The dispersion interaction between quantum mechanics and effective fragment potential molecules

Quentin A. Smith, Klaus Ruedenberg, Mark S. Gordon, and Lyudmila V. Slipchenko

Citation: J. Chem. Phys. **136**, 244107 (2012); doi: 10.1063/1.4729535
View online: http://dx.doi.org/10.1063/1.4729535
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v136/i24
Published by the American Institute of Physics.

The dispersion interaction between quantum mechanics and effective fragment potential molecules

Quentin A. Smith,[1] Klaus Ruedenberg,[1] Mark S. Gordon,[1,a)]
and Lyudmila V. Slipchenko[2,a)]
[1]*Department of Chemistry and Ames Laboratory, Iowa State University, Ames, Iowa 50011, USA*
[2]*Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, USA*

(Received 6 April 2012; accepted 31 May 2012; published online 26 June 2012)

A method for calculating the dispersion energy between molecules modeled with the general effective fragment potential (EFP2) method and those modeled using a full quantum mechanics (QM) method, e.g., Hartree-Fock (HF) or second-order perturbation theory, is presented. C_6 dispersion coefficients are calculated for pairs of orbitals using dynamic polarizabilities from the EFP2 portion, and dipole integrals and orbital energies from the QM portion of the system. Dividing by the sixth power of the distance between localized molecular orbital centroids yields the first term in the commonly employed London series expansion. A C_8 term is estimated from the C_6 term to achieve closer agreement with symmetry adapted perturbation theory values. Two damping functions for the dispersion energy are evaluated. By using terms that are already computed during an ordinary HF or EFP2 calculation, the new method enables accurate and extremely rapid evaluation of the dispersion interaction between EFP2 and QM molecules. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4729535]

I. INTRODUCTION

The dispersion interaction is an attractive force between atoms and molecules that is caused by the interactions of induced multipoles. Dispersion arises from the correlated movement of electrons; an instantaneous multipole on one molecule may induce a multipole on another molecule.[1–4] While dispersion is generally a weak intermolecular force, it is the major attractive force between neutral atoms and molecules that lack permanent multipole moments. For example, dispersion is singularly responsible for the attraction between atoms of the noble gases,[5] allowing for their condensation at low temperatures. Because dispersion depends on the atomic or molecular dynamic (frequency dependent) polarizability, the dispersion energy is significant between atoms or molecules with large, diffuse electron clouds, such as π clouds.

The general formula for the dispersion interaction between two quantum mechanical (QM) molecules comes from the second-order term in intermolecular perturbation theory.[1] Based on this functional form, a formula[6] for the dispersion interaction between two molecules modeled with the general effective fragment potential (EFP2) (Refs. 7–9) method was previously derived and implemented in the general atomic and molecular electronic structure system (GAMESS) package.[10] However, a formula for the dispersion energy in mixed systems—those in which one molecule is modeled with EFP2 and another with a full *ab initio* (AI) QM method, e.g., Hartree-Fock (HF) or second-order perturbation theory (MP2) (Ref. 11)—has been lacking. This paper presents a derivation and implementation of the EFP2-AI dispersion interaction, as well as a comparison of resulting EFP2-AI dispersion energy calculations with those of EFP2-EFP2 and symmetry adapted perturbation theory (SAPT).[12] The EFP2-AI dispersion term is important in order to describe solute-solvent interactions, where the solute may be modeled with an *ab initio* method and the solvent with EFP2. Dispersion is the primary means of solute-solvent interactions with nonpolar solvents, e.g., cyclohexane.

The effective fragment potential (EFP) method is a discrete solvent method in GAMESS. The original EFP method, EFP1, involves fitted parameters explicitly designed for water, while the general method, EFP2, contains no fitted parameters and may be applied to any molecule. Parameters necessary to describe the fragment potential are generated in a MAKEFP calculation for a given isolated molecule. The resulting potential may be used in future calculations on any number of molecules of the specified type, with the same internal geometry. Potentials can be used to compute Coulomb, exchange-repulsion, polarization, charge transfer, and dispersion energy components in the interaction between two EFP2 fragments.[9] With the current development of the EFP2-AI dispersion term, all interaction energy terms except charge transfer may be calculated between an EFP2 fragment and a molecule modeled with a full *ab initio* method.

The EFP2 Coulomb term is calculated via the distributed multipolar analysis (DMA) method of Stone,[1] carried out through octopole moments. The DMA, a classical pointwise model, cannot account for the overlap of charge densities between molecules that occurs at short distances. As such, charge penetration effects are modeled by a distance-dependent cutoff function.[13] Polarization, or induction, is the interaction between an induced dipole on one fragment and the electric field of another. In the EFP2 method,

a)Authors to whom correspondence should be addressed. Electronic addresses: mgordon@iastate.edu and lslipchenko@purdue.edu.

induction is expressed in terms of localized molecular orbital (LMO) polarizabilities, with polarizability points at each bond and lone pair. The induced dipoles are iterated to self-consistency. The exchange-repulsion term, derived from first principles,[1,14] is expressed as a function of intermolecular overlap integrals.[15,16] Charge transfer is an interaction between the occupied valence orbitals of one molecule and the virtual orbitals of another. The implementation of EFP2-EFP2 charge transfer is described in Ref. 17; EFP2-AI charge transfer is currently under development. The EFP2-EFP2 dispersion term, discussed further in Secs. II and III, is a function of frequency-dependent dynamic polarizabilities summed over the entire (imaginary) frequency range.[6] Extensive descriptions of how the potentials are generated and used to compute intermolecular interaction energies may be found elsewhere.[6–9,13,15–18]

Previous work in deriving interaction terms for discrete QM/MM methods (that is, the interaction between quantum mechanical and classical molecular mechanical systems) includes the direct reaction field approach of van Duijnen and co-workers.[19] Distributed approaches for both multipoles[14] and polarizabilities[20] have been proposed by Stone.[1] Claverie and Rein[21] have discussed a distributed model for intermolecular interactions involving localized orbitals, and Karlström and co-workers[22] have used a localized orbital approach for constructing intermolecular potentials with the non-empirical molecular orbital method (NEMO). The EFP2-AI dispersion term proposed in the present work involves distributed polarizabilities on the EFP part (which are calculated for localized orbitals) and localized orbital energies and dipole integrals on the AI part of the system.

The organization of this paper is as follows. Section II presents a derivation of the EFP2-AI dispersion energy term beginning from the dispersion energy equation in its most general form, from intermolecular perturbation theory. Section III describes the implementation of the EFP2-AI dispersion code in GAMESS. Computational details concerning the systems used to test the code appear in Sec. IV. Section V compares dispersion energy values obtained from the EFP2-AI dispersion code with those of EFP2-EFP2 and SAPT. Conclusions are presented in Sec. VI.

II. THEORY

A. Dispersion interaction

The general formula for the dispersion interaction between two closed-shell nondegenerate ground state molecules can be derived with Rayleigh-Schrödinger perturbation theory (RSPT), following the method of Stone.[1] While RSPT does not account for intermolecular antisymmetry effects, such as exchange-dispersion, that predominate at short intermonomer distances, it adequately accounts for mid- to long-range effects like the dispersion interaction that is of interest here.[22,23] The unperturbed Hamiltonian is given by the sum of the Hamiltonians for the individual molecules A and B

$$\widehat{H}_0 = \widehat{H}_A + \widehat{H}_B. \tag{1}$$

The perturbation operator embodies all electrostatic interactions between the molecules. It can be represented in numerous ways, the simplest of which is

$$\widehat{V} = \sum_{i\in A}\sum_{j\in B}\frac{q_i q_j}{R_{ij}}, \tag{2}$$

where q_i is the charge on particle i (electron or nucleus) on molecule A and R_{ij} is the distance between i and j. A more convenient representation is the multipolar expansion:

$$\widehat{V} = T^{AB}q^A q^B + \sum_{a}^{x,y,z} T_a^{AB}\left(q^A\mu_a^B - \mu_a^A q^B\right) - \sum_{a,b}^{x,y,z} T_{ab}^{AB}\mu_a^A\mu_b^B + \ldots, \tag{3}$$

where q^A is the net charge on molecule A, μ_a^B is the ath directional (x, y, z) component of the dipole moment on B, etc. The quantities T, T_a, T_{ab}, etc., are electrostatic tensors of rank indicated by the number of subscripts. T corresponds to charge-charge interactions, T_a corresponds to charge-dipole interactions, T_{ab} contains dipole-dipole and charge-quadrupole interactions, T_{abc} contains dipole-quadrupole and charge-octopole interactions, and so on. The formulae of the first three tensors appear below, where R_{ij}^a denotes the ath directional component of the distance between i and j,

$$T^{ij} = \frac{1}{R_{ij}}, \tag{4}$$

$$T_a^{ij} = \nabla_a\frac{1}{R} = -\frac{\vec{R}_{ij}^a}{R_{ij}^3}, \tag{5}$$

$$T_{ab}^{ij} = \nabla_a\nabla_b\frac{1}{R} = \frac{3\vec{R}_{ij}^a\vec{R}_{ij}^b - R_{ij}^2\delta_{ab}}{R_{ij}^5}. \tag{6}$$

When the perturbation expansion is carried out, the zeroth-order term gives the sum of the ground state energies of A and B, and the first-order term gives the Coulomb (electrostatic) interaction energy. The second-order energy term in the perturbation expansion is given by

$$W_0'' = -\sum_{m,n}\frac{\langle 0_A 0_B|\,\widehat{V}\,|mn\rangle\,\langle mn|\,\widehat{V}\,|0_A 0_B\rangle}{W_{mn} - W_{0_A 0_B}}, \tag{7}$$

where m and n are states of molecules A and B, respectively, 0_A represents the ground state of molecule A, and $W_{mn} = E_m + E_n$ is the energy of the system in state $|mn\rangle$. In Eq. (7), either m or n may be zero, but not both. Formally, since antisymmetry is not imposed on the intermolecular wavefunction, this expression is valid for exact eigenstates of the isolated Hamiltonians.[23]

This second-order perturbative expression, Eq. (7), encompasses both the induction (polarization) and dispersion energies. Induction arises from the interactions between permanent multipoles on one molecule and induced multipoles on the other. As such, its representation in Eq. (7) occurs in those terms in the summation in which either m or n is equal to 0, i.e., one of the interacting molecules is in its ground state, while the other is in an (induced) excited state. The remainder of the terms in the summation, those in which both m and n

refer to excited states, correspond to the dispersion energy

$$E^{disp} = -\sum_{\substack{m\neq 0\\ n\neq 0}} \frac{\langle 0_A 0_B|\,\widehat{V}\,|mn\rangle\,\langle mn|\,\widehat{V}\,|0_A 0_B\rangle}{E_m^A + E_n^B - E_0^A - E_0^B}. \tag{8}$$

The multipolar expansion of Eq. (3) may then be substituted for the perturbation operator in Eq. (8). Since the charges q in Eq. (3) are scalar values, integrals involving q will be of the form, for example, $\langle 0_A|q^A|m\rangle = q^A\langle 0_A|m\rangle$; these integrals are zero because the ground and excited states are orthogonal to each other. Therefore, the multipole expansion, when used in this context, properly begins with the dipole-dipole term. The dispersion energy corresponding to the dipole-dipole interaction, labeled E_6^{disp}, is given by

$$E_6^{disp} = -\sum_{\substack{m\neq 0\\ n\neq 0}} \frac{\langle 0_A 0_B|\sum_{a,b}^{x,y,z} T_{ab}^{AB}\widehat{\mu}_a^A\widehat{\mu}_b^B\,|mn\rangle\,\langle mn|\sum_{c,d}^{x,y,z} T_{cd}^{AB}\widehat{\mu}_c^A\widehat{\mu}_d^B\,|0_A 0_B\rangle}{E_m^A + E_n^B - E_0^A - E_0^B}. \tag{9}$$

The dipole operators apply only to the states of their respective molecules (m on A or n on B), and tensors T_{ab} do not depend on the states. Assuming that the states are separable, i.e., $|mn\rangle = |m\rangle|n\rangle$, Eq. (9) may then be simplified to

$$E_6^{disp} = -\sum_{abcd}^{x,y,z} T_{ab}^{AB}T_{cd}^{AB}\sum_{\substack{m\neq 0\\ n\neq 0}} \frac{1}{E_{m0}^A + E_{n0}^B}\langle 0_A|\widehat{\mu}_a^A|m\rangle\langle m|\widehat{\mu}_c^A|0_A\rangle\langle 0_B|\widehat{\mu}_b^B|n\rangle\langle n|\widehat{\mu}_d^B|0_B\rangle, \tag{10}$$

where $E_{m0} = E_m - E_0$. The energy factors do not permit the easy separation of Eq. (10) into portions relating to molecule A and portions relating to molecule B. The Casimir-Polder identity[1,24]

$$\frac{1}{A+B} = \frac{2}{\pi}\int_0^\infty \frac{AB}{(A^2+\omega^2)(B^2+\omega^2)}d\omega \tag{11}$$

may be applied to the denominator of Eq. (10) to obtain

$$E_6^{disp} = -\frac{2\hbar}{\pi}\sum_{abcd}^{x,y,z} T_{ab}^{AB}T_{cd}^{AB}\int_0^\infty d\omega \sum_{\substack{m\neq 0\\ n\neq 0}} \frac{\omega_{m0}^A\,\langle 0_A|\,\widehat{\mu}_a^A|m\rangle\langle m|\widehat{\mu}_c^A|0_A\rangle}{\hbar\left(\left(\omega_{m0}^A\right)^2+\omega^2\right)}\frac{\omega_{n0}^B\langle 0_B|\widehat{\mu}_b^B|n\rangle\langle n|\widehat{\mu}_d^B|0_B\rangle}{\hbar\left(\left(\omega_{n0}^B\right)^2+\omega^2\right)} \tag{12}$$

after the energies in Eq. (10) are expressed in terms of frequencies, e.g., $E_{m0}^A = \hbar\omega_{m0}^A$.

From time-dependent (TD) perturbation theory, it can be shown that the response of a molecule to an oscillating electric field is an oscillating dipole moment. If a field F_b with frequency ω is given by $F_b e^{-i\omega t}$, then the dipole moment is $\mu_a = \alpha_{ab}(\omega)F_b e^{-i\omega t}$. The components of the frequency-dependent dynamic polarizability tensor, α, are given by[1]

$$\alpha_{ab}(\omega) = \sum_m \frac{\omega_{m0}\{\langle 0|\,\widehat{\mu}_a\,|m\rangle\,\langle m|\,\widehat{\mu}_b\,|0\rangle + \langle 0|\,\widehat{\mu}_b\,|m\rangle\,\langle m|\,\widehat{\mu}_a\,|0\rangle\}}{\hbar\left(\omega_{m0}^2-\omega^2\right)} = 2\sum_m \frac{\omega_{m0}\,\langle 0|\,\widehat{\mu}_a\,|m\rangle\,\langle m|\,\widehat{\mu}_b\,|0\rangle}{\hbar\left(\omega_{m0}^2-\omega^2\right)}, \tag{13}$$

α describes the propagation of a density fluctuation within a molecule.[2] Expressing the factor ω^2 in Eq. (13) as $\omega^2 = -(-\omega^2) = -(i\omega)^2$, Eq. (12) can be recast in terms of the dynamic polarizability tensor at an imaginary frequency. The concept of imaginary frequencies is purely a mathematical construct, with no actual physical meaning.[1] Regardless, it can be used to construct functions that are well-behaved and that decrease monotonically to 0 as $\omega \to \infty$.

B. EFP-EFP dispersion

If the conversion into terms of dynamic polarizability tensors is performed for both molecule A and molecule B, as in the case of EFP2-EFP2 interactions, the resulting expression is

$$E_6^{disp} = -\frac{\hbar}{2\pi}\sum_{abcd}^{x,y,z} T_{ab}^{AB}T_{cd}^{AB}\int_0^\infty \alpha_{ac}^A(i\omega)\,\alpha_{bd}^B(i\omega)\,d\omega. \tag{14}$$

The dynamic polarizability tensors in Eq. (14) are implicitly calculated at a single point on each molecule. However, there are known issues with convergence when using this approach; to be effective, it may require higher terms in the multipolar expansion[7] (i.e., quadrupole polarizabilities might need to be considered). A distributed polarizability model, in which polarizability tensors are calculated at multiple expansion points, has the advantage of more successful convergence and possibly giving a more realistic portrayal of the response of a molecule to the nonuniform fields arising from

the interacting molecules.[7] To improve numerical accuracy without introducing additional complication, the form of Eq. (14) is used to construct an approximate distributed polarizability model:

$$E_6^{disp} = -\frac{\hbar}{2\pi}\sum_{k\in A}\sum_{j\in B}\sum_{abcd}^{x,y,z} T_{ab}^{kj} T_{cd}^{kj} \int_0^\infty \alpha_{ac}^{k}(i\omega)\,\alpha_{bd}^{j}(i\omega)\,d\omega. \tag{15}$$

In Eq. (15), k and j are LMOs on molecules A and B, respectively, and the approximation

$$\alpha_{bd}^{B}(i\omega) \rightarrow \sum_{j\in B}\alpha_{bd}^{j}(i\omega) \tag{16}$$

is used for the distributed polarizabilities.

The EFP2-EFP2 dispersion energy expression in Eq. (15) is anisotropic, as the directional component summation is over all possible x, y, z pairs. In most computational methods, an isotropic expression—one that does not include xy, xz, and other off-diagonal terms—is preferred. The off-diagonal terms of the frequency-dependent polarizability tensor typically do not contribute greatly to the dispersion energy. This is because, in the case of the EFP method, the polarizability tensor is constructed using the principal orientation for a given molecule: the axes are chosen such that the diagonal components (xx, yy, and zz) are always dominant. Thus, omitting the off-diagonal terms from the calculation significantly reduces the total number of terms that must be calculated in Eq. (15) without serious loss of accuracy. Additionally, because the majority of the C_6 values reported in the literature are isotropic,[6] introducing an isotropic expression for the EFP2-EFP2 dispersion coefficient allows straightforward comparison with other computational and theoretical methods. Therefore, the following approximation is made:

$$\begin{aligned} E_6^{disp} &\approx -\frac{\hbar}{2\pi}\sum_{k\in A}\sum_{j\in B}\sum_{abcd}^{x,y,z}\delta_{ac}\delta_{bd}T_{ab}^{kj}T_{cd}^{kj} \\ &\quad\times\int_0^\infty \alpha_{ac}^{k}(i\omega)\,\alpha_{bd}^{j}(i\omega)\,d\omega \\ &= -\frac{\hbar}{2\pi}\sum_{k\in A}\sum_{j\in B}\sum_{ab}^{x,y,z} T_{ab}^{kj}T_{ab}^{kj}\int_0^\infty \alpha_{aa}^{k}(i\omega)\,\alpha_{bb}^{j}(i\omega)\,d\omega. \end{aligned} \tag{17}$$

A further simplification may be achieved by introducing the isotropic dynamic polarizability $\bar{\alpha}^j(i\omega)$ for LMO j and frequency ω as 1/3 of the trace of the polarizability tensor at a given imaginary frequency and employing the spherical atom approximation

$$\alpha_{xx}^{j}(i\omega) \approx \alpha_{yy}^{j}(i\omega) \approx \alpha_{zz}^{j}(i\omega) = \bar{\alpha}^{j}(i\omega). \tag{18}$$

Taking into account that

$$\sum_{ab}^{x,y,z} T_{ab}^{kj}T_{ab}^{kj} = \frac{6}{R_{kj}^6}, \tag{19}$$

the isotropic dispersion energy expression in atomic units becomes

$$E_6^{disp} = -\frac{3}{\pi}\sum_{k\in A}\sum_{j\in B}\frac{1}{R_{kj}^6}\int_0^\infty \bar{\alpha}^{k}(i\omega)\,\bar{\alpha}^{j}(i\omega)\,d\omega. \tag{20}$$

Equation (20) is the final expression for the dipole-dipole term of the dispersion interaction energy between two EFP2 molecules, where k and j are expansion points for distributed polarizabilities on EFP2 molecules A and B, respectively, and R_{kj} is the distance between these expansion points. The integral over the imaginary frequency range is evaluated via a 12-point Gauss-Legendre numerical quadrature.[6] The calculation of the dynamic polarizability tensors α at the necessary (predetermined) frequencies is part of the MAKEFP procedure, which generates the fragment potential parameters before the calculation of interaction energy components begins.

C. EFP-AI dispersion

In constructing an expression for the EFP2-AI dispersion interaction, it would be appealing to obtain an equation analogous to Eq. (20). However, the procedure used to obtain the dynamic polarizability tensors α in a MAKEFP calculation is too time consuming to be appropriate for the AI molecule, especially since many energy and gradient evaluations may be needed in the calculation. Following the approach of Amos and co-workers,[3,4] values for α associated with each polarizable point are computed via the time-dependent analog of the coupled perturbed Hartree-Fock (CPHF)[25] equations. (Details may be found in Refs. 3, 4, 6, and 25.) The time requirement for this procedure is acceptable in the case of EFP2-EFP2 because the MAKEFP calculation occurs prior to, and as a separate step from, the EFP2-EFP2 dispersion calculation. Constructing the α tensors for the *ab initio* molecule would necessarily occur during the calculation of the EFP2-AI interaction, thus increasing the run time significantly. For example, if a Hartree-Fock calculation on phenol with the 6-311++G(3df,2p) basis set (equivalent to 375 basis functions) takes approximately 6 min on a given computer, a CPHF calculation on the same chemical system requires 10 min, and a time-dependent Hartree-Fock (TDHF) calculation requires 12 min. To avoid these time-consuming calculations, the EFP2-AI dispersion derivation must diverge from the EFP2-EFP2 derivation at Eq. (12). Only the portion of Eq. (12) relating to the EFP2 part (molecule B) is recast in terms of α, giving

$$\begin{aligned} E_6^{EFP-AI} &= -\frac{\hbar}{\pi}\sum_{abcd}^{x,y,z} T_{ab}^{AB}T_{cd}^{AB}\int_0^\infty d\omega \sum_{m\neq 0}\frac{\omega_{m0}^{A}\langle 0_A|\hat{\mu}_a^A|m\rangle\langle m|\hat{\mu}_c^A|0_A\rangle}{\hbar\left(\left(\omega_{m0}^{A}\right)^2+\omega^2\right)}\alpha_{bd}^{B}(i\omega) \\ &= -\frac{1}{\pi}\sum_{abcd}^{x,y,z} T_{ab}^{AB}T_{cd}^{AB}\sum_{m\neq 0}\langle 0_A|\hat{\mu}_a^A|m\rangle\langle m|\hat{\mu}_c^A|0_A\rangle\int_0^\infty d\omega\frac{\omega_{m0}^{A}}{\left(\omega_{m0}^{A}\right)^2+\omega^2}\alpha_{bd}^{B}(i\omega). \end{aligned} \tag{21}$$

To convert from a sum-over-states approach to an orbital-based approach, let i correspond to occupied canonical molecular orbitals (MOs) and r to virtual canonical MOs on the AI molecule. This changes the dipole integrals in Eq. (21) to

$$\sum_{m\neq 0} \langle 0_A|\,\hat{\mu}_a^A\,|m\rangle\,\langle m|\,\hat{\mu}_c^A\,|0_A\rangle \rightarrow \sum_{k}^{occ}\sum_{r}^{vir} \langle k|\,\hat{\mu}_a^A\,|r\rangle\,\langle r|\,\hat{\mu}_c^A\,|k\rangle \tag{22}$$

(see Appendix A). The integral over the imaginary frequency range in Eq. (21) becomes

$$\int_0^\infty d\omega \frac{\omega_{m0}^A}{\left(\omega_{m0}^A\right)^2+\omega^2}\alpha_{bd}^B(i\omega) \rightarrow \int_0^\infty d\omega \frac{\omega_{rk}^A}{\left(\omega_{rk}^A\right)^2+\omega^2}\alpha_{bd}^B(i\omega), \quad \text{where } \omega_{rk}=\omega_r-\omega_k \tag{23}$$

(see Appendix B). Invoking Eqs. (22) and (23) to change from sum-over-states to the orbital approximation and Eq. (16) to change to distributed polarizability tensors on fragment B, Eq. (21) becomes

$$E_6^{EFP-AI} = -\frac{1}{\pi}\sum_{j\in B}\sum_{abcd}^{x,y,z}\sum_{k}^{occ}\sum_{r}^{vir} T_{ab}^{kj}T_{cd}^{kj}\,\langle k|\,\hat{\mu}_a^A\,|r\rangle\,\langle r|\,\hat{\mu}_c^A\,|k\rangle \int_0^\infty d\omega \frac{\omega_{rk}^A}{\left(\omega_{rk}^A\right)^2+\omega^2}\alpha_{bd}^j(i\omega). \tag{24}$$

The isotropic approximation is now employed. First, off-diagonal directional component terms (which are assumed to be negligible) are eliminated

$$\begin{aligned} E_6^{EFP-AI} &\approx -\frac{1}{\pi}\sum_{j\in B}\sum_{abcd}^{x,y,z}\sum_{k}^{occ}\sum_{r}^{vir} \delta_{ac}\delta_{bd}T_{ab}^{kj}T_{cd}^{kj}\,\langle k|\,\hat{\mu}_a^A\,|r\rangle\,\langle r|\,\hat{\mu}_c^A\,|k\rangle \int_0^\infty d\omega \frac{\omega_{rk}^A}{\left(\omega_{rk}^A\right)^2+\omega^2}\alpha_{bd}^j(i\omega) \\ &= -\frac{1}{\pi}\sum_{j\in B}\sum_{ab}^{x,y,z}\sum_{k}^{occ}\sum_{r}^{vir} T_{ab}^{kj}T_{ab}^{kj}\,\langle k|\,\hat{\mu}_a^A\,|r\rangle\,\langle r|\,\hat{\mu}_a^A\,|k\rangle \int_0^\infty d\omega \frac{\omega_{rk}^A}{\left(\omega_{rk}^A\right)^2+\omega^2}\alpha_{bb}^j(i\omega). \end{aligned} \tag{25}$$

The product $T_{ab}^{kj}T_{ab}^{kj}$ in Eq. (25) can be expressed in terms of R_{kj}^{-6} as in Eq. (19), where R_{kj} is the distance between the centroids of orbitals k (on the AI molecule) and j (on the EFP2 fragment). Similar to the direction-independent quantity $\bar{\alpha}^j(i\omega)$ from Eq. (18), the Cartesian component-averaged dipole integrals, given by $\langle k|\,\hat{\mu}\,|r\rangle\,\langle r|\,\hat{\mu}\,|k\rangle = \frac{1}{3}\sum_a^{x,y,z}\langle k|\,\hat{\mu}_a^A\,|r\rangle\,\langle r|\,\hat{\mu}_a^A\,|k\rangle$, are introduced. The resulting fully isotropic EFP2-AI dispersion equation is

$$E_6^{EFP-AI} = -\frac{6}{\pi}\sum_{j\in B}\sum_{k}^{occ}\sum_{r}^{vir}\frac{1}{R_{kj}^6}\,\langle k|\,\hat{\mu}\,|r\rangle\,\langle r|\,\hat{\mu}\,|k\rangle \int_0^\infty d\omega \frac{\omega_{rk}^A}{\left(\omega_{rk}^A\right)^2+\omega^2}\bar{\alpha}^j(i\omega). \tag{26}$$

The dispersion expansion is commonly expressed in the form of the series

$$E^{disp} = \frac{C_6}{R^6}+\frac{C_8}{R^8}+\frac{C_{10}}{R^{10}}+\ldots, \tag{27}$$

where C_6/R^6 corresponds to the instantaneous dipole-induced dipole interaction, C_8/R^8 primarily to the instantaneous dipole-induced quadrupole and instantaneous quadrupole-induced dipole interaction, etc. (Odd-order terms are almost always neglected, although they are strictly equal to zero only for atoms and for molecules with inversion centers.[1]) The isotropic EFP2-AI distributed dipole-dipole dispersion term can be expressed in the form

$$E_6^{EFP-AI} = -\sum_{k\in A}\sum_{j\in B}\frac{C_6^{kj}}{R_{kj}^6}, \tag{28}$$

where, by comparison with Eq. (26),

$$C_6^{kj} = \frac{6}{\pi}\sum_{r}^{\text{vir}}\langle k|\,\hat{\mu}\,|r\rangle\,\langle r|\,\hat{\mu}\,|k\rangle \int_0^\infty d\omega \frac{\omega_{rk}^A}{\left(\omega_{rk}^A\right)^2+\omega^2}\bar{\alpha}^j(i\omega). \tag{29}$$

III. LMO FORMULATION AND IMPLEMENTATION

While Eqs. (26) and (29) are formulated in the canonical MO basis in the AI region, previous studies with the EFP method suggest that formulation in terms of LMOs provides faster convergence and more accurate results.[6,7,15] Therefore, in the following, distributed LMO-based C_6 coefficients and EFP-AI dispersion energy are derived. In analogy with Eq. (28), the LMO reformulation of the EFP2-AI dispersion energy term is proposed to have the form

$$E_6^{EFP-AI} = -\sum_{\ell\in A}\sum_{v\in B}\frac{C_6^{\ell v}}{R_{\ell v}^6}, \tag{30}$$

where $|\ell\rangle$ and $|v\rangle$ refer to LMOs on the AI and EFP molecules, respectively. $R_{\ell v}$ is the distance between the centroids of these LMOs. The C_6 coefficient between the two LMOs, averaged over the directional components, $C_6^{\ell v}$, is

$$C_6^{\ell v} = \frac{1}{3} \cdot \frac{6}{\pi} \sum_a^{x,y,z} C_{6,a}^{\ell v} = \frac{2}{\pi} \sum_a^{x,y,z} C_{6,a}^{\ell v}, \tag{31}$$

where the directional components are

$$C_{6,a}^{\ell v} = \sum_r^{vir} \langle \ell | \hat{\mu}_a | r \rangle \sum_{\ell'} \langle r | \hat{\mu}_a | \ell' \rangle \sum_k^{valence} T_k^{\ell \ell'} \times \int_0^\infty d\omega \frac{\omega_{rk}^A}{\left(\omega_{rk}^A\right)^2 + \omega^2} \alpha^v(i\omega). \tag{32}$$

In Eq. (32), k sums over the orbitals in the valence space and

$$T_k^{\ell \ell'} = L_{k\ell} L_{k\ell'}. \tag{33}$$

The matrix $L_{k\ell}$, obtained by performing Boys localization[26] on the valence MOs, is the orthogonal transformation that expresses the localized AI orbitals $|\ell\rangle$ in terms of the canonical AI orbitals $|k\rangle$ according to

$$|\ell\rangle = \sum_k |k\rangle L_{k\ell}, \quad |\ell\rangle = \text{localized}, \quad |k\rangle = \text{canonical}. \tag{34}$$

For convenience, the integral over imaginary frequencies in Eq. (32), which includes canonical MO terms ω_{rk} in the AI region, is kept in the MO basis. Substituting Eq. (34) for ℓ and ℓ' into Eq. (32) and taking into account the orthogonality of $L_{k\ell}$ yields

$$C_{6,a}^{\ell v} = \sum_k^{valence} \sum_{k'}^{valence} L_{k\ell} \left[\sum_r^{vir} \langle k | \hat{\mu}_a | r \rangle \langle r | \hat{\mu}_a | k' \rangle \times \int_0^\infty d\omega \frac{\omega_{rk'}^A}{\left(\omega_{rk'}^A\right)^2 + \omega^2} \alpha^v(i\omega) \right] L_{k'\ell}. \tag{35}$$

The LMO-distributed quantities $C_6^{\ell v}$ obtained by this method are not identical to the MO-distributed C_6^{kj} of Eq. (29), but they are equivalent through the transformation to the LMO basis. To ensure fast execution of the code, only the valence orbitals (omitting core orbitals) are used in summations in Eq. (35). Equation (35) is implemented in the GAMESS code for EFP-AI dispersion.

The distributed LMO-based C_6 coefficient in Eq. (35) is comprised of dipole integrals over AI orbitals and an integral over the imaginary frequency range. The imaginary frequency integral for EFP2 LMO v with AI occupied orbital k and virtual orbital r is given by

$$I^{v,kr} = \int_0^\infty d\omega \frac{\omega_{rk}^A}{\left(\omega_{rk}^A\right)^2 + \omega^2} \bar{\alpha}^v(i\omega), \quad \text{where } \omega_{rk} = \omega_r - \omega_k. \tag{36}$$

This integral is computed using a 12-point Gauss-Legendre quadrature, in analogy with the EFP2-EFP2 dispersion expression.[1] To convert this integral to a range that is conducive to the use of a Gauss-Legendre numerical quadrature, the same transformation of variables that is used in the EFP2-EFP2 dispersion procedure[6] may be applied:

$$\omega = \nu_0 \frac{(1+t)}{(1-t)}, \quad d\omega = \frac{2\nu_0 dt}{(1-t)^2}. \tag{37}$$

Equation (37) converts the range of integration to be from -1 to $+1$. By using Gauss-Legendre abscissas for t, this procedure also determines the values of the (imaginary) frequencies at which the polarizability tensors must be calculated. The substitution shown in Eq. (37) was determined in Ref. 27, where the optimal value of ν_0 was found to be 0.3. The imaginary frequency integral in Eq. (36) becomes

$$I^{v,kr} = \sum_{n=1}^{12} w_n \frac{2\nu_0}{(1-t_n)^2} \frac{\omega_{rk}}{\omega_{rk}^2 + \omega_n^2} \bar{\alpha}^v(i\omega_n), \quad \text{where} \quad \omega_n = \nu_0 \frac{(1+t_n)}{(1-t_n)}. \tag{38}$$

Here, t_n and w_n are Gauss-Legendre abscissas and weights, and $\bar{\alpha}^v(i\omega_n)$ is 1/3 of the trace of the polarizability tensor for LMO v at frequency ω_n (Eq. (18)). Since the MO energies and the factors ω_{rk} in Eq. (38) differ by $\hbar$, which is 1 in atomic units, the energy difference $\boldsymbol{\varepsilon_{rk}}$ between virtual MO r and occupied MO k is used in place of the frequency difference ω_{rk}

$$I^{v,kr} = \sum_{n=1}^{12} w_n \frac{2\nu_0}{(1-t_n)^2} \frac{\varepsilon_{rk}}{\varepsilon_{rk}^2 + \omega_n^2} \bar{\alpha}^v(i\omega_n), \quad \text{where} \quad \boldsymbol{\varepsilon_{rk}} = \boldsymbol{\varepsilon_r} - \boldsymbol{\varepsilon_k}. \tag{39}$$

Upon computation of the integral $I^{v,kr}$ in Eq. (39) for a given EFP2 LMO v and AI MOs k and r, $I^{v,kr}$ is multiplied by the canonical MO-basis dipole integrals obtained by the standard transformation

$$\mu_a^{A,\text{occ}\times\text{vir MOs}} = OCC^T \cdot \mu_a^{A,\text{AOs}} \cdot VIR, \tag{40}$$

where $\mu_a^{A,\text{AOs}}$ is the matrix of dipole integrals with elements $\langle \chi_1 | \hat{\mu}_a^A | \chi_2 \rangle$ over all pairs of AOs χ_1 and χ_2, OCC consists of the $N_{valence} = N_{occupied} - N_{core}$ (number of valence orbitals) columns of the MO coefficient matrix, and VIR consists of the remaining $N_{total} - N_{occupied}$ virtual columns. This produces a matrix $\mu_a^{A,\text{occ}\times\text{vir MOs}}$ with dimensions $N_{valence} \times (N_{total} - N_{occupied})$ corresponding to the dipole integrals between valence (k) $\times$ virtual (r) canonical MOs; its elements are $\langle k | \hat{\mu}_a^A | r \rangle$. The product of the imaginary frequency integral $I^{v,kr}$ with the dipole integrals is transformed to the LMO basis using the Boys transformation matrix $L_{k\ell}$ as shown in Eq. (35). Finally, each $C_6^{\ell v}$ coefficient is multiplied by $R_{\ell v}^{-6}$, calculated between the centroids of *ab initio* LMO ℓ and EFP2 LMO v. Summing over all ℓ and v gives the total dipole-dipole contribution to the EFP2-AI dispersion energy, Eq. (30).

Two further approximations must be invoked in order to obtain the total dispersion energy from the dipole-dipole dispersion term. First, numerous theoretical studies suggest that, in order to compare with experimental or high-level *ab initio* calculations, a damping function must be used with the dispersion energy term.[28–35] The damping function, $F_6^{\ell v}$, is employed to account for both exchange-dispersion and charge

penetration effects, which predominate at short intermolecular separations. Additionally, as with EFP2-EFP2 dispersion,[6] the total EFP2-AI dispersion energy expansion in Eq. (27) is truncated after the C_8/R^8 term, which is approximated as 1/3 of the C_6/R^6 term. This approximation was determined in Ref. 6 by comparison of EFP-EFP dispersion energies truncated at the C_6/R^6 term with the total SAPT dispersion energies for a selection of dimers. Thus, the final form for the EFP2-AI dispersion energy becomes

$$E_{disp}^{EFP-AI} = -\frac{4}{3}\sum_{\ell\in A}\sum_{v\in B}\frac{F_6^{\ell v}C_6^{\ell v}}{R_{\ell v}^6}. \tag{41}$$

As was done for the EFP2-EFP2 dispersion interaction,[6] both Tang-Toennies (TT)[6,36,37] and overlap-based[37] damping functions have been explored. The sixth-order Tang-Toennies damping function has the form

$$F_6^{TT}(R_{\ell v}) = 1 - \exp(-\beta R_{\ell v})\sum_{n=1}^{6}\frac{(\beta R_{\ell v})^n}{n!}, \tag{42}$$

where β is a damping parameter chosen to be 1.5, based on a study of several small dimers at their equilibrium distances.[6] The overlap-based damping function is given by[37]

$$F_6^{overlap}(S_{\ell v}) = 1 - S_{\ell v}^2(1 - 2\ln|S_{\ell v}| + 2(\ln|S_{\ell v}|)^2), \tag{43}$$

where $S_{\ell v}$ is the overlap integral calculated between LMOs $|\ell\rangle$ and $|v\rangle$. Dispersion energies obtained using each type of damping are presented in Sec. V.

IV. COMPUTATIONAL DETAILS

EFP2-EFP2 and EFP2-AI dispersion calculations were performed with the GAMESS[10] software package. SAPT[12] interaction energies were calculated using SAPT-2006.

The dimers chosen for this study—argon, methane, H_2, HF, water, ammonia, methanol (MeOH), dichloromethane, and the sandwich and T-shaped benzene dimers—are identical to those used in a previous study of EFP2 damping functions[37] (Fig. 1). Equilibrium geometries of CH_4, H_2, MeOH, and NH_3 are taken from Ref. 6. The CH_2Cl_2 dimer geometry is taken from Ref. 38. The HF dimer geometry, from Ref. 39, was obtained with coupled cluster with singles, doubles, and perturbative triples[40] [CCSD(T)] in the complete basis set limit. The Ar dimer geometry, from Ref. 41, was obtained with CCSD(T)/aug-cc-pVQZ.[42,43] The geometry of the H_2O monomer was obtained in Ref. 37 by optimizing with CCSD(T)/cc-pVTZ (Ref. 43); this was followed by a constrained optimization with MP2/6-311++G(3df,2p)[44–46] to obtain the dimer geometry.[37] The sandwich and T-shaped benzene dimer geometries are taken from Ref. 47.

A. Comparison of C_6 coefficients

The total EFP2-EFP2 and EFP2-AI C_6 coefficients, equal to the sum over all LMO pairs of the distributed C_6 values, were obtained for all dimers. The 6-311++G(3df,3p) basis set[44–46] was used to generate the EFP2 potentials and to perform the AI Hartree-Fock calculations for all monomer types except benzene. For benzene, the 6-311++G(3df,2p) basis set[44–46] was used. Experimental C_6 values from Refs. 1 and 27 are used for comparison.

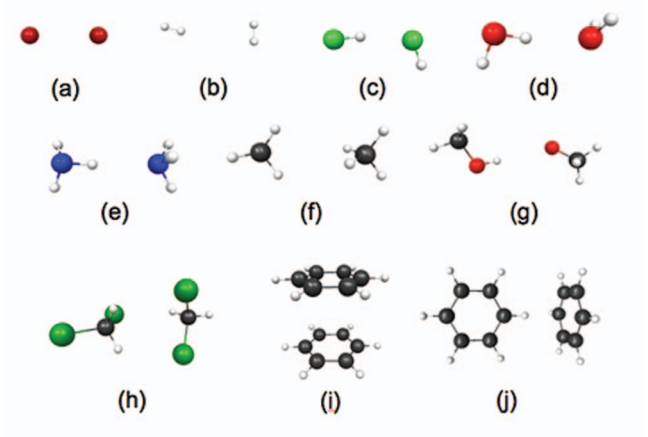

FIG. 1. Dimers examined in the present study. Equilibrium geometries are shown for (a) argon, (b) H_2 dimer, (c) HF, (d) water, (e) ammonia, (f) methane, (g) methanol (MeOH), and (h) dichloromethane. Dimer geometries for (i) sandwich benzene dimer and (j) T-shaped benzene dimer are constrained structures.

B. Comparison of potential energy curves

As in Ref. 37, potential energy curves were generated for all dimers except the benzene dimers by varying intermonomer distances in increments of 0.2 Å from –0.8 Å to +0.8 Å with respect to the equilibrium distance, while keeping the internal monomer geometries fixed. For the benzene sandwich dimer, the distance between the benzene rings was varied from 3.3 Å to 6.0 Å. For the T-shaped benzene dimer, the ring center to ring center distance was varied from 4.7 Å to 6.9 Å. The 6-311++G(3df,2p) basis set was used to generate the EFP2 parameters for EFP2-EFP2 and EFP2-AI calculations on the Ar, H_2, H_2O, CH_4, NH_3, HF, and sandwich and T-shaped benzene dimers. The 6-311++G(2d,2p) basis[44] was used with the MeOH dimers. For CH_2Cl_2 dimers, the 6-31+G(d) basis[48] was used. In EFP2-AI calculations, the same basis used for the EFP2 monomer was also used on the AI molecule. For all dimer types other than the benzene dimers, second-level SAPT calculations were carried out using the same basis set employed for the EFP2-EFP2 and EFP2-AI calculations. SAPT data for the benzene dimers come from Ref. 47, in which the aug-cc-pVDZ basis[42] was used. The SAPT reference values reported in Sec. V are the sum of the SAPT second-order dispersion and exchange-dispersion values for a given dimer geometry.

V. RESULTS AND DISCUSSION

The neon dimer provides a concrete example of the process for determining the C_6 coefficients for an EFP2-AI system. Having four valence orbitals (lone pairs), a neon atom is modeled with EFP2 as four α polarizability tensors on the

TABLE I. EFP2-EFP2 and EFP2-AI isotropic C_6 coefficients (a.u.)[a,b]

	AI (left)-EFP2 (right) C_6	EFP2 (left)-AI (right) C_6	EFP2-EFP2 C_6	Experimental[c]
Ar	67.4 (+4.8%)	(Identical)	60.6 (−5.8%)	64.3[d]
H_2	10.2 (−15.7%)	10.2 (−15.7%)	11.4 (−5.8%)	12.1[e]
HF	17.0 (−10.5%)	15.4 (−18.9%)	15.3 (−19.5%)	19.0[e]
H_2O	43.0 (−5.3%)	40.6 (−10.6%)	39.3 (−13.4%)	45.4[e]
NH_3	81.2 (−7.0%)	82.1 (−6.0%)	78.1 (−10.5%)	87.3[e]
CH_4	120.3 (−7.2%)	(Identical)	120.4 (−7.1%)	129.6[e]
MeOH	197.7 (−11.1%)	197.2 (−11.3%)	195.8 (−11.9%)	222.2[e]
CH_2Cl_2	843.0	843.0	755.8	N/A
C_6H_6 (*)[f]	2087 (+21.1%)	2087 (+21.1%)	1805 (+4.8%)	1723[d]

[a]Calculated using the 6-311++G(3df,3p) basis set for all dimers except the benzene dimer (*), which was calculated with the 6-311++G(3df,2p) basis set; for EFP2-EFP2 and EFP2-AI, the percent error compared to the experimental C_6 value (if available) is shown in parentheses. "(Identical)" indicates that there is no difference between AI (left) and AI (right).

[b]For nonsymmetrical EFP-AI dimers, "left" and "right" correspond to the monomer on the left or right in the dimers pictured in Fig. 1; in hydrogen bonded species, "left" corresponds to the hydrogen bond donor and "right" to the hydrogen bond acceptor.

[c]Experimental values for footnotes (d) and (e).

[d]From Ref. 27.

[e]From Ref. 1.

[f]EFP2-AI C6 values for sandwich and T-shaped benzene dimers are identical to four significant figures.

orbital centroids. The EFP2-EFP2 treatment of dispersion entails integrals between each α on fragment A with every α on fragment B, giving 16 C_6^{pj} integrals. The EFP2-AI approach outlined above produces four matrices (one for each α^j on the EFP2 part) of dimension 4 × 4 (indexed by the AI valence LMOs). The diagonal elements of these four matrices give the distributed C_6^{pj} values between the valence orbital p on the AI molecule and the EFP2 expansion point j. For comparison, the distributed EFP2-EFP2 C_6^{pj} coefficients for the neon dimer, calculated with the 6-311++G(3d) basis set, are all approximately 0.286984097, differing from one another beginning in the ninth decimal place. The distributed EFP2-AI coefficients, with both the AI and the EFP2 part calculated using the 6-311++G(3d) basis set, are all approximately 0.2857, differing from one another after the third decimal place.

Unlike the EFP2-EFP2 C_6 coefficient, the EFP2-AI C_6 coefficient varies somewhat with intermonomer separation. This is because the AI orbitals respond to the perturbative presence of the EFP2 fragment. The EFP2 Coulomb, exchange-repulsion, and induction energies are iterated into the AI Hartree-Fock calculation (the dispersion energy itself is not iterated, but added as a correction to the final Hartree-Fock energy). In contrast, EFP2 LMOs are held fixed. The EFP2-EFP2 C_6 coefficient is calculated from dynamic polarizabilities that are determined entirely from a MAKEFP calculation on the individual monomers. Since these values do not change in the construction of the dimer, the resulting C_6 value is completely independent of distance. The distance dependence of the EFP2-AI C_6 coefficient is generally not significant except at very close intermonomer separations, where it becomes very large. The use of a distance- or overlap-dependent damping function helps to mitigate this instability of C_6 as the intermonomer separation approaches zero.

A comparison of the EFP2-AI and EFP2-EFP2 total C_6 coefficients (the sum of the $C_6^{\ell v}$ values over every LMO $|\ell\rangle$and $|v\rangle$) appears in Table I. For EFP2-AI, the reported C_6 value was calculated at the equilibrium separation for each dimer, except the benzene dimer. There are two possible ways to model the nonsymmetric dimers, depending on which monomer is AI and which is EFP2. Both possibilities are reported in Table I, where "left" and "right" refer to the relative positions of the monomers as they appear in Fig. 1. In hydrogen bonded dimers, the "left" monomer is the hydrogen bond donor and the "right" monomer is the hydrogen bond acceptor. The calculated EFP2-AI benzene dimer C_6 values are the same, to the given number of significant figures, for the sandwich and T-shaped configurations. Compared to experimental values,[1,27] EFP2-EFP2 C_6 coefficients are underestimated (by 5.8% to 19.5%) for all dimer types except the benzene dimer, for which C_6 is slightly (4.8%) overestimated. For all dimer types other than argon and benzene, the EFP2-AI C_6 coefficients are similarly underestimated compared to the experimental values (5.3% to 18.9%). The EFP2-AI C_6 value for the argon dimer is slightly overestimated (4.8%), and that for the benzene dimer is rather overestimated compared to the experimental value (21.1%).

The benzene molecule is a rare case in which the Hartree-Fock method overestimates the static polarizability.[6] Typically, HF underestimates this quantity.[6] In Ref. 6, the static polarizability for benzene, determined with CPHF, was found to be 13% greater than the experimental value with the largest basis set examined [6-311G(3df,3pd) in that study[6]]. The TDHF dynamic polarizability was shown[6] to exhibit a clear basis set dependence as well. Since the C_6 coefficient is a function of the polarizability and the static polarizability (at zero frequency) is the leading term in the integration over the imaginary frequency range, it is unsurprising that both the EFP2-EFP2 and EFP2-AI benzene dimer C_6 values are overestimated (Table I) with the large 6-311++G(3df,2p) basis set used.

For all dimers in Table I other than benzene (and excluding CH_2Cl_2, for which experimental data are not available), the average absolute error in the EFP2-EFP2 C_6 value is 10.6% compared to experiment. The average absolute error in C_6 over all unique EFP2-AI dimers other than benzene is 10.3%. The largest EFP2-AI error (besides that for

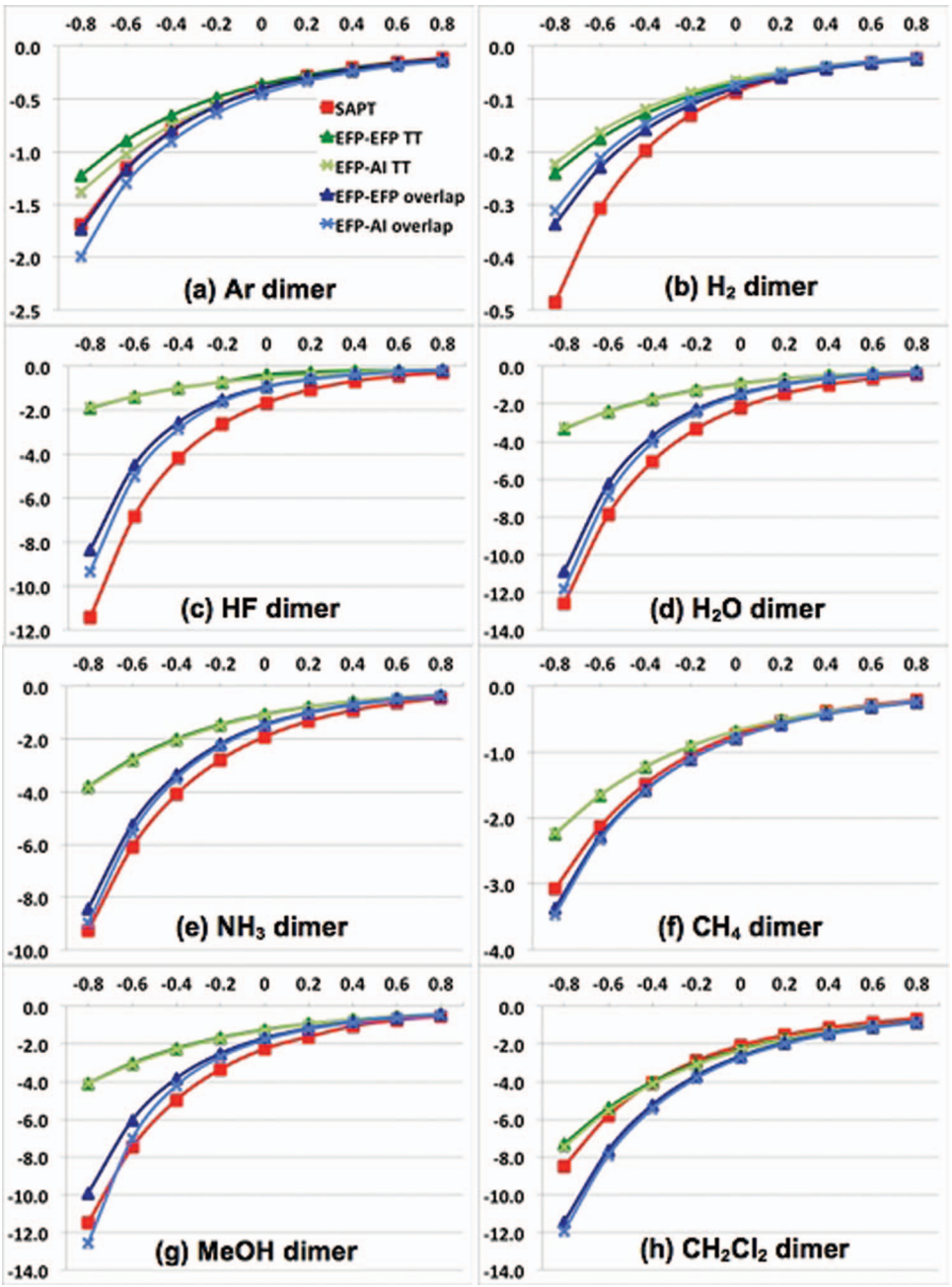

FIG. 2. Dimer dispersion energies (kcal/mol) as a function of displacement (in Å) from the equilibrium separation (denoted 0 Å). The reference SAPT values (red lines) are the sums of dispersion and exchange-dispersion energies at each dimer geometry. Dispersion energies plotted for both EFP2-EFP2 and EFP2-AI were calculated with each of the two possible damping functions: Tang-Toennies (dark green for EFP2-EFP2, light green for EFP2-AI) and overlap-based damping (dark blue for EFP2-EFP2, light blue for EFP2-AI).

benzene) occurs with the HF dimer when the hydrogen bond acceptor monomer is modeled with Hartree-Fock and the hydrogen bond donor monomer with EFP2; the magnitude of this error is 18.9%. When the EFP2 and *ab initio* monomers are switched, so that the hydrogen bond acceptor is modeled with EFP2 and the hydrogen bond donor with Hartree-Fock, this error decreases to 10.5%. Similarly, for the water dimer, the calculated C_6 coefficient differs from the experimental value by 10.6% when Hartree-Fock is used to model the hydrogen bond acceptor, but by only 5.3% when Hartree-Fock is used to model the hydrogen bond donor. However, in the methanol dimer, the errors are very similar regardless of which monomer is modeled with EFP2 and which with Hartree-Fock (both are ∼11%). For the ammonia dimer, the trend reverses: there is a difference of 7.0% between calculated and experimental C_6 values when the hydrogen bond donor is modeled with Hartree-Fock, but the difference decreases slightly to 6.0% when the hydrogen bond acceptor is modeled with Hartree-Fock. For other nonsymmetric dimers (H_2, CH_2Cl_2, T-shaped benzene), the C_6 value does not depend on which monomer is modeled with EFP2 and which is modeled with Hartree-Fock.

The difference between EFP2-EFP2 and EFP2-AI C_6 coefficients has two origins. The first origin arises from the

TABLE II. Effect on EFP2-AI dispersion energy (kcal/mol) for selected nonsymmetrical dimers when EFP2 and *ab initio* monomers are switched.

	AI (left)-EFP (right)	EFP (left)-AI (right)
H_2	−0.10	−0.10
HF	−1.11	−1.01
H_2O	−1.78	−1.71
NH_3	−1.56	−1.70
MeOH	−0.50	−0.50
CH_2Cl_2	−4.34	−4.44

Equilibrium dimer geometries. 6-311++G(3df,3p) basis set used for both EFP2 and *ab inito* (Hartree-Fock) monomers. "Left" and "right" correspond to the monomer on the left or right in the dimers pictured in Fig. 1. (In hydrogen bonded species, "left" corresponds to the hydrogen bond donor and "right" to the hydrogen bond acceptor.)

mathematical formulae used with the respective methods. The EFP2-EFP2 C_6 coefficient is a function of polarizability tensors calculated via the time-dependent Hartree-Fock equations, whereas the EFP2-AI approach relies on applying the orbital approximation to a sum-over-states formula. The second difference between EFP2-EFP2 and EFP2-AI C_6 values results from the use of perturbed orbitals in EFP2-AI, as discussed above. EFP2 orbitals are held fixed, but the perturbing field of the EFP2 fragment is taken into account in constructing the orbitals on the *ab initio* monomer; these orbitals are, in turn, used in the calculation of C_6. Since monomers subject to a stronger electrostatic field will have correspondingly more perturbed orbitals, this aspect of the EFP2-AI C_6 calculation is distance dependent. As a result of these two differences, the EFP2-AI method is not necessarily invariant to the choice of EFP2 and AI monomers (as noted above). However, despite the differences in C_6 values, the difference in the final dispersion energy appears to be small: on the order of a tenth of a kcal/mol for the nonsymmetrical dimers examined (Table II).

The total dipole-dipole term of the dispersion energy is given by dividing each of the distributed $C_6^{\ell v}$ values by $R_{\ell v}^6$, where R is the distance between the centroids of LMOs $|\ell\rangle$ and $|v\rangle$, then summing over all ℓ and v. Given the general similarity of EFP2-EFP2 and EFP2-AI C_6 values (Table I), it is reasonable to expect that EFP2-AI damping functions analogous to those of EFP2-EFP2 (Ref. 37) may be used. Additionally, the approximation[6] utilized with EFP2-EFP2 to account for higher order terms in the dispersion expansion of Eq. (27) appears to apply equally well for the EFP2-AI dispersion: the expansion in Eq. (27) is truncated after C_8/R^8, which is then approximated[6] as 1/3 of the total dipole-dipole dispersion term C_6/R^6. With the use of a damping function and the approximation for C_8/R^8, EFP2-AI dispersion energy values compare well with those predicted by SAPT (Figs. 2 and 3).

Figures 2 and 3 compare EFP2-EFP2 and EFP2-AI dispersion energies, each with the two possible damping functions, and SAPT values. Distances in Fig. 2, for the dimers shown in Figs. 1(a)–1(h), are given as displacements with respect to the equilibrium intermonomer separation (denoted by 0 Å). Distances in Fig. 3, for the sandwich and T-shaped benzene dimers (Figs. 1(i) and 1(j)), are the distances between ring centers. The SAPT values (red line) are the sum of SAPT dispersion and exchange-dispersion values, as the EFP2-EFP2 and EFP2-AI methods do not explicitly model exchange-dispersion as a quantity separate from the total dispersion energy. For all dimers examined, EFP2-EFP2 Tang-Toennies damped values (dark green line) and EFP2-AI TT damped values (light green line) are very similar to one another, as are EFP2-EFP2 overlap-based damped values (dark blue line) and EFP2-AI overlap-based damped values (light blue line). The difference between the two damping functions in Fig. 2 is more significant for hydrogen-bonded dimers (HF, H_2O, MeOH) than for dimers bound primarily by dispersive forces (Ar, H_2, CH_4; benzene dimers in Fig. 3). This is due to the $R_{\ell v}$ or overlap dependence of the respective damping functions: hydrogen-bonded dimers tend to have shorter equilibrium intermonomer separations compared to dispersion-bound dimers. The Tang-Toennies damping function tends to produce dispersion energies that are too weak compared to SAPT values. Thus, due to the superior overall agreement of overlap-based damping with the SAPT values (Figs. 2 and 3), as well as the fact that overlap-based damping is free of arbitrary parameters that must otherwise be fitted in some manner, this is the preferred damping function for use with both EFP2-EFP2 (Ref. 37) and EFP2-AI dispersion.

The speed of the EFP2-AI dispersion calculation makes this method appealing for calculations involving large numbers of molecules modeled with EFP2, as in solute-solvent interactions. For all of the dimers examined, calculation of

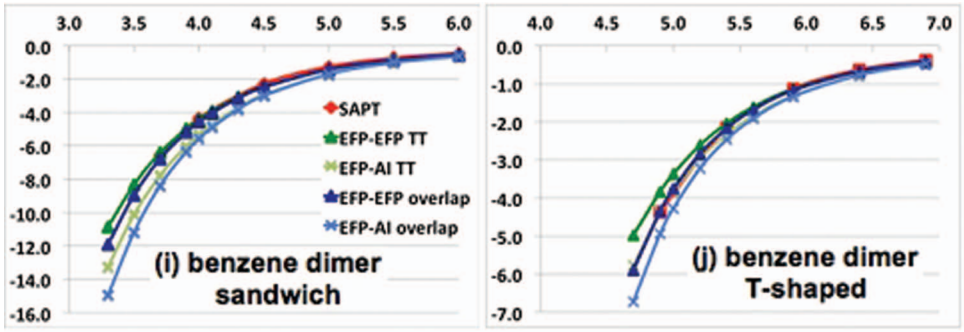

FIG. 3. Sandwich (i) and T-shaped (j) benzene dimer dispersion energies (kcal/mol) as a function of distance between ring centers (Å). The reference SAPT values (red lines) are the sums of dispersion and exchange-dispersion energies at each dimer geometry. Dispersion energies plotted for both EFP2-EFP2 and EFP2-AI were calculated with each of two possible damping functions: Tang-Toennies (dark green for EFP2-EFP2, light green for EFP2-AI) and overlap-based damping (dark blue for EFP2-EFP2, light blue for EFP2-AI).

the EFP2-AI dispersion term requires a fraction of a second. Even for the largest system examined (benzene dimer, 330 basis functions on the AI monomer, no symmetry imposed), Hartree-Fock convergence requires ~100 min on a given computer, while the dispersion energy calculation, including orbital localization, requires less than a second. The computational cost scales linearly with the number of EFP2 fragments (more precisely, with the total number of EFP2 LMOs).

VI. CONCLUSION

EFP2-AI distributed C_6 coefficients can be obtained by multiplying AI dipole integrals with an integral that consists of the EFP2 dynamic polarizability tensor and a function of AI orbital energies. The product is transformed from the canonical to the localized molecular orbital basis using a Boys localization[26] matrix. This gives a distributed C_6 value for each pair of LMOs $|\ell\rangle$ on the *ab initio* molecule and $|v\rangle$ on the EFP2 fragment. For all systems examined, this method yields values similar to the distributed C_6 coefficients in the EFP2-EFP2 dispersion interaction. The total EFP2-AI dipole-dipole dispersion interaction energies are obtained by multiplying the distributed $C_6^{\ell v}$ values by $R_{\ell v}^{-6}$, where $R_{\ell v}$ is the distance between the centroids of LMOs $|\ell\rangle$ and $|v\rangle$. As with the EFP2-EFP2 dispersion, the dispersion expansion is truncated after C_8/R^8; the latter term is approximated as 1/3 of the C_6/R^6 contribution.

This EFP2-AI method is shown to yield good agreement with both experimentally determined C_6 coefficients and theoretically determined (SAPT) dispersion energies for a variety of dimers examined. The EFP2-AI dispersion energies also agree well with those predicted by the EFP2-EFP2 method. Compared to experimental values, the average absolute error in the C_6 coefficient is 10.3% for dimers other than benzene. The C_6 coefficient is underestimated for most dimers, although it is overestimated by more than 20% in benzene dimers. This is a result of the observation that the coupled perturbed and time-dependent Hartree-Fock methods overestimate the benzene static and dynamic polarizability tensors, which are used in the calculation of the C_6 coefficient.

Two damping functions, an overlap-based damping function and a Tang-Toennies damping function, are evaluated for use with the EFP2-AI dispersion energy calculation. A damping function is necessary because the multipole expansion is not guaranteed to converge at short intermonomer separations. As previously recommended for EFP2-EFP2 dispersion,[37] the parameter-free overlap-based damping function is found to give better overall agreement with SAPT results.

ACKNOWLEDGMENTS

This work was supported in part by a grant from the U.S. Air Force Office of Scientific Research (US AFOSR) (Q.A.S., M.S.G.), by a National Science Foundation (NSF) Multiscale Modeling grant (M.S.G., L.V.S.), by a National Science Foundation Petascale Applications grant (Q.A.S., M.S.G.), by a National Science Foundation CAREER grant (L.V.S.), and by the Division of Chemical Sciences, Office of Basic Energy Sciences, U.S. Department of Energy (DOE) under Contract No. DE-AC02-07CH11358 with Iowa State University through the Ames Laboratory (K.R.). The authors are grateful for many helpful discussions with Dr. Ivana Adamovic.

APPENDIX A: CONVERSION FROM A SUM OVER STATES TO AN ORBITAL BASED APPROACH

$$\sum_{m\neq 0} \langle 0_A|\,\widehat{\mu}_a^A\,|m\rangle\,\langle m|\,\widehat{\mu}_c^A\,|0_A\rangle \rightarrow \sum_k^{occ}\sum_r^{vir} \langle k|\,\widehat{\mu}_a\,|r\rangle\,\langle r|\,\widehat{\mu}_c\,|k\rangle \tag{A1}$$

Each of the (**singly**) excited states m in Eq. (A1) differs from the ground state 0_A by one spin orbital.

From Ref. 49:

Let $|K\rangle = |\ldots mn\ldots\rangle$ and $|L\rangle = |\ldots pn\ldots\rangle$, i.e., $|K\rangle$ and $|L\rangle$ are Slater determinants differing by one spin orbital (denoted m and p for each determinant, respectively).

Also let the sum of one-electron operators be $\vartheta_1 = \sum_{i=1}^N h(i)$.

Then

$$\langle K|\,\vartheta_1\,|L\rangle = \langle m|\,h\,|p\rangle\,.$$

Let excited state $m = |\ldots r\ldots\rangle$ differ from the ground state $0_A = |\ldots k\ldots\rangle$ in a single spin orbital ($k \rightarrow r$). (Note that *assuming single excitations does not result in any loss of generality*; Slater determinants with doubly, triply, and higher excited states give integrals $\langle K|\vartheta_1|L\rangle$ equal to zero.) Then, according to the above,

$$\langle 0_A|\,\mu_a\,|m\rangle = \langle k|\,\mu_a\,|r\rangle\,.$$

Since the sum on the LHS of Eq. (A1) goes over all excited states (all possible values of spin orbital k in ground state 0_A are individually "excited" to all possible values of virtual spin orbitals r in excited state m), sums over all occupied and all virtual orbitals are included

$$\sum_{m\neq 0} \langle 0_A|\,\mu_a\,|m\rangle = \sum_k^{occ}\sum_r^{vir} \langle k|\,\mu_a\,|r\rangle.$$

Doing this for both integrals on the LHS of Eq. (A1) gives the RHS as indicated.

APPENDIX B: INTEGRAL OVER IMAGINARY FREQUENCY RANGE

$$\int_0^\infty d\omega \frac{\omega_{m0}^A}{\left(\omega_{m0}^A\right)^2+\omega^2}\alpha_{bd}^B(i\omega) \rightarrow \int_0^\infty d\omega \frac{\omega_{rk}^A}{\left(\omega_{rk}^A\right)^2+\omega^2}\alpha_{bd}^B(i\omega)\,,$$

$$\text{where} \quad \omega_{rk} = \omega_r - \omega_k. \tag{B1}$$

The quantity ω_{m0} was previously defined by $E_{m0}^A = \hbar\omega_{m0}^A$, where $E_{m0} = E_m - E_0$ is the excited state energy (so ω is a frequency multiplied by 2π).

Equation (B1) can be justified by a generalization of Koopman's theorem. The difference between the ground and excited states is the subtraction of an electron from an occupied orbital (with energy ε_i or associated frequency $\times 2\pi$ ω_i) and the addition of an electron to a virtual orbital (with energy ε_k or associated frequency $\times 2\pi$ ω_k). Invoking the **frozen orbital approximation** (neglecting the effect of correlation with changing numbers of electrons), subtracting an electron from an occupied orbital changes the Hartree-Fock energy E_{HF} by $-\varepsilon_k$; adding an electron to a previously unoccupied orbital changes the HF energy by $+\varepsilon_r$. Thus,

$$\begin{aligned}\omega_{m0} &= \omega_m - \omega_0 = \frac{1}{\hbar}(E_m - E_0)\\ &= \frac{1}{\hbar}([E_{HF} + \varepsilon_r - \varepsilon_k] - E_{HF}) = \frac{1}{\hbar}(\varepsilon_r - \varepsilon_k)\\ &= \omega_r - \omega_k = \omega_{rk}.\end{aligned}$$

[1] A. J. Stone, *The Theory of Intermolecular Forces* (Clarendon, Oxford, 1996).
[2] R. McWeeny, *Methods of Molecular Quantum Mechanics*, 2nd ed. (Academic, London, 1989).
[3] R. D. Amos, N. C. Handy, P. J. Knowles, J. E. Rice, and A. J. Stone, J. Phys. Chem. **89**, 2186 (1985).
[4] A. G. Ioannou, S. M. Colwell, and R. D. Amos, Chem. Phys. Lett. **278**, 278 (1997).
[5] R. Eisenschitz and F. Z. London, Physik **60**, 491 (1930); F. Z. London, Physik **63**, 245 (1930); F. Z. London, Phys. Chem. **33**, 8 (1937).
[6] I. Adamovic and M. S. Gordon, Mol. Phys. **103**, 379 (2005).
[7] P. N. Day, J. H. Jensen, M. S. Gordon, S. P. Webb, W. J. Stevens, M. Krauss, D. Garmer, H. Basch, and D. Cohen, J. Chem. Phys. **105**, 1968 (1996).
[8] M. S. Gordon, M. A. Freitag, P. Bandyopadhyay, J. H. Jensen, V. Kairys, and W. J. Stevens, J. Phys. Chem. A **105**, 293 (2001).
[9] M. S. Gordon, L. V. Slipchenko, H. Li, and J. H. Jensen, Annu. Rep. Comp. Chem. **3**, 177 (2007).
[10] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993); M. S. Gordon and M. W. Schmidt, in *Theory and Applications of Computational Chemistry*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, 2005).
[11] C. Moller and S. Plesset, Phys. Rev. **46**, 618 (1934).
[12] B. Jeziorski, R. Moszynski, and K. Szalewicz, Chem. Rev. **94**, 1887 (1994).
[13] M. A. Freitag, M. S. Gordon, J. H. Jensen, and W. J. Stevens, J. Chem. Phys. **112**, 7300 (2000).
[14] A. J. Stone, Chem. Phys. Lett. **83**, 233 (1981).
[15] J. H. Jensen and M. S. Gordon, Mol. Phys. **89**, 1313 (1996).
[16] D. D. Kemp, J. M. Rintelman, M. S. Gordon, and J. H. Jensen, Theor. Chem. Acc. **125**, 481 (2010).
[17] H. Li, M. S. Gordon, and J. H. Jensen, J. Chem. Phys. **124**, 214108 (2006).
[18] D. Ghosh, D. Kosenkov, V. Vanovschi, C. F. Williams, J. M. Herbert, M. S. Gordon, M. W. Schmidt, L. V. Slipchenko, and A. I. Krylov, J. Phys. Chem. A **114**, 12739 (2010).
[19] B. T. Thole and P. Th. van Duijnen, Chem. Phys. **71**, 211 (1982); P. Th. van Duijnen and A. H. de Vries, Int. J. Quantum Chem. **60**, 1111 (1996).
[20] A. J. Stone, Mol. Phys. **56**, 1065 (1985).
[21] P. Claverie and R. Rein, Int. J. Quantum Chem. **3**, 537 (1969).
[22] O. Engkvist, P. O. Åstrand, and G. Karlström, Chem. Rev. **100**, 4087 (2000).
[23] P. Claverie, Int. J. Quantum Chem. **5**, 273 (1971).
[24] H. B. G. Casimir and D. Polder, Phys. Rev. **73**, 360 (1948).
[25] Y. Yamaguchi, J. D. Goddard, Y. Osamura, and H. F. Schaefer, *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory* (Oxford, New York/Oxford, 1994).
[26] S. F. Boys, Rev. Mod. Phys. **32**, 306 (1960); C. Edmiston and K. Ruedenberg, *ibid.* **35**, 457 (1963).
[27] E. K. Gross, C. A. Ullrich, and U. J. Gossmann, NATO ASI Ser., Ser. B **337**, 149 (1995).
[28] J. T. Ajit, J. Chem. Phys. **89**, 2092 (1988).
[29] D. Cvetko, A. Lausi, A. Morgante, F. Tommasini, P. Cortona, and M. G. Dondi, J. Chem. Phys. **100**, 2052 (1994).
[30] A. G. Donchev, Phys. Rev. B **74**, 235401 (2006).
[31] G. Ihm, W. C. Milton, T. Flavio, and G. Scoles, J. Chem. Phys. **87**, 3995 (1987).
[32] S. H. Patil, K. T. Tang, and J. P. Toennies, J. Chem. Phys. **116**, 8118 (2002).
[33] M. Wilson, P. A. Madden, N. C. Pyper, and J. H. Harding, J. Chem. Phys. **104**, 8068 (1996).
[34] A. J. Stone and A. J. Misquitta, Int. Rev. Phys. Chem. **26**, 193 (2007).
[35] R. J. Wheatley and W. Meath, J. Mol. Phys. **80**, 25 (1993).
[36] K. T. Tang and J. P. Toennies, J. Chem. Phys. **80**, 3726 (1984).
[37] L. V. Slipchenko and M. S. Gordon, Mol. Phys. **107**, 999 (2009).
[38] L. V. Slipchenko and M. S. Gordon, J. Comput. Chem. **28**, 276 (2007).
[39] K. A. Peterson and T. H. Dunning, J. Chem. Phys. **102**, 2032 (1995).
[40] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[41] D. E. Woon, J. Chem. Phys. **100**, 2838 (1994).
[42] T. H. Dunning, J. Chem. Phys. **157**, 479 (1989).
[43] R. A. Kendall, T. H. Dunning, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[44] T. Clark, J. Chandrasekhar, G. W. Spitznagel, and P. V. Schleyer, J. Comput. Chem. **4**, 294 (1983).
[45] M. J. Frisch, J. A. Pople, and J. S. Binkley, J. Chem. Phys. **80**, 3265 (1984).
[46] R. Krishnan, J. S. Binkley, R. Seeger, and J. A Pople, J. Chem. Phys. **72**, 650 (1980).
[47] M. O. Sinnokrot and C. D. Sherrill, J. Phys. Chem. A **108**, 10200 (2004).
[48] M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. Defrees, and J. A. Pople, J. Chem. Phys. **77**, 3654 (1982).
[49] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (Dover, Mineola, NY, 1996), pp. 68–70.