Valence electron momentum distributions of ethylene; comparison of EMS measurements with near Hartree–Fock limit, configuration interaction and density functional theory calculations

B.P. Hollebone a, J.J. Neville a, Y. Zheng a, C.E. Brion a, Y. Wang b, E.R. Davidson b

a Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, B.C., Canada V6T 1Z1
b Department of Chemistry, Indiana University, Bloomington, IN 47405, USA

Received 24 January 1995

Abstract

The valence binding energy spectra and orbital electron momentum profiles of C2H4 have been measured by energy dispersive, multichannel, symmetric non-coplanar electron momentum spectroscopy at an impact energy of 1200 eV + binding energy. The effects of final state correlation on the valence binding energy spectrum are investigated using multi-reference singles and doubles configuration interaction calculations and the results are compared with calculations from the literature and with experiment. The measured momentum profiles are compared with cross sections calculated using a range of wavefunctions from minimal basis to essentially Hartree–Fock limit in quality. In addition, the effects of correlation and relaxation on the calculated momentum profiles are investigated using multi-reference singles and doubles configuration interaction calculations of the full ion–neutral overlap distributions. Several different configuration interaction results are investigated in the case of the 1b3g HOMO momentum profile. Electron correlation effects in the ground state are also examined using several density functional approaches for all the momentum profiles. The experimental momentum resolution is accounted for in all of the calculated momentum profiles by using the Gaussian-weighted planar grid procedure. The effects of vibrational averaging are investigated in the case of the 1b3g orbital of ethylene.

1. Introduction

The electronic structure of unsaturated carbon–carbon bonds poses a challenging problem for quantum chemistry. As the simplest doubly bonded hydrocarbon, ethylene (ethene, C2H4) serves as an excellent prototype of this functional group, amenable to both ab initio theoretical and gas-phase experimental study. In addition to its fundamental interest as a quantum chemical benchmark system, ethylene is also an important compound as a feedstock for applications as diverse as drug synthesis and polymer production. The carbon–carbon double bond is also a highly reactive site for addition reactions in aliphatic alkane systems. Clearly, an accurate theoretical model of the electronic structure of double bonds is highly desirable in order to interpret and predict the chemical and physical behavior of C2H4.
and related alkene systems. However for the development of a ‘universal wavefunction’ suitable for a wide range of purposes, the theoretical model systems must be evaluated by comparison of calculated properties with a carefully selected set of measurements, such as the total energy, quadrupole moment and momentum profile which emphasize small $r$, medium $r$ and large $r$ regions of space respectively [1–3].

Electron momentum spectroscopy (EMS) has become a powerful technique in experimental quantum chemistry [1–6] for orbital imaging and for the detailed study of electronic structure and electron motion in molecules. In particular consideration of the experimental momentum profiles (XMPs) of valence electrons, selected according to their binding energies, has been found to be extremely useful in theoretical chemistry for the evaluation and design of Hartree–Fock basis sets and correlated molecular wavefunctions [1–3,5,7]. Since the EMS measurements are most sensitive to the chemically important larger spatial (i.e. low momentum) parts of the wavefunction they provide an effective counterbalancing consideration to the necessary, but usually insufficient, variational minimization of the total energy commonly used in wavefunction evaluation and design. This is because, in practice, the variational theorem (at a given level of energy convergence) is often an insufficient criterion for basis set or method evaluation because the total energy is dominated by the short range (i.e. high momentum) part of the molecular potential. Consequently, the consideration of both the total energy and of the accurate reproduction of experimental momentum profiles (selected by binding energy) has been found to provide a much more effective means of testing and developing ‘universal’ molecular wavefunctions, suitable for the highly accurate calculation of a wide range of electronic properties which emphasize different regions of space [1,2,7]. For example, while it has long been realized [8] that variationally determined (i.e. energy optimized) wavefunctions do not necessarily lead to accurately calculated dipole moments (for polar molecules), this situation is dramatically improved if the additional consideration of momentum profile is applied, as can be seen for example in the studies of the water molecule [1,2,7]. With the enhanced capabilities and improvements in sensitivity provided by recently reported multichannel electron momentum spectrometers [9–12] much more precise information can be expected.

The first reported studies of ethylene by EMS were the early work of Dixon et al. [13] and of Coplan et al. [14]. Subsequent work has involved studies of the lowest binding energy momentum profiles of substituted carbon–carbon double bonds, including measurements of the $1b_{3u}$ of ethylene [15,16], but no further studies of the complete valence shell have been reported. In the earlier valence shell measurements of ethylene [13,14], the shapes of the observed momentum profiles agreed qualitatively with the predictions of modest levels of theory, however no attempts were made to account for the effects of instrumental momentum resolution in the calculations. A further important limitation in both of these studies [13,14] was that the individual orbital momentum profiles were each separately height normalized to theory. Subsequently many studies of other systems have clearly demonstrated that collecting all of the experimental momentum profiles on a common intensity scale, with a single-point normalization to one calculation provides a much more stringent and informative quantitative comparison of theory and experiment. In addition, the XMPs reported by Dixon et al. [13] were obtained only for momenta above ~ 0.3 au, which limited the usefulness of this study for the evaluation of theoretical distributions in the chemically important, low momentum regions (below 0.5 au). Independently height normalized momentum profiles [13,14], the comparatively poor statistics obtained in both studies and the limited momentum range used by Dixon et al. [13], limit the usefulness of these early measurements for the evaluation and development of theoretical models.

Early X-ray photoelectron studies [17,18] have shown the existence of strong satellite structure in the inner valence region of the binding energy spectrum of $\text{C}_2\text{H}_4$. The main satellite final ion states were initially assigned to be $(2a_g)^{-1}$ at 23.7 eV and $(2p_{1u})^{-1}$ at 27.2 eV. However, the 2ph Tamm–Dancoff Approximation (TDA) Greens' function calculation of Cederbaum et al. [19] and the CI results of Martin and Davidson [20] predicted that the entire inner valence region was dominated by $(2a_g)^{-1}$ poles, with very little $(2b_{1u})^{-1}$ intensity. The earlier
EMS studies [13,14] assigned the two main satellite features as $(2\alpha_g)^{-1}$ processes by measurement of the momentum profiles at binding energies of 23.7 and 27.2 eV, thus supporting the general ab initio predictions. However, no definitive assignment of the other wide-ranging satellite intensity was made [13,14]. Martin and Davidson [20] first assigned the main satellite as a $(2\alpha_g)^{-1}$ feature based on a limited CI calculation. The Green’s function [19] calculation predicted a ‘twinning’ of the $(2\alpha_g)^{-1}$ poles at the energy of the most intense experimentally observed peak at 23.7 eV in the inner valence region. The earlier work of Martin and Davidson [20] had omitted the class of configurations responsible for this twinning. However, earlier measurements of this region of the binding energy spectra by both EMS [13,14] and photoelectron spectroscopy [18,21,22] as well as a more recent high resolution synchrotron radiation PES measurement [23] have not revealed any splitting of this satellite peak. Other more recent many-body calculations [24–28] report a single pole for this feature, but it should be noted that they do not quantitatively reproduce the intensities or positions of the experimentally observed inner valence manifold.

The above history clearly suggests that more detailed experimental and theoretical investigations of $\text{C}_2\text{H}_4$ are desirable. The present work reports multichannel energy dispersive electron momentum spectroscopy measurements of the momentum profiles and binding energy spectra for ethylene over a greater binding energy range and with better energy resolution and better statistical precision than in the previous studies. Careful attention has been given to obtaining measurements of the energetically close $1b_3g$, $3\alpha_g$ and $1b_{2u}$ momentum profiles. The EMS binding energy spectra are compared with the results of a new multi-reference singles and doubles configuration interaction (MRSD-CI) calculation as well as with previously published Greens’ function [19] and symmetry-adapted cluster CI calculations [26]. The measured momentum profiles are compared with a range of SCF treatments ranging from a minimal STO-3G basis to near Hartree-Fock limit basis sets including a 196 contracted Gaussian type orbital (196-GTO) calculation developed in the course of the present collaborative work. The effects of correlation and relaxation on all of the valence momentum profiles are also investigated with a new MRSD-CI overlap calculation (196-G(CI)) using the 196-GTO basis. For the lowest binding energy ($1b_{3u}$) momentum profile, the effects of even further increasing the size of both the basis set and the CI configuration space are studied. Many-electron effects on the ethylene profiles are also studied by density functional methods, using the recently developed target Kohn–Sham approximation of the EMS cross section [29,30]. Finally, in order to assess the effects of nuclear motion, vibrational averaging is investigated in the $1b_{3u}$ HOMO momentum density at the 196-GTO SCF level.

2. Method

The energy-dispersive multichannel symmetric non-coplanar EMS spectrometer has been described in detail recently [11] and thus only a brief description will be given here. The gas-phase target molecules are ionized by impact with a high energy electron beam ($E_0 = 1200$ eV + binding energy). The outgoing electrons (scattered and ionized) are angle and energy selected and are then detected in coincidence. The experimental geometry is referred to as symmetric non-coplanar, that is the two outgoing electrons are selected to have equal polar angles ($\theta_1 = \theta_2 = 45^\circ$) relative to the forward scattered electron beam. Each electron energy analyser accepts a range of kinetic energies from 596 to 604 eV simultaneously, but only those coincident electron pairs with summed energies in the range of $1200 \pm 3.5$ eV are recorded (see Ref. [11] for details). The relative azimuthal (out of plane) angle $\phi$ between the two outgoing electrons is variable over the range 0° to ±30°. Under the binary encounter [31] requirements of high impact energy, high momentum transfer and negligible kinetic energy transfer to the resultant ion, the initial momentum $p$ of the knocked-out (ionized) electron can be shown to be related to the azimuthal angle by:

$$p = \left\{(2p_1 \cos \theta_1 - p_0)^2 + \left[2p_1 \sin \theta_1 \sin\left(\frac{\phi}{2}\right)\right]^2\right\}^{1/2},$$

where $p_1 = p_2 = \sqrt{2E_1}$ is the magnitude of the momentum of each of the two outgoing electrons and
\[ p_0 = \sqrt{2E_0} \] is the momentum of the incident electron (in atomic units).

The EMS binary \((e, 2e)\) differential cross section in the plane wave impulse approximation (PWIA) for randomly oriented gas-phase molecules \([6]\) is given by

\[
\sigma_{\text{EMS}} \propto \int \frac{d\Omega}{4\pi} |\langle \mathbf{p} \Psi_{N-1} \mid \Psi_N \rangle|^2, \tag{2}
\]

where \(\mathbf{p}\) is the momentum of the target electron state prior to knockout and \(\Psi_{N-1}\) and \(\Psi_N\) are the final ion and molecular target wavefunctions respectively. This generalized formulation of the EMS cross section is useful when using many-body calculations for the initial and final wavefunctions, such as configuration interaction treatments.

Eq. (2) is greatly simplified by the assumption that the ground state is a single Slater determinant. In this target Hartree-Fock approximation (THFA), the initial state wavefunction is typically approximated by an SCF wavefunction. If the ion state, expressed with the same molecular orbital basis set as the neutral, is dominated by a single hole in only one orbital, Eq. (2) can be simplified to:

\[
\sigma_{\text{EMS}} \propto S_i \int \frac{d\Omega}{4\pi} |\psi_i(\mathbf{p})|^2, \tag{3}
\]

where \(\psi_i(\mathbf{p})\) is the momentum representation of the more familiar position space molecular orbital wavefunction, \(\psi_i(\mathbf{r})\), which corresponds to the orbital from which the electron was ionized. The spherically averaged square of the \(p\)-space molecular orbital (Eq. (3)) is referred to as a momentum distribution. The quantity \(S_i\) is called the spectroscopic factor and is the probability of the ionization event producing a \((\psi_i)^{-1}\) one-hole configuration of the final ion state, \(\Psi_{N-1}\).

Recently, Eq. (2) has been re-interpreted in the context of Kohn–Sham density functional theory \([29,30]\). In a manner very similar to that described for the target Hartree–Fock approximation above (Eq. (3)), the ion–neutral overlap term in Eq. (2) can be factored in terms of the target Kohn–Sham orbitals \(\psi_{\text{KS}}\) to give the target Kohn–Sham approximation (TKSA) \([29,30]\):

\[
\sigma_{\text{EMS}} \propto S_i \int \frac{d\Omega}{4\pi} |\psi_{\text{KS}}(\mathbf{p})|^2. \tag{4}
\]

It should be noted that the TKSA formulation of the EMS cross section provides for the inclusion of correlation and exchange in the target ground state only through the exchange–correlation potential. The results of such TKSA density functional theory calculations have recently been compared with near Hartree–Fock limit and MRSD-CI overlap calculations and EMS measurements for the valence orbital momentum profiles of acetone \([11]\).

In EMS the individual orbitals are selected by the choice of the binding (or ionization) energy. With the multichannel energy dispersive spectrometer used in the present work, binding energy spectra (BES) are collected at a series of azimuthal angles \(\phi\). Distributions as a function of angle are obtained by deconvolution of these binding energy spectra using Gaussian functions located at each ionization energy in the BES. The widths of the Gaussian functions can be determined from a consideration of published PES vibronic manifolds and the instrumental energy resolution function (1.6 eV fwhm). For each ionization process, the area of the fitted peak (or the integral of a spectral region, where appropriate) is plotted as a function of momentum (calculated from \(\phi\) using Eq. (1)). The set of areas as a function of momentum for a specific binding energy is referred to as an experimental momentum profile or XMP. To compare the XMPs with the relative cross sections calculated as a function of momentum using expressions (2)–(4) above, the effects of the finite spectrometer acceptance angles in both \(\theta\) and \(\phi\) (\(\Delta\theta \approx \pm 0.6^\circ\) and \(\Delta\phi \approx \pm 1.2^\circ\)) must be included. This is achieved in the present work using the Gaussian-weighted planar grid method of Duffy et al. \([32]\). The theoretical cross sections, including the contributions from the instrumental angular resolution are called theoretical momentum profiles (TMPs).

The ethylene sample used in the present work was taken from a cylinder (Matheson Gas Company) of spectroscopic grade, with stated impurities of less than 1%. The gaseous sample was used directly, without further purification. No impurities were observed in any of the spectra.

3. Calculations

Spherically averaged theoretical momentum profiles have been calculated for \(\text{C}_2\text{H}_4\) from several
basis sets of varying quality using the plane wave impulse and the target Hartree–Fock approximations described in Section 2, Eq. (3). In addition, full ion–neutral overlap distributions have been calculated in the plane wave impulse approximation of the EMS cross section (Eq. (2)) using several multi-reference singles and doubles excitations configuration interaction (MRSD-CI) wavefunctions for both the ground state and molecular ion states to investigate the effects of electron correlation and relaxation on the wavefunctions. The MRSD-CI calculations were based on the respective near Hartree–Fock limit basis sets. Target Kohn–Sham approximation calculations have also been performed for a variety of correlation and relaxation potentials using a variation of the basis sets recently developed by Dunning and co-workers [33,34]. As described previously, the instrumental angular (momentum) resolution was included in the calculations using the Gaussian-weighted planar grid method [32].

Due to the many possible choices of symmetry axes in $D_{2h}$, the orbital notation found in the literature is often inconsistent and can be confusing. For clarity, the so-called Mulliken co-ordinate system has been used in the present study as described by Herzberg [35]. In this set of co-ordinates, the $z$ axis is oriented along the C=C double bond. The $y$ axis lies in the molecular plane, bisecting the double bond; the $x$ axis is then perpendicular to the plane of the molecule. The molecular geometry used for all of the calculations in this work is that described by McMurchie and Davidson [36].

Details of the calculation methods are described below. Selected properties for these various calculations are given in Table 1. Some additional calculations are discussed in Sections 4.2 and 4.3.

(1) STO-3G: This is the standard minimal basis set designed by Pople and co-workers [37]. For ethylene this is a 14-GTO basis.

(2) 6-311G: This triple valence basis developed by Krishnan et al. [38] is a 38-GTO basis for $C_2H_4$.

(3) 6-311 + + G * * : Basis set 2 has been further augmented with the standard additional diffuse [39] (sp on C, s on H) and polarization functions [40] (d on C, p on H) to increase the basis set flexibility. The final contracted basis set is 74-GTO.

(4) aug-cc-pVTZ: Developed by T.H. Dunning and his co-workers [33,34], this large basis set description adds both polarization and diffuse functions to an already triple-$\zeta$ description of the ethylene

Table 1

<table>
<thead>
<tr>
<th>Wavefunction</th>
<th>Ref.</th>
<th>Contracted basis set [C/H]</th>
<th>Total energy (hartree)</th>
<th>$\Theta_{zz}$ a (au)</th>
<th>$\rho_{MAX}$ b (au)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>STO-3G</td>
<td>[37]</td>
<td>[2s, 1p/1s]</td>
<td>-77.072233</td>
<td>0.484</td>
</tr>
<tr>
<td>2</td>
<td>6-311G</td>
<td>[38]</td>
<td>[4s, 3p/3s]</td>
<td>-78.018534</td>
<td>1.320</td>
</tr>
<tr>
<td>3</td>
<td>6-311 + + G * *</td>
<td>[39,40]</td>
<td>[5s, 4p, 1d/4s, 1p]</td>
<td>-78.055448</td>
<td>1.506</td>
</tr>
<tr>
<td>4</td>
<td>aug-cc-pVTZ</td>
<td>[33,34]</td>
<td>[6s, 7p, 3d, 1f/6s, 3p, 1d]</td>
<td>-78.064257</td>
<td>1.530</td>
</tr>
<tr>
<td>5</td>
<td>196-GTO</td>
<td>c</td>
<td>[6s, 7p, 3d, 1f/5s, 2p, 1d]</td>
<td>-78.069343</td>
<td>1.532</td>
</tr>
<tr>
<td>5a</td>
<td>196-G(CI) c</td>
<td>K orbitals</td>
<td>-78.408495</td>
<td>1.327</td>
<td>0.47</td>
</tr>
<tr>
<td>5c</td>
<td>196-G(CI) c</td>
<td>natural orbitals (NO)</td>
<td>-78.408870</td>
<td>1.241</td>
<td>0.47</td>
</tr>
<tr>
<td>6</td>
<td>228-GTO</td>
<td>c</td>
<td>[14s, 12p, 3d, 1f/7s, 3p, 1d]</td>
<td>-78.096964</td>
<td>1.520</td>
</tr>
<tr>
<td>6c</td>
<td>228-G(CI) c</td>
<td>K orbitals</td>
<td>-78.402716</td>
<td>1.312</td>
<td>0.47</td>
</tr>
<tr>
<td>7</td>
<td>234-GTO</td>
<td>c</td>
<td>[15s, 12p, 3d, 1f/8s, 3p, 1d]</td>
<td>-78.069578</td>
<td>1.520</td>
</tr>
<tr>
<td>7a</td>
<td>234-G(CI)a</td>
<td>K orbitals</td>
<td>-78.402701</td>
<td>1.312</td>
<td>0.47</td>
</tr>
<tr>
<td>7b</td>
<td>234-G(CI)b</td>
<td>neutral natural orbitals</td>
<td>-78.410454</td>
<td>1.303</td>
<td>0.47</td>
</tr>
<tr>
<td>7c</td>
<td>234-G(CI)c</td>
<td>neutral NO, low threshold cut-off</td>
<td>-78.426308</td>
<td>1.258</td>
<td>0.47</td>
</tr>
<tr>
<td>7d</td>
<td>234-G(CI)d</td>
<td>1b$_{3u}$ cation NOs</td>
<td>-78.424229</td>
<td>1.320</td>
<td>0.47</td>
</tr>
</tbody>
</table>

* Quadrupole moment, $\Theta_{zz} = \frac{1}{2}(3z^2 - r^2)$, the component parallel to the C–C bond. These values have been calculated at the static equilibrium geometry for a non-rotating, non-vibrating molecule.

b Momentum value for maximum in the 1b$_{3u}$ momentum profile. From Fig. 4.

c These SCF and CI calculations were developed in the course of the present work.

d Recommended value, uncorrected for zero-point motion, obtained by induced birefringence [66].
valence orbitals. The Hartree–Fock calculations were performed with a spherical harmonic basis set to avoid linear dependence problems in the d and f functions. This is a 184-GTO basis for C₂H₄.

(5) 196-GTO: This near Hartree–Fock limit [6s, 7p, 3d, 1f/5s, 2p, 1d] highly diffuse 196-GTO basis set was developed at Indiana University in the course of the present collaborative work. For the carbon centres, the primitive basis was chosen to be the (18s, 13p) Partridge basis set [41]. Partridge’s first fourteen s functions were contracted into two s functions using the 1s and 2s atomic orbital coefficients from an SCF calculation of the (atomic carbon) \(^3\)P state. Similarly, the first seven p functions were contracted into one p function using the 2p atomic orbital coefficients. Partridge’s (10s) basis was also used for hydrogen [42]. The first six s functions were contracted into a single s basis function using the 1s atomic orbital coefficients for the free atom. The rest of the functions were left uncontracted. This contraction scheme lost less than 0.1 kcal mol\(^{-1}\) in a trial SCF calculation on CH₄ [43].

For better reproduction of the ‘large-r’ properties, Dunning’s polarization functions [33] were used to supplement the Partridge contracted basis sets. For carbon, four additional (3d, 1f) polarization functions (\(\alpha_d = 1.848, 0.649, 0.228; \alpha_f = 0.761\)) were added. For hydrogen three additional polarization functions were added: (2p, 1d) \(\alpha_p = 1.257, 0.355; \alpha_d = 0.916\). This basis was further augmented by two p and two d diffuse Rydberg functions at the center of the C=C bond with exponents of 0.052 and 0.104 for both the p and d functions. All the Cartesian components of the d and f functions were kept.

(5a) 196-G(CI): Frozen-core MRSD configuration interaction calculations for both the parent neutral molecule and the radical cation were performed with the full virtual space. For the neutral molecule, the important configurations of a Hartree–Fock singles and doubles configuration interaction (HFSD-CI) calculation were selected using the second-order Rayleigh–Schrödinger perturbation theory (RSPT) with the Epstein–Nesbet partitioning. For the cation, a Koopmans’ theorem singles and doubles configuration interaction (KTSD-CI) was performed, in which the leading configuration was constructed from the neutral molecular orbitals through Koopmans’ approximation. Only configurations having a RSPT contribution to the energy exceeding \(0.25 \times 10^{-6}\) hartree were kept for the subsequent variational calculation. Then a full multi-reference singles and doubles configuration interaction (MRSD-CI) calculation was performed with a reference space cut-off of 0.030 for the coefficient contribution and a cut-off of 0.020 hartree for the energy contribution. For the neutral molecule, 21 reference configurations were included in the final MRSD-CI calculation which generated 23706 spin-adapted configurations and recovered about 80% of the empirical valence correlation energy.

(5c) 196-G(CI): Average natural orbitals were used for a further set of MRSD-CI calculations since it was found that the use of the ground state K orbitals did not result in good reproduction of the experimental ionization potentials. In particular, for poles of A\(_g\) symmetry the second root of the CI matrix did not necessarily correspond to ionization of the 2a\(_g\) molecular orbital when using the canonical Hartree–Fock orbitals. Transformation to average natural orbitals ensures that each root receives about the same correlation contribution. Separate average natural orbitals were obtained from the average density matrix for the first fifteen roots of each symmetry of the cation. The Dyson orbitals, used to calculate the final theoretical momentum profiles, were derived from the overlap between the correlated ground state of the neutral molecule and each root of the cation for each symmetry. The ground state of the neutral molecule was calculated separately with ANOs for each symmetry. All of these ground state energies were in the range \(-78.4084\) to \(-78.4080\) hartree. The calculated pole energies and intensities (\(S_i\)) are reported in Table 2 and are shown graphically in Fig. 2 (see Section 4.1).

Further CI results for the 1P₃u momentum profile are described in Section 4.2 below. A discussion of the Kohn–Sham DFT methods and results is presented in Section 4.3.

4. Results and discussion

Ethylene belongs to the D\(_{2h}\) point group and has a \(^1\)A\(_g\) electronic ground state. In the Mulliken co-
ordinate system the Hartree–Fock ground state configuration is written as

\[ 1a_g^2 1b_{1u}^2 2a_g^2 2b_{1u}^2 1b_{2u}^2 3a_g^2 1b_{3g}^2 1b_{2u}^2 \]

core valence

The valence shell ionization potentials reported in the He(II) PES study of Bieri et al. [22] and the XPS results of Banna and Shirley [18] are presented in Table 2. The earlier EMS [13,14] and XPS studies [17,18] have shown that the inner valence region of

<table>
<thead>
<tr>
<th>Orbital</th>
<th>Experimental [22]</th>
<th>[18]</th>
<th>this work</th>
<th>MRSD-CI</th>
<th>SAC-CI [26]</th>
<th>2ph-TDA [19]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1b_{3u}</td>
<td>10.68</td>
<td>10.51</td>
<td>(10.51)</td>
<td>10.50</td>
<td>[0.84] B_{3u} 10.83 [0.95] B_{3u} 9.76 [0.91] c B_{3u}</td>
<td></td>
</tr>
<tr>
<td>1b_{3g}</td>
<td>12.8</td>
<td>12.85</td>
<td>12.8</td>
<td>12.90</td>
<td>[0.81] B_{3g} 13.26 [0.93] B_{3g} 12.16 [0.90] B_{3g}</td>
<td></td>
</tr>
<tr>
<td>3a_{g}</td>
<td>14.8</td>
<td>14.66</td>
<td>14.7</td>
<td>14.70</td>
<td>[0.83] A_g  15.07 [0.92] A_g  13.79 [0.88] A_g</td>
<td></td>
</tr>
<tr>
<td>1b_{2u}</td>
<td>16.0</td>
<td>15.87</td>
<td>15.9</td>
<td>15.90</td>
<td>[0.73] B_{2u} 16.48 [0.87] B_{2u} 15.06 [0.81] B_{2u}</td>
<td></td>
</tr>
<tr>
<td>2b_{1u}</td>
<td>19.1</td>
<td>19.23</td>
<td>19.1</td>
<td>19.20</td>
<td>[0.62] B_{1u} 19.93 [0.81] B_{1u} 18.34 [0.76] B_{1u}</td>
<td></td>
</tr>
<tr>
<td>23.5</td>
<td></td>
<td></td>
<td></td>
<td>21.90</td>
<td>d B_{2u} 21.90 [0.02] B_{2u}</td>
<td></td>
</tr>
<tr>
<td>24.64</td>
<td>[0.30] A_g</td>
<td></td>
<td></td>
<td>24.64</td>
<td>[0.02] A_g  24.71 [0.64] A_g  25.63 [0.08] B_{1u}</td>
<td></td>
</tr>
<tr>
<td>25.14</td>
<td>[0.02] A_g</td>
<td></td>
<td></td>
<td>25.14</td>
<td>[0.02] A_g  25.14 [0.06] B_{2u}</td>
<td></td>
</tr>
<tr>
<td>26.20</td>
<td>d B_{2u}</td>
<td></td>
<td></td>
<td>26.20</td>
<td>d B_{2u} 26.20 [0.06] B_{2u}</td>
<td></td>
</tr>
<tr>
<td>27.2</td>
<td></td>
<td></td>
<td></td>
<td>27.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27.36</td>
<td>[0.01] A_g</td>
<td></td>
<td></td>
<td>27.36</td>
<td>[0.01] A_g  27.97 [0.03] A_g  26.96 [0.07] A_g</td>
<td></td>
</tr>
<tr>
<td>28.84</td>
<td>[0.05] A_g</td>
<td></td>
<td></td>
<td>28.84</td>
<td>[0.05] A_g  28.84 [0.05] A_g  28.05 [0.08] A_g</td>
<td></td>
</tr>
<tr>
<td>30.54</td>
<td>[0.01] B_{1u}</td>
<td></td>
<td></td>
<td>30.54</td>
<td>[0.01] B_{1u} 30.54 [0.01] B_{1u}</td>
<td></td>
</tr>
<tr>
<td>30.93</td>
<td>[0.05] A_g</td>
<td></td>
<td></td>
<td>30.93</td>
<td>[0.05] A_g  30.93 [0.05] A_g  33.14 [0.01] A_g</td>
<td></td>
</tr>
</tbody>
</table>

a All ionization energies are in eV.
b From table 3 of Ref. [19], non-diagonal Green's function results.
c Poles above 35 eV were not reported for the 2ph-TDA calculation [19].
d Pole strength less than 0.01.
e The peak fitted to this position represents a continuum of states – see text for discussion, Section 4.1.
the binding energy spectrum of ethylene is split into many 'satellite' states in the region above 23 eV up to at least 40 eV. The most intense $(2a_g)^{-1}$ peaks are reported to have binding energies of 23.7 and 27.2 eV [13,14,18,21,22].

### 4.1. Binding energy spectra

Fig. 1 shows the binding energy spectra of C$_2$H$_4$ from 5 to 50 eV for measurements at relative azimuthal angles of $\phi = 0^\circ$ and $8^\circ$, obtained at an impact energy of (1200 eV + binding energy) on a common intensity scale and at an energy resolution of 1.6 eV FWHM. The energy scale was calibrated with respect to the $(1b_3u)^{-1}$ vertical ionization potential as measured by high resolution photoelectron spectroscopy [22,44]. Gaussian peak shapes have been fitted to the main peaks throughout the spectrum in Fig. 1 using vertical ionization potentials and Franck-Condon widths (folded with the EMS instrumental energy width of 1.6 eV FWHM) estimated from photoelectron spectroscopy measurements [18,22,44]. The relative energy spacings of the Gaussian peaks were estimated from the vertical ionization potentials, with small adjustments to compensate for the asymmetries in the shapes of the Franck-Condon envelopes. The absolute binding energy scale was set by assigning the energy of the $(1b_3u)^{-1}$ peak to the vertical ionization potential as measured by high-resolution PES [44]. Measured and calculated ionization potentials from the literature and those used in the present work are shown in Table 2.

In the outer valence region of the $\phi = 0^\circ$ and $8^\circ$ experimental binding energy spectra, several features can be seen below 21 eV in each spectrum. Peaks due to the $(1b_3u)^{-1}$ and $(2b_1u)^{-1}$ ionization processes are at binding energies of 10.51 and 19.1 eV respectively, a spacing which is consistent with the results of photoelectron spectroscopy [18,22,44]. These two peaks display characteristic 'p-type' behavior, having a greater intensity at $\phi = 8^\circ$ than at $\phi = 0^\circ$. In the case of the $(2b_1u)^{-1}$ ionization process, two functions, located at binding energies of 19.0 and 19.8 eV respectively, have been used to fit the asymmetry in the $(2b_1u)^{-1}$ ionization peak. The sum of these two functions, with a centroid at 19.1 eV, is shown in Fig. 1, rather than the individual functions, as it is not clear whether this asymmetry is due to the vibrational structure of the $(2b_1u)^{-1}$ ionization process or the presence of a $(2b_3g)^{-1}$ pole at approximately 20 eV (see Table 2 and Fig. 2). The energies of the $(1b_3g)^{-1}$, $(3a_g)^{-1}$ and $(1b_2u)^{-1}$ ionizations are too closely spaced to permit complete resolution of the individual peaks in the present EMS binding energy spectra, however these processes can be identified using the Gaussian functions fitted at energies consistent with high resolution photoelectron spectra [22,44]. In the binding energy spectrum recorded at $\phi = 0^\circ$ the intense feature at 14.7 eV is predominantly due to the $(3a_g)^{-1}$ ionization. The peak in the $\phi = 8^\circ$ binding energy spectrum at 15.9 eV is mostly attributable to the $(1b_2u)^{-1}$ ionization process. The feature at 12.8 eV between the $(1b_3u)^{-1}$ (10.5 eV) and $(1b_2g)^{-1}$ (15.9 eV) peaks is mainly due to the $(1b_3g)^{-1}$ process. Care has been taken when performing the Gaussian fits to the data to avoid cross-contamination between the s-type $(3a_g)^{-1}$ and p-type $(1b_3g)^{-1}$ and $(1b_2u)^{-1}$ processes.
The XPS spectrum of ethylene reported by Gelius [17] in 1974 was one of the first observations of satellite structure in molecular binding energy spectra. In a subsequent EMS study of C2H4, Dixon et al. [13] assigned the binding energy region of ~ 21–30 eV to the (2ag)⁻¹ inner valence process on the basis of the experimental momentum profiles of the two main features, at 24 and 27 eV. These profiles have 's-type' behavior, with a maximum intensity at \( \phi = 0^\circ \) (near \( p = 0 \)), decreasing monotonically at higher momenta, consistent with the shapes of calculated 2ag momentum profiles. Even with the modest statistics of these earlier studies [13,14], it was clear from the shapes of the observed momentum profiles that these two large satellite structures were due to (2ag)⁻¹ and not the (3ag)⁻¹ ionization process, which has a much narrower s-type momentum distribution. This assignment is supported by the CI calculations of Martin and Davidson [20], the Greens' function treatment of Cederbaum et al. [19], the CI results of Wasada et al. [26] and by the MRSD-CI calculated ionization energies reported in the present work. All three theoretical results predict several (2ag)⁻¹ poles concentrated in the 23–28 eV region. The relative intensities observed in this region in the binding energy spectra at \( \phi = 0^\circ \) and \( \phi = 8^\circ \) (Fig. 1) suggest that the majority of the inner valence structure out to the limit of the present measurements at 50 eV is due to (2ag)⁻¹ ionization channels. This assignment is further supported by (a) the quality of the fit of the relative integrated areas from 21–50 eV of the \( \phi = 0^\circ \) and \( \phi = 8^\circ \) spectra to the theoretical 2ag momentum profiles (see Fig. 9, discussed in the next section); and (b) the XMP's shapes at binding energies of 23.5, 27.2 and 30.4 eV (see discussion in Section 4.2 and Fig. 9).

In Fig. 2, the binding energy spectra summed over fifteen \( \phi \) angles (in panel (a) – see the following section for details), is compared with the results of several many body calculations (panels (b), (c) and (d)). In order to compare with the theoretical pole strengths, the intensities in the summed binding energy spectra have been corrected for the variable angle spacings and multiplied by \( 4\pi p^2 \) (calculated for each binding energy and angle \( \phi \)) to provide an approximate integration over momentum from 0 to 2 au. The calculations have been folded with the experimental resolution and the same Franck–Condon widths obtained from the PES spectra [18,22,44] that were used to fit Fig. 1, to obtain Fig. 2b, 2c and 2d. One of the earlier ab initio many-body calculations for ethylene, a Greens' function treatment, is shown in Fig. 2d. This is a non-diagonal two particle-hole Tamm–Dancoff calculation taken from the work of Cederbaum et al. [19]. The results of the symmetry-adapted cluster calculation (SAC-CI) of Wasada et al. [26] are shown in Fig. 2c. Also shown (in Fig. 2b) are the energies and spectroscopic factors of the MRSD-CI roots calculated in the course of the present work. It can be seen from Fig. 2 and the pole positions in Table 2 that all three calculations give a reasonable reproduction of the energy positions of the outer valence peaks, however the Greens' function results seem to be uniformly shifted by -0.3 eV. The relative intensities are reasonably reproduced by the calculations, with the MRSD-CI giving the best overall description. While the three treatments show some differences in their descriptions of the inner valence region, all attribute most of the inner valence intensity to the (2ag)⁻¹ process (see
Table 2). The MRSD-CI calculation (Fig. 2b) reproduces the relative shape of the inner valence region quantitatively. While the SAC-CI treatment [26] (Fig. 2c) gives a qualitative description of the inner valence shapes with good reproduction of the energy of the major feature at 23.5 eV, it clearly gives incorrect relative intensities for the poles. In particular, the intensity of the (2a_g)^{-1} peak found at 27.2 eV in the experimental spectra is significantly underestimated. The 2ph-TDA Greens' function results [19] (Fig. 2d) predict a twinned splitting of the A_g ion states in the region of 23 eV instead of one 'main line' (see Table 2). Murray and Davidson [27] concluded that this structure was an artifact of the truncation of the virtual orbital space in the Greens’ function calculation. The present MRSD-CI calculation, however also predicts a twinning due to a near degeneracy of the (2a_g)^{-1} configuration and the (1b_3g)^{-2} 3a_g configuration. In addition, the large number of small A_g poles (see Table 2) in the present MRSD-CI calculation predicts that many states populate this energy region.

4.2. Comparison of experimental and theoretical momentum profiles

In order to obtain the experimental momentum profiles corresponding to the main peaks in the outer and inner valence regions, fifteen binding energy spectra over the energy range of 5–32 eV were collected at the out-of-plane azimuthal angles φ = 0°, 1°, 2°, 3°, 4°, 5°, 6°, 7°, 8°, 9°, 11°, 13°, 15°, 20° and 25° in a series of sequential repetitive scans. These binding energy spectra (total accumulation time: 380 h) are shown in Fig. 3. Nine functions were fitted to each spectrum, with the same energy positions and widths used for the wider energy range spectra fits in Fig. 1 (see Section 4.1). The fitted functions are shown with broken lines on Fig. 3 with their sum as a solid curve. Momentum profiles for each of the peaks observed in the binding energy spectra were obtained from the deconvoluted peak areas of the fitted Gaussian functions using the φ to p transformation of Eq. (1). Since these momentum profiles were obtained from a series of sequential scans, they share a common intensity scale.

The experimental data (shown as filled circles with error bars estimated from the quality of the fit) and also a range of calculated spherically averaged momentum profiles (lines) corresponding to the 1b_3u, 1b_3g, 3a_g, 1b_2u, 2b_1u and 2a_g orbitals of ethylene are shown in the upper panels of Figs. 4–9 respectively. The instrumental angular resolution has been accounted for in all the theoretical momentum profiles using the GW-PG method [32]. In the lower panels the corresponding momentum and position space orbital electron density maps are shown for an oriented ethylene molecule, calculated using the 196-GTO wavefunction. The 1b_3u experimental momentum profile (BE = 10.51 eV) has been normalized to the 196-G(CI) calculation (Fig. 4). It should be noted that the 196-G(CI) TMPs are shown renormalized to unit pole strength. This assignment assumes that all of the missing pole strength, in excess of the main pole (see Table 2) for each ionization process is contained within the broad envelope of the respec-
tive peaks fitted to the binding energy spectra. With this single point normalization all theoretical moment profiles and experimental measurements share a common intensity scale. The relative intensities obtained from the fits to the wider energy range spectra (Fig. 1) at 0° and 8° are also shown on these figures as indicated by the filled squares. A single point normalization of the long (5–50 eV) binding energy range data to the short (5–32 eV) range data was performed at ϕ = 8° for the 1b3u momentum profiles (Fig. 4).

The 1b3u experimental and theoretical momentum profiles are shown in Fig. 4. There is generally good agreement of the data with the calculated profiles with the exception of the smaller basis set HF results, in particular the STO-3G and 6-311G basis sets. As in the case of the 1τ HOMO of acetylene [45], the Hartree–Fock models describe this orbital in C2H4 as a carbon–carbon π-bond with no s-type character (see also description in Section 4.5).

In the case of the 1b3u momentum profile in Fig. 5, there is a clear discrepancy between the experimental data and all of the theoretical predictions, particularly at low momentum. A ‘p-type’ momentum distribution with a node at zero momentum is predicted by all of the calculations, due to the symmetry of this predominantly C–H bonding orbital. Also, while all of the theoretical models provide reasonable descriptions of the the shape of the 3ag experimental momentum profile (Fig. 6), in all cases the calculated profiles predict substantially less intensity below 1 au than is observed experimentally. Assuming that the 196-G(CI) calculation provides a good description of both the shapes and relative intensities of the ethylene valence momentum profiles, there are several possible reasons which could
account for these two discrepancies. Any impurity in
the ethylene sample with ionization potential be-
tween 13 and 14 eV having an 's-type' momentum
profile would cause increases in intensity at low
momentum in both of the 1b3g and 3ag XMPs.
However, no commonly expected impurities would
give rise to a peak of such symmetry at this energy
and, in addition, if an impurity was present, addi-
tional structures would have been observed in the
binding energy spectra at other energies. In fact it is
found that all of the features observed in both the
short (Fig. 3) and long range (Fig. 1) binding energy
spectra, as discussed in Section 4.1 earlier, can be
assigned to C2H4 ionization process, with no traces
of other species. The binding energy spectra are also
in good agreement with earlier (lower precision)
single channel EMS measurements [46].

A further possibility is that inadequacies in the
peak fitting and deconvolution procedure, possibly
due to the close energy spacing of the (1b3g)−1 and
(3ag)−1 ionization energies and the considerable in-
tensity of the 3ag profile at low momentum, could
cause some contamination of the fitted areas of the
1b3g momentum profile from the 3ag profile in the
low momentum range (Fig. 6). If this is the case, the
large discrepancy between theory and experiment
should disappear if the sum of all the outer valence
XMPs is compared with a sufficiently high level
model. Fig. 10 shows the sum of the fitted peaks
corresponding to the four lowest binding energy
momentum profiles (1b3g, 1b3u, 3ag and 1b2u) as
filled circles and squares for the short binding energy
range (Fig. 3) and long binding energy range data
(Fig. 1) respectively. These summed intensities are
plotted as a function of $\phi$ because of the slight variation of momentum with binding energy (see Eq. (1)). The open circles and squares are the corresponding intensities obtained by summing the binding energy spectral data points over the 5–17 eV energy range for each angle $\phi$. Also shown on Fig. 10 is the curve obtained by summing the four (resolution-folded) outer valence 196-G(CI) theoretical angle profiles. It is evident that there is a high degree of consistency between the summed data points and the summed deconvolutions in the outer valence region and that the discrepancy between the theoretical predictions and experiment at low momentum remains, indicating that this difference is not due to errors in the deconvolution procedures.

An additional possibility to consider is that the single point normalization factor between the experimental and theoretical momentum profiles could be chosen differently. While the normalization factor could be adjusted so that the 3ag TMPs agree for intensity with the XMP on Fig. 6, the intensities of the calculated momentum profiles for many of the other orbitals, especially the 1b3g and 3ag profiles (Figs. 4, 5 and 7), would be in much poorer agreement with the experimental data.

While the mismatches for the 1b3g and 3ag momentum profiles may be due to experimental or data reduction difficulties, it is also possible that the 196-G(CI) calculation is an insufficiently complete model for the momentum profiles of ethylene. For the 3ag momentum profile, the mismatch between theory and experiment could be resolved by an increase in the predicted intensity at low momentum. The presence of further s-type poles in the 11–13 eV region, not predicted by the present calculations,
could be the reason for the intensity mismatches. The total intensity of a single pole would not be required to be more than 5% of either of the \((3a_g)^{-1}\) or \((2a_g)^{-1}\) processes, indeed a number of very small poles over a narrow energy range would have the same effect. Any such s-type poles in the 11–13 eV region would contribute intensity in the energy region of the \(1b_{3g}\) and the \(3a_g\) processes, and might account for the respective disparities between the calculated and measured intensities. It is interesting to note that in two high resolution photoelectron studies [21,44] of ethylene, the \(1b_{3g}\) band has a structure which could suggest the existence of multiple poles in this region. An alternative explanation for the large mismatch between experiment and theory is that the \((1b_{3g})^{-1}\) ionization process could involve a symmetry breaking possibly due to vibrational coupling. This seems unlikely however since the \(3a_g\) and \(2a_g\) XMPs, which would be expected to donate intensity to the lower energy ionization in this picture, both have intensities greater than predicted by the 196-G(CI) calculations.

The double maximum observed in the \(3a_g\) momentum profile of \(C_2H_4\) (Fig. 6) is a result of the \(\sigma_{C-\text{-}C}\) and \(\sigma_{C-\text{-}H}\) multi-center bonding behavior of this orbital (see Section 4.5). While an intensity mismatch occurs (see above discussion) between the \(3a_g\) XMP and the TMMs, the ‘sp-type’ shape of this momentum profile is nevertheless quite well reproduced by all the theoretical methods (Fig. 6) over the entire momentum range. The situation for the multi-center \(3\sigma\) momentum profile of acetylene [45] is similar. In contrast, for many other systems it has been found that ‘secondary’ higher momentum maxima are very sensitive to the quality and size of the basis sets at the Hartree–Fock level of treatment (see, for example Refs. [11,47–49]).

The shapes of the \(1b_{2u}\) and the \(2b_{1u}\) XMPs are well reproduced by all of the respective calculation methods, as is the intensity of the \(1b_{2u}\) XMP. The \(2b_{1u}\) experimental profile, however, is significantly lower in intensity than all of the theoretical profiles. For example, it can be seen that the 196-G(CI) result, scaled by 93% (shown as a dashed line on Fig. 8) fits the \(2b_{1u}\) XMP very well. This small drop in intensity is consistent with the many-body calculations, discussed in Section 4.1 earlier, which predict a rather smaller pole strength for the main \((2b_{1u})^{-1}\)

\[
\begin{array}{c}
\text{C}_2\text{H}_4 \\
\text{outer valence} \\
5–17\text{ eV}
\end{array}
\]

Fig. 10. Summed angle distributions for the outer valence region. The unfilled circles and squares represent the summed counts in the short range (see Fig. 3) and long range (see Fig. 1) binding energy spectra for the four outermost ionization processes; the solid symbols represent the respective summed areas from the deconvolution procedure — see text for details. The lines are the sum of the indicated theoretical (resolution folded) angle profiles for the four outermost ionization processes.
process than for the other outer valence ionization processes (see Table 2).

Fig. 9 shows the experimental and theoretical momentum profiles for the \(2a_g\) manifold. There is reasonable agreement for intensity for all \(2a_g\) TMPs with experiment in Fig. 9. It should be noted (see Section 4.1) that all of the experimentally observed intensity (in the BES of Fig. 1) above 21 eV has been assigned to the \(2a_g\) ionization manifold. The unfilled squares on Fig. 9 represent the relative summed areas from 21–50 eV in the 0° and 8° binding energy spectra (Fig. 1). The filled squares show the deconvoluted areas of the 0° and 8° BES for the 23.5, 27.2 and 30.4 eV \((2a_g)^{-1}\) peaks. The filled circles on Fig. 9 are the fitted peak areas of the 23.5 eV peak in the shorter range, multi-angles BES (Fig. 3). Finally, this filled circle data set has been scaled up by a factor equal to the ratio of the unfilled squares to the filled squares for the 23.5 eV peak at \(\phi = 0°\), and these scaled data points are shown on Fig. 9 as unfilled circles. If there are any contributions from ionization processes of p-type symmetry in the 21–50 eV energy region, then the intensity of the scaled momentum profile (open circles) would be greater than expected from the \((2a_g)^{-1}\) process alone. Furthermore, if some \(2a_g\) poles are outside of the integrated BE spectral region (21–50 eV), then the relative intensity of the \(2a_g\) experimental momentum profile would be lower than expected from a good calculation. In fact, the \(2a_g\) XMP has a somewhat higher intensity than predicted even by the 196-G(CI) calculation. This difference may be due to experimental uncertainties or to small contaminations from poles of other ionization manifolds or some residual limitation of the calculations. Nevertheless the agreement for intensity between theory and experiment is reasonable in Fig. 9 and generally confirms the overall assignment of the 20–50 eV BES region to the \((2a_g)^{-1}\) ionization process as made earlier in Section 4.1. This assignment is further supported by the high degree of consistency between the \(\phi\) of the inner valence experimental profiles obtained from the fitted areas at the binding energies of 23.5, 27.2 and 30.4 eV and the scaled 196-G(CI) calculations as shown in Fig. 9. Note that, while the 23.5 and 27.2 eV profiles correspond to groupings of a few stronger poles as predicted by the 196-G(CI) calculation (see Fig. 2 and Table 2), the distribution at 30.4 eV most likely represents a number of closely spaced, very low intensity \((2a_g)^{-1}\) poles rather than a small set of strong ones. It should be noted that the integrated area of the \(\phi = 8°\) BE spectrum (Fig. 1), plotted as an open square at ~ 0.75 au on Fig. 9, is high when compared with the scaled momentum profile (open circles). This may be indicative of small contributions in the ionization energy range above 20 eV from p-type processes. This is consistent with the small depletion in measured intensity of the \((2b_{1u})^{-1}\) process relative to theory as seen in Fig. 8 and discussed in the preceding paragraph. It is thus possible that higher energy \((2b_{1u})^{-1}\) p-type poles contribute to the long range binding energy sum at 8° (open square) plotted on Fig. 9. In fact, the 196-G(CI) calculation predicts at least one significant \(B_{1u}\) pole (> 1%) at 24.42 eV.

Turning now to a more detailed examination of the calculated momentum profiles, a comparison of the profiles calculated from the 196-G(CI) ion–neutral overlap using correlated wavefunctions and those of the 196-GTO Hartree–Fock calculation reveals very good agreement for shape in each case (Fig. 4–9). When compared with the experimental results, these two calculations also give a good reproduction of most of the experimental momentum profiles for both intensity and shape, with the exception of the lower momentum regions of the \(1b_{3g}\) and \(3a_g\) profiles discussed earlier.

The Hartree–Fock calculations of the momentum profiles (see Figs. 4–9 and also Table 1 for selected properties) generally show progressively improving agreement with the experimental data as basis set size and diffuseness increases. The minimal basis set STO-3G results (curve 1) have generally the poorest agreement with the experimental momentum profiles, especially for the \(1b_{3u}\) and \(3a_g\) XMPs (see Fig. 4 and 6 and also \(P_{MAX}\) of the \(1b_{3u}\) in Table 1). However, this limited basis set also gives a surprisingly good description of the rest of the momentum profiles (Figs. 5–9). It should be noted, however, that the STO-3G quadrupole moment is very low compared with the experimental value (Table 1), which is a reflection of the limited flexibility of this minimal basis set. In addition, the total energy is very poor indicating the low degree of basis set saturation in the STO-3G SCF treatment.

The 6-311G calculation (momentum profile 2 on
Figs. 4–9) provides results that are generally intermediate between the minimal and larger results. The 6-311G basis set provides a particularly poor description of the 1b3u HOMO XMP. Curiously, the calculated quadrupole moment for this basis set is the closest to both the experimental and the many-body calculation values (see Table 1), although this likely reflects a fortuitous cancellation of errors, rather than a truly good model of the charge densities. This kind of result emphasizes the need for consideration of an appropriate range of properties for the evaluation of theoretical models and the development of 'universal wavefunctions'.

It is interesting to note that the 6-311 + + G** basis set (curve 3 on Figs. 4–9) produces similar results to the near Hartree–Fock limit calculations for many of the theoretical momentum profiles, with much less computational effort. This standard basis set gives a reasonably converged quadrupole moment, however it also results in a total energy significantly above those of the better Hartree–Fock calculations (see Table 1). These results illustrate the fact that while good reproduction of certain properties can be attained with limited basis sets, it is often at the expense of others.

The momentum profiles calculated using the aug-cc-pVTZ and 196-GTO near Hartree–Fock limit basis sets (curves 4 and 5 on Figs. 4–9), both provide an excellent description of the shapes of all of the XMPs except the 1b3g profile, and a reasonable reproduction of the observed intensities. There is also a large improvement in the total energy (Table 1) using these basis sets when compared with the 6-311G and 6-311 + + G** results. As expected the 196-GTO basis provides a somewhat better total energy than the slightly smaller 182-GTO aug-cc-pVTZ basis set. However, it should be noted that none of the Hartree–Fock calculations provide a very accurate quadrupole moment although the values appear to be converging to ~ 1.53 au. It is clear that additional efforts beyond Hartree–Fock limit treatments must be made in order to predict a more accurate value of the quadrupole moment.

As discussed above, there is little change in the TMPs with the inclusion of electron correlation and relaxation effects for example in the 196-G(CI) overlap calculation. This is similar to behaviour observed for acetylene [45] and methane [50] for which the inclusion of electron correlation and relaxation did not have a large effect on the shapes of the valence orbital momentum profiles, even though, as is the case for ethylene, significant improvements were found in the total energies in both cases. Note however, that there are large improvements in the calculated total energy and quadrupole moment of ethylene (Table 1) with the inclusion of correlation and relaxation. Nevertheless a small discrepancy remains for the quadrupole moment. However, it should be noted that all of the calculations in the present work (with the exception of those in Section 4.4) have been performed at a fixed geometry, while the experimental value reflects a vibrating and rotating molecule. In general, the 196-G(CI) TMPs are very similar in shape to the results of the 196-GTO Hartree–Fock calculation.

In the course of the present work several even larger basis sets then the 196-GTO were developed and used for Hartree–Fock and MRSD-CI calculations of the 1b3u momentum profile of $C_2H_4$. All calculations include the angular resolution.
found in earlier work for NH₃, H₂O and HF, that when using MRSD-CI methods, the basis set size and especially the extreme diffuseness and flexibility of the SCF wavefunctions are very important for reproduction of the low momentum region of the lowest energy ('HOMO') XMPS and for obtaining other properties with high accuracy [1-3,7,51]. It was found for these cases that achieving a high degree of basis set saturation and convergence not only for total energies but also for other properties such as the dipole moment and momentum profiles at the Hartree–Fock level was an essential precursor to meaningful inclusion of electronic correlation and relaxation. Once effective convergence to the Hartree–Fock limits had been achieved for a suitable range of test properties (most often total energy, dipole moment and the momentum distributions – see the discussion in Refs. [1,2,7]) the importance of including correlation and relaxation effects in the wavefunction for reproduction of the XMPS for these second row hydrides was clearly demonstrated. For NH₃, H₂O and HF, the ‘HOMO’ experimental momentum profile was not well reproduced at low momentum even at the near Hartree–Fock limit level. Good agreement was only obtained by a Dyson orbital (i.e. the full overlap of ion and neutral correlated wavefunctions). The results of similar investigations for C₂H₄ are shown in Fig. 11. The effects of finite instrumental momentum resolution have been accounted for in the profiles using the Gaussian-weighted planar grid method [32].

In these further investigations, calculations for ethylene were carried out with the same even-tempered [14s, 10p, 3d, 1f/6s, 3p, 1d] basis set used earlier for C₂H₂ [45]. However, trial MRSD-CI calculations with this basis set resulted in a poorer total energy than for the 196-G(CI) (5c) results. After carefully comparing the 196-GTO (5) and the original acetylene basis sets, a second, intermediate basis set was formed from the 196-GTO (5) basis set by uncontracting some of the s and p functions and adding one more diffuse s ( αₜ = 0.033728) and p (αₜ = 0.020323) function on the carbon atoms using Partridge’s recommended C⁻ anion supplements [41]. As Partridge did not report a supplementary function for hydrogen, an additional s function with an exponent of 0.018644 was found by minimizing the SCF energy of H⁺.

All the possible combinations of the whole s, p and d sets of the intermediate basis and the original acetylene basis set were evaluated in atomic and molecular fragment SCF trial calculations. The fragments used for carbon were C⁺, C, C⁻ and C₂ (3Σ⁻). Neutral H, H⁺, H₂ and H₂⁺ were used for hydrogen. The properties of the final 228-GTO basis, [14s, 12p, 3d, 1f/7s, 3p, 1d], which corresponds to the lowest-energy combination, can be found in Table 1. The Cartesian s and p components were removed from d and f functions respectively. This Gaussian basis set was then used for the 228-G(CI) (6c) calculations. For the neutral molecule, 21 reference configurations were included in the final MRSD-CI calculation which generated 20443 spin-adapted configurations and recovered about 80% of the empirical valence correlation energy. Twenty-three reference configurations were included for the cation which generated 35619 spin-adapted configurations and also recovered about 80% of the expected valence correlation energy. The resulting properties for the 228-G(CI) (6c) calculations are shown in Table 1. Very little change occurred in the resulting TMP (see Fig. 11). In fact, the 228-GTO (6) and 228-G(CI) (6c) TMPs are almost identical to the 196-GTO (5) and 196-G(CI) (5c) results respectively. It is interesting to note (see Table 1) that while a slightly better SCF energy is achieved using the 228-GTO (6) basis set than with the 196-GTO (5), the 228-G(CI) (6c) result is slightly worse than the total energy obtained by the 196-G(CI) (5c) treatment. It is clear that the nature of the SCF basis set as well as the size are important considerations for CI calculations.

In view of this situation, the final 228-GTO (6) basis was expanded to a 234-GTO (7) basis, [15s, 12p, 3d, 1f/8s, 3p, 1d], by replacing the last three s functions on carbon with four even-tempered s functions and uncontracting the last s function from the 1s hydrogen function. These 234-GTO (7) SCF and corresponding (molecular orbital based) CI results (the first CI calculation on this basis set, 234-G(CI) (7a)) are collected in Table 1. To ensure that the wavefunctions were sufficiently converged, the CI calculation was redone with natural orbitals obtained from the original MRSD-CI wavefunction of the neutral molecule. The first natural orbital calculation (the second CI calculation on this basis set, 234-G(CI) (7b)) was performed using the same con-
figuration space cut-offs as the 196-G(CI) (5c) treatment. A second natural orbital calculation (the third CI calculation on this basis set, 234-G(CI) (7c)) was obtained by lowering the perturbation selection threshold to \(0.32 \times 10^{-7}\) hartree, which recovered about 85% of the empirical valence correlation energy, a 5% improvement. Both of the two natural orbital calculations show a slight decrease in the intensity of the TMP but the positions of the maxima \(P_{\text{MAX}}\) are found to be unchanged in both cases. A final CI calculation (the fourth CI calculation on this basis set, 234-G(CI) (7d)) was performed using natural orbitals which were obtained from the original MRSD-CI wavefunction of the cation. From Table 1, it can be seen that this calculation obtained the second lowest energy for the neutral molecule. However, the corresponding TMP (7d in Fig. 11) is very similar to the other 234-G(CI) results.

Overall, of all the various treatments, the third 234-G(CI) calculation (7c) has the best characteristics, giving the lowest value for total energy and the best calculated quadrupole moment together with an excellent reproduction of the XMP. In the case of the quadrupole moment it should be noted that the calculations involve the equilibrium geometry whereas the experimental results are for a rotating and vibrating molecule.

4.3. Calculation of the momentum profiles of ethylene using density functional theory

An alternative approach to the Hartree–Fock and configuration interaction methods for calculating momentum profiles has recently been proposed by Casida [30] and Duffy and co-workers [29] using Kohn–Sham density functional theory. Density functional theory (DFT) has been remarkably successful in describing the electronic structure of molecular and solid state systems [52–54], but until recently [11,29] had not been employed for EMS structure factor calculations. Three density functional theory calculations are presented here using a truncated form of the aug-cc-pVTZ basis set (number 4) described in Section 3. In addition, a Hartree–Fock calculation using the same basis set has been carried out for the purposes of comparison. The f functions of Kendall et al. [34] for the carbon nuclei were not included due to limitations of the present density functional program (deMon [55,56]). To maintain balance in the basis sets, the d augmentations on the hydrogen centers were also removed, forming the final 122-pVTZ basis used for both the Hartree–Fock and density functional SCF calculations discussed here. It should be noted that all calculations (HF and DFT) which were performed using this truncated 122-pVTZ basis were carried out with six-membered Cartesian d functions. The Hartree–Fock result for this truncated basis set is shown on Table 3 and Fig. 12 as calculation 122-pVTZ (4f). The local density approximation for the exchange and correlation potential, \(v_{\text{xc}}\), of Vosko et al. [57] (122-pVTZ-VWN, 4f) and two non-local potential functionals, the gradient corrected exchange term of Perdew and Yue [58] with Perdew’s gradient correction to the correlation energy [59] (122-pVTZ-P, 4p) and a combination of the Perdew correlation correction and Becke’s gradient correction to the exchange energy [60] (122-pVTZ-BP, 4b), have all been used to generate molecular densities for ethylene from the 122-pVTZ basis set, using the deMon program [55,56]. The

<table>
<thead>
<tr>
<th>Table 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparison of density functional theory calculations and experimental properties for ethylene</td>
</tr>
<tr>
<td>Wavefunction</td>
</tr>
<tr>
<td>4t</td>
</tr>
<tr>
<td>4f</td>
</tr>
<tr>
<td>4p</td>
</tr>
<tr>
<td>4b</td>
</tr>
<tr>
<td>experimental</td>
</tr>
</tbody>
</table>

\(^a\) Momentum value for maximum in the \(1b_{3u}\) momentum profile, Fig. 12.

\(^b\) Hartree–Fock calculation using equivalent basis set.

\(^c\) Ref. [66].
Kohn–Sham orbitals have been extracted from the result, and EMS cross sections calculated using Eq. (4). The instrumental angular resolution functions were then incorporated to produce the theoretical momentum profiles shown on Fig. 12. Other properties for these calculations are shown in Table 3.

It can be seen from Fig. 12 that overall there is very good agreement of all of the TKSA momentum profiles with those of the 196-G(CI) calculations discussed in Section 4.2 above. Also, the medium-sized 122-pVTZ Hartree–Fock calculation appears to be converged to the very large 196-GTO results for all of the momentum profiles. However there are some small but interesting differences between the 196-G(CI) TMPs and the density functional results. In particular, the intensities predicted for valence momentum profiles vary by as much as 5% from the 196-G(CI) results. The local density approximation results, 122-pVTZ-VWN (4l), appear to be at least as good a model of the ethylene momentum profiles as the more sophisticated non-local potential treatments, 122-pVTZ-P (4p) and 122-pVTZ-BP (4b), which include density-gradient corrections to the exchange and correlation energies. It is interesting to note that in the case of the 3a_g momentum distributions all of the DFT profiles are higher at lower momentum than those given by the CI or HF treatments. Although not in complete agreement with the 3a_g experimental profile, nonetheless, the DFT results appear to offer slightly improved agreement with the XMP over the 196-G(CI) calculations. However, the DFT results also predict intensities that are somewhat higher than both the XMPs and 196-G(CI) TMPs for each of the 1b_2u and 2b_1u profiles. The quadrupole moments, shown on Table 3 are also very reasonable for the DFT results and are at least as good as all of the CI results reported in the previous sections. Although the DFT result using the Becke–Perdew potential offers a quadrupole moment closest to the reported value, given the uncertainties in the measurement and the necessary corrections for nuclear motion, it is not clear that the DFT value is any better or worse that those derived from the large 234-G(CI) calculations. The total energies resulting from the DFT calculations should be used with some caution due to the well known difficulty of obtaining accurate absolute energies [52–54]. However, it must be said that all of the DFT treatments provide a generally excellent description of all of the other properties evaluated in the present work.

4.4. Vibrational effects in the 1b_3u momentum distribution

In the results discussed in the previous sections, all of the theoretical calculations have been performed at the equilibrium geometry [36]. This completely ignores vibrational effects on the momentum profiles which may include a non-zero cross section at zero momentum for ‘p-type’ orbitals (even before resolution folding). The effects of vibrational averaging on momentum profiles have not been widely investigated, although studies of water [61,62] have shown negligible vibrational effects on the 1b_1u non-bonding HOMO momentum distribution. It is therefore of some interest to investigate the effects of
vibrational averaging for a bonding orbital, such as the HOMO \( \pi \) 1b\(_{3u} \) orbital of ethylene where vibrational effects might be expected to be more significant.

Ethylene has 12 normal vibrational modes [35]. Even assuming that in the initial state only the lowest vibrational levels are occupied, the final state will be vibrationally hot due to the Franck-Condon overlap for the ionization process. Summing over all the possible vibrational states, one can rewrite the momentum distribution (\( \rho(p) = |\langle p \psi_f^{N-1} | \psi_0^N \rangle|^2 \) from Eq. (2)) as

\[
\rho(p) = \sum_j |\langle \psi_0^N \sum_{\nu} \phi_{\nu} | p \psi_f^{N-1} \sum_{\nu'} \phi_{\nu'} \rangle| \times |\langle \psi_0^N \sum_{\nu} \phi_{\nu} | p \psi_f^{N-1} \sum_{\nu'} \phi_{\nu'} \rangle|,
\]

where \( \psi_f^{N-1} \) and \( \psi_0^N \) are the final and initial state electronic wavefunctions; \( p \) is a spin-orbital representing the target electron; \( \phi_{\nu} \) is the vibrational wavefunction (see Eq. (9 below)), where \( \nu \) (\( \nu' \)) denotes a normal mode in the initial (final) state, and \( j \) indicates the \( j \)th vibration level of the mode \( \nu \).

Ignoring the dependence of \( \psi_f^{N-1} \) on nuclear coordinate the following identity is true for the final state vibrational wavefunction,

\[
\sum_j |\langle \sum_{\nu} \phi_{\nu} | p \psi_f^{N-1} \sum_{\nu} \phi_{\nu} \rangle|^2 = I.
\]

Thus, from Eq. (5),

\[
\rho(p) = |\langle p \psi_f^{N-1} | \sum_{\nu} \phi_{\nu} \psi_0^N \rangle|^2.
\]

One of the consequences of Eq. (7) is that only those modes which have the same symmetry as the final electronic state will induce non-zero contributions to the calculated momentum distribution at \( p = 0 \). Since ethylene has only one vibrational mode of \( B_{3u} \) symmetry [35], integration over all normal modes reduces to a single vibrational average for the 1b\(_{3u} \) orbital. In this case, the momentum density can then be simplified to

\[
\rho(p) = \int \rho_0(p, Q) |\phi_{B_{3u}}|^2 \, dQ,
\]

\[
\rho_0(p, Q) = |\langle p \psi_f^{N-1} | \psi_0^N \rangle|^2,
\]

where the integral over the normal coordinate \( Q \) indicates averaging over the \( B_{3u} \) vibrational mode. Note that \( \rho_0 \) is not the usual electronic structure factor for EMS calculations (in Eq. (2)) because of its dependence on the normal coordinate, \( Q \).

Assuming a harmonic oscillator, the vibrational wavefunction \( \phi_{\nu} \) is given by

\[
\phi_{\nu} = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha Q^2/2},
\]

where \( \alpha = (2\pi/\mu/v)^2/\hbar, \mu \) is the reduced mass of the molecule and \( \nu \) is the vibration frequency of the mode.

The matrix elements of Eq. (8) can be evaluated exactly (with a coordinate transform) with the assumption of a harmonic vibrational potential using Gauss–Hermite integration. However, the normal coordinate \( Q \) for the vibrational motion must be known in terms of the atomic displacements of the vibrational mode; \( \delta x_N \) for center \( N \). The \( \nu_{B_{3u}} \) [35] mode \(^1\) of ethylene is an out-of-plane 'umbrella' motion with the hydrogen centers all moving in opposite phase to the carbon nuclei:

\[
2m_c \delta x_c + 4m_H \delta x_H = 0,
\]

\[
\frac{\delta x_c}{\delta x_H} = - \frac{4m_H}{2m_c}.
\]

The normal coordinate \( Q \) for this mode is then defined by

\[
\delta X = QC,
\]

where

\[
X^T = (x_H, x_H, x_H, x_H, x_C, x_C),
\]

\[
C^T = \lambda(2m_c, 2m_c, 2m_c, 2m_c, -4m_H, -4m_H),
\]

and where \( \lambda \) is the reduced-mass factor which makes matrix \( C \) unitless.

The vibrationally and spherically averaged \( \rho(p) \) for \( \nu_{B_{3u}} = 949.2 \text{ cm}^{-1} \) from Eq. (8) and the equilibrium geometry momentum density from Eq. (3), both calculated using the 196-GTO SCF basis set are shown in Fig. 13. The vibrationally averaged density is shown with a broken line, the equilibrium geometry density.
try MD with a solid curve. It is clear from Fig. 13 that the effects of averaging over this normal mode are very small and do not significantly change the shape of the momentum density. Note that neither curve in Fig. 13 includes any consideration of the instrumental resolution function.

4.5. Electron density maps for C$_2$H$_4$ in position and momentum space

Also shown on Figs. 4–9, below the panels of the momentum profiles, are the orbital electron density distribution maps in both position and momentum space for an oriented ethylene molecule. They are presented here for pedagogical interest as they compare the electronic density distributions in position and momentum space for this simplest doubly bonded organic molecule. These results complement those presented earlier in the studies of the other small hydrocarbons CH$_4$ [50] and C$_2$H$_2$ [45] and also aid in the understanding of the shapes of the momentum profiles.

In the lower panels of Fig. 4, the 1b$_{3u}$ density maps, cut perpendicular to the plane of the molecule along the double bond show the characteristic double lobed shape of the carbon–carbon π-bonding orbital in both position and momentum space. The multiple lobes of the momentum density map are the result of constructive interference (commonly called ‘bond oscillations’ [63–65]) of the wavefunctions on the carbon centres reflecting the bonding nature of this orbital. The node in the plane of the molecule is reflected in the spherically averaged momentum profile as an intensity minimum at $p = 0$, giving rise to a ‘p-type’ TMP. The small $P_{\text{MAX}}$ value (Table 1) for the 1b$_{3u}$ momentum profile is a reflection of the very diffuse nature of this orbital in position space. In the lower right hand panel of Fig. 5, the 1b$_{3g}$ orbital can be seen to be $\sigma_{\text{C–C}}$ bonding and at the same time $\sigma^*_{\text{C–C}}$ anti-bonding. Again, the participation of this orbital in molecular bonding is reflected as an ‘interference pattern’ in momentum space. Note that the p-type structure of the 1b$_{3g}$ momentum profile arises from nodal planes in the plane of and perpendicular to the C–C bond, in contrast to the 1b$_{3u}$ HOMO. The 3a$_g$ electron densities (Fig. 6) show that this orbital is both strongly $\sigma_{\text{C–C}}$ bonding and weakly $\sigma_{\text{C–H}}$ bonding. The s–p character of the 3a$_g$ momentum profile is a reflection of this multiple center bonding. The spatially contracted $\sigma_{\text{C–C}}$ density gives rise to a strong contribution at $p = 0$ in the momentum profile with a broad high momentum tail, while the much more spatially diffuse $\sigma_{\text{C–H}}$ bonding behavior corresponds to a highly contracted maximum in momentum space at $p = 0$. The combination of both the $\sigma_{\text{C–C}}$ and $\sigma_{\text{C–H}}$ contributions thus give a large maximum at low momentum and a smaller, higher momentum tail in the 3a$_g$ momentum profile.

The 1b$_{2u}$ density distributions (Fig. 7) show that this orbital is strongly $\pi_{\text{C–C}}$ bonding with some $\pi_{\text{C–C}}$ bonding characteristics. This type of orbital is often labeled a $\pi'_{\text{C–C}}$ bond. Note that it lies in the molecular plane, perpendicular to the 1b$_{3u}$ π-system. The 1b$_{2u}$ TMP is p-type due to a nodal plane along the carbon–carbon axis, perpendicular to the plane of the molecule. The maps of the 2b$_{1u}$ charge distribution (Fig. 8), in both position and momentum space, show both the strongly $\sigma_{\text{C–C}}$ anti-bonding nature and the weak $\sigma_{\text{C–H}}$ bonding character of this orbital. The 2b$_{1u}$ orbital primarily involves the C s basis functions, with some hydrogen s contributions. This is a highly contracted orbital in position space and is correspondingly diffuse in momentum space. Finally, the 2a$_g$ orbital (Fig. 9) is $\sigma_{\text{C–C}}$ bonding, like the 3a$_g$ orbital, but has less $\sigma_{\text{C–H}}$ bonding character. The 2a$_g$ is composed of mainly carbon s contributions of
similar magnitude to, but in opposite phase of those making up the $2b_{1u}$ orbital, and has a similarly broad momentum profile.

Acknowledgement

This work received financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the National Science Foundation (USA). One of us (JJN) gratefully acknowledges an NSERC Postgraduate Scholarship. We would also like to thank Dr. N. Cann for very helpful discussions and her many thoughtful comments.

References


