Universitat de Girona

IMPLEMENTATION AND APPLICATION OF BASIS SET SUPERPOSITION ERROR-CORRECTION SCHEMES TO THE THEORETICAL MODELING OF WEAK INTERMOLECULAR INTERACTIONS

Pedro SALVADOR SEDANO

ISBN: 84-8458-143-8
Dipòsit legal: Gi.254-2002
http://hdl.handle.net/10803/7935

IMPLEMENTATION AND APPLICATION OF BASIS SET SUPERPOSITION ERROR-CORRECTION SCHEMES TO THE THEORETICAL MODELING OF WEAK INTERMOLECULAR INTERACTIONS

A dissertation submitted by Pedro Salvador Sedano in partial fulfillment of the requirements for the degree of Doctor in Philosophy

Department of Chemistry
and
Institute of Computational Chemistry
University of Girona

Girona, December 2001

El sotasignat **Prof. Miquel Duran i Portas**, Catedràtic d' Universitat del Departament de Química de la Universitat de Girona

CERTIFICO:

Que en **Pedro Salvador Sedano**, llicenciat en Ciències, secció de Químiques, ha realitzat sota la meva direcció, a l'Institut de Química Computacional i el Departament de Química d'aquesta Universitat, el treball titulat " *On the implementation and application of Basis Set Superposition Error-correction schemes for the theoretical modeling of weak intermolecular interactions* " que es troba recollit en aquesta memòria i que es presenta en pública defensa per optar al grau de Doctor en Ciències Químiques.

I perquè consti als efectes legals escaients, signo aquest certificat a Girona, a 4 de Octubre de 2001.

M. Duran

CONTENTS

ABBREVIATIONS

ρ: electron density

$\nabla^2\rho$: Laplacian of the electron density

AIM: Atoms in Molecules theory

AO: Atomic Orbitals

BCP: Bond critical point

BSIE: Basis set incompleteness error

BSSE: Basis set superposition error

CECA: Chemical Energy Component Analysis

CHA/DFT: Chemical Hamiltonian Approach applied to the Kohn-Sham Fockian

CHA/F: Chemical Hamiltonian Approach applied to the Fock matrix

CHA: Chemical Hamiltonian Approach

CHA-MP2: Extension of the CHA-SCF method to second-order perturbation theory.

CHA-SCF: Chemical Hamiltonian Approach applied to the Hamiltonian

CP: Counterpoise method

CPHF: Coupled-perturbed Hatree-Fock equations

DCBS: Dimer-centered basis set.

DFT: Density Functional Theory

HF: Hartree-Fock

MCBS: Monomer-centered basis set.

MO: Molecular Orbitals

MP2: Second-order Møller-Plesset perturbation theory

PAFC: Pairwise additive function counterpoise

QMS: Quantum Molecular Similarity

RCP: Ring critical point

SAPT: Symmetry-adapted perturbation theory

SCF: Self- Consistent Field

SQ: Second Quantization

SSFC: Site-Site function counterpoise

VMFC: Valiron-Mayer (or hierarchical) function counterpoise

ZPVE: Zero-point vibrational energy correction

INDEX OF INTERMOLECULAR COMPLEXES

Name	Formula	BSSE	Opt	Freq	SAPT	ρ	Section
Acethylene–Ozone	$C_2H_2\cdots O_3$	CP	MP2	Y	N	N	II.1.1.3.b
Boron trifluoride–Ammonia	$BF_3\cdots NH_3$	CP	MP2	N	N	N	II.1.2.1.a
Ethylene–Sulfur dioxide	$C_2H_4\cdots SO_2$	CP	MP2	N	N	N	II.1.2.1.b
Formic acid dimer	$(HCOOH)_2$	CHA	HF	N	N	Y	II.1.3.1c, II.1.3.2.c, II.1.3.3, II.1.3.4
Hydrogen cyanide–Ozone	$HCN\cdots O_3$	CP	MP2	Y	N	N	II.1.1.3.a
Hydrogen fluoride dimer	$(HF)_2$	CP CHA SAPT	MP2 DFT	N	N	Y	II.1.1.1.a, II.1.1.1.d, II.1.1.2.a, II.1.3.1.a, II.1.3.2.a, II.1.3.3.
Hydrogen fluoride tetramer (chain)	$(HF)_4$	CP SAPT	MP2	Y	Y	N	II.1.4.1.b
Hydrogen fluoride tetramer (cyclic)	$(HF)_4$	CP SAPT	MP2	Y	Y	N	II.1.4.1.a
Hydrogen fluoride trimer (chain)	$(HF)_3$	CP SAPT	MP2	Y	Y	N	II.1.4.1.b
Hydrogen fluoride trimer (cyclic)	$(HF)_3$	CP SAPT	MP2	Y	Y	N	II.1.4.1.a
Water–Ammonium	$H_2O\cdots NH_4^+$	CP	MP2	Y	N	N	II.1.1.4.c
Water–Hydroxide	$H_2O\cdots OH^-$	CP	MP2	Y	N	N	II.1.1.4.a
Water–Uracil	$H_2O\cdots C_4H_4N_2O_2$	CHA	HF	N	N	Y	II.1.3.1.c, II.1.3.2.c, II.1.3.3, II.1.3.4
Water dimer	$(H_2O)_2$	CP CHA SAPT	MP2 DFT	N	N	Y	II.1.1.1.b, II.1.1.1.d, II.1.1.2.b, II.1.3.1.b, II.1.3.2.b, II.1.3.3, II.1.3.4
Water-Hydrogen flouride	$H_2O\cdots HF$	CP CHA	MP2 DFT	N	N	N	II.1.1.1.c, II.1.1.1.d II.1.1.2.c
Water-Hydronium	$H_2O\cdots H_3O^+$	CP	MP2	Y	N	N	II.1.1.4.b

I Introduction

I.1 The Basis Set Superposition Error (BSSE)

I.1.1 Historical overview

The application of quantum mechanics techniques based on the time-independent Schrödinger equation to the study of chemical systems lead to different differential equations that can only be solved exactly for monoelectronic systems. In the most general case of the study of polyatomic molecules, even upon the use of the so-called Born-Oppenheimer[1] approximation to decouple the motion of the nuclei and the electrons of the system, the electronic repulsion term forces the introduction of further approximations that, roughly, consist in treating this interaction in a averaged fashion.

The well-known Hartree-Fock equations are obtained when the molecular orbital (MO) approximation is introduced. This approximation supposes that the electrons of the system are described by single-particle wavefunctions, spin-orbitals, extended over the whole system. The total wavefunction of the system is later constructed from these MO's. In this way, a polyelectronic problem is transformed in a set of monoelectronic (coupled, non-linear) equations where the electron-electron repulsion can only be included in the monoelectronic (Fock) operator through an averaged potential due to the field created by the other electrons of the system.

The introduction of monoelectronic functions where those orbitals can be expanded in, carried out by Roothaan[2], represent a key point from a computational point of view. The complicated Hartree-Fock equations can be transformed now in an approximate matricial pseudo-secular equation in order to obtain the coefficient of the expansion of the molecular orbitals in the basis functions.

This technique was fist introduced independently by Hund[3] and Mulliken[4], as an

---

[1] A. Szabo, N.S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory. MacMillan Publishing Co. Inc.* (1982).

[2] C. C. J. Roothaan, Rev. Mod. Phys. 23, 69 (1951).

[3] F. Hund, Z. Phys. 36, 657 (1926)

[4] R. S. Mulliken Phys. Rev. 32, 186 (1927)

alternative to the valence bond theory[5]. It was later applied also to homonuclear diatomic molecules in by Lennard-Jones[6] and was finally generalized in a series of excellent papers by Mulliken[7], who was awarded the Nobel Prize in Chemistry in 1966 for his "fundamental work concerning chemical bonds and the electronic structure of molecules."

During the last fifty years an unnumerable amount of new ab intio methodological improvements have appeared in the literature with two goal that are often antagonistic: the accurate description of the nature of the chemical systems and their properties, and the reduction in the computational cost of the calculations.

The Hartree-Fock model has often represented the starting point for other sophisticated methods in order to include the so-called electron correlation. One can roughly split those methods into two families: those based on a single Slater determinant, and the multiconfigurational ones, where a linear combination of Slater determinants are used as wavefunction in the Hartree-Fock procedure. The second type of methods include the non-dynamic correlation, i. e., that due to the presence of nearly degenerate configuration state functions to be combined with the single-determinant Hartree-Fock ground state wavefunction. In all cases, however, the wavefunction is finally expressed as linear combination of Slater determinants; the coefficients and the structure of the expansion differ for one method to another.

Another approach which has been extensively developed and used during the last decade with great success has been the Density Functional Theory (DFT). In this method, the Hohenberg-Kohn theorem[8] is used to express the energy of the chemical system in terms of the electron density so that the concept of wavefunction is strictly not necessary. However, the later Kohn-Sham[9] formulation results in an algorithm conceptually very similar to that of the Hartree-Fock method, with the introduction of molecular orbitals and basis functions. The main advantage of such a methodology is that the electron correlation is introduced with very low computational effort and large systems like oligopeptides or organometallic compounds can be nowadays

[5] W. Heitler and F. London, Z. Phys. 44, 455 (1927)

[6] J. E. Lennard-Jones, Trans. Faraday Soc. 25, 668 (1929)

[7] R. S. Mulliken, Phys. Rev. 43, 279 (1933), R. S. Mulliken, Phys. Rev. 41, 49 (1932), R.S. Mulliken, J. Chim. Phys. 46, 497 (1949)

[8] P. Hohenberg, W. Kohn, Phys. Rev. B 136, 864 (1964).

[9] W. Kohn, L. J. Sham, Phys. Rev. A 140, 1133 (1965)

routinely studied with pseudo-*ab initio* techniques.

The ambiguous selection of the exchange-correlation functional and the lack of a systematic methodology to improve these functionals are the main drawbacks of DFT.

In short, there is a wide range of ab initio methodologies for the understanding of the properties and nature of the chemical systems. State-of-the–art methods can yield quantitative predictions within the so-called chemical precision, that is, of the order of ca. 1kcal/mol. The fast development of the computational capabilities is also helping us to set closer to the challenge of connecting the understanding of the physico-chemical processes present at the molecular scale with the properties of the bulk matter.

Even though it is not conceptually necessary, most of the present available *ab initio* methods are based on the use of atomic orbitals centered on the nuclei where the wavefunction or in general the electron density, is expanded. For the Hartree-Fock method, the difficulties found in order to obtain analytically the exact solution implies that the exact spin-orbitals to be used to build the Slater determinant must be approximated by linear combination of monoelectronic functions (orbitals). This is the so-called MO-LCAO approximation, i.e., molecular orbitals expressed as a linear combination of atomic orbitals.

These set of functions centered on the nuclei simulate the shape of the exact solution for the hydrogenoid atoms, and hence should represent a good approximation from a both physical and chemical point of view. The atomic orbitals generally correspond to Slater-type, Gaussian or linear combination of Gaussian functions. The former are more suitable to describe the electron density at the atomic positions but are much more expensive from a computational point of view. The expressions of these atomic functions are obtained by optimizing a given molecular property with respect to some parameters like the exponential factor or coefficients of the linear combination. In this way, depending upon the property one is interested in and the methodology used, one can use an specific set of atomic orbitals specifically optimized for this purpose.

When dealing with molecules the set of functions formed by all the atomic orbitals of the molecule is called basis set, which obviously must be linearly independent.

There are other options, of course. One can use the so-called floating functions,

that is, a set of basis functions that are free to move in space. Their main advantage is that the polarization of the atoms due to the presence of external perturbations (other atoms, electric field, etc) can be reproduced without the inclusion of high-order angular momentum nuclei-centered functions. In a similar manner, the molecular basis set can be further increased with some extra functions located in any place of the space, commonly in the middle of a given intermolecular bond (bond functions).

In any case, the use of these atomic set of functions, both floating or fixed, cause a mathematical inconsistency which is essentially ignored in the calculation of polyatomic systems. For instance, when one is comparing different conformations of a molecule, the position of the nuclei vary from one conformer to another an so does for the position of the basis functions. Therefore, strictly speaking, the functional space where the wavefunction of the system is expressed is different in both cases so that, for instance, any comparison of the energy of the conformers can be considered inconsistent from a purely mathematical point of view. The same would occur during geometry optimizations, where the positions of the nuclei also change.

Only in a recent paper, Jensen examined[10] this effect for the inversion barrier of the ammonia and the rotational barrier for the ethane by using an extended basis set including the basis functions of the two structures involved in the process, i.e. the C_{3v} and planar structures of the ammonia and the eclipsed and alternated conformers of the ethane. The errors supposed more than 10% of the barrier for the smallest basis set and converged to zero with much larger basis sets.

Historically, there is a case where the use of truncated basis sets under the LCAO approximation has been recognized to produce large inconsistencies, being in many cases the main source of error. This is clearly the case of the calculation of weak intermolecular interactions.

Let us consider a chemical system AB composed of two interacting fragments A and B. The interaction energy can be expressed simply as the difference of the energy of the complex AB and the energy of its fragments A and B. Even though there are methods to compute directly the interaction energy by means of Intermolecular

[10] F. Jensen, Chem. Phys. Lett. 216, 633 (1996)

Perturbation Theory[11], this straightforward scheme, called supermolecular approximation, has been and nowadays still is the most common procedure for the quantitative determination of interaction energies. The problem is that, conceptually, the description of the fragment A *within* the complex can be improved by the basis functions of the fragment B and viceversa, whereas such an extension is not possible in the calculation of the isolated fragments A and B. This unbalance provokes that the overall description of the AB system is improved with respect to the fragments description. Hence, the interaction energy, expressed as the difference between the energies of the complex and its components is biased by the fact that the basis set where the corresponding energies are expanded are different. The variational principle implies that the computed energetic difference is artificially increased, as the complex's wavefunction is expanded in a larger basis set compared to that of the fragments forming the complex. This effect was called Basis Set Superposition Error (BSSE) and was firstly pointed out by Jansen and Ros[12] in 1969, although the terminology BSSE was first introduced by Liu and McLean[13] in 1973.

Since this problem was first addressed, a number of methods and strategies have been proposed to eliminate the BSSE or, at least, minimize it. Obviously, the most natural way of eliminate it would be the use of exact wave functions, thus avoiding the truncation of the basis set. The use of an infinite basis set is not feasible from a computational point of view but the basis sets of the interacting fragments could be improved in such a way that the presence of the basis functions of the partner would not further improve the description of the considered fragment.

Another rather unexplored possibility is to use a set of functions centered at some given points in the space to compute the energies of the complex and its fragments. The three-dimensional space might be saturated of basis functions whose position and parameters should be kept constant for each calculation. However, the problem in this case would be that the wavefunction obtained will not be neither translationally nor rotationally invariant. Also, linear dependencies could easily

[11] B. Jeziorki, W. Kolos. *Molecular Interactions*. H. Ratajczak, W. J. Orville-Thomas, Eds.; Wiley, New York, 1982; Vol 3.

[12] H. B. Jansen, P. Ross, Chem. Phys. Lett. 3, 140 (1969)

[13] B. Liu, A. D. McLean, J. Chem. Phys. 59, 4557 (1973)

appear in the basis set if the space is too saturated with basis functions.

A promising tool to obtain good quality ab-initio DFT results within the local spin density approximation (LSDA) without the use of nuclei centered basis sets are the so-called plane waves. They are nowadays widely used in Car-Parrinello molecular dynamics[14] but there is not yet much work in electronic structure calculations. In this particular case, the Kohn-Sham molecular orbitals are expanded on a set of imaginary functions (plane waves) independent of the position and number of nuclei of the molecule. The real space representation of the wavefunction is obtained after the application of Fast Fourier transform techniques. Unfortunately, recent work[15] is pointing towards the combination of both nuclei-centered and plane waves so that an eventual future application to intermolecular interactions will bear a similar basis set unbalance.

I.1.1.1 The counterpoise philosophy

The discussion about the existence, definition, consequences or usefulness of the correction schemes of the BSSE in the literature has been very active during the last twenty years and it is still under debate. Most of these controversies are still due to the BSSE-correction scheme based on the earlier work of Boys and Bernardi in 1970. They proposed the use of a function counterpoise (CP)[16] in order to calculate the interaction energy of a given AB system in such a way that the separate energies of the fragments A and B are calculated using the full set of basis functions used in the calculation of the energy of the AB system. Practically speaking, for each fragment calculation, the electrons belonging to the other fragment were omitted and its nuclear charges were set to zero.

After the counterpoise method was introduced, there was a widespread opinion that the BSSE was being overestimated with this straightforward procedure. The arguments given by several authors[17,18] were that with the original function

[14] R. Car and M. Parrinello, Phys. Rev. Lett. 55, 2471 (1985)

[15] M. Valiev and J. H. Weare, J. Phys. Chem. A 103, 10588 (1999)

[16] S. F. Boys, F. Bernardi, Mol. Phys. 19, 553 (1970)

[17] A. Johansson, P. Kollman, S. Rothenberg, Theor. Chim. Acta 29, 167 (1973)

[18] J. P. Daudey, P. Claverieand J. P. Malrieu, Int. J. Quantum Chem. 8, 1 (1974)

counterpoise scheme, the fragment energies were evaluated using the whole basis set of the complex, thus including the orbitals of the other fragment (ghost orbitals) *to be occupied* in the complex formation. It was proposed another counterpoise scheme where the fragment energies were calculated extending the monomer basis set with just the ghost virtual orbitals of the partner. However, the BSSE is clearly an unphysical effect, hence no arguments based on physical effects like the Pauli exclusion principle in this particular case can be applicable in our opinion. Further numerical results[19,20] on the helium dimer showed that the use of the whole basis set for the fragment calculations was necessary in order to reproduce the inherently BSSE-free Intermolecular Perturbation Theory results.

The overcorrection debate has continued in the literature mainly due to results obtained upon inclusion of electron correlation. It was argued that the counterpoise method should be able to improve the description of the monomer's orbitals yet avoiding an increase of the number of virtual orbitals used for the correlation[21]. This opinion is connected to the observation that, in some cases, when using small basis sets, the counterpoise-corrected results seem to disagree more with the experimental values than the uncorrected results; see for instance recent papers of Liedl[22], Oliphant et al.[23] or Masamura[24]. It has been also proposed to scale the BSSE estimated with the counterpoise method, in order to correct for its supposed overcorrection nature. Kim et al.[25] have been using half of the counterpoise correction as the estimate of the BSSE, obviously with no solid argument to support this, more than ambiguous, choice.

Some authors[26,27] also concluded that the best strategy would be to use the

---

[19] M. Gutowski, J. J. Van Lenthe, J. Verbeek, F. B. Van Duijneveldt and G. Chalasinski, Chem. Phys. Lett. 124, 370 (1986)

[20] M. Gutowski, F. B Van Duijneveldt, G. Chalasinski and L. Piela, Mol. Phys. 61, 223 (1987)

[21] D. B. Cook, J. A. Sordo and T. Sordo, Intl. J. Quantum Chem. 48, 385 (1993)

[22] K. R. Liedl, J. Chem. Phys. 108, 3199 (1998)

[23] N. Oliphant, M. Rosenkrantz and D. D. Konowalow, Chem. Phys. Lett 223, 7 (1994)

[24] M. Masamura, Theor. Chem. Acc. 106, 301 (2001)

[25] K. S. Kim, P. Tarakashemar, J. Y. Lee, Chem. Rev. 100 4145 (2000)

[26] D. W. Schwenke and D. G. Truhlar, J. Chem. Phys. 82, 2418 (1985); 84, 4113 (1986); 87, 3760 (1987)

[27] M. J. Frisch, J. E. Del Bene, J. S. Binkley and H. F. Schaeffer III, J. Chem. Phys. 84, 2279 (1986)

largest basis set possible and simply ignore the BSSE. However, it has been shown in many benchmark studies[24,28] that the convergence of the BSSE correction is very slow at the correlated level and no BSSE-free results have been obtained so far.

At the same time, there have been other attempts to prove the correctness of the counterpoise methods, both numerically[19,20,29,30] and with theoretical arguments[31]. The counterpoise method has even been considered a theorem as in case of the full-CI wavefunctions[31]. However, none of these efforts have helped to end with the controversy.

I.1.1.2 Beyond the CP-correction to the interaction energy

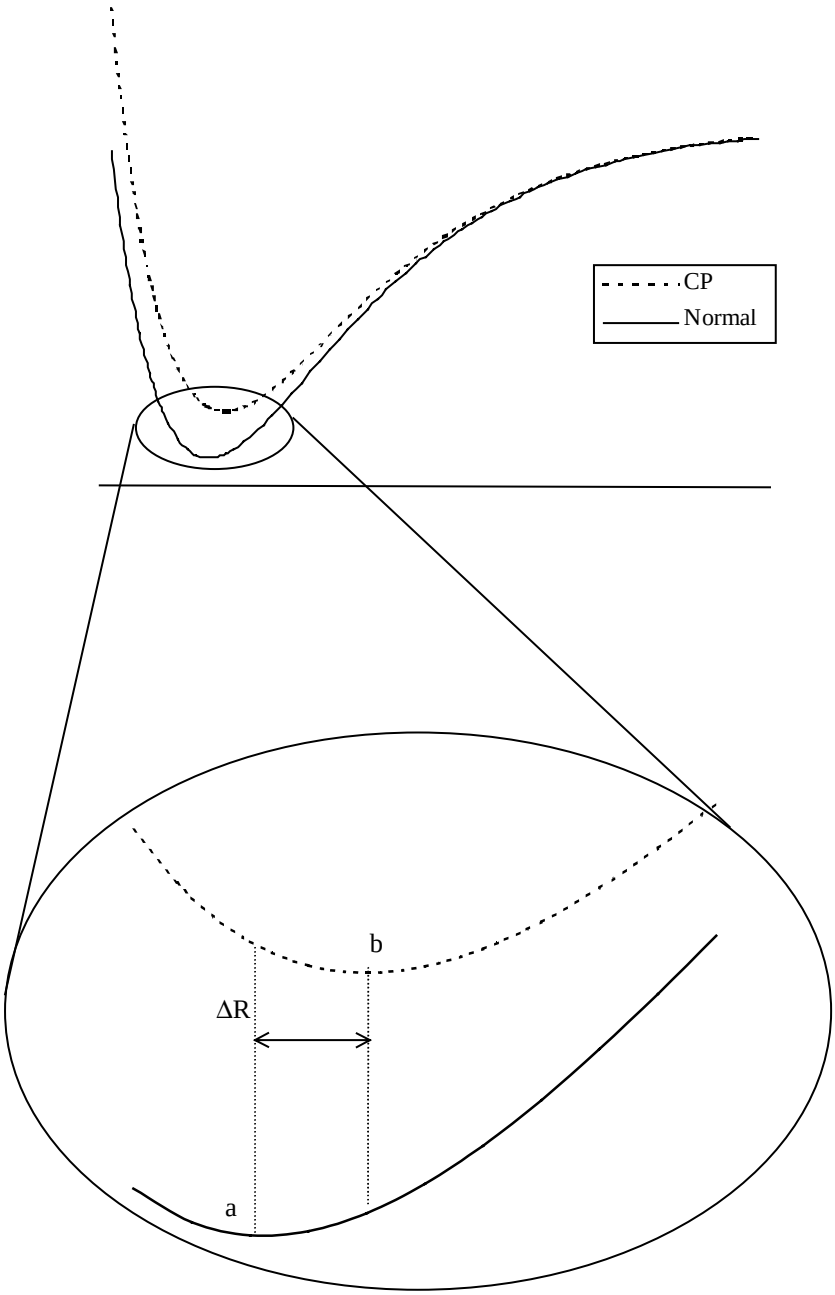
At the time when the BSSE was pointed out, the *ab initio* calculations in molecular complexes were mainly used to explore the interaction potentials. It was recognized that the BSSE induced to curves which are too attractive and that the counterpoise-corrected ones lie above the uncorrected ones. The obvious anharmonicity of the interaction potential curves and the fact that the basis set extensions increases when the fragment get closer induced unavoidably to another side effect: the position of the minimum of the curve, i. e., the optimum interaction energy, had to change upon counterpoise correction too. This can be easily seen in the following qualitative interaction potential picture

[28] T. Van Mourik, A. K. Wilson, K. A. Peterson, D. E. Woon, T. H. Dunning Jr, Adv. Quantum Chem. 31, 105 (1999)

[29] S. M. Cybulski anf G. Chalasinski, Chem. Phys. Lett. 197, 591 (1992)

[30] M. Gutowski, J. G. C. M. van Duijneveldt-van de Rijdt and J. H. van Lenthe, F. B. Van Duijneveldt, J. Chem. Phys. 98, 4728 (1993)

[31] F. B. Van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt and J. H. van Lenthe, Chem. Rev. 94 1873 (1994)



SCHEME 1

The counterpoise-corrected interatomic potential lies above the uncorrected one and the position of the optimum interatomic distance is displaced by ΔR. The shape of the PES is modified as well. Furthermore, the counterpoise-corrected interaction energy computed at the uncorrected minimum (a) is always larger than (b) the optimum interaction energy calculated on the corrected PES. This difference depends

on the shape of the PES. The more shallow the potential well the larger the effect on the interatomic distance.

Obviously, the interacting systems could be molecules as well. In this case, the molecular complex must be partitioned into two or more interacting molecules or fragments. One of the first attempts to assess the effect of the BSSE on the geometry of the molecular complexes was carried out by Diercksen et al.[32] in a SCF study of the water dimer. They determined numerically that the intermolecular distance increased by 0.04Å upon counterpoise-correction. Other authors[33] reported similar effects on both hydrogen-bonded and van der Waals complexes calculations including electron correlation, also by performing numerical optimizations of the intermolecular distance.

Soon, the simultaneous optimization of many geometrical parameters became feasible so that the effect on the intramolecular parameters due to the interaction could be assessed. This is essential in order to connect the theoretical values with experiment. Indeed, one of the typical fingerprints of a molecular complex is the shifting on the IR absorption bands of the isolated complexes. In the case of the hydrogen bonded complexes X-H···Y, this red (blue) shift is due to the elongation (shortening) of the intramolecular X-H bond upon interaction with Y[34].

However, the fact that the geometries of the interacting molecules are distorted by the physical interaction was not taken into account in the counterpoise scheme in its original formulation, because only atomic interactions were considered. The simultaneous consideration of this geometry relaxation, usually referred to as relaxation energy, and the BSSE using the counterpoise method is not trivial. This problem was addressed by Emsley et. al.[35] in an early study of the hydrogen bonding in the biformate anion. The total counterpoise-corrected interaction energy was

[32] G. H. F. Diercksen, W. P. Kraemer, B. O. Roos, Theor. Chim. Acta 36, 249 (1975)

[33] J. M. Leclerq, M. Allavena and Y. Bouteiller, J. Chem. Phys. 78, 460 (1983), R. A. Kendall, J. Simons, M. Gutowski, G. Chalasinski, J. Phys. Chem. 93, 621 (1989); M. D. Newton, N. R. Kestner, Chem. Phys. Lett. 94, 198 (1983); R. Eggenberger, S. Gerber, H. Hubert, D. Searles, chem. Phys. Lett. 183, 223 (1991); J. G. C. M. van Duijneveldt-van der Rijdt, F. B. van Duijneveldt , J. Chem. Phys. 97, 6104 (1992).

[34] This topic is discussed later in Section II.1.1.3.

[35] J. Emsley, O. P. A. Hoyte, R. E. Overill, J. Am. Chem. Soc. 100, 3303 (1978)

calculated as the sum of the counterpoise-corrected interaction energy at the supermolecular geometry and the fragment relaxation energy. Therefore, seven energy calculations had to be carried out, instead of the three energy evaluations needed for both an uncorrected or a counterpoise-corrected with the fragments geometries frozen, for the determination of the interaction energy. The authors deduced that their results were approximate since the fragment relaxation energy was carried out using the fragment basis set, in contrast with the counterpoise-corrected interaction energy calculated with the dimer basis set. The key point is that the conventional counterpoise method cannot be applied unambiguously for fragment geometries different from those within the complex, because the position of the ghost orbitals for each fragment calculation becomes undetermined. Mayer and Surján[36] showed that the procedure outlined above was the only consistent way of introducing the fragment relaxation in the counterpoise scheme.

Therefore, when dealing with molecular complexes, one must distinguish between the interaction energy and the so-called stabilization energy. The former stands for the energetic difference between the complex energy and that of their subunits computed at the supermolecular geometry. The latter represents the total stabilization energy resulting from bringing the fragments from an infinite separation to the equilibrium distance. Both terms are often used in the literature ambiguously.

The interaction energy is, nevertheless, a useful theoretical concept. It can be exactly computed by means of Intermolecular Perturbation Theory (provided that the series converges). It allows for a further partitioning into different terms of more or less clear physical significance and therefore gives insight into the nature of the interaction. Another useful concept that can be derived is the so-called interaction density, that is the difference between the electron density distribution of the molecular complex and that of the free interacting molecules. A typical isosurface map of the interaction density shows the polarization of the electron density of each fragment induced by the presence of the other. Note that since the electron density is mainly localized in the atomic positions, a density difference map comparing electron densities of different geometries provides no meaningful chemical information in terms of redistribution of the electron density due to charge densities overlap,

[36] I. Mayer, P. R. Surján, Chem. Phys. Lett. 191, 497 (1992).

induction, polarization, etc.

On the other hand, the stabilization energy provides the comparison with the experimental results. Neither the thermometric nor the spectroscopical techniques can provide such insight in the interaction energy, not even measure the relaxation energy.

There is a substantial difference between those experimental techniques. Direct measures of the formation or dissociation energies yield no information of the structural features of the complex. Of course, no *pure* energies can be measured; instead, thermodynamic quantities such as enthalpies or free energies are determined. That is, translational, rotational, vibrational thermal corrections and zero-point vibrational energy contributions must be added to the computed stabilization energies.

Spectroscopic techniques do provide indirect insight in the geometry of the complexes and the dynamical properties of the interaction. In the case of solid state, X-ray and neutron diffraction give a averaged picture of the position of the atoms in the crystal. Microwave and IR techniques are used to provide reliable interaction potentials by fitting the experimental data to a parametrized potential, which will provide information about the equilibrium configuration and zero-point oscillation amplitudes[37]. NMR techniques can be used also to investigate proton transfers or tautomerism equilibrium in solution and solid state.

From such a point of view, the theoretical modeling of the weak interactions should provide not only accurate interaction energies, but also structural parameters of the complexes, in order to correctly reproduce the anisotropy of the interaction, dipole moments, electron density distribution, etc. When haunting for the accuracy, the BSSE can represent and important source of error, yet keeping in mind that a correct elimination of the BSSE does not ensure a good agreement with the experiment. Indeed, efficient error cancellation sometimes induces to consider a given methodology as accurate. Hence, correction of BSSE, anharmonicity, size consistency and extensivity and proper inclusion of dynamical and statical correlation should be added to the widely used *black-box* methodologies when looking for the chemical precision.

In this work we focus on the efficient elimination of the BSSE in the determination of *any* molecular property of the molecular complexes. That is, not just

[37] K. R. Leopold, G. T. Fraser, S. E. Novick, W. Kemplerer, Chem. Rev. 94, 1807 (1994), and references therein

BSSE-corrected interaction energies, but also formation enthalpies[38], geometrical parameters[39,40], force constants[41], dipole moments[42], rotational[43] and tunneling barriers[41], wavefunction and electron densities[44,45], one and two-center component energies[46] or many-body energetic contributions[47].

To this end, the counterpoise method has been considered as a correction term to the description of the molecular complex. This new interpretation[48] of the counterpoise method will be reviewed in Section I.1.3.

I.1.1.3 Aprioristic methodologies

Another subject being discussed has been the so-called *a priori* correction methodologies. In this case, rather than recalculating the fragments in the complex's basis set, these methods try to eliminate the BSSE at the complex's calculation stage by imposing some restrictions on the molecular orbitals, depending on the method. The first attempt was carried out by Mayer[49] in 1983, with the so-called Chemical Hamiltonian Approach (CHA), which will be discussed in more detail in Section I.1.4. Briefly, at the SCF level of theory, Mayer used second quantization techniques to split the Hamiltonian into the sum of all the *intramolecular* operators and the pure-interaction intermolecular operator. Then, the BSSE is removed by projecting all the intramolecular terms into the subspace spawned by the basis functions of the

[38] P. Salvador, S. Simon, M. Duran, J. J. Dannenberg, J. Chem. Phys., submitted

[39] P. Salvador, B. Paizs, M. Duran, S. Suhai, J. Comput. Chem. 22, 765 (2001)

[40] B. Paizs, P. Salvador, A. G. Császár, M. Duran, S. Suhai, J. Comput. Chem 22, 196 (2001)

[41] P. Salvador, S. Simon, M. Duran, J. J. Dannenberg, J. Chem. Phys. 113, 5666 (2000)

[42] M. C. Daza, J. A. Dobado, J. Molina, P. Salvador, M. Duran, J. L. Villaveces, J. Chem. Phys. 110, 11806 (1999).

[43] P. Salvador, M. Duran, J. Chem. Phys. 111, 4460 (1999)

[44] P. Salvador, X. Fradera, M. Duran, J. Chem. Phys. 112, 10106 (2000)

[45] P. Salvador, X. Fradera, M. Duran, Adv. Mol. Sim.,Vol 4 *(accepted).*

[46] P. Salvador, M. Duran, X. Fradera, J. Chem. Phys., *(submitted.)*

[47] P. Salvador, M. M. Szczesniak, J. Chem. Phys., *(submitted)*

[48] S. Simon, M. Duran, J.J. Dannenberg, J. Chem. Phys. 105, 11024 (1996)

[49] I. Mayer, Int. J. Quantum Chem. 23, 341 (1983)

corresponding molecular fragment. The final result is a description of the complex where the BSSE has been eliminated, according to the defined fragments, with no extra *a posteriori* calculations.

So far, the method has been applied at the SCF, MP2, DFT and CI levels of theory with satisfactory results. The CHA and counterpoise methods correct for the BSSE in a very similar way, and the most important fact is that "as the basis set improves both methods converge to each other much faster than the BSSE disappears from the interaction energies"[50]. Despite these findings, the non-hermitian nature of the resulting BSSE-free Hamiltonian and the fact that the energy must be computed using a different Hamiltonian (the conventional one) produced an skeptical reaction in the scientific community[31,51].

In another method, the so-called Constrained Dimer Function approach (CDF)[52], the BSSE is supposedly removed by including some extra constraints, with the corresponding set of Lagrange multipliers, to the energy expression of the supermolecule at the SCF level of theory. These constrains ensure that the occupied orbitals of the supermolecule are not mixed with the approximate solutions of the monomers. The solution of the final equations was derived using a perturbational scheme. Only the leading term of the expansion was analyzed and no proof of the convergence of the series was given. Moreover, Gutowski and Chalasinski[51] proved that the constraints imposed implied that the component of the occupied orbitals of the fragment A which is orthogonal to the orbitals of fragment B is also excluded from the dimer orbital space. Numerical calculations[53] also pointed out the degradation of the supermolecule description and hence no further development of the method was performed.

Later on, other approaches have faced the BSSE problem in a similar manner but all of them based in the (wrong) assumption that at the SCF level the BSSE can be eliminated by imposing the molecular orbitals of each fragment to be expanded exclusively on its own basis set.

---

[50] I. Mayer, Int. J. Quantum Chem. 70, 41 (1998)

[51] M. Gutowski, G. Chalasinski, J. Chem. Phys. 98, 5540 (1993)

[52] A. J. Sadlej, J. Chem. Phys. 95, 6705 (1991)

[53] I. Mayer, J. Chem. Phys. 97, 5257 (1992); A. J. Sadlej, *ibid.* 97, 5259 (1992)

The first attempt, the Strictly Monomer Molecular Orbital (SMMO) SCF approach, was developed by Cullen[54] in 1991. In this method, the BSSE-corrected supermolecule description is obtained by imposing to the SCF equations that each molecular orbital must be localized into an individual fragment, and then expanded on the fragment basis set. The author observed that the interaction energies obtained with this method were systematically less attractive than the counterpoise-corrected ones. Also, the difference was much larger at short intermolecular distances. He assigned this effect to the neglect of the charge transfer contribution due to the special form of the wavefunction. Indeed, Gutowski and Chalasinski[51] pointed out that not just the fragment energies but also the interaction energy are expanded strictly on the fragment basis sets. That means that not only the charge-transfer contribution is neglected but also other components of the interaction like the electrostatic and the exchange are affected.

An analogous version, the co-called SCF for molecular interactions (SCF-MI) was proposed by Gianinietti et al.[55] a few years ago and very recently reformulated using a locally projected technique by Iwata[56] and coworkers. Again, the lack of some of the true-interaction terms leads to BSSE-overcorrection. This can be clearly observed when increasing the basis set size. For the water dimer complex , it was observed[57] that for a basis set like the TZP++, and systematically increasing the basis set size to up to 548 functions, the SCF-MI interaction energy was 3.33kcal/mol, whereas the uncorrected one was converged to 3.71kcal/mol. The difference in the intermolecular distance was also larger than 0.1Å. Hence, the desired asymptotic behavior of the uncorrected and the BSSE-corrected energies or any parameter in

[54] J. M. Cullen, Int. J. Quantum Chem. Symp. 25, 193 (1991)

[55] E. Gianinetti, M. Raimondi, E. Tornaghi, Int. J. Quantum Chem. 60, 157 (1996);A. Famulari, E. Gianinetti, M. Raimondi, M .Sironi, I Vandoni, Theor. Chem. Acc. 99, 358 (1998)

[56] T. Nagata, O. Takahashi, K. Saito S. Iwata, J. Chem. Phys. 115, 3553 (2001)

[57] C. Gatti A. Famulari, C. Gatti and A. Famulari, Interaction Energies and Densities. A Quantum Theory of Atoms in Molecules Insight on the Effect of Basis Set Superposition Error Removal, Kluwer book series, *Understanding Chemical Reactivity: Electron, Spin and Momentum Densities and Chemical Reactivity*, Vol 2; Mezey P. G.; Robertson, B, eds; Kluwer Dordrecht, The Netherlands, 1998.

general, was not observed. It is worth noting that both the counterpoise and the CHA results have proved to converge to the uncorrected ones when using large basis sets, as will be shown in this work.

In the spirit of the two methods discussed above, Muget and Robinson[58] proposed a much simpler self-consistent algorithm based on the localization of the molecular orbitals. In this case, the orbital localization is not imposed in the SCF equations. Contrarily, the canonic orbitals obtained from a conventional dimer SCF calculation are localized on the monomers, reorthogonalized and further used to build up the new Fockian. This process is carried out until self-consistency of the wavefunction is achieved. Their interaction energy results for the ammonia dimer showed the expected pathological behavior of such BSSE-correction methods. Whereas the counterpoise-correction was converged to zero upon increasing the basis set size, their BSSE values were far from being converged. Another rather surprising effect was that the inclusion of diffuse functions in the basis set greatly enhanced the BSSE extent in the interaction energies, which is in total disagreement with the counterpoise and CHA results.

As for the electron correlation level, several methods[59,60,61] have been also proposed in order to minimize the BSSE. The most successful ones have been the so-called local correlation methods[61]. Curiously, the main goal of these methods was to reduce the computational cost of the application of high-level electron correlation methods to large systems, while the BSSE reduction can be considered as a side effect. In this method, the supermolecule occupied orbitals are localized and kept orthonormal. Then, a set of virtual orbitals obtained by projecting out the occupied space from the atomic orbitals is assigned to each localized molecular orbital (LMO). The excitations involved in the corresponding correlated method are restricted to the local domains so that the double excitations of the electrons of a fragment into the localized virtual orbitals of the partner are automatically excluded. The method has

[58] F. F. Muguet, G. W. Robinson, J. Chem. Phys. 102, 3648 (1995)

[59] P. Wind, J. –L. Heully, Chem. Phys. Lett. 230, 35 (1994)

[60] I. Røeggen, H. Skullerud, J. Phys. B 25, 1795 (1992)

[61] S. Sæbø, P. Pulay, J. Chem. Phys 86, 914 (1987); S. Sæbø, W. Tong, P. Pulay, J. Chem. Phys 98, 2170 (1993); C. Hampel, H. –J. Werner, J. Chem. Phys. 104, 6286 (1996) and references therein.

been applied at the MPn, n=2-4 and CCSD levels of theory. Some recent calculations at the MP2 level of theory (LMP2)[62] have shown that the BSSE is dramatically decreased, compared to the full-counterpoise method. The electron correlation contribution is actually negative, that is, the estimated CP-correction with the LMP2 method is smaller than the HF+MP2 counterpoise correction. This shows that, obviously, not only the BSSE is eliminated from the correlation energy but also some true-correlation contribution is lost in the local approximation.

In a recent investigation, Schütz[63] et al. studied small water clusters at LMP2 level of theory, where they showed that the geometrical parameters and stabilization energies obtained with the LMP2 method for the water dimer were comparable to the counterpoise-corrected ones obtained by Xantheas[64].

This method, however, cannot be considered as the ultimate solution for the BSSE correction. Indeed, the LMP2 method does represent a valuable way to reduce the BSSE at the correlated level, although, by construction, it is unable to deal with the BSSE at the HF level, since the localized molecular orbitals are obtained from an *uncorrected* SCF calculation. Therefore, in case of a large BSSE effect at the uncorrelated level, the LMP2 method might fail to predict reliable results, and a method that eliminates the BSSE at both the uncorrelated and correlated level, like the CHA-MP2[65], may be desirable.

In this work, as a paradigm of the *aprioristic* methodologies, the Chemical Hamiltonian Approach has been used in order to obtain BSSE-corrected wavefunctions at HF and DFT levels of theory. The results are given in Section II.1.3.

On the other hand, the restricted and unrestricted versions of the CHA-SCF equations for an arbitrary number of fragments have also been implemented in a series of external programs interfaced to the HONDO package. The details of the novel developed algorithm are presented in Section II.2.

---

[62] A. El Azhary, G. Rauhut, P.Pulay, H. –J. Werner, J. Chem. Phys. 108, 5185 (1998)

[63] M. Schütz, G. Rauhut, H. –J. Werner, J. Phys. Chem. A 102, 5997 (1998)

[64] S. S. Xantheas, J. Chem. Phys. 104, 8821 (1996)

[65] I. Mayer, P. Valiron, J. Chem. Phys. 109, 3360 (1998)

I.1.1.4 Current trends

Nowadays, the counterpoise method is still the most widely used in the literature. Most authors take the BSSE into account when performing calculations of intermolecular complexes. In some particular cases, however, there is a marked tendency to try to use the largest basis set possible or simply neglect the effect of the BSSE. The most controversial situations occur when dealing with open-shell complexes, charged molecular complexes, molecular aggregates composed by more than two fragments, or intramolecular processes like rotational barriers or proton transfer reactions. These particular cases have been addressed in this work and will be discussed in more detail in Section I.1.6.

Regarding the *a priori* schemes, only the CHA and SCF-MI methods are currently under development. During the last months prior to the completion of the present work, Mayer succeeded to reformulate his CHA-SCF equations by using a Hermitian Fock matrix[66], which will hopefully lead to significant simplification on the further treatment of the electron correlation, as well as facilitate the use of gradient techniques with a reduced computational cost.

Also, the SCF-MI method has been implemented recently in standard *ab initio* programs but the limitations of this approach, commented above, have prevented the scientific community at the present time to adopt this methodology in their calculations.

It is worth to note also that BSSE-free Intermolecular Perturbation Theory methodologies like the SAPT have been also improved during the last years and they are becoming more and more attractive, mainly for the study van der Waals complexes. The implementation of the methodology for three fragments, the derivation of equations for open-shell systems,[67] and the inclusion of the intramonomer correlation corrections to the interaction energy are the most important achievements in the last few years. Furthermore, the SAPT method, together with the conventional (counterpoise-corrected) supermolecular MPn results, can be seen also as an energy decomposition scheme at the correlated level.

---

[66] I. Mayer, *Hermitized Fockian in the Chemical Hamiltonian Approach: Fulfilling a Löwdin Prediction,* submitted for publication.

[67] S. M. Cybulski, R. Burcl, G. Chalasinski, M. M. Szczesniak, J. Chem. Phys. 103, 10116 (1995)

Even though it is a methodology which is free of BSSE by construction, the SAPT results depend upon the considered geometry of the complex, which in turn can be strongly affected by the BSSE. In this work this dependence has been also numerically investigated. The results obtained are presented in Section II.1.4.1.

I.1.2 The actual implications of the BSSE

As commented in the previous section, the counterpoise-correction has been traditionally related intimately to the concept of the interaction energy. Furthermore, the existence of the BSSE has been connected *only* to the calculation of interaction energies. The fact that the BSSE also affects the shape of the PES and thus the position of the stationary points and curvature of the PES forces to a reinterpretation of the problem.

Let us assume that one might be interested in the structure of a given molecular complex, but not in the interaction energy that brings it weakly bonded. One can argue that a conventional geometry optimization, followed by the corresponding vibrational analysis, is free of the BSSE since the molecular complex is considered as a whole. That is, no partitioning of the complex seems necessary if there is no intention of computing the interaction energy.

In a second step, let us consider that the given complex is actually composed by two monomers (fragments), and we do are interested in the stabilization energy. Once the monomeric species are defined, the next step would be to compute the respective absolute energies of the monomers and substract them from the complex energy. However, in this case the BSSE appears because one is substrancting quantities obtained using a projection of the exact values on different Hilbert spaces. The unbalance is due to the monomer calculations so that the counterpoise philosophy is to compute the monomer energies in the dimer centered basis set (DCBS), instead of the monomer centered basis set (MCBS). For the sake of simplicity, let us assume that nuclear relaxation effects upon complexation are negligible. Now, since the monomer energies depend on the intermolecular parameters, because they determine the position of the DCBS, the optimized complex geometry do not longer correspond to the stationary point of the counterpoise-corrected interaction energy. In other words, the counterpoise-correction induces to a new geometry of the molecular complex whose interaction energy is optimum (point b in SCHEME 1), different from the former geometry that was assumed to be *correct* because no MCBS were defined (point a in SCHEME 1).

One has two options now; either the DCBS description of the complex bears

some BSSE or the MCBS monomer energies are the only inconsistency. If one sticks to the latter case, two different essentially correct geometries need to be defined, depending upon whether the determination of the interaction energy is the goal or it is not. Actually, it could be argued that whenever the definition of fragments is needed, in essence any MCBS, the counterpoise-corrected geometry (PES) should be used. Otherwise, the DCBS dimer geometry is the correct one.

Now let us go one step further. Suppose that one is interested in the electron density distribution (wavefunction) of the complex, let us say to obtain the interaction density. The arising question is obvious. Should the wavefunction be computed at the uncorrected geometry of the complex or at the counterpoise-corrected one? In the first case, the resulting wavefunction will include the BSSE effect indirectly, as it is computed at a geometry which is inconsistent with the monomer DCBS calculations, that is, with the one we are using to determine monomer-dependent properties of the complex. In the latter case, the wavefunction of the complex will not be that of the equilibrium geometry because it is not computed at the complex's DCBS optimum geometry.

This contradiction forces us to consider the first option proposed, that is, the DCBS description of the dimer includes some BSSE. Even if there is no need to define the MCBS basis sets, the BSSE appears in the calculation of the dimer. Therefore, it is essential to assign *any* BSSE correction to the energy (description) of the molecular complex, namely counterpoise or *aprioristic*, which by definition already acts on the dimer.

However, the implications of this reasoning generate another ambiguity. The DCBS description of the dimer is already BSSE-contaminated with no previous definition of the MCBS or fragments, whereas different partitioning of the complex leads to different BSSE-corrected descriptions of the system. In other words, there can be several BSSE-corrected descriptions of the complex depending upon the definition of the MCBS. At this point one may wonder what makes the weakly bound complexes different from the strongly covalent or ionic bonded molecules. In fact, there is no physical difference, other than the nature and strength of the interaction between them. And obviously there is no mathematical difference either. *Ab initio* calculations do not require a previous knowledge of the nature and number of bonds of the chemical system.

All these considerations point out that the existence of the BSSE is not restricted

to the molecular interaction calculations. The BSSE is a *intramolecular* phenomenon due to the use of truncated atomic centered basis sets and it is present in any calculation. Moreover, since the calculation of a chemical system is inherently BSSE-contaminated, the definition and exact correction of the BSSE must come from the chemical system itself, and not out of any extra subunit contribution.

This means that the counterpoise correction cannot be considered as an exact methodology, but a very good approximation and efficient simulation of the BSSE effects within a molecular complex. This does not imply that the counterpoise theorem is wrong. The counterpoise method do can provide an *exact* BSSE-free interaction energy. In a sense, the interaction energy is only a specific characteristic value that comes from a more general entity, i. e., that can be extracted from the wavefunction of the quantum mechanical description of the systems. In our opinion, since the counterpoise method can not provide this wavefunction, the counterpoise-corrected description of a chemical system can not be considered exact within the given basis set.

The only definition of the BSSE consistent with the considerations expressed above was given by Mayer in his Chemical Hamiltonian Approach[49]. In his work, the existence of an intramolecular BSSE is revealed by the analysis of the effects of the truncation of the basis set expansion on each atom of the chemical system.

Unfortunately, the exploration of the CHA at the atomic level did not provide satisfactory results for a main reason. First of all, the existing basis sets are optimized in order to reproduce a given molecular property, namely dipole moments, polarizabilities, formation enthalpies, etc... For the construction of these basis sets, a set of different molecules are used to fit the experimental or desired results to the predicted ones obtained from a given *ab initio* method. However, those molecules were not free of the atomic BSSE so that its elimination results in poor results compared to the uncorrected ones. In the case of weakly bonded systems, or molecular aggregates, the correction of the *intermolecular* BSSE usually results in a improvement of the molecular complex description because the sample molecules from which the basis sets were parametrized were free of this specific BSSE.

This is why the intramolecular basis set extensions are necessary for a good description of the chemical systems by means of *ab initio* techniques. It is meaningless to try to eliminate all the basis set extension in a calculation as long as an essential component of the *ab initio* methods, like the basis sets, are already taking

them into account.

If the basis sets were constructed with a proper elimination of the intramolecular BSSE, then there would be no ambiguities in the definition of the fragments in a molecular aggregate since each atom would become a fragment. Chemical reactions where the number and definition of the *classical* molecular fragments changes could then be easily studied. Probably, the much-sought transferability of the atoms within molecules could be more accurately reached.

This is however out of the scope of the present work and it will be limited to the particular, yet very important, case of the elimination of the *intermolecular* BSSE on molecular complexes.

I.1.3 An alternative interpretation of the CP-correction

Now, it will be shown how to trivially consider the counterpoise-correction as a correction of the complex energy, and how this can be used to define counterpoise-corrected gradients, vibrational frequencies, dipole moments and any derivative of the molecular complex energy.

Let us consider a supermolecule AB made up of two interacting subsystems A and B. The stabilization energy can be expressed as:

$$\Delta E(AB) = E_{AB}^{AB}(AB) - E_{A}^{A}(A) - E_{B}^{B}(B) \qquad (1)$$

We define $E_{Y}^{Z}(X)$ as the energy of a subsystem X at geometry Y with basis set Z. The stabilization energy can be split in the following way

$$\Delta E(AB) = \Delta E_{int}(AB) + \Delta E_{rel}(A,B) \qquad (2)$$

The first term represents the interaction energy contribution, which depends only on the supermolecule geometrical parameters, {AB}

$$\Delta E_{int}(AB) = E_{AB}^{AB}(AB) - E_{AB}^{A}(A) - E_{AB}^{B}(B) \qquad (3)$$

whereas the second (positive definite) term represents the relaxation contribution, which compensates for the geometry distortion of the subsystems in the supermolecule, $E_{AB}^{A}(A)$ and $E_{AB}^{B}(B)$, with regard to the isolated optimum geometry, $E_{A}^{A}(A)$ and $E_{B}^{B}(B)$.

$$\Delta E_{rel}(A,B) = E_{AB}^{A}(A) - E_{A}^{A}(A) + E_{AB}^{B}(B) - E_{B}^{B}(B) \tag{4}$$

Note that $\Delta E_{rel}(A,B)$ depends on both the supermolecule and subsystem parameters, {AB, A, B}.

According to the counterpoise idea, since the same basis set is used in the relaxation term for each subsystem, only the interaction energy contribution term brings about BSSE. Thus, the counterpoise-corrected interaction energy should be written as:

$$\Delta E^{CP}(AB) = \left[E_{AB}^{AB}(AB) - E_{AB}^{AB}(A) - E_{AB}^{AB}(B)\right] + \left[E_{AB}^{A}(A) + E_{AB}^{B}(B) - E_{A}^{A}(A) - E_{B}^{B}(B)\right] \tag{5}$$

Rearranging terms of the expression above one obtains

$$\begin{aligned}\Delta E^{CP}(AB) = &\left[E_{AB}^{AB}(AB) - E_{A}^{A}(A) - E_{B}^{B}(B)\right] + \\ &+ \left[E_{AB}^{A}(A) + E_{AB}^{B}(B) - E_{AB}^{AB}(A) - E_{AB}^{AB}(B)\right] = \Delta E(AB) + \delta_{AB}^{BSSE}\end{aligned} \tag{6}$$

where the CP-correction expressed as δ_{AB}^{BSSE} presents the following properties:

- Tends to zero as the fragment's basis set (MCBS) approaches completeness.
- Due to the variational principle, it is a positive definite value (definite no negative in case of complete MCBSs)
- Depends upon the geometrical parameters of the complex, as shown by the subindex AB

The third property implies that the BSSE is not an additive term to the

stabilization energy. Actually can be strongly geometry-dependent so that it should be taken into account at every point of the potential energy surface when looking for stationary points.

Indeed, any stationary point of the uncorrected supermolecule PES determines a stationary point of the interaction energy surface, because there is no variation in the isolated subsystems. Differentiating Eq. (5)

$$\frac{\partial\left(\Delta E^{CP}(AB)\right)}{\partial R_i}=\frac{\partial(\Delta E(AB))}{\partial R_i}+\frac{\partial\left(\delta_{AB}^{BSSE}\right)}{\partial R_i}=\frac{\partial\left(E_{AB}^{AB}(AB)\right)}{\partial R_i}+ \frac{\partial\left(E_{AB}^{A}(A)\right)}{\partial R_i}+\frac{\partial\left(E_{AB}^{B}(B)\right)}{\partial R_i}-\frac{\partial\left(E_{AB}^{AB}(A)\right)}{\partial R_i}-\frac{\partial\left(E_{AB}^{AB}(B)\right)}{\partial R_i} \qquad \forall R_i \in \{AB\} \tag{7}$$

it can be seen that the uncorrected supermolecular geometry is only valid under the assumption that the BSSE defined as δ_{AB}^{BSSE} is stationary at the current nuclear arrangement.

$$\frac{\partial\left(\delta_{AB}^{BSSE}\right)}{\partial R_i}=0 \qquad \forall R_i \in \{AB\}$$

In order to obtain corrected interaction energies, one has to deal with a corrected supermolecular PES. The counterpoise corrected PES for the supermolecule can be defined as follows

$$E^{CP}(AB)=E_{AB}^{AB}(AB)+\delta_{AB}^{BSSE}=E_{AB}^{AB}(AB)+\left[E_{AB}^{A}(A)+E_{AB}^{B}(B)-E_{AB}^{AB}(A)-E_{AB}^{AB}(B)\right] \tag{8}$$

The equation above represents another point of view of the CP-correction. In our opinion, it should be more generally assigned to the supermolecule description, rather than to the interaction energy. Some authors have argued that BSSE is a pure interaction energy term; however, BSSE exists even thought we are not interested in interaction energy. In fact, $E_{AB}(AB)$ and $\Delta E_{AB}(AB)$ differ by a BSSE-free constant term which depends only on the system and the definition of the fragments. Therefore it seems quite coherent to assign the BSSE correction to $E_{AB}(AB)$.

Eq. (8) can be easily generalised to Nth-order energy derivatives.

$$\frac{\partial^n E^{CP}(AB)}{\partial R^n} = \frac{\partial^n E_{AB}^{AB}(AB)}{\partial R^n} + \frac{\partial^n \delta_{AB}^{BSSE}}{\partial R^n} = \\ = \frac{\partial^n E_{AB}^{AB}(AB)}{\partial R^n} + \left[\frac{\partial^n E_{AB}^{A}(A)}{\partial R^n} + \frac{\partial^n E_{AB}^{B}(B)}{\partial R^n} - \frac{\partial^n E_{AB}^{AB}(A)}{\partial R^n} - \frac{\partial^n E_{AB}^{AB}(B)}{\partial R^n} \right] \tag{9}$$

Second and third derivatives to be used for both harmonic and anharmonic vibrational frequency analysis are expressed as linear combinations of the contributions of each term. Any property defined as a derivative of the energy can be corrected for the BSSE by means of the counterpoise method. Unfortunately, there is no definition for a CP-corrected electronic density for the supermolecule.

I.1.3.1 The role of the fragment relaxation term

In order to calculate the CP-corrected stabilization energy neglecting the relaxation of the fragments geometry, one can take simply

$$\Delta E_{no-rel}^{CP}(AB) = E_{AB}^{AB}(AB) - E_{AB}^{AB}(A) - E_{AB}^{AB}(B) \tag{10}$$

that is, the CP-corrected interaction energy.

In principle, from Eq. (2), the fragment relaxation contribution is additive to the interaction energy. Indeed, it is additive only in the sense that it switches from interaction energies to stabilization energies at any point of the complex PES. However, the geometries corresponding to the stationary points of both quantities can be very different, so that the fragment relaxation should not be considered as an additive term. The only exception is for single-point calculation at frozen complex geometry, like most of the Intermolecular Perturbation Theory analysis in the literature.

It is a common practice in the intermolecular interaction calculations to freeze the fragment geometries and optimize complex energy for the rest of intermolecular parameters. A test calculation of a full optimization is used to assess the magnitude of the relaxation term. In many cases, it is claimed that the validity of Eq. (3) can be

acceptable if the relaxation contribution is negligible or smaller than the desired accuracy.

As shown in Eq. (4), its value depends both on the supermolecule parameters and those of the isolated fragments. However, derivative of the fragment relaxation term

$$\frac{\partial\left(\Delta E_{rel}(AB)\right)}{\partial R_i}=\frac{\partial\left(E_{AB}^{A}(A)+E_{AB}^{B}(B)-E_{A}^{A}(A)-E_{B}^{B}(B)\right)}{\partial R_i}= \\ =\frac{\partial\left(E_{AB}^{A}(A)\right)}{\partial R_i}+\frac{\partial\left(E_{AB}^{B}(B)\right)}{\partial R_i} \qquad \forall R_i \in \{A,B\} \tag{11}$$

only depends on the intramolecular parameters. Indeed, derivatives with respect to $E_A^A(A)$ and $E_B^B(B)$ vanish by definition. Derivatives involving intermolecular parameters also vanish because both fragment contributions are calculated using only their own basis set. Therefore, the value of the fragment relaxation term does not depend on the variation of the intermolecular parameters.

Neglecting the fragment relaxation is particularly dangerous when considering transition states. In many cases, the geometries of the fragments in the supermolecule transition state structures are much more stretched, leading to a large fragment relaxation contribution. This is particularly evident for chemical reactions involving intermolecular complexes, like the concerted hydrogen exchange. In case of internal rotational barriers, the difference between the fragment relaxation contributions of the minimum and transition state is usually very small. Nevertheless, in Section II.1.2 it will be shown that these differences are not additive to the total barrier for internal rotation, mostly when the fragment relaxation is clearly nonadditive to the interaction energies.

However, when dealing with both the BSSE and the fragment relaxation another problem arise. One must take into account that, if we stick to the old definition of the CP correction, where it is assigned to the interaction energy, we will not be able to obtain stationary points for the CP-corrected stabilization energy. Indeed, the optimization of the CP-corrected interaction energy is only possible taking into account the variation of the intermolecular parameters

$$\frac{\partial\left(\Delta E^{CP}_{no-rel}(AB)\right)}{\partial R_i}=\frac{\partial\left(E^{AB}_{AB}(AB)\right)}{\partial R_i}-\frac{\partial\left(E^{AB}_{AB}(A)\right)}{\partial R_i}-\frac{\partial\left(E^{AB}_{AB}(B)\right)}{\partial R_i}\qquad \forall R_i\in\{AB'\}\qquad (12)$$

{AB'} represent the variables corresponding to the intermolecular parameters of the complex. For a full CP-corrected optimization, the CP correction must be formally assigned to the complex's energy and hence the fragment relaxation term must be included.

The value of the fragment relaxation contribution is always determined at the uncorrected supermolecule geometry, which is indeed different when corrected for BSSE by the counterpoise method. One must assume that the relaxation contribution at the geometry of both the corrected and uncorrected supermolecular stationary points is of the same order of magnitude. Hence, even though it is a quantity free of intermolecular BSSE, the fragment relaxation term can indirectly affect the CP-correction value if the geometries of the uncorrected and CP-corrected stationary points are very different. In this way, one could *a priori* predict a given fragment relaxation contribution based on that of the uncorrected complex stationary point, and then obtain a totally different value at the CP-corrected geometry.

It will be shown in this work that the effect of the BSSE on the intramolecular parameters is negligible compared to the modification of the intermolecular distance and bond angles. Hence, the inclusion of the CP-correction doesn't change the aforementioned problematic when the fragment relaxation is neglected.

The most important thing to bear in mind is that the monomer relaxation term is additive only when switching from interaction energies to stabilization energies at any point of the complex PES. As mentioned above, the CP-correction is indirectly affected upon inclusion of the monomer relaxation due to the change in the location of the stationary point. However, this change is by far less important than the induced change on the total energy of the complex. In principle, total energies are not of much interest. However, when calculating energetic barriers, i. e., differences between states, the effects of neglecting the monomer relaxation are much more important than merely the monomer relaxation energy. In other words, the monomer relaxation contribution is not additive at all neither for energy differences between total energies nor for interaction energies.

I.1.4 The Chemical Hamiltonian Approach

In 1983 Mayer introduced the so-called Chemical Hamiltonian Approach (CHA)[49]. In this method, the terms responsible for the BSSE are eliminated from the Hamiltonian, yet keeping all true interaction delocalizations with physical meaning. This method, however, is much more than a scheme to eliminate the BSSE in molecular complexes as, indeed, it reveals the existence of an *intramolecular* basis set superposition error.

The Hamiltonian operator, within the Born-Oppenheimer framework, contains only one and two-particle operators, so that the three and four-center integrals that appear as a consequence of the MO-LCAO approximation have no physical correspondence. The CHA, in its original formulation, is intended to be a partitioning scheme of the SCF Hamiltonian in different atomic and interatomic contributions expressed in terms of the one and two-center integrals. One of these contributions is assigned to basis set extensions due to the incompleteness of the atomic basis sets.

This methodology can be easily applied to the BSSE problem of the intermolecular interactions simply by defining molecular fragments, instead of the atomic ones. In this case, the Hamiltonian can be partitioned intro an (effective) intramolecular, pure intermolecular and intermolecular basis set extension terms. This particular case has been the most widely developed, in terms of using the *physical* Hamiltonian, i.e., without the basis set extension term, to derive SCF-like equations (CHA-SCF[68]) and further introduction of electron correlation through full-CI[69] (CHA-CI) or MP2[65] (CHA-MP2) CHA-adapted methodologies.

At the SCF and DFT level of theory a simpler way to eliminate the BSSE was proposed[70] and called CHA/F and CHA/DFT, respectively. Instead of acting on the one and two-electron integrals, that is on the Hamiltonian, the essence of the CHA method is applied at each step of a conventional SCF algorithm to the conventional

---

[68] I. Mayer, Á. Vibók, Chem. Phys. Lett. 136, 115 (1987)

[69] I. Mayer, Á. Vibók, P. Valiron, Chem. Phys. Lett. 223, 166 (1994)

[70] I. Mayer, Á. Vibók, Int. J. Quantum Chem. 40, 694 (1991); Á. Vibók, I. Mayer, *ibid.* 43, 801 (1992)

Fock matrix. In this case, the same HF solution than that of the isolated fragments is ensured for a hypothetical ghost-orbital fragment calculation. The implementation of this algorithm is much simpler than the CHA-SCF one and can be extended to DFT calculations based on the Kohn-Sham equations.

As for the accuracy, both the CHA/F and CHA-SCF yield almost equivalent results. However, only the CHA-SCF canonic orbitals (both occupied and virtuals) can be safely used for further inclusion of electron correlation[50] energy.

In this work, the CHA/F and CHA/DFT versions have been used for the study of several complexes. On the other hand, the CHA-SCF equations for an arbitrary number of fragments have been coded using a novel algorithm for both the restricted and unrestricted versions. Further extension to the CHA-MP2 and CHA-UMP2 methodologies, respectively, is in progress and no results will be presented here.

In the following section, a deeper insight on the CHA philosophy, problems and implementation will be given.

I.1.4.1 CHA-SCF

Let us consider a molecular aggregate composed of N interacting fragments. In the case of the isolated fragments, one deals with *intramolecular* operators (with respect to the molecular aggregate) acting on *intramolecular* basis functions, i. e. the MCBS of the given fragment. The resulting function will have also an intramolecular character because the final elements or integrals will be expanded in the Hilbert subspace spawned by the respective MCBS. However, in the description of a given fragment *within* the complex, i. e., when the whole basis functions of the complex are available, the result of the action of an intramolecular operator on the same intramolecular basis functions can be freely expanded in the whole basis set. These basis set extensions have a purely BSSE character.

The basic idea is to analyze how these undesirable expansions can be avoided in order to ensure that the resulting wavefunction is the same than that of the MCBS calculation. And then, when doing the calculation of the molecular aggregate, the same changes must be applied on the *intramolecular* part of the Hamiltonian.

Therefore, let us define h^A as the monoelectronic term of the Hamiltonian

associated to the fragment A of the complex,

$$\hat{h}^A = -\tfrac{1}{2}\sum_{i=1}^{Na}\nabla_i^2 - \sum_{i=1}^{Na}\sum_{\alpha\in A}\frac{Z_\alpha}{r_{i\alpha}} \tag{13}$$

the total Hamiltonian being

$$\hat{H}^A = -\frac{1}{2}\sum_{i}^{N_A}\nabla_i^2 - \sum_{i}^{N_A}\sum_{\alpha}^{M_A}\frac{Z_\alpha}{r_{i\alpha}} + \sum_{i\neq i}^{N_A}\frac{1}{r_{ij}} + \sum_{\alpha\neq\beta}^{M_A}\frac{Z_\alpha Z_\beta}{R_\alpha} \tag{14}$$

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

The action of this operator, of intramolecular nature with respect to the molecular complex Hamiltonian, on a basis function assigned to this fragment can be formally expressed as

$$\hat{h}^A\varphi^A \equiv \hat{P}^A\hat{h}^A\varphi^A + \left(1-\hat{P}^A\right)\hat{h}^A\varphi^A \tag{15}$$

being

$$\hat{P}^A = \sum_{i,j\in A}\chi_i^A\, S^{-1}_{(A)ij}\chi_j^A \tag{16}$$

the projector associated to the subspace spawned by the atomic orbitals of the fragment A. The inverse matrix of the metric, S^{-1}, must be included for the general case of non-orthogonal basis sets.

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. However, the most important point is that, in the case of the isolated fragment calculation (MCBS) or the exact solution one has trivially

$$\hat{h}^A\varphi^A = P^A\hat{h}^A\varphi^A \tag{17}$$

Whereas this does not hold when the fragment dispose of the whole molecular basis set or atomic orbitals. Thus, the term $\left(1-\hat{P}^A\right)\hat{h}^A\varphi^A$, the projection on the orthogonal complement of the A^{th} atomic orbital's subspace of the result of applying the intramolecular operator over a atomic orbitals of the same subspace, is the responsible for the BSSE. In order to obtain a molecular complex description consistent with that of the composing fragments this term must be dropped from the Hamiltonian of the molecular aggregate. That means one must perform the following substitutions

$$\begin{gathered}\hat{h}^A\varphi^A \Rightarrow \hat{P}^A\hat{h}^A\varphi^A \\ \frac{1}{r_{12}}\varphi_i^A(1)\varphi_j^A(2) \Rightarrow \hat{P}^A(1)\hat{P}^A(2)\,\frac{1}{r_{12}}\varphi_i^A(1)\varphi_j^A(2)\end{gathered} \tag{18}$$

for the mono- and bielectronic term, respectively. Note that in the later case a projector for each electron is needed.

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.

$$\begin{gathered}\left\langle \chi_i \middle| \hat{h}^A \middle| \chi_j \right\rangle \Rightarrow \sum_{\mu,\nu\in A} S_{i\mu}\, S^{-1}_{(A)\mu\upsilon} \left\langle \chi_\nu \middle| \hat{h}^A \middle| \chi_j \right\rangle = \left\{\chi_i \mid \hat{h}^A \mid \chi_j\right\} \qquad \forall \mathrm{j} \in \mathrm{A},\ \forall \mathrm{i} \\ \left[\chi_i\chi_j \mid \chi_k\chi_l\right] \Rightarrow \sum_{\mu,\nu,\sigma,\rho\in A} S_{i\mu}\, S_{j\nu} S^{-1}_{(A)\mu\sigma} S^{-1}_{(A)\upsilon\rho} \left[\chi_\sigma\chi_\rho \mid \chi_k\chi_l\right] = \left\{\chi_i\chi_j \mid \chi_k\chi_l\right\} \\ \forall \mathrm{k,l} \in \mathrm{A},\ \forall \mathrm{i,j}\end{gathered} \tag{19}$$

S_{ij} corresponds to the elements of the metric matrix of the molecular aggregate basis set. Note that not all the integrals are transformed; only when the atomic orbital of the right–hand side of the integral belongs to the A^{th} fragment basis set. The projectors act on the right-hand side of the integrals, so that the CHA-transformed integrals do not satisfy the usual symmetry. In fact, they have the following properties

$$\begin{aligned} &\{\chi_i \mid \hat{h}^A \mid \chi_j\} = \{\chi_j \mid \hat{h}^A \mid \chi_i\} = [\chi_i \mid \hat{h}^A \mid \chi_j] \quad \forall\{i, j \in A\} \vee \forall\{i, j \notin A\} \\ &\{\chi_i \mid \hat{h}^A \mid \chi_j\} \neq \{\chi_j \mid \hat{h}^A \mid \chi_i\} \qquad \forall\{i \in A, j \notin A\} \vee \forall\{i \notin A, j \in A\} \end{aligned} \tag{20}$$

In case of the two-electron integrals it reduces to

$$\begin{aligned} \{\chi_i\chi_j \mid \chi_k\chi_l\} &= [\chi_i\chi_j \mid \chi_k\chi_l] \quad \forall\{i, j, k, l \in A\} \vee \forall\{i, j, k, l \notin A\} \\ \{\chi_i\chi_j \mid \chi_k\chi_l\} &= \{\chi_j\chi_i \mid \chi_l\chi_k\} \neq \{\chi_k\chi_l \mid \chi_i\chi_j\} \quad \forall\{k, l \in A; i, j \notin A\} \vee \forall\{i, j \in A; k, l \notin A\} \end{aligned} \tag{21}$$

In short, if all the atomic orbitals involved in the integral belong to the same subset, the original integral remains unchanged. If both belong to the same fragment that the corresponding operator, the integral has a pure intramolecular character whereas if both belong to a different fragment, the integral is purely intermolecular. Also, the two-electron integrals remain unchanged with respect to the interchange of the electrons. Unfortunatelly, neither the one- nor the two-electron integrals are symmetric, which means that the transformed intramolecular Hamiltonian is not hermitian. As a consequence, the whole Hamiltonian of the A^{th} fragment in the whole basis set complex will not be hermitian either. This will introduce some difficulties, both technical and conceptual, when dealing with the CHA Hamiltonian of the complex.

The previous analysis showed how obtain the description of a fragment in the whole basis set consistent with that of the fragment with its own basis set. The question now is the following: *how the molecular complex Hamiltonian can be partitioned into intra and intermolecular components?*

It is not a trivial task, indeed. For instance, if one tries to separate the kinetic energy into the contributions of the N fragments, one must distribute the electrons and assign each of them to a fragment. For instance, in the simplest case of two fragments A and B one has

$$\sum_{i=1}^{N} -\tfrac{1}{2}\nabla_i^2 = \sum_{i=1}^{N_A} -\tfrac{1}{2}\nabla_i^2 + \sum_{i=1}^{N_B} -\tfrac{1}{2}\nabla_i^2 \tag{22}$$

whereas it is well-known that by the Pauli antisymmetry principle, the electrons must be indistinguishable and the wavefunction must be antisymmetric with respect to the interchange of two electrons.

There is, however, a formalism that does allow for such a partitioning, the so-called Second Quantization (SQ)[71]. The details of this formalism will not be discussed here. In the particular but usual case of using a finite, truncated, basis set, the Hamiltonian of a system can be expressed explicitly by means of the SQ formalism in terms of the so-called annihilation and creation operator and the one and two-electron integrals over the basis set. These operators act on the wavefunction in such a way that the electron described by the corresponding molecular orbital is added or removed. Therefore, the expression of the Hamiltonian is usually given with the corresponding integrals expressed in terms of the molecular orbitals.

However, our goal is to work with the atomic orbitals since they can be partitioned into different subsets, one for each fragment[72]. As mentioned before, these atomic orbitals do not form an orthogonal set, so that a variant of the classical SQ formalism that takes into account this overlap must be used. Basically, biorthogonal orbitals are used to define the creation and annihilation operator whereas the one and two-electron integrals are expressed on the AO basis set. A biorthogonal set $\{\phi_i\}$ is defined in such a way that

$$\begin{aligned} &\langle \phi_i \mid \phi_i \rangle = S_{ij}^{-1} \;;\quad \langle \phi_i \mid \chi_j \rangle = \delta_{ij} \\ &\langle \chi_i \mid \chi_j \rangle = S_{ij} \quad \textit{and} \quad S_{ij}^{-1} \equiv \left(S^{-1}\right)_{ij} \end{aligned} \tag{23}$$

where the $\{\chi_i\}$ correspond to the original, non-orthogonal set. The corresponding creation and annihilation operators are formally expressed in terms of the AO set and the biorthogonal set, respectively, so that the necessary anticonmmutation rule holds for the non-orthogonal case.

The interesting advantage of the SQ formalism is that the Hamiltonian does not

[71] P. R. Surján, Second Quantized Approach to Chemistry, Springer-Verlag Berlin Heidelberg (1989)

[72] In the case of using bond function they can be either assigned to a given fragment or they can be considered as a (dummy) fragment itself.

depend on the number of electrons of the system, but is determined uniquely by the elements of the one and two-electron terms Hamiltonian. This dependence is moved onto the wavefunction.

Having said that, the Hamiltonian of the complex in terms of this *mixed* formalism is expressed as

$$\hat{H} = \sum_{\mu,\nu,\lambda} S^{-1}_{\lambda\mu} \left\langle \chi_\mu \left| \hat{h} \right| \chi_\nu \right\rangle \hat{\lambda}^+ \hat{\tilde{\nu}}^- + \tfrac{1}{2} \sum_{\mu,\nu,\sigma,\rho,\lambda,\gamma} S^{-1}_{\lambda\mu} S^{-1}_{\gamma\nu} \left[\chi_\mu \chi_\nu \mid \chi_\rho \chi_\sigma \right] \hat{\lambda}^+ \hat{\gamma}^+ \hat{\tilde{\sigma}}^- \hat{\tilde{\rho}}^- + \sum_{\alpha \neq \beta}^{M} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} \qquad (24)$$

whereas for the A^{th} fragment description of the same basis set, i. e. including the ghost orbitals one has

$$\begin{aligned} \hat{H}^A = & \sum_{\mu,\nu,\lambda} S^{-1}_{\lambda\mu} \left\langle \chi_\mu \left| \hat{h}^A \right| \chi_\nu \right\rangle \hat{\lambda}^+ \hat{\tilde{\nu}}^- + \\ & + \tfrac{1}{2} \sum_{\sigma,\rho \in A} \sum_{\mu,\nu,\lambda,\gamma} S^{-1}_{\lambda\mu} S^{-1}_{\gamma\nu} \left[\chi_\mu \chi_\nu \mid \chi_\rho \chi_\sigma \right] \hat{\lambda}^+ \hat{\gamma}^+ \hat{\tilde{\sigma}}^- \hat{\tilde{\rho}}^- + \sum_{\alpha \neq \beta}^{M_A} \frac{Z_\alpha Z_{\beta\alpha}}{R_{\alpha\beta}} \end{aligned} \qquad (25)$$

where $\hat{h}$ and $\hat{h}^A$ correspond to the monoelectronic part of the total and A^{th} intramolecular Hamiltonian, respectively, as defined previously in Eq.(13).

Finally, the expression of the A^{th} fragment Hamiltonian expanded on its own basis set (MCBS) can be written, analogously, in the following way

$$\begin{aligned} \hat{H}^{A(eff)} = & \sum_{\mu,\nu,\lambda \in A} S^{-1}_{(A)\lambda\mu} \left\langle \chi_\mu \left| \hat{h}^A \right| \chi_\nu \right\rangle \hat{\lambda}^+ \hat{\tilde{\nu}}^- + \\ & + \tfrac{1}{2} \sum_{\mu,\nu,\sigma,\rho,\lambda,\gamma \in A} S^{-1}_{(A)\lambda\mu} S^{-1}_{(A)\gamma\nu} \left[\chi_\mu \chi_\nu \mid \chi_\rho \chi_\sigma \right] \hat{\lambda}^+ \hat{\gamma}^+ \hat{\tilde{\sigma}}^- \hat{\tilde{\rho}}^- + \sum_{\alpha \neq \beta}^{M_A} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} \end{aligned} \qquad (26)$$

The operator above is called *effective* intramolecular Hamiltonian, in the sense that it acts on a molecular function expanded in the whole basis set in the same manner the original Hamiltonian acts in the MCBS case.

All the integrals and creation operators ($\hat{\lambda}^+, \hat{\gamma}^+$) are expressed in terms of the atomic orbitals, whereas the annihilation ones ($\hat{\tilde{\sigma}}^-, \hat{\tilde{\rho}}^-$) are given by its biorthogonal set.

Now, we can easily express the Hamiltonian of the molecular complex as the sum of the intramolecular Hamiltonians associated to each fragment and the pure intermolecular interaction operator .

$$\hat{H} = \sum_{A=1}^{N} \hat{H}^{A} + \sum_{A<B}^{N} \hat{H}_{AB}^{\text{int}} \tag{27}$$

were the intermolecular interaction term can be easily obtained substracting Eq.(25) for each fragment from Eq.(24).

$$\begin{aligned} \hat{H}_{AB}^{\text{int}} = & -\sum_{\nu\in A}\sum_{\mu,\lambda} S_{\lambda\mu}^{-1} \left\langle \chi_\mu \left| \sum_{B\notin A} \frac{Z_B}{r_B} \right| \chi_\nu \right\rangle \hat{\lambda}^{+}\hat{\tilde{\nu}}^{-} - \\ & -\sum_{\nu\in B}\sum_{\mu,\lambda} S_{\lambda\mu}^{-1} \left\langle \chi_\mu \left| \sum_{A\notin B} \frac{Z_A}{r_A} \right| \chi_\nu \right\rangle \hat{\lambda}^{+}\hat{\tilde{\nu}}^{-} + \\ & +\sum_{\rho\in A}\sum_{\sigma\in B}\sum_{\mu,\nu,\lambda,\gamma} S_{\lambda\mu}^{-1} S_{\gamma\nu}^{-1} \left[\chi_\mu\chi_\nu \mid \chi_\rho\chi_\sigma\right] \hat{\lambda}^{+}\hat{\gamma}^{+}\hat{\tilde{\sigma}}^{-}\hat{\tilde{\rho}}^{-} + \sum_{\substack{\alpha\in A \\ \beta\in B}} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} \end{aligned} \tag{28}$$

It includes the interaction of the electrons of each fragment with the nuclei of the other fragments and the electron repulsion between electrons belonging to different fragments.

Now, if one substitutes the expressions of the intramolecular operators H^A by that of the effective intramolecular Hamiltonian $H^{A(eff)}$ for each fragment, and keeps the intermolecular interaction contribution unchanged, the CHA Hamiltonian is retrieved.

$$
\begin{aligned}
\hat{H}^{CHA} = \sum_{A}^{N} \Bigg\{ & \sum_{\mu,\nu,\lambda \in A} S^{-1}_{(A)\lambda\mu} \left\langle \chi_\mu \left| \hat{h}^A \right| \chi_\nu \right\rangle \hat{\lambda}^+ \hat{\bar{\nu}}^- + \\
& + \tfrac{1}{2} \sum_{\mu,\nu,\sigma,\rho,\lambda,\gamma \in A} S^{-1}_{(A)\lambda\mu} S^{-1}_{(A)\gamma\nu} \left[\chi_\mu \chi_\nu \mid \chi_\rho \chi_\sigma \right] \hat{\lambda}^+ \hat{\gamma}^+ \hat{\bar{\sigma}}^- \hat{\bar{\rho}}^- \Bigg\} - \\
& - \sum_{A \neq B}^{N} \Bigg\{ \sum_{\nu \in A} \sum_{\mu,\lambda} S^{-1}_{\lambda\mu} \left\langle \chi_\mu \left| \sum_{B \notin A} \frac{Z_B}{r_B} \right| \chi_\nu \right\rangle \hat{\lambda}^+ \hat{\bar{\nu}}^- - \\
& - \sum_{\nu \in B} \sum_{\mu,\lambda} S^{-1}_{\lambda\mu} \left\langle \chi_\mu \left| \sum_{A \notin B} \frac{Z_A}{r_A} \right| \chi_\nu \right\rangle \hat{\lambda}^+ \hat{\bar{\nu}}^- + \\
& + \sum_{\rho \in A} \sum_{\sigma \in B} \sum_{\mu,\nu,\lambda,\gamma} S^{-1}_{\lambda\mu} S^{-1}_{\gamma\nu} \left[\chi_\mu \chi_\nu \mid \chi_\rho \chi_\sigma \right] \hat{\lambda}^+ \hat{\gamma}^+ \hat{\bar{\sigma}}^- \hat{\bar{\rho}}^- \Bigg\} + \\
& + \sum_{A \neq B}^{N} \sum_{\substack{\alpha \in A \\ \beta \in B}}^{N_A, N_B} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}}
\end{aligned}
\tag{29}
$$

The CHA philosophy is applied directly to the Hamiltonian operator and hence to the corresponding one and two-center integrals. Therefore, no localization of the molecular orbitals o restriction on their expansion coefficients is introduced, as opposite to most of the other *a priori* BSSE-correction methods.

With this methodology, the total Hamiltonian of the molecular complex is BSSE-free, but it has the main drawback that the hermiticity is lost. There are several implications of this fact. The most important one is that the application of the variational principle in order to find the expression for the spin-orbitals of the CHA wavefunction is not at all trivial.

However, it was shown that a pseudo-secular equation analogous to the conventional SCF can be derived by applying the Brillouin theorem. Starting from the Schrödinger equation

$$
\hat{H}^{CHA} \mid \Psi > - E \mid \Psi >= 0 \tag{30}
$$

The trial wavefunction is approximated by a single-Slater determinant built up of orthonormalized molecular spin-orbitals ϕ_i , and the expectation value is

$$
E = \langle \Psi \mid \hat{H}^{CHA} \mid \Psi > \tag{31}
$$

Now, one must require that the difference in Eq.(30) has no component including any singly-excited configurations, according to the Brillouin theorem. In order to do that, the coefficient for all $|\Psi_1(\phi_i^{occ} \rightarrow \phi_p^{vir})\rangle$ terms in this difference are forced to vanish. The set of resulting equations[68] provide the appropriate optimization of the spin-orbitals that form the Slater determinant.

The final expressions closely resemble the conventional Hartree-Fock pseudo-secular equations, with the difference that the Fock matrix being diagonalized is non-hermitian. One has

$$F^{CHA} c_i = \varepsilon_i S c_i \tag{32}$$

where the modified Fock matrix can be written as the conventional one

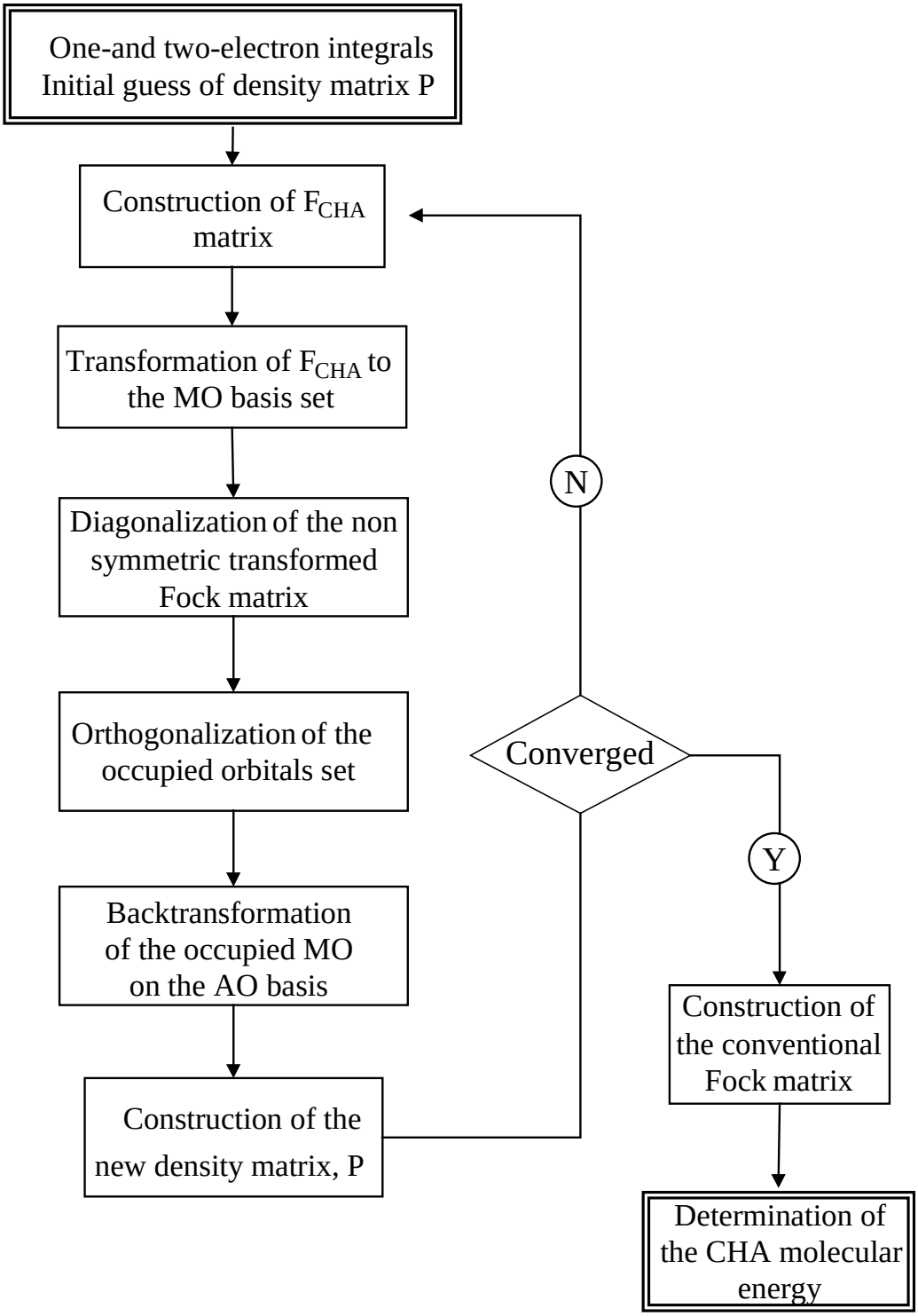
$$\begin{gathered} F_{\mu\nu}^{CHA} = \{\chi_\mu | \hat{h} | \chi_\nu\} + \sum_{\sigma,\rho} P_{\rho,\sigma} \left(2\{\chi_\mu\chi_\sigma | \chi_\nu\chi_\rho\} - \{\chi_\mu\chi_\sigma | \chi_\rho\chi_\nu\}\right) \\ P_{\mu\nu} = \sum_i^{occ} c_{\nu i} c_{\mu i}^* \end{gathered} \tag{33}$$

but using the CHA-modified one and two-electron integrals[65]

$$\begin{gathered} \{\chi_\nu | \hat{h} | \chi_\mu\} = \sum_{\sigma,\rho} S_{\upsilon\rho} (S_A)^{-1}_{\rho\sigma} \langle \chi_\sigma | \hat{h}^A | \chi_\mu \rangle - \sum_{\alpha \notin A} \langle \chi_\nu | \frac{Z^A}{r_A} | \chi_\mu \rangle \\ \forall \mu \in A \wedge \nu \notin A \\ \{\chi_\mu\chi_\sigma | \chi_\nu\chi_\rho\} = \sum_{\kappa,\lambda,\tau,\varepsilon} S_{\mu\varepsilon} (S_A)^{-1}_{\varepsilon\tau} S_{\sigma\kappa} (S_A)^{-1}_{\kappa\lambda} [\chi_\tau\chi_\lambda | \chi_\nu\chi_\rho] \\ \forall \nu, \rho \in A \wedge \mu, \sigma \notin A \end{gathered} \tag{34}$$

Since the CHA Fockian is not hermitian, its eigenvectors are not orthogonal and the virtual orbitals can overlap with the occupied ones. Also, both the eigenvectors and the eigenvalues may become complex.

The CHA energy and wavefunctions of a system can be obtained the following iterative self-consistent algorithm



SCHEME 2

Once the CHA Fockian is contructed, the algorithm above is analogous to that of the conventional SCF procedure except for two important details. First, since the Eq.(31) holds for Slater determinants build up from orthonormalized orbitals, the

occupied orbitals must be orthonormalized before they are used to construct the density-matrix for the next iteration. And second, note that when the self-consistency is achieved, the conventional Fockian must be built and used to compute the CHA energy.

However, the determination of the CHA Fockian at every iteration is not a trivial task. Also, the diagonalization of a non-symmetric Fock matrix needs for a robust method since the eigenvalues and eigenvectors may occur in complex pairs.

In Section II.2, these two problematic steps are examined with more detail. Particularly, an algorithm for the efficient construction of the CHA Fockian is presented. It will be shown that the computational cost of the described CHA modifications is very small. The CPU time required for each CHA-SCF cycle is almost the same than that of the conventional SCF. However, since the Fock matrix is non hermitian, the application of the conventional convergence acceleration techniques such as the DIIS[73] is not trivial. As a consequence, the total number of cycles necessary to achieve self-consistency is usually larger than for the conventional SCF.

I.1.4.2 The CHA/CE concept

One of the typical aspects of CHA is that the total energy is calculated as an expectation value of the full Hamiltonian over the CHA wave function (Ψ^{CHA}):

$$E^{CHA} = \frac{\left\langle \Psi^{CHA} \left| \hat{H} \right| \Psi^{CHA} \right\rangle}{\left\langle \Psi^{CHA} \left| \Psi^{CHA} \right\rangle \right.} \qquad (35)$$

This equation defines the BSSE-free CHA potential energy surface. One of the advantages of CHA over the counterpoise scheme is that in the former case the BSSE-free PES can be determined by a single energy calculation for every geometrical

[73] P. Pulay, Chem. Phys. Lett. 73, 393 (1980)

arrangement of the system. The same holds for the case of supermolecules with an arbitrary number of monomers.

It is rather surprising that the CHA Hamiltonian is only used to obtain the molecular orbitals and that the CHA energy is computed as the expectation value of the full Hamiltonian. This must be done in this way for two reasons: first, the preliminary numerical results computed with the CHA Hamiltonian were not satisfactory. The computed energy was far too small (in absolute value) than the uncorrected one and the BSSE-corrected by other methods, whereas if the energy was computed with the conventional Hamiltonian the results were much more satisfactory. And second, since the molecular orbital energies may be complex, the total energy computed with the CHA Hamiltonian could be complex as well, which is unacceptable from a physical point of view.

One has to mention again that during the elaboration of this memory, a new reformulation of the CHA has allowed to reproduce the CHA-SCF results by using a new CHA hermitic Hamiltonian. Hence, the non hermiticity of the current CHA-SCF procedure cannot be seen as formally incorrect but just a rather problematic methodology.

Indeed, one has to face some rather serious difficulties calculating derivatives of Eq. (35). Let us consider an SCF-type CHA method in the following. The first derivative of the CHA-SCF energy with respect to a geometrical parameter can be given as

$$\frac{dE^{CHA}}{dx} = \frac{\partial E^{CHA}}{\partial x} + \sum_i \frac{\partial E^{CHA}}{\partial C_i}\frac{\partial C_i}{\partial x} \qquad (36)$$

where $\frac{\partial E^{CHA}}{\partial x}$ represents how the CHA energy explicitly depends on the geometry parameter x and C_i are SCF coefficients. As we mentioned above the SCF parameters (wave function) are determined by using $\hat{H}^{CHA}$ but the energy is calculated from the complete Hamiltonian $\hat{H}$. This has the consequence that in a SCF-type CHA model the partial derivatives $\frac{\partial E^{CHA}}{\partial C_i}$ are not equal to zero. This fact is in contrast to the usual Hartree-Fock theory where

$$\frac{\partial E}{\partial C_i} = 0 \qquad (37)$$

because the SCF parameters are determined by using the variational principle. Hereby, to calculate gradients in the CHA framework, one must determine the gradient of the SCF coefficients (C_i) with respect to the geometry parameters (x). This step involves the solution of the CHA version of the coupled-perturbed Hartree-Fock (CPHF) equations[74]. If the partial derivatives $\frac{\partial C_i}{\partial x}$ are known, calculation of the CHA-SCF gradient is straightforward. However, the calculation of the derivatives of the SCF parameters is a rather time-consuming procedure. First, the CHA CPHF equations have to be solved for all the geometrical parameters. Second, the solution of the CHA CPHF equations does involve a full SCF-type procedure making the practical calculation of the CHA gradient insufficient from the computational point of view. Since the CHA models are not based on variational methods, it is obvious to apply techniques which make the calculation of gradients of non-variational correlation methods (MPn, Coupled Cluster) efficient. These techniques are based on the so-called Z-vector method[75] or algorithms that use fully variational correlation functionals instead of the traditional non-variational ones[76]. The application of such techniques is possible in the CHA framework, but the CHA-SCF gradient formulas based on the Z-vector or on fully variational methods seem still too complicate for practical computations at the SCF level of theory. Therefore, at present time, the CHA-SCF gradient has not been implemented and the location of stationary points on the CHA-SCF PES must be carried out by means of numerical first derivatives.

I.1.4.3 CHA/F and CHA/DFT

[74] B. Paizs and I. Mayer, *Chem. Phys. Lett.*, **220**, 97 (1994)

[75] N. C. Handy and H. F. Schaefer III, *J. Chem. Phys.*, **81**, 5031 (1984)

[76] T. Helgaker and P. Jorgensen, Theoret. Chim. Acta, **75**, 111 (1989)

The easiest way of implementing the CHA ideas to correct for the BSSE has been called CHA/F[77]. In this case, instead of modifying the total Hamiltonian by the use of the corresponding projectors, the conventional Fock matrix is first constructed and then properly modified in order to ensure the description of the fragments within the complex is consistent with that of the fragments with the MCBS. Since it does not act on the one- and two-electron integrals it can be easily generalized to the Kohn-Sham case, and hence the elimination of the BSSE can be reached also at the DFT level of theory (CHA/DFT[78]).

Let us start again with a molecular complex composed by N fragments. The Roothaan equations for the isolated, non interacting, A^{th} fragment in terms of the spin-orbitals can be written in this way

$$F_A^{AA} c_i^A = \varepsilon_i^A S^{AA} c_i^A \qquad \forall A \in N;\ \forall i \in n_A \tag{38}$$

where n_A is the number of molecular spin orbitals of fragment A and the Fock matrix is a diagonal block of the Fock matrix of the given fragment expanded in the whole basis set, i.e., using the ghost orbitals.

$$F_A = \begin{pmatrix} F_A^{11} & F_A^{12} & \dots & F_A^{1A} & \dots & F_A^{1N} \\ F_A^{21} & F_A^{22} & \dots & F_A^{2A} & \dots & F_A^{2N} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ F_A^{A1} & F_A^{A2} & \dots & F_A^{AA} & \dots & F_A^{AN} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ F_A^{N1} & F_A^{N2} & \dots & F_A^{NA} & \dots & F_A^{NN} \end{pmatrix} \tag{39}$$

This Fockian is obtained from the Hamiltonian associated to the A^{th} fragment, given in Eq. (13).

Therefore, for the case of the calculation of this fragment using the whole basis set one has

---

[77] B. Paizs, S. Suhai, J. Comput. Chem 18, 694 (1997); G. Halász, Á. Vibók, P. Valiron, I. Mayer 6332, 100 (1996)

[78] M. Kieninger, S. Suhai, I. Mayer, Chem. Phys. Lett., 485, 230 (1994)

$$F_A c_i^{A*} = \varepsilon_i^{A*} S c_i^{A*} \qquad \forall \mathrm{A} \in \mathrm{N};\ \forall \mathrm{i} \in \mathrm{n} \tag{40}$$

In this calculation, n molecular orbitals and corresponding energies are obtained, the first n_A being, in general, different to that of the obtained for the isolated fragment (Eq. (38)), both occupied and virtuals, due to purely BSSE effects.

The goal now is to determine how F_A must be modified in order to obtain the same orbitals and energies derived from F_A^{AA}, that is

$$c_i^{A*} = c_i^A \qquad ; \qquad \varepsilon_i^{A*} = \varepsilon_i^A \tag{41}$$

On can formally write the following equation for each fragment

$$F_A d_i^A = \varepsilon_i^A S d_i^A \tag{42}$$

where each pseudo-eigenvector has the following structure

$$d_i^A = \begin{pmatrix} 0 \\ 0 \\ \cdots \\ c_i^A \\ \cdots \\ 0 \end{pmatrix} \tag{43}$$

Expanding Eq. (42) in terms of the block structure given in Eq. (39), for both the Fockian and the overlap matrix, the following matrix equations are derived

$$\begin{aligned} F_A^{1A} c_i^A &= \varepsilon_i^A S^{1A} c_i^A \\ F_A^{2A} c_i^A &= \varepsilon_i^A S^{2A} c_i^A \\ &\cdots \\ \boldsymbol{F_A^{AA} c_i^A} &= \boldsymbol{\varepsilon_i^A S^{AA} c_i^A} \\ F_A^{NA} c_i^A &= \varepsilon_i^A S^{NA} c_i^A \end{aligned} \tag{44}$$

The equation marked in bold is just the HF equation associated to the isolated fragment A using the MCBS, i.e that of Eq. (38). It is easy to see that the other N-1 equations do not hold unless

$$F_A^{BA} = S^{BA}(S^{AA})^{-1}F_A^{AA} \qquad \forall \mathrm{B} \neq \mathrm{A} \tag{45}$$

The same considerations hold for the N-1 fragments composing the molecular complex. In general, in order that the description of a given fragment X within the complex is not affected by the basis sets of the other fragments (ghost orbitals), the following substitution must be performed for each block

$$F_X^{YX} \Leftarrow S^{YX}(S^{XX})^{-1}F_X^{XX} \;;\; \forall \mathrm{Y} \in \mathrm{N} \tag{46}$$

For an efficient implementation of the algorithm, the CHA Fockian can be expressed as the sum of the conventional Fockian plus a correction matrix

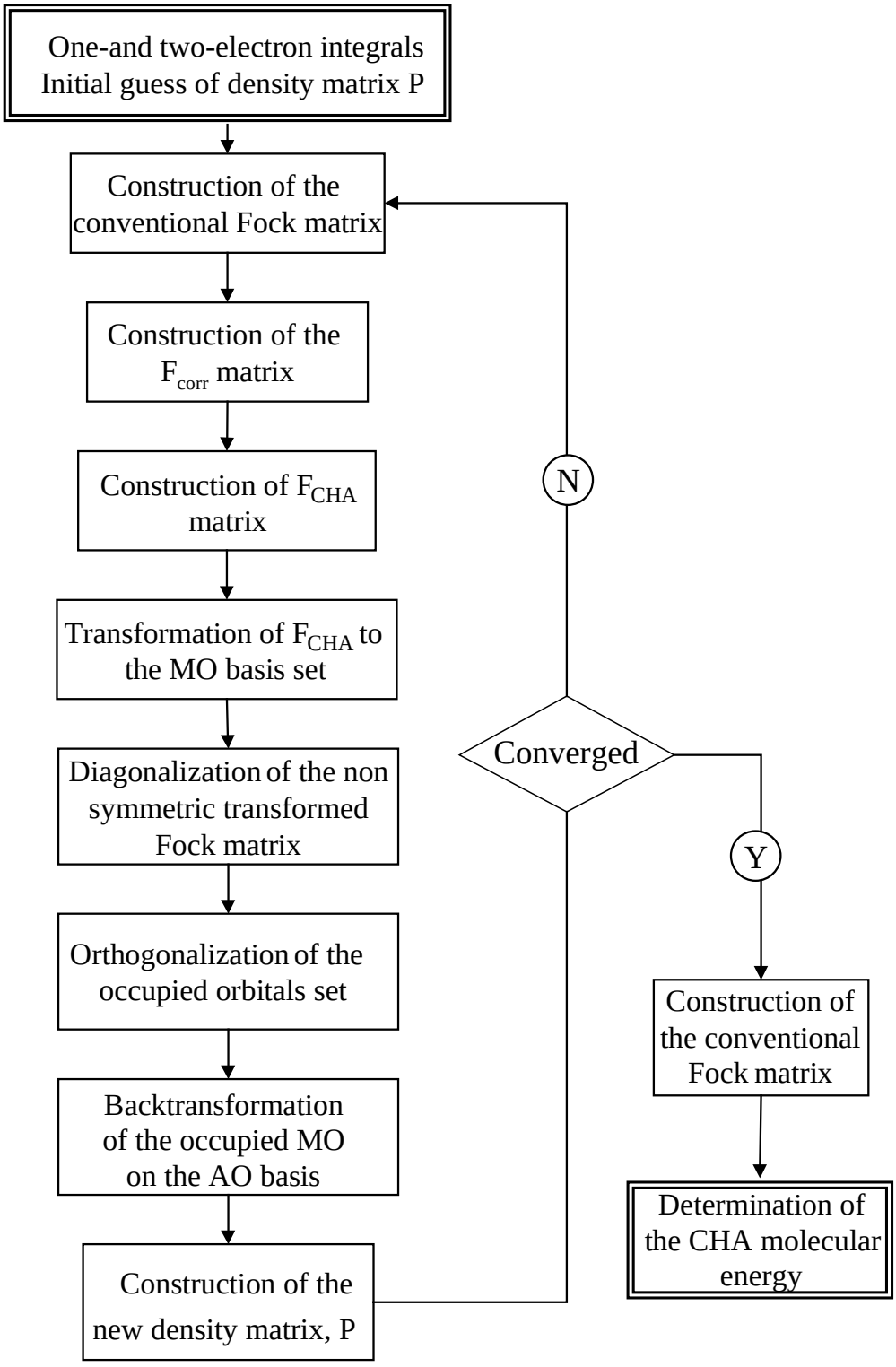
$$F_{CHA/F} = F + F_{corr} \tag{47}$$

where

$$F_{corr} = \begin{pmatrix} 0 & S^{12}S_2^{-1}F_2^{22} - F_2^{12} & \ldots & S^{1N}S_N^{-1}F_N^{NN} - F_N^{1N} \\ S^{21}S_1^{-1}F_1^{11} - F_1^{21} & 0 & \ldots & S^{2N}S_N^{-1}F_N^{NN} - F_N^{2N} \\ \ldots & \ldots & \ldots & \ldots \\ S^{N1}S_1^{-1}F_1^{11} - F_1^{N1} & S^{N2}S_2^{-1}F_2^{22} - F_2^{N2} & \ldots & 0 \end{pmatrix} \tag{48}$$

Once the transformed Fock matrix is obtained, the same strategy than that of the CHA-SCF case can be used now to obtain the CHA/F orbitals, and finally the wavefunction. Moreover, as mentioned before, since the modifications are not carried out on the integrals, the determination of the CHA Fockian is simplified.

The CHA/F procedure for each cycle of the self-consistent algorithm can be summarized in the following way:



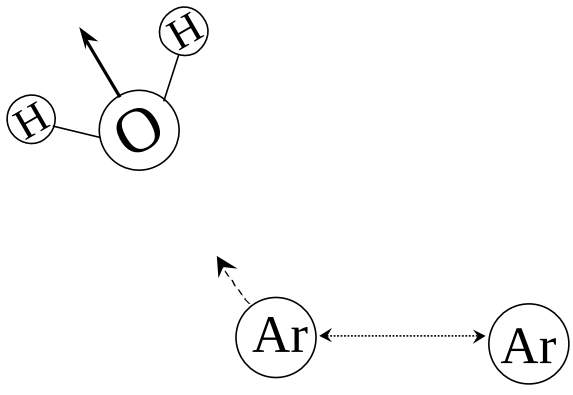
SCHEME 3

Note that, as in the CHA-SCF case, the CHA energy is computed as the expectation value of the conventional Hamiltonian over the CHA wavefunction. This ensures the molecular energy to be real.

I.1.5 Molecular Aggregates

The properties of the macroscopic bulk mater are entirely determined by the intermolecular forces acting at the molecular scale. These interactions play an essential role in the packing of the molecular crystals or the structure of large macromolecules like DNA. These effects appear because of the nonadditivity of the forces in the molecular aggregates. Indeed, the intermolecular potential of a cluster cannot be expressed as the sum over of all the pairwise interactions, because the nearest molecules not only interact with the considered pair, but also modify the way they interact with each other.

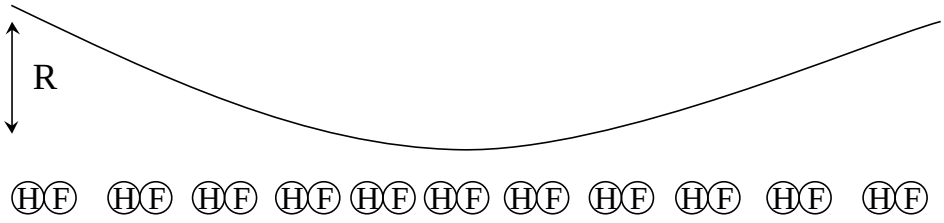
This can be easily understood in the example of the interaction of two rare gas atoms and a polar molecule like water. The two atoms interact with each other with dispersion forces due to the mutual induced polarization of the charge density clouds, that is, induced dipole-induced dipole interaction. When a water molecule approaches one of the rare gas atoms, its permanent dipole moment polarizes the atom and allows for a permanent dipole-induced dipole interaction. As a consequence, the interaction between the two rare gas atoms is enhanced due to the polarization of the charge distribution of one of them.



SCHEME 4

Obviously a two-body intermolecular potential will not be able to describe the actual effects between the three interacting systems, as the three-body contribution is not negligible at all.

The many-body contributions are also responsible for the so-called cooperative effects. Let us take for instance the case of the hydrogen fluoride. Rotational-vibrational spectroscopy techniques have determined for the dimeric species an intermolecular distance of 2.72Å in gas-phase experiments[79], whereas neutron diffraction experiments[80] have revealed that in the crystalline solid state structure this distance is shortened by ca. 0.3Å. Therefore the presence of many hydrogen fluoride monomers contributes to the mutual polarization of the neighbor molecules and the interaction is greatly enhanced.



SCHEME 5

The *ab initio* calculations have helped to the development of accurate interaction potentials, which are the starting point and the basis of both the molecular mechanics and the molecular dynamics simulations. Not only is the use of accurate pairwise interaction potentials very important but, in order to properly describe the anisotropy of the interactions, taking into account explicitly the many body effects is also necessary in order to obtain results comparable to the experiment.

It has been shown that the BSSE is one of the main sources of error in the calculations of intermolecular interactions. In the case of molecular aggregates, in analogy to the cooperative effects enhancing the stabilization energy, the magnitude of the BSSE could also be affected in a non-additive fashion. However, the removal of the BSSE in molecular complexes composed by more than two fragments has not been discussed much in the literature.

A few years ago, Turi and Dannenberg[81] pointed out the ambiguity of the counterpoise correction when studying growing chains of hydrogen fluoride. They

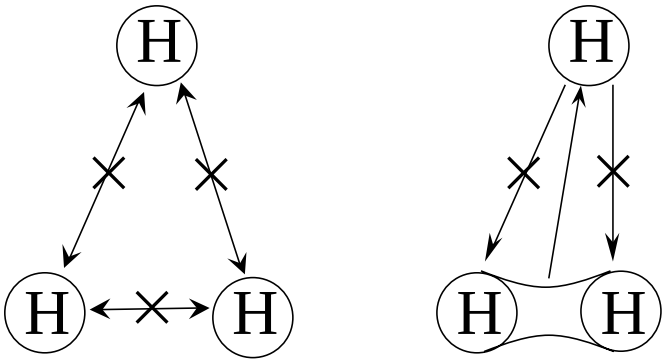
[79] T. R. Dyke, B. J. Howard, W. Kemplerer, J. Chem. Phys. 56, 2442 (1972)

[80] M. W. Johnson, E. Sandor, E. Arzi, Acta Cryst. B31, 1998 (1975)

[81] L. Turi and J. J. Dannenberg, J. Phys. Chem. 97 2488 (1993)

showed that the BSSE computed for the insertion of a new HF monomer to the $(HF)_n$ aggregate depends upon whether the incoming monomer is added to the H or to the F end of the aggregate. Hence, one can obtain different stabilization energies for the same chemical process, which is unacceptable. They proposed the use of the counterpoise method by defining as many fragments as monomeric subunits present in the molecular aggregate, the BSSE being the differences between the energies of each monomer with their own basis set and the whole molecular aggregate. This is just a trivial generalization of the earlier counterpoise method.

This method clearly solves the problem of the ambiguity of the CP correction, but is unable to explain all the effects of the incoming monomer on the interaction (and BSSE) *already* present in the molecular aggregate. Valiron and Mayer[82] illustrated this deficiency with the example of three interacting H atoms described by a Schrödinger 1s orbital. In this particular case, the counterpoise scheme above will not predict any BSSE in the system whereas all the H-H + H interactions bear some BSSE. Hence, there is a *second-order* BSSE due to the basis set extensions of all the H dimer descriptions within in the H trimer.



SCHEME 6

Indeed, these diatomic basis set extensions are as natural as the atomic ones. The BSSE is due to the improvement of the description of the atoms (fragments) within the complex by using the other fragment basis sets to expand the genuine atomic (single fragment) contributions to the Hamiltonian. Analogously, the genuine

diatomic (fragment pair) descriptions, including the respective interaction contribution within the atom (fragment) pair, are also artificially improved due to their expansion in the whole complex basis set (this is the particular case of the basis set extensions present on the H-H + H interaction commented above). In this sense, the hierarchical partition of an aggregate into atomic (single fragment), diatomic (fragment pairs), etc, arises naturally.

One way to take into account those high-order BSSE effects within the counterpoise framework was first introduced by White and Davidson[83] and later generalized by Valiron and Mayer[82]. They proposed a hierarchical counterpoise scheme for N-body cluster that treats the basis set extension effects of the monomers but also all the dimers, trimers, and so on, present in the aggregate.

In a recent paper, Mierzwicki and Latajka[84] analyzed the behavior of these two counterpoise methods in the calculation of many-body interactions of $Li(NH_4)_n$ and $Li(NH_4)_n^+$ clusters at several levels of theory. They also used another rather unusual scheme introduced by Wells and Wilson[85] where the counterpoise correction is carried out over pairs of fragments.

Therefore, there is not a unique way of correcting the BSSE, even within the counterpoise framework. In this section, the three counterpoise schemes mentioned above will be derived from the well known many-body expansion of the interaction energy.

Let us consider first a dimeric complex AB. The energy of the system at a given geometry[86] can be expressed simply as

$$E_{AB} = E_A + E_B + \Delta E_{AB} \tag{49}$$

[82] P. Valiron and I. Mayer, Chem. Phys. Lett. 275 46 (1997)

[83] J. C. White, and E. R. Davidson, J. Chem. Phys. 93 8029 (1990)

[84] K. Mierzwicki and Z. Latajka, Chem. Phys. Lett. 325 465 (2000)

[85] B. H. Wells, and S. Wilson, Chem. Phys. Lett. 101 429 (1983)

[86] Since we are interested in the supermolecule energy there is no need to introduce now the relaxation energy term. Once the counterpoise-corrected energy of the complex is obtained both the interaction and stabilization (including monomer relaxation) energies can be easily calculated.

where ΔE_{AB} represents the two-body interaction energy. According to the counterpoise philosophy, this value must be computed using the same basis set for all the terms involved

$$E_{AB}^{CP} = E_A + E_B + E_{AB} - E_A^{AB} - E_B^{AB} = E_{AB} + \left(E_A - E_A^{AB} + E_B - E_B^{AB}\right) \qquad (50)$$

where the superscript AB means the whole complex basis set is used for the calculation (if no superscript is used, it is assumed that the energy is computed with the subscript's basis set).

In this way, the counterpoise-corrected complex energy is recovered. Note that the one-body interaction energies, i.e., the fragment energies, are computed with the MCBS, whereas only the interaction energy term is computed with the DCBS. It is very important to point out that this is conceptually similar to the case of the *a priori* methods, like the CHA, where the diagonal (fragment-only) blocks of the Hamiltonian are maintained, and the BSSE-correction takes place only in the off-diagonal blocks (intermolecular interaction).

When the molecular complex is composed by three interacting units, ABC, the energy of the system can be expressed as the sum of one-, two- and three-body interaction energies,

$$E_{ABC} = E_A + E_B + E_C + \Delta E_{AB} + \Delta E_{AC} + \Delta E_{BC} + \Delta E_{ABC} \qquad (51)$$

the later due to the non-additivity of the interaction.

Similarly, in order to obtain a counterpoise-corrected energy of the trimer, the three-body energy term must be computed following the counterpoise receipt.

$$\Delta E_{ABC} = E_{ABC} - E_{AB}^{ABC} - E_{BC}^{ABC} - E_{AC}^{ABC} + E_A^{ABC} + E_B^{ABC} + E_C^{ABC} \qquad (52)$$

The point now is to determine which two-body interaction energies must be used. If no counterpoise-correction is taken into account at all for those terms, the following expression is obtained by substituting Eq (49) and (52) into Eq (51)

$$\begin{aligned} E_{ABC}^{CP} &= E_A + E_B + E_C + \left(E_{AB} - E_A - E_B\right) + \left(E_{BC} - E_B - E_C\right) + \left(E_{AC} - E_A - E_C\right) + \\ &+ E_{ABC} - E_{AB}^{ABC} - E_{BC}^{ABC} - E_{AC}^{ABC} + E_A^{ABC} + E_B^{ABC} + E_C^{ABC} = \\ &= E_{ABC} + \left(E_{AB} - E_{AB}^{ABC}\right) + \left(E_{BC} - E_{BC}^{ABC}\right) + \left(E_{AC} - E_{AC}^{ABC}\right) - \\ &- \left(E_A - E_A^{ABC}\right) - \left(E_B - E_B^{ABC}\right) - \left(E_C - E_C^{ABC}\right) \end{aligned} \tag{53}$$

There are three counterpoise terms related to basis set extension for all the dimers and monomers, respectively, the latter of different sign. The application of this scheme, however, yields meaningless results because the monomer basis set extensions are larger than the dimer ones and hence the BSSE is negative. In other words, the description of the complex is improved upon counterpoise correction, which is unacceptable.

On the other hand, one can consider using counterpoise-corrected two-body interaction terms in Eq. (51) but using the whole trimer's basis set, ABC.

$$\begin{aligned} E_{ABC}^{CP} &= E_A + E_B + E_C + \left(E_{AB}^{ABC} - E_A^{ABC} - E_B^{ABC}\right) + \left(E_{BC}^{ABC} - E_B^{ABC} - E_C^{ABC}\right) + \\ &+ \left(E_{AC}^{ABC} - E_A^{ABC} - E_C^{ABC}\right) + E_{ABC} - E_{AB}^{ABC} - E_{BC}^{ABC} - E_{AC}^{ABC} + E_A^{ABC} + E_B^{ABC} + E_C^{ABC} = \\ &= E_{ABC} + \left(E_A - E_A^{ABC}\right) + \left(E_B - E_B^{ABC}\right) + \left(E_C - E_C^{ABC}\right) \end{aligned} \tag{54}$$

In this case the conventional counterpoise scheme is obtained, which includes only the basis set extensions of the monomers in the whole basis set. Wells and Wilson[85] called this approach site-site function counterpoise.

However, the same considerations than that of the dimer case imply that the two-body interaction energy terms must be described with the respective DCBS basis set. That is, the exclusively dimer energy contributions must be expanded in its own DCBS, the same as the one-body monomer's contributions are expressed in the MCBS. According to these considerations, the counterpoise-corrected trimer energy will have the following expression

$$\begin{aligned} E_{ABC}^{CP} &= E_A + E_B + E_C + \left(E_{AB} - E_A^{AB} - E_B^{AB}\right) + \left(E_{BC} - E_B^{BC} - E_C^{BC}\right) + \\ &\left(E_{AC} - E_A^{AC} - E_C^{AC}\right) + E_{ABC} - E_{AB}^{ABC} - E_{BC}^{ABC} - E_{AC}^{ABC} + E_A^{ABC} + E_B^{ABC} + E_C^{ABC} = \\ &= E_{ABC} + \left(E_{AB} - E_{AB}^{ABC}\right) + \left(E_{BC} - E_{BC}^{ABC}\right) + \left(E_{AC} - E_{AC}^{ABC}\right) + (E_A + E_A^{ABC} - \\ &- E_A^{AB} - E_A^{AC}) + \left(E_B + E_B^{ABC} - E_B^{AB} - E_B^{BC}\right) + \left(E_C + E_C^{ABC} - E_C^{BC} - E_C^{AC}\right) \end{aligned} \tag{55}$$

Rearranging the terms, the Valiron and Mayer's hierarchical counterpoise

expression for the energy of the molecular complex is gained

$$E_{ABC}^{CP} = E_{ABC} + \left(E_A - E_A^{ABC}\right) + \left(E_B - E_B^{ABC}\right) + \left(E_C - E_C^{ABC}\right) + \\ + \left(\Delta E_{AB} - \Delta E_{AB}^{ABC}\right) + \left(\Delta E_{BC} - \Delta E_{BC}^{ABC}\right) + \left(\Delta E_{AC} - \Delta E_{AC}^{ABC}\right) \tag{56}$$

The last three extra terms with respect to the conventional counterpoise scheme of Eq. (54), correspond to the differences, for each dimer in the aggregate, between the dimer interaction energy computed with the dimer DCBS basis set and the same value computed using the whole basis set of the complex.

The other counterpoise scheme previously proposed by Wells and Wilson[85], the pairwise additive function counterpoise, can also be obtained in a systematic manner like the other two schemes discussed above. In this case, the three-body (or higher in case of an N-body cluster) interaction terms are not corrected following the counterpoise scheme. Instead, only the two-body interaction energies are corrected by using the DCBS. In the case of the considered trimer, the expression of the corrected energy can be easily obtained from Eq. (51)

$$\begin{aligned} E_{ABC}^{CP} &= E_A + E_B + E_C + \left(E_{AB} - E_A^{AB} - E_B^{AB}\right) + \left(E_{BC} - E_B^{BC} - E_C^{BC}\right) + \\ &+ \left(E_{AC} - E_A^{AC} - E_C^{AC}\right) + E_{ABC} - E_{AB} - E_{BC} - E_{AC} + E_A + E_B + E_C = \\ &= E_{ABC} + \left(E_A - E_A^{AB}\right) + \left(E_A - E_A^{AC}\right) + \left(E_B - E_B^{AB}\right) + \\ &+ \left(E_B + E_B^{BC}\right) + \left(E_C + E_C^{AC}\right) + \left(E_C + E_C^{BC}\right) \end{aligned} \tag{57}$$

The main feature of this approach is that the whole aggregate basis set is never used for any subunit calculation, except for the trivial case of a dimer. The counterpoise-correction is obtained by summing for all the subunits of the cluster, the differences between the MCBS and *all* the different DCBS descriptions of the given fragment.

Therefore, one can predict that this scheme will have problems to correct for the high-order BSSE but also to deal properly with the conventional, first-order, BSSE. Indeed, for a given N-body cluster, the energetic difference between the MCBS and the whole-complex basis set description of each fragment, as defined in the conventional counterpoise correction, is substituted by N-1 energy differences calculated using only the corresponding DCBS. Therefore, one might expect that this

scheme may have problems in dealing with cyclic or highly packed clusters where the presence of many close-by DCBS representations for each fragment may lead to an overestimation of BSSE.

The generalization for N-body clusters of these three function counterpoise schemes is straightforward.. The final expressions for the so-called site site, pairwise-additive and Valiron-Mayer hierarchical function counterpoise schemes, SSFC, PAFC, and VMFC, respectively, are the following

$$E_{i_1 i_2 ..i_N}^{SSFC} = E_{i_1 i_2 ..i_N} + \sum_{i_1}^{N} \left(E_{i_1}^{i_1} - E_{i_1}^{i_1 i_2 ..i_N} \right) \tag{58}$$

$$E_{i_1 i_2 ..i_N}^{PAFC} = E_{i_1 i_2 ..i_N} + \sum_{i_1 \neq i_2}^{N} \left(E_{i_1}^{i_1} - E_{i_1}^{i_1 i_2} \right) \tag{59}$$

$$\begin{aligned} E_{i_1 i_2 ..i_N}^{VMFC} = E_{i_1 i_2 ..i_N} + \sum_{i_1}^{N} \left(E_{i_1}^{i_1} - E_{i_1}^{i_1 i_2 ..i_N} \right) + \sum_{i_1 < i_2}^{N} \left(\Delta E_{i_1 i_2}^{i_1 i_2} - \Delta E_{i_1 i_2}^{i_1 i_2 ..i_N} \right) + \\ + \sum_{i_1 < i_2 < i_3}^{N} \left(\Delta E_{i_1 i_2 i_3}^{i_1 i_2 i_3} - \Delta E_{i_1 i_2 i_3}^{i_1 i_2 ..i_N} \right) + ... + \sum_{i_1 < i_2 < ... < i_{N-1}}^{N} \left(\Delta E_{i_1 i_2 ... i_{N-1}}^{i_1 i_2 ... i_{N-1}} - \Delta E_{i_1 i_2 ... i_{N-1}}^{i_1 i_2 ... i_N} \right) \end{aligned} \tag{60}$$

In Eq. (60), the third, fourth, and N^{th} term on the right hand side will be referred to as the *second-*, *third-*, and *N^{th}-order* CP contributions.

An important point is the scalability of these methods. Obviously, VMFC is the most accurate but the number of extra calculations needed rapidly increases with the cluster size. The SSFC method needs 2N+1 calculations, including the supermolecular one. For the PAFC, N times N-1 DCBS calculations plus N MCBS calculations must be carried out, that is N^2 extra energy calculations. In the case of the VMFC, it can be proved that the total number of calculations needed are determined by

$$\sum_{i=1}^{N} 2^{N-i} \binom{N}{i}$$

That means that a full hierarchical treatment of the BSSE for non-symmetric trimer, up to hexamer would involve 19, 65, 211 and 665 energy evaluations,

respectively. The correction just up to second-order BSSE will involve N (N+1) single-fragment plus $2\binom{N}{2}$ dimer calculations, that is a total of $2N^2 + 1$ energy evaluations. In this case *only* 19, 33, 51, and 73 calculations are needed for the trimer trough hexamer cluster series.

The use of the hierarchical scheme is clearly prohibitive even for small oligomers. However, in the particular case of highly symmetric clusters, the assessment of high-order BSSE effects is still possible since many fragment calculations are equivalent. Furthermore, due to the simplicity of the counterpoise correction terms, and as all of them are computed at a given cluster's geometry, the corrected gradients, Hessian and in general any derivative of the energy can be obtained by a linear combination of all the terms properly differentiated.

In Section II.1.4.1, the results obtained for some small clusters of hydrogen fluoride will be presented.

I.1.6 Other controversies and paradoxes

As mentioned before, the counterpoise method is nowadays widely used to estimate the BSSE in most *ab initio* calculations. Several studies have shown that:

a) It removes the BSSE in a way very similar to other sophisticated a priori methods, like the CHA at different levels of theory.

b) It tends to the values obtained with an inherent BSSE-free method like SAPT[29].

c) It asymptotically (but slowly for correlated methods) tends to zero when the basis set is increased[40] and the CP-corrected results using small basis sets are closer to the results with much larger basis sets.

However, most of these comparisons have been made upon calculations on benchmark systems like water or hydrogen fluoride clusters, van der waals complexes involving rare gas atoms, etc. With some exceptions[87], no extensive calculations have been carried out for open-shell interactions, charged complexes, or chemical reactions, and indeed some researchers do not include the counterpoise correction in those particular cases[88].

In this section, these problematic cases will be critically examined.

---

[87] N. Kobko, J. J. Dannenberg, J. Phys. Chem. A 105, 1944 (2001)

[88] a) Y. Xie, R. B. Remington and H. F. Schaeffer III, J. Chem.Phys. 101 4878 (1994), b) E. F. Valeev and H. F. Schaeffer III, J. Chem. Phys. 108 7197 (1998), c) J. R. Pliego Jr. and J. M. Riveros, J. Chem. Phys. 112 4045 (2000), d) H.-H. Bueker and E. Uggerud, J. Phys. Chem. 99 5945 (1995).

I.1.6.1 Charged and/or open-shell complexes

It has been pointed out that one has to be very careful when using the counterpoise method for open shell systems[89]. First, DCBS and MCBS fragment calculations must lead to the same spin state in order to assess properly the basis set extension effects. However, even when these aspects are taken into account properly, some ambiguities may rise when the PES exhibits two different states of the same symmetry, as pointed out by Alexander[90]

When the complex is charged another problem arise. In the case of ion-neutral molecule interaction, one can expect that the charge is located in one of the fragments, and thus define the fragments in this way. This approach was used in a previous study[42] about the interaction between several X^+ and X^- ions and hydrogen peroxide with satisfactory results. However, when the charge is not clearly located on one fragment, i.e., when charged molecules are involved in the complex, the definition of the fragments is sometimes not straightforward and unambiguous at all. This is the case, for instance, of protonated (deprotonated) dimers, nucleophilic substitution reactions or ionized dimers. The latter case is very peculiar because it involves a charged radical complex. As an example, let us consider a protonated dimer like $A\cdots AH^+$. One can define two fragments, A and AH^+ , or consider the proton as an independent moiety and define three fragments, the two neutral molecules and the proton. In a CP-corrected optimization, both fragment definitions may force artificially the optimized structure to be consistent with the complex partitioning selected. Therefore, if the complex would tend naturally to a symmetric interaction where the hydrogen atom involved in the hydrogen bond is equally shared by both A moieties, the use of the CP correction with the two fragments as commented above may artificially force the system towards a non symmetric interaction. The definition of three fragments could provoke the opposite effect for nonsymmetric interactions. One of the objectives of the present work is to check to which extent this assumption is true.

---

[89] Chalasinski and M. M. Szczesniak, Chemical Reviews 100, 4227 (2000).

[90] M. H. Alexander, J. Chem. Phys. 99 6104 (1993).

It is important to mention that those problems in the definition of the fragments are intimately related to the counterpoise method, not being directly related with the BSSE. Indeed, within the CHA framework, one has only to select which atoms belong to a given fragment but not deal with how many electrons or which spin state the fragment possesses. Because no extra fragment calculation is needed, the elimination of the BSSE is carried out consistently with the supermolecule calculation. This is not, however, a feature of all the so-called a priori methods. In the SCF-MI method of Gianninetti et. al.[55] , two modified SCF-like coupled equations must be solved (for a dimer), whose dimensions depend on the number of occupied molecular orbitals of each fragment. Therefore, in this framework, like in the counterpoise method, one has to deal with the ambiguity of the definition of the fragments.

Finally, another conceptual inconsistency of some of the BSSE-correction schemes is worth being pointed out. One has to bear in mind that the BSSE appears at the supermolecule level, which is the reason why the a priori methods do not need to deal with MCBS calculations. The BSSE is caused by the basis set extension that arise only on the part of the intermolecular complex's Hamiltonian assigned exclusively to each of the fragments. The BSSE cannot be eliminated properly by restricting the orbitals of each fragment to be expanded in their own basis set, because the pure interaction terms will be underestimated, as discussed in Section I.1.1.3.

In principle, one cannot partition the Hamiltonian operator in terms of intramolecular (fragment) and intermolecular terms because of the indistinguishability of the electrons. However, it has been shown in Section I.1.4 that in the framework of the second quantization this is possible as one deals with basis functions, instead. The CHA method is based on the elimination of those terms of the Fockian which account for the BSSE.

When using the counterpoise method one estimates those basis set extensions by computing the energy difference (and derivatives) of each fragment computed using the MCBS and DCBS, respectively. One assumes, hence, that this artificial energetic difference is equivalent to the effect on the supermolecule Fockian of those aforementioned basis set extensions. It has been shown that both somewhat antagonistic methods, behave in a very similar way, and somehow can be considered as complementary (while CP-corrected optimizations are straightforward, only the CHA provides BSSE-corrected wavefunctions). We are not going to discuss here to which extent the difference between the MBCS and DCBS calculations of the

fragments accounts for the BSSE of a given calculation, but we can show under which circumstances it cannot be exact.

A requisite that at least one can impose is that all the calculations, i.e., the MCBS of the fragments and DCBS of the fragments and complex, must be performed at the same level of theory. However, this rather trivial issue is not always fulfilled. For instance, let us consider a typical SCF calculation of a homonuclear dimer A···A, where A has an even number of electrons. In this particular case, with the CP method, the BSSE is estimated using a different method (UHF or ROHF) of calculation than that of the whole complex (RHF). An analogous situation occurs for ionized dimers like $(C_2H_4)_2^{\cdot +}$.

With these considerations in mind, the open questions are now: can we still use safely the counterpoise method or is it better not to correct for the BSSE in those particular cases? How do the usually ambiguous definition of the fragments affect the geometries, vibrational frequencies and interaction energies of those complexes?

These questions will be addressed in Section II.1.1.4

I.1.6.2 The intramolecular rotational barriers

The treatment of the BSSE in chemical reactivity is not trivial. In this particular case, the definition of the fragments changes during the process so that the continuous treatment of the BSSE is often possible only by taking each atom as a fragment. Another option is to define different fragments for reactants and products, but then the definition of the fragments of the transition state is rather ambiguous.

There have been a few attempts in the literature to take into account the BSSE in chemical reactivity, with different conclusions. Mayer et. al.[91] studied a proton transfer reaction at several levels of theory by using the counterpoise method. Also, whenever three fragments where defined, the hierarchical counterpoise method was used. They concluded that when the fragment definition is kept constant during the process, the results are not satisfactory, whereas the use of different fragment definitions for the reactants, products and TS's is inconsistent from a mathematical

[91] G. Lendvay, I. Mayer, Chem. Phys. Lett. 297, 365 (1998).

point of view.

In a recent paper, Kobko and Dannenberg[87] tested the use of the counterpoise correction in the study of several reactions, including a Diels-Alder condensation, which is particularly interesting since the product is usually a single molecule. They determined reaction barriers by using a different fragment definition for reactants and transition states. Finally, they concluded that the counterpoise-corrected results were better than the uncorrected ones and that the use of rather small basis sets in conjunction with the counterpoise correction to the complex gives satisfactory results.

On the other hand, one of the chemical processes where fragments are defined continuously constant along the reaction coordinate is the internal rotation in weakly bonded systems.

In the last years, Sordo and collaborators[92-94] have been questioning the validity of the counterpoise correction in the evaluation of energy barriers to internal rotation including the fragment relaxation term. The main argument was that since the same basis set is used to determine the minima and the corresponding transitions states connecting them, there must be no inconsistency in the calculation and therefore no BSSE.

The authors incorrectly claimed the nuclear relaxation to be responsible for the inconsistency of the counterpoise correction in this case. It has been shown in Section I.1.3 that the concept of BSSE fragment relaxation term is misleading. The expression for the BSSE-correction within the counterpoise approach is defined disregarding the fragments of the system being allowed or forbidden to relax their own geometry. One cannot split the BSSE-correction term into a *relaxation term* and an *intermolecular term*. When the intramolecular parameters are frozen, the CP-correction depends only on the intermolecular distances and angles, but the value of BSSE at a given geometry depends conceptually on all the parameters involved in the calculation. Nevertheless, experience shows that the intramolecular parameters are not too sensitive to the BSSE correction.

The papers by Sordo et al. have assumed that there is no BSSE at all in the

[92] V. M. Rayón, J. L. Sordo, Theor. Chem. Acc. 99, 68 (1998)

[93] V. M. Rayón, J. A. Sordo, J. Chem. Phys. 110, 377 (1999)

[94] J.A. Sordo, J. Chem. Phys. 106, 6204 (1997)

evaluation of energy barriers to internal rotation processes. For a system S, the value of the energy barrier is easily calculated as

$$E_b(S) = E_{TS}(S) - E_{min}(S) \tag{61}$$

E_{min} and E_{TS} being the energy of the minima and the energy of the connecting transition state structure involved in the process, respectively. Consistent results can be obtained provided that the system is properly described. However, both E^{TS} and E^{min} and hence E_b are BSSE-contaminated, i.e., if the system S is made up of two subsystems A and B. This is the case for weakly bonded complexes. Thus, the energy barrier should be calculated as

$$E_b^{CP}(S) = E_{TS}^{CP}(S) - E_{min}^{CP}(S) = E_{TS}(S) - E_{min}(S) + \delta_{TS}^{CP}(S) - \delta_{min}^{CP}(S) \tag{62}$$

The assumption of zero BSSE is only valid if the last two terms vanish, i.e., if BSSE were independent of the geometry, which is not at all the case. Therefore, it can be important to relocate the structures in the CP-corrected PES. Note that both $E_{TS}^{CP}(S)$ and $E_{min}^{CP}(S)$ can be obtained with the counterpoise receipt regardless the fragment relaxation being included or excluded.

To gain a deeper insight on this point, the results obtained for the rotational barriers of two weakly bonded systems will be discussed in Section II.1.2.1.

I.2 Methodological Aspects

In this section, the standard *ab initio* methods for solving the Schrödinger equation that have been used in this work are briefly overviewed. First, the equations resulting from the separation of the electronic and nuclear motion on the Hamiltonian through the Born-Oppenheimer approximation are given. The most commonly used technique to approximately solve the electronic Schrödinger equation is the Hartree-Fock, or Self-Consistent Field method, based on the variational theorem. Moreover, it represents the starting point for other sophisticated *ab initio* methods including the so-called electron correlation. Among these methods, only the Many-Body Perturbation Theory, and particularly the Møller-Plesset formulation, have been used in the calculations presented in this work.

A conceptually different methodology to obtain the energy and electron density of the chemical systems is the Density Functional Theory. In this work we have used the most commonly applied Kohn-Sham formulation, which is practically analogous to the Hartree-Fock equations. We have used the BLYP and B3LYP functionals for our calculations.

I.2.1 The Born-Oppenheimer approximation

The starting point is the non-relativistic time-independent Schrödinger equation.

$$\hat{H}\Psi(r,R)_n = \varepsilon_n \Psi(r,R)_n \tag{63}$$

The set of solutions [ε_n, ψ_n] describe stationary states of the system, the fundamental being the one with the lowest energy, ε_i, associated. The Hamiltonian of the system, $\hat{H}$, depends explicitly on the coordinates of the nuclei (R_α) and electrons (r_i)

$$\hat{H} = -\sum_{i}^{N}\frac{1}{2}\nabla_i^2 - \sum_{\alpha}^{M}\frac{1}{2M_\alpha}\nabla_\alpha^2 - \sum_{i}^{N}\sum_{\alpha}^{M}\frac{Z_\alpha}{r_{i\alpha}} + \sum_{i\neq i}^{N}\frac{1}{r_{ij}} + \sum_{\alpha\neq\beta}^{M}\frac{Z_\alpha Z_\beta}{R_\alpha} \qquad (64)$$

where all the terms are expressed in atomic units. N, M represent the number of electrons and nuclei, respectively. Z_α and M_α hold for the atomic number of nucleus α and the ratio between the mass of nucleus α and the mass of an electron, respectively. The equation above may be written more compactly as

$$\hat{H} = \hat{T}_e(r) + \hat{T}_N(R) + \hat{V}_{eN}(r,R) + \hat{V}_{NN}(R) + \hat{V}_{ee}(r) \qquad (65)$$

The first two term are the operators for the kinetic energy of electrons and nuclei; the third term corresponds to the electron-nuclei coulombic attraction; the last two terms represent the repulsion between the nuclei and electrons, respectively.

If the electronic and nuclear part could be separated, the molecular wavefunction would be expressed as a product of nuclear and electronic terms. However, the $\hat{V}_{eN}(r,R)$ term is responsible for the coupling of the electronic and the nuclear parts and the exact separation is not possible.

In the Born- Oppenheimer approximation, the nuclear kinetic energy term $\hat{T}_N(R)$, is neglected, and the nuclear repulsion, $\hat{V}_{NN}(R)$, is considered as a constant value. That means it only adds a constant value to the eigenvalues and has no effect on the corresponding eigenfunctions, so it can be dropped from the Hamiltonian. The remaining terms form the so-called electronic Hamiltonian

$$\hat{H}_{el} = -\sum_{i}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i}^{N}\sum_{\alpha}^{M}\frac{Z_\alpha}{r_{i\alpha}} + \sum_{i\neq i}^{N}\frac{1}{r_{ij}} \qquad (66)$$

That depends explicitly on the electron coordinates and parametrically on the nuclear positions. The solution of the Schrödinger equation involving the electronic Hamiltonian yield the electronic wavefunction and the corresponding electronic energy. Additionally, if the spin-orbit coupling effects are important, they can be included in the electronic Hamiltonian at each nuclear geometry using the corresponding spin-orbit Hamiltonian.

$$\hat{H}_{el}^{'} = \hat{H}_{el} + \hat{H}_{so} \tag{67}$$

The total energy for fixed nuclei is obtained by adding the nuclear repulsion to the electronic energy

$$\varepsilon_{total} = \varepsilon_{elec} + \sum_{\alpha \neq \beta}^{M} \frac{Z_\alpha Z_\beta}{R_\alpha} \tag{68}$$

Within the Born-Oppenheimer approximation, the solution for the motion of the nuclei can be obtained in a similar way using the nuclear Hamiltonian

$$\hat{H}_{nuc} = -\sum_{i}^{M} \frac{1}{2M_\alpha} \nabla_i^2 + \sum_{\alpha \neq \beta}^{M} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} + < \varepsilon_{elec} > \tag{69}$$

that includes the kinetic and coulombic energy repulsion for the nuclei and the electronic energy averaged over the electronic wavefunction at a given, fixed, nuclear geometry. Hence, the nuclei within this approximation are considered to move on a potential energy surface obtained by solving the electronic problem. The solutions of the nuclear motion give information about the vibrations, rotations, and translations of the molecule.

When wavefunctions that diagonalize the electronic Hamiltonian are used, the potential energy surfaces obtained are called adiabatic. This is the general procedure. In some cases, namely when studying low lying excited states that may result in conical intersections, it is preferable to minimize some coupling terms resulting from the expansion of the total wavefunction in terms of electronic and nuclear wavefunctions. In this case, the potential energy surfaces are called diabatic and are free to cross.

In short, the Born-Oppenheimer approximation is of capital importance in applied quantum chemistry. It is based on the assumption that the motion of the nuclei is much slower than that of the electrons since the nuclei are much heavier than the electrons. Therefore, one can consider that the nuclei are fixed in the space and the

electrons move in the field of the frozen nuclei and that of the electrons themselves. In this way, an electronic Schrödinger equation, that depends on both the nuclear and electronic coordinates, is solved at a given nuclear position only for the electronic motion. In other words, the initial equations depends parametrically on the nuclear coordinates. By solving the electronic Schrödinger equation at different nuclear arrangements one can obtain a potential energy hypersurface of the total energy, where only the nuclear kinetic energy term has been ignored. This term contributes to the vibrational, rotational and translational molecular energy and can be determined by solving the corresponding nuclear Schrödinger equation.

In the present work, only ground-state electronic wavefunctions and energies have been considered.

I.2.2 The Hartree-Fock method

The most commonly used technique to find approximate solutions to the electronic Schrödinger equation is the so called molecular orbital approximation. Within this approximation, the electrons are considered to occupy molecular spin-orbitals of a given energy. In the single-determinant approximation, the approximate wavefunction to the ground-state of the molecular system is expressed as a determinant formed from these one-electron molecular spin-orbitals, which diagonalize an approximate one-electron operator, the Fock operator. This single determinant, the so-called Slater determinant, ensures the necessary antisymmetry with respect to the exchange of two electrons of the molecular wavefunction.

According to the variational principle, the best spin-orbitals are the ones that minimize the electronic energy computed as the expectation value of the electronic Hamiltonian. Let us consider a trial wavefunction

$$|\Psi_0>=|\chi_1\chi_2\cdots\chi_N> \tag{70}$$

made up of N spin-orbitals. Without loss of generality, these spin-orbitals can be chosen to form an orthonormal set. In order to obtain the molecular spin-orbitals to build the wavefunction, the energy expression for a single Slater determinant is minimized with respect to the variation of the spin-orbitals, subject to the constraint

that the spin-orbitals remain orthonormal

$$< \chi_i \mid \chi_j >= \delta_{ij} \tag{71}$$

Under these conditions, the expectation value of the electronic Hamiltonian over the trial wavefunction can be expressed as

$$E_0 =< \Psi_0 \mid \hat{H}_{elec} \mid \Psi_0 >= \sum_i^N < \chi_i \mid \hat{h} \mid \chi_i > + \frac{1}{2} \sum_{i,j}^N \left[\chi_i \chi_j \mid \chi_i \chi_j\right] - \left[\chi_i \chi_j \mid \chi_j \chi_i\right] \tag{72}$$

where the two electron integrals are written following the [12|12] formalism.

It can be shown that the minimization of Eq. (72) with respect to the spin-orbitals and imposing the constraints given in Eq. (71) for each pair of spin-orbital leads to a set of equations of the following form

$$f \mid \chi_i >= \sum_j^N \varepsilon_{ij} \mid \chi_j >; \forall i = 1, N \tag{73}$$

The Fock operator, *f*, is a single-particle operator expressed as the sum of a core Hamiltonian, including the kinetic energy and nuclear attraction, and an effective one-electron potential due to the N-1 electrons in the molecular system. Indeed, the form of this operator depends explicitly on the spin-orbitals. Thus, the solution of these equations must be performed iteratively until self-consistency. This is the reason why the Hartree-Fock equations are sometimes referred as Self-Consistent Field equations.

However, it is convenient to formally perform unitary transformation among the spin-orbitals to obtain the so-called canonical spin-orbitals, where the matrix of the Lagrange multipliers ε_{ij} is diagonal. The final expression take the following form

$$f \mid \chi_i >= \varepsilon_i \mid \chi_i > ; \forall i = 1, N \tag{74}$$

where the χ_i and ε_i represent the canonical spin-orbital and the orbital energy associated. By substituting Eq. (74) into Eq. (72), the Hartree-Fock energy is obtained in terms of the spin-orbital energies and two-electron electron repulsion integrals

$$E_0 = \sum_i^N \varepsilon_i - \frac{1}{2}\sum_{i,j}^N [\chi_i \chi_j \mid \chi_i \chi_j] - [\chi_i \chi_j \mid \chi_j \chi_i] \tag{75}$$

Note that the total energy is not just the sum of the one-electron orbital energies, as the electron-electron interactions are counted twice.

In practical implementations, for closed-shell systems, the spin is integrated and the spin-orbitals are transformed into doubly-occupied spatial molecular orbitals. These orbitals are expressed as linear combinations of a given set of pre-established functions and the coefficients of the linear combination are obtained in a self-consistent fashion. These set of functions are usually referred as atomic orbitals basis set. In the present work only Gaussian-type nuclear centered basis sets have been used, but also Slater-type or the so-called plane waves could be also applied. Once a basis set is defined, the Fock operator is transformed into a Fock matrix expanded on the atomic orbitals and the set of Eq. (74) transform into a pseudo-secular equation of the following form

$$F(C)C = \varepsilon S C \tag{76}$$

where matrix S represents the metric matrix on the given atomic orbital basis set, which, in general, is not diagonal. The coefficients of the molecular orbitals on the atomic orbital basis are collected in the columns of matrix C with the corresponding orbitals energies on the diagonal of the diagonal matrix ε. The dimensions of the involved matrices is M×M, where M is the total number of atomic orbitals. Therefore, upon self-consistency, the N/2 molecular orbitals with the corresponding lowest orbital energies are considered as occupied orbitals and are taken to build the ground-state Slater determinant. The rest of the molecular orbitals, the virtual orbitals, do not influence the Hartree-Fock energy or wavefunction, but are used for further inclusion of the so-called correlation energy.

The correlation energy is defined as the difference between the exact non-relativistic Born-Oppenheimer energy of the system and the Hartree-Fock energy obtained in the basis set limit. Indeed, the Hartree-Fock method has its limitations. The main problem is that only the electrons with the same spin are correlated, that is,

for a Slater determinant wavefunction, the probability of finding simultaneously two electrons of the same spin in the same position in the space is zero, but this is not the case for electrons of different spin.

The restricted Hartree-Fock method fails to describe the dissociation of molecules into open-shell fragments and it is unsuitable also for the description of excited states. These problems can be solved in some cases by using more than a single determinant to express the approximate wavefunction, and hence dealing with the Multi-Configurational Self- Consistent Field equations. However, none of these particular cases have been treated in this work so we will not discuss the details of this methodology here.

However, another drawback directly related to the chemical systems studied in this work is that it does not account for the dispersion part of the interaction energy. The dispersion energy results from the mutual polarization of the electron densities of the molecules and it is essential for the intermolecular bonding of van der Waals complexes and non-polar weakly bound complexes in general. In the case of hydrogen bonding, the inclusion of the dispersion energy is still important in order to correctly describe the anisotropy of the interaction and, in principle, to obtain more reliable molecular geometries.

The easiest way to introduce the electron correlation from a Hartree-Fock calculation, and thus include, among other contributions, the dispersion energy, is the Many-Body Perturbation Theory (MBPT). In the next section, the most commonly formulation of the MBPT is briefly discussed.

I.2.3 Møller-Plesset Perturbation Theory

In quantum chemistry, the Perturbation Theory is based on the partition of a given Hamiltonian into a zero-order part, $\hat{H}_0$, which has known solutions in terms of eigenfunctions and eigenvalues, and a perturbation operator, *V*, so that

$$\hat{H} = \hat{H}_0 + \hat{V} \tag{77}$$

Within the framework of the so-called Rayleigh-Schrödinger Perturbation

Theory (RSPT), the solutions of the initial Hamiltonian are expressed as a summation of infinite contributions for both the wavefunction and molecular energy.

$$\begin{aligned} |\Psi_i> &= |\Psi_i^{(0)}> + |\Psi_i^{(1)}> + |\Psi_i^{(2)}> + ... + |\Psi_i^{(\infty)}> \\ E_i &= E_i^{(0)} + E_i^{(1)} + E_i^{(2)} + ... E_i^{(\infty)} \end{aligned} \tag{78}$$

In general, the zero-order eigenfunctions, exact solutions, Ψ_i , of the unperturbed Hamiltonian, are used as a basis to expand the higher-order corrections. These eigenfunctions are constructed from both the occupied and virtual molecular spin-orbitals obtained in the previous Hartree-Fock calculation. The singly-excited zero-order wavefunctions are constructed by substituting one occupied spin-orbital by another from the set of virtual orbitals. The doubly-excited wavefunctions involve double substitutions of occupied spin-orbitals and so on.

Practically, one often starts from the Hartree-Fock wavefunction of the system and wishes to improve the description of the system. Within the framework of the RSPT, if the zero-order Hamiltonian is taken as the Fock operator, the so-called Møller-Plesset Perturbation Theory (MP) equations are obtained.

Let us start from the Hartree-Fock description of the molecular system

$$\hat{H}_0 |\Psi_0^{(0)}> = E_0^{(0)} |\Psi_0^{(0)}> \tag{79}$$

Now, the “zero” subscripts indicates that we are dealing with the ground-state wavefunction whereas the “zero” superindex gives the order of correction. In this case, we are considering obviously the zero-order correction to the exact solution. From the RSPT, it is well-known that the N^{th}-order correction to the energy is expressed simply as

$$\begin{aligned} E_i^{(0)} &= <\Psi_i^0 | \hat{H}_0 | \Psi_i^{(0)}> \\ E_i^{(n)} &= <\Psi_i^0 | \hat{V} | \Psi_i^{(N-1)}> ; \forall n > 0 \end{aligned} \tag{80}$$

It is easy to see that the summation up to first-order correction on the energy corresponds to the Hartree-Fock energy

$$E_0^{(0)} = < \Psi_0^0 \mid \hat{H}_0 \mid \Psi_0^0 > = \sum_i^N \varepsilon_i \tag{81}$$

$$\begin{aligned} E_0^{(1)} &= < \Psi_0^{(0)} \mid \hat{V} \mid \Psi_0^{(0)} > = < \Psi_0^{(0)} \mid \sum_{i>j} r_{ij}^{-1} \mid \Psi_0^{(0)} > - < \Psi_0^{(0)} \mid \hat{v}^{HF} \mid \Psi_0^{(0)} > = \\ &= \frac{1}{2} \sum_{i,j}^N \left[\chi_i \chi_j \parallel \chi_i \chi_j \right] - \sum_{i,j}^N \left[\chi_i \chi_j \parallel \chi_i \chi_j \right] = -\frac{1}{2} \sum_{i,j}^N \left[\chi_i \chi_j \parallel \chi_i \chi_j \right] \end{aligned} \tag{82}$$

$$E_0^{(0)} + E_0^{(1)} = \sum_i^N \varepsilon_i - \frac{1}{2} \sum_{i,j}^N \left[\chi_i \chi_j \parallel \chi_i \chi_j \right] = E_0^{HF} \tag{83}$$

being

$$\left[\chi_i \chi_j \parallel \chi_i \chi_j \right] = \left[\chi_i \chi_j \mid \chi_i \chi_j \right] - \left[\chi_i \chi_j \mid \chi_j \chi_i \right] \tag{84}$$

Note that the perturbation operator is taken as the difference between the actual two-electron repulsion operator and the sum of the averaged one-electron potential operators.

The first correction to the Hartree-Fock energy occurs in the second order of the perturbation expansion. The general expression for $E_i^{(2)}$, obtained by expanding the first-order correction to the wavefunction in terms of the zero-order solutions is as follows

$$E_0^{(2)} = < \Psi_0^{(0)} \mid \hat{V} \mid \Psi_0^{(1)} > = \sum_{k \neq 0} \frac{\left| < \Psi_0^{(0)} \mid \hat{V} \mid \Psi_k^{(0)} > \right|^2}{E_0^{(0)} - E_k^{(0)}} \tag{85}$$

Among all the *k* zero-order wavefunctions, only the ones involving double excitations contribute to the second-order correction to the energy. Hence, the second order correction to the energy can be finally expressed in terms of the molecular spin-orbital energies and the two-electron integrals of the spin-orbitals.

$$E_0^{(2)} = \sum_{\substack{a<b \\ r<s}} \frac{\left| \left[\chi_a \chi_b \| \chi_r \chi_s \right] \right|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} \tag{86}$$

The so-called MP2 energy is obtained by truncating the expansion of the exact energy up to the second-order correction. In contrast to the Hartree-Fock energy, it is not variational, that is, it might be *below* the exact energy. Indeed, the perturbational series might not be convergent. However, the great advantage of this method with respect to other variational methods that introduce electron correlation is that it is size-consistent. This property is essential when dealing with intermolecular interactions since the errors due to size-inconsistency and size-extensitivity may be larger than the BSSE itself.

I.2.4 Density Functional Theory

Density Functional methods represent an alternative to introduce electron correlation on a single-determinant wavefunction at a very low computational cost. Methods derived from the Density Functional Theory have been used for years for the physicists to study solid state and electron gas, but only in the last decades it has been applied to isolated molecular systems.

The original formulation of the DFT tells us that the energy of a system can be exactly expressed as a functional of the one-electron density, ρ. Formally, it exists a functional E[ρ] that determines the ground-state molecular electronic energy. Unfortunately, this *universal* functional is not known. Obviously, the goal is to obtain a good approximation of the universal functional to be applied to any one-electron density.

In principle, since the energy can be extracted from the one-electron density, there is no need to introduce the molecular orbital approximation in the DFT methodology. However, in the most successful DFT methodology introduced by Kohn and Sham, the electron density is expressed as a sum over some one electron occupied molecular orbitals

$$\rho(r) = \sum_{i}^{N_{occ}} \left| \chi_i \right|^2 \tag{87}$$

and the application of the variational principle to the energy functional lead to a set of one-electron equations analogous to the Hartree-Fock ones, where the molecular spin-orbitals are obtained from a Kohn-Sham-derived Hamiltonian.

$$\hat{h}_{KS} \mid \chi_i >= \varepsilon_i \mid \chi_i > ; \forall i = 1, N \tag{88}$$

From the non-interacting particles model, the Kohn-Sham Hamiltonian can be expressed formally as

$$\hat{h}_{KS}(r) = -\frac{1}{2}\nabla^2 + v(r) + \frac{1}{2}\int \frac{\rho(r')}{\mid r - r' \mid} dr' + v_{xc}(r) \tag{89}$$

where the first term represents the exact kinetic energy for a non-interacting electronic system; the second term is the external potential due to the nuclei; the third term is the coulombic electron-electron interaction, and the last term represents the so-called exchange-correlation functional. It must include the true effect of the electron correlation and exchange on the kinetic energy and the electron-electron repulsion.

Since the Kohn-Sham Hamiltonian depends explicitly on the one-electron density (second and third terms) , the solution of Eq. (88) must be carried out iteratively until self-consistency is reached.

For practical applications, a DFT calculation is analogous to a Hartree-Fock one, with the exception that, due to the complexity of the exchange-correlation functional, the necessary integrations are carried out numerically.

The election of the exchange-correlation functional is the key point in a DFT calculation. There is a wide number of functionals for the exchange and for the correlation counterparts that can be chosen in a rather arbitrary way. There are basically two families of functionals for the exchange-correlation energy: the ones based on the Local Density Approximation (LDA) or including gradient corrections, like the Generalized Gradient Approximation (GGA). The former depend only on the

value of the density at each point of the space, whereas the later include also dependencies of the gradient of the density and hence a two-electron character is simulated. The GGA functionals seem to describe more properly the exchange-correlation hole than the LDA functionals and are the most widely used in the literature. In this work, only the GGA-type BLYP (Becke 88 exchange[95] and Lee, Yang, Parr correlation[96]) and B3LYP (Becke's three parameter exchange[97] and Lee, Yang, Parr correlation) functionals have been applied. for the calculations of intermolecular complexes.

These two functional belong to the so-called hybrid[98] exchange-correlation functionals. The hybrid model rests on a linear combination of HF exchange with DFT exchange-correlation contributions:

$$E_{xc} = a_0 (E_x^{HF} - E_{xc}^{GGA}) \tag{90}$$

In the particular case of the BLYP functional, the exchange-correlation functional is taken from the Becke 88 exchange and LYP correlation.

The most popular implementation, the B3LYP method, uses three parameters to combine the B exchange functional, the LYP correlation functional, the expression for the exchange of a uniform spin-polarized electron gas, and the correlation component, represented by the Vosko-Wilk-Nusair parametrization (VWN).

$$E_{xc}^{B3LYP} = a_{0x} E_x^{LSD} + (1 - a_{0x}) E_x^{HF} + a_{1x} E_x^{B} + a_c E_c^{LYP} + (1 - a_c) E_c^{VWN} \tag{91}$$

The constants a_{0x}, a_{1x}, and a_c, are those determined by Becke by fitting to the G1 molecule set. Becke determined the values of the three parameters by fitting to the 56 atomization energies, 42 ionization potentials, 8 proton affinities, and 10 first-row atomic energies in the G1 molecule set, computing values of a_{0x} = 0.80, a_{1x}, = 0.72 and a_c = 0.81.

---

[95] A.D. Becke, *Phys. Rev. A,* **38**, 3098 (1988)

[96] C. Lee, W. Yang, and R.G. Parr, *Phys. Rev. B,* **37**, 785 (1988)

[97] A.D. Becke, *J. Chem. Phys.*, **98**, 5648 (1993)

[98] In the hybrid functionals, the exact exchange part of the exchange-correlation energy is taken from the Hartree-Fock exchange.

The eminently local character of the existing functionals, even for the GGA-type ones, in combination with the single-determinant nature of the electron density, prevents, a priori, for a correct description of the dispersion energy. It has been pointed out that the existing DFT methodology fail to describe van der Waals complexes, where the dispersion interaction is the dominant term. However, for hydrogen bonded complexes the results are comparable to the ones obtained with more sophisticated and much more computational costly *ab initio* methodologies.

I.3 Objectives

This thesis was originally thought to deal with the Basis Set Superposition Error from both a methodological and a practical point of view.

The purpose of the present thesis is twofold: (a) to contribute step ahead in the correct characterization of weakly bound complexes and, (b) to shed light the understanding of the actual implications of the basis set extension effects in the *ab intio* calculations and contribute to the BSSE debate.

One of the main goals in the field of intermolecular interactions is to be able to obtain reliable *ab initio* descriptions of the intermolecular interactions occurring on large systems with a low computational cost. In this work we have explored whether the use of the counterpoise method, or any BSSE-correction scheme, in conjunction with affordable basis sets can be used for this purpose. Prior to fulfill such a broad purpose, the existing BSSE-correction procedures must be analyzed, compared, validated and, if necessary, improved.

In this sense, in the workgroup of the Institute of Computational Chemistry a novel methodology was developed for the calculation of counterpoise corrected geometry optimizations and vibrational frequencies.

The first goal to accomplish was to improve the optimization algorithm in order to reduce the computational cost of the calculations. Several methodologies were implemented and tested, like the DIIS, C2-DIIS, and the respective combination with variable metric methods like the BFGS or DFP and linear search techniques. Numerical results showed an excellent performance of the combined DIIS-BFGS method, as long as a initial hessian was provided.

The work on the code has been carried out continuously along the duration of this thesis. Finally, the whole code was rewritten in order to implement the several counterpoise schemes for the study of molecular clusters. The culmination of this effort arrived when the method was implemented into the Gaussian98 package for the use of the whole scientific community. At this point, the CP methodology can automatically be applied for IRC calculations, PES scans, third derivatives of the energy, etc...

Nevertheless, the last version of the CP program is still useful in order to take the molecular symmetry into account to drastically reduce the computational time and also for future development of the counterpoise methodology. A detailed explanation of the features of the program an its applicability is given in Section II.2.1.

Along with the development of the counterpoise methodology, the use of an aprioristic method was necessary for a further exploration of the BSSE effects. After a critical analysis of the existing methodologies, the Chemical Hamiltonian Approach seemed the best option from a conceptual point of view. Furthermore, the difficulties in the practical implementation of the method, mainly at the CHA-MP2 level of theory, represented a challenge for the author of this work. In collaboration with Prof. Mayer, the CHA methodology was successfully implemented at the SCF level both restricted and unrestricted. Even though the CHA-MP2 methodology for an arbitrary number of fragments has been implemented, the particular case of complex virtual orbitals and energies has not been solved yet and hence no results are presented here. Application to intramolecular BSSE was tested with unsatisfactory results.

Once the we disposed of a set of efficient programs, the next step was obviously to explore the BSSE effects in any property of a molecular complex. For the energy-related properties like Potential energy surfaces, the counterpoise method was used. When dealing with the wavefunction and density-related properties, the CHA methodology was the chosen one.

Obviously, the first step was to compare both BSSE-correction schemes by means of a comparative systematic studies at different levels of theory and basis sets. The results are presented in Section II.1.1.1.

Then, it was observed that the convergence towards zero for the BSSE correction when applied to electron correlation methods like the MP2. An exploration of the Complete basis set limit of both uncorrected and CP-corrected molecular properties like stabilization energies and intermolecular distances. It was shown that the CP-corrected results monotonically behave with respect to the increasing basis set size, whereas this is not always true for the uncorrected values. The results of these large calculations are given in Section II.1.1.2.

Once the validity of the CP-correction was verified, we decided to apply our

methodology to actual controversial topics.

First, the application of the counterpoise method was applied to internal rotational barriers. It was shown that the BSSE-correction was essential for these chemical processes, contrary to the existing opinion that there is no BSSE involved at all. The results are presented in Section II.1.2. The viability of the CP method for dealing with charged complexes and the BSSE effects on the double-well PES blue-shifted hydrogen bonds was also studied in detail. Our findings are given in Sections II.1.1.3 and II.1.1.4

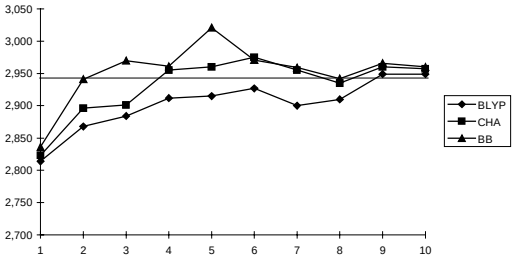
On the other hand, the lack of studies in the literature covering the effect of the BSSE on the molecular orbitals and, in essence, on the wavefunction, motivated us to investigate the effect of the BSSE on the electron density-related properties. For practical reasons, the CHA/F and CHA/DFT methodologies, implemented by Dr. Paizs into Gaussian 92, were the chosen to retrieve the corresponding BSSE-corrected molecular orbitals. The first-order electron density was used to assess, both graphically and numerically, the redistribution of the charge density upon BSSE-correction. Several different tools were used to deeply analyze, for the first time, the BSSE effects on the electron density of several hydrogen bonded complexes of increasing size. The results can be found in Section II.1.3.

Finally, in order to complement this work, the intermolecular perturbation theory was also explored during a collaboration with Prof. Szczesniak. The BSSE-free SAPT method was used for several studies of molecular aggregates, like hydrogen fluoride and carbon dioxide cluster. Despite the BSSE-free features of the method, it was recognized that the use of a counterpoise-corrected PES was essential in order to determine the molecular geometry to perform the SAPT analysis. These differences between the uncorrected and CP-corrected geometries yield differences on the SAPT expansion terms of up to 100%. These results are collected in Section II.1.4

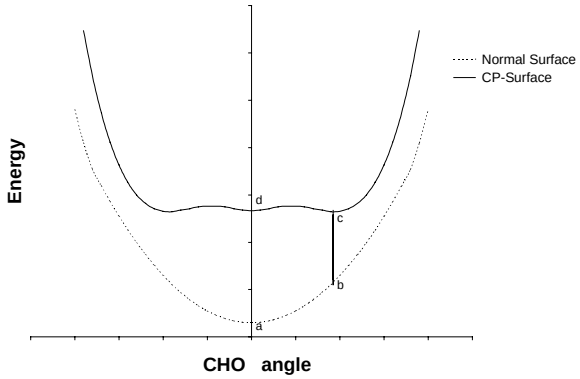
II Results

II.1 On the effect of the BSSE on the properties of molecular complexes

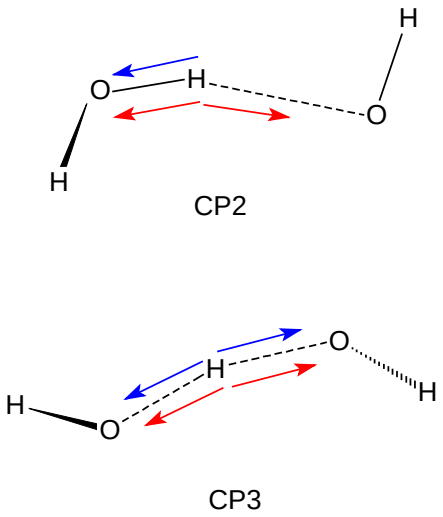
II.1.1 Potential Energy Surfaces and Stabilization energies



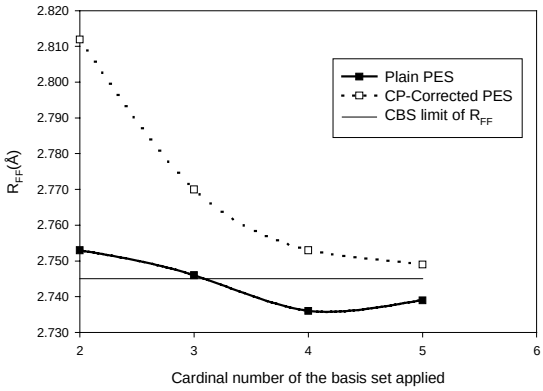
(A)



(B)



(C)



(D)

A) CP and CHA Intermolecular distances for the complex $(H_2O)_2$ vs. basis sets

B) Comparison between conventional and CP-corrected double-well PES

C) Qualitative picture of the acting forces on the central H atom of the $H_2O\cdots OH^-$ complex

D) Comparison of the plain and CP-corrected r_{FF} distances for $(HF)_2$ in the CBS limit

II.1.1.1 Comparison and validation of the CP and CHA methods

The main features of the *a priori* and *a posteriori* correction methods have already been discussed in the sections I.3 and I.4. As mentioned before, it has been widely accepted that the BSSE represents more than just an unbalance between the energies of the complex and its fragments when computing the interaction energy. The BSSE is a phenomenon related to the LCAO approximation that affects the whole description of the complex, i.e. stationary points, vibrational frequencies, wavefunction, etc.

The appearance of the so-called *a priori* BSSE-correction methods helped to understand this point of view and to recognize that the well-known and widely used Boys-Bernardi (BB) or counterpoise correction is not an energetic correction to the interaction energy but rather to the complex's energy. This interpretation permits the definition of counterpoise-corrected derivatives of the energy in order to obtain the stationary points on the counterpoise-corrected potential energy surface, and also vibrational frequencies, dipolar moments, spin-spin coupling constants, etc. In the case of the *a priori* methods, besides all of the above molecular properties, the important concept of the BSSE-corrected wavefunction is also gained.

Despite this, no exhaustive comparison of the two antagonistic methods in terms of molecular geometry have been carried out so far. Indeed, no geometry optimization have been reported at the CHA level of theory.

In this section we address the problems of intermolecular structure determination described above. One of the most important questions is how the BSSE distort the calculated PES and how the *a posteriori* and *a priori* BSSE correction schemes perform compared to each other. To evaluate these problems we carried out calculations at the SCF, CHA/F-SCF, SCF-CP, DFT, CHA/F-DFT, DFT-CP, MP2, and MP2-CP levels of theory in conjunction with twelve different basis sets of increasing size. Geometry parameters and stabilization energies were determined for the $(HF)_2$, $(H_2O)_2$, and H_2O-HF complexes by using all the possible theoretical models which can be derived from the sets of methods and basis functions listed above. We chose these complexes because

accurate experimental results are available and it is well known that DFT methods predict quite reasonable results for these particular cases.

All the *ab initio* calculations were carried out using the Gaussian92[99] and Gaussian94[100] program systems. For the CHA/F and CHA/DFT computations we used a modified version of Gaussian92[77], from Prof. Paizs. For the CHA/F geometry optimizations the gradients were determined by the finite difference method and the eigenvector following method was used to locate the stationary points. For the CP-corrected geometry optimizations we used a program based on the Ref. 48 with some small modifications in some Gaussian links (for further details see Section II.2.1). For the calculations Pople's 6-31G, 6-31G(d), 6-31G(d,p), 6-31++G(d,p), 6-311G(d,p), 6-311++G(d,p), 6-311++G(2df,2p), and 6-311++G(3df,2pd) and Dunning's TZV(d,p), TZV(d,p)++, aug-cc-pVDZ, and aug-cc-pVTZ basis sets were used. In the DFT calculations the BLYP and potentials were applied.

The results for the three studied systems will be discussed separately. Also, the rather unexpected values obtained when using the aug-cc-VXZ basis set series will be analyzed.

II.1.1.1.a Hydrogen fluoride dimer

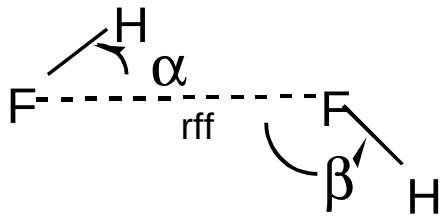
The HF dimer has a linear structure which is known from both experiment[101] and

[99] M.J. Frisch, G.W. Trucks, M. Head-Gordon, P.M.W. Gill, M.W. Wong, J.B. Foresman, B.G. Johnson, H.B. Schlegel, M.A. Robb, E.S. Reploge, R. Gomperts, J.L. Andres, K. Raghavachari, J.S. Binkley, C. Gonzalez, R.L. Martin, D.J. Fox, D.J. Defrees, J. Baker, J.J.P. Stewart and J.A. Pople, Gaussian, Inc., Pittsburgh PA, 1992.

[100] M.J. Frisch, G.W. Trucks, H.B. Schlegel, P.M.W. Gill, B.G. Johnson, M.A. Robb, J.R. Cheeseman, T.A. Keith, G.A. Petersson, J.A. Montgomery, K. Raghavachari, M.A. Al-Laham, V.G. Zakrzewski, J.V. Ortiz, J.B. Foresman, J. Cioslowski, B.B. Stefanov, A. Nanayakkara, M. Challacombe, C.Y. Peng, P.Y. Ayala, W. Chen, M.W. Wong, J.L. Andres, E.S. Reploge, R. Gomperts, R.L. Martin, D.J. Fox, J.S. Binkley, D.J. Defrees, J. Baker, J.P. Stewart, M. Head-Gordon, C. Gonzalez and J.A. Pople, Gaussian, Inc., Pittsburgh PA, 1995.

[101] B. J. Howard, T. R. Dyke, and W. Klemperer, *J. Chem. Phys.*, 81, 5417 (1984), A. S. Pine, B. J. Howard, J. Chem. Phys. 84, 590 (1983)

high level calculations[102] However, it is also well known[103,104] that DFT methods in conjunction with small basis sets tend to predict the cyclic structure to be the only stable one on the PES.



SCHEME 7:Geometrical parameters of the $(HF)_2$

In the following, both these pathological and the linear cases will be examined from the point of view of BSSE. Tables I and II summarize selected geometrical parameters of the HF dimer (see SCHEME 7) calculated at the uncorrected and corrected SCF, BLYP, B3LYP, and MP2 levels of theory using the Pople and the TZP** and TZP++** basis sets. The corresponding stabilization energies are shown graphically in Figure 1. The dependence of the distance between the fluorine atoms on the applied theoretical model is also shown graphically in Figure 2.

[102] C. Maerker, P. v. R. Schleyer, K. R. Liedl, T.-K. Ha, M. Quack, and M. A. Suhm, *J. Comput. Chem.*, **18**, 1695 (1997) and references therein.

[103] J. E. Del Bene, W. B. Person, and K. Szczepaniak, *J. Phys.. Chem.*, **99**, 10705 (1995)

[104] P. Hobza, J. Sponer, and T. Reschel, *J. Comput. Chem.*, **11**, 1315 (1995)

TABLE I.

Geometrical parameters for the HF-HF dimer calculated in ten different basis sets at the SCF, CHA-SCF, SCF-CP, MP2 and MP2-CP levels of theory. The number of basis functions is given in parentheses. The experimental values of *rff* (r_e), α and β are 2.72 Å, 10 ± 6 degrees and 117 ± 6 degrees, respectively. For notation, see SCHEME 7.

Basis Set	Method	rff (Å)	α	β	Method	rff (Å)	α	β
6-31G	SCF	2.706	8.3	126.0	MP2	2.719	12.5	109.0
(22)	CHA-SCF	2.713	5.1	135.6	MP2-CP	2.789	4.5	133.8
	SCF-CP	2.741	3.7	143.1				
6-31G*	SCF	2.709	17.1	96.6	MP2	2.535	45.3	51.8
(34)	CHA-SCF	2.756	7.6	114.8	MP2-CP	2.790	6.5	113.7
	SCF-CP	2.798	6.9	117.7				
6-31G**	SCF	2.725	14.4	101.7	MP2	2.541	47.8	48.9
(40)	CHA-SCF	2.760	8.3	116.7	MP2-CP	2.799	6.6	115.2
	SCF-CP	2.811	6.7	120.1				
6-31++G**	SCF	2.812	8.0	120.0	MP2	2.776	7.6	115.3
(50)	CHA-SCF	2.831	7.4	120.8	MP2-CP	2.836	7.7	114.9
	SCF-CP	2.842	7.5	120.8				
6-311G**	SCF	2.773	11.6	112.6	MP2	2.710	17.8	94.7
(50)	CHA-SCF	2.822	7.7	122.8	MP2-CP	2.850	6.0	121.9
	SCF-CP	2.850	6.3	126.3				
6-311++G**	SCF	2.825	6.4	126.7	MP2	2.791	6.6	120.9
(60)	CHA-SCF	2.868	6.8	126.3	MP2-CP	2.880	7.3	120.6
	SCF-CP	2.871	7.0	126.4				
TZV**	SCF	2.796	5.9	128.1	MP2	2.756	5.6	122.8
(52)	CHA-SCF	2.824	5.7	129.4	MP2-CP	2.842	5.5	124.9
	SCF-CP	2.827	5.7	129.9				
TZV++**	SCF	2.809	5.9	129.5	MP2	2.782	5.6	125.1
(62)	CHA-SCF	2.829	6.6	127.9	MP2-CP	2.851	7.4	121.8
	SCF-CP	2.833	6.8	128.0				
6-311++G(2df,2p)	SCF	2.837	7.6	120.4	MP2	2.762	7.7	111.0
(98)	CHA-SCF	2.851	7.9	120.5	MP2-CP	2.817	7.7	113.5
	SCF-CP	2.860	7.8	120.8				
6-311++G(3df,2pd)	SCF	2.821	7.0	118.7	MP2	2.749	6.5	112.2
(122)	CHA-SCF	2.840	7.2	119.9	MP2-CP	2.793	7.1	112.2
	SCF-CP	2.846	7.1	119.9				

Concerning the effect of the BSSE on the energetics and geometry of the HF dimer the following conclusions can be drawn. In all cases part of the binding is due to the BSSE, the uncorrected stabilization energies are always larger than the corrected ones. The corrected intermolecular distances (*rff*, the distance between the fluorine atoms) are always longer than the corresponding uncorrected values. This finding is in line with literature data[48]. The difference between the uncorrected and corrected *rff* distances is rather large in the case of small basis sets but gets smaller by using larger basis sets. Analyzing the data presented in Tables I and II it can be seen that even in the case of the 6-311++G(3df, 2pd) basis set the corrected and uncorrected *rff* distances differ from each other by 0.015-0.025, 0.005-0.015, and 0.044-0.055 Å at the SCF, DFT, and MP2 levels, respectively. Except for a few cases when the optimization leads to distorted structures, the uncorrected SCF *rff* distances are longer than the corresponding experimental[101] value of 2.72 Å. Correction for BSSE even lengthens these distances, the corrected parameters are farther from the experimental value than the uncorrected ones. Disregarding the

pathological cases discussed below in detail, DFT methods predict reasonable *rff* distances especially in the case of the B3LYP fuctional.

TABLE II.

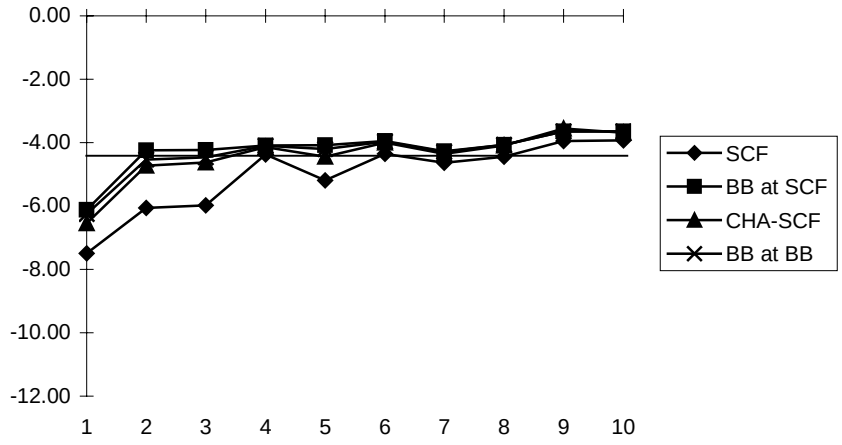
Geometrical parameters for the HF-HF dimer calculated in ten different basis sets at the BLYP, CHA-BLYP, BLYP-CP, B3LYP, CHA-B3LYP and B3LYP-CP levels of theory. The number of basis functions is given in parentheses. The experimental values of *rff* (r_e), α and β are 2.72 Å, 10 ± 6 degrees and 117 ± 6 degrees, respectively. For notation, see SCHEME 7.

Basis Set	Method	rff (Å)	α	β	Method	rff (Å)	α	β
6-31G	BLYP	2.480	46.7	47.2	B3LYP	2.478	48.5	49.8
(22)	CHA-BLYP	2.662	5.6	114.9	CHA-B3LYP	2.647	5.2	120.0
	BLYP-CP	2.735	5.4	120.1	B3LYP-CP	2.701	4.7	125.2
6-31G*	BLYP	2.485	44.9	45.2	B3LYP	2.485	46.0	47.0
(34)	CHA-BLYP	2.652	5.2	105.9	CHA-B3LYP	2.647	5.7	107.5
	BLYP-CP	2.762	7.1	107.0	B3LYP-CP	2.732	6.8	109.5
6-31G**	BLYP	2.494	44.9	44.9	B3LYP	2.493	46.0	46.9
(40)	CHA-BLYP	2.626	7.6	105.2	CHA-B3LYP	2.627	7.8	107.5
	BLYP-CP	2.781	7.2	108.7	B3LYP-CP	2.749	7.0	111.3
6-31++G**	BLYP	2.760	7.5	111.6	B3LYP	2.732	7.6	113.1
(50)	CHA-BLYP	2.782	7.4	110.6	CHA-B3LYP	2.752	7.4	112.5
	BLYP-CP	2.786	7.8	109.9	B3LYP-CP	2.758	7.7	112.0
6-311G**	BLYP	2.572	46.9	48.3	B3LYP	2.567	47.2	51.2
(50)	CHA-BLYP	2.730	8.8	107.9	CHA-B3LYP	2.725	7.0	113.9
	BLYP-CP	2.809	7.2	113.6	B3LYP-CP	2.777	6.9	116.3
6-311++G**	BLYP	2.778	6.9	116.2	B3LYP	2.747	7.3	117.5
(60)	CHA-BLYP	2.822	6.3	115.6	CHA-B3LYP	2.790	7.0	116.9
	BLYP-CP	2.816	8.0	113.1	B3LYP-CP	2.785	7.9	115.5
TZV**	BLYP	2.740	6.6	115.4	B3LYP	2.716	6.5	117.8
(52)	CHA-BLYP	2.783	6.1	116.5	CHA-B3LYP	2.751	6.2	118.8
	BLYP-CP	2.786	6.5	116.5	B3LYP-CP	2.756	6.4	119.0
TZV++**	BLYP	2.764	7.2	117.1	B3LYP	2.734	7.2	119.1
(62)	CHA-BLYP	2.787	6.4	116.0	CHA-B3LYP	2.755	6.7	117.9
	BLYP-CP	2.786	7.8	114.1	B3LYP-CP	2.755	7.7	116.7
6-311++G(2df,2p)	BLYP	2.775	9.0	107.0	B3LYP	2.750	8.1	111.0
(98)	CHA-BLYP	2.784	8.4	107.5	CHA-B3LYP	2.756	6.0	113.7
	BLYP-CP	2.790	7.7	108.9	B3LYP-CP	2.760	7.8	110.9
6-311++G(3df,2pd)	BLYP	2.768	5.3	112.0	B3LYP	2.734	5.8	112.5
(122)	CHA-BLYP	2.778	4.8	113.0	CHA-B3LYP	2.747	5.4	113.9
	BLYP-CP	2.780	6.9	109.1	B3LYP-CP	2.750	7.0	110.9

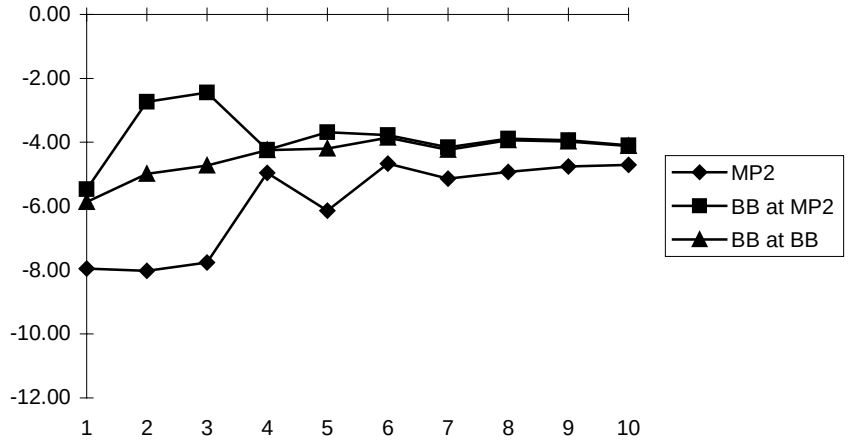
The BSSE has an enormous effect on the *rff* distances determined from MP2 calculations. The uncorrected *rff* values are close to the experimental data. The CP corrected MP2 *rff* distances are always larger than the experimental value, the difference between the corrected and uncorrected MP2/6-311++G(3df,2pd) distances is still 0.044 Å. These results show that BSSE effects are significant in the MP2 calculations even using the largest Pople basis sets.

Furthermore, the BSSE can have smaller or larger effect on the intermolecular bond angles. This effect is quite large using the SCF method in conjunction with small basis sets and results highly distorted uncorrected structures. In these cases correction for the BSSE leads to reasonable geometrical parameters. By the appearance of diffuse functions the corrected and uncorrected bond angles gets very close to each other. In the case of the correlation methods the differences between the corrected and uncorrected

bond angles are larger than the corresponding SCF values even if we do not consider the totally pathological cases which will be discussed later in detail. In the case of the largest Pople basis sets the corrected bond angles can still differ from the uncorrected ones by a few degrees. It is to be noted in this respect that one can not find any tendency even in the case of the CP or CHA surfaces regarding the bond angles. Basis set effects are rather important as well, however, it is not easy to justify which theoretical model is preferable in this respect because of the large uncertainties of the experimental results. Since the intramolecular parameters (F-H distances) are not affected by the BSSE, their actual values are not reported in order to save space.



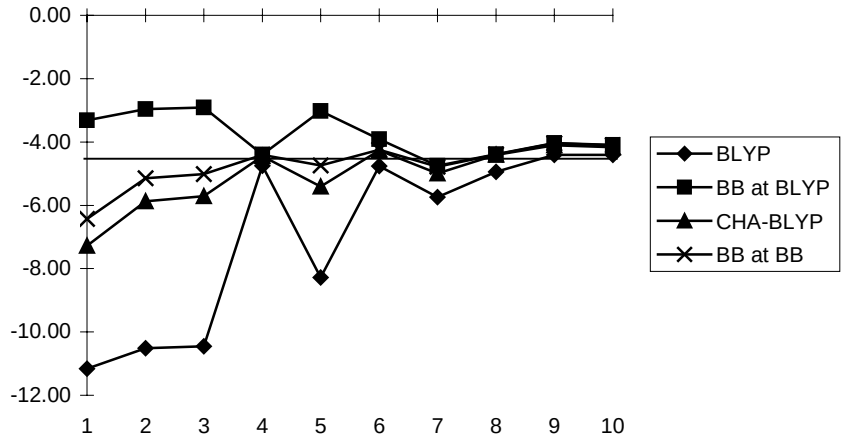
(a)



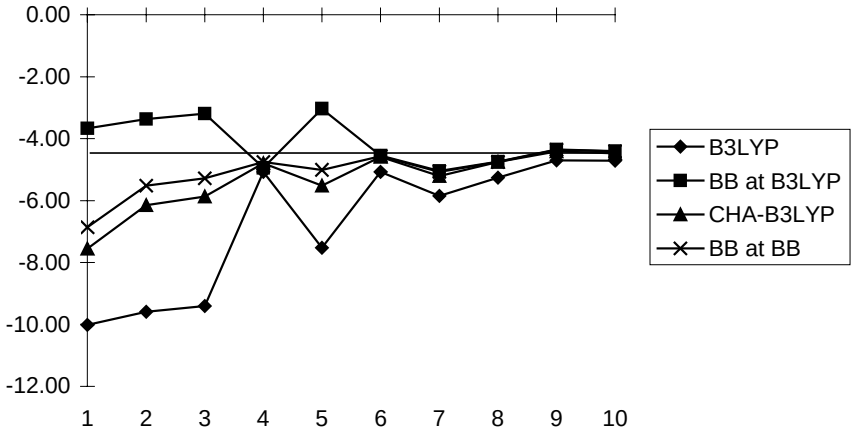
(b)

FIGURE 1: Electronic stabilization energy (kcal/mol) for the $(HF)_2$ in ten basis sets. The experimental stabilization energy is -4.56 ± 0.29 kcal/mol (horizontal solid line).

Comparing the behavior of the *a priori* and *a posteriori* correction schemes, one can find large differences when small or moderate basis sets are used in the computations. By comparing CP and CHA results one can explore the adequacy of the basis sets in a given computational situation.



(c)

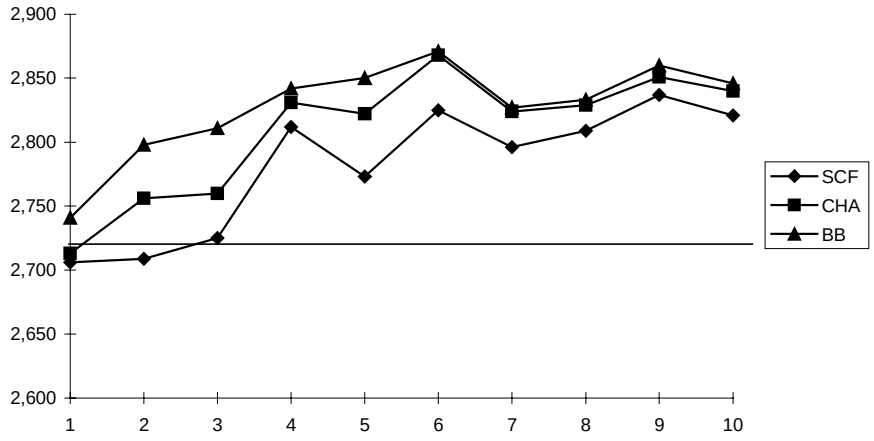


(d)

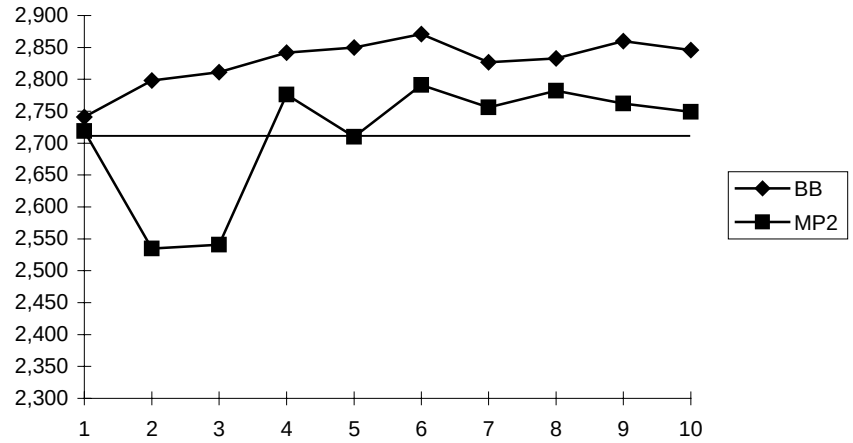
FIGURE 1 (cont.): Electronic stabilization energy (kcal/mol) for the $(HF)_2$ in ten basis sets. The experimental stabilization energy is -4.56 ± 0.29 kcal/mol (horizontal solid line).

Except for a few cases when the CP and CHA values are very close to each other, the CP intermolecular distances (*rff*) are longer than the corresponding CHA values. In the case of small basis sets the difference between the CP and CHA values can be rather large (for example, 0.105 Å at the B3LYP/6-31G(d) level of theory) but the difference gets

smaller as the applied basis set is improved. From this point of view the role of diffuse functions is rather important, by their appearance the difference between the CP and CHA corrected *rff* values gets close to 0.005 Å. The largest deviations between the CP and CHA corrected *rff* values can be found in those cases when the uncorrected model fails to describe even qualitatively the PES of the HF dimer. (For detailed discussion of this problem, see below.) In these cases the CP method shows smoother convergence than that of the CHA.



(a)



(b)

FIGURE 2: F-F distance (Å) for the $(HF)_2$ in ten basis sets. Expt: 2.72 Å (horizontal solid line).

As we mentioned above, the DFT methods in conjunction with small basis sets tend to predict the cyclic structure to be the only stable one on the PES of the HF dimer. This pathological behavior of the DFT methods can be seen also in our results in the case of

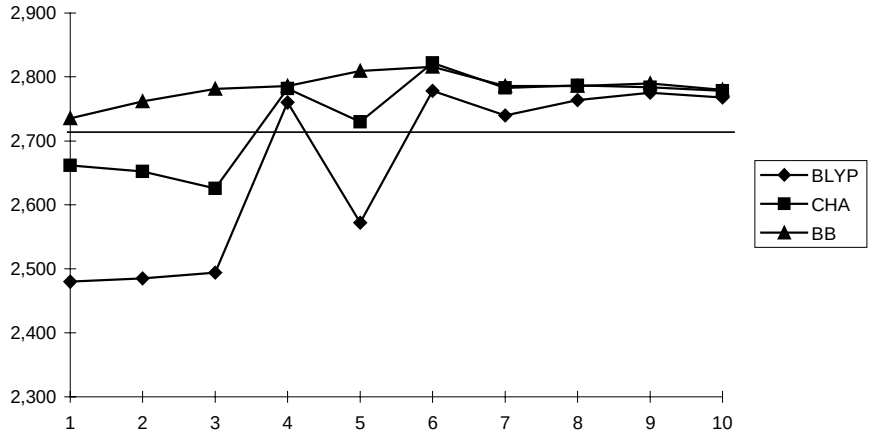
the 6-31G, 6-31G(d), 6-31G(d,p), and even 6-311G(d,p) basis sets. Here, the α and β bond angles are close to 45° and the *rff* distance is shorter than the corresponding experimental value of 2.72 Å. On the other hand, both the CHA and the CP optimization lead to the bent structure in these cases correcting the pitfall of the underlying uncorrected models. One has to note in this respect that this pathological behavior described above is believed in the literature as a problem of DFT methods. However, optimizations at the MP2/6-31G(d) and MP2/6-31G(d,p), levels also lead to the cyclic structure while the MP2/6-31G and MP2/6-311G(d,p) models predict distorted bent structures. Optimizations on the corresponding CP-corrected potential energy surfaces lead to the bent structure and predict (except for MP2/6-31G(d)) reasonable bond angles.

One can calculate the stabilization energy of an intermolecular complex considering both BSSE and geometry effects in a few ways. First, the stabilization energy can be calculated after geometry optimization of the complex without any correction for BSSE. Second, one can calculate the BSSE-corrected stabilization energy at the optimized geometry ("single-point CP, SP"). Furthermore, one can correct (CP or CHA) for the BSSE during the geometry optimization, too, and determine BSSE-corrected stabilization energies using geometries obtained from BSSE-free potential energy surfaces.

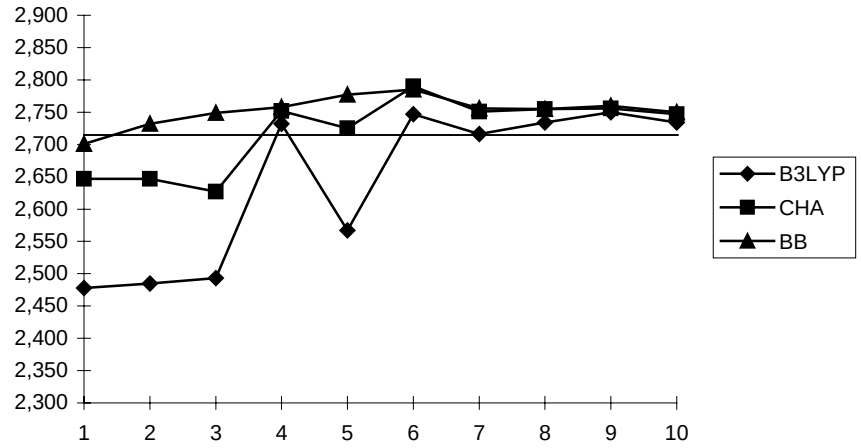
In Figure 1 we show our results for the stabilization energy of the HF dimer obtained by using the four procedures described above. The behavior of the first two procedures is not satisfactory, calculations without any BSSE correction can lead to high stabilization energies while single-point CP corrections at the final (uncorrected) geometries can predict too weak interaction for small basis sets. Of course this behavior is due to the tendency that some of the uncorrected DFT and MP2 optimizations lead to the cyclic structure.

This can be graphically seen in the Figure 2. On the other hand, corrected stabilization energies computed at the BSSE-corrected geometries converge fast, independent whether the CP or CHA method is used in the calculations. It is to be noted that the difference between the CHA and CP stabilization energies is always smaller than the difference between the uncorrected and corrected values. Using basis sets with diffuse functions the single-point CP-corrected energies get closer to the fully corrected ones. The energy data suggest that it is easier to reach convergence in the stabilization energy by improving the quality of the applied basis set than getting close to the limit regarding the

geometric parameters.



(c)



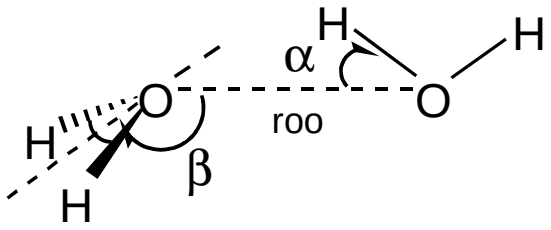
(d)

FIGURE 2 (cont.): F-F distance (Å) for the $(HF)_2$ in ten basis sets. Expt: 2.72 Å (horizontal solid line).

Of course, this is due to the fact that these intermolecular potential surfaces are much flatter than the intramolecular ones. Correction for BSSE seems to be important for both energies and geometries.

II.1.1.1.b Water dimer

The water dimer (SCHEME 8) has a linear structure which is known from both experiment[105] and theoretical calculations[106].



SCHEME 8: Geometrical parameters of the $(H_2O)_2$

Our geometry results are summarized in Tables III and IV while the energetics of the water dimer and the dependence of the *roo* distance on the applied theoretical model are shown graphically in Figures 3 and 4, respectively.

[105] J. A. Odutola and T. R. Dyke, *J. Chem. Phys*, **72**, 5062 (1980)

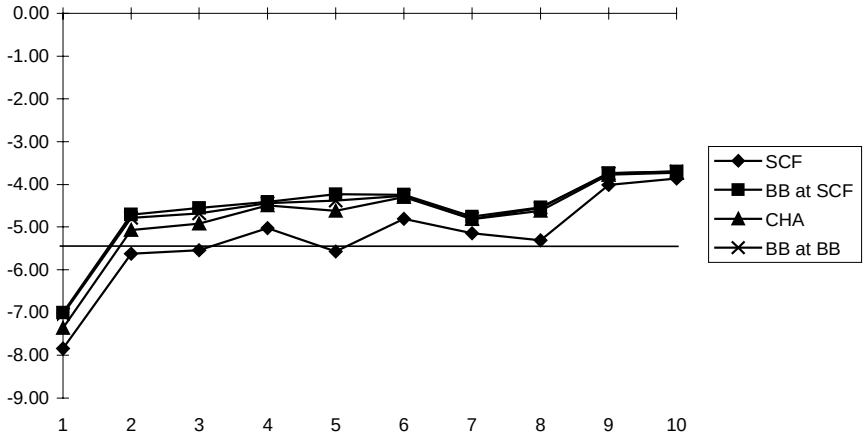
[106] A. Halkier, H. Koch, P. Jorgensen, O. Christianses, I. M. Beck Nielsen, and T. Helgaker, *Theor. Chem. Acc*., **97**, 150 (1997)

TABLE III

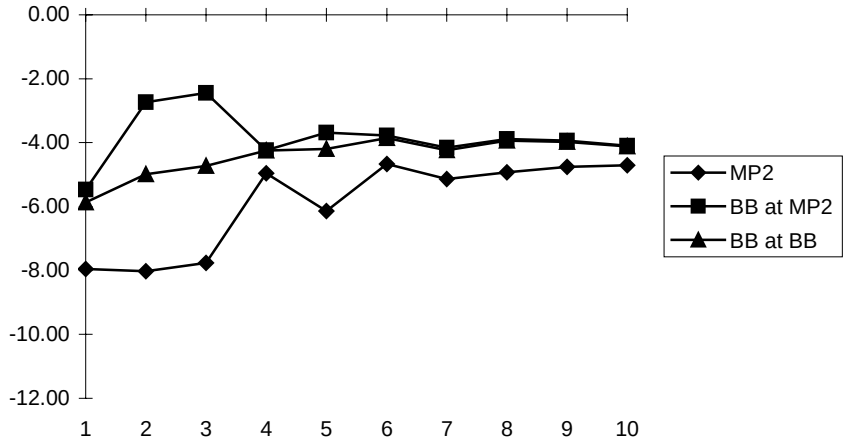
Geometrical parameters for the $(H_2O)_2$ dimer calculated in ten different basis sets at the SCF, CHA-SCF, SCF-CP, MP2 and MP2-CP levels of theory. The number of basis functions is given in parentheses. The experimental values of *roo* (r_e), α and β are 2.946 Å, 123 ± 10 degrees and 2 ± 10 degrees, respectively. For notation, see SCHEME 8.

Basis Set	Method	roo (Å)	α	β	Method	roo (Å)	α	β
6-31G	SCF	2.843	-0.3	152.0	MP2	2.867	2.6	136.8
(26)	CHA-SCF	2.866	0.0	155.9	MP2-CP	2.901	0.8	153.2
	SCF-CP	2.862	0.0	157.5				
6-31G*	SCF	2.973	4.6	118.4	MP2	2.916	9.2	101.5
(38)	CHA-SCF	2.985	2.6	128.4	MP2-CP	2.978	3.0	126.5
	SCF-CP	3.002	1.4	134.1				
6-31G**	SCF	2.983	5.2	117.3	MP2	2.910	9.8	99.3
(50)	CHA-SCF	2.999	2.2	133.1	MP2-CP	2.990	2.8	129.9
	SCF-CP	3.017	1.2	137.2				
6-31++G**	SCF	2.987	2.6	136.2	MP2	2.921	3.5	133.4
(62)	CHA-SCF	3.030	1.5	140.6	MP2-CP	3.007	4.6	129.9
	SCF-CP	3.046	1.7	140.2				
6-311G**	SCF	2.975	2.2	129.5	MP2	2.907	4.3	117.0
(62)	CHA-SCF	3.036	-0.1	142.8	MP2-CP	3.038	-0.1	140.2
	SCF-CP	3.052	-0.7	146.8				
6-311++G**	SCF	3.000	0.9	142.9	MP2	2.922	2.2	135.7
(74)	CHA-SCF	3.048	0.7	146.0	MP2-CP	3.019	3.1	135.7
	SCF-CP	3.049	0.8	146.0				
TZV**	SCF	2.970	2.4	138.2	MP2	2.886	0.0	142.1
(64)	CHA-SCF	3.006	0.4	148.9	MP2-CP	2.987	1.6	144.5
	SCF-CP	3.007	0.2	150.3				
TZV++**	SCF	2.976	0.4	147.2	MP2	2.900	1.6	140.2
(76)	CHA-SCF	2.996	0.9	148.0	MP2-CP	2.984	3.3	139.5
	SCF-CP	3.006	0.9	148.0				
6-311++G(2df,2p)	SCF	3.039	2.7	134.1	MP2	2.919	4.5	126.2
(118)	CHA-SCF	3.056	2.7	137.0	MP2-CP	2.966	5.2	125.7
	SCF-CP	3.060	2.9	136.5				
6-311++G(3df,2pd)	SCF	3.036	2.9	134.2	MP2	2.911	4.8	125.4
(154)	CHA-SCF	3.048	2.7	137.3	MP2-CP	2.950	5.3	125.0
	SCF-CP	3.049	3.0	136.4				

Because many of the conclusions drawn in the preceding section are also valid for the water dimer we do not discuss them in details. Because part of the binding is due to the BSSE, all the corrected intermolecular distances (*roo*) are longer than the corresponding uncorrected ones. The difference between the corrected and uncorrected *roo* distances is similar to those values found in the case of the HF dimer. The SCF *roo* distances are nearly always longer than the corresponding experimental[105] value (2.946 Å) and correction for the BSSE further lengthens this parameter.



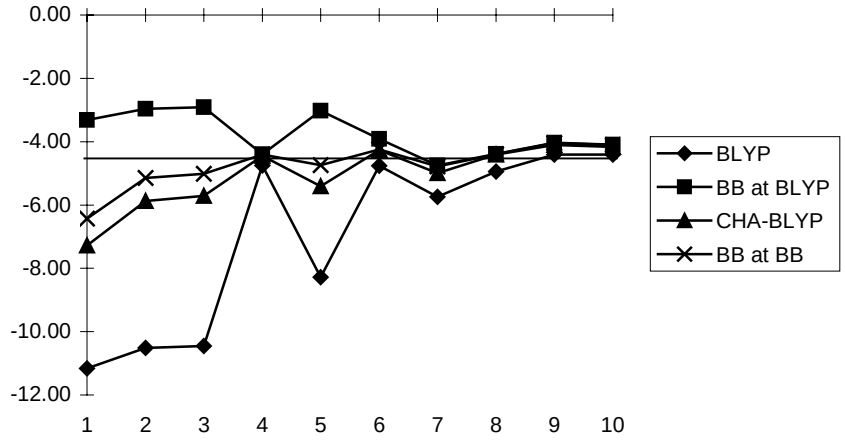
(a)



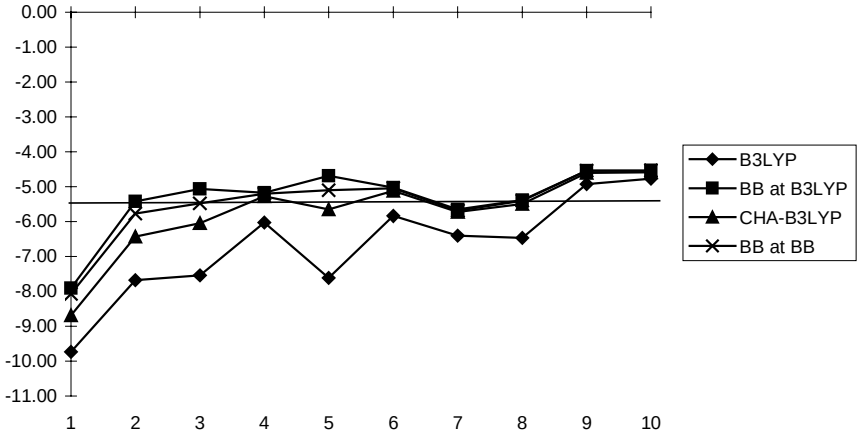
(b)

FIGURE 3: Stabilization energy (kcal/mol) for the $(H_2O)_2$ in ten basis sets Expt: -5.4 ± 0.2 kcal/mol (horizontal solid line).

DFT methods usually underestimate the *roo* distance, the corresponding BSSE-corrected values are closer to experiment (Figure 3, Table IV). The uncorrected MP2 intermolecular distances are also too short, while the CP-corrected values are usually too large. It is very promising, however, that by improving the quality of the basis both the corrected and uncorrected *roo* distances get close to the experimental value.



(c)



(d)

FIGURE 3(cont.): Stabilization energy (kcal/mol) for the $(H_2O)_2$ in ten basis sets Expt: -5.4 ± 0.2 kcal/mol (horizontal solid line).

Both the BSSE and the quality of the basis set have a large effect on the optimized bond angles. For example, nearly all the optimizations at the SCF level overestimate the β bond angle. The corresponding CP and CHA optimizations do not improve these results. In the case of the DFT methods the predicted β bond angles are close to the experimental value when one uses the largest basis sets. However, this is not the case for the small and moderate basis sets but both CP and CHA are able to correct these pitfalls. For example, optimization at the B3LYP/6-31G(d,p) level led to 94.5 degrees, the corresponding CP and CHA values are 114.3 and 119.1 degrees, respectively. The MP2 results also vary on a wide range, one has to use at least moderate size basis sets and corrected models in order to get reasonable bond angles.

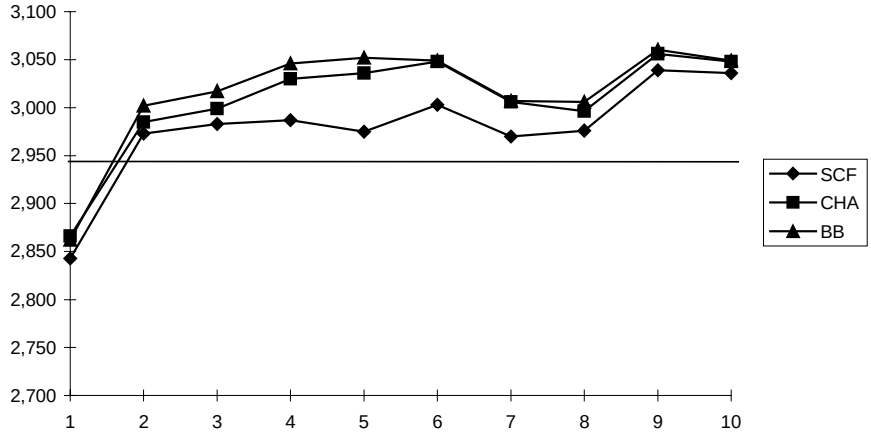
TABLE IV.

Geometrical parameters for the $(H_2O)_2$ dimer calculated in ten different basis sets at the BLYP, CHA-BLYP, BLYP-CP, B3LYP, CHA-B3LYP and B3LYP-CP levels of theory. The number of basis functions is given in parentheses. The experimental values of *roo* (r_e), α and β are 2.946 Å, 123 ± 10 degrees and 2 ± 10 degrees, respectively.

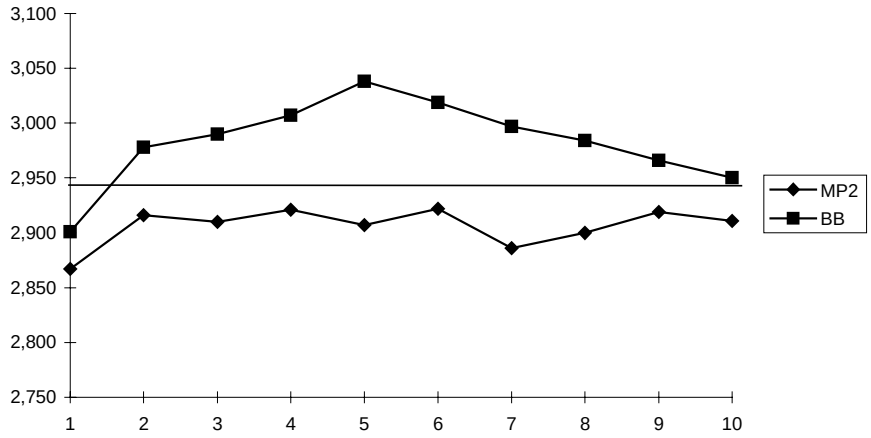
Basis Set	Method	roo (Å)	α	β	Method	roo (Å)	α	β
6-31G	BLYP	2.814	6.1	116.0	B3LYP	2.776	3.7	130.2
(26)	CHA-BLYP	2.823	4.0	132.1	CHA-B3LYP	2.795	3.2	140.5
	BLYP-CP	2.836	2.0	142.2	B3LYP-CP	2.797	1.9	147.7
6-31G*	BLYP	2.868	16.3	82.1	B3LYP	2.861	12.1	93.1
(38)	CHA-BLYP	2.896	4.6	111.7	CHA-B3LYP	2.878	4.9	114.3
	BLYP-CP	2.941	4.6	115.4	B3LYP-CP	2.911	4.4	119.1
6-31G**	BLYP	2.884	15.2	83.3	B3LYP	2.876	10.9	94.5
(50)	CHA-BLYP	2.901	5.5	114.0	CHA-B3LYP	2.883	5.6	116.5
	BLYP-CP	2.970	4.5	118.0	B3LYP-CP	2.936	3.9	122.4
6-31++G**	BLYP	2.912	4.2	127.3	B3LYP	2.887	4.2	128.8
(62)	CHA-BLYP	2.955	3.5	127.1	CHA-B3LYP	2.924	3.7	128.9
	BLYP-CP	2.961	6.2	119.8	B3LYP-CP	2.931	5.7	123.4
6-311G**	BLYP	2.915	7.6	100.5	B3LYP	2.887	8.3	105.1
(62)	CHA-BLYP	2.960	2.9	123.2	CHA-B3LYP	2.941	2.8	127.3
	BLYP-CP	3.021	1.5	131.1	B3LYP-CP	2.979	1.1	135.3
6-311++G**	BLYP	2.927	3.9	129.5	B3LYP	2.900	3.8	131.3
(74)	CHA-BLYP	2.975	3.5	128.3	CHA-B3LYP	2.944	3.5	131.1
	BLYP-CP	2.971	5.3	124.8	B3LYP-CP	2.941	4.6	129.5
TZV**	BLYP	2.900	1.7	130.4	B3LYP	2.877	1.0	136.5
(64)	CHA-BLYP	2.955	3.7	129.7	CHA-B3LYP	2.920	3.7	133.1
	BLYP-CP	2.959	3.6	132.8	B3LYP-CP	2.925	2.9	137.6
TZV++**	BLYP	2.910	3.4	131.2	B3LYP	2.883	2.5	136.1
(76)	CHA-BLYP	2.935	3.8	129.7	CHA-B3LYP	2.905	3.7	132.4
	BLYP-CP	2.942	5.3	127.2	B3LYP-CP	2.911	4.8	131.3
6-311++G(2df,2p)	BLYP	2.949	5.3	120.4	B3LYP	2.922	4.1	125.9
(118)	CHA-BLYP	2.960	5.8	118.4	CHA-B3LYP	2.931	4.2	125.6
	BLYP-CP	2.966	6.0	118.7	B3LYP-CP	2.937	5.5	122.3
6-311++G(3df,2pd)	BLYP	2.949	5.7	119.0	B3LYP	2.919	5.1	122.3
(154)	CHA-BLYP	2.957	5.7	119.5	CHA-B3LYP	2.928	4.0	126.7
	BLYP-CP	2.960	5.7	119.7	B3LYP-CP	2.931	5.2	123.2

Comparing the behavior of the *a priori* and *a posteriori* correction schemes one can draw similar conclusions as those obtained for the HF dimer. Except for a very few cases the CP method overcompensates the BSSE with respect to the CHA. values Usually, the differences between the CP and CHA intermolecular parameters are smaller than the differences between the corrected and uncorrected values. Larger differences between the *a priori* and *a posteriori* corrected *roo* distances can be found in those cases when the corresponding uncorrected model predicts distorted structures. As an overall conclusion on the behavior of the CHA and CP correction schemes one can argue that when correction for the BSSE is important both the *a priori* and *a posteriori* methods do the same job at least in a qualitative manner.

Previous studies showed[104] that the linear structures determined at the DFT level with BLYP and B3LYP functionals in conjunction with the 6-31G* and 6-31G** basis sets actually correspond to saddle points. Table IV presents the geometry data calculated at the DFT level with both functionals.



(a)

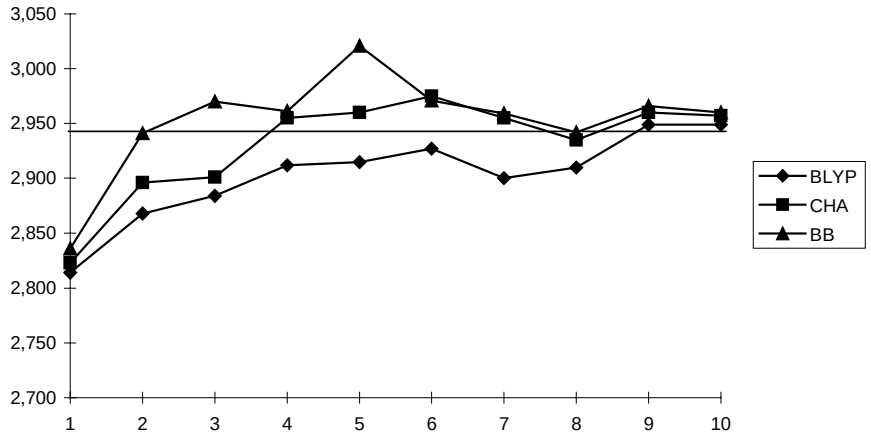


(b)

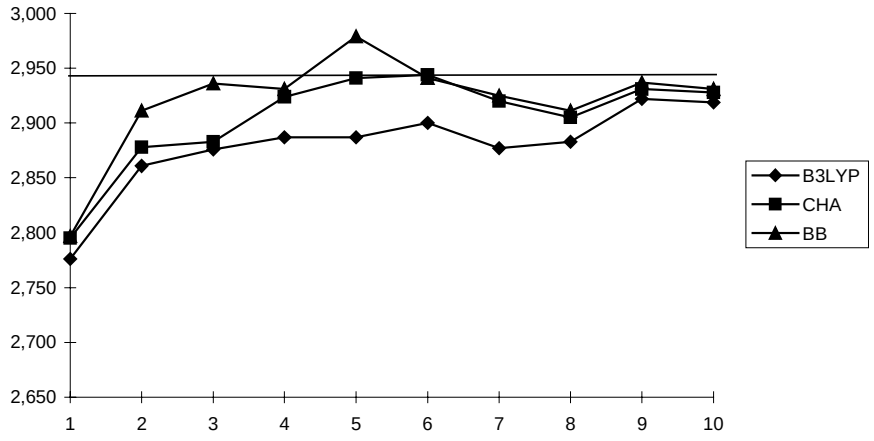
FIGURE 4: O-O distance (Å) in ten basis sets. Expt: 2.946 Å (horizontal solid line).

It can be seen that both BSSE-free methods correct the rather distorted uncorrected structures, the angular features of these structures are closer to the large basis results than those of the uncorrected ones. We carried out frequency calculations at the corrected and uncorrected BLYP/6-31G(d,p) levels to explore the curvature of the PES from the BSSE point of view. Similarly to the literature data the Hessian calculated at the uncorrected BLYP/6-31G(d,p) level has a negative eigenvalue. The CHA/DFT(BLYP) optimized structure was characterized as a minimum by calculating the Hessian using numerical derivatives. The adequacy of the numerical differentiation was checked by comparing the accuracy of full numerical and full analytical methods for the case of the uncorrected BLYP/6-31G(d,p) level. Furthermore, analytical CP corrected derivatives were also used to carry out frequency calculation at the CP corrected structure at the same level. The CP

calculations also result a minimum, so the CP method can also correct the deficiency of the BLYP/6-31G(d,p) model discussed above. Similar calculations were performed for other pathological cases listed above resulting the same tendency obtained for the BLYP/6-31G(d,p) model.



(c)



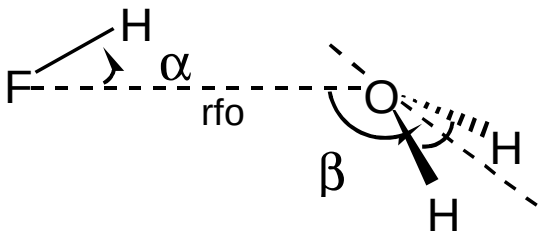
(d)

FIGURE 4(cont.): O-O distance (Å) in ten basis sets. Expt: 2.946 Å (horizontal solid line).

We have already mentioned that some of the uncorrected DFT and MP2 models resulted the cyclic structure of the HF dimer. To check the performance of the MP2 method in the case of the water dimer we carried out frequency calculations using the 6-31G(d) and 6-31G(d,p) basis sets. Both the MP2/6-31G(d) and MP2/6-31G(d,p) calculations lead to a minimum, indicating that the small basis MP2 models perform better than the corresponding DFT ones from the PES curvature point of view.

II.1.1.1.c HF-H_2O complex

The HF-H_2O complex is a rather special system, the determination of its structure is a challenge for both experimentalists and theoreticians. Theoretical studies showed[107] that one can find stationary points on the HF-H_2O PES with both C_s and C_{2v} symmetries. (In the case of the C_{2v} structure the α and β angles (SCHEME 9) are 0 and 180 degrees, respectively.)



SCHEME 9: Geometrical parameters of the HF-H_2O complex

The C_s structure is more stable than the C_{2v} one, the energy difference between them is 0.1 and 0.5 kcal/mol at the SCF and MP2 levels, respectively. The experimental[108] estimate of the barrier is 0.4 kcal/mol. Based on these data one can understand that the experimental determination of the angular features of the HF-H_2O complex is not a trivial task. The barrier is very close to the first vibrational level, most probably the two C_s structures rapidly interconvert to each other and the underlying double well potential can not be measured.

[107] M. M. Szczesniak, S. Scheiner, and Y. Bouteiler, *J. Chem. Phys*, **81**, 5024 (1984)

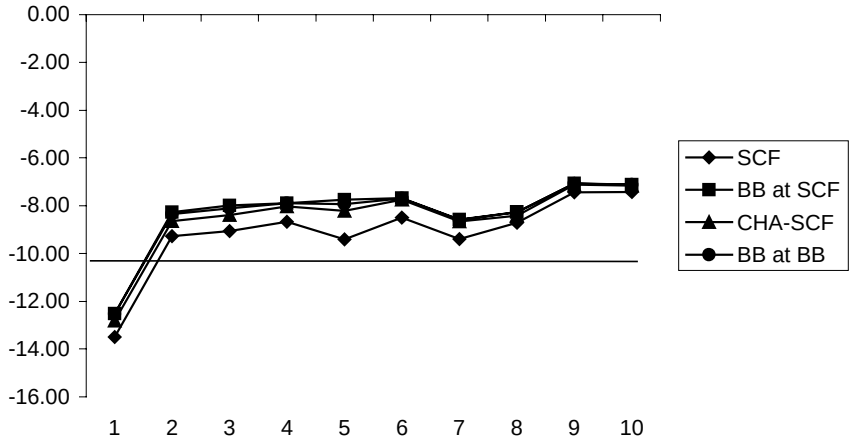
[108] a) A.C. Legon and D. J. Millen, *Faraday Discuss. Chem. Soc.*, **73**, 71 (1982), b) J.W. Bevan, Z. Kisiel, A.C. Legon, D.J. Millen and S.C. Rogers, *Proc. Roy. Soc. A*, **372**, 441 (1980)

TABLE V.

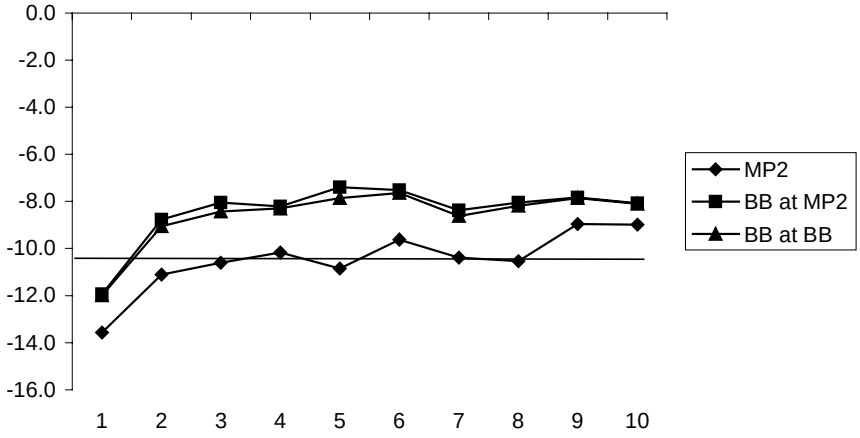
Geometrical parameters for the HF-H_2O dimer calculated in ten different basis sets at the SCF, CHA-SCF, SCF-CP, MP2 and MP2-CP levels of theory. The number of basis functions is given in parentheses. The experimental value of *rfo* (r_0) is 2.66 Å. For notation, see SCHEME 9.

Basis Set	Method	rfo (Å)	α	β	Method	rfo (Å)	α	β
6-31G	SCF	2.616	0.0	180.0	MP2	2.646	0.0	180.0
(24)	CHA-SCF	2.657	0.0	180.0	MP2-CP	2.686	0.0	180.0
	SCF-CP	2.641	0.0	180.0				
6-31G*	SCF	2.719	4.7	132.3	MP2	2.689	7.1	116.9
(36)	CHA-SCF	2.749	2.7	140.8	MP2-CP	2.748	2.5	135.8
	SCF-CP	2.753	2.1	146.3				
6-31G**	SCF	2.718	4.0	135.8	MP2	2.675	6.3	117.6
(45)	CHA-SCF	2.745	2.2	147.4	MP2-CP	2.750	2.2	135.8
	SCF-CP	2.761	1.7	152.6				
6-31++G**	SCF	2.713	1.6	151.7	MP2	2.659	1.7	138.4
(56)	CHA-SCF	2.741	1.1	158.1	MP2-CP	2.727	1.7	140.6
	SCF-CP	2.753	1.1	159.0				
6-311G**	SCF	2.700	2.7	145.4	MP2	2.645	4.1	128.7
(56)	CHA-SCF	2.752	0.5	170.9	MP2-CP	2.762	1.0	152.5
	SCF-CP	2.759	0.0	180.0				
6-311++G**	SCF	2.720	0.9	162.5	MP2	2.664	1.4	138.9
(67)	CHA-SCF	2.751	0.0	180.0	MP2-CP	2.739	1.4	149.1
	SCF-CP	2.750	0.0	180.0				
TZV**	SCF	2.695	0.0	180.0	MP2	2.637	1.4	146.3
(58)	CHA-SCF	2.712	0.0	180.0	MP2-CP	2.727	0.3	172.0
	SCF-CP	2.725	0.0	180.0				
TZV++**	SCF	2.690	0.0	180.0	MP2	2.648	0.9	145.0
(69)	CHA-SCF	2.717	0.0	180.0	MP2-CP	2.716	0.8	163.9
	SCF-CP	2.718	0.0	180.0				
6-311++G(2df,2p)	SCF	2.725	1.3	147.0	MP2	2.661	1.3	133.6
(108)	CHA-SCF	2.733	1.1	150.6	MP2-CP	2.693	1.4	134.9
	SCF-CP	2.737	1.0	153.7				
6-311++G(3df,2pd)	SCF	2.709	1.2	146.1	MP2	2.635	1.1	134.2
(138)	CHA-SCF	2.724	1.1	149.8	MP2-CP	2.671	1.4	133.2
	SCF-CP	2.725	1.1	150.0				

Our main reason for investigating the C_s structure of the HF-H_2O complex was to present a system which, up to our knowledge, is free from DFT or MP2 pitfalls described in the previous chapters for the cases of the HF and H_2O dimers, respectively. In our opinion this pitfall-free status of the HF-H_2O complex can be connected with the strength of the interaction in this system. As we go farther from the HF dimer to the H_2O dimer and to the HF-H_2O complex the stabilization energies become larger.

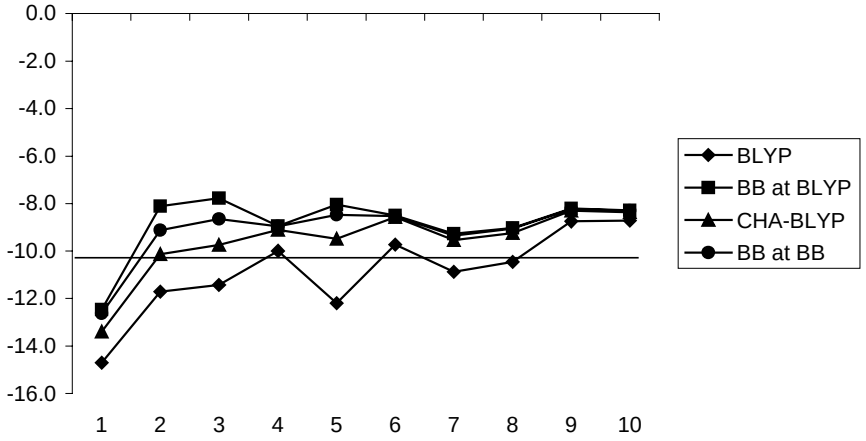


(a)

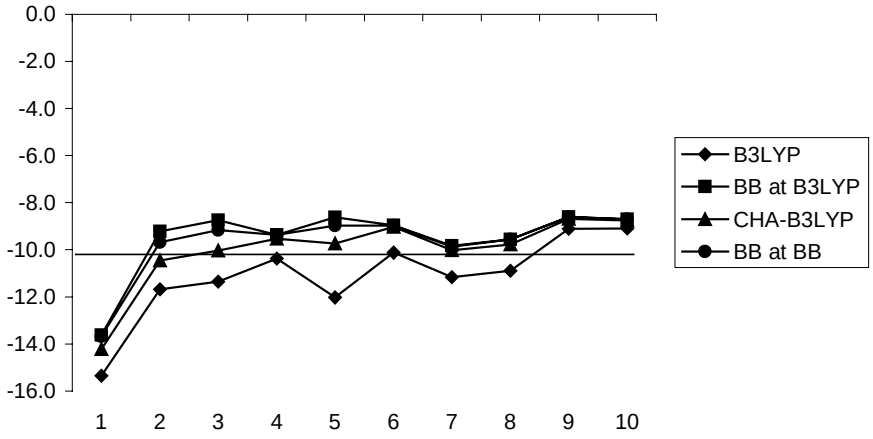


(b)

FIGURE 5: Stabilization energy (kcal/mol) in ten basis sets. Expt: -10.2 kcal/mol (horizontal solid line).



(c)



(d)

FIGURE 5(cont.): Stabilization energy (kcal/mol) in ten basis sets. Expt: -10.2 kcal/mol (horizontal solid line)

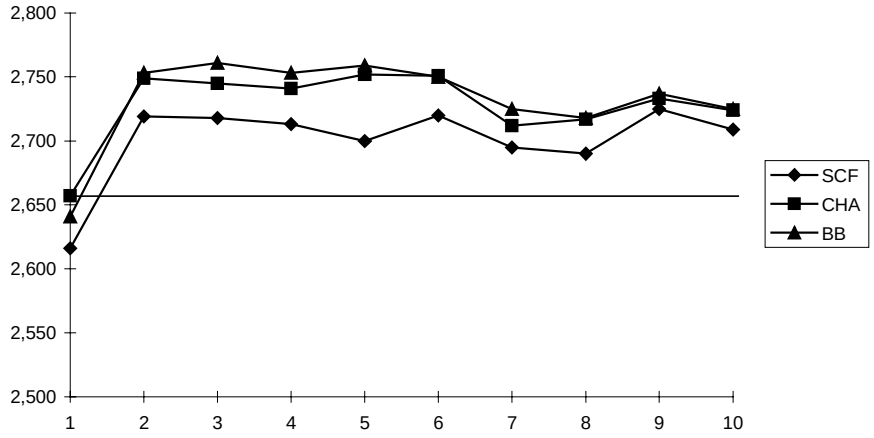
In the case of the HF dimer the DFT and MP2 methods using small basis sets can totally fail to describe the PES. In the case of the water dimer we found some problems with respect to the curvature of the PES. For the HF-H_2O complex, disregarding the smallest 6-31G basis set, the C_s structure is predicted to be the more stable by all the DFT and MP2 models.

TABLE VI.

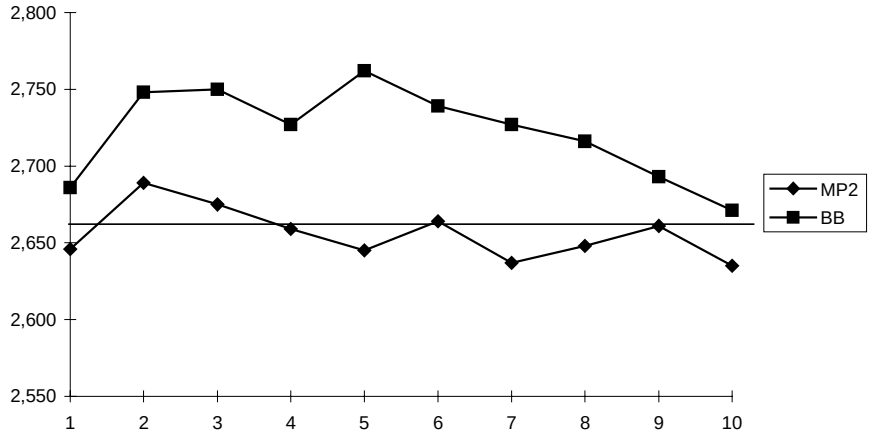
Geometrical parameters for the HF-H_2O dimer calculated in ten different basis sets at the BLYP, CHA-BLYP, BLYP-CP, B3LYP, CHA-B3LYP and B3LYP-CP levels of theory. The number of basis functions is given in parentheses. The experimental value of *rfo* (r_0) is *2.66* Å.

Basis Set	Method	rfo (Å)	α	β	Method	rfo (Å)	α	β
6-31G	BLYP	2.610	2.8	149.4	B3LYP	2.576	0.0	180.0
(24)	CHA-BLYP	2.642	0.0	180.0	CHA-B3LYP	2.619	0.0	180.0
	BLYP-CP	2.638	0.0	180.0	B3LYP-CP	2.608	0.0	180.0
6-31G*	BLYP	2.652	15.2	92.2	B3LYP	2.650	10.4	106.0
(36)	CHA-BLYP	2.701	3.3	119.2	CHA-B3LYP	2.682	3.2	123.9
	BLYP-CP	2.730	3.2	124.2	B3LYP-CP	2.701	3.0	128.5
6-31G**	BLYP	2.663	13.1	96.2	B3LYP	2.652	8.7	109.5
(45)	CHA-BLYP	2.678	3.1	122.1	CHA-B3LYP	2.660	3.3	126.1
	BLYP-CP	2.733	2.9	127.3	B3LYP-CP	2.700	2.7	132.4
6-31++G**	BLYP	2.642	2.6	129.4	B3LYP	2.623	2.4	133.8
(56)	CHA-BLYP	2.672	2.6	128.9	CHA-B3LYP	2.650	2.0	135.4
	BLYP-CP	2.665	2.3	129.3	B3LYP-CP	2.645	2.2	133.9
6-311G**	BLYP	2.657	7.3	113.2	B3LYP	2.636	6.4	119.5
(56)	CHA-BLYP	2.716	3.2	126.9	CHA-B3LYP	2.690	2.4	135.3
	BLYP-CP	2.744	2.4	135.8	B3LYP-CP	2.708	1.5	145.2
6-311++G**	BLYP	2.668	2.3	132.4	B3LYP	2.645	2.0	137.6
(67)	CHA-BLYP	2.689	1.8	133.0	CHA-B3LYP	2.663	1.6	139.4
	BLYP-CP	2.691	2.5	131.6	B3LYP-CP	2.667	1.8	141.0
TZV**	BLYP	2.648	2.3	135.3	B3LYP	2.631	1.2	146.4
(58)	CHA-BLYP	2.682	1.0	145.8	CHA-B3LYP	2.654	1.3	149.1
	BLYP-CP	2.694	1.3	144.4	B3LYP-CP	2.666	1.3	150.8
TZV++**	BLYP	2.654	1.5	138.2	B3LYP	2.625	1.2	146.7
(69)	CHA-BLYP	2.662	0.8	143.8	CHA-B3LYP	2.640	1.2	148.1
	BLYP-CP	2.675	2.7	136.7	B3LYP-CP	2.652	1.7	146.4
6-311++G(2df,2p)	BLYP	2.664	2.1	127.0	B3LYP	2.640	1.9	131.1
(108)	CHA-BLYP	2.675	2.0	126.7	CHA-B3LYP	2.650	1.7	131.4
	BLYP-CP	2.672	1.8	127.4	B3LYP-CP	2.650	1.9	131.3
6-311++G(3df,2pd)	BLYP	2.655	1.9	126.5	B3LYP	2.632	1.8	130.9
(138)	CHA-BLYP	2.667	1.9	126.8	CHA-B3LYP	2.644	1.7	131.4
	BLYP-CP	2.666	1.8	127.5	B3LYP-CP	2.642	1.8	131.1

In most of the cases when the underlying uncorrected DFT and MP2 methods predict wrong α and β parameters, the corrected angular features are close to the large basis results. Concerning the SCF results one has to mention that in some cases we could find only the C_{2v} structure on the PES. Correction for the BSSE does not change this situation.



(a)

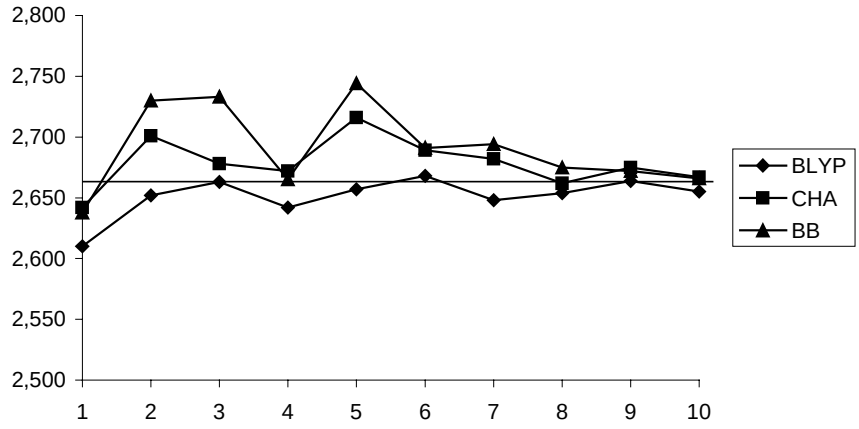


(b)

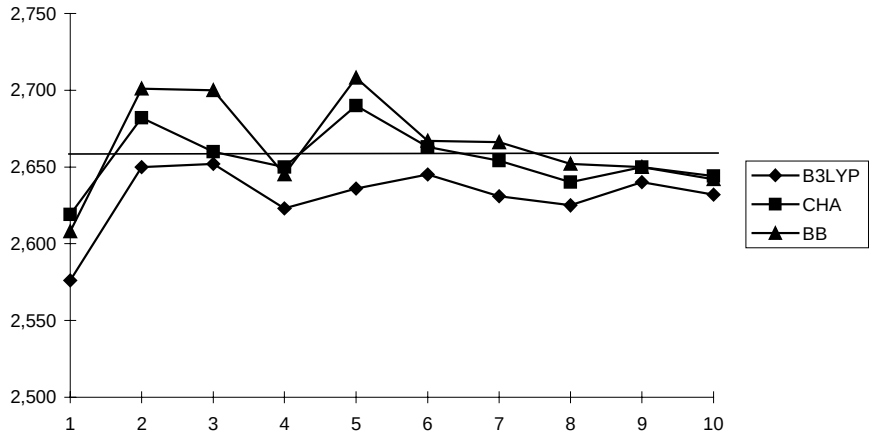
FIGURE 6: F-O distance (Å) in ten basis sets. Expt: 2.66 Å (horizontal solid line).

Turning to the *rof* geometry parameter one can draw the following conclusions. Because part of the binding is due to the BSSE, the corrected *rof* distances are always longer than the uncorrected ones. The SCF *rof* distances are nearly always longer than the experimental[109] value of 2.66 Å. Correction for the BSSE further lengthens this parameter. The uncorrected DFT *rof* distances are surprisingly close to the experimental value. In the case of small basis sets the corrected *rof* distances are farther from the experimental value than the uncorrected ones but for the largest basis sets they get close to each other.

[109] A. C. Legon, D. J. Millen, and H. M. North, *Chem. Phys. Lett.*, **135**, 303 (1987)



(c)



(d)

FIGURE 6(cont.): F-O distance (Å) in ten basis sets. Expt: 2.66 Å (horizontal solid line).

The largest BSSE effects can be found again in the case of the MP2 geometries. The uncorrected *rof* distances usually closer to the experimental value than that of predicted by the CP corrected MP2 model. However, the convergence of the *rof* distance on the corrected PESs is smoother than for the uncorrected PESs.

Comparing the behavior of the CHA and CP correction methods one can find again that CP tends to predict larger BSSE than CHA in the case of small basis sets. However, if balanced basis sets are used for the calculations, both the CP and CHA methods yield almost undistinguishable results.

II.1.1.1.d Structures optimized using Dunning's aug-cc-pVDZ and aug-cc-pVTZ basis sets

Our primary goal was to carefully investigate the effect of BSSE on the energetics and structure of intermolecular complexes when only relatively small basis sets are applied in the calculations. However, it was suggested to include results obtained from calculations using Dunning's correlation consistent aug-cc-pVDZ and aug-cc-pVTZ basis sets[110]. In the following we briefly summarize our findings obtained at the SCF, BLYP, B3LYP, and MP2 levels in conjunction with the aug-cc-pVDZ and aug-cc-pVTZ basis sets on the uncorrected and BSSE-corrected potential energy surfaces. The corresponding energetical and structural data are listed in Tables VII-X.

The correlation consistent [(aug)-cc-p(C)V*X*Z, *X* = 2(D), 3(T), 4(Q), …)] basis sets systematically extend the atomic radial and angular spaces as a function of the cardinal number *X*. These basis sets were designed to be applied in conjunction with traditional correlation methods where the electron-electron cusp is explicitly described. Because of this feature one expects that the BSSE content of various properties are small using the aug-cc-pVDZ and aug-cc-pVTZ basis sets at the HF and DFT levels. Going further to largest cardinal numbers the effect of the BSSE on structural and energetical parameters calculated at the HF and DFT levels should decrease enormously. This is not the case, however, for traditional correlation methods like MP2 where the BSSE content of the investigated properties decreases as the cardinal number of the applied basis set is enlarged but it is practically impossible to reach that status when the effect of BSSE becomes negligible.

[110] T. H. Dunning Jr., *J. Chem. Phys.* **90**, 1007 (1989); R. A. Kendall, T. H. Dunning Jr. and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992); D. E. Woon and T. H. Dunning Jr., *J. Chem. Phys.* **98**, 1358 (1993).

TABLE VII.

Geometrical parameters for the HF-HF dimer calculated using the aug-cc-pVDZ and aug-cc-pVTZ basis sets at the SCF, CHA-SCF, SCF-CP, MP2, MP2-CP, BLYP, CHA-BLYP, BLYP-CP, B3LYP, CHA-B3LYP and B3LYP-CP levels of theory. The number of basis functions is given in parentheses.

Basis Set	Method	rff	α	β	Method	rff	α	β
aug-cc-pVDZ	SCF	2.826	6.7	118.7	MP2	2.751	6.2	110.6
(68)	CHA-SCF	2.854	7.2	119.0	MP2-CP	2.812	6.9	111.6
	SCF-CP	2.846	7.1	118.6				
aug-cc-pVDZ	BLYP	2.761	5.9	109.1	B3LYP	2.734	6.1	110.6
(68)	CHA-BLYP	2.805	8.9	107.0	CHA-B3LYP	2.774	8.1	109.7
	BLYP-CP	2.772	6.2	109.2	B3LYP-CP	2.745	6.5	110.8
aug-cc-pVTZ	SCF	2.825	6.2	121.0	MP2	2.739	5.5	112.8
(160)	CHA-SCF	2.831	6.7	120.2	MP2-CP	2.764	6.6	111.8
	SCF-CP	2.828	6.5	120.4				
aug-cc-pVTZ	BLYP	2.755	5.8	109.9	B3LYP	2.727	6.1	111.4
(160)	CHA-BLYP	2.773	4.4	112.7	CHA-B3LYP	2.740	5.0	113.6
	BLYP-CP	2.759	6.6	108.7	B3LYP-CP	2.730	6.5	110.8

In the light of these facts one can easily rationalize most of our data obtained using the aug-cc-pVDZ and aug-cc-pVTZ basis sets. All the BSSE-free intermolecular distances are longer than the corresponding uncorrected ones using the aug-cc-pVDZ basis set. However, contrary to the tendency observed for the Pople basis sets, the CHA distances are longer than the CP ones. The HF and DFT intermolecular distances corrected using the CP algorithm are usually very close to the corresponding uncorrected values, see for example the actual *rfo* values, 2.643 and 2.647 Å (HF-H_2O complex, Table IX.) obtained at the B3LYP/aug-cc-pVDZ level on the plain and CP-corrected PESs, respectively. This behavior is a little bit surprising since the aug-cc-pVDZ basis set is the first and smallest member of the aug-cc-pVXZ family with weaker performance compared to that of the larger (aug-cc-pVTZ, aug-cc-pVQZ, etc.) Dunning basis sets. The CHA distances obtained at the DFT and HF levels are more reasonable in this respect and are in accordance with the size of the actual basis set.

TABLE VIII.

Geometrical parameters for the $(H_2O)_2$ complex calculated using the aug-cc-pVDZ and aug-cc-pVTZ basis sets at the SCF, CHA-SCF, SCF-CP, MP2, MP2-CP, BLYP, CHA-BLYP, BLYP-CP, B3LYP, CHA-B3LYP and B3LYP-CP levels of theory. The number of basis functions is given in parentheses.

Basis Set	Method	roo	α	β	Method	roo	α	β
aug-cc-pVDZ	SCF	3.037	4.0	130.9	MP2	2.921	6.4	119.6
(86)	CHA-SCF	3.062	3.3	132.8	MP2-BB	2.978	5.7	122.5
	SCF-BB	3.053	3.3	133.0				
aug-cc-pVDZ	BLYP	2.950	6.0	118.0	B3LYP	2.920	5.8	120.7
(86)	CHA-BLYP	2.995	3.5	126.1	CHA-B3LYP	2.961	3.7	127.8
	BLYP-BB	2.958	5.7	119.0	B3LYP-BB	2.929	5.3	122.1
aug-cc-pVTZ	SCF	3.038	2.8	137.7	MP2	2.902	4.7	125.6
(210)	CHA-SCF	3.044	3.0	138.0	MP2-BB	2.930	5.4	125.1
	SCF-BB	3.041	2.9	137.9				
aug-cc-pVTZ	BLYP	2.945	5.0	121.0	B3LYP	2.916	4.9	123.8
(210)	CHA-BLYP	2.992	8.0	118.0	CHA-B3LYP	-	-	-
	BLYP-BB	2.951	5.0	120.3	B3LYP-BB	2.920	5.3	123.5

Our results obtained at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels are consistent from the point of view of BSSE and reproduce the available literature data.[39-41] (The small differences between our and the literature data are due to the fact that because of program limitations we used Cartesian d and f functions in our calculations.) The BSSE-content of the investigated properties decreases very much using the larger basis set, however, the MP2/aug-cc-pVTZ data are still significantly polluted by the BSSE.

TABLE IX.

Geometrical parameters for the HF-H_2O complex calculated using the aug-cc-pVDZ and aug-cc-pVTZ basis sets at the SCF, CHA-SCF, SCF-CP, MP2, MP2-CP, BLYP, CHA-BLYP, BLYP-CP, B3LYP, CHA-B3LYP and B3LYP-CP levels of theory. The number of basis functions is given in parentheses.

Basis Set	Method	rfo	α	β	Method	rfo	α	β
aug-cc-pVDZ	SCF	2.725	1.5	142.8	MP2	2.660	1.6	127.7
(77)	CHA-SCF	2.760	1.2	149.2	MP2-BB	2.702	1.5	130.3
	SCF-BB	2.734	1.4	144.0				
aug-cc-pVDZ	BLYP	2.664	1.6	126.3	B3LYP	2.643	1.6	130.1
(77)	CHA-BLYP	2.684	2.5	128.6	CHA-B3LYP	2.664	1.6	135.1
	BLYP-BB	2.668	1.7	126.0	B3LYP-BB	2.647	1.7	129.8
aug-cc-pVTZ	SCF	2.716	1.1	149.5	MP2	2.640	1.3	131.1
(185)	CHA-SCF	2.731	1.1	152.3	MP2-BB	2.662	1.4	132.3
	SCF-BB	2.718	1.1	149.8				
aug-cc-pVTZ	BLYP	2.658	1.6	126.4	B3LYP	2.636	1.6	130.9
(185)	CHA-BLYP	-	-	-	CHA-B3LYP	-	-	-
	BLYP-BB	2.662	1.6	127.1	B3LYP-BB	2.639	1.6	131.2

Our HF and especially the DFT calculations using the aug-cc-pVTZ basis set led to some unexpected results. The difference between the CP-corrected and the corresponding uncorrected parameters is usually smaller than what we found in the case of the aug-cc-pVDZ basis set. For example, the actual values of *rfo* are 2.716 and 2.718 Å (HF-H_2O complex, Table IX.) calculated at the HF/aug-cc-pVTZ level on the uncorrected and CP-corrected PESs, respectively. The CHA method performs better in this respect whenever we could locate minima on the corresponding PESs. For example, the actual value of *rfo* is 2.731 Å obtained at the CHA-SCF/aug-cc-pVTZ level of theory. However, in three cases (water dimer at the B3LYP/aug-cc-pVTZ and the HF-H_2O complex at the BLYP/aug-cc-pVTZ and B3LYP/aug-cc-pVTZ levels) we could not optimize the geometry of the investigated complexes at the CHA levels. Beside improving our implementation of CHA (application of better convergence accelerators) we tried many numerical tricks to obtain self-consistency of the applied CHA models. For example, the geometry optimizations were started at large intermolecular distances allowing the optimizer to get slowly close to the particular minimum. At large intermolecular distances the CHA wave functions were easily obtained but we found serious convergence problems getting close to the desired minima again. In the next step we removed that

component of the applied aug-cc-pVTZ basis set which bears the smallest exponent. In this case our convergence problems disappeared, we could easily obtain optimized geometries using the modified basis set. Somewhat similar problems have been met at small distances by Valiron and Mayer[111], too.

TABLE X.

Uncorrected, CHA-corrected and Counterpoise-corrected stabilization energies (Kcal/mol) of $(HF)_2$, $(H_2O)_2$ and HF-H_2O dimers calculated using the aug-cc-pVDZ and aug-cc-pVTZ (in parenthesis) basis sets.

HF-HF			**H_2O-H_2O**			**HF-H_2O**		
Method	**Interaction Energy (Kcal/mol)**		**Method**	**Interaction Energy (Kcal/mol)**		**Method**	**Interaction Energy (Kcal/mol)**	
SCF	3.82	(3.72)	SCF	3.86	(3.74)	SCF	7.31	(7.25)
CHA-SCF	3.68	(3.65)	CHA-SCF	3.71	(3.68)	CHA-SCF	6.98	(7.15)
SCF-BB	3.67	(3.65)	SCF-BB	3.71	3.69)	SCF-BB	7.10	(7.18)
BLYP	4.23	(4.25)	BLYP	4.24	(4.27)	BLYP	8.49	(8.54)
CHA-BLYP	3.87	(4.06)	CHA-BLYP	3.89	(4.10)	CHA-BLYP	7.86	(----)
BLYP-BB	4.04	(4.13)	BLYP-BB	4.06	(4.15)	BLYP-BB	8.26	(8.37)
B3LYP	4.56	(4.56)	B3LYP	4.64	(4.64)	B3LYP	8.89	(8.91)
CHA-B3LYP	4.24	(4.41)	CHA-B3LYP	4.34	(----)	CHA-B3LYP	8.63	(----)
B3LYP-BB	4.38	(4.46)	B3LYP-BB	4.48	(4.53)	B3LYP-BB	8.67	(8.78)
MP2	4.69	(4.76)	MP2	5.24	(5.24)	MP2	8.97	(9.03)
MP2-BB	4.04	(4.27)	MP2-BB	4.45	(4.74)	MP2-BB	7.86	(8.36)

In our opinion, the above described strange behavior of the CP (BB) and CHA methods is related to some fundamental aspects these BSSE-correction models. In both cases one has to specify subsystems of the investigated molecular complexes. This partition is straightforward in the investigated hydrogen bonded complexes but is clearly not unambiguous when one explores, for example, the effect of BSSE on proton transfer reactions. Another difficulty arises when the applied basis set does consist of basis functions with very small exponents. These functions represent very non-local objects making the partition of the investigated supermolecule to subsystems questionable.

The observed strange behavior of the CP and the CHA methods can be explained on the above basis. The CP approach predicts rather small BSSE-content of the investigated properties using the aug-cc-pVDZ basis set. In this case the CHA performs significantly better. However, the observed CHA convergence problems are clearly due to the fact that the investigated minima are located on that region of the PES where, due to the functions with small exponents present in the basis set, the partition of the supermolecule to subsystems is not feasible. This explanation is well supported by the

[111] I. Mayer, private communication.

facts that we had no convergence problems at geometries with large intermolecular distances and after removing that basis function of the basis set which bears the smallest exponent. Furthermore, convergence problems were not found in the case of the CHA-SCF optimizations where the optimized intermolecular distances are longer than the corresponding DFT values. The same effect can be seen in the case of the BLYP/aug-cc-pVTZ and B3LYP/aug-cc-pVTZ optimizations on the water dimer. The BLYP/aug-cc-pVTZ optimization converged easily while we had serious convergence problems at the B3LYP/aug-cc-pVTZ level. Analyzing our geometry data it is apparent that the B3LYP intermolecular distances are always shorter than the corresponding BLYP values. One has to note, however, that the validity of the CHA-BLYP/aug-cc-pVTZ geometry for the water dimer is questionable since the BSSE content of the uncorrected *roo* distances (Table VIII) are very similar at the BLYP/aug-cc-pVDZ and BLYP/aug-cc-pVTZ levels predicted by the CHA method. These facts indicate that there is a rather sharp change of behavior of CHA with decreasing intermolecular distances.

Based on our results obtained at the HF and DFT levels in conjunction with the aug-cc-pVDZ and aug-cc-pVTZ basis sets one can clearly state that the effect of the BSSE on energetical and structural parameters calculated at the above level is not significant and its correct description is a difficult task for both the CP and the CHA approaches. However, this is evidently not the case if traditional correlation methods are applied for which the CP approach works reasonably well. Future work to test the CHA-MP2 method in conjunction with the aug-cc-pVXZ basis sets is desired.

II.1.1.2 Complete basis set limit extrapolations (CBS)

In the previous section, the Dunning and co-workers family of correlation consistent [(aug)-cc-p(C)V*X*Z, *X*=2(D),3(T),4(Q),…)] basis sets[110] have been used with some convergence problems concerning the CHA method. As for the counterpoise-corrected results, the BSSE is reduced considerably with respect to other standard Pople of Huzinaga basis sets. As commented before, these basis sets systematically extend the atomic radial and angular spaces as a function of the cardinal number *X*. Therefore, results obtained with them seem to provide an excellent opportunity to extrapolate energies as well as properties to the complete basis set (CBS) limit. In most systems studied, the extrapolation of the total energy and many properties has been achieved with simple functional forms. The exponential form

$$A_X = A_{\mathrm{CBS}} + ae^{-bX} \tag{92}$$

and polynomials of the form

$$A_X = A_{\mathrm{CBS}} + \sum_{k=3}^{k_{\max}} \alpha_k (X + \beta_k)^{-k} \tag{93}$$

where $k_{\max}$ is small, or similar polynomial forms with non integer exponents have been extensively employed to estimate the CBS limit.[106,112-114]

The form of Eq. (92), for example, suggests that properties obtained with three basis

---

[112] T. van Mourik, A. K. Wilson, K. A. Peterson, D. E. Woon, T. H. Dunning Jr., *Adv. Quant. Chem.*, **31**, 105, (1998).

[113] a) D. Feller, *J. Chem. Phys.* **96**, 6104 (1992), b) K. A. Peterson and T. H. Dunning Jr., J. Chem. Phys. **102**, 2032 (1995), c) M. W. Feyereisen, D. Feller and D. A. Dixon, J. Phys. Chem. **100**, 2993 (1996)

[114] a) A.G. Császár, W.D. Allen and H.F. Schaefer III, *J. Chem. Phys.* **108**, 9751 (1998), b) J.M.L. Martin, *Chem. Phys. Lett.* **259**, 669 (1996), c) D. Feller, *J. Chem. Phys.* **98**, 7059 (1993), d) Gy. Tarczay, A.G. Császár, W. Klopper, V. Szalay, W.D. Allen and H.F. Schaefer III, *J. Chem. Phys.* **110**, 11971 (1999), e) T. Helgaker, W. Klopper, H. Koch and J. Noga, *J. Chem. P*hys. **106**, 9639 (1997), f) Y. Chuang and D.G. Truhlar, J. Phys. Chem. A **103**, 651 (1999)

sets (*e.g.*, *X*=D, T, and Q) gives an opportunity to extrapolate to the CBS limit of the investigated property.

Much experience[112] has been accumulated on how the correlation-consistent basis sets can be employed in calculations on various properties. For example, it is clear that for calculations of intermolecular complexes[113] one should apply the augmented version (aug-cc-pV*X*Z) of the cc-pV*X*Z sets containing diffuse functions, which are essential for the description of long-range interactions. However, the general trend is that the uncorrected aug-cc-pV*X*Z results lie close to the corresponding experimental values and the convergence of the CP-corrected aug-cc-pV*X*Z results to the CBS limit is slow. Based on these findings many of the authors of the pioneering studies criticised[113] the CP method and questioned the applicability of the CP scheme in conjunction with the aug-cc-pV*X*Z basis sets.

Dunning and co-workers have recently published a review[112] about their activity devoted to the exploration of the limits of the CP scheme. These authors have shown that, in many cases, the convergence behaviour of various molecular properties is significantly improved if the calculations are corrected for the BSSE. According these authors, for many investigated properties the smooth convergence behaviour of the results obtained with aug-cc-pV*X*Z basis sets is a pure illusion as it is due to a fortuitous cancellation of the BSSE and the basis set incompleteness error (BSIE)[31]. Correction for the BSSE destroys the balance of the two errors; consequently, the CP-corrected data lie farther away from experiment than the uncorrected ones. However, the use of CP-corrected data is much safer for CBS limit extrapolation: the corrected data suffer only from the BSIE, which can be taken into account by a suitable extrapolation to the CBS limit. Anomalies related to the extrapolation to the CBS limit were most pronounced for weakly-bound intermolecular complexes (van der Waals and hydrogen-bonded systems), although similar behaviour was observed even for certain strongly bound systems.

Dunning and co-workers have investigated various molecular properties like stabilization energies and equilibrium distances of di- and triatomic molecules. However, their work on real many-dimensional potential energy hypersurfaces (PESs) was limited to $(HF)_2$, partly due to the lack of an efficient automated procedure for BSSE-free geometry optimizations. [The optimized geometries were computed from numerical (CP-corrected or uncorrected) gradients.] Those studies, which expressed scepticism[113] about the usefulness of the CP scheme when applied in connection with the aug-cc-pV*X*Z basis sets were devoted to the characterization of PESs of hydrogen-bonded systems like

the HF dimer and the water dimer.

To resolve this apparent controversy, we decided to carry out geometry optimisations on the prototypical hydrogen-bonded systems chosen before. Our goal is to investigate the behaviour of the CP method in conjunction with the aug-cc-pV*X*Z basis sets for equilibrium geometric parameters. Therefore, we optimized the geometry of $(HF)_2$ and $(H_2O)_2$ employing analytical gradients at both the uncorrected and BSSE-corrected MP2/aug-cc-pV*X*Z [*X*=D(2), T(3), Q(4), for $(HF)_2$ also 5] levels of theory. The geometrical data obtained provide evidence in support of Dunning's opinion described above, that is, the use of the counterpoise-corrected PES is essential in order to properly apply the interpolation formulae to obtain the CBS limit of any molecular property.

From the comparison of our data with the literature results available on $(H_2O)_2$ we evaluate the reliability of various approximations used in geometry optimizations (*e.g.*, freezing a subset of the parameters) and show the importance of carrying out full geometry optimizations when extrapolation of geometrical parameters to the CBS limit is desired. Finally, we present optimized geometries and equilibrium stabilization energies of the HF–H_2O complex obtained on both uncorrected and BSSE-corrected MP2/aug-cc-pV*X*Z (*X* = D, T, Q) PESs.

Again, the discussion of the results will be carried out separately for each system.

II.1.1.2.a HF dimer

The structure of $(HF)_2$ was studied at the MP2/aug-ccpV*X*Z (*X* = D, T and Q) levels by Peterson and Dunning (henceforth PD)[113b]. In that paper the geometry of the complex was fully optimized at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels on both the uncorrected and CP-corrected potential energy hypersurfaces. (PD used numerical gradients for the optimization.) With the aug-cc-pVQZ basis set PD carried out full geometry optimization on the uncorrected PES, while they optimized only the intermolecular F–F distance (r_{FF}, see SCHEME 7) on the CP-corrected PES, keeping the other parameters fixed at their estimated MP2/aug-cc-pVTZ values. (For further details, see Ref. 113b.)

The most important conclusions drawn by PD are as follows:

(a) the uncorrected MP2/aug-cc-pV*X*Z stabilization energies and geometries are closer to experiment than the corresponding corrected ones

(b) the convergence of the CP-corrected binding energies and geometries is more regular than that of the uncorrected parameters, and

(c) the CBS limit of the CP-corrected r_{FF} distance, 2.737 Å, is numerically indistinguishable from the uncorrected aug-cc-pVQZ optimized value.

Overall, PD stressed the importance of the CP correction for the investigated properties. As a first step in our investigation, we reoptimized the geometry of $(HF)_2$ at the levels for which PD published data. We were able to reproduce their results with only slight deviations by using our automated CP optimization algorithm (see Table XI for the results).

TABLE XI.

Geometrical parameters (angles/deg and lengths/Å) and total energies (E_{tot}/E_h) of the HF dimer calculated at the uncorrected and CP-corrected MP2/aug-cc-pV*X*Z (*X* = 2, 3, 4, 5) levels of theory.

Uncorrected PES				
N	r_{FF}	α	β	E_{tot}
D	2.753[a]	6.5 (6.6[a])	110.2 (110.1[a])	–200.519075[a]
T	2.746[a]	6.4[a]	111.1[a]	–200.689285[a]
Q	2.736 (2.737[a])	6.4[a]	111.5 (111.6[a])	–200.746906[a]
5	2.739	6.3	112.0	–200.768417
Corrected PES				
N	r_{FF}	α	β	E_{tot}
D	2.812 (2.813[a])	7.0[a]	111.3[a]	–200.518060
T	2.770 (2.772[a])	6.7[a]	111.7 (111.8[a])	–200.688525
Q	2.753[a,b]	6.7	111.7	–200.746495
5	2.749	6.7	111.7	–200.768158

[a]Ref.114a. [b]Partial optimization. For details, see Ref. 114a. For the definition of the variables, see SCHEME 7. In those cases, when two values are presented for a variable, our optimized value does not agree with that of Ref. 114a given in parentheses.

For the CP-corrected MP2/aug-cc-pVQZ case PD carried out partial optimizations by fixing most of the internal parameters. Our full optimization at the same level confirmed the adequacy of the approximations applied by PD, the fully optimized parameters (α = 6.7°, β = 111.7°, r_{FF} = 2.753 Å) are close to the corresponding values (α = 6.4°, β = 111.15°, r_{FF} = 2.753 Å) of PD. As mentioned above, the uncorrected

MP2/aug-cc-pVQZ r_{FF} distance seems to be converged, and agrees with the CBS limit, 2.737Å, obtained on the CP-corrected surface by PD. (One has to note here that the CBS limit of r_{FF} determined by our geometry data, 2.741 Å, differs slightly from the corresponding value of PD, 2.737Å.) At the same time, the actual value of the r_{FF} distance (2.753 Å) on the corrected PES is far from both the MP2/aug-cc-pVQZ and the corrected CBS limit values. This situation provides an opportunity to further investigate the importance of CP correction in estimating the CBS limit of geometrical parameters, like the r_{FF} distance in $(HF)_2$. It seems to be worth investigating the geometry of $(HF)_2$ at the MP2/aug-cc-pV5Z level for the following reasons. If convergence of r_{FF} was really manifested at the MP2/aug-cc-pVQZ level, the MP2/aug-cc-pV5Z value would coincide with the MP2/aug-cc-pVQZ value. On the other hand, any other MP2/aug-cc-pV5Z r_{FF} value would question the extrapolation to the CBS limit using the uncorrected data. As it turns out, both the uncorrected (α = 6.3 °, β = 112.0°, r_{FF} = 2.739 Å) and CP-corrected (α = 6.7 °, β = 111.7°, r_{FF} = 2.749 Å) geometric parameters lie close to the corresponding MP2/aug-cc-pVQZ values. It is noted, that the difference between the corrected and uncorrected r_{FF} values is substantial, 0.010 Å, at the MP2/aug-cc-pV5Z level of theory. The series of MP2/aug-cc-pV*X*Z (*X* = D, T, Q, 5) r_{FF} data represent a minimum curve with the minimum between *X* equal to Q and 5, (see Figure 7). On the other hand, the CP-corrected MP2/aug-cc-pV*X*Z (*X* = D, T, Q, 5) r_{FF} distances follow a monotonic curve.

One could extrapolate to the CBS limit of the r_{FF} distance in three ways using the available data obtained on the corrected PES; *e.g.*, one could use the {2,3,4}, {3,4,5}, and {2,3,4,5} series for determining extrapolated geometric parameters according to Eq. (92). Obviously, more reasonable values are expected for the latter choices. The CBS limit values of the r_{FF} distance of $(HF)_2$ for the above listed series are reasonably close to each other at 2.741Å, 2.748Å, and 2.745Å, respectively. (For the 2–5 fitting we used our own CP-corrected data, which slightly differ from those of PD.)

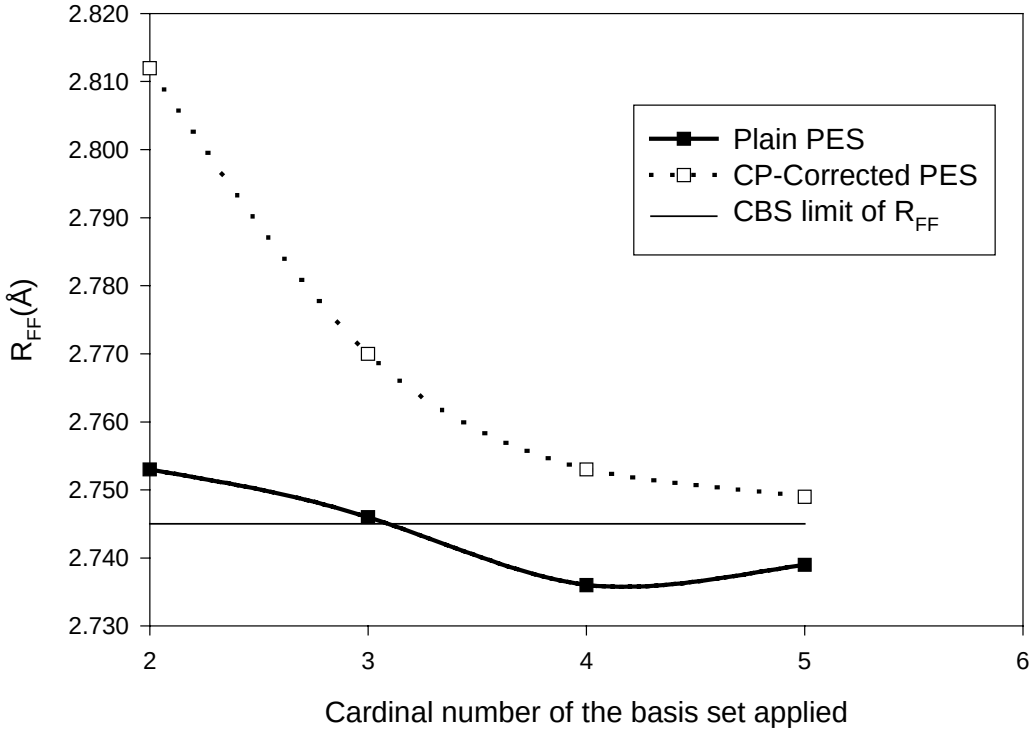


FIGURE 7: Comparison of the plain and CP-corrected r_{FF} distances for $(HF)_2$.

One has to note that the MP2/aug-cc-pVDZ level, which represents the lowest level of theory employed, provides a much poorer approximation than that obtained at the MP2/aug-cc-pVTZ level. This effect can be seen, for example, in the huge BSSE content of the geometry. After correction for BSSE, the r_{FF} distance is lengthened by 0.059 Å, 0.024 Å, 0.017 Å, and 0.010 Å at the MP2/aug-cc-pVDZ, MP2/aug-cc-pVTZ, MP2/aug-cc-pVQZ and MP2/aug-cc-pV5Z levels of theory, respectively. Inferiority of the MP2/aug-cc-pVDZ geometry parameters, as compared to the larger basis set results, is basically responsible for the differences between the CBS limit values obtained from the three fittings.

Nevertheless, as Chuang and Truhlar[114f] explored, one is tempted to use the inexpensive DZ and TZ data to obtain extrapolated geometric parameters corresponding to the basis set limit. However, instead of using the scheme of Chuang and Truhlar, we employed Eq. (93) in the form of

$$A_X = A_{\mathrm{CBS}} + aX^{-3} \tag{94}$$

to obtain estimates of the extrapolated r_{FF}. Our noteworthy result is that the extrapolated r_{FF} distances, 2.743 Å and 2.752 Å in the uncorrected and BSSE-corrected cases, respectively, are very close to the corresponding MP2/aug-cc-pV5Z numbers. This is especially notable for the CP-corrected r_{FF} distance, in which case the MP2/aug-cc-pVTZ value deviates from the MP2/aug-cc-pV5Z value by 0.021 Å, while the extrapolated distance deviates only by 0.003 Å.

In summary, the smooth convergence behavior of the geometric parameters obtained directly with the aug-cc-pV*X*Z basis sets is due to a fortuitous cancellation of BSSE and BSIE. The simple exponential or polynomial functions employed to extrapolate to the CBS limit cannot work reliably if the points do not follow a monotonic curve, as is the case for the uncorrected MP2/aug-cc-pV*X*Z results. On the other hand, although the CP-corrected r_{FF} distances are usually farther away from the extrapolated values than the uncorrected distances (this is true for all but the MP2/aug-cc-pV5Z data), changes in the corrected geometric parameters are monotonic and thus are in better accordance with the design philosophy of the aug-cc-pVXZ basis sets.

II.1.1.2.b Water dimer

The first study devoted to the investigation of $(H_2O)_2$ (see SCHEME 8) employing the aug-cc-pV*X*Z (*X* = 2, 3, 4, 5) basis sets at the MP2 level was carried out by Feller[113a], who determined the stabilization energy of the complex at various correlated levels at a fixed geometry. The first geometry optimizations on the water dimer at the MP2(FC)/aug-cc-pVXZ (X = D, T, Q) were carried out by Feller et al.[115] on the uncorrected PES. Importance of fragment relaxation terms in the CP scheme was investigated by Xantheas[64], who determined the fully relaxed geometry of $(H_2O)_2$ at the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels on the uncorrected potential energy hypersurfaces. Furthermore, Xantheas carried out partial optimizations at the MP2/aug-cc-pVQZ and MP2/aug-cc-pV5Z levels varying the intermolecular r_{OO} distance (Figure 8) and keeping other parameters fixed at their MP2/aug-cc-pVTZ value. Xantheas's MP2/aug-cc-pV5Z value for the r_{OO} distance is 2.905 Å on the uncorrected and 2.913 Å on the CP-corrected PESs, respectively. In a recent paper, Halkier et al.[106]

[115] D. Feller, E. D. Glendening, R. A. Kendall and K. A. Peterson, *J. Chem. Phys.*, **100**, 4981 (1994)

also investigated the structure of the water dimer employing the MP2/aug-cc-pV*X*Z models. These authors performed constrained optimizations, *e.g.*, they kept the water monomer parameters frozen at the corresponding experimental values. With this approximation, the geometry of $(H_2O)_2$ was determined at the MP2/aug-cc-pV*X*Z (*X* = 2, 3, 4) levels at both the uncorrected and the CP-corrected surfaces. (As a further approximation, the intermolecular valence angles were also frozen in the optimizations on the CP-corrected PES at their corresponding uncorrected values.) One should note here that, based on the analysis of their data, Halkier and co-workers argued against the computation of CP-corrected geometries at the MP2/aug-cc-pVXZ levels.

TABLE XII.

Geometrical parameters (deg, Å) and total energies (E_{tot}/E_h) of the water dimer calculated at the uncorrected and CP-corrected MP2/aug-cc-pVXZ (X = 2, 3, 4, 5) levels of theory.

	Ref. 34 and 35		Ref. 15				Present work							
	Plain	CP-cor.	Plain			CP-cor.	Plain				CP-cor.			
X	r_{OO}	r_{OO}	r_{OO}	α	β	r_{OO}	r_{OO}	α	β	E_{tot}	r_{OO}	α	β	E_{tot}
2	2.916[b]	2.975[c]	2.912[a]	5.3[a]	124.7[a]	2.977[a]	2.917	5.8	122.4	–152.530207	2.976	5.8	122.5	–152.528936
3	2.907[b]	2.933[c]	2.891[a]	4.7[a]	124.1[a]	2.935[a]	2.907	5.5	123.4	–152.666241	2.932	5.5	124.7	–152.665503
4	2.903[a,b]	2.918[a,c]	2.895[a]	5.4[a]	124.3[a]	2.915[a]	2.903	5.5	124.4	–152.711954	2.917	5.5	125.0	–152.711954
5	2.905[a,c]	2.913[a,c]	–	–	–	–	–	–	–	–	–	–	–	–

[a]Partial optimizations, for details of the constraints employed during the optimizations, see text. For the definition of the variables, see SCHEME 8.
[b] Ref 115
[c] Ref 64

The results of our full geometry optimizations carried out on the uncorrected and CP-corrected PESs are shown in Table XII. At the MP2/aug-cc-pVDZ and MP2/aug-cc-pVTZ levels we could reproduce Feller's and Xantheas's results with slight differences on both the uncorrected and corrected PESs. On the other hand, the r_{OO} distances obtained by Halkier et al. for the uncorrected surfaces differ considerably from the corresponding fully relaxed values. For example, the difference between the r_{OO} distances of Feller and Halkier is 0.016 Å at the MP2/aug-cc-pVTZ level. This discrepancy questions the validity of the approximations utilized by Halkier et al. Furthermore, Xantheas's data determined at the MP2/aug-cc-pVQZ level from partial optimizations agree reasonably well with both Feller's and our values obtained from full optimizations. Hereby, we can confirm the adequacy of the approximations employed by Xantheas in his partial optimizations on $(H_2O)_2$, *e.g.*, keeping all the intramolecular and the valence bending intermolecular parameters at their optimized MP2/aug-cc-pVTZ values in the subsequent aug-cc-pVQZ

and aug-cc-pV5Z optimizations. Because of the reliability of the approximations employed by Xantheas and the cost of the MP2/aug-cc-pV5Z optimization we decided not to determine the geometry of $(H_2O)_2$ at the MP2/aug-cc-pV5Z level.

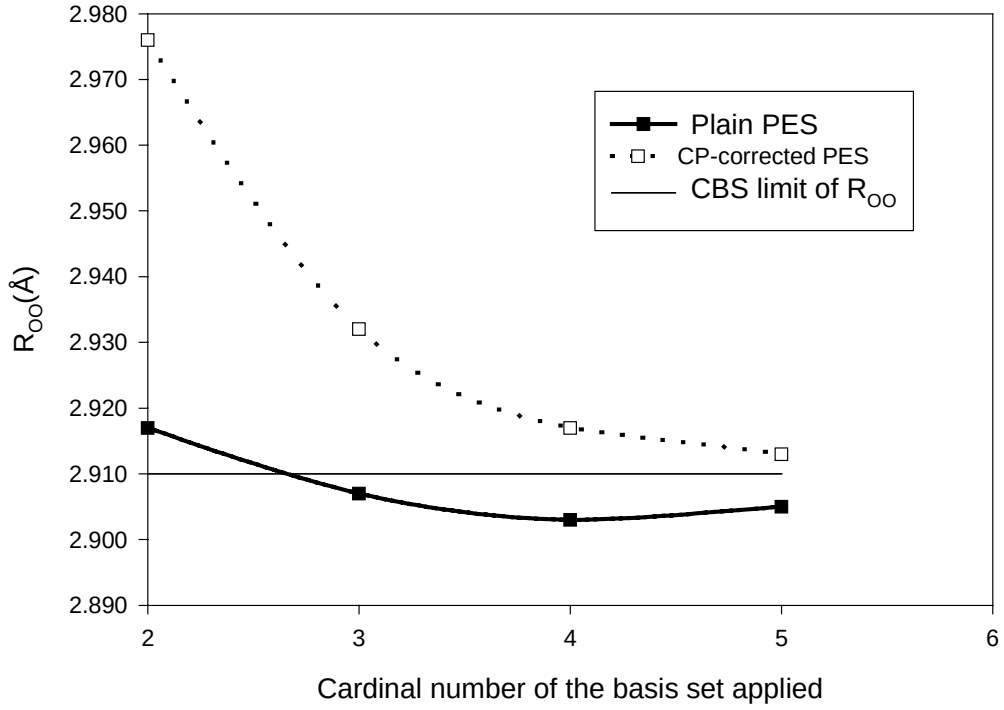


FIGURE 8: Comparison of the plain and CP-corrected r_{OO} distances for $(H_2O)_2$.

Concerning the extrapolation of the r_{OO} distance to the CBS limit, one can draw the following conclusions. The two series of r_{OO} distances, determined by Feller et al., Xantheas and Halkier et al. employing different approximations during the optimizations, follow a curve with a minimum. That is, Halkier's aug-cc-pVTZ r_{OO} distance at 2.891 Å is shorter than the corresponding aug-cc-pVDZ (2.912 Å) and aug-cc-pVQZ (2.895 Å) values. This behavior is presumably due to the inadequacy of the constraints employed by Halkier et al. during the optimizations. In the data series of Feller et al. and Xantheas, the aug-cc-pVQZ r_{OO} distance (2.903 Å) is shorter than the corresponding aug-cc-pVDZ (2.916 Å), aug-cc-pVTZ (2.907 Å), and aug-cc-pV5Z (2.905 Å) values. This behavior of the aug-cc-pV*X*Z basis sets in the case of uncorrected calculations is very similar to what we observed in full geometry optimizations on $(HF)_2$. On the other hand, the CP-corrected r_{OO} distances converge more regularly, all three data series obtained by Xantheas, Halkier

et al. and the present work ascend gradually when enlarging the cardinal number of the employed basis set, (see Figure 8). The extrapolated CBS limit of the r_{OO} distance obtained from Halkier's data seems to be too short at 2.897 Å, indicating again the inadequacy of the constraints employed during the geometry optimizations. On the other hand, the extrapolated r_{OO} CBS limit values obtained from Xantheas's partially optimized aug-cc-pV*X*Z {*X* = 2, 3, 4, 5} and our fully optimized aug-cc-pV*X*Z {*X* = 2 , 3, 4} data using Eq. (92) practically coincide at 2.910 Å. It is worth noting in this respect that the CBS limit values determined from the {*X* = 2, 3, 4} and {*X* = 3, 4, 5} series of Xantheas are also the same at 2.910 Å.

Similarly to the case of $(HF)_2$, use of Eq. (94) for the extrapolation of aug-cc-pVDZ and aug-cc-pVTZ data gives very satisfactory results for the CP-corrected r_{OO} distances. While the difference between the aug-cc-pVTZ and aug-cc-pV5Z r_{OO} distances is 0.020 Å, the extrapolated value differs by only 0.005 Å.

As a summary of the above considerations, we strongly advocate to perform full geometry optimizations in those studies whose aim is the extrapolation of geometrical parameters to the CBS limit. The reliability of the CBS limit of the r_{OO} distance determined from the fully relaxed CP-corrected aug-cc-pV*X*Z {*X* = 2, 3, 4} potential energy hypersurfaces is appealing compared to the aug-cc-pV*X*Z {*X* = 2, 3, 4, 5} data obtained by Xantheas.

II.1.1.2.c HF-H_2O complex

The most recent theoretical investigation devoted to the structure of HF–H_2O, including geometry optimizations, was carried out by Novoa *et al.*[116] at the MP2/6-311++G(2d,2p) level of theory resulting in an energy difference of 0.45 kcal/mol between the C_s and C_{2v} structures (see Section II.1.1.1.c). They have also calculated the equilibrium r_{FO} distance (2.663 Å) using a constrained optimization keeping the monomer parameters fixed. Finally, their stabilization energies computed at the uncorrected and CP-corrected MP2/6-311++G(2d,2p) levels are –9.20 and –7.87 kcal/mol, respectively.

Because of the available experimental and theoretical information, the HF–H_2O complex is an ideal subject to test the CBS limit values of various properties determined

[116] J.J Novoa, M. Planas, M. Whangbo and J.M. Williams, *Chem. Phys.*, **186**, 175 (1994)

from calculations performed at the MP2/aug-cc-pV*X*Z {*X*=2, 3, 4} levels. To explore the quality of theoretical results obtained on the uncorrected and CP-corrected surfaces, we decided to calculate the equilibrium dissociation energy, the energy difference between the C_s and C_{2v} species, and the equilibrium geometry of HF–H_2O. The calculated total energies, energy barriers and geometrical parameters obtained at the MP2/aug-cc-pV*X*Z {*X*=2, 3, 4} levels are given in Table XIII. The CBS value for the total energy of H_2O is –76.363558 E_h, obtained from the total energies (–76.260910, –76.328992, and –76.351919 E_h) calculated at the MP2/aug-cc-pVDZ, MP2/aug-cc-pVTZ, and MP2/aug-cc-pVQZ levels, respectively.

TABLE XIII.

Geometrical parameters (angles/deg and lengths/Å) and total energies (E_{tot}/E_h) of HF–H_2O calculated at the uncorrected and CP-corrected MP2/aug-cc-pV*X*Z (*X* = 2, 3, 4) levels of theory.

	C_s							
	Plain				CP-corrected			
X	E_{tot}	r_{FO}	α	β	E_{tot}	r_{FO}	α	β
2	–176.531107	2.658	129.5	1.4	–176.529305	2.701	130.6	1.5
3	–176.684156	2.643	130.7	1.4	–176.683198	2.662	132.0	1.5
4	–176.735847	2.640	132.4	1.4	–176.735257	2.654	132.6	1.4
CBS	–176.762208	2.639	–	–	–176.761869	2.652	–	–
	C_{2v}							
	Plain				CP-corrected			
X	E_{tot}	r_{FO}	α	β	E_{tot}	r_{FO}	α	β
2	–176.530314	2.666	0.0	180.0	–176.528674	2.710	0.0	180.0
3	–176.683503	2.655	0.0	180.0	–176.682605	2.670	0.0	180.0
4	–176.735318	2.648	0.0	180.0	–176.734751	2.661	0.0	180.0
CBS	–176.761803	2.636	–	–	–176.761465	2.658	–	–

The experimental[109] equilibrium stabilization energy (E_{int}) of HF–H_2O is –10.2 kcal/mol, determined by Legon and co-workers from absolute intensities of rotational transitions. As mentioned before, the best theoretical values[116] available are –9.20 and –7.87 kcal/mol obtained from uncorrected and CP-corrected calculations using a medium-size basis set and the MP2 method. The uncorrected aug-cc-pV*X*Z {*X* = 2, 3, 4} stabilization energies shown in Table XIV are appealingly close to each other. The CBS limit value of E_{int} obtained using the extrapolated total energies of the monomers and the complex, differs only slightly from these values. The stabilization energies obtained at the CP-corrected PESs considerably differ from the corresponding uncorrected energies. However, this difference gradually decreases considering the series of basis sets, going

from the aug-cc-pVDZ set to the aug-cc-pVQZ one. The CP-corrected CBS limit of E_{int} at –8.6 kcal/mol is remarkably close to the corresponding uncorrected value at –8.8 kcal/mol. Comparing the experimental and theoretical results, one has to note that the CBS limit ΔE values obtained at the MP2 level considerably differ from the experimental[108a] value (–10.2 kcal/mol). Perhaps, truncation of the correlation energy expansion at the MP2 level does not represent a satisfactory description of the investigated problem.

TABLE XIV

Equilibrium stabilization energies (E_{int} in kcal/mol) and energy splitting (ΔE in kcal/mol) between the C_s and C_{2v} species of HF–H_2O calculated at the uncorrected and CP-corrected MP2/aug-cc-pV*X*Z (*X* = 2, 3, 4) levels of theory.

	E_{int}		ΔE	
X	Plain	CP-corrected	Plain	CP-corrected
2	–9.0	–7.9	0.50	0.40
3	–9.0	–8.4	0.41	0.37
4	–8.9	–8.5	0.33	0.32
CBS[a]	–8.8	–8.6	0.25	0.25

[a]The CBS limit values were calculated using the respective CBS limit total energies shown in Table XIII.

The calculated energy barriers (ΔE) belonging to the transition from the C_s to the C_{2v} species of HF–H_2O are also shown in Table XIV. ΔE gradually decreases going from the smallest to the larger basis sets. It is worth noting that the ΔE values obtained at the plain and CP-corrected surfaces are very close to each other at the MP2/aug-cc-pVQZ level of theory. The CBS limit values of the same parameter obtained at the plain and the CP-corrected PESs numerically coincide at 0.25 kcal/mol. The CBS limit of ΔE is considerably smaller than the corresponding equilibrium experimental[108a] data (0.36 kcal mol^{-1}), again showing the importance of inclusion of higher order correlation terms. It is worth noting that the good agreement with experiment obtained in previous theoretical studies[107,116] is clearly due to cancellation of errors, *e.g.*, the quality of the basis sets employed in those investigations was unsatisfactory to consistently deal with the problems of BSSE and BSIE.

Analyzing the geometrical data (mainly the parameter r_{FO}) presented in Table XIII, one can find trends similar to those already detected in the cases of $(HF)_2$ and $(H_2O)_2$. The r_{FO} values obtained on the uncorrected PES change less than the corresponding CP-

corrected values with change in the basis set. The CBS limit of r_{FO} at the uncorrected PES, 2.639 Å, is very close to both the MP2/aug-cc-pVTZ (2.643 Å) and MP2/aug-cc-pVQZ (2.640 Å) values. On the other hand, the CBS limit of r_{FO} is 2.652 Å on the CP-corrected PES, representing a difference of 0.013 Å between the two extrapolated values. Legon and co-workers have determined the experimental value of r_{FO} at 2.662 Å from the microwave rotational spectrum[108b] of HF–H_2O. Direct comparison of the experimental r_0 and theoretical r_e values is not fully valid but it is clear that the CP-corrected CBS limit r_{FO} distance lies closer to experiment than the corresponding uncorrected value, (see SCHEME 9). However, one has to note here that the vibrational effects on hydrogen bond lengths can be substantial resulting in a case when the uncorrected CBS limit value is closer to the unmeasured experimental r_e value. Another interesting anomaly concerns the change of the actual value of r_{FO} going from the C_s to the C_{2v} species. One expects that the intermolecular bond length lengthens when climbing the transition state (C_{2v} species) region. All the data listed in Table XIII supports this statement but the CBS limit of r_{FO} obtained at the plain PESs contradicts it. In our opinion, this strange behavior is again due to the unreliability of the CBS limit values of geometry parameters obtained on the uncorrected PES.

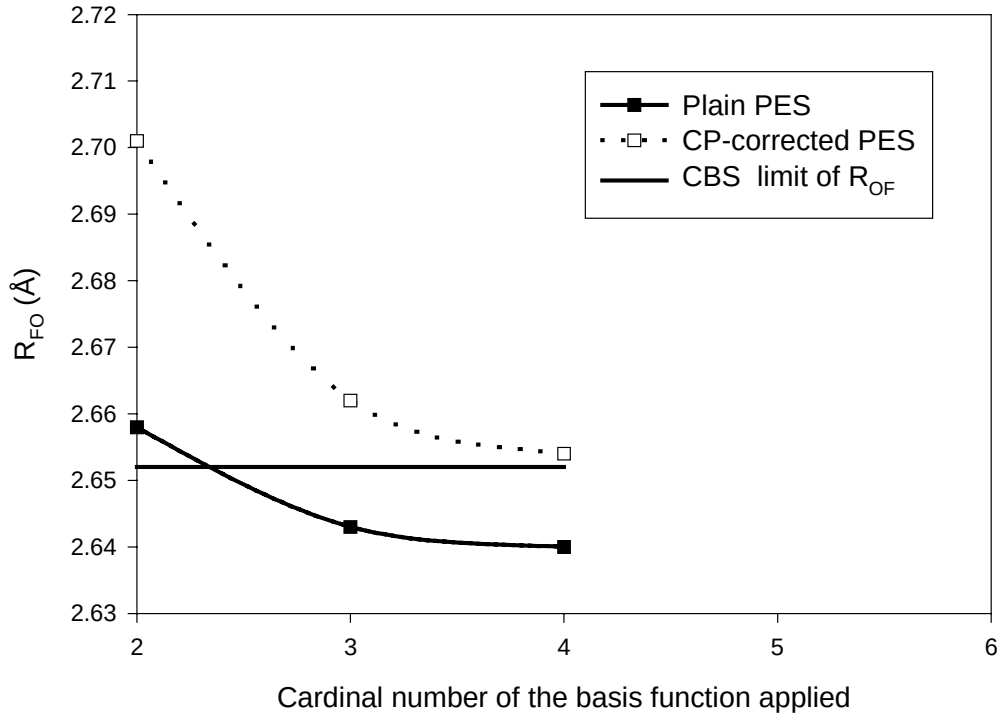


FIGURE 9: Comparison of the plain and CP-corrected r_{FO} distances for HF–H_2O

Concerning the extrapolated r_{FO} distances obtained by using Eq. (94), one finds tendencies similar to those observed in the cases of $(HF)_2$ and $(H_2O)_2$. The extrapolated C_{2v} r_{FO} distances, 2.650 Å and 2.653 Å, in the uncorrected and CP-corrected cases, respectively, are very close to the corresponding CBS value (2.658 Å) determined by using Eq. (92) on the CP-corrected PESs. The extrapolated C_s r_{FO} distances, 2.637 Å and 2.646 Å determined on the plain and CP-corrected PESs lies close to the corresponding CBS limit values (2.639 Å and 2.652 Å) obtained by using Eq. (92). It is worth noting here that the extrapolated (Eq. (94)) C_{2v} r_{FO} value is longer than the corresponding C_s value in both the plain and CP-corrected cases. This means that bond length extrapolation based on Eq. (94) is more reliable in the case of the HF–H_2O complex than application of Eq. (92) since the expected change of parameter r_{FO} is at least qualitatively given back by the former technique.

In short, the application of the extrapolation to the CBS limit of intermolecular geometry parameters seems to be fruitful if various species are to be compared on the PES of the complex investigated. Also, fixing intramolecular parameters at their experimental values could cause difficulties during the extrapolation. As the available literature data and our results clearly show, the MP2/aug-cc-pV*X*Z {*X* = 2, 3, 4} data series of intermolecular distances obtained from the CP-corrected surfaces can be safely used for the purpose of CBS extrapolations.

II.1.1.3 Blue-Shifting and double vs single-well PES

It has just been pointed out, that the hydrogen fluoride-water complex exhibits a double-well PES, with a tunneling barrier of the order of the first vibrational level. The interaction nergy of the complex was estimated to be around 10kcal/mol. Therefore, it shows a quite strong hydrogen bond interaction which seemed to prevent for drastic changes due to BSSE in the PES.

In this section we will see another type of weakly bonded complexes with a C-H···O interaction which often present this double-well PES. Due to the particular nature of the interaction one can expect the BSSE have a important role in the accurate description of the PES, mainly with small basis sets. Another interesting point is concerning the so-called blue-shifted hydrogen bonds[117]. This a rather hot issue in the literature in the last few years and is generating important controversies about the nature of these hydrogen bonds (see for instance Refs 118,119 and references therein). Roughly, in these particular complexes, the *intramolecular* X-H, where X is usually C, shortens upon interaction and hence the stretching frequency associated increases. This is in complete disagreement with the typical situation in hydrogen bonded and van der Waals complexes where there is a characteristic red-shift of the intramolecular frequency, which is often referred as the fingerprint of the interaction.

In a recent paper, Gu et al.[118] showed that the features of the C-H···O interactions are very analogous to the standard O-H···O hydrogen bond interactions, though they are much weaker as a consequence of the smaller acidity of the proton involved. By comparing a C-H hydrogen-bonded complex to the water dimer in terms of the energetic contributions to the bonding, these authors found that there is no actual distinction between the two types of hydrogen bonds. In both cases, the main opposite forces operating in the interactions are the electrostatic and the exchange ones. Thus, the blue shift effect observed for some of these complexes is just a consequence of the balance of two sets of forces acting in opposite directions.

In a previous paper[119], Turi and Dannenberg studied several complexes containing

---

[117] P. Hobza, Z. Havlas, Chem. Rev. 100, 4253 (2000)

[118] Y. Gu, T. Kar and S. Scheiner, *J. Am. Chem. Soc.*, **121**, 9411 (1999)

[119] L.Turi and J.J. Dannenberg, *J. Phys. Chem.*, **97**, 7899 (1993)

intermolecular C-H···O interactions. They dealt with complexes between acetylene or hydrogen cyanide (as hydrogen donors) and water, formaldehyde or ozone molecules. For each system, full geometry optimizations and harmonic vibrational frequency calculations were carried out. Stabilization energies corrected for BSSE, zero-point vibrational energy (ZPVE) and thermal correction to enthalpy at 298K were also computed. The calculations were made at the Hartree-Fock and MP2 levels of theory. For the complexes involving ozone, two geometries were found on the PES, depending on the basis set used. With the 6-31G(d,p) basis sets, a three-center hydrogen bond between the hydrogen atom of the donor molecule and the oxygen atoms of ozone was predicted. On the contrary, using the larger d95++(d,p) basis set, non-symmetric structures were found, where the ozone molecule rotated in the molecular plane so that one hydrogen bond was ca. 0.5 - 0.7Å shorter than the other bond. Further ZPVE, BSSE and enthalpy corrections stabilized the symmetric structure with respect to the non-symmetric complex, the largest differences being ca. 0.1Kcal/mol.

Therefore the goal now is twofold. To investigate the BSSE effects on the PES by means of the counterpoise method and to assess possible effects on the shift of the intramolecular vibrational frequencies.

We perform a deeper study of some of the C-H···O complexes studied by Turi and Dannenberg[119], but now taking into account the BSSE effect in both the geometry and the vibrational frequencies. We assess how the corresponding PES is modified through correction by BSSE, and whether the counterpoise method is still able to yield reliable results using small basis sets. In that case, the CP correction would be a good compromise to obtain accurate descriptions of large systems presenting hydrogen bonds or weak intermolecular interactions like crystals structures or biological molecules.

We have chosen as benchmarks the complexes of cyanide and acetylene with ozone for two reasons. First, the interactions between the donor and acceptor moieties are small enough to foresee that a correct estimation of the BSSE effect is essential to discern between the possible structures the complexes have proved to adopt, namely the symmetric three-center hydrogen bond (single well) and the non-symmetric bond (double well). Second, because the early study of by Turi and Dannenberg[119] already suggested that correction for BSSE may transform a double well into a single well. For this reason we have performed the CP-corrected calculation at the same level of theory than in Ref 119, that is, RHF and MP2 calculations with two different basis set, namely 6-31G(d,p)

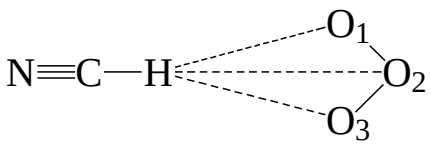
and D95++(d,p).

We expect that the deeper analysis of the CP correction by means of a CP-corrected PES will help us to explain under which circumstances a single-well on the PES of some weakly bonded complexes transforms into a double well, depending on the basis set and methodology used, and whether this situation is due to a poor behavior of the method or to a pure BSSE effect.

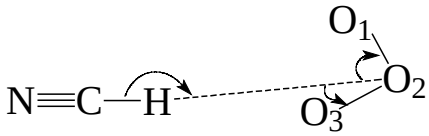
In the following we discuss the CP-corrected results obtained in terms of geometries, stabilization energies and vibrational frequencies for the complexes, and compare them with both the uncorrected results and the experimental data available.

II.1.1.3.a The $CNH\cdots O_3$ complex

The results obtained for the $CNH\cdots O_3$ complex are collected in Tables XV and XVI. Two different planar geometries are considered for this complex. One of them has a symmetric structure, where the rH_1O_1 and rH_1O_3 distances are equivalent and the CHO_2 angle has a value of 180 degrees (see SCHEME 10). Uncorrected optimizations lead to these symmetric structures with the 6-31G(d,p) basis set, whereas with Dunning's d95++(d,p) basis set, only non-symmetric minima are found, which corresponds to the other planar geometry considered.



Symmetric



Non-symmetric

SCHEME 10: $CNH\cdots O_3$ structures

Uncorrected stabilization energies range from -1.51 Kcal/mol, obtained at the

HF//d95++(d,p) level, to the largest stabilization energy of -3.84Kcal/mol at the MP2//6-31G(d,p) level of theory. ZPVE and enthalpy corrections decrease the stabilization energy by roughly 0.5-0.8 and 1.0-1.1 Kcal/mol, respectively. Inclusion of correlation energy through the MP2 method increases this value by more than 60% in all cases.

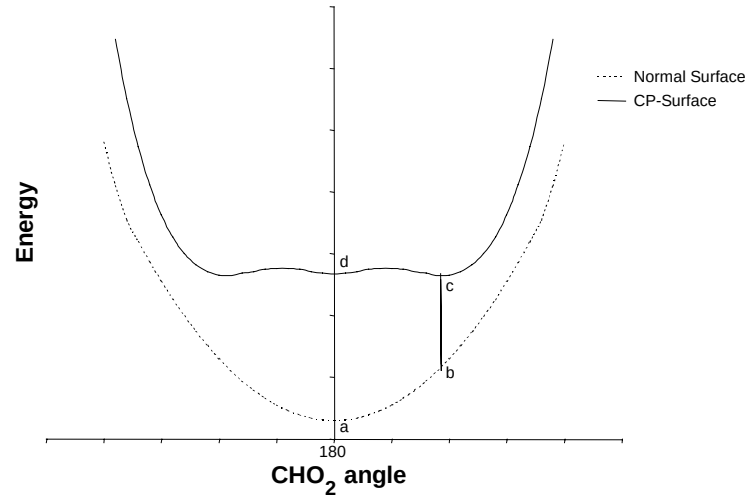
TABLE XV.

Comparison of the geometric parameters of $CNH\cdots O_3$ on the normal surface, the CP-optimized surface and the symmetric CP-optimized. Distances in Å and angles in degrees.

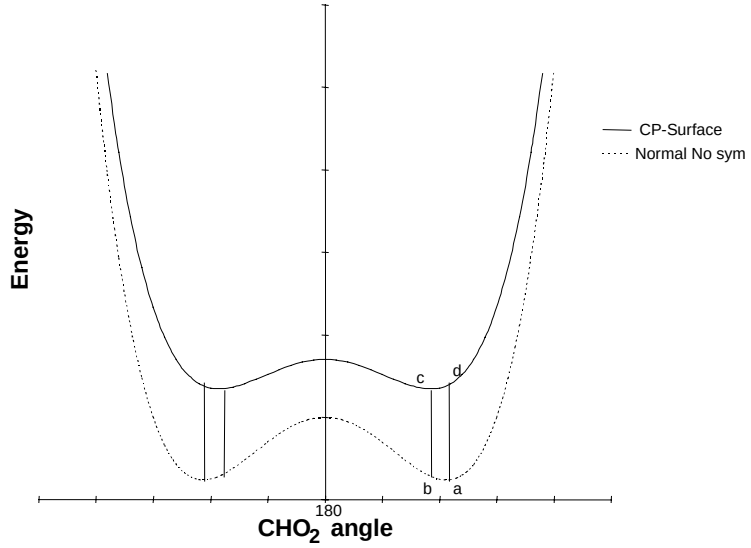
		6-31G(d,p)		**D95++(d,p)**	
		HF	**MP2**	**HF**	**MP2**
rHO_1	Normal	2.704	2.536	3.265	2.993
	CP opt	3.163	2.983	3.292	3.020
	CP opt (sym.)	2.771	2.631	2.822	2.755
rHO_2	Normal	3.111	2.972	3.313	3.168
	CP opt	3.244	3.167	3.338	3.252
	CP opt (sym.)	3.185	3.108	3.238	3.213
rHO_3	Normal	2.704	2.536	2.540	2.448
	CP opt	2.507	2.452	2.560	2.570
	CP opt (sym.)	2.771	2.631	2.822	2.755
$a(HO_1O_2)$	Normal	59.4	58.1	77.2	70.2
	CP opt	75.4	70.0	77.4	68.2
	CP opt (sym.)	59.3	58.1	59.4	58.0
$a(HO_2O_3)$	Normal	59.4	58.1	41.7	45.8
	CP opt	43.4	46.1	41.5	47.9
	CP opt (sym.)	59.3	58.1	59.4	58.0
$a(CHO_2)$	Normal	180.0	180.0	168.6	163.7
	CP opt	168.1	167.3	168.2	168.2
	CP opt (sym.)	180.0	180.0	180.0	180.0

When the CP correction is added to the uncorrected minima, the numerical dispersion in the different stabilization energies clearly decreases. SP-corrected results range from -1.23 to -1.66, -0.70 to -0.86, and -0.28 to -0.51 Kcal/mol for the electronic, ZPVE-corrected and enthalpy-corrected stabilization energies, respectively. SP-correction has proved to be larger at the MP2 level for both basis sets; however, the use of diffuse functions clearly decreases its value.

Differing from conventional, BSSE-contaminated results, CP-corrected optimizations yield non-symmetric minima in all cases. The ozone molecule is displaced farther from the hydrogen and becomes slightly rotated in the molecular plane, so that the intermolecular distances rHO_1 and rHO_2 increase, whereas O_3 approaches the hydrogen atom (see SCHEME 10).



(a)



(b)

SCHEME 11: Comparison between conventional and CP-corrected PES. a) for the 6-31G(d,p) basis set. b) for the d95++(d,p) basis set. Points *a* and *c* are optimized structures on normal and CP-corrected surfaces, respectively, while points *b* and *d* represent optimized geometries on the other surface.

Since symmetric structures were located on the uncorrected surface with the 6-31G(d,p) basis set, larger geometric changes appear in this case. Differences of up to 0.4Å in the intermolecular distances and 10-15 degrees in the hydrogen bond angles are observed upon CP-correction. It is interesting to see that, even though the rHO_1 and rHO_2 distances clearly increase, the distance to the closest oxygen atom decreases by more than 0.1Å, thus leading to a stronger interaction.

Concerning energetics, CP and SP electronic stabilization energies are reasonably close for both basis sets, despite the large difference in the geometries. This is due to the large flatness of the PES studied. The interaction is so weak that large displacements on the quadratic region of a minimum translate into small energy penalties. However, the SP correction slightly overestimates the effect of BSSE, because it is computed at the uncorrected minimum. SCHEME 11a depicts the situation in this case: the distance between the points c and b (CP-correction) is always shorter than the distance between d and a (SP-correction).

TABLE XVI.

Energetic results of the NCH···O_3 complex. The values labeled ZPVE corr. and Enthal. corr. hold for the stabilization energy including the ZPVE correction and both the ZPVE and the thermal correction to the enthalpy, respectively.

Basis Set	**Method**	**Geom.**	**Stabilization energy (Kcal/mol)**			**ΔE_{O3}**	**ΔE_{CNH}**
			Electronic	**ZPVE corr**	**Enth. corr**	**(Kcal/mol)**	
6-31G(d,p)	HF[a]	Symm.	-2.28	-1.69	-1.26		
	SP-HF[b]	Symm.	-1.39	-0.80	-0.38	0.83	0.05
	CP-HF[c]	No Symm.	-1.48	-0.83	-0.47	0.61	0.05
	CP_{sym}-HF[d]	Symm.	-1.41	-0.79	-0.41	0.81	0.04
	MP2[a]	Symm.	-3.84	-3.04	-2.69		
	SP-MP2[b]	Symm.	-1.66	-0.86	-0.51	2.03	0.14
	CP-MP2[c]	No Symm.	-1.78	-0.97	-0.64	1.66	0.11
	CP_{sym}-MP2[d]	Symm.	-1.74	-0.96	-0.60	1.91	0.10
d95++(d,p)	HF[a']	No Symm.	-1.51	-1.01	-0.56		
	SP-HF[b']	No Symm.	-1.35	-0.85	-0.41	0.09	0.05
	CP-HF[c']	No Symm.	-1.36	-0.87	-0.41	0.10	0.06
	CP_{sym}-HF[d]	Symm.(TS)	-1.28	-0.89	*-0.94*	0.10	0.05
	MP2[a]	No Symm.	-2.20	-1.67	-1.24		
	SP-MP2[b]	No Symm.	-1.23	-0.70	-0.28	0.72	0.24
	CP-MP2[c]	No Symm.	-1.27	-0.82	-0.34	0.67	0.23
	CP_{sym}-MP2[d]	Symm.(TS)	-1.26	-0.88	*-0.93*	0.64	0.23

[a] Point a in SCHEME 11a
[b] Point d in SCHEME 11a
[c] Point c in SCHEME 11a
[d] Point c in a symmetric surface
[a'] Point a in SCHEME 11b
[b'] Point d in SCHEME 11b
[c'] Point c in SCHEME 11b

The behavior of CP-correction changes upon use of diffuse functions; as mentioned above, with the d95++(d,p) basis sets, both the conventional and the CP-corrected PES yield non-symmetric minima. A qualitative graphical representation of both PES is given in SCHEME 10b. At the HF level, the BSSE effect on geometry is very small: maximum differences of 0.03Å in the intermolecular distances are observed, whereas the angles remain practically unchanged. At the MP2 level, the CP-corrected intermolecular

distances increase by more than 0.1 Å for the rHO_3 bond, and the ozone molecule rotates about 3 degrees. Regarding stabilization energies, as both the SP and the CP corrections are performed on a non-symmetric surface, the differences are very small. Again, predicted SP-corrected stabilization energies are more repulsive than the corresponding CP-corrected ones.

Because all CP-corrected structures are unsymmetrical, we performed CP-corrected optimizations on a symmetry-restricted PES (CP_{sym}) too, i.e., keeping the rHO_1 and the rHO_2 distances the same, and freezing the angle $a(C_1H_1O_2)$ to be linear. The nature of the corresponding CP_{sym} stationary point will clarify whether the CP-corrected PES exhibits either a double-well minimum separated by a small energetic barrier, or multiple minima on an extremely flat surface. In the first case, the two equivalent CP-corrected non-symmetric minima would be connected through a saddle point, which would correspond to the CP_{sym}. stationary point.

Harmonic frequencies computed at the corresponding stationary points of the three different PES considered (uncorrected, CP and CP_{sym}) are gathered in Table XVII. The stationary points of the uncorrected PES are well-defined minima in all cases, showing symmetric and non-symmetric features for the 6-31G(d,p) and d95++(d,p) basis sets, respectively. However, when correcting for the BSSE the situation changes. For the smaller basis set, both the symmetric and the non-symmetric structures located on the CP and CP_{sym} PES, respectively, exhibit real vibrational frequencies, i.e., both are local minima. The energy difference between both geometries represents less than 0.1kcal/mol in all cases, and even decreases after ZPVE and enthalpy correction. This difference is smaller than that of the first vibrational level, and hence, the averaged atomic positions correspond to a symmetric structure, which is actually what experimental data[120] have shown.

[120] Z. Mielke, A Andrews, J. Phys. Chem. 94, 3519 (1990)

TABLE XVII

Comparison of the vibrational frequencies of $CNH\cdots O_3$ on the normal surface, the CP-optimized surface and the symmetric CP-optimized surface. (cm^{-1})

		6-31G(d,p)		D95++(d,p)	
		HF	**MP2**	**HF**	**MP2**
Normal		28	53	24	24
CP opt	**δ**	32	36	25	17
CP opt (sym)		30	43	*47i*	*48i*
Normal		51	63	71	79
CP opt	**δ_0**	81	94	72	65
CP opt (sym)		58	73	38	31
Normal		54	67	44	41
CP opt	**π_1**	65	80	45	42
CP opt (sym)		68	81	42	39
Normal		64	87	64	66
CP opt	**ν**	68	77	61	45
CP opt (sym)		56	67	52	53
Normal		150	162	121	126
CP opt	**π_2**	144	160	120	123
CP opt (sym)		156	164	129	124
Normal		859	737	850	719
CP opt		856	735	850	711
CP opt (sym)		860	738	853	698
Normal		907	772	885	738
CP opt		916	791	884	730
CP opt (sym)		908	776	869	726
Normal		932	845	889	739
CP opt		928	833	887	737
CP opt (sym)		930	836	884	736
Normal		1447	1176	1424	1185
CP opt		1436	1174	1424	1185
CP opt (sym)		1443	1176	1432	1186
Normal		1539	2041	1557	1997
CP opt		1548	2039	1557	1997
CP opt (sym)		1537	2040	1546	1997
Normal		2434	2392	2405	2440
CP opt		2433	2389	2404	2440
CP opt (sym)		2434	2392	2405	2445
Normal		3642	3510	3619	3472
CP opt		3633	3503	3621	3477
CP opt (sym)		3641	3510	3627	3481

On the contrary, for the CP_{sym} d95++(d,p) calculations, the CP_{sym} geometries are

found to be a saddle point, which must be related to two equivalent CP non-symmetric structures. Therefore, one can determine the barrier height to internal rotation, estimated as the energy difference between the CP_{sym} and the CP minima. It is interesting to note that this energy difference becomes negative upon ZPVE or enthalpy corrections, i.e., the symmetric structure is favored respect to the non-symmetric one. In the latter case, this gap is ca 0.5Kcal/mol. However, this large difference is spurious, as there is one vibrational mode (that assigned to the imaginary frequency) that does not contribute to the enthalpic thermal corrections for the CP_{sym} structure[121], and thus leads to an artificial lowering in the energy of the CP_{sym} stabilization energy.

Vibrational analysis shows the presence of up to five intermolecular vibrational modes having an associated vibrational frequency below $200cm^{-1}$. Indeed, such vibrations are those where the CP-correction has a largest effect, not just concerning the vibrational frequency value (the maximum differences observed are ca. $40cm^{-1}$), but especially on the relative ordering of the vibrational modes. We have sorted these intermolecular modes from smaller to larger vibrational frequency, with respect to the values obtained using the 6-31G(d,p) basis set. These modes are assigned as follows: modes labeled δ and δ_0 correspond to the oscillation of the H-donor molecule respect to the ozone moiety in the molecular plane (for the δ_0 case, the motion of the ozone molecule is practically negligible). Two out-of-plane bending modes are marked as π_1 and π_2 , and finally, ν holds for the $XH\cdots O_3$ stretching frequency.

The overall effect of BSSE correction is similar to the effect of improving the basis set; for instance, for the 6-31G(d,p) basis set, the δ_0 frequency is smaller than one of the out-of-plane bending (π_1) and the $CNH\cdots O_3$ stretching (ν). Correcting by CP or improving the basis set both reverse this trend, so their effect is similar.

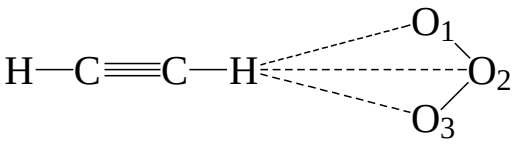
Regarding the intramolecular modes, the uncorrected and CP-corrected frequencies are very similar, in agreement with previous results for other intermolecular complexes[48].

The comparison of calculated (uncorrected) and experimental vibrational frequencies was already carried out by Turi and Dannenberg[119]. The same considerations hold after CP-correction, as all experimental frequencies were assigned to intramolecular modes, and the CP correction has proved to have no meaningful effect on them.

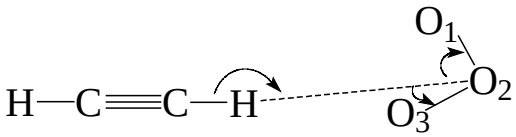
[121] Each vibrational mode contributes with ca. 0.6kcal/mol (kT) to the thermal correction to the enthalpy

II.1.1.3.b The $HCCH\cdots O_3$ complex

For this system, two planar geometries have been considered again. First, a symmetric structure presenting two H···O interactions, and second a non-symmetric structure (see SCHEME 12). The features of this system turn out to be similar to those of $CNH\cdots O_3$ complex in terms of geometries and vibrational analysis. Most details discussed above can apply to the acetylene complex. However, the acidic character of the hydrogen atom of the acetylene is not so strong as that of the hydrogen cyanide hydrogen, and thus the intermolecular interaction is weaker. Probably due to this situation, we know of no experimental data for this complex. In practice, we studied the same two planar conformations considered for the $CNH\cdots O_3$ complex, namely the symmetric and the non-symmetric.



Symmetric



Non-symmetric

SCHEME 12: $HCCH\cdots O_3$ structures

The results collected on Tables XVIII-XX show trends similar to those of the first complex studied. With the 6-31G(d,p) basis set, symmetric stationary points have been located on the uncorrected PES at both the HF and MP2 levels of theory. CP-correction changes this situation and non-symmetric minima are obtained in all cases. Inclusion of diffuse functions dramatically decreases the BSSE effect on the PES.

TABLE XVIII.

Comparison of the geometric parameters of HCCH···O_3 on the normal surface, the CP-optimized surface and the symmetric CP-optimized. Distances in Å and angles in degrees.

		6-31G(d,p)		D95++(d,p)	
		HF	**MP2**	**HF**	**MP2**
RH_1O_1	Normal	2.888	2.642	3.420	3.271
	CP opt	3.358	3.203	3.472	3.277
	CP opt (sym.)	3.002	2.819	3.067	2.888
RH_1O_2	Normal	3.307	3.086	3.531	3.357
	CP opt	3.474	3.507	3.550	3.434
	CP opt (sym.)	3.430	3.282	3.499	3.357
RH_1O_3	Normal	2.888	2.642	2.794	2.529
	CP opt	2.747	2.579	2.782	2.659
	CP opt (sym.)	3.002	2.819	3.067	2.888
$a(H_1O_2O_1)$	Normal	59.5	58.1	74.8	75.0
	CP opt	74.4	72.2	76.5	72.1
	CP opt (sym.)	59.4	58.1	79.5	58.1
$a(H_1O_2O_3)$	Normal	59.5	58.1	44.2	41.2
	CP opt	44.5	44.0	42.5	44.1
	CP opt (sym.)	59.4	58.1	59.5	58.1
$a(C_1H_1O_2)$	Normal	180.0	180.0	174.5	180.0
	CP opt	171.5	168.9	172.5	170.9
	CP opt (sym.)	180.0	180.0	180.0	180.0

Compared to the hydrogen cyanide-ozone complex, the uncorrected intermolecular distances are systematically larger, whereas the stabilization energy decreases. However, all different correction terms to the energy, namely the CP and SP, ZPVE and enthalpy corrections, are of the same magnitude, so the corresponding corrected stabilization energies are less negative and even in some cases, e.g., the enthalpy-corrected ones, the interaction turns out to be repulsive.

TABLE XIX.

Energetic results of HCCH···O_3 complex. The values labelled ZPVE corr. and Enthal. corr. hold for the stabilization energy including the ZPVE correction and both the ZPVE and the thermal correction to the enthalpy, respectively.

Basis Set	**Method**	**Geom.**	**Stabilization energy (Kcal/mol)**			**ΔE_{O3}**	**ΔE_{HCCH}**
			Electronic	ZPVE corr	Enthal. corr	**(Kcal/mol)**	
6-31G(d,p)	HF[a]	Symm.	-1.37	-0.82	-0.34		
	SP-HF[b]	Symm.	-0.55	0.00	0.48	0.80	0.02
	CP-HF[c]	No Symm.	-0.60	-0.07	0.41	0.59	0.02
	CP_{sym}-HF[d]	Symm.	-0.57	0.01	0.46	0.76	0.02
	MP2[a]	Symm.	-2.92	-2.16	-1.76		
	SP-MP2[b]	Symm.	-0.83	-0.07	0.33	1.98	0.11
	CP-MP2[c]	No Symm.	-0.97	-0.22	0.18	1.49	0.07
	CP_{sym}-MP2[d]	Symm.	-0.93	-0.12	0.25	1.80	0.06
d95++(d,p)	HF[a']	No Symm.	-0.82	-0.39	0.09		
	SP-HF[b']	No Symm.	-0.58	-0.15	0.33	0.09	0.15
	CP-HF[c']	No Symm.	-0.58	-0.15	0.33	0.09	0.14
	CP_{sym}-HF[d]	Symm.(TS)	-0.55	-0.18	*-0.23*	0.10	0.16
	MP2[a']	No Symm.	-1.75	-1.21	-0.76		
	SP-MP2[b']	No Symm.	-0.62	-0.08	0.36	0.75	0.37
	CP-MP2[c']	No Symm.	-0.67	-0.21	0.30	0.69	0.35
	CP_{sym}-MP2[d]	Symm.(TS)	-0.67	-0.27	*-0.30*	0.66	0.37

[a] Point a in SCHEME 11a
[b] Point d in SCHEME 11a
[c] Point c in SCHEME 11a
[d] Point c in a symmetric surface
[a'] Point a in SCHEME 11b
[b'] Point d in SCHEME 11b
[c'] Point c in SCHEME 11b

The energetic difference between the symmetric and non-symmetric structures on the CP-corrected PES is also inverted when the ZPVE correction is included. Larger differences on the enthalpy-corrected energies are found for the d95++(d,p) calculations for the same reasons brought above for NCH···O_3.

TABLE XX

Comparison of the vibrational frequencies of HCCH···O_3 on the normal surface, the CP-optimized surface and the symmetric CP-optimized surface. (cm^{-1})

		6-31G(d,p)		D95++(d,p)	
		HF	**MP2**	**HF**	**MP2**
Normal		33	44	21	17
CP opt	δ	27	28	22	8
CP opt (sym)		35	39	*24i*	*30i*
Normal		41	49	48	49
CP opt	π_1	51	65	43	48
CP opt (sym)		57	68	50	50
Normal		51	74	42	58
CP opt	ν	46	61	47	45
CP opt (sym)		43	54	36	42
Normal		58	67	57	72
CP opt	δ_0	63	78	60	60
CP opt (sym)		61	72	44	32
Normal		118	130	102	110
CP opt	π_2	113	127	100	108
CP opt (sym)		123	133	104	110
Normal		817	175	784	541
CP opt		816	490	783	537
CP opt (sym)		819	495	781	533
Normal		821	512	786	542
CP opt		818	510	786	538
CP opt (sym)		823	521	786	537
Normal		855	733	848	731
CP opt		854	732	847	731
CP opt (sym)		856	735	849	733
Normal		896	779	872	747
CP opt		896	788	871	743
CP opt (sym)		897	785	866	736
Normal		910	823	876	750
CP opt		904	812	875	746
CP opt (sym)		910	819	874	744
Normal		1449	1174	1433	1183
CP opt		1446	1173	1433	1182
CP opt (sym)		1447	1173	1434	1183
Normal		1537	1997	1547	1951
CP opt		1539	1996	1547	1951
CP opt (sym)		1536	1996	1544	1951

TABLE XX (Cont.)

	6-31G(d,p)		D95++(d,p)	
	HF	**MP2**	**HF**	**MP2**
Normal	2241	2394	2206	2444
CP opt	2240	2391	2205	2443
CP opt (sym)	2241	2394	2206	2446
Normal	3583	3496	3567	3451
CP opt	3581	3492	3568	3452
CP opt (sym)	3582	3494	3570	3454
Normal	3695	3584	3679	3544
CP opt	3693	3581	3680	3545
CP opt (sym)	3694	3583	3681	3546

Concerning the vibrational features, it is interesting to see that the frequency labeled δ_0, in this case assigned to the oscillation of the HCCH molecule respect to the ozone, is larger than both π_1 and ν for the HF/6-31G(d,p) calculation, contrary to the behavior found in the NCH···O_3 complex.

II.1.1.3.c Single vs. Multiple PES well

One may wonder why and under which circumstances a single well on the PES transforms into a double well (or shallow multiple well) when either BSSE is eliminated at the geometry level or the basis set size is increased. To get a better insight into this point, in the last column of Tables XVI and XIX, we collect the energy differences between the fragment calculation with its own basis set and with the basis functions of the whole complex for the ozone and the corresponding donor molecules, labeled ΔE_{O3} and ΔE_{CNH} or ΔE_{HCCH} , respectively. The energy differences are calculated at the uncorrected, CP-corrected and CP_{sym}-corrected stationary points. Note that the sum of both energy differences yields the BSSE value. The bigger the energy difference, the larger BSSE-like delocalizations are found for the corresponding fragment. From the reported results several conclusions can be drawn; first of all, the most relevant contribution to the BSSE term is caused by the partial delocalizations of the ozone orbitals onto the donor molecule atomic orbitals, mainly due to the proximity of the s and p hydrogen atomic orbitals to the oxygen atoms labeled O_1 and O_3.

At the MP2 level of theory, the BSSE extent is much more important. In this case, besides the artificial stabilization of the fragment molecular orbitals as a consequence of BSSE-like delocalizations, a large number of virtual orbitals contribute to the MP2 correction energy term when the whole basis set is used for the fragment calculations. Comparing the values obtained with the 6-31G(d,p) and d95++(d,p) basis sets, the main contribution at the MP2 level of theory is found for the largest basis set.

A closer look at the reported data shows that the ΔE_{O3} is strongly dependent on the geometrical structure of the complex when using the 6-31G(d,p) basis set. The energetic stabilization due to BSSE is ca. 0.2kcal/mol smaller for the non-symmetric structures with respect to the symmetric ones. At the MP2 level of theory, the difference increases up to 0.4kcal/mol. However, the ratio between the MP2 and HF ΔE_{O3} values ranges between 2.3-2.7, independently of the geometry complex. When diffuse functions are included in the basis set, the dependence on the geometry is very small, even at the MP2 level. Again, the ratio between the MP2 and HF ΔE_{O3} energies is rather constant in all cases. Thus, it seems that the energy stabilization of the ozone molecule on the ghost-orbital calculation at the MP2 level with respect to the HF level is independent of the molecular geometry considered.

All in all, the underlying reason for the different topology of the uncorrected and the CP-corrected PES obtained with the 6-31G(d,p) basis set is the strong dependence of the ΔE_{O3} value on the molecular geometry.

The behavior of the ΔE_{O3} can be understood by an analysis of atomic orbital interplay. In the symmetric case, both O_1 and O_3 atoms are close to the hydrogen atom of the donor, so in the ghost-orbitals calculation of the ozone molecule, the p atomic orbitals of the hydrogen atom play an important role on BSSE-like delocalizations[122].

However, in the non-symmetric structure, the O_3 atom is closer to the hydrogen atom but O_1 is too far to interact with the hydrogen atomic orbitals. The final balance is that the difference on the BSSE term is large enough to compensate for the energy penalty caused by loss of symmetry, due to the flatness of the PES. Therefore, the CP-corrected PES exhibits a second energy minimum, which is not symmetric.

[122] The use of the 6-31G(d) basis would probably decrease the BSSE extent for this system. However, the lack of p orbitals on the hydrogen would also rebound on the correct description of the whole complex.

The inclusion of diffuse basis functions prevents this problem indeed. The BSSE extent does not depend on the angular disposition of the fragments, but mainly on the intermolecular rH-O_3 distance. In this case, both the uncorrected and the CP-corrected show similar topologies, the symmetric structure being saddle-point.

As a conclusion, the CP correction has proved to be important at the geometry level for the two complexes studied with basis sets having no diffuse functions. In these cases, the uncorrected optimizations lead only to symmetric structures, whereas the CP-corrected PES exhibit both the symmetric and the non-symmetric geometries. Despite the large differences in the intermolecular parameters, the CP correction only modifies meaningfully the intermolecular vibrational frequencies, so that the ZPVE and thermal corrections to enthalpy are almost not modified.

Numerical dispersion on the reported stabilization energies decrease upon CP correction. SP and CP-corrected stabilization energies show very similar values for the complexes studied. For such weakly bonded systems, the CP correction applied to the geometry is not mandatory when one is interested merely on stabilization energies.

However, the BSSE correction term has proved to be strongly dependent on the molecular geometry. The large energy stabilization in the ghost orbital calculations for the ozone molecule at the symmetric arrangement of the complex accounts for the overestimation of the stabilization energy at this geometry with respect to a non-symmetric one, which is hidden by BSSE. When BSSE is taken into account by the CP correction, both geometric minima are located on the PES, stabilizing slightly the non-symmetric geometry with respect to the symmetric one.

When diffuse functions are included the situation changes, so both the uncorrected and the CP-corrected PES show the same double-well profile. The BSSE correction term does not depend on the angular features of the complex. In this case, the CP_{sym} structures turn out to be saddle points connecting two equivalent non-symmetric minima. This situation suggests that, using the 6-31G(d,p) basis set, the CP-correction may not fully account for BSSE; the PES is modified, as non-symmetric structures are found, but the topology of the symmetric stationary points remains unchanged.

The most important fact is that the corresponding CP-corrected results reproduce, at least qualitatively, the PES obtained with larger basis including diffuse functions. Thus, the CP correction involving geometry reoptimizations can be a very useful tool in order to yield good descriptions of larger molecular aggregates (i.e., crystal structures or

biomolecules), where the use of large basis sets including diffuse functions is at the moment prohibitive.

II.1.1.4 Charged intermolecular complexes

In Sections II.1.1.1 to II.1.1.3, the validity of the counterpoise-correction in the study of intermolecular complexes has been shown. The interpretation of the correction term as a correction to the complex's description has proved to be extremely useful in order to study BSSE-corrected PES by means of analytical derivatives.

However, as noted in the introduction, under certain circumstances the counterpoise-correction may fail to describe the molecular complex due to the limitations of a purely correction method.

In this section we want to explore the limits of the counterpoise method by performing a systematic study, using several basis sets and *ab initio* methods, of charged molecular complexes. We have studied three charged closed shell molecular complexes: the deprotonated and protonated water dimers and the protonated ammonia-water complex. These are quite strong hydrogen bonds, the so-called charge-assisted hydrogen bonds, where the molecular charge enhances the interaction. They behave, however, in a quite different way. The deprotonated water dimer shows a non symmetric interaction. The geometry of the protonated water dimer, contrarily, is symmetric, with the central proton equally shared by two water molecule moieties. Thus, the charge delocalizations for this complex is much larger than that of the deprotonated water dimer. Finally, in the ammonium-water complex, the central proton involved in the intermolecular bond is clearly attached to the ammonia moiety and the positive charge is mainly localized in one of the fragments.

Since the definition of the fragments is ambiguous in all the cases, we have tried two different strategies to compute the counterpoise correction. For the three model systems studied, i.e. $(H_2O\cdots OH)^-$, $(H_2O)_2^+$ and $(NH_4\cdots H_2O)^+$, the counterpoise method has been applied in two different ways, defining two and three fragments, respectively. In one case we have defined two fragments for each complex (henceforth CP2), arbitrarily assigning the charge of the complex to one of them. The corresponding fragments are H_2O and OH^- , H_2O and H_2O^+ ,and NH_4^+ and H_2O, respectively. On the other hand, one can define three fragments for each system (henceforth CP3), i.e., $2\times OH^-$ and H^+, $2\times H_2O$ and H^+, and NH_3, H^+ and H_2O, respectively In all the cases, there is a ghost fragment corresponding to the hydrogen atom involved in the hydrogen bond. This fragment has no electrons and therefore only the two other fragments contribute to the counterpoise

correction.

Note that in all cases, the fragments have been chosen as closed-shell so that no extra ambiguities due to the use of open-shell fragments will be introduced. For each complex, three PES have been studied at each level of theory. The uncorrected results have been obtained by means of standard gradient optimization techniques. The counterpoise-corrected gradients and second derivatives have been calculated using a modified version of Gaussian 98[123], which will be discussed in detail in Section II.2.1. Also, for the calculations involving two fragments but having no electrons one of them, a our program to externally drive Gaussian has also been used (see Section II.2.1).

We have used the 6-31G, 6-31G(d,p), 6-31++G(d,p), 6-311G(d,p), 6-311++G(d,p), d95(d,p), d95++(d,p), and 6-311++G(3df,2pd) basis sets, together with the *ab initio* methods Hartree-Fock (HF) and second order Møller-Plesset perturbation theory (MP2), where all orbitals have been included for the correlation.

The results obtained for each complex will be discussed separately. Next, we include some brief comments about the thermal correction to the energy.

II.1.1.4.a Deprotonated water dimer

The interaction of OH^- with water molecules in gas-phase has been recently studied[88d,124,125]. As for the $OH^-(H_2O)$, it has been found to form an asymmetric complex with an almost collinear hydrogen bond. The central proton can be transferred in a isomerization reaction with almost no energetic barrier. Experimental studies[126] have

[123] Gaussian 98 (Revision A.11), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J. J. Dannenberg, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, 2001.

[124] S. S. Xantheas, J. Am. Chem. Soc. 117 10373 (1995), and references therein.

[125] M. Masamura, J. Comput. Chem. 22 31 (2001).

[126] M. Meot-Ner (Mautner) and C. V. Speller, J. Phys. Chem. 90 6616 (1986)

estimated a value of 26.5 ±1 kcal/mol for the formation enthalpy (ΔH^o) of the complex.

In some of the papers devoted to this system the BSSE was neglected[88d,124] , whereas others included single-point counterpoise calculations at the (uncorrected) optimized stationary points[125].

TABLE XXI

Formation Enthalpies (kcal/mol) and H-bond distances (Å) for the $H_2O\cdots OH^-$ complex at RHF and MP2 levels of theory. The CP-corrected results (CP2 and CP3) have been obtained at the CP-corrected PES (see text for fragment definitions). The symmetric structures are designed by *s*.

Method	**Formation Enthalpy**			**rO···H distance (Å)**		
	Uncorr	***CP2***	***CP3***	***Uncorr.***	***CP2***	***CP3***
RHF/6-31G	38.41	31.39	31.05	1.400	1.457	1.445
RHF/6-31G**	34.52	27.21	26.44	1.459	1.494	1.501
RHF/6-31++G**	23.78	22.89	22.98	1.569	1.578	1.577
RHF/6-311G**	33.74	24.42	23.93	1.501	1.566	1.567
RHF/6-311++G**	23.23	22.43	22.46	1.581	1.588	1.586
RHF/D95 **	31.21	27.66	26.91	1.452	1.489	1.487
RHF/D95++**	23.78	22.95	23.04	1.567	1.571	1.570
RHF/6-311++G(3df,2pd)	22.77	22.10	22.10	1.555	1.566	1.562
MP2/6-31G	45.15	32.79	33.24	1.244 *s*	1.435	1.362
MP2/6-31G**	44.52	30.96	31.92	1.223 *s*	1.388	1.256 *s*
MP2/6-31++G**	28.67	24.64	25.46	1.396	1.487	1.412
MP2/6-311G**	45.17	26.39	26.56	1.216 *s*	1.494	1.418
MP2/6-311++G**	28.60	23.69	24.31	1.380	1.515	1.468
MP2/D95 **	40.30	31.94	33.15	1.217 *s*	1.382	1.235 *s*
MP2/D95++**	28.83	24.69	25.88	1.405	1.489	1.430
MP2/6-311++G(3df,2pd)	29.23	25.76	26.66	1.332	1.439	1.384

Our BSSE-corrected geometry optimizations results are gathered in the Table XXI.

The variation of ΔH^o and the intermolecular hydrogen bond distance with respect to the basis sets and methods used are depicted also in Figures 10 and 11, respectively.

As expected, after CP correction the hydrogen bond length increases and thus the stabilization energy lessens.

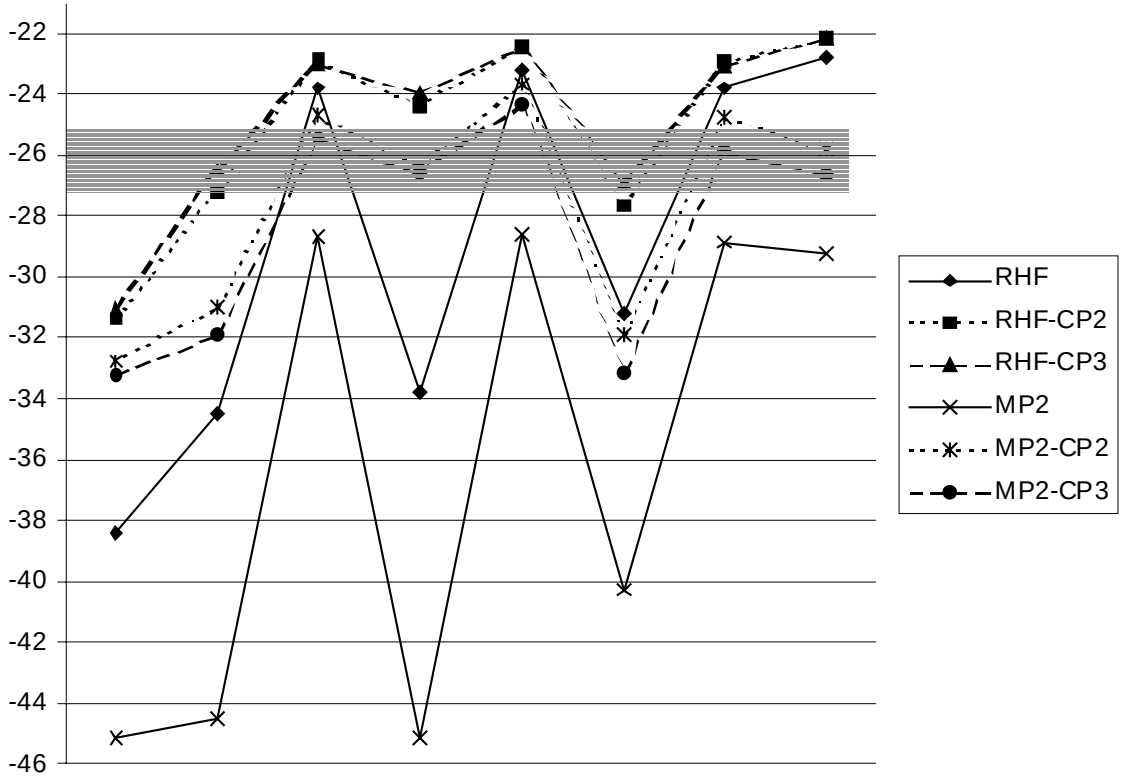


FIGURE 10: Convergence of RHF and MP2 calculated formation enthalpies (ΔH^o) in kcal/mol of the $(H_2O\cdots OH)^-$ complex vs. basis set. The shadowed region shows the experimental range

At the RHF level of theory, the effect of the BSSE in the hydrogen bond length is not dramatic even for the smallest basis set used. This is not the same case, however, for ΔH^o. When diffuse functions are not included in the basis sets, the uncorrected ΔH^o values are clearly overestimated. The CP-corrected results are much closer to the experimental values and the ones obtained with the largest basis set used (see Figure 10). Also, the CP3 scheme yields lower ΔH^o values than the CP2, the difference being less than 1 kcal/mol. The differences between both CP methods and the uncorrected values are dramatically decreased when diffuse functions are included. In those cases CP2 and CP3 are within 0.1 kcal/mol, but in this case the CP2 values are lower. For the largest basis set used, the 6-

311++G(3df,2pd), the uncorrected ΔH^o is still 0.67 kcal/mol above the predicted with the CP-corrected methods, which yield almost equivalent results.

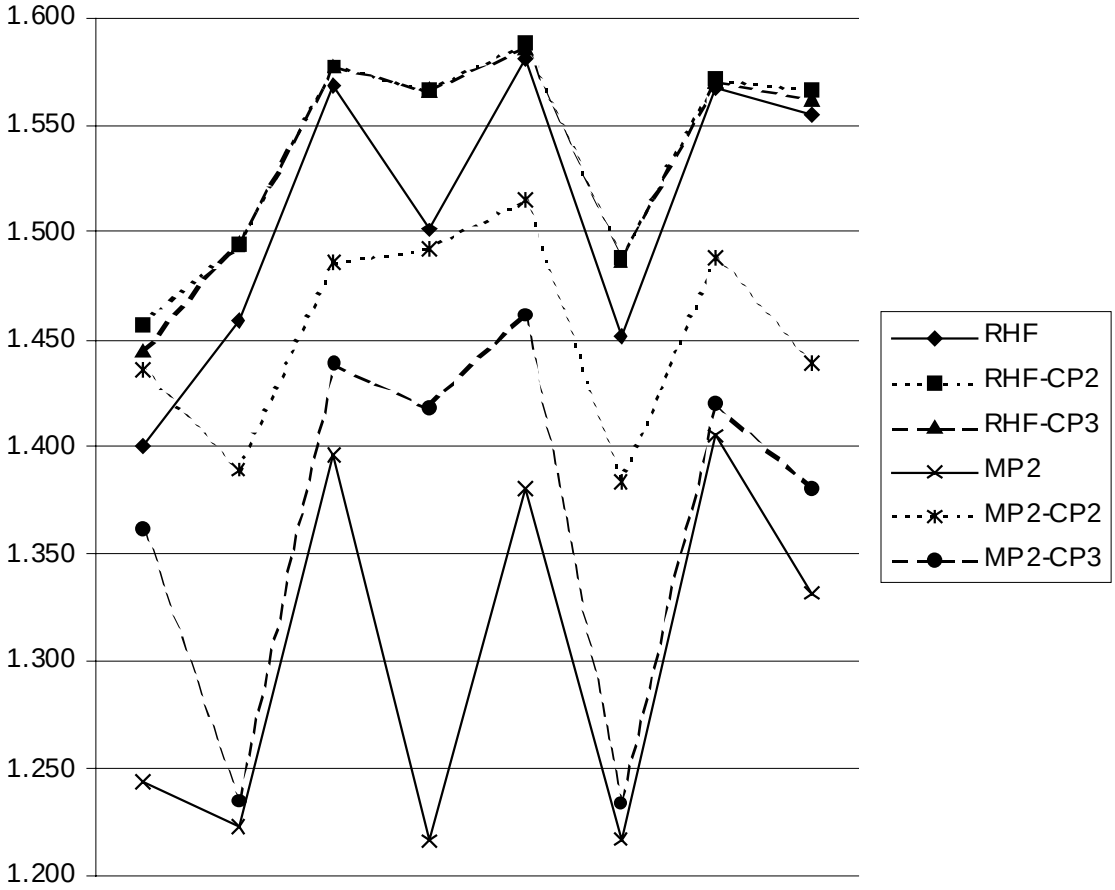


FIGURE 11: Convergence of the RHF and MP2 O-H intermolecular distances (Å) of the $(H_2O\cdots OH)^-$complex vs. basis set.

The leading term of the CP correction is by far the basis set extension effect of the OH^- moiety. In the CP2 calculations without diffuse functions this value stands for up to 95% of the BSSE. In the CP3 calculations, there are two contributions to the BSSE due to basis set extensions of two OH^- fragments. The leading term corresponds to the H-bond donor OH^- moiety, as the central proton is much closer to this OH^- fragment. Obviously, this basis set extension will depend on the position of the proton. Hence, the BSSE of the complex for the CP3 calculations is more than twice the CP2 one. The reason why this is not translated also into the ΔH^o values is that to compute the ΔH^o for the CP3 method one has to take into account the three-fragment nature of the complex. That is, the H^o values of the H_2O fragment must also be CP-corrected using the same fragment definition of the

complex, i.e. $H^+ + OH^-$.

The explanation above can be expressed qualitatively in this way. The CP-corrected absolute enthalpies can be expressed as

$$H^o_{CP2}(H_3O_2^-) = H^o(H_3O_2^-) + \delta_{CP2}(OH^-)^A + \delta_{CP2}(H_2O) \quad (95)$$

$$H^o_{CP3}(H_3O_2^-) = H^o(H_3O_2^-) + \delta_{CP3}(OH^-)^A + \delta_{CP3}(OH^-)^D \quad (96)$$

$$H^o_{CP}(H_2O) = H^o(H_2O) + \delta_{CP}(OH^-)^D \quad (97)$$

where $\delta_{CP2}(H_2O)$, $\delta(OH^-)^A$ and $\delta(OH^-)^D$ represent the positive-definite counterpoise contributions due to basis set extensions of the H_2O moiety and the OH^- H-bond acceptor and donor, respectively. Accordingly, the CP-corrected ΔH^o values for the complex can be obtained easily.

$$\Delta H^o_{CP2}(H_3O_2^-) = \Delta H^o(H_3O_2^-) + \delta_{CP2}(OH^-)^A + \delta_{CP2}(H_2O) \quad (98)$$

$$\Delta H^o_{CP3}(H_3O_2^-) = \Delta H^o(H_3O_2^-) + \delta_{CP3}(OH^-)^A + \delta_{CP3}(OH^-)^D - \delta_{CP}(OH^-)^D \quad (99)$$

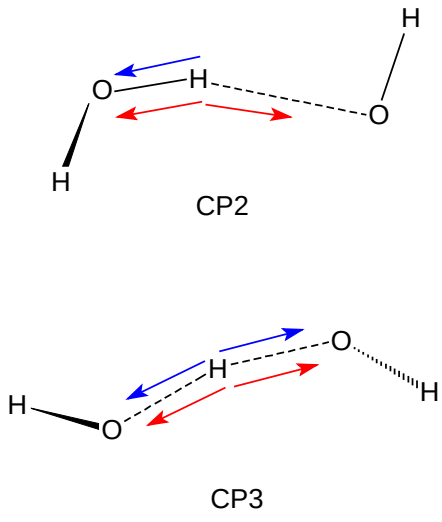
The CP2 calculation bears no BSSE in the fragments. However, the BSSE contribution of the H_2O fragment, $\delta_{CP}(OH^-)^D$, is large enough to compensate for the $\delta_{CP3}(OH^-)^D$ term of the CP3 complex calculation. The final result is that both CP2 and CP3 ΔH^o values are very close.

The qualitative differences between the results obtained with basis sets including diffuse functions can be explained as follows. When no diffuse functions are included, the intermolecular distance are ca. 0.1Å larger than the ones obtained using diffuse functions. Consequently, the central proton's basis functions are closer to the H-bond donor OH^- moiety because the intramolecular bond shortens. This means that, in the CP3 calculation, the term $\delta_{CP3}(OH^-)^D$ is greater than $\delta_{CP}(OH^-)^D$ so that the BSSE is larger than the respective CP2 one and thus the ΔH^o is smaller. When diffuse functions are included the contributions mentioned above compensate and the difference between the CP2 and CP3 values can be assigned to the $\delta_{CP2}(H_2O)$ contribution, which is no longer negligible with respect to the $\delta(OH^-)$ ones. The final consequence is that, in these cases, the CP2 ΔH^o

tend to be slightly smaller than the CP3 one.

At MP2 level of theory the picture is different. The BSSE effects are much larger both energetically and geometrically. The intermolecular distances of the CP2 optimized structures are much larger than both the uncorrected and the CP3 ones. Even for the largest basis set, the intermolecular distance lengthens by more than 0.1A. The effect of BSSE in the geometry for the CP3 calculations is, in general, less important. These large discrepancies between both CP methods from a structural point of view are not translated, however, into large differences in the formation enthalpies calculated. In all the cases, both ΔH^o values are within 0.6-1.2 kcal/mol, whereas the differences with respect to the uncorrected values range from 2.5 to up to 18 kcal/mol. Contrary to the RHF case, now the ΔH^o values are smaller than the CP3 ones, obviously determined by the enlargement of the intermolecular bond length.

From the Figure 11 it seems clear that there are two rather strange values, corresponding to the CP3 MP2/6-31G** and MP2/d95** calculations. In these particular cases, the effect of the BSSE in the geometrical parameters is almost negligible. This is quite unexpected mostly in the case of the MP2/6-31G** level of theory, which is known to bear large BSSE and even modify the topology of the PES of some weak hydrogen bonded complexes[38,39].



SCHEME 13: Qualitative picture of the acting forces on the central H atom

A more detailed inspection to the optimized structures gives the answer. When diffuse functions were not used, all the uncorrected calculations predicted stationary

points where the central H nucleus is equally shared by both OH moieties. This effect is correctly avoided in all the CP2 calculations. However, in the case of the CP3 optimizations, the already mentioned MP2/6-31G** and MP2/d95** incorrectly yield the symmetric structure. One can rationalize these findings graphically as shown in SCHEME 13. As explained before, the main contribution to the BSSE in the counterpoise-corrected calculations is due to the OH^- fragments. In the uncorrected calculation, due to the lack of diffuse functions, the O atoms tend to use the basis functions of the central H atom. As a result the atom lay in the middle of both oxygen atoms under the effect of two BSSE-like forces (lower arrows) of the same strength and opposite direction. In the CP3 calculations, these artificial BSSE forces are equally annihilated by the counterpoise forces (upper arrows) and the symmetry is maintained. However, in the CP2 calculations, it only appears the force due to one OH^- fragment, corresponding to the H-bond acceptor moiety, therefore the symmetry is broken and the correct asymmetric structure is obtained. When using diffuse functions or at the RHF level of theory, the true interaction forces are much larger than the BSSE strength, and the uncorrected optimizations lead to the correct geometries.

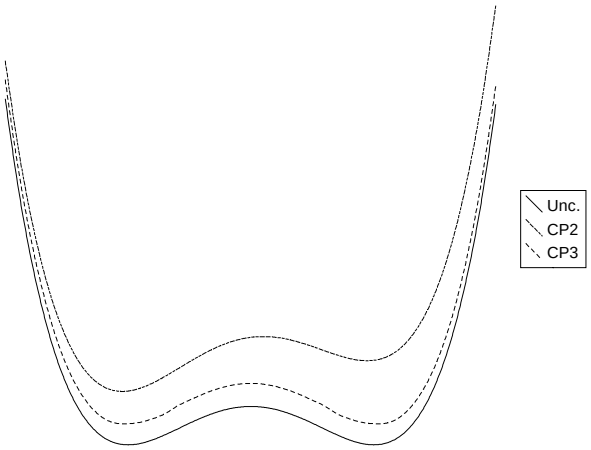
Obviously, this reasoning cannot explain why the MP2/6-31G and the MP2/6-311G** CP3 calculations do produce the correct geometries. In these cases the true interaction forces slightly overcome the BSSE effects. From another point of view, the obtention of an asymmetric structure may imply that the stability of the water molecule is enhanced with respect to the hydroxide ion at this level of theory. Indeed, at the MP2/6-311G** level, the optimized O-H distance of the water molecule is the shortest of the basis sets studied (0.957Å) and in the case of the 6-31G basis set, this distance is much larger (0.974Å) but, at the same time, it is much shorter than the OH^- bond distance (1.005 Å).

II.1.1.4.b Protonated water dimer

The protonated water dimer, $H_5O_2^+$, is one of the most studied charged hydrogen-bonded complexes. Its potential energy surface is very flat and strongly dependent of the basis sets and methodology. At the RHF level[88a], the gradient optimization leads to a C_s asymmetric structure where the hydrogen atom involved in the hydrogen bond is closer to one of the water molecules of the complex. The inclusion of electron correlation is

necessary to obtain the C_2 structure[88a,127] , where the proton is equally shared by both water molecules. Another intermediate C_1 structure was also obtained using the TZ2P basis set together with the CCSD or CCSD(T) methods. Recently, state of the art *ab initio* calculations, namely Brueckner coupled cluster[88b] or explicitly correlated methods[128] together with large basis sets removed this spurious structure and determined a global minimum of C_2 symmetry only about 0.4 kcal/mol above the C_s geometry. Furthermore, the geometrical parameters and energies seems not to smoothly converge with the increasing size of the basis sets[88b].

All these findings make this system very interesting for the purpose of this paper. Our aim is to test how the correction of BSSE, often ignored in those studies, affect the topology of this extremely sensitive PES and the convergence of the energetic and structural features of the complex. Also, contrary to the deprotonated water dimer, the best *ab initio* calculations predict a global minimum where the proton involved in the hydrogen bond is shared by both water moieties. Hence, the CP3 scheme here seems, a priori, more appropriate to preserve the symmetry of the complex.



SCHEME 14: Uncorrected, CP2 and CP3 RHF PESs for the movement of the central proton of the $(H_2O)_2 H^+$ complex.

We have performed CP-corrected geometry optimizations for both structures to determine the convergence of the energy gap between them and also possible inversion of

[127] C. Tuma, A. D. Boesse and N. C. Handy, Phys. Chem. Chem. Phys. 1 3939 (1999)

[128] A. A. Auer, T. Helgaker and W. Klopper, Phys. Chem. Chem. Phys., 2, 2235 (2000)

the stability upon BSSE correction. The ΔH^o and H-bond distances obtained for the C_2 and C_s structures are collected in Tables XXII and XXIII, respectively. The convergence of their calculated ΔH^o with respect to the basis set used is also depicted in Figure 12 and 13. When focusing with the asymmetric C_s structure both CP2 and CP3 methods can be easily applied and actually behave in a very similar way for all the basis sets. It is important to note that the CP-corrected optimization of the symmetric structure, C_2, is not possible when two fragments are defined, as the symmetry of the molecule is broken. The situation is qualitatively shown in the SCHEME 14. It represents the energy profile for the movement of the central H atom between the oxygen atoms, at the RHF level of theory. The uncorrected and CP3 curves show the typical double well PES where the C_s and C_2 structures correspond to the minimum and the TS, respectively. The CP2 method forces the PES to be asymmetric because the BSSE is larger when the H atoms is closer to the water fragment than to the H_3O^+ one. Also, the stationary point corresponding to the TS does not have the C_2 symmetry.

In principle, at any (asymmetric) point of the PES, the partition of the complex into H_3O^+ and H_2O fragments can always be done following just geometrical criteria. The central proton can be assigned to the water moiety with shorter O-H distance. That is, one can easily study half of the PES depicted in SCHEME 14. If one is interested in the whole PES, the fragment definition should change according to the relative position of the proton with respect the two water moieties. In the symmetric case, the single-point CP2 calculation is independent of the definition of the fragments, as both water molecules are equivalent. The problem is that an eventual CP2 geometry optimization of the TS at this point will push the H atom to be closer to the other water fragment, hence forcing a dynamic redefinition of the fragments. Therefore, our CP2 values shown in the Table XXII are obtained by applying the single-point correction at the respective uncorrected stationary point.

TABLE XXII

Formation Enthalpies (kcal/mol) and H-bond distances (Å) for the C_2 symmetric $(H_2O)_2H^+$ structure at RHF and MP2 levels of theory. The CP2 values have been obtained by single-point CP correction on the uncorrected PES. The transition states are designed by *i*.

Method	**Formation Enthalpy**			**H-O distance**	
	Uncorr	***CP2***	***CP3***	***Uncorr.***	***CP3***
RHF/6-31G	42.64	41.24	41.77	1.183 [a]	1.185 [a]
RHF/6-31G**	35.62	34.46	34.66	1.181 *i*	1.182 *i*
RHF/6-31++G**	32.99	31.75	31.56	1.182 *i*	1.183 *i*
RHF/6-311G**	34.87	32.4	33.12	1.179 *i*	1.180 *i*
RHF/6-311++G**	32.48	31.24	31.37	1.180 *i*	1.181 *i*
RHF/D95 **	34.29	33.82	34.13	1.181 *i*	1.181 *i*
RHF/D95++**	32.78	31.83	31.79	1.182 *i*	1.183 *i*
RHF/6-311++G(3df,2pd)	31.27	30.72	30.98	1.183 *i*	1.184 *i*
MP2/6-31G	44.79	41.87	43.30	1.200 [a]	1.208 [a]
MP2/6-31G**	39.71	36.14	37.58	1.193	1.200
MP2/6-31++G**	36.35	32.82	33.61	1.194	1.203
MP2/6-311G**	39.28	33.35	35.37	1.189	1.200
MP2/6-311++G**	35.92	31.94	33.07	1.191	1.199
MP2/D95 **	37.96	35.15	36.80	1.192	1.200
MP2/D95++**	36.11	32.49	33.82	1.194	1.206
MP2/6-311++G(3df,2pd)	35.28	33.19	34.15	1.191	1.199

[a] D_{2d} symmetry

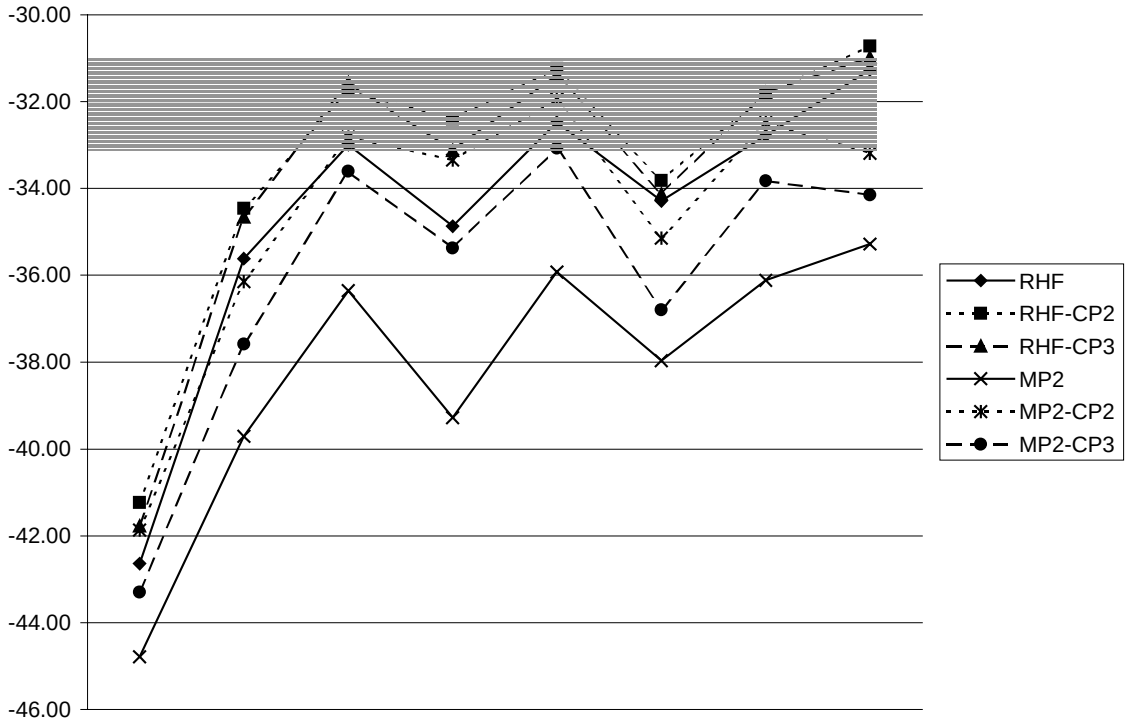


FIGURE 12: Convergence of RHF and MP2 calculated formation enthalpies (ΔH^o) in kcal/mol of the symmetric C_2 $(H_2O)_2H^+$ complex vs. basis set. The shadowed region shows the experimental range.

At the RHF level of theory, except for the 6-31G basis set (in this case a D_{2d} symmetry structure was found to be the minimum even after CP correction), the C_2 structure is a transition state. The counterpoise correction does not change the topology of the PES. The CP3 optimized O-H distances are very similar to the uncorrected ones due to the symmetry of this structure. The differences are larger for the C_s structure, the minima in all cases but the 6-31G. The effect on the intermolecular distances are not, however, as large as for the deprotonated water dimer. In some cases, for instance with the D95** basis set, the hydrogen bond distance shortens upon CP3 correction because the intramolecular O-H distance lengthens. This is because both O-H bonds can be considered as intermolecular ones when defining three fragments.

TABLE XXIII

Formation Enthalpies (kcal/mol) and H-bond distance (Å) for the C_s asymmetric $(H_2O)_2H^+$ structure at RHF and MP2 levels of theory. The transition states are designed by *i*.

Method	**Formation Enthalpy**			**H-bond distance**		
	Uncorr	***CP2***	***CP3***	***Uncorr.***	***CP2***	***CP3***
RHF/6-31G		41.14			1.215	
RHF/6-31G**	34.05	32.62	32.95	1.314	1.329	1.324
RHF/6-31++G**	31.21	29.88	29.74	1.348	1.365	1.372
RHF/6-311G**	33.14	29.99	31.15	1.342	1.384	1.366
RHF/6-311++G**	30.63	29.23	29.48	1.363	1.386	1.381
RHF/D95 **	32.46	31.82	32.13	1.341	1.347	1.336
RHF/D95++**	30.98	29.93	29.95	1.361	1.370	1.372
RHF/6-311++G(3df,2pd)	29.54	28.94	29.23	1.374	1.383	1.377
MP2/6-31G		41.98			1.255	
MP2/6-31G**	39.99	36.74	38.10	1.230 *i*	1.251 *i*	1.231 *i*
MP2/6-31++G**	36.65	33.44	33.84	1.246 *i*	1.220 *i*	1.270 *i*
MP2/6-311G**	39.45	33.88	35.76	1.244 *i*	1.193 *i*	1.261 *i*
MP2/6-311++G**	36.10	32.53	33.24	1.255 *i*	1.285 *i*	1.284 *i*
MP2/D95 **	38.21	35.82	37.39	1.247 *i*	1.229 *i*	1.234 *i*
MP2/D95++**	36.20	33.04	33.98	1.267 *i*	1.230 *i*	1.277 *i*
MP2/6-311++G(3df,2pd)	35.42	33.60	34.30	1.253 *i*	1.220 *i*	1.263 *i*

Concerning the energetics, the C_s structure is from 0.15 to 0.5 kcal/mol below the C_2 one. The inclusion of standard thermal corrections to the enthalpy (including ZPVE) inverts the relative stability of both structures. Upon correction, the calculated ΔH^o for the C_2 structure is about 1.5-2.0 kcal/mol larger. The counterpoise correction further increases the energy gap between both structures because the BSSE is larger for the C_s structure. In the CP3 method this is due to the proximity of the central H atom to one of the water

fragments. With the CP2 method, the BSSE is computed on the uncorrected stationary point, which has been proved to be the cause of the sometimes-referred overcorrecting of the BSSE by means of the counterpoise scheme[24,25].

Except for the 6-31++G** basis set, the CP2 calculated ΔH^o values are smaller than the CP3 ones. Again, in the CP3 case, one has to correct for BSSE as well the ΔH^o of H_3O^+ in order to be consistent with the number of fragments in the complex. This compensates for some of the BSSE of the complex so that the final ΔH^o values are closer to the uncorrected ones.

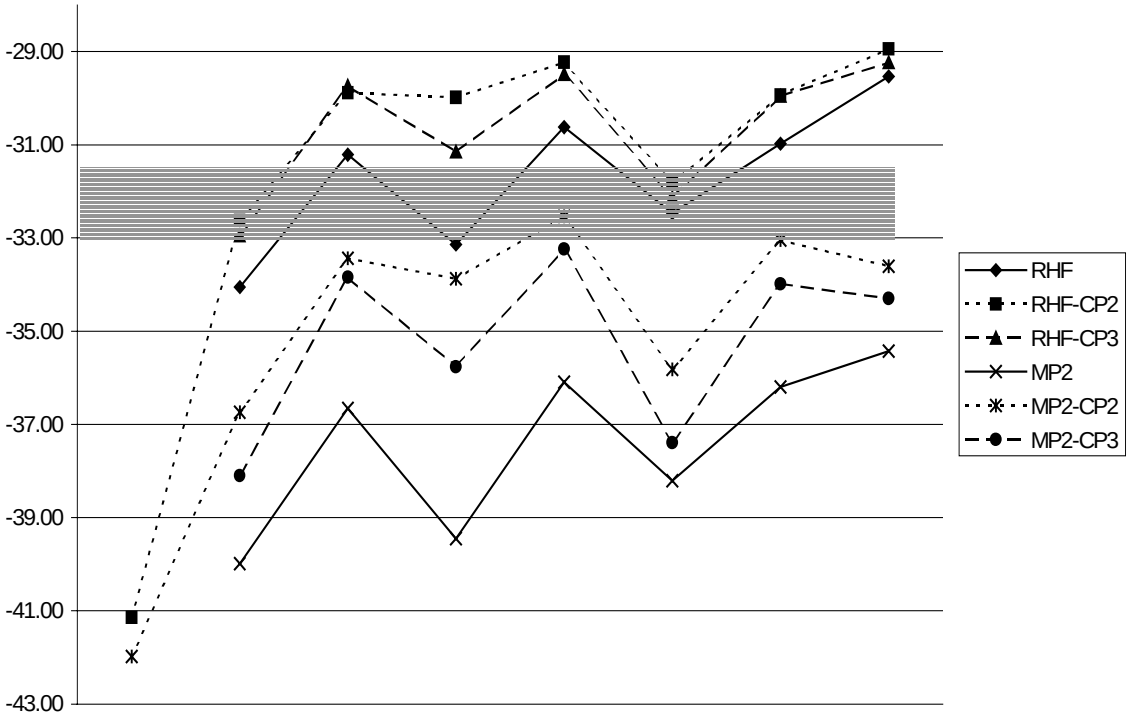


FIGURE 13: Convergence of RHF and MP2 calculated formation enthalpies (ΔH^o) in kcal/mol of the symmetric C_s $(H_2O)_2H^+$ complex vs. basis set. The shadowed region shows the experimental range.

At the MP2 level the differences are larger, as expected. The inclusion of electron correlation changes the topology of the PES and now the symmetric C_2 structure is the minima whereas the C_s stands for a transition state. The counterpoise correction does not modify the nature of the stationary points but at the CP2 level the changes in the geometrical parameters are dramatic. For the C_s structure, the intermolecular bond length shortens upon CP2 correction and the intramolecular O-H lengthens. Indeed, in some

cases like for instance with the 6-31++G** ,D95++** and 6-311++G(3df,2pd) basis sets, both O-H distances are almost equivalent. This rather strange shortening of the CP-corrected intermolecular distances was already reported in a previous paper of Kobko and Dannenberg[87]. Basically, it is due to the fact that we are dealing with a transition state instead of a minimum. The CP2 energies for the symmetric C_2 structure are computed at the uncorrected stationary points for the reasons given above. However, if one lets the molecule to fully optimize, the symmetry is broken and the central H migrates towards the oxygen atom of the H_3O^+ moiety. This spurious geometry corresponds to the C_1 spurious minimum already found by Xie et .al[88a]. The use of three fragments, more chemically meaningful, solves this problem. The CP3 geometrical parameters of both structures are more similar to the uncorrected ones and asymptotically tend to them by using diffuse functions and increasing the basis set size.

The BSSE is again larger for the CP2 calculations. The differences are larger at the C_2 geometry because the CP2 values are computed at the uncorrected stationary point. Nevertheless, the differences between both counterpoise calculations are smaller than the BSSE, even in the most favorable cases.

Similarly to the RHF results, the thermal correction to the enthalpy including ZPVE changes the relative stability of both structures and the C_s is favored with respect to the symmetrical one. The CP-correction does not change the situation even though, in the case of the CP2 method, one can reasonable expected the spurious C_1 structure be the more stable.

Experimental studies of the enthalpy of solvation of the H_3O^+ ion report values of 33.0 kcal/mol[129] and 31.6 kcal/mol[130], respectively. In general, our RHF results are closer to the experimental values than the MP2 ones. Also, the C_2 structures seem to produce ΔH^o values closer to the experiment. Figure 12 and 13 shows the variation of the computed ΔH^o vs. the increasing basis set size. Except for the 6-31G basis set, which totally failed to properly describe the system, the counterpoise-corrected MP2 values for small basis sets are much better than the uncorrected ones.

---

[129] M. Meot-Ner and F. H. Field, J. Am. Chem. Soc. 99 998 (1977)

[130] E. P. Grimsrud and P. Kerbale, J. Am. Chem. Soc. 95 7939 (1973).

II.1.1.4.c Protonated ammonia-water

The case of the protonated ammonia-water complex is also interesting because there is no possible equivalence between the fragments of the complex. Therefore, neither the CP2 ($NH_4^+ + H_2O$) nor the CP3 ($NH_3 + H^+ + H_2O$) fragment definitions will break any symmetry of the complex. It has been shown[131] that the most stable structure of this complex corresponds to a C_1 direct H-bond interaction where the central proton involved in the intermolecular bond is closer to the NH_3 moiety and forms an almost linear angle with the oxygen atom. Similarly to the protonated water dimer case, the BSSE is expected to be rather small compared to the neutral of negatively charged complexes so both CP2 and CP3 calculations should not be very different. Since the proton is clearly attached to the NH_3 moiety, it seems more coherent the CP2 partitioning of the complex. However, if one is interested in studying the proton transfer process

$$NH_4^+ + H_2O \rightarrow NH_3 + H_3O^+$$

only the CP3 method can yield a continuous PES. Hence, it is interesting to see whether the CP3 calculations can also produce meaningful results when the central H atom is clearly attached to the ammonia

The calculated ΔH^o and intermolecular O-H distance values are gathered in Table XXIV. The convergence of both parameters with respect to the basis set used is also depicted in Figures 14 and 15. Two different experimental estimations of the solvation enthalpy of the ammonium cation are 19.9 kcal/mol[132] and 17.2 kcal/mol[133], respectively. The two counterpoise-corrected calculations nicely converge for all the basis sets from both an energetic and structural point of view (see Figures 14 and 15). The effect of the BSSE is not dramatic either. Maximum differences of 3 kcal/mol and 0.05Å are observed at the MP2 level of theory for ΔH^o and O-H distance, respectively.

---

[131] J. C. Jiang, H.-C. Chang, Y. T. Lee and S. H. Lin, J. Phys. Chem 103 3123 (1999)

[132] M. Meot-Ner, J. Am. Chem. Soc. 106 1265 (1984)

[133] P. Kerbale, Annu. Rev. Phys. Chem. 16 267 (1977)

TABLE XXIV

Formation Enthalpies (kcal/mol) and H-bond distance (Å) for the $(NH_4\cdots H_2O)^+$ complex at SCF and MP2 levels of theory. The CP-corrected results (CP2 and CP3) have been obtained at the CP-corrected PES (see text for fragment definitions).

Method	**Formation Enthalpy**			**rO···H distance (Å)**		
	Uncorr	***CP2***	***CP3***	***Uncorr.***	***CP2***	***CP3***
RHF/6-31G	25.32	25.15	24.43	1.629	1.641	1.636
RHF/6-31G**	20.14	19.96	19.11	1.719	1.730	1.729
RHF/6-31++G**	18.26	17.49	18.01	1.759	1.779	1.781
RHF/6-311G**	19.97	18.61	18.63	1.723	1.751	1.749
RHF/6-311++G**	18.06	18.01	17.99	1.751	1.767	1.767
RHF/D95 **	19.12	19.34	18.53	1.729	1.738	1.739
RHF/D95++**	18.18	18.11	17.53	1.748	1.761	1.762
RHF/6-311++G(3df,2pd)	16.93	17.32	16.76	1.748	1.754	1.752
MP2/6-31G	26.87	25.79	25.80	1.603	1.629	1.616
MP2/6-31G**	23.76	21.86	21.26	1.616	1.648	1.636
MP2/6-31++G**	21.14	19.11	19.02	1.672	1.718	1.714
MP2/6-311G**	23.86	20.86	20.45	1.607	1.675	1.666
MP2/6-311++G**	21.03	19.06	18.45	1.647	1.698	1.695
MP2/D95 **	21.91	21.17	20.90	1.632	1.660	1.652
MP2/D95++**	21.17	19.20	18.64	1.664	1.700	1.693
MP2/6-311++G(3df,2pd)	19.72	19.33	19.53	1.635	1.662	1.656

The CP-correction is very important in order to obtain more reliable ΔH^o values, closer to the experiment, mainly for the small basis sets. Again the inclusion of diffuse functions is essential for a good description and, usually a lower BSSE. Curiously, the addition of diffuse functions to the D95** basis set further increases the BSSE. Nevertheless, a smaller absolute value of BSSE does not ensure a better description of the system and indeed, this is clearly shown in Figure 14. The CP-corrected values at the MP2/D95++** are better compared to the experiment than the MP2/D95** ones.

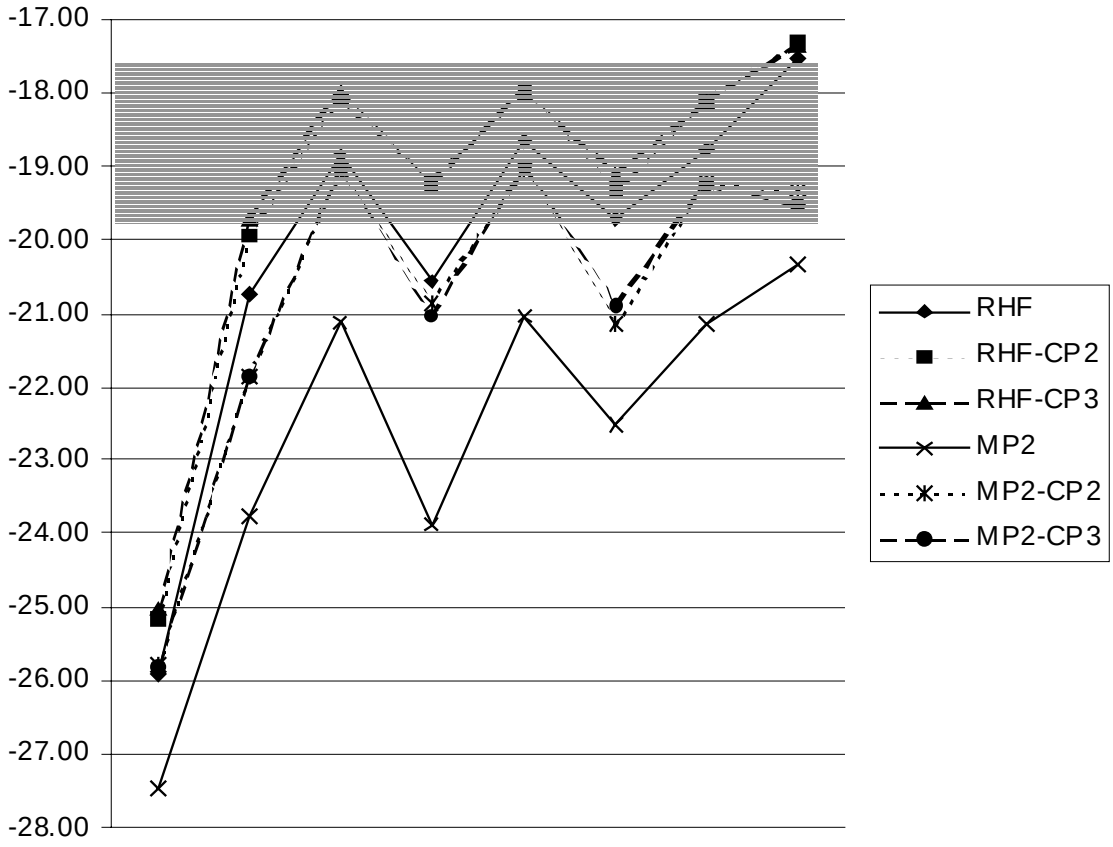


FIGURE 14: Convergence of the RHF and MP2 calculated formation enthalpies (ΔH^o) in kcal/mol of the $(NH_4 \cdots H_2O)^+$ complex vs. basis set. The shadowed region shows the experimental range.

It is noticeable also that with the largest basis set used the CP-corrected ΔH^o values are still 1 kcal/mol lower than the uncorrected ones. The 6-31G basis set again shows that the BSSE is not the most important source of error and predicts too strong interaction even after counterpoise-correction.

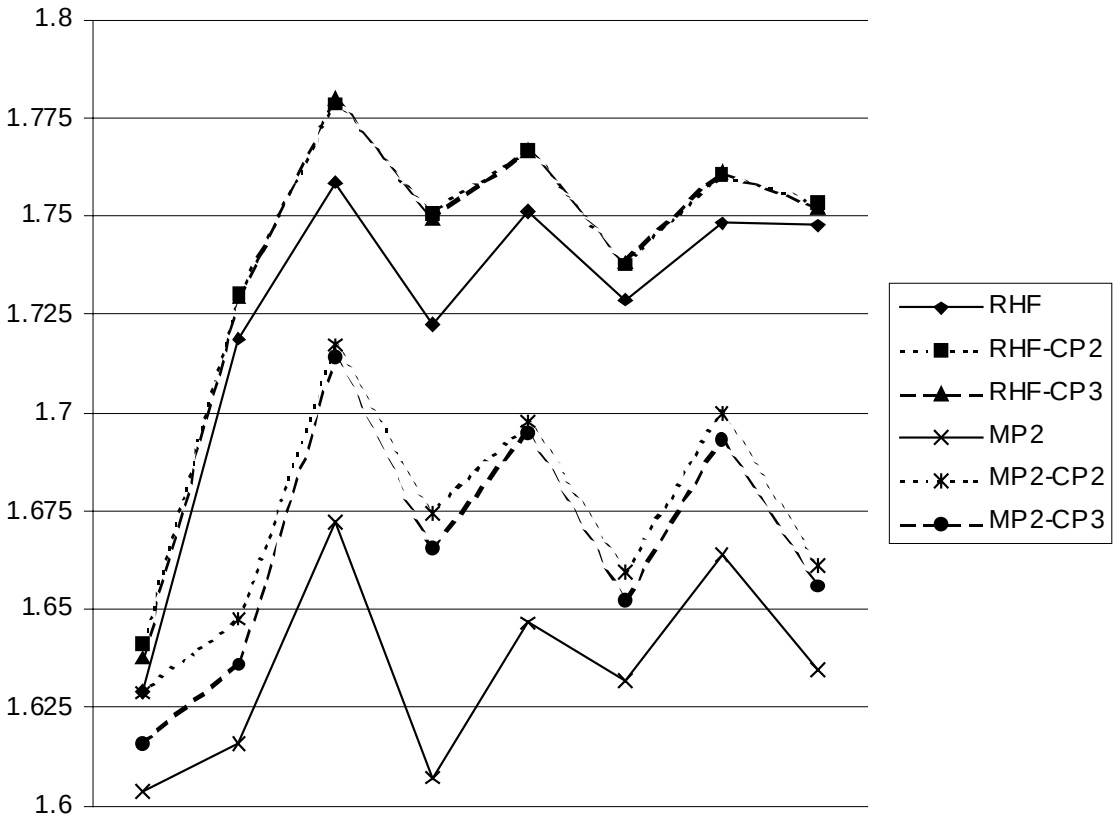


FIGURE 15: Convergence of the RHF and MP2 O-H intermolecular distances (Å) of the $(NH_4 \cdots H_2O)^+$ complex vs. basis set.

II.1.1.4.d The thermal correction to Enthalpy

The PES of the NH_4^+-H_2O complex is extremely flat. The water molecule can rotate almost freely around the intermolecular bond. For instance, at the MP2/6-31G** level, the energetic barrier for the rotation (without any thermal correction) is less than 10^{-4} kcal/mol. In most cases, the lowest vibrational frequencies for the minima and TSs are about $10cm^{-1}$ and $10i\ cm^{-1}$, respectively, which indicates that the harmonic approximation must be very rough for this PES. Indeed, the standard thermal correction to the energy sums ~0.6 kcal/mol[121] (KT) for the lowest *real* vibrational energy, whereas any imaginary vibrational frequency is ignored. Therefore, in this case, the calculation of the rotational barrier including the thermal corrections would destabilize by ca. 0.6 kcal/mol the minimum with respect to the TS. The values of ΔH^o presented in Table XXIV are

computed ignoring the contribution of the lowest vibrational frequency, i.e. they are ~0.6 kcal/mol shifted with respect to the standard thermal correction calculation. Indeed, this problem was already observed[38] for some C-H···O hydrogen bonded complexes that exhibit a double well PES.

A similar situation occurs in case of the protonated water dimer. At the MP2 level of theory, the C_s structure corresponds to a TS and hence is ca. 0.6 kcal/mol stabilized with respect to the C_2 minimum, resulting in an artificial inversion of the stability of both structures. At the SCF level of theory, the C_s structure is the minimum and the C_2 the TS. After thermal corrections to the enthalpy the relative stability was also inverted. The difference in this case is that the estimated zero-point vibrational energy is also about 2 kcal/mol larger for the C_s structure so that, even after removing the 0.6 kcal/mol due to the contribution of the lowest vibrational frequency, the inversion of stability is conserved. Furthermore, now the energy gap between the considered structures is in better agreement with the predicted at MP2 level. Since the lowest vibrational frequencies were about 150cm^{-1} and 350i cm^{-1} for the minima and the TS, much larger than the respective values for the NH_4^+-H_2O complex, the values given in Tables XXII and XXIII are computed taking into account the contribution of all the vibrational modes. Nevertheless, our calculated ΔH^o values should be taken with care.

The main conclusion that can be derived is that both CP2 and CP3 methods yield very similar results at the RHF level from both an energetic and structural point of view. Only appreciable differences in the geometry have been observed for the $(H_2O\cdots OH)^-$ complex when using the 6-311G** and D95** basis sets. At the MP2 level of theory the differences increase but, in general, are less important than the BSSE itself. There is a good agreement in the formation enthalpies for non symmetric complexes, like the $(H_2O\cdots OH)^-$ and the $(NH_4\cdots H_2O)^+$.

The protonated water dimer shows more difficulties due to the symmetric nature of the complex but the CP3 method correctly eliminates the BSSE. The counterpoise-corrected results are much closer to the experimental values than the uncorrected ones. The CP2 method shows some limitations when dealing with symmetric complexes. In this case, the definition of two fragments allows the use of the single-point correction but do prevents form counterpoise-corrected optimizations due to the break of he symmetry.

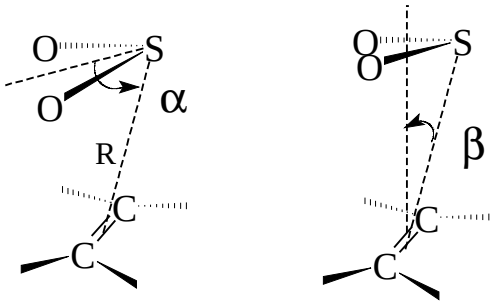
With small basis sets without diffuse functions the BSSE correction is essential. The

bad results of the 6-31G basis set even after CP-correction show that the surprisingly good results that sometimes are obtained must arise from an astonishing error compensation.

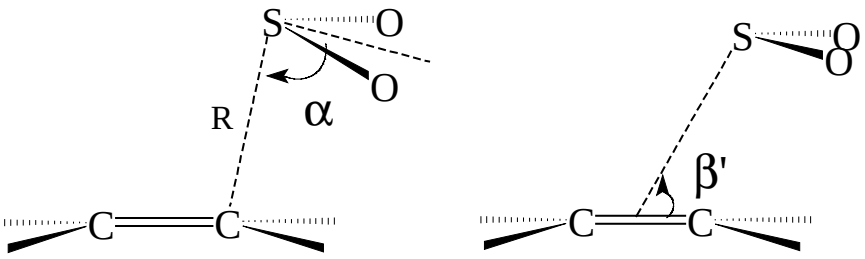
The use of single-point counterpoise-correction, that is the counterpoise correction to the stabilization energy at stationary point of the uncorrected PES, yields larger BSSE values. This is the reason why some authors claim for an overcorrection of the BSSE. We show that the counterpoise-corrected values are better than the uncorrected ones provided that the corrected PES is used throughout.

Finally, the CP3 methods has been shown to perform equally well for both symmetric and non symmetric complexes. Hence, it may become a useful receipt in order to deal with the BSSE in a continuos way for proton transfer reactions but further systematic studies on other reactions are desirable.

II.1.2 Chemical Processes and Reactivity



Minimum



Transition State

.Structures involved in the internal rotation of the $C_2H_4\cdots SO_2$. complex. The considered intermolecular parameters are also indicated.

II.1.2.1 Internal rotation barriers

The problematic of the BSSE correction on chemical reactivity has already been stressed. With the present methodology (a new heuristic approach in this direction is proposed in this work), only chemical processes where the number and of fragments is conserved along the reaction can be treated. This is the case of the internal rotation barriers of weakly bonded complexes.

In the literature there has been some controversy about the existence of BSSE in the determination of the energetic barriers for the internal rotation, as the structures involved in the calculation, the minima and the transition state connecting them, are evaluated with the same basis set.

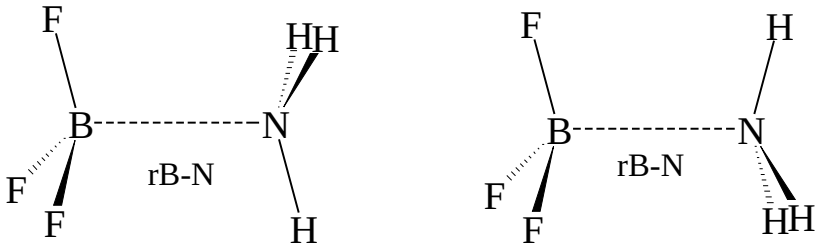
In the following section we will show the results obtained for a couple of intermolecular complexes already studied by Sordo *et al.*[92,93], where the use of the counterpoise-correction was criticized.

The two molecular complexes we are concerned are the $BF_3\cdots NH_3$ and $C_2H_4\cdots SO_2$. In our opinion the BSSE must be taken into account in the calculations of the energetic barriers to internal rotation and, again, not only single-point corrections to en energy but full counterpoise-corrected optimizations in order to avoid the BSSE dependence on the molecular geometry.

Thus, we have performed both uncorrected and CP-corrected geometry optimizations for the two weakly bonded subsystems and we will compare our values obtained for the energy barriers using Eq. (61) and Eq. (62) with those obtained by Sordo[92]. We have analyzed also the differences between the single-point counterpoise correction, i.e. the energy correction at the uncorrected PES (SP-correction), and the CP-corrected optimization. Finally, we decided to study the effect of the fragment relaxation to both the energy and the geometry for these systems, in order to shed light into Sordo's results and criticism of the counterpoise method[94]. Therefore we have re-optimized the geometry with the intramolecular parameters frozen at the optimum value of the free monomers.

II.1.2.1.a $BF_3\cdots NH_3$ complex.

SCHEME 15 depicts the structures involved in the internal rotation along the C_3 axis for the $BF_3\cdots NH_3$ complex. Tables XXV and XXVI collect the total energies and relevant geometric parameters for the minimum and transition state, respectively. In both cases, uncorrected and CP-corrected optimizations have been carried out. Uncorrected numbers are in perfect agreement with those obtained by Sordo. As expected, the CP-corrected intermolecular distance (rB-N) is larger than the uncorrected one. Differences of the order of 0.05Å are observed for both the eclipsed and alternated geometry. However, the effect of BSSE on the intramolecular distances and angles is very small. A maximum difference of 0.7 degrees and 0.003 Å is found for the aBFN angle and rB-F distance, respectively in the TS structure. SP-correction overcorrects the BSSE by about 0.2-0.3Kcal/mol.



Minimum Transition State

SCHEME 15: Structures involved in the internal rotation of the $BF_3\cdots NH_3$ system. The intermolecular distance is indicated

One of the main goals of this paper is the analysis of the relaxation term in both the uncorrected and CP-corrected energy and geometry. In Tables XXV and XXVI we also present the result of the optimization of the system with intramolecular variables fixed to the values they exhibit in free fragments (e.g., BF_3 being planar). Thus, intermolecular interaction is studied also, by keeping fixed intramolecular parameters, to assess the importance of fragment geometry relaxation on the intermolecular geometries and energetics.

The relaxation energy is BSSE-independent in this system. The energy penalty falls in the range of 54-59 cm^{-1} at both the minimum and TS geometry, independently of the CP-correction being applied. However, the effect on the B-N distance is dramatic. For both structures, rB-N is lengthened by about 0.7-0.8Å when the fragments are not allowed to relax. The CP-correction increases this number even

more. Thus, the BSSE effect on the geometry is more important without the relaxation term, differences of *ca.* 0.13 being observed here.

TABLE XXV.

Geometrical parameters, electronic energies and relaxation contribution for the minimum energy structure of $BF_3\cdots NH_3$ complex calculated at MP2, SP-MP2 and CP-MP2 levels of theory. The second half collects the values obtained in the calculations where the intramolecular parameters where frozen to the values they have in the free monomers. Relaxation energy is calculated using Eq.(4).

Method	**Energy (a.u.)**	**rB-N (Å)**	**rB-F (Å)**	**aFBN (deg.)**	**rN-H (Å)**	**aHNB (deg.)**	**E_{rel} (cm^{-1})**
MP2/ d95 (d,p)	-380.294 541 2	1.671	1.377	104.1	1.019	110.5	59
SP-corrected	-380.281 615 0						
CP-MP2/d95 (d,p)	-380.282 071 8	1.725	1.375	103.5	1.018	110.4	55
MP2/ d95 (d,p)	-380.268 241 7	2.383	1.321*	90.0*	1.015*	112.2*	0**
SP-corrected	-380.263 448 6						
CP-MP2/d95 (d,p)	-380.263 922 7	2.519	1.321*	90.0*	1.015*	112.2*	0**

*Optimized parameters for BF_3 and NH_3 systems. **Zero by definition.

TABLE XXVI.

Geometrical parameters, electronic energies and relaxation contribution for the transition state structure of $BF_3\cdots NH_3$ complex calculated at MP2, SP-MP2 and CP-MP2 levels of theory.

Method	**Energy (a.u.)**	**rB-N (Å)**	**rB-F (Å)**	**aFBN (deg.)**	**rN-H (Å)**	**aHNB (deg.)**	**E_{rel} (cm^{-1})**
MP2/ d95 (d,p)	-380.292 457 2	1.687	1.377	104.2	1.018	110.4	58
SP-corrected	-380.279 754 8						
CP-MP2/d95 (d,p)	-380.280 267 3	1.747	1.374	103.5	1.018	110.4	54
MP2/ d95 (d,p)	-380.267 710 8	2.411	1.321*	90.0*	1.015*	112.2*	0**
SP-corrected	-380.263 299 4						
CP-MP2/d95 (d,p)	-380.263 730 3	2.533	1.321*	90.0*	1.015*	112.2*	0**

*Optimized parameters for BF_3 and NH_3 systems. **Zero by definition.

Table XXVII shows the calculated values for the barrier height to internal rotation. It can be seen that it is not strongly affected by the CP-correction. That means BSSE-contamination is similar for both the minimum and the TS. Using Eq. (61) we obtained 457 and 408 cm^{-1} for the uncorrected and SP-corrected energies, respectively, in good agreement with Sordo's reported values[92]. Use of the CP-corrected receipt (Eq. (62)) decreases that value to 396 cm^{-1}. The effect of the

relaxation term is again very important. The calculated values are much smaller when the fragments are undistorted than in the full optimization. The uncorrected value decreases to 117cm^{-1} whereas for SP- and CP-corrected rotations the barrier height is only 33 and 42 cm^{-1}, respectively. Note that whereas the relaxation energy at both the eclipsed and alternated geometries is *ca.* 58 cm^{-1}, the difference in the barrier height turns out to be *ca.* 350 cm^{-1}.

TABLE XXVII

Electronic energies (cm^{-1}) for the internal rotation barrier in the $BF_3\cdots NH_3$ complex. In parenthesis the values obtained neglecting the relaxation term.

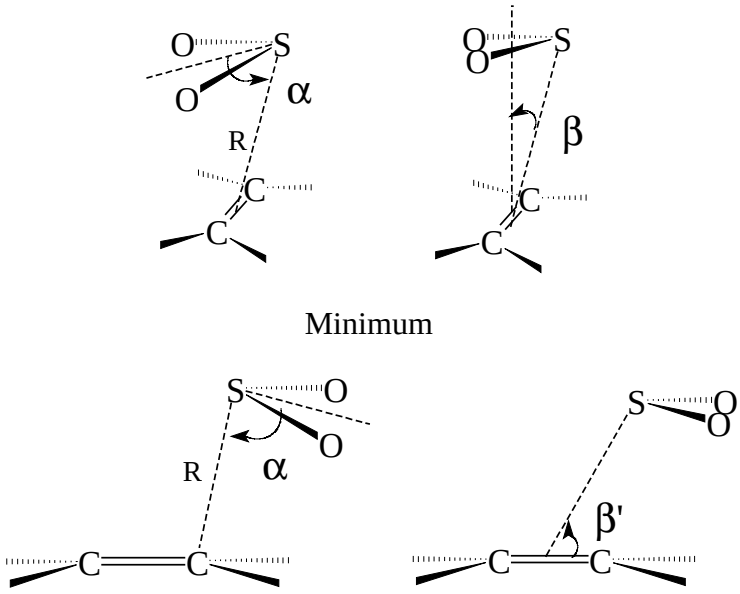
Method	**Rotational Barrier (cm^{-1})**	
MP2/ d95 (d,p)	457	(117)
SP-corrected	408	(33)
CP-MP2/ d95 (d,p)	396	(42)

The same happens for the CP-corrected calculations. These results show clearly the importance of the effect of the relaxation term on the geometry.

II.1.2.1.b $C_2H_4\cdots SO_2$ complex.

Tables XXVIII and XXIX collect the geometrical parameters obtained for the $C_2H_4\cdots SO_2$ system at both the minimum and transition state structures (see SCHEME 16). Only selected intramolecular parameters are shown. Regarding the BSSE effect, a trend similar to that found for the first system is observed. The SP-correction overestimates BSSE in all the cases. CP-corrected intermolecular distance R is *ca.* 0.2 Å longer than the uncorrected value for the minimum geometry. Differences in the order of 1.2 Å are obtained for the TS. The effect on the angular parameters is meaningful, mostly at the TS, where α and β' parameters change from 71.6 and 71.7 to 86.7 and 79.5 degrees , respectively. Neglect of the relaxation term does not change this situation. Relaxation energy terms are negligible. The largest distortions are found in the bond angles aSOO and aHCH, belonging to SO_2 and C_2H_4 molecule

respectively. However, differences of 0.009 and 0.022 Å in the intermolecular distance are obtained. Note that neglect of the relaxation term lengthens slightly the value of R for the minimum, whereas the effect is opposite for the transition state.



SCHEME 16 : Structures involved in the internal rotation of the $C_2H_4\cdots SO_2$. system. The intermolecular distance is indicated

TABLE XXVIII

Geometrical parameters, electronic energies and relaxation contribution for the minimum energy structure of $C_2H_4\cdots SO_2$ complex calculated at MP2, SP-MP2 and CP-MP2 levels of theory. The second half collects the values obtained in the calculations where the intramolecular parameters where frozen to the values they have in the free monomers.

Method	**Energy (a.u.)**	**R (Å)**	**α (deg.)**	**β (deg.)**	**aSOO (deg)**	**aHCH (deg.)**	**E_{rel} (cm^{-1})**
MP2/ d95 (d,p)	-626.044 177 6	3.465	85.0	19.7	108.6	121.3	1
SP-corrected	-626.038 665 5						
CP-MP2/d95 (d,p)	-626.039 011 8	3.677	91.7	14.8	108.7	121.3	0.5
MP2/ d95 (d,p)	-626.043 800 2	3.474	84.0	20.1	108.9*	120.0*	0**
SP-corrected	-626.038 287 3						
CP-MP2/d95 (d,p)	-626.038 633 1	3.687	91.0	15.0	108.9*	120.0*	0**

*Optimized parameters for C_2H_4 and SO_2 isolated systems. **Zero by definition.

Table XXX shows the values obtained for the rotational barrier; in this case, the relaxation term does not have a large effect. Small differences of *ca.* 4-5 cm^{-1} are

observed.

TABLE XXIX

Geometrical parameters, electronic energies and relaxation contribution for transition state structure of $C_2H_4\cdots SO_2$ complex calculated at MP2, SP-MP2 and CP-MP2 levels of theory.

Method	**Energy (a.u.)**	**R (Å)**	**α (deg.)**	**β (deg.)**	**aSOO (deg)**	**aHCH (deg.)**	**E_{rel} (cm^{-1})**
MP2/ d95 (d,p)	-626.043 422 0	3.619	71.6	71.8	108.7	121.3	0.5
SP-corrected	-626.038 455 9						
CP-MP2/d95 (d,p)	-626.038 798 3	3.732	86.7	79.5	108.7	121.3	0.5
MP2/ d95 (d,p)	-626.043 068 4	3.597	73.4	73.4	108.9*	120.0*	0**
SP-corrected	-626.038 118 7						
CP-MP2/d95 (d,p)	-626.038 437 7	3.726	87.1	80.3	108.9*	120.0*	0**

*Optimized parameters for C_2H_4 and SO_2 isolated systems. **Zero by definition.

However, the CP-correction is mandatory: uncorrected values were three times larger than both the SP- and CP-corrected. CP-corrected values are in much better agreement with the estimated experimental value of $30cm^{-1}$ proposed by Andrews et. al[134].

TABLE XXX

Electronic energies (cm^{-1}) for the internal rotation barrier in the $C_2H_4\cdots SO_2$ complex. In parenthesis the values obtained neglecting the relaxation term.

Method	**Rotational Barrier (cm^{-1})**	
MP2/ d95 (d,p)	166	(161)
SP-corrected	46	(37)
CP-MP2/ d95 (d,p)	47	(43)

In short, the counterpoise correction is found to be mandatory for these weakly bonded systems. The effects on the barrier to internal rotation energy and geometrical

[134] A.M. Andrews, K.W. Hillig II, R.L. Kuczkowski, J. Chem. Phys **96**, 1784 (1992)

parameters have been pointed out. It has been shown again that the CP-correction term is strongly geometry dependent; so that CP-corrected optimization must be carried out to obtain accurate BSSE-free geometry for further vibrational analysis.

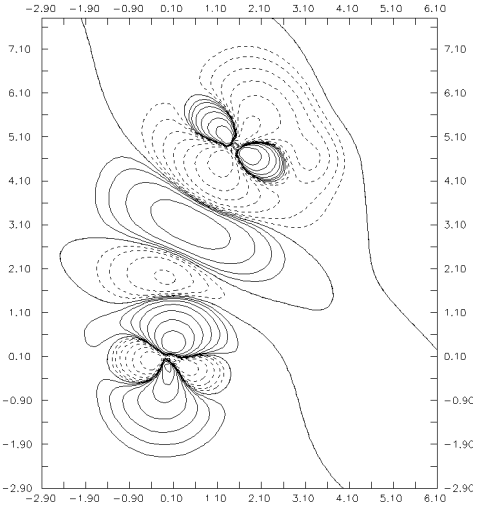
Furthermore, the fragment relaxation energy cannot be seen as an additional term to the rotational barrier. The effect of the fragment relaxation on the intermolecular parameters probed to be very important for the $BF_3 \cdots NH_3$ complex, the CP-correction not changing this situation. The CP-correction scheme can be successfully applied despite the relaxation contribution being not taken into account.

In terms of rotational barrier height, the single-point CP-corrected and CP-corrected values are very similar. However, anharmonic effects can be very important in the hydrogen-bonded and van der Waals complexes. For instance, for several water clusters, Jung et al[135]. have shown that the anharmonic correction modifies by 100 % the vibrational frequencies of intermolecular modes. Moreover, large amplitude modes related to flat potential energy surfaces are poorly described using the rigid rotor-harmonic oscillator approach[136]. This is actually the case of internal rotation motions. Thus, ZPVE corrections using CP-corrected harmonic and anharmonic frequencies for the intermolecular vibrational modes should be considered when looking for high accuracy.

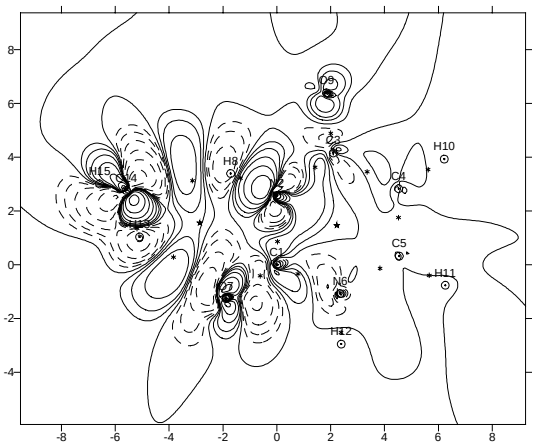
[135] O. Jung, R.B. Gerber, J. Chem. Phys **105**, 23 (10332)

[136] C. Muñoz-Caro, A. Niño, J. Phys.Chem. A **101**, 4128 (1997)

II.1.3 Wavefunction and electron densities



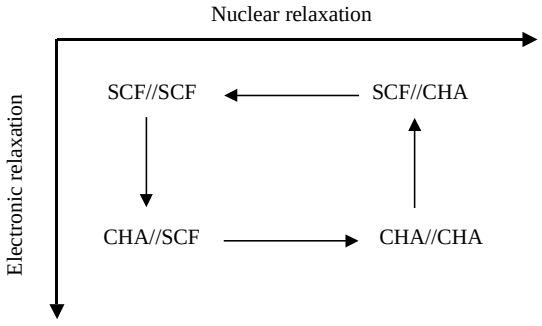
(A)



(B)

Basis Set	Method	Conventional Geometry	Corrected Geometry
6-31G	SCF	0.007101	0.006734
	B3LYP	0.021325	0.011470
6-31G(d,p)	SCF	0.008113	0.007450
	B3LYP	0.020160	0.012080
6-31G++(d,p)	SCF	0.000971	0.000936
	B3LYP	0.001490	0.001385
6-311G(d,p)	SCF	0.004174	0.003740
	B3LYP	0.013927	0.006043
6-311G++(3df,2pd)	SCF	0.001538	0.001459
	B3LYP	0.001844	0.001796

(C)



(D)

A) ρ_{SCF} - ρ_{CHA} *isocontour maps for hydrogen fluoride dimer in the molecular plane*

B) ρ_{SCF} - ρ_{CHA} *isocontour maps for uracil-water complex*

C) QMS distance indices between conventional and CHA densities for the hydrogen fluoride dimer

D) Scheme of the different BSSE-effects in the electron density

In this section we will deal with the BSSE effects on the wavefunction of the molecular complexes, and specifically on the electron density. As mentioned before, unfortunately, the CP recipe does not furnish a corrected wavefunction (and thus a corrected electron density) from which one can obtain the CP-corrected energy as the expected value of the Hamiltonian operator of the system.

These shortcomings can be avoided by the aprioristic methods, since they act directly on the molecular complex description, contrarily to the counterpoise method which is only an error compensation scheme.

Among the *a priori* methods, the Chemical Hamiltonian Approach (CHA) represents the best option, both from the numerical evidence and from a conceptual point of view. The details of this methodology have already been given in the introduction of this work.

Recent studies[57,137] using the SCF-MI[55] method have also dealt with the obtention of BSSE-corrected electron densities for the water dimer complex. Since we have already shown that the charge transfer is neglected in this kind of projection methodologies, we believe that their results are, at least, questionable.

In this work we analyzed for the first time the CHA density at different levels of theory, CHA/F[77] and CHA/DFT[78]. The main goal is to proceed one step ahead on the analysis of the effect of BSSE on molecular complexes, and to assess how the artificial delocalizations caused by BSSE modify the first order electron density of some hydrogen bonded complexes which have been previously studied in this work, namely hydrogen fluoride dimer and water dimer. Short incursions on larger systems in order to assess the possible long range effects on the electron density redistribution upon BSSE correction have also carried out for the formic acid dimer and the uracil-water complex.

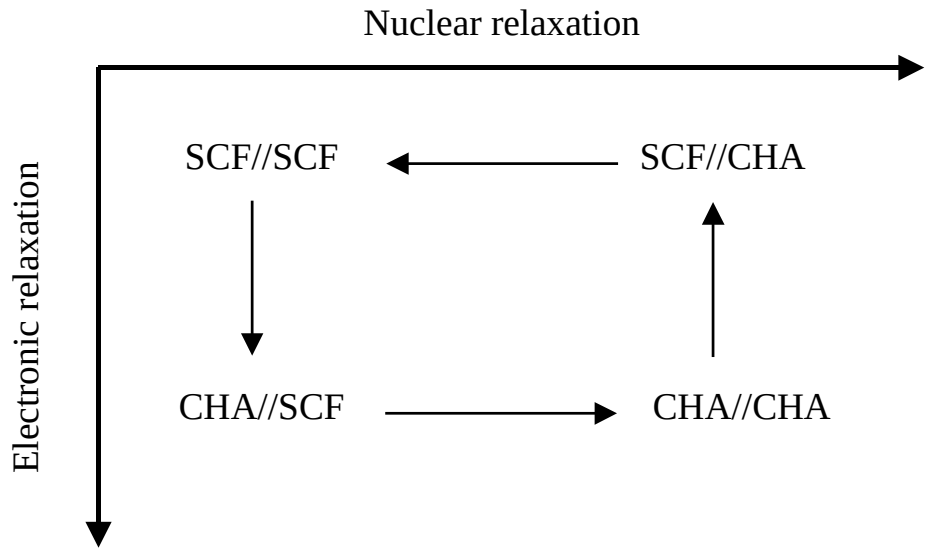
There are several tools for the analysis of the charge densities of molecules. In this case we have used density difference contour maps, topological analysis of the density provided by the Atoms in Molecules (AIM) theory[138], and Quantum

[137] M. Raimondi, A. Famulari and E. Gianinetti, Int. J. Q. Chem **74**, 259 (1999)

[138] R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*(Clarenton, Oxford, 1990)

Molecular Similarity[139] (QMS) measures for the quantitative comparison of conventional and CHA densities. The results obtained by these techniques will be presented in the next subsections.

Now, in order to gain insight into the changes between BSSE-corrected and uncorrected densities, we have considered the BSSE-correction as a *perturbation* on the system, in analogy with the inclusion of electronic correlation. We have considered the conventional, uncorrected calculation as the unperturbed situation. At the given geometry, the corresponding CHA calculation will affect the electronic distribution, as the inclusion of correlation does on a system. A subsequent nuclear relaxation will lead to the final BSSE-corrected description for the given method. The situation is depicted on SCHEME 17, where X//Y represents a single-point calculation using method X at the geometry optimized with method Y. We will use this notation throughout the present section.



SCHEME 17

The analysis of the BSSE effect on the electron density will be carried out in the following way: the SCF//SCF point represents the unperturbed situation, where no BSSE correction has been taken into account. When the corresponding CHA method is applied over the SCF optimized geometry, the electron density is redistributed

[139] R. Carbó, L. Leyda, M. Arnau, Int. J. Quantum Chem. 17, 1185 (1980); E. Besalú, R. Carbó, J. Mestres, M. Solà, Top. Curr. Chem. 173, 31 (1995)

according to the new Hamiltonian introduced by the CHA. We denote this transition as SCF//SCF→ CHA//SCF. The influence on the electron density will be assigned here to electronic relaxation effects. In a second step, the system is allowed to relax to its optimal geometry, induced by the new electrostatic potential. This process is represented by the CHA//SCF→ CHA//CHA transition, and accounts for nuclear relaxation effects. The elimination of the perturbation (CHA//CHA→ SCF//CHA) followed by the final nuclear relaxation (SCF//CHA→ SCF//SCF) would lead to the starting point. An analogous pathway holds in the case of the DFT level of theory.

It is important to note that the CHA method is not a perturbation itself but by considering the BSSE-correction as a perturbation one can study the BSSE-correction in terms of two consecutive effects: the electronic redistribution and the nuclear relaxation.

The effect of the electronic relaxation can be easily visualized by means of contour maps of density differences between the conventional and the BSSE-corrected densities. For nuclear relaxation effects, electron densities computed at different nuclear arrangements can be compared through the AIM theory[138] by means of the characterization of the critical points on the electron density. Note that tackling the analysis of the BSSE-correction is similar to the way that introduction of electronic correlation on a system is usually studied. Indeed, the analysis of the influence of the correlation energy on the bonding or the electron density is generally carried out considering also the electronic and nuclear relaxation effects separately[140].

In short, we will analyze the results obtained in terms of the electronic and nuclear relaxation effects caused by the BSSE at different levels of theory. Numerical geometry optimizations of the considered systems at the CHA/F and CHA/DFT with two different functionals have been already presented and discussed in Section II.1.1.1. Now, for each method we have computed the first-order electron density at the stationary points of the conventional method and its CHA counterpart.

In the next section, the analysis of the critical points of the density will shed light into the electron and nuclear relaxation processes induced by the correction of

[140] M. Solà, J. Mestres, R. Carbó and M. Duran, J. Chem. Phys. **104,** 636 (1996); K. B. Wiberg, C. M. Hadad, T. J. LePage, C. M. Breneman and M. J. Frisch, J. Phys. Chem. **96**, 671 (1992)

the BSSE. The AIMPAC package[141] was used for the characterisation of critical points of the electron density. Next,. the local effects of the BSSE on the density will be visualized by means of electron density difference plots. Finally, the QMS-derived distance indices will give a quantitative measure of the similarities between the uncorrected and BSSE-corrected densities.

Also, a novel methodology recently developed to split the total HF energy of a molecular system into one- and two-center contributions will be used also to assess the local effects of the BSSE in terms of energetic contributions. In this way, the local BSSE effects qualitatively observed by means of density difference plots are complemented with the quantitative counterpart.

[141] F. W. Biegler-König, R. F. W. Bader and T. H. Tang, J. Comp. Chem. **3**, 317 (1982)

II.1.3.1 Topologic analysis of the CHA density

II.1.3.1.a Hydrogen fluoride dimer

The geometrical parameters and electronic stabilization energies of the hydrogen fluoride dimer (see SCHEME 7) for selected basis sets at the SCF, BLYP, and B3LYP levels of theory together with their respective CHA counterparts are collected in Tables XXXI-XXXIII, respectively. For the sake of completeness, we have included the values of the intramolecular parameters. We will comment here just the general trends concerning the effect of BSSE on the geometry in order to understand better the effect of nuclear relaxation on the electron density. For a deeper discussion, see Section II.1.1.1.

TABLE XXXI

Geometrical Parameters (Å and degrees) and electronic stabilization energy (Kcal/mol) for the $(HF)_2$ calculated with different basis sets at the SCF and CHA/F corrected levels of theory. The experimental values of rF-F, α and β are 2.72 Å, 10±6 and 117±10 degrees, respectively. The number of basis functions is showed in parenthesis.

Basis Set	**Method**	**rfh1**	**rF-F**	**rfh2**	**α**	**β**	**Stabilization energy**
6-31G	SCF	0.925	2.706	0.923	8.3	126.0	-7.49
(22)	CHA/F	0.925	2.713	0.923	5.1	135.6	-6.55
6-31G(d)	SCF	0.915	2.709	0.915	17.1	96.6	-6.06
(34)	CHA/F	0.915	2.756	0.914	7.6	114.8	-4.73
6-31G(d,p)	SCF	0.905	2.725	0.904	14.4	101.7	-5.97
(40)	CHA/F	0.904	2.760	0.903	8.3	116.7	-4.62
6-31++G(d,p)	SCF	0.906	2.812	0.905	8.0	120.0	-4.37
(50)	CHA/F	0.906	2.831	0.905	7.4	120.8	-4.14
6-311G(d,p)	SCF	0.900	2.773	0.899	11.6	112.6	-5.20
(50)	CHA/F	0.900	2.822	0.899	7.7	122.8	-4.45
6-311++G(3df,2pd)	SCF	0.902	2.821	0.900	7.0	118.7	-3.93
(122)	CHA/F	0.901	2.840	0.900	7.2	119.9	-3.69

It can be seen that intermolecular distances optimized on the CHA PES (rF-F) are always longer that the corresponding BSSE-uncorrected ones, whereas the

intramolecular parameters (rfh1 and rfh2) remain almost unchanged. The larger intermolecular distance affects to the electronic stabilization energy, which decreases its value after correction for BSSE. Tables XXXI-XXXIII also show the importance of using diffuse functions for the correct description of the systems and the minimization of BSSE. However, the use of the CHA methodology seems to ensure correct descriptions[142] of the systems even with small basis sets. A better agreement with the experimental parameters is obtained in all cases, but for the SCF/6-31G calculation. Finally, as expected, the uncorrected and the CHA molecular parameters (energies, atomic distances) tend to each other as the size of the basis set increases, converging to the same value in the complete basis set (CBS) limit. All these general trends have been observed for all systems studied so far, both including hydrogen bonds or Van der Waals interactions.

TABLE XXXII

Geometrical Parameters (Å and degrees) and electronic stabilization energy (Kcal/mol) for the $(HF)_2$ calculated with different basis sets at the BLYP and CHA/DFT(BLYP) levels of theory.

Basis Set	**Method**	**rfh1**	**RF-F**	**rfh2**	**α**	**β**	**Stabilization energy**
6-31G	BLYP	0.972	2.480	0.972	46.9	47.0	-11.16
(22)	CHA/BLYP	0.972	2.662	0.964	5.6	114.9	-7.27
6-31G(d)	BLYP	0.954	2.485	0.954	45.0	45.0	-10.52
(34)	CHA/BLYP	0.954	2.652	0.949	5.2	105.9	-5.88
6-31G(d,p)	BLYP	0.946	2.494	0.946	44.8	45.0	-10.45
(40)	CHA/BLYP	0.946	2.626	0.940	7.6	105.2	-5.71
6-31++G(d,p)	BLYP	0.947	2.760	0.942	7.5	111.6	-4.75
(50)	CHA/BLYP	0.947	2.782	0.942	7.4	110.6	-4.45
6-311G(d,p)	BLYP	0.938	2.572	0.938	46.8	48.6	-8.28
(50)	CHA/BLYP	0.939	2.730	0.934	8.8	107.9	-5.41
6-311++G(3df,2pd)	BLYP	0.941	2.768	0.937	5.3	112.0	-4.40
(122)	CHA/BLYP	0.941	2.778	0.936	4.8	113.0	-4.16

As mentioned before, an interesting situation arises in the case of the hydrogen fluoride dimer. The results obtained for the angular parameters indicate that

[142] In the sense that molecular properties found with the BSSE-corrected procedures are close to the ones obtained with much larger basis sets.

optimizations using the 6-31G(d), 6-31G(d, p) and even 6-311G(d, p) basis sets lead to linear distorted structures for the SCF calculations, and to cyclic structures when using both the BLYP and the B3LYP functionals. Note that the corresponding CHA calculations change this situation and lead to the correct structures and stabilization energies as compared to the experiment. These kind of dramatic changes in the topology of the PES upon the BSSE-correction have also been discussed for other weakly bound systems like the water dimer or some C-H···O hydrogen bonded complexes. Both CHA and CP results point out that this effect is directly caused by the BSSE rather than by a supposed poor behavior of the actual functionals used to describe these kind of systems.

TABLE XXXIII

Geometrical Parameters (Å and degrees) and electronic stabilization energy (Kcal/mol) for the $(HF)_2$ calculated with different basis sets at the B3LYP and CHA/DFT(B3LYP) levels of theory.

Basis Set	**Method**	**rfh1**	**RF-F**	**rfh2**	**α**	**β**	**Stabilization energy**
6-31G	B3LYP	0.956	2.478	0.956	48.7	49.6	-10.20
(22)	CHA/B3LYP	0.957	2.647	0.951	5.2	120.0	-7.56
6-31G(d)	B3LYP	0.941	2.485	0.941	46.4	46.6	-9.60
(34)	CHA/B3LYP	0.942	2.647	0.938	5.7	107.5	-6.15
6-31G(d,p)	B3LYP	0.933	2.493	0.933	46.4	46.5	-9.42
(40)	CHA/B3LYP	0.933	2.627	0.929	7.8	107.5	-5.87
6-31++G(d,p)	B3LYP	0.935	2.732	0.931	7.6	113.1	-5.08
(50)	CHA/B3LYP	0.935	2.752	0.931	7.4	112.5	-4.79
6-311G(d,p)	B3LYP	0.926	2.567	0.926	47.9	50.5	-7.53
(50)	CHA/B3LYP	0.927	2.725	0.924	7.0	113.9	-5.52
6-311++G(3df,2pd)	B3LYP	0.930	2.734	0.926	5.8	112.5	-4.71
(122)	CHA/B3LYP	0.930	2.747	0.926	5.4	113.9	-4.47

By studying the PES one can extract information mainly in terms of nuclear relaxation, i.e., how the nuclear positions move according to the electronic potential. To gain a deeper insight in how the BSSE affects the electronic distribution at a given nuclear arrangement (electronic relaxation) one is driven to study the electron density and its topology. In order to compare both the uncorrected and the CHA first order electron densities we will use the well-know Atoms in Molecules (AIM) theory,

characterizing the topology of the electronic distributions in terms of the nature of its critical points and values of the electron density and its Laplacian. Note that since no BSSE-corrected densities can be obtained by means of the CP method, this analysis can be carried out only by using a priori methods for the elimination of the BSSE.

TABLE XXXIV

Electron density and its Laplacian at the intermolecular bond critical points located on the SCF and CHA/F electron densities calculated with several basis sets. The distance of the bcp to the fluorine atom of the H-acceptor molecule is also reported.

Basis Set	**Method // Geometry**	**$\rho(r)$ (e/a.u.^{3})**	**$\nabla^2\rho(r)$ (e/a.u.^{5})**	**Distance to F (Å)**
6-31G	SCF//SCF	0.0285	0.1242	1.153
	CHA//SCF	0.0278	0.1302	1.150
	CHA//CHA	0.0272	0.1304	1.147
6-31G(d)	SCF//SCF	0.0252	0.1041	1.195
	CHA//SCF	0.0241	0.1114	1.192
	CHA//CHA	0.0231	0.1065	1.193
6-31G(d,p)	SCF//SCF	0.0222	0.0946	1.228
	CHA//SCF	0.0213	0.0961	1.223
	CHA//CHA	0.0199	0.0914	1.231
6-31G++(d,p)	SCF//SCF	0.0175	0.0763	1.262
	CHA//SCF	0.0177	0.0750	1.263
	CHA//CHA	0.0169	0.0713	1.273
6-311G(d,p)	SCF//SCF	0.0227	0.1189	1.201
	CHA//SCF	0.0235	0.1129	1.201
	CHA//CHA	0.0167	0.0875	1.258
6-311G++(3df,2pd)	SCF//SCF	0.0173	0.0779	1.246
	CHA//SCF	0.0180	0.0725	1.246
	CHA//CHA	0.0170	0.0697	1.258

The results for the topological analysis of SCF electron densities are collected in Table XXXIV. In all cases a bond critical point between the hydrogen donor (H-donor) and the hydrogen acceptor (H-acceptor) molecules is located on the electron density. The values of the electron density and its Laplacian at this point are typical of a closed-shell weak interaction. The large distance of the bcp to the F atom of the H-

acceptor molecule also reveals this situation. Deviations caused by the BSSE are small in magnitude and, in general, represent about 5-10% of the value of $\rho(r)$ and $\nabla^2\rho(r)$ in the bcp. Note that the improvement of the quality of the basis set seems not to be the main factor as, for instance, the differences represent up to 25% for the large triple-zeta 6-311G(d, p) calculation. As expected, both the 6-31G++(d, p) and the 6-311G++(3df, 2p), the largest basis set used for this study, present the smallest differences, although these are still appreciable for the Laplacian, which is known to be a parameter very sensitive to small perturbations in the electron density.

Concerning the electronic relaxation, the following trends have been observed. For the small basis sets, the SCF//SCF $\rightarrow$ CHA//SCF transition causes a slight displacement of the bcp towards the F atom of the H-acceptor molecule, a subsequent reduction of the value of the electron density and an increase of charge depletion at the bcp. However, when using larger basis sets including diffuse functions, or even with the 6-311G(d, p) basis set, the opposite effect is found.

On the other hand, nuclear relaxation effects can be understood by the subsequent CHA//SCF $\rightarrow$ CHA//CHA transition. The general trend, observed for all cases but the 6-31G basis set, is the following: after BSSE correction, charge density is more concentrated in the lone pairs of the F atoms. Therefore the electron repulsion increases, and when nuclear relaxation is allowed, the intermolecular rF-F parameter increases. The distance of the bond critical point to the F atom of the H-acceptor also augments and finally, both the electron density and its Laplacian at this bcp decrease to their minimum value.

When considering the whole process, SCF//SCF $\rightarrow$ CHA//CHA, it is worth to note that the electron density at the bcp decreases in all cases, in good agreement with the weakening of the intermolecular interaction after BSSE correction, whereas the behavior of its Laplacian depends upon the precise location of this bcp. Hence, the Laplacian decreases slightly when the bcp sets closer to the F atom when correcting for BSSE. However, the opposite trend is found when the bcp is located farther of the F atom after the BSSE correction.

TABLE XXXV

Electron density and its Laplacian at the intermolecular critical point located on the BLYP and CHA/BLYP first order electron densities calculated with several basis sets. The distance of the corresponding critical point to the fluorine atom of the H-acceptor molecule is also reported.

Basis Set	Method // Geometry	Critical Point	$\rho(r)$ (e/a.u.^{3})	$\nabla^2\rho(r)$ (e/a.u.^{5})	Distance to F (Å)
6-31G	BLYP//BLYP[a]	(3, -1)[b]	0.0260	0.1111	
		(3, +1)	0.0250	0.1262	1.245
	CHA//BLYP[a]	(3, -1)[b]	0.0245	0.1218	
		(3, +1)	0.0240	0.1289	1.249
	CHA//CHA	(3, -1)	0.0398	0.1555	1.095
6-31G(d)	BLYP//BLYP[a]	(3, -1)[b]	0.0270	0.1283	
		(3, +1)	0.0268	0.1403	1.246
	CHA//BLYP[a]	(3, -1)[b]	0.0257	0.1399	
		(3, +1)	0.0256	0.1449	1.252
	CHA//CHA	(3, -1)	0.0384	0.1406	1.108
6-31G(d,p)	BLYP//BLYP[a]	(3, -1)[b]	0.0271	0.1345	
		(3, +1)	0.0271	0.1402	1.254
	CHA//BLYP[a]	(3, -1)	0.0259	0.1443	1.238
	CHA//CHA	(3, -1)	0.0380	0.1277	1.120
6-31G++(d,p)	BLYP//BLYP	(3, -1)	0.0264	0.0845	1.192
	CHA//BLYP	(3, -1)	0.0267	0.0807	1.191
	CHA//CHA	(3, -1)	0.0254	0.0760	1.203
6-311G(d,p)	BLYP//BLYP[a]	(3, -1)	0.0221	0.0968	1.260
	CHA//BLYP[a]	(3, -1)	0.0214	0.0996	1.264
	CHA//CHA	(3, -1)	0.0286	0.1077	1.179
6-311G++(3df,2pd)	BLYP//BLYP	(3, -1)	0.0263	0.0848	1.185
	CHA//BLYP	(3, -1)	0.0271	0.0789	1.186
	CHA//CHA	(3, -1)	0.0265	0.0775	1.191

[a] Cyclic structures.

[b] Only the features of one of the bcp's are reported, as both are almost equivalent by symmetry.

Dynamic correlation has been included through Density Functional Theory with the BLYP and B3LYP exchange-correlation functionals. As commented in Section

II.1.1.1.a, the effect of BSSE on the molecular geometry is very important. All BSSE-uncorrected optimizations involving basis sets without diffuse functions lead to cyclic structures (see SCHEME 7), which should normally correspond to a transition state connecting two equivalent linear conformations in an internal rotation process[143]. Moreover, all CHA optimized structures yield the correct, nearly linear arrangement, in good agreement with experimental data. Therefore, for this system, CHA deals well with BSSE-like delocalizations occurring even when rather incomplete basis sets are used, modifying the topology of the PES.

The results of a Bader analysis of the charge density are collected in Tables XXXV and XXXVI. Again the values of the electron density and its Laplacian at the bond critical points reveal a closed-shell interaction between the monomers. Interestingly, for uncorrected BLYP, with all the basis sets without diffuse functions (except the 6-311G(d, p)), one finds a ring critical point (rcp) and two bond critical points, almost equivalent due to the symmetry, between both H-F molecules. That is, both monomers act as H-donor and H-acceptor at the same time and the bonding is described by two intermolecular interactions between both hydrogen fluoride monomers. For the rest of calculations only a bond critical point is located on the charge density in the intermolecular region. Using the uncorrected B3LYP functional changes this situation dramatically. Then, only the smallest basis set, 6-31G, yields a ring critical point, even though cyclic structures are found for the rest of basis sets without diffuse functions. In terms of the geometrical parameters obtained for both functionals, one could argue that both yielded bad descriptions of the bonding when BSSE is not taken into account. However, Bader analysis shows that a cyclic structure is not always associated with the existence of two hydrogen bonds in the BSSE-uncorrected system.

[143] For instance, at the SCF/6-31G(d) level of theory, a cyclic structure is located on the PES with an imaginary frequency of ca. 150i cm^{-1}. Similarly, with the CHA/DFT method using the same basis set, another cyclic structure is located on the PES and characterized as a transition state with a similar imaginary frequency.

TABLE XXXVI

Electron density and its Laplacian at the intermolecular critical point located on the B3LYP and CHA/B3LYP first order electron densities calculated with several basis sets. The distance of the corresponding critical point to the fluorine atom of the H-acceptor molecule is also presented.

Basis Set	Method // Geometry	Critical Point	$\rho(r)$ (e/a.u.^{3})	$\nabla^2\rho(r)$ (e/a.u.^{5})	Distance to F (Å)
6-31G	B3LYP//B3LYP[a]	(3, -1)[b]	0.0240	0.1106	1.259
		(3, +1)	0.0233	0.1207	
	CHA//B3LYP[a]	(3, -1)	0.0229	0.1186	1.199
	CHA//CHA	(3, -1)	0.0387	0.1588	1.096
6-31G(d)	B3LYP//B3LYP[a]	(3, -1)	0.0256	0.1287	1.207
	CHA//B3LYP[a]	(3, -1)	0.0244	0.1380	1.210
	CHA//CHA	(3, -1)	0.0366	0.1409	1.114
6-31G(d,p)	B3LYP//B3LYP[a]	(3, -1)	0.0254	0.1312	1.212
	CHA//B3LYP[a]	(3, -1)	0.0242	0.1399	1.222
	CHA//CHA	(3, -1)	0.0353	0.1278	1.131
6-31G++(d,p)	B3LYP//B3LYP	(3, -1)	0.0263	0.0929	1.189
	CHA//B3LYP	(3, -1)	0.0267	0.0885	1.188
	CHA//CHA	(3, -1)	0.0254	0.0837	1.199
6-311G(d,p)	B3LYP//B3LYP[a]	(3, -1)	0.0210	0.0950	1.246
	CHA//B3LYP[a]	(3, -1)	0.0204	0.0978	1.248
	CHA//CHA	(3, -1)	0.0272	0.1103	1.183
6-311G++(3df,2pd)	B3LYP//B3LYP	(3, -1)	0.0268	0.0905	1.179
	CHA//B3LYP	(3, -1)	0.0277	0.0849	1.180
	CHA//CHA	(3, -1)	0.0267	0.0827	1.186

[a] Cyclic structures

[b] Only the features of one of the bcp's are reported, as both are almost equivalent by symmetry

In fact, up to four different connectivities are observed for these cyclic structures. When a single intermolecular bcp is located on the charge density, the intermolecular bond path links both F atoms. This molecular graph is also observed for the cyclic transition state structures obtained at the SCF level of theory. When the electron density exhibits a rcp there are several possible connectivities. For instance, with the 6-31G basis set using the BLYP functional, both the uncorrected and the CHA corrected electron density show a molecular graph where the two intermolecular bcp connect each F atom with the H atom of the partner. When the basis set size is

increased with polarization functions two bond paths connecting the F atoms are observed. Finally, the B3LYP/6-31G electron density exhibits a rather unphysical molecular graph with one bond path connecting the two F atoms and another one linking one of the F atoms with the H atom of the other monomer. Indeed, the electron density surface is so flat in the intermolecular region that small perturbations can induce dramatic changes on its topology and connectivity. The main conclusion is that the cyclic structure is very unstable and that the electronic relaxation caused by the BSSE correction is unable to overcome this circumstance.

To investigate the effect of the CHA correction on the charge density we refer again to the electronic relaxation process described by the transition DFT//DFT $\rightarrow$ CHA//DFT. In general, rather small differences in the values of the charge density and its Laplacian are found again and trends similar to those found above for the SCF calculations are observed for some basis sets. However, for basis sets yielding ring critical points, the effect of the electronic relaxation is very important. Regarding BLYP results, it can be seen that in all cases the ring critical point is slightly moved away from the F atom of the original H-acceptor molecule, which induces a decrease of the electron density and an increase of its Laplacian. However, whereas the rank of the critical point is not modified for the BLYP/6-31G and BLYP/6-31G(d, p) basis sets after CHA correction, it is modified when adding polarization functions to H atoms. In such a case, the critical point is displaced along the opposite direction. Therefore, the two new bond critical points observed between both BSSE-uncorrected F···H intermolecular interactions before the CHA correction is applied disappear, so the hydrogen bonding becomes described by a unique bcp corresponding to a single hydrogen bond.

Concerning B3LYP calculations, the ring critical point is found only for the smallest basis set. Again, after electronic relaxation the topology of the charge density changes and a single bond critical point is found. In this case, both the electron density and its Laplacian decrease; the most important effect is, however, the displacement of the critical point away from the F atom of the H-acceptor molecule by more than 0.05 Å, The final distance of 1.199 Å is much closer to those obtained with larger basis sets.

Finally, nuclear relaxation process is very important. In all cases the final CHA//CHA structures have a linear arrangement and the bonding is represented by a

single bond critical point in the intermolecular region. When nuclear relaxation is allowed, the intermolecular rF-F distance increases by ca 0.2 Å, whereas the intermolecular F···H distance decreases in the same way due to the changes in the angular disposition of both monomers. Of course, the main differences in terms of the location of the bond critical point and its density and Laplacian are observed for the basis sets that yielded cyclic structures on the uncorrected PES. The final location of the bcp is much closer to the F atom of the H-acceptor molecule whereas, contrary to the linear arrangement cases, the value of the charge density increases by up to 35% in some cases. There is a simple explanation for this effect. In the cyclic structure, the intermolecular F···H distances are much longer than in the linear case. The uncorrected stabilization energies are of the order of 8-10Kcal/mol, but this value can be seen as the sum of the contributions of two intermolecular bonds, even though the Bader analysis does not always detect their presence, e.g., at the BLYP/6-31G(d), BLYP/6-31G(d, p) and BLYP/6-311G(d, p) levels of theory. Under this assumption, the stabilization energy per intermolecular bond is always larger for the CHA//CHA calculation than for the DFT//DFT in the cyclic cases; therefore its seems reasonable to expect a larger value of the density at the corresponding bond critical point. Changes in the Laplacian are less important in general and much more erratic; is not easy to define a tendency from the tabulated values.

The differences observed after the BSSE correction for the calculations including diffuse functions are almost negligible, showing again the asymptotic approach between the uncorrected and the CHA results when the quality of the basis set is improved.

II.1.3.1.b Water dimer

All calculations refer to the trans-linear water dimer, having C_s symmetry and a single H bond between the two water molecules. Although the general structure and molecular connectivity are preserved along all calculations, the level of theory, basis set and BSSE correction have a significant impact on the molecular geometry (see Section II.1.1.1.b). Table XXXVII and XXXVIII reports the selected parameters extracted from the AIM analysis of the critical point of the electron density for the water dimer, at SCF and B3LYP levels of theory, with and without BSSE correction.

The BLYP results are very similar to the B3LYP ones so, for the sake of simplicity, they have been omitted.

TABLE XXXVII

Electron density, Laplacian and elypticity at the intermolecular bond critical points located on the SCF and CHA/F electron densities calculated with several basis sets. The distance of the bcp to the oxygen atom of the H-acceptor molecule is also reported.

Basis Set	**Method // Geometry**	**ρ(r) (e/a.u.^{3})**	**$\nabla^2\rho$(r) (e/a.u.^{5})**	**Elypticity**	**Distance to O_1 (Å)**
6-31G	SCF//SCF	0.0297	0.1119	0.069	1.220
	CHA//SCF	0.0300	0.1125	0.070	1.217
	CHA//CHA	0.0283	0.1068	0.073	1.229
6-31G(d)	SCF//SCF	0.0221	0.0743	0.035	1.299
	CHA//SCF	0.0216	0.0770	0.035	1.296
	CHA//CHA	0.0209	0.0749	0.042	1.299
6-31G(d,p)	SCF//SCF	0.0199	0.0621	0.028	1.334
	CHA//SCF	0.0195	0.0619	0.028	1.329
	CHA//CHA	0.0186	0.0597	0.039	1.333
6-31G++(d,p)	SCF//SCF	0.0183	0.0610	0.046	1.327
	CHA//SCF	0.0191	0.0579	0.043	1.329
	CHA//CHA	0.0173	0.0525	0.048	1.350
6-311G(d,p)	SCF//SCF	0.0186	0.0824	0.032	1.325
	CHA//SCF	0.0193	0.0773	0.031	1.325
	CHA//CHA	0.0163	0.0675	0.043	1.356
6-311G++(3df,2pd)	SCF//SCF	0.0157	0.0658	0.043	1.345
	CHA//SCF	0.0159	0.0636	0.042	1.344
	CHA//CHA	0.0154	0.0621	0.045	1.349

It has been determined that in general, the rO-O distance is longer at the SCF than at the B3LYP level of theory. In both cases, the effect of increasing the basis set size is to lengthen the rO-O bond. Furthermore, correcting the geometry for BSSE does also yield systematically larger rO-O distances. At both levels of theory, results for the 6-311++G(3df,2pd) basis set show that, when the complete basis set limit is

approached, the corrected and uncorrected geometrical parameters tend to converge to the same values.

TABLE XXXVIII

Electron density, Laplacian and elypticity at the intermolecular critical point located on the B3LYP and CHA/B3LYP first order electron densities calculated with several basis sets. The distance of the corresponding critical point to the oxygen atom of the H-acceptor molecule is also presented.

Basis Set	**Method // Geometry**	**ρ(r) (e/a.u.^{3})**	**$\nabla^2\rho$(r) (e/a.u.^{5})**	**Elypticity**	**Distance to O_1 (Å)**
6-31G	B3LYP//B3LYP	0.0399	0.1305	0.054	1.166
	CHA//B3LYP	0.0400	0.1318	0.053	1.164
	CHA//CHA	0.0377	0.1268	0.062	1.173
6-31G(d)	B3LYP//B3LYP	0.0311	0.0879	0.027	1.233
	CHA//B3LYP	0.0302	0.0930	0.025	1.231
	CHA//CHA	0.0304	0.0924	0.032	1.226
6-31G(d,p)	B3LYP//B3LYP	0.0288	0.0756	0.021	1.256
	CHA//B3LYP	0.0281	0.0753	0.019	1.252
	CHA//CHA	0.0284	0.0753	0.030	1.247
6-31G++(d,p)	B3LYP//B3LYP	0.0264	0.0769	0.042	1.245
	CHA//B3LYP	0.0276	0.0709	0.040	1.245
	CHA//CHA	0.0255	0.0646	0.040	1.264
6-311G(d,p)	B3LYP//B3LYP	0.0265	0.0953	0.019	1.260
	CHA//B3LYP	0.0277	0.0871	0.018	1.263
	CHA//CHA	0.0245	0.0806	0.032	1.285
6-311G++(3df,2pd)	B3LYP//B3LYP	0.0244	0.0796	0.032	1.263
	CHA//B3LYP	0.0247	0.0774	0.033	1.261
	CHA//CHA	0.0241	0.0763	0.035	1.265

For all calculations, the water dimer maintains the same molecular connectivity with and without BSSE correction. Thus, a bcp corresponding to an intermolecular H bond is always observed between the two water monomers, and the properties associated to this bcp can be used to assess the characteristics of the H_2O-H_2O interaction for different calculations For the B3LYP/6-311++G(3df,2dp) calculations, the value of the density at the bcp, $\rho_{bcp}(\mathbf{r})$, is ca. 0.024, and the rO-O distance is ca.

2.93 Å; these values are characteristic of a strong O-H·····O bond.[144] Moreover, the Laplacian of the density at the bcp, $\nabla^2\rho_{bcp}(\mathbf{r})$, is ca. 0.08, also in agreement with the values expected for this kind of interactions.

According to the positive value of $\nabla^2\rho_{bcp}(\mathbf{r})$, there is no accumulation of charge density at the intermolecular region, so the water-water interaction might be classified as closed-shell; however, recent experimental and theoretical data suggest that, in general, H bonds should be considered at least partially covalent.[145,146]

The $\rho_{bcp}(\mathbf{r})$ and $\nabla^2\rho_{bcp}(\mathbf{r})$ values are nearly unchanged after correcting for the BSSE without reoptimizing the molecular geometry. Only $\nabla^2\rho_{bcp}(\mathbf{r})$ displays significant changes. However, there is no definite trend for the effect of BSSE correction. In general, the BSSE correction leads to larger $\nabla^2\rho_{bcp}(\mathbf{r})$ values for the smaller basis sets (6-31G and 6-31G(d,p)), and to smaller values for the rest of basis sets.

The effects of BSSE on the molecular density are more important when the nuclear relaxation is also taken into account, i.e., for the CHA//SCF→CHA//CHA transition. This is reflected on the corresponding values of $\rho_{bcp}(\mathbf{r})$ and $\nabla^2\rho_{bcp}(\mathbf{r})$ for the intermolecular bcp. For all the approximations and basis sets, $\rho_{bcp}(\mathbf{r})$ always decreases significantly after BSSE correction, which agrees with a weakening of the intermolecular interaction. As expected, the corrected and uncorrected values tend to converge to the same limit for large basis sets. The BSSE correction also lowers the $\nabla^2\rho_{bcp}(\mathbf{r})$ value for most calculations.

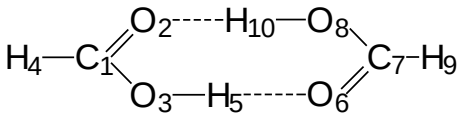
[144] E. Espinosa, M. Souhassou, H. Lachekar, C. Lecomte, Acta Cryst. B **55**, 563 (1999).

[145] E. D. Isaacs, A. Shukla, P. M. Platzman, R. D. Hamann, B. Barbiellini, C. A. Tulk, Phys. Rev. Lett. **82**, 600 (1999)

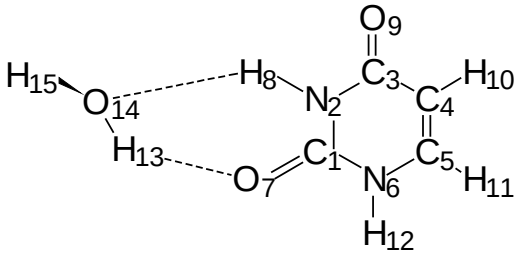
[146] J. M. Ugalde, I. Alkorta, J. Elguera, Angew. Chem. Int. Ed. **39**, 7171 (2000).

II.1.3.1.c Formic acid dimer and uracil-water complex

In order to perform a BSSE-free calculation, first of all one must define which are the fragments to be considered. For the molecular dimers discussed so far, the choice is straightforward, and the fragments are coincident with the monomers within each complex. Thus, it is expected that the non-physical terms removed by the CHA method will correspond to interactions between atoms involved in the intermolecular bonding. However, analysis of density difference maps for the water and hydrogen fluoride dimers has shown that the effect of BSSE is not strictly localized in the intermolecular region (see next subsection). Furthermore, in many cases, the density redistribution associated to the elimination of the BSSE are located mainly in the valence shells of the heavy atoms. It can be interesting to analyze the scope of the BSSE-related density redistribution for larger complexes. For this purpose, we will analyze briefly the effect of BSSE on the formic acid dimer and the uracil-water complex. The latter example is of considerable interest because the hydration of nucleic acids and biomolecules in general is important in many biological processes.



(a)



(b)

SCHEME 18: Formic acid dimer (a) and Uracil-water (b) complex

TABLE XXXIX

Electron density and its Laplacian at the intermolecular critical points located on the first-order electron density of the formic acid. SCF and CHA/F values (in parenthesis) are reported. All the calculations have been performed at the geometry optimized on the SCF PES

Basis Set	**Critical Point**	**$\rho(r)$ (e/a.u.^{3})**	**$\nabla^2\rho(r)$ (e/a.u.^{5})**
6-31G(d, p)	(3, +1)	0.0059 (0.0056)	0.0282 (0.0291)
	(3, -1)	0.0286 (0.0290)	0.1034 (0.0976)
6-31G+(d,p)	(3, +1)	0.0060 (0.0060)	0.0273 (0.0274)
	(3, -1)	0.0269 (0.0278)	0.0972 (0.0902)

TABLE XL

Electron density and its Laplacian at the intermolecular critical points located on the first-order electron density of the uracil-water complex (see SCHEME 18b). SCF and CHA/F values (in parenthesis) are reported. All the calculations have been performed at the geometry optimized on the SCF PES

Basis Set	**Critical Point**	**$\rho(r)$ (e/a.u.^{3})**	**$\nabla^2\rho(r)$ (e/a.u.^{5})**
6-31G(d)	(3, +1)	0.0086 (0.0084)	0.0460 (0.0465)
	O_{14}-H_8 (3, -1)	0.0203 (0.0197)	0.0699 (0.0726)
	O_{13}-H_7 (3, -1)	0.0193 (0.0188)	0.0700 (0.0730)
6-31G+(d)	(3, +1)	0.0079 (0.0081)	0.0427 (0.0424)
	O_{14}-H_8 (3, -1)	0.0173 (0.0179)	0.0672 (0.0658)
	O_{13}-H_7 (3, -1)	0.0182 (0.0183)	0.0691 (0.0692)

SCF and CHA/F calculations for the formic acid dimer and uracil-water complex have been performed with the 6-31G(d,p) and the 6-31G(d) basis sets respectively, as well as adding diffuse functions on the heavy atoms, at the molecular geometries optimized on the conventional PES. All optimized structures exhibit a C_{2h} symmetry, for the formic acid dimer, and C_s, for uracil-water. For the latter complex, the optimized structure reported corresponds in fact to one of several local minima which

are close in energy.[147] In all cases, two intermolecular H bonds are found, and the corresponding bcp's can be characterized on the electron density, together with a rcp (see Table XXXIX and XL, for formic acid dimer and uracil-water, respectively).

As expected, since nuclear relaxation has not been allowed, the effect of the BSSE correction on $\rho_{bcp}(\mathbf{r})$ and $\nabla^2\rho_{bcp}(\mathbf{r})$ is always very small. The patterns observed in the density difference maps (see next section) are consistent with those found for the water and hydrogen fluoride dimers.

[147] T. van Mourik, S. L. Price, D. C. Clary, J. Phys. Chem. A **103**, 1611 (1999)

II.1.3.2 Electron density difference maps

One can get further insight on the local electronic redistribution caused after correcting for BSSE at a fixed geometry by depicting density difference maps between conventional and BSSE-corrected densities. These maps have been created using a program that properly handles the electron density obtained from the Gaussian package and creates a grid of points in a given plane.

For the hydrogen fluoride and formic acid dimers, the maps correspond to the molecular plane. In case of the water dimer, the selected plane is that defined by the H-bond water moiety and the oxygen atom of the partner. For the uracil-water complex, the selected plane only excludes that hydrogen atom of the water molecule that does not interact directly with the uracil moiety.

In all cases, the quantity plotted is the difference between the uncorrected density and the corresponding CHA density, *computed at the same, fixed, molecular geometry*, namely

$$\Delta\rho_{SCF} = \rho_{SCF} - \rho_{CHA/F}$$

$$\Delta\rho_{DFT} = \rho_{DFT} - \rho_{CHA/DFT}$$

Solid lines represent the zones where the electron density is overestimated by BSSE, whereas dashed lines hold for the zones where the CHA density is larger than the uncorrected one.

In order to follow the perturbative scheme aforementioned, the density difference grids have been generated at the *uncorrected* geometry of the given level of theory. Hence, the electronic relaxation contribution is the one that will be visualized.

II.1.3.2.a Hydrogen fluoride dimer

Figure 16 depicts these density difference maps for the hydrogen fluoride dimer for several basis sets. The maps obtained with basis sets without diffuse functions are qualitatively very similar (see Figures 16.a, b, and c.). The main differences between the conventional and the CHA densities are found in the atomic basins of the F atoms, mostly for the H-acceptor moiety.

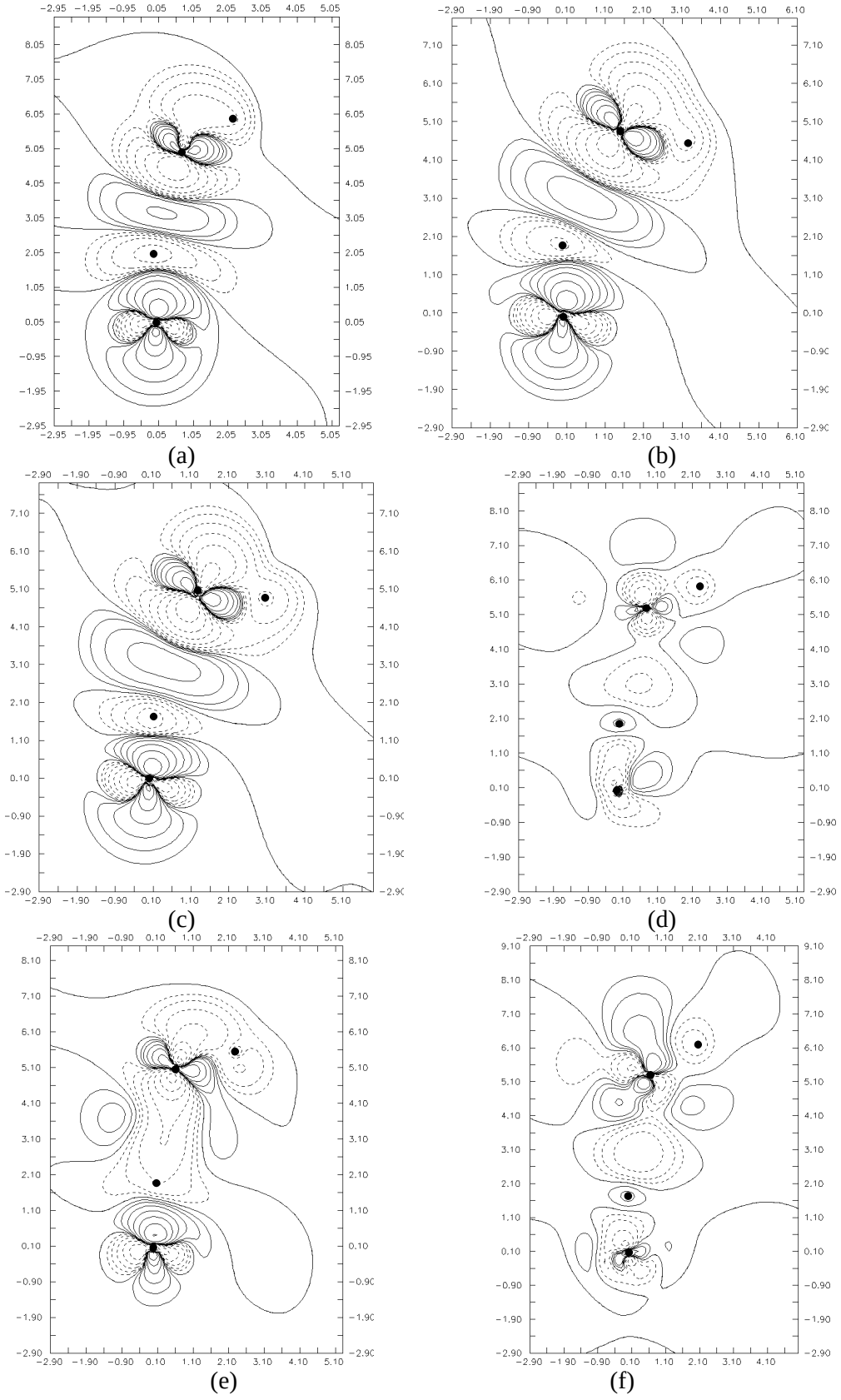


FIGURE 16: SCF//SCF- CHA//SCF Density difference isocontour maps. (a) 6-31G, (b) 6-31G(d), (c) 6-31G(d,p), (d) 6-31++G(d,p), (e) 6-311G(d,p) and (f) 6-311++G(3df,2pd). The nuclear positions are indicated by solid dots. Isodensity contours at $1.e^{-4}$, $2.e^{-4}$, $4.e^{-4}$, $8.e^{-4}$,etc...

It can be observed that the elimination of the BSSE transfers electron density from the intramolecular F-H bonds to the F lone electron pairs. The electron density on the H atom regions is also increased, allowing for a minor charge transfer, and hence decreasing the intermolecular bond strength. Finally, the electron density in the intermolecular bond region is slightly overestimated by BSSE, except for the 6-311G(d, p) basis set, for which this positive region in the map is not observed, (see Figure 16.e). This overestimation is connected with the increasing value of $\rho(r)$ at the bcp after electronic relaxation, as commented above.

Adding diffuse functions changes the situation dramatically, because differences between the corrected and uncorrected densities become less important. Isocontour lines are more separated in the vicinity of the F atoms, (see Figures 16d and f). An interesting point is that the BSSE effect tends to underestimate the electron density in the intermolecular bond region. This situation, opposite to that observed in the other cases, can be explained in the following way: the atomic orbitals of the hydrogen involved in the intermolecular bond can be used by the F atom of the H-acceptor molecule to become more stable, causing the BSSE-like delocalizations. The CHA correction tends to increase the electron density on the F atom, thus avoiding a larger charge transfer and hence a larger interaction. Therefore, after correction for the BSSE, a decrease of $\rho(r)$ is observed along the intermolecular bond zone, together with an increase of the density in the valence region of both atoms. The situation should be the same when diffuse functions are available for the H atom. However, in this case, as the intermolecular bcp falls into the van der Waals radius of the H atom, the main part of the $\rho(r)$ at the bcp is provided by the orbitals with a larger contribution of the diffuse functions. The final effect on the electron density is the opposite; the charge density decreases at the H atom but increases slightly at the intermolecular bond region after correcting for BSSE. One must say that the differences are of the order of 0.0002 a.u. and have to be seen as a consequence of the changes on the atomic basins of the F atoms, which have proved to be more important.

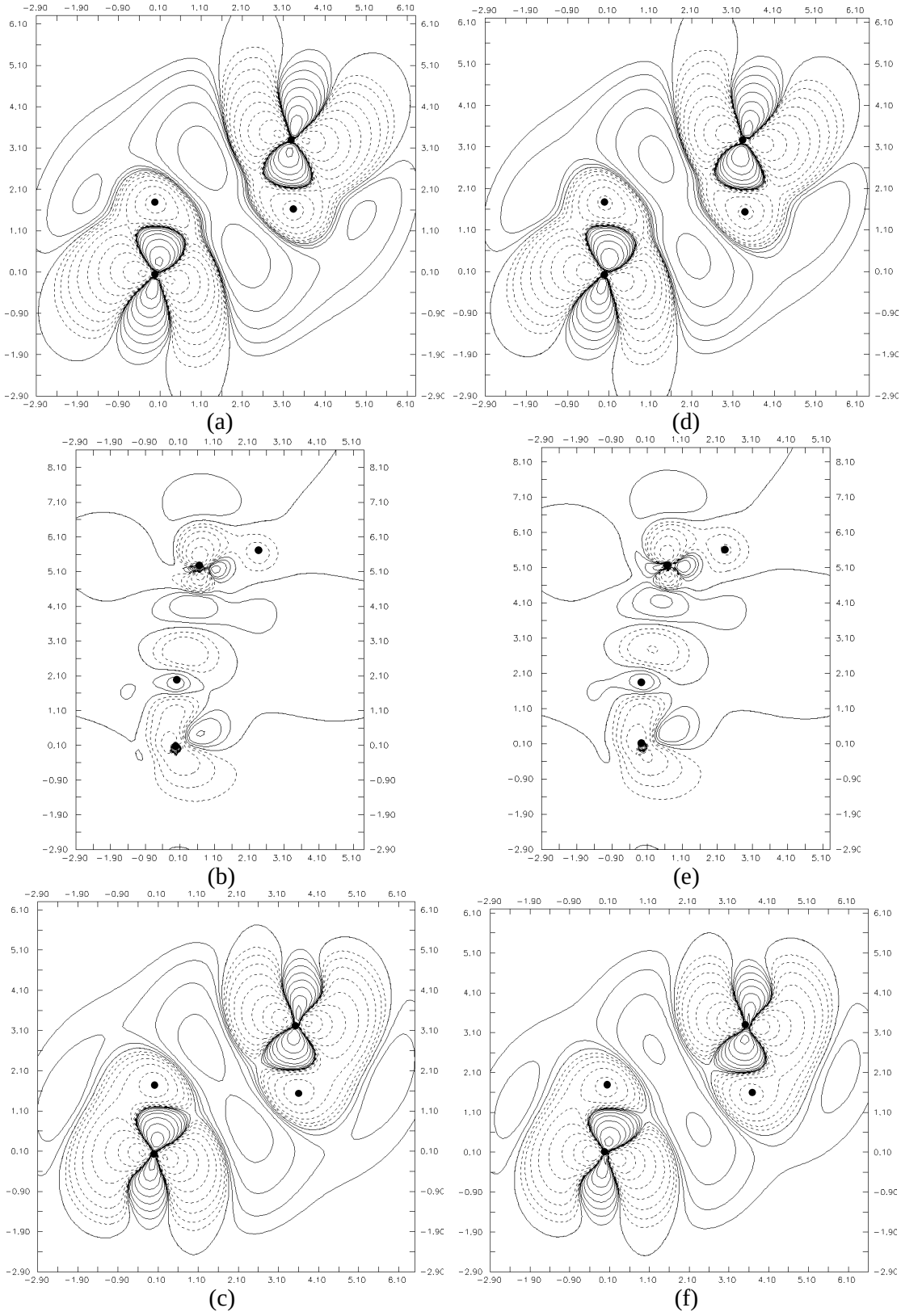


FIGURE 17: DFT//DFT-CHA//DFT Density difference isocontour maps. (a) BLYP/6-31G(d,p), (b) BLYP/6-31++G(d,p), (c) BLYP/6-311G(d,p), (d) B3LYP/6-31G(d,p), (e) B3LYP/6-31++G(d,p) and (f) B3LYP/6-311G(d,p). The nuclear positions are indicated by solid dots. Isodensity contours at $1.e^{-4}$, $2.e^{-4}$, $4.e^{-4}$, $8.e^{-4}$,etc...

Density difference maps at DFT levels of theory are collected in Figure 17. All maps have been obtained at the uncorrected geometry for each basis set. Maps for basis sets with and without diffuse functions were very similar among them, so in order to save space, only those corresponding to the 6-31G(d, p), 6-31G++(d, p) and 6-311G(d, p) are presented.

Interestingly, there are almost no differences between the maps obtained with BLYP or B3LYP functionals, despite their rather different effect on the topology of the charge density (see previous section), and the effect of the electronic relaxation.

Graphical representations for the basis sets including diffuse functions are as well very similar to the SCF ones (see Figures 17b and e). The main differences are located in the vicinity of the F atoms, mainly those of the H-acceptor molecule. Furthermore, the large effect of the electronic relaxation for the rest of the basis sets leading to cyclic structures is obvious from the maps (see Figures 17a, c, d f). Isocontour lines increase in the vicinity of both F atoms.

Again, the CHA density is more concentrated in the lone pairs regions of the F atoms and in both hydrogen atoms, whereas BSSE overestimates the charge density in the intermolecular region and along the intramolecular F···H bond paths. This overestimation of the charge density in the intermolecular zone is more concentrated in the zone where the two new bond critical points associated to the ring critical point are expected to appear in the uncorrected density. Nevertheless, all maps show qualitatively the same trends, disregarding the existence of the rcp on the uncorrected density or whether CHA correction changes this situation.

II.1.3.2.b Water dimer

The isodensity maps for the water dimer are depicted in Figures 18 and 19. They are very similar to the ones obtained for the hydrogen fluoride dimer. The same polarization pattern is observed on the heavy atoms and in the intermolecular bond region

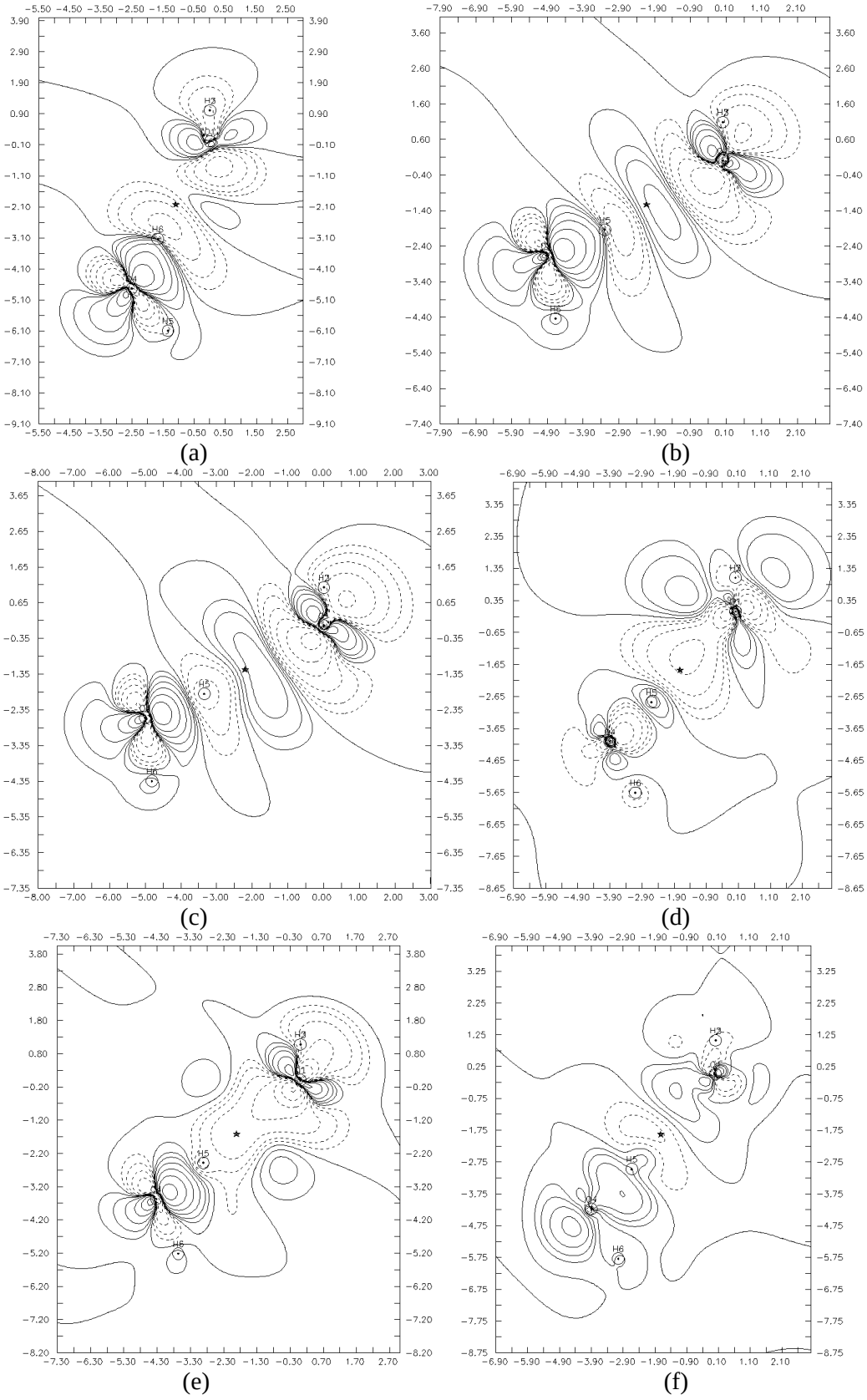


FIGURE 18: Water dimer SCF//SCF- CHA//SCF Density difference isocontour maps. (a) 6-31G, (b) 6-31G(d), (c) 6-31G(d,p), (d) 6-31++G(d,p), (e) 6-311G(d,p) and (f) 6-311++G(3df,2pd). Isodensity contours at $1.e^{-4}$, $2.e^{-4}$, $4.e^{-4}$, $8.e^{-4}$,etc...

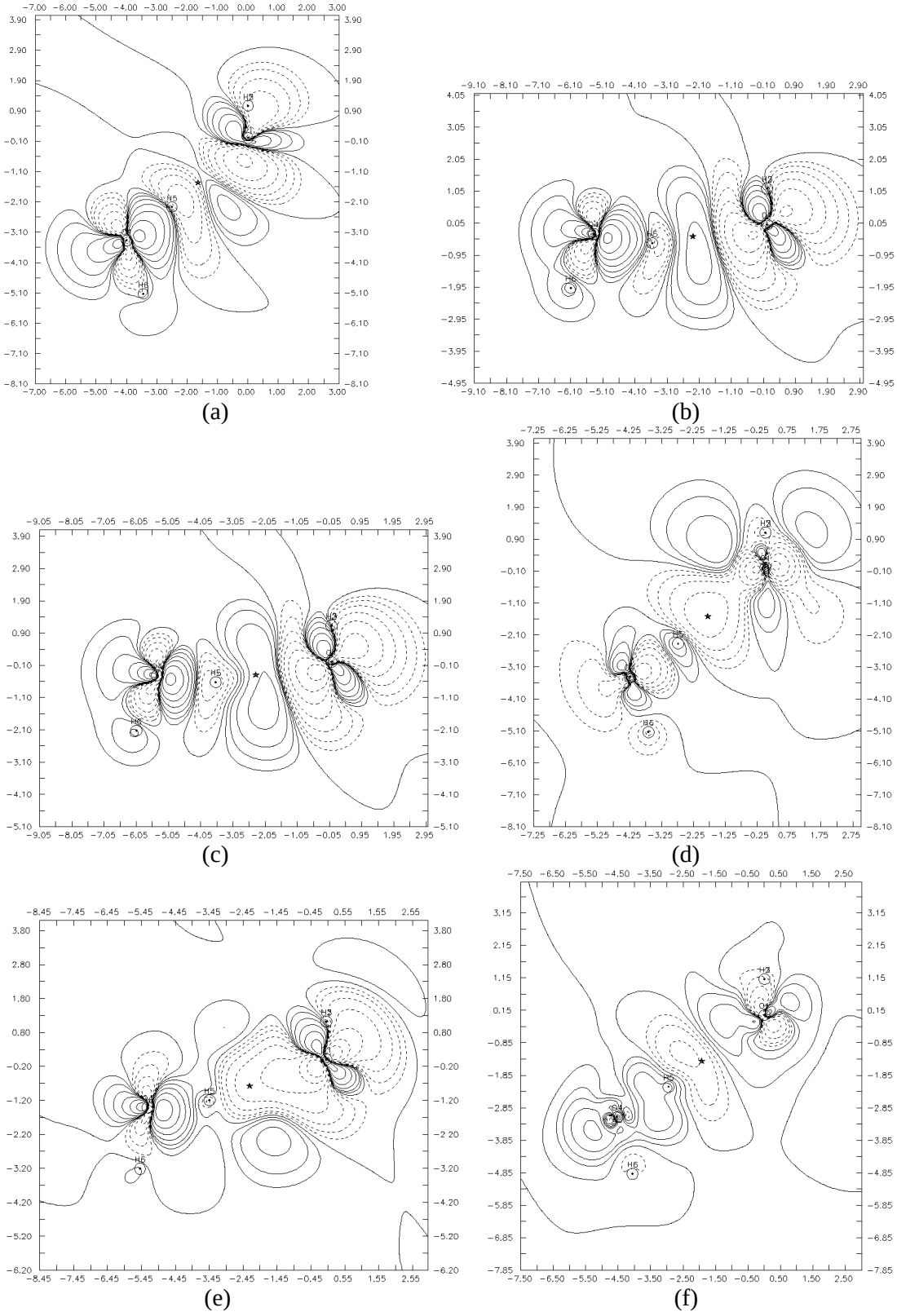


FIGURE 19: Water dimer B3LYP//B3LYP- CHA//B3LYP Density difference isocontour maps. (a) 6-31G, (b) 6-31G(d), (c) 6-31G(d,p), (d) 6-31++G(d,p), (e) 6-311G(d,p) and (f) 6-311++G(3df,2pd). Isodensity contours at $1.e^{-4}$, $2.e^{-4}$, $4.e^{-4}$, $8.e^{-4}$,etc...

SCF and DFT results are very similar. It can be seen that for the small 6-31G basis set the CHA correction has a strong influence on the electron density around each of the O atoms: with respect to a conventional calculation, CHA increases the electron density along the axes corresponding to the three intramolecular O-H bonds which are not involved in the intermolecular interaction, while there is a clear decrease of density along the axis corresponding to the remaining intramolecular O-H bond. Moreover, it is revealed that BSSE leads actually to an exaggeration of the density within the intermolecular region. The difference map corresponding to the 6-31G(d,p) basis set exhibits similar density redistributions for the two O atoms. However, in this case, there is a region around the intermolecular bcp where the density actually decreases after applying the CHA correction. In fact, this is the most common situation observed when using small basis sets, i.e. double zeta without diffuse functions. Finally, the difference map for the largest basis set used, 6-311++G(3df,2p), reveals that there is still some redistribution of density around the heavy atoms, and a region around the intermolecular bcp where CHA accumulates more density.

However, the density differences are rather small, as reflected in the $\rho_{bcp}(\mathbf{r})$ values obtained at the SCF//SCF and CHA//SCF levels of theory. Similar trends are found for density difference maps calculated at the B3LYP level of theory with the same basis sets.

In short, according to the density difference maps, the main effects of the BSSE correction in the electron density take place in the valence shells of the O atoms, both at the HF and B3LYP levels of theory. Thus, for basis sets with no diffuse functions, the BSSE correction leads to redistribution of electron density along the O-H intermolecular axis and centered around each of the heavy atoms. For the O of the donor moiety, the BSSE removes electron density from the O-H axis to an axis perpendicular to it, pointing towards the other intramolecular O-H bond. For the O of the acceptor molecule, the effect of BSSE correction is just the opposite. Indeed, the subtle density differences found in the intermolecular region may actually be just a consequence of the redistributions taking place around the heavy atoms. Finally, density difference maps corresponding to calculations with diffuse functions, Figures 18,19,(d) and (f) exhibit also maximal density differences around the heavy atoms, but

not the polarization patterns characteristic of the maps in Figure 18,19,(a), (b), (c) and (e).

II.1.3.2.c Formic acid dimer and uracil-water complex

Results for the hydrogen fluoride and water dimers suggest that inclusion of diffuse functions in the basis set is the main factor influencing the magnitude of the BSSE, while the level of theory and inclusion of more valence or polarization functions has a minor impact. Therefore, only the HF method and two different basis sets were used for each system, namely, 6-31G(d,p) and 6-31++G(d,p) for the formic acid dimer, and 6-31G(d) and 6-31+G(d) for the uracil-water complex. In this case, both the critical points of the density and the atomic positions have been represented (see SCHEME 18 for the atomic numbering). The maps for the calculations without diffuse functions correspond to the Figures 20a and 21a. As compared to the ones with diffuse functions (Figures 20b and 21b) the differences between the uncorrected density and the CHA-corrected are much more important.

The formic acid dimer exhibits the main trends found for the water dimer and hydrogen fluoride dimer. Thus, for the 6-31G(d,p) basis set, figure 20(a), there is a narrow intermolecular region where the BSSE overestimates the electron density. This region includes the rcp but not the two bcp's. That is, the electron density at the intermolecular rcp is increased upon correction. The main density redistribution effects take place in the valence shells of the heavy atoms. In particular, the O atoms in the hydroxyl and carbonyl moieties exhibit the density redistribution patterns directed along the bonding axes characteristic of H-donor and acceptor atoms, respectively.

Now, let us focus on the atoms not directly involved in the hydrogen bond formation, namely C_1, C_7, H_4, and H_9. The only appreciable difference is observed for the 6-31G(d,p), where the electron density slightly increases upon BSSE correction on the carbon atom domains. However, no polarization pattern similar to that of the atoms involved in the interaction is observed in any case.

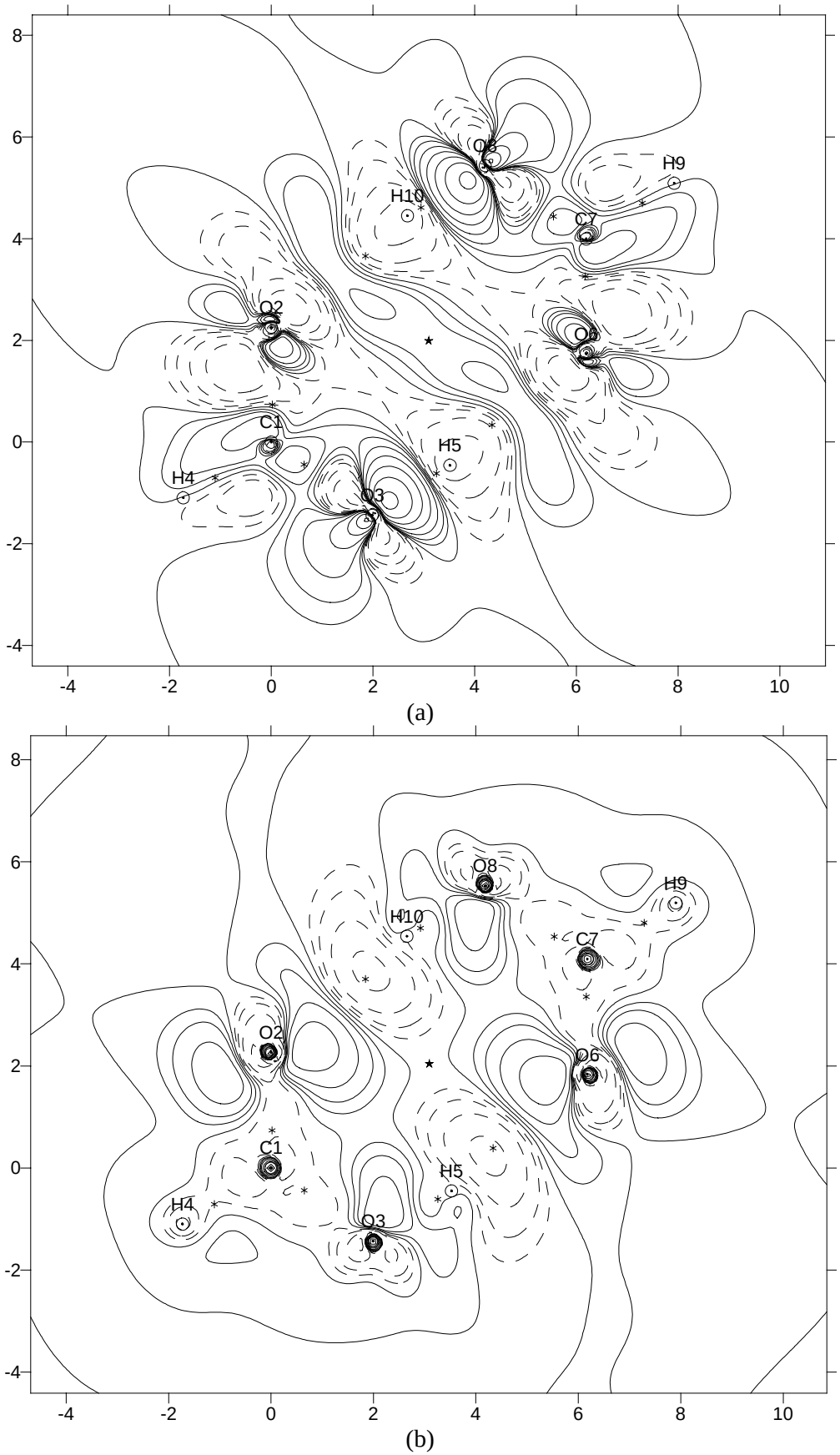


FIGURE 20: Formic acid dimer SCF//SCF- CHA//SCF density difference isocontour maps. (a) 6-31G(d,p), (b) 6-31++G(d,p). Isodensity contours at $1.e^{-4}$, $2.e^{-4}$, $4.e^{-4}$, $8.e^{-4}$,etc...

It seems that the effects on these skeleton atoms are residual. The $\Delta\rho$ value is influenced by the difference pattern on the atoms involved on the interaction, in order to counterpoise the electron redistribution upon BSSE correction. Indeed, the zone corresponding to the intramolecular C-O bonds seems to be much more affected by the polarization of the O atoms than that of the carbon atom.

When diffuse functions are added (see Figure 20(b)), SCF//SCF – CHA//SCF density differences in the intermolecular region become slightly negative. In this case, all the intermolecular cp's fall into this negative zone. The strong redistribution patterns associated to the H-donor and acceptor atoms in figure 20(a) are not found in this case. It appears that the negative region in the intermolecular zone is followed by alternating positive and negative regions at each side, with some positive regions focused strictly on the heavy nuclei. Nevertheless, the density difference decreases dramatically when including the diffuse functions. The maximum density differences observed with the 6-31G(d,p) basis set were -0.0190 and 0.0076 a.u., whereas for the 6-31++G(d,p) these values decrease to –0.0012 and 0.0015 a.u., respectively.

The results for the uracil-water dimer complex show a similar behavior. Once again, the use of diffuse functions dramatically decrease the difference between the uncorrected and the CHA densities.

However, the uracil-water complex has some features that may add interesting insights. First of all, it is a relatively large system, which allows to study the scope of the BSSE effects on molecular electron densities. Second, the O in the water moiety acts as H-donor and acceptor at the same time. Figure 21(a) corresponds to the SCF//SCF – CHA//SCF map with the 6-31G(d) basis set. In this case, a positive and a negative region are found in the intermolecular region. All the intermolecular cp's fall into the negative one. In the uracil molecule, the O_7 and the N_2 atoms exhibit the directional density redistribution patterns characteristic of H-acceptor and donor systems, respectively. Thus, the BSSE overestimates the electron density along the N_2-H_8 bond and underestimates it along the O_7-H_{13} intermolecular H bond.

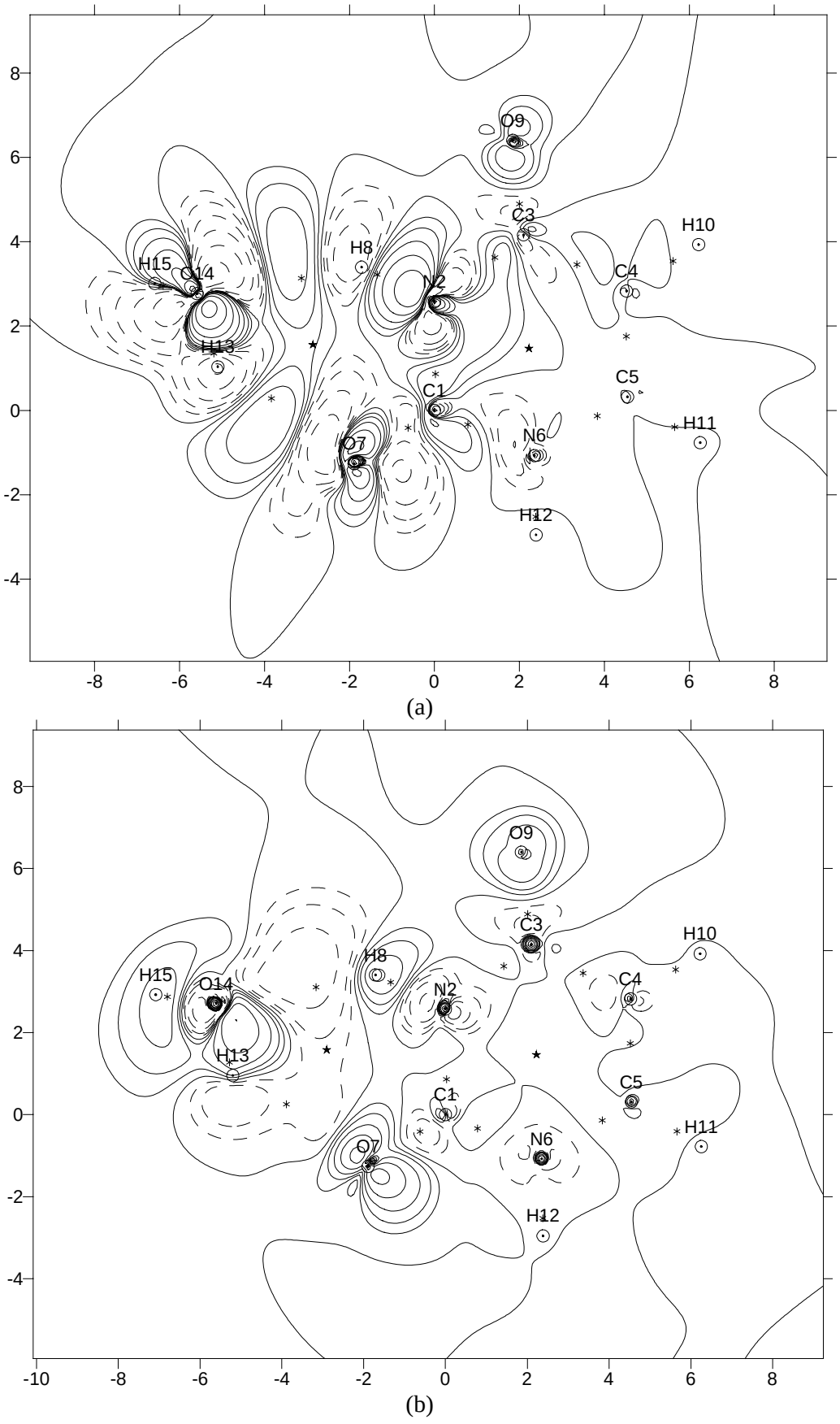


FIGURE 21: Uracil-water complex SCF//SCF- CHA//SCF density difference isocontour maps. (a) 6-31G(d,p), (b) 6-31++G(d,p). Isodensity contours at $1.e^{-4}$, $2.e^{-4}$, $4.e^{-4}$, $8.e^{-4}$,etc...

The O atom in the water molecule combines both features: the BSSE overestimates the density along the O_{14}-H_{13} intramolecular bond, and underestimates it along the intermolecular O_{14}-H_8 bond. The C_1 atom, which is bonded to an acceptor and to a donor atom, exhibits minor density redistributions, similar to the ones in the C atom in $(HCOOH)_2$. Finally, the effect of the BSSE in the rest of atoms is practically negligible, except for the carbonyl O_9 atom.

The 6-31+G(d,p) difference map (Figure 21(b)) presents a relatively large intermolecular region with negative values, which encloses all the intermolecular cp's. Significant density redistribution takes place only around the atoms directly involved in the H-bond interactions. As usual, the highly directional density redistribution patterns around heavy atoms found in the 6-31G(d) difference map are lacking in the 6-31+G(d) one. Atoms not involved in the intermolecular interaction do not exhibit meaningful density redistributions, except for O_9.

It is noticeable that this atom is not even directly bonded to the heavy atoms involved in the intermolecular bond. We believe that his effect is only due to the fact that in the vicinity of this atom the accumulation of electron density is much larger than for the rest of the molecule, basically due to the existence of two free electron pairs. The absolute density difference is larger than for other atoms but relative to the total electron density on the considered domain should be of the same order.

Nevertheless, the maximum density differences are about one order of magnitude larger when no diffuse functions are included.

II.1.3.3 Quantitative determination of the BSSE effect by means of Quantum Molecular Similarity techniques

The differences between the conventional and CHA density have been graphically and numerically pointed out by means of isodensity plots and the AIM theory. Another possibility is to use Quantum Molecular Similarity[139] (QMS) measures to quantify the differences between conventional and BSSE-corrected densities. QMS has been previously used for the analysis of atomic and molecular one-electron densities in a variety of contexts, e.g., to assess the effects of electron correlation on molecular densities[140a], to quantify the electronic reorganisation taking place in several reactions[148,149], to analyze the quality of several basis sets[150] and to evaluate the interactions of the electron density with external electric fields[150] or continuos solvents[151]. Finally, QMS has recently been used to compare contracted second-order electron densities[152].

For any molecule, one can define an overlap-like similarity measure between the conventional or uncorrected density, $\rho_U(\mathbf{r})$, and the CHA or BSSE-corrected density at the same geometry, $\rho_C(\mathbf{r})$,

$$Z_{UC} = \int \rho_U(\mathbf{r})\rho_C(\mathbf{r})d\mathbf{r}\,. \qquad (100)$$

The value of Z_{UC} is proportional to the degree of overlap between $\rho_U(\mathbf{r})$ and $\rho_C(\mathbf{r})$. A normalized distance index is defined as,

---

[148] M. Solà, J. Mestres, R. Carbó and M. Duran, J. Am. Chem. Soc. **116**, 5909 (1994).

[149] X. Fradera, L. Amat, M. Torrent, J. Mestres, P. Constans, E. Besalú, J. Martí, S. Simon, M. Lobato, J.M. Oliva, J.M. Luis, J.L. Andrés, M. Solà, R. Carbó, M. Duran, J. Mol. Struct. (Theochem), **37**, 1171 (1996)

[150] S. Simon, M. Duran, J. Chem. Phys. **107**, 1529 (1997)

[151] J. Mestres, M. Solà, R. Carbó, F.J. Luque and M. Orozco, J. Phys. Chem. **100**, 606 (1996)

[152] X. Fradera, M. Duran, J. Mestres, Theor. Chem. Acc. **99**, 44 (1998), X. Fradera, M. Duran, J. Mestres , "Comparison of Quantum Similarity Measures Derived from One-Electron, Intracule, and Extracule Densities", *Adv. Molec. Sim.*, Vol. 2, 215-243, JAI Press, 1998.

$$D_{UC} = \left(Z_{UU} + Z_{CC} - 2Z_{UC}\right)^{1/2}. \tag{101}$$

where Z_{UU} and Z_{CC} and the self-similarities associated to $\rho_U(\mathbf{r})$ and $\rho_C(\mathbf{r})$, respectively, that is, the overlap of each of these densities with itself. D_{UC} in Eq. (101) tends to zero for the limiting case of identical electron densities, which corresponds to a null BSSE, and has no general upper limit.

Our aim is thus to introduce Quantum Molecular Similarity as a suitable tool for the quantitative evaluation of the electron redistribution that takes place in a molecular complex after correcting for BSSE. The hydrogen fluoride, water and formic acid dimers, and the uracil-water complex are used to introduce this kind of analysis. Since the effects of BSSE on the electron densities of these complexes have been already analyzed at the SCF and DFT levels of theory, we will simply present the similarity-based analysis.

All similarity calculations are performed between molecular densities at the same molecular geometry (both the conventional and CHA geometries have been used). Therefore, similarity measures and indices account merely for the electronic relaxation effects associated to the BSSE.

For the water dimer complex, Table XLI collects all distance indices (D_{CU}) obtained from the comparison of conventional and CHA densities by means of overlap similarity measures, for all the combinations of *ab initio* methodology and basis set. For each case, the comparisons have been carried out both at the conventional and BSSE-corrected optimized molecular geometries. In general, the D_{CU} values are very small, showing that BSSE is in fact a subtle error which leads only to small deformations of the electron density.

The same trends are found for the comparisons at the conventional and CHA geometries; however, D_{CU} values are always slightly larger at the uncorrected geometries, both for HF and B3LYP calculations. By considering the electronic and relaxation redistributions caused by the CHA correction as perturbations with respect to a conventional calculation (SCHEME 17), one can think that, at the conventional geometry, CHA tends to exaggerate the density redistribution needed to correct the

BSSE. In contrast, at the CHA optimized minimum, nuclear relaxation may account for the major part of the BSSE correction; therefore, the density redistribution associated to the BSSE correction is less important in this case. Altogether, this trends confirm the importance of BSSE corrections to the molecular geometry for a proper description of this kind of hydrogen-bonded complexes.

TABLE XLI

Distance indices between conventional and BSSE-corrected densities calculated with several basis sets at the SCF, and B3LYP levels of theory, for the water dimer.

Basis Set	**Method**	**Conventional Geometry**	**Corrected Geometry**
6-31G	SCF	0.004139	0.003911
	B3LYP	0.006283	0.005800
6-31G(d,p)	SCF	0.004592	0.003973
	B3LYP	0.008562	0.006749
6-31G++(d,p)	SCF	0.002654	0.002499
	B3LYP	0.003405	0.003236
6-311G(d,p)	SCF	0.004011	0.003714
	B3LYP	0.006918	0.005750
6-311G++(3df,2pd)	SCF	0.000948	0.000894
	B3LYP	0.001457	0.001422

For the SCF and B3LYP calculations, the D_{CU} values at the conventional and CHA geometries decrease upon increase of basis set size, in agreement with the corresponding decrease of the error in the stabilization energies. In fact, the inclusion of diffuse functions and high order polarization functions in the basis set appears to be the main factor for decreasing D_{CU}. For instance, at the SCF level of theory, D_{CU} is ca. 0.004, for all the basis sets without diffuse functions, 0.003 for the 6-31++G(d,p) basis set, and 0.0009 for the 6-311G++(3df,2pd) basis set. Similar trends are found at the B3LYP level: D_{CU} lies between 0.006 and 0.008 for basis sets without diffuse functions, while it is ca. 0.003 for 6-31++G(d,p), and ca. 0.0015 for the largest basis set. From these values it is also apparent that, according to the distance indices obtained, the effect of BSSE on the electron density is larger at the DFT than at the HF level of theory.

As for the hydrogen fluoride dimer, Table XLII gathers the distance indices corresponding to the comparisons between corrected and uncorrected densities, for all the combinations of level of theory and basis sets. The distances obtained are always very small, confirming again that the electron redistribution associated to the CHA *perturbation* is very small. For instance, the largest indices are obtained for the calculations with the 6-31G basis set at the conventional molecular geometry. In this case, D_{UC} is 0.0071 and 0.0213, at the SCF and B3LYP levels, respectively. Taking into account that the distance index between the conventional SCF and B3LYP densities of the $(HF)_2$ dimer, at the same geometry, is 0.0277, one can assume that the electron redistributions due to the elimination of the BSSE are comparable to those caused by considering electron correlation, with respect to a conventional HF density.

TABLE XLII

Distance indices between conventional and BSSE-corrected densities calculated with several basis sets at the SCF and B3LYP levels of theory, for the hydrogen fluoride dimer.

Basis Set	**Method**	**Conventional Geometry**	**Corrected Geometry**
6-31G	SCF	0.007101	0.006734
	B3LYP	0.021325	0.011470
6-31G(d,p)	SCF	0.008113	0.007450
	B3LYP	0.020160	0.012080
6-31G++(d,p)	SCF	0.000971	0.000936
	B3LYP	0.001490	0.001385
6-311G(d,p)	SCF	0.004174	0.003740
	B3LYP	0.013927	0.006043
6-311G++(3df,2pd)	SCF	0.001538	0.001459
	B3LYP	0.001844	0.001796

Regarding the dependence of the distance indices on the level of theory, basis set and molecular geometry, the trends are the same found for the water dimer. Thus, distances between corrected and uncorrected densities are always larger at the B3LYP level of theory. Moreover, distances calculated at minima on the conventional PES are always larger than the equivalent distances at the CHA minima. Finally, as expected,

distance indices decrease when improving the quality of the basis set. Again, diffuse functions appear to be very important in order to minimize the BSSE on the density.

Finally, the distance indices between corrected and uncorrected densities calculated at the SCF optimized geometries for the larger complexes are collected in Table XLIII. Inclusion of diffuse functions halves the distance indices between both densities. Distances calculated for both systems are quite similar even though the difference on the strength of the respective intermolecular bonds (~14 Kcal/mol for the formic acid dimer vs. ~8 Kcal/mol for the uracil-water complex) and their molecular size.

TABLE XLIII

Distance indices between conventional and BSSE-corrected densities at the SCF levels of theory, for formic acid dimer and uracil-water complex.

Basis Set	**Formic acid dimer**	**Uracil-water**
6-31G(d, p)	0.006620	0.06402
6-31G+(d,p)	0.003195	0.03239

The effect of BSSE on the density of these complexes is quantitatively comparable to the effect on the water and hydrogen fluoride dimers, which agrees with the observation that the BSSE effect is mostly localized in the intermolecular bond zone and the atoms involved in it.

As a conclusion, the distance indices for the studied complexes are in overall agreement with the analyses of the critical points and the density difference maps, and provide a quantitative measure of changes caused by correction or removal of BSSE.
In general, all the analyses performed show that the effects of BSSE on the electron density are parallel to those on the molecular geometry and stabilization energies. Improvement of the quality of the basis set leads always to smaller differences

between conventional and CHA molecular properties and densities; in particular, inclusion of diffuse functions is of great importance for minimization of the BSSE. The same trends are found at the SCF and DFT levels of theory; however, distance indices reveal that the effects of BSSE on the electron density are normally larger for the DFT than for the SCF calculations. Moreover, distance indices are always larger at minima optimized at the conventional PES, respect to indices corresponding to minima on the CHA PES.

All in all, Quantum Molecular Similarity has been found to be useful for evaluating quantitatively the differences between conventional and CHA-corrected densities.

II.1.3.4 Chemical Energy Component Analysis (CECA).

With semiempirical quantum methods, the molecular energy can be expressed in terms of one- and two-center contributions only. This property allows for a straightforward decomposition of the molecular energy into atomic (one-center) and interatomic (two-center) contributions. In contrast, calculation of molecular energies with *ab-initio* methods involves also three- and four-center terms. These terms may contribute significantly to the total energy and cannot be generally ignored; however, it is difficult to attach a direct chemical significance to these multicenter terms.

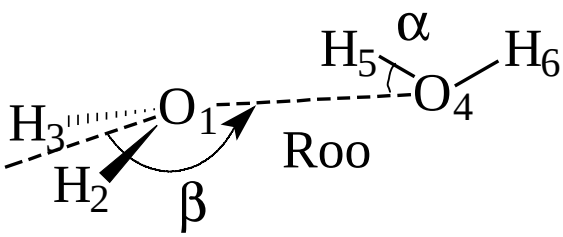
The basic idea in the CECA is to use a projective integral expansion scheme that allows to express approximately each multicenter term as a summation of one- and two-center terms. The theory behind the CECA is tightly related to the CHA (see Ref. 153 for a detailed description). The main advantage of this approach compared to the classical bond order analysis is that bonding interactions with formally the same multiplicity (single, double, triple bonds) can be now clearly differentiated in terms of energetic (static) contributions to the overall energy of the system. Also, it allows to distinguish between bonding and antibonding interactions. However, one must take into account that this decomposition is exact only for diatomic molecules, where no multicenter contributions are possible. In the general case, the sum of all the energy contributions does not match exactly the total molecular energy. Mayer states that the CECA reproduces HF energies of small molecules with a precision between 10 and 40 mHartrees ($\approx$ 6 to 25 kcal·mol^{-1})[153]. This is a relatively small error, compared to the total molecular energy. Furthermore, one should take into account that this error arises from the summation of *all* the one- and two-center energy contributions. One can expect that the individual components have much smaller errors. Therefore, the accuracy of the CECA decomposition scheme should be sufficient in most cases.

A second problem of the CECA is related to the fact that the energy decomposition is performed in the Hilbert space. That is, each atomic orbital or basis function is assigned to a single atom and the partition of each energetic term is carried out by projecting over the subspace spawned by these atomic orbitals. Practically, the

[153] I. Mayer, Chem. Phys. Lett. 332, 381 (2000).

basis functions are supposed to belong to the atom in which they are centered. Then, one- and two-center contributions are related to atomic and diatomic components. However, the results of this kind of decomposition schemes can be very dependent on the basis set used for the calculation. Particularly, these analysis may lose significance when diffuse functions are included in the basis set. For instance, it is well known that Mulliken charges, also based on the formal partition of the atomic orbital's space, have little chemical meaning when diffuse basis functions are used.

This problem can be solved by performing a similar decomposition in the Euclidean space, for instance in the frame of the AIM theory[154]. It can be shown that a partitioning of that kind, that can be connected to the CECA one by a simple mapping of the integrals[155], can provide an exact decomposition of the HF molecular energy. Several test calculations revealed some differences between the CECA and the AIM-based energy partitioning[154]. Unfortunately, the huge computational cost that implies the evaluation of double integrations over two AIM atomic basins prevents the use of this methodology for the present work, and only CECA approximated results will be shown.



SCHEME 19

Tables XLIV-XLVII gather the results of the CECA decomposition for all the HF calculations of water dimer, formic acid dimer and uracil-water complex. Table XLIV collects the one- and two-center energy components obtained for the water dimer with the 6-31G(d,p) and 6-31++G(d,p) basis sets. Atomic energies are always negative (stabilizing), as well as the interaction between bonded atom pairs. Some terms like the O-O and H-H interactions are repulsive, which agrees with the chemical intuition. Direct comparison of the energy components obtained with different basis

[154] P. Salvador, M. Duran, I. Mayer, J. Chem. Phys. 115, 1153 (2001).

[155] I. Mayer and A. Hamza, Theor. Chem. Acc. 105, 374 (2001).

sets is not very convenient since the total molecular energy can be very different. Hence, since we are interested in the analysis of the effects of the BSSE in the energy, only selected differences between the SCF//SCF and CHA//SCF values for each basis set are discussed.

TABLE XLIV

SCF one- and two-center energy components for the $(H_2O)_2$ complex for the 6-31G(d,p) and 6-31++G(d,p) (lower triangle, in cursive) in a.u.. The values in parenthesis correspond to the H_2-H_3 diatomic term.

Atom	**O_1**	**H_2**	**O_4**	**H_5**	**H_6**
O_1	-74.3319 *-74.3758*	-0.7854 (0.0450)	0.0942	-0.1016	-0.0331
H_2	*-0.7819* *(0.0609)*	0.0450 *-0.0736*	-0.0380	0.0280	0.0142
O_4	*0.1087*	*-0.0422*	-74.3423 *-74.3884*	-0.7859	-0.7909
H_5	*-0.0970*	*0.0332*	*-0.7853*	-0.0612 *-0.0599*	0.0469
H_6	*-0.0406*	*0.0164*	*-0.7781*	*0.0674*	-0.0810 *-0.0828*

We use $E^{BSSE}(A)$ and $E^{BSSE}(A,B)$ to denote the BSSE effect in one- and two-center energy components involving atom A and the atomic pair A,B, respectively. Negative $E^{BSSE}(A)$ and $E^{BSSE}(A,B)$ values correspond to energy components that are too stabilizing in the uncorrected calculation, because of the BSSE. That is, the given one or two-center component is more stabilizing (less destabilizing) for the SCF//SCF than for the CHA//SCF calculation. Inversely, positive values correspond to energy components that are lower in energy for the CHA than for the SCF calculations. Tables XLV-XLVII also lists the total energy difference for each monomer as well as the correction to the static interaction energy, computed by summing up all the corresponding CECA one- and two-center terms.

Note that the sum of all the CECA intermolecular energy components must be

clearly distinguished from the conventional stabilization and interaction energies, the former reported in Section II.1.1.1. In the supermolecular approach, the interaction energy is defined as the difference between the energy of the complex and the energies of the monomers at the complex's geometry. The stabilization energy is obtained by taking into account also the energetic penalty due to the geometrical distortions of the monomers in the complex. Hence, both the interaction and the stabilization energies take into account the electronic relaxation, as the wavefunction of the monomers is computed to obtain the corresponding energies. In the case of the stabilization energy, the nuclear relaxation of the monomers is also taken into account. In contrast, the static interaction energy account only for local energetic interactions extracted uniquely from the complex's wavefunction, and is easily obtained as the summation of all the CECA energy components associated to intermolecular two-center interactions.

TABLE XLV

CECA analysis of the $(H_2O)_2$ at the SCF level of theory for several basis sets. Given values represent energetic differences between SCF//SCF and CHA//SCF calculations in kcal/mol. E^{BSSE} $(O_1\text{-}H_5)$, $(O_4\text{-}H_5)$ and (H_5), hold for the BSSE contribution on selected two- and one-center interactions (see SCHEME 19). ΔE_D , ΔE_A , and ΔE_{int} are the static BSSE contributions on donor, acceptor and interaction energies, respectively, computed from the CECA one- and two-center terms. Last two columns give the exact and the CECA approximated BSSE

Basis	$\mathbf{E^{BSSE}(O_1,H_5)}$	$\mathbf{E^{BSSE}(O_4,H_5)}$	$\mathbf{E^{BSSE}(H_5)}$	$\mathbf{\Delta E_{int}}$	$\mathbf{\Delta E_D}$	$\mathbf{\Delta E_A}$	**BSSE**	$\mathbf{BSSE_C}$
6-31G	-9.4	-7.0	10.0	-8.5	3.6	4.6	-0.50	-0.23
6-31G(d)	-12.4	-1.3	8.0	-11.9	3.5	8.2	-0.59	-0.02
6-31G(d,p)	-11.1	3.1	3.1	-10.4	2.6	8.2	-0.70	-0.04
6-31++G(d,p)	6.5	-1.0	-1.3	9.7	-4.4	-7.4	-0.55	-2.17
6-311G(d,p)	-9.5	3.0	1.1	-10.4	2.1	8.0	-1.03	-0.26
6-311++G(d,p)	15.6	-14.7	6.4	9.0	-4.7	-8.2	-0.54	-3.94

In a similar way, the static monomer energies can also be obtained by collecting all the one- and two-center CECA components involving the atoms of the given monomer. The summation of all BSSE corrections to each (static) monomer energy

and to the (static) intermolecular component yields the total correction to the complex energy. The overall BSSE calculated as the difference between the uncorrected (SCF//SCF) and corrected (CHA//SCF) energies is also reported. Comparison of these values gives a measure of the accuracy of the CECA partition in each case.

For the 6-31G basis sets, the BSSE is manifested mainly in the energy components related to H_5, which is the H participating in the intermolecular bond. The principal destabilizing contribution comes from the one-center component in H_5, while the major stabilizing contributions correspond to two-center components involving H_5 and other atoms. Thus, $E^{BSSE}(H_5)$ is +10.0 kcal·mol^{-1}, while $E^{BSSE}(O_1,H_5)$ and $E^{BSSE}(O_4,H_5)$ are -9.4 and -7.0 kcal·mol^{-1}, respectively. However, these trends are not general for all the calculations. For all the basis sets without diffuse functions, (a), (b) (c) and (e), $E^{BSSE}(O_1,H_5)$ is ca. -10 kcal·mol^{-1}. $E^{BSSE}(H_5)$ and $E^{BSSE}(O_4,H_5)$ also contribute to the BSSE, but to a small extend, compared to the 6-31G results. $E^{BSSE}(H_5)$ is always positive, but $E^{BSSE}(O_4,H_5)$ can be positive or negative, depending on the basis set. In some cases, other components exhibit also significant BSSE. In general, for all these basis sets, the BSSE destabilizes the two water monomers, especially the donor one, but makes the intermolecular component more stabilizing. The overall effect of the BSSE in the static interaction energy is always stabilizing. This clearly shows that the interaction between the monomers is artificially enhanced by the BSSE.

The results of the analysis are quite different when diffuse functions are included. For the 6-31++G(d,p) basis set, $E^{BSSE}(O_1,H_5)$ is +6.5 kcal·mol^{-1}, and $E^{BSSE}(H_5,H_6)$ component is +4.9 kcal·mol^{-1}. This is compensated mainly in the one-center components of the O atoms, which are ca. -6.6 kcal·mol^{-1} each. For the 6-311++G(d,p) basis set, $E^{BSSE}(O_1,H_5)$ is +15.6 kcal·mol^{-1} and $E^{BSSE}(H_5)$ is +6.4 kcal·mol^{-1}, while $E^{BSSE}(O_4,H_5)$, $E^{BSSE}(O_1,O_4)$ and $E^{BSSE}(O_1)$ are -14.7, -6.6 and -4.8 kcal·mol^{-1}, respectively. In terms of molecular and intermolecular components, the BSSE contribution to the intermolecular term is always unfavorable (ca. +9 kcal·mol^{-1} in both cases). For the 6-31++G(d,p) and 6-311++G(d,p) basis sets, the overall BSSE contribution is negative for both the donor and acceptor molecules. Altogether, the effect of BSSE on the molecular static interaction energy is always stabilizing, but the sign of the contributing terms is reversed, compared to the calculation with no diffuse functions.

Table XLIV also lists also the BSSE calculated i) as the difference between SCF//SCF and CHA//SCF energies, and ii) as the summation of the BSSE in each one- and two-center energy component. The difference between the two values can be used to estimate the accuracy of the CECA. In general, the differences are significant, taking into account that the BSSE is generally small. For the basis sets with no diffuse functions, the difference is always less than 0.8 kcal·mol^{-1}, and the CECA always underestimates the magnitude of the BSSE. On the contrary, for the 6-31++G(d,p) and 6-311++G(d,p) basis set, the CECA overestimates the magnitude of the BSSE by ca. 1.5 and 3.5 kcal·mol^{-1}, respectively. Similar conclusions can be drawn when comparing SCF//SCF and CHA//CHA energies, so the results are not reported.

TABLE XLVI

CECA analysis of the formic acid dimer . Given values represent energetic differences between SCF and CHA at the SCF (uncorrected) geometries in kcal/mol. See SCHEME 18a for the selected one- and two-center energy differences.$\Delta E_{A\text{-}D}$, and ΔE_{int} are the static differences on the formic acid moiety and the interaction energy, respectively, computed from the CECA one- and two-center terms. BSSE and $BSSE_C$ give the exact and the CECA approximated difference between the CHA and the SCF energies.

Basis	$\mathbf{E^{BSSE}(C_1,O_2)}$	$\mathbf{E^{BSSE}(H_5,O_6)}$	$\mathbf{E^{BSSE}(H_5)}$	$\mathbf{\Delta E_{int}}$	$\mathbf{\Delta E_{A\text{-}D}}$	**BSSE**	$\mathbf{BSSE_C}$
6-31G(d,p)	15.6	-14.1	5.0	-29.2	13.8	-1.63	-1.42
6-31++G(d,p)	-3.6	-7.3	8.4	-16.6	7.6	-0.78	-1.17

The results for the formic acid dimer are presented in Table XLVI. For the 6-31G(d,p) basis set the main contributions to the BSSE are $E^{BSSE}(C_1,O_2)$ (15.6 kcal·mol^{-1}) and $E^{BSSE}(O_2,H_{10})$ (-14.1 kcal·mol^{-1}). Note that due to the symmetry, equivalent contributions arise from the C_7, O_6 and H_5 atoms. $E^{BSSE}(H_{10})$ makes a smaller but significant contribution (5.0 kcal·mol^{-1}). For the 6-31++G(d,p) basis set, $E^{BSSE}(C_1,O_2)$ and $E^{BSSE}(O_2,H_{10})$ are smaller (-3.6 and -7.3 kcal·mol^{-1}, respectively). In contrast, $E^{BSSE}(H_{10})$ is significantly larger, +8.4 kcal·mol^{-1}. The overall picture is the same in both calculations: the BSSE stabilizes the two-center components related to the H-bond interactions, as well as the one-center components in the acceptor atoms, but destabilizes the H atoms participating in the intermolecular bond. Altogether, the BSSE destabilizes each formic acid monomer but increases the attractive intermolecular energy component. Hence, the overall contribution of the BSSE to the

molecular interaction is stabilizing for both basis sets. For this complex, the error in the CECA analysis is quite small, compared to $(H_2O)_2$. For the small basis set, the BSSE calculated as the sum of the CECA components is ≈ 0.2 kcal·mol^{-1} lower in magnitude than the real value. For the 6-31++G(d,p), CECA overestimates the magnitude of the BSSE error by ≈ 0.4 kcal·mol^{-1}.

TABLE XLVII

CECA analysis of the uracil-water complex . Given values represent energetic differences between SCF and CHA at the SCF (uncorrected) geometries in kcal/mol. See SCHEME 18b for the selected one- and two-center energy differences.ΔE_U , ΔE_W, and ΔE_{int} are the static differences on the uracil, water and the interaction energy, respectively, computed from the CECA one- and two-center terms.

Basis	$\mathbf{E^{BSSE}(C_1,O_7)}$	$\mathbf{E^{BSSE}(H_8,O_{14})}$ ***(H13,O7)***	$\mathbf{E^{BSSE}(H_8)}$ ***(H13)***	$\mathbf{\Delta E_{int}}$	$\mathbf{\Delta E_U}$	$\mathbf{\Delta E_W}$	**BSSE**	$\mathbf{BSSE_C}$
6-31G(d)	18.5	-16.0 *-12.6*	7.6 *4.8*	-23.9	10.4	11.7	-1.48	-1.45
6-31+G(d)	-1.3	-2.3 *-11.8*	3.2 *9.2*	1.0	-8.3	-2.3	-0.78	-8.71

The results of the CECA analysis for the uracil-water complex are collected in Table XLVII. The main trends are very similar to those found for the formic acid dimer. Thus, for both basis sets, the BSSE has a large stabilizing contribution to the two-center components related to the H bonds, reflected in the large negative values of $E^{BSSE}(O_7,H_{13})$ and $E^{BSSE}(O_{14},H_8)$, while $E^{BSSE}(H_8)$ and $E^{BSSE}(H_{13})$ are positive. The main difference between the 6-31G(d) and 6-31+G(d) results is that the BSSE in the energy components associated to the O_{14}-H_8 interaction are much smaller when diffuse functions are used. Furthermore, $E^{BSSE}(C_1,O_7)$ has a significant destabilizing contribution, for the 6-31G(d) results, but small and negative for 6-31+G(d). There are other components that have important contributions to the BSSE in the 6-31+G(d) calculation. Some of these contributions come from atom pairs that are not directly bonded, but are connected through a common atom. However, most of all the significant contributions involve the atoms that directly participate in a H bond: C_1, N_2, O_7, H_8, H_{13} or O_{14}. In terms of intramolecular and intermolecular components, the BSSE stabilizes the intermolecular component and destabilizes the intramolecular ones, for the 6-31G(d) calculation, and inversely for the 6-31+G(d) one. As usual, the

overall contribution of the BSSE to the interaction energy is attractive. For the 6-31G(d) results, the CECA partition is nearly exact, and the error in the calculation of the BSSE correction to the energy is only 0.03 kcal·mol^{-1}. In contrast, for the 6-31+G(d) case, the difference between the BSSE calculated by using the CECA and the supermolecular approach is ca. 8 kcal·mol^{-1}.

II.1.3.5 Concluding Remarks.

It is interesting to remark that the main effects of the BSSE correction on the electron density of the water dimer are very similar to those found for the hydrogen fluoride dimer. The patterns of electron redistribution caused by the removal of the BSSE at frozen geometries for $(HF)_2$ and $(H_2O)_2$ are very similar. Indeed, for calculations without diffuse functions, the main feature of the difference maps is the redistribution of electron density in the valence shells of the heavy atoms in both cases. Moreover, similar trends are found for the 6-31G(d,p) and 6-31G(d) calculations on the formic acid dimer and uracil-water complexes, respectively. Furthermore, addition of diffuse functions leads to similar effects for all the systems analyzed: an overall decrease of the differences between corrected and uncorrected densities, negative differences in the intermolecular region, and lack of the highly directional density redistribution patterns in heavy donor and acceptor atoms that are observed with smaller basis sets.

In fact, some of the differences between the SCF//SCF and CHA//SCF electron densities appear to be at odds with simple chemical intuition. For instance, it might be expected that the BSSE correction should weaken the intermolecular interaction and therefore lead to a decrease of the electron density in the intermolecular region. Actually, in many cases, the BSSE correction works in the opposite direction, leading to an accumulation of electron density in the intermolecular region. Moreover, the BSSE correction also decreases the electron density in the intramolecular bonds of the donor moieties. In general, it should be taken into account that the CHA//SCF results used in the difference maps do not correspond to stationary points on the BSSE-corrected surface. It is well-known that geometry relaxation is necessary for fully correcting the BSSE. In fact, when nuclear relaxation is taken into account, there is

always a depletion of the electron density in the intermolecular region, as reflected in the properties of the intermolecular bcp's.

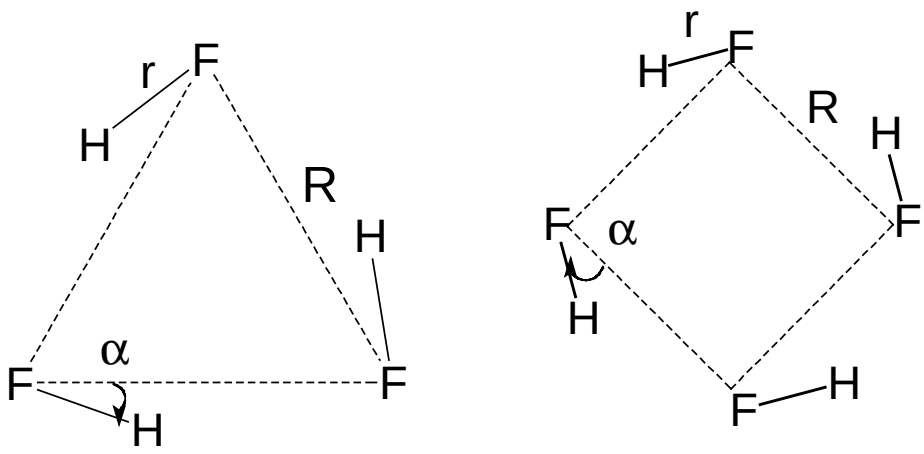
The study of larger systems, like the formic acid dimer, and especially the uracil-water complex, reveals that the effects of BSSE on the electron density are generally restricted to the intermolecular region and especially to the atoms directly involved in the intermolecular interaction and their first-neighbors.

The CECA decomposition scheme has been found to be a valuable tool for analyzing the local effects of BSSE in terms of atomic and interatomic contributions. However, one has to be aware that the CECA decomposition is not exact. Therefore, the applicability of this method to analyze the subtle effects of the BSSE on the molecular energy depends on the accuracy of the approximation. In general, for the calculations reported in this paper, the accuracy of the decomposition, calculated as the difference to the true BSSE, is good or acceptable when basis sets without diffuse functions are used. In these cases, the results of the CECA analysis are in agreement with chemical intuition: the BSSE generally destabilizes the purely intramolecular energies of the two molecules forming the complex, but it enhances the intermolecular energy component. The final result is that BSSE always leads to more attractive interaction energies. In general, it is worth to note that, although the BSSE in the total molecular energy is usually small, the individual atomic or interatomic contributions can be quite large.

When diffuse functions are taken into account, the results of the analysis are just the opposite. That is, in general, small BSSE energy stabilization results from the combination of a large stabilization of the intramolecular energies and a destabilization of the intermolecular term. However, the validity of the CECA analysis in these cases is questionable for two reasons. First, the accuracy of the CECA decomposition is very low (the BSSE is overestimated by several $kcal \cdot mol^{-1}$, except for the formic acid dimer). Second, the identification of the one- and two-center components with atomic and interatomic contributions is questionable when diffuse functions are involved. Nevertheless, the CECA results in these particular cases seem to agree with the density difference maps in the sense that the differences tend to be smaller when diffuse functions are included but they are more delocalized. Indeed, the H atoms involved in the intermolecular H-bonds have similar BSSE effect

on the two-center components with the acceptor atom and the neighbors of this atom. In other words, since the diffuse functions are so spread in the space and hardly assigned to a given nuclei, the BSSE is not only energetically localized in the bonds but also in the same extent in non-bonded interactions. This effect is in agreement with the observations of corresponding isocontour density difference maps

II.1.4 Molecular Clusters



(A)

CYCLIC $(HF)_3$	**Uncorrected**	**VMFC-corrected**
$\varepsilon_{ind.r}^{(20)}$ (2-body)	-19.22	-10.23
$\varepsilon_{ind.r}^{(20)}$ (3-body)	-2.13	-1.52
$\varepsilon_{ind.r}^{(30)}$ (2-body)	-0.23	-0.17
$\varepsilon_{ind.r}^{(30)}$ (3-body)	-0.58	-0.37
$\varepsilon_{disp}^{(20)}$ (2-body)	-8.64	-5.82
$\varepsilon_{disp}^{(30)}$ (3-body)	0.04	0.02
$\varepsilon_{es}^{(10)}$	-31.79	-23.23
ε_{exch}^{HL}	37.50	20.12
ε_{exch}^{HL} (3-body)	-0.49	-0.20
ΔE^{HL}	*5.71*	-3.11
ΔE_{def}^{SCF}	-13.09	-7.66
ΔE_{int}^{SCF}	-7.38	-10.77
ΔE_{int}^{SCF} (3-body)	-4.92	-2.78
$\Delta E_{int}^{(2)}$	-5.46	-3.15
ΔE_{int}^{MP2}	-13.86	-14.54
ΔE_{rel}	0.93	0.53
ΔE_{int}^{MP2} (3-body)	-5.10	-2.91

(B)

(A) Geometrical parameters of the cyclic hydrogen fluoride oligomers

(B) SAPT results for the hydrogen fluoride trimer at the uncorrected and CP-corrected cluster geometries with the Sadlej's medium polarized basis set.

II.1.4.1 Gas-Phase hydrogen fluoride clusters $(HF)_n$

The counterpoise correction when there are more than two molecules interacting is not trivial from a conceptual point of view. It has already been pointed out that high order BSSE delocalizations arising from the dimers, trimers, tetramers, etc descriptions within the molecular aggregate may contribute to the BSSE in analogy to the non-additivity of the many-body interaction present in clusters.

In a recent paper, Mierzwicki and Latajka[84] analyzed the behavior of several counterpoise methods already proposed in the literature in the calculation of many-body interactions of $Li(NH_4)_n$ and $Li(NH_4)_n^+$ clusters at several levels of theory. Unfortunately, only stabilization energies were computed and not even nuclear relaxation effects were taken into account. The three counterpoise-correction schemes they use were the so called site-site, pairwise additive and hierarchical (Valiron-Mayer) function counterpoise; SSFC, PAFC and VMFC, respectively.

In the following we want to go one step forward. As commented throughout this work, in order to properly take into account the BSSE, the counterpoise-correction must be seen as a correction to be added to the molecular aggregate description. This allows us to compute not just interaction energies but also gradients and harmonic frequencies for the three different counterpoise schemes. Furthermore, the location of the stationary points on the BSSE-corrected PES is essential to obtain reliable counterpoise-corrected energies and avoid the sometimes referred as overcorrection when using small basis sets.

We want to assess the differences between the different CP methods in terms of molecular geometries, vibrational frequencies and stabilization energies. For the first time, we will perform full geometry optimizations using both the pairwise additive and the hierarchical counterpoise methods. The use of the hierarchical counterpoise scheme will tell us the effects of the high-order BSSE terms and to which extent they can be considered negligible. The validity of the pairwise additive scheme will also be critically analyzed.

We will apply these methods to the hydrogen fluoride trimer and tetramer. The

hydrogen fluoride clusters have received much attention in the last years. Last experimental[156,157] and theoretical[158-160] studies predict planar ring structures of C_{nh} symmetry for the $(HF)_n$ $3 \leq n < 6$ gas phase oligomers. In the case of the tetramer and pentamer, however, there is still some debate[161]. X-ray and neutron diffraction experiments[80] have shown that solid HF tends to form infinite zig-zag chains with very large cooperative effects. Therefore there must be an inversion of the relative stability of the cyclic and chain isomers as the molecular aggregate grows.

In this work, both the cyclic and chain-like arrangements are considered in order to compare the BSSE effect for two structures where the importance of the cooperative effects is very different. Also, the high symmetry of the cyclic aggregates will allow us to perform geometry optimizations with a relatively large basis set using the hierarchical counterpoise scheme even for the tetramer.

For all the geometry optimizations and frequency calculations, we have used our code to automatically generate all the necessary input files and repeatedly call a slightly modified Gaussian94[100] package. For this particular study, the program was updated in order to cope with the molecular symmetry and use the VMFC and PAFC methods. A detailed explanation of the capabilities of the code will be give in the next section.

As for the computational details, we have performed MP2 calculations (frozen core) with the 6-31G(d,p), 6-31++G(d,p) and a medium polarized Sadlej's basis set with a (10s,6p,4d,1f/6s,4p)→[5s,3p,2d,1f/3s,2p] contraction scheme. The aug-cc-pVTZ basis set has also been used in order to compare our results with some of the last uncorrected calculations.

---

[156] L. Andrews, V. E. Bondybey and J. H. English, J. Chem. Phys. 81, 3452 (1984)

[157] D. W. Michael and J. M. Lisy, J. Chem. Phys. 85, 2528 (1986)

[158] L. Rincón, R. Almeida, D. Garica-Aldea and H. Diez y Riega, J. Chem. Phys. 114 5552 (2001)

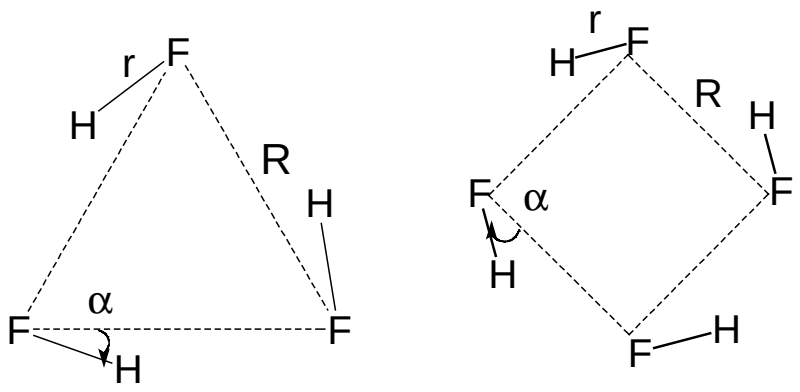
[159] K. R. Liedl, J. Chem. Phys. 108, 3199 (1998)

[160] a) G. S. Tschumper, Y. Yamaguchi and H. F. Schaeffer III, J. Chem. Phys. 106, 9627 (1997); b) T. A. Blake, S. W. Sharpe and S. S. Xantheas, J. Chem. Phys. 113, 707 (2000); c) B. L. Grigorenko, A. A. Moskovsky and A. V. Nemukhin, J. Chem. Phys. 111, 4442 (1999); d) W. Klopper , M. Quack and M. A. Suhm, Mol. Phys. 94 105 (1998).

[161] F. Huisken, M. Kaloudis, A. Kulcke, C. Laush and J. M. Lisy, J. Chem. Phys. 103, 5367 (1995)

II.1.4.1.a $(HF)_n$ cyclic

The results of the geometry optimizations of the HF cyclic trimer and tetramer are gathered in Table XLVIII. As shown in SCHEME 20, both trimer and tetramer structures are determined by three parameters: the intermolecular F-F (R_{F-F}) and the intramolecular F-H (R_{F-H}) distances and the angle between each HF unit and the center or the ring structure (α). We have studied only the C_{3h} and C_{4h} cluster arrangements for the trimer and the tetramer, respectively.



SCHEME 20: Geometrical parameters of the C_{3h} and C_{4h} cyclic hydrogen fluoride trimer and tetramer

As a general trend for the HF trimer results, it can be seen that intermolecular distance increases upon correction for BSSE, the larger differences being observed for the PAFC method. Upon CP-correction, the intramolecular F-H distance shortens by less than 0.01Å. However, this difference is still larger than the variation of this distance with respect to basis sets. The cyclic nature of the complexes prevents from large effect of BSSE correction on the angular parameter. In all cases, α increases (up to 3 degrees), inducing to a larger deviation from the triangular arrangement and hence a large intermolecular distance and cluster size.

The effect of adding diffuse functions dramatically decreases the effect of BSSE for the 6-31G(d, p) basis set. The differences between the uncorrected and CP-corrected intermolecular distances, larger than 0.1Å for the small basis set, decrease to ca. 0.05Å upon inclusion of diffuse functions. The medium polarized basis set, specifically design for the description of intermolecular interactions yields large BSSE. Indeed, both the uncorrected and the CP-corrected geometrical parameters are

close to the values obtained with the 6-31G(d, p) basis set.

TABLE XLVIII

Geometrical parameters (Å, degrees), total (a.u.) and stabilization energies (kcal/mol),and BSSE corrections (kcal/mol) for the cyclic HF trimer and tetramer in several basis sets and counterpoise methods. See SCHEME 20 for the definition of the geometrical parameters.

	R_{F-F}	R_{F-H}	α	Energy	BSSE	SP BSSE	ΔE
TRIMER							
				MP2/6-31G(d,p)			
Uncorr.	2.530	0.9432	20.6	-300.626538			-26.75
SSFC	2.651	0.9355	22.8	-300.608461	11.34	12.23	-15.40
PAFC	2.676	0.9345	23.6	-300.607189	12.14	13.47	-14.60
VMFC	2.666	0.9339	23.6	-300.607143	12.17	13.39	-14.57
				MP2/6-31++G(d,p)			
Uncorr.	2.649	0.9411	24.2	-300.672298			-15.40
SSFC	2.700	0.9391	24.5	-300.668919	2.12	2.21	-13.28
PAFC	2.708	0.9389	24.4	-300.668461	2.41	2.53	-12.99
VMFC	2.700	0.9390	24.7	-300.668901	2.13	2.23	-13.27
				MP2/Sadlej			
Uncorr.	2.521	0.9489	19.6	-300.843118			-23.87
SSFC	2.658	0.9423	22.2	-300.827375	9.88	10.94	-14.00
PAFC	2.680	0.9417	22.4	-300.825724	10.91	12.26	-12.96
VMFC	2.660	0.9422	22.2	-300.827249	9.94	11.03	-13.92
TETRAMER							
				MP2/6-31G(d,p)			
Uncorr.	2.471	0.9563	9.0	-400.848922			-44.16
SSFC	2.580	0.9440	11.1	-400.824190	15.51	16.70	-28.64
PAFC	2.607	0.9422	11.7	-400.822892	16.33	18.26	-27.83
VMFC(2)	2.604	0.9408	11.1	-400.821037	17.50	19.36	-26.66
VMFC(3)	2.611	0.9400	12.0	-400.82055	17.80	19.91	-26.36
				MP2/6-31++G(d,p)			
Uncorr.	2.563	0.9516	11.5	-400.908758			-28.29
SSFC	2.623	0.9467	11.9	-400.902411	3.98	4.22	-24.31
PAFC	2.635	0.9462	11.8	-400.901641	4.47	4.79	-23.82
VMFC(2)	2.621	0.9468	12.0	-400.902673	3.82	4.03	-24.47
VMFC(3)	2.621	0.9468	12.1	-400.902642	3.84	4.06	-24.45
				MP2/Sadlej			
Uncorr.	2.479	0.9586	8.5	-401.140773			-42.26
SSFC	2.579	0.9515	9.9	-401.114300	16.61	17.55	-25.65
PAFC	2.604	0.9499	10.3	-401.110746	18.48	20.15	-23.42
VMFC(2)	2.583	0.9511	10.0	-401.114102	16.73	17.73	-25.52
VMFC(3)	2.583	0.9511	10.0	-401.114097	16.73	17.73	-25.52

Adding another HF unit to the complex results in a shortening of the

intermolecular distance by ca. 0.06Å. The intramolecular H-F distance decreases whereas the angular parameter slightly increases. Cooperative effects are also evident from an energetical point of view. The stabilization energy per hydrogen bond increases by more than 1kcal/mol, hence yielding an extra stabilization energy of ca. 6Kcal/mol for the tetramer

The same effect of the BSSE on the geometrical parameters is observed for the tetrameric structure. However, even though the BSSE correction to the stabilization energy increases respect to the trimer complex, the differences between the uncorrected and the CP-corrected geometrical parameters are comparable to the ones observed for the trimeric cluster.

As for the energies, the BSSE correction to the trimer and tetramer energies is always overestimated at the uncorrected geometry. The CP-corrected stabilization energies computed at the uncorrected minima (single-point counterpoise calculation) are smaller than the ones computed at the respective CP-corrected stationary point. The differences in the case of the trimer range from ca. 0.1Kcal/mol for the 6-31++G(d, p) basis set to more than 1kcal/mol for both the 6-31G(d, p) and the Sadlej basis set. For the tetramer these differences are twice as large. It is important to note that, after CP-correction, the basis set dependence of both the calculated stabilization energies and geometrical parameters decreases. The uncorrected stabilization energies obtained with the 6-31G(d, p) and the Sadlej basis set are far too large. All the CP-corrected values are within 3 and 5Kcal/mol for the trimer and the tetramer, respectively. The same situation has been observed in previous studies of weakly bonded complexes.

On the other hand, it can be seen that the differences between the SSFC and the VMFC corrected values are still appreciable for the smallest basis set used, the 6-31G(d, p). The intermolecular distance is again the most sensitive geometrical parameter to the BSSE. The inclusion of high-order terms in the CP method leads to larger intermolecular distances, the differences being of 0.015Å and 0.024Å for the trimer and tetramer, respectively. The effect on the cluster energy is much more relevant. The SSFC method overestimates the stabilization energy by ca. 1 Kcal/mol and 2 Kcal/mol, for the trimer and tetramer, respectively. In the later case, the inclusion of the third order CP-correction terms, VMFC(3), whose determination is feasible in this case due to the high symmetry of the complex, shows no significant

effect on neither the geometry nor the energy of the complex, provided that the basis set used is flexible enough. Only for the 6-31G(d, p) basis set the intermolecular distance still increases by 0.007Å and decreases the stabilization energy by 0.3 kcal/mol.

Fortunately, for the other basis sets including diffuse functions, the high-order correction terms of the VMFC method have practically no effect on both geometrical parameters and energies. It is worth to point out two facts. First, for the 6-31++G(d, p) tetramer calculation, the inclusion of high-order CP-correction terms induces to a smaller BSSE. The intermolecular distance slightly shortens and the stabilization energy increases by 0.16Kcal/mol upon correction. Even though it is a rather unexpected result, we would like to emphasize that the high-order terms in the VMFC method can be actually of opposite sign, in other words, the BSSE can be overestimated if one is restricted to the SSFC method. The fact that the dimer correction term is negative does not mean that the dimer description is better with the DCBS than with the TCBS (from an energetical point of view). Instead, it is the dimer stabilization energy, which is larger (more negative). The reason why this happens is because the lowering of the monomers energies are larger than the dimer counterparts as the basis set increases. Secondly, it is remarkable that, even though the large BSSE exhibit at the monomer level by the Sadlej basis set, the effect of high-order BSSE correction is rather irrelevant. It is confirmed again that a basis set can not be considered *bad* or *unbalanced* just because it bears large BSSE. Indeed, we will show that the Sadlej basis set provides very accurate results, provided that the BSSE is properly taken into account.

Finally, the PAFC method clearly overestimates the BSSE in all the cases except for the calculations involving the 6-31G(d, p) basis set. In this case, the PAFC results show that the method seems to mimic the effect of the high-order CP-correction terms. However, the differences observed in both the geometrical parameters and cluster energies when using more suitable basis sets make this method not advisable.

TABLE XLIX

Harmonic frequencies (cm^{-1}), ZPVE correction (Kcal/mol) to the stabilization energy and frequency shifts (cm^{-1}) for the cyclic HF trimer. Values in parenthesis correspond to the uncorrected harmonic frequencies and ΔZPVE computed at the CP-corrected geometry. Red shifts calculated respect the monomer H-F stretching frequencies obtained at each level of theory (4193.4, 4118.7 and 4082.4 cm^{-1} for the 6-31G(d, p), 6-31++G(d, p) and Sadlej basis sets, respectively)

	Uncorrected	**SSFC**	**PAFC**	**VMFC**
		MP2/6-31G(d,p)		
ω_6 (E')	259.6	203.1 (193.6)	189.1 (180.8)	194.7 (185.0)
ω_3 (A')	276.5	224.7 (202.2)	217.9 (186.9)	219.7 (193.1)
ω_8 (E'')	539.6	505.8 (474.5)	489.5 (467.8)	476.7 (456.6)
ω_5 (E')	755.2	631.6 (607.1)	619.4 (577.2)	589.4 (571.1)
ω_7 (A'')	765.9	711.6 (669.0)	689.4 (651.4)	688.9 (644.2)
ω_2 (A')	1153.4	988.3 (963.6)	954.0 (919.1)	948.3 (917.8)
ω_1 (A')	3668.4	3863.5 (3869.1)	3892.6 (3899.0)	3905.1 (3905.0)
ω_4 (E')	3841.4	3967.3 (3972.4)	3985.6 (3990.7)	3995.7 (3998.6)
ΔZPVE	5.84	5.48 (5.19)	5.36 (5.03)	5.30 (5.03)
Shift	-352	-226	-188	-198
		MP2/6-31++G(d,p)		
ω_6 (E')	184.0	172.6 (160.4)	169.1 (157.3)	171.3 (160.0)
ω_3 (A')	207.4	197.3 (181.3)	194.1 (177.6)	196.0 (181.0)
ω_8 (E'')	465.0	446.4 (432.1)	438.6 (427.7)	444.8 (432.5)
ω_5 (E')	553.2	531.2 (516.4)	524.6 (514.2)	531.2 (512.9)
ω_7 (A'')	680.7	654.6 (648.3)	642.6 (645.8)	648.0 (646.0)
ω_2 (A')	936.5	884.0 (897.2)	884.1 (895.2)	876.0 (890.8)
ω_1 (A')	3785.8	3835.8 (3833.5)	3839.3 (3837.2)	3838.7 (3836.6)
ω_4 (E')	3878.6	3915.2 (3913.2)	3917.6 (3916.0)	3917.3 (3915.1)
ΔZPVE	4.89	4.78 (4.64)	4.72 (4.61)	4.76 (4.63)
Shift	-240	-204	-202	-202
		MP2/Sadlej		
ω_6 (E')	259.7	186.2 (169.8)	174.3 (157.1)	185.5 (168.6)
ω_3 (A')	285.7	208.7 (181.1)	200.8 (166.1)	208.3 (179.8)
ω_8 (E'')	579.2	468.0 (532.4)	443.4 (524.9)	468.8 (530.8)
ω_5 (E')	752.4	575.0 (582.9)	557.6 (564.4)	573.2 (580.9)
ω_7 (A'')	771.7	654.0 (689.1)	625.7 (678.9)	654.8 (687.7)
ω_2 (A')	1166.0	936.4 (940.8)	912.3 (916.4)	933.1 (938.8)
ω_1 (A')	3533.0	3730.0 (3723.1)	3751.3 (3741.8)	3732.3 (3725.5)
ω_4 (E')	3711.4	3832.3 (3823.7)	3846.0 (3835.0)	3833.8 (3825.5)
ΔZPVE	5.72	4.71 (4.85)	4.54 (4.72)	4.70 (4.84)
Shift	-371	-250	-236	-249

It has been shown how the BSSE and the several CP methods affect the PES in terms of the location of the stationary points. Obviously, the harmonic frequencies on

each PES will differ too. To which extent depends on two main factors. First, the geometrical parameters of the stationary point are different, so that the frequencies will vary, depending upon how large was the CP-correction on the geometry. And second, the higher-order derivatives of the BSSE-correction term are non-zero, so that the CP-corrected second derivatives will differ from the uncorrected supermolecule derivatives. By comparing the corrected and uncorrected frequencies at the CP-corrected stationary point one can determine to which extent the rather expensive calculation of the CP-corrected second order derivatives are necessary. The differences between the frequencies properly computed on the corresponding uncorrected and CP-corrected PES contain both the *geometrical* and the *differential* factor.

The uncorrected and CP-corrected harmonic frequencies for the cyclic HF trimer calculated for the three basis sets and CP methods are gathered in Table XLIX. In all the cases, the uncorrected low-frequencies are overestimated whereas the frequencies of the two stretching modes are underestimated respect to the CP-corrected ones. For the 6-31G(d, p) and Sadlej basis sets the differences between the uncorrected and the CP-corrected frequencies range from 60cm^{-1} for the lowest frequency to more than 200cm^{-1} for the frequencies labeled ω_1 and ω_2. In general, the BSSE modifies the low frequencies by 10-25%. The differences in the fundamental stretching frequency ω_4 are >100cm^{-1}. As expected, the 6-31++G(d, p) frequencies aren't modified in a significant manner. The maximum differences are ca. 50cm^{-1}, even for the most sensitive frequencies ω_1 and ω_2.

The inclusion second-order CP-correction seems not to induce any appreciable changes in the frequencies. Only for the smallest basis set the frequencies are further shifted by up to 7% respect to the SSFC values. The PAFC frequencies are very similar to both the SSFC or VMFC values.

The differences between the uncorrected and CP-corrected frequencies at a given geometry are much smaller than the disagreement between the uncorrected and CP-corrected values calculated on the stationary point of their corresponding PES. In other words, out of the two factors explained above, the "geometrical" is clearly the most important. It is worth to note that, in general, the addition of the second derivatives of the BSSE term to the supermolecule derivatives shifts the frequencies in the opposite direction. For instance, the uncorrected 6-31G(d, p) low frequencies

decrease when computed on the CP-corrected PES minima, but tend to increase when computed from the CP-corrected second derivatives. However, the opposite trend is observed for the Sadlej basis set.

The effect on the ZPVE correction can be explained also in the basis of this opposite effect. For the small basis sets, the ZPVE decreases when computed at the CP-corrected PES but then increases after inclusion of BSSE-term's second derivatives. The opposite occurs for the Sadlej basis set, which shows the largest effect in the ZPVE correction. In this case, the uncorrected ZPVE correction of 5.72 kcal/mol decreases to 4.84 when computed on the VMFC PES and further reduces to 4.54Kcal/mol.

In most of the last papers devoted to the study of the cyclic HF trimer, the BSSE correction is not considered at all, and even criticized[159]. For instance, Tschumper et al.[160a] carried out high level CCSD(T) *ab initio* calculations of the cyclic trimer and compare them with the results obtained by using a empirically refined SC-2.9+HF3BG potential from Quack and Suhm[162]. Their binding energies, ZPVE correction and red shifts of the HF stretching frequency were systematically larger than the ones obtained with the accurate potential. In our opinion this is probably due because the authors neglect the effect of the BSSE. Our best CP-corrected results obtained with the Sadlej basis set provide lower binding energies but closer to the ones obtained with the SC-2.9+HF3BG potential after inclusion of the ZPVE correction. Furthermore, the predicted redshift of the HF stretching frequency ($249cm^{-1}$) agrees with the experimental ($249cm^{-1}$) and the one obtained with the potential ($250cm^{-1}$) of Quack and Suhm. Even though this perfect agreement can be considered fortuitous, since our reported redshift does not include anharmonicity corrections, the BSSE correction seems essential for this basis set in order to obtain accurate results.

Recently, Liedl[159] studied the concerted hydrogen exchange process of the HF trimer at the MP2/aug-cc-pVXZ, X=2,4. He found that the uncorrected energies for the C_{3h} and D_{3h} structures were less dependent on the basis set than the counterpoise-corrected ones. He claimed the counterpoise-corrected results were useless and particularly the CP-corrected interaction energies, that is without the monomer

[162] M. A. Quack, M. A. Suhm, (unpublished)

relaxation contribution. Actually, his results show that the uncorrected MP2/aug-cc-pVDZ energies are closer to the ones obtained with the larger basis set, but there is no monotonic trend in the uncorrected values to properly extrapolate to the basis set limit. For both the minimum and the TS, the MP2/aug-cc-pVTZ values are higher than the MP2/aug-cc-pVDZ but lower than the MP2/aug-cc-pVQZ ones. His estimated CBS values are, therefore, completely arbitrary. It is true that the CP-corrected MP2/aug-cc-pVDZ interaction energies are too low compared with the respective values obtained with larger basis sets, but in this case the energies monotonically decrease and a proper extrapolation can be carried out.

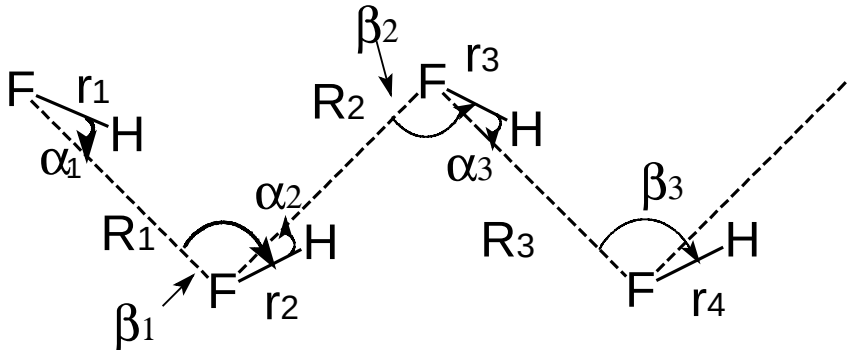
Even though the CP-corrected values might be further from the extrapolated or experimental results than the respective uncorrected ones, they still represent a better description of the properties of the system. For instance, reported barriers for the hydrogen exchange for the aug-cc-pVXZ, X=2,4 series are 20.17, 17.89 and 18.61kcal/mol, 23.83, 20.48 and 20.21kcal/mol for the uncorrected and the CP-corrected values, respectively. The best estimate of this barrier using explicitly correlated coupled cluster calculations is 20.33 kcal/mol. It is evident that the CP-correction is absolutely necessary to obtain accurate CBS results for this process.

Another effect that was not taken into account was the use of CP-corrected optimizations, that is, only single-point BSSE corrections were performed. Even though it has been shown that the effect on the geometries can be very important, we have observed that these differences are much lower for aug-cc-pVDZ basis set than for other basis sets of comparable size (see Section II.1.1.1.d). In this particular case, the stabilization energy for the C_{3h} structure computed on the CP-corrected stationary point is -13.30kcal/mol, only 0.16kcal/mol lower than the single-point CP-corrected one.

Finally, the usefulness of the CP-corrected interaction energy without fragment relaxation can not be questioned by direct comparison with the uncorrected, fully optimized, stabilization energies. This is obvious for the D_{3h} structure where the fragment relaxation energies (~60kcal/mol) are thirty times larger than the stabilization energy itself.

II.1.4.1.b $(HF)_n$ linear

The zig-zag linear arrangement of both the trimer and tetramer HF clusters have been studied also at the same level of theory. The results obtained for the uncorrected and the CP-corrected geometry optimizations are shown in Table L and LI for the trimer and tetramer, respectively. The defined geometrical parameters are depicted in the SCHEME 21. In the later case, only the SSFC method was used for the corrected optimization, since including the second-order CP or full VMFC corrections involves 33 and 65 gradient calculations, respectively.



SCHEME 21: Geometrical parameters of the zig-zag linear hydrogen fluoride trimer and tetramer

Similar trends to the cyclic clusters are observed. Upon CP correction, intermolecular distances increase while intramolecular HF bonds shorten, leading to a weaker interaction. The differences between the uncorrected and CP-corrected parameters are of the same order that for the cyclic complexes, except for the intermolecular bond angles. In this case, the 6-31G(d, p) basis set poorly describes the anisotropy of the interaction. Intermolecular bond angles α_1 and α_2 are overestimated whereas β_1 and β_2 are clearly underestimated by up to 20 degrees. This is not surprising since the uncorrected geometry optimization at this level of theory for the hydrogen fluoride dimer leads to a spurious cyclic structure. The corresponding CP-corrected optimizations, however, overcome this problem. Indeed, the CP-corrected angular parameters are in good agreement with the ones obtained using more flexible basis sets. Again, the effect of BSSE is minimized by adding diffuse functions. Discrepancies between uncorrected and CP-corrected values are less than 1 degree for the angles and 0.06Å for intermolecular distances.

TABLE L

Geometrical parameters (Å, degrees), total (a.u.) and stabilization energies (kcal/mol),and BSSE corrections (kcal/mol) (single-point BSSE in parenthesis) for the linear HF trimer in several basis sets and counterpoise methods. See SCHEME 21 for the definition of the geometrical parameters.

	Uncorrected	**SSFC**	**PAFC**	**VMFC**
		MP2/6-31G(d,p)		
r_1	0.9293	0.9277	0.9277	0.9275
r_2	0.9356	0.9310	0.9311	0.9303
r_3	0.9266	0.9246	0.9246	0.9244
R_1	2.647	2.752	2.754	2.755
R_2	2.630	2.722	2.727	2.733
α_1	8.98	2.87	2.94	3.05
β_1	98.80	118.7	118.8	119.2
α_2	5.24	3.22	2.85	3.48
β_2	108.6	121.2	121.7	121.0
Energy	-300.6104703	-300.60154618	-300.6015036	-300.6011282
ΔE	-16.66	-11.06	-11.04	-10.80
BSSE		5.60 (6.41)	5.63 (6.47)	5.86 (6.71)
		MP2/6-31G++(d,p)		
r_1	0.9342	0.9331	0.9331	0.9331
r_2	0.9377	0.9361	0.9360	0.9361
r_3	0.9303	0.9298	0.9298	0.9299
R_1	2.728	2.789	2.789	2.787
R_2	2.700	2.762	2.764	2.760
α_1	3.16	3.16	3.28	3.20
β_1	119.31	119.1	119.1	119.4
α_2	3.65	4.05	3.80	3.92
β_2	121.2	121.0	121.5	121.9
Energy	-300.6658512	-300.66329844	-300.6632270	-300.66335745
ΔE	-11.35	-9.75	-9.71	-9.79
BSSE		1.60 (1.70)	1.65 (1.75)	1.56 (1.66)
		MP2/Sadlej		
r_1	0.9361	0.9346	0.9344	0.9346
r_2	0.9398	0.9378	0.9374	0.9379
r_3	0.9312	0.9306	0.9305	0.9306
R_1	2.616	2.749	2.759	2.749
R_2	2.596	2.722	2.734	2.722
α_1	2.62	2.90	2.90	2.91
β_1	110.8	114.7	115.9	114.7
α_2	2.72	3.40	4.45	3.39
β_2	114.7	117.0	117.2	117.2
Energy	-300.8323145	-300.8205844	-300.8201539	-300.8205622
ΔE	-17.10	-9.73	-9.46	-9.72
BSSE		7.36 (7.93)	7.63 (8.32)	7.37 (7.95)

The Sadlej basis set bears the largest BSSE. However, whereas the intermolecular distances are underestimated by more than 0.12Å without CP-correction, the angular features of the complex are well described even at the uncorrected level.

Regarding the CP-corrected stabilization energies, the linear structures are about 4kcal/mol and 8Kcal/mol less stable compared to the cyclic trimer and tetramer, respectively. The BSSE correction in the trimer ranges from 1.6 and 2.76kcal/mol for the 6-31G(d, p) basis set to 7.95 and 12.31kcal/mol for the Sadlej basis set for the trimer and the tetramer complexes, respectively. The dependence of the BSSE correction to the energy on the geometry seems to be less important than for the cyclic case. However, the BSSE can still be overestimated by up to 1kcal/mol (depending on the basis set) when computed the uncorrected geometry. The cooperative effects are obviously less important than in the cyclic structures. However, the inclusion another HF unit to the linear trimer still enhances the stabilization energy per hydrogen bond by ca. 0.5Kcal/mol

The performance of the two first-order CP methods, i.e. SSFC and PAFC, is similar. Both methods modify the values of the geometrical parameters in the same direction, even though the PAFC method seems to overcorrect the BSSE when using larger basis sets.

Our results show again that the differences between the SSFC and VMFC approaches are only appreciable when using small basis sets. The inclusion of second-order CP terms for the 6-31G(d, p) basis set increases the intermolecular distances and angles by up to 0.01Å and 0.5 degrees, whereas the stabilization energy is lowered by 0.26Kcal/mol. The effect of this term in the other basis sets is negligible. Nevertheless, it is worth to note that in the case of the 6-31++G(d, p) basis set, analogously to the situation in the cyclic structure, the sign of the correction is the opposite, i.e. the stabilization energy increases upon correction.

The inclusion of high-order CP-correction terms was computationally feasible only for the trimer, therefore only the SSFC method was used to compute the CP-corrected geometry of the linear tetramer. Nevertheless, we performed a single-point second-order CP-correction at the SSFC corrected geometry with the 6-31G(d, p) basis set, in order to estimate the effect of BSSE on the stabilization energy. The BSSE increased by 0.64Kcal/mol upon correction. The final stabilization energy of

17.59 Kcal/mol is closer to the ones obtained with larger basis sets.

TABLE LI

Uncorrected and SSFC-corrected Geometrical parameters (Å, degrees), total (a.u.) and stabilization energies (kcal/mol),and BSSE corrections (kcal/mol) (single-point BSSE in parenthesis) for the linear HF tetramer in several basis sets. See SCHEME 21 for the definition of the geometrical parameters.

	MP2/6-31G(d,p)	**MP2/6-31++G(d,p)**	**MP2/Sadlej**
	UNCORRECTED		
r_1	0.9307	0.9355	0.9376
r_2	0.9400	0.9416	0.9442
r_3	0.9379	0.9398	0.9422
r_4	0.9271	0.9308	0.9318
R_1	2.643	2.706	2.602
R_2	2.575	2.639	2.550
R_3	2.607	2.674	2.575
α_1	6.85	2.40	2.17
β_1	100.8	119.3	110.7
α_2	2.15	1.16	1.22
β_2	111.1	122.1	115.0
α_3	4.04	2.72	2.31
β_3	110.5	122.4	115.6
Energy	-400.82141532	-400.8932367	-401.1172646
ΔE	-26.89	-18.55	-27.50
	SSFC		
r_1	0.9287	0.9341	0.9358
r_2	0.9344	0.9391	0.9413
r_3	0.9331	0.9378	0.9399
r_4	0.9251	0.9303	0.9311
R_1	2.731	2.767	2.727
R_2	2.660	2.699	2.662
R_3	2.695	2.735	2.695
α_1	2.2	2.44	2.13
β_1	118.5	118.9	115.0
α_2	1.0	1.37	1.11
β_2	121.8	121.9	118.3
α_3	2.4	3.03	2.50
β_3	122.1	122.3	118.3
Energy	-400.80761515	-400.8890763	-401.098932
ΔE	-18.23	-15.94	-16.00
BSSE	8.66 (9.69)	2.61 (2.76)	11.50 (12.31)

As pointed out recently by Rincón et al.[158], the open chain structures for the HF trimer and tetramer are first-order saddle points connecting two equivalent cyclic configurations. Our results are consistent with their findings in all the cases. The CP-

correction does not change the topology of the PES in any case, not even for the 6-31G(d, p) basis set, where the effect on the geometry was very large.

In short, it has been shown that the effect of the second and third-order basis set extensions is negligible for these systems, although they are still considerable when using small basis sets. The use of diffuse functions is essential for the saturation of the high-order BSSE.

The PAFC results obtained with small basis set tend to the VMFC values. However, with larger basis sets it clearly overestimates the BSSE and large differences with respect to the other counterpoise methods are observed. Therefore, in our opinion it does not represent valid correction scheme.

The medium polarized basis set used bears large BSSE, but the results obtained are very accurate compared to other empirical potential or high level *ab initio* calculations, provided that the CP-corrected PES is used. For this basis set, the high-order BSSE effects are almost negligible even though the first-order BSSE correction is ca. the 70-80% of the stabilization energy.

II.1.4.1.c Symmetry-Adapted Perturbation Theory results.

The differences between the geometries of corrected and uncorrected surfaces can be further emphasized by dissecting the interaction energy terms into perturbative components obtained from the symmetry-adapted perturbation theory (SAPT). These corrections are free from BSSE and, as shown by Cybulski and Chalasinski, their sum converges asymptotically to the BSSE-corrected supermolecular terms. The supermolecular SCF interaction energy, ΔE_{int}^{SCF}, may be divided into the Heitler-London interaction energy, ΔE^{HL} and the SCF deformation part, $\Delta E_{\mathrm{def}}^{SCF}$. The former originates from the unperturbed monomer wavefunctions and can be further divided into its electrostatic and exchange components (see Table LII). The latter involves effects of electronic polarization, which is restrained by the exchange effects. Its exchangeless contributions can be represented by the sum of the SAPT induction corrections $\varepsilon_{ind,r}^{(n0)}$. Out of these terms only the electrostatic term is additive; the remaining ones contribute three-body components.

TABLE LII.

Description of SAPT corrections $\varepsilon^{(ij)}$ (where i and j correspond to the interaction and the intramonomer correlation operators, respectively) which are implicitly present in the two- and three-body supermolecular Møller-Plesset interaction energy terms ΔE_{int} at the SCF level and in the second order.

Supermolecular Møller-Plesset	SAPT	Physical interpretation
		Two-body
ΔE_{int}^{SCF}	$\varepsilon_{es}^{(10)}$	Electrostatic energy between SCF monomers
	ε_{exch}^{HL}	Heitler-London exchange effect between SCF monomers
	ΔE_{def}^{SCF}	SCF-deformation energy =Induction effect restrained by exchange (includes $\varepsilon_{ind,r}^{(20)}$, $\varepsilon_{ind,r}^{(30)}$ etc.)
	$\varepsilon_{ind,r}^{(20)}$	2nd-order induction energy with response effects
	$\varepsilon_{ind,r}^{(30)}$	3rd-order induction energy with response effects
$\Delta E_{int}^{(2)}$	$\varepsilon_{disp}^{(20)}$	2nd-order dispersion energy between SCF monomers.
		Three-body
ΔE_{int}^{SCF}	ε_{exch}^{HL}	Heitler-London exchange nonadditivity
	ΔE_{def}^{SCF}	SCF-deformation nonadditivity (includes $\varepsilon_{ind,r}^{(20)}$, $\varepsilon_{ind,r}^{(30)}$, etc.)
	$\varepsilon_{ind,r}^{(20)}$	2nd-order induction nonadditivity with response effects
	$\varepsilon_{ind,r}^{(30)}$	3rd-order induction nonadditivity with response effects
$\Delta E_{int}^{(2)}$		
$\Delta E_{int}^{(3)}$	$\varepsilon_{disp}^{(30)}$	3rd-order dispersion nonadditivity

In the second-order of Møller-Plesset theory, one of the dominating SAPT terms is the second order dispersion term $\varepsilon_{disp}^{(20)}$. This term is additive and thus only contribute the two-body components. The first dispersion nonadditivity appears in the third order of perturbation theory as the $\varepsilon_{disp}^{(30)}$ term. This term usually dominates the nonadditivity of the $\Delta E_{int}^{(3)}$ supermolecular Møller-Plesset term. The physical sense of the SAPT corrections considered in this work, and their correspondence to the

supermolecular Møller-Plesset terms are summarized in Table LVII.

TABLE LIII

Energetic SAPT contributions (kcal/mol) to the interaction energy for the cyclic and linear HF trimer with the 6-31G(d, p) basis set at the uncorrected and VMFC-corrected geometries. The difference between the corrected and uncorrected values and its percentage with respect to the total corrected stabilization energy is also reported.

CYCLIC	**6-31G(d, p) unc**	**6-31G(d, p)**	**Δcorr.-unc.**	**%total**
$\varepsilon_{ind,r}^{(20)}$ (2-body)	-17.55	-9.08	8.47	58.1
$\varepsilon_{ind,r}^{(20)}$ (3-body)	-1.87	-1.35	0.52	3.5
$\varepsilon_{ind,r}^{(30)}$ (2-body)	-0.29	-0.17	0.12	0.8
$\varepsilon_{ind,r}^{(30)}$ (3-body)	-0.62	-0.35	0.26	1.8
$\varepsilon_{disp}^{(20)}$ (2-body)	-5.84	-3.57	2.27	15.5
$\varepsilon_{disp}^{(30)}$ (3-body)	0.00	0.00	0.00	0.0
$\varepsilon_{es}^{(10)}$	-35.86	-26.47	9.39	64.4
ε_{exch}^{HL}	36.55	19.40	-17.15	117.7
ε_{exch}^{HL} (3-body)	-0.51	-0.22	0.30	2.0
ΔE^{HL}	0.69	-7.07	-7.76	53.2
ΔE_{def}^{SCF}	-11.22	-5.89	5.33	36.5
ΔE_{int}^{SCF}	-10.53	-12.96	-2.44	16.7
ΔE_{int}^{SCF} (3-body)	-4.59	-2.48	2.10	14.4
$\Delta E_{int}^{(2)}$	-2.83	-1.61	1.23	8.4
ΔE_{int}^{MP2}	-15.49	-15.70	0.21	1.4
ΔE_{rel}	0.97	0.41	-0.56	3.8
ΔE_{int}^{MP2} (3-body)	-4.86	-2.63	2.23	15.3
LINEAR	**6-31G(d, p)unc**	**6-31G(d, p)**	**Δcorr.-unc.**	**%total**
$\varepsilon_{ind,r}^{(20)}$ (2-body)	-8.38	-5.32	3.06	28.4
$\varepsilon_{ind,r}^{(20)}$ (3-body)	-0.31	-0.38	-0.07	0.6
$\varepsilon_{ind,r}^{(30)}$ (2-body)	-0.13	-0.10	0.04	0.3
$\varepsilon_{ind,r}^{(30)}$ (3-body)	-0.16	-0.13	0.03	0.3
$\varepsilon_{disp}^{(20)}$ (2-body)	-3.03	-2.18	0.85	7.9
$\varepsilon_{disp}^{(30)}$ (3-body)	0.00	0.00	0.00	0.0
$\varepsilon_{es}^{(10)}$	-20.58	-17.10	3.48	32.2
ε_{exch}^{HL}	17.49	10.95	-6.54	60.5
ε_{exch}^{HL} (3-body)	-0.16	-0.10	0.06	0.6
ΔE^{HL}	-3.09	-6.15	-3.05	28.3
ΔE_{def}^{SCF}	-5.62	-3.99	1.63	15.1
ΔE_{int}^{SCF}	-8.71	-10.14	-1.43	13.2
ΔE_{int}^{SCF} (3-body)	-1.09	-0.88	0.21	1.93
$\Delta E_{int}^{(2)}$	-1.24	-0.66	0.58	5.3
ΔE_{int}^{MP2}	-10.47	-11.14	0.67	6.2
ΔE_{rel}	0.22	0.08	-0.14	1.3
ΔE_{int}^{MP2} (3-body)	-1.14	-0.91	0.23	2.13

The SAPT contributions were calculated for the cyclic and linear trimer at three basis sets. The calculations were carried out at uncorrected and VMFC-corrected minima of the trimers and the results are shown in Tables LIII-LV. As a trimer of highly polar molecules, $(HF)_3$ is dominated, at the level of two-body interactions, by the electrostatic attraction. These effects are counterbalanced, to a certain degree, by the repulsive exchange effects. The two-body induction effects are also important in view of the fact that HF has a considerable dipole moment and an appreciable polarizability. The two-body dispersion interaction is the third in importance. At the level of three-body interactions the bulk of nonadditive interaction originates from the SCF-deformation term. The exchange nonadditivity is quite small while the three-body dispersion is nearly zero. The cyclic configuration is stabilized over the linear one at the level of two body interactions, because of more favorable electrostatic and induction effects in the cyclic arrangement. The three-body terms also favor the cyclic structure.

TABLE LIV

Energetic SAPT contributions (kcal/mol) to the interaction energy for the cyclic and linear HF trimer with the 6-31++G(d, p) basis set at the uncorrected and VMFC-corrected geometries.

CYCLIC	**6-31++G(d,p)$_{unc}$**	**6-31++G(d, p)**	**Δcorr.-unc.**	**%total**
$\varepsilon_{ind,r}^{(20)}$ (2-body)	-9.69	-7.78	1.91	14.3
$\varepsilon_{ind,r}^{(20)}$ (3-body)	-1.38	-1.22	0.15	1.1
$\varepsilon_{ind,r}^{(30)}$ (2-body)	-0.15	-0.13	0.02	0.1
$\varepsilon_{ind,r}^{(30)}$ (3-body)	-0.35	-0.29	0.06	0.4
$\varepsilon_{disp}^{(20)}$ (2-body)	-4.07	-3.47	0.60	4.5
$\varepsilon_{disp}^{(30)}$ (3-body)	0.00	0.00	0.00	0.0
$\varepsilon_{es}^{(10)}$	-24.97	-22.60	2.37	17.8
ε_{exch}^{HL}	20.02	16.07	-3.96	29.8
ε_{exch}^{HL} (3-body)	-0.23	-0.17	0.06	0.4
ΔE^{HL}	-4.95	-6.54	-1.59	11.9
ΔE_{def}^{SCF}	-6.82	-5.50	1.32	9.9
ΔE_{int}^{SCF}	-11.77	-12.04	-0.27	2.0
ΔE_{int}^{SCF} (3-body)	-2.55	-2.10	0.45	3.3
$\Delta E_{int}^{(2)}$	-1.40	-1.23	0.17	1.2
ΔE_{int}^{MP2}	-13.61	-13.59	0.02	0.2
ΔE_{rel}	0.42	0.32	-0.10	0.8
ΔE_{int}^{MP2} (3-body)	-2.66	-2.20	0.47	3.5

TABLE LIV (cont.)

Energetic SAPT contributions (kcal/mol) to the interaction energy for the cyclic and linear HF trimer with the 6-31++G(d, p) basis set at the uncorrected and VMFC-corrected geometries.

LINEAR	**6-31++G(d,p)$_{unc}$**	**6-31++G(d, p)**	**Δcorr.-unc.**	**%total**
$\varepsilon_{ind,r}^{(20)}$ (2-body)	-6.27	-4.96	1.31	13.4
$\varepsilon_{ind,r}^{(20)}$ (3-body)	-0.36	-0.34	0.02	0.2
$\varepsilon_{ind,r}^{(30)}$ (2-body)	-0.10	-0.09	0.01	0.1
$\varepsilon_{ind,r}^{(30)}$ (3-body)	-0.15	-0.12	0.03	0.3
$\varepsilon_{disp}^{(20)}$ (2-body)	-2.62	-2.21	0.41	4.2
$\varepsilon_{disp}^{(30)}$ (3-body)	0.00	0.00	0.00	0.0
$\varepsilon_{es}^{(10)}$	-16.25	-14.79	1.47	15.0
ε_{exch}^{HL}	11.93	9.32	-2.61	26.7
ε_{exch}^{HL} (3-body)	-0.13	-0.10	0.03	0.3
ΔE^{HL}	-4.32	-5.47	-1.15	11.7
ΔE_{def}^{SCF}	-4.81	-3.91	0.90	9.2
ΔE_{int}^{SCF}	-9.13	-9.37	-0.25	2.5
ΔE_{int}^{SCF} (3-body)	-0.95	-0.78	0.16	1.6
$\Delta E_{int}^{(2)}$	-0.57	-0.42	0.16	1.6
ΔE_{int}^{MP2}	-9.79	-9.84	-0.05	0.5
ΔE_{rel}	0.13	0.09	-0.04	0.4
ΔE_{int}^{MP2} (3-body)	-0.98	-0.81	0.17	1.72

The calculations carried out in three different basis sets give insights into the basis-set dependence of the SAPT terms. In the 6-31G(d,p) basis set the electrostatic effects are overestimated and the two-body dispersion contributions are underestimated. The 6-31++G(d,p) leads to the improvement in the electrostatic interaction. Only in the Sadlej basis set, however, all terms, including dispersion energy, become correctly represented.

Calculations of SAPT terms at two different geometries, one uncorrected and a counterpoise-corrected (at the VMFC level) are displayed in Tables LIII-LV. In the 6-31G(d,p) basis set the evaluation of SAPT terms at the uncorrected minimum geometry leads to large discrepancies in SAPT terms resulting in considerable per cent errors in the electrostatic, exchange, and induction terms. For example, in the cyclic configuration, the errors in these terms range from 50% to over 100%. These errors are reduced in the 6-31++G(d,p) basis set to 15-30%.

TABLE LV

Energetic SAPT contributions (kcal/mol) for the cyclic and linear HF trimer with the Sadlej basis set

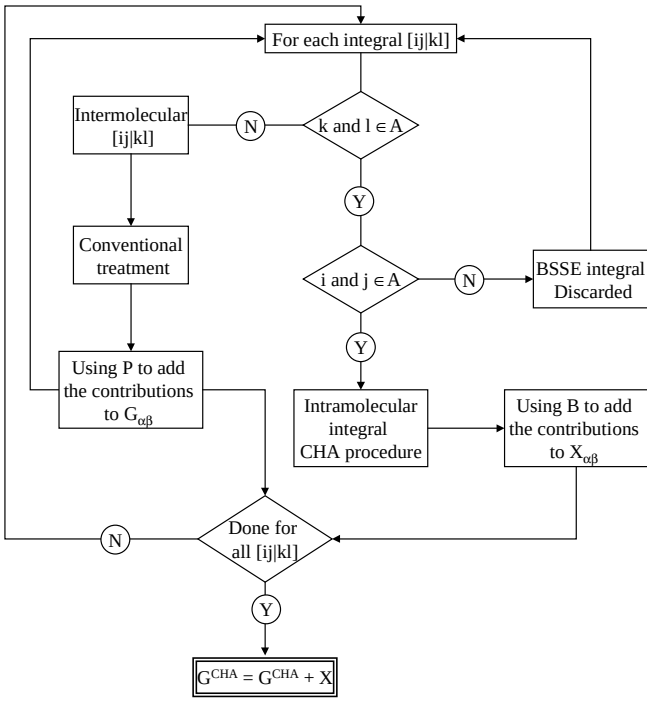
CYCLIC	**Sadlej $_{unc.}$**	**Sadlej $_{corr.}$**	**Δcorr.-unc.**	**%total**
$\varepsilon_{ind,r}^{(20)}$ (2-body)	-19.22	-10.23	9.00	64.6
$\varepsilon_{ind,r}^{(20)}$ (3-body)	-2.13	-1.52	0.61	4.3
$\varepsilon_{ind,r}^{(30)}$ (2-body)	-0.23	-0.17	0.06	0.4
$\varepsilon_{ind,r}^{(30)}$ (3-body)	-0.58	-0.37	0.20	1.4
$\varepsilon_{disp}^{(20)}$ (2-body)	-8.64	-5.82	2.82	20.2
$\varepsilon_{disp}^{(30)}$ (3-body)	0.04	0.02	-0.02	0.1
$\varepsilon_{es}^{(10)}$	-31.79	-23.23	8.56	61.4
ε_{exch}^{HL}	37.50	20.12	-17.37	124.8
ε_{exch}^{HL} (3-body)	-0.49	-0.20	0.29	2.0
ΔE^{HL}	*5.71*	-3.11	-8.82	63.3
ΔE_{def}^{SCF}	-13.09	-7.66	5.44	39.0
ΔE_{int}^{SCF}	-7.38	-10.77	-3.38	24.3
ΔE_{int}^{SCF} (3-body)	-4.92	-2.78	2.14	15.3
$\Delta E_{int}^{(2)}$	-5.46	-3.15	2.30	16.5
ΔE_{int}^{MP2}	-13.86	-14.54	-0.68	4.9
ΔE_{rel}	0.93	0.53	-0.40	2.9
ΔE_{int}^{MP2} (3-body)	-5.10	-2.91	2.19	15.7
LINEAR	Sadlej$_{unc}$	Sadlej$_{corr.}$	Δcorr.-unc.	%total
$\varepsilon_{ind,r}^{(20)}$ (2-body)	-10.03	-5.98	4.05	41.7
$\varepsilon_{ind,r}^{(20)}$ (3-body)	-0.36	-0.33	0.02	0.2
$\varepsilon_{ind,r}^{(30)}$ (2-body)	-0.13	-0.11	0.02	0.2
$\varepsilon_{ind,r}^{(30)}$ (3-body)	-0.22	-0.15	0.07	0.7
$\varepsilon_{disp}^{(20)}$ (2-body)	-4.76	-3.42	1.35	13.8
$\varepsilon_{disp}^{(30)}$ (3-body)	0.00	0.00	0.00	0.0
$\varepsilon_{es}^{(10)}$	-18.70	-14.68	4.02	41.4
ε_{exch}^{HL}	18.99	11.04	-7.95	81.8
ε_{exch}^{HL} (3-body)	-0.19	-0.11	0.08	0.8
ΔE^{HL}	0.30	-3.64	-3.93	40.5
ΔE_{def}^{SCF}	-8.96	-4.78	4.18	43.0
ΔE_{int}^{SCF}	-8.67	-8.41	0.25	2.6
ΔE_{int}^{SCF} (3-body)	-1.27	-0.86	0.41	4.2
$\Delta E_{int}^{(2)}$	-0.48	-1.31	-0.82	8.4
ΔE_{int}^{MP2}	-9.34	-9.84	-0.50	5.1
ΔE_{rel}	0.17	0.10	-0.07	0.7
ΔE_{int}^{MP2} (3-body)	-1.30	-0.89	0.41	4.3

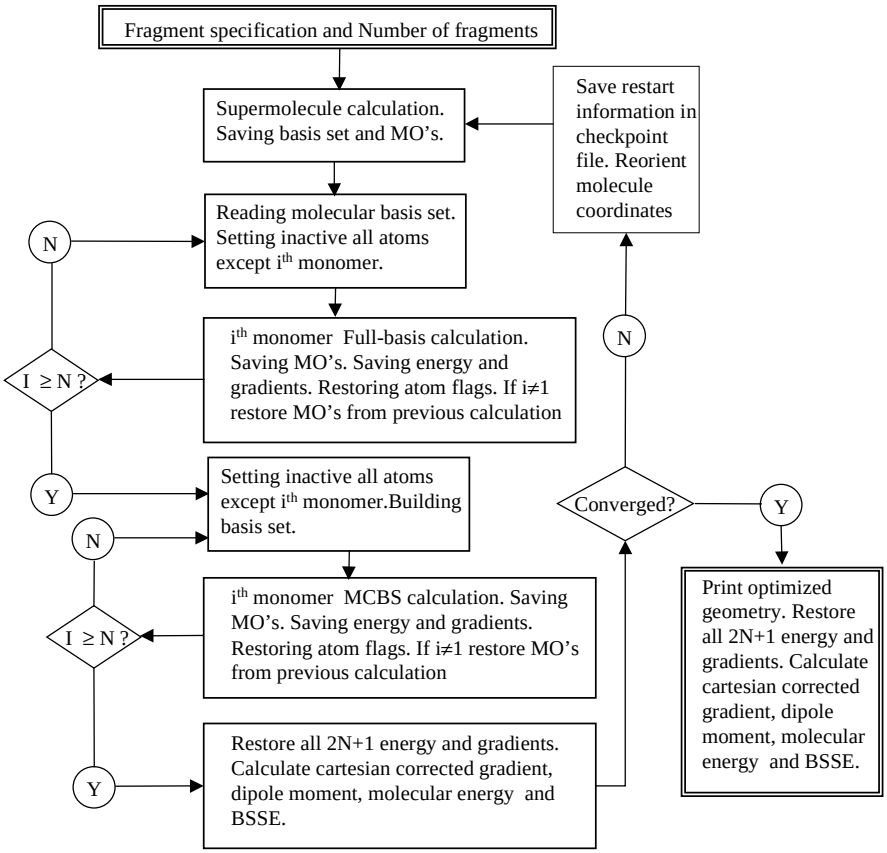
The results obtained in Sadlej basis set also indicate large discrepancies in SAPT terms derived using these two geometries. Although, this basis set produces

reliable values of SAPT terms, it also generates large values of BSSE which results in large distortions of geometrical parameters. It is worth noting that because of this difference in geometry it is possible to obtain a false picture of the interaction energy composition if the calculations are done for the uncorrected minimum. For example, in Sadlej basis set the Heitler-London interaction energy, ΔE^{HL} is negative (-3.11 kcal/mol) in the VMFC-corrected minimum whereas in the uncorrected minimum it has a repulsive (5.71 kcal/mol) value for the cyclic configuration. A similar sign reversal of ΔE^{HL} also takes place in the linear configuration.

In short, the presented SAPT results show that already at the level of two-body interactions the cyclic configuration is stabilized over the linear one. The three-body terms also favor the cyclic structure. The large differences between the results obtained at the uncorrected and counterpoise-corrected geometries underscore the need for performing the analysis of the interaction energy at the counterpoise-corrected minimum geometries.

II.2 Methodologic Developments





II.2.1 The application of the counterpoise method in order to obtain corrected molecular energies, gradients and vibrational frequencies.

II.2.1.1 Implementation into Gaussian98 Rev A10 (and further).

The counterpoise-procedure has been implemented into the Gaussian 98 package, in order to automatically carry out single-point, geometry optimizations, vibrational frequency and even third order derivatives (were available) calculations at any level of theory including the counterpoise correction. This work has been carried out in collaboration with Prof. J. J. Dannenberg from the City University of New York, and Dr. Mike Frisch, President of Gaussian Inc.

The main goal was to provide the scientific community with an easy and simple tool to perform CP-corrected geometry optimizations, where the BSSE may cause major errors when unbalanced basis sets are used. The first operative version of the *cp* program was already given to several research groups that were interested in taking BSSE into account not only as a single-point correction. This program was designed to drive Gaussian in order to automatically perform counterpoise-corrected single-point, geometry optimizations and vibrational frequency calculations. However, due to some restrictions in Gaussian and also some bugs on the code (already fixed in the Rev A10), some Gaussian links had to be modified and recompiled in order to properly run the program. The reduced number of geometry optimization methods and options prevented the use of the program for location of transitions states or carry out Intrinsic Reaction Coordinate (IRC) calculations. Due to the increasing interest of the intermolecular interactions research groups on a efficient tools for taking BSSE into account, the generalized counterpoise method was implemented into Gaussian in such a way that most of the options of the program can be applied in conjunction with the counterpoise calculation.

Briefly, a new link (l122) on Overlay 1 was created. It is used to control the type of calculation (single-point, gradient, second derivative), the number of fragments and which particular fragment (or supermolecule) calculation is going to be computed next. The only keyword necessary in the Route section is

counterpoise=N

where N holds for the total (integer) number of fragments composing the complex. Its value is limited to fifty fragments.

Obviously, one must assign each atom of the complex to the corresponding single fragment. Any dummy atom should not be assigned to any fragment. It is possible also for one (or more) atoms not to be assigned to any fragment. In this way, bond functions can be defined as single, dummy, fragments with no electrons assigned to. Proton transfer reactions can eventually be studied also by assigning the transferred proton to a dummy fragment without electrons, as shown in Section II.1.1.4.

Therefore, in order to specify the fragments, the input file also requires an additional integer to be placed at the end of each atom specification indicating which fragment/monomer it is part of.

Here are examples using a Z-matrix (left) and Cartesian coordinates (right):

```
# MP2/6-31G Counterpoise=2 Opt

Counterpoise with Z-matrix

0,1,0,3,1,2
O,0.0,0.0,0.0,1
O,1,ROO,2
H,1,RO1H,3,HOX3,2,90.,0,1
H,1,RO1H,3,HOX3,2,-90.,0,1
X,2,1.,1,52.5,3,180.,0
H,2,RO2H1,6,H7OX,1,180.,0,2
H,2,RO2H2,6,H8OX,1,0.,0,2
```

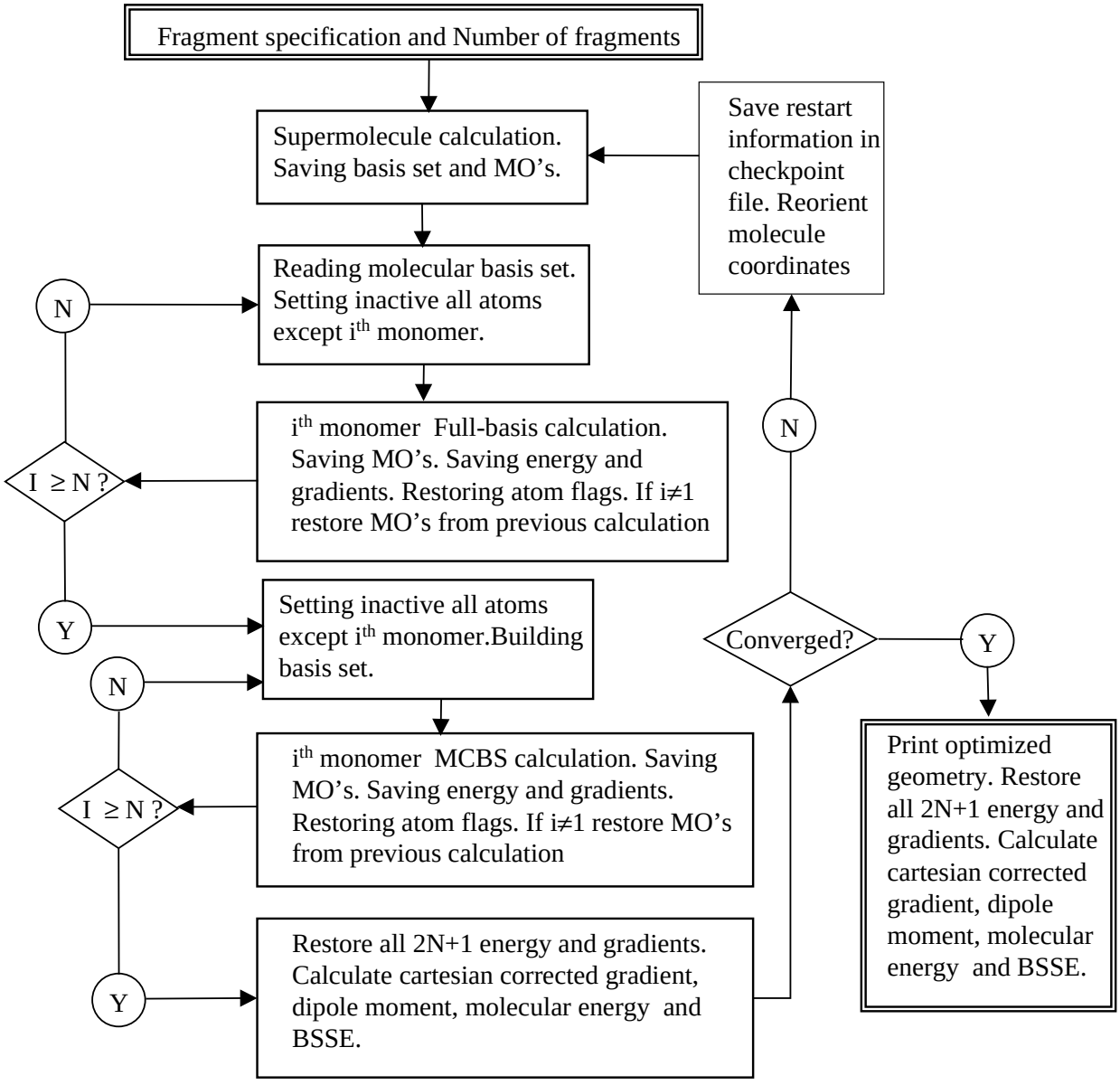
```
# MP2/6-31G Counterpoise=2 Opt

Counterpoise with Cartesians

0,1
1  0.00  0.00  0.92  1
9  0.17  0.00  2.73  2
9  0.00  0.00  0.00  1
```

Note that the Z-matrix input requires a 0 after the dihedral angle value/variable (to indicate that the final angle is a dihedral) prior to the fragment number. Also, the first atom in the Z-matrix must be given in Cartesian coordinates. Clearly, using Cartesian coordinates for such jobs makes specifying fragment numbers in the input much more straightforward. The preceding Z-matrix also illustrates the use of fragment-specific charge and spin multiplicity specifications. The format of the corresponding input line in this case is:

molecule-charge, molecule-spin multiplicity, frag.1-charge, frag.1 multiplicity, frag. 2 charge, frag. 2 multiplicity,....

Note that if all the fragments present the same charge and multiplicity than that of the molecular complex, the specification of all the monomer charge and multiplicities is no longer necessary.

Practically, a gradient optimization is carried out as follows:



The current implementation presents some limitations:

a) No ECP pseudopotentials or general basis set can be used.
b) If an atom is not assigned to any fragment, only cartesian coordinates optimization can be performed.
c) In general, the use of molecular symmetry messes up the calculation, unless the symmetry of the monomers is equal or higher of the whole molecule.
d) The molecular symmetry can not be used to save monomer calculations in symmetric clusters.
e) No particular options for each of the 2N+1 can be given, unless specified using a non-standard route.

With the last implementation of the *CP* program, all these drawbacks are overcome, as will be discussed in the next section. The molecular symmetry can be handled and used to save fragment calculations. Moreover, different function counterpoise are implemented for the study of molecular clusters. The structure of the program also allows for the definition of user-defined BSSE corrections by linear combination of fragment calculations.

Only using the *CP* program, the analysis of BSSE in hydrogen fluoride clusters presented in the Section II.1.4 could be performed.

Next, a short manual of the *CP* program is attached.

II.2.1.2 Manual of the CP program.

CP-DRIVER MANUAL

Pedro Salvador
Dept. of Chemistry, Oakland University, Rochester, Michigan, US
Institute of Computational Chemistry, University of Girona, Spain

CONTENTS:

1-SETTING UP

The CP program automatically calls Gaussian94/98 in order to perform either energy, gradient optimizations or frequency corrected for Basis Set Superposition Error. Three methods can be used automatically to correct for BSSE. Moreover, the user can define which fragment calculations are necessary in each case in terms of the fragment symmetry of the supermolecule or any further approximation.

In order to carry on geometry optimizations with high accuracy, link l103 of Gaussian must be slightly modified to output (to the file fort.99) the gradient vector with full precision.
The modification consists in adding the following lines in subroutine *GrdOpt*,

```
Open(unit=99,status='unknown',form='unformatted')
Rewind 99
Write(99) Nvar
Write(99) (aprint(2,iii),iii=1,Nvar)
Close (99)
```

after the line:

```
515     Continue
```

A bug in link l301 (*Massge* subroutine) must be fixed too. The massage of an atom located in the z-matrix just before a dummy atom induces to an error that stops the gradient calculation. To fix the bug substitute the following lines

```
Do 210 J = 1, NZ
If(IAnZ(J).ge.0) I = I + 1
        If(I.eq.ICntr) IAnZ(J) = IAn(ICntr)
 210   Continue
```

by

```
Do 210 J = 1, NZ
                If(IAnZ(J).ge.0) I = I + 1
If(I.eq.ICntr) then
                  IAnZ(J) = IAn(ICntr)
                  goto 213
                endif
 210   Continue
Continue
```

(Note that in Gaussian98 label 210 is changed to 20)

To create the new l103.exe and l301.exe files, first rename the original l103.exe and l301.exe , use the update utility of Gaussian from the $g9*root/g9* directory,

```
mg l103.exe
mg l301.exe
```

then move the new l103.exe and l301.exe files to the desired directory and finally restore the original l103.exe and l301.exe files.
The new links must be used for the CP-corrected geometry optimizations, hence, before the route card of the Gaussian calculation the following lines must be included:

```
% subst l301 /home/usr/CP
% subst l103 /home/usr/CP
```

where */home/usr/C*P is, in this case, the full path of the directory where the l301.exe and l103.exe files are placed

2-COUNTERPOISE-CORRECTION METHODS

The following counterpoise methods have been implemented in this version:

Site-Site CP method:
The BSSE is defined as the difference between the energy (gradient) of each fragment with its own basis set and the basis set of the whole molecule, and summed up to the uncorrected supermolecule respective values.

Pairwise Aditive CP method:
In this case, the BSSE is defined as the difference, for all pairwise interactions, between the energy (gradient) of each fragment with its own basis set and the basis set of each dimer in the supermolecule geometry. Again, the resulting energy(gradient) is summed up to the uncorrected supermolecule respective values.

Hierarchical CP method:
The BSSE is split in one ,two-body, three-body etc… contributions to the supermolecule energy. The number of calculations increases dramatically with the number of fragments. Note that 17 calculations are needed for a three-body calculation.

3-FILES

a) cp.f -> Source code of the driver program.

Input file: cp.met

Output files: cp.out, cp.res (info to restart calculation), cpfreq.out (only for frequency calculations)

Scratch files: *.in, cp.geo, cp.com, cp.fchk, cpfreq.fchk (only for frequency calculations)

Use Fortran77 compiler to create the executable, e.g. cp.x

b) hess.f -> Program to yield inverse hessian from a Gaussian output file. The matrix is created in file HESS and can be read by the driver program.

Usage : hess.x < gaussian_file

Use Fortran77 compiler to create the executable, e.g. hess.x

c) glocal.com -> Unix script to invoke Gaussian and formchk Gaussian utility. For instance:

```
#!/bin/csh
setenv g98root /usr/soft
 source $g98root/g98/bsd/g98.login
 setenv GAUSS_SCRDIR  /scratch
 g98 < $1.in  > cp.qfi
 formchk $1.chk  cp.fchk
 rm cp.qfi
```

Note that formchk output MUST be cp.fchk whereas Gaussian output for $1.in calculation can be overwritten.

d) freqlocal.com -> Unix script to invoke Freqchk Gaussian utility. For instance:

```
#!/bin/csh
setenv g98root /usr/soft
 source $g98root/g98/bsd/g98.login
freqchk cpfreq.chk <fc.inp> cpfreq.out
```

Also, input file for freqchk MUST be named cpfreq.chk. File fc.inp is used to include the desired options for the freqchk utility (temperature, pressure, etc..). For instance

N	*(write Hyperchem files?)*
0	*(Temperature (K)? [0=>298.15])*
0	*(Pressure (Atm)? [0=>1 atm])*
1	*(Scale factor for frequencies during thermochemistry? [0=>1/1.12])*
Y	*(use default isotopes?)*

The cpfreq.fchk Gaussian formatted checkpoint file is obtained from the respective supermolecule calculation checkpoint file (cp.fchk), where ONLY the second derivative matrix has been substituted by the CP-corrected one. Hence, one can further use the freqchk utility to extract the uncorrected frequencies from the cp.fchk file (at the given geometry).

Files cp.x, glocal.com and freqlocal.com must be in the same directory and this directory should be included in the user's PATH, in order to allow user to run the calculations in another directory.

When executing the driver program several gaussian input files (*.in) will be created, as well as a the file cp.geo, containing the current geometry.
The cp.res file is either read upon request in cp.met file or created in order to save information to resume the calculation. Unix script cp.com is also created for each calculation and called by the previously set glocal.com script.

Note that only one calculation can be performed correctly in a given directory. Use different directories for simultaneous calculations.

4-INPUT FILE DESCRIPTION

File consists in several sections, each starting with the following set of characters:
NAME_OF_SECTION #.
Each section will be discussed separately

A-# METODE # section:

Main section of the input file. Includes several fields in the form:
NAME_OF_FIELD= integer (I) or real (R) numbers
Each field represents either a parameter, like the number of fragments or number of variables, or a flag to control other options of the program, spefified in the next sections.

-*Input parameters*
NFRAG= I, being I the number of fragments of the calculation
NVAR= I, being I the number of variables for the calculation. (up to 100)

-*Options for the geometry optimization:*
TOL= R, being R the threshold of the RMS of the gradient for the geometry optimization.
MAXIT= I, being I the maximum number of iterations for the geometry optimization
NDIIS= I, being I the maximum dimension of the DIIS subspace for the geometry optimization (up to 10)
HESS= 1 if file HESS is to be used as starting hessian for the optimization, 0 if no initial hessian is given (default)
UPDATE= 1 to activate the BFGS method to update the hessian for the optimization, 0 otherwise (default)
HESS= 1 if file HESS is to be used as starting hessian for the optimization, 0 if no initial hessian is given (default)

-*Options for the CP method and fragment definition*
BSSE = 0 for the SSCP method, 1 for the hierarchical NBODY method, or 2 for the PACP method, default 0.
SYM= 1 to use symmetry, included in # SYMMETRY # section (see below), 0 otherwise, default 0
RESTART= 1 to use cp.res file data to resume the calculation, 0 otherwise (cp.res file will be overwritten), default 0

SPIN= 0 to use supermolecule charge and multiplicity for all fragments, 1 to use specific values for each fragment calculation, either in # FRAG # or # SYMMETRY # sections (see below). If BSSE=1 then specific charge and multiplicity must be introduced in # SYMMETRY # section.

-Known problems and restrictions:

Using the hierarchical counterpoise method, the number of fragment calculations needed increase enormously with the number of fragments. For this reason, automatic calculations of either energy or geometry optimizations are not allowed if NFRAG > 3 and BSSE=1. However, in this case, if symmetry between the fragments exists, the calculation can be performed by including in the # SYMMETRY # section the necessary fragment calculations and their respective scale factors for both energy and gradient.

B-# GEOM # section:

Enter Z-matrix or Cartesian coordinates in Gaussian format. Dummy atoms can be used, however BQ should be avoided. For instance,

 # GEOM ####
X
F1 1 r1
F2 1 r1 2 120.
F3 1 r1 3 120. 2 180.
H1 2 rh1 1 alpha 4 180.
H2 3 rh1 1 alpha 2 180.
H3 4 rh1 1 alpha 3 180.

C-# FRAGS # section:

Free format input. Fist line must include the number of atoms included in each fragment. The next N lines include, for each fragment, the number of the atoms (as specified in the z-matrix or cartesian coordinates in # GEOM # section) belonging to it.

Finally, if SPIN=1 then two extra lines are needed. First one including the charge and the second one the multiplicity of each fragment. Note that, in this case, if BSSE=1

the charge and multiplicity must be included instead in the # SYMMETRY # section.

D-# VARIABLES # section:

Free format. Enter initial values for the variables specified in the z-matrix or cartesian coordinates. Do not use "=" between variable name and its value

E-# VARTYP # section:

Free format. Single line including, for each variable 1 if it represents a distance or 2 if is an angle.

F-# GAUSSIAN # section:

Enter route card and charge and multiplicity for Gaussian calculation. Nosymm keyword must be included. If force keyword is present CP-geometry optimization is performed. If freq is present CP-frequency calculation is carried out (see below). If none of the above keywords are present then a single-point CP-energy is performed

If a general basis is used for the calculation, the basis set specification MUST be included in file cp.gbs but not in this section. For instance,

 # GAUSSIAN ###
%mem=100Mb
%subst l103 /usr/users/pedros/bin
 # mp2/6-31++G(d,p) 6d nosymm force

MP2 HF-trimer CP-corrected optimization

0 1

-Restrictions and problems:

Input is CASE SENTITIVE. Please, use lower case for force, freq and gen Gaussian keywords.

-Frequency calculations:

No symmetry can be used. Note that the second derivatives for the fragment calculations are obtained using the following Gaussian keywords:

opt=CalcFC optcyc=1

Hence, most probably, an error message will appear at the end of each Gaussian calculation, which, however, will not affect the CP-corrected frequency calculation.

G-# SYMETRY # section

Must be included if SYM=1 on # METODE # section.

First, enter number of fragment calculations will be performed. Then enter, separated by commas, name of calculation to be performed, scale factor for the gradient and energies. Charge and multiplicity for each calculation must also be included if BSSE=1 and SPIN=0 in # METODE # section.

The name of each calculation is set in this way:

(fragments included) _ (basis set used)

For instance, the following line

12_123 , -2.0 , 0 ,1

specifies that fragments 1 and 2 (as defined in # FRAGS # section) using basis set of fragments 1, 2 and 3 are included in the calculation. Also, both energy and gradient will be scaled by -2.0 and summed up to the corresponding uncorrected supermolecule energy and gradient respectively. Charge=0 and multiplicity=1 will be used.

-Restrictions and problems:

The uncorrected supermolecule calculation must be included in the last line.

eg. for N fragments:

12...N_12...N, 1.0, 0 , 1

H- Ending card:

Line including "---" set of characters indicates the end of the input file.

5-GEOMETRY OPTIMIZATIONS

Geometry optimization on the corresponding CP-corrected PES can be carried out automatically by using the *force* keyword in the *#GAUSSIAN#* section of the input file. The optimization starts at the given geometry and in the next iteration, a new guess of the geometry is computed from the linear combination of the fragmetn calculation gradients and it is written in the *cp.geo* file. For every iteration, the energies and gradients of each fragment calculation are printed, as well as the RMS of the gradient. When this parameter reaches a value below the introduced tolerance (see *#METODE#* section in the input file) or the total number of iterations is exceeded the program stops.

The final geometry can be extracted both from the *cp.geo* file or from the *cp.out* output file. In the later, the gradients of each calculation and the RMS is also given.

By choosing the different options in the first section of the input file, many different optimization procedures can be selected.(see below)

Method	NDIIS	HESS	UPDATE
DIIS	N	0/1	0/1
Steepest-descent	1	0	0
BFGS	1	0/1	1
DIIS-VM	N	0/1	0/1

If HESS is set to 1 , HESS file containing initial inverse hessian must be on current directory. This file can be generated using the *hess.f* program, that extracts from a Gaussian output the first second derivative matrix found. In order to generate the HESS file simply use

hess.x< gaussian_output

Note that is the uncorrected hessian at the given geometry.

6- FREQUENCY CALCULATIONS

Frequency calculations using the corresponding CP-corrected second derivatives can be carried out automatically by using the *freq* keyword in the *#GAUSSIAN#* section of the input file.

The computations is carried out in the following way: for each fragment calculation, the second derivative matrix in cartesian coordinates is extracted form the corresponding Gaussian Formatted CheckPoint file. The CP-corrected second derivatives are computed as linear combination of the necessary fragment calculations. Then, the Formatted CheckPoint file of the supermolecule calculation, *cp.fchk* file, is duplicated and the uncorrected second derivative matrix is substituted by the CP-corrected one, generating the *cpfreq.fchk* file. Finally, the *freqchk* Gaussian keyword is used automatically to extract form the properly modified *cpfreq.fchk* file the vibrational frequencies and thermal analysis on a separated output file, in this case the *cpfreq.out*. The parameters for the thermal analysis, namely pressure, temperature and isotops can be easily selected in the *fc.inp* file (see FILES section of this manual)

7-SAMPLE INPUT FILES

A few examples of the applicability of the program are described below. Some comment lines have been introduced in parenthesis.

EXAMPLE 1: HF trimer SCF/SSCP geometry optimization using symmetry:

EXAMPLE 2: HF trimer SCF/PACP single-point energy calculation:

EXAMPLE 3:HF trimer MP2/ Hierarchical CP-corrected frequencies:

EXAMPLE 1

HF trimer SCF/SSCP geometry optimization using symmetry:

```
# METODE ##
NVAR=3 NFRAG=3 NDIIS=5  MAXIT=9 UPDATE=1 HESS=1  SPIN=0
BSSE=0 RESTART=0 SYM=1 TOL=0.0001
# VARIABLES #####
rh1       0.941
r1        1.550
alpha             52.3
# VARTYP ####
1 1 2                    (Type of variable, in the same order as in previous section)
# FRAGS #####
2 2 2
1 4
2 5
3 6
0 0 0                              (charge optional as SPIN=0)
1 1 1                              (multiplicity optional as SPIN=0)
# GEOM ####
X                                (dummy atom not considered in previous section)
F1 1 r1
F2 1 r1  2 120.
F3 1 r1 3 120. 2 180.
H1 2  rh1 1  alpha 4 180.
H2 3  rh1 1 alpha 2 180.
H3 4 rh1  1 alpha 3 180.
# GAUSSIAN ###
%subst l103 /usr/users/pedros/bin
%subst l301 /usr/users/pedros/bin
# MP2/6-31G(d,p) nosymm force                (CP-geometry optimization)

MP2 calculation of HF trimer using symmetry

0 1
# SYMETRY ##
3                                (three calculations are needed)
1_1 , 3.0                (HF basis calculation, equiv. to the other HF's in molecule)
1_123 ,-3.0              (HF full basis calc., equiv. to the other HF's in molecule)
123_123   ,1.0           (Trimer calculation introduced last)

---                              (end of file line)
```

EXAMPLE 2:

HF trimer SCF/PACP single-point energy calculation:

```
# METODE ##
NVAR=3 NFRAG=3 NDIIS=5  MAXIT=9 UPDATE=1 HESS=1  SPIN=0  TOL=0.0001
BSSE=2  SYM=0                          (any field  not present set to default value)

# VARIABLES #####
rh1       0.941
r1        1.550
alpha             52.3
# VARTYP ####
1 1 2                    (Type of variable, in the same order as in previous section)
# FRAGS #####
2 2 2
1 4
2 5
3 6
# GEOM ####
X                                (dummy atom not considered in previous section)
F1 1 r1
F2 1 r1  2 120.
F3 1 r1 3 120. 2 180.
H1 2  rh1 1  alpha 4 180.
H2 3  rh1 1 alpha 2 180.
H3 4 rh1  1 alpha 3 180.
# GAUSSIAN ###
%subst l301 /usr/users/pedros/bin
%subst l103 /usr/users/pedros/bin
# SCF/6-31G(d,p) nosymm 6d             (CP-corrected energy calculation)
gfinput

HF calculation of HF trimer

0 1

---                              (end of file line)
```

EXAMPLE 3:

HF trimer MP2/ Hierarchical CP-corrected frequencies:

```
# METODE ##
NVAR=3 NFRAG=3 NDIIS=5  MAXIT=9 UPDATE=1 HESS=1  SPIN=0
BSSE=1 RESTART=0 SYM=0
# VARIABLES #####
rh1       0.941
r1        1.550
alpha             52.3
# VARTYP ####
1 1 2                     (Type of variable, in the same order as in previous section)
# FRAGS #####
2 2 2
1 4
2 5
3 6
# GEOM ####
X                                 (dummy atom not considered in previous section)
F1 1 r1
F2 1 r1  2 120.
F3 1 r1 3 120. 2 180.
H1 2  rh1 1  alpha 4 180.
H2 3  rh1 1 alpha 2 180.
H3 4 rh1  1 alpha 3 180.
# GAUSSIAN ###
%subst l103 /usr/users/pedros/bin
%subst l301 /usr/users/pedros/bin
# MP2/6-31G(d,p) nosymm freq          (CP-frequency calculation)

MP2 calculation of HF trimer

0 1
# SYMETRY ##                  (No symmetry can be used for frequencies)

---                               (end of file line)
```

II.2.2 Implementation of the CHA method

The CHA versions of the Restricted (Closed-shell) and Unrestricted Hartree-Fock equations, CHA-RHF and CHA-UHF, respectively, have been implemented using the FORTRAN 77 language in collaboration with Prof. I. Mayer, during a three-month stay at the Hungarian Academy of Sciences in Budapest.

The main goal was to design efficient CHA-RHF and CHA-UHF algorithms for the general case of N interacting fragments. In a next step, the CHA canonical orbitals would be further used for the calculation of the corresponding CHA versions of the second-order Møller-Plesset energy, CHA-MP2 and CHA-UMP2, respectively. Finally, the so-called *intramolecular* formulation of the CHA equations for dealing with the intramolecular BSSE problem was also explored. In this particular case, the unsatisfactory preliminary results prevented for further investigation on this direction.

The CHA-SCF equations represent, thus, a preliminary but necessary step for the implementation of the CHA-MP2 methodology. Furthermore, prior to the codification of the CHA equations, it is desirable to implement the conventional restricted closed-shell and unrestricted Hartree-Fock equations and apply the necessary modifications introduced by the CHA equations on these programs.

The RHF, UHF, CHA-RHF and CHA-UHF equations have been coded as external programs to the HONDO package[163]. The one and two-electron integrals are extracted from a previous HONDO run and are utilized to reproduce the Hartree-Fock orbitals and energy with full precision.

As commented in the introductory overview of the CHA methodology, the key point at the SCF level of theory is the construction of the CHA Fockian. Once the CHA Fockian is gained, the protocol to obtain the molecular orbitals and total energy is well established (see SCHEME 2). The details of the implementation of the designed algorithm for the efficient construction of the one- and two-electron part of the CHA Fockian, in the restricted and unrestricted versions, are given in this Section.

---

[163] 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

II.2.2.1 RHF and CHA-RHF

Let us assume first a typical restricted closed-shell problem of molecular system composed of N electrons and M nuclei (see Section I.2.2 for a brief overview of the SCF method). The matrix representation of the Fock operator in the atomic orbital basis can be written as

$$F_{\mu\nu} = H_{\mu\nu}^{core} + G_{\mu\nu} \quad ;\forall\, \mu,\nu \tag{II.1}$$

were

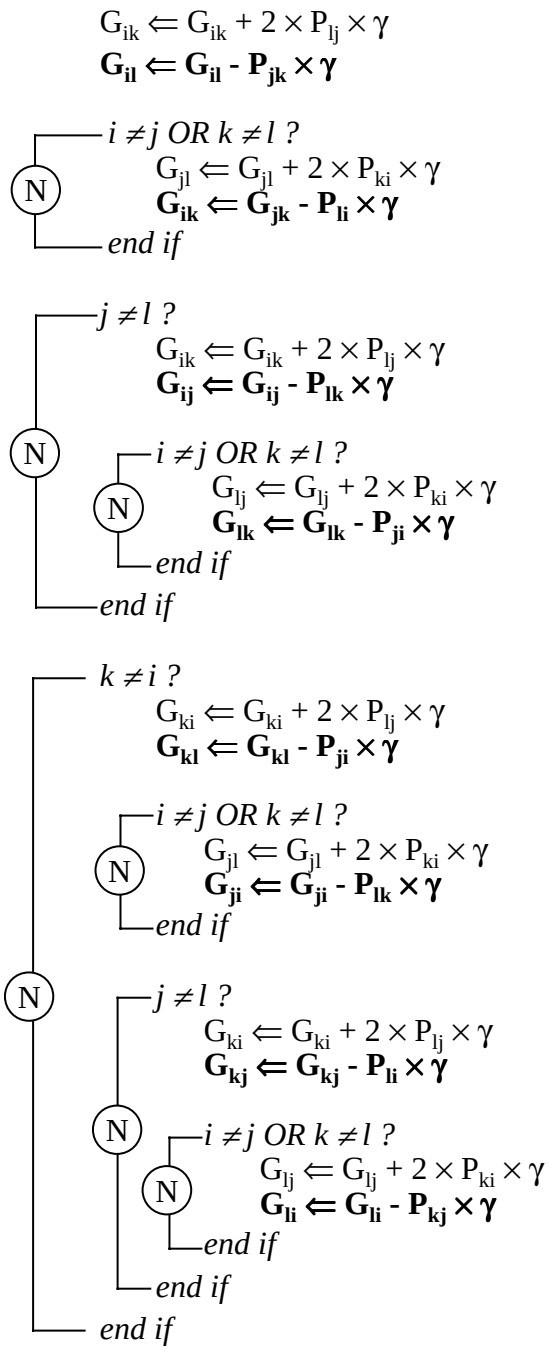
$$H_{\mu\nu}^{core} = T_{\mu\nu} + V_{\mu\nu} = <\varphi_\mu\,|-\frac{1}{2}\nabla^2\,|\,\varphi_\nu> + <\varphi_\mu\,|-\sum_{\alpha}^{M}\frac{Z_\alpha}{|r-R_\alpha|}\,|\,\varphi_\nu> \tag{II.2}$$

and

$$G_{\mu\nu} = \sum_{\kappa,\lambda} P_{\lambda\kappa}\left(2\left[\varphi_\mu\varphi_\kappa\,|\,\varphi_\nu\varphi_\lambda\right] - \left[\varphi_\mu\varphi_\nu\,|\,\varphi_\kappa\varphi_\lambda\right]\right) \tag{II.3}$$

The two-electron integrals over the atomic orbital basis are written following the [12|12] notation and the $P_{\mu\nu}$ represent the elements of the density matrix.

In HONDO, the two-electron integrals, henceforth [ij|kl], are classified in fourteen different classes according to the relations between the four indexes i j, k, l. When saved to the disk, five integers and a double-precision variables, corresponding to the four indices, a tag to assign the integral to a class and the value of the integral itself, respectively, are packed in 12 byte records. Then, depending on the assigned class, the integral is multiplied by a factor and contributes in a determinate way to the two-electron part of the Fock matrix. In our case, however, we felt more convenient not to *a priori* classify the two-electron integrals, since they should be further reclassified when constructing the two-electron part of the CHA Fockian. The alternative algorithm designed for the calculation of the two-electron part of the Fockian is as follows:



SCHEME 22

where γ represents the actual value of the [ij|kl] integral. Depending upon the relationship between the indexes, the given integral contributes to the two-electron part of the Fockian in one or more cases. With this algorithm, the coulombic and exchange (marked bold) contributions are computed simultaneously.

In the case of CHA-SCF, the SCF algorithm must be slightly modified. As we discussed in the first section of the present work, some of the one and two-center integrals are modified or discarded in order to eliminate the pure BSSE-type extension from the Fockian. Mayer showed that the expresion of the CHA Fockian can be expressed simply as[65]

$$F_{\mu\nu}^{CHA} = H_{\mu\nu}^{core(CHA)} + G_{\mu\nu}^{CHA} \quad ;\forall\, \mu,\nu \tag{II.4}$$

where the core and two-electron parts are obtained using the CHA-modified one and two-electron integrals (see Eq. (34))

$$\begin{aligned} H_{\mu\nu}^{core(CHA)} = \left\{\chi_\mu \mid \hat{h} \mid \chi_\nu\right\} = \sum_{\lambda\in A} A_{\mu\lambda}^{A} < \chi_\lambda \mid -\frac{1}{2}\nabla^2 - \sum_{\alpha\in A}\frac{Z_\alpha}{\left|r-R_\alpha\right|} \mid \chi_\nu > - \\ - < \chi_\mu \mid \sum_{\alpha\notin A}\frac{Z_\alpha}{\left|r-R_\alpha\right|} \mid \chi_\nu > \qquad (\nu \in A) \end{aligned} \tag{II.5}$$

and

$$\begin{aligned} G_{\mu\nu}^{CHA} &= \sum_{\kappa,\lambda} P_{\lambda\kappa}\left(2\left\{\varphi_\mu\varphi_\kappa \mid \varphi_\nu\varphi_\lambda\right\} - \left\{\varphi_\mu\varphi_\nu \mid \varphi_\kappa\varphi_\lambda\right\}\right) = \\ &= \sum_{\kappa,\lambda} P_{\lambda\kappa}\left(\underbrace{2\sum_{\rho,\sigma\in A} A_{\mu\rho}^{A} A_{\kappa\sigma}^{A}\left[\varphi_\rho\varphi_\sigma \mid \varphi_\nu\varphi_\lambda\right]}_{\nu,\lambda\in A} - \underbrace{\sum_{\rho,\sigma\in A} A_{\mu\rho}^{A} A_{\nu\sigma}^{A}\left[\varphi_\rho\varphi_\sigma \mid \varphi_\kappa\varphi_\lambda\right]}_{\kappa,\lambda\in A}\right) \end{aligned} \tag{II.6}$$

The superscript A represents a given fragment composing the molecular complex. Also, the auxiliary A^A matrix is defined for each fragment as

$$A_{\mu\nu}^{A} = \sum_{\kappa\in A} S_{\mu\kappa}(S_A^{-1})_{\kappa\nu} \qquad (\nu \in A) \tag{II.7}$$

where S and S_A^{-1} represent the overlap matrix of the atomic orbitals of the whole complex, and the inverse of the overlap matrix of the atomic orbitals assigned to the A^{th} fragment, respectively.

Note that the matrix elements of the core and two-electron contributions to the Fockian are not symmetric, unless the two indices refer to atomic orbitals assigned to the same fragment. In this particular case the CHA integrals are equivalent to the conventional ones, as pointed out by Eqs. (20) and (21).

The proposed algorithm to obtain the CHA Fockian matrix is be presented next. We will focus now on the one-electron (core) contribution.

First of all, as one can see in Eq. (II.5), the nuclear repulsion matrix elements must be obtained for each fragment, separately, in order to build the core Hamiltonian matrix for each atom (h^A). The overlap matrix must also be split into block in order to obtain the S_A^{-1} elements. However, all these fragment matrix elements can be put into $N_b \times N_b$ square matrices (being N_b the total number of AO's) and the CHA core Hamiltonian can be obtained with the following matrix equation

$$H^{core(CHA)} = S\, S_0^{-1}\, H_0^{eff} + V^{int} \tag{II.8}$$

where:

$$S_0^{-1} = \begin{pmatrix} (S_{11})^{-1} & 0 & \ldots & 0 \\ 0 & (S_{22})^{-1} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & (S_{NN})^{-1} \end{pmatrix} \tag{II.9}$$

is a block diagonal matrix gathering the fragment inverse overlap $N_b^A \times N_b^A$ matrices;

$$H_0^{eff} = \begin{pmatrix} h_{11}^1 & 0 & \ldots & 0 \\ 0 & h_{22}^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & h_{NN}^N \end{pmatrix} = \begin{pmatrix} T_{11} & 0 & \ldots & 0 \\ 0 & T_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & T_{NN} \end{pmatrix} + \begin{pmatrix} V_{11}^1 & 0 & \ldots & 0 \\ 0 & V_{22}^2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & V_{NN}^N \end{pmatrix} \tag{II.10}$$

is the block diagonal matrix containing the $N_b^A \times N_b^A$ matrices of the kinetic

energy and electrostatic interaction of the electrons and nuclei of each fragment;

$$V^{int} = \begin{pmatrix} V_{11}^{\bar{1}} & V_{12}^{\bar{2}} & \cdots & V_{1N}^{\bar{N}} \\ V_{21}^{\bar{1}} & V_{22}^{\bar{2}} & \cdots & V_{2N}^{\bar{N}} \\ \cdots & \cdots & \cdots & \cdots \\ V_{N1}^{\bar{1}} & V_{N2}^{\bar{2}} & \cdots & V_{NN}^{\bar{N}} \end{pmatrix} \tag{II.11}$$

represents, in columns, the intermonomer electron-nuclei interaction, that is for the i^{th} block column, the electron-nuclei contribution due to the nuclei of the i^{th} fragment have been subtracted.

$$V^{int} = \begin{pmatrix} V_{11} & V_{12} & \cdots & V_{1N} \\ V_{21} & V_{22} & \cdots & V_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ V_{N1} & V_{N2} & \cdots & V_{NN} \end{pmatrix} - \begin{pmatrix} V_{11}^{1} & V_{12}^{2} & \cdots & V_{1N}^{N} \\ V_{21}^{1} & V_{22}^{2} & \cdots & V_{2N}^{N} \\ \cdots & \cdots & \cdots & \cdots \\ V_{N1}^{1} & V_{N2}^{2} & \cdots & V_{NN}^{N} \end{pmatrix} \tag{II.12}$$

The first term or the left-hand side of Eq. (II.8) can be further compacted by computing the A matrix defined in Eq. (II.7)

$$\begin{aligned} A = S\,S_0^{-1} &= \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1N} \\ S_{21} & S_{22} & \cdots & S_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ S_{N1} & S_{N2} & \cdots & S_{NN} \end{pmatrix} \begin{pmatrix} (S_{11})^{-1} & 0 & \cdots & 0 \\ 0 & (S_{22})^{-1} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & (S_{NN})^{-1} \end{pmatrix} = \\ &= \begin{pmatrix} 1 & S_{12}(S_{22})^{-1} & \cdots & S_{1N}(S_{NN})^{-1} \\ S_{21}(S_{11})^{-1} & 1 & \cdots & S_{2N}(S_{NN})^{-1} \\ \cdots & \cdots & \cdots & \cdots \\ S_{N1}(S_{11})^{-1} & S_{N2}(S_{22})^{-1} & \cdots & 1 \end{pmatrix} \end{aligned} \tag{II.13}$$

Hence, the elements of the CHA core Hamiltonian are simply computed by straightforward matrix products and summation of $N_b \times N_b$ matrices. The final expression results

$$H_0^{eff} = \begin{pmatrix} h_{11}^1 + V_{11}^{\bar{1}} = H_{11}^{core} & S_{12}\left(S_{22}\right)^{-1} h_{22}^2 + V_{12}^{\bar{2}} & \ldots & S_{1N}\left(S_{NN}\right)^{-1} h_{NN}^N + V_{1N}^{\bar{N}} \\ S_{21}\left(S_{11}\right)^{-1} h_{11}^1 + V_{21}^{\bar{1}} & h_{22}^2 + V_{22}^{\bar{2}} = H_{22}^{core} & \ldots & S_{2N}\left(S_{NN}\right)^{-1} h_{NN}^N + V_{2N}^{\bar{N}} \\ \ldots & \ldots & \ldots & \ldots \\ S_{N1}\left(S_{11}\right)^{-1} h_{11}^1 + V_{N1}^{\bar{1}} & S_{N2}\left(S_{22}\right)^{-1} h_{22}^2 + V_{N2}^{\bar{2}} & \ldots & h_{NN}^N + V_{NN}^{\bar{N}} = H_{NN}^{core} \end{pmatrix} \quad \text{(II.14)}$$

For the two-electron contribution, a similar algorithm to that of SCHEME 22 could be used by substituting the conventional [ij|kl] integrals by the CHA ones {ij|kl}. However, in this way, a two-index transformation of the conventional integrals by means of the A matrices would be necessary to get *each* CHA integral to be used in the spirit of the previous algorithm. In other words, it would represent a $O(N_b^2)$ transformation carried out $O(N_b^6)$ times; altogether $O(N_b^6)$ calculations, as well as the storage of a four index array, $O(N_b^4)$. This procedure would be clearly inefficient from a computational point of view compared to a conventional SCF calculation.

One can circumvent this transformation by using another strategy. The key point is, given a conventional [ij|kl] integral, to compute all the partial contributions to the elements of G^{CHA} trough the CHA integrals. That is, the {ij|kl}CHA integrals are never computed explicitly, but their contribution to the appropriate elements of G^{CHA} are stepwise added by performing the summation with all the [αβ|κλ] integrals that intervene in the computation of {ij|kl}.

In the Introduction it was shown that only that [ij|kl] integrals where k and l are AO's that belong to the same fragment A and both i and j do not belong A, must be CHA-transformed. Moreover, Eq. (19) shows that the only [αβ|κλ] integrals that contribute to the CHA integral {ij|kl} are the ones of the form [i'j'|kl], ∀i', j' ∈ A[164]. By using the A matrix

$$\{\chi_\mu \chi_\sigma \mid \chi_\nu \chi_\rho\} = \sum_{\tau,\lambda} A_{\mu\tau}^A A_{\sigma\lambda}^A [\chi_\tau \chi_\lambda \mid \chi_\nu \chi_\rho] \qquad (\nu, \rho, \lambda, \tau \in A) \qquad \text{(II.15)}$$

Hence the contribution of the CHA integrals to G^{CHA} can be expressed as

[164] Due to the structure of matrix A, in the particular case that either i or j belong to fragment A, the only non-zero contribution to the CHA integral would occur for either i'=i or j'=j, respectively.

$$G_{\mu\nu}^{CHA}(\nu \in A) \leftarrow \sum_{\rho \in A,\sigma} P_{\rho\sigma}\{\chi_\mu\chi_\sigma \mid \chi_\nu\chi_\rho\} = \sum_{\rho \in A,\sigma} P_{\rho\sigma} \sum_{\tau,\lambda \in A} A_{\mu\tau}^{A} A_{\sigma\lambda}^{A}[\chi_\tau\chi_\lambda \mid \chi_\nu\chi_\rho] \qquad \text{(II.16)}$$

The direct application of such an scheme is again too expensive. Now each [ij|kl] integral would be used *once* but again a two-index transformation would be necessary to determine all their contributions to the Fockian. That is, it requires to perform $O(N_b^4)$ times a $O(N_b^2)$ transformation; again a $O(N_b^6)$ algorithm but with no storage of a four-index array since the CHA integrals will not be explicitly computed.

One can do the following[165]:

In the first step, an auxiliary matrix can be defined as

$$B_{\rho\lambda}^{A} = \sum_{\sigma} P_{\rho\sigma} A_{\sigma\lambda}^{A} \quad (\rho,\lambda \in A) \qquad \text{(II.17)}$$

for each fragment A and stored.

Then, using the previously computed B matrices, a second set of matrices can be obtained using the two-electron integrals

$$X_{\tau\nu}^{A} = \sum_{\rho,\lambda \in A} B_{\rho\lambda}^{A}\left[\chi_\tau\chi_\lambda \mid \chi_\nu\chi_\rho\right] \quad (\tau,\nu \in A) \qquad \text{(II.18)}$$

Finally, the product of both matrices gives the elements of G^{CHA} :

$$G_{\mu\nu}^{CHA} = \sum_{\tau \in A} A_{\mu\tau}^{A} X_{\tau\nu}^{A} \quad (\nu \in A) \qquad \text{(II.19)}$$

All the operations above must be carried out for each fragment in the molecular system, as remarked by the corresponding subscripts. However, the use of block diagonal and block off-diagonal $N_b \times N_b$ matrices again allows for an easy computation of all the fragment contributions with a single calculation. Now, the B and X matrices above mentioned take the form

[165] I. Mayer, Á. Vibók, Chem. Phys. Lett. 140, 558 (1987)

$$B = P\,A = \begin{pmatrix} P_{11} & P_{12} & \ldots & P_{1N} \\ P_{21} & P_{22} & \ldots & P_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ P_{N1} & P_{N2} & \ldots & P_{NN} \end{pmatrix} \begin{pmatrix} 1 & A_{12} & \ldots & A_{1N} \\ A_{21} & 1 & \ldots & A_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ A_{N1} & A_{N2} & \ldots & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} P_{11} + \sum_{i \neq 1}^{N} P_{1i} A_{i1} & 0 & \ldots & 0 \\ 0 & P_{22} + \sum_{i \neq 2}^{N} P_{2i} A_{i2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & P_{NN} + \sum_{i \neq N}^{N} P_{Ni} A_{iN} \end{pmatrix} \qquad \text{(II.20)}$$

$$X = \begin{pmatrix} B_{11} & 0 & \ldots & 0 \\ 0 & B_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & B_{NN} \end{pmatrix} \begin{pmatrix} [ij]_{11} & 0 & \ldots & 0 \\ 0 & [ij]_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & [ij]_{NN} \end{pmatrix}_{\tau,\nu} =$$

$$= \begin{pmatrix} B_{11}[ij \mid \tau\nu]_{11} & 0 & \ldots & 0 \\ 0 & B_{22}[ij \mid \tau\nu]_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & B_{NN}[ij \mid \tau\nu]_{NN} \end{pmatrix} \qquad \text{(II.21)}$$

$$G^{CHA} = \begin{pmatrix} 1 & A_{12} & \ldots & A_{1N} \\ A_{21} & 1 & \ldots & A_{2N} \\ \ldots & \ldots & \ldots & \ldots \\ A_{N1} & A_{N2} & \ldots & 1 \end{pmatrix} \begin{pmatrix} X_{11} & 0 & \ldots & 0 \\ 0 & X_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & X_{NN} \end{pmatrix} =$$

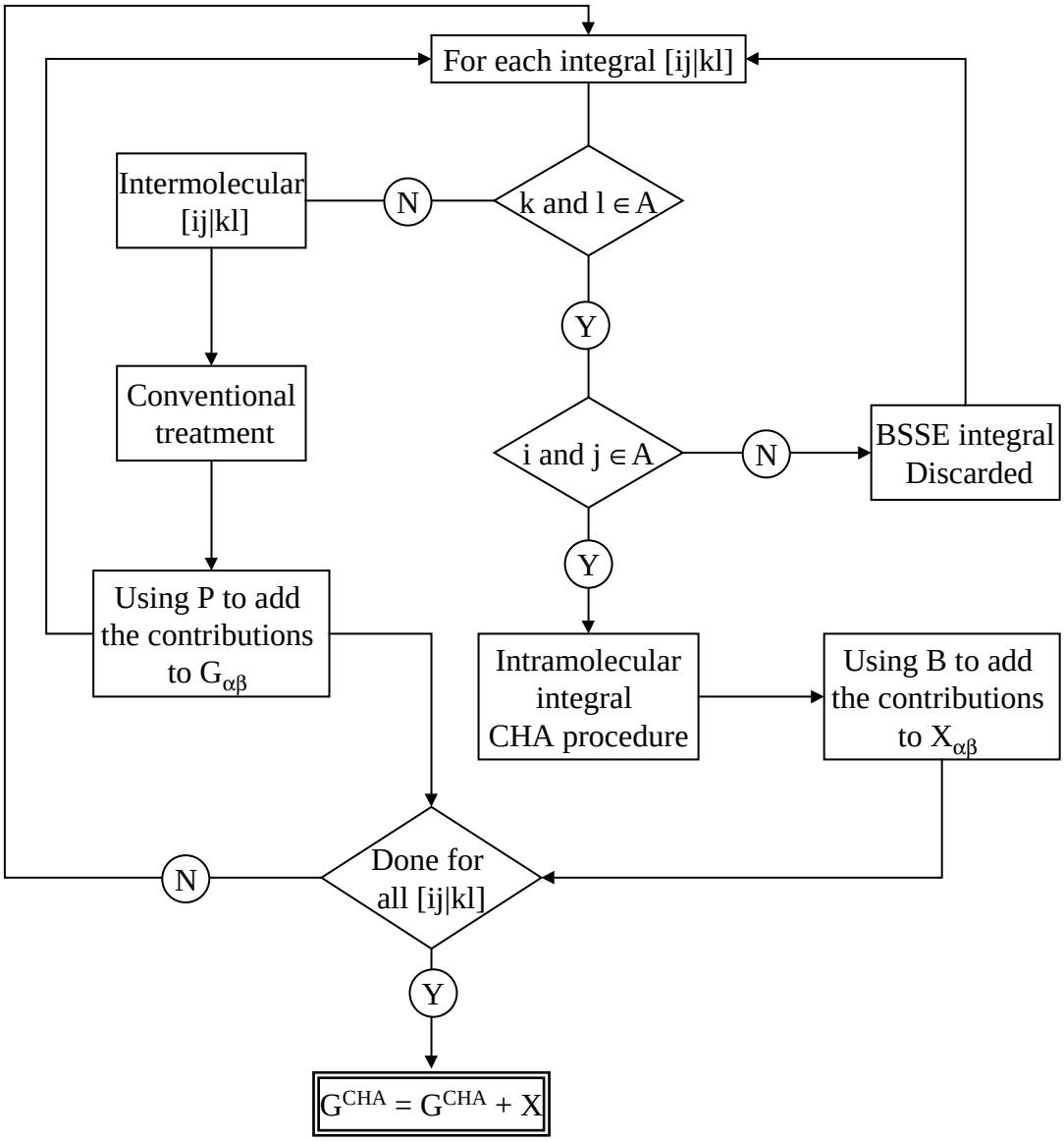
$$= \begin{pmatrix} X_{11} & A_{12}X_{22} & \ldots & A_{1N}X_{NN} \\ A_{21}X_{11} & X_{22} & \ldots & A_{2N}X_{NN} \\ \ldots & \ldots & \ldots & \ldots \\ A_{N1}X_{11} & A_{N2}X_{22} & \ldots & X_{NN} \end{pmatrix} \qquad \text{(II.22)}$$

In this way, the $O(N^6)$ algorithm is transformed in $2\times O(N_b^3)$ operations and the unavoidable $O(N_b^4)$ scaling, which is also necessary in case of the conventional SCF. Hence, as the number of basis set increases the extra computational effort introduced by the CHA tends to be negligible.

There are still an important point to take into account. Since, in general, each CHA two-electron integral contributes to more than one element of the two-electron

Fockian, G^{CHA}, all the possible contributions (for instance, the exchange) must be taken into account when building the X matrix in an analogous procedure to that of the conventional SCF. However, in this case, only that conventional integrals involving all four indexes of the same fragments will be taken into account. On the other hand, all [ij|kl] integrals with k and l indices referring to different fragments are pure intermolecular integrals. That means they are not CHA-modified and contribute to G^{CHA} in the same fashion they would contribute on a conventional SCF procedure.

Hence the final CHA algorithm for the two-electron contribution is the following:



SCHEME 23

At this point, the CHA Fockian is gained by adding the CHA one and two-electron contributions, as given in Eq. II.4. Once we have the CHA Fockian computed with some given molecular orbitals (by means of the density matrix, P), the algorithm to solve the CHA-SCF equations is that presented before in Section I.1.4.3 (See SCHEME 2).

II.2.2.2 UHF and CHA-UHF

For the unrestricted equations the same strategy is followed. The only difference now is that the alpha and beta contributions are computed separately. An extra set of matrices must be used except for that containing one and two-electron integrals, like overlap, kinetic energy, electron-nuclei attraction or two-electron repulsion.

The UHF equations can be expressed simply as

$$\begin{aligned} F^{\alpha}_{\mu\nu} &= H^{core}_{\mu\nu} + G^{\alpha}_{\mu\nu} \\ F^{\beta}_{\mu\nu} &= H^{core}_{\mu\nu} + G^{\beta}_{\mu\nu} \end{aligned} \qquad \text{(II.23)}$$

were H^{core} is that given by Eq. II.2 and

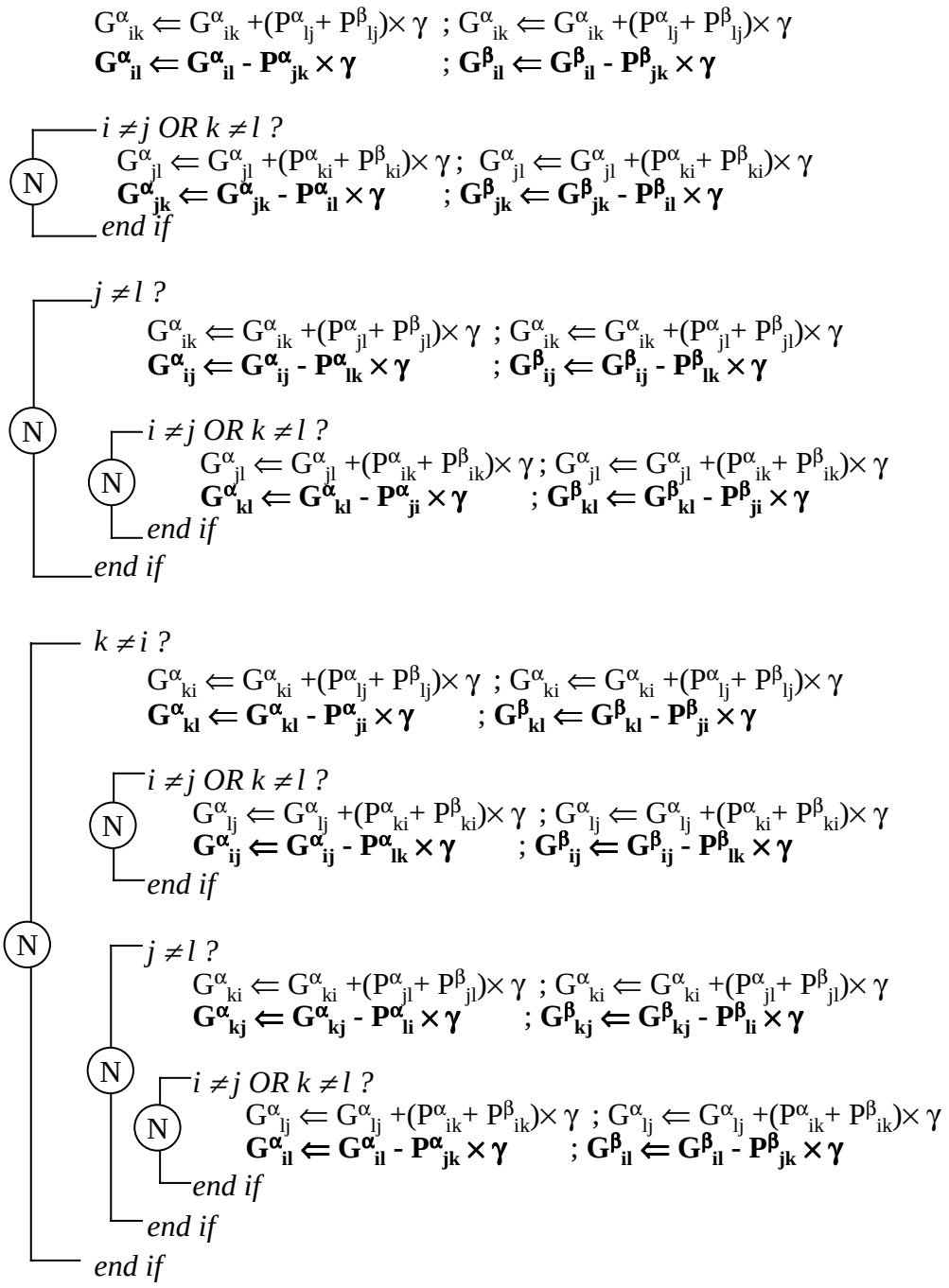
$$\begin{aligned} G^{\alpha}_{\mu\nu} &= \sum_{\kappa,\lambda} (P^{\alpha}_{\lambda\kappa} + P^{\beta}_{\lambda\kappa}) \left[\varphi_\mu \varphi_\kappa \mid \varphi_\nu \varphi_\lambda\right] - P^{\alpha}_{\lambda\kappa} \left[\varphi_\mu \varphi_\nu \mid \varphi_\kappa \varphi_\lambda\right] \\ G^{\beta}_{\mu\nu} &= \sum_{\kappa,\lambda} (P^{\alpha}_{\lambda\kappa} + P^{\beta}_{\lambda\kappa}) \left[\varphi_\mu \varphi_\kappa \mid \varphi_\nu \varphi_\lambda\right] - P^{\beta}_{\lambda\kappa} \left[\varphi_\mu \varphi_\nu \mid \varphi_\kappa \varphi_\lambda\right] \end{aligned} \qquad \text{(II.24)}$$

The core Hamiltonian does not depend on the density matrix so that it is calculated once and it is used for both alpha and beta contributions.

The G^{α} and G^{β} contributions must be computed at the same SCF step because both depend explicitly on the density matrices of the alpha and beta part, P^{α} and P^{β}. Eq. (II.24) shows that the coulombic contribution (first term) is equivalent for both alpha and beta parts. Only the exchange term can induce to different alpha and beta Fockians and hence to different alpha and beta molecular orbitals (spin density matrices). Obviously, when the number of alpha and beta electrons is different, the resulting alpha and beta density matrices differ. If the number of alpha and beta electrons is the same, the RHF solution is always a solution. There may exist other UHF solutions lower in energy with different alpha and beta orbitals. In general, for normal geometries the truly UHF solution may not exist. However, when a bond is stretched with respect to the equilibrium distance, the UHF solution will always be

observed at large bond lenghts.

The unrestricted open-shell algorithm for the determination of the two-electron contribution to the conventional Fockian is easily derived from SCHEME 22:



SCHEME 24

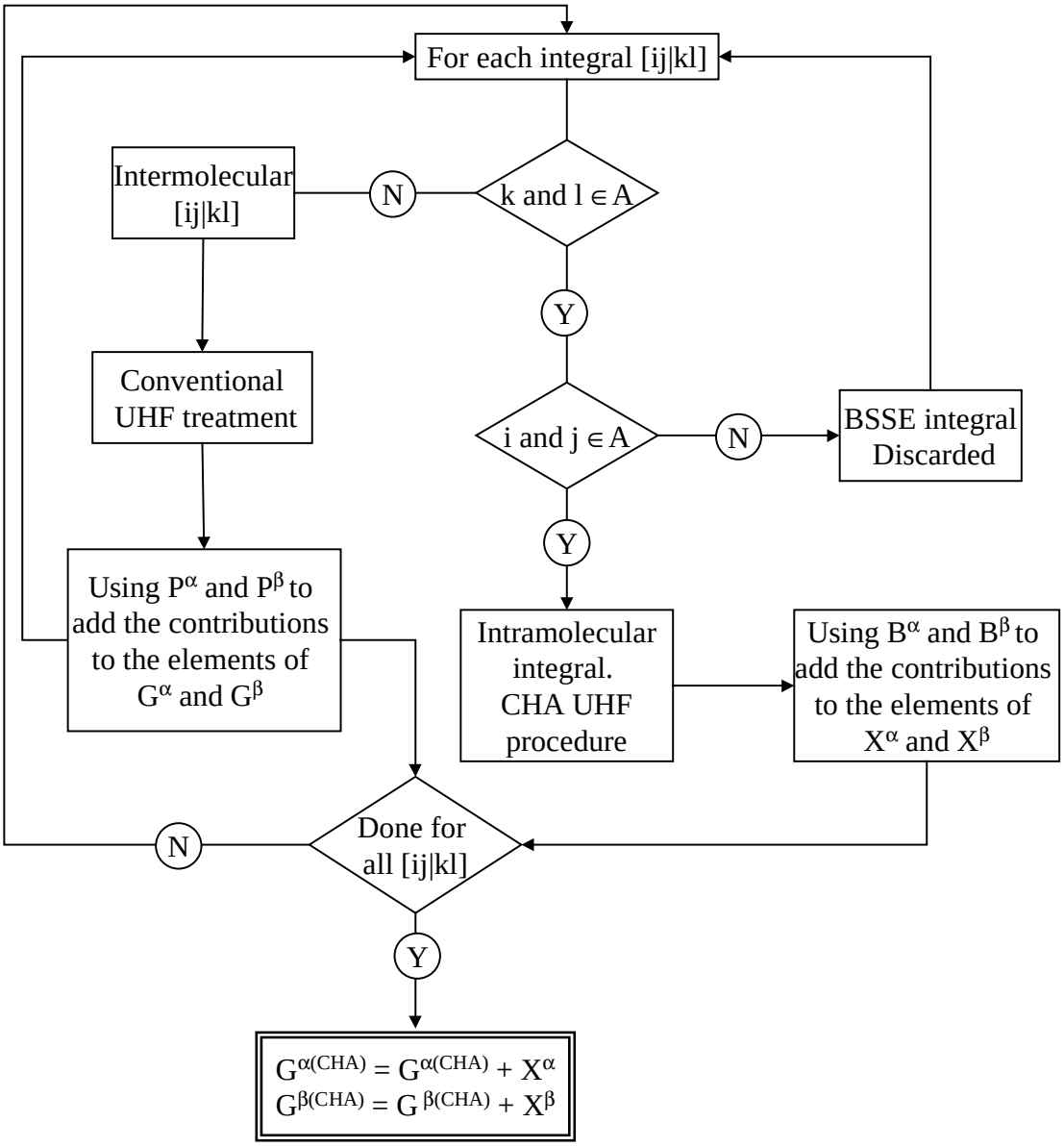
The corresponding CHA version of the UHF equations have the same peculiarities. The CHA core Hamiltonian does not depend on the density matrix so that it is calculated once following Eqs. (II.8)- (II.14) and it is used for both alpha and beta Fockians. The two-electron alpha and beta Fockians must be obtained for the alpha and beta contributions separately. At each iteration, $G^{\alpha(CHA)}$ and $G^{\beta(CHA)}$ are constructed using the previous P^{α} and P^{β}matrices and a new guess for the corresponding density matrices is obtained after diagonalization of the alpha and beta Fockians until self-consistency is reached.

$$G_{\mu\nu}^{\alpha(CHA)} = \sum_{\kappa,\lambda}(P_{\lambda\kappa}^{\alpha}+P_{\lambda\kappa}^{\beta})\underbrace{\sum_{\rho,\sigma\in A}A_{\mu\rho}^{A}A_{\kappa\sigma}^{A}\left[\varphi_{\rho}\varphi_{\sigma}\mid\varphi_{\nu}\varphi_{\lambda}\right]}_{\nu,\lambda\in A}-P_{\lambda\kappa}^{\alpha}\underbrace{\sum_{\rho,\sigma\in A}A_{\mu\rho}^{A}A_{\nu\sigma}^{A}\left[\varphi_{\rho}\varphi_{\sigma}\mid\varphi_{\kappa}\varphi_{\lambda}\right]}_{\kappa,\lambda\in A}$$

$$G_{\mu\nu}^{\beta(CHA)} = \sum_{\kappa,\lambda}(P_{\lambda\kappa}^{\alpha}+P_{\lambda\kappa}^{\beta})\underbrace{\sum_{\rho,\sigma\in A}A_{\mu\rho}^{A}A_{\kappa\sigma}^{A}\left[\varphi_{\rho}\varphi_{\sigma}\mid\varphi_{\nu}\varphi_{\lambda}\right]}_{\nu,\lambda\in A}-P_{\lambda\kappa}^{\beta}\underbrace{\sum_{\rho,\sigma\in A}A_{\mu\rho}^{A}A_{\nu\sigma}^{A}\left[\varphi_{\rho}\varphi_{\sigma}\mid\varphi_{\kappa}\varphi_{\lambda}\right]}_{\kappa,\lambda\in A} \quad \text{(II.25)}$$

Analogously to the conventional UHF equations, the only difference with respect to the RHF equations appears when computing the exchange contribution. In the case of the coulombic contribution, the alpha or beta density matrices contribute both $G^{\alpha(CHA)}$ and $G^{\beta(CHA)}$; that is, instead of using the corresponding P^{α} or P^{β} elements, the total spin matrix, $P = P^{\alpha} + P^{\beta}$, is necessary.

Hence, the implemented algorithm for the CHA-UHF case is easily derived from the CHA-RHF one given in SCHEME 23.


For each integral [ij|kl]
k and l ∈ A
N
Intermolecular [ij|kl]
Y
Conventional UHF treatment
i and j ∈ A
N
BSSE integral Discarded
Y
Using Pα and Pβ to add the contributions to the elements of Gα and Gβ
Intramolecular integral. CHA UHF procedure
Using Bα and Bβ to add the contributions to the elements of Xα and Xβ
Done for all [ij|kl]
N
Y
Gα(CHA) = Gα(CHA) + Xα
Gβ(CHA) = Gβ(CHA) + Xβ

SCHEME 25

II.2.2.3 Test results

A key point of the code is the diagonalization of a non-symmetric matrix. As commented in the introductory section, the eigenvectors are not orthogonal to each other. Since the actual derivation of the CHA-SCF equations imposed the orthonormality condition for the occupied molecular orbitals, they must be orthogonalized at each CHA-SCF cycle before using them for the next iteration. The most important consequence of the non-orthogonality of the eigenvectors is that the virtual and occupied spaces overlap (mix). In other words, they do not obey the conventional Brillouin theorem. This problem can be circumvented by using the biorthogonality properties of the left and right eigenvectors. Indeed, one can compute both left and right eigenvectors (in columns and rows, respectively) and take the occupied orbitals from the left eigenvectors and the virtual from the corresponding right eigenvectors.

Another typical feature is that the diagonalization of a non-symmetric matrix may produce complex eigenvectors and eigenvalues, that would occur in pairs, and convergence problems due to singularities. Hence, it is essential to use a robust algorithm capable to deal with complex solutions and avoid numerical instabilities. As pointed out by Mayer[68], the cases when the non-symmetric matrix is not diagonalizable have no practical relevance since infinitesimal changes on the matrix elements are sufficient to avoid this singularity. Several matrix diagonalization techniques were tested by Mayer years ago. In the case of the implementation presented here, a subroutine written by Mayer[166] that performs the diagonalization of a real non-symmetric matrix by a Jacobi-type method has been used. Also, both left and right eigenvectors can be computed.

In the general case of CHA-SCF, only the right eigenvectors are used because only the occupied orbitals are necessary to build the corresponding Fockian. The

[166] Written by I. Mayer (Budapest 1986) on the basis of an algol procedure in Wilkinson-Reinsch book: Handbook for automatic computation.Linear algebra,v.2,Springer,Berlin 1971.

particular case of complex orbitals do not represent any complication as long as both components of the complex pair are assigned to the occupied space. In that case the density matrix remains real. Indeed, the normalized real and imaginary parts can be used to build the density matrix.

Both methods have been successfully implemented and tested for several molecular systems, namely H and He clusters, diatomic molecules and B_2H_6. The results for up to three fragments have been compared with the existing programs of Mayer showing perfect agreement. It is remarkable that the algorithm presented here and that implemented by Mayer were completely different (A and B matrices were computed using a different strategy that resulted efficient for the particular case of two fragments). Since both programs produce identical results programming errors in the code are highly unlikely.

Test calculations have revealed that CHA calculations of the same molecular system but using different number of fragments (up to eight) show almost no extra computational cost with respect to the number of fragments.

Also, for a given molecular orbital guess, the number of cycles to reach self-consistency is, in general, similar for the CHA and for the conventional SCF. However, the symmetric character of the conventional Fockian allows for the application of convergence acceleration techniques, namely the DIIS[73]. Therefore, when compared to a optimized SCF code including those convergence acceleration techniques, the plain CHA calculation needs about twice the number of cycles to reach convergence.

Obviously, the best strategy for a CHA calculation is to compute first the SCF orbitals and use them as the initial guess for the CHA procedure. In general, about 5-15 extra cycles are needed to obtain the CHA canonical orbitals. Nevertheless, the derivation of efficient convergence acceleration techniques for non-symmetric SCF equations would be very interesting from both a mathematical and a practical point of view.

The study of the BH diatomic molecule with the unrestricted methodology provides both the RHF and the UHF solutions at large diatomic distances for both the conventional and the CHA calculation. The CHA PES obtained for the BH system is depicted in the following figure

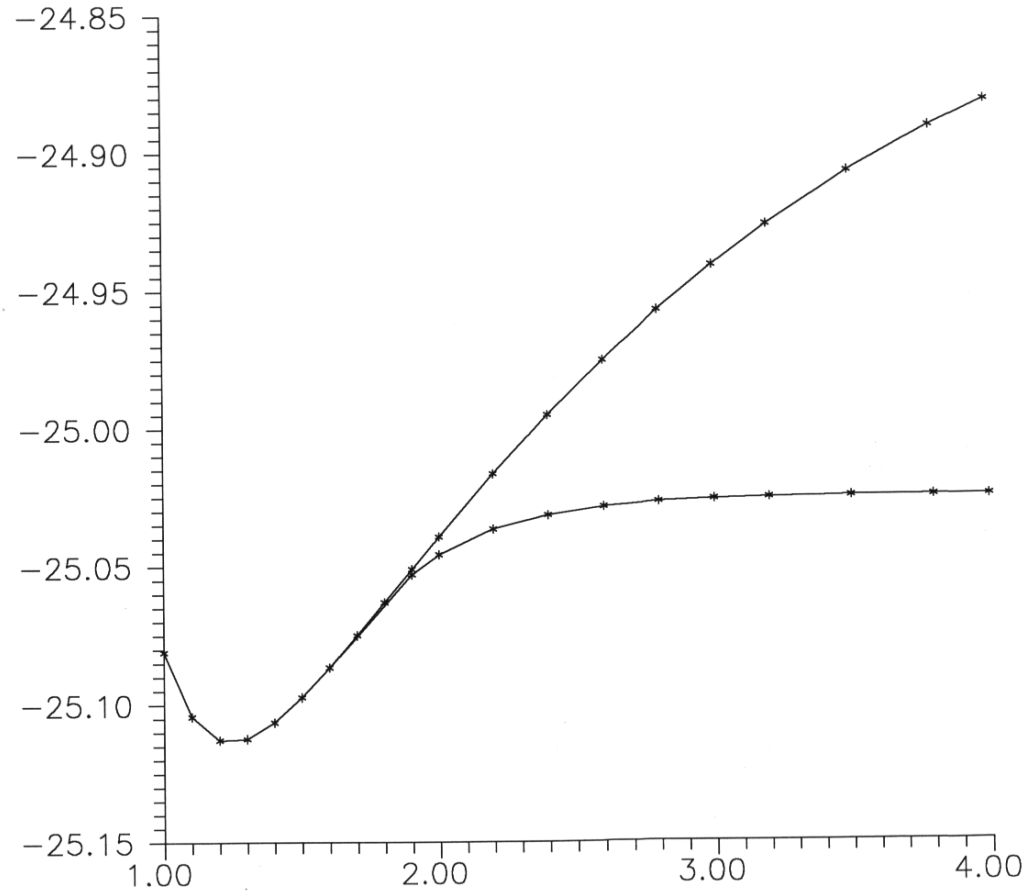


FIGURE 22: RHF (upper line) and UHF CHA energies for the BH molecule

The CHA–UHF results correctly describe the dissociation of the atoms into open-shell fragments, whereas the CHA-RHF does not. A closer inspection of the PES on the zone where the RHF and UHF curves split, shows that the true UHF solution appears at shorter distances for the conventional PES. The CHA PES softly splits into two, showing no numerical instability problems.

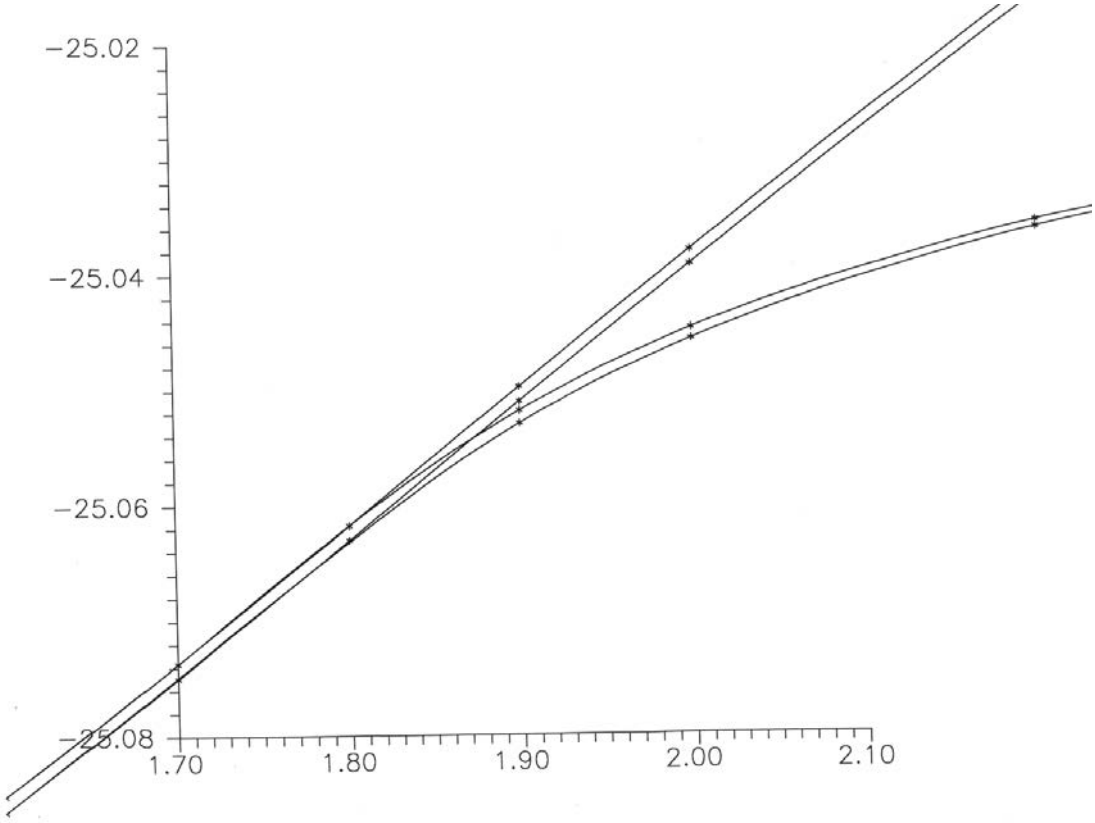


FIGURE 23 Comparison of CHA and conventional RHF (upper curve) and UHF Potential Energy Surfaces. The CHA PES lies slightly above the uncorrected one.

The actual implementation of the CHA-SCF methods has some limitations that are summarized below:

a) No gradients for eventual CHA geometry optimizations are available. As pointed out in Section I.1.4, the equations for the CHA gradient were derived by Paizs and Mayer, but they have not been implemented yet. As in the case of the CHA/F , numerical optimizations can be carried out by finite differences but this has not been explored in this work.

b) As the program is external to the *ab initio* code, HONDO-8 in this particular case, the CHA-SCF procedure must be carried out in a non-direct fashion. That is, the two-electron integrals must be computed once with HONDO and must be stored for its further use by the CHA subroutines. In this way, the storage requirement increase enormously with the size of the problem and soon become the bottleneck

of the algorithm. Hence, a direct algorithm were the two-electron integrals are computed each time they are needed is highly desirable. Moreover, the intramolecular and intermolecular two-electron integrals could be generated separately. The discarded integrals on the CHA algorithm will be computed only after self-consistency, in order to calculated the CHA energy[167].

c) Our modified HONDO-8 version can handle only up to 127 basis functions for a direct job so the use of the actual CHA programs for practical applications is very restricted. We are working towards the implementation of the CHA programs into a commercial package like Gaussian, in the spirit of the counterpoise methodology.

d) An inherent drawback of the CHA-SCF methods is clearly that the electron correlation is not taken into account. The accurate study of weakly bounded complexes is not possible without the inclusion of electron correlation. Only calculations for strong and moderate hydrogen-bonded complexes can be considered reliable. The inclusion of electron correlation through second-order Perturbation Theory is in progress. The preliminary results for both the restricted and unrestricted formulations without complex orbitals are promising.

[167] Note that the CHA energy is computed as the expectation value of the conventional Hamiltonian over the CHA wavefunction. Hence, in order to compute the CHA energy an extra SCF conventional cycle is necessary.

III Conclusions

The most relevant results derived from this work on the efficient treatment of the Basis Set Superposition Error can be summarized as follows:

FIRST: A new interpretation of the counterpoise method has been used in order to define counterpoise-corrected description of the molecular complexes. This novel point of view allowed us to study the BSSE-effects not only in the interaction energy but also on the potential energy surface and, in general, in any property derived from the molecular energy and its derivatives

SECOND: The usefulness of the combination of the DIIS and the variable metric BFGS methods for the location of the stationary points on the CP-corrected surfaces has been demonstrated. The use of an initial inverse Hessian matrix has been proved essential for a good convergence.

THIRD: A program to perform counterpoise-corrected optimizations and vibrational analysis at any level of theory and for three different counterpoise procedures has been codified. The possible symmetry of the complex can be as well taken into account in order to reduce the computational cost of the calculation. This program has allowed us to perform hierarchical counterpoise correction to clusters to asses the high order BSSE effects.

FOURTH: The BSSE induces to the overestimation of the interaction energy of the molecular complexes. The magnitude can be strongly dependent on the geometrical parameters so that the Potential Energy Surface and hence the vibrational frequencies and zero point energy contributions are also affected.

FIFTH: Optimized molecular geometries and stabilization energies obtained with the CHA/F and CHA/DFT methodologies have been compared with those obtained with the counterpoise method for several hydrogen bonded complexes. It has been shown that both methods correct for the BSSE in a very similar fashion and that the

differences between them are much lesser than the BSSE itself. Contrary to other approaches for the a priori correction of the BSSE, like the SCF-MI or SMMO methods, both the CHA and counterpoise-corrected stabilization energies tend asymptotically to the uncorrected ones as the basis set is increased.

SIXTH: The location of stationary points on this BSSE-corrected PES by means of both counterpoise or CHA methodologies provides molecular geometries corrected for BSSE. The BSSE-corrected geometries and vibrational frequencies obtained with rather small basis sets tend to the values obtained for much larger basis sets. Particularly the CP-correction involving geometry reoptimizations can be a very useful tool in order to yield good descriptions of larger molecules where the use of large basis sets including diffuse functions is prohibitive.

SEVENTH: The application of the extrapolation to the CBS limit of intermolecular geometry parameters and energies seems to be fruitful as long as the corresponding CP-corrected values are used. Also, fixing intramolecular parameters at their experimental values could cause difficulties during the extrapolation. As the available literature data and our results clearly show, the MP2/aug-cc-pV*X*Z {*X* = 2, 3, 4} data series of intermolecular distances obtained from the CP-corrected surfaces can be safely used for the purpose of CBS extrapolations.

EIGHTH: The fragment relaxation energy cannot be seen as an additional term to the energy, mostly when studying chemical processes like rotational barriers. The effect of the fragment relaxation on the intermolecular parameters probed to be very important for the $BF_3 \cdots NH_3$ complex, the CP-correction not changing this situation. The CP-correction scheme can be successfully applied despite the relaxation contribution being not taken into account.

NINTH: When dealing with charged molecular complexes both CP2 and CP3 proposed methods yield very similar results at the RHF level from both an energetic and structural point of view. Only appreciable differences in the geometry have been observed for the $(H_2O \cdots OH)^-$ complex when using the 6-311G** and D95** basis sets. At the MP2 level of theory the differences increase but, in general, are less important than the BSSE itself. There is a good agreement in the formation enthalpies

for non symmetric complexes, like the $(H_2O\cdots OH)^-$ and the $(NH_4\cdots H_2O)^+$.

TENTH: The corresponding expressions for the counterpoise-corrected cluster energies using three different correction schemes, namely SSFC, PAFC and VMFC have been derived by using a many-body partitioning of the energy of the aggregate. It has been shown that the effect of the second and third-order basis set extensions given by the VMFC method is negligible for some hydrogen fluoride cluster, although they are still meaningful when using small basis sets. The use of diffuse functions is essential for the saturation of the high-order BSSE.

ELEVENTH: The effect of the geometry on the energetic terms obtained from a SAPT analysis of the hydrogen fluoride trimer reveal that that already at the level of two-body interactions the cyclic configuration is stabilized over the linear one. The main conclusion is that the large differences between the results obtained at the uncorrected and counterpoise-corrected geometries underscore the need for performing the analysis of the interaction energy at the counterpoise-corrected minimum geometries.

TWELVETH: The CHA method applied at the Hartree-Fock (CHA/F) and DFT (CHA/DFT) levels of theory has been used obtain BSSE-corrected wavefunctions of several hydrogen bonded complexes. A systematic study of the effect of the BSSE correction in terms of electronic and nuclear relaxation contributions has been carried out for several basis sets and functionals.

THIRTEENTH: When the uncorrected PES is qualitatively well described without the BSSE correction, both electronic and nuclear relaxation lead to deviations of the order of 10% in the intermolecular critical point location, its electron density and Laplacian. The convergence of uncorrected and CHA results for the largest basis set shows again the validity of this BSSE-correction scheme.

FOURTEENTH: The effect of BSSE on the electron density has proved to be more relevant at the DFT level than at the SCF level of theory, and to be strongly dependent on the inclusion of diffuse functions in the basis set. Addition of diffuse functions leads to similar effects for all the systems analyzed: an overall decrease of the

differences between corrected and uncorrected densities, negative differences in the intermolecular region, and lack of the highly directional density redistribution patterns in heavy donor and acceptor atoms that are observed with smaller basis sets.

FIFTEENTH: Density difference maps at frozen geometry reveal that the effects of the BSSE are not limited to the intermolecular region. Rather, the main redistribution effects take place in the valence shells of the heavy atoms directly involved in the intermolecular interaction. The study of larger systems, like the formic acid dimer, and especially the uracil-water complex, reveals that the effects of BSSE on the electron density are generally restricted to the intermolecular region and especially to the atoms directly involved in the intermolecular interaction and their first-neighbors.

SIXTEENTH: The origin of the BSSE in terms of one- and two-center contributions is very different depending on the inclusion of diffuse function. When no diffuse functions are considered, the BSSE is stabilizing for the intermolecular component and destabilizing for the intramolecular components. When diffuse functions are used, exactly the opposite is found. The density difference maps do also reveal systematic differences between calculations with and without diffuse functions, thus a CECA analysis may well be meaningful as a complement for the understanding of intermolecular interactions.

SEVENTEENTH: The Chemical Hamiltonian Approach methodology has been successfully implemented at the CHA-RHF and CHA-UHF levels of theory. The use of block diagonal matrices allows for an efficient implementation of the methodology for an arbitrary number of fragments. The programs extract from a previous HONDO run the necessary integrals and perform the CHA transformation on the fly. Hence, only nondirect type of calculations can be performed at the present time.

EIGHTEENTH: The implementation of the corresponding CHA-MP2 methodology is not yet operative due to the complications introduced when complex molecular orbitals and energies appear. The results for the intramolecular CHA formulation are not satisfactory and have been omitted.

IV Appendix

IV.1 Related Publications

Most of the results collected in this work have already been published in international journals of chemistry and physics. The complete reference of these published papers are given below:

1

The effect of counterpoise correction and relaxation energy term to internal rotation barriers: Application to the $BF_3 \cdots NH_3$ and $C_2H_4 \cdots SO_2$ dimers.

P. Salvador, M. Duran

J. Chem. Phys. 111 (1999) 4460-4465

2

Effect of Basis Set Superposition Error on the electron density of molecular complexes.

P. Salvador, X. Fradera and M. Duran

J. Chem. Phys. 112 (2000) 10106-10115

3

C-H···O H-Bonded complexes. How does basis set superposition error change their potential energy surfaces?

P. Salvador, S. Simon, M Duran, J. J: Dannenberg

J. Chem. Phys. 113 (2000) 5666-5674

4

Intermolecular bond lengths: extrapolation to the basis set limit on uncorrected and BSSE-corrected potential energy hypersurfaces.

B. Paizs, P. Salvador, A. G. Császár, M. Duran, S. Suhai.

J. Comput. Chem. 22 (2001) 196-207

5

On the effect of the BSSE on Intermolecular Potential Energy Surfaces. Comparison of a priori and a posteriori BSSE Correction Schemes

P. Salvador, B. Paizs, M. Duran, S. Suhai.

J. Comput. Chem. 22 7 (2001) 765-786

IV.2 Unpublished contributions

Some of the results shown in this Thesis have not been published at the present time. The partial references of the accepted and submitted manuscripts are given below:

6

Quantitative assessment of the effect of BSSE on the electron density of molecular complexes by means of Quantum Molecular Similarity Measures.

P. Salvador, M. Duran, X. Fradera

Adv. Mol. Sim., vol 4 (accepted)

7

On the counterpoise correction of charged intermolecular complexes

P. Salvador, M. Duran, J. J. Dannenberg

J. Chem. Phys., (submitted for publication)

8

Counterpoise-corrected geometries and harmonic frequencies of N-body clusters. Application to $(HF)_n$ (n = 3, 4).

P. Salvador and M. M. Szczesniak

J. Chem. Phys., (submitted for publication)

9

A Chemical Hamiltonian Approach study of the Basis Set Superposition Error changes on electron densities and one- and two-center energy components.

Pedro Salvador, Miquel Duran, Xavier Fradera

J. Chem. Phys., (submitted for publication)

IV.3 Other publications

A copy of two published papers corresponding to the references 42 and 154 is provided. The results included in these publications have not been explicitly discussed in this Thesis.

The first contribution, co-authored with Mrs M. Daza and Profs. J. Molina, J. A. Dobado and J. L. Villaveces is part of the Ph. D. Thesis of the former.

The second paper describes the implementation of the partitioning scheme in terms of one- and two-center energy components of the HF energy in the framework of the Atoms in Molecules theory. As discussed in the Section II.1.3.4, the computational cost of the actual implementation is so elevate that prevented us for systematically apply this methodology to the study of the considered hydrogen bonded complexes. Moreover, test calculations for the smallest H-bonded complex showed integration problems that weren't observed for single molecule calculations.

10

BSSE-error-counterpoise corrected Potential Energy Surfaces. Application to the Hydrogen Peroxide ...X (X = F-, Cl-, Br-, Li+, Na+) Complexes.

M. Daza, J.A. Dobado, J. Molina, P. Salvador, M. Duran, J.L. Villaveces

J. Chem. Phys. 110 (1999) 11806-11813

11

One- and two-center energy components in the AIM theory.

P. Salvador, M. Duran and I. Mayer

J. Chem. Phys. 115 (2001) 1153-1157