Theoret. Chim. Acta (Berl.) 58, 81–96 (1981)

THEORETICA CHIMICA ACTA

Original Investigations

Accuracy of Energy Extrapolation in Multireference Configuration Interaction Calculations

Charles F. Jackels

Battelle, Columbus Laboratories, Columbus, OH 43201, USA
and
Department of Chemistry, Wake Forest University, Winston-Salem, NC 27109, USA*

Isaiah Shavitt

Batelle, Columbus Laboratories, Columbus, OH 43201, USA
and
Department of Chemistry, The Ohio State University, Columbus, OH 43210, USA

The accuracy of extrapolation procedures in conjunction with energy-based configuration selection in CI calculations is examined. The normally high accuracy of such extrapolation can deteriorate in multireference CI calculations when configuration functions of low weight are included in the root (reference) set. This is due to the inadequacy of second-order energy contribution estimates for the very large number of discarded low-contribution functions generated as single and double excitations from the minor members of the root set. The problem may be overcome by increasing the number of configurations included in the zero-order function used for the energy contribution estimation process. Illustrative results are presented for excited states of the H_2O molecule and the H_2O^+ ion.

Key words: Configuration interaction – Selection of configurations – Energy extrapolation.

1. Introduction

The configuration interaction (CI) method [1] has become a standard tool for the evaluation of correlated electronic wave functions and energies of molecular

---

* Present address

0040-5744/81/0058/0081/$03.20

systems within a variational framework. A major advantage of this method is its generality, enabling consistent treatments of different electronic states and different regions of potential energy surfaces. Its principal limitation is the slow covergence of the CI wave function expansion, requiring a very large number of terms (often as many as 10^4–10^5 or more) for satisfactory accuracy, even for fairly small molecules.

This slow convergence has led to the increasing use of configuration selection techniques [1–11], in which individual terms (configuration functions, or CF's) are selected for inclusion in the wave function expansion by an automatic procedure on the basis of appropriate estimates of the magnitude of their expected contributions to the solution. An important contribution to the effectiveness of such methods has been the introduction of an extrapolation technique by Buenker and Peyerimhoff [9, 12, 13]. This technique makes it possible to account for the effect on the energy of many more configuration functions than can actually be included in the *CI* expansion, and is thus a very useful tool for increasing the accuracy and reliability of quantum chemical calculations.

The present contribution describes some tests of the accuracy of the Buenker–Peyerimhoff (BP) extrapolation procedure, and discusses some cases in which such extrapolation is somewhat less accurate than otherwise expected. It is not intended as a criticism of the BP procedure, which has been found to be extremely useful, but merely to note certain situations which require extra caution if highly accurate results are to be obtained.

2. Configuration Selection

In general, the selection of configuration functions (CF's) for inclusion in a CI expansion consists of two stages [1]. In the first stage, which may be called preselection, a certain class of CF's is chosen as potentially significant, and in the second, which may be called screening or individual selection, the members of that class are tested individually according to some criterion, and are either accepted or rejected for the final CF list.

The most typical and generally useful form of preselection is based on the choice of a set of "root" or "reference" configuration functions [1, 6, 9–11, 13], with the candidates for final selection being all single and double excitations (SD) relative to the root set, i.e. all CF's which differ by at most two orbitals from at least one of the root CF's. The root set consists of one or more CF's which are judged to be the principal contributors to the wave function, and their choice is often refined on the basis of subsequent tests [10]. Particularly important in calculations of potential energy surfaces and in simultaneous calculations of several states of the same symmetry and spin species [10, 11] is to choose a root set which provides a balanced zero-order description of the several states under consideration and of the different regions of the potential surface, including proper description of dissociation processes.

The final screening (individual selection) can be carried out in several ways [1–11], generally depending on some estimate of the energy contribution of each CF in the final CI wave function, this estimate being based in most cases on perturbation theory (PT) ideas. Given some normalized zero-order wave function constructed from a "primary set" of CF's (which often is, but need not be, the same as the root set),

$$\Psi_0=\sum_{i=1}^{k} a_i\Phi_i, \tag{1}$$

a perturbation theory estimate of the energy contribution ε_j of a CF Φ_j $(j>k)$ can be obtained from

$$\varepsilon_j=|H_{j0}|^2/(E_0-E_j), \tag{2}$$

where

$$\begin{aligned} &H_{j0}=\langle\Phi_j|\hat{\boldsymbol{H}}|\Psi_0\rangle=\sum_{i=1}^{k} H_{ji}a_i \qquad (j>k),\\ &H_{ij}=\langle\Phi_i|\hat{\boldsymbol{H}}|\Phi_j\rangle \qquad (i,j=1,2,\ldots),\\ &E_j=H_{jj},\\ &E_0=\langle\Psi_0|\hat{\boldsymbol{H}}|\Psi_0\rangle=\sum_{i=1}^{k}\sum_{j=1}^{k} a_i^{*}H_{ij}a_j, \end{aligned} \tag{3}$$

and it has been assumed that the CF's are orthonormal,

$$\langle\Phi_i|\Phi_j\rangle=\delta_{ij} \qquad (i,j=1,2,\ldots).$$

The coefficients a_i and the zero-order energy E_0 can be obtained as an eigenvector and corresponding eigenvalue of the primary Hamiltonian matrix of order k,

$$\sum_{j=1}^{k} H_{ij}a_j=E_0a_i \qquad (i=1,2,\ldots,k). \tag{4}$$

This procedure has been referred to as "A_k" by Gershgorn and Shavitt [1, 5], and is essentially the same as that used by several other investigators to obtain energy contribution estimates [3, 6, 10, 11]. In calculations involving more than one state of a given symmetry and spin species, this procedure provides energy contribution estimates for each state separately by using an appropriate eigenvector of Eq. (4) as the corresponding zero-order function for each state [1, 10, 11].

Buenker and Peyerimhoff [9] have employed a somewhat different scheme for estimating energy contributions, in which a separate secular equation of order $k+1$ is solved for each secondary CF. This equation involves the k primary CF's plus the one secondary function under consideration, and the energy contribution is taken as the difference between the eigenvalue obtained and the corresponding zero-order eigenvalue E_0. There is little extra work involved in this procedure (compared to the A_k scheme), since the primary block (of order k) has been prediagonalized at the outset, but the results are not expected to differ significantly from those of the A_k procedure described above. (The two methods give identical results for $k = 1$.)

The actual selection is usually determined by a pre-chosen threshold value T, with each CF being retained or rejected depending on whether the magnitude of its estimated energy contribution exceeds, or does not exceed, the value of T, respectively. (For a different approach see Raffenetti et al. [11].)

In the present work, as in some of the work of Davidson and co-workers [7, 14] and of Buenker and Peyerimhoff, selection is carried out by "spin blocks", with an entire spin block being selected or rejected as a unit depending on its *combined* energy contribution. The term spin block refers to a set of CF's which are constructed from the same orbital product ("space configuration"), multiplied by different members of a set of spin eigenfunctions and antisymmetrized. The spin eigenfunctions are all those which can be combined with the given orbital product to produce a set of orthonormal and antisymmetric CF's of the desired spin species.

Labeling each CF $\Phi_{s\nu}$ by two indices, s and ν, which identify the orbital product and spin function, respectively, the combined spin block energy contribution is estimated in this work as [7]

$$\varepsilon_s = \sum_\nu |H_{s\nu,0}|^2/(E_0 - \bar{E}_s) \tag{5}$$

where $\bar{E}_s$ is the average of the energies $E_{s\nu}$ over the spin block. A spin block is retained if its combined energy contribution estimate ε_s exceeds the threshold T.

Selection by spin block is not quite optimal, since it often includes complete spin blocks of which only a fraction of the members have significant energy contributions [8, 15], but this is not expected to affect the conclusions to be described below. This approach was used because it was the most feasible for the set of CI programs (developed by E. R. Davidson and co-workers at the University of Washington) which were employed in this work.

3. Extrapolation

Rather than simply truncating the CI expansion on the basis of a fixed selection threshold T, Buenker and Peyerimhoff [9, 12] (BP) have proposed extrapolating the CI energy to $T = 0$ on the basis of a series of truncations at different levels, T_1, T_2, ... This extrapolation is facilatated by considering the family of curves

$$E_\lambda(T) = E(T) + \lambda\,\Delta E^r(T), \tag{6}$$

where

$$\Delta E^r(T) = \sum_{s\in\{r(T)\}} \varepsilon_s, \tag{7}$$

the sum being over the set of all CF's (or spin blocks) $\{r(T)\}$ rejected at threshold level T. If $\Delta E^r(T)$ were to provide an accurate estimate of the truncation error (the true total energy contribution of the rejected CF's), then the $\lambda = 1$ curve would be a straight horizontal line at $E(0)$, the correct zero-threshold energy. In actual practice this line is not quite horizontal, and usually not quite straight, but a plot of $E_\lambda(T)$ curves for several values of λ facilitates a fairly accurate graphical extrapolation [12] to $T = 0$.

An alternative graphical extrapolation can however be performed, using a single curve, by plotting $E(T)$ *vs.* $\Delta E^r(T)$. Again, if $\Delta E^r(T)$ is an accurate estimate of the truncation error, then the proposed plot would be a straight line with unit slope. In actual fact the line is not quite straight and deviates from unit slope, but extrapolation to $T = 0$ is fairly easy if no pathological behavior is found near the $T = 0$ limit (as discussed further below).

Another extrapolation method has been employed by Langhoff and Davidson [14], based on the assumption that the fractional error made in estimating the energy contribution of the rejected CF's by $\Delta E^r(T)$ is the same as in the corresponding estimate

$$\Delta E^a = \sum_{s \in \{a(T)\}} \varepsilon_s \tag{8}$$

for a set $\{a(T)\}$ of accepted CF's. The resulting extrapolated energy $E(\rightarrow 0)$ is

$$E(\rightarrow 0) = E(T) + [E(T) - E_0]\Delta E^r(T)/\Delta E^a(T), \tag{9}$$

E_0 being the reference energy. This is not unrelated to the BP extrapolation approach, and is essentially equivalent to the assumption that for some value of λ, given by

$$\lambda = [E(T) - E_0]/\Delta E^a(T), \tag{10}$$

The $E_\lambda(T)$ curve is a horizontal straight line. However, the BP approach based on Eq. (6) is more general, since it allows for some variation in the ratio on the r.h.s. of Eq. (10), and takes that variation into account in the extrapolation.

Buenker and Peyerimhoff indicate that by using a minimum threshold of $T = 20$ μhartree, the value of $E(0)$ can be estimated in their approach with an uncertainty of 0.5 m hartree or less [16]. Somewhat larger errors are expected for larger systems [17].

4. Calculations

4.1. Extrapolation tests on H_2O

In order to verify that the effects described in the next subsection are not a result of any differences in the details of the extrapolation procedure compared to the procedures used by BP, the calculations on the ${}^3B_2$ excited state of H_2O reported by Buenker and Peyerimhoff [12, 16] were repeated with the selection procedures based on Eq. (5). An additional objective in repeating these calculations was to test the extrapolation based on a plot of $E(T)$ *vs.* $\Delta E^r(T)$ described in the previous section.

In these calculations, the molecular geometry and Gaussian lobe basis set were chosen to be the same as those in Ref. [16]. The orbitals used in the CI calculations were the canonical SCF orbitals computed for the $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^1(1b_1)^2(2b_2)^1(^3B_2)$ configuration of H_2O. The root and primary set consisted of just the SCF configuration ($k = 1$), and the preselected CF list consisted of all single and double excitations which kept the $(1a_1)^2$ core orbital doubly occupied and its high-energy $11a_1$ complement unoccupied, just as in the BP work [12, 16]. This produced a list of 2151 CF's (for $T = 0$). The BP-type extrapolation plots obtained from calculations at several threshold values are shown in Fig. 1. They are fairly smooth, and can be extrapolated from $T = 20$ or 10 μhartree with an accuracy of 0.1–0.2 mhartree. The extrapolated result based on $T = 40$, 20, 10 or 20, 10, 5 μhartree is $E(\rightarrow 0) = -75.7856 \pm 0.0001$ hartree, compared with the exact $T(0) = -75.785634$ hartree. (This $T(0)$ value is about 0.9 mhartree lower than the BP value, probably because of some minor differences in the basis set data, such as in the Gaussian lobe representation of the d functions.)

While Fig. 1 is qualitatively similar to the corresponding Fig. 3 of BP [12], there are quantitative differences which are traceable to the use of spin-block selection in the present work. As a result, the size of the secular equation solved at each T value (other than $T = 0$) is larger than that of BP, and the $\lambda = 0$ curve is less steep. The value of λ which produces a nearly horizontal curve at small T is $\lambda = 0.97$, compared to 0.838 for BP, indicating that the retention of entire spin blocks on the basis of the energy contribution formula of Eq. (5) leads to truncation error

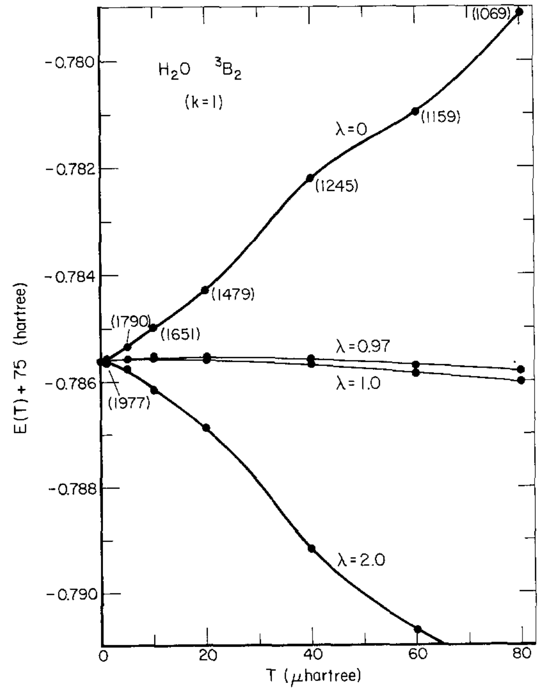

Fig. 1. Extrapolation curves $E_\lambda(T)$ for the $^3B_2$ state of the water molecule. Both the root set and primary set consisted of just the SCF configuration. The numbers in parentheses are the dimensions of the secular equations solved at the corresponding threshold values T. The unselected expansion contained 2151 terms

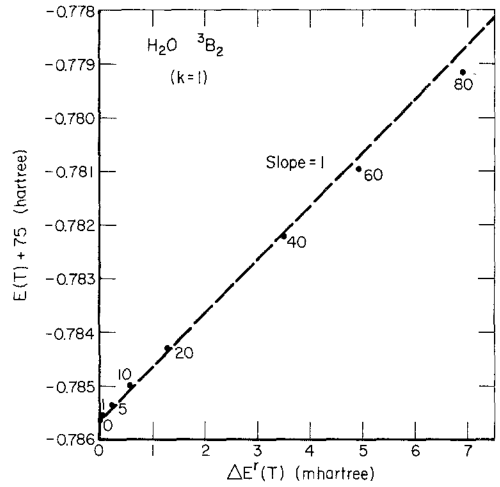

Fig. 2. Alternative extrapolation plot $E(T)$ *vs* $\Delta E^r(T)$ for the example of Fig. 1. The numbers next to the dots are the corresponding threshold values T (in μhartrees)

estimates $\Delta E^r(T)$ which were closer (at least in this example) to the true truncation error than in the BP procedure. This was done, however, at the cost of solving larger secular equations. The principal conclusion from the comparison is that the extrapolation procedure used here, based on perturbation theory energy contribution estimates and entire spin block selection, gives results quite similar to those of the BP procedure, and that the difficulties in the extrapolation behavior described in the next subsection cannot be due to the differences in these procedures.

The alternative extrapolation plot of $E(T)$ *vs.* $\Delta E^r(T)$ is shown for this example in Fig. 2. The points are seen to lie fairly close to the dashed line, which is the unit slope straight line through $E(0)$. A least squares line fitted to the $T = 40, 20, 10$ or 20, 10, 5 μhartree data gives an extrapolated result with an error less than 0.1 mhartree, quite similar to the case of the BP-type plots of Fig. 1.

The principal difference between the calculations described above and those to be described in the next subsection is in the size of the root set. To verify the assumption that this aspect of the calculation is related to the problems discussed below, another set of calculations was carried out for the $^3B_2$ state of H_2O using two configuration functions ($k = 2$) in the root and primary set. The second configuration (coefficient −0.04 in the CI expansion) was $(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^1(2b_2)^1(3b_1)^2$. In this case the preselected list contained 4017 CF's, and yielded an unselected energy of $E(0) = -75.786620$ hartree. The extrapolation curves for this case are shown in Fig. 3, and are seen to have a fairly steep drop in $E(T)$ close to $T = 0$, making extrapolation less accurate than in the above example. This is indeed symptomatic of the problems described below for H_2O^+.

4.2. Multireference extrapolations for H_2O^+

The problems described in this subsection were encountered in a study [18] on potential energy surface crossings in H_2O^+. The two surfaces involved are those

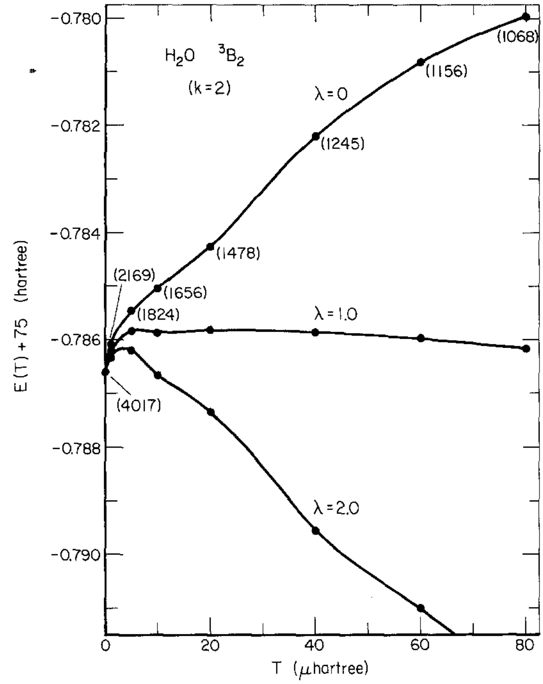

Fig. 3. Same as Fig. 1, but with two functions in the root and primary sets. Note the large number of CF's with energy contributions smaller than 1 μhartree

for the 1^2A_1 and 1^2B_2 states, which interact and mix upon asymmetric distortion of the ion. The calculations described here are limited to the symmetric (C_{2v}) geometries of H_2O^+, but the choices of orbitals and root sets were dictated by the needs of the surface crossing study.

The basis set used was a Dunning [19] double zeta contraction of the Huzinaga [20] ($9s5p$) basis for oxygen and ($4s$) basis for hydrogen (with a scale factor [19] of $\eta = 1.275$). A full set (five) of d polarization functions on oxygen (with exponent 0.85) and a set of p functions on each hydrogen atom (exponent 1.0) were added [21]. An initial set of orbitals was obtained by solving the open shell SCF equations by Davidson's method [22] for either the $1^2A_1(1a_1)^2(2a_1)^2(1b_2)^2(3a_1)^1(1b_1)^2$ or the 1^2B_2 $(1a_1)^2(2a_1)^2(1b_2)^1(3a_1)^2(1b_1)^2$ state (the choice of which SCF configuration to use was based on considerations having to do with the overall potential surface study [18]). The orbitals were obtained in Davidson's "internally consistent" SCF form [23]. A small CI calculation (about 400–500 CF's) was then carried out in the B_k approximation [5] (i.e. ignoring all off-diagonal matrix elements which do not involve any of the primary set configurations), in order to obtain approximate natural orbitals for the same state as that of the SCF calculation [24]. The inner shell $1a_1$ orbital was constrained to contain two electrons in all the subsequent treatments, and its high-energy counterpart $12a_1$ was discarded. Using these orbitals and the root sets of 6 (for $^2A_1$) and 9 (for $^2B_2$) CF's listed in Table 1, preselected configuration lists were generated for the two states, containing 6182 and 8186 CF's, respectively.

Table 1. Root configurations and dimensions of the preselected configuration lists for the 1^2A_1 and 1^2B_2 states of the H_2O^+ ion

		Preselected list dimension	
	Root configurations[a]	Spatial configurations	Spin-adapted CF's
1.	State 1^2A_1 $\cdots(1b_2)^2(3a_1)^1(1b_1)^2$ $\cdots(1b_2)^1(3a_1)^1(1b_1)^2(2b_2)^1$ $\cdots(1b_2)^2(3a_1)^2(4a_1)^1$ $\cdots(3a_1)^2(1b_1)^2(4a_1)^1$ $\cdots(1b_2)^2(3a_1)^1(4a_1)^2$	2085	6182
2.	State 1^2B_2 $\cdots(1b_2)^1(3a_1)^2(1b_1)^2$ $\cdots(1b_2)^1(1b_1)^2(4a_1)^2$ $\cdots(1b_2)^1(3a_1)^1(1b_1)^2(4a_1)^1$ $\cdots(1b_2)^1(3a_1)^1(1b_1)^1(4a_1)^1(2b_1)^1$	1776	8186

[a] The $1a_1$ and $2a_1$ orbitals are doubly occupied in all the root configurations, and are indicated by the ellipsis ($\cdots$). The five and four "spatial" root configurations for the 1^2A_1 and 1^2B_2 states give rise to 6 and 9 CF's, respectively, when all appropriate spin couplings are applied.

Selections were then carried out at a range of thresholds, using the energy contribution estimates of Eq. (5), for the electronic states and geometries listed in Table 2. The primary set used in the selection process (see Eqs. (1, 3, 5)) in these calculations was identical to the root set. Both selected and unselected ($T = 0$) CI calculations were carried out, and graphical extrapolation, using the BP technique with thresholds of 20, 10, and 5 μhartree, was also done. The zero threshold results and the extrapolation errors are listed in Table 2. It is seen that even

Table 2. Extrapolation errors for several geometries and electronic states for the H_2O^+ ion (in C_{2v} symmetry)[a]

Case	HOH angle (degrees)	OH distance (bohrs)	State	NO's from state	$E(0)$ (hartrees)	Extrapolation error (mhartrees)
1	60.0	2.20	$^2A_1$	$^2B_2$	−75.5560	+1.5
2	55.2	2.15	$^2B_2$	$^2B_2$	−75.6371	+1.8
3	104.52	1.8088	$^2A_1$	$^2B_2$	−75.7139	+1.1
4	7.65	15.00	$^2B_2$	$^2B_2$	−75.5078	+1.7
5	170.0	1.875	$^2A_1$	$^2A_1$	−75.7648	+1.5

[a] The root and primary set contained 6 (for $^2A_1$) and 9 (for $^2B_2$) CF's (Table 1). The orbitals were approximate natural orbitals (NO's) obtained, as described in the text, for the states indicated. The $E(0)$ results quoted for $^2A_1$ were actually obtained using 2000 out of the 2085 spatial configurations (corresponding to a threshold of $T \approx 10^{-9}$ hartree) due to program limitations. The extrapolation error is defined as $E(\rightarrow 0) - E(0)$, where $E(\rightarrow 0)$ is the result of BP-type graphical extrapolation based on $T = 20, 10, 5$ μhartree.

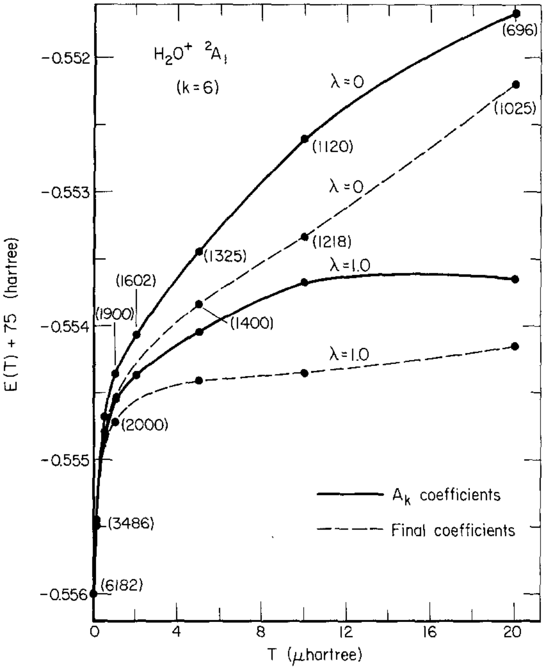

Fig. 4. Extrapolation curves $E_\lambda(T)$ for the $^2A_1$ state of the H_2O^+ ion with six functions in the reference and root sets. The full lines correspond to the usual A_k procedure (case A of Table 3), while the dashed lines correspond to the use of the renormalized final CI vector components in the zero-order primary space function used for energy contribution estimation (case F). The numbers in parentheses are the corresponding secular equation dimensions

though the lowest threshold used in the extrapolations was quite small (5 μhartree compared to the more typical 20 μhartree), the extrapolation errors, ranging from 1.1 to 1.8 mhartree, were considerably larger than may have been expected on the basis of previous trends.

The BP-type extrapolation plot for case 1 (first line) of Table 2 is shown in Fig. 4 (full lines). The reason for the larger extrapolation error is seen to lie in the rather sharp drop in the plots very near the zero-threshold value. The same data is plotted as $E(T)$ *vs.* $\Delta E^r(T)$ in Fig. 5, and it is seen that the computed points do not lie on an approximately straight line, and that significant extrapolation errors would remain even at rather low thresholds. In fact, a linear fit to the points at $T = 20$, 10, 5 μhartree gives an extrapolation error of 1.9 mhartree. Essentially the same behavior is found for the other calculations in Table 2.

In order to explore the reasons for the extrapolation difficulties described above, a series of additional calculations were carried out for the first example of Table 2. The results of these experiments are summarized in Table 3. These calculations explore the effects of changes in the size of the root set (the set of CF's used to generate the preselected list of CF's) and/or the primary set (used as the zero-order CF's for the estimation of energy contributions).

Case A in Table 3 corresponds to the original calculation (case 1 of Table 2). Case B explores the effect of augmenting both the root set and primary set to include all other CF's with coefficient magnitudes which exceeded 0.03 in case A. The extrapolation curves obtained in this experiment are similar to those of case A,

Table 3. Extrapolation errors for the 1^2A_1 state of the H_2O^+ ion with different choices of the root and primary sets[a]

Case	Root set size	Primary set size	Preselected set size	$E(0)$ (hartrees)	Extrapolation error (mhartrees)
A	6	6	6182[b]	−75.5560[b]	+1.5
B	33	33	20138[c]	−75.5613[c]	+1.1
C	3	3	3339	−75.5550	+0.7
D	1	1	1342	−75.5522	−0.3
E	6	33	6182[b]	−75.5560[b]	+0.3
F[d]	6	6	6182[b]	−75.5560[b]	+1.5

[a] The geometry and orbitals correspond to case 1 of Table 2. Set sizes are the numbers of CF's in each set. Root sets are used to generate the preselected set, while the primary sets are used in Ψ_0 for energy contribution estimates (see text). Root sets of 1, 3, 6, and 33 CF's correspond to 1, 2, 5, and 18 spatial configurations, respectively. See Table 2 for definition of the extrapolation error.

[b] The $E(0)$ result quoted actually corresponds to $T \approx 10^{-9}$ hartree, using 5983 CF's (see Table 2).

[c] The $E(0)$ result quoted actually corresponds to an extrapolation based on $T = 20, 10, \ldots, 0.24$ μhartree, using a maximum of 7531 CF's.

[d] Primary set expansion coefficients taken from final CI vector (see text).

and so is the extrapolation error (the zero-threshold calculation could not be carried out in this case, and $E(0)$ was estimated from an extrapolation based on calculations at $T = 20, 10, \ldots, 0.24$ μhartree). However, the extrapolation error (at $T = 20, 10, 5$ μhartree) is a smaller fraction of $\Delta E^r(T)$ in this case, because of the larger size of $\Delta E^r(T)$.

The following two entries in Table 3 explore the effect of reducing the size of the root and primary sets first to three CF's (case C) and then to one CF (case D). The second and third CF's in case C belong to the same spatial configuration, and have a combined coefficient of 0.12 in the final wave function. While the extrapolation error is reduced in case C, it is still fairly large (particularly when we consider that $\Delta E^r(T)$ is smaller in this than in the previous cases), and the shape of the extrapolation curves is qualitatively similar to those of case A. However, the use of a single root and primary configuration in case D changes the situation markedly, producing a smaller extrapolation error (and of opposite sign) and resulting in relatively normal extrapolation plots (Fig. 6).

The obvious question is then whether it is the size of the root set or the primary set which determines the above behavior. This is explored in example E, in which the root set is that of case A, but the primary set is the 33-term set of case B. The extrapolation error is seen to be considerably smaller in this case than in either of the cases A and B, indicating that some question of balance between the root set and primary set may be involved. This is discussed further in the next section.

Finally, one further calculation (example F) was carried out in order to explore the sensitivity of the selection process to the choice of primary set mixing coefficients

Table 4. Comparison of primary set coefficients used in estimating energy contributions in cases A and F of Table 3

CF[a]	Case A[b]	Case F[c]
1	0.995	0.992
2	0.078	0.112
3	−0.020	−0.044
4	−0.036	−0.018
5	0.015	0.020
6	−0.035	−0.028

[a] Corresponding to the configuration list for $^2A_1$ in Table 1. CF's 2 and 3 belong to the same configuration.

[b] From primary set eigenvector, Eq. (4) (A_k procedure).

[c] Renormalized coefficients from final (unselected) CI eigenvector.

a_i, Eq. (1), used in the evaluation of the energy contributions. This calculation was similar to that of case A, except that the mixing coefficients a_i were taken from the corresponding components of the final unselected CI eigenvector of case A (and renormalized), instead of being obtained from the solution of the primary space

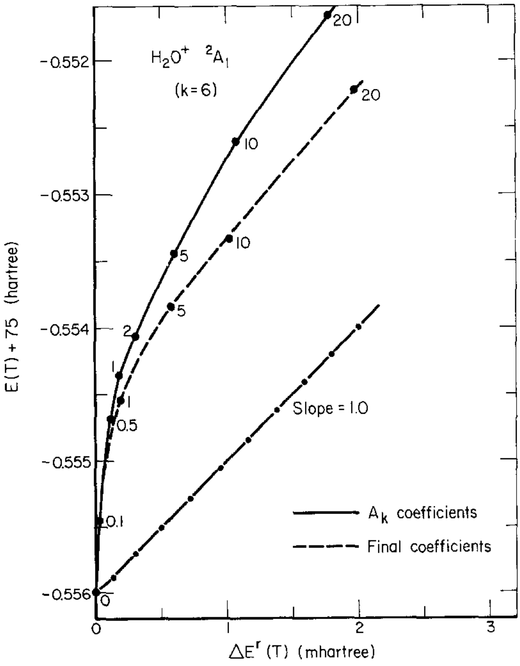

Fig. 5. Alternative extrapolation plots $E(T)$ *vs.* $\Delta E^r(T)$ for the examples of Fig. 4. The numbers next to the dots are the corresponding threshold values (in μhartrees)

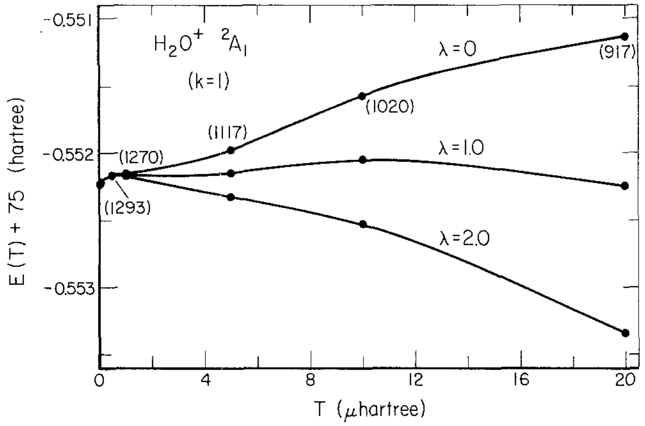

Fig. 6. Extrapolation curves $E_\lambda(T)$ for the $^2A_1$ state of the H_2O^+ ion with a single root function (case D of Table 3). The unselected expansion contained 1342 functions in this case

eigenvalue problem, Eq. (4). (The two sets of coefficients are compared in Table 4.) Since this presumably corresponds to the use of a more accurate zero-order function for the perturbation-theory estimation of the energy contributions of secondary CF's, better selection results may have been expected. The corresponding BP-type extrapolation curves are shown by the dashed lines in Fig. 4, and for $T \geq 5$ μhartree, they are indeed found to be smoother, with a linear and nearly horizontal $\lambda = 1$ curve in this region, and with energies which are lower (and closer to the extrapolated limit) than the corresponding points for case A (full lines). However, the sharp dip at smaller T and the resulting extrapolation error are still essentially the same as in case A, indicating that the extrapolation difficulties are not significantly related to the choice of the primary set mixing coefficients in the A_k selection process. (The corresponding $E(T)$ *vs.* $\Delta E^r(T)$ curves for the two cases are compared in Fig. 5.)

5. Discussion

Clearly, the accuracy of the extrapolation procedures depends on how well $\Delta E^r(T)$, Eq. (7), reflects the actual truncation error $E(T) - E(0)$. It is not necessary for $\Delta E^r(T)$ to be equal to $E(T) - E(0)$, but successful extrapolation requires that the ratio between these quantities be essentially constant as a function of T. The use of $\Delta E^r(T)$ as an estimate of the truncation error is based on perturbation theory (Section 2), and assumes that a second order energy expression based on a zero-order function obtained from the primary space eigenvalue problem is adequate for that purpose.

As seen in Table 4, the relative weights of the primary space functions in Ψ_0 in the A_k procedure are significantly different from their relative weights in the final CI wave function. Furthermore, the results in Figs. 4 and 5 show that this difference is a significant factor in the generation of nonlinear extrapolation plots. But it is also clear that this difference is not responsible for the anomalous small-T behavior of those plots. The reason for that behavior must therefore be sought in the use of second-order energy contribution estimates in $\Delta E^r(T)$. The form of the extrapolation plots also indicates that the principal problem is in the energy contribution estimates for the large number of CF's with very small estimated contributions, primarily those with $\varepsilon_s \leq 1$ μhartree. In fact, the large number of small contribution CF's is a principal characteristic of CI expansions which are generated from large root sets. An examination of the ε_s values for cases D, C, A, and B of Table 3, with root set sizes of 1, 3, 6, and 33, respectively, finds 6, 39, 68, and 76 pet cent of them, respectively, to be smaller than 1 μhartree.

Obviously, the additional small-contribution CF's in the multiroot expansions (A–C) are functions which have no direct interaction with the principal CF (No. 1 in Table 4). Their energy contribution estimates are therefore obtained from their direct interaction with the less important primary CF's. However, due to the large difference in magnitude between the coefficients of these primary CF's and that of the principal function, it is quite likely that higher-order (indirect) interactions of the small-contribution CF's with the principal function are at least as important as their direct, second-order interactions with the other primary functions. Those higher-order interactions are ignored in the selection and extrapolation procedures, and this must be the principal cause of the anomalies in the extrapolation plots.

Further support for this analysis is provided by case E, in which the primary set size of 33 is significantly greater than the root set size of 6. This has the effect of including in the energy contribution estimates interactions which would have been of higher order in case A, and results in a significantly smaller extrapolation error.

The principal source of the extrapolation anomalies is thus found in the inclusion in the root set of CF's with small coefficients. In most applications, such CF's would not be included in the root set, and no difficulties should then be encountered in the extrapolations. However, the inclusion of such functions in the root set is sometimes desirable for reasons extraneous to the selection and extrapolation process. This happens, for example, when more than one electronic state of a given symmetry are to be studied simultaneously [10, 11] or, in some cases, when potential curves or surfaces with correct dissociation behavior or with accurate relative positions [18] are to be determined. In such cases it appears that improved extrapolation behavior can be restored by the use in the energy contribution estimation of a primary CF set which is substantially larger than the root set.

The use of larger primary sets can substantially increase the matrix element computation time in the selection process. In the present example, 60% more central processor time was used in the selection stage in case E compared to case A. However, the selection stage generally accounts for a relatively small fraction of the total CI computation effect, and thus the above increase is not particularly significant.

It should be abundantly clear that this study is concerned solely with extrapolation accuracy, not with the effect of root set size on the corresponding $E(0)$ total energies. In any actual application, the importance of root set choice may greatly exceed the significance of the extrapolation errors, but it is still important to be aware of the magnitude and causes of such errors. This may be particularly important in those cases in which the relative positions of close-lying electronic states (or different regions of a potential surface) are of primary interest.

It should also be noted that several alternative approaches to the reduction of computational effort in CI calculations, often based on the B_k method of Gershgorn and Shavitt [5] and its extensions, have been advocated and used [1, 25–31].

6. Summary

Configuration selection has been an important tool in large-scale CI calculations. Buenker and Peyerimhoff have demonstrated that extrapolation can greatly enhance the effectiveness of that tool, often producing results which are within 0.5 mhartree or the unselected (zero threshold) energy. It has been shown here that the accuracy of the extrapolation is sensitive to the choice of the root set, the set of configuration functions used to generate (by single and double excitations) the preselected configuration list. Quite apart from the question of the effect of the root set choice on the $T = 0$ (unselected) energy, it is seen that unexpectedly large extrapolation errors can result if that root set contains relatively unimportant functions.

The root set is usually chosen on the basis of criteria (such as correct dissociation behavior or the needs of other electronic states) which have nothing to do with the extrapolation process. Second-order energy expressions used in the estimation of energy contributions do not provide reliable estimates for functions which interact directly only with minor components of the root set, since their ignored higher-order, indirect interactions with the major components may be considerably greater than their small estimated contributions. This leads to steep drops in the extrapolation curves at low threshold values and significantly increased extrapolation errors. Improved extrapolation accuracy can be restored in such cases by increasing the size of the primary set, the set of CF's which make up the zero-order function in the energy contribution calculations, well beyond the size of the root set.

Acknowledgement. This work was partially supported by Battelle Memorial Institute.

References

1. Shavitt, I. in: Methods of electronic structure theory, Schaefer, H. F., III, ed. p. 189, New York: Plenum Press 1977
2. Bernal, M. J. M., Boys, S. F.: Phil. Trans. R. Soc. (London) A **245**, 139 (1952)
3. Nesbet, R. K.: Proc. R. Soc. (London) A **230**, 312 (1955)
4. Bender, C. F., Davidson, E. R.: J. Chem. Phys. **47**, 4972 (1967); Phys. Rev. **183**, 23 (1969)
5. Gershgorn, Z., Shavitt, I.: Int. J. Quantum Chem. **2**, 751 (1968)

6. Whitten, J. L., Hackmeyer, M.: J. Chem. Phys. **51**, 5584 (1969); Hackmeyer, M., Whitten, J. L.: J. Chem. Phys. **54**, 3739 (1971); Whitten, J. L.: J. Chem. Phys. **56**, 5458 (1972)
7. Langhoff, S. R., Davidson, E. R.: Int. J. Quantum Chem. **7**, 759 (1973); Langhoff, S. R., Elbert, S. T., Davidson, E. R.: Int. J. Quantum Chem. **7**, 999 (1973)
8. Chu, S.-I., Yoshimine, M., Liu, B.: J. Chem. Phys. **61**, 5389 (1974)
9. Buenker, R. J., Peyerimhoff, S. D.: Theor. Chim. Acta (Berl.) **35**, 33 (1974)
10. Kahn, L. R., Hay, P. J., Shavitt, I.: J. Chem. Phys. **61**, 3530 (1974)
11. Raffenetti, R. C., Hsu, K., Shavitt, I.: Theor. Chim. Acta (Berl.) **45**, 33 (1977)
12. Buenker, R. J., Peyerimhoff, S. D.: Theor. Chim. Acta (Berl.) **39**, 217 (1975)
13. Buenker, R. J., Peyerimhoff, S. D., Butscher, W.: Mol. Phys. **35**, 771 (1978)
14. Langhoff, S. R., Davidson, E. R.: J. Chem. Phys. **64**, 4699 (1976)
15. McLean, A. D., Liu, B.: J. Chem. Phys. **58**, 1066 (1973)
16. Buenker, R. J., Peyerimhoff, S. D.: Chem. Phys. Letters **29**, 253 (1974)
17. Buenker, R. J., Shih, S., Peyerimhoff, S. D.: Chem. Phys. Letters **44**, 385 (1976)
18. Jackels, C. F.: J. Chem. Phys. **72**, 4873 (1980)
19. Dunning Jr., T. H.: J. Chem. Phys. **53**, 2823 (1970)
20. Huzinaga, S.: J. Chem. Phys. **42**, 1293 (1965)
21. Dunning Jr., T. H., Hay, P. J. in: Methods of electronic structure theory, H. F. Schaefer III, ed., p. 1. New York: Plenum Press 1977
22. Davidson, E. R.: Chem. Phys. Letters **21**, 565 (1973)
23. Davidson, E. R.: J. Chem. Phys. **57**, 1999 (1972)
24. Hay, P. J.: J. Chem. Phys. **59**, 2468 (1973)
25. Bagus, P. S., Liu, B., McLean, A. D., Yoshimine, M. in: Computational methods for large molecules and localized states in solids, Herman, F., McLean, A. D., Nesbet, R. K. eds., p. 87 New York: Plenum Press 1973
26. Eberhardt, J. J., Moccia, R., Zandomeneghi, M.: Chem. Phys. Letters **24**, 524 (1974)
27. Segal, G. A., Wetmore, R. W.: Chem. Phys. Letters **32**, 556 (1975)
28. Iwata, S., Freed, K. F.: Chem. Phys. **11**, 433 (1975)
29. Fortune, P. J., Rosenberg, B. J: Chem. Phys. Letters **37**, 110 (1976).
30. Nitzsche, L. E., Davidson, E. R.: J. Chem. Phys. **68**, 3103 (1978)
31. Dunning Jr., T. H.: Chem. Phys. **42**, 249 (1979)

Received August 10, 1980