Efficient evaluation of triple excitations in symmetry-adapted perturbation theory via second-order Møller–Plesset perturbation theory natural orbitals

Edward G. Hohenstein and C. David Sherrill[a]

Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, and School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, USA

An accurate description of dispersion interactions is required for reliable theoretical studies of many noncovalent complexes. This can be obtained with the wave function-based formulation of symmetry-adapted perturbation theory (SAPT) provided that the contribution of triple excitations to dispersion is included. Unfortunately, this triples dispersion correction limits the applicability of SAPT due to its $\mathcal{O}(N^7)$ scaling. The efficiency of the evaluation of this correction can be greatly improved by removing virtual orbitals from the computation. The error incurred from truncating the virtual space is reduced if second-order Møller–Plesset perturbation theory (MP2) natural orbitals are used in place of the canonical Hartree–Fock molecular orbitals that are typically used. This approximation is further improved if the triples correction to dispersion is scaled to account for the smaller virtual space. If virtual MP2 natural orbitals are removed according to their occupation numbers, in practice, roughly half of the virtual orbitals can be removed (with the aug-cc-pVDZ basis set) with negligible errors if the remaining triples dispersion contribution is scaled. This typically leads to speedups of 15–20 times for the cases considered here. By combining the truncated virtual space with the frozen core approximation, the triples correction can be evaluated approximately 50 times faster than the canonical computation. These approximations cause less than 1% error (or at most 0.02 kcal mol^{-1}) for the cases considered. Truncation of greater fractions of the virtual space is possible for larger basis sets (leading to speedups of over 40 times before additional speedups from the frozen core approximation). © *2010 American Institute of Physics.*
[doi:10.1063/1.3479400]

I. INTRODUCTION

In many noncovalent interactions, dispersion is the primary attractive component. The stacking of nucleobases,[1–5] the interaction of rare gas atoms,[6–8] and the lattice energy of many organic crystals[9–11] are all dominated by dispersion interactions. Computationally, these interactions require a high level of electron correlation to describe properly. Low levels of theory, such as density functional theory or Hartree–Fock (HF), fail completely to capture long-range dispersion interactions. Second-order Møller–Plesset perturbation theory (MP2) is the simplest method that describes dispersion qualitatively; however, MP2 is known to significantly overestimate binding in many dispersion bound complexes.[12–14] Among standard electronic structure theory methods, the best for computing dispersion dominated interaction energies is coupled-cluster with singles, doubles, and perturbative triples [CCSD(T)].[15] A drawback of this method is its computational cost; CCSD(T) scales formally as $\mathcal{O}(N^7)$, which limits its applicability.

To gain a more complete understanding of the interactions within a noncovalent complex, it is often useful to examine the individual components of the interaction energy (electrostatics, exchange, induction, and dispersion). A supermolecular approach to compute interaction energies, such as CCSD(T), does not provide such a decomposition. The most rigorous method of obtaining these components separately is symmetry-adapted perturbation theory (SAPT).[16] SAPT is a perturbative approach to directly compute an interaction energy. Accurate interaction energies can be obtained from the wave function-based formulation of SAPT provided that the effect of triple excitations on the dispersion interaction is included. In this work, we will only discuss the wave function-based formulation of SAPT. The inclusion of triple excitations in SAPT is similar to the perturbative triples correction to CCSD and also scales as $\mathcal{O}(N^7)$.[17]

Recently, we have developed a SAPT program that uses the density-fitting (DF) approximation to evaluate the necessary two-electron integrals.[18] With this DF-SAPT program, we have performed the largest SAPT computations to date that account for the triples correction to dispersion.[19] While these advances have made SAPT computations much faster, the computations nevertheless remain very demanding if triples are included, and for such computations, our DF-SAPT program remains limited to systems roughly the size of a nucleic acid base pair (with a double-ζ basis set). The triples dispersion correction includes terms that scale as $\mathcal{O}(o^3 v^4)$ and $\mathcal{O}(o^4 v^3)$ (where o is the number of occupied

a)Electronic mail: sherrill@gatech.edu.

orbitals and v is the number of virtual or unoccupied orbitals). Since the number of virtual orbitals is usually much larger than the number of occupied orbitals (this is required for an accurate description of dispersion), the overall scaling will be $\mathcal{O}(o^3 v^4)$.

In this work, we apply approximations to the evaluation of the triples correction to dispersion that allow it to be evaluated more efficiently. Here, we use MP2 naturals orbitals (NOs)[20,21] instead of HF molecular orbitals (MOs) to evaluate the triples correction; this allows an appreciable fraction of the virtual orbitals to be removed from the computation without significant loss of accuracy. Natural orbitals are those orbitals that diagonalize the one-particle density matrix (OPDM).[20] For a two-electron system, natural orbitals comprise the basis that requires the fewest configurations to reach a given accuracy in the energy,[22] and, in general, the natural orbitals tend to concentrate the electron correlation energy into the those virtual NOs with the largest occupation numbers (one-particle density matrix eigenvalues). Conversely, the virtual NOs with the smallest eigenvalues contribute very little to the correlation energy and may be neglected. Natural orbitals have been used to select active spaces or as guess orbitals in multiconfigurational self-consistent-field (MCSCF) computations,[23–25] or as replacements for fully optimized MCSCF orbitals.[26,27] They have also been used in highly correlated configuration interaction computations[28–34] and coupled-cluster computations.[35–37] The optimized virtual orbital subspace approach of Adamowicz and Bartlett[38] is an alternative technique with the same goal of limiting the number of virtual orbitals for highly correlated computations; this approach has been reformulated by Urban and co-workers[39] and used to reduce the cost of CCSD(T) computations[40] including an impressive recent study of the benzene dimer.[41] A related approach to truncate the virtual space using pair natural orbitals[42–44] has been recently explored by Neese *et al.*[45,46] in the context of the coupled-electron pair approximation, CCSD, and quadratic configuration interaction with single and double excitations. Although shown to have many promising applications, these methods may not be well suited for the study of weakly interacting systems.[46]

Using the natural orbital cutoffs and the scaling procedure introduced here, we are able to neglect about half of the virtual NOs in the triples dispersion term in SAPT when using an aug-cc-pVDZ basis set, with an expected reduction in the computational time of approximately a factor of $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$. Larger fractions of the virtual NOs may be neglected for larger basis sets. In addition, we also employ the frozen core approximation [which has been previously applied to SAPT (Ref. 47)] allowing an additional two to three times the speedup in the triples correction. The combination of the frozen core and MP2 NO approximations in our DF-SAPT program allow the triples correction to dispersion to be evaluated much faster than in any previous implementation.

II. THEORETICAL METHODS

In SAPT, one defines a Hamiltonian for each monomer, H_A (for monomer A), as $H_A = F_A + W_A$, where F_A is the usual Fock operator for nuclei and electrons associated with monomer A and W_A is the fluctuation potential accounting for electron correlation on monomer A. An intermolecular operator V is defined as the difference between the full dimer Hamiltonian and the sum of two monomer Hamiltonians so that

$$H = F_A + W_A + F_B + W_B + V. \tag{1}$$

For a thorough description of SAPT, we refer the reader to Ref. 16. Dispersion interactions are at least second-order in V. Through second-order in W, the dispersion energy is given as

$$E_{\text{dispersion}} = E_{\text{disp}}^{(20)} + E_{\text{disp}}^{(21)} + E_{\text{disp}}^{(22)}, \tag{2}$$

where the first number in parentheses denotes the perturbation order in V and the second number in parentheses denotes the perturbation order in W. When three numbers appear in parentheses, the first number still denotes the perturbation order in V, while the second and third numbers denote perturbation order in W_A and W_B, respectively. Definitions and physical interpretations of the individual terms shown above are given in Ref. 17.

In this work we will focus on the $E_{\text{disp}}^{(22)}$ term; specifically, the triples contribution to this term, $E_{\text{disp}}^{(22)}(T)$. This is the most computationally demanding SAPT term through second-order in V and W. $E_{\text{disp}}^{(220)}(T)$ is given by[17]

$$E_{\text{disp}}^{(220)}(T) = \frac{(4 W_{r_1 r_2 s_1}^{a_1 a_2 b_1} - 2 W_{r_1 r_2 s_1}^{a_2 a_1 b_1}) W_{a_1 a_2 b_1}^{r_1 r_2 s_1}}{\epsilon_{a_1} + \epsilon_{a_2} + \epsilon_{b_1} - \epsilon_{r_1} - \epsilon_{r_2} - \epsilon_{s_1}}, \tag{3}$$

where

$$W_{r_1 r_2 s_1}^{a_1 a_2 b_1} = \nu_{r_2 s_1}^{r_3 b_1} t_{r_1 r_3}^{a_1 a_2} + \nu_{r_1 s_1}^{r_3 b_1} t_{r_3 r_2}^{a_1 a_2} - \nu_{a_3 s_1}^{a_2 b_1} t_{r_1 r_2}^{a_1 a_3} - \nu_{a_3 s_1}^{a_1 b_1} t_{r_1 r_2}^{a_3 a_2}$$
$$+ \nu_{r_1 r_2}^{r_3 a_2} t_{r_3 s_1}^{a_1 b_1} + \nu_{r_1 r_2}^{a_1 r_3} t_{r_3 s_1}^{a_2 b_1} - \nu_{a_3 r_2}^{a_1 a_2} t_{r_1 s_1}^{a_3 b_1} - \nu_{r_1 a_3}^{a_1 a_2} t_{r_2 s_1}^{a_3 b_1}. \tag{4}$$

Here, ϵ is an orbital energy, $\nu_{jl}^{ik} = (ij|kl)$, and $t_{r_1 r_2}^{a_1 a_2} = \nu_{r_1 r_2}^{a_1 a_2}/(\epsilon_{a_1} + \epsilon_{a_2} - \epsilon_{r_1} - \epsilon_{r_2})$. The indices a and r correspond to occupied and virtual orbitals of monomer A, respectively. Likewise, b and s correspond to occupied and virtual orbitals of monomer B, respectively. $E_{\text{disp}}^{(220)}(T)$ corresponds to the correction for the intramonomer electron correlation of monomer A. The correction for the intramonomer electron correlation of monomer B, $E_{\text{disp}}^{(202)}(T)$, can be found by interchanging the indices corresponding to monomer A with those for monomer B. The frozen core approximation can be applied to this correction by simply restricting the occupied indices to run over only active orbitals. Likewise, virtual orbitals can be removed by restricting the virtual indices.[47]

Reducing the number of virtual orbitals in the evaluation of $E_{\text{disp}}^{(22)}(T)$ is extremely beneficial due to its $\mathcal{O}(o^3 v^4)$ scaling. However, only a small number of virtual HF MOs can be removed before the accuracy of the computation is severely impacted. To avoid this problem, we will use MP2 NOs in place of the HF MOs. The (unrelaxed) MP2 OPDM is given as

$$P_{ij} = -2 \frac{[2(ia|kb) - (ib|ka)](ja|kb)}{D_{ik}^{ab} D_{jk}^{ab}}, \tag{5}$$

$$P_{ab} = 2 \frac{[2(ia|jc) - (ic|ja)](ib|jc)}{D_{ij}^{ac} D_{ij}^{bc}}, \tag{6}$$

where i, j, k correspond to occupied orbitals and a, b, c correspond to unoccupied orbitals and $D_{ij}^{ab} = \epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b$. Summation over repeated indices is implied. In an SAPT computation, the MP2 OPDM is used to compute the $E_{\text{elst}}^{(12)}$ term, so it will already be available. In this work, we will correlate all electrons in the MP2 OPDM formation regardless of whether or not they are correlated in the $E_{\text{disp}}^{(22)}(T)$ evaluation.

The equation for $E_{\text{disp}}^{(22)}(T)$ presented above assumes that the Fock matrix for each monomer is diagonal. MP2 NOs do not diagonalize the Fock matrix, so they must be modified before they can be used to evaluate $E_{\text{disp}}^{(22)}(T)$. Our procedure for generating a set of usable MP2 NOs is summarized below.

(1) Form the MP2 OPDM in the HF MO basis.
(2) Diagonalize the MP2 OPDM to obtain MP2 NOs (in the HF MO basis).
(3) Truncate the MP2 NO virtual space.
(4) Transform the MO based Fock matrix into the truncated MP2 NO basis.
(5) Diagonalize the NO based Fock matrix to obtain semicanonical MP2 NOs (in the MP2 NO basis) and orbital energies.
(6) Express the semicanonical MP2 NOs in terms of the AOs.
(7) Using the above NO/AO transformation matrix, transform the integrals needed to evaluate $E_{\text{disp}}^{(22)}(T)$ from the AO basis into the semicanonical MP2 NO basis and evaluate this term as usual.

The eigenvalues of the MP2 OPDM are occupation numbers that represent the number of electrons in each NO. These eigenvalues are a convenient metric for removing virtual orbitals from the computation. Since the unrelaxed MP2 OPDM is used, the occupied HF orbitals are recovered (the NO based Fock matrix is block diagonal); this is often called the "frozen natural orbital" procedure. If no virtual orbitals are removed, the virtual HF MOs are also recovered in the semicanonicalization process.

Our DF-SAPT program has been modified to compute MP2 NOs and use them in the evaluation of $E_{\text{disp}}^{(22)}(T)$.[18,19] This program has been developed within the framework of PSI 3.4.[48] All two-electron integrals are computed under the DF approximation. Also, all computations are performed in the dimer basis. The accuracy of the MP2 NO approximation to $E_{\text{disp}}^{(22)}(T)$ will be tested for the S22 test set of Hobza and co-workers.[49] We compute interaction energies for these complexes with the aug-cc-pVDZ basis set.[50,51] The corresponding aug-cc-pVDZ-RI basis set[52] is employed to approximate the two-electron integrals. Recently, we have shown that the aug-cc-pVDZ basis is sufficient to obtain accurate interaction energies (mean absolute deviation of 0.33 kcal mol^{-1}) for the S22 test set[49,53] when using second-
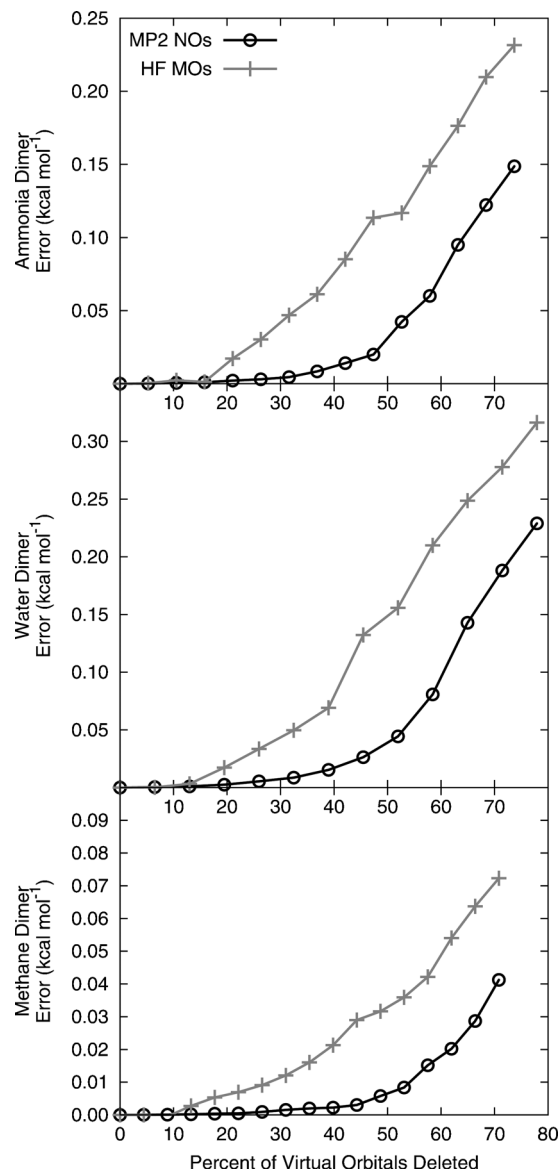

FIG. 1. Errors (in kcal mol^{-1}) of the $E_{\text{disp}}^{(22)}(T)$ correction evaluated with the aug-cc-pVDZ basis set as virtual orbitals (HF MO or MP2 NO) are removed from the computation. The total $E_{\text{disp}}^{(22)}(T)$ corrections for these test cases are -0.281 (ammonia dimer), -0.344 (water dimer), and -0.102 kcal mol^{-1} (methane dimer).

order SAPT including the $E_{\text{disp}}^{(21)}$ and $E_{\text{disp}}^{(22)}$ terms (which we have denoted by SAPT2+).[19] All of the geometries used in this work are taken from the S22 test set.[49]

III. RESULTS AND DISCUSSION

The use of MP2 NOs is based on the experience that NOs are more efficient than HF MOs for capturing electron correlation in the low-lying (or most heavily occupied) orbitals. This can be demonstrated in the present case for the evaluation of $E_{\text{disp}}^{(22)}(T)$ by removing virtual HF MOs and MP2 NOs at the same rate and determining how much error arises from the reduced virtual space. The results of such a test are shown in Fig. 1 for the ammonia dimer (95 virtual orbitals), water dimer (77 virtual orbitals), and methane dimer (113 virtual orbitals). The MP2 NOs are clearly superior to HF MOs for reproducing the $E_{\text{disp}}^{(22)}(T)$ correction with a smaller

virtual space. However, the error associated with removing virtual MP2 NOs still increases too rapidly to remove more than roughly one-third of the virtual orbitals if the error is to be kept negligible (greater fractions may be removed if small to modest errors may be tolerated). This is a great improvement over the HF MOs; only the few most high lying virtual HF MOs can be removed before significant errors begin to accrue.

As previously mentioned, the scaling of the $E_{disp}^{(22)}(T)$ correction is $\mathcal{O}(o^3 v^4)$. Assuming ideal behavior, removing one-third of the virtual orbitals would lead to roughly a five times speedup. If one-half of the virtual orbitals could be removed, it would result in a 16 times speedup. In Fig. 1, the error created by removing half of the NOs is below 0.05 kcal mol^{-1} [or less than 15% of the total $E_{disp}^{(22)}(T)$ contribution], which is probably acceptable in many or most applications. However, here we wish to explore ways in which we may achieve this level of computational savings while allowing truly negligible errors. To improve this approximation, we will assume that the magnitude of the $E_{disp}^{(22)}(T)$ correction changes at the same rate as $E_{disp}^{(20)}$ when virtual orbitals are removed, viz.,

$$\frac{E_{disp}^{(22)}(T)_{exact}}{E_{disp}^{(22)}(T)_{approx}} \approx \frac{E_{disp,exact}^{(20)}}{E_{disp,approx}^{(20)}}. \tag{7}$$

The $E_{disp}^{(20)}$ term scales as $\mathcal{O}(o^2 v^2)$, so for a given system, if it is possible to evaluate the $E_{disp}^{(22)}(T)$ correction, it is trivial to evaluate $E_{disp}^{(20)}$. We will denote the $E_{disp}^{(22)}(T)$ correction evaluated using this approximation as Est.$E_{disp}^{(22)}(T)$ to signify that the value is estimated, not computed explicitly. In order to determine whether or not this approximation is valid, $E_{disp}^{(22)}(T)$ and Est.$E_{disp}^{(22)}(T)$ are computed for the 11 smallest complexes from the S22 test set. Virtual orbitals are removed to determine how effective the removal of virtual orbitals will be for increasing the computational efficiency. The results of this analysis are shown in Fig. 2. The Est.$E_{disp}^{(22)}(T)$ correction performs much better than the unscaled correction. Simply removing a certain fraction of the virtual orbitals (as is the case in Fig. 2) can provide a 16 times speedup with only 1.25% error. The unscaled $E_{disp}^{(22)}(T)$ can only provide a three times speedup with the same accuracy. Clearly, the approximation of Eq. (7) significantly increases the number of virtual orbitals that can be removed while keeping the error negligible.

To this point, the number of virtual orbitals removed was not physically motivated. Each natural orbital has an occupation associated with it (eigenvalues of the MP2 OPDM). We will use these values as a metric to determine which virtual orbitals can be removed. In Fig. 3, again, $E_{disp}^{(22)}(T)$ and Est.$E_{disp}^{(22)}(T)$ are computed for the smallest 11 complexes in the S22 test set with different cutoffs based on the number of electrons in an orbital. Once again, the Est.$E_{disp}^{(22)}(T)$ correction is far superior to the unscaled correction. With this scaling, a cutoff of 10^{-6} electrons creates less than 1% error. In addition to the reduced virtual space, the computations shown in Fig. 3 were performed under the frozen core approximation. It should be noted that $E_{disp,approx}^{(20)}$ in Eq. (7) should include all relevant approximations (i.e., in this case
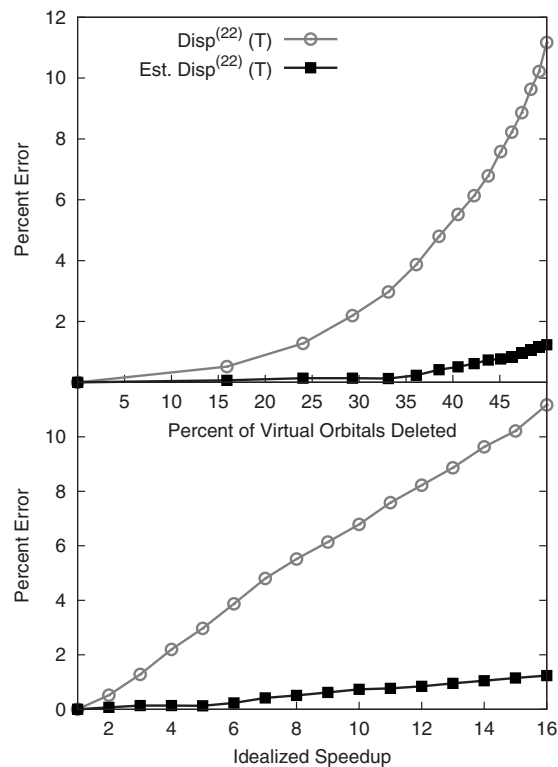

FIG. 2. Percent errors for the $E_{disp}^{(22)}(T)$ correction (scaled and unscaled) averaged over the 11 smallest complexes from the S22 test set (Ref. 49) as virtual MP2 NOs are removed from the computation to achieve a certain percentage of deleted virtual orbitals (top panel) or to obtain a specified idealized speedup (bottom panel).

it was computed with the core electrons frozen). This leads to our recommendation for evaluating the $E_{disp}^{(22)}(T)$ correction: virtual orbitals with less than 10^{-6} electrons should be removed, core electrons should be frozen, and the result should be scaled according to Eq. (7). This procedure introduces only negligible errors but greatly reduces computational costs. Of course, in very large molecules, one may wish to employ somewhat more aggressive truncations.

The accuracy of the approximations outlined above is assessed in Table I for the entire S22 test set. The $E_{disp}^{(22)}(T)$ values in this table were all computed using the DF approxi-
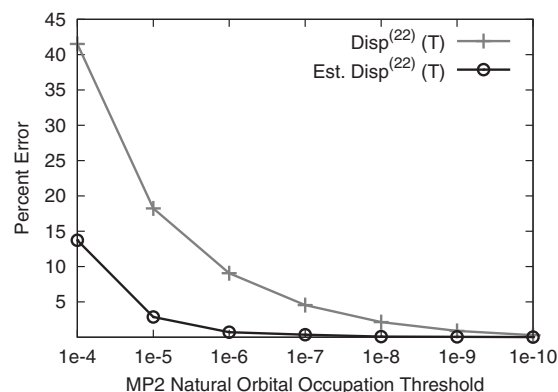

FIG. 3. Percent errors for the $E_{disp}^{(22)}(T)$ correction (scaled and unscaled) as virtual orbitals with less than a specified number of electrons are removed from the computation for the 11 smallest complexes from the S22 test set (Ref. 49).

TABLE I. The effect of the MP2 NO and frozen core approximations on the $E_{\text{disp}}^{(22)}(T)$ correction evaluated with the aug-cc-pVDZ basis set for the complexes of the S22 test set (Ref. 49). (MP2 NOs with occupancies less than 10^{-6} electrons are removed. Errors are given in kcal mol^{-1}.)

	$E_{\text{disp}}^{(22)}(T)_{\text{exact}}$	Error[a]	HF MOs[b]		MP2 NOs[c]	
			A	B	A	B
H-bonded complexes						
$(NH_3)_2$	−0.281	0.003	95	95	46	46
$(H_2O)_2$	−0.344	0.003	77	77	36	39
Formic acid dimer	−1.767	0.008	162	162	83	83
Formamide dimer	−1.389	0.010	180	180	90	90
Uracil dimer	−1.838	0.005	411	411	194	194
2-pyridoxine·2-aminopyridine	−2.098	0.006	396	396	183	191
Adenine·thymine WC	−2.153	0.004	501	503	241	229
Dispersion dominated complexes						
$(CH_4)_2$	−0.102	0.002	113	113	53	53
$(C_2H_4)_2$	−0.342	0.005	156	156	75	75
Benzene·CH$_4$	−0.514	0.006	230	246	165	55
PD benzene dimer	−2.396	0.008	363	363	174	174
Pyrazine dimer	−2.601	0.004	327	327	159	159
Uracil dimer	−2.765	0.020	411	411	231	231
Stacked indole·benzene	−3.708	0.004	441	431	179	242
Stacked adenine·thymine	−4.598	0.020	501	503	255	241
Mixed complexes						
Ethene·ethine	−0.232	0.001	138	139	74	54
Benzene·H$_2$O	−0.568	0.008	212	228	164	40
Benzene·NH$_3$	−0.552	0.007	221	237	165	48
Benzene·HCN	−0.736	0.005	226	240	167	52
T-shaped benzene dimer	−1.119	0.007	363	363	168	172
T-shaped indole benzene	−1.624	0.009	441	431	174	230
Phenol dimer	−1.445	0.005	405	405	190	190

[a]Signed error computed as Est.$E_{\text{disp}}^{(22)}(T)_{\text{approx}} - E_{\text{disp}}^{(22)}(T)_{\text{exact}}$.
[b]Number of unoccupied HF MOs for monomers A and B.
[c]Number of virtual MP2 NOs remaining after orbitals are removed.

mation of the two-electron integrals; errors due to the DF approximation have been considered previously[18,19] and are generally in the hundredths of 1 kcal mol^{-1} or less. The errors reported in Table I reflect the removal of virtual orbitals, the frozen core approximation, and the scaling shown in Eq. (7). These approximations introduce only modest errors into the computation; the largest error is only 0.02 kcal mol^{-1} [appearing twice in the table, for molecules with total $E_{\text{disp}}^{(22)}(T)$ contributions of −2.765 and −4.598 kcal mol^{-1}]; such a small error seems acceptable for nearly any application. For these cases and the aug-cc-pVDZ basis set, a threshold of 10^{-6} electrons for the removal of virtual orbitals typically removes about half of the virtual orbitals. However, for cases where one monomer is much larger than the other (e.g., benzene-methane), the majority of the smaller monomer's virtual orbitals can be removed. Note that in this implementation of SAPT, all computations are performed in the dimer basis. Because of this, some of the virtual orbitals will be composed mainly of basis functions centered on the other monomer. The procedure of removing MP2 NOs is a way of removing these extraneous virtual orbitals while retaining the important orbitals. It follows from this consideration that the hydrogen bonded and mixed complexes in the

S22 will benefit the most from the removal of virtual orbitals since they will have more spatially distant basis functions.

To evaluate this correction more efficiently, we use threaded Intel® MKL BLAS routines to form the triples amplitudes in Eq. (4); the energy evaluation is threaded using OpenMP. The timings shown in Fig. 4 were run on dual
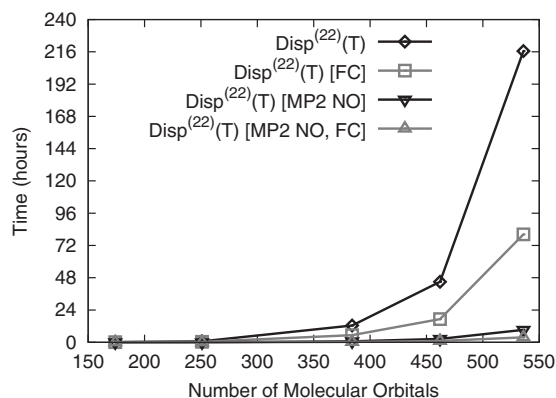

FIG. 4. Timings of the evaluation of $E_{\text{disp}}^{(22)}(T)$ correction with the aug-cc-pVDZ basis set as various approximations are applied for (from left to right) formic acid dimer, methane-benzene, T-shaped benzene dimer, T-shaped indole-benzene, and hydrogen bonded adenine-thymine.
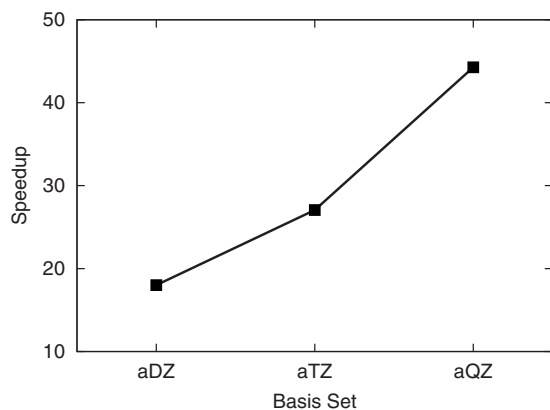
FIG. 5. Average speedup for the $E_{disp}^{(22)}(T)$ correction of water dimer, ammonia dimer, and methane dimer computed with various basis sets when virtual MP2 NOs are removed using a threshold of 10^{-6} electrons. An additional factor of approximately two times can be gained by the frozen core approximation.

IV. CONCLUSIONS

We have developed an approximation to the triples correction to dispersion in SAPT that uses MP2 NOs to reduce the number of virtual orbitals and a scaling relation to reduce the size of the error incurred. By truncating the virtual space and scaling the resulting energy according to Eq. (7), roughly half of the virtual orbitals can be removed with negligible errors. When this approximation is used in conjunction with the frozen core approximation, $E_{disp}^{(22)}(T)$ can be evaluated 50–60 times faster for the aug-cc-pVDZ basis set, with even greater speedups for larger basis sets. For the S22 test set, these approximations result in errors of, at most, only a few hundredths of 1 kcal mol^{-1}. With the approximations developed in this work, it should now be possible to include the important $E_{disp}^{(22)}(T)$ term for most systems where SAPT2 is applicable.

ACKNOWLEDGMENTS

quad-core Intel® Xeon E5430 processors clocked at 2.66 GHz with 8 threads. The largest computation shown in this figure corresponds to hydrogen bonded adenine-thymine. Run with 1 thread, without any approximations (other than the DF integrals), this computation would take roughly 2 months. Our threaded code reduces this to 9 days, and with the frozen core approximation, the computation takes a more manageable 3 days. When the virtual space is reduced, $E_{disp}^{(22)}(T)$ can be computed in less than 4 h. For the systems considered, the combination of the frozen core and MP2 NO approximations result in a remarkable 50–60 times speedup.

It is possible to realize even greater speedups for larger basis sets (to this point, we have only shown results for the modest aug-cc-pVDZ basis). For larger basis sets, truncation with a certain occupation threshold will remove a larger fraction of the virtual orbitals. The speedups for the $E_{disp}^{(22)}(T)$ evaluation with the aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ bases are shown in Fig. 5. Due to the expense of triples corrections in an aug-cc-pVQZ basis, we only show results for three small dimers from the S22 test set. The average speedups that result from the truncation of MP2 NOs increase with the size of the basis set to an impressive 45 times with the aug-cc-pVQZ basis. When the truncated virtual space is combined with the frozen core approximation, the overall speedup increases to 85 times with the aug-cc-pVQZ basis. In our limited test cases, for a particular system, the amount of error introduced by the deletion of NOs with occupation numbers smaller than 10^{-6} remains similar as one goes to larger basis sets.

As defined above, the truncation of the virtual space using MP2 NO occupations will not result in a continuous potential energy surface. At two adjacent points, it is possible that different numbers of virtual orbitals will be removed. Due to the excellent performance of this approximation, however, discontinuities are likely to be unnoticeably small. Additionally, derivatives of SAPT energies are not typically computed. If continuity becomes an issue, we recommend determining a reasonable number of virtual orbitals to remove and fixing that number for the entire surface.

[1] P. Hobza and J. Šponer, J. Am. Chem. Soc. **124**, 11802 (2002).
[2] V. R. Cooper, T. Thonhauser, A. Puzder, E. Schroder, B. I. Lundqvist, and D. C. Langreth, J. Am. Chem. Soc. **130**, 1304 (2008).
[3] A. Fiethen, G. Jansen, A. Hesselmann, and M. Schütz, J. Am. Chem. Soc. **130**, 1802 (2008).
[4] S. Li, V. R. Cooper, T. Thonhauser, B. I. Lundqvist, and D. C. Langreth, J. Phys. Chem. B **113**, 11166 (2009).
[5] D. Svozil, P. Hobza, and J. Sponer, J. Phys. Chem. B **114**, 1191 (2010).
[6] R. A. Aziz and M. J. Slaman, Mol. Phys. **58**, 679 (1986).
[7] T. van Mourik and J. H. van Lenthe, J. Chem. Phys. **102**, 7479 (1995).
[8] R. Podeszwa and K. Szalewicz, Chem. Phys. Lett. **412**, 488 (2005).
[9] R. Podeszwa, B. M. Rice, and K. Szalewicz, Phys. Rev. Lett. **101**, 115503 (2008).
[10] A. L. Ringer and C. D. Sherrill, Chem.-Eur. J. **14**, 2542 (2008).
[11] R. Podeszwa, B. M. Rice, and K. Szalewicz, Phys. Chem. Chem. Phys. **11**, 5512 (2009).
[12] R. L. Jaffe and G. D. Smith, J. Chem. Phys. **105**, 2780 (1996).
[13] S. Tsuzuki, T. Uchimaru, K. Matsumura, M. Mikami, and K. Tanabe, Chem. Phys. Lett. **319**, 547 (2000).
[14] S. Tsuzuki and H. P. Lüthi, J. Chem. Phys. **114**, 3949 (2001).
[15] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[16] B. Jeziorski, R. Moszynski, and K. Szalewicz, Chem. Rev. (Washington, D.C.) **94**, 1887 (1994).
[17] B. Jeziorski, R. Moszynski, A. Ratkiewicz, S. Rybak, K. Szalewicz, and H. L. Williams, in *Methods and Techniques in Computational Chemistry: METECC94*, edited by E. Clementi (STEF, Cagliari, 1993), Vol. B, p. 79.
[18] E. G. Hohenstein and C. D. Sherrill, J. Chem. Phys. **132**, 184111 (2010).
[19] E. G. Hohenstein and C. D. Sherrill, J. Chem. Phys. **133**, 014101 (2010).
[20] P. O. Löwdin, Phys. Rev. **97**, 1474 (1955).
[21] E. R. Davidson, Rev. Mod. Phys. **44**, 451 (1972).
[22] P.-O. Löwdin and H. Shull, Phys. Rev. **101**, 1730 (1956).
[23] K. Ruedenberg, L. M. Cheung, and S. T. Elbert, Int. J. Quantum Chem. **16**, 1069 (1979).
[24] H. J. Aa Jensen, P. Jørgensen, H. Ågren, and J. Olsen, J. Chem. Phys. **88**, 3834 (1988).
[25] P. Pulay and T. P. Hamilton, J. Chem. Phys. **88**, 4926 (1988).
[26] J. M. Bofill and P. Pulay, J. Chem. Phys. **90**, 3637 (1989).

[27] M. L. Abrams and C. D. Sherrill, Chem. Phys. Lett. **395**, 227 (2004).

[28] R. S. Grev and H. F. Schaefer, J. Chem. Phys. **96**, 6850 (1992).

[29] C. D. Sherrill and H. F. Schaefer, J. Phys. Chem. **100**, 6069 (1996).

[30] M. L. Abrams and C. D. Sherrill, J. Chem. Phys. **118**, 1604 (2003).

[31] L. Bytautas, J. Ivanic, and K. Ruedenberg, J. Chem. Phys. **119**, 8217 (2003).

[32] L. Bytautas and K. Ruedenberg, J. Chem. Phys. **121**, 10905 (2004).

[33] L. Bytautas and K. Ruedenberg, J. Chem. Phys. **121**, 10919 (2004).

[34] L. Bytautas and K. Ruedenberg, J. Chem. Phys. **121**, 10852 (2004).

[35] W. Klopper, J. Noga, H. Koch, and T. Helgaker, Theor. Chem. Acc. **97**, 164 (1997).

[36] A. G. Taube and R. J. Bartlett, Collect. Czech. Chem. Commun. **70**, 837 (2005).

[37] A. Landau, K. Khistyaev, S. Dolgikh, and A. I. Krylov, J. Chem. Phys. **132**, 014109 (2010).

[38] L. Adamowicz and R. J. Bartlett, J. Chem. Phys. **86**, 6314 (1987).

[39] P. Neogrády, M. Pitoňák, and M. Urban, Mol. Phys. **103**, 2141 (2005).

[40] M. Pitoňák, F. Holka, P. Neogrády, and M. Urban, J. Mol. Struct.: THEOCHEM **768**, 79 (2006).

[41] M. Pitoňák, P. Neogrády, J. Řezáč, P. Jurečka, M. Urban, and P. Hobza, J. Chem. Theory Comput. **4**, 1829 (2008).

[42] C. Edmiston and M. Krauss, J. Chem. Phys. **42**, 1119 (1965).

[43] W. Meyer, Int. J. Quantum Chem., Symp. **5**, 341 (1971).

[44] W. Meyer, J. Chem. Phys. **58**, 1017 (1973).

[45] F. Neese, F. Wennmohs, and A. Hansen, J. Chem. Phys. **130**, 114108 (2009).

[46] F. Neese, A. Hansen, and D. G. Liakos, J. Chem. Phys. **131**, 064103 (2009).

[47] K. Patkowski and K. Szalewicz, J. Chem. Phys. **127**, 164103 (2007).

[48] T. D. Crawford, C. D. Sherrill, E. F. Valeev, J. T. Fermann, R. A. King, M. L. Leininger, S. T. Brown, C. L. Janssen, E. T. Seidl, J. P. Kenny, and W. D. Allen, J. Comput. Chem. **28**, 1610 (2007).

[49] P. Jurečka, J. Šponer, J. Černý, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).

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

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

[52] F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).

[53] T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall, and C. D. Sherrill, J. Chem. Phys. **132**, 144104 (2010).