JCTC
Journal of Chemical Theory and Computation

Article

How Much Dispersion Energy Is Included in the Multiconfigurational Interaction Energy?

Michał Hapka, Agnieszka Krzemińska, and Katarzyna Pernal*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We demonstrate how to quantify the amount of dispersion interaction recovered by supermolecular calculations with the multiconfigurational self-consistent field (MCSCF) wave functions. For this purpose, we present a rigorous derivation which connects the portion of dispersion interaction captured by the assumed wave function model—the residual dispersion interaction—with the size of the active space. Based on the obtained expression for the residual dispersion contribution, we propose a dispersion correction for the MCSCF that avoids correlation double counting. Numerical demonstration for model four-electron dimers in both ground and excited states described with the complete active space self-consistent field (CASSCF) reference serves as a proof-of-concept for the method. Accurate results, largely independent of the size of the active space, are obtained. For many-electron systems, routine CASSCF interaction energy calculations recover a tiny fraction of the full second-order dispersion energy. We found that the residual dispersion is non-negligible only for purely dispersion-bound complexes.
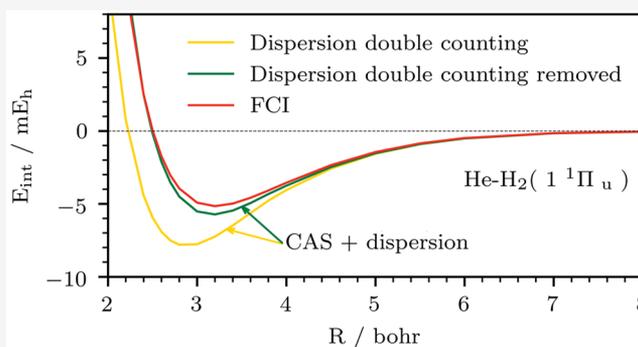
1. INTRODUCTION

Calculations of the intermolecular interaction energy typically involve the supermolecular approach. The perturbation theory is also a common choice and is often used to expose the nature of the interaction. In hybrid approaches, the supermolecular results are supplemented with energy contributions obtained in a perturbative manner, the most popular example being density functional theory (DFT) corrected for the dispersion energy.[1,2]

A given wave function theory (WFT) approximation is useful in describing interactions either if it accounts for high-order components in the intermolecular interaction operator or when it is reasonably accurate in the lowest orders and it is known which higher-order components are missing. A representative example from the first group are the coupled cluster (CC) methods[3,4] which predict interaction energies with high accuracy when effects of single, double, and triple excitations are accounted for. The Hartree−Fock (HF) method belongs to the second group. The HF energy contains electrostatic, exchange, and induction interaction energy components and has been rigorously shown to miss the second-order dispersion component.[5] In spite of this deficiency, the supermolecular HF result can be corrected for the dispersion energy which leads to a low-cost hybrid approach.[6−8] The accuracy of the hybrid approach can be significantly improved by including intramonomer correlation effects in a perturbative manner which is known as the HFDc[(1)] method.[8,9]

HF calculations assume a single-determinantal description of the interacting species. Although high-order single-reference CC methods may be applied to multireference systems,[10] the cost of going to full triple or quadruple excitations is too high in practice. The multireference CC extensions[11,12] are expensive, and their applications to intermolecular interactions remain scarce (e.g., see refs 11, 13, 14). Going beyond single-determinantal methods in the intermolecular interaction theory is a challenge. The multiconfigurational (MC) wave function methods are a good starting point and often the required one for open-shell systems, systems dominated by strong correlation, excited states, and transition-metal complexes. Although including multiple configurations guarantees proper representation of static correlation and partial recovery of the intramonomer correlation effects,[15,16] it is not a viable way to restore the remaining part of the dynamic correlation which in the case of intermolecular interactions manifests as dispersion. Notwithstanding significant progress in this field, none of the existing MC methods have yet achieved the levels of both accuracy and efficiency comparable to what is now

within reach of black-box interaction-energy calculations for single-reference systems. Therefore, the search for new and improved MC approaches remains active.

The most popular strategies to capture dispersion forces in the MC framework are the configuration interaction (CI) expansion and the perturbation theory. Each of these approaches involves certain limitations both in terms of computational cost and in terms of accuracy. A significant problem of CI is the lack of size consistency. Although it can be ameliorated with approximate corrections,[17,18] the residual size inconsistency typically still has to be addressed. This is done by obtaining the asymptotic interaction energy limit in a dimer calculation with the monomers separated at a large distance and shifting the results with respect to this value.[19] Note that in this scheme, the counterpoise correction (CP)[20] cannot be applied and the basis set superposition error (BSSE) is not removed. Additionally, because of the lack of triple excitations, multireference CI may underestimate the dispersion energy in van der Waals complexes by as much as 30%.[21,22]

A common choice to account for dynamic correlation via the multireference perturbation theory are the complete active space (CAS) perturbation theory (CASPT2)[23,24] and n-electron valence state perturbation theory (NEVPT2)[25–27] methods. Truncated at the second order, both require access to four-electron density matrices, which quickly becomes a computational bottleneck with respect to the size of the active space. Moreover, CASPT2 and multireference formulations of the Møller–Plesset theory[28,29] are plagued with intruder states[30] which have to be removed to guarantee a smooth potential energy surface. Various techniques that eliminate intruders exist, yet one should be cautious when choosing the optimal value of the shift parameter.[31] The intruder-state problem is by design avoided in the NEVPT2 method. The accuracy of single-reference MP2 calculations is often insufficient for quantitative predictions. A major weakness is that MP2 restores the dispersion energy at the uncoupled level of theory[32–34] which leads, for example, to the well-known overestimation of interaction energy in π–π complexes[35,36] or underestimation in the case of alkali-metal dimers.[37] This implies that the accuracy of the multireference second-order perturbation theory is seriously limited.

Several variants of the multiconfiguration DFT (MC DFT)[38,39] capable of treating dispersion forces have also been developed. In MC DFT, the connection between the DFT and wave function descriptions of the system is realized by partitioning of the Coulomb interaction operator into the short- and long-range components. The short-range (sr) electron–electron correlation effects are recovered by a density functional approximation, whereas the multiconfiguration wave function model acts in the long range. The usual strategy is to restrict the wave function to a few determinants, which is sufficient to recover the static correlation at a relatively low cost but not enough to capture the long-range dynamic correlation. This translates into the absence of the dispersion contribution in the interaction energy. As a remedy, Fromager and coauthors[40] proposed a long-range adaptation of the NEVPT2 approach. Recently, Hapka et al.[41] introduced a less expensive alternative based on the multireference formulation of the adiabatic connection formalism[42] which relies only on one- and two-electron reduced density matrices of the system. The cost-efficient route in which the complete active space self-consistent field (CASSCF) srPBE functional[43,44] is simply

corrected with the semi-classical D3 dispersion correction of Grimme[45] has been advocated by Stein and Reiher.[46]

The direct addition of the dispersion energy to the supermolecular CASSCF results, without correcting for the double counting, was attempted several times in the past. In the study of the Hg⋯Hg dimer, Kunz et al.[15] used this approach to partially account for the dynamic correlation within the monomers—an effect that cannot be recovered in the simple "HF plus dispersion" model. Rajchel et al.[47] verified the performance of the "CAS plus dispersion" method against the RCCSD(T) results for the potential energy curves of the Sc⋯Cr complex in the $^8\Sigma^+$, $^8\Pi$, and $^8\Delta$ states. The authors found that "CAS plus dispersion" consequently underestimated the magnitude of the interaction and predicted the location of the minima at ca. 0.2−0.3 bohr longer distances compared to the CC reference. The CAS-based approach was also applied to describe the unusually strong binding in the He⋯BeO($^1\Sigma^+$) dimer and showed excellent agreement with the MRCISD + Q potential.[48] In all of the mentioned examples, it was argued that the compact representation of the active space selected in CASSCF calculations is too small to recover the dispersion energy. Nevertheless, to the best of our knowledge, no systematic study of this claim has been presented to date and expressions for the residual dispersion energy recovered in the supermolecular CASSCF have not been derived.

In this work, we present how to rigorously extract the dispersion term in the interaction energy calculated with MC wave functions so that the double counting of the intermonomer correlation is avoided. As will be discussed, any wave function dissociating to MC wave functions for both monomers leads to recovering a small portion of the dispersion energy, denoted as $E_{\text{DISP}}^{\text{MC}}$. Being able to explicitly calculate the $E_{\text{DISP}}^{\text{MC}}$ contribution, we propose the "CAS plus dispersion" approach which strictly avoids the double counting of the dispersion contribution and recovers the total interaction energy as

$$E_{\text{int}} = E_{\text{int}}^{\text{MC}} + (E_{\text{DISP}}^{(2)} - E_{\text{DISP}}^{\text{MC}}) \tag{1}$$

where $E_{\text{int}}^{\text{MC}}$ is the supermolecular interaction energy obtained within the assumed WFT and $E_{\text{DISP}}^{(2)}$ is the full second-order dispersion energy. We present the results for wave functions of the CAS type and use the symmetry-adapted perturbation theory (SAPT)[49,50] to obtain the $E_{\text{DISP}}^{(2)}$ energy contributions.

It is worth noticing that Stein and Reiher[46] raise the concern of dispersion double counting in the context of large active spaces available with the density-matrix renormalization group wave functions. They propose a procedure that could approximately remove the residual dispersion inherently captured by an MC function. This approach requires orbital localization, followed by the removal of configurations (determinants) corresponding to excitations of a dispersion-like character.[51–53] In contrast, our protocol for $E_{\text{disp}}^{\text{MC}}$ evaluation avoids orbital localization and stems directly from the exact definition of the dispersion energy.

In the next section, we derive an exact expression for the dispersion energy recovered by supermolecular MC WFT calculation and discuss a method for its approximation. In Section 3, we present the details of the implementation and calculations. Section 4 provides numerical demonstration of the proposed "CAS plus dispersion" method on the example of model He⋯He and He⋯H$_2$ dimers as well as representative

noncovalently bound, many-electron complexes. In Section 5, we summarize our findings.

2. THEORY

2.1. Setting the Stage. For two weakly interacting electronic subsystems A and B forming a dimer AB, the interaction energy is defined as a difference of energies of the dimer E_{AB} and isolated monomers E_A and E_B

$$E_{int} = E_{AB} - E_A - E_B \tag{2}$$

By definition, the interaction energy vanishes when subsystems are infinitely separated

$$\lim_{R_{AB} \to \infty} E_{int} = 0 \tag{3}$$

where R_{AB} measures the distance between A and B. Let $\hat{H}_{AB}$, $\hat{H}_A$, and $\hat{H}_B$ denote the dimer Hamiltonian and the Hamiltonians of the isolated monomers, respectively. In the assumed WFT model, computation of the interaction energy in the supermolecular way requires finding wave functions for a dimer Ψ_{AB} and isolated monomers Ψ_A and Ψ_B, and making use of the pertinent Hamiltonians as follows

$$E_{int} = \langle \Psi_{AB} | \hat{H}_{AB} | \Psi_{AB} \rangle - \langle \Psi_A | \hat{H}_A | \Psi_A \rangle - \langle \Psi_B | \hat{H}_B | \Psi_B \rangle \tag{4}$$

A primary condition which must be satisfied by a physically relevant WFT is size consistency, that is, vanishing of the supermolecular interaction energy in the infinite-separation limit, which can be written as

$$\lim_{R_{AB} \to \infty} \langle \Psi_{AB} | \hat{H}_{AB} | \Psi_{AB} \rangle = \langle \Psi_A | \hat{H}_A | \Psi_A \rangle + \langle \Psi_B | \hat{H}_B | \Psi_B \rangle \tag{5}$$

which in turn imposes the separability condition for the (antisymmetric) wave function Ψ_{AB} reading

$$\lim_{R_{AB} \to \infty} \Psi_{AB} = \hat{A} \Psi_A \Psi_B \tag{6}$$

where the antisymmetrization operator $\hat{A}$ is applied on the product $\Psi_A \Psi_B$ of monomer functions.

In the Rayleigh–Schrödinger (RS) perturbation theory applied to molecular interactions, a dimer N_{AB}-electron Hamiltonian is written as a sum of N_A- and N_B-electron Hamiltonians of the monomers ($N_{AB} = N_A + N_B$) and a remainder, which is the interaction operator $\hat{V}_{int}$

$$\hat{H}_{AB} = \hat{H}_A(1, ..., N_A) + \hat{H}_B(N_A + 1, ..., N_{AB}) + \hat{V}_{int} \tag{7}$$

A sum of $\hat{H}_A + \hat{H}_B$ is taken as the zero-order Hamiltonian, $\hat{V}_{int}$ is taken as a perturbation, and the $\Psi_A \Psi_B$ product of monomer functions is the zero-order dimer eigenfunction. It is important to notice that the antisymmetrized wave function given in eq 6 is an eigenfunction of $\hat{H}_{AB}$ in the limit $R_{AB} \to \infty$, but it is not an eigenfunction of the zero-order Hamiltonian $\hat{H}_A + \hat{H}_B$. As is well known, the interaction energy, eq 4, follows from perturbation theory as a sum of the first-, second-, and higher-order terms in the perturbation[54]

$$E_{int} = E_{elst}^{(1)} + E_{ind}^{(2)} + E_{disp}^{(2)} + ... \tag{8}$$

where $E_{elst}^{(1)}$ denotes the electrostatic energy, while $E_{ind}^{(2)}$ and $E_{disp}^{(2)}$ denote the second-order induction and dispersion energy contributions, respectively. Each term is well defined and has a clear physical interpretation. In particular, the second-order dispersion energy is given by the expression

$$E_{disp}^{(2)} = -\sum_{i,j}{}' \frac{|\langle \Psi_A \Psi_B | \hat{V}_{int} | \Psi_{A,i} \Psi_{B,j} \rangle|^2}{E_{A,i} - E_A + E_{B,j} - E_B} \tag{9}$$

where i (j) refers to states $\Psi_{A,i}$ ($\Psi_{B,j}$) of the energy $E_{A,i}$ ($E_{B,j}$) corresponding to $\hat{H}_A$ ($\hat{H}_B$) and a prime indicates that summation indices run through all states with the exception of the states of interest Ψ_A and Ψ_B for which the dispersion is computed. The leading term of the multipole expansion of the dispersion energy of molecular systems vanishes as R_{AB}^{-6} and dominates in the asymptotic regime of the total interaction energy in van der Waals complexes.

The problem that we address in the next section is whether one can rigorously show to what extent a given wave function model accounts for the dispersion interaction in the supermolecular approach. The positive answer will be provided by singling out from the interaction energy expression terms of a dispersion character, which decay asymptotically in the same (R_{AB}^{-6}) fashion as the full dispersion energy, cf. eq 9, and which are equivalent to the latter if WFT corresponds to the full CI (FCI) description of a dimer.

2.2. Dispersion Interaction in the MC Wave Function Description of the Interaction Energy. In this section, we present the expression for the dispersion energy, E_{disp}^{MC}, which is included in the supermolecular interaction energy, obtained with the MC wave function description of a dimer. We begin by considering a general N-electron wave function which can be expressed as the antisymmetrized product of wave functions

$$\Psi = \hat{A} \prod_P \Psi_P \tag{10}$$

where $\hat{A}$ is the antisymmetrization operator (including a normalization constant). Ψ_P is an N_P-electron antisymmetric wave function, and $\sum_P N_P = N$. In addition, it is assumed that the whole set of spin orbitals is partitioned into disjoint subsets and a given function Ψ_P is expanded in a set of Slater determinants built from spin orbitals belonging to a subset P. The assumed ansatz for Ψ is therefore a group product function.[55] We refer to Ψ_P as a group function, where P denotes a group, indexes a group function, and refers to a pertinent subset of spin orbitals. Ψ_P can be a single determinant or a combination of many determinants, for example, an MC wave function. MC self-consistent field (MCSCF) wave functions, which are in the focus of this work, conform to the ansatz given in eq 10. They can be written as an antisymmetrized product of the inactive part Ψ_{inact} and the active part Ψ_{active}

$$\Psi^{MCSCF} = \hat{A} \Psi_{inact} \Psi_{active} \tag{11}$$

where Ψ_{inact} is a single Slater determinant built of one-electron groups, that is, spin orbitals φ_p

$$\Psi_{inact} = \hat{A} \prod_{p=1}^{N_{inact}} \varphi_p \tag{12}$$

while Ψ_{active} is a combination of Slater determinants.

If Ψ of the form as in eq 10 is found variationally by minimizing the energy $\langle \Psi | \hat{H} | \Psi \rangle$, where $\hat{H}$ is the electronic Hamiltonian, then, for each group wave function Ψ_P, one can form a group electronic Hamiltonian $\hat{H}_P$ reading

$$\hat{H}_P = \hat{H}_{P,intra} + \hat{H}_{P,SCF} \tag{13}$$

The intragroup Hamiltonian $\hat{H}_{P,\text{intra}}$ includes the kinetic energy operator, the interaction of N_P electrons, assigned to a group P, with all nuclei in the system (denoted as M_{nuc}), and the appropriate electron−electron interaction

$$\hat{H}_{P,\text{intra}} = -\frac{1}{2}\sum_{i=1}^{N_P}\nabla_{\mathbf{r}_i}^2 - \sum_{i=1}^{N_P}\sum_{a=1}^{M_{\text{nuc}}}\frac{Z_a}{|\mathbf{r}_i - \mathbf{R}_a|} + \sum_{i<j}^{N_P}\frac{1}{r_{ij}}$$

(14)

The other part of the group Hamiltonian given in eq 13 describes the intergroup mean field interaction of electrons assigned to a group P with electrons in the other groups. It arises as a result of self-consistent field (SCF) optimization and depends on the self-consistently obtained group spin orbitals. The mean field intergroup interaction operator, $\hat{H}_{P,\text{SCF}}$, is a sum of one-electron Coulomb and exchange operators pertaining to all groups except for P, namely

$$\hat{H}_{P,\text{SCF}} = \sum_{i=1}^{N_P}\sum_{Q \neq P}\int dx_1 \frac{\rho_Q(x_1)}{r_{1i}} - \sum_{i=1}^{N_P}\sum_{Q \neq P}\int dx_1 \frac{\gamma_Q(x_i, x_1)}{r_{1i}}\hat{P}_{1i}$$

(15)

where $r_{1i} = |\mathbf{r}_1 - \mathbf{r}_i|$, $x = (\mathbf{r}, \sigma)$ stands for a combined spatial-spin coordinate, and the $\hat{P}_{ij}$ operator interchanges x_i with x_j; $\rho_Q(x) = \langle\Psi_Q|\hat{\rho}(x)|\Psi_Q\rangle$ and $\gamma_Q(x,x') = \langle\Psi_Q|\hat{\gamma}(x,x')|\Psi_Q\rangle$ are, respectively, the density and one-electron reduced density matrix functions corresponding to Ψ_Q. A wave function Ψ_P is one of the eigenfunctions of $\hat{H}_P$ with the corresponding energy ε_P, that is, if we write the eigenproblem for the group Hamiltonian $\hat{H}_P$

$$\hat{H}_P\Psi_{P,i} = \varepsilon_{P,i}\Psi_{P,i}$$

(16)

then, $\exists i\Psi_P = \Psi_{P,i}$. Within the assumed wave function approximate model, $\Psi_{P,i}$ is an MC wave function given as a combination of M_P Slater determinants $\Phi_{P,K}$

$$\Psi_{P,i} = \sum_{K \in P}^{M_P} d_{P,iK}\Phi_{P,K}$$

(17)

and the Hamiltonian $\hat{H}_P$ is the diagonal in the M_P-dimensional space spanned by $\Phi_{P,K}$ functions.

It is straightforward to notice that the total electronic energy is given as a sum of group energies, $\varepsilon_P = \langle\Psi_P|\hat{H}_P|\Psi_P\rangle$, with the intergroup Coulomb-exchange interaction energy subtracted to avoid its double counting

$$E = \langle\Psi|\hat{H}|\Psi\rangle = \sum_P \varepsilon_P - \frac{1}{2}\sum_P\sum_{Q \neq P}E_{\text{Coul-Exch}}^{PQ}$$

(18)

$$E_{\text{Coul-Exch}}^{PQ} = \iint\frac{\rho_P(x_1)\rho_Q(x_2)}{r_{12}}dx_1 dx_2 - \iint\frac{\gamma_P(x_1, x_2)\gamma_Q(x_2, x_1)}{r_{12}}dx_1 dx_2$$

(19)

Now, consider a perturbed electronic Hamiltonian, linear in the perturbation parameter λ

$$\hat{H}(\lambda) = \hat{H}^{(0)} + \lambda\hat{H}^{(1)}$$

(20)

A corresponding perturbed group Hamiltonian, cf. eqs 13 and 15, will include a direct perturbing term (linear in λ) and

its SCF part will acquire a λ dependence because of the λ dependence of both the density and density matrices, namely

$$\hat{H}_P(\lambda) = \hat{H}_{P,\text{intra}} + \lambda\hat{H}_P^{(1)} + \hat{H}_{P,\text{SCF}}(\lambda)$$

(21)

where

$$\hat{H}_{P,\text{SCF}}(\lambda) = \sum_{i=1}^{N_P}\sum_{Q \neq P}\int dx_1 \frac{\rho_Q^\lambda(x_1)}{r_{1i}} - \sum_{i=1}^{N_P}\sum_{Q \neq P}\int dx_1 \frac{\gamma_Q^\lambda(x_i, x_1)}{r_{1i}}\hat{P}_{1i}$$

(22)

The total electronic energy, eq 18, expanded in terms of λ, is a sum of terms in different orders of the perturbation. Because our interest lies in the dispersion energy, which appears in the second order of the interaction operator, we only focus on the second-order term reading

$$E^{(2)} = \sum_P \varepsilon_P^{(2)} - \frac{1}{2}\sum_P\sum_{Q \neq P}(E_{\text{Coul-Exch}}^{PQ})^{(2)}$$

(23)

where using the second Hellmann−Feynman theorem,[56] the second-order group energy is given as

$$\varepsilon_P^{(2)} = \left\langle\Psi_P^{(0)}\left|\hat{H}_P^{(1)} + \frac{d\hat{H}_{P,\text{SCF}}(\lambda)}{d\lambda}\right|\Psi_P^{(1)}\right\rangle + \left\langle\Psi_P^{(0)}\left|\frac{d^2\hat{H}_{P,\text{SCF}}(\lambda)}{d\lambda^2}\right|\Psi_P^{(0)}\right\rangle$$

(24)

Notice that we have assumed hermiticity of the perturbing operators and that wave functions are real-valued.

Let us now turn to the molecular interaction theory, where the interaction operator $\hat{V}_{\text{int}}$, introduced in eq 7, is treated as a perturbation. The zero-order part is given as a sum of monomer Hamiltonians

$$\hat{H}^{(0)} = \hat{H}_A + \hat{H}_B$$

(25)

and the explicit form of the perturbation, $\hat{H}^{(1)} = \hat{V}_{\text{int}}$, reads

$$\hat{V}_{\text{int}} = -\sum_{i=1}^{N_A}\sum_{b=1}^{M_{\text{nuc,B}}}\frac{Z_b}{|\mathbf{r}_i - \mathbf{R}_b|} - \sum_{j=1}^{N_B}\sum_{a=1}^{M_{\text{nuc,A}}}\frac{Z_a}{|\mathbf{r}_j - \mathbf{R}_a|} + \sum_{i=1}^{N_A}\sum_{j=1}^{N_B}\frac{1}{r_{ij}}$$

(26)

where the subscripts a and b label the nuclei and i and j denote the electrons of monomers A and B, respectively. A monomer X is composed of N_X electrons and $M_{\text{nuc,X}}$ nuclei. Throughout the text, X will stand for either A or B.

When monomers are infinitely separated, the size-consistency condition given in eq 5 imposes that the total dimer wave function acquires the form of a product of monomer functions, as seen in eq 6. This can only be satisfied if each N_P-electron wave function Ψ_P dissociates into a product of monomer wave functions Ψ_{P_A} and Ψ_{P_B}, which are, respectively, N_{P_A}- and N_{P_B}-electron and $N_P = N_{P_A} + N_{P_B}$. Neglecting the electron exchange, the zero-order group function reads

$$\Psi_P^{(0)} = \Psi_{P_A}(x_1, ..., x_{N_{P_A}})\Psi_{P_B}(x_{N_{P_A}+1}, ..., x_{N_{P_A}+N_{P_B}})$$

(27)

In the $R_{AB} \rightarrow \infty$ limit, the group density and the density matrix corresponding to Ψ_P separate into monomer group functions, namely

$$\rho_P^{(0)}(x) = \rho_{P_A}(x) + \rho_{P_B}(x) \qquad (28)$$

$$\gamma_P^{(0)}(x, x') = \gamma_{P_A}(x, x') + \gamma_{P_B}(x, x') \qquad (29)$$

which is a consequence of vanishing overlap between Ψ_{P_A} and Ψ_{P_B} in the dissociation limit. Partitioning of the total electronic Hamiltonian assumed in eqs 25 and 26 results in the following decomposition of the group Hamiltonian $\hat{H}_P$, cf. eq 21

$$\hat{H}_P(\lambda) = \hat{H}_{P_A,\text{intra}} + \hat{H}_{P_B,\text{intra}} + \lambda \hat{V}_{P_{AB}} + \hat{H}_{P,\text{SCF}}(\lambda) \qquad (30)$$

where $\hat{H}_{P_A,\text{intra}}$ collects intragroup operators for a group P in monomer A

$$\hat{H}_{P_A,\text{intra}} = -\frac{1}{2}\sum_{i=1}^{N_{P_A}} \nabla_{\mathbf{r}_i}^2 - \sum_{i=1}^{N_{P_A}} \sum_{a=1}^{M_{\text{nuc,A}}} \frac{Z_a}{|\mathbf{r}_i - \mathbf{R}_a|} + \sum_{i<i'}^{N_{P_A}} \frac{1}{r_{ii'}} \qquad (31)$$

analogously for $\hat{H}_{P_B,\text{intra}}$, and the interaction operator for a group P

$$\hat{V}_{P_{AB}} = -\sum_{i=1}^{N_{P_A}} \sum_{b=1}^{M_{\text{nuc,B}}} \frac{Z_b}{|\mathbf{r}_i - \mathbf{R}_b|} - \sum_{j=1}^{N_{P_B}} \sum_{a=1}^{M_{\text{nuc,A}}} \frac{Z_a}{|\mathbf{r}_j - \mathbf{R}_a|} + \sum_{i=1}^{N_{P_A}} \sum_{j=1}^{N_{P_B}} \frac{1}{r_{ij}} \qquad (32)$$

includes the interaction of electrons in A assigned to a group P with nuclei in monomer B, the interaction of electrons in monomer B assigned to a group P with nuclei in monomer A, and the interaction among electrons in monomers A and B within a group P. At infinite intermonomer separation (the $\lambda = 0$ limit), the SCF part of the group Hamiltonian, eq 22, is partitioned as

$$\hat{H}_{P,\text{SCF}}^{(0)} = \hat{H}_{P_A,\text{SCF}} + \hat{H}_{P_B,\text{SCF}} \qquad (33)$$

$$\hat{H}_{P_X,\text{SCF}} = \sum_{i=1}^{N_{P_X}} \sum_{Q\neq P} \int dx_1 \frac{\rho_{Q_X}(x_1)}{r_{i1}} - \sum_{i=1}^{N_{P_X}} \sum_{Q\neq P} \int dx_1 \frac{\gamma_{Q_X}(x_i, x_1)}{r_{i1}} \hat{P}_{i1} \qquad (34)$$

where ρ_{Q_X} and γ_{Q_X} are the monomer group density and the density matrix, as seen in eqs 28 and 29, respectively. This allows one to write the group Hamiltonian, eq 30, as a sum of monomer group Hamiltonians

$$\hat{H}_P(0) = \hat{H}_P^{(0)} = \hat{H}_{P_A} + \hat{H}_{P_B} \qquad (35)$$

where

$$\hat{H}_{P_X} = \hat{H}_{P_X,\text{intra}} + \hat{H}_{P_X,\text{SCF}} \qquad (36)$$

Our goal is to identify the dispersion terms in the second-order energy given in eq 23 corresponding to the intermolecular perturbation operator given in eq 26. First, one should notice that the last term in eq 23, which includes second-order perturbations of the Coulomb-exchange inter-

group interactions, $(E_{\text{Coul-Exch}}^{PQ})^{(2)}$, cannot contribute to the dispersion energy. These terms describe the interactions of the perturbed one-electron density and density matrices. Next, let us consider the second-order group energy $\varepsilon_P^{(2)}$, presented in eq 24, where the group interaction operator $\hat{H}_P^{(1)} = \hat{V}_{P_{AB}}$ is given in eq 32. The derivatives of $\hat{H}_{P,\text{SCF}}(\lambda)$ with respect to λ, involved in $\varepsilon_P^{(2)}$, are given by the pertinent derivatives of densities and density matrices

$$\hat{H}_{P,\text{SCF}}^{(n)} = \sum_{i=1}^{N_P} \sum_{Q\neq P} \int dx_1 \frac{\rho_Q^{(n)}(x_1)}{r_{1i}} - \sum_{i=1}^{N_P} \sum_{Q\neq P} \int dx_1 \frac{\gamma_Q^{(n)}(x_i, x_1)}{r_{1i}} \hat{P}_{1i} \qquad (37)$$

Thus, they are one-electron operators and as such cannot contribute to the dispersion interaction either. Therefore, the only term in eq 23 which may give rise to the dispersion interaction is the one involving the operator $\hat{V}_{P_{AB}}$. Retaining solely this term in $E^{(2)}$ and denoting the resulting energy as $E_d^{(2)}$, we write

$$E_d^{(2)} = \sum_P \langle \Psi_P^{(0)} | \hat{V}_{P_{AB}} | \Psi_P^{(1)} \rangle \qquad (38)$$

To see under which condition dispersion terms emerge from $E_d^{(2)}$, we need to find the first-order correction for group wave functions, $\Psi_P^{(1)}$. For this purpose, consider zero-order group functions, $\Psi_{P,i}^{(0)}$, that are eigenfunctions of the $\hat{H}_P^{(0)}$ Hamiltonian, eq 35

$$\hat{H}_P^{(0)} \Psi_{P,i}^{(0)} = \varepsilon_{P,i}^{(0)} \Psi_{P,i}^{(0)} \qquad (39)$$

describing a group P in the noninteracting, $R_{AB} \rightarrow \infty$, limit. Let $\Psi_{P_X,i}$ denote monomer functions corresponding to a group monomer Hamiltonian $\hat{H}_{P_X}$ given in eq 36

$$\hat{H}_{P_X} \Psi_{P_X,i} = \varepsilon_{P_X,i} \Psi_{P_X,i} \qquad (40)$$

Then, any (N_P-electron) ith eigenfunction $\Psi_{P,i}^{(0)}$ is a product of (N_{P_A}- and N_{P_B}-electron) monomer-group functions

$$\Psi_{P,i}^{(0)} = \Psi_{P_A,k} \Psi_{P_B,l} \qquad (41)$$

with the corresponding energy

$$\varepsilon_{P,i}^{(0)} = \varepsilon_{P_A,k} + \varepsilon_{P_B,l} \qquad (42)$$

The index i, labeling $\varepsilon_P^{(0)}$ and $\Psi_P^{(0)}$, is understood as a combined index (kl), where k and l label group states of A and B, respectively.

In the assumed multiconfiguration approximation, a monomer wave function can be written as

$$\Psi_{P_X,k} = \sum_{K\in P_X} d_{P_X,kK} \Phi_{P_X,K} \qquad (43)$$

where $\{\Phi_{P_X,K}\}$ is a set of single determinants built from spin orbitals belonging to a group P_X. In the multireference approximations considered in this work, orbital and configurational spaces are finite, and the wave functions are obtained with the optimal orbitals and configuration expansion coefficients. Having defined a group zero-order Hamiltonian, eq 35, with the corresponding eigenfunctions shown in eq 41, applying the standard perturbation theory to eq 16 using the

first-order Hamiltonian $\hat{V}_{P_{AB}} + \hat{H}^{(1)}_{P,SCF}$ as the perturbation (see eqs 30, 32 and 37) leads to obtaining first-order perturbation to an ith state of the group P, $\Psi^{(1)}_{P,i}$, which can be written in a general form as

$$\Psi^{(1)}_{P,i} = \sum_{j\neq i} C_{P,ij}\Psi^{(0)}_{P,j} + \Psi_{P_{B},l}\sum_g c_{P_{A},ig}\Phi^{(1)}_{P_{A},g} + \Psi_{P_{A},k}\sum_g c_{P_{B},ig}\Phi^{(1)}_{P_{B},g} \tag{44}$$

The first term results from admitting state mixing and the sum runs over all unperturbed states of the Hamiltonian $\hat{H}^{(0)}_P$ orthogonal to $\Psi^{(0)}_{P,i}$. The second and third terms in eq 44 are due to orbital optimization, that is, allowing for mixing of orbitals between groups. If, for example, $\varphi_p(x)$ is an unperturbed spin orbital belonging to a group P and assigned to monomer A, $p \in P_A$, then, its first-order perturbation $\varphi^{(1)}_p$ is expanded in the space of all spin orbitals except those belonging to P_A. Such first-order perturbations contribute to $\Psi^{(1)}_{P,i}$ as combinations of singly excited determinants $\Phi^{(1)}_{P_A,g}$ and $\Phi^{(1)}_{P_B,g}$. In other words, $\Phi^{(1)}_{P_A,g}$ denotes a determinant built of $N_{P_A} - 1$ orbitals from a group P_A and one spin orbital from another group $Q \neq P_A$ (the same for monomer B). Assume from now on that the considered dimer is in its ground state, thus the indices i, l, and k in eq 44 will be set to 0. Notice, however, that the presented analysis is general, not limited to ground states. The explicit expression for $C_{P,0j}$ coefficients in eq 44 reads

$$\forall_{j\neq 0} \qquad C_{P,0j} = \frac{\langle\Psi^{(0)}_{P,j}|\hat{V}_{P_{AB}} + \hat{H}^{(1)}_{P,SCF}|\Psi^{(0)}_{P,0}\rangle}{\varepsilon^{(0)}_{P,0} - \varepsilon^{(0)}_{P,j}} \tag{45}$$

Coefficients $\{c_{P_X,0g}\}$ arising from orbital perturbations cannot be given explicitly. They follow as solutions of linear equations which are generalizations of the coupled-perturbed HF equations.[57,58] As in the case of the latter equations, $c_{P_X,0g}$ coefficients are linearly dependent on the elements $\langle\Phi^{(1)}_{P_A,f}\Psi_{P_B,0}|\hat{V}_{P_{AB}} + \hat{H}^{(1)}_{P,SCF}|\Psi_{P_A,0}\Psi_{P_B,0}\rangle$, which allows us to write

$$c_{P_A,0g} = \sum_f \langle\Phi^{(1)}_{P_A,f}\Psi_{P_B,0}|\hat{V}_{P_{AB}} + \hat{H}^{(1)}_{P,SCF}|\Psi_{P_A,0}\Psi_{P_B,0}\rangle F_{P_A,gf} \tag{46}$$

similar for $c_{P_B,0g}$, where the functions $F_{P_X,gf}$ depend on the energies $\varepsilon_{P_X,0}$ and the matrix elements of $\hat{H}^{(0)}_{P_X}$

$$F_{P_X,gg'} = F_{P_X,gg'}(\varepsilon_{P_X,0}, \{\langle\Phi^{(1)}_{P_X,f}|\hat{H}^{(0)}_{P_X}|\Phi^{(1)}_{P_X,f'}\rangle\}) \tag{47}$$

Let us return to the energy expression in eq 38 obtained after eliminating these terms from the second-order energy of the dimer, which cannot give rise to the dispersion energy. The $C_{P,0j}$ coefficients (eq 45) which enter in the first term of eq 44 depend on the interaction operator $\hat{V}_{P_{AB}}$ and a one-electron operator $\hat{H}^{(1)}_{P,SCF}$. Only the former is retained because, as we discussed, the latter does not contribute to the dispersion interaction. The second and third terms in eq 44 (expression for $\Psi^{(1)}_{P,i}$), which arise from the orbital perturbation in SCF methods, cannot recover the dispersion interaction either. Their insertion into eq 38 leads to products of terms of the form $\langle\Psi_{P_A,0}\Psi_{P_B,0}|\hat{V}_{P_{AB}}|\Phi^{(1)}_{P_A,g}\Psi_{P_B,0}\rangle$, in which the interaction

operator $\hat{V}_{P_{AB}}$ does not connect the ground state with both-monomer-excited states. Such terms, corresponding to one monomer remaining in the ground state, contribute to the induction components of the interaction energy, not to dispersion.

Thus, out of all terms in $E^{(2)}_d$, we retain only the ones involving $C_{P,0j}$ and the operator $\hat{V}_{P_{AB}}$, cf. eq 45. After expressing both the zero-order energy and wave function in terms of its monomer counterparts, eqs 41 and 42, we arrive at the energy expression which is a part of the dimer energy and gives rise to the dispersion interaction. It reads

$$E^{MC}_{disp} = -\sum_P \sum_{ij\neq 0} \frac{\langle\Psi_{P_A,0}\Psi_{P_B,0}|\hat{V}_{P_{AB}}|\Psi_{P_A,i}\Psi_{P_B,j}\rangle^2}{\omega_{P_A,i} + \omega_{P_B,j}} \tag{48}$$

where $\omega_{P_X,i}$ denotes a transition energy for a group P and a monomer X

$$\omega_{P_X,i} = \varepsilon_{P_X,i} - \varepsilon_{P_X,0} \tag{49}$$

and only the terms for which $i \neq 0$ and $j \neq 0$ have been kept (naturally, the terms for which $i = 0$ and $j \neq 0$ and vice versa will be of the induction type; therefore, they have been excluded from eq 48).

A more convenient form of E^{MC}_{disp} follows after introducing transition density matrices. Notice that in the assumed wave function, eq 10, a set of occupied spin orbitals is divided into disjoint subsets P corresponding to group functions Ψ_P. In the dimer dissociation limit, each subset decomposes into sets P_A and P_B of the orbitals assigned to the pertinent monomers. A transition density matrix for a monomer X in the representation of the orbitals from P_X reads

$$\forall_{pq\in P_X} \gamma^{X,i}_{pq} \equiv \langle\Psi_{X,0}|\hat{a}^\dagger_q\hat{a}_p|\Psi_{X,i}\rangle = \langle\Psi_{P_X,0}|\hat{a}^\dagger_q\hat{a}_p|\Psi_{P_X,i}\rangle \tag{50}$$

which allows one to turn the expression in eq 48 into

$$E^{MC}_{disp} = -\sum_P \sum_{ij\neq 0} \frac{\left(\sum_{pq\in P_A}\sum_{ab\in P_B}\gamma^{A,i}_{pq}\gamma^{B,j}_{ab}\langle pa|qb\rangle\right)^2}{\omega_{P_A,i} + \omega_{P_B,j}} \tag{51}$$

where $\langle pa|qb\rangle$ are two-electron integrals in the 1212 convention. This is a central equation of this work.

E^{MC}_{disp} yields the amount of dispersion interaction energy recovered by the employed MC wave function model conforming to the ansatz in eq 10. Let us list the most relevant conclusions that the obtained result leads to

(a) A given group P has nonzero contributions to the interaction energy of the dispersion type only if in the $R_{AB} \to \infty$ limit, the wave function Ψ_P dissociates into a product $\Psi_{P_A}\Psi_{P_B}$, where both Ψ_{P_A} and Ψ_{P_B} are MC functions. The reason is that only the states excited in the configurational space of the P_A and P_B groups enter E^{MC}_{disp}, which obviously implies that configurational spaces, cf. eq 43, must be at least two-dimensional.

(b) The previous point immediately implies that $E^{MC}_{disp} = 0$ and no dispersion is recovered in supermolecular calculations in the case of the HF or CAS(2,2) description of a dimer or if upon dissociation, one monomer is described with a single determinant.

(c) In a given wave function model, the inactive orbitals (doubly occupied in each Slater determinant) do not play a role in recovering the dispersion. A group of

inactive orbitals forms a single determinant Ψ_{inact}, as seen in eq 11, which does not contribute to $E_{\mathrm{disp}}^{\mathrm{MC}}$.

(d) Orbital optimization should not play a significant role, so CAS-CI and CASSCF at the same CAS(n,m) level are expected to reproduce the same amount of dispersion.

2.3. Dispersion Energy Recovered in the CAS Model.

In this work, we focus our attention on the CAS model. A dimer wave function takes a product form, as shown in eq 11, where the active function Ψ_{active} is an n-electron full configuration function obtained from m orbitals called active. Such a function is denoted as CAS(n,m). The dispersion energy recovered by a dimer CAS function dissociating into CAS functions for monomers can be written as

$$E_{\mathrm{disp}}^{\mathrm{CAS}} = -\sum_{ij \neq 0} \frac{\left(\sum_{pq}^{\mathrm{active_A}} \sum_{ab}^{\mathrm{active_B}} \gamma_{pq}^{\mathrm{A},i} \gamma_{ab}^{\mathrm{B},j} \langle pa|qb \rangle \right)^2}{\omega_{\mathrm{active_A},i} + \omega_{\mathrm{active_B},j}} \tag{52}$$

pq (ab) are active orbitals of monomer A (B). Notice that in fact, eq 52 is more general and applies to any MCSCF model. By $\omega_{\mathrm{active_X},i}$, we denote transition energies obtained by diagonalizing a Hamiltonian for the active group, cf. eqs 31, 34, and 36,

$$\hat{H}_{\mathrm{act_X}} = -\frac{1}{2} \sum_{i=1}^{N_{\mathrm{active,X}}} \nabla_{\mathbf{r}_i}^2 - \sum_{i=1}^{N_{\mathrm{active,X}}} \sum_{a=1}^{M_{\mathrm{nuc,X}}} \frac{Z_a}{|\mathbf{r}_i - \mathbf{R}_a|} + \sum_{i<j}^{N_{\mathrm{active,X}}} \frac{1}{r_{ij}}$$
$$+ \sum_{i=1}^{N_{\mathrm{active,X}}} \int \mathrm{d}x_1 \frac{\rho_{\mathrm{inact,X}}(x_1)}{r_{i1}} - \sum_{i=1}^{N_{\mathrm{active,X}}} \int \mathrm{d}x_1 \frac{\gamma_{\mathrm{inact,X}}(x_i, x_1)}{r_{i1}} \hat{P}_{i1} \tag{53}$$

where $N_{\mathrm{active,X}}$ is the number of active electrons in monomer X, $M_{\mathrm{nuc,X}}$ denotes the number of nuclei in X, and $\rho_{\mathrm{inact,X}}$ and $\gamma_{\mathrm{inact,X}}$ are the density and the density matrix corresponding to inactive orbitals, respectively.

We have shown that the supermolecular CAS interaction energy recovers a portion of the second-order dispersion energy, quantified by eq 52, which directly relates to the size of the active space. In the extreme case, when all orbitals are active and the CAS turns into the FCI, the $E_{\mathrm{disp}}^{\mathrm{CAS}}$ term becomes identical to the full second-order dispersion energy shown in eq 9. In other cases, the CAS supermolecular interaction energy can be corrected for the missing portion of the dispersion interaction by subtracting $E_{\mathrm{disp}}^{\mathrm{CAS}}$ and adding the full dispersion energy recovered by a method of choice, which leads to the "CAS corrected for the dispersion" interaction energy method

$$E_{\mathrm{int}}^{\mathrm{CAS+dispCAS}} = E_{\mathrm{int}}^{\mathrm{CASSCF}} - E_{\mathrm{disp}}^{\mathrm{CAS}} + E_{\mathrm{disp}}^{(2)} \tag{54}$$

When defining monomer CAS wave functions, which are dissociation products of the assumed CAS(n,m) dimer function, one should be cautious. It is clear that the number of active electrons in the monomers, n_A and n_B, should sum up to n to satisfy the size-consistency condition. However, if $m = m_A + m_B$ active orbitals are considered, m_A of which are localized on A and m_B of which are localized on B, then in the limit $R_{AB} \to \infty$, a product of CAS(n_A,m) and CAS(n_B,m) functions yields the same energy as a sum of monomer energies computed with the pertinent CAS(n_X,m_X) functions. Thus, the aforementioned product is a valid CAS zero-order wave function of a dimer, satisfying the size-consistency condition. In practical terms, this implies that in the

supermolecular computation of the CAS interaction energy, the energy of monomers should be obtained with CAS(n_X,m) functions, where m includes m_X active orbitals on X and $m - m_X$ orbitals on the other monomer. This problem is similar to, but not the same as, the counterpoise correction for the BSSE. We call it a counterpoise CAS-superposition error (CASSE) correction. It vanishes when the CAS function turns into the FCI and a complete basis set is used.

It should be noticed that the dispersion correction, $E_{\mathrm{disp}}^{\mathrm{CAS}}$, discussed in this section follows from the standard RS perturbation theory, which misses exchange contributions. If perturbation theory with symmetry forcing[59,60] were applied, it would reveal that the CAS interaction energy includes only a fraction of the second-order exchange-dispersion term, whose magnitude depends on the size of the active space. One would therefore arrive at the conclusion that the CAS must be corrected for the exchange-dispersion energy in a way similar to that for the dispersion energy, that is, in analogy to eq 54. In the next section, an approximate exchange-dispersion correction is proposed.

3. COMPUTATIONAL DETAILS

The second-order dispersion energies calculated in this work include both the polarization and exchange components. We use the $E_{\mathrm{DISP}}^{(2)}$ notation for the energy terms calculated directly within the SAPT[49,50] framework

$$E_{\mathrm{DISP}}^{(2)} = E_{\mathrm{disp}}^{(2)} + E_{\mathrm{exch-disp}}^{(2)} \tag{55}$$

Unlike $E_{\mathrm{disp}}^{\mathrm{CAS}}$, the second-order exchange-dispersion energy recovered in supermolecular CASSCF calculations, $E_{\mathrm{exch-disp}}^{\mathrm{CAS}}$, is not obtained directly but approximated by scaling of the full term

$$E_{\mathrm{exch-disp}}^{\mathrm{CAS}} = E_{\mathrm{exch-disp}}^{(2)} \times \frac{E_{\mathrm{disp}}^{\mathrm{CAS}}}{E_{\mathrm{disp}}^{(2)}} \tag{56}$$

In analogy to eq 55, we use the

$$E_{\mathrm{DISP}}^{\mathrm{CAS}} = E_{\mathrm{disp}}^{\mathrm{CAS}} + E_{\mathrm{exch-disp}}^{\mathrm{CAS}} \tag{57}$$

notation to represent the complete second-order dispersion energy recovered by the CAS (the $E_{\mathrm{disp}}^{\mathrm{CAS}}$ term is given in eq 52). In other words, $E_{\mathrm{DISP}}^{\mathrm{CAS}}$ denotes the dispersion energy which is accounted for in the supermolecular interaction energy, eq 2.

The "CAS corrected for dispersion" energy takes the form

$$E_{\mathrm{int}}^{\mathrm{CAS+DispCAS}} = E_{\mathrm{int}}^{\mathrm{CASSCF}} - E_{\mathrm{DISP}}^{\mathrm{CAS}} + E_{\mathrm{DISP}}^{(2)} \tag{58}$$

For comparison, we also present the results obtained by taking the sum of the CASSCF interaction energy and the $E_{\mathrm{DISP}}^{(2)}$ term (the dispersion energy included in the CASSCF interaction energy is not removed), the approach which we denote as $E_{\mathrm{int}}^{\mathrm{CAS+DISP}}$. An analogous notation, $E_{\mathrm{int}}^{\mathrm{HF+DISP}}$, is used for the HF method supplemented with full dispersion.

Evaluation of the $E_{\mathrm{disp}}^{\mathrm{CAS}}$ formula, eq 52, requires access to both transition energies and transition density matrices of the monomers corresponding to the active group Hamiltonian, eq 53. These may be obtained either by direct diagonalization of the Hamiltonian or from full linear response equations. In this work, we followed the latter approach for the model He···He and He···H$_2$ dimers using time-dependent linear response equations of ref 61. For these systems, therefore, $E_{\mathrm{disp}}^{\mathrm{CAS}}$ is computed exactly, within a given basis set. For many-electron systems, we approximate the required response properties

using solutions of extended random phase approximation[62,63] (ERPA) equations. The ERPA eigenproblem for each monomer is formulated for the $\hat{H}_{\mathrm{act_x}}$ Hamiltonian, as seen in, for example, eqs 6–9 in ref 16, which leads to a Hessian matrix of a simple form[64] (for details, see the Supporting Information). Both construction and diagonalization of the Hessian for a CAS wave function require steps which scale as M_{act}^6, where M_{act} denotes the number of the active orbitals. Thus, the cost of computing $E_{\mathrm{disp}}^{\mathrm{CAS}}$ in the ERPA approximation is small considering that regular CAS calculations employ an active space that is considerably smaller compared to the virtual space.

The CASSE introduced in the previous section was accounted for in $E_{\mathrm{disp}}^{\mathrm{CAS}}$ energy calculations according to a three-step protocol. First, monomer $\mathrm{CAS}(n_A,m_A)\mathrm{SCF}$ and $\mathrm{CAS}(n_B,m_B)\mathrm{SCF}$ calculations were converged in the dimer-centered basis set. Next, a set containing inactive and active natural orbitals of a given monomer was extended with active orbitals from its interacting partner and orthonormalized according to the Schmidt procedure. The new sets containing $m = m_A + m_B$ active orbitals for each monomer were completed by adding CASSCF secondary orbitals from monomer calculations. Finally, these orbital sets entered CAS-CI calculations.

The CAS wave function captures dynamic correlation effects only within the active space. This affects the quality of both the density and density matrices[16,65] and translates into a rather poor representation of the intramonomer correlation effects in the supermolecular CAS energy. To correct this deficiency, one may describe the first-order interaction energy terms, that is, the electrostatic and exchange energies, at the fully correlated level of theory instead of the CASSCF one. This may be realized in an approximate manner by subtracting the first-order terms calculated with CASSCF density matrices and replacing them with more accurate ones. When applied together with the discussed dispersion correction, this leads to

$$E_{\mathrm{int}}^{\mathrm{CASc+DispCAS}} = E_{\mathrm{int}}^{\mathrm{CAS+DispCAS}} + [E_{\mathrm{elst}}^{(1)} - E_{\mathrm{elst}}^{(1)}(\mathrm{CAS})]$$
$$+ [E_{\mathrm{exch}}^{(1)} - E_{\mathrm{exch}}^{(1)}(\mathrm{CAS})] \qquad (59)$$

where the first term is given by eq 58. The first and second terms in square brackets represent corrections of electrostatic and exchange interaction components, respectively, where the $E_{\mathrm{elst}}^{(1)}$ and $E_{\mathrm{exch}}^{(1)}$ terms account for intramonomer correlation effects. The $E_{\mathrm{elst}}^{(1)}(\mathrm{CAS})$ term is computed according to the known expression for electrostatic electron interaction, as seen in, for example, ref 49, with the electron density obtained from the pertinent CAS wave function. It is assumed that first-order exchange terms are calculated in the S^2 approximation (e.g., see eq 9 in ref 66) which requires access only to one- and two-electron reduced density matrices of the monomers. The CASc + DispCAS approach can be interpreted as a multireference counterpart of the HFDc$^{(1)}$ method.[67]

For the sake of convenience, acronyms and notations used in figures and tables throughout the article have been listed in Table 1.

Two model systems studied in this work are the He···He dimer and the He···H$_2$ dimer. The reference total interaction energy curve for He···He was taken from the work of Przybytek et al.[68] The benchmark interaction energies for excited states of the He···H$_2$ dimer were obtained at the FCI level of theory in the aug-cc-pVTZ basis set.[69] For many-electron systems from the A24 data set, the coupled-cluster

Table 1. Notations Used for Dispersion Energy and Interaction Energy Expressions in This Work

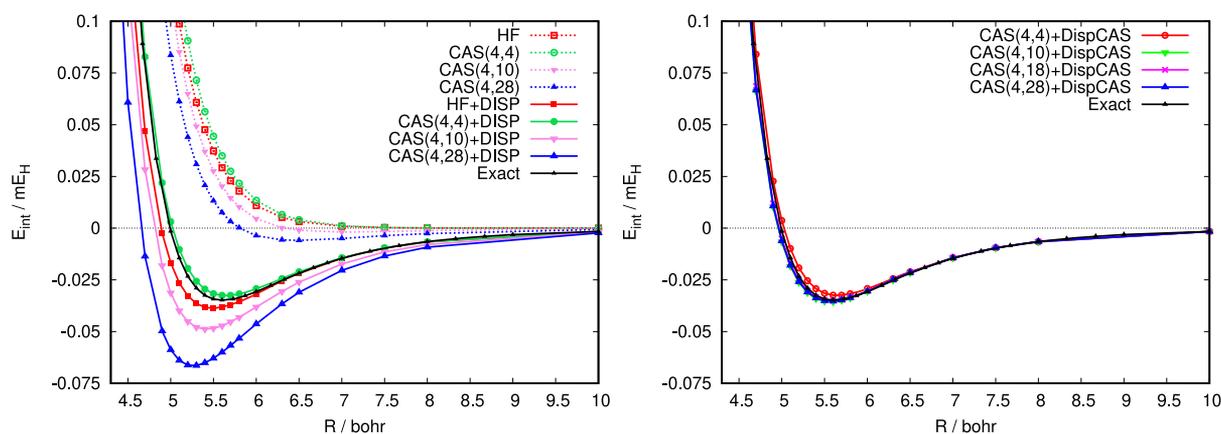|	dispersion energy
$E_{\mathrm{disp}}^{(2)}$	second-order dispersion energy, eq 9
$E_{\mathrm{exch\text{-}disp}}^{(2)}$	second-order exchange-dispersion energy
$E_{\mathrm{DISP}}^{(2)}$	sum of second-order dispersion and exchange-dispersion energies, eq 55
$E_{\mathrm{disp}}^{\mathrm{CAS}}$	residual CAS dispersion energy, eq 52
$E_{\mathrm{exch\text{-}disp}}^{\mathrm{CAS}}$	residual CAS exchange-dispersion energy, eq 56
$E_{\mathrm{DISP}}^{\mathrm{CAS}}$	sum of residual CAS dispersion and exchange-dispersion energies, eq 57^a
	interaction energy
$E_{\mathrm{int}}^{\mathrm{CASSCF}}$	CAS: CASSCF supermolecular interaction energy
$E_{\mathrm{int}}^{\mathrm{CAS+DISP}}$	CAS + DISP: CASSCF supermolecular interaction energy with the $E_{\mathrm{DISP}}^{(2)}$ dispersion added (dispersion double counting present)
$E_{\mathrm{int}}^{\mathrm{CAS+DispCAS}}$	CAS + DispCAS: CASSCF supermolecular interaction corrected for the dispersion (dispersion double counting removed), eq 58
$E_{\mathrm{int}}^{\mathrm{CASc+DispCAS}}$	CASc + DispCAS: CASSCF supermolecular interaction corrected for both dispersion and first-order (electrostatic and exchange) energy, eq 59

$^a E_{\mathrm{DISP}}^{\mathrm{CAS}}$ terms were corrected for CASSE in He···He and He···H$_2$ calculations (see the article).

single and double excitation and the perturbative triples/complete basis set [CCSD(T)/CBS] values from ref 70 served as a benchmark.

The $E_{\mathrm{disp}}^{\mathrm{CAS}}$ energy calculations were performed in the in-house code. All necessary integrals and CASSCF one- and two-electron reduced density matrices were obtained from the Molpro[71] program. The same software was used for super-molecular CASSCF calculations. The convergence threshold for SCF orbital optimization was set to 10^{-10} in the gradient norm. All reported interaction energies were corrected for the BSSE using the CP scheme[20] of Boys and Bernardi. In contrast to $E_{\mathrm{disp}}^{\mathrm{CAS}}$, the supermolecular CASSCF results were not corrected for the CASSE.

CASSCF wave functions for systems from the A24 data set were obtained with MP2 natural orbitals employed as an initial guess for the CAS optimization. The supermolecular CAS interaction energy may suffer from the lack of size consistency if dimer active orbitals, obtained for the assumed $\mathrm{CAS}(n,m)$ model and yielding the lowest energy of a dimer in the dissociation limit, do not correspond to the optimal active orbitals of the monomers [it is assumed that monomers are described with $\mathrm{CAS}(n_X,m_X)$ functions such that $n_A + n_B = n$ and $m_A + m_B = m$]. To eliminate the risk of size-consistency violation, for each dimer, the CASSCF energy was computed at an intermonomer distance of 200 bohr and it was checked that the dimer energy agrees with the sum of monomer energies within at least 10^{-4} mHa.

For both the He···He and He···H$_2$ dimers, the reference values of SAPT energy components were calculated at the FCI level of theory in a house-developed code.[65,72] In the case of He···He, both dispersion and exchange-dispersion components were extrapolated to the CBS limit according to the aug-cc-pVQZ → aug-cc-pV5Z two-point scheme of Bak et al.[73] For He···H$_2$, all calculations employed the aug-cc-pVTZ basis and CBS extrapolation was not performed. For systems from the A24 data set, the SAPT energy components entering eqs 58 and 59 were obtained at the DFT-SAPT[74−76] level of theory based on the asymptotically corrected[77] PBE0[78,79] functional using the dimer-centered basis for monomer calculations. The

Figure 1. He···He ground-state interaction energy curves. Left: $E_{int}^{CAS+DISP} = E_{int}^{CAS(n,m)} + E_{DISP}^{(2)}$; right: $E_{int}^{CAS+DispCAS}$ (see eq 58). "Exact" refers to the benchmark nonrelativistic potential of ref 68. The basis set is aug-cc-pV5Z.

exchange energy components were calculated in the S^2 approximation. All DFT-SAPT calculations were performed in the Molpro[71] program. The aug-cc-pVTZ basis set supplemented with (3s3p2d2f) midbond functions was used. Both midbond exponents and midbond location procedures were taken from ref 67.

4. RESULTS

4.1. Model Systems: He···He and He···H₂. In this section, we examine the performance of the introduced "CAS plus dispersion" models for the interaction between two-electron monomers. As mentioned, in this case, we use accurate CAS transition density matrices which enter the E_{disp}^{CAS} expression. In calculations of the latter, the CASSE is removed according to the protocol described in the previous section. The second-order dispersion energy is obtained from SAPT calculations at the FCI level of theory.

As we have rigorously shown in Section 2.2, CASSCF recovers a portion of the dispersion energy, as seen in eq 52. Therefore, a simple addition of the dispersion component, cf. eq 55, to the supermolecular CASSCF interaction energy leads to double counting of intermonomer correlation effects, the magnitude of which depends on the size of the active space. This is illustrated in Figure 1 (see "CAS + DISP" curves) on the example of the He···He dimer in the ground state. While a small CAS(4,4) wave function of the dimer [dissociating to a product of CAS(2,2) monomer functions] supplemented with dispersion deviates from the benchmark only by 6.6% in the van der Waals minimum (R_{eq} = 5.6 bohr), going to the CAS(4,28) model leads to 72% overestimation of the interaction energy. Additionally, dispersion double counting shifts the minimum from the correct distance of 5.6 bohr for CAS(4,4) to 5.3 bohr for CAS(4,28). In the proposed CAS + DispCAS approach, we improve the CAS + DISP results by removing the partial dispersion contained in the supermolecular CASSCF energy, cf. eq 58. Consequently, the extension of the active space does not lead to overcorrelation and the errors in the He···He minimum region do not exceed 2% regardless of the active space size (Figure 1). The only exception is the CAS(4,4) + DispCAS curve where the error reaches 7%, which is mainly due to the poor representation of the first-order energy terms (see Figure S1 in the Supporting Information).

In the asymptotic regime the He···He interaction is dominated by the dispersion energy which decays with the

inverse sixth power of the interatomic distance. Therefore, at large distances, the CASSCF interaction energy and the E_{disp}^{CAS} term calculated directly according to eq 52 should become identical. Indeed, in Figure 2 we demonstrate a perfect
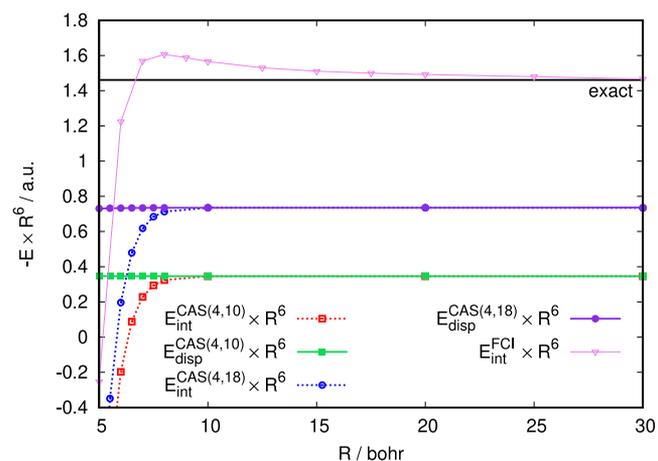

Figure 2. Comparison of the long-range behavior of the CASSCF interaction energy ($E_{int}^{CAS} \times R^6$) and the CASSCF dispersion energy ($E_{disp}^{CAS} \times R^6$) for the ground state of the He···He dimer based on the CAS(4,10) and CAS(4,18) references. "Exact" refers to the benchmark C_6 value of ref 80. The basis set is aug-cc-pVTZ.

agreement on the example of CAS(4,10) and CAS(4,18) wave functions. Even in the larger active space, CASSCF recovers only approximately 50% of the C_6 coefficient, that is, 0.735 a.u. compared to the reference value of 1.461 a.u.[80] A closer comparison with the long-range behavior of the FCI curve reveals that neither CAS(4,10) nor CAS(4,18) recovers higher C_n coefficients. This is not surprising because only p orbitals are included in the active space.

A more challenging model system is the He···H₂ dimer in which the hydrogen molecule is in one of the two lowest excited states, that is, either the $^1\Sigma_u^+$ or the $^1\Pi_u$ state. Note that the dispersion corrections employed in "CAS plus dispersion" approaches may be applied to excited states provided that the zero-order wave function is nondegenerate. We present the results for the T-shaped geometry where we vary the distance R between the center of mass of H₂ and the He atom. The equilibrium bond lengths in the H₂ molecule are 2.443 bohr for the $^1\Sigma_u^+$ state and 1.951 bohr the $^1\Pi_u$ state. We begin by

comparing the performance of the "CAS plus dispersion" models. In Figure 3, we show the results for the He···H$_2$($^1\Pi_u$)
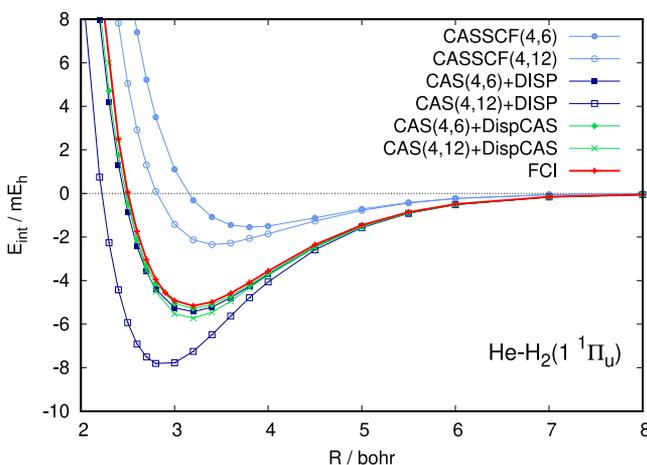
Figure 3. He···H$_2$($^1\Pi_u$) dimer interaction energy curve. The complex is kept in a T-shaped configuration. R denotes the distance between the He atom and the center of the H$_2$ molecule. The basis set is aug-cc-pVTZ.

complex obtained by choosing 2 active orbitals on the He atom and either 4 or 10 active orbitals on the H$_2$ molecule. Clearly, increasing the number of active orbitals leads to a severe overestimation of the interaction energy if the $E_{\text{DISP}}^{(2)}$ term (eq 55) is directly added to the CASSCF interaction energy (cf. "CAS + DISP" curves in Figure 3). Because of the correlation double counting in the larger CAS space, the minimum is predicted at 2.8 bohr, while the smaller CAS + DISP model correctly locates it at 3.2 bohr. When the amount of the added dispersion is tailored to the employed multireference model, as proposed in the CAS + DispCAS approach, interaction energy curves greatly improve. For CAS(4,6), the error in the van der Waals minimum amounts to 2.3%. In CAS(4,12), it is larger and reaches 11%. Considering that the dispersion energy is calculated exactly, this should be attributed to the erroneous representation of the remaining energy components in the supermolecular CASSCF approach.

Because applications of multireference methods to interacting excited systems are scarce, it is instructive to compare the performance of CAS + DispCAS with that of other approaches.
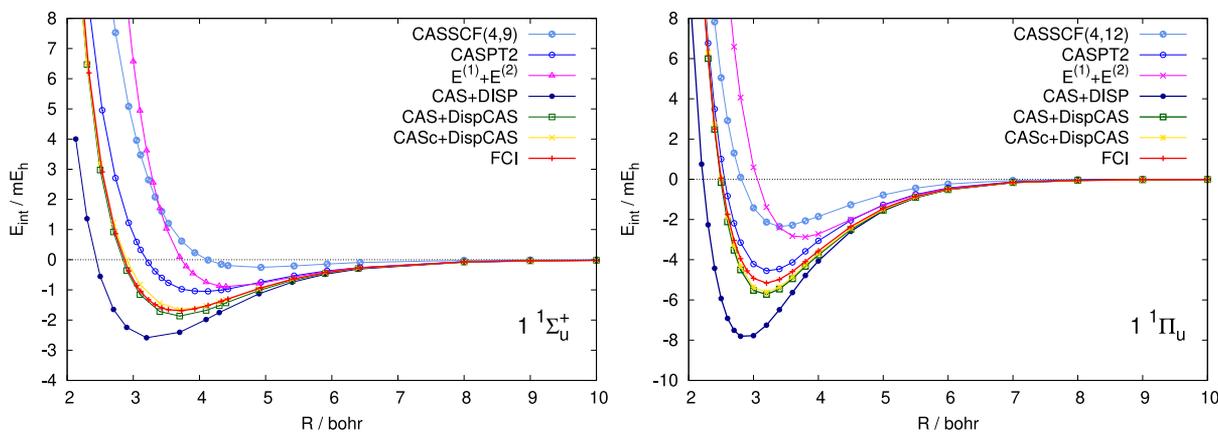
In Figure 4, the results from CASPT2 and SAPT(FCI) calculations up to the second order in $\hat{V}_{\text{int}}$ (the latter denoted as $E^{(1)} + E^{(2)}$) are shown next to the CASSCF and "CAS plus dispersion" models. The He atom and H$_2$($^1\Sigma_u^+$) molecule were described with the CAS(2,5) and CAS(2,4) wave functions, respectively. In the He···H$_2$($^1\Pi_u$) complex, we chose the CAS(2,2) active space for He and CAS(2,10) for H$_2$. The corresponding active spaces assumed for a dimer CASSCF calculation were (4,9) and (4,12) for the considered states. By confronting Figure 4, one observes a strikingly poor performance of the second-order SAPT energy (see "$E^{(1)} + E^{(2)}$" curves). Although the corresponding curves account for the second-order dispersion energy and in consequence show a correct asymptotic behavior, at intermediate- and short intermonomer distances second-order SAPT becomes even less accurate than CASSCF. This indicates that higher than second order induction terms, accounted for in CASSCF but missed in SAPT, are comparable in magnitude to the considered $E^{(1)} + E^{(2)}$ energy contributions. In the CASPT2 method the perturbation correction is added to the CASSCF energy, so the aforementioned higher-order terms are included. In spite of this, CASPT2 significantly underestimates the interaction energy for the $^1\Sigma_u^+$ state. In particular, the minimum is predicted at $R = 4.1$ bohr and 1.0 mE$_h$ deep, which is too shallow compared to the FCI benchmark that is 1.7 mE$_h$ deep and located at 3.7 bohr. For the $^1\Pi_u$ state, CASPT2 performs better and the curve has a correct shape. The interaction energy at the minimum ($R = 3.2$ bohr) is underestimated by 12.3%, which is similar in magnitude to the −11.0% overestimation from CAS + DispCAS calculations.

As already discussed, direct addition of the dispersion energy in the CAS + DISP approach leads to severe overcorrection in both the He···H$_2$($^1\Pi_u$) and He···H$_2$($^1\Sigma_u^+$) states (Figure 4). Contrary to this, CAS + DispCAS avoids dispersion double counting and performs well in both cases. The positions of van der Waals minima match the FCI benchmark and are predicted with −9.6 and −11.0% relative percent errors for the Σ_u^+ and Π_u states, respectively. The improvement of the first-order interaction energy components according to the procedure proposed in eq 59, cf. the CASc + DispCAS curves, has a marginal effect on the CAS + DispCAS energies in the $^1\Pi_u$ state. For example, the error in the minimum is reduced from −11% to −8.6%. This points toward the inaccurate treatment of higher than second order terms. In the $^1\Sigma_u^+$ state, the

Figure 4. He···H$_2$($^1\Sigma_u^+$) and He···H$_2$($^1\Pi_u$) interaction energy curves. The complex is kept in the T-shaped geometry. R denotes the distance between the He atom and the center of the H$_2$ molecule. The basis set is aug-cc-pVTZ.

Table 2. Interaction Energies (kcal/mol) and Error Statistics (in Percent) for Selected Systems of the A24 Data Set[a]

	HF	CAS	HF + DISP	HFc + DISP	CAS		CASc		ref.
					+DISP	+DispCAS	+DISP	+DispCAS	
H_2O–H_2O	−3.628	−2.710	−6.078	−4.859	−5.045	−5.012	−4.752	−4.719	−5.006
HCN–HCN	−4.100	−3.408	−5.866	−4.850	−5.108	−5.084	−4.695	−4.671	−4.745
HF–HF	−3.742	−2.741	−5.635	−4.294	−4.526	−4.505	−4.109	−4.088	−4.581
HF–CH_4	−0.257	0.168	−1.839	−1.725	−1.347	−1.328	−1.628	−1.609	−1.654
HCOH–HCOH	−2.458	−0.656	−6.659	−4.743	−4.749	−4.718	−4.410	−4.379	−4.554
C_2H_2–C_2H_2(1)	−0.511	−0.254	−1.922	−1.724	−1.637	−1.637	−1.577	−1.577	−1.524
C_2H_4–C_2H_4(1)	0.902	0.801	−1.232	−1.280	−1.299	−1.281	−1.204	−1.186	−1.090
Ar–CH_4	0.355	0.350	−0.494	−0.456	−0.470	−0.470	−0.425	−0.425	−0.405
Ar–C_2H_4	0.430	0.356	−0.348	−0.415	−0.409	−0.409	−0.390	−0.390	−0.364
C_2H_4–C_2H_2[b]	3.400	2.643	1.116	0.540	0.368	0.412	0.557	0.601	0.821
C_2H_4–C_2H_4(2)[b]	3.877	2.975	1.304	0.598	0.403	0.447	0.624	0.667	0.934
C_2H_2–C_2H_2(2)[b]	3.476	2.752	1.448	0.924	0.738	0.784	0.928	0.974	1.115
$\bar{\Delta}$	140	128	−6.4	−12.2	−16.2	−14.6	−7.3	−5.7	
$\bar{\Delta}_{abs}$	141	128	24.7	13.7	19.4	18.1	10.8	9.7	
σ	112	71.5	27.7	13.1	22.5	20.8	13.9	12.2	
Δ_{max}	315	222	46.2	35.9	56.8	52.2	33.2	28.5	

[a]HFc and CASc denote supermolecular HF and CASSCF interaction energies, respectively, corrected for the first-order SAPT(PBE0AC) energy. "CAS plus dispersion" models include the dispersion energy obtained at the SAPT(PBE0AC) level of theory in the aug-cc-pVTZ + (3322) basis. For the "CAS + DispCAS" and "CASc + DispCAS" methods, see eqs 58 and 59, respectively. "Ref." is the CCSD(T)/CBS benchmark of ref 70. [b]Nonequilibrium geometry.

difference is more pronounced—at the minimum geometry (R = 3.7 bohr) CASc + DispCAS deviates from the FCI by 2.2%, while CAS + DispCAS is off by − 9.6%.

4.2. Correcting Dispersion and First-Order Interaction Energy Components in the CAS Model for Many-Electron Systems. For model, few-electron systems discussed in Section 4.1, it is possible to employ large active spaces and restore a significant portion of the dispersion energy in size-consistent supermolecular CASSCF calculations. To verify what amount of the dispersion energy is captured by CASSCF for many-electron systems and examine the effect of correcting first-order interaction energy according to eq 59, in this section, we focus on a set of small, noncovalently bound dimers selected from the A24 database.[70] The analyzed complexes may be grouped according to the dominating interaction energy contributions into hydrogen-bonded dimers ($H_2O \cdots H_2O$, HCN$\cdots$HCN, and HF$\cdots$HF), mixed electrostatic/dispersion dimers [HF$\cdots CH_4$, HCOH$\cdots$HCOH, $C_2H_2 \cdots C_2H_2$(1), and $C_2H_4 \cdots C_2H_4$(1)], and dispersion-dominated (van der Waals) dimers [Ar$\cdots CH_4$, Ar$\cdots C_2H_4$, $C_2H_4 \cdots C_2H_2$, $C_2H_4 \cdots C_2H_4$(2), and $C_2H_2 \cdots C_2H_2$(2)]. In Table 2, we report the interaction energy values following from "CAS plus dispersion" and "HF plus dispersion" calculations [CCSD-(T)/CBS is taken as the reference]. Table 2 also contains error statistics in terms of the mean error $\bar{\Delta}$, the standard deviation σ, the mean absolute error $\bar{\Delta}_{abs}$, and the maximum absolute error Δ_{max}. We found that for the selected many-electron systems, the effect of the CASSE on the E_{disp}^{CAS} energy was negligible and the results presented in this section do not account for it (see Table S6 in the Supporting Information).

As expected for single-reference systems, both the CASSCF and HF methods not corrected for dispersion show similar, poor performance with mean errors that amount to 128% and 141%, respectively (Table 2). Addition of the $E_{DISP}^{(2)}$ dispersion term from SAPT(DFT) calculations drastically improves the results ($\bar{\Delta}_{abs}$ = 19% for CAS + DISP and $\bar{\Delta}_{abs}$ = 25% for HF + DISP). Note that the CAS + DISP model clearly outperforms its HF counterpart for hydrogen-bonded dimers. In contrast,

HF + DISP performs better for systems dominated by dispersion. The CAS-based approach may be corrected further by removing from CAS + DISP the dispersion energy recovered by CAS, that is, the E_{DISP}^{CAS} term, as done in the CAS + DispCAS model (see eq 58). This has a weak effect—the values of $|E_{DISP}^{CAS}|$ are smaller than 0.05 kcal/mol for the investigated systems. The only notable error reduction in CAS + DispCAS with respect to CAS + DISP occurs for the van der Waals dimers and does not exceed 5% (see also Table S4 in the Supporting Information). We conclude that in many-electron systems described with moderate active spaces, the double counting of the dispersion energy does not play a significant role, with the only exception being van der Waals complexes where it introduces discernible errors.

Both HF and CASSCF clearly benefit when the first-order electrostatic and exchange interaction energy components are treated at the SAPT(DFT) level of theory. The resulting HFc and CASc models provide a nearly 2-fold reduction in terms of both mean errors and the standard deviation when compared with their uncorrected HF and CAS equivalents (Table 2). In particular, the CASc + DispCAS approach, as formulated in eq 59, performs the best in the group ($\bar{\Delta}$ = −5.7% and σ = 12.2%). Still, relatively large errors between 13 and 27% persist for dispersion-bounded dimers for which the interaction energy has a repulsive character.

5. CONCLUSIONS

We presented a way to quantify the amount of the dispersion energy recovered in the supermolecular interaction energy, the residual dispersion energy, calculated with MC wave functions. Applying the RS perturbation theory to a group product wave function, we arrived at a general expression that rigorously yields a portion of the dispersion energy in the MCSCF interaction energy (eq 51). It is different from 0 if a dimer wave function is such that it dissociates into monomer wave functions both of which are at least two-configurational. The dispersion-in-MCSCF energy increases upon active space

expansion and reaches the full second-order dispersion limit when all orbitals are active.

By deriving the expression for the residual dispersion, we were able to rigorously address the problem of dispersion double counting which occurs when the dispersion energy is added directly to the supermolecular MCSCF interaction energy. We developed a dispersion correction for wave functions of the CAS type. The correction is tailored to the employed active space so that its addition to the CAS interaction energy, as shown in eq 54, accounts for exactly the amount of the second-order dispersion energy which is missed by the CAS. The resulting "CAS plus dispersion" method requires computing the $E_{\text{disp}}^{\text{CAS}}$ energy, eq 52, either in an exact manner by diagonalization of the active Hamiltonians for the monomers or approximately, by solving ERPA equations formulated with the active Hamiltonian for each monomer. One should notice that the proposed route of avoiding dispersion double counting in supermolecular CAS calculations is rigorous and does not rely on orbital localization, contrary to the physically motivated scheme outlined in ref 46. In the latter approach, the degree to which double counting of dispersion is avoided will depend on how well orbitals are localized on fragments. The most expensive step of computing $E_{\text{disp}}^{\text{CAS}}$ scales only with the sixth power of the number of active orbitals if ERPA equations are used, so it is modest comparing to the cost of CASSCF calculation. To assure the proper physical damping of the dispersion energy, we include exchange-dispersion terms approximated with second-order SAPT energy expressions, eqs 55−58. Therefore, in the CAS + DispCAS scheme proposed in this work, the CASSCF interaction energy is corrected for both dispersion and exchange-dispersion components.

On the example of the model He⋯He and He⋯H$_2$ systems in the ground and excited states, we demonstrated that the direct addition of the full second-order dispersion to the CAS interaction energy leads to overcorrelation which increases with the number of active orbitals. In contrast, the tailored dispersion correction applied in the CAS + DispCAS approach yields an accurate interaction energy regardless of the size of the active orbital set which validates the method. For the He⋯H$_2$($^1\Sigma_u^+$) and He⋯H$_2$($^1\Pi_u$) complexes, CAS + DispCAS proved to be superior to CASPT2 and second-order SAPT. This shows that the method may be successfully applied to study excited-state interactions provided that an accurate representation of the dispersion energy is available.

To examine the extent of the dispersion double counting in CASSCF calculations for many-electron systems, we studied several representative dimers selected from the A24 data set. We found out that in small active spaces feasible for size-consistent calculations [up to CAS(12,12) for the dimer], CASSCF recovers only 2−3% of the true dispersion. The overcorrelation becomes relevant in van der Waals (dispersion-dominated) systems where it translates into interaction energy errors at the level of 4−5%. This effect may be larger in the case of systems in excited states, when a few active diffuse orbitals contribute significantly to the dispersion interaction. Taking into account the low cost of computing the residual dispersion (as shown in this work, it depends only on the number of active orbitals, which is always much lower than the total number of orbitals), it is advisable to remove it from the supermolecular energy, even when the employed active space is relatively small.

Applications of the proposed CAS + DispCAS method for truly multireference systems require an adequate representation of the dispersion energy. Currently, it is possible to calculate both the polarization and exchange components of the second-order dispersion energy for many-electron CAS wave functions within the ERPA formalism.[16,65] A substantial improvement of the CAS + DispCAS accuracy is achieved if both the electrostatic and exchange first-order components recovered by the CAS are corrected for intramonomer correlation effects, as shown in eq 59. This approach is analogous to the HF-based HFDc$^{(1)}$ method of ref 8. Improving the quality of 1-RDMs from CASSCF calculations to compensate for the missing intramonomer correlation is less straightforward compared to the single-reference HFDc$^{(1)}$ approach. The multiconfiguration DFT theory could be explored in this context. Works on further verification and improvement of the accuracy provided by the CAS + DispCAS method, particularly in complexes involving excited states, are being carried out in our group.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.0c00681.

> ERPA equations for the active group Hamiltonian; second-order $E_{\text{disp}}^{\text{CAS}(n,m)}$ energy values for the He⋯He interaction energy curve; CASSCF, CASPT2, SAPT, and "CAS plus dispersion" interaction energy curves for the He⋯H$_2$($^1\Sigma_u^+$, $^1\Pi_u$) dimers; and absolute CASSCF energies and individual relative percent errors of "CAS plus dispersion" models for the systems from the A24 data set (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Katarzyna Pernal − *Institute of Physics, Lodz University of Technology, 90-924 Lodz, Poland;* ⓞ orcid.org/0000-0003-1261-9065; Email: pernalk@gmail.com

Authors

Michał Hapka − *Institute of Physics, Lodz University of Technology, 90-924 Lodz, Poland; Faculty of Chemistry, University of Warsaw, 02-093 Warsaw, Poland;* ⓞ orcid.org/0000-0001-7423-3198

Agnieszka Krzemińska − *Institute of Physics, Lodz University of Technology, 90-924 Lodz, Poland*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.0c00681

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105−5154.

(2) Hermann, J.; DiStasio, R. A.; Tkatchenko, A. First-Principles Models for van der Waals Interactions in Molecules and Materials: Concepts, Theory, and Applications. *Chem. Rev.* **2017**, *117*, 4714−4758.

(3) Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291−352.

(4) Bartlett, R. J. Coupled-cluster theory and its equation-of-motion extensions. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 126−138.

(5) Jeziorski, B.; van Hemert, M. Variation-perturbation treatment of the hydrogen bond between water molecules. *Mol. Phys.* **1976**, *31*, 713−729.

(6) Hepburn, J.; Scoles, G.; Penco, R. A simple but reliable method for the prediction of intermolecular potentials. *Chem. Phys. Lett.* **1975**, *36*, 451−456.

(7) Ahlrichs, R.; Penco, R.; Scoles, G. Intermolecular forces in simple systems. *Chem. Phys.* **1977**, *19*, 119−130.

(8) Podeszwa, R.; Pernal, K.; Patkowski, K.; Szalewicz, K. Extension of the Hartree−Fock Plus Dispersion Method by First-Order Correlation Effects. *J. Phys. Chem. Lett.* **2010**, *1*, 550−555.

(9) Garcia, J.; Szalewicz, K. Ab Initio Extended Hartree-Fock plus Dispersion Method Applied to Dimers with Hundreds of Atoms. *J. Phys. Chem. A* **2020**, *124*, 1196−1203.

(10) Pastorczak, E.; Shen, J.; Hapka, M.; Piecuch, P.; Pernal, K. Intricacies of van der Waals Interactions in Systems with Elongated Bonds Revealed by Electron-Groups Embedding and High-Level Coupled-Cluster Approaches. *J. Chem. Theory Comput.* **2017**, *13*, 5404−5419.

(11) Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. Multireference Nature of Chemistry: The Coupled-Cluster View. *Chem. Rev.* **2012**, *112*, 182−243.

(12) Evangelista, F. A. Perspective: Multireference coupled cluster theories of dynamical electron correlation. *J. Chem. Phys.* **2018**, *149*, 030901.

(13) Cui, Z.-h.; Lischka, H.; Beneberu, H. Z.; Kertesz, M. Rotational Barrier in Phenalenyl Neutral Radical Dimer: Separating Pancake and van der Waals Interactions. *J. Am. Chem. Soc.* **2014**, *136*, 5539−5542.

(14) Cui, Z.-h.; Lischka, H.; Beneberu, H. Z.; Kertesz, M. Double Pancake Bonds: Pushing the Limits of Strong π-π Stacking Interactions. *J. Am. Chem. Soc.* **2014**, *136*, 12958−12965.

(15) Kunz, C. F.; Hättig, C.; Hess, B. A. Ab initio study of the individual interaction energy components in the ground state of the mercury dimer. *Mol. Phys.* **1996**, *89*, 139−156.

(16) Hapka, M.; Przybytek, M.; Pernal, K. Second-Order Dispersion Energy Based on Multireference Description of Monomers. *J. Chem. Theory Comput.* **2019**, *15*, 1016−1027.

(17) Langhoff, S. R.; Davidson, E. R. Configuration interaction calculations on the nitrogen molecule. *Int. J. Quantum Chem.* **1974**, *8*, 61−72.

(18) Pople, J. A.; Seeger, R.; Krishnan, R. Variational configuration interaction methods and comparison with perturbation theory. *Int. J. Quantum Chem.* **1977**, *12*, 149−163.

(19) Buchachenko, A. A.; Szczęśniak, M. M.; Chałasiński, G. Van der Waals interactions and dipole polarizabilities of lanthanides: $Tm(^2F)$-He and $Yb(^1S)$-He potentials. *J. Chem. Phys.* **2006**, *124*, 114301.

(20) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553−566.

(21) Kłos, J. A.; Chałasiński, G.; Szczęśniak, M. M.; Werner, H.-J. Ab initio calculations of adiabatic and diabatic potential energy surfaces of $Cl(^2P)$-$HCl(^1\Sigma^+)$ van der Waals complex. *J. Chem. Phys.* **2001**, *115*, 3085−3098.

(22) Rode, J. E.; Klos, J.; Rajchel, L.; Szczęśniak, M. M.; Chałasiński, G.; Buchachenko, A. A. Interactions in Open-Shell Clusters: Ab Initio Study of Pre-reactive Complex $O(^3P)$ + HCl. *J. Phys. Chem. A* **2005**, *109*, 11484−11494.

(23) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483−5488.

(24) Andersson, K.; Malmqvist, P.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218−1226.

(25) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252−10264.

(26) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. N-electron valence state perturbation theory: a fast implementation of the strongly contracted variant. *Chem. Phys. Lett.* **2001**, *350*, 297−305.

(27) Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. n-electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.* **2002**, *117*, 9138−9153.

(28) Hirao, K. Multireference Møller-Plesset method. *Chem. Phys. Lett.* **1992**, *190*, 374−380.

(29) Nakano, H. Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions. *J. Chem. Phys.* **1993**, *99*, 7983−7992.

(30) Choe, Y.-K.; Witek, H. A.; Finley, J. P.; Hirao, K. Identifying and removing intruder states in multireference Møller-Plesset perturbation theory. *J. Chem. Phys.* **2001**, *114*, 3913−3918.

(31) Chang, S.-W.; Witek, H. A. Choice of Optimal Shift Parameter for the Intruder State Removal Techniques in Multireference Perturbation Theory. *J. Chem. Theory Comput.* **2012**, *8*, 4053−4061.

(32) Chałasiński, G.; Szczęśniak, M. On the connection between the supermolecular Møller-Plesset treatment of the interaction energy and the perturbation theory of intermolecular forces. *Mol. Phys.* **1988**, *63*, 205−224.

(33) Cybulski, S. M.; Chałasiński, G.; Moszyński, R. On decomposition of second-order Møller-Plesset supermolecular interaction energy and basis set effects. *J. Chem. Phys.* **1990**, *92*, 4357−4363.

(34) Chałasiński, G.; Szczęśniak, M. M. Origins of structure and energetics of van der Waals clusters from ab initio calculations. *Chem. Rev.* **1994**, *94*, 1723−1765.

(35) Podeszwa, R.; Szalewicz, K. Accurate interaction energies for argon, krypton, and benzene dimers from perturbation theory based on the Kohn-Sham model. *Chem. Phys. Lett.* **2005**, *412*, 488−493.

(36) Cybulski, S. M.; Lytle, M. L. The origin of deficiency of the supermolecule second-order Møller-Plesset approach for evaluating interaction energies. *J. Chem. Phys.* **2007**, *127*, 141102.

(37) Żuchowski, P. S.; Bussery-Honvault, B.; Moszynski, R.; Jeziorski, B. Dispersion interaction of high-spin open-shell complexes in the random phase approximation. *J. Chem. Phys.* **2003**, *119*, 10497−10511.

(38) Stoll, H.; Savin, A. In *Density Functional Methods in Physics*; Dreizler, R., da Providencia, J., Eds.; Plenum: New York, 1985; pp 177−207.

(39) Savin, A. In *Recent Developments of Modern Density Functional Theory*; Seminario, J. M., Ed.; Elsevier: Amsterdam, 1996; pp 327−357.

(40) Fromager, E.; Cimiraglia, R.; Jensen, H. J. A. Merging multireference perturbation and density-functional theories by means of range separation: Potential curves for Be_2, Mg_2, and Ca_2. *Phys. Rev. A: At., Mol., Opt. Phys.* **2010**, *81*, 024502.

(41) Hapka, M.; Pastorczak, E.; Krzemińska, A.; Pernal, K. Long-range-corrected multiconfiguration density functional with the on-top pair density. *J. Chem. Phys.* **2020**, *152*, 094102.

(42) Pernal, K. Electron Correlation from the Adiabatic Connection for Multireference Wave Functions. *Phys. Rev. Lett.* **2018**, *120*, 013001.

(43) Goll, E.; Werner, H.-J.; Stoll, H. A short-range gradient-corrected density functional in long-range coupled-cluster calculations for rare gas dimers. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3917−3923.

(44) Fromager, E.; Toulouse, J.; Jensen, H. J. A. On the universality of the long-/short-range separation in multiconfigurational density-functional theory. *J. Chem. Phys.* **2007**, *126*, 074111.

(45) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(46) Stein, C. J.; Reiher, M. Semiclassical Dispersion Corrections Efficiently Improve Multiconfigurational Theory with Short-Range Density-Functional Dynamic Correlation. *J. Phys. Chem. A* **2020**, *124*, 2834−2841.

(47) Rajchel, Ł.; Żuchowski, P. S.; Kłos, J.; Szczęśniak, M. M.; Chałasiński, G. Interactions of transition metal atoms in high-spin states: Cr_2, Sc-Cr, and Sc-Kr. *J. Chem. Phys.* **2007**, *127*, 244302.

(48) Hapka, M.; Kłos, J.; Korona, T.; Chałasiński, G. Theoretical Studies of Potential Energy Surface and Bound States of the Strongly Bound $He(^1S)$-BeO $(^1\Sigma^+)$ Complex. *J. Phys. Chem. A* **2013**, *117*, 6657−6663.

(49) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chem. Rev.* **1994**, *94*, 1887−1930.

(50) Patkowski, K. Recent developments in symmetry-adapted perturbation theory. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2020**, *10*, No. e1452.

(51) Schütz, M.; Rauhut, G.; Werner, H.-J. Local Treatment of Electron Correlation in Molecular Clusters: Structures and Stabilities of $(H_2O)_n$, n=2-4. *J. Phys. Chem. A* **1998**, *102*, 5997−6003.

(52) Thirman, J.; Head-Gordon, M. An energy decomposition analysis for second-order Møller-Plesset perturbation theory based on absolutely localized molecular orbitals. *J. Chem. Phys.* **2015**, *143*, 084124.

(53) Schneider, W. B.; Bistoni, G.; Sparta, M.; Saitow, M.; Riplinger, C.; Auer, A. A.; Neese, F. Decomposition of Intermolecular Interaction Energies within the Local Pair Natural Orbital Coupled Cluster Framework. *J. Chem. Theory Comput.* **2016**, *12*, 4778−4792.

(54) Kaplan, I. G. *Intermolecular Interactions: Physical Picture, Computational Methods and Model Potentials*; John Wiley & Sons, 2006.

(55) McWeeny, R. The Density Matrix in Many-Electron Quantum Mechanics. I. Generalized Product Functions. Factorization and Physical Interpretation of the Density Matrices. *Proc. R. Soc. A* **1959**, *253*, 242−259.

(56) Singh, S. B.; Singh, C. A. Extensions of the Feynman-Hellman theorem and applications. *Am. J. Phys.* **1989**, *57*, 894−899.

(57) Caves, T. C.; Karplus, M. Perturbed Hartree-Fock Theory. I. Diagrammatic Double-Perturbation Analysis. *J. Chem. Phys.* **1969**, *50*, 3649−3661.

(58) Stevens, R. M.; Pitzer, R. M.; Lipscomb, W. N. Perturbed Hartree-Fock Calculations. I. Magnetic Susceptibility and Shielding in the LiH Molecule. *J. Chem. Phys.* **1963**, *38*, 550−560.

(59) Jeziorski, B.; Szalewicz, K.; Chałasiński, G. Symmetry forcing and convergence properties of perturbation expansions for molecular interaction energies. *Int. J. Quantum Chem.* **1978**, *14*, 271−287.

(60) Szalewicz, K.; Patkowski, K.; Jeziorski, B. In *Intermolecular Forces and Clusters II*; Wales, D. J., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2005; pp 43−117.

(61) Pernal, K.; Chatterjee, K.; Kowalski, P. H. How accurate is the strongly orthogonal geminal theory in predicting excitation energies? Comparison of the extended random phase approximation and the linear response theory approaches. *J. Chem. Phys.* **2014**, *140*, 014101.

(62) Chatterjee, K.; Pernal, K. Excitation energies from extended random phase approximation employed with approximate one- and two-electron reduced density matrices. *J. Chem. Phys.* **2012**, *137*, 204109.

(63) Pernal, K. Intergeminal Correction to the Antisymmetrized Product of Strongly Orthogonal Geminals Derived from the Extended Random Phase Approximation. *J. Chem. Theory Comput.* **2014**, *10*, 4332−4341.

(64) Pastorczak, E.; Pernal, K. Correlation Energy from the Adiabatic Connection Formalism for Complete Active Space Wave Functions. *J. Chem. Theory Comput.* **2018**, *14*, 3493−3503.

(65) Hapka, M.; Przybytek, M.; Pernal, K. Second-Order Exchange-Dispersion Energy Based on a Multireference Description of Monomers. *J. Chem. Theory Comput.* **2019**, *15*, 6712−6723.

(66) Korona, T. First-order exchange energy of intermolecular interactions from coupled cluster density matrices and their cumulants. *J. Chem. Phys.* **2008**, *128*, 224104.

(67) Podeszwa, R.; Bukowski, R.; Szalewicz, K. Potential Energy Surface for the Benzene Dimer and Perturbational Analysis of $\pi-\pi$ Interactions. *J. Phys. Chem. A* **2006**, *110*, 10345−10354.

(68) Przybytek, M.; Cencek, W.; Jeziorski, B.; Szalewicz, K. Pair Potential with Submillikelvin Uncertainties and Nonadiabatic Treatment of the Halo State of the Helium Dimer. *Phys. Rev. Lett.* **2017**, *119*, 123401.

(69) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(70) Řezáč, J.; Hobza, P. Describing Noncovalent Interactions beyond the Common Approximations: How Accurate Is the "Gold Standard," CCSD(T) at the Complete Basis Set Limit? *J. Chem. Theory Comput.* **2013**, *9*, 2151−2155.

(71) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 242−253.

(72) Korona, T.; Moszynski, R.; Jeziorski, B. Convergence of Symmetry-Adapted Perturbation Theory for the Interaction between Helium Atoms and between a Hydrogen Molecule and a Helium Atom. In *Advances in Quantum Chemistry*; Löwdin, P.-O., Sabin, J. R., Zerner, M. C., Karwowski, J., Karelson, M., Eds.; Academic Press, 1997; Vol. *28*; pp 171−188.

(73) Bak, K. L.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Klopper, W. Accuracy of atomization energies and reaction enthalpies in standard and extrapolated electronic wave function/basis set calculations. *J. Chem. Phys.* **2000**, *112*, 9229−9242.

(74) Heßelmann, A.; Jansen, G.; Schütz, M. Density-functional theory-symmetry-adapted intermolecular perturbation theory with density fitting: A new efficient method to study intermolecular interaction energies. *J. Chem. Phys.* **2005**, *122*, 014103.

(75) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional calculations. *J. Chem. Phys.* **2005**, *123*, 214103.

(76) Jansen, G. Symmetry-adapted perturbation theory based on density functional theory for noncovalent interactions. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 127−144.

(77) Grüning, M.; Gritsenko, O. V.; van Gisbergen, S. J. A.; Baerends, E. J. Shape corrections to exchange-correlation potentials by gradient-regulated seamless connection of model potentials for inner and outer region. *J. Chem. Phys.* **2001**, *114*, 652−660.

(78) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(79) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.

(80) Zhang, J.-Y.; Yan, Z.-C.; Vrinceanu, D.; Babb, J. F.; Sadeghpour, H. R. Long-range interactions for $He(nS)−He(n'S)$ and $He(nS)−He(n'P)$. *Phys. Rev. A: At., Mol., Opt. Phys.* **2006**, *74*, 014704.