VIRT&L–COMM.9.2016.

Università degli Studi di Perugia (IT)
Dottorato di Ricerca in Scienze Chimiche
XXV Ciclo
Settore Scientifico Disciplinare CHIM/03
in cotutela con
Université Paul Sabatier, Toulouse (FR)

Molecular Simulations as test beds for bridging High Throughput and High Performance computing

Marco Verdicchio

Relatore Italiano:	Relatore Francese:
Prof. Antonio Laganà	Prof. Stefano Evangelisti
Università degli Studi di Perugia	*Université Paul Sabatier, Toulouse*

Anno Accademico 2011/2012

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

ISSN: 2279-8773

Contents

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

ISSN: 2279-8773

Introduction

The advances in computational sciences allow nowadays the *a priori* study of complex chemical systems and enable the detailed evaluation of macroscopic properties of the matter without resorting to empirical considerations or to the mixing of pieces of information coming from different (heterogeneous and sometimes inconsistent) sources. This has been made possible by the progress in theoretical chemistry (that has produced (and is producing) a large variety of methods and techniques enabling the foundation of the simulation of the molecular structures and processes of real systems on solid procedures based on first principles (from this the appellation *ab initio*)) and the evolution of networked computing platform(that is increasingly allowing the combined usage of a large amount of heterogeneous distributed computing resources).

The strong connotation of computational chemistry in terms of computer technologies is at the same time the strength and the weakness of molecular simulations. As a matter of fact, in order to perform such studies (even for few atom systems) we first need to carry out high level electronic structure calculations. These calculations typically require nodes (or clusters of nodes) equipped with large (of the order of many GB) memories and processors performing at the level of several Gigaflops because the whole Potential Energy Surface (PES) governing the nuclear motion needs to be worked out first. On the mentioned High Performance Computing (HPC) platforms with enanched parallel capabilities we can run concurrently, on several single multicore (or clusters of) processors, the calculations required by the (large number of) potential energy values necessary to describe the PES explored by a reactive chemical process. The real bottleneck in carrying out related computational campaigns, indeed, is repre-

VIRT&L-COMM.9.2016.4

sented by the availability of a computing platform having the proper computational requirements in terms of computing time and physical memory. The (limited) computing capabilities in general available to the scientific community, in fact, still set severe limitations to the development of full a priori computational simulations of molecular processes. Fortunately, innovative computing technologies combining concurrency and networking (such as distributed computing, virtual laboratories, supercomputing, Grid computing) are opening new prospects to the possibility of achieving significant computational throughputs and, therefore, of developing a priori molecular simulations of real systems.

In several scientific fields computational researchers need platforms providing access to HPC machines from High Throughput Computing (HTC) ones enabling the accurate modeling of real-like systems as well as virtual reality simulations based on multi-scale and multi-physics approaches. Within the activities carried out by the Chemistry and Molecular & Material Sciences and Technology Communities (CMMST), COMPCHEM [1] (the EGI (European Grid Infrastructure) [2] Virtual Organization (VO) for computational chemistry) is engaged in designing and implementing accurate realistic multiscale CMMST applications like the Grid Empowered Molecular Simulator (GEMS) [3, 4, 5]. GEMS requires the cooperation of complementary expertise exploiting the coordinated usage of various HPC from HTC computational platforms to run packages whose performances vary from platform to platform. The use of HPC and HTC e-infrastructures and the interoperation of large computational applications not only will allow the overcoming of the present highly unsatisfactory situation, but it will also allow an optimization of the usage of HPC computing resources for tightly coupled computational tasks (avoiding so far their use as a bunch of loosely coupled processors with the consequent underutilization of their fast dedicated network) and encourage the usage of HTC for purely distributed tasks avoiding so far ending up by wasting a large amount of time in data transfer. A coordination of the two types of platforms to interoperate via a single workflow and properly manage the various components on the most suitable of them, would instead allow a clever composition of complex applications optimizing the use of the various computing resources and offering to the users the best level of performance.

The theoretical foundationsand the computing paradigms employed for the assemblage of the components of the Grid Empowered Molecular Simulator GEMS are described in Chapter 1. In that chapter the development of grid based workflows allowing the *ab initio* evaluation of the observable properties of small chemical systems starting from the calculation of the electronic properties is illustrated.

In Chapter 2 the issue of the of interoperability between computational codes across different stages of the workflow is faced. The Chapter proposes Q5cost and D5cost common data modesl as de facto standard formats for quantum chemistry calculations.

Chapter 3 relates to the results of standalone *ab initio* calculations performed on different small chemical systems (X_4 clusters and BeH^- dimer). The Chapter discusses particular and interesting chemical bonds requiring high-level quantum methods to the end of being rationalized..

Finally Chapter 4 and Chapter 5 report the results of our work on two combustion and atmospheric chemistry problems ($CH_3CH_2OO^\bullet$ isomerization and N_2+N_2 reaction) respectively. They both aim at constructing the PES for a reactive process. Once a PES is generated, the kinetic and dynamical data needed to be calculated for a large number of initial conditions can be computed on HTC platforms. The assemblage of the computational workflows for the coupled use of HPC and HTC systems is also dealt there.

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

1

Molecular Simulators

From the very beginning the foundations of chemistry have been built upon the concepts of atoms and molecules. Accordingly also theoretical and computational chemistry take fractions of nanometers (tipycally the Bohr radius) as reference units and atomic postulates as *ab initio* laws. As a consequence, all phenomena and properties are re conducted to atomic and molecular level through a multi-scale approach and the construction of *ab initio* Molecular simulators is the starting block of any rationalization of the behavior of all physical systems. A Molecular Simulator is a set of software programs and tools driving the user from first principles down to observables through the accurate calculation of elementary quantities and the appropriate averaging of the unobserved parameters of a physical system.

1.1 THE GENERAL SCHRÖDINGER EQUATION

A physical system specified by a set of spatial nuclear coordinates ($\boldsymbol{R}$), electronic coordinates ($\boldsymbol{r}$) and by a time coordinate t is completely represented by a function $\Psi(\boldsymbol{R}, \boldsymbol{r}, t)$ called *system wavefunction* that has the property of describing the spatial and temporal distribution of the system. The quantum mechanical equation that governs the dynamics of the system, i.e. how the system evolves in time, is the time-dependent (TD) *Schrödinger equation*:

$$i\hbar \frac{\partial}{\partial t} \Psi(\boldsymbol{R}, \boldsymbol{r}, t) = \hat{H}\, \Psi(\boldsymbol{R}, \boldsymbol{r}, t) \tag{1.1}$$

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

where $\hat{H}$ is the *Hamilton operator*. For a general N-particle system the $\hat{H}$ operator consists of kinetic ($\hat{\boldsymbol{T}}$) and potential ($\hat{\boldsymbol{V}}$) term.

$$\hat{H} \quad = \quad \hat{T} + \hat{V} \tag{1.2}$$

which are formulated as a sum of the divergence operator ∇_i^2 and of all particles i and of all pair ij

$$\begin{aligned} \hat{T} &= -\sum_{i=1}^{N} \frac{\hbar^2}{2m_i} \nabla_i^2 \qquad (1.3) \\ \hat{V} &= \sum_{i}^{N} \sum_{j>i}^{N} V_{ij} \end{aligned}$$

Coulomb potential. If the Hamilton operator is independent of time, the time dependence of Ψ can be factorized as a simple phase factor to obtain the so called time-independent *Schrödinger equation*:

$$\hat{H}\,\Psi(\boldsymbol{R},\boldsymbol{r}) = E\,\Psi(\boldsymbol{R},\boldsymbol{r}). \tag{1.4}$$

Such equations are postulates. By itself, the wavefunction has no physical meaning. At each instant, however, its square modulus is proportional to the probability of the system to be in the spatial configuration ($\boldsymbol{R}$,$\boldsymbol{r}$).

1.2 THE ELECTRONIC STRUCTURE

The solution of equation 1.1 cannot in general be expressed in a closed form. In order to compute suitable numerical solutions to equation 1.1 one may consider to adopt decoupling and/or simplification schemes.

The first step is the decoupling of the equations for the center of mass based on the fact that its motion remains unaltered when no external forces act on the system. A more popular and efficient simplification is, however, the Born-Oppenheimer (BO) approximation. As the nuclei are much heavier than electrons and their speeds much smaller, the Schrödinger equation can be split into two parts. One part which describes the electronic wavefunction ($\phi^e(\boldsymbol{R};\boldsymbol{r})$) for a fixed value of $\boldsymbol{R}$ (nuclear geometry) and another part for the nuclear wavefunction ($\psi^n(\boldsymbol{R},t)$) in which the electronic energy

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

is formulated as a function of $\boldsymbol{R}$, $V(\boldsymbol{R})$. More in details adopting an approach typical of linear algebra methods, the wavefunction $\Psi(\boldsymbol{R}, \boldsymbol{r}, t)$ can be expanded in a complete set of $\phi_l^e(\boldsymbol{R}; \boldsymbol{r})$ functions:

$$\Psi(\boldsymbol{R}, \boldsymbol{r}, t) \quad = \quad \sum_l^{\infty} \psi_l^n(\boldsymbol{R}, t)\, \phi_l^e(\boldsymbol{R}; \boldsymbol{r})$$

where n refers to nuclei and e to electrons. At the same time by splitting the kinetic operator $\hat{T}$ into a nuclear ($\hat{T}_n$) and an electronic ($\hat{T}_e$) kinetic operator one has

$$\hat{T} \quad = \quad \hat{T}_n + \hat{T}_e \tag{1.5}$$

Accordingly the Schrödinger equation becomes:

$$i\hbar \frac{\partial}{\partial t} \sum_l^{\infty} \psi_l^n(\boldsymbol{R}, t)\, \phi_l^e(\boldsymbol{R}; \boldsymbol{r}) = \left[\hat{T}_n + \hat{T}_e + \hat{V}\right] \sum_l^{\infty} \psi_l^n(\boldsymbol{R}, t)\, \phi_l^e(\boldsymbol{R}; \boldsymbol{r}) \tag{1.6}$$

Following the already mentioned BO approximation, the electronic wavefunction $\phi_l^e(\boldsymbol{R}; \boldsymbol{r})$ which depends on electronic coordinates and, only parametrically, on nuclear coordinates, are usually chosen to be eigenfunctions of the operator $\hat{T}_e + \hat{V}$ satisfying the eigenvalue equation:

$$(\hat{T}_e + \hat{V})\, \phi_l^e(\boldsymbol{R}; \boldsymbol{r}) = V_l(\boldsymbol{R}) \phi_l^e(\boldsymbol{R}; \boldsymbol{r}) \tag{1.7}$$

The function $V_l(\boldsymbol{R})$ is the lth PES, function of the nuclear coordinates, of the lth electronic state on which nuclei move under the combined effect of its gradients and the kinetic operator $\hat{T}_n$.
The task of determining the PES can be accomplished using one of the available *ab initio* methods. Most of these methods are based on Hartree-Fock (HF) or post Hartree-Fock approaches. The essence of the Hartree Fock approximation is to solve the complicated many-electron problem in equation 1.7 using one-electron functions in which the electron-electron repulsion is treated in an average way. The application of this procedure leads to the Hartree Fock eigenvalue equations

$$\hat{F}_i\, \chi(\boldsymbol{r}_i) = \epsilon_i\, \chi(\boldsymbol{r}_i) \tag{1.8}$$

where $\hat{F}_i$ is the Fock operator for the ith electron and $\chi(\boldsymbol{r}_i)$ and ϵ_i are the Molecular Orbital (MO) and the MO energy respectively.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

The crucial complication in the HF procedure is the presence of the electron-electron potential energy that makes the Hartree Fock eigenvalue equation non linear and soluble only iteratively. The Potential part of the Fock operator, or equivalently the "field" seen by the ith electron, depends on the MO of the other electrons (i.e. the Fock operator depends on its eigenfunctions). The procedure for solving the HF equation is called Self Consistent Field (SCF) [6]. The Hartree Fock method, however, does not consider the instantaneous electrostatic interactions between electrons nor does it take into account the quantum mechanical effects on electrons distributions because the effect of other electrons is treated in an average way (ignoring *electron correlation*). A great deal of modern work in the field of electronic structure calculation is aimed at taking electron correlation into account. Hartree Fock, in fact, usually constitutes the first step towards the exact solution of the electronic Schrödinger equation and can be improved either by adding additional determinants to the wavefunction (Multi Reference, Configuration Interaction) or by perturbing the Hamiltonian (Perturbation Theory, Coupled Cluster), leading to the so called post-HF methods [7]. Additional approximations can be applied to the HF equations with the introduction of empirical parameters in the Fock operator matrix, leading to semiempirical quantum chemistry methods.
The methods considered may also include density functional theory (DFT) or Molecular Mechanics (MM). DFT lays its foundations in the concept that the ground-state electronic energy is determined completely by electron density ρ, while MM calculates the potential energy of the system using *force fields*.
The programs adopted for this type of calculations include both open source and commercial software. Most of them are large packages, often articulated into several separate programs, and have been assembled over many years. A list of such programs is given in Table 1.2.

Being our attention focused onto the motion of the nuclei on a single potential energy surface (electronically adiabatic motion) that is usually the electronic ground state $l = 0$, by rearranging equation 1.6 and neglecting, as from the BO approximation, cross terms (nucleus-electron), we get the following nuclei equation:

$$i\hbar\frac{\partial}{\partial t}\psi^n(\boldsymbol{R},t) = \left[\hat{T}_n + V(\boldsymbol{R})\right]\psi^n(\boldsymbol{R},t) \tag{1.9}$$

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Table 1.1: List of the most diffused *ab initio* codes and their principals capabitlities.

Package	License[a]	Lang.	Basis	Periodic[b]	Mol. Mech	Semi-empir.	HF	Post-HF	DFT
ABINIT	GPL	Fortran	PW	3d	Yes	No	No	No	Yes
ADF	Commercial	Fortran	STO	Any	Yes	Yes[1]	Yes	No	Yes
COLUMBUS	Academic	Fortran	GTO	No	No	No	Yes	Yes	No
CP2K	GPL	Fortran 95	Hybrid GTO / PW	Any	Yes	Yes	Yes	No	Yes
CPMD	Academic	Fortran	PW	Any	Yes	No	Yes	No	Yes
CRYSTAL	Academic (UK) / Commercial	Fortran	GTO	Any	Yes	No	Yes	Yes[2]	Yes
DALTON	Academic	Fortran	GTO	No	No	No	Yes	Yes	Yes
DIRAC	Academic	Fortran 77, Fortran 90, C	GTO	No	No	No	Yes	Yes	Yes
Firefly / PC GAMESS	Academic	Unknown	GTO	No	Yes[3]	Yes	Yes	Yes	Yes
GAMESS (UK)	Academic (UK) / Commercial	Fortran	GTO	No	No	Yes	Yes	Yes	Yes
GAMESS (US)	Academic	Fortran	GTO	No	Yes[4]	Yes	Yes	Yes	Yes
GAUSSIAN	Commercial	Fortran	GTO	Any	Yes	Yes	Yes	Yes	Yes
JAGUAR	Commercial	Unknown	GTO	Unknown	Yes	No	Yes	Yes	Yes
MOLCAS	Commercial	Fortran	GTO	No	Yes	Yes	Yes	Yes	Yes
MOLPRO	Commercial	Fortran	GTO	No	No	No	Yes	Yes	Yes
MOPAC	Academic / Commercial	Fortran	Unknown	Unknown	Unknown	Yes	No	No	No
NWChem	ECL v2	Fortran 77 / C	GTO, PW	Yes(PW) No (GTO)	Yes	Yes	Yes	Yes	Yes
ORCA	Academic	C++	GTO	No	Yes	Yes	Yes	Yes	Yes
Q-Chem	Commercial	Fortran / C++	GTO	No	Yes	Yes	Yes	Yes	Yes
Quantum ESPRESSO	GPL	Fortran	PW	3d	Yes	No	Yes	No	Yes
SIESTA	Academic	Fortran	NAO	3d	Yes	No	No	No	Yes
TURBOMOLE	Commercial	Fortran	GTO	No	Yes	No	Yes	Yes	Yes

[a] "Academic": academic (no cost) license possible upon request; "Commercial": commercially distributed.

[b] Support for periodic systems (3d-crystals, 2d-slabs, 1d-rods and isolated molecules):
3d-periodic codes always allow the simulation of systems with lower dimensionality within a supercell.
Specified here is the capability for actual simulation within lower periodicity.

[1] Trough interface to MOPAC.

[2] Through CRYSCOR program.

[3] Through Ascalaph.

[4] Through interface to TINKER.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

1.3 THE NUCLEI DYNAMICS

The physical observable traditionally measure for reactive processes are thermal rate coefficients $k(T)$. Thermal rate coefficients $k(T)$ can be usually decomposed in state-to-state rate coefficients $k_{ij}(T)$

$$k(T) = \sum_{if} k_{if}(T) \tag{1.10}$$

with i being the initial and f the final states of the system and T its temperature. In bimolecular processes state-to-state rate coefficients can be evaluated by properly integrating the state-to-state cross section $\sigma_{ij}(E_{tr})$ over the translation energy E_{tr}.

$$k_{if} = \left(\frac{8}{\pi\mu k_B^3 T^3}\right)^{\frac{1}{2}} \int_0^\infty E_{tr}\sigma_{if}(E_{tr}) e^{-\frac{E_{tr}}{k_B T}} dE_{tr} \tag{1.11}$$

where k_B is the Boltzmann constant and μ the reduced mass of the collisional system. In turn, $\sigma_{ij}(E_{tr})$ can be evaluated by summing over all the contributing values of the total angular momentum quantum number J the properly weighed detailed partial (fixed J) reactive probabilities $P^J_{if}(E_{tr})$ as follows

$$\sigma_{ij}(E_{tr}) = \sum_{J=0}^{\infty} \frac{\pi}{k_{if}^2} (2J+1)\ P^J_{if}(E_{tr}) \tag{1.12}$$

$P^J_{if}(E_{tr})$ is equal to the square modulus of the corresponding scattering matrix $\boldsymbol{S}$.

The scattering $\boldsymbol{S}$ matrix elements, that contain the dynamical information we are looking for, are evaluated by considering that the nuclei Schrödinger equation given in 1.6 is a first order differential equation in t and its solution must have the form

$$\psi^n(\boldsymbol{R}, t) = \hat{U}(t, t_0)\psi^n(\boldsymbol{R}, t_0). \tag{1.13}$$

The operator $\hat{U}(t, t_0)$ is called *time evolution operator*, or *wave operator*. It is a linear operator, in accordance with the principle of superposition, and satisfies the initial condition $\hat{U}(t_0, t_0) = \hat{1}$.

If the system is not subject to external forces (that is if the Hamilton operator is independent of time as we have already assumed before) the time

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

evolution operator reads

$$\hat{U}(t, t_0) = e^{\frac{-i\hat{H}\Delta t}{\hbar}}, \tag{1.14}$$

where $\Delta t = t - t_0$. From this, equation 1.13 can be reformulated as

$$\psi(\boldsymbol{R}, t) = e^{\frac{-i\hat{H}\Delta t}{\hbar}} \psi(\boldsymbol{R}, t_0). \tag{1.15}$$

Introducing Eq. 1.15 in the Schrödinger equation 1.6, we can drop the time dependence and get the time independent (TI) form of the Schrödinger equation:

$$\hat{H}\psi^n(\boldsymbol{R}, t_0) = E\psi^n(\boldsymbol{R}, t_0) \tag{1.16}$$

that is one dimension less than TD one.
This means that we can indifferently use two quantum techniques to solve the problem: the TD (*time dependent*) and the TI (*time independent*) ones. A key difference between the two approaches is that in the first case one solves equations like 1.6 using time as the continuity variable whereas, in the second case, one solves equations like 1.16 for which a new continuity variable has to be worked out. Another difference is that in TI methods ψ^n of reactants is propagated from the reactant to the product region (or from an intermediate point towards both reactant and product region) and then by comparison with the unperturbed solutions the $\boldsymbol{S}$ matrix is worked out. On the contrary in the TD method the wavepacket, initially shaped to describe the reactant system, is mapped at each time step on the product wavefunction and the time dependent coefficient of the projection are then Fourier transformed to evaluate the $\boldsymbol{S}$ matrix. Nowadays quantum calculations are routinely performed for atom diatom systems whose Hamiltonian is three dimensional and for the particular arrangement τ reads

$$\hat{H}_n = -\frac{\hbar^2}{2\mu_{3\tau}}\nabla^2_{R_\tau} - \frac{\hbar^2}{2\mu_{2\tau}}\nabla^2_{r_\tau} + V(R_\tau, r_\tau, \theta_\tau) \tag{1.17}$$

with $\mu_{3\tau}$ being the atom-diatom reduced mass and $\mu_{2\tau}$ the diatomic reduced mass. R and r are the Jacobi distances for the three body case and θ the angle formed by R and r as illustrated in Figure 1.1. For four body systems (like the diatom diatom ones) the Hamiltonian is six dimensional (see Chapter 4) and the computations becomes more difficult. Exact Quantum computations of larger systems are still out of reach even with presently

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

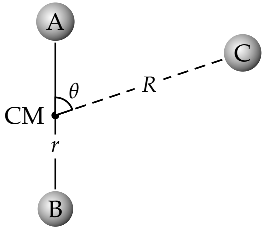

Figure 1.1: Sketch of the Jacobi internal coordinates for an atom diatom system.

available computers.

When the system is too large to afford using quantum dynamical means, classical trajectories calculations can be adopted in order to compute kinetic observables. These allow the study of the dynamics at microscopic level (differential and total cross sections, product energy distributions, etc.) as well as at macroscopic level (thermal rate constants) by a Monte Carlo integration of Eq. 1.11 in which the classical equations of motion

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \tag{1.18}$$

$$\dot{q}_i = +\frac{\partial H}{\partial p_i} \tag{1.19}$$

need to be integrated in time for each particle i (in the BO scheme only nuclei) for a large set of randomly generated initial position q and momenta p.
In order to run the trajectories, the potential energy surface needs to be supplied. To adequately sample the initial phase space (and therefore to get meaningful results) many trajectories (usually several thousand) have to be run.

1.4 THE CHEMICAL KINETICS SYSTEMS

The quantum or classical simulations of the kinetics of chemical systems do not necessarily require the knowledge of the PES over the all range

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

of configuration space. Thermal reaction rates, thanks to their largely averaged nature, can be also satisfactorily estimated using approximate kinetic models such as Transition State Theory (TST) [8], RRKM [9] and Master Equation [10] which require limited informations on the PES. The improvement of the efficiency of program devoted to the calculation of electronic structure and potential energy values combined with the progress in reaction rate theory has significantly enhanced the possibility of calculating rate coefficients with an accuracy comparable with experimental uncertainties. The precision of the kinetic data calculated in this way depends simultaneously on the accuracy of the potential energy surface (in particular regions of the phase space associated with reactive events) up to fractions of a $kcal\ mol^{-1}$, of the partition function, and of the rate theory level used to describe the reaction.

Transition State Theory

In the case of a temperature dependent rate coefficients based on a potential energy surface with a well-defined saddle point, TST is often sufficient to calculate the high-pressure limit rate coefficients. TST focus on the activation process assuming that thermal rate coefficients are determined only wby what happens at the saddle point with a substantial ininfluence what happens before or afterward. The transition state divides the phase space into a reactant region and a product region separated by a "dividing hypersurface" normal to the "reaction coordinate" (a coordinate connecting smoothly reactants to products in TST usually taken as a minimum energy path). TST also assumes that the motion along reaction coordinate can be separated from that on all other degrees of freedom.
The classical TST expression to evaluate thermal rate constants:

$$k(T) = \kappa\, \frac{k_b T}{h} \frac{Q_{\ddagger}}{Q_A Q_B}\, e^{-E_a/RT} \tag{1.20}$$

has been proposed by Eyring [11] in 1935 and applied with considerable success to a wide variety of reactive processes. In this equation $Q_{\ddagger}$ is the transition state partition function, Q_A and Q_B are respectively the reactants partition functions, κ is the transmission coefficient, E_a the activation energy and k_b, R and h the well known physical constants.
Variational version of TST (VTST) improves over it the accuracy of the kinetic parameters (for reactions with a defined barrier) by taking into ac-

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

count entropic effects. Moreover, its variational nature is appropriate for reactions with no defined saddle point on the PES, *i.e.* no reaction barrier.

RRKM / ME

The rate coefficient of a large number of reaction, such as the unimolecular ones, show a strong dependence also on pressure. For this type of reactions, different theoretical refinements are necessary to properly account for the pressure effects on the rate coefficients. Elementary reactions involving one isomer (single well) and one reaction pathway (single potential channel) were first described using RRKM theory. The fundamental idea that underlies this theory is that the degrees of freedom of a highly excited, isolated molecule or collision complex are so strongly coupled that the intramolecular distribution of the excess energy occurs faster than the characteristic time for reaction. The RRKM assumption is that the reactant of each elementary step is in microcanonical equilibrium; therefore one gets the same rates regardless of how energy is deposited in the various degree of freedom of a molecule (or complex). It is thus possible to define universally applicable elementary rate coefficient (for unimolecular elementary steps) $k(E)$ as:

$$k(E) = \frac{1}{h}\frac{N^{\ddagger}(E)}{\rho(E)} \tag{1.21}$$

where $N^{\ddagger}(E)$ is the sum of states of the transition state with energy lower than or equal to E and $\rho(E)$ is the reactant density of states per unit energy, bypassing entirely the problem of computing the intramolecular dynamics of the complexes.

For multiple-channel and multiple-well systems, more accurate Master Equation (ME) methods have to be envisaged to accurately describe the competition between reaction and collisional stabilization. These approaches constitute the state of the art for the description of pressure dependent multi-well and multi-channel system. ME simulations are based on the formulation of a set of integro-differential equations that describes the evolution in time of the concentration of each species. ME calculations require microcanonical rate constants $k(E)$ (computed from RRKM) and energy transfer parameters. It is useful to note that accurate predictions of pressure dependent rate coefficient obtained from ME/RRKM calculations still depend on semi-empirical factors such as collisional energy transfer.

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

The accuracy of the PES is another factor that critically influences the precision of the calculated rate parameters. For example the effect of the deviation of the vibrating rotating fragments from rigid rotor harmonic oscillator (RRHO) on the calculation of the partition function is crucial. Vibrational anharmonicities can affect the accuracy of the rate coefficients not only through the calculation of partition function but also through the tunneling transmission coefficient. A proper description of the low frequency motions (such as internal rotations) is therefore necessary to achieve good accuracy. Several methods have been developed to accurately calculate hindered rotor contributions [12, 13].

1.5 GEMS THE GRID SIMULATOR

At the core of the problems discussed in the previous section is computational efficiency and the heaviness of the associated computational campaigns.

A solution to this was found in the adoption of concurrent computing (and in particular Grid computing) that is the current paradigm for *high throughput computing* (HTC) as the ones of the present thesis. This has requested from our side to dig to some extent into the structure of the used computer codes and to reorganize the codes in a concurrent fashion. Our efforts for a concurrent reorganization of the programs calculating the properties of reactive systems has mainly focused on parameter sweeping approaches exploiting the fact that either the dynamical or the kinetic equations starting from different sets of initial conditions.

Accordingly, the computation is articulated as a coarse grained uncoupled loop over total or internal energies (for quantum dynamics), coordinates and momenta (for classical trajectories) or temperature and pressure for kinetic calculations. To this end, the various components of our treatment were gathered together in a single workflow implemented on the Grid (from this the name GEMS (Grid Empowered Molecular Simulator)) and a procedure able to handle concurrently large sets of jobs was developed. Each job execution requires the sending to the Grid of an execution script, of a specific input file and of the program adopted for the calculation. The execution script is the same for all jobs while the input file is different for each job. In order to better cope with the heterogeneous nature of both the computing hardware and software (compilers, libraries, submission systems, etc.) of the Grid, executable rather than source programs were

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

distributed over the net. In fact, in spite of the fact that the time required for sending the source code is considerably shorter than that required for sending its executable (such procedure is also more selective in terms of the type of machine to adopt) this approach exploits the fact that there is no need for identifying the compiler of each machine, selecting the optimal options for compilation, compiling the code and verifying that all runs give exactly the same results as the ones obtained on the original machine. The net result is (including networking time) a speed up factor equal to about half the number of processors used.

However, as we shall discuss in detail in the next section, the power of Grid computing goes much further than just distributing jobs execution and gaining a very large speed up. GEMS has in fact enormously increased our capability to simulate chemical processes at the molecular level and to leap forward in modeling natural phenomena, designing new materials and products, mastering new technologies and carrying out innovative experiments. This progress typically requires the assemblage of various pieces of software, the convergence of the competences of different experts, the concurrence of the elaboration on several processors. The GEMS ability of virtually gathering to coperate in an easy way all the necessary hardware, software and human resources is its true added value
To exploit this we have established COMPCHEM [1] a virtual organization (VO) that gathers together the expertise necessary to GEMS. GEMS is made, in fact, of three computational blocks: Interaction (evaluate the electronic interaction among the atoms involved), Dynamics (solve the equations of motion of the nuclei), Observables (perform the averaging over multiple valued parameters to compose theoretical estimates of observable properties). As shown in Figure 1.2 the related workflow has been generalized by articulating the procedure in the three modules:

- **Interaction** is the module that carries out full or partial evaluation, if not otherwise available, of the investigated system potential energy values. If needed also a suitable functional form is fitted to the *Ab Initio* (AI) single point potential energy values. As we shall discuss later, in order to facilitate interoperation among different quantum chemistry codes, Q5 (quantum chemistry) and D5 (dynamics)de facto standard formats have been developed to store data in related files (circles in the figure).

- **Dynamics** is the module that carries out the solutions of the equa-

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

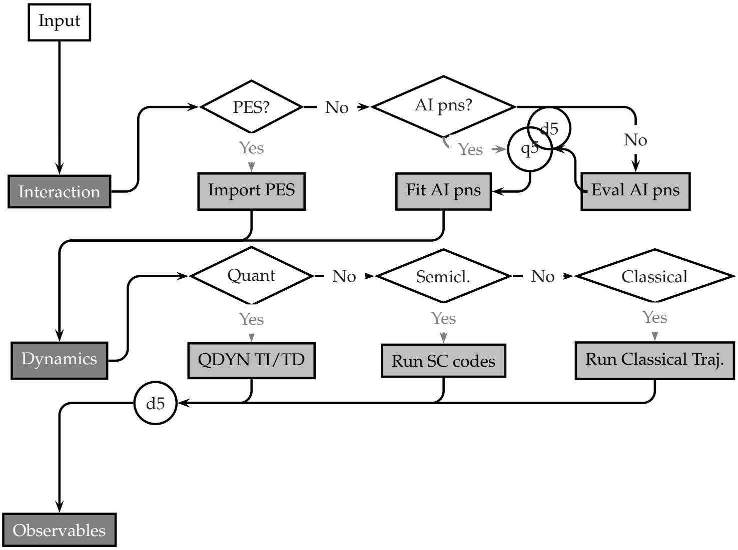

Figure 1.2: The workflow of GEMS for the a priori modeling of chemical reactions.

tions determining the dynamics of the nuclei of the system. In a rigorous approach the problem is dealt by using full dimensional quantum mechanics means which are the method of election for an exact calculation of chemical reactivity. However, except for small systems, reduced dimensionality quantum or classical or semi- (or quasi-) classical mechanics methods need to be used (or even a combination of them and model treatments).

- **Observable** is the module that carries out the necessary statistical (and model) treatments of the outcomes of the theoretical calculations to provide a priory estimates of the measured properties of the system. This implies a proper sampling of both initial and final conditions as well as an additional integration over some unobserved variables.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

1.6 THE KINETICS VERSION OF THE WORKFLOW

As already mentioned in the previous section, when dynamic calculations become unfeasible the kinetic thermal rate coefficients can be worked out using the previously mentioned model treatments. In Figure 1.3 the version of the workflow modified for kinetic studies of a general reactive system is shown.

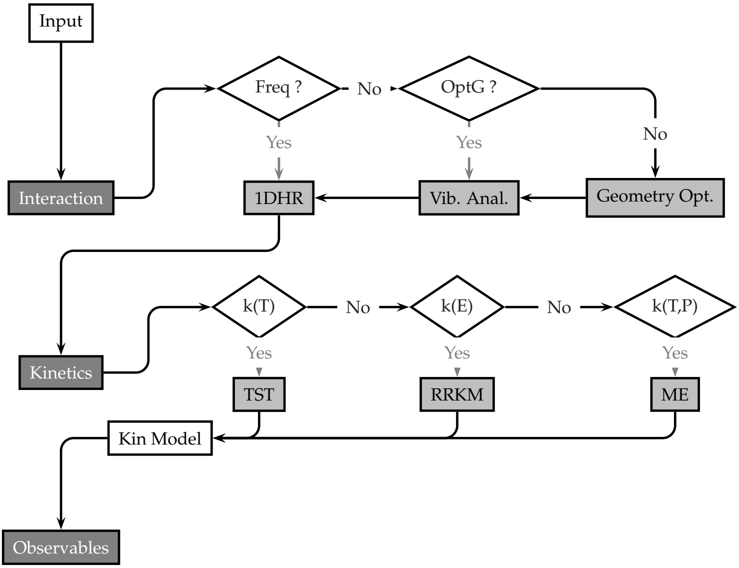

Figure 1.3: The workflow of GEMS for the calculation of the kinetic parameter of chemical reactions.

As apparent from Fig. 1.3 the main stream of the workflow remains almost unaltered with the division into three computational blocks. However the path followed inside each module is quite different. As already mentioned, in fact, generally applicable methods for kinetics, unlike detailed dynamics, do not require the knowledge of the whole PES. They only use the stationary points lying along the reaction coordinate. Yet, in the Interaction block, for each stationary point optimized geometries, vi-

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

brational frequencies and partition functions need to be calculated with high accuracy. Once these informations have been collected, the thermal rate coefficients can be worked out, within the Kinetics block, with different kinetics methods depending on the informations needed the model adopted.

1.7 BRIDGING HTC TO HPC

The development of GEMS requires not only the cooperation between complementary expertise and softwares but it has also prompted the coordinated usage of HPC and HTC computational platforms to run packages whose performance vary as a function of the calculation parameters The advent of large grids made of off-the-shelf High Throughput Computing (HTC) elements has in fact provided scientists with increasing processing power based on small memories and moderately fast architectures. However several computational chemistry programs make heavy usage of HPC resources. Yet HPC and HTC platforms pursue different strategies in supplying computing capabilities. Namely:

- HPC networks (like PRACE [14] in Europe and TeraGrid [15] in the United States) offer a suitable distributed infrastructure for tightly coupled calculations requiring large memory sizes, MPI libraries [16], a high speed interconnection network, large and high throughput storage devices. The interconnection of the different supercomputers is mainly intended in this case for facilitating the job management and offering unified resources (like storage).
- HTC grids (like that of the European Grid Infrastructure EGI) offer a highly cost-effective computational platform exploiting the concurrent elaboration of a huge number of small-middle sized computers most often of the rackable CPU cluster type. The typical job exploiting the advantage of this infrastructure consists of a huge amount of substantially uncoupled computational tasks that can be distributed on independent CPUs.

These HPC and HTC computing paradigms have developed in the past separately, and sometimes even conflictingly, to enable two totally different approaches to concurrent computing. As a consequence, they are

ISSN: 2279-8773

also managed by middlewares quite different in nature and target different classes of applications, numerical algorithms and computational methods. Yet, from the researchers' point of view, it is difficult to cooperatively handle the different HPC and HTC infrastructures even if many scientific fields and researchers would greatly benefit from combining the two paradigms together. This will, in fact, enable the accurate modeling of real-like systems and the assemblage of virtual reality simulations based on multi-scale and multi-physics approaches.
For these reasons, COMPCHEM, in the last years, has been engaged in designing and implementing accurate realistic multi-scale GEMS applications, involving the coordinated utilization of HPC and HTC resources. This is meant to overcome the shortcomings of the present unsatisfactory situation whereby none of the available resources is suited for complex requests, in terms of the diversity of requested computations, as is the case of GEMS. What is lacking, in fact, is the coordination of different systems, such as HPC and HTC, to interoperate via a single Workflow Management System that properly isolates into known patterns the user Workflow and dispatches the associated building blocks to the most suitable resource. This would allow a more efficient execution of complex applications.
This is, indeed, the main work of GriF (Grid Framework) [17, 18] that is a Workflow Management System (WfMS) designed in our laboratory to support users of the COMPCHEM VO, mainly, in optimizing the return of their parameter sweeping studies submitted to the Grid, to split the Workflow execution into single blocks whose execution is most suited either to HPC or HTC platforms. The elaboration of a project for implementing such a High Performance Grid (called HIPEG) has requested a tight collaboration among experts of IGI (the Italian Grid Initiative) [19] for what concerns the gLite middleware, of CINECA (the largest Supercomputing Center of Italy) [20] for the UNICORE software stack and of COMPCHEM for the structuring of the use-case.

As a prototype HIPEG application we have considered the implementation of the first block of GEMS devoted to the *ab initio* calculation of the electronic structure of molecular systems (in our case we considered a high level few atom one). Quantum detailed calculations of this type require, in fact, nodes (or clusters of nodes) equipped with large (of the order of many GB) physical or virtual memories and processors performing at the level of several Gigaflops to work out the whole Potential Energy Surface.

VIRT&L-COMM.9.2016.4

Moreover, in order to work out a global PES, we need to repeat single point calculations for many geometries of the system. To this end use has been made of the package GAMESS (US) [21] and GriF has been structured so as to allow the launch of jobs on both the Grid and the IBM SP6 machine of CINECA.

In order to distribute grid-oriented GAMESS jobs on a HPC platform, monitor their evolution and gather the results back to the user, the whole package was implemented as a service offered by the SP6 machine of CINECA that makes use of MPI and OpenMP libraries. Furthermore, GriF has been implemented as a modular structure so as to allow, in the near future, th utilization the EMI [22] Execution Service [23], presently under development, that will offer a new, common interface for job submission, specifically conceived with the intent of encompassing both HTC and HPC environments.

In addition to the user benefits, on the resource providers side there are also good reasons for coordinating the use of HPC and HTC infrastructures and interoperate large computational applications allowing an optimization of the different computing resources usage. It is, in fact, not infrequent the case in which a user utilizes HPC platforms not as such but as a bunch of loosely coupled processors underutilizing their fast dedicated network. At the same time, HTC users may utilize massively distributed HTC platforms to solve tightly coupled computational tasks ending up by wasting a large amount of time in transferring data. A coordination of the two types of platforms to interoperate via a single Workflow and properly manage the various components on the most suitable of them, would instead allow a clever composition of complex applications that optimizes the use of the various computing resources and offers to the users the best level of performance. The prototype applications chosen by us to test these execution patterns (that we will call HTPC *skeletons*) typically consist of two COMPCHEM codes which combine the use of HPC and HTC facilities. These HTPC skeletons are, indeed, typical of several Computational Chemistry applications of COMPCHEM. Figure 1.7 shows the implementation of a general scientific workflow based on an execution flow distributing high computational demanding independent jobs (A*n*) on a HPC platform. The produced outputs are then collected and processed to become part of the input of a large number of subsequent independent tasks (B*n*) that can be submitted and monitored on the HTC infrastructure. Each job both in the HPC and HTC section can be further composed

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

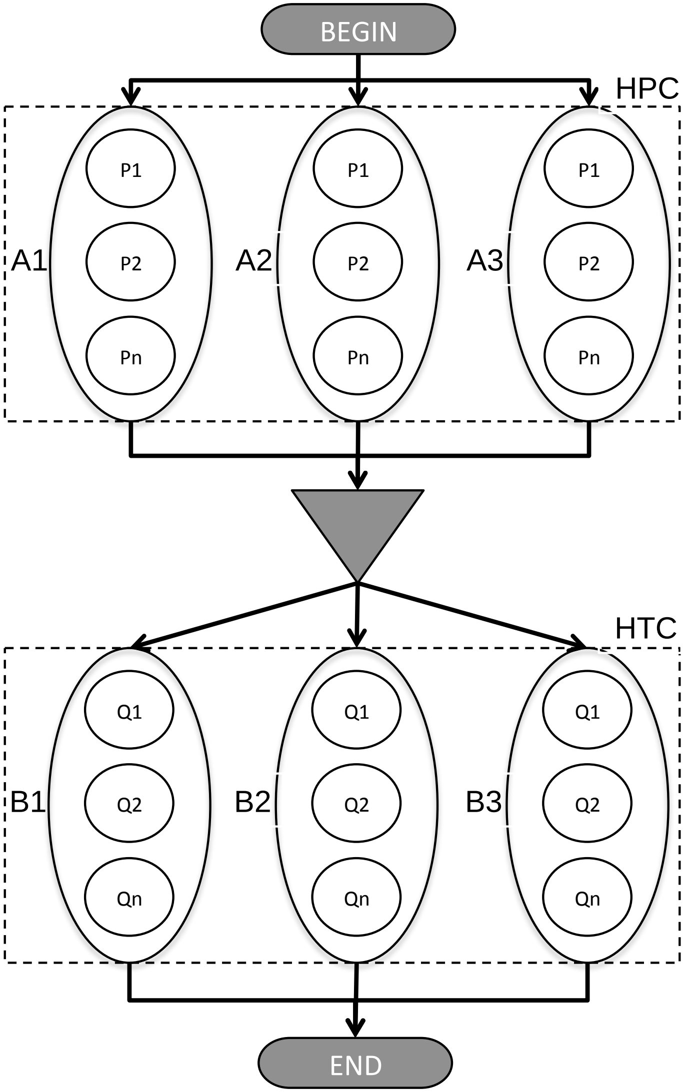

Figure 1.4: HPTC workflow for a n-HPC to m-HTC general HPTC *skeleton*.

ISSN: 2279–8773

VIRT&L-COMM.9.2016.4

by several sub-jobs which make the entire flow very complex and versatile. As can be seen, the workflow is of much more general validity than before said because its building blocks constitutes the basic components of many of the presently used computational applications developed not only in Chemistry and Materials Sciences but also in several other disciplines.

ISSN: 2279-8773

2

Code interoperability in Quantum chemistry and Dynamics

A fascinating problem (additional to that of bridging HTC and HPC resources) prompted by the cooperative nature of the grid, especially for complex workflows interconnecting programs belonging to different domain of chemistry, is data interoperability. The possibility for user applications to run on the different machines of heterogeneous and distributed platforms and utilize data coming from different packages and programs in a cooperative fashion (so that the outcome of a program can be straightforwardly managed by other programs) is a difficult task to achieve.

More specifically, in GEMS the topic is of great interest not only for both the purely theoretical modules (INTERACTION and DYNAMICS) but also for the one dealing more directly with the assemblage of the virtual signal (OBSERVABLES). In GEMS, in fact, code sharing is used both for developing and testing new theories and hypotheses and for managing modern experiments. Despite that there is still a significant time lag into standardization with respect to other scientific domains, probably due to the particular nature of the data to be treaded, that is usually large in size and complex in structure.

Along this direction, significant efforts have been already spent, within the D37 COST action "GridChem" [24], for the development and implementation of the two common models for QC and QD called respectively Q5cost [25, 26, 27] and D5cost [28, 29, 30].

The proposed Q5cost and D5cost formats are extensible data models which represent the first move towards standard representations of data in Computational Chemistry and are a clear candidate to become a universal model for designing new data formats for inter- and intra- operability between

VIRT&L-COMM.9.2016.4

different chemical domains. Together with these models a library was also developed allowing the users to set up an interface program (wrapper) for the conversion of the information from the code-specific language to the Q5/D5 standard and vice versa.
As a result, these formats and libraries enhance data cooperation among different packages, specially for the large binary datasets peculiar of the field: Hamiltonian matrices, general operator elements (integrals), expansion coefficients for orbitals and determinants and so on.

2.1 Data of Ab initio Quantum Chemistry

In order to design a successful data format for code interoperability, it is necessary to take into account the scientific nature of the disciplines involved, the possible connections among different programs, and the data objects that can flow among them. As already mentioned we defined for QC and QD the Q5cost and D5cost data models respectively. These two models are interconnected with QD deriving from QC.
Ab initio computational tools are used to evaluate the electronic structure of a molecule through suitable QC methods. As already mentioned all the QC calculations considered in this work are performed in the framework of Born-Oppenheimer (BO) approximation. Several methods have been proposed and are currently used to obtain approximate or high level of theory solutions of Eq. 1.7. Correspondingly a large number of software packages have been developed for this purpose. Generally speaking, these codes start from the definition of a molecular species in terms of nature of atoms, geometry, and atomic basis set; then one- and two-electron integrals are computed and a preliminary Hartree-Fock or DFT calculation is usually performed in order to produce Molecular Orbitals and to get a first approximation of the energy (if needed, a post-HF treatment can follow to introduce electron correlation (static or dynamic)). Finally, using the computed wavefunction, general properties are evaluated.
For the sake of simplicity, we can divide available QC codes into general codes and in-house made codes, although the border between the two classes cannot be precisely set.

General codes, usually distributed under a license agreement, are standalone programs that do not necessitate preliminary data objects. They can start simply by the definition of the molecular system and compute the

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

wavefunction and different properties at a given level of theory. Since they are used outside the developers community, particular care is put in designing user interfaces. Due to software engineering considerations, they usually have a modular structure.

In-house made codes implement quite advanced or experimental post-HF treatments (Configuration Interaction, Perturbation Theory, etc.). Their use is often restricted to the group of developers, and they are under constant development. For this reasons the user interface is usually rough and seldom user-friendly. Software engineering techniques are not usually applied, the codes are smaller and much less modular. This type of codes are not stand-alone: they usually rely on general-purpose codes in order to pre compute input quantities required for the treatment (one- and two-electron molecular integrals, orbitals coefficients, etc.).

Clearly both kinds of codes could take a great advantage from data format standardization because in this way they can share information and produced data making interoperability much easier. Moreover the use of a standardized format will allow information produced by in-house made codes to become much more accessible, and at the same time will foster usability of such codes.

On the contrary, as a matter of fact each code has presently its own way of dealing with input, intermediate and final data with little or no care for portability and compatibility. The first step is, therefore, to define a common policy for the representation of the produced data. This is the key idea behind the Q5cost that for this purpose divides data into two classes according to their size (there is no sharp border though between this two classes): Small Data (SD) or Large Data (LD).

Small Data contain all the information necessary to define in a unique way the molecular system, as well as the expectation values for the relevant properties, total energies, orbital energies, occupation numbers, basis-set overlaps, Molecular Orbitals, etc. In principle, SD can be stored in a formatted -human-readable way, since the amount of information associated with them is quite limited. In some cases, QC codes organize this type of information in xml structures (as in the case of Molpro [31], for instance).

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Large Data are the (often huge) data produced by the calculation and worth to be stored for further use. They are preferably given as binary unformatted files because this is the most performing and effective way for writing and reading such data. For their very nature large binary data need not to be treated by xml-like models (too verbose) or stored into databases (too structured). Nevertheless, the adoption of a common format for this kind of data is crucial, since they are required for interoperability.

2.2 DATA OF QUANTUM DYNAMICS

Besides the accurate characterization of the chemical systems provided by the QC packages, QD programs take care of determining the details of their evolution in time. To investigate the connection between QC and QD data we considered the well established prototype atom-diatom case.

$$A + BC(v, j) \rightarrow AB(v', j') + C \tag{2.1}$$

in which an atom A collides with a molecule BC (being in the quantum rovibrational state v, j) yielding a new molecule AB (in the v', j' rovibrational state) and an isolated atom C or a new molecule AC (in the v'', j'' rovibrational state) and an isolated atom B. This process is also the simplest model of a reaction in which a chemical bond breaks and a new one is formed.

The most straightforward approaches to the description of the atom diatom reactions are the time dependent or time independent formulation of the equation of motion for the nuclei (see Equations 1.15 and 1.16) in which the central quantities that need to be stored are S matrix elements, whose square moduli are the detailed state to state reaction probabilities from which cross sections and rate coefficients can be evaluated. In the present study we mainly focus on Quantum Dynamics. However the Q5/D5cost model can be easily extended to other possible schemas. For example, thanks to the fact that the potential energy gradient is retained, this approach can be used also for on-the-fly dynamics, i.e. to perform *ab initio* calculations for each molecular geometry reached during dynamical calculations and directly derive from them the forces acting on the nuclei without pre-computing a large set of potential energy values and fitting them before starting dynamical calculations. Moreover, the extension to mixed Quantum Mechanics: Molecular Mechanics Born-Oppenheimer Dynamics

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

will be straightforward. The only difference with respect to the previous case being that nuclei are propagated following Newton equations of motion. The ab initio derived gradients needed for this approach are already among the information retained by the D5cost format.

The information to be passed from QC to QD are those on the potential energy values of all the relevant molecular geometries. These information are produced by the electronic structure ab initio packages and, if calculated on the proper grid and stored in a predetermined format, they can flow directly to the dynamics programs. It is however important to point out that QC data refers to a fixed geometry scheme, producing one output result for each different geometrical organization of a given molecular species. On the contrary, as already mentioned, dynamical calculations consider all possible geometries contributing to the evolution of the system. Accordingly, they need the detailed definition of the potential energy surface of the system under investigation. QD needs therefore, one comprehensive file collecting information from multiple QC calculations. Further information that must flow from QC to QD concerns the set of adopted coordinates, the gradients, the Hessians and the non adiabatic coupling matrices.
To channel such information from the QC to the QD domain, each QC calculation is associated with a unique atomic specification and geometry whose contents relevant to dynamics (atomic coordinates, energies, gradients, Hessians, non adiabatic coupling matrix elements) are stored in a Q5cost file. A set of Q5cost files, each related to different points for the same molecular species, are then packed into a single D5cost file, ready to be used as input by a fitting or a QD program.

2.3 THE DATA MODEL: Q5COST

Q5cost and D5cost are data models containing most of the quantities used and produced respectively by QC and QD calculations, that are important for interoperability. The main issue to face is that data types are not completely defined, models are continuously evolving and the same quantities can be represented in different ways. The data scheme has to be general, flexible and easy to upgrade and therefore the above mentioned information have been integrated and organized into a logical hierarchical schema optimizing containment.

ISSN: 2279-8773

For these reasons, HDF5 [32] was selected as the base technology for the data format assuring portability among different computing platforms and efficiency, both in space (good compression algorithms) and time (good I/O performance). Moreover the use of the HDF5 technology makes the format very flexible and extensible, so that the file can be easily designed in an incremental way. These hierarchical data models can be discussed in terms of container (group in HDF5 terminology), metadata (attributes) and data (data sets).

Data is the set of large binary quantities (such as integral, properties and wavefunctions) that can be stored using matrices with an arbitrary number of indices (rank-n array). These quantities scale aggressively with the size of the system and are normally accessed with a chunked approach (i.e., using a well-defined buffer to read a block of data).

Metadata is, instead, a set of simple and small pieces of data describing and better defining the data. These quantities represent well-known chemical entities like nuclear energy, molecular orbital labels, molecular symmetry and can be stored as scalars, vectors or matrices.
Domain is the container for quantities defined on common function basis. The role of the domain is to group together entities whose indices conceptually refer to the same kind of functions.

The Q5cost and D5cost data models are, therefore, common HDF5-based formats made of a collection of chemical relevant objects related within a hierarchical structure (see Fig. 2.1 and Fig. 2.2).

In the Q5cost model, a first (root) container, named **System**, has been defined. It represents the molecular system as defined by its structural data (chemical composition, spatial geometry and basis functions). All the related general metadata can be associated with this container. A system can contain several "domains". For the Q5cost data format the following domain have been defined:

- **AO (Atomic Orbitals)**: refers to the data defined on the AO basis: overlap matrix and Hamiltonian matrix elements, i.e. one- and two-electron integrals, in addition to the generic property integrals. With "generic property" we refer to any other property that can be described on the AO basis by the application of a suitable operator (for instance dipole moment integrals, magnetic dipole, etc..). More than one AO domain is allowed, referring to different combinations of the basis set.

- **MO (Molecular Orbitals)**: refers to the data defined on the MO basis, one- and two-electron integrals and the generic property. This domain contains also the transformation matrix needed to define the MO on the AO basis. More than one MO domain is allowed, referring to different combinations of the AO basis. Relevant metadata such as canonical orbital energies or orbital occupation numbers are possible as well for every MO set.
- **WF (wavefunction)**: refers to the definition of the wavefunction, in terms of linear combination of determinants weighted by the proper coefficients. More than one WF domain is allowed, referring to different levels of theory and different choices of the MO basis.

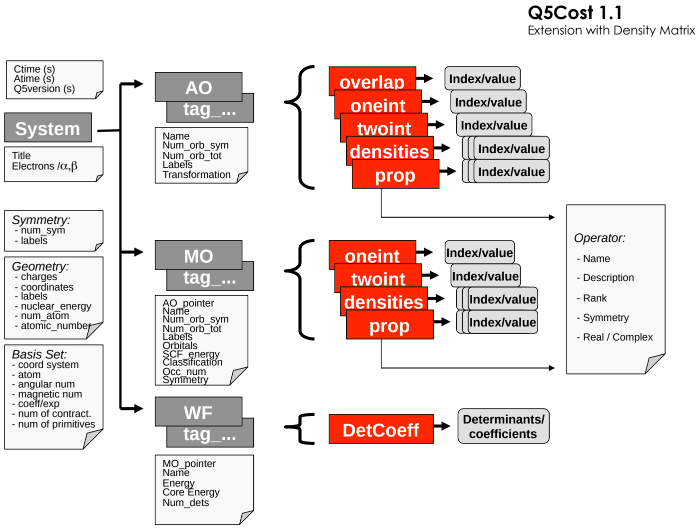

Figure 2.1: Functional schema of the Q5cost data model.

VIRT&L-COMM.9.2016.4

The latest version of the Q5cost (1.1) includes also the recent extension to "electronic density" entities (in particular the one-particle and the two-particle matrices, either on the MO or the AO basis), useful for linear scaling algorithms.

2.4 THE D5COST DATA MODEL

The D5cost format has been tailored on the model of Q5cost and maintains a similar structure. It is designed for Quantum Dynamical applications and therefore must present different domains. A System (root) container has been defined to store the information identifying the molecular system (number of atoms, nuclear charge etc.). No information on molecular coordinates are present in the root container, since the model allows multiple geometries in a single file. At present, the following domains have been included in the D5cost data model:

- **Interaction**: deals with the nuclear coordinates of a set of different nuclear geometries storing for each of them the Coulomb interaction as well as the potential energy values and related gradient and hessian matrices.
- **Dynamics**: deals the information computed by the QD algorithms (***S*** matrix elements, reaction probabilities and rate coefficients).
- **Observable**: deals with measurable dynamic and kinetic quantities.

2.5 THE LIBRARY

The Q5cost and D5cost libraries provide read and write access to files defined in accordance with the data models described before. By convention, the files have a .q5 and .d5 extension respectively. The two libraries are written in Fortran95, are based on the HDF5 library and are made of several hundred routines. The library is made of three layers, each one providing different facilities:

- **Q5core** is a low level module designed to provide a wrapping facility for HDF5 routines. It also provides simplified routines to perform frequently used low-level tasks. End users in general should not access Q5core routines.

ISSN: 2279-8773

- **Q5error** provides the routines needed for debugging and monitoring the behavior of the library and of the application code. A ring buffer is provided to keep track of error messages generated by the library.
- **D5/Q5error** defines the high-level application programming interface. It provides access to the Q5cost/D5cost format with a high level of abstraction. Accordingly, it is also the main reference for the final users.

The library D5cost relies upon the same low level routines (Q5core and Q5error) of the other library even if the user level (D5cost itself) has been completely rewritten.

The library routine names are chemically meaningful and closely recall their functions. The user interface devoted to the management of the data format is defined as follow:

```
Q5cost_<Domain>_<Action>_<Object>
D5cost_<Domain>_<Action>_<Object>
```

where `<Domain>` represent the conceptual domain inside which the routine acts, `<Actions>` specify the purpose of the routine (create, get, append etc.)

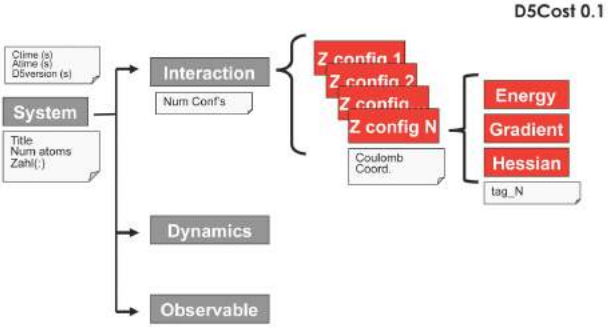

Figure 2.2: Functional schema of the D5cost data model.

VIRT&L-COMM.9.2016.4

and `<Object>` is the data object affected by the operation. A complete list of all the routines and how to call them, is embedded directly into the Fortran source code as comments using a custom tag system. The list can be extracted with a simple parser to HTML either for web presentation or for plain text.
The libraries can be directly used within a Fortran90 program and several bindings have been developed for Fortran77, C, C++ and Python which are automatically produced by a python script from the native Fortran95 library.

2.6 Tools and Wrappers

The libraries come with a number of tools and utilities, useful for both developers and final users. Among tools designed for developers it can be mentioned the already described tools for extracting documentation from the source files and for building bindings for languages different from Fortran90. As far as user tools are concerned, they mainly allow users to inspect the content of the Q5cost file. Indeed, even if the file is binary coded, it is possible to inspect the content thanks to the presence of metadata.

- **q5dump**: is a simple tool that reports the content of a Q5cost file by listing the data and the metadata contained in it.
- **q5edit**: is a more complete Python tool that allow the user to navigate through the file, list the content and modify the data using his/her favorite text editor.

Furthermore, since Q5cost files can contain relevant information on the molecular system (geometry, molecular orbitals, etc...), an useful application is related to the visualization of the molecular structure and of the densities iso-surfaces (as a visual description of the file itself). Indeed, different tools have been designed for the visualization of Q5cost files through both molden format files and cube files which can be easily read from specialized tools like Molekel [33], GaussView [34] or Avogadro [35] to produce visual rendering of orbitals and molecular structures as the example showed in Fig. 2.3.

In order to make a QC or QD program to interact with the Q5/D5cost data model, i.e. to read and write data in the common format, there are two possibilities:

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

1. create "wrappers", special programs using the Q5/D5cost libraries, and converting the data from/to the proprietary format and the common one;

2. call the Q5/D5cost routines directly into the program for managing the data I/O

Although the presence of wrappers requires a somehow heavier infrastructure somehow heavier than the use of direct calls, the first choice has the crucial advantage of permitting the evolution of the codes with little or no consequences on the wrapper side. At present, several QC/QD codes are interfaced with the Q5cost routines both through wrappers (Dalton [36], Molcas [37], GAMESS (US) [21], COLUMBUS [38], ABC [39], GFIT3C [40, 41]) and direct routine calls (Neptunus [42], QMC=CHEM [43]).

The structure of a wrapper: GAMCOST

For a given QC code, the output wrapper reads the output written in the code specific language and converts it into the common format file

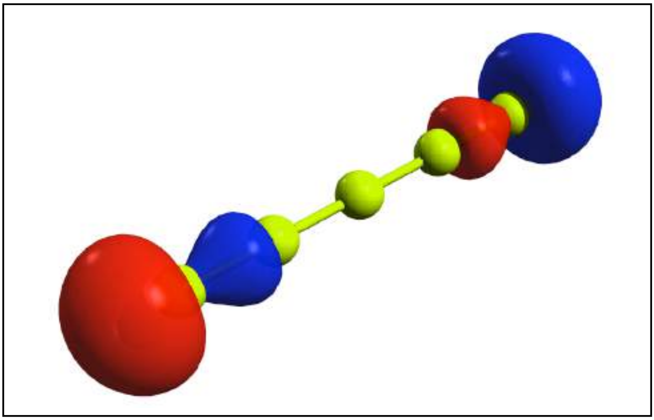

Figure 2.3: One of the two edge orbital iso-density surface of a linear Beryllium chain. The image is produced by the Avogadro program from a cube file obtained from a Q5cost file.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

whereas the input wrapper makes the conversion from the common format to the specific language of the QC code.

GAMCOST is a specific external output wrapper written in Fortran90 and working on the two output (".out" and ".dat") file produced by the GAMESS-US quantum chemistry code.
The program is articulated in four components:

- `GAMCOST.F90` is the main file. In this file all the calls to the various subroutines of the program are placed as well as to both the subroutines of the Q5cost library and the internal ones for reading the GAMESS outputs. Looking into this file a user can easily figure out what is written into Q5cost by looking at the different calls to the its library. The actual version of the program shows on the monitor the actions performed on the q5 file.
- `open.F90` is the file providing the reading of the GAMCOST input and the opening of the other needed files.
- `read_out.F90` is the file providing the reading of the ".out" file. It is a collection of subroutines that are invoked by the main file. Each of these subroutines reads the output file for the required property and passes it to the GAMCOST file which provides the writing into the q5 file. This file contains most of the useful information produced by GAMESS: geometry, energy, symmetries, etc.
- `read_dat.F90` is the file providing the reading of the “dat” file. As the previous one also this file is a collection of several subroutines devoted to the inspection of the dat file to work out the different information contained in it. The dat file stores the matrix of the molecular orbitals vectors and other quantities such as gradient and Hessian matrices and localized orbitals matrix.

The program is compiled by a Makefile which includes the compiler flags needed by the Q5cost library to work. After the compilation the user must insert into the "input" file the informations needed by the program and run the executable "gamcost". Here is an example of the GAMCOST input:

```
Q5 FILE NAME  =filename.q5
Q5 TITLE      =Title
```

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

```
GAMESS OUTPUT =outputname.out
GAMESS DATFILE=datfilename.dat

AO USER TAG   =aotag
MO USER TAG   =motag

NUMBER OF WF  =2
WF1 USER TAG  =wf1_tag
WF2 USER TAG  =wf2_tag
```

After these simple steps the user obtains a ".q5" file which contains all the relevant quantities produced by GAMESS and can easily use this file for the subsequent steps of his computational workflow.

Web site and SVN repository

A website for the dissemination of the Q5/D5cost data model has been designed and is available at the address: `http://q5cost.org/`. The most recent stable version (1.1) of the library is downloadable from the download section of the site where full documentation and a list of related publications are also available.
Moreover Trac [44] page and a Subversion [45] instances are available, for the project developers, at the address:

`https://hpc-forge.cineca.it/trac/Q5cost`

Trac provides an interface to subversion, an integrated Wiki and reporting facilities for managing the software. Moreover an efficient ticket system is provided in order to keep track of the issues and bugs within the project. The services are hosted by the software development infrastructure HPC-Forge at CINECA (`https://hpc-forge.cineca.it`).

2.7 PERFORMANCE AND BENCHMARKS

Due to the huge size of the *Large Data* stored in the Q5/D5cost files (mainly integrals and wavefunction coefficients) a careful optimization of the performances, both in term of disk space and access time, is crucial. Moreover the necessity for transferring files over the network in computational GRID applications makes this problem very important for computational

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Table 2.1: Times (in seconds) for writing $15,000,064$ integrals as a function of the chunk/buffer size. The total size of the two files is similar: 343 Mbytes (conventional file) and 346 Mbytes (Q5cost file).

No. of integrals: 15 000 064		
Buffer Size	**Time Fortran Binary (s)**	**Time Q5cost (s)**
1 024	265.23	226.62
2 048	121.13	114.53
4 096	62.38	59.02
8 192	34.39	31.46
16 384	18.86	17.04
32 768	8.56	6.09
131 072	6.19	4.86
262 144	5.84	4.08

intensive applications of such a data model. The choice of HDF5 as base technology let us to take advantage of its highly-optimized I/O features with the use of the embedded compression facilities provided by the zip libraries.

A number of simple benchmark tests were carried out in order to verify that the Q5/D5cost library infrastructure did not impose a significant overhead on the I/O operations. The tests consist in the production of a Q5cost file, with its internal structure, and a conventional Fortran file with unformatted sequential access. A varying number of two-electron integral (one-dimensional array of real numbers and a four-dimensional array of integer numbers) are then written on these files using a chunk of variable size and finally the time for the writing and the disk occupation for the two formats have been compared.

In Table 2.1 the the effect of the chunk size on the time required for writing a given large amount of integrals has been reported. As expected, an increase in the chunk size brings to a limitation in the I/O access to the file, resulting in a significant reduction of the overall writing time, both for the conventional and Q5cost files. The time for the Q5cost file is lower

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

than the time for the conventional file, regardless of the chunk size. This confirms the optimized approach of the HDF5 library in the management of large amount of data.

Writing times and disk occupation for an increasing number of integrals, by using a fixed value for the chunk/buffer size, have been reported in Table 2.2. The sizes of the two files are comparable with the conventional file being slightly smaller than the Q5cost file. The difference is mainly evident for smaller sizes and is about 1% of the total size. It can be rationalized in terms of the more complex structure of the HDF5 file containing also metadata as compared to a conventional Fortran binary file. The writing time is significant reduced by using the Q5cost library, and the difference becomes important as the file size increases. For instance, in the case of a 2.3 GB file the difference becomes about 18%.

These results were obtained without the use of the compression option of the HDF5. For these reason the dimension of the written files for the Q5cost case need to be intended as an upper limit to the real size. Using the HDF5 compression option, in fact, the disk occupation reduces by a factor from 3 to 10, depending on the system characteristics. Timings are also reduced since the overhead due to the compression activity is more than counterbalanced by the reduction in the I/O volume.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Table 2.2: Disk occupation and writing time (in seconds) as a function of the number of integrals. The chunk/buffers size is fixed.

Buffer size: 16,384 bytes				
	Q5cost		**Fortran**	
Num. Integral	**Size**	**Writing Time**	**Size**	**Writing Time**
16 384	397 KB	5.00×10^{-2}	384 KB	5.00×10^{-2}
65 536	1.5 MB	1.00×10^{-1}	1.5 MB	1.00×10^{-1}
114 688	2.7 MB	0.15	2.6 MB	0.17
507 904	12.0 MB	0.62	12 MB	0.68
1 015 808	23.0 MB	1.21	23 MB	1.37
5 013 504	115.0 MB	5.88	115 MB	6.41
10 010 624	231.0 MB	11.11	229 MB	12.12
50 003 968	1.1 GB	56.19	1.1 GB	64.21
100 007 936	2.34 GB	125.32	2.2 GB	148.53

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

3

High level *ab initio* calculations of small systems

As discussed in the first chapter the starting module of GEMS is INTERACTION that evaluates the Potential Energy Surface of the investigated system. The INTERACTION module offers a choice of *ab initio* packages devoted to the calculation not only of the microscopic properties of the molecular system but also of the kinetics and thermodynamics parameters. For our purpose high-level wavefunction based theoretical methods, such as CCSD(T) and Full CI, were used as will discussed in detail in the present Chapter for some study cases.

3.1 COUPLED-CLUSTER STUDY: THE ELECTRONIC WAVEFUNCTION OF TETRAHEDRAL CLUSTERS

Alkali and noble-metal tetrahedral clusters are interesting systems that exhibit particular and sometimes exotic electronic properties. Such systems are known to have bound states in which all the valence electrons have the same spin (${}^5A_2$), giving rise to the so called "no-paired" chemical bonds, a situation which seems to be in contrast with the common chemical bonding model.

For these reasons, these systems have been intensively studied by several groups using both density functional theory (DFT) [46, 47, 48, 49, 50] and wave-function based methods [51, 52, 53]. Moreover it seems that, at low temperature, these highest-spin cluster states are stabilized by helium droplets, with a very high total spin selectivity: in such conditions, they give rise to aggregates so stable that it was possible to determine experimentally many of their spectroscopic parameters.

In spite of the obvious interest for the quintet with tetrahedral geometry,

ISSN: 2279–8773

VIRT&L-COMM.9.2016.4

the other electronic states also deserve attention, in order to better understand the peculiarity of these systems. This is particularly true for the lowest lying state, bearing a triplet multiplicity. The global-minimum geometry of these systems is, in fact, characterized by a rhomboidal structure, with an electronic wavefunction having a singlet spin multiplicity. At a slightly higher energy we find a manifold of electronic states with tetrahedral geometries, which are characterized by different total spins culminating in the no-pair bonding quintet. The lowest-energy states within this manifold are characterized by triplet spin multiplicities, and are the main object of the present study. In these systems, because of the presence, in the undistorted tetrahedral geometry, of three degenerate-energy levels that host two electrons in the triplet state, the system undergoes a first-order Jahn-Teller distortion [54] that flattens the tetrahedron. To rationalize this effect, we performed high-quality Coupled-Cluster (CC) calculations on the Li_4, Na_4, K_4 and Cu_4 clusters. The electronic structure of the isolated atom is characterized by a closed-shell structure ([R] and $[R](n-1)d^{10}$, for Alkali and Noble-metal, respectively, where [R] indicates a rare-gas core), on top of which there is an ns^1 unpaired electron. Since the ns and np orbital energies are close, and much higher than those of the underlying electrons, the electronic structure of these clusters is characterized mainly by the behavior of the four valence electrons in the four ns orbitals. In this study we restricted the analysis to the lightest atomic species, up to the fourth period. Heavier atoms cannot be treated at a non-relativistic level. One has to take into account, in fact, spin-orbit effects (which, however, go beyond the scope of the present work). We performed geometry optimizations of the singlet (rhombus), triplet (distorted and undistorted tetrahedron) and quintet (tetrahedron) lowest states. Moreover, in order to get an insight on the nature of the calculated stationary points, we computed the frequencies of the optimized geometries for the singlet and triplet minima. Additional MRCI calculations on the D_{2d} optimized geometries were performed for different states and wavefunction symmetries.

The symmetry group of regular (*i.e.*, undistorted) tetrahedral X_4 is T_d. The lowest states are obtained by distributing the four valence electrons over the four valence ns orbitals of the four atoms, while the remaining doubly occupied orbitals form a totally symmetric core (A_1 symmetry), and do not contribute to the total symmetry. The four singly occupied ns atomic orbitals give rise to one a_1 molecular orbital ϕ_0, and a three-fold de-

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

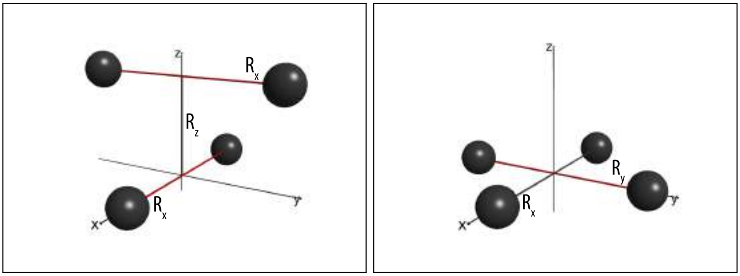

Figure 3.1: Representation of the structures for the singlet D_{2d} (left panel) and triplet D_{2h} (right panel) configurations.

generate t_2 set, ϕ_1, ϕ_2 and ϕ_3. Notice that these last orbitals have a slightly higher energy than ϕ_0. A total of six singlets, five triplets and one quintet states can be obtained from the four valence electrons in the $\{a_1, t_2\}$ valence orbitals set. We shall concentrate here on the lowest state, which is expected to be a triplet. In fact, by adopting a monoelectronic approximation, one can guess that the ground state of this system will be obtained by having the a_1 orbital doubly occupied, while two electrons are placed into two of the three t_2 orbitals. In this way, we have $(a_1)^2\,(t_2)^2$ occupation. Because of orbital degeneracy, the ground state of the system is expected to be a triplet. Three degenerate triplets can be built by distributing two electrons in the three degenerate orbitals, and the system will be Jahn-Teller distorted. A distortion from a T_d to D_{2d} structure will keep the degeneration of two out of the three originally degenerate orbitals. Since a triplet state has two singly occupied orbitals, this distortion pattern, lowering the energy of the two degenerate orbitals and lifting the energy level of the third one, is expected to be the most favorable one.

The type of distortion (oblate versus prolate tetrahedron) can be inferred by the following energy considerations. Let us consider the tetrahedron as composed of two interacting fragments (two dimers), placed at a variable distance R_z (see Figure 3.1). The orbitals of each fragment will be be indicated as σ_g and σ_u, accordingly to their (local) symmetry with respect to the dimer central point. The four valence Molecular Orbitals of the tetramer can be expressed as different combinations of the bonding and antibonding orbitals of the dimers as shown in Figure 3.2. In particu-

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

lar, the a_1 orbital ϕ_0 will result from the in-phase combination of the two bonding orbitals σ_g,

- $\phi_0 = \sigma_g(1) + \sigma_g(2)$.

The three t_2 orbitals, on the other hand, can be expressed as

- $\phi_1 = \sigma_g(1) - \sigma_g(2)$
- $\phi_2 = \sigma_u(1) + \sigma_u(2)$
- $\phi_3 = \sigma_u(1) - \sigma_u(2)$

(notice that, because of orbital degeneracy in the T_d symmetry, different expression are also possible for these last three orbitals). The last orbital ϕ_3 has an antibonding character with respect to the interaction between the two fragments. Let us suppose it is this orbital that is kept empty in the triplet wavefunction (the three t_2 being equivalent, this assumption is by no means restrictive). By "squeezing" the tetrahedron (*i.e.*, by reducing the value of d), the two fragments will become closer, and the energy of the empty orbital ϕ_3 will increase. As a consequence, the energy of the two remaining degenerates orbitals ϕ_1 and ϕ_2 will lower in energy. In a regular tetrahedron, the relation $R_z = R_x/\sqrt{2}$ holds, where R_x is the atom-atom distance in each dimer, while R_z is the distance between the middle points of the two dimers. In analogy with ellipsoids, a tetrahedron will be called "oblate" if R_z is smaller than the ideal value, and "prolate" if it is larger. The triplet wavefunction will be stabilized by a distortion toward an oblate tetrahedron. Indeed, our calculations invariably show a distortion of the tetrahedron at the energy minimum that *reduces* the value of d for the ground-state triplet wavefunction with respect to the regular tetrahedron.

So we can infer the correspondence between orbitals belonging to the T_d and D_{2d} point group, respectively:

- $a_1 \rightarrow a_1$
- $t_2 \rightarrow \{e, b_2\}$

Because of first-order Jahn-Teller distortion the three degenerate t_2 orbitals split in two degenerate e orbitals and one orbital of symmetry b_2, higher in

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

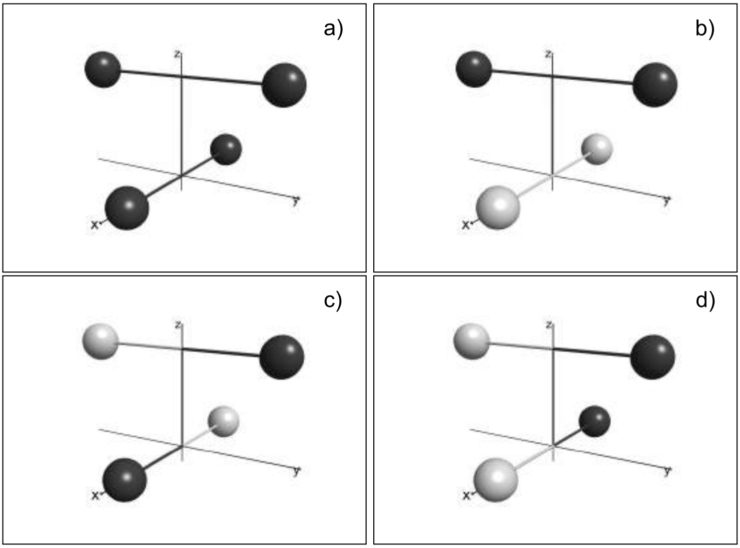

Figure 3.2: Schematic representation of the four valence orbitals ϕ_0 (a), ϕ_1 (b), ϕ_2 (c), ϕ_3 (d) expressed as combinations of the bonding and antibondig orbitals of the interacting fragments. The two colors stand for the positive (black) or negative (gray) phase.

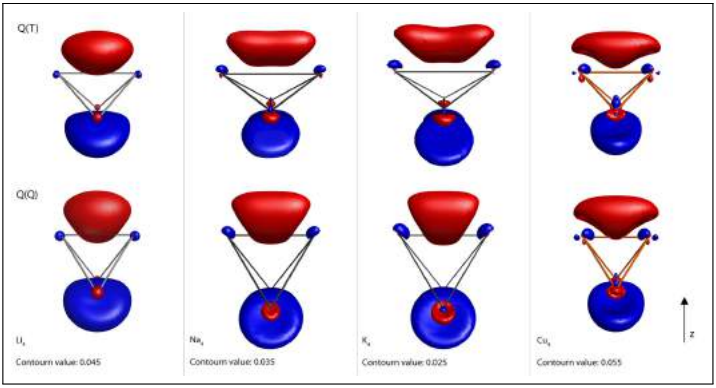

Figure 3.3: Contour plots of the ϕ_3 orbital for the four investigated clusters at the Q(T) (top) and Q(Q) (bottom) geometries. See text for more details.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

energy. Hence the distorted triplet wavefunction will have a ${}^3A_2$ symmetry.

This analysis is confirmed by the orbital energies of the triplet and the quintet states. They are reported in Table 3.1, for the four types of clusters and three different states/geometries: Q(Q) indicates the quintet at the quintet equilibrium geometry, while the notation T(T) is used for the triplet at the triplet equilibrium geometry (RHF wavefunction). However, the comparison of orbital energies for orbitals having a different occupation is somehow difficult at Hartree-Fock level. For this reason, we reported also the quintet energies at the triplet equilibrium energies, Q(T), to compare to Q(Q). From Table 3.1, it is seen that the energy of ϕ_3, ϵ_3 rises passing from Q(Q) to Q(T), while the energy of the degenerate pair ϕ_1 and ϕ_2, $\epsilon_{1,2}$ lowers. The energy of ϕ_0, ϵ_0 also decrease, because the bonding character of this orbital is more pronounced at the triplet geometry than at the quintet one.

The shift of the ϕ_3 orbital is also confirmed by the plotted isosurfaces in Figure 3.3. As we can see, in fact, passing from Q(Q) to Q(T) we have not only a rising of the energy but also a change of the orbital shape which lengthens and goes up along the z-axis. In Table 3.2, the oblateness index, defined as $\frac{R_x}{R_z\sqrt{2}}$, is reported for the triplet equilibrium geometries of the four clusters. Lithium turns out to be the less distorted cluster, while Sodium is the more distorted.

The *ab initio* code MOLPRO is restricted to use Abelian point groups only. Therefore, in the case of tetrahedron, all our computations were performed by assuming a C_{2v} symmetry on the electronic wavefunction, while the full D_{2d} symmetry was imposed on the nuclei. The true D_{2h} group, on the other hand, was used for the rhombus.

The Vibrational Normal Modes

The internal vibrations of the tetrahedral cluster, once the three translations and three rotations have been projected out, give six internal modes that have to be classified according to the irreducible representations of the D_{2d} group. This gives

$$\Gamma_{vib} \; = \; 2A_1 + B_1 + B_2 + E$$

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Table 3.1: Orbital energies (hartree) of the valence orbitals for the different clusters and at different geometries and states (see text for details).

	Li			Na		
state	ϵ_0	$\epsilon_{1,2}$	ϵ_3	ϵ_0	$\epsilon_{1,2}$	ϵ_3
Q(Q)	-0.3098	-0.1727	-0.1727	-0.2502	-0.1659	-0.1659
Q(T)	-0.3283	-0.1866	-0.1685	-0.2820	-0.1768	-0.1348
T(T)	-0.2341	-0.1575	-0.0034	-0.1992	-0.1484	-0.0016
	K			Cu		
state	ϵ_0	$\epsilon_{1,2}$	ϵ_3	ϵ_0	$\epsilon_{1,2}$	ϵ_3
Q(Q)	-0.2006	-0.1344	-0.1344	-0.2207	-0.1416	-0.1416
Q(T)	-0.2207	-0.1416	-0.1146	-0.4130	-0.2382	-0.1789
T(T)	-0.1549	-0.1190	-0.0051	-0.3074	-0.1980	0.0001

Table 3.2: Oblateness index of the different clusters at the triplet equilibrium geometry (CCSD(T)). The regular tetrahedron has an index equal to one.

Li	Na	K	Cu
1.425	2.367	2.313	1.496

In Table 3.3, the harmonic internal frequencies are reported for the ^1A_g (rhombus) and $^3A_2$ states for the four different species.
We can then define a set of six internal coordinates, three radial and three angular, so that each normal mode can be easily associated to a particular motion. In this work we take the two distances R_x (in this section we distinguish R_{xa} from R_{xb}) and R_z as the radial coordinates whereas the two plain angles θ_a and θ_b (formed by R_z and the relative R_x vector) and the

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

dihedral angle Φ (formed by the planes containing the two θ angles) are the three angular coordinates. With this choice of internal coordinates each vibrational motion can be classified as follow:

- $\mathbf{A_1}$ (higher frequency): asymmetrical stretch of R_x with respect to R_z
- $\mathbf{A_1}$ (lower frequency): symmetrical stretch of R_x with respect to R_z
- $\mathbf{B_2}$: asymmetrical stretching of R_{xa} with respect to R_{xb} (R_z fixed)
- $\mathbf{B_1}$: bending of Φ angle
- **E**: bending of the angles θ_a and θ_b

3.2 GEOMETRY OPTIMIZATIONS OF THE TETRAHEDRAL CLUSTERS

For the above mentioned clusters, four states have been considered:

- **SINGLET** : $^1A_1$ global minimum
- **TRIPLET** (D_{2d}) : $^3A_2$ distorted (D_{2d}) tetrahedron minimum
- **TRIPLET** (T_d) : $^3A_2$ tetrahedral (T_d) state
- **QUINTET** : $^5A_2$ tetrahedral minimum

The calculation of the equilibrium geometries of both the singlet (rhombus) and triplet (tetrahedron) lowest states were computed at the Coupled-Cluster level of theory, with single and double excitations and perturbative correction for the effect of the triples (CCSD(T) [55]), using the MOLPRO 2010 [31] code. The optimization of both the triplet and singlet states are performed using ANO-RCC basis sets proposed by Roos et al [56, 57]. The bases were retrieved from the EMSL public database [58, 59]. The use of ANO basis allows to recover a larger fraction of the correlation energy, due to the more correct representation of the nodal structure. Moreover these are the largest basis sets available for all the four atoms investigated in this work Li, Na, K and Cu, thus permitting a homogeneous treatment of the different clusters. In order to analyze the dependence of the results on the basis-set size and to reduce the computational effort of our calculations we choose to work also with reduced versions of the above mentioned

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Table 3.3: Harmonic frequencies of the internal vibration (cm^{-1}) for the different clusters at the rhombus (top) and at distorted tetrahedral (bottom) optimized geometries.

Rhombus(D_{2h})

Li	Na	K	Cu
103.74 (B_{1u})	42.19 (B_{1u})	18.39 (B_{1u})	52.27 (B_{1u})
182.48 (B_{2u})	55.11 (B_{1g})	35.67 (B_{1g})	119.27 (B_{2u})
202.78 (B_{1g})	78.45 (B_{2u})	44.38 (B_{2u})	148.70 (B_{1g})
234.02 (A_g)	100.52 (A_g)	53.27 (A_g)	153.52 (A_g)
325.38 (B_{3u})	142.96 (B_{3u})	81.95 (B_{3u})	230.95 (B_{3u})
332.73 (A_g)	154.09 (A_g)	83.26 (A_g)	261.84 (A_g)

Distorted Tetrahedron (D_{2d})

Li	Na	K	Cu
139.68 (A_1)	29.83 (A_1)	19.32 (A_1)	40.84 (B_2)
163.80 (B_2)	32.25 (B_2)	23.47 (B_2)	45.25 (A_1)
246.43 (B_1)	106.40 (E)	63.77 (E)	176.97 (B_1)
252.54 (E)	119.17 (B_1)	65.44 (B_1)	189.58 (E)
320.68 (A_1)	129.69 (A_1)	72.99 (A_1)	238.79 (A_1)

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

basis. We have called this reduced basis sets ANO-x where x is the number of removed contractions for each shell. Among these sets the ANO-2 was found to be the best compromise between computational performance and quality of results (compared with the full ANO-RCC). The above mentioned set involved a total of 124 contractions for the Li_4, 216 contractions for Na_4, 204 contractions for K_4 and 420 contractions for Cu_4. Since the ANO-RCC basis set is optimized to include scalar relativistic effects, all the calculations were performed with the use of the Douglas-Kroll-Hess Hamiltonian.

For all the investigated clusters the four valence electrons and the outer closed shell for each atom have been correlated. Following this approach, only the $4s$ and $3d$ electrons for Copper, the $4s$ and the $3p$ electrons for Potassium, the $3s$ and the $2p$ electrons for Sodium, and all the electrons for Lithium have been allowed to fully correlate. The inner electrons were kept frozen at the corresponding self-consistent field (SCF) level. We performed full geometry optimizations for the D_{2h} rhombus singlet state and the D_{2d} distorted-tetrahedral triplet state. Our results are therefore expressed as a function of the two rhombus diagonals (R_x and R_y) for the singlet configuration and as a function of the interatomic distances of the two fragments composing the distorted tetrahedron, R_x, and their separation distance, R_z, for the triplet state (see Figure 3.1).

The second derivatives of the potential energy surface have been calculated at the CCSD(T) level for the singlet and distorted-triplet optimized minima. The hessian is computed numerically by finite differences with the use of above mentioned ANO-2 basis sets for Li_4, Na_4 and K_4 clusters whereas for the Cu_4 system, due to the high computational effort required, only frequencies calculations with the ANO-4 basis-set are available. All the computed harmonic frequencies being real the stationary points can be safely considered as minima. Moreover vibrational normal modes analysis is presented.

The low energy electronic spectra were also analyzed in detail, in order to confirm that the triplet is the lowest energy state at the D_{2d} symmetry, and to get a clearer insight on the electronic structure of the different clusters. For this reason, we also performed Multi Reference Configuration-Interaction (MRCI) calculations on the lowest part of the energy spectrum. An active space comprised of 4 electrons in 4 orbitals was used to obtain the reference Complete-Active-Space Self-Consistent-Field (CAS-SCF) wavefunction. Different combination of spin multiplicity and spatial sym-

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Table 3.4: MRCI energies (hartree) for distorted tetrahedral states.

State	**Li$_4$**	**Na$_4$**	**K$_4$**	**Cu$_4$**
$^3A_2$	-29.8205	-648.3222	-2405.8653	-6614.1580
$^1A_2$	-29.8042	-648.3090	-2405.8542	-6614.1414
$^3B_1$	-29.8060	-648.2915	-2405.8427	-6614.1275
$^1A_1$	-29.8109	-648.3132	-2405.8589	-6614.1398

Geometry is kept fixed at the CCSD(T) level.

metry have been calculated.
The result for the vertical excitation for different spin state and wavefunction symmetries at the distorted geometry, carried out at the MRCI level, confirm that the $^3A_2$ triplet is the lowest state at that geometry. The obtained energy values for the different states are shown in Table 3.4.

The main results for the CCSD(T) geometry optimizations of the various metal clusters are collected in Table 3.5. From an energetic point of view our CCSD(T) calculations show that, for all the investigated clusters, the Jahn-Teller distorted $^3A_2$ minima lie very close to the corresponding singlet ground state. At a CCSD(T) level, the energy difference between the singlet and triplet (D_{2d}) states is only of 0.19 eV for K$_4$, slightly higher in energy we find Na$_4$ and Li$_4$, with a difference respectively of 0.34 and 0.43 eV. The Cu$_4$ cluster shows the greatest energy gap, with 0.94 eV. In order to gain an insight into the nature of the Jahn-Teller effect for these particular clusters, we perform CCSD(T) calculations also for the tetrahedral triplet state. At T_d symmetry the investigated clusters present a triple degenerate triplet state. Although this geometry, because of the Jahn-Teller distortion, is not a minimum on the triplet potential energy surface, its analysis allows the evaluation of the energy lowering occurring when distorting the T_d symmetry into the D_{2d}. Comparison of the results for the two

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Table 3.5: X_4 CCSD(T) optimized geometries (see Figure 3.1) and relative energies for the singlet, triplet and quintet states.

	SINGLET			**TRIPLET (D_{2d})**		
	R_x (Bohr)	R_y (Bohr)	E (Hartree)	R_x (Bohr)	R_z (Bohr)	E (Hartree)
Li_4	10.130	5.002	-30.0012	6.214	3.082	-29.9848
Na_4	11.836	5.934	-649.2808	8.274	2.526	-649.2682
K_4	14.822	7.788	-2406.9061	10.540	3.222	-2406.8988
Cu_4	7.932	4.280	-6614.6929	5.158	2.438	-6614.6584

	QUINTET			**TRIPLET (T_d)**		
	R_x (Bohr)	R_z (Bohr)	E (Hartree)	R_x (Bohr)	R_z (Bohr)	E (Hartree)
Li_4	5.630	3.982	-29.9436	5.478	3.874	-29.9804
Na_4	7.312	5.170	-649.2275	6.622	4.682	-649.2596
K_4	9.360	6.626	-2406.8752	8.482	5.998	-2406.8924
Cu_4	4.704	3.326	-6614.5463	4.512	3.190	-6614.6516

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

triplet states reported in Table 3.5 shows that the Na_4 is the most energy-stabilized cluster, with a difference between the T_d and D_{2d} triplet states of 0.23 eV, while the Cu_4, K_4 and Li_4 show gaps of 0.19, 0.17 and 0.12 eV, respectively. Because of the different electronic structure of the four investigated atoms, the energy differences of the different clusters could not be directly compared. It is clear, in fact, that the presence of the 3*d* orbitals in the Copper cluster, for example, spaces out its energy levels and makes all the energy differences much greater than those of the alkali-metals clusters. For this reason, we decided to analyze the position of the two triplet states with respect to the singlet (ground state, lower limit) and the quintet (highest-energy tetrahedral state, upper limit).
In Figure 3.4, we report the relative positions of the D_{2d} and T_d triplet minima (respectively red and dashed lines) for the four metal clusters. The represented energy levels are not in scale, but the relative distances from the two (upper and lower) limits are respected. As one can see from that picture, all the four distorted triplets lie at about 25%, of the respective quintet-singlet energy gap (Q-S gap). On the other hand, the position of the tetrahedral undistorted triplet depends on the nature of the atoms involved and vary from that of the K_4 (45% of the relative Q-S gap) to the lowest Cu_4 triplet state (28% of the relative Q-S gap). From this point of view the clusters with the greater stability gain seems to be K_4, which halves its distances from the singlet state, and Na_4 for which we have the strongest energy stabilization.
On the other hand, for Cu_4, despite an energy lowering greater in magnitude than those obtained by Potassium, the Jahn-Teller distortion has a weaker effect on the structure because of the large separation of its electronic states. Different considerations should be made on Lithium cluster which, despite the Q-S gap very similar to the Na_4, is weakly stabilized by the distortion.

A further insight into the Jahn-Teller effect can be made analyzing the structural parameters reported in Table 3.5. Accordingly to the symmetry considerations reported in Section 2, in order to remove the degeneracy the system undergoes a distortion from the T_d symmetry into a D_{2d} geometry. If we compare the geometries of the optimized triplet states we can see that the distances of the two interacting fragments (R_z) become shorter moving from the regular to the distorted tetrahedron, at the contrary the distances of the atoms composing the dimers (R_x) are increased. These

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

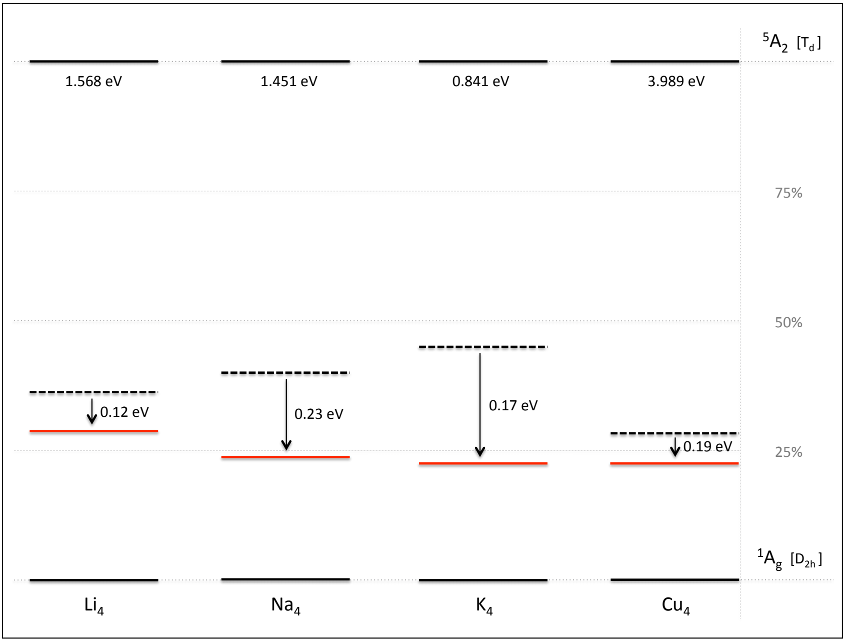

Figure 3.4: Energy levels diagram for the four X_4 clusters. The diagram shows the relative positions, with respect to the singlet (lower limit) and quintet (upper limit) states, for both the distorted (red line) and regular (dashed line) tetrahedral triplet state. The Q-S energy gap and the Jahn-Teller energy lowering have been reported.

results again confirm the hypothesis that the triplet wavefunction is stabilized by a distortion toward an oblate tetrahedron.
The effects on the geometry for the different clusters can be compared by examining the "oblateness indexes" reported in Table 3.2. These values show that, as for the energy considerations made above, the Na_4 and the K_4 clusters are the most affected by the distortion with an index of about 2.3 (the regular tetrahedron having an index of 1). On the other hand, the Li_4 and Cu_4 clusters show a weaker flattening with an index of 1.4 and 1.5 respectively. Additional important considerations should be made by looking at the optimized geometries of the two regular tetrahedral states (triplet and quintet). As shown in Figure 3.4, and explained above, at the

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

T_d symmetry the triplet states for Cu_4 and Li_4 are most stable with respect to the Q-S gap. By analyzing the geometries of these states, we can see that the R_z and R_x structural parameters are very similar to that of the respective tetrahedral quintets. On the other hand, Potassium and Sodium clusters, whose triplet is the closest in energy to the corresponding quintet state, vary their geometries considerably by passing from the quintet to the tetrahedral triplet state (K_4 in particular). This characteristic can be linked directly to the Jahn-Teller effect. As explained above, indeed, these last two clusters are the most affected by the distortion in regards to both the energy stabilization and the tetrahedron crushing.
At the light of the above considerations it seems that these clusters show less affinity for the tetrahedral configuration (the corresponding X_4 cluster being the less stable) and when they are submitted to a driving force (as in the case of the Jahn-Teller distortion) they try to keep two of their atoms closer together.

3.3 NO-PAIR BONDING: CU_4 BASIS SETS AND CORRELATED ORBITALS

No-pair bonding is a situation that strongly defies the common picture of the chemical bond, usually drawn since Lewis time. Such a situation is characterized by a system that shows a bound state although all the valence electrons are characterized by having parallel spins. It is, therefore, easy to see how such highest multiplicity systems represent a very interesting and challenging intellectual problem, in particular to clarify the nature of the bonding scheme.
As already mentioned in the previous section, alkali and noble metal clusters have been know to exhibit a very important and strong no pair bond. In particular, no-pair bonding strength is very dependent on the number of atoms in the cluster, as well as on the geometry and the atomic species composing the cluster. In this section an high-level *ab initio* study of the behavior of the tetrahedral Cu_4 quintet no-pair bounded state is reported. The four-atom cluster and the tetrahedral geometry have been chosen since in the case of alkali metals this arrangement gives the most stable high spin state, a result which was also confirmed at DFT level in the case of copper by Shaik and coworkers [60].

In order to to have an insight of the no-pair bond we have performed a systematic study of the quintet state using different basis sets for two different set of correlated orbitals. The potential energy curve for the sym-

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

metric expansion of the Cu_4 cluster (tetrahedral geometry) was computed at CCSD(T) with three different basis sets obtained from the ANO-RCC proposed by Roos *et al.*. This let us, indeed, to reduce the computational effort of our calculations and, in particular, to study the dependence of the results on the basis-set size. In the present work we will call ANO-4 the smaller set, ANO-3 the medium, and ANO-2 the bigger one. These three sets are composed as follows:

- ANO-4: (21s,15p,10d,6f)/[6s,5p,4d,2f]
- ANO-3: (21s,15p,10d,6f,4g)/[7s,6p,5d,3f,1g]
- ANO-2: (21s,15p,10d,6f,4g)/[8s,7p,6d,4f,2g]

In the case of Cu4 these basis sets involve a total of 220 contractions for the ANO-4, 320 for the ANO-3 and 420 for the ANO-2. Since the ANO-RCC basis set is optimized to include scalar relativistic effects, all the calculations were performed with the use of the Douglas-Kroll-Hess Hamiltonian. In order to ensure a more complete description of the bond the potential energy curves have been determined with two different choices of the correlated orbitals: a large core (LC) and a small core (SC). In the case of the LC calculations, only the four valence electrons have been correlated, while all the others were frozen at the corresponding self-consistent field (SCF) level. In the case of SC, on the other hand, we correlated also the five 3d orbitals of each metal atom. Symmetry-constrained geometry optimizations have also been performed in both cases. In order to define the role of the n = 3 shell orbitals in a better way (n being the principal quantum number) into the 'no-pair' bonding formation, we also performed a geometry optimization with an even smaller core (in the following very small core (VSC)). This was obtained by correlating also the 3s and the 3p orbitals of the copper atoms. Notice that in the latter case only geometry optimizations and energy differences were performed, without the explicit computation of the complete dissociation curve, due to the extremely high computational effort that would have been required.
The basis-set superposition error (BSSE) [61] effect was evaluated by using the standard counterpoise correction (CP correction) proposed by Boys and Bernardi [62]. BSSE corrected potential energy curves, in the region close to the minimum, have been interpolated to obtain BSSE corrected

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

equilibrium geometries and the corresponding dissociation energies by using a cubic function of general equation

$$y = ax^3 + bx^2 + cx + d \tag{3.1}$$

These results are also compared with the geometrical parameters obtained from geometry optimization that, on the other hand, have not been corrected for the BSSE.

3.4 THE NO-PAIR BOND STATE

Contrary to the lithium cluster [51] and coherently to the other alkali-metal cluster [52] Cu_4 gives always a repulsive potential energy curve at the SCF level and therefore the inclusion of dynamical correlation is essential in order to observe a bounded quintet state.

The results of the CCSD(T) symmetry-constrained geometry optimizations (non BSSE-free) with the three different basis sets, as well as the BSSE corrected interpolations, are summarized in Table 3.6 for both the LC and SC. As can be seen from the reported dissociation energies, the choice to correlate only the four valence single-occupied *s* orbitals (LC) leads to a very swallow energy well of about 0.2 eV which underlines the inability of this core to correctly represent bond formation. For this reason one can safely conclude that the role played by, at least 3d orbitals appears to be extremely crucial in the formation of the bound state. BSSE effects, for this core, are very small and their magnitude become nearly zero for the ANO-2 basis set, which shows a a remarkable agreement between corrected and uncorrected parameters.

When one increases the number of correlated orbitals, by including also the 3*d* orbitals of each atom (SC), the dissociation energy becomes considerably larger. We can see, indeed, that the energy well becomes about seven times deeper than in the case of the LC and the equilibrium distance decreases of about 1.2 bohr. With this choice of the core the effect of triple correction, negligible for the LC calculations, become large and BSSE much more relevant. This is can be clearly inferred from the potential energy curves for the SC and ANO-2 basis set reported in Figure 3.5. In this figure the differences from the curves with and without BSSE correction are, indeed, very important even if the position of the minima are quite similar. As it was done with the former LC, the geometrical parameters

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Table 3.6: Dissociation energies and equilibrium internuclear distances for the LC and SC calculations. The results of the geometries optimizations are compared with those obtained from the relative BSSE corrected interpolated curves.

	LC		**INTERP**	
BASIS SET	ΔE (eV)	R_e (Bohr)	ΔE (eV)	R_e (Bohr)
ANO-4	-0.1764	5.954	-0.1551	5.962
ANO-3	-0.2055	5.911	-0.1986	5.917
ANO-2	-0.2146	5.898	-0.2122	5.903
	SC		**INTERP**	
BASIS SET	ΔE (eV)	R_e (Bohr)	ΔE (eV)	R_e (Bohr)
ANO-4	-1.3606	4.752	-1.0721	4.803
ANO-3	-1.4355	4.722	-1.2735	4.752
ANO-2	-1.5102	4.704	-1.4068	4.723

reported in Table 3.6 have been obtained by using a cubic interpolation of the BSSE corrected curve. The second derivatives of the potential energy surface have been computed at CSSD(T) level with all the basis set for LC and SC. All the harmonic frequencies being real, the stationary point can safely be considered as a minimum, thus confirming the bounded nature of the quintet state.

The geometrical parameters obtained, at CCSD(T) level, for the VSC are reported in Table 3.7. As already stated in section 3.3, because of the very high computational effort that would have been required, we did not explore the whole symmetric dissociation curve. Instead, we performed a direct minimization without correcting the results for the BSSE. Since the

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

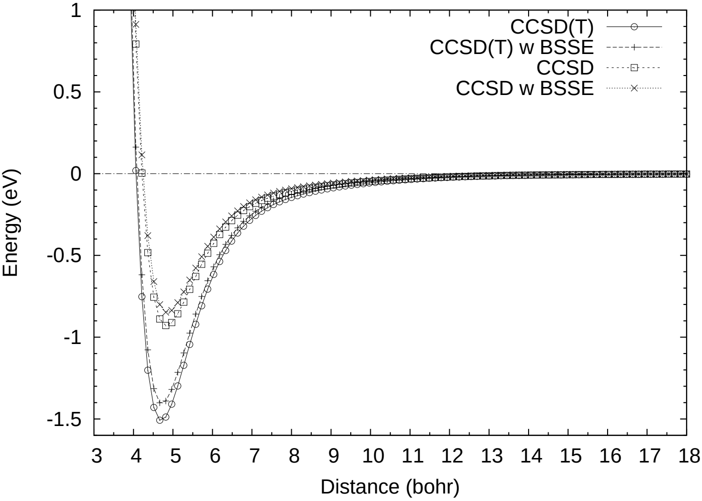

Figure 3.5: Cu_4 potential energy curves for the small core with the ANO-2 basis set. Both the CCSD and CCSD(T) results are reported. Distances are given in bohr and energies in eV.

influence of the BSSE error appears much less important on the position of the minimum than on the value of the energy well depth, we assumed that the obtained minimum distance could be considered as the correct one. The BSSE counterpoise correction was than applied to this fixed geometry only. As can be seen in Table 3.7, by reducing the size of the frozen core the BSSE error becomes quite important, and in the case of the smallest ANO-4 basis set it is so strong that before correction this basis gives the deepest energy well. After correction, the right behavior between the three basis set is restored. Moreover, it appears evident that the values of dissociation energy are in a very good agreement with those of the SC set: the bigger ANO-2 basis, for instance, gives a difference energy well depth of only about 65 meV between SC and VSC.

Finally on the ANO-2 optimized geometry we also performed a single point computation of the dissociation energy using the complete ANO-RCC basis set. This give a value of 1.63 meV which confirm the stability of

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

BASIS SET	ΔE (eV)	ΔE_{bsse} (eV)	R_e (Bohr)
ANO-4	-2.1562	-1.0114	4.684
ANO-3	-1.8074	-1.2062	4.694
ANO-2	-1.5958	-1.3422	4.694

Table 3.7: Geometry optimization results for the VSC core. The values of the dissociation energies and the equilibrium internuclear distances are reported for the three different basis sets.

the high-spin no-pair bonding state in the case of copper Cu_4.

In Figure 3.6 we have reported the isosurface density of the totally symmetric Cu_4 molecular orbital computed at restricted open shell Hartree-Fock level for the quintet no-pair bond state. As it was already observed in the case of lithium [51], the orbital seems to attract some electronic den-

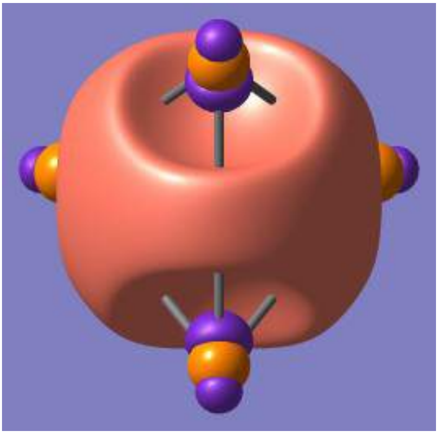

Figure 3.6: Density isosurface (for a cut-off value of 0.06) for the total symmetric Cu_4 valence orbital calculated at the ROHF level for the ANO-2 basis set.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

sity toward the center of the tetrahedron, in such a way that could explain the formation of the bond. Anyway, the analysis of the orbital alone is definitely not sufficient to completely characterize the bond formation but it gives confirmation of the nature of the effect that characterize this state.

3.5 THE BEH$^-$ ANION: A FULL CI INVESTIGATION

Beryllium is a four-electron atom having a $1s^2\,2s^2$ closed-shell electronic structure, a fact that explains why this species is not particularly reactive. This is related to a "negative" Electron Affinity (EA) (isolated Be$^-$ anions do not exist as stable species), and a large Ionization Potential (IP), larger than for both Lithium and Boron. However, it is well known that the Be atom is able to form two sp hybrid orbitals, that are placed in a linear arrangement at the opposite sides with respect to the Be atom. A system showing this feature is the BeH_2 molecule, often quoted in textbooks as a prototype for the sp molecular-orbital hybridization. This possibility is due to the quasi-degenerate nature of the $2s$ and $2p$ atomic orbitals in the isolated Be atom [63]. Indeed, the first excited triplet state (^3P) is only 2.725192 eV above the ground state, while the corresponding singlet (^1P) lies 5.277430 eV above the fundamental [64].

Experimentally, despite its popularity in textbooks, BeH_2 is a relatively poorly known system. In fact, the synthesis and accurate measurement of the isolate-molecule spectroscopic constants have been the object of a recent study [65]. At equilibrium, the Be-H bond length in this species is found to be 1.3324 Å, while from the computed energetics of the reaction $Be(g) + H_2(g) \rightarrow BeH_2(g)$ one can deduce a bond energy of 2.55 eV. Both values are very close to those of the BeH free radical, whose bond length is 1.3417 Å.

The bond in Beryllium dimer, whose existence as a relatively stable molecule has been a surprise among the chemistry community, is also a result of sp hybridization.
Another interesting case is the anionic species BeH$^-$, having also four valence electrons. Since the Beryllium atom is unable to accept an extra electron (accordingly to the previously reported negative EA), while the H$^-$ ion is a well known stable species (EA=0.75419 eV) [66], the BeH$^-$ anion will dissociate as $Be + H^-$. This means that, for large values of the inter-

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

atomic distance R, we are in the presence of the interaction between two closed-shell systems, one of which bearing a negative charge. The large-R situation is therefore very much similar to the structure of the HeH^- anion, which is a weakly bonded system with an interatomic potential going to the asymptote as R^{-4} (R being the inter-nuclear distance) [67]. At short distance, on the other hand, the Be atom is expected to form the two previously mentioned sp hybrid. One of them combines with the H 1s orbital to form a σ Be-H orbital. The two electrons located in this bonding orbital come (formally) one from Be and the other one from H. The other sp hybrid orbital, the one located at the opposite side with respect to the H atom, can host the second Be valence electron and the anionic negative charge (again, we stress the fact that this partition of the electrons is purely formal). If this scenario is correct, we can predict that, for the distances that give rise to the formation of sp hybrids, the Lewis structure of the anionic dimer will be $(:\text{Be-H})^-$, in which the negative charge in excess is formally located on the Beryllium atom. On the other hand, for large values of the interatomic distance R, we will have a Be...H^- weakly bonded system, with the negative charge located on the Hydrogen atom. Notice that the short-distance result is somehow surprising, in view of the large difference between the electronegativity of the two atoms (2.20 for Hydrogen and only 1.57 for Beryllium in the Pauling scale[68]).

3.6 THE FCI COMPUTATIONAL STRATEGY

In order to confirm the previously presented qualitative picture, we decided to perform a high-quality Full-Configuration-Interaction investigation on this system. The electronic energies for the BeH^- dimer have been evaluated at the Full Configuration Interaction level of theory using the NEPTUNUS code [42].

The Hartree-Fock preliminary steps for the evaluation of the one and two-electrons integrals have been carried out with the DALTON quantum chemistry package [36], while the following four-index transformation has been performed with the Ferrara transformation code [69]. The interface between the two codes as well as with the FCI code has been realized by the Q5Cost data format and libraries (see Chapter 2).

The FCI calculations presented in this work have been performed by keeping the 1s orbital on the beryllium atom doubly occupied and frozen at the

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Table 3.8: Total number of basis functions and FCI space size for the five employed basis sets.

Basis Set	Contractions Be	Contractions H	Total Number of functions	No. of FCI determinants
D	$7s6p5d$	$6s5p$	71	$1,521,853$
F	$7s6p5d4f$	$6s5p4d$	119	$12,086,733$
G	$7s6p5d4f3g$	$6s5p4d3f$	167	$47,208,513$
H	$7s6p5d4f3g2h$	$6s5p4d3f2g$	207	$111,918,625$
I	$7s6p5d4f3g2h1i$	$6s5p4d3f2g1h$	231	$173,923,093$

SCF level (fc-FCI), with the 4 valence electrons correlated in the FCI space without taking into account core-core and core-valence contributions. In fact, the contribution of core-related effects to the bond structure is expected to be very small. Relativistic effects and spin-orbit coupling has also been neglected, a fact justified because of the weakness of such effects for light atoms like beryllium and hydrogen. The size of the FCI space in terms of symmetry adapted determinants, as well as the total number of orbitals for each basis set, are reported in Table 3.8.

We perform a systematic study of the $X^1\Sigma_g^+$ state of BeH$^-$ using a cc-PV6Z basis set. In particular two different variants of this set has been adopted for the two atoms: an unpublished version, courtesy of Kirk Peterson [70] for the Be atom ($[16s9p5d4f3g2h1i]/[7s6p5d4f3g2h1i]$) and the version available at the EMSL public database [58, 59] for the H atom ([10s5p4d3f2g]/[6s5p4d3f2g1h]).
In order to extrapolate the dissociation energy of the dimer at the infinite basis set limit and to analyze the dependence of our results on the basis-set size we perform also a whole set of calculations with smaller basis sets. These were obtained from the complete set by taking off higher angular-momentum functions up to the small contraction $[7s6p5d]$ for Be and and $[6s5p]$ for H. In this article each basis set will be denoted hereinafter with the name of the largest angular function which is taking into account for

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

the Be atom (*i.e.* I, H, G, F and D). Details of each basis set adopted in this work are reported in Table 3.8.

The PES of the system was sampled at different values of the interatomic distance (R) up to the total dissociation, for the D, G and F basis-sets. In particular from R=1.0 to 2.0 bohr with a step size of 0.2, from 2.0 to 5.0 in step of 0.1, from 5.0 to 6.0 in step of 0.2, from 6.0 to 10.0 in step of 0.5, from 10 to 15 in step of 1 and finally from 20 to 100 with step of 20 bohr. For the two largest basis sets H and I only five points around the minimum, the R=100 asymptotic point and a few points in the region where the dipole momentum changes sign, have been calculated. The calculated potential energy curves, in the region close to the minima, have been finally interpolated for all the employed basis set in order to obtain the equilibrium geometries and dissociation energies using a cubic function.

The center of charge of the system was computed, in order to have an indication about the average position of the electrons. We also computed the FCI Natural Orbitals (NO), their occupation numbers, and the center of charge associated to each single orbital. In this way, we were able to decompose the total center of charge as a sum of contributions coming from each NO. The occupations of the NO's give a good indication of the character (single reference versus multi reference) of the wavefunction.

3.7 DISSOCIATION ENERGY AND EQUILIBRIUM DISTANCE

Following the computational strategy illustrated in the previous section, we were able to compute the dissociation energy (D_e) and equilibrium distance (R) of the dimer. Figure 3.7 shows the PES of the ground X Σ_g^+ electronic state of the BeH$^-$ dimer for the five adopted basis sets (only five points around the minimum have been computed for the H and I sets). The calculated equilibrium distance is around 2.7 bohr for all the employed basis set, with a dissociation energy varying from about. 2.05 (D basis set) to 2.12 eV c.a. (CBS extrapolated energy) confirming the formation of a fairly stable true chemical bond for short internuclear distances. As already mentioned, in order to estimate the contribution of higher angular momenta function in the basis set, we have extrapolated the dissociation energies to the infinite basis-set limit with the interpolation expression:

$$D_e(l) = D_\infty + A\, l^{-\beta} \tag{3.2}$$

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

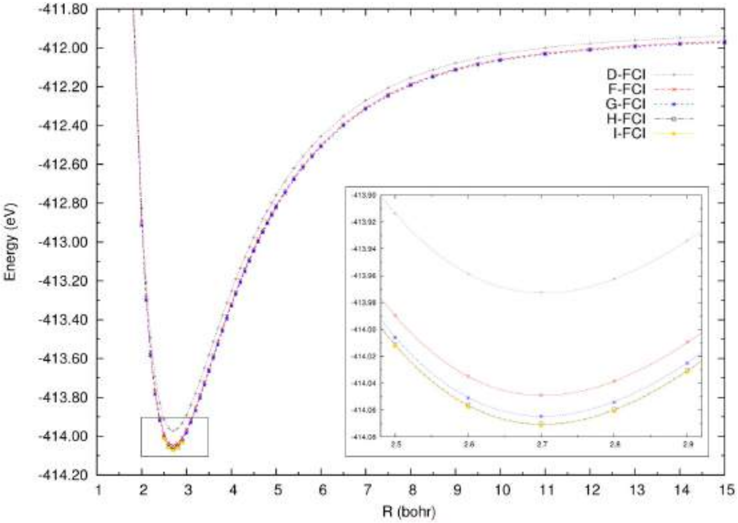

Figure 3.7: Potential energy curves of BeH^- with five different basis set of increasing dimension (see text for more details). In the inset a zoom of the minima region has been reported (the lines corresponds to the fitted cubic functions).

where $D_e(l)$ represents the dissociation energy calculated with the basis-set truncated to the l-th angular momentum, A and β are suitable constants and D_∞ is the extrapolated dissociation energy for the infinite basis-set. This equation is a generalization of an extrapolation expression proposed in references [71, 72], to extrapolate CI energies as a function of the maximum basis-set angular momentum. The interpolating function has been reported in Figure 3.8, showing an extremely accurate reproduction of the fitted values. As we can see from the energetic profiles and from the interpolated minima, reported in Table 3.9, the inclusion of the higher angular momenta functions doesn't improve considerably the dissociation energy which is correctly represented already by the F basis set (with this basis set we can reproduce about the 99% of the extrapolated infinite basis-set D_e).

The same considerations can be made also on the interpolated equilibrium distance (see Table 3.9) that vary only by about 0.001 bohr passing from D to I basis set.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

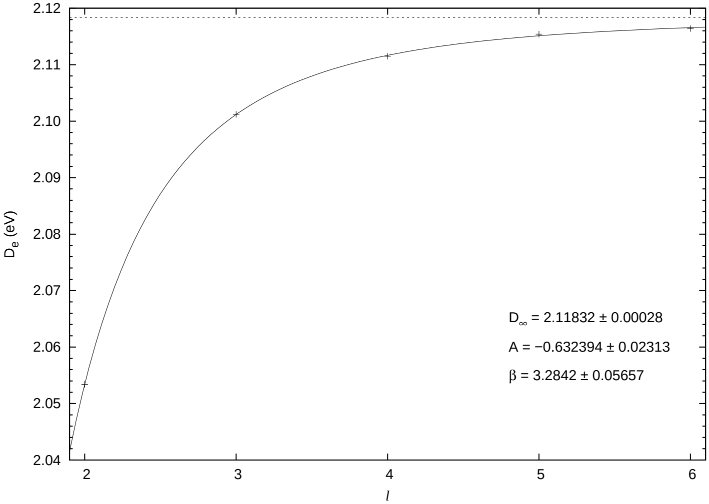

Figure 3.8: Extrapolated dissociation energy using equation 3.2 for the BeH$^-$ dimer. The dashed horizontal line represents D_∞ at 2.118 eV.

The long range part of the potential energy curve has also been analyzed for the *G* basis set. Figure 3.9 shows that the calculated FCI dissociation energies, for R=40, 60, 80 and 100 bohr of separation of the nuclei, are perfectly fitted by a function $-k/R^4$.

This form is, in fact, the expression expected for the interaction between two isotropic closed shells, a neutral and a charged ones. It can be shown that the value of the constant k is related to the polarizability of the neutral fragment, and we have $U(R) = q^2\alpha/2R^4$ where U is the interaction energy of the two atoms, q is the charge of the ion (1 in a.u.) and α the polarizability of the neutral atom [73, 74].

In our case, the polarizability of the Be atom, computed with the *G* basis set, is 37.662 (a.u.), while the same quantity, obtained as $\alpha = 2k$ and using the same basis set, is 37.932 (a.u.). The agreement between the two values is remarkable.

ISSN: 2279-8773

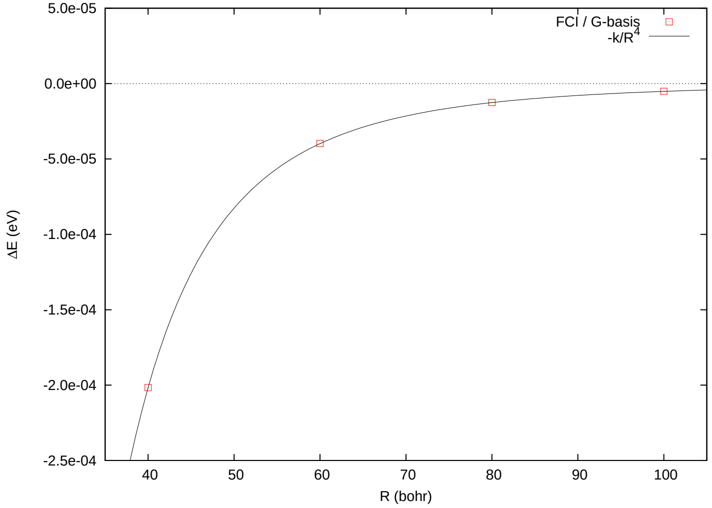

Figure 3.9: Long range potential energy curve fitted by a $-k/R^4$ function for the *G* basis-set. See text for more details.

3.8 ANALYSIS OF THE CHARGE DISTRIBUTION

As already mentioned, for large values of the interatiomic distance the BeH^- can be described as two interacting closed shell. When the two atoms approach each other the negative charge that at large distance is totally concentrated on the H atom moves gradually through the Be, despite the difference in electronegativity between the two atoms that should favor a negative Hydrogen.

In order to follow the movement of the charge we defined the Charge Shift (CS), η, as the ratio between the electronic center of charge, calculated at the FCI level, and the interatomic distance R.In all the calculations performed in this work, the Be atom is kept fixed at the origin of the coordinate system, while the H moves along the z axis. For this reason, at large value of R the contribution of the electrons on beryllium is practically zero, while the two electrons on the hydrogen give the largest contribution to the total electronic dipole moment. For our system, therefore, a value η=2 is obtained when two negative charges are located on the H nucleus; a value

VIRT&L-COMM.9.2016.4

Table 3.9: FCI interpolated dissociation energies and equilibrium internuclear distances for the different basis-sets.

Basis Set	$\mathbf{R}_{eq}$ **(bohr)**	$\mathbf{D}_{e}$ **(eV)**
D	2.7034	2.0534
F	2.7028	2.1012
G	2.7020	2.1115
H	2.7022	2.1154
I	2.7021	2.1164
∞	–	2.1183

of η=1 is obtained when the negative charge in excess is located at the middle of the distance between the two nuclei (the other electrons being on the nuclei); η=0 is obtained if all the electrons are located symmetrically around the Be atom; finally, a negative η means that there is an excess of negative charge on the opposite side with respect to he Hydrogen position. The η and the system energy (with respect to dissociation) as a function of the interatomic distance R for the G basis set have been reported in Figure 3.10.

As we can see from the plot, for large value of R the charge is completely localized on the hydrogen atom (η=2). At R=10 bohr, the initial stage of the formation of the Be-H bond, the η value starts to decrease considerably, and for geometries close to the equilibrium distance we found that the negative charge in excess is closer to the beryllium than to the hydrogen atom (η=0.6). By decreasing even more the value of R, the η passes through zero and then reaches negative values/ for $R = 1.8$ bohr, the η assumes a zero value, while at 1.0 bohr it is equal to about -0.7. This means that a relevant part of the negative charge in excess is located *on the opposite position* of Be with respect to the H atom.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

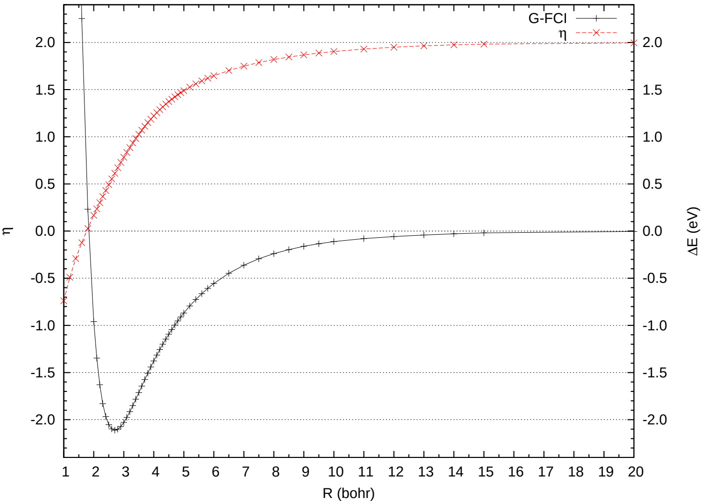

Figure 3.10: Charge shift (η) and potential energy curve comparison for the G basis-set.

In order to reach a better understanding of this peculiar feature of the system, we analyzed in detail the five most relevant FCI Natural Orbitals (NO) (the $1s$ orbital of Beryllium being kept frozen at the SCF level, it was not considered in the FCI expansion), and the center of charge associated to each orbital, from $R = 1$ to $R = 20$ bohr. These orbitals, when the interatomic distance becomes small, give rise to the formation of three σ orbitals and two degenerate π orbitals, that are labeled herein after as σ_1, σ_2, σ_3, π_x and π_y. Figure 3.11 shows the occupation and the contribution to the total dipole moment of the orbitals σ_1 (panel A) and σ_2 (panel B) at various value of R. As can be seen from the reported curves, the σ_1 orbital corresponds to the $1s$ orbital on the H atom at long distance of the nuclei giving the largest contribute to the total electronic dipole moment that, for large values of R, has a linear behavior as a function of R. The σ_2 orbital, on the contrary, shows a contribution that is practically zero when the two atoms are well separated and becomes very important at small value of

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

R. Its negative contribution to the charge average position is, in fact, the principal cause of the negative value reached by the CS. This orbital can be associated to the 2s orbital on the Beryllium which at large R gives a vanishing contribution to the total dipole, but, when the interatomic distance becomes small, it gives the formation of the sp hybrid which, following our interpretation, host the negative charge.

This is confirmed by the two-dimensional contour maps of the lowest NO's we reported in Figure 3.12. The 2D contour maps of the σ_1 and σ_2 orbitals, calculated at the FCI level with the F basis set, for R=1.0, 2.7, 10.0 and 15.0 bohr have been reported. As we can see from these plots, the σ_1 orbital is well localized on the H atom for large R and at short distance combines with one sp hybrid to form the σ bonding Be-H orbital. The σ_2 orbital, on the other hand, is localized on the Be when the two atoms are well separated but, with the shortening of R, changes gradually its shape and moves to negative z.

When the interatomic distance is small, also the influence of other higher orbitals becomes important and the contributions of these orbitals become relevant to understand the movement of the charge. In Figure 3.13 we have reported the same information of Figure 3.11 but for the π (because of degeneracy π_y and π_x have the same behavior) and σ_3 orbitals.
As can be seen from the plot, for large separation of the nuclei, these orbitals show a contribution to the total dipole moment equal to zero (despite an occupation different from zero) that suggests that both are localized on the Be atom. Their contribution, however, change drastically when the two atoms approach each other. The π orbital, indeed, gives a small negative contribution to the total dipole moment that starts to slightly decrease for values of R shorter than 13 bohr and reach a minimum (-0.055) for $R = 2$. This is also attested by the relative 2D contour plots reported in Figure 3.14 which show that this orbitals don't take part to the hybridization and therefore to the bond formation.

More interesting is, instead, the behavior of the σ_3 orbital. At long distance this orbital corresponds to the p_z atomic orbital, since it has an occupation equal to that of the p_x and p_y. Its contribution to the total dipole moment shows a maximum (0.4) at around 11 bohr that suggests a movement of the orbital towards the H atom for that distances. When the atoms become closer, however, the contribution strongly decreases and assumes

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

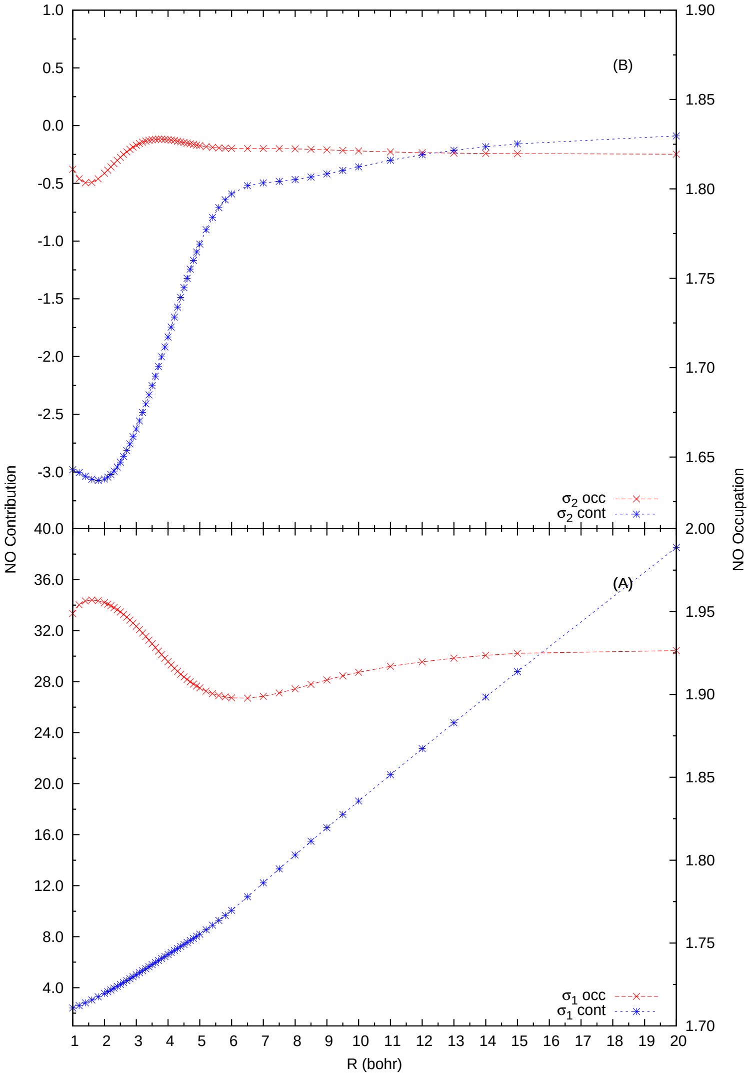

Figure 3.11: Occupations and contributions to the total dipole moment for the σ_1 (A) and σ_2 (B) natural orbitals for the *G* basis-set.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

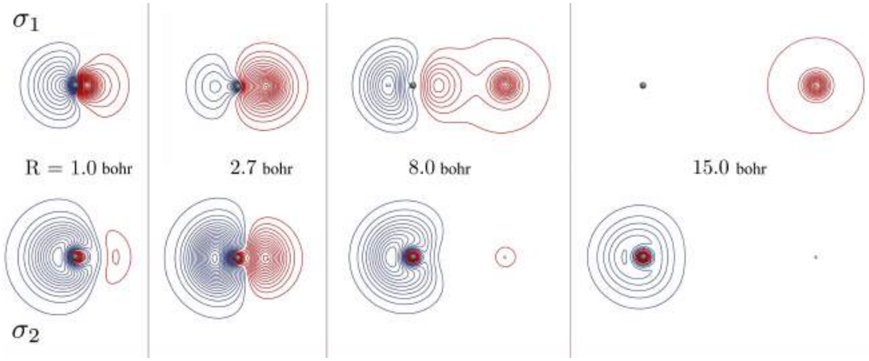

Figure 3.12: Two dimensional plots of the σ_1 (top) and σ_2 (bottom) natural orbitals, for different interatomic distances, calculated with the F basis-set.

also negative value for $R \leq 4$ bohr. Unlike the π orbitals, the contribution of σ_3 does not tend rapidly to its asymptotic value, confirming that this orbital has an active role in the description of the movement of the charge from H to Be which mainly take place in the σ space.
Looking at the occupation of the σ_3 orbital, one can notice the presence of an extremely sharp minimum close to $R = 3$ bohr. The corresponding charge contribution has a sudden change, although a careful analysis shows that both quantities are continuous. By analyzing the occupation of the other natural orbitals, we found the presence of an avoided crossing *in the occupation number* between the σ_3 and σ_4 (the next in occupation after σ_3). This has been reported in Figure 3.15, where one can see a weakly avoided crossing between the two orbital occupations. Correspondingly, we can notice a switch in the charge contributions, indicating the presence of two diabatic states (in the occupation number).

Finally, we give here an interpretation of these rather unexpected results in term of the *sp* hybrid orbitals of Beryllium. At long internuclear distance, the valence orbitals of Be are the atomic 2*s*, doubly occupied, and the three 2*p* orbitals, empty. For short internuclear distances, there is the formation of two *sp* hybrids. One of them combines with the Hydrogen 1*s* orbital to give rise to the bond orbital. The other one, that points on the opposite direction with respect to the H atom, is a lone pair. When these

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

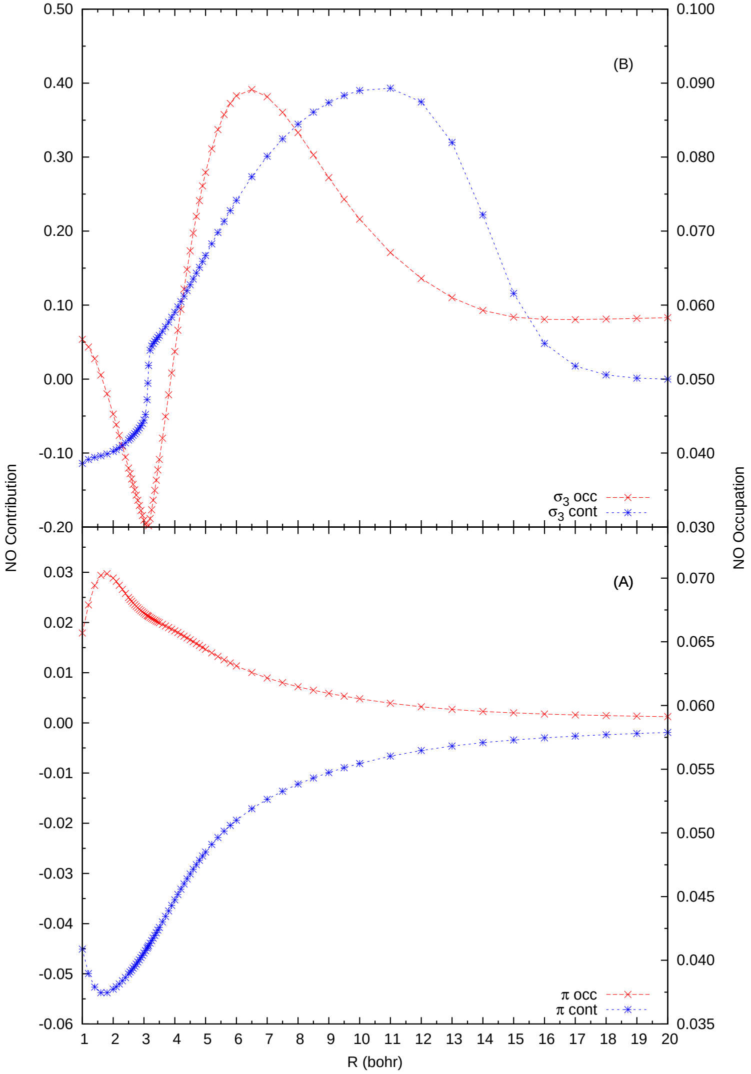

Figure 3.13: Occupations and contributions to the total dipole moment for the π (A) and σ_3 (B) natural orbitals for the G basis-set.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

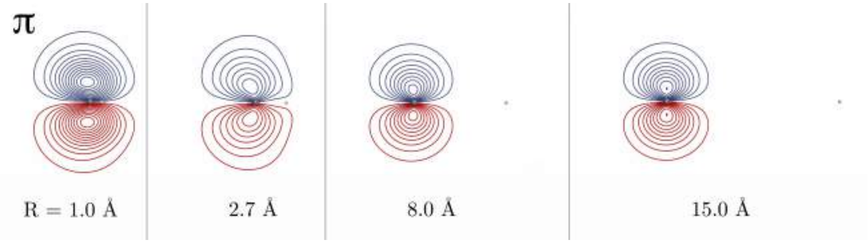

Figure 3.14: Two dimensional plots of the π natural orbital, for different interatomic distances, calculated with the F basis-set.

two orbitals have roughly the same size, the total charge contribution is close to zero. It is the case for the region close to equilibrium (actually, a vanishing η is obtained for slightly shorter distances). However, the size of these orbitals has a very different evolution as a function of R. The lone pair σ_2 is not very sensitive to variations in R, since the H nucleus is far from this orbital. The bond orbital σ_1, on the other hand, tends to reduce its size as the two nuclei become closer. This explains why, in the region between 1.0 and 2.0 bohr, the charge contribution of σ_1 strongly decreases as R is reduced, while the corresponding value of σ_2 is almost constant (see Figure 3.11, where the different scales should be noticed).

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

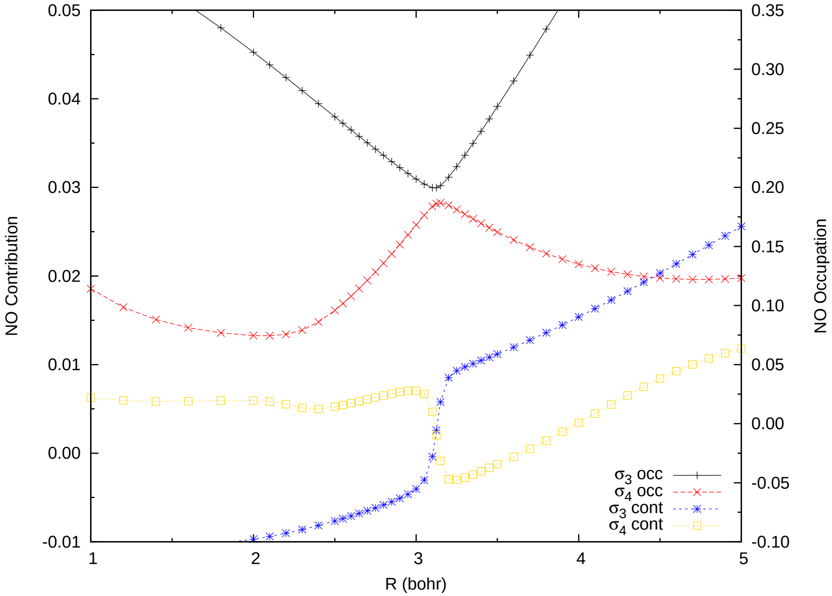

Figure 3.15: Occupations and contributions to the total dipole moment, in the avoided crossing region, for the σ_3 and σ_4 natural orbitals.

ISSN: 2279–8773

ISSN: 2279-8773

4

HPTC for a Dynamical problem

Significant theoretical and experimental work has been reported on weakly bound molecular clusters and complexes [75, 76]. Some of these studies have been performed in order to better characterize the dynamics of formation of highly symmetric simple molecular clusters (like the dimer $(H_2)_2$) and their main properties. Recently, growing efforts have been devoted to the determination of the interaction energy of two nitrogen molecules, $(N_2)_2$, a system that is important in high temperature atmospheric chemistry of Earth (like that occurring in spacecraft reentry [77, 78]). Moreover, due to the importance of $(N_2)_2$ in astrochemistry (it has been found, for example, that the N_2 dimer plays an important role in the atmospheric chemistry of Titan, the largest moon of Saturn), the investigation of $(N_2)_2$ has also been extended to low temperature [79].

In the past, several models potentials for the nitrogen dimer have been proposed by combining *ab initio* and experimental data [80, 81, 82]. However, despite the interest shown for the characterization of $(N_2)_2$ its intermolecular interaction still remains a subject of controversy. The system is optically inactive (for this reason it is unsuitable for experimental investigations) and the potential strongly depends on both the distance and the reciprocal orientation of the two molecules.

In this work we have employed the first block of GEMS to determine in an *ab initio* fashion the electronic energy of several arrangements of the $N_2(^1\Sigma_g^+) + N_2(^1\Sigma_g^+)$ system all the way through from reactants to products in nitrogen atom exchange reactions [83]. The steps of the development of a PES for N_2 + N_2 by carrying out a campaign of *ab initio* calculations of the potential energy values and the assemblage of the workflow exploiting in a combined way both the High Throughput (HTC) and the High

Performance (HPC) computing capabilities offered by the GRID will be described in detail.

4.1 THE COORDINATE SYSTEM AND THE INVESTIGATED ARRANGEMENTS

A first aspect we had to take care of was the formulation of a coordinate framework tailored to preferentially describe the subspace relevant to the investigated processes. In this work we adopted a space fixed axis system with the z axis passing through the centers of mass of the two nitrogen molecules. Using such coordinates the relative positions of the nuclei of a 4 atom system, like the N_4 one considered in this work, can be uniquely defined using the 6 variables (three radial and three angular) illustrated in Figure 4.1.

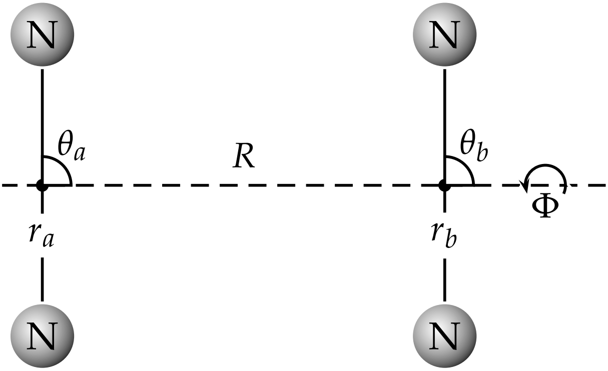

Figure 4.1: Schema of the adopted coordinate system for $(N_2)_2$.

In this scheme the orientation of the internuclear axis of the two separated molecules with respect to the z axis is specified by the angles θ_a and θ_b, while the dihedral angle, ϕ, is the one formed by the planes defined by one of the two N-N bonds and the z axis. The distance between the centers of mass of the two nitrogen molecules is given by R, while r_a and r_b represent the two intramolecular distances. The sampling of the various reaction channels was carried out by calculating the potential energy

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

value at five different θ_a, θ_b, ϕ triples for R varying from 1 to 8 Å and setting the bond length of the two monomers at the equilibrium distance of 1.094 Å [84] (each (θ_a, θ_b, ϕ; R) combination will be called arrangement hereinafter). The five arrangements belong to different symmetry groups (namely: D_{2h} (90, 90, 0; R), D_{2d} (90, 90, 90; R), $D_{\infty h}$ (0, 0, 0; R), C_{2v} (0, 90, 0; R), C_{2h} (45, 45, 0; R)) and are labeled also H (parallel), X (X-shaped), L (linear) and T (T-shaped) and Z (Z-shaped), respectively, as illustrated in Figure 4.2.

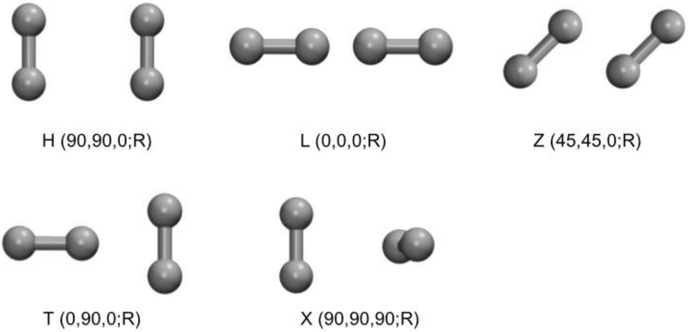

Figure 4.2: Summary of the five investigated arrangements.

Moreover in order to characterize the angular dependence of the potential energy, we carried out additional calculations by varying the Φ angle from 0° (H shape) to 90° (X shape), by a step size of 15° so that to fully characterize the change from parallel to crossed conformation (P15, P30, P45, P60, P75). At the same time, to better sample the effect of stretching the bonds, further calculations were performed by varying the internuclear distances r_a and r_b as follows: r_a = 1.094 Å, r_b =1.694 Å; r_a =1.494 Å, r_b =1.694 Å; r_a =1.694 Å, r_b =1.694 Å. The calculations provided us with a more accurate and a more detailed picture of the interaction of the $(N_2)_2$ system. A summary of the structural parameters for all the investigated configuration can be found in Table 4.1.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

Table 4.1: Summary of the structural parameters of the investigated configurations.

	Φ [degree]	θ_a [degree]	θ_b [degree]	$\mathbf{r_a}$ [Å]	$\mathbf{r_b}$ [Å]
H	0	90	90	1.094	1.094
	0	90	90	1.094	1.694
	0	90	90	1.494	1.694
	0	90	90	1.694	1.694
P15	15	90	90	1.094	1.094
P30	30	90	90	1.094	1.094
P45	45	90	90	1.094	1.094
P60	60	90	90	1.094	1.094
P75	75	90	90	1.094	1.094
X	90	90	90	1.094	1.094
	90	90	90	1.094	1.694
	90	90	90	1.494	1.694
	90	90	90	1.694	1.694
L	0	0	0	1.094	1.094
T	0	90	0	1.094	1.094
Z	0	45	45	1.094	1.094

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

4.2 THE LOWER LEVEL OF THEORY CALCULATIONS

To obtain a first systematic lower level *ab initio* evaluation of the electronic energies of N_4 we adopted a relatively small basis set called cc-pVTZ (correlation consistent polarized triple valence) involving 140 gaussian functions retrieved from the EMSL public database [58, 59] that is too small for the system considered and the final desired accuracy. Using such basis set we carried out extended MP2 (Second order Møller-Plesset perturbation theory) [85] *ab initio* calculations by correcting the BSSE (Basis Set Superposition Error) via the full counterpoise procedure (FCP) [86] as well as CCSD(T) (Coupled Cluster with Single and Double excitations and perturbative correction to the Triple) [87, 88, 89] ones. They were both carried out using the serial version of GAMESS-US [21] offered as a Grid service by the COMPCHEM VO, that thanks to its limited usage of memory and cpu time was run on the typical Grid single-core machines.

The adoption of a quite small (for this system) basis set has made the use of the GRID particularly suited for this type of study since single geometry calculations are platform independent and can be efficiently distributed on different nodes reducing considerably the total amount of computing time.

Table 4.2: CPU time and Memory allocations for CCSD(T) calculations with four different basis set of increasing dimension. The results for different number of processors have been reported.

Basis Set	**CPU time** [**seconds**]				**Memory** [**MBytes**]			
	CPU[1] number				CPU* number			
	1	2	4	8	1	2	4	8
cc-pVDZ	11[2]	16	9	7	36	44	59	90
cc-pVTZ	368	291	182	105	292	322	383	505
cc-pVQZ	7837	4394	2762	1698	1922	2181	2700	3737
cc-pV5Z	36610	24724	12744	8492	5540	6112	7256	9545

[1] Calculations performed on quad-core Nehalem EX (2,8 Ghz) cpus.

[2] Single core CPU time is shorter than Two core CPU time because, due to the small size of the basis-set, communications between processors slow down the calculation.

ISSN: 2279-8773

Table 4.2 shows memory allocation and CPU time for a stantard CCSD(T) calculation for the N_2-N_2 system. The performance of the GAMESS code for an increasing number of processors and a growing basis set are reported.
The potential energy values have been computed only for the L, T, H and X arrangements as a function of *R* and the results for the short range interaction region (from *R*=1 to *R*=3 Å) are reported in Figure 4.3.

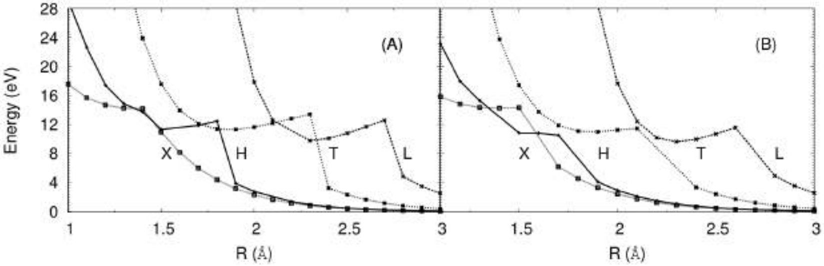

Figure 4.3: MP2 (panel A) and CCSD(T) (panel B) short range (*R*) potential energy values (connected by straight segments) for the H, L, T and X arrangements (see text for more details) plotted as a function of *R*. Distances are given in Angstrom, energies in eV with the zero set at total dissociation of the two nitrogen molecules.

Obviously, the plots of Fig. 4.3 can be used only as a qualitative guide to the actual reaction path because the diatomic bond lengths are not relaxed to search for an energy minimum [90]. The plots show a close similarity between X and H arrangements (there is also a close similarity between the T and Z ones not shown here) as well as a clearly decreasing repulsive behavior of the potential energy along the series L, T, H, X indicating that collinear encounters are quite unlikely to drive the system to reactive close encounters. The calculations show also that close encounters are in any case difficult to occur (though less difficult than for linear geometries) for perpendicular approaches. For H and X arrangements, instead, the repulsive barrier is located at such a small value of *R* to suggest the formation of a stable $(N_2)_2$ adduct and an opening of a side (relaxed bond length) window to reaction as possible. Indeed, the formation of bound states on the singlet ($^1\Sigma_g^+$) PES was already pinpointed by previous theoretical studies [91]. These are two relative minima of symmetry D_{2h} and T_d (the D_{2d} is

VIRT&L-COMM.9.2016.4

the largest Abelian subgroup of the T_d) that favourably shape the potential channel for an exchange process.

4.3 THE SUPERCOMPUTER BASED HIGHER-LEVEL OF THEORY CALCULATIONS

After having investigated on the Grid at a lower level of theory the arrangements more meaningful for understanding reaction, we turned our efforts into using a higher level ab initio CCSD(T) approach and the cc-pV5Z [58, 59] basis set that makes use of a total of 420 gaussian contractions. This suppresses the large BSSE that unphysically lowers the energy in the short range region (with respect to the long range one).
Using the (larger) cc-pV5Z basis set we repeated the calculations for the H, X, T, L and Z arrangements. The use of the cc-pV5Z basis set led inevitably to matrices of much larger dimensions that, while significantly improving the accuracy of the calculations, made the cost of the related computations three orders of magnitude larger and prompted the use of supercomputers (see Table 4.2). This motivated the cross submission from the EGI Grid (using gLite middleware) to the PRACE supercomputer platform of CINECA [20] (using Unicore middleware) on which parallel versions of electronic structure *ab initio* packages like GAMESS-US 2009 matching our calculation requirements are available.
With this larger basis set we repeated the calculations for all the arrangements reported in Table 4.1. The calculated intermolecular pair potential of N_2-N_2 for the H, X, T, L and Z arrangements considered by us are shown in Figure 4.4, in which the improved long range interaction energy is plotted as a function of R in the range 3-8 Å. As can be seen from the long range curves shown in Figure 4.4, which confirm the close similarity between the H and X and T and Z arrangements, all the investigated arrangements form van der Waals clusters having the energy minimum in the interval of R values raging from ca. 3.5 Å (for the H-shape) to ca. 5 Åfor the L-shape. In order to get a more accurate estimate of the energy well depth the five different curves have been interpolated, in the region close to the minimum, using a cubic function:

$$y = ax^3 + bx^2 + cx + d \tag{4.1}$$

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

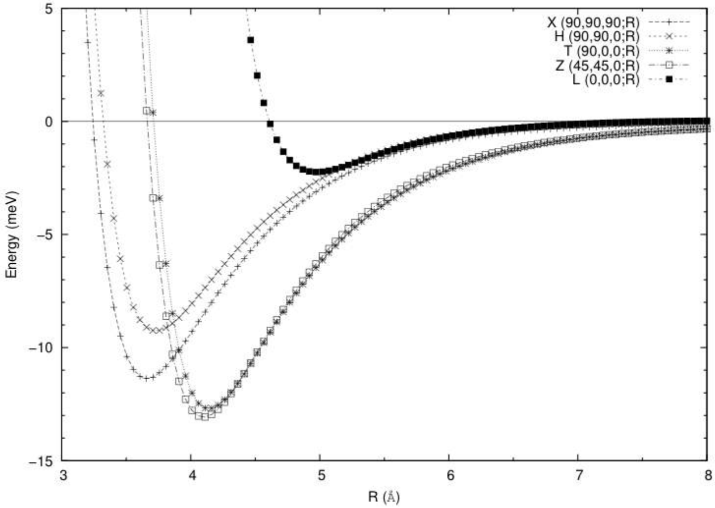

Figure 4.4: Long range CCSD(T)/cc-pV5Z potential energy curves plotted as a function of R for the H-shape, X-shape, T-shape, Z-shape and L-shape arrangements.

The minimum energy and the associated intermolecular distance R value obtained from the interpolation are showed in Table 4.3 where ΔE, the difference between the minimum energy and the corresponding asymptotic energy value calculated at the total separation of the two monomers -218.82612 Hartree) is reported. In this way we obtained that the T and Z arrangements are the most stable geometries and show also very close dissociation energies. On the contrary, the less stable bound geometry is the linear one, with a stabilization energy of few meV that is related to the head-to-head encounter of the two molecules. In addition to a comparable energetic stability, T and Z structures show also a very similar position of the minimum that is for both located at about 4.1 Å.
Due to their importance for reactive exchange processes, as singled out by the lower level *ab initio* preliminary investigation, we focused our attention on the parallel (H-shape) and crossed (X-shape) arrangements. These lead to wells deep 9.38 and 11.55 meV (less deep than for the T and Z-shape arrangements) whose minimum is located at R=3.74 Å (for H) and

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

R=3.64 Å (for X), respectively. Further information can be obtained from Figure 4.5, where the long range potential energy curves for the arrangements obtained by varying the Φ angle value from 0 (H-shape) to 90 (X-shape) degree have been reported. As can be clearly seen from the Figure,

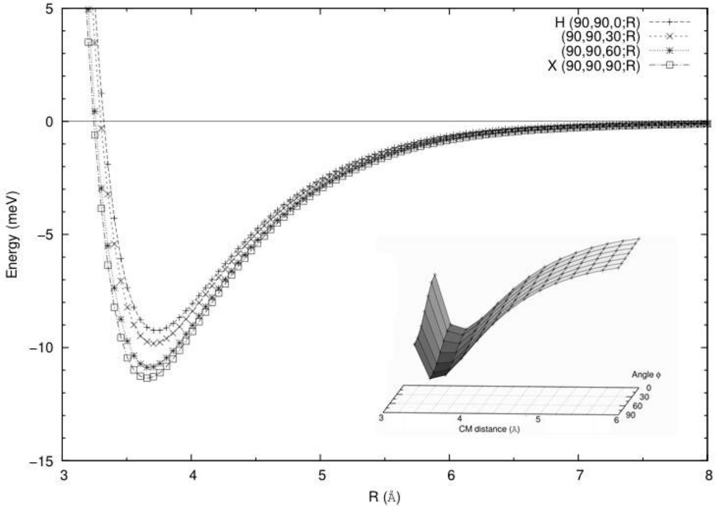

Figure 4.5: Long range CCSD(T) potential energy curves plotted as a function of R at different values of Φ (a pseudo 3D representation is given in the right lower corner).

Table 4.3: Interpolated Δ*E* and intermolecular distance value positions and energies for the five investigated configurations. The intramolecular bonds are fixed at their experimental equilibrium distance (1.094 Å).

	H	**X**	**T**	**Z**	**L**
R [Å]	3.74	3.64	4.15	4.11	4.98
ΔE [meV]	9.3782	11.552	12.833	13.186	2.2914

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

in going from the H to the X arrangement the location on *R* of the minimum lowers gradually with Φ and the well depth deepens. The pseudo 3D sketch highlights the smooth curvature of the PES going from $\Phi = 0°$ to $\Phi = 90°$.

Unlike the long range region, the short range one does not show any stable structures. Figure 4.6 suggests that when *R* decrease, the potential energy increases. However, the plots seem to indicate that a shift from a X to a H a channel is likely to allow exchange reaction via a change of arrangement. The X, T and L arrangements exhibit a local minimum that follows (as *R* decreases) a first maximum (located at larger distances) and moves to lower *R* values in going from L to X, while this is not true for the H and Z ones which exhibit only a repulsive wall. The dependence of the potential energy value from Φ in the same short range interval of *R* values is plotted in Figure 4.11.

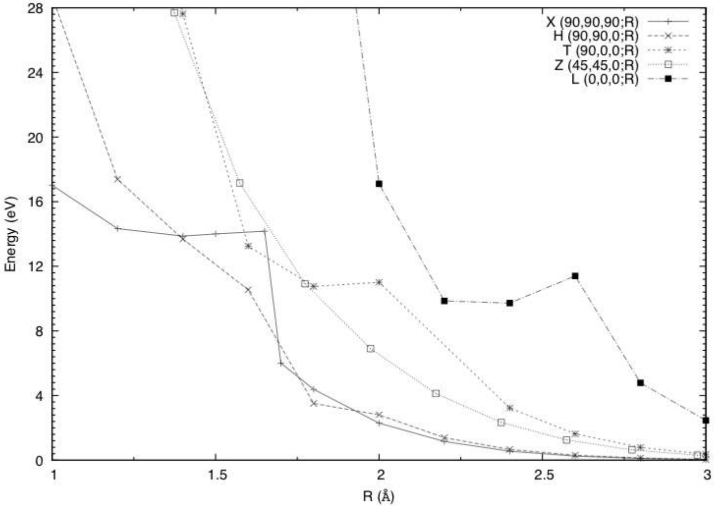

Figure 4.6: Short range CCSD(T) potential energy values plotted as a function of R, like in panel (B) of Fig. 4.3.

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

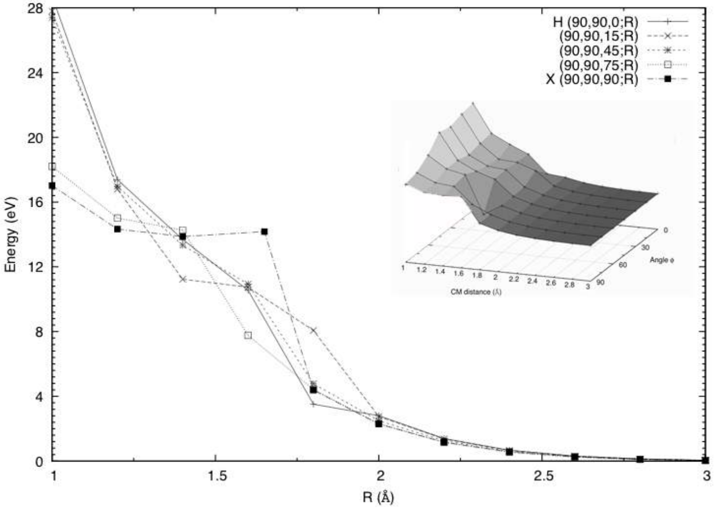

Figure 4.7: Short range CCSD(T) potential energy curves plotted as a function of R for different values of Φ.

In order to better analyze the PES, we also compared the behavior of the H and X arrangements when one or both intramolecular distances are stretched with respect to the N_2 equilibrium distance ($R_{eq} = 1.094$ Å). We checked the potential energy for the considered arrangements as a function of R while stretching the internuclear distances from 0.4 to 0.6 Å. For illustrative purposes, we plot in Figure 4.8 the potential energy curves of the H and X arrangements varying the intermolecular distances. However, when the intramolecular distance is increased from equilibrium the role of non-dynamical correlation energy becomes important and the wavefunction of the system could not be properly described by a single determinant.

ISSN: 2279-8773

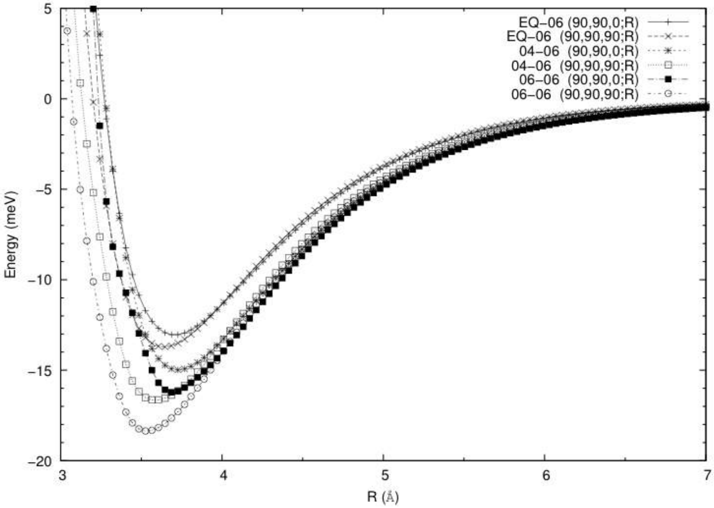

Figure 4.8: CCSD(T) Van der Waals potential energy curves plotted as a function of R of the three stretched geometries for the H and X arrangements. The plus and cross symbols refer to the arrangements with one intramolecular distance at the equilibrium and the other one stretched by 0.6 Å (EQ-06); the star and white-square symbols refer to the arrangements with one distance stretched by 0.4 Å and the other one by 0.6 Å (04-06) whereas the circle and the black-square symbols refer to the arrangements with both intramolecular distances stretched by 0.6 Å (06-06).

4.4 THE FIT OF THE AB INITIO POTENTIAL ENERGY VALUES

The last part of the work has been concerned with the fitting of the calculated CCSD(T)/cc-PV5Z results using a suitable functional form in order to allow a fast evaluation of the potential energy and its derivatives at any given geometry accessed during dynamical calculations for the $N_2(^1\Sigma_g^+)$ + $N_2(^1\Sigma_g^+)$ reaction. A further clear advantage of adopting the procedure of fitting the *ab initio* points using a functional form is associated with the fact that this allows the incorporation of corrections when the values of the *ab initio* points are found to be inadequate either during the fitting process or during dynamical calculations. The functional form adopted for this purpose is made of polynomials in physical coordinates. The theo-

VIRT&L–COMM.9.2016.4

retical ground for such representation is the Many Body Expansion (MBE) method developed by Sorbie and Murrell [92]. The MBE method is able, in principle, to fit potential energy surfaces for systems made of any number of atoms. The expression of the many-body expansion for a four-atom system is given as a sum of one-body (1), two-body (2), three-body (3) and four-body (4) terms as follows:

$$\begin{aligned} V(r_a, r_b, \theta_a, \theta_b, \Phi, R) \quad = \quad & V_A^{(1)} + V_B^{(1)} + V_C^{(1)} + V_D^{(1)} + \\ & V_{AB}^{(2)}(\xi_{AB}) + V_{AC}^{(2)}(\xi_{AC}) + V_{BC}^{(2)}(\xi_{BC}) + \\ & V_{AD}^{(2)}(\xi_{AD}) + V_{BD}^{(2)}(\xi_{BD}) + V_{CD}^{(2)}(\xi_{CD}) + \\ & V_{ABC}^{(3)}(\xi_{AB}, \xi_{BC}, \xi_{AC}) + V_{ABD}^{(3)}(\xi_{AB}, \xi_{BD}, \xi_{AD}) + \\ & V_{ACD}^{(3)}(\xi_{AC}, \xi_{AD}, \xi_{CD}) + V_{BCD}^{(3)}(\xi_{BC}, \xi_{CD}, \xi_{BD}) + \\ & V_{ABCD}^{(4)}(\xi_{AB}, \xi_{AC}, \xi_{AD}, \xi_{BC}, \xi_{BD}, \xi_{CD}) \end{aligned} \tag{4.2}$$

in which the four atoms are called A, B, C, D and ξ is an appropriate function of the relative positions of the nuclei. The one-body terms $V^{(1)}$ are the electronic energies of the atoms in the appropriate dissociation configuration. For the N_2+N_2 electronic ground state potential energy surface these terms were set equal to zero. In the Paniagua approach embodied in the gfit4c routines [93] adopted by us, ξ is defined as a function of the internuclear distances r_{ij} and the two-body components (say $V_{AB}^{(2)}(\xi_{AB})$) of the potential are expressed as:

$$V_{AB}^{(2)}(\xi_{AB}) = \frac{c_{0AB} e^{-\alpha_{AB} r_{AB}}}{r_{AB}} + \sum_{s=1}^{S} c_{sAB} \rho_{AB}^{s} \tag{4.3}$$

in which ρ_{AB} is a Rydberg-like variable formulated as:

$$\rho_{AB} = r_{AB} e^{-\beta_{AB}^{(p)} r_{AB}}, p = 2, 3, or\ 4 \tag{4.4}$$

with p being chosen as in [94]. The c_{sAB} coefficients (including s=0) as well as the non linear parameters $\beta_{AB}^{(p)}$ were evaluated by fitting the two-body *ab initio* values and S is the chosen order of the polynomial. The three-body terms (say $V_{ABC}^{(3)}(\xi_{AB}, \xi_{AC}, \xi_{BC})$) were expressed in terms of polynomials

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

of order N in the same Rydberg-like variables (ρ_{AB}, ρ_{AC} and ρ_{BC}):

$$V^{(3)}_{ABC}(\xi_{AB}, \xi_{AC}, \xi_{BC}) = \sum_{lmn}^{N} d_{lmn} \rho^{l}_{AB} \rho^{m}_{AC} \rho^{n}_{BC} \tag{4.5}$$

with the d_{lmn} coefficients being evaluated by fitting the three body ABC *ab initio* values. Finally, also the functional form chosen to express the four-body term $V^{(4)}_{(ABCD)}(\xi_{AB}, \xi_{AC}, \xi_{AD}, \xi_{BC}, \xi_{BD}, \xi_{CD})$ was written, similarly to what has been done for the three-body term, in terms of a polynomial of order N in the Rydberg-like variables (ρ_{AB}, ρ_{AC}, ρ_{AD}, ρ_{BC}, ρ_{BD} and ρ_{CD}):

$$V^{(4)}_{ABCD} = \sum_{ijklmn}^{N} e_{ijklmn} \rho^{i}_{AB} \rho^{j}_{AC} \rho^{k}_{AD} \rho^{l}_{BC} \rho^{m}_{BD} \rho^{n}_{CD} \tag{4.6}$$

where appropriate constraints are imposed on the indices in order to guarantee the right behavior of the polynomial in all regions of the configuration space.
In our case we have exploited the symmetry of the system by adopting a version of the routine designed for the A_4 system class of tetratomic molecules, giving potentials that are invariant with respect to the permutations of three and four equivalent nuclei, respectively. If the system under consideration has two or more identical nuclei, one must add further constraints in the linear (e_{ijklm} in Eq. 4.6) and nonlinear ($\beta(4)$ in 4.4) parameters to get that the global potential be invariant with respect to permutations of all the equivalent nuclei.

4.5 THE CCSD(T) POTENTIAL ENERGY SURFACE

As already mentioned, the fitting procedure of the PES for the $N_2(^1\Sigma_g^+)$ + $N_2(^1\Sigma_g^+)$ system has been based on three different sets of *ab initio* points: 1440 two-body values, 4320 three-body values and 1917 four-body values. Whenever possible *ab initio* values were also corrected using experimental information. This has been possible for both the two and three body terms thanks also to the work previously performed for the assemblage of the PES for the N + N_2 system [95, 96]. Moreover, all the *ab initio* values were scaled with respect to the dissociation energy of the dimer. The scaling values were calculated at CCSD(T) level of theory. In particular, we

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

calibrated the calculated potential energy values with respect to the single energy point obtained by increasing *R* till its large distance convergency and comparing the obtained value with the one obtained for an isolated nitrogen molecule (the value used is -218.82612523 Hartree).

Figure 4.9 reports the comparison of the *ab initio* calculated point and the one obtained from the fitted functional form. The plots refer to the short range region of the PES (R=1-3 Å) for the H and X arrangements. The comparison shows good agreement between the fitted PES and the calculated *ab initio* energy points. The point R=1.8 Å of the X arrangement is a clear exception. The fitted functional form, in fact, can't account for the strong discontinuity showing up in that zone of the surface. We have also tested the ability of the fitted PES to describe the N_2 dissociation. With our PES we obtain a dissociation energy of 0.361 hartree and an equilibrium distance of 1.099Å in good agreement with the experiment (1.098Å with and 0.358 hartree [97, 98]).

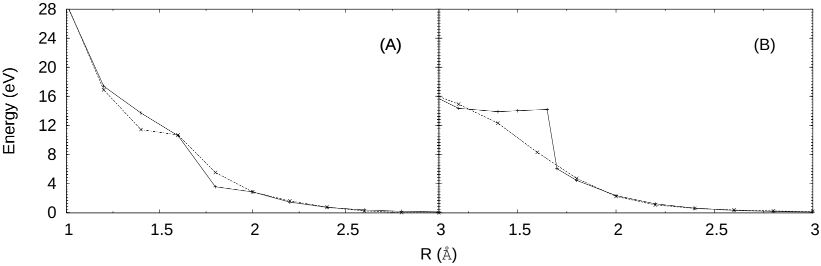

Figure 4.9: Comparison of the fitted (dashed line) and calculated (solid line) potential energy curves for the H arrangement (panel A) and for the X arrangement (panel B). Only the short range zone (R=1-3Å) is shown.

In order to better inspect the key features of the produced surface, we report also the isoenergetic contours of the fitted PES.

Figure 4.10 shows the contour maps of the PES for the H arrangement when keeping fixed the intermolecular distance *R* and varying the two intramolecular bonds r_a and r_b. This zone of the PES was not covered by our *ab initio* calculations because of the inability of the CC scheme to describe stretched configurations. The plots show that when the intermolecular distance is short ($R = 3$), the interaction of the two stretching molecules

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

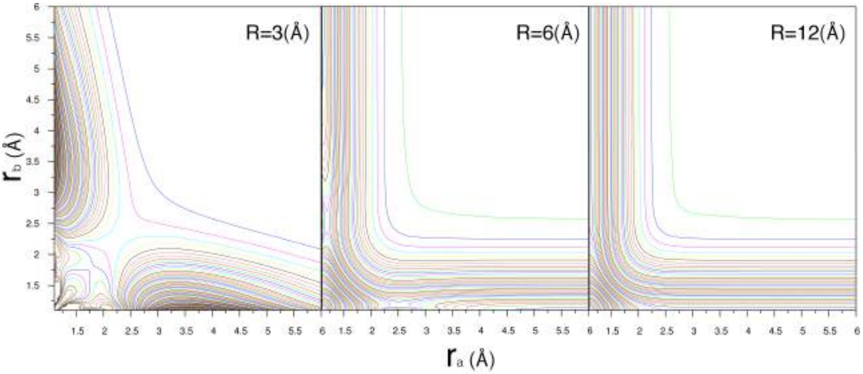

Figure 4.10: Isoenergetic contour plot for the H arrangement at R=3, 6 and 12 Å.

gives rise to the formation of particular energy profiles. When R increases the interaction between the two molecules decreases even if at $R = 6$ two small barriers show up both at the equilibrium and a stretched distance of 2.5 Å. For $R = 12$, on the other hand, all the interactions between the two molecule have disappeared and only the minimum associated with both distances at the equilibrium occurs.

In order to better understand the characteristic of the short range zone possibly leading to exchange reactive events, we have plotted the isoenergetic contours for the H arrangement by varying R while displacing the two intramolecular distances by the same quantity ($r_a = r_b$). Related contour plots are given in Figure 4.11. This figure shows the occurrence of several minima at short separation of the nuclei. In particular for R and r_a equal to about 2.5 Å a large well associable to the well know square minima.

4.6 THE MULTI REFERENCE REPRESENTATION OF THE PES

As already mentioned, when the interatomic distances are stretched from their equilibrium distances the wavefunction of the N_4 system is not properly described by a single determinant. When a triple bond, like the one in the N_2 dimer, is stretched the role of non-dynamical correlation becomes important and the electronic structure needs to be investigated with Multi Reference (MR) methods. This is particularly relevant for the short range

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

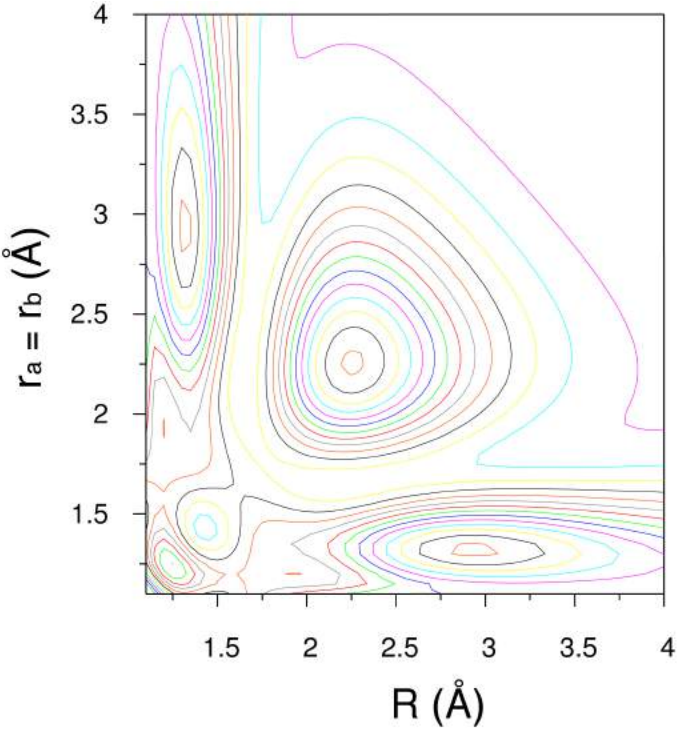

Figure 4.11: Isoenergetic contour plots for the H arrangement with $r_a = r_b$.

zone of the PES where the two molecules are very close and the stretched configurations are important for the description of bond breaking and formation.

In order to have a better description of the interactions governing the dynamics of possible dissociation and recombination processes, we have repeated the calculations with the MRPT2 computational method implemented in Gamess [99, 100]. This method makes use, in fact, of a CASSCF reference wavefunction, that lets us to recover the statical correlation energy, and a second order perturbation theory method that lets us to recover the dynamical correlation energy. In particular for the $N_2 + N_2$ case treated in this thesis, we chose to correlate the two degenerate π_{2p} orbitals and the

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

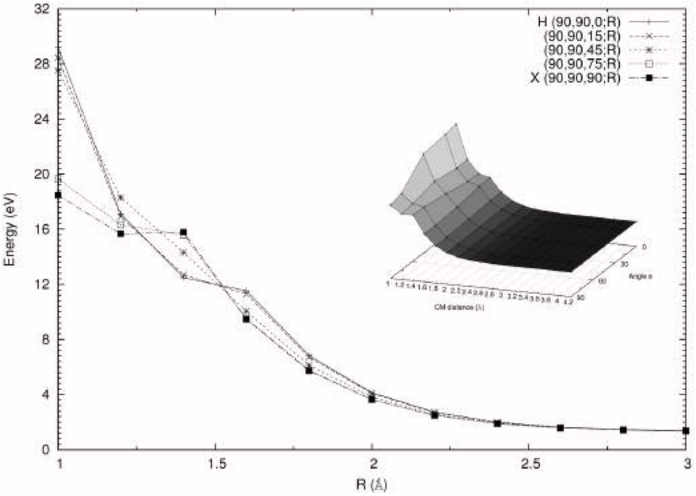

Figure 4.12: Short range MRPT2 potential energy curves plotted as a function of R for different values of Φ. The interatomic distances of each N_2 molecule are keep fixed at the equilibrium value ($r_a = r_b = 1.094$ Å).

σ_{2p} orbital of each molecule keeping the inner σ orbitals frozen at the SCF level. At this level of theory we repeated the calculations for the following arrangements (see Table 4.1 for more details):

- H, X, P15, P45 and P75 for $r_a = r_b = 1.094$ Å
- H, X, P15, P30, P45, P60 and P75 for $r_a = 1.094$ Å $r_b = 1.694$ Å
- H, X, P15, P30, P45, P60 and P75 for $r_a = 1.494$ Å $r_b = 1.694$ Å
- H, X, P15, P30, P45, P60 and P75 for $r_a = 1.694$ Å $r_b = 1.694$ Å

for values of distances of the center of mass, R, from 1 Å to 5 Å . The outcomes of these calculations are shown in Figures 4.12, 4.13, 4.14 and 4.15 where the potential energy curves referred to the total separation of the two monomers have been reported. A pseudo three-dimensional sketch of the potential energy surface is also reported in rhs of the figures for a better understanding of the topology of the potential energy values as a

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

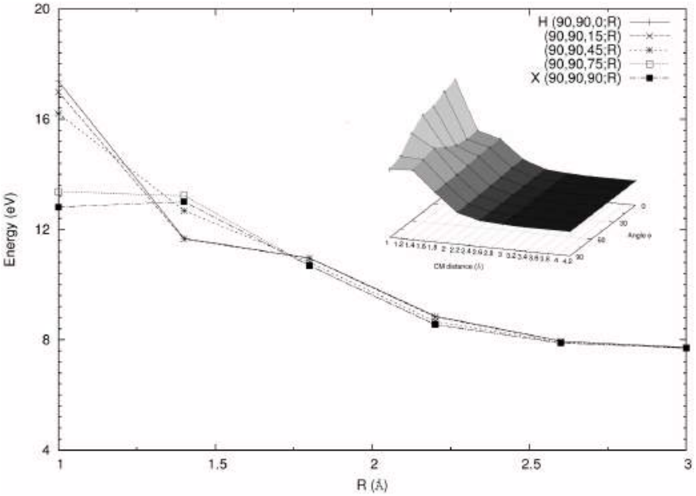

Figure 4.13: Short range MRPT2 potential energy curves plotted as a function of R for different values of Φ. One interatomic distance is stretched by 0.6 Å and the other is kept fixed at the equilibrium value ($r_a = 1.094$ Å , $r_b = 1.694$ Å).

function of the Φ angle and R.

As shown by Figure 4.12 the profile of the curves when both bonds are at the equilibrium value is similar to the one obtained with the CCSD(T) method (Figure 4.11). The positions and heights of the energy barriers for the two methods are also comparable, confirming the good quality of the coupled cluster calculations for these arrangements. Looking at the curve in the zone $2.5 < R < 3$ we can see that the energy does not tend to zero. This behavior can be explained with the differences in the theoretical approach followed by the methods adopted for the calculation of the zero of the energy (CCSD(T)) and for the illustrated potential energy curves (MRPT2). In the optimal conditions, in fact, coupled cluster methods are able to recover a greater fraction of dynamical correlation than MRPT2, giving the result closer to the real energy. Comparing the different curves in the long range zone of the non stretched configurations we have estimated an average energy shift of about 1.3 eV that can be used for the possible matching of MRPT2 and CCSD(T) values.

ISSN: 2279-8773

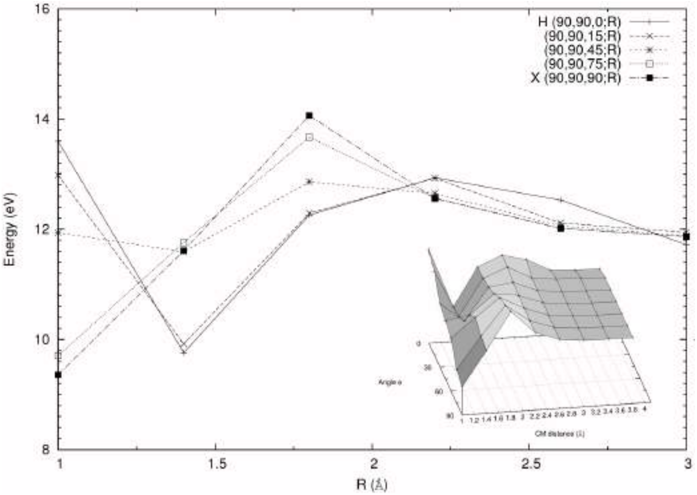

Figure 4.14: Short range MRPT2 potential energy curves plotted as a function of R for different values of Φ. One interatomic distance is stretched by 0.6 Å and the other by 0.4 Å ($r_a = 1.494$ Å , $r_b = 1.694$ Å).

The differences between the outcomes of the two approaches become more marked when we try to stretch the intramolecular distances. As mentioned above, CC calculations are not adequate to describe these situations and a Multi Reference picture is essential to have a correct description of the behavior of the system. Short range MRPT2 potential energy curves for the arrangement with one distance stretched by 0.6 Å and the other at the equilibrium value, is reported in Figure 4.13. As we can see, the energy of the PES at $R = 3$ Å is higher than that of the non stretched arrangement. This is due to the fact that the stretched molecule has a higher energy than that at the equilibrium and this increases the potential energy of the whole system. On the other hand, at very short distances $1 < R < 2$ the potential energies are lower than the previous case resulting in a decrease of the barriers height and therefore in a less repulsive behavior. The change is particularly evident for the crossed (X) arrangement whose energy at $R = 1$ is about 7 eV lower than in the non-stretched one confirming the possibility of a the existence of a reactive path.

VIRT&L-COMM.9.2016.4

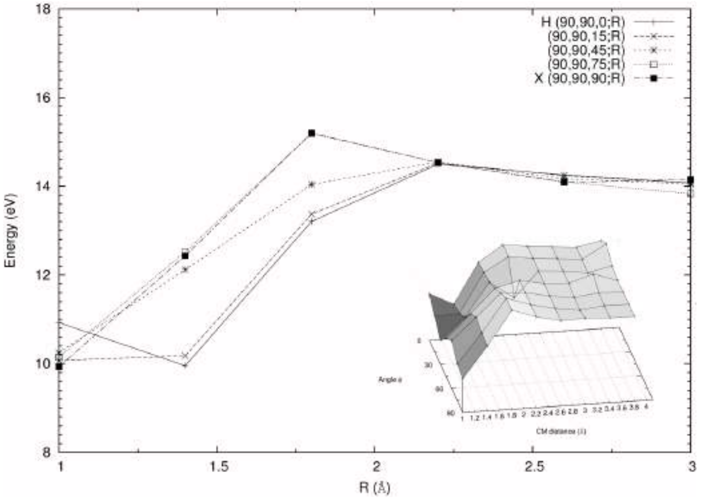

Figure 4.15: Short range MRPT2 potential energy curves plotted as a function of R for different values of Φ. Both the interatomic distances are stretched by 0.6 Å ($r_a = 1.694$ Å , $r_b = 1.694$ Å).

This is clearer if we analyze the PES when both the intermolecular distances are stretched. Figures 4.14 and 4.15 report the potential energy curves for respectively the unsymmetrical ($r_a = 1.494$ Å, $r_b = 1.694$ Å) and the symmetrical ($r_a = 1.494$ Å, $r_b = 1.694$ Å) stretching of the two intramolecular bonds. As for the previous case, the energy at $R = 3$ Å grows when the two intramolecular distances increase and at the same time the energy barriers fall off and move to larger R values. Very interesting is also the behavior of these arrangements for $1 < R < 2$. In this region in fact the potential energies decrease almost linearly with the shortening of R (the energy goes up again only for the H and P15 configurations) confirming the hypothesis of the opening of a possible reactive channel in this zone although further evidence needs to be found from additional calculations.

ISSN: 2279-8773

ISSN: 2279-8773

5

HPTC for a Kinetic problem

Traditional fossil fuels are considered to be largely responsible for causing important atmospheric degradations [101] such as global warming [102], acid rain, and tropospheric ozone increase. Is therefore of high importance to limit the environmental impact of using these fuels [103]. In the recent years an important effort as been made by industry to develop both more efficient types of engines and cleaner types of fuels.

This has prompted a recent growth of industrial projects targeting the development of detailed chemical kinetic models for the combustion of a wide range of hydrocarbons as well as the improvement of physical modelling of combustion in engines (turbulence, spray evaporation...). Integration of the chemical models for combustion and pollutant emissions in computational fluid dynamics simulations should be particularly useful in allowing the study of actual engine emissions over a wide range of operating conditions.

In order to model the kinetic of combustion systems we must be able to reproduce at the same time the global mechanism of the reaction and the formation of all the intervening species (both final and intermediate). To build these models, it is necessary to know with a good accuracy thermochemical and kinetic data, and therefore to develop a well structured high-level PES for each involved reaction.

Within the fully a priori approach followed in our group, one can perform such studies by going through the scheme of GEMS, whose first block carries out high level electronic structure calculations (see Section 1.6 for detailed information about kinetic workflow). These calculations need to be run on the High Performance Grid that, thanks to the availability of supercomputers, is able to adequately describe the Potential Energy Surface of

VIRT&L-COMM.9.2016.4

the most relevant stationary points involved in the reaction.
Once the PES is generated, the thermochemical and kinetic data of the related process, can be evaluated by the kinetic equations reported in the first Chapter. Such calculation do not require neither large memory nor long execution time and are therefore practicable also on a common desktop PC. However, in order to get a description of the dependence of the calculated kinetic data on pressure (P) and temperature (T), the computation of the kinetic constant needs to be repeated several times by changing the initial value of T and P. This makes these calculations, thanks to the independent nature of the various computational tasks, well suited for a massively distribution over different HTC nodes of the Grid.
In particular, this Chapter is focused on the *ab initio* evaluation of the kinetic properties of the reaction of formation of the Hydroperoxy Ethyl radical ($^{\bullet}CH_2CH_2OOH$) and its subsequent decomposition into cyclic ether and Hydroxyl radical ($^{\bullet}OH$) (relevant for the characterization of low temperature oxidation of alkanes).

5.1 MODEL THE KINETIC OF COMBUSTION REACTIONS

Basically, combustion reactions involve the oxidation of the fuel leading to the formation of carbon dioxide and water accompanied by the production of heat and light in the form of either a glow or flame. The simple overall chemical equation, however, does not describe the way in which the reaction occurs. A combustion mechanism involves, in fact, a sequence of elementary reactions (one and bi-molecular reactions) many of which include atoms or radicals that are short-living species with an high reactivity. In order to understand the fundamental mechanism of combustion reactions in gas-phase thermal reactions in general, different models are used for numerical simulations of the phenomena. The model for a given reaction system usually consists of the chemical species taking part to the reaction, their thermochemistry and a set of equations (usually Ordinary Differential Equations ODE) which mathematically describe the chemical process. The parameters of these equations are the rate coefficients which may be function of temperature, pressure and nature of the medium. In order to obtain a satisfactory agreement between model and experimental data the construction of a detailed kinetic mechanism needs to provide the most complete description possible including all chemical reactions and

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

species contributing to the observed quantity.

The creation of the kinetic model is the core and is the most computational intensive part of a combustion simulation. In order to handle the full complexity of a fuel mixture, a large number of species should be included in the model. However to avoid the explosion of its size only a limited number of secondary products must be taken into account. Models have been developed for a large number of hydrocarbons families [104], diesel surrogates [105], for several methyl esters [106] and for a few alcohols [107]. To generate a kinetic model usually a set of realistic reaction rules, able to describe the chemistry of the investigated system, are initially defined. This rules are then applied to a set of reactant molecules (usually only fuel) and all the possible combinations are performed to produce a set of specific reactions and a new pool of product molecules is formed that can be used in a following iteration. The set of rules adopted, the selection procedure of the molecules that are taken into account in a single iteration, the filtering of reactions and the way in which isomers are represented (lumped species [108]) distinguishes the various generation systems. Various codes have been proposed for the automatic generation of detailed combustion mechanisms such as MAMOX [109], NetGen [110], EXGAS [111], REACTION [112], etc.

Once the mechanism has been generated, the thermochemical and kinetic data of each reaction included into the detailed model must to be evaluated. Several experimental techniques are available for the measurement of new kinetic and thermochemical data [113, 114, 115, 116]. Recent years, however, have seen an increasing role of theoretical methods in the development of detailed chemical kinetic models. A priori thermochemical and kinetic parameters are nowadays routinely calculated using *ab inito* quantum techniques especially because experimental studies of elementary reactions are difficult to perform at temperature and pressure conditions of interest for combustion. The interplay between theory and experiments becomes therefore very important for the extrapolation of experimental results to a wide range of temperature and pressure conditions.

The *ab initio* calculations of the kinetic parameters that can be directly incorporated into the model, require the calculation of the PES using high-level theoretical methods. Because of the exponential dependence of the kinetic constants on the barrier height (see TST and RRKM theories in section 1.4), in fact, *ab initio* methods uncertainties can lead to large errors in the prediction of the rate coefficients. Moreover vibrational anharmonici-

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

ties, that can affect kinetic parameter through the calculation of the partition function and transmission coefficients, need to be evaluated carefully. This is the most computational expensive step of the simulation because one has to work out not only the *ab initio* PES, but also the optimized geometry and the expensive vibrational frequencies for each stationary point on the reaction coordinate.
Finally the produced model must be validated through complex computational simulations to obtain the most detailed and accurate experimental data to check the correctness of the mechanism. Moreover sensitivity analysis can be an efficient means for detecting the reactions having the largest impact on the combustion process. Their theoretical properties can be in this way refined and adjusted to fit experimental results and validate the modelling of the formation of minor products. For combustion systems, the underlying energy, transport, and mass interaction equations can be solved using available software designed for this purpose: e.g. CHEMKIN [117], COSILAB [118] or CANTERA [119].

5.2 LOW TEMPERATURE OXIDATION OF ALKANES

Kinetic models for the low temperature oxidation of alkanes, have been largely developed in the past years [120]. The chemistry of this fuel is, in fact, of high importance in the contest of limiting the environmental impact of traditional fossil fuels and for the industrial design of new type of fuel and engines. Although the processes involved in the alkanes combustion have been intensively investigated, large discrepancies still occur between the kinetic parameters used in most of the developed models [121].
The simplified schema of the main primary reactions included in the model is summarized in Figure 5.1. At temperatures below 900 K, alkanes react with $^{\bullet}OH$ radicals to produce alkyl radicals ($^{\bullet}R$). These radicals, after O_2 addition, form peroxy radicals $ROO^{\bullet}$ that can isomerize to give hydroperoxyalkyl radicals ($^{\bullet}QOOH$). These molecules react again with O_2 to form hydroperoxides which are degenerate branching agent, explaining the high reactivity of alkanes at low temperatures.
For temperatures between 700 - 800 K, however, the reaction of formation of the $ROO^{\bullet}$ is displaced back to reactants and the $\mathrm{RO_2}$ is less favored than formation of olefin and $^{\bullet}HO_2$. Moreover, $^{\bullet}QOOH$ radicals can decomposes into cyclic ether, ketons or $^{\bullet}HO_2$ radicals and smaller species reducing even more the formation of branching agents. This characteris-

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

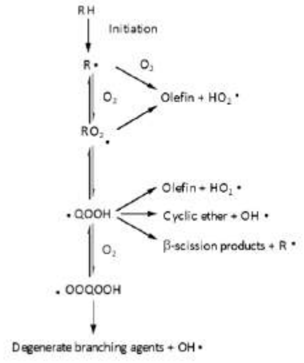

Figure 5.1: General schema of the primary reactions for the oxidation of alkanes (RH) generated with the EXGAS code [111].

tic temperature range takes the name of Negative Temperature Coefficient (NTC) zone, a particular temperature region in which reactivity decreases as the temperature increases. At temperatures higher than 900 K, the decomposition of $H_2O_2(H_2O_2(+M) = 2^\bullet OH(+M))$ becomes a new source of chain branching leading to an increase of reactivity.

5.3 AB INITIO STUDY OF AN ELEMENTARY PROCESS

We tried to refine the model by improving the thermochemical and kinetic data for a reaction involved in the NTC zone using quantum chemistry calculations.
This first part of the work focused, on the assemblage of an *ab initio* PES for the formation of the $^\bullet QOOH$ radical and its subsequent unimolecular decomposition into Cyclic Ether (CE) and $^\bullet OH$ radical. In particular we investigated the reaction

$$CH_3CH_2OO^\bullet(A) \rightarrow^\bullet CH_2CH_2OOH(B) \rightarrow CH_2OCH_2(C) +^\bullet OH \tag{5.1}$$

The ethyl peroxy radical (A) is the smallest compound which, after isomerization (B), gives the formation of the cyclic ether oxirane (C) and is considered, thanks to its small size, a good prototype for the description

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

of the formation of cyclic ethers. The two reactions pass through two transition states (TS) called TS$_{iso}$ (for the isomerization) and TS$_{oxi}$ (for the formation of oxirane). The isomerization involves the transfer of an H-atom that is a very fast reaction strongly affected by tunnel effects. The related correlation diagram is sketched in Figure 5.2.

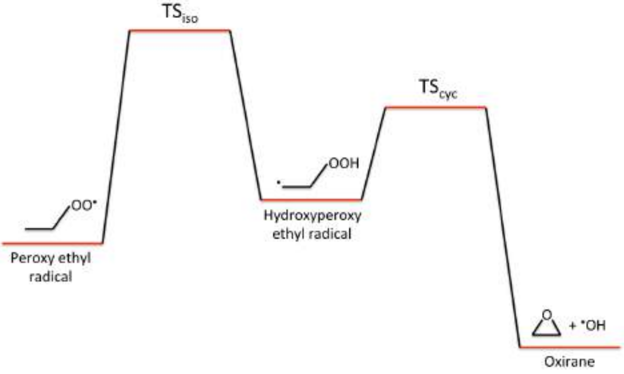

Figure 5.2: Potential energy correlation diagram for the $^{\bullet}QOOH$ formation and dissociation reactions.

Beyond its relevance for combustion chemistry this reaction is an interesting case study also for the multi reference nature of the TS$_{oxi}$ wavefunction. It has been pointed out, in fact, that the transition state structures of the CE formation from $^{\bullet}QOOH$ radicals generally suffer from spin- contamination [122]. Most of the combustion kinetic studies on these molecules, are based on DFT and/or single reference methods (usually of the multi-level type such as CBS-QB3 [123, 124, 125], Gn-X [126, 127]). Table 5.3 show the results of different DFT version and single reference calculations determining the energy of the investigated stationary points with respect to the energy of $^{\bullet}C_2H_5 + O_2$ the $ROO^{\bullet}$ radical precursor. From an analysis of these results we can conclude that DFT (B2PLYP, BMK, M06, X3LYP, and B3LYP) methods generally lead to semi-quantitative results. As a matter of fact the M06 computed PES is, in fact, in good agreement with the FCC/CBS values (which are considered as reference) and its use

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

Table 5.1: Energies (kcal/mol) at 0 K calculated with respect to $^{\bullet}C_2H_5 + O_2$. The spin contaminated species is colored in red.

	B2PLYP[1]	BMK[1]	M06[1]	X3LYP[1]	B3LYP[1]	CBS-QB3	ROCBS-QB3	G4	CC /CBS[2]	FCC /CBS[3]	G2like[4]
RO_2•	-28.8	-34.8	-32.9	-29.6	-28.6	-34.1	-32.7	-33.4	-33.0	-33.3	-33.9
TS_{iso}	9.6	5.2	4.7	7.8	9.4	1.8	3.5	3.7	4.0	3.1	3.1
•QOOH	-10.6	-14.3	-16.0	-10.6	-8.0	-17.2	-15.8	-16.7	-16.8	-17.1	-17.0
TS_{oxi}	6.4	1.3	-1.6	1.8	4.7	-3.4	-3.8	-0.3	-0.6	-1.3	-0.6
cC_2H_4O+•OH	-27.4	-33.7	-35.6	-26.1	-21.0	-33.2	-31.8	-32.7	-33.0	-33.0	-34.0

[1] B2PLYP, BMK, M06, X3LYP, and B3LYP calculations performed with 6-311+G(3df,2p) basis set.

[2] Geometries optimized at the CCSD(T)/cc-pVDZ level of theory;
CBS extrapolation performed with cc-pVTZ and cc-pVQZ basis sets.

[3] Same as for CC/CBS with correction for CI truncation errors as described in [130].

[4] Approximate E[QCISD(T)/6-311++G(3df, 2pd)] as described in [131].

may be appropriate also for larger systems. All stationary points are well described using a single-reference wave-function, except TS_{oxi}. As shown in Table 5.3 these methods are unable to correctly describe the electronic wavefunction of TS_{oxi}. The expectation value of the spin-squared operator and the T1 diagnostic for the CC calculations show that the tested methods

Table 5.2: $< S^2 >$ value and T1 diagnostic results for DFT and single-reference methods.

	B2PLYP[1]	BMK[1]	M06[1]	X3LYP[1]	B3LYP[1]	CBS-QB3[2]	G4[2]	CCSD(T)/ cc-pVDZ	T1 diagnostic
RO_2•	0.759	0.754	0.754	0.755	0.753	0.764	0.764	0.760	0.025
•QOOH	0.756	0.754	0.755	0.754	0.754	0.763	0.763	0.763	0.012
TS_{iso}	0.774	0.761	0.759	0.760	0.759	0.810	0.808	0.817	0.026
TS_{oxi}	0.932	0.830	0.826	0.936	0.816	1.286	1.259	1.295	0.041

[1] B2PLYP, BMK, M06, X3LYP, and B3LYP calculations performed with 6-311+G(3df,2p) basis set.

[2] Eigenvalues of S^2 operator from post-HF calculation steps in composite methods

ISSN: 2279-8773

overestimate the correct spin value (0.75) and lead, therefore, to incorrect energies.
Because of this we made use of the CASPT2 method [128, 129] as implemented in the Molpro code, to work out an *ab initio* multi reference PES for the above mentioned reaction and in this way evaluate the effect of the spin contamination on the TS_{oxi} barrier height.
In order to ground our work on good starting points, trial configurations of the different stationary points were first optimized at the DFT level. Then, CASPT2 geometry optimizations for the $ROO^\bullet$ and $^\bullet QOOH$ minima as well as for the two transition states TS_{oxi} and TS_{iso} were performed. A representation of the obtained structures is given in Figure 5.3.
However, transition state optimizations are usually definitely more difficult than for minima finding and require good starting points and a high level approximations of the hessian. The latter is evaluated numerically before the first optimization step (and when the Quadratic Steepest Descend method is near to a trap region) making each geometry optimization very CPU time demanding. For the calculations performed in this work we, have employed an aug-ccpVTZ [58, 59] basis set that provides a good compromise between speed of computations and quality of results. The active space for the CASPT2 method has been also accurately designed in order to keep the calculations within reasonable CPU time and memories requirements, but at the same time to fully account for the physics of the problem. The reaction of formation of the ring involves, indeed, the breaking of the O-O bond of the $^\bullet QOOH$ molecule and the formation of the C-O one while the radical center "moves" from carbon to the oxygen of the hy-

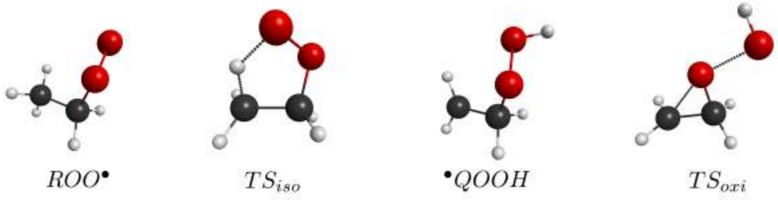

Figure 5.3: CASPT2(3,4) optimized structures for the four stationary points calculated with the aug-cpVTZ basis set. Black balls represent carbon atoms, red ones represent oxygen atoms and the white balls hydrogen atoms.

VIRT&L-COMM.9.2016.4

droxyl radical. On the other hand the reaction of formation of the $ROO^{\bullet}$ involves the breaking of an C-H bond and the formation of the O-H one. In order to get a good description of the process and consistency between the results obtained from the different points (energy differences must be calculated), we have localized the molecular orbitals (through Pipek-Mezey localization) for each input structure and carefully selected the orbitals which take part to the reaction. We have in this way included in the active space of the $ROO^{\bullet}$ radical the C-H σ orbital and the single-electron occupied orbital whereas for $^{\bullet}QOOH$ we took into account the orbital hosting the unpaired electron and the σ O-O orbital. Figure 5.4 shows a three dimensional view of the orbitals included in the active space for a contour value of 0.1. In both cases also the first two virtual orbitals and three electrons have been correlated obtaining a (3,4) active space.

In order to calculate the energies at the Complete Basis Set (CBS) limit

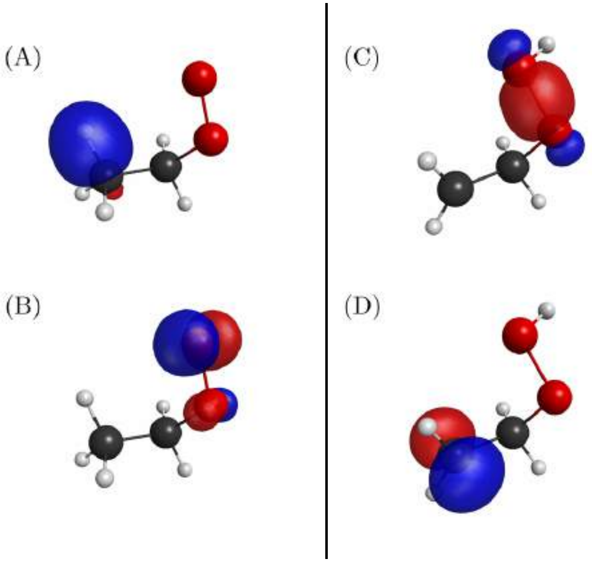

Figure 5.4: Three dimensional plot of the occupied orbitals included into the (3,4) active space. A and B are respectively the σ C-H and the single-electron orbital for the $ROO^{\bullet}$ radical while C and D are σ O-O and the single occupied orbital for the $^{\bullet}QOOH$ radical.

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

we extrapolated the potential following the formula proposed by You *et. al.* [130]

$$E(x) = E(\infty) + Ax^{-3} \tag{5.2}$$

where $E(\infty)$ is the energy at the CBS limit, x is the order of the higher angular momentum function contained in the basis-set and $E(x)$ the energy of the relative basis set.

For each CASPT2/aug-ccPVTZ optimized point we have performed additional single point calculations with aug-ccPQZ ($x = 4$) and aug-ccP5Z ($x = 5$) that have been interpolated with equation 5.2. The resulting energies are summarized in Table 5.3 where the $\Delta E(kcal/mol)$ with respect to the $ROO^{\bullet}$ energy is reported for different computational methods. As can be shown by the table, the CASPT2/CBS results are in good agreement with the FCC/CBS values for both the TS_{iso} and the $^{\bullet}QOOH$. On the other hand, the relative energy of the TS_{oxi} computed at the CASPT2/CBS level of theory is significantly lower than the other result. This can be understood as an the effect of spin contamination which increases the calculated energy of the TS_{oxi} to higher barriers.

Table 5.3: Energy (kcal/mol) at 0 K calculated with respect to the $ROO^{\bullet}$ energy.

	M06	CBS-QB3[1]	ROCBS-QB3	G4	CC/CBS	FCC/CBS	CASPT2/CBS
TS_{iso}	37.6	36.0	36.2	37.2	37.0	36.5	**36.7**
•QOOH	16.9	16.9	16.9	16.7	16.2	16.2	**16.5**
TS_{oxi}	31.2	30.7	28.9	33.1	32.3	32.0	**26.3**

[1] CBS-QB3 method features an empirical correction for spin contamination that lowers energy of TS_{oxi} by about 3 kcal/mol.

In order to evaluate the effect of the active space size on the computed energies, we have performed additional single point calculations on the CASPT2(3,4)/aug-ccpVTZ geometry with a (7,8) active space. Using such an active space we obtain a barrier for TS_{oxi} of 10.13 kcal/mol while with the (3,4) active space, for the same barrier, we obtain the value of 10.35 kcal/mol. The resulting energies barriers are therefore in good agreement, confirming the ability of the adopted active space to recover most of the static correlation effects.

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

5.4 THE EVALUATION OF KINETIC AND THERMOCHEMICAL DATA

As already mentioned, the assemblage of a combustion model requires the evaluation of the kinetic and thermochemical data of each elementary reaction involved in the process. In recent years computational methods have demonstrated to be valuable and indispensable tools in order to work out these quantities. High-level *ab initio* studies of chemical systems offer, in fact, not only a qualitative insight but also a quantitative evaluation of the thermochemical and kinetic properties characterizing reactive processes. Although the continuous increase in accuracy of quantum chemistry calculations the evaluation of the partition function and transmission coefficient, necessary for estimating both the thermochemical and kinetic data, often involves additional considerations and the introduction of approximation based on the scientist intuition and ability. Examples of this are the treatment of internal rotations and quantum tunnelling effects which can be crucial to minimize errors during the calculation of rate coefficients.

In this work we have employed the MultiWell [132] software to work out the thermochemical and kinetic parameters for the above mentioned process (5.1). This code is distributed together with various tools designed for specific purposes:

- **MomInert**: calculates principal and approximated reduced moment of inertia for molecules starting from the coordinates of the optimized geometry.
- **Thermo**: calculates entropy, heat capacity, and $H(T) - H(0)$ for individual species.
- **DenSum**: carries out exact sum and density of states using the Stein-Rabinovitch algorithm [133].
- **MultiWell**: calculates, within the RRKM/ME (see section 1.4) approach, time-dependent concentration, yields, vibrational distribution, and rate coefficients as a function of temperature and pressure.

Thermochemical properties

Thermochemical data are essential quantities for the understanding of chemical processes especially for combustion reactions. Standard-state enthalpies

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

of formation (H°), entropies (S°) and heat capacities (C_p°) are, indeed, used for the evaluation of the heat release and of the consistency between forward and reverse reactions properties. Thermodynamic data must be therefore calculated for all the species involved in the process. The common way to represent H°, S° and C_p as a function of temperature is given by the specific format used by the NASA Complex Chemical Equilibrium Program [134]. The thermodynamic data are stored as polynomial fits to specific heat C_P/R, enthalpy H°/RT, and entropy S°/R:

$$\begin{aligned} \frac{C_p^\circ}{R} &= a_1 + a_2T + a_3T^2 + a_4T^3 + a_5T^4 \\ \frac{H^\circ}{RT} &= a_1 + \frac{a_2}{2}T + \frac{a_3}{3}T^2 + \frac{a_4}{4}T^3 + \frac{a_5}{5}T^4 + \frac{a_6}{T} \\ \frac{S^\circ}{T} &= a_1 lnT + a_2T + \frac{a_3}{3}T^2 + \frac{a_4}{4}T^3 + \frac{a_5}{5}T^4 + a_7. \end{aligned} \tag{5.3}$$

Usually, two sets of coefficients are prepared for two temperature ranges, 300-1000 K and 1000-5000 K. The thermodynamic data consists of seven coefficients ($a_1, a_2 ... a_7$) for the high temperature range and other seven ($a_1', a_2' ... a_7'$) for the low temperature one, for a total of 14 coefficients.

We have therefore employed the above mentioned Thermo program, to compute entropy, enthalpy and heat capacities at different temperatures for each stationary point represented in Figure 5.3. Following the statistical mechanics formulae [135] adopted by the thermo code, all the thermodynamic molecular properties can be expressed in term of the molecular partition function, with a vibrational, rotational, translational and electronic contribution. To obtain the molecular entropy, for example, the following equations need to be evaluated [136]:

- *Vibration*:

$$S_{vib} = R\sum_i \Big[\frac{(h\nu_i/kT)exp(-h\nu_i/kT)}{1-exp(-h\nu_i/kT)} - ln[1-exp(-h\nu_i/kT)]\Big] \tag{5.4}$$

- *Rotation*:

$$S_{rot} = \frac{R}{2}\, ln\Big[\Big(\frac{\sqrt{\pi}}{6}\Big)\sqrt{\Big(\frac{2I_AkT}{\hbar^2}\Big)\Big(\frac{2I_BkT}{\hbar^2}\Big)\Big(\frac{2I_CkT}{\hbar^2}\Big)}\Big] + \frac{3}{2}R \tag{5.5}$$

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

- *Translation*:

$$S_{tra} = R\Big[\frac{5}{2} + \frac{3}{2}\, ln\Big(\frac{2\pi k}{h^2}\Big)\Big] \tag{5.6}$$

- *Electronic motion*:

$$S_{ele} = R\, ln\, s \tag{5.7}$$

with I_A, I_B and I_C being the moments of inertia of the global molecule, R the gas constant, k the Boltzmann's constant, h the Plank's constant, T the absolute temperature, σ the symmetry number involving rotations of the entire molecule [137], s the electronic spin multiplicity and ν_i the vibrational frequency associated with the ith normal mode.

The moments of inertia have been calculated using the Mominert code starting from the optimized geometries of each stationary point. Harmonic vibrational frequencies and normal modes have been then calculated using the Molpro code. The Hessian was computed numerically using finite differences in 3N cartesian coordinates and the already described CASPT2(3,4)/aug-ccpVTZ method.

The vibrational contribution (Eq. 5.4) to the entropy is the most significant part and needs to be evaluated with high accuracy. However, the above mentioned expressions, particularly the vibrational term S_{vib}, are only valid for molecule slightly deviating from equilibrium. This is true at low temperatures or for internal motions characterized by steep potentials. The low vibrational spectrum in most molecules is, however, characterized by some large-amplitude vibrations which give rise to large deviations from the equilibrium configuration. The Rigid Rotor Harmonic Oscillator (RRHO) approximation usually associated with low frequency vibrational modes, can therefore introduce large errors since the corresponding quadratic formulation of the potential is no longer valid (see n-butane example in Figure 5.5). As both RRHO and free rotor assumptions fail to accurately represent these low frequency rotational motions, several methods, like the Pitzer and Gwinn's [13] tabulation, have been developed to more accurately calculate the hindered rotor contributions.

In order to correctly treat internal rotations, we have adopted the 1-D hindered rotor (1-DHR) model [12]. Following this approach, the potential energy surface for each rotation is determined starting from a TS or minimum potential energy structure and then progressively rotating the top over 360°. The obtained energy profile is fit to a trigonometric function and the reduced moment of inertia for the rotating top is calculated from

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

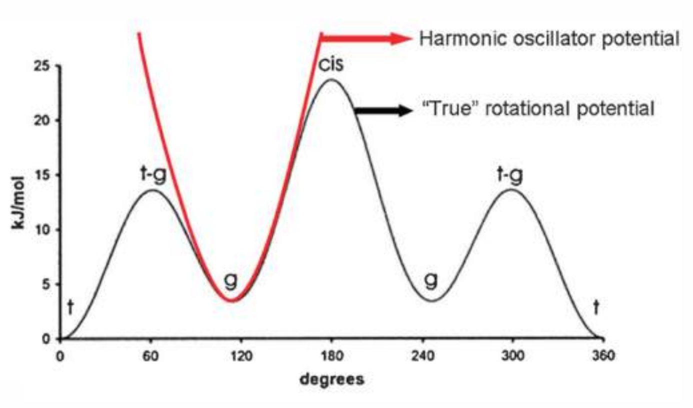

Figure 5.5: Potential energy for the ethyl group rotation in a n-butane molecule.

the geometry of the stationary point. Once the potential is obtained, the energy levels for the rotation are calculated by solving a 1-D Shrödinger equation,

$$-\frac{h^2}{8\pi^2 I_{red}}\frac{d^2}{d\theta^2}\Psi + V(\theta)\Psi = E\Psi \tag{5.8}$$

where I_{red} is the reduced moment of inertia for the rotating top, θ the rotation angle and $V(\theta)$ is the associated rotational hindered potential. The partition function is finally corrected to account for the contribution of each internal rotation.

We have therefore analyzed the results of the *ab initio* vibrational analysis to identify the internal rotators and their associated frequencies. As shown in Figure 5.6, for the four investigated stationary points, we have identified six different internal rotations:

- $ROO^\bullet$
  - CH_3 group rotation around $C-C$ bond axis (a1)
  - OO group rotation around $C-O$ bond axis (a2)

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

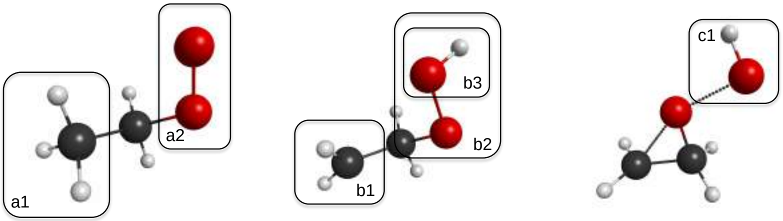

Figure 5.6: Representation of the hindered rotors for the $ROO^\bullet$, $^\bullet QOOH$ and TS_{oxi} structures. (See text for more details).

- $^\bullet QOOH$
    - CH_2 group rotation around $C-C$ bond axis (b1)
    - OOH group rotation around $C-O$ bond axis (b2)
    - OH group rotation around $O-O$ bond axis (b3)
- TS_{oxi}
    - OH group rotation around $C-O$ bond axis (c1)

For the TS_{oxi} transition state we have purposely kept fixed the CCOO chain to better represent the ring formation process while, because of its cyclic structure, the TS_{iso} does not show any internal rotation. For each identified hindered rotor, we have then performed a scan of the rotation angle (step size of 10°) using the UB3LYP DFT procedure implemented in the Gaussian code [138]. Constrained geometry optimizations have been performed at each step of the scan and the resulting energy profile fitted using the following series:

$$V(\theta) = V_0 + \sum_{n=1}^{N} V_n^c \, cos\big(n\,\sigma_V(\theta+\phi_V)\big) + \sum_{n=1}^{N} V_n^s \, sin\big(n\sigma_V(\theta+\phi_V)\big) \quad (5.9)$$

with θ being the rotation angle, σ_v the internal symmetry number of the rotation, ϕ a phase factor and V_n^c and V_n^s the fitted coefficient. In our fit

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

we have employed a total 11 coefficients (N=5) and a value of ϕ equal to 0. The results of our calculations are summarized in Figure 5.7.

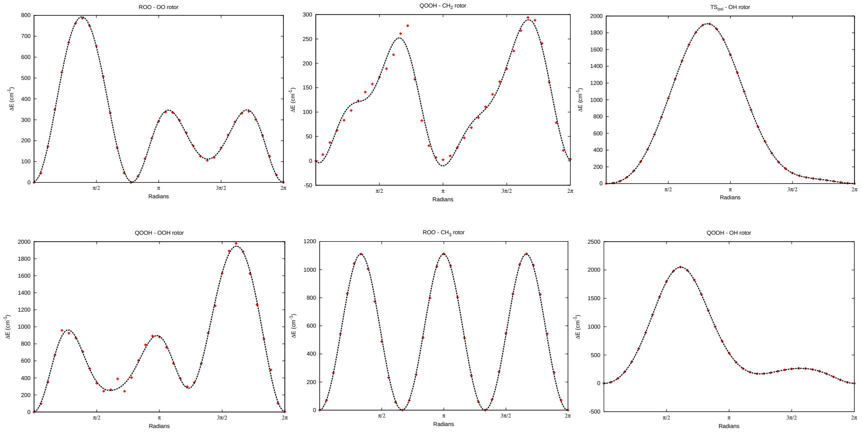

Figure 5.7: Plot of the calculated rotational potential energy values (red points) and related fitted curves (solid line) for the 6 identified internal rotations.

The calculation of the moments of inertia has been performed for the most stable geometry and is assumed fixed along the rotational path. Once each hindered rotor has been characterized and the associated harmonic vibrational frequency corrected with the proper functional form, the Thermo code evaluates the contribution of the anharmonicities to the total molecular partition function and finally calculates the required thermodynamical properties.

Kinetic parameters

Kinetic parameters are the key quantities for the development of the kinetic models. For each reaction involved in the mechanism, in fact, we have to provide the rate coefficient of the process and its dependence from temperature. The forward rate constant for the *i*th reaction is generally assumed to have the following Arrhenius temperature dependence:

$$k_{fi} = A_i \, T^{\beta_i} \, exp\left(\frac{-E_i}{RT}\right) \tag{5.10}$$

ISSN: 2279-8773

VIRT&L–COMM.9.2016.4

where the exponential factor, A_i, the temperature exponent, β_i, and activation energies E_i are specified. These are obtained by calculating the high-pressure limit kinetic coefficients at different initial T, with one of the available rate theory methods illustrated in section 1.4, and then fitting the results using Eq. 5.10. The reverse rate coefficient are then related to the forward ones via the equilibrium constant determined from the thermodynamic properties worked out previously.
In the high-pressure regime the kinetic coefficients are independent form pressure, however in certain conditions, pressure effect must be taken into account. Depending on how the rate coefficients changes with pressure, several methods of representing the rate expressions can be adopted [139, 140] and auxiliary data for the pressure dependence are required.

In this work the kinetic parameters for the investigated combustion process have been calculated using the RRKM/ME method implemented in the MultiWell program. This code is based on a modified version of the Gillespie stochastic method [141] for the resolution of 1D "energy grained" Master Equation. In order to carry out Master equation simulations, the energy distribution of the initiated species (in a multi-well system) and collision rate coefficients for energy transfer must be included. Moreover sum and densities of states of each stationary point must be also provided. These are commonly computed by assuming that vibrations, rotations, and internal rotations are separable. To carry out efficient "exact counts" of states for separable degrees of freedom we have employed the DenSum code. The "active" degrees of freedom include vibrations, internal rotations, the "K-rotor" (the rotation associated with the symmetry axis (the K quantum number) of a symmetric top). As for the Thermo code, we had to correct the vibrational frequencies associated with internal rotation (see above). It is useful to note that accurate predictions of pressure dependent rate coefficients obtained from ME/RRKM calculations still depend on semi-empirical factors such as collisional energy transfer. The real time needed for a calculation of this type depends mostly upon the number of collisions occurring during a simulated time period and on the number of stochastic trials needed to achieve the desired precision.
These results, together with the previous calculated energy barriers (see Tabel 5.1), have been then used to build the MultiWell input file for the investigated process. The dependence from temperature of the infinite pressure limit rate coefficients, k_∞, have been worked out for the forward and

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

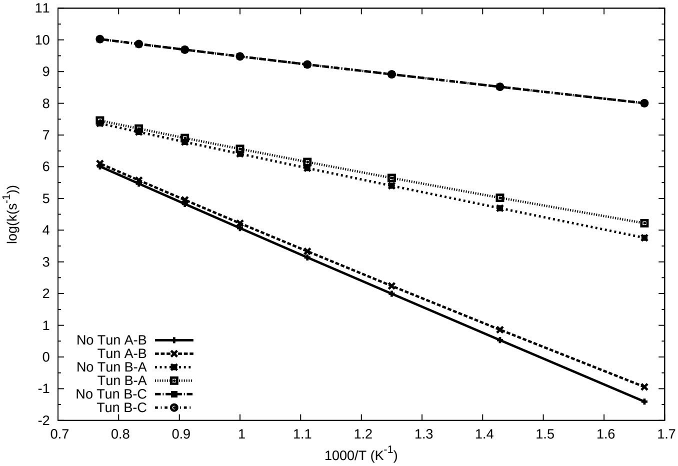

Figure 5.8: Infinite pressure limit kinetic coefficients plotted as a function of temperature for the *ROO-QOOH* (A-B), *QOOH-ROO* (B-A) and *QOOH-CE* (B-C) reactions. The effect of the inclusion (Tun) or not-inclusion (No Tun) of quantum tunneling is also shown.

reverse $ROO^{\bullet}$ isomerization reactions as well as for the cyclic ether (CE) formation.
Figure 5.8 shows the calculated k_{∞} at initial T=600, 700, 800, 900, 1000, 1100, 1200 and 1300 K, for the A-B, B-A and B-C reactions of equation 5.1. Quantum mechanical tunneling corrections to the microcanonical rate coefficients have been introduced using an asymmetric Eckart potential [142]. As we can seen from the plots, in fact, the forward and reverse isomerization reactions (A-B and B-A), which involve a hydrogen atom transfer, are strongly affected by quantum tunnel effects and the Tun and No Tun rate coefficients are well separated especially for low temperature where the reactivity increases considerably if this effect is taken into account. On the other hand the two lines for the Cyclic Ether formation reaction (B-C) superpose perfectly suggesting that this process is not affected by quantum tunel.

The results of this work show clearly that the interleaving of HP and

ISSN: 2279-8773

VIRT&L-COMM.9.2016.4

HT computing can improve considerably the efficiency of the work done for evaluating the kinetic and thermochemical parameters of simple reactions. For this purpose the GEMS molecular simulator is found to be a good instrument for the assemblage of a complex kinetic models, and the facilities offered by COMPCHEM, such as GriF, turn out to be very useful for the treatment of the large number of reactions which compose that model.

Moreover, additional effort has been spent on the implementation, within the GEMS workflow, of the apparatuses adopted for the modelling of global gas phase combustion processes. In collaboration with the *Laboratoire Réactions et Génie des Procédés* of Nancy (FR), we have, in fact, implemented on the Grid computing a software for the simulation of the time-dependent chemical kinetics behavior of a homogeneous gas mixture in a closed system. Usually, the observable properties obtained from these softwares must be calculated for a large set of initial conditions (starting temperature, composition of the fuel, morphology of the burner, etc.). The availability of a large number of computing elements, as typical of the Grid, enables the distribution of independent calculations on different nodes and drastically reduce the elevated time of a complete simulation. This job is particularly suited for "sensitivity analysis" of combustion kinetic models. These studies require, in fact, the execution of a large number of independent jobs in which each time a different kinetic parameter, of a different reaction, is modified. Comparing the resulting observable properties (such as the ignition delay time, outcomes molar fraction, etc.) with the one obtained from the original mechanism, the impact of the single elementary process on the whole reactivity can be extrapolated and a better understanding of the combustion process can be achieved.

Currently, a code for the brute-force sensitivity analysis on the GRID is available trough the GriF science gateway, to facilitate massive calculations that larger models, including thousand reactions, require.

ISSN: 2279-8773

ISSN: 2279-8773

Bibliography

[1] A. Laganà, A. Riganelli, O. Gervasi, On the structuring of computational chemistry Virtual Organization COMPCHEM, Lecture Notes in Computer Science 3980, 665-674 (2006).

[2] The European Grid Initiative, http://web.eu-egi.eu

[3] A. Laganà Towards a grid based universal molecular simulator, in A. Laganà , G. Lendvay Eds "Theory of the dynamics of elementary chemical reactions", 363-380 (Kluwer, 2004).

[4] A. Laganà, A. Costantini, O. Gervasi, N. Faginas Lago, C. Manuali, S. Rampino. Journal of Grid Computing 8, (2010) 571.

[5] S. Rampino, A. Monari, E. Rossi, S. Evangelisti, A. Laganà, Chem. Phys. 398, (2012) 192.

[6] Molecular Quantum Mechanics Parts I and II: An Introduction to QUANTUM CHEMISTRY (Volume 1), P.W. Atkins, Oxford University Press, 1977.

[7] A. Szabo, N.S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory, Courier Dover, 1996.

[8] D. G. Truhlar, B. C. Garrett and S. J. Klippenstein, J. Chem. Phys., 100, (1996) 12771.

[9] J. A. Miller, S. J. Klippenstein, S. H. Robertson, M. J. Pilling and N. J. B. Green, Phys. Chem. Chem. Phys., 11, (2009) 1128.

VIRT&L-COMM.9.2016.4

[10] D. M. Golden, J. R. Barker, Combust flame, 158, (2011) 602.

[11] H. Eyring, J. Chem. Phys., 3 (1935) 107.

[12] J. Pfaendtner, X. Yu, L. J. Broadbelt, Theor. Chem. Account, 118 (2007) 881.

[13] K. S. Pitzer and W. D. Gwinn, J. Chem. Phys., 10 (1942) 428.

[14] PRACE Research Infrastructure, http://www.prace-project.eu (cited February 13, 2012).

[15] TeraGrid Archives, https://www.xsede.org/tg-archives (cited February 13, 2012).

[16] The Message Passing Interface (MPI) standard, http://www.mcs.anl.gov/research/projects/mpi (cited February 13, 2012).

[17] Manuali, C., Rampino, S., Laganà, Comp. Phys. Comm. 181, (2010) 1179.

[18] Manuali, C., Lagana‘, A., 27 (2011) 315.

[19] Italian Grid Infrastructure, http://grid.infn.it (cited February 13, 2012).

[20] CINECA, http://www.cineca.it (cited February 13, 2012).

[21] GAMESS-US. M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S. Su, T.L. Windus, M. Dupuis, J.A. Montgomery J. Comput. Chem., 14, 1347-1363(1993). See http://www.msg.ameslab.gov/gamess/.

[22] European Middleware Initiative, http://www.eu-emi.eu (cited February 13, 2012).

[23] EMI Execution Service page, https://twiki.cern.ch/twiki/bin/view/EMI/EmiE
(cited February 13, 2012).

[24] The COST/D37 action:
http://www.cost.eu/domains_actions/cmst/Actions/Grid_Computing_in_Ch

ISSN: 2279-8773

[25] C. Angeli, G. L. Bendazzoli, S. Borini, R. Cimiraglia, A. Emerson, S. Evangelisti, D. Maynau, A. Monari, E. Rossi, J. Sanchez-Marin, P. G. Szalay, A. Tajti, The problem of interoperability: a common data format for Quantum Chemistry codes, International Journal of Quantum Chemistry 107, 2082-2091 (2007).

[26] Borini, S., Monari, A., Rossi, E., Tajti, A., Angeli, C., Bendazzoli, G.L., Cimiraglia, R., Emerson, A., Evangelisti, S., Maynau, D., Sanchez-Marin, J., Szalay, P.G.: J. Chem. Inf. Model. 47, (2007) 1277.

[27] A. Scemama, A. Monari, C. Angeli, S. Borini, S. Evangelisti, E. Rossi, Lecture Notes in Computer Science, 5072, (2008) 1094.

[28] S. Rampino COST-STSM-D37-5218 The issue of interoperability in GEMS: designing a D5cost data model.

[29] Rampino S., Monari A., Evangelisti S., Rossi E., Ruud K., Laganà A. Proceedings of the Cracow Grid Workshop CGW09, 164 (2010).

[30] Rampino S., Pirani F., Garcia E., Laganà A., IJWGS 6 (2010) 196.

[31] MOLPRO is a package of ab initio programs written by H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklaß, D. P. O'Neill, P. Palmieri, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, A. Wolf , see http://www.molpro.net/.

[32] The HDF Group. Hierarchical data format version 5, 2000-2010. http://www.hdfgroup.org/HDF5.

[33] Ugo Varetto; Swiss National Supercomputing Centre: Lugano.

[34] GaussView, Version 5, Roy Dennington, Todd Keith and John Millam, Semichem Inc., Shawnee Mission KS, 2009.

[35] Avogadro: an open-source molecular builder and visualization tool. http://avogadro.openmolecules.net/

VIRT&L-COMM.9.2016.4

[36] T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, J. Olsen, K. Ruud, H.Ågren, T. Andersen, K. L. Bak, V. Bakken, O. Christiansen, P. Dahle, E. K. Dalskov, T. Enevoldsen, H. Heiberg, H. Hettema, D. Jonsson, S. Kirpekar, R. Kobayashi, H. Koch, K. V. Mikkelsen, P. Norman, M. J. Packer, T. Saue, P. R. Taylor and O. Vahtras, DALTON, an ab initio electronic structure program, Release 1.0, 1997.

[37] MOLCAS, G. Karlström, R. Lindh, P.-Å. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrády, L. Seijo, Computational Material Science, 28, 222 (2003). See http://www.molcas.org/.

[38] H. Lischka, R. Shepard, I. Shavitt, R. M. Pitzer, M. Dallos, Th. Müller, P. G. Szalay, F. B. Brown, R. Ahlrichs, H. J. Böhm, A. Chang, D. C. Comeau, R. Gdanitz, H. Dachsel, C. Ehrhardt, M. Ernzerhof, P. Höchtl, S. Irle, G. Kedziora, T. Kovar, V. Parasuk, M. J. M. Pepper, P. Scharf, H. Schiffer, M. Schindler, M. Schüler, M. Seth, E. A. Stahlberg, J.-G. Zhao, S. Yabushita, Z. Zhang, M. Barbatti, S. Matsika, M. Schuurmann, D. R. Yarkony, S. R. Brozell, E. V. Beck, and J.-P. Blaudeau, M. Ruckenbauer, B. Sellner, F. Plasser, and J. J. Szymczak, COLUMBUS, an ab initio electronic structure program, release 7.0 (2012).

[39] D. Skouteris, J.F. Castillo, D.E. Manolopoulos, Computer Physics Communications 133 (2000) 128.

[40] A. Aguado, C: Tablero, M. Paniagua, Computer Physics Communications, 108 (1998) 259.

[41] A. Aguado, M. Paniagua, Journal of Chemical Physics, 96 (1992) 1265 .

[42] NEPTUNUS is a Quantum-Chemistry FORTRAN code for the calculation of FCI and CAS- CI energies and properties written by G.L. Bendazzoli and S. Evangelisti, with contributions from L. Gagliardi, E. Giner, A. Monari and M. Verdicchio.

[43] QMC=CHEM. See http://qmcchem.ups-tlse.fr/

[44] The TRAC project hompage http://trac.edgewall.org.

ISSN: 2279-8773

[45] The online home of the Apache Subversion^{TM} software project http://subversion.apache.org

[46] M.E. Alikhani, S. Shaik, Theor. Chem. Acc. 116 (2006) 390.

[47] S.P. de Visser, D. Danovich, S. Shaik, Phys. Chem. Chem. Phys. 5 (2002) 158.

[48] S. P. de Visser, Y. Alpert, D. Danovich, S. Shaik, J. Phys. Chem. A 104 (2000) 11223. .

[49] D. Danovich, W. Wu, and S. Shaik, J. Am. Chem. Soc. 121, (1999) 3165.

[50] S. P. de Visser, D. Danovich, W. Wu, and S. Shaik, J. Phys. Chem. A 106, (2002) 4961.

[51] A. Monari, J. Pitarch-Ruiz, G.L. Bendazzoli, S. Evangelisti, J. Sanchez-Marin, J. Chem. Theory Comp. 4 (2008) 404.

[52] A. Monari, J. Pitarch-Ruiz, G.L. Bendazzoli, S. Evangelisti, J. Sanchez-Marin, Int. J. Quant. Chem. 874 (2010) 110.

[53] M. Verdicchio, S. Evangelisti, T. Leininger, A. Monari, Chem. Phys. Lett. 503 (2011) 215 .

[54] H. A. Jahn and E. Teller, Proc. Roy. Soc. 161 (1937) 220 .

[55] P. J. Knowles, C. Hampel, and H.-J. Werner, J. Chem. Phys. 99 (1993) 5219 .

[56] B. O. Roos, V. Veryazov, and P. O. Widmark, Theor. Chem. Acc. 111 (2003) 345.

[57] B. O. Roos, R. Lindh, P. A. Malmqvist, V. Veryazov, and P. O. Widmark, J. Phys. Chem. 108 (2005) 2851 .

[58] D. Feller, J. Comp. Chem. 17 (1996) 1571 .

[59] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, J. Chem. Inf. Model 47 (2007) 1045.

[60] D. Danovich, S. Shaik, J. Chem. Theor. Comp. 6 (2010) 1479.

VIRT&L–COMM.9.2016.4

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

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

[63] Salomonson, S.; Wartson, H.; Lindgren, I. Phys. Rev. Lett. 76, (1996) 3092.

[64] Kramida, A.; Martin, W. C. J. Phys. Chem. Ref. Data 26, (1997) 1185.

[65] Bernath, P. F.; Shayesteh, A.; Tereszchuk, K.; Colin, R. Science, 297, (2002) 1323.

[66] Rienstra-Kiracofe, J. C.; Tschumper, G. S.; Schaefer III, H. F. Chem. Rev., 102 (2002) 231.

[67] Bendazzoli, G. L.; S., E.; Passarini, F. Chem. Phys., 215 (1997) 217.

[68] Pauling, L. The Chemical Bond; Cornell University Press, Ithaca, New York, 1967.

[69] Angeli, C.; Cimiraglia, R. *private communication*.

[70] K. Peterson, *private communication*.

[71] Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. J. Chem. Phys., 106 (1997) 9639.

[72] Halkier, A.; Helgaker, T.; Jørgensen, W., P. and Klopper; Koch, H.; Olsen, J.; Wilson, A. K. Chem. Phys. Lett., 286 (1998) 243.

[73] Stone, A. J. The theory of intermolecular forces; Claredon Press, Oxford, 1996.

[74] Bottcher, C. J., Bordewijk, P. Theory of electric polarization vols. II; Elsvier, New York, 1980.

[75] Nesbitt, D. J., Chem. Rev., 88 (1988) 843.

[76] Weber, A. Structure and Dynamics of Weakly Bound Molecular Complexes (NATO Science Series C) (1987).

[77] Capitelli, M., Springer-Verlag, Berlin (1986).

ISSN: 2279-8773

[78] Armenise, I., Capitelli, M., Garcia, E., Gorse, C., Lagana', A. and Longo, S. Chem. Phys. Lett., 200 (1992) 597.

[79] Knauth, D. C., Andersson, B. G., McCandliss, S. R. and Moos, H. W. Nature, 429 (2004) 636.

[80] Cappelletti, D., Vecchiocattivi, F., Pirani, F. and McCourt, F. R. W., Chem. Phys. Lett., 248 (1996) 237.

[81] Huo, S. W. M. and Green, S., J. Chem. Phys., 104 (1996) 7572.

[82] Stallcop, J. R. and H., P. Chem. Phys. Lett., 281 (1997) 212.

[83] Lagana', A., Balucani, N., Crocchianti, S., Casavecchia, P., Garcia, E. and Saracibar, A., Lecture Notes on Computer Science, 6784 (2011) 453.

[84] A. van der Avoird, P.E.S. Wormer, A.P.J. Jansen, J. Chem. Phys. 84 (1986) 1629.

[85] Møller, C. and Plesset, M. S., Phys. Rev., 46 (1934) 618.

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

[87] Piecuch, P., Kucharski, S. A., Kowalski, K. and Musial, M., Comput. Phys. Comm., 149 (2002) 71.

[88] Bentz, J. L., Olson, R. M., Gordon, M. S., Schmidt, M. W. and Kendall, R. A., Comput. Phys. Comm., 176 (2007) 589.

[89] Olson, R. M., Bentz, J. L., Kendall, R. A., Schmidt, M. W. and Gordon, M. S. A, J. Comput. Theo. Chem., 3 (2007) 1312.

[90] Laganà, A., Computer and Chemistry, 4 (1980) 137.

[91] Lee, T. J. and Rice, J. E., J. Chem. Phys., 94 (1991) 1215.

[92] Sorbie, K. S. and Murrell, J. N., Mol. Phys., 52 (1975) 1387.

[93] Aguado, A., Tablero, C. and Paniagua, M. Comput. Phys. Comm., 134 (2001) 97.

[94] Aguado, A., Suarez, C. and Paniagua, M., J. Chem. Phys., 101 (1994) 404.

VIRT&L-COMM.9.2016.4

[95] Garcia, E., Saracibar, A., Gomez-Carrasco, S. and Laganá, Phys. Chem. Chem. Phys., 10 (2008) 2552.

[96] Caridade, P.J.S.B., Galvao, B.R.L. and Varandas, A.J.C. J. Phys. Chem. A, 114 (2010) 6063.

[97] E. J. Baerends, D. E. Ellis, P. Ros, Chem. Phys., 2 (1973) 41.

[98] T. Ziegler, J. G. Snijders and E. J. Baerends, J. Chem. Phys., 74,(1981) 1271.

[99] K.Hirao, Chem.Phys.Lett., 190 (1992) 374.

[100] K.Hirao, Int.J.Quant.Chem., S26 (1992) 517.

[101] R. P. Wayne, Chemistry of Atmosphere, 3rd ed., University Press, Oxford, 2000.

[102] A. C. Köne and T. Buke, Renew. Sustain. Energy Rev.,14 (2010) 2906.

[103] F. Battin-Leclerc, E. Blurock, R. Bounaceur, R. Fournet, P.-A. Glaude, O. Herbinet, B. Sirjean, and V. Warth. Chemical Society Reviews, 40 (2011) 4762.

[104] J. Simmie, Prog. Energy Combust. Sci., 29 (2003) 599.

[105] W. J. Pitz and C. J. Mueller, Prog. Energy Combust. Sci., (2011) 330.

[106] J. Y. W. Lai, K. C. Lin and A. Violi, Prog. Energy Combust. Sci., 37 (2011) 1.

[107] K. Kohse-Höinghaus, P. Oswald, T. A. Cool, T. Kasper, N. Hansen, F. Qi, C. K. Westbrook and P. R. Westmoreland, Angew. Chem., Int. Ed., 49 (2010) 3572.

[108] A. S. Tomlin, T. Turanyi and M. J. Pilling, Comprehensive Chemical Kinetics: Oxidation Kinetics and Autoignition of Hydro- carbons, ed. M. J. Pilling, Elsevier, 1997, p. 35.

[109] E. Ranzi, T. Faravelli, P. Gaffuri, and A. Sogaro. Combustion and Flame, 102 (1995) 179.

ISSN: 2279-8773

[110] Roberta G. Susnow, Anthony M. Dean, William H. Green, P.Peczak, and Linda J. Broadbelt. The Journal of Physical Chemistry A, 101 (1997) 3731.

[111] S. Touchard, R. Fournet, P. A. Glaude, V. Warth, F. Battin- Leclerc, G. Vanhove, M. Ribaucour and R. Minetti, Proc. Combust. Inst., 30 (2005) 1073.

[112] E. S. Blurock, J. Chem. Inf. Comput. Sci., 35 (1995) 607.

[113] M. J. Pilling, Proc. Combust. Inst., 32 (2009) 27.

[114] B. J. Finlayson-Pitts and J. N. Pitts, Atmospheric chemistry: fundamentals and experimental techniques, Wiley, New York, 1986.

[115] S. H. Robertson, P. W. Seakins and M. J. Pilling, in Comprehensive Chemical Kinetics: Oxidation Kinetics and Autoignition of Hydrocarbons, ed. M. J. Pilling, Elsevier, 1997.

[116] K. A. Bhaskaran and P. Roth, Prog. Energy Combust. Sci., 28 (2002) 151.

[117] S. Drennan, Chem. Eng. Prog., 104 (2008) 43, http://www. reactiondesign.com/products/open/chemkin.html

[118] COSILAB, The Combustion Simulation Laboratory, Version 3.0., 2009, Rotexo GmbH & Co. KG, Haan, Germany. http://www.rotexo.com

[119] D. G. Goodwin, Cantera C++ Users Guide, 2002. http:// sourceforge.net/projects/cantera.

[120] W. J. Pitz , R. D. Wilk , C. K. Westbrook, N.P. Cernansky N. P. Paper No. WSSCI 88-51, Western States Sections/The Combustion Institute Spring Meeting, 1988.

[121] F. Battin-Leclerc, Prog. Energy Combust. Sci., 34 (2008) 440.

[122] M. Cord, B. Sirjean, R. Fournet, A. Tomlin, M.F. Ruiz-Lopez, F. Battin-Leclerc, J. Phys. Chem. A, 116 (2012) 6142.

[123] J. A. Montgomery Jr., M. J. Frisch, J. W. Ochterski and G. A. Petersson, J. Chem. Phys.,112 (2000) 6532.

VIRT&L–COMM.9.2016.4

[124] J. A. Montgomery Jr., M. J. Frisch, J. W. Ochterski and G. A. Petersson, J. Chem. Phys., 110 (1999) 2822.

[125] J. W. Ochterski, G. A. Petersson and J. A. Montgomery Jr., J. Chem. Phys., 104 (1996) 2598.

[126] L. A. Curtiss, K. Raghavachari, G. W. Trucks and J. A. Pople, J. Chem. Phys., 94, (1991) 7221.

[127] L. A. Curtiss, P. C. Redfern and K. Raghavachari, J. Chem. Phys., 126 (2007) 084108.

[128] H.-J. Werner, Mol. Phys. 89 (1996) 645-661.

[129] P. Celani and H.-J. Werner, J. Chem. Phys. 112 (2000) 5546.

[130] X.. You, H. Wang, E. Goos, C.J. Sung, S.J. Klippenstein, J. Phys. Chem. A, 111 (2007) 4031.

[131] J.A. Miller, S.J. Klippenstein, S.H. Robertson, Proc. Comb. Inst., 28 (2000) 1479.

[132] MultiWell-2011.1 Software, 2011, designed and maintained by John R. Barker with contributors Nicholas F. Ortiz, Jack M. Preses, Lawrence L. Lohr, Andrea Maranzana, Philip J. Stimac, T. Lam Nguyen, and T. J. Dhilip Kumar; University of Michigan, Ann Arbor, MI; http://aoss.engin.umich.edu/multiwell/.

[133] S. E. Stein, B. S. Rabinovitch, J. Chem. Phys. 58 (1973) 2438.

[134] S. Gordon and B. J. McBride, NASA Report SP-273 (1971).

[135] D. A. McQuarrie, J. D. Simon, Molecular Thermodynamics, University Science Books, 1999.

[136] V. van Speybroeck, , R. Gani, R. J. Meier. Chemical Society Reviews 39, (2010) 1764.

[137] W. J. Moore, Physical Chemistry, Longman, London, 1972.

ISSN: 2279–8773

[138] Gaussian 09, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2010.

[139] F. Lindemann, Trans. Faraday Soc., 17 (1922) 598.

[140] R. G. Gilbert, K. Luther, J. Troe, Ber. Bunsenges. Phys. Chem., 87 (1983) 169.

[141] J. R. Barker, Int. J. Chem. Kinetics, 33 (2001) 232.

[142] C. Eckart, Phys. Rev., 35 (1930) 1303.