# W2X and W3X-L: Cost-Effective Approximations to W2 and W4 with kJ mol^{-1} Accuracy

Bun Chan* and Leo Radom*

School of Chemistry and ARC Centre of Excellence for Free Radical Chemistry and Biotechnology, University of Sydney, Sydney, New South Wales 2006, Australia

S *Supporting Information*

ABSTRACT: We have formulated the W2X and W3X-L protocols as cost-effective alternatives to W2 and W3/W4, respectively, and to supplement our previously developed set of W1X and W3X procedures. The W2X procedure provides an accurate approximation to the all-electron scalar-relativistic CCSD(T)/CBS energy, with a mean absolute deviation (MAD) of 0.6 kJ mol^{-1} from benchmark energies provided by the CCSD(T) component in the W4 protocol. Such a performance is comparable to that of W2w (0.5 kJ mol^{-1}) but comes at a significantly lower cost. Comparison of computational requirements shows that W2X should be applicable to systems that can be treated by the W1w method. Thus, W2X provides an accurate means for the treatment of medium-sized systems such as naphthalene. For the calculation of post-CCSD(T) effects, we propose a slight modification to the method used in our previously devised W3X procedure. Our new W3-type protocol (W3X-L) combines this new post-CCSD(T) treatment with our new W2X procedure. It has an MAD from benchmark values of 0.8 kJ mol^{-1} for the W4-11 set, which is comparable to that for the computationally more demanding W3.2 method (0.6 kJ mol^{-1}). However, the use of the even relatively modest post-CCSD(T) calculations in W3X-L still represents a computational bottleneck, and this currently restricts its application to systems up to the size of benzene with our current computing resources.
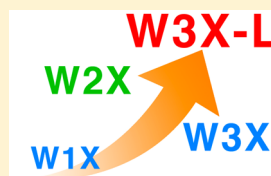
INTRODUCTION

The development of accurate quantum chemistry procedures has come a long way since the introduction of the W1 procedure,[1] which is arguably the first widely applicable protocol to achieve the chemical accuracy of 1 kcal mol^{-1} (4.2 kJ mol^{-1}). Further evolution of the W1 and closely related W2 methods[1] has led to the W3 and W4 procedures,[2,3] which are capable of achieving kJ mol^{-1} accuracy. The W1 and W2 procedures provide approximations to the CCSD(T) method at the complete-basis-set (CBS) limit. The W3 and W4 protocols additionally incorporate post-CCSD(T) correlation terms up to CCSDTQ5.

While the Wn series of composite procedures conceivably represent feasible means for attaining such excellent accuracy, they are computationally demanding. At one end of the spectrum, the W1 protocol is currently restricted to systems with a size roughly up to that of naphthalene (by exploiting its symmetry). At the other extreme, W4 is restricted to molecules with only three or four heavy atoms and even then places a considerable demand on computational resources. The high cost is associated with the computationally demanding coupled-cluster procedures and the often taxing dependence on basis-set size.

In recent years, advances in explicitly correlated techniques[4] have enabled the use of smaller basis sets for coupled-cluster calculations with little compromise on the accuracy. The W1-F12 procedure[5] is the first W1-type protocol to make use of explicitly correlated methods to reduce the cost of W1. Also building on explicitly correlated coupled cluster, we have formulated the W1X-1 and the related W1X-2 procedures[6] to further lower the computational expense of W1-type methods.

The W1X-1 procedure is approximately an order of magnitude less costly than W1 and is also about five times less demanding than W1-F12.

Recently we have built on the foundations of W1X-1 to devise the W3X procedure,[7] with the aim of retaining the general accuracy of W1X-1 for typical systems, while being more robust in challenging cases where CCSD(T) itself is not adequate. This is a somewhat different objective from that for W3, which improves the performance over W1 not only for difficult cases but also for standard systems, at the expense of a significant increase in the demand for computational resources. The W3X procedure, in a similar manner to W3, contains post-CCSD(T) components but relies on a more finely tuned balance between computational requirements and general accuracy.

During the course of development of W3X, we also designed a set of guidelines using new and existing diagnostics to probe the likely accuracy of W3X for a particular system.[7] We found that, for typical systems, the limiting factor for W3X is the quality of the underlying CCSD(T)/CBS component, i.e., W1X-1. Thus, to achieve an improved accuracy in these cases, we recommended the use of the more rigorous W2w procedure to obtain the underlying CCSD(T) energies. We also found that, while the post-CCSD(T) term in W3X is generally accurate for the treatment of higher-order excitation effects, the truncated basis sets employed in this protocol do lead to larger deviations in a small number of cases.

In the present study, we devise new CCSD(T)-based and post-CCSD(T) composite protocols in the spirit of W1X and W3X. In doing so, we aim to further improve the general accuracy for typical systems and the robustness for the more challenging cases. For our new CCSD(T)-based method (W2X), we aim to achieve a similar performance to the W2 protocol. For the new post-CCSD(T) procedure (W3X-L), the target accuracy is that of the W3.2 procedure,[2] which, like W3X, employs the CCSDT(Q) procedure as the highest-level component. At the same time, we hope to keep the additional cost over W1X and W3X to a minimum, such that the new protocols would still be applicable to systems of similar size to those that may be investigated with W1X and W3X.

■ COMPUTATIONAL DETAILS

Standard *ab initio* molecular orbital theory and density functional theory calculations were carried out with the Gaussian 09,[8] Molpro 2012,[9] and ORCA 3.0[10] programs. Geometries were optimized with the B3-LYP/cc-pVTZ+d procedure, as in W1.[1] Following each geometry optimization, harmonic frequency analysis was carried out to confirm the nature of the stationary point as an equilibrium structure. Two-point extrapolation schemes with the formula $E_L = E_{CBS} + AL^{-\alpha}$ were used in the evaluation of the CBS limit,[11] where L is the cardinal number of the basis set, i.e., 3 for triple-ζ and 4 for quadruple-ζ basis sets, etc., and α is an adjustable parameter. Following previous practice,[6,7,12] optimization of α and adjustable scale coefficients involved minimizing the average of the mean absolute deviation (MAD) from benchmark values and the standard deviation (SD) of these deviations. Appropriate components of the W4 energies in the W4-11 set[13] were used for the fitting of the parameters. Literature values for corrections for spin−orbit coupling, where available,[14] were included for atomic species and a number of molecules. To obtain the zero-point vibrational energies (ZPVEs) and thermal corrections for enthalpies at 298 K (ΔH_{298}), we used scaled B3-LYP/cc-pVTZ+d harmonic vibrational frequencies. A frequency scale factor of 0.985 has been proposed for the original W1,[1] and this value is also used in the W1X-type and W3X procedures. In comparison, a value of 0.9896 has been put forward for the related B3-LYP/aug′-cc-pV(T+d)Z method,[15] while for B3-LYP/cc-pVTZ, separate scale factors of 0.9878, 0.9960, and 1.0010 have been suggested, respectively, for ZPVEs, ΔH_{298} and S_{298}.[16] In the present study, we have determined scale factors (for B3-LYP/cc-pVTZ+d) of 0.9886, 0.9926, and 0.9970 for ZPVEs, ΔH_{298}, and S_{298}, respectively, using the methodology of Merrick et al.[16] All relative energies are given in kJ mol^{-1}.

■ RESULTS AND DISCUSSION

Overview of the W3X Procedure. The detailed descriptions of the W1X-1 and W1X-2 protocols, and that for the W3X procedure, are given in refs 6 and 7. Here we briefly summarize the key features of the single-point-energy components in these methods in order to provide a basis for the discussion that follows.

The W1X-1 procedure employs the CCSD-F12b method[17] and the underlying HF(CABS) reference,[17] in conjunction with the cc-pVDZ-F12 and cc-pVTZ-F12 basis sets (denoted in the present study as VnZ-F12),[18] for the extrapolation to the Hartree−Fock limit and to obtain the CCSD/CBS correlation energy. The $\Delta(T)$ component of CCSD(T) within W1X-1, i.e.,

$E[CCSD(T)] - E[CCSD]$, is calculated using conventional CCSD(T) and is extrapolated to the CBS limit with the aug′-cc-pVDZ and aug′-cc-pVTZ basis sets (i.e., cc-pVnZ[19] for hydrogen and aug-cc-pVnZ[20] for heavy atoms, denoted A′VnZ). To include core−valence correlation and scalar-relativistic effects [$\Delta(C+R)$], MP2/CVTZ calculations (CVTZ = cc-pCVTZ[21]) are carried out using both nonrelativistic frozen-core and all-electron DKH approaches. The four components, namely HF(CABS)/CBS, ΔCCSD-F12b/CBS, $\Delta(T)$/CBS, and $\Delta(C+R)$, are combined to give the total W1X-1 electronic energy.

The W1X-2 procedure provides a simpler protocol for the approximation of the all-electron scalar-relativistic CCSD(T)/CBS energy. It uses the A′VDZ and A′VTZ basis sets for the extrapolation of the HF(CABS) energies and the correlation energies for the ΔCCSD-F12b and $\Delta(T)$-F12b components. Summation of these three terms gives the frozen-core CCSD(T)/CBS energy. The combined core−valence correlation and scalar-relativistic-effect component is calculated in the same manner as that for W1X-1, i.e., with MP2/CVTZ.

The W3X procedure is based on W1X-1 but includes additional post-CCSD(T) terms to account for higher-order excitation effects. It contains a CCSDT − CCSD(T) component (ΔT) and a CCSDT(Q) − CCSDT component [$\Delta(Q)$]. The ΔT term is obtained with a truncated VTZ basis set, in which the polarization functions are removed. In a similar manner, a truncated VDZ basis set that contains no polarization functions is used for the calculation of the $\Delta(Q)$ term.

An Improved Approximation to Frozen-Core CCSD(T)/CBS. We now proceed to devise a theoretically more rigorous estimate of valence CCSD(T)/CBS than those used in the W1X-type procedures. The valence components in the W1X-type methods are similar to those used in the W1-F12 protocol. It has been demonstrated that the W2-F12 protocol[5] represents a significantly better approximation of the CCSD-(T)/CBS energy than W1-F12 because of the use of larger basis sets. Thus, we would expect that the terms used for the valence-CCSD(T)/CBS energy in the W2-F12 method will also form an excellent basis for improving the performance of W1X.

In order to minimize the impact on the computational costs associated with the use of larger basis sets, we use the A′VTZ and A′VQZ basis sets instead of the VTZ-F12 and VQZ-F12 basis sets used in W2-F12. The VnZ-F12 basis sets are formulated specifically for CCSD-F12-type procedures, but they are computationally more demanding than the corresponding A′VnZ basis sets. We note that, while extrapolation with the V[D,T]Z-F12 basis sets provides a more accurate estimate for the CCSD-F12b/CBS correlation energies than those with AV[D,T]Z, when triple- and quadruple-ζ basis sets are used, there is only a minimal difference between the two types of basis sets.[22] Karton and Martin have examined a number of additivity and extrapolation schemes for the estimation of the CBS limit in W2-F12.[5] They found that the use of the nonextrapolated HF(CABS) energy, in combination with the extrapolated CCSD-F12b correlation energy and a scaled $\Delta(T)$, provides the most accurate estimation in their test cases. In the present study, we have adopted this same general approach, as our re-examination of a number of alternative procedures has led to the same conclusion. Thus, in our new method, which we will refer to as W2X, the frozen-core CCSD(T)/CBS is calculated as

$$E[\text{CCSD(T)}] = E[\text{HF}] + E[\Delta\text{CCSD}] + E[\Delta(T)]$$

where

$$E[\text{HF}] = E[\text{HF(CABS)/A'VQZ}]$$

$$E[\Delta\text{CCSD}] = E[\Delta\text{CC/Q}] - \{E[\Delta\text{CC/Q}]$$
$$- E[\Delta\text{CC/T}]\} \times 4^{-5.88}/(4^{-5.88} - 3^{-5.88})$$

i.e., CBS extrapolation using $E_L = E_{\text{CBS}} + AL^{-5.88}$ with $L = 3, 4$
where

$$E[\Delta\text{CC/T}] = E[\text{CCSD-F12b/A'VTZ}]$$
$$- E[\text{HF(CABS)/A'VTZ}]$$

and

$$E[\Delta\text{CC/Q}] = E[\text{CCSD-F12b/A'VQZ}]$$
$$- E[\text{HF(CABS)/A'VQZ}]$$

and

$$E[\Delta(T)] = 1.06\{E[\text{CCSD(T)-F12b/A'VTZ}]$$
$$- E[\text{CCSD-F12b/A'VTZ}]\}$$

The exponent in the ΔCCSD-F12b/CBS extrapolation formula and the scale coefficient for the $\Delta(T)$ term are obtained by fitting the CCSD(T)/CBS energies to the valence CCSD(T) component of the W4 energies in the W4-11 set.

We have compared the performance of the frozen-core CCSD(T)/CBS components of various W1-type and W2-type protocols for the W4-11 set, using the corresponding component values in W4 as the benchmark. The statistical results are shown in Table 1. The W1-type methods have MAD values

Table 1. Mean Absolute Deviations (MADs) and Mean Deviations (MDs) (kJ mol^{-1}) from Benchmark Frozen-Core CCSD(T)/CBS Energies in W4 for the Corresponding Component Energies for the Various W1-Type and W2-Type Procedures for the W4-11 Set[a]

method	MAD	MD
W1w	2.1	+1.4
W1-F12	1.2	−1.0
W1X-1	2.4	+1.9
W1X-2	2.6	+2.0
W2w	0.6	+0.5
W2-F12	0.5	+0.1
W2X	0.6	−0.1

[a]Values for W1w, W1-F12, W2w, and W2-F12 are obtained from ref 5.

of 2.1 kJ mol^{-1} (W1w),[2] 1.2 kJ mol^{-1} (W1-F12), 2.4 kJ mol^{-1} (W1X-1), and 2.6 kJ mol^{-1} (W1X-2). In comparison, the W2-type procedures give lower MADs of 0.6 kJ mol^{-1} (W2w),[2] 0.5 kJ mol^{-1} (W2-F12), and 0.6 kJ mol^{-1} (W2X). Thus, there is a notable improvement in accuracy in going from W1-type to W2-type methods. Importantly, the CCSD(T)/CBS component that we use in the W2X protocol is able to achieve a level of performance similar to those for the W2w and W2-F12 procedures.

The Core−Valence Correlation Plus Scalar-Relativistic-Effect Component. In the formulation of the W1X-1 and W1X-2 procedures, the major reduction in computational requirements compared with W1 was achieved by using MP2 instead of CCSD(T) for the calculation of Δ(C+R). With this

consideration in mind, we have sought improvement by examining the use of a number of higher-level procedures for this term but keeping in mind that the computational cost should not be unduly increased.

The methodologies examined included MP2, MP3, MP4SDQ, MP4, and NCEPA1,[23] and these are computed in conjunction with the CVTZ basis set. We have also investigated the use of CCSD(T) with smaller basis sets. The basis sets used are (1) CVDZ, (2) a modified CVTZ basis set [CVTZ(d,p)] with the 2df,2pd polarization functions replaced by the corresponding d,p functions from the CVDZ basis set, and (3) an enlarged CVDZ basis set [CVDZ(2d)] in which the d polarization functions for heavy atoms are replaced by a larger set of 2d polarization functions taken from the CVTZ basis set. Finally, we have also used a composite approach with CCSD(T) and MP2 to obtain the Δ(C+R) term.

It can be seen from Table 2 that, when compared with the sum of the core−valence correlation and the scalar-relativistic

Table 2. Mean Absolute Deviations (MADs) and Mean Deviations (MDs) from Benchmark Core−Valence Correlation Plus Scalar-Relativistic-Effect Component Energies from W4 for the W4-11 Set

method	basis set	MAD	MD
MP2	CVTZ	2.1	−1.9
MP3	CVTZ	2.2	−2.2
MP4SDQ	CVTZ	2.0	−2.0
MP4	CVTZ	1.3	+0.2
NCEPA1	CVTZ	1.1	+0.1
CCSD(T)	CVDZ	2.2	−1.2
CCSD(T)	CVTZ(d,p)	3.2	−3.0
CCSD(T)	CVDZ(2d)	1.2	+0.7
CCSD(T)	∼CVTZ(3d)[a]	0.4	+0.1
CCSD(T)	MTsmall	0.4	−0.4

[a]Obtained using a composite approach, see text for details.

terms in W4 for the W4-11 set, the use of MP2, MP3, and MP4SDQ all lead to similar performance, with MAD values of ∼2 kJ mol^{-1}. It is also apparent that these three methods show systematic negative deviations from the benchmark values (MD ∼ −2 kJ mol^{-1}), which, in the case of MP2 within the W1X-1 and W1X-2 protocols, somewhat compensate for the systematic positive deviations in the valence components (Table 1, MD ∼ +2 kJ mol^{-1}). The inclusion of triple excitation effects in the full MP4 method for the treatment of Δ(C+R) yields notably better agreement with the benchmark values, with an MAD of 1.3 kJ mol^{-1}. The NCEPA1 procedure gives rise to a similar accuracy to MP4, with an MAD of 1.1 kJ mol^{-1}.

For the Δ(C+R) energies calculated with CCSD(T), the use of the CVDZ basis set leads to an MAD of 2.2 kJ mol^{-1}, which does not represent an improvement over the MP2/CVTZ method used in W1X-1 and W1X-2. We note that the Δ(C+R) component in W1, which is obtained using CCSD(T) with a triple-ζ basis set (MTsmall[1]), is quite accurate. It thus appears that to obtain a high-quality Δ(C+R) term with CCSD(T) may require a basis set that is larger than CVDZ. For this reason, we have examined two basis sets that are larger than CVDZ but still smaller than CVTZ, whose use can be expected to be computationally quite restrictive. The CVTZ(d,p) basis set provides additional flexibility for the valence orbitals over CVDZ, while the CVDZ(2d) basis set has a larger set of d polarization functions for heavy atoms. We can see that the CVTZ(d,p)

basis set in fact leads to a deterioration in the performance (MAD = 3.2 kJ mol^{-1}). On the other hand, the inclusion of a larger set of polarization functions in CVDZ(2d) gives an MAD of 1.2 kJ mol^{-1}, which is similar to those for MP4/CVTZ and NCEPA1/CVTZ.

We now combine the CCSD(T)/CVDZ(2d) and MP2/CVTZ(3d) approaches in an attempt to improve the performance for the calculation of Δ(C+R) with minimal additional computational expense over CCSD(T)/CVDZ(2d) itself.[24] The CVTZ(3d) basis set is defined in a similar manner to the CVDZ(2d) basis set. Thus, the 2d polarization functions for heavy atoms are substituted by the corresponding 3d functions in the CVQZ basis set. The combined Δ(C+R) is obtained in a similar manner to the composite scheme employed in, for example, G4(MP2)-type procedures,[25] i.e., using MP2 with a triple-ζ basis set and a correction term for CCSD(T) obtained with a double-ζ basis set. The Δ(C+R) term in the present study is calculated with the formula

$$E[\Delta(C+R)] = 1.14E[\Delta(C+R):MP2/TZ]$$
$$+ 1.56E[\Delta(C+R):\Delta CCSD(T)]$$

where

$$E[\Delta(C+R):MP2/TZ] = E[MP2(Full, DKH)/CVTZ(3d)]$$
$$- E[MP2(FC)/CVTZ(3d)]$$

and

$$E[\Delta(C+R):\Delta CCSD(T)]$$
$$= E[\Delta(C+R):CCSD(T)] - E[\Delta(C+R):MP2]$$

where

$$E[\Delta(C+R):CCSD(T)]$$
$$= E[CCSD(T)(Full, DKH)/CVDZ(2d)]$$
$$- E[CCSD(T)(FC)/CVDZ(2d)]$$

and

$$E[\Delta(C+R):MP2] = E[MP2(Full, DKH)/CVDZ(2d)]$$
$$- E[MP2(FC)/CVDZ(2d)]$$

The coefficients for the scaling of the CCSD(T) and MP2 components are obtained by fitting to the W4 Δ(C+R) energies for the W4-11 set, in a similar manner to obtaining the parameters for the CCSD(T)/CBS term in W2X. Using this composite approach, the MAD against the benchmark values is further lowered to 0.4 kJ mol^{-1}. This is comparable to that for the Δ(C+R) component in W1w, which requires CCSD(T) calculations using a completely decontracted cc-pVTZ-type basis set (MTsmall). As a result of the superior performance of the composite Δ(C+R) protocol, we use this approach for W2X. We note that the calculation of this term requires two CCSD(T) and two MP2 calculations. In comparison, the corresponding Δ(C+R) component in W2-F12 requires four CCSD(T) and two CCSD calculations, with all the basis sets used being of triple-ζ size.

Performance of W2X as an Approximation to Relativistic All-Electron CCSD(T)/CBS. The complete W2X result is obtained by combining the valence CCSD(T)/CBS and composite-Δ(C+R) components. We will now examine its accuracy, and we will discuss the computational costs of W2X, and that of the new W3-type procedure to be introduced, at a

Table 3. Mean Absolute Deviations (MADs) and Mean Deviations (MDs) (kJ mol^{-1}) from Benchmark Relativistic All-Electron CCSD(T)/CBS Energies in W4 for the Various W1-Type and W2-Type Procedures for the W4-11 Set[a]

method	MAD	MD
W1w	1.8	+1.0
W1-F12	1.6	−1.4
W1X-1	2.0	−0.2
W1X-2	2.2	−0.2
W2w	0.5	+0.2
W2-F12	0.7	−0.3
W2X	0.6	−0.1

[a]Values for W1w, W1-F12, W2w, and W2-F12 are obtained from ref 5.

later stage. Table 3 shows the performance of W2X against benchmark relativistic all-electron CCSD(T)/CBS energies obtained from the relevant components of W4 energies, together with the MADs and MDs of various W1- and W2-type procedures. The W1-type methods generally have MAD values of ∼1.5−2 kJ mol^{-1}. Interestingly, while the W1w values are systematically too high (MD = +1.0 kJ mol^{-1}) and those for W1-F12 are generally too low (−1.4 kJ mol^{-1}), the MD values for W1X-1 and W1X-2 are fairly close to zero (−0.2 kJ mol^{-1}).

In comparison with the W1-type methods, all three W2-type procedures have considerably smaller MAD values of ∼0.5 kJ mol^{-1}. There is little systematic bias in the deviations for these three procedures, with MD values of +0.2 (W2w), −0.3 (W2-F12), and −0.1 (W2X) kJ mol^{-1}. Overall, the W2X procedure appears to have an accuracy that is very close to those for the other two W2-type protocols, and this seems to be a result of its good performance for both the valence CCSD(T)/CBS (Table 1) and Δ(C+R) (Table 2) components.

A Post-CCSD(T) Component with Improved Accuracy. With a refined composite-CCSD(T)/CBS protocol established in the form of W2X, we now turn our attention to post-CCSD(T) contributions to the energies. The relatively good computational efficiency of the $\Delta(T)+$ term in W3X is a result of the complete omission of polarization functions in the VnZ basis sets used in its calculation. These basis sets are denoted VnZ(NP) in the present study, where NP stands for "No Polarization". In order to yield an improved performance for this component with only a small impact on the computational efficiency, we have examined two approaches in which small sets of polarization functions are reintroduced into these basis sets.

In the first approach [mW3X(a), with m signifying a modified W3X], the ΔT component, i.e., $E[CCSDT] - E[CCSD(T)]$, is obtained using a modified VTZ basis set. It contains no polarization functions on hydrogen and has the 2df functions for heavy atoms replaced by the d polarization functions taken from the VDZ basis set. This will be referred to as the VTZ(d) basis set. The $\Delta(Q)$ component is calculated with the formula $E[CCSDT(Q)/VDZ(d)] - E[CCSDT/VDZ(d)]$, with VDZ(d) representing the VDZ basis set with no polarization functions on hydrogen. The total post-CCSD(T) contribution $[\Delta(T)+]$ is obtained as the sum of scaled ΔT and $\Delta(Q)$ terms:

$$E[\Delta(T)+:mW3X(a)]$$
$$= 1.06E[\Delta T/VTZ(d)] + 0.69E[\Delta(Q)/VDZ(d)]$$

The scaling coefficients are obtained by optimization, using as a benchmark the total post-CCSD(T) contributions in W4 for the

W4-11 set. This formulation of the $\Delta(T)+$ energy is completely analogous to that for W3X, with the differences being that the $VnZ(d)$ basis sets are used in place of $VnZ(NP)$ and the scaling parameters are refitted.

The second procedure [mW3X(b)] is identical to the first one for the $\Delta(Q)$ component, i.e., $\Delta(Q) = E[CCSDT(Q)/VDZ(d)] - E[CCSDT/VDZ(d)]$. For the ΔT term, in addition to computations with the VTZ(d) basis set, CCSDT and CCSD(T) energies are also obtained with the VDZ(NP) basis set. These additional calculations enable basis-set extrapolation because, while the basis sets have truncated polarization functions, the progression from VDZ(NP) to VTZ(d) can nonetheless be considered systematic. The extrapolation is performed with the standard formula:

$$E[\Delta T/CBS] = E[\Delta T/VTZ(d)] - \{E[\Delta T/VTZ(d)]$$
$$- E[\Delta T/VDZ(NP)]\} \times 3^{-2.61}/(3^{-2.61} - 2^{-2.61})$$

The total $\Delta(T)+$ energy contains this term and a scaled $\Delta(Q)$ energy:

$$E[\Delta(T)+:mW3X(b)]$$
$$= E[\Delta T/CBS] + 0.78E[\Delta(Q)/VDZ(d)]$$

The exponent for the ΔT term and the scaling coefficient for $\Delta(Q)$ are determined by simultaneous fitting to the W4 total post-CCSD(T) contributions for the W4-11 set. We note that mW3X(b) is similar to the post-CCSD(T) term employed for W4lite,[27] with key differences being the use of smaller basis sets for mW3X(b) and accordingly different values for the ΔT exponent and the $\Delta(Q)$ coefficient.

Let us now look at the performance of these approaches for the calculation of post-CCSD(T) effects in comparison with those obtained from the $\Delta(T)+$ terms of W3X and the more rigorous W3.2 procedure (Table 4). The MAD from W4 values

Table 4. Mean Absolute Deviations (MADs) and Mean Deviations (MDs) (kJ mol⁻¹) from Benchmark Post-CCSD(T) Energies in W4 for the Various W3-Type Procedures for the W4-11 Set

method	MAD	MD
W3X	1.0	+0.3
mW3X(a)	0.8	+0.4
mW3X(b)	0.6	+0.2
W3.2	0.7	−0.3

for the $\Delta(T)+$ term for W3X is 1.0 kJ mol⁻¹, and that for W3.2 is 0.7 kJ mol⁻¹. Thus, the highly truncated $VnZ(NP)$ basis sets used in the post-CCSD(T) calculations for W3X already yield fairly reasonable energies. The use of the mW3X(a) approach with the addition of minimal sets of d polarization functions, i.e., $VnZ(d)$, lowers the MAD value from 1.0 to 0.8 kJ mol⁻¹. Supplementing this with the use of VDZ(NP) calculations to enable basis-set extrapolation for the ΔT term in mW3X(b) further lowers the MAD to 0.6 kJ mol⁻¹. It is noteworthy that, in mW3X(b), the CCSDT/VDZ(NP) calculation represents the lowest-cost post-CCSD(T) computation within the total composite $\Delta(T)+$ protocol, and its inclusion is therefore a cost-effective means to improve the performance over mW3X(a).

The W3X-L Protocol Based on an Underlying W2X Energy. The results in Table 3 show that W2X (MAD = 0.6 kJ mol⁻¹) represents a significant improvement over W1X-1 (2.0 kJ mol⁻¹) as an approximation for the all-electron

scalar-relativistic CCSD(T)/CBS energy. Table 4 indicates that the mW3X(b) approach has a somewhat improved performance (0.6 kJ mol⁻¹) compared with that in the original W3X (1.0 kJ mol⁻¹) for the computation of post-CCSD(T) effects. We now combine the $\Delta(T)+$ energy derived from mW3X(b) with the all-electron scalar-relativistic CCSD(T)/CBS energy of W2X to yield an improved W3X approach. We anticipate that this new W3-type method would have a considerably improved performance over the original W3X, primarily due to the significantly smaller MAD for the CCSD(T)/CBS component. We term the new procedure W3X-L, with L signifying the use of larger basis sets (L), both with the use of W2X to provide the underlying CCSD(T)/CBS component, as well as mW3X-(b) for the post-CCSD(T) term.

The MADs and MDs from the total W4 energies for the various WnX procedures are shown in Table 5. The W1-type

Table 5. Mean Absolute Deviations (MADs) and Mean Deviations (MDs) (kJ mol⁻¹) from Benchmark W4 Energies for the Various Wn-Type Procedures for the W4-11 Set[a]

| method	MAD	MD
W1w	2.0	+0.5
W1-F12	2.1	−2.1
W1X-1	2.1	−1.5
W1X-2	2.2	−1.5
W2w	1.1	−0.6
W2-F12	1.3	−0.9
W2X	1.8	−1.3
W3.2	0.6	+0.0
W3-F12	0.8	−0.3
W3X	1.9	+0.1
W3X-L	0.8	+0.2

[a]Values for W1w, W1-F12, W2w, W2-F12, W3.2, and W3X are obtained from ref 5.

protocols typically have MAD values of ~2 kJ mol⁻¹. For the procedures that employ explicitly correlated methods, namely W1-F12, W1X-1, and W1X-2, the deviations are, in general, fairly systematic, with MD values of ~−1.5 to −2 kJ mol⁻¹. The MADs for the W2-type methods are in the range of 1–2 kJ mol⁻¹, and the corresponding MD values are between −0.6 and −1.3 kJ mol⁻¹. In comparison with the W1- and W2-type procedures, the W3-type methods all have MD values that are close to zero, which suggests that systematic deviations have largely been eliminated. Among these, W3X has an MAD of 1.9 kJ mol⁻¹. Such a modest improvement over W1X-1 (2.1 kJ mol⁻¹), on which W3X is based, has been attributed to the relatively modest performance of W1X-1 as an approximation to CCSD(T)/CBS.[7] This proposition is further supported by the results of the present study (Table 1). The three more rigorously formulated W3-type procedures all have notably smaller MADs of 0.6 (W3.2), 0.8 (W3-F12[5]), and 0.8 (W3X-L) kJ mol⁻¹. Thus, the W3X-L procedure shows a significant improvement over W3X, with a performance that approaches that of the W3.2 protocol, which is the most rigorous W3-type procedure in Table 5, but involves a significantly higher computational cost.

In our previous development of W3X, the protocol was tested against several challenging problems. They included the automerization of cyclobutadiene, the dissociation of F_2, and the atomization of chromium oxides. Among these, the W3X procedure was shown to provide an accurate means for the

computation of the barrier for cyclobutadiene automerization. For F_2 dissociation, it mostly corrected for the deficiency of CCSD(T)-based methods but still did not completely match the performance of the more rigorous W3.2 method. For the most difficult case of the atomization of chromium oxides,[26] even W3.2 did not yield a satisfactory result.

Having demonstrated the general accuracy of W3X-L for the W4-11 set, we now proceed to further evaluate its applicability to the dissociation of F_2 (Figure 1), for which the use of the
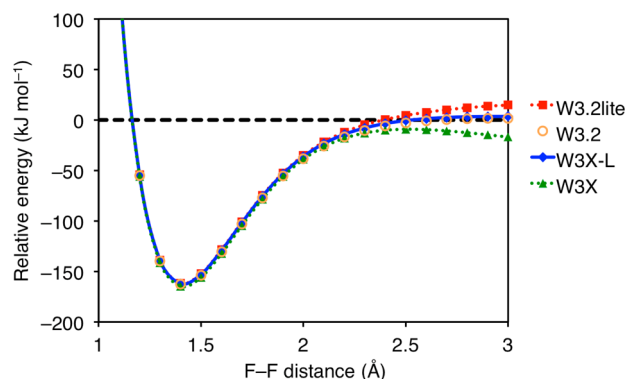
Figure 1. F_2 homolytic dissociation curves for W3-type procedures with zero defined as the energy for $2 \times F\bullet$ at the respective level.

more rigorous $\Delta(T)+$ term may lead to an improvement compared with W3X. We have previously shown that both the W3X and W3.2lite[27] procedures start to deviate from the correct dissociation behavior at an F–F distance of ~2 Å. On the other hand, it has been demonstrated that W3.2 properly reproduces the potential energy curve. It can be seen that the W3X-L procedure also shows the correct qualitative behavior for F_2 dissociation and in fact almost completely matches the W3.2 profile at all points.

Timing Comparisons for the Wn-Type Procedures. In terms of computational requirements, how does the new W2X protocol compare with other W1- and W2-type procedures, and how does W3X-L fare among W3-type methods? We use benzene for our comparison of resource usage for the various Wn-type methods. In all cases, the memory provided to the calculations is limited to 8 GB, which is a typical amount used in regular workstations at present. The memory and scratch disk usage, and relative timings, are shown in Table 6.

First of all, it is noteworthy that all of the calculations are computationally feasible with the 8 GB memory constraint. Nonetheless, W2w has the largest memory requirement with a maximum usage of 7.3 GB. For larger species or molecules with lower symmetry, a computer system with 8 GB of memory may not be sufficient for W2w computations.

There is a large range of disk usage for the different Wn-type procedures. The most demanding calculations in W1X-1 and W1X-2 involve CCSD(T) with a triple-ζ basis set. For benzene, such a combination leads to disk usage of ~10 GB. The demand on scratch disk size increases notably to ~60 GB for W1w and W2X, for which CCSD computations with quadruple-ζ basis sets are required. For W2w, the use of a pentuple-ζ basis set in the CCSD computation leads to a further, very dramatic, rise in the disk usage to 413.1 GB. Such a severe demand on the scratch disk is again likely to limit the applicability of W2w. For the $\Delta(T)+$ terms in W3-type procedures, the sizes of 106.0 and 284.8 GB for the disk usages for W3X and W3X-L are also quite substantial, but there is still room for the treatment of slightly

Table 6. Timings (Relative to That of W1X-1) and Memory and Scratch Disk Usage (GB) for the Various Wn-Type Procedures

	memory	disk	relative time
W1 and W2-Type Procedures			
W1X-2	0.6	6.5	0.7
W1X-1	0.7	10.0	1.0^a
W1w	2.2	59.2	4.6
W2X	2.2	59.2	3.0
W2w	7.3	413.1	38.3
$\Delta(T)+$ Component for W3-Type Procedures			
W3X	0.7	106.0	27.9
W3X-L	1.9	284.8	75.5
CCSDT/cc-pVDZ^b	1.3	196.5	35.0^c

^{a}Corresponds to 27 min on one core. ^{b}The computationally least demanding post-CCSD(T) calculation in W3.2. ^{c}The CCSDT/VTZ and CCSDT(Q)/VDZ component calculations both require >200 h.

larger systems with modern workstations, which typically have ~1−2 TB of disk space.

In terms of relative timing, the W1-type methods and W2X are fairly similar in this regard for the moderately sized system of benzene. Nonetheless, a significant observation is that the W2X procedure is in fact slightly less costly than W1w. In contrast, the W2w procedure is more than an order of magnitude more costly than W2X, and such an observation resonates with our findings for other resource types. Post-CCSD(T) [$\Delta(T)+$] calculations dominate the time required for W3X and W3X-L computations. Nonetheless, they are attainable with actual consumed times of 13 and 35 h, respectively, for W3X and W3X-L (shown as relative times of 27.9 and 75.5 in Table 6). In comparison, the CCSDT/VDZ calculation, which is the least demanding component within the $\Delta(T)+$ term of W3.2, already requires 17 h of computational time, and the more expensive CCSDT/VTZ and CCSDT(Q)/VDZ calculations within the W3.2 protocol did not complete within our preset limit of 200 h.

Description of W2X and W3X-L and Application to the G2/97 Set. We now provide a summary description of the W2X and W3X-L protocols for the sake of convenience to the readers, and we present W3X-L values for the G2/97 set[28] of thermochemical properties. The timing analysis shown in Table 6 suggests that the W2X protocol is likely to be applicable to systems that are considerably larger than benzene. However, the use of the $\Delta(T)+$ component as defined by W3X-L may be computationally quite demanding. The largest systems in G2/97 are the radical cations $C_6H_5Me^{\bullet+}$, $C_6H_5OH^{\bullet+}$, and $C_6H_5NH_2^{\bullet+}$, and we find that the computation of their W3X-L $\Delta(T)+$ terms are too demanding with our present computational resources. Thus, the $\Delta(T)+$ component of W3X is combined with W2X in the calculation of the associated ionization energies in these three cases.

The W2X protocol involves the following computations:
- The geometries and vibrational frequencies are obtained at the B3-LYP/cc-pVTZ+d level. ZPVE, 298 K enthalpy, and entropy thermal correction values are obtained with frequency scale factors of 0.9886, 0.9926, and 0.9970, respectively.
- The HF energy is provided by HF(CABS)/A′VQZ.
- The ΔCCSD/CBS component is obtained with CCSD-F12b using the A′VTZ and A′VQZ basis sets, and extrapolated using the formula $E_L = E_{CBS} + AL^{-5.88}$.

Table 7. W3X-L Energies for the G2/97 Test Set (kJ mol^{-1})^{a,b}

Heats of Formation (298 K)

LiH	139.7	F_2	2.0	C_2Cl_4	−21.1	CH_3CHO	−166.8
BeH	337.3	CO_2	−395.1	CF_3CN	−499.2	glyoxal	−214.5
CH	595.3	Na_2	139.6	propyne	185.6	EtOH	−236.5
CH_2 triplet	391.2	Si_2	585.8	allene	189.5	Me_2O	−185.5
CH_2 singlet	428.6	P_2	145.5	cyclopropene	283.6	thiooxirane	75.3
CH_3	146.0	S_2	126.0	propene	18.9	DMSO	−155.7
CH_4	−74.7	Cl_2	2.2	cyclopropane	52.9	EtSH	−48.7
NH	358.8	NaCl	−179.3	propane	−107.1	Me_2S	−41.4
NH_2	186.0	SiO	−99.7	butadiene	110.5	CH_2CHF	−143.6
NH_3	−45.2	CS	281.8	2-butyne	147.0	EtCl	−113.6
OH	36.8	SO	4.5	CH_2=cyclopropane	192.2	CH_2CHCl	21.3
H_2O	−241.8	ClO	103.1	bicyclobutane	223.5	CH_2CHCN	188.3
HF	−273.1	ClF	−55.0	cyclobutene	159.8	Me_2CO	−218.3
SiH_2 singlet	267.7	Si_2H_6	68.6	cyclobutane	25.0	CH_3CO_2H	−432.5
SiH_2 triplet	356.3	CH_3Cl	−83.2	isobutene	−18.4	CH_3COF	−441.6
SiH_3	194.9	CH_3SH	−25.3	butane	−129.1	CH_3COCl	−242.6
SiH_4	28.6	HOCl	−75.9	isobutane	−136.5	chloropropane	−135.6
PH_2	132.7	SO_2	−299.2	spiropentane	183.4	isopropanol	−275.4
PH_3	5.5	BF_3	−1131.5	C_6H_6	81.5	MeOEt	−220.4
H_2S	−22.3	BCl_3	−399.4	CH_2CF_2	−453.2	Me_3N	−26.1
HCl	−92.4	AlF_3	−1204.2	CHF_3	−699.9	furan	−35.8
Li_2	216.0	$AlCl_3$	−584.9	CH_2Cl_2	−94.7	thiophene	111.5
LiF	−339.5	CF_4	−937.4	$CHCl_3$	−100.4	pyrrole	107.2
C_2H_2	229.7	CCl_4	−92.6	CH_3NH_2	−21.7	pyridine	139.3
C_2H_4	52.2	COS	−143.6	CH_3CN	74.3	H_2	0.6
C_2H_6	−85.5	CS_2	113.7	CH_3NO_2	−74.2	SH	140.5
CN	440.7	CF_2O	−608.9	CH_3ONO	−66.9	CCH	568.3
HCN	131.0	SiF_4	−1615.3	CH_3SiH_3	−30.9	C_2H_3	296.5
CO	−110.2	$SiCl_4$	−663.7	HCO_2H	−379.5	CH_3CO	−11.2
HCO	41.4	N_2O	83.4	HCO_2CH_3	−360.4	CH_2OH	−17.5
CH_2O	−109.8	ClNO	55.2	CH_3CONH_2	−236.5	CH_3O	18.6
CH_3OH	−202.2	NF_3	−133.7	aziridine	127.1	EtO	−15.7
N_2	2.3	PF_3	−957.3	NCCN	311.6	CH_3S	120.4
N_2H_4	97.5	O_3	146.1	Me_2NH	−17.4	C_2H_5	119.8
NO	92.2	F_2O	25.8	$EtNH_2$	−51.0	Me_2CH	86.9
O_2	1.9	ClF_3	−163.5	ketene	−49.1	Me_3C	50.6
H_2O_2	−134.3	C_2F_4	−676.3	oxirane	−53.3	NO_2	34.5

Ionization Energies (0 K)

H	1312.6	HF	1547.4	B_2F_4	1131.3	CH_3Cl	1085.5
He	2372.6	SiH_4	1061.0	CO_2	1332.4	EtOH	1002.6
Li	519.9	PH	979.1	CF_2	1102.3	CH_3CHO	986.7
Be	898.4	PH_2	947.0	COS	1079.3	CH_3OF	1094.3
B	802.1	PH3	950.2	CS_2	973.5	thiooxirane	872.3
C	1084.5	SH	1002.8	CH_2	1000.4	NCCN	1291.8
N	1401.6	SH_2 $^2B_1$	1007.2	CH_3	948.9	furan	851.0
O	1312.9	SH_2 $^2A_1$	1230.5	C_2H_5	783.6	pyrrole	789.1
F	1680.5	HCl	1227.1	cyclopropene	928.7	C_6H_5OH	820.7^b
Ne	2080.2	C_2H_2	1098.6	allene	934.4	$C_6H_5NH_2$	741.9^b
Na	493.9	C_2H_4	1013.6	C_3H_7	717.6	B_2H_4	923.0
Mg	734.8	CO	1350.8	C_6H_6	887.0	NH	1299.0
Al	577.1	N_2 Σ	1504.9	C_6H_5Me	848.2^b	NH_2	1077.5
Si	782.5	N_2 Π	1608.3	CN	1344.0	N_2H_2	925.4
P	1009.5	O_2	1165.2	HCO	786.8	N_2H_3	734.3
S	996.7	P_2	1011.4	CH_2OH	727.4	HOF	1225.0
Cl	1248.0	S_2	901.0	CH_3O	1033.0	SiH_2	879.7
Ar	1518.0	Cl_2	1106.6	CH_3OH	1047.6	SiH_3	783.1
CH_4	1219.0^a	ClF	1217.9	CH_3F	1211.1	Si_2H_2	788.5
NH_3	981.5	CS	1090.8	CH_2S	902.8	Si_2H_4	778.4
OH	1255.4	BF_3	1515.1	CH_2SH	727.4	Si_2H_5	739.0
H_2O	1216.4	BCl_3	1133.0	CH_3SH	909.5	Si_2H_6	930.3

Table 7. continued

Electron Affinities (0 K)

Li	59.2	NH_2	71.6	CCO	221.3	CH_3O	148.6
B	23.4	OH	174.5	CF_2	14.4	CH_3S	178.8
C	120.3	SiH	119.6	NCO	347.7	CH_2S	47.2
O	138.7	SiH_2	103.9	NO_2	218.4	CH_2CN	148.1
F	327.6	SiH_3	133.6	O_3	203.9	CH_2NC	108.6
Na	52.8	PH	97.3	OF	219.0	HCCO	224.0
Al	38.5	PH_2	120.2	SO_2	106.1	CH_2CHO	175.7
Si	132.4	SH	222.0	S_2O	178.4	CH_3CO	37.8
P	70.4	$O2$	42.1	CCH	284.4	EtO	161.1
S	199.2	NO	2.2	C_2H_3	64.0	EtS	187.6
Cl	347.9	CN	371.6	CH_2CC	172.9	LiH	33.2
CH	115.7	PO	103.7	CH_2CCH	87.4	HNO	31.0
CH_2	61.5	S_2	161.3	C_3H_5	46.4	HO_2	102.0
CH_3	2.1	Cl_2	229.7	HCO	27.3		
NH	32.6	C_2	313.2	HCF	50.8		

Proton Affinities (298 K)

NH_3	853.7	C_2H_2	645.2	PH_3	785.1	HCl	561.4
H_2O	690.1	SiH_4	640.9	H_2S	709.3	H_2	423.8

[a]Based on CCSD(T)/VTZ geometry and ZPVE for CH_4^+. [b]The $\Delta(T)$+ of W3X is used in conjunction with W2X in the calculation of this value.

- The $\Delta(T)$ portion of the valence CCSD(T) correlation energy is obtained from the (T) correlation energy of CCSD(T)-F12b/A'VTZ and scaled by 1.06.
- A combined core-correlation and scalar-relativistic component is obtained using a composite approach with the following procedure:

  ○ An MP2 term is taken as the difference between the frozen-core and all-electron-DKH MP2/CVTZ(3d) energies, where CVTZ(3d) is the CVTZ basis set with the d components replaced by those from CVQZ for heavy atoms.

  ○ A ΔCCSD(T) correction to the above MP2 relative energy is obtained with the CVDZ(2d) basis set, which contains the sp components of CVDZ and d functions from CVTZ for heavy atoms.

  ○ The total core-correlation and scalar-relativistic term is obtained by the linear combination of these two terms: $1.14 \times E[MP2] + 1.56 \times E[\Delta CCSD(T)]$.

For W3X-L, the W2X energy is combined with the following valence post-CCSD(T) components:

- A CCSDT − CCSD(T) component is calculated by extrapolation using the formula $E_L = E_{CBS} + AL^{-2.61}$ with the VDZ(NP) and VTZ(d) basis sets, where VDZ(NP) contains only the sp functions from VDZ, and VTZ(d) is comprised of the sp component of VTZ and the d components from VDZ.
- A CCSDT(Q) − CCSDT term is obtained with the VDZ(d) basis set and scaled by 0.78, where VDZ(d) represents the VDZ basis set but without the p function for hydrogen.

The W3X-L energies for the G2/97 set are shown in Table 7. A major source of error is presumably from the use of B3-LYP/cc-pVTZ+d geometries and, more importantly, frequencies for the calculation of ZPVEs and thermal corrections for enthalpy. For the B3-LYP/cc-pVTZ+d procedure, we find a root-mean-square deviation (RMSD) of 0.35 kJ mol^{-1} from benchmark experimental values[16] for the ZPVEs for a set of small molecules (with an average ZPVE of 28.1 kJ mol^{-1}). If we assume that the deviation grows linearly with the size of the molecule, for the somewhat larger species in G2/97 (average scaled B3-LYP/cc-pVTZ+d ZPVE = 85.6 kJ mol^{-1}), we can expect a larger RMSD value of ~1.1 kJ mol^{-1}. The RMSD for thermal corrections is 0.13 kJ mol^{-1} for the small molecule set, which have an average value of 12.8 kJ mol^{-1}. For the G2/97 set, for which the average thermal correction is 11.6 kJ mol^{-1}, we assume an RMSD value of 0.2 kJ mol^{-1}. Based on standard error propagation principles, we assume a total RMSD value of 1.1 kJ mol^{-1} $[(1.1^2 + 0.2^2)^{1/2}]$ in association with the use of ZVPEs and thermal corrections obtained with scaled B3-LYP/cc-pVTZ+d frequencies. This corresponds to an MAD of 0.9 kJ mol^{-1}, assuming a typical value of $(2/\pi)^{1/2}$ for the MAD/RMSD ratio.[29]

For the electronic energies, the MAD for W3X-L for the W4-11 set is 0.8 kJ mol^{-1} (Table 5). This set of molecules on average contains small molecules of ~3.6 atoms. Assuming a linear increase in the deviations with system size, we propose an MAD of 1 kJ mol^{-1} for the electronic energies for the G2/97 set (~4.3 atoms on average), and on average a total presumed MAD of ~1.5 kJ mol^{-1} with the inclusion of the ZPVEs and thermal corrections. We believe that our W3X-L method provides the highest-level theoretical values currently available for the thermochemical quantities in the full G2/97 set and complements another recent study on the atomization energies in this set.[30] We also believe the W3X-L values to be somewhat more accurate in many cases than the experimental ones, particularly for a number of cases that will be discussed in the following section.

Outliers in the G2/97 Set. We note that four of our W3X-L energies show deviations that are more than 10 kJ mol^{-1} from designated experimental values,[28] which is surprising given the estimated accuracy of the W3X-L procedure. These are the 298 K heats of formation ($\Delta_f H_{298}$s) for C_2F_4 and CH_2CHCl and the ionization energies (IEs) for B_2F_4 and B_2H_4 (Table 8).

Among these four values, based on the results of a number of high-level theoretical calculations, it has previously been suggested[25c] that the experimental $\Delta_f H_{298}$ values of C_2F_4 (−658.6 kJ mol^{-1}, versus −676.3 kJ mol^{-1} for W3X-L) and CH_2CHCl (37.2 kJ mol^{-1}, versus 21.3 kJ mol^{-1} for W3X-L) warrant re-examination. The large deviation for the IE of B_2F_4 has been attributed to structural changes upon ionization and their impact on the accuracy of the experimental value (1164.6 kJ mol^{-1} versus 1131.3 kJ mol^{-1} for W3X-L).[31] For the

Table 8. Energies (kJ mol^{-1}) for the Outliers in the G2/97 Set for Which the Deviations Between W3X-L and Designated Experimental Values are Larger than 10 kJ mol^{-1}

	expt	W3X-L	remarks
$\Delta_f H_{298}$ (C_2F_4)	−658.6	−676.3	−672.0 (G4, ref 25c)
$\Delta_f H_{298}$ (CH_2CHCl)	37.2	21.3	22.1 (G4, ref 25c)
IE (B_2F_4)	1164.6	1131.3	see ref 31
IE (B_2H_4)	935.9	923.0	

IE of B_2H_4 (935.9 kJ mol^{-1} from experiment and 923.0 kJ mol^{-1} for W3X-L), we have examined the original report for the experimental value.[32] We believe that the data, which were based on the onset for the appearance of $B_2H_4^+$ in the photo-electron ionization experiment, in reality suggests an IE that is somewhat smaller and closer to our W3X-L value. We thus believe that in these four cases of deviations between W3X-L and experiment of more than 10 kJ mol^{-1}, the W3X-L energies are likely to be the more accurate.

We also note that there are 11 instances in which the deviations lie between 6 and 10 kJ mol^{-1} (see Table S3). These include five $\Delta_f H$s, namely those of PH_2, C_2Cl_4, methylene-cyclopropane, thiooxirane, and CH_2CHCN. Five IEs (P_2, C_3H_7, CH_3CH_2OH, furan, and Si_2H_6) and one EA (CH_2NC) also fall into this category. It is likely that in many of these 11 cases, our W3X-L energies represent somewhat more accurate values than the quoted experimental ones. For example, there is a range of reported $\Delta_f H_{298}$ values for C_2Cl_4, and a recent reviewed estimate[33] of −24 kJ mol^{-1} is in better agreement with our W3X-L value (−21.1 kJ mol^{-1}) than with the designated G2/97 experimental value (−12.4 kJ mol^{-1}).

It is noteworthy that, when the B3-LYP/cc-pVTZ+d geometry and associated ZPVE of CH_4^+ are used to obtain the IE of CH_4, the resulting W3X-L value (1223.8 kJ mol^{-1}) shows a notable deviation from experiment (1216.7 kJ mol^{-1}). It has been pointed out previously[34] that B3-LYP does not give a satisfactory geometry for CH_4^+. Indeed, the W3X-L IE of CH_4 based on the CCSD(T)/VTZ geometry and scaled (by 0.9878)[16] ZPVE for CH_4^+ (1219.0 kJ mol^{-1}) is in significantly better agreement with the experimental value. We have thus included this in Table 7 instead of the value calculated with the standard B3-LYP/cc-pVTZ+d geometry for CH_4^+.

General Comments and Recommendations for the WnX-Type Procedures. This study, together with our previous investigations, has established a series of WnX procedures. They build on the foundation of the Wn-type protocols[1−3,5] of Martin and co-workers but use a combination of modern explicitly correlated coupled-cluster techniques, the less expensive MP2 method, and appropriately truncated basis sets to reduce the demand on computational resources.

The sequence W1X-2 → W1X-1 → W2X forms a series of progressively more accurate but computationally increasingly demanding approximations to all-electron scalar-relativistic CCSD(T)/CBS. As such, they provide a good means of obtaining relative energies for systems with dominant single-reference character. Their general uncertainties in this regard correspond to MADs from accurate values of roughly 3 kJ mol^{-1} for W1X-1 and W1X-2, with W1X-1 being somewhat more robust,[6] and approximately 2 kJ mol^{-1} for W2X. With our present computational resources, these methods can be applied to medium-sized systems such as C_6F_6 and fluoranthene (W1X-1 and W1X-2), benzophenone (W1X-2), and naphthalene (W2X).[35] We have previously formulated a diagnostic for the adequacy of

W1X-type procedures as approximations to CCSD(T)/CBS,[7] such that users would have a tool for judging the potential advantage of using W2X over W1X.

The W3X and W3X-L protocols are based on CCSD(T) energies obtained with W1X-1 and W2X, respectively, and additionally include post-CCSD(T) terms up to the CCSDT(Q) level. The W3X procedure is less expensive than W3X-L, and it is capable of tackling problems with moderate multireference character such as the automerization of cyclobutadiene. It mostly but not completely rectifies the inadequate description of F_2 dissociation by CCSD(T). In comparison, W3X-L more fully corrects for the shortcomings of CCSD(T) in its treatment of F_2 dissociation. The post-CCSD(T) calculations dominate the computational requirements for W3X and W3X-L for medium-sized systems. At present, with our current computational hardware, these methods are limited to treating systems approximately the size of benzene (W3X-L) and toluene (W3X). To this end, we have recommended the use of a diagnostic originally proposed by Martin and co-workers[2] as a rough guide to the importance of post-CCSD(T) effects.[7]

The set of WnX protocols, namely W1X-2,[6] W1X-1,[6] W2X, W3X,[7] and W3X-L, thus represents a series of cost-effective composite methods of increasing accuracy with MADs from benchmark values in the range of ∼1.5−3.0 kJ mol^{-1}. Our recommended diagnostics[7] can assist readers in choosing an appropriate WnX method for a particular problem.

CONCLUDING REMARKS

In the present study, we have built on our previous endeavors in putting together cost-effective Wn-type procedures to formulate the W2X and the W3X-L protocols. Together with W1X[6] and W3X,[7] the collection of WnX procedures provides a cost-effective means for the accurate treatment of medium-sized systems.

The W2X procedure provides an accurate approximation to the all-electron scalar-relativistic CCSD(T)/CBS energy, with an MAD of 0.6 kJ mol^{-1} for the W4-11 set against the corresponding benchmark CCSD(T) component energies in the W4 protocol. Such a performance is comparable to that of W2w (0.5 kJ mol^{-1}) and represents a notable improvement over the W1X-1 method, which has a corresponding MAD of 2.0 kJ mol^{-1}.

For the calculation of benzene, the W2X procedure is approximately three times more time-consuming than W1X-1, and it has a disk requirement (59.2 GB) that is larger than the 10.0 GB required for W1X-1. Nonetheless, these demands on the computational resources are still only comparable to those for W1w and are much lower than those for W2w, which is nearly 40 times more time-consuming than W1X-1 and has a disk requirement of 413.1 GB. The W2X procedure is applicable to systems that are somewhat larger than benzene, e.g., naphthalene. In particular, systems that can be treated by the W1w method are likely to be readily computed with W2X. We thus deem W2X a worthwhile upgrade to W1X-1 for medium-sized systems.

For the calculation of post-CCSD(T) effects, we use slightly larger basis sets in W3X-L than the ones used in W3X to yield an improved theoretical treatment. For the W4-11 set of W4 post-CCSD(T) energies, the more rigorous approach yields an MAD of 0.6 kJ mol^{-1}, which is close to that for the component used in W3.2 (0.7 kJ mol^{-1}) and indeed represents an improvement over the corresponding W3X value (1.0 kJ mol^{-1}). Comparison of the computational resources consumed shows that benzene may at present represent one of the largest systems that can be readily treated with our current com-

putational resources by the newly developed W3X-L approach to post-CCSD(T) effects, while the more economical W3X method is applicable to slightly larger species.

Our new W3-type protocol (W3X-L) combines W2X with the new post-CCSD(T) treatment. It has an MAD of 0.8 kJ mol^{-1} for the W4-11 set, which is comparable to that for the much more expensive W3.2 method (0.6 kJ mol^{-1}) and represents a notable improvement over W3X (1.9 kJ mol^{-1}). For the dissociation of F$_2$, for which W3X shows non-negligible deviations from the accurate potential energy curve, W3X-L remedies the problem and has a dissociation curve that is virtually indistinguishable from that of W3.2. We have applied W3X-L to the G2/97 set of thermochemical properties and have estimated that these values have an MAD from accurate energies of approximately 1.5 kJ mol^{-1}. This deviation comes both from the single-point energy component and from the use of B3-LYP zero-point vibrational energies and thermal corrections. Our W3X-L energies represent the highest-level theoretical values currently available for the full G2/97 set of thermochemical quantities.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Instructions for carrying out W2X and W3X-L calculations, sample Molpro input and output files, incorporated in a single zip archive, total electronic energies for Wn-type procedures for the W4-11 set (Table S1), deviations from W4 benchmark values from ref 13 (Table S2), components of W2X and W3X-L energies for the G2/97 set corresponding to the W3X-L values in Table 7 together with experimental energies, as well as theoretical atomization energies (Table S3), and experimental atomic $\Delta_f H$ values employed in the present study (Table S4). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors
*E-mail: bun.chan@sydney.edu.au.
*E-mail: radom@chem.usyd.edu.au.

Notes
The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Martin, J. M. L.; de Oliveira, G. *J. Chem. Phys.* **1999**, *111*, 1843−1856.

(2) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kállay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 4129−4141.

(3) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108−1−17.

(4) Ten-no, S.; Noga, J. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 114−125.

(5) Karton, A.; Martin, J. M. L. *J. Chem. Phys.* **2012**, *136*, 124114−1−12.

(6) Chan, B.; Radom, L. *J. Chem. Theory Comput.* **2012**, *8*, 4259−4269.

(7) Chan, B.; Radom, L. *J. Chem. Theory Comput.* **2013**, *9*, 4769−4778.

(8) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision C.01*; Gaussian, Inc.: Wallingford, CT, 2009.

(9) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 242−253.

(10) Neese, F. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73−78.

(11) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243−252.

(12) (a) Chan, B.; Radom, L. *J. Chem. Theory Comput.* **2011**, *7*, 2852−2863. (b) Chan, B.; Radom, L. *Theor. Chem. Acc.* **2014**, *133*, 1426−1−10.

(13) Karton, A.; Daon, S.; Martin, J. M. L. *Chem. Phys. Lett.* **2011**, *510*, 165−178.

(14) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221−7230.

(15) Karton, A.; Yu, L.-J.; Kesharwani, M. K.; Martin, J. M. L. *Theor. Chem. Acc.* **2014**, *133*, 1483−1−15.

(16) Merrick, J. P.; Moran, D.; Radom, L. *J. Phys. Chem. A* **2007**, *111*, 11683−11700.

(17) (a) Adler, T. B.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2007**, *127*, 221106−1−4. (b) Knizia, G.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2009**, *130*, 054104−1−20.

(18) Peterson, K. A.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2008**, *128*, 084102−1−12.

(19) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(20) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796−6806.

(21) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1995**, *103*, 4572−4585.

(22) Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2009**, *131*, 194105−1−13.

(23) Wennmohs, F.; Neese, F. *Chem. Phys.* **2008**, *29*, 217−230.

(24) We have also investigated the use of MP2 with the standard CVTZ basis set in our composite scheme for Δ(C+R). However, we find that the improvement over the use of CCSD(T)/CVDZ(2d) alone is notably less substantial with this approach.

(25) (a) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *126*, 124105-1−8. (b) Chan, B.; Coote, M. L.; Radom, L. *J. Chem. Theory Comput.* **2010**, *6*, 2647−2653. (c) Chan, B.; Deng, J.; Radom, L. *J. Chem. Theory Comput.* **2011**, *7*, 112−120.

(26) Chan, B.; Karton, A.; Raghavachari, K.; Radom, L. *J. Chem. Theory Comput.* **2012**, *8*, 3159−3166.

(27) Karton, A.; Kaminker, I.; Martin, J. M. L. *J. Phys. Chem. A* **2009**, *113*, 7610−7620.

(28) (a) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063−1079. (b) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 42−55.

(29) Geary, R. C. *Biometrika* **1935**, *27*, 310−332.

(30) Haunschild, R.; Klopper, W. *J. Chem. Phys.* **2012**, *136*, 164102-1−7. We note that the frozen-core vibrationless atomization energies given in this study are in fair agreement with the corresponding component energies of our W3X-L values, with an MAD of 1.6 kJ mol^{-1}, see the Supporting Information.

(31) Chan, B.; Trevitt, A.; Blanksby, S. J.; Radom, L. *J. Phys. Chem. A* **2012**, *116*, 9214−9215.

(32) Ruscic, B.; Schwarz, M.; Berkowitz, J. *J. Chem. Phys.* **1989**, *91*, 4576−4581.

(33) Manion, J. A. *J. Phys. Chem. Ref. Data* **2002**, *31*, 123−172.

(34) Baboul, A. G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **1999**, *110*, 7650−7657.

(35) Our test W2X calculation on naphthalene took 27 h of computer time and consumed 320 GB of scratch disk space under the same conditions as for the calculations shown in Table 6.