Understanding the many-body expansion for large systems. II. Accuracy considerations

Ka Un Lao, Kuan-Yu Liu, Ryan M. Richard,[a] and John M. Herbert[b]
Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, USA

(Received 16 February 2016; accepted 5 April 2016; published online 22 April 2016)

To complement our study of the role of finite precision in electronic structure calculations based on a truncated many-body expansion (MBE, or "*n*-body expansion"), we examine the accuracy of such methods in the present work. Accuracy may be defined either with respect to a supersystem calculation computed at the same level of theory as the *n*-body calculations, or alternatively with respect to high-quality benchmarks. Both metrics are considered here. In applications to a sequence of water clusters, $(H_2O)_{N=6-55}$ described at the B3LYP/cc-pVDZ level, we obtain mean absolute errors (MAEs) per H_2O monomer of ∼1.0 kcal/mol for two-body expansions, where the benchmark is a B3LYP/cc-pVDZ calculation on the entire cluster. Three- and four-body expansions exhibit MAEs of 0.5 and 0.1 kcal/mol/monomer, respectively, without resort to charge embedding. A generalized many-body expansion truncated at two-body terms [GMBE(2)], using 3–4 H_2O molecules per fragment, outperforms all of these methods and affords a MAE of ∼0.02 kcal/mol/monomer, also without charge embedding. GMBE(2) requires significantly fewer (although somewhat larger) subsystem calculations as compared to MBE(4), reducing problems associated with floating-point roundoff errors. When compared to high-quality benchmarks, we find that error cancellation often plays a critical role in the success of MBE(*n*) calculations, even at the four-body level, as basis-set superposition error can compensate for higher-order polarization interactions. A many-body counterpoise correction is introduced for the GMBE, and its two-body truncation [GMBCP(2)] is found to afford good results without error cancellation. Together with a method such as ωB97X-V/aug-cc-pVTZ that can describe both covalent and non-covalent interactions, the GMBE(2)+GMBCP(2) approach provides an accurate, stable, and tractable approach for large systems. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4947087]

I. INTRODUCTION

Electronic structure methods based on molecular fragmentation are an increasingly popular way to sidestep the non-linear scaling of computational cost with respect to system size.[1–4] Such methods rely, at some level, on Kohn's principle of the "near-sightedness" of electronic matter,[5,6] and attempt to decompose a large calculation into a (potentially large number of) small subsystem calculations that are independent of one another and thus lend themselves to trivially parallel distributed computing. High accuracy is reported in many applications (see Ref. 1 for a review), and thus molecular fragmentation would seem to offer the proverbial "free lunch," enabling high-level *ab initio* methods to be applied to large systems at a fraction of the cost (in wall time, memory, and disk requirements) that would otherwise be required.

Due to the high cost of obtaining benchmark results in large systems, the accuracy of fragment-based methods has primarily been evaluated in small systems (typically ≲25 heavy atoms) and/or at low levels of electronic structure theory (e.g., self-consistent field theory with minimal or double-ζ basis sets). It is unclear whether such benchmarks are representative of the performance that can be expected when high levels of theory are applied to larger systems. Consider that even the smallest naturally occurring protein (the 20-residue Trp cage) has a total electronic energy approaching 10^4 hartree, which must therefore be predicted to a precision of about 0.000 01% in order to achieve an accuracy of ∼1 kcal/mol in the total energy. This is the famous "weighing the captain" problem in electronic structure theory,[7,8] i.e., determining the captain's weight based on measuring the ship's displacement when she is, or is not, on board. The situation is arguably somewhat worse for fragment-based methods, which require huge numbers of electronic structure calculations that must be performed at significantly higher precision than is required in conventional quantum chemistry.[9] A more apt analogy for fragment-based quantum chemistry might be the notion of determining the weight of a pilot on an aircraft carrier based on measuring the displacement when the pilot sits in his plane, versus the displacement for various combinations of aircraft and pilots on board the ship.

In part I of this series,[9] we examined the role of finite-precision arithmetic in methods based on a truncated many-body expansion (MBE), otherwise known as an "*n*-body expansion." The analysis in Ref. 9 focused on systematic error as system size was increased, and we found that the uncertainty in the *n*-body approximation to the energy is strongly influenced by the self-consistent field (SCF)

[a]Present address: School of Chemistry & Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332, USA.
[b]herbert@chemistry.ohio-state.edu

convergence threshold used for the individual subsystem calculations (much more so than is the total supersystem energy), as well as by accumulation of floating-point rounding errors. Implementations of the MBE that rely on external calls to an electronic structure program (rather than being fully integrated into such a program) suffer from an additional source of uncertainty, namely, the fact that discrepancies as small as 10^{-6} a.u. between the driver routine and the electronic structure program (due to six-digit roundoff in the electronic structure output, for example) can translate into errors of several kcal/mol, for systems not much larger than $(H_2O)_{30}$.[9] Owing to the combinatorial nature of the MBE, these problems are compounded as the system size increases, and also compounded as one moves to higher n in pursuit of greater accuracy.

Although these precision problems appear to be surmountable using a combination of tight SCF convergence thresholds, arbitrary-precision arithmetic to sum the terms in the MBE, and a consistent internal precision in all electronic structure calculations,[9] these considerations do increase the cost of the calculations relative to what has previously been discussed in the literature. The worst cases manifest only when $n \geq 4$, and for very large collections of fragments ($N \gtrsim 40$) that may be avoidable in applications to polyatomic molecules, and might be sidestepped in applications to molecular liquids via some kind of distance-based criterion for discarding or approximating well-separated n-mers. Thus, two- and three-body approaches might still prove useful, provided that good accuracy can be obtained.

Using water clusters as exemplary cases where many-body polarization effects are important, we demonstrate in the present work that a four-body expansion [MBE(4)] is required in order to obtain accurate values for both total interaction energies as well as relative energies of various isomers. On the other hand, a *generalized* many-body expansion[2,10,11] (GMBE) that utilizes overlapping subsystems can provide results of comparable accuracy at the two-body level. This is not a "free lunch" relative to MBE(4), because the size of the subsystem calculations increases in this "GMBE(2)" approach, but the *number* of subsystem calculations is dramatically reduced, staving off numerical precision problems while maintaining the trivial parallelizability of the traditional MBE approach.

Following up on our study of precision problems in Ref. 9, here we seek to evaluate the accuracy of various n-body approximations, using the same set of water clusters that was examined in Ref. 9. Regrettably, the overwhelming majority of our present results fail to achieve 1 kcal/mol accuracy in the total energy as approximated at the three-body level, but do reveal some interesting trends as these methods are pushed towards the large-system ($N \to \infty$) limit. Here, we focus on: (1) whether the supersystem energy at a given level of theory can be accurately approximated when the same level of theory is used for the subsystem calculations; (2) the oscillatory nature of the n-body expansion as the truncation order, n, is increased; (3) size-dependent errors in the n-body total energy and what impact these have on the prediction of relative energies; and (4) the ability of these methods to reproduce high-level benchmarks. As in our previous study,[9] our findings unearth potential pitfalls that have not been widely discussed in the rapidly growing literature on fragment-based quantum chemistry.

II. THEORY

A. Traditional MBE

The basic idea of the n-body expansion is straightforward and has been reviewed elsewhere.[1–4] The total energy

$$E = \sum_{I=1}^{N} E_I + \sum_{I=1}^{N} \sum_{J<I} (E_{IJ} - E_I - E_J) + \cdots \tag{1}$$

is expressed as a sum of monomer energies (E_I), dimer energies (E_{IJ}), etc., becoming exact (by tautological definition) when $n = N$. Consult Ref. 9 for compact formulas. To place Eq. (1) in the context of what is to come, note that this equation can alternatively be expressed as[12]

$$E = \sum_{I=1}^{N} \Delta E_I^{(1)} + \sum_{J=1}^{\binom{N}{2}} \Delta E_J^{(2)} + \sum_{K=1}^{\binom{N}{3}} \Delta E_K^{(3)} + \cdots, \tag{2}$$

where

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} \tag{3}$$

and

$$\Delta E_I^{(n)} = \sum_{m=1}^{n} (-1)^{n-m} \sum_{J \subset I} E_J^{(m)} \tag{4}$$

is a correction to the energy of the Ith n-body subsystem. The quantity $E_J^{(m)}$ in Eq. (4) is the energy of the Jth m-body sub-cluster formed from the Ith n-body cluster. As such, the second summation in Eq. (4) ranges from $J = 1, 2, \ldots, \binom{n}{m}$, and this is the meaning of the $J \subset I$ restriction in that equation.

In cases where fragmentation does not sever any covalent bonds, either of Eqs. (1) or (2) is a formally exact expression for the total energy, but their appeal comes in dropping terms beyond some given level of n-body interaction. Neglecting terms involving $(n+1)$-body and larger sub-clusters defines the so-called n-body expansion, which we will call MBE(n). A compact formula for the energy within the MBE(n) approximation is[9]

$$E^{(n)} = \sum_{m=1}^{n} (-1)^{n-m} \binom{N-m-1}{n-m} \sum_{K=1}^{\binom{N}{m}} E_K^{(m)}. \tag{5}$$

The summations in this equation run over all unique sub-clusters containing up to n fragments.

B. Counterpoise correction in a many-body system

It is well known that electronic structure calculations converge slowly to the basis-set limit[13] (see Ref. 12 for just one example), which arises from a combination of basis-set incompleteness and basis-set superposition error (BSSE). In the present work, we address the incompleteness issue by means of complete-basis extrapolations, but the convergence of these extrapolations is sensitive to the presence

of BSSE, which we therefore attempt to eliminate. The interaction energy, ΔE, subject to a generalized Boys-Bernardi counterpoise (CP) correction is[14,15]

$$\Delta E = E^{IJK\ldots N}_{IJK\ldots N} - \sum_I E^{IJK\ldots N}_I, \tag{6}$$

where, following the convention of previous work,[12,16] the subscripts represent real molecules whereas superscripts indicate where basis functions are centered. Thus $E^{IJK\ldots N}_I$ denotes the energy of fragment I computed in a basis set having basis functions located in the positions of fragments $I, J, K, \ldots, N$; that is, $E^{IJK\ldots N}_I$ is the energy of fragment I computed in the cluster basis, as in the original Boys-Bernardi scheme.[14] The quantity $E^{IJK\ldots N}_{IJK\ldots N}$ is the normal supersystem energy, which for brevity we will henceforth denote simply as E.

Equation (6) can be trivially rewritten as

$$\begin{aligned}\Delta E &= \left(E - \sum_I E^I_I\right) + \left(\sum_I E^I_I - \sum_I E^{IJK\ldots N}_I\right) \\ &= \Delta E^{(\mathrm{uncorr})} + \delta E^{\mathrm{CP}},\end{aligned} \tag{7}$$

where we have separated the interaction energy ΔE into an "uncorrected" (and thus BSSE-contaminated) energy difference

$$\Delta E^{(\mathrm{uncorr})} = E - \sum_I E^I_I \tag{8}$$

and subsumed all BSSE corrections into a counterpoise correction

$$\delta E^{\mathrm{CP}} = \sum_{I=1}^{N} \delta E^{\mathrm{CP}}_I = \sum_{I=1}^{N} \left(E^I_I - E^{IJK\ldots N}_I\right). \tag{9}$$

The notation $\delta E^{\mathrm{CP}}_I = E^I_I - E^{IJK\ldots N}_I$ is introduced in order to make contact with approximate counterpoise corrections introduced below, in which the quantity $E^{IJK\ldots N}_I$ is approximated by means of a MBE. Equation (9) looks like the traditional Boys-Bernardi correction,[14] applied in turn to each fragment.

C. Generalized MBE

The ground-state energy can also be approximated using a GMBE that employs overlapping fragments and is derived based on the set-theoretical principle of inclusion/exclusion;[10,11] application of the GMBE requires calculations on subsystems that are formed from intersections of fragments. In an n-body GMBE, which we call GMBE(n), the approximate energy is

$$\begin{aligned}\varepsilon^{(n)} = &\sum_{i=1}^{\binom{N_f}{n}} E^{(n)}_i - \sum_{i=1}^{\binom{N_f}{n}} \sum_{j>i}^{\binom{N_f}{n}} E^{(n)}_{i\cap j} \\ &+ \cdots + (-1)^{\binom{N_f}{n}+1} E^{(n)}_{i\cap j\cap\cdots\cap\binom{N_f}{n}}.\end{aligned} \tag{10}$$

We have previously called this the *intersection-corrected energy* at order n.[10] Lower-case indices $i, j, \ldots$ in Eq. (10) refer to n-mers of fragments, whose energies are $E^{(n)}_i, E^{(n)}_j, \ldots$, and $i \cap j$ is the subsystem formed from the intersection of n-mers i and j, with energy $E^{(n)}_{i\cap j}$. For general, macromolecular applications, construction of $i \cap j$ requires severing covalent bonds and capping the severed valencies (as in a recent application of the GMBE to proteins[17]), but in this work we only consider non-covalent clusters, in an intentional effort to sidestep this complexity. Note also that in this generalized approach the number of fragments, N_f, is generally larger than the number of monomers, N. As in previous studies,[10,11,18–20] for the GMBE we use a distance cutoff of 3 Å between atoms to define the fragments.

The advantage of GMBE(n) relative to MBE(n) is that multiple monomers are included in a single fragment and the system is tessellated into overlapping fragments based on a simple distance criterion, but in a manner whose set-theoretical derivation prevents over- or undercounting of interactions despite the use of overlapping fragments.[10,11] Obviously, one could assign more than one monomer to a single fragment in a traditional MBE(n) calculation, but we anticipate that distance-based (rather than cardinality-based) fragmentation will make it easier to adjust the fragment definitions on-the-fly in a continuous way based on distance cutoffs, during simulations or geometry optimizations, although we leave this question to a future study. The presence of multiple monomers per fragment means that some many-body effects are included already at the level of GMBE(1), so this is a non-trivial approximation, in contrast to MBE(1). (All MBE calculations performed here use single-monomer fragments.) The GMBE(1) method is equivalent to the *generalized energy-based fragmentation* approach of Li and co-workers.[18,19] Up to the selection of fragments, GMBE(1) is also equivalent to the *cardinality-guided molecular tailoring approach* of Gadre and co-workers.[70,71]

We find that GMBE(2) is generally sufficient to reproduce total interaction energies in non-covalent clusters, even in very challenging cases such as $F^-(H_2O)_N$.[2,10,11] At the same time, the number of subsystem calculations remains manageable as compared to the four-body approximation that we will see is necessary to obtain accurate MBE(n) results in clusters composed of polar monomers. For water clusters, the 3 Å threshold for GMBE fragmentation that is used here typically results in fragments containing 3–4 H_2O molecules. Application of GMBE(2) to the largest cluster considered here, $(H_2O)_{55}$, requires 1469 dimer calculations (consisting of 6–9 H_2O molecules for the dimer) and 17 883 calculations on intersections (4–6 H_2O molecules each). These numbers are dwarfed by the demands of an MBE(4) calculation, which requires that subsystem calculations be performed on 341 055 tetramers (four H_2O molecules), 26 235 trimers (three H_2O molecules), and 1486 dimers. On the other hand, the GMBE(2) subsystems are larger and require ∼9 times more computational time (ωB97X-V/aug-cc-pVTZ level). Operationally, GMBE(2) sidesteps the precision problems associated with MBE(4), but at increased computational cost. Considerations of cost are revisited in Section IV C.

D. Approximate many-body counterpoise corrections

There has long been a perceived inadequacy with the Boys-Bernardi counterpoise correction, although the evidence

for this inadequacy is debatable.[21] The Valiron-Mayer function counterpoise (VMFC) approach[22] was later introduced as an attempt to quantify and correct this inadequacy. It is based on the exact energy expression in Eqs. (1) and (2), with a counterpoise correction applied order-by-order in the MBE. The resulting BSSE correction can be written

$$\delta E^{\mathrm{VMFC}} = \delta E^{\mathrm{CP}} + \sum_{I<J} \left(\Delta E_{IJ}^{IJ} - \Delta E_{IJ}^{IJK\cdots N}\right) + \cdots, \tag{11}$$

where $\Delta E_{IJ}^{XY\cdots}$, for example, is the two-body interaction between fragments I and J, computed in a basis with functions on centers $XY\cdots$. Analogous expressions exist for the three-body and higher-order terms.[12,16] Note that the monomer term is equivalent to the original Boys-Bernardi correction, δE^{CP} in Eq. (9).

In large clusters, where one might expect more significant BSSE effects, even the original Boys-Bernardi correction is expensive to evaluate, and evaluation of VMFC becomes intractable even more rapidly. Indeed, the cost of these calculations has impeded any final reckoning on the (in)adequacy of the Boys-Bernardi approach. Several approximations have been proposed that are consistent with a truncated MBE, and in particular we have suggested an n-body approximation to Eq. (9) that we have called a many-body counterpoise correction truncated at order n, or MBCP(n).[12,23] This approximation consists of applying an n-body expansion to $E_I^{IJK\cdots N}$ in Eq. (9). For n = 2–4, the MBCP(n) approximations for monomer I are

$$\delta E_I^{\mathrm{MBCP(2)}} = (N-1)E_I^I - \sum_{J\neq I}^{N} E_I^{IJ}, \tag{12}$$

$$\delta E_I^{\mathrm{MBCP(3)}} = \delta E_I^{\mathrm{MBCP(2)}} + (N-2)\sum_{J\neq I}^{N} E_I^{IJ} - \sum_{J\neq I}^{N}\sum_{\substack{K>J\\K\neq I}}^{N} E_I^{IJK} - \tfrac{1}{2}(N-2)(N-1)E_I^I, \tag{13}$$

and

$$\delta E_I^{\mathrm{MBCP(4)}} = \delta E_I^{\mathrm{MBCP(3)}} + (N-3)\sum_{J\neq I}^{N}\sum_{\substack{K>J\\K\neq I}}^{N} E_I^{IJK} - \sum_{J\neq I}^{N}\sum_{\substack{K>J\\K\neq I}}^{N}\sum_{\substack{L>K\\L\neq I}}^{N} E_I^{IJKL} - \tfrac{1}{2}(N-3)(N-2)\sum_{J\neq I}^{N} E_I^{IJ} + \tfrac{1}{6}(N-3)(N-2)(N-1)E_I^I. \tag{14}$$

Analogous to Eq. (9), the overall MBCP(n) counterpoise correction is then

$$\delta E^{\mathrm{MBCP}(n)} = \sum_{I=1}^{N} \delta E_I^{\mathrm{MBCP}(n)}. \tag{15}$$

A somewhat different counterpoise technique, called the "many-ghost many-body expansion," has been suggested recently,[24] which consists of applying a MBE to each of the terms containing ghost functions in Eq. (11). In practice, the higher-order terms are found to be negligible and only the leading term was retained in Ref. 24; this leading term is precisely what we call MBCP(n). We use this correction together with an n-body expansion of the energy in a method that we call MBE(n)+MBCP(n), meaning that n-body approximations are applied separately to both $\Delta E^{(\mathrm{uncorr})}$ and δE^{CP} in Eq. (7).

In a similar vein, Kamiya *et al.*[16] introduced an n-body approximation to the VMFC procedure, which we call VMFC(n). The MBCP(2) and VMFC(2) corrections are equivalent but the two approaches differ starting at n = 3, with VMFC(n) involving a significantly larger number of calculations. Based on comparisons to complete-basis benchmarks in Ref. 12, for systems no larger than $F^-(H_2O)_{10}$, it is not clear that the increased expense of VMFC(3) is justified, but those are the largest VMFC(3) calculations reported to date and results in larger clusters may paint a different picture.

Alternatively, δE^{CP} in Eq. (9) can be approximated using the GMBE. When truncated at order n, we will refer to this as a generalized many-body counterpoise correction, GMBCP(n), the formula for which is introduced here for the first time:

$$\delta E^{\mathrm{GMBCP}(n)} = \sum_{I=1}^{N}\left[\sum_{i=1}^{\binom{N_f}{n}} E_I^i - \sum_{i=1}^{\binom{N_f}{n}}\sum_{j>i}^{\binom{N_f}{n}} E_I^{i\cap j} + \cdots + (-1)^{\binom{N_f}{n}+1} E_I^{i\cap j\cap\cdots\cap\binom{N_f}{n}}\right]. \tag{16}$$

The notation here requires some explanation. For each fragment I, the first term inside the brackets ($\sum_i E_I^i$) represents the energy of I computed in all of the possible n-mer basis sets that contain this fragment. In the second term, the energy of monomer I is computed in all possible two-fragment intersection basis sets that contain fragment I, and so on. Like MBCP(n), there is only one "real" fragment in any of these subsystem calculations, with the others serving simply as centers for basis functions. VMFC(n), in contrast, requires calculations with up to $n-1$ real fragments in an n-body basis.

III. COMPUTATIONAL DETAILS

In some previous studies, the MBE has been used in a "multi-level" or "stratified" way, with different n-body terms computed at different levels of theory.[25–29] In other cases, efforts have been made to identify a subsystem level of theory that will reproduce high-accuracy supersystem benchmarks, computed at an altogether different level of theory.[30] In contrast, the aim here is to test the accuracy and convergence of the n-body expansion itself, so we will compare n-body calculations to supersystem benchmarks computed at the same level of theory.

In the first part of this study, we examine a sequence of water clusters, $(H_2O)_{N=6-55}$, whose structures are putative global minima on the TIP4P potential surface at each cluster size.[31] The same set of clusters was used in our previous investigation of precision considerations for the MBE,[9] and as in that study, we use the affordable B3LYP/cc-pVDZ level of theory since our goal is to understand the size- and n-body-dependence of errors in total interaction energies. In view of a recommendation by Ouyang *et al.*[32] that the MBE should never be used with augmented basis sets (due to the appearance of serious BSSE problems when diffuse functions are added), we will also test aug-cc-pVDZ and 6-31+G(d,2p). The SG-1 quadrature grid[33] was used for all B3LYP calculations. A detailed study in Ref. 9 suggests that while MBE results are sensitive to certain numerical thresholds, as discussed below, they are only very weakly dependent on the choice of grid.

Various forms of electrostatic embedding (EE) have been used in an attempt to hasten the convergence (and thus improve the accuracy) of the n-body expansion,[1] and in some cases EE-MBE results are surprisingly insensitive to the precise details of the embedding charges,[3,34,35] at least in systems such as small water clusters. Here, we test the effects of including supersystem-derived atom-centered point charges computed at the B3LYP/cc-pVDZ level. (This is not ideal for production calculations, as it requires the very supersystem calculation that we are trying to avoid, but is used here for testing purposes and in the end we will conclude that embedding charges provide very little advantage anyway.) We consider Mulliken charges, Hirshfeld charges,[36] charges from natural population analysis (NPA),[37] charges derived from the electrostatic potential ("ChElPG"),[38] and charges from the empirical CM5 model.[39]

To avoid precision issues associated with the reading and writing of electronic structure input and output files, all calculations are carried out completely internally within a locally-modified version of Q-Chem,[40] where the subsystem calculations are parallelized via the message-passing interface, MPI. (This will be released with Q-Chem v. 4.4.) By performing all MBE calculations internally within Q-Chem, we avoid the need to compute and subtract out the self-interaction of the embedding charges. Although this may seem like a trivial point, errors due to finite precision can accumulate rapidly in the self-energy calculation, especially for the larger systems considered here, if the self-interaction calculation is not performed using all digits of double-precision arithmetic.[9] GMBE and GMBCP calculations were performed using our driver program, Fragme∩t,[10,11,17] to prepare and execute the necessary Q-Chem input files in full double precision. Note that this requires reading Q-Chem's binary scratch files and thus is also tightly interconnected with the electronic structure program.

In the second part of this study, we focus on establishing, and then attempting to reproduce, high-accuracy benchmarks for a set of $(H_2O)_{20}$ clusters representing four different structural motifs. In Ref. 42, structures for 20 different isomers of $(H_2O)_{20}$ are reported, representing the five lowest-lying isomers in each of the four families of structures, according to basin-hopping Monte Carlo calculations. For reasons of cost, we consider only 8 of these structures, representing the highest- and lowest-energy structures of the five reported for each structural motif. (Structures of the 8 that we consider are provided in the supplementary material.[41]) Benchmark energetics are computed using second-order Møller-Plesset perturbation theory (MP2) within the resolution-of-identity (RI) approximation,[43–45] RIMP2, as implemented in Q-Chem.[46] The aug-cc-pVTZ (aTZ) and aug-cc-pVQZ (aQZ) basis sets are used in conjunction with their corresponding auxiliary basis sets,[45,47] and only valence orbitals are correlated. We assume that the Hartree-Fock/aQZ energy is sufficiently close to the complete basis set (CBS) limit to be used without extrapolation, whereas we extrapolate the RIMP2 correlation energies using a two-point ("aTZ/aQZ") X^{-3} scheme.[13] Standard Boys-Bernardi counterpoise corrections are applied,[14] which for a many-body cluster means that we compute each monomer energy using the full cluster basis,[15] as in Eq. (9).

In principle these RIMP2/CBS benchmarks could be extended to the CCSD(T)/CBS level by means of a triples correction

$$\delta_{\mathrm{CCSD(T)}} = E_{\mathrm{CCSD(T)}} - E_{\mathrm{MP2}}, \tag{17}$$

and good results based on two- and three-body approximations to Eq. (17) have been demonstrated.[12,48–50] The $\delta_{\mathrm{CCSD(T)}}$ correction for $(H_2O)_{20}$ has been estimated to be around −2.7 kcal/mol,[20,51] but in the interest of performing a larger number of exploratory calculations, we omit this correction from the MP2 calculations reported here. As an alternative, we will report calculations using the ωB97X-V density functional,[52] as it has been shown to provide accurate non-covalent interaction energies, when used in conjunction with a triple-ζ basis set.[51–53] In particular, for $(H_2O)_{20}$, counterpoise-corrected ωB97X-V/def2-TZVPPD calculations differ from CCSD(T)/CBS benchmarks only by an average of 1.7 kcal/mol (see the supplementary material[41]). Per the recommendation in Ref. 52, we used an Euler-Maclaurin-Lebedev grid with 75 radial points and 302 angular points to integrate the semi-local parts of ωB97X-V, and the SG-1 grid[33] for the nonlocal correlation part.

IV. RESULTS AND DISCUSSION

In Section IV A below, we focus on how well the MBE reproduces a supersystem calculation at the same level of theory. It will be seen that this issue is complicated by the oscillatory (with respect to n) and size-extensive (with respect

to N) nature of the resulting errors. In Section IV B, we focus on the ability of the MBE to reproduce high-accuracy benchmarks. There, we will see that error cancellation between MBE(n) and MBCP(n) plays a major role in the accuracy of the results, whereas the good performance of GMBE(2)+GMBCP(2) does not rely on cancellation of errors. As noted above, all MBE calculations use a single monomer per fragment whereas GMBE calculations define fragments based on a 3 Å distance threshold, resulting in 3–4 monomers per fragment.

A. Absolute performance of the many-body expansion

1. Errors versus system size

Previously, we observed that the total energy predicted by an n-body expansion can change dramatically as a function of both the SCF convergence criterion, τ_{SCF}, as well as the integral screening threshold, τ_{ints}.[3,9] Figure 1 presents a more thorough, size-dependent analysis of the effects of numerical thresholds, comparing results obtained with "loose" thresholds (defined here as $\tau_{\mathrm{SCF}} = 10^{-5}$ a.u. and $\tau_{\mathrm{ints}} = 10^{-9}$ a.u.) to those garnered from "tight" thresholds ($\tau_{\mathrm{SCF}} = 10^{-7}$ a.u. and $\tau_{\mathrm{ints}} = 10^{-14}$ a.u.). A statistical summary of the mean absolute errors (MAEs) is given in Table I. For the absolute benchmarks presented in Section IV A, we define "error" as

$$\text{error} = E^{(n)} - E_{\text{supersystem}}, \tag{18}$$

where all calculations (supersystem and subsystem) are performed at the same level of theory. Negative errors indicate that the n-body approximation is overbound with respect to the benchmark.

From Fig. 1, the difference between MBE(2) using loose and tight thresholds cannot be differentiated, and is only barely discernible for MBE(3), except for the largest clusters. When using a four-body expansion, however, the accumulation of roundoff is enough to change the sign of the error for $N \gtrsim 40$. In the largest cluster, $(H_2O)_{55}$, the differences between results with loose and tight thresholds are

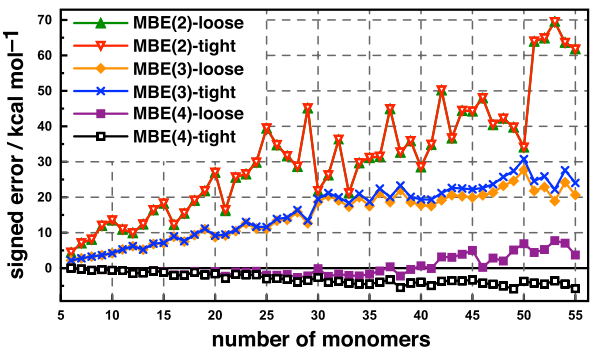

FIG. 1. Signed errors in two-, three-, and four-body total interaction energies for a series of $(H_2O)_{6-55}$ clusters. As defined in Eq. (18), "error" means the difference in the n-body approximation and a supersystem calculation carried out at the same level of theory as the subsystem calculations, which is B3LYP/cc-pVDZ in this case. The "loose" thresholds are $\tau_{\mathrm{SCF}} = 10^{-5}$ a.u. and $\tau_{\mathrm{ints}} = 10^{-9}$ a.u., whereas "tight" thresholds are $\tau_{\mathrm{SCF}} = 10^{-7}$ a.u. and $\tau_{\mathrm{ints}} = 10^{-14}$ a.u.

TABLE I. MAEs and (in parentheses) MAEs per monomer, for (G)MBE(n) calculations of clusters in the range $(H_2O)_{6-55}$, with energetics computed at the B3LYP/cc-pVDZ level.

	MAE (kcal/mol)			
Embedding	$n = 1$	$n = 2$	$n = 3$	$n = 4$
	— MBE (loose thresholds) —			
None	...	31.31 (1.04)	14.73 (0.48)	2.11 (0.07)
	— MBE (tight thresholds) —			
None	...	31.31 (1.04)	15.88 (0.51)	3.01 (0.10)
Mulliken	...	5.38 (0.20)	14.15 (0.45)	2.14 (0.07)
ChElPG	...	47.14 (1.45)	12.68 (0.40)	0.79 (0.03)
NPA	...	57.74 (1.77)	12.57 (0.39)	0.75 (0.03)
Hirshfeld	...	6.92 (0.25)	14.29 (0.46)	2.24 (0.07)
CM5	...	30.59 (0.92)	12.91 (0.41)	1.11 (0.04)
	— GMBE (tight thresholds) —			
None	26.68 (0.78)	0.68 (0.02)	...	...

−0.11 kcal/mol (two-body), 3.40 kcal/mol (three-body), and −9.47 kcal/mol (four-body). This observation complements the propagation-of-errors analysis that we presented in Ref. 9, and underscores the fact that each subsystem energy in the MBE is multiplied by a binomial coefficient that is growing factorially with respect to both n and N [see Eq. (5)]. All remaining calculations in this work will be performed using the tight thresholds. For a B3LYP/cc-pVDZ calculation of $(H_2O)_{55}$, this requires about 2.4 times more computer time as compared to the looser thresholds.

The other obvious message from these data is that, even with tight thresholds, errors at the MBE(2) and MBE(3) level are unacceptably large, even for clusters containing only a few water molecules. Electrostatic embedding is designed to rectify this, so in Fig. 2 we plot size-dependent errors in MBE(n) using a variety of different embedding charges. In stark contrast to the conventional wisdom that embedding charges should improve the accuracy of truncated MBEs, no such trend is observed at the two-body level, where the errors can change dramatically (even in sign) depending upon the choice of point charges. In the three- and four-body cases, use of any one of five different choices for the embedding charges does lead to a systematic reduction in the error across all cluster sizes, and the results are consistent with previous observations that the results are rather insensitive to the particulars of how the embedding charges are chosen.[3,34,35] Moreover, they are consistent with the notion that the details of the embedding should matter less as n increases and ever-larger sub-clusters are described quantum-mechanically.[34] On the other hand, in the three-body case the error reduction engendered by the embedding charges is quite small relative to the overall error with respect to the supersystem benchmark. Only in the four-body case—which we would like to avoid, by virtue of its disastrous combinatorics—do we see a meaningful reduction in the errors by virtue of the embedding charges.

Of all the EE-MBE(n) methods examined in Fig. 2, only the four-body expansion with NPA, ChElPG, or CM5 charges achieves the so-called "chemical accuracy" of ∼1 kcal/mol.

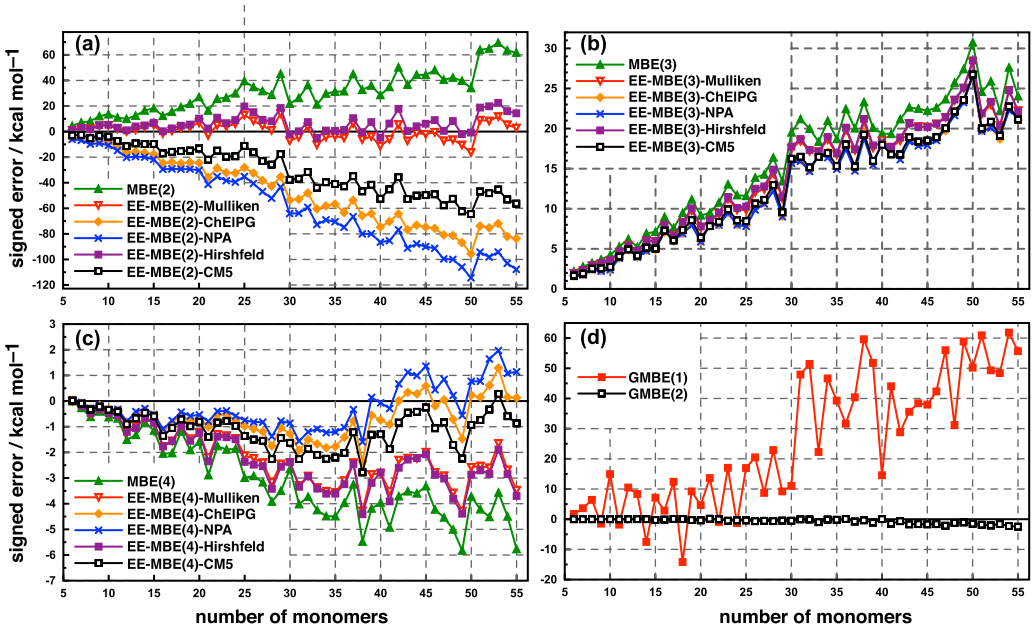

FIG. 2. Signed errors in total interaction energies for a series of water clusters using (a) two-body, (b) three-body, and (c) four-body expansions, at the B3LYP/cc-pVDZ level, either with or without electrostatic embedding. In (d), we plot the signed errors for GMBE(1) and GMBE(2). Error is measured with respect to a supersystem calculation at the same level of theory as the subsystem calculations. Note that the vertical energy scales are different in each panel.

Note that the error *per monomer* (Table I) is a size-intensive quantity, unlike the total errors in Fig. 2, and it has been argued that for dynamical studies of clusters, it is sufficient to achieve a "dynamical accuracy" per monomer equal to 10% of the average molecular kinetic energy at room temperature, $(3/2)k_BT$.[24] This amounts to an accuracy threshold of 0.09 kcal/mol/monomer, and unfortunately none of the two- or three-body expansions achieve even this level of accuracy, although all of the four-body methods do, including MBE(4) with no embedding charges at all. This is a useful fact to note, given that evaluation of the Coulomb self-energy of the embedding charges is a serious obstacle to obtaining reproducible, high-precision results when the n-body expansion is implemented via an external driver program or script.[9]

Several previous studies have noted that the "success" of n-body expansions often relies heavily on error cancellation,[16,23,25,32] and the EE-2B data in Fig. 2(a) suggest that MBE(2) with either Mulliken or Hirshfeld embedding represents a "Pauling point"[54] at the two-body level. This is almost certainly not for any physically meaningful reason, given that the average Mulliken charge on oxygen is $-0.26e$ in the case of $(H_2O)_{40}$, and $-0.22e$ when Hirshfeld charges are used, as compared to $-0.80e$ for ChElPG (which is similar to force-field charges designed to reproduce the presumed H_2O dipole moment) and $-0.99e$ for NPA (consistent with chemical intuition). It is interesting to note that for the EE-MBE(2) and EE-MBE(3) plots in Figs. 2(a) and 2(b), the ordering of the errors as a function of the choice of embedding charges follows the trend

$$\text{none} > \text{Hirshfeld} > \text{Mulliken} > \text{CM5} > \text{ChElPG} > \text{NPA},$$

i.e., exactly the same order as the magnitude of the embedding charge on the oxygen atom, and thus the same order as the H_2O dipole moments. On the other hand, for EE-MBE(4) [Fig. 2(c)], the order is precisely reversed.

Finally, results for GMBE(n) are plotted in Fig. 2(d). As noted above, the GMBE includes some many-body effects even for $n = 1$, but these turn out to be insufficient to afford reasonable accuracy even in small clusters, at least with the 3 Å threshold for fragment formation that is used here. Using the same threshold, however, GMBE(2) results are outstanding, with a MAE per monomer of only 0.02 kcal/mol across the whole range of cluster sizes, $N = 6$–55. This is more accurate than any of the MBE(4) methods, and this accuracy does not require the use of embedding charges. While it is likely true that the total error in GMBE(2) calculations is size-extensive, the plot in Fig. 2(d) suggests that the rate of growth with N is so small that this extensivity is unlikely to prove problematic for values of N that are likely to be used in practical calculations, and GMBE(2) can safely be employed to predict *total* cluster binding energies.

2. Basis-set dependence

Much of the rationale for examining water clusters in this work is these have been popular test cases for various fragment-based methods, and in the supplementary material[41] we compare our EE-MBE(2) and EE-MBE(3) results to a large number of previous benchmarks from the literature. Care must be taken in comparing these literature values either to one another or to the numbers reported here, because it is uncommon to find a detailed specification of how the thresholds were set in a given calculation, and these settings can significantly alter the results, as can the choice of electrostatic embedding, in the two-body case.

That said, a survey of the literature suggests that the large MBE(2) errors documented above are not unprecedented. Our

MBE(3) results are noticeably worse than most literature values, however. With only a few exceptions, the literature reports errors of $\lesssim$3 kcal/mol for clusters in the size range N = 16–32, with numerous reports of errors <1 kcal/mol in basis sets such as STO-3G, 3-21G, and 6-31G* (see Table S2).[41]

In an attempt to understand this discrepancy, we examined size-dependent trends in the MBE(3) errors in three different basis sets, as plotted in Fig. 3 with error statistics summarized in Table II. In contrast to B3LYP/cc-pVDZ results, where the choice of embedding charges made only a minor difference in the error, results using aug-cc-pVDZ show a dramatic dependence on the choice of embedding charges. Starting from the MBE(3) results without charge embedding [green symbols in Fig. 3(b)], the various EE-MBE(3) plots diverge in an order that correlates precisely with the average embedding charge on the oxygen atoms, which is small for the Hirshfeld and Mulliken embeddings but nearly $-1e$ for NPA charges. Use of the latter leads to errors that approach 50 kcal/mol for the largest clusters, whereas errors are <10 kcal/mol when no embedding charges are used at all.

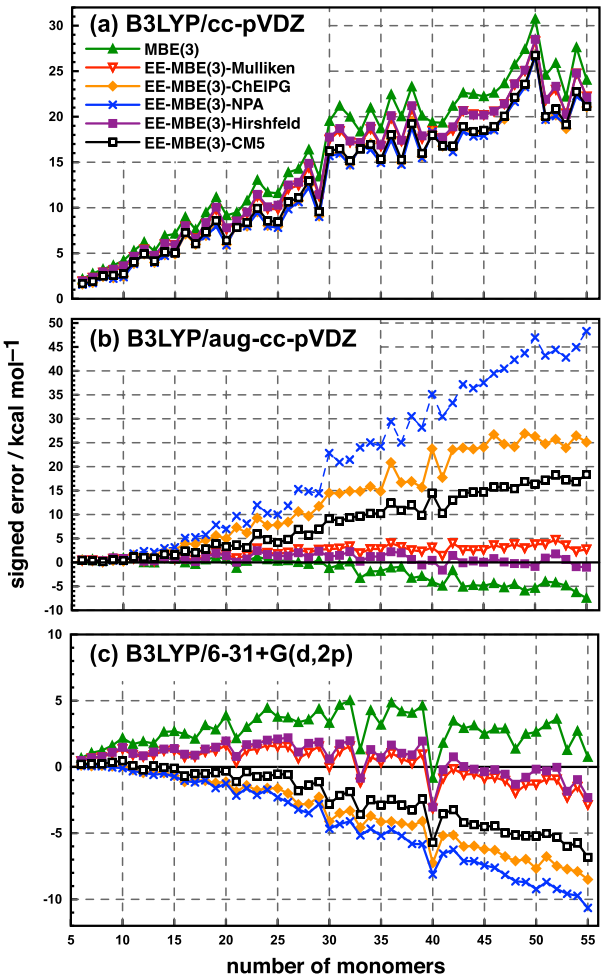

FIG. 3. Signed errors in the MBE(3) total interaction energies for $(H_2O)_{6-55}$, using the B3LYP functional with three different basis sets. Error is measured with respect to a supersystem calculation at the same level of theory as the subsystem calculations.

TABLE II. Mean absolute errors per monomer for EE-MBE(3) calculations of $(H_2O)_{6-55}$ at the B3LYP level in several basis sets. Supersystem charges computed at the B3LYP/cc-pVDZ are used for charge embedding.

	MAE per monomer (kcal/mol)		
Embedding	cc-pVDZ	aug-cc-pVDZ	6-31+G(d,2p)
None	0.51	0.06	0.11
Mulliken	0.45	0.07	0.05
ChElPG	0.40	0.37	0.09
NPA	0.39	0.56	0.11
Hirshfeld	0.46	0.04	0.05
CM5	0.41	0.23	0.06

Very accurate results for small water clusters (N < 10) were reported in Ref. 26 using EE-MBE(3) at the B3LYP/6-31+G(d,2p) level, so we have applied the same methodology to the $(H_2O)_{6-55}$ sequence, with errors reported in Fig. 3(c). MAEs in Table II show that each choice of embedding charges except NPA leads to reduction in the errors, and even the errors using NPA charges are comparable in magnitude (though different in sign) to those without embedding. The errors are also much smaller using 6-31+G(d,2p) as compared to either of the other two basis sets, although total errors remain $\gg$1 kcal/mol for the largest clusters. Despite smaller absolute errors in the case of 6-31+G(d,2p), however, note from Fig. 3 that the choice of which particular embedding charges to use (or none at all) matters much more for the two basis sets that include diffuse functions than it does for cc-pVDZ. The larger variations in the case of aug-cc-pVDZ and 6-31+G(d,2p) are likely due to particularly strong interactions between point charges and diffuse basis functions, i.e., the "overpolarization" problem that is well known in QM/MM calculations.[55–58]

This dramatic sensitivity to basis sets is the main conclusion from Fig. 3. A recent review of fragment-based methods[1] mentions the need for a comprehensive study of how the accuracy of each method depends upon the level of electronic structure theory used for the subsystem calculations, yet few systematic studies exist in the literature.[3,9,23,32] Here, we use B3LYP in conjunction with six different double- and triple-ζ basis sets to study the performance of (G)MBE(n) for $(H_2O)_{40}$, both in the absence of embedding charges and when ChElPG embedding charges are used. Signed errors are reported in Table III.

For MBE(2), results are improved by employing charge embedding, in all basis sets except cc-pVDZ and cc-pVTZ. For MBE(3), results are improved by employing charge embedding in all basis sets except aug-cc-pVDZ and 6-31+G(d,2p). For MBE(4), results are improved only for cc-pVDZ and cc-pVTZ by employing charge embedding. In short, the results are rather erratic, and in none of the six basis sets tested does charge embedding consistently reduce the errors for two-, three-, or four-body expansions. Moreover, in the absence of embedding charges we observe a monotonic decrease in the errors (going from n = 2 to 4) only for cc-pVDZ and 6-31G(d,2p). This is consistent with the observation by Ouyang *et al.*[32] that only in the absence of diffuse functions does one obtain monotonic convergence of the MBE, although

TABLE III. Signed errors for $(H_2O)_{40}$ using B3LYP in various basis sets. (Embedding charges are from a supersystem calculation at the level of B3LYP/cc-pVDZ.)

		Signed error (kcal/mol)			
MAE	Embedding	MBE(2)	MBE(3)	MBE(4)	GMBE(2)
cc-pVDZ	None	28.56	19.39	−3.96	−0.79
cc-pVDZ	ChElPG	−74.74	18.36	−0.74	...
aug-cc-pVDZ	None	105.00	−4.05	−9.26	0.23
aug-cc-pVDZ	ChElPG	−22.06	23.74	−75.35	...
cc-pVTZ	None	20.09	32.59	−9.85	0.05
cc-pVTZ	ChElPG	−68.78	28.71	−6.50	...
aug-cc-pVTZ	None	105.54	−8.92	9.26	0.33
aug-cc-pVTZ	ChElPG	−0.40	4.19	−23.40	...
6-31G(d,2p)	None	74.18	5.83	−0.82	0.07
6-31G(d,2p)	ChElPG	−32.47	5.77	1.60	...
6-31+G(d,2p)	None	84.24	−0.84	5.19	0.73
6-31+G(d,2p)	ChElPG	6.68	−7.26	5.65	...

our results for cc-pVTZ are an exception demonstrating that the absence of diffuse functions alone does not guarantee monotonic convergence. In summary, and in view of all the data presented so far, it is difficult to predict whether MBE(n) results will be improved or degraded by MBE($n+1$) for a given basis set or charge-embedding scheme.

Fortunately, GMBE(2) performs well for all six basis sets tested, and is superior to (or in one case, comparable to) MBE(4) results with or without charge embedding. Recent work has shown that the performance of GMBE(2) does depend on the nature of the embedding charges,[17] so in this study we choose *not* to use these charges in GMBE calculations.

3. Convergence of the expansion

Ultimately, one needs a metric for gauging the accuracy of a given approach that does not require a supersystem calculation, else the utility of the MBE is lost. Given the apparent extensivity of errors in total interaction energies, it is not immediately clear how (or even whether) small-cluster benchmarks can be used to assess accuracy in applications to much larger clusters. A reasonable metric is to examine how the errors converge (or fail to converge) with respect to the truncation order, n, but unfortunately the convergence behavior of the MBE is complicated,[24,32,59–62] as seen above. For atomic clusters, the MBE is oscillatory and slowly convergent, often precluding truncation.[59–61] For small molecular clusters, the situation may not be so dire, as the oscillations tend to settle down much more quickly,[25,61,63] possibly owing to the weaker nature of intermolecular interactions as compared to interatomic interactions.

Consider the n-body interaction energy $\Delta E_I^{(n)}$ in Eq. (4). As a simple, qualitative model, let us assume that the error in each subsystem energy $E_J^{(m)}$ is the same, and denote this error as δE. Then the total error in $\Delta E_I^{(n)}$ would be

$$\delta\left(\Delta E_I^{(n)}\right) = (\delta E)\sum_{m=1}^{n}(-1)^{n-m}\binom{n}{n-m} = (-1)^{n+1}(\delta E). \tag{19}$$

This suggests there is reason to expect that errors in MBE(n) calculations may oscillate as a function of N, although the result is only rigorous if the errors are truly identical for each n, which likely requires that $|\Delta E_I^{(n)}| \approx |\Delta E_J^{(n+1)}|$. If the magnitude of the $(n+1)$-body interactions is significantly different than those of the n-body interactions, then one cannot say with certainty whether the error will oscillate or not.

In light of this analysis, let us reconsider the $n = 2$ results in Fig. 2(a). According to Eq. (19), the error should be negative, and indeed large, negative errors are obtained using CM5, ChElPG, and NPA charges, but on the other hand large *positive* errors are obtained in the absence of charge embedding. (For reasons ultimately having to do with error cancellation, the Hirshfeld and Mulliken embeddings lead to errors much closer to zero, and which change sign as a function of cluster size.) These observations suggest that the analysis in Eq. (19) is overly simplistic in this case, consistent with the fact that $|\Delta E_I^{(1)}| \gg |\Delta E_I^{(2)}|$.

On the other hand, errors for EE-MBE(3) [Fig. 2(b)] are all positive, consistent with Eq. (19), and for EE-MBE(4) [Fig. 2(c)], most of the errors are negative, and those that are positive are relatively small. Thus, we conclude there is reason to expect that the MBE converges (if at all) in an oscillatory way, behavior that may prove troublesome for high-accuracy applications if the oscillations do not decay rapidly enough. In the present calculations, it is not clear that convergence is achieved even by $n = 4$, necessitating five-body calculations to check for convergence. As noted in Ref. 23, and in Section IV B below, these oscillations may play a role in error cancellation. BSSE is a significant component of these oscillations.[32]

B. High-accuracy calculations of relative energies

To this point we have focused on the intricacies of replicating a low level of theory in a large system. We now turn our attention to an investigation of the accuracy of the MBE as compared to high-quality RIMP2/CBS and ωB97X-V/aTZ benchmarks, specifically for relative energies of $(H_2O)_{20}$ isomers. Unlike the calculations in Section IV A, where "error" is defined so as to measure how faithful the n-body approximation is to the supersystem calculation [Eq. (18)], here we are trying to reproduce accurate benchmarks but also to understand the role of any BSSE errors that we observe. In this section we therefore define

$$\text{error} = E^{(n)} - E_{\text{benchmark}}, \tag{20}$$

where the benchmark may be counterpoise-corrected or not. Counterpoise-corrected supersystem benchmarks will be compared to counterpoise-corrected n-body results, using the (G)MBCP(n) corrections in the latter case, and uncorrected supersystem benchmarks will be compared to uncorrected (G)MBE(n) calculations. As in Section IV A, we always compare sub- and supersystem calculations computed using the same electronic structure method.

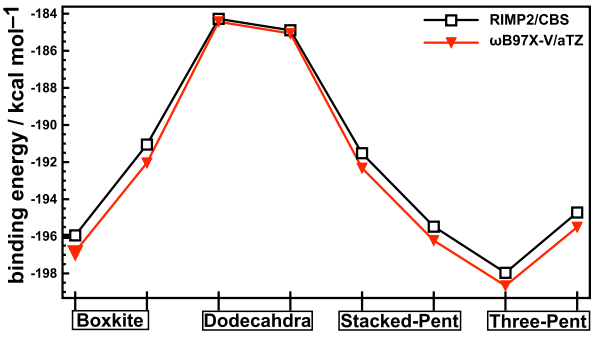

FIG. 4. Benchmark total interaction energies for eight different isomers of $(H_2O)_{20}$, including two from each of the four families of structural motifs.[42] All calculations are counterpoise-corrected and the RIMP2/CBS extrapolation uses the aTZ and aQZ basis sets.

1. Accuracy of the benchmarks

Total interaction energies for eight different $(H_2O)_{20}$ clusters are plotted in Fig. 4. In these calculations, we attempted to reach the RIMP2/CBS limit using T/Q extrapolation of RIMP2/aug-cc-pVXZ data, both with and without counterpoise corrections. Counterpoise corrections are quite large, on average 21 kcal/mol for the aTZ basis set and 10 kcal/mol for the aQZ basis set, and the extrapolated RIMP2/CBS values in the absence of counterpoise correction are shifted to larger interaction energies by ≈3.5 kcal/mol relative to an extrapolation of the counterpoise-corrected values. The counterpoise-corrected RIMP2/CBS extrapolations differ by an average of 2.7 kcal/mol as compared to counterpoise-corrected CCSD(T)/CBS benchmarks.[41]

The ωB97X-V functional[52] combined with triple-ζ basis sets gives good performance for interaction energies in $(H_2O)_{20}$ with a MAE of 1.7 kcal/mol as compared to CCSD(T)/CBS benchmarks.[41] In the absence of counterpoise correction, ωB97X-V/aTZ values are shifted to higher interaction energies by ≈2.8 kcal/mol, relative to counterpoise-corrected results. In contrast to the RIMP2/aTZ case, counterpoise corrections at the ωB97X-V/aTZ level are <1 kcal/mol, consistent with more rapid basis-set convergence of DFT as compared to MP2. Counterpoise-corrected ωB97X-V/aTZ interaction energies agree with (counterpoise-corrected) RIMP2/CBS results to within 1 kcal/mol, as shown in Fig. 4, suggesting that these two methods are reliable levels of theory for relative energies of $(H_2O)_{20}$ isomers. In Sec. IV B 2, we examine how well *n*-body approximations can reproduce these benchmarks.

2. Accuracy of the (G)MBE

In Fig. 5, we examine the errors in three- and four-body versions of MBE(*n*)+MBCP(*n*) as compared to counterpoise-corrected RIMP2/CBS and ωB97X-V/aTZ benchmarks. We find that MBE(2)+MBCP(2) performs quite poorly, with errors >40 kcal/mol, so this approach will not be discussed here. In contrast, the GMBE(2)+GMBCP(2) performs well but the GMBCP(2) correction is very demanding in terms of the number of subsystem calculations that are required, hence we have only used this method in the aTZ basis set, and will not extrapolate to the CBS limit.

In each of the MBE(3)+MBCP(3) calculations shown in Fig. 5, the MAE is >1 kcal/mol, and this three-body approach consistently underestimates the interaction energies in the two dodecahedral isomers of $(H_2O)_{20}$, as compared to those in the other three structural motifs. (In a previous study, we noted that BSSE effects are quite different for the dodecahedrons as compared to the other families of isomers.[23]) This is somewhat disturbing, in that it suggests that a three-body expansion might compare favorably to benchmark calculations in one region of the potential surface, only to

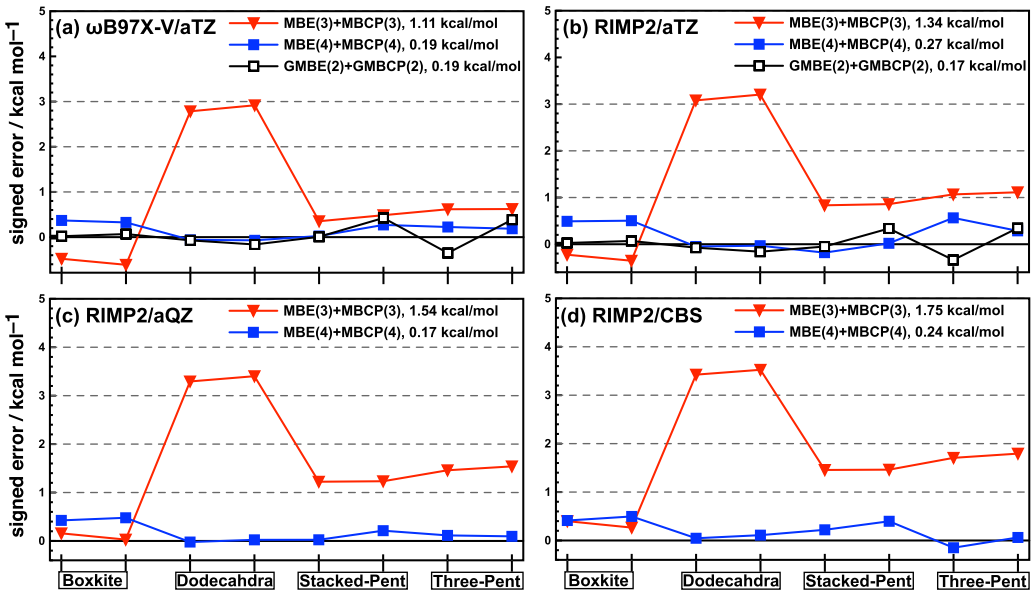

FIG. 5. Signed errors in relative energies of $(H_2O)_{20}$ isomers at the indicated levels of theory, using the MBE(*n*)+MBCP(*n*) approach for $n = 3$ and 4. MAEs for each fragmentation method are also shown, with respect to supersystem benchmarks computed at the same level of theory. All supersystem benchmarks are counterpoise-corrected, hence the MBE approximation with MBCP correction, or GMBE with GMBCP correction, is the appropriate comparison.

perform in a less-favorable manner for other isomers of the same cluster. On the other hand, the MBE(4)+MBCP(4) and GMBE(2)+GMBCP(2) methods perform consistently well for all four classes of $(H_2O)_{20}$ isomers, with MAEs <0.3 kcal/mol.

In an attempt to separate the role of BSSE corrections from the performance of the underlying n-body approximation to the total, uncorrected interaction energy, we plot in Fig. 6 the errors in MBE(3), MBE(4), and GMBE(2) interaction energies for the aTZ basis set, the aQZ basis set, and the CBS extrapolation, without applying counterpoise corrections to either the supersystem calculation or the n-body calculation. (This is therefore a test of how well a BSSE-contaminated n-body expansion reproduces the results of a BSSE-contaminated supersystem calculation.) In the same figure, we also isolate the BSSE calculations and compare the MBCP(3), MBCP(4), and GMBCP(2) corrections to the "full" counterpoise correction, the latter being defined as the difference between traditional counterpoise-corrected and uncorrected supersystem calculations.

Previously, Ouyang *et al.*[32] have suggested that, for reasons related to BSSE, MBE calculations should be performed using either the full cluster basis set (which is intractable in large systems) or else using basis sets of at least aug-cc-pVTZ quality, and these are the only basis sets considered in Fig. 6. RIMP2 results indicate that the accuracy of MBE(n) increases systematically as the basis set is improved, especially for $n = 4$, whereas the errors start small for GMBE(2) calculations—even as compared to MBE(4)—and are not substantially different between aTZ, aQZ, and CBS. GMBE(2) is also the best-performing of these three methods at the ωB97X-V level.

Likewise for the counterpoise corrections, the GMBCP(2) errors are smaller than those observed for MBCP(3) or MBCP(4) in both the RIMP2 and the DFT calculations. Because GMBE(2), without counterpoise correction, agrees well with BSSE-contaminated supersystem results, while GMBCP(2) agrees well with the supersystem counterpoise correction, we conclude that the composite GMBE(2)+GMBCP(2) method offers good accuracy for the right reasons, and does not rely on error cancellation. In contrast, MBE(3) tends to underestimate the BSSE-contaminated interaction energy while MBCP(3) overestimates the counterpoise correction, while at the four-body level the reverse is true, consistent with the oscillatory nature of the MBE. In these cases, the composite MBE(n)+MBCP(n) method is relying on error cancellation between the neglect of higher-order terms in the MBE and spurious BSSE effects. This example is a more incisive analysis of a similar error cancellation that we first noted in Ref. 23.

In studies of non-covalent clusters, relative energies are usually more important than total interaction energies, the latter of which are only measurable in small clusters. Having demonstrated that error cancellation plays a pivotal role in the accuracy of MBE(n)+MBCP(n) calculations, it is conceivable that this could be turned into a feature, by exploiting error cancellation to obtain high accuracy at affordable cost. As such, we report MAEs for relative energies of the eight $(H_2O)_{20}$ clusters in Table IV, where we also compare them to MAEs in the total interaction energies.

Interestingly, the error statistics for total interaction energies are not significantly different from those for relative isomer energies, vindicating the use of the former as the metric by which MBEs have been assessed in our work and many previous studies. (We rationalize this observation in terms of the fact that the total interaction energy is, from a certain point of view, merely the relative energy between two very different points on the potential energy surface, namely, a deep well versus an exit channel.) The MBE(3)+MBCP(3) approach does not lead to sub-kcal/mol accuracy, for which the four-body analogue is required, or alternatively one

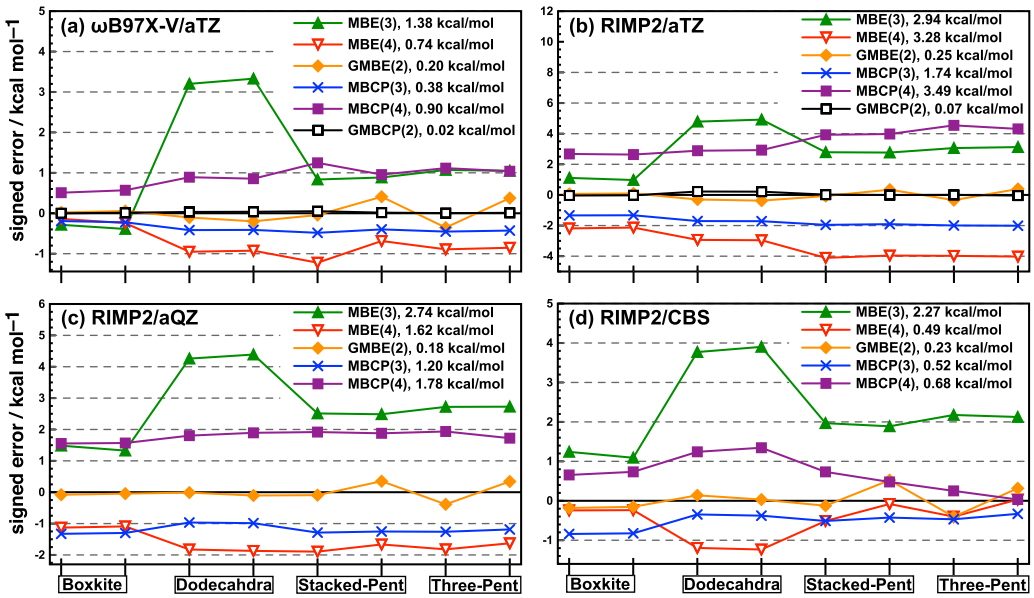

FIG. 6. Signed errors in relative energies of $(H_2O)_{20}$ isomers at the indicated levels of theory. Errors in MBE(n) and GMBE(2) results are defined relative to a supersystem calculation that has *not* been corrected for BSSE, and errors in the MBCP(n) and GMBCP(2) counterpoise corrections are defined relative to the "full" (supersystem) counterpoise correction. Note that the vertical energy scales are different in each panel.

TABLE IV. MAEs in relative energies of $(H_2O)_{20}$ isomers and in total interaction energies for the same isomers.

	MAE (kcal/mol)		
Method	MBE(3)+ MBCP(3)	MBE(4)+ MBCP(4)	GMBE(2)+ GMBCP(2)
	— Interaction energy —		
ωB97X-V/aTZ	1.11	0.19	0.19
RIMP2/aTZ	1.34	0.27	0.17
RIMP2/aQZ	1.54	0.17	...
RIMP2/CBS	1.75	0.24	...
	— Relative energy —		
ωB97X-V/aTZ	1.03	0.16	0.45
RIMP2/aTZ	1.05	0.42	0.41
RIMP2/aQZ	1.01	0.16	...
RIMP2/CBS	0.98	0.40	...

can bypass the need for four-body calculations and use GMBE(2)+GMBCP(2) instead, whose favorable performance does not appear to rest on error cancellation.

More exhaustive testing is required in order to obtain a comprehensive understanding of which subsystem levels of theory can be expected to perform well as compared to high-level supersystem benchmarks. Ideally, one would like to establish a protocol (in terms of the choice of embedding, the level of truncation, the size of the fragments, and any BSSE corrections) such that the accuracy of a given electronic structure model for calculations in small clusters would be in some way indicative of its accuracy when applied to large clusters via the MBE. This is needed especially since RIMP2/CBS, for example, is a reasonable level of theory for water clusters but would not be reasonable for applications to systems where a sizable fraction of the interaction energy comes from dispersion. The same can be said for the M11-based[64] density-functional models that are used for water clusters in Ref. 30, whose success for dispersion-bound systems appears to rest on a delicate cancellation of errors.[65] Counterpoise-corrected ωB97X-V/aTZ, on the other hand, appears to be accurate across a broad spectrum of noncovalent interactions,[51–53] and the MBE(4)+MBCP(4) and GMBE(2)+GMBCP(2) approximations to this supersystem method accurately reproduce both total interaction energies and relative isomers energies for water clusters. Given a sufficient number of cores over which to distribute such calculations, the time-to-solution can be made quite small.

C. Computational cost

It is often tacitly assumed that fragment-based methods are always less expensive than the corresponding supersystem calculations, but in terms of aggregate computer time (rather than time-to-solution or "wall time"), this is often not the case.[17] In this section, we consider the cost of the (G)MBE calculations reported here, with the caveat that we have made no attempt to perform any sort of thresholding, by means of which subsystem calculations involving spatially distant monomers might be neglected or approximated. Preliminary tests suggest that a significant fraction of the subsystem calculations are negligible in many cases, and that this fraction increases with n, but since the purpose of this work is to evaluate the intrinsic accuracy of the (G)MBE we retain all subsystems. Given the cancellation of errors that is often inherent in applications of the MBE,[9] it seems wise to establish exact (G)MBE(n) benchmarks before proceeding to discard terms.

Table V summarizes the number of subsystem calculations required for a MBE(4)+MBCP(4) or GMBE(2)+GMBCP(2) calculation, for both $(H_2O)_{20}$ and $(H_2O)_{55}$. For $N = 20$, GMBE(2) affords a modest reduction in the number of subsystem calculations, as compared to MBE(4), but the reduction is quite dramatic for $N = 55$. This is especially true for the counterpoise correction, which is far more expensive than a simple calculation of the supersystem energy. For $N = 55$, use of GMBCP(2) to approximate $E_I^{IJK\ldots N}$ generates 205 dimers (with 5–8 ghost molecules each) and 1430 intersections (3–5 ghost molecules). Since there are N separate monomer energies to correct, the full GMBCP(2) calculation consists of 11 275 dimers and 78 650 intersections. In contrast, for MBCP(4) there are 1 364 220 tetramers (with three ghost molecules each), with a large number of smaller subsystems as well.

The most time-consuming calculations for MBE(4)+MBCP(4) are tetramers in the tetramer basis set and monomers in the tetramer basis set, which are approximately equally expensive at the DFT level because they require all the same electron repulsion integrals. (Although the cost

TABLE V. Number of subsystem calculations required for several different fragment-based approaches, for two different cluster sizes.

	$(H_2O)_{20}$				$(H_2O)_{55}$			
Size	MBE(4)[a]	MBCP(4)[b]	GMBE(2)[a]	GMBCP(2)[b]	MBE(4)[a]	MBCP(4)[b]	GMBE(2)[a]	GMBCP(2)[b]
$n = 4$	4845	19 380	...	...	341 055	1 364 220	...	...
$n = 3$	1140	3 420	...	...	26 235	78 705	...	...
$n = 2$	190	380	...	...	1 485	2 970	...	...
$n = 6$–9	...	...	150	1 110	...	...	1 469	11 275
$n = 4$–6	...	...	4113	16 040	...	...	17 883	78 650
Total	6175	23 180	4263	17 150	368 775	1 445 895	19 352	89 925

[a]Calculations involving n monomers in an n-mer basis.
[b]Calculations involving one monomer in an n-mer basis.

analysis changes somewhat for correlated wave functions, we will see below that timing profiles for DFT and RIMP2 are actually rather similar.) For MBE(4)+MBCP(4) and $N = 20$ there are over 1.7×10^6 terms requiring a tetramer basis set, whereas the full GMBE(2)+MBCP(2) calculation requires fewer than 110 000 subsystem calculations. On the other hand, the subsystems are larger in the latter case, and for $(H_2O)_{20}$ at the ωB97X-V/aTZ level, the average computer time per subsystem job is 601 s for MBE(4)+MBCP(4) but 5534 s for GMBE(2)+GMBCP(2). These figures reflect the total aggregate computing time across all processors rather than wall time. In practice, we carry out these calculations across as many as 500 processors, a number that is limited by our available resources rather than by a lack of scalability, since the fragment-based calculation should scale well at least to the regime where the number of processors is comparable to the number of subsystem calculations required.

Figure 7 shows actual timing data for one ωB97X-V/aTZ or RIMP2/aTZ calculation on $(H_2O)_{20}$, where we have separated the cost of the total energy calculation from that of the counterpoise correction. In terms of total computer time aggregated across all processors, a MBE(4) calculation is 8–9 times more expensive than simply performing a calculation on the entire cluster, depending on the level of theory, and a GMBE(2) calculation is up to 74 times more expensive! From another point of view, however, these ratios (which are listed in Table VI) provide a rough estimate of the number of processors that would be required in order to make the time-to-solution the same for both the supersystem and the fragment-based calculation, assuming that all jobs are run in serial mode. Looking at it in this way, the highly accurate GMBE(2) method will outperform the supersystem approach already on fewer than 100 processors.

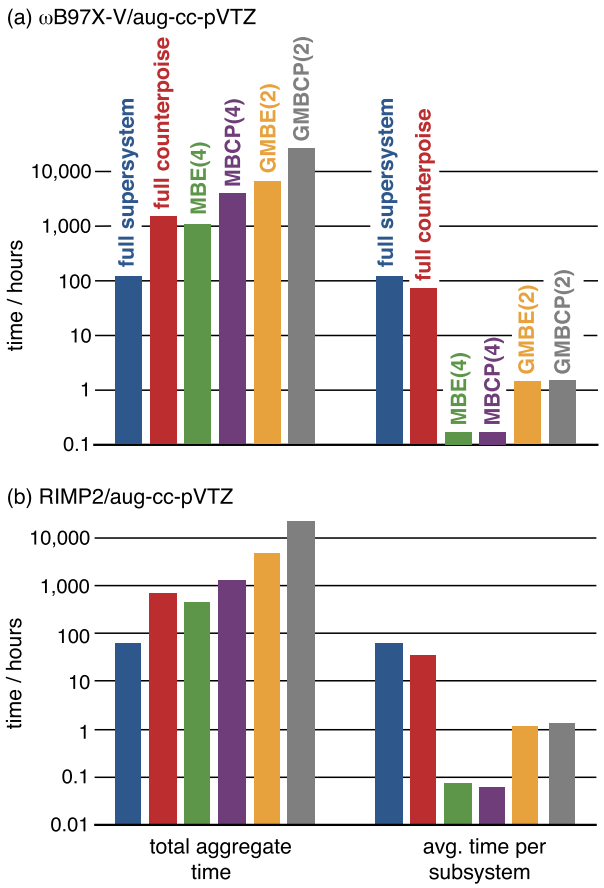

FIG. 7. Total computer time (summing all processors) for one isomer of $(H_2O)_{20}$, presented on a logarithmic scale. On the left are data for (a) counterpoise-corrected ωB97X-V/aTZ calculations and (b) RIMP2/aTZ calculations. On the right, these timings are divided by the number of subsystem calculations that is required in each case. The bars labeled "full supersystem" represent a calculation on the entire cluster, which does not decompose into trivially parallelizable subsystems, although the corresponding "full counterpoise" correction *can* be sub-divided by a factor of $N = 20$, and this reduction is reflected in the "full counterpoise" bar in the data sets on the right. Supersystem calculations were multithreaded across 20 cores, whereas (G)MBE subsystem calculations were each run in serial but distributed across as many as 500 processors.

TABLE VI. Ratio of the total aggregate computing time for a given fragment-based calculation to that required for the supersystem calculation that it is meant to approximate, rounded up to the nearest integer. The system is the same $(H_2O)_{20}$ isomer used to generate the timing data in Fig. 7.

Fragment method	ωB97X-V/aTZ	RIMP2/aTZ
MBE(4)	9	8
MBCP(4)	3	2
GMBE(2)	53	74
GMBCP(2)	18	34

As a measure of how favorable the wall times could be made, given the availability of a very large number of processors, the data on the right side of Fig. 7 show the average time per subsystem calculation. (We obtain this simply by dividing the total aggregate computer time by the total number of subsystem calculations from Table V, reasoning that the largest and most expensive subsystem calculations are also the most numerous. The full Boys-Bernardi counterpoise correction can be trivially parallelized into $N = 20$ separate calculations, and the timings on the right side of Fig. 7 reflect this.) Speed-ups approaching factors of 700–800 for MBE(4), and 60–80 for GMBE(2), are possible, although to realize these will require use of a number of processors comparable to the number of subsystem calculations, i.e., $\sim 10^4$ or so. Scalability should be good into this regime, meaning that (G)MBE-based calculations are good applications for peta- and exa-scale computing.

V. CONCLUSION

We report a systematic study of the accuracy of truncated MBEs as applied to water clusters, $(H_2O)_{N=6-55}$. Elaborating upon a previous study of finite-precision problems associated with these methods,[9] we demonstrate that for systems containing 30–40 fragments or more, results from three- and four-body expansions are quite sensitive to the numerical thresholds used in the subsystem calculations, such that common default settings in an electronic structure program may afford supersystem energies that differ by several kcal/mol as compared to results obtained using very tight thresholds. The same is not true of the supersystem calculation itself, but arises from the very large number of subsystem energies

that must be summed. It is therefore important to use tight thresholds in all subsystem calculations in order to minimize propagation of errors and ensure that the result is not an artifact of roundoff error.

These problems are dramatically worse for the four-body expansion than they are for the three-body expansion, which is unfortunate because only the former appears to achieve what has been called "dynamical accuracy" for cluster calculations,[24] defined as 10% of mean thermal energy at room temperature, or 0.09 kcal/mol per monomer. Although atom-centered embedding point charges are often assumed to accelerate convergence of the MBE, we find no compelling evidence that this is true at the three- or four-body level. (In previous three-body calculations we showed that density embedding, as used in the fragment molecular orbital method,[66,67] does not improve the results relative to simple point-charge embeddings.[3]) Two-body results are highly erratic, depending on the particular choice of embedding charges.

In contrast, a *generalized* two-body expansion based on overlapping fragments containing 3–4 water molecules per fragment achieves an accuracy of 0.02 kcal/mol per monomer, without resorting to embedding charges. The fact that high accuracy can be achieved without charge embedding is significant for several reasons. First, precision problems associated with computing the Coulomb self-energy of the embedding charges present a serious challenge to any implementation of the MBE that uses a script or driver program that is external to the quantum chemistry program itself, because the magnitude of this term can easily exceed 10^5 hartree in examples such as the ones considered here, so that six- or eight-digit roundoff in the output files of the electronic structure program can make a significant difference.[9] Second, applications to more heterogeneous clusters probably require that the embedding charges be iterated to self-consistency, a process that destroys the variational nature of the subsystem SCF calculations and significantly complicates the formulation of analytic energy gradients for the MBE.[67,68] Finally, the use of embedding charges leads to dramatic variations in the accuracy of the MBE from one basis set to the next, especially in the presence of diffuse functions where overpolarization problems are severe.

Having established the general characteristics of the (G)MBE using modest levels of theory, we examined isomers of $(H_2O)_{20}$ at several more respectable levels, namely RIMP2/CBS and counterpoise-corrected ωB97X-V/aTZ. To achieve high accuracy, one must deal with BSSE present in the n-body calculations, and we have introduced many-body counterpoise corrections for both the traditional MBE (in Refs. 12 and 23) and for the GMBE (in the present work). The three-body approach with three-body counterpoise corrections, MBE(3)+MBCP(3), fails to achieve 1 kcal/mol accuracy, which can be achieved using the corresponding four-body calculation. As above, however, the GMBE approach—GMBE(2)+GMBCP(2), in this case—performs well at the two-body level, with a MAE <0.3 kcal/mol as compared to counterpoise-corrected supersystem benchmarks. Separate analysis of the counterpoise corrections themselves suggests that this success does not rest on error cancellation, whereas for MBE(4)+MBCP(4) some error cancellation is indeed involved, wherein BSSE compensates for five-body and higher-order terms that are neglected.

Together, this study and a previous one[9] raise serious concerns about the efficacy, or at least the generality, of the n-body expansion. These difficulties arise not only from the factorial increase in the number of subsystem calculations, with respect to both n and N, but also from the fact that to maintain consistent precision, tighter SCF convergence thresholds and arbitrary precision floating-point arithmetic are required. Non-monotone convergence with respect to n, coupled with size-extensive errors in total interaction energies, means that higher-order terms become more important as system size grows. Together, these problems certainly detract from the "free lunch" reputation of fragment-based methods. On the other hand, GMBE(2)+GMBCP(2) has a number of advantages, as suggested above, and in conjunction with the ωB97X-V/aTZ level of theory offers an accurate and stable method for application to large systems, which can be competitive in cost on only a few hundred processors. Performance will improve given a thresholding procedure for discarding irrelevant subsystem calculations *a priori*, and we hope to report on systematic tests of such a procedure in the future.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences under Award No. DE-SC0008550. Calculations were performed at the Ohio Supercomputer Center under Project No. PAA-0003.[69] K.U.L. acknowledges support from a Presidential Fellowship awarded by The Ohio State University. J.M.H. is a Camille Dreyfus Teacher-Scholar and a fellow of the Alexander von Humboldt Foundation.

[1]M. S. Gordon, D. G. Fedorov, S. R. Pruitt, and L. V. Slipchenko, Chem. Rev. **112**, 632 (2012).
[2]L. D. Jacobson, R. M. Richard, K. U. Lao, and J. M. Herbert, Annu. Rep. Comput. Chem. **9**, 25 (2013).
[3]R. M. Richard, K. U. Lao, and J. M. Herbert, Acc. Chem. Res. **47**, 2828 (2014).
[4]M. A. Collins and R. P. Bettens, Chem. Rev. **115**, 5607 (2015).
[5]W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).
[6]E. Prodan and W. Kohn, Proc. Natl. Acad. Sci. U. S. A. **102**, 11635 (2005).
[7]C. A. Coulson, *Valence*, 2nd ed. (Oxford University Press, 1961).
[8]C. J. Cramer, *Essentials of Computational Chemistry: Theories and Methods* (Wiley, New York, 2002).
[9]R. M. Richard, K. U. Lao, and J. M. Herbert, J. Chem. Phys. **141**, 014108 (2014).
[10]R. M. Richard and J. M. Herbert, J. Chem. Phys. **137**, 064113 (2012).
[11]R. M. Richard and J. M. Herbert, J. Chem. Theory Comput. **9**, 1408 (2013).
[12]R. M. Richard, K. U. Lao, and J. M. Herbert, J. Phys. Chem. Lett. **4**, 2674 (2013).
[13]T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, New York, 2000).
[14]S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).
[15]G. S. Tschumper, "Reliable electronic structure computations for weak non-covalent interactions in clusters," in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and T. R. Cundari (Wiley-VCH, 2009), Vol. 26, Chap. 2, pp. 39–90.

[16]M. Kamiya, S. Hirata, and M. Valiev, J. Chem. Phys. **128**, 074103 (2008).
[17]J. Liu and J. M. Herbert, J. Chem. Theory Comput. **12**, 572 (2016).
[18]W. Li, S. Li, and Y. Jiang, J. Phys. Chem. A **111**, 2193 (2007).
[19]S. Hua, W. Hua, and S. Li, J. Phys. Chem. A **114**, 8126 (2010).
[20]K. Wang, W. Li, and S. Li, J. Chem. Theory Comput. **10**, 1546 (2014).
[21]F. B. van Duijneveldt, J. G. C. M. van Duijneveldt-van de Rijdt, and J. H. van Lenthe, Chem. Rev. **94**, 1873 (1994).
[22]P. Valiron and I. Mayer, Chem. Phys. Lett. **275**, 46 (1997).
[23]R. M. Richard, K. U. Lao, and J. M. Herbert, J. Chem. Phys. **139**, 224102 (2013).
[24]J. F. Ouyang and R. P. A. Bettens, J. Chem. Theory Comput. **11**, 5132 (2015).
[25]G. J. O. Beran, J. Chem. Phys. **130**, 164115 (2009).
[26]E. E. Dahlke and D. G. Truhlar, J. Chem. Theory Comput. **3**, 1342 (2007).
[27]U. Góra, R. Podeszwa, W. Cencek, and K. Szalewicz, J. Chem. Phys. **135**, 224102 (2011).
[28]N. J. Mayhall and K. Raghavachari, J. Chem. Theory Comput. **7**, 1336 (2011).
[29]H. W. Qi, H. R. Leverentz, and D. G. Truhlar, J. Phys. Chem. A **117**, 4486 (2013).
[30]J. Friedrich, H. Yu, H. R. Leverentz, P. Bai, J. I. Siepmann, and D. G. Truhlar, J. Phys. Chem. Lett. **5**, 666 (2014).
[31]S. Kazachenko and A. J. Thakkar, J. Chem. Phys. **138**, 194302 (2013).
[32]J. F. Ouyang, M. W. Cvitkovic, and R. P. A. Bettens, J. Chem. Theory Comput. **10**, 3699 (2014).
[33]P. M. W. Gill, B. G. Johnson, and J. A. Pople, Chem. Phys. Lett. **209**, 506 (1993).
[34]E. E. Dahlke and D. G. Truhlar, J. Chem. Theory Comput. **3**, 46 (2007).
[35]H. R. Leverentz and D. G. Truhlar, J. Chem. Theory Comput. **5**, 1573 (2009).
[36]F. L. Hirshfeld, Theor. Chem. Acc. **44**, 129 (1977).
[37]A. E. Reed, R. B. Weinstock, and F. Weinhold, J. Chem. Phys. **83**, 735 (1985).
[38]C. M. Breneman and K. B. Wiberg, J. Comput. Chem. **11**, 361 (1990).
[39]A. V. Marenich, S. V. Jerome, C. J. Cramer, and D. G. Truhlar, J. Chem. Theory Comput. **8**, 527 (2012).
[40]Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kús, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, Jr., H. Dop, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, P. A. Pieniazek, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, N. Sergueev, S. M. Sharada, S. Sharmaa, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, V. Vanovschi, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhou, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xua, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, Mol. Phys. **113**, 184 (2015).
[41]See supplementary material at http://dx.doi.org/10.1063/1.4947087 for structures of the $(H_2O)_{20}$ isomers that we consider, for benchmark interaction energies for these clusters, and for comparisons to other fragment-based quantum chemistry results for various water clusters.
[42]D. J. Wales and M. P. Hodges, Chem. Phys. Lett. **286**, 65 (1998).
[43]M. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).
[44]R. A. Kendall and H. A. Früchtl, Theor. Chem. Acc. **97**, 158 (1997).
[45]F. Weigend, M. Häser, J. Patzelt, and R. Ahlrichs, Chem. Phys. Lett. **294**, 143 (1998).
[46]R. A. DiStasio, Jr., R. P. Steele, Y. M. Rhee, Y. Shao, and M. Head-Gordon, J. Comput. Chem. **28**, 839 (2007).
[47]F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
[48]D. M. Bates, J. R. Smith, T. Janowski, and G. S. Tschumper, J. Chem. Phys. **135**, 044123 (2011).
[49]K. U. Lao and J. M. Herbert, J. Chem. Phys. **139**, 034107 (2013); Erratum **140**, 119901 (2013).
[50]R. O. Ramabhadran and K. Raghavachari, J. Chem. Theory Comput. **9**, 3986 (2013).
[51]K. U. Lao and J. M. Herbert, J. Phys. Chem. A **119**, 235 (2015).
[52]N. Mardirossian and M. Head-Gordon, Phys. Chem. Chem. Phys. **16**, 9904 (2014).
[53]K. U. Lao, R. Schäffer, G. Jansen, and J. M. Herbert, J. Chem. Theory Comput. **11**, 2473 (2015).
[54]P.-O. Löwdin, Int. J. Quantum Chem., Symp. **28**(S19), 19 (1986).
[55]D. Das, K. P. Eurenius, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodošček, and B. R. Brooks, J. Chem. Phys. **117**, 10534 (2002).
[56]P. H. König, M. Hoffmann, T. Frauenheim, and Q. Cui, J. Phys. Chem. B **109**, 9082 (2005).
[57]T. Vreven, K. S. Byn, I. Komáromi, S. Dapprich, J. A. Montgomery, K. Morokuma, and M. J. Frisch, J. Chem. Theory Comput. **2**, 815 (2006).
[58]G. A. Cisneros, J.-P. Piquemal, and T. A. Darden, J. Phys. Chem. B **110**, 13682 (2006).
[59]I. G. Kaplan, R. Santamaria, and O. Novaro, Mol. Phys. **84**, 105 (1995).
[60]B. Paulus, K. Rosciszewski, N. Gaston, P. Schwerdtfeger, and H. Stoll, Phys. Rev. B **70**, 165106 (2004).
[61]A. Hermann, R. P. Krawcyzk, M. Lein, P. Schwerdtfeger, I. P. Hamilton, and J. J. P. Stewart, Phys. Rev. A **76**, 013202 (2007).
[62]T. J. Mach and T. D. Crawford, Theor. Chem. Acc. **133**, 1449 (2014).
[63]M. P. Hodges, A. J. Stone, and S. S. Xantheas, J. Phys. Chem. A **101**, 9163 (1997).
[64]R. Peverati and D. G. Truhlar, J. Phys. Chem. Lett. **2**, 2810 (2011).
[65]N. Mardirossian and M. Head-Gordon, J. Chem. Theory Comput. **9**, 4453 (2013).
[66]D. G. Fedorov and K. Kitaura, J. Phys. Chem. A **111**, 6904 (2007).
[67]T. Nagata, D. G. Fedorov, and K. Kitaura, "Mathematical formulation of the fragment molecular orbital method," in *Linear-Scaling Techniques in Computational Chemistry and Physics*, edited by R. Zaleśny, M. G. Papadopoulos, P. G. Mezey, and J. Leszczynski, Challenges and Advances in Computational Chemistry and Physics (Springer, New York, 2011), Vol. 13, Chap. 2, pp. 17–64.
[68]T. Nagata, K. Brorsen, D. G. Fedorov, K. Kitaura, and M. S. Gordon, J. Chem. Phys. **134**, 124115 (2011).
[69]See http://osc.edu/ark:/19495/f5s1ph73 for Ohio Supercomputer Center.
[70]V. Ganesh, R. K. Dongare, P. Balanarayan, and S. R. Gadre, J. Chem. Phys. **125**, 104109 (2006).
[71]S. R. Gadre, K. V. J. Jose, and A. P. Rahalkar, J. Chem. Sci. **122**, 47 (2010).