

Basis Set Superposition Errors in the Many-Body Expansion of Molecular Properties

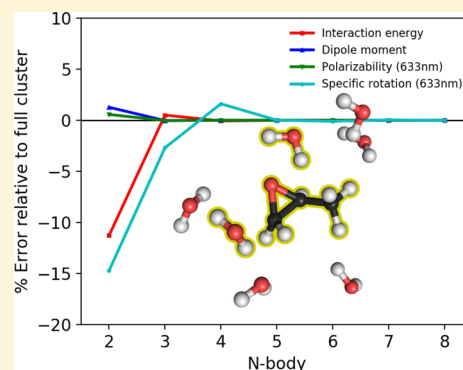
Published as part of *The Journal of Physical Chemistry virtual special issue "Leo Radom Festschrift"*.

Benjamin G. Peyton and T. Daniel Crawford*

Department of Chemistry, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, United States

Supporting Information

ABSTRACT: The underlying reasons for the poor convergence of the venerated many-body expansion (MBE) for higher-order response properties are investigated, with a particular focus on the impact of basis set superposition errors. Interaction energies, dipole moments, dynamic polarizabilities, and specific rotations are computed for three chiral solutes in explicit water cages of varying sizes using the MBE including corrections based on the site–site function counterpoise (or “full-cluster” basis) approach. In addition, we consider other possible causes for the observed oscillatory behavior of the MBE, including numerical precision, basis set size, choice of density functional, and snapshot geometry. Our results indicate that counterpoise corrections are necessary for damping oscillations and achieving reasonable convergence of the MBE for higher order properties. However, oscillations in the expansion cannot be completely eliminated for chiroptical properties such as specific rotations due to their inherently nonadditive nature, thus limiting the efficacy of the MBE for studying solvated chiral compounds.



1. INTRODUCTION

The accurate simulation of solvent effects on molecular optical response is exceptionally challenging because of the numerous dynamical factors required for a physically faithful model.^{1–9} Whereas many properties may be viewed as intrinsic to the solute subjected to a perturbation by the solvent environment (perhaps as represented by a dielectric continuum), others—particularly mixed electric-/magnetic-field responses^{10–15}—are, in fact, correctly viewed as inherent to the combined solute/solvent system. Schemes based on implicit solvent^{5,16–18} and frozen density embedding (FDE)^{19–23} have been used, but with limited success. Concomitant non-additivity effects, dynamic configurational sampling, and molecule-specific interactions serve to exacerbate the computational and theoretical demands on robust simulations.^{6,7} Thus, reducing the cost of these calculations for explicitly solvated systems has become a major concern.^{18,24–26}

The many-body expansion (MBE) formalism has seen widespread success in predicting the energies of large molecular systems at a reduced cost, though recently the limits of these successes have been tested with clusters of increasing size and complexity.^{27–30} These limits are bounded by well-explored challenges like loss of precision^{28,30,31} and basis set superposition error (BSSE).^{27,30–39} Electrostatic embedding (EE) has been applied to MBE treatments of water clusters alongside BSSE corrections to improve convergence,³⁷ as well as the N -body:many-body QM:QM technique⁴⁰ to reduce the cost of including higher-order

effects. The latter has also been applied to small water clusters for predicting vibrational frequencies.⁴¹

The fundamental concept behind the MBE is the decomposition of the energy (or other properties) into a sum of smaller contributions from subcomponents of a complex, interacting system.^{42,43} In the MBE, the energy (or energy derivative) of n interacting fragments may be expanded in orders of interaction energies:

$$E_{1,2,\dots,n} = \sum_{i \in N_1} E_i + \sum_{ij \in N_2} \epsilon_{ij} + \sum_{ijk \in N_3} \epsilon_{ijk} + \dots \quad (1a)$$

where N_1 , N_2 , N_3 , etc., denote the sets of unique monomers, dimers, trimers, etc. The interaction energies are defined as, for example, for dimers

$$\epsilon_{ij} = E_{ij} - E_i - E_j \quad (1b)$$

and for trimers

$$\epsilon_{ijk} = E_{ijk} - (E_{ij} + E_{ik} + E_{jk}) + (E_i + E_j + E_k) \quad (1c)$$

where the subtraction of energies of subcomponents is necessary to avoid overcounting of energy contributions. While the untruncated MBE scheme is formally exact, large computational savings result from the truncation of eq 1a to k -body terms where $k < n$. Truncation to two- or three-body terms only has proved sufficient for interaction energies in

Received: April 24, 2019

Published: May 6, 2019

many examples. However, limiting cases exist for increasingly complex systems.^{27–30}

While interaction energies are the popular target of MBE applications, properties computed using the MBE have received comparatively little attention. Though some work has been carried out concerning induced electronic properties in linear species,^{44–48} solvated systems such as those studied in ref 49 are of particular interest for chiroptical property prediction. A study by Mach and Crawford⁵⁰ showed that the MBE for a number of solvated chiral systems exhibited oscillatory convergence in specific rotation with respect to n -body truncation, while interaction energies, dipole moments, and dynamic polarizabilities still converged well by the three-body approximation as is typical for interaction energies with the MBE for systems with fewer than 12 monomers.^{27,30}

The successive sign-flips of k -body terms, such as E_i changing from positive to negative in eqs (1a) and (1b), have been suggested to be the cause of oscillations in the MBE for large clusters,²⁷ and BSSE has been identified as a major contributing factor to these oscillating errors. BSSE is a result of an imbalance of basis functions and thus will be present in any electronic structure calculation using a finite number of functions. The BSSE of a dimer interaction term in eq 1b is relatively straightforward to correct using the Boys and Bernardi counterpoise method (BBCP):³²

$$\epsilon_{ij} = E_{ij} - E_i^{(ij)} - E_j^{(ij)} \quad (2)$$

where the superscript denotes the basis set used, and any terms without a superscript are calculated in their own basis. When E_i and E_j are calculated in just their own basis sets, each monomer does not benefit from the nearby basis functions placed on the partner monomer. However, when calculating E_{ij} , the basis functions are shared between the monomers, which increases flexibility in the wave function and changes the energy despite having little bearing on the interaction between monomers i and j , only their (incomplete) basis sets. By using the same basis (that of the dimer ij) for all three calculations, there is no imbalance in the wave function, hence the term “counterpoise (CP) correction”.

The BBCP method was originally intended for correcting dimer interaction energies, but two fundamentally different generalizations of the BBCP correction scheme for an n -body interaction term are currently in use: the site–site function counterpoise (SSFC) and Valiron-Mayer function counterpoise (VMFC) methods.^{33,34} The former, also referred to as the “full cluster basis,” simply uses the basis functions of the full n -body cluster for each fragment calculation,

$$E_{1,2,\dots,n} = \sum_{i \subseteq N_1} E_i^{(n)} + \sum_{ij \subseteq N_2} \epsilon_{ij}^{(n)} + \sum_{ijk \subseteq N_3} \epsilon_{ijk}^{(n)} + \dots \quad (3)$$

with similar generalizations for eqs (1b) and (1c). While often prohibitively expensive, it has been shown to be effective in eliminating oscillations in the MBE.^{27,30}

The VMFC method is based on correcting for BSSE at each k -body level such that the k -mer basis is used for each k -body interaction term,

$$E_{1,2,\dots,n} \approx \sum_{i \subseteq N_1} E_i^{(i)} + \sum_{ij \subseteq N_2} \epsilon_{ij}^{(ij)} + \sum_{ijk \subseteq N_3} \epsilon_{ijk}^{(ijk)} + \dots \quad (4)$$

By correcting at each k -body term, the VMFC method prevents spurious “ghost dipoles” from appearing in the k -body approximations, which can occur with the SSFC method:

asymmetrically distributed basis functions placed around a real fragment can cause electron density to move to those locations, causing net dipoles (and possibly other properties) which are solely dependent on the placement of ghost functions, though such effects should only appear significantly for close-packed fragments. The VMFC method requires that each calculation be performed in multiple basis sets, resulting in many more calculations (an additional binomial coefficient, in fact) than the classic MBE. The exact nature of the MBE at n -body is also lost, e.g., the monomer energies in eq 1a no longer cancel exactly with the monomer energies in eq 1b, due to the difference in basis sets. The resulting energy is a “counterpoise corrected” energy, which makes benchmarking relative to complete cluster calculations difficult. In practice, results from the VMFC method vary only slightly relative to those from the SSFC method and other correction schemes.^{30,35,37}

The goal of the present study is to understand the cause of the oscillations in the MBE for specific rotation observed in the earlier work by Mach and Crawford⁵⁰ and to determine whether or not they can be removed using a BSSE correction scheme. We will focus on the SSFC method due to its convergence to the correct n -body limit. We have chosen to exclude the VMFC method at present because of its high computational cost and ambiguity in benchmarking calculations. We further extend the previous work by employing larger solvation clusters and testing additional configurations of explicitly solvated systems.

2. COMPUTATIONAL DETAILS

The dynamic polarizability at an external field frequency, ω , is computed as the isotropic average of the dynamic polarizability property tensor:^{10,12}

$$\alpha_{\alpha\beta}(\omega) = \frac{2}{\hbar} \sum_{j \neq n} \frac{\omega}{\omega_{jn}^2 - \omega^2} \text{Re}(\langle n | \mu_\alpha | j \rangle \langle j | \mu_\beta | n \rangle) \quad (5)$$

where n is the electronic ground state, j is an electronic excited state with excitation frequency ω_{jn} , and μ is the electric dipole operator. Similarly, the specific rotation (in $\text{deg dm}^{-1}(\text{g/mL})^{-1}$) is related to the isotropic average of the electric-dipole/magnetic-dipole property tensor (also called the Rosenfeld tensor):⁵¹

$$G'_{\alpha\beta}(\omega) = -\frac{2}{\hbar} \sum_{j \neq n} \frac{\omega}{\omega_{jn}^2 - \omega^2} \text{Im}(\langle n | \mu_\alpha | j \rangle \langle j | m_\beta | n \rangle) \quad (6)$$

The α and G' property tensors were computed in this work using the linear response formalism,⁵² the latter using gauge-including atomic orbitals (GIAOs).⁵³ Dynamic response properties such as polarizabilities or specific rotations may be formulated in terms of the time-averaged quasi-energy,

$$Q = \frac{1}{T} \int_0^T \langle \bar{\Psi} | \left(\hat{H} - i \frac{\partial}{\partial t} \right) | \bar{\Psi} \rangle dt \quad (7)$$

where $\bar{\Psi}$ is the regular part of the phase-isolated wave function⁵⁴ and $\hat{H}$ includes the time-dependent external field. For periodic potentials, the time-averaged quasi-energy is uniquely defined, and both variational and Hellmann–Feynman theorems apply.⁵⁵ Indeed, in the limit of a time-independent Hamiltonian, the quasi-energy reduces to the energy of the stationary state. Thus, just as the many-body expansion applies to the time-independent energy, it also applies to the time-averaged quasi-energy. Furthermore, since

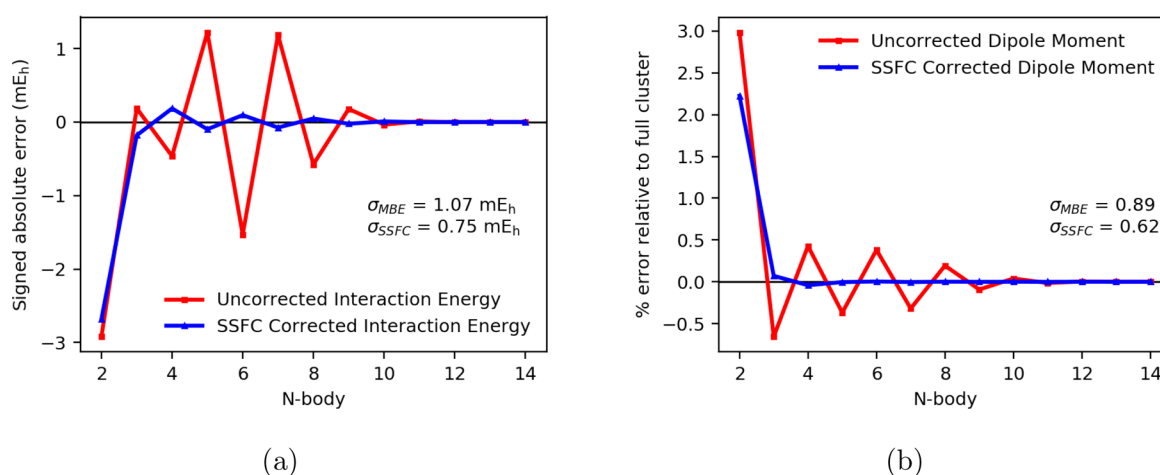


Figure 1. MBE and SSFC of (a) interaction energies and (b) dipole moments for (S)-methyloxirane in a 13-water solvent shell.

response functions are computed as derivatives of Q with respect to the perturbation coefficients (e.g., the polarizability is the negative of the second derivative of the time-averaged quasi-energy with respect to the strength parameter of the external electric field), response functions should also be subject to the many-body expansion, just as in the time-independent case. This was the driving force behind the Mach paper, which also showed that this was correct for one linear response property, the polarizability.⁵⁰ In the following, it will be shown that while this expansion is valid, its effective truncation depends highly on the additivity of the subjected property. CP corrections should remove BSSE errors which do not cancel in the MBE equations, and other potential issues will be explored for the nonconvergent specific rotation.

Geometries for (S)-methyloxirane in a cage of seven water molecules, (S)-methylthiirane surrounded by six water molecules, and (M)-dimethylallene with seven water molecules were reused from the previous study⁵⁰ for consistency. These geometries were generated by Gromacs⁵⁶ simulations of each solute in water, and snapshots with 5.5 Å solvent shells were extracted from the resulting trajectories. Additional geometries for (S)-methyloxirane in seven- and 13-water-molecule cages and (S)-methylthiirane in a six water-molecule cage were generated to test the effects of geometry and solvent shell size on MBE convergence and BSSE. All geometries are available in the Supporting Information.

Interaction energies, dipole moments, and specific rotations were calculated using the B3LYP^{57,58} functional in the aug-cc-pVDZ (aDZ) basis^{59,60} as in the previous study. Additionally, the aug-cc-pVTZ (aTZ) basis set was explored in selected examples to determine the effects of basis set size on the MBE convergence and the BSSE. The CAM-B3LYP⁶¹ functional was also employed to explore the effects of long-range interactions on the G' tensor. Dipole polarizabilities for all systems (and all properties for the 13-water/S-methyloxirane system) were computed using only CAM-B3LYP/aug-cc-pVDZ. Specific rotations and dynamic polarizabilities were computed at four common wavelengths: 355, 436, 589, and 633 nm. All calculations were performed using Gaussian 09⁶² using inputs generated by a Psi4⁶³ plugin, which also carried out the subsequent data collection and analysis. This plugin gathers the data from formatted checkpoint files generated by Gaussian 09 to address precision issues brought about by propagation of error as noted in recent studies.^{28,30,31} Default SCF and CPHF

convergence criteria (10^{-7} and 10^{-10} , respectively) were used with standard pruned integration grids, which include 75 and 35 radial shells, and 302 and 110 angular points per shell for SCF and CPHF, respectively. When compared to tighter convergence (10^{-12}) and “fine” grids (pruned 75 shell, 302 nodes) for both SCF and CPHF, no appreciable difference was found for the (S)-methyloxirane system in a seven-water solvent shell (compare Figures S7 and S13 of the Supporting Information). Additionally, unpruned “fine” grids were also tested with the same system; similarly, this had little effect on the convergence of the expansion (compare Figures S7 and S14 of the Supporting Information).

We will use both graphical evaluations and standard deviations of the error relative to the converged results to assess the oscillations present in the MBE. We calculated standard deviations only for the two-body and higher approximations, due to the much larger, nonrepresentative errors associated with the one-body approximation. We report standard deviations and percent errors for dynamic polarizabilities and specific rotations at 633 nm unless otherwise noted. Plots of data not discussed, such as the interaction energies and electric dipole moments of smaller solvent cages, are available in the Supporting Information.

3. RESULTS AND DISCUSSION

3.1. Interaction Energies and Electric Dipole Moments. Interaction energies and electric dipole moments computed with the MBE were reported in the previous study⁵⁰ and extensive reviews of the general behavior of the MBE for such properties are available elsewhere (e.g., ref 30 and references therein). We found that the SSFC correction was insignificant for all cases (see the Supporting Information), except for methyloxirane in a 13-water solvent shell, Figure 1. As noted in previous studies of ~ten-body expansions,^{27,30} oscillating errors in the convergence of the interaction energy in the millihartree range were observed for the methyloxirane system considered. While these oscillations are not as significant as those seen for other properties, the improvement in convergence by the SSFC correction as evidenced by Figure 1a is still noteworthy. Note that, by definition, the MBE and SSFC interaction energies are different due to the different basis sets used for the monomer terms, so the absolute errors relative to the MBE or SSFC interaction energy are reported. The dipole moment in Figure 1b also exhibits oscillations

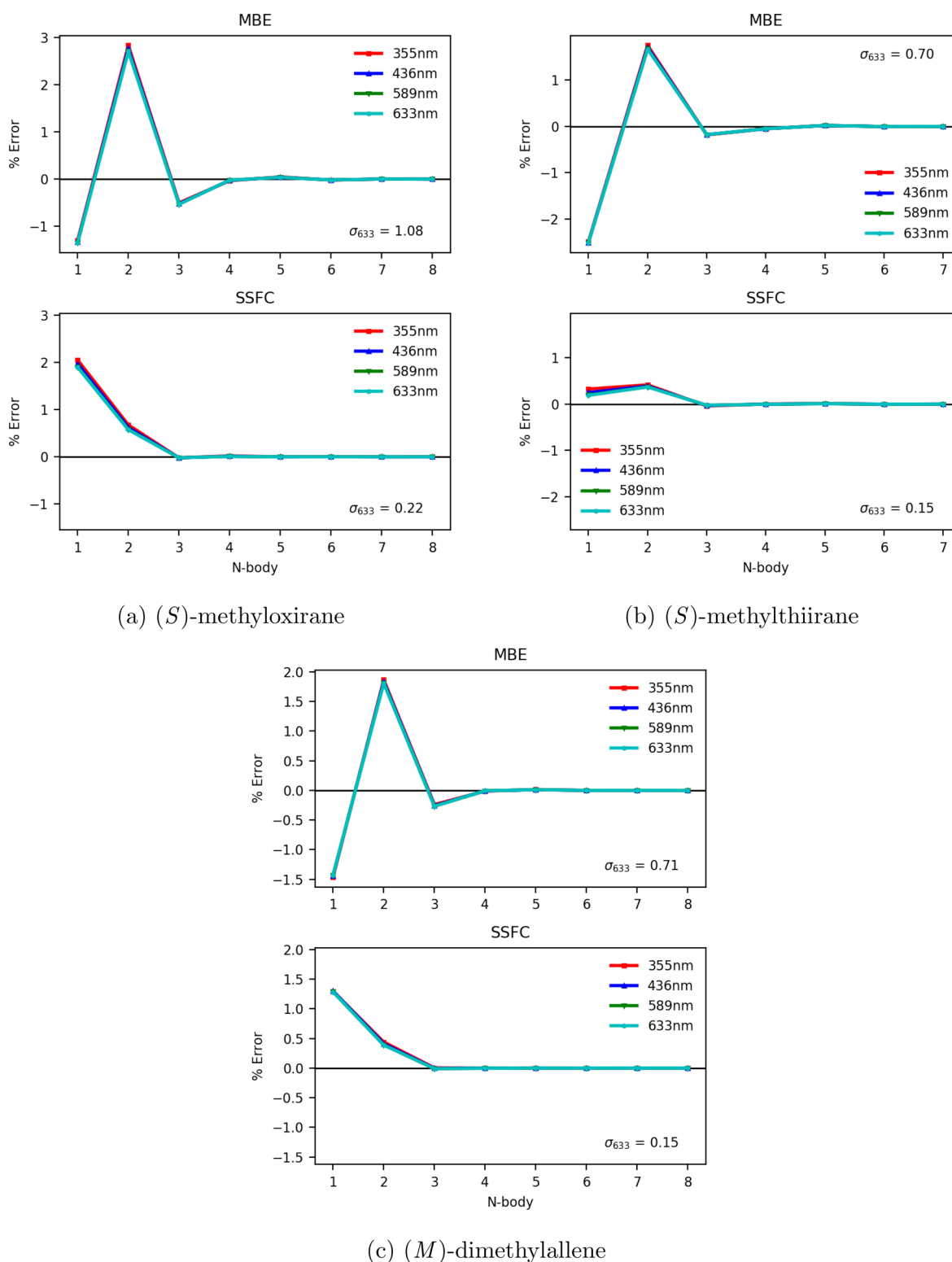


Figure 2. MBE and SSFC of dynamic polarizabilities for (a) (S)-methyloxirane in a seven-water solvent shell, (b) (S)-methylthiirane in a six-water solvent shell, and (c) (M)-dimethylallene in a seven-water solvent shell. Computed with CAM-B3LYP/aDZ.

uncharacteristic of the smaller solvent shells (nearly 0.5% error even at six-body), and again the SSFC correction dampens the oscillations dramatically. This suggests that the BSSE inherent in the MBE for interaction energies—and its characteristic increase relative to system size—is present for other properties as well. If the oscillations observed for higher-order properties

in the previous study are indeed indicative of BSSE, then CP corrections should play a major role in damping them.

3.2. Dynamic Dipole Polarizabilities. Dipole polarizabilities calculated with the MBE were reported for solvated (S)-methyloxirane, (M)-dimethylallene, and (S)-methylthiirane in the previous study⁵⁰ using the B3LYP functional. Their oscillatory behavior has been reproduced here with the CAM-

B3LYP functional and CP-corrected with the SSFC method, as shown in Figure 2. For all three systems, the relatively large oscillations of the MBE for one- to four-body contributions were almost completely removed by SSFC, replacing the convergence trend with a nearly monotonic decaying function. Analysis of a second snapshot taken from a molecular dynamics trajectory for (S)-methyloxirane and (S)-methylthiirane produce similar results (see the Supporting Information), with only that of (S)-methyloxirane exhibiting a slightly negative three-body term before reaching convergence. In all three systems, the SSFC correction yields substantial improvement in the convergence of the MBE as compared to the uncorrected results, e.g. the standard deviation for (S)-methyloxirane decreases from 1.08 to only 0.22.

Dynamic polarizabilities were also calculated for the larger 13-water (S)-methyloxirane system, as shown in Figure 3.

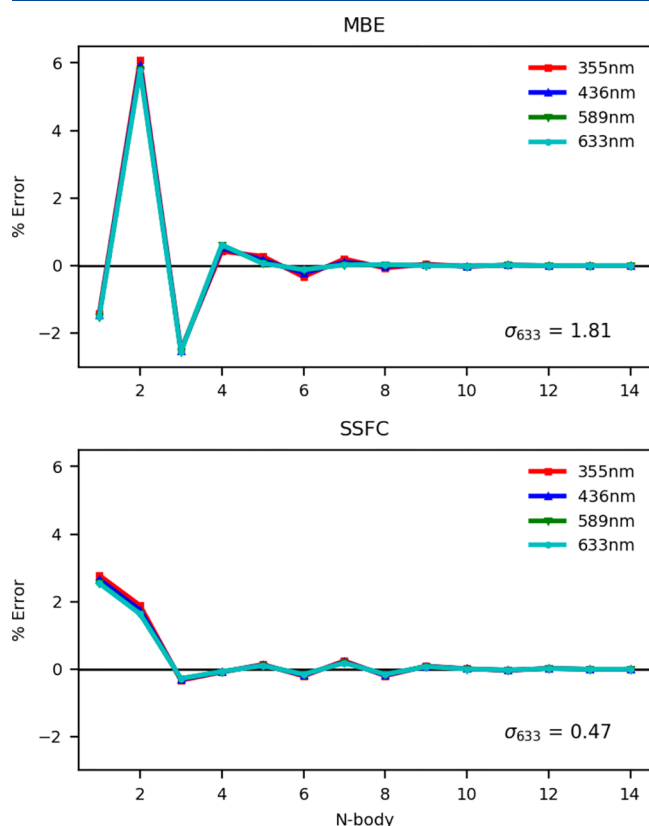


Figure 3. MBE and SSFC of dynamic polarizabilities for (S)-methyloxirane in a 13-water solvent shell. Computed with CAM-B3LYP/aDZ.

Oscillations to a higher error (up to 6%) resulted in a higher standard deviation in this case than for the smaller solvent systems considered for the MBE. However, the SSFC still dampened these oscillations significantly, with some small oscillations remaining. As with the electric dipole moment data considered earlier, BSSE is shown to be present and perhaps more important for properties than for interaction energies of the same system, with SSFC being a consistent method of reducing oscillations and speeding up convergence regardless of system size.

3.3. Specific Rotations. As stated previously, the properties considered thus far can be thought of as intrinsic to the solute being perturbed by the solvent. MBEs for such

properties, even beyond the energy, still converge, and oscillations in the convergence for large clusters are mostly due to BSSE and therefore can be corrected with CP-corrections such as SSFC or truncated schemes (such as VMFC(n)³⁶ or MBCP(n)³⁷). It is shown that BSSE is perhaps more important in response property calculations; however, CP corrections still recover convergence within a three body approximation for even the largest solvent cluster considered. Next, the oscillatory convergence of the highly nonadditive specific rotation of these clusters will be investigated, with the goal of deciding whether an MBE is an appropriate approximation even with a CP correction.

The earlier work by Mach and Crawford⁵⁰ identified specific rotations as a particular challenge for the MBE and its failure to converge with even small solvent shells suggested that BSSE was a potentially significant source of error. However, a number of other potential parameters exist that could also cause or exacerbate the observed erratic behavior for this property. As Ouyang et al. pointed out,²⁷ oscillations in the MBE can occur due to successive sign-flips of terms with large errors; as a result, oscillations could potentially occur due to any large bias in the subsystem calculations. To that end, we selected (S)-methyloxirane and (S)-methylthiirane as representative cases for investigating various strategies for addressing these oscillations, though we also report results for (M)-dimethylallene at the B3LYP/aDZ and CAM-B3LYP/aDZ levels of theory.

One possible source of convergence problems for the MBE is noise due to the precision of the collected G' property tensor elements. Recent publications^{28,30,31} have taken notice of the effects of subsystem calculation convergence criteria and precision on higher-order terms in the MBE. Richard, Lao, and Herbert demonstrated in ref 28 that precision errors in the energy rise rapidly for four- and five-body calculations as the total system size increases. We accounted for this by using Gaussian's formatted checkpoint files for the data collection to ensure that we maintain full machine precision until the final result. Nevertheless, significant oscillations in the MBE results for specific rotations remained with little to no difference from the previous results,⁵⁰ as shown in Figure 4. (NB the significant difference in the vertical scale for the three test cases.) Indeed, the specific rotation of (S)-methylthiirane approaches 500% error at the three-body contributions as compared to the converged results for the three longer wavelengths considered, in agreement with Table 8 of ref 50., and thus the precision of the individual terms is ruled out as a contributor to this behavior.

A second potential source of MBE convergence error is the use of the relatively small aDZ basis set. Although this basis set has been found to be adequate for many applications in studies of optical activity, for some systems (notably, methylthiirane), larger basis sets are needed.^{18,64} To investigate this issue, we carried out B3LYP/aTZ specific rotation calculations for both (S)-methyloxirane and (S)-methylthiirane, the results of which are presented in Figure 5. For (S)-methyloxirane, the four- and five-body errors with the aDZ basis (Figure 4a) exhibit a significant oscillation from $\sim 7\%$ error to $\sim -3\%$ error; the aTZ basis (Figure 5a) reduces this somewhat to $\sim 5\%$ to -0.7% error, and the two- to n -body standard deviation slightly decreases from 8.1 to 6.8% error. Comparing Figures 4b and 5b for (S)-methylthiirane, we note that increasing the basis set from aDZ to aTZ similarly offers no significant qualitative improvement, as errors vary wildly and are greater than 15%

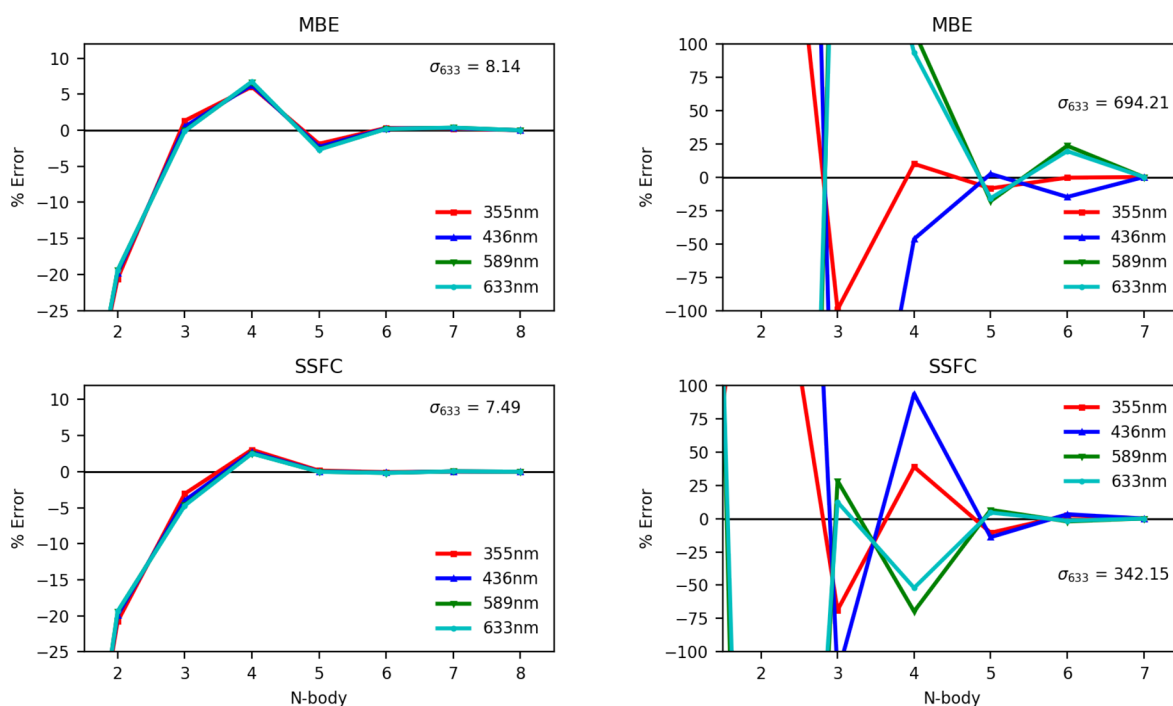
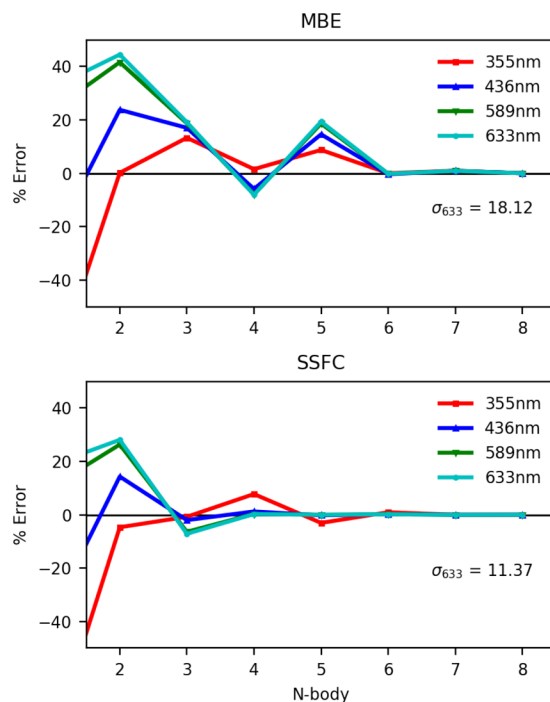
(a) (*S*)-methyloxirane(b) (*S*)-methylthiirane(c) (*M*)-dimethylallene

Figure 4. MBE and SSFC of specific rotation for (a) (*S*)-methyloxirane in a seven-water solvent shell, (b) (*S*)-methylthiirane in a six-water solvent shell, and (c) (*M*)-dimethylallene in a seven-water solvent shell. Computed with B3LYP/aDZ.

for all wavelengths even at five-body contributions (and worse than the aDZ basis in some cases). The standard deviation at 633 nm does decrease from 694 to 159% error with the improved basis set, though this is still well beyond acceptable limits and convergence with respect to *n*-body truncation remains strongly dependent on wavelength.

What about the choice of density functional? While most studies to date have utilized the popular B3LYP functional, CAM-B3LYP has been proven useful to reduce errors in the near-resonance regions^{8,65} and to produce more consistent results with respect to wavelength, as shown in Figure 6b. While CAM-B3LYP for (*S*)-methyloxirane does not yield significant differences in the MBE (cf. Figure 4a with Figure 7

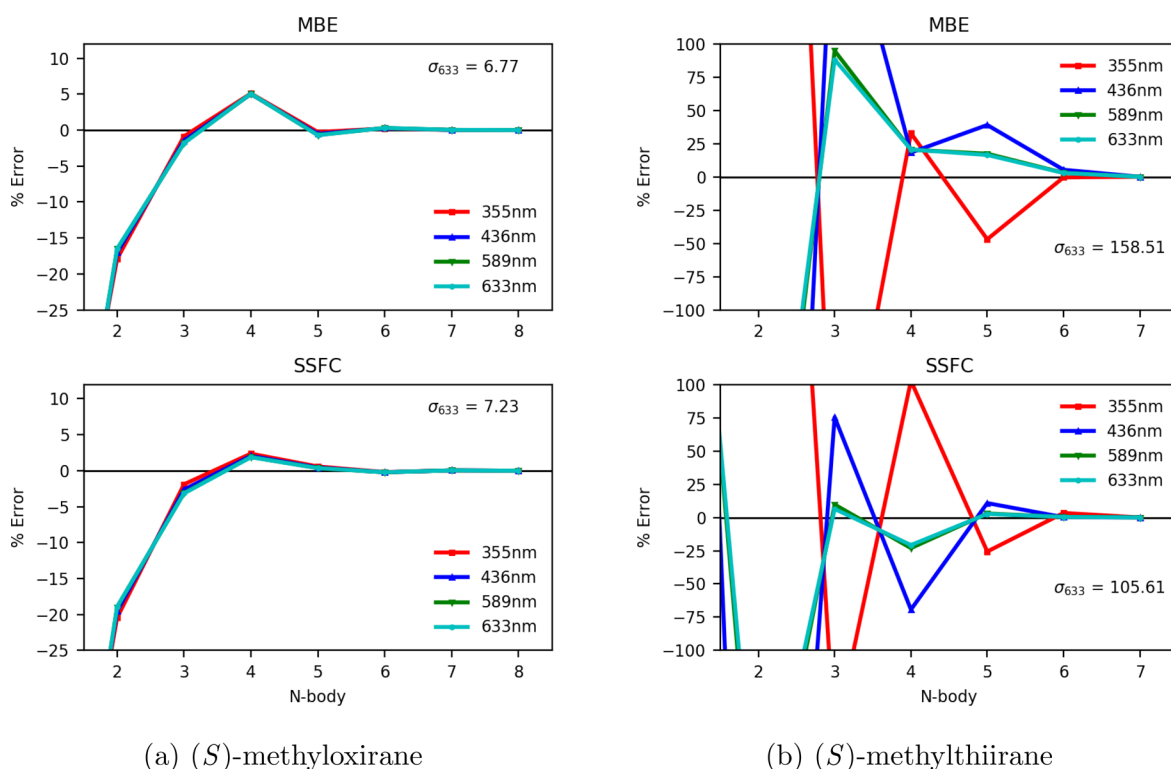


Figure 5. MBE and SSFC of specific rotation for (a) *(S)*-methyloxirane in a seven-water solvent shell and (b) *(S)*-methylthiirane in a six-water solvent shell. Computed with B3LYP/aTZ.

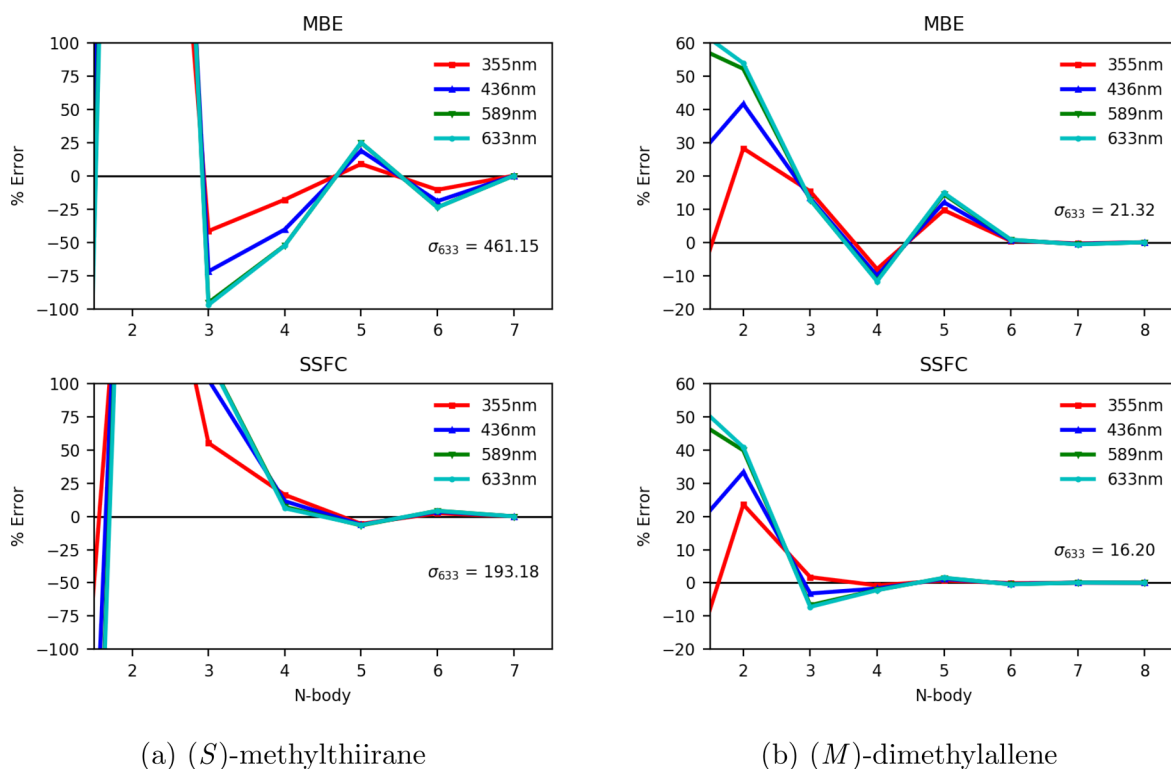


Figure 6. MBE and SSFC of specific rotation for (a) *(S)*-methylthiirane in a six-water solvent shell and (b) *(M)*-dimethylallene in a seven-water solvent shell. Computed with CAM-B3LYP/aDZ.

of the Supporting Information), the change in functional makes a considerable difference for *(S)*-methylthiirane (Figure 6a), as convergence trends were more consistent across all wavelengths tested. Nevertheless, the MBE still never

converges to below 5% error until the full expansion is reached, and a standard deviation of 461% error was still observed (primarily due to the two-body error which is over 1000% error). *(M)*-dimethylallene also exhibits more con-

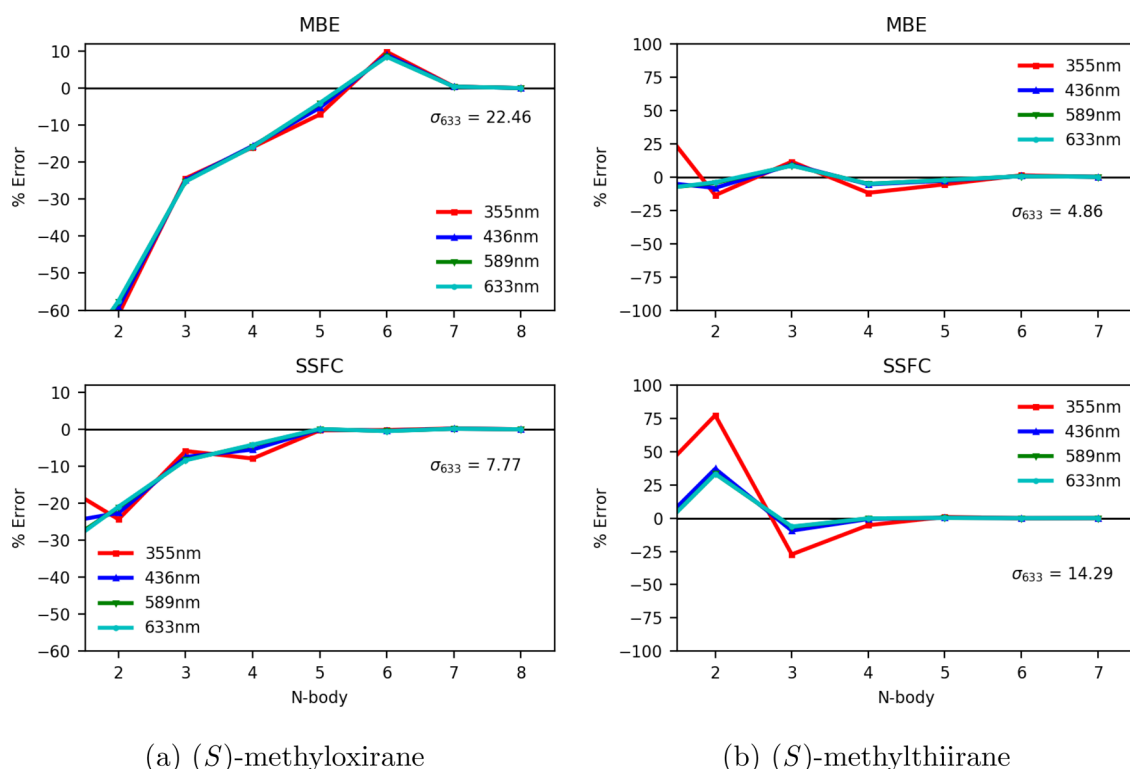


Figure 7. MBE and SSFC of specific rotation for additional snapshots of (a) (S)-methyloxirane in a seven-water solvent shell and (b) (S)-methylthiirane in a six-water solvent shell. Computed with CAM-B3LYP/aDZ.

sistent convergence trends with CAM-B3LYP, with only a slightly increased standard deviation (18% to 21% error).

The final non-BSSE issue considered is the choice of snapshot geometry. Calculations of solvent-phase specific rotation based on MD trajectories typically employ a large number of such snapshots, each with a distinct geometry that may be far from an energetic local minimum. Thus, we considered an additional snapshot for both (S)-methyloxirane and (S)-methylthiirane in their seven- and six-water solvation shells respectively, using CAM-B3LYP/aDZ (Figure 7). For (S)-methyloxirane, the MBE performs somewhat worse for the new snapshot, with errors approaching from far below zero and changing sign at the six-body level with errors of nearly 10% still remaining, and a considerably larger 22% error standard deviation, though the same qualitative trends appear. The new snapshot for (S)-methylthiirane, on the other hand, yields somewhat better results, but again with the same qualitative trends. The large errors near (and over) 100% seen previously are not present for this snapshot, though oscillations are still clearly visible with respect to n -body truncation, as evidenced by the 4.9% error standard deviation. Additionally, the original 7-water (S)-methyloxirane snapshot was edited to contain two short O...H bonds to test the effects of hydrogen bonding distance on the convergence (compare Figures S7 and S12 of the Supporting Information). As with the second snapshot, the same qualitative trends were observed, with very little difference in percent errors.

Next, in order to focus on the impact of BSSE on the MBE, we employed the SSFC CP-correction scheme in each of the test cases discussed above. In all cases, the corrections reduce—but do not eliminate—the oscillations in the relative errors of the MBE, as evidenced by the lower frame of every subfigure of Figures 4–7. Standard deviations also generally

agree with this trend, with exceptions being Figure 5a and 7b. In those two cases, the slightly higher two-body error in the SSFC case inflates the standard deviation, but from three-body onward the oscillations are reduced. For the smaller solvation shells of (S)-methyloxirane, (M)-dimethylallene, and (S)-methylthiirane, the convergence was also accelerated: by three- or four-body contributions, errors reasonable enough for predicting the sign of the optical rotation (<5%) were achieved for most wavelengths with CAM-B3LYP/aDZ, though (S)-methylthiirane remains a challenge (16% error at the four-body truncation for 355 nm), as well as 355 and 436 nm for the second methyloxirane snapshot (−8% and −6% at four-body). However, despite the relative improvements compared to the uncorrected MBE, these results do not imply that rapid convergence of even a CP-corrected MBE should be expected for highly nonadditive properties such as specific rotation. Furthermore, a qualitative description of the specific rotation will not suffice: proper simulations of the property should be averaged over possibly hundreds of snapshots from an MD trajectory, and for complex systems with multiple chiral centers even a small error in the specific rotation can predict the wrong stereoisomer.⁶⁶ The fact that these trends also hold for (M)-dimethylallene speaks to the generality of the conclusions made thus far. Chirality in this compound is induced by a stereogenic axis, rather than a stereogenic center such as that in methyloxirane. In addition, the lowest energy excited states for dimethylallene are $\pi \rightarrow \pi^*$ valence transitions localized near the double bonds, as opposed to methyloxirane whose lowest excitation energies are Rydberg states.⁶⁷

The 13-water solvated (S)-methyloxirane in Figure 8, however, presents an even larger challenge, despite including the SSFC correction. Encouragingly, errors slightly above 5%

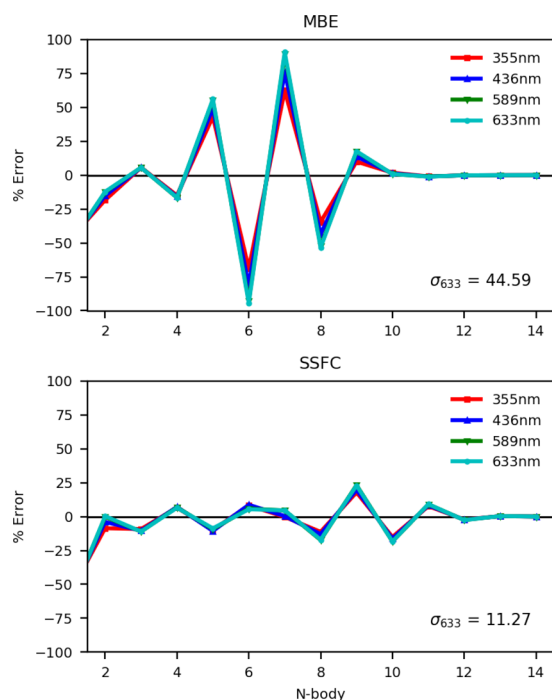


Figure 8. MBE and SSFC of specific rotation for (*S*)-methyloxirane in a 13-water solvent shell. Computed with CAM-B3LYP/aDZ.

were found at the four-body truncation, but then the error continues to oscillate without converging until the 12-body truncation. While the SSFC greatly outperforms the uncorrected MBE in the five- to seven-body range (the MBE presents errors between 50 and 100%, while SSFC presents

errors <10% in this range), it is actually worse from the nine-body truncation onward. The eight-, nine-, and ten-body truncations increase substantially to roughly $\pm 20\%$ error, while the MBE converged to <10% by nine-body. This is one of the largest solvated property calculations performed with the MBE to date, and the results undermine its usefulness. This, combined with the substantial cost of the SSFC correction, greatly hinders the feasibility of the approach for production-level calculations of specific rotation for such large clusters, which may be necessary to properly model the effects of the solvent.

In cases of achiral solvents, such as water, one way to reduce the expense of the MBE for larger solvent shells is to include only contributions due to the solute and solute–solvent interactions. This is equivalent to restricting the sums in eq 1a to only include fragments that contain the chiral solute. In principle, this should be advantageous for specific rotation calculations, as a pure achiral solvent (such as water) should give an optical rotation of zero upon averaging over a large number of snapshots and/or sufficiently large solvation clusters. Thus, neglecting solvent–solvent contributions should be of little consequence in the context of a sufficient MD trajectory while perhaps reducing the cost of the MBE through the removal of roughly half of the calculations needed for the full expansion of a solvated system (and much more for truncated expansions). For the smaller solvent shells of (*S*)-methyloxirane and (*S*)-methylthiirane (Figure 9), the removal of solvent–solvent interactions has little impact from five-body interactions onward. However, the four-body approximation is slightly worsened even with the SSFC correction for (*S*)-methylthiirane (though the standard deviation decreases). Using this approach, the total number of calculations necessary for the full expansion decreases from 255 to 128 unique

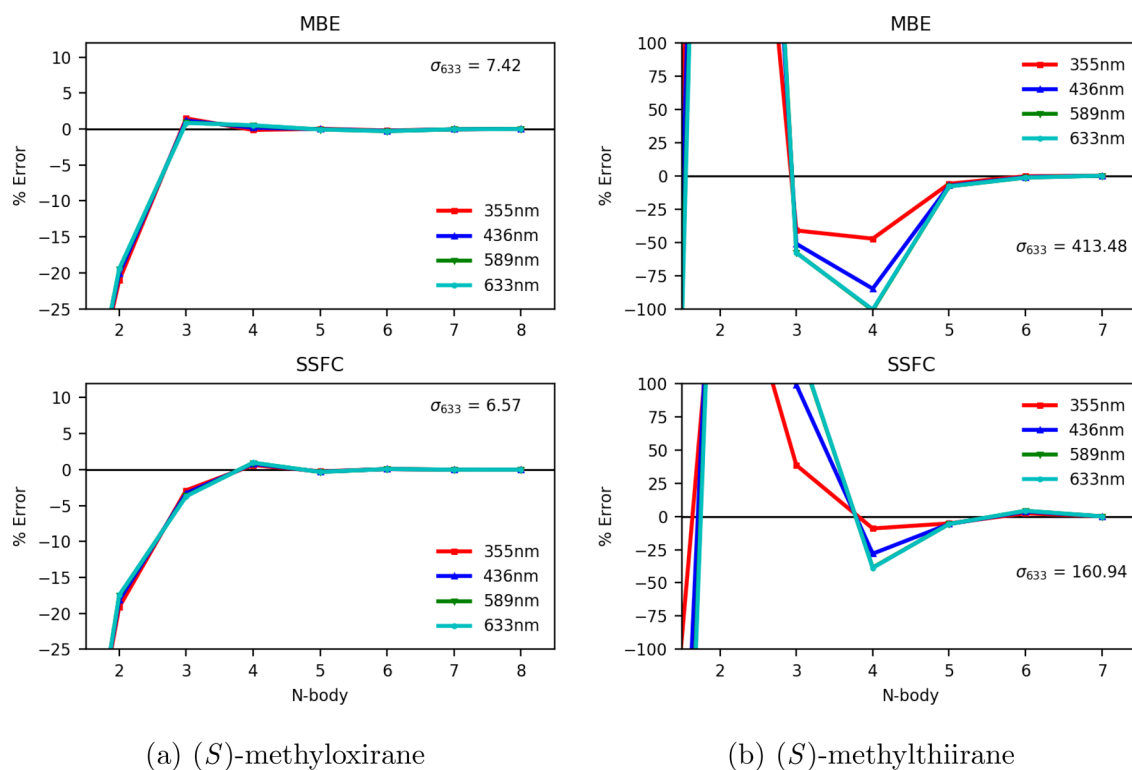


Figure 9. MBE and SSFC of solute-fragment-only specific rotation for (a) (*S*)-methyloxirane in a seven-water solvent shell and (b) (*S*)-methylthiirane in a six-water solvent shell. Computed with CAM-B3LYP/aDZ.

electronic structure calculations for (S)-methyloxirane, and, at any truncated n -body approximation, this number would decrease even more. For large systems in which only a qualitative description of the specific rotation is necessary, these errors may be acceptable for such an improvement in the computational cost.

In the case of the 13-water solvent shell for (S)-methyloxirane, excluding solvent–solvent interactions has a significant positive impact on the magnitude of oscillations in the MBE (Figure 10). While the standard deviation increases

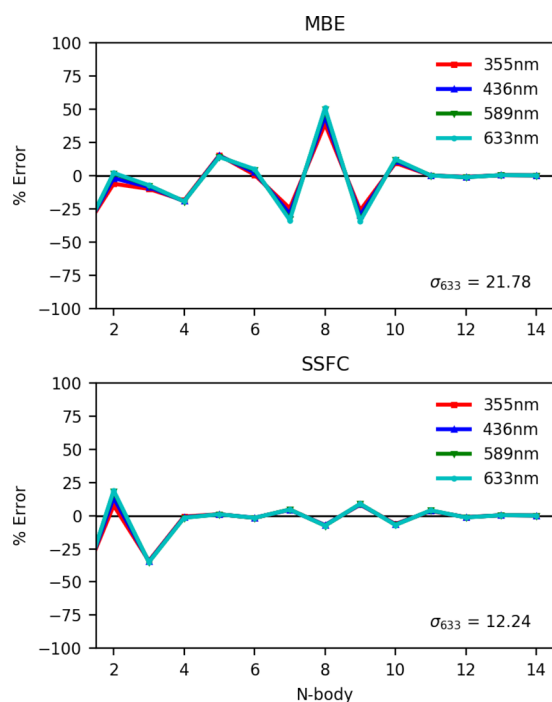


Figure 10. MBE and SSFC of solute-fragment-only specific rotation for (S)-methyloxirane in a 13-water solvent shell. Computed with CAM-B3LYP/aDZ.

for the SSFC, the convergence in the troublesome eight- to ten-body range is improved, as compared to Figure 8. Furthermore, the number of unique electronic structure calculations for the full expansion reduces by roughly a factor of 2—16383 to 8192—with massive savings possible for truncated expansions. Still, due to the significant non-BSSE oscillations in the MBE, the possibility of false convergence, and the prohibitive cost of the SSFC correction, a CP-corrected MBE cannot be recommended as a reliable low-cost quantitative approximation to the specific rotation of the full-cluster.

4. CONCLUSIONS

We find that the SSFC serves to reduce significantly the BSSE in the MBE of higher-order properties such as dipole polarizabilities and specific rotations, resulting in somewhat dampened oscillations and thus smoother convergence in most cases. Computed standard deviations decrease on average with inclusion of SSFC corrections, with the only exceptions being (S)-methylthiirane in a six-water solvent shell computed at the CAM-B3LYP/aDZ level of theory and (S)-methyloxirane in a seven-water solvent shell computed at the B3LYP/aTZ level of theory due to the erratic behavior of the MBE at the two-body correction. Furthermore, these trends held for variations in the

choice of MD snapshot, density functional, and basis set. We thus conclude that, in agreement with previous studies on interaction energies,^{27,30,31} counterpoise corrections are essential for the MBE calculation of response properties, whose magnitude (and sign) often depend heavily on the diffuse regions of electron density most prone to BSSE. However, the convergence difficulties characteristic of properties such as specific rotations remain a challenge, exhibiting oscillations in the MBE that cannot be remedied by a counterpoise correction alone. As illustrated clearly by the example of (S)-methyloxirane in a solvent shell of 13 water molecules, significant oscillations remain even after BSSE corrections, and these are inherent to nonadditive, non-perturbative properties such as mixed electric-/magnetic-field responses. Significant changes in the subsystem calculations as n increases can cause sign flips of successive terms, and small changes in the subsystem can cause large changes in the computed specific rotation contribution. In short, the rapid convergence of the MBE inherently assumes that the property in question is fundamentally additive, in this case that the property may be viewed as that of the solute with relatively small perturbations arising from the nearby solvent molecules. This requirement does not hold for specific rotations (and related chiroptical responses), leading to the conclusion that the MBE is of very limited utility for such cases.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpca.9b03864.

Interaction energies and dipoles for smaller solvent shells and convergence/grid pruning data as well as coordinates for each system studied (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*(T.D.C.) E-mail: crawdad@vt.edu.

ORCID

T. Daniel Crawford: 0000-0002-7961-7016

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by Grants CHE-1465149 and ACI-1450169 from the U.S. National Science Foundation. The authors gratefully acknowledge Advanced Research Computing at Virginia Tech for providing computational resources and technical support that have contributed to the results reported within the paper. The authors also acknowledge helpful discussions with Dr. Taylor Mach (Concordia University, St. Paul).

■ REFERENCES

- (1) Tomasi, J.; Persico, M. Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. *Chem. Rev.* **1994**, *94*, 2027–2094.
- (2) Gao, J. Hybrid QM/MM Simulations: An Alternative Avenue to Solvent Effects in Organic Chemistry. *Acc. Chem. Res.* **1996**, *29*, 298–305.
- (3) Cramer, C. J.; Truhlar, D. G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999**, *99*, 2161–2200.

- (4) Tomasi, J.; Cammi, R.; Mennucci, B.; Cappelli, C.; Corni, S. Molecular properties in solution described with a continuum solvation model. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5697–5712.
- (5) Mennucci, B.; Tomasi, J.; Cammi, R.; Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Gabriel, S.; Stephens, P. J. Polarizable continuum model (PCM) calculations of solvent effects on optical rotation of chiral molecules. *J. Phys. Chem. A* **2002**, *106*, 6102–6113.
- (6) Mukhopadhyay, P.; Zuber, G.; Goldsmith, M.-R.; Wipf, P.; Beratan, D. N. Solvent Effect on Optical Rotation: A Case Study of Methyloxirane in Water. *ChemPhysChem* **2006**, *7*, 2483–2486.
- (7) Mukhopadhyay, P.; Zuber, G.; Wipf, P.; Beratan, D. N. Contribution of a solute's chiral solvent imprint to optical rotation. *Angew. Chem., Int. Ed.* **2007**, *46*, 6450–6452.
- (8) Lipparini, F.; Egidi, F.; Cappelli, C.; Barone, V. The optical rotation of methyloxirane in aqueous solution: A never ending story? *J. Chem. Theory Comput.* **2013**, *9*, 1880–1884.
- (9) Egidi, F.; Giovannini, T.; Del Frate, C.; Lemler, P. M.; Vaccaro, P. H.; Cappelli, C. A combined experimental and theoretical study of optical rotatory dispersion for (R)-glycidyl methyl ether in aqueous solution. *Phys. Chem. Chem. Phys.* **2019**, *21*, 3644.
- (10) Barron, L. D. *Molecular Light Scattering and Optical Activity* **2004**.
- (11) Pecul, M.; Ruud, K. The Ab Initio Calculation of Optical Rotation and Electronic Circular Dichroism. *Adv. Quantum Chem.* **2005**, *50*, 185–212.
- (12) Crawford, T. D. Ab initio calculation of molecular chiroptical properties. *Theor. Chem. Acc.* **2006**, *115*, 227–245.
- (13) Crawford, T. D.; Tam, M. C.; Abrams, M. L. The current state of ab initio calculations of optical rotation and electronic circular dichroism spectra. *J. Phys. Chem. A* **2007**, *111*, 12057–68.
- (14) Crawford, T. D. High-Accuracy Quantum Chemistry and Chiroptical Properties. *Comprehensive Chiroptical Spectroscopy* **2012**, *1*, 675–697.
- (15) Crawford, T. D. Frontiers of Coupled Cluster Chiroptical Response Theory. *Front. Quantum Chem.* **2018**, 49–68.
- (16) Kongsted, J.; Pedersen, T. B.; Strange, M.; Osted, A.; Hansen, A. E.; Mikkelsen, K. V.; Pawłowski, F.; Jørgensen, P.; Hättig, C. Coupled cluster calculations of the optical rotation of S-propylene oxide in gas phase and solution. *Chem. Phys. Lett.* **2005**, *401*, 385–392.
- (17) Kongsted, J.; Pedersen, T. B.; Jensen, L.; Hansen, A. E.; Mikkelsen, K. V. Coupled cluster and density functional theory study of the vibrational contribution to the optical rotation of (S)-propylene oxide. *J. Am. Chem. Soc.* **2006**, *128*, 976–982.
- (18) Howard, J. C.; Sowndarya, S. V.; Ansari, I.; Mach, T. J.; Baranowska-Laczowska, A.; Crawford, T. D. On the Performance of Property-Optimized Basis Sets for Optical Rotation With Coupled Cluster Theory. *J. Phys. Chem. A* **2018**, *122*, 5962–5969.
- (19) Govind, N.; Wang, Y. A.; da Silva, A. J. R.; Carter, E. A. Accurate ab initio energetics of extended systems via explicit correlation embedded in a density functional environment. *Chem. Phys. Lett.* **1998**, *295*, 129–134.
- (20) Neugebauer, J.; Jacob, C. R.; Wesolowski, T. A.; Baerends, E. J. An Explicit Quantum Chemical Method for Modeling Large Solvation Shells Applied to Aminocoumarin C151. *J. Phys. Chem. A* **2005**, *109*, 7805–7814.
- (21) Jacob, C. R.; Neugebauer, J.; Visscher, L. A Flexible Implementation of Frozen-Density Embedding for Use in Multilevel Simulations. *J. Comput. Chem.* **2008**, *29*, 1011–1018.
- (22) Gomes, A. S. P.; Jacob, C. R.; Visscher, L. Calculation of local excitations in large systems by embedding wave-function theory in density-functional theory. *Phys. Chem. Chem. Phys.* **2008**, *10*, 5353–5362.
- (23) Crawford, T. D.; Kumar, A.; Hannon, K. P.; Höfener, S.; Visscher, L. Frozen-Density Embedding Potentials and Chiroptical Properties. *J. Chem. Theory Comput.* **2015**, *11*, 5305–5315.
- (24) Crawford, T. D. Reduced-Scaling Coupled-Cluster Theory for Response Properties of Large Molecules. *Recent Prog. Coupled Clust. Methods Theory Appl.* **2010**, 37–55.
- (25) McAlexander, H. R.; Crawford, T. D. A Comparison of Three Approaches to the Reduced-Scaling Coupled Cluster Treatment of Non-Resonant Molecular Response Properties. *J. Chem. Theory Comput.* **2016**, *12*, 209–222.
- (26) Kumar, A.; Crawford, T. D. Frozen virtual natural orbitals for coupled-cluster linear-response theory. *J. Phys. Chem. A* **2017**, *121*, 708–716.
- (27) Ouyang, J. F.; Cvitkovic, M. W.; Bettens, R. P. A. Trouble with the Many-Body Expansion. *J. Chem. Theory Comput.* **2014**, *10*, 3699–3707.
- (28) Richard, R. M.; Lao, K. U.; Herbert, J. M. Understanding the many-body expansion for large systems. I. Precision considerations. *J. Chem. Phys.* **2014**, *141*, 014108.
- (29) Ouyang, J. F.; Bettens, R. P. A. When are Many-Body Effects Significant? *J. Chem. Theory Comput.* **2016**, *12*, 5860–5867.
- (30) Richard, R. M.; Bakr, B. W.; Sherrill, C. D. Understanding the Many-Body Basis Set Superposition Error: Beyond Boys and Bernardi. *J. Chem. Theory Comput.* **2018**, *14*, 2386–2400.
- (31) Liu, K.-Y.; Herbert, J. M. Understanding the many-body expansion for large systems. III. Critical role of four-body terms, counterpoise corrections, and cutoffs. *J. Chem. Phys.* **2017**, *147*, 161729.
- (32) Boys, S.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553–566.
- (33) Wells, B. H.; Wilson, S. Van der Waals Interaction Potentials: Many-Body Basis Set Superposition Effects. *Chem. Phys. Lett.* **1983**, *101*, 429–434.
- (34) 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.
- (35) Salvador, P.; Szczesniak, 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.
- (36) Kamiya, M.; Hirata, S.; Valiev, M. Fast electron correlation methods for molecular clusters without basis set superposition errors. *J. Chem. Phys.* **2008**, *128*, 074103.
- (37) Richard, R. M.; Lao, K. U.; Herbert, J. M. Achieving the CCSD(T) Basis-Set Limit in Sizable Molecular Clusters: Counterpoise Corrections for the Many-Body Expansion. *J. Phys. Chem. Lett.* **2013**, *4*, 2674.
- (38) Ouyang, J. F.; Bettens, R. P. A. Many-body basis set superposition effect. *J. Chem. Theory Comput.* **2015**, *11*, 5132–5143.
- (39) Mayer, I.; Bakó, I. Many-Body Energy Decomposition with Basis Set Superposition Error Corrections. *J. Chem. Theory Comput.* **2017**, *13*, 1883–1886.
- (40) Hopkins, B. W.; Tschumper, G. S. A multicentered approach to integrated QM/QM calculations. Applications to multiply hydrogen bonded systems. *J. Comput. Chem.* **2003**, *24*, 1563–1568.
- (41) Howard, J. C.; Tschumper, G. S. N-body: Many-body QM: QM vibrational frequencies: Application to small hydrogen-bonded clusters. *J. Chem. Phys.* **2013**, *139*, 184113.
- (42) Hankins, D.; Moskowitz, J. W.; Stillinger, F. H. Water molecule interactions. *J. Chem. Phys.* **1970**, *53*, 4544–4554.
- (43) Kaplan, I. G. *Theory of Molecular Interactions*; Elsevier: Amsterdam, 1986.
- (44) Skwara, B.; Zawada, A.; Bartkowiak, W. On the Many-Body Components of Interaction-Induced Electric Properties: Linear Fluoroacetylene Trimer as a Case Study. *Comput. Lett.* **2007**, *3*, 175–182.
- (45) Skwara, B.; Bartkowiak, W.; Da Silva, D. L. On the basis set superposition error in supermolecular calculations of interaction-induced electric properties: Many-body components. *Theor. Chem. Acc.* **2009**, *122*, 127–136.
- (46) Baranowska, A.; Zawada, A.; Fernández, B.; Bartkowiak, W.; Kedziera, D.; Kaczmarek-Kedziera, A. Interaction-induced electric properties and cooperative effects in model systems. *Phys. Chem. Chem. Phys.* **2010**, *12*, 852–62.

- (47) Zawada, A.; Kaczmarek-Kędziera, A.; Bartkowiak, W. On the potential application of DFT methods in predicting the interaction-induced electric properties of molecular complexes. Molecular H-bonded chains as a case of study. *J. Mol. Model.* **2012**, *18*, 3073–3086.
- (48) Zawada, A.; Bartkowiak, W. Many-body interactions and the electric response of hydrogen-bonded molecular chains. *Comput. Theor. Chem.* **2011**, *967*, 120–128.
- (49) Leverentz, H. R.; Maerzke, K. A.; Keasler, S. J.; Siepmann, J. I.; Truhlar, D. G. Electrostatically embedded many-body method for dipole moments, partial atomic charges, and charge transfer. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7669–7678.
- (50) Mach, T. J.; Crawford, T. D. Computing optical rotation via an N-body approach. *Theor. Chem. Acc.* **2014**, *133*, 1–9.
- (51) Rosenfeld, L. Quantenmechanische Theorie der natürlichen optischen Aktivität von Flüssigkeiten und Gasen. *Eur. Phys. J. A* **1929**, *52*, 161–174.
- (52) Norman, P. A perspective on nonresonant and resonant electronic response theory for time-dependent molecular properties. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20519.
- (53) Wolinski, K.; Hinton, J. F.; Pulay, P. Efficient Implementation of the Gauge-Independent Atomic Orbital Method for NMR Chemical Shift Calculations. *J. Am. Chem. Soc.* **1990**, *112*, 8251–8260.
- (54) Langhoff, P. W.; Epstein, S. T.; Karplus, M. Aspects of Time-Dependent Perturbation Theory. *Rev. Mod. Phys.* **1972**, *44*, 602–644.
- (55) Helgaker, T.; Coriani, S.; Jørgensen, P.; Kristensen, K.; Olsen, J.; Ruud, K. Recent Advances in Wave Function-Based Methods of Molecular Property Calculations. *Chem. Rev.* **2012**, *112*, 543–631.
- (56) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845–854.
- (57) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.
- (58) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (59) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (60) Woon, D. E.; Dunning, T. H. J. Gaussian Basis Sets for Use in Correlated Molecular Calculations. IV. Calculation of Static Electrical Response Properties. *J. Chem. Phys.* **1994**, *100*, 2975–2988.
- (61) Yanai, T.; Tew, D. P.; Handy, N. C. A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (62) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; et al. *Gaussian 16* Revision A.03; 2016.
- (63) Parrish, R. M.; Burns, L. A.; Smith, D. G.; Simmonett, A. C.; DePrince, A. E.; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M.; et al. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197.
- (64) Mach, T. J.; Crawford, T. D. Basis Set Dependence of Coupled Cluster Optical Rotation Computations. *J. Phys. Chem. A* **2011**, *115*, 10045–10051.
- (65) Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. Excitation energies in density functional theory: An evaluation and a diagnostic test. *J. Chem. Phys.* **2008**, *128*, 044118.
- (66) Fuentes, R.; Pearce, K.; Du, Y.; Rakotondrarafa, A.; Valenciano, A. L.; Cassera, M. B.; Rasamison, V. E.; Crawford, T. D.; Kingston, D. G. I. Phloroglucinols from the Roots of *Garcinia dauphinensis* and Their Antiproliferative and Antiplasmodial Activities. *J. Nat. Prod.* **2019**, *82*, 431–439.
- (67) Tam, M. C.; Russ, N. J.; Crawford, T. D. Coupled cluster calculations of optical rotatory dispersion of (S)-methyloxirane. *J. Chem. Phys.* **2004**, *121*, 3550–7.