JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

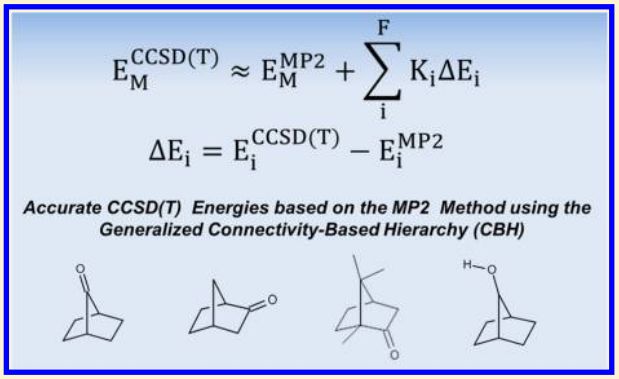
Extrapolation to the Gold-Standard in Quantum Chemistry: Computationally Efficient and Accurate CCSD(T) Energies for Large Molecules Using an Automated Thermochemical Hierarchy

Raghunath O. Ramabhadran and Krishnan Raghavachari*

Department of Chemistry, Indiana University, Bloomington, Indiana 47405, United States

Supporting Information

ABSTRACT: The CCSD(T) method is known as the gold-standard in quantum chemistry and has been the method of choice in quantum chemistry for over 20 years to obtain accurate bond energies and molecular properties. Its computational cost formally scales as the seventh power of the size of the system and can be prohibitive for large molecules. As part of our efforts to reduce the computational cost of the CCSD(T) method yet retain its accuracy, we present a simple, efficient, and user-friendly protocol to extrapolate to CCSD(T) energies in conjunction with MP2 energies. The method is based on the automated error-canceling thermochemical hierarchy previously developed by us called the Connectivity-Based Hierarchy (CBH). For a test set containing 30 diverse nonaromatic organic molecules and biomonomers, we obtain highly accurate extrapolated CCSD(T) energies (with a mean absolute error of only 0.2−0.3 kcal/mol with different basis-set). Additionally, the work also features the successful extrapolation to CCSD energies using a similar protocol.



1. INTRODUCTION

Increasing the accuracy and computational applicability of electronic structure calculations has been a primary objective in computational quantum chemistry for decades.[1] Starting with the Hartree−Fock method,[2] the development of Moller−Plesset (MP*n*) perturbation methods,[3] and a range of progressively more sophisticated electron correlation methods[4] culminating in the highly accurate coupled cluster methods[5] (CC), aids in the accurate determination of molecular energies.[1,6] However, the steep computational scaling of these accurate high level methods (for example, ~N^7 scaling for CCSD(T) method) renders them to be prohibitively expensive for medium sized and large molecules.[1]

Fragmentation-based methods offer significant promise to perform high level *ab initio* computations on medium sized and larger molecules.[7−21] The general idea in all the fragment-based methods is to divide a large molecule into small fragments, perform electronic structure calculations on each fragments, and assemble them to yield the energy of the large molecule.[7−21] Prominent fragment-based methods include (a) the fragment molecular orbital method (FMO) of Kitura,[7] (b) Gadre's cardinality-guided molecular tailoring approach (CG-MTA),[8] (c) the molecular fractionation with conjugate caps (MFCC) method of Zhang,[9] (d) the systematic molecular fragmentation (SMF) of Collins,[10] (e) the effective fragment potential (EFP) method of Gordon,[11] (f) Truhlar's electrostatically embedded many-body method (EE-MB),[12] (g) the generalized energy-based fragmentation method (GEBF) of Li,[13] (h) the kernel energy method (KEM) of Karle,[14] (i) the hybrid many-body interaction (HMBI) method of Beran,[15] (j) the multilevel fragment-based approach (MFBA) of Salahub,[16] (k) the XPOL and multilevel X-pol methods of Gao,[17] (j) the generalized many-body expansion of Herbert (GMBE),[18] (l) the three-body:many-body integrated fragmentation and the two-body/many-body QM:QM fragmentation methods of Tshumper,[19] and (m) the molecules-in-molecules (MIM)[20] and the many-overlapping-body (MOB)[21] methods from our own group.

Theoretical thermochemistry is another dominant area of research in quantum chemistry.[1,6] Error-canceling reaction schemes in theoretical thermochemistry permit the use of modest levels of theory routinely employed by both experts and nonexperts in quantum chemistry to accurately compute the enthalpies of formations of organic molecules. The isodesmic-bond separation reaction scheme, introduced by Pople and coworkwers in 1970 was the first efficient error-canceling scheme.[22] Subsequently, several clever schemes have been proposed to improve upon the isodesmic-bond separation reaction scheme.[23] However, as Wheeler, Houk, Schlyer, and Allen pointed out in 2009,[24] widespread definition-based inconsistencies in these reaction schemes had for a long time prevented the development of an automated, well-defined, and reliable error-canceling thermochemical hierarchy applicable to all classes of organic molecules—hydrocarbons as well as nonhydrocarbons. We developed the generalized connectivity-

Received: June 4, 2013
Published: July 26, 2013

based hierarchy (CBH) in 2011 to solve this long-standing problem. CBH is the first automated error canceling thermochemical hierarchy to obtain accurate enthalpies of formations of organic molecules using modest levels of theories.[25]

Automated error-canceling thermochemical hierarchies can be thought as being analogous to the fragment based-methods, but with key differences. Typically in fragment based methods, the geometries of all the fragments are extracted as in the same geometry of the parent molecule. In contrast, in a thermochemical hierarchy using experimental reference values, all the fragments obtained need to be optimized in their equilibrium geometries. Since many large molecules share the same smaller fragments, it is not necessary to repeatedly carry out electronic structure calculations. This feature significantly reduces the total number of electronic structure calculations needed to be performed as part of a thermochemical hierarchy such as CBH, relative to the various fragment-based methods. In this paper, we bridge the gap between thermochemical hierarchies and fragment-based methods.

Herein, we use the connectivity-based hierarchy (CBH) to extrapolate to accurately obtain the CCSD(T) energies for organic compounds and biomonomers. The CCSD(T) method,[5u] which iteratively includes single and double substitutions and perturbatively incorporates the triple substitutions, is known to be very successful in obtaining accurate energies and molecular properties.[1,6,26] However, the CCSD(T) method formally scales with the number of basis sets as $O(N^7)$. Therefore, it is very much desirable to reduce the cost involved in a CCSD(T) calculation while maintaining the impeccable accuracy it offers. We found (*vide infra*) that the MP2 method[3a,27] is an excellent starting point for extrapolation[28] to CCSD(T). MP2 is the simplest *ab initio* method to estimate the electron correlation energy, and it formally scales as the fifth power of the size of the system,[6] (the practical scaling is much better) thus providing a lucrative basis for our extrapolation. Along with obtaining accurate CCSD(T) energies (accurate to about 0.19 kcal/mol with the aug-cc-pvDZ basis-set for a test set of 30 varied organic molecules), we herein also extrapolate to the reasonably accurate CCSD method (formal scaling ~ N^6) again based on the MP2 method to demonstrate the robustness of our procedure.

The rest of the paper is organized as follows. We first present our methodology in section 2. This is followed by a critical evaluation of our method in section 3 and, finally, a discussion of the scope of our method in section 4.

2. METHODOLOGY

(a). A Brief Account of CBH. Herein, we present only the central ideas involved in constructing CBH. Full details regarding the construction of the hierarchy are provided elsewhere.[25]

CBH has several rungs denoted CBH-1, CBH-2, CBH-3, etc. CBH-1 is formally the same as the isodesmic bond separation scheme. At the isoatomic CBH-2 rung, the immediate chemical bonding environment of all the heavy atoms is preserved in the reaction scheme. At CBH-3, the immediate chemical environment of all the heavy atom bonds is preserved. Additional higher rungs can also be defined in a similar manner, if required. Thus, it is readily seen that the basic idea behind the construction of CBH is to increasingly preserve the chemical environment (i.e., the atom-connectivity) in a molecule on ascending up the rungs of the hierarchy leading to superior error-cancellation.

In the context of theoretical thermochemistry,[25] once the automated reaction schemes are obtained, we can compute the reaction energies at any given level of theory. Then, on the basis of the Hess's law, using the computed reaction energies and experimental reference values, accurate enthalpies of formations are calculated.

(b). Physical Basis of our Method. In the context of the present work, once we derive the automated CBH reaction schemes and obtain the reaction energies at any given level of theory, we first write down the energy of the larger molecule in terms of the smaller fragments generated at any particular rung of CBH. This is clearly illustrated using the example of methionine (Figure 1), an amino-acid for which we previously accurately computed the enthalpy of formation using CBH.[25c]

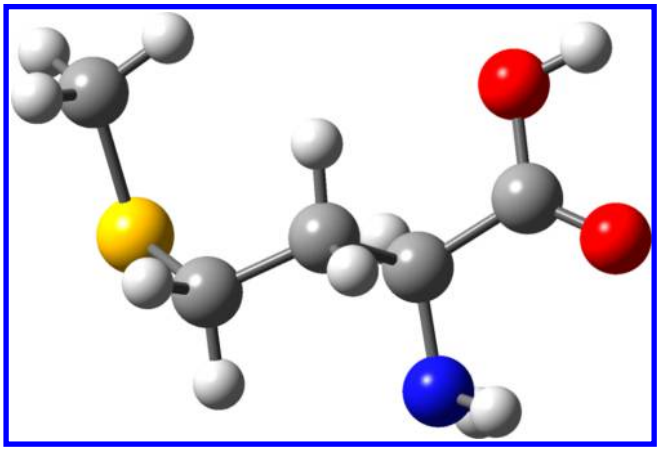


Figure 1. Ball and stick representation of methionine, used to illustrate the construction of CBH. The enthalpy of formation of methionine was previously accurately computed using CBH.[25c]

CBH-1 reaction scheme:

$$C_5H_{11}NO_2S + 6CH_4 + H_2S \rightarrow 3C_2H_6 + H_3CNH_2 + 2H_3CSH + H_2CO + H_3COH \quad (1)$$

CBH-2 reaction scheme:

$$C_5H_{11}NO_2S + 3C_2H_6 + H_3CSH \rightarrow C_2H_5SH + C_3H_9N\ (\text{isopropyl amine}) + H_3CCOOH + C_3H_8 + H_3CSCH_3\ (\text{dimethyl sulfide}) \quad (2)$$

CBH-3 reaction scheme:

$$C_5H_{11}NO_2S + C_3H_8 + C_2H_5SH + C_3H_9N\ (\text{isopropyl amine}) \rightarrow C_3H_7NO_2\ (\text{alanine}) + C_4H_{11}N\ (\text{2-butylamine}) + C_3H_7SH\ (\text{1-propanethiol}) + C_3H_8S\ (\text{ethyl methyl sulfide}) \quad (3)$$

Based on previous experience,[25] CBH-3 usually suffices for the organic molecules considered in this study, as, at CBH-4, the size of the fragments increases even further. Thus, we limit the construction of our hierarchy to CBH-3 in this study.

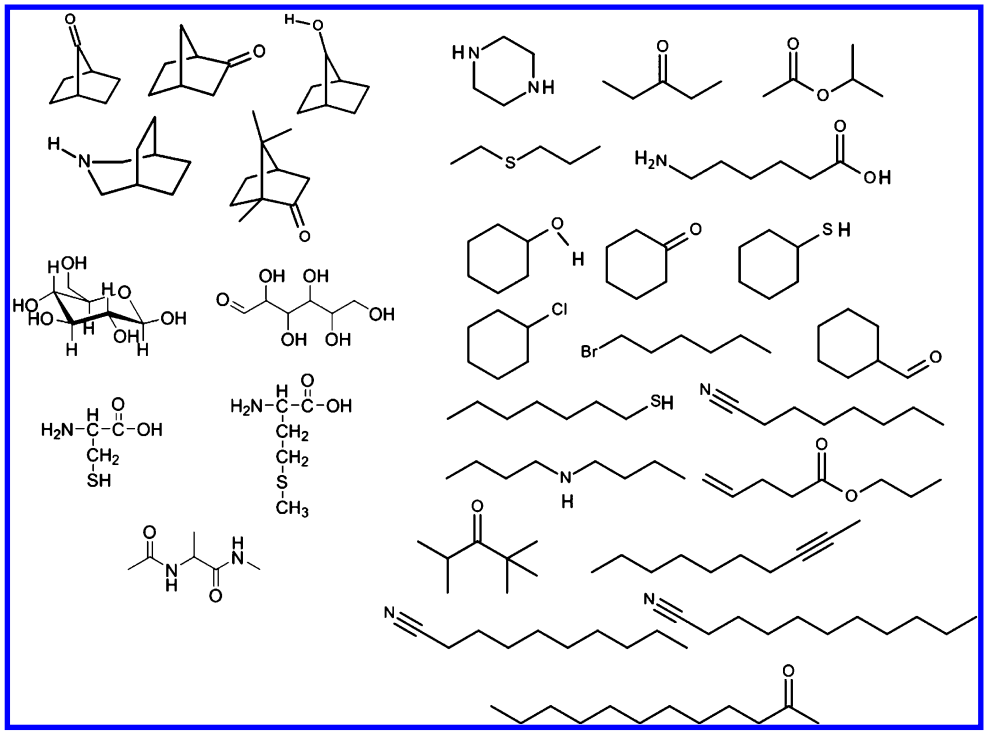


Figure 2. Structures of all the organic compounds used in our diverse 30 molecule test set.

Next, generalizing such reaction schemes for any generic molecule M at any given *n*th rung of CBH, we have

$$\text{Reaction energy (CBH-}n) = \sum_{i}^{F} K_i E_i - E_M \tag{4}$$

where E_M is the energy of the full molecule M at a given level of theory, E_i is the energy of the *i*th fragment (each reactant and product) at the same level of theory, and K_i is the signed stoichiometric coefficient of the *i*th fragment. (i.e., K_i is a positive integer if *i* is a product fragment and is a negative integer if *i* is a reactant fragment). The summation runs over the total number of fragments generated from molecule M.

If we specifically choose the CCSD(T) and the MP2 methods for any given basis set, based on our previous works,[25b] we have observed that, at CBH-2 and higher rungs, for a wide variety of organic molecules (6−13 heavy atoms), superior error cancellation yielded very similar reaction energies for the CCSD(T) and the MP2 methods.

$$\text{Reaction Energy}^{\text{CCSD(T)}}_{\text{CBH-}n} \approx \text{Reaction Energy}^{\text{MP2}}_{\text{CBH-}n} \tag{5}$$

Therefore,

$$E_M^{\text{CCSD(T)}} \approx E_M^{\text{MP2}} + \sum_{i}^{F} K_i \Delta E_i \tag{6}$$

where

$$\Delta E_i = E_i^{\text{CCSD(T)}} - E_i^{\text{MP2}} \tag{7}$$

Thus, we exploit the similar reaction energies noticed between the MP2 and CCSD(T) methods (and the similar error-cancellations in high level thermochemical hierarchies) for extrapolation using the additivity approximation implicit in eq 6.[28] That is, the MP2 method is a suitable starting point for approximating to the highly accurate CCSD(T) energies and the CCSD(T) energy of a larger molecule M can be accurately obtained (vide infra) without having to perform the expensive CCSD(T) calculation on M. The bottleneck CCSD(T) calculation now involves the largest fragment generated in the selected CBH scheme.

As an example, using the CBH-2 rung for methionine (9 heavy atoms), the most expensive CCSD(T) calculation is that for isopropylamine (4 heavy atoms). Similarly, at the CBH-3 rung, the most expensive CCSD(T) calculation is that for 2-butylamine (5 heavy atoms). Since CCSD(T) scales as the seventh power of the size of the system, it is straightforward to note that there is an enormous reduction in the computational expense. As the molecule gets larger, the computational savings increase dramatically since the largest fragment size is independent of the size of the full molecule, though the number of fragments increases. In the next section, we will see that this drastic reduction in the computational expense hardly affects the accuracy of the CCSD(T) energy obtained.

Besides extrapolation to get CCSD(T) energies, in the next section, we shall successfully extrapolate to get CCSD energies of organic molecules as well. Thus,

$$E_M^{\text{CCSD}} \approx E_M^{\text{MP2}} + \sum_{i}^{F} K_i \Delta E_i \tag{8}$$

where

$$\Delta E_i = E_i^{\text{CCSD}} - E_i^{\text{MP2}} \tag{9}$$

(c). Overall Protocol. To summarize the overall protocol involved in obtaining extrapolated CCSD(T) (or CCSD) energies ($E_M^{\text{CCSD(T)}}$ or E_M^{CCSD}), for a generic organic molecule M at any given basis-set at the selected CBH-n rung is very easy and is as follows:

(i) Generate the CBH reaction scheme for M and obtain the fragments.

(ii) Obtain the equilibrium geometries of M and the fragments at a reasonable level of theory (*vide infra* for

Table 1. Listing the Errors (kcal/mol) between the Full CCSD(T) Energies and CCSD(T) Energies Obtained by Extrapolation in this Work[a]

molecular formula	chemical name	CBH-2 631+G(d,p)	CBH-3 6-31+G(d,p)	CBH-2 aug-cc-pVDZ	CBH-3 aug-cc-pVDZ
$C_4H_{10}N_2$	piperezine	0.36	0.18	0.40	0.12
$C_5H_{10}O$	3-pentanone	−0.07	−0.01	−0.10	0.00
$C_5H_{10}O_2$	isopropyl acetate	−0.05	−0.01	0.00	−0.01
$C_5H_{12}S$	ethyl propyl sulfide	0.03	−0.05	0.08	−0.07
$C_6H_{13}NO_2$	6-aminohexanoic acid	−0.15	−0.04	−0.24	−0.10
$C_6H_{12}O$	cyclohexanol	0.37	0.30	0.29	0.14
$C_6H_{10}O$	cyclohexanone	0.28	0.38	0.12	0.22
$C_6H_{12}S$	cyclohexanethiol	0.56	0.33	0.46	0.17
$C_6H_{11}Cl$	cyclohexyl chloride	0.49	0.42	0.36	0.24
$C_6H_{13}Br$	*n*-hexyl bromide	−0.02	−0.05	−0.04	−0.08
$C_7H_{12}O$	cyclohexanal	0.38	0.45	0.19	0.26
$C_7H_{16}S$	1-heptanethiol	−0.08	−0.10	−0.06	−0.15
$C_8H_{15}N$	octanenitrile	−0.04	−0.01	−0.06	−0.07
$C_8H_{19}N$	dibutylamine	−0.11	0.18	−0.03	−0.16
$C_8H_{14}O_2$	propyl pent-4-enoate	−0.26	−0.11	−0.4	−0.23
$C_8H_{16}O$	t-butyl isopropyl ketone	0.08	0.04	−0.06	0.09
$C_{10}H_{18}$	2-decyne	0.07	0.01	0.00	−0.07
$C_{10}H_{19}N$	caprinitrile	−0.11	−0.06	−0.12	−0.13
$C_{11}H_{21}N$	1-cyanodecane	−0.14	−0.08	−0.15	−0.16
$C_{12}H_{24}O$	decyl methyl ketone	−0.31	−0.25	−0.29	−0.25
$C_{10}H_{16}O$	camphor	1.60	1.31	1.21	0.81
$C_6H_{12}O_6$	glucose (ring)	−0.63	−0.04	−0.74	−0.04
$C_6H_{12}N_2O_2$	alanine dipeptide	−0.21	0.06	−0.31	0.09
$C_6H_{12}O_6$	glucose (open)	−0.25	0.14	−0.5	0.06
$C_3H_7NO_2S$	cysteine	0.28	0.30	0.12	0.28
$C_5H_{11}NO_2S$	methionine	0.00	0.10	−0.09	0.06
$C_7H_{12}O$	7-norbornanol	0.98	0.85	0.74	0.47
$C_7H_{10}O$	2-norbornanone	1.16	1.11	0.83	0.76
$C_7H_{10}O$	7-norbornanone	0.87	0.79	0.46	0.41
$C_8H_{15}N$	3-azabicyclo [3.2.2] nonane	0.64	0.39	0.52	0.07
mean absolute error		0.35	0.27	0.30	0.19

[a]Error is defined as the energy obtained with full calculation − energy obtained using an extrapolated approach.

computational details). Since many large molecules share the same fragments, this step can be avoided for fragments that have been optimized previously. Even though we use the fragments in their equilibrium geometries, we find that there is no loss of accuracy (*vide infra*, next section).

(iii) Perform an MP2 calculation on the full molecule, M to obtain E_M^{MP2}

(iv) Perform CCSD(T) (or CCSD) calculations on the fragments to get ΔE_i (defined in eq 7 for CCSD(T) and in eq 9 for CCSD). Since MP2 and CCSD energies are obtained as subsets of a CCSD(T) calculation, only a single calculation on each fragment is needed. In addition, since the same fragments are present for many molecules, repeated calculations are easily avoided with a look-up table.

(v) Finally, use eq 6 to get the extrapolated CCSD(T) energy, $E_M^{CCSD(T)}$, (or eq 8 to get the extrapolated CCSD energy, E_M^{CCSD}).

(d). Computational Details. All the electronic structure calculations have been performed using the Gaussian 09 suite of programs.[29] Prior to obtaining CCSD(T) and CCSD energies based on the MP2 method, we use the B3LYP/6-31G(2df,p) level of theory to obtain the geometries as in the popular G4 theory.[30] Throughout the paper, we intentionally use modest sized double-ζ quality basis sets (one Pople style, 6-31+g(d,p), and one Dunning style, aug-cc-pVDZ) for the following reasons: (a) Our goal in this paper is to propose an accurate method easily accessible for large molecules, and these double-ζ basis-sets satisfy this requirement. Therefore, using these basis-sets demonstrates the accuracy and applicability of our method. Besides, previously in the context of thermochemistry, such basis sets are usually quite good in deriving accurate enthalpies of formation due to superior error cancellation.[25] (b) Given the expensive scaling of CCSD(T), we found out that it is usually computationally expensive (and in some cases prohibitive) to perform such reference calculations using triple-ζ quality basis-sets such as 6-311++G(3df,2p) on some of the larger molecules in our test set. However, our extrapolation method is very robust and is not basis-set dependent. To prove this, we include the excellent performance of our method with the 6-311++G(3df,2p) basis-set in the Supporting Information for a subset of these molecules (those with 8 or fewer number of heavy atoms) where reference CCSD(T) calculations were possible.

(e). CBH and MIM. It is worthwhile to mention here that, at a broad level, CBH[25] shares a few similar features with general fragment based methods such as MIM,[20] also developed in our group. For instance, $E_M^{CCSD(T)}$ obtained at CBH-2 is conceptually similar to the E^{MIM2}(CCSD(T):MP2) energy obtained using one specific way (out of many different possibilities) of subsystem formation in MIM. This similarity is however limited to CBH-2 and MIM2, and there are notable

Table 2. Mean Absolute Errors (kcal/mol) between the Full CCSD(T) Energies and CCSD(T) Energies Obtained by Extrapolation in This Work for Test Sets A and B[a,b]

test set	CBH-2 6-31+G(d,p)	CBH-3 6-31+G(d,p)	CBH-2 aug-cc-pVDZ	CBH-3 aug-cc-pVDZ
A(20 molecules)	0.19	0.15	0.17	0.14
B(10 molecules)	0.66	0.51	0.55	0.30

[a]Error is defined as the energy obtained with full calculation − energy obtained using an extrapolated approach. [b]Test set A corresponds to the first 20 molecules (piperezine through decyl methyl ketone) in Table 1, and test set B corresponds to the next 10 molecules (camphor through 3-azabicyclo [3.2.2] nonane).

differences in most other aspects. Most prominently, all the fragments in the MIM method are extracted at the same geometry as the parent molecule. However, since CBH was originally designed for accurate thermochemistry and used experimental reference values for the fragments generated, all the fragments obtained using CBH were optimized at their equilibrium geometries. In this study as well, we continue to use the fragments in their equilibrium geometries and yet find that there is no loss of accuracy (*vide infra*, next section). The other major difference arises from the fact that, due to the lack of appropriate link atoms for double bonds, they are not cleaved in MIM. In CBH, however, since we optimize all the hydrogen terminated fragments to their equilibrium geometries, we do not worry about the link atoms and cleave the double bonds as well. Thus, the fragments obtained from MIM and CBH can be significantly different.

3. CRITICAL EVALUATION OF THE METHOD

(a). Test Set. The first step in the assessment of our method involves assembling a credible test set that represents the enormous structural variety of organic molecules. Our diverse 30 molecule test set comprises of 20 organic molecules used in references 25a and 25b, 5 molecules possessing ring strain (7-norbornanone, 2-norbornanone, 7-norbornanol, camphor, and 3-azabicyclo[3.2.2.] nonane), 4 biomonomers (cysteine, methionine, glucose (ring), glucose (open)), and alanine dipeptide. The structures of all the molecules used in our test set are provided in Figure 2. The test set was chosen such that it had molecules containing different functional groups (alcohol, ether, aldehyde, ketone etc.), different hetero atoms (N, O, C, S, Cl, Br etc.), different numbers of hetero atoms in the same molecule to test the efficacy of error-cancellations (up to a maximum of 6 hetero atoms), different molecular architectures (cyclic vs acyclic, branched vs linear etc.), biologically relevant molecules (amino acids, sugars, and dipeptides), and molecules possessing certain special features such as ring strain (camphor, etc.). The minimum number of heavy atoms present in the molecules in our test set is 6 and the maximum number 13.

(b). Performance of the Method—Extrapolated CCSD(T) Energies. Columns 3 and 4 in Table 1 list the errors (full CCSD(T)/6-31+G(d,p) − extrapolated CCSD(T)/6-31+G(d,p) energy) for all the 30 molecules in our test set with the CBH-2 and CBH-3 rungs. With a mean absolute error of 0.35 kcal/mol, it can be readily seen that the method performs very well at CBH-2. An incremental but clear improvement is noticed at CBH-3 as the mean absolute error falls to 0.27 kcal/mol. It must, however, be stressed that the incremental improvement in performance at CBH-3 over CBH-2 comes at a computational cost as the fragments involved in CBH-3 are larger.

It is extremely useful to divide the 30 molecule test set into 2 different test sets—set A and set B—to further gauge the performance of the method. Test set A contains the first 20 molecules listed in Table 1 (piperezine through decyl methyl ketone), which are common organic molecules. Test set B contains the next 10 molecules (camphor through 3-azabicyclo [3.2.2] nonane) in Table 1, which are molecules containing certain special features (such as ring strain, possessing 3 or more heteroatoms, etc.).

Glancing at the individual performance for the molecules (Table 1) and the mean absolute error (Table 2) clearly shows that *our method works with impressive accuracy for the common organic molecules* in test set A. The numbers for test set B (Tables 1 and 2) indicate very good performance as well, but the errors are relatively larger than those seen with test set A (mean absolute errors of 0.19 vs 0.66 kcal/mol at CBH-2 and 0.15 vs 0.51 kcal/mol at CBH-3, respectively, Table 2). This is to be expected since the more complicated molecules in test set B might not provide as sophisticated an error-cancellation as the simpler molecules in test set A.

Among all the molecules in test set B, camphor results in the maximum error of about 1.5 kcal/mol at CBH-2 and, notably,

Table 3. Listing of the Errors between the Full CCSD(T) Energies and CCSD(T) Energies Obtained by Extrapolation Using CBH-1 (Column 3) and the Errors between the Full CCSD Energies and CCSD Energies Obtained by Extrapolation Using CBH-1 (Column 4)[a,b]

molecular formula	chemical name	error for CCSD(T)	error for CCSD
$C_4H_{10}N_2$	piperezine	1.93	3.21
$C_5H_{10}O$	3-pentanone	1.29	2.38
$C_5H_{10}O_2$	isopropyl acetate	2.81	4.90
$C_5H_{12}S$	ethyl propyl sulfide	1.33	2.43
$C_6H_{13}NO_2$	6-aminohexanoic acid	2.95	5.21
$C_6H_{12}O$	cyclohexanol	2.09	3.69
$C_6H_{10}O$	cyclohexanone	2.36	3.91
$C_6H_{12}S$	cyclohexanethiol	2.69	4.62
$C_6H_{11}Cl$	cyclohexyl chloride	2.55	4.36
$C_6H_{13}Br$	*n*-hexyl bromide	1.36	2.68
$C_7H_{12}O$	cyclohexanal	2.79	4.75
$C_7H_{16}S$	1-heptanethiol	1.50	3.07
$C_8H_{15}N$	octanenitrile	1.90	3.40
$C_8H_{19}N$	dibutylamine	1.63	3.37
$C_8H_{14}O_2$	propyl pent-4-enoate	3.68	6.25
$C_8H_{16}O$	t-butyl isopropyl ketone	3.14	6.33
$C_{10}H_{18}$	2-decyne	2.77	4.55
$C_{10}H_{19}N$	caprinitrile	2.31	4.33
$C_{11}H_{21}N$	1-cyanodecane	2.51	4.79
$C_{12}H_{24}O$	decyl methyl ketone	2.72	5.58

[a]Error is defined as the energy obtained with full calculation − energy obtained using an extrapolated approach. [b]6-31+G(d,p) basis set was used throughout.

Table 4. Errors (kcal/mol) between the Full CCSD Energies and CCSD Energies Obtained by Extrapolation in This Work[a]

molecular formula	chemical name	CBH-2 6-31+G(d,p)	CBH-3 6-31+G(d,p)	CBH-2 aug-cc-pVDZ	CBH-3 aug-cc-pVDZ
$C_4H_{10}N_2$	piperezine	0.15	−0.11	0.17	−0.09
$C_5H_{10}O$	3-pentanone	0.11	0.01	0.11	0.01
$C_5H_{10}O_2$	isopropyl acetate	0.25	0.09	0.41	0.09
$C_5H_{12}S$	ethyl propyl sulfide	0.20	−0.06	0.26	−0.08
$C_6H_{13}NO_2$	6-aminohexanoic acid	0.22	0.07	0.18	0.02
$C_6H_{12}O$	cyclohexanol	0.35	0.01	0.27	−0.26
$C_6H_{10}O$	cyclohexanone	0.34	0.13	0.15	−0.16
$C_6H_{12}S$	cyclohexanethiol	0.67	0.01	0.57	−0.32
$C_6H_{11}Cl$	cyclohexyl chloride	0.58	0.08	0.46	−0.26
$C_6H_{13}Br$	*n*-hexyl bromide	0.24	0.02	0.31	0.02
$C_7H_{12}O$	cyclohexanal	0.55	0.12	0.37	−0.18
$C_7H_{16}S$	1-heptanethiol	0.26	−0.04	0.35	−0.07
$C_8H_{15}N$	octanenitrile	0.28	0.08	0.32	0.04
$C_8H_{19}N$	dibutylamine	0.11	0.08	0.28	−0.07
$C_8H_{14}O_2$	propyl pent-4-enoate	0.23	0.01	0.18	−0.14
$C_8H_{16}O$	t-butyl isopropyl ketone	0.90	0.24	0.90	0.31
$C_{10}H_{18}$	2-decyne	0.51	0.15	0.53	0.11
$C_{10}H_{19}N$	caprinitrile	0.35	0.09	0.44	0.05
$C_{11}H_{21}N$	1-cyanodecane	0.74	0.09	0.49	0.05
$C_{12}H_{24}O$	decyl methyl ketone	0.28	−0.05	0.44	0.03
$C_{10}H_{16}O$	camphor	2.33	−0.24	2.00	−1.30
$C_6H_{12}O_6$	glucose (ring)	−0.17	−1.05	−0.42	−1.37
$C_6H_{12}N_2O_2$	alanine dipeptide	0.82	0.34	0.75	0.27
$C_6H_{12}O_6$	glucose (open)	1.07	0.29	0.86	0.15
$C_3H_7NO_2S$	cysteine	1.24	0.80	1.14	0.80
$C_5H_{11}NO_2S$	methionine	0.74	0.40	0.71	0.39
$C_7H_{12}O$	7-norbornanol	0.91	−0.24	0.67	−0.91
$C_7H_{10}O$	2-norbornanone	1.19	0.13	0.83	−0.52
$C_7H_{10}O$	7-norbornanone	0.89	−0.27	0.43	0.97
$C_8H_{15}N$	3-azabicyclo [3.2.2] nonane	0.76	−0.22	0.70	−0.77
mean absolute error		0.58	0.19	0.52	0.33

[a]Error is defined as the energy obtained with full calculation − energy obtained using an extrapolated approach.

Table 5. Mean Absolute Errors (kcal/mol) between the Full CCSD Energies and CCSD Energies Obtained by Extrapolation in This Work for Test Sets A and B[a,b]

test set	CBH-2 6-31+G(d,p)	CBH-3 6-31+G(d,p)	CBH-2 aug-cc-pVDZ	CBH-3 aug-cc-pVDZ
A (20 molecules)	0.37	0.08	0.36	0.13
B (10 molecules)	1.01	0.40	0.85	0.74

[a]Error is defined as the energy obtained with full calculation − energy obtained using an extrapolated approach. [b]Test set A corresponds to the first 20 molecules (piperezine through decyl methyl ketone) in Table 1, and test set B corresponds to the next 10 molecules (camphor through 3-azabicyclo [3.2.2] nonane).

only a modest improvement resulting in a 1.3 kcal/mol error at CBH-3. 2-norbnornanone is the molecule with the next largest errors (Table 1). It is interesting to note that, between the isomers 2-norbornanone and 7-norbornanone, 2-norbornanone, wherein one ring is more strained than the other (due to the asymmetric positioning of the carbonyl functional group), results in larger errors than 7-norbornanone, wherein the ring strain is equally shared between the two rings (Table 1). It is very likely that the slightly larger errors observed in the cases of molecules such as camphor arise because the fragments are not maintained in the same geometry as the strained molecule itself. This interesting aspect warrants careful scrutiny in the future.

The presence of a large number of the same heteroatoms in an organic molecule presents the possibility of accumulation of errors rather than error cancellation. With this in mind, the excellent performance of the ring form of glucose as well as the open form of glucose with both CBH-2 and CBH-3 (Table 1) is particularly satisfying given the fact that these molecules contain six heteroatoms of the same kind (oxygen). Finally, for the amino acids (cysteine and methionine) as well as the alanine peptide, our method offers accurate CCSD(T) energies at a drastically reduced computational cost.

The method readily works with other basis-sets as well. Extrapolated CCSD(T)/aug-cc-pVDZ energies and the errors associated are also provided in Table 1 for all the 30 molecules in the test set (columns 5 and 6). Throughout, it is noticed the performance is enhanced with the Dunning style basis sets relative to the Pople-style basis sets. Additionally, larger basis set CCSD(T)/6-311++G(3df,2p) calculations were carried out for many of the molecules in the test set. The performance is very similar, as documented in the SI.

Finally, Table 3 contains the CBH-1 based extrapolated CCSD(T) energies that yield much larger errors (mean

absolute error for just the test set A alone ~2−3 kcal/mol). This is evidence that eq 9 is valid only at CBH-2 and higher rungs. A very important point to note here is that the CBH-1 is formally the same as the isodesmic-bond separation reaction scheme that has been used for over 40 years in quantum chemistry. In conjunction with our previous thermochemistry results,[25] Table 3 presents conclusive evidence that error cancellation with the isodesmic bond separation reactions are inferior. Hence, we recommend the use of the CBH-2 and higher rungs for appropriate error-cancellation in quantum chemistry.

(c). Performance of the Method—Extrapolated CCSD Energies. The errors (full CCSD/6-31+G(d,p) − extrapolated CCSD/6-31+G(d,p) energy) for all the 30 molecules are provided in Table 4. The mean absolute error at the CBH-2 method is 0.58 kcal/mol, and it improves to 0.19 kcal/mol at CBH-3 (Table 4). An interesting feature in all the CBH-2 errors in Table 4 is that, barring one number (that for the ring form of glucose), all other errors are positive; that is, at CBH-2, our method is consistently slightly underestimating the actual CCSD energies for most of the molecules.

Dissecting the test set in Table 4 into test sets A and B as described earlier, we immediately see from Table 5 that the mean absolute error for test set B at CBH-2 is slightly large (1.01 kcal/mol). This is very similar to the trend noticed with CCSD(T) extrapolation (*vide supra*). Interestingly, in comparison to the CCSD(T) extrapolation, for both test sets A and B, CCSD extrapolation is marginally worse at CBH-2 and marginally better at CBH-3 (compare Tables 2 and 5).

Inspection of the individual molecules in test set B (columns 3 and 4, Table 4) reveals that for all molecules barring the ring form of glucose, there is a notable improvement on going from CBH-2 to CBH-3. Even for camphor, a massive error of 2.33 kcal/mol at CBH-2 is dramatically reduced to −0.24 kcal/mol at CBH-3. This trend is unlike the earlier observation in the case of CCSD(T) (*vide supra*) wherein the improvement on going from CBH-2 to CBH-3 in most cases was only modest.

Analogous to section 3c, extrapolation to CCSD using different basis-sets works extremely well too. Extrapolated CCSD/aug-cc-pVDZ (for all the 30 molecules in the test set, Table 4) energies and errors help confirm the usefulness of our method. For the sake of completion, we have also provided the very large errors (mean absolute error for just the test set A alone ~4 kcal/mol) when CBH-1 is used to get extrapolated CCSD energies in Table 3.

(d). Aromatic Molecules. Our method does not result in accurate CCSD or CCSD(T) energies for aromatic molecules. With both CBH-2 or CBH-3, using Pople-style or Dunning style basis sets, extrapolation to either CCSD or CCSD(T) generally led to errors of 6−8 kcal/mol for even the simplest of aromatic compounds—benzene.

We, however, anticipated this result, as we know from our previous work in thermochemistry[25b] that the MP2 method does not perform well in the case of aromatic compounds, whereas CCSD(T) does quite well. Therefore, there is a mismatch in the error-cancellations between CCSD(T) and MP2 (i.e., eq 5 is not valid for aromatic molecules). Hence, extrapolations using eqs 6 and 8 are not expected to be accurate.

4. CONCLUSIONS AND SCOPE

In summary, we have devised an elegant and easy-to-use protocol to obtain extrapolated CCSD(T) and CCSD energies based on the MP2 method for organic molecules by tweaking our automated thermochemical hierarchy CBH to be applicable as a fragment-based method. The method is robust, and it substantially diminishes the computational cost involved in a CCSD(T) (or CCSD) calculation without any significant loss of accuracy. Our approach of treating the fragments in their equilibrium geometries severely reduces the total number of electronic structure calculations necessary to assemble the overall energy of a larger molecule. The success of our method is based on similar error-cancellations observed between the CCSD(T) and MP2 methods at the higher rungs of CBH. For a test set of 30 diverse nonaromatic organic molecules and biomonomers, the mean absolute error for extrapolation to CCSD(T) (full CCSD(T)/6-31+G(d,p) − extrapolated CCSD(T)/6-31+G(d,p)) energy is only 0.35 kcal/mol at CBH-2, and further, it slightly improves to 0.27 kcal/mol at CBH-3. Similarly, the mean absolute error for extrapolation to CCSD is only 0.58 kcal/mol at CBH-2 and greatly improves to 0.19 kcal/mol at CBH-3 (both with the 6-31+G(d,p) basis-sets). These results are very good despite the fact that the fragments involved have different geometries than that of the parent molecule.

For extrapolation to CCSD(T), since there is only a moderate improvement in the result upon going from CBH-2 to CBH-3, yet the size of the fragments in CBH-3 increases, we recommend using CBH-2 to balance the computational cost with desired accuracy. For the extrapolation to CCSD, however, there is a sizable improvement in the accuracy upon going from CBH-2 to CBH-3. Hence, we recommend using CBH-3 for extrapolating to CCSD. We currently do not recommend our method to obtain extrapolated CCSD(T) or CCSD energies for aromatic molecules. This is due to a previously noted mismatch in the performance of CCSD(T) and MP2 methods for such molecules. We are currently working toward developing a more accurate extrapolation method for aromatic molecules.

The method described here lays the foundation to carry out very accurate CCSD(T) calculations on even larger molecules and carry out sophisticated electronic structure applications using a combination of black-box approaches (such as MIM) and chemically intuitive methods (such as CBH).

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Performance of our method with a larger basis-set and the coordinates of the optimized geometries of all the 30 molecules used in this work. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kraghava@indiana.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge support from NSF grant CHE-1266154 at Indiana University.

REFERENCES

(1) *Theory and Applications of Computational Chemistry The First Forty Years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds; Elsevier: Amsterdam, The Netherlands, 2005; pp 1−7.

(2) Fock, V. *Z. Phys.* **1930**, *61*, 126.

(3) (a) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618. (b) Krishnan, R.; Pople, J. A. *Int. J. Quantum Chem.* **1978**, *14*, 91. (c) Krishnan, R.; Frisch, M. J.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 4244. (d) Frisch, M. J.; Krishnan, R.; Pople, J. A. *Chem. Phys. Lett.* **1980**, *75*, 66. (e) Bartlett, R. J.; Shavitt, I. *Chem. Phys. Lett.* **1977**, *50*, 190. (f) Bartlett, R. J.; Purvis, G. D. *J. Chem. Phys.* **1978**, *68*, 2114. (g) Bartlett, R. J.; Sekino, H.; Purvis, G. D. *Chem. Phys. Lett.* **1983**, *98*, 66. (h) Amos, R. D.; Andrews, J. S.; Handy, N. C.; Knowles, P. J. *Chem. Phys. Lett.* **1991**, *185*, 256. (i) Knowles, P. J.; Andrews, J. S.; Amos, R. D.; Handy, N. C.; Pople, J. A. *Chem. Phys. Lett.* **1991**, *186*, 130. (j) Lauderdale, W. J.; Stanton, J. F.; Gauss, J.; Watts, J. D.; Bartlett, R. J. *Chem. Phys. Lett.* **1991**, *187*, 21. (k) Murray, C.; Davidson, E. R. *Chem. Phys. Lett.* **1991**, *187*, 451. (l) Lee, T. J.; Jayatilaka, D. *Chem. Phys. Lett.* **1993**, *201*, 1. (m) Wilson, S.; Saunders, V. R. *J. Phys. B* **1979**, *12*, 403; **1979**, *13*, 1505−(E). (n) Kucharski, S. A.; Bartlett, R. J. *Adv. Quantum Chem.* **1986**, *18*, 281. (o) Kucharski, S. A.; Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1989**, *90*, 7282. (p) Bartlett, R. J.; Watts, J. D.; Kucharski, S. A.; Noga, J. *Chem. Phys. Lett.* **1990**, *165*, 513; **1990**, *167*, 609−(E). (q) Raghavachari, K.; Pople, J. A.; Replogle, E. S.; Head-Gordon, M. *J. Phys. Chem.* **1990**, *94*, 5579. (r) He, Z.; Cremer, D. *Theor. Chim. Acta* **1993**, *85*, 305.

(4) (a) Shavitt, I. In *Modern Theoretical Chemistry*, Schaefer, H. F., Ed.; Plenum Press: New York, 1977; Vol. 3, p 189. (b) For one of the early applications of the CI method, see Boys, , S. F. *Proc. R. Soc. London* **1950**,, *A201*,, 125. (c) For an excellent review of FCI applications, see Bauschlicher, C. W.; Langhoff, S. R.; Taylor, P. R. *Adv. Chem. Phys.* **1990**, *77*, 103. (d) Brooks, B. R.; Laidig, W. D.; Saxe, P.; Goddard, J. D.; Yamaguchi, Y.; Schaefer, H. F. *J. Chem. Phys.* **1980**, *72*, 4652. (e) Krishnan, R.; Schlegel, H. B.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 4654. (f) Schaefer, H. F.; Thomas, J. R.; Yamaguchi, Y.; DeLeeuw, B. J.; Vacek, G. In *Modern Electronic Structure Theory*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; p 1. (g) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968. (h) Raghavachari, K.; Pople, J. A.; Head-Gordon, M. In *Many-Body Methods in Quantum Chemistry, Lecture Notes in Chemistry*; Kaldor, U., Ed.; Springer-Verlag: Berlin, 1989; Vol. *52*, p 215. (i) Scuseria, G. E.; Schaefer, H. F. *J. Chem. Phys.* **1989**, *90*, 3700. (j) Paldus, J.; Cizek, J.; Jeziorski, B. *J. Chem. Phys.* **1989**, *90*, 4356. (k) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1989**, *90*, 4635. (l) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Replogle, E. *Chem. Phys. Lett.* **1989**, *158*, 207. (m) He, Z.; Kraka, E.; Cremer, D. *Int. J. Quantum Chem.* **1996**, *57*, 157.

(5) (a) Cizek, J. *J. Chem. Phys.* **1966**, *45*, 4256; *Adv. Chem. Phys.* **1969**, *14*, 35. (b) Bartlett, R. J.; Dykstra, C. E.; Paldus, J. In *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules*; Dykstra, C. E., Ed.; Reidel: Dordrecht, 1984; p 127. (c) Bartlett, R. J. *J. Phys. Chem.* **1989**, *93*, 1697. (d) Bartlett, R. J. In *Modern Electronic Structure Theory*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; p 1047. (e) Lee, T. J.; Scuseria, G. E. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer: Dordrecht, 1995; p 47. (f) Taylor, P. R.; Bacskay, G. B.; Hush, N. S.; Hurley, A. C. *Chem. Phys. Lett.* **1976**, *41*, 444. (g) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem.* **1978**, *14*, 545. (h) Bartlett, R. J.; Purvis, G. D. *Int. J. Quantum Chem.* **1978**, *14*, 561. (i) Chiles, R. A.; Dykstra, C. E. *J. Chem. Phys.* **1981**, *74*, 4544. (j) Paldus, J.; Cizek, J.; Shavitt, I. *Phys. Rev. A* **1972**, *5*, 50. (k) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910. (l) Lee, T. J.; Rice, J. E. *Chem. Phys. Lett.* **1988**, *150*, 406. (m) Scuseria, G. E.; Scheiner, A. C.; Lee, T. J.; Rice, J. E.; Schaefer, H. F. *J. Chem. Phys.* **1987**, *86*, 2881. (n) Scuseria, G. E.; Janssen, C. L.; Schaefer, H. F. *J. Chem. Phys.* **1988**, *89*, 7382. (o) Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1987**, *86*, 7041; **1988**, *89*, 3401−(E). (p) Hoffmann, M. R.; Schaefer, H. F. *Adv. Quantum Chem.* **1986**, *18*, 207. (q) Scuseria, G. E.; Schaefer, H. F. *Chem. Phys. Lett.* **1988**, *152*, 382. (r) Watts, J. D.; Bartlett, R. J. *J. Chem. Phys.* **1990**, *93*, 6104. (s) Lee, Y. S.; Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1984**, *81*, 5906; **1984**, *82*, 5761−(E). (t) Noga, J.; Bartlett, R. J.; Urban, M. *Chem. Phys. Lett.* **1987**, *134*, 126. (u) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.

(6) (a) Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons: Chichester, U.K., 2007. (b) Raghavachari, K.; Anderson, J. B. *J. Phys. Chem.* **1996**, *100*, 12960.

(7) (a) Fujimoto, H.; Koga, N.; Fukui, K. *J. Am. Chem. Soc.* **1981**, *103*, 7452. (b) Fedorov, D. G.; Ishida, T.; Kitaura, K. *J. Phys. Chem. A* **2005**, *109*, 2638. (c) Fedorov, D. G.; Ishida, T.; Uebayasi, M.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 2722. (d) Fedorov, D. G.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 6904. (e) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. *Chem. Phys. Lett.* **1999**, *313*, 701. (f) Fedorov, D. G.; Slipchenko, L. V.; Kitaura, K. *J. Phys. Chem. A* **2010**, *114*, 8742. (g) Pruitt, S. R.; Fedorov, D. G.; Kitaura, K.; Gordon, M. S. *J. Chem. Theory Comput.* **2010**, *6*, 1.

(8) (a) Gadre, S. R.; Shirsat, R. N.; Limaye, A. C. *J. Phys. Chem.* **1994**, *98*, 9165. (b) Gadre, S. R.; Ganesh, V. *J. Theor. Comput. Chem.* **2006**, 5835. (c) Ganesh, V.; Dongare, R. K.; Balanarayan, P.; Gadre, S. R. *J. Chem. Phys.* **2006**, *125*, 104109. (d) Kavathekar, R.; Khire, S.; Ganesh, V.; Rahalkar, A. P.; Gadre, S. R. *J. Comput. Chem.* **2009**, *30*, 1167. (e) Rahalkar, A. P.; Ganesh, V.; Gadre, S. R. *J. Chem. Phys.* **2008**, *129*, 234101. (f) Rahalkar, A. P.; Katouda, M.; Gadre, S. R.; Nagase, S. *J. Comput. Chem.* **2010**, *31*, 2405. (g) Yeole, S. D.; Gadre, S. R. *J. Chem. Phys.* **2010**, *132*, 094102. (h) Rahalkar, A. P.; Mishra, B. K.; Ramanathan, V.; Gadre, S. R. *Theor. Chem. Acc.* **2011**, *130*, 491. (i) Sahu, N.; Yeole, S. D.; Gadre, S. R. *J. Chem. Phys.* **2013**, *138*, 104101.

(9) (a) Mei, Y.; Ji, C. G.; Zhang, J. Z. H. *J. Chem. Phys.* **2006**, *125*, 094906. (b) Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2003**, *119*, 3599.

(10) (a) Addicoat, M. A.; Collins, M. A. *J. Chem. Phys.* **2009**, *131*, 104103. (b) Collins, M. A.; Deev, V. A. *J. Chem. Phys.* **2006**, *125*, 104104. (c) Collins, M. A. *J. Chem. Phys.* **2007**, *127*, 024104. (d) Deev, V.; Collins, M. A. *J. Chem. Phys.* **2005**, *122*, 154102. (e) Netzloff, H. M.; Collins, M. A. *J. Chem. Phys.* **2007**, *127*, 134113. (f) Collins, M. A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7744.

(11) (a) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293. (b) Slipchenko, L. V.; Gordon, M. S. *J. Comput. Chem.* **2007**, *28*, 276.

(12) (a) Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 46. (b) Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 1342. (c) Dahlke, E. E.; Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 33. (d) Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 1. (e) Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 1573. (f) Sorkin, A.; Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 683. (g) Kurbanov, E, K.; Leverentz, H. R.; Truhlar, D. G.; Amin, E. A. *J. Chem. Theory Comput.* **2013**, *9*, 2617 DOI: 10.1021/ct4001872. (h) Isegava, M.; Wang, Bo.; Truhlar, D. G. *J. Chem. Theory Comput.* **2013**, *9*, 1381.

(13) (a) Li, W.; Li, S.; Jiang, Y. *J. Phys. Chem. A* **2007**, *111*, 2193. (b) Hua, W. J.; Fang, T.; Li, W.; Yu, J. G.; Li, S. H. *J. Phys. Chem. A* **2008**, *112*, 10864. (c) Hua, S. G.; Hua, W. J.; Li, S. H. *J. Phys. Chem. A* **2010**, *114*, 8126.

(14) (a) Huang, L.; Massa, L.; Karle. *J. Proc. Natl. Acad. Sci.* **2006**, *103*, 1233. (b) Huang, L.; Massa, L.; Karle, J. *Int. J. Quantum Chem.* **2005**, *103*, 808.

(15) (a) Beran, G. J. O. *J. Chem. Phys.* **2009**, *130*, 164115. (b) Beran, G. J. O.; Nanda, K. *J. Phys. Chem. Lett.* **2010**, *1*, 3480. (c) Wen, S.; Nanda, K.; Hunag, Y.; Beran, G. J. O. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7578. (d) Nanda, K.; Beran, G. J. O. *J. Chem. Phys.* **2012**, *137*, 174106.

(16) Rezac, J.; Salahub, D. R. *J. Chem. Theory Comput.* **2010**, *6*, 91.

(17) (a) Wang, Y.; Sosa, C. P.; Cembran, A.; Truhlar, D. J.; Gao, J. *J. Phys. Chem. B* **2012**, *116*, 6781. (b) Gao, J. *J. Phys. Chem. B* **1997**, *101*, 657. (c) Gao, J. *J. Chem. Phys.* **1998**, *109*, 2346. (d) Xie, W.; Gao, J. *J. Chem. Theory Comput.* **2007**, *3*, 1890. (e) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Chem. Phys.* **2008**, *128*, 234108. (f) Song, L.; Han, J.;

Lin, Y. L.; Xie, W.; Gao, J. *J. Phys. Chem. A* **2009**, *113*, 11656. (g) Wierzchowski, S. J.; Kofke, D. A.; Gao, J. *J. Chem. Phys.* **2003**, *119*, 7365. (h) Xie, W.; Orozco, M.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 459. (i) Cembran, A.; Song, L.; Mo, Y.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 2702.

(18) (a) Richard, R. M.; Herbert, J. M. *J. Chem. Phys.* **2012**, *137* (064113), 1−17. (b) Richard, R. M.; Herbert, J. M. *J. Chem. Theory Comput.* **2013**, *9*, 1408.

(19) (a) Bates, D. M.; Smith, J. R.; Tshumper, G. S. *J. Chem. Phys.* **2011**, *135*, 044123. (b) Bates, D. M.; Smith, J. R.; Tshumper, G. S. *J. Chem. Theory Comput.* **2011**, *7*, 2753.

(20) Mayhall, N. J.; Raghavachari, K. *J. Chem. Theory Comput.* **2011**, *7*, 1336.

(21) Mayhall, N. J.; Raghavachari, K. *J. Chem. Theory Comput.* **2012**, *8*, 2669.

(22) (a) Hehre, W. J.; Ditchfield, R.; Radom, L.; Pople, J. A. *J. Am. Chem. Soc.* **1970**, *92*, 4796. (b) Radom, L.; Hehre, W. J.; Pople, J, A. *J. Am. Chem. Soc.* **1971**, *93*, 289. (c) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley-Interscience: New York, 1986.

(23) (a) George, P.; Trachtman, M.; Bock, C. W.; Brett, A. M. *Theor. Chem. Acc.* **1975**, *38*, 121. (b) George, P.; Trachtman, M.; Bock, C. W.; Brett, A. M. *Tetrahedron* **1976**, *32*, 317. (c) George, P.; Trachtman, M.; Bock, C. W.; Brett, A. M. *J. Chem. Soc., Perkin Trans. 2* **1976**, 1222. (d) Hess, B. A., Jr.; Schaad, L. J. *J. Am. Chem. Soc.* **1983**, *105*, 7500. (e) Nyulaszi, L.; Vamai, P.; Veszpremi, T. *J. Mol. Struct. THEOCHEM* **1995**, *358*, 55. (f) Vianello, R.; Liebman, J. F.; Maksic, Z. B. *Chem.—Eur. J.* **2004**, *10*, 5751. (g) Vianello, R.; Liebman, J. F.; Maksic, Z. B.; Muller, T. J. *J. Phys. Chem. A* **2005**, *109*, 10594. (h) Chestnut, D. B.; Davis, K. M. *J. Comput. Chem.* **1997**, *18*, 584. (i) El-Nahas, A. M.; Bozelli, J. W.; Simmie, J. M.; Navarro, M. V.; Black, G.; Curran, H. J. *J. Phys. Chem. A* **2006**, *110*, 13618. (j) George, P.; Bock, C. W.; Trachtman, M. *Theor. Chim. Acta* **1987**, *71*, 289. (k) Gimarc, B. M.; Zhao, M. *J. Phys. Chem.* **1994**, *98*, 1596. (l) Zhao, M.; Gimarc, B. M. *J. Phys. Chem.* **1993**, *97*, 4023. (m) Warren, D. S.; Gimarc, B. M. *J. Phys. Chem.* **1993**, *97*, 4031.

(24) Wheeler, S. E.; Houk, K. N.; Schlyer, P. v. R.; Allen, W. D. *J. Am. Chem. Soc.* **2009**, *131*, 2547.

(25) (a) Ramabhadran, R. O.; Raghavachari, K. *J. Chem. Theory Comput.* **2011**, *7*, 2094. (b) Ramabhadran, R. O.; Raghavachari, K. *J. Phys. Chem. A* **2012**, *116*, 7531. (c) Ramabhadran, R. O; Sengupta, A.; Raghavachari, K. *J. Phys. Chem. A* **2013**, *117*, 4973.

(26) (a) Lee, T. J.; Scuseria, G. E. In *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*; Langhoff, S. R., Ed.; Kluwer: Dordrecht, 1995; p 47. (b) Lee, T. J.; Scuseria, G. E. *J. Chem. Phys.* **1990**, *93*, 489. (c) Scuseria, G. E.; Lee, T. J. *J. Chem. Phys.* **1990**, *93*, 5851. (d) Tew, D. P.; Klopper, W.; Helgaker, T. *J. Comput. Chem.* **2007**, *28*, 1307. (e) Rezac, J.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 2151. (f) Rezac, J.; Simoa, L.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 364.

(27) For efficient direct and semidirect MP2 algorithms, see (a) Head-Gordon, M.; Pople, J. A.; Frisch, M. *J. Chem. Phys. Lett.* **1988**, *153*, 503. (b) Saebø, S.; Almlo, F. J. *Chem. Phys. Lett.* **1989**, *154*, 83. (c) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 275.

(28) Our extrapolation to CCSD(T) energies based on MP2 energies method is actually based on an additivity approximation. Thus, our method can also be described as an additivity scheme.

(29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E..; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, A.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.1; Gaussian Inc.: Wallingford, CT, 2009.

(30) (a) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *127*, 124105. (b) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *127*, 084108.