RESEARCH ARTICLE

A basis set superposition error-free second-order perturbation theory from Hermitian chemical Hamiltonian approach self-consistent field canonic orbitals

Pedro Salvador[1] [ID] | István Mayer[2†]

[1]Institute of Computational Chemistry and Catalysis and Department of Chemistry, University of Girona, Girona, Spain

[2]Chemical Research Center, Hungarian Academy of Sciences, Budapest, Hungary

Correspondence
Pedro Salvador, Institute of Computational Chemistry and Catalysis and Department of Chemistry, University of Girona, c/M Aurelia Capmany 69, 17003 Girona, Spain
Email: pedro.salvador@udg.edu

Abstract

We present an alternative perturbational approach free of basis set superposition error (BSSE) within the framework of the chemical Hamiltonian approach (CHA). The new formulation (CHA-S-MP2) is based on canonic (and orthogonal) CHA orbitals obtained from a hermitized CHA Fock operator. The final expression shows a considerable simplification of the method as compared to the previous CHA-MP2 formalism. Numerical full geometry optimizations of water and hydrogen fluoride dimers and potential energy surfaces for helium and argon dimers for several basis sets are presented. The present method is compared to both the counterpoise and previous CHA-MP2 BSSE correction schemes, showing a remarkable agreement between all three methods. However, the wrong behavior using the aug-cc-pVDZ basis set indicates that the present method is not as robust as the original non-hermitian CHA-MP2 formulation.

KEYWORDS
basis set superposition error, chemical Hamiltonian approach, water dimer

1 | INTRODUCTION

István Mayer's legacy in the field of quantum chemistry is overwhelming. While being most known for the general expression of bond order and valences in non-orthogonal basis, his paper "Towards a Chemical Hamiltonian" published in this journal almost 40 years ago remain in my opinion one of his most memorable achievements [1]. In that study he paves the road for an energy decomposition into genuine chemical terms (i.e., as interaction between atoms occurring within the molecule), which gained much interest many years later, including both its Hilbert-space [2] and real-space [3] realizations. That was the main focus of István's interests, namely extracting chemical information from well-founded wavefunction analysis techniques. Curiously enough, the so-called chemical Hamiltonian approach (CHA) soon reached particular interest in the field of inter-molecular interactions, as it represented a novel way of dealing with the basis set superposition error (BSSE) [4] in an a priori fashion. So, perhaps deviating from the original purpose, István's research focused for quite a while on developing the CHA-version of different electronic structure methods to achieve BSSE-free description of intermolecular complexes [5] (and even some early incursions into the less known intramolecular BSSE problems) [6, 7]. Thus, starting from the CHA-self consistent field (SCF) method [8, 9], a full-CI version [10] was derived and successfully tested, followed by a delicate perturbational treatment leading to the CHA-MP2 scheme [11, 12].

The alternative was, and still is, the so-called counterpoise (CP) scheme by Boys and Bernardi [13]. In CP, the energy of a complex is corrected from BSSE in an a posteriori fashion, by evaluating the energy lowering of the monomers when borrowing the basic functions of the other moieties (ghost orbitals). At that time, it was shown [12, 14] that (i) both CP and CHA o strategies lead to very similar results in terms of interaction energies and geometrical parameters and (ii) the differences between both approaches are smaller than the BSSE itself.

A mandatory feature of any BSSE-correction method is that the BSSE correction must tend asymptotically to zero when appropriately increasing the basis set size, that is, in the complete basis set limit. It is worth to recall that another a priori BSSE-correction scheme, the so-called SCF for molecular interactions (SCF-MI) [15], failed at this test, as true charge-transfer effects were also removed from the interaction. A reformulation form Nagata et al. [16] that included one-electron corrections overcame this pitfall but this avenue was not further explored, leaving CHA as the only a priori alternative to CP correction.

Both CP and CHA methodologies exhibits correct asymptotics, however, even though it is relatively easy to reach basis set converged results at Hartree–Fock level (or for single-determinant wavefunctions in general), when including electron correlation the situation is quite different. The accumulated experience [17–21] shows that BSSE tends to zero asymptotically but slowly, and hence a proper BSSE correction is mandatory even for very large basis sets. For instance, Wieczorek et al. [18] showed that a residual BSSE (CP-corrected) of 0.23 kcal/mol is still present in a MP2(FC)/aug-cc-pVQZ calculation for the water dimer, whereas at the Hartree–Fock level the BSSE is just 0.02 kcal/mol with the same basis set.

As far as the computational cost is concerned, a conventional CP-corrected calculation of a molecular complex composed by N interacting fragments involves $2N + 1$ calculations, where $N + 1$ of them involve the whole basis set of the molecular system. Therefore, for a dimer, the computational cost of a CP calculation is about three times that of the corresponding conventional uncorrected one. The advantage of the CP method is that any property expressed from the total molecular energy can be potentially CP-corrected. That means analytical gradients and harmonic frequencies can be easily derived (provided they are available for the given level of theory) [22].

On the other hand, the CHA-SCF equations can be efficiently implemented so that the computational cost is virtually independent on the number of monomers and their nature [23]. However, the CHA generalization of the MP2 theory (CHA-MP2) [11,12] involves three four-index transformations so that the computational cost represents again roughly three times that of the conventional E2 calculation. Thus, a single-point CHA-MP2 calculation is computationally equivalent to a CP-MP2 one for a dimer and superior for larger molecular aggregates. However, the lack of analytical gradients and higher derivatives still makes the CP method more efficient for potential energy surface (PES) exploration of molecular systems with large number of degrees of freedom.

The main feature of the CHA methods is that they use to non-hermitian operators. In particular, CHA-SCF uses a non-hermitian Fock operator to obtain the molecular orbitals and a conventional hermitian Fock operator to obtain the energy (the so-called CHA with conventional energy [CHA/CE] recipe [8]). Moreover, the CHA-MP2 method itself proved to be really challenging [11]. Here, a zeroth order Hamiltonian is built up on the not orthogonal (and not necessarily real) canonic CHA-SCF orbitals and their orbital energies. Then, biorthogonal perturbation theory is applied to obtain the first-order wave function correction, while forced by the CHA/CE concept, the second order energy correction is calculated by using the conventional Hamiltonian, by means of a generalized Hylleraas functional [24] to guarantee the energy to be real even in the case of complex CHA-SCF orbitals.

István justified this by pointing out that the BSSE itself is a non-physical phenomenon, hence no hermitian operator is expected to be associated to it. However, two decades after his seminal CHA paper, Mayer showed (again on this journal) that a set of orthogonal CHA-SCF canonic orbitals can be obtained as eigenvectors of a hermitian *effective* Fock matrix [25]. The energy obtained with the Slater determinant built out of those orbitals is *exactly* the same one can get by orthogonalizing the occupied space spawned by the CHA-SCF non-orthogonal eigenvectors of the non-hermitian CHA Fockian. Both sets of molecular orbitals are related trough a unitary transformation and provide the same density matrix (P-matrix). The story behind the paper is charming. He had the opportunity to discuss the CHA-SCF methodology with P. O. Löwdin, while visiting his group in Florida. Lowdin was not satisfied with the fact the theory used non-hermitian operators and told to István "*I am sure that you will be able to get the same results also by using a hermitian Fock matrix.*" Hence the title of the paper: "Fulfilling Lowdin's prediction" [25].

This breakthrough allowed a more efficient determination of the total energy since, contrary to the earlier formulations, the Fock matrix to be diagonalized in every step of the self-consistent calculation was now hermitian. But, more important, is the fact that since the CHA-SCF canonic orbitals are now eigenvectors of a hermitian Fock matrix (and hence orthogonal), the electron correlation could be potentially introduced in a much simpler way than the quite involved original CHA-MP2 theory.

I will now introduce what most probably was the last effort of István devoted to the BSSE problem. We had worked on this alternative CHA-MP2 formulation based on canonic (and orthogonal) CHA orbitals back in 2005 but never published it, for the reason may be apparent by the end of this manuscript. In the next section the new method is described in more detail. We will briefly describe the hermitized CHA-SCF equations and the novel CHA-S-MP2 perturbative extension. Details of the previous non-hermitian CHA-SCF and CHA-MP2 methods are given elsewhere [9, 11]. Finally, numerical full geometry optimizations of prototypical molecular complexes like water and hydrogen fluoride dimers are presented, where the new scheme is compared to both the CP-MP2 and CHA-MP2 results for several basis sets. The performance of the method is further explored by calculating the potential energy surface for the helium dimer.

2 | THEORY

2.1 | The CHA-SCF scheme

The CHA affords a partitioning of the Fock operator in intraatomic and interatomic contributions by a "mixed" second quantization formalism [1] using the (non-orthogonal) atomic orbital basis. Let us define h^A as the monoelectronic term of the Hamiltonian associated to the monomer A of a molecular complex,

$$\hat{h}^A = -\frac{1}{2}\sum_{i=1}^{Na}\nabla_i^2 - \sum_{i=1}^{Na}\sum_{\alpha \in A}\frac{Z_\alpha}{r_{i\alpha}},\tag{1}$$

the total Hamiltonian being

$$\hat{H}^A = -\frac{1}{2}\sum_i^{N_A}\nabla_i^2 - \sum_i^{N_A}\sum_\alpha^{M_A}\frac{Z_\alpha}{r_{i\alpha}} + \sum_{i \neq i}^{N_A}\frac{1}{r_{ij}} + \sum_{\alpha \neq \beta}^{M_A}\frac{Z_\alpha Z_\beta}{R_\alpha},\tag{2}$$

where N_A, M_A, and Z_α correspond to the number of electrons, number of atoms and their corresponding charge, of the monomer A.

The action of operator (1), of intramolecular nature with respect to the molecular complex, on a basis function assigned to A can be formally expressed as

$$\hat{h}^A|\varphi^A> \equiv \hat{P}^A\hat{h}^A|\varphi^A> + \left(1 - \hat{P}^A\right)\hat{h}^A|\varphi^A>,\tag{3}$$

where

$$\hat{P}^A = \sum_{i,j \in A}|\chi_i^A> S_{(A)ij}^{-1} <\chi_j^A|,\tag{4}$$

is the projector on the subspace spawned by the atomic orbitals of A. As the basis set increases, the second term hopefully diminishes, being zero in the hypothetical case of the complete basis set. It is expected that the contribution of the first term is larger than that of the second. But, most importantly, in the case of the isolated monomer calculation or for an exact solution, one has trivially

$$\hat{h}^A|\varphi^A> = P^A\hat{h}^A|\varphi^A>.\tag{5}$$

Hence, the second term on the rhs of Equation (3) is responsible for the BSSE. In order to obtain a molecular complex description consistent with that of the composing monomer this term must be dropped from the Hamiltonian of the molecular aggregate. That means one must perform the following substitutions

$$\hat{h}^A|\varphi^A> \Rightarrow \hat{P}^A\hat{h}^A|\varphi^A>,$$

$$\frac{1}{r_{12}}|\varphi_i^A(1)\varphi_j^A(2)> \Rightarrow \hat{P}^A(1)\hat{P}^A(2)\frac{1}{r_{12}}|\varphi_i^A(1)\varphi_j^A(2)>,\tag{6}$$

for the mono- and bi-electronic term, respectively. It can be seen that, by performing these substitutions, the intramolecular integrals that correspond to the representation of the operator in the given basis set are transformed in the following way.

$$\langle\chi_i|\hat{h}^A|\chi_j\rangle \Rightarrow \sum_{\mu,\nu \in A} S_{i\mu}S_{(A)\mu\nu}^{-1}\langle\chi_\nu|\hat{h}^A|\chi_j\rangle = \left\{\chi_i|\hat{h}^A|\chi_j\right\}\forall j \in A, \quad \forall i$$

$$[\chi_i\chi_j|\chi_k\chi_l] \Rightarrow \sum_{\mu,\nu,\sigma,\rho \in A} S_{i\mu}S_{j\nu}S_{(A)\mu\sigma}^{-1}S_{(A)\nu\rho}^{-1}[\chi_\sigma\chi_\rho|\chi_k\chi_l] = \left\{\chi_i\chi_j|\chi_k\chi_l\right\}\forall k, \quad 1 \in A, \forall i,j.\tag{7}$$

Note that not all integrals are CHA-transformed; only when the atomic orbital of the right-hand side of the integral (ket) belongs to fragment A. The projectors act on the right-hand side of the integrals, so that the CHA-transformed integrals do not satisfy the usual symmetry, leading to a non-Hermitian SFC problem.

2.2 | The hermitized CHA-SCF equations

In the original derivation [8, 9] of the CHA-SCF equations within the LCAO framework, the following condition was used to obtain the solutions

$$c_q^+ F^{CHA} c_i = 0; \quad q \in \text{virt.}, \quad i \in \text{occ.} \tag{8}$$

where c_q and c_i are the vectors of the corresponding ith occupied and qth virtual orbitals and F^{CHA} is the non-hermitian CHA Fockian, whose elements can be expressed as (for a closed-shell case)

$$F_{\mu\nu}^{CHA} = \left\{ \mu | \hat{h} | \nu \right\} + \sum_{\kappa, \tau} (2\{\mu\kappa | \nu\tau\} - \{\mu\kappa | \tau\nu\}) P_{\tau\kappa}. \tag{9}$$

The quantities $\left\{ \mu | \hat{h} | \nu \right\}$ and $\{\mu\kappa | \nu\tau\}$ correspond to the CHA-transformed one- and two-electron integrals over the AO's of Equation (7). Due to the particular form of the CHA transformation, the one and two-electron integrals used to build the CHA Fockian do not have the usual symmetry properties and hence, F^{CHA} is non-hermitian.

It can be readily seen [8, 9] that Equation (1) is equivalent to solve the corresponding eigenvalue problem in an iterative fashion,

$$F^{CHA} c_i = S c_i \varepsilon_i, \tag{10}$$

were the non-orthogonal right eigenvectors must be orthogonalized in order to obtain the occupied orbitals to build the CHA Fockian for the next iteration, and the virtual orbitals can be obtained from the set of left eigenvectors, $\tilde{c}_q$, by solving

$$\tilde{c}_q^+ F_{CHA} = \varepsilon_q \tilde{c}_q^+ S. \tag{11}$$

However, as pointed out in Reference [25], Equation (1) refers only to the virtual- occupied block of the F^{CHA} matrix expressed in terms of the MO basis. Thus, one can construct the following effective hermitian Fock matrix, F_{eff}

$$F_{eff} = \begin{pmatrix} F_{occ.occ.} & F_{virt.occ.}^{CHA+} \\ F_{virt.occ.}^{CHA} & F_{virt.virt} \end{pmatrix}, \tag{12}$$

from the corresponding blocks of the conventional, F, and F^{CHA} matrices expressed in the MO basis. In the next step, this hermitian matrix is back-transformed to the AO basis and diagonalized in order to get the orbital coefficients to be used in the next iteration. Upon convergence, the condition in Equation (8) is fulfilled and the total energy can be calculated in the usual manner (expectation value of the conventional Hamiltonian) using the molecular orbitals thus obtained. It is worth to note again that the same total energy obtained with the previous non-hermitian CHA-SCF method is recovered with full machine accuracy [25].

2.3 | The hermitized CHA-MP2 (CHA-S-MP2) formalism

In the CHA theory, the energy must always be calculated using the conventional Hamiltonian, whereas the CHA Hamiltonian is used to provide the BSSE-free wavefunction. For that reason, in the perturbative formalisms within the CHA framework, two sets of unperturbed and perturbation operators are introduced, that is, those of the conventional Hamiltonian

$$\hat{H} = \hat{H}^0 + \hat{V}, \tag{13}$$

and those of the CHA Hamiltonian

$$\hat{H}^{CHA} = \hat{H}^0 + \hat{W}. \tag{14}$$

In both cases, the unperturbed Hamiltonian is chosen to be the CHA Fockian, being the unperturbed eigenfunctions the single-determinant wavefunctions built up of the respective eigenvectors, and the unperturbed eigenvalues the sums of eigenvalues.

In the CHA-MP2 [11] formalism, the unperturbed Hamiltonian is the non-hermitian CHA Fockian, so neither the unperturbed nor the perturbation operators involved are hermitian. Consequently, the second-order energy correction is obtained by means of a generalized Hylleras functional [24] that takes into account the non-hermiticity of the operators involved, yet ensuring the second order energy term to be real. Furthermore, the unperturbed wavefunctions (Slater determinants built up of the canonic non-orthogonal CHA-SCF orbitals) are non-orthogonal in general, and hence the different doubly excited determinants overlap with the single-excited and ground state one.

As shown in Reference [11] (we wish not enter into detail into the previous CHA-MP2 formalism), the final expression for the second-order energy correction contains contributions from single-excited and ground-state determinants, apart from the double-excited ones. Besides, since an auxiliary orthonormal basis is introduced, up to three four-index transformations are needed to actually compute the energy.

We will show now how the problem is greatly simplified if we make use of the effective Fock matrix described in Equation (12). In this case, the unperturbed Hamiltonian is constructed from the eigenvalues and eigenvectors of the effective CHA Fock matrix. Since it is hermitian, the unperturbed eigenfunctions (Slater determinants built up of the orthogonal CHA orbitals) are now orthogonal.

Our aim is to get perturbational approximations to the CHA/CE full-CI problem, in which the expectation value of the conventional Hamiltonian is calculated over the eigenfunction of the CHA Hamiltonian. Hence, the strategy is to calculate the first-order wave function with the CHA perturbation operator and then expand the expectation value of the conventional energy for the sum of the zero- and first-order wave functions by keeping terms up to second order. It can be easily seen that in this way the energy can be expressed as the sum of the zero-th and first-order terms

$$E^0 = \frac{\left\langle \Psi^0 | \hat{H}^0 | \Psi^0 \right\rangle}{\left\langle \Psi^0 | \Psi^0 \right\rangle} \quad E^1 = \frac{\left\langle \Psi^0 | V | \Psi^0 \right\rangle}{\left\langle \Psi^0 | \Psi^0 \right\rangle}, \tag{15}$$

and a second-order energy contribution, given by the Hylleraas functional [26]

$$J^2 = \left\langle \psi^1 | \hat{H}^0 - E^0 | \psi^1 \right\rangle + 2\mathrm{Re}\left\langle \psi^1 | \hat{V} - E^1 | \psi^0 \right\rangle. \tag{16}$$

Note that the sum of the zero and first order energies is nothing but the CHA/CE SCF energy so the total CHA-S-MP2 energy will be

$$E^{CHA\text{-}S\text{-}MP2} = E^0 + E^1 + J^2 = E^{CHA\text{-}SCF} + J^2. \tag{17}$$

At this point, we just need a first-order correction to the wavefunction. A simple way to obtain it is to use the expression given by the Rayleigh-Schrödinger perturbation theory,

$$\left| \psi^1 \right\rangle = \sum_k \frac{\left\langle \psi^0 | \hat{W} | \psi_k^0 \right\rangle}{E^0 - E_k^0} \left| \psi_k^0 \right\rangle, \tag{18}$$

with the following perturbation operator

$$\hat{W} = \hat{H}^{CHA} - \hat{H}^0 = \sum_{i<j} \left\{ r_{ij}^{-1} \right\} - \sum_i v^{eff}(i). \tag{19}$$

In the equation above, expression $\sum_{i<j}\left\{r_{ij}^{-1}\right\}$ stands for the CHA-transformed two-electron operator, whereas $\sum_i v^{eff}(i)$ is the Hartree-Fock coulomb and exchange potential of the effective CHA Fockian.

The next step is simply to substitute Equation (18) into the Hylleraas functional of Equation (16)

$$
\begin{aligned}
J^2 &= \sum_l \sum_k \left(\frac{\langle \Psi^0 | \widehat{W} | \Psi_k^0 \rangle}{E^0 - E_k^0} \right)^* \langle \Psi_k^0 | \widehat{H}^0 - E^0 | \Psi_l^0 \rangle \left(\frac{\langle \Psi^0 | \widehat{W} | \Psi_l^0 \rangle}{E^0 - E_l^0} \right) \\
&\quad + 2\mathrm{Re} \sum_k \frac{\langle \Psi^0 | \widehat{W} | \Psi_k^0 \rangle}{E^0 - E_k^0} \langle \Psi_k^0 | \widehat{V} - E^1 | \Psi^0 \rangle \\
&= \sum_l \sum_k \left(\frac{\langle \Psi^0 | \widehat{W} | \Psi_k^0 \rangle}{E^0 - E_k^0} \right)^* (E_l^0 - E^0) \langle \Psi_k^0 | \Psi_l^0 \rangle \left(\frac{\langle \Psi^0 | \widehat{W} | \Psi_l^0 \rangle}{E^0 - E_l^0} \right) \\
&\quad + 2\mathrm{Re} \sum_k \frac{\langle \Psi^0 | \widehat{W} | \Psi_k^0 \rangle}{E^0 - E_k^0} \left(\langle \Psi_k^0 | \widehat{V} | \Psi^0 \rangle - E^1 \langle \Psi_k^0 | \Psi^0 \rangle \right),
\end{aligned}
\tag{20}
$$

where the perturbation operator is

$$
\widehat{V} = \widehat{H} - \widehat{H}^0 = \sum_{i<j} r_{ij}^{-1} - \sum_i v^{eff}(i)..
\tag{21}
$$

Taking into account that the CHA-SCF equations fulfill the Brillouin theorem, and hence the first-order correction to the wavefunction contains only doubly excited determinants, we obtain after some manipulations

$$
J^{(2)} = -\sum_k \frac{\left| \langle \psi^0 | \widehat{W} | \psi_k^0 \rangle \right|^2}{E^0 - E_k^0} + 2\sum_k \frac{\langle \psi^0 | \widehat{W} | \psi_k^0 \rangle \langle \psi_k^0 | \widehat{V} | \psi^0 \rangle}{E^0 - E_k^0}.
\tag{22}
$$

The integrals involved in the previous equations can be expressed, similarly to the Møller–Plesset theory, in terms of the CHA transformed

$$
\left\langle \psi^0 | \widehat{W} | \psi_{ab}^{rs} \right\rangle = \left\langle \psi^0 | \sum_{i<j} \left\{ r_{ij}^{-1} \right\} | \psi_{ab}^{rs} \right\rangle - \left\langle \psi^0 | \sum_i v^{eff} | \psi_{ab}^{rs} \right\rangle = \left\{ ab \| rs \right\},
\tag{23}
$$

and conventional

$$
\left\langle \psi^0 | \widehat{V} | \psi_{ab}^{rs} \right\rangle = \left\langle \psi^0 | \sum_{i<j} r_{ij}^{-1} | \psi_{ab}^{rs} \right\rangle - \left\langle \psi^0 | \sum_i v^{eff} | \psi_{ab}^{rs} \right\rangle = \left[ab \| rs \right],
\tag{24}
$$

two-electron integrals over the molecular spinorbitals.

Substituting Equations (23) and (24) into Equation (22) we obtain the final expression for the second-order correction to the energy

$$
J^2 = \sum_{\substack{a < b \in occ \\ r < s \in virt}} \frac{2\{ab \| rs\}[rs \| ab] - \left| \{ab \| rs\} \right|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s},
\tag{25}
$$

and, in the case of a closed-shell system

$$
\begin{aligned}
J^2 &= 2\sum_{\substack{a,b \\ r,s}} \frac{[ab|rs](2\{ab|rs\} - \{ab|sr\})}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} - \sum_{\substack{a,b \\ r,s}} \frac{[ab|rs](2\{ab|rs\} - \{ab|sr\})}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} \\
&\quad \sum_{\substack{a,b \\ r,s}} \frac{(2[ab|rs] - \{ab|rs\})(2\{ab|rs\} - \{ab|sr\})}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s},
\end{aligned}
\tag{26}
$$

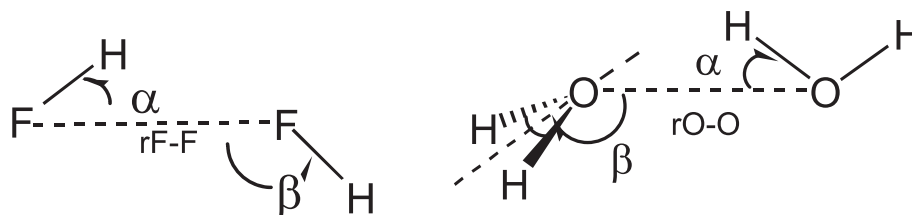where the summation runs over spatial molecular orbitals.

FIGURE 1 Geometrical parameters of the water and hydrogen fluoride dimers

TABLE 1 Interaction energies (in kcal/mol) and geometrical parameters (Å and degrees, see Figure 1) of the $(H_2O)_2$ complex with several basis sets at the MP2, CP-MP2, CHA-MP2, and CHA-S-MP2 levels of theory

Basis set	Method	ΔE_{int}	rO–O	α	β
6-31G**	MP2	−7.12	2.910	9.1	100.2
	CP-MP2	−5.07	2.990	1.4	134.6
	CHA-MP2	−5.34	2.959	4.4	123.6
	CHA-S-MP2	−5.26	2.962	3.1	129.3
6-31++G**	MP2	−6.47	2.907	3.5	133.1
	CP-MP2	−4.87	3.013	2.3	136.3
	CHA-MP2	−4.82	3.029	2.2	135.4
	CHA-S-MP2	−5.01	2.981	1.8	140.3
6-311G**	MP2	−7.00	2.903	4.6	116.0
	CP-MP2	−4.59	3.038	2.4	135.1
	CHA-MP2	−5.02	2.983	0.8	134.5
	CHA-S-MP2	−5.12	2.976	0.5	134.9
6-311++G**	MP2	−6.07	2.916	2.0	136.0
	CP-MP2	−4.56	3.018	3.1	136.4
	CHA-MP2	−4.65	2.999	1.6	138.1
	CHA-S-MP2	−4.96	2.979	1.9	138.0
6-311++G(2df,2p)	MP2	−5.49	2.914	3.6	127.1
	CP-MP2	−4.55	2.960	2.2	133.4
	CHA-MP2	−4.63	2.958	5.4	123.8
	CHA-S-MP2	−4.71	2.945	5.2	124.3
cc-pvdz	MP2	−7.53	2.907	5.4	105.0
	CP-MP2	−4.21	3.047	0.1	138.2
	CHA-MP2	−4.70	2.956	2.2	128.4
	CHA-S-MP2	−4.65	2.956	1.6	130.2
aug-cc-pvdz	MP2	−5.50	2.903	6.3	120.7
	CP-MP2	−4.45	2.977	5.7	122.8
	CHA-MP2	−4.48	2.970	3.9	126.1
	CHA-S-MP2	−3.55	3.017	4.1	130.0
	MP2 limit[a]	−4.98	2.925	4.3	128.2

Abbreviations: CHA, chemical Hamiltonian approach; CP, counterpoise.
[a]Reference [21].

It can be easily seen that Equation (26) reduces to the conventional Møller–Plesset second-order correction when the BSSE perturbation is switched off (i.e., $\{ab\|rs\} \equiv [ab\|rs]$). Furthermore, Equation (26) implies only two four-index transformations so that the computational cost of the CHA-S-MP2 calculation is just twice that of a conventional uncorrected MP2 calculation. It should be stressed that, contrary to the CP method, the computational cost does not depend on the number of interacting fragments.

TABLE 2 Interaction energies (in kcal/Mol) and geometrical parameters (Å and degrees, see Figure 1) of the $(HF)_2$ complex with several basis sets at the MP2, CP-MP2, CHA-MP2, and CHA-S-MP2 levels of theory

Basis set	Method	ΔE_{int}	rF–F	α	β
6-31G**	MP2	−7.80	2.540	46.9	49.5
	CP-MP2	−4.72	2.799	6.9	114.8
	CHA-MP2	−4.85	2.734	7.8	113.1
	CHA-S-MP2	−4.57	2.741	10.1	111.7
6-31++G**	MP2	−5.02	2.772	7.5	115.4
	CP-MP2	−4.25	2.838	7.0	115.9
	CHA-MP2	−4.06	2.847	7.1	117.1
	CHA-S-MP2	−3.98	2.843	7.8	116.2
6-311G**	MP2	−6.16	2.708	17.5	95.2
	CP-MP2	−4.21	2.849	5.9	122.5
	CHA-MP2	−4.54	2.791	8.2	116.9
	CHA-S-MP2	−4.61	2.781	8.0	116.6
6-311++G**	MP2	−4.69	2.788	6.5	121.0
	CP-MP2	−3.86	2.877	8.6	118.7
	CHA-MP2	−3.92	2.865	8.1	118.4
	CHA-S-MP2	−4.06	2.844	7.9	118.2
6-311++G(2df,2p)	MP2	−4.92	2.758	7.3	111.1
	CP-MP2	−3.99	2.817	7.2	114.5
	CHA-MP2	−4.03	2.805	6.9	114.3
	CHA-S-MP2	−4.07	2.787	7.6	111.8
cc-pvdz	MP2	−7.06	2.589	47.2	49.9
	CP-MP2	−4.17	2.834	5.7	119.7
	CHA-MP2	−4.61	2.746	7.6	114.9
	CHA-S-MP2	−4.50	2.746	6.6	116.4
aug-cc-pvdz	MP2	−4.91	2.741	5.5	111.5
	CP-MP2	−4.04	2.810	6.9	111.6
	CHA-MP2	−3.91	2.832	5.4	116.4
	CHA-S-MP2	−3.04	2.885	7.4	113.3
	MP2 limit	−4.37[a]	2.72[b]	6.7[b]	111.7[b]

Abbreviations: CHA, chemical Hamiltonian approach; CP, counterpoise.
[a]Reference [28].
[b]Reference [17].

3 | COMPUTATIONAL DETAILS

The hermitized CHA-SCF and CHA-S-MP2 methods were implemented as independent FORTRAN-77 code interfaced with a slightly modified version of the HONDO8 package [27] from which proper one and two-electron integrals are obtained. We were limited to up to 255 basis functions and up to nine primitives per basis function.

Concerning the PT correction term, two four index transformations must be computed, namely, a conventional and a CHA-modified one. The key point for an efficient implementation of the method is to compute on the fly the contribution of each four-center integral of the AO basis to the CHA-transformed integrals over the MO. That is, the CHA-modified integrals over the AO-s are never computed explicitly. One only must distinguish between intermolecular and intramolecular contributions and, obviously, the transformation should be carried out in four consecutive steps to achieve N^5 scaling.

Since the CHA-S-MP2 is a simplification of the more involved original CHA-MP2 method, the actual implementation of our code can compute either of the two PT methods simply by selecting the proper zeroth-order Hamiltonian (canonic non-orthogonal orbitals eigenvectors of the CHA Fockian or the canonic eigenvectors of the effective hermitian Fockian).
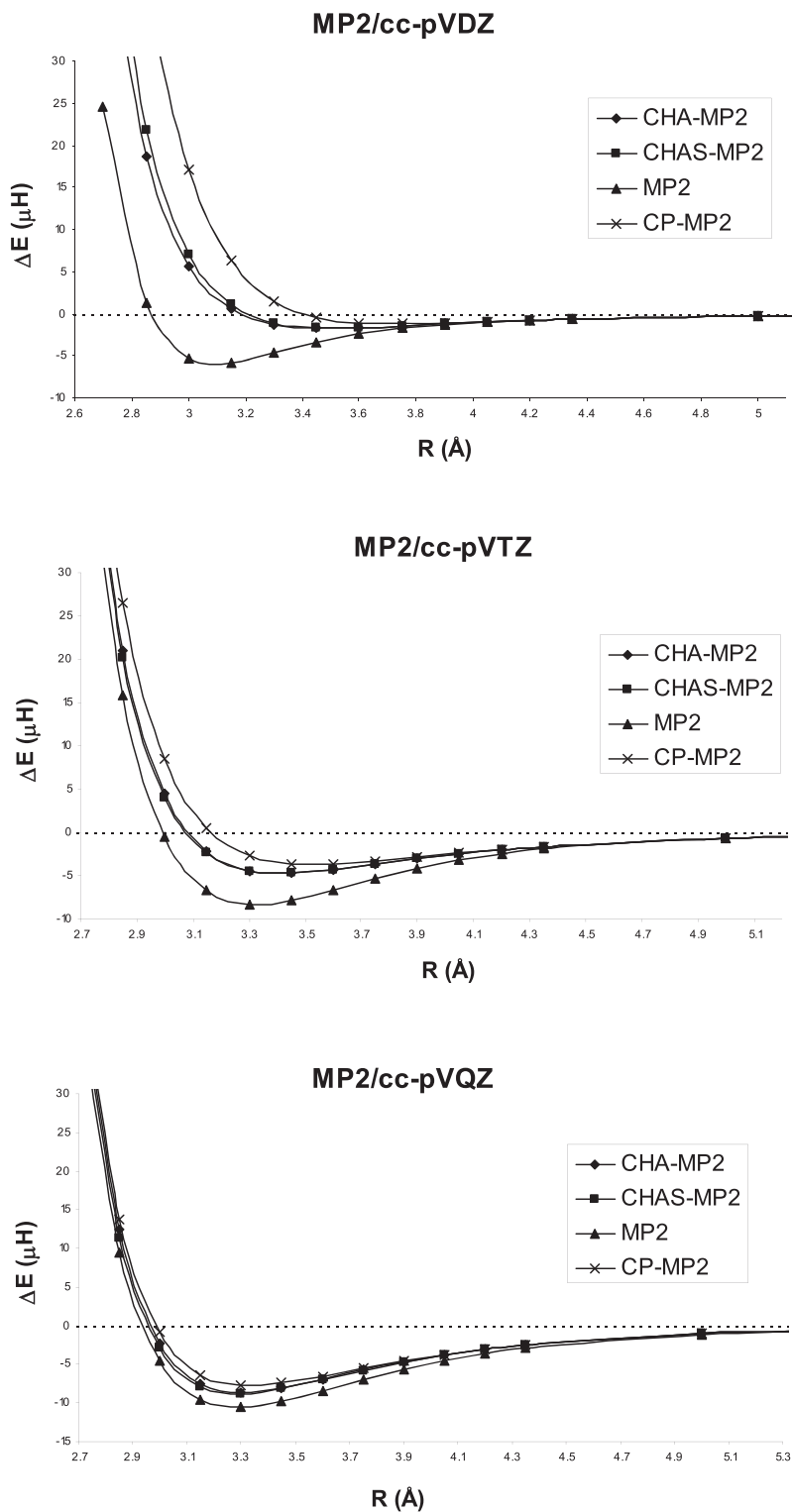
MP2/cc-pVDZ

MP2/cc-pVTZ

MP2/cc-pVQZ

FIGURE 2 Potential energy surfaces of the helium dimer at the MP2/cc-pVXZ, (X = D, T, Q) levels of theory

4 | ILLUSTRATIVE CALCULATIONS

We present full numerical geometry optimizations at the CHA-S-MP2 level (all electrons were correlated) of two prototype hydrogen-bonded complexes, namely water and hydrogen fluoride dimers. The interaction energies and geometrical parameters (see Figure 1) are gathered in Tables 1 and 2, respectively. The geometrical parameters and interaction energies are compared with previous CHA-MP2 (numerically) optimized [23] and both CP-MP2 and MP2 (analytically) optimized [14, 23] structures.
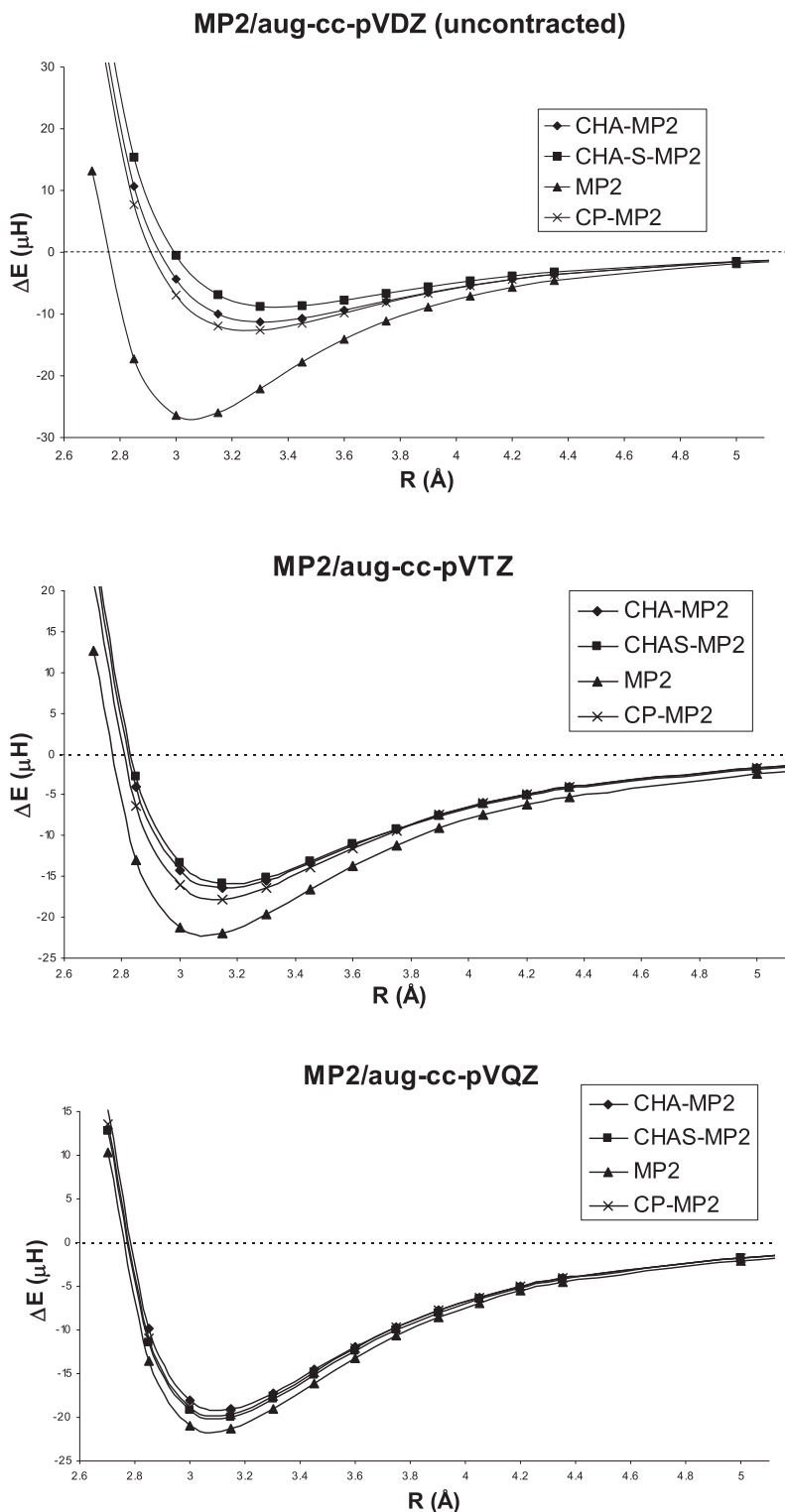
MP2/aug-cc-pVDZ (uncontracted)

MP2/aug-cc-pVTZ

MP2/aug-cc-pVQZ

FIGURE 3   Potential energy surfaces of the helium dimer at the MP2/aug-cc-pVXZ, (X = D, T, Q) levels of theory

The effect of BSSE correction on interaction energies and intermolecular distances is well-known [14] and will not be addressed here in detail. Our aim is to explore the capabilities of the proposed BSSE-correction method at the correlated level. Hence, we present results carried out with rather small basis sets, too.

First of all, we will focus on differences between the three BSSE-correction schemes. Except for the 6-31++G** basis set, the CHA-MP2 interaction energies lie between the uncorrected and CP-corrected one. The largest difference between both BSSE-corrected energies is less than
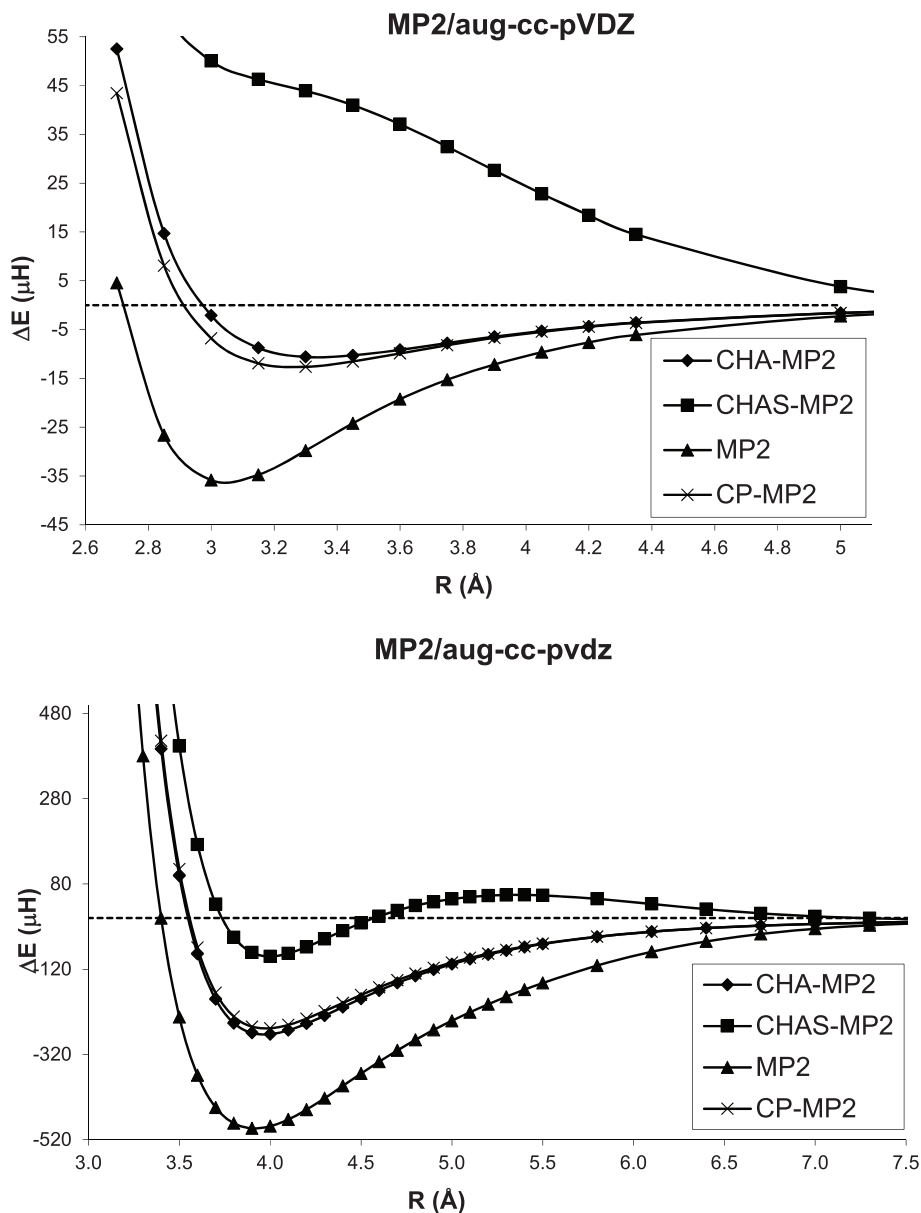
FIGURE 4 MP2/aug-cc-pVDZ potential energy surfaces of the helium dimer (top) and argon dimer (bottom)

0.5 kcal/mol (water dimer, MP2/6-311G**). In this particular case, the difference with respect to the uncorrected energy is about 2 kcal/mol, that is 40% of the overall value.

Concerning the CHA-S-MP2 results, they are very close to the CHA-MP2 ones. The largest discrepancy is about 0.3 kcal/mol, except for the aug-cc-pVDZ basis set (this case is discussed in more detail below). Also, the effect on the intermolecular R_{O-O} and R_{F-F} distances matches that of the interaction energy.

Interestingly, the inclusion of diffuse functions in the basis sets drives to a larger difference between both CHA methods, whereas with the basis sets that do not include such special functions the results almost coincide. In our opinion, this effect is due to the fact that these diffuse functions are too spread over the molecule and it is rather questionable to assign them only to a given monomer. It is also known that the conventional CHA method exhibits difficulties when the fragments strongly overlap, as is the case of intramolecular interactions [6] or strong ionic molecular complexes [29].

A much weaker intermolecular interaction such as He_2 has also been considered. The potential energy surfaces of the He⋯He interaction in the range of 2.5–5 Å for the cc-pVXZ and aug-cc-pVXZ, (X = D, T, Q) basis set families are depicted in Figures 2 and 3, respectively. In a previous study, Lee et al. [30] have estimated a MP2 limit value for the interaction energy of this system of −21.4 μH, still far from the −31.4 μH experimental value [31]. Our results are in excellent agreement with those of Lee et al. The curves corresponding to the cc-pVXZ family show that the

BSSE is responsible for most of the binding energy, especially for the cc-pVDZ basis set. In this case, BSSE correction lead to unbound curves in all cases. The differences between both CHA approaches is negligible, whereas the CP-MP2 curve lies always above the rest, showing a larger BSSE correction.

The inclusion of diffuse function (see Figure 3) is essential for the correct description of the interaction. In this case, the CHA-MP2 curves lie above the CP-MP2 ones, but the difference is smaller than the BSSE itself. By increasing the basis set to triple zeta this effect is diminished whereas the aug-cc-pVQZ curves are practically coincident. Note, however, that the results corresponding to the aug-cc-pVDZ basis set have been obtained after uncontracting the 1s orbital of this basis.

As mentioned above, the CHA-S-MP2/aug-cc-pVDZ results strongly deviate from the respective CHA-MP2 and CP-MP2 values for the two hydrogen-bonded systems. In fact, the difference between the CHA-S-MP2 and CHA-MP2 interaction energies are of the same order of the BSSE itself, $\sim$1 kcal/mol. Accordingly, the O–O and F–F distances are also overestimated. It is to be mentioned that others have also found problems associated with this particular basis set [28]. We repeated the calculations by uncontracting the basis set and the problem disappeared.

In the case of He_2 and Ar_2 the results are particularly odd, as illustrated in Figure 4. The CHA-S-MP2 interaction energies converge to the dissociation limit from above, as an unphysical shoulder (much significant in the case of He_2) is observed. The limitation of this basis set is confirmed by the fact that the uncorrected SCF curve exhibits a rather deep minimum. The CHA-SCF method properly corrects for this situation, however, the further inclusion of electron correlation though CHA-S-MP2 cannot lead to a proper description of the weak interaction. We have identified the diffuse s-type orbital of the basis as responsible for this odd behavior. If removed, the shoulder in the curve disappears. However, that essentially leads back to the cc-pVDZ basis. On the other hand, if more flexibility is given to the basis set just by uncontracting the 1s orbital, the problem is also solved, as shown in Figure 3 in the case of He_2. These kind of problems are reminiscent of the known limitation of conventional CHA to deal with strong overlap (e.g., the intramolecular BSSE problem) [7].

A careful analysis of the situation seems to indicate that the CHA-S-MP2 scheme, though much more efficient computationally than the more involved CHA-MP2, is not as robust as the latter when dealing with strong overlap. The origin of the difference is on the definition of the zero-th order Hamiltonian. In fact, in the case of the CHA-MP2 method, some of the effects due to the BSSE are already included in the zero-th order Hamiltonian, since it is chosen to be non-hermitian. On the contrary, when using the effective Fockian as zero-th order Hamiltonian in the CHA-S-MP2, the non-hermiticity is necessarily contained into the perturbation operator. In other words, the eigenvectors of the effective Fockian are such that the virtual-occupied block of the CHA Fockian vanishes, but this is not true in general for the occupied-virtual one, for instance. This non-hermiticity might be at the heart of the failure of the CHA-S-MP2 method, albeit only observed for the problematic aug-cc-pVDZ basis set. Yet, similar problems might also appear for strongly bound complexes described with diffuse functions, where overlap effects are larger.

A good compromise might be to include the non-hermiticity into the zeroth order Hamiltonian, yet keeping the benefits of the occupied and virtual orthogonal molecular orbitals. This will lead to a non-diagonal zero-th order Hamiltonian, similarly to other MP2 formulations with non-canonic orbitals [32,33]. It would be important to determine whether such a formulation could be implemented efficiently enough to compete with the non-hermitian CHA-MP2 method.

5 | CONCLUSIONS

We describe an alternative perturbational approach free of BSSE within the framework of the CHA. The new formulation is based on the canonic and orthogonal CHA orbitals obtained from a hermitized CHA Fock operator. Significant simplification of the method as compared to the previous CHA-MP2 formalism is achieved in the present formulation, so that the computational burden of the CHA-S-MP2 calculation is just twice that of a conventional uncorrected MP2 calculation. Also, the computational cost is independent of the number of interacting fragments. The results obtained are in excellent agreement with both the CHA-MP2 and CP-MP2 methods for both hydrogen bonded and weak molecular interactions with one exception. The difficulties arisen with the aug-cc-pVDZ basis set indicate that the present method is not as robust as the original non-hermitian CHA-MP2 formulation.

AUTHOR CONTRIBUTIONS

Pedro Salvador: Data curation; formal analysis; software; writing - original draft. **István Mayer:** Conceptualization; supervision.

ORCID

Pedro Salvador https://orcid.org/0000-0003-1823-7295

REFERENCES

 [1] I. Mayer, *Int. J. Quantum Chem.* **1983**, *23*, 341.
 [2] I. Mayer, *Chem. Phys. Lett.* **2003**, *382*, 265.
 [3] P. Salvador, M. Duran, I. Mayer, *J. Chem. Phys.* **2001**, *115*, 1153.
 [4] H. B. Jansen, P. Ross, *Chem. Phys. Lett.* **1969**, *3*, 140.
 [5] I. Mayer, *Int. J. Quantum Chem.* **1998**, *70*, 41.
 [6] F. Jensen, *J. Chem. Theory Comput* **2010**, *6*(1), 100.
 [7] G. J. Halasz, A. Vibok, S. Suhai, I. Mayer, *Int. J. Quantum Chem.* **2002**, *89*, 190.
 [8] I. Mayer, Á. Vibók, *Chem. Phys. Lett.* **1987**, *140*, 558.
 [9] I. Mayer, Á. Vibók, *Chem. Phys. Lett.* **1988**, *148*, 68.
[10] I. Mayer, A. Vibok, P. Valiron, *Chem. Phys. Lett.* **1994**, *224*, 166.
[11] I. Mayer, P. Valiron, *J. Chem. Phys.* **1998**, *109*, 3360.
[12] P. Salvador, I. Mayer, *J. Chem. Phys.* **2004**, *120*, 5882.
[13] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553.
[14] P. Salvador, B. Paizs, M. Duran, S. Suhai, *J. Comput. Chem.* **2001**, *22*, 765.
[15] A. Famulari, M. Raimondi, M. Sironi, E. Gianinetti, *Chem. Phys.* **1998**, *232*, 275.
[16] T. Nagata, S. Iwata, *J. Chem. Phys.* **2004**, *120*, 3555.
[17] B. Paizs, P. Salvador, A. G. Császár, M. Duran, S. Suhai, *J. Comput. Chem.* **2001**, *22*, 196.
[18] R. Wieczorek, L. Haskamp, J. J. Dannenberg, *J. Phys. Chem. A* **2004**, *108*, 6713.
[19] J. G. C. M. Van Duijneveldt-van de Rijdt, F. B. van Duijneveldt, *J. Chem. Phys.* **1992**, *97*, 5019.
[20] T. Van Mourik, A. K. Wilson, K. A. Peterson, D. E. Woon, T. H. Dunning Jr., *Adv. Quantum Chem.* **1998**, *31*, 105.
[21] M. Shütz, S. Brdarski, P. O. Wildmark, R. Lindh, G. Karlström, *J. Chem. Phys* **1997**, *107*, 4597.
[22] S. Simon, M. Duran, J. J. Dannenberg, *J. Chem. Phys.* **1996**, *105*, 11024.
[23] P. Salvador, D. Asturiol, I. Mayer, *J. Comput, Chem* **2006**, *27*, 1505.
[24] I. Mayer, *Mol. Phys.* **1996**, *89*, 515.
[25] I. Mayer, *Int. J. Quantum Chem.* **2002**, *90*, 89.
[26] E. A. Hylleraas, *Z. Physik* **1928**, *48*, 469.
[27] HONDO-8, from MOTECC-91, contributed and documented by M. Dupuis and A. FFArazdel, IBM Corporation, Center for Scientific and Engineering Computations, Kingstom, New York, **1991**.
[28] G. Weck, E. Kochanski, J.-F. Gouyet, *Mol. Phys.* **2005**, *103*, 825.
[29] P. Salvador, M. Duran, J. J. Dannenberg, *J. Phys. Chem. A.* **2002**, *106*, 6883.
[30] J. S. Lee, S. Y. Park, *J. Chem. Phys.* **2000**, *112*, 10746.
[31] R. A. Aziz, M. J. Slaman, *J. Chem. Phys.* **1991**, *94*, 8047.
[32] P. Pulay, S. Saebø, *Theor. Chim. Acta* **1986**, *69*, 357.
[33] S. Saebø, P. Pulay, *J. Chem. Phys.* **1987**, *86*, 914.