Extrapolating potential energy surfaces by scaling electron correlation at a single geometry

A.J.C. Varandas *, P. Piecuch 1

Departamento de Química, Universidade de Coimbra, 3004-535 Coimbra, Portugal

Received 17 July 2006; in final form 28 August 2006
Available online 12 September 2006

Abstract

It is shown that the molecular potential energy surface corresponding to a high level of ab initio theory can be accurately predicted by performing calculations with smaller basis sets and then scaling the electron correlation at a single point calculated with the larger target basis set.

© 2006 Elsevier B.V. All rights reserved.

Knowledge of a molecular potential energy surface (PES) is key in many fields, ranging from spectroscopy and mechanisms of chemical reactions to force fields in biological molecules and chemical kinetics. Although a limited portion of its τ-dimensional configuration space (τ = 3N – 6 for a N-atom species) near the equilibrium geometry often suffices for studying small-amplitude vibrations, the full PES or the large part of it is required for spectroscopic studies in floppy molecules, chemical reaction dynamics, and molecular dynamics simulations. Whether using a global PES [1] or performing direct dynamics calculations by determining the energies and energy gradients (forces) on the fly [2], it is crucial that the PES or its part relevant to the dynamics simulation is calculated with high accuracy. In both cases, the main challenge is to generate a relatively large number of points on the PES along the relevant nuclear coordinates with high accuracy ab initio methods of quantum chemistry. Because of the large cost of accurate electronic structure calculations, which rapidly grows with the size of the basis set, it is essential to develop a procedure for obtaining high-quality ab initio energies and forces from inexpensive calculations employing smaller basis sets.

Despite major advances in ab initio electronic structure methods, the conventional process of solving the electronic Schrödinger equation point-by-point to generate a molecular PES is a painstaking computational effort which grows as \(m^\tau\), where \(m\) is a typical number of points required to represent a one-dimensional PES cut (typically, \(m \sim 10–30\) for bond distances and 5–20 for bond angles). This compounds the total cost of generating a PES for a given molecular system to at least \(m^\tau n^4\), since accurate ab initio methods scale with at least the fourth power of the number of unoccupied orbitals in a basis set (\(n\)). The calculation of an accurate PES using large basis sets is then affordable only for few-electron molecular systems, where \(n\) is already on the order of 100, with medium to large polyatomic systems usually calculated using small and therefore unrealistic basis sets.

A significant progress in calculations for small molecules has become possible after the introduction of the so-called correlation-consistent (cc) basis sets [3,4], including the aug-cc-pVXZ and aug-cc-pCVXZ basis sets used in this work (abbreviated by AVXZ and ACVXZ, respectively). Built in a systematic manner, that is intended to relate the correlation energy to the cardinal number \(X\), such basis sets prompted the search for laws to extrapolate the correlation energy to the complete basis set (CBS) limit (see, e.g.,
Although low-order perturbation-type laws are known for the dependence of the correlation energy on the partial wave quantum number in two-electron atomic systems and for second-order pair energies in many-electron atoms [9,10], it is not a rigorous quantum number for scaling correlation energy in many-electron atoms and molecules. Several empirical laws defining the dependence of the correlation energy on \( X \) have been suggested (see, e.g., [5–8]), such as \( \Delta E_X = \Delta E_{\infty} + c X^{-3} \), where \( \Delta E_X \) is the correlation energy obtained with the \( X \)-tuple basis set of (aug)cc quality and \( \Delta E_{\infty} \) and \( c \) are parameters determined from the calculations for the two highest affordable values of \( X \), often 5 and 6. An obvious criticism is that this scheme utilizes as many values of \( \Delta E_X \) as the number of parameters in a situation, where the \( X^{-3} \)-dependence of the correlation energy is only approximate. Moreover, because the scaling of \( \Delta E_X \) with \( X \) is geometry-dependent, the CBS extrapolation of a molecular PES must be repeatedly carried out for a large number of nuclear geometries.

This is complicated by the fact that for AVXZ and ACVXZ basis sets, the number of basis functions grows rapidly with \( X \). For example, for the \( N_2, O_2 \), and \( F_2 \) molecules considered in this work, the dimensions of the AVXZ basis sets grow from 46 for \( X = 2 \) to 254 for \( X = 5 \). For larger molecules, the numbers of basis functions are in thousands. Thus, the basis-set extrapolation of a PES for a polyatomic system is a mammoth task even for the fastest computers, without a guarantee that the predicted PES is accurate since there will be a numerical noise in the extrapolated energies at different geometries, and the extrapolation laws employed may not be precise enough.

In this Letter, a simple, yet a robust and accurate extrapolation scheme is proposed that allows a drastic reduction in the labor required to generate a high-quality global PES corresponding to larger basis sets. The proposed scheme shows three prominent features: (a) the calculations of a global PES or PES cut using basis sets of ranks \( X - 2 \) and \( X - 1 \), or smaller, and the calculation of a single point with the target basis set of rank \( X \), accurately predicts the global PES or PES cut corresponding to the basis set of rank \( X \); (b) it works well with basis sets that do not belong to the (aug)cc hierarchy and that are not defined by a cardinal number \( X \); (c) it is applicable to scaling both the dynamical and full correlation energies in a variety of electronic structure calculations, using methods such as the multi-reference configuration interaction (MRCI) and coupled-cluster (CC) approaches. For example, from the PESs obtained in the MRCI/AVDZ and MRCI/AVTZ calculations, a global MRCI/AVQZ PES can be predicted using only one MRCI/AVQZ calculation at some, in principle arbitrary, reference (e.g., equilibrium) geometry. The extrapolation error amounts typically 1–2% of the change in total energy, when going from the PESs obtained with the \( X - 2 \) and \( X - 1 \) basis sets to the PES obtained with the target \( X \) or \( X + 1 \) basis set (the same applies to basis sets, which do not belong to the (aug)cc series, and which will be labeled by an index \( m \)).

The proposed procedure is based on extrapolating the difference between the total energy \( E \) and the energy of some reference wave function, \( E_{\text{ref}} \), which represents the correlation energy or part of it. In the case of the MRCI method, including the popular MRCI(Q) approach [11], the reference wave function can be obtained in the preceding CASCC (complete-active-space self-consistent-field) calculations. In the case of CC approach, the reference wave function can be obtained with the Hartree–Fock method, although other choices are possible. In the CC calculations reported here using the CR–CC(2,3) method [12], abbreviated as CRCC, which can be used to describe single-bond breaking, the energy obtained in CCSD (CC singles and doubles) [13] calculations with the restricted Hartree–Fock (RHF) reference is corrected for triple excitations. Thus, the RHF or CCSD or even the CASCC wave functions can provide reference energies for the extrapolation of the CRCC PES.

Suppose that MRCI(Q) or CRCC or some other correlated calculations using basis sets indexed by \( m - 1 \) and \( m \) are performed \( \{ m \} \) for AVXZ and ACVXZ basis sets and their cc-p(C)VXZ analogs, and is an arbitrary index for basis sets that do not belong to the (aug)cc family. The extrapolated MRCI(Q) or CRCC or other PESs for the \( (m + 1) \)th basis set are obtained as

\[
E_{m+1}(R) = E_{\text{ref}}^{m+1}(R) + \chi_{m+1,m}(R) \Delta E_m(R).
\]

Here, \( E_{\text{ref}}^{m+1}(R) \) is the energy of the reference wave function calculated with the \( (m + 1) \)th basis set, \( \Delta E_m(R) = E_m(R) - E_{\text{ref}}^m(R) \), \( R \) denotes the \( \tau \)-dimensional vector of the nuclear space coordinates defining the PES, and

\[
\chi_{m+1,m}(R) = \frac{1}{S(R)_{m,m-1}} - \frac{1}{S(R)_{m+1,m-1}} [S(R)_{m+1,m} - 1],
\]

with \( S(R)_{m,m-1} = \Delta E_m(R)/\Delta E_{m-1}(R) \). The reference geometry \( R_c \) can be the equilibrium geometry or any other point on the PES. Thus, the MRCI(Q) or CRCC or some other PES \( E_{m+1}(R) \) is extrapolated from the PESS \( E_{m-1}(R) \) and \( E_{m+1}(R) \), the reference energies \( E_{\text{ref}}^{m-1}(R) \), \( E_{\text{ref}}^m(R) \), and \( E_{\text{ref}}^{m+1}(R) \), and a single correlated energy \( E_{m+1}(R_c) \) calculated at the reference geometry \( R_c \). Eq. (1) represents the simplest mathematical expression that one can propose to extrapolate the energy based on scaling the correlation energy or the part of it from basis set \( m \) to basis set \( m + 1 \). In particular, the scaling function \( \chi_{m+1,m}(R) \) in Eq. (1) satisfies the following desirable properties: (i) \( \chi_{m+1,m}(R) \to 1 \) for all values of \( R \) when \( m \to \infty \); (ii) the value of \( \chi_{m+1,m}(R) \) at the reference geometry \( R = R_c \) equals the ratio of \( \Delta E_{m+1}(R_c) \) and \( \Delta E_m(R_c) \). The extrapolation scheme proposed here is analytic, resulting in a smooth PES. Thus, it should be usable for calculating enough information about the PES to fit it to some functional form that could subsequently be utilized in dynamics calculations. One can even envision the development of analytic gradients for the energy resulting from Eq. (1) that could be exploited in direct dynamics calculations. Our scheme does not rely on any empirical
data, which contrasts it with the earlier schemes [14–18] that use readily available experimental atomization or dissociation energies.

Table 1 shows the results obtained from MRCI(Q)/AVXZ calculations employing CASSCF reference (performed using MOLPRO [19]) for the N₂, O₂, and F₂ molecules, which represent three most frequent types of chemical bonds (single in F₂, double in O₂, and triple in N₂). For brevity, only the results obtained for a few geometries are given. The notable feature is that the above extrapolation scheme (in each case, the reference bond length Rₑ was chosen as the experimental equilibrium geometry [20], although similar results are obtained for any other choice) recovers the changes in the total energies over the entire range of bond lengths (0.7 ≤ R/Rₑ ≤ 5.0), when going from the smaller to larger AVXZ basis sets, to within a fraction of a millihartree (0.8–2.5% on average) for ΔE₄₃, and 0.1 millihartree (0.5–0.9% on average) for ΔE₅₄, where ΔEₗₖ = Eₖ − Eₗ is a change in energy, when going from basis set k to basis set l [for simplicity, we have dropped R from E(R)]. In particular, the entire MRCI(Q)/AVSZ potential curves can be predicted with a millihartree or sub-millihartree accuracy through much less expensive MRCI(Q)/AVDZ and MRCI(Q)/AVTZ calculations, as long as one calculation at the MRCI(Q)/AVQZ level and another one at MRCI(Q)/AVSZ level can be performed. Extrapolation to the CBS-limit could be performed using the CBS energy extrapolated at Rₑ and the MRCI(Q)/AVDZ and MRCI(Q)/AVTZ [or MRCI(Q)/AVTZ and MRCI(Q)/AVQZ] curves, but this is not elaborated on here. The key point is that our scheme accurately predicts entire potential curves corresponding to larger basis sets in a smooth manner.

Table 2 summarizes the results for two cuts of the ground-state PES of H₂O, calculated with MRCI(Q) and ACVXZ basis sets. They cover distinct regions of configuration space and energy, and have only one point in common, the equilibrium geometry Rₑ taken from [21]. The first cut (I) corresponds to breaking of one OH bond. The second cut (II) describes the simultaneous dissociation of two OH bonds, which is a much higher energy process than I. In both cuts, the angle ∠HOH was fixed at its equilibrium value (a and b label the H atoms). As the results show, the predictive capability of our extrapolation scheme is as good as in the cases of N₂, O₂, and F₂. The results in Table 2 have been obtained by fixing the reference geometry Rₑ at equilibrium, which is a common point in both cuts, but they are almost identical when a point that belongs to only one of the two cuts (e.g., RᵢH₋ = RᵢH₊ = 5Rₑ, which belongs to cut II only) is used as the reference geometry for extrapolation. This suggests that a global high-level PES can be generated for H₂O from only one or two MRCI(Q) energies obtained with the largest X-tuple basis sets and PESs calculated with smaller

**Table 1**

<table>
<thead>
<tr>
<th>System</th>
<th>R/Rₑ</th>
<th>E₄ᵃ</th>
<th>ΔE₄₃ᵇ</th>
<th>ε₄ᶜ</th>
<th>E₅ᵃ</th>
<th>ΔE₅₄ᵇ</th>
<th>ε₅ᵈ</th>
<th>ε₅ᵉ</th>
</tr>
</thead>
<tbody>
<tr>
<td>N₂</td>
<td>0.8</td>
<td>−109.173214</td>
<td>−32.707</td>
<td>−0.556</td>
<td>−109.182846</td>
<td>−9.632</td>
<td>−0.604</td>
<td>−0.033</td>
</tr>
<tr>
<td></td>
<td>1.3</td>
<td>−109.276761</td>
<td>−21.967</td>
<td>−0.048</td>
<td>−109.283649</td>
<td>−6.888</td>
<td>0.003</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td>1.8</td>
<td>−109.087088</td>
<td>−17.980</td>
<td>−0.500</td>
<td>−109.092710</td>
<td>−5.622</td>
<td>−0.458</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>4.0</td>
<td>−109.050342</td>
<td>−16.604</td>
<td>−0.658</td>
<td>−109.055528</td>
<td>−5.186</td>
<td>−0.590</td>
<td>0.078</td>
</tr>
<tr>
<td>rmsdᵇ</td>
<td>0.430</td>
<td></td>
<td></td>
<td>0.443</td>
<td>0.120</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O₂</td>
<td>0.8</td>
<td>−150.018095</td>
<td>−41.344</td>
<td>0.114</td>
<td>−150.031102</td>
<td>−13.007</td>
<td>0.047</td>
<td>−0.069</td>
</tr>
<tr>
<td></td>
<td>1.3</td>
<td>−150.096957</td>
<td>−33.501</td>
<td>−0.458</td>
<td>−150.108083</td>
<td>−11.126</td>
<td>−0.417</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>1.8</td>
<td>−149.994763</td>
<td>−31.993</td>
<td>−1.137</td>
<td>−150.005352</td>
<td>−10.590</td>
<td>−1.057</td>
<td>0.104</td>
</tr>
<tr>
<td></td>
<td>4.0</td>
<td>−149.986743</td>
<td>−31.857</td>
<td>−1.302</td>
<td>−149.997359</td>
<td>−10.616</td>
<td>−1.191</td>
<td>0.140</td>
</tr>
<tr>
<td>rmsdᵇ</td>
<td>0.841</td>
<td></td>
<td></td>
<td>0.777</td>
<td>0.086</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F₂</td>
<td>0.8</td>
<td>−199.268096</td>
<td>−54.217</td>
<td>−0.089</td>
<td>−199.286432</td>
<td>−18.336</td>
<td>−0.191</td>
<td>−0.100</td>
</tr>
<tr>
<td></td>
<td>1.3</td>
<td>−199.324432</td>
<td>−49.677</td>
<td>−0.183</td>
<td>−199.341478</td>
<td>−17.046</td>
<td>−0.157</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>1.8</td>
<td>−199.300262</td>
<td>−49.089</td>
<td>−0.561</td>
<td>−199.317259</td>
<td>−16.997</td>
<td>−0.467</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>4.0</td>
<td>−199.299614</td>
<td>−49.307</td>
<td>−0.544</td>
<td>−199.316695</td>
<td>−17.080</td>
<td>−0.448</td>
<td>0.111</td>
</tr>
<tr>
<td>rmsdᵇ</td>
<td>0.379</td>
<td></td>
<td></td>
<td>0.332</td>
<td>0.087</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ᵃ Total energies, in hartree, calculated at MRCI(Q)/AVQZ (E₄) and MRCI(Q)/AVSZ (E₅) levels.
ᵇ Differences, in millihartree, between the actual MRCI(Q)/AVQZ and MRCI(Q)/AVTZ (ΔE₄₃) or MRCI(Q)/AVSZ and MRCI(Q)/AVQZ (ΔE₅₄) energies.
ᶜ Deviation, in millihartree, between the actual MRCI(Q)/AVQZ energy and the extrapolated one from MRCI(Q)/AVTZ and MRCI(Q)/AVDZ calculations.
ᵈ Deviation, in millihartree, between the actual MRCI(Q)/AVSZ energy and the extrapolated one from MRCI(Q)/AVTZ and MRCI(Q)/AVDZ calculations.
ᵉ Deviation, in millihartree, between the actual MRCI(Q)/AVQZ energy and the extrapolated one from MRCI(Q)/AVQZ and MRCI(Q)/AVTZ calculations.
ᶠ Experimental equilibrium geometries [20]. Rₑ, for N₂, O₂ and F₂ are 1.09768, 1.20752, and 1.41193 Å, respectively. All calculations have been performed for 0.7 ≤ R/Rₑ ≤ 5.0.
ᵍ Root-mean-square deviation, in millihartree, for the 29 calculated points of each of the N₂, O₂, and F₂ systems.
Table 2

<table>
<thead>
<tr>
<th>PES cut</th>
<th>${R/R_e}^a$</th>
<th>$E^b$</th>
<th>$\Delta E_{43}^c$</th>
<th>$\epsilon_4^d$</th>
<th>$E^5_e$</th>
<th>$\Delta E_{45}^f$</th>
<th>$\epsilon_5^g$</th>
<th>$\epsilon_5^i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HO–H^h</td>
<td>0.75</td>
<td>-76.319568</td>
<td>-25.940</td>
<td>0.100</td>
<td>-76.327540</td>
<td>-7.973</td>
<td>0.148</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>1.125</td>
<td>-76.412167</td>
<td>-24.143</td>
<td>0.000</td>
<td>-76.419533</td>
<td>-7.366</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>-76.260564</td>
<td>-22.852</td>
<td>-0.250</td>
<td>-76.267564</td>
<td>-7.000</td>
<td>-0.231</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>-76.223480</td>
<td>-22.472</td>
<td>-0.476</td>
<td>-76.230392</td>
<td>-6.912</td>
<td>-0.449</td>
<td>0.037</td>
</tr>
<tr>
<td>rmsd^h</td>
<td></td>
<td>0.299</td>
<td>0.284</td>
<td>0.028</td>
<td>0.284</td>
<td>0.028</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td>H−O−H^h</td>
<td>0.75</td>
<td>-76.217614</td>
<td>-27.189</td>
<td>0.561</td>
<td>-76.226078</td>
<td>-8.465</td>
<td>0.661</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td>1.05</td>
<td>-76.422840</td>
<td>-24.627</td>
<td>0.000</td>
<td>-76.430320</td>
<td>-7.480</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>-76.114238</td>
<td>-20.899</td>
<td>-0.553</td>
<td>-76.120718</td>
<td>-6.480</td>
<td>-0.509</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>5.0</td>
<td>-76.054710</td>
<td>-19.984</td>
<td>-0.971</td>
<td>-76.060973</td>
<td>-6.263</td>
<td>-0.901</td>
<td>0.090</td>
</tr>
<tr>
<td>rmsd^h</td>
<td></td>
<td>0.604</td>
<td>0.575</td>
<td>0.058</td>
<td>0.575</td>
<td>0.058</td>
<td>0.058</td>
<td></td>
</tr>
</tbody>
</table>

---

*a* $R_e = 0.95785$ Å stands for the OH bond length in the equilibrium C$_{2v}$ structure of the water molecule [21].

*b* Total energies, in hartree, calculated at MRCI(Q)/ACVQZ ($E_4$) and MRCI(Q)/ACVTZ ($E_5$) levels.

*c* Differences, in millihartree, between the actual MRCI(Q)/ACVQZ and MRCI(Q)/ACVTZ ($\Delta E_{43}$) or MRCI(Q)/ACVQZ and MRCI(Q)/ACVQZ ($\Delta E_{45}$) energies.

*d* Deviation, in millihartree, between the actual MRCI(Q)/ACVQZ energy and the extrapolated one from MRCI(Q)/ACVDZ and MRCI(Q)/ACVTZ calculations.

*e* Deviation, in millihartree, between the actual MRCI(Q)/ACVTZ energy and the extrapolated one from MRCI(Q)/ACVTZ and MRCI(Q)/ACVDZ calculations.

*f* Deviation, in millihartree, between the actual MRCI(Q)/ACVQZ energy and the extrapolated one from MRCI(Q)/ACVQZ and MRCI(Q)/ACVTZ calculations.

*g* Breaking of a single OH bond (cut I), while keeping the angle $\angle$HOH fixed at the equilibrium value of 104.51°.

*h* Root-mean-square deviation, in millihartree, for all calculated points from a given PES cut (11 points for cut I, and 12 points for cut II).

*i* Symmetric breaking of the two OH bonds (cut II), while keeping the angle $\angle$HOH fixed at the equilibrium value of 104.51°.

---

basis sets. Fig. 1 highlights the predictive capability of the method by comparing the actual changes in the total energies, when going from one basis set to another, with those obtained from extrapolation. A direct comparison of the calculated and predicted $\Delta E_{kl}$ values resulting from MRCI(Q) calculations is shown in panels (a) for O$_2$ and (b) for H$_2$O, while the deviations between the predicted ($E_k^p$) and actual ($E_k$) MRCI(Q) energies, $\epsilon_k = E_k^p - E_k$, are shown in Tables 1 and 2. The agreement between the predicted and actual values is very good, irrespective of nuclear geometry.

For wide applicability, the method should also work with basis sets that are not as systematic as AVXZ and ACVXZ. In many applications, particularly when larger systems are studied, Pople-type basis sets are often the most affordable choice. To test the robustness of the proposed extrapolation scheme, the following four basis sets of the Pople-type have been used for O$_2$: 6-311G (basis set index $m = 2$, 26 functions), 6-311+G(d) ($m = 3$, 44 functions), 6-311+G(2d) ($m = 4$, 54 functions), and 6-311+G(3d) ($m = 5$, 64 functions). As shown in Fig. 1c, the curves representing the actual energy changes $\Delta E_{kl}$ are highly non-monotonic, and yet the errors in the $\Delta E_{kl}$ values resulting from extrapolation are small. In particular, the proposed extrapolation scheme correctly reproduces the shapes of the actual $\Delta E_{kl}$ curves.

The procedure based on Eqs. (1) and (2) works equally well, when methods other than MRCI(Q), such as, e.g., CRCC, are used to calculate total energies. This is illustrated in Fig. 1d and Table 3, where the results of CRCC/ACVXZ calculations ($X = 2$–$4$), performed using the CC routines [12,22] in GAMESS [23], for the breaking of a single OH bond in the water molecule are summarized. Three schemes have been adopted to extrapolate the PES. The first scheme employs the RHF wave function to provide the reference energy $E_{ref}$, and hence is based on the extrapolation of total correlation (non-dynamical plus dynamical) energy at the equilibrium geometry $R_e$. In the second scheme, CCSD provides the reference energy $E_{ref}$, and the extrapolation is performed for the remaining part of the correlation energy that defines the triples correction of CRCC. The third scheme uses the difference between the CRCC and CASSCF energies to define $\Delta E_{ref}$ energies in Eqs. (1) and (2). In all three cases, the substantial changes in the total energies $\Delta E_{43}$, when going from ACVQZ to ACVQZ basis sets are predicted with high accuracy, typically within 1–3% of the value of $\Delta E_{43}$ obtained in the corresponding CRCC/ACVTZ and CRCC/ACVQZ calculations.

To conclude, an ab initio scheme has been proposed that predicts the global PES corresponding to a calculation with a large basis set from small basis set calculations. The crucial idea is to scale the correlation energy or part of it using the scaling factor obtained at a single geometry. The ability to predict the smooth PES corresponding to larger basis sets from smaller basis set calculations with (sub)millihartree accuracy is the key to successfully extrapolate PESs to the CBS-limit. If combined with the recently developed method of correlation energy extrapolation by intrinsic scaling [24], which enables one to achieve a nearly exact solution of the non-relativistic electronic Schrödinger equation in a basis set, the method proposed in this work may
Fig. 1. A comparison of the actual and predicted changes in total energies, when going from basis set \( l \) to basis set \( k \), \( \Delta E_{kl} = E_k - E_l \): (a) MRCI(Q)/AVXZ results for \( \text{O}_2 \); (b) MRCI(Q)/ACVXZ results for \( \text{H}_2\text{O} \) (cuts I and II); (c) MRCI(Q) results for \( \text{O}_2 \) using Pople-type basis sets; (d) CRCC/ACVXZ results for single-bond breaking in \( \text{H}_2\text{O} \) (cut I). In panels (a), (b), and (d), \( k \) and \( l \) are the cardinal numbers of the corresponding AVXZ (a) and ACVXZ ((b) and (d)) basis sets (\( k = 2 \) for AVDZ and ACVDZ, \( k = 3 \) for AVTZ and ACVTZ, etc). In panel (c), \( k \) and \( l \) label Pople-type basis sets as described in the text. The actual values of \( \Delta E_{kl} \) are represented by circles, triangles, and squares. The predicted \( \Delta E_{kl} \) values are represented by lines. The solid lines in panels (a)–(c) represent the extrapolations from the \( (k - 2) \) and \( (k - 1) \) bases to the basis set \( k \) (\( k = 4,5 \)) or, in the case of \( \Delta E_{53} \), the result of two consecutive extrapolations, from the actual calculations with basis sets 2 and 3 to 4 and from the actual calculations using basis sets 3 and 4 to 5, i.e. \( \Delta E_{53} = \Delta E_{43} + \Delta E_{54} \), where \( \Delta E_{43} \) and \( \Delta E_{54} \) are the predicted energy changes. The dashed lines representing the predicted \( \Delta E_{53} \) values in panels (a)–(c) correspond to the extrapolation from the actual calculations using basis sets 2 and 3 to 4, followed by the extrapolation from the actual calculations using basis set 3 and the predicted result from 4 to 5. The curves in blue in panel (b) for \( \text{H}_2\text{O} \) refer to single-bond breaking (cut I), while those in red refer to double bond breaking (cut II); for both cuts, \( \angle \text{HOH} \) has been fixed at the equilibrium value. The lines in panel (d) have the following meaning: solid line, predicted \( \Delta E_{43} \) from CRCC/ACVXZ calculations using the RHF energy as \( E^{\text{ref}} \); dashed line, predicted \( \Delta E_{43} \) using the CASSCF energy as \( E^{\text{ref}} \); dashed-dotted line, predicted \( \Delta E_{43} \) using the CCSD energy as \( E^{\text{ref}} \).

Table 3
Energies for single-bond breaking in \( \text{H}_2\text{O} \) from CRCC/ACVXZ calculations at selected geometries from cut I

<table>
<thead>
<tr>
<th>( R/R_e ) a</th>
<th>( E_b^b )</th>
<th>( \Delta E_{kl}^c )</th>
<th>( \varepsilon_{\text{RHF}}^d )</th>
<th>( \varepsilon_{\text{CASSCF}}^e )</th>
<th>( \varepsilon_{\text{CCSD}}^f )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.25</td>
<td>-76.390403</td>
<td>-23.964</td>
<td>0.148</td>
<td>-0.034</td>
<td>0.003</td>
</tr>
<tr>
<td>1.5</td>
<td>-76.377124</td>
<td>-23.597</td>
<td>0.301</td>
<td>-0.002</td>
<td>0.021</td>
</tr>
<tr>
<td>2.0</td>
<td>-76.260215</td>
<td>-22.861</td>
<td>0.322</td>
<td>-0.327</td>
<td>0.059</td>
</tr>
<tr>
<td>3.0</td>
<td>-76.223179</td>
<td>-22.141</td>
<td>0.876</td>
<td>-0.410</td>
<td>0.328</td>
</tr>
<tr>
<td>5.0</td>
<td>-76.221884</td>
<td>-22.003</td>
<td>1.031</td>
<td>-0.486</td>
<td>0.395</td>
</tr>
<tr>
<td>rmsd f</td>
<td></td>
<td>0.636</td>
<td>0.307</td>
<td>0.232</td>
<td></td>
</tr>
</tbody>
</table>

a \( R_e = 0.95785 \) Å stands for the OH bond length in the equilibrium \( C_2 \) structure of the water molecule [21].
b Total energy, in hartree, calculated at CRCC/ACVQZ level.
c Difference, in millihartree, between the actual CRCC/ACVQZ and CRCC/ACVTZ energies.
d Deviation, in millihartree, between the actual CRCC/ACVQZ energy and the extrapolated one from CRCC/ACVTZ and CRCC/ACVDZ energies using the RHF wave function to provide the reference energy for extrapolation.
e Deviation, in millihartree, between the actual CRCC/ACVQZ energy and the extrapolated one from CRCC/ACVTZ and CRCC/ACVDZ energies using the CASSCF wave function to provide the reference energy for extrapolation.
f Deviation, in millihartree, between the actual CRCC/ACVQZ energy and the extrapolated one from CRCC/ACVTZ and CRCC/ACVDZ energies using the CCSD wave function to provide the reference energy for extrapolation.
g Root-mean-square deviation, in millihartree, for the 9 calculated geometries.

enable one to obtain the virtually exact PESs from first principles.

Acknowledgements

Supported by Fundação para a Ciência e Tecnologia, Portugal, and the U.S. Department of Energy.

References