

Molecular potential energy surfaces constructed from interpolation of systematic fragment surfaces

Michael A. Collins^{a)}

Research School of Chemistry, Australian National University, Canberra ACT 0200, Australia

(Received 23 March 2007; accepted 9 May 2007; published online 10 July 2007)

A systematic method for approximating the *ab initio* electronic energy of molecules from the energies of molecular fragments has previously been presented. Here it is shown that this approach provides a feasible, systematic method for constructing a global molecular potential energy surface (PES) for reactions of a moderate-sized molecule from the corresponding surfaces for small molecular fragments. The method is demonstrated by construction of PESs for the reactions of a hydrogen atom with propane and *n*-pentane. © 2007 American Institute of Physics.
[DOI: 10.1063/1.2746025]

I. INTRODUCTION

Many chemical reactions take place while the molecule remains in a single electronic state, so that the Born-Oppenheimer approximation holds. A complete theoretical description of the mechanism of such a reaction is a description of how the molecule moves during such reactions. To calculate this motion we need to know the total electronic energy of the molecule as a function of the molecular shape; that is, the potential energy surface (PES) which governs the motion of the atomic nuclei. *Ab initio* quantum chemistry provides the practical means to calculate this energy for moderate-sized molecules.

However, the computational time required to calculate the total electronic energy, even at a single molecular geometry, increases rapidly with the number of electrons, N_e , in the molecule, and with the level of *ab initio* theory employed. For example, the reliably accurate coupled cluster approximation, using single and double excitations, and including the triple excitations noniteratively [CCSD(T)], requires a computer time which scales as N_e^7 . Current research is directed towards developing new computational algorithms which reduce this “scaling problem” to linearity, N^1 [see, for example, Ref. 1 and references therein]. Nevertheless, the computational cost of *ab initio* calculations currently represents a substantial barrier to the theoretical study of reactions involving large molecules.

The problem is exacerbated by the fact that the total electronic energy must be known over a wide range of molecular geometries. For a molecule of N atoms, there are $3N-6$ internal molecular coordinates, so the PES is a function defined over a $3N-6$ dimensional space. Hence, to study chemical reaction mechanisms and dynamics for molecules in general, we must overcome two severe “scaling” problems: As the number of atoms increases, so does the number of electrons, and hence the computational cost of a reliable *ab initio* estimate of the PES at any one geometry increases

rapidly, and the number of relevant molecular geometries increases exponentially.

This paper sets out to establish the feasibility of one approach to overcoming these twin difficulties. Zhang and co-workers developed² and applied³ a very simple idea to estimate the *ab initio* interaction energy between two large molecules from the interactions of fragments of the two molecules. These fragments are not disjoint but overlapping segments of the molecules, similar to the fragments employed in “molecular tailoring.”⁴ Motivated by this work, it soon became clear to a number of authors that the total electronic energy of a single molecule could be estimated very simply from combinations of the energies of such tailored molecular fragments.⁵⁻⁷ Two previous papers^{7,8} presented a systematic version of this approach. A molecule is viewed as a collection of bonded functional groups, and a computer algorithm automatically and systematically decomposes the molecule into connected groups or fragments, in a hierarchy of fragment sizes. This provides a sequence, denoted as levels 1, 2, 3, and so on, of increasingly reliable estimates of the total electronic energy of a molecule. The interaction of functional groups with α substituents is accounted for at level 1; α and β substituents at level 2, and α , β , and γ substituents at level 3. In addition, the total energy of a molecule is corrected to account for “nonbonded” interactions between functional groups which are close together in space, even though they may be well separated in terms of bonding.

Here we show that a simple reorganization of this fragmentation hierarchy provides an initial description of the PES for the whole (reacting) molecule in terms of the PES for small, “level 1,” molecular fragments. This initial estimate of the global molecular PES can then be systematically improved by the addition of molecular potential energy correction surfaces (PECSs) which account for the influence of β and γ substituents on the energetics (and reactivity) of the small fragments.

The initial, level 1, description of the PES for the whole (reacting) molecule requires PES for small molecular fragments which typically contain several atoms. It has been

^{a)}Electronic mail: collins@rsc.anu.edu.au

demonstrated over the last several years that a modified form of Shepard interpolation can be used to construct PES for such molecules. These PESs are constructed from *ab initio* calculations at a sparse set of molecular geometries scattered over the relevant configuration space.^{9–11} The interpolation method used to construct the PES is combined with classical trajectory simulations to provide an iterative scheme for successively and systematically improving the accuracy of the PES with additional *ab initio* calculations. The PES for such small molecules can be refined in this way to an accuracy at which experimental data can be accurately simulated.¹²

To attain the same accuracy for larger molecules, it will be necessary to systematically include the PECSs in the total description of the energy surface. The potential energy correction surfaces, derived below, can be expected to depend on the coordinates of about ten or more atoms. Accurate interpolation, from sparse scattered data, of functions defined on such large dimensional spaces can be expected to present a formidable problem. However, these PECSs are much simpler objects than PESs. Conceptually, a PES which describes a chemical reaction is reminiscent of the Himalayan mountains: A landscape characterized by high peaks, deep valleys, and sudden change. In contrast, as demonstrated below, a PECS is a surface with most if not all of these violent variations removed; a landscape more reminiscent of rolling downs, if not the Netherlands. Heuristically, we can expect that relatively flat functions are relatively easily estimated by interpolation of sparse data, even if the space of coordinates is large. Hence, it may be feasible to construct accurate energy surfaces to describe the chemical reaction dynamics of moderate-sized molecules.

The purpose of this paper is to describe the general approach and to demonstrate the feasibility of the method with two simple examples. The examples chosen are the abstraction reactions of a hydrogen atom with propane and *n*-pentane. The paper is set out as follows. Section II presents the general form for the expansion of the PES in terms of level 1 PESs and higher level PECSs. Section III presents some results for the reactions of a hydrogen atom with propane and pentane. The final section contains some discussion and concluding remarks.

II. PES AND PECS

A molecule is viewed herein as a set of bonded functional groups; with the usual definition of functional groups, defined in detail in Papers I (Ref. 7) and II (Ref. 8). Atoms connected by double bonds belong to the same functional group. So ethanol contains three functional groups (CH₃, CH₂, and OH) while ethanal contains just two functional groups (CH₃ and CHO). The total electronic energy of a molecule is expressed as an energy due to bonding interactions between functional groups and that due to nonbonded interactions between functional groups.

A. Systematic fragmentation

The bonding energy can be estimated in a systematic sequence of approximations which result from breaking the molecule into fragments of increasing size. When we break a

bond between functional groups in a molecule, we restore the bonding by appending hydrogen atoms to each group. Thus breaking all the intergroup bonds in ethanol produces three isolated fragment molecules: CH₄, CH₄, and H₂O.

It is important to emphasize that in all other respects, the shape of the fragment molecules is *identical* to the shape they had in the parent molecule.

To illustrate the basic fragmentation approach, consider the simple example of a chain molecule of *K* groups,⁷

$$M = G_1 G_2 G_3 \cdots G_K. \quad (2.1)$$

Suppose the molecule is broken between groups *n*–1 and *n*,

$$M \rightarrow M_1 + M_2, \quad (2.2)$$

where

$$M_1 = G_1 G_2 G_3 \cdots G_{n-1} H^{(n-1)}$$

and

$$M_2 = H^{(n)} G_n G_{n+1} \cdots G_K.$$

The total electronic energies of these three molecules are trivially related by

$$E(M) = E(M_1) + E(M_2) + dE_1, \quad (2.3)$$

where the net energy change, *dE*₁, is unknown. Breaking *M* at some *other* place, say, between groups *i*–1 and *i*, gives

$$E(M) = E(M_3) + E(M_4) + dE_2, \quad (2.4)$$

where *M*₃ = *G*₁*G*₂*G*₃⋯*G*_{*i*–1}*H*^(*i*–1), *M*₄ = *H*^(*i*)*G*_{*i*}*G*_{*i*+1}⋯*G*_{*K*} and *dE*₂ is also unknown. Now, if we break *M* at *both* places,

$$E(M) = E(M_1) + E(M_5) + E(M_4) + dE_3, \quad (2.5)$$

where *M*₅ = *H*^(*n*)*G*_{*n*}*G*_{*n*+1}⋯*G*_{*i*–1}*H*^(*i*–1). Now, if the *G*_{*n*–1}*G*_{*n*} bond is far in space from the *G*_{*i*–1}*G*_{*i*} bond, then the energy change from the simultaneous fragmentation, *dE*₃, will be very nearly equal to the sum of the energy changes from each separate fragmentation. That is,

$$dE_3 \approx dE_1 + dE_2. \quad (2.6)$$

The more separated the two fragmentation sites are, the closer should be the equality in Eq. (2.6). From the definitions of these energy changes,

$$\begin{aligned} E(M) - E(M_1) - E(M_5) - E(M_4) \\ \approx E(M) - E(M_1) - E(M_2) + E(M) - E(M_3) - E(M_4) \end{aligned}$$

or

$$E(M) \approx E(M_2) + E(M_3) - E(M_5). \quad (2.7)$$

Note that *M*₅ is the “overlapping” segment, common to *M*₂ and *M*₃, and in this sense

$$M \rightarrow M_2 + M_3 - M_5. \quad (2.8)$$

To make this fragmentation into a systematic procedure we simply define a sequence of rules for deciding which pair of bonds to break, one at a time and simultaneously. By definition, at level 1, the bonds are separated by just one functional group (both bonds have a common group). A molecule is said to be fragmented at level 1, when we break *all* possible pairs of bonds that satisfy this condition. At level 2,

the bonds are separated by two functional groups; at level 3, the bonds are separated by three functional groups, and so on. The details can be found in Papers I and II. Pentane, a chain of five functional groups, numbered 12345, provides a very simple example of fragmentation. The final fragments are as follows: at level 1, 12+23+34+45-2-3-4-5; at level 2, 123+234+345-23-34; and at level 3, 1234+2345-234. In general, the molecule is represented by a sum (and difference) of molecular fragments as

$$M \rightarrow \sum_{n=1}^{N_{\text{frag}}^{(L)}} c_n^{(L)} F_n^{(L)}, \quad (2.9)$$

where $F_n^{(L)}$ denotes a molecular fragment at level L and $c_n^{(L)}$ is a numerical coefficient.

The total energy due to the functional groups and the bonding between them is given by

$$E_b(M) = \sum_{n=1}^{N_{\text{frag}}^{(L)}} c_n^{(L)} E(F_n^{(L)}). \quad (2.10)$$

We recall that the molecular geometry of each fragment, $F_n^{(L)}$, is the same as that for the corresponding atoms in the original molecule, with the addition of hydrogen atom ‘‘caps.’’ These H atoms are placed along the direction of the broken bond, at a distance which approximately corresponds to the appropriate bond length,

$$\mathbf{x}(\text{H}) = \mathbf{x}(i) + \frac{\text{rad}(i) + \text{rad}(\text{H})}{\text{rad}(i) + \text{rad}(j)} [\mathbf{x}(j) - \mathbf{x}(i)], \quad (2.11)$$

where $\mathbf{x}(i)$ denotes the Cartesian position of the atom to which H is attached, $\mathbf{x}(j)$ denotes the position of the ‘‘missing’’ bonded atom in the original molecule, and rad denotes the covalent radius of the corresponding atom. If one sketches the molecules resulting from the simple fragmentations of pentane (for example), it is easy to see that the capping H atoms in the positive ($c_n^{(L)} > 0$) fragments have exact counterparts in the negative fragments, so that there are no ‘‘net’’ additional H atoms on the right hand side of Eq. (2.9).

The gradient of the energy with respect to the coordinates of the atoms in the original molecule can be obtained as a sum (and difference) of the gradients of the energies of the fragments. The capping H atoms contribute to the energy gradients for the original atoms since the positions of the extra H atoms only depend on the coordinates of the original atoms, as shown in Eq. (2.11). Similarly, the second and higher energy derivatives can be estimated.

Paper II presented a substantial body of calculations which showed that Eq. (2.10) approaches chemical accuracy for moderate-sized organic molecules by level 3.

B. Energy corrections

Suppose that a molecule has been fragmented at each of the levels 1–3. Then we can combine these fragmentations in a tautology as follows:

$$M = \left[M - \sum_{n=1}^{N_{\text{frag}}^{(3)}} c_n^{(3)} F_n^{(3)} \right] + \left[\sum_{n=1}^{N_{\text{frag}}^{(3)}} c_n^{(3)} F_n^{(3)} - \sum_{n=1}^{N_{\text{frag}}^{(2)}} c_n^{(2)} F_n^{(2)} \right] + \left[\sum_{n=1}^{N_{\text{frag}}^{(2)}} c_n^{(2)} F_n^{(2)} - \sum_{n=1}^{N_{\text{frag}}^{(1)}} c_n^{(1)} F_n^{(1)} \right] + \sum_{n=1}^{N_{\text{frag}}^{(1)}} c_n^{(1)} F_n^{(1)}. \quad (2.12)$$

Now, each of the level 3 fragments may itself be further fragmented at level 2 as follows:

$$F_n^{(3)} \rightarrow \sum_{m=1}^{N_{\text{frag}}^{(2)}(n,m)} c_m^{(2)}(n) F_m^{(2)}(n). \quad (2.13)$$

The level 3 fragmentation rule does not violate the level 2 rule, as the bonds broken at level 3 are further apart than they must be at level 2. A level 3 fragmentation can be thought of as an incomplete level 2 fragmentation. Fragmentation of all the level 3 fragments at level 2 is equivalent to completing the (exhaustive) application of the level 2 rule. Thus, fragmentation of the molecule at (say) level 2 is equivalent to the sum of all the level 3 fragments fragmented at level 2,

$$\sum_{n=1}^{N_{\text{frag}}^{(2)}} c_n^{(2)} F_n^{(2)} = \sum_{n=1}^{N_{\text{frag}}^{(3)}} c_n^{(3)} \left[\sum_{m=1}^{N_{\text{frag}}^{(2)}(n,m)} c_m^{(2)}(n) F_m^{(2)}(n) \right]. \quad (2.14)$$

Hence, the second collected term in Eq. (2.12) can be rewritten as

$$\sum_{n=1}^{N_{\text{frag}}^{(3)}} c_n^{(3)} F_n^{(3)} - \sum_{n=1}^{N_{\text{frag}}^{(2)}} c_n^{(2)} F_n^{(2)} = \sum_{n=1}^{N_{\text{frag}}^{(3)}} c_n^{(3)} \left[F_n^{(3)} - \sum_{m=1}^{N_{\text{frag}}^{(2)}(n,m)} c_m^{(2)}(n) F_m^{(2)}(n) \right]. \quad (2.15)$$

Now, each of the fragments at level 3 which have negative coefficients cannot be decomposed further at level 2, so that all the terms with coefficients $c_n^{(3)} < 0$ on the right hand side (rhs) of Eq. (2.15) vanish. The same argument applies to fragmentation of level 2 fragments at level 1, so that Eq. (2.12) can be written as

$$M = \left[M - \sum_{n=1}^{N_{\text{frag}}^{(3)}} c_n^{(3)} F_n^{(3)} \right] + \sum_{n=1}^{N_{\text{frag}}^{(3+)}} c_n^{(3)} \left[F_n^{(3)} - \sum_{m=1}^{N_{\text{frag}}^{(2)}(n,m)} c_m^{(2)}(n) F_m^{(2)}(n) \right] + \sum_{n=1}^{N_{\text{frag}}^{(2+)}} c_n^{(2)} \left[F_n^{(2)} - \sum_{m=1}^{N_{\text{frag}}^{(1)}(n,m)} c_m^{(1)}(n) F_m^{(1)}(n) \right] + \sum_{n=1}^{N_{\text{frag}}^{(1)}} c_n^{(1)} F_n^{(1)}, \quad (2.16)$$

where $N_{\text{frag}}^{(3+)}$ and $N_{\text{frag}}^{(2+)}$ denote the number of ‘‘positive’’ fragments at levels 3 and 2, respectively, and the associated sums only involve the positive fragments at the corresponding

level. The corresponding tautology for the energy of the molecule is given by

$$\begin{aligned}
 E(M) = & \left[E(M) - \sum_{n=1}^{N_{\text{frag}}^{(3)}} c_n^{(3)} E[F_n^{(3)}] \right] \\
 & + \sum_{n=1}^{N_{\text{frag}}^{(3+)}} c_n^{(3)} \left[E[F_n^{(3)}] - \sum_{m=1}^{N_{\text{frag}}^{(2)}(n,m)} c_m^{(2)}(n) E[F_m^{(2)}(n)] \right] \\
 & + \sum_{n=1}^{N_{\text{frag}}^{(2+)}} c_n^{(2)} \left[E[F_n^{(2)}] - \sum_{m=1}^{N_{\text{frag}}^{(1)}(n,m)} c_m^{(1)}(n) E[F_m^{(1)}(n)] \right] \\
 & + \sum_{n=1}^{N_{\text{frag}}^{(1)}} c_n^{(1)} E[F_n^{(1)}]. \quad (2.17)
 \end{aligned}$$

We can now draw an important distinction between a molecular (PES) and a molecular (PECS). Each term in the last sum in Eq. (2.17), $E[F_n^{(1)}]$, represents, for all possible configurations of fragment $F_n^{(1)}$, the molecular potential energy surface for the (small) molecule represented by $F_n^{(1)}$. However, each term in Eq. (2.17) of the form

$$E[F_n^{(2)}] - \sum_{m=1}^{N_{\text{frag}}^{(1)}(n,m)} c_m^{(1)}(n) E[F_m^{(1)}(n)] \equiv \text{PECS}(F_n^{(2)}; 2) \quad (2.18)$$

represents, for all possible configurations of fragment $F_n^{(2)}$, only the surface which describes the energy difference between the energy of fragment $F_n^{(2)}$ and the level 1 approximation to this energy. Hence, we can rewrite the tautology for the molecular energy surface yet again as

$$\begin{aligned}
 E(M) = & \text{PECS}(M; 4) + \sum_{n=1}^{N_{\text{frag}}^{(3+)}} c_n^{(3)} \text{PECS}(F_n^{(3)}; 3) \\
 & + \sum_{n=1}^{N_{\text{frag}}^{(2+)}} c_n^{(2)} \text{PECS}(F_n^{(2)}; 2) + \sum_{n=1}^{N_{\text{frag}}^{(1)}} c_n^{(1)} E[F_n^{(1)}]. \quad (2.19)
 \end{aligned}$$

A very simple example of this reexpression of the energy surface for the molecule in terms of PESs and PECSs is provided by the pentane fragmentation described above. Denoting pentane as the 12345 molecule, we can write

$$\begin{aligned}
 E(12345) = & \{E(12345) - [E(1234) + E(2345) - E(234)]\} + \{E(1234) - [E(123) + E(234) - E(23)]\} + \{E(2345) - [E(234) \\
 & + E(345) - E(34)]\} + \{E(123) - [E(12) + E(23) - E(2)]\} + \{E(234) - [E(23) + E(34) - E(3)]\} + \{E(345) \\
 & - [E(34) + E(45) - E(4)]\} + E(12) + E(23) + E(34) + E(45) - E(2) - E(3) - E(4). \quad (2.20)
 \end{aligned}$$

Each term collected in $\{\}$ in Eq. (2.20) is a PECS. The last seven terms are a sum and difference of PESs for level 1 fragments.

It is important to realize that a PES, such as the PES for ethane, $E(12)$, has a value on the order of (minus) several tens of hartrees, and varies by some hundreds of kJ mol^{-1} as the molecular geometry changes during a chemical process. In contrast, a PECS such as that for propane at level 2,

$$\text{PECS}(123; 2) = E(123) - [E(12) + E(23) - E(2)], \quad (2.21)$$

has a value which spans only tens of kJ mol^{-1} as the molecular geometry changes during a chemical process. This PECS describes the error in a level 1 fragmentation approximation

to the energy of propane. We can also interpret this PECS as the energy correction due to the β substituent interaction, which is not accounted for at level 1. Similarly, the level 3 PECS for butane, $\text{PECS}(1234; 3)$, describes the energy correction for the γ substituent interaction that is not accounted for at level 2,

$$\text{PECS}(1234; 3) = E(1234) - [E(123) + E(234) - E(23)]. \quad (2.22)$$

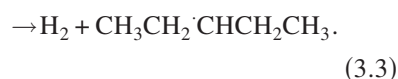
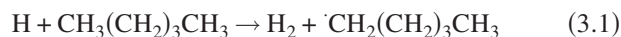
In order to appreciate the general result in Eq. (2.19), it is also useful to consider a slightly more complicated example, provided by a branched molecule like 2-methylbutane. In this case, the level 3 fragmentation returns the whole molecule and Eq. (2.19) becomes

$$\begin{aligned}
 E\left(\begin{matrix} 3 \\ 1245 \end{matrix}\right) = & \left\{ E\left(\begin{matrix} 3 \\ 1245 \end{matrix}\right) - \left[E\left(\begin{matrix} 3 \\ 124 \end{matrix}\right) + E(245) - E(24) \right] \right\} + \left\{ E\left(\begin{matrix} 3 \\ 124 \end{matrix}\right) - [E(12) + E(23) + E(24) - E(2) - E(2)] \right\} \\
 & + \{E(245) - [E(24) + E(45) - E(4)]\} + E(12) + E(23) + E(24) + E(45) - E(2) - E(2) - E(4). \quad (2.23)
 \end{aligned}$$

Since an automated computer algorithm (see Papers I and II) can fragment any arbitrary molecule at each level, the PES for any molecule can be expressed in terms of the PES and PECS of smaller molecules in this way.

III. H+CH₃CH₂CH₃ AND H+CH₃(CH₂)₃CH₃

To test the feasibility of constructing a PES which describes reactions of a moderate-sized polyatomic molecule, we consider the abstraction reactions of a hydrogen atom with pentane,



Since the purpose here is to test the accuracy of a global surface constructed from PES and PECS, rather than to construct a chemically accurate *ab initio* surface for these particular reactions, a low level of electronic structure theory has been employed for computational convenience. This allows rapid evaluation of the various surfaces and facilitates testing the fragmentation approach. All energy calculations have been carried out using density functional theory with the B3LYP functional¹³ and a Dunning cc-pVDZ basis set.¹⁴ All calculations were carried out using the GAUSSIAN program.¹⁵

The accuracy of the fragmentation approximations for the classical barrier heights and reaction energies in reactions

TABLE I. The UB3LYP/cc-pVDZ values (kJ mol⁻¹) of the classical barrier heights, ΔE^\ddagger , and classical energy difference (products minus reactants), ΔE , for hydrogen atom abstraction in H+CH₃(CH₂)₃CH₃ [Eqs. (3.1)–(3.3)], H+CH₄ [Eq. (3.6)], H+CH₃CH₃ [Eq. (3.7)], and H+CH₃CH₂CH₃ [Eqs. (3.9) and (3.10)] are compared with the corresponding errors in the estimates at fragmentation levels 1, 2, and 3.

Reaction energetics	Value	Errors		
		Level 1	Level 2	Level 3
3.1 ΔE^\ddagger	27.38	-0.87	-0.05	-0.05
3.2 ΔE^\ddagger	17.12	-1.60	0.12	-0.05
3.3 ΔE^\ddagger	17.72	-2.32	0.27	-0.00
3.6 ΔE^\ddagger	39.90
3.7 ΔE^\ddagger	26.80
3.9 ΔE^\ddagger	27.40	-0.78	0	0
3.10 ΔE^\ddagger	16.63	-0.89	0	0
3.1 ΔE	-5.62	-0.92	0.02	-0.01
3.2 ΔE	-21.58	-5.32	0.45	-0.11
3.3 ΔE	-20.64	-6.43	0.98	0.05
3.6 ΔE	12.97
3.7 ΔE	-6.40
3.9 ΔE	-5.44	-0.96	0	0
3.10 ΔE	-22.22	-4.33	0	0

(3.1)–(3.3) is indicated in Table I. It is clear that these energies are reproduced to better than 1 kJ mol⁻¹ at level 2. Hence, to study the competing reactions (3.1)–(3.3), it may suffice to construct a PES at level 2.

The PES for H+CH₃(CH₂)₃CH₃ can be obtained from Eq. (2.20) by reacting a hydrogen atom with each and every fragment,

$$\begin{aligned} E(\text{H} + 12345) = & \{E(\text{H} + 12345) - [E(\text{H} + 1234) + E(\text{H} + 2345) - E(\text{H} + 234)]\} + \{E(\text{H} + 1234) - [E(\text{H} + 123) + E(\text{H} + 234) \\ & - E(\text{H} + 23)]\} + \{E(\text{H} + 2345) - [E(\text{H} + 234) + E(\text{H} + 345) - E(\text{H} + 34)]\} + \{E(\text{H} + 123) - [E(\text{H} + 12) \\ & + E(\text{H} + 23) - E(\text{H} + 2)]\} + \{E(\text{H} + 234) - [E(\text{H} + 23) + E(\text{H} + 34) - E(\text{H} + 3)]\} + \{E(\text{H} + 345) \\ & - [E(\text{H} + 34) + E(\text{H} + 45) - E(\text{H} + 4)]\} + E(\text{H} + 12) + E(\text{H} + 23) + E(\text{H} + 34) + E(\text{H} + 45) - E(\text{H} + 2) \\ & - E(\text{H} + 3) - E(\text{H} + 4). \end{aligned} \quad (3.4)$$

The level 2 approximation to this PES is given in terms of PES and PECS by

$$\begin{aligned} E(\text{H} + 12345) \approx & \{E(\text{H} + 123) - [E(\text{H} + 12) + E(\text{H} + 23) - E(\text{H} + 2)]\} + \{E(\text{H} + 234) - [E(\text{H} + 23) + E(\text{H} + 34) - E(\text{H} + 3)]\} \\ & + \{E(\text{H} + 345) - [E(\text{H} + 34) + E(\text{H} + 45) - E(\text{H} + 4)]\} + E(\text{H} + 12) + E(\text{H} + 23) + E(\text{H} + 34) + E(\text{H} + 45) \\ & - E(\text{H} + 2) - E(\text{H} + 3) - E(\text{H} + 4). \end{aligned} \quad (3.5)$$

From Eq. (3.5), it is clear that the level 1 PESs required for this case must describe the reactions of hydrogen with methane and ethane as follows:



PESs for these reactions have been constructed by modified Shepard interpolation of B3LYP/cc-pVDZ data, using the GROW program.¹¹ The relative energies of the various reactants, products, and saddle points for the relevant reactions are also shown in Table I.

From Eq. (3.5), it is also clear that the level 2 PECS required for this case is

$$\begin{aligned} \text{PECS}(\text{H} + \text{CH}_3\text{CH}_2\text{CH}_3; 2) \\ = E(\text{H} + \text{CH}_3\text{CH}_2\text{CH}_3) - [E(\text{H} + \text{CH}_3\text{CH}_3) \\ + E(\text{H} + \text{CH}_3\text{CH}_3) - E(\text{H} + \text{CH}_4)]. \end{aligned} \quad (3.8)$$

A. PES for H+CH₄

The optimized geometries for the reactants, products, and saddle point for reaction (3.6) were determined at UB3LYP/cc-pVDZ. The classical energy barrier is 39.90 kJ mol⁻¹, and the classical reaction energy is endoergic by 12.97 kJ mol⁻¹ (see Table I). The minimum energy path for reaction (3.6) was approximately evaluated at the B3LYP/cc-pVDZ level of theory, and a set of 62 points along this path was employed as an initial data set. This approximate reaction path was determined by simple energy minimization, beginning with configurations displaced slightly to either side of the saddle point. Typically these minimizations terminate at moderate separations of the reactants and products, and configurations in the asymptotes are constructed simply by further separation of the fragments in steps of about 0.5 bohr. Since these configurations only provide an initial data set, it is not necessary to construct an accurate intrinsic reaction path. The PES was then constructed using the standard iterative GROW procedure with data sampled from classical trajectories. These trajectories were initiated for reaction (3.6) and the reverse reaction, and for molecular configurations initially in the vicinity of the saddle point. A range of initial vibrational and relative translational energies were employed. A data set of 1048 points was thereby accumulated. The interpolation is carried out including every version of these data points allowed by the molecular symmetry group. Here the symmetry group is all 120 permutations of the hydrogen atoms. The confidence lengths, used in the interpolation weight function, were evaluated with $M=200$ and $E_{\text{tot}}=1.8$ kJ mol⁻¹.⁹ To probe the accuracy of this interpolated PES, an additional sample of 755 molecular configurations was sampled from classical trajectories and the B3LYP/cc-pVDZ value of the energy was compared to that calculated from the interpolated PES at each configuration. Figure 1 presents the variation of the average absolute interpolation error with the size of the data set. The convergence of the PES accuracy with data set size is also indicated by the variation of the reaction probability (at zero impact parameter) with the size of the data set, also shown in Fig. 1. The reaction probability was evaluated from a sample of 2000 trajectories, for which CH₄ initially has zero angular momentum and a microcanonical distribution of vibrational energy of 112.9 kJ mol⁻¹ (corresponding approximately to the zero point energy). The initial relative collision energy was 39.4 kJ mol⁻¹.

B. PES for H+CH₃CH₃

The optimized geometries for the reactants, products, and saddle point for reaction (3.7) were determined at UB3LYP/cc-pVDZ. The classical energy barrier is

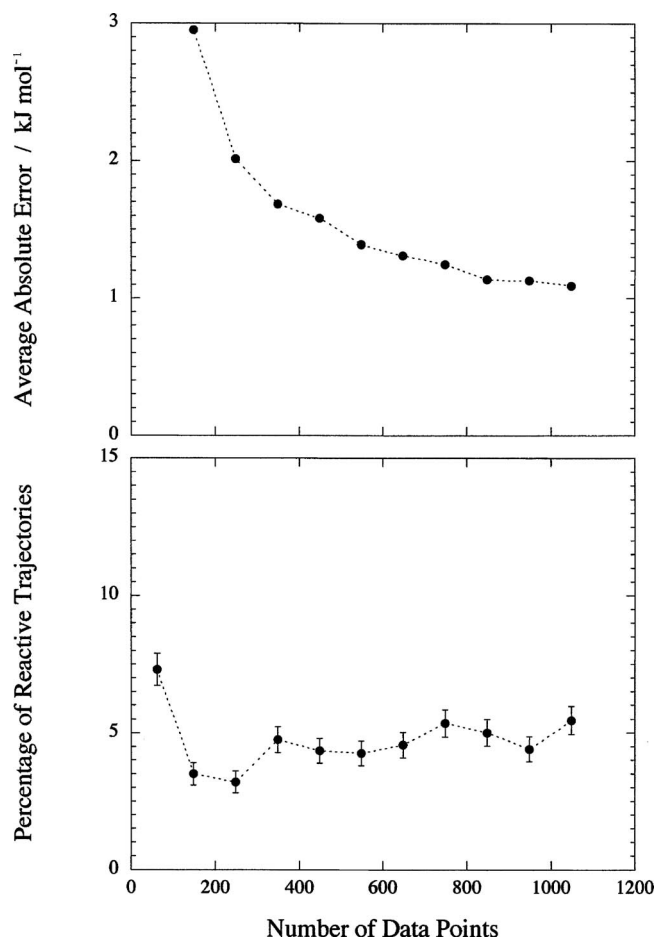


FIG. 1. The average absolute interpolation error, calculated for a sample of 755 configurations for H+CH₄, is shown as a function of the size of the data set (upper panel). The lower panel shows the classical reaction probability (see text for details) as a function of the size of the data set. The error bars denote one standard deviation due to the finite trajectory sample. The lines are merely visual aids.

26.80 kJ mol⁻¹, and the classical reaction energy is -6.40 kJ mol⁻¹ (see Table I). The minimum energy path for reaction (3.7) was approximately evaluated at the B3LYP/cc-pVDZ level of theory, and a set of 47 points along this path was employed as an initial data set. The PES was then constructed using the standard iterative GROW procedure with data sampled from classical trajectories, as for reaction (3.6) above. A total of 2297 data points were accumulated. Here the molecular symmetry group¹⁶ is taken to contain 504 permutations of the hydrogen and carbon atoms. This group allows for all 72 permutations within CH₃CH₃ which do not involve bond breaking. The total group is obtained as the product of this group with the six permutations of the colliding H atom with each H atom in ethane (plus the identity). The confidence lengths, used in the interpolation weight function, were evaluated with $M=150$ and $E_{\text{tot}}=1.8$ kJ mol⁻¹.⁹ The accuracy of this interpolated PES was estimated by evaluating the B3LYP/cc-pVDZ energy at each of a set of 750 molecular configurations which were sampled from classical trajectories. Figure 2 presents the variations of the average interpolation error and the reaction probability (at zero impact parameter) with data set size for this surface. The reaction probability was evaluated from a sample of

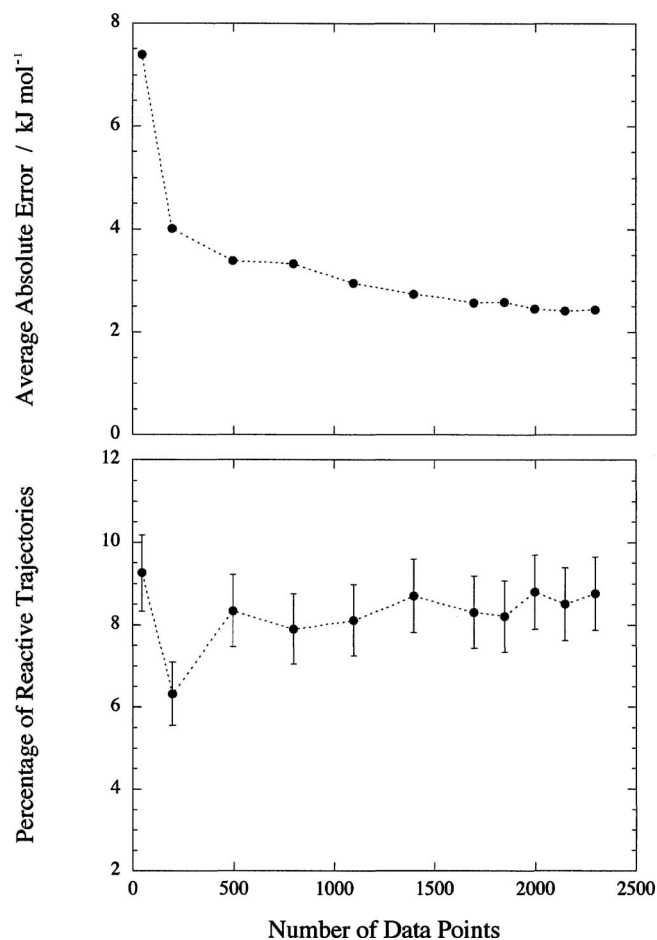


FIG. 2. The average absolute interpolation error, calculated for a sample of 750 configurations for $\text{H}+\text{CH}_3\text{CH}_3$, is shown as a function of the size of the data set (upper panel). The lower panel shows the classical reaction probability (see text for details) as a function of the size of the data set. The error bars denote one standard deviation due to the finite trajectory sample. The lines are merely visual aids.

1000 trajectories, for which CH_3CH_3 initially has zero angular momentum and a microcanonical distribution of vibrational energy of $170.7 \text{ kJ mol}^{-1}$ (corresponding approximately to the zero point energy). The initial relative collision energy was 39.4 kJ mol^{-1} .

C. PES and PECS for $\text{H}+\text{CH}_3\text{CH}_2\text{CH}_3$

There are two possible products for hydrogen atom abstraction from propane:



The optimized geometries for the reactants, products, and saddle points for these reactions were determined at UB3LYP/cc-pVDZ. The classical energy barriers and reaction energies are shown in Table I. The minimum energy paths for both reactions (3.9) and (3.10) were approximately evaluated by simple energy minimization at the UB3LYP/cc-pVDZ level of theory, and a set of 87 points along these paths was selected. An interpolated UB3LYP/cc-pVDZ PES for these reactions was evaluated using these 87 configura-

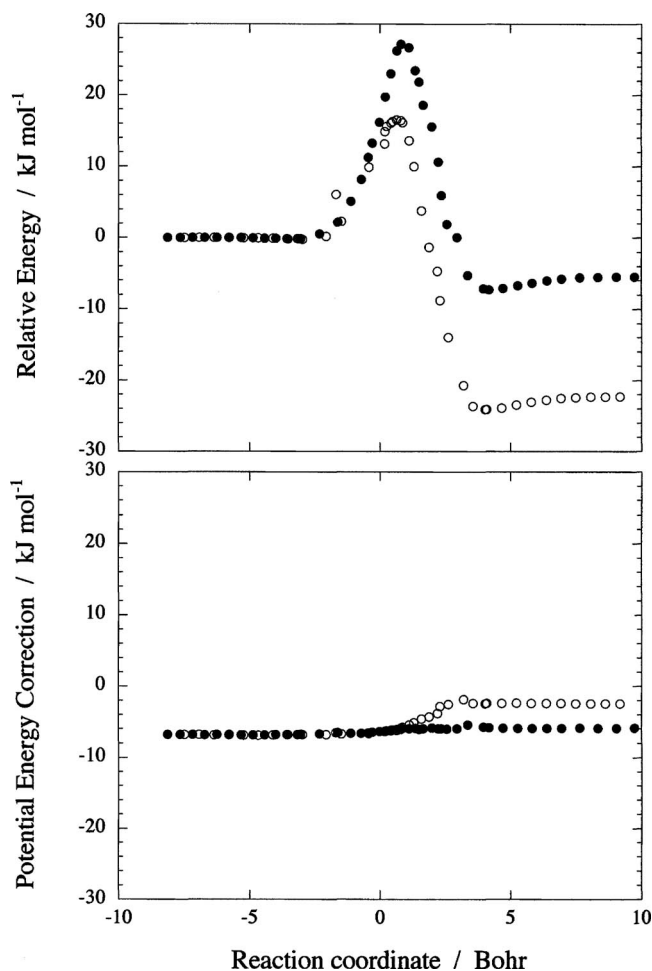


FIG. 3. The energy profiles (relative to the reactants) along approximate minimum energy paths for reaction (3.9) (●) and reaction (3.10) (○) are shown in the upper panel. The lower panel shows the corresponding values of the PECS of Eq. (3.8).

tions as the data set. Using the same data set, a level 2 PECS for $\text{H}+\text{CH}_3\text{CH}_2\text{CH}_3$ was also constructed according to Eq. (3.8). The molecular symmetry group for $\text{H}+\text{CH}_3\text{CH}_2\text{CH}_3$ has 1296 elements. This group allows for all 144 permutations within $\text{CH}_3\text{CH}_2\text{CH}_3$ which do not involve bond breaking. The total group is obtained as the product of this group with the eight permutations of the colliding H atom with each H atom in propane (plus the identity).

The variations of the PES and the level 2 PECS along the reactions paths are shown in Fig. 3. This figure clearly indicates that the PECS is a much “flatter” function than the corresponding PES.

D. Interpolation accuracy

Equation (3.5) implies that the PES at one configuration of $\text{H}+\text{CH}_3(\text{CH}_2)_3\text{CH}_3$ might be approximated from the PES for $\text{H}+\text{CH}_4$, evaluated at three configurations of CH_5 , the PES for $\text{H}+\text{CH}_3\text{CH}_3$, evaluated at four configuration of C_2H_7 , and the level 2 PECS for $\text{H}+\text{CH}_3\text{CH}_2\text{CH}_3$, evaluated at three configurations of C_3H_9 . Each PES or PECS evaluation is subject to an error which results from the finite size of the data set in each case. The utility of Eq. (3.5) depends on

TABLE II. The average absolute interpolation errors (in kJ mol^{-1} , see text for details) are shown for the PES, and including level 2 PECS, for reactions of a hydrogen atom with four alkanes. The energy range denotes the largest energy differences between two geometries in the samples of configurations used.

PES and PECS	Number of data points	Average absolute error (kJ mol^{-1})	Energy range (kJ mol^{-1})
H+CH ₄	1098	1.1	148
H+CH ₃ CH ₃	2297	2.4	150
H+CH ₃ CH ₂ CH ₃	143	4.7	169
H+CH ₃ CH ₂ CH ₂ level 1	0	5.2	169
H+CH ₃ CH ₂ CH ₃ level 1+PECS	143	3.4	169
H+CH ₃ (CH ₂) ₃ CH ₃ level 1	0	12.6	206
H+CH ₃ (CH ₂) ₃ CH ₃ level 1+PECS	0	5.7	206

the magnitude of the resulting composite interpolation error. Table II presents a summary of the average interpolation errors for each component of Eq. (3.5).

As described above, the average interpolation errors for H+CH₄ and H+CH₃CH₃ were estimated from samples of several hundred configurations taken from trajectories for each reaction. Table II shows the average absolute errors for the full data sets, and Figs. 2 and 3 suggest that the reaction dynamics for H+CH₄ and H+CH₃CH₃ are insensitive to average errors of these magnitudes.

A sample of 750 configurations of C₃H₉ was obtained from trajectories for H+CH₃CH₂CH₃. These trajectories were evaluated at the level 1 approximation to the PES, that is, using just the PES for H+CH₄ and H+CH₃CH₃. A sample of 750 configurations of C₅H₁₃ was similarly obtained from trajectories for H+CH₃(CH₂)₃CH₃. It was found that the level 1 approximation to the energies for C₃H₉ and C₅H₁₃ was in very large error for some of these sampled configurations. By comparing the structures of these “high-error” configurations with the PES data sets, it was clear that these errors arose from the fact that the level 1 fragments corresponding to these high-error configurations were far from all configurations in the H+CH₄ and H+CH₃CH₃ data sets. This suggests that the normal sampling scheme employed by the GROW program to construct PES is inadequate for this composite approach to the PES for larger systems.

The data sets for H+CH₄ and H+CH₃CH₃ were selected from configurations encountered by classical trajectories for reactions (3.6) and (3.7), respectively. The domains of the data set configurations are limited by the maximum total energy at which the corresponding trajectories were calculated. The data distribution also reflects the relative probability with which configuration space is sampled by classical trajectories for the reactions studied. When classical trajectories are evaluated for H+CH₃CH₂CH₃ and H+CH₃(CH₂)₃CH₃, the maximum potential energy (relative to the lowest energy configuration) may be higher than for H+CH₄ and H+CH₃CH₃. Moreover, the relative probability with which CH₅ and C₂H₇ configuration spaces are sampled must be different in C₃H₉ and C₅H₁₃ trajectories than in trajectories for CH₅ and C₂H₇. Hence, PES data sets constructed for H+CH₄ and H+CH₃CH₃ may not adequately sample the space of CH₅ and C₂H₇ configurations encountered in trajectories for C₃H₉ and C₅H₁₃.

To account for this sampling problem, 56 high-error con-

figurations of C₃H₉ were selected from the C₃H₉ and C₅H₁₃ samples. The corresponding 56 level 1 configurations of CH₅ and 112 configurations of C₂H₇ were added to the PES data sets for H+CH₄ and H+CH₃CH₃. The 56 configurations for C₃H₉ were also added to the data sets for the PES of H+CH₃CH₂CH₃ and the corresponding level 2 PECS, bringing the number of data points in both sets to 143. New samples of 750 configurations of C₃H₉ and C₅H₁₃ were generated from level 1 trajectories for H+CH₃CH₂CH₃ and H+CH₃(CH₂)₃CH₃. This additional sampling of CH₅ and C₂H₇ data configurations resulted in *no* unphysical high-error configurations in the new samples. Hence, it appears that some limited additional sampling of the C₃H₉ and C₅H₁₃ configuration spaces is necessary and sufficient to obtain data sets for PESs for CH₅ and C₂H₇.

Table II presents the average absolute errors for these new samples of C₃H₉ and C₅H₁₃ configurations. Since the PES and PECS for H+CH₃CH₂CH₃ now contain some (56) scattered data points in addition to the configurations along the minimum energy paths, it is possible to employ the more accurate two-part weight function and confidence lengths to interpolate the PES and PECS. The confidence lengths were evaluated with $M=150$ and $E_{\text{tol}}=1.8 \text{ kJ mol}^{-1}$.⁹

Let us first consider the errors for H+CH₃CH₂CH₃. Table II shows that the PES for H+CH₃CH₂CH₃ with just 143 data points gives an average absolute interpolation error of 4.7 kJ mol^{-1} . The level 1 approximation alone yields an average error of 5.2 kJ mol^{-1} . Inclusion of the level 2 PECS of Eq. (3.8) reduces this error to 3.4 kJ mol^{-1} . These errors are consistent with the following interpretation. Interpolation using the 30 dimensional PES is subject to *unsystematic* errors, with a standard deviation we might roughly estimate as 4.7 kJ mol^{-1} . The level 1+PECS approximation is similarly subject to an *unsystematic* interpolation error with a standard deviation we might estimate as 3.4 kJ mol^{-1} . This standard deviation is a composite of the variances arising from the interpolation of the individual terms in Eq. (3.8). If we denote the standard deviation of the level 2 PECS by σ , then the composite standard deviation of the rhs of Eq. (3.8) is

$$\sigma_{3.8} = [2(2.4)^2 + (1.1)^2 + \sigma^2]^{1/2}, \quad (3.11)$$

where we have estimated the standard deviations for interpolation of the H+CH₄ and H+CH₃CH₃ PESs as the average absolute errors in Table II, 1.1 and 2.4 kJ mol^{-1} , respectively. Setting $\sigma=0$ in Eq. (3.11) yields $\sigma_{3.8}=3.6 \text{ kJ mol}^{-1}$, which is

in close agreement with the measured value of 3.4 kJ mol^{-1} . This suggests that the standard deviation of the level 2 PECS is much less than 2.4 kJ mol^{-1} , which is the standard deviation for the $\text{H}+\text{CH}_3\text{CH}_3$ PES that dominates the value of Eq. (3.11). Hence, σ (the standard deviation of the PECS) is also much less than the average absolute error of the PES for $\text{H}+\text{CH}_3\text{CH}_2\text{CH}_3$ (4.7 kJ mol^{-1}).

It would be possible to reduce the latter error by adding more points to the $\text{H}+\text{CH}_3\text{CH}_3$ PES data set. However, the high dimensionality (30) of the configuration space would ensure that the interpolation error reduces only slowly with the size of the data set. The average error of 5.2 kJ mol^{-1} for the level 1 approximation cannot be reduced to zero by adding more data to the $\text{H}+\text{CH}_4$ and $\text{H}+\text{CH}_3\text{CH}_3$ PESs, since there is a *systematic* component of this error, arising from the neglect of the level 2 correction.

Now let us consider the errors for $\text{H}+\text{CH}_3(\text{CH}_2)_3\text{CH}_3$. With no more *ab initio* calculations, the PES for $\text{H}+\text{CH}_3(\text{CH}_2)_3\text{CH}_3$ can be estimated. As Table II shows, the level 1 error is 12.6 kJ mol^{-1} , reduced with the inclusion of the level 2 PECS to just 5.7 kJ mol^{-1} . Table I suggests that the systematic error in this level 1+level 2 PECS approximation is very small (higher level corrections are negligible, at least at the stationary points on the PES). From Eq. (3.5), the composite standard deviation expected in this case for the level 1+level 2 PECS approximation is

$$\sigma_{3,5} = [4(2.4)^2 + 3(1.1)^2 + 3\sigma^2]^{1/2}. \quad (3.12)$$

Setting $\sigma=0$ in Eq. (3.12) yields $\sigma_{3,5}=5.2 \text{ kJ mol}^{-1}$, which is in close agreement with the measured value of 5.7 kJ mol^{-1} . Again, this suggests that the standard deviation for the level 2 PECS is much less than 2.4 kJ mol^{-1} , which is the standard deviation for the $\text{H}+\text{CH}_3\text{CH}_3$ PES that dominates the value in Eq. (3.12).

IV. DISCUSSION

In light of the above results, we can begin to discuss the feasibility of constructing a PES for the reaction of H with a general linear alkane, $\text{CH}_3(\text{CH}_2)_n\text{CH}_3$. The bonding energy for such a molecule can be expanded in terms of a level 1 fragmentation plus a hierarchy of correction terms, as in Eq. (3.4). As Table I indicates, the magnitude of level 3 correction terms is very small for the alkanes, much less than 1 kJ mol^{-1} . Higher level corrections are much smaller still. The total energy can therefore be estimated accurately using a level 1 fragmentation plus level 2 corrections (ignoring nonbonded interactions for the moment). Thus the energy of $\text{H}+\text{CH}_3(\text{CH}_2)_n\text{CH}_3$ is given by a sum of $(n+1)$ C_2H_7 configuration energies, n CH_5 configuration energies, and level 2 PECS for $(n-2)$ C_3H_9 configurations. Recalling that σ denotes the standard deviation for interpolation of the level 2 PECS, then the expected error in estimating the bonding energy for the alkane is given by $[(n+1)(2.4)^2 + n(1.1)^2 + (n-1)\sigma^2]^{1/2}$. The value of the level 2 PECS varies over a range of only about 20 kJ mol^{-1} for $\text{H}+\text{propane}$. Hence, even an average interpolation error as large as 5% of this range would ensure that σ does not exceed 1 kJ mol^{-1} . The results above for $\text{H}+\text{CH}_3\text{CH}_2\text{CH}_3$ and $\text{H}+\text{CH}_3(\text{CH}_2)_3\text{CH}_3$ support

the view that σ is indeed sufficiently small that it does not contribute significantly to the composite error. The composite interpolation error is therefore dominated by the interpolation error for each C_2H_7 configuration. Hence, we might expect that the interpolation error for $\text{H}+\text{CH}_3(\text{CH}_2)_n\text{CH}_3 \approx 2.4\sqrt{n+1}$. This simple estimate of the composite error assumes that the interpolation errors are independent for each fragment and that the individual interpolation errors are normally distributed. Although these assumptions may not hold exactly, the point here is that interpolation of the level 2 PECS, even using very few data points, is accurate enough to account for the *systematic* level 2 correction to the total PES, so that the residual error is dominated by the interpolation error of the PES for C_2H_7 . By comparison, the average absolute error of a PES for $\text{H}+\text{CH}_3\text{CH}_2\text{CH}_3$, with a data set of only 143 points, is relatively large.

It is difficult to draw firm conclusions about the significance of these estimated errors in PESs for the chemistry of large molecules, since there is no simple map which connects the PES (and errors in the PES) with dynamical observables, such as reaction rates or cross sections. From the arguments above, one might expect that estimates of reaction enthalpies involving $\text{CH}_3(\text{CH}_2)_n\text{CH}_3$ would be subject to an error which increases in proportion to \sqrt{n} . However, the total energy of $\text{CH}_3(\text{CH}_2)_n\text{CH}_3$ must be proportional to n , so the relative error in the enthalpy is decreasing as $1/\sqrt{n}$. Hence, the enthalpy calculation might be considered accurate or inaccurate, depending on the context. Similarly, a chemical process involving $\text{CH}_3(\text{CH}_2)_n\text{CH}_3$ might only involve relatively local changes in the structure of the molecule. The energy difference between two structures of the molecule which are essentially the same structure except in some small region might be estimated with an error that is independent of n . Further investigations will be needed to explore the effect of PES error on many observables of interest. At least such investigations could now be carried out in the context of a homologous series of molecules for which the PES errors vary in a systematic way.

The important inference suggested by the results obtained herein is that interpolation of level 2 PECS with relatively few data points is sufficiently accurate that the accuracy of a large molecule PES is limited by the accuracy of the associated level 1 PESs. If true, this suggests that we can overcome the “tyranny of dimension” which seemed to preclude the construction of high dimensional PES because of the exponential explosion in the number of relevant molecular configurations with the dimension of the configuration space. The bottleneck to construction of accurate PES for large molecules now appears to be the accuracy of PES for the relatively small level 1 fragments.

One appealing feature of constructing the PES for moderately large molecules from level 1 PES plus PECS is the possibility of setting up a library of fragment PES and PECS from which the total PES of many large molecules could be put together in a “Lego-type” fashion. The reactions studied here have revealed one factor which will complicate such an approach. For the $\text{H}+\text{CH}_3\text{CH}_2\text{CH}_3$ and $\text{H}+\text{CH}_3(\text{CH}_2)_3\text{CH}_3$ reactions, it was necessary to append a few additional data points to the PESs for CH_5 and C_2H_7 , to account for level 1

configurations that were not sampled by the trajectories for $H+CH_4$ and $H+CH_3CH_3$ (at least at the energies employed). This suggests that the relevant configuration space for level 1 fragments, here CH_5 and C_2H_7 , may partly depend on the molecular context in which they are used. For example, we could use the PES and PECS presented above to study the reaction of a H atom with cyclohexane at the level 2 fragmentation approximation. However, it may be that there are important configurations of the fragments in $H+cyclohexane$ that have not been adequately sampled in reactions involving the isolated fragments. Further study is required to establish an efficient approach to sampling the configuration space which is relevant to reactions of the whole molecules.

Nevertheless, the results obtained here indicate that the level 1+PECS approach to large molecule PESs is feasible, at least for the n -alkanes. In order to construct an accurate PES for the (bonding) energy of a large reacting molecule, we must accurately construct PESs for small fragments such as C_2H_7 , and in addition, we need only estimate PECS, for systems such as C_3H_9 , using a relatively small number of data points.

Of course, for larger molecules, the magnitude of the non-bonded energy must become larger and more chemically significant. An accurate but efficient method to estimate such non-bonded PESs is the focus of current research.

ACKNOWLEDGMENT

The author wishes to acknowledge discussions with Vitali Deev in the early stages of this work.

¹W. Yang, Phys. Rev. A **44**, 7823 (1991); Phys. Rev. Lett. **66**, 1438 (1991); W. Yang and T. Lee, J. Chem. Phys. **103**, 5674 (1995); S. Liu, J. M. Pérez-Jordá, and W. Yang, *ibid.* **112**, 1634 (2000); N. Flocke and R. J. Bartlett, *ibid.* **121**, 10935 (2004); T. Yanai, G. I. Fann, Z. Gan, R. J.

Harrison, and G. Beylkin, *ibid.* **121**, 6680 (2004); C. K. Gan, G. J. Tymczak, and M. Challacombe, *ibid.* **121**, 6608 (2004); X. Li, J. M. Millam, G. E. Scuseria, M. J. Frisch, and H. B. Schlegel, *ibid.* **119**, 7651 (2003); G. E. Scuseria, J. Chem. Phys. A **103**, 4782 (1999); W. Liang, C. Saravanan, Y. Shao, R. Baer, A. T. Bell, and M. Head-Gordon, J. Chem. Phys. **119**, 4117 (2003); H.-J. Werner, F. R. Manby, and P. J. Knowles, *ibid.* **118**, 8149 (2003); P. Ordejon, Comput. Mater. Sci. **12**, 157 (1998); S. Goedecker, Rev. Mod. Phys. **71**, 1087 (1999); S. Dixon and K. Merz, Jr., J. Chem. Phys. **104**, 6643 (1996); W. Li and S. Li, *ibid.* **122**, 194109 (2005); W. Li, T. Fang, and S. Li, *ibid.* **124**, 154102 (2006); D. G. Fedorov and K. Kitaura, J. Chem. Phys. **121**, 2483 (2004); Chem. Phys. Lett. **389**, 129 (2004).

²D. W. Zhang and J. Z. H. Zhang, J. Chem. Phys. **119**, 3599 (2003); Y. Mei, D. W. Zhang, and J. Z. H. Zhang, J. Phys. Chem. A **109**, 2 (2005).

³X. H. Chen, D. W. Zhang, and J. Z. H. Zhang, J. Chem. Phys. **120**, 839 (2004); X. H. Chen and J. Z. H. Zhang, *ibid.* **120**, 11386 (2004); D. W. Zhang and J. Z. H. Zhang, J. Theor. Comput. Chem. **3**, 43 (2004).

⁴S. R. Gadre, R. N. Shirsat, and A. C. Limaye, J. Phys. Chem. **98**, 9165 (1994); K. Babu and S. R. Gadre, J. Comput. Chem. **24**, 484 (2003).

⁵X. He and J. Z. H. Zhang, J. Chem. Phys. **122**, 031103 (2005); X. Chen, Y. Zhang, and J. Z. H. Zhang, *ibid.* **122**, 184105 (2005).

⁶S. Li, W. Li, and T. Fang, J. Am. Chem. Soc. **127**, 7215 (2005).

⁷V. Deev and M. A. Collins, J. Chem. Phys. **122**, 154102 (2005).

⁸M. A. Collins and V. A. Deev, J. Chem. Phys. **125**, 104104 (2006).

⁹R. P. A. Bettens and M. A. Collins, J. Chem. Phys. **111**, 816 (1999).

¹⁰K. C. Thompson, M. J. T. Jordan, and M. A. Collins, J. Chem. Phys. **108**, 8302 (1998); K. C. Thompson and M. A. Collins, J. Chem. Soc., Faraday Trans. **93**, 871 (1997); M. J. T. Jordan, K. C. Thompson, and M. A. Collins, J. Chem. Phys. **102**, 5647 (1995); J. Ischtwan and M. A. Collins, *ibid.* **100**, 8080 (1994).

¹¹M. A. Collins, Theor. Chem. Acc. **108**, 313 (2002).

¹²D. H. Zhang, M. A. Collins, and S.-Y. Lee, Science **290**, 961 (2000); D. H. Zhang, M. Yang, M. A. Collins, and S.-Y. Lee, Proc. Natl. Acad. Sci. U.S.A. **99**, 11579 (2002).

¹³A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

¹⁴T. H. Dunning and K. A. Peterson, in *Encyclopedia of Computational Chemistry*, edited by P. v. R. Schleyer (Wiley, Chichester, 1998), Vol. 1, p. 88.

¹⁵G. W. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Gaussian, Inc., Pittsburgh, PA, 2003.

¹⁶P. R. Bunker, *Molecular Symmetry and Spectroscopy* (Academic, New York, 1979).