



**HAL**  
open science

# The electronic excited states of ethylene with large-amplitude deformations: A dynamical symmetry group investigation

Benjamin Lasorne, Michael A. Robb, H.-D. Meyer, Fabien Gatti

► **To cite this version:**

Benjamin Lasorne, Michael A. Robb, H.-D. Meyer, Fabien Gatti. The electronic excited states of ethylene with large-amplitude deformations: A dynamical symmetry group investigation. *Chemical Physics*, 2010, 377, pp.30. 10.1016/j.chemphys.2010.08.011 . hal-00538796

**HAL Id: hal-00538796**

**<https://hal.science/hal-00538796>**

Submitted on 14 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# The electronic excited states of ethylene with large-amplitude deformations: A dynamical symmetry group investigation

B. Lasorne <sup>a,\*</sup>, Michael A. Robb <sup>b</sup>, Hans-Dieter Meyer <sup>c</sup>, F. Gatti <sup>a</sup>

<sup>a</sup> Institut Charles Gerhardt Montpellier, UMR 5253 (CNRS-UM2-ENSCM-UM1), CTMM, Université Montpellier 2, CC 1501, Place Eugène Bataillon, 34095 Montpellier, France

<sup>b</sup> Department of Chemistry, Imperial College, SW7 2AZ London, United Kingdom

<sup>c</sup> Ruprecht-Karls Universität, Physikalisch Chemisches Institut, Theoretische Chemie, Im Neuenheimer Feld 229, 69120 Heidelberg, Germany

We show that the lowest manifold of electronic states of ethylene (ethene, C<sub>2</sub>H<sub>4</sub>) can be described correctly with a complete active space of 17 quasidiabatic configurations built on state-averaged orbitals. This space is stable upon large-amplitude deformations, such as torsion, pyramidalization, CC stretching and HCH bending. The properties of the nuclear coordinates and valence and Rydberg electronic states are investigated within the framework of nuclear-permutation-inversion group theory. This systematic analysis is compared to a previous model of the valence states of ethylene (R.P. Krawczyk, A. Viel, U. Manthe, W. Domcke, J. Chem. Phys. 119 (2003) 1397). Our approach is intended to be generalized to the non-adiabatic photochemistry of organic molecules where large-amplitude deformations require global vibronic Hamiltonian models to be expressed in terms of simple functions of polyspherical valence coordinates.

## 1. Introduction

Excited-state dynamics concerns molecular processes induced upon UV-visible light absorption. Simulations of photochemical reactions and calculations of electronic spectra involve nuclear dynamics methods to determine the time evolution of the molecular geometry used in concert with electronic structure methods capable of computing electronic excited state potential energy surfaces. Radiationless decay processes are governed by non-adiabatic transitions that are intense around conical intersections where potential energy surfaces are degenerate [1]. The Born-Oppenheimer approximation cannot be used, and the molecular Hamiltonian must account for non-adiabatic coupling terms. Quasidiabatic representations [2-4] allow the singular kinetic non-adiabatic couplings to be replaced by potential non-adiabatic couplings that are smooth-varying functions of the nuclear coordinates. In addition, the diabatic potential energy surfaces (diagonal elements) also are simple functions that can cross with no restriction, as opposed to the two-dimensional cusp at the conical intersection be-

tween two adiabatic surfaces. Quasidiabatic functions are thus easier to fit to regular mathematical expressions.

Adiabatic-to-diabatic transformations, also called diabatizations, are based on a smoothness condition of either the electronic wavefunctions or the expectation values of some physical properties. They often relate to the dominant electronic configurations underlying the adiabatic electronic wavefunctions in the spirit of correlation diagrams or valence-bond theory [5]. A large number of approaches have been developed over the years (see, e.g., Refs. [3,4,6] and references therein).

The approach we are interested in is called diabatization by *ansatz*, where the eigenvalues of a quasidiabatic Hamiltonian model are fitted to the adiabatic energies calculated with an electronic structure method over a range of geometries. The diabatization criterion is not explicit in terms of the electronic states, but is somewhat related to the compactness of the mathematical expressions assumed for the quasidiabatic functions. Simple expressions mean that the electronic states vary as smoothly as possible with the nuclear coordinates. In particular, the vibronic coupling Hamiltonian model developed by Köppel and co-workers [7] is based on a low-order local expansion of the quasidiabatic potential energy surfaces and coupling functions around the Franck-Condon point in terms of normal coordinates. In concert with the

\* Corresponding author.

E-mail address: benjamin.lasorne@univ-montp2.fr (B. Lasorne).

multiconfiguration time-dependent Hartree (MCTDH) quantum dynamics method [8], the vibronic coupling Hamiltonian approach has been used with great success to the calculation of many photo-absorption and photoelectron spectra [9].

Calculating electronic spectra mostly relies on short-term dynamics, especially when the conical intersection is close to the Franck–Condon point. However, describing photochemical events implies to deal with longer processes, large-amplitude nuclear motions, and complicated reaction pathways connecting several potential energy wells through transition barriers and crossings between different electronic states. There is thus a need for generalizing the vibronic coupling Hamiltonian model by making it a global description capable of dealing with large-amplitude deformations of the molecular geometry. This implies using internal coordinates that are curvilinear rather than rectilinear to get simple expressions for the potential energy functions. However, curvilinear coordinates lead to complicated expressions for kinetic energy operators. Polyspherical coordinates, as described in Ref. [10], are a good compromise, well-adapted to run MCTDH quantum dynamics calculations.

As first shown by Keating and Mead [11] and recently pointed out in Ref. [12], building a truly global representation of the electronic Hamiltonian for a flexible molecule [13] should be achieved within the framework of dynamical symmetry, also called molecular symmetry [14], when there are identical nuclei in the molecular system. This implies using the complete nuclear-permutation-inversion (CNPI) group, which is the symmetry group relevant for the vibronic Hamiltonian of the flexible molecule (as opposed to the point group of any given geometry), or a subgroup of it if some deformations are considered unfeasible. In this, the operations correspond to identical nuclei permutations and inversion of the whole molecule. See also Refs. [12,15,16] for related approaches based on CNPI projectors.

This treatment was used by Peyerimhoff and co-workers [17] to build a one-dimensional quasidiabatic representation for ethylene (ethene,  $C_2H_4$ ) undergoing a pure torsional deformation. The aim of the present work is to extend this approach to more electronic states and more nuclear coordinates to write the elements of the effective Hamiltonian matrix in a quasidiabatic representation adapted to dynamical symmetry. Thinking further ahead, our actual purpose is to generalize the current vibronic Hamiltonian model of Köppel and co-workers by combining it to a dynamical symmetry treatment using simple basis functions of polyspherical coordinates (rather than polynomials of normal coordinates with point-group symmetry adapted to the geometry at the origin of the expansion).

The  $\pi$ -to- $\pi^*$  photo-excitation of CC bonds and subsequent *cis*-*trans* photo-isomerization is one of the most fundamental processes in organic photochemistry. In particular, electronically excited ethylene experiences large-amplitude deformations that eventually lead to ground-state regeneration through a conical intersection. Despite the apparent simplicity of this system, the assignment of the photoabsorption spectrum and the radiationless decay mechanism of this prototype molecule still remain controversial. In this work, we focus on the first step that follows the photoabsorption. H migrations are not considered at this time scale, which allows a reduced-dimensionality treatment whereby six internal coordinates are frozen. We refer to Ref. [18] for a short review of experimental and theoretical investigations. See also Refs. [19–23].

In the present work, we present the first step in building such a global model. We show that the lowest manifold of adiabatic electronic states of ethylene can be described correctly with multiconfiguration self-consistent field (MCSCF) calculations restricted to 17 selected quasidiabatic configurations built on state-averaged orbitals. This space is stable upon large-amplitude deformations,

such as torsion, pyramidalization, CC stretching and HCH bending. The properties of the nuclear coordinates and valence and Rydberg electronic states are investigated within the framework of dynamical symmetry. This systematic analysis is compared to a previous model of the valence states of ethylene [18]. The  $\pi$ -to- $\pi^*$  state of ethylene is still today a computational challenge, as discussed in Ref. [23] and references therein. Although very cheap, our 17-selected-configuration approach is capable of describing the  $V$  state of ethylene quite remarkably.

## 2. Nuclear coordinates

We start with the definition of the nuclear coordinates that will be used to express the Hamiltonian matrix dependence on the molecular geometry. They are taken as symmetrized valence coordinates, directly related to the polyspherical coordinates used in Refs. [18,24,25]. The former are adapted to the expression of the potential energy functions within the framework of dynamical symmetry. The latter lead to a tractable expression for the kinetic energy operator and are suitable for further MCTDH calculations [10]. Our choice of coordinates is adapted to describing deformations that preserve the  $\sigma$  bonds, in particular the torsion around the CC bond involved in the *cis*-*trans* photo-isomerization of substituted ethylene. The dimensionality is reduced to a set of six active internal degrees of freedom.

### 2.1. Vector parametrization of the nuclei

The first step in building polyspherical coordinates is the definition of the vectors to be parametrized [10]. The relative motion of the six nuclei of  $C_2H_4$  is parametrized by five independent vectors and three angular constraints. Two sets are considered here (see Fig. 1): a mixture of valence and Jacobi vectors for defining the polyspherical coordinates used in previous work [18,24,25] to which we want to compare; and valence vectors only, which will serve to define the symmetrized valence coordinates to be used here to investigate the potential energy surfaces.

The first set of valence–Jacobi vectors,  $\vec{R}_i (i = 1$  to 5), is defined as

$$\vec{R}_1 = C_5 \vec{G}_R, \quad (1a)$$

$$\vec{R}_2 = G_L \vec{C}_6, \quad (1b)$$

$$\vec{R}_3 = H_2 \vec{H}_1, \quad (1c)$$

$$\vec{R}_4 = H_4 \vec{H}_3, \quad (1d)$$

$$\vec{R}_5 = C_6 \vec{C}_5, \quad (1e)$$

where  $G_R$  is the center of mass of  $H_1H_2$  and  $G_L$  that of  $H_3H_4$  [18,24,25].

The second set of valence vectors,  $\vec{R}_i (i = 5$  to 9), corresponds to:

$$\vec{R}_6 = C_5 \vec{H}_1, \quad (2a)$$

$$\vec{R}_7 = C_5 \vec{H}_2, \quad (2b)$$

$$\vec{R}_8 = C_6 \vec{H}_3, \quad (2c)$$

$$\vec{R}_9 = C_6 \vec{H}_4. \quad (2d)$$

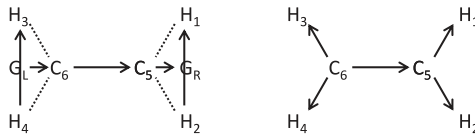


Fig. 1. The two sets of five vectors (left: Jacobi-valence; right: valence).

The two sets are related through:

$$\vec{R}_1 = \frac{\vec{R}_6 + \vec{R}_7}{2}, \quad (3a)$$

$$\vec{R}_2 = -\frac{\vec{R}_8 + \vec{R}_9}{2}, \quad (3b)$$

$$\vec{R}_3 = \vec{R}_6 - \vec{R}_7, \quad (3c)$$

$$\vec{R}_4 = \vec{R}_8 - \vec{R}_9, \quad (3d)$$

which exhibits that Jacobi vectors simply are symmetry-adapted combinations of the valence vectors of the CH<sub>2</sub> fragments.

These two types of vector parametrization are well-suited for describing internal deformations of the molecular geometry that preserve the  $\sigma$  bonds. H migrations and formation of ethylidene structures will not be considered here, although they have been recognized by some authors as important in the interpretation of ethylene photochemistry [21,26–30,20]. For these, satellite coordinates (where H atoms are represented by vectors originated from the center of mass of CC) would be better-suited.

## 2.2. Polyspherical coordinates

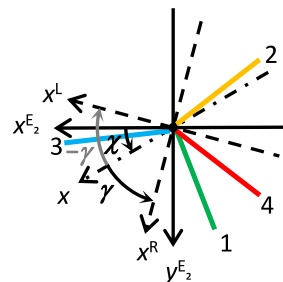
In a second step, each of the vectors defined above are specified in terms of their spherical coordinates in the E<sub>2</sub> frame [10], ( $R_i, \theta_i, \phi_i^{E_2}$ ). All radii, polar angles, and azimuthal angles are restricted according to  $0 \leq R_i \leq \infty$ ,  $0 \leq \theta_i \leq \pi$ , and  $0 \leq \phi_i < 2\pi$  (mod  $2\pi$ ), respectively. The orientation of the E<sub>2</sub> frame with respect to the axes of the laboratory frame is specified by two Euler angles, ( $\phi, \theta$ ), where  $0 \leq \phi < 2\pi$  (mod  $2\pi$ ) and  $0 \leq \theta \leq \pi$ . Their definition corresponds to a constraint involving  $\vec{R}_5 = C_6 C_5 : \theta_5 = 0$  and  $\phi_5^{E_2} = 0$ . The quantization axis,  $z$ , is thus parallel to the CC bond, oriented from C<sub>6</sub> to C<sub>5</sub>.

Radii and polar angles are unambiguously related to internal deformations on one hand, whereas  $\{\phi, \theta\}$  describe the body rotation of the CC fragment in the laboratory frame on the other hand. Things are more complicated for the four dihedral angles. Only three of these are internal degrees of freedom, the remaining one being the third body rotation.

The orientation of any body-fixed (BF) frame relative to the E<sub>2</sub> frame having in common the same  $z$ -axis is determined by the third Euler angle,  $\chi^{BF}$ , where  $0 \leq \chi^{BF} < 2\pi$  (mod  $2\pi$ ). For polyspherical coordinates,  $\chi^{BF}$  must be identified to one of the four unconstrained  $\phi_i^{E_2}$  angles. Once this choice is made, the relative geometry of the molecule is defined by the remaining twelve independent polyspherical coordinates. The radii and polar angles are the same in both frames, hence the absence of superscript for these variables. In contrast, azimuthal angles are related through  $\phi_i^{BF} = \phi_i^{E_2} - \chi^{BF}$  ( $i \neq 5$ ). The relative orbiting motion of the four H nuclei around the CC fragment is thus parametrized by three internal rotations and one body rotation.

In previous work [25], Jacobi-valence vectors were used and the corresponding polyspherical coordinates were defined in the single-rotor-fixed BF frame labeled L, ( $x^L, y^L, z$ ), attached to the left-hand-side CH<sub>2</sub> fragment (see Fig. 2). This particular choice corresponds to a third constraint involving  $\vec{R}_4 = H_4 H_3 : \phi_4^{E_2} = \chi^L$ , such that the  $x^L$ -axis is set according to  $x_4^L \geq 0$  and  $y_4^L = 0$ . Internal deformations and body rotations are thus parametrized by twelve polyspherical coordinates and three Euler angles as required,  $\{R_1, R_2, R_3, R_4, R_5, \theta_1, \theta_2, \theta_3, \theta_4, \phi_1^L, \phi_2^L, \phi_3^L\} \oplus \{\phi, \theta, \chi^L\}$ .

This system had originally been introduced in relation with the internal coordinate system used in Refs. [18,24], which consisted of only six selected degrees of freedom (active coordinates),  $\{r, \alpha_R, \alpha_L, \vartheta_R, \vartheta_L, \varphi\}$ . The latter had been used to parametrize a three-electronic-state, six-dimensional, vibronic Hamiltonian model, while the former had been derived to re-write an exact kinetic energy



**Fig. 2.** The three sets of BF axes used in this work, ( $x^L, y^L$ ) (related to the coordinates used in Refs. [18,24,25]), ( $x^R, y^R$ ), and ( $x, y$ ) (bisecting frame of the L and R frames, related to the coordinates used in the present work), and their relationships with the E<sub>2</sub> axes, ( $x^{E_2}, y^{E_2}$ ). All three frames are attached to  $\vec{R}_5$ , and the common  $z$ -axis is orthogonal to the figure plane. BF L is further attached to the left rotor,  $\vec{R}_4$  (third Euler angle:  $\chi^L = \chi - \gamma$ ); BF R, to the right rotor,  $\vec{R}_3$  (third Euler angle:  $\chi^R = \chi + \gamma$ ), such that  $x^R \parallel H_2 H_1$  and  $x^L \parallel H_4 H_3$  within the figure plane (no  $y$ -component considered). The third frame is the bisecting frame of the former two (third Euler angle:  $\chi = (\chi^R + \chi^L) / 2$ ; torsional angle:  $\gamma = (\chi^R - \chi^L) / 2$ ).

operator for the nuclei. In this,  $r$  describes CC stretching;  $\alpha$ , HCH bending (left or right);  $\vartheta$ , C(HCH) pyramidalization (left or right), defined as the planar angle between CC and CG<sub>R</sub> or CG<sub>L</sub>;  $\varphi$ , (HH)CC(HH) torsion, defined as the dihedral angle between (CG<sub>R</sub>, CC) and (CC, CG<sub>L</sub>).

The coordinate transformation is given through six independent relationships (see Ref. [25]):

$$R_1 = R_{CH} \cos \alpha_R / 2, \quad (4a)$$

$$R_2 = R_{CH} \cos \alpha_L / 2, \quad (4b)$$

$$R_5 = r, \quad (4c)$$

$$\vartheta_1 = \vartheta_R, \quad (4d)$$

$$\vartheta_2 = \vartheta_L, \quad (4e)$$

$$\phi_3^L = \varphi, \quad (4f)$$

where the  $\alpha$ -angles vary from 0 to  $\pi$ , and  $R_{CH}$  is frozen to a constant value of the CH bond length (stretching). The distinction between  $\vartheta_i$  and  $\theta_i$  ( $i = 1, 2$ ) is discussed below.

These six variables are completed with six additional constraints:

$$R_3 = 2R_{CH} \sin \alpha_R / 2, \quad (5a)$$

$$R_4 = 2R_{CH} \sin \alpha_L / 2, \quad (5b)$$

$$\theta_3 = \pi / 2, \quad (5c)$$

$$\theta_4 = \pi / 2, \quad (5d)$$

$$\phi_1^L = 3\pi / 2 + \varphi, \quad (5e)$$

$$\phi_2^L = \pi / 2. \quad (5f)$$

Because of the constraints on  $\phi_1^L$  and  $\phi_2^L$ , the range of variation of the corresponding polar angles are set to the whole circle [25],  $-\pi < \vartheta_i \leq \pi$  ( $i = 1, 2$ ), in order to describe full oscillations in the pyramidalizations of the CH<sub>2</sub> groups. Alternatively, one could use two half circles by allowing for two discrete values of the corresponding azimuthal angles,  $\vartheta_i \geq 0 \equiv \{\theta_i \geq 0, \phi_i^L\}$  and  $\vartheta_i \leq 0 \equiv \{\theta_i \geq 0, \phi_i^L + \pi\}$  ( $i = 1, 2$ ).

## 2.3. Symmetrized valence coordinates

Let us consider now the spherical coordinates of the four CH valence vectors in the E<sub>2</sub> frame. These are more easily related to the geometry parametrizations used in quantum chemistry calculations, and they will be symmetrized in order to make a direct connection between the symmetry of the electronic states and the geometry.

The six-dimensional constraint defined above translates here as four constraints on the four  $R_i$  and two constraints on the four  $\theta_i$  ( $i = 6$  to  $9$ ),

$$R_6 = R_7 = R_8 = R_9 = R_{\text{CH}}, \quad (6a)$$

$$\theta_6 = \theta_7 = \beta_R, \quad (6b)$$

$$\theta_8 = \theta_9 = \pi - \beta_L, \quad (6c)$$

where  $0 \leq \beta_i \leq \pi$  ( $i = R, L$ ) are two of the six active degrees of freedom and correspond to HCC bending. The four azimuthal angles,  $\phi_i^{E_2}$  ( $i = 6$  to  $9$ ) correspond thus to three internal degrees of freedom (two pyramidalizations and torsion) plus the body rotation around the CC bond. The remaining degree of freedom is  $R_5 = r$ .

For the expression of the quasidiabatic functions (potential energies and couplings), it is convenient to symmetrize the valence coordinates with respect to the two  $\text{CH}_2$  fragments. First, for the bending (scissoring) angles:

$$\beta_S = \frac{\beta_R + \beta_L}{2}, \quad (7a)$$

$$\beta_A = \frac{\beta_R - \beta_L}{2}, \quad (7b)$$

Second, for the pyramidalization angles:

$$\tau_U = \frac{\tau_R + \tau_L}{2}, \quad (8a)$$

$$\tau_G = \frac{\tau_R - \tau_L}{2}, \quad (8b)$$

where

$$\tau_R = \frac{\phi_6^{E_2} - \phi_7^{E_2}}{2} + \frac{\pi}{2}, \quad (9a)$$

$$\tau_L = \frac{\phi_8^{E_2} - \phi_9^{E_2}}{2} + \frac{\pi}{2}. \quad (9b)$$

These are illustrated in Fig. 3.

The corresponding valence spherical coordinates read in the  $E_2$  frame as

$$R_5 = r, \quad (10a)$$

$$\theta_6 = \theta_7 = \beta_S + \beta_A, \quad (10b)$$

$$\theta_8 = \theta_9 = \pi - \beta_S + \beta_A, \quad (10c)$$

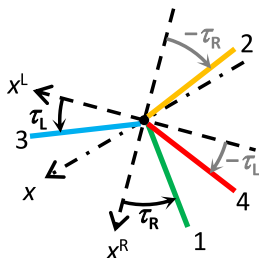
$$\phi_6^{E_2} = \tau_U + \tau_G + \gamma + \chi, \quad (10d)$$

$$\phi_7^{E_2} = \pi - \tau_U - \tau_G + \gamma + \chi, \quad (10e)$$

$$\phi_8^{E_2} = \tau_U - \tau_G - \gamma + \chi, \quad (10f)$$

$$\phi_9^{E_2} = \pi - \tau_U + \tau_G - \gamma + \chi. \quad (10g)$$

The  $\gamma$  coordinate corresponds to the torsion around CC and is identical to  $\varphi/2$  ( $\varphi$  has been introduced above as the dihedral angle (HH)CC(HH)); we prefer to use  $\gamma$  for symmetrized valence coordinates because it leads to a more symmetrical relationship in Eq. (10g)),  $\tau_U$  and  $\tau_G$  to the *ungerade* and *gerade* pyramidalizations,



**Fig. 3.** The left and right pyramidalization angles:  $\tau_R$  or  $\pi - \tau_R$  (right pyramidalization: from  $x^R$  to the projections of the  $\text{C}_5\text{H}_1$  or  $\text{C}_5\text{H}_2$  bonds) and  $\tau_L$  or  $\pi - \tau_L$  (left pyramidalization: from  $x^L$  to the projections of the  $\text{C}_6\text{H}_3$  or  $\text{C}_6\text{H}_4$  bonds).

respectively (non-pyramidalized geometries correspond to  $\tau_R = \tau_L = \tau_U = \tau_G = 0$ ). These motions are illustrated in Fig. 4.

The reverse transformation for the azimuthal angles reads:

$$\chi = \frac{\phi_6^{E_2} + \phi_7^{E_2} + \phi_8^{E_2} + \phi_9^{E_2}}{4} - \frac{\pi}{2}, \quad (11a)$$

$$\gamma = \frac{\phi_6^{E_2} + \phi_7^{E_2} - \phi_8^{E_2} - \phi_9^{E_2}}{4}, \quad (11b)$$

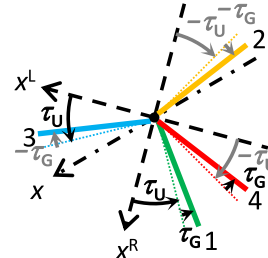
$$\tau_U = \frac{\phi_6^{E_2} - \phi_7^{E_2} + \phi_8^{E_2} - \phi_9^{E_2}}{4} + \frac{\pi}{2}, \quad (11c)$$

$$\tau_G = \frac{\phi_6^{E_2} - \phi_7^{E_2} - \phi_8^{E_2} + \phi_9^{E_2}}{4}. \quad (11d)$$

The respective domains of definition of the non-symmetrized coordinates are not ambiguous ( $0 \leq \phi_i^{E_2} < 2\pi \pmod{2\pi}$  ( $i = 6$  to  $9$ )), whereas using symmetrized coordinates may lead to multi-valued arrangements in the  $E_2$  frame (see discussion in Ref. [14]). Here, a single-valued, but not periodic, domain is obtained by setting  $-\pi/2 \leq \gamma$ ,  $\tau_U$ ,  $\tau_G < \pi/2$  (and  $0 \leq \chi < 2\pi$ ). Symmetrized valence coordinates will be used to sample geometries on which potential energies will be calculated for the electronic states of interest. The  $\chi$  angle will be set to zero, as the relative geometry is not affected by this parameter.

The  $\chi$  angle introduced here corresponds to the third Euler angle of the bisecting BF frame,  $(x, y, z)$  (see Fig. 2), halfway between both single-rotor-fixed frames:  $(x^R, y^R, z)$  and  $(x^L, y^L, z)$ , rotated about  $z$  through  $\chi^R = \chi + \gamma$  and  $\chi^L = \chi - \gamma$  with respect to  $E_2$ , respectively (see Refs. [14,31,32] for more details). The corresponding valence coordinates in the bisecting BF frame have identical expressions as in Eq. (10g) without the  $\chi$  term. An important remark must be made: the spherical coordinates of the valence vectors expressed in the bisecting BF frame are related to, but do not fulfill the definition of polyspherical coordinates (polyspherical coordinates as defined in Ref. [10] need the BF frame to be attached to two of the five vectors:  $\vec{R}_5 = \text{C}_6\text{C}_5$  and, e.g.,  $\vec{R}_9 = \text{C}_6\text{H}_4$ ).

Before going further, it is worth making a few comments about the naming convention on the  $(x, y)$ -axes. Let us consider a  $D_{2h}$  geometry. The names of the three principal axes ( $C_2$ ) and corresponding symmetry species ( $B_{1g/u}, B_{2g/u}, B_{3g/u}$ ) that appear in the character table of the  $D_{2h}$  group are ambiguous. As often done in the literature, we will follow the Mulliken recommendation for symmetry labels [33]. We label the Mulliken axes  $x_M$ ,  $y_M$ , and  $z_M$ . The  $z_M$ -axis is parallel to CC and is identical to our choice of  $z$ -axis. However, the  $x_M$ -axis should be perpendicular to the plane of the molecule. As pointed out in Ref. [31], this may induce some confusion between  $B_{2g/u}$  and  $B_{3g/u}$  symmetry species, and the orientation of the current  $x$  and  $y$  according to the value of the torsion angle,  $\varphi$  or  $\gamma$ . For instance, they coincide with  $x_M$  and  $y_M$  when  $\gamma = \pi/2 \pmod{2\pi}$  but are inverted when  $\gamma = 0 \pmod{\pi}$ . The same convention can be generalized to  $D_2$  geometries. For a general value of  $\gamma$ ,  $x_M$  and  $y_M$  correspond to the greater and lesser of the two perpendicular



**Fig. 4.** The symmetrized pyramidalization angles:  $\tau_U$  or  $\pi - \tau_U$  (*ungerade* pyramidalization: from  $x^R$  or  $x^L$  to the four dotted lines) and  $\pm \tau_G$  (*gerade* pyramidalization: from the dotted lines to the projections of the four CH bonds).

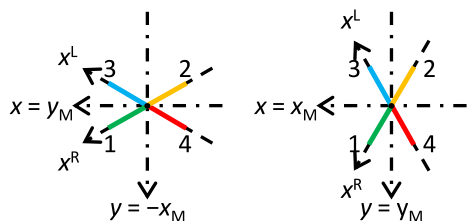


Fig. 5. Comparison between the bisecting BF axes and Mulliken's axes for  $D_2$  geometries (left panel:  $0 < \gamma < \pi/4$ ; right panel:  $\pi/4 < \gamma < \pi/2$ ).

principal moments of inertia. Our choice of  $(x,y)$ -axes is such as they vary smoothly with  $\gamma$ , whereas the Mulliken axes experience an abrupt relabeling when the molecule passes through a  $D_{2d}$  geometry as illustrated in Fig. 5. Finally, note that, at a  $D_{2d}$  geometry, there is no need to distinguish between  $x$  and  $y$  because they form a degenerate representation that transforms as an  $E$  symmetry species.

### 3. Electronic states

We now give details on the calculation of the electronic states of ethylene. Starting from the  $D_{2h}$  equilibrium geometry in the electronic ground state, electronic energies were calculated along scans involving the six internal degrees of freedom defined in Section 2,  $\{r, \beta_S, \beta_A, \tau_U, \tau_G, \gamma\}$ . We are interested in seven singlet electronic states: the first three valence states,  $N$  (normal,  $\pi^2$ ),  $V$  (valence,  $\pi\pi^*$ ), and  $Z$  (zwitterionic,  $\pi^*\pi^2$ ) in Merer-Mulliken's notation [34], and the first four Rydberg states,  $R(3s)$ ,  $R(3p_x)$ ,  $R(3p_y)$ , and  $R(3p_z)$  [34,32,21]. The notation  $R(nf)$  means that the Rydberg state corresponds essentially to a  $\pi$ -to- $nf$  single excitation where  $nf$  is a virtual orbital in the semi-united atom (SUA) formalism. Note that these notations refer to single configurations and will label either the electronic states where these configurations are dominant or the electronic configurations themselves.

The four Rydberg states are energetically embedded around the  $V$  and  $Z$  valence states around the equilibrium geometry and separate only later, around twisted geometries. These seven states are suspected to dominate the photochemistry of ethylene because a series of real and avoided crossings occurs along the reaction path that lead to ground-state regeneration after light absorption [21]. The first two bright states are  $1^1B_{3u}(R(3s))$  and  $1^1B_{1u}(V)$  under the electric-dipole approximation [19], although  $1^1B_{1g}(R(3p_y))$  may borrow intensity as it couples with  $V$  at non-planar geometries [32,17] (both states belong to  $B_1$  in  $D_2$ ). Although it has been claimed that the effect of the non-adiabatic couplings between valence and Rydberg states could be neglected on the ground that the overlap between compact and diffuse states is small, the received assumption in the most recent works favors the need for a correct treatment of Rydberg-valence interactions. Here, we show that we need to consider more than seven electronic configurations to get a description consistent at all geometries. As a result, the electronic states of interest can be described qualitatively with the interaction of 17 selected configurations involving two electrons in nine active orbitals, a good reference for more accurate calculations.

#### 3.1. Active orbitals and selected configurations

We explain here the approach we followed to identify a selected active space of 17 configurations adapted to run MCSCF calculations stable with respect to large-amplitude deformations. This proved to be non-trivial for reasons related to dynamical symmetry. State-averaged orbitals optimized for all electronic states with equal weights were used to get a correct description of the many

state-crossings encountered along all considered deformations. This also allowed an easier control of the stability of the active orbitals when these have very small weights or are almost doubly occupied in the configurations involved in the electronic states of interest. All calculations presented below were made with a development version of the Gaussian quantum chemistry program [35].

For the electronic states we are interested in, the 15 most relevant molecular orbitals of  $D_{2h}$  ethylene are gathered in Table 1 where they are given (i) with their  $D_{2h}$  labels, (ii) as approximate linear combinations of atomic orbitals (LCAO) centered on C, (iii) with mention of the bonding/antibonding overlap with the  $1s(H)$  atomic orbitals, (iv) in an SUA form, (v) with their dominant bonding character, and (vi) with the restricted active space (RAS) to which they belong in the MCSCF calculations presented below (the 15 orbitals will be partitioned in three groups: RAS1 for orbitals almost doubly occupied; RAS2 for orbitals with varying occupancies; RAS3 for orbitals almost vacant). The SUA formalism correlates the molecular orbitals of ethylene with the atomic orbitals of Si (the four core electrons are treated as a  $K$ -shell with two electrons) and is used to label the Rydberg orbitals ( $ns, np_z, nd_{z^2}, np_x, np_y, nd_{xz},$  and  $nd_{yz}$  are also termed  $ns\sigma, np\sigma, nd\sigma, np\pi_x, np\pi_y, nd\pi_x,$  and  $nd\pi_y$  by some authors). Note that  $nd$  orbitals start here with  $n=2$  only to label the Rydberg states  $R(nd)$  with  $n=3$  as usually done in the literature.

The relevant  $D_{2d}$  orbitals are listed in Table 2. From  $D_{2h}$  to  $D_{2d}$ ,  $a_g$  and  $b_{1u}$  orbitals correlate with  $a_1$  and  $b_2$  orbitals, respectively. The  $D_{2h}$  orbitals  $b_{2g}$  and  $b_{3u}$  on one hand, and  $b_{2u}$  and  $b_{3g}$  on the other hand become degenerate pairs of  $e$  orbitals that mix together through hyperconjugation.

Using the Mulliken  $(x_M, y_M)$ -axes to label the  $p$  orbitals in  $D_{2d}$  would be confusing. Axes corresponding to the rotor-fixed BF frames are simpler to deal with. In  $D_{2d}$ ,  $\gamma = \pi/4 \pmod{2\pi}$ , and  $x^R$  is orthogonal to  $x^L$  (in the plane of the  $CH_2$  groups) and  $y^R$  to  $y^L$  (out of the plane of the  $CH_2$  groups). We introduce a simplified notation for  $e$  orbitals:  $(\sigma_R, \sigma_L) = (2p_{x^R}, 2p_{x^L})$  (with positive CH bonding);  $(p_R, p_L) = (2p_{y^R}, 2p_{y^L})$ ;  $(\sigma'_R, \sigma'_L) = (2p_{x^R}, 2p_{x^L})$  (with negative CH bonding);  $(p'_R, p'_L) = (3p_{y^R}, 3p_{y^L})$ . The R/L components of a pair of valence-bond-like fragment orbitals are not orthogonal in general, except in perpendicular geometries such as  $D_{2d}$ .

The relationships with the  $D_{2h}$  delocalized orbitals (using the Mulliken convention) read:

$$2p_y = \frac{\sigma_R + \sigma_L}{\sqrt{2}}, \quad (12a)$$

$$2d_{yz} = \frac{\sigma_R - \sigma_L}{\sqrt{2}}, \quad (12b)$$

Table 1

The 15 most relevant orbitals in  $D_{2h}$  ( $x$  and  $y$  here correspond to Mulliken's convention:  $x$  is the out-of-plane direction).

| $D_{2h}$  | LCAO on C centers | CH bonding | SUA        | Global bonding | RAS |
|-----------|-------------------|------------|------------|----------------|-----|
| $2a_g$    | $2s + 2s$         |            | $2s$       | $\sigma_{CC}$  | 1   |
| $2b_{1u}$ | $2s - 2s$         | (+)        | $2p_z$     | $\sigma_{CH}$  | 1   |
| $1b_{2u}$ | $2p_y + 2p_y$     | (+)        | $2p_y$     | $\sigma_{CH}$  | 1   |
| $3a_g$    | $2p_z - 2p_z$     | (+)        | $2d_{z^2}$ | $\sigma_{CH}$  | 1   |
| $1b_{3g}$ | $2p_y - 2p_y$     | (+)        | $2d_{yz}$  | $\sigma_{CH}$  | 1   |
| $1b_{3u}$ | $2p_x + 2p_x$     |            | $2p_x$     | $\pi_{CC}$     | 2   |
| $1b_{2g}$ | $2p_x - 2p_x$     |            | $2d_{xz}$  | $\pi'_{CC}$    | 2   |
| $4a_g$    | $2p_z - 2p_z$     | (-)        | $3s$       | $\sigma'_{CH}$ | 2   |
| $3b_{1u}$ | $2p_z + 2p_z$     | (-)        | $3p_z$     | $\sigma'_{CH}$ | 2   |
| $2b_{2u}$ | $2p_y + 2p_y$     | (-)        | $3p_y$     | $\sigma'_{CH}$ | 2   |
| $2b_{3g}$ | $2p_y - 2p_y$     | (-)        | $3d_{yz}$  | $\sigma'_{CH}$ | 2   |
| $2b_{3u}$ | $3p_x + 3p_x$     |            | $3p_x$     | $\pi'_{CC}$    | 2   |
| $2b_{2g}$ | $3p_x - 3p_x$     |            | $3d_{xz}$  | $\pi'_{CC}$    | 2   |
| $5a_g$    | $3p_z - 3p_z$     |            | $3d_{z^2}$ | $\sigma'_{CC}$ | 2   |
| $4b_{1u}$ | $2p_z + 2p_z$     |            | $4p_z$     | $\sigma'_{CC}$ | 3   |

**Table 2**The 15 most relevant orbitals in  $D_{2d}$  ( $x$  and  $y$  refer to Mulliken's convention in  $D_{2h}$ ).

| $D_{2d}$   | LCAO on C centers            | CH bonding | SUA                                     | Global bonding            | RAS |
|------------|------------------------------|------------|---|---------------------------|-----|
| $2a_1$     | $2s + 2s$                    |            | $2s$                                    | $\sigma_{CC}$             | 1   |
| $2b_2$     | $2s - 2s$                    | (+)        | $2p_z$                                  | $\sigma_{CH}$             | 1   |
| $3a_1$     | $2p_z - 2p_z$                | (+)        | $2d_{z^2}$                              | $\sigma_{CH}$             | 1   |
| $(1e, 2e)$ | $(\sigma_R, \sigma_L)$       | (+)        | $(2p_y \pm 2d_{yz})$                    | $\sigma_{CH}\pi_{CC}$     | 1   |
| $(3e, 4e)$ | $(p_R, p_L)$                 |            | $(2p_x \pm 2d_{xz})$                    | $\sigma_{CH}\pi_{CC}$     | 2   |
| $4a_1$     | $2p_z - 2p_z$                | (-)        | $3s$                                    | $\sigma_{CH}^+$           | 2   |
| $3b_2$     | $2p_z + 2p_z$                | (-)        | $3p_z$                                  | $\sigma_{CH}^+$           | 2   |
| $(5e, 6e)$ | $(\tilde{p}_R, \tilde{p}_L)$ | (-)        | $(3p_x - 3d_{yz} \pm 3p_y \pm 3d_{xz})$ | $\sigma_{CH}^+\pi_{CC}^+$ | 2   |
| $(7e, 8e)$ | $(\tilde{d}_R, \tilde{d}_L)$ | (-)        | $(3d_{yz} + 3p_x \pm 3d_{xz} \mp 3p_y)$ | $\sigma_{CH}^+\pi_{CC}^+$ | 2   |
| $5a_1$     | $3p_z - 3p_z$                |            | $3d_{z^2}$                              | $\sigma_{CC}^+$           | 2   |
| $4b_2$     | $2p_z + 2p_z$                |            | $4p_z$                                  | $\sigma_{CC}^+$           | 3   |

$$2p_x(\pi) = \frac{p_R + p_L}{\sqrt{2}}, \quad (12c)$$

$$2d_{xz}(\pi^*) = \frac{p_R - p_L}{\sqrt{2}}, \quad (12d)$$

$$3p_y = \frac{\sigma_R^* + \sigma_L^*}{\sqrt{2}}, \quad (12e)$$

$$3d_{yz} = \frac{\sigma_R^* - \sigma_L^*}{\sqrt{2}}, \quad (12f)$$

$$3p_x(\pi') = \frac{p'_R + p'_L}{\sqrt{2}}, \quad (12g)$$

$$3d_{xz}(\pi'^*) = \frac{p'_R - p'_L}{\sqrt{2}}. \quad (12h)$$

Hyperconjugation between  $(\sigma_R, \sigma_L)$  (in RAS1) and  $(p_R, p_L)$  (in RAS2) actually gives some  $\sigma_{CH}$  character to the  $\pi_{CC}$  and  $\pi_{CC}^*$  orbitals (and reciprocally). This implies that the orbitals used to define RAS1 and RAS2 in  $D_{2h}$  undergo a smooth adaptation along the torsional deformation that leads to  $D_{2d}$ . However, the mixture is small enough for the original bonding character to stay dominant and RAS1/RAS2 to be considered stable. This is not obvious when looking at the orbitals in Fig. 6, but examining the atomic orbital coefficients shows that the  $\sigma$  component does not involve orbitals on the other C center and only a bit on the two H centers.

Hyperconjugation happens also to the virtual orbitals  $(\sigma_R^*, \sigma_L^*)$  and  $(p'_R, p'_L)$ . They all belong to RAS2, so there is no issue related to the stability of RAS2. The orbitals shown in Fig. 6 correspond to a mixture of about half-half when correlating back to  $D_{2h}$  orbitals. We introduce the following notation:

$$\tilde{p}'_R = \frac{p'_R + \sigma_L^*}{\sqrt{2}}, \quad (13a)$$

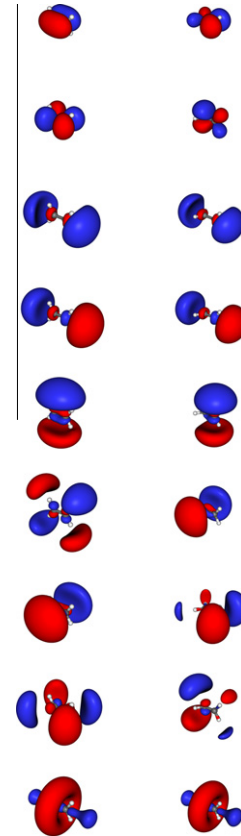
$$\tilde{p}'_L = \frac{p'_L - \sigma_R^*}{\sqrt{2}}, \quad (13b)$$

$$\tilde{d}_R = \frac{p'_R - \sigma_L^*}{\sqrt{2}}, \quad (13c)$$

$$\tilde{d}_L = \frac{p'_L + \sigma_R^*}{\sqrt{2}}, \quad (13d)$$

where the overlaps between  $(\sigma_R^*, \sigma_L^*)$  and  $(p'_R, p'_L)$  are neglected.

RAS1 corresponds to 10 electrons in  $\sigma_{CC}$  and the four  $\sigma_{CH}$  orbitals and RAS3 to the vacant  $\sigma_{CC}^*$  orbital. Accurate calculations (see below) involving a large RAS1/RAS2/RAS3 active space showed that RAS1/RAS3 orbitals always had occupancies of about 1.99/0.01 for the electronic states of interest. These could thus be neglected in a first approximation. In contrast, RAS2 is made of two electrons distributed in nine orbitals with much-varying occupancies. This minimal set of active orbitals (see Fig. 6) was found after many unsuccessful attempts to get a self-contained set of electronic configurations stable for all geometries reached by the six internal coordinates presented in Section 2. Some are built on valence



**Fig. 6.** The nine RAS2 orbitals in  $D_{2h}$  (left column; experimental geometry) and  $D_{2d}$  (right column; same geometry, except for  $\varphi = 90^\circ$ ) optimized at the SA-17-SASSCF(2,9)/6-31+G\* level of theory (see below) and given in the order used in Tables 1 and 2.

atomic orbitals ( $1s(H), 2s(C), 2p(C)$ ), whereas others are purely Rydberg (built on  $3s(C)$  or  $3p(C)$  atomic orbitals). The  $\pi_{CC}$  and  $\pi_{CC}^*$  orbitals are called valence molecular orbitals while the  $\sigma_{CH}$  orbitals are called Rydberg molecular orbitals by analogy with the  $\sigma_{HH}$ ,  $\sigma_{HH}^*$ , and Rydberg orbitals of  $H_2$  [34]. This correspondence could be disputed because the so-called Rydberg states are remarkably low-lying in ethylene and are embedded in the valence-state manifold. Except for  $R(3p_x)$  they all have a  $\pi\sigma^*$  valence character.

We now explain how RAS2 was obtained. Six out of nine of our RAS2 orbitals are an obvious choice to generate a minimal complete active space (CAS) capable of describing the seven electronic states of interest (see, e.g., Refs. [36,21]). Five molecular orbitals ( $\pi, \pi^*, 3s, 3p_y$ , and  $3p_z$ ) are mostly built on valence atomic orbitals and are used to describe the configurations corresponding to  $N$ ,

$V$ ,  $Z$ ,  $R(3s)$ ,  $R(3p_y)$  and  $R(3p_z)$ . The additional molecular orbital built on the two Rydberg atomic orbitals  $3p_x(C)$  (labeled  $\pi'$  because it is similar to  $\pi$  with an extra nodal surface) is used to describe the other low-lying Rydberg state  $R(3p_x)$ . Note that an antibonding interaction between C and H atoms, such as in  $3p_y$ , means an extra node, so even though this orbital is built on  $2p(C)$  it looks somewhat similar to  $3p_x$  built on  $3p(C)$  (see Fig. 6).

We first tried an SA-7-CASSCF(2,6)/6-31+G\* calculation of the electronic energies scanned along the torsional angle (SA-7 means that the orbitals are state-averaged over the first seven states). This corresponds to the same reference space as in the multireference configuration interaction (MRCI) calculations made by Barbatti et al. [21]. By comparison, our active space will lack some dynamical correlation, but it provides the static correlation needed to describe the seven quasidegenerate states qualitatively. That was obtained indeed, but the CAS(2,6) generated at the ground-state equilibrium geometry proved to be unstable when approaching perpendicular geometries due to the replacement of one of the states by an intruder state, thus inducing a discontinuity in the state-averaged orbitals.

The origin of the problem was found in the stability of the set of active orbitals with respect to dynamical symmetry. In particular, the  $3d_{xz}(\pi^*)$  and  $3d_{yz}$  orbitals become degenerate with  $3p_x(\pi')$  and  $3p_y$ , respectively, in  $D_{2d}$  (see Table 2) and should thus be added to the minimal active space to preserve its stability with respect to the torsional deformation. Adding these orbitals created two new Rydberg states ( $R(3d_{xz})$  and  $R(3d_{yz})$ ) between  $R(3p_x)$  and  $Z$ , thus implying to run SA-9-CASSCF(2,8) calculations.

This description proved insufficient to get stable results because of the Jahn–Teller degeneracy experienced by the electronic states  $R(3s)$  or  $R(3p_z)$  at  $D_{2d}$  geometries. Again, this is a consequence of dynamical symmetry, as  $\pi$  and  $\pi^*$  correlate to a pair of degenerate  $e$  orbitals in  $D_{2d}$  (see Table 2). Thus, correlating back to  $D_{2h}$  geometries implies to add the electronic configurations  $\pi^*3s$  and  $\pi^*3p_z$ . We will note such symmetry partners  $R^*(3s)$  and  $R^*(3p_z)$ . Barbatti et al. [21] did not meet this issue, as their MRCI calculations involved all necessary configurations not present in their reference CAS.

In addition, excitation from  $(\pi, \pi^*)$  to either  $(3p_y, 3d_{yz})$  or  $(3p_x, 3d_{xz})$  leads to two groups of four non-degenerate configurations ( $E \otimes E = A_1 \oplus A_2 \oplus B_1 \oplus B_2$ ). Treating this in a balanced way implies to run SA-15-CASSCF(2,8) calculations involving the three valence configurations and the 12 Rydberg configurations  $R(nf)$  and  $R^*(nf)$  corresponding to single excitations from  $\pi$  or  $\pi^*$  to each of the remaining six virtual orbitals  $nf$ .

Calculations made at other geometries (pyramidalized, etc.) showed the need for an extra Rydberg orbital,  $3d_{z^2}(\sigma_{CC}')$ , because the corresponding Rydberg state  $R(3d_{z^2})$  strongly interacts with the other states. This thus defines the nine active orbitals of RAS2 (see Tables 1 and 2) for SA-17-CASSCF(2,9) calculations. The dominant configurations of the 17 electronic states are gathered in Table 3 in  $D_{2h}$ , and in Table 4 in  $D_{2d}$ , where they also are given as combinations of  $D_{2h}$ -adapted configurations.

Finally, we noticed the appearance of high-lying intruder states at some geometry corresponding to double excitations to Rydberg orbitals. For example,  $3s^2$  was shown to be the 18th state in  $D_{2d}$  but the 17th state in  $D_{2h}$ , thus crossing  $R^*(3d_{xz})$  partway through the torsional reaction path. The complete active space CAS(2,9) (termed RAS2) made of 38 configurations was thus restricted by hand to the 17 dominant configurations given in Table 3. This configuration selection proved to make the subspace really stable with respect to all considered deformations for the 17 adiabatic electronic states calculated. No significant discrepancy was found regarding the energetics between the two approaches, thus supporting the validity of the configuration selection. This approach will be termed SA-17-SASSCF(2,9)/6-31+G\*, where SAS stands for

**Table 3**

Dominant configurations of the 17 electronic states in  $D_{2h}$  ( $x$  and  $y$  here correspond to Mulliken's convention:  $x$  is the out-of-plane direction).

| Configuration    | SUA orbitals           | $D_{2h}$ orbitals             | $D_{2h}$ symmetry |
|------------------|------------------------|-------------------------------|-------------------|
| $N$              | $\pi^2$                | $1b_{3u}^2$                   | $^1A_g$           |
| $R(3s)$          | $\pi 3s$               | $1b_{3u} 4a_g$                | $^1B_{3u}$        |
| $R(3p_z)$        | $\pi 3p_z$             | $1b_{3u} 3b_{1u}$             | $^1B_{2g}$        |
| $R(3p_y)$        | $\pi 3p_y$             | $1b_{3u} 2b_{2u}$             | $^1B_{1g}$        |
| $V + R(3d_{xz})$ | $\pi(\pi^* + 3d_{xz})$ | $1b_{3u} (1b_{2g} + 2b_{2g})$ | $^1B_{1u}$        |
| $R(3p_x)$        | $\pi 3p_x$             | $1b_{3u} 2b_{3u}$             | $^1A_g$           |
| $R(3d_z)$        | $\pi 3d_z$             | $1b_{3u} 5a_g$                | $^1B_{3u}$        |
| $R(3d_{yz})$     | $\pi 3d_{yz}$          | $1b_{3u} 2b_{3g}$             | $^1A_u$           |
| $R(3d_{xz}) - V$ | $\pi(3d_{xz} - \pi^*)$ | $1b_{3u} (2b_{2g} - 1b_{2g})$ | $^1B_{1u}$        |
| $Z$              | $\pi^2$                | $1b_{2g}^2$                   | $^1A_g$           |
| $R^*(3s)$        | $\pi^* 3s$             | $1b_{2g} 4a_g$                | $^1B_{2g}$        |
| $R^*(3p_y)$      | $\pi^* 3p_y$           | $1b_{2g} 2b_{2u}$             | $^1A_u$           |
| $R^*(3p_z)$      | $\pi^* 3p_z$           | $1b_{2g} 3b_{1u}$             | $^1B_{3u}$        |
| $R^*(3d_z)$      | $\pi^* 3d_z$           | $1b_{2g} 5a_g$                | $^1B_{2g}$        |
| $R^*(3p_x)$      | $\pi^* 3p_x$           | $1b_{2g} 2b_{3u}$             | $^1B_{1u}$        |
| $R^*(3d_{yz})$   | $\pi^* 3d_{yz}$        | $1b_{2g} 2b_{3g}$             | $^1B_{1g}$        |
| $R^*(3d_{xz})$   | $\pi^* 3d_{xz}$        | $1b_{2g} 2b_{2g}$             | $^1A_g$           |

**Table 4**

Dominant configurations of the 17 electronic states in  $D_{2d}$  ( $x$  and  $y$  refer to Mulliken's convention in  $D_{2h}$ ).

| Configuration                                     | $D_{2d}$ orbitals                 | $D_{2d}$ symmetry |
|---|-----------------------------------|-------------------|
| $N - Z$   | $p_R p_L$                         | $^1B_1$           |
| $V$   | $p_R p_R - p_L p_L$               | $^1B_2$           |
| $Z + N$   | $p_R p_R + p_L p_L$               | $^1A_1$           |
| $R(3s) \pm R^*(3s)$                               | $(p_R 3s, p_L 3s)$                | $^1E$             |
| $R(3p_z) \pm R^*(3p_z)$                           | $(p_R 3p_z, p_L 3p_z)$            | $^1E$             |
| $R^*(3p_x) - R(3d_{xz}) - R(3p_y) - R^*(3d_{yz})$ | $p_R \bar{p}'_L - p_L \bar{p}'_R$ | $^1A_2$           |
| $R(3p_x) + R^*(3d_{xz}) + R^*(3p_y) - R(3d_{yz})$ | $p_R \bar{p}'_R + p_L \bar{p}'_L$ | $^1A_1$           |
| $R(3d_{xz}) + R^*(3p_x) + R(3p_y) - R^*(3d_{yz})$ | $p_R \bar{p}'_R - p_L \bar{p}'_L$ | $^1B_2$           |
| $R(3p_x) - R^*(3d_{xz}) - R(3d_{yz}) - R^*(3p_y)$ | $p_R \bar{p}'_L + p_L \bar{p}'_R$ | $^1B_1$           |
| $R(3d_z) \pm R^*(3d_z)$                           | $(p_R 3d_z, p_L 3d_z)$            | $^1E$             |
| $R^*(3p_x) - R(3d_{xz}) + R(3p_y) + R^*(3d_{yz})$ | $p_R \bar{d}_L - p_L \bar{d}_R$   | $^1A_2$           |
| $R(3p_x) - R^*(3d_{xz}) + R(3d_{yz}) + R^*(3p_y)$ | $p_R \bar{d}_L + p_L \bar{d}_R$   | $^1B_1$           |
| $R(3d_{xz}) + R^*(3p_x) - R(3p_y) + R^*(3d_{yz})$ | $p_R \bar{d}_R - p_L \bar{d}_L$   | $^1B_2$           |
| $R(3p_x) + R^*(3d_{xz}) - R^*(3p_y) + R(3d_{yz})$ | $p_R \bar{d}_R + p_L \bar{d}_L$   | $^1A_1$           |

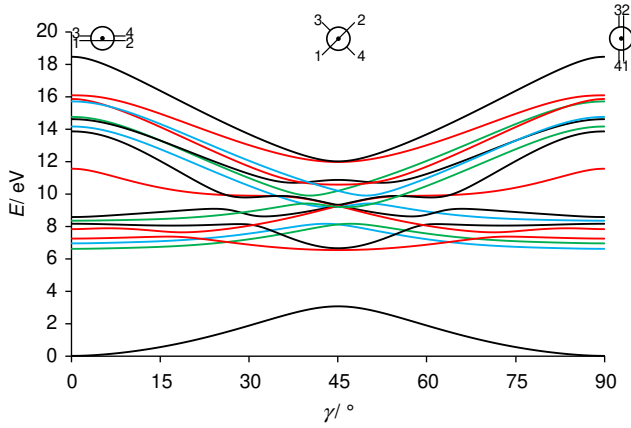
selected active space. Note that no significant improvement was obtained using the larger 6-311+G\* basis set at this level of theory.

### 3.2. Adiabatic energy scans

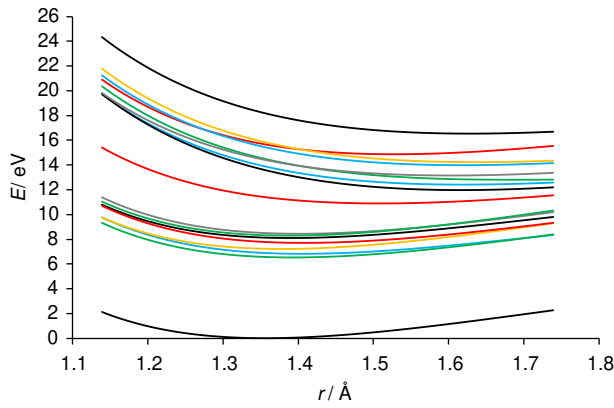
We now show the energies of the first 17 electronic states of ethylene scanned along the six internal degrees of freedom  $\{r, \beta_S, \beta_A, \tau_U, \tau_C, \gamma\}$  (see Section 2). Calculations were made at the SA-17-SASSCF(2,9)/6-31+G\* level of theory. Two initial starting points were taken for the scans: the experimental planar equilibrium geometry ( $D_{2h}$ , CC:  $r=1.339$  Å, CH:  $R_{CH}=1.086$  Å, HCH:  $\alpha=117.6^\circ$ ) [37] and the corresponding perpendicular structure with the same bond lengths and polar angles ( $D_{2d}, \gamma=45^\circ$ ). The back and forth scans along the torsional angle  $\gamma$  were checked to give identical energies and states, irrespectively of whether the active orbitals in RAS2 were originally generated at either one or the other starting points (see Fig. 6). An extra cut was added along  $\tau_R$  (or  $\tau_L$ , which is equivalent) for reasons that will be made clearer in Section 4. The adiabatic energy curves are given in Figs. 7–18 where they are distinguished according to their symmetry labels in the current point group. Ambiguities in the symmetry labeling of  $D_{2h}$ ,  $D_2$ , and  $C_{2v}$  point groups are specified in the captions.

For the scan along  $\gamma$  (see Fig. 7) varying from 0 to  $\pi/4$ , the orbitals correlating with  $\pi$  and  $\pi^*$  ( $\gamma=0$ ) are  $p_R + p_R$  and  $p_R - p_R$  (omitting the normalization factor). In  $D_{2d}$  ( $\gamma=\pi/4$ ), the orbitals become degenerate, which induces a Jahn–Teller crossing of  $R(nf)$  and

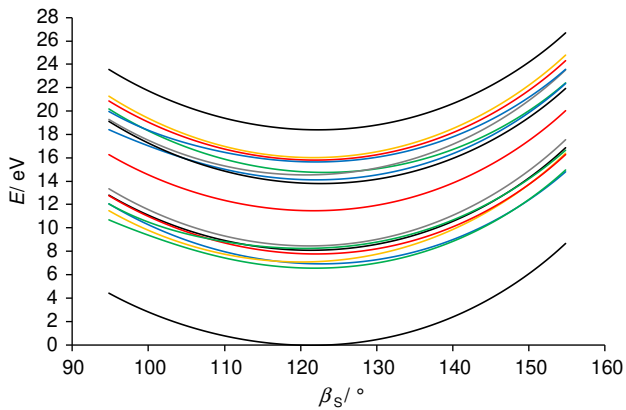




**Fig. 7.** Adiabatic energy scan along  $\gamma$  connecting the  $D_{2h}$  origin to its equivalent version through the  $D_{2d}$  origin.  $D_2$  symmetry, black:  $A$  states; red:  $B_1$  states; blue:  $B_2$  states; green:  $B_3$  states (Mulliken's convention is used for  $0 < \gamma < \pi/4$ , and a reversed convention for  $\pi/4 < \gamma < \pi/2$ ).

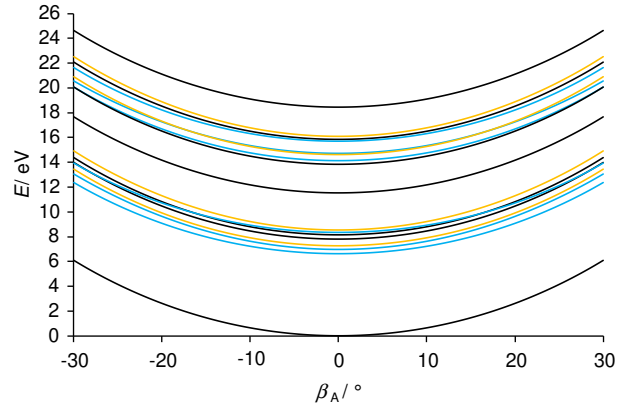


**Fig. 8.** Adiabatic energy scan along  $r$  from  $D_{2h}$  origin.  $D_{2h}$  symmetry, black:  $A_g$  states; gray:  $A_u$  states; red:  $B_{1u}$  states; orange:  $B_{1g}$  states; blue:  $B_{2g}$  states; green:  $B_{3u}$  states (Mulliken's convention).

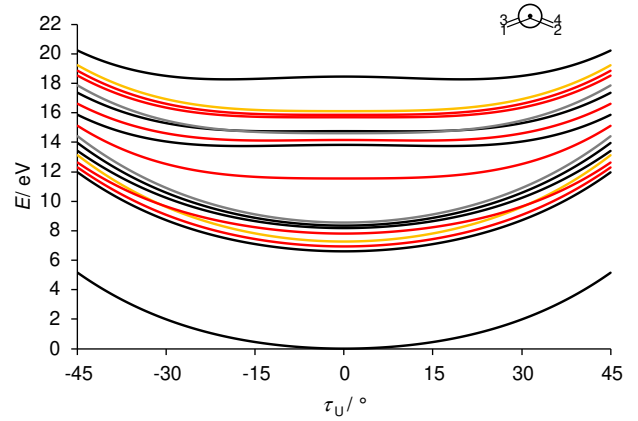


**Fig. 9.** Adiabatic energy scan along  $\beta_S$  from  $D_{2h}$  origin.  $D_{2h}$  symmetry, black:  $A_g$  states; gray:  $A_u$  states; red:  $B_{1u}$  states; orange:  $B_{1g}$  states; blue:  $B_{2g}$  states; green:  $B_{3u}$  states (Mulliken's convention).

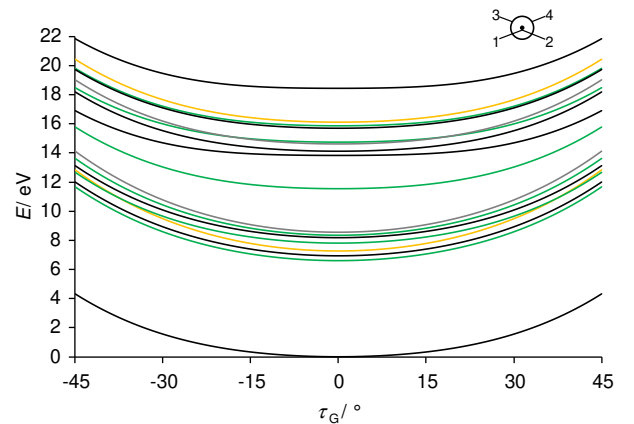
$R^*(nf)$  for  $nf = 3s, 3p_z, 3d_{z^2}$ . From  $\pi/4$  to  $\pi/2$ , a diabatic correlation of  $p_R$  and  $p_L$  (smooth rotation following the two  $\text{CH}_2$  fragments) shows that  $p_R + p_R$  is now antibonding while  $p_R - p_R$  has become bonding. If we keep calling them  $\pi$  and  $\pi^*$  for reference with the  $D_{2h}$  origin ( $\gamma = 0$ ), the energy of the  $R(nf)$  component is now higher than that of the  $R^*(3s)$  component.



**Fig. 10.** Adiabatic energy scan along  $\beta_A$  from  $D_{2h}$  origin.  $C_{2v}$  symmetry, black:  $A_1$  states; orange:  $A_2$  states; blue:  $B_1$  states (same axes as in Mulliken's convention).

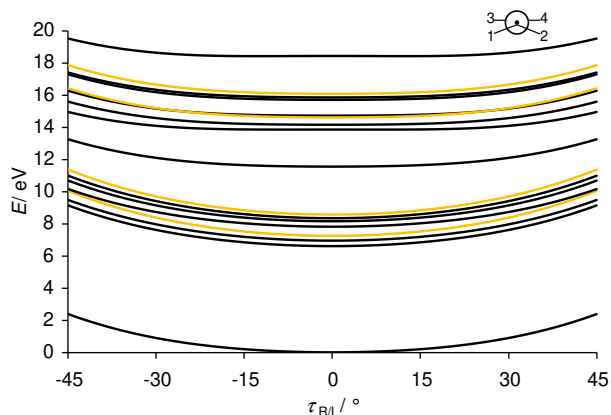


**Fig. 11.** Adiabatic energy scan along  $\tau_U$  from  $D_{2h}$  origin.  $C_{2v}$  symmetry, black:  $A_1$  states; gray:  $A_2$  states; orange:  $B_1$  states; red:  $B_2$  states ( $y$ : internuclear axis,  $z$ : out-of-plane axis,  $x$ : in-plane axis).

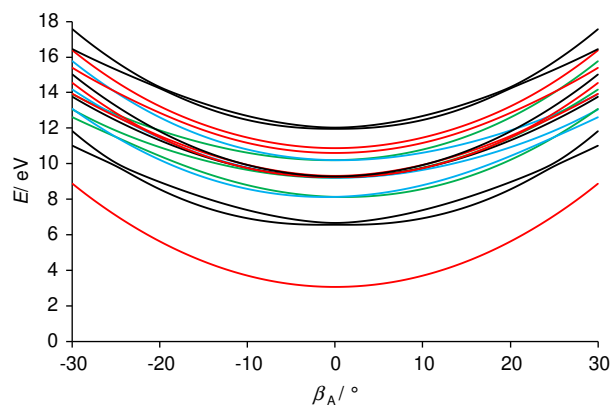


**Fig. 12.** Adiabatic energy scan along  $\tau_G$  from  $D_{2h}$  origin.  $C_{2h}$  symmetry, black:  $A_g$  states; orange:  $B_g$  states; gray:  $A_u$  states; green:  $B_u$  states.

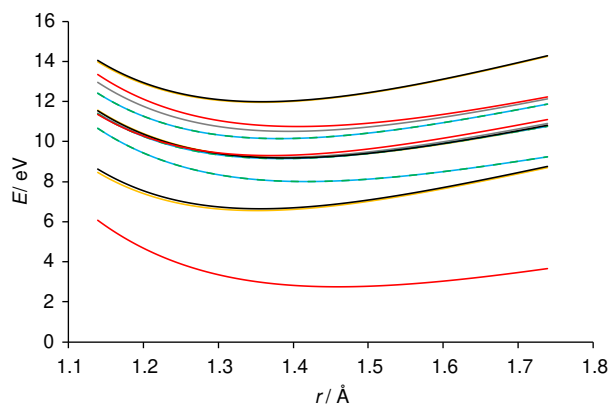
Mulliken's convention is used for  $0 < \gamma < \pi/4$ , and a reversed convention for  $\pi/4 < \gamma < \pi/2$ . This allows a common BF frame to be used on both sides of the  $D_{2d}$  point with consistent irreducible representations, and this reveals better the features of the Jahn-Teller crossing. For example, with our choice,  $R(3s)$  and  $R^*(3s)$  are  $B_3$  and  $B_2$ , respectively, all along. They cross smoothly in  $D_{2d}$  beyond which their energy curves swap. However, Mulliken's



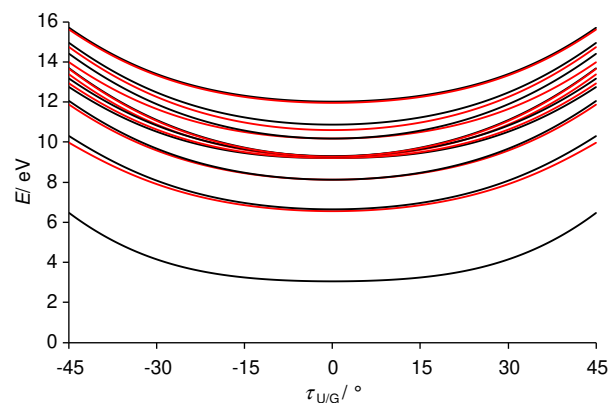
**Fig. 13.** Adiabatic energy scan along  $\tau_R$  or  $\tau_L$  from  $D_{2h}$  origin.  $C_s$  symmetry, black:  $A_1$  states; orange:  $A^+$  states.



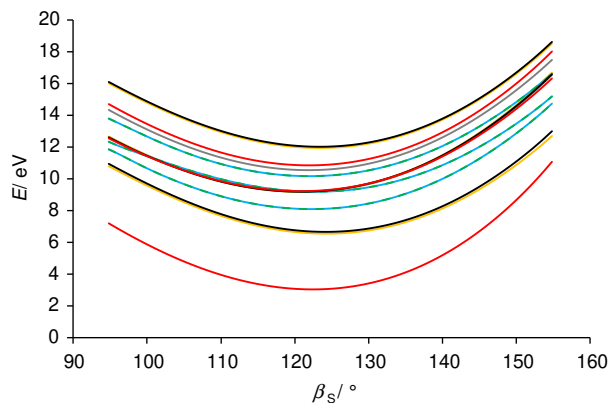
**Fig. 16.** Adiabatic energy scan along  $\beta_A$  from  $D_{2d}$  origin.  $C_{2v}$  symmetry, black:  $A_1$  states; red:  $A_2$  states; blue:  $B_1$  states; green:  $B_2$  states ( $z$ : internuclear axis,  $x||H_4H_3$ ,  $y||H_2H_1$ ).



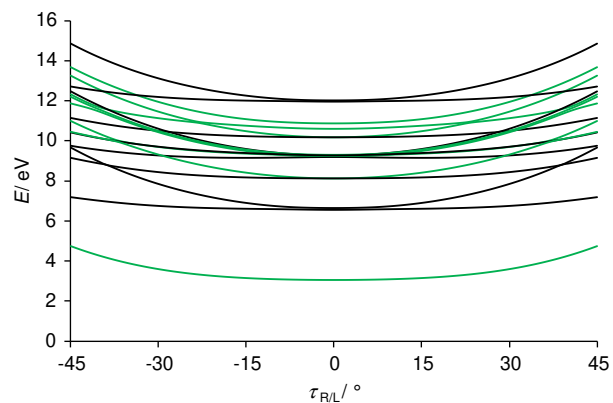
**Fig. 14.** Adiabatic energy scan along  $r$  from  $D_{2d}$  origin.  $D_{2d}$  symmetry, black:  $A_1$  states; gray:  $A_2$  states; red:  $B_1$  states; orange:  $B_2$  states; dashed blue-green:  $E$  states.



**Fig. 17.** Adiabatic energy scan along  $\tau_U$  or  $\tau_G$  from  $D_{2d}$  origin.  $C_2$  symmetry, black:  $A_1$  states; blue:  $B$  states.



**Fig. 15.** Adiabatic energy scan along  $\beta_S$  from  $D_{2d}$  origin.  $D_{2d}$  symmetry, black:  $A_1$  states; gray:  $A_2$  states; red:  $B_1$  states; orange:  $B_2$  states; dashed blue-green:  $E$  states.



**Fig. 18.** Adiabatic energy scan along  $\tau_R$  or  $\tau_L$  from  $D_{2d}$  origin.  $C_s$  symmetry, black:  $A_1$  states; green:  $A^+$  states.

convention would have led them to be  $B_2$  and  $B_3$  from  $\pi/4$  to  $\pi/2$ , thus inducing a cusp in energy curves of the same symmetry label.

The other first-order active coordinate with respect to the Jahn–Teller crossing is  $\beta_A$  (see Fig. 16). Here too, the features of the crossing are better-understood with a common BF frame on both sides. For example, with the convention chosen here for the  $C_{2v}$  group ( $z$ : internuclear axis,  $x||H_4H_3$ ,  $y||H_2H_1$ ), configurations  $p_R 3s = R(3s) + R^*(3s)$  and  $p_L 3s = R(3s) - R^*(3s)$  (omitting the normalization factor) are  $B_1$  and  $B_2$ , respectively, all along. For  $\beta_A < 0$ ,

the energy of  $p_R 3s$  is lower than that of  $p_L 3s$ . They cross smoothly in  $D_{2d}$  ( $\beta_A = 0$ ) beyond which their energy curves swap.

### 3.3. Accuracy tests

The  $V$  state of ethylene is still today a computational challenge. Its partly-valence/ionic/Rydberg character makes any description

**Table 5**

Vertical excitation energies (eV) for the valence and some of the Rydberg states at the  $D_{2h}$  equilibrium geometry with respect to the ground state (absolute energies given in footnotes). SAS: SA-17-SASSCF(2,9); MRCI: MR-CISD(12,15).

|               | SAS<br>6-31+G*    | MRCI<br>6-31+G*   | SAS<br>6-311+G*   | MRCI<br>6-311+G*  | [21]                      | [38] | Exp.  |
|---------------|-------------------|-------------------|-------------------|-------------------|---------------------------|------|---|
| $N$           | 0.00 <sup>a</sup> | 0.00 <sup>b</sup> | 0.00 <sup>c</sup> | 0.00 <sup>d</sup> | 0.00 <sup>e</sup>         | 0.00 |   |
| $R(3s)$       | 6.61              | 6.88              | 6.37              | 6.61              | 7.16                      | 7.09 | 7.11 <sup>gh</sup>  |
| $R(3p_z)$     | 6.95              | 7.25              | 6.84              | 7.10              | 7.85                      | 7.96 | 7.90 <sup>h</sup>   |
| $R(3p_y)$     | 7.25              | 7.58              | 7.07              | 7.35              | 7.82                      | 7.76 | 7.80 <sup>gh</sup>  |
| $V$           | 7.83              | 7.90              | 7.80              | 7.86              | 7.91 (7.80 <sup>f</sup> ) | 7.90 | 7.66 <sup>g</sup> , 7.8 <sup>i</sup> , 7.9 ± 0.1 <sup>j</sup> |
| $R(3p_x)$     | 8.17              | 8.29              | 7.97              | 8.09              | 8.19                      | 8.33 | 8.26 <sup>g</sup> , 8.28 <sup>h</sup>                         |
| $R(3d_{z^2})$ | 8.35              | 8.63              | 8.21              | 8.46              |                           | 8.59 | 8.62 <sup>k</sup>   |
| $R(3d_{yz})$  | 8.55              | 8.88              | 8.30              | 8.59              |                           | 9.12 |   |
| $R(3d_{xz})$  | 11.54             |                   | 11.29             |                   |                           | 9.35 | 9.33 <sup>l</sup>   |
| $Z$           | 13.84             |                   | 13.82             |                   | 12.73                     |      |   |

<sup>a</sup> -78.029923  $E_h$ .

<sup>b</sup> -78.052602  $E_h$ .

<sup>c</sup> -78.039781  $E_h$ .

<sup>d</sup> -78.059639  $E_h$ .

<sup>e</sup> -78.363611  $E_h$ .

<sup>f</sup> MR-CISD+Q/SA-3-RDP/aug-cc-pVTZ.

<sup>g</sup> From Ref. [39].

<sup>h</sup> From Ref. [40].

<sup>i</sup> From Ref. [41].

<sup>j</sup> From Ref. [42].

<sup>k</sup> From Ref. [43].

<sup>l</sup> From Ref. [44].

critical in terms of dynamical correlation ( $\sigma - \pi$ -correlation,  $p$ -contraction, etc.) as discussed, e.g., in Ref. [23] (see also references therein). We found out that, although very cheap, our SA-17-SASSCF(2,9)/6-31+G\* approach was capable of giving the energy of the  $V$  state quite remarkably (see Table 5). We suspect this comes from using orbitals forced to be state-averaged over an (unusually) large number of states, i.e., optimized for  $V$  as well as for other states that involve  $\pi$ -to- $\sigma^*$  and  $\pi$ -to- $\pi^*$  excitations, thus allowing relaxation of  $V$  (mainly a  $\pi$ -to- $\pi^*$  excitation) through configuration interaction with the other states. In particular, our calculations show that the effective  $V$  state ( $1^1B_{1u}$ ) is a balanced mixture of  $\pi\pi^*$  (ideal  $V$  state) and  $\pi\pi'^*$  (see Table 3), in other words, a  $\pi\tilde{\pi}^*$  dominant configuration where  $\tilde{\pi}^*$  is a mixture of  $\pi^*$  and  $\pi'^*$  orbitals. This is partly induced by state-averaging the orbitals over  $V$  at the same time as  $R(3d_{xz})$  and  $R^*(3p_x)$  ( $1^1B_{1u}$ ,  $2^1B_{1u}$ , and  $3^1B_{1u}$  in  $D_{2h}$ ), in a spirit similar to the recent work of Angeli [23].

The energies of the ten lower states at the experimental equilibrium geometry [37] are given in Table 5 at the SA-17-SASSCF(2,9) level of theory with the 6-31+G\* and 6-311+G\* basis sets. The quality of  $V$  is indicated by the vertical excitation energy, in agreement with the best theoretical estimates and experiments (see Table 5). For the out-of-plane diffuseness, we obtained  $\langle x^2 \rangle = 23a_0^2$  with 6-31+G\* and 24 with 6-311+G\*. Although this value lies within the range of values found with most published calculations (see Ref. [23] and references therein), it is larger than the best theoretical estimates, 16-18  $a_0^2$  [23]. This could be attributed to an overestimation of the contribution of the  $\pi'^*$  in the effective  $\tilde{\pi}^*$  orbital.

The lower Rydberg states (up to  $R(3d_{yz})$ ) are qualitatively correct but there is no quantitative agreement with experimental values, except for  $R(3p_x)$  and  $R(3d_{z^2})$ . The energetic order of  $R(3p_z)$  and  $R(3p_y)$  is reversed, and the energy can be underestimated by about 1 eV. This is to be expected because special basis sets are needed to describe such diffuse states accurately. In contrast, the higher-lying Rydberg states are useful only as mathematical states, here for giving flexibility to the lower-lying states and making the active space stable and self-contained with respect to geometrical deformations. Their energies are not listed in Table 5, as the states may not have any physical meaning.

SA-17-SASSCF(2,9) calculations were checked against subsequent MRCI-like calculations, noted MR-CISD(12,15) (SA-17-SASS-

CF(2,9)//CISD(12,15)), and involving all single and double excitations from 5 doubly-occupied orbitals and to one extra virtual orbital (see Table 5). This approach corresponds in fact to a RAS(12,15) divided into RAS1/RAS2/RAS3 (see Tables 1 and 2), with orbitals optimized from a previous SA-17-RASSCF(2,9) calculations and kept unchanged. The energy of the first few states at the experimental equilibrium geometry have been compared to experimental results and two kinds of accurate computations: (i) Ref. [21] (MR-CISD+Q/SA-7-CAS(2,2)+aux/d-aug-cc-pVDZ); (ii) Ref. [38] (IDDCI with a large basis set including atomic natural orbitals).

These calculations confirmed that  $V$  was already well-described at the SA-17-SASSCF(2,9) level of theory. The energy did not change significantly (see Table 5) nor the diffuseness ( $\langle x^2 \rangle = 22 a_0^2$  with both basis sets). The lower Rydberg states proved to be more sensitive than  $V$  to further dynamical correlation. Their energies get closer to the experimental values but are not fully converged, thus confirming again the need for basis sets adapted to Rydberg states for running more accurate calculations. The  $Z$  state and Rydberg states higher than  $R(3d_{yz})$  were not obtained with MR-CISD(12,15) calculations because states beyond about 9 eV appear embedded in a series of double excitations.

We also calculated the barrier between the experimental equilibrium geometry and the perpendicular structure with the same bond lengths and polar angles (this is not the optimized transition state). At the SA-17-SASSCF(2,9) level of theory, the barrier is equal to 3.06 (3.02) eV with the 6-31+G\* (6-311+G\*) basis set. MR-CISD(12,15) calculations give about the same value: 2.97 (2.95) eV with the 6-31+G\* (6-311+G\*) basis set. The same value is given in Ref. [21]). This confirms that no further dynamical correlation was needed to describe correctly the mixture of inactive and active orbitals through hyperconjugation.

#### 4. Dynamical symmetry analysis

In what follows, we will determine how nuclear coordinates and electronic configurations transform through nuclear-permutation-inversion operations (dynamical symmetry) in ethylene. The nuclei have been labeled such as  $H_1$  and  $H_2$  are chemically bound to  $C_5$ , and  $H_3$  and  $H_4$  to  $C_6$  (see Fig. 1). The type of deformations

considered here preserve all  $\sigma$  bonds, but  $\pi$  bonds may break upon electronic excitation from the ground electronic state, thus inducing possible large-amplitude motion in the torsion around the CC bond (barrierless process in some of the excited electronic states as opposed to tunneling in the electronic ground state, which can be too weak to be observed).

#### 4.1. The $G_{16}$ group

The symmetry group of the non-relativistic vibronic Hamiltonian is called the molecular symmetry group [14]. For rigid systems, where only small deformations around the equilibrium geometry are considered, the molecular symmetry group is isomorphic to the point group at this geometry, and all symmetry properties can be derived from the corresponding irreducible representations. However, flexible molecules can undergo large-amplitude deformations that connect various versions of the equilibrium geometry (geometries that are distinct only because identical nuclei are labeled individually). In addition, such deformations can lead to intermediate geometries where the point group is larger than the subgroup followed from the origin. See, for example, Fig. 7:  $D_{2h} (\gamma = 0) \rightarrow D_2 \rightarrow D_{2d} (\gamma = \pi/4) \rightarrow D_2 \rightarrow D_{2h} (\gamma = \pi/2)$ . Correlating adiabatic energy scans along the same coordinates and originated from  $D_{2h}$  and  $D_{2d}$  points (e.g., Figs. 8 and 14) is not obvious when working with point-group symmetry. In contrast, dynamical symmetry gives global irreducible representations that facilitate the connection of electronic states at different geometries.

For ethylene with feasible torsion the dynamical group is termed  $G_{16}$ , the subgroup made of the operations of the CNPI group  $G_{96}$  [14,31,32] that preserve the bond connectivity of the molecule. This allows a consistent description of symmetry properties for all the point groups that can be explored through large-amplitude deformations ( $D_{2h}$ ,  $D_2$ ,  $D_{2d}$ , etc.). With the choice of connectivity made in Fig. 1, the equilibrium geometry of ground-state ethylene can be represented by two distinct versions regarding nuclear labeling (see Fig. 19) that interconvert through torsion (see Fig. 2). When considered rigidly planar, each version is structurally twelve-fold degenerate and described by the molecular symmetry group  $D_{2h}(M)$ , subgroup of  $G_{16}$  isomorphic to the rigid point group  $D_{2h}$ . Note that distorted geometries may be less symmetrical and be described by a smaller point group than  $D_{2h}$ , but they always belong to the same molecular symmetry group,  $G_{16}$ . This property will be fully exploited below, in the determination of the electronic-state symmetry and their dependences on the nuclear coordinates. Finally, geometries resulting from H migration, e.g.,  $D_{4h}(C(H_4)C)$  or  $C_{3v}(H_3CC(H))$ , will not be considered here.

Now, we will consider the action of each operation of  $G_{16}$  on the nuclear coordinates and the electronic configurations mentioned above. The notations for the operations are those used in [14]. In particular, the four generating operations of  $G_{16}$  are  $E$ , the identity, and  $a = (1423)(56)^*$ ,  $b = (13)(24)(56)$ , and  $c = E^*$ , the parity (space inversion). They satisfy

$$a = (1423)(56)^*, a^2 = (12)(34), a^3 = (1324)(56)^*, a^4 = E, \quad (14a)$$

$$b = (13)(24)(56), b^2 = E, \quad (14b)$$

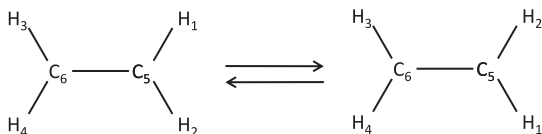


Fig. 19. The effect of the feasible permutation (12) between two  $D_{2h}$  versions of ethylene that can interconvert through torsion.

$$c = E^*, c^2 = E, \quad (14c)$$

$$a^2b = ba^2 = (14)(23)(56), ab = ba^3 = (34)^*, a^3b = ba = (12)^*. \quad (14d)$$

Coordinate transformations are summarized in Table 6. In this, we used the following properties for degenerate sets:

$$E^-(R^-, L^-) \otimes E^-(R^-, L^-) = A_1^+(R^-R^- + L^-L^-) \oplus A_2^+(R^-L^- - L^-R^-) \\ \oplus B_1^+(R^-L^- + L^-R^-) \oplus B_2^+(R^-R^- - L^-L^-), \quad (15a)$$

$$E^-(G^-, U^-) \otimes E^-(G^-, U^-) = A_1^+(G^-G^- + U^-U^-) \oplus A_2^+(G^-U^- - U^-G^-) \\ \oplus B_1^+(G^-G^- - U^-U^-) \oplus B_2^+(G^-U^- + U^-G^-), \quad (15b)$$

where (U, G) or (R, L) are two kinds of pairs of components for  $E$  representations such that (U, G) correlate with  $D_{2h}$  and (R, L) correlate with  $C_{2v}$ , the largest Abelian subgroup of  $D_{2d}$  ( $U \propto R + L$ ;  $G \propto R - L$ ).

The symmetry properties of the nuclear coordinates derive directly from the definitions of the nuclear-permutation-inversion operations. Those of the electronic configurations derive ultimately from the transformations of the atomic orbitals. This may need a few comments. First, only the space inversion (parity,  $E^*$ ) acts on the electronic coordinates in the space-fixed frame. The atomic orbitals change according to the coordinate transformation induced on the electronic and nuclear coordinates by  $E^*$ . The  $s$ ,  $p$ ,  $d$ , etc. orbitals behave as tensors of order 0, 1, 2, etc. with respect to  $E^*$  acting on the electronic coordinates, e.g.,  $s$  orbitals: scalars (positive parity);  $p$  orbitals: vectors (negative parity), etc. In addition, they move with their respective nuclear centers. Second, although the electrons themselves do not follow the nuclei through a pure nuclear permutation, the nuclear centers and the BF axes that label the atomic orbitals are transformed according to the operation. So, when labeled with BF directions, even orbitals simply transform as the corresponding monomials built on  $x, y$ , and  $z$ , whereas odd orbitals do the same for operations that do not contain  $E^*$  (e.g.  $E$  and  $b$ ) and are further transformed into their negatives for operations that contain  $E^*$  (e.g.,  $a$  and  $c$ ).

Let us consider as an example  $b = (13)(24)(56)$  and  $bc = (13)(24)(56)^*$ ,

$$bp_{xR}(C_5) = p_{xL}(C_6), \quad (16a)$$

$$bp_{yR}(C_5) = -p_{yL}(C_6), \quad (16b)$$

$$bp_z(C_5) = -p_z(C_6), \quad (16c)$$

Table 6

$G_{16}$  character table with indication of the coordinate transformations.

| $G_{16}$ | $E$ | $a$      | $b$      | $c$    | Coordinates  |
|----------|-----|----------|----------|--------|--|
| $A_1^+$  | 1   | 1        | 1        | 1      | $r, (\alpha_R + \alpha_L)/2, \beta_S, \vartheta_R^2 + \vartheta_L^2, \tau_R^2 + \tau_L^2, \tau_C^2 + \tau_U^2$ |
| $A_2^+$  | 1   | 1        | -1       | 1      |  |
| $B_1^+$  | 1   | -1       | 1        | 1      | $\cos\varphi, \cos 2\gamma, \vartheta_R\vartheta_L, \tau_R\tau_L, \tau_C^2 - \tau_U^2$                         |
| $B_2^+$  | 1   | -1       | -1       | 1      | $(\alpha_R - \alpha_L)/2, \beta_A, \vartheta_R^2 - \vartheta_L^2, \tau_R^2 - \tau_L^2, \tau_C\tau_U$           |
| $E^*$    | 2   | 0        | 0        | 2      |  |
| $(U^+)$  | (1) | $(-G^+)$ | (1)      | (1)    |  |
| $(G^+)$  | (1) | $(U^+)$  | $(-1)$   | (1)    |  |
| $(R^+)$  | (1) | $(L^+)$  | $(L^+)$  | (1)    |  |
| $(L^+)$  | (1) | $(-R^+)$ | $(R^+)$  | (1)    |  |
| $A_1^-$  | 1   | 1        | 1        | -1     | $\sin\varphi, \sin 2\gamma$  |
| $A_2^-$  | 1   | 1        | -1       | -1     |  |
| $B_1^-$  | 1   | -1       | 1        | -1     |  |
| $B_2^-$  | 1   | -1       | -1       | -1     |  |
| $E^-$    | 2   | 0        | 0        | -2     |  |
| $(G^-)$  | (1) | $(U^-)$  | (1)      | $(-1)$ | $\tau_C$   |
| $(U^-)$  | (1) | $(-G^-)$ | $(-1)$   | $(-1)$ | $\tau_U$   |
| $(R^-)$  | (1) | $(L^-)$  | $(-L^-)$ | $(-1)$ | $\vartheta_R, \tau_R$  |
| $(L^-)$  | (1) | $(-R^-)$ | $(-R^-)$ | $(-1)$ | $\vartheta_L, \tau_L$  |

$$bcp_{x^R}(C_5) = p_{x^L}(C_6), \quad (16d)$$

$$bcp_{y^R}(C_5) = p_{y^L}(C_6), \quad (16e)$$

$$bcp_z(C_5) = -p_z(C_6), \quad (16f)$$

because  $b$  swaps  $C_5$  and  $C_6$ , and  $b(x^{R,L}, y^{R,L}, z) = (x^{L,R}, -y^{L,R}, -z)$ , whereas  $c$  keeps the atomic orbitals on their centers and only changes the sign of  $p_{y^R}(C_5)$  (combined action of  $c(x^{R,L}, y^{R,L}, z) = (-x^{R,L}, y^{R,L}, -z)$  and the change of sign of  $p$ -orbitals). Note that, L and R BF  $(x, y)$ -axes do not necessarily coincide with the Mulliken axes. Also, as opposed to point-group symmetry based on invariance properties, here there is no reason for orbitals transformed into each other to be identical if the geometry does not correspond to any point-group symmetry.

## 4.2. Basis functions of the nuclear coordinates

Different sets of nuclear coordinates have been introduced in Section 2. For what follows, we need to determine simple functions of these that belong to the irreducible representations of the  $G_{16}$  group.

In the work presented in Ref. [18], symmetry considerations were drawn consistently with respect to some of the nuclear-permutation-inversion operations. However, dynamical symmetry was not used explicitly, which would have simplified the derivation greatly. In particular, functions of  $\vartheta_R$  and  $\vartheta_L(E^-(R^-, L^-)$  pair; see Table 6) were recombined in a second stage, after examination of how these two coordinates transform between each other, whereas functions of symmetrized coordinates could have been built directly (see Table 7). We now show how dynamical symmetry is to be used to generate basis functions of the nuclear coordinates adapted to  $G_{16}$ . We will show further on how these enter the electronic Hamiltonian matrix naturally when a quasidiabatic basis set adapted to  $G_{16}$  is used.

In Ref. [18], the potential energy matrix for the  $(N, V, Z)$  space was expanded in terms of two-dimensional functions of  $\varphi$  and any coordinate in  $\{r, (\alpha_R + \alpha_L)/2, (\alpha_R - \alpha_L)/2, \vartheta_R, \vartheta_L\}$ . From the dynamical symmetry properties of the nuclear coordinates in  $G_{16}$  (see Table 7), one gets the symmetry labels of product functions using the multiplication table of  $G_{16}$ . Also, from Eq. (15b), it is straightforward to show that  $\vartheta_R^4 + \vartheta_L^4 = (\vartheta_R^2 + \vartheta_L^2)^2 - 2\vartheta_R^2\vartheta_L^2$  is  $A_1^+$  and  $\vartheta_R^4 - \vartheta_L^4 = (\vartheta_R^2 + \vartheta_L^2)(\vartheta_R^2 - \vartheta_L^2)$  is  $B_2^+$ , but higher-order terms (not considered here) need more careful consideration. The basis functions used in Ref. [18] are given in Table 7 along with their  $G_{16}$  irreducible representations. Note that no  $A_2^+$  one-dimensional function is given because no non-zero  $A_2^+$  matrix element was considered in Ref. [18], although using  $B_1^+ \otimes B_2^+ = A_2^+$  would have led to either  $[(\alpha_R - \alpha_L)/2]^{2n+1} \cos^{2n+1} \varphi$  or  $[\vartheta_R^{2,4} - \vartheta_L^{2,4}] \cos^{2n+1} \varphi$  functions.

Now, using the symmetrized valence coordinates  $\{r, \beta_S, \beta_A, \tau_U, \tau_G, \gamma\}$  defined in Section 2, we can proceed to the same analysis. The  $G_{16}$  basis functions needed for building the potential energy matrix (see below) corresponding to the 17 valence and Rydberg states defined in Section 3 are given in Table 8.

We prefer to use real Fourier basis functions rather than powers of  $\cos \varphi$  and  $\sin \varphi$  for simplicity, although both kinds transform into

**Table 7**

Dynamical-symmetry-adapted basis functions of the nuclear coordinates used in Ref. [18] ( $m = 1, 2$ ;  $n = 1$  to 4;  $p = 1, 2$ ;  $q = 1$ ;  $r_0$  and  $\alpha_0$  are particular values defining the  $D_2$  equilibrium geometry along with  $R_{CH}$ ).

| $G_{16}$ | $\varphi$             | $r$           | $(\alpha_R + \alpha_L)/2$                | $(\alpha_R - \alpha_L)/2$          | $(\vartheta_R, \vartheta_L)$            |
|----------|-----------------------|---------------|--|------------------------------------|---|
| $A_1^+$  | $\sin^{2m} \varphi$   | $(r - r_0)^n$ | $[(\alpha_R + \alpha_L)/2 - \alpha_0]^p$ | $[(\alpha_R - \alpha_L)/2]^{2q}$   | $\vartheta_R^{2,4} + \vartheta_L^{2,4}$ |
| $B_1^+$  | $\cos^{2m-1} \varphi$ |               |  | $[(\alpha_R - \alpha_L)/2]^{2q-1}$ | $\vartheta_R^{2,4} - \vartheta_L^{2,4}$ |

each other directly. Note that some restrictions must be made on the values of the powers to avoid redundancies in the latter case. Also, using Eq. (15b) and the same derivations as for  $(\vartheta_R, \vartheta_L)$  leads to showing that  $\tau_C^4 + \tau_U^4$  is  $A_1^+$  and  $\tau_C^4 - \tau_U^4$  is  $B_1^+$ .

## 4.3. Active orbitals

The active orbitals involved in the 17 reference configurations selected within the RAS2 space have been presented in Section 3. Transformations between localized and delocalized orbitals are given in Eq. (12h). Table 9 summarizes their symmetry properties with respect to  $G_{16}$  where they also are correlated to three point groups:  $D_{2h}$ ,  $D_{2d}$ , and  $C_{2v}$ . The latter corresponds to the largest Abelian subgroup of  $D_{2d}$  with left-right broken-symmetry ( $B_1$  and  $B_2$  are symmetric with respect to the left and right HCH planes, respectively). Note that the  $B_{2(3)u(g)}$  irreducible representations of  $D_{2h}$  are correlated assuming  $y^R = y^L = \pm x_M$  and  $x^R = x^L = \mp y_M$ .

We now enunciate the hypotheses that will be the ground of our quasidiabatic description. First, we will consider the set of nine active orbitals in RAS2 self-contained everywhere. In practice, the active orbitals adapt themselves smoothly along geometrical deformations: (i) they translate and rotate with the atom centers and the BF axes, which we consider as a diabatic adaptation; (ii) they mix with orbitals not originally in RAS2 (in particular, hyperconjugation between  $(2p_y, 2d_{yz})$  and  $(2p_x, 2d_{xz})$  gives some  $\sigma_{CH}$  character to the  $\pi_{CC}$  and  $\pi_{CC}^*$  orbitals). Although we treat the problem formally as if this mixing did not occur, we do not prevent it numerically as it allows the configuration space to keep the same size and still be stable with respect to geometrical deformations. The ‘‘diabaticity’’ criterion is the continuity of the 17 eigenvalues over geometrical deformations. As already mentioned, finding the right active space was not a trivial task. Many unsuccessful attempts led to discontinuities in the energies at some geometries.

A very important difference between dynamical symmetry and point-group symmetry is the electronic-integral rule such as orbital overlaps. In the latter case, the overlap is not zero only if the product of the two orbitals contains the totally symmetric representation. In dynamical symmetry, the overlap simply is a function of the nuclear coordinates that has the same symmetry as the orbital product. The electronic-integral rule at a given geometry actually comes from the fact that only functions belonging to the totally symmetric representation are non-zero in the corresponding point group (a finite value of any other function can only be due to a geometrical displacement that breaks the local symmetry). As a consequence, it is not possible to assume zero overlaps in a global way when using dynamical symmetry.

However, as a second hypothesis, we will neglect such overlaps between orbitals belonging to different irreducible representations in  $G_{16}$ . The orbitals listed in Tables 1 and 2, shown in Fig. 6, have been calculated at either one of two starting points:  $D_{2h}$  (experimental ground-state equilibrium geometry) and  $D_{2d}$  (same geometry, except for  $\gamma = 45^\circ$ ), respectively. The  $D_{2h}$  orbitals already belong to irreducible representations of  $G_{16}$ . The MCSCF active orbitals propagated from  $D_{2h}$  to  $D_{2d}$  have adapted themselves to the geometrical deformation. They still are orthogonal, but they have mixed and no longer belong to the irreducible representations of  $G_{16}$ . Hyperconjugation can be neglected for  $(p_R, p_L)$  ( $E^-$ ) but not between  $(\sigma_R^*, \sigma_L^*)(E^+)$  and  $(p_R', p_L')(E^-)$  that both are  $E$  in  $D_{2d}$  (see Eq. (13d)). Here, we assume that it is possible to find a smooth transformation that separates contributions from different irreducible representations of  $G_{16}$  while keeping the orbitals orthogonal. The ‘‘diabaticity’’ criterion is that orbitals change as smoothly as possible from the starting point where they were defined.

It is to be understood that such hypotheses are needed for simplifying the interpretation of results but not to produce them numerically. The aim is to describe the adiabatic electronic states

**Table 8**  
Dynamical-symmetry-adapted basis functions of the symmetrized valence coordinates ( $j, m, n, p$ , and  $q$  are any positive integers;  $r_0$  and  $\beta_0$  are particular values defining the  $D_2$  equilibrium geometry along with  $R_{CH}$ ).

| $G_{16}$      | $\gamma$  | $r$           | $\beta_S$               | $\beta_A$        | $(\tau_G, \tau_U)$            |
|---------------|---|---------------|-------------------------|------------------|-------------------------------|
| $A_1^+$       | $\cos^{2j}2\gamma \sin^{2m}2\gamma$ or $\cos 4m\gamma$      | $(r - r_0)^n$ | $(\beta_S - \beta_0)^p$ | $\beta_A^{2q}$   | $\tau_G^{2.4} + \tau_U^{2.4}$ |
| $B_1^+$       | $\cos^{2j+1}2\gamma \sin^{2m}2\gamma$ or $\cos(4m+2)\gamma$ |               |                         |                  | $\tau_G^{2.4} - \tau_U^{2.4}$ |
| $B_2^+$       |   |               |                         | $\beta_A^{2q+1}$ | $\tau_G \tau_U$               |
| $A_1^-$       | $\cos^{2j}2\gamma \sin^{2m+1}2\gamma$ or $\sin(4m+2)\gamma$ |               |                         |                  |                               |
| $B_1^-$       | $\cos^{2j+1}2\gamma \sin^{2m+1}2\gamma$ or $\sin 4m\gamma$  |               |                         |                  |                               |
| $E(G^-, U^-)$ |   |               |                         |                  | $(\tau_G, \tau_U)$            |

**Table 9**  
Active orbitals in  $G_{16}$  (SUA notation given with the Mulliken axis convention).

| $G_{16}$        | Approximate LCAO    | Short Name                 | SUA                | $D_{2h}$           | $D_{2d}$ | $C_{2v}$     |
|-----------------|---------------------|----------------------------|--------------------|--------------------|----------|--------------|
| $E^-(R^-, L^-)$ | $(2p_y^R, 2p_y^L)$  | $(p_R, p_L)$               | $2p_x \pm 2d_{xz}$ |                    | $E$      | $(B_1, B_2)$ |
| $E^-(U^-, G^-)$ | $2p_y^R \pm 2p_y^L$ | $(\pi, \pi^*)$             | $(2p_x, 2d_{xz})$  | $(B_{3u}, B_{2g})$ | $E$      |              |
| $A_1^+$         | $2p_z - 2p_z$       | $3s$                       | $3s$               | $A_g$              | $A_1$    | $A_1$        |
| $B_2^+$         | $2p_z + 2p_z$       | $3p_z$                     | $3p_z$             | $B_{1u}$           | $B_2$    | $A_1$        |
| $E^+(R^+, L^+)$ | $(2p_x^R, 2p_x^L)$  | $(\sigma_R^+, \sigma_L^+)$ | $3p_y \pm 3d_{yz}$ |                    | $E$      | $(B_2, B_1)$ |
| $E^+(U^+, G^+)$ | $2p_x^R \pm 2p_x^L$ | $(3p_y, 3d_{yz})$          | $(3p_y, 3d_{yz})$  | $(B_{2u}, B_{3g})$ | $E$      |              |
| $E^-(R^-, L^-)$ | $(3p_y^R, 3p_y^L)$  | $(p'_R, p'_L)$             | $3p_x \pm 3d_{xz}$ |                    | $E$      | $(B_1, B_2)$ |
| $E^-(U^-, G^-)$ | $3p_y^R \pm p_y^L$  | $(3p_x, 3d_{xz})$          | $(3p_x, 3d_{xz})$  | $(B_{3u}, B_{2g})$ | $E$      |              |
| $A_1^+$         | $3p_z - 3p_z$       | $3d_{z^2}$                 | $3d_{z^2}$         | $A_g$              | $A_1$    | $A_1$        |

of interest as mixtures of 17 quasidiabatic configurations of well-defined symmetry in  $G_{16}$ . Each configuration can thus be written at all geometries in terms of the original set of active orbitals that we approximate as being strictly diabatic over the range of deformations considered. In other words, any non-adiabatic coupling is reduced to the contribution due to configuration interaction within the 17-configuration space. Let us remark that this is similar in spirit to calculations of the local non-adiabatic coupling between two states, where it is often assumed that the orbital contribution can be neglected.

#### 4.4. Valence states

We show here how to build a set of symmetry-adapted combinations of electronic configurations that (i) belong to irreducible representations of  $G_{16}$ , (ii) are orthogonal, (iii) can serve as a basis set to expand the three valence states,  $N$ ,  $V$ , and  $Z$ .

As shown by Krawczyk et al. in Ref. [18], it is possible to qualitatively understand the nature of these states in terms of simple configurations involving either localized or delocalized molecular orbitals. In fact, none of the two kinds of basis sets is adapted to  $G_{16}$ . Objects corresponding to  $E^- \otimes E^-$  must be written in terms of symmetrized combinations of (R,L)- or (G,U)-components (see Eq. (12h)).

Using the ‘‘short names’’ for the orbitals listed in Table 9, the three singlet-adapted configuration-state functions corresponding to the three valence states in  $D_{2h}$  expressed in terms of *delocalized* orbitals read:

$$N = |^1\pi^2\rangle = \pi\pi \frac{\alpha\beta - \beta\alpha}{2}, \quad (17a)$$

$$V = |^1\pi\pi^*\rangle = \frac{\pi\pi^* + \pi^*\pi}{\sqrt{2}} \frac{\alpha\beta - \beta\alpha}{2}, \quad (17b)$$

$$Z = |^1\pi^{*2}\rangle = \pi^*\pi^* \frac{\alpha\beta - \beta\alpha}{2}. \quad (17c)$$

Note that  $S_{\pi\pi^*} = \langle \pi | \pi^* \rangle$ , the overlap between the two orbitals, is zero by construction (even at less symmetrical geometries than  $D_{2h}$ ).

Using the *localized* orbitals,  $p_R$  and  $p_L$ , where

$$\pi = \frac{p_R + p_L}{\sqrt{2(1+S)}}, \quad (18a)$$

$$\pi^* = \frac{p_R - p_L}{\sqrt{2(1-S)}}, \quad (18b)$$

and  $S = S_{p_R p_L} = \langle p_R | p_L \rangle$ , leads to non-orthogonal valence-bond configurations:

$$|^1p_R^2\rangle = p_R p_R \frac{\alpha\beta - \beta\alpha}{2}, \quad (19a)$$

$$|^1p_R p_L\rangle = \frac{p_R p_L + p_L p_R}{\sqrt{2(1+S^2)}} \frac{\alpha\beta - \beta\alpha}{2}, \quad (19b)$$

$$|^1p_L^2\rangle = p_L p_L \frac{\alpha\beta - \beta\alpha}{2}. \quad (19c)$$

Spatial-symmetry-adapted configurations involving doubly-occupied orbitals are defined as

$$|^1(p_R^2 \pm p_L^2)\rangle = \frac{p_R p_R \pm p_L p_L}{\sqrt{2(1 \pm S^2)}} \frac{\alpha\beta - \beta\alpha}{2}, \quad (20a)$$

$$|^1(\pi^2 \pm \pi^{*2})\rangle = \frac{\pi\pi \pm \pi^*\pi^*}{\sqrt{2}} \frac{\alpha\beta - \beta\alpha}{2}. \quad (20b)$$

From there, one shows that there are two kinds of  $G_{16}$ -symmetry-adapted sets of configurations:

$$N' = |^1(\pi^2 - \pi^{*2})\rangle, \quad (21a)$$

$$V = |^1\pi\pi^*\rangle, \quad (21b)$$

$$Z' = |^1(\pi^2 + \pi^{*2})\rangle, \quad (21c)$$

or

$$N'' = |^1p_R p_L\rangle, \quad (22a)$$

$$V = |^1(p_R^2 - p_L^2)\rangle, \quad (22b)$$

$$Z'' = |^1(p_R^2 + p_L^2)\rangle. \quad (22c)$$

The transformations between the three sets read:

$$N = \frac{Z' + N'}{\sqrt{2}} = \frac{\sqrt{1+S^2}}{\sqrt{2}(1+S)}(Z'' + N''), \quad (23a)$$

$$Z = \frac{Z' - N'}{\sqrt{2}} = \frac{\sqrt{1+S^2}}{\sqrt{2}(1-S)}(Z'' - N''), \quad (23b)$$

$$N' = \frac{N - Z}{\sqrt{2}} = \frac{\sqrt{1+S^2}}{1-S^2}(N'' - SZ''), \quad (23c)$$

$$Z' = \frac{N + Z}{\sqrt{2}} = \frac{\sqrt{1+S^2}}{1-S^2}(Z'' - SN''), \quad (23d)$$

$$N'' = \frac{(1+S)N - (1-S)Z}{\sqrt{2(1+S^2)}} = \frac{N' + SZ'}{\sqrt{1+S^2}}, \quad (23e)$$

$$Z'' = \frac{(1+S)N + (1-S)Z}{\sqrt{2(1+S^2)}} = \frac{Z' + SN'}{\sqrt{1+S^2}}. \quad (23f)$$

If  $S$  is zero (e.g., in  $D_{2d}$ ), or is neglected as in Ref. [18], Eq. (23f) simplifies as

$$N = \frac{Z' + N'}{\sqrt{2}}, \quad (24a)$$

$$Z = \frac{Z' - N'}{\sqrt{2}}, \quad (24b)$$

$$N' = \frac{N - Z}{\sqrt{2}} = N'', \quad (24c)$$

$$Z' = \frac{N + Z}{\sqrt{2}} = Z''. \quad (24d)$$

The configuration symmetries are easily obtained from the expansion of  $E^- \otimes E^-$  in terms of (R,L)- or (G,U)-components as already mentioned (see Eq. (12h)).  $N'$  and  $N''$  belong to  $B_1^+$ ,  $V$  to  $B_2^+$ , and  $Z$  and  $Z'$  to  $A_1^+$ . In addition, the symmetry of the overlap,  $S$ , must be taken into account as can be noticed in Eq. (23f). It is easy to show that it belongs to  $B_1^+$ . Terms involving  $S^2$  do not matter as they belong to  $A_1^+$ .

The basis set  $(N', V, Z')$  is more convenient for building a quasidegenerate representation as all three configurations are orthogonal with each other (in contrast, the overlap between  $N''$  and  $Z'$  is equal to  $2S/(1+S^2)$ ). Using the same notations as Krawczyk et al. for the elements of the potential energy matrix, we get, in the  $(N', V, Z')$ -representation:

$$\mathbf{V} = V_{\text{pot}} \mathbf{1} + \begin{pmatrix} -V_A & V_Z & V_\pi \\ V_Z & -V_C & V_{\text{pyr}} \\ V_\pi & V_{\text{pyr}} & V_C + V_A \end{pmatrix}, \quad (25)$$

where  $\mathbf{1}$  is the unit matrix. The structure of this matrix is simpler than the two versions given in Ref. [18], either in the delocalized basis set,  $(\pi^2, \pi\pi^*, \pi^{*2})$ :

$$\mathbf{V}^{\text{deloc}} = V_{\text{pot}} \mathbf{1} + \begin{pmatrix} V_C/2 + V_\pi & \frac{V_{\text{pyr}} + V_Z}{\sqrt{2}} & V_A + V_C/2 \\ \frac{V_{\text{pyr}} + V_Z}{\sqrt{2}} & -V_C & \frac{V_{\text{pyr}} - V_Z}{\sqrt{2}} \\ V_A + V_C/2 & \frac{V_{\text{pyr}} - V_Z}{\sqrt{2}} & V_C/2 - V_\pi \end{pmatrix}, \quad (26)$$

or in the localized basis set,  $(p_R^2, p_R p_L, p_L^2)$ ,

$$\mathbf{V}^{\text{loc}} = V_{\text{pot}} \mathbf{1} + \begin{pmatrix} V_A/2 + V_{\text{pyr}} & \frac{V_\pi + V_Z}{\sqrt{2}} & V_A/2 + V_C \\ \frac{V_\pi + V_Z}{\sqrt{2}} & -V_A & \frac{V_\pi - V_Z}{\sqrt{2}} \\ V_A/2 + V_C & \frac{V_\pi - V_Z}{\sqrt{2}} & V_A/2 - V_{\text{pyr}} \end{pmatrix}. \quad (27)$$

thus revealing how  $G_{16}$  symmetry conditions the natural functions entering the Hamiltonian. Each entry in the matrix  $\mathbf{V}$  belongs to a well-defined irreducible representation of  $G_{16}$  as shown in Table 10.

From then, the basis functions that can be used and multiplied together to fit this matrix are to be taken from either Table 7 or Table 8.

#### 4.5. Rydberg states

We now generalize the analysis started for the three valence configurations to the 14 Rydberg configurations. As already mentioned, we make the hypothesis that overlaps between orbitals belonging to different representations in  $G_{16}$  can be neglected to simplify the derivation. We will use  $a$  for representing orbitals belonging to one-dimensional irreducible representations in  $G_{16}$  ( $a=3s, 3p_z$ , or  $3d_{z^2}$ ), ( $u, g$ ) for either  $(3p_y, 3d_{yz})$  ( $E^+$ ) or  $(3p_x, 3d_{xz})$  ( $E^-$ ), and  $(r, \ell)$  for either  $(\sigma_R^+, \sigma_L^+)$  ( $E^+$ ) or  $(p_R^-, p_L^-)$  ( $E^-$ ) (see Table 9).

The expressions of the electronic states in terms of dominant configurations have been analyzed in Section 3 within the framework of point-group theory at  $D_{2h}$  and  $D_{2d}$  geometries. Here we give a consistent representation based on dynamical symmetry, valid at all geometries.

Let us start with the Jahn–Teller pairs. The singlet-adapted configuration-state functions read:

$$R(a) = |^1\pi a\rangle = \frac{|^1p_R a\rangle + |^1p_L a\rangle}{\sqrt{2}}, \quad (28a)$$

$$R^*(a) = |^1\pi^* a\rangle = \frac{|^1p_R a\rangle - |^1p_L a\rangle}{\sqrt{2}}. \quad (28b)$$

These two configurations are degenerate at geometries where  $G_{16}$  corresponds to a point group where  $E^-$  correlates to a degenerate irreducible representation (e.g.,  $D_{2d}$ ).  $|^1\pi a\rangle$  and  $|^1\pi^* a\rangle$  transform as  $E^-(U^-) \otimes \Gamma_a$  and  $E^-(G^-) \otimes \Gamma_a$ , respectively ( $\Gamma_a=A_1^+$  or  $B_2^+$ ). Within the Jahn–Teller pair, the potential energy matrix elements are thus such that  $V_{R(a)R(a)} + V_{R^*(a)R^*(a)}$  is  $A_1^+$ ,  $V_{R(a)R(a)} - V_{R^*(a)R^*(a)}$  is  $B_1^+$ , and  $V_{R(a)R^*(a)}$  is  $B_2^+$ . The individual functions  $V_{R(a)R(a)}$  and  $V_{R^*(a)R^*(a)}$  can be expressed easily in a second stage. Matrix elements  $V_{R(a)C}$  and  $V_{R^*(a)C}$ , where  $C$  is a non-degenerate configuration form a degenerate pair of functions belonging to either  $E^+$  or  $E^-$  depending on the symmetry of  $C$ . They behave as  $E^-(U^-) \otimes \Gamma_C$  and  $E^-(G^-) \otimes \Gamma_C$ , respectively (where  $C$  belongs to the irreducible representation  $\Gamma_C$ ).

The remaining configurations involve two pairs of degenerate orbitals. The combinations symmetrized with respect to  $G_{16}$  read:

$$\frac{|^1\pi u\rangle + |^1\pi^* g\rangle}{\sqrt{2}} = \frac{|^1p_R r\rangle + |^1p_L \ell\rangle}{\sqrt{2}} \rightarrow A_1^+ \quad \text{or} \quad A_2^-, \quad (29a)$$

$$\frac{|^1\pi u\rangle - |^1\pi^* g\rangle}{\sqrt{2}} = \frac{|^1p_L r\rangle + |^1p_R \ell\rangle}{\sqrt{2}} \rightarrow B_1^+ \quad \text{or} \quad B_2^-, \quad (29b)$$

$$\frac{|^1\pi g\rangle + |^1\pi^* u\rangle}{\sqrt{2}} = \frac{|^1p_R r\rangle - |^1p_L \ell\rangle}{\sqrt{2}} \rightarrow B_2^+ \quad \text{or} \quad B_1^-, \quad (29c)$$

$$\frac{|^1\pi g\rangle - |^1\pi^* u\rangle}{\sqrt{2}} = \frac{|^1p_L r\rangle - |^1p_R \ell\rangle}{\sqrt{2}} \rightarrow A_2^+ \quad \text{or} \quad A_1^-, \quad (29d)$$

where representations with a “+” and “−” correspond to  $(\pi, \pi^*)(3p_x, 3d_{xz})$  and  $(\pi, \pi^*)(3p_y, 3d_{yz})$ , respectively. The symmetries of the corresponding matrix elements are readily obtained as the products of the irreducible representations of two configurations involved.

**Table 10**  
Potential energy functions in  $G_{16}$ .

| $G_{16}$ | Matrix element              | Function                   |
|----------|-----------------------------|----------------------------|
| $A_1^+$  | $V_{N'N'}, V_{VV}, V_{ZZ'}$ | $V_{\text{pot}}, V_C, V_A$ |
| $B_1^+$  | $V_{Z'N'}$                  | $V_\pi$                    |
| $A_2^+$  | $V_{NV}$                    | $V_Z$                      |
| $B_2^+$  | $V_{VZ'}$                   | $V_{\text{pyr}}$           |

**Table 11**  
The 17  $G_{16}$ -adapted quasidiabatic configurations.

| Configuration                  | $G_{16}$ Irreducible representation |
|--------------------------------|-------------------------------------|
| $N'$                           | $B_1^+$                             |
| $V$                            | $B_2^+$                             |
| $Z'$                           | $A_1^+$                             |
| $(R(3s), R^*(3s))$             | $E^-$                               |
| $(R(3p_z), R^*(3p_z))$         | $E^-$                               |
| $(R(3d_{z^2}), R^*(3d_{z^2}))$ | $E^-$                               |
| $R(3p_x) + R^*(3d_{xz})$       | $A_1^+$                             |
| $R(3p_x) - R^*(3d_{xz})$       | $B_1^+$                             |
| $R(3d_{xz}) + R^*(3p_x)$       | $B_2^+$                             |
| $R(3d_{xz}) - R^*(3p_x)$       | $A_2^+$                             |
| $R(3p_y) + R^*(3d_{yz})$       | $A_2^-$                             |
| $R(3p_y) - R^*(3d_{yz})$       | $B_2^-$                             |
| $R(3d_{yz}) + R^*(3p_y)$       | $B_1^-$                             |
| $R(3d_{yz}) - R^*(3p_y)$       | $A_1^-$                             |

#### 4.6. Summary

The final list of 17  $G_{16}$ -adapted quasidiabatic configurations is recalled in Table 11. As already pointed out in Section 3, the active orbitals and/or the dominant configurations may mix depending on the the local point groups. The most significant cases are: (i) the adiabatic  $N$  and  $Z$  states in  $D_{2h}$ , which are balanced mixtures of  $N'$  and  $Z'$  (see Eq. (23f)); (ii) the adiabatic  $V$  and  $R(3d_{xz})$  states in  $D_{2h}$ , which effectively mix the actual  $V$  and  $R(3d_{xz})$  with almost equal weights (see Table 3); (iii) the balanced mixtures (due to hyperconjugation) of the adiabatic  $E \otimes E$  states in  $D_{2d}$ :  $R(3p_x) + R^*(3d_{xz})$  with  $R(3d_{yz}) - R^*(3p_y)(A_1)$ ,  $R(3p_x) - R^*(3d_{xz})$  with  $R(3d_{yz}) + R^*(3p_y)(B_1)$ ,  $R(3d_{xz}) + R^*(3p_x)$  with  $R(3p_y) - R^*(3d_{yz})(B_2)$ , and  $R(3d_{xz}) - R^*(3p_x)$  with  $R(3p_y) + R^*(3d_{yz})(A_2)$  (see Table 4). Other mixtures are far less significant, so that this this defines a convenient starting point for further diabaticization of the adiabatic states.

## 5. Conclusions

Quantum dynamics methods such as MCTDH are now capable of treating molecular systems of about 10 atoms (the multilayer formulation [45], ML-MCTDH, is expected to make even larger systems tractable soon). Although quantum dynamics can be applied in the form of approximate direct dynamics treatments (the potential energies and derivatives are calculated on the fly; see, e.g., Ref. [46]), more accurate calculations still require grid-based representations of the potential energy surfaces. These must be given in the most compact form. Photochemical applications involve large-amplitude deformations and several coupled electronic states. There is thus a real need for generalizing the current vibronic Hamiltonian models now limited to ultrafast processes involving small deformations. This implies using curvilinear coordinates for the nuclear motion and determining the electronic configurations that dominate in the regions of the reactant, the product(s), as well as all intermediate species. This point of view is supported by the idea that complicated shapes in the resulting adiabatic potential energy surfaces often result from coupling between simpler quasidiabatic surfaces through configuration interaction.

In this work, we proposed an approach based on nuclear-permutation-inversion groups to analyze the underlying structure of the electronic states and nuclear coordinates. Dynamical symmetry properties are always satisfied, irrespectively of the current point group for a given geometry. Our strategy directly provides the most simple basis functions of the nuclear coordinates on which a natural quasidiabatic representation can be built for all explored geometries in a flexible molecule.

The natural nuclear coordinates for treating the large-amplitude deformations that occur during a chemical reaction are curvilinear.

For ethylene photochemistry, we showed how to start from poly-spherical coordinates and symmetrize them with respect to the dynamical group  $G_{16}$ . In terms of electronic configurations, we started the analysis by determining a minimal set of valence-bond-like fragment orbitals that are necessary to describe the valence states of ethylene as well as the lowest Rydberg states embedded in these (note that the valence/Rydberg distinction can be disputed: the  $V$  state is partly Rydberg  $\pi^*$  and the lowest Rydberg states are partly valence  $\sigma^*$ ). Using dynamical symmetry helped in the determination of the self-contained set of nine active orbitals that is stable along all considered deformations of the geometry, except for some minor and smooth mixtures with the space of inactive orbitals. This led to the definition of a stable set of 17 configurations that was used to run SA-17-SASSCF(2,9) calculations, i.e. CAS(2,9) calculations restricted to these 17 selected configurations and using orbitals state-averaged over 17 states. This strategy gave a remarkably good description of the  $V$  state, although this remains today a computational challenge. We also got a qualitatively correct description of the lowest Rydberg states. This stable space of 17 configurations proved to be the most natural multireference for subsequent MRCI calculations. Further tests are to be made with better-suited atomic orbital basis sets in order to define a level of calculation capable of reaching high accuracy for all geometries and states.

This work is the first stage in building an accurate and global analytical fit of the  $N$  and  $V$  potential energy surfaces of ethylene and those of all electronic states that couple to them in this energy range. This in turn will be used to run quantum dynamics calculations and (i) assign the  $\pi$ -to- $\pi^*$  absorption spectrum, (ii) decipher the radiationless decay mechanism.

## Acknowledgements

F. Gatti thanks Prof. Horst Köppel for many important discussions about diabaticization and the vibronic coupling model. This work was supported by projet blanc ANR-09-BLAN-0417 of the ANR of the French Centre National de la Recherche Scientifique (CNRS). Financial support by the Deutsche Forschungsgemeinschaft is also gratefully acknowledged.

## References

- [1] W. Domcke, D.R. Yarkony, H. Köppel (Eds.), *Conical Intersections*, World Scientific, New Jersey, 2004.
- [2] C.A. Mead, D.G. Truhlar, *J. Chem. Phys.* 77 (1982) 6090.
- [3] T. Pacher, L.S. Cederbaum, H. Köppel, *Adv. Chem. Phys.* 84 (1993) 293.
- [4] D.R. Yarkony, *J. Chem. Phys.* 112 (2000) 2111.
- [5] M.J. Bearpark, M. Boggio-Pasqua, M.A. Robb, F. Ogliaro, *Theor. Chem. Acc.* 116 (2006) 670.
- [6] B.N. Papas, M.S. Schuurman, D.R. Yarkony, *J. Chem. Phys.* 129 (2008) 124104.
- [7] H. Köppel, W. Domcke, L.S. Cederbaum, *Adv. Chem. Phys.* 57 (1984) 59.
- [8] H.-D. Meyer, F. Gatti, G.A. Worth (Eds.), *Multidimensional Quantum Dynamics: MCTDH Theory and Applications*, Wiley-VCH, Weinheim, 2009.
- [9] G.A. Worth, H.-D. Meyer, H. Köppel, L.S. Cederbaum, *Int. Rev. Phys. Chem.* 27 (2008) 569.
- [10] F. Gatti, *C. Iung, Phys. Rep.* 484 (2009) 1.
- [11] S.P. Keating, C.A. Mead, *J. Chem. Phys.* 82 (1985) 5102.
- [12] X. Zhu, D.R. Yarkony, *J. Chem. Phys.* 132 (2010) 104101.
- [13] H.C. Longuet-Higgins, *Mol. Phys.* 6 (1963) 445.
- [14] P.R. Bunker, P. Jensen, *Molecular Symmetry and Spectroscopy*, NRC Research Press, Ottawa, 1998.
- [15] O. Godsi, C.R. Evenhuis, M.A. Collins, *J. Chem. Phys.* 125 (2006) 104105.
- [16] B.J. Braams, J.M. Bowman, *Int. Rev. Phys. Chem.* 28 (2009) 577.
- [17] C. Petrongolo, R.J. Buenker, S.D. Peyerimhoff, *J. Chem. Phys.* 78 (1983) 7284.
- [18] R.P. Krawczyk, A. Viel, U. Manthe, W. Domcke, *J. Chem. Phys.* 119 (2003) 1397.
- [19] A.M. Mebel, Y.-T. Chen, S.-H. Lin, *J. Chem. Phys.* 105 (1996) 9007.
- [20] J.M. Mestdagh, J.P. Visticot, M. Elhanine, B. Soep, *J. Chem. Phys.* 113 (2000) 237.
- [21] M. Barbatti, J. Paier, H. Lischka, *J. Chem. Phys.* 121 (2004) 11614.
- [22] K. Kosma, S.A. Trushin, W. Fuss, W.E. Schmid, *J. Phys. Chem. A* 112 (2008) 7514.
- [23] C. Angeli, *J. Comp. Chem.* 30 (2009) 1319.
- [24] A. Viel, R.P. Krawczyk, U. Manthe, W. Domcke, *J. Chem. Phys.* 120 (2004) 11000.
- [25] M.R. Brill, F. Gatti, D. Lauvergnat, H.-D. Meyer, *Chem. Phys.* 338 (2007) 186.



- [26] J.J. Lin, C.C. Wang, Y.T. Lee, X. Yang, *J. Chem. Phys.* 113 (2000) 9668.
- [27] A. Toniolo, M. Ben-Nun, T.J. Martínez, *J. Phys. Chem. A* 106 (2002) 4679.
- [28] G. Granucci, M. Persico, A. Toniolo, *J. Chem. Phys.* 114 (2001) 10608.
- [29] A.H.H. Chang, A.M. Mebel, X.-M. Yang, S.H. Lin, Y.T. Lee, *J. Chem. Phys.* 109 (1998) 2748.
- [30] J. O'Reilly, S. Douin, S. Boyé, N. Shafizadeh, D. Gauyacq, *J. Chem. Phys.* 119 (2003) 820.
- [31] A.J. Merer, J.K.G. Watson, *J. Mol. Spectrosc.* 47 (1973) 499.
- [32] C. Petrongolo, R.J. Buenker, S.D. Peyerimhoff, *J. Chem. Phys.* 76 (1982) 3655.
- [33] R.S. Mulliken, *J. Chem. Phys.* 23 (1955) 1997.
- [34] A.J. Merer, R.S. Mulliken, *Chem. Rev.* 69 (1969) 639.
- [35] 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, 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, P.V. Parandekar, N.J. Mayhall, A.D. Daniels, O. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian Development Version, Revision H.01, Gaussian, Inc., Wallingford, CT, 2009.
- [36] M. Ben-Nun, T.J. Martínez, *Chem. Phys.* 259 (2000) 237.
- [37] G.H. Herzberg, *Electronic Spectra of Polyatomic Molecules*, Van Nostrand Reinhold, Toronto, 1966.
- [38] V.M. Garcia, R. Caballol, J.P. Malrieu, *Chem. Phys. Lett.* 261 (1996) 98.
- [39] R.J. Sension, B.S. Hudson, *J. Chem. Phys.* 90 (1989) 1377.
- [40] B.A. Williams, T.A. Cool, *J. Chem. Phys.* 94 (1991) 6358.
- [41] E.R. Davidson, A.A. Jarzecki, *Chem. Phys. Lett.* 285 (1998) 155.
- [42] R. Lindh, B.O. Roos, *Int. J. Quantum. Chem.* 35 (1989) 813.
- [43] R. McDiarmid, *J. Phys. Chem.* 84 (1980) 64.
- [44] E.H. van Veen, *Chem. Phys. Lett.* 41 (1976) 540.
- [45] H. Wang, M. Thoss, *J. Chem. Phys.* 119 (2003) 1289.
- [46] G.A. Worth, M.A. Robb, B. Lasorne, *Mol. Phys.* 106 (2008) 2077.