

Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics†

Yan Zhao, Benjamin J. Lynch and Donald G. Truhlar

Department of Chemistry and Supercomputing Institute, University of Minnesota,
207 Pleasant Street S.E., Minneapolis, MN 55455-0431, USA

Received 4th November 2004, Accepted 11th November 2004
First published as an Advance Article on the web 10th December 2004

We have developed a new kind of multi-coefficient correlation method (MCCM) by empirically mixing correlated wave function methods and density functional methods. The new methods constitute a generalization of hybrid density functional theory and may be called multi-coefficient extrapolated density functional theory. Results by the new methods are compared to those obtained by G3SX, G3SX(MP3), CBS-Q and MCCM/3 for calculations of atomization energies, barrier heights, ionization potentials and electron affinities. These results show that the multi-coefficient extrapolated density functional theory is more accurate for thermochemistry and thermochemical kinetics than the pure wave function methods of comparable cost. As a byproduct of this work we optimized a new hybrid meta density functional theory called TPSS1KCIS, which has excellent performance for thermochemistry.

1. Introduction

The continued development of improved electronic structure methods is essential for extending the power of computational thermochemistry to new practical applications. Over the past decades, much progress has been made in two directions: methods based on wave function theory (WFT) and methods based on density functional theory (DFT). John Pople and Walter Kohn shared the 1998 Nobel Prize for their contributions to these two areas.

Among the wavefunction-based approaches, the Gaussian-*n* theories and their variants developed by Pople and coworkers,^{1–3} the complete basis set (CBS) family of methods by Petersson and coworkers,^{4,5} the Weizmann-*n* theories and their variants of Martin and coworkers,^{6,7} and the multi-coefficient correlation methods (MCCMs)^{8–15} of our group are the most popular schemes which achieve “chemical accuracy” (± 1 kcal mol⁻¹ for bond energies and barrier heights). These methods are classified as multilevel methods because they involve combining more than one level of electronic structure theory and/or different one-electron basis sets, usually with empirical parameters, to try to extrapolate to a more accurate result than the most accurate component calculation, and this multilevel approach is a very powerful alternative to single-level *ab initio* methods for many types of calculations. Unfortunately, the computational cost of most of the above mentioned methods such as G2,¹ G3,² CBS-Q,⁴ CBS/APNO,⁴ W1,⁶ W2,⁶ and multi-coefficient G3 (MCG3)^{11,15} formally scales as N^7 , where N is the number of atoms. If one wants to calculate consistent gradients or Hessians (for example, in dynamics calculations or geometry optimizations), these methods are very expensive, and sometimes prohibitively so.

A previous paper¹⁵ developed a suite of MCCMs of varying accuracy and cost, namely the MCCM/3 suite, that is suitable for a variety of problems with different sizes of molecules and different accuracy requirements. The recommended methods in the MCCM/3 suite are MC-CO/3, MC-UT/3, MC-QCISD/3

and MCG3/3. A particularly powerful version of the MC method is the MC-QCISD method,^{13,15} which scales as N^6 . In the limit of only a single coefficient, the MCCM methods reduce to the older scaling-all-correlation (SAC) method,¹⁶ which scales as N^5 . Although the MCCM and SAC methods have empirical parameters, the guiding principle of scaling the correlation energy and extrapolating to an infinite one-electron basis set provide a rational basis for the functional forms in which the parameters are embedded.

In addition to these WFT approaches, DFT-type methods^{17–63} have been widely used for thermochemistry and thermochemical kinetics due to their excellent cost-to-performance ratio. Hybrid DFT is a particularly powerful subset of DFT methods in which nonlocal Hartree–Fock exchange is combined with density functionals in the one-electron Hamiltonian (*i.e.*, the Fock–Kohn–Sham operator) that determines the orbitals. Like the Gaussian-*n* and MCCM theories, hybrid DFT theories are all semiempirical. Recently we optimized a 1-parameter DFT model for kinetics, namely BB1K based on Becke’s 1988 gradient corrected exchange functional (Becke88 or B)²⁰ and Becke’s 1995 kinetic-energy-dependent dynamical correlation functional (Becke95 or B95).²⁶ BB1K is a hybrid meta DFT method which has better performance for kinetics than the previous successful non-meta MPW1K model.⁴² (The descriptor “meta” means that the “density functional” depends not just on the density and its gradient but also on the orbitals in the form of a kinetic energy density.) More recently we optimized two hybrid DFT models based on the modified Perdew and Wang exchange functional³¹ (MPW) and the B95 meta correlation functional.²⁶ One is called MPW1B95,⁶³ and it is a hybrid DFT model for general applications in thermochemistry. Our assessments show that MPW1B95 has better performance for noncovalent interactions than the also very successful B1B95²⁶ and B97-2⁴⁷ models. The other model we developed is called MPWB1K,⁶³ and it is a one-parameter hybrid meta DFT method for kinetics. The MPWB1K method gives good results for thermochemistry, thermochemical kinetics, hydrogen bonding and weak interactions. Although one sometimes makes a distinction between pure DFT, where the density functionals depend explicitly on the density and its derivatives, and meta, hybrid, and hybrid meta DFT, where the

† Electronic supplementary information (ESI) available: Databases for parametrization and testing. See <http://www.rsc.org/suppdata/cp/b4/b416937a/>

functionals also depend explicitly on Kohn–Sham orbitals and (in meta methods) on their derivatives, we note that technically these are all DFT since the generalized Kohn–Sham orbitals are functionals of the Kohn–Sham potential, which is a functional of the density.^{19,38} When, for either conceptual or operational reasons, we want to remain mindful of the explicit orbital dependence density functionals that depend on orbitals may be called orbital dependent DFT (or ODDFT) functionals. Nevertheless, when we call a method DFT, we do exclude ODDFT; we need to keep the terminology manageable.

The inclusion of the Hartree–Fock component in hybrid DFT is not the only way to combine the WFT methods with DFT methods. Another combination is to use hybrid DFT to optimize geometries and calculate zero point vibrational energy (ZPVE). For example, the wave function-based G3SX,³ W1,⁶ and CBS-QB3⁵ methods use the B3LYP²⁵ hybrid DFT method to optimize geometries and calculate the ZPVE. Recently we generalized the multi-coefficient approach to directly mix WFT energies and gradients with DFT ones.⁶⁴ In particular, we used the multi-coefficient approach to mix the SAC method^{8,12,15,16,65} with a hybrid meta DFT or hybrid DFT method. The resulting methods are called the multi-coefficient three-parameter Becke88–Becke95 (MC3BB) method⁶⁴ and multi-coefficient three-parameter modified Perdew–Wang (MC3MPW) method.⁶⁴ These two methods can also be viewed as a semiempirical hybrid of the SAC method and the hybrid meta DFT or hybrid DFT method, respectively, so we also labeled the MC3BB method as a doubly hybrid meta DFT method, and we call the MC3MPW method a doubly hybrid DFT method. The MC3BB and MC3MPW methods give better performance for the atomization energies and barrier heights calculations than the hybrid meta DFT and hybrid DFT methods, but at only slightly increased cost. A key distinction between these MC methods and previous empirical DFT methods that involve more than one linear parameters (such as B3PW91²³ and B3LYP²⁵) is the explicit attempt to extrapolate to a complete basis set and full configuration interaction; for this reason we call the new methods multi-coefficient extrapolated DFT.

An important conclusion of our recent paper⁶³ mentioned above is that the modified Perdew and Wang exchange functional (MPW)³¹ is the overall best available exchange functional, and the Becke’s 1995 correlation functional (B95)²⁶ is the best correlation functional for complementing it. Therefore, in the present paper, we will develop a new doubly hybrid meta DFT method based on combining MPW, B95 and SAC. Following the spirit of standard procedures for naming DFT methods, the new method is called the multi-coefficient three-parameter modified Perdew–Wang–Becke95 (MC3MPWB) method, similarly we optimize another MC3 method called MC3TS in which the exchange is treated by the functional of Tao *et al.*⁵⁹ (called TPSS), and correlation is treated by the method of Savin and coworkers^{36,37,51} (sometimes called KCIS, but here abbreviated as S).

The present article is, however, *primarily* addressed to the question of whether we can profitably combine higher-level WFT methods with DFT. To answer this, we will develop a series of doubly hybrid methods that directly mix the WFT-based MCCMs with DFT using semiempirical parameters. We will compare the accuracy of these new multi-coefficient extrapolated DFT methods to wave-function based methods such as the CBS-Q, G3SX and the MCCM/3 suite and to nonextrapolated DFT methods such as the popular B3LYP hybrid DFT method and a new DFT method, presented in the present article, called TPSS1KCIS (which is explained in section 3).

One line of research seeks DFT methods where the parameters are obtained from general principles rather than semiempirically, but the present work de-emphasized that distinction in light of the thin line³⁵ between theoretical con-

siderations and empiricism and in light of lack of a unique theoretical justification for the functional forms themselves. As clarified by Becke,²⁹ the fundamental value of the semiempirical hybrid approach is not diminished by parametrization. In fact, a flexible parametrization has implications for the limits achievable by a particular combination of elements and functional dependences, and it provides a guide into the areas where further development and new insights are required.²⁹ This is the spirit of the present work. At the same time, the resulting parametrized methods can be very useful for practical work until such (unknown) time when additional breakthroughs may allow a more fundamental approach, and the development of useful methods for practical applications is the immediate goal of our work.

Section 2 presents the databases used in the present work. Section 3 discusses the theory and parametrization of the new methods. Section 4 presents results and discussion.

2. Databases and parametrization

We will use several sets of data for parametrization and testing, and these data and their usage are explained in this section.

2.1. MGAE109/04 database

The MGAE109/04 database consists of 109 atomization energies (AEs) of main group (MG) molecules. This AE test set contains a diverse set of molecules including organic and inorganic compounds. All 109 data are pure electronic energies, *i.e.*, zero-point energies and thermal vibrational-rotational energies have been removed. The 109 zero-point-exclusive atomization energies are part of Database/3 and are identical to those used previously,^{15,53,64} except that we update the atomization energies of 12 molecules by using the anharmonic zero point energies published by Martin and Oliveira.⁶ The MGAE109/04 database is given in the supporting information. The average number of bonds per molecule in this database is 4.71, where single, double and triple bonds are all counted as one bond.

2.2. HTBH38/04 database

The HTBH38/04 database consists of 38 transition state barrier heights of hydrogen transfer reactions and it is a subset of the previous BH42/04 database. The best estimates for the barrier heights were obtained, as explained elsewhere,^{14,15,42,46,61,64} from a combination of experimental and theoretical kinetics data. The HTBH38/04 database is also listed in the supporting information.

2.3. Ionization potential and electron affinity databases

The zero-point-exclusive ionization potential (IP) and electron affinity (EA) databases are taken from a previous paper.⁵³ These databases are also part of Database/3, and they consist of six atoms and seven molecules for which the IP and EA are both present in the G3 data set.² These databases are called IP13/3 and EA13/3, respectively, and they are listed in the supporting information.

2.4. Parametrization

We optimized a hybrid DFT method, namely TPSS1KCIS, against the root mean square error (RMSE) for the MGAE109/04 Database. The optimization of parameters for the two new MC3-type methods, namely MC3MPWB95 and MC3TS (mentioned in Section 1 and described in Section 3), are carried out by minimizing the following function:

$$F = 0.5[\text{MUEPB}(\text{MGAE109/04}) + \text{MUE}(\text{HTBH38/04})] \quad (1)$$

where MUEPB is the mean unsigned error (MUE, same as mean absolute deviation (MAD)) on a per bond basis (which is obtained by dividing the MUE for MGAE109/04 data base by the average number of bonds per molecule). All other multi-coefficient hybrid methods are optimized against the RMSE of the 173 data in the above mentioned three databases. The optimization of the coefficients was carried out by using the Solver routine⁶⁶ in the Microsoft Excel program.

2.5. TAE6/04 test set

The TAE6/04 test set contains atomization energies of six difficult molecules that are not in any of our training sets. The molecules are ClCN, OCS, O₃, P₄, C₈H₁₈ (*n*-octane) and C₁₀H₈ (naphthalene). The ClCN and OCS molecules are in the training set of Martin and Oliveria's W1 and W2 methods,⁶ but they are not in the G3/99 test set,⁶⁷ whereas O₃, P₄, C₈H₁₈ (*n*-octane) and C₁₀H₈ (naphthalene) are in the G3/99 test set.⁶⁷ O₃ is a difficult case for W1 and W2 theory due to its intrinsic multireference character. Curtiss and coworkers found that P₄ is a difficult case for their scaled G3 methods and its variants.^{3,68} We selected C₈H₁₈ (*n*-octane) and C₁₀H₈ (naphthalene) because they are respectively the largest molecule (in terms of number of atoms) and largest non-alkane in the G3/99 test set.⁶⁷ We will test our methods against the TAE6/04 database to see the transferability of the parameters and methods developed in this paper.

3. Theory

Our description of the new methods presented here use the pipe notation for the level (L) and basis set (B); this was introduced elsewhere.⁹ The energy difference between two basis sets will be represented as

$$E(L/B2|B1) = E(L/B2) - E(L/B1) \quad (2)$$

where L is a particular electronic structure method, and B1 is smaller than B2. The energy change that occurs upon improving the treatment of the correlation energy will be represented by

$$E(L2|L1/B) = E(L2/B) - E(L1/B) \quad (3)$$

where L1 is a lower many-electron level of theory than L2, and B is a common basis set.

The new methods are constructed by taking linear combinations of DFT methods with WFT single-level methods such as Hartree-Fock (HF) theory, Møller-Plesset perturbation theory (MP2,⁶⁹ MP4SDQ⁷⁰), quadratic configuration interaction with single and double excitations (QCISD),⁷¹ and QCISD with quasiperturbative connected triples⁷¹ (QCISD(T)). The geometries of all species in all calculations on the three databases of section 2.1–2.3 were optimized at the QCISD level with the MG3 basis set. The MG3 basis set,¹¹ also called G3LargeMP2,⁷² is the same as 6-311++G(3d2f, 2df, 2p)⁷³ for H–Si, but improved² for P–Ar. The QCISD/MG3 geometries for all calculations in this paper can be obtained from the Truhlar group database website.⁷⁴

We also used four basis sets in the parametrization of the new multi-coefficient hybrid methods, namely, the 6-31G(d)^{70,75} 6-31+G(d,p),^{70,75} 6-31G(2df,p)^{70,75} and MG3S basis set. The MG3S basis set⁵³ is the same as MG3 (explained above) except that it omits diffuse functions on hydrogens. The following abbreviations for standard basis sets are used throughout this paper:

$$\begin{aligned} \text{Dd} & \text{6-31G(d)} \\ \text{DIDZ} & \text{6-31+G(d,p)} \\ \text{D2dfp} & \text{6-31G(2df,p)} \end{aligned}$$

We will describe the new methods in the next three sections, classified according to their scaling properties.

3.1. N^7 methods

We developed three methods that scale as N^7 . The three new methods are based on the MCG3/3 method¹⁵ as combined with an ODDFT method. Here and in the creation of the other new methods in this paper we consider three DFT methods, which are either hybrid DFT or hybrid meta DFT, as specified below. We replace the HF/MG3S component of the MCG3/3 method by one or another of these DFT methods, resulting in three multi-coefficient extrapolated DFT methods of a type that we will call doubly hybrid MCG3 methods, respectively. We also found that adding an MP2/DIDZ calculation improves the performance of the new methods. The new methods can be written as:

$$\begin{aligned} E(\text{MCG3-DFT}) = & c_8 \{ E[\text{HF/Dd}] + c_1 E[\text{MP2|HF/Dd}] \\ & + c_2 E[\text{MP2/DIDZ|Dd}] \\ & + c_3 E[\text{MP2 /D2dfp|DIDZ}] \\ & + c_4 E[\text{MP2/MG3S|D2dfp}] \\ & + c_5 E[\text{MP4SDQ|MP2/Dd}] \\ & + c_6 E[\text{MP4SDQ/D2dfp|Dd}] \\ & + c_7 E[\text{QCISD(T)|MP4SDQ/Dd}] \} \\ & + (1 - c_8) E(\text{DFTX/MG3S}) + E_{\text{SO}} \end{aligned} \quad (4)$$

where c_1, \dots, c_8 are parameters, DFTX denotes the chosen DFT method with a parameter X , and E_{SO} is the spin-orbit energy (which is often zero).

The three DFTX methods we used are: MPWBX, MPWX and TSX. MPWBX is a hybrid meta DFT method which uses modified Perdew-Wang exchange and the Becke95 correlation functional with $X\%$ HF exchange. MPWX is a hybrid DFT method which uses MPW exchange and Perdew-Wang 1991 (PW91) correlation²² again with $X\%$ HF exchange. TSX is a hybrid meta DFT method which uses the TPSS exchange⁵⁹ and the correlation functional of Savin and coworkers^{36,37,51} (as stated above, this is abbreviated S in the present work although it is sometimes called⁵¹ KCIS) Note that X is a parameter that is optimized along with the c_i parameters. We optimized the c_1 – c_8 coefficients and X simultaneously against the databases described in section 2, and they are listed in Table 1.

3.2. N^6 methods

We developed three methods that scale as N^6 and that are based on the MC-QCISD/3 method.¹⁵ We added a MP2/DIDZ calculation, and we replaced the HF/MG3S component of the MC-QCISD/3 method in turn by each of the three DFTX methods used in section 3.1, resulting respectively in three multi-coefficient extrapolated DFT methods of a type that we will call doubly hybrid MC-QCISD methods. The new methods can be written as:

$$\begin{aligned} E(\text{MCQCISD-DFT}) = & c_5 \{ E[\text{HF/Dd}] + c_1 E[\text{MP2|HF/Dd}] \\ & + c_2 E[\text{MP2/DIDZ|Dd}] \\ & + c_3 E[\text{MP2/MG3S|DIDZ}] \\ & + c_4 E[\text{QCISD|MP2/Dd}] \} \\ & + (1 - c_5) E(\text{DFTX/MG3S}) + E_{\text{SO}} \end{aligned} \quad (5)$$

Table 1 Parameters in the hybrid multi-coefficient methods ^a

Scaling	Method	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	X
N^7	MCG3-MPWB	1.085	0.276	0.465	1.329	1.255	0.889	1.057	0.727	34
N^7	MCG3-MPW	1.096	0.422	0.622	1.299	1.270	0.742	1.078	0.815	32
N^7	MCG3-TS	1.082	0.551	0.734	1.295	1.185	0.626	0.933	0.776	19
N^6	MCQCISD-MPWB	1.138	1.161	1.260	1.282	0.640				34
N^6	MCQCISD-MPW	1.163	1.106	1.324	1.331	0.711				32
N^6	MCQCISD-TS	1.127	1.140	1.309	1.144	0.673				19
N^6	MCUT-MPWB	1.102	1.171	1.346	1.404	0.526				29
N^6	MCUT-MPW	1.134	1.092	1.416	1.545	0.643				17
N^6	MCUT-TS	1.104	1.125	1.391	1.241	0.598				13
N^5	MCCO-MPWB	0.828	1.678	0.267						34
N^5	MCCO-MPW	0.957	1.546	0.386						32
N^5	MCCO-TS	0.952	1.435	0.401						22
N^5	MC3BB	1.332	0.205							39
N^5	MC3MPW	1.339	0.266							38
N^5	MC3MPWB95	1.196	0.257							38
N^5	MC3TS	1.256	0.336							31
N^4	TPSS1KCIS									13

^a See text for the definition of these parameters and also see text for the exchange and correlation functionals used in each methods.

We optimized c_1 – c_5 and X simultaneously against the databases described in section 2, and they are listed in Table 1.

Note that the three DFTX methods used in this work each has its own advantages. The use of MPW exchange and PW91 correlation does not require a computer code that includes kinetic energy density, but combining MPW exchange with B95 correlation is more accurate. Combining TPSS exchange and S correlation is even more accurate and leads to more stable SCF iterations.

We also developed three methods that scale as N^6 and that are based on the MC-UT/3 method.¹⁵ In each of them, we replace the HF/MG3S component of the MC-UT/3 method by one of the three DFTX methods, resulting respectively in three multi-coefficient extrapolated DFT methods of a type that we will call doubly hybrid MC-UT methods. Again we also add a MP2/DIDZ calculation. The new methods can be written as:

$$\begin{aligned}
 E(\text{MCUT-DFT}) = & c_5\{E[\text{HF/Dd}] + c_1E[\text{MP2|HF/Dd}] \\
 & + c_2E[\text{MP2/DIDZ|Dd}] \\
 & + c_3E[\text{MP2/MG3S|DIDZ}] \\
 & + c_4E[\text{MP4SDQ|MP2/Dd}]\} \\
 & + (1 - c_5)E(\text{DFTX/MG3S}) + E_{\text{SO}}
 \end{aligned} \quad (6)$$

We optimized c_1 – c_5 and X simultaneously against the databases described in section 2, and they are also listed in Table 1.

3.3. N^5 methods

We developed five multi-coefficient extrapolated methods that scale as N^5 . Three of them are based on the MC-CO/3 method¹⁵ and are called doubly hybrid MC-CO methods. We replace the HF/MG3S component of the MC-CO/3 method by the MPWBX, MPWX and TSX methods, respectively. We did not add a MP2/DIDZ calculation because we found that it does not improve the performance of these methods. The new methods can be written as:

$$\begin{aligned}
 E(\text{MCCO-DFT}) = & c_3\{E[\text{HF/Dd}] + c_1E[\text{MP2|HF/Dd}] \\
 & + c_2E[\text{MP2/MG3S|Dd}]\} \\
 & + (1 - c_3)E(\text{DFTX/MG3S}) + E_{\text{SO}}
 \end{aligned} \quad (7)$$

We optimized c_1 – c_3 and X simultaneously against the databases, and they are listed in Table 1.

We also optimized two new MC3-type methods, namely MC3MPWB and MC3TS. Both of them are analogs to MC3BB and MC3MPW, and they scale as N^5 . The MC3-type methods can be described as:

$$\begin{aligned}
 E(\text{MC3-type}) = & c_2\{E(\text{HF/DIDZ}) + c_1\Delta E(\text{MP2|HF/DIDZ})\} \\
 & + (1 - c_2)E(\text{DFTX/MG3S}) + E_{\text{SO}}
 \end{aligned} \quad (8)$$

As mentioned in the introduction, the DFTX method in MC3MPWB is MPWBX, whereas in MC3TS it is TSX. The optimized c_1 , c_2 and X for MC3MPWB and MC3TS as well as the parameters in the MC3BB and MC3MPW methods are listed in Table 1.

3.4. N^4 methods

We also optimized a hybrid meta DFT method that we call TPSS1KCIS, and it scales as N^4 . TPSS1KCIS uses TPSS⁵⁹

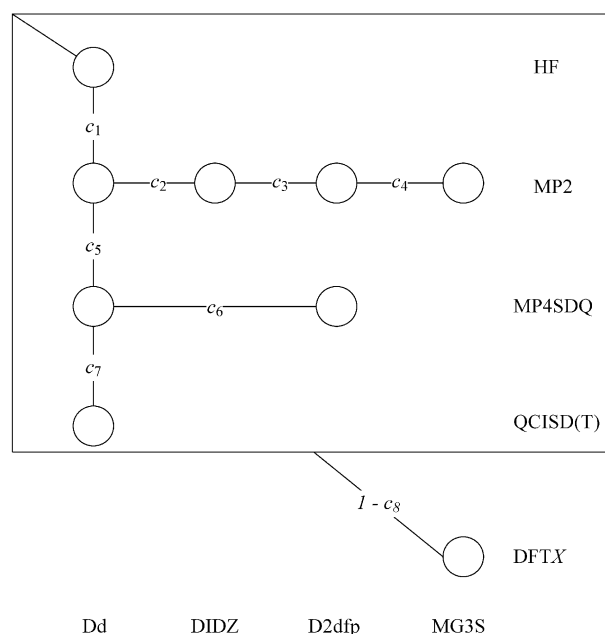


Fig. 1 Coefficient tree for the new N^7 multi-coefficient hybrid methods.

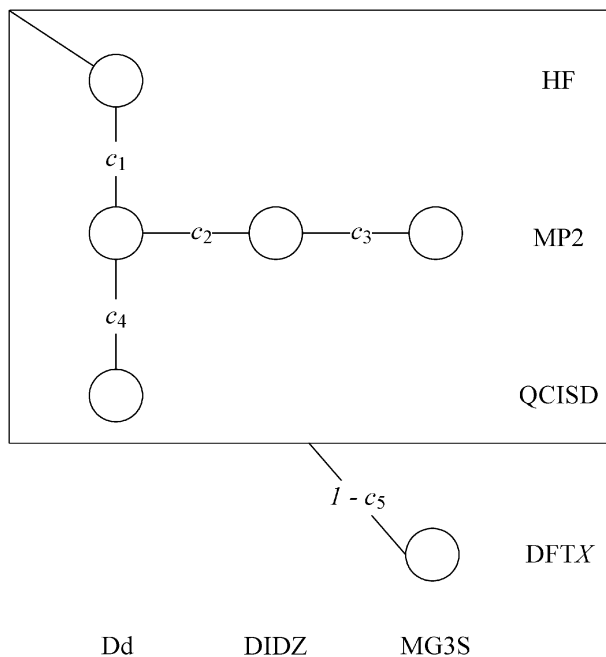


Fig. 2 Coefficient tree for the new N^6 multi-coefficient hybrid methods based on MC-QCISD/3.

exchange and S correlation.^{36,37,51} We optimized the percentage of HF exchange, X , against the RMSE of the MGAE109 database. The optimized parameter for TPSS1KCIS is listed in Table 1.

Figs. 1–5 are the coefficient trees for the new multi-coefficient hybrid methods.

All calculations were performed with the GAUSSIAN03⁷⁶ computer program. In all of the calculations, the spin-orbit stabilization energy was added to atoms and open-shell molecules for which it is nonzero, as described previously.¹¹

4. Results and discussion

We discuss the results in the next four sections, with the discussion divided according to the scaling properties of the methods.

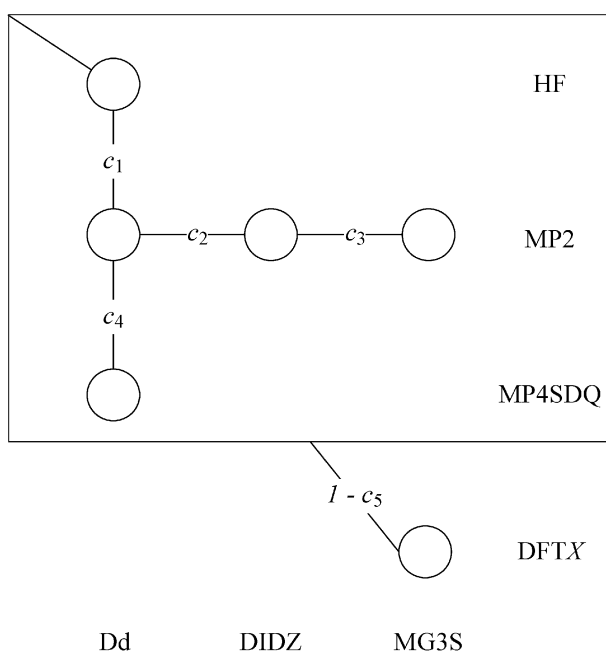


Fig. 3 Coefficient tree for the new N^6 multi-coefficient hybrid methods based on MC-UT/3.

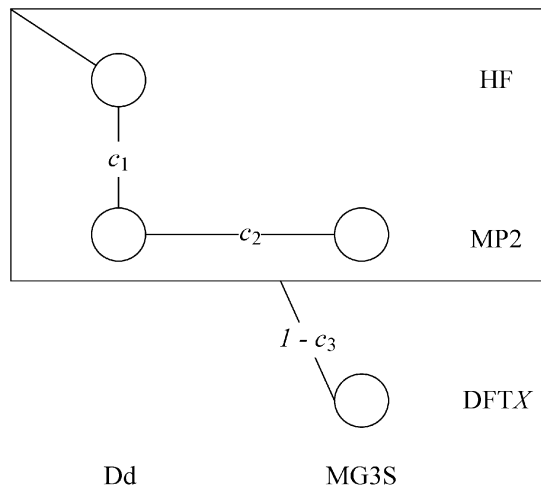


Fig. 4 Coefficient tree for the new N^5 multi-coefficient hybrid methods based on MC-CO/3.

4.1. N^7 methods

Table 2 gives the errors for the new N^7 hybrid multi-correlation methods and for G3SX, G3SX(MP3), CBS-Q and MCG3/3. The errors in all tables are calculated against the electronic-energy database described in section 2. To distinguish the three methods, we use a suffix to specify the exchange–correlation functional, and the abbreviations of the suffixes we used are:

MPWB MPW Exchange + B95 correlation

MPW MPW Exchange + PW91 correlation

TS TPSS Exchange + KCIS correlation

The cost (in all tables of this article) is the sum of the times to calculate the gradients for the two molecules, 1-phosphino-methanol and 2,2-dichloro-1-ethanol, with a single 500 MHz R14000 processor on a Silicon Graphics Origin 3800 normalized by dividing by the sum of the times for MP2/6-31G(2df,p) gradient calculations on the same computer.

Table 2 summarizes the errors of various methods. In the tables, “HCO compounds” are compounds containing at most hydrogen, carbon and oxygen atoms. From Table 2, we can see that all three new N^7 hybrid multi-coefficient methods outperform the G3SX method, which is one of the most accurate and most expensive Gaussian- n type methods. Note that the G3SX

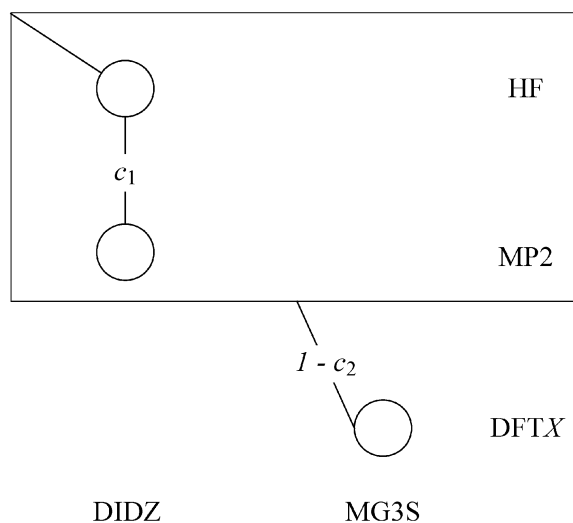


Fig. 5 Coefficient tree for the N^5 MC3-type multi-coefficient hybrid methods.

Table 2 Mean errors and cost for N^7 methods ^a

Quantity ^b	Item ^c	BPM ^b	MCG3-MPWB	MCG3-TS	MCG3-MPW	G3SX	MCG3/3	G3SX (MP3)	CBS-Q
MUEPB	Atomization energy (109)	4.71	0.16	0.16	0.16	0.18	0.22	0.21	0.29
	HCO compound (54)	6.23	0.08	0.09	0.08	0.11	0.13	0.13	0.21
	Containing second row atom (34)	2.78	0.37	0.39	0.39	0.44	0.52	0.50	0.55
	Other (21)	3.84	0.24	0.20	0.21	0.18	0.23	0.19	0.30
MUE	Barrier height (38)		0.54	0.56	0.54	0.67	0.84	0.81	0.73
MUE	Ionization potential (13)		0.67	0.73	0.69	1.07	0.95	1.34	1.26
MUE	Electron affinity (13)		0.86	0.82	0.97	1.06	0.92	1.00	1.12
AMUE	All data (173)		0.49	0.50	0.51	0.64	0.66	0.73	0.74
RMSE	All data (173)		0.95	1.00	1.00	1.19	1.27	1.32	1.67
Max (–)	All data (173)		–2.86 (BCl ₃)	–3.16 (PF ₃)	–3.55 (PF ₃)	–5.17 (CH ₃ CH ₂ O)	–4.79 (NF ₃)	–5.71 (CH ₃ CH ₂ O)	–4.09 (C ₄ H ₆) ^e
Max (+)	All data (173)		3.12 (O ₂ [–])	3.00 (O ₂ [–])	3.07 (O ₂ [–])	3.75 (AlCl ₃)	2.94 (CN)	3.92 (AlCl ₃)	8.33 (C ₂ Cl ₄)
Cost ^d			46.2	46.2	46.1	460	45.0	66	79

^a QCISD/MG3 geometries are used. ^b All mean unsigned error (MUE) in atomization energies are in a per bond basis (MUEPB). These errors are computed by dividing the MUE for the entire data set or the subset corresponding to a given row by the average number of bonds per molecules (BPM) for that set. AMUE is defined as 1/3 times mean unsigned error per bond in atomization energies plus 1/3 times mean signed error in barrier heights plus 1/6 each times the mean unsigned errors in ionization potentials and electron affinities. RMSE is the root mean square error for the 173 data without weighting. ^c Number in parentheses is number of data for that row. ^d The cost is the sum of the times to calculate the gradients for the two molecules, 1-phosphinomethanol and 2,2-dichloro-1-ethanol, with a single 500 MHz R14000 processor on a Silicon Graphics Origin 3800 normalized by dividing by the sum of the times for MP2/6-31G(2df,p) gradient calculations on the same computer. ^e C₄H₆ (bicyclobutane).

method is about 10 times more expensive than the new multi-coefficient hybrid methods.

If we compare the new hybrid methods to the MCG3/3 method, we can see that the costs of the new hybrid multi-coefficient methods are about the same as the MCG3/3 method, but the MUE for atomization energy calculated by the new methods is reduced about 25% as compared to the MCG3/3 method.

Table 2 also shows that CBS-Q and G3SX(MP3) are significantly more expensive than the new hybrid methods, but much less accurate. The DFT component in the new kind of method contains information (such as static correlation, which

is present in DFT exchange) that is not present in the perturbative triple excitation method, QCISD(T). In principle, the DFT-type calculation also explicitly takes into account core correlation effects, but the basis sets used here do not have enough flexibility in the core region to treat this reliably. The semiempirical parameters determined in this paper combine WFT methods with DFT in a way that empirically blends these features to extrapolate toward complete configuration interaction.

The magnitude of worst errors for the new methods is much less than that for the G3SX, G3SX(MP3) and CBS-Q methods.

Table 3 Mean errors for N^6 methods based on MC-QCISD ^{ab}

Quantity	Item	MCQCISD-TS	MCQCISD-MPWB	MCQCISD-MPW	MC-QCISD	QCISD/MG3
MUEPB	Atomization energy (109)	0.17	0.20	0.20	0.36	5.01
	HCO compound (54)	0.09	0.10	0.09	0.21	4.44
	Containing second row atom (34)	0.41	0.43	0.48	0.70	6.05
	Other (21)	0.26	0.38	0.34	0.56	6.26
MUE	Barrier height (38)	0.78	0.66	0.69	1.22	2.83
MUE	Ionization potential (13)	0.99	1.28	0.87	1.95	3.28
MUE	Electron affinity (13)	0.69	0.72	0.85	1.38	5.20
AMUE	All data (173)	0.60	0.62	0.59	1.08	4.03
RMSE	All data (173)	1.07	1.22	1.24	2.02	21.70
Max (–)	All data (173)	–3.71 (P ₂)	–4.06 (N ₂)	–5.62 (Si ₂ H ₆)	–7.39 (P ₂)	–61.50 (C ₅ H ₅ N)
Max (+)	All data (173)	3.28 (C ₂ Cl ₄)	4.06 (C ₂ Cl ₄)	4.05 (C ₂ Cl ₄)	6.24 (AlCl ₃)	9.68 (O [–])
Cost		6.0	6.0	6.0	5.5	55

^a QCISD/MG3 geometries are used. ^b See notes below Table 2 for the definition of MUEPB, AMUE and cost.

Table 4 Mean errors for N^6 methods based on MC-UT ^{ab}

Quantity	Item	MCUT-TS	MCUT-MPWB	MCUT-MPW	MC-UT
MUEPB	Atomization energy (109)	0.23	0.26	0.26	0.45
	HCO compound (54)	0.13	0.12	0.12	0.25
	Containing second row atom (34)	0.46	0.52	0.57	0.85
	Other (21)	0.37	0.52	0.49	0.84
MUE	Barrier height (38)	0.79	0.71	0.78	2.52
MUE	Ionization potential (13)	1.05	1.22	1.40	1.93
MUE	Electron affinity (13)	0.77	0.74	1.12	1.43
AMUE	All data (173)	0.64	0.65	0.77	1.55
RMSE	All data (173)	1.50	1.71	1.77	3.05
Max (–)	All data (173)	–10.65 (CN)	–11.88 (CN)	–11.77 (CN)	–19.15 (P ₂)
Max (+)	All data (173)	3.88 (C ₂ Cl ₄)	5.05 (C ₂ Cl ₄)	4.80 (C ₆ Cl ₆)	5.32 (AlF ₃)
Cost		5.5	5.5	5.5	5.3

^a QCISD/MG3 geometries are used. ^b See notes below Table 2 for the definition of MUEPB, AMUE and cost.

From Table 2, we can see that the main sources of error in atomization energies for the new methods are due to the molecules containing second row atoms. This is also the problem for other methods in Table 2. The relatively worse performance for the second row molecules is assumed to be due largely to the imperfection of the basis set.

4.2. N^6 methods based on MC-QCISD

The mean errors for the N^6 methods based on MC-QCISD/3 are given in the Table 3. The results for the MC-QCISD/3 and QCISD/MG3 methods are also listed in Table 3.

The performances of the three hybrid methods in Table 2 are quite similar, however, the three new N^6 methods based on MC-QCISD give different performances. The MCQCISD-TS method, which is based on the TS DFT method, gives the overall best performance as indicated by its low AMUE and RMSE for all data. Note that this N^6 method also outperforms the G3SX method and all other previous N^7 methods listed in Table 2, and the cost of G3SX is about 100 times bigger than the new N^6 hybrid methods in Table 3 even for the two modest size molecules on which we base our timings. The magnitude of worst errors of the MCQCISD-TS method is also smaller than the G3SX, G3SX(MP3) and CBS-Q methods.

From Table 1, we can see that the percentage of HF exchange in the DFT calculation in the MCQCISD-TS method

is 19%, and this low HF exchange percentage causes the barrier heights calculated by the MCQCISD-TS method to be slightly less accurate than the other two hybrid methods in Table 3. All new hybrid methods in Table 3 have analytical gradients in the GAUSSIAN03⁷⁶ program.

The single-level N^6 calculation in Table 3, QCISD/MG3, gives a big systematic error as compared to all the multilevel methods in this paper.

The mean errors for the new hybrid N^6 methods based on MC-UT/3 are given in Table 4. All three new hybrid methods in Table 4 do not require a post-SCF iterative process, and they outperform the MC-QCISD/3 method.

If we compared the new hybrid methods in Table 4 to those in Table 3, we find that the hybrid methods in Table 3 significantly outperform the ones in Table 4. This result shows the superiority of QCISD to MP4SDQ for use as a component in multilevel methods. However the MCUT methods have the advantage that the N^6 step in Table 4 is noniterative, whereas the N^6 step in Table 3 is iterative.

4.3. N^5 methods

The mean errors for the new hybrid N^5 methods based on MC-CO/3 as well as the mean errors for the MC3BB, MC3MPW, and MC3MPWB methods are listed in Table 5. All methods in Table 5 have analytical Hessians in GAUSSIAN03.⁷⁶

Table 5 Mean errors for N^5 methods ^{ab}

Quantity	Item	MCCO-TS	MCCO-MPWB	MCCO-MPW	MC3BB	MC3MPWB	MC3TS	MC-CO	MC3MPW
MUEPB	Atomization energy (109)	0.41	0.49	0.57	0.61	0.67	0.69	0.67	0.82
	HCO compound (54)	0.29	0.30	0.33	0.39	0.42	0.36	0.30	0.41
	Containing second row atom (34)	0.74	0.91	1.23	1.17	1.30	1.81	1.52	2.12
	Other (21)	0.55	0.80	0.81	0.88	1.01	0.77	1.19	1.00
MUE	Barrier height (38)	0.97	1.12	1.18	0.73	0.79	0.71	3.12	0.76
MUE	Ionization potential (13)	1.43	1.56	2.34	2.17	2.42	2.34	2.09	2.34
MUE	Electron affinity (13)	1.77	1.59	2.93	4.25	4.42	4.57	2.09	3.54
AMUE	All data (173)	0.99	1.06	1.46	1.51	1.63	1.62	1.96	1.51
RMSE	All data (173)	2.27	2.57	3.14	3.30	3.48	3.87	4.14	4.27
Max (–)	All data (173)	–9.88 (CN)	–10.58 (CN)	–12.10 (CN)	–10.30 (CN)	–10.82 (CN)	–15.52 (SiF ₄)	–21.66 (CN)	–16.14 (SiF ₄)
Max (+)	All data (173)	6.70 (C ₅ H ₅ N)	8.01 (C ₂ Cl ₄)	9.36 (C ₆ H ₆)	9.94 (C ₅ H ₈) ^c	9.63 (C ₅ H ₈) ^c	12.32 (C ₆ H ₆)	12.01 (AlCl ₃)	12.32 (C ₆ H ₆)
Cost		5.2	5.2	5.2	2.0	2.0	2.0	4.9	1.8

^a QCISD/MG3 geometries are used. ^b See notes below Table 2 for the definition of MUEPB, AMUE, and cost. ^c C₅H₈ (spiropentane).

Table 6 Mean errors for N^4 methods ^{ab}

Quantity	Item	MPW1B95	TPSS1KCIS	MPW1PW91	MPWB1K	B3LYP	TPSSh
MUEPB	Atomization energy (109)	0.63	0.67	0.89	0.98	0.92	0.98
	HCO compound (54)	0.48	0.40	0.51	0.60	0.51	0.59
	Containing second row atom (34)	0.97	1.49	2.06	1.56	2.39	2.28
	Other (21)	0.84	0.86	1.13	1.94	0.92	1.12
MUE	Barrier height (38)	3.02	4.69	3.55	1.29	4.23	5.97
MUE	Ionization potential (13)	2.14	2.63	3.72	2.05	4.72	3.17
MUE	Electron affinity (13)	2.91	2.81	2.62	4.11	2.29	2.81
AMUE	All data (173)	2.06	2.69	2.54	1.78	2.88	3.31
RMSE	All data (173)	3.84	4.46	4.97	5.15	5.79	5.90
Max (–)	All data (173)	–10.38 (SiF ₄)	–21.61 (SiF ₄)	–24.59 (SiF ₄)	–16.83 (SiF ₄)	–26.71 (SiCl ₄)	–23.47 (SiF ₄)
Max (+)	All data (173)	12.73 (C ₅ H ₈) ^c	8.96 (O ⁺)	9.05 (O ₂ ⁺)	10.23 (OH [–])	12.95 (O ⁺)	13.54 (Si ₂ H ₆)
Cost		1.8	1.8	1.7	1.8	1.7	1.8

^a QCISD/MG3 geometries are used. All DFT calculations are performed with MG3S basis set. ^b See notes below Table 2 for the definition of MUEPB, AMUE and cost. ^c C₅H₈ (spiropentane).

The overall performances of the three hybrid methods based on MC-CO are better than the MC3-type hybrid methods. However, in part because the way they were optimized, the MC3-type methods are better for barrier heights calculations, and they have lower cost, so the MC3-type methods are very suitable for dynamics calculation.

Notice that for all the doubly hybrid methods in Table 1, the last coefficient (c_8 , c_5 , c_3 , or c_2 , depending on the method) is the fraction of WFT in the doubly hybrid method, and this fraction decreases as the level of WFT is lowered. Thus the highest-level methods are 72.7–81.5% WFT, plus the additional Hartree–Fock exchange in the DFT part, whereas the lowest-level ones are 20.5–33.6% WFT, plus the additional Hartree–Fock exchange in the DFT part. This is a very reasonable trend, and it increases our confidence in the physicality of the doubly hybrid methods.

4.4. N^4 methods

The mean errors for the new N^4 method, TPSS1KCIS, are listed in Table 6 with mean errors for some previous developed DFT methods, in particular, MPW1B95,⁶³ MPW1PW91,³¹ MPWB1K,⁶³ B3LYP²⁵ and TPSSh.⁵⁷ TPSSh is a hybrid meta DFT method using TPSS exchange, TPSS correlation and 10% HF exchange; Table 6 shows that TPSS1KCIS outperform TPSSh⁵⁷ for all properties tested in this paper. TPSS1KCIS is slightly worse than B3LYP for barrier height and electron

affinity calculations, but TPSS1KCIS significantly outperform B3LYP for atomization energy and ionization potential calculations.

4.5. Test of the methods for molecules not in the training set

As described in section 2.5, we used the TAE6/04 database of atomization energies of six molecules that are not in our training set. We tested an N^7 method, namely MCG3-TS, and a N^6 method, MCQCISD-TS, and we compare them to G3SX and G3SX(MP3). The results are shown in Table 7, which gives references for the accurate data.^{6,74,77–79}

O₃ is a very difficult case for CCSD(T) calculations with basis set extrapolation; the W2 theory gives an error of about 3 kcal mol^{–1}.⁷ Helgaker and coworkers⁸⁰ used CCSD(T)/cc-pcV(56)Z calculations with core correlation and relativistic contribution, and they still had an error about 2.6 kcal mol^{–1}. Table 7 shows that G3SX and G3SX(MP3) work well for this molecule. The MCG3-TS method gives a 1.25 kcal mol^{–1} error and MCQCISD-TS method gives a 5.27 kcal mol^{–1} error. Apparently triple excitations are important for treating this molecule.

For the two molecules, ClCN and OCS, that are not in our training set and not in G3/99 set, our two methods perform better than G3SX and G3SX(MP3).

P₄ is a difficult case for the scaled G3 methods, and it is the source of the maximum error for all scaled G3 methods and

Table 7 Comparison of methods for TAE6/04 test set ^a

Molecule	Experimental ^b atomization energy	Deviation (calculation – experiment)			
		MCG3-TS	MCQCISD-TS	G3SX	G3SX(MP3)
O ₃	146.10	–1.25	–5.27	0.87	0.08
ClCN	284.52	0.21	–0.03	0.16	0.38
OCS	334.24	1.79	0.77	2.06	2.61
P ₄	289.90	–0.90	–0.75	–8.96	–11.21
C ₈ H ₁₈ (<i>n</i> -octane)	2478.23	–1.07	–1.54	0.69	0.09
C ₁₀ H ₈ (naphthalene)	2162.16	–0.96	–2.59	0.14	0.76
MUE ^c		1.03	1.83	2.14	2.52
MUEPB ^c		0.13	0.24	0.28	0.33

^a QCISD/MG3 geometry is used for O₃, and B3LYP/6-31(2df,p) geometries are used for P₄, ClCN, OCS, C₈H₁₈, and C₁₀H₈. ^b The experimental atomization energy of O₃ is from our Database/4.^{74,77} The experimental value for ClCN and OCS are taken from Martin and Oliveira.⁶ The experimental value for P₄ is obtained by adding the scaled fundamental zero point vibrational energy (ZPVE) to D_0 and D_0 and fundamental ZPVE are taken from the Computational Chemistry Comparison and Benchmark Data Base (CCCBDB).⁷⁸ The experimental values for C₈H₁₈ and C₁₀H₈ are obtained by subtracting the thermal vibrational-rotational-translational contribution and the scaled ZPVE (0.9854) at the B3LYP/6-31(2df,p) level and the experimental heat of formation⁷⁹ of both molecules. ^c See notes below Table 2 for the definition of MUE, MUEPB. The average number of bonds per molecule is 7.67 for the TAE6/04 database.

their variants.⁶⁸ Table 7 shows that our two new methods perform well for this molecule, and the errors for both multi-coefficient extrapolated DFT methods are less than 1 kcal mol⁻¹.

For the two largest molecules in the G3/99 test, C₈H₁₈ (*n*-octane) and C₁₀H₈ (naphthalene), our new methods perform worse than G3SX and G3SX(MP3).

The MUEs and MUEPBs in Table 7 are consistent with previous tables. Both our *N*⁷ method and our *N*⁶ method outperform G3SX and G3SX(MP3).

From Table 7, we can conclude that our parametrizations are not overfitted by the training set; that is, they can usefully be applied to the systems that are not in the training set.

5. Concluding remarks

In this paper, we developed several doubly hybrid multi-coefficient methods which empirically mix WFT methods with DFT methods. The new methods have different costs, and they are suitable for various applications.

The new methods are more accurate than pure WFT multi-level methods and they have lower cost. In particular we developed several new methods (MCG3-MPWB, MCG3-MPW, MCG3-TS and MCQCISD-TS) that outperform G3SX and several less expensive methods (MCQCISD-MPWB, MCQCISD-MPW) that outperform G3SX(MP3). An even less expensive method (MCUT-TS) still outperforms CBS-Q.

During the course of the development of modern DFT methods, it has clearly been established that including some HF exchange provides accuracy that can not be obtained by pure DFT exchange functionals, Becke's hybrid DFT methods, justified by adiabatic connection theory,²⁴ revolutionized quantum chemistry. Now hybrid DFT and hybrid meta DFT have become the most popular DFT methods in computational chemistry. On the other hand, in the present study, we found that the DFT calculations contain information which cannot be easily mimicked by WFT methods even with quasiperturbative triple excitations. Our empirical multi-coefficient approach combines the two different types of contributions together, and the new resulting methods are more accurate and more efficient than either type of methods used separately. We have named the new methods as extrapolated versions of DFT, but one can also consider them to involve improved version (based on DFT) of the high-level correction^{1,2} of conventional multilevel WFT theory.

Given that the number and quality of the exchange and correlation functionals increase year by year, we can expect that even more efficient and accurate hybrid multi-coefficient methods can be constructed for theoretical chemistry in the future.

Acknowledgements

This work was supported in part by the US Department of Energy, Office of Basic Energy Sciences.

References

- 1 L. A. Curtiss, K. Raghavachari, G. W. Trucks and J. A. Pople, *J. Chem. Phys.*, 1991, **94**, 7221.
- 2 L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov and J. A. Pople, *J. Chem. Phys.*, 1998, **109**, 7764.
- 3 L. A. Curtiss, P. C. Redfern, K. Raghavachari and J. A. Pople, *J. Chem. Phys.*, 2001, **114**, 108.
- 4 J. W. Ochterski, G. A. Petersson and J. A. Montgomery, *J. Chem. Phys.*, 1996, 2598.
- 5 J. A. Montgomery, M. J. Frisch, J. W. Ochterski and G. A. Petersson, *J. Chem. Phys.*, 2000, 6532.
- 6 J. M. L. Martin and G. d. Oliveira, *J. Chem. Phys.*, 1999, **111**, 1843.

- 7 A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay and J. Gauss, *J. Chem. Phys.*, 2004, **120**, 4129.
- 8 P. L. Fast, J. Corchado, M. L. Sanchez and D. G. Truhlar, *J. Phys. Chem. A*, 1999, **103**, 3139.
- 9 P. L. Fast, J. Corchado, M. L. Sanchez and D. G. Truhlar, *J. Phys. Chem. A*, 1999, **103**, 5129.
- 10 P. L. Fast, J. Corchado, M. L. Sanchez and D. G. Truhlar, *J. Chem. Phys.*, 1999, **110**, 11679.
- 11 P. L. Fast, M. L. Sanchez and D. G. Truhlar, *Chem. Phys. Lett.*, 1999, **306**, 407.
- 12 C. M. Tratz, P. L. Fast and D. G. Truhlar, *PhysChemComm*, 1999, **2**, 14.
- 13 P. L. Fast and D. G. Truhlar, *J. Phys. Chem. A*, 2000, **104**, 6111.
- 14 B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2002, **106**, 842.
- 15 B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2003, **107**, 3898.
- 16 M. S. Gordon and D. G. Truhlar, *J. Am. Chem. Soc.*, 1986, 5412.
- 17 P. Hohenberg and W. Kohn, *Phys. Rev. B*, 1964, **136**, 864.
- 18 W. Kohn and L. J. Sham, *Phys. Rev. A*, 1965, **140**, 1133.
- 19 J. P. Perdew and A. Zunger, *Phys. Rev. B*, 1981, **23**, 5048.
- 20 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098.
- 21 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785.
- 22 J. P. Perdew, in *Electronic Structure of Solids '91*, ed. P. Ziesche and H. Eschig, Akademie Verlag, Berlin, 1991, p. 11.
- 23 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372.
- 24 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.
- 25 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623.
- 26 A. D. Becke, *J. Chem. Phys.*, 1996, **104**, 1040.
- 27 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.
- 28 J. P. Perdew, M. Ernzerhof and K. Burke, *J. Chem. Phys.*, 1996, **105**, 9982.
- 29 A. D. Becke, *J. Chem. Phys.*, 1997, **107**, 8554.
- 30 M. Filatov and W. Thiel, *Mol. Phys.*, 1997, 847.
- 31 C. Adamo and V. Barone, *J. Chem. Phys.*, 1998, **108**, 664.
- 32 H. L. Schmider and A. D. Becke, *J. Chem. Phys.*, 1998, **108**, 9624.
- 33 T. V. Voorhis and G. E. Scuseria, *J. Chem. Phys.*, 1998, **109**, 400.
- 34 F. A. Hamprecht, A. J. Cohen, D. J. Tozer and N. C. Handy, *J. Chem. Phys.*, 1998, **109**, 6264.
- 35 J. B. Krieger, J. Chen and G. J. Iafrate, *Int. J. Quantum Chem.*, 1998, **69**, 255.
- 36 J. Rey and A. Savin, *Int. J. Quantum Chem.*, 1998, **69**, 581.
- 37 J. B. Krieger, J. Chen, G. J. Iafrate and A. Savin, in *Electron Correlations and Materials Properties*, ed. A. Gonis and N. Kioussis, Plenum, New York, 1999.
- 38 E. Engel and R. M. Dreizler, *J. Comput. Chem.*, 1999, **20**, 31.
- 39 B. Hammer, L. B. Hansen and J. K. Nørskov, *Phys. Rev. B*, 1999, **59**, 7413.
- 40 J. P. Perdew, S. Kurth, A. Zupan and P. Blaha, *Phys. Rev. Lett.*, 1999, **82**, 2544.
- 41 A. D. Becke, *J. Chem. Phys.*, 2000, **112**, 4020.
- 42 B. J. Lynch, P. L. Fast, M. Harris and D. G. Truhlar, *J. Phys. Chem. A*, 2000, **104**, 4811.
- 43 E. Proynov, H. Chermette and D. R. Salahub, *J. Chem. Phys.*, 2000, **113**, 10013.
- 44 N. C. Handy and A. J. Cohen, *Mol. Phys.*, 2001, **99**, 403.
- 45 W.-M. Hoes, A. J. Cohen and N. C. Handy, *Chem. Phys. Lett.*, 2001, **341**, 319.
- 46 B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2001, **105**, 2936.
- 47 P. J. Wilson, T. J. Bradley and D. J. Tozer, *J. Chem. Phys.*, 2001, **115**, 9233.
- 48 C. Adamo and V. Barone, *J. Chem. Phys.*, 2002, **116**, 5933.
- 49 A. D. Boese and N. C. Handy, *J. Chem. Phys.*, 2002, **116**, 9559.
- 50 J. Baker and P. Pulay, *J. Chem. Phys.*, 2002, **117**, 1441.
- 51 J. Toulouse, A. Savin and C. Adamo, *J. Chem. Phys.*, 2002, **117**, 10465.
- 52 G. Menconi and D. J. Tozer, *Chem. Phys. Lett.*, 2002, **360**, 38.
- 53 B. J. Lynch, Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2003, **107**, 1384.
- 54 V. V. Karasiev, *J. Chem. Phys.*, 2003, **118**, 8576.
- 55 A. D. Boese, J. M. L. Martin and N. C. Handy, *J. Chem. Phys.*, 2003, **119**, 3005.
- 56 V. Guner, K. S. Khuong, A. G. Leach, P. S. Lee, M. D. Bartberger and K. N. Houk, *J. Phys. Chem. A*, 2003, **107**, 1445.
- 57 V. N. Staroverov, G. E. Scuseria, J. Tao and J. P. Perdew, *J. Chem. Phys.*, 2003, **119**, 12129.
- 58 J. Y. Hsu, *Phys. Rev. Lett.*, 2003, **91**, 133001.

- 59 J. Tao, J. P. Perdew, V. N. Staroverov and G. E. Scuseria, *Phys. Rev. Lett.*, 2003, **91**, 146401.
- 60 X. Xu and W. A. Goddard, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 2673.
- 61 Y. Zhao, B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 2715.
- 62 Q. S. Li, X. D. Xu and S. Zhang, *Chem. Phys. Lett.*, 2004, 20.
- 63 Y. Zhao and D. G. Truhlar, *J. Phys. Chem.*, 2004, **108**, 6908.
- 64 Y. Zhao, B. J. Lynch and D. G. Truhlar, *J. Phys. Chem. A*, 2004, **108**, 4786.
- 65 I. Rossi and D. G. Truhlar, *Chem. Phys. Lett.*, 1995, **234**, 64.
- 66 D. Fylstra, L. Lasdon, J. Watson and A. Waren, *Interfaces*, 1998, **28**, 29.
- 67 L. A. Curtiss, K. Raghavachari, P. C. Redfern and J. A. Pople, *J. Chem. Phys.*, 2000, **112**, 7374.
- 68 L. A. Curtiss, P. C. Redfern, K. Raghavachari and J. A. Pople, *Chem. Phys. Lett.*, 2002, **359**, 390.
- 69 C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618.
- 70 W. J. Hehre, L. Radom, P. v. R. Schleyer and J. A. Pople, *Ab initio Molecular Orbital Theory*, Wiley, New York, 1986.
- 71 J. A. Pople, M. Head-Gordon and K. Raghavachari, *J. Chem. Phys.*, 1987, **87**, 5968.
- 72 L. A. Curtiss, P. C. Redfern, K. Raghavachari, V. Rassolov and J. A. Pople, *J. Chem. Phys.*, 1999, **110**, 4703.
- 73 M. J. Frisch, J. A. Pople and J. S. Binkley, *J. Chem. Phys.*, 1984, **80**, 3265.
- 74 B. J. Lynch, Y. Zhao and D. G. Truhlar, <http://comp.chem.umn.edu/database>.
- 75 W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257.
- 76 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, T. V. Jr., K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *GAUSSIAN 03 (Revision C.01 ed.)*, Gaussian, Inc., Pittsburgh PA, 2003.
- 77 B. J. Lynch, Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A*, 2004, submitted.
- 78 NIST, Computational Chemistry Comparison and Benchmark DataBase, <http://srdata.nist.gov/cccbdb/>.
- 79 NIST Chemistry Webbook, <http://webbook.nist.gov/chemistry>.
- 80 T. Helgaker, W. Klopper, A. Halkier, K. L. Bak, P. Jorgensen and J. Olsen, in *Quantum Mechanical Prediction of Thermochemical Data*, ed. J. Cioslowski, Kluwer, Dordrecht, 2001.