Achieving the CCSD(T) Basis-Set Limit in Sizable Molecular Clusters: Counterpoise Corrections for the Many-Body Expansion

Ryan M. Richard, Ka Un Lao, and John M. Herbert*

Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States

S *Supporting Information*

ABSTRACT: An efficient procedure is introduced to obtain the basis-set limit in electronic structure calculations of large molecular and ionic clusters. This approach is based on a Boys–Bernardi-style counterpoise correction for clusters containing arbitrarily many monomer units, which is rendered computationally feasible by means of a truncated many-body expansion. This affords a tractable way to apply the sequence of correlation-consistent basis sets (aug-cc-pVXZ) to large systems and thereby obtain energies extrapolated to the complete basis set (CBS) limit. A three-body expansion with three-body counterpoise corrections is shown to afford errors of $\lesssim 0.1-0.2$ kcal/mol with respect to traditional MP2/CBS results, even for challenging systems such as fluoride–water clusters. A triples correction, $\delta_{CCSD(T)} = E_{CCSD(T)} - E_{MP2}$, can be estimated accurately and efficiently as well. Because the procedure is embarrassingly parallelizable and requires no electronic structure calculations in systems larger than trimers, it is extendible to very large clusters. As compared to traditional CBS extrapolations, computational time is dramatically reduced even without parallelization.

SECTION: Molecular Structure, Quantum Chemistry, and General Theory
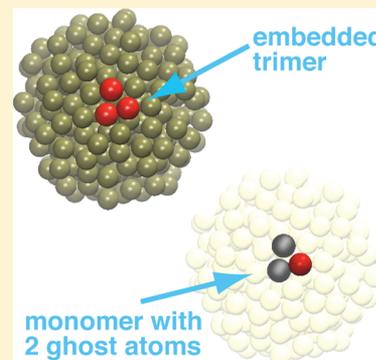
For electronic structure calculations of dimer binding energies (BEs), it is well-known that naïve attempts to use the formula

$$BE = E_{AB} - E_A - E_B \qquad (1)$$

usually result in significant overestimation of the $A\cdots B$ binding energy, owing to the basis-set superposition error (BSSE) that disappears only very slowly as the monomer basis sets approach completeness. BSSE is a manifestation of the fact that eq 1 is an unbalanced approximation in the sense that E_{AB} is computed using a more flexible basis set than is used to compute the monomer energies. The widely used Boys–Bernardi counterpoise (CP) procedure[1] corrects this imbalance by computing all three energies using the dimer (AB) basis set. The obvious generalization of this CP correction to N-body clusters is[2–4]

$$BE = E_{IJK\cdots} - \sum_{i=I,J,K,\cdots}^{N} E_i^{IJK\cdots} \qquad (2)$$

where $E_i^{IJK\cdots}$ denotes the energy of monomer i computed using a basis set with functions on monomers $I,J,K,\cdots$ Operationally, this means placing "ghost atoms" on all monomers $J \neq i$.

Although the generalized CP correction in eq 2 has sometimes been criticized for neglecting "basis-set extension" effects,[5] which are further discussed below, we will show that this approach correctly reproduces known complete basis set (CBS) energetics if the individual energies are evaluated in large basis sets and extrapolated to the CBS limit. We will furthermore demonstrate how the cost of such calculations can be dramatically reduced, without substantial loss of accuracy, by

consistent application of a truncated many-body expansion (MBE) to each of the energies in eq 2. We thereby aim to reproduce the whole sequence of CP-corrected MP2/aug-cc-pVXZ energies (abbreviated MP2/aXZ, for X = D, T, or Q) and thereby perform MP2/CBS extrapolations in a small fraction of the computer time that is required for traditional calculations. Finally, we will show that a triples correction

$$\delta_{CCSD(T)} = E_{CCSD(T)} - E_{MP2} \qquad (3)$$

can also be accurately approximated by means of truncated MBEs. Together, these techniques extend well-established techniques for estimating CCSD(T)/CBS-quality energetics to clusters that are far larger than what was previously possible.

Equation 2 nominally requires $N + 1$ calculations in the full N-body cluster basis set. Our strategy is to replace each individual energy in this equation with a MBE and then apply a consistent truncation of each. For the cluster energy $E_{IJK\cdots}$, this looks like a traditional MBE

$$E_{IJK\cdots} = \sum_{i=1}^{N} E_i + \sum_{i=1}^{N}\sum_{j>i}^{N} \Delta E_{ij} + \sum_{i=1}^{N}\sum_{j>i}^{N}\sum_{k>j}^{N} \Delta E_{ijk} + \cdots \qquad (4)$$

where

$$\Delta E_{ij} = E_{ij} - E_i - E_j \qquad (5a)$$

$$\Delta E_{ijk} = E_{ijk} - E_{ij} - E_{ik} - E_{jk} + E_i + E_j + E_k \qquad (5b)$$

For the monomer energies $E_i^{IJK\cdots}$, however, some of the "bodies" $I, J, K,\dots$ are simply ghost atoms, and the energy of any term in the MBE that contains only ghost atoms is zero. The nonvanishing terms can be written

$$E_I^{IJ\cdots N} = E_I^I + \Delta E_I^{(2)} + \Delta E_I^{(3)} + \cdots \qquad (6)$$

where $\Delta E_I^{(2)}$ and $\Delta E_I^{(3)}$ represent the nonvanishing two- and three-body terms in the MBE of $E_I^{IJ\cdots N}$. After rearranging some terms, one obtains the following explicit formulas for these corrections:

$$\Delta E_I^{(2)} = -(N - 1)E_I^I + \sum_{J \neq I}^{N} E_I^{IJ} \qquad (7a)$$

$$\Delta E_I^{(3)} = \frac{1}{2}(N - 1)(N - 2)E_I^I - (N - 2)\sum_{J \neq I}^{N} E_I^{IJ}$$
$$+ \sum_{J \neq I}^{N} \sum_{\substack{K > J \\ K \neq I}}^{N} E_I^{IJK} \qquad (7b)$$

Here, $E_I^{IJK\cdots}$ denotes the energy of monomer I, computed in the $IJK\cdots$ basis, which therefore contains ghost atoms in place of monomers $J, K, \cdots$

The basic idea of this many-body counterpoise correction is illustrated in Figure 1, which provides a pictorial representation
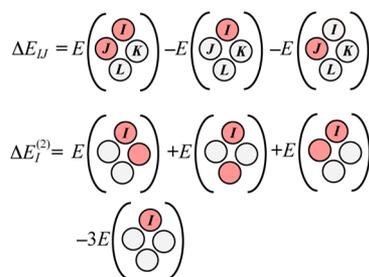

Figure 1. Illustration of one particular two-body term (ΔE_{IJ}, eq 5a) that contributes to the binding energy of cluster $IJKL$ and also a two-body term $\Delta E_I^{(2)}$ that appears in the monomer energy E_I^{IJKL} (eq 7a). Monomers labeled with letters are actual molecules containing electrons, whereas unlabeled monomers are comprised of ghost atoms. Monomers shaded red are included in the calculation, whereas monomers shaded gray are not.

of some two-body contributions to the binding energy of a cluster $IJKL$. The term ΔE_{IJ} is a traditional two-body energy correction, where dimer IJ is computed using the IJ basis set, and an estimate of the binding energy is obtained by subtracting the monomer energies, computed using their respective monomer basis sets. The term $\Delta E_I^{(2)}$ in the lower portion of Figure 1 represents a two-body correction for the energy of monomer I computed in the $IJKL$ basis. This term consists of three calculations involving a dimer basis with one ghost monomer, along with a one-body calculation in a one-body basis set.

An n-body, CP-corrected approximation to the binding energy is obtained as the difference between the cluster energy (eq 4) and the monomer energies $E_I^{IJK\cdots}$ (eq 6), each truncated at the n-body level. We refer to this procedure as an nth-order many-body CP correction, MBCP(n), and its aim is to

reproduce the CP-corrected binding energy in eq 2, albeit at significantly reduced cost. As such, extrapolation is still required in order to obtain CBS-quality energetics, but this can be accomplished by traditional means insofar as a high-quality approximation to eq 2 can be obtained for a whole sequence of correlation-consistent basis sets. It is worth noting that alternative fragment-based approaches such as the fragment molecular orbital (FMO) method[6] appear to work best in small basis sets[7] and therefore cannot be exploited for what we aim to accomplish here.

The CP correction that is embodied by eq 2 has been criticized[5] for neglecting so-called "basis-set extension" effects, which are best defined in the context of the projection-operator formalism of the chemical Hamiltonian approach[8] but that should disappear in the CBS limit. An alternative CP correction that directly addresses basis-set extension effects has been proposed by Valiron and Mayer,[5] but the number of independent electronic structure calculations required for this Valiron–Mayer function counterpoise (VMFC) approach grows factorially with the number of monomers. Truncated two- and three-body versions, which we will call VMFC(2) and VMFC(3), were introduced later.[9,10]

In fact, VMFC(2) is completely equivalent to MBCP(2), though the two approaches differ starting at $n = 3$. For $n \geq 3$, our MBCP(n) approach requires fewer independent electronic structure calculations; at the $n = 3$ level, for example, MBCP(3) requires only monomer calculations in the trimer basis set, whereas VMFC(3) requires both monomer and dimer calculations in the trimer basis. For a cluster with $N = 11$ monomer units, as in some of the numerical examples that follow, the VMFC(2)/MBCP(2) method requires 110 distinct electronic structure calculations, whereas three-body approaches require an additional 990 calculations for VMFC(3) but only half of that number for MBCP(3).

Before considering CBS extrapolations, which is the main topic of this Letter, we first consider a numerical comparison of the VMFC(n) and MBCP(n) schemes. Figure 2 shows the binding energy errors for various two- and three-body methods applied to two different systems, $F^-(H_2O)_{10}$ and $Gly(H_2O)_{10}$, where "Gly" denotes the zwitterionic tautomer of glycine.
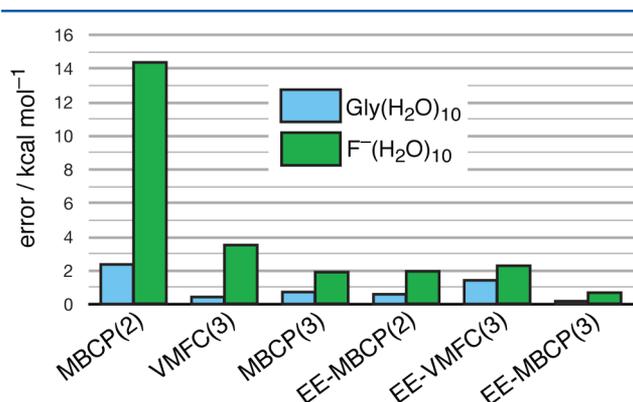

Figure 2. Comparison of two- and three-body counterpoise corrections, with and without electrostatic embedding (EE), for MP2/aDZ calculations on two different systems. Errors are reported with respect to a calculation performed on the entire cluster, for which the CP-corrected binding energy is defined by eq 2. To better resemble the self-consistent charge-embedding used in the VMFC(n) calculations in ref 10, embedding charges are computed using the self-consistent XPol-CHELPG procedure.[16,17]

(These are challenging test cases owing to large many-body polarization effects, and we have previously used them to test a variety of fragment-based approaches.[11−16]) Ten different isomers of both systems were used to generate error statistics, and "error" is defined in this work as deviation from the binding energy computed using eq 2.

Some of the monomer-based methods considered in Figure 2 utilize electrostatic embedding (EE) for the n-body (dimer and trimer) electronic structure calculations, and in the VMFC(n) calculations originally reported in ref 10, this was done in a sophisticated way based on charges and point dipoles that were iterated to self-consistency. In the interest of making a fair comparison, the EE results in Figure 2 also employ self-consistent embedding charges, determined using our CHELPG-based version[16,17] of the "explicit polarization" (XPol) procedure.[18]

Examining the MP2/aDZ results in Figure 2, we see that the two-body method is simply unacceptable unless electrostatic embedding is employed, consistent with the results of MBEs in which no attempt at CP correction is made.[19] The difference here is that CP-corrected MP2/aDZ is an acceptable (if qualitative) level of theory for clusters, whereas small-basis MP2 calculations with no correction for BSSE whatsoever should not be trusted. The fact that fragment-based methods can accurately approximate supersystem energies in clusters is largely meaningless unless BSSE is addressed by means of large basis sets and/or CP corrections.

The VMFC(3) method, with or without electrostatic embedding, exhibits larger errors for the very challenging $F^-(H_2O)_{10}$ system, as compared to MBCP(3), and the EE-MBCP(3) method appears quite promising. For this reason, our primary focus will be on extending EE-MBCP(n) methods to larger basis sets and evaluating their performance in CBS extrapolations, although a few additional results using the VMFC(n) approach will be considered below, for comparative purposes.

Electrostatically embedded MBEs are often remarkably insensitive to the details of the embedding scheme,[20] at least for clusters composed of small, polar monomers, and in view of this observation, the remaining calculations presented herein employ a very simple embedding in which those monomer units that are not included in the electronic structure calculation are replaced with gas-phase Mulliken charges computed at the B3LYP/6-31G* level. (Ghost atoms do not contribute embedding charges.) This approach is admittedly more simplistic than a self-consistent embedding procedure, and it is possible that a more sophisticated embedding might improve the EE-VMFC(n) results. In view of the favorable performance of EE-MBCP(n) with this simple embedding, however, we have not considered such alternatives.

Figure 3 shows the convergence toward the basis set limit of MP2/aXZ binding energies for the "bag" isomer of $(H_2O)_6$. (Analogous plots for seven other isomers can be found in the Supporting Information, but the behavior of the methods discussed here is very similar in each case.) Extrapolating to the CBS limit using standard techniques (see the Computational Methods section for details), we obtain the same binding energy, to four significant digits, that was previously reported[21] as the MP2/CBS limit, based on MP2-R12 calculations in a very large basis (modified aug-cc-pV5Z). CBS extrapolation using aug-cc-pVXZ through $X = 5$ agrees with these results as well,[22] which gives us confidence that the MP2/CBS binding energy has been accurately determined for this system. Results
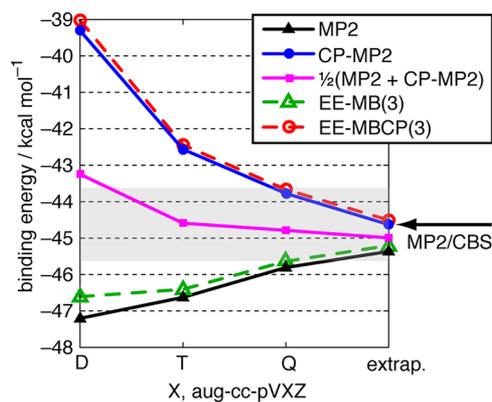

Figure 3. Convergence to the CBS limit for the bag isomer[21] of $(H_2O)_6$. CP-MP2 denotes traditional counterpoise-corrected MP2 (eq 2), whereas EE-MBCP(3) is a three-body approximation to CP-MP2. The EE-MB(3) method is a three-body approximation to the uncorrected MP2 energy. The shaded region delineates ± 1 kcal/mol of the MP2/CBS binding energy.

in Figure 3 demonstrate that we can accurately approximate this limit via a low-order MBE, provided that CP corrections are included.

When attempting to compute cluster binding energies in the CBS limit, it is commonly assumed that the CP-corrected and uncorrected results place error bars on the value that would be obtained in the CBS limit.[22,23] For this $(H_2O)_6$ example, we do obtain usable error bars, but even at the quadruple-ζ level, these bounds are no better than ± 1 kcal/mol. On the other hand, the average of the CP-corrected and uncorrected binding energies affords a quite accurate estimate of the CBS limit at the triple-ζ level (<0.04 kcal/mol error). The same cannot be said at the double-ζ level, where even this average is in error by 1.4 kcal/mol. This is important to consider in the context of fragment-based electronic structure methods, which have often[10,19,24−27] (though not always[28−34]) been performed using double-ζ basis sets.

Here, we observe monotonic convergence of the many-body approximations to the binding energy, as a function of X (= D,T,Q). The EE-MBCP(3) method affords a very accurate approximation to the CP-corrected MP2 binding energy in each basis set (errors < 0.3 kcal/mol and improving with basis size) and can be used on its own to obtain the MP2/CBS result (error = 0.12 kcal/mol). Better still, having done that calculation, one has all of the information required for a more traditional three-body approximation *without* CP correction. This method, which we call EE-MB(3), affords a highly accurate approximation to the *uncorrected* MP2 binding energy (except at the double-ζ level, where the error is 0.8 kcal/mol). Together, the EE-MBCP(3) and EE-MB(3) methods can be used to bracket the CBS limit in the usual way but without the need for any electronic structure calculations on systems larger than trimers.

Furthermore, these MP2/CBS binding energies can be extended to CCSD(T)/CBS quality by means of a MBE applied to the difference $E_{CCSD(T)} - E_{MP2}$, eq 3. This difference converges more rapidly with respect to basis expansion than do either of the individual energies,[35−38] and basis sets of triple-ζ quality are sufficient to converge this correction term. In the case of $(H_2O)_6$, the haTZ basis set (aTZ without diffuse functions on hydrogen) has been used for this purpose, and

values of $\delta_{CCSD(T)}$ for eight different isomers are reported in ref 21.

Table 1 compares these benchmark values to corrections obtained based on a two- or a three-body approximation to

Table 1. Triples Corrections, $\delta_{CCSD(T)}$, for Isomers of $(H_2O)_6$ and Errors Engendered by Low-Order Many-Body Approximations to $\delta_{CCSD(T)}$[a]

isomer	$\delta_{CCSD(T)}$[b]	error[c]	
		EE-MB(2)	EE-MB(3)
bag	0.33	−0.08	0.00
boat1	0.53	0.04	0.01
boat2	0.51	0.04	0.01
book1	0.33	−0.04	0.00
book2	0.33	−0.04	0.01
cage	0.13	−0.20	0.00
cyclic	0.53	0.05	0.01
prism	−0.06	−0.20	0.00
MUE[d]		0.09	0.00

[a]All values in kcal/mol. [b]Values obtained from CCSD(T)/haTZ calculations in ref 21. [c]Errors engendered by two- and three-body CCSD(T)/haTZ calculations. [d]Mean unsigned error for all eight isomers.

$\delta_{CCSD(T)}$. Both approximations prove to be quite accurate, with mean unsigned errors (MUEs) of only 0.09 kcal/mol for the EE-MB(2) approximation and <0.01 kcal/mol for the EE-MB(3) approximation, with respect to an exact calculation of $\delta_{CCSD(T)}$ requiring CCSD(T)/haTZ calculations on the full hexamer. In terms of accuracy, the EE-MBCP(3) method, applied to approximate the sequence of MP2/aXZ energies and thereby to extrapolate the MP2/CBS limit, augmented with a two-body CCSD(T) approximation, affords results that differ by <0.1 kcal/mol as compared to the best existing estimates of the CCSD(T)/CBS binding energies for isomers of $(H_2O)_6$.

It is certainly true that the triples correction for $(H_2O)_6$ is small relative to values obtained for some dispersion-dominated systems. An example of the latter is the benzene dimer, for which $\delta_{CCSD(T)} \approx 2.0-2.5$ kcal/mol.[39] On the other hand, three-body effects are no larger than 0.33 kcal/mol for any isomer of the benzene trimer[40] and are more often an order-of-magnitude smaller than that.[40,41] In acetylene clusters, the triples correction is not much larger than it is for the water hexamer.[33,42] In view of the general success of two-body approximations to the correlation energy,[24,29,30,43] we expect the combination of fragment-based and CP-corrected MP2/CBS extrapolations, in conjunction with a two-body CCSD(T) correction, to be a robust and accurate approach.

Moreover, this approach requires only a fraction of the computer time that is required for traditional supersystem calculations. Considering only the MP2 calculations, Figure 4 shows the total aggregate computer time required to perform CP-corrected MP2/aXZ calculations for the $F^-(H_2O)_{10}$ system, for X = D, T, and Q. For X = Q, the 12 different calculations required by eq 2 consume a total of 87 days of wall time. Setting aside, for the moment, the fact that the MBCP(n) calculations are embarrassingly parallelizable and instead focusing simply on the total wall time required for all of the individual calculations, we find that the MBCP(3) approach reduces the total time for X = Q by more than 70%, and MBCP(2) reduces it by more than 98%! Moreover, no single electronic structure calculation is more expensive than 1737 s
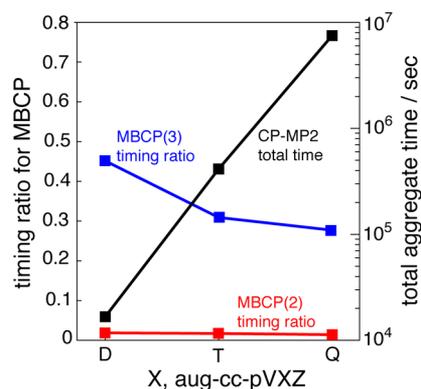

Figure 4. Timing data for RI-MP2/aXZ single-point calculations on $F^-(H_2O)_{10}$. The total aggregate wall time required for a CP-MP2 calculation is shown on the logarithmic scale at the right, whereas the linear axis on the left shows the ratio of the MBCP(2) or MBCP(3) time to the CP-MP2 time, which indicates the speed-up for the fragment-based calculations. All calculations were run in serial on a single 2.5 GHz Opteron processor, meaning that we have *not* taken advantage of the embarrassingly parallel nature of the MBCP(n) methods.

[MBCP(2)] or 8510 s [MBCP(3)]; therefore, the total waiting time can be made as short as that, given a sufficient number of processors. (For N = 11 monomers, that "sufficient number" is 44S, which nowadays does not even constitute "massively parallel".)

To test the robustness of these procedures, we examine errors (with respect to MP2/CBS benchmarks[21]) for eight isomers of $(H_2O)_6$. Figure 5a shows error statistics obtained using a variety of MBE-based methods to approximate the MP2/aXZ energies, which are then extrapolated to the CBS limit and compared to the benchmarks to determine the error. (Several of these methods represent "control experiments" designed to understand the limits of the methodology rather than practical procedures intended for production calculations, as will become clear in the discussion below.)

Consistent with the results for the bag isomer in Figure 3, where extrapolation of *uncorrected* MP2 binding energies does not afford the correct MP2/CBS limit, we observe MUEs > 0.4 kcal/mol for the EE-MB(2) and EE-MB(3) methods that do not employ CP correction. In fact, the three-body approximation actually affords the larger error, owing to fortuitous error cancellation in the two-body case. Specifically, three-body effects contribute ∼10 kcal/mol to the binding energy of $(H_2O)_6$; therefore, a two-body approximation would underbind by about this much except that BSSE overstabilizes the cluster by approximately the same amount. This error cancellation is also the reason why CP correction slightly *increases* the mean error at the two-body level, but when the important three-body interactions are included, CP correction serves to significantly *decrease* the mean error. Together, these results demonstrate the importance of CP correction even when CBS extrapolation is employed, as is also seen in the basis-set convergence behavior shown in Figure 3, which is smoother for the CP-corrected results than for the uncorrected ones.

As another experiment, we use an n-body [EE-MB(n)] approximation to the cluster energy and subtract from this a "full" CP correction, in which each monomer energy is computed using the entire cluster basis set. (In other words, we use an n-body approximation for $E_{IJK...}$ in eq 2 but not for $E_i^{IJK...}$.) We consider this to be an unbalanced approximation
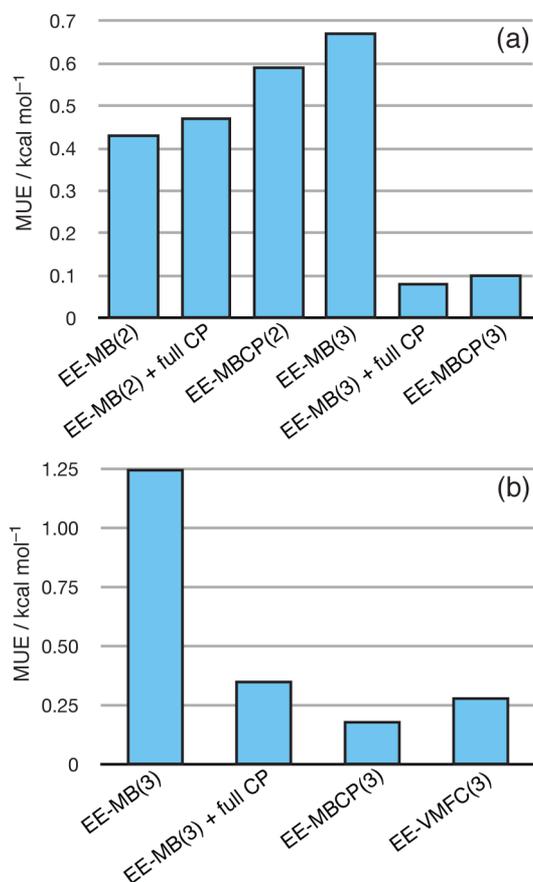
Figure 5. MUEs with respect to MP2/CBS binding energies for (a) 8 isomers of $(H_2O)_6$ and (b) 10 isomers of $F^-(H_2O)_{10}$, computed by basis-set extrapolation using an assortment of low-cost many-body methods to approximate the MP2/aXZ energies (X = D, T, Q). For $(H_2O)_6$, the benchmark MP2/CBS results are taken from ref 21. The "full CP" correction means that the entire cluster basis set is used to compute the monomer energies.

and one that is rather expensive as compared to EE-MBCP(n), but we include it here simply to demonstrate that such a correction cannot rescue the $n = 2$ calculations from errors > 0.4 kcal/mol. On the other hand, a balanced, two-body approximation to eq 2, that is, the EE-MBCP(2) method, actually affords slightly larger errors as compared to this unbalanced "full CP" correction.

To achieve an accuracy of ≲0.1 kcal/mol, three-body terms are evidently required, as has been found in previous studies where CP-corrected two- and three-body terms have been used in a MBE.[44−46] For the EE-MBCP(3) method, the MUE is an acceptable 0.1 kcal/mol. A "full" CP correction reduces this MUE only very slightly (and not in all cases, as shown below) and does not justify the greatly increased cost of adding a cluster's worth of extra virtual orbitals to each monomer calculation.

In Figure 5b, we consider a larger and more challenging test case, namely, 10 isomers of $F^-(H_2O)_{10}$, where binding energies for our unrelaxed structures are on the order of −115 kcal/mol and span a range of 10 kcal/mol. Here, the performance of EE-MBCP(3) exceeds even that of EE-MB(3) with a full CP correction—possibly owing the unbalanced nature of the latter approach—and affords a MUE of 0.18 kcal/mol with respect to CP-corrected MP2/CBS benchmarks. The EE-VMFC(3) approach affords a MUE that is only slightly larger, 0.28

kcal/mol, demonstrating that any "basis extension" effects[5,8] still present in the CP-corrected binding energy defined by eq 2 must be quite small. The MUE for the relative isomer energies, computed at the EE-MBCP(3) level, is 0.24 kcal/mol, which demonstrates that not only total binding energies but also relative conformational energies are predicted accurately by this low-cost approximation.

In summary, we have shown that the EE-MBCP(3) method, which incorporates counterpoise corrections in a simple way and requires electronic structure calculations on subsystems no larger than trimers, can be used to approximate the sequence of CP-corrected MP2/aug-cc-pVXZ calculations (X = D,T,Q). The MP2/CBS limit can thus be obtained (via traditional extrapolation techniques) at low cost, even in challenging systems such as $F^-(H_2O)_{10}$. This many-body counterpoise correction may also be useful in studies of noncovalent interactions in even larger systems, where no attempt may be made to extrapolate to the CBS limit but where CP correction is nevertheless useful for avoiding large artifacts in potential energy surfaces that are caused by BSSE.[47−51]

Total computational time is dramatically reduced by our EE-MBCP(3) procedure, even if no parallelism is exploited at all. Owing to the embarrassingly-parallel nature of the truncated MBE, wall times can be reduced even further by farming out the subsystem calculations to a large number of processors. For rough screening of the energy landscape at a slightly lower level of accuracy, two-body EE-MBCP(2) calculations can be performed at a total cost that is dramatically reduced even in comparison to the three-body method.

In conjunction with *two*-body CCSD(T) calculations in a triple-ζ basis set, binding energies within ≲0.2 kcal/mol of CCSD(T)/CBS benchmarks are obtained for hydrogen-bonded clusters whose total binding energies exceed 40 kcal/mol, which represents an error of <0.5%. Previous work at much lower levels of theory (e.g., B3LYP/double-ζ) has shown that we can reproduce lower-level binding energy benchmarks quite accurately in systems as large as $(H_2O)_{57}$,[11−13] which makes us optimistic about our ability to scale up the techniques presented here to significantly larger cluster sizes. This paves the way to obtaining CCSD(T)/CBS-quality energetics in clusters of hitherto unthinkable size, as well as to geometry optimizations on BSSE-free potential energy surfaces. Such applications, along with a fuller characterization of the accuracy of the methodology in large systems, are currently underway in our group.

■ COMPUTATIONAL METHODS

For $(H_2O)_6$, we use the MP2/haTZ geometries from ref 21, whereas $Gly(H_2O)_{10}$ and $F^-(H_2O)_{10}$ structures are taken from our own simulations, using the AMOEBA force field,[52] and were not relaxed. Only valence electrons were correlated in MP2 and CCSD(T) calculations, and MP2 calculations for $Gly(H_2O)_{10}$ and $F^-(H_2O)_{10}$ employed the resolution-of-identity (RI) approximation, in conjunction with standard auxiliary basis sets.[53] For CBS extrapolations, we first extrapolate the Hartree−Fock energy using a three-point (D, T, Q) exponential ansatz[54] and then use a two-point (T, Q) extrapolation of the MP2 correlation energy, using an X^{-3} ansatz.[55] All electronic structure calculations were performed using Q-Chem.[56,57] A home-built code, Fragm∩t (as described in refs 12 and 13) was used as a driver.

■ ASSOCIATED CONTENT

S **Supporting Information**

Additional numerical data for the systems considered here and binding enery convergence plots. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: herbert@chemistry.ohio-state.edu.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Boys, S. F.; Bernardi, F. The Calculation of Small Molecular Interactions by the Differences of Separated Total Energies. Some Procedures with Reduced Errors. *Mol. Phys.* **1970**, *19*, 553−566.

(2) Turi, L.; Dannenberg, J. J. Correcting for Basis Set Superposition Error in Aggregates Containing More than Two Molecules: Ambiguities in the Calculation of Counterpoise Correction. *J. Phys. Chem.* **1993**, *97*, 2488−2490.

(3) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. State of the Art in Counterpoise Theory. *Chem. Rev.* **1994**, *94*, 1873−1885.

(4) Tschumper, G. S. Reliable Electronic Structure Computations for Weak Noncovalent Interactions in Clusters. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Cundari, T. R., Eds.; Wiley-VCH: New York, 2009; Vol. 26, Chapter 2, pp 39−90.

(5) Valiron, P.; Mayer, I. Hierarchy of Counterpoise Corrections for *N*-Body Clusters: Generalization of the Boys−Bernardi Scheme. *Chem. Phys. Lett.* **1997**, *275*, 46−55.

(6) Fedorov, D. G.; Kitaura, K. Extending the Power of Quantum Chemistry to Large Systems with the Fragment Molecular Orbital Method. *J. Phys. Chem. A* **2007**, *111*, 6904−6914.

(7) Fedorov, D. G.; Kitaura, K. Coupled-Cluster Theory Based upon the Fragment Molecular Orbital Method. *J. Chem. Phys.* **2005**, *123*, 134103/1−134103/11.

(8) Mayer, I. Towards a "Chemical" Hamiltonian. *Int. J. Quantum Chem.* **1983**, *23*, 341−363.

(9) Salvador, P.; Szczęśniak, M. M. Counterpoise-Corrected Geometries and Harmonic Frequencies of *N*-Body Clusters: Application to HF$_n$ (*n* = 3, 4). *J. Chem. Phys.* **2003**, *118*, 537−549.

(10) Kamiya, M.; Hirata, S.; Valiev, M. Fast Electron Correlation Methods for Molecular Clusters without Basis Set Superposition Errors. *J. Chem. Phys.* **2008**, *128*, 074103/1−074103/11.

(11) Jacobson, L. D.; Richard, R. M.; Lao, K. U.; Herbert, J. M. Efficient Monomer-Based Quantum Chemistry Methods for Molecular and Ionic Clusters. *Annu. Rep. Comput. Chem.* **2013**, in press.

(12) Richard, R. M.; Herbert, J. M. A Generalized Many-Body Expansion and a Unified View of Fragment-Based Methods in Quantum Chemistry. *J. Chem. Phys.* **2012**, *137*, 064113:1−17.

(13) Richard, R. M.; Herbert, J. M. Many-Body Expansion with Overlapping Fragments: Analysis of Two Approaches. *J. Chem. Theory Comput.* **2013**, *9*, 1408−1416.

(14) Lao, K. U.; Herbert, J. M. Breakdown of the Single-Exchange Approximation in Third-Order Symmetry-Adapted Perturbation Theory. *J. Phys. Chem. A* **2012**, *116*, 3042−3047.

(15) Lao, K. U.; Herbert, J. M. An Improved Treatment of Empirical Dispersion and a Many-Body Energy Decomposition Scheme for the Explicit Polarization Plus Symmetry-Adapted Perturbation Theory (XSAPT) Method. *J. Chem. Phys.* **2013**, *139*, 034107/1−034107/16.

(16) Herbert, J. M.; Jacobson, L. D.; Lao, K. U.; Rohrdanz, M. A. Rapid Computation of Intermolecular Interactions in Molecular and Ionic Clusters: Self-Consistent Polarization Plus Symmetry-Adapted Perturbation Theory. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7679−7699.

(17) Jacobson, L. D.; Herbert, J. M. An Efficient, Fragment-Based Electronic Structure Method for Molecular Systems: Self-Consistent Polarization with Perturbative Two-Body Exchange and Dispersion. *J. Chem. Phys.* **2011**, *134*, 094118/1−094118/17.

(18) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. The Variational Explicit Polarization Potential and Analytical First Derivative of Energy: Towards a Next Generation Force Field. *J. Chem. Phys.* **2008**, *128*, 234108/1−234108/9.

(19) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Expansion for Large Systems, with Applications to Water Clusters. *J. Chem. Theory Comput.* **2007**, *3*, 46−53.

(20) Leverentz, H. R.; Truhlar, D. G. Electrostatically Embedded Many-Body Approximation for Systems of Water, Ammonia, and Sulfuric Acid and the Dependence of Its Performance on Embedding Charges. *J. Chem. Theory Comput.* **2009**, *5*, 1573−1584.

(21) Bates, D. M.; Tschumper, G. S. CCSD(T) Complete Basis Set Limit Relative Energies for Low-Lying Water Hexamer Structures. *J. Phys. Chem. A* **2009**, *113*, 3555−3559.

(22) Xantheas, S. S.; Burnham, C. J.; Harrison, R. J. Development of Transferable Interaction Models for Water. II. Accurate Energetics of the First Few Water Clusters from First Principles. *J. Chem. Phys.* **2002**, *116*, 1493−1499.

(23) Clarkson, J. R.; Herbert, J. M.; Zwier, T. S. Infrared Photodissociation of a Water Molecule from a Flexible Molecule−H$_2$O Complex: Rates and Conformational Product Yields Following XH Stretch Excitation. *J. Chem. Phys.* **2007**, *126*, 134306/1−134306/15.

(24) Dahlke, E. E.; Truhlar, D. G. Electrostatically Embedded Many-Body Correlation Energy, with Applications to the Calculation of Accurate Second-Order Møller−Plesset Perturbation Theory Energies for Large Water Clusters. *J. Chem. Theory Comput.* **2007**, *3*, 1342−1348.

(25) Dahlke, E. E.; Truhlar, D. G. Application of the Electrostatically Embedded Many-Body Expansion to Microsolvation of Ammonia in Water Clusters. *J. Chem. Theory Comput.* **2008**, *4*, 1−6.

(26) Fedorov, D. G.; Kitaura, K. The Importance of Three-Body Terms in the Fragment Molecular Orbital Method. *J. Chem. Phys.* **2004**, *120*, 6832−6840.

(27) Fedorov, D. G.; Kitaura, K. On the Accuracy of the 3-Body Fragment Molecular Orbital Method (FMO) Applied to Density Functional Theory. *Chem. Phys. Lett.* **2004**, *389*, 129−134.

(28) Hermann, A.; Schwerdtfeger, P. Complete Basis Set Limit Second-Order Møller−Plesset Calculations for the fcc Lattices of Neon, Argon, Krypton, and Xenon. *J. Chem. Phys.* **2009**, *131*, 244508/1−244508/7.

(29) Beran, G. J. O. Approximating Quantum Many-Body Intermolecular Interactions in Molecular Clusters Using Classical Polarizable Force Fields. *J. Chem. Phys.* **2009**, *130*, 164115/1−164115/9.

(30) Beran, G. J. O.; Nanda, K. Predicting Organic Crystal Lattice Energies with Chemical Accuracy. *J. Phys. Chem. Lett.* **2010**, *1*, 3480−3487.

(31) Bates, D. M.; Smith, J. R.; Janowski, T.; Tschumper, G. S. Development of a 3-Body:Many-Body Integrated Fragmentation Method for Weakly Bound Clusters and Application to Water clusters (H$_2$O)$_{n=3−10,16,17}$. *J. Chem. Phys.* **2011**, *135*, 044123/1−044123/8.

(32) Bates, D. M.; Smith, J. R.; Tschumper, G. S. Efficient and Accurate Methods for the Geometry Optimization of Water Clusters: Application of Analytic Gradients for the Two-Body:Many-Body QM:QM Fragmentation Method to (H$_2$O)$_n$, *n* = 3−10. *J. Chem. Theory Comput.* **2011**, *7*, 2753−2760.

(33) Rahalkar, A. P.; Mishra, B. K.; Ramanathan, V.; Gadre, S. R. "Gold Standard" Coupled-Cluster Study of Acetylene Pentamers and Hexamers via Molecular Tailoring Approach. *Theor. Chem. Acc.* **2011**, *130*, 491−500.

(34) Bygrave, P. J.; Allan, N. L.; Manby, F. R. The Embedded Many-Body Expansion for the Energetics of Molecular Crystals. *J. Chem. Phys.* **2012**, *137*, 164102/1−164102/9.

(35) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. Estimates of the Ab Initio Limit for π−π Interactions: The Benzene Dimer. *J. Am. Chem. Soc.* **2002**, *124*, 10887−10893.

(36) Tsuzuki, S.; Honda, K.; Azumi, R. Model Chemistry Calculations of Thiophene Dimer Interactions: Origin of π-Stacking. *J. Am. Chem. Soc.* **2002**, *124*, 12200−12209.

(37) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. Origin of Attraction and Directionality of the π/π Interaction: Model Chemistry Calculations of Benzene Dimer Interaction. *J. Am. Chem. Soc.* **2002**, *124*, 104−112.

(38) Tschumper, G. S.; Leininger, M. L.; Hoffman, B. C.; Valeev, E. F.; Schaefer, H. F., III; Quack, M. Anchoring the Water Dimer Potential Energy Surface with Explicitly Correlated Computations and Focal Point Analysis. *J. Chem. Phys.* **2002**, *116*, 690−701.

(39) Sinnokrot, M. O.; Sherrill, C. D. Highly Accurate Coupled Cluster Potential Energy Curves for the Benzene Dimer: Sandwich, T-Shaped, and Parallel-Displaced Configurations. *J. Phys. Chem. A* **2004**, *108*, 10200−10207.

(40) Tauer, T. P.; Sherrill, C. D. Beyond the Benzene Dimer: An Investigation of the Additivity of π−π Interactions. *J. Phys. Chem. A* **2005**, *109*, 10475−10478.

(41) Ringer, A. L.; Sherrill, C. D. First Principles Computation of Lattice Energies of Organic Solids: The Benzene Crystal. *Chem.—Eur. J.* **2008**, *14*, 2542−2547.

(42) Hopkins, B. W.; Tschumper, G. S. Ab Initio Studies of π···π Interactions: The Effects of Quadruple Excitations. *J. Phys. Chem. A* **2004**, *108*, 2941−2948.

(43) Nanda, K. D.; Beran, G. J. O. Prediction of Organic Molecular Crystal Geometries from MP2-Level Fragment Quantum Mechanical Molecular Mechanical Calculations. *J. Chem. Phys.* **2012**, *137*, 174106/1−174106/11.

(44) White, J. C.; Davidson, E. R. An Analysis of the Hydrogen Bond in Ice. *J. Chem. Phys.* **1990**, *93*, 8029−8035.

(45) Pedulla, J. M.; Kim, K.; Jordan, K. D. Theoretical Study of the *n*-Body Interaction Energies of the Ring, Cage and Prism Forms of $(H_2O)_6$. *Chem. Phys. Lett.* **1998**, *291*, 78−84.

(46) Christie, R. A.; Jordan, K. D. *n*-Body Decomposition Approach to the Calculation of Interaction Energies of Water Clusters. In *Intermolecular Forces and Clusters II*; Wales, D., Christie, R. A., Eds.; Springer: New York, 2005; Vol. *116*, pp 27−41.

(47) Shields, A. E.; van Mourik, T. Comparison of Ab Initio and DFT Electronic Structure Methods for Peptides Containing an Aromatic Ring: Effect of Dispersion and BSSE. *J. Phys. Chem. A* **2007**, *111*, 13272−13277.

(48) van Mourik, T. Basis Set Superposition Error Effects Cause the Apparent Nonexistence of the Ethene/Benzenium Ion Complex on the MP2 Potential Energy Surface. *J. Phys. Chem. A* **2008**, *112*, 11017−11020.

(49) van Mourik, T. Comment on 'To stack or Not To Stack: Performance of a New Density Functional for the Uracil and Thymine Dimers' [Chem. Phys. Lett. 459 (2008) 164]. *Chem. Phys. Lett.* **2009**, *473*, 206.

(50) Cao, J.; van Mourik, T. Performance of the M06-L Density Functional for a Folded Tyr−Gly Conformer. *Chem. Phys. Lett.* **2010**, *485*, 40−44.

(51) Toroz, D.; van Mourik, T. Structure of the Gas-Phase Glycine Tripeptide. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3463−3473.

(52) Grossfield, A.; Ren, P.; Ponder, J. W. Ion Solvation Thermodynamics from Simulation with a Polarizable Force Field. *J. Am. Chem. Soc.* **2003**, *125*, 15671−15682.

(53) Weigend, F.; Häser, M.; Patzelt, J.; Ahlrichs, R. RI-MP2: Optimized Auxiliary Basis Sets and Demonstration of Efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143−152.

(54) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. Basis-Set Convergence of the Energy in Molecular Hartree−Fock Calculations. *Chem. Phys. Lett.* **1999**, *302*, 437−446.

(55) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley: New York, 2000.

(56) Shao, Y.; et al. Advances in Methods and Algorithms in a Modern Quantum Chemistry Program Package. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172−3191.

(57) Krylov, A. I.; Gill, P. M. W. Q-Chem: An Engine for Innovation. *WIREs Comput. Mol. Sci.* **2013**, *3*, 317−326.