. The corresponding orbital pairs have to be excluded from the expansion of the excitation operator in the entire geometry range studied, in order to get a continuous potential surface.

Fragment spin by ERPA-APSG follows closely the result of APSG. The ERPA scheme resorting to spin conserving single excitations, triplet geminals necessary for curing atomic spin can appear via squared amplitudes of elementary excitations between dissociating geminals. These constituents are however neglected for avoiding discontinuity. Inspecting the numerical effect of

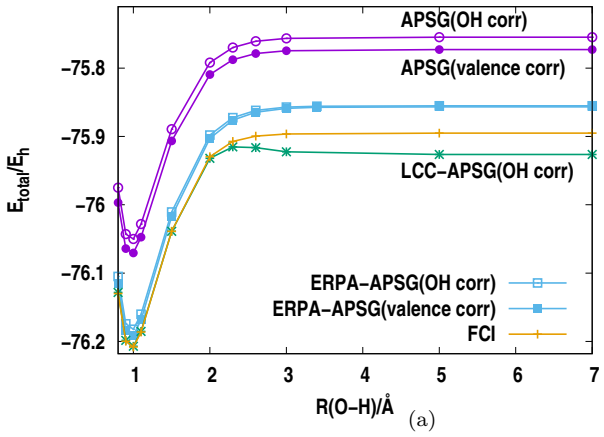
intergeminal elementary excitations on atomic spin at geometries where they are still admissible only a marginal improvement is revealed.

Omitting intrafragment excitations, based on Boys localized orbitals is found to produce remarkably good fragment spin in the case of the H_2O molecule with OH geminals correlated only at the level of the reference. Unfortunately the source appears to be a fortuitous incompensation in local spin. Moreover, this approach brings considerable violation in total spin while that of ERPA-APSG remained below 0.1 in absolute value in the examples studied.

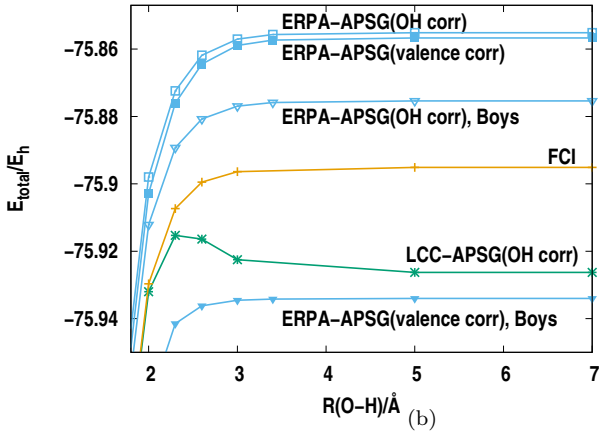
Acknowledgements Discussions with prof. Péter R. Surján (Budapest) are gratefully acknowledged. The work has been supported by the National Science Center of Poland under Grant No. 2016/23/B/ST4/02848 and by the National Research, Development and Innovation Office of Hungary (NKFIH), under grant number K115744. Acknowledgement is expressed by Á.Sz. for the János Bolyai Research Scholarship provided the Hungarian Academy of Sciences (MTA). Support of the New National Excellence Program (ÚNKP) provided by the Ministry of Human Capacities of Hungary (EMMI) is acknowledged by Á.M. The FCI results and density matrices were computed by an implementation of Olsen's FCI algorithm[73], written by Zoltán Rolik (Budapest University of Technology and Economics).

This is a post-peer-review, pre-copyedit version of an article published in Theoretical Chemistry Accounts. The final authenticated version is available online at: http://dx.doi.org/10.1007/s00214-018-2355-4 .

5 Figures

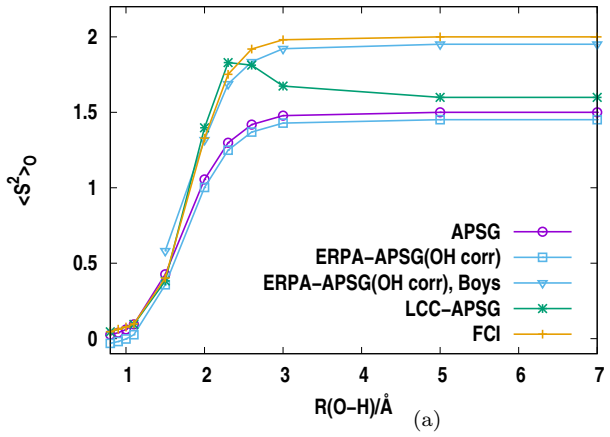


(a)

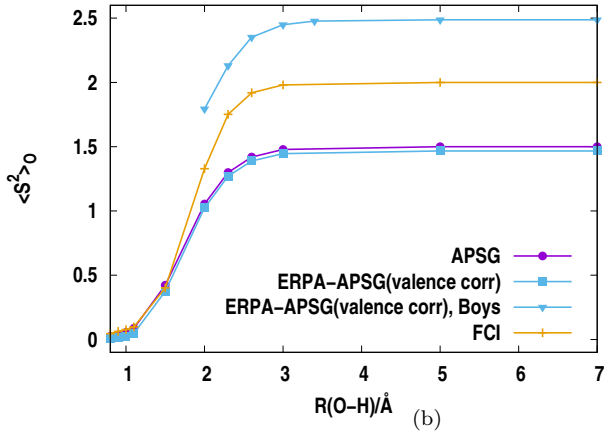


(b)

Fig. 1 Symmetric dissociation of the water molecule in 6-31G* basis, $\angle(HOH) = 109.5^o$. Total energy by APSG, ERPA-APSG and LCC-APSG. Full CI (FCI) is shown as benchmark. Label 'APSG(OH corr)' refers to bonding OH geminals being correlated only at the APSG level, with 2 orbitals assigned to each bond. All valence geminals are assigned 2 orbitals in the model 'APSG(valence corr)'. See text for elementary excitations participating in Eq.(21) of ERPA-APSG schemes.



(a)



(b)

Fig. 2 Symmetric dissociation of the water molecule in 6-31G* basis, $\angle(HOH) = 109.5^o$. Spin-squared expectation value of atom oxygen by APSG, ERPA-APSG and LCC-APSG. Benchmark is provided by FCI. Results obtained with OH geminals correlated only in the reference (APSG(OH corr)) are shown in panel (a). Panel (b) depicts results based on all valence geminals correlated in the reference (APSG(valence corr)). See text for elementary excitations participating in Eq.(21) of ERPA-APSG schemes.

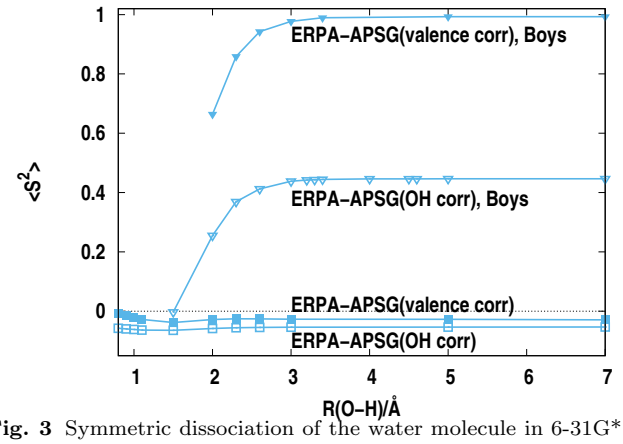


Fig. 3 Symmetric dissociation of the water molecule in 6-31G* basis, $\angle(HOH) = 109.5^o$. Spin-squared expectation value by ERPA-APSG. Acronym APSG(OH corr) refers to OH geminals correlated only in the reference. All valence geminals are correlated in the reference for APSG(valence corr). See text for elementary excitations participating in Eq.(21) of ERPA-APSG schemes.

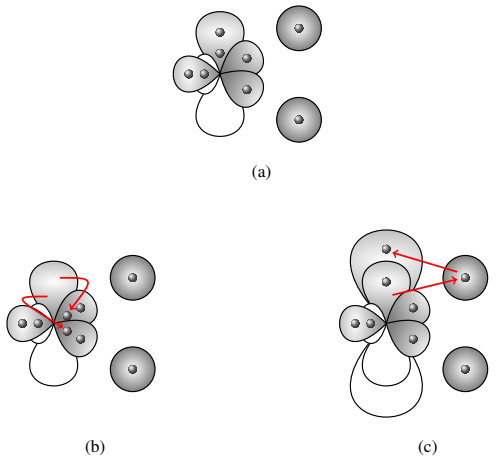


Fig. 4 Lewis-structures implied by ERPA-APSG schemes for the water molecule, both OH bonds stretched. Orbitals on OH geminals are assumed localized. Panel (a) depicts the structure involved by APSG. Panel (b) shows a structure implied by ERPA-APSG but not by "ERPA-APSG, Boys". Structure of panel (c) is involved both in ERPA-APSG and "ERPA-APSG, Boys".

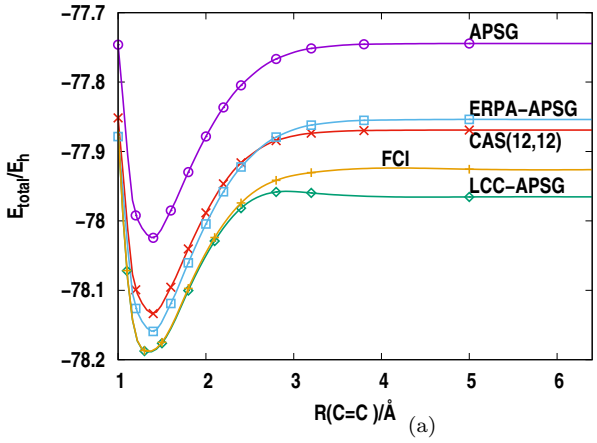
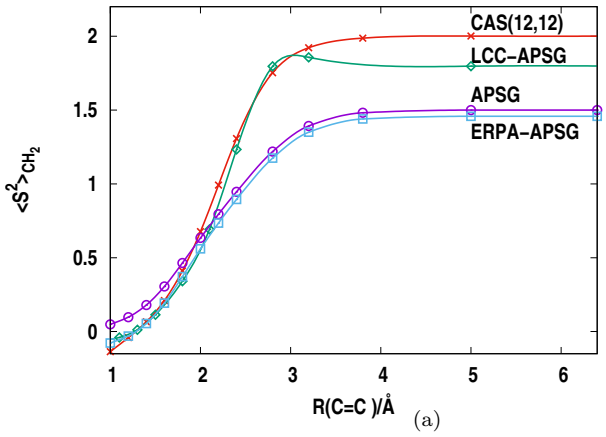
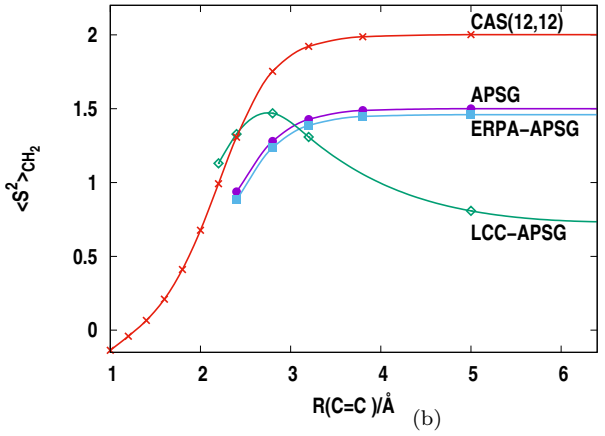


Fig. 5 C=C bond stretching of the ethylene molecule with 6-31G basis on atom C and STO-3G on atom H, $\angle(HCH) = 116.6^o$, R(C-H)=1.076 Å . Total energy by APSG, ERPA-APSG and LCC-APSG. The only geminals correlated at the APSG level are C=C bonds, with two orbitals assigned to each. For comparison FCI and CAS(12,12) are displayed. Core orbitals are frozen in LCC-APSG and FCI. Open symbols in panel (a) correspond to C=C bonding geminals build of σ or π orbitals. Filled symbols, shown in panel (b) indicate the solution with two equivalent C=C bonding geminals, built of $\sigma - \pi$ mixed, banana type orbitals. See text for elementary excitations participating in Eq.(21) of ERPA-APSG schemes.



(a)



(b)

Fig. 6 C=C bond stretching of the ethylene molecule with 6-31G basis on atom C and STO-3G on atom H, $\angle(HCH) = 116.6^o$, R(C-H)=1.076 Å . Spin-squared expectation value of fragment CH_2 by APSG, ERPA-APSG and LCC-APSG. The only geminals correlated at the APSG level are C=C bonds, with two orbitals assigned to each. For comparison CAS(12,12) is shown. Open symbols in panel (a) correspond to C=C bonding geminals built of σ or π orbitals. Filled symbols in panel (b) indicate the solution with two equivalent C=C bonding geminals, built of $\sigma - \pi$ mixed, banana type orbitals. See text for elementary excitations participating in Eq.(21) of ERPA-APSG schemes.

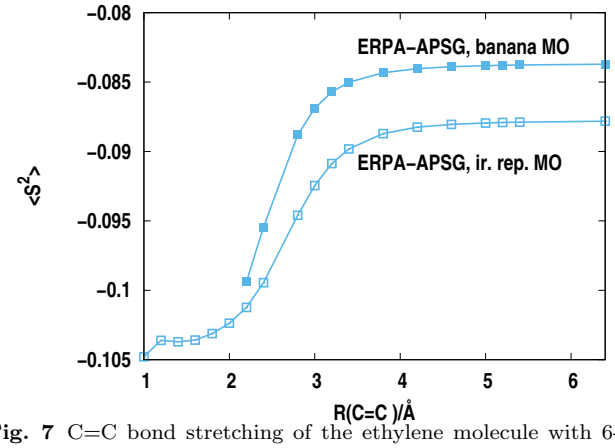
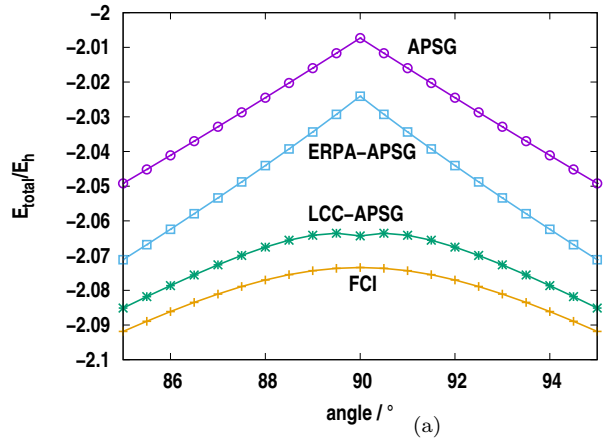
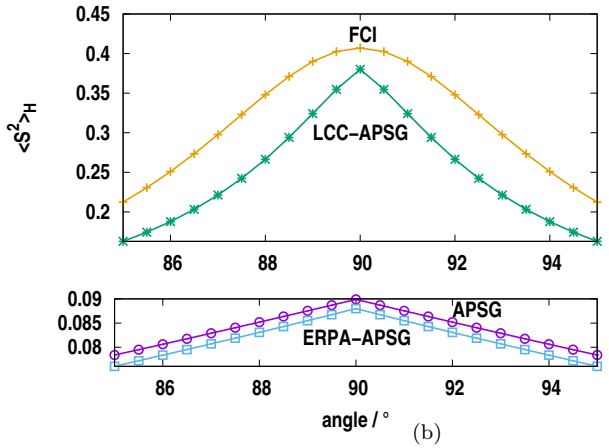


Fig. 7 C=C bond stretching of the ethylene molecule with 6-31G basis on atom C and STO-3G on atom H, $\angle(HCH) = 116.6^o$, R(C-H)=1.076 Å . Spin-squared expectation value by ERPA-APSG. The only geminals correlated at the APSG level are C=C bonds, with two orbitals assigned to each. Open symbols correspond to C=C bonding geminals build of σ or π orbitals. Filled symbols indicate the solution with two equivalent C=C bonding geminals, built of $\sigma - \pi$ mixed, banana type orbitals. See text for elementary excitations participating in Eq.(21) of ERPA-APSG schemes.



(a)



(b)

Fig. 8 Rectangle to square transformation of H_4 in 6-31G** basis. Atoms H move on a circle of 0.75 Å radius, angle of two H atoms, viewed from the origin is displayed on axis x. Total energy (a) and spin of atom H (b) by APSG, ERPA-APSG and LCC-APSG. Benchmark is provided by FCI.

References

1. P.E.M. Siegbahn, J. Almlöf, A. Heiberg, B.O. Roos, J. Chem. Phys. **74**, 2384 (1981)
2. K.D. Vogiatzis, D. Ma, J. Olsen, L. Gagliardi, W.A. de Jong, J. Chem. Phys. **147**, 184111 (2017)
3. G.K.L. Chan, M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002)
4. S. Keller, M. Dolfi, M. Troyer, M. Reiher, J. Chem. Phys. **143**, 244118 (2015)
5. G. Li Manni, S.D. Smart, A. Alavi, J. Chem. Theory Comput. **12**, 1245 (2016)
6. P.A. Malmqvist, A. Rendell, B.O. Roos, J. Phys. Chem. **94**, 5477 (1990)
7. D. Ma, G.L. Manni, L. Gagliardi, J. Chem. Phys. **135**, 044128 (2011)
8. J. Ivanic, J. Chem. Phys. **119**, 9364 (2003)
9. F.W. Bobrowicz, W.A. Goddard-III, in *Methods of Electronic Structure Theory*, ed. by H.F. Schaefer-III (Plenum, New York, 1977), p. 79
10. W. Kutzelnigg, in: *Modern Theoretical Chemistry, The Methods of Electronic Structure Theory* (Plenum, New York, 1977), vol. 3, p. 129
11. K.V. Lawler, D.W. Small, M. Head-Gordon, J. Phys. Chem. A **114**, 2930 (2010)
12. T.V. Voorhis, M. Head-Gordon, J. Chem. Phys. **112**, 5633 (2000)
13. V.A. Rassolov, J. Chem. Phys. **117**, 5978 (2002)
14. V.A. Rassolov, F. Xu, J. Chem. Phys. **127**, 044104 (2007)
15. P.A. Johnson, P.W. Ayers, P.A. Limacher, S.D. Baerdemacker, D.V. Neck, P. Bultinck, Comput. Theor. Chem. **1003**, 101 (2013)
16. P.A. Limacher, P.W. Ayers, P.A. Johnson, S. De Baerdemacker, D. Van Neck, P. Bultinck, J. Chem. Theory Comput. **9**, 1394 (2013)
17. P.R. Surján, Topics in current chemistry **203**, 63 (1999)
18. T. Arai, J. Chem. Phys. **33**, 95 (1960)
19. I.I. Ukrainskii, Theoret. Math. Phys. **32**, 816 (1978)
20. M. Tarumi, M. Kobayashi, H. Nakai, Int. J. Quantum Chem. **113**, 239 (2013)
21. K. Chatterjee, Theor. Chem. Acc. **135**, 246 (2016)
22. J.J. McDouall, M.A. Robb, Chem. Phys. Letters **132**, 319 (1986)
23. R.B. Murphy, R.P. Messmer, J. Chem. Phys. **98**, 7958 (1993)

24. S.O. Odoh, G.L. Manni, R.K. Carlson, D.G. Truhlar, L. Gagliardi, Chem. Sci. **7**, 2399 (2016)
25. S. Keller, K. Boguslawski, T. Janowski, M. Reiher, P. Pulay, J. Chem. Phys. **142**, 244104 (2015)
26. G.J.O. Beran, M. Head-Gordon, S.R. Gwaltney, J. Chem. Phys. **124**(11), 114107 (2006)
27. J.A. Parkhill, M. Head-Gordon, J. Chem. Phys. **133**(12), 124102 (2010)
28. E. Rosta, P.R. Surján, J. Chem. Phys. **116**, 878 (2002)
29. E. Xu, S. Li, J. Chem. Phys. **139**, 174111 (2013)
30. S. Li, J. Ma, Y. Jiang, J. Chem. Phys. **118**, 5736 (2003)
31. V.A. Rassolov, F. Xu, S. Garashchuk, J. Chem. Phys. **120**, 10385 (2004)
32. Á. Szabados, Á. Margócsy, Mol. Phys. **115**, 2731 (2017)
33. T. Zoboki, Á. Szabados, P.R. Surján, J. Chem. Theory Comput. **9**, 2602 (2013)
34. P.A. Limacher, P.W. Ayers, P.A. Johnson, S. De Baerdemacker, D.V. Neck, P. Bultinck, Phys. Chem. Chem. Phys. **16**, 5061 (2014)
35. K. Pernal, J. Chem. Theory Comput. **10**, 4332 (2014)
36. M. Piris, X. Lopez, F. Ruipérez, J.M. Matxain, J.M. Ugalde, J. Chem. Phys. **134**, 164102 (2011)
37. M. Piris, J.M. Matxain, X. Lopez, J. Chem. Phys. **139**, 234109 (2013)
38. K. Pernal, Comput. Theor. Chem. **1003**, 127 (2013)
39. W. Kutzelnigg, Chem. Phys. **401**, 119 (2012)
40. K. Chatterjee, K. Pernal, J. Chem. Phys. **137**, 204109 (2012)
41. E. Pastorczak, K. Pernal, Phys. Chem. Chem. Phys. **17**, 8622 (2015)
42. K. Chatterjee, E. Pastorczak, K. Jawulski, K. Pernal, J. Chem. Phys. **144**, 244111 (2016)
43. P.A. Schultz, R.P. Messmer, J. Amer. Chem. Soc. **115**(23), 10925 (1993)
44. F. Faglioni, W.A. Goddard, Int. J. Quantum Chem. **73**, 1 (1999)
45. V.A. Rassolov, F. Xu, J. Chem. Phys. **126**, 234112 (2007)
46. D.W. Small, K.V. Lawler, M. Head-Gordon, J. Chem. Theory Comput. **10**, 2027 (2014)
47. P. Jeszenszki, P. Surján, Á. Szabados, J. Chem. Theor. Comput. **11**, 30963103 (2015)
48. P. Surján, P. Jeszenszki, Á. Szabados, Mol. Phys. **113**, 2960 (2015)

49. P. Jeszenszki, V. Rassolov, P.R. Surján, Á. Szabados, Mol. Phys. **113**(3-4), 249 (2015)
50. E. Ramos-Cordoba, P. Salvador, M. Piris, E. Matito, J. Chem. Phys. **141**, 234101 (2014)
51. A.E. Clark, E.R. Davidson, J. Chem. Phys. **115**, 7382 (2001)
52. A.E. Clark, E.R. Davidson, J. Phys. Chem. A **106**, 6890 (2002)
53. I. Mayer, Faraday Discuss. **135**, 146 (2007)
54. D. Alcoba, A. Torre, L. Lain, R. Bochicchio, J. Chem. Theory Comput. **7**, 3560 (2011)
55. E. Ramos-Cordoba, E. Matito, I. Mayer, P. Salvador, J. Chem. Theory Comput. **8**, 1270 (2012)
56. E. Ramos-Cordoba, E. Matito, I. Mayer, P. Salvador, Phys. Chem. Chem. Phys. **14**, 15291 (2012)
57. D. Rowe, Rev. Mod. Phys **40**, 153 (1968)
58. J.T. Golab, D.L. Yeager, P. Jørgensen, Chem. Phys. **78**, 175 (1983)
59. W.D. Laidig, P. Saxe, R.J. Bartlett, J. Chem. Phys. **86**, 887 (1987)
60. W. Kutzelnigg, Chem. Phys. Letters **35**, 283 (1975)
61. P. R. Surján and Á. Szabados, J. Chem. Phys. **112**, 4438 (2000)
62. H.A. Witek, H. Nakano, K. Hirao, J. Chem. Phys. **118**, 8197 (2003)
63. R.F. Fink, Chem. Phys. **356**, 39 (2009)
64. I. Mayer, Theor. Chem. Acc. **104**, 163 (2000)
65. P. Nagy, P. Surján, Á. Szabados, Theor. Chem. Acc. **131**, 1109 (2012)
66. G.E. Scuseria, T.M. Henderson, D.C. Sorensen, J. Chem. Phys. **129**, 231101 (2008)
67. X. Ren, P. Rinke, C. Joas, M. Scheffler, J. Mater. Sci. **47**, 7447 (2012)
68. G. Jansen, R.F. Liu, J.G. Ángyán, J. Chem. Phys. **133**, 154106 (2010)
69. J. Toulouse, W. Zhu, A. Savin, G. Jansen, J.G. Ángyán, J. Chem. Phys. **135**, 084119 (2011)
70. J.G. Ángyán, R.F. Liu, J. Toulouse, G. Jansen, J. Chem. Theory Comput. **7**, 3116 (2011)
71. K. Pernal, Int. J. Quantum Chem. **118**, e25462 (2017)
72. K. Pernal, Phys. Rev. Lett. **120**, 013001 (2018)
73. J. Olsen, B.O. Roos, P. Jørgensen, H.J.A. Jensen, J. Chem. Phys. **89**, 2185 (1988)