Extrapolation in Theoretical Chemistry: Approximating Reality from Matters-of-fact

A. J. C. VARANDAS

ACADEMIA DAS CIÊNCIAS
DE LISBOA

Extrapolation in Theoretical Chemistry: Approximating Reality from Matters-of-fact

A. J. C. Varandas*

RESUMO

The use of extrapolation methods in Quantum Chemistry is overviewed. Two topics are addressed, both lying at the heart of chemistry: electronic structure and quantum reaction dynamics. In the first, we address the problem of extrapolating the energy obtained by solving the Schrödinger equation to the limit of a complete one-electron basis set. With the uniform-singlet-and-triplet-extrapolation (USTE) scheme at the focal point, the emphasis is on recent updates. Secondly, we address the extrapolation of quantum mechanical reactive scattering probabilities from zero total angular momentum to all J values by running quasiclassical trajectories, QCT/QM-αJ. Some implications and prospects for future work conclude the overview.

Analiza-se o uso de métodos de extrapolação em Química Quântica focando em duas áreas fundamentais: estructura electrónica e dinâmica quântica. Sobre a primeira, tal necessidade advém do facto dos valores próprios da energia serem obtidos por solução da equação de Schrödinger com bases finitas, requerendo por isso extrapolação para o limite da base completa (infinita). Com ênfase no método USTE (Uniform-Singlet-and-Triplet-Extrapolation), limita-se por brevidade a análise aos estudos mais recentes. Sobre a segunda, discute-se o método QCT/QM-αJ com vista ao cálculo de secções eficazes e constantes de velocidade, *i.e.* extrapolação pelo método das trajectórias quasi-clássicas (QCT) para qualquer valor do momento angular total das probabilidades de reacção calculadas para momento angular total nulo J = 0 por métodos mais elaborados (e por isso dispendiosos) da mecânica quântica (QM). Sumariam-se as implicações e linhas de actuação futura.

I. INTRODUCTION

It is well known that no problem in physics and chemistry can ever be solved exactly. Approximations are unavoidable, with the adiabatic approximation due to Born and Oppenheimer[1] (BO) being most fundamental. Owing to the disparity of masses (nuclei are at least 1837 times heavier than electrons), BO proposed that their motions could be disentangled and considered separately: the nuclei are said to move adiabatically governed by the potential energy surface (PES) created by the electrons. In fact, even a smaller disparity of masses than the one encountered between a proton and an electron may suffice to justify such an approximation. By considering four equally charged fermions, two positive and two negative, we have shown[2] that a mass ratio of 200 (which is approximately the muon to electron mass ratio) validates the separation of their motions up to ~ 80 %.

* School of Physics and Physical Engineering, Qufu Normal University, 273165 Qufu, China and Department of Chemistry, and Chemistry Centre, University of Coimbra, 3004-535 Coimbra, Portugal. Electronic address: varandas@uc.pt

As an outcome of the BO approximation, the electronuclear Schrödinger equation splits into two: one for the electrons moving in a fixed arrangement of the nuclei (electronic Schrödinger equation: eSE), the other for the nuclei moving in the PES created by the electrons (nSE). In this adiabatic formulation of molecular quantum mechanics, one solves pointwise the eSE to get the PES, and then the nSE to obtain the ro-vibrational energy levels of the supermolecule (thence any observable of the system depending on them: bulk properties, cross sections in molecular scattering, and rate constants in chemical kinetics just to mention a few). Because the nuclei have commonly heavy masses, their motion is also often approximated by solving the fundamental laws of classical mechanics, *i.e.*, Hamilton's equations.[3]

Two major difficulties (known as "explosions") arise in theoretical and computational chemistry.[4]
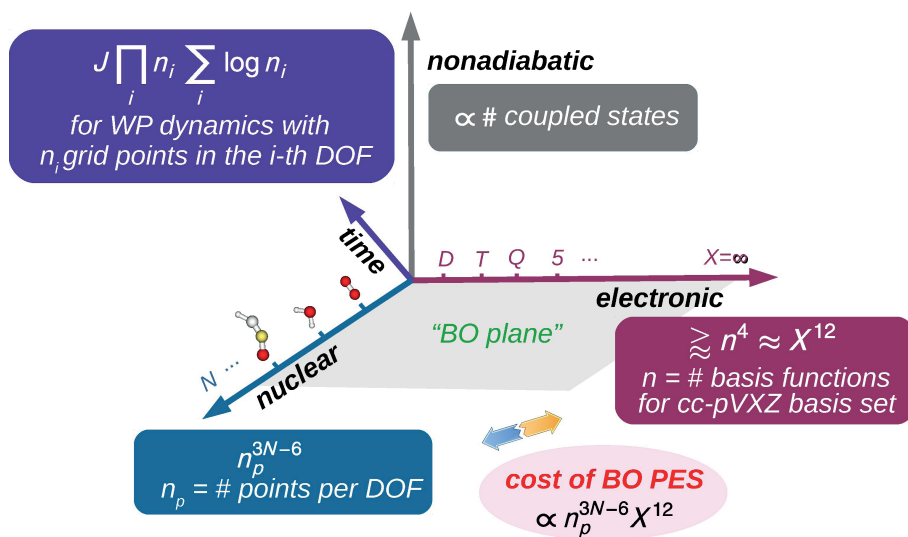

Fig. 1: "Explosions" in molecular computational chemistry. The formulas illustrate the cost scaling with physical variables such as the total angular momentum (J), number of points n_p calculated per degree of freedom (DOF), and hierarchical number of the one-electron basis set (X).

As shown in Figure 1, the so-called n_p^{3N-6} explosion indicates how many points need to be calculated to represent the PES of a N-atom system if n_p points are required per dimension. In turn, the X^{12} explosion tells how the cost/point raises with basis set enhancement which is commonly implied by its cardinal number X.[5] Added to such "explosions" is the fact that reliable results in reaction dynamics demand PESs with chemical accuracy (< 1 kcal mol^{-1}), which implies the use of large basis sets (typically quadruple-zeta, QZ, or larger). If the aim is accurate vibrational calculations, then a PES with spectroscopic accuracy (≤ 1 cm^{-1}) is needed, which is an even more difficult task. Ideally, both require PESs calculated at the one-electron complete basis set (CBS) level. Extrapolation methods are then mandatory as schematized in Figure 2. Embedding in principle no empiricism, this can be done at a pure mathematical level (using polynomials and rational fractions) or, preferably, guided by physically motivated asymptotic theories as here endorsed.

As Figure 2 indicates, extrapolation is also key in reaction dynamics[6], a topic at the core of chemical kinetics which is discussed in section III. Because the correspondence principle states that for very large
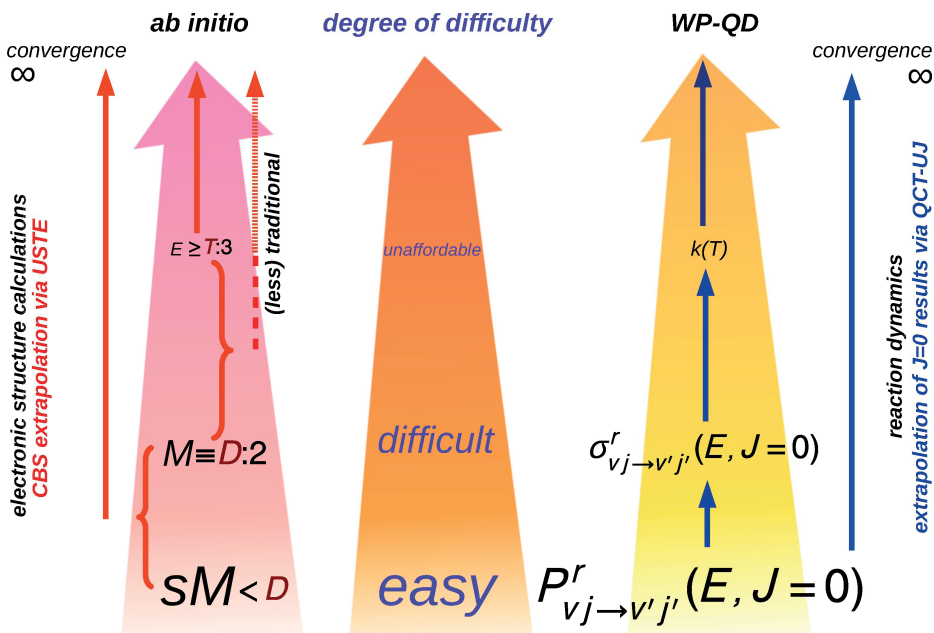
Fig. 2: Extrapolation in electronic structure calculations and quantum reaction dynamics. The arrow in dashed indicates what is traditionally done in electronic structure.

quantum numbers the laws of quantum theory merge with those of classical physics[7], classical mechanics is shown to be of help on this endeavour. In the following, we provide a brief survey of our progress in both such topics.

II. EXTRAPOLATION IN ELECTRONIC STRUCTURE

Methods for solving the eSE have long been of utmost importance in computational molecular science.[8,9] Without any attempt to completeness, the most popular ones are sketched in Figure 3, including valence-bond (VB) type and semi-empirical. The simplest is Hartree-Fock (HF), a mean-field theory where electron correlation is ignored. The error due to its disregard is significant, with other more sophisticated single-reference (SR) MO-based approaches being of variational (configuration interaction, CI), perturbative Møller-Plesset (or many-body perturbation theory type assumed at second-order, MP2), and couple-cluster (CC) types. Among them, the CC singles and doubles with perturbative triples method, CCSD(T), is commonly viewed as the golden rule of quantum chemists due to its accuracy and affordability even for moderate-sized systems. Yet, SR methods miss the static (nondynamical) correlation, which can only be recovered with multireference (MR) methods. These may use complete-active--space-self-consistent-field (CASSCF or CAS) or MRCI wave functions, the latter commonly accounting for the dynamical correlation by inclusion of single and double excitations (MRCISD). Because integration of the eSE with accuracy at demand is unreachable, extrapolation of raw energies to converged values will then be key and the focal point here.

As noted above, the need for going beyond HF arises from the correlated motion of the electrons. Because the electron-electron repulsion operator shows a singularity at $r_{12}=0$, the exact wave function must have a discontinuous derivative as implied by Kato's[10] cusp condition. However, conventional
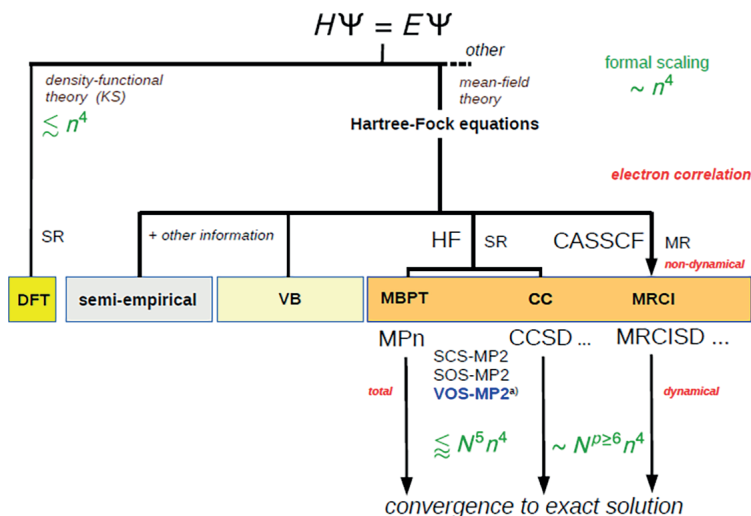
Fig. 3: Some popular methods in electronic structure theory. For acronyms, see text and references therein.

methods expand the wave function in terms of Slater determinants built from orbitals, and hence fail to obey such a condition. This may explain the extremely slow convergence of the conventional correlated methods. Although the purist approach to solve this problem would be to make the many-electron wave function depend explicitly on r_{12}[9,11], the alternative is to systematize the error made by conventional methods and seek an extrapolation to predict the inherent error.

From a partial-wave expansion for two-electron atoms, Kutzelnigg and Morgan[12] established the following: a) for natural-parity singlet states, the leading contribution to the energy at second-order of perturbation theory is $\propto (\ell+\frac{1}{2})^{-4}$ with no odd-terms either $\propto (\ell+\frac{1}{2})^{-5}$ or $\propto (\ell+\frac{1}{2})^{-7}$; b) for triplet states, the leading term is $\propto (\ell+\frac{1}{2})^{-6}$. The above results were also shown to be valid for the MP2 energy of atoms with any number of electrons.[12,13] Such findings followed Schwartz' pioneering work[14] who analyzed the partial-wave expansion of the second-order energy expression and obtained only the coefficients of powers -4 and -6, since the $(l + 1/2)^{-5}$ term vanishes. If $\Delta E_l \propto \sum_{m=4} a_m(l + 1/2)^{-m}$, the convergence error when all the terms with $l \geq L$ are omitted will then assume the form:

$$\Delta E = \sum_l^\infty \Delta E_l \sim \sum_{m=4} a_m \int_{L+\frac{1}{2}}^\infty \left(l+\frac{1}{2}\right)^{-m} dl$$
$$= \sum_{m=4} A_{m-1}(L+1)^{-m+1} \tag{1}$$

where the first two leading asymptotic terms are $A_3(L + 1)^{-3}$ and $A_5(L + 1)^{-5}$. Considering just the first may be accuracy-limiting,[13] an issue recently dissected through MP2 calculations on He and Ne using basis

sets with converged radial extent and values of l up to $14^{[15]}$. While the form $(L + 1)^{-3}$ was reached for Ne, higher-order terms were normally important in molecular calculations.

Largely motivated by the above, modern basis sets[16–22] are commonly built according to a principal expansion (thus by systematic inclusion of all $\sim n^2$ functions in the next shell), with each shell contributing an amount of energy $\propto n^{-4}$. An example are the popular correlation-consisted basis sets cc-pVXZ (or VXZ), aug-cc-pVXZ (AVXZ) when augmented with diffuse functions, etc., of the Dunning family.[16,17,23] Because the cardinal number ($X=2$:D, 3:T, 4:Q…) is identifiable with $L + 1$, the popular X^{-3} power-law error[24,25] arises. Its slow convergence compounds with the time required for MP2, CCSD, and CCSD(T) calculations which scale as[26] $N^5 N_b^4$, $N^6 N_b^4$, and $N^7 N_b^4$; N_b is the number of basis functions per atom, which for a VXZ basis is $N_b \simeq X^{3.5}$.[25] Note that correlated calculations beyond QZ become unaffordable for many interesting systems, which leave the raw energies far from a safe point for extrapolation with some popular CBS schemes.

We focus on the current work on the correlation energy, with the reader being addressed elsewhere for extrapolation of the Hartree-Fock component.[4] Suffice it to note here that the Hartree-Fock energy is found numerically and analytically to converge exponentially.[27] In fact, CBS extrapolation from many basis sets beyond the minimal is feasible,[4,28] with direct convergence of the raw *ab initio* energy being also affordable in many cases.

A. The USTE scheme: update 2018

Two ways stand to obtain accurate energies: solution of the eSE after explicit introduction of correlation in the wave function,[9] and exploitation of the convergence of hierarchized correlation consistent basis sets toward the CBS limit. Despite the fast convergence of the former (believed to be as high as $\propto X^{-7}$ but also noted as $\propto X^{-3.8}$), such R12/F12 methods appear to perform inefficiently with small basis sets.[9] Besides, both conventional and the so-called R12/F12 methods converge to the same asymptotic energy, with CBS extrapolation doing well and often outperforming, a merit recognized[29] by the number of CBS schemes (see elsewhere[30] for an extrapolation calculator with some popular ones) vying the above purist techniques.

CBS extrapolation of the electronic energy is best performed by first splitting the total energy into its Hartree-Fock and correlation components. The latter varies asymptotically as an inverse cubic power of X for opposite spin electron pairs and as an inverse quintic power for pairs with the same spin. The USTE scheme[31] accounts for both as

$$E_X^{\mathrm{cor}} = E_\infty^{\mathrm{cor}} + \frac{A_3}{(X + \alpha)^3} \left[1 + \frac{\tau_{53}}{(X + \alpha)^2} \right] \qquad (2)$$

here $\tau_{53} = (A_5^\circ / A_3) + cA_5^{\mathrm{m}-1}$; α, A_5°, and c are parameters to be determined from *ab initio* data. Thus, it is an empirical-free dual-level scheme which shows the correct asymptotic behavior.[4] Besides allowing to determine the CBS energy from only two raw energies for distinct cardinal numbers, it allows a prediction of the (mostly unknown) ratio $\tau_{53} = A_5 / A_3$. When based on the highest affordable cardinal numbers, it allows an accuracy for the CBS extrapolated energy as high as one possibly can get. Because use of the

(5, 6) pair is affordable only for few electron systems, a CBS scheme using at most (D, T) is key for larger systems. This was the goal of GUSTE,[32] where τ_{53} is kept invariant over the configuration space once determined from raw data calculated with $X \geq Q$. Although only one extra raw point is required, the task gets out of reach for large molecules.

Eq. (2) may be rewritten[33,34] as

$$E_X^{cor} = \eta E_\infty^{cor} + \frac{A_3}{\tilde{X}^3} \tag{3}$$

where the hierarchical number x= $\tilde{X}$ is defined by

$$\tilde{X} = (X + \alpha) \left[1 + \frac{A_5/A_3}{(X + \alpha)^2} \right]^{-1/3} \tag{4}$$

and η is a tolerance factor to be defined later, also from the raw *ab initio* data. The concept is then to "educate" the basis such as to account for its deficiencies in composition according to the recovered correlation energy. Stated differently, a hierarchical staircase as straight as possible in X^{-3} will enhance the reliability in extrapolating from any step(s) of the ladder. Although various possibilities exist for the above hierarchical reassignment,[33,35] we have determined[35] the new hierarchical numbers as statistical averages of the values obtained from the condition that the $X \leq 6$ values fall on the straight line obtained by fitting USTE(5, 6) correlation energies.[31] The method has kept the same acronym[35] but specifies the hierarchical number-pair used for the extrapolation: USTE(x_1, x_2); typically, $x_1 = x_2 - 1$. The hierarchical numbers so obtained, $x = d, t, q, p, h\ldots$, are real positive but still universal as they apply equally well to any correlation-consistent-type basis sets.[36] For subminimal (sM) basis sets [smaller than VDZ or AV DZ, which are minimal (M), with larger ones called extended (E)], such hierarchical numbers may not be easy to identify (they are $X = D$ for minimal and $x \geq T$ for extended), the basis may alternatively be indicated. The average deviations $X - x$ so obtained are well mimicked by[4]

$$\Delta\chi = p_0 + \frac{1}{2} \tanh \left[\frac{(X - X_0)}{p_1} \right] \tag{5}$$

where $\Delta\chi = X - x$ and $X_0 = 1.75$; with p_0 constrained at -0.2057, one gets $p_1 = 1.7816$ for MP2 and $p_1 = 0.3678$ for CC. Note that the cardinal number always exceeds the hierarchical number for the CC method, but is smaller for MP2 with the DZ basis (and for MP2 and CC with even smaller X values). Equivalently stated, the DZ basis is statistically better than a true double-zeta for MP2, with a corresponding implication applying to sM basis sets at both MP2 and CC levels of theory. Conversely, larger basis sets perform poorer than the bases they represent. Furthermore, the above results show that x tends to a value ~0.26 smaller than X, both for CC and MP2. Indeed, both methods tend to show the same $X - x$ asymptotic deviation. Thus, Eq. (5) entails the concept that $\Delta\chi$ should not vary with the method (Møller-Plesset *vs* couple-cluster) at the CBS limit.

The most recent version of USTE includes the η tolerance factor.[34] Named $USTE_\alpha(x-1, x)$, where α stands for "analytic", the protocol yields high quality predictions while allowing to CBS extrapolate from any pair of hierarchical numbers. Its reliability has actually been checked by comparing predictions with the best available estimates, also often employed as reference to scrutinize raw energies obtained from MP2-F12 and CCSD(T)-F12 calculations.[37] To enhance further the agreement and delve into subchemical accuracy (< 1 kcal mol^{-1}), the correlation versus x^{-3} straight lines were allowed to slide down by 1.5‰ for MP2 and up by 1.0‰ for CCSD(T) in E_∞^{cor} (*i.e.*: η≈1.0015, and 0.999). Such a slight scaling not only helps to level off the effect of having used CBS($V5Z$, $V6Z$) energies as reference, but also the fact that the calculations were not performed at optimized geometries but at all-electron CCSD(T)/CVTZ ones.[35] Clearly, they enhance the $USTE_\alpha$ agreement with the reference raw F12 energies:[37] rmsd of 0.180 and 0.086 cal mol^{-1} for MP2/CBS and CCSD(T)/CBS. To go beyond this (*i.e.*, to attain spectroscopic accuracy: cm^{-1}) would imply including many other corrections (core and core-valence effects, perturbative contributions for connected quadruple excitations, relativistic effects, etc) which lies outside the scope of model. Because it is highly reliable, the $USTE_\alpha(x-1, x)$ scheme may itself be used to assign a cardinal number to any basis, a key capability just recently explored.[34] With it, basis sets from Huzinaga's MINI to the most advanced extended VXZ-F12 anstazes have been ranked from their ability to capture correlation energy. USTE has also been extensively applied, and much of the early work has been recently reviewed.[4] Most recently, it has been used[38] to assess how correlated MO calculations can perform vs Kohn-Sham density functional theory (DFT) by testing the performance of both methods on the calculation of 38 hydrogen transfer barrier heights and 38 non-hydrogen transfer barrier heights/isomerizations extracted from extensive databases, in addition to four 2p isomerization reactions and six others for large organic molecules.
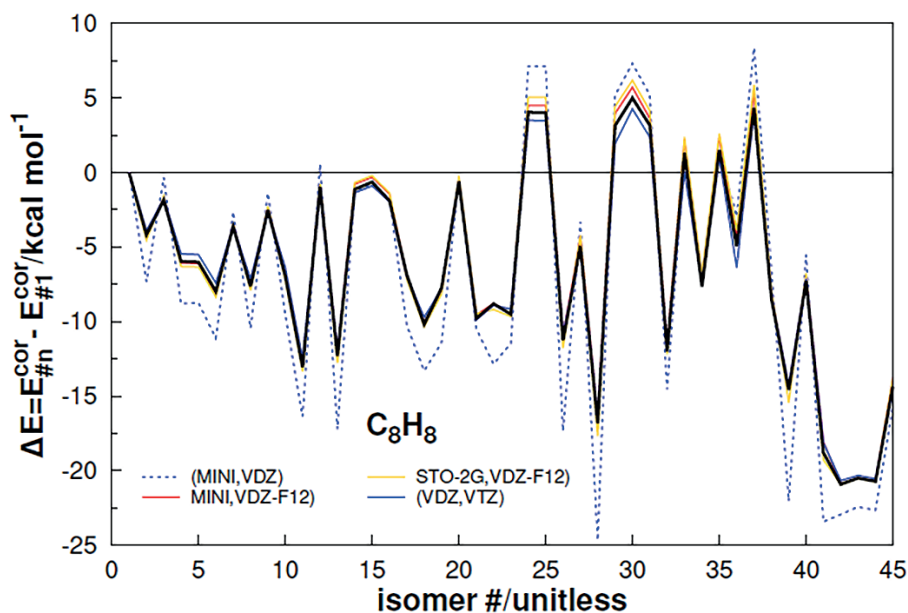

Fig. 4: Energy separation of C_8H_8 isomers. Shown in solid black is the accurate, yet expensive, CCSD(T)/CBS(VDZ-F12,VT Z-F12) data;[45] see reference for names of all 45 C_8H_8 isomers and their geometries. Adapted from the literature.[34]

All Kohn-Sham DFT calculations employed the popular M06-2X functional, while the correlated MO-based ones used single-reference MP2 and CCSD(T) methods with the raw MO energies subsequently CBS (*d, t*) extrapolated. It was found[38] that MP2/CBS(*d, t*) is as cost effective as DFT/M06-2X and often slightly more while showing a satisfactory accuracy when compared with the reference data. A similar performance has been demonstrated[39] in the study of even-numbered carbon clusters (C_n; n=4–10). Moreover, once such MP2/CBS(*d, t*) calculations were combined with variable-scaling opposite spin[40] (VOS) theory, the results approximated couple cluster quality at no additional cost.[39]

As a case study, we consider here an organic molecule with multiple isomers since these are known to pose a challenging problem to DFT.[41–44] Specifically, we consider the full set of 45 isomers of C_8H_8 for which accurate isomerization energies have been reported[45] by means of accurate high-level *ab initio* methods and the W1-F12 thermochemical protocol.[46] Besides covering a whole range of hydrocarbon functional groups [which include (linear and cyclic) polyacetylene, polyyne, and cumulene moieties, as well as aromatic, anti-aromatic, and highly-strained rings] they have been studied by a variety of composite semiempirical procedures, and a panoply of DFT functionals. Figure 4 compares some of our CBS extrapolated correlation energies[34] with the most accurate results thus far reported. Suffice it to note that the trends observed with MP2 are similar to the ones found in CCSD(T) calculations, but obviously at a much lower cost. Specifically, we highlight the following:[34] (a) use of a (sM,M) basis-set pair is enough to mimic the CCSD(T) reference data[46] with high accuracy [rmsd of 3.01, 0.42, and 0.64 kcal mol^{-1}, respectively from (MIN,V*DZ*), (MINI,V*DZ*-F12), and (STO-2G,V*DZ*-F12) pairs], which compares with the value 0.49 kcal mol^{-1} from our recently[38] recommended CBS(V*DZ*,V*TZ*) scheme; (b) the wall-times for the first three pairs [in the same order: (761±103), (17357±1783) and (11803 ± 1294) s] are much smaller than the references,[45] and are essentially the time spent with the extended basis set calculation, which is a fiftyfold factor less expensive than CCSD(T)/CBS(V*DZ*-F12,V*TZ*-F12) for the cheapest combination; (c) the CBS results slightly outperform in one case the CCSD(T)/CBS(V*DZ*,V*TZ*) ones, although they are far less expensive and outperform DFT/M06-2X by 2.8 kcal mol^{-1}, which performs itself similarly to MP2/CBS(MINI,V*DZ*-F12); (c) CBS(sM,M) schemes are general and pseudo-single-level type, thus extending and probably outperforming genuine single level[4,47] ones; (d) CBS(sM,sM) extrapolation schemes show somewhat modest performances, but at a drastically smaller cost while occasionally performing at levels of accuracy comparable to some DFT functionals. Such a performance is partly expected from the excessive proximity of the involved hierarchical numbers, which tend to spoil the accuracy of the CBS extrapolation. Yet, a reliable prediction is observed for the ups and downs observed in the evolution of the isomers' energy from styrene (isomer #1) to ethynyl-bicyclo [3.1.0] hex-2-ene[45] (#45). Such "educated" prediction contrasts with the highly mismatched pattern that is observed at raw *ab initio* level with sM bases, whose correct ordering is judged difficult to get also at DFT level of theory.[45]

III. EXTRAPOLATION IN REACTION DYNAMICS

Despite all progress made in performing quantum reactive scattering calculations, many practical difficulties persist. The most prominent concerns the calculation of the cross-section which requires converged quantum mechanical (QM) reaction probabilities for all contributing total angular momenta partial waves, thence a large number of total angular momenta *J* with Coriolis coupling taken into

account. This is extremely computationally intensive due to the proliferation of angular momentum states as J increases. It implies therefore a huge labor, particularly in obtaining the thermal rate coefficient,

$$k(T) = \frac{1}{\hbar \mathcal{Z}(T)} \int_0^\infty dE P(E) \exp(-E/k_B T) \qquad (6)$$

where E is the total energy, $\hbar$ the Planck constant, k_B the Boltzmann constant, and $Z(T)$ the total partition function (translational, rotational and vibrational). Since there are no $J = 0$ experimental rate constants, it is desirable to be able to obtain accurate rate constants without having to carry out explicit quantum dynamics calculations for many total angular momentum partial waves. A common approach is to employ the so-called sudden approximations[48,49] for dynamics calculations at high J values which significantly reduces the amount of computational effort. An even simpler approach is to use the so called J-shifting method.[50] This enables to evaluate cross-sections and rate constants using only the accurate results for $J = 0$. It consists of assuming that non-zero total angular momentum probabilities can be derived from $J = 0$ ones using the approximation

$$P^J_{v_i,j_i;v_f,j_f} \sim P^{J=0}_{v_i,j_i;v_f,j_f}(E^J_{tr}) \qquad (7)$$

where v_i (v_f) and j_i (j_f) are the reactant (product) vibrational and rotational quantum numbers, and

$$E^J_{tr} = E_{tr} - B^\dagger J(J+1) \qquad (8)$$

In principle, the energy shift should be estimated from the rotational energy of the intermediate complex, and hence $B^\dagger$ should be the rotational constant at the transition state. A more rational J-shifting approach is to compute the reaction probabilities at a number of nonzero total angular momentum values of J. By comparing the reaction probabilities computed explicitly at these nonzero values of J with those obtained by applying the J-shifting method, one may obtain more realistic values for the shifting constant(s) B. However, if the reactions are not dominated by a simple transition state[51], it is not obvious which geometry should be used. Moreover, for nonlinear molecules, there are generally three distinct rotational constants, and it is not easy to decide which (if any) is the most appropriate. The shifting constant so obtained can therefore produce large errors in the calculated rate constant and it is generally difficult to ascertain its accuracy without knowing the exact result. This is the case particularly for reactions governed by deep potential wells, an example being the prototypical $H + O_2$ reaction. For such reactions, time-independent methods require large basis sets representing the open and closed states to solve accurately the dynamical problem while the time-dependent ones need long integration times for propagation of the wave packet until the reaction is complete. An alternative idea that we have long suggested[52] is whether the higher-J reaction probabilities can be extrapolated from $P^{J=0}$ to values as large as they are necessary at high energies (these are predicted to reach $J_c \sim 50$ for the reaction $H + O_2 \rightarrow OH + O$) such as to attain full convergence.

Classical mechanics is deemed ideal as the extrapolation tool, since the QM and CM results should approach each other at high energies by the correspondence principle, which states that quantum mechanics reproduces classical physics in the limit of large quantum numbers.

A. Extrapolation with QCT/QM-αJ method

Consider the quantum mechanical reactive cross section for the transition from the initial (v, j) state to v' (*i.e.*, over all final j' states)

$$\sigma_{vj,v'}(E_{tr}) = \frac{1}{g}\frac{\pi}{k^2}\sum_{J=0}^{\infty}(2J+1)P_{vj,v'}^J(E_{tr}) \qquad (9)$$

where $k^{-2} = (2j+1)k^{-2}_{vj} = 2\mu E_{tr}$, $1/g$ is the electronic degeneracy factor, and μ is the reactants reduced mass. By summing over all final vibrational states, one gets

$$\sigma_{vj}(E_{tr}) = \frac{1}{g}\frac{\pi}{k^2}\sum_{J=0}^{\infty}(2J+1)P_{vj}^J(E_{tr}) \qquad (10)$$

where

$$P_{vj}^J(E_{tr}) = \sum_{\Omega}\sum_{v'j'\Omega'}^{\infty} P_{vj\Omega,v'j'\Omega'}^J(E_{tr}) \quad J = 0, 1, 2, \ldots$$

$$(11)$$

with the summation in this equation being over all energetically open channels; Ω (or Ω') is the projection of J on the body-fixed z axis for the reactant (product), and its range is $-\min(j<, J)$ to $\min(j<, J)$ where $j<$ is the lesser of j and j'.[48,53]

Consider next the following approximation

$$P_{vj}^J(E_{tr}) = P_{vj}^0(E_{tr})\mathcal{F}(J) \qquad J \leq J_c \qquad (12)$$
$$= 0 \qquad J > J_c \qquad (13)$$

where $\mathcal{F}(J)$ is some distribution of the reaction probability with J at a given value of the energy E_{tr}. The simplest approximation is to assume it uniform, QM-UJ, thence unit probability for all values of J below a critical value J_c (which approximately mimics the maximum impact parameter in the trajectory approach). Indeed, a more flexible expression is a distribution such as

$$\mathcal{F}(J) = 1 - \left(\frac{J}{J_c}\right)^{\alpha} \qquad (14)$$

where α is some power to be specified by the user. Possible shapes are illustrated in Figure 5 for various integer powers, including the uniform α = ∞ distribution.

$$\mathcal{F}(J) = 1 - \left(\frac{J}{J_c}\right)^\alpha$$

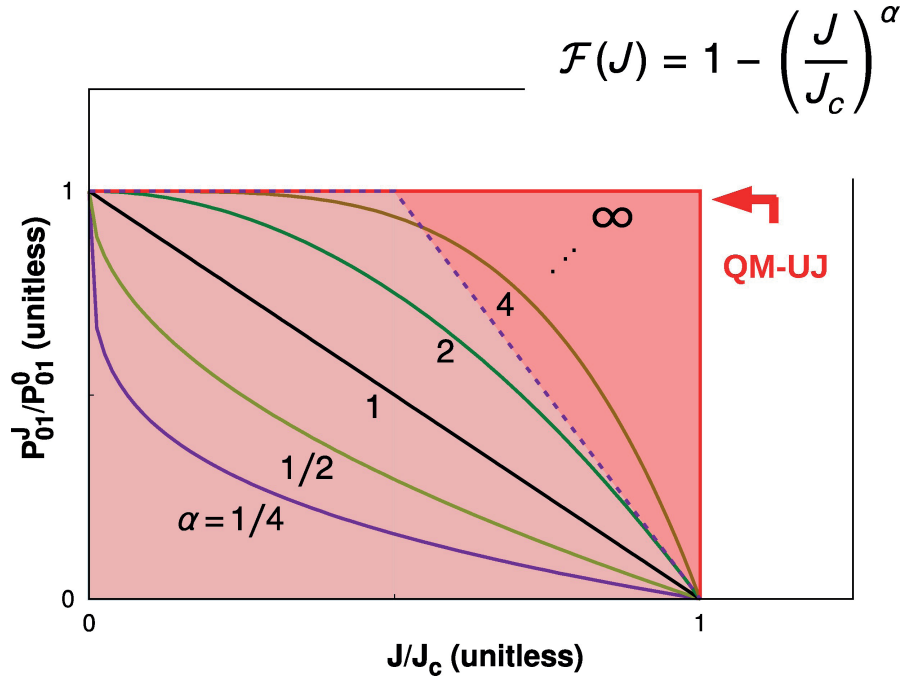**Fig. 5:** The model distribution $\mathcal{F}(J)$ in Eq. (14).

Although some $vj \to v'$ partial reactive probabilities may display a maximum at some intermediate value of J,[54,55] such details more likely average out during the summation procedure, and may therefore be of no great concern. By using the Euler-MacLaurin summation formula,

$$\sum_{\ell=1}^{n-1} = \int_0^n f(\ell)d\ell - \frac{1}{2}\left[f(0) + f(n)\right]$$
$$+ \frac{1}{12}\left[f'(n) - f'(0)\right] - \frac{1}{720}\left[f'''(n) - f'''(0)\right] + \dots$$
$$+ (-1)^{p-1}\frac{B_p}{(2p)!}\left[f^{(2p-1)}(n) - f^{(2p-1)}(0)\right] + \dots$$

$$(15)$$

where B_p is a Bernoulli number, a prime denotes a derivative of f with respect to ℓ, and $f^{(2p-1)}(n)$ its $(2p-1)$--th derivative. The general extrapolation scheme is then denoted QCT/QM-αJ, which for α=2 (Q from quadratic) assumes the form:

$$\sigma_{v,j}(E_{tr}) = \frac{1}{g}\frac{\pi}{k^2}\left(\frac{1}{2}J_c^2 + \frac{2}{3}J_c - \frac{1}{6J_c}\right)P_{vj}^{J=0}(E_{tr}) \quad (16)$$

The results for this case ($\alpha = 2$) and ∞ should perhaps bracket most expectations and give essentially the same results, even if necessarily predicting different J_c values.

The above extrapolation scheme has been originally utilized[52] for the H + O$_2$ reaction using the DMBE IV PES,[56] and the cut-off total angular momentum J_c obtained from the requirement that the $P_{01}^{J=0}$ QCT reactive probabilities when replaced in Eq. (16) should mimic the corresponding calculated total reactive cross-sections. As the results suggest, the simplest assumption of an E_{tr}-independent α appears justified for this reaction. For example, the well known semiclassical result $J \sim (2\mu E_{tr})^{1/2} b/\hbar$, suggests that J_c and the maximum impact parameter used for the trajectory calculations (b_{max}) should have a related dependence on E_{tr}, which was actually verified numerically to be the case.[52] In fact, for such a reaction,[57] and probably others,[54,58,59] a value $\alpha \sim 2$ turns out to be satisfactory, although some $vj \rightarrow v'$ partial reactive probabilities[54,59] may call for the use a form of $F(J)$ showing a peak at some intermediate value of J. Suffice it to emphasize that the value of α is not decisive for predicting the cross-section. Of course, it has influence on the cut-off total angular momentum J_c but this may not be so relevant in the context of the proposed extrapolation technique.

The performance of the J-shifting[50] and extrapolation[60] schemes in calculating integral cross sections and rate constants has been recently revisited[61] for the H + O$_2$ reaction using the DMBE IV and CHIPR PESs.[62] The QD $J = 0$ calculations where performed using a three-dimensional time-dependent wave packet formalism based on hyperspherical coordinates approach developed by Billing and Markovic[63] for $J = 0, 1$, and extended to arbitrary J by Adhikari and Varandas.[64] Figure 6 depicts the results obtained when using J-shifting[50] versus extrapolation with the help of QCT calculations for specific J values.[60] As shown, the integral cross sections calculated on the CHIPR PES by J-shifting are in excellent agreement with the integral cross sections obtained via extrapolation. Moreover, Figure 6 shows that the cross sections calculated by those two approaches on the DMBE IV PES are also in close agreement with each other, except for the fact that the J-shifted cross sections slightly overestimate the extrapolated results for E_{tr} values beyond 1.5 eV. In fact, we are led to believe that the extrapolation scheme is probably more reliable for high collision energies, since the correspondence principle tells that quantum results should approach their classical counterparts. Of course, it should not be surprising that unlike J-shifting, the extrapolated cross section curve exhibits a marked oscillatory pattern, which originates from the rich resonance structure of the $P^{J=0}(v = 0, j = 1)$ profile. Thence, it should be viewed on an averaged sense.

Once obtained the reaction probability for $J=0$, one can easily calculate the integral cross section. The results are displayed in Figure 7 for the $v=0$, $j=1$ rovibrational state of O$_2$ on both the CHIPR and DMBE PESs over the collision energy range 0.0–2.5eV. Other calculations by the author[60] on the DMBE IV PES[65], Quéméner et al. on the DIMKP PES[66] as well as XXZLG PES[67] are also shown together with the experimental findings by Keβler et al.[68], Abu Bajeh et al.[69] and Seeger et al.[70]. Generally, there is good agreement between the calculated profile on the most recent and likely accurate CHIPR PES which shows a broad maximum around ~1.7 eV, and the experimental measurements by Keβler et al.[68] and Seeger et al.[70]. Notably, the present profile calculated on the CHIPR PES overestimates previous theoretical calculations by Quéméner et al., who had performed time-dependent quantum mechanical scattering calculations based on hyperspherical coordinates and the J-shifting approximation to calculate integral cross-sections. On the other hand, the calculated integral cross section on the DMBE IV PES are in reasonably good
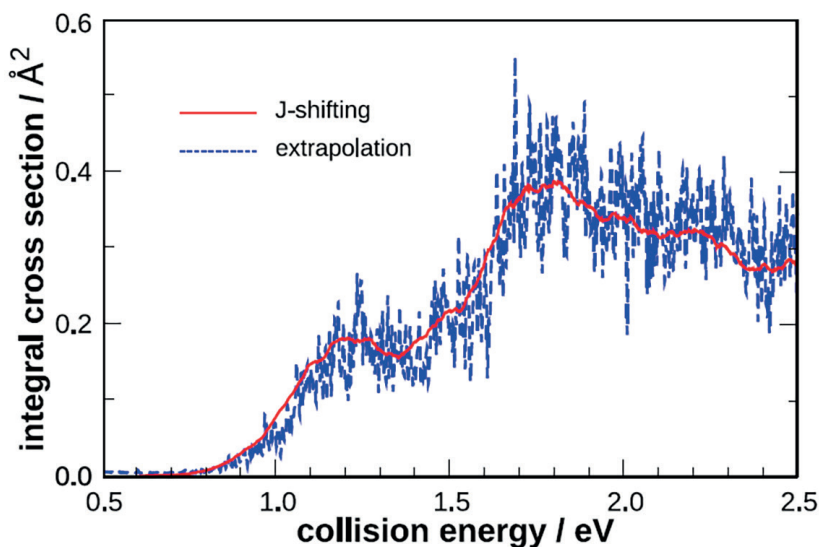
Fig. 6: Total integral cross section for H + O$_2$($v = 0$, $j = 1$) reaction as a function of the collision energy computed on the CHIPR PES by utilizing both the J-shifting approximation and the extrapolation method. Adapted from the literature.[61]

agreement with the most recent results. In addition, the calculations on CHIPR seem to be in good agreement with the experimental measurements of Keβler *et al.* and Seeger *et al.*, specifically in the lower collision energy range, whereas the profile for DMBE IV case matches well with the same experimental profiles[68,70] in the higher energy regime. Interestingly, it also agrees in shape with our earlier results,[60] with any quantitative differences attributable both to the methodology and PESs that have been utilized. The reader is addressed to the original papers for further details as well as for thermal rate coefficient calculations.

Expectedly, a more rigorous approach could be to obtain *v*-specific values of J_c as a function of the translational energy. This could be done by choosing at each value of E_{tr} the value of J_c that once replaced in Eq. (16) could mimic the value of σ_{vj} calculated at the QCT level of theory. From this, a sequence of J_c^v (E_{tr}) would be obtained that could be used together with the QM values of $P_v^j = 0$ to predict the corresponding QM vibrational specific cross sections. Boltzmann-averaging would then yield a prediction of the total QM cross section and subsequently the thermal rate coefficient. Such a procedure might have the further advantage of more effectively washing out some of the rotational structure inherent to the calculated $P_{v1}^j = 0$ quantum reaction probabilities, thus yielding a more realistic representation of the excitation function prior to any data smoothing. Of course, the values of J_c may show some dependence on E_{tr}, although the results obtained thus far show that they vary little with the type of quasiclassical trajectory approach that has been employed in the calculations. Thus, the dependence of J_c on E_{tr} has been neglected, although this is not a model requirement.

IV. CONCLUDING REMARKS

Saturating a basis of functions is key to achieve chemical accuracy in electronic structure calculations but unaffordable for medium and large molecules. CBS extrapolation offers a cost-effective and reliable
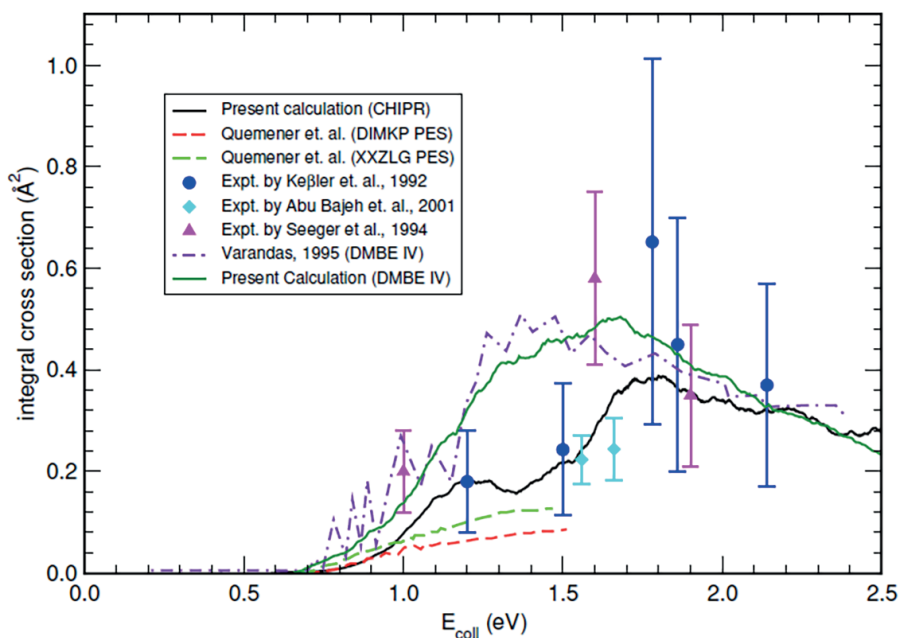
Fig. 7: Integral cross section for $H + O_2(v = 0, j = 1)$ reaction as a function of collision energy in comparison with previous theoretical results and experimental measurements. Adapted from the literature.[61]

way out, and we have recently[4,34] shown how correlated calculations with cost-effective sM and M basis sets can help on the endeavour. This has been made easier with our novel USTE$_\alpha$ scheme,[34] which allows to hierarchize any available basis sets from the recovered correlation energy. Because the joint use of a SM and M or E basis sets costs about as much as the calculation of a single-point with the latter, USTE$_\alpha$ has become as efficient (and likely more reliable) than true single-level schemes despite being dual level.[4] The same reasoning applies to the Hartree-Fock energy component although in a less severe manner, since CBS extrapolation from larger basis sets may be feasible[4,28] and even affordable a direct convergence to satisfactory accuracy of the raw *ab initio* energy. Due to their low-cost and reliability, MO-based MP2/CBS(sM,M) and MP2/CBS(M,M) methods get comparable if not less expensive than KS DFT calculations with popular functionals. This opens a wealth of research topics to revisitation, ranging from CBS extrapolations in large-sized systems to on-the-fly dynamics calculations. They may then, *per se* or jointly with cost-free MP2-VOS[40] theory, have a broad impact in chemistry and even materials science.

Regarding reaction dynamics, QCT calculations have been suggested to offer a convenient way to extrapolate QM reactivities. This is not unexpected given the good agreement found (away from the threshold regions where tunneling dominates) between the QCT and QM results even in the most unfavorable case of the $H + H_2$ reaction. In fact, such an expectation bears a solid ground on the correspondence principle, with applications of the method to polyatomic reactions[6] being highly desirable.

ACKNOWLEDGMENTS

(Comunicação apresentada à Classe de Ciências
na sessão de 15 de março de 2018)

REFERENCES

1   M. Born and J. R. Oppenheimer, *Ann. Phys.*, 1927, 84, 457-484.

2   A. J. C. Varandas, J. P. da Providência, M. Brajczewska, and J. da Providência, *Chem. Phys. Lett.*, 2014, 610, 167-172.

3   A. J. C. Varandas, *Chem. Phys. Lett.*, 2007, 439, 386-392.

4   A. J. C. Varandas, *Ann. Rev. Phys. Chem.*, 2018, 69, 177-203.

5   A. J. C. Varandas, *J. Chem. Phys.*, 2000, 113, 8880-8887.

6   B. Fu, X. Shan, D. H. Zhang, and D. C. Clary, *Chem. Soc. Rev.*, 2017, 46(24), 7625-7649.

7   N. Bohr, *Zeitschrift fu¨r Physik*, 1920, 423, 423-469.

8   T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, Chichester, 2000.

9   C. H¨attig, W. Klopper, A. Köhn, and D. P. Tew, *Chem. Rev.*, 2012, 112, 4-74.

10 T. Kato, *Commun. Pure Appl. Math.*, 1957, 10, 151-177.

11 W. Kutzelnigg and W. Klopper, *J. Chem. Phys.*, 1991, 94, 1985-2001.

12 W. Kutzelnigg and J. D. Morgan III, *J. Chem. Phys.*, 1992, 96, 4484-4508.

13 J. R. Flores, R. Slupski, and K. Jankowski, *J. Chem. Phys.*, 2006, 124, 104107.

14 C. Schwartz, *Phys. Rev.*, 1962, 126, 1015-1019.

15 D. W. Schwenke, *Mol. Phys.*, 2012, 110, 2557-2567.

16 T. H. Dunning, Jr., *J. Chem. Phys.*, 1989, 90, 1007-1023.

17 T. H. Dunning Jr., K. A. Peterson, and D. E. Woon in *Encyclopedia of Computational Chemistry*, ed. P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kolman, and H. F. Schaefer III; Wiley, Chichester, 1998; p. 88.

18 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, 7(18), 3297-3305.

19 A. C. Neto, E. P. Muniz, R. Centoducatte, and F. E. Jorge, *J. Mol. Struct*.:{*THEOCHEM*}, 2005, 718, 219-224.

20 S. Zhong, E. C. Barnes, and G. A. Petersson, *J. Chem. Phys.*, 2008, 129, 184116.

21 D. S. Ranasinghe and G. A. Petersson, *J. Chem. Phys.*, 2013, 138, 144104.

22 F. Jensen, *Advanced Review*, WIREs Comput. Mol. Sci., p. 1.

23 T. H. Dunning Jr., *J. Phys. Chem. A*, 2000, 104, 9062-9080.

24 T. Helgaker, W. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.*, 1997, 106(23), 9639-9646.

25 A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, *Chem. Phys. Lett.*, 1998, 286, 243-252.

26 K. Raghavachari and J. B. Anderson, *J. Phys. Chem.*, 1996, 100, 12960-12973.

27 W. Kutzelnigg, *Int. J. Quantum Chem.*, 1994, 51, 447-463.

28 F. N. N. Pansini, A. C. Neto, and A. J. C. Varandas, *Theor. Chem. Acc.*, 2016, 135, 261.

29 D. Feller, K. A. Peterson, and T. D. Crawford, *J. Chem.Phys.*, 2006, 124, 054107.

30 V. Vasilyev, *Comput. Theor. Chem.*, 2017, 1115, 1-3.

31 A. J. C. Varandas, *J. Chem. Phys.*, 2007, 126, 244105.

32 A. J. C. Varandas, *J. Phys. Chem. A*, 2008, 112, 1841-1850.

33 A. J. C. Varandas, *J. Phys. Chem. A*, 2010, 114, 8505-8516.

34 A. J. C. Varandas, *Phys. Chem. Chem. Phys.*, 2018, 20, 22084-22098.

35 A. J. C. Varandas and F. N. N. Pansini, *J. Chem. Phys.*, 2014, 141, 224113.

36 F. N. N. Pansini, A. C. Neto, and A. J. C. Varandas, *Chem. Phys. Lett.*, 2015, 641, 90-96.

37 J. G. Hill, K. A. Peterson, G. Knizia, and H.-J. Werner, *J. Chem. Phys.*, 2009, 131, 194105.

38 A. J. C. Varandas, M. M. González, L. A. Montero-Cabrera, and J. M. G. de la Vega., *Chem. Eur. J.*, 2017, 23, 9122-9129.

39   A. J. C. Varandas, *Eur. Phys. J. D*, 2018, 72(10), 180.

40   A. J. C. Varandas, *J. Chem. Phys.*, 2010, 133, 064104.

41   S. Grimme, *Angew. Chem. Int. Ed.*, 2006, 45, 4460-4464.

42   P. R. Schreiner, A. A. Fokin, R. A. Pascal Jr, and A. de Meijere, *Org. Lett.*, 2006, 8, 3635-3638.

43   P. R. Schreiner, *Angew. Chem. Int. Ed.*, 2007, 46, 4217-4219.

44   M. S. S. Grimme and M. Korth, *J. Org. Chem.*, 2007, 72, 2118.

45   A. Karton and J. M. L. Martin, *Mol. Phys,*, 2013, 110, 2477-2491.

46   A. Karton and J. M. L. Martin, *J. Chem. Phys.*, 2012, 136, 124114.

47   F. N. N. Pansini and A. J. C. Varandas, *Chem. Phys. Lett.*, 2015, 631-632, 70-77.

48   R. T. Pack, *J. Chem. Phys.*, 1974, 60(2), 633-639.

49   P. McGuire and D. J. Kouri, *J. Chem. Phys.*, 1974, 60, 2488-2499.

50   J. M. Bowman, *J. Phys. Chem.*, 1991, 95, 4960-4968.

51   C. Leforestier and W. H. Miller, *J. Chem. Phys.*, 1994, 100, 733-735.

52   A. J. C. Varandas, *Mol. Phys.*, 1995, 85, 1159-1164.

53   G. C. Schatz and A. Kuppermann, *J. Chem. Phys.*, 1976, 65, 4642-4667.

54   F. J. Aoiz, L. Banãres, V. J. Herrero, and V. S. Rábanos, *Chem. Phys. Lett.*, 1994, 218, 422-432.

55   J. M. Launay and M. L. Dourneuf, *Chem. Phys. Lett.*, 1990, 169, 473-481.

56   M. R. Pastrana, L. A. M. Quintales, J. Brandão, and A. J. C. Varandas, *J. Phys. Chem.*, 1990, 94, 8073-8080.

57   A. J. C. Varandas, *J. Chem. Phys.*, 1993, 99, 1076-1085.

58   F. Aoiz, L. Banãres, M. D'Mello, V. Herrero, V. S. Rábanos, L. Schnieder, and R. Wyatt, *J. Chem. Phys.*, 1994, 101, 5781-5791.

59   H. Szichman, A. J. C. Varandas, and M. Baer, *Chem. Phys. Lett.*, 1994, 231, 253-256.

60   A. J. C. Varandas, *Molec. Phys.*, 1995, 85(6), 1159-1164.

61   S. Ghosh, R. Sharma, S. Adhikari, and A. J. C. Varandas, *Phys. Chem. Chem. Phys.*, 2018, 20, 478-488.

62   A. J. C. Varandas, *J. Chem. Phys.*, 2013, 138, 054120.

63   G. D. Billing and N. Markovic, *J. Chem. Phys.*, 1994, 100(2), 1085-1093.

64   S. Adhikari and A. J. C. Varandas, *Comput. Phys. Comm.*, 2013, 184(2), 270-283.

65   M. R. Pastrana, L. A. M. Quintales, J. Brandao, and A. J. C. Varandas, *J. Phys. Chem.*, 1990, 94(21), 8073-8080.

66   B. K. Kendrick and R. Pack, *J. Chem. Phys.*, 1995, 102, 1994-2012.

67   G. Quemener, B. K. Kendrick, and N. Balakrishnan, *J. Chem. Phys.*, 2010, 132(1), 014302/1-10.

68   K. Keβler and K. Kleinermanns, *J. Chem. Phys.*, 1992, 97(1), 374-377.

69   M. A. Bajeh, E. M. Goldfield, A. Hanf, C. Kappel, A. J. H. M. Meijer, H. R. Volpp, and J. Wolfrum, *J. Phys. Chem. A*, 2001, 105(13), 3359-3364.

70   S. Seeger, V. Sick, H. R. Volpp, and J. Wolfrum, *Isr. J. Chem.*, 1994, 34(1), 5-18.