High-accuracy extrapolated *ab initio* thermochemistry. III. Additional improvements and overview

Michael E. Harding,[1] Juana Vázquez,[1,2] Branko Ruscic,[3] Angela K. Wilson,[4] Jürgen Gauss,[1] and John F. Stanton[2,a)]

[1]*Institut für Physikalische Chemie, Universität Mainz, D-55099 Mainz, Germany*
[2]*Institute for Theoretical Chemistry, Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712, USA*
[3]*Chemical Sciences and Engineering Division, Argonne National Laboratory, Argonne, Illinois 60439, USA*
[4]*Department of Chemistry, University of North Texas, Denton, Texas 76203, USA*

Effects of increased basis-set size as well as a correlated treatment of the diagonal Born-Oppenheimer approximation are studied within the context of the high-accuracy extrapolated *ab initio* thermochemistry (HEAT) theoretical model chemistry. It is found that the addition of these ostensible improvements does little to increase the overall accuracy of HEAT for the determination of molecular atomization energies. Fortuitous cancellation of high-level effects is shown to give the overall HEAT strategy an accuracy that is, in fact, higher than most of its individual components. In addition, the issue of core-valence electron correlation separation is explored; it is found that approximate additive treatments of the two effects have limitations that are significant in the realm of <1 kJ mol^{-1} theoretical thermochemistry. © *2008 American Institute of Physics*.
[DOI: 10.1063/1.2835612]

I. INTRODUCTION

Converging advances in theoretical methods and technology available for computation as well as precision in thermochemical knowledge has recently produced several theoretical approaches that can reliably predict molecular enthalpies of formation (and, by consequence, the much more important property of bond energies) to an accuracy that exceeds the standard of "chemical accuracy" (<1 kcal mol^{-1}). At this point, we need to parenthetically add a word of caution. Namely, in experimental thermochemistry, the "accuracy" or, in the vernacular, the "error bar," is customarily expressed as a 95% confidence limit (rather than, for example, a standard deviation). The same holds for the chemical accuracy definition: the target accuracy of ± 1 kcal mol^{-1} should have the intrinsic meaning of a 95% confidence limit. The general use of a 95% confidence limit (approximately equal to two standard deviations) in thermochemistry was proposed originally by Rossini[1] and subsequently embraced by virtually all thermochemical compilations, such as CODATA,[2] Janaf,[3] Gurvich *et al.*,[4] etc., and, as a consequence, it is the expected level of coverage when attaching an uncertainty to thermochemical quantities such as enthalpies of formation or atomization energies. Contrary to this, other measures of fidelity of a class of electronic structure results have proliferated in theoretical thermochemistry; a favorite one being the mean absolute deviation (or MAD) of a computational method, which severely underestimates the thermochemical uncertainty of the computed result. In fact, the MAD typically needs to be multi-plied by a factor of ~ 2.5 (or, more generally, a factor between 2 and 3) in order to achieve a 95% confidence coverage.

Recent methods based on the focal-point extrapolation,[5] the Weizmann (Wn) family of methods,[6–8] and the high-accuracy extrapolated *ab initio* thermochemistry (HEAT) protocol[9,10] have exhibited or exceeded thermochemical accuracies of <1 kJ mol^{-1}, i.e., about five times better than chemical accuracy. Common to these latter and most accurate methods is a treatment of electron correlation effects that goes beyond the so-called gold standard of quantum chemistry—the now perhaps venerable coupled-cluster singles and doubles model with a perturbative treatment of triple excitations[11] [CCSD(T)]—and these have indeed been shown to be essential to surpass the <1 kJ mol^{-1} barrier. It also seems necessary to perform basis-set extrapolations[12,13] of the components of the Born-Oppenheimer nonrelativistic electronic energy [the Hartree-Fock self-consistent field (HF-SCF) and, especially, the electron correlation contribution] to achieve high accuracy. None of this work would have been possible (or, perhaps more precisely, able to be calibrated) without the Active Thermochemical Tables (ATcT) paradigm[14] that has been developed and applied by Ruscic *et al.* to establish precise and tight bounds on the thermochemical properties of a large number of small molecules. This has allowed the quantitative accuracy of computational thermochemistry to be judged with the precision that is demanded to gauge the accuracy of such high-level calculations.

This is the third paper on HEAT, a model chemistry that is similar, and indeed seems to be mutually convergent with the Wn family of methods[6–8] developed by Martin and co-workers in Israel. Differences between the most elaborate HEAT methods and the W4.n approaches[8] of Martin and

---

a)Electronic mail: jfstanton@mail.utexas.edu.

co-workers are almost negligible in theory and so small numerically as to be of no consequence for analysis of experimental data; all give accuracies of better than <0.5 kJ mol^{-1}. The point of this paper is twofold. First, we document results obtained with the HEAT protocol employing larger basis-set extrapolations than those that have been used previously as well as a size-consistent and correlated treatment of the diagonal Born-Oppenheimer correction (DBOC).[15] After this, the results are discussed, with an emphasis on remaining sources of error and a discussion of how rather small (ca. 0.1–0.3 kJ mol^{-1}) intrinsic errors in the calculations have a tendency toward fortuitous cancellation. We also present a table of molecular electronic energies (both HF-SCF and correlation contributions) that we believe to be accurate to ca. 1 mhartree (2.5 kJ mol^{-1}) (Ref. 16) that can be used for benchmarking purposes by others using methods such as quantum Monte Carlo.

We believe that the HEAT protocol has been well documented in the previous literature,[9,10] and this paper will complete this process. Others in the field of computational thermochemistry and chemical physics are encouraged to use these data; the contributions to the total HEAT atomization energies for the molecules presented here are listed in Table I of this paper; raw absolute energies of atoms and molecules are available upon request from any of the authors. Furthermore, others are encouraged to carry out calculations on species of interest that are not listed in this paper.

II. METHODOLOGY

The total ground-state energy of atoms and molecules in the HEAT paradigm[9] is given by the sum

$$E = E_{HF}^{\infty} + \Delta E_{CCSD(T)}^{\infty} + \Delta E_{CCSDT} + \Delta E_{HLC} + \Delta E_{rel} + \Delta E_{SO}$$
$$+ \Delta E_{DBOC} + \Delta E_{ZPE}, \quad (1)$$

where E_{HF}^{∞} is the estimate of the Hartree-Fock limit energy (restricted HF for closed shells and unrestricted HF for open-shell systems), $\Delta E_{CCSD(T)}^{\infty}$ is the estimate for the CCSD(T) correlation energy, $\Delta E_{CCSDT} + \Delta E_{HLC}$ collectively account for deficiencies in the CCSD(T) treatment of electron correlation, ΔE_{rel} accounts for so-called scalar relativistic effects via perturbation theory, ΔE_{SO} is the difference between the true ground-state energy of the system and the weighted average of spin-orbit levels that are represented by the nonrelativistic electronic Schrödinger equation (this vanishes for closed-shell systems and all nondegenerate states), ΔE_{DBOC} is the diagonal Born-Oppenheimer correction, and ΔE_{ZPE} is the vibrational zero-point energy. All of these contributions, and the procedure used to calculate them, are described in detail in Parts I and II of this series[9,10] and the interested reader is referred there for details.

The nonrelativistic electronic energy contributions E_{HF}^{∞}, $\Delta E_{CCSD(T)}^{\infty}$, and ΔE_{CCSDT} involve extrapolation of energies obtained with various basis sets. Like most others in the field of computational thermochemistry, the HEAT protocol is based on the hierarchy of correlation-consistent basis sets developed by Dunning and co-workers.[17–21] Moreover, but unlike most if not all other theoretical model chemistries, the HEAT paradigm does not treat "valence" and "core" correlation separately. Rather, to avoid the approximations inherent in any such separation, both are determined together with the aug-cc-pCVXZ series of basis sets[21] that were specifically designed to treat the correlation of electrons associated with the core (1s-type) molecular orbitals. The choice of which members of the sequence are selected for the extrapolation differentiates various HEAT energies: HEAT-345 is based on extrapolation of HF-SCF energies using the X=T, Q, and 5 members of the sequence of basis sets and CCSD(T) energies obtained with the X=4 and 5 basis sets. Similarly, HEAT-456 involves E_{HF}^{∞} based on X=4, 5, and 6 and $\Delta E_{CCSD(T)}^{\infty}$ based on X=5 and 6, HEAT-567 would come from E_{HF}^{∞} based on X=5, 6, and 7 with $\Delta E_{CCSD(T)}^{\infty}$ based on X=6 and 7, etc. In this paper, the HEAT-456 results are presented for the first time. ΔE_{CCSDT} is always determined in the manner described in Ref. 9, while various methods can be used for ΔE_{HLC}. In principle, ΔE_{HLC} is given by the difference between the full configuration interaction (FCI) result and the full CCSDT (Ref. 22) correlation energy. However, since FCI calculations are not possible for most chemically interesting systems, these energies are determined from methods that treat only quadruple or quadruple and quintuple excitations. The original HEAT approach now classified as HEAT-345Q used $\Delta E_{HLC} = \Delta E_{CCSDTQ} - \Delta E_{CCSDT}$, both obtained with the cc-pVDZ basis set in the frozen-core approximation. In Ref. 10, HEAT-345(Q) and HEAT-345QP were introduced, in which the much more economical CCSDT(Q) approach[23,24] with a perturbative treatment of quadruple excitations was used instead of the costly CCSDTQ approach. Even more expensive CC singles, doubles, triples, quadruples, pentuples (CCSDTQP) calculations were used for the ΔE_{HLC} in order to obtain the "benchmark" best results.

The auxiliary corrections to the energy ΔE_{rel}, ΔE_{SO}, ΔE_{DBOC}, and ΔE_{ZPE} have also been discussed in some detail in Refs. 9 and 10. In this work, the DBOC, which can be viewed as the first-order correction to the electronic energy associated with the nuclear (mass-dependent) kinetic energy operator $T_n(R)$,

$$\Delta E_{adiabatic} = \langle \Psi(r;R) | T_n(R) | \Psi(r;R) \rangle, \quad (2)$$

has been calculated at the coupled-cluster level, exploiting the recent theoretical (and computational) advances detailed in Ref. 15. This contribution is negligible in most practical cases when 1 kcal mol^{-1} accuracy is desired, but becomes important—especially for hydrides—when one approaches the <0.5 kJ mol^{-1} level that both HEAT and the W4.n methods have recently achieved. It is also responsible for some other phenomena usually associated with the Born-Oppenheimer breakdown, such as the principal contribution to the dipole moment of a molecule such as HD.[25]

All calculations have been performed using the ACES II (Mainz-Austin-Budapest version)[26] and MRCC program[27] interfaced to ACESII. Some of the calculations have been performed using the recently developed parallel version of ACESII.[28]

TABLE I. Contributions to atomization energies (in kJ mol^{-1}) for the present set of molecules. The various terms contributing to the total atomization energies are described in the text; the DBOC is that calculated at the CCSD level of theory. Also included are ATcT values based on the Core (Argonne) Thermochemical Network, C(A)TN, Version 1.064. The listed uncertainties for the ATcT benchmark values correspond to 95% confidence limits (approximately two standard deviations).

Species	method	E_{HF}^{∞}	$\Delta E_{CCSD(T)}^{\infty}$	ΔE_{CCSDT}	ΔE_{HLC}	ΔE_{rel}	ΔE_{SO}	ΔE_{DBOC}	ΔE_{ZPE}	Total	ATcT
N$_2$	HEAT-345(Q)	482.94	472.26	−3.14	4.25	−0.48	0.03	−0.00	−14.12	941.74	941.14 ± 0.03
	HEAT-345Q	482.94	472.26	−3.14	3.61	−0.48	0.03	−0.00	−14.12	941.09	941.14 ± 0.03
	HEAT-345QP	482.94	472.26	−3.14	4.08	−0.48	0.03	−0.00	−14.12	941.56	941.14 ± 0.03
	HEAT-456(Q)	483.07	471.49	−3.14	4.25	−0.48	0.03	−0.00	−14.12	941.10	941.14 ± 0.03
	HEAT-456Q	483.07	471.49	−3.14	3.61	−0.48	0.03	−0.00	−14.12	940.46	941.14 ± 0.03
	HEAT-456QP	483.07	471.49	−3.14	4.08	−0.48	0.03	−0.00	−14.12	940.93	941.14 ± 0.03
H$_2$	HEAT-345(Q)	350.81	107.41	−0.00	−0.00	−0.00	0.06	−0.00	−26.07	432.22	432.07 ± 0.00
	HEAT-345Q	350.81	107.41	−0.00	−0.00	−0.00	0.06	−0.00	−26.07	432.22	432.07 ± 0.00
	HEAT-345QP	350.81	107.41	−0.00	−0.00	−0.00	0.06	−0.00	−26.07	432.22	432.07 ± 0.00
	HEAT-456(Q)	350.81	107.31	−0.00	−0.00	−0.00	0.06	−0.00	−26.07	432.11	432.07 ± 0.00
	HEAT-456Q	350.81	107.31	−0.00	−0.00	−0.00	0.06	−0.00	−26.07	432.11	432.07 ± 0.00
	HEAT-456QP	350.81	107.31	−0.00	−0.00	−0.00	0.06	−0.00	−26.07	432.11	432.07 ± 0.00
F$_2$	HEAT-345(Q)	−154.98	316.00	−1.31	3.87	−0.17	0.01	−3.01	−5.50	154.91	154.51 ± 0.13
	HEAT-345Q	−154.98	316.00	−1.31	3.43	−0.17	0.01	−3.01	−5.50	154.47	154.51 ± 0.13
	HEAT-345QP	−154.98	316.00	−1.31	3.59	−0.17	0.01	−3.01	−5.50	154.63	154.51 ± 0.13
	HEAT-456(Q)	−155.08	315.95	−1.31	3.87	−0.17	0.01	−3.01	−5.50	154.76	154.51 ± 0.13
	HEAT-456Q	−155.08	315.95	−1.31	3.43	−0.17	0.01	−3.01	−5.50	154.32	154.51 ± 0.13
	HEAT-456QP	−155.08	315.95	−1.31	3.59	−0.17	0.01	−3.01	−5.50	154.48	154.51 ± 0.13
CO	HEAT-345(Q)	729.86	356.91	−2.37	2.63	−0.65	0.02	−1.20	−12.99	1072.22	1072.13 ± 0.09
	HEAT-345Q	729.86	356.91	−2.37	2.22	−0.65	0.02	−1.20	−12.99	1071.81	1072.13 ± 0.09
	HEAT-345QP	729.86	356.91	−2.37	2.35	−0.65	0.02	−1.20	−12.99	1071.94	1072.13 ± 0.09
	HEAT-456(Q)	730.13	356.53	−2.37	2.63	−0.65	0.02	−1.20	−12.99	1072.11	1072.13 ± 0.09
	HEAT-456Q	730.13	356.53	−2.37	2.22	−0.65	0.02	−1.20	−12.99	1071.70	1072.13 ± 0.09
	HEAT-456QP	730.13	356.53	−2.37	2.35	−0.65	0.02	−1.20	−12.99	1071.83	1072.13 ± 0.09
O$_2$	HEAT-345(Q)	140.36	362.56	−2.07	4.65	−0.79	0.01	−1.64	−9.56	493.52	493.69 ± 0.00
	HEAT-345Q	140.36	362.56	−2.07	4.13	−0.79	0.01	−1.64	−9.56	493.00	493.69 ± 0.00
	HEAT-345QP	140.36	362.56	−2.07	4.56	−0.79	0.01	−1.64	−9.56	493.43	493.69 ± 0.00
	HEAT-456(Q)	140.68	362.54	−2.07	4.65	−0.79	0.01	−1.64	−9.56	493.81	493.69 ± 0.00
	HEAT-456Q	140.68	362.54	−2.07	4.13	−0.79	0.01	−1.64	−9.56	493.29	493.69 ± 0.00
	HEAT-456QP	140.68	362.54	−2.07	4.56	−0.79	0.01	−1.64	−9.56	493.72	493.69 ± 0.00
C$_2$H$_2$	HEAT-345(Q)	1228.98	468.20	−3.03	2.56	−1.16	0.33	−0.76	−69.08	1626.04	1626.16 ± 0.24
	HEAT-345Q	1228.98	468.20	−3.03	2.24	−1.16	0.33	−0.76	−69.08	1625.72	1626.16 ± 0.24
	HEAT-345QP	1228.98	468.20	−3.03	2.56	−1.16	0.33	−0.76	−69.08	1626.04	1626.16 ± 0.24
	HEAT-456(Q)	1229.01	467.41	−3.03	2.56	−1.16	0.33	−0.76	−69.08	1625.29	1626.16 ± 0.24
	HEAT-456Q	1229.01	467.41	−3.03	2.24	−1.16	0.33	−0.76	−69.08	1624.96	1626.16 ± 0.24
	HEAT-456QP	1229.01	467.41	−3.03	2.56	−1.16	0.33	−0.76	−69.08	1625.29	1626.16 ± 0.24
CCH	HEAT-345(Q)	777.35	333.01	0.56	2.46	−1.16	0.10	−0.76	−36.47	1075.10	1075.05 ± 0.25
	HEAT-345Q	777.35	333.01	0.56	2.29	−1.16	0.10	−0.76	−36.47	1074.94	1075.05 ± 0.25
	HEAT-345QP	777.35	333.01	0.56	2.60	−1.16	0.10	−0.76	−36.47	1075.25	1075.05 ± 0.25
	HEAT-456(Q)	777.35	332.29	0.56	2.46	−1.16	0.10	−0.76	−36.47	1074.38	1075.05 ± 0.25
	HEAT-456Q	777.35	332.29	0.56	2.29	−1.16	0.10	−0.76	−36.47	1074.21	1075.05 ± 0.25
	HEAT-456QP	777.35	332.29	0.56	2.60	−1.16	0.10	−0.76	−36.47	1074.53	1075.05 ± 0.25
CF	HEAT-345(Q)	338.46	216.63	−0.60	1.23	−0.71	0.01	−1.42	−7.89	545.72	545.61 ± 0.63
	HEAT-345Q	338.46	216.63	−0.60	0.99	−0.71	0.01	−1.42	−7.89	545.47	545.61 ± 0.63
	HEAT-345QP	338.46	216.63	−0.60	0.93	−0.71	0.01	−1.42	−7.89	545.41	545.61 ± 0.63
	HEAT-456(Q)	338.53	216.55	−0.60	1.23	−0.71	0.01	−1.42	−7.89	545.70	545.61 ± 0.63
	HEAT-456Q	338.53	216.55	−0.60	0.99	−0.71	0.01	−1.42	−7.89	545.46	545.61 ± 0.63
	HEAT-456QP	338.53	216.55	−0.60	0.93	−0.71	0.01	−1.42	−7.89	545.40	545.61 ± 0.63
CH$_2$	HEAT-345(Q)	649.11	149.13	−0.02	0.13	−0.65	0.06	−0.38	−44.95	752.43	752.70 ± 0.26
	HEAT-345Q	649.11	149.13	−0.02	0.14	−0.65	0.06	−0.38	−44.95	752.44	752.70 ± 0.26
	HEAT-345QP	649.11	149.13	−0.02	0.14	−0.65	0.06	−0.38	−44.95	752.45	752.70 ± 0.26
	HEAT-456(Q)	649.11	148.82	−0.02	0.13	−0.65	0.06	−0.38	−44.95	752.12	752.70 ± 0.26

TABLE I.   (*Continued.*)

Species method		E_{HF}^{∞}	$\Delta E_{CCSD(T)}^{\infty}$	ΔE_{CCSDT}	ΔE_{HLC}	ΔE_{rel}	ΔE_{SO}	ΔE_{DBOC}	ΔE_{ZPE}	Total	ATcT
	HEAT-456Q	649.11	149.82	−0.02	0.14	−0.65	0.06	−0.38	−44.95	752.13	752.70 ± 0.26
	HEAT-456QP	649.11	148.82	−0.02	0.14	−0.65	0.06	−0.38	−44.95	752.13	752.70 ± 0.26
CH	HEAT-345(Q)	238.28	113.56	0.43	0.11	−0.18	−0.44	−0.38	−17.00	334.39	334.66 ± 0.23
	HEAT-345Q	238.28	113.56	0.43	0.12	−0.18	−0.44	−0.38	−17.00	334.40	334.66 ± 0.23
	HEAT-345QP	238.28	113.56	0.43	0.13	−0.18	−0.44	−0.38	−17.00	334.41	334.66 ± 0.23
	HEAT-456(Q)	238.30	113.44	0.43	0.11	−0.18	−0.44	−0.38	−17.00	334.29	334.66 ± 0.23
	HEAT-456Q	238.30	113.44	0.43	0.12	−0.18	−0.44	−0.38	−17.00	334.30	334.66 ± 0.23
	HEAT-456QP	238.30	113.44	0.43	0.13	−0.18	−0.44	−0.38	−17.00	334.31	334.66 ± 0.23
CH_3	HEAT-345(Q)	1017.30	271.11	−0.15	0.24	−0.74	0.06	−0.38	−77.96	1209.48	1209.63 ± 0.13
	HEAT-345Q	1017.30	271.11	−0.15	0.24	−0.74	0.06	−0.38	−77.96	1209.48	1209.63 ± 0.13
	HEAT-345QP	1017.30	271.11	−0.15	0.25	−0.74	0.06	−0.38	−77.96	1209.50	1209.63 ± 0.13
	HEAT-456(Q)	1017.30	270.69	−0.15	0.24	−0.74	0.06	−0.38	−77.96	1209.06	1209.63 ± 0.13
	HEAT-456Q	1017.30	270.69	−0.15	0.24	−0.74	0.06	−0.38	−77.96	1209.06	1209.63 ± 0.13
	HEAT-456QP	1017.30	270.69	−0.15	0.25	−0.74	0.06	−0.38	−77.96	1209.08	1209.63 ± 0.13
CN	HEAT-345(Q)	379.38	371.64	3.14	5.08	−0.61	0.06	−0.38	−12.40	745.90	745.52 ± 0.25
	HEAT-345Q	379.38	371.64	3.14	3.50	−0.61	0.06	−0.38	−12.40	744.33	745.52 ± 0.25
	HEAT-345QP	379.38	371.64	3.14	3.98	−0.61	0.06	−0.38	−12.40	744.81	745.52 ± 0.25
	HEAT-456(Q)	379.48	371.03	3.14	5.08	−0.61	0.06	−0.38	−12.40	745.40	745.52 ± 0.25
	HEAT-456Q	379.48	371.03	3.14	3.50	−0.61	0.06	−0.38	−12.40	743.82	745.52 ± 0.25
	HEAT-456QP	379.48	371.03	3.14	3.98	−0.61	0.06	−0.38	−12.40	744.30	745.52 ± 0.25
CO_2	HEAT-345(Q)	1033.08	599.65	−4.49	5.03	−2.02	0.15	−2.02	−30.47	1598.91	1598.27 ± 0.09
	HEAT-345Q	1033.08	599.65	−4.49	4.12	−2.02	0.15	−2.02	−30.47	1598.00	1598.27 ± 0.09
	HEAT-345QP	1033.08	599.65	−4.49	4.28	−2.02	0.15	−2.02	−30.47	1598.16	1598.27 ± 0.09
	HEAT-456(Q)	1033.62	598.98	−4.49	5.03	−2.02	0.15	−2.02	−30.47	1598.77	1598.27 ± 0.09
	HEAT-456Q	1033.62	598.98	−4.49	4.12	−2.02	0.15	−2.02	−30.47	1597.86	1598.27 ± 0.09
	HEAT-456QP	1033.62	598.98	−4.49	4.28	−2.02	0.15	−2.02	−30.47	1598.02	1598.27 ± 0.09
H_2O_2	HEAT-345(Q)	562.97	562.77	−2.37	3.35	−1.60	0.36	−1.64	−68.80	1055.04	1055.23 ± 0.07
	HEAT-345Q	562.97	562.77	−2.37	2.95	−1.60	0.36	−1.64	−68.80	1054.64	1055.23 ± 0.07
	HEAT-345QP	562.97	562.77	−2.37	3.16	−1.60	0.36	−1.64	−68.80	1054.84	1055.23 ± 0.07
	HEAT-456(Q)	563.19	562.12	−2.37	3.35	−1.60	0.36	−1.64	−68.80	1054.62	1055.23 ± 0.07
	HEAT-456Q	563.19	562.12	−2.37	2.95	−1.60	0.36	−1.64	−68.80	1054.22	1055.23 ± 0.07
	HEAT-456QP	563.19	562.12	−2.37	3.16	−1.60	0.36	−1.64	−68.80	1054.42	1055.23 ± 0.07
H_2O	HEAT-345(Q)	652.40	323.02	−0.97	1.09	−1.13	0.41	−0.82	−55.79	918.21	917.83 ± 0.03
	HEAT-345Q	652.40	323.02	−0.97	0.99	−1.13	0.41	−0.82	−55.79	918.11	917.83 ± 0.03
	HEAT-345QP	652.40	323.02	−0.97	1.02	−1.13	0.41	−0.82	−55.79	918.14	917.83 ± 0.03
	HEAT-456(Q)	652.58	322.47	−0.97	1.09	−1.13	0.41	−0.82	−55.79	917.84	917.83 ± 0.03
	HEAT-456Q	652.58	322.47	−0.97	0.99	−1.13	0.41	−0.82	−55.79	917.74	917.83 ± 0.03
	HEAT-456QP	652.58	322.47	−0.97	1.02	−1.13	0.41	−0.82	−55.79	917.77	917.83 ± 0.03
HCN	HEAT-345(Q)	833.50	477.62	−3.21	3.62	−0.92	0.20	−0.38	−41.68	1268.75	1268.26 ± 0.17
	HEAT-345Q	833.50	477.62	−3.21	3.04	−0.92	0.20	−0.38	−41.68	1268.17	1268.26 ± 0.17
	HEAT-345QP	833.50	477.62	−3.21	3.46	−0.92	0.20	−0.38	−41.68	1268.59	1268.26 ± 0.17
	HEAT-456(Q)	833.63	476.83	−3.21	3.62	−0.92	0.20	−0.38	−41.68	1268.09	1268.26 ± 0.17
	HEAT-456Q	833.63	476.83	−3.21	3.04	−0.92	0.20	−0.38	−41.68	1267.51	1268.26 ± 0.17
	HEAT-456QP	833.63	476.83	−3.21	3.46	−0.92	0.20	−0.38	−41.68	1267.92	1268.26 ± 0.17
HCO	HEAT-345(Q)	764.55	403.92	−1.78	2.69	−1.16	−0.21	−1.20	−33.76	1133.05	1132.87 ± 0.19
	HEAT-345Q	764.55	403.92	−1.78	2.16	−1.16	−0.21	−1.20	−33.76	1132.53	1132.87 ± 0.19
	HEAT-345QP	764.55	403.92	−1.78	2.32	−1.16	−0.21	−1.20	−33.76	1132.68	1132.87 ± 0.19
	HEAT-456(Q)	764.79	403.50	−1.78	2.69	−1.16	−0.21	−1.20	−33.76	1132.87	1132.87 ± 0.19
	HEAT-456Q	764.79	403.50	−1.78	2.16	−1.16	−0.21	−1.20	−33.76	1132.35	1132.87 ± 0.19
	HEAT-456QP	764.79	403.50	−1.78	2.32	−1.16	−0.21	−1.20	−33.76	1132.50	1132.87 ± 0.19
HF	HEAT-345(Q)	405.69	187.35	−0.67	0.79	−0.83	0.28	−1.51	−24.70	566.39	565.97 ± 0.01
	HEAT-345Q	405.69	187.35	−0.67	0.73	−0.83	0.28	−1.51	−24.70	566.33	565.97 ± 0.01
	HEAT-345QP	405.69	187.35	−0.67	0.73	−0.83	0.28	−1.51	−24.70	566.34	565.97 ± 0.01
	HEAT-456(Q)	405.84	187.13	−0.67	0.79	−0.83	0.28	−1.51	−24.70	566.33	565.97 ± 0.01

TABLE I.   (*Continued.*)

Species method		E_{HF}^{∞}	$\Delta E_{CCSD(T)}^{\infty}$	ΔE_{CCSDT}	ΔE_{HLC}	ΔE_{rel}	ΔE_{SO}	ΔE_{DBOC}	ΔE_{ZPE}	Total	ATcT
	HEAT-456Q	405.84	187.13	−0.67	0.73	−0.83	0.28	−1.51	−24.70	566.26	565.97 ± 0.01
	HEAT-456QP	405.84	187.13	−0.67	0.73	−0.83	0.28	−1.51	−24.70	566.27	565.97 ± 0.01
HNO	HEAT-345(Q)	331.69	528.35	−2.45	4.10	−1.15	−0.40	−0.82	−36.00	823.32	823.65 ± 0.12
	HEAT-345Q	331.69	528.35	−2.45	3.54	−1.15	−0.40	−0.82	−36.00	822.76	823.65 ± 0.12
	HEAT-345QP	331.69	528.35	−2.45	3.89	−1.15	−0.40	−0.82	−36.00	823.10	823.65 ± 0.12
	HEAT-456(Q)	332.02	527.95	−2.45	4.10	−1.15	−0.40	−0.82	−36.00	823.25	823.65 ± 0.12
	HEAT-456Q	332.02	527.95	−2.45	3.54	−1.15	−0.40	−0.82	−36.00	822.69	823.65 ± 0.12
	HEAT-456QP	332.02	527.95	−2.45	3.89	−1.15	−0.40	−0.82	−36.00	823.04	823.65 ± 0.12
HO$_2$	HEAT-345(Q)	300.93	431.57	−0.60	2.97	−1.17	−0.04	−1.64	−36.92	695.10	694.46 ± 0.22
	HEAT-345Q	300.93	431.57	−0.60	2.50	−1.17	−0.04	−1.64	−36.92	694.64	694.46 ± 0.22
	HEAT-345QP	300.93	431.57	−0.60	2.71	−1.17	−0.04	−1.64	−36.92	694.85	694.46 ± 0.22
	HEAT-456(Q)	301.18	431.21	−0.60	2.97	−1.17	−0.04	−1.64	−36.92	694.99	694.46 ± 0.22
	HEAT-456Q	301.18	431.21	−0.60	2.50	−1.17	−0.04	−1.64	−36.92	694.53	694.46 ± 0.22
	HEAT-456QP	301.18	431.21	−0.60	2.71	−1.17	−0.04	−1.64	−36.92	694.73	694.46 ± 0.22
NH$_2$	HEAT-345(Q)	492.93	270.76	0.06	0.48	−0.65	−0.14	−0.00	−49.61	713.83	713.32 ± 0.19
	HEAT-345Q	492.93	270.76	0.06	0.48	−0.65	−0.14	−0.00	−49.61	713.82	713.32 ± 0.19
	HEAT-345QP	492.93	270.76	0.06	0.50	−0.65	−0.14	−0.00	−49.61	713.85	713.32 ± 0.19
	HEAT-456(Q)	493.02	270.31	0.06	0.48	−0.65	−0.14	−0.00	−49.61	713.48	713.32 ± 0.19
	HEAT-456Q	493.02	270.31	0.06	0.48	−0.65	−0.14	−0.00	−49.61	713.47	713.32 ± 0.19
	HEAT-456QP	493.02	270.31	0.06	0.50	−0.65	−0.14	−0.00	−49.61	713.49	713.32 ± 0.19
NH$_3$	HEAT-345(Q)	841.38	406.03	−0.63	0.79	−1.07	0.42	−0.00	−89.44	1157.48	1157.25 ± 0.04
	HEAT-345Q	841.38	406.03	−0.63	0.72	−1.07	0.42	−0.00	−89.44	1157.41	1157.25 ± 0.04
	HEAT-345QP	841.38	406.03	−0.63	0.77	−1.07	0.42	−0.00	−89.44	1157.46	1157.25 ± 0.04
	HEAT-456(Q)	841.48	405.36	−0.63	0.79	−1.07	0.42	−0.00	−89.44	1156.90	1157.25 ± 0.04
	HEAT-456Q	841.48	405.36	−0.63	0.72	−1.07	0.42	−0.00	−89.44	1156.84	1157.25 ± 0.04
	HEAT-456QP	841.48	405.36	−0.63	0.77	−1.07	0.42	−0.00	−89.44	1156.89	1157.25 ± 0.04
NH	HEAT-345(Q)	214.88	132.40	0.29	0.20	−0.30	−0.24	−0.00	−19.49	327.74	328.43 ± 0.29
	HEAT-345Q	214.88	132.40	0.29	0.22	−0.30	−0.24	−0.00	−19.49	327.75	328.43 ± 0.29
	HEAT-345QP	214.88	132.40	0.29	0.22	−0.30	−0.24	−0.00	−19.49	327.76	328.43 ± 0.29
	HEAT-456(Q)	214.94	132.20	0.29	0.20	−0.30	−0.24	−0.00	−19.49	327.59	328.43 ± 0.29
	HEAT-456Q	214.94	132.20	0.29	0.22	−0.30	−0.24	−0.00	−19.49	327.61	328.43 ± 0.29
	HEAT-456QP	214.94	132.20	0.29	0.22	−0.30	−0.24	−0.00	−19.49	327.61	328.43 ± 0.29
NO	HEAT-345(Q)	225.52	411.50	−1.54	3.59	−0.79	0.02	−0.10	−11.46	626.74	626.85 ± 0.08
	HEAT-345Q	225.52	411.50	−1.54	2.98	−0.79	0.02	−0.10	−11.46	626.13	626.85 ± 0.08
	HEAT-345QP	225.52	411.50	−1.54	3.32	−0.79	0.02	−0.10	−11.46	626.47	626.85 ± 0.08
	HEAT-456(Q)	225.80	411.23	−1.54	3.59	−0.79	0.02	−0.10	−11.46	626.76	626.85 ± 0.08
	HEAT-456Q	225.80	411.23	−1.54	2.98	−0.79	0.02	−0.10	−11.46	626.14	626.85 ± 0.08
	HEAT-456QP	225.80	411.23	−1.54	3.32	−0.79	0.02	−0.10	−11.46	626.48	626.85 ± 0.08
OF	HEAT-345(Q)	−64.04	282.44	1.05	2.69	−0.38	0.00	−1.24	−6.37	214.16	213.47 ± 0.47
	HEAT-345Q	−64.04	282.44	1.05	2.08	−0.38	0.00	−1.24	−6.37	213.55	213.47 ± 0.47
	HEAT-345QP	−64.04	282.44	1.05	2.22	−0.38	0.00	−1.24	−6.37	213.69	213.47 ± 0.47
	HEAT-456(Q)	−64.00	282.40	1.05	2.69	−0.38	0.00	−1.24	−6.37	214.16	213.47 ± 0.47
	HEAT-456Q	−64.00	282.40	1.05	2.08	−0.38	0.00	−1.24	−6.37	213.55	213.47 ± 0.47
	HEAT-456QP	−64.00	282.40	1.05	2.22	−0.38	0.00	−1.24	−6.37	213.69	213.47 ± 0.47
OH	HEAT-345(Q)	286.41	162.12	−0.13	0.47	−0.53	−0.01	−0.04	−22.25	426.05	425.62 ± 0.03
	HEAT-345Q	286.41	162.12	−0.13	0.47	−0.53	−0.01	−0.04	−22.25	426.05	425.62 ± 0.03
	HEAT-345QP	286.41	162.12	−0.13	0.48	−0.53	−0.01	−0.04	−22.25	426.06	425.62 ± 0.03
	HEAT-456(Q)	286.51	161.88	−0.13	0.47	−0.53	−0.01	−0.04	−22.25	425.91	425.62 ± 0.03
	HEAT-456Q	286.51	161.88	−0.13	0.47	−0.53	−0.01	−0.04	−22.25	425.91	425.62 ± 0.03
	HEAT-456QP	286.51	161.88	−0.13	0.48	−0.53	−0.01	−0.04	−22.25	425.92	425.62 ± 0.03

III. RESULTS

A. Effects due to use of improved basis sets

Total atomization energies (TAEs) are listed for 26 molecules in Table I. Results are listed for the HEAT-345- and HEAT-456-based methods using the CCSDT(Q),[23,24] CCSDTQ,[29] and CCSDTQP (Ref. [30]) treatments of the ΔE_{HLC} contribution. All of the latter were obtained with the cc-pVDZ basis set in the frozen-core approximation. The trends within a given basis-set extrapolation procedure have been discussed elsewhere;[9] it is sufficient to focus here on the effects that arise when the basis sets used for the extrapolation are expanded from the 345 to the 456 protocols. It is evident that the HF-SCF contribution to the TAE stays largely the same with a tendency toward a very small increase when the larger extrapolation sequence is used. This is due primarily to the fact that the aug-cc-pCVTZ basis set is quite far from convergence; indeed the straightforward and unextrapolated aug-cc-pCV5Z basis set (that is used in the 345 extrapolation sequence) is *by itself* better than the 345 extrapolation. For example, the 345- and 456-based values of E_{HF}^{∞} for CO are 729.86 and 730.13 kJ mol^{-1}, respectively, while the results with the individual basis sets are 727.39, 730.20, 730.14, and 730.13 (X=T, Q, 5, and 6, respectively). The convergence beyond aug-cc-pCVTZ is rapid and monotonic, but it is clear that the 345-based extrapolation is spoiled by the quite out-of-line aug-cc-pCVTZ basis. Similar behavior is observed for all other molecules where the 345- and 456-based E_{HF}^{∞} contributions are significantly different. Clearly, and somewhat amusingly, it is better to use HF-SCF energies obtained in a quadruple- or quintuple-zeta basis set than extrapolated HF-SCF energies that include triple-zeta or smaller basis-set contributions! In fact, it seems that there is little point in doing extrapolations of the HF-SCF energy at all.

The correlation contribution to the atomization energies exhibits a more dramatic dependence on the extrapolation procedure, as is expected since electron correlation is more profoundly influenced by basis sets than are electronic energies in the HF approximation. In *all* cases, the 56-based CCSD(T) contributions to the atomization energy are smaller than those based on 34 extrapolations. In other words, expansion of the basis tends to confer greater stability to atoms than it does to molecules. From yet another point of view, since electron correlation effects always increase bond energies, the 45-based treatment of CCSD(T) tends to overshoot this effect. It is somewhat remarkable that the magnitude of the differences (often ca. 0.5 kJ mol^{-1}) is, in fact, larger than the statistical error of the HEAT 345-based approaches. This point will be addressed again in Sec. IV of this paper.

Let us again take CO as an example. Here the values of $\Delta E_{CCSD(T)}^{\infty}$ from the 45- and 56-based extrapolations are 356.91 and 356.53 kJ mol^{-1}, respectively. To underscore the fact that correlation converges much more slowly with basis set than does the HF-SCF energy, the individual values of the TAE from the aug-cc-pCVXZ basis sets are (in kJ mol^{-1}) the following:

> aug-cc-pCVTZ 331.35,
> aug-cc-pCVQZ 346.12,
> aug-cc-pCV5Z 351.59,
> aug-cc-pCV6Z 353.55.

It is clear that extrapolation is absolutely essential for correlation energies obtained with traditional methods (the $R12$ and $F12$ treatments[31] being a possible exception) to obtain anything resembling 1 kJ mol^{-1} accuracy. It is difficult to assess the overall accuracy, but it is at least comforting that differences between TAE contributions calculated from 45 and 56 extrapolations for CCSD(T) never exceed 1 kJ mol^{-1} for the molecules in our dataset.

As mentioned in Part I of this series, the partial-wave expansion formula used to extrapolate the CCSD(T) energies is valid only for two-electron atoms in the limit of a basis saturated with functions of each angular momentum. This theoretical foundation becomes quite precarious when one uses the correlation-consistent series of basis sets (which are clearly not saturated with respect to functions of the highest angular momentum, which is represented by a single Gaussian function!), and the success of extrapolations must really be viewed as a triumph of a justifiable, but certainly nonrigorous approach. Quite recently, Schwenke[32] has embraced empiricism and advocated a two-point extrapolation of the form

$$E^{\infty} = E^{N-1} + \mathcal{F}_{\mathcal{N}}[E^N - E^{N-1}], \qquad (3)$$

where N is an energy obtained with a cc-pVXZ (X=N) basis set, and the constants $\mathcal{F}_N$ depend on N. While not mentioned in his paper, this approach is precisely equivalent to an inverse-power formula, except that the exponent is not 3, nor is it the same for all pairs of basis sets. Precisely, it is quite easy to show that the formula above can be written as

$$E^{\infty} = E^N - \frac{a}{N^{-X}}, \qquad (4)$$

where

$$X = \frac{\ln\left(1 - \frac{1}{\mathcal{F}_n}\right)}{\ln\left(\frac{N-1}{N}\right)}. \qquad (5)$$

For example, Schwenke's recommendation for $\mathcal{F}_6$ within the aug-cc-pVXZ hierarchy is 2.265 620 6 (sic) for extrapolating the CCSD correlation energy obtained with the aug-cc-pVXZ basis, which is equivalent to using Eq. (4) with X=3.1937, while the formula of Helgaker *et al.*[13] is the same as that obtained with $\mathcal{F}_n$=2.3736. However, Schwenke makes the additional recommendations that the CCSD and (T) correlation energies should be extrapolated with different coefficients $\mathcal{F}$. We have used Schwenke's coefficients (or equivalently exponents) within the HEAT scheme to see what effect this approach has on the predicted thermochemical properties. It should be mentioned here that Schwenke was careful

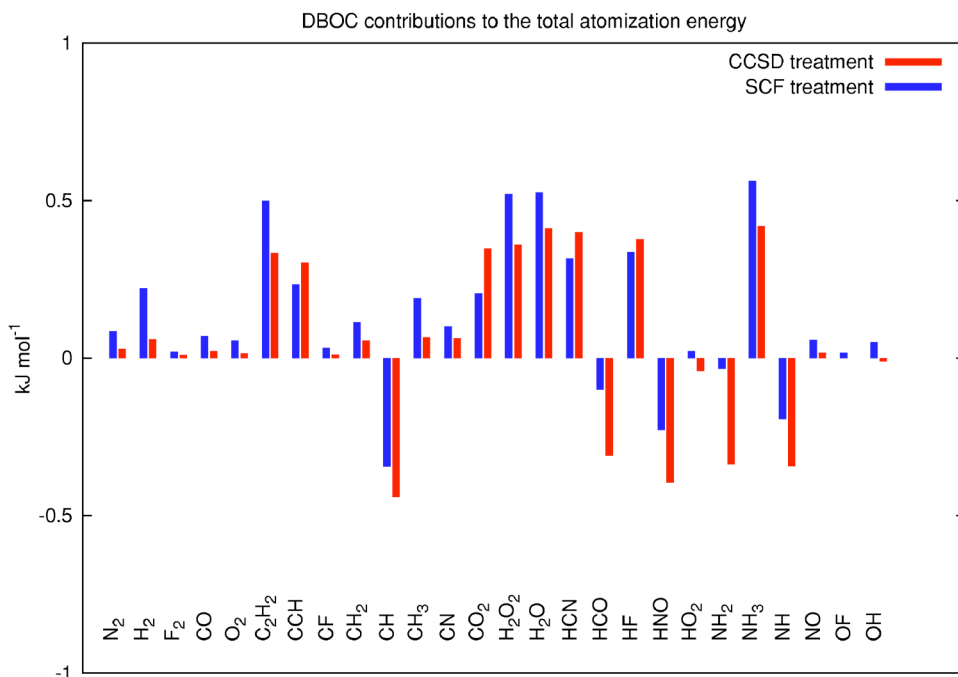DBOC contributions to the total atomization energy

FIG. 1. (Color online) Contributions to the atomization energy based on correlated and HF-SCF treatments of the diagonal Born-Oppenheimer calculation.

to mention that he had optimized these parameters only for valence correlation energies within the aug-cc-pVXZ hierarchy, and we are glibly using the same constants to extrapolate energies obtained within the aug-cc-pCVXZ series. However, the approach is empirical, and the fact is that using exponents greater than 3 will tend to reduce the extrapolated correlation energies. We view this as favorable, because it is our belief that the CCSD(T) correlation energies used in HEAT-456 methods tend to be overestimated.

In any event, for the CO molecule, the Schwenke-based extrapolation of the aug-cc-pCVXZ CCSD and (T) energies (done separately, per his recommendation) gives CCSD(T) contributions of 356.23 and 356.29 kJ mol^{-1} to the TAE for the 45 and 56 procedures, respectively. Similar behavior (smaller differences between 45- and 56-based Schwenke extrapolations and an overall reduction in the TAE contribution) is seen almost uniformly [for example, the $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ TAEs of N$_2$ with the 45- and 56-based extrapolations are 472.26 and 471.49 kJ mol^{-1} with the partial-wave ($X=3$) formula and 471.49 and 471.29 kJ mol^{-1} with the Schwenke parameters]. This overall pattern, and the demonstrated accuracy of the Schwenke parametrization of extrapolation, suggests that the HEAT $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ contributions to the TAE are overestimated still at the HEAT-456 level of theory. This appears to be a consistent deficiency in the protocol; it would seem evident at this point to conclude that the overall CCSD(T) contributions to atomization energies calculated with all HEAT protocols to date are overestimated.

B. Improved treatment of the DBOC

The diagonal Born-Oppenheimer contribution is the simplest way to correct for the usual use of the Born-Oppenheimer approximation in which the nuclear motion dependence of the electronic wavefunction is ignored.[33] However, this first-order treatment of nuclear motion does not spoil the separability of nuclear and electronic wavefunc-

tions and represents a mass-dependent adiabatic approximation. It is a small correction for all of the molecules studied here; the maximum contribution to the TAEs is less than 0.5 kJ mol^{-1}, and the correction is entirely negligible for molecules that do not contain hydrogen atoms. However, as mentioned in Ref. 10, the maximum error in HEAT-456 thermochemistry for H$_2$ actually comes from correlation contributions to the DBOC, which have been neglected up to now. The reason for this is that, while treatments of the DBOC within the HF-SCF approximation have been around for some time (see, for example, Ref. 34), correlated calculations of this correction within a size-consistent methodology are much less prevalent. In fact, the theory—which proved surprisingly subtle—for the coupled-cluster approximation has been worked out and implemented only very recently. In this work, we have used DBOC contributions calculated with the CCSD treatment of electron correlation,[15] using the aug-cc-pCVQZ basis set. Figure 1 shows a schematic representation of how the DBOC contribution to the TAEs is affected by electron correlation effects. For molecules without hydrogen, this contribution is entirely neglibile and the correlation effects are not even worthy of discussion. For hydrides, the trend is relatively systematic; inclusion of electron correlation in the DBOC leads to an overall reduction in atomization energies (molecular destabilization) in most cases, including all examples in which the correlation effect is greater than 0.1 kJ mol^{-1}. Hence, improved treatments of both the CCSD(T) and DBOC contributions to the energy generally lead to overall reduction in the TAE.

C. Comparison with ATcT

Total atomization energies for the molecules that have historically constituted the statistical dataset for HEAT are compared with recent values from ATcT (Ref. 35) in Table II. It is clear from the error analysis that "the best" HEAT ap-

TABLE II. Errors (in kJ mol^{-1}) in atomization energies, compared with ATcT values based on the Core (Argonne) Thermochemical Network, C(A)TN, Version 1.064. Molecules here are those that come from the original HEAT dataset.

Species	HEAT345-Q	HEAT456-Q	HEAT345-(Q)	HEAT456-(Q)	HEAT345-QP	HEAT456-QP
N_2	−0.05	−0.69	0.60	−0.04	0.42	−0.22
H_2	0.15	0.04	0.15	0.04	0.15	0.04
F_2	−0.04	−0.19	0.40	0.25	0.12	−0.03
CO	−0.33	−0.44	0.09	−0.02	−0.19	−0.31
O_2	−0.69	−0.39	−0.17	0.13	−0.26	0.03
C_2H_2	−0.44	−1.20	−0.12	−0.87	−0.12	−0.87
CCH	−0.11	−0.84	0.05	−0.67	0.20	−0.52
CH_2	−0.26	−0.57	−0.27	−0.58	−0.25	−0.57
CH	−0.26	−0.36	−0.27	−0.37	−0.25	−0.35
CH_3	−0.15	−0.57	−0.15	−0.57	−0.13	−0.55
CO_2	−0.27	−0.41	0.63	0.49	−0.11	−0.25
H_2O_2	−0.59	−1.01	−0.19	−0.61	−0.39	−0.81
H_2O	0.28	−0.10	0.38	0.01	0.31	−0.06
HCO	−0.34	−0.52	0.18	0.00	−0.19	−0.37
HF	0.36	0.30	0.43	0.36	0.37	0.31
HO_2	0.18	0.07	0.64	0.53	0.39	0.27
NO	−0.73	−0.71	−0.11	−0.10	−0.39	−0.37
OH	0.43	0.29	0.43	0.29	0.44	0.30
rms error	0.37	0.57	0.35	0.42	0.28	0.42
Mean error	−0.16	−0.41	0.15	−0.10	0.01	−0.24
Max error	0.73	1.20	0.64	0.87	0.44	0.87
95% confidence limits[a]	0.74	1.14	0.70	0.84	0.56	0.84

[a]Estimated 95% confidence limits have been approximated by twice the rms value.

proaches do not necessarily have the best statistical properties. For example, the largest error for all methods using a particular treatment for the ΔE_{HLC} is *larger* when the 456-based extrapolations are used than the simpler HEAT-345 variants. It is not just the maximum errors that are affected; the rms errors increase as well! However, all rms errors are less than 0.5 kJ mol^{-1}, and a detailed analysis of them is anything but straightforward. Let us demonstrate now that such an analysis is largely a pointless exercise and, as a corollary, that efforts to distinguish between the quality of the various HEAT approaches are simply not possible in general on the basis of such a *statistical comparison*. To this end, we choose a few selected molecules for a detailed investigation of systematic deficiencies that exist even at the HEAT-456QP level, which is—unquestionably—the most theoretically complete approach. It is self-evident that none of the contributions in the HEAT protocol associated with the non-relativistic and scalar relativistic energies is experimentally determinable,[36] and a certain degree of guessing will always be involved as to the accuracy of these terms. One can only base the level of confidence in a particular correlation energy on the convergence of the energy contributions; that is, molecules with a low HLC probably have more accurate correlation energies than those with larger ΔE_{HLC} values. A similar faith must be associated with the DBOC contribution but fortunately this is usually of a magnitude such that a 10%–20% error in the calculation (a conservative estimate, in our view) would amount to no more than 5 cm^{-1} (ca. 0.05 kJ mol^{-1}) in the TAE. However, the remaining contributions, the spin-orbit stabilization energy, and the zero-point vibrational energy can be determined in the laboratory, subject to certain assumptions. For a few of the simpler systems

in the test suite, the corresponding data are available. Let us first investigate the simple diatomic hydrides which have extremely tight ATcT error bars (HF, OH, and H_2O) of ≤ 0.1 kJ mol^{-1}. For all of these systems, the error in the HEAT-based SO and ZPE corrections can be quantified, and improved TAEs can be calculated using the remaining HEAT-456QP contributions. All of these molecules are addressed now, beginning with those that have precisely known TAEs.

1. HF

Hydrogen fluoride represents a somewhat puzzling case. The straightforward HEAT-456QP atomization energy is 0.31 kJ mol^{-1} above the current ATcT value. Use of the exact zero-point correction of 2050.8 cm^{-1} (Ref. 37) (which is 14.3 cm^{-1} below the HEAT value) acts to further increase this discrepancy—the molecule is now predicted to be overbound by 0.48 kJ mol^{-1}. As the HLC for HF is less than 1 kJ mol^{-1}, experience suggests that basis-set effects would have only a small effect and actually make the agreement worse. However, the HEAT-based spin-orbit stabilization of the fluorine atom undershoots the experimental value of 134.7 cm^{-1} (Ref. 38) by 8.7 cm^{-1}, and application of the appropriate adjustment brings the theoretical value closer; exact ZPE and spin-orbit contributions lead to a TAE that is 0.37 kJ mol^{-1} above the ATcT value. The recent work of Karton *et al.*,[39] who have admirably explored subtleties and sources of error in the high-level correlation correction ΔE_{HLC}, offers the answer: they find that the HLC for HF is predicted surprisingly poorly with a small basis set and in fact *decreases* with basis-set size. In general, the HLC deter-

mined in a small basis set will undershoot [the same is true for *individual contributions* to the CCSD(T) contributions; it is just the extrapolations that tend to overshoot the correct result] the correct result, meaning that the contribution to the TAE obtained with HEAT-based values of ΔE_{HLC} will be slightly too small. However, for HF, Karton *et al.*[8] find that the cc-pVDZ basis set *overshoots* the cc-pV5Z result for the CCSDT(Q) contribution to the TAE by roughly 0.3 kJ mol^{-1}. This accounts, nearly quantitatively, for the remaining error in the TAE for hydrogen fluoride.

2. OH radical

The HEAT-456QP atomization energy is calculated to be 425.92 kJ mol^{-1}, which is 0.30 kJ mol^{-1} above the quite precise ATcT value of 425.62 ± 0.03 kJ mol^{-1}. One expects the electron correlation effects to be quite well converged at this level of theory, as the HLC contribution is a mere 0.48 kJ mol^{-1}. Moreover, the work of Karton *et al.*[8] indicates that the basis-set error in the HLC is less than 0.1 kJ mol^{-1} (and again too large, which seems to be part of a pattern in which HLC corrections to the TAE obtained in the cc-pVDZ basis set are too large for molecules with very polar HX bonds), meaning that the error is still on the order of 0.2 kJ mol^{-1}. Given the relative simplicity of the electronic structure of OH and the small size of the radical, such a "large" error (at least relative to some of the other molecules in the test suite) is perhaps surprising. Why, one might wonder, would OH have a larger error at the HEAT456-QP level than do N$_2$ and CO$_2$, which are considerably more complicated molecules for theory? However, let us now see what discrepancy remains when the ZPE and SO contributions are replaced by their experimental values. How much will this change the HEAT-456QP result? In which direction?

It turns out that largest in magnitude amongst these is the error due to the HEAT-based SO contribution, which corresponds to the center of gravity between the nonrelativistic energy (which corresponds to that associated with the eigenvalue of the Born-Oppenheimer electronic Hamiltonian) and the true ground-state *electronic* energy, a quantity that we refer to here as the spin-orbit stabilization energy (SOSE). The situation that is operative in OH is depicted schematically in Fig. 2. The HEAT-based SOSE for the oxygen atom is 68.5 cm^{-1}, while the value based on the precise positions of the three ^3P sublevels is 78.0 cm^{-1} (Ref. 38), a difference of 0.11 kJ mol^{-1}. For OH, the spin-orbit constant A_0 describing the splitting of $^2\Pi_{3/2}$ and $^2\Pi_{1/2}$ levels is 139.2 cm^{-1}, placing the weighted average at +69.6 cm^{-1} relative to the ground state under the most simplistic assumption. While this is quite close to the HEAT calculation of +65.2 cm^{-1}, OH has a fairly substantional *rotational* zero-point energy due to electronic-rotational coupling, and the simple treatment $E_{ground} = E_{nonrelativistic} - \frac{1}{2}A_0$ does a rather poor prediction of estimating the position of the lowest allowed rotational state. Application of the Hill–Van Vleck equation[40] shows that, in fact, the lowest existing rotational level is 38.2 cm^{-1} below the center of gravity, rather than the 69.6 cm^{-1} that is associated with the simple formula. Thus, the exact SOSE for OH is -38.2 cm^{-1} (-0.46 kJ mol^{-1}), which is 0.32 kJ mol^{-1} smaller in magnitude than the HEAT result of
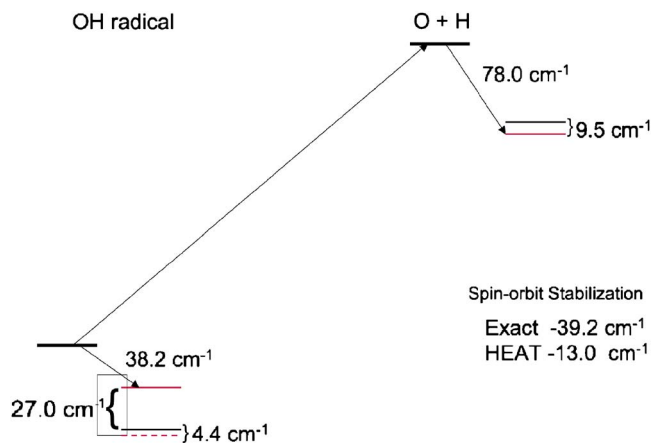

FIG. 2. (Color online) Energy levels of the OH atomization process. The thick levels represent the center-of-gravity nonrelativistic energies obtained in the *ab initio* calculations. The red levels (dotted for OH) represent the simplistic center of gravity (not taking into account the detailed coupling of rotational, spin, and electronic angular momenta, such as that occurring in a $^2\Pi$ state of a diatomic molecule) of the individual spin-orbit sublevels, and the black thin lines are the corresponding HEAT estimates. The solid red line for OH is the position of the lowest allowed rotational term.

-65.2 cm^{-1}. Hence, the error in the HEAT-based SOSE for the atomization energy of OH is $0.32+0.11=0.43$ kJ mol^{-1} (see Fig. 2), which decreases the theoretical TAE to 425.48 kJ mol^{-1}, which is just 0.14 kJ mol^{-1} below the ATcT value. It is particularly noteworthy that the *error* in the calculated quantity is, in fact, about 30% larger in magnitude than the discrepancy between HEAT456-QP and ATcT atomization energies. It is perhaps amusing that the quantity under consideration is one that is ignored in most theoretical model chemistries. Nonetheless, with a residual error of 0.13 kJ mol^{-1} after correcting the SOSE, things are getting better for the ostensibly simple OH radical in the brave new world of highly precise theoretical thermochemistry.

While one might (rationally) think it naive to attempt a further analysis, let us indulge this interest and focus on the other potentially problematic contribution that can be (within the Born-Oppenheimer approximation) determined experimentally: the ZPE contribution. For OH, the experimental value of the ZPE is 1850.69 cm^{-1} (22.139 kJ mol^{-1}),[37] which is 0.11 kJ mol^{-1} below the HEAT value. Thus, the theoretical atomization energy is increased by this amount, which results in a value of 425.61 kJ mol^{-1}, which is in fortuitously perfect agreement with ATcT and is the most accurate calculation of this quantity to date.[41] The accuracy issue alluded to above—why OH had a larger discrepancy than the difficult molecules N$_2$ and CO$_2$—is seen to be a consequence mostly of the complicated angular momentum coupling in the OH radical, with a smaller contribution from an overestimation of the ZPE by theory, rather than deficiencies in the electronic energy contribution. OH is apparently treated very well indeed by the nonrelativistic energy calculations associated with HEAT456-QP.

3. H$_2$O

The water molecule is another interesting case, where the complexities associated with treating the spin-orbit inter-

action in hydroxyl are not present, water being a closed-shell singlet molecule. That the electronic structure aspect of the calculation is well converged is evidenced by the total non-relativistic energy (see Table IV) of -76.4388 a.u., which is in perfect agreement with a recent estimate of Bytautas and Ruedenberg[42] (-76.4390 ± 0.0004 a.u.), suggesting that the HEAT-456QP energy is within 20 cm^{-1} of the exact nonrelativistic result. Indeed, the HEAT456-QP atomization energy of 917.76 kJ mol^{-1} is but 0.07 kJ mol^{-1} below the ATcT estimate of 917.84 ± 0.03 kJ mol^{-1}. However, this agreement is somewhat fortuitous and degrades when proper account is taken for the "wild-card" spin-orbit and ZPE contributions. For the former, the error of 0.11 kJ mol^{-1} in the HEAT value for the oxygen atom spin-orbit stabilization energy[38] must be added to the HEAT-456QP result, raising it to 917.88 kJ mol^{-1}, this value being in better agreement with ATcT. The situation is further improved when one includes the ZPE error; the semiexperimental value of 55.76 kJ mol^{-1} (Ref. [43]) is 0.03 kJ mol^{-1} below the ATcT value, lowering the theoretical value to 917.85 kJ mol^{-1}. However, again there is something of a Pauling point even at this level; the calculations of Karton *et al.*[8] suggest that the ΔE_{HLC} contribution to the TAE H$_2$O—like HF and OH and, it would seem, most molecules with highly polar HX bonds—is overestimated with the cc-pVDZ basis set by about 0.12 kJ mol^{-1}. This, together with basis-set effects of the $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ and $\Delta E_{\mathrm{CCSDT}}$ corrections, clearly shows that one must be very careful in analyzing theoretical thermochemical results at such advanced levels of theory. Perfect, or nearly perfect, agreement is always fortuitous and it is totally unclear whether improving a particular contribution is really beneficial. We return to this issue at the conclusion of this paper.

IV. DISCUSSION

The previous section makes clear that the overall error in HEAT-based schemes for thermochemistry is so small that even more sophisticated treatments of individual contributions and/or consideration of effects such as the angular momentum coupling in the hydroxyl radical change results in subtle ways that do not necessarily improve results. Indeed, it seems clear that easily the most practical and useful general HEAT method is probably HEAT-345(Q), which is also the least expensive approach. At this point, it is appropriate to go through each contribution to the HEAT energy and discuss some general trends. First, E_{HF}^{∞} is apparently quite well converged; differences between TAE contributions calculated from aug-cc-pCV5Z, aug-cc-pCV6Z, and HEAT-456 HF-SCF energies are so small as to suggest that E_{HF}^{∞} from HEAT-456 is probably in error by less than 0.05 kJ mol^{-1} in all cases. The same is not true for the HEAT-345 E_{HF}^{∞} results, simply because of the poor performance of the triple-zeta basis discussed earlier in this manuscript.

$\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ is of course one of the largest sources of error and the principal term of importance for the correlation energy. Here, the error is apparently systematic. While CCSD(T) energies calculated in all of the basis sets used here lead to TAE contributions that are too small (correlation

effects always give a positive contribution to the TAE), the partial-wave extrapolations tend to overshoot the results. The finite basis-set calculations converge (slowly) from below; the extrapolations appear to converge from above. Interestingly, the extrapolations based on the Schwenke formulas appear to converge from below, and—based on a few examples studied here—more rapidly than do the partial-wave extrapolations. However, the latter lead to overestimated TAEs, on the order of a few tenths of a kJ mol^{-1}. The $\Delta E_{\mathrm{CCSDT}}$ and ΔE_{HLC} contributions have been analyzed in detail by Karton *et al.*,[39] and it seems that the strategy for $\Delta E_{\mathrm{CCSDT}}$ that is used in the HEAT approach (equivalent to the second column in Table II of Ref. [8])—which is also based on extrapolation—also gives a slightly too positive contribution to the TAE as well. However, the magnitude of the difference is smaller than that of $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ and typically on the order of 0.1 kJ mol^{-1}.

The overestimation of the correlation contribution to the TAE through the CCSDT level is, in HEAT-based thermochemistry, offset to a large degree by a compensating error in the ΔE_{HLC} term. Boese *et al.* were the first to demonstrate the basis-set effects in CCSDTQ and related methods vis-a-vis the TAE. Boese *et al.* have advocated various formulas, ranging from a simple scaling of (frozen-core) CCSDTQ energies obtained in a cc-pVDZ basis set by a factor of 1.2532 (Ref. [7]) to considerably more complicated approaches (the so-called "W4.3 combo" and "W4 combo"). However, here one is faced with two fundamental options. First, one can do calculations with larger and larger basis sets, and thereby approach the *exact* $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ and $\Delta E_{\mathrm{CCSDT}}$ contributions, albeit with considerable computational demands, and then attempt to also account quantitatively for the ΔE_{HLC} using large basis sets and/or the empiricism of the "combos." There is no doubt that this is the only way to establish definitive results and to make predictions of fundamental thermochemical parameters when extraordinarily high precision is required. However, it is very costly, and it seems that such an approach will always encounter severe restrictions with regard to its scope of application. The other approach, which is clearly more practical, is to use the empirical cancellation between the overshooting of the $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ and $\Delta E_{\mathrm{CCSDT}}$ contributions and the undershooting of the ΔE_{HLC} contribution. Each has its merits, and it is always satisfying to understand why a method performs "better than it ought to," as does HEAT345-(Q).

Most of the remaining contributions all have rather smaller error bars. First, the ΔE_{rel} contribution is based on first-order perturbation theory, and contains both one- and two-electron Darwin and mass-velocity contributions. It is believed that this is quite satisfactory for first-row molecules. We have done calculations with two different basis sets (aug-cc-pCVTZ, which is used in all HEAT calculations and a larger aug-cc-pCVQZ basis), and it is found that the larger basis set tends to affect TAE contributions by at most 0.07 kJ mol^{-1} (for, not surprisingly, F$_2$). The spin-orbit contribution ΔE_{SO} is a somewhat larger source of error, as the calculations for the atoms C, O, and F have errors (with respect to the exact spectroscopic stabilizations) of -0.02,

TABLE III. Guide to typical sizes and signs of errors in the contributions to HEAT-based atomization energies for molecules with two to four first-row atoms.

Contribution to TAE	Sign of typical error in HEAT-based theory	Greater than 0.2 kJ mol⁻¹?
E_{HF}^{∞}	Essentially exact	No
$\Delta E_{CCSD(T)}^{\infty}$	+	Can be[a]
ΔE_{CCSDT}	+	No
ΔE_{HLC}	−	Can be
	+	Can be
ΔE_{rel}	Varies	No
ΔE_{SO}	+ for hydrocarbons	No
	− otherwise	No
ΔE_{DBOC}	Unknown	No
ΔE_{ZPE}	−	Can be

[a]Especially for tetra-atomics.

0.11, and 0.10, with the HEAT-based stabilization energies larger in magnitude for O and F. Hence, this error—which can easily be quantified—leads to an error of the associated TAE contribution for closed-shell molecules by a magnitude that is easily calculated from stoichiometric considerations. The DBOC contribution changes little when one goes from the HF-SCF-based treatment of this correction that was used in the first two papers in this series; it is expected to be accurate to well within 0.1 kJ mol⁻¹.

Finally, the zero-point vibrational correction can be determined essentially exactly from spectroscopic data on diatomics and is also occasionally well known for molecules such as water. Results reported earlier in this paper, together with the empirical knowledge that quantum chemical calculations of vibrational energy levels tend to approach the correct answer from above, suggest that the calculated ΔE_{ZPE} contribution to the atomization energies is too large in magnitude; a more accurate treatment of this contribution would tend to increase atomization energies by, again, roughly 0.1 kJ mol⁻¹ for most of the molecules studied here.

A summary of the effectively systematic errors encountered in the HEAT scheme is shown in Table III, where the tendency toward error cancellation is evident, as is the cause of the seemingly anomalous behavior of HF. It should be emphasized that the error bars for HEAT—roughly 0.5 kJ mol⁻¹ rms (or ∼1 kJ mol⁻¹ in terms of 95% confidence limits)—apply *only* to molecules containing a very few first-row atoms. Calculating atomization energies is very difficult, and the magnitude of the error ultimately will become proportional to the size of the system. For example, it is entirely unrealistic to expect a HEAT-based calculation of the total atomization energy of benzene (which we are currently undertaking) to have this sort of accuracy. It is a necessary consequence of size extensivity that the characteristic *relative* error in atomization energies will remain constant, but the absolute error is clearly going to grow linearly with the size of the system. Hence, it would be rather more realistic to assume that the TAE of benzene computed with HEAT would have an error larger by a factor of 10 or so. This is a point that we feel is made explicit only rarely—if at all—in the high-accuracy theoretical thermochemistry litera-

ture and an important qualification—"1 kJ mol⁻¹ accuracy" applies only within the small molecule (two or three heavy atoms, at most) approximation.[44] Hence, the achievement of chemical accuracy is really a relative goal; if one is interested in TAEs, then chemical accuracy for a molecule with, say, a dozen or so atoms is probably still not an achievable goal. However, it would be better to calculate the enthalpy of formation of benzene via some sort of isodesmic or nearly isodesmic scheme, because that process would tend to lead toward error cancellation. In this regard, we remind the reader that HEAT was never intended to be a method that was entirely for the purpose of calculating TAEs, but rather for estimating the *total* energies of molecules. Thermochemistry is best calculated via reaction-based schemes (as beautifully epitomized by the work of Allen and co-workers using the focal-point treatment[5,45]); TAEs are simply a useful quantity to calculate in benchmarking the performance of a method.

To this end, we give a table of Hartree-Fock limit energies and HEAT-456QP correlation energies that we propose to be the most accurate such results yet presented in the literature. The values in Table IV were calculated using the unrestricted Hartree-Fock method for open-shell molecules, which vitiates comparison with some previous works in this area, but the total energies compare favorably with the best available results, as discussed for water.

Another issue that is sometimes asked during the question session after a presentation about high-accuracy theoretical thermochemistry is that of geometry, and we would like to address that here. Clearly, the CCSD(T)/cc-pVQZ geometries used in the HEAT protocol are not the true equilibrium geometries. Rather, there is an intrinsic error of approximately 0.005 Å which must be considered. Obviously, any error in geometry will result in an atomization energy that is a bit too small; it is always further from the bottom of the well to the asymptote than from some point above the minimum. However, the fact that the energy is stationary in this vicinity means that the resulting error will be small. But should it be considered when one is striving for accuracies on the sub-kJ mol⁻¹ level? A typical force constant is 0.5 hartree bohr⁻², which is equivalent to about 4000 kJ mol⁻¹ Å⁻². Hence, an error of 0.005 Å in a bond distance translates to 0.05 kJ mol⁻¹. Therefore it is unlikely that such effects will be important for diatomic molecules, but certainly need to be considered for larger systems.

Finally, we address the issue that differentiates HEAT from the focal-point and Wn family of methods. Specifically, it is of interest to know just how important it is to include core correlation explicitly in HEAT, rather than to save the computational expense that accompanies the significantly larger basis sets that are needed to correlate the core electrons. Is it worth it?

Valence correlation effects are usually obtained by extrapolating energies obtained with very large basis sets, while core correlation effects are assumed to be less important. It is, however, important to define what precisely is meant by "less important" here. The general viewpoint is that core electron correlation is largely an atomic phenomenon; the perturbation of the atomic environment in the formation of

TABLE IV. Estimated Hartree-Fock limit (HFL) energies, correlation energies, and total nonrelativistic electronic energies for molecules in the test suite. For open-shell systems, the partitioning of Hartree-Fock and correlation energies is that associated with the unrestricted Hartree-Fock method, and the energies refer to the spin-orbit averaged energy for atomic ^2P and ^3P states and diatomic $^2\Pi$ states. All values are in a.u.

Species	HFL	E_{corr}	E_{nr}
N_2	−108.9931	−0.5494	−109.5425
H_2	−1.1336	−0.0409	−1.1745
F_2	−198.7735	−0.7570	−199.5306
O_2	−149.6915	−0.6360	−150.3275
C	−37.6937	−0.1513	−37.8450
F	−99.4163	−0.3179	−99.7342
H	−0.5000	0.0000	−0.5000
N	−54.4045	−0.1847	−54.5893
O	−74.8190	−0.2485	−75.0674
CO	−112.7908	−0.5355	−113.3264
C_2H_2	−76.8556	−0.4804	−77.3360
CCH	−76.1835	−0.4303	−76.6139
CH_2	−38.9410	−0.2080	−39.1490
CH	−38.2845	−0.1947	−38.4792
CH_3	−39.5812	−0.2544	−39.8356
CO_2	−187.7254	−0.8763	−188.6017
H_2O_2	−150.8525	−0.7113	−151.5638
H_2O	−76.0675	−0.3713	−76.4388
HCO	−113.3040	−0.5536	−113.8577
HF	−100.0709	−0.3892	−100.4601
HO_2	−150.2527	−0.6620	−150.9147
NO	−129.3095	−0.5905	−129.9000
OH	−75.4281	−0.3103	−75.7384
HNO	−129.8500	−0.6348	−130.4848
CN	−92.2428	−0.4800	−92.7229
HCN	−92.9158	−0.5177	−93.4335
CF	−137.2390	−0.5518	−137.7908
NH_2	−55.5923	−0.2879	−55.8802
NH_3	−56.2250	−0.3392	−56.5642
NH	−54.9864	−0.2353	−55.2217
OF	−174.2109	−0.6752	−174.8861

chemical bonds is obviously less profound than that involving the valence electrons. This is an idea that goes to the very crux of bonding and descriptive chemistry, and we adhere to this belief. The purpose of this section, however, is to explore just how much error is introduced when core and valence electron correlation is separated, and how can we best do this separation.

Let us take a fairly straightforward yet pedagogical approach. It is obviously and manifestly true that the total correlation energy associated with any correlated level of theory can be expressed as

$$E^{\text{exact}} = E_v^\infty + [E_{\text{all}}^{\text{exact}} - E_v^\infty], \tag{6}$$

where E_v^∞ and E_{exact}^∞ are the valence-only and total correlation energies, both evaluated with a complete basis set. The term in brackets represents what is usually known as the core correlation contribution. Let us now assume that all terms on the right-hand side of the equation are to be approximated by means of basis-set extrapolation techniques, viz.,

$$E^{\text{exact}} \sim E_v^A + [E_{\text{all}}^B - E_v^C], \tag{7}$$

where A, B, and C designate (potentially different) extrapolation schemes. Clearly, if scheme B is sufficiently accurate, one can dispense with the entire separation and make the assumption

$$E^{\text{exact}} \sim E_{\text{all}}^B, \tag{8}$$

which is precisely the strategy employed in HEAT. However, if the calculations required by the scheme B are too demanding, one can then invoke the core-valence separation strategy motivated by the considerations above. Then, we have to make the approximation

$$E^{\text{exact}} \sim E^v + E^{\text{cv}}, \tag{9}$$

with

$$E^v \sim E_v^A, \tag{10}$$

and

$$E^{\text{cv}} \sim E_{\text{all}}^B - E_v^C. \tag{11}$$

In the usual scheme, the extrapolation strategy A employs very large basis sets and the chemically more important valence correlation effects are adequately appoximated by E_v^A. The core-valence contribution is assumed to be less important for the property in question, and both B and C involve less demanding calculations. Clearly if B and C provide the exact results, then the total correlation energy error is simply that of the valence contribution. Similarly, if the errors in the valence correlation approximation E_v^C and the total correlation approximation E_v^B are the same, then their difference gives—exactly—the correct core-valence contribution to the correlation energy.

To this point, we have made a completely general discussion that makes no assumptions about the way that the extrapolations are performed, nor which basis sets are used. At this point let us assume the correlation-consistent basis sets popularized by Dunning and co-workers are used and that a two-point extrapolation formula such as the partial-wave (inverse cube) approach advocated by Helgaker *et al.* or the method of Schwenke is employed. The extrapolation schemes A, B, and C then differ only in the choice of the basis sets used, and we shall designate them as TQ, Q5, and 56, these being extrapolations made on the basis of energies obtained with the corresponding VXZ (VQZ and V5Z are designated by Q5, etc.) realizations of one of the hierarchies of basis sets known as cc-pVXZ, aug-cc-pVXZ, or aug-cc-pCVXZ.

The valence correlation energy should be obtained with the largest possible basis sets, such as E_v^{56} obtained in, say, the aug-cc-pVXZ hierarchy. It should be pointed out here that the aug-cc-pCVXZ hierarchy could also be used, as—in principle—the valence correlation energies in the limit of completeness in either hierarchy are the same, but the other choice is clearly preferred on practical grounds. The major question is then how best to calculate the core-valence contribution. Here, there are really a few choices, even within the framework that has been imposed on this discussion. It is common to do identical extrapolations of both contributions,
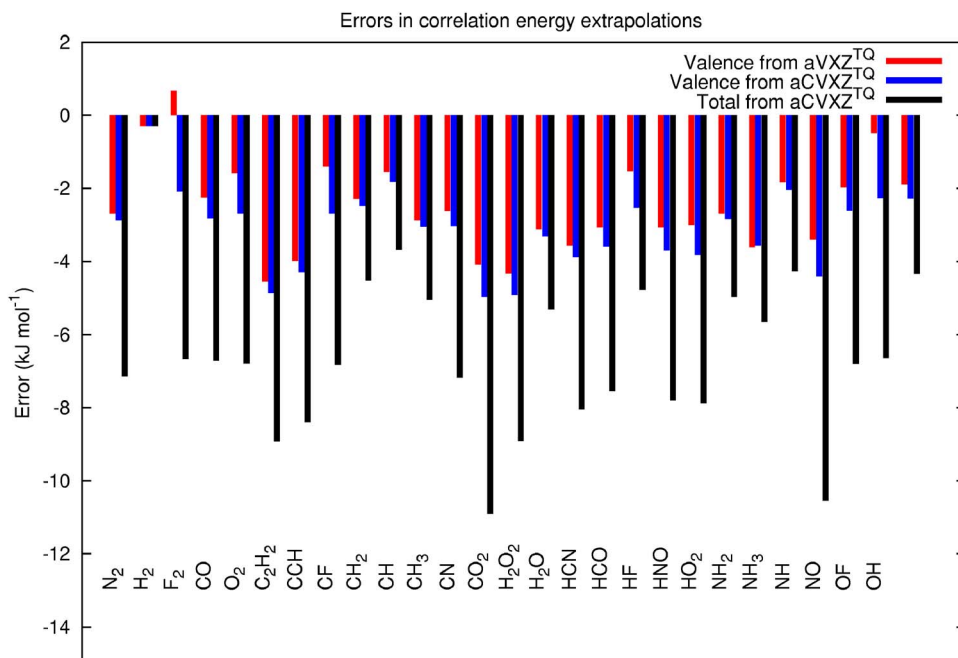
FIG. 3. (Color online) Errors in kJ mol^{-1} in components of the correlation energy obtained with different TQ extrapolation schemes. The red lines are the errors in the valence energy obtained by extrapolating the aug-cc-pVXZ sequence; the blue lines come from extrapolating the valence energy in the aug-cc-pCVXZ sequence, and the black lines are the errors in the total correlation energy obtained from the aug-cc-pCVXZ extrapolation. The points of reference are the 56-extrapolated aug-cc-pVXZ and aug-cc-pCVXZ valence and total energies.

i.e., $A=B$, with the same hierarchy of basis sets. For example, the Wn family of methods developed by Martin and co-workers assumes

$$E^{cv} \sim E_{all}^{TQ} - E_v^{TQ} \qquad (12)$$

using the aug-cc-pCVXZ hierarchy. However, there is a practical issue that should be considered here. Particularly for the smaller elements of this series such as aug-cc-pCVDZ and aug-cc-pCVTZ the inclusion of tight basis functions, ostensibly and commonly thought of as only assisting in the correlation of the core electrons, also tends to assist the valence electron correlation. As a result, the way that the valence correlation scales with the basis set is expected (and is) to be fairly different than if a purely valence basis-set hierarchy like aug-cc-pVXZ series is used. Indeed (see Fig. 3), E_v^{TQ} obtained with the aug-cc-pCVXZ basis sets is systematically inferior to the same quantity obtained with the aug-cc-pVXZ series. Both overshoot the exact value (here approximated with E_v^{56}), with the aug-cc-pVXZ extrapolations generally closer. Hence, if one is interested in a low-level estimate of the valence correlation energy, the latter hierarchy should be used in extrapolations. But recall that it is not the accuracy of the individual contributions in Eq. (11) that are important, but rather the overall accuracy of the difference. There is no question about the hierarchy of basis sets that should be used for the E_{all}^{TQ} term—the aug-cc-pCVXZ series is the only choice. Now, as can be seen again in Fig. 3, the TQ extrapolations of the total correlation energy again overshoot the exact results (again here approximated by E_{all}^{56}) by a magnitude of roughly twice that seen in the valence-only extrapolations. Hence, when the two terms are subtracted from one another, the difference remains negative and is generally *smaller in magnitude* when the less accurate (for the valence energy) E_v^{TQ} based on the aug-cc-pCVXZ series is used. Total errors in correlation energies are seen to be best when this approach (that of Martin) is used for the core-valence con-

tribution, which transfers over to atomization energies (see Fig. 4).

Nonetheless we see that the residual errors from the TQ extrapolation of the core-valence energies using the scheme of Martin lead to errors that are generally small but reach 0.5 kJ mol^{-1} in a few cases. While this is certainly acceptable for the overwhelming number of applications (recall that the definition of chemical accuracy is 1 kcal mol^{-1}—almost an order of magnitude greater—and this is usually an acceptable level of accuracy), it does underscore the idea that those working in this area *must* begin to entertain the idea of abandoning separate treatments of core and valence correlation when the highest levels of accuracy are sought.

V. SUMMARY

The results of this work serve to make a general statement about very high-accuracy theoretical thermochemistry. To wit, the accuracies that can now be obtained—via improvements in quantum chemistry made in the past decade (general coupled-cluster approaches, basis-set extrapolation techniques, and the development of large basis sets necessary to adequately use the extrapolations)—are at the point where it is extremely difficult to judge the quality of comparably accurate strategies, but also sometimes verifiably detrimental to "do a better job." As a result, a critical analysis of the calculated numbers beyond the relatively qualitative discussion epitomized by the content of Table III is not straightforward in general, but must take the form of the brief discussions of OH, HF, and H$_2$O earlier in this paper. For other systems, where the uncertainty in the ATcT data is much greater, it is really not possible to make any sort of detailed analysis at all.

What is clear, and has been implicitly pointed out by Karton *et al.*,[8] is that the time-honored idea of "Pauling points" is valuable here, even at what might seem to be an absurdly high level of theory. The rather simple (in a relative

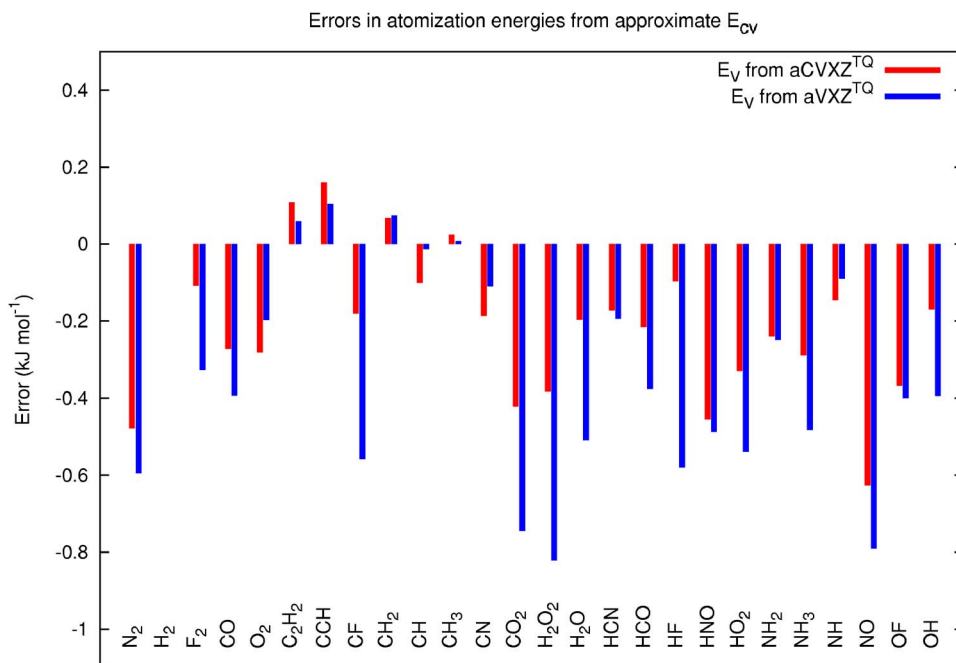Errors in atomization energies from approximate E_{CV}

FIG. 4. (Color online) Errors in the TAE due to the separation of core and valence correlation, as treated approximately by the two schemes mentioned in the text. The point of reference are the all-electron 56-extrapolated aug-cc-pCV*XZ* correlation contributions to the total atomization energies.

sense) HEAT345-(Q) method is, in fact, statistically the *most* accurate of the various approaches, despite being the cheapest. While the argument can be made that there are cases in which HEAT345-(Q) might be prone to larger errors [difficult molecules where the perturbation theory used to justify the CCSDT(Q) method[23] is asked to do "too much"], the conclusion that this simplest HEAT method is the most pragmatic and offers the best compromise between accuracy and cost cannot be realistically refuted. To improve systematically upon HEAT345-(Q), the results of this work and that of Ref. 39 suggest that (a) some treatment of pentuple excitations should be used, (b) the quadruple excitation effects will have to be extrapolated which necessitates large basis-set CCSDTQ and/or CCSDT(Q) calculations, and (c) better extrapolation of CCSD(T) is required, with the implication that the Schwenke approach should be used, perhaps together with septuple-zeta calculations. The careful reader might think at this point: "well, condition (c) is not an obstacle, especially if the sextuple-zeta (or even pentuple-zeta) calculations are sufficient," but the lesson learned in this work and Refs. 8 and 39 is that *all* of these improvements must be made to effect an overall improvement in the quality of the results. The field of high-accuracy thermochemistry has clearly reached an impasse, but quite a satisfactory one since the level of accuracy achievable now is sufficient for almost all needs.

ACKNOWLEDGMENTS

[1] F. D. Rossini, J. Res. Natl. Bur. Stand. **6**, 1 (1931).
[2] J. D. Cox, D. D. Wagman, and V. A. Medvedev, *CODATA Key Values for Thermochemistry* (Hemisphere, New York, 1989).
[3] M. W. Chase, C. A. Davies, J. R. Downey, Jr., D. J. Frurip, R. A. McDonald, and A. N. Syverud, J. Phys. Chem. Ref. Data **14**, 1 (1985).
[4] L. V. Gurvich, I. V. Veyts, and C. B. Alcock, *Thermodynamic Properties of Individual Substances* (Hemisphere, New York, 1989), Vol. 1, Pts. 1 and 2; *ibid.*1991), Vol. , Pts. 1 and 2; *ibid.*1996), Vol. , Pts. 1 and 2.
[5] A. G. Császár, W. D. Allen, and H. F. Schaefer III, J. Chem. Phys. **108**, 9751 (1998).
[6] J. M. L. Martin and G. de Oliveira, J. Chem. Phys. **111**, 1843 (1999).
[7] A. D. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay, and J. Gauss, J. Chem. Phys. **120**, 4129 (2004).
[8] A. Karton, E. Rabinovich, J. M. L. Martin, and B. Ruscic, J. Chem. Phys. **125**, 144108 (2006).
[9] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton, J. Chem. Phys. **121**, 11599 (2004).
[10] Y. J. Bomble, J. Vázquez, M. Kállay, C. Michauk, P. G. Szalay, A. G. Császár, J. Gauss, and J. F. Stanton, J. Chem. Phys. **125**, 064108 (2006).
[11] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[12] D. Feller, J. Chem. Phys. **96**, 6104 (1992).
[13] T. Helgaker, W. Klopper, H. Koch, and J. Noga, J. Chem. Phys. **106**,

9639 (1997).

[14] B. Ruscic, R. E. Pinzon, M. L. Morton, G. von Laszevski, S. J. Bittner, S. G. Nijsure, K. A. Amin, M. Minkoff, and A. F. Wagner, J. Phys. Chem. A **108**, 9979 (2004); B. Ruscic, R. E. Pinzon, G. von Laszewski, D. Kodeboyina, A. Burcat, D. Leahy, D. Montoya, and A. F. Wagner, J. Phys.: Conf. Ser. **16**, 561 (2005); B. Ruscic, R. E. Pinzon, M. L. Morton, N. K. Srinivasan, M.-C. Su, J. W. Sutherland, and J. V. Michael, J. Phys. Chem. A **110**, 6592 (2006).

[15] J. Gauss, A. Tajti, M. Kállay, J. F. Stanton, and P. G. Szalay, J. Chem. Phys. **125**, 144111 (2006).

[16] It is important to realize here that this statement refers to *total* electronic energies, i.e., which is equal to the sum of successive ionization potentials. This, of course, is even more demanding than the total atomization energies, as the error cancellation between nonbonding and, especially, core electrons in an atomization energy scheme is not operative. This is why the estimated error bar is so much greater than that associated with TAEs.

[17] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).

[18] A. K. Wilson, T. van Mourik, and T. H. Dunning, Jr., J. Mol. Struct.: THEOCHEM **388**, 339 (1998).

[19] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[20] T. van Mourik, Mol. Phys. **96**, 529 (1999).

[21] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **103**, 4572 (1995).

[22] J. Noga and R. J. Bartlett, J. Chem. Phys. **86**, 1041 (1987); G. E. Scuseria and H. F. Schaefer, Chem. Phys. Lett. **132**, 382 (1988); J. D. Watts and R. J. Bartlett, J. Chem. Phys. **93**, 6104 (1993).

[23] Y. Bomble, M. Kállay, J. Gauss, and J. F. Stanton, J. Chem. Phys. **123**, 054101 (2005).

[24] M. Kállay and J. Gauss, J. Chem. Phys. **123**, 214105 (2005).

[25] J. F. Stanton and E. F. Valeev (unpublished).

[26] ACESII Mainz-Austin-Budapest, Version 2006, a package of *ab initio* programs. J. F. Stanton, J. Gauss, J. D. Watts, P. G. Szalay, R. J. Bartlett, A. A. Auer, D. E. Bernholdt, O. Christiansen, M. E. Harding, M. Heckert, O. Heun, C. Huber, D. Jonsson, J. Jusélius, W. J. Lauderdale, F. Schiffmann, A. Tajti, M. E. Varner, J. Vázquez; the integral packages MOLECULE (J. Almlöf and P. R. Taylor); PROPS (P.R. Taylor); ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen). For current version, see www.aces2.de.

[27] M. Kállay and P. R. Surján, J. Chem. Phys. **115**, 2945 (2001); for the current version, see http://www.mrcc.hu.

[28] M. E. Harding, T. Metzroth, J. Gauss, and A. A. Auer, J. Chem. Theory Comput. **4**, 6474 (2008).

[29] S. A. Kucharski and R. J. Bartlett, Theor. Chim. Acta **80**, 387 (1991); J. Chem. Phys. **97**, 4282 (1992); N. Oliphant and L. Adamowicz, *ibid.* **94**, 1229 (1991).

[30] M. Kallay and P. Surjan, J. Chem. Phys. **115**, 2945 (2001).

[31] W. Klopper, F. R. Manby, S. Ten-No, and E. F. Valeev, Int. Rev. Phys. Chem. **25**, 427 (2006).

[32] D. W. Schwenke, J. Chem. Phys. **122**, 014107 (2005).

[33] For an authoritative reference, see W. Kutzelnigg, Mol. Phys. **90**, 909 (1997).

[34] N. C. Handy, Y. Yamaguchi, and H. F. Schaefer III, J. Chem. Phys. **84**, 4481 (1986).

[35] All ATcT results reported in this paper are based on the Core (Argonne) Thermochemical Network Version 1.064, 2007.

[36] It may be of some interest to readers that the correlation energy is itself potentially determinable by experiment. See R. M. Gavin and L. S. Bartell, J. Chem. Phys. 45, 4700 (1966).

[37] For reasons of simplicity, all ZPEs, with the exception of OF, were calculated from spectroscopic constants given in K.-P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979). For several diatomic molecules discussed here, there are newer and more accurate spectroscopic data, but the resulting ZPEs would differ by 0.1 cm^{-1} or less. For OF, the spectroscopic constants are from C. E. Miller and B. J. Drouin, J. Mol. Spectrosc. **205**, 312 (2001). All ZPEs include the Dunham Y_{00} term and, where needed, additional correction that reflect spin orbit effects and rotational/nuclear statistics.

[38] C. E. Moore, *Atomic Energy Levels, Natl. Bur. Stand. (U.S.) Circ. No. NSRDS-NBS-35* (U.S. Gp., Washington, DC, 1970), Vol. 1.

[39] A. Karton, P. R. Taylor, and J. M. L. Martin, J. Chem. Phys. **127**, 064104 (2007).

[40] The lowest rotational (F_1) term for an inverted (negative value of the spin-orbit coupling constant, A_0) $^2\Pi_{3/2}$ state is given by the equation $B_0\left(3 - \frac{1}{2}[16 + Y_0(Y_0 - 4)]^{1/2} - (81D_0/16)\right)$, where B_0 and D_0 are the ground-state rotational and centrifugal distortion constants, respectively, and $Y \equiv A_0/B_0$. For details, see G. Herzberg, *Spectra of Diatomic Molecules*, 2nd ed., *12th Printing* (Van Nostrand, Princeton, NJ, 1950), Eqs. (5) and 28, p. 232.

[41] Technically, this is not an *ab initio* result, as is the HEAT456-QP value of 425.92 kJ mol^{-1}, since experimental information has been used. It is, nonetheless, the best result that can be determined by "theory," as the other contributions to the atomization energy used in the HEAT approach are not experimental observables.

[42] L. Bytautas and K. Ruedenberg, J. Chem. Phys. **122**, 014107 (2005).

[43] A. G. Csaszar and I. M. Mills, Spectrochim. Acta, Part A **53**, 1101 (1997).

[44] This argument applies only to the electronic energy contribution and not to the zero-point energy, which presents a different set of problems.

[45] See, for example, S. E. Wheeler, W. D. Allen, and H. F. Schaefer, J. Chem. Phys. **121**, 8800 (2004); M. S. Schuurman, S. R. Muir, W. D. Allen, and H. F. Schaefer, *ibid.* **120**, 11586 (2004).